Jednotkový koreň (unit root), diferencovanie časového radu, ADF test na testovanie jednotkového koreňa

Beáta Stehlíková

Časové rady

Fakulta matematiky, fyziky a informatiky, UK v Bratislave

Obsah

- Čo je jednotkový koreň, prečo neumožňuje použiť ARMA metodológiu a ako ho odstrániť
- Ako z dát zistiť, či má proces jednotkový koreň alebo nie $\rightarrow$ testy jednotkového koreňa (unit root testy)
- Ako spraviť tento test v R-ku
- Kompletný postup hľadania ARIMA modelu pre zadané dáta

Čo je jednotkový koreň, prečo neumožňuje použiť ARMA metodológiu a ako ho odstrániť

Príklad 1

Majme proces

$$y_t = y_{t-1} + u_t$$

resp. po zápise pomocou operátora posunu

$$(1 - L)y_t = u_t$$

- je to nestacionárny AR(1) proces
- polynóm $1 - L$ má koreň $L = 1$, t. j. **jednotkový koreň**
- keď hovoríme o jednotkovom koreni, ide nám koreň $L = 1$ autoregresného polynómu - spôsobuje nestacionaritu
- pre diferencie tohto procesu $\Delta y_t = y_t - y_{t-1}$ platí $\Delta y_t = u_t$
- teda diferencie Δy_t sú stacionárne

Príklad 2

- Majme nestacionárny proces s jednotkovým koreňom

$$(1 - \frac{1}{2}L)(1 - L)y_t = 1 + (1 - \frac{1}{3}L)u_t$$

- Potom pre diferencie tohto procesu

$$\Delta y_t = y_t - y_{t-1} = (1 - L)y_t$$

  platí $(1 - \frac{1}{2}L)\Delta y_t = 1 + (1 - \frac{1}{3}L)u_t$
- Teda diferencie Δy_t sú stacionárne

Príklad 3

- Majme nestacionárny proces s dvojnásobným jednotkovým koreňom

$$(1 - \frac{1}{2}L)(1 - L)^2 y_t = 1 + (1 - \frac{1}{3}L)u_t$$

- Potom pre druhé diferencie tohto procesu

$$\Delta^2 y_t = \Delta(\Delta y_t) = (1 - L)(1 - L)y_t = (1 - L)^2 y_t$$

  platí $(1 - \frac{1}{2}L)\Delta^2 y_t = 1 + (1 - \frac{1}{3}L)u_t$
- Teda druhé diferencie Δy_t sú stacionárne

Vo všeobecnosti

- Majme proces s jednotkovým koreňom násobnosti k , pričom ostatné korene sú mimo jednotkového kruhu
- Potom k-te diferencie procesu sú stacionárne

ARIMA modely, terminológia

- Ak treba proces k krát diferencovať, aby sme z neho dostali stacionárny proces, nazýa sa integrovaný proces rádu k a označuje sa $I(k)$

Canarella, G., Gupta, R., Miller, S. M., & Pollard, S. K. (2019). **Unemployment rate hysteresis and the great recession: exploring the metropolitan evidence.** Empirical Economics, 56(1), 61-79.

Abstract Standard unit-root tests of the hysteresis hypothesis specify a unit root under the null against the stationary alternative of the natural-rate hypothesis, making the two theories of unemployment mutually exclusive over the sample period. In this paper, we allow switches between hysteresis and natural-rate theory using the Kejriwal, Perron, and Zhou test. The null hypothesis of the test is that the unemployment rate is I(1) throughout the sample, and the alternative hypothesis is that the unemployment rate changes persistence [i.e., switches between I(0) and I(1) regimes]. We apply the test to the unemployment rate of 20 metropolitan statistical areas (MSAs) and the USA. We use monthly observations over the period 1990:1–2016:12 and apply the test to seasonally unadjusted and seasonally adjusted data.

