UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



Modelovanie dynamickej korelácie vo finančných modeloch

DIPLOMOVÁ PRÁCA

2019 Bc. Márius Kostroš

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Modelovanie dynamickej korelácie vo finančných modeloch

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. RNDr. Beáta Stehlíková, PhD.

Bratislava 2019 **Bc. Márius Kostroš**





Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Márius Kostroš
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Modelovanie dynamickej korelácie vo finančných modeloch
Modelling dynamic correlation in financial models

Anotácia: Práca sa zaoberá vybranými prístupmi k modelovaniu dynamickej korelácie vo finančných modeloch:
a) modelovanie dynamickej korelácie v GARCH modeloch
b) meranie vzdialenosti korelačných matíc a následné zhlukovanie časových období
c) dynamická korelácia v konvergenčných modeloch úrokovej miery a jej vplyv na oceňovanie dlhopisov

Vedúci: doc. RNDr. Mgr. Beáta Stehlíková, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 24.01.2018

Dátum schválenia: 25.01.2018 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie V prvom rade sa chcem veľmi pekne poďakovať doc. RNDr. Beate Stehlíkovej, PhD. za množstvo cenných rád a stráveného času pri vyhotovení diplomovej práce. Moja vďaka patrí tiež rodine za podporu a Bohu za odhodlanie počas doterajšieho štúdia.

Abstrakt v štátnom jazyku

KOSTROŠ, Márius: Modelovanie dynamickej korelácie vo finančných modeloch [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky;
školiteľ: doc. RNDr. Beáta Stehlíková, PhD., Bratislava, 2019, 56 s.

Práca sa zaoberá konvergenčným modelom úrokových mier Vašíčkovho typu a rozdielom pri použití konštantnej a dynamickej korelácie na oceňovanie dlhopisov.
Na riešenie parciálnych diferenciálnych rovníc odvodených z modelu je použitý matematický softvér `R` a práca ukazuje postup, ako prostredníctvom neho používať `Wolfram|Alpha` API.

Pre odhad dynamickej korelácie medzi časovými radmi je využívaný DCC-GARCH model, ktorý je následne aplikovaný na odhad korelácie medzi finančnými dátami, ako napríklad akcie, či výmenné kurzy.

Na základe odhadnutej korelácie sú vytvárané zhluky období, v ktorých mala korelácia porovnateľný vývoj. Pre Vyčíslenie podobnosti sa využíva "vzdialenosť korelačných matíc" a takzvaný "DTW algoritmus".

Kľúčové slová: Dynamická podmienená korelácia, DCC-GARCH model,
Konvergenčné modely úrokových mier, Vašíčkov model, Hierarchické zhlukovanie,
DTW algoritmus

Abstract

KOSTROŠ, Márius: Modelling dynamic correlation in financial models [Diploma Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics;
supervisor: doc. RNDr. Beáta Stehlíková, PhD., Bratislava, 2019, 56p.

The diploma deals with Vasicek convergence model of interest rates and refer the difference between constant and dynamic correlation in bond pricing. Mathematical software `R` is used to solve differential equation calculation derived from the model and the procedure how to use `Wolfram|Alpha` API is shown.

Dynamic correlation between time series is estimated with DCC-GARCH model, which is subsequently applicated for correlation estimation in financial data, such as shares or currency rates.

According to the estimated correlation, years are segmented to clusters, where the correlation has a resemblance. The similarity is quantified by the "correlation matrix distance" and "Dynamic time warping".

Keywords: Dynamic conditional correlation, DCC-GARCH model, Convergence models of interest rates, Vasicek model, hierarchical clustering, Dynamic time warping

Obsah

Úvod **9**

1 Teoretický základ **10**

1.1 Časové rady 10

1.1.1 Základné pojmy 10

1.1.2 ARMA proces 12

1.1.3 (G)ARCH model 13

1.1.4 DCC-GARCH model 14

1.2 Oceňovanie dlhopisov pomocou PDR 16

1.2.1 Dlhopis 16

1.2.2 Jednofaktorové modely 16

1.2.3 Dvojfaktorové modely 17

1.3 Hierarchické zhlukovanie 19

2 Výpočet výrazov pomocou R **20**

2.1 WolframAlpha API 20

3 Odhad dynamickej korelácie **23**

3.1 Algoritmus pre odhad DCC 23

3.2 Modelový príklad 24

4 Praktická časť **31**

4.1 Oceňovanie dlhopisov 31

4.2 Rozdiel cien dlhopisu pre rôzne funkcie korelácie 32

4.3 Rozdiel cien pre konkrétne parametre 34

4.4 Odhad DCC pre výmenné kurzy 37

4.4.1 Modelovanie DCC pre viaceré časové rady 41

4.5 Vzdialenosť korelačných matíc 43

4.5.1 Zhlukovanie časových radov na základe CMD 43

4.6 Podobnosť časových radov s rozdielnymi dĺžkami 47

4.6.1 Zhlukovanie časových radov na základe DTW 48

Záver **51**

Zoznam použitej literatúry **52**

A Príloha **55**

Príloha A **55**

A.1 Wolfram Alpha API prostredníctvom R 55

A.1.1 Inicializácia klienta 56

Úvod

Kovergenčným modelom bola v minulosti venovaná nemalá pozornosť. Uplatnenie našli napríklad pri opise vývoja krátkodobých úrokových mier a cien dlhopisov v krajinách Európskej Únie, kedy vývoj úrokových mier štátu vstupujúceho do EÚ konvergoval k EURIBORu. Zaoberať sa budeme konvergenčným modelom Vašíčkovho typu, pričom budeme uvažovať dynamickú koreláciu medzi úrokovými mierami. Taktiež si ukážeme jeden zo spôsobov, ako možno pristupovať k odhadovaniu korelácie medzi dvoma, či viacerými časovými radmi.

V niektorých prácach bola v konvergenčných modeloch korelácia zanedbaná pre možnosť vyjadrenia explicitného riešenia parciálnej diferenciálnej rovnice. Ďalšie práce predpokladali dynamickú koreláciu. Neskôr bola venovaná pozornosť nekonštantnej korelácii, ktorá bola zastúpená nejakou "peknou" funkciou. Takáto funkcia bola vyjadrená tak, aby opisovala predpokladaný vývoj korelácie.

Prvá kapitola obsahuje teoretický základ, ktorý využijeme počas práce. Zhrnieme základné poznatky z modelovania časových radov, ako aj rozšírenie GARCH modelu, takzvaný *DCC-GARCH model*, ktorý použijeme pre odhad dynamickej korelácie. Taktiež spomenieme základné vzťahy pre oceňovanie dlhopisov a *hierarchické zhlukovanie*, ktoré využijeme pri práci s reálnymi dátami.

To, čo niekedy nedokázali počítače, dnes už dokážu mobily, preto sa v druhej kapitole budeme snažiť využiť dostupný softvér a čo najviac zjednodušiť a urýchliť proces riešenia parciálnych diferenciálnych rovníc (ďalej len PDR) daných konvergenčným modelom, či časovo nenáročne nájsť numerické riešenie rovníc, kde nie je možné explicitné riešenie. Zameriame sa na matematický softvér `R`.

V ďalšej časti sa budeme venovať dynamickej korelácii. Bližšie predstavíme algoritmus na odhad dynamickej korelácie medzi časovými radmi, *DCC-GARCH model*, a ukážeme jeho použitie na názornom príklade.

Posledná kapitola ukazuje rozdiel pri zanedbaní korelácie, uvažovaní konštantnej alebo dynamickej korelácie v konvergenčnom modeli úrokových mier. Ďalej odhadneme dynamickú koreláciu na medzi výmennými kurzami a vytvoríme zhluky rokov, v ktorých mala korelácia podobný vývoj.

1 Teoretický základ

V prvej kapitole zhrnieme potrebné pojmy a postupy známe z modelovania časových radov a finančnej matematiky, ktoré budeme v našej práci využívať.

1.1 Časové rady

Časové rady bývajú analyzované z dôvodu pochopenia minulosti a predikovaniu, respektíve odhadovaniu budúcnosti, aby umožnili manažérom, či politikom byť dostatnočne informovaní pri rozhodovaní. Analýza časových radov kvantifikuje hlavné znaky a náhodné odchýlky v dátach. Tieto dôvody, v kombinácii s neustálym zvyšovaním výkonov počítačov, umožnili metódam časových radov širokú aplikovateľnosť v štátnej správe, priemysle a obchode. V tejto časti textu spomenieme základné pojmy a modely, čerpať budeme z [7] a [3], kde môžeme nájsť teoretické poznatky ako aj ich používanie prostredníctvom softvéru `R`. Ako základné zdroje pre modelovanie dynamickej podmienenej korelácie sme využili [1] a [2].

1.1.1 Základné pojmy

Definícia 1.1. *Nech* $\{X_t\}$ *je časový rad so strednou hodnotou*

$$\mu_x(t) = E(X_t),$$

kde $E(X_t^2) < \infty$ *a kovariančnou funkciou*

$$\gamma_x(i,j) = Cov(X_i, X_j) = E[(X_i - \mu_x(i))(X_j - \mu_x(j))]$$

pre každé i, j*. Potom* $\{X_t\}$ *nazývame (slabo)* ***stacionárny****, ak sú splnené nasledujúce podmienky:*

1. $\mu_x(t)$ *nezávisí od času* t*,*
2. $\gamma_x(t, t+k)$ *nezávisí od času* t *pre každé* k*.* [3]

Definícia 1.2. *Časový rad* $\{w_t : t = 1, 2, ...\}$ *nazývame* ***biely šum****, ak premenné* $\{w_1, w_2, ...\}$ *sú nekorelované náhodné premenné so strednou hodnotou* 0 *a varianciou* $\sigma^2 < \infty$*. Potom je zrejme stacionárny a platí* $\gamma_w(i,j) = 0$ *pre každé* $i \neq j$*.* [3]

Stacionaritu časových radov môžeme dosiahnuť integráciou, čo v našom kontexte predstavuje diferencovanie. Takéto rady označujeme $I(d)$, kde d predstavuje stupeň integrácie. $I(1)$ je teda označenie pre časový rad integrovaný rádom jedna, ak je stacionárny po prvých diferenciách. Každej analýze dát by preto mali predchádzať testy jednotkového koreňa. V našej práci budeme používať **rozšírený Dickey-Fullerov test**, angl. *Augmented Dickey-Fuller test* (ADF), ktorý je aplikovaný na model:

$$\Delta x_t = \delta + \beta t + \alpha x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + w_t,$$

kde δ je konštanta (prítomnosť driftu v dátach), β je koeficient časového trendu (prítomnosť trendu v dátach), α vyjadruje autoregresný parameter, $\sum_{i=1}^{p} \gamma_i \Delta x_{t-i}$ predstavuje autoregresné členy a w_t je biely šum. V rámci testovania určujeme prítomnosť členov δ a βt, od čoho závisia aj kritické hodnoty testu. Testujeme hypotézy

$$H_0 : \alpha = 0,$$

$$H_1 : \alpha < 0,$$

kde z $H_0 : \alpha = 0$ vyplýva, že časový rad obsahuje jednotkový koreň a proces je nestacionárny, $H_1 : \alpha < 0$ potom zrejme značí, že časový rad neobsahuje jednotkový koreň. Testovacia štatistika, ktorá sa riadi neštandardným rozdelením Dickeyho a Fullera, má tvar

$$DF_t = \frac{\hat{\alpha}}{SE(\hat{\alpha})},$$

kde $\hat{\alpha}$ je odhad parametra α a $SE(\cdot)$ štandardnú odchýlku.

Autokorelácia je korelácia premennej so sebou samou v rozdielnych časových úsekoch, lagoch.

Definícia 1.3. ***Autokorelačná funkcia*** *(ďalej len ACF), ρ_k, definovaná pre stacionárny časový rad x_t ako funkcia k-teho lagu má tvar*

$$\rho_k = Cor(x_t, x_{t+k}) = \frac{Cov(x_t, x_{t+k})}{\sigma^2}. \quad [7]$$

V prípade, že X_t je biely šum, odhad autokorelácie $\hat{\rho}_k$ má v asymptoticky normálne rozdelenie s varianciou $\sim \frac{1}{T}$. Autokorelačnú funkciu využijeme pri modelovaní časových radov pomocou ARMA a GARCH modelov ako test, či bol ARMA model vhodne zvolený a test, či je GARCH model potrebný.

1.1.2 ARMA proces

Rad $\{X_t\}$ je autoregresný proces rádu p, značíme AR(p), ak

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} + \omega_t, \tag{1.1}$$

kde α_i sú parametre modelu, pričom $\alpha_p \neq 0$.
Moving average process MA(q), takzvané kĺzavé priemery rádu q, sú lineárnou kombináciou q posledných bielych šumov definované ako

$$X_t = \omega_t + \beta_1 \omega_{t-1} + \beta_2 \omega_{t-2} + ... + \beta_q \omega_{q-1}. \tag{1.2}$$

Z rovnice (1.1) vyjadríme polynóm pomocou Woldovej reprezentácie, pričom korene rovnice

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

musia byť mimo jednotkového kruhu kvôli stacionarite procesu [3].
Užitočnú triedu modelov môžeme získať spojením AR a MA členov, zlúčením (1.1) a (1.2) do jedného výrazu. Časový rad $\{X_t\}$ nazývaný *autoregressive moving average process* ARMA rádu (p, q), značíme ARMA(p, q), ak teda

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \omega_t + \beta_1 \omega_{t-1} + ... + \beta_q \omega_{q-1}.$$

Pri modelovaní časových radov pomocou ARMA je potrebný test jednotkového koreňa, napríklad vyššie spomínaný ADF test. Ak je jednotkový koreň prítomný, dáta musíme diferencovať toľkokrát, koľkonásobný je jednotkový koreň. Výsledný model potom označujeme ARIMA(p, d, q), kde d označuje stupeň integrácie - násobnosť diferencovania. Po odhadnutí AR a MA členov je potrebné skontrolovať autokorelácie rezíduí, ktoré testujeme pomocou ACF. Tieto hodnoty by mali byť pod 5% hranicou. Model je potom konzistentný s predpokladom, že rezíduá sú realizáciou bieleho šumu, čo podporuje správnosť použitého modelu.

1.1.3 (G)ARCH model

Pri modelovaní volatility potrebujeme model, ktorý umožňuje podmienené zmeny variancie. Jeden z prístupov je použiť autoregresný model pre varianciu, čo vedie k nasledovnej definícii. Rad $\{\epsilon_t\}$ nazývame autoregresno-podmieneno-heteroskedastický (*autoregressive conditional heteroskedastic*) rádu p, značíme ARCH(q), ak

$$\epsilon_t = \omega_t \sqrt{\sigma_t^2},$$
$$\sigma^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2,$$

kde ω_t je biely šum, $\omega_t \sim N(0, \sigma^2)$ a α_i sú parametre modelu pre ktoré platí: $\alpha_0 > 0; \alpha_1, ..., \alpha_{q-1} \geq 0; \sum_{i=1}^{q} \alpha_i < 1$ a $\alpha_q > 0$. Pri odvodení variancie pre ϵ_t by sme videli, že variancia ARCH(1) procesu sa správa práve tak, ako AR(1) model . Preto pri fitovaní modelu sledujeme autokoreláciu štvorcov rezíduí, ktorá indikuje, či je ARCH model vhodný alebo nie. [7]

Rozšírením tohto modelu, ktorý nachádza široké spektrum aplikovateľnosti vo financiách, je tzv. zovšeobecnený ARCH model (*generalised ARCH*). Rad $\{\epsilon_t\}$ nazývame GARCH(p, q), ak

$$\epsilon_t = \omega_t \sqrt{\sigma_t^2},$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

kde α_i a β_j sú parametre modelu, pričom platí: $\alpha_0 > 0; \alpha_1, ..., \alpha_{q-1} \geq 0$ a $\alpha_q > 0; \beta_1, ..., \beta_{p-1} \geq 0; \beta_p > 0$ a $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$.

Po určení modelu znovu testujeme štandardizované rezíduá pomocou autokorelačnej funkcie, kde by sa heteroskedasticita nemala nachádzať, ak je model dobre definovaný. Parametre ARIMA a GARCH modelov ako aj výsledky testov môžeme získať napríklad pomocou softvéru `R`, knižníc `astsa` [24] a `urca` [25].

1.1.4 DCC-GARCH model

Od začiatku 21. storočia si môžeme všimnúť podstatný a narastajúci záujem o analýzu dynamickej kovariancie a korelácie v rámci investičných nástrojov. Obzvlášť sa kládol dôraz na analýzu finančných akcií, neskôr sa zvyšoval záujem v energetike [4]. V tomto smere je najrozšírenejšia variácia viacrozmerného GARCH modelu opisujúca dynamickú podmienenú koreláciu (*Dynamic conditional correlation*, ďalej len DCC), DCC-GARCH model.

V práci sa budeme zaoberať modelmi, ktoré vykazujú štatisticky signifikantné výkyvy v koreláciách v čase, teda kovariančná matica sa mení v čase. Pre odhad DCC využijeme spomínaný model aj v našej práci.

