

VYUŽITÍ WAVELETŮ PŘI ANALÝZE ČASOVÝCH ŘAD

1. TEORETICKÁ ČÁST

USING WAVELETS BY TIME SERIES ANALYSIS

1. THEORETICAL PART

Vratislava Mošová

Moravská vysoká škola Olomouc, Ústav informatiky a aplikované matematiky
vratislava.mosova@mvsso.cz

Abstrakt:

Wavelety jsou moderní matematický prostředek, který se využívá především při zpracování nejrůznějších signálů. Waveletová transformace umožňuje oddělit od sebe aproximace a detaily signálu a pak prostřednictvím prahování očistit signál od šumu popř. signál komprimovat. Schopnost odlišit od sebe aproximace a detaily lze využít i při analýze časových řad. Při waveletovém rozkladu časové řady aproximace korespondují se systematickou složkou a detaily zase s n systematickou složkou řady. V kombinaci s Boxovými-Jenkinsovými modely waveletová transformace představuje alternativní nástroj k analýze a prognózování časových řad. Článek je zaměřen na teoretickou základnu této problematiky.

Klíčová slova:

Časové řady, Boxovy–Jenkinsovy modely, wavelety, víceúrovňová analýza, waveletové koeficienty.

Abstract:

Wavelets are a recent mathematical tool that is used in the signal processing. Wavelet transformation gives possibility to separate approximations and details of a signal and then by using thresholding to clear noise from the signal or to compress the signal. It is possible to use the ability to distinguish approximations from details by analysis of the time series. The wavelet approximations correspond to the systematic part of the series and the details correspond to the non systematic part of the series by signal processing. The wavelet transformation in combination with Box-Jenkins models represents an alternative tool for analysis and prediction in time series. The paper presents the theoretical basis of the corresponding issues.

Keywords:

Time series, Box–Jenkins models, wavelets, multiresolution analysis, wavelet coefficients.

1 Úvod

Analýza časových řad je pro ekonoma důležitou činností, protože umožňuje odhalit zákonitosti, které fungují v dané řadě. Pokud se stávající podmínky radikálně nezmění, lze na základě získaných poznatků odhadnout, jakým směrem se bude ubírat další vývoj řady v nejbližším období. Vedle klasických metod tvorby regresních nebo Boxových-Jenkinsových modelů řady lze při zkoumání časových řad využít také Fourierovu a waveletovou transformaci.

Cílem tohoto článku je seznámit čtenáře s teoretickou základnou pro využití waveletové transformace při prognózování časových řad. První část textu se týká časových řad a jejich zpracování prostřednictvím Boxovy–Jenkinsovy metodologie (viz [1], [5]). Druhá část je věnována waveletům a těm

vlastnostem waveletové transformace, které jsou použitelné při práci s časovými řadami (viz [2], [3], [4], [6]).

2 Časové řady

2.1 Časové řady a korelogramy

Časovou řadu můžeme vymezit jako množinu chronologicky uspořádaných hodnot řady určitého ukazatele. Hodnoty časové řady budeme značit

$$y_t, \quad t = 1, 2, \dots, T.$$

Budeme předpokládat, že řada se skládá z nesystematické náhodné složky I a ze systematických složek: trendu T (dlouhodobé tendence), cyklické složky C (opakuje se v intervalu delším než jeden rok), sezónní složky S (opakuje se v intervalu kratším než jeden rok). V případě, že variabilita hodnot časové řady je zhruba konstantní v čase, můžeme psát

$$y_t = T_t + C_t + S_t + I_t.$$

Z různých typů grafického znázornění časové řady si lze utvořit hrubou představu o jednotlivých charakteristikách řady i o jejím vývojovém trendu. K zevrubnějšímu studiu pak slouží metody umožňující sestavit matematický model, který odpovídá dané řadě.

Jedním ze způsobů, jak takový model vytvořit, je využití Boxovy-Jenkinsovy metodologie. Boxovy-Jenkinsovy modely se konstruují pro stacionární procesy, tj. takové časové řady, které mají pro všechna t konstantní střední hodnotu i konstantní rozptyl a jejichž korelační i kovarianční funkce jsou závislé pouze na časové vzdálenosti náhodných veličin.

Konstruované modely se opírají o informace získané z korelogramů tj. grafů reziduální autokorelační funkce (ACF) a parciální autokorelační funkce (PACF) se zpožděním k . Autoregresi k -tého řádu definujeme vztahem

$$y_t = \phi_{k1} y_{t-1} + \phi_{k2} y_{t-2} + \dots + \phi_{kk} y_{t-k} + a_t,$$

ve kterém ϕ_{kk} představuje parciální regresní koeficient a a_t je veličina nekorelovaná s hodnotami $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$. Zatímco ACF vypovídá o rozsahu lineární závislosti mezi y_t a y_{t-k} , která může být ovlivněna korelací s veličinami $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$, PACF podává informaci o rozsahu lineární závislosti od vlivu veličin $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$ očištěnou. V grafech ACF se na jednu osu vynášejí zpoždění a na druhou osu hodnoty výběrové autokorelace

$$r_k = \frac{\sum_{t=1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2},$$

pro $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$, $k = 1, 2, \dots, T-1$. V grafech PACF se na vodorovnou osu vynášejí jednotlivá zpoždění a na svislou osu hodnoty výběrové parciální autokorelační funkce

$$f_{k,k} = \frac{r_k - \sum_{j=1}^{k-1} f_{k-1,j} r_{k-1}}{1 - \sum_{j=1}^{k-1} f_{k-1,j} r_j},$$

kde $f_{k,j} = f_{k-1,j} - f_{k,k} f_{k-1,k-j}$, $j = 1, 2, \dots, k-1$.

