

Spektrální vlastnosti stacionárních modelů

Luboš Marek*

Cílem tohoto článku je popsat chování spektrální hustoty stacionárních modelů, konkrétně autoregresního modelu AR(p), modelu klouzavých součtů MA(q) a smíšeného modelu ARMA(p, q). Bude ukázán vztah mezi spektrální hustotou a autokorelační funkcí těchto procesů. Hlavní částí práce je grafické znázornění průběhu hustoty pro různé kombinace parametrů výše uvedených procesů.

Spektrální reprezentace časových řad

Periodogram časové řady

Jedním z možných přístupů k analýze časových řad je její analýza ve spektrální doméně. V takovém případě považujeme hodnoty časové řady za směs sinusovek a kosinusovek o různé frekvenci. Nástrojem, pomocí kterého analyzujeme periodické chování časové řady, je periodogram. Za předpokladu lichého počtu pozorování ($N = 2k + 1$) můžeme, časovou řadu vyjádřit pomocí Fourierovy transformace – viz Marek (1999) ve tvaru

$$y_t = \alpha_0 + \sum_{i=1}^k (\alpha_i c_{it} + \beta_i s_{it}) + \varepsilon_t$$

kde

$$c_{it} = \cos(2\pi f_i t) \text{ a } s_{it} = \sin(2\pi f_i t)$$

a

$$f_i = \frac{i}{N}$$

je i -tý násobek základní frekvence $1/N$ (tzv. *harmonické frekvence*). Odhady parametrů α_0 , α_i , β_i získané MNČ mají tvar – viz Box, Jenkins, Reinsel (1994):

$$a_0 = \bar{y}$$

$$a_i = \frac{2}{N} \sum_{t=1}^N y_t c_{it} \quad \text{pro } i = 1, 2, \dots, k$$

$$b_i = \frac{2}{N} \sum_{t=1}^N y_t s_{it}$$

* Doc. RNDr. Luboš Marek, CSc.; Katedra statistiky a pravděpodobnosti, Fakulta informatiky a statistiky, VŠE v Praze, marek@vse.cz

Periodogram pak sestává z $k=(N-1)/2$ hodnot

$$I(f_i) = \frac{N}{2}(a_i^2 + b_i^2), \quad i = 1, 2, \dots, k$$

kde $I(f_i)$ je intenzita na frekvenci f_i . Je třeba podotknout, že různí autoři definují periodogram různě – většinou se liší hodnotou konstanty.

Pro $N = 2k$ a pro $i = 1, 2, \dots, (k-1)$ je

$$a_k = \frac{1}{N} \sum_{i=1}^N (-1)^i y_i$$

$$b_k = 0$$

a

$$I(f_k) = I(0,5) = N a_k^2.$$

Spektrum

Pro $f_i = \frac{i}{N}$ můžeme na intervalu $(0;0,5)$ přepsat periodogram do tvaru

$$I(f) = \frac{N}{2}(a_f^2 + b_f^2)$$

přičemž $I(f)$ nazýváme výběrové spektrum. Lze dokázat – viz Box, Jenkins, Reinsel (1994), že výběrové spektrum a odhad autokovarianční funkce c_k stochastického procesu jsou vázány vztahem

$$I(f) = 2 \left[c_0 + 2 \sum_{k=1}^{N-1} c_k \cos(2\pi f k) \right]$$

To znamená, že výběrové spektrum je Fourierovou kosinovou transformací odhadu autokovarianční funkce, přičemž platí

$$\lim_{N \rightarrow \infty} E(c_k) = \gamma_k$$

kde γ_k je teoretická autokovarianční funkce. Pak definujeme spektrum $p(f)$ jako

$$p(f) = \lim_{N \rightarrow \infty} E[I(f)] = 2 \left[\gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos(2\pi f k) \right] \quad 0 \leq f \leq 0,5$$

Spektrum je tedy Fourierovou kosinovou transformací autokovarianční funkce. Z vyjádření je patrné, že znalost autokovarianční funkce je matematicky ekvivalentní znalosti spektra a naopak.