Ako testovať prítomnosť jednotkového koreňa v dátach - simulácie

Cieľ

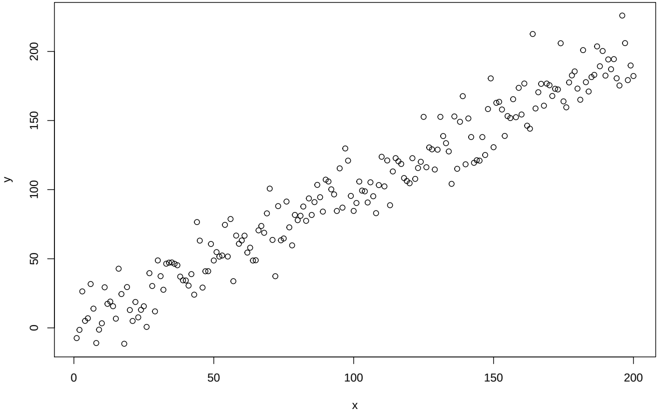
- Uvažujme najskôr AR(1) proces $x_t = \delta + \rho x_{t-1} + u_t$
- Chceme:
  - testovať hypotézu o jednotkovom koreni (vtedy je proces nestacionárny), teda $H_0 : \rho = 1$
  - zistiť, či sa dá zamietnuť v prospech stacionarity, teda $H_0 : \rho < 1$
- Skúsme použiť testovanie hypotéz o koeficientoch regresného modelu známe z ekonometrie.
  - spravíme simulácie
  - uvidíme, že tento postup bude treba upraviť

Simulácia 1: klasická regresia a t-štatistika

Štandardný postup - vieme spraviť:

- Majme vektor `x <- 1:200`
- Vygenerujeme `y <- x + rnorm(200) * sigma`
- Odhadneme model $y = c + \rho x + \varepsilon$
- Zaznamenávame:
  - odhad parametra ρ
  - hodnotu t-štatistiky zodpovedajúcej hypotéze $H_0 : \rho = 1$ (ktorá platí)
- Zopakujeme 10^5 krát a vykreslíme histogram

- Ukážka dát:



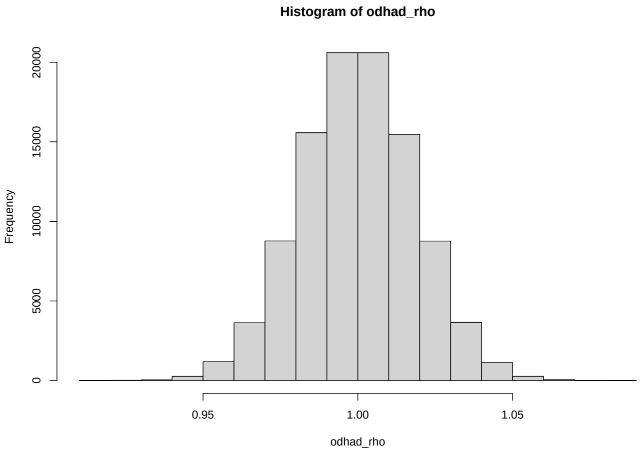
- Výstup z regresie:

```
summary(lm(y ~ x))$coefficients
```

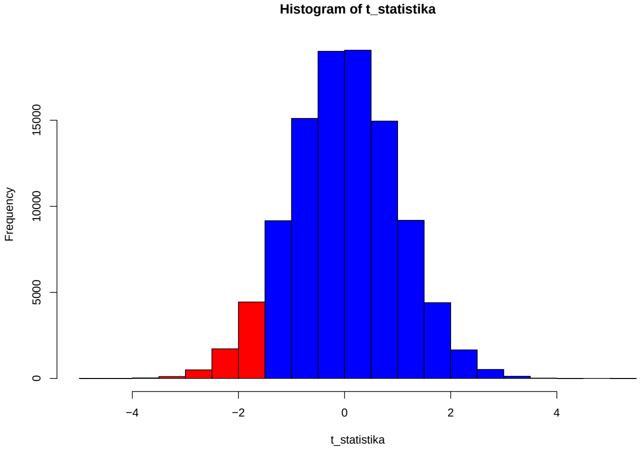
```
##              Estimate Std. Error    t value     Pr(>|t
## (Intercept) 1.0330020 2.01108356  0.5136544  6.080663e-0
## x           0.9884422 0.01735144 56.9660036 9.490913e-12
```

- Odhadnutý koeficient ρ je 0.98844.
- T-štatistika k hypotéze $\rho = 1$ sa počíta ako $\frac{0.98844-1}{0.01735}$

- Výsledok zo simulácií: odhad parametra ρ (normálne rozdelenie)



- Výsledok zo simulácií: t-štatistika (Studentovo rozdelenie, vyznačený 5% kvantil tohto rozdelenia)

