

MASARYKOVA UNIVERZITA
Fakulta sportovních studií
Katedra kineziologie

**ČASOVÉ ŘADY
V KINANTROPOLOGICKÉM VÝZKUMU**

Habilitační práce

Brno, 2012

Mgr. Martin Sebera, Ph.D.

Prohlašuji, že jsem uvedenou habilitační práci vypracoval samostatně, na základě originálního výzkumu a s využitím uvedených literárních a internetových zdrojů.

Souhlasím, aby práce byla uložena na Masarykově univerzitě v knihovně Fakulty sportovních studií a zpřístupněna ke studijním účelům.

Mgr. Martin Sebera, Ph.D.

Obsah

OBSAH	3
SEZNAM OBRÁZKŮ	4
SEZNAM TABULEK	4
ÚVOD	5
1 ZÁKLADNÍ POJMY	6
1.1 Definice a druhy časových řad.....	6
1.2 Grafická analýza časových řad	9
1.3 Základní charakteristiky časových řad.....	12
1.3.1 Průměry	12
1.3.2 Míry dynamiky.....	15
1.4 Metody analýzy časových řad.....	18
1.4.1 Dekompoziční metoda	19
Nejčastěji používané <i>funkce lineární z hlediska parametrů</i> jsou:.....	20
1.4.2 Adaptivní metody.....	24
1.4.3 Boxovy-Jenkinsonovy metody.....	24
1.4.4 Metoda spektrální analýzy	24
2 ANALÝZA NEPERIODICKÝCH ČASOVÝCH ŘAD	26
2.1 Popis trendu analytickými modely	26
2.1.1 Trendové funkce	26
2.1.2 Metody odhadu parametrů trendových funkcí	27
2.1.3 Volba vhodného modelu trendové funkce	30
3 ZÁKLADY A APARÁT BOXOVY-JENKINSONOVY METODOLOGIE	37
3.1 Základní pojmy	37
3.1.1 Stacionarita	37
3.1.2 Autokorelační funkce (ACF)	38
3.1.3 Parciální autokorelační funkce (PACF)	39
3.1.4 Výběrové ACF a PACF	40
3.2 Základní reprezentace stochastických procesů	41
3.2.1 Proces bílého šumu	41
3.2.2 Woldova a Box-Jenkinsonova reprezentace stochastického procesu	42
3.3 Modely stacionárních časových řad.....	43
3.3.1 Autoregresní proces řádu p [AR(p)]	43
3.3.2 Proces klouzavých průměrů řádu q [MA(q)]	44
3.3.3 Smíšené procesy ARMA [ARMA(p , q)]	45
3.4 Modely nestacionárních a sezónních časových řad	46
3.4.1 Nestacionární procesy ve střední hodnotě.....	46
3.4.2 Proces náhodné procházky („random walk“)... ..	47
3.4.3 Procesy nestacionární v rozptylu	48
3.4.4 Sezónní procesy	48
3.5 Identifikace a ověřování modelu, odhady parametrů, konstrukce předpovědí	49
3.5.1 Identifikace modelu.....	49
3.5.2 Odhady parametrů modelu.....	51
3.5.3 Ověřování modelu.....	57
3.6 Konstrukce předpovědí	63
3.6.1 Výpočet předpovědí	63
3.7 Literatura.....	65
4 WAVELET TRANSFORMACE VE ZPRACOVÁNÍ SIGNÁLŮ	66
4.1 Úvod.....	66
4.2 Matematický základ.....	67

4.2.1 Fourierova a waveletová řada	67
4.2.2 Časově-frekvenční spektrální charakteristika	71
4.2.3 Waveletové vyhlazování (filtrace)	73
4.3 MR-analýza	76
4.3.1 Dekompozice a rekonstrukce signálu	80
4.4 Aplikace	81
4.4.1 Detekce nelinearit a segmentace časových řad	81
4.5 Informace o software	82
4.6 Závěr	83
4.7 Literatura	85
PŘÍPADOVÉ STUDIE	87
Analýza ekonomických dat - inflace	87
Kinematická analýza	102
Distance-Time Dependencies of Spatial Coordinates for Gait Recognition	110
Economical time series prediction via wavelets (wavelet games)	118
Modelování a prognózování časových řad pomocí waveletů	125

Seznam obrázků

Obr. 1 Vývoj světového rekordu v desetiboji mužů	10
Obr. 2 Srovnání pohybu těžiště v ose y dvou testovaných osob při chůzi	12
Obr. 3 Graf kvadratického regresního modelu	35
Obr. 4 Graf kubického regresního modelu	36
Obr. 5 Výběrová ACF a PACF procesu AR(2)	44
Obr. 6 Výběrová ACF a PACF procesu MA(2)	45
Obr. 7 Ukázky waveletů	70
Ob. 8 Ukázka vyhlazení měkkým práhováním	75

Seznam tabulek

Tab. 1 Průměrné denní tržby (tis. Kč)	8
Tab. 2 Vývoj světového rekordu v desetiboji mužů	10
Tab. 3 Srovnání pohybu těžiště v ose y dvou testovaných osob (TO) při chůzi	11
Tab. 4 Počet vyšetřených osob v laboratoři biomotoriky na FSpS v roce 2011	14
Tab. 5 Průměrná hrubá mzda zaměstnanců v ČR v letech 1989 až 2000	17
Tab. 6 Popisné charakteristiky jako kritérium volby trendu časové řady	30
Tab. 7 Vstupní data kvadratického regresního modelu	34
Tab. 8 Výsledky kvadratické regrese	34
Tab. 9 Vstupní data	35
Tab. 10 Výsledky kubické regrese	36
Tab. 11 Tvary ACF a PACF modelu AR, MA a ARMA	50

ÚVOD

Práce předkládá možnosti využití matematicko-statistického aparátu pro analýzu časových řad, které lze identifikovat v kinantropologickém výzkumu na Fakultě sportovních studií. Tato data pochází zejména z trojdimenzionálních kinematických analýz prováděných za účelem identifikace a diagnostiky technického provedení daného pohybu, dále z pedobarometrických plošin v rámci analýzy plantárního tlaku nebo z přístrojů EEG, EKG nebo EMG. Speciální využití časových řad lze nalézt ve výzkumu, který se zabývá problematikou identifikace osob podle dynamického stereotypu chůze.

Analýza časových řad se tak v poslední době stala velmi dynamicky rozvíjející disciplínou. Kromě standardních postupů analýzy časových řad vznikla řada nových efektivních a netradičních postupů a metod modelování časových řad. S rostoucí dostupností výkonné měřicí a výpočetní techniky se sofistikované metody začaly ve stále větším měřítku prosazovat v kinantropologické praxi. Smyslem analýzy časových řad je porozumět mechanismu, na jehož základě jsou hodnoty časové řady vytvářeny. Cílem analýzy je pak konstrukce vhodného modelu a konstrukce předpovědi vývoje časové řady. Jednotlivé přístupy mají při aplikaci své výhody a nevýhody, které jsou na vybraných problémech popsány.

Předkládaný text se skládá z 5 kapitol – Základní pojmy časových řad, Analýza neperiodických řad, Základy a aparát Boxovy-Jenkinsonovy metodologi a Wavelet transformace ve zpracování signálů a Případové studie. V prvních 5 kapitolách převažuje zejména teoretický aparát, který je doplněn tematickými příklady, poslední kapitola je zaměřena výhradně na konkrétní využití a analýzy časových řad v kinantropologickém výzkumu.

Předložené výpočty a grafické výstupy pocházejí z výpočetních systémů STATISTICA verze 10 firmy Statsoft a software pro komplexní matematické výpočty MATLAB firmy MathWorks.

1 ZÁKLADNÍ POJMY

Analýza časových řad je jednou z důležitých oblastí současné statistiky. S daty, která jsou uspořádána chronologicky v čase, se setkáváme snad ve všech oblastech lidské činnosti, sport nevyjímaje. Jako příklady můžeme uvést:

- sledování fyziologických parametrů (změna tepové frekvence, teploty kůže, hladiny laktátu při zátěži aj.)
- záznam EMG při zátěži určité svalové skupiny,
- sledování jednotlivých bifurkačních bodů definovaných z 3D kinematické analýzy

Smyslem analýzy časových řad je porozumět mechanismu, na jehož základě jsou hodnoty časové řady vytvářeny. Cílem analýzy je:

- konstrukce vhodného modelu
- konstrukce předpovědi vývoje časové řady
- zjednodušení časové řady pro další analýzu
- transformace časové řady do jiné datové struktury vhodné pro další analýzu (např. porovnání)

1.1 Definice a druhy časových řad

Časovou řadou rozumíme řadu věcně a prostorově srovnatelných hodnot určitého statistického znaku (ukazatele), uspořádanou z hlediska času ve směru od minulosti do přítomnosti.

Časovou řadu lze nejjednodušším způsobem vyjádřit jako soubor pozorování

$$y_1, y_2, \dots, y_n, \quad (1.1)$$

kde y_t je hodnota posuzovaného znaku Y , $t = 1, \dots, n$, značí časovou proměnnou (časový interval nebo okamžik) a n je počet pozorování (délka řady). Dále se předpokládá, že časová vzdálenost mezi sousedními pozorováními je shodná.

Časové řady dělíme podle různých hledisek:

- Jestliže se hodnota zkoumaného znaku vztahuje k určitému časovému intervalu nenulové délky, pak se **časová řada** nazývá **intervalová** (např. při cyklickém pohybu chůze to je změna definovaná začátkem každého kroku). Pro tento typ časové řady je charakteristická sčitatelnost hodnot za jednotlivé intervaly.
- Jestliže se hodnota zkoumaného znaku vztahuje k určitému okamžiku, pak se **časová řada** nazývá **okamžiková** (např. aktuální tepová frekvence na začátku tréninkové jednotky). Typickým rysem okamžikových časových řad je nesčitatelnost hodnot pro jednotlivé časové okamžiky.
- Podle délky časových úseků s jakými jsou data sledována dělíme **časové řady** na **dlouhodobé** (délka intervalu nebo časové rozpětí mezi rozhodnými okamžiky je alespoň jeden rok, např. vývoj světového rekordu za jednotlivé roky) a na **časové řady krátkodobé** (délka intervalu nebo časové rozpětí mezi rozhodnými okamžiky je menší než jeden rok, např. výkonnost v motorickém testu během přípravného období daného roku).
- Podle druhu sledovaných dat dělíme časové řady na **časové řady původních hodnot** (např. časová řada celkové rychlosti těžiště při běhu) a na **časové řady odvozených charakteristik** (např. relativní rychlost těžiště vůči pevnému bodu).

Při zpracování statistických dat typu časových řad se vyskytuje **řada specifických problémů**, které je třeba při jejich statistické analýze zohlednit:

- V první řadě jde o problém *srovnatelnosti*. Chceme-li v intervalových krátkodobých časových řadách porovnávat jednotlivé hodnoty, musí se tyto hodnoty vztahovat ke stejně dlouhým časovým intervalům, čili přistupují problémy s kalendářem (různá délka kalendářních měsíců, 4 nebo 5 víkendů v měsíci, různý počet pracovních dnů v měsíci, pohyblivé svátky).
- Před analýzou řad, které jsou srovnatelné věcně i prostorově, je nutné provést *očistění časové řady* o důsledky kalendářních variací.

a) **očistění na standardní měsíc** provedeme podle vzorce

$$y_t^0 = \frac{y_t}{k_t} \bar{k}_t, \quad (1.2)$$

kde y_t je hodnota očištěvaného ukazatele, k_t je počet kalendářních dnů v příslušném období, $\bar{k}_t = 30$ nebo $365/12 = 30,4167$ dnů je průměrný počet kalendářních dní.

b) Obdobně jako v případě (1.2) provádíme **očištění na pracovní dny** podle vzorce

$$y_t^0 = \frac{y_t}{p_t} \bar{p}_t, \quad (1.3)$$

kde y_t je hodnota očišťovaného ukazatele, p_t je počet kalendářních dnů v příslušném období, \bar{p}_t je průměrný počet pracovních dnů ve stejném období.

Příklad 1.1

Laboratoř podpory zdraví na FSpS má otevřeno 7 dní v týdnu s otevírací dobou:

Po – Pá: 7 - 18 hod., So: 7 – 11 hod., Ne: 14 – 18 hod.

Pravidelně sleduje tržby v jednotlivých dnech. Průměrné tržby jsou uvedeny v tab. 1.1. Zjistěte, který ze dnů v týdnu je z hlediska tržeb „nejefektivnější“. Co byste na základě zjištěných skutečností vedení laboratoře poradili?

Tab. 1 Průměrné denní tržby (tis. Kč)

Dny	Po	Út	St	Čt	Pá	So	Ne
Tržby	23	33	28	30,8	52	28	4,5

Řešení:

Abychom mohli porovnat efektivnost tržby z hlediska „dne v týdnu“, je třeba očistit velikost tržeb od závislosti na délce otevírací doby. V týdnu je otevřeno 11 hod. denně, proto přepočítáme tržby v So a v Ne také na 11-ti hodinovou otevírací dobu.

$$y_{So} = \frac{28}{4} \cdot 11 = 77, \quad y_{Ne} = \frac{4,5}{4} \cdot 11 = 12,4.$$

Porovnáme-li přepočítané tržby, zjistíme, že „obchodně nejefektivnějším“ dnem v týdnu je sobota, nejméně „obchodně efektivním“ dnem je neděle. Vedení laboratoře by se měl zamyslet nad tím, zda by stálo za to v sobotu prodloužit otevírací dobu a pokud chce mít otevřeno i v neděli, otevírací dobu posunout.

- V mnoha případech je nutné porovnávat údaje časové řady pomocí relativních hodnot. Např. v příkladech z finanční matematiky vzhledem ke změnám cen je nutné často údaje vyrovnat pomocí cenových indexů na tzv. **relativní údaje, srovnatelné ceny**. Cenové změny jsou eliminovány prostřednictvím tzv. stálých cen (např. ceny z roku 1995, ceny z roku 2010, atd.).
- Dalším problémem při analýze časových řad je **zastarávání údajů**, projevující se především v časových řadách značné délky. V těchto případech je třeba hledat **optimální délku časové řady**. Při volbě délky časových řad se střetávají dvě protichůdné tendence: Při nízké hustotě riskujeme, že nám unikne charakteristický rys řady, při velkém množství údajů, které některé metody vyžadují, je nebezpečí, že se s průběhem času změní obsahová náplň ukazatele.
- Specifickým problémem časových řad je **autokorelace**, kdy každé pozorování je statisticky závislé na předchozím. Například kladná autokorelace znamená, že po sobě jdoucí pozorování mají sklon podobat se svým předchůdcům a tudíž přinášejí jen velmi málo čerstvých informací. Takové řady se analyzují mnohem obtížněji než běžný regresní model, ve kterém jsou jednotlivá pozorování navzájem nezávislá. Test pro zjištění, zda data vykazují autokorelaci uvedeme později.
- Dalším rysem, který se může u časových řad vyskytnout, je **sezónnost**. Jestliže řada vykazuje sezónní chování, musíme mít tento fakt stále na paměti, hlavně při hodnocení jednotlivých pozorování a při predikci. Těmito problémy se budeme zabývat v následujících kapitolách.

1.2 Grafická analýza časových řad

Prvním krokem při analýze časových řad je grafické znázornění jejich průběhu. Z grafu lze většinou vyčíst spoustu důležitých informací. V první řadě lze z obrázku odhadnout typ trendové funkce, zda řada vykazuje periodicitu apod. Někdy lze z obrázku vyčíst, zda máme použít aditivní nebo multiplikativní model. Nejčastěji se znázorňují původní hodnoty časové řady pomocí spojnicových a plošných grafů.

