ANALÝZA ČASOVÝCH ŘAD

Radka Vokřálová

Ústav matematiky a statistiky
Masarykova univerzita, Brno

Seminář z finanční matematiky

Obsah

1. Náhodné procesy
   - Časové řady
   - Popis časových řad
2. Klasická dekompozice časových řad
   - Modelování trendu
3. Box-Jenkinsonova metodologie
   - Odhady v $ARMA$ procesech
   - Autokorelační funkce
4. Spektrální analýza časových řad
   - Spektrální hustota
   - Odhady spektrální hustoty

Náhodné procesy

- Náhodný proces je reálnou funkcí dvou proměnných – elementárního jevu a jedné reálné proměnné, kterou obvykle bývá čas. Značení $\{Y(\omega, t), \omega \in \Omega, t \in T\}$ a zkráceně $\{Y_t, t \in T\}$.
- Náhodný proces, kde indexová množina $T = \mathbb{Z}$ nebo je $T \subset \mathbb{Z}$, se nazývá proces s diskrétním časem nebo náhodná posloupnost nebo také **časová řada**.
- Příklady náhodných procesů v ekonomice
  - Změny poptávky po určitém výrobku
  - Analýza vývoje kurzu akcií na burze
  - Objem zemědělské produkce

Popis časových řad

- Časové řady lze popsat pomocí tzv. aditivního modelu:

$$Y_t = Tr_t + Sz_t + \varepsilon_t \tag{1}$$

  nebo multiplikativního modelu: $Y_t = Tr_t \cdot Sz_t \cdot \varepsilon_t$, který se ovšem může logaritmováním transformovat na aditivní model.
- Jednotlivé složky jsou
  - Tr_t – trend, který odráží dlouhodobé působení vlivů
  - Sz_t – sezónní složka, která popisuje periodické změny, např. vliv ročních období
  - ε_t – náhodné fluktuace (kolísání), které modeluje vlivy, které působí nepravidelně

I. Klasická dekompozice časových řad

- Mezi základní přístupy analýzy časových řad patří klasická dekompozice, která je založena na regresní analýze. Vychází z předpokladu, že náhodný proces, který generuje časovou řadu, je **závislý pouze na čase**.
- Jednotlivá pozorování se obvykle považují za navzájem **nekorelovaná**, tj. $\{\varepsilon_t, t \in \mathbb{Z}\}$ se chápe jako **bílý šum** s nulovou střední hodnotou a varianční maticí $(C(\varepsilon_i, \varepsilon_j))_{i,j=1}^{n} = \sigma^2 \mathbf{I}_n$.
- Dekompozicí rozumíme rozklad na **deterministickou** a **stochastickou** (náhodnou) složku, kde deterministickou složku ještě dále rozkládáme na sezónní složku a trend.

I. Klasická dekompozice časových řad

MODELOVÁNÍ TRENDU

- Trend v časové řadě představuje tendenci dlouhodobého vývoje sledovaného ukazatele v čase. V rámci ekonomického využití časových řad je trend nejdůležitější složkou, která nás zajímá jak z hlediska současného stavu, tak i **predikce budoucího vývoje**.
- Klasická dekompozice trend modeluje pomocí regresních modelů, kde se do trendu obvykle zahrnují i cyklické složky s periodou.
- Ke stanovení trendů lze v závislosti na typu a charakteru náhodného kolísání použít dva různé přístupy:
  1. **Parametrický přístup**, který předpokládá určitý typ rozdělení, obvykle pak normální rozdělení bílého šumu
  2. **Neparametrický přístup**, kam patří různé metody založené na jádrových odhadech

I. Klasická dekompozice časových řad

MODELY GLOBÁLNÍHO TRENDU

- Regresní modely dále můžeme rozdělit na modely globálního trendu a modely postupného lokálního trendu.
- U modelů **globálního trendu**, uvažujeme časovou řadu $\{Y_t, t \in T\}$ a n pozorování této řady v časových okamžicích $t_1 < t_2 < \ldots < t_n$. Předpokládáme, že jednotlivá pozorování v čase t_i vyhovují modelu

$$Y_{t_i} = f(t_i) + \varepsilon_i, \quad i = 1, \ldots, n \tag{2}$$

kde $f(t_i)$ je neznámá **trendová funkce** vybraná z předem dané třídy funkcí, tj. jedná se o parametrický přístup.

