Sesión 11: Corrección de Tarea del Análisis de Componentes Principales

Alejandra Lelo de Larrea Ibarra

24 de febrero de 2019.

Índice

1. Introducción **1**

2. Datos **1**
2.1. Análisis exploratorio 2
2.2. México vs India 4

3. Inferencia de SPCA **8**
3.1. Estandarización de los datos 9
3.2. Actualizacion bayesiana de parámetros 10
3.3. Simulación 11

4. Resultados **12**
4.1. Eigenvalores 12
4.2. ¿Qué economía tiene el mayor peso esperado en la descomposición PCA? 13
4.3. ¿Qué economía tiene la mayor consistencia en estimacion de los cjs correspondientes? 15

1. Introducción

Se tienen datos con los tipos de cambio reales (respecto a USD) de varias economias (periodo 1970-2010). El objetivo es implementar el procedimiento inferencial PCA considerando distribuciones iniciales no informativas para (μ, Λ) y contestar lo siguiente:

- ¿Qué economía tiene el mayor peso esperado en la descomposicion PCA?
- ¿Qué economía tiene la mayor consistencia en estimacion de los cjs correspondientes?

2. Datos

```r
# Se cargan los paquetes
library("fields")
library("mnormt")
library("MCMCpack")
library("actuar")
library("ggplot2")
library("kernlab")
library("tidyverse")
library("readr")
library("psych")
library("mvtnorm")
library("MASS")
```

```r
library("xlsx")
library("knitr")
```

```r
# Función para extraer modas
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
```

Leemos los datos correspondientes a los tipos de cambio de distintas economías. Se tienen 492 observaciones mensuales para 80 economías.

```r
# Cargamos los datos)
data<-read.xlsx("../01_Notas_Ovando/est46114_s06_data.xls",sheetName = 'RealXR_Data')

# Obtenemos las dimensiones de los datos
dim(data)
```

```
## [1] 492  81
```

```r
# Extraemos las fechas
fechas<-data$Date

data<-select(data,-Date)
```

2.1. Análisis exploratorio

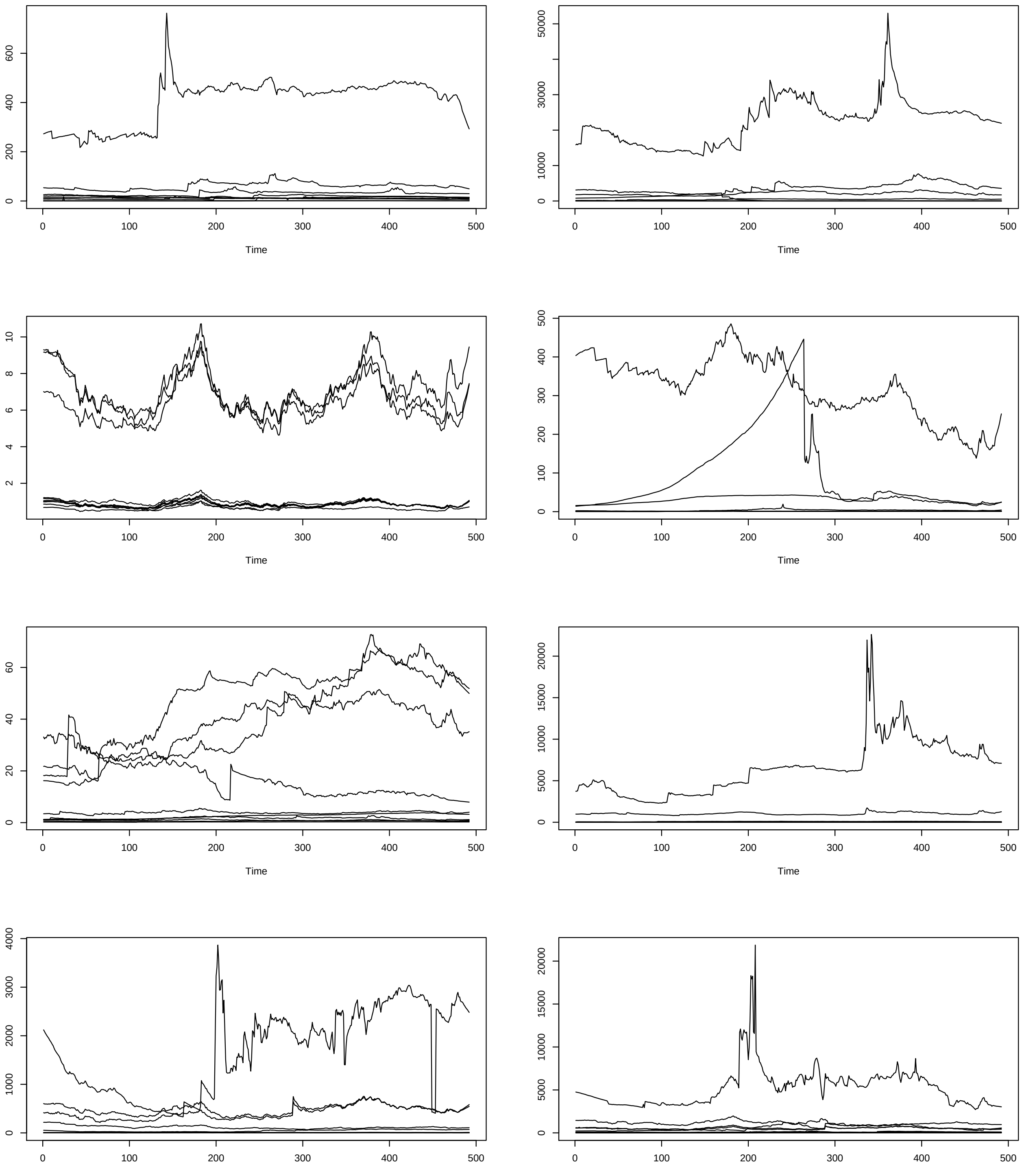
Vemos que países están en la muestra:

```r
# Vemos la lista de países
colnames(data)
```

```
##  [1] "Canada"        "Mexico"        "Guatemala"     "El.Salvador"
##  [5] "Honduras"      "Nicaragua"     "Costa.Rica"    "Panama"
##  [9] "Jamaica"       "Dominican.Rep" "Trin.Tobago"   "Colombia"
## [13] "Venezuela."    "Ecuador"       "Peru"          "Chile"
## [17] "Brazil."       "Paraguay"      "Uruguay"       "Argentina"
## [21] "EU12"          "Sweden"        "Norway"        "Finland"
## [25] "Denmark"       "U.K."          "Ireland"       "Luxembourg"
## [29] "Netherlands"   "France"        "Germany"       "Austria"
## [33] "Czech.Rep"     "Hungary"       "Switzerland"   "Poland"
## [37] "Russia"        "Spain"         "Portugal"      "Italy"
## [41] "Greece"        "Turkey"        "Syria"         "Israel"
## [45] "Jordan"        "Kuwait"        "Saudi.Arabia"  "India"
## [49] "Pakistan"      "Bangladesh"    "Sri.Lanka."    "Thailand"
## [53] "Malaysia"      "Singapore"     "Indonesia"     "Philippines"
## [57] "China.PR"      "Korea"         "Hong.Kong"     "Taiwan"
## [61] "Japan"         "Australia"     "New.Zealand"   "Morocco"
## [65] "Algeria"       "Tunisia"       "Egypt"         "Cameroon"
## [69] "Senegal"       "Sierra.Leone"  "Cote.d.Ivoire" "Ghana"
## [73] "Nigeria"       "Benin"         "Congo"         "Kenya"
## [77] "Tanzania"      "Mozambique"    "South.Africa"  "Zambia"
```

