

MODELOVANIE ČASOVÝCH RADOV V PROSTREDÍ R

LANÍK PETER, MANE

Školiteľ: Doc. RNDr. Marta Vrábelová Csc.

Abstrakt. Analýza časových radov je jednou z najvýznamnejších aplikácií štatistických metód v ekonómii. Rozsiahlosť a náročnosť výpočtov si žiada zapojenie výpočtovej techniky do analýzy. Existuje množstvo štatistických softvérov, väčšina je však ekonomicky nákladná. Avšak v Bell Laboratories John Chambers a jeho kolegovia vyvinuli programovací jazyk R a výpočtové prostredie. R je komplexný systém na manipuláciu s údajmi, ich spracovanie, analýzu a grafické zobrazenie. Program je voľne šírený vďaka licencií GNU~GPL. Cieľom práce je demonštrovať komplexnú analýzu časového radu, ktorá vedie k získaniu kvalitnej predikcii a priblížiť čitateľovi použitý softvér.

Kľúčové slová: Časový rad. Dekompozícia. Vyrovnávanie. Prognózovanie. Fourierova analýza. Box Jenkinsova metodológia. R.

TIME SERIES MODELING IN R ENVIRONMENT

Abstract. Time series analysis is one of the most important application of statistical methods in economics. Extent and severity calculation requires the involvement of computer technology in the analysis. There is a lot of statistical software, but most are expensive. However, in the Bell Laboratories John Chambers and his colleagues developed the R programming language and computing environment. R is a comprehensive system for data manipulation, processing, analysis and graphical display. The program is freely distributed thanks to GNU ~ GPL license. The aim of this work is to demonstrate a comprehensive analysis of time series, leading obtain quality predictions and to bring the reader to use the software.

Keywords: Time series. Decomposition. Smoothing. Forecasting. Fourier analysis. Box Jenkins methodology. R.

1 Úvod

Práca je rozdelená na 2 časti. Prvá časť stručne opisuje teóriu, ktorá je podkladom k analýze. Druhá časť je analýza akciového indexu NIKKEI 225, čo je najviac sledovaným indexom ázijskej akciovej burzy. Je tvorený 225 spoločnosťami, ktoré sú do tohto indexu vyberané zo skupiny 450 spoločností s najvyššou likviditou. NIKKEI 225 by mal v prvej sekcii reprezentovať výkonnosť akcií podľa trhovej kapitalizácie. Pretože všeobecne dochádza k rozširovaniu trhov, skladba komponentov býva čas od času vyvažovaná, aby sa zabezpečilo, že všetky položky indexu sú nielen vysoko likvidné, ale aj reprezentujú štruktúru japonského priemyslu. Akciové indexy sú dôležitým ukazovateľom vývoja kurzov cenných papierov. Odrážajú súčasný stav a vývoj kurzov cenných papierov, monitorujú dlhodobé trendy, preto sa pokúsime odhaliť vývoj indexu NIKKEI 225. Súčasťou textu je aj zdrojový kód s podrobným popisom.

2 Teoretické predpoklady analýzy

Časový rad je chronologicky usporiadaná postupnosť hodnôt, ktorú zapisujeme jedným z nasledujúcich spôsobov y_1, y_2, \dots, y_n , alebo $y_t, t=1, 2, \dots, n$. Základným predpokladom je, že časový rad sa skladá z nasledujúcich zložiek: trend, sezónna, cyklická a náhodná zložka. Podľa toho, akou matematickou operáciou sú tieto zložky spojené, rozlišujeme modely aditívne, multiplikatívne a zmiešané. V nasledujúcom budeme predpokladať nasledovný aditívny model: $Y_t = C_t + E_t$. Teda model obsahujúci cyklickú a náhodnú zložku. Obe zložky budeme modelovať samostatne a taktiež ich budeme samostatne predikovať. Na záver ich sčítame, čím dostaneme výsledný model a predikcie.

Modelovanie cyklickej zložky

Vhodnou metódou modelovania cyklickej zložky je Fourierova analýza, ktorá skúma časový rad v spektrálnej doméne. Vychádza z predpokladu, že časový rad je nekonečnou zmesou sínusových a kosínusových kriviek s rôznymi frekvenciami a amplitúdami. Teda,

$$y_t = \mu + \sum_{k=1}^q [U_{k1} \cos(2\pi \omega_k t) + U_{k2} \sin(2\pi \omega_k t)],$$

kde μ je priemer časového radu, $U_{k1}, U_{k2}, k=1, 2, \dots, q$ sú nezávislé parametre s nulovou strednou hodnotou a rozptylom σ_k^2 , ω_k sú navzájom rôzne frekvencie.

Definícia. Pre ľubovoľnú funkciu $\{a_t; t=0, \pm 1, \pm 2, \dots\}$, spĺňajúcu podmienku

$$\sum_{t=-\infty}^{\infty} |a_t| < \infty,$$

definujeme Fourierov transformačný pár ako

$$A(\omega) = \sum_{t=-\infty}^{\infty} a_t e^{-2\pi\omega t},$$
$$a_t = \int_{-\pi/2}^{\pi/2} A(\omega) e^{2\pi\omega t} d\omega.$$

Definícia. Pre dané $\{y_1, \dots, y_n\}$ (napr. časový rad) definujeme diskretnú Fourierovu transformáciu ako

$$d(\omega_j) = \sqrt{n} \sum_{t=1}^n x_t e^{-2\pi i \omega_j t}, \text{ pre } j=0, 1, \dots, n-1$$

a kde frekvencie $\omega_j = j/n$ sa nazývajú Fourierove (alebo fundamentálne) frekvencie.

Definícia. Pre dané $\{y_1, \dots, y_n\}$ (napr. časový rad) definujeme periodogram ako

$$I(\omega_j) = |d(\omega_j)|^2 \text{ pre } j=0, 1, \dots, n-1.$$

Definícia. Pre dané $\{y_1, \dots, y_n\}$ (napr. časový rad) definujeme kosínusovú transformáciu

$$d_c(\omega_j) = \sqrt{n} \sum_{t=1}^n x_t \cos(2\pi \omega_j t)$$

a sínusovú transformáciu

$$d_s(\omega_j) = \sqrt{n} \sum_{t=1}^n x_t \sin(2\pi \omega_j t),$$

kde $\omega_j = j/n$ pre $j=0, 1, \dots, n-1$.

Vychádzajúc z nasledujúcich definícií Fourierovho transformačného páru, diskkrétnej Fourierovej transformácie, periodogramu a sínusovej, kosínusovej transformácie sa dá dokázať, že $\{y_1, \dots, y_n\}$ (napr. časový rad) kde n je nepárne platí

$$y_t = a_0 + \sum_{j=1}^m [a_j \cos(2\pi \omega_j t) + b_j \sin(2\pi \omega_j t)], \quad (1)$$

kde $m = (n-1)/2$ presne pre $t=1, \dots, n$. Využitím vzorcov viacnásobnej regresie, dostaneme :

$$a_0 = \bar{x},$$

$$a_j = \frac{2}{n} \sum_{t=1}^n x_t \cos(2\pi \omega_j t) = \frac{2}{\sqrt{n}} d_c(\omega_j),$$

$$b_j = \frac{2}{n} \sum_{t=1}^n x_t \sin(2\pi \omega_j t) = \frac{2}{\sqrt{n}} d_s(\omega_j).$$

Avšak my chceme odhadnúť iba cyklickú zložku, takže využijeme, že periodogram je dobrý ukazovateľ periodicity. Nájdeme 4 najväčšie hodnoty periodogramu a k nim prislúchajúce frekvencie. Tieto frekvencie potom dosadíme do (1) a získame odhad cyklickej zložky radu.