Spektrální hustota

Spektrální hustota je založena na hodnotě spektra a autokorelační funkce ρ_k

$$g(f) = \frac{p(f)}{\sigma_y^2} = 2 \left[1 + 2 \sum_{k=1}^{\infty} \rho_k \cos(2\pi f k) \right] \quad 0 \leq f \leq 0,5$$

Platí pro ni vztah

$$\int_0^{1/2} g(f) df = 1$$

Odhadem spektra je

$$\hat{p}(f) = 2 \left[c_0 + 2 \sum_{k=1}^{N-1} \lambda_k c_k \cos(2\pi f k) \right]$$

kde λ_k jsou vhodně zvolené váhy.

Spektrum lineárního stacionárního procesu

Spektrum lineárního stacionárního procesu má obecný tvar (Box, Jenkins, Reinsel, 1994)

$$p(f) = 2\sigma_\varepsilon^2 \psi(e^{-i2\pi f}) \psi(e^{i2\pi f}) = 2\sigma_\varepsilon^2 \left| \psi(e^{-i2\pi f}) \right|^2 \quad 0 \leq f \leq 0,5$$

kde $\psi(B)$ je *přenosová funkce* obecného lineárního procesu y_t (ten totiž můžeme vyjádřit jako $y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$). Pak $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$. Konkrétní tvar $\psi(B)$ pak bude záležet na typu procesu, jak uvidíme dále.

Spektrum AR procesu

Pro AR(p) proces s parametry ϕ má spektrum tvar

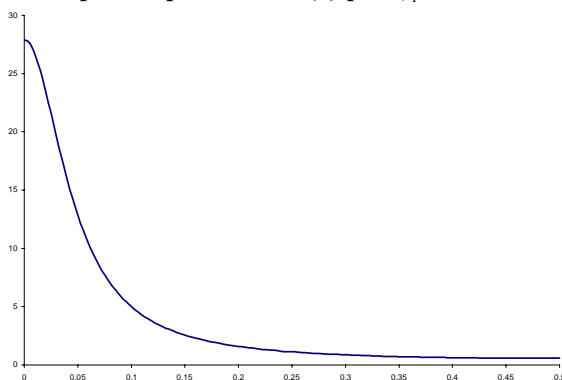
$$p(f) = \frac{2\sigma_\varepsilon^2}{\left| 1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f} - \dots - \phi_p e^{-i2\pi p f} \right|^2} \quad 0 \leq f \leq 0,5$$

V praxi se většinou používají pouze AR procesy do řádu 2. Jejich spektrum má potom konkrétní tvar:

$$\text{AR(1)} \quad p(f) = \frac{2\sigma_\varepsilon^2}{\left| 1 - \phi_1 e^{-i2\pi f} \right|^2} = \frac{2\sigma_\varepsilon^2}{1 + \phi_1^2 - 2\phi_1 \cos(2\pi f)} \quad 0 \leq f \leq 0,5$$

Vyjádříme-li grafický průběh spektra pro proces AR(1) pro $\phi_1 = 0,75$, obdržíme:

Obr. č. 1: Průběh spektra procesu AR(1) pro $\phi_1 = 0,75$ – simulovaná data



Z obrázku je zřejmé, že pro $\phi_1 = 0,75$ (sousední pozorování jsou kladně korelovaná), jsou v procesu AR(1) s největší intenzitou zastoupeny právě periodické složky o nižších frekvencích. Pokud by byl parametr ϕ_1 záporný, převládaly by naopak složky o frekvencích vyšších, což je ostatně patrné z přehledu průběhů spektra v závěru tohoto článku.

$$\begin{aligned} \text{AR(2)} \quad p(f) &= \frac{2\sigma_\varepsilon^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f}\right|^2} = \\ &= \frac{2\sigma_\varepsilon^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos(2\pi f) - 2\phi_2\cos(4\pi f)} \quad 0 \leq f \leq 0,5 \end{aligned}$$

Spektrum MA procesu

Pro MA(q) proces s parametry θ má spektrum tvar

$$\begin{aligned} p(f) &= 2\sigma_\varepsilon^2 \left|1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f} - \dots - \theta_q e^{-i2\pi qf}\right|^2 \quad 0 \leq f \leq 0,5 \\ \text{MA(1)} \quad p(f) &= 2\sigma_\varepsilon^2 \left|1 - \theta_1 e^{-i2\pi f}\right|^2 = 2\sigma_\varepsilon^2 [1 + \theta_1^2 - 2\theta_1 \cos(2\pi f)] \quad 0 \leq f \leq 0,5 \\ \text{MA(2)} \quad p(f) &= 2\sigma_\varepsilon^2 \left|1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f}\right|^2 = \\ &= 2\sigma_\varepsilon^2 [1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos(2\pi f) - 2\theta_2\cos(4\pi f)] \quad 0 \leq f \leq 0,5 \end{aligned}$$