Nasledovný text sme čerpali z [1] a [2]. Majme pre nejaké časové rady dobre špecifikované ARMA a GARCH modely [1]. Odhadnutú korelačnú maticu získame pomocou DCC-GARCH modelu postupnosťou nasledovných krokov:

$$Cov(r_t r_t^T) = H_t = D_t R_t D_t, \tag{1.3}$$

kde r_t je n rozmerný vektor rezíduí získaných z ARMA modelu s nulovým vektorom stredných hodnôt, D_t je diagonálna matica podmienených štandardných odchýlok (i-ty člen diagonály je definovaný ako $\sqrt{\epsilon_{it}}$) a R_t je matica podmienených korelácií. Matica R_t je v prípade DCC získaná preškálovaním:

$$R_t = diag\{Q_t\}^{-\frac{1}{2}} Q_t diag\{Q_t\}^{-\frac{1}{2}}, \tag{1.4}$$

kde $Q_t = q_{ij,t}$ je podmienená kovariančná matica rezíduí s rozmerom $n \times n$:

$$Q_t = (1 - \alpha - \beta)\overline{Q} + \alpha(\varepsilon_{t-1}\varepsilon_{t-1}^T) + \beta Q_{t-1}, \tag{1.5}$$

kde ε_t sú štandardizované rezíduá získané ako $\varepsilon_t = D_t^{-1} r_t$ a $\overline{Q}$ je nepodmienená (v čase konštantná) matica $n \times n$ rezíduí ε_t a α, β sú nezáporné skaláry, kvôli stacionarite spĺňajúce podmienku $\alpha + \beta < 1$. Prvok korelačnej matice R_t má potom tvar

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t} q_{jj,t}}}, \quad i, j = 1, 2, ..., n; i \neq j.$$

[1] Vhodne zvolený ARMA a GARCH model - zamietnutá prítomnosť jednotkového koreňa v dátach, zamietnutá autokorelácia v rezíduách získaných z ARMA modelu, prítomnosť heteroskedasticity v dátach a zamietnutý dodatočný výskyt heteroskedasticity v rezíduách modelovaných pomocou GARCH

Test konštantnosti korelácie

"Dobre špecifikované ARMA a GARCH modely" nie sú postačujúcou podmienkou k modelovaniu DCC. Potrebujeme naviac overiť, či má korelácia dynamickú štruktúru. K tomu môžeme použiť napríklad test na konštantnosť korelácie opísaný v práci [9], odkiaľ s miernou úpravou čerpáme aj nasledujúci text.

Testujeme nulovú hypotézu o konštantnosti korelácie oproti alternatívnej hypotéze o dynamickej podmienenej korelácii:

$$H_0 : R_t = \bar{R},$$
$$H_1 : L(R_t) = L(\bar{R}) + \beta_1 L(R_{t-1}) + ... + \beta_s L(R_{t-s}) \quad \forall t \in T. \qquad [2]$$

V rámci testovacej procedúry odhadneme koreláciu štandardizovaných rezíduí získaných z GARCH modelu a zároveň štandardizujeme vektor rezíduí odmocninou z matice $\bar{R}$. Za platnosti H_0 by mali byť tieto rezíduá nezávislé, rovnako rozdelené s identickou kovariančnou maticou rovnou identickej matici I_n. Premenná, ktorú budeme modelovať pomocou regresie pre testovanie hypotéz má tvar

$$Y_t = U\Big[(\hat{\bar{R}}^{-\frac{1}{2}}\hat{D_t}^{-1}\hat{\epsilon}_t)(\hat{\bar{R}}^{-\frac{1}{2}}\hat{D_t}^{-1}\hat{\epsilon}_t)^T - I_n\Big]. \qquad [3]$$

Autoregresia má tvar $Y_t = \alpha + \beta_1 Y_{t-1} + ... + \beta_s Y_{t-s} + \omega_t$, respektíve v maticovom zápise:

$$Y_{T\times 1} = X_{T\times(s+1)} \times \delta_{(s+1)\times 1} + \omega_{T\times 1}$$

$$\begin{bmatrix} | \\ Y_t \\ | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ 1 & Y_{t-1} & \cdots & Y_{t-s} \\ | & | & & | \end{bmatrix} \times \begin{bmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_s \end{bmatrix} + \omega$$

Metódou najmenších štvorcov odhadneme koeficienty lineárnej regresie a testujeme ich signifikantnosť. V prípade platnosti H_0 by malo platiť: $\alpha = 0; \beta_i = 0, \ \forall i \in \{1, ..., s\}$. Na testovanie tejto hypotézy môžeme použiť napríklad Waldov test, ktorého testová štatistika

$$W = \frac{\hat{\delta} X^T X \hat{\delta}^T}{\mathrm{Var}(\hat{\delta})}$$

má za platnosti H_0 asymptoticky $\chi^2_{(s+1)}$ rozdelenie.

[2] V nasledujúcich poznámkach predpokladajme rozklad matice $A = L+D+U$. Potom $L(\cdot)$ označuje dolnú trojuholníkovú maticu, T je rozmer časových radov.

[3] $U[\cdot]$ označuje hornú trojuholníkovú maticu.

1.2 Oceňovanie dlhopisov pomocou PDR

V nasledujúcej časti zhrnieme vzťahy, ktoré našli uplatnenie pri modelovaní úrokových mier. Teoretické znalosti sme prebrali z [14]. Všetkým, čo budeme v tejto časti opisovať sa zaoberali mnohé práce pred nami (napríklad [19], [17] a ďalšie), preto uvedieme len základné myšlienky a najdôležitejšie vzťahy. Detailný postup odvodenia a riešenia takejto PDR, ako aj odvodenie trhovej ceny rizika, či ďalšie mnohokrát opisované teoretické znalosti nebudeme znovu "predstavovať".

1.2.1 Dlhopis

Cenu bezkupónového dlhopisu v čase t so splatnosťou v čase T s koncovou podmienkou $P(0,T) = 1$ môžeme definovať ako

$$P(t,T) = e^{-R(t,T)(T-t)}, \tag{1.6}$$

kde $R(t,T)$ je časová štruktúra úrokových mier určujúca závislosť úrokovej miery (výnosu) od maturity dlhopisu. Z predchádzajúceho vzťahu potom vyplýva:

$$R(t,T) = -\frac{\ln(P(t,T))}{T-t}. \tag{1.7}$$

1.2.2 Jednofaktorové modely

V jednofaktorových modeloch sa predpokladá, že okamžitá úroková miera r je charakterizovaná pomocou stochastickej diferenciálnej rovnice, ktorú môžeme zapísať vo všeobecnom tvare

$$dr = \mu(r,t)dt + \sigma(r,t)dw$$

Deterministická časť procesu určuje trend vo vývoji úrokovej miery, volatilita $\sigma(r,t)$ určuje charakter náhodných fluktuácií úrokovej miery v okolí deterministickej zložky.

Medzi často spomínané modely patrí **CKLS model**: $dr = \kappa(\theta - r)dt + \sigma r^{\gamma}dw$ [5], či špeciálne prípady: **Vašíčkov model**, kedy $\gamma = 0$ [18] alebo **Cox-Ingersoll-Rossov model** pre $\gamma = \frac{1}{2}$ [8].

Existencia *rizikovo neutrálnej miery* značí, že na trhu neexistuje arbitráž, pričom prevod modelov do rizikovo neutrálnej miery umožňuje oceňovanie derivátov. CIR model v rizikovo neutrálnej miere má tvar

$$dr = (\kappa(\theta - r) - \lambda\sigma r)dt + \sigma\sqrt{r}dw$$

kde $\lambda\sqrt{r}$ predstavuje *trhovú cenu rizika*. Z modelu v takomto tvare potom vieme vyjadriť cenu dlhopisu $P = P(r, \tau)$ ako riešenie parciálnej diferenciálnej rovnice, kde $\tau = T - t$ označuje čas do expirácie. PDR má tvar

$$-\frac{\partial P}{\partial \tau} + (\kappa(\theta - r) - \lambda\sigma r)\frac{\partial P}{\partial r} + \frac{\sigma^2}{2} r \frac{\partial^2 P}{\partial r^2} - rP = 0, \quad r > 0, \tau \in (0, T)$$

a začiatočnú podmienku $P(r, 0) = 1$ pre všetky $r > 0$.

1.2.3 Dvojfaktorové modely

Základnou myšlienkou viacfaktorových modelov úrokovej miery je závislosť úrokovej miery r od viacerých faktorov. V prípade dvojfaktorových modelov označme tieto faktory x, y, teda $r = r(x, y)$. Vo všeobecnom prípade dvojfaktorového modelu budeme predpokladať, že faktory x, y vyhovujú nasledovným stochastickým diferenciálnym rovniciam:

$$dx = \mu_x(x, y)dt + \sigma_x(x, y)dw_1,$$
$$dy = \mu_y(x, y)dt + \sigma_y(x, y)dw_2,$$

pričom korelácia prírastkov dw_1 a dw_2 Wienerových procesov w_1 a w_2 je konštanta ρ, t.j. $E(dw_1 dw_2) = \rho dt$.

Konvergenčné modely

Špeciálnym prípadom dvojfaktorových modelov, ktorým sa venovala práca [19] a iné, sú konvergenčné modely. Využitie našli pri modelovaní vývoja úrokových mier štátu, ktorý vstupuje do menovej únie, napríklad štát Európskej únie prijímajúci euro. Úrokové miery takého štátu konvergujú k úrokovým mieram menovej únie. Ako príklad uvádzame dvojfaktorový konvergenčný Vašíčkov model:

$$\begin{aligned} dr_d &= (a + b(r_e - r_d))dt + \sigma_d dw_d, \\ dr_e &= (c(d - r_e))dt + \sigma_e dw_e, \\ Cov(dw_d, dw_e) &= \rho dt \quad [6] \end{aligned} \tag{1.8}$$

kde r_e reprezentuje európsku úrokovú mieru, r_d domácu. Koeficient b, respektíve c určujú rýchlosť konvergencie domácej úrokovej miery k európskej s odchýlkou danou parametrom a, respektíve európskej úrokovej miery k limitnej hodnote d.

Prevedením rovníc do rizikovo neutrálnej miery a následným preznačením parametrov dostávame parciálnu diferenciálnu rovnicu pre cenu domáceho dlhopisu $P(r_d, r_e, \tau)$

$$-\frac{\partial P}{\partial \tau}+(a_1+a_2 r_d+a_3 r_e)\frac{\partial P}{\partial r_d}+(b_1+b_2 r_e)\frac{\partial P}{\partial r_e}+\frac{\sigma_d^2}{2}\frac{\partial^2 P}{\partial r_d^2}+\frac{\sigma_e^2}{2}\frac{\partial^2 P}{\partial r_e^2}+\rho(T-\tau)\sigma_d\sigma_e\frac{\partial^2 P}{\partial r_d \partial r_e}-r_d P = 0 \tag{1.9}$$

s počiatočnou podmienkou

$$P(r_d, r_e, \tau) = 1 \text{ pre } \forall r_e, r_d > 0 \tag{1.10}$$

Konvergenčné modely s dynamickou koreláciou

Predpokladáme dlhopisy s rôznymi dynamickými korelačnými funkciami času $\rho_1(t)$ a $\rho_2(t)$, respektíve $\rho_1(T-\tau)$ a $\rho_2(T-\tau)$. Ceny týchto dlhopisov hľadáme v tvare:

$$P_i(\tau) = e^{A_i(\tau)-D(\tau)r_d-U(\tau)r_u}, \qquad i = 1, 2 \tag{1.11}$$

Z počiatočnej podmienky (1.10) zrejme $A_i(0) = 0, D(0) = 0, U(0) = 0$. Do PDR (1.9) dosadíme jednotlivé derivácie hľadaného tvaru riešenia (1.11), následnými úpravami získame systém obyčajných diferenciálnych rovníc:

$$\begin{aligned}
\dot{D}(\tau) &= 1 + a_2 D(\tau), \\
\dot{U}(\tau) &= a_3 D(\tau) + b_2 U(\tau), \\
\dot{A}_i(\tau) &= -a_1 D(\tau) - b_1 U(\tau) + \frac{\sigma_d^2 D^2(\tau)}{2} + \frac{\sigma_e^2 U^2(\tau)}{2} + \rho_i(T-\tau)\sigma_d\sigma_e D(\tau)U(\tau).
\end{aligned} \tag{1.12}$$

čoho riešením je:

$$\begin{aligned}
D(\tau) &= \frac{-1+e^{a_2\tau}}{a_2}, \\
U(\tau) &= -\frac{a_3\big(a_2(e^{b_2\tau}-1)-b_2(e^{a_2\tau}-1)\big)}{a_2 b_2(a_2-b_2)}, \\
A_i(\tau) &= \int_0^\tau -a_1 D(s) - b_1 U(s) + \frac{\sigma_d^2 D^2(s)}{2} + \frac{\sigma_e^2 U^2(s)}{2} + \rho_i(T-s)\sigma_d\sigma_e D(s)U(s)ds.
\end{aligned} \tag{1.13}$$

Použitím dynamickej korelácie v konvergenčnom modeli sa zaoberá aj práca [17].

1.3 Hierarchické zhlukovanie

Zhlukovanie je dôležité technika strojového učenia "bez učiteľa" (ku vstupným dátam nie je známy výstup) používaná pri zhlukovaní objektov podľa ich podobných vlastností. Hierarchické zhlukovanie možno rozdeliť do dvoch kategórií: *aglomeratívne zhlukovanie "zdola-nahor"* a *divízne zhlukovanie "zhora-nadol"*.

V našej práci využijeme aglomeratívne zhlukovanie, ktoré môže byť charakterizované ako sekvenčné, aglomeratívne, hierarchické a neprekrývajúce sa. Objekty sú najprv rozdelené po samom, teda ak máme n objektov, dostávame n zlukov. Následne sa v každom kroku algoritmu vytvorí jeden pár najpodobnejších zhlukov, až kým nakoniec neostane len jeden zhluk. Existuje sedem zvyčajne využívaných metód na tvorbu väzieb: *single*, *complete*, *average* (UPGMA), *weighted* (WPGMA), *centroid* (UPGMC), *median* (WPGMC) and *Wardova metóda*. Podrobnejší rozbor týchto metód môžeme nájsť v [10]. Kritériá, ktoré sa využívajú pri týchto metódach môžeme všeobecne reprezentovať rekurentným vzťahom publikovaným v práci [11].

Majme dva zhluky C_i a C_j, ktoré boli spojené do zhluku C_k. Vzdialenosť d_{km} medzi zhlukom C_k a akýmkoľvek ďalším zhlukom C_m je definovaná ako

$$d_{km} = d(C_i \cup C_j, C_m) = \alpha_1 d_{im} + \alpha_2 d_{jm} + \beta d_{ij} + \gamma |d_{im} - d_{jm}|, \tag{1.14}$$

kde α_1, α_2, β a γ sú špecifické parametre pre jednotlivé metódy, ich reprezentácia je zhrnutá v nasledujúcej tabuľke:

Väzba	α_1	α_2	β	γ	Alternatívny vzorec
Single	0,5	0,5	0	$-0{,}5$	$d_{ij} = \min_{x \in C_i, y \in C_j} d_{xy}$
Complete	0,5	0,5	0	0,5	$d_{ij} = \max_{x \in C_i, y \in C_j} d_{xy}$
Average	$\frac{\lvert C_i\rvert}{\lvert C_i\rvert+\lvert C_j\rvert}$	$\frac{\lvert C_j\rvert}{\lvert C_i\rvert+\lvert C_j\rvert}$	0	0	$d_{ij} = \frac{1}{\lvert C_i\rvert + \lvert C_j\rvert} \sum_{x \in C_i, y \in C_j} d_{xy}$
Weighted	0,5	0,5	0	0	
Centroid	$\frac{\lvert C_i\rvert}{\lvert C_i\rvert+\lvert C_j\rvert}$	$\frac{\lvert C_j\rvert}{\lvert C_i\rvert+\lvert C_j\rvert}$	$-\frac{\lvert C_i\rvert\lvert C_j\rvert}{(\lvert C_i\rvert+\lvert C_j\rvert)^2}$	0	
Median	0,5	0,5	-0.25	0	
Ward	$\frac{\lvert C_i\rvert+\lvert C_m\rvert}{\lvert C_i\rvert+\lvert C_j\rvert+\lvert C_m\rvert}$	$\frac{\lvert C_j\rvert+\lvert C_m\rvert}{\lvert C_i\rvert+\lvert C_j\rvert+\lvert C_m\rvert}$	$-\frac{\lvert C_m\rvert}{\lvert C_i\rvert+\lvert C_j\rvert+\lvert C_m\rvert}$	0	

Tabuľka 1: Parametre Lance-Williamsovej rekurentného vzorca pre sedem zvyčajne používaných väzbových schém. Text a tabuľka prebrané z [15].

2 Výpočet výrazov pomocou R

V tejto časti sa pozrieme na možnosti využitia softvéru `R` pri riešení algebraických výrazov, čo využijeme napríklad pri výpočtoch vyplývajúich z konvergenčných modelov s dynamickou koreláciou.

Prvotným cieľom bol výpočet parciálnej diferenciálnej rovnice (ďalej len PDR), či už kvôli efektivite práce, no taktiež v prípadoch, kde nie je možný prevod do systému obyčajných diferenciálnych rovníc (ďalej len ODR) a následné explicitné riešenie.

K výpočtu diferenciálnych rovníc môžeme použiť knižnice `deSolve` a `ReacTran`, pomocou ktorých sme riešili systému ODR, vedeli sme nájsť riešenie pre parabolickú, či hyperbolickú PDR (rovnica vedenia tepla), no nepodarilo sa nám to pre PDR, ktorú dostaneme ako vzťah pre cenu dlhopisu v konvergenčnom modeli.

Skúsili sme preto zamerať aj na ďalšie matematické softvéry. Prostredníctvom API[4] sme prostredníctvom R používali `Wolfram|Alpha` (ďalej len WA), no ani pomocou neho sme nezískali riešenie vyššie spomenutej PDR . WA ale vieme využiť napríklad pre získanie explicitného riešenia integrálov, ktoré vznikajú pri konvergenčných modeloch s dynamickou koreláciou, taktiež pri ďalších problémoch v našej práci, preto považujeme toto prepojenie za užitočné a postup "Ako na to" bude obsahom tejto kapitoly.