Modelovanie náhodnej zložky

Pri modelovaní náhodnej zložky budeme vychádzať z odhadu náhodnej zložky, ktorý získame ako rozdiel pôvodných hodnôt radu a odhadu cyklickej zložky. Náhodná zložka musí spĺňať podmienky bieleho šumu. Teda musí mať nulovú strednú hodnotu, hodnoty musia byť navzájom nekorelované a musia mať normálne rozdelenie pravdepodobnosti. Po overení týchto podmienok budeme modelovať náhodnú zložku ARMA modelom, ktorý predpokladá že hodnota časového radu v čase t je daná vzťahom

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q},$$

kde y_t sú pozorované hodnoty v čase t , α_i, β_i sú neznáme parametre a ε_t je biely šum. Po odhadnutí parametrov testujeme vhodnosť modelu pomocou Portmanteauvej štatistiky

$$Q = n \sum_{k=1}^K [r_k(\hat{\varepsilon})]^2,$$

kde $r_k(\hat{\varepsilon})$, je odhadnutá autokorelácia odhadnutého bieleho šumu. Ak $Q < X_{K-p-1}^2(\alpha)$, tak model ARMA(p,q) je správne definovaný. $K \sim \sqrt{(n)}$.

3 Analýza indexu NIKKEI 225

Načítanie dát je možné priamo z internetu, alebo zo súboru, uvedieme obe možnosti.

Zo súboru

```
>raw<-  
read.csv(file="mn225.csv", skip=4, colClasses=c("character",NULL,"numeric")  
)
```

Z internetu

```
>raw<-  
read.csv(file="http://wrenresearch.com.au/downloads/files/mn225.csv", skip  
=4, colClasses=c("character",NULL,"numeric"))
```

K analýze nám stačí tretí stĺpec tabuľky, pre zmenšenie rozptylu budeme dáta transformovať pomocou logaritmu a ďalej budeme pracovať s transformovanými údajmi. Program považuje načítané hodnoty za reťazec, takže ich treba zmeniť na čísla.

```
> data<-log(as.numeric(raw[,3]))
```

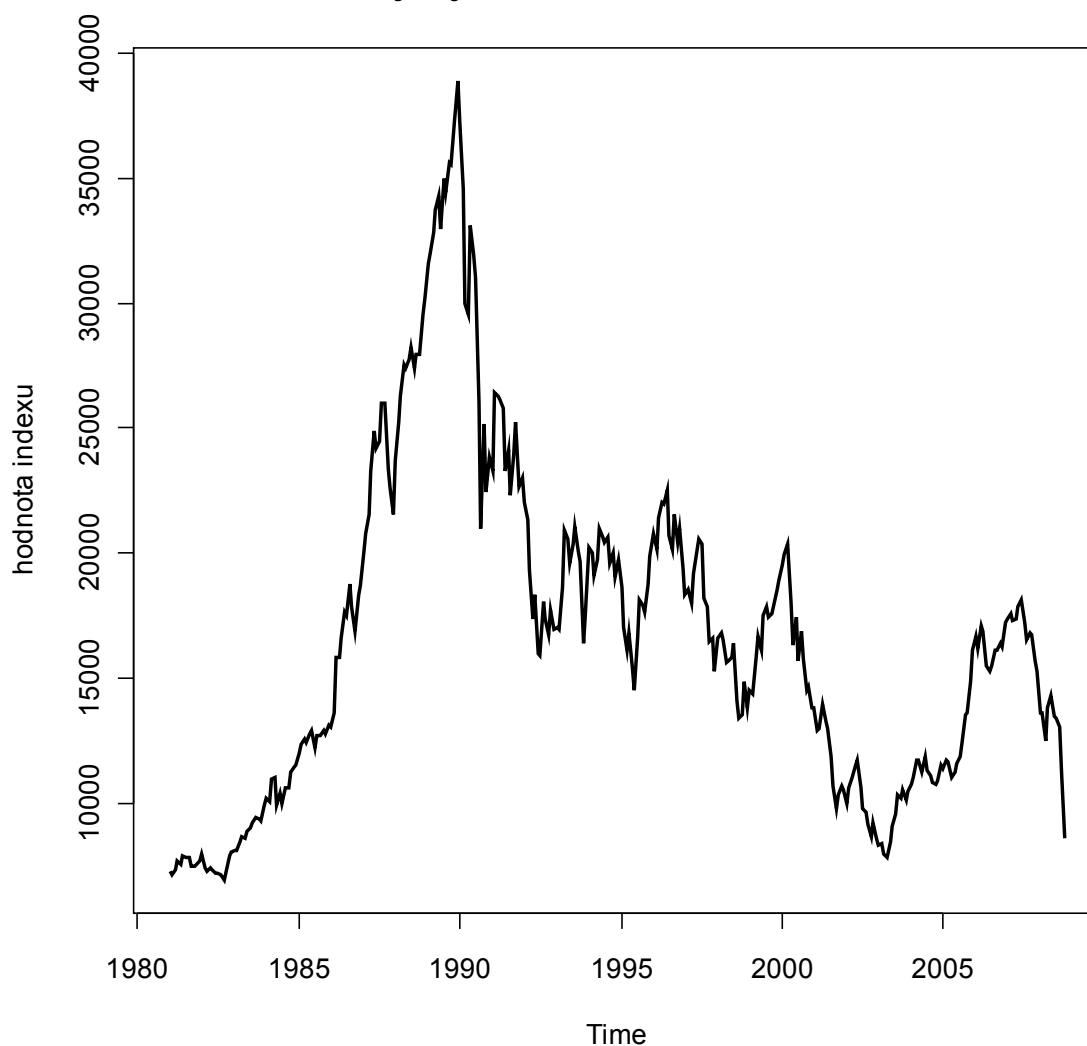
Teraz vektor prevedieme na časový rad, ktorý začína januárom 1981.

```
> nikkei=ts(data[133:466], start=1981, f=12)
```

Tento časový rad si graficky znázorníme, pri zobrazení použijeme inverznú transformáciu, aby sme zobrazili reálne dáta.

```
> plot(exp(nikkei), col="black", ylab="hodnota indexu", lwd=2, main="Vývoj  
akciového indexu Nikkei")
```

Vývoj akciového indexu Nikkei



Vycentrujeme časový rad odpočítaním priemeru.

```
> d.nikkei=nikkei-mean(nikkei)
```

Teraz môžeme zahájiť samotné modelovanie. Vypočítame periodogram využitím fourierovej transformácie.

```
> P = abs(fft(d.nikkei)/334)^2# funkcia fft -fast Fourier transform
```

V prvej polovici periodogramu nájdeme 4 najväčšie hodnoty. Prvú hodnotu však ignorujeme, keďže tá prislúcha 0-ovej frekvencii a predstavuje priemer časového radu.

```
> hladane= sort(P[2:167],decreasing=T)[1:4]/334
```

Už vieme, ktoré hodnoty periodogramu v ňom hľadáme, ale nevieme, ktorým periódam tieto hodnoty prislúchajú. Vieme však, že poradie v periodograme a hodnota periodogramu spolu súvisia a to, ak p je hľadaná hodnota periodogramu, ktorej prislúcha poradie i ,

tak $P[i]/n=p$ a $perioda=\frac{n-1}{i-1}$. Náš postup bude nasledovný: zoberieme jednu hľadanú hodnotu periodogramu, porovnáme ju s každou hodnotou periodogramu z prvej polovice okrem prvej hodnoty. Po nájdení tejto pozície si ju zaznamenáme. Postup opakujeme pre ďalšiu hľadanú hodnotu periodogramu. Ručné hľadanie by bolo extrémne náročné, preto využijeme cyklus s pevným počtom opakovaní, pomocou ktorého tieto pozície nájdeme.

```
> pozicia=0 # vytvorenie premennej
> for (j in 1:4){ for (i in 2:167){
+   if (isTRUE(all.equal(P[i]/334,hladane[j]))) {pozicia[j]=i-1}}
> period=333/pozicia # 333 kvoli tomu ze 1.prvok reprezentuje priemer
radu
```

Štyri najvýznamnejšie periódy.

```
> period
[1] 333.00 111.00 166.50 66.60
```

Keď už máme periódy, môžeme namodelovať cyklickú zložku radu podľa vzťahu (1). Problém sa obmedzil len na rozsahovú náročnosť výpočtov. Na výpočet sumy využijeme cyklus s pevným počtom opakovaní, ktorý umiestnime do funkcie. Toto nám umožní univerzálnejšie rátať Fourierove koeficienty a tiež hodnotu aproximácie Fourierového radu. Stačí raz definovať funkcie a potom ich len zavolať so správnymi parametrami. Inak by sme museli vždy zapisovať sumačný cyklus.

```
> furier.koef<-function(period,rad,Trig){
#vlastna funkcia na vypocet Fourierovych koeficientov
n=length(rad)
k=rep(0,length(period))
for (i in 1:length(period)){for (t in 1:n){
k[i]=k[i]+rad[t]*Trig(2*pi*t/period[i])}}
furier.koef<-k*2/n}
```