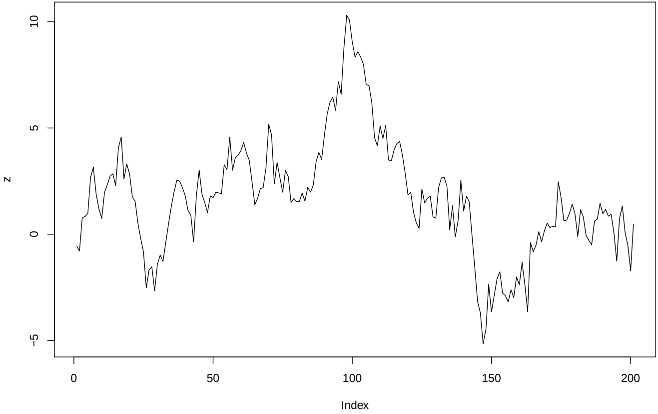


Simulácia 2: jednotkový koreň a t-štatistika

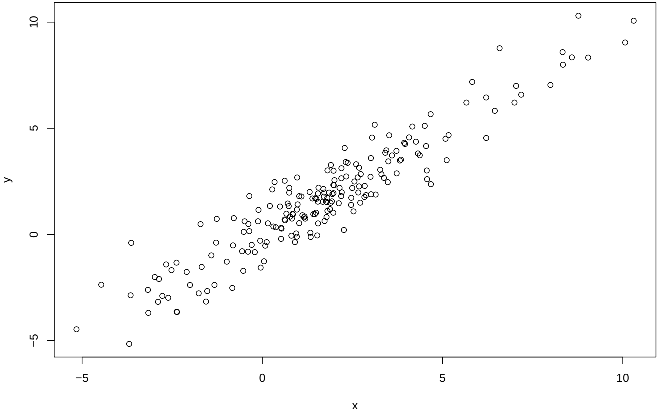
Druhá simulácia:

- Majme vektor `z` vygenerovaný ako $z_t = z_{t-1} + \varepsilon_t$
- Zoberieme `x <- z[1:200], y <- z[2:201]`, teda $x_t = z_{t-1}$, $y_t = z_t$
- Odhadneme model $y = c + \rho x + \varepsilon$, teda $z_t = c + \rho z_{t-1} + \varepsilon_t$
- Zaznamenávame:
  - odhad parametra ρ
  - hodnotu t-štatistiky zodpovedajúcej hypotéze $H_0 : \rho = 1$ (ktorá platí)
- Zopakujeme 10^5 krát a vykreslíme histogram

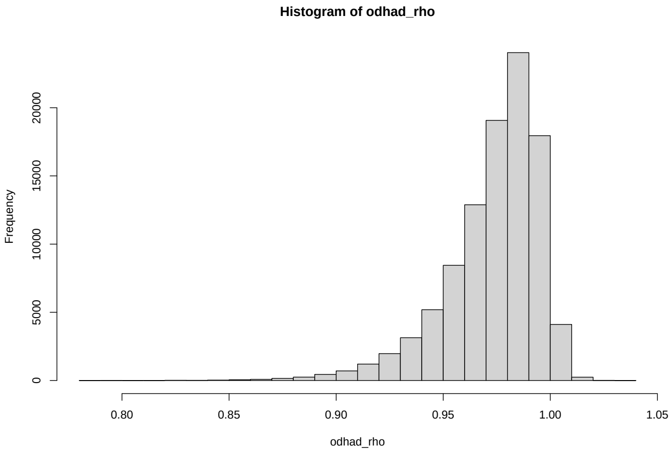
- Ukážka vygenerovaných dát: proces z_t do regresie $z_t = c + \rho z_{t-1} + \varepsilon_t$)



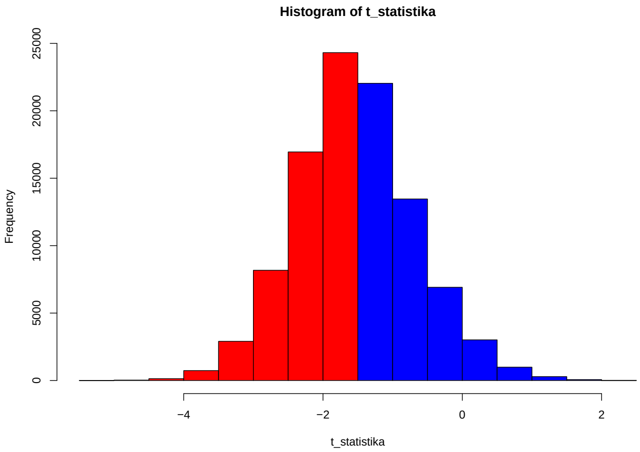
- Ukážka vygenerovaných dát: dáta do regresie



- Výsledok zo simulácií: odhad parametra ρ (nemá normálne rozdelenie)



- Výsledok zo simulácií: "t-štatistika" (nemá Studentovo rozdelenie, vyznačený 5% kvantil tohto rozdelenia)



Ako testovať prítomnosť jednotkového koreňa v dátach - ADF test

Základná myšlienka:

- Ponecháme výpočet testovacej štatistiky
- Ale budeme používať **iné kritické hodnoty**
- Približne aké by mali byť kritické hodnoty (podľa našich simulácií):

```
quantile(t_statistika, 0.05)
```

```
##        5%
## -2.888873
```

```
quantile(t_statistika, 0.01)
```

```
##        1%
## -3.471761
```