MODELY GLOBÁLNÍHO TRENDU

Regresní model trendu

- Předpokládáme, že trendová funkce je určena konečným počtem parametrů
$$f(t) = \beta_0 + \beta_1 \varphi_1(t) + \cdots + \beta_p \varphi_p(t), \tag{3}$$
kde $\beta_0, \beta_1, \ldots, \beta_p$ jsou neznámé parametry a $\varphi_1(t), \ldots, \varphi_p(t)$ jsou známé funkce.
- Takto nám vznikne lineární regresní model tvaru
$$\boxed{\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}}$$
kde matice plánu $\mathbf{X} = ((\varphi_j(t_i))_{i=1,\ldots,n, j=0,\ldots,p}$ je plné hodnosti, tj. $h(\mathbf{X}) = p+1$, funkce $\varphi_0 \equiv 1$.
- Neznámé parametry $\boldsymbol{\beta}$ odhadneme pomocí metody nejmenších čtverců (MNČ), kde platí
$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{4}$$

MODELY GLOBÁLNÍHO TRENDU

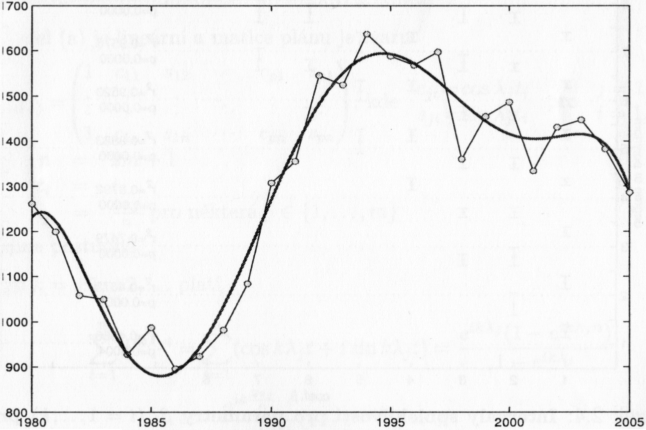
Regresní model trendu

- Příkladem trendové funkce, která vede na lineární regresní model, je tzv. **polynomický trend**

$$f(t) = \beta_0 + \beta_1 t + \cdots + \beta_p t^p \tag{5}$$

  kde kromě neznámých parametrů $\boldsymbol{\beta}$ musíme ještě určit vhodný stupeň polynomu p.
- Pro odhad stupně polynomu můžeme využít různé metody. Například metodu „od nejnižšího stupně k nejvyššímu“, penalizační metodu nebo Akaikeovo informační kritérium. Viz. [1, str. 66]

Ukázka časové řady



Vstupní data spolu s polynomickým trendem 6. řádu.

I. Klasická dekompozice časových řad

MODELY LOKÁLNÍHO POSTUPNÉHO TRENDU

- Hlavní myšlenka **lokální metody** nejmenších čtverců spočívá v tom, že provedeme odhad trendu Tr_t polynomem na lokálním intervalu
$$\langle t-s, t+s \rangle$$
na rozdíl od klasické metody, kdy trend odhadujeme polynomem na celém intervalu možných hodnot parametru t.
- Parametr $s > 0$ se nazývá **šířka vyhlazovacího okénka** a jeho vhodná volba hraje při odhadu důležitou roli.