Graficamos las series de tiempo de los países para tener una idea de qué esté pasando.

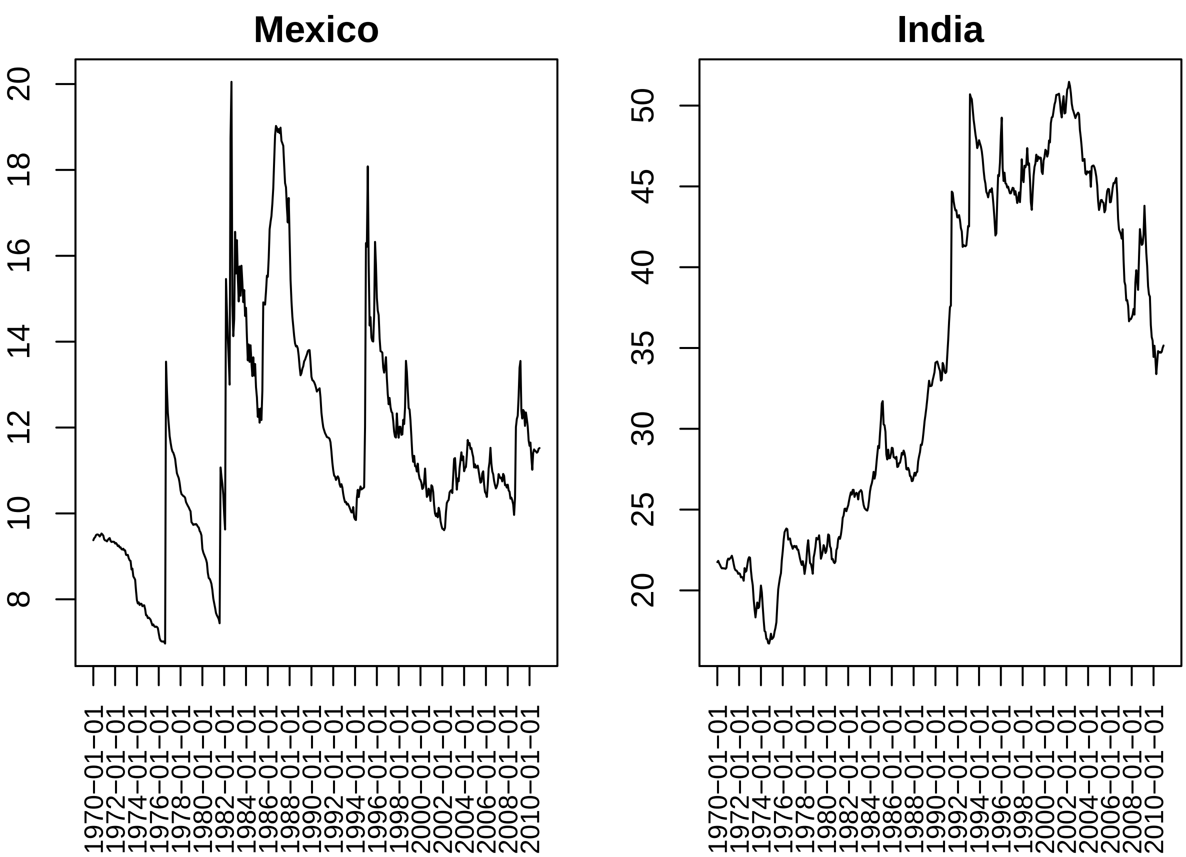
```r
# SE parte el plot en 8 pedazos
par(mfrow=c(4,2))

# Se grafican series de tiempo de los tipos de cambio
ts.plot(as.ts(data[,1:10],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,11:20],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,21:30],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,31:40],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,41:50],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,51:60],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,61:70],start=c(1970,1),end=c(2010,12),frequency=12))
ts.plot(as.ts(data[,71:80],start=c(1970,1),end=c(2010,12),frequency=12))
```