Testovanie jednotkového koreňa pre AR(1)

- AR(1) proces
$$y_t = \rho y_{t-1},$$
jednotkový koreň znamená, že $\rho = 1$
- Ekvivalentne:
$$\Delta y_t = (\rho - 1) y_{t-1} + u_t$$
a zaujíma nás **t-štatistika zo signifikancie koeficienta pri** y_{t-1}, ale **s inou kritickou hodnotou**
- Bolo zistené, že tá kritická hodnota
  - závisí od počtu dát
  - zmení sa, ak proces obsahuje konštantu a/alebo lineárny trend
- Vo všeobecnosti: $\Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + u_t$

Testovanie jednotkového koreňa pre AR(p)

- AR(p) proces

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_p \dots y_{t-p} + u_t$$

- Prepíšeme pomocou posunu

$$(1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p) y_t = u_t$$

- To, že $L = 1$ je koreňom, znamená:

$$1 - \alpha_1 1 - \alpha_2 1^2 - \dots - \alpha_p 1^p = 0$$

- Jednotkový koreň teda predstavuje podmienku

$$\alpha_1 + \dots + \alpha_p = 1$$

- Máme teda AR(p) proces
  $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + u_t$
- Upravíme ho do tvaru

$$y_t = \rho y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \ldots \theta_{p-1} \Delta y_{t-p+1} + u_t,$$

  kde $\rho = \sum_{j=1}^{p} \alpha_j, \theta_i = -\sum_{j=i+1}^{p} \alpha_j$ pre $i = 1, \ldots, p-1$
- Teda na pravej strane je:
  - predchádzajúca hodnota procesu y_{t-1} s koeficientom $\alpha_1 + \cdots + \alpha_p$
  - diferencie v starších časoch $\Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-p+1}$

- Napríklad pre AR(3):

$$
\begin{aligned}
y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + u_t \\
&= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + (-\alpha_3)(y_{t-2} - y_{t-3}) + \alpha_3 y_{t-2} + u_t \\
&= \alpha_1 y_{t-1} + (\alpha_2 + \alpha_3) y_{t-2} + (-\alpha_3)(y_{t-2} - y_{t-3}) + u_t \\
&= \alpha_1 y_{t-1} + (-\alpha_2 - \alpha_3)(y_{t-1} - y_{t-2}) + (-\alpha_3)(y_{t-2} - y_{t-3}) \\
&\quad + (\alpha_2 + \alpha_3) y_{t-1} + u_t \\
&= (\alpha_1 + \alpha_2 + \alpha_3) y_{t-1} + (-\alpha_2 - \alpha_3)(y_{t-1} - y_{t-2}) \\
&\quad + (-\alpha_3)(y_{t-2} - y_{t-3}) + u_t \\
&= (\alpha_1 + \alpha_2 + \alpha_3) y_{t-1} + (-\alpha_2 - \alpha_3)\Delta y_{t-1} + (-\alpha_3)\Delta y_{t-2} \\
&\quad + u_t
\end{aligned}
$$

- AR(p) proces v tvare

$$y_t = \rho y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \theta_{p-1} \Delta y_{t-p+1} + u_t,$$

sa dá ekvivalentne zapísať ako

$$\Delta y_t = (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \theta_{p-1} \Delta y_{t-p+1} + u_t$$

a zaujíma nás potom t-štatistika z koeficienta pri y_{t-1}.
- Vo všeobecnosti: proces môže obsahovať konštantu a/alebo lineárny trend

$$\Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \theta_{p-1} \Delta y_{t-p+1} + u_t$$

Augmented Dickey-Fuller test (ADF)

- Wayne A. Fuller (1976), David A. Dickey, Wayne A. Fuller (1979, 1981)
- Odhadujeme model

  $$\Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \cdots + \theta_k \Delta y_{t-k} + u_t,$$

  pričom musíme
  - rozhodnúť, či zahrnúť konštantu α a lineárny trend βt (podľa toho, či ich obsahuje proces y)
  - určiť k (podľa informačných kritérií)
- Zaujíma nás potom t-štatistika zo sigifikancie koeficienta pri y_{t-1}, ale so správnymi kritickými hodnotami

ADF test - kritické hodnoty

- James G. MacKinnon (1991) - dostupné ako súčasť doplnenej verzie z roku 2010: http://ideas.repec.org/p/qed/wpaper/1227.html
- Simulačne získané hodnoty:

Table 1. Response Surface Estimates of Critical Values

N	Variant	Level	Obs.	β_∞	(s.e.)	β_1	β_2
1	no constant	1%	600	−2.5658	(0.0023)	−1.960	−10.04
		5%	600	−1.9393	(0.0008)	−0.398	
		10%	560	−1.6156	(0.0007)	−0.181	
1	no trend	1%	600	−3.4336	(0.0024)	−5.999	−29.25
		5%	600	−2.8621	(0.0011)	−2.738	−8.36
		10%	600	−2.5671	(0.0009)	−1.438	−4.48
1	with trend	1%	600	−3.9638	(0.0019)	−8.353	−47.44
		5%	600	−3.4126	(0.0012)	−4.039	−17.83
		10%	600	−3.1279	(0.0009)	−2.418	−7.58

ADF test - kritické hodnoty

- Ak v regresii použijeme T dát, kritická hodnota je $\beta_\infty + \beta_1/T + \beta_2/T^2$
- V našom prípade zo simulácií: konštanta bez trendu, $T = 200$:

N	Variant	Level	Obs.	β_∞	(s.e.)	β_1	β_2
1	no trend	1%	600	$\boxed{-3.4336}$	(0.0024)	$\boxed{-5.999}$	$\boxed{-29.25}$
		5%	600	$\boxed{-2.8621}$	(0.0011)	$\boxed{-2.738}$	$\boxed{-8.36}$
		10%	600	-2.5671	(0.0009)	-1.438	-4.48

- Dostaneme:
  - pre 1 percento: $-3.4336 - 5.999/100 - 29.25/200^2 = -3.451$
  - pre 5 percent: $-2.8621 - 2.738/200 - 8.36/200^2 = -2.879$
- Porovnajme s t-rozdelením (úplne iné) a s kvantilmi zo simulácií (ok)

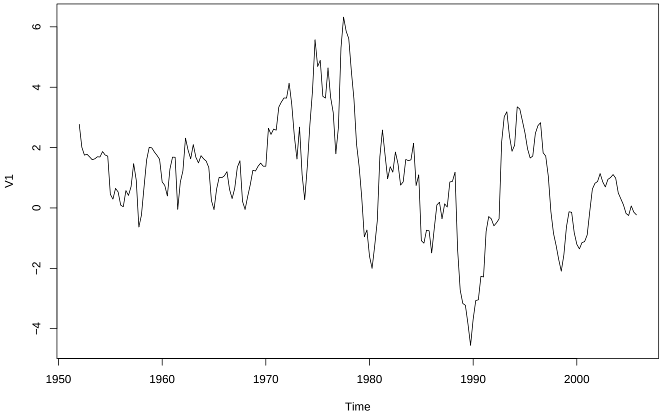
ADF test v R-ku

Funkcia ur.df

- Balík `urca` (**ur** = unit root, **ca** = cointegration)
- Funkcia `ur.df` (**ur** = unit root, **df** = Dickey-Fuller) s parametrami:
  - `type`: možnosti sú `drift` (konštanta bez lineárneho trendu) `trend` (konštanta aj lineárny trend) `none` (nič)
  - `lags`: maximálny počet lagov
  - `selectlags`: kritérium, podľa ktorého sa vyberá počet lagov (informačné kritériá `AIC`, `BIC`)

Príklad použitia

- Dáta `spread` z predchádzajúcich prednášok (rozdiel dlhodobej a krátkodobej úrokovej miery)



- Spravíme:

```
library(urca)
ur.df(spread,              # data
      type = "drift",      # konstanta bez trendu
      lags = 8,            # 8 lagov = 2 roky
      selectlags = "BIC" # Bayesovo kriterium
      )
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root / Cointegratio
## ###############################################
##
## The value of the test statistic is: -3.9112 7.6595
```

- Vypíšeme `summary`, aby sme dostali aj kritické hodnoty

```r
summary(ur.df(spread, type = "drift",
              lags = 8, selectlags = "BIC"))
```

- Odhadnutá regresia a testovacia štatistika (vo všeobecnosti je konštantný člen označený ako `Intercept` a lineárny trend `tt`):

y_{t-1}

Δy_{t-1}

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.10035    0.05592   1.794 0.074240 .
z.lag.1      -0.10720    0.02741  -3.911 0.000125 ***
z.diff.lag    0.29007    0.06706   4.326 2.38e-05 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7006 on 204 degrees of freedom
Multiple R-squared:  0.1216,    Adjusted R-squared:  0.113
F-statistic: 14.12 on 2 and 204 DF,  p-value: 1.806e-06


Value of test-statistic is: -3.9112 7.6595
```

- Testovacia štatistika a kritické hodnoty

prvá štatistika