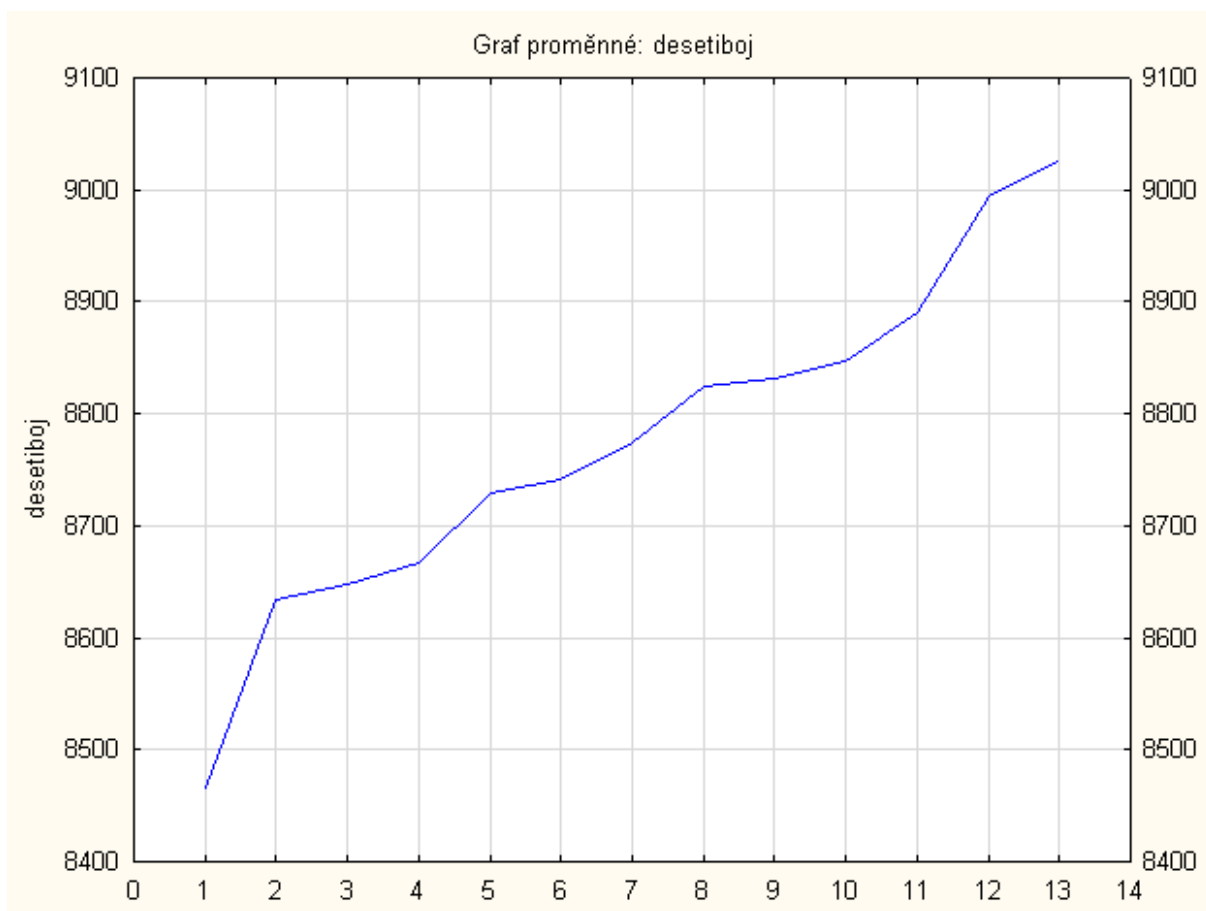
Spojnicové grafy

Princip spojnicových grafů spočívá v zakreslení jednotlivých hodnot časové řady do souřadných os, na kterých jsou vyznačeny příslušné stupnice. Na osu horizontální se vynášší

časová proměnná t a na osu vertikální hodnoty y_t časové řady. Takto vzniklé body jsou spojeny lomenou čarou – viz obr. 1.3.

Tab. 2 Vývoj světového rekordu v desetiboji mužů

Hodnota	Jméno	Rok
8466	Nikolaj Avilov	1972
8634	Bruce Jenner	1976
8648	Daley Thompson	1980
8667	Quido Kratchmer	1980
8730	Daley Thompson	1982
8741	Jurgen Hingsen	1982
8774	Daley Thompson	1982
8825	Jurgen Hingsen	1983
8832	Jurgen Hingsen	1984
8847	Daley Thompson	1984
8891	Dan O'Brien	1992
8994	Tomáš Dvořák	1999
9026	Roman Šebrle	2001



Obr. 1 Vývoj světového rekordu v desetiboji mužů

Plošné grafy

Pro zobrazení a následné porovnání dvou a více časových řad používáme plošné grafy – viz tab. 1.3 a obr. 1.4.

Tab..3 Srovnání pohybu těžiště v ose y dvou testovaných osob (TO) při chůzi

čas	TO1	TO2
1	292,2	305,1
2	288,2	306,8
3	286,0	307,1
4	284,7	309,2
5	283,4	311,8
6	283,3	312,6
7	283,5	313,3
8	283,0	314,9
9	282,7	316,4
10	283,6	318,4
11	282,9	319,9
12	285,0	322,7
13	284,3	324,1
14	285,9	326,9
15	286,9	329,0
16	288,5	331,7
17	288,9	334,3
18	290,6	337,4
19	292,7	340,2
20	295,2	344,3
21	297,5	348,4
22	299,3	352,1
23	301,8	356,8
24	305,0	361,0
25	307,9	365,4
26	312,2	369,5
27	316,1	374,1
28	318,9	378,8
29	322,9	382,8
30	326,5	385,5
31	331,2	388,0
32	334,6	390,2
33	338,6	391,0
34	343,0	391,6
35	346,5	392,4
36	349,7	392,3
37	351,0	392,1
38	352,3	391,6
39	353,3	390,7
40	353,4	390,0
41	352,5	389,6
42	351,8	389,1
43	351,1	388,9
44	349,9	388,1
45	349,4	387,4
46	347,4	386,0
47	346,3	384,9
48	345,1	383,4
49	343,6	382,3
50	342,1	379,7
51	340,9	378,5
52	339,4	376,3
53	338,2	374,1

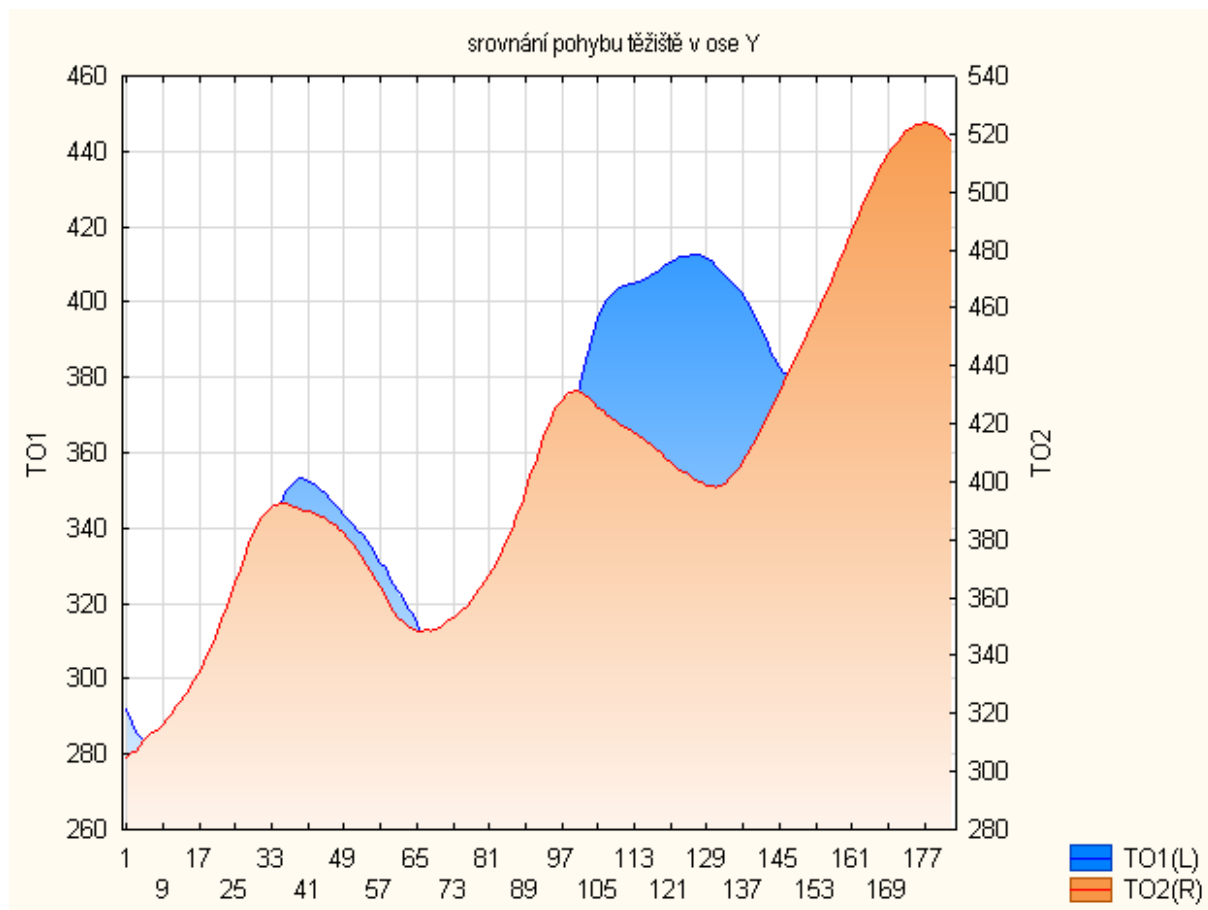
čas	TO1	TO2
54	336,6	371,1
55	335,2	369,0
56	332,9	365,8
57	330,7	363,8
58	329,6	360,5
59	326,7	357,5
60	324,7	355,2
61	322,7	352,5
62	321,2	351,6
63	318,8	350,2
64	317,3	349,0
65	314,8	348,6
66	312,6	348,5
67	310,6	348,7
68	308,1	348,6
69	305,5	349,0
70	304,7	349,5
71	303,6	350,7
72	303,4	352,1
73	303,9	353,1
74	304,8	354,6
75	305,0	356,0
76	305,4	357,3
77	306,9	359,3
78	308,2	361,3
79	309,1	363,6
80	310,7	365,6
81	312,1	367,9
82	313,7	370,6
83	316,0	373,2
84	318,2	376,6
85	319,9	380,2
86	320,3	383,5
87	324,6	387,7
88	327,1	391,8
89	329,7	396,6
90	332,6	401,7
91	335,6	406,1
92	338,7	410,5
93	342,6	415,4
94	346,4	419,5
95	350,4	423,3
96	354,3	426,3
97	358,6	428,3
98	363,2	429,8
99	368,2	430,6
100	373,5	431,1
101	378,0	430,7
102	382,9	430,0
103	387,9	428,8
104	392,2	427,7
105	395,4	425,7
106	398,5	424,8

čas	TO1	TO2
107	400,8	422,7
108	402,0	421,8
109	403,3	420,5
110	403,9	419,1
111	404,5	418,6
112	404,6	417,8
113	404,8	416,7
114	405,1	415,8
115	405,7	414,4
116	406,3	413,6
117	407,4	412,5
118	408,0	410,8
119	409,1	409,9
120	410,1	408,1
121	410,8	406,6
122	411,1	405,2
123	411,9	404,3
124	412,2	403,1
125	412,0	402,3
126	412,4	401,1
127	412,5	400,2
128	411,9	399,7
129	411,7	398,7
130	411,0	398,6
131	409,6	398,1
132	408,5	398,7
133	406,9	399,5
134	406,0	401,1
135	404,7	402,8
136	403,6	404,7
137	402,2	406,8
138	400,0	409,6
139	397,8	412,7
140	395,7	415,0
141	392,8	418,0
142	390,4	421,2
143	387,3	424,4
144	385,2	427,6
145	382,9	430,5
146	381,2	433,9
147	380,4	437,5
148	379,1	441,0
149	378,6	444,4
150	378,4	447,8
151	378,3	451,1
152	378,3	454,5
153	378,5	457,8
154	378,8	461,2
155	379,3	464,4
156	379,9	467,9
157	380,0	471,4
158	380,6	475,2
159	381,1	478,9

čas	TO1	TO2
160	382,3	482,8
161	383,1	486,6
162	384,2	490,2
163	386,1	494,2

čas	TO1	TO2
164	388,4	497,6
165	390,9	501,2
166	393,2	504,7
167	396,4	507,8

čas	TO1	TO2
168	399,6	510,4
169	403,0	513,1
170	406,3	515,5



Obr. 2 Srovnání pohybu těžiště v ose y dvou testovaných osob při chůzi

1.3 Základní charakteristiky časových řad

1.3.1 Průměry

Prostý aritmetický průměr

počítáme pouze u intervalových časových řad, protože v intervalových časových řadách má součet hodnot znaku věcný smysl. Průměrná hodnota intervalové časové řady

$$\bar{y} = \frac{\sum_{t=1}^n y_t}{n}, \text{ kde } n \text{ je počet časových intervalů.} \quad (1.4)$$

Chronologický průměr

počítáme z hodnot okamžikové časové řady. Průměrná hodnota okamžikové časové řady y_t , $t = 1, \dots, n$, kde vzdálenost mezi jednotlivými časovými okamžiky je stejná, má tvar

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + \dots + \frac{y_{n-1} + y_n}{2}}{n-1} = \frac{\frac{1}{2}y_1 + \sum_{t=2}^{n-1} y_t + \frac{1}{2}y_n}{n-1}. \quad (1.5)$$

Při různé délce mezi jednotlivými okamžiky se používá vážený chronologický průměr

$$\bar{y} = \frac{\frac{y_1 + y_2}{2}d_2 + \frac{y_2 + y_3}{2}d_3 + \dots + \frac{y_{n-1} + y_n}{2}d_n}{d_2 + d_3 + \dots + d_n}, \quad (1.6)$$

kde d_i , $i = 2, \dots, n$, je délka jednotlivých časových intervalů sledování daného okamžikového ukazatele.

Klouzavé průměry

Klouzavé průměry jsou určitou lineární kombinací vždy m původních hodnot v časové řadě.

Prosté klouzavé průměry

Časovou řadu rozdělíme na úseky po lichém počtu hodnot. Období $m = 2p+1$ nazýváme *klouzavou částí*, kterou posunujeme (kloužeme) po časové ose tak, že z prvních $2p+1$ hodnot vypustíme první hodnotu a přibereme následující hodnotu. Z lichého počtu ($m=2p+1$) hodnot počítáme prostý aritmetický průměr.

$$\bar{y}_t = \frac{1}{2p+1} \sum_{i=-p}^p y_{t+i}, \quad t = p+1, p+2, \dots, n-p. \quad (1.7)$$

Protože klouzavá část má lichý počet členů, je pořadí prostředního z nich celočíselné. Průměr (1.7) přiřadíme prostřední empirické hodnotě y_t .

Centrované klouzavé průměry

Má-li klouzavá část sudý počet hodnot, tj. $m = 2p$, pak prostřední hodnota nemá celočíselné pořadí a musíme prosté klouzavé průměry centrovat.

$$\bar{y}_t = \frac{1}{4p} (y_{t-p} + 2y_{t-p+1} + \dots + 2y_{t+p-1} + y_{t+p}), \quad t = p+1, p+2, \dots, n-p. \quad (1.8)$$

Počítáme aritmetický průměr dvou sousedních klouzavých průměrů a hodnotu centrovaného klouzavého průměru (1.8) přiřadíme empirické hodnotě y_t , $t = p+1, p+2, \dots, n-p$.

Vážené klouzavé průměry

Při užití klouzavých průměrů počítáme pro každý klouzavý úsek časové řady vážený aritmetický průměr

$$\bar{y} = \sum_{i=-p}^p w_i y_{t+i}, \quad t = p+1, p+2, \dots, n-p, \quad (1.9)$$

kde w_i jsou váhy, pro které platí $\sum_{i=-p}^p w_i = 1$ a mohou být konstruovány různým způsobem. Podrobněji se budeme zabývat váženými klouzavými průměry v části věnované popisu trendu časové řady.

Řady klouzavých průměrů mají vždy o $2p$ hodnot méně (chybí p hodnot na začátku a p hodnot na konci řady), než časové řady absolutních ukazatelů. Řady centrovaných klouzavých průměrů jsou dokonce o $2p+1$ hodnot kratší.

Příklad 1.2

Určete průměrné hodnoty pohybu těžiště TO 1 a TO 2 (tab. 1.3).

Řešení:

Protože uvedená časová řada je intervalová, vypočítáme prostý aritmetický průměr (1.4).

$$\bar{y}_{TO1} = \frac{65323}{183} = 356,9 \quad \text{a} \quad \bar{y}_{TO2} = \frac{73911,1}{183} = 403,9.$$

Příklad 1.3

Vypočítejte průměrný měsíční počet vyšetřených osob v laboratoři biomotoriky na FSpS v roce 2012 (stav ke konci období). Vstupní údaje jsou v tab. 4.

Tab. 4 Počet vyšetřených osob v laboratoři biomotoriky na FSpS v roce 2011

Období	Počet uchazečů	Délka intervalu
leden	151	31
únor	111	28
březen	442	31
duben	200	30
květen	243	31
červen	196	30
červene	128	31

c		
srpen	264	31
září	272	30
říjen	174	31
listopad	232	30
prosinec	169	31

Řešení:

Pro výpočet použijeme vážený chronologický průměr (1.6), protože počet osob tvoří okamžikovou časovou řadu s nestejně dlouhými úseky mezi okamžiky zjišťování.

$$\bar{y} = \frac{\frac{151+111}{2} \cdot 28 + \frac{111+442}{2} \cdot 31 + \dots + \frac{232+169}{2} \cdot 31}{28 + 31 + \dots + 31} = 220,6$$

V roce 2011 bylo průměrný počet vyšetřených osob 220,6.

Příklad 1.4

Z tab. 1.4 vypočítáme čtvrtletní klouzavé průměry.

$$\frac{151+111+442}{3}, \frac{111+442+200}{3}, \dots, \frac{174+232+169}{3}$$

Nová časová řada je pak 234,7; 251,0; 295,0; 213,0; 189,0; 196,0; 221,3; 236,7; 226,0; 191,7

1.3.2 Míry dynamiky

Dynamikou vývoje časové řady rozumíme změny hodnot sledovaného ukazatele v čase, a to buď v absolutním nebo v relativním pojetí. Charakteristiky dynamiky vývoje časových řad lze počítat jak u intervalových, tak u okamžikových řad. Nutnou podmínkou pro věcný smysl a správnou interpretaci charakteristik je shoda délky časových intervalů, resp. shoda vzdáleností mezi okamžiky zjišťování.

