



Študentská vedecká aktivita

Rekurentné zobrazovanie a klastrová analýza ako alternatívy pri analýze finančných časových radov

Bc. Peter Tomko

Sekcia: Kvantitatívne metódy a informatika

Konzultant: Doc. RNDr. Denis Horváth, CSc.

Banská Bystrica
rok 2013

Abstrakt práce

1. Téma práce: Rekurentné zobrazovanie a klastrová analýza ako alternatívy pri analýze finančných časových radov.

2. Meno autora: Bc. Peter Tomko

3. Ročník a odbor štúdia: 1. ročník 2.stupeň; Financie, bankovníctvo a investovanie

4. Konzultant/vedúci práce: Doc. RNDr. Denis Horváth, CSc.

5. Obsah:

a) cieľ riešenia:

Lineárne modely vo finančných časových radoch nedostatočne modelujú priebehy, ktoré sú zložitejšieho charakteru. V práci sme predstavili metódu rekurentného zobrazovania, ktorá je nelineárneho charakteru. Cieľom pomocou tejto metódy je identifikácia nelinearity a určitých vzorov v časových radoch. V práci sme taktiež použili klastrovú analýzu, keďže vzory budúcnosti sú determinované minulosťou. Pomocou Markovovho reťazca sme ukázali, že je možné

b) obsah riešenia:

Práca podáva alternatívny prístup pri analyzovaní časových radov. Tento nový prístup je založený na zobrazovaní dynamiky časových radov. Metóda rekurentného zobrazovania podáva istý obraz o tom, aké systémy sa v danom časovom rade nachádzajú. Rekurentne kvantifikujúca analýza nadväzuje na rekurentné zobrazovanie. RQA kvantifikuje štruktúru získaných grafov. Ukazovatele RQA použité v práci určujú komplexnosť systému.

c) dosiahnuté výsledky

Klasické modely Box-Jenkinsa nepodali v práci dostatočne významné výsledky. V práci sme taktiež poukázali na to, že v rámci rekurentne kvantifikujúcej analýzy sa ukazovatele DET a LAM výrazne znížili pred krízovým obdobím. Zhoršenie predikcie sme mohli vidieť aj pomocou Theilovej štatistiky. AICc ukázalo, že v týchto časových oknách nadobúdalo vysoké hodnoty.

6. Kľúčové slová: rekurentné zobrazovanie, klastrová analýza, ARMA, VAR

1 Úvod

V práci sú rozoberané alternatívne prístupy pri analyzovaní časových radov. Medzi základné metódy analýzy časových radov patrí metodológia Box-Jenkinsa, dekompozícia časových radov a iné. Ako hlavná problematika časových radov je nelinearita. Tá spôsobuje problém, ktorý spočíva v tom, že klasické lineárne modely nedávajú štatisticky významné výsledky. V kontexte týchto poznatkov sa v práci budeme venovať nelineárnym prístupom, ktoré sa snažia zobrazit nelinearitu. Ide predovšetkým o rekurentné zobrazovanie. Klastrová analýza bola použitá ako istý doplnok k spomenutým metódam. Zodpovedanie otázky čo je nelinearita môže byť veľmi komplikované. V ekonometrickej praxi existuje nelineárna metóda najmenších štvorcov alebo taktiež modely nelineárne v parametroch, či nelineárne modely volatility (EGARCH, IEGARCH, GJR-GARCH a pod.). Rekurentné zobrazovanie prezentované v práci je metóda, ktorá zobrazuje dynamiku systému. Nie je to klasická metóda založená na odhadovaní parametrov, ale ide o metódu, ktorá spracováva údaje do grafu. V grafoch sa následne dajú identifikovať obrazce, ktoré sú v niektorých prípadoch pravidelného charakteru. Skúmaním rekurentných grafov sa dostávame k rekurentne kvantifikujúcej analýze. Súčasťou RQA sú ukazovatele, ktoré kvantifikujú štruktúru rekurentného obrazca. Nelinearitu budeme chápať ako neschopnosť použitia klasických lineárnych metód metodológie Boxa-Jenkinsa. Nemožnosť použitia týchto metód bude v kontexte rekurentne kvantifikujúcej analýzy podporená tým, že v časovom rade sa ukazovatele determinizmu a laminarity zhoršia. Následne by tiež kritérium $AICc$ malo rásť. To budeme sledovať na časovom okne akciových indexov. Tým, že LAM poklesne, znamená to toľko, že klesá predvídateľnosť systému. DET určuje percento rekurentných bodov formujúcich diagonálnu líniu. Pokles tohto ukazovateľa znamená, že predošlý stav systému menej determinuje nasledujúce stavy systému. Poukážeme na to, že tieto ukazovatele pri zhoršení podmienok na trhu, klesajú. Nedávne krízové udalosti na trhu sú len dôkazom toho, že je potrebné hľadať nové prístupy pri analyzovaní časových radov.

2 Úvod do analýzy časových radov

Časové rady sú charakteristické vlastnosťami akými sú trend, sezónnosť, podmienená heteroskedasticita a nelinearita [1], [2]. Pod pojmom časový rad sa rozumie akákoľvek postupnosť dát chronologicky usporiadaná v čase. V našom prípade sa budeme venovať časovým radom finančného charakteru. Nelinearita v časovom rade môže mať za následok to, že sa bude v čase meniť autokorelačná štruktúra časového radu [2]. Na druhej strane problematika nelinearity je veľmi zložitá a zďaleka nie preskúmaná oblasť [2].

2.1 Autokorelačné vlastnosti časových radov

Box-Jenkinsova metodológia pristupuje k analýze časových radov pomocou špeciálnych stochastických modelov založených na predpoklade linearity. Ide napríklad o modely *ARIMA*, *ARMA*, *VAR* a podobne. Postupy, ktoré boli prezentované v monografii Box a Jenkinsa ¹, sú schopné modelovať priebehy, ktoré sú zložitejšieho charakteru a nezávládnuteľné dekompozičným prístupom [1]. Modely Box-Jenkinsa majú špecifické tvary autokorelačných a parciálnych autokorelačných funkcií [2]. Metodológia Box-Jenkinsa bola v knihe [1] analyzovaná pre stacionárne časové rady. Stacionarita je, všeobecne povedané, stochasticky ustálené správanie časového radu. Rozlišujeme slabú ² a striktnú ³ stacionaritu [1].

¹Box a Jenkins Time series analysis, Forecasting and Control

²Slabá stacionarita je stochasticky ustálené správanie strednej hodnoty a kovariancie, z čoho vyplýva invariantnosť rozptylu.

³Striktná stacionarita znamená to, že pravdepodobnostné rozdelenie časového radu je nezávislé od posunov v čase.

