

Análisis Factorial Fiel

Melissa Ortega Galarza

2022-05-21

Se instalan las paqueterías

```
install.packages("datos")
```

```
library(datos)
```

1.- Lectura de la matriz de datos

```
M<-data.frame(datos::fiel)
```

```
M<-as.data.frame(fiel)
```

2.- Quitar los espacios de los nombres

```
colnames(M)[1]="Life.Exp"
```

```
colnames(M)[2]="HS.Grad"
```

3.- Separa n (estados) y p (variables)

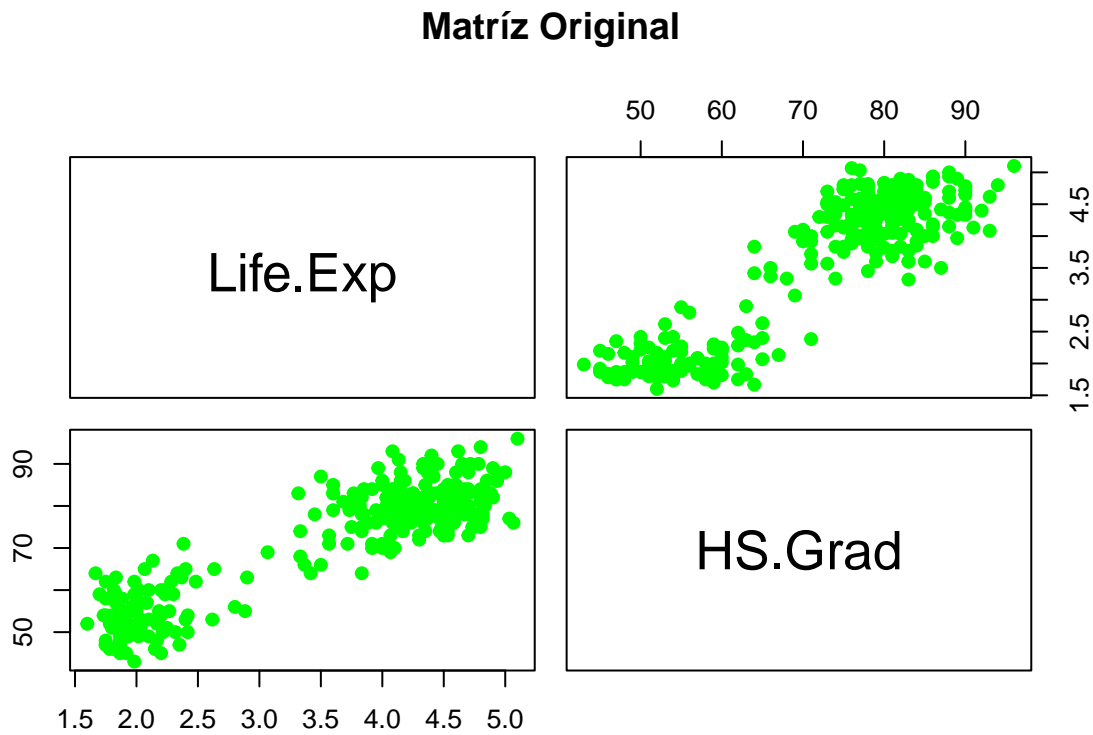
```
n<-dim(M)[1]
```

```
p<-dim(M)[2]
```

4.- Generación de un scatter plot

Para la Visualización de variables originales

```
pairs(M, col="green", pch=19, main="Matriz Original")
```



Transformación de alguna variables

1.- Aplicamos logaritmo para las columnas 1,3 y 8

```
M[,1]<-log(M[,1])  
colnames(M)[1]<-"Log-Population"
```

