Series Temporales y Predicción

Práctica 4

Polinomio característico y estimación de los parámetros

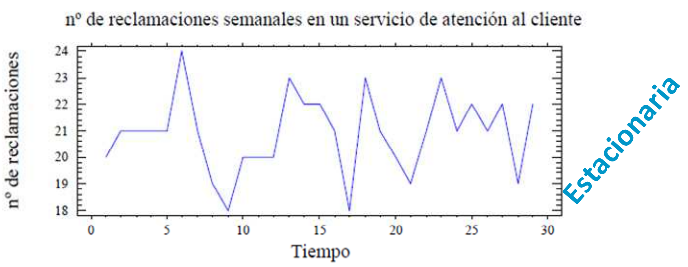
Solución orientativa a los ejercicios propuestos

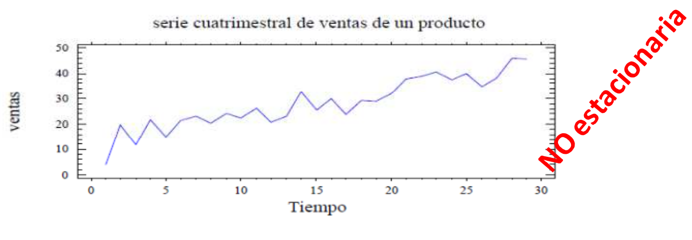
1. Raíces del polinomio característico y estacionariedad

La idea de estacionariedad es que el comportamiento (probabilístico) de la serie en el futuro será análogo al comportamiento en el pasado.

Una serie es estacionaria cuando la media y la variabilidad son constantes a lo largo del tiempo.

Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.





Un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ direm que és *autoregressiu de primer ordre* ($AR(1)$) si compleix una relació del tipus següent:

$$\boxed{X_n = \phi X_{n-1} + Z_n, \quad (1)}$$

on $\{Z_n, n \in \mathbf{Z}\}$és un soroll blanc i ϕ és una constant, $|\phi| < 1$ Noteu que si $\phi = 1$ s'obtindria una passejada aleatòria, que com hem dit no és estacionari.

Per treballar amb aquests processos és convenient introduir *l'operador de retard* (*backward*) $\boldsymbol{B}$ definit per

$$BX_n = X_{n-1}.$$

L'equació d'autoregressió (1) s'escriurà:

$$(I - \phi B)X_n = Z_n.$$

Si posem $\Phi(x) = 1 - \phi x$, podem escriure:

$$\Phi(B)X = Z.$$

El polinomi $\Phi(x) = 1 - \phi x$ s'anomena *polinomi característic* del procés. Noteu que la condició $|\phi| < 1$ implica que l'arrel d'aquest polinomi, $\eta = 1 / \phi$ valor absolut $|\eta| < 1$ (òbviament suposem $\phi \neq 0$.)

Si introduïm ara els operadors de retard iterats:

$$B^k X_n = B^{k-1}(BX_n) = X_{n-k},$$

l'expressió (2) pot escriure's

$$X_n = \sum_{k=0}^{\infty} \phi^k B^k (Z_n).$$

Notem que, formalment, el que hem fet ha estat invertir l'operador $I - \phi B$, ja que com que $|\phi| < 1, \forall x \in [-1, 1]$, $|\phi x| < 1$, aleshores,

$$(1 - \phi x)^{-1} = \sum_{k=0}^{\infty} \phi^k x^k$$

(suma d'una progressió geomètrica de raó de valor absolut < 1).

3.2.4 Procés autoregressiu d'ordre p ($AR(p)$)

Es tracta d'un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ que compleix una equació de l'estil

$$\boxed{X_n = \phi_1 X_{n-1} + \cdots + \phi_p X_{n-p} + Z_n,}$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc.

S'escriu també

$$\Phi(B)X = Z,$$

on

$$\Phi(x) = 1 - \phi_1 x - \cdots - \phi_p x^p.$$

Les autocorrelacions han de complir l'equació en diferències

$$\Phi(B)\rho_k = 0,$$

que per a $k = 1, \ldots, p$ dóna les equacions de Yule-Wolker

$$\left.\begin{array}{l} \rho_1 = \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\ \rho_2 = \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\ \vdots \\ \rho_p = \phi_1\rho_{p-1} + \cdots + \phi_p \end{array}\right\}$$

Raonant igual que al procés AR(2), perquè les solucions de l'equació en diferències no explotin, cal que les arrels del polinomi caracterpistic $\Phi(x)$ estiguin fora del cercle unitat. També en aquest cas serà un procés causal.

Bibliografia: https://www.raco.cat/index.php/ButlletiSCM/article/view/149407/201308

Práctica 1.1

Considera el proceso AR(1) dado por $X_t = 0.6 \cdot X_{t-1} + s_t$. ¿Es un proceso estacionario?

Solución Orientativa

Para que un proceso AR(1) sea estacionario es necesario que $|\phi| < 1$. En nuestro caso $\phi = 0{,}6 < 1$ Por lo que el proceso es estacionario.

Práctica 1.2

Considera el proceso AR(2) dado por $X_t = 1.095445\ X_{t-1} - 0.3\ X_{t-2} + s_t$. ¿Qué valores toma las raíces de su polinomio característico? ¿Es un proceso estacionario? ¿Cómo es su correlograma (gráfica ACF)?

