MARTES 18 JUNIO 2019 — Series de Tiempo

BJ. Aplicación 2

autoarima $(2,1,0)(2,0,0)[4]$ — modelo 1

(1) IDENTIFICACIÓN

KPSS. e $d=0$ — FAC: ext. / FAP: trunc. } $p=1$, $q=0$ | autoarima $p=2$, $q=0$

DF / PP } NE → $d>0$ $\xrightarrow{dif}$ modelo 3 $\boxed{d=1}$ → modelo 4 $\boxed{d=2}$ → modelo 6 $\boxed{d=3}$ ← modelo 2

FAC / FAP > Ruido blanco $p=0$, $q=0$

modelo 4 — FAC / FAP ruido blanco } $p=0$, $q=0$ | autoarima $p=0$, $q=1$ — modelo 5

...

con 3 diferencias se hace estacionario (excepcional)

ggtsdisplay (Y, lag=12) — FAC parece extingue / FAP trunca en desfase 1 } $p=1$, $q=0$ | autoarima $p=2$, $q=0$

autoarima (Y, d=1)

autoarima (Y, d=2) ruido blanco | autoarima $p=0$, $q=1$

(Y, d=3) parece ext. - trunc. desf. 2 | $p=2$, $q=0$

Cuando hay 2 o más diferencias la constante es ignorada.

(2) Modelamiento

$$Y_t = \delta + \phi Y_{t-1} + \varepsilon_t \qquad AR(1) \text{ con intercepto}$$

Modelo 1 (sin intercepto). No estacionario

Modelo 1 (con ")

$$Y_t = 0.5755 + 0.8135\, Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, 0.04057)$$

$$\mu = \delta / 1-\phi_1 = \delta = 5.0858\,(1-0.8135)$$
$$= 0.5755$$

Forma general (que permite llegar a ...)

$$\phi(B)\,(1-B)^d\,(Y_t - \delta) = \Theta(B)\,\varepsilon_t$$

p. autoregresivo — polinomio de medias móviles

B operador de sus índices

$$(1-0.8135B)\,(Y_t - 0.5755) = \varepsilon_t$$

$$Y_t - 0.5755 - 0.8135\, Y_t = \varepsilon_t$$

reordenando

Vemos con AR(2) con y sin constante.

modelo 2a Sin media Y depende de Y_{t-1} y Y_{t-2} pero necesita una media. AR(2) con intercepto

modelo 2 $(1-1.1316B)-(-0.4084)B^2)(1-B)^0(Y_t-\delta) = \boxed{?}\,\varepsilon_t$

$$\delta = 3.1136(1-1.1316+0.4084) = 0.8618$$

$$Y_t = 0.8618 + 1.1316Y_{t-1} - 0.4084Y_{t-2} + \varepsilon$$

$\varepsilon_t \sim N(0, 0.0354)$

modelo 3 Sin constante $I(1)$ porqué se le llama integrado? es: modelo simple random walk

$$(1-B)^{d=1}Y_t = \varepsilon_t$$

$$Y_t - Y_{t-1} = \varepsilon_t$$

$$Y_t = Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, 0.0403)$$

Con ~~constante~~ drift

$$(1-B)(Y_t - \beta t) = \varepsilon_t \qquad \beta = 0.0355$$

$$Y_t - \beta t - Y_{t-1} = \varepsilon_t \qquad \varepsilon_t \sim N(0, 0.0\ \ )$$

$$Y_t = Y_{t-1} + 0.0355t + \varepsilon_t$$

modelo 4 es un $I(2)$ integrado de orden 2

$$(1-B)^2Y_t = \varepsilon_t$$

$$(1-2B+B^2)Y_t = \varepsilon_t$$

$$Y_t = 2Y_{t-1} - Y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0, 0.0595)$$

modelo 5 IMA(2,1) parecido a modelo 4: izquierda igual a modelo 4 derecha diferente.

$$(1-B)^2Y_t = \theta(B)\varepsilon_t$$

$$(1-2B+B^2)Y_t = (1-0.6427B)\varepsilon_t$$

$$Y_t = \underbrace{2Y_{t-1} - Y_{t-2} + \varepsilon_t}_{\text{igual}} - 0.6427\varepsilon_{t-1} \qquad \varepsilon_t \sim N(0, 0.4511)$$

modelo 6 ... todos los modelos han sido simples

arima (2,3,0)

$$\underbrace{(1-(-1.0219)B-(-0.7148)B^2)}_{\text{p. autor.}}(1-B)^3 Y_t = \varepsilon_t$$