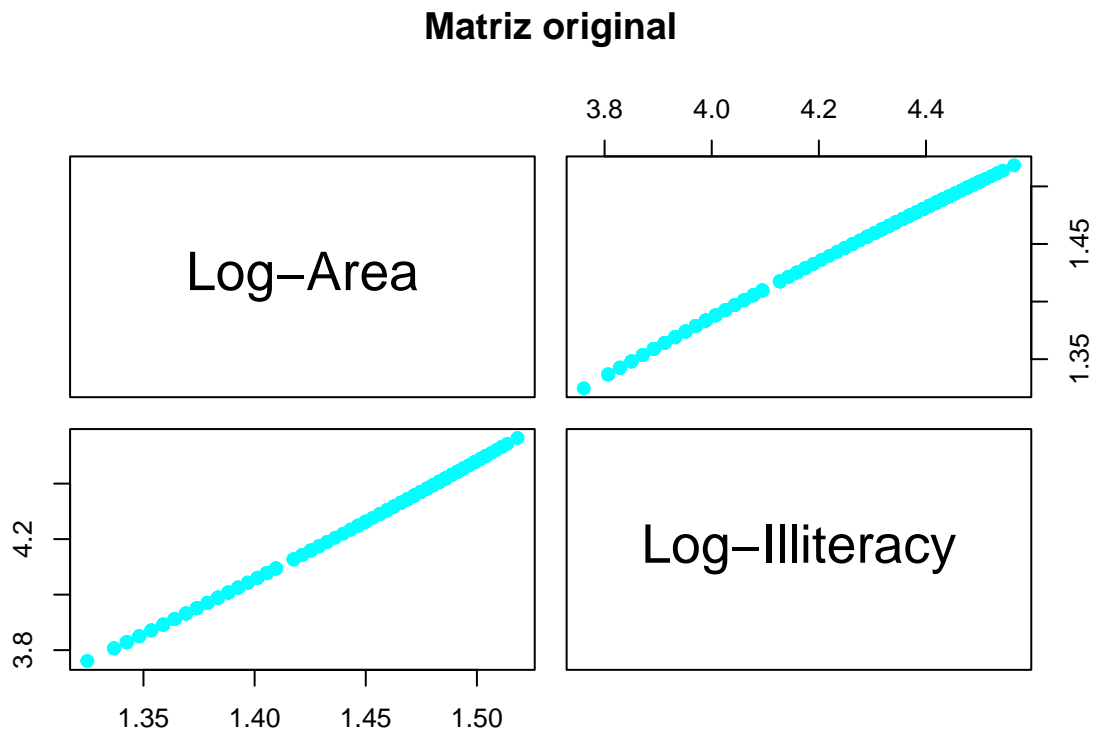
```
M[,2]<-log(M[,2])  
colnames(M)[2]<-"Log-Illiteracy"
```

```
M[,1]<-log(M[,2])  
colnames(M)[1]<-"Log-Area"
```

Gráfico scater

Para la visualización de la matriz original con 3 variables que se incluyeron

```
pairs(M,col="cyan", pch=19, main="Matriz original")
```



Nota:

Como las variables tienen diferentes unidades de medida, se va a implementar la matriz de correlaciones para estimar la matriz de carga.

Reduccion de la dimensionalidad

Análisis Factorial de componentes principales (PCFA)

1.- Calcular la matriz de medias y de correlaciones

Matriz de medias

```
mu<-colMeans(M)
```

#Matriz de correlaciones

```
R<-cor(M)
```

2.- Reducción de la dimensionalidad mediante

Análisis factorial de componentes principales (PCFA).

1.- Calcular los valores y vectores propios.

```
eR<-eigen(R)
```

2.- Valores propios

```
eigen.val<-eR$values
```

3.- Vectores propios

```
eigen.vec<-eR$vectors
```

4.- Calcular la proporcion de variabilidad

```
prop.var<-eigen.val/sum(eigen.val)
```

5.- Calcular la proporcion de variabilidad acumulada

```
prop.var.acum<-cumsum(eigen.val)/sum(eigen.val)
```

```
L.est.1<-eigen.vec[,1:2] %*% diag(sqrt(eigen.val[1:2]))
```

Rotación varimax

```
L.est.1.var<-varimax(L.est.1)
```

Estimación de la matriz de los errores

1.- Estimación de la matriz de perturbaciones

```
Psi.est.1<-diag(diag(R-as.matrix(L.est.1.var$loadings))%% t(as.matrix(L.est.1.var$loadings))))
```