Se puede ver que los tipos de cambio siguen trayectorias muy distintas. Exploremos más a fondo un par de éstas:

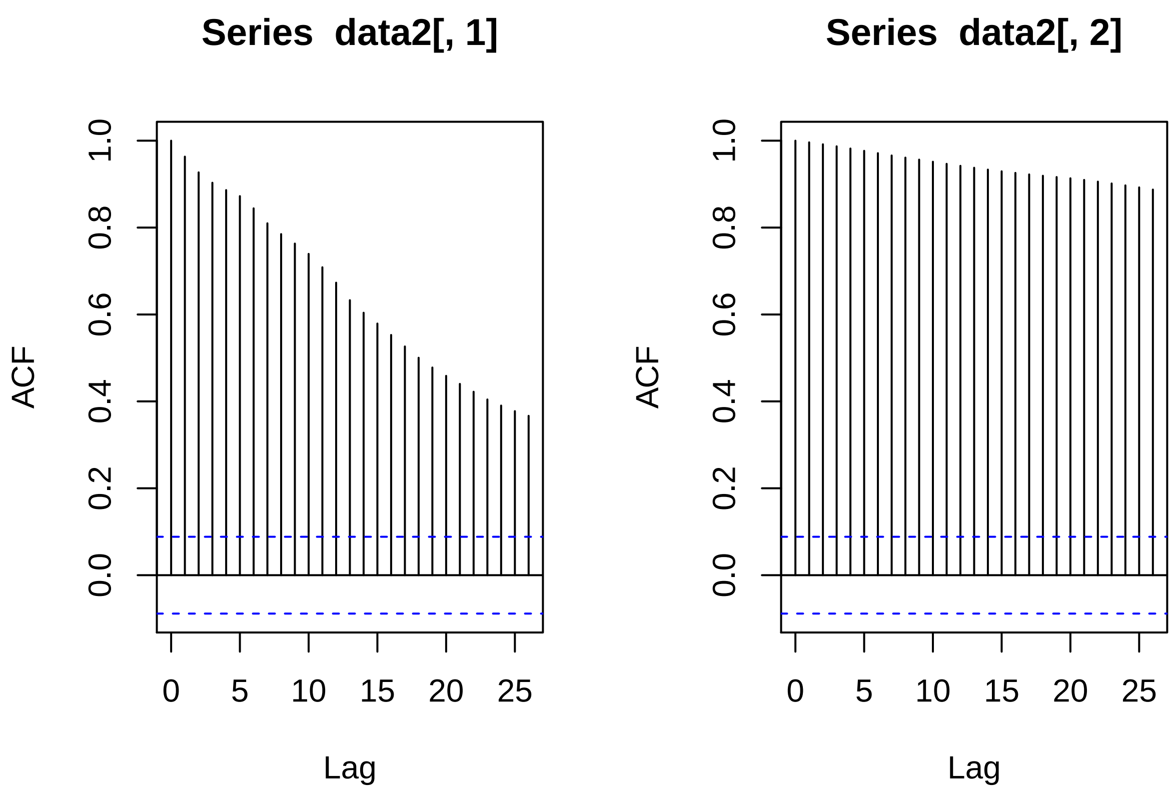
2.2. México vs India

```
# Encontramos las columnas correspondientes a México y a la India
cols<-match(c('Mexico','India'),colnames(data))

# Extraemos los datos
data2<-data[,cols]

# Graficamos las series de tiempo por serparado
par(mfrow=c(1,2))
par(mar=c(5.1,2.1,1.6,1.6))
plot(as.ts(data2$Mexico,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main="Mex
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
plot(as.ts(data2$India,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main="India
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
```



Veamos la FAC y la FACP de las series de tiempo

```
par(mfrow=c(1,2))
acf(data2[,1])
acf(data2[,2])
```



ambas series presentan una fuerte autocorrelación de orden alto. Esto indica que sí hay dependencia entre los renglones de los datos.

