

SARIMA modely – poznámky k softvéru R

Beáta Stehlíková, Časové rady FMFI UK, 2011

Znovu budeme potrebovať súbor **tsa3.rda** zo stránky www.stat.pitt.edu/stoffer/tsa3. Funkcie **sarima** a **sarima.for** sme už používali pri odhadovaní ARIMA modelov. Teraz si ukážeme ich použitie na všeobecnejšie SARIMA modely. Budeme pracovať s dátami **spanielsko.txt** o počte turistov v Španielsku (mesačné dáta od januára 1970 do marca 1989), pre ktoré sme hľadali model na konci prednášky.

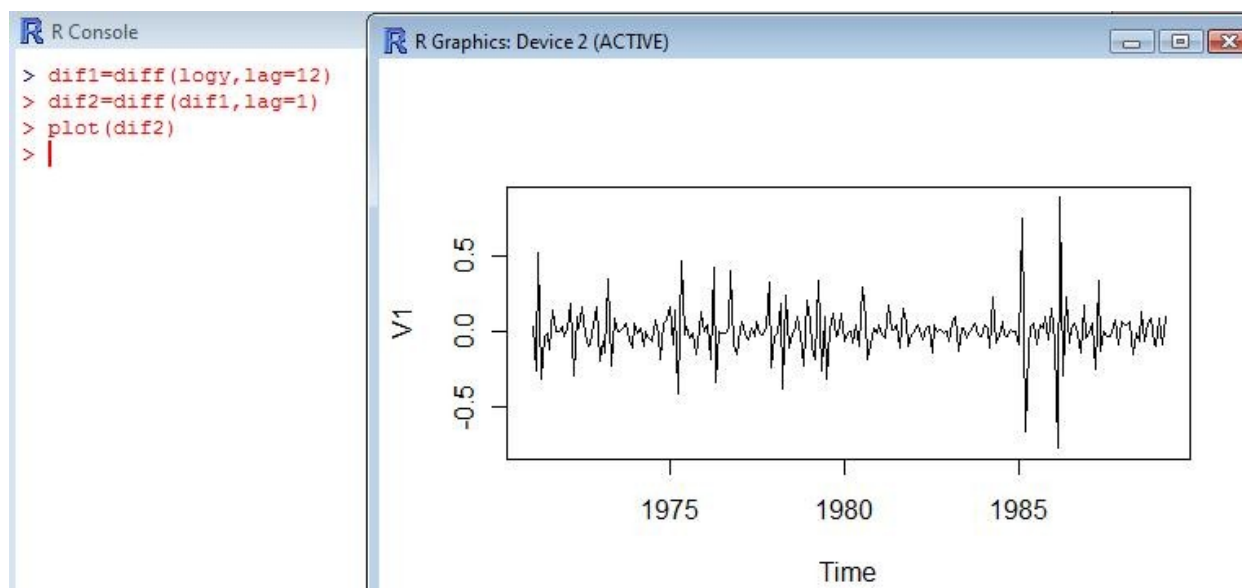
Dáta

Načítame skripty z **tsa3.rda** a dáta. Do premennej **y** uložíme počty turistov a do premennej **logy** ich logaritmy.

```
R Console
> setwd('e://skola/r')
> load('tsa3.rda')
> y=read.table('spanielsko.txt')
> y=ts(y,frequency=12,start=c(1970,1))
> logy=log(y)
```

Diferencie

Diferencie vytvárame príkazom **diff**, má tvar **diff(premenna, lag=...)**. Ak si teda chceme zobrazit' priebeh modelovanej premennej, ktorou v našom prípade boli zdiferencované logaritmy – sezónne ročné diferencie a klasické prvé diferencie – zadáme:



Pre tieto diferencie si môžeme zobrazit' ACF a PACF príkazom **acf2(dif2)** alebo pomocou pôvodných funkcií R **acf** a **pacf**.

SARIMA model

Odhadujeme ho príkazom **sarima**, ktorý sme doteraz používali v tvare **sarima(premenna,p,d,q)** na odhadovanie ARIMA (p,d,q) modelu. Ak chceme odhadnúť model SARIMA (p,d,q)x(P,D,Q)_s, spravíme to príkazom **sarima(premenna,p,d,q,P,D,Q,s)**. Ak máme aj klasické aj sezónne diferencie, takto odhadnutý model neobsahuje konštantu.