Samotné volanie funkcií, v zátvorke uvedieme najvýznamnejšie periódy, skúmaný rad a či chceme počítat' sínusové alebo kosínusové koeficienty.

```
> s.koef<-furier.koef(period,d.nikkei,sin)
> c.koef<-furier.koef(period,d.nikkei,cos)
```

Keď máme koeficienty, môžeme vyrátať vyrovnaný rad tiež použitím vlastnej funkcie.

```
> furier.rad<-function(rad,s,c,period){
+ #funkcie ktora vytvori tzv Fourierov rad, je volana parametrami:
vyrovnany rad,sinusove Fourierove koeficienty , kosinusove Fourierove
koeficienty, najvyznamnejsie periody
+ furier<-rep(0,length(rad))
+ for (t in 1:length(rad)){ for (i in 1:length(period)){
+   furier[t]<-furier[t]+c[i]*cos(2*pi*t/period[i])
+s[i]*sin(2*pi*t/period[i])}}
+ furier.rad<-furier}
```

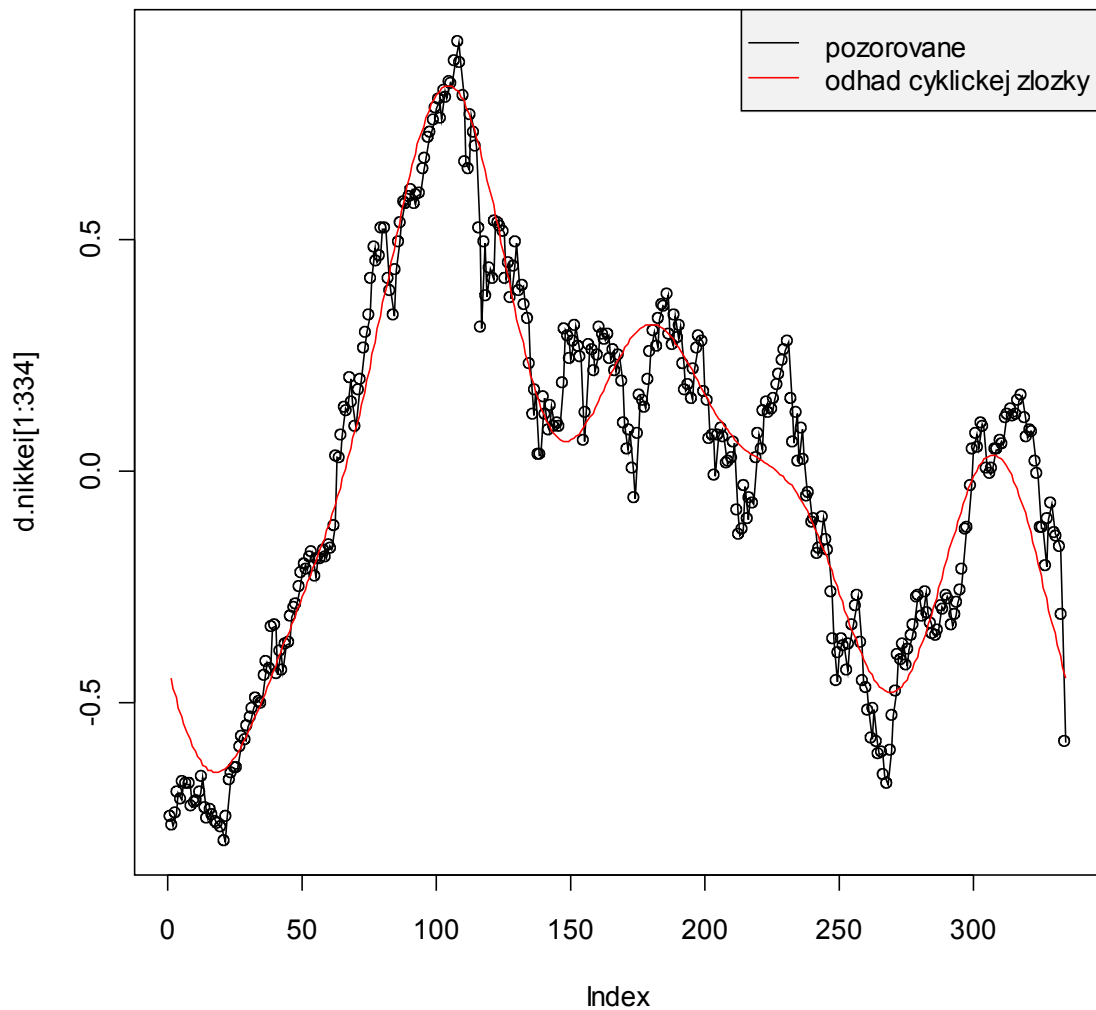
Volanie funkcie pre vyrovnaný rad. Výsledok graficky znázorníme.


```

> f.nikkei<-furier.rad(d.nikkei,s.koef,cos.koef,period)

> plot(d.nikkei[1:334],type="o")
> lines(f.nikkei, col="red")
> legend("topright", c("pozorovane","odhad cyklickej zlozky"),
lty=c("solid","solid"), pch=c(NA,NA), col=c("black","red"), bg="gray95")

```



Modelovanie cyklickej zložky je týmto ukončené. Dostali sme nasledovný model.

$$\begin{aligned}
\hat{C}_t = & 9,642834 - 0,35973895 \cos(2\pi t/333) + 0,12689725 \cos(2\pi t/111) \\
& - 0,09756944 \cos(2\pi t/166,5) - 0,09317855 \cos(2\pi t/66,6) \\
& + 0,20558181 \sin(2\pi t/333) - 0,23808357 \sin(2\pi t/111) \\
& - 0,15425599 \sin(2\pi t/166,5) - 0,08409671 \sin(2\pi t/66,6)
\end{aligned}$$