Přírůstky (1. difference)

Absolutní přírůstek

$$\Delta_t^{(1)} = y_t - y_{t-1}, \quad t = 2, 3, \dots, n. \quad (1.10)$$

Průměrný absolutní přírůstek

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=2}^n \Delta_t^{(1)} = \frac{y_n - y_1}{n-1}. \quad (1.11)$$

Relativní přírůstek

$$\delta_t = \frac{\Delta_t^{(1)}}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1, \quad t = 2, 3, \dots, n. \quad (1.12)$$

Relativní přírůstek vyjádřený v % se nazývá *tempo růstu* sledovaného ukazatele v původní časové řadě.

První diference umožňují charakterizovat směr, velikost absolutních změn zkoumaného znaku. Řadu složenou z prvních diferencí lze dále diferencovat a získáme řadu druhých diferencí atd.

Druhé diference

$$\Delta_t^{(2)} = \Delta_t^1 - \Delta_{t-1}^1, \quad t = 3, 4, \dots, n \quad (1.13)$$

Koeficienty růstu

koeficient růstu

$$k_t = \frac{y_t}{y_{t-1}}, \quad t = 2, 3, \dots, n. \quad (1.14)$$

Koeficienty růstu vyjádřené v % se označují jako *tempa růstu*.

průměrný koeficient růstu

$$\bar{k} = \sqrt[n-1]{k_1 k_2 \dots k_n} = \sqrt[n-1]{\frac{y_n}{y_1}}. \quad (1.15)$$

Charakteristiky (1.11) a (1.15) závisí bez ohledu na délku řady jen na krajních hodnotách. Proto je užíváme jen na řady s monotónním vývojem. V ostatních případech tyto charakteristiky vývoj zkreslují.

Příklad 1.5

Pro časovou řadu hodnot průměrné měsíční hrubé mzdy v České republice v letech 1989 až 2000 vypočítáme:

- absolutní přírůstky a průměrný absolutní přírůstek,
- koeficienty růstu a průměrný koeficient růstu.

Vstupní hodnoty i výpočty jsou v tab. 1.6.

Tab. 5 Průměrná hrubá mzda zaměstnanců v ČR v letech 1989 až 2000

Rok	y_t	$\Delta_t^{(1)}$	k_t
1989	3170	-	-
1990	3286	116	1,037
1991	3792	506	1,154
1992	4644	852	1,225
1993	5817	1173	1,253
1994	6894	1077	1,185
1995	8172	1278	1,185
1996	9676	1504	1,184
1997	10696	1020	1,105
1998	11693	997	1,094
1999	12666	973	1,083
2000	13490	824	1,065

Řešení:

- První diference jsou ve třetím sloupci tab.1.6. Průměrný absolutní přírůstek (1.11) říká, že měsíční hrubá mzda stoupala v letech 1989-2000 průměrně o 938 Kč ročně.
- Koeficienty růstu (1.14) jsou v posledním sloupci tabulky 1.6.
- Průměrný koeficient růstu $\bar{k} = \sqrt[11]{\frac{13\,490}{3\,170}} = 1,141$ říká, že hrubá mzda rostla v letech

1989- 2000 v průměru ročně o 14%.

1.4 Metody analýzy časových řad

Pro analýzu časových řad existuje dnes velmi rozsáhlá teorie, která popisuje velké množství metod a postupů. Volba vhodné metody závisí na následujících faktorech, které analýzu podmiňují:

- **účel analýzy** – musíme vědět, zda nám jde pouze o tvorbu modelu, o konstrukci předpovědi, o vzájemné vztahy s jinými řadami atd.
- **typ časové řady** – ne všechny metody jsou vhodné pro všechny řady.
- **znalosti statistika** – některé metody jsou teoreticky nenáročné, jiné kladou na vzdělání statistika nemalé nároky.
- **softwarové vybavení** - analýza časových řad se dnes neobejde bez vybavení výpočetní technikou a hlavně programy vhodnými pro analýzu.

Analýza časové řady vychází z představy, že každou *empirickou hodnotu* y_t pro $t = 1, \dots, n$, lze vyjádřit ve tvaru

$$y_t = f(Y_t, \varepsilon_t), \quad t = 1, \dots, n, \quad (1.15)$$

kde $Y_t, t = 1, \dots, n$, představuje *teoretickou (systematickou) složku* časové řady a $\varepsilon_t, t = 1, \dots, n$, *náhodnou (nepravidelnou) složku*.

Základní obtíž při analýze časové řady je úsudek o funkci $f()$. Nejjednodušší je předpoklad, že obě složky časové řady empirických hodnot – systematická a náhodná – jsou ve vztahu součtu., kde

$$y_t = Y_t + \varepsilon_t, \quad t = 1, \dots, n. \quad (1.16)$$

V takovém případě mluvíme o *aditivním modelu* časové řady.

V případě *multiplikativního modelu* předpokládáme, že obě složky (systematická a náhodná) jsou ve vztahu součinu, kde

$$y_t = Y_t \cdot \varepsilon_t, \quad t = 1, \dots, n. \quad (1.17)$$

Multiplikativní model (1.17) za předpokladu, že empirické hodnoty $y_t > 0$, můžeme logaritmováním převést na aditivní model časové řady logaritmů. Protože známe pouze

empirické hodnoty y_t , nahrazujeme systematickou složku Y_t v aditivním modelu (1.16) odhadem \hat{Y}_t a rozdíl

$$e_t = y_t - \hat{Y}_t, \quad t = 1, \dots, n \quad (1.18)$$

nazýváme **reziduem**. Reziduum $e_t = \hat{\varepsilon}_t$ je **odhadem náhodné složky** ε_t .

1.4.1 Dekompoziční metoda

Dekompoziční metoda je klasickou metodou analýzy časových řad. Vychází předpokladu, že časovou řadu lze rozložit na čtyři složky:

- *trendovou složku* T_t , která vyjadřuje dlouhodobou změnu v chování časové řady,
- *cyklickou složkou* C_t , která popisuje dlouhodobé pravidelně se opakující výkyvy kolem trendu v rámci několika let,
- *sezónní složku* S_t , která vyjadřuje pravidelně se opakující výkyvy v rámci jednoho roku,
- *náhodnou složku* ε_t , která je tvořena náhodnými výkyvy. Pod tuto složku řadíme všechny vlivy, které na časovou řadu působí a které nedokážeme systematicky podchytit a popsat.

Hodnoty systematické složky lze tedy zapsat jako funkci trendu, cyklické složky a sezónní složky, tj.

$$Y_t = f(T_t, C_t, S_t), \quad t = 1, \dots, n. \quad (1.19)$$

Měření a matematický popis systematické složky se nazývá **vyrovnáním (vyhlazováním)** časových řad. Časovou řadu, která má pouze trendovou složku nazýváme **neperiodickou** časovou řadou. Řadu, která kromě trendu obsahuje pravidelně se opakující výkyvy hodnot zkoumaného znaku od hlavního vývojového směru, nazýváme **periodickou** časovou řadou.

Dekompoziční metoda tedy spočívá v izolování trendu a periodické (nejčastěji sezónní) složky jako neměnných prvků vývoje v celém průběhu zkoumané časové řady. Tento aparát modelování vývoje je často příliš zjednodušující a tudíž nedostatečný k zachycení složitějšího vývoje.

Regresní metoda

Regresní přístup objasňuje vývoj jedné veličiny prostřednictvím změn jiné nebo jiných veličin. Uplatnění regresního přístupu předpokládá eliminaci rušivých faktorů tak, aby mohla být odhalena skutečná závislost mezi vysvětlovanou a vysvětlující veličinou. K popisu vývoje se nejčastěji používá nějaká *spojitá funkce času*.

Pokud jde o funkci *lineární v parametrech*, používáme k výpočtu jejich parametrů **metodu nejmenších čtverců**. Vedle jednoduchých lineárních funkcí (polynomická, lomená, logaritmická) se používají *funkce nelineární vzhledem k parametrům* (exponenciální, logistická křivka, Gompertzova křivka, S-křivky), jejichž parametry možno počítáme různými *iteračními metodami*.

Lineární regresní funkce má tvar

$$Y = f(x, \beta) = \beta_0 f_0(x) + \beta_1 f_1(x) + \beta_2 f_2(x) + \dots + \beta_p f_p(x), \quad (1.20)$$

kde $\beta_j, j = 0, 1, \dots, p$, jsou regresní parametry a regresory $f_j(x)$ jsou známé funkce proměnné X .

Nejčastěji používané *funkce lineární z hlediska parametrů* jsou:

- *regresní přímka* $Y = \beta_0 + \beta_1 x$,
- *regresní parabola* $Y = \beta_0 + \beta_1 x + \beta_2 x^2$,
- *regresní log. funkce* $Y = \beta_0 + \beta_1 \ln x$,
- *regresní hyperbola* $Y = \beta_0 + \beta_1 \frac{1}{x}$.

V opačném případě, kdy regresní funkce má tvar rozdílný od (1.20), mluvíme o **nelineární regresní funkci**. Například:

- *regresní exponenciální funkce* $Y = \beta_0 \beta_1^x$,
- *regresní mocninná funkce* $Y = \beta_0 x^{\beta_1}$,
- *posunutá exponenciální funkce* $Y = \beta_0 \beta_1^x + \beta_2$.

Exponenciální regrese

Mezi nejužívanější nelineární regresní modely patří exponenciální regresní funkce

$$Y = \beta_0 \beta_1^x, \quad (1.21)$$

kde $\mathbf{Y} = (y_1, \dots, y_n)'$ je náhodný vektor pozorovaných hodnot, β_0, β_1 jsou neznámé parametry, $\mathbf{x} = (x_1, \dots, x_n)'$ je vektor známých čísel.

Logaritmováním (linearizací) exponenciální funkce $Y = \beta_0 \beta_1^x$ získáme tvar lineární vzhledem k parametrům, tj.

$$\ln Y = \ln \beta_0 + x \ln \beta_1 \quad (1.22)$$

Odhady parametrů $\ln \beta_0$ a $\ln \beta_1$ získáme řešením soustavy rovnic

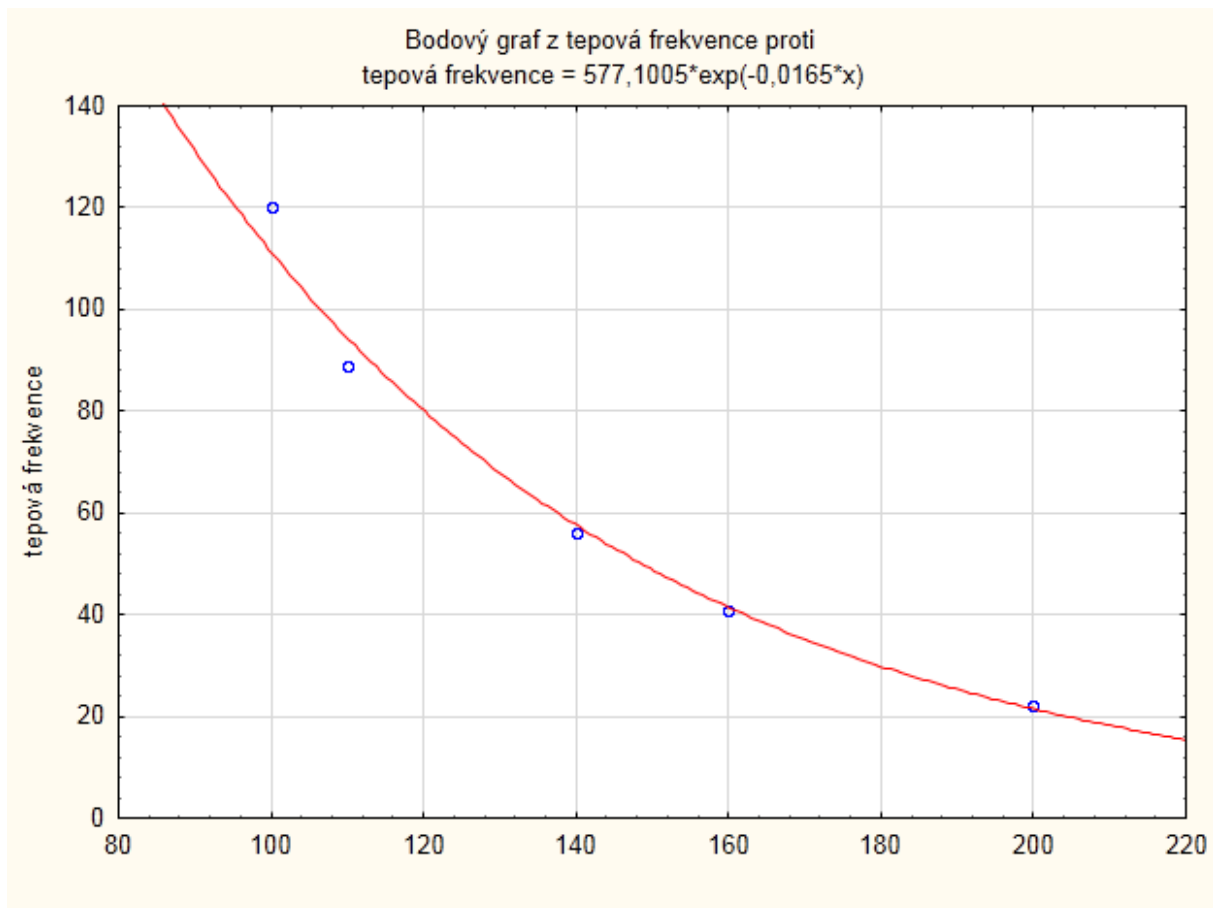
$$n \ln b_0 + \ln b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n \ln y_i, \quad (1.23)$$

$$\ln b_0 \sum_{i=1}^n x_i + \ln b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \ln y_i.$$

Příklad 1.6

Data v následující tabulce představují tepovou frekvenci (Y) po anaerobním tréninku v závislosti na čase (X). Vyrovnějte data exponenciální regresní funkcí a tepovou frekvenci v čase 120 s.

x_i (čas v s)	100	110	140	160	200
y_i (tepová frekvence)	120	89	56	41	22



$$[\hat{Y} = 577,101 \cdot e^{-0,0165x}, \hat{Y}(120) = 79,6]$$

Mocninná regrese

Vedle exponenciální regresní funkce se často používá mocninná funkce

$$Y = \beta_0 x^{\beta_1} \quad (1.24)$$

Multiplikativní model $Y = \beta_0 x^{\beta_1}$ je nelineární z hlediska parametrů, však logaritmováním jej můžeme linearizovat. Rovnice

$$\ln Y = \ln \beta_0 + \beta_1 \ln x,$$

je rovnicí přímky v logaritmických souřadnicích a parametry $\ln \beta_0$ a β_1 .

Odhady parametrů $\ln \beta_0$ a β_1 získáme řešením soustavy rovnic

$$n \ln b_0 + b_1 \sum_{i=1}^n \ln x_i = \sum_{i=1}^n \ln y_i \quad (1.25)$$

$$\ln b_0 \sum_{i=1}^n \ln x_i + b_1 \sum_{i=1}^n (\ln x_i)^2 = \sum_{i=1}^n \ln x_i \ln y_i .$$

Příklad 1.7

Data z příkladu 1.6 vyrovnejte mocninnou regresní funkcí. Na základě tohoto multiplikativního modelu odhadněte velikost tepové frekvence při 20 s.

x_i (čas v s)	100	110	140	160	200
y_i (tepová frekvence)	120	89	56	41	22

Řešením soustavy (1.25) získáme odhady

$$\ln b_0 = 15,646, \quad b_1 = -2,361, \quad \text{čili } \ln \hat{Y} = 15,646 - 2,361 \ln x .$$

Po zpětné transformaci

$$\hat{Y} = \exp(15,646 - 2,361 \ln x), \quad \hat{Y}(120) = \exp(15,646 - 2,361 \ln 120) = 76,917 .$$

Při čase 120 s můžeme očekávat tepovou frekvenci 77.

Vzhledem k tomu, že hodnoty regresních koeficientů byly odhadnuty pomocí výběrových (naměřených) hodnot, lze **výsledky používat k odhadům pouze v rozsahu těchto naměřených hodnot!**

1.4.2 Adaptivní metody

Někdy není možné vyrovnat časovou řadu jedinou funkcí v celé délce řady. Nelze-li vyrovnat delší časovou řadu jedinou křivkou, lze oprávněně předpokládat úspěšnost vyrovnání dílčích částí této řady systémem křivek, které se přizpůsobují (**adaptují**) lokálnímu průběhu časové řady.

Základní metodou adaptivního přístupu k systematické složce *je technika klouzavých průměrů*, kdy se vyrovnávání časových řad provádí po kratších klouzavých úsecích.

Další adaptivní technikou je skupina metod tzv. *exponenciálního vyrovnávání*, které jsou založeny na zohlednění procesu stárnutí dat. Název mají tyto metody podle toho, že váha pozorování je exponenciálně klesající funkcí věku pozorování. Tyto metody se rozlišují podle předpokládaného trendu a přítomnosti či nepřítomnosti sezónní složky.

1.4.3 Boxovy-Jenkinsovy metody

Tyto metody představují jisté pokračování a zdokonalení adaptivního přístupu. Boxova-Jenkinsova metodologie je založena na hypotéze, že všechny složky časové řady mají náhodný (stochastický) charakter. Většina z těchto metod (ARMA, ARIMA, SARIMA) je početně náročná, vyžaduje poměrně vysoký počet pozorování a předpokládá využití počítače. Podrobněji se o této metodě zmíníme v další kapitole.