2.2 Vybrané modely jednorozmerných a viacrozmerých časových radov

2.2.1 Proces ARMA p, q

Medzi základné procesy Box-Jenkinsovej metodológie sa radí aj proces ARMA(p, q), ktorý je zmiešaným procesom autoregresného procesu AR(p) a procesu kľzavých súčtov MA(Q). Autoregresný proces AR(p) ⁴ možno symbolicky zapísať:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-3} + \dots + \varphi_p y_{t-p} + \varepsilon_t, \quad (2.1)$$

kde $\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_p$ sú parametre autoregresného procesu a ε_t je postupnosť nekorelovaných náhodných veličín s vlastnosťami $E(\varepsilon_t) = 0$ a $\sigma^2 > 0$. Následne vyjadrením rovnice pomocou operátora lagu získame nasledujúci tvar

$$y_t - \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-3} + \dots + \varphi_p y_{t-p} = \varepsilon_t, \quad (2.2)$$

$$(1 - \varphi_1 B + \varphi_2 B^2 + \varphi_3 B^3 + \dots + \varphi_p B^p) y_t = \varepsilon_t, \quad (2.3)$$

$$\varphi(B) y_t = \varepsilon_t. \quad (2.4)$$

Autoregresný proces radu p je stacionárny ⁵, ak korene polynómu $\varphi(B)$ ležia vnútri jednotkového kruhu v komplexnej rovine a teda platí $|\varphi_1|, |\varphi_2|, |\varphi_3|, |\varphi_4|, \dots, |\varphi_p| < 1$. Za tohto predpokladu má autoregresný proces AR(p) strednú hodnotu $E(y_{t-p}, \varepsilon_t = 0) = 0$ a rozptyl $\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - \varphi_1^2 - \varphi_2^2 - \varphi_3^2 - \dots - \varphi_p^2}$. Autokorelačnú funkciu autoregresného procesu ⁶ získame riešením diferenčnej rovnice tvaru [1]

$$y_t = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + \dots + \phi_p \rho_{k-p}. \quad (2.5)$$

⁴ $y_t - \mu = \varphi_1(y_{t-1} - \mu) + \varphi_2(y_{t-2} - \mu) + \varphi_3(y_{t-3} - \mu) + \dots + \varphi_p(y_{t-p} - \mu) + \varepsilon_t$ Platí však, že stredná hodnota má v čase nemenný charakter $E(y) = \mu$.

⁵Za predpokladu stacionarity je autoregresný proces invertovaný lineárny proces useknutý v ráde p . Lineárny proces s nulovou strednou hodnotou môžeme vyjadriť ako $y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \psi_3 \varepsilon_{t-3} + \dots + \psi_j \varepsilon_{t-j}$.

⁶ p v rovnici predstavuje bod useknutia autoregresného procesu AR(p). Ide o bod, za ktorým je autokorelačná funkcia rovná nule. Tento bod sa dá získať z výberovej ACF. Na tieto účely sa používa Bartlettova aproximácia, viď [1], [2].

Parciálna autokorelačná funkcia je pre oneskorenia $k = 1, 2, 3, 4, 5, \dots, p$ rôzna od nuly, avšak za bodom useknutia je nulová. Keďže proces ARMA (p, q) sa skladá z dvoch procesov, je potrebné spomenúť proces kľavých súčtov MA (q) . Formálne ho môžeme vyjadriť [1], [2]

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.6)$$

$$(1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q) y_t = \varepsilon_t, \quad (2.7)$$

$$y_t = \theta(B^q) \varepsilon_t. \quad (2.8)$$

Proces MA (q) je vždy stacionárny a vzniká useknutím lineárneho procesu v bode q . Pre stacionárny proces MA (q) platí, že má nulovú strednú hodnotu ⁷. Rozptyl procesu vieme zapísať ako $\sigma_y^2 = (1 + \theta_1^2 + \theta_2^2 + \theta_3^2 + \dots + \theta_q^2)$. Autokorelačná funkcia procesu MA (q) má tvar ⁸

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \theta_3^2 + \dots + \theta_q^2} & \text{ak } k = 1, 2, 3, 4, \dots, q \\ 0 & \text{ak } k > q \end{cases} \quad (2.9)$$

Proces ARMA (p, q) je teda kombináciou hore uvedených procesov a môžeme ho zapísať ako ⁹ [1], [2]

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}. \quad (2.10)$$

Autokorelačná funkcia takto definovaného procesu splňuje diferenčnú rovnicu

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \varphi_3 \rho_{k-3} + \dots + \varphi_p \rho_{k-p}. \quad (2.11)$$

⁷Taktiež vieme pre nenulovú strednú hodnotu zapísať tento proces ako $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \dots + \theta_q \varepsilon_{t-q}$

⁸Bod q je bodom, za ktorým je autokorelačná funkcia procesu kľavých súčtov rovná 0. Opäť to vieme dokázať Bartlettovou aproximáciou.

⁹Takto vyjadrený tvar predpokladá nulovú strednú hodnotu. Pre strednú hodnotu rôznu od nuly a nezávislú od času prechádza tento tvar do tvaru rovnice: $y_t = \mu + \varphi_1(y_{t-1} - \mu) + \varphi_2(y_{t-2} - \mu) + \dots + \varphi_p(y_{t-p} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$

Autokorelačná funkcia modelu ARMA (p, q) nemá bod useknutia a pozostáva z klesajúcich geometrických postupností a sínusoid rôznych frekvencií s klesajúcimi amplitúdami. Box-Jenkinsova metodológia priniesla istý algoritmus do analýzy časových radov a zosumarizovala dovtedy známe fakty. Analýza časových radov pozostáva z identifikácie modelu, jeho odhadu, diagnostiky a predpovedí ¹⁰ [1].

2.2.2 Základný model viacrozmerných časových radov VAR

Doteraz sme pracovali s jednorozmerným časovým radom, keď sme predstavili model ARMA (p, q) . V ekonometrii sa však často pracuje s viacrozmernými časovými radmi, kde medzi najznámejšie modely patrí vektorová autoregresia alebo inak napísané VAR (p) . Vlastnosti, ktoré sme definovali pre jednorozmerný časový rad dokážeme zovšeobecniť aj pre viacrozmerné časové rady¹¹. Vychádzajme z viacrozmerného časového radu $\mathbf{y}_t$. Slabá stacionarita viacrozmerného časového radu znamená invarianťnosť prvého a druhého momentu viacrozmerného náhodného procesu voči posunom v čase. Rozdiel oproti jednorozmerným radom bude v zápise tohto predpokladu a teda pre ľubovoľné h platí [1], [2]

$$E(\mathbf{y}_t) = \boldsymbol{\mu} = konst., \quad (2.12)$$

$$cov(\mathbf{y}_s, \mathbf{y}_t) = E[(\mathbf{y}_s - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})^T] = cov(\mathbf{y}_{s+h}, \mathbf{y}_{t+h}), \quad (2.13)$$

z čoho vyplýva, že $var(\mathbf{y}_t) = \boldsymbol{\Sigma}yy$. Pre viacrozmerné časové rady sa taktiež uvádza autokorelačná a autokovariančná funkcia¹² v maticovom tvare.

$$\boldsymbol{\rho}_k = \mathbf{D}^{-1/2} \boldsymbol{\Gamma}_k \mathbf{D}^{-1/2}, \quad (2.14)$$

pre $k = \dots, -1, 0, 1, \dots$, pričom $\mathbf{D} = \mathbf{diag} \boldsymbol{\Sigma}yy$. Matica $\mathbf{D}$ predstavuje diagonálnu líniu matice $\boldsymbol{\Sigma}yy$.

$$\boldsymbol{\gamma}_k = cov(\mathbf{y}_t, \mathbf{y}_{t-k}) = E(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-k} - \boldsymbol{\mu})^T, \quad (2.15)$$

¹⁰v práci budeme aplikovať ARMA (p, q) model na časové rady indexov.

¹¹Viacrozmerné časové rady – realizácia viacrozmerného náhodného procesu

¹²Samozrejme v skutočnosti pracujeme s odhadmi autokorelačných a autokovariančných funkcií.

Pri zápise sa to zmení o odhady matice strednej hodnoty a rozptylovej matice.

pre $k = \dots, -1, 0, 1, \dots$

Základným modelom viacrozmerných časových radov je model VAR (p). Model môžeme zapísať v tvare

$$\mathbf{y}_t = \phi_0 + \phi_1 \mathbf{y}_{t-1} + \phi_2 \mathbf{y}_{t-2} + \phi_3 \mathbf{y}_{t-3} + \dots + \phi_p \mathbf{y}_{t-p} + \varepsilon_t, \quad (2.16)$$

kde ε_t je m -rozmerný biely šum¹³, ϕ_0 je intercept a $\phi_1, \phi_2, \phi_3, \dots, \phi_p$ sú matice parametrov VAR(p).. Ako vidíme, model VAR (p) je založený na predpoklade linearity.