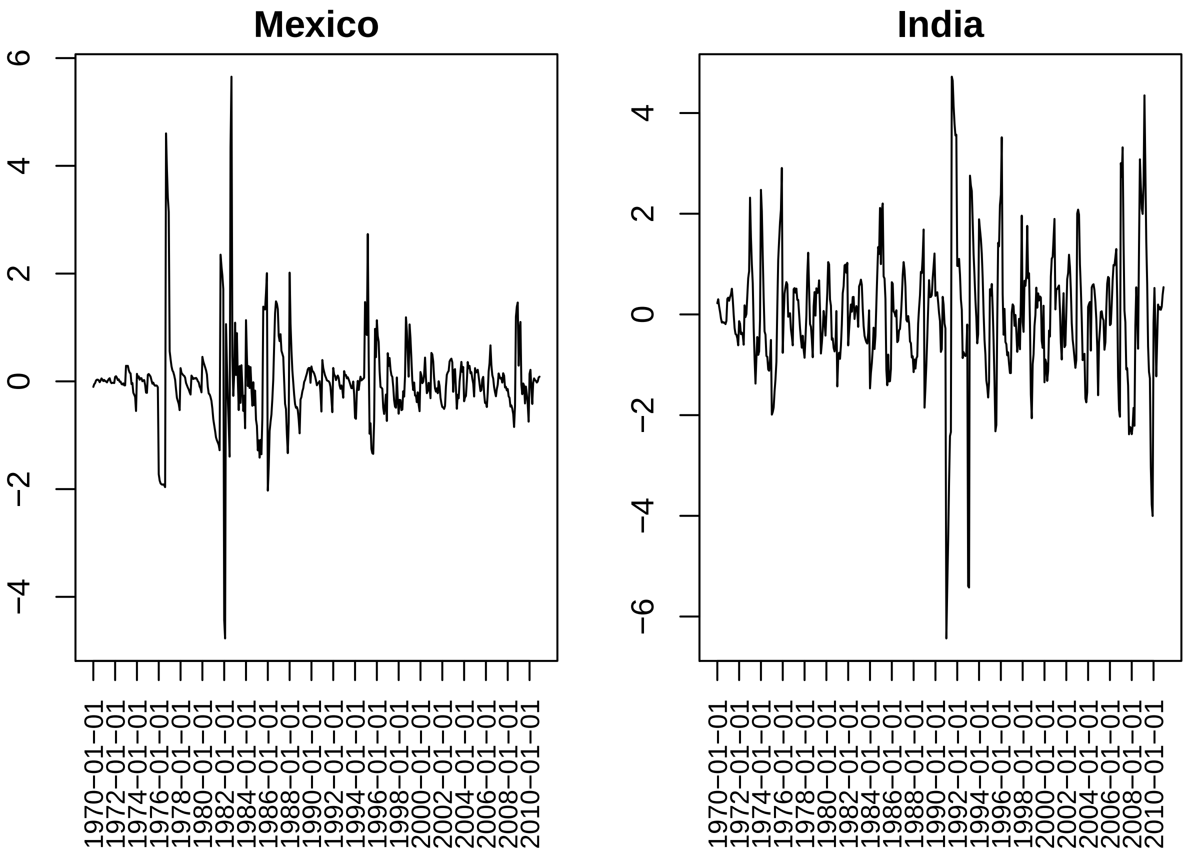
De esta manera, como se puede notar, las dos series no parecen cumplir con los supuestos establecidos para realizar ACP que son:

- Media constante por renglón (estas dos series no tienen la misma media a lo largo del tiempo)
- Varianza constante por renglón (la volatilidad de los tipos de cambio ha variado a lo largo del tiempo, sobre todo para México)
- Homogeneidad en los renglones (si se cambia el orden de las observaciones no obtiene la misma estructura para las series de tiempo.)
- Simetría estocástica (lo que pase entre 2 renglones no es ajeno entre sí ya que existe autorcorrelación)

Tratemos de extraer la media por año y centrar los datos con este valor para estabilizar la series.

```
# Se agrega la variable del año
data2$year<-as.numeric(format(fechas,"%Y"))

# Se extraen las medias por año
data_means<-data2%>%group_by(year)%>%summarise(Mean.Mex=mean(Mexico),Mean.India=mean(India))

# Se agregan las medias dependiendo del año
data2<-left_join(data2,data_means, by="year")

# Se restan las medias
data2<-data2%>%
  mutate(Mex.Est=Mexico-Mean.Mex,
         India.Est=India-Mean.India)

# Graficamos los datos centrados por año
```