2.1 WolframAlpha API

Naša práce nie je technického zamerania, preto sa nebudeme zaoberať podrobnosťami a vysvetľovaním jednotlivých kódov a ich "pozadiu". Uvádzame len postup, ako získať výsledky z `WA` do `R`.

Najhlavnejším krokom celého procesu je pochopitelne samotné volanie `WA` API funkcie, vďaka ktorej `R` dokáže "komunikovať" s `WA`. Skript, pomocou ktorého ju voláme a používame je dostupný na webovej stránke [26].
Používateľ sa taktiež musí zaregistrovať na webovej stránke online aplikácie WA [31], kde si v profile vygeneruje voľne dostupný "App ID", ktorý je nutný pre využívanie API.

[4] *Application programming interface*, v programátorskej terminológii to je zbierka funkcií a tried, ktoré určujú akým spôsobom sa majú funkcie knižnic volať zo zdrojového kódu programu. Sú to programové celky, ktoré voláme namiesto naprogramovania celého postupu.

Nasledujúci R-kovský kód je akousi kostrou nášho programu, kde definujeme vygenerované App ID, priraďujeme nami zadefinovanú funkciu do premennej a získavame výstup z `WA` na základe nami zadaného výrazu:

```
app_id <- "TuVlozteZiskanyAppID"
wa <- WolframClient(app_id)
result <- wa$query("TuVlozteAlgebraickyVyraz")
```

Kompletný kód uvádzame v prílohe A.1.

Výstupom, ktorý sme získali je tabuľka[5] vo formáte XML, preto bolo našou ďalšou úlohou skonvertovať tento výsledok do výrazu, s ktorým vie `R` ďalej pracovať. Tabuľku sme pomocou knižníc `knitr` a `htmltab` previedli do R-kovej tabuľky s hodnotami vo formáte `character` a v nej sme identifikovali pole s výsledným výrazom. Takto získaný výraz sme postupne upravovali nasledovne:

- každé "`e^`" zamenili za "`exp`",
- doplnili znamienko násobenia "`*`",
- v prípade výpočtu neurčitého intregrálu sme odstránili konštantu "`+ constant`",
- goniometrické funkcie zmenili z tvaru `sin^X(f(Y))` do tvaru `(sin(f(Y)))^X`, respektíve z technických dôvodov do tvaru `(function(Y) sin(f(Y))^X)(Y)`

Poznámka: Tieto úpravy nepokrývajú všetky možnosti výstupov, preto je v prípade komplexnejších výpočtov a funkcií potrebné doplniť ďalšie úpravy finálneho výrazu.

[5] Tabuľka obsahuje všetky "výrazové výsledky", ktoré sú obvykle výstupom pri používaní webovej aplikácie, t.j. výsledok výrazu, alternatívne úpravy výsledného výrazu, rozšírené tvary, Taylorov rozvoj, limita, či špecifických prípadoch hodnoty určitého integrálu, lokálne extrémy, parciálne derivácie, definičný obor a obor hodnôt a iné. Tabuľka neobsahuje grafy ani iné "nevýrazové výsledky".

Modelový príklad

Povedzme, že máme nasledujúce zadanie:

$$\frac{d}{dx}\left(\frac{\sin(x)\cos(x)}{e^x}\right)$$

Po zadefinovaní funkcie `wa` podľa návodu v prílohe A.1 a bude vstup v `R` vyzerať nasledovne [6]:

```
result <- wa$query("d/dx (sin(x) cos(x)/e^x)")
```

Výstup

```
1/2 e^(-x) (2 cos(2 x) - sin(2 x))
```

nakoniec skonvertujeme, aby sme s ním mohli ďalej pracovať v `R`. Dostávame tvar

```
1/2 * exp(-x) * (2 * cos(2 * x) - sin(2 * x))
```

matematicky vyjadrený ako

$$\frac{1}{2}\, e^{-x}\big(2\cos(2x) - \sin(2x)\big)\ .$$

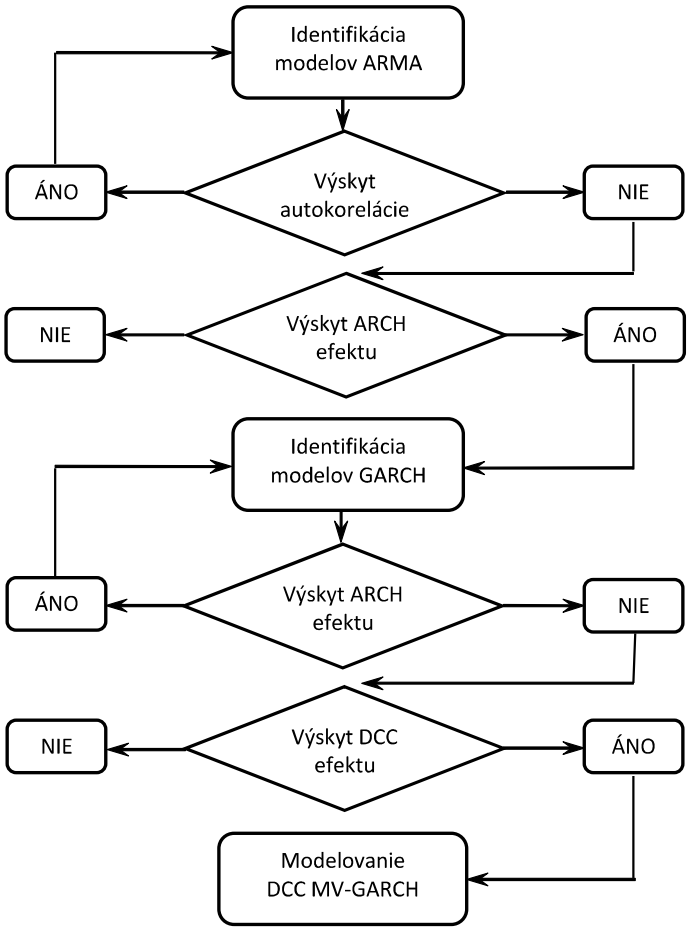
[6] Znamienko "$*$" ani zátvorky vo výraze "$\sin(x)$" `R` vstupu nie sú potrebné, no odporúčame vkladať jednoznačné zadanie výrazu vo formáte určenom pre WA.

3 Odhad dynamickej korelácie

V práci budeme odhadnovať dynamickú podmienenú koreláciu, k čomu využijeme nižšie uvedený algoritmus.

3.1 Algoritmus pre odhad DCC

Pre modelovanie dynamickej korelácie využijeme postupnosť viacerých krokov, pričom samotnú koreláciu odhadneme pomocou DCC-GARCH modelu. Podobným modelovaním sa zaoberali v zahraničí [13], ale aj na slovenských univerzitách [1]. Pre lepšiu predstavu uvádzame schému



Obr. 1: Schéma algoritmu [1]

Majme časové rady popisujúce vývoj nejakých premenných. Najprv teda odhadneme modely ARMA. Testujeme prítomnosť jednotkového koreňa pomocou Dickey-Fullerovho testu. Ak jeho prítomnosť nezamietneme, časové rady diferencujeme, čo opa-

kujeme dovtedy, kým nebude prítomnosť jednotkového koreňa zamietnutá. Následne určíme počet AR a MA členov. Testujeme získané rezíduá pomocou autokorelačnej funkcie a Ljung-Boxovho testu. Pre modelovanie korelácie totižto potrebujeme rezíduá očistené od autokorelácie.

Ďalším krokom je test štvorcov rezíduí na výskyt ARCH efektu. Tu overujeme, či má zmysel modelovať autoregresnú podmienenú heteroskedasticitu. Pre vyhovujúce rezíduá modelujeme ich podmienenú heteroskedasticitu pomocou jednorozmerných GARCH modelov. Tie sú dobre definované, ak sa v nich ďalšie ARCH efekty nenachádzajú.

Po identifikovaní GARCH modelov, v ktorých sa nepreukázal dodatočný výskyt heteroskedasticity, pokračujeme testovaním výskytu DCC efektu v získaných štandardizovaných rezíduách, ktorý bol opísaný v časti 1.1.4. Testujeme nulovú hypotézu o rovnosti konštantných podmienených korelácií a dynamických podmienených korelácií oproti alternatívnej hypotéze, ktorá hovorí o výskyte dynamickej štruktúry v koreláciách. Jednoducho povedané, testujeme, či sa korelácia medzi časovými radmi mení v čase. Ak nulovú hypotézu zamietneme, pristúpime k modelovaniu DCC-GARCH, ktorého postup sme priblížili v prvej časti práce, podkapitola 1.1.4. Samotné výpočty a test prevedieme v softvéri `R`, knižnice `rmgarch` [20] a `rugarch` [21].

3.2 Modelový príklad

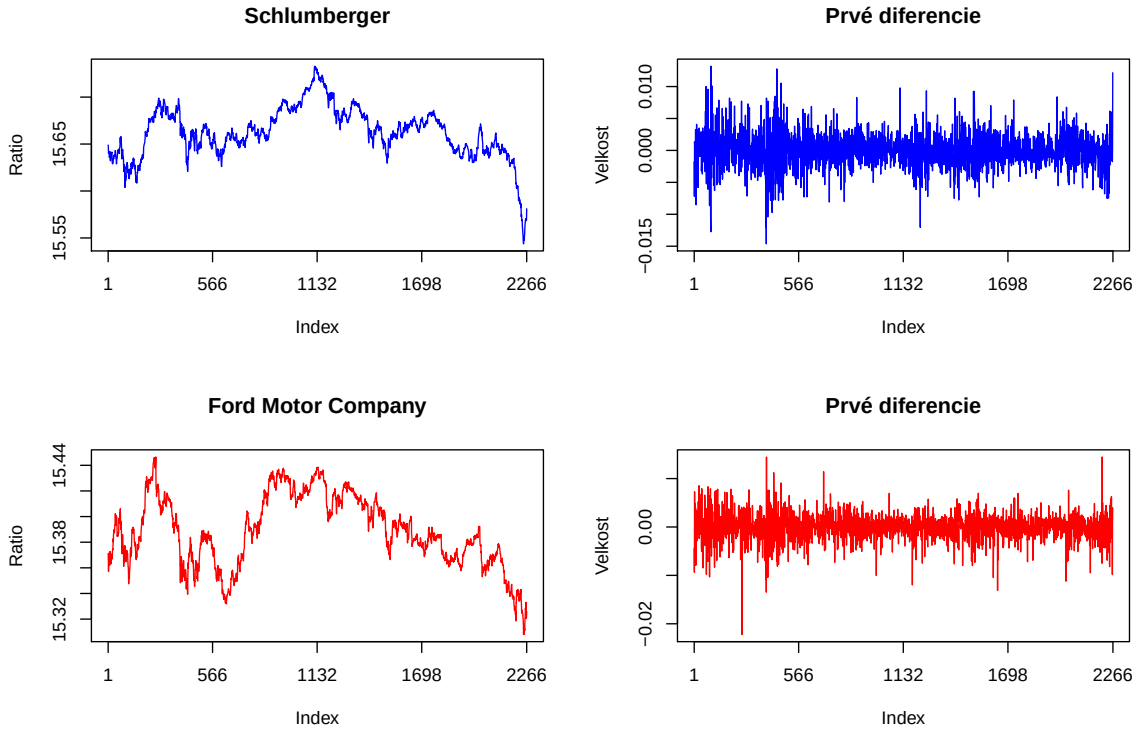
Podrobnejšie predstavíme výpočet DCC na konkrétnom príklade. Tiež uvedieme všetky výstupy z testov, či už vo forme grafov, testových štatistík alebo p-hodnôt.

Odhadneme dynamickú koreláciu medzi dvomi vybranými akciami. Jedná sa o akcie spoločností *Schlumberger*, ktorá prevádzkuje ropné služby a automobilovú nadnárodnú korporáciu *Ford Motor Company*. Rady obsahujú denné dáta z rokov $2010 - 2018$, čo predstavuje 2266 hodnôt.

Prvým krokom je identifikácia ARMA modelu, ktoré určujeme pre každý rad zvlášť. Rozšírený Duckey-Fullerov test nezamietol ani v jednom prípade prítomnosť jednotkového koreňa, preto časové rady diferencujeme. Vývoj cien a diferencií vidíme na obrázku 2.

ADF následne jasne zamietol prítomnosť jednotkového koreňa. Testujeme na hladine

významnosti 5%, pričom pre zamietnutie nulovej hypotézy musí byť hodnota testovej štatistiky nižšia ako kritická hodnota. Výsledky testových štatistík vidíme v tabuľke 2. Poznamenajme, že 5% kritické hodnoty $-2,86$ a $-1,95$ prislúcha dátam s driftom a dátam neobsahujúcim drift ani trend.



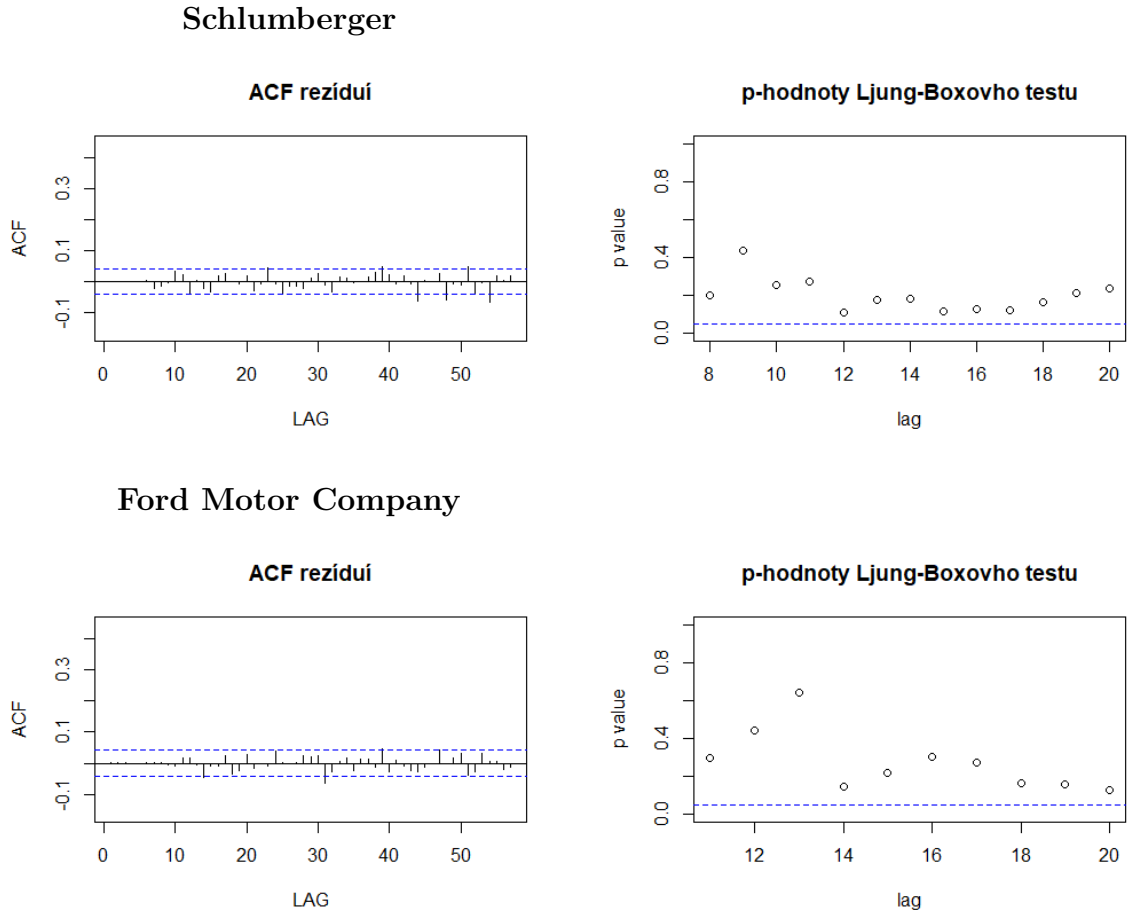
Obr. 2: Vývoj radov a ich prvých diferencií

	Pôvodné dáta	**Prvé diferencie**
5% kritická hodnota	$-2{,}86$	$-1{,}95$
Schlumberger	$-1{,}76$	$-32{,}14$
Ford Motor Company	$-2{,}05$	$-32{,}59$

Tabuľka 2: Výsledky ADF testu

Môžeme pristúpiť k určeniu počtu AR a MA členov. V získaných rezíduách skontrolujeme výskyt autokorelácie pomocou ACF a Ljung-Boxovho testu. Výsledky vidíme na obrázku 3.

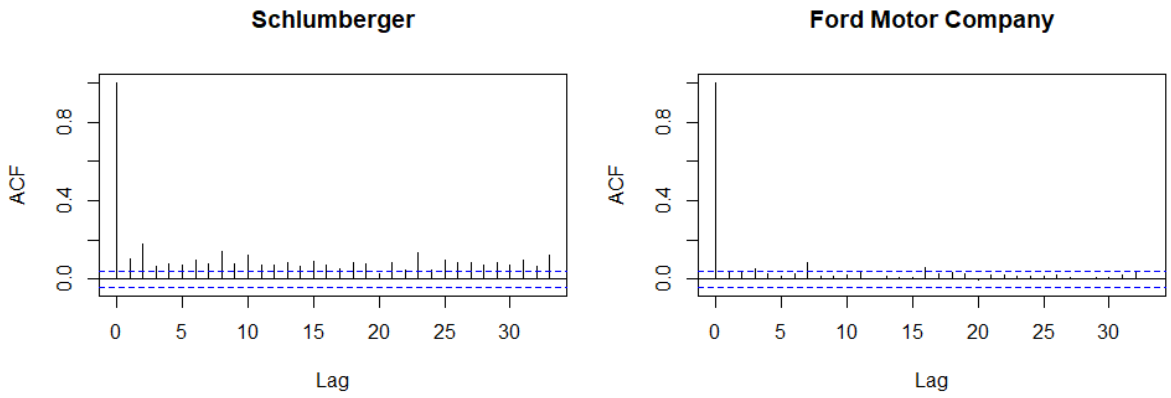
Testy zamietli autokoreláciu a teda sme našli dobre špecifikované modely ARMA pre naše dáta. Pre Schlumberger to je ARIMA$(6, 1, 1)$ a pre Ford Motor Company



Obr. 3: Testy určených ARMA modelov

ARIMA$(9, 1, 1)$.