2.3 Komplexnosť systému a nelinearita

Lineárne metódy interpretujú pravidelnú štruktúru v dátach pomocou lineárnej korelácie. Paradigma linearity spočíva v tom, že malé zmeny vedú k malým konečným efektom. Ako sme uviedli, finančný trh môžeme považovať za komplexný systém, v ktorom sa interakcie medzi agentmi chápu nelineárne [6]. Štúdie realizované na rozsiahlych finančných dátach skúmali komplexnosť v dynamike cien finančných nástrojov. Ako je napísané v článku [6], výsledkom týchto štúdií je definovanie komplexity na troch úrovniach. Prvá úroveň komplexity je časový rad vo vzťahu ku svojej autokorelačnej funkcii [6]. Štúdie ukázali, že či už ide o vysoko aktívny trh akým je NYSE, alebo emerging trh akým je Budapest stock Exchange, či regionálny trh akým je Milan stock Exchange, autokorelačná funkcia týchto trhov má monotónne klesajúci charakter s veľmi krátkym časom korelovanosti¹⁴. Druhý stupeň komplexity vyjadruje cross-korelácia medzi časovo synchronizovanými časovými radmi finančných nástrojov. Pre časový horizont jedného obchodného dňa je charakteristický vysoký korelačný koeficient [6]. Tretia úroveň komplexity predstavuje spoločné správanie počas extrémnych udalostí na trhu. V štúdii boli realizované analýzy distribučných

¹³Angl. white noise

¹⁴Pri vysoko aktívnych trhoch to nie je dokonca viac ani ako 5 minút.

funkcií výnosnosti finančných nástrojov. Boli zistené dramatické zmeny v distribučných funkciách výnosov počas extrémnych udalostí v porovnaní s obvyčajnými obchodnými dňami [6]. V kontexte komplexity môžeme tieto stupne chápať aj ako isté prekážky voči aplikácií klasických lineárnych modelov na dáta systematického charakteru. V posledných rokoch sme svedkami systematizácie nástrojov nelineárneho charakteru vyúsťujúcich do nelineárnej analýzy časových radov. Alternatívou klasických lineárnych metód je taktiež klastrová analýza, ktorú taktiež v nasledujúcich kapitolách ešte predstavíme. Ak by sme to chceli zhrnúť, tak v tejto časti sme naznačili, že v súčasnosti existuje, podľa nášho názoru, opodstatnenosť inklinácie ku metódam nelineárneho charakteru vyplývajúca zo stále zvyšujúcej sa komplexity reálneho sveta. V kontexte komplexity môžeme tieto stupne chápať aj ako isté prekážky voči aplikácií klasických lineárnych modelov na dáta systematického charakteru. V posledných rokoch sme svedkami systematizácie nástrojov nelineárneho charakteru vyúsťujúcich do nelineárnej analýzy časových radov. Tieto metódy môžeme rozdeliť na metrické, topologické a dynamické [18]. V nasledujúcich kapitolách si z týchto metód predstavíme rekurentné zobrazovanie, ako príklad topologického prístupu¹⁵.

¹⁵Metrické metódy skúmajú vzdialenosti systémových atraktorov, medzi dynamické metódy patrí napríklad Lyapunov atraktor. Topologické metódy skúmajú organizáciu podivných atraktorov a medzi tieto metódy patrí spomínané rekurentné zobrazovanie. Pre viac informácií odporúčame [18].

3 Rekurentné zobrazovanie a rekurentne kvantifikujúca analýza

3.1 Charakteristika rekurentného zobrazovania

Metóda rekurentného zobrazovania [11] bola predstavená v článku Recurrence plots in dynamical systems¹⁶. Ako sme uviedli, metóda rekurentného zobrazovania patrí medzi topologické postupy nelineárneho charakteru. Metóda je skonštruovaná na lokalizáciu skrytých opakujúcich sa vzorov, nestacionarity a štrukturálnych zmien [18], [7]. Inými slovami, rekurentné zobrazovanie odhaľuje všetky časové okamihy, kedy trajektória fázového priestoru¹⁷ navštívi približne rovnaký fázový priestor [8]. Podstatou rekurentného zobrazovania je výpočet $N * N$ – rozmernej matice. Vezmime jednorozmerný časový rad, ktorý je časťou väčšieho n - dimenzionálneho modelu. Ako sa uvádza v [18], na základe Takensovho teorému je možné pomocou jednorozmerného časového radu vytvoriť topologicky ekvivalentný obraz dynamického systému. Zostrojíme vektor¹⁸

$$\mathbf{x}_i^m = (\mathbf{x}_t, \mathbf{x}_{t+d}, \mathbf{x}_{t+2d}, \dots, \mathbf{x}_{t+(m-1)d}), \quad (3.1)$$

kde m je embedding dimenzia a d je časové oneskorenie.

$$\mathbf{R}_{i,j} = \theta(\varepsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (3.2)$$

pre $i, j = 1, 2, 3, 4, \dots, N$ Rekurentné zobrazovanie je matica bodov i, j kde každý bod je označený bodkou ak vzdialenosť¹⁹ medzi oneskorenými vektormi x_i a x_j je

¹⁶Predstavili ju autori – J.P. ECKMANN, S. OLIFFSON KAMPHORST a D. RUELLE

¹⁷fázový priestor (angl. phase space). Body fázového priestoru slúžia vo fyzike na popis stavov sústavy hmotných bodov v každom okamihu. Tieto body vytvárajú trajektóriu (x_i^m) , ktorá keď vytvára určitú štruktúru, hovoríme o atraktoroch. Zobrazenie fázového priestoru podľa [4] mapuje stavy vektoru X_i oproti oneskorenému vektoru $X_{i+delay}$, formálne $X_i \approx X_{i+delay}$.

¹⁸Angl. sa to uvádza ako embedded vector. Ide o trajektóriu fázového priestoru zostrojenú metódou časového oneskorenia – viď [18].

¹⁹ $\|\cdot\|$ je Euklidova vzdialenosť dvoch bodov, ε_i je preddefinovaná vzdialenosť, θ je Heavisideova funkcia alebo inak povedané jednotkový skok definovaná zápisom $\theta(x) = 0$ if $x > 0$, $\theta(x) = 1$

menej ako preddefinovaná vzdialenosť [12]. Parametre rekurentného zobrazovania – embedding dimenzia m a oneskorenie d , sa v kontexte použitia tejto metódy musia adekvátne odhadnúť²⁰.

3.2 Komplexnosť v kontexte RQA

Nadväzne na rekurentné zobrazovanie boli predstavené ukazovatele, ktoré špecifikujú štruktúru získanú v rekurentných grafoch [10]. Tieto ukazovatele sú založené na hustote rekurentných bodoch a na diagonálnej respektíve vertikálnej línii rekurentného obrazca. Autori článkov [9], [13] uvádzajú, že tieto ukazovatele sú ukazovatele komplexity. Hodnoty rekurentných mier sú schopné lokalizovať skryté dynamické štruktúry v rekurentných grafoch. Majme teda bod x_i vektora, ku ktorému prislúcha množina rekurentných bodov [13]

$$\mathbf{S}_i := \{\mathbf{x}_k : R_{i,k} = 1, k \in 1, 2, 3, 4, \dots, N-1\}. \quad (3.3)$$

Podmnožinou množiny $\mathbf{S}_i$ je s_i

$$s_i := \{x_l \in \mathbf{S}_i : (R_{i,l}R_{i,l+1} + R_{i,l}R_{i,l-1}) > 0; l \in \langle 1, N \rangle; R_{i,0} = R_{i,N+1} = 0\}, \quad (3.4)$$

a definujeme ju ako vertikálnu líniu²¹ v rekurentnom obraze v stĺpci i . Je to dôležité definovať kvôli ukazovateľu LAM ²², ktorý môžeme definovať ako

$$LAM = \frac{\sum_{v=v_{min}}^N P(v)}{\sum_{v=v_1}^N P(v)}. \quad (3.5)$$

²⁰*if* $x \geq 0$ [8]

²⁰Mutual information je ekvivalentné k autokorelačnej funkcii, avšak na rozdiel od ACF , mutual information zisťuje aj nelineárne korelácie.