Solución Orientativa

$X_t - 1.095445\ X_{t-1} + 0.3\ X_{t-2} = St$

$(\ I - 1.095445\ B + 0.3\ B^2)\ X_t = S_t$

Por lo que el polinomio característico del proceso es

$\Phi(x) = 1 - 1.095445\ x + 0.3\ x^2$

Para encontrar las raíces en R utilizaremos la función polyroot:

https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/polyroot

Código R

```
phi1=1.095445; phi2=-0.3;
raices = polyroot(c(1, -phi1, -phi2))
raices
Mod(raices)
```

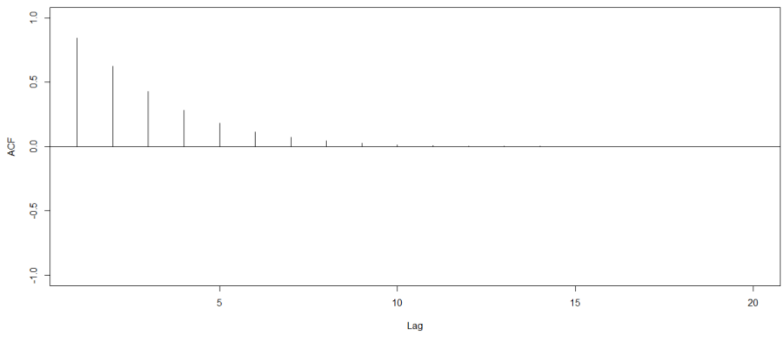
```
> raices = polyroot(c(1, -phi1, -phi2))
> raices
[1] 1.825742+0.000837i 1.825742-0.000837i
> Mod(raices)
[1] 1.825742 1.825742
```

Como las raíces se salen del círculo unidad sí es un proceso estacionario.

Correlograma

Código R

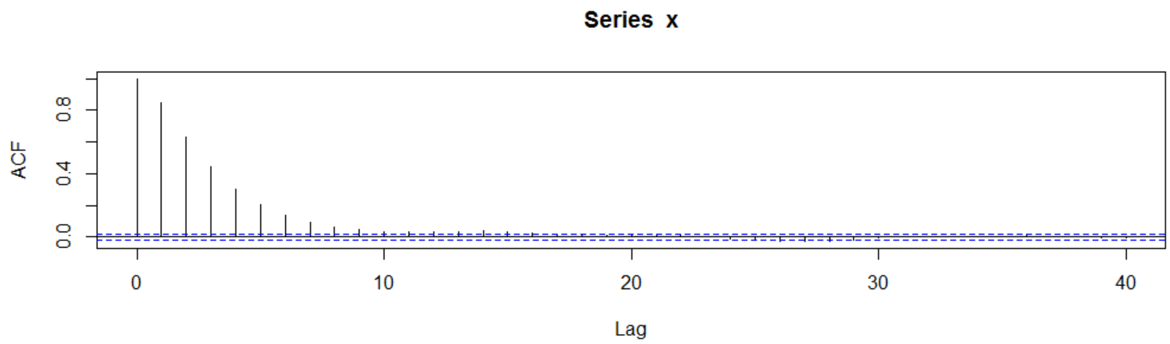
```
rho=NULL;
rho1=phi1/(1-phi2); rho2=(phi2*(1-phi2)+phi1^2)/(1-phi2)
rho[1]=rho1; rho[2]=rho2
max.lag=20
for (k in 3:max.lag) rho[k]=phi1*rho[k-1]+phi2*rho[k-2]
rho # presentar los valores
plot(y=rho,x=1:max.lag,type='h',ylab='ACF',xlab='Lag',ylim=c(-1,+1)); abline(h=0)
```

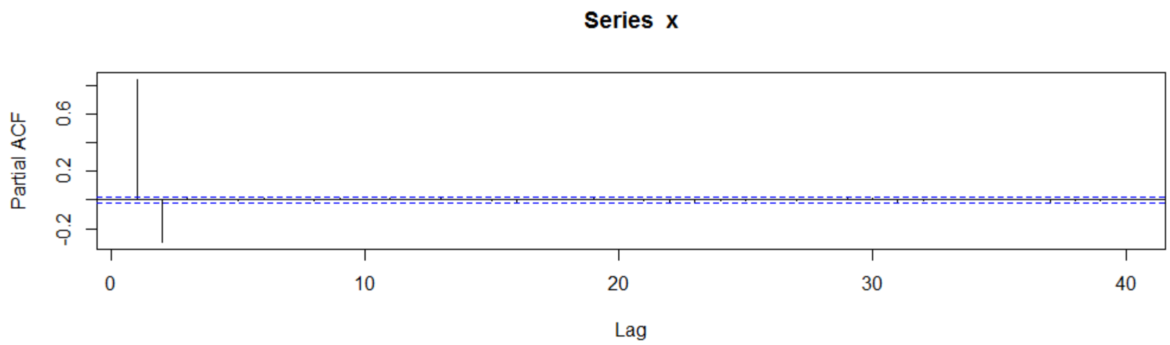



```
> rho # presentar los valores
 [1] 0.84265000000 0.62307672925 0.42975128767 0.28384588055 0.18201216432
 [6] 0.11423055118 0.07052963684 0.04299217267 0.02593666954 0.01551454316
[11] 0.00921432787 0.00543942645 0.00319429414 0.00186734561 0.00108728617
[16] 0.00063085852 0.00036488496 0.00021045385 0.00012107513 0.00006949499
```