Ďalej budeme modelovať náhodnú zložku, ktorú získame ako $E_t = y_t - \hat{C}_t$

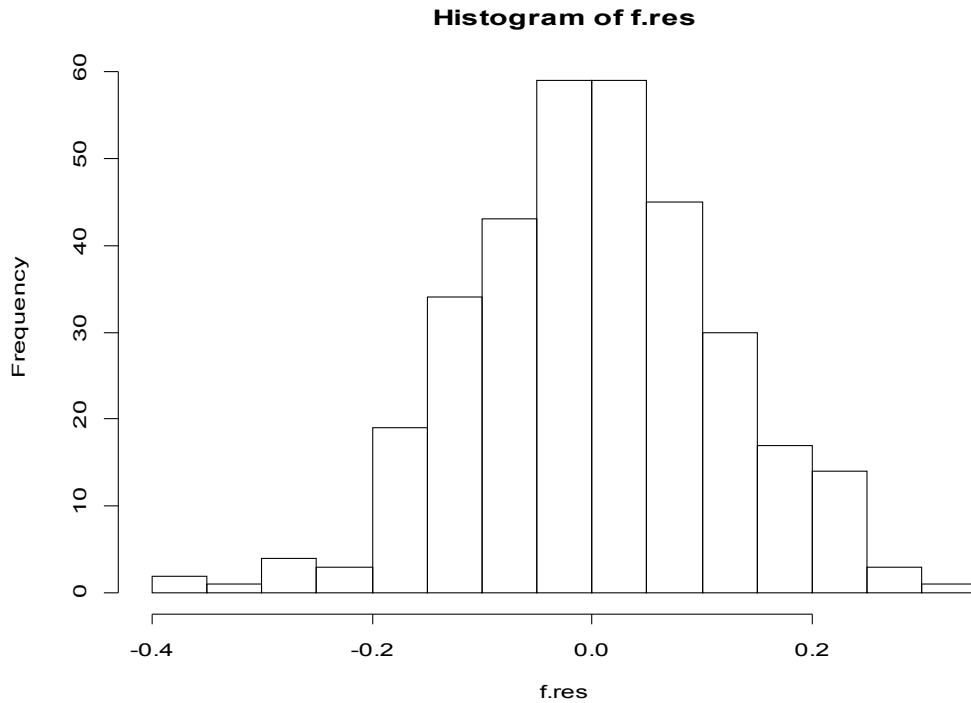
```

> f.res<-d.nikkei-f.nikkei

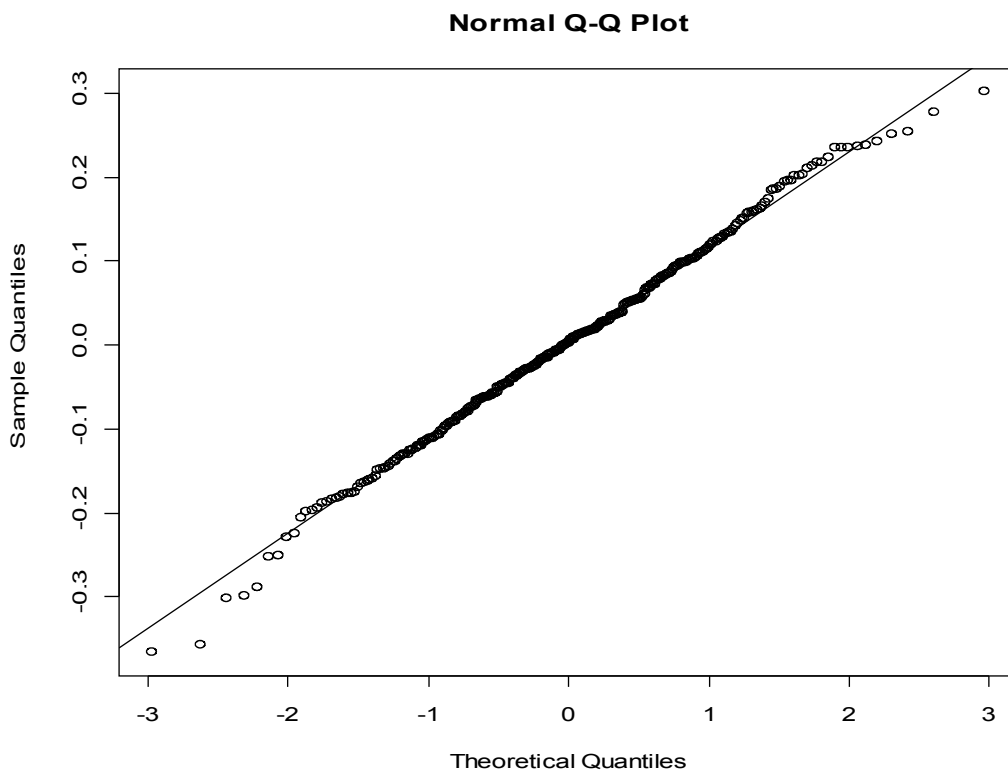
```

Pohľadom na normálny pravdepodobnostný graf a histogram overíme normalitu náhodnej zložky.

```
> hist(f.res)  
> qqnorm(f.res)
```



```
> qqline(f.res)
```



Normalitu overíme ešte Shapirovým-Wilkovým testom.

```
> shapiro.test(f.res)
```

```
Shapiro-Wilk normality test
```

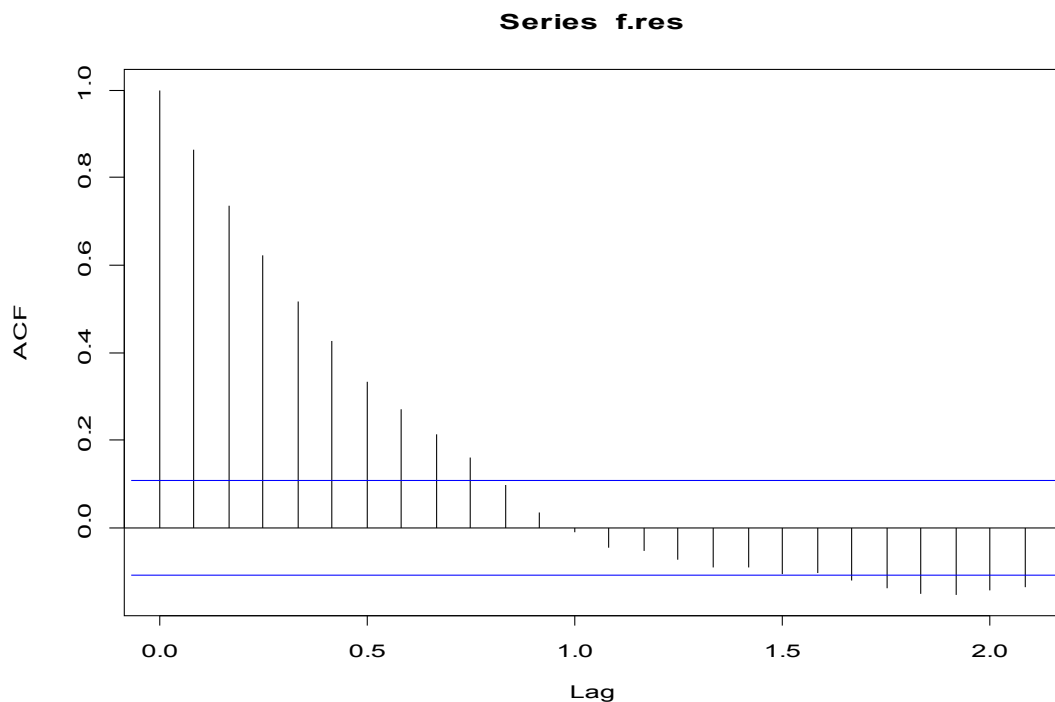
```
data: f.res
```

```
W = 0.996, p-value = 0.5587
```

Test nezamietol nulovú hypotézu, ktorá predpokladá normálne rozdelenie. Náhodnú zložku môžeme považovať za normálne rozdelenú náhodnú veličinu.

Vykreslíme autokorelačnú funciu.

```
> acf(f.res)
```



Z pohľadu na graf je ťažké posúdiť, či je autokorelácia významná alebo nie, preto toto posúdenie necháme na KPSS test. Zároveň určíme ARIMA model s odhadom parametrov.

```
> library(forecast)
```

```
> arima<-auto.arima(f.res,max.P=0,max.Q=0,stationary=F, trace=T,  
ic="aic")
```

```
ARIMA(2,0,2) with non-zero mean : 1e+20  
ARIMA(0,0,0) with non-zero mean : -487.1367  
ARIMA(1,0,0) with non-zero mean : -974.4252  
ARIMA(0,0,1) with non-zero mean : -740.6399  
ARIMA(2,0,0) with non-zero mean : -973.217  
ARIMA(1,0,1) with non-zero mean : -973.5426  
ARIMA(2,0,1) with non-zero mean : 1e+20  
ARIMA(1,0,0) with zero mean : -976.369  
ARIMA(0,0,0) with zero mean : -489.0921  
ARIMA(2,0,0) with zero mean : -975.1507
```

```

ARIMA(1,0,1) with zero mean      : -975.4964
ARIMA(2,0,1) with zero mean      : 1e+20
ARIMA(1,0,0) with non-zero mean : -974.4252

Best model: ARIMA(1,0,0) with zero mean

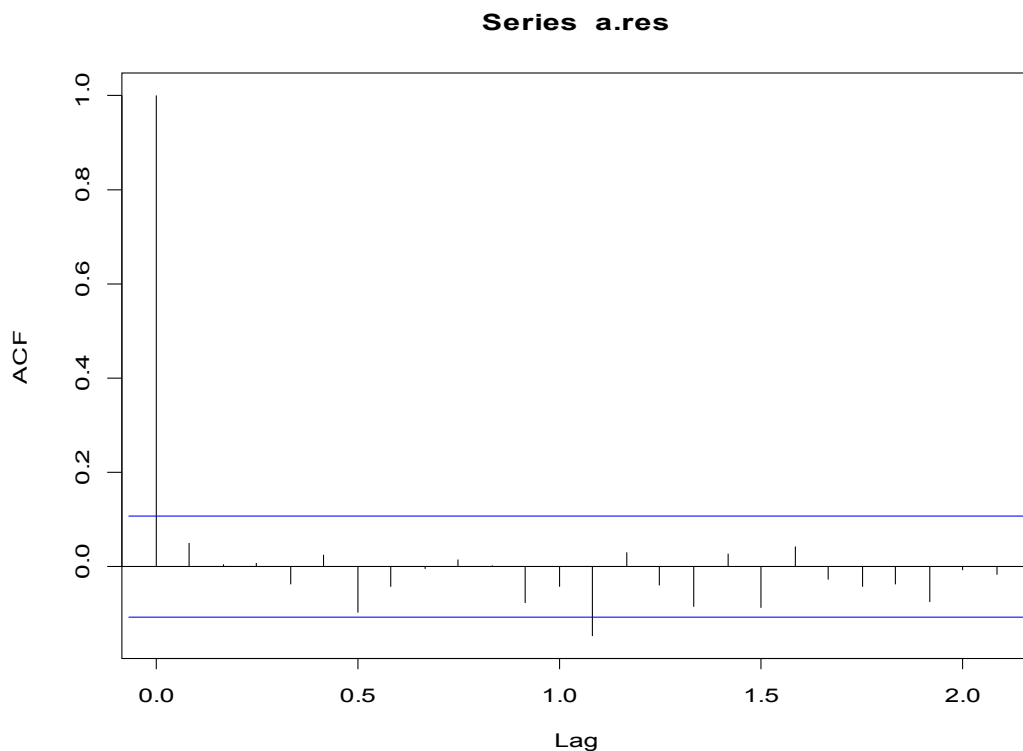
> summary(arima)
Series: f.res
ARIMA(1,0,0) with zero mean

Coefficients:
      ar1
      0.8826
s.e.    0.0265

sigma^2 estimated as 0.003150:  log likelihood = 487.27
AIC = -970.54   AICc = -970.51   BIC = -962.92