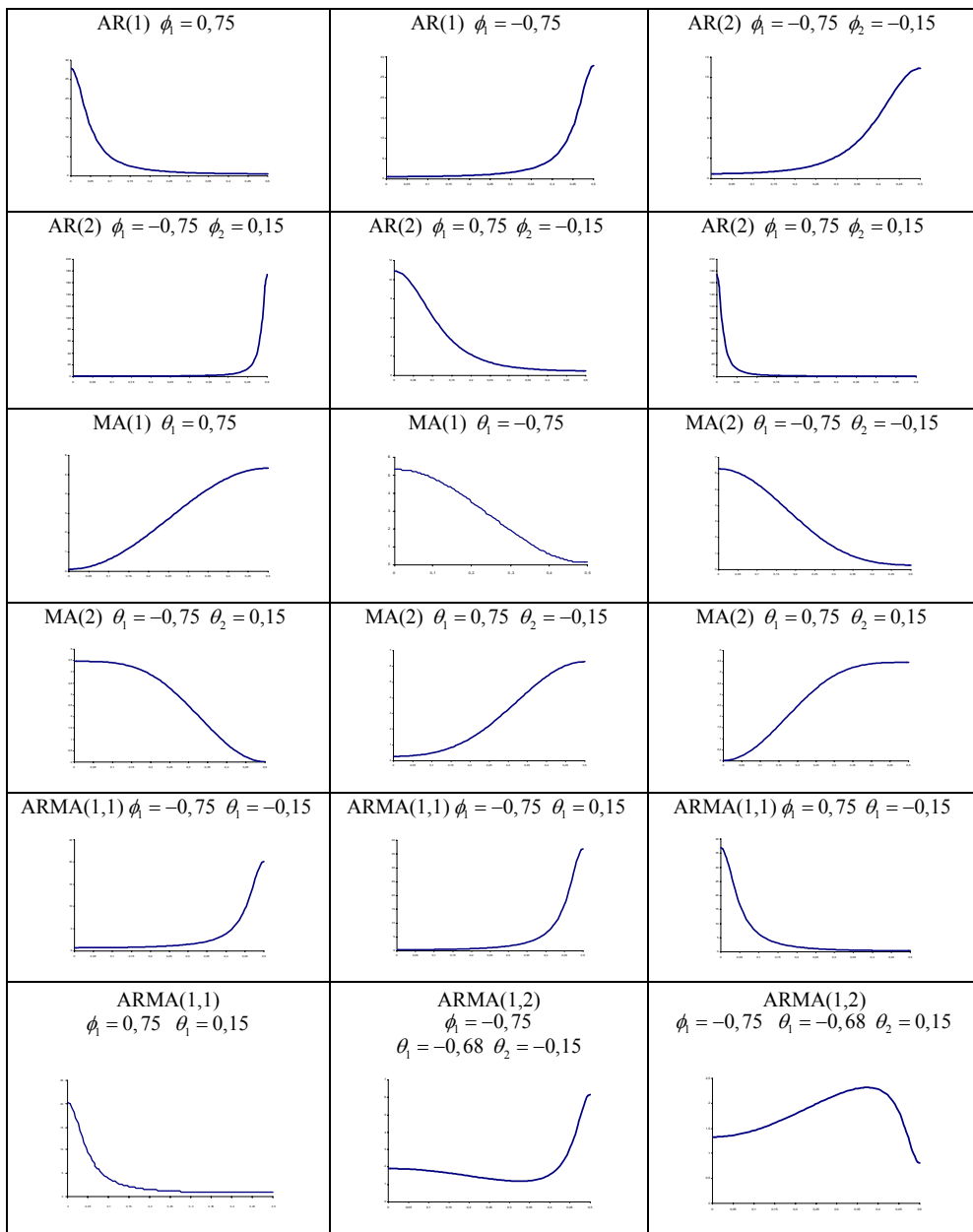
Spektrum ARMA procesu