Otro camino

```
x <- vector()
x[1] <- 0
x[2] <- 0
for (i in 3:10000)
{
  x[i] <- phi1*x[i-1] + phi2*x[i-2]+rnorm(1)
}
par(mfrow = c(2,1))
acf(x)
pacf(x)
```





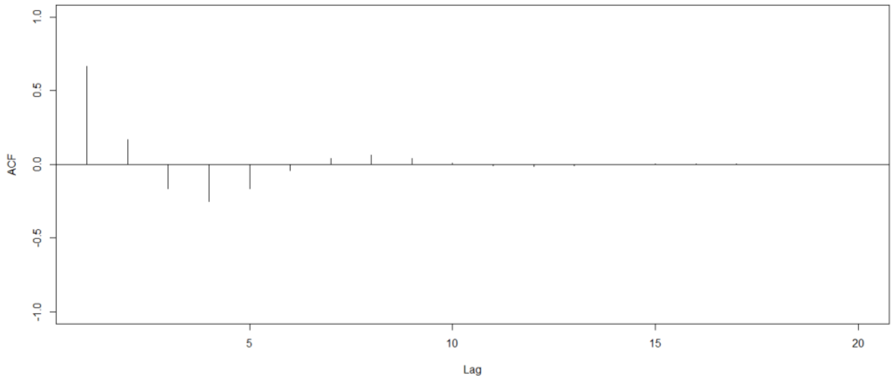
Práctica 1.3

Considerad el proceso AR(2) dado por $X_t = X_{t-1} - 0.5 \cdot X_{t-2} + S_t$. ¿Cuáles son las raíces de su polinomio característico? ¿Es un proceso estacionario? ¿Cómo es su correlograma (gráfica ACF)?

```
> raices
[1] 1+1i 1-1i
> Mod(raices)
[1] 1.414214 1.414214
```

Como las raíces se salen del círculo unidad sí es un proceso estacionario.

Correlograma




```
> rho # presentar los valores
 [1]  0.6666666667  0.1666666667 -0.1666666667 -0.2500000000 -0.1666666667
 [6] -0.0416666667  0.0416666667  0.0625000000  0.0416666667  0.0104166667
[11] -0.0104166667 -0.0156250000 -0.0104166667 -0.0026041667  0.0026041667
[16]  0.0039062500  0.0026041667  0.0006510417 -0.0006510417 -0.0009765625
```

Práctica 1.4

Suponed que las raíces del polinomio característico de un proceso AR(2) son 0.6 y 0.3. ¿Es un proceso estacionario?

No es estacionario dado que las raíces están dentro del círculo unidad.

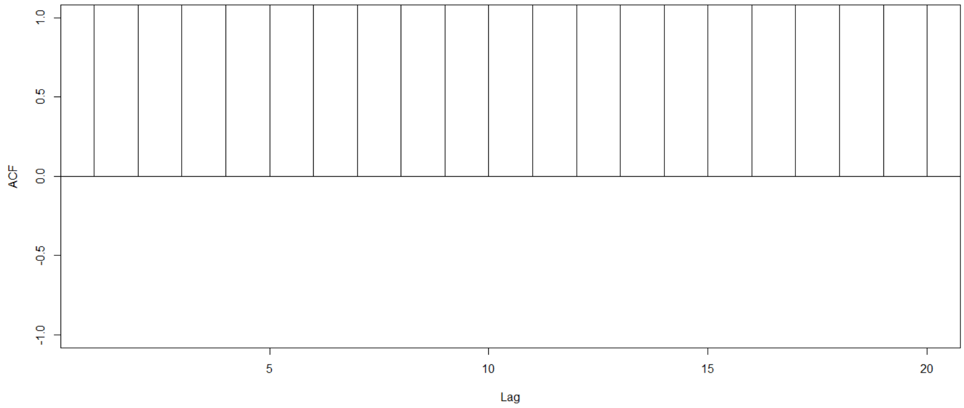
Práctica 1.5

Considerad el proceso AR(2) dado por $X_t = 1.6 \cdot X_{t-1} + 0.3 \cdot X_{t-2} + S_t$. ¿Cuáles son las raíces de su polinomio característico? ¿Es un proceso estacionario? ¿Cómo es su correlograma (gráfica ACF)? Simula el proceso y ejecuta la función arima(sim, order=c(2,0,0)). ¿Cuál es su resultado?

```
> raices
[1]  0.5651199+0i -5.8984532-0i
> Mod(raices)
[1] 0.5651199 5.8984532
```

No es estacionario dado que las raíces están dentro del círculo unidad.