I. Klasická dekompozice časových řad

MODELY LOKÁLNÍHO POSTUPNÉHO TRENDU

- Lokální metoda nejmenších čtverců se někdy nazývá **klouzavá polynomická metoda**, protože kolem bodu t, v němž má být trend odhadnut, je umístěno vyhlazovací okénko a odhad trendu se mění (pohybuje) spolu s t.
- Uvnitř intervalu aproximujeme trend polynomem stupně m

$$\varrho(x) = \sum_{j=0}^{m} \beta_j(t)(x-t)^j \tag{6}$$

a neznámé koeficienty $\beta_j(t)$ odhadujeme metodou nejmenších čtverců (případně váženou MNČ), tyto koeficienty budou pro každé t jiné.

I. Klasická dekompozice časových řad

MODELY LOKÁLNÍHO POSTUPNÉHO TRENDU

- Různé metody, které jsou založené na lokální metodě nejmenších čverců, jsou například - klouzavé průměry, jednoduché klouzavé průměry, centrované klouzavé průměry
- Dále například **modely exponenciálního vyrovnávání** vychází z lokální **vážené** metody nejmenších čtverců, kde jednotlivým údajům v časové řadě jsou přiřazeny rozdílné váhy, které směrem do minulosti exponenciálně klesají.

II. Box-Jenkinsonova metodologie

- Box-Jenkinsonova metodologie na rozdíl od klasické dekompozice předpokládá, že všechny složky časové řady, tj. trend i cyklická složka, mají **náhodný charakter**. A jejím těžištěm je **korelační analýza**.
- Výhodou této metody je její **flexibilita** a to, že se rychle adaptuje na změnu v charakteru modelovaného procesu.
- Nevýhodou bývá **požadavek na dostatečně dlouhé realizace** časové řady. Ztrácí se také možnost jednoduché interpretace výsledných modelů.

II. Box-Jenkinsonova metodologie

- V dalším předpokládejme, že náhodná posloupnost $\{Y_t, t \in \mathbb{Z}\}$ je **stacionární**, **centrovaná** a **druhého řádu**, tj. je stacionární ve střední hodnotě, která je rovna nule, je kovariančně stacionární, a má konečné druhé momenty.
- Základní modelová schémata:
  - **Autoregresní procesy** $AR(p)$
  - **Procesy klouzavých průměrů** $MA(q)$
  - Případně kombinace předchozích, $ARMA(p, q)$
- $ARMA$ proces řádu p, q je definovám vztahem

$$Y_t - \varphi_1 Y_{t-1} - \cdots - \varphi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}, \qquad (7)$$

kde $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

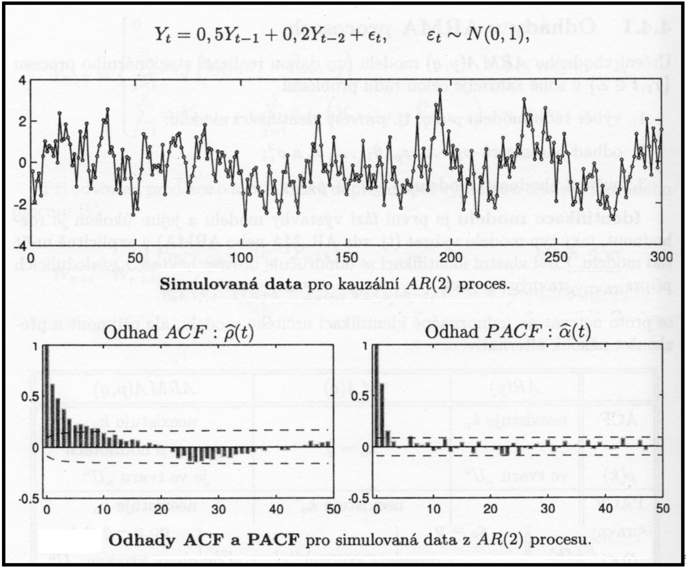
Odhady v $ARMA$ procesech

- Určení vhodného $ARMA(p,q)$ modelu pro danou realizaci stacionárního procesu v sobě zahrnuje
  1. výběr řádu modelu p a q, tj. provedení **identifikace modelu**
  2. odhad parametrů $\varphi_1, \ldots, \varphi_p, \theta_1, \ldots, \theta_q$ a σ_ε^2
  3. ověření vhodnosti modelu.
- Před provedením identifikace modelu se doporučuje vykreslit graf řady a pokud by byla střední hodnota nenulová, tak provést centrování řady.
- Představu o struktuře studovaného procesu získáme na základě zkoumání průběhu odhadnuté **autokorelační funkce** (ACF) a odhadnuté **parciální autokorelační funkce** (PACF).