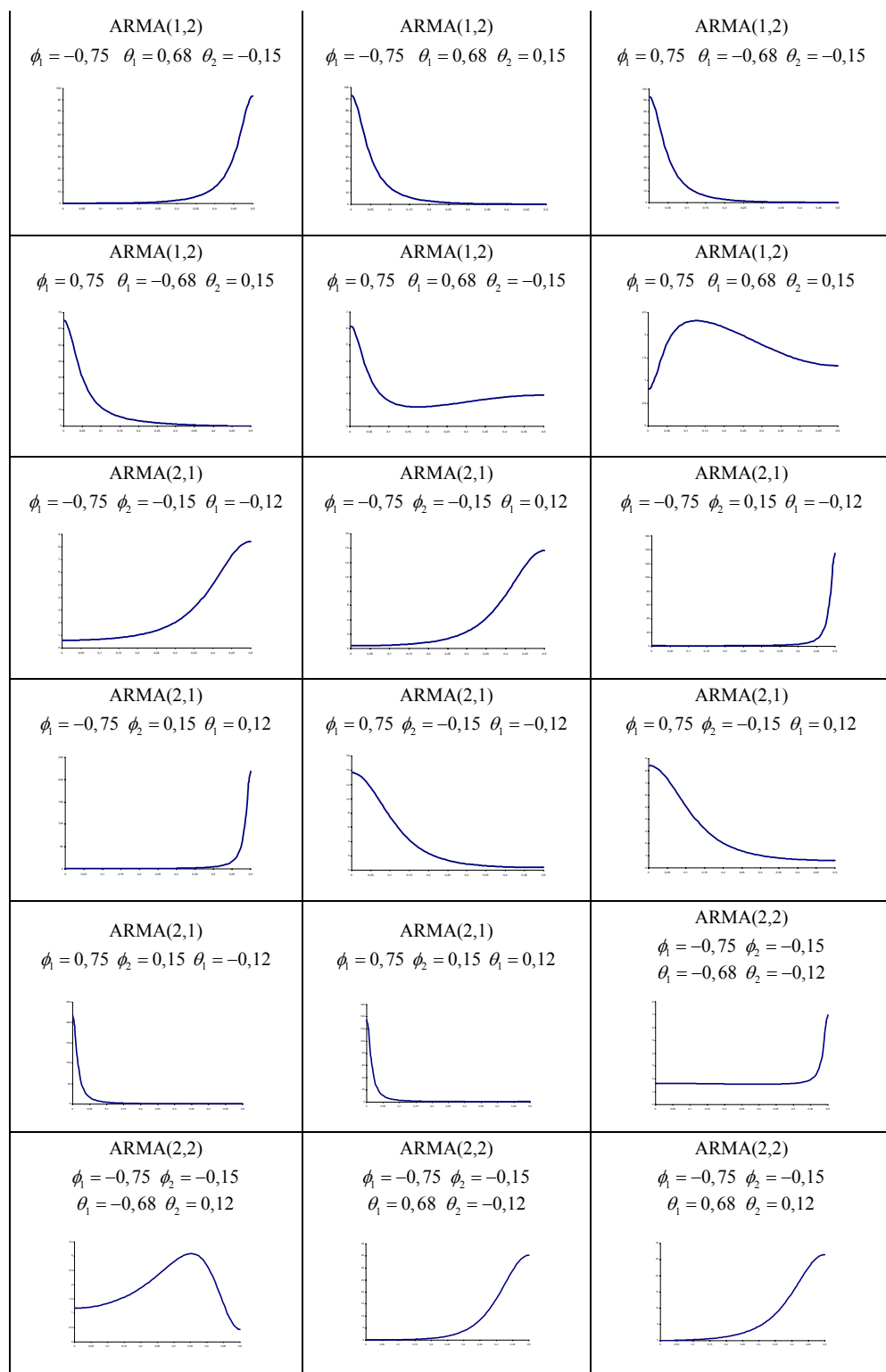
Pro ARMA(p,q) proces s parametry ϕ a θ má spektrum tvar