```

Najlepší model je ARIMA(1,0,0) $\hat{e}_t = 0,8826 e_{t-1} + \varepsilon_t$, $e_t \sim N(0; 0,00315)$, takže autokorelácia bola nevýznamná. O tomto rozhodol KPSS test, ktorý na hladine významnosti $\alpha=0,05$ nezamietol nulovú hypotézu o stacionarite časového radu. Pozrime sa ešte na ARMA reziduá, konkrétne na autokorelačnú.



Vhodnosť ARMA modelu overíme pomocou Portmanteauvej štatistiky.

```

> K = trunc(sqrt(length(a.res)))
> n=length(a.res)
> ac = as.vector(acf(a.res)$acf) # odhadute autokorelacie
> Q = n*sum(ac[1:K+1]^2) # posun o 1 kvoli r(0)
> Q # testova statistika
[1] 17.34596
> kh = qchisq(0.95,K-arma$arma[1]-arma$arma[2]) # kriticka hodnota
> kh
[1] 27.58711
> c("nezamietame H0 (H0:model ARMA je spravne spefikovany)", "zamietame H0
(H0:model ARMA je spravne spefikovany)")[(Q > kh)+1]
[1] "nezamietame H0 (H0:model ARMA je spravne spefikovany)"

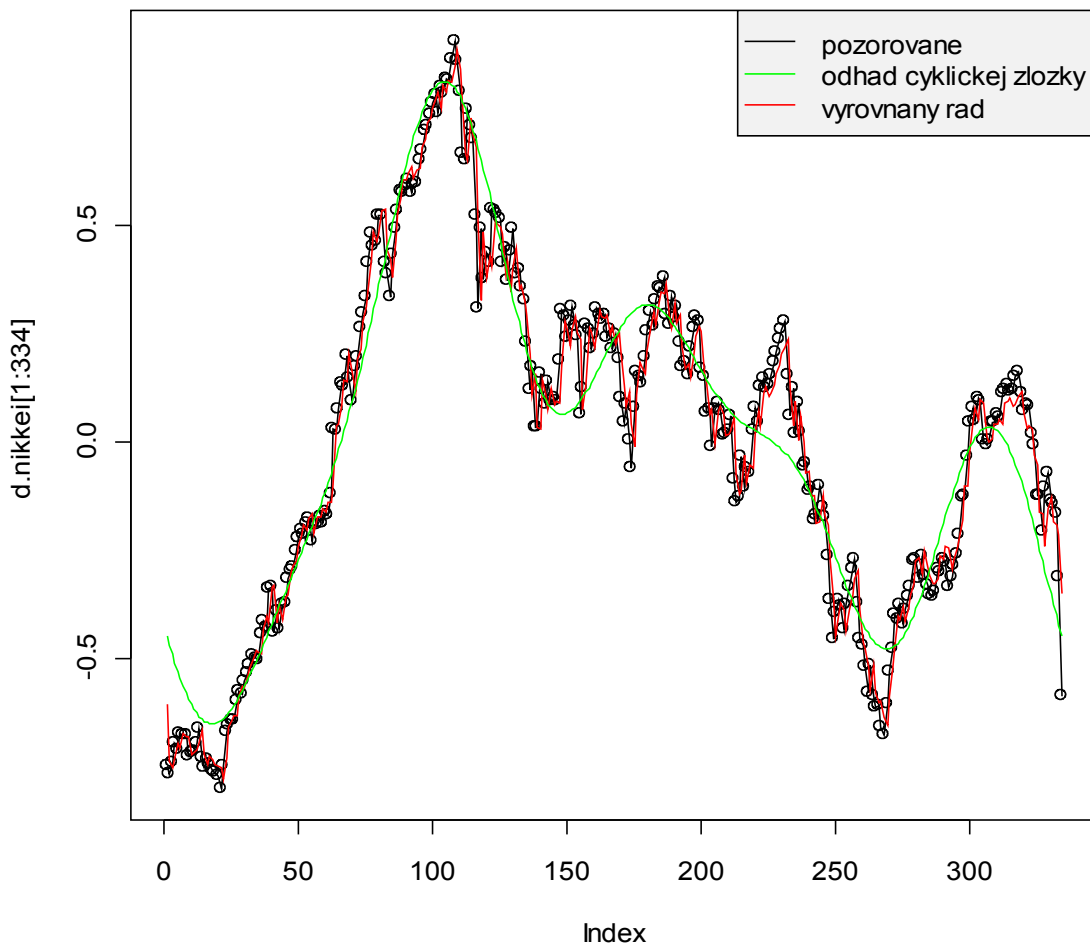
```

Teraz už môžeme skonštruovať náš vyrovnaný rad, graficky ho porovnať s pôvodným radom.

```

> vyr.rad<-d.nikkei-a.res
> plot(d.nikkei[1:334],type="o")
> lines(vyr.rad[1:334],col="red")
> lines(f.nikkei[1:334],col="green")
> legend("topright", c("pozorovane","odhad cyklickej zlozky","vyrovnaný
rad"), lty=c("solid","solid","solid"), pch=c(NA,NA,NA),
col=c("black","green","red"), bg="gray95")

```



Vypočítame priemernú relatívnu absolútnu chybu reziduí MAPE. Avšak MAPE vypočítame pre pôvodný rad (nezlogaritmovaný).

```
> mape<-function(rad,vyr.rad) {
+ m<-0
+ res<-rad-vyr.rad
+ for (i in 1:length(rad)) {m <- m + abs(res[i]/rad[i])}
+ mape<-(100*m)/length(rad)}
> rad<-exp(mean(nikkei)+vyr.rad)
> MAPE<-mape(exp(nikkei),rad)
> MAPE
[1] 4.312583
```

Priemerná absolútna chyba je 4,31%. To znamená, že priemerná hodnota vyrovnaného radu má odchýlku od skutočného priebehu časového radu +/- 4,31%.