Následne testujeme výskyt heteroskedasticity v dátach. Testujeme nulovosť autokorelácie štvorcov rezíduí získaných z ARMA modelov. Signifikantnosť autokorelácií vidíme na obrázku 4.



Obr. 4: ACF test štvorcov rezíduí

Keďže ich prítomnosť nebola zamietnutá, pristúpime k určeniu GARCH modelov. V oboch prípadoch sa ukazujú ako vyhovujúce GARCH$(1,1)$ modely, čo potvrdzuje aj Ljung-Boxov test štandardizovaných rezíduí a ich druhých mocnín, nakoľko sa dodatočný výskyt heteroskedasticity nepreukázal. Výsledné *p-hodnoty* boli pre akcie spoločnosti *Schlumberger* v rozmedzí $0{,}55 - 0{,}99$ a pre *Ford Motor Company* $0{,}72 - 1$.

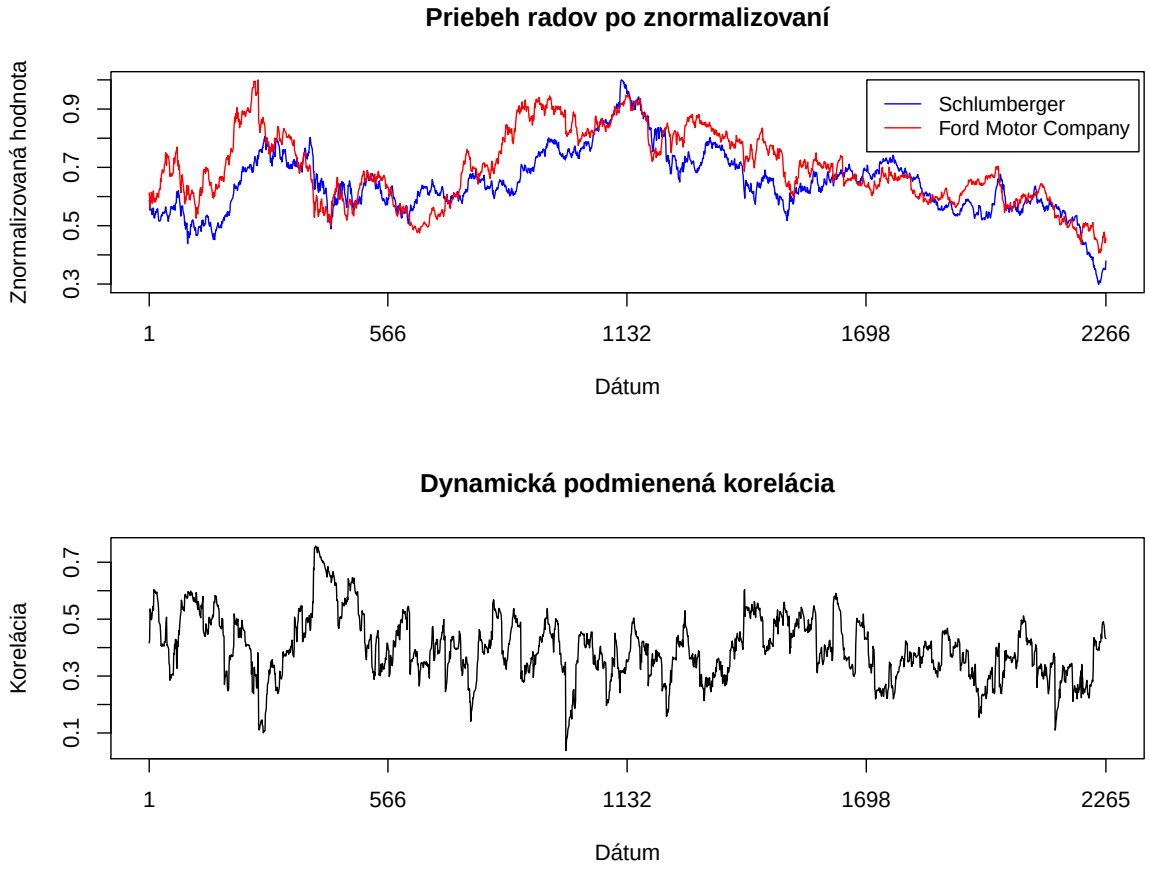
Podarilo sa nám teda dobre špecifikovať ARMA a GARCH modely pre dáta a ich varianciu, ich zhrnutie vidíme v tabuľke 3:

	Schlumberger	**Ford Motor Company**
ARIMA	ARIMA(6,1,1)	ARIMA(9,1,1)
GARCH	GARCH(1,1)	GARCH(1,1)

Tabuľka 3: Odhadnuté ARIMA a GARCH modely

Pokračujeme testovaním prítomnosti DCC efektu. Test zamietol hypotézu o konštantnosti korelácií s *p-hodnotou* 0,0054. Môžeme teda pristúpiť k odhadu dynamickej podmienenej korelácie, ktorej priebeh vidíme na Obr. 5.

Obrázok 5 zobrazujúci dynamickú podmienenú koreláciu hovorí o "uveriteľnosti" odhadu. Z vývoja hodnôt v čase je vidieť celková vzájomná korelovanosť radov, o čom vypovedá odhadnutá kladná korelácia fluktuujáca okolo hodnoty 0,4.



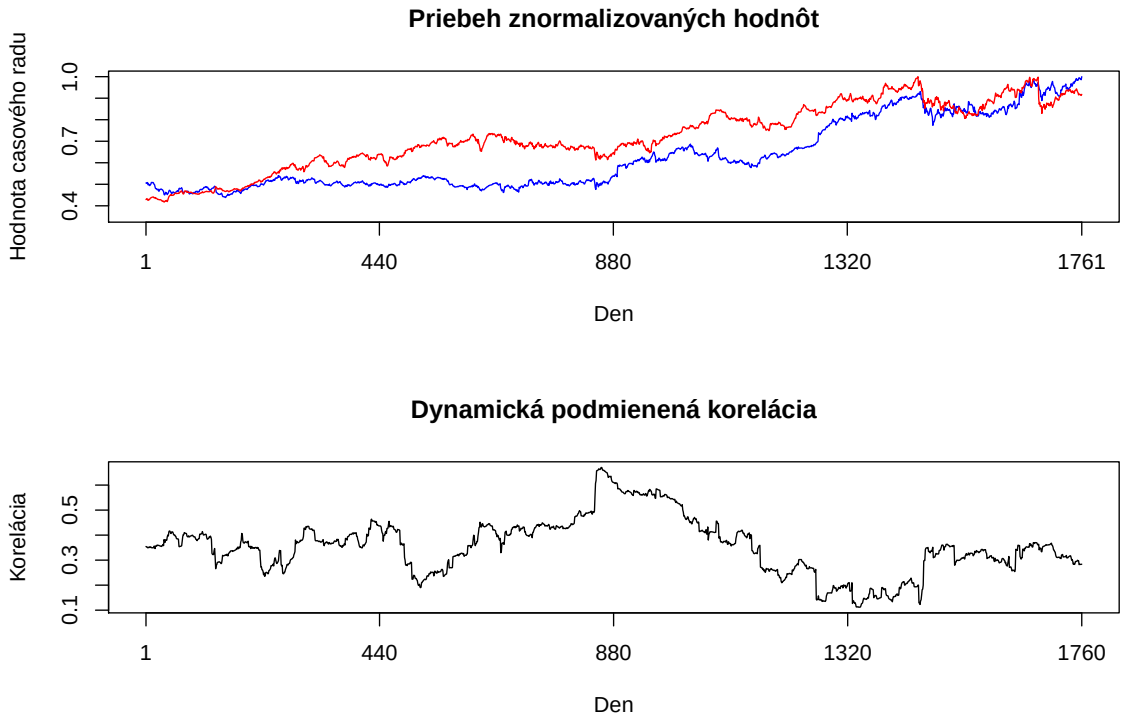
Obr. 5: Priebeh korelácie

Postrehy k DCC-GARCH modelu

DCC-GARCH model považujeme za dobrý prístup k modelovaniu dynamickej korelácie medzi časovými radmi. Naviac nie sme obmedzení len na dvojicu časových radov, model môže odhadovať korelácie pre viaceré rady súčasne. Za ďalšiu pozitívnu vlastnosť považujeme rýchle reagovanie na zmenu korelácie, čoho dôkazom je aj modelový príklad.

Je ale potrebné poukázať na konzervatívnosť Englovho testu prítomnosti dynamickej korelácie. Pre odhad korelácie je potrebný veľký počet dát, pri nízkom počte sa nám nepodarilo nájsť takú dvojicu časových radov, medzi ktorými by bola konštantná korelácia zamietnutá.

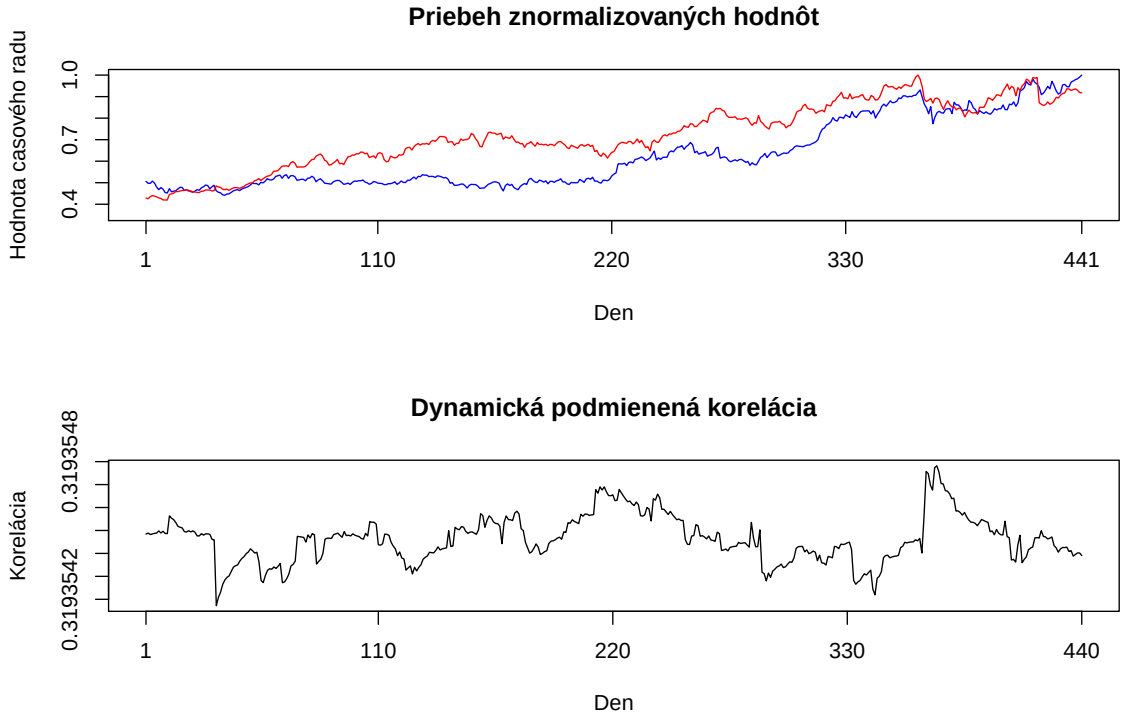
Uvádzame ukážke vplyvu počtu hodnôt na odhad DCC. Priebeh časových radov vidíme na obrázku 6.



Obr. 6: Priebeh časových radov a odhadnutej korelácie

Dané časové rady už na prvý pohľad naznačujú určitú závislosť, čo sa odzrkadlilo aj v odhadnutej korelácii kolísajúcej v rozmedzí [0,11; 0,67]. Englov DCC test (pochopiteľne) zamietol konštantnosť s *p-hodnotou* rádovo 10^{-6}. Pozrime sa ale ako sa zmení hodnota pri skrátení dĺžky časových radov.

Použijeme rovnaké časové rady, vezmeme len každú druhú hodnotu. V tomto prípade sme dostali veľmi podobný výsledok v teste konštantnosti korelácie aj pri jej odhade. Keď ale vezmeme len každú štvrtú hodnotu, čím dostávame časové rady popisujúce rovnaké obdobie, no obsahujúce len 441 hodnôt, dostávame úplne iné výsledky. Englov DCC test nezamietol konštantnosť korelácie s *p-hodnotou* 0,68 a ak sme sa aj napriek tomu pokúsili odhadnúť dynamickú koreláciu, dostali sme "dynamickú" koreláciu fluktuujúcu v rozmedzí rádovo 10^{-7}, priebeh ilustruje obrázok 7.



Obr. 7: Priebeh časových radov a odhadnutej korelácie

Taktiež nie sú všetky časové rady vhodné na modelovanie pomocou ARMA a GARCH modelov. V niektorých prípadoch sa nám stalo, že po určení AR a MA členov sa v získaných rezíduách nevyskytoval ARCH efekt a teda nebol splnený predpoklad k modelovaniu korelácie pomocou DCC-GARCH modelu.

4 Praktická časť

V tejto časti použijeme Vašíčkov dvojfaktorový model na ocenenie domáceho dlhopisu, pričom porovnáme vplyv korelácie na výslednú cenu.

Ďalej využijeme DCC-GARCH model opísaný v kapitole 3 k odhadu dynamickej korelácie priamo z dát. Pre finančné dáta taktiež využijeme *vzdialenosť korelačných matíc* na určenie podobných korelácií pre jednotlivé roky. Na základe tejto vzdialenosti rozdelíme roky do skupín prostredníctvom *hierarchického zhlukovania*.

4.1 Oceňovanie dlhopisov

Je niekoľko prístupov, ako zakomponovať koreláciu do modelu. Tie najjednochšie ju zanedbajú, prípadne odhadujú ako konštantu. Týmto sa zaoberala napríklad práca [19].

Ďalšou možnosťou je nahradenie korelácie funkciou času, ktorá je navrhnutá tak, aby opisovala jej odhadovaný vývoj v čase. V prípade vývoja úrokovej miery štátu, ktorého delí niekoľko mesiacov od vstupu do menovej únie (model (1.8) z kapitoly 1) je zrejme korelácia rastúca. Takýmto príkladom môže byť napríklad korelácia odhadnutá funkciou času

$$\rho_1(t) = \frac{c_1}{1 + e^{-\frac{t}{c_2}}}, \tag{4.1}$$

ktorá je konkávna a funkčná hodnota konverguje ku nami určenej konštante c_1 z intervalu $(0, 1)$, konštanta c_2 určuje "rýchlosť" konvergencie. Analogický postup bol použitý v práci [17], kde bola navrhnutá korelačná funkcia

$$\rho_2(t) = 1 - c_3 e^{-c_4 t}. \tag{4.2}$$

Pre takýto prístup numericky porovnáme výsledky pre dve rôzne funkcie korelácie.