2.- Se utiliza el método Análisis de factor principal (PFA)

Para estimación de autovalores y autovectores

```
RP<-R-Psi.est.1
```

Calculo de la matriz de autovalores y autovectores

```
eRP<-eigen(RP)
```

Autovalores

```
eigen.val.RP<-eRP$values
```

Autovectores

```
eigen.vec.RP<-eRP$vectors
```

Proporcion de variabilidad

```
prop.var.RP<-eigen.val.RP/ sum(eigen.val.RP)
```

Proporcion de variabilidad acumulada

```
prop.var.RP.acum<-cumsum(eigen.val.RP)/ sum(eigen.val.RP)
```

Estimación de la matriz de cargas con rotación varimax

```
L.est.2<-eigen.vec.RP[,1:2] %%% diag(sqrt(eigen.val.RP[1:2]))
```

Rotación varimax

```
L.est.2.var<-varimax(L.est.2)
```

Estimación de la matriz de covarianzas de los errores.

```
Psi.est.2<-diag(diag(R-as.matrix(L.est.2.var$loadings))%% t(as.matrix(L.est.2.var$loadings))))
```

Obtención de los scores de ambos métodos

PCFA

```
FS.est.1<-scale(M)%*% as.matrix(L.est.1.var$loadings)
```

PFA

```
FS.est.2<-scale(M)%*% as.matrix (L.est.2.var$loadings)
```

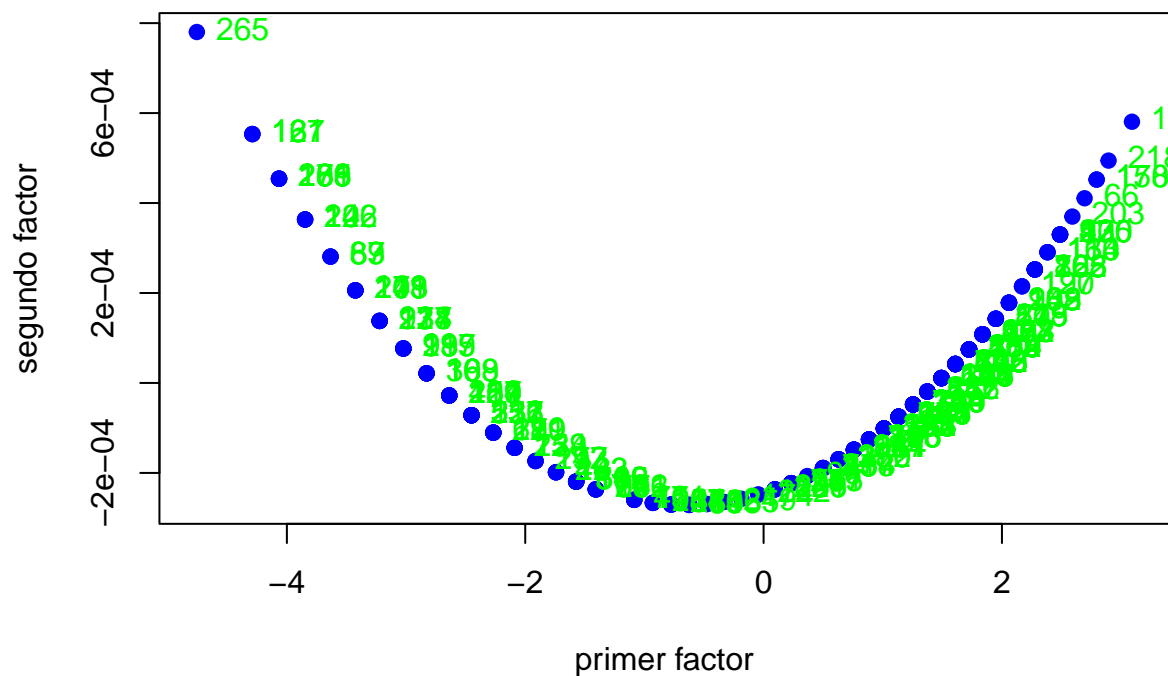
Graficamos ambos scores

```
par(mfrow=c(2,1))
```