Predpoveď cyklickej zložky získame dosadením príslušného $T > t$ do vzorca. Predpoveď náhodnej zložky získame použitím funkcie predict na premennú typu arima model.

```
> furier.rad.predpoved<-function(rad,s,c,period,T){
+ #funkcie ktora vytvori tzv Fourierov rad, je volana paramatrami
vyrovnavany rad,sinusove Fourierove koeficienty
+ # kosinusove Fourierove koeficienty, najvyznamnejsie periody
+ furier<-rep(0,length(rad)+T)
+ for (t in 1:length(rad)+T){ for (i in 1:length(period)){
+ furier[t]<-furier[t]+c[i]*cos(2*pi*t/period[i])
+s[i]*sin(2*pi*t/period[i])}}
+ furier.rad.predpoved<-furier[(length(rad)+1):(length(rad)+T)]}
>
>predpoved<furier.rad.predpoved(d.nikkei,s.koef,c.koef,period,12)+predict
(arima,12)$pred
```

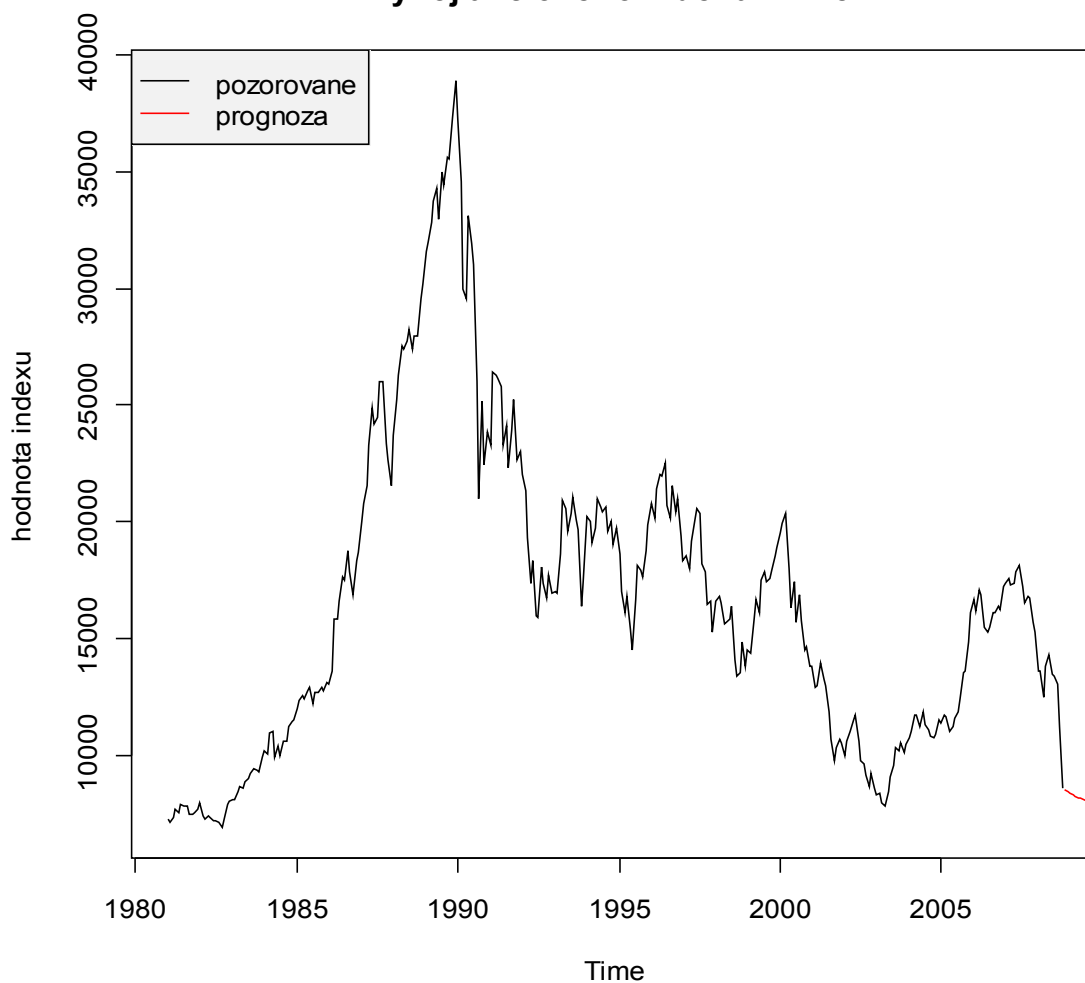
Na záver ešte spravíme inverzné transformácie v opačnom poradí, aby sme mohli predikcie doplniť do pôvodných dát.

```
> delog.predpoved<-exp(mean(nikkei)+predpoved)
```

2008 Nov	2008 Dec	2009 Jan	2009 Feb
8527,28	8469,35	8406,13	8340,17
2009 Mar	2009 Apr	2009 Maj	2009 Jun
8273,72	8208,69	8146,71	8089,16
2009 Jul	2009 Aug	2009 Seb	2009 Okt
8037,18	7991,7	7953,48	7923,13

```
> plot(exp(nikkei),ylab="hodnota indexu",main="Vývoj akciového indexu
Nikkei")
> lines(delog.predpoved,col="red")
> legend("topleft", c("pozorovane","prognoza"), lty=c("solid","solid"),
pch=c(NA,NA), col=c("black","red"), bg="gray95")
```

Vývoj akciového indexu Nikkei



Keďže časom sú už dostupné reálne hodnoty indexu, môžeme ešte porovnať naše predpovede z realitou. Hodnoty indexu sa dajú nájsť na

<http://finance.yahoo.com/q/hp?s=%5E%5EN225&a=10&b=4&c=2008&d=03&e=5&f=2009&g=m>
(4.4.2009), sledujeme stĺpec close.

mar 2009 8109,5

feb 2009 7568,42

jan 2009 7994,05

dec 2008 8859,56

nov 2008 8512,27

```
> realne<-c(8512.27,8859.56,7994.05,7568.42,8109.5)
```

```
> MAPE<-mape(realne,delog.predpoved[1:length(realne)])
```

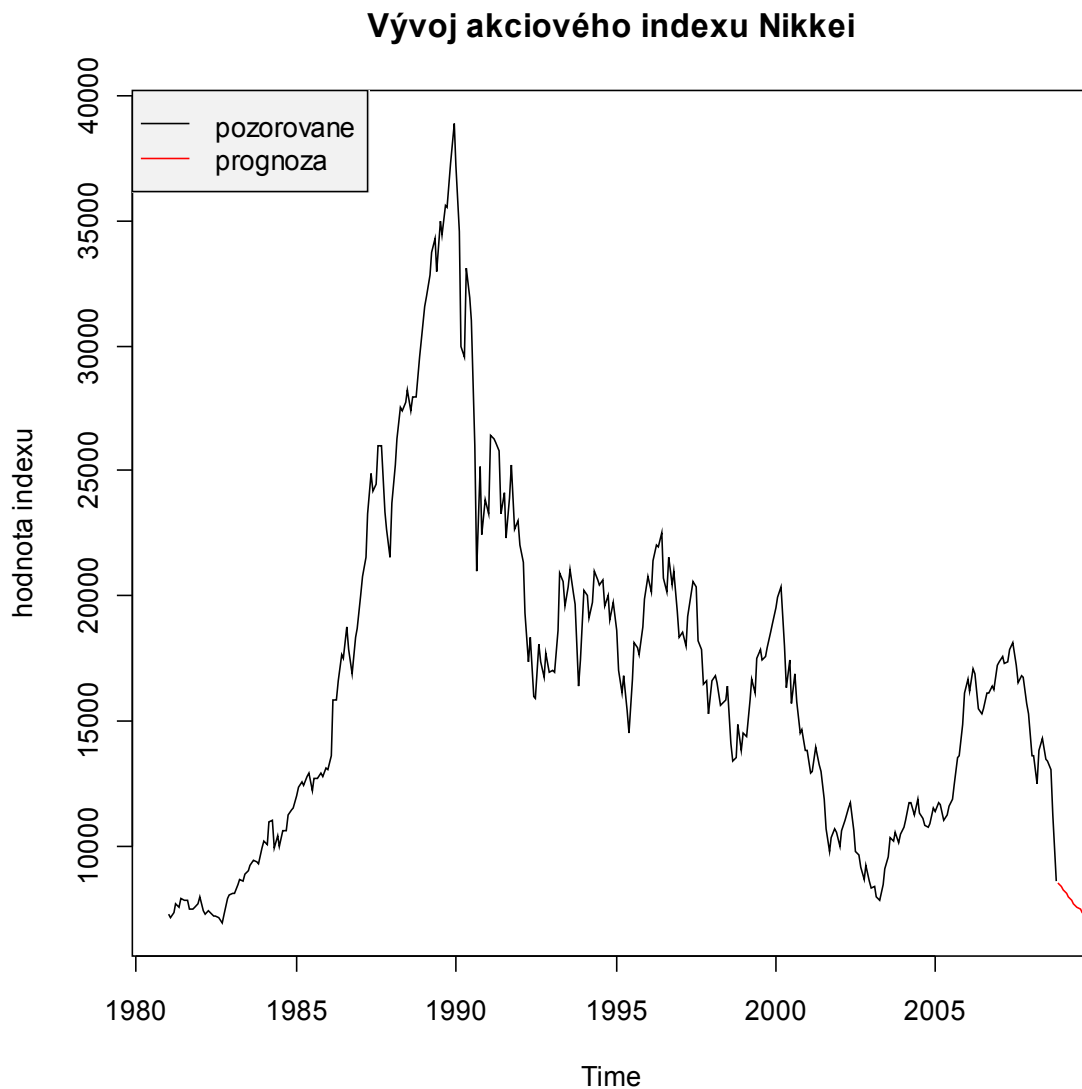
```
> MAPE
```

```
[1] 4.39149
```

Priemerná absolútna chyba je 4,39%. To znamená, že priemerná predpoveď modelu má odchýlku od skutočného priebehu časového radu +/- 4,39%.