Correlograma



Observamos un correlograma que no tiene sentido, con valores de autocorrelación mayores que 1!!!

```
> rho # presentar los valores
 [1]      2.285714      3.957143      7.017143     12.414571     21.968457
 [6]     38.873903     68.788782    121.724222    215.395389    381.149889
[11]    674.458439   1193.478469   2111.903083   3737.088473   6612.912482
[16]  11701.786512  20706.732164  36641.307416  64838.111515 114733.370649
```

Código R

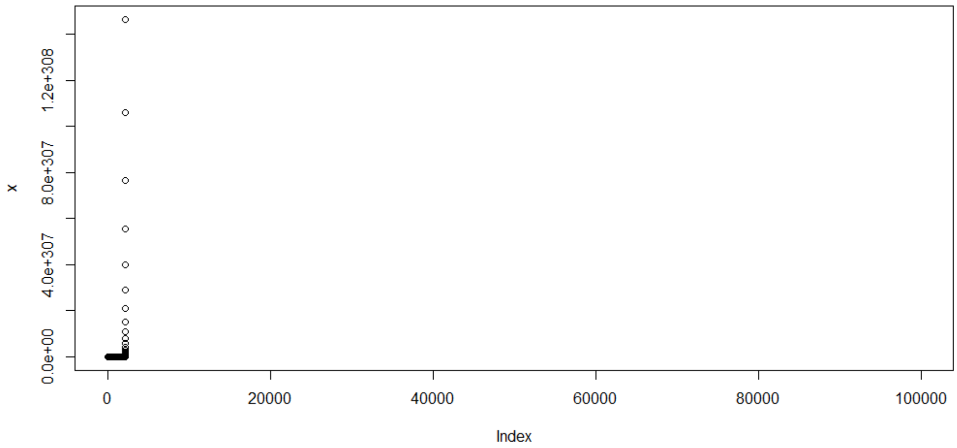
```
ar2.r <- arima.sim(model=list(ar=c(phi1, phi2)), n=100)
fit <- arima(ar2.r, order=c(2,0,0))
```

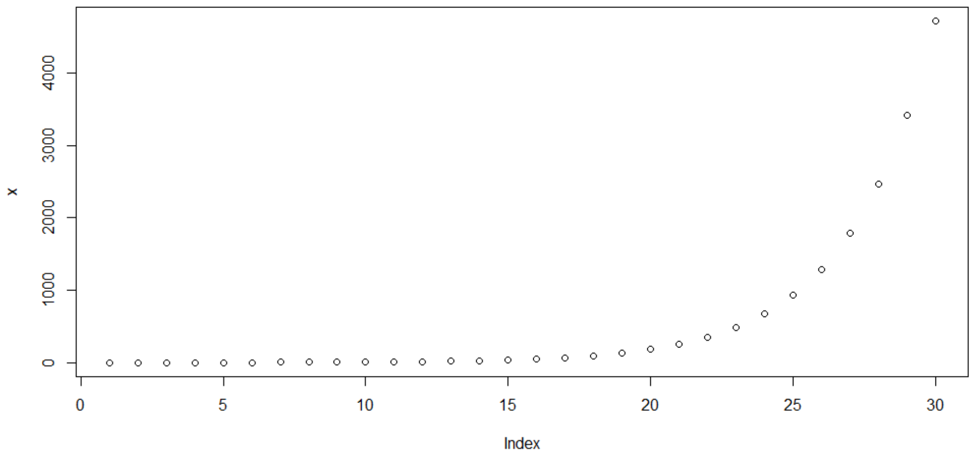
```
> ar2.r <- arima.sim(model=list(ar=c(phi1, phi2)), n=100)
Error in arima.sim(model = list(ar = c(phi1, phi2)), n = 100) :
  'ar' part of model is not stationary
> fit <- arima(ar2.r, order=c(2,0,0))
Error in NCOL(x) : objeto 'ar2.r' no encontrado
```

No podemos simular directamente con la función arima.sim dado que detecta que el proceso no es estacionario. Si realizamos la simulación, obtenemos una gráfica que crece rápidamente a números muy grandes.

Código R

```
x <- vector()
x[1] <- 0
x[2] <- 0
for (i in 3:30)
{
  x[i] <- phi1*x[i-1]-phi2*x[i-2]+rnorm(1)
}
head(x)
plot(x)
fit <- arima(x, order=c(2,0,0))
```





2. Estimación de los parámetros

Práctica 2.1

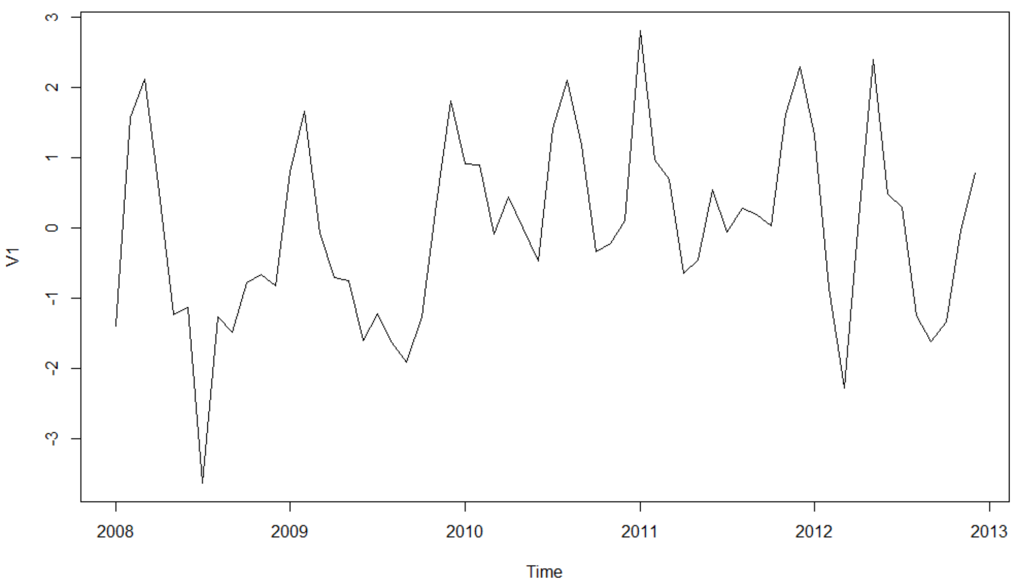
En el Campus Virtual encontraréis el fichero de datos “prac4TS.txt”, que contiene información sobre los beneficios mensuales medios de ciertas operaciones bursarias realizadas entre Enero de 2008 y Diciembre de 2012. Proponed un modelo que ajuste bien estos datos.