1.4.4 Metoda spektrální analýzy

Tyto metody přistupují k časové řadě jako ke směsi určitého konečného nebo nekonečného počtu sinusových a kosinových křivek s různými frekvencemi, amplitudami a fázovými posuny. Hlavním nástrojem spektrální analýzy je *peridiogram* umožňující grafickou formou identifikovat nejvýznamnější frekvence, obsažené v časové řadě. Tato metoda umožňuje vystopovat významné složky periodicity, které se podílí na věcných vlastnostech zkoumaného procesu. Společně s fourierovskou transformací popíšeme Wavelet transformaci (WT), která představuje nový matematický prostředek pro analýzu signálů

pomocí Wavelet funkcí s aplikací na široké spektrum reálných signálů, které zahrnuje technologické časové řady, biomedicínské signály a obrazy, družicové snímky, ekonomická data i lidskou řeč. V řadě případů je problémem efektivní kódování, komprese, potlačování rušivých složek, modelování a detekce dílčích složek signálu. Wavelet transformace představuje alternativní přístup k diskrétní Fourierově transformaci (FFT) pro analýzu nestacionárních signálů a detekci bodů, pro které signál mění vlastnosti. Hlavní výhoda WT spočívá v její schopnosti analýzy signálů s proměnným časově-frekvenčním rozlišením.

Podrobněji se o těchto metodách zmíníme v kapitole 4.

2 ANALÝZA NEPERIODICKÝCH ČASOVÝCH ŘAD

Neperiodickou časovou řadou nazýváme řadu, která má pouze trendovou a náhodnou složku, tj. řadu

$$y_t = T_t + \varepsilon_t, \quad t = 1, \dots, n. \quad (2.1)$$

V neperiodické řadě je systematická složka $Y_t = T_t$ a její odhad $\hat{Y}_t = \hat{T}_t$.

2.1 Popis trendu analytickými modely

Analytickým modelováním rozumíme nalezení funkce, která nejlépe popisuje trend časové řady. Při této metodě vyrovnáváme jednou trendovou funkcí všechna empirická pozorování.

2.1.1 Trendové funkce

Při popisu trendu využíváme spojité funkce času. Mezi nejčastěji používané patří

➤ **přímka (lineární trend)**

$$T_t = \beta_0 + \beta_1 t, \quad t = 1, \dots, n, \quad (2.2)$$

➤ **parabola druhého řádu (kvadratický trend)**

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2, \quad t = 1, \dots, n, \quad (2.3)$$

➤ **parabola k-tého řádu (polynomický trend)**

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_k t^k, \quad 1 \leq k < n, \quad (2.4)$$

➤ **dvouparametrická exponenciála (ryzí exponenciální trend)**

$$T_t = \beta_0 e^{\beta_1 t}, \quad \text{kde } \beta_1 > 0, \quad t = 1, \dots, n, \quad (2.5)$$

➤ **tříparametrická exponenciála (modifikovaný exponenciální trend)**

$$T_t = \beta_0 + \beta_1 e^{\beta_2 t}, \text{ kde } \beta_2 > 0, \quad t = 1, \dots, n, \quad (2.6)$$

➤ **logistická funkce (logistický trend)**

$$T_t = \beta_0 / (1 + \beta_1 \beta_2^t), \text{ kde } \beta_0 > 0, \beta_2 > 0, \quad t = 1, \dots, n. \quad (2.7)$$

2.1.2 Metody odhadu parametrů trendových funkcí

Pro odhad parametrů trendové funkce, která je lineární z hlediska parametrů, se používá:

- **Metoda nejmenších čtverců.**

Tuto metodu jsme používali v regresní analýze, kdy jsme minimalizovali výraz

$$\sum_{t=1}^n (y_t - \hat{Y}_t)^2, \quad (2.8)$$

kde y_t je naměřená hodnota a funkce

$$\hat{Y} = b_0 + b_1 f_1(t) + b_2 f_2(t) + \dots + b_p f_p(t) \quad (2.9)$$

je odhad trendové funkce, která je lineární vzhledem k parametrům.

Z regresní analýzy plyne, že odhad metodou nejmenších čtverců vychází z následujících **předpokladů** o náhodné složce:

➤ Náhodná složka ε_t má v každém pozorování nullovou střední hodnotu, čili

$$E(\varepsilon_t) = 0, \quad t = 1, \dots, n \quad (2.10)$$

➤ Rozptyl náhodné složky modelu je konstantní, tj.

$$E(\varepsilon_t^2) = \sigma^2, \quad t = 1, \dots, n. \quad (2.11)$$

➤ Hodnoty náhodné složky z různých pozorování jsou nekorelované, tj.

$$\text{Cov}(\varepsilon_t, \varepsilon_{t+j}) = 0 \text{ pro } t = 1, \dots, n \quad \text{a } j \neq 0. \quad (2.12)$$

➤ Náhodná složka ε_t má normální rozdělení.

V ostatních případech (2.5) – (2.7), kdy trendová funkce není lineární v parametrech, používáme k získání odhadů parametrů následující metody:

- **Metoda vybraných bodů**

spočívá v tom, že na časové ose zvolíme tolik bodů, kolik má trendová funkce parametrů, a v těchto bodech položíme empirické hodnoty časové řady rovny teoretickým hodnotám trendové funkce. Získáme soustavu rovnic, kterou vyřešíme vzhledem k neznámým parametřům. Řešením „počáteční“ odhady, které pak pomocí iteračních metod zpřesňujeme.

Chceme-li např. odhadnout 3 *parametry logistické funkce*, zvolíme na časové ose 3 body $t, t+m, t+2m$ a ze soustavy

$$y_t = \frac{b_0}{1 + b_1 b_2^t}, \quad y_{t+m} = \frac{b_0}{1 + b_1 b_2^{t+m}}, \quad y_{t+2m} = \frac{b_0}{1 + b_1 b_2^{t+2m}} \quad (2.13)$$

pomocí základních matematických úprav vypočítáme b_0, b_1, b_2 .

$$b_2^m = \frac{\frac{1}{1 + b_1 b_2^t} - \frac{1}{1 + b_1 b_2^{t+m}}}{\frac{1}{1 + b_1 b_2^{t+m}} - \frac{1}{1 + b_1 b_2^t}}, \quad (2.14)$$

$$b_1 = \frac{y_t - y_{t+m}}{b_2^t (b_2^m y_{t+m} - y_t)}, \quad (2.15)$$

$$b_0 = y_t (1 + b_1 b_2^t). \quad (2.16)$$

- **Metoda částečných součtů**

spočívá v tom, že časovou osu na tolik stejně velkých disjunktních částí, kolik má trendová funkce parametrů. Někdy je nutné několik hodnot na počátku řady vynechat, aby mohly být vytvořeny stejně velké části. V jednotlivých částech sečteme empirické hodnoty a tyto součty položíme rovny součtům teoretických hodnot trendové funkce. Získáme soustavu rovnic, kterou vyřešíme vzhledem k neznámým parametřům. Rovněž tato metoda poskytuje „počáteční“ odhady, které pak pomocí iteračních metod zpřesňujeme.

Chceme-li např. odhadnout *parametry modifikované exponenciály*, zvolíme počáteční bod t a tři na sebe navazující disjunktní časová období délky $m=n/3$. Utvoříme tři částečné součty empirických hodnot a položíme tyto součty rovny součtům teoretických hodnot.

$$\begin{aligned} S_1 &= \sum_{t=1}^m y_t = \sum_{t=1}^m (b_0 + b_1 b_2^t), \\ S_2 &= \sum_{t=1}^m y_{t+m} = \sum_{t=1}^m (b_0 + b_1 b_2^{t+m}), \\ S_3 &= \sum_{t=1}^m y_{t+2m} = \sum_{t=1}^m (b_0 + b_1 b_2^{t+2m}). \end{aligned} \quad (2.17)$$

Řešením soustavy (2.17) získáme odhady jednotlivých parametrů.

$$b_2 = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{1/m}, \quad (2.18)$$

$$b_1 = \frac{b_2 - 1}{b_2 (b_2^m - 1)^2} (S_2 - S_1), \quad (2.19)$$

$$b_0 = \frac{S_1 - \frac{b_1 b_2 (b_2^m - 1)}{b_2 - 1}}{m}. \quad (2.20)$$

- **Iterační metody**

postupně zlepšují výchozí odhad parametrů, které jsme získali např. metodou částečných součtů nebo metodou vybraných bodů. Přitom si buď předem stanovíme počet kroků (iterací), které pro zlepšení odhadů provedeme, nebo ukončíme proces postupného zlepšování v okamžiku, kdy je absolutní hodnota rozdílu vektorů parametrů ve dvou po sobě následujících iteracích menší, než předem zvolené číslo ε .

Užití iteračních metod (jsou velmi složité) je podmíněné statistickým softwarem.

Při volbě trendové funkce i při následném hodnocení výstižnosti konkrétní vypočtené funkce vycházíme z následujících informací:

- **věcný rozbor vývoje** sledovaného jevu,
- **vizuální posouzení grafu**, srovnání průběhu časové řady a průběhu trendové čáry,
- **analýza popisných charakteristik** vývoje časové řady,
- **hodnoty interpolačních a extrapolačních kritérií.**

2.1.3 Volba vhodného modelu trendové funkce

Jak bylo uvedeno v minulém odstavci, při výběru modelu trendu bychom měli vycházet především z věcné analýzy údajů v časové řadě. Výběr vhodného modelu trendu usnadní analýza diferencí a popisných charakteristik zkoumané časové řady – viz tab. 6.

Tab. 6 Popisné charakteristiky jako kritérium volby trendu časové řady

Trend	Test
Lineární	První diference jsou přibližně konstantní.
Kvadratický	Druhé diference jsou přibližně konstantní.
Ryzí exponenciální, resp. modifikovaná exponenciální	Koeficienty růstu jsou přibližně konstantní.
Logistický	Křivka prvních diferencí se podobá křivce hustoty normálního rozdělení.

Testy uvedené v tab. 6 jsou sice výpočetně jednoduché, ale málo použitelné při rozhodování mezi složitějšími typy trendových křivek. Pokud analýza popisných charakteristik nevede k jednoznačnému výsledku, vybereme několik trendových funkcí a jejich vhodnost ověřujeme až po odhadnutí parametrů. Používáme dva druhy kritérií - **interpolační kritéria** a **indexy** popisující kvalitu vytvořeného modelu.

Interpolační kritéria jsou založena na zkoumání vztahu skutečných hodnot y_t a tzv. hodnot interpolovaných \hat{Y}_t , tj. hodnot odhadnutých na základě dané trendové funkce.

- **střední chyba odhadu** (průměr hodnot reziduí – M.E.)

$$M.E. = \frac{\sum_{t=1}^n (y_t - \hat{Y}_t)}{n}, \quad (2.21)$$

- **střední kvadratická chyba odhadu** (průměr čtverců reziduí – M.S.E.)

$$M.S.E. = \frac{\sum_{t=1}^n (y_t - \hat{Y}_t)^2}{n}, \quad (2.22)$$

Protože kritérium (2.22) zvyhodňuje víceparametrické funkce, užívá se

- **upravená střední kvadratická chyba** ($M.S.E._M$)

$$M.S.E._M = \frac{\sum_{t=1}^n (y_t - \hat{Y}_t)^2}{n - c}, \quad (2.23)$$

kde c je počet parametrů trendové funkce. Podle vzorce (2.23) provádí výpočet MSE program STATISTICA 10.

- **střední absolutní chyba odhadu** (M.A.E.)

$$M.A.E. = \frac{\sum_{t=1}^n |y_t - \hat{Y}_t|}{n}, \quad (2.24)$$

- **střední procentuální chyba odhadu** (M.P.E.)

$$M.P.E. = \frac{1}{n} \sum_{t=1}^n \frac{(y_t - \hat{Y}_t)}{y_t} \cdot 100, \quad (2.25)$$

- **střední absolutní procentuální chyba** (M.A.P.E.)

$$M.A.P.E. = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{Y}_t|}{y_t} \cdot 100. \quad (2.26)$$

Čím menší je hodnota statistik (2.21) – (2.26), tím je model lepší. Z uvedených kritérií se **nejčastěji** používá kritérium (2.23). **Obecně dáváme přednost modelu, u něhož je hodnota $M.S.E._M$ nejnižší.**

Další interpolační kritérium, které používáme u funkcí lineárních v parametrech, je

➤ **index determinace** (R^2)

$$R^2 = \frac{\sum_{t=1}^n (\hat{Y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2}. \quad (2.27)$$

Trendová funkce je tím vhodnější, čím je její index determinace bližší jedné. Nedostatkem indexu determinace je jeho závislost na počtu parametrů zvolené funkce.

Proto při rozhodování mezi různými (nejčastěji polynomickými) funkcemi počítáme

➤ **upravený index determinace** (R_{adj}^2)

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n-1}{n-c}, \quad (2.28)$$

kde n je počet pozorování a $c = p+1$ je počet parametrů regresní funkce.

Dalším interpolačním kritériem je

➤ **celkový F-test.**

Testové kritérium

$$F = \frac{\frac{\sum_{t=1}^n (\hat{Y}_t - \bar{y})^2}{c-1}}{\frac{\sum_{t=1}^n (y_t - \hat{Y}_t)^2}{n-c}}, \quad (2.29)$$

má při platnosti nulové hypotézy F- rozdělení s $\nu_1 = c-1$ a $\nu_2 = n-c$ stupni volnosti.

Kritérium (2.29) je podíl vysvětlené variability a nevysvětlené variability. Při srovnávání více modelů vybereme ten, pro který je F- statistika (2.22) největší.

Při testování hypotézy, že rezidua (odhady náhodných poruch) nejsou autokorelovaná používáme

➤ **Durbinův-Watsonův test.**

Testové kritérium má tvar

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}, \quad (2.30)$$

kde $e_t = y_t - \hat{Y}_t$ jsou rezidua.

Statistika (2.30) nabývá hodnot od 0 do 4. V případě nezávislosti reziduí se pohybuje okolo čísla 2, v případě přímé závislosti jsou její kladné hodnoty blízké 0 a v případě nepřímé závislosti se blíží zleva 4. Kritické hodnoty testového kritéria (2.22) jsou uvedeny ve speciálních tabulkách.

Chceme-li konstruovat prognózy vývoje sledovaného ukazatele, používáme při rozhodování o modelu trendové funkce **extrapolační kritéria**. Nejčastější způsob použití extrapolačních kritérií je založen na **simulaci**. Zkoumaná časová řada se nejprve zkrátí a provede se volba trendové funkce. Zvolená trendová funkce se použije pro výpočet předpovědí, jenž se následně porovnají se skutečnými hodnotami časové řady. Nejčastěji používané extrapolační kritérium je

➤ **Theilův koeficient nesouladu**

$$T^2 = \frac{\sum_{j=1}^m (y_{n-m+j} - \hat{P}_j)^2}{\sum_{j=1}^m y_{n-m+j}^2}, \quad (2.31)$$

kde $(n-m)$ je délka zkrácené časové řady, $j=1, 2, \dots, m$, \hat{P}_j je předpověď na j -období dopředu modelem vypočteným z $(n-m)$ hodnot. Je-li $0 \leq T^2 \leq 1$, je daný model z extrapolačního hlediska vhodný. Je-li $T^2 > 1$, je třeba volit jiný model. Při srovnání více modelů preferujeme model, který má menší koeficient nesouladu.

Příklad 2.1

Bylo změřeno procentuální zlepšení startovní reakce v závislosti na počtu tréninkových jednotek. Sestavte regresní model, a vyjádřete přesnost modelu.