Factor I y II

```
pl1<-plot(FS.est.1[,1], FS.est.1[,2], xlab="primer factor",  
          ylab="segundo factor", main="scores con factor I y II con PCFA",  
          pch=19, col="blue")  
text(FS.est.1[,1], FS.est.1[,2], labels = rownames(M), pos=4, col="green")
```

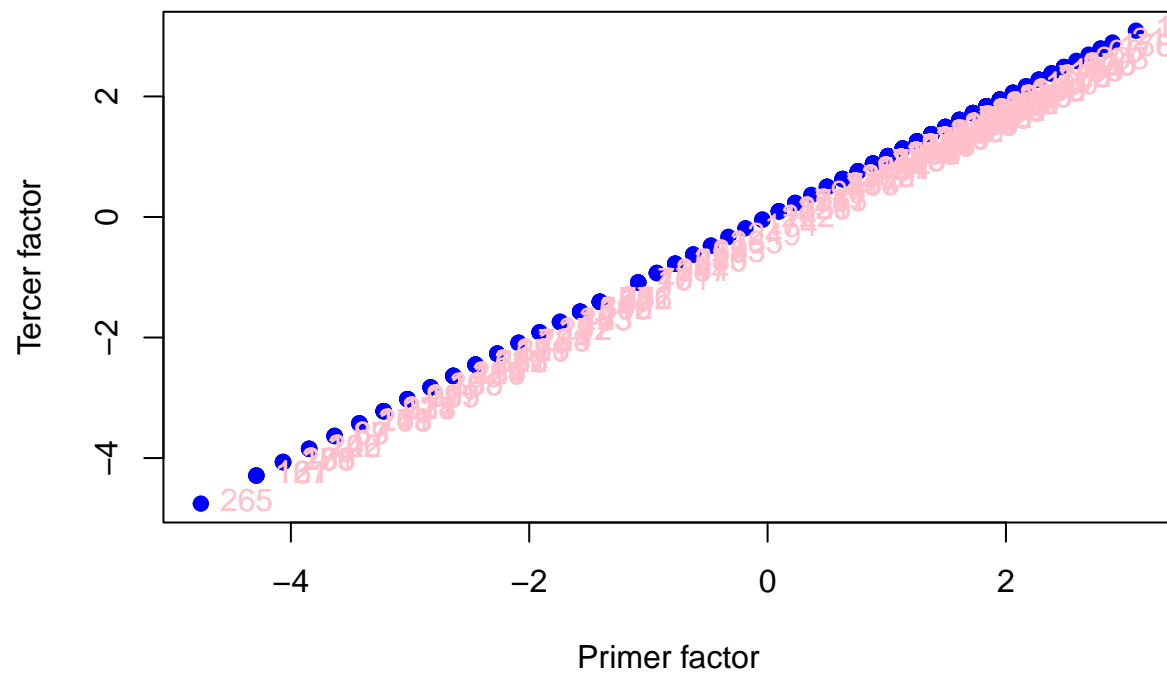
scores con factor I y II con PCFA



Factor II y I

```
pl2<-plot(FS.est.2[,1], FS.est.1[,1], xlab="Primer factor",  
          ylab="Tercer factor", main="scores con factor II y I con PCFA",  
          pch=19, col="blue")  
text(FS.est.2[,1], FS.est.2[,1], labels = rownames(M), pos=4, col="pink")
```

scores con factor II y I con PCFA



Factor II y II

```
pl3<-plot(FS.est.2[,2], FS.est.2[,2], xlab="Segundo factor",  
          ylab="Tercer factor", main="scores con factor II y II con PCFA",  
          pch=19, col="cyan")  
text(FS.est.2[,2], FS.est.2[,2], labels = rownames(M), pos=4, col="red")
```

scores con factor II y II con PCFA