```
Value of test-statistic is: -3.9112 7.6595

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

prvý riadok kritických hodnôt

- Kritérium: Hypotéza o jednotkovom koreni sa zamieta, ak je štatistika menšia ako kritická hodnota
- V našom prípade
  - hypotézu o jednotkovom koreni zamietame
  - dáta teda netreba diferencovať

Cvičenie 1

- *E. Ghaemi et al.: Improving the ARIMA Model Prediction for Water Quality Parameters of Urban Water Distribution Networks (Case Study: CANARY Dataset). International Journal of Environmental Research 16, 2022.*
- Vysvetlite, čo vyplýva z výsledkov v tabuľke pre daný časový rad

Case Study

The real data collected from US water infrastructure in CANARY (2012) were analyzed in this study. The water quality data, including the Cl (mg/L), pH, and TOC (ppb), were measured in 4 months with 5-min intervals and indicated the normal operation of the water network. Of the entire data, 70% were used for training and the remaining for testing the predictive model. The training data were used to extract the ARIMA model under normal conditions of the network. The test data were used for examining the performance of the extracted model on raw data with Static and Rolling Sample methods.

Table 1 Augmented Dickey Fuller test result

Test critical values	Cl	
	t-Statistic	Prob.
	– 10.7476	0.0000
1% level	– 3.9584	
5% level	– 3.4100	
10% level	– 3.1267	

Cvičenie 2: Interpretácia regresie z výstupu

- Napíšte regresiu, ktorá sa odhadla v tomto ACF teste.
- Aká hypotéza o parametroch regresie sa testuje?
- Odvoďte AR model pre dáta, ktorý regresia po úprave definuje.
- Odvoďte, že testovaná hypotéza zodpovedá jednotkovému koreňu

```
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-2.62437 -0.35488 -0.00071  0.35322  2.54618

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.10051    0.05552   1.810 0.071692 .
z.lag.1     -0.10514    0.02800  -3.755 0.000225 ***
z.diff.lag1  0.29195    0.06743   4.330 2.32e-05 ***
z.diff.lag2 -0.01416    0.06898  -0.205 0.837603
```

Cvičenie 3: Interpretácia regresie z výstupu

- Zopakujte pre nasledujúci výstup:

```
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.9585 -0.4996  0.2749  0.7940  3.1786 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
z.lag.1     -0.17906    0.09112  -1.965 0.052426 .  
z.diff.lag1 -0.59969    0.11658  -5.144 1.51e-06 ***
z.diff.lag2 -0.40873    0.11889  -3.438 0.000882 ***
z.diff.lag3 -0.30658    0.10005  -3.064 0.002863 ** 
```

Hľadanie ARIMA modelu: zhrnutie

Postup

- ARMA modely sú modelmi pre stacionárne dáta $\Rightarrow$
  - Čo sme vedeli: Ak majú trend, nie sú stacionárne.
  - Nové: **Ak majú jednotkový koreň, nie sú stacionárne a treba ich zdiferencovať** (a znovu sa pozrieť na diferencie).
- Postup:
  - Ak treba, zdiferencujeme najskôr dáta kvôli odstráneniu trendu
  - Potom otestujeme prítomnosť jednotkového koreňa
  - Ak v dátach je jednotkový koreň, zdiferencujeme ich a znovu otestujeme na jednotkový koreň (a diferencujeme ďalej, kým nedostaneme dáta bez jednotkového koreňa)
  - Pre diferencie, v ktorých už nie je trend ani jednotkový koreň, hľadáme ARMA model
- Ak sa dáta diferencujú, tak väčšinou len raz, výnimočne dvakrát.
- Ak sa diferencuje aspoň dvakrát, v modeli sa uvažuje nulový konštantný člen

ARIMA modely: poznámky

Potreba diferencovania: empiricky z odhadu ACF

- Nutnosť diferencovania sa dá čakať, ak ACF klesá príliš pomaly
- Ukážka z článku: *J. Lin: Improved Markowitz portfolio investment model based on ARIMA model and BP neural network. In 2022 IEEE 2nd International Conference on Data Science and Computer Application.*

ARIMA modeling requires the data to be stationary. After time series plot and autocorrelation plot tests, the data were found to be unstable, and the results are shown in Figure 1.

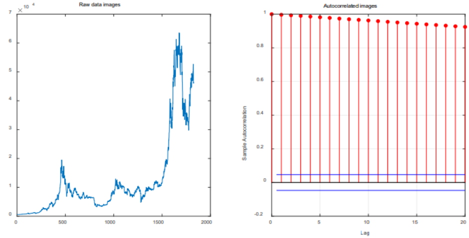


Fig. 1. Timing diagram before differencing and autocorrelation diagram

- Ukážka z článku *N. H. Hussin et al. (2021). Forecasting Wind Speed in Peninsular Malaysia: An Application of ARIMA and ARIMA-GARCH Models. Pertanika Journal of Science & Technology, 29(1).*
- Tiež sa tieto dáta diferencovali.

