

Actividad 3: Componentes Principales (1)

Code ▼

Yolanda Elizondo Chapa A01137848: 11 de octubre del 2022

Parte 1

Construcción del data frame con los datos de x1 y x2

Hide

```
x1 = c(2.5, 0.5, 2.2, 1.9, 3.1, 2.3, 2, 1, 1.5, 1.1)
x2 = c(2.4, 0.7, 2.9, 2.2, 3.0, 2.7, 1.6, 1.1, 1.6, 0.9)

M = data.frame(x1,x2)
```

Obtener la matriz de los datos centrados

Hide

```
m1 = c(rep(mean(x1),10))
m2 = c(rep(mean(x2),10))

M1 = data.frame(m1,m2)
MatrizCentrada = M - M1
```

Obtener matriz de varianza-covarianza de la matriz de datos centrados

Hide

```
mcov = cov(MatrizCentrada)
mcov
```

```
      x1      x2
x1 0.6165556 0.6154444
x2 0.6154444 0.7165556
```

Obtener valores propios de la matriz de varianza-covarianza

Hide

```
propios = eigen(mcov)$values #valores propios
v = eigen(mcov)$vectors #vectores propios
```

Hide

```
t_v = t(v) #Transpuesta de vectores propios
t_v
```

```

      [,1]      [,2]
[1,]  0.6778734 0.7351787
[2,] -0.7351787 0.6778734

```

Hide

```

t_M1 = t(MatrizCentrada) #Transpuesta de matriz de datos centrados
t_M1

```

```

      [,1] [,2] [,3] [,4] [,5]
x1 0.69 -1.31 0.39 0.09 1.29
x2 0.49 -1.21 0.99 0.29 1.09
      [,6] [,7] [,8] [,9]
x1 0.49 0.19 -0.81 -0.31
x2 0.79 -0.31 -0.81 -0.31
      [,10]
x1 -0.71
x2 -1.01

```

Multiplicación de la matriz transpuesta de los vectores propios con la transpuesta de la matriz de datos centrados.

Hide

```

CP = t_v%%t_M1
CP

```

```

      [,1]      [,2]
[1,]  0.8279702 -1.7775803
[2,] -0.1751153 0.1428572
      [,3]      [,4]
[1,] 0.9921975 0.2742104
[2,] 0.3843750 0.1304172
      [,5]      [,6]
[1,] 1.6758014 0.9129491
[2,] -0.2094985 0.1752824
      [,7]      [,8]
[1,] -0.09910944 -1.14457216
[2,] -0.34982470 0.04641726
      [,9]      [,10]
[1,] -0.43804614 -1.2238206
[2,] 0.01776463 -0.1626753

```

Resultados en forma de tabla

Hide

```

rownames(CP)= c("CP1", "CP2")
t(CP)

```

```
      CP1      CP2
[1,]  0.82797019 -0.17511531
[2,] -1.77758033  0.14285723
[3,]  0.99219749  0.38437499
[4,]  0.27421042  0.13041721
[5,]  1.67580142 -0.20949846
[6,]  0.91294910  0.17528244
[7,] -0.09910944 -0.34982470
[8,] -1.14457216  0.04641726
[9,] -0.43804614  0.01776463
[10,] -1.22382056 -0.16267529
```

Parte 2: Fórmulas de R para componentes principales

[Hide](#)

```
cpa <- prcomp(M, scale=TRUE)
cpa
```

```
Standard deviations (1, ..., p=2):
[1] 1.3877785 0.2721594
```

```
Rotation (n x k) = (2 x 2):
      PC1      PC2
x1 -0.7071068  0.7071068
x2 -0.7071068 -0.7071068
```

[Hide](#)

```
names(cpa)
```

```
[1] "sdev"      "rotation" "center"   "scale"    "x"
```

[Hide](#)

```
print("desviaciones estándar: ")
```

```
[1] "desviaciones estándar: "
```

[Hide](#)

```
cpa$sdev
```

```
[1] 1.3877785 0.2721594
```

[Hide](#)

```
print("medias: ")
```

```
[1] "medias: "
```

Hide

```
print("center y scale dan las medias y desv estándar previa estandarización: ")
```

```
[1] "center y scale dan las medias y desv estándar previa estandarización: "
```

Hide

```
cpa$center
```

```
  x1  x2  
1.81 1.91
```

Hide

```
cpa$scale
```

```
      x1      x2  
0.7852105 0.8464960
```

Hide

```
print("Los coeficientes de la combinación lineal normalizada de componete")
```

```
[1] "Los coeficientes de la combinación lineal normalizada de componete"
```