Na prednáške sme sa dopracovali k modelu, ktorý pre logaritmus počtu turistov používa model SARIMA (0,1,1)x(0,1,1)₁₂:

```

> sarima(logy,0,1,1,0,1,1,12,details=FALSE)
$fit

Call:
arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
      include.mean = !no.constant, optim.control = list(trace = trc, REPORT = 1,
        reltol = tol))

Coefficients:
      ma1      sma1
-0.7354  -0.7267
s.e.   0.0455   0.0515

sigma^2 estimated as 0.01118:  log likelihood = 175.52,  aic = -345.04

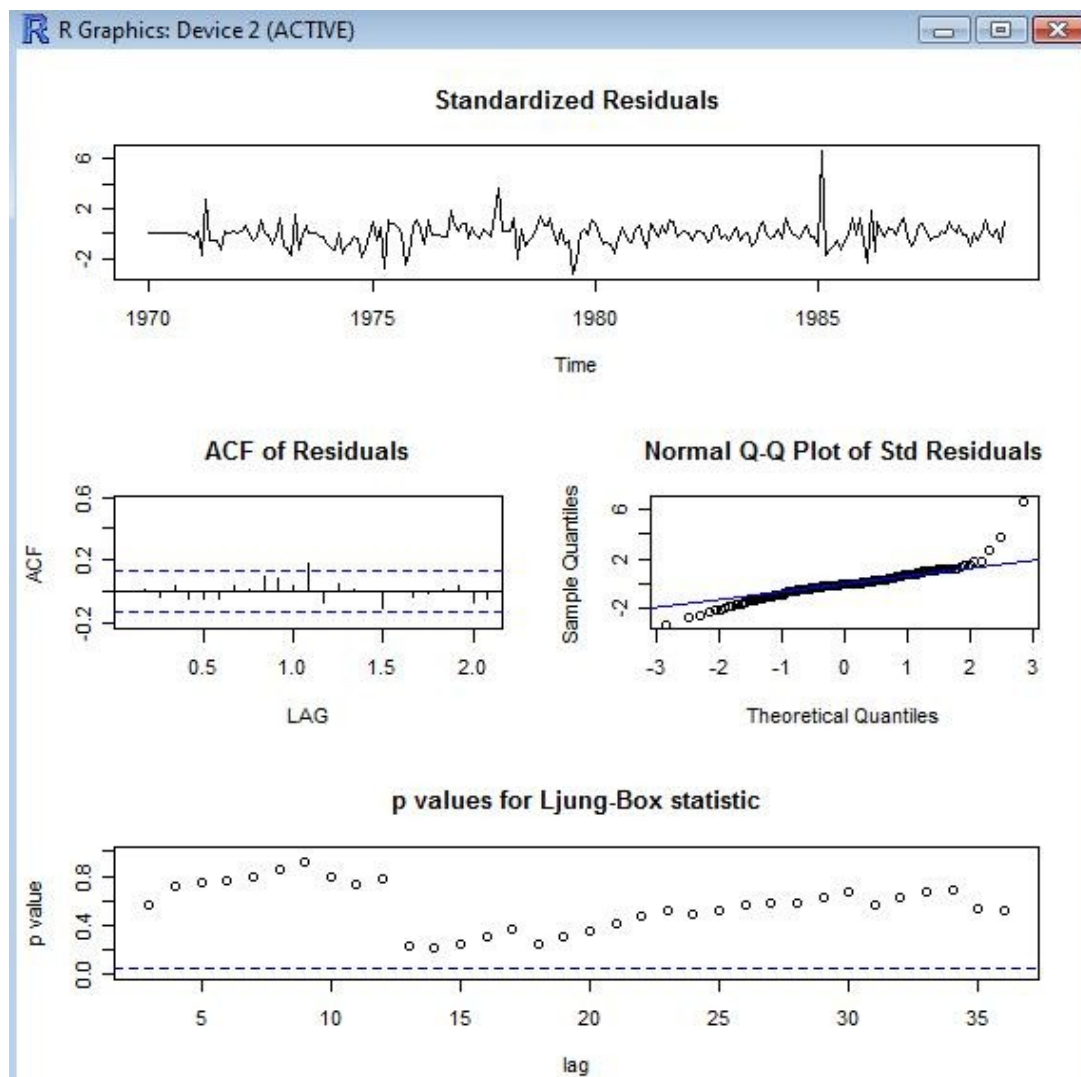
$AIC
[1] -3.475887

$AICc
[1] -3.466771

$BIC
[1] -4.446082

```

Znovu dostávame okrem horeuvedeného aj grafický výstup:



Predikcie v SARIMA modeloch

Predikcie zostrojime príkazom **sarima.for**, ktorý sme doteraz používali v tvare pre ARIMA modely. Jeho použitie pre SARIMA modely je analogické ako v prípade príkladu **sarima**.

Zopakujeme výpočet z prednášky – odhadneme model z dát do konca roku 1986, pre zvyšné dáta (roky 1987, 1988 a prvé tri mesiace roku 1989 – čiže spolu 27 pozorovaní) spravíme predikcie a porovnáme ich so skutočným vývojom.

```
> logy2=ts(logy,frequency=12,start=c(1970,1),end=c(1986,12))
> sarima.for(logy2,27,0,1,1,0,1,1,12)
$pred
      Jan      Feb      Mar      Apr      May      Jun      Jul      Aug      Sep
1987 14.61796 14.64774 14.77975 15.00071 15.12291 15.35394 15.94320 16.10161 15.54438
1988 14.67033 14.70010 14.83211 15.05308 15.17528 15.40631 15.99556 16.15398 15.59674
1989 14.72270 14.75247 14.88448
      Oct      Nov      Dec
1987 15.10393 14.71868 14.96013
1988 15.15630 14.77105 15.01250
1989
$se
      Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
1987 0.1107684 0.1148056 0.1187056 0.1224814 0.1261443 0.1297038 0.1331681 0.1365447
1988 0.1611891 0.1657268 0.1701436 0.1744486 0.1786499 0.1827546 0.1867692 0.1906992
1989 0.2169747 0.2219760 0.2268670
      Sep      Oct      Nov      Dec
1987 0.1398397 0.1430588 0.1462071 0.1492889
1988 0.1945499 0.1983258 0.2020312 0.2056698
1989
> lines(logy)
```