Solución orientativa

El primer paso es ver cómo son nuestros datos. Con el código que se muestra a continuación observamos que la serie es mensual, por lo que convertimos la serie a clase ts con frecuencia mensual y la visualizamos gráficamente.

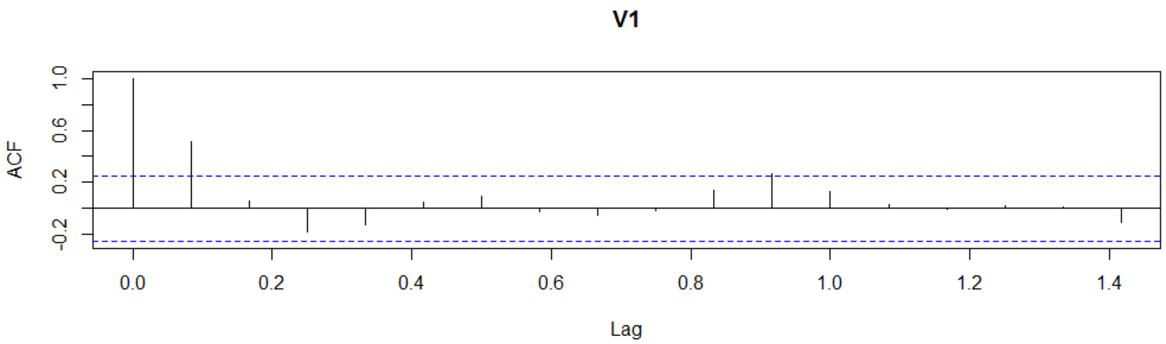
Código R

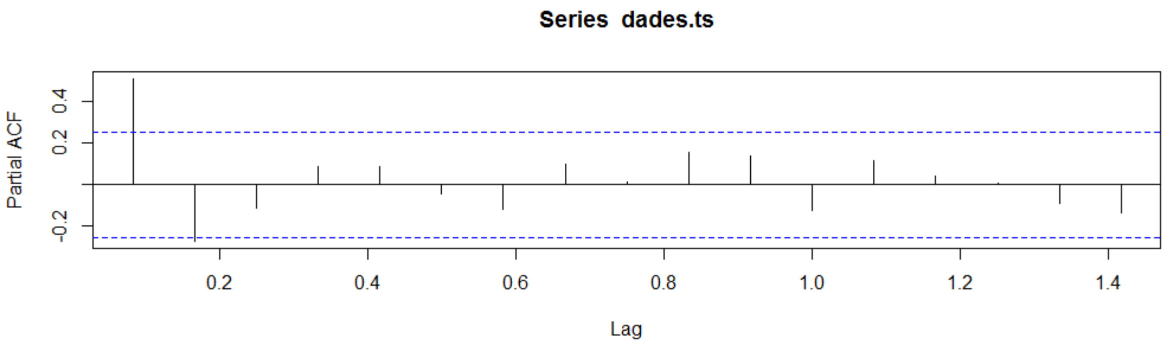
```
# Investigamos como es la serie y la visualizamos gráficamente
dades <- read.table("/Users/POR740051/Desktop/UOC/SeriesTemporalsUAB/prac4TS.txt")
head(dades)
dades.ts <- ts(dades, start=c(2008, 1), end=c(2012, 12), frequency=12)
dades.ts
plot(dades.ts)
```



En el gráfico observamos una serie que parece estacionaria, con media cercana o igual a 0 y cuya varianza podría ser considerada constante (la serie se mueve dentro de unas bandas tocándolas repetidamente en las diferentes subidas y bajadas). Debido a su estacionariedad podemos plantearnos un modelo AR o MA.

```
# Construimos el correlograma y la función de autocorrelación parcial
acf(dades.ts)
pacf(dades.ts)
```





El correlograma presenta la primera autocorrelación positiva y distinta a 0 (ρ(1) > 0) y el resto de autocorrelaciones se encuentran dentro de la banda de confianza del valor 0. Este correlograma parece indicar un proceso MA(1).

La función de autocorrelación parcial presenta la primera función negativa y distinta a 0 (α(1) > 0) y el resto de la función se encuentra dentro de la banda de confianza del valor 0. Este correlograma parece indicar un proceso AR(2).