Hide

```
cpa$rotation
```

```
      PC1      PC2  
x1 -0.7071068 0.7071068  
x2 -0.7071068 -0.7071068
```

Hide

```
print("Los datos por sustituidos en la combinación lineal de vectores propios:")
```

```
[1] "Los datos por sustituidos en la combinación lineal de vectores propios:"
```

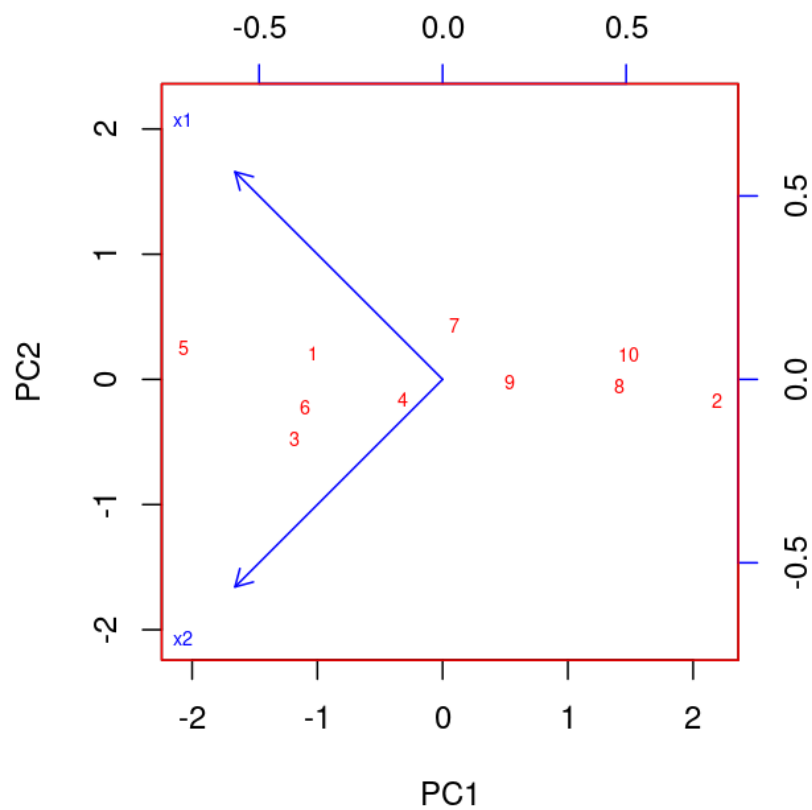
Hide

```
cpa$x
```

	PC1	PC2
[1,]	-1.03068029	0.21205314
[2,]	2.19045016	-0.16894230
[3,]	-1.17818776	-0.47577321
[4,]	-0.32329464	-0.16119898
[5,]	-2.07219947	0.25117173
[6,]	-1.10117414	-0.21865330
[7,]	0.08785251	0.43005447
[8,]	1.40605089	-0.05281009
[9,]	0.53811824	-0.02021127
[10,]	1.48306451	0.20430982

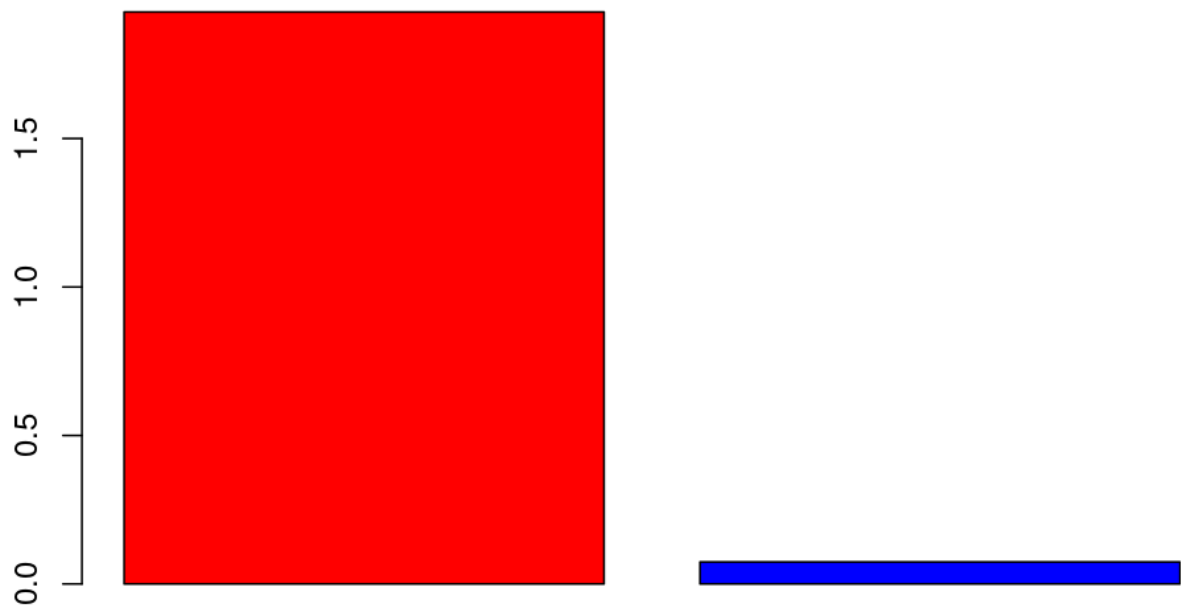
Hide

```
biplot(x = cpa, scale = 0, cex = 0.6, col = c("red", "blue"))
```