4 Porovnanie presnosti vybraných iných modelov

Model zostrojený identicky ako predchádzajúci bez logaritmickkej transformácie.



Mape modelu je 4,342299 %.

Mape prvých 5 predikcií je 3,555794 %.

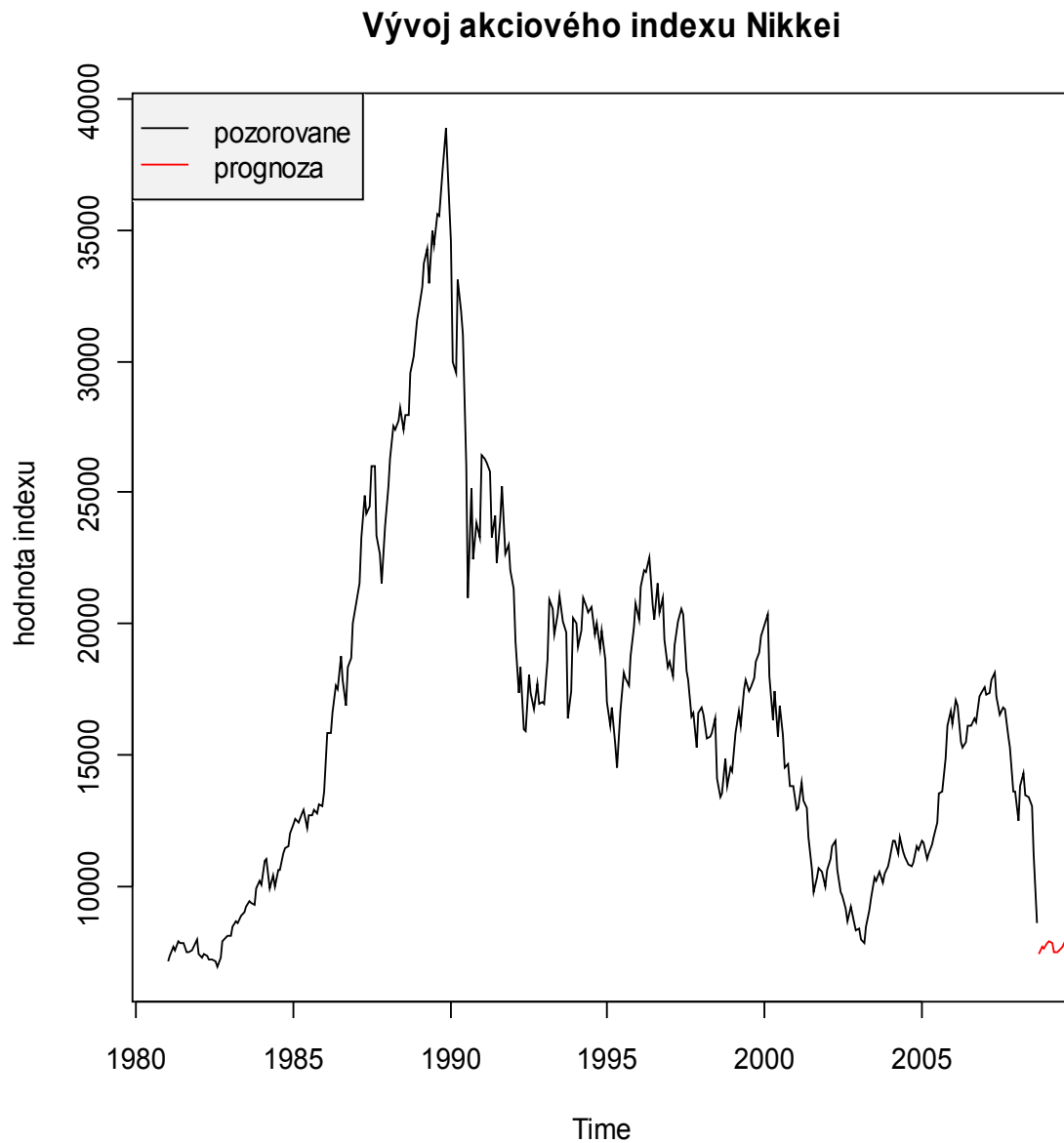
Avšak model ARMA je zostrojený nekorektne, keďže bola zamietnutá nulová hypotéza o normalite, taktiež Portmanteauva štatistika zamietla vhodnosť modelu.

```
Shapiro-Wilk normality test
data: f.res
W = 0.9832, p-value = 0.0006153
> Q # testova statistika
[1] 38.90599
> kh # kriticka hodnota
```


[1] 27.58711

[1] "zamietame H0 (H0:model ARMA je spravne spifikovany)"

Model bez logaritmickej transformácie zostrojený výlučne Fourierovým radom.



Mape modelu je 0,8947706 %.

Mape prvých 5 predikcií je 7,585572 %.

5 Záver

Podarilo sa nám zostaviť vhodný model, ktorý opisuje vývoj akciového indexu NIKKEI 225. Keďže príklad bol riešený s odstupom času, mali sme k dispozícii reálne hodnoty indexu, ktoré sme mohli porovnať s našimi predikciami. Model poskytol kvalitné predikcie minimálne na prvých 5 mesiacov.

Zoznam použitej literatúry

- [1] Anděl, J.: Matematická statistika. SNTL, Praha 1978.
- [2] Anděl, J.: Statistická analýza časových řad. SNTL, Praha 1976.
- [3] Anděl, J.: Statistické metody. MATFYZPRESS, Praha 1998.
- [4] Anděl, J.: Základy matematické statistiky. MATFYZPRESS, Praha 2005.
- [5] Arlt, J. & A.: Ekonomické časové řady. Grada Publishing, Praha 2007.
- [6] Arlt -- Arltová -- Rublíková: Analýza ekonomických časových řad s příklady. Vysoká škola ekonomická, Praha, 2002.
- [7] Arlt, J.: Moderní metody modelování ekonomických časových řad. Grada, Praha, 1999.
- [8] Arlt – Arltová: Finanční časové řady: vlastnosti, metody modelování, příklady a aplikace. Grada, Praha, 2003.
- [9] Baltagi, Badi H.: Econometrics, Springer-Verlag Berlin Heidelberg, 2008.
- [10] Brockwell J. Peter – Davis A. Richard : Time series: Theory and methods. New York : Springer-Verlag, 1991.
- [11] Buša - Ševčovič : Open source systém na spracovanie údajov, Košice 2007, <http://people.tuke.sk/ladislav.sevcovic/esf/Rs.pdf> (30.9.2008).
- [12] Cipra, T.: Analýza časových řad s aplikacemi v ekonomii. SNTL, Praha 1986.
- [13] Cipra, T.: Ekonometrie, SNP, Praha 1986.
- [14] Cook Dianne - Deborah F. Swayne: Graphics for Data Analysis Interactive and Dynamic With R and Ggobi. Springer, NY 2007.
- [15] Forbelská, M.: Stochastické modelování jednorozměrných časových řad. MUNIPRESS, Brno, 2009.
- [16] Garaj – Šujan: Ekonometria. Alfa, Bratislava 1980.
- [17] Greene H. William: ECONOMETRIC ANALYSIS. Upper Saddle River, New Jersey 2002.
- [18] Hamilton James: Time series analysis, New Jersey 2004.
- [19] Chajdiak - Komorník - Komorníková.: Štatistické metódy, Statis Bratislava, 1999.