2.2 Boxovy-Jenkinsovy modely

V roce 1976 Box a Jenkins navrhli nový nástroj pro analýzu časových řad - model ARIMA. Uvedme si jednotlivé základní typy procesů, ze kterých se ARIMA skládá, i typy, které jsou rozšířením modelu ARIMA.

AR(p) - autoregresní model řádu p má podobu

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t,$$

a_t zde zastupuje bílý šum. Pokud B je operátor zpětného posunutí, pro který platí $B^i y_t = y_{t-i}$, je možné předchozí vztah přepsat na tvar

$$\phi_p(B)y_t = a_t, \quad \phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p).$$

Uvažovaný proces je stacionární, jestliže kořeny polynomu $\phi_p(B)$ leží vně jednotkového kruhu. V případě, že střední hodnota procesu $\mu \neq 0$, je proces AR(p) popsán vztahem

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t, \quad c = \mu - \phi_1 \mu - \phi_2 \mu - \dots - \phi_p \mu.$$

Proces AR(p) si pamatuje p kroků zpět, kde byl, a od toho se také odvíjejí další predikované hodnoty časové řady.

MA(q) - proces klouzavých průměrů řádu q je určen diferenční rovnicí

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots + \theta_q a_{t-q}.$$

Při použití operátoru zpětného posunutí má uvedený vztah tvar

$$y_t = \theta_q(B)a_t, \quad \theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q).$$

Proces je invertibilní, když kořeny polynomu $\theta_q(B)$ leží vně jednotkového kruhu. Proces MA(q) si pamatuje, kde byl a také ještě dalších q posledních složek bílého šumu. Předpověď se děje na základě odhadu z bílého šumu.

ARMA(p,q) - kombinovaný proces, který je popsán rovnicí

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}.$$

Pomocí operátoru zpětného posunutí jej lze psát ve formě

$$\phi_p(B)y_t = \theta_q(B)a_t.$$

Proces je stacionární, když kořeny polynomu $\phi_p(B)$ leží vně jednotkového kruhu a invertibilní, když kořeny polynomu $\theta_q(B)$ leží vně jednotkového kruhu. Proces si pamatuje, kde byl, spolu s dalšími p hodnotami časové řady a q složkami bílého šumu. Předpověď je proto kombinací hodnot predikovaných z autoregresní funkce a z odhadů pro bílý šum. Model se používá v případě, kdy řada má tendenci zůstat blízko dlouhodobého průměru, tj. při zpracování takových údajů jako jsou míra nezaměstnanosti, změny v cenovém indexu, úrokové sazby, poměr dluhu k HDP firmy.

Integrovaný proces I - nestacionární proces určený rovnicí

$$y_t = \delta + y_{t-1} + a_t,$$

kde konstanta $\delta \neq 0$ ovlivňuje zrychlení nebo zpomalení procesu. Pro $\delta = 0$ hovoříme o procesu náhodné procházky. Analýza procesu se provádí užitím diferencí $\Delta y_t = y_t - y_{t-1}$. Proces si pamatuje kde je, ale zapomněl, jak se sem dostal, a dále se pohybuje náhodně. Předpověď se realizuje přidáním odhadnuté hodnoty $\hat{\delta}$ ke stávajícímu pozorování. Proces se využívá např. při modelování chování akciových trhů.

ARIMA(p,d,q) - autoregresní integrovaný proces klouzavých průměrů řádu p, d, q . Model tohoto procesu při použití operátoru zpětného posunutí má tvar

$$\phi_p(B)(1-B)^d y_t = \theta_q(B)a_t.$$

K analýze studovaného procesu se přitom využívá diference řádu d . Připomeňme si, že diference jsou definovány vztahy

$$\Delta y_t = y_t - y_{t-1}, \quad t = 2, 3, \dots, T, \dots, \Delta^d y_t = \Delta^{d-1} y_t - \Delta^{d-1} y_{t-1}, \quad t = d, \dots, T.$$

Model ARIMA(p,d,q) se po transformaci pomocí diference řádu d chová jako invertibilní model ARMA(p,q). Proces ARIMA je vhodný aplikovat na hladká, pomalu se měnící data, u nichž se neočekává, že hodnoty zůstanou blízko dlouhodobého průměru. Model se používá při studiu HDP nebo vah indexu spotřebitelských cen.

SARIMA(p,d,q) (P,D,Q) - sezonní integrovaný proces s délkou sezóny s . Při použití operátoru zpětného posunutí lze model vyjádřit ve formě

$$\phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\theta_Q(B^s)a_t.$$

Zde p je řád procesu AR, d je řád prosté diference, q je řád procesu MA, s je délka sezónní periody, P je řád sezónního procesu AR, D je řád sezónní diference a Q je řád sezónního procesu MA.