²¹vertikálna línia je istý časový interval, v ktorom sa stav systému nemení alebo mení veľmi pomaly. V rekurentnom obraze sa môže vytvoriť diagonálna línia, ktorá sa objaví ak segmenty trajektórie (body vektora x_i^m) narazia paralelne do iného segmentu – trajektória prechádza tým istým fázovým priestorom ale v rozličnom čase. Pre viac informácií odporúčame [8].

²²Z angl. laminarity.

Vo vzorci môžeme v definovať ako dĺžku všetkých spojených podmnožín $\mathbf{s}_i$. $P(v)$ je histogram vertikálnej línie na celom rekurentnom obraze [13]. LAM určuje percento rekurentných bodov formujúcich vertikálnu líniu rekurentného grafu [9]. V článku [14] používajú autori tento ukazovateľ v súvislosti s trhovými pádmi a krízami. Poukazujú na to, že pred takýmito situáciami nastane v ukazovateli LAM ²³ pokles. Na zmenu ukazovateľov rekurentnej analýzy v dôsledku exogénnych či endogénnych príčin poukazujú taktiež v článku [8]. V praxi sa požívajú aj iné ukazovatele ako napríklad percento rekurencie. Čím je percento rekurencie vyššie, tým viac je dynamika periodickejšia [9]. Vzťah vyzerá nasledovne

$$RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}(\varepsilon), \quad (3.6)$$

kde N je dĺžka časového radu. Ďalším ukazovateľom je determinizmus označovaný ako DET. Tento ukazovateľ určuje percento rekurentných bodov, ktoré patria diagonálnej línii aspoň o dĺžke l_{min} . Vzťah na výpočet tohto ukazovateľa je

$$DET = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=1}^N lP(l)}, \quad (3.7)$$

kde $P(l)$ je histogram diagonálnej línie dĺžky l v rekurentnom zobrazení. Táto miera určuje predvídateľnosť daného systému [9]. Rekurentne kvantifikujúca analýza bola úspešne aplikovaná v mnohých vedeckých oblastiach. V poslednom období sa čoraz viac dostáva do povedomia v oblasti ekonómie ako metóda na báze nelinearity. V tejto časti sme si ukázali niečo z nelineárnych metód na jednorozmernom časovom rade. V nasledujúcej časti si ukážme situáciu, ak budeme analyzovať dvojicu časových radov.

3.3 Cross rekurentné zobrazovanie

Pri rekurentnom zobrazovaní sme uvažovali len jednu trajektóriu $\mathbf{x}_t^m$ o dĺžke N_x . Cross rekurentné zobrazovanie²⁴ už uvažuje s pridaním ďalšej trajektórie $\mathbf{y}_t^m$ o dĺžke

²³Klasicky sa uvažuje s $v_{min} = 2$.

²⁴Angl. CRP – cross recurrence plots.

N_y . CRP symbolicky môžeme zapísať ako

$$\mathbf{CRP}_{i,j} = \theta(\varepsilon_i - \|\mathbf{x}_i - \mathbf{y}_j\|), \quad (3.8)$$

pre $i = 1, 2, 3, 4, \dots, N_x$ a pre $j = 1, 2, 3, 4, \dots, N_y$, pričom $\mathbf{x}_i, \mathbf{y}_j \in \mathbf{R}^m$. Na pozícii v matici $N_x \times N_y$ sa zobrazí rekurentný bod, pokiaľ je bod v druhej trajektórii v čase j blízko stavu v prvej trajektórii v čase i . Cross rekurentné obrazce zisťujú časovo závislé správanie procesov zaznamenaných vo dvoch jednorozmerných časových radoch. CRP taktiež umožňujú štúdium synchronizácie časových radov [15], [16]. Čo sa týka porovnania s rekurentne kvantifikujúcou analýzou, pri CRP dochádza k menším zmenám²⁵.

²⁵Zmeny sa týkajú ukazovateľov v súvislosti s pridaním novej trajektórie. spomenieme napr. RR (miera rekurencie) – $RR = \frac{1}{N-1} \sum_{j-1=k} CR_{i,j}$

4 Vybrané metódy klastrovej analýzy

Podstatou zhlukovania je rozdelenie dát do skupiny podobných objektov, pričom objekty v rámci skupiny majú podobné vlastnosti zatiaľ čo objekty dvoch odlišných skupín majú rozdielne vlastnosti [17]. Skupinu v tomto zmysle voláme zhluk. Zatriedenie veľkého počtu dát do niekoľkých skupín spôsobí stratu niektorých detailov, ale na druhej strane získame zjednodušenie. Zhuková analýza je nástroj na objavenie asociácií a štruktúry v dátach rôzneho typu.

4.1 K-means zhlukovanie

Táto metóda patrí medzi najpopulárnejší nástroj zhukovej analýzy. Meno pochádza z toho, že každý zhluk je reprezentovaný stredom jeho bodov. Takýto stred nazývame centroid. Je zrejmé, že pri použití tejto metódy je nevyhnutné určiť počet takýchto centroidov a priori, do ktorých chceme roztriediť dáta [17]. Hlavná myšlienka k-means zhlukovania je definovanie k počet centroidov pre každý zhluk. V konečnom dôsledku k-means metóda minimalizuje cieľovú funkciu v tvare

$$\mathbf{J} = \sum_{j=1}^k \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{c}_j\|. \quad (4.1)$$

Výraz $\|\dots\|$ predstavuje vzdialenosť medzi dátovým bodom x_i a centroidom c_j .

4.2 Fuzzy k-means zhlukovanie

Fuzzy k-means klastrovanie sa trochu od predošlého líši. Pri tejto metóde určujeme príslušnosť, s akou objekt patrí danému zhuk. Je teda zrejmé, že objekt môže patriť viacerým zhukom s rozdielnym participačným koeficientom.

$$J = \sum_{j=1}^k \sum_{i=1}^n \mathbf{u}_{ij}^m \|\mathbf{x}_i - \mathbf{c}_j\|^2; 1 \leq m < \infty. \quad (4.2)$$

V tomto prípade je m reálne číslo väčšie ako 1, u_{ij} predstavuje mieru príslušnosti k danému zhluku a c_j je n - rozmerný stred zhluku. Túto príslušnosť danému zhluku (participáciu) a n -rozmerný stred zhluku vieme zdefinovať nasledujúcim spôsobom

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|\mathbf{x}_i - \mathbf{c}_j\|}{\|\mathbf{x}_i - \mathbf{c}_k\|} \right)^{\frac{2}{m-1}}}, \quad (4.3)$$

kde $c_j = \frac{\sum_{i=1}^N \mathbf{u}_{ij}^m \vec{\mathbf{x}}_i}{\sum_{i=1}^N \mathbf{u}_{ij}^m}$. Klastrová analýza predstavuje alternatívu ku klasickým metódami a budeme ju používať na finančnom rade trhového indexu akým je DJIA.