Autokorelační funkce

- Na základě teoretických výpočtů známe základní vlastnosti ACF a PACF v procesech AR, MA a $ARMA$
- AR proces
  - ACF funkce exponenciálně klesá k nule pro $t \to \infty$ a PACF pro $t > k_0$, kde k_0 je tzv. **identifikační faktor**, se výrazně neliší od nuly
- MA proces
  - Chování ACF a PACF je opačné oproti funkcím v AR procesu
- $ARMA$ proces
  - ACF i PACF exponenciálně klesají k nule pro $t \to \infty$, ovšem nelze určit identifikační bod, od kterého by se některá z korelačních funkcí výražně lišila od nuly.

Příklad $AR(p)$ procesu



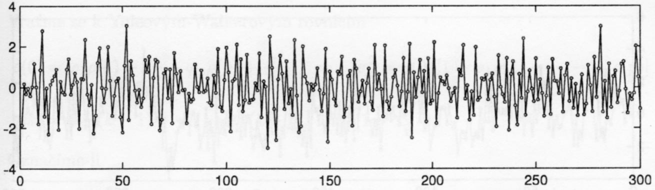
$$Y_t = 0,5Y_{t-1} + 0,2Y_{t-2} + \epsilon_t, \qquad \varepsilon_t \sim N(0,1),$$

Simulovaná data pro kauzální $AR(2)$ proces.

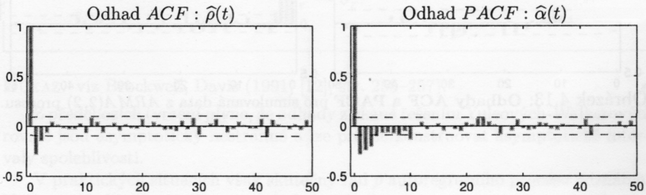
Odhady ACF a PACF pro simulovaná data z $AR(2)$ procesu.

Příklad $MA(q)$ procesu

$$Y_t = \epsilon_t - 0,5\epsilon_{t-1} - 0,2\epsilon_{t-2}, \qquad \varepsilon_t \sim N(0,1),$$

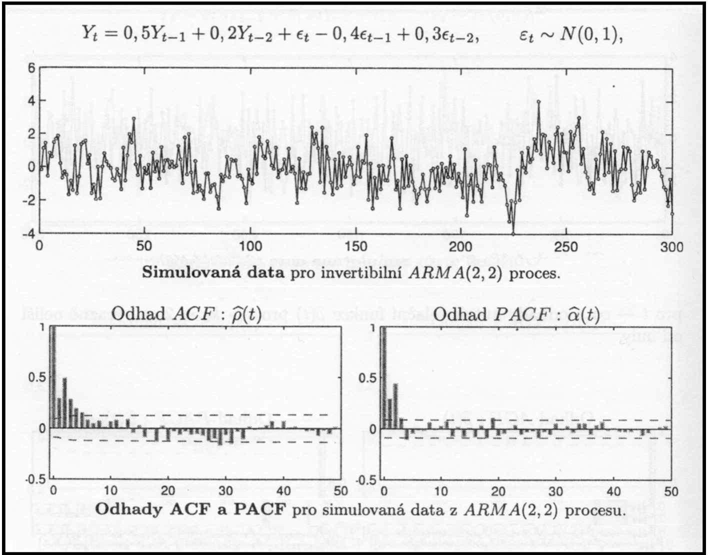


Simulovaná data pro invertibilní $MA(2)$ proces.



Odhady ACF a PACF pro simulovaná data z $MA(2)$ procesu.