- [20] Kozák – Hindls – Arlt: Úvod do analýzy ekonomických časových řad. Vysoká škola ekonomická, Praha, 1994.
- [21] Manual R : <http://cran.r-project.org/doc/manuals/fullrefman.pdf> (13.10.2008).
- [22] Shumway Robert - Stoffer David : Time Series Analysis and Its Applications With R Examples. New York: Springer, 2006.
- [23] Schelter Björn - Mathias Winterhalder - Jens Timmer : Handbook of Time Series Analysis. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2006.
- [24] Urbaníková - Vrábelová: Modelovanie ekonomických a finančných procesov. FPV UKF Nitra, 2005.
- [25] Wiley John : Regression Analysis by Example Fourth Edition. Hoboken, New Jersey, 2006.
- [26] Zongwu Cai : Analysis of Time Series Data Using R. UNC Charlotte, 2006.
- [27] Zvára K.: Regresní analýza. Academia, Praha 1986.

Prílohy

Zdrojový kód pre model zostrojený bez logaritmickej transformácie.

```
raw<-
read.csv(file="mN225.csv", skip=4, colClasses=c("character",NULL,"numeric")
)
data<-as.numeric(raw[,3])

nikkei=ts(data[133:466], start=1981,f=12)
plot(nikkei,col="black",ylab="hodnota indexu",lwd=2,main="Vývoj
akciového indexu Nikkei")

d.nikkei=nikkei-mean(nikkei)
P = abs(fft(d.nikkei)/334)^2
hladane= sort(P[2:167],decreasing=T)[1:4]/334

pozicia=c(0,0,0,0) #tolko 0 ,kolko najvyznamnejsich period chceme pouzit
for (j in 1:4){ for (i in 2:167){
if (isTRUE(all.equal(P[i]/334,hladane[j]))){pozicia[j]=i-1}}
period=333/pozicia # 479 kvoli tomu ze lprvok reprezentuje priemer radu
period

furier.coef.s<-function(n,period,rad) {
s=rep(0,length(period))
for (i in 1:length(period)){for (t in 1:n){
s[i]=s[i]+rad[t]*sin(2*pi*t/period[i])}}
furier.coef.s<-s*2/n}

furier.coef.c<-function(n,period,rad) {
c=rep(0,length(period))
for (i in 1:length(period)) {for(t in 1:n){c[i]=c[i]
+rad[t]*cos(2*pi*t/period[i])}}
furier.coef.c<-c*2/n}

sin<-furier.coef.s(334,period,d.nikkei)
cos<-furier.coef.c(334,period,d.nikkei)

furier.rad<-function(rad,s,c,period) {
furier<-rep(0,length(rad))
for (t in 1:length(rad)){ for (i in 1:length(period)){
furier[t]<-furier[t]+c[i]*cos(2*pi*t/period[i])
+s[i]*sin(2*pi*t/period[i])}}
furier.rad<-furier}

f.nikkei<-furier.rad(d.nikkei, sin, cos, period)
plot(d.nikkei[1:334],type="o")
lines(f.nikkei, col="red")

f.res<-d.nikkei-f.nikkei

hist(f.res)
qqnorm(f.res)
qqline(f.res)
acf(f.res)
shapiro.test(f.res)
```

```

arima<-auto.arima(f.res,max.P=0,max.Q=0,stationary=F, trace=T, ic="aic")
a.res<-arima$res

K = trunc(sqrt(length(a.res)))
n=length(a.res)
ac = as.vector(acf(a.res)$acf) # odhadute autokorelacie
Q = n*sum(ac[1:K+1]^2) # posun o 1 kvoli r(0)
Q # testova statistika
kh = qchisq(0.95,K-arima$arma[1]-arima$arma[2]) # kriticka hodnota
kh
c("nezamietame H0 (H0:model ARMA je spravne spefikovany)","zamietame H0
(H0:model ARMA je spravne spefikovany)")[(Q kh)+1]

vyr.rad<-d.nikkei-a.res
print(mape(nikkei,mean(nikkei)+vyr.rad))
plot(d.nikkei[1:334],type="o")
lines(vyr.rad[1:334],col="red")
lines(f.nikkei[1:334],col="green")

furier.rad.predpoved<-function(rad,s,c,period,T){
furier<-rep(0,length(rad)+T)
for(t in 1:length(rad)+T){ for(i in 1:length(period)){
furier[t]<-furier[t]+c[i]*cos(2*pi*t/period[i])
+s[i]*sin(2*pi*t/period[i])}}
furier.rad.predpoved<-furier[(length(rad)+1):(length(rad)+T)]}

predpoved<-
furier.rad.predpoved(d.nikkei,sin,cos,period,12)+predict(arima,12)$pred
plot(d.nikkei)
lines(predpoved,col="red")

demean.predpoved<-mean(nikkei)+predpoved
plot(nikkei,ylab="hodnota indexu",main="Vývoj akciového indexu Nikkei")
lines(demean.predpoved,col="red")
legend("topleft", c("pozorovane","prognoza"), lty=c("solid","solid"),
pch=c(NA,NA), col=c("black","red"), bg="gray95")

realne<-c(8512.27,8859.56,7994.05,7568.42,8109.5)
print(mape(realne,demean.predpoved[1:length(realne)]))

```

Zdrojový kód pre model bez logaritmickej transformácie zostrojený výlučne Fourierovým radom.

```

raw<-
read.csv(file="mN225.csv",skip=4,colClasses=c("character",NULL,"numeric")
)
data<-as.numeric(raw[,3])

nikkei=ts(data[134:466], start=1981,f=12)
plot(nikkei,col="black",ylab="hodnota indexu",lwd=2,main="vývoj
akciového indexu Nikkei")

d.nikkei=nikkei-mean(nikkei)

furier.coef.s<-function(rad){
n<-length(rad)
s=rep(0,166)
for(i in 1:166){for(t in 1:n){

```

```

s[i]=s[i]+rad[t]*sin(2*pi*t*i/(n-1))}}
furier.coef.s<-s*2/n}

furier.coef.c<-function(rad){
n<-length(rad)
s=rep(0,166)
for (i in 1:166){for (t in 1:n){
s[i]=s[i]+rad[t]*cos(2*pi*t*i/(n-1))}}
furier.coef.s<-s*2/n}

sin<-furier.coef.s(d.nikkei)
cos<-furier.coef.c(d.nikkei)

furier.rad<-function(rad,s,c){
n<-length(rad)
furier<-rep(0,length(rad))
for (t in 1:length(rad)){ for (i in 1:length(s)){
furier[t]<-furier[t]+c[i]*cos(2*pi*t*i/(n-1))+s[i]*sin(2*pi*t*i/(n-1))}}
furier.rad<-furier}

f.nikkei<-furier.rad(d.nikkei, sin, cos)

furier.rad.predpoved<-function(rad,s,c,T){
n<-length(rad)
furier<-rep(0,length(rad)+T)
for (t in 1:length(rad)+T){ for (i in 1:166){
furier[t]<-furier[t]+c[i]*cos(2*pi*t*i/(n-1))+s[i]*sin(2*pi*t*i/(n-1))}}
furier.rad.predpoved<-furier[(length(rad)+1):(length(rad)+T)]}

predpoved<-furier.rad.predpoved(d.nikkei, sin, cos, 12)

f=ts(mean(nikkei)+f.nikkei, start=1981, f=12)
demean.predpoved<-ts(mean(nikkei)+predpoved, start=c(2008,10), f=12)

plot(nikkei, ylab="hodnota indexu", main="Vývoj akciového indexu Nikkei")
lines(demean.predpoved, col="red")
legend("topleft", c("pozorovane", "prognoza"), lty=c("solid", "solid"),
pch=c(NA, NA), col=c("black", "red"), bg="gray95")
print(mape(nikkei, f))

realne<-c(8512.27, 8859.56, 7994.05, 7568.42, 8109.5)
print(mape(realne, demean.predpoved[1:length(realne)]))

```