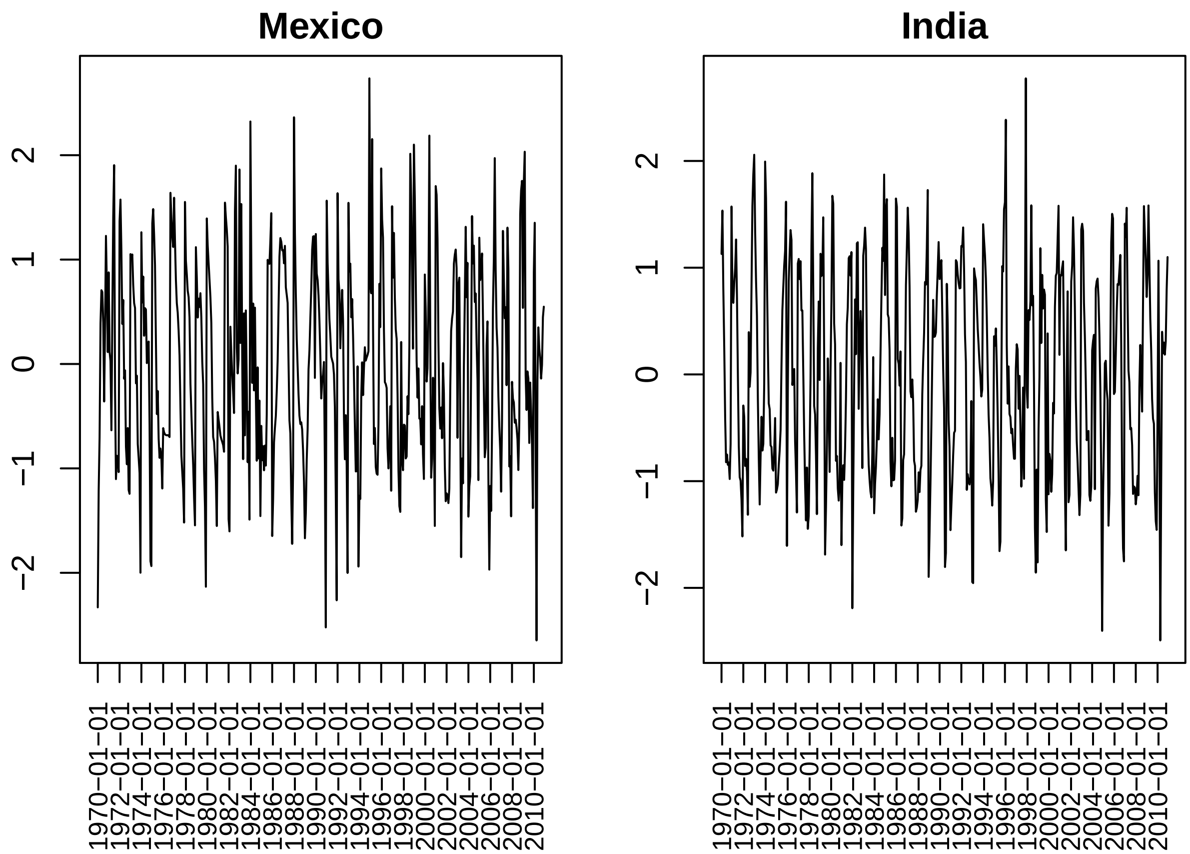
```
par(mfrow=c(1,2))
par(mar=c(5.1,2.1,1.6,1.6))
plot(as.ts(data2$Mex.Est,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main="Me
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
plot(as.ts(data2$India.Est,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main="
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
```



Como podemos ver, ya se solucionó el hecho de que la serie oscile alrededor del cero, sin embargo seguimos teniendo problemas con la volatilidad de las mismas. Apliquemos la misma estrategia estandarizando los datos con la desviación estándar de cada año.

```
# Se extraen las desviaciones estándar por año
data_sd<-data2%>%
  select(year,Mexico, India)%>%
  group_by(year)%>%
  summarise(sd.Mex=sd(Mexico),sd.India=sd(India))

# Se agregan las desviaciones estándar dependiendo del año
data2<-left_join(data2,data_sd, by="year")

# Se restan las medias
data2<-data2%>%
  mutate(Mex.Est2=(Mexico-Mean.Mex)/sd.Mex,
         India.Est2=(India-Mean.India)/sd.India)

# Graficamos los datos centrados por año
par(mfrow=c(1,2))
```

```r
par(mar=c(5.1,2.1,1.6,1.6))
plot(as.ts(data2$Mex.Est2,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main="Me
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
plot(as.ts(data2$India.Est2,start=c(1970,1),end=c(2010,12),frequency=12),ylab="",xaxt='n',xlab="",main=
axis(1,at=seq(1,length(fechas),24),labels=fechas[seq(1,length(fechas),24)],las=2,cex.axis=0.8)
```



Ahora si, las medias de las variables parecen ser constantes en el tiempo y su variabilidad también. De esta manera tenemos dos opciones:

- modificar los datos para que se cumplan los supuestos
- modificar los supuestos del modelo para que se adecúen a los datos.

En esta ocasion, vamos a modificar los supuestos del análisis de componentes principales para que se adecúen a los datos.

3. Inferencia de SPCA

Sea $x_{m,t}$ una observación de los datos, donde $m = 1, 2, ...12$ y $t = 1970, 1971, ..., 2010$.

Supuestos:

- $\mu_t = E[X_{mt}]$ tal que $m = 1, 2, ..., 12$.
- $\Sigma_t = var[X_{mt}]$ tal que $m = 1, 2, ...12$.
- $X_{mt,\cdot} \sim N_p(\boldsymbol{X}|\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t})$ donde $\boldsymbol{\mu_t}$ y $\boldsymbol{\Lambda_t}$ son desconocidos $\forall t = 1970, ..., 2010$.

Con base en estos supuestos, esto es equivalente a aplicar un ACP por año para los datos en la muestra. Vamos a estandarizar todas las variables por año.

3.1. Estandarización de los datos

```
# Extrae el año de cada observación
data$year=as.numeric(format(fechas,"%Y"))

# Se extraen las medias por año por variable
data_mean<-data%>%
  group_by(year)%>%
  summarise_all(mean)

colnames(data_mean)[2:ncol(data_mean)]<-paste("mean",colnames(data_mean)[2:ncol(data_mean)],sep=".")


# Se extraen las desviaciones estándar por año por variable
data_sd<-data%>%
  group_by(year)%>%
  summarise_all(sd)

colnames(data_sd)[2:ncol(data_sd)]<-paste("sd",colnames(data_sd)[2:ncol(data_sd)],sep=".")

# Funcion para estandarizar datos.
estandarizar<-function(country){

  aux<-select(data,year,contains(country))
  aux.mean<-select(data_mean,year,contains(country))
  aux.sd<-select(data_sd,year,contains(country))

  aux<-left_join(aux,aux.mean,by="year")%>%
    left_join(aux.sd,by="year")

  colnames(aux)<-c("year","obs","mean","sd")

  aux<-aux%>%
    mutate(estand=(obs-mean)/sd)

  return(aux$estand)
}

# Se estandariza cada uno de los tipos de cambio con media y varianza anual.
data_estand<-lapply(colnames(data)[1:(ncol(data)-1)],estandarizar)

# Se asignan nombres a los elementos de la lista
names(data_estand)<-colnames(data)[1:(ncol(data)-1)]

# Se convierte a dataframe y se elimina nicaragua.
data_estand<-as_data_frame(data_estand)%>%select(-Nicaragua)
```

Nota: la serie de tiempo de Nicaragua tiene faltantes en el inicio de la muestra que están llenados como ceros, por lo tanto generan errores al estandarizar. Se elimina nicaragua de la muestra.