Procés	ACF (ρ)	PACF (α)
AR(p)	Decreixement ràpid a zero	$\alpha(k) \neq 0, \; k = 1,\dots,p$ $\alpha(k) = 0, \; k = p+1,\dots$
MA(q)	$\rho(k) \neq 0, \; k = 1,\dots,q$ $\rho(k) = 0, \; k = q+1,\dots$	Decreixement ràpid a zero

 # ¿Qué proceso propone la función auto.arima?
fitarima = auto.arima(dades.ts)
coeftest(fitarima)

```
> auto.arima(dades.ts)
Series: dades.ts
ARIMA(2,0,0) with zero mean

Coefficients:
         ar1      ar2
      0.6948  -0.3221
s.e.  0.1264   0.1265

sigma^2 estimated as 1.131:  log likelihood=-88.08
AIC=182.16   AICc=182.58   BIC=188.44


z test of coefficients:

    Estimate Std. Error z value     Pr(>|z|)
ar1  0.69484    0.12636   5.499 0.0000000382 ***
ar2 -0.32210    0.12651  -2.546       0.0109 *
---
```

La función auto.arima propone un modelo AR(2):

Xt = 0,6948 · Xt-1 – 0.3221 · Xt-2 + Et

El test realizado a que los componentes son igual o distinto a 0 nos indica que con un nivel de significación de 0.05 ambos coeficientes son distintos de cero. Por lo que el modelo contemplado sería adecuado.

Ahora bien, con un nivel de significatividad del 0.01, el coeficiente que acompaña al componente Xt-2 es igual a 0. Por lo que podríamos proponer mejorar el modelo proporcionado por la auto-arima considerando un AR(1).

arima2<-arima(dades.ts,c(1,0,0), include.mean = FALSE)
arima2
coeftest(arima2)

```
> arima2<-arima(dades.ts,c(1,0,0),  include.mean = FALSE)
> arima2

Call:
arima(x = dades.ts, order = c(1, 0, 0), include.mean = FALSE)

Coefficients:
         ar1
      0.5158
s.e.  0.1103

sigma^2 estimated as 1.215:  log likelihood = -91.13,  aic = 186.26
> coeftest(arima2)

z test of coefficients:

    Estimate Std. Error z value    Pr(>|z|)
ar1  0.51579    0.11026  4.6779 0.000002898 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

El resultado obtenido nos da el siguiente proceso:

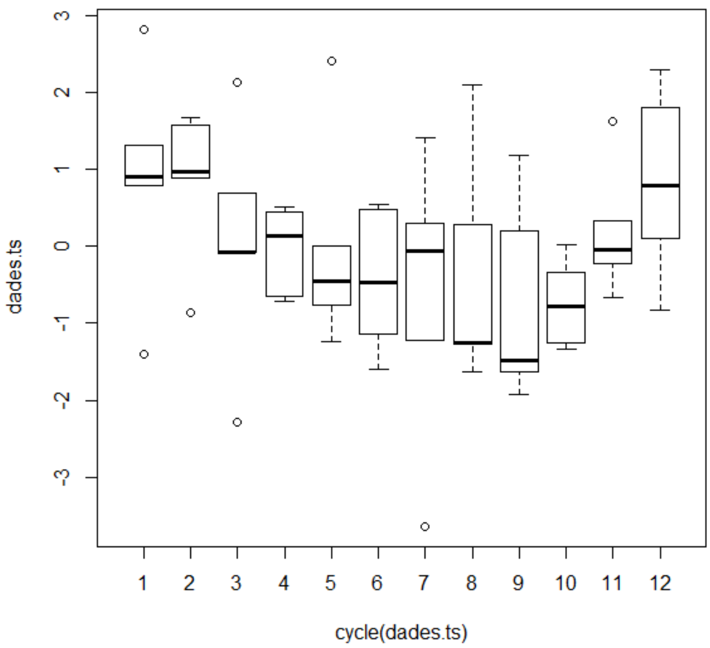
$$X_t = 0{,}51579 \cdot X_{t-1} + E_t$$

donde el coeficiente que acompaña a la componente X_{t-1} es significativamente distinto a 0. Por lo que podemos considerar un AR(1) como un modelo adecuado para describir el proceso.

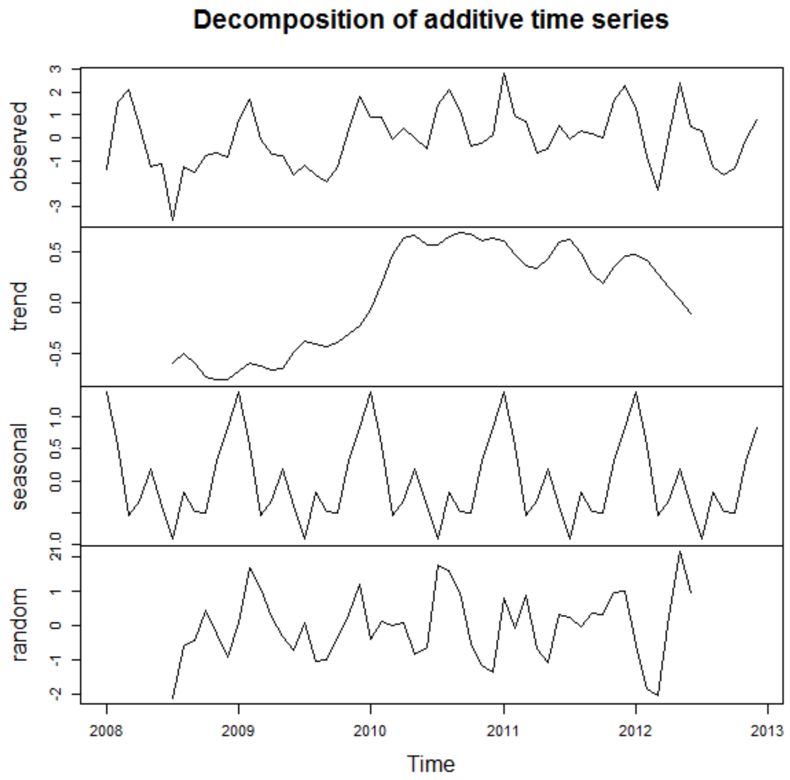
En todo el estudio no nos hemos preguntado si tiene estacionalidad.

¿Tiene estacionalidad?

par(mfrow = c(1,1))
boxplot(dades.ts ~ cycle(dades.ts))



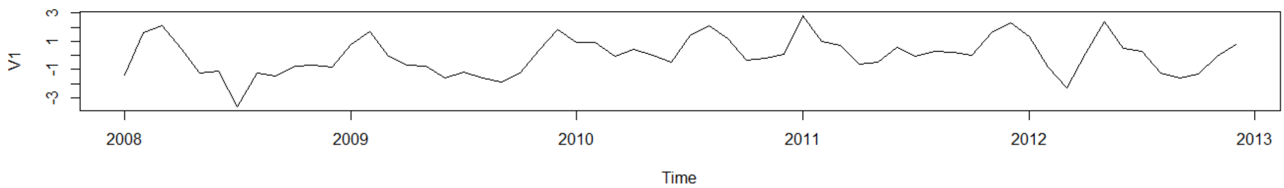
decompose = decompose(dades.ts)
plot(decompose(dades.ts))

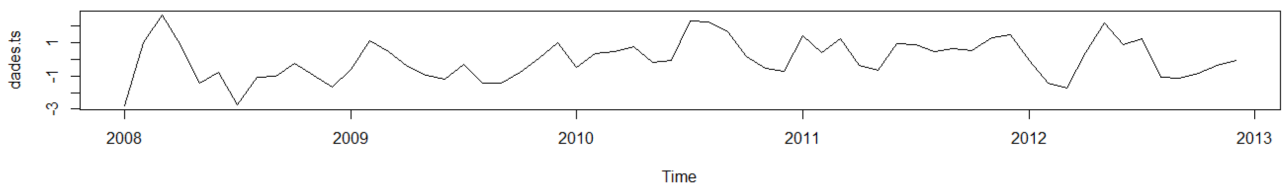