5 Analýza a použité dáta

V práci budeme pracovať s dennými uzatváracími cenami akciových indexov – DJIA, PX50, NIKKEI, NASDAQ, SP500 a FTSE100²⁶. Uzatváracie ceny v jednotlivých obchodných dňoch sme transformovali na percentuálne logaritmické výnosnosti

$$r_t = 100(\ln P_t - \ln P_{t-1}), \quad (5.1)$$

kde P_t a P_{t-1} sú uzatváracie ceny indexov v čase t respektíve v čase $t-1$. A priori sme zadefinovali maximálny počet autoregresných parametrov p a parametrov kľzavého súčtu q .

$$p, q \in \langle 1, 5 \rangle. \quad (5.2)$$

Na jednotlivé indexy sme identifikovali ARMA a odhadnuté parametre tohto procesu sme určili minimalizáciou Akaikeho korigovaného kritéria²⁷ [1].

$$(\hat{p}, \hat{q}) = \min AICc(p, q), \quad (5.3)$$

kde $\hat{p}$, $\hat{q}$ sú odhadnuté parametre procesu ARMA a $AICc(p, q)$ je Akaikeho informačné kritérium pre zvolené p a q v tvare:

$$AICc(p, q) = AIC + \frac{2k(k+1)}{n-k-1}, \quad (5.4)$$

kde $k = p + q$, n je počet pozorovaní a AIC je Akaikeho informačné kritérium v tvare:

$$AIC = 2k - 2\ln(ARMA^L(p, q)), \quad (5.5)$$

kde $\ln(ARMA^L(p, q))$ je logaritmus hodnoty maximálnej vierohodnosti daného modelu. Takto sme získali modely na celom časovom rade jednotlivých indexov. Následne sme 3800 pozorovaní sme rozčlenili na časové okná o dĺžke 200. Na takto

²⁶Problém, s ktorým sme sa pri uzatváracích cenách stretli, spočíval v nesúlade medzi obchodnými dňami v USA, Európe a Japonsku. Časové obdobie, ktoré sme brali do úvahy začína 8.1.1996. Po zosúladení pozorovaní v jednotlivých indexoch sme takto získali 3800 pozorovaní, ktoré sa končili dátumom 27.9.2012.

²⁷Určuje sa pre konečný počet pozorovaní.

určenom okne sme pre odhadnuté ARMA modely určili *AICc* kritérium na celkovo 180 pozorovaní. Na zvyšných pozorovaniach sme určili ex-post predikciu. Predikciu sme hodnotili pomocou Theilovej štatistiky pre jednotlivé časové okná [1]:

$$THEIL = \frac{\sqrt{\sum_{t=n+1}^{n+h} (r_t - \hat{r}_t)^2}}{\sqrt{\sum_{t=n+1}^{n+h} r_t^2} + \sqrt{\sum_{t=n+1}^{n+h} \hat{r}_t^2}}. \quad (5.6)$$

kde $\hat{r}_t$ sú predikcie logaritmických výnosností na základe modelu ARMA (p, q) . Ako sa v [3] uvádza, pri dobre určených modeloch nadobúda táto štatistika hodnoty blízko 0 a naopak hodnoty blízke sa 1 vypovedajú o úplnom nesúlade medzi predikciou a skutočnosťou [1]. Pomocou rekurentného zobrazovania transformovaných časových radov sme získali obraz o dynamike systému. Ako je v predošlých kapitolách spomenuté, z rekurentne kvantifikujúcej analýzy sme vybrali niekoľko ukazovateľov – *DET*, *LAM* a *RR*²⁸. Tieto ukazovatele sme kvantifikovali pre každé časové okno. V kapitole rekurentného zobrazovania sme naznačili, že parametre trajektórie je potrebné určiť. Ako súčasť softvéru VRA je metóda FNN (angl. False nearest neighbours), ktorá určuje optimálnu dimeziu m trajektórie. Oneskorenie (lag) určíme aplikáciou vzájomnej informácie (angl. mutual information). Tieto parametre trajektórie sú taktiež dôležité pri rekurentne kvantifikujúcej analýze. Na testovanie významnosti parametrov príslušných modelov na celom časovom rade sme použili *t*-test²⁹. Testovanie modelu ako celku sme realizovali *F*-testom. Určili sme taktiež koeficient determinácie, ktorý určuje koľko percent variability vysvetľovanej premennej sme vysvetlili pomocou vysvetľujúcich premenných [3]. Z alternatívnych prístupov k analýze časových radov taktiež aplikujeme klastrovú analýzu na časový rad. A priori nás priebeh časového radu nebude zaujímať. Na časovom okne o dĺžke 28 definujeme momenty, ktoré budú vstupovať do zhlukovej analýzy. Na celkovom počte 29148 pozorovaní indexu DJIA je teda zrejmé, že celkový počet vytvorených

²⁸DET – determinizmus, LAM – laminarita, RR – percento rekurencie

²⁹Hodnotu príslušného odhadnutého parametra (autoregresného alebo parametra kľzavých súčtov) sme podelili štandardnou odchýlkou odhadnutého parametra. *T*-štatistika je zo Studentovho rozdelenia s $(n - k - 1)$ [3] stupňami voľnosti.

okien bude 1041. Časové okno sa nebude prekrývať. Momenty, ktoré budú vstupovať do klastrovania, definujeme dvoma spôsobmi. V prvom prípade pôjde o klasické štatistické vlastnosti – stredná hodnota, smerodajná odchýlka, medián, šikmosť a špicatosť. Tieto vlastnosti určíme pre jednotlivé časové okno. Keďže každá vlastnosť má iné rozpätie, pred ich použitím musíme tieto údaje normalizovať. Urobíme tak pre každú vlastnosť a dôvodom je to, aby do analýzy vstupovali s rovnakou váhou. Použijeme teda dátovú štandardizáciu a konkrétne v našom prípade budeme dáta štandardizovať nasledujúcim spôsobom:

$$z_t = \frac{x_t - x_{min}}{x_{max} - x_{min}}. \quad (5.7)$$

Pre druhý spôsob definujeme štatistické momenty iným spôsobom. A teda:

$$\mu_q = \left\{ \sum_{i=1}^n \left| x_i - \frac{1}{n} \sum_{j=1}^n x_j \right|^q \right\}^{1/q}, \quad (5.8)$$

pre $q = 1, 2, 3, \dots$. Ako v predošlom prípade, pred vstupom do metriky je potrebné dáta štandardizovať. Pri zhľukovaní budeme pracovať s piatimi momentmi t.j. $q = 1, 2, 3, 4, 5$. Štandardizáciu vykonáme vzhľadom k druhému momentu:

$$\left(\frac{\mu_1}{\mu_2}, \frac{\mu_3}{\mu_2}, \frac{\mu_4}{\mu_2}, \frac{\mu_5}{\mu_2} \right). \quad (5.9)$$

K-means zhľukovaním sme získali priradenie zhľuku k časovému radu. Vznikla určitá postupnosť (napr. 112145133364...). Na túto postupnosť môžeme aplikovať analýzu Markovovho reťazca. V našom prípade sa bavíme o náhodnom procese, ktorý má diskkrétne stavy [1]. Tento proces predchádza medzi stavmi 1 až 10 (podľa toho aký počet zhľukov pre k-means zhľukovanie určíme – označíme ich ako Y_t). Platí tzv. markovovská vlastnosť: $P(Y_{t+1} = j | Y_t = i_t, Y_{t-1} = i_{t-1}, Y_0 = i_0 \dots) = P(Y_{t+1} = j | Y_t = i)$, čo znamená, že správanie Markovho reťazca v budúcnosti $t + 1$ nezávisí od minulosti $t - 1$, ale len od prítomnosti (čas označíme ako: t ako prítomnosť, budúcnosť ako $t + 1$, minulosť $t - 1$ a stavy systému ako Y_t). Vo vzorci figuruje na ľavej strane pravdepodobnosť prechodu zo stavu j v čase t do stavu i v čase $t + 1$.

$$p_{ij} = P(Y_{t+1} = j | Y_t = i) \quad (5.10)$$

Túto pravdepodobnosť je možné odhadnúť: $\hat{p}_{ij} = \frac{n_{ij}}{\sum_{k=-\infty}^{\infty} n_{ik}}$, kde n_{ij} je počet všetkých prechodov zo stavu i do stavu j , ktorý podelíme sumou všetkých pozorovaných prechodov zo stavu i [1].