3.2. Actualizacion bayesiana de parámetros

El desconocimiento acerca de $(\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t})$ se expresa como

$$(\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t}) \sim \text{N-Wi}(\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t} | \boldsymbol{m}_{0t}, s_{0t}, a_{0t}, \boldsymbol{B}_{0t}),$$

donde $\boldsymbol{m}_{0t} = 0, s_{0t} = 1/2, a_{0t} = 1, \boldsymbol{B}_{0t} = I$; es decir, se tiene una distribución inicial no informativa.

De esta manera, la distribución posterior de (μ_t, Λ_t) a partir de la consolidación de la información contenida en los datos y la información complementaria está dada por

$$(\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t} | \boldsymbol{X_t}) \sim \text{N-Wi}(\boldsymbol{\mu_t}, \boldsymbol{\Lambda_t} | \boldsymbol{m}_{nt}, s_{nt}, a_{tn}, \boldsymbol{B}_{nt}),$$

con

$$\begin{aligned}
\boldsymbol{m}_{nt} &= \frac{s_{0t} m_{0t} + n\bar{x}}{s_{nt}}, \\
s_{nt} &= s_{0t} + n, \\
a_{nt} &= a_{0t} + \frac{n}{2}, \\
\boldsymbol{B}_{nt} &= \boldsymbol{B}_{0t} + \frac{n}{2}\left[S + \frac{s_{0t}}{s_{nt}}(\bar{x} - m_{0t})(\bar{x} - m_{0t})'\right]
\end{aligned}$$

```r
# Fijar hiperparámetros
  a0 <- 1
  s0 <- 1/2
  m0 <- matrix(0,ncol=1,nrow=ncol(data_estand))
  B0 <- diag(1,ncol=ncol(data_estand),nrow=ncol(data_estand))


# Función para calcular la posterior
gaussian.posterior <- function(data,m0,s0,a0,B0){

  # Media de los datos
  xbar <- as.matrix(colMeans(data))

  # Mat. de var y cov de los datos.
  S <- cov(data)

  # No. de obs.
  n <- nrow(data)

  # No. de variables.
  p <- ncol(data)

  # Parámetros actualizados de la posterior.
  sn <- s0 + n
  an <- a0 + n/2
  mn <- (s0*m0 + n*xbar)/sn
  Bn <- B0 + (n/2)*(S + (s0/sn)*(xbar-m0)%*%t(xbar-m0))
```

```
  # Salida (parámetros actualizados)
  output <- list(mn=mn,sn=sn,an=an,Bn=Bn)
  return(output)
}

# Calculamos los hiperparams actualizados para la posterior
output <- gaussian.posterior(data_estand,m0,s0,a0,B0)
```

3.3. Simulación

Una vez realizada la actualización bayesiana de los hiperparámetros, simulamos M observaciones para (μ_t, Λ_t) como

$$
\begin{aligned}
\boldsymbol{\Lambda_t}^{(m)} &\sim \mathrm{Wi}(\boldsymbol{\Lambda_t}|a_{nt}, \boldsymbol{B}_{nt}), \\
\boldsymbol{\mu_t}^{(m)}|\boldsymbol{\Lambda_{nt}}^{(m)} &\sim \mathrm{N}(\boldsymbol{\mu_t}|\boldsymbol{m}_{nt}, s_{nt}\boldsymbol{\Lambda_t}^{(m)}).
\end{aligned}
$$

Con estos parámetros, obtenemos simulaciones de los eigenvalores y eigenvectores $(e_{jt}^{(m)}, \boldsymbol{v}_{jt}^{(m)})_{j=1}^{p}$, de la matriz de varianzas y covarianzas de los datos simulada en el paso anterior (inverso de $\Lambda_t^{(m)}$.

Por último, obtenemos simulaciones de las componentes principales como $\boldsymbol{c}_{jt,\cdot}^{(m)} = \boldsymbol{X}\boldsymbol{v}_{jt}^{(m)}$.

```
# Función para simular los datos

  # Se fija el no de simulaciones
  M <- 10000

  # Matriz para guardar medias
  mu.sim <- matrix(NA,nrow=M, ncol=ncol(data_estand))

  # Arreglo para guardar precisiones
  Lambda.sim <- array(NA,dim=c(M,ncol(data_estand),ncol(data_estand)))

  # Matriz para guardar valores propios.
  e.sim <- matrix(NA,nrow=M, ncol=ncol(data_estand))

  # Arreglo para gaurdar vectores propios
  V.sim <- array(NA,dim=c(M,ncol(data_estand),ncol(data_estand)))

  # Arreglo para guardar componentes principales
  C.sim <- array(NA,dim=c(M,nrow(data_estand),ncol(data_estand)))

  # Se convierten los datos a matriz.
  X <- as.matrix(data_estand)

  # En cada iteración:
  for(m in 1:M){

    # Se simulan valores (mu,Lambda)
    Lambda.sim[m,,] <- rWishart(1, output$an, output$Bn)
    mu.sim[m,] <- mvrnorm(1, mu=output$mn, Sigma=solve(output$sn*Lambda.sim[m,,]), tol = 1e-6)
```

```
    # Simulacion de eigenvalores y eigenvectores (e,V)
    eigen_aux <- eigen(solve(Lambda.sim[m,,]))
    e.sim[m,] <- eigen_aux$values
    V.sim[m,,] <- eigen_aux$vectors

    # Simulación de componentes principales.
    C.sim[m,,] <- X %*% V.sim[m,,]
  }
```