```
> decompose = decompose(dades.ts)
> plot(decompose(dades.ts))
> decompose$seasonal
            Jan        Feb        Mar        Apr        May        Jun
2008  1.3918323  0.5588248 -0.5460843 -0.2961595  0.1979111 -0.3826392
2009  1.3918323  0.5588248 -0.5460843 -0.2961595  0.1979111 -0.3826392
2010  1.3918323  0.5588248 -0.5460843 -0.2961595  0.1979111 -0.3826392
2011  1.3918323  0.5588248 -0.5460843 -0.2961595  0.1979111 -0.3826392
2012  1.3918323  0.5588248 -0.5460843 -0.2961595  0.1979111 -0.3826392
            Jul        Aug        Sep        Oct        Nov        Dec
2008 -0.9142927 -0.1712322 -0.4786339 -0.5087150  0.3132622  0.8359263
2009 -0.9142927 -0.1712322 -0.4786339 -0.5087150  0.3132622  0.8359263
2010 -0.9142927 -0.1712322 -0.4786339 -0.5087150  0.3132622  0.8359263
2011 -0.9142927 -0.1712322 -0.4786339 -0.5087150  0.3132622  0.8359263
2012 -0.9142927 -0.1712322 -0.4786339 -0.5087150  0.3132622  0.8359263
```

dades_sense = dades.ts - decompose$seasonal
dades_sense

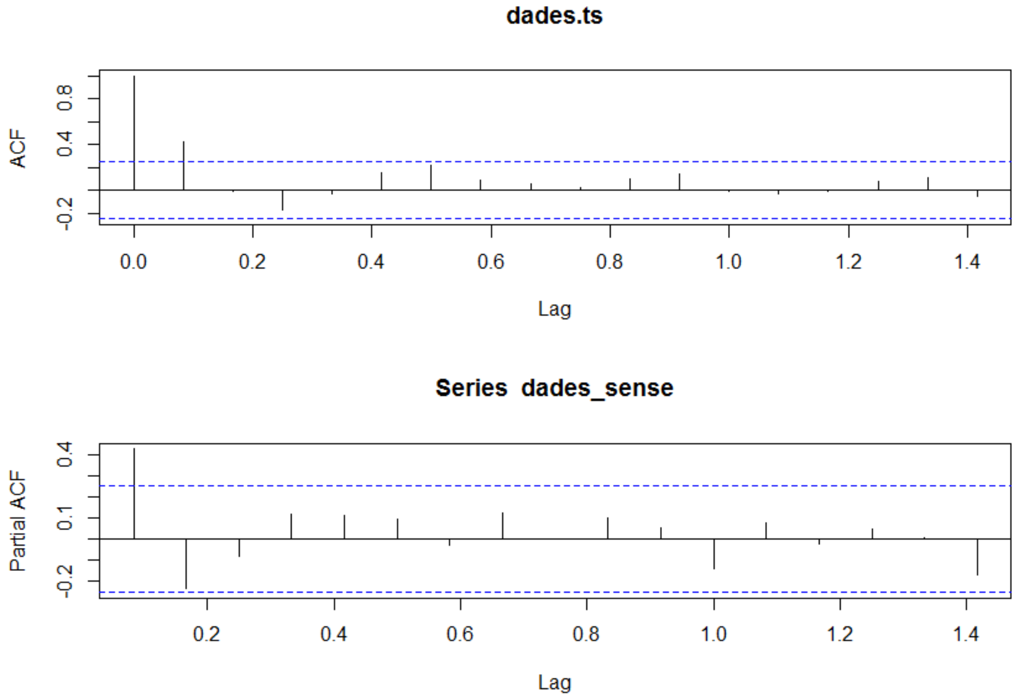
par(mfrow = c(2,1))
plot(dades.ts)
plot(dades_sense)





En el gráfico observamos que al extraer la estacionalidad hemos suavizado un poco (muy poco) la serie.

par(mfrow = c(2,1))
acf(dades_sense)
pacf(dades_sense)



fitarima = auto.arima(dades_sense)
fitarima