Příklad $ARMA(p, q)$ procesu



$$Y_t = 0,5Y_{t-1} + 0,2Y_{t-2} + \epsilon_t - 0,4\epsilon_{t-1} + 0,3\epsilon_{t-2}, \qquad \varepsilon_t \sim N(0,1),$$

Simulovaná data pro invertibilní $ARMA(2,2)$ proces.

Odhady ACF a PACF pro simulovaná data z $ARMA(2,2)$ procesu.

III. Spektrální analýza časových řad

- Klasická dekompozice a Box-Jenkinsonova metodologie jsou základní přístupy k analýze časových řad **v časové doméně.** **Ve spektrální doméně** patří mezi základní přístupy spektrální analýza, která je založena na Fourierově analýze.
- Spektrální analýza časových řad předpokládá, že lze časovou řadu vyjádřit pomocí funkcí $\sin(x)$ a $\cos(x)$ o rozličných amplitudách a frekvencích. Stěžejním faktorem zde není časová proměnná, ale faktor frekvenční.
- Vhodná při srovnávání chování několika řad, kdy můžeme porovnat řady v rámci jednotlivých frekvencí.

III. Spektrální analýza časových řad

- Předpokládejme opět, že náhodná posloupnost $\{Y_t, t \in \mathbb{Z}\}$ je **stacionární**, **centrovaná** a **druhého řádu**, tj. je stacionární ve střední hodnotě, která je rovna nule, je kovariančně stacionární, a má konečné druhé momenty.
- Důležitou vlastností stacionární náhodné posloupnosti je vlastnost, že její **autokovarianční funkci** lze vyjádřit jako nespočetný součet harmonických funkcí s různými amplitudami a frekvencemi.

III. Spektrální analýza časových řad

Spektrální hustota

- Podle Herglotzovy věty [1, str. 28] známe spektrální rozklad autokovarianční funkce stacionární posloupnosti

$$\gamma(t) = \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} \mathrm{d}F(\lambda), \tag{8}$$

kde funkce $F(\lambda)$ se nazývá **spektrální distribuční funkce** a je-li absolutně spojitá, pak existuje taková funkce $f(\lambda)$, že pro náhodné stacionární posloupnosti platí

$$F(\lambda) = \int_{-\pi}^{\lambda} f(x) \mathrm{d}x \tag{9}$$

a funkce $f(\lambda)$ se pak nazývá **spektrální hustota**.

III. Spektrální analýza časových řad

Spektrální hustota

- Existuje-li spektrální hustota, pak můžeme psát

$$\gamma(t) = \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} f(\lambda) d\lambda \tag{10}$$

- Pomocí Fourierovy transformace lze pak vyjádřit spektrální hustotu pomocí autokovarianční funkce

$$f(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} \mathrm{e}^{-it\lambda} \gamma(t) \tag{11}$$

Spektrální hustota

Odhady spektrální hustoty

- PERIODOGRAM
  - Periodogram je funkce definovaná pro n pozorování náhodné posloupnosti vztahem

$$I_n(\omega) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} Y_t \mathrm{e}^{-it\omega}\right|^2 \quad \omega \in \langle -\pi, \pi \rangle \tag{12}$$

  - $I_n(\omega)$ se dá považovat za **empirický odhad spektrální hustoty**, který je asymptoticky nestranným odhadem, ovšem není konzistentním – jeho rozptyl nekonverguje k nule, pokud vzrůstá délka posloupnosti pozorování.
  - Díky svým vlastnostem je užitečným prostředkem pro vyhledávání významných periodických složek v časové řadě.

- [1] FORBELSKÁ, Marie. *Stochastické modelování jednorozměrných časových řad*. Brno: Masarykova univerzita, 2009.
- [2] KRIŠTOF, Aleš. *Nové metody a přístupy k analýze a prognóze ekonomických časových řad*. Praha, 2006. Disertační práce. Česká zemědělská univerzita v Praze.

Děkuji za pozornost.