Tab. 7 Vstupní data kvadratického regresního modelu

Pořadové číslo	počet tréninků	reakce
1	0	1,000
2	10	1,000
3	20	0,997
4	30	0,996
5	40	0,993
6	50	0,985
7	60	0,983
8	70	0,978
9	80	0,973
10	90	0,961
11	100	0,958

Tab. 8 Výsledky kvadratické regrese

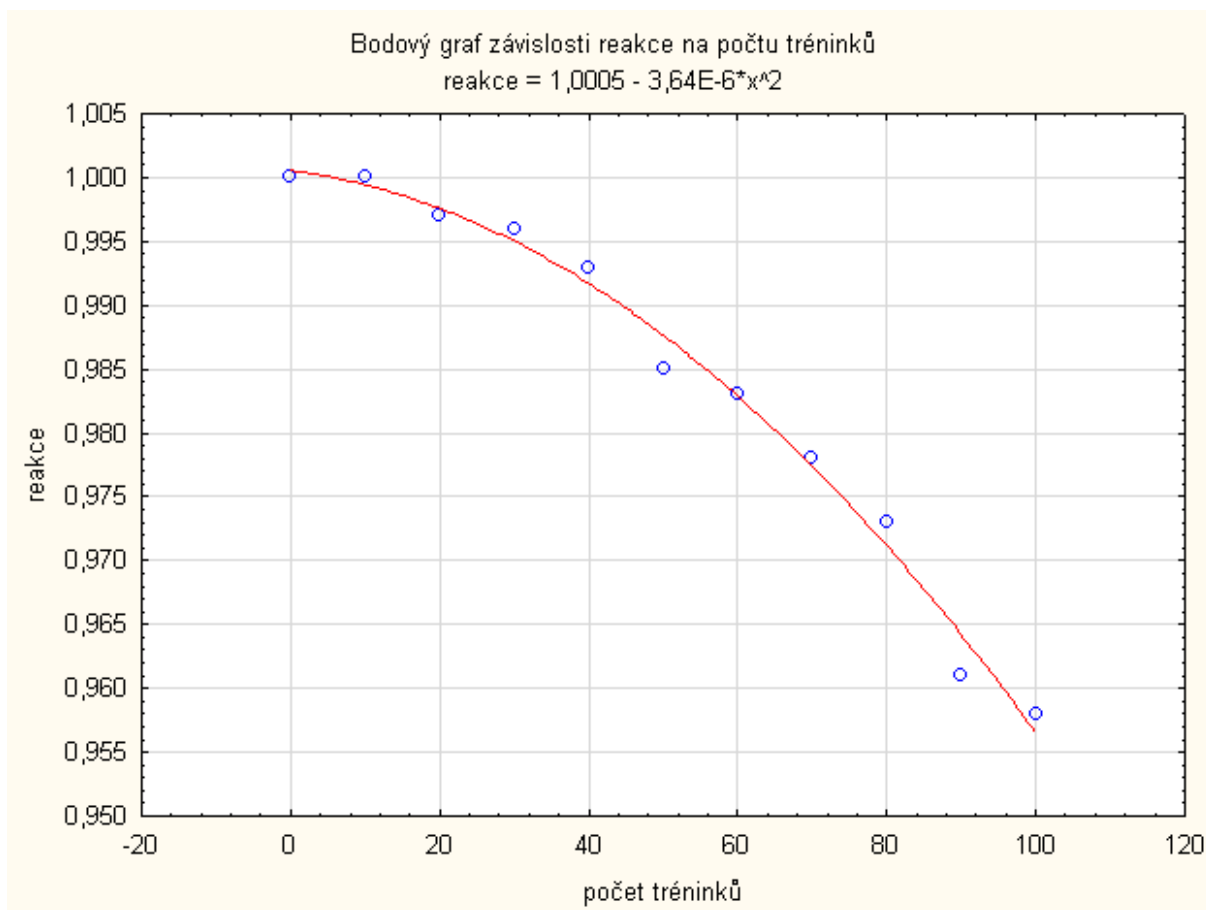
Efekt	Odhady parametrů (reakce) Sigma-omezená parametrizace									
	reakce Param.	reakce Sm.Ch.	reakce t	reakce p	-95,00% LmtSpol.	+95,00% LmtSpol.	reakce Beta (β)	reakce Sm.Ch. β	-95,00% LmtSpol.	+95,00% LmtSpol.
Abs. člen	1,000545	0,001398	715,5962	0,000000	0,997321	1,003770				
počet tréninků	-0,000076	0,000065	-1,1739	0,274204	-0,000226	0,000074	-0,168101	0,143201	-0,49832	0,162122
počet tréninků ²	-0,000004	0,000001	-5,8038	0,000403	-0,000005	-0,000002	-0,831113	0,143201	-1,16134	-0,500890

Závislé Proměnná	Test SČ celého modelu vs. SČ reziduí (reakce)										
	Vícenás. R	Vícenás. R2	Upravené R2	SČ Model	sv Model	PČ Model	SČ Rezid.	sv Rezid.	PČ Rezid.	F	p
reakce	0,994047	0,988130	0,985162	0,002243	2	0,001122	0,000027	8	0,000003	332,9771	0,000000

Nejvýhodnější je kvadratický regresní model ve tvaru:

$$r = 1,0005 (\pm 1,3982 \cdot 10^{-3}) - 3,64 \cdot 10^{-6} (6,3 \cdot 10^{-7}) t^2$$

s koeficientem determinace $R^2 = 0,98813$, $MEP = 5,0464 \cdot 10^{-6}$, $s_e = 1,8353 \cdot 10^{-3}$



Obr. 3 Graf kvadratického regresního modelu

Příklad 2.2

Úspěšnost v podávání grantů na FSpS MU v letech 2004 - 2011 (v % z celkového počtu všech podaných grantů)

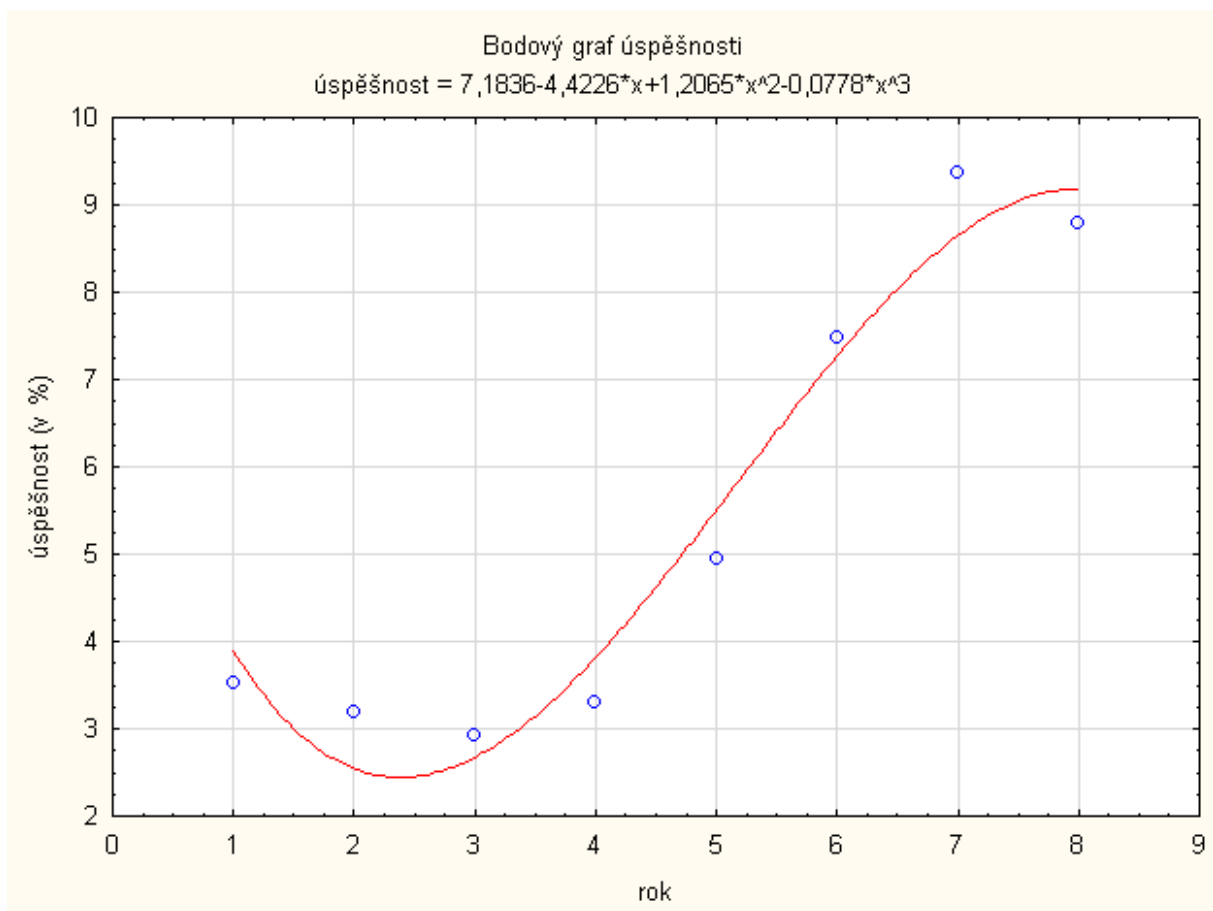
Tab. 9 Vstupní data

rok	úspěšnost
2004	3,52
2005	3,19
2006	2,93
2007	3,31
2008	4,94
2009	7,48
2010	9,37
2011	8,78

Tab. 10 Výsledky kubické regrese

Efekt	Odhady parametrů (data-regrese-3) Sigma-omezená parametrizace									
	počet Param.	počet Sm.Ch.	počet t	počet p	-95,00% LmtSpol.	+95,00% LmtSpol.	počet Beta (β)	počet Sm.Ch. β	-95,00% LmtSpol.	+95,00% LmtSpol.
Abs. člen	7,18357	1,704605	4,21421	0,013540	2,45083	11,91631				
rok	-4,42256	1,542414	-2,86730	0,045593	-8,70499	-0,14013	-4,02932	1,405268	-7,9310	-0,12767
rok^2	1,20652	0,386893	3,11847	0,035579	0,13233	2,28070	10,13445	3,249816	1,1115	19,15739
rok^3	-0,07783	0,028385	-2,74190	0,051804	-0,15664	0,00098	-5,32667	1,942690	-10,7204	0,06710

Závislé Proměnná	Test SČ celého modelu vs. SČ reziduí (data-regrese-3)										
	Vícenás. R	Vícenás. R2	Upravené R2	SČ Model	sv Model	PČ Model	SČ Rezid.	sv Rezid.	PČ Rezid.	F	p
počet	0,980901	0,962166	0,933790	48,68367	3	16,22789	1,914329	4	0,478582	33,90825	0,002650



Obr. 4 Graf kubického regresního modelu

Nejvýhodnější je kubický regresní model ve tvaru:

$$\text{úspěšnost} = 7,1836 - 4,4226 x + 1,2065 x^2 - 0,0778 x^3$$

s koeficientem determinace $R^2 = 0,962$

3 ZÁKLADY A APARÁT BOXOVY-JENKINSONOVY METODOLOGIE

Tato kapitola se zabývá jedním z mnoha způsobů rozboru časových řad. V první kapitole je popsána Boxova-Jenkinsonova metodologie. Tato je založena na myšlence, že časová řada může být chápána jako řada stochastického charakteru. Na jejím základě lze modelovat systematickosti v reziduální složce. Na druhé straně je možné každou časovou řadu chápat jako realizaci stochastického procesu a Boxova-Jenkinsonova metodologie se stává relativně samostatným přístupem k analýze časových řad.

3.1 Základní pojmy

Stochastický proces $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$, resp. $\{X_t\}$ je třída v čase t uspořádaných náhodných veličin. Časovou řadou rozumíme v čase t uspořádanou řadu hodnot stochastického procesu. Je to tedy náhodný výběr získaný tak, že z každého rozdělení náhodné veličiny je vybrána právě jedna hodnota. Někdy se hovoří o *výběrové funkci* nebo také *realizaci* daného stochastického procesu.

3.1.1 Stacionarita

Každá náhodná veličina má jisté pravděpodobnostní rozdělení, současně však také každá kombinace náhodných veličin má jisté pravděpodobnostní rozdělení, a tak i skupina náhodných veličin má pravděpodobnostní rozdělení (n -rozměrné). Pravděpodobnostní rozdělení lze charakterizovat distribuční funkcí, takže například n -rozměrná distribuční funkce je dána vztahem

$$F(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = P(X_{t_1} < x_{t_1}, X_{t_2} < x_{t_2}, \dots, X_{t_n} < x_{t_n}),$$

tj. pravděpodobností, že všechny náhodné veličiny nabudou současně hodnoty menší než jsou hodnoty x_{t_i} , $i = 1, \dots, n$.

Stochastický proces se nazývá *striktně stacionární*, jestliže pro jakoukoliv indexní část (t_1, t_2, \dots, t_n) z $\mathbf{T} = \{0, \pm 1, \pm 2, \dots\}$ a jakékoliv reálné číslo k , pro které $t_i + k \in \mathbf{T}$, $i = 1, 2, \dots, n$ platí

$$F(X_{t_1}, X_{t_2}, \dots, X_{t_n}) = F(X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}),$$

kde $F(\cdot)$ je n -rozměrná distribuční funkce. Striktní stacionarita tedy říká, že pravděpodobnostní chování daného stochastického procesu je časově invariantní.

Pro stochastický proces $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ definujeme funkci středních hodnot

$$\mu_t = E(X_t)$$

variační funkci

$$\sigma_t^2 = D(X_t) = E(X_t - \mu_t)^2$$

kovarianční funkci mezi X_{t_i} a X_{t_j} , $i, j = 1, 2, \dots, n; i \neq j$

$$\gamma(t_i, t_j) = E(X_{t_i} - \mu_{t_i})(X_{t_j} - \mu_{t_j})$$

a korelační funkci mezi X_{t_i} a X_{t_j} , $i, j = 1, 2, \dots, n; i \neq j$

$$\rho(t_i, t_j) = \frac{\gamma(t_i, t_j)}{\sigma_{t_i} \sigma_{t_j}}$$

Platí-li pro všechna t , že $\mu_t = \mu$, $\sigma_t^2 = \sigma^2$ a kovarianční a korelační funkce závisí pouze na časové vzdálenosti náhodných veličin, tj. pro $t_i = t - k$ a $t_j = t$, pak

$$\gamma(t_i, t_j) = \gamma(t - k, t) = \gamma(t, t + k) = \gamma_k$$

a

$$\rho(t_i, t_j) = \rho(t - k, t) = \rho(t, t + k) = \rho_k$$

potom se daný proces nazývá *slabě stacionární* nebo také *kovariančně stacionární*.

Striktně stacionární proces, který má první dva obecné momenty konečné, je také slabě stacionární proces. V praktické analýze časových řad se operuje výhradně se slabou stacionaritou, neboť je relativně snadné odhadovat první dva momenty. V případě, že budeme dále hovořit o stacionaritě, budeme mít na mysli právě stacionaritu slabou.

3.1.2 Autokorelační funkce (ACF)

V případě stacionárního stochastického procesu $\{X_t\}$ lze vyjádřit autokovarianční funkci mezi veličinami X_t a X_{t-k} jako

$$\gamma_k = C(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu)$$

a autokorelační funkci jako

$$\rho_k = \frac{C(X_t, X_{t-k})}{\sqrt{D(X_t)} \sqrt{D(X_{t-k})}} = \frac{\gamma_k}{\gamma_0}$$

kde vzhledem ke stacionaritě procesu $D(X_t) = D(X_{t-k}) = \gamma_0$. Autokorelační funkce se označuje jako ACF.

V případě stacionárního stochastického procesu má autokorelační funkce následující vlastnosti:

- a) $\rho_0 = 1$.
- b) $|\gamma_k| \leq \gamma_0$; $|\rho_k| \leq 1$
- c) $\gamma_k = \gamma_{-k}$ a $\rho_k = \rho_{-k}$ pro všechna k , autokovarianční a autokorelační funkce je tedy symetrická kolem $k = 0$. Graf autokorelační funkce se nazývá korelogram.

3.1.3 Parciální autokorelační funkce (PACF)

Korelace mezi dvěma náhodnými veličinami je často způsobena tím, že obě tyto veličiny jsou korelovány s veličinou třetí. Velká část korelace mezi veličinami X_t a X_{t-k} může být tedy způsobena jejich korelací s veličinami X_{t-1} , X_{t-2} , ..., X_{t-k+1} . *Parciální autokorelace* podávají informaci o korelaci veličin X_t a X_{t-k} očištěné o vliv veličin ležících mezi nimi.

Parciální autokorelaci se zpožděním k vyjadřuje parciální regresní koeficient ϕ_{kk} v autoregresi k -tého řádu

$$X_t = \phi_{k1} X_{t-1} + \phi_{k2} X_{t-2} + \dots + \phi_{kk} X_{t-k} + e_t$$

kde veličina e_t je nekorelovaná s veličinami X_{t-j} , $j \geq 1$. Tato matice je vždy regulární, v důsledku čehož můžu použít Cramerovo pravidlo. Vynásobíme-li obě strany rovnice veličinou X_{t-j} , potom střední hodnota této rovnice má tvar

$$\gamma_j = \phi_{k1} \rho_{j-1} + \phi_{k2} \rho_{j-2} + \dots + \phi_{kk} \rho_{j-k}$$

takže platí

$$\rho_j = \phi_{k1} \rho_{j-1} + \phi_{k2} \rho_{j-2} + \dots + \phi_{kk} \rho_{j-k}$$

Pro $j = 1, 2, \dots, k$ dostaneme následující systém rovnic

$$\rho_1 = \phi_{k1} \rho_0 + \phi_{k2} \rho_1 + \dots + \phi_{kk} \rho_{k-1}$$

$$\rho_2 = \phi_{k1} \rho_1 + \phi_{k2} \rho_0 + \dots + \phi_{kk} \rho_{k-2}$$

...

$$\rho_k = \phi_{k1} \rho_{k-1} + \phi_{k2} \rho_{k-2} + \dots + \phi_{kk} \rho_0$$

Použijeme-li postupně pro $k = 1, 2, \dots$ Cramerovo pravidlo, získáme

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \\ &\vdots \\ \phi_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}, \end{aligned}$$

V literatuře se často parciální autokorelace označuje jako ϕ_{kk} . Jako funkce zpoždění k je ϕ_{kk} nazývána *parciální autokorelační funkcí* (PACF).

3.1.4 Výběrové ACF a PACF

Stacionární stochastické procesy umožňují provést odhady ACF a PACF, neboť časová řada se v tomto případě stává výběrem z jednoho pravděpodobnostního rozdělení. Předpokládejme časovou řadu $X_t, t = 1, 2, \dots, n$.

Odhadem střední hodnoty je výběrový průměr

$$\bar{X} = \frac{1}{n-1} \sum_{t=1}^n X_t$$

Podobně lze odhadnout autokovarianční funkci pomocí výběrové autokovarianční funkce

$$\hat{\gamma}_k = \frac{1}{n-1} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})$$

Výběrová autokorelační funkce (výběrová ACF) je definována jako

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^n (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}, k = 1, 2, \dots$$

Graf $\hat{\rho}_k$ proti k se nazývá výběrový korelogram.

