

Porovnanie predikcií s reálnymi dátami

Beáta Stehlíková
FMFI UK Bratislava

Porovnanie predikcií s reálnymi dátami

Porovnanie predikcií s reálnymi dátami

Základná myšlienka:

- ▶ Vynecháme z našich dát niekoľko posledných pozorovaní
- ▶ Model vytvoríme a otestujeme pomocou zostávajúcich dát
- ▶ Spravíme predikcie
- ▶ Porovnáme s reálnymi dátami, ktoré sme si na začiatku odložili

Príklad - dáta

Ešte raz sa vrátíme k týždenným dátam cien benzínu:

```
library(astsa)
data(gas)
gas <- window(gas, start=c(2006,1))
```

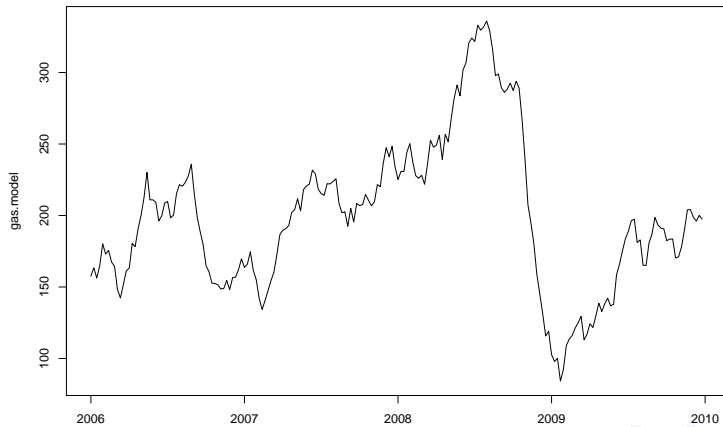
Zobrazte priebeh týchto dát.

Na modelovnie použijeme len dáta do konca roku 2009

```
gas.model <- window(gas, end=c(2009,52))
```

Príklad - dáta

```
plot(gas.model)
```



Príklad - testovanie jednotkového koreňa a ARMA model

Minule sme povedali, že budeme modelovať diferencie tejto premennej. Prečo?

Spravte ADF test, vysvetlite, čo sa testuje a s akými výsledkami. Nájdite dáta, pre ktoré budeme hľadať ARMA model.

Na zopakovanie:

- ▶ zobrazte ACF a PACF týchto dát
- ▶ odhadnite AR(4) a MA(3) model pre diferencie

Príklad - predikcie

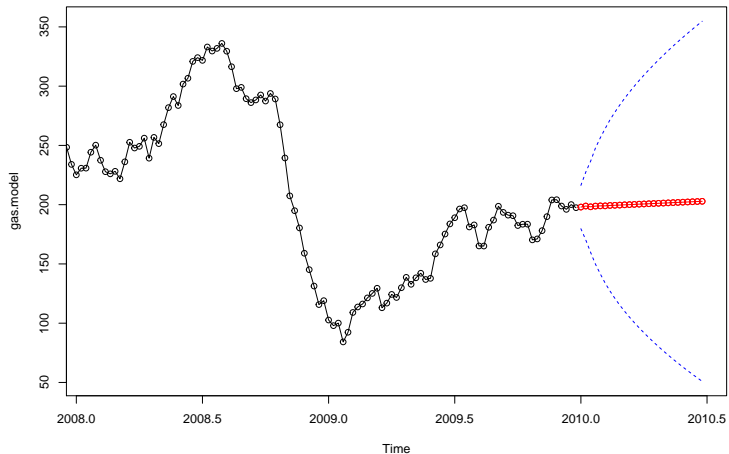
Spravíme predikcie pre ARIMA(4,1,0) model na najbližší polrok, ale teraz ich uložíme do premennej:

```
predikcie<-sarima.for(gas.model,26,4,1,0)
```

Teraz vieme pristupovať k hodnotám predikcií a ich štandardným odchýlkam pomocou `predikcie$pred` a `predikcie$se`.