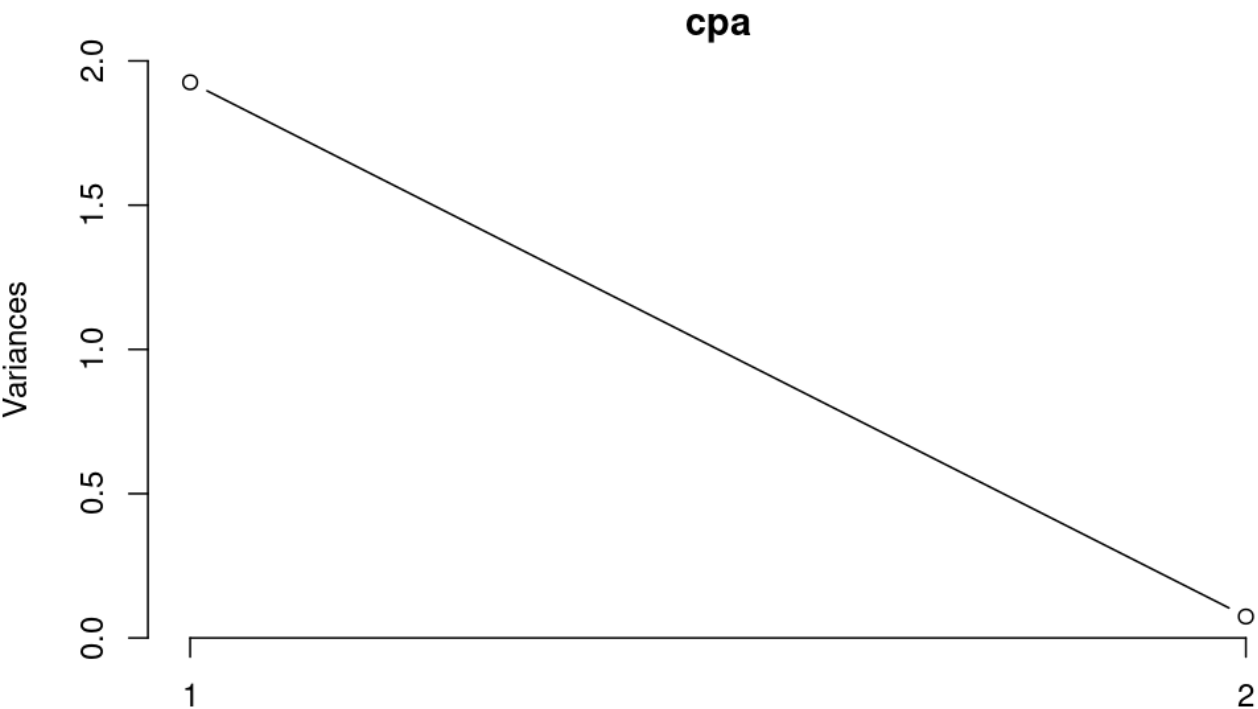
Hide

```
barplot(cpa$sdev^2, col = c("red", "blue"))
```



Hide

```
plot(cpa, type="l")
```



Hide

```
summary(cpa)
```

Importance of components:

	PC1	PC2
Standard deviation	1.388	0.27216
Proportion of Variance	0.963	0.03704
Cumulative Proportion	0.963	1.00000

Solo con la primera componente llegamos a un 96% por lo que podemos eliminar la segunda variable y aun así se logra explicar el 96% de la variabilidad de los datos. Esto pensando que el dataset fuera mucho más grande.