4.2 Rozdiel cien dlhopisu pre rôzne funkcie korelácie

Pre pripomenutie znovu uvádzame riešenie systému ODE (1.13) získaného z Vašíčkovho modelu:

$$D(\tau) = \frac{-1 + e^{a_2\tau}}{a_2},$$
$$U(\tau) = -\frac{a_3\big(a_2(e^{b_2\tau} - 1) - b_2(e^{a_2\tau} - 1)\big)}{a_2 b_2(a_2 - b_2)},$$
$$A(\tau) = \int_0^\tau -a_1 D(s) - b_1 U(s) + \frac{\sigma_d^2 D^2(s)}{2} + \frac{\sigma_e^2 U^2(s)}{2} + \rho(T-s)\sigma_d\sigma_e D(s)U(s)ds.$$

Z neho dostávame explicitné vyjadrenie rozdielu cien dlhopisov, resp. ich logaritmov:

$$\ln P_1(\tau) - \ln P_2(\tau) = A_1(\tau) - A_2(\tau) = \int_0^\tau \Big(\rho_1(T-s) - \rho_2(T-s)\Big)\sigma_d\sigma_e D(s)U(s)ds. \tag{4.3}$$

Uvažujme korelačnú funkciu času $\rho_1(T-\tau), t = T-\tau$, napríklad horeuvedenú funkciu (4.1). Rozoberieme 2 prípady:

1. $\boldsymbol{\rho_2(T-\tau) = c_0}$, kde $\mathbf{c_0}$ je konštanta

a) Nech $c_0 = 0$, teda rozdiel vyjadrený vzťahom (4.3) pri použití $\rho_1(T-\tau)$ a zanedbaní korelácie. Explicitným výpočtom integrálu (pomocou `Wolfram|Alpha`) dostávame netriviálny výraz, ktorého súčasťou je aj niekoľko hypergeometrických funkcií v tvare

$${}_2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!}, \text{ kde } (x)_n = \frac{x!}{(x-n)!}.$$

Nás ale zaujíma rozdiel ceny dlhopisu pre malé maturity, preto kvôli vyčísleniu rozdiel vyjadríme pomocou Taylorovho rozvoja 5. rádu:

$$\ln P_1(\tau) - \ln P_2(\tau) = \Big(A_1(0) - A_2(0)\Big) + \Big(A_1^{(1)}(0) - A_2^{(1)}(0)\Big)\tau + \Big(A_1^{(2)}(0) - A_2^{(2)}(0)\Big)\frac{\tau^2}{2}$$
$$+\Big(A_1^{(3)}(0) - A_2^{(3)}(0)\Big)\frac{\tau^3}{6} + \Big(A_1^{(4)}(0) - A_2^{(4)}(0)\Big)\frac{\tau^4}{24} + \Big(A_1^{(5)}(0) - A_2^{(5)}(0)\Big)\frac{\tau^5}{120} + \mathcal{O}(\tau^6) \tag{4.4}$$

Pre $\tau = 0$ dostávame $D(0) = 0, U(0) = 0$. Ďalej platí:

$$\begin{aligned}
&A_1^{(i)}(0) - A_2^{(i)}(0) = 0 \qquad \text{pre } i \in \{1,2,3\},\\
&A_1^{(4)}(0) - A_2^{(4)}(0) = \frac{3c_1a_3\sigma_d\sigma_e e^{\frac{T}{c_2}}}{e^{\frac{T}{c_2}}+1},\\
&A_1^{(5)}(0) - A_2^{(5)}(0) = \frac{c_1a_3\sigma_d\sigma_e}{e^{\frac{T}{c_2}}+1}\Bigg(\frac{e^{\frac{T}{c_2}}(-3a_2b_2+4a_2^2-b_2^2)}{(a_2-b_2)},\\
&\qquad\qquad - \frac{14a_2b_2e^{\frac{2T}{c_2}}(a_2-b_2)\big(2a_2(e^{\frac{T}{c_2}}+1)c_2+b_2(e^{\frac{T}{c_2}}+1)c_2-1\big)}{(e^{\frac{T}{c_2}}+1)^3c_2^2}\Bigg).
\end{aligned} \tag{4.5}$$

Dosadením rovností (4.5) do (4.4) vyjadríme rozdiel pri použití dynamickej korelačnej funkcie (4.1) a pri zanedbaní korelácie potom dostávame

$$\begin{aligned}
\ln P_1(\tau) - \ln P_2(\tau) = \frac{c_1a_3\sigma_d\sigma_e}{e^{\frac{T}{c_2}}+1}\Bigg(e^{\frac{T}{c_2}}\frac{\tau^4}{8} + \Bigg(\frac{e^{\frac{T}{c_2}}(-3a_2b_2+4a_2^2-b_2^2)}{(a_2-b_2)}\\
- \frac{14a_2b_2e^{\frac{2T}{c_2}}(a_2-b_2)\big(2a_2(e^{\frac{T}{c_2}}+1)c_2+b_2(e^{\frac{T}{c_2}}+1)c_2-1\big)}{(e^{\frac{T}{c_2}}+1)^3c_2^2}\Bigg)\frac{\tau^5}{120}\Bigg) + \mathcal{O}(\tau^6).
\end{aligned} \tag{4.6}$$

b) Nech c_0 je ľubovoľná konštanta z intervalu $(0,1)$, vypočítame teda rozdiel medzi $\rho_1(T-\tau)$ a konštantnou koreláciou.
Z rovnice (4.3) dostávame vzťah

$$\ln P_1(\tau) - \ln P_2(\tau) = \int_0^\tau \frac{c_1}{1+e^{-\frac{s}{c_2}}}\sigma_d\sigma_e D(s)U(s)ds - c_0\int_0^\tau \sigma_d\sigma_e D(s)U(s)ds. \tag{4.7}$$

Integrály nahradíme Taylorovým rozvojom, čím dostávame rozdiel medzi použitím dynamickej korelačnej funkcie $\rho_1(T-\tau)$ a konštantnej korelácie:

$$\begin{aligned}
\ln P_1(\tau) - \ln P_2(\tau) = \frac{c_1a_3\sigma_d\sigma_e}{e^{\frac{T}{c_2}}+1}\Bigg(e^{\frac{T}{c_2}}\frac{\tau^4}{8} + \Bigg(\frac{e^{\frac{T}{c_2}}(-3a_2b_2+4a_2^2-b_2^2)}{(a_2-b_2)}\\
- \frac{14a_2b_2e^{\frac{2T}{c_2}}(a_2-b_2)\big(2a_2(e^{\frac{T}{c_2}}+1)c_2+b_2(e^{\frac{T}{c_2}}+1)c_2-1\big)}{(e^{\frac{T}{c_2}}+1)^3c_2^2}\Bigg)\frac{\tau^5}{120}\Bigg)\\
- c_0a_3\sigma_d\sigma_e\left(\frac{\tau^4}{8} + (5a_2+2b_2)\frac{\tau^5}{60}\right) + \mathcal{O}(\tau^6).
\end{aligned} \tag{4.8}$$

2. $\boldsymbol{\rho_2(\mathbf{T}-\boldsymbol{\tau}) = \mathbf{1} - \mathbf{c_3 e^{-c_4(T-\tau)}}}$, teda vyjadríme rozdiel ceny dlhopisu pri použití odlišných korelačných funkcií.

Pri výpočte rozdielu logaritmov cien dlhopisov dostávame podobný vzťah ako v (4.7), integrály nahradíme Taylorovym rozvojom, čím pre konkrétne korelačné funkcie dostávame vzťah:

$$\begin{aligned}\ln P_1(\tau) - \ln P_2(\tau) = \frac{c_1 a_3 \sigma_d \sigma_e}{e^{\frac{T}{c_2}}+1}\Bigg(e^{\frac{T}{c_2}}\frac{\tau^4}{8} + \Bigg(\frac{e^{\frac{T}{c_2}}(-3a_2b_2+4a_2^2-b_2^2)}{(a_2-b_2)} \\ - \frac{14a_2b_2e^{\frac{2T}{c_2}}(a_2-b_2)\big(2a_2(e^{\frac{T}{c_2}}+1)c_2 + b_2(e^{\frac{T}{c_2}}+1)c_2 - 1\big)}{(e^{\frac{T}{c_2}}+1)^3 c_2^2}\Bigg)\frac{\tau^5}{120}\Bigg) \\ -a_3\sigma_d\sigma_e\left(\frac{\tau^4}{8} + (5a_2+2b_2)\frac{\tau^5}{60}\right) + c_3a_3\sigma_d\sigma_e e^{-c_4T}\left(\frac{\tau^4}{8} + (5a_2+2b_2+6c_4)\frac{\tau^5}{60}\right) + \mathcal{O}(\tau^6)\end{aligned} \tag{4.9}$$

4.3 Rozdiel cien pre konkrétne parametre

V tejto časti ukážeme rozdiel vo vyššie spomenutých prístupoch k použitej korelácii na numerických príkladoch. Pre rozdiel časových štruktúr úrokových mier s koreláciami $\rho_1(T-\tau)$ a $\rho_2(T-\tau)$ zo vzťahu (1.7) dostávame

$$R_1(\tau) - R_2(\tau) = -\frac{\ln P_1(\tau) - \ln P_2(\tau)}{\tau}. \tag{4.10}$$

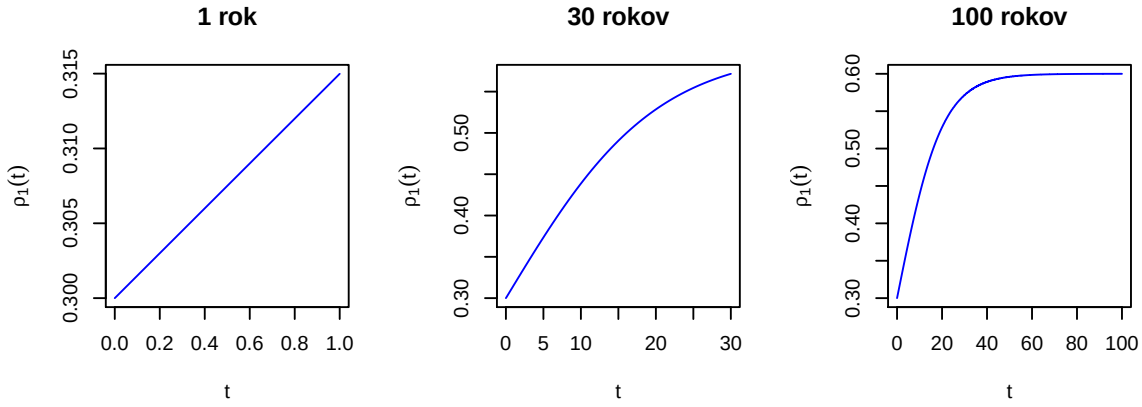
Uvažujme spomínanú korelačnú funkciu $\rho_1(T-\tau) = \dfrac{c_1}{1+e^{-\frac{T-\tau}{c_2}}}$ s parametrami z tabuľky 4. Jej priebeh vidíme na nasledujúcich grafoch:

Parametre Vašíčkovho modelu po transformácii premenných a parametre korelačných funkcií $\rho_1(T-\tau)$, $\rho_2(T-\tau)$ (4.1) (4.2) vidíme v tabuľke 4.

a_1	b_1	a_2	a_3	b_2	σ_d	σ_e	c_1	c_2	c_3	$c_4(T)$ [7]	T
0,01	0,003	−2	2	−0,2	0,04	0,01	0,6	10	0,7	$-\dfrac{\ln\left(\frac{1+e^{-\frac{T}{c_2}}-c_1}{c_3\left(1+e^{-\frac{T}{c_2}}\right)}\right)}{T}$	30

Tabuľka 4: Vstupné parametre

[7] Hodnota parametra c_4 sa mení v závislosti od dĺžky splatnosti T. Je vyjadrený tak, aby boli hodnoty korelačných funkcií v čase splatnosti domáceho dlhopisu totožné.



Obr. 8: Priebeh korelačnej funkcie pre rôzne obdobia

Naviac odhadneme rád konvergencie pomocou *experimentálneho rádu konvergencie* (experimental order of convergence, EOC), ktorým aproximujeme exponent α v rovnici

$$\ln P_1(\tau) - \ln P_2(\tau) = \mathcal{O}(\tau^\alpha), \text{ kde } \tau \to \infty. \tag{4.11}$$

Označme $err_i = \ln P_1(\tau_i) - \ln P_2(\tau_i)$. Exponent α potom odhadneme vzťahom:

$$EOC_i = \frac{\ln(err_i) - \ln(err_{i+1})}{\ln(\tau_i) - \ln(\tau_{i+1})}. \quad [12] \tag{4.12}$$

Rozobereme možnosti korelačnej funkcie $\rho_2(T - \tau)$ spomínané v predchádzajúcej podkapitole:

1. $\boldsymbol{\rho_2(T - \tau) = c_0}$, kde $\mathbf{c_0}$ je konštanta

a) Nech $c_0 = 0$. Pre $\tau \in \{\frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \frac{6}{12}, \frac{9}{12}, 1\}$, teda 1 rok pred splatnosťou domáceho dlhopisu a v niektorých mesiacov počas posledného roku dostávame hodnoty pre rozdiel vyjadrený v rovnici (4.3):

$\boldsymbol{\tau}$	$\frac{\lvert\mathbf{lnP_1(\tau)-lnP_2(\tau)}\rvert}{\boldsymbol{\tau}}$	$\mathbf{EOC_\tau}$
$\frac{1}{12}$	$2{,}9501 \times 10^{-8}$	3,8393
$\frac{2}{12}$	$2{,}1113 \times 10^{-7}$	3,7324
$\frac{3}{12}$	$6{,}3931 \times 10^{-7}$	3,6327
$\frac{6}{12}$	$3{,}7558 \times 10^{-6}$	3,3669
$\frac{9}{12}$	$9{,}5300 \times 10^{-6}$	3,1438
1	$1{,}7352 \times 10^{-5}$	2,9568

Tabuľka 5: Rozdiel pri použití $\rho_1(T-\tau)$ a zanedbaní korelácie

b) Nech $c_0 \doteq 0{,}57$. Tá je zvolená tak, aby mala v čase spatnosti $T = 30$ rovnakú hodnotu ako dynamická korelácia $\rho_1(30)$. Pre hodnoty τ z obdobia posledného roka do splatnosti dostávame hodnoty rozdielu:

$\boldsymbol{\tau}$	$\frac{\lvert\mathbf{lnP_1(\tau)-lnP_2(\tau)}\rvert}{\boldsymbol{\tau}}$	$\mathbf{EOC_\tau}$
$\frac{1}{12}$	$9{,}3149 \times 10^{-12}$	4,8374
$\frac{2}{12}$	$1{,}3315 \times 10^{-10}$	4,7292
$\frac{3}{12}$	$6{,}0399 \times 10^{-10}$	4,6284
$\frac{6}{12}$	$7{,}0710 \times 10^{-9}$	4,3597
$\frac{9}{12}$	$2{,}6829 \times 10^{-8}$	4,1352
1	$6{,}4971 \times 10^{-8}$	3,9483

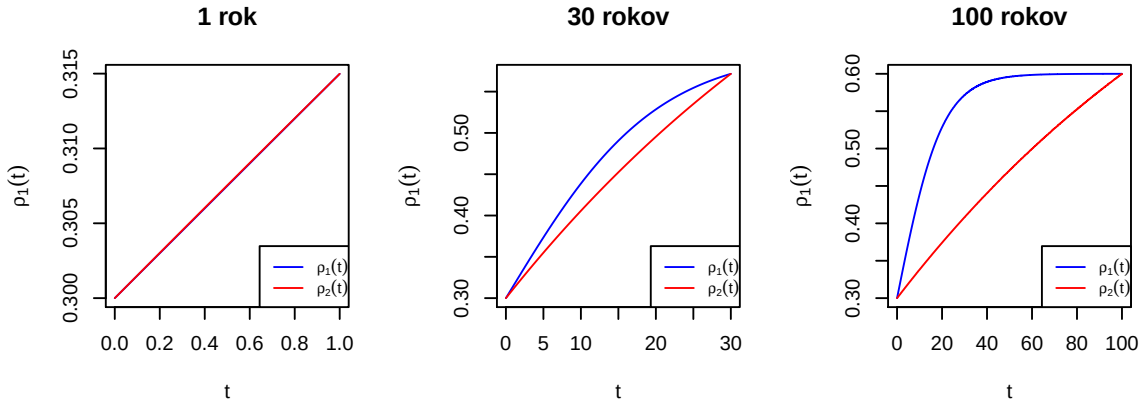
Tabuľka 6: Rozdiel pri použití $\rho_1(T-\tau)$ a konštantnej korelácie

2. $\boldsymbol{\rho_2(\mathbf{T}-\boldsymbol{\tau}) = \mathbf{1} - \mathbf{c_3 e^{-c_4(T-\tau)}}}$

Na obrázku 9 vidíme porovnanie korelačných funkcií $\rho_1(t)$ a $\rho_2(t)$, pričom parameter $c_4(T)$ nadobúda pre niektoré časy splatnosti T hodnoty približne:

$\mathbf{c_4(1)}$	$\mathbf{c_4(30)}$	$\mathbf{c_4(100)}$
0,0216	0,0164	0,0056

Pri takto zvolených korelačných funkciách uvádzame hodnoty rozdielu zo vzťahu 4.3 počas posledného roka do splatnosti v tabuľke 7.



Obr. 9: Priebeh korelačných funkcií pre rôzne obdobia

τ	$\frac{\lvert \ln P_1(\tau) - \ln P_2(\tau) \rvert}{\tau}$	EOC_τ
$\frac{1}{12}$	$1{,}4138 \times 10^{-11}$	4,8604
$\frac{2}{12}$	$2{,}0535 \times 10^{-10}$	4,7353
$\frac{3}{12}$	$9{,}3378 \times 10^{-10}$	4,6255
$\frac{6}{12}$	$1{,}0869 \times 10^{-8}$	4,3393
$\frac{9}{12}$	$4{,}0827 \times 10^{-8}$	4,1009
1	$9{,}7785 \times 10^{-8}$	3,9012

Tabuľka 7: Rozdiel pri použití $\rho_1(T-\tau)$ a $\rho_2(T-\tau)$

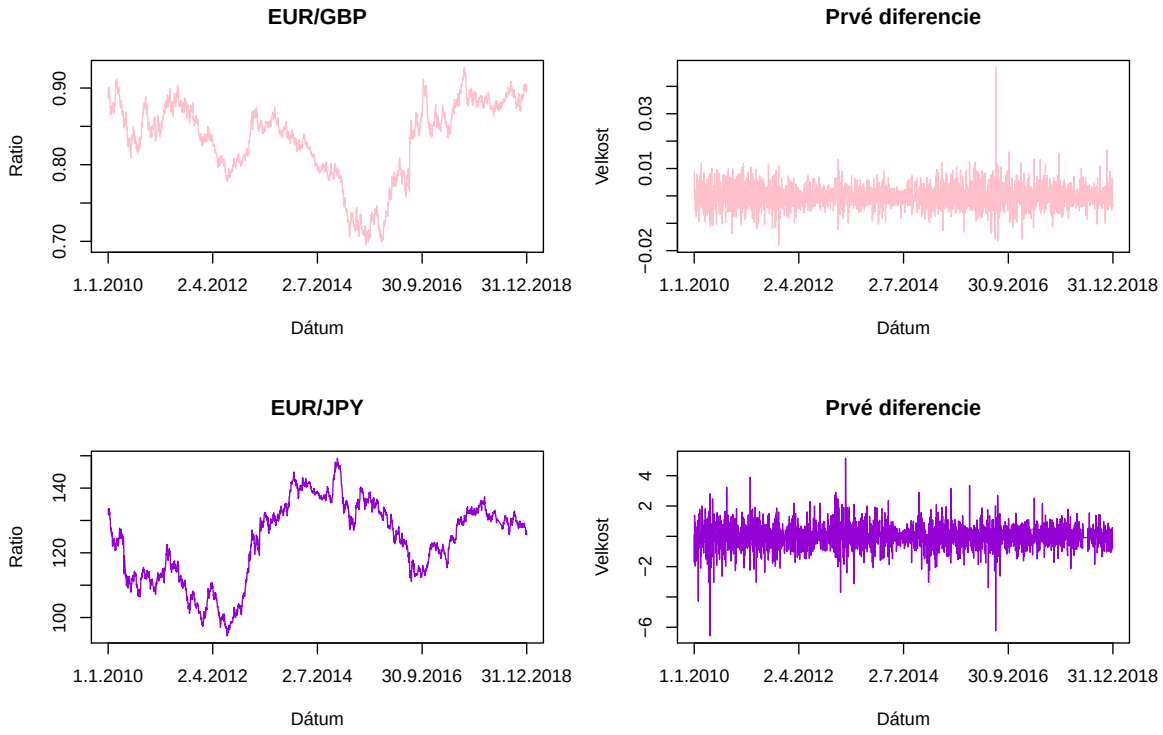
Rád konvergencie rovný 4 sa ukázal pri analytickom odvodení v rovniciach 4.6, 4.8 a 4.9, ako aj numerickým výpočtom EOC v tabuľke 5. V ďalších dvoch prípadoch, tabuľky 6 a 7, sme dostali rád rovný 5 z dôvodu, že korelácie boli zvolené tak, aby sa v čase splatnosti rovnali. Inak by sa s klesajúcim časom do splatnosti τ EOC blížilo k 4.