2.3 Konstrukce modelu

Konstrukce vhodného modelu časové řady sestává ze tří fází: výběru modelu, ověření modelu a stanovení odhadu dalšího vývoje řady. V rámci výběru modelu

- 1) Odhadneme ze spojnicového grafu řady, zda je řada stacionární a jestli obsahuje cyklickou složku.
- 2) Pokud je to nutné, řadu upravíme pomocí vhodné transformace. Např. linearizaci řady a stabilizaci rozptylu je možné realizovat prostřednictvím logaritmické transformace.
- 3) Výskyt cyklické složky včetně její periody ověříme pomocí
 - periodogramu (grafického znázornění rozkladu řady na sinusové vlny s různými frekvencemi). Výrazné vrcholy v periodogramu upozorňují na přítomnost cyklu, jejich počet odpovídá počtu cyklů v řadě. Délku jednotlivých cyklů odhadneme pomocí nejvyšších hodnot v periodogramu;
 - pomocí korelogramů ACF a PACF, pokud se v nich vyskytují vzory, které se periodicky opakují.
- 4) Na nutnost stacionarizovat řadu usuzujeme
 - z grafů výběrové ACF a PACF, když se v prvním zpoždění vyskytují hodnoty blízké jedné a ostatní hodnoty klesají pomalu;
 - pokud jsme zjistili, že proces má sezonní složku, v grafu ACF se objevují vysoké hodnoty v nesezónních frekvencích a v grafu PACF jsou vysoké hodnoty v sezónních frekvencích;
 - z periodogramu, pokud se jeho výrazný vrchol nachází v nulové frekvenci.
 Stacionarizaci řady provedeme prostřednictvím diferencování.
- 5) Rozhodnutí, který z Boxových – Jenkinsových modelů si vybereme, lze podepřít o tvar grafů výběrové ACF a PACF. Pokud ve stacionarizované řadě
 - výběrová ACF má pro $k > p$ exponenciální nebo exponenciálně sinusoidní tvar a pro výběrovou PACF je $\phi_{kk} = 0$ uvažujeme proces typu AR(p);
 - když pro $k > q$ v ACF je $\rho_k = 0$ a PACF je omezená exponenciálním nebo exponenciálně sinusoidním poklesem, rozhodneme se pro model MA(q);

- když pro $q > p$ má ACF od zpoždění $q-p$ exponenciální nebo exponenciálně sinusoidní pokles a pro $p > q$ je PACF od zpoždění $p-q$ omezená exponenciálním nebo exponenciálně sinusoidním poklesem, zvolíme ARMA(p,q).

Při ověřování adekvátnosti modelu ARIMA se k testování autokorelace nesystematické složky používají buď hodnoty výběrové autokorelační funkce

$$r_k = \frac{\sum_t \hat{a}_t \hat{a}_{t-k}}{\sum_t \hat{a}_t^2},$$

(stříškou je označen odhad příslušné veličiny) nebo je možné využít Box-Pearsonova testu s testovacím kritériem

$$Q = T \sum_{k=1}^K \hat{r}_k^2,$$

kteří má $\chi^2(K - p - q)$ rozložení, a testovat hypotézu, že autokorelace nesystematické složky jsou nulové. V případě, že nesystematická složka není autokorelovaná, testují se ještě parametry modelu μ , ϕ_i , θ_i pomocí t-testů

$$t_{\mu_i} = \frac{\hat{\mu}_i}{\hat{S}_{\mu}}, \quad t_{\phi_i} = \frac{\hat{\phi}_i}{\hat{S}_{\phi_i}}, \quad i = 1, 2, \dots, p, \quad t_{\theta_i} = \frac{\hat{\theta}_i}{\hat{S}_{\theta_i}}, \quad i = 1, 2, \dots, q.$$

Po ověření modelu je možné přistoupit ke konstrukci předpovědi pomocí podmíněné střední hodnoty $E(y_{T+h} | y_T, y_{T-1}, y_{T-2}, \dots)$. Konkrétní konstrukci předpovědi pro ARIMA model je možné najít v [1].

K porovnání kvality jednotlivých modelů i kvality predikce lze využít některý z následujících typů chyb:

Průměrná absolutní procentuální chyba

$$MAPE = \frac{1}{T} \sum_{t=1}^T 100 \cdot \frac{|y_t - \hat{y}_t|}{y_t}$$

průměrná procentuální chyba

$$MPE = \frac{1}{T} \sum_{t=1}^T 100 \cdot \frac{(y_t - \hat{y}_t)}{y_t}$$

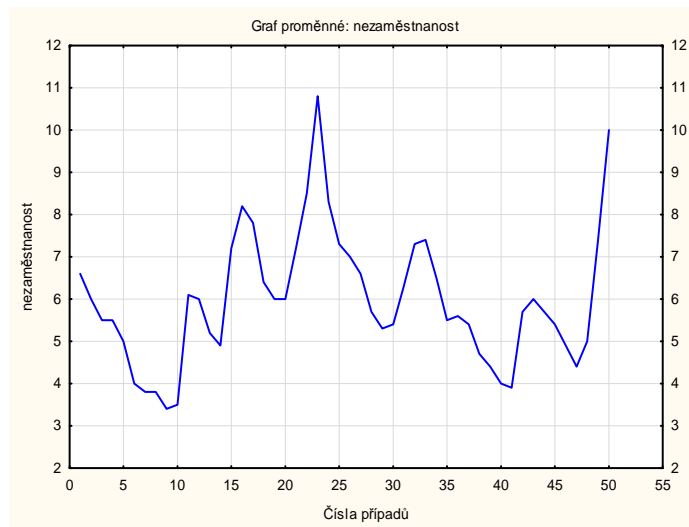
nebo průměrná čtvercová chyba

$$MSE = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2.$$

2.1 Příklad

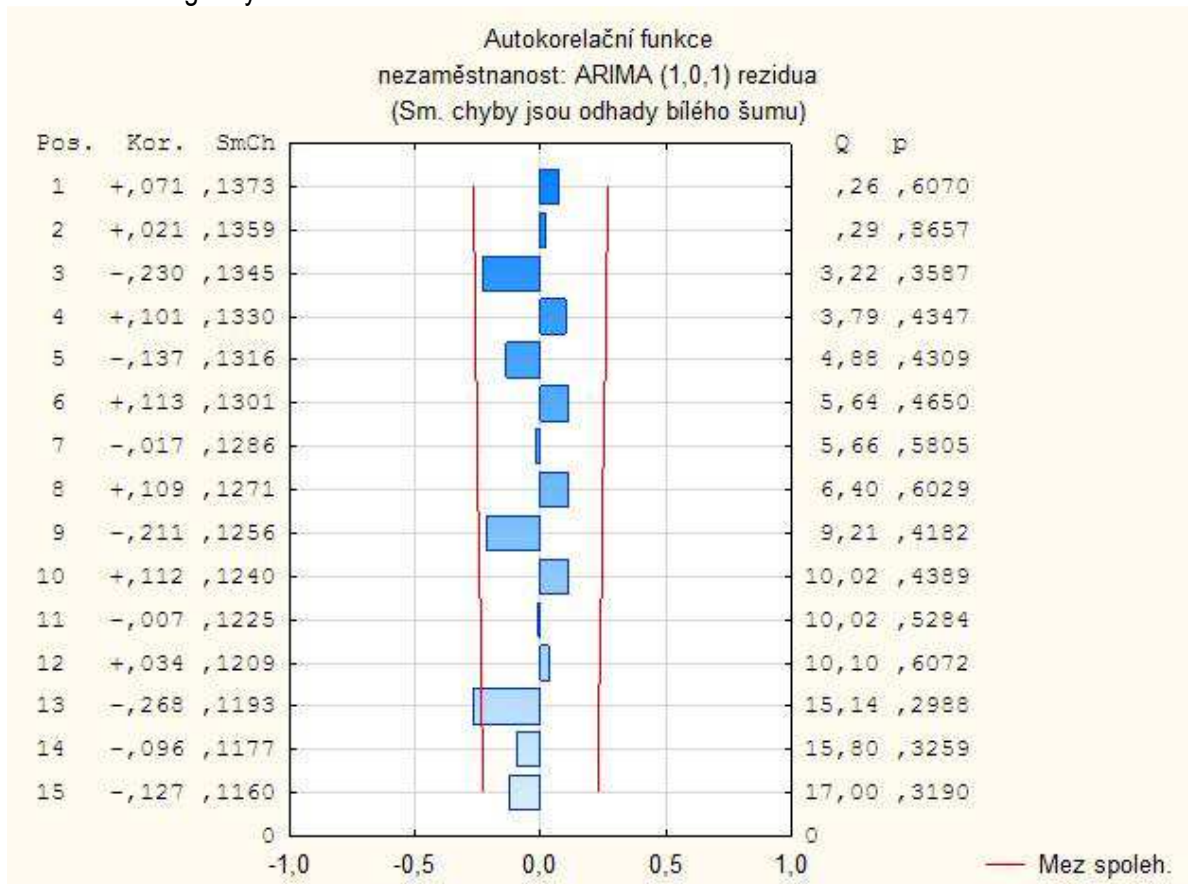
Na následujících obrázcích jsou zachyceny některé grafické výstupy z programu Statistica, které byly získány při analýze časové řady reprezentující vývoj míry nezaměstnanosti v USA v letech 1960 až 2009. Na Obr. 2.1 je spojnicový graf této řady

