Tarea 2

Nelson Muriel

10 de septiembre de 2016

Modelos ARIMA

Ejercicio 1: Simulación de ARIMA(p,d,q)

Consulta la documentación de `arima.sim` para este ejercicio.

1. Genera y grafica 100 observaciones del proceso AR(1) con $\phi = 0.9$. Haz lo mismo con $\phi = -0.9$.
2. Haz lo mismo con $\phi = 1$ y mira lo que pasa. Utiliza la función `rnorm` (`?rnorm`) para escribir un programa que simule un proceso AR(1) con cualquier parámetro. Comienza con

```
ar1.sim<- function(n.sim, phi){
 x <- rep(0, n.sim)
 ...
 ...
}
```

Los argumentos de la función `ar1.sim` son `n.sim` que se refiere al número de variables que quieres generar y `phi` que es el parámetro del proceso que simularás. La primera línea del código inicializa una variable `x` con `n.sim` ceros. Completa el código…

a. Mantén la primera coordenada de `x` como cero.

b. Utiliza la recursión que define el AR(1) para asignar los valores de `x[2]` hasta `x[n.sim]`

3. Ahora encuentra una manera de generar este mismo proceso con la función `arima.sim`

4. Simularemos ahora 500 observaciones del proceso AR(2) con parámetros $\phi_1 = 0.3, \phi_2 = 0.4$. Inmediatamente antes del comando `arima.sim()` utiliza `set.seed(12)`. Esto hace que la *semilla* del generador de números aleatorios sea 12 y puedas recuperar exactamente esta simulación (o alguien más). Analiza la serie que obtienes y argumenta a favor de dos modelos posibles para los datos. Hay algún elemento de diagnóstico que te lleva a decidirte por el AR(2)?

Ejercicio 2: Ajustando modelos

En este ejercicio veremos algunos detalles del ajuste de modelos ARIMA(p, d, q) en R. Genera un proceso `x` que sea una simulación de tamaño 100 del AR(1) con parámetro 0.9 y ruidos normales. Define `y <- x+50`.

1. Cuál es la media teórica para el proceso $\{Y_t\}$?

2. En la forma $Y_t = \mu + 0.9Y_{t-1} + \varepsilon_t$, cuál es el valor que debería tomar *mu*?

3. Ahora utiliza la función `arima` para ajustar el modelo AR(1) a los datos `y`. Cómo deberías interpretar el coeficiente `intercept`

4. Según los resultados arrojados por este código, escribe el modelo en la firma del inciso 2.

5. Ajusta el modelo con la función `sarima`. Busca los coeficientes y compara sus nombres. Cuál función te parece más clara?

Ejercicio 3

Visite la página de "FRED" (Federal Reserve Economic Data) y consiga los datos del PIB de Estados Unidos en el período Enero de 1947 hasta la fecha en formato *.csv*. Utilice la función `read.csv` para importarlos a R, teniendo cuidado con el formato del documento obtenido. Consulte `?read.csv` y utilice la opción `skip = n` para saltarse n renglones durante la importación.

- Examine la serie y sus diferencias gráficamente. Utilice la función `acf` para ver la función de autocorrelación o de autocorrelación parcial. Consulte `?acf`.
- Ajuste dos modelos ARIMA(p, d, q) a estos datos, utilizando la función `arima` y justificando sus elecciones.
- Ahora ajuste estos modelos con la función `sarima` (de la paquetería `astsa`). Compare los resultados de las dos funciones. Hay diferencias en la estimación? Las funciones dan distintas conclusiones? De ser así, cuál de estas funciones deberíamos usar y por qué?
- Utilice la respuesta que recibió con la función `sarima` para diagnosticar sus modelos. Con base en esto, puede elegir uno?
- Compare los criterios de información para los modelos elegidos. Con base en esto puede elegir uno?