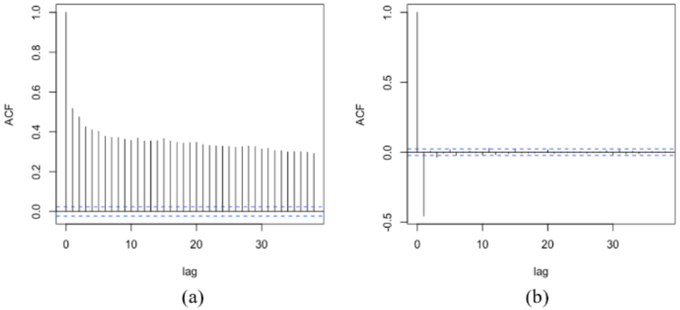


Figure 4. ACF plot for station NS1; (a) observation data and (b) after first difference.

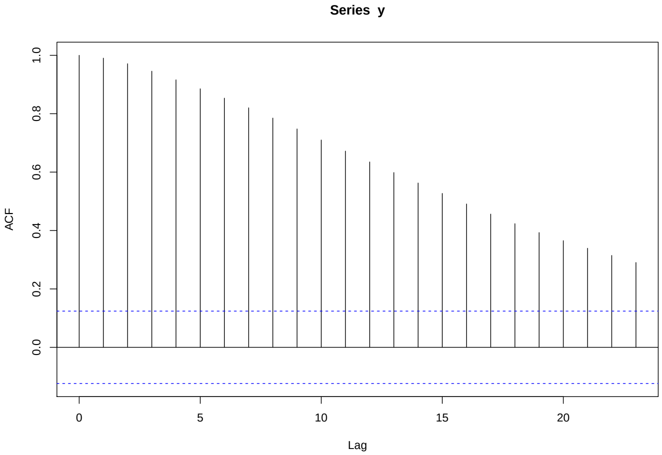
Simulované dáta

- Vygenerujeme simuláciu ARIMA(2,1,0) procesu (diferencie sú rovné `x`, či je AR(2) proces):

```
set.seed(123)
x <- arima.sim(model = list(ar=c(0.5, 0.2)), n = 250)
y <- cumsum(x)  # nase data na modelovanie
```

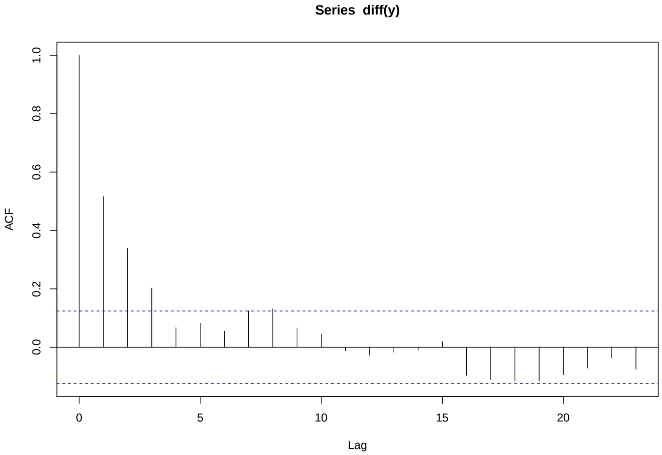
- Odhad ACF pre nestacionárne dáta:

```
acf(y)
```



- Odhad ACF pre stacionárne dáta:

```
acf(diff(y))
```



ARIMA modely pre dáta a pre diferencie

```
sarima(data, p, 1 , q)
sarima.for(diff(data) n.ahead = ..., p, 0, q)

sarima(diff(data), p, 0, q)
sarima.for(data n.ahead = ..., p, 1, q)
```

- Rozdiel sa prejaví pri predikciách:
  - v prvom prípade sa budú predikovať hodnoty premennej data
  - v druhom prípade sa budú predikovať diferencie
- Pri zápise modelu treba dať pozor na konštantný člen:
  - xmean je stredná hodnota procesu
  - constant je konštantný člen, v našom značení delta

Rozdiel v konštantnom člene

- Rozdiel medzi `xmean` a `constant`:

```
model1 <- sarima(y, 2, 1, 0, details = FALSE)
model2 <- sarima(diff(y), 2, 0, 0, details = FALSE)
```

```
model1$fit$coef
```