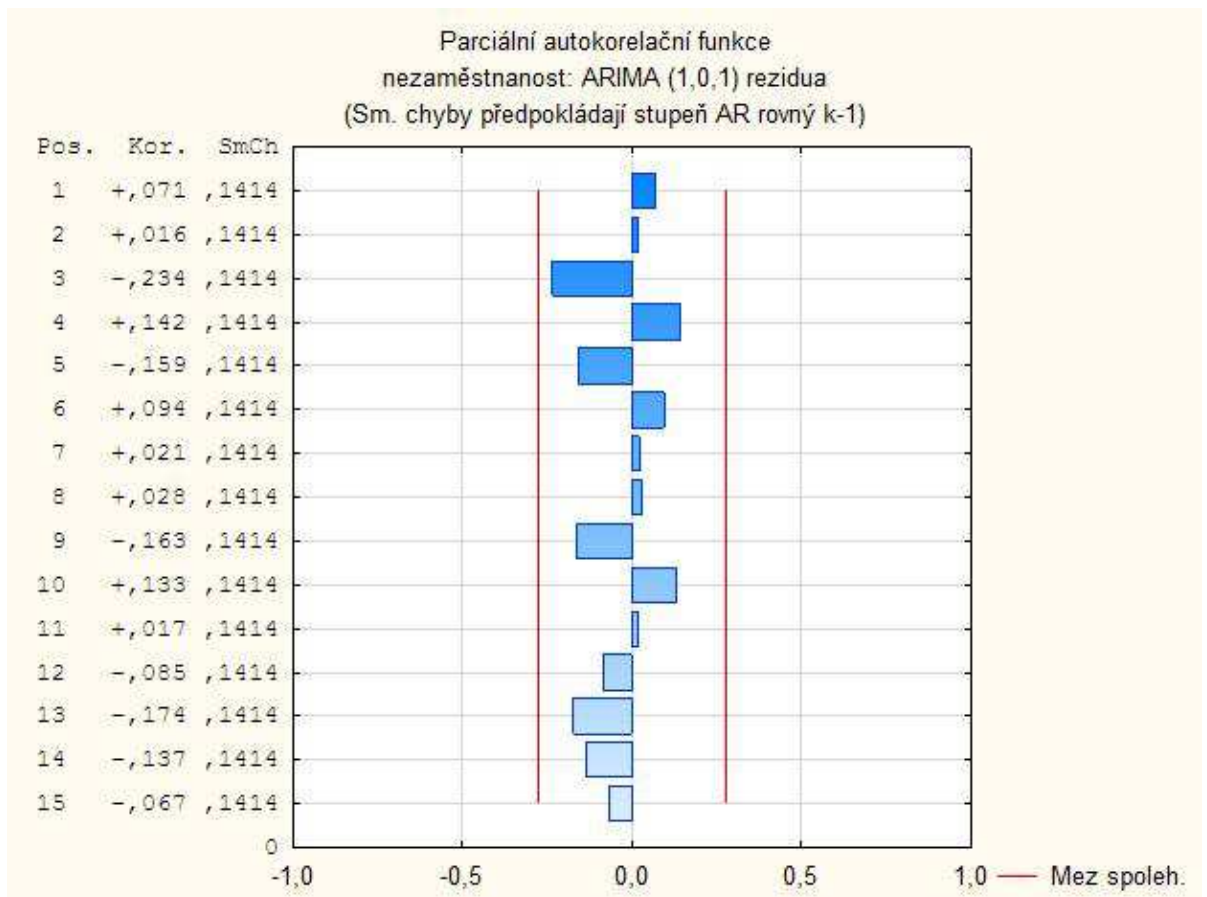
Obr. 2.1 Spojnicový graf



Na Obr. 2.2. jsou zachyceny korelogramy příslušné k modelu ARIMA(1,0,1)

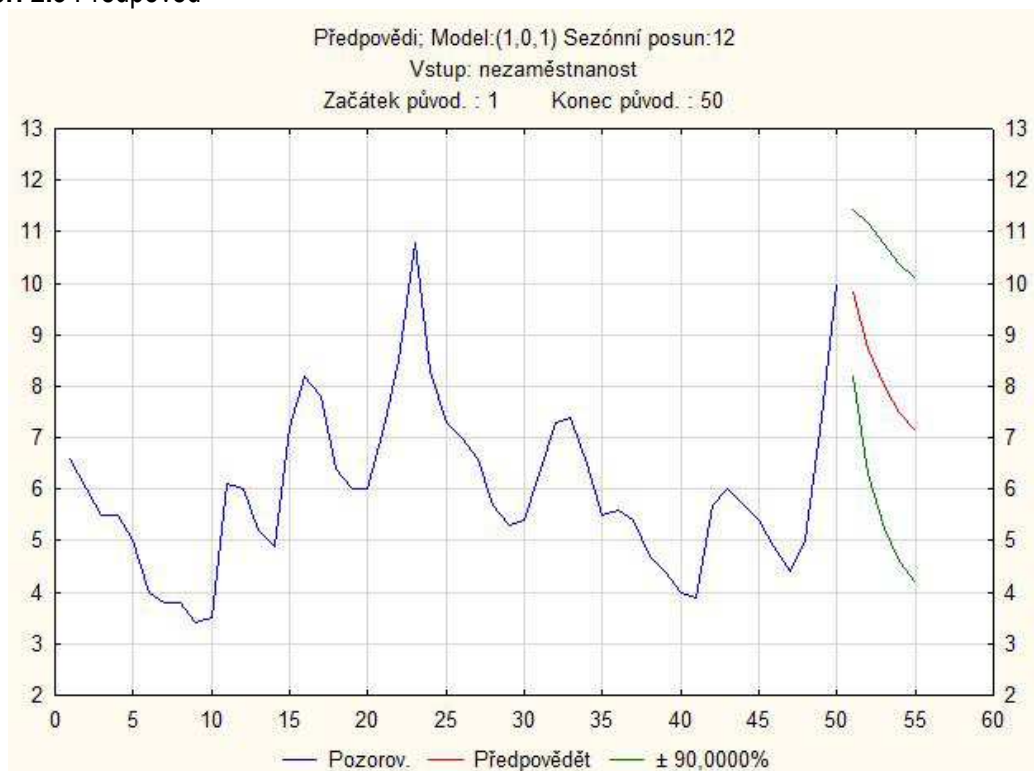
Obr. 2.2 Korelogramy





Předpověď (průměrná křivka v koncové části grafu) na dalších 5 let s využitím modelu ARIMA(1,0,1) je zachycena na Obr 2.3.