Pro výpočet výběrové parciální autokorelační funkce navrhl Durbin rekurzivní vztah

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j}$$

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j} \quad j = 1, 2, \dots, k-1$$

Jako počáteční odhad se bere $\hat{\phi}_{11} = \hat{\rho}_1$.

3.2 Základní reprezentace stochastických procesů

3.2.1 Proces bílého šumu

Jestliže je stochastický proces $\{a_t\}$ řadou nekorelovaných náhodných veličin s konstantní střední hodnotou $E(a_t) = \mu_a$ (obvykle nulovou), konstantním rozptylem $D(a_t) = \sigma_a^2$ a $\gamma_k = C(a_t, a_{t-k}) = 0$, pro všechna $k \neq 0$, potom se označuje jako *proces bílého šumu*.

Z této definice vyplývá, že proces bílého šumu $\{a_t\}$ je stacionární, s autokorelační funkcí

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

a parciální autokorelační funkcí

$$\phi_{kk} = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

Vzhledem k tomu, že podle definice platí $\rho_0 = \phi_{00} = 1$ pro jakýkoliv proces, budeme uvažovat autokorelace a parciální autokorelace pro $k \neq 0$. Základním rysem procesu bílého šumu tedy je, že ACF a PACF jsou identicky nulové. I když se tento proces prakticky nevyskytuje, hraje důležitou roli jako základní stavební prvek při výstavbě modelů časových řad.

3.2.2 Woldova a Box-Jenkinsova reprezentace stochastického procesu

Wold (1938) dokázal, že stacionární proces, který neobsahuje žádnou deterministickou složku (rozumí se složku, kterou by bylo možné přesně predikovat na základě jejího minulého vývoje nějakou funkcí) lze vyjádřit jako lineární kombinaci řady nekorelovaných náhodných veličin ve formě

$$X_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j} \quad (3.1)$$

kde $\psi_0, \psi_1, \psi_2, \dots$ jsou parametry, $\psi_0 = 1$, $\{a_t\}$ je proces bílého šumu a $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Jestliže zavedeme tzv. *operátor zpětného posunutí* B , který je definován jako $BX_t = X_{t-1}$ a tedy $B^j X_t = X_{t-j}$, potom lze vztah (3.1) zapsat zkráceně jako

$$X_t = \mu + \psi(B) a_t$$

kde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \psi_j B^j$$

Vztah (3.1) se nazývá MA („moving average“) reprezentace stochastického procesu (reprezentace klouzavých průměrů) nebo také *lineární proces*.

Lineární proces může být také vyjádřen ve formě AR („autoregressive“) reprezentace (autoregresivní reprezentace), které se někdy také říká Box-Jenkinsova reprezentace. Tato reprezentace má tvar

$$X_t - \mu = \pi_1 (X_{t-1} - \mu) + \pi_2 (X_{t-2} - \mu) + \dots + a_t \quad (3.2)$$

resp.

$$\pi(B)(X_t - \mu) = a_t$$

kde π_1, π_2, \dots jsou parametry autoregresivního modelu, $\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$ a $\sum_{j=1}^{\infty} |\pi_j| < \infty$.

Lineární proces (proces MA) lze zapsat také ve tvaru (3.2), tj. jeho současnou hodnotu lze vyjádřit pomocí minulých hodnot a současné hodnoty bílého šumu, takový proces se nazývá *invertibilní*. Avšak ne každý lineární (stacionární) proces může být invertibilní. Aby

bylo možné zapsat MA reprezentaci ve formě AR reprezentace, musí kořeny polynomu $\psi(B)$, což jsou hodnoty proměnné B , ležet vně jednotkového kruhu, tj. jestliže β je kořenem polynomu $\psi(B)$, potom $|\beta| > 1$. A naopak, aby byl AR proces stacionární, musí jej být možné zapsat ve formě MA reprezentace, tj.

$$\pi^{-1}(B) = \psi(B)$$

přičemž musí být splněna podmínka $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Aby toto bylo splněno, je nutné, aby kořeny polynomu $\pi(B)$ ležely vně jednotkového kruhu.

3.3 Modely stacionárních časových řad

3.3.1 Autoregresní proces řádu p [AR(p)]

Autoregresní model řádu p je možné zapsat ve formě

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t$$

kde ϕ_1, \dots, ϕ_p jsou parametry autoregresivního modelu.

Pomocí operátoru zpětného posunutí jej lze zapsat jako

$$(1 - \phi_1 B - \dots - \phi_p B^p) X_t = a_t$$

tj.

$$\phi_p(B) X_t = a_t,$$

kde $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$.

Proces AR(p) je stacionární, jestliže kořeny polynomiální rovnice

$$(1 - \phi_1 B - \dots - \phi_p B^p) = 0$$

leží vně jednotkového kruhu.

Rozptyl procesu AR(p) má formu

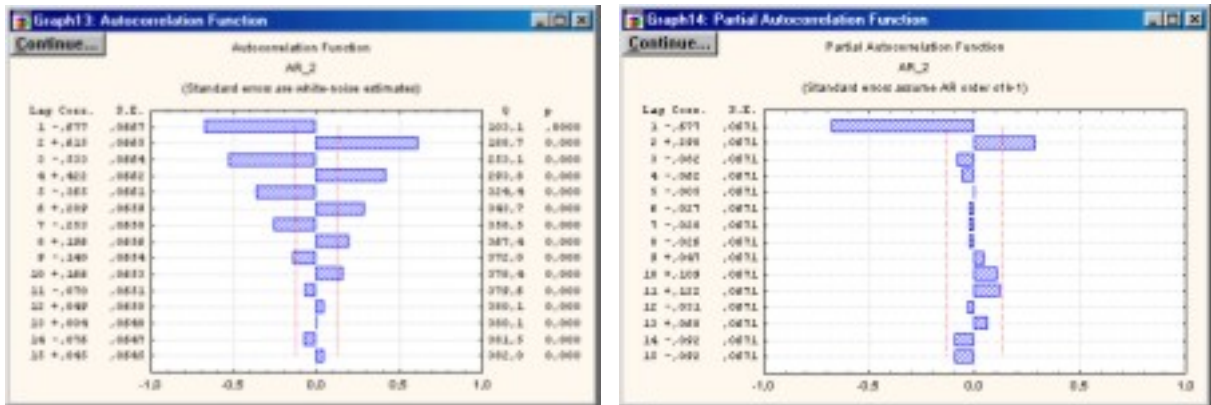
$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p}$$

Hodnoty autokorelační funkce lze získat na základě diferenční rovnice

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k = 1, 2, \dots$$

ACF tvoří kombinace exponenciálně klesajících pohybů (v případě reálných kořenů rovnice $\phi_p(B) = 0$) a exponenciálně klesajících sinusoidních pohybů (v případě komplexních

kořenů rovnice $\phi_p(B) = 0$). Hodnoty PACF pro zpoždění $k = 1, 2, \dots, p$ jsou různé od nuly, pro další zpoždění se potom rovnají nule (viz obr. 5).



Obr. 5 Výběrová ACF a PACF procesu AR(2)

3.3.2 Proces klouzavých průměrů řádu q [MA(q)]

Proces klouzavého průměru řádu q je možné zapsat ve formě

$$X_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

kde $\theta_1, \dots, \theta_q$ jsou parametry modelu klouzavých průměrů, nebo také

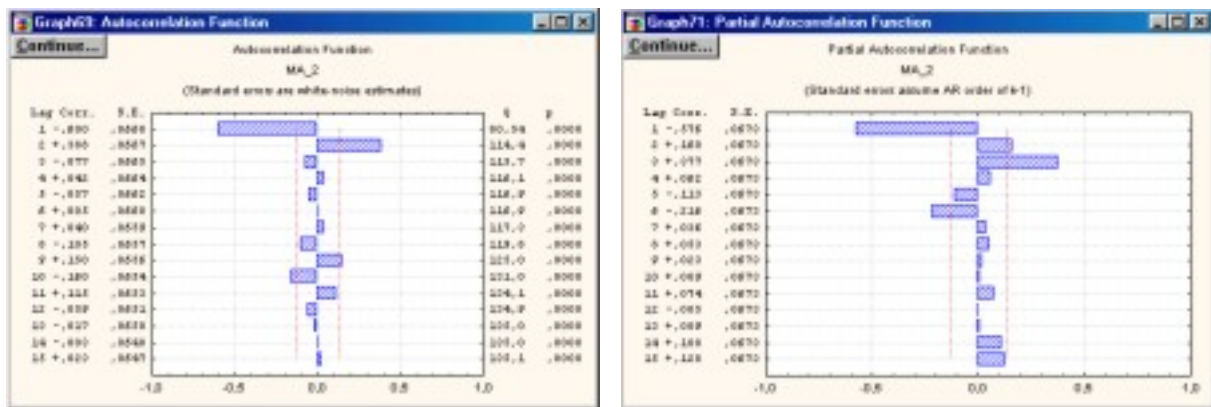
$$X_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta_q(B) a_t$$

Proces MA(q) je vždy stacionární, protože $\sum_{i=1}^q \theta_i^2 < \infty$. Jeho střední hodnota je nulová a rozptyl je roven $\gamma_0 = (1 + \theta_1^2 + \dots + \theta_p^2) \sigma_a^2$, kde σ_a^2 je rozptyl bílého šumu.

Autokorelační funkce procesu MA(q) má tvar

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & k = 1, 2, \dots, q \\ 0, & k > q \end{cases}$$

Hodnoty ACF jsou tedy pro hodnoty $k = 1, 2, \dots, q$ různé od nuly, pro další zpoždění se potom rovnají nule. PACF tvoří kombinace exponenciálně klesajících pohybů (v případě reálných kořenů rovnice $\theta_q(B) = 0$) a exponenciálně klesajících sinusoidních pohybů (v případě komplexních kořenů rovnice $\theta_q(B) = 0$) (viz obr. 6).



Obr. 6 Výběrová ACF a PACF procesu MA(2)

3.3.3 Smíšené procesy ARMA [ARMA(p, q)]

Proces ARMA(p, q) je možné vyjádřit ve tvaru

$$\phi_p(B)X_t = \theta_q(B)a_t$$

Lineární proces MA(∞) lze vyjádřit jako

$$X_t = \psi(B) a_t, \text{ kde } \psi(B) = \frac{\theta_q(B)}{\phi_p(B)}$$

proces AR(∞) lze vyjádřit jako

$$\pi(B) X_t = a_t, \text{ kde } \pi(B) = \frac{\phi_p(B)}{\theta_q(B)} = \psi(B)^{-1}$$

Aby byl proces ARMA(p, q) stacionární, musí kořeny rovnice $\phi_p(B) = 0$ ležet vně jednotkového kruhu, aby byl invertibilní, musí kořeny rovnice $\theta_q(B) = 0$ ležet vně jednotkového kruhu.

Autokovariance a autokorelace se získají z rovnic

$$\gamma_k - \phi_1 \gamma_{k-1} - \dots - \phi_p \gamma_{k-p} = 0 \text{ pro } k > q$$

a

$$\rho_k - \phi_1 \rho_{k-1} - \dots - \phi_p \rho_{k-p} = 0 \text{ pro } k > q$$

Z těchto rovnic plyne, že tvar ACF procesu ARMA(p, q) bude obdobný jako v případě procesu AR(p), tj. bude mít formu kombinace exponenciálně klesajících pohybů a exponenciálně klesajících sinusoidních pohybů. Tento tvar však bude následovat až po prvních $q - p$ hodnotách jestliže $q > p$. Hodnoty $\rho_0, \rho_1, \dots, \rho_{q-p}$ tento tvar mít nebudou.

Pro $k > p$ se PACF u procesu ARMA(p, q) bude v případě, že $p > q$, chovat stejně jako u procesu MA(q). Pro $k \leq p - q$ je však tvar této funkce odlišný.

3.4 Modely nestacionárních a sezónních časových řad

3.4.1 Nestacionární procesy ve střední hodnotě

Až dosud jsme všechny úvahy vedli o procesech stacionárních v kovariancích, tj. o procesech s konstantní střední hodnotou a rozptylem a s kovariancí závisící pouze na vzdálenost $|t - k|$. Stacionární procesy se však v realitě vyskytují zřídka, zejména v ekonomické praxi existují téměř vždy procesy nestacionární. Nestacionarita procesu může být způsobena v čase se měnící střední hodnotou procesu či v čase se měnícím rozptylem procesu.

Procesy AR a ARMA jsou stacionární, leží-li všechny kořeny polynomu $\phi(B)$ vně jednotkového kruhu. Není-li tato podmínka splněna, tj. leží-li alespoň jeden kořen na jednotkovém kruhu, proces není stacionární, ale obsahuje stochastický trend. Platí následující: leží-li jeden kořen na jednotkové kružnici a ostatní jsou vně, proces je stacionární po první diferenci, leží-li dva kořeny na jednotkové kružnici a ostatní jsou vně, proces je stacionární po druhé diferenci atd.

Nestacionární procesy se stochastickým trendem mohou být na stacionární převedeny diferencováním. Proces, který neobsahuje po d -té diferenci žádnou deterministickou složku (trend, sezónní složku) a která lze popsat stacionární a invertibilní reprezentací ARMA, se nazývá integrovaným procesem d -tého řádu a značí se $I(d)$.

Proces, jenž je možné po d -té diferenci popsat stacionární a invertibilní reprezentací ARMA(p, q) lze zapsat jako

$$\phi_p(B)(1-B)^d X_t = \theta_0 + \theta_q(B)a_t \quad (3.3)$$

kde $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ je autoregresní operátor, $\theta_p(B) = 1 - \theta_1 B - \dots - \theta_p B^p$ je operátor klouzavých průměrů, $(1-B)^d X_t$ je stacionární stochastický proces dosažený d -tou diferencí. Je zřejmé, že polynom $(1-B)^d X_t = (1-B)(1-B) \dots (1-B) X_t$ má d jednotkových kořenů (kořenů ležících na jednotkovém kruhu), takže pro $d \geq 1$ proces obsahuje stochastický polynomický trend řádu d . θ_0 je volný parametr, který má různou úlohu pro $d = 0$ a $d \geq 1$.

V případě $d = 0$ (stacionární proces) je θ_0 střední hodnotou daného procesu. V případě nestacionarity ($d \geq 1$) se tato střední hodnota stává měnlivou v závislosti na čase a stává se tak lineárním, parabolickým atd. deterministickým trendem. Model (3.3) se označuje jako $ARIMA(p, d, q)$ (Autoregressive Integrated Moving Average). Jestliže $d = 0$, $ARIMA(p, d, q)$ se redukuje na $ARMA(p, q)$, v případě, že $p = 0$, redukuje se na $IMA(d, q)$ a v případě $q = 0$ na $ARI(p, d)$.

3.4.2 Proces náhodné procházky („random walk“)

Zvláštním případem je proces $ARIMA(p, d, q)$, kde $p = 0, q = 0$. Je charakteristickým tím, že po d -té diferenci z něho vznikne proces bílého šumu. Lze jej zapsat jako

$$(1-B)^d X_t = a_t$$

Prakticky je významný zejména proces *náhodné procházky*, kdy $d = 1$, takže k získání procesu bílého šumu stačí první diference. Má formu

$$(1-B) X_t = a_t$$

tedy

$$X_t = X_{t-1} + a_t$$

Vzhledem k tomu, že

$$(1-B)^{-1} = (1 + B + B^2 + \dots)$$

lze proces $(1-B) X_t = a_t$ přepsat do formy

$$X_t = (1 + B + B^2 + \dots) a_t = a_t + a_{t-1} + a_{t-2} + \dots$$

Proces náhodné procházky je limitním případem procesu $AR(1)$ pro $\phi \rightarrow 1$. Hodnoty ACF procesu náhodné procházky budou klesat velmi pomalu. Proces náhodné procházky je tvořen kumulováním náhodných veličin tvořících proces bílého šumu. Náhodná procházka je *nestacionární proces*. Zdrojem její nestacionarity je *stochastický trend* $\sum_{j=1}^t a_j$.

Parciální autokorelační funkce procesu $AR(1)$ blíží se procesu náhodné procházky je logicky velmi podobná parciální autokorelační funkci procesu $AR(1)$.

3.4.3 Procesy nestacionární v rozptylu

Není-li splněna podmínka neměnnosti rozptylu v čase, je proces nestacionární v rozptylu. Může se stát, že i kovariance závisí na čase (nejen na k), stacionární proces ve střední hodnotě tak může mít časově závislý rozptyl i kovarianci.

V některých případech lze stacionarity (časově nezávislá střední hodnota, rozptyl, kovariance) dosáhnout pouze diferencemi. V mnoha případech tomu ale tak není. Potom je třeba provést vhodné transformace procesu. Často se stává, že stabilizací rozptylu je stabilizována i kovariance.

Obecně je možné použít Box-Coxovu transformaci (Box-Cox, 1964), která má následující tvar

$$T(X_t) = \begin{cases} X_t^{(\lambda)} = \frac{X_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln(X_t) & \lambda = 0 \end{cases}$$

kde λ je tzv. transformační parametr. Tento parametr je odhadován metodou maximální věrohodnosti (Anděl, 1976): jde o to odhadnout takové λ , které vede k transformaci časové řady minimalizující reziduální součet čtverců zvoleného modelu.