```
Series: dades_sense
ARIMA(0,0,2) with zero mean

Coefficients:
         ma1     ma2
      0.6543  0.2896
s.e.  0.1279  0.1693

sigma^2 estimated as 1.069:  log likelihood=-86.35
AIC=178.7   AICc=179.13   BIC=184.98
> 
```

```
> fitarima = auto.arima(dades_sense)
> fitarima
Series: dades_sense
ARIMA(0,0,2) with zero mean

Coefficients:
         ma1     ma2
      0.6543  0.2896
s.e.  0.1279  0.1693

sigma^2 estimated as 1.069:  log likelihood=-86.35
AIC=178.7   AICc=179.13   BIC=184.98
> coeftest(fitarima)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)    
ma1  0.65429    0.12786  5.1171 3.102e-07 ***
ma2  0.28961    0.16932  1.7104   0.08719 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

arima2<-arima(dades.ts,c(0,0,1), include.mean = FALSE)
arima2
coeftest(arima2)

```
> arima2<-arima(dades.ts,c(0,0,1),  include.mean = FALSE)
> arima2

Call:
arima(x = dades.ts, order = c(0, 0, 1), include.mean = FALSE)

Coefficients:
         ma1
      0.5294
s.e.  0.0929

sigma^2 estimated as 1.191:  log likelihood = -90.54,  aic = 185.08
> coeftest(arima2)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ma1 0.529371   0.092858  5.7009 1.192e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Un posible proceso seria:

Xt = 0,5294 Et-1 + Et + Componente Estacional del mes

Práctica 2.2

Estimad los parámetros del modelo que propone la función auto.arima() realizada directamente sobre el fichero de datos “prac4TS.txt” de acuerdo a las ecuaciones de Yule-Walker.

Solución orientativa

- Para el caso $p = 2$ resulta

$$\gamma(0) = \gamma(1)\phi_1 + \gamma(2)\phi_2 + \sigma^2$$
$$\gamma(0)\phi_1 + \gamma(1)\phi_2 = \gamma(1)$$
$$\gamma(1)\phi_1 + \gamma(0)\phi_2 = \gamma(2)$$

sustituyendo $\gamma(k)$ por $\hat{\gamma}(k)$, donde

Código R

```
p0 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[2]
p1 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[3]
p2 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[4]
a <- rbind(c(1,p0),c(p0,1))
a
b <- c(p0,p1)
b
solve(a,b)
```

```
> p0 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[2]
> p1 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[3]
> p2 <- as.numeric(acf(dades.ts, plot=FALSE)[[1]])[4]
> a <- rbind(c(1,p0),c(p0,1))
> a
          [,1]      [,2]
[1,] 1.0000000 0.5109971
[2,] 0.5109971 1.0000000
> b <- c(p0,p1)
> b
[1] 0.51099712 0.06025986
> solve(a,b)
[1]  0.6499070 -0.2718407
```

Por lo que el proceso AR(2) tiene la siguiente ecuación:

Xt = 0,6949 · Xt-1 – 0.2718407 · Xt-2 + Et

```
#Equivalentmente
autocors <- as.numeric(acf(dades.ts, plot=F)[[1]])[2:3]
A <- matrix(c(1, autocors[1],autocors[1], 1), 2, 2)
A
B <- matrix(c(autocors[1], autocors[2]), 2, 1)
B
phis <- solve(A, B)
phis
```

```
> autocors <- as.numeric(acf(dades.ts, plot=F)[[1]])[2:3]
> A <- matrix(c(1, autocors[1],autocors[1], 1), 2, 2)
> A
          [,1]      [,2]
[1,] 1.0000000 0.5109971
[2,] 0.5109971 1.0000000
> B <- matrix(c(autocors[1], autocors[2]), 2, 1)
> B
           [,1]
[1,] 0.51099712
[2,] 0.06025986
> phis <- solve(A, B)
> phis
           [,1]
[1,]  0.6499070
[2,] -0.2718407
```

```
#Equivalentmente
ar.yw(dades.ts)
```

```
Call:
ar.yw.default(x = dades.ts)

Coefficients:
      1        2
 0.6499  -0.2718

Order selected 2  sigma^2 estimated as  1.195
```

Bibliografía:

https://rstudio-pubstatic.s3.amazonaws.com/442799_ad56f31da23842a98c62ba4688a3d36d.html