Obr. 2.3 Předpověď



3 WAVELETY A WAVELETOVÁ TRANSFORMACE

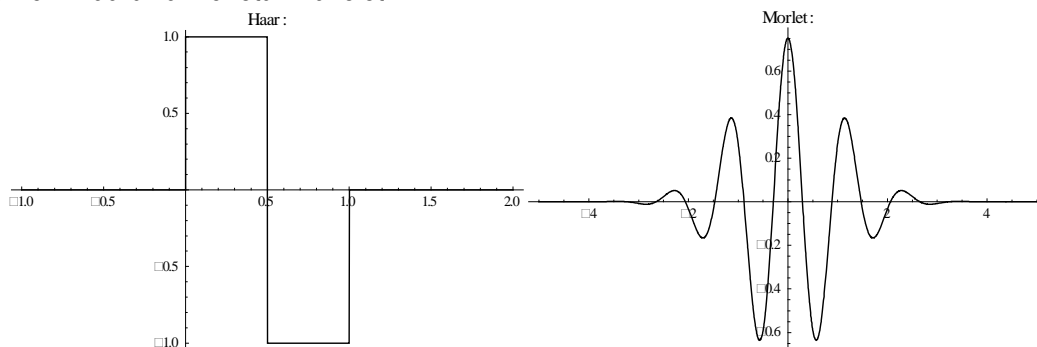
3.1 Co jsou wavelety?

Efektivním nástrojem využívaným při analýze časových řad jsou také integrální transformace. Fourierova transformace umožňuje pomocí periodogramu odhalit sezónnost v časové řadě. Waveletová transformace zase oddělí systematickou složku časové řady od nesystematické.

Waveletovou transformaci lze chápat jako alternativu k transformaci Fourierově. Fourierova transformace se odvíjí od Fourierovy báze $\{e^{inx}\}_{n \in \mathbb{Z}}$, která představuje v prostoru $L^2(0, 2\pi)$ systém dilatovaných a posunutých sinusových vln. V případě waveletové transformace požadujeme, aby báze konstruovaná v $L^2(\mathbb{R})$ měla podobné vlastnosti jako báze Fourierova. Pro tento účel je vhodné použít funkce, kterým se říká wavelety. To proto, že se dají zobrazit jako malé vlnky, které v nekonečnu rychle klesají k nule.

Podnětem ke zkoumání waveletů byly nejenom důvody teoretické (vytvoření báze v prostoru funkcí integrovatelných s kvadrátem), ale i důvody praktické (zpracování signálu). První – Haarův wavelet - byl použit už v roce 1908 s cílem vytvořit ortogonální bázi v prostoru $L^2(\mathbb{R})$. V roce 1984 Morlet za pomoci waveletu, který je součinem funkce s kompaktním nosičem (tzv. okenní funkce) a funkce e^{inx} , zavedl spojitou waveletovou (tzv. okénkovou) transformaci. V roce 1986 zkonstruovala Ingrid Doubechiová třídu waveletů, které mají pro daný nosič maximální možný počet nulových momentů. Tyto wavelety pak využila při realizaci diskrétní waveletové transformace. Na Obr. 3.1 jsou první dva ze zmíněných waveletů zachyceny graficky.

Obr. 3.1 Haarův a Morletův wavelet



3.2 Waveletová transformace

Pro wavelet ψ definujeme waveletovou transformaci funkce $f \in L^2(\mathbb{R})$ vztahem

$$W_\psi(f)(a, b) = \frac{1}{|a|} \int_{\mathbb{R}} f(x) \psi\left(\frac{x-b}{a}\right) dx,$$

kde a je škála (měřítko) a b je translace. Hodnoty $W_\psi(f)(a, b)$ pro pevně daná (a, b) se nazývají waveletové koeficienty.

V případě, že $(a, b) \in \mathbb{R}^2$, hovoříme o spojitě waveletové transformaci. Když a a b jsou diskrétní hodnoty, mluvíme o diskrétní waveletové transformaci. Z důvodu efektivity výpočtu je zvykem položit $a = 2^j$, $b = k$, pro $j, k \in \mathbb{Z}$. Waveletová transformace má pak tvar

$$W_{j,k} = 2^{\frac{j}{2}} \frac{1}{|a|} \int_{\mathbb{R}} f(x) \psi(2^j x - k) dx.$$

Diskrétní rekonstrukce funkce se realizuje z těchto koeficientů pomocí diskrétní inverzní waveletové transformace

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} 2^{\frac{j}{2}} W_{j,k} \psi(2^j x - k).$$

3.3 Multirozkad

Standardizovaný systém $\{2^{\frac{j}{2}}(2^j x - k)\}_{j,k \in \mathbb{Z}}$ ovšem nemusí být pro obecnou funkci ψ ortonormální. Jednou z možností, jak získat ortonormální bázi v $L^2(\mathbb{R})$, je realizace víceúrovňového multirozkladu (MRA). Tento způsob konstrukce waveletů, který v roce 1988 navrhl francouzský matematik Mallat, je postaven na konstrukci takových prostorů $V_j \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$, pro které platí

$$\begin{aligned} V_j &\subset V_{j+1}, \\ \bigcap_{j \in \mathbb{Z}} V_j &= \{0\}, \\ \bigcup_{j \in \mathbb{Z}} V_j &= L^2(\mathbb{R}), \end{aligned}$$

existuje $\varphi \in V_0$ tak, že $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$ je úplná ortogonální množina v $L^2(\mathbb{R})$,
 $f \in V_0 \Leftrightarrow f(2^j x) \in V_j$.