4.4 Odhad DCC pre výmenné kurzy

Odhadneme koreláciu pre výmenné kurzy eura voči britskej libre (EUR/GBP) a japonskému jenu (EUR/JPY). Využijeme denné dáta z rokov 2010 až 2018, t.j. časové rady obsahujú 2347 hodnôt. Dáta sú voľne dostupné na [28].

Prvým krokom je identifikácia ARMA modelu, ktorý určujeme pre každý kurz zvlášť. Rozšírený Duckey-Fullerov test ani v jednom prípade prítomnosť jednotkového koreňa

nezamietol, preto časové rady diferencujeme. Vývoj cien a diferencií vidíme na obrázku 10. ADF následne jasne zamietol prítomnosť jednotkového koreňa. Testujeme na hla-



Obr. 10: Vývoj kurzov a prvých diferencií

dine významnosti 5%, pričom hodnota testovej štatistiky musí byť nižšia ako kritická hodnota. Výsledky testových štatistík vidíme v tabuľke 8.

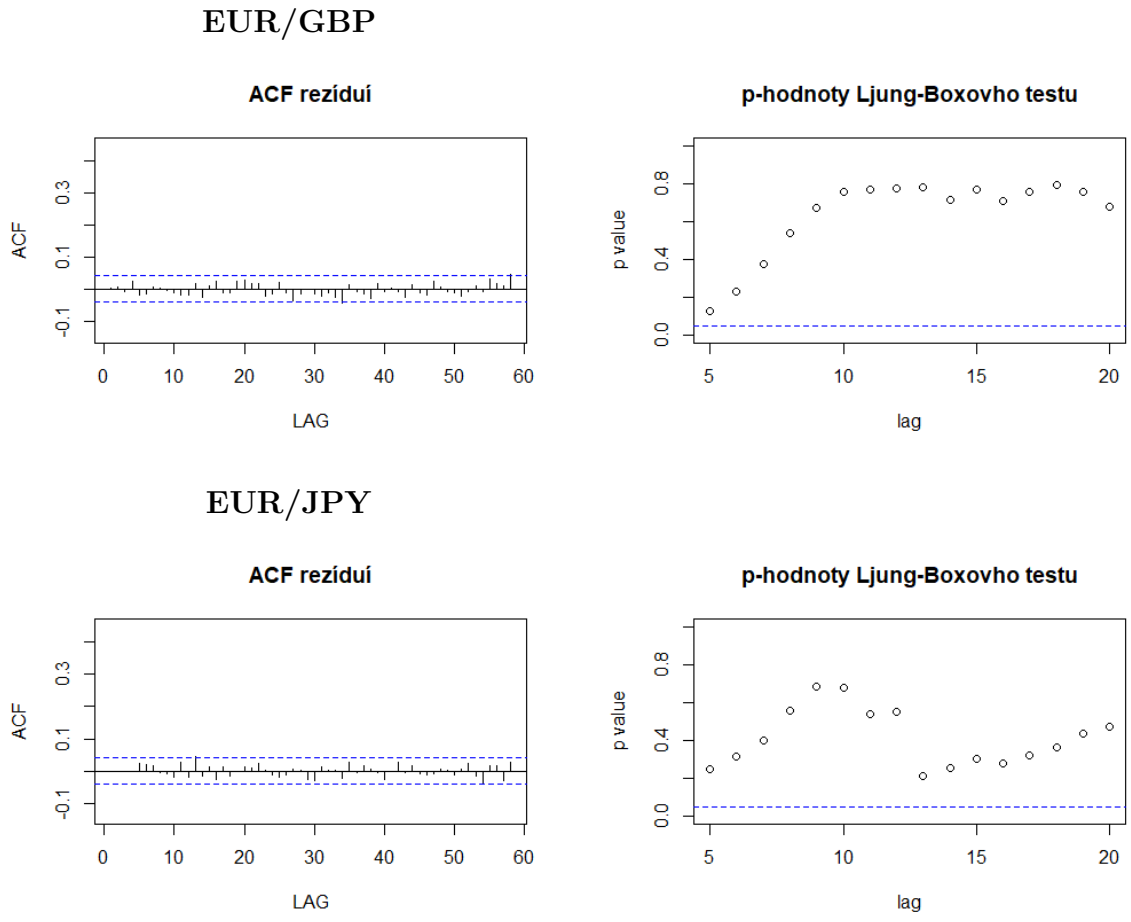
	Pôvodné dáta	**Prvé diferencie**
5% kritická hodnota	−2,86	−1,95
EUR/GBP	−2,07	−35,1
EUR/JPY	−1,74	−34,48

Tabuľka 8: Výsledky ADF testu

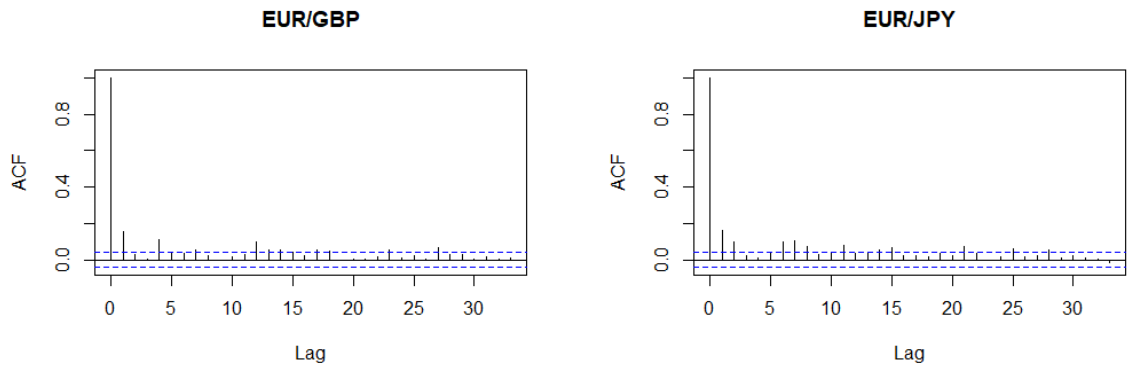
Môžeme pristúpiť k určeniu počtov AR a MA členov. V získaných rezíduách skontrolujeme výskyt autokorelácie pomocou ACF a Ljung-Boxovho testu. Výsledky vidíme na obrázku 11.

Testy zamietli autokoreláciu a teda sme našli dobre špecifikované modely ARMA pre naše dáta. Pre kurz EUR/GBP to je ARMA$(1, 3)$ a pre EUR/JPY ARMA$(3, 1)$.

Následne testujeme výskyt heteroskedasticity v dátach. Testujeme nulovosť autokorelácie štvorcov rezíduí získaných z ARMA modelov. Signifikantnosť autokorelácií vidíme na obrázku 12.



Obr. 11: Testy určených ARMA modelov

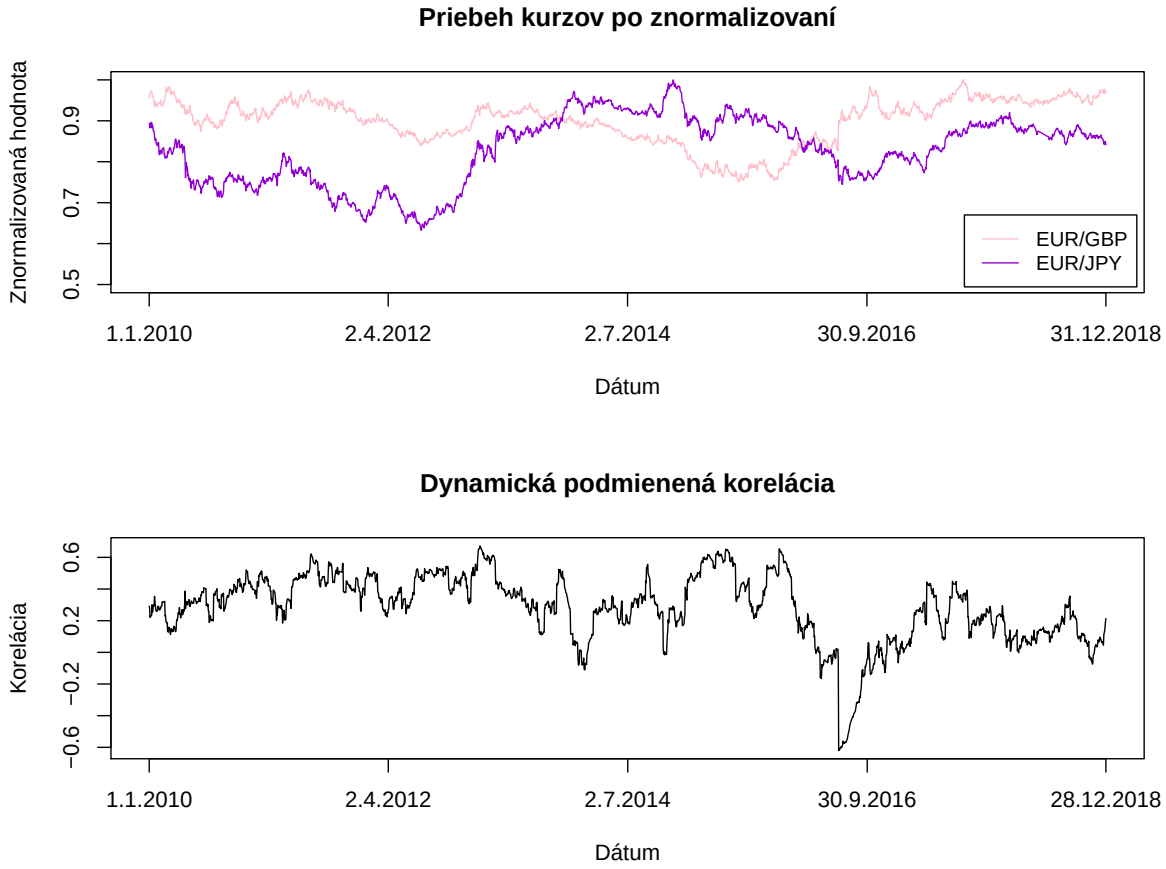


Obr. 12: ACF štvorcov rezíduí

Keďže ich prítomnosť nebola zamietnutá, pristúpime k určeniu GARCH modelov. V oboch prípadoch sa ukazujú ako vyhovujúce GARCH$(1,1)$ modely, čo potvrdzuje aj Ljung-Boxov test štandardizovaných rezíduí a ich druhých mocnín. Dodačný výskyt heteroskedasticity sa teda nepreukázal, výsledné *p-hodnoty* boli pre kurz EUR/GBP v rozmedzí $0{,}17 - 1$ a pre EUR/JPY $0{,}36 - 0{,}75$.

Pokračujeme testovaním prítomnosti DCC efektu. Test jasne zamietol hypotézu o konštantnosti korelácií s *p-hodnotou* rádovo 10^{-5}. Môžeme preto pristúpiť k odhadu dynamickej podmienenej korelácie, ktorej priebeh vidíme na obrázku 13.

Od začiatku roka 2010 po koniec roka 2013 majú kurzy podobný vývoj cien, čo sa odrazilo aj na odhadnutej kladnej korelácii s rastúcim tredom. Naproti tomu vidíme pokles od začiatku roku 2016, kedy bol ich vývoj opačný, kedy bola korelácii odhadnutá hodnota rovná $-0{,}6$. Náhla zmena korelácie potvrdzuje rýchlu odozvu modelu na zmenu vývoja časových radov.

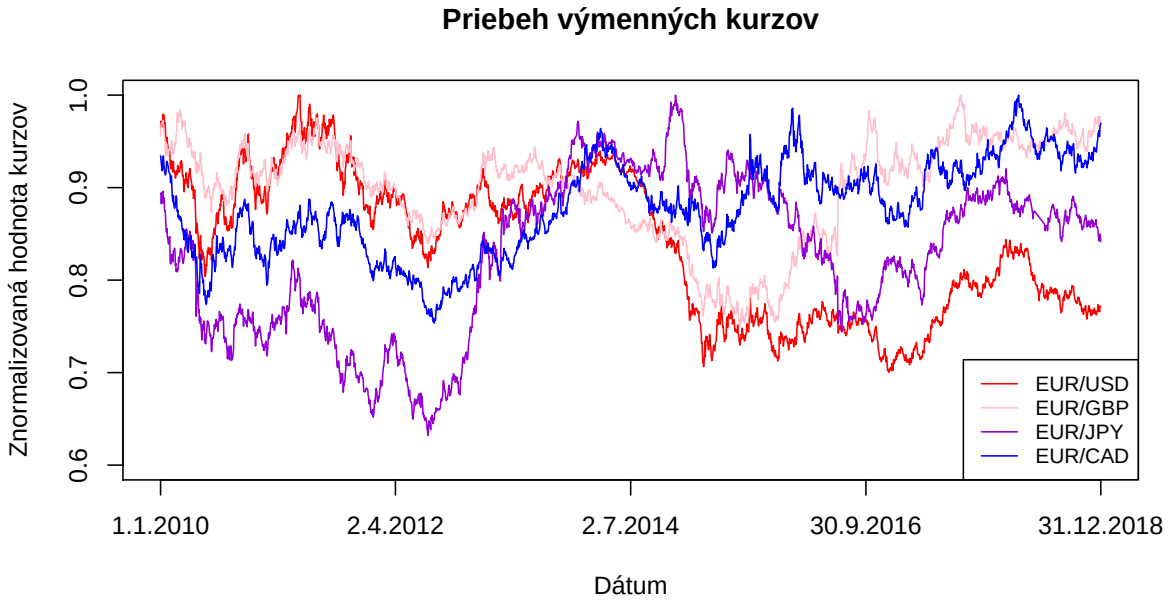


Obr. 13: Priebeh korelácie

4.4.1 Modelovanie DCC pre viaceré časové rady

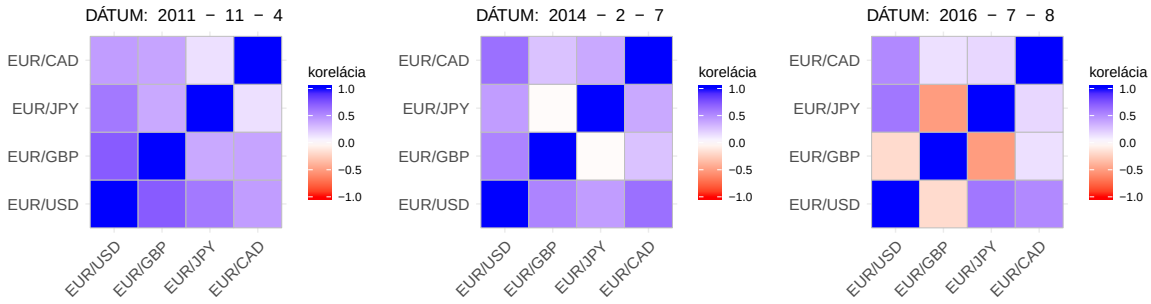
Časové rady z predchádzajúcej časti doplníme o ďalšie dva výmenné kurzy. Na dátovej sade zloženej zo štyroch časových radov ukážeme výpočet dynamickej korelácie pre viaceré časové rady súčasne.

Použijeme denné dáta z rokov 2010 až 2018 (2347 hodnôt) pre výmenné kurzy eura voči americkému doláru (EUR/USD), kanadskému doláru (EUR/CAD) a vyššie uvedené EUR/GBP a EUR/JPY. Priebeh znormalizovaných kurzov vidíme na obrázku 14



Obr. 14: Priebeh výmenných kurzov počas 13 rokov

Rovnako ako v predchádzajúcich prípadoch sme odhadli ARIMA a GARCH modely, Englov test jasne zamietol konštantnosť korelácií a tak sme pristúpili k odhadnu vzájomných dynamických korelácií. Celkovo 6 získaných korelácií následne vizualizujeme pomocou *korelogramov* [22]. Na obrázku 15 uvadzame vybrané korelogramy. Kompletnú animáciu možno vidieť na webovej stránke *https://s3.gifyu.com/images/gif78f52f8246af2448.gif*.