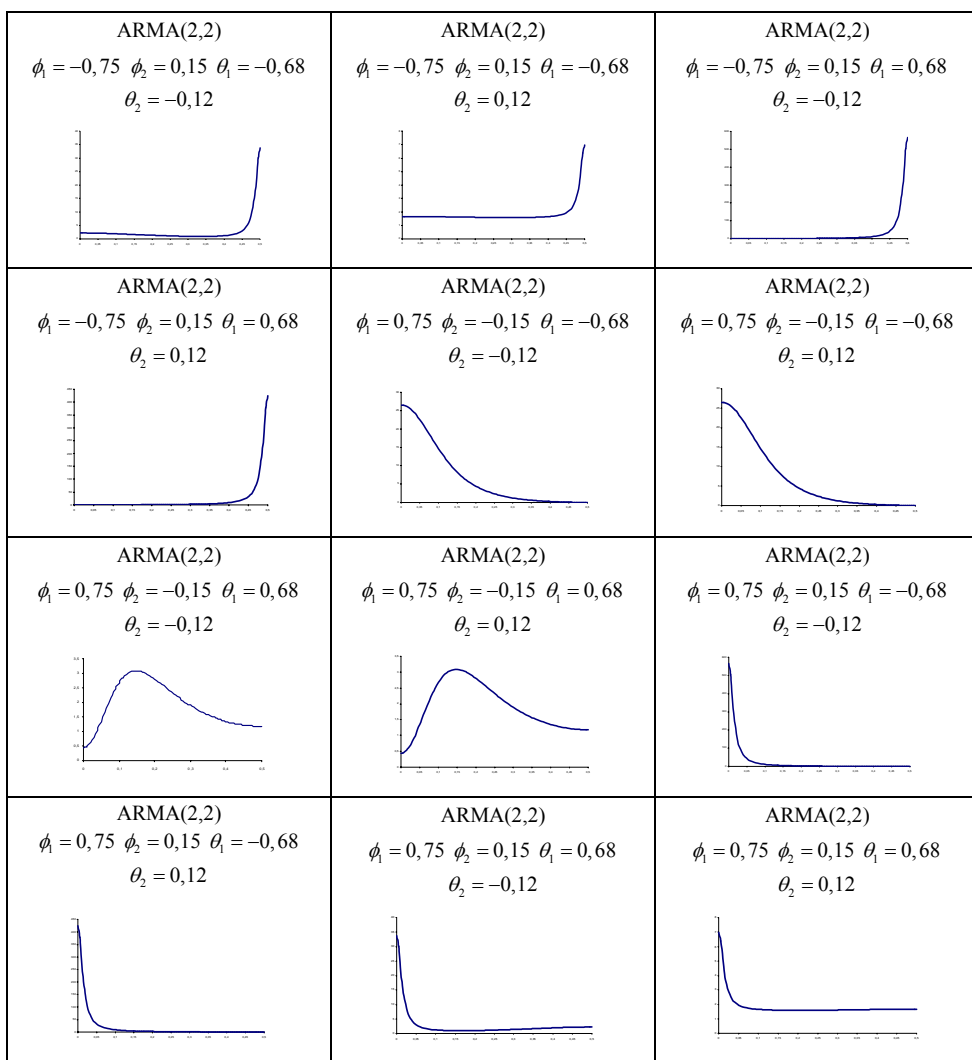
$$\begin{aligned} p(f) &= 2\sigma_\varepsilon^2 \frac{\left|1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f} - \dots - \theta_q e^{-i2\pi qf}\right|^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f} - \dots - \phi_p e^{-i2\pi pf}\right|^2} \quad 0 \leq f \leq 0,5 \\ \text{ARMA(1,1)} \quad p(f) &= 2\sigma_\varepsilon^2 \frac{\left|1 - \theta_1 e^{-i2\pi f}\right|^2}{\left|1 - \phi_1 e^{-i2\pi f}\right|^2} = \frac{1 + \theta_1^2 - 2\theta_1 \cos(2\pi f)}{1 + \phi_1^2 - 2\phi_1 \cos(2\pi f)} \quad 0 \leq f \leq 0,5 \\ \text{ARMA(2,1)} \quad p(f) &= 2\sigma_\varepsilon^2 \frac{\left|1 - \theta_1 e^{-i2\pi f}\right|^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f}\right|^2} = \\ &= \frac{1 + \theta_1^2 - 2\theta_1 \cos(2\pi f)}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos(2\pi f) - 2\phi_2\cos(4\pi f)} \quad 0 \leq f \leq 0,5 \\ \text{ARMA(1,2)} \quad p(f) &= 2\sigma_\varepsilon^2 \frac{\left|1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f}\right|^2}{\left|1 - \phi_1 e^{-i2\pi f}\right|^2} = \\ &= \frac{1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos(2\pi f) - 2\theta_2\cos(4\pi f)}{1 + \phi_1^2 - 2\phi_1 \cos(2\pi f)} \quad 0 \leq f \leq 0,5 \\ \text{ARMA(2,2)} \quad p(f) &= 2\sigma_\varepsilon^2 \frac{\left|1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f}\right|^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f}\right|^2} = \\ &= \frac{1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos(2\pi f) - 2\theta_2\cos(4\pi f)}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos(2\pi f) - 2\phi_2\cos(4\pi f)} \quad 0 \leq f \leq 0,5 \end{aligned}$$