3.4.4 Sezónní procesy

Myšlenka sezónního procesu je následující: jako v případě klasického procesu ARIMA předpokládáme vzájemnou závislost mezi veličinami $\dots, X_{t-3}, X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2}, X_{t+3}, \dots$. Protože tento proces obsahuje ještě sezónní kolísání, lze očekávat i závislost mezi sobě odpovídajícími veličinami v jednotlivých sezónách, tj. mezi veličinami $\dots, X_{t-2s}, X_{t-1s}, X_{ts}, X_{t+1s}, X_{t+2s}, \dots$, kde s je délka sezónní periody (u měsíčních časových řad 12, u čtvrtletních 4 atd.).

Předpokládejme, že proces obsahuje oba typy závislostí. Závislost uvnitř period je potom zachycena klasickým ARIMA modelem

$$\phi_p(B)(1-B)^d X_t = \theta_0 + \theta_q(B) b_t \quad (3.4)$$

kde θ_0 charakterizuje deterministický trend v případě, že $d \neq 0$. Proces $\{b_t\}$ obsahuje závislost mezi periodami, tento proces může být popsán modelem

$$\Phi_p(B^s)(1-B^s)^D b_t = \Theta_q(B^s) a_t \quad (3.5)$$

kde $\Phi_p(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_p B^{ps}$ je sezónní autoregresivní operátor a $\Theta_q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_q B^{qs}$ je sezónní operátor klouzavých průměrů. Prostřednictvím členu $(1-B^s)$ se konstruuji sezónní diference, $\{a_t\}$ je proces bílého šumu. Jestliže se procesy (3.4) a (3.5) spojí, získá se proces

$$\Phi_p(B^s) \phi_p(B) (1-B)^d (1-B^s)^D X_t = \theta_0 + \theta_q(B) \Theta_q(B^s) a_t \quad (3.6)$$

který je označován jako SARIMA(p, d, q)(P, D, Q) $_s$, kde p je řád procesu AR, q je řád procesu MA, d je řád prosté diference, P je řád sezónního procesu AR, Q je řád sezónního procesu MA, D je řád sezónní diference, s je délka periody.

Základní rysy autokorelační a parciální autokorelační funkce procesu (3.6) jsou podobné jako v případě procesu ARIMA s tím rozdílem, že tvar ACF a PACF se periodicky opakuje v modifikacích odpovídajících klasickému modelu ARIMA. Jestliže tedy v bodech k_i , které se pravidelně opakují po s sezónách, dosahuje ACF významných hodnot, jenž se od sebe liší jen málo (povolně klesají), je třeba uvažovat příslušnou sezónní diferenci. Pokud jsou i ostatní hodnoty významné, je třeba diferencí nesezónních, prostých. Jestliže je proces diferencován sezónně i prostě, tvoří ACF a PACF v první periodě stejné tvary jako v případě procesů nesezónních, v ostatních periodách se pak transformují v závislosti na typu sezónního procesu (SAR, SMA, SARMA, SARIMA). Tvary ACF a PACF jsou poměrně komplikované a lze je těžko obecně jednoznačně popsat. Složitost ACF a PACF je dána tím, že se zde současně projevuje působení čtyř druhů parametrů (AR, MA, SAR, SMA) a navíc se někdy jedná o procesy jenž jsou nestacionární.

3.5 Identifikace a ověřování modelu, odhady parametrů, konstrukce předpovědí

3.5.1 Identifikace modelu

Jednou z nejtěžších úloh při výstavbě Boxových-Jenkinsových modelů je jejich identifikace. Tato úloha spočívá v rozhodnutí, jaký typ modelu vybrat. Identifikace je pouze

první fázi konstrukce Boxových-Jenkinsových modelů. Výsledný model je dost často odlišný od modelu identifikovaného (tento model je třeba ještě ověřit a upravit). Jednotlivé kroky identifikace modelu $ARIMA(p, d, q)$ jsou:

- 1) Prozkoumání grafu časové řady. V mnoha případech je možné již na první pohled rozpoznat přítomnost trendu, často lze vidět i „outliers“ (odlehlá pozorování). V této fázi jde především o subjektivní zhodnocení situace. Nicméně, již nyní je vhodné stacionarizovat časovou řadu či provést jiné úpravy. Především je třeba odstranit odlehlá pozorování jejich vynecháním a provést transformaci, jež stabilizuje rozptyl. Následně se musí řada diferencovat a stabilizovat ve střední hodnotě. Vhodné je provést transformace ještě před stanovením řádu diferencování a vlastních diferencí.
- 2) Dalším krokem je výpočet odhadů ACF a PACF. Na jejich základě je možné revidovat, zda je řada diferencována dostatečně. Jak již bylo výše poznamenáno, pro řadu, která obsahuje trend, je charakteristické, že odhady hodnot ACF jsou vysoké a klesají velmi pomalu, zatímco u PACF je vysokých pouze několik málo prvních hodnot.
- 3) Po stacionarizaci řady se použijí odhady ACF a PACF pro identifikaci modelů AR a MA (nalezení hodnot p a q). Tato identifikace je založena na principu podobnosti odhadnutých ACF a PACF s teoretickými ACF a PACF. Následující tabulka obsahuje popis tvarů ACF a PACF pro modely AR, MA a ARMA

Tab. 11 Tvary ACF a PACF modelu AR, MA a ARMA

Model	ACF	PACF
$AR(p)$	exponenciálně klesající nebo sinusoida s exponenciálně klesající amplitudou	po p posunutích výrazně klesá
$MA(q)$	po q posunutích výrazně klesá	exponenciálně klesající nebo sinusoida s exponenciálně klesající amplitudou
$ARMA(p, q)$	po q posunutích jako $AR(p)$	po p posunutích jako $MA(q)$

Aby bylo možné konstruovat věrohodný model, je třeba mít alespoň $T = 50$ pozorování a je vhodné mít odhad ACF a PACF tvořený asi $T / 4$ hodnotami.

Obsahuje-li časová řada také sezónní složku, je třeba identifikovat model typu SARIMA tvaru (3.6). Princip je stejný jako v předchozím případě.

- 1) Nejprve je třeba časovou řadu stacionarizovat. Pokud je to nutné, stabilizuje se řada z hlediska rozptylu. Poté se řada diferencuje: nejprve se provádí sezónní diference (sezónní výkyvy jsou často patrné již na obrázku časové řady) a potom diference prosté.
- 2) V další fázi se vypočítají odhady ACF a PACF, pomocí kterých se určí typ sezónních modelů SAR, SMA, či SARMA. Identifikací těchto modelů se však ještě nekončí. Po

identifikaci a modelování sezónní složky zůstává dále stacionární řada, která může obsahovat ještě další systematickosti. Je tedy třeba ještě jednou vypočítat odhady ACF a PACF pro takto transformovanou řadu a posoudit, zda ji lze dále modelovat nesezónními AR, MA, či ARMA modely.

3.5.2 Odhady parametrů modelu

Pro identifikaci ARIMA modelu, tj. pro nalezení hodnot p , d , q je třeba odhadnout jeho parametry. Pro tento účel existuje několik postupů. Uvede zde klasickou metodu, která se stala základem pro metody ostatní. Procedura odhadu parametrů modelu ARIMA je založena na metodě maximální věrohodnosti.

Podmíněná metoda maximální věrohodnosti

Uvažujme obecný nesezónní model ARIMA(p , q , d) ve formě

$$\phi_p(B)(1-B)^d X_t = \theta_q(B) a_t \quad (3.7)$$

takže proces $(1-B)^d X_t$ je stacionární ve střední hodnotě μ a lze jej popsat stacionární a invertibilní reprezentací ARMA(p , q)

$$\phi_p(B) X_t^* = \theta_q(B) a_t, \text{ kde } X_t^* = (1-B)^d X_t, \quad (3.8)$$

$\{a_t\}$ je gausovský bílý šum, tj. jednotlivé veličiny jsou nekorelované a mají normální rozdělení s nulovou střední hodnotou a konstantním rozptylem σ_a^2 , $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ je autoregresivní operátor a $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ je operátor klouzavých průměrů.

Úlohou je odhadnout parametry $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \mu$ a σ_a^2 . Při odhadu těchto parametrů se vychází z předpokladu, že $a = (a_1, \dots, a_T)$ je gausovský bílý šum, tj. každá veličina $a_t = \phi^{-1}(B) \theta(B) X_t^*$, $t = 1, \dots, T$ ($T = n - d$ je počet pozorování po diferencování) má rozdělení $N(0, \sigma_a^2)$ a T -rozměrná veličina $a = (a_1, \dots, a_T)$ má hustotu pravděpodobnosti

$$f(a | \phi, \mu, \theta, \sigma_a^2) = (2\pi \sigma_a^2)^{-T/2} \cdot \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=1}^T a_t^2 \right] \quad (3.9)$$

kteřá je v tomto případě rovněž funkcí věrohodnostní, tj. $f(a | \phi, \mu, \theta, \sigma_a^2) = L(a | \phi, \mu, \theta, \sigma_a^2)$. Myšlenka odhadu parametrů metodou maximální věrohodnosti spočívá v nalezení takových odhadů parametrů $\phi, \mu, \theta, \sigma_a^2$, které při daném $a(\phi, \mu, \theta | X^*)$, $X^* = (X_1^*, \dots, X_T^*)$ maximalizují věrohodnostní funkci (3.9).

Věrohodnostní funkcí se nazývá sdružená hustota pravděpodobnosti náhodného výběru $Z = (Z_1, \dots, Z_T)$, jenž je při daném $z = (z_1, \dots, z_T)$ funkcí vektoru parametrů $\mathbf{G} = (G_1, \dots, G_T)$, lze ji zapsat ve formě $L(\mathbf{G}) = \prod_{j=1}^T f(z_j; \mathbf{G})$.

Řešení bude nalezeno, při maximalizaci logaritmu věrohodnostní funkce

$$\ln L(\phi, \mu, \theta, \sigma_a^2) = -\frac{T}{2} \ln 2\pi \sigma_a^2 - \frac{S(\phi, \mu, \theta)}{2\sigma_a^2}, \quad (3.10)$$

kde

$$S(\phi, \mu, \theta) = \sum_{t=1}^T a_t^2(\phi, \mu, \theta | X^*). \quad (3.11)$$

Maximální hodnoty věrohodnostní funkce se dosáhnou při minimální hodnotě reziduálního součtu čtverců, takže k odhadu parametrů ϕ, μ, θ stačí minimalizovat reziduální součet čtverců (3.11). Odhad reziduálního rozptylu $\hat{\sigma}_a^2$ se pak vypočítá jako podíl

$$\hat{\sigma}_a^2 = \frac{S(\hat{\phi}, \hat{\mu}, \hat{\theta})}{T - (p + q + 1)}. \quad (3.12)$$

Při této proceduře však vzniká následující problém.

Přepišme nyní proces (3.8) do tvaru

$$a_t = \phi_1 a_{t-1} + \dots + \phi_q a_{t-q} + X_t^* - \phi_1 X_{t-1}^* - \dots - \phi_p X_{t-p}^*, \quad (3.13)$$

z něhož je vidět, že pro rekurentní „odstartování“ procesu výpočtu veličin a_t , je třeba mít k dispozici dodatečné informace. To je zřetelnější, zapíše-li se rovnice pro první reziduální veličinu

$$a_1 = \phi_1 a_0 + \dots + \phi_q a_{1-q} + X_1^* - \phi_1 X_0^* - \dots - \phi_p X_{1-p}^*, \quad (3.14)$$

Podmínkou pro výpočet odhadů parametrů modelu jsou vhodně zvolené veličiny $a_* = (a_0, a_{-1}, \dots, a_{1-q})$ a $X^* = (X_0^*, X_{-1}^*, \dots, X_{1-p}^*)$; proto se hovoří o podmíněné metodě maximální věrohodnosti a (3.11) je podmíněný reziduální součet čtverců.

Existuje několik možností volby hodnot a_* a X^* . První vychází z předpokladu, že proces $\{X_t^*\}$ je stacionární a $\{a_t\}$ je gausovský proces bílého šumu, takže je možné neznámé hodnoty X_t^* nahradit průměrem $\overline{X_t^*}$ hodnot známých. Je-li $\mu = 0$, dosadí se nuly. Za neznámé hodnoty a_t se dosadí rovněž nuly. Leží-li ale kořeny polynomu $\phi_p(B)$ blízko jednotkové kružnice, může dojít k tomu, že skutečně naměřená hodnota X_t se od odhadů hodnot minulých značně liší, což by odhadovací proceduru zkreslilo. Další možností je položit rovné nule veličiny a_{p-q}, \dots, a_p . Potom se lze omezit pouze na výpočet veličin a_{p+1}, \dots, a_n , tj. a_t se počítá podle vzorce (3.13) až pro $t \geq (p+1)$, takže problém volby počátečních hodnot X_t zcela odpadá. Potom ale podmíněný reziduální součet čtverců má podobu

$$S_*(\phi, \mu, \theta) = \sum_{t=p+1}^T a_t^2(\phi, \mu, \theta | X^*). \quad (3.15)$$

a reziduální rozptyl se odhadne jako

$$\hat{\sigma}_a^2 = \frac{S_*(\hat{\phi}, \hat{\mu}, \hat{\theta})}{(T-p) - (p+q+1)}, \quad (3.16)$$

neboť odhad podmíněného reziduálního součtu čtverců závisí na $p+q+1$ odhadovaných parametrech v $(T-p)$ sčítancích.

Nepodmíněná metoda maximální věrohodnosti

Nepodmíněná metoda maximální věrohodnosti při odhadu neznámých hodnot X_0^*, X_{-1}^*, \dots atd. spočívá v následujícím postupu. Nejprve se položí $X_0^*, X_{-1}^*, \dots, X_{1-p}^*$ a $a_0, a_{-1}, \dots, a_{1-q}$ rovny nule. Potom se odhadne model ARMA pomocí podmíněné metody maximální věrohodnosti. Dále se použije odhadnutý model ARMA k odhadu hodnot X_0^*, X_{-1}^*, \dots následujícím způsobem. Model ARMA může být zapsán ve formě

$$(1 - \theta_1 B - \dots - \theta_p B^p) X_t^* = (1 - \theta_1 B - \dots - \theta_p B^p) a_t \quad (3.17)$$

nebo ve formě

$$(1 - \theta_1 F - \dots - \theta_p F^p) X_t^* = (1 - \theta_1 F - \dots - \theta_p F^p) e_t \quad (3.18)$$

kde F je inverzní operátor k operátoru B a nazývá se tzv. *operátor posunutí vpřed*, takže platí $F^j X_t = X_{t+j}$. Vzhledem ke stacionaritě procesu X_t^* , mají oba tyto modely stejnou autokovarianční strukturu, z čehož plyne, že $\{e_t\}$ je proces bílého šumu s nulovou střední hodnotou a konstantním rozptylem, tj. lze psát

$$X_t^* = \phi_1 X_{t+1}^* + \phi_2 X_{t+2}^* + \dots + \phi_p X_{t+p}^* + e_t - \theta_1 e_{t+1} - \dots - \theta_q e_{t+q} \quad (3.19)$$

takže v případě, že jsou k dispozici počáteční odhady parametrů modelu ARMA, a tedy podmíněné střední hodnoty $E(e_t | X_1^*, \dots, X_T^*)$, $t = 1, \dots, q$, odhady hodnot X_0^*, X_{-1}^*, \dots lze vypočítat následujícím způsobem

$$\begin{aligned} \hat{X}_0^* &= \hat{\phi}_1 X_1^* + \dots + \hat{\phi}_p X_p^* - \hat{\theta}_1 \hat{e}_1 - \dots - \hat{\theta}_q \hat{e}_q \\ \hat{X}_{-1}^* &= \hat{\phi}_1 X_0^* + \dots + \hat{\phi}_p X_{p-1}^* - \hat{\theta}_2 \hat{e}_1 - \dots - \hat{\theta}_q \hat{e}_{q-1} \\ \hat{X}_{-2}^* &= \hat{\phi}_1 X_{-1}^* + \hat{\phi}_2 X_0^* + \dots + \hat{\phi}_p X_{p-2}^* - \hat{\theta}_3 \hat{e}_1 - \dots - \hat{\theta}_q \hat{e}_{q-2} \\ &\vdots \end{aligned} \quad (3.20)$$

kde $E(e_t | X_1^*, \dots, X_T^*)$ se značí jako \hat{e}_t . Pomocí tzv. *zpětné extrapolace* získané odhady X_0^*, X_{-1}^*, \dots jsou podmíněné střední hodnoty $E(X_t^* | X_1^*, \dots, X_T^*)$, $t = 0, -1, -2, \dots$. Při výpočtu se pokračuje tak dlouho, až absolutní rozdíl $|E(X_t^* | X_1^*, \dots, X_T^*) - E(X_{t-1}^* | X_1^*, \dots, X_T^*)|$ je menší, než nějaká malá hodnota ε pro $t \leq -M$. Výpočtem hodnot X_0^*, X_{-1}^*, \dots vzhledem ke vztahu (3.13), jsou dány odhady reziduí $\hat{a}_t = E(a_t | X_{-M}^*, X_{-M+1}^*, \dots, X_T^*)$ pro kompletní řadu $X_1^*, X_2^*, \dots, X_T^*$.