Obr. 15: Ukážka vybraných korelogramov

4.5 Vzdialenosť korelačných matíc

Vzdialenosť korelačných matíc (correlation matrix distance) R_1, R_2 je definovaná ako

$$\mathrm{CMD}(R_1, R_2) = 1 - \frac{\mathrm{tr}(R_1 R_2)}{\|R_1\|_f \|R_2\|_f} \in [0, 1]. \ [8] \tag{4.13}$$

Táto vzdialenosť je nulová, ak sú korelačné matice totožné z hľadiska použitej miery (4.13) a jednotková, ak sa líšia v maximálnom rozsahu. Prečo je takto definovaná vzdialenosť môže byť odôvodnené tak, že vzťah (4.13) prepíšeme ako skalárny súčin matíc prepísaných do vektorov:

$$\mathrm{CMD}(R_1, R_2) = 1 - \frac{\langle \mathrm{vec}(R_1), \mathrm{vec}(R_2) \rangle}{\|R_1\|_2 \|R_2\|_2}.$$

Z tohto je zrejmé, že CMD$(\cdot, \cdot)$ meria "ortogonalitu" daných matíc v $n \times n$ rozmernom priestore. [30]

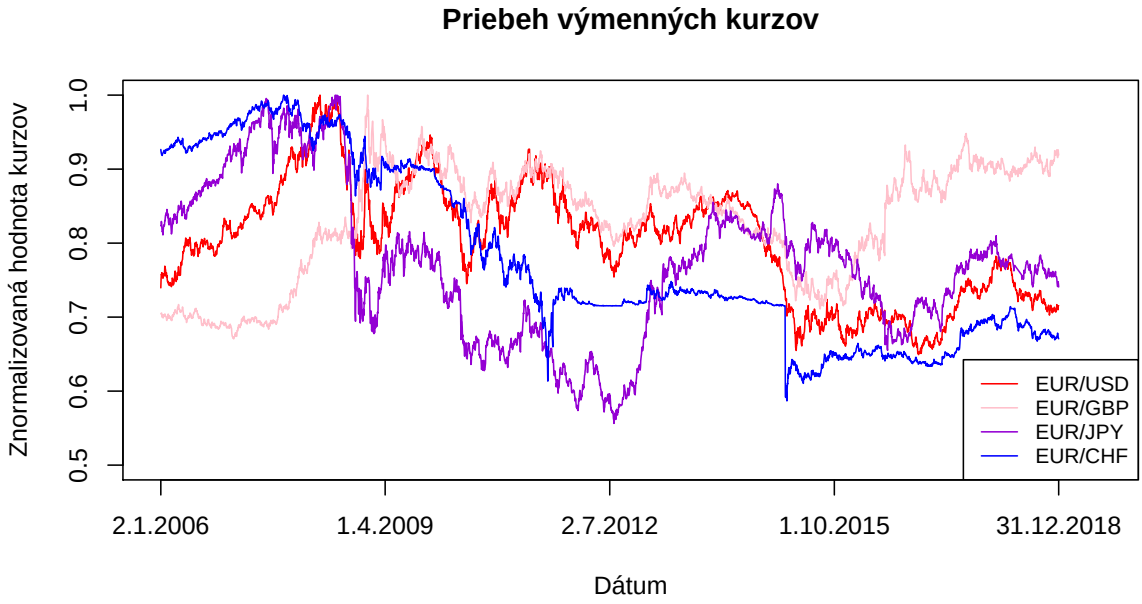
4.5.1 Zhlukovanie časových radov na základe CMD

Použijeme kurzy z predchádzajúcej časti s tým, že zameníme kurz EUR/CAD za kurz eura voči švajčiarskemu franku (EUR/CHF) a časové obdobie predĺžime o štyri roky, čo predstavuje 3391 hodnôt z rokov 2006 až 2018. Odhadneme koreláciu medzi transformovanými kurzami. Na základe vzdialeností korelačných matíc a hierarchického zhlukovania vytvoríme skupiny rokov s podobnými koreláciami. Rozsah hodnôt jednotlivých kurzov uvádzame v tabuľke 9 pretože kvôli ilustrácii sme hodnoty všetkých kurzov znormalizovali, ich priebeh zobrazujeme na obrázku 16.

Kurz	EUR/USD	EUR/GPB	EUR/JPY	EUR/CHF
Rozsah	[1,039; 1,600]	[0,655; 0,977]	[94,32; 169,5]	[0,985; 1,679]

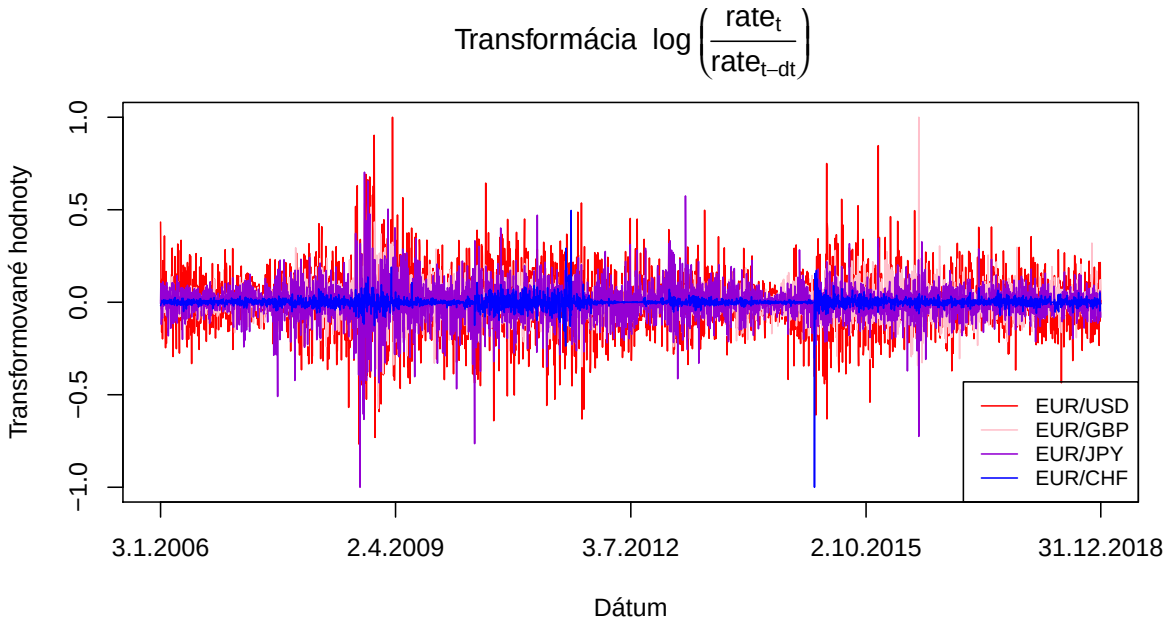
Tabuľka 9: Rozsah kurzov

[8] tr$(\cdot)$ označuje stopu matice, $\|.\|_f$ je tzv. Frobeniova norma: $\|A\|_f = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\mathrm{tr}(A^T A)}$



Obr. 16: Priebeh výmenných kurzov počas 13 rokov

Pre zachovanie stacionarity použijeme k odhadu korelačných matíc transformované hodnoty definované vzťahom $Y_t = \log\left(\frac{X_t}{X_{t-dt}}\right)$. Priebeh transformovaných časových radov po znormalizovaní vidíme nižšie (pôvodný rozsah bol $0{,}008 - 0{,}08$)



Obr. 17: Priebeh transformovaných výmenných kurzov

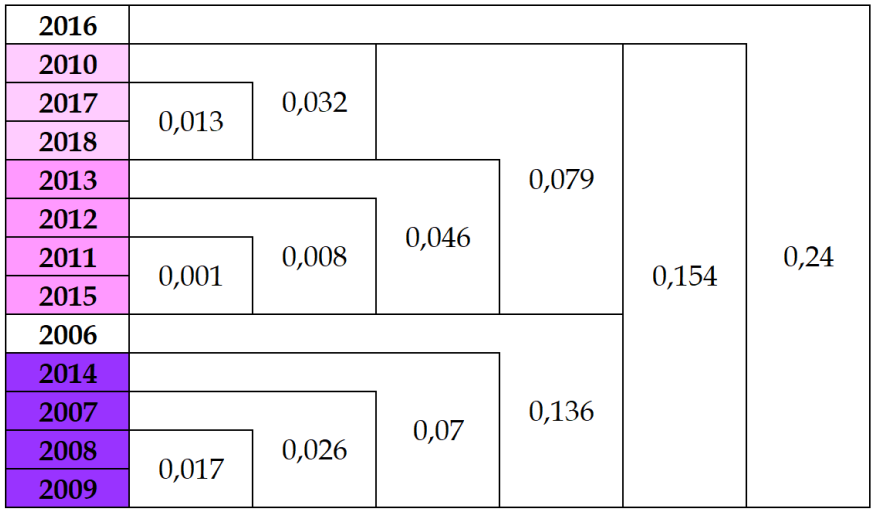
Z takto získaných dát odhadneme korelačné matice $R_{4\times4}$ pre jednotlivé roky odhadneme vzdialenosť korelačných matíc podľa vzťahu (4.13). Výsledné vzdialenosti vidíme

v tabuľke 10

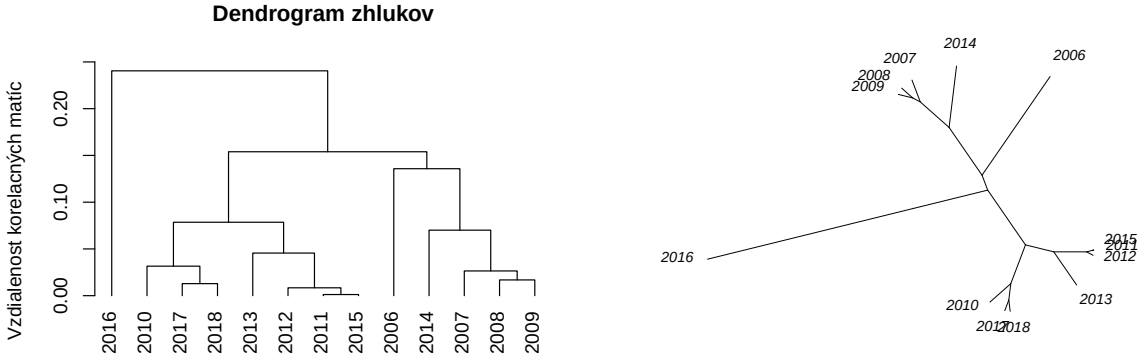
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
2007	0,136											
2008	0,134	0,025										
2009	0,080	0,026	0,017									
2010	0,154	0,104	0,034	0,056								
2011	0,119	0,126	0,046	0,058	0,017							
2012	0,092	0,134	0,060	0,050	0,039	0,008						
2013	0,055	0,111	0,064	0,067	0,051	0,035	0,046					
2014	0,070	0,070	0,068	0,044	0,131	0,091	0,071	0,078				
2015	0,120	0,148	0,061	0,069	0,023	0,001	0,007	0,039	0,100			
2016	0,179	0,046	0,101	0,071	0,168	0,218	0,214	0,198	0,140	0,240		
2017	0,115	0,061	0,027	0,039	0,017	0,043	0,064	0,037	0,103	0,054	0,105	
2018	0,126	0,037	0,020	0,021	0,032	0,065	0,078	0,074	0,106	0,079	0,063	0,013

Tabuľka 10: Vzdialenosť korelačných matíc pre roky 2006-2018

Následne vytvoríme zhluky na základe ich podobností. Použijeme hierarchické zhlukovanie opísané v kapitole 1, časti 1.3. Vzdialenosti zhlukov vidíme nižšie v tabuľke 11, ako aj ilustrované pomocou dendogramu a takzvaného "nezakoreneného" stromu.



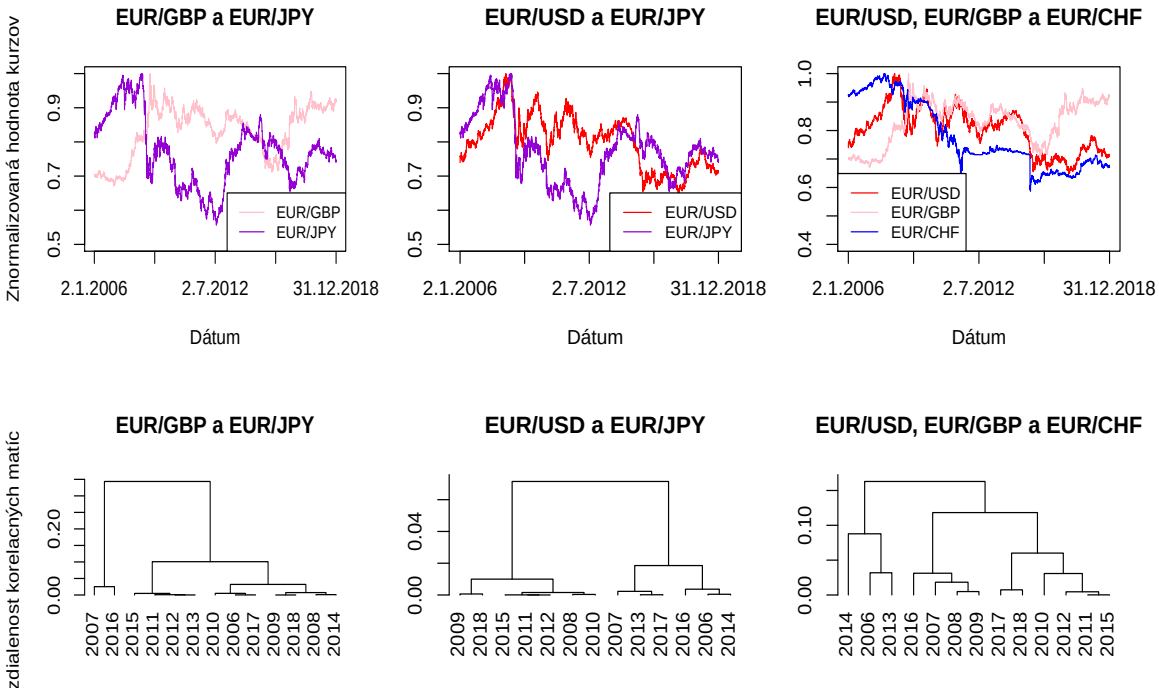
Tabuľka 11: Priebeh výmenných kurzov počas 13 rokov



Obr. 18: Grafická interpretácia zhlukov

Medzi všeobecné dôvody, ktoré ovplyvňujú vývoj jednotlivých kurzov patrí inflácia, úrokové sadzby, pomer importu a exportu, či politická stabilita a hospodárska výkonnosť jednotlivých štátov. Veľký vplyv mali tiež ekonomické udalosti ako Finančná kríza v roku 2008 a Európska dlhová kríza v rokoch 2009 − 2013, najsilnejšie zemetrasenie v Japonsku v roku 2011, či rozhodnutie o Brexite v roku 2016.

Nakoniec uvádzame podobnosť korelačných matíc v priebehu rokov pre vybrané výmenné kurzy. Celkovo boli najväčšie rozdiely v koreláciách medzi kurzami EUR/GBP a EUR/JPY. Naopak, najpodobnejšie korelácie v jednotlivých rokoch boli vypočítané pre kurzy EUR/USD a EUR/JPY. Pri vynechaní jedného zo 4 spomínaných výmenných kurzov sme dostali najmenšie vzdialenosti medzi zhlukmi pre kurzy EUR/USD, EUR/GBP a EUR/CHF.

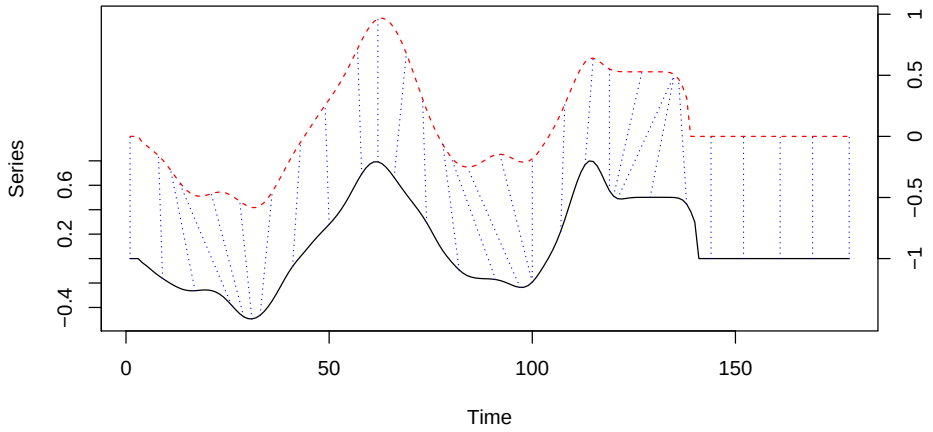


Obr. 19: Podobnosť korelačných matíc pre vybrané kurzy

4.6 Podobnosť časových radov s rozdielnymi dĺžkami

Medzi najpoužívaniejšie vektorové normy patria l_1 a l_2, tiež známe ako Manhattanská a Euklidovská vzdialenosť. Sú ale definované pre časové rady rovnakej dĺžky a taktiež sú "citlivé" na prítomnosť šumu (výkyvy hodnôt) v dátach, mierku a časové posuny. Aj to bol jeden z dôvodov pre vývoj mnohých ďalších metód na meranie vzdialenosti časových radov.

Jednou z takýchto algoritmov je DTW (Dynamic time warping), ktorý porovnáva dva časové rady a hľadá medzi nimi optimálne deformácie pri určitých podmienkach, ako napríklad monotónnosť. Obrázok 20 ilustruje spomínanú deformáciu medzi dvoma časovými radmi. V tomto prípade sa musia navzájom porovnávať počiatočné a konečné hodnoty časových radov, ostatné hodnoty môžu byť spárované ľubovoľne.