Průběh spektra pro simulované stacionární procesy

Na následujících obrázcích je zachycen grafický průběh spektra pro výše uvedené typy stochastických procesů a různé hodnoty jejich parametrů. U parametrů jsme měnili pouze znaménka (nikoliv jejich absolutní hodnoty) tak, aby bylo patrné, jak změna se změna znaménka u parametrů procesu projeví v průběhu hustoty. Protože se v praxi nejčastěji používají procesy do řádu 2, pro vyšší řády modelů jsme spektrum nepočítali. Všechny procesy, se kterými jsme pracovali, byly vygenerovány pro danou kombinaci parametrů v programu SAS verze 8.2, řady byly převedeny do MS Excel XP, ve kterém byly i napočítány hodnoty spektra a vyrobeny grafy.







Závěr

Na předchozích obrázcích jsme mohli sledovat průběh spektra pro různé typy stacionárních modelů a různé hodnoty jejich parametrů (většinou se mění pouze znaménko parametrů). Na jedné straně je zřejmé, že pro různé hodnoty parametrů má spektrum různý, pro daný model typický průběh. Na druhé straně je ale vidět, že z průběhu spektra nelze jednoznačně určit tvar modelu, neboť různé typy modelů mají pro různé hodnoty parametrů podobný průběh. Nelze tedy z průběhu spektra usuzovat jednoznačně na ten který model a založit na tvaru spektra identifikaci modelu. Průběh spektra však tvoří důležitou doplňkovou informaci, která o tvaru modelu může zkušenému statistikovi již mnohé napovědět.

Literatura

- [1] BOX, G. E. P. –JENKINS, G. M. –REINSEL G. C., 1994: *Time Series Analysis, Forecasting and Control, Third Edition*. New Jersey, Englewood Cliffs, PRENTICE HALL, 1994.
- [2] MAREK, L., 1999: Fourier Analysis. In *Acta Oeconomica Pragensia*, 1999, roč. 7, č. 4, s. 107–118.

Spektrální vlastnosti stacionárních modelů

Luboš Marek

Abstrakt

Obsahem článku je popis chování spektrální hustoty stacionárních modelů, konkrétně autoregresního modelu $AR(p)$, modelu klouzavých součtů $MA(q)$ a smíšeného modelu $ARMA(p,q)$. Na základě jednoznačného vztahu mezi spektrální hustotou a autokorelační funkcí těchto procesů byly odvozeny konkrétní tvary spektra pro různé typy stacionárních procesů. Hlavní částí práce je grafické znázornění průběhu hustoty pro různé kombinace parametrů výše uvedených modelů. Pro různé hodnoty parametrů má spektrum různý, pro daný model typický průběh. Z průběhu spektra však nelze jednoznačně určit tvar modelu, neboť různé typy modelů mají pro různé hodnoty parametrů průběh velmi podobný. Nelze tedy na základě tvaru spektra jednoznačně identifikovat model. Průběh spektra však tvoří důležitou doplňkovou informaci, která o tvaru modelu může mnohé napovědět.

Klíčová slova: stacionární proces; spektrální hustota; spektrum.

Spectral properties of Stationary Models

Abstrakt

The content of this paper is characterization the spectral properties of stationary models – namely autoregression model $AR(p)$, model of moving averages $MA(q)$ and mixed models $ARMA(p,q)$. There is the clear relationship between spectral density and autocorrelation function of these stationary models. The spectrum has the typical shape for different models. This shape depends on sign of parameters. On other side, from shape of spectrum we cannot derive the accurate type of model, because the different models have the similar shape of spectrum. But the shape of spectrum is very important complementary information that can through many things prompt.

Key words: stationary models; spectral density; spectrum.

JEL classification: G30