Ejercicio 4

La serie `cmort` está disponible en su espacio de trabajo. Ajuste a esta serie un modelo ARIMA(p, d, q) basado en las funciones de autocorrelación y de autocorrelación parcial. Comente sobre su ajuste (si no lo hizo así, utilice `sarima`).

Haga lo mismo con la serie `jj`. Busque en los residuales alguna evidencia de efecto GARCH. Explique sus hallazgos.

Ejercicio 5: Una probadita...

Si no tienes la paquetería `forecast` o tienes problemas para instalarla, utiliza estas líneas para crear la función `tsdisplay`

```
tsdisplay <- function(x, lag.max = NULL){
  title <- deparse(substitute(x))
  x <- as.ts(x)
  M <- matrix(c(1,2,1,3),2)
  layout(M)
  plot(x, main = title)
  acf(x, lag.max = lag.max)
  pacf(x, lag.max = lag.max)
  layout(1)
}
```

Nota Esta función toma dos argumentos, `x` y `lag.max`. Como en la definición dice `lag.max = NULL` el argumento es opcional y puedes escribir simplemente `tsdisplay(x)`. El argumento `lag.max` sirve para indicar cuántos rezagos quieres ver en el correlograma.

En tu espacio de trabajo está disponible la serie `prodn` que contiene los datos del índice de prodution mensual desde 1948 y hasta 1978 (un total de 372 meses). Los datos no están ajustados por estacionalidad.

1. Grafica la serie con `plot`. Ves los comportamientos estacionales?

2. Despliega los logaritmos de la serie con `tsdisplay`. Observa que la raíz unitaria domina todo el correlograma y no permite leer los efectos estacionales.

3. Diferencia el logaritmo de la serie y despliégala con `tsdisplay`. En este despliegue, los comportamientos estacionales son evidentes, los puedes ver? Repite el despliegue pero con la opción `lag.max = 36` de forma que puedas ver 3 años completos.

4. Concéntrate en el patrón que forma la autocorrelación y la parcial pero sólo en los múltiplos del período estacional. De ser necesario, incrementa `lag.max` para que los puedas apreciar. Esto debería indicarte si debes o no diferenciar la serie (de primeras diferencias) de forma estacional. De ser necesario hacerlo, puedes aplicar la función `diff` con argumento `lag = 12` que es la frecuencia de la serie. De no ser necesario, elige los parámetros del modelo ARMA pero sólo sobre los patrones estacionales. Guíate con el correlograma: Por ejemplo, si la función de autocorrelación parcial tiene valores significativos para los rezagos $h = 0, 12, 24, 36$ pensarías en un modelo AR(3).

5. Si tuviste que diferenciar estacionalmente las primeras diferencias simples, despliega esta nueva serie. Vuelve a concentrarte sólo en los patrones estacionales y piensa cuál modelo ARMA daría esta conducta.

6. Por último, examina lo que pasa en los períodos inter-ciclo, es decir en los rezagos $h = 1, \ldots, 11$. Piensa cuál ARMA te daría esta forma del correlograma.

7. Ponlo todo en orden: El modelo que acabas de idear es un SARIMA(p, d, q, P, D, Q). El significado de "S" es "Seasonal" y se refiere a la presencia de estacionalidades. En nuestro caso, hemos diferenciado de forma simple una vez, así que d=1. Si diferenciaste la serie (diferenciada simplemente) de forma estacional, entonces D=1. Los parámetros del inciso 5 son P y Q. Los del inciso 6 son p y q. Utiliza la función `sarima` para ajusar tu modelo y mira los diagnósticos. Intenta que queden lo más decentes posibles.

8. Intenta hacer lo mismo por tu cuenta con la serie `unemp` que también está disponble en tu espacio de trabajo. Utiliza este modelo y la función `sarima.for` para hacer la predicción del desempleo en un horizonte de 12 meses.

Referencia Bibliográfica: El ejemplo guiado viene del libro *Time Series Analysis and Its Applications: With R Examples*, sección 3.9