Ze způsobu konstrukce prostorů V_j je vidět, že existují podprostory W_j ortogonální k V_j tak, že $V_{j+1} = V_j \oplus W_j$. Také role škálové funkce φ je v MRA podstatná. Pokud V_j je multirozklad se škálová funkcí φ , která splňuje dilatační rovnici

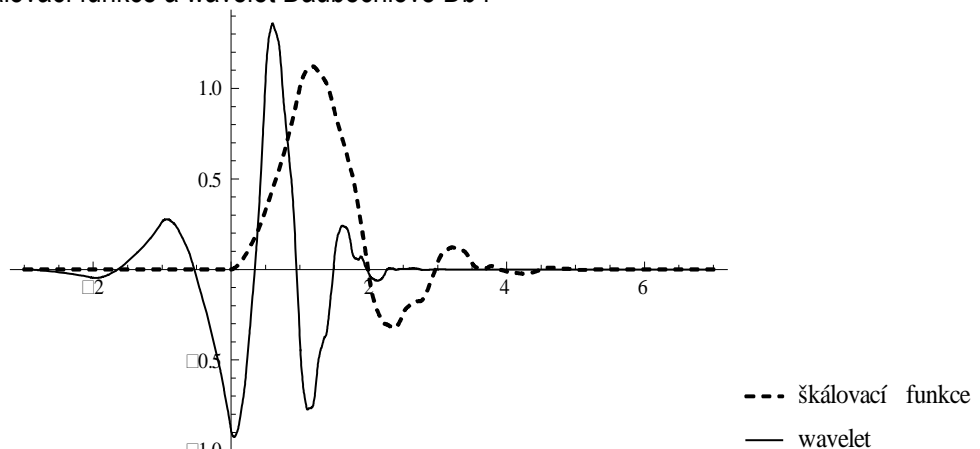
$$\varphi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} a_k \varphi(2^j x - k),$$

pak přidruženým waveletem příslušným k uvažovanému multirozkladu je funkce

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} b_k \varphi(2^j x - k), \quad b_k = (-1)^{k-1} \bar{a}_{1-k}.$$

(Proužek nad a značí komplexně sdružené číslo.) Na Obr. 3.2 je zobrazen wavelet Daubechiové spolu s příslušnou škálovou funkcí.

Obr. 3.2 Škálovací funkce a wavelet Daubechiové Db4



Pomocí dilatací a translací škálových funkcí a přidružených waveletů lze generovat prostory V_j resp. W_j . Definujeme

$$V_j = \text{span}\{\varphi_{j,k}\}_{j,k \in \mathbb{Z}}, \text{ kde } \varphi_{j,k}(x) = 2^{-\frac{j}{2}} \varphi(2^j x - k),$$

$$W_j = \text{span}\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}, \text{ kde } \psi_{j,k}(x) = 2^{-\frac{j}{2}} \psi(2^j x - k).$$

Prostor $V_{j+1} = V_0 \oplus W_0 \oplus W_1 \oplus \dots \oplus W_j$ lze interpretovat jako aproximační prostor prostoru $L^2(\mathbb{R})$. To znamená, že libovolnou funkci $f \in L^2(\mathbb{R})$ je možné rozložit a

$$f(x) = \sum_{k \in \mathbb{Z}} a_{0k} \varphi_{0k}(x) + \sum_{i=1}^j \sum_{k \in \mathbb{Z}} b_{ik} \psi_{ik}(x).$$

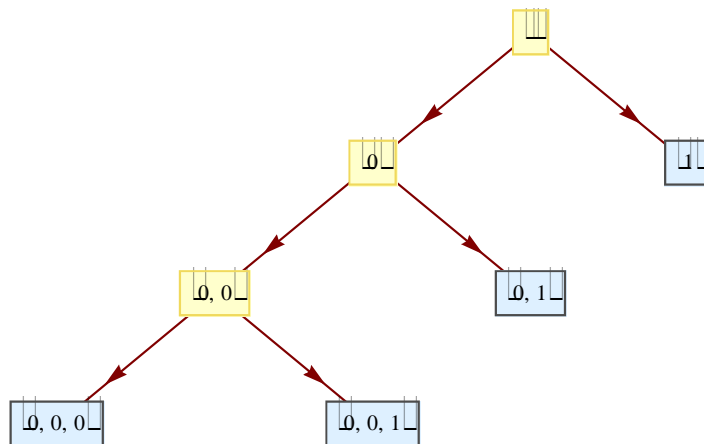
V uvedeném rozkladu jsou $a_{0,k} = \langle f, \varphi_{0,k} \rangle$ aproximační (škálové) koeficienty a $b_{j,k} = \langle f, \psi_{j,k} \rangle$ waveletové koeficienty funkce f na úrovni j a definujeme je prostřednictvím skalárního součinu v $L^2(\mathbb{R})$.

Výpočet koeficientů probíhá na základě Mallatova algoritmu (viz [5]) ve dvou fázích - dekompozice a rekonstrukce. Během procesu dekompozice se spočítají hodnoty aproximačních koeficientů $\{a_{j,k}\}$ a detailních koeficientů $\{b_{j,k}\}$. Aproximační koeficienty přísluší nízkým frekvencím a v časových řadách odpovídají trendu. Detailní koeficienty pak patří k vyšším frekvencím, které mohou být interpretovány jako šum.

3.4 Příklad

Pro datový soubor $\{1,2,3,4,5,6,7,8\}$ chceme spočítat koeficienty waveletové transformace realizované prostřednictvím Haarova waveletu. Protože soubor obsahuje 2^3 hodnot, je možné uskutečnit tříúrovňový waveletový rozklad. Schéma rozkladu se dá graficky zachytit ve formě stromu (viz Obr.3.3).