6 Záver

Analýzou časových radov pomocou Box-Jenkinovej metodológie sme dospeli k záveru, že odhadnuté modely ARMA nie sú veľmi dobrým nástrojom. Z pohľadu významnosti parametrov usudzujeme, že na hladine významnosti 0,05 sa v každom modeli nachádzajú práve tri parametre, ktoré sú štatisticky významné. Celkový model je na základe p-hodnoty získanej z F-štatistiky nevýznamný. Tento záver sme získali zo všetkých indexov. Koeficient determinácia bol na veľmi nízkej úrovni. Celkovo nevysvetlil ani 2 percentá variability vysvetľujúcej premennej. Rekurentným zobrazením časových radov sme zistili to, že v prípade krízovej situácie sa na vertikále zobrazí línia, ktorá je farebne zvýraznená. Ako vidíme v prílohe č. 2, v poradí pätnáste časové okno súvisí s veľmi významnou udalosťou na trhu. Na časových radoch všetkých indexoch v časovom okne č.15 vidíme ako logaritmické výnosnosti prechádzajú do vyššej volatility. Ide o obdobie medzi májom 2008 a koncom marca 2009. Je ešte zaujímavé poukázať na index NASDAQ. Zvýšeniu volatility zodpovedá aj časové okno č. 5,6 a 7, ktoré zodpovedajú obdobiu konca roka 1999 až 2001. V tomto období sa prejavila DOT.COM kríza internetových spoločností.

Príloha č.1: Identifikácia ARMA procesov

Identifikované ARMA procesy na vybraných trhovách indexov na základe minimalizácie AICc:

Trhový index	Odhadnutý ARMA (p,q)	AICc
DJIA	(5,2)	12567,9
FTSE100	(5,1)	12797,5
PX50	(2,5)	14254,8
NIKKEI 225	(4,2)	14386,2
NASDAQ	(5,4)	15139,3

Odhadnuté koeficienty ARMA procesov:

	DJIA	FTSE	NIKKEI	NASDAQ	PX
Koeficienty AR (p)					
ar1	1,001022	-0,21722	0,503363	0,008738	-0,6882
ar2	-0,9181	-0,08438	-1,01065	-0,46195	-0,85814
ar3	-0,05219	-0,05636	-0,00598	0,011283	x
ar4	0,002362	0,008814	-0,06482	-0,93731	x
ar5	-0,04095	-0,09362	x	-0,03266	x
Koeficienty MA(q)					
ma1	-1,06069	0,196002	-0,54728	-0,01588	0,720611
ma2	0,960372	x	0,982257	0,450786	0,860622
ma3	x	x	x	-0,03084	-0,00122
ma4	x	x	x	0,935139	-0,02664
ma5	x	x	x	x	-0,04937
intercept	0,022355	0,011814	-0,0217	0,027677	0,018809

T – štatistiky odhadnutých koeficientov modelu (uvedené pomocou p-hodnôt):

	DJIA	FTSE	NIKKEI	NASDAQ	PX
ar1	1	0,072464	1	0,608314	8,07E-06*
ar2	3,63E-10*	0,00052*	2,08E-11*	8,01E-08*	6,82E-08*
ar3	0,046351*	0,009715*	0,375993	0,633243	x
ar4	0,538939	0,689655	0,002156*	4,54E-11*	x
ar5	0,019233	0,000273*	x	0,046936*	x
ma1	2E-12*	0,908047	6,05E-12*	0,280535	0,999994
ma2	1	x	1	1	1
ma3	x	x	x	0,182568	0,480142
ma4	x	x	x	1	0,122625
ma5	x	x	x	x	0,01011*
intercept	0,871976	0,740659	0,192892	0,8254	0,762129

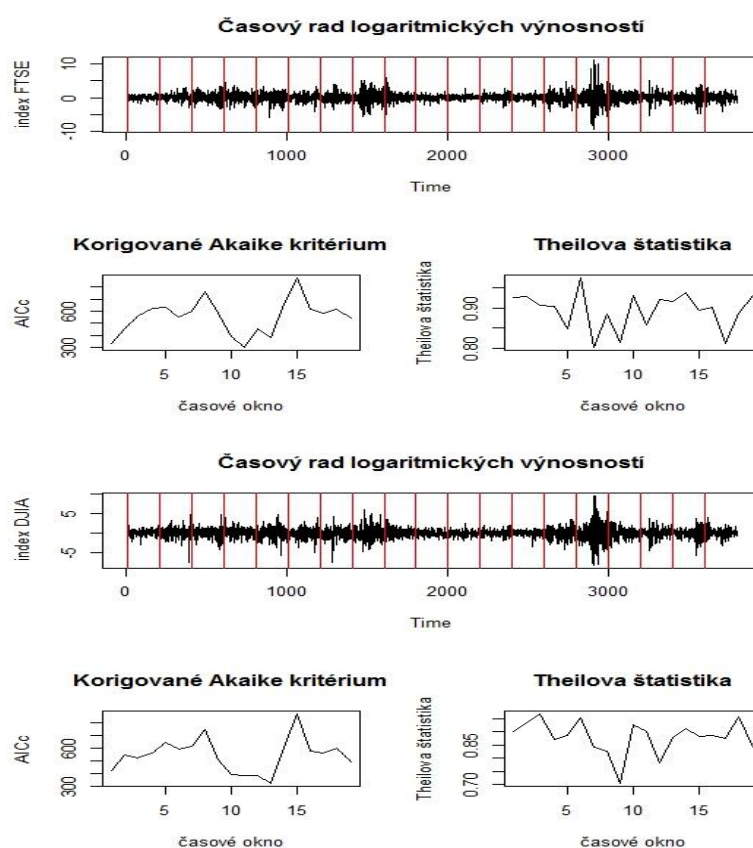
*p<0,05

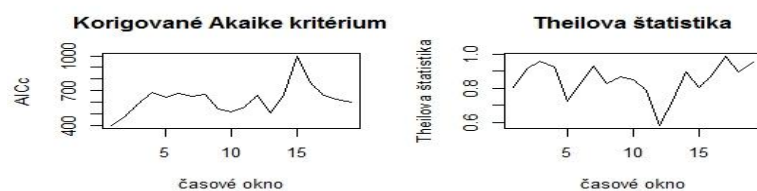
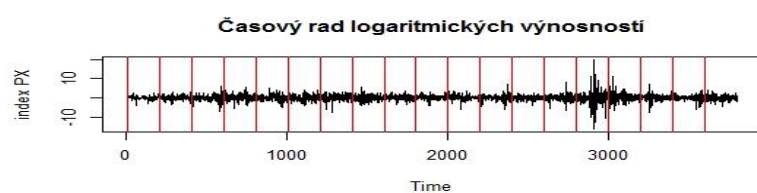
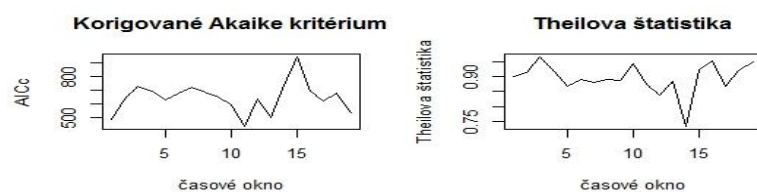
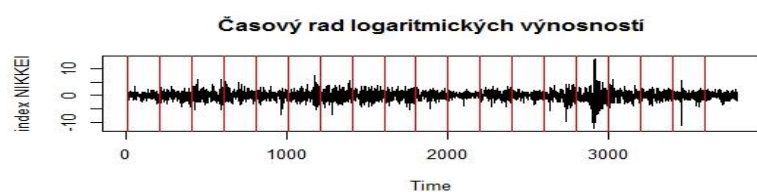
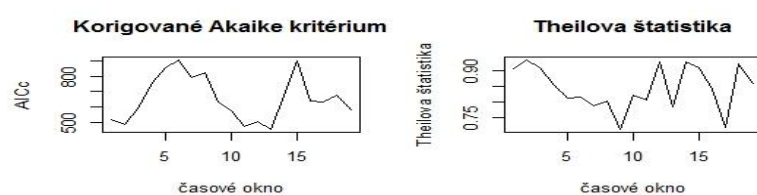
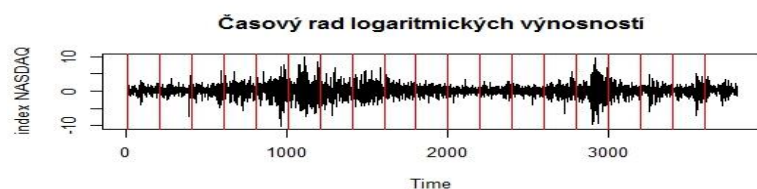
F – štatistiky (ich p-hodnoty) a koeficienty determinácie odhadnutých ARMA procesov:

Index	DJIA	FTSE	NASDAQ	PX	NIKKEI
R^2	0,01	0,017	0,0078	0,0074	0,0082
p-hodnota	0,52	0,95	0,086*	0,33	0,61

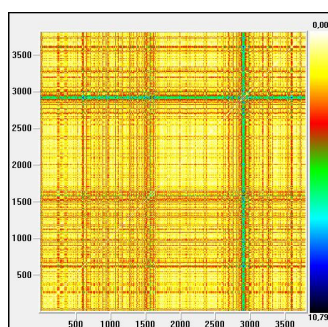
* $p < 0,1$

Príloha č.2: Obrázky – výstupy programu R

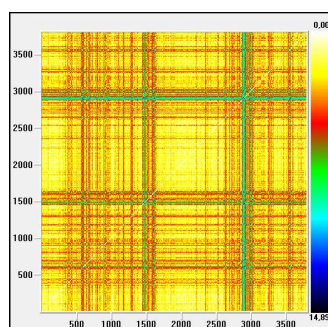




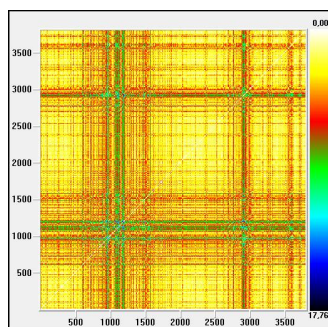
Príloha č.3: Rekurentné zobrazenie vybraných indexov



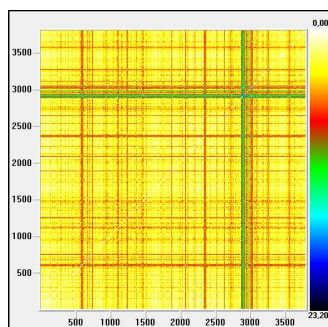
Obrázok: DJIA



Obrázok: FTSE100



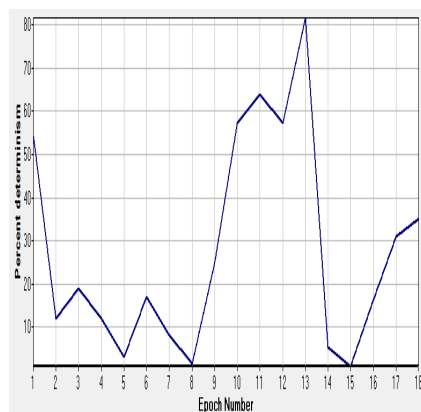
Obrázok: NASDAQ



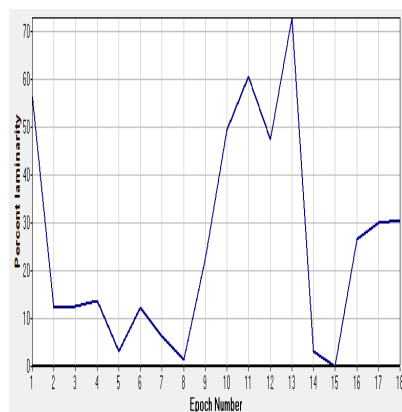
Obrázok: PX

Príloha č.4: Rekurentne kvantifikujúca analýza

DJIA

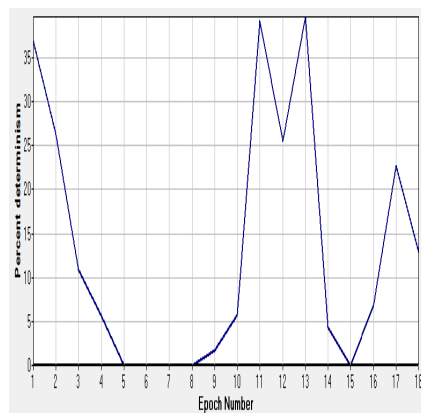


Obrázok: DET

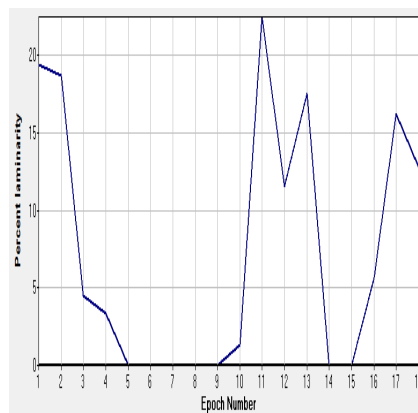


Obrázok: LAM

NASDAQ

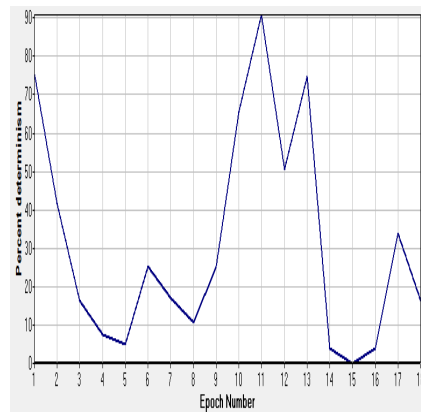


Obrázok: DET

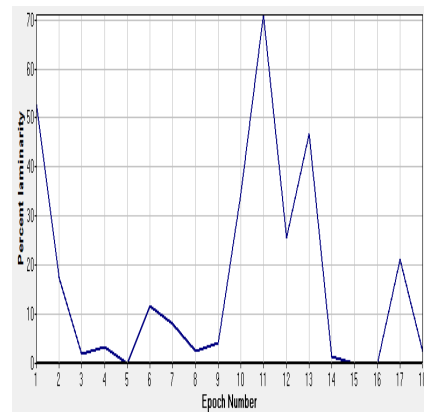


Obrázok: LAM

FTSE100

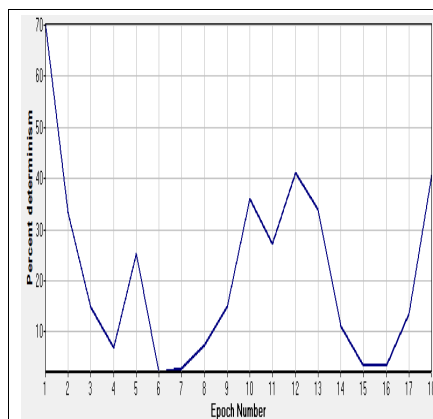


Obrázok: DET

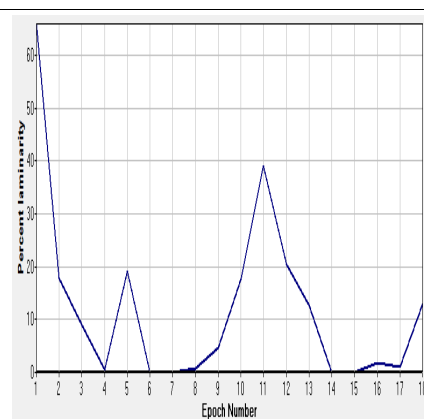


Obrázok: LAM

PX50



Obrázok: DET



Obrázok: LAM

Príloha č.5: Zhluková analýza

Klastrovanie na základe definovaných momentov:

stavA/stavB	1	2	3	4	5
1	0.24644550	0.24644550	0.13744076	0.29857820	0.07582938
2	0.18309859	0.28521127	0.15845070	0.31690141	0.05281690
3	0.19496855	0.30817610	0.12578616	0.27672956	0.09433962
4	0.17532468	0.26623377	0.17857143	0.29220779	0.08766234
5	0.28205128	0.25641026	0.12820513	0.26923077	0.06410256