Hranice intervalov spoľahlivosti potom sú `predikcie$pred - 2*predikcie$se` a `predikcie$pred + 2*predikcie$se`

Príklad - predikcie



Príklad - predikcie

Napríklad:

```
predikcie$pred
```

```
## Time Series:  
## Start = c(2010, 1)  
## End = c(2010, 26)  
## Frequency = 52  
## [1] 197.9833 198.8581 198.2043 198.7984 198.9949 199.0000  
## [8] 199.4834 199.6536 199.8796 200.0392 200.2316 200.4200  
## [15] 200.7922 200.9779 201.1624 201.3500 201.5355 201.7200  
## [22] 202.0938 202.2800 202.4661 202.6522 202.8383
```

Príklad - predikcie a reálne dáta

Začneme vykreslením dát (všetkých):

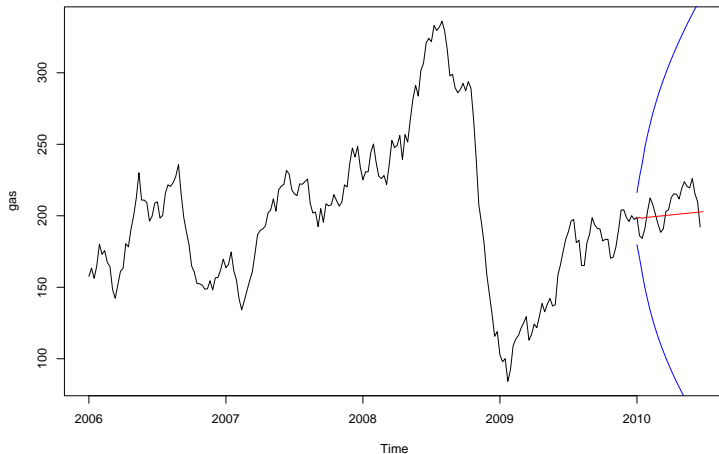
```
plot(gas)
```

Pridáme predikcie (červenou):

```
lines(predikcie$pred, col=c("red"))
```

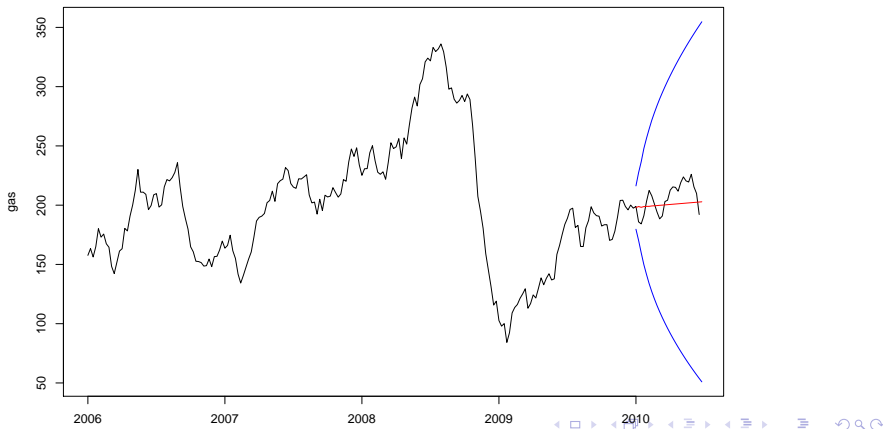
Pridajte intervaly spoľahlivosti pre predikcie (modrou).

Príklad - predikcie a reálne dáta



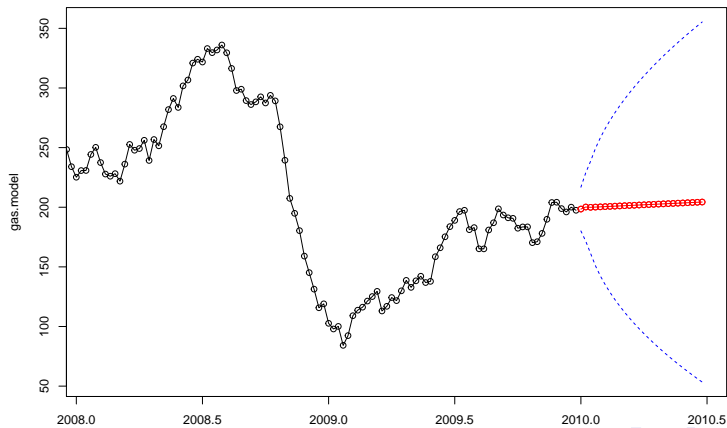
Príklad - predikcie a reálne dáta

Upravte rozsah y-ovej osi grafu tak, aby obsahoval všetky hodnoty (najlepšie automaticky, nie zadaním pevných hraníc).



Príklad - porovnanie predikcií modelov

```
predikcie2<-sarima.for(gas.model,26,0,1,3) # ARIMA(0,1,3)
```



Príklad - porovnanie predikcií modelov

Porovnajzte predikcie ARIMA(4,1,0) a MA(0,1,3) modelu (môžete zmenšiť rozsh x-ovej osi, aby boli predikcie lepšie viditeľné):