4. Resultados

4.1. Eigenvalores

```
multi.hist(e.sim)
```

Todas las distribuciones de las simulaciones correspondientes a los valores propios se ven mucho más homogenes y simétricas que antes.

4.2. ¿Qué economía tiene el mayor peso esperado en la descomposición PCA?

Los pesos de las componentes están representados por los eigenvectores. Utilizamos la moda de la distribución para obtener la matriz de eigenvectores promedio.

```
# Estimador puntual de la matriz de eigenvectores.
V.mode<-matrix(NA,ncol(data_estand),ncol(data_estand))

# Se extrae la moda para cada elemento de la matriz de eigenvectores
for(i in 1:ncol(data_estand)){
```

```
  for(j in 1:ncol(data_estand)){

    V.mode[i,j]<-getmode(V.sim[,i,j])
  }
}

colnames(V.mode)<-paste("PC",1:ncol(data_estand),sep="_")
rownames(V.mode)<-colnames(data_estand)

# Se busca el peso más grande en cada una de las componentes.
peso_max<-data_frame(CP=colnames(V.mode),
                     Economia=apply(V.mode,2,function(x){rownames(V.mode)[which.max(x)]}))

kable(cbind(peso_max[1:20,],
            peso_max[21:40,],
            peso_max[41:60,],rbind(peso_max[61:79,],rep(NA,2))),
      type='latex',
      caption='Economia con mayor peso por componente principal')
```

Tabla 1: Economia con mayor peso por componente principal

CP	Economia	CP	Economia	CP	Economia	CP	Economia
PC_1	Germany	PC_21	Spain	PC_41	Egypt	PC_61	Russia
PC_2	Austria	PC_22	Switzerland	PC_42	Jordan	PC_62	Canada
PC_3	EU12	PC_23	Cote.d.Ivoire	PC_43	Argentina	PC_63	Paraguay
PC_4	Netherlands	PC_24	Sweden	PC_44	Philippines	PC_64	Poland
PC_5	Denmark	PC_25	Algeria	PC_45	Peru	PC_65	Sierra.Leone
PC_6	France	PC_26	Korea	PC_46	Poland	PC_66	Egypt
PC_7	Tunisia	PC_27	Colombia	PC_47	Guatemala	PC_67	Guatemala
PC_8	Denmark	PC_28	Switzerland	PC_48	China.PR	PC_68	El.Salvador
PC_9	Tunisia	PC_29	Spain	PC_49	Kenya	PC_69	Uruguay
PC_10	Finland	PC_30	Greece	PC_50	India	PC_70	Guatemala
PC_11	Portugal	PC_31	Singapore	PC_51	Ghana	PC_71	Tanzania
PC_12	Senegal	PC_32	Algeria	PC_52	South.Africa	PC_72	Venezuela.
PC_13	Morocco	PC_33	El.Salvador	PC_53	Trin.Tobago	PC_73	South.Africa
PC_14	Norway	PC_34	Malaysia	PC_54	Hong.Kong	PC_74	Paraguay
PC_15	Benin	PC_35	Trin.Tobago	PC_55	Jamaica	PC_75	Colombia
PC_16	Thailand	PC_36	Chile	PC_56	El.Salvador	PC_76	Sierra.Leone
PC_17	Congo	PC_37	Sri.Lanka.	PC_57	Kenya	PC_77	Saudi.Arabia
PC_18	Finland	PC_38	Canada	PC_58	Mozambique	PC_78	Japan
PC_19	Chile	PC_39	Argentina	PC_59	Israel	PC_79	Austria
PC_20	Israel	PC_40	Guatemala	PC_60	Australia	NA	NA

```
# Tabla de frecuencias para pesos por país
frec_peso_max<-table(peso_max$Economia)
frec_peso_max<-data_frame(Economia=names(frec_peso_max),
                          Frecuencia=frec_peso_max)%>%
  arrange(desc(Frecuencia))

if(nrow(frec_peso_max)%%2==0){
kable(cbind(frec_peso_max[1:ceiling(nrow(frec_peso_max)/2),],
            frec_peso_max[(ceiling(nrow(frec_peso_max)/2)+1):nrow(frec_peso_max),]),
```

```
      type='latex',
      caption='Frecuencias de Economías con Mayor Peso')
}else{
  kable(cbind(frec_peso_max[1:ceiling(nrow(frec_peso_max)/2),],
            rbind(frec_peso_max[(ceiling(nrow(frec_peso_max)/2)+1):nrow(frec_peso_max),],rep(NA,2))),
      type='latex',
      caption='Frecuencias de Economías con Mayor Peso')
}
```