Obr. 20: Príklad zosúladenia dvoch časových radov pomocou DTW algoritmu. Bodkované modré čiary ukazujú, ako boli napárované niektoré hodnoty, čo ukazuje, že nemusia byť za rovnaký čas. Poznamenajme, že časový rad zobrazený červenou prerušovanou krivkou bol vertikálne posunutý kvôli vizualizácii. [23]

Predpokladajme dva časové rady x a y dĺžky m a n. Prvý krok DTW zahŕňa vytvorenie matice lcm (local cost matrix) rozmeru $m \times n$. Jednotlivé členy matice obsahujú vzdialenosti každej hodnoty radu x od každej hodnoty radu y, teda $lcm(i,j) = \mid x_i - y_j \mid$.

Následne algoritmus "prechádza" maticou lcm a napáruje hodnoty časových radov tak, aby bola vzdialenosť medzi nimi minimálna. Súčet týchto vzdialeností je výsledná vzdialenosť časových radov. Podrobný opis algoritmu možno nájsť v [23].

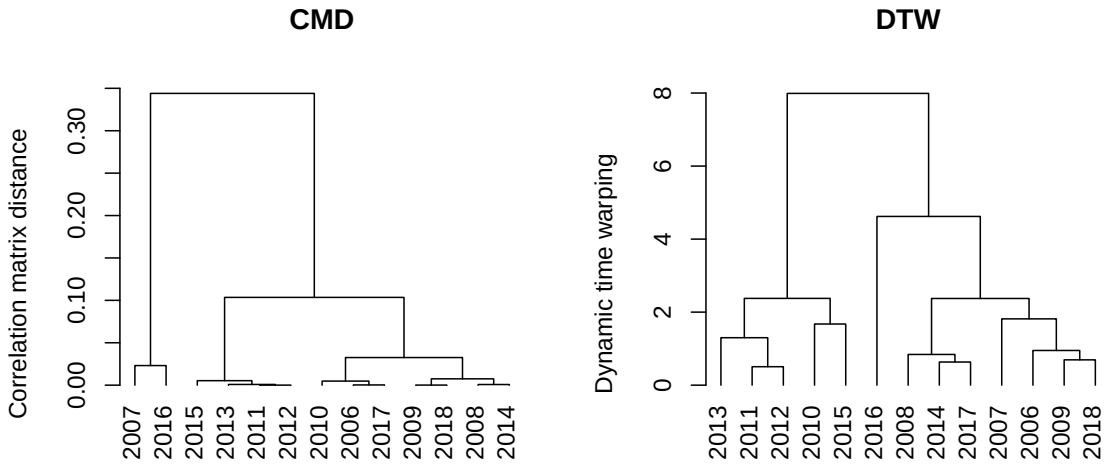
4.6.1 Zhlukovanie časových radov na základe DTW

Znova budeme pracovať s dátami predstavenými vyššie. Pre kurzy EUR/GBP a EUR/JPY sme odhadli dynamickú koreláciu v podkapitole 4.4. Tú rozdelíme podľa rokov a vypočítame vzdialenosti pomocou DTW. Nakoniec roky zaradíme do zhlukov.

Na výpočet vzdialeností časových radov sme použili funkciu `dtw_basic` z knižnice `dtwclust` [23], ktorá je implementáciou DTW algoritmu využívanou najmä pri práci s dvoma časovými radmi. Vstupnými parametrami sú časové rady (v našom prípade výmenné kurzy za jednotlivé roky), miera použitá pri výpočte DTW (Euklidovská norma) a spôsob, akým algoritmus "prechádza" maticou lcm pri určení najmenšej vzdialenosti medzi časovými radmi (`step.pattern = symmetric1`). Alternatívou je "`symmetric2`", nanešťastie ale veľmi málo článkov zameraných na *data mining* uvádza, ktorý zo spôsobov použili. Autor knižnice "zo svojej skúsenosti" považuje za štandard

`symmetric1`. Pri našom porovnaní sme dostávali v tomto prípade zanedbateľne menšie vzdialenosti, konečné rozdelenie zhlukov by to ale pravdepodobne ovplyvnilo až v prípade komplexnejších príkladov.

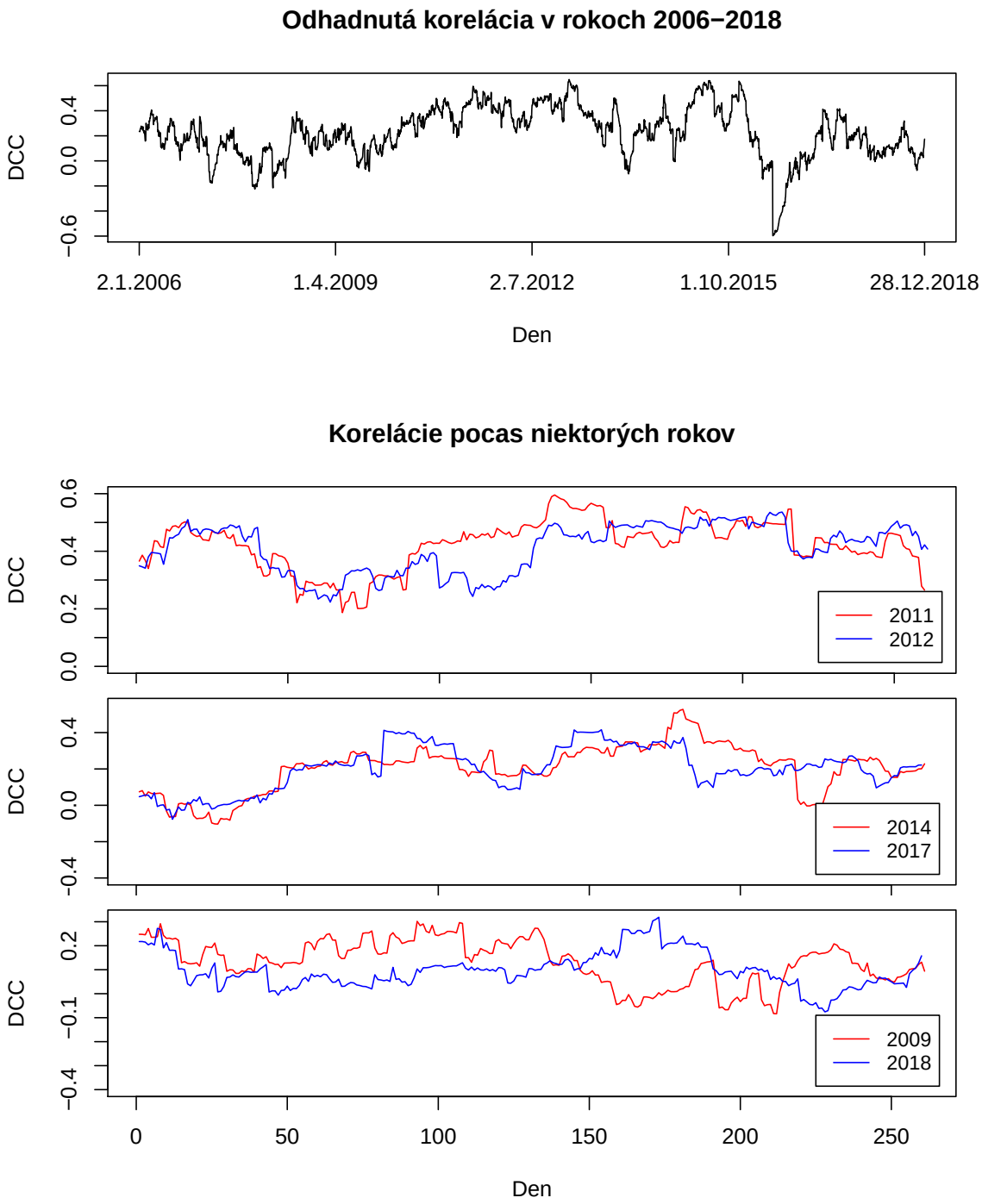
Obrázok 21 porovnáva zhluky, kde bol v jednom prípade pre porovnanie časových radov použitý DTW a v druhom prípade pre porovnanie matíc CMD (opísané v podkapitole 4.5.1).



Obr. 21: Výsledok zhlukovania pri odlišnom porovnávaní korelácií

Pre niektoré roky vidíme značné rozdiely v zaradení do zhlukov, no dá sa povedať, že väčšina je zoradená v podobnej štruktúre. Môžeme ale predpokladať, že zhluky vytvorené pomocou DTW sú presnejšie, nakoľko sme roky nezhlukovali len na základe ich "priemernej"korelácie, ale porovnávali sme celý jej priebeh.

Na obrázku 22 znovu pripomíname odhadnutú koreláciu medzi kurzami EUR/GBP a EUR/JPY a taktiež ilustrujeme vývoj korelácie počas rokov, ktoré boli na základe DTW algoritmu najpodobnejšie, čo do istej miery potvrdzuje správnosť jeho použitia.



Obr. 22: Grafická interpretácia zhlukov

Záver

Cieľom našej práce bolo nadviazať na doterajšie výsledky z okruhu konvergenčných modelov úrokovej miery, pričom budeme predpokladať dynamickú koreláciu. Porovnali sme jej vplyv na výslednú cenu dlhopisu oproti zanedbaní a uvažovaní konštantnej korelácie. Tento rozdiel sme ukázali pre niekoľko konkrétnych prípadov. Ukázalo sa, že vypočítané ceny dlhopisu pre malé maturity sú takmer rovnaké, či už koreláciu zanedbáme alebo uvažujeme dynamickú koreláciu.

Ďalej sme sa snažili zjednodušiť proces riešenia parciálnej diferenciálnej rovnice pomocou štatistického softvéru `R`. Síce sme zatiaľ nedospeli k priamemu výpočtu zadanej PDR, sme schopní efektívne vypočítať systém obyčajných diferenciálnych rovníc získaný z tejto PDR a s tým ďalej pracovať. Využili knižnice dostupné v tomto softvéri. Taktiež sme používali webovú aplikáciou `Wolfram|Alpha` priamo z `R`, pomocou ktorej vieme vyjadriť explicitné riešenie pre rôzne matematické problémy, čo sme využili v našej práci, no môže mať široké využitie aj mimo nej.

Zaoberali sme sa teóriou časových radov, čo sa ukázalo ako dobrý podklad pre algoritmus na odhad dynamickej korelácie. Využili sme DCC-GARCH model, ktorý bol prvýkrát spomenutý v práci Engleho a Steppardov práce z roku 2001 [9]. Celý postup uvedený v kapitole 3 a následný modelový príklad poukazuje na funkčnosť tohto prístupu, čo sa ukázalo aj pri ďalších aplikáciách na reálne dáta.

Algoritmus sme ďalej využili k odhadu korelácie medzi výmennými kurzami eura voči niektorým svetovým menám počas poslednej dekády, kde sme taktiež ukázali možnosť modelovania korelácie medzi viacerými časovými radmi súčasne. Takto odhadnuté dynamické korelácie vo forme korelačných matíc sme ilustrovali pomocou animácie zloženej z korelogramov.

Nakoniec sme využili aglomeratívne zhlukovanie k zgrupovaniu rokov na základe podobnosti korelácie. K určeniu tejto podobnosti sme využili vzdialenosť korelačných matíc a Dynamic time warping algoritmus, ktorý porovnáva časové rady s rôznou dĺžkou. Ako komplexnejší prístup k takémuto porovnávaniu sa ukázal DTW algoritmus, ktorý zohľadňuje celý priebeh časových radov.

Zoznam použitej literatúry

[1] Baumöhl, E., Farkašovská, M., Výrost, T. : *Integrácia akciových trhov: DCC MV-GARCH model*, Politická ekonómia vol.4 (s.488-503), University of Economics, Prague, 2010

[2] Berg, N., Gu, A. Y., Lien, D. : *Dynamic Correlation:: A Tool for Hedging House Price Risk?*, The Journal of Real Estate Portfolio Management, Vol.13, No.1 (s.17-28), 2007

[3] Brockwell, P .J., Davis, R. A. : *Introduction to Time Series and Forecasting*, Springer, 2002

[4] Caporin, M., McAleer, M. : *Ten Things You Should Know about the Dynamic Conditional Correlation Representation*, KIER Working Papers 870, Kyoto University, Institute of Economic Research, 2013

[5] Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B. : *An Empirical Comparison of Alternative Models of the Short-Term Interest Rate*, The Journal of Finance vol. 47, s. 1209-1227

[6] Corzo, T. S., Schwartz, E. S. : *Convergence within the EU: Evidence from Interest Rates*, Economic Notes, vol. 29, 243-266

[7] Cowpertwait, P. S. P., Metcalfe, A. V. : *Introductory Time Series with R*, Springer, 2008

[8] Cox, J. C., Ingersoll, J. E., Ross, S. A. : *A Theory of the Term Structure of Interest Rates*, Econometrica vol. 53, s. 385-407

[9] Engle, R. F.,Sheppard, K. : *Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH*, NBER Working Paper, Cambridge, October 2001

[10] Everitt, B. S., Landau, S., Leese, M., Stahl, D. : *Cluster Analysis, 5th Edition*, Wiley Series in Probability and Statistics, 2011

[11] Lance, G. N., Williams, W. T. : *A General Theory of Classificatory Sorting Strategies: 1. Hierarchical Systems*, The Computer Journal, vol. 9, s. 373–380, 1967

[12] Leveque, R. J. : *Finite Difference Methods for Ordinary and Partial Differential Equations:Steady-State and Time-Dependent Problems*, Society for Industrial and Applied Mathematics, 2007

[13] Matteson, D. S., Ruppert, D. : *Garch models of dynamic volatility and correlation*, IEEE Signal Processing Magazine (s.72-82), New York, 2011

[14] Mikula, K., Stehlíková, B., Ševčovič, D. : *Analytické a numerické metódy oceňovania finančných derivátov*, Nakladateľstvo STU, 2009, Bratislava

[15] Nguyen, T. D., Schmidt, B., Kwoh, CH. K. : *SparseHC: A Memory-efficient Online Hierarchical Clustering Algorithm*, Procedia Computer Science, vol. 29, s. 8-19, 2014

[16] Nogová, J. : *Úrokové sazby jako indikátor vývoje cen akcií*, Diplomová práce, VŠB-TUO, 2015

[17] Stehlíková, B., Bučková, Z. : *Dynamic correlation in a convergence model of interest rates*, Preprint, FMFI UK, 2018

[18] Vasicek, O. A. : *An equilibrium characterization of the term structure*, Journal of Financial Exonomics, vol. 5, s. 177-188

[19] Zíková, Z. : *Konvergenčné modely úrokových mier*, Diplomová práce, FMFI UK, 2011

[20] Ghalanos, A. : *rmgarch: Multivariate GARCH models*, R package version v1.3-5, 2018

[21] Ghalanos, A. : *rugarch: Univariate GARCH models*, R package version v1.4-0, 2018

[22] Kassambara, A. : *ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'*, R package version 0.1.2, 2018

[23] Sarda-Espinosa, A. : *dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance*, R package version 5.5.2, 2019

[24] Stoffer, D. : *astsa: Applied Statistical Time Series Analysis.* R package version 1.8, https://CRAN.R-project.org/package=astsa, 2017

[25] Stoffer, D. : *urca: Analysis of Integrated and Cointegrated Time Series with R*, Second Edition. Springer, New York, ISBN 0-387-27960-1, 2008

[26] https://gist.github.com/sebkopf/a8bf2033f3eda527fa61

[27] https://moneytransfercomparison.com/major-economical-events-which-impacted-currency-rates/

[28] http://www.global-view.com

[29] https://www.investopedia.com/trading/factors-influence-exchange-rates/

[30] https://www.researchgate.net/publication/4194743_Correlation_matrix_distance_a_meaningful_measure_for_evaluation_of_non-stationary_MIMO_channels

[31] https://www.wolframalpha.com

A Príloha

A.1 Wolfram Alpha API prostredníctvom R

Nasledujúci kód poskytuje jednoduché základné rozhranie pre Wolfram Alpha API:

```
library(XML)
library(plyr)
WolframClient <- setRefClass(
  "WolframClient",
  fields = list(
    app_id = "character"
  ),
  methods = list(
    initialize = function(app_id) {
      "initialize Wolfram Alpha Client"
      app_id <<- app_id
    },
    query = function(query, quiet = TRUE) {
      "run a query against the wolfram alpha database"
      message("Querying wolfram alpha...")
      wa_url <- paste0("http://api.wolframalpha.com/v2/query?input=",
                       URLencode(query), "&appid=", app_id)
      destfile <- tempfile()
      download.file(wa_url, destfile,
                    method = "libcurl",
                    quiet = quiet)
      result <- WolframResult(xmlInternalTreeParse(destfile))
      info <- result$get_info()
      if (info$success != "true" || info$error != "false")
        stop("There was a problem with this query. Please check the query infos:\n",
             paste(paste(names(info), info, sep = " = "), collapse = "\n"))
      return(result)
    }
  )
)

WolframResult <- setRefClass(
  "WolframResult",
  fields = list(
    doc = "XMLInternalDocument"
  ),
  methods = list(
    initialize = function(doc) {
      doc <<- doc
    },
```

```
    xml = function() doc,
    get_info = function() {
      as.list(xpathApply(doc, "//queryresult", xmlAttrs)[[1]])
    },
    get_pods = function() {
      xpathApply(doc, "//pod")
    },
    get_plaintext = function() {
      ldply(get_pods(), function(pod) {
        data.frame(
          pod = xmlGetAttr(pod, "title"),
          id = xmlGetAttr(pod, "id"),
          plaintext = unlist(xpathApply(pod, "subpod/plaintext", xmlValue))
        )
      })
    }
  )
)
```

A.1.1 Inicializácia klienta

Tu sa vyžaduje app ID, ktorú si používateľ môže vygenerovať po prihlásení na webstránku *https://www.wolframalpha.com*. Uvádzame len jeho formát.

```
app_id <- "0XXXXX-XXXXXXX000"
wa <- WolframClient(app_id)
result <- wa$query("pi")
```