DIAGNOSTICO

normalidad de residuales.   todos!

t.test. media cero para que predicciones sean insesgadas   todos!

ggtsdisplay(   )   ver que este dentro de bandas (No haya patrón significativo) sobre todo de FAC

"Sin autocorrelaciones significativas, dentro de cada

AR(2) con media parece ser el mejor de acuerdo a 2 indicadores de accuracy. Parece tener relación con los residuales menores.

Validación cruzada con 23 datos (años)

se puede usar k=18 (10% de 23)

ho = 2 años

RMSE para cada modelo:

((ts(predi   ) - ts(test))^2)

Gráficos: modelo 2.

modelo 4, 5 y 6: son menos simples

modelos ARIMA tienden a ser parsimoniosos: por lo tanto más simples son mejores.

Pronóstico

forecast(modelo2, h=2)   volver a correr modelo 2!

D.

Se trabaja con p, q, d, f y (P, D, Q)$_s$ subíndice de estacionalidad.

cuando e. aditiva: podríamos decir estacionalidad estacionaria patrón constante (No hay que hacer ____)

e. multiplicativa: hay que estacionalizar por medio de diferenciaciones con desfases que correspondan a estación ?

Diferenciación Estacional (D)

Aplicación 3.

library(fpp2) euretail: índice trimestral con base 2005

gráfico: predominio de tendencia FAC

pruebas de estacionariedad: hay valores divididos. NO estacionariedad en PH con diferenciación vemos en silueta que por estacionalidad

La estacionalidad se revela cada 4 desfases

diff(Y) una dif. simple

" (Y, lag=12) una dif de orden 12

" (Y, differences=2) dos difer. simples

" (Y, lag=2) una diferenc. de orden 2

" (Y, lag=2, differences=2) dos diferenciaciones de orden

" (Y, lag=4, differences=2) dos diferenciaciones de orden 4

como hay tendencia ⇒ 1 diferenciación

debemos ver FAC, PAP } autoarima

Si $d=1 \Rightarrow D=0$

ext - tru
tru - ext
ext - ext
dentro de bandas
otro ?

$\boxed{D=0}$ FAC, FAP solo desfases múltiplos de 4 { ext. trunc = 8 pero s=4 ⇒ lag estacional = 2 ⇒ $P=2, Q=0$

autoarima:
arima(2,1,0)(2,0,0)[4]
P Q

ARIMA $\phi(B)(1-B)^d(Y_t-\delta) = \theta(B)\varepsilon_t$

SARIMA $\underbrace{\phi(B)}_{\text{P. Estacional}}\underbrace{\phi B}_{\text{P. Autor.}}(1-B)^d(1-B^s)^D(Y_t-\delta) = \underbrace{\Theta(B)}_{\text{P.MM Estacional}}\underbrace{\theta(B)}_{\text{P.MM.}}\varepsilon_t$

5

coeficientes sar1 y sar2   ar1, ar2

$d=1$, $D=0$ $p=2$, $P=2$, $q=0$, $Q=0$

SARIMA (2,1,0)(2,0,0)4

$(1-0.3338B^4-0.4101B^8)(1-0.3021B-0.1823B^2)(1-B)Y_t = \varepsilon_t$ $\varepsilon_t \sim N(0, 0.1761)$ sigma2^2

operando habría 18 términos: 3 x 3 x 2 términos

Con $\boxed{d=1}$ y $\boxed{D=1}$ < FAC truncada orden 1, desfase 4: MA(1) $P=0$, $\Theta=1$ | autoArima $P=0$, $Q=1$
FAP extinción

FAC FAP
autoarima $p=0$, $q=3$

No Autor. Simple Ni estacional ⇒

$(1-B)^1(1-B^4)^1(Y_t-0) = (1-0.6636B^4)(1+0.263\theta+0.369\theta^2+0.4200^3)\varepsilon_t$

sma1 $\varepsilon_t \sim N(0, 0.156)$ sigma2^2

aquí depende de varios ruidos blancos previos. No, de tantos Y_t

PC2 add 3 puntos. Impuesto Tributario.