Tabla 2: Frecuencias de Economas con Mayor Peso

Economia	Frecuencia	Economia	Frecuencia
Guatemala	4	Germany	1
El.Salvador	3	Ghana	1
Algeria	2	Greece	1
Argentina	2	Hong.Kong	1
Austria	2	India	1
Canada	2	Jamaica	1
Chile	2	Japan	1
Colombia	2	Jordan	1
Denmark	2	Korea	1
Egypt	2	Malaysia	1
Finland	2	Morocco	1
Israel	2	Mozambique	1
Kenya	2	Netherlands	1
Paraguay	2	Norway	1
Poland	2	Peru	1
Sierra.Leone	2	Philippines	1
South.Africa	2	Portugal	1
Spain	2	Russia	1
Switzerland	2	Saudi.Arabia	1
Trin.Tobago	2	Senegal	1
Tunisia	2	Singapore	1
Australia	1	Sri.Lanka.	1
Benin	1	Sweden	1
China.PR	1	Tanzania	1
Congo	1	Thailand	1
Cote.d.Ivoire	1	Uruguay	1
EU12	1	Venezuela.	1
France	1	NA	NA

Los resultados son completamente diferentes a los obtenidos sin estandarizar las variables con sus medias y varianzas anuales. Ahora las simulaciones son mucho más estables.

4.3. ¿Qué economía tiene la mayor consistencia en estimacion de los cjs correspondientes?

```
# Varianza de los eigenvectores.
V.var<-matrix(NA,ncol(data_estand),ncol(data_estand))

# Se calcula la varianza de cada elemento de los eigenvectores
```

```r
for(i in 1:ncol(data_estand)){

  for(j in 1:ncol(data_estand)){

    V.var[i,j]<-var(C.sim[,i,j],na.rm=TRUE)
  }
}

colnames(V.var)<-paste("PC",1:ncol(data_estand),sep="_")
rownames(V.var)<-colnames(data_estand)

# Se busca el menor varianza en cada una de las componentes.
var_min<-data_frame(CP=colnames(V.var),
                    EcoVarMin=apply(V.var,2,function(x){rownames(V.var)[which.min(x)]}))

kable(cbind(var_min[1:20,],var_min[21:40,],var_min[41:60,],rbind(var_min[61:79,],rep(NA,2))),
      type='latex',
      caption='Economia con Menor Varianza por Componente Principal')
```

Tabla 3: Economia con Menor Varianza por Componente Principal

CP	EcoVarMin	CP	EcoVarMin	CP	EcoVarMin	CP	EcoVarMin
PC_1	Ghana	PC_21	Egypt	PC_41	Cameroon	PC_61	Cameroon
PC_2	Ghana	PC_22	Cote.d.Ivoire	PC_42	Cameroon	PC_62	Cameroon
PC_3	Ghana	PC_23	Cote.d.Ivoire	PC_43	Cameroon	PC_63	Turkey
PC_4	Ghana	PC_24	Cote.d.Ivoire	PC_44	Cameroon	PC_64	Turkey
PC_5	Ghana	PC_25	Cote.d.Ivoire	PC_45	Cameroon	PC_65	Turkey
PC_6	Senegal	PC_26	Cote.d.Ivoire	PC_46	Cameroon	PC_66	EU12
PC_7	Ghana	PC_27	Cote.d.Ivoire	PC_47	Cameroon	PC_67	EU12
PC_8	Ghana	PC_28	Cote.d.Ivoire	PC_48	Cameroon	PC_68	EU12
PC_9	Turkey	PC_29	Cote.d.Ivoire	PC_49	Cameroon	PC_69	EU12
PC_10	Turkey	PC_30	Cote.d.Ivoire	PC_50	Cameroon	PC_70	EU12
PC_11	Turkey	PC_31	Cameroon	PC_51	Cameroon	PC_71	South.Africa
PC_12	Turkey	PC_32	Cameroon	PC_52	Cameroon	PC_72	South.Africa
PC_13	Turkey	PC_33	Cameroon	PC_53	Cameroon	PC_73	South.Africa
PC_14	Turkey	PC_34	Cameroon	PC_54	Cameroon	PC_74	Greece
PC_15	Ghana	PC_35	Morocco	PC_55	Cameroon	PC_75	Greece
PC_16	Egypt	PC_36	Morocco	PC_56	Cameroon	PC_76	Turkey
PC_17	Cote.d.Ivoire	PC_37	Morocco	PC_57	Cameroon	PC_77	Turkey
PC_18	Egypt	PC_38	Morocco	PC_58	Cameroon	PC_78	Turkey
PC_19	Cote.d.Ivoire	PC_39	Morocco	PC_59	Cameroon	PC_79	Tanzania
PC_20	Cote.d.Ivoire	PC_40	Greece	PC_60	Cameroon	NA	NA

```r
# Tabla de frecuencias para pesos por país
frec_var_min<-table(var_min$EcoVarMin)
frec_var_min<-data_frame(Economia=names(frec_var_min),
                         Frecuencia=frec_var_min)%>%
  arrange(desc(Frecuencia))

kable(frec_var_min,
      type='latex',
      caption='Frecuencia de Economías con menor varianza')
```

Tabla 4: Frecuencia de Economas con menor varianza

Economia	Frecuencia
Cameroon	26
Cote.d.Ivoire	12
Turkey	12
Ghana	8
EU12	5
Morocco	5
Egypt	3
Greece	3
South.Africa	3
Senegal	1
Tanzania	1

Disminuyó el número de paises con mayor inconsistencia en los componentes principales. Esto es gracias a que se estandarizan los datos por variable.