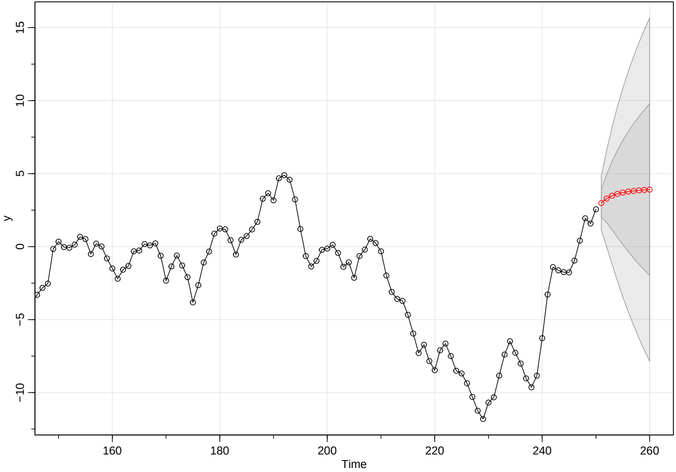
```
##        ar1        ar2   constant
## 0.46660982 0.09806022 0.02055461
```

```
model2$fit$coef
```

```
##        ar1        ar2      xmean
## 0.46660981 0.09806022 0.02055457
```

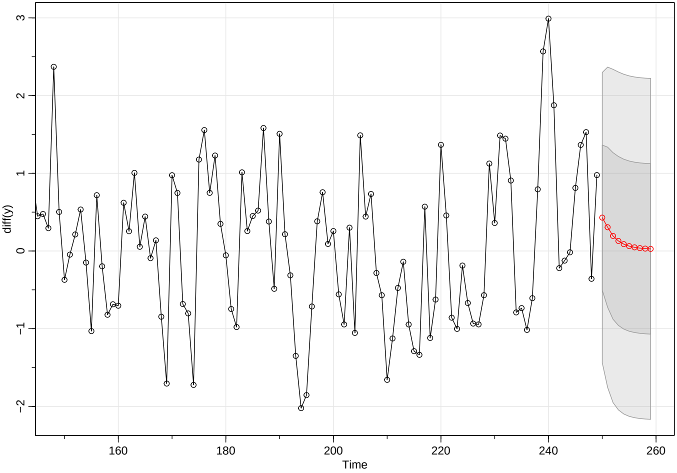
Rozdiel medzi v predikciách

```
sarima.for(y, n.ahead = 10, 2, 1, 0) # data
```



- Predikcie diferencií:

```
sarima.for(diff(y), n.ahead = 10, 2, 0, 0) # diferencie
```



Automatizovaný výber modelu

- Funkcia `auto.arima` z balíka `forecast`
- Nekontroluje rezíduá, uvažuje zvolené informačné kritérium
- Môže sa hodiť
  - ako inšpirácia pre vyskúšanie modelu pre dáta (napr. ak na základe ACF a PACF nemáme konkrétny tip)
  - pri automatickom výbere (ak si ho určujeme my, treba jasne popísať použité parametre, ak sme nepoužili defaultné)
- Treba si pozrieť výstup, či sa odhadol model s konštantou (tiež je to podľa informačného kritéria), resp. nastaviť parametre

```r
library(forecast)
auto.arima(y)
```

```
## Series: y
## ARIMA(1,1,1)
##
## Coefficients:
##          ar1      ma1
##       0.6416  -0.1697
## s.e.  0.0873   0.1100
##
## sigma^2 estimated as 0.8803:  log likelihood=-336.6
## AIC=679.19   AICc=679.29   BIC=689.75
```

```r
auto.arima(y, stepwise = FALSE, approximation = FALSE)
```

```
## Series: y
## ARIMA(2,1,0)
##
## Coefficients:
##          ar1     ar2
##       0.4667  0.0982
## s.e.  0.0631  0.0631
##
## sigma^2 estimated as 0.8796:  log likelihood=-336.51
## AIC=679.02   AICc=679.11   BIC=689.57
```

```r
auto.arima(y, ic = "bic")
```

```
## Series: y
## ARIMA(1,1,0)
##
## Coefficients:
##          ar1
##       0.5178
## s.e.  0.0542
##
## sigma^2 estimated as 0.8847:  log likelihood=-337.71
## AIC=679.42   AICc=679.47   BIC=686.46
```

```r
auto.arima(y, include.mean = TRUE)
```

```
## Series: y
## ARIMA(1,1,1) with drift
##
## Coefficients:
##          ar1      ma1   drift
##       0.6412  -0.1694  0.0206
## s.e.  0.0874   0.1100  0.1362
##
## sigma^2 estimated as 0.8838:  log likelihood=-336.59
## AIC=681.17   AICc=681.34   BIC=695.24
```