Klastrovanie pre štatistické momenty

	stavB				
stavA	1	2	3	4	5
1	0.22222222	0.00000000	0.00444444	0.15555556	0.62222222
2	0.00000000	0.97196261	0.01869158	0.00000000	0.00000000
3	0.00000000	0.05769230	0.94230769	0.00000000	0.00000000
4	0.30147058	0.00000000	0.00000000	0.21323529	0.48529411
5	0.25818882	0.00000000	0.00000000	0.13872832	0.60308285

Príloha: Niektoré príkazy v programe R použité v práci

Identifikácia modelu ARMA na základe minimalizácie AICc

```
BEST.ARMA<-function(x){  
  for(p in 0:5){  
    for(q in 0:5){  
      myfit<-arima(x,order=c(p,0,q), method="ML", include.mean=F, optim.control = list(maxit = 1000))  
      AICC<-aicc(myfit$loglik, p, q, length(x))  
      print(c(p,q,AICC))  
    }  
  }  
}
```

AICc kritérium

```
aicc<-function(loglik,p,q,n)  
{ k<-p+q  
  AICC<-2*loglik+2*k+2*k*(k+1)/(n-k-1)  
  return(AICC)  
}
```

AICc kritérium a Theilova štatistika

```
AICc1.FTSE<-numeric()
Theil.FTSE <- numeric()

FTSE <- FTSE1

w<-200

imax<-19

for (i in 1:imax){

  dn<-i*w-199

  up<-dn+199

  dn1<-up-20

  dn2 <- dn1+1

  vektor<-FTSE[dn:dn1]

  arima1.FTSE<-arima(vektor, order=c(5, 0, 1), method=c("ML"), optim.control = list(maxit =
10000000))

  AICc1.FTSE[i]<--2*arima1.FTSE$loglik+2*6+2*6*(6+1)/(180-6-1)

  prediction <- predict(arima1.FTSE,20)

  Theil.FTSE[i] <- ((sum((FTSE[dn2:up]-
prediction$pred)^2))^(1/2))/((sum((prediction$pred)^2))^(1/2)+(sum((FTSE[dn2:up])^2))^(1/2))

}
```

Klastrová analýza

```
vlast1<-numeric(1040)
vlast2<-numeric(1040)
vlast3<-numeric(1040)
vlast4<-numeric(1040)
vlast5<-numeric(1040)
w<-28
imax<-1040
for (i in 0:imax){
  dn<-i*w+1
  up<-dn+w-1
  print(paste("dn=",dn," ", "up=",up))
  vektor<-DJ[dn:up]
  print(vektor)
  vlast1[i]<-mean(vektor)
  vlast2[i]<-sd(vektor)
  vlast3[i]<-median(vektor)
  vlast4[i]<-skewness(vektor)
  vlast5[i]<-kurtosis(vektor)}
vla1stand<-(vlast1-min(vlast1))/(max(vlast1)-min(vlast1))
vla2stand<-(vlast2-min(vlast2))/(max(vlast2)-min(vlast2))
vla3stand<-(vlast3-min(vlast3))/(max(vlast3)-min(vlast3))
vla4stand<-(vlast4-min(vlast4))/(max(vlast4)-min(vlast4))
vla5stand<-(vlast5-min(vlast5))/(max(vlast5)-min(vlast5))
vlastnosti<-cbind(vla1stand, vla2stand, vla3stand, vla4stand, vla5stand)
stavB<-retazec[2:1041]
stavA<-retazec[1:1040]
vektor.spolocny<-cbind(stavA, stavB)
TM<-table(stavA, stavB)
TM.prob<-TM/colSums(TM)
```

Literatúra

- [1] Cipra, T., 2008. Finanční ekonometrie. Praha: EKOPRESS, 2009, ISBN 978-80-86929-43-9
- [2] Arlt, J., Arltová, M., 2009. Ekonomické časové rady. Praha: Professional Publishing, 2009, ISBN 978-80-86946-85-6
- [3] Hatrák, M., 2007. Ekonometria. Bratislava: IURA EDITION, 2007, ISBN 978-80-8078-150-7
- [4] Kantz, H., Schreiber, T., 2004. Nonlinear time series analysis. Cambridge University Press
- [5] Mnategna, R. N., Stanley, H. E., 2000. Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge University Press.
- [6] Bonnano, G., Lillo, F., Mantegna, R. N., 2001. Levels of complexity in financial markets. IN: arXiv.org
- [7] Pecar, B., 2003. The use of visual recurrence analysis and hurst exponents as qualitative tools for analysing financial time series.
- [8] Guhathakurta, K., Bhattacharya, B., Roy, A.CH., 2010. Using recurrence plot analysis to distinguish between endogenous and exogenous stock market crashes IN: Physica A: Statistical Mechanics and its Applications, ISSN: 0378-4371, 2010, vol.389, no. 9, pp.K 1874 - 1882
- [9] Bigdeli, N., Afshar, K., 2009. Characterization of Iran electricity market indices with pay-as-bid payment mechanism IN: Physica A: Statistical Mechanics and its Applications, ISSN: 0378-4371, 2009, vol.388, no. 8, pp.K 1577 - 1592
- [10] Zou, Y., Donner, R. V., Dognes, J. F., Marvan, N., Kurths, J., Identifying complex periodic windows in continuous-time dynamical systems using recurrence-based methods IN: Chaos, ISSN 1054-1500, 2010 , vol. 20, no. 4.

-
- [11] Eckmann, J. P., Kamphorst, S. O., Ruelle, D., Recurrence Plots of Dynamical Systems IN: World scientific series on nonlinear science sereis A, 1995, vol. 16, pp. K 441-446.
 - [12] Fabretti, A., Ausloos, M., Recurrence analysis of the NASDAQ crash of April 2000, IN: arXiv.org, 2005.
 - [13] Marvan, N., Wessel, N., Meyerfeldt, U., Schidewan, A., Kurths, J., Recurrence-plot-based measures of complexity and their application to heart-rate-variability data IN: Phys. Rev. E, 2002, vol. 66.
 - [14] Piskun, O., Piskun, S., Recurrence Quantification Analysis of Financial Market Crashes and Crises, 2011, IN: arXiv.org.
 - [15] Marvan, N., Kurths, J., Cross recurrence plots and their applications. IN: Mathematical Physics Research at the Cutting Edge, 2004.
 - [16] Marvan, N., Thiel, M., Nowaczyk, N. R., Cross recurrence plot based synchronization of time series, 2002.
 - [17] Berkhin, P., A survey of clustering dat mining techniques, 2006, Springer, pp. K 25-71
 - [18] Belaire, F. B., Contreras, D., Recurrence Plots in Nonlinear Time Series Analysis: Free Software, IN: Journal of statistical software, 2002, vol. 07, no. i09.