Obr. 3.3 Schéma waveletového rozkladu



Hodnoty aproximačních koeficientů se v tomto případě počítají jako součet dvou sousedních hodnot z předchozí fáze rozkladu, který se vydělí $\sqrt{2}$, a hodnoty detailů se stanoví jako rozdíl dvou sousedních hodnot dělený $\sqrt{2}$. V prostředí softwaru Mathematica obdržíme:

$\{0\} \rightarrow \{2.12132, 4.94975, 7.77817, 10.6066\}$, $\{1\} \rightarrow \{-0.707107, -0.707107, -0.707107, -0.707107\}$,
 $\{0,0\} \rightarrow \{5., 13.\}$, $\{0,1\} \rightarrow \{-2., -2.\}$,
 $\{0,0,0\} \rightarrow \{12.7279\}$, $\{0,0,1\} \rightarrow \{-5.65685\}$.

Získané waveletové koeficienty je možné ještě před rekonstrukcí modifikovat. Cílem je redukovat nebo odstranit nežádoucí nebo nadbytečná data.

- Odstranění nechtěného trendu ze signálu se dosáhne vymazáním složek s nízkou frekvencí tzn. odstraní se aproximační waveletové koeficienty.
- Šum ze signálu zmizí, když se položí rovny nule waveletové koeficienty $b_{i,j}$, které mají menší frekvenci než vybraný práh λ . Vedle tohoto tvrdého prahování je také možné použít měkkého prahování, při kterém se waveletové koeficienty modifikují např. následovně:

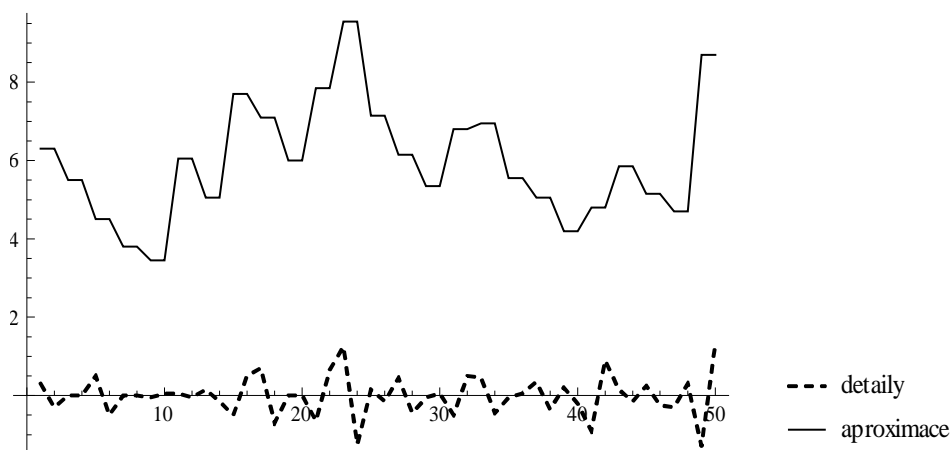
$$\tilde{b}_{j,k} = \begin{cases} 0 & \text{pro } b_{j,k} < \lambda, \\ \text{sgn } b_{j,k} |b_{j,k} - \lambda| & \text{pro } b_{j,k} \geq \lambda. \end{cases}$$

Po provedení příslušných úprav realizujeme zpětnou rekonstrukci pomocí inverzní waveletové transformace aplikované na upravené koeficienty.

3.5 Příklad

Na následujícím obrázku je znázorněn waveletový rozklad datového souboru z Příkladu 2.1.

Obr. 3.4 Rozklad dat na aproximace a detaily



4 ZÁVĚR

Wavelety jsou moderní matematický aparát, který nachází uplatnění při zpracování zvukových a obrazových signálů, v numerické matematice i v rámci analýzy statistických dat. Předností waveletové transformace je její myšlenková jednoduchost, která je blízká osvědčeným Fourierovským přístupům. Navíc Mallatův algoritmus výpočtu waveletových koeficientů umožňuje zrychlení výpočtu pomocí rychlé Fourierovy transformace. Vlastnosti waveletového rozkladu přímo vybízejí k jeho využití při studiu časových řad. Jak uvidíme ve druhé části článku, je waveletová transformace ve spojení s Boxovými-Jenkinsovými modely užitečným nástrojem k predikci dalšího vývoje časové řady.

Poděkování

Tento příspěvek vznikl s finanční podporou a v rámci řešení projektu GAČR P403/12/1811: Vývoj nekonvenčních modelů manažerského rozhodování v podnikové ekonomice a veřejné ekonomii.

LITERATURA

- [1] Artl, J., Artlová, M., Rublíková, E. *Analýza ekonomických časových řad s příklady*. VŠE Praha, 2004. ISBN:978-80-247-3243-5
- [2] Jansen, M., Oonincx P. *Second Generation Wavelets and Applications*. Springer – Verlag London, 2005. ISBN 1-85233-916-0
- [3] Najzar, K. *Základy teorie waveletů*. Karolinum Praha 2004. ISBN 80-246-0957-6.
- [4] Segeth, K *Numerický software I*. Karolinum Praha, 1998. ISBN 80-7184-627-9
- [5] Siegel, A. F. *Practical Business Statistics*. Elsevier, 2011. ISBN 9780123852083,
- [6] Švec, M. *Waveletové transformace*. UJEP Ústí nad Labem, 2008. ISBN 80-246-0957-6.