Při odhadu parametrů se vychází opět z logaritmu věrohodnostní funkce, tj. provádí se minimalizace součtu čtverců podmíněných středních hodnot reziduí.

$$S(\phi, \mu, \theta) = \sum_{t=-M}^T [E(a_t | X_{-M}^*, X_{-M+1}^*, \dots, X_T^*)]^2. \quad (3.21)$$

Odhad reziduálního rozptylu se pak vypočítá jako

$$\hat{\sigma}_a^2 = \frac{S(\hat{\phi}, \hat{\mu}, \hat{\theta})}{T}.$$

Empiricky je prokázáno, že čím je časová řada kratší, tím má volba počátečních hodnot řady větší význam, takže má význam použít zpětné extrapolace. Jsou-li časové řady dlouhé, odhady získané na základě podmíněné a nepodmíněné metody maximální věrohodnosti jsou téměř stejné. Prakticky se ale doporučuje zpracovávat řady minimálně o 50 hodnotách. V tom případě výsledky odhadů na počátečních podmínkách prakticky nezávisí. Bylo prokázáno, že metodu zpětné extrapolace je výhodné používat především při odhadu parametrů sezónních modelů a modelů nestacionárních časových řad.

Metoda nejmenších čtverců

Výše bylo uvedeno, že pro odhad parametrů ϕ, μ, θ postačí minimalizovat reziduální součet čtverců $S(\phi, \mu, \theta)$. V této souvislosti se nabízí použít k odhadu parametrů metodu nejmenších čtverců.

Uvažujme na chvíli následující regresní model

$$Y_t = \beta Z_t + e_t \quad t = 1, \dots, n$$

Metoda nejmenších čtverců pro odhad β vychází z následujících podmínek o reziduální složce e_t (podmínky klasického lineárního regresního modelu):

1. $E(e_t) = 0$, střední hodnota je nulová,
2. $E(e_t^2) = \sigma_e^2$, rozptyl je konstantní,
3. $E(e_t e_k) = 0$ pro $t \neq k$, rezidua jsou nezávislá,
4. $E(Z_t e_k) = 0$, rezidua a vysvětlující proměnné jsou nezávislé.

$\{e_t\}$ je tedy proces bílého šumu, někdy se předpokládá, že $e_t \sim N(0, \sigma_e^2)$, tj. jde o gausovský bílý šum. Odhad parametrů β metodou nejmenších čtverců má za uvedených podmínek formu

$$\hat{\beta} = \frac{\sum_{t=1}^n Z_t Y_t}{\sum_{t=1}^n Z_t^2} \quad (3.22)$$

Lze dokázat, že jde o nejlepší (nejmenší rozptyl), nestranný, konzistentní odhad.

Vraťme se nyní k modelům našeho typu. Uvažujme stacionární model AR(1), který má formu

$$X_t = \phi X_{t-1} + e_t, \quad t = 1, \dots, n$$

Za podmínek klasického lineárního regresního modelu je možné získat odhad parametrů ϕ za pomoci metody nejmenších čtverců, ve tvaru

$$\hat{\phi} = \frac{\sum_{t=2}^n X_{t-1} X_t}{\sum_{t=2}^n X_{t-1}^2} \quad (3.23)$$

I tento odhad je lineární, nejlepší, nestranný a konzistentní.

Maximum věrohodnostní funkce se nachází ve stejném bodě (= vektor parametrů) v jakém je minimální reziduální součet čtverců.

Podívejme se na případ obecnější, na invertibilní model AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t \quad (3.24)$$

Opět předpokládejme, že budou splněny předchozí podmínky 1.-3., v případě 4. podmínky předpokládáme nezávislost e_t a X_t v každém uvažovaném posunu. Protože se však jedná o případ mnohonásobné lineární regrese, je třeba uvažovat ještě další podmínku, že vysvětlující proměnné jsou vzájemně lineárně nezávislé, tj. vzájemně nekorelované. Pokud tomu tak není, hovoří se o multikolinearitě, která způsobuje nadhodnocení odhadů směrodatných chyb odhadů parametrů ϕ_1, \dots, ϕ_p a tím zkreslení výsledků testů týkajících se těchto parametrů.

Až dosud jsme se v této části zabývali modely AR, kdy rezidua jsou bílým šumem (případně gaussovským bílým šumem), tj. $e_t = a_t$. Ale právě v případě regresních modelů uvažovaného typu, tomu tak často není a nastává situace, že rezidua jsou na sobě závislá, autokorelovaná. Není tak splněná třetí podmínka klasického lineárního regresního modelu, což při použití metody nejmenších čtverců vede k problémům s odhady parametrů ϕ_1, \dots, ϕ_p . Přítomnost autokorelovaných reziduí se sice nedotkne nestrannosti a konzistence odhadů parametrů modelu AR, jejich vydatnost však klesne (vzroste jejich směrodatná chyba). Jestliže je v modelu přítomna multikolinearita, pokles vydatnosti odhadu parametrů je ještě výraznější.

Autokorelovaná rezidua je možné modelovat následujícím způsobem

$$e_t = \theta_q(B) a_t \quad (3.25)$$

V případě existence autokorelovaných reziduí se problém odhadu parametrů ϕ_1, \dots, ϕ_p modelu (3.24) rozšiřuje na problém odhadu dalších parametrů $\theta_1, \dots, \theta_q$. Jde o úlohu odhadu

parametrů modelu ARMA(p, q) formy (19) prostřednictvím minimalizace reziduálního součtu čtverců $S(\phi, \mu, \theta)$. Zde již s obyčejnou metodou nejmenších čtverců nevystačíme.

Nelineární odhady parametrů

Pro odhad parametrů modelu ARMA(p, q) nelze použít metodu nejmenších čtverců, neboť na rozdíl od modelu AR(p) tento model není lineární v parametrech, což lze ilustrovat na příkladu stacionárního modelu ARMA(1, 1).

$$(1 - \phi_1 B) X_t^* = (1 - \theta_1 B) a_t \quad (3.26)$$

veličinu a_t je možné získat následujícím způsobem

$$\begin{aligned} a_t &= X_t^* - \phi_1 X_{t-1}^* + \theta_1 a_{t-1} \\ &= X_t^* - \phi_1 X_{t-1}^* + \theta_1 (X_{t-1}^* - \phi_1 X_{t-2}^* + \theta_1 a_{t-2}) \\ &= X_t^* - (\phi_1 - \theta_1) X_{t-1}^* + \phi_1 \theta_1 X_{t-2}^* + \theta_1^2 a_{t-2} \\ &\vdots \end{aligned}$$

Parametry nemohou být odhadnuty pomocí metody nejmenších čtverců, neboť nejsou lineární. Pro odhad se používá následující iterační nelineární odhadovací procedura.

Princip spočívá v linearizaci funkce nelineární v parametrech pomocí Taylorova rozvoje funkce okolo počátečních odhadů pro parametry $\phi = (\phi_1, \dots, \phi_p)$, $\theta = (\theta_1, \dots, \theta_q)$. Odhady parametrů jsou následně získány použitím obyčejné metody nejmenších čtverců. Dalším krokem je opět stejná linearizace původní funkce nelineární v parametrech provedená tentokrát na základě nově odhadnutých parametrů. Proces se iterativně opakuje do té doby než se odhady parametrů ϕ a θ a μ budou lišit v jednotlivých iteracích o nějaké malé číslo ε .

3.5.3 Ověřování modelu

Testování stacionarity

Při výstavbě modelů třídy ARIMA je velmi důležité určení řádu diferencování d . Při analýze časových řad se prakticky setkáváme s integrovanými časovými řadami maximálně řádu dva, tj. řadami typu I(2). Nejčastěji se však můžeme setkat s řadami I(1). Časové řady se tedy stacionarizují většinou prostřednictvím první či druhé diference.

Velmi jednoduchou subjektivní metodou je posouzení grafu časové řady. Vhodné je porovnat původní nediferencovanou časovou řadu s časovou řadou prvních či druhých diferencí. I když má tato metoda subjektivní charakter, je v mnoha případech velmi efektivní.

Další velmi jednoduchou metodou je také posouzení tvaru výběrové autokorelační funkce. Klesá-li tato funkce pomalu, přibližně lineárním tempem, jedná se o situaci, kdy alespoň jeden kořen rovnice $\phi_p(B) = 0$ je blízký jedné a je tedy vhodné provést diferencování. Použití výběrové autokorelační funkce může někdy vést k „přediferencování“ časové řady. I když diferencování stacionárních časových řad vede opět k stacionárním časovým řadám, tato operace může způsobovat vážné problémy.

Třetí možností je Dickey-Fullerův test, který testuje, zda je stacionární řada vzniklá po první diferenci. Takové řady se nazývají řady s jednotkovým kořenem. Teorie týkající se testování řad s jednotkovým kořenem je poměrně rozsáhlá, uvedu pouze základní verzi Dickey-Fullerova testu.

Uvažujme, že jsme modelovali časovou řadu modelem AR(1)

$$X_t = \delta_0 + \rho X_{t-1} + a_t$$

kde $\{a_t\}$ je proces gaussovského bílého šumu. Pokud $\rho = 1$ a $\delta_0 = 0$, potom se tento model stává modelem náhodné procházky. Dickey-Fullerův test ověřuje hypotézu, že $H_0: \rho = 1$. Alternativní hypotézou je $H_1: \rho < 1$. Testuje se tedy hypotéza, že časová řada vzniklá po první diferenci je stacionární proti hypotéze, že časová řada je stacionární. Jako testové kritérium se nabízí t -statistika vypočítaná na základě metody nejmenších čtverců, jenž má tvar

$$\hat{\tau} = \frac{\hat{\rho} - 1}{S\hat{\rho}}$$

kde $\hat{\rho}$ je odhad pořízený metodou nejmenších čtverců a $S\hat{\rho}$ je odhad směrodatné chyby odhadu $\hat{\rho}$. Tato statistika má tzv. Fullerovo rozdělení.

Vraťme se k obecnému modelu ARIMA(p, d, q). Budeme uvažovat, že $d = 1$, tj. že časová řada vzniklá po první diferenci je stacionární. Potom se používá pro testování jednotkového kořenu tzv. rozšířený Dickey-Fullerův test, při kterém se vychází z modelu ve tvaru

$$X_t = \delta_0 + \rho X_{t-1} + Z_t$$

kde Z_t lze modelovat stacionárním a invertibilním modelem ARMA(p, q) tvaru

$$Z_t + \sum_{i=1}^p \phi_i Z_{t-i} = a_t + \sum_{j=1}^q \theta_j a_{t-j}$$

kde $\{a_t\}$ je proces gaussovského bílého šumu. Tento model lze převést na model tvaru

$$X_t = \delta_0 + \rho X_{t-1} + \sum_{i=1}^k b_i (X_{t-i} - X_{t-1-i}) + a_t$$

Odhady parametrů se získají metodou nejmenších čtverců. Said a Dickey (1984) rovněž dokazují, že limitní rozdělení statistiky \bar{x} je stejné jako v předchozím jednoduchém případě a lze tedy použít stejné tabulky kvantilů tohoto rozdělení.

Testování reziduí

V modelu ARIMA(p, d, q) typu (3.3) předpokládáme, že $\{a_t\}$ je proces bílého šumu. Odhadnutá rezidua na základě identifikovaného modelu

$$\hat{a}_t = \hat{\theta}_q(B) \hat{\phi}_p(B) y_t$$

kde $\hat{\phi}_p(B) = 1 - \hat{\phi}_1 B - \dots - \hat{\phi}_p B^p$ je autoregresivní operátor a $\hat{\theta}_q(B) = 1 - \hat{\theta}_1 B - \dots - \hat{\theta}_q B^q$ je operátor klouzavých průměrů, by měla být nekorelována. Jestliže jsou rezidua nekorelována, musí být odhady autokorelační funkce a parciální autokorelační funkce blízké nule pro posunutí $k \geq 1$, tyto odhady by měly ležet uvnitř tolerančních mezí ($\pm 2S\hat{\rho}_k$ či $\pm 2S\hat{\phi}_{kk}$ při $\alpha = 0,05$)

Další možností, jak zjistit, zda odhadnutá rezidua jsou nekorelována, je použití *portmanteau testu*. Testuje se hypotéza $H_0: r_1 = r_2 = \dots = r_K = 0$ proti hypotéze $H_1: \text{non } H_0$, kde r_k $k = 1, \dots, K$ jsou autokorelační koeficienty reziduí pro posunutí k . Hodnoty výběrové autokorelační funkce reziduí \hat{r}_k jsou počítány jako

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

Platí-li hypotéza H_0 , jsou reziduální výběrové autokorelace pro velká k , normálně rozdělené náhodné veličiny s nulovou střední hodnotou a rozptylem $1/T$, kde $T = n - d$ je počet pozorování diferencované časové řady.

Uvažujme nyní statistiku

$$Q = T \sum_{k=1}^K \hat{r}_k^2 \quad (3.27)$$

kteřá má přibližně rozdělení χ^2 (rozdělení je přibližné proto, že pro malá k je rozptyl těchto náhodných veličin poněkud podhodnocen a veličiny mohou být závislé). Bylo dokázáno, že

tato statistika (portmanteau statistika) má pro velká K asymptoticky rozdělení χ^2 s $(K - p - q)$ stupni volnosti. Takže porovnáním hodnoty testového kritéria (3.27) s příslušnými kvantily rozdělení $\chi^2(K - p - q)$ lze testovat hypotézu o nezávislosti reziduí.

Bylo dokázáno, že pro rozsahy výběrů používaných v praxi má statistika (3.27) jiné pravděpodobnostní rozdělení, než je její rozdělení asymptotické (limitní). Proto byla navržena následující statistika

$$Q' = T(T+2) \sum_{k=1}^K (T-k)^{-1} \hat{r}_k^2$$

která má rozdělení $\chi^2(K - p - q)$. Nazývá se *modifikovaná portmanteau statistika* (Ljung-Boxova statistika)

Testování bílého šumu v případech sezónních modelů je obdobné jako u modelů nesezónních s tím, že se mění počet stupňů volnosti, portmanteau test má rozdělení $\chi^2(K - p - q - P - Q)$.

Testování náhodné složky

Náhodná složka ε_t časové řady představuje působení nepodchycených drobných vzájemně nezávislých vlivů, které se v rámci časové řady vzájemně vykompenzují, tj. platí $E(\varepsilon_t) = 0, t=1,2,\dots, n$.

Různé typy předpokladů na náhodné složce:

- Homoskedasticita* – jednotlivé náhodné složky jsou lineárně nezávislé s konstantním rozptylem: $D(\varepsilon_t) = \sigma^2, t = 1, 2, \dots, n$ a $E(\varepsilon_i \varepsilon_j) = 0, \forall i \neq j$.
- Heteroskedasticita* – náhodné složky jsou nezávislé s měnlivými rozptyly: $D(\varepsilon_t) = \sigma^2 / w_t, t = 1, 2, \dots, n$ kde váhy w_t splňují podmínku $\sum w_t^{-1} = n$ a $E(\varepsilon_i \varepsilon_j) = 0, \forall i \neq j$.
- Autokorelace* – náhodná porucha v čase t závisí lineárně na poruše v předcházejícím čase $t-1$, tj. $\varepsilon_t = \rho \varepsilon_{t-1} + u_t, t = 1, 2, \dots, n, 0 < \rho < 1$, kde ρ se nazývá koeficient autokorelace, který je považován za konstantu a u_t jsou opět nezávislé náhodné poruchy s nulovými středními hodnotami a konstantními rozptyly.

Ve všech testech se jako nulová hypotéza H_0 testuje, zda předložená pozorování jsou realizace vzájemně nezávislých stejně rozdělených náhodných veličin, které nemusí mít jako bílý šum nulovou střední hodnotu (může se tedy jednat o bílý šum kolísající kolem nenulové

úrovně). Při zamítnutí nulové hypotézy zřejmě analyzovaná řada tím spíše klasickým bílým šumem být nemůže.

a) *Znaménkový test*

Vypočteme první diference časové řady reziduí. Označíme S počet kladných diferencí časové řady ε_t . K testu nulové hypotézy o náhodnosti uspořádání reziduí, tj. hypotézy $H_0: E(S) = (n - 1)/2$ proti alternativě $H_1: E(S) \neq (n - 1) / 2$ použijeme testové kritérium

$$U = \frac{S - \frac{1}{2}(n-1)}{\sqrt{n+1}} \sqrt{12},$$

kteřá má za platnosti H_0 asymptoticky (pro $n > 12$) normální rozdělení. Kritický obor odpovídající hladině významnosti α je vymezen nerovností $|U| > u_{1-\alpha/2}$, kde $u_{1-\alpha/2}$ je kvantil rozdělení $N(0, 1)$.

b) *Test bodů obrátů*

Body obrátů jsou vrcholy a dolíky řady. Vrcholem je pozorování s hodnotou vyšší, než jsou hodnoty dvou sousedních pozorování, dolíkem je pozorování s hodnotou nižší, než hodnoty dvou sousedních pozorování. Celkový počet obrátů u časové řady reziduí ε_t označme P . Lze ukázat, že P je náhodná veličina se střední hodnotou $E(P) = 2(n-2) / 3$ a rozptylem $D(P) = (16n-29) / 90$. Testujeme tedy hypotézu $H_0: P = 2(n-2) / 3$ proti alternativě $H_1: P \neq 2(n-2) / 3$. Užijeme testového kritéria

$$U = \frac{P - \frac{2}{3}(n-2)}{\sqrt{16n+29}} \sqrt{90},$$

kteřé má za platnosti H_0 asymptoticky normální rozdělení $N(0,1)$. Kritický obor odpovídající hladině významnosti α je vymezen nerovností $|U| > u_{1-\alpha/2}$, kde $u_{1-\alpha/2}$ je kvantil rozdělení $N(0,1)$.

c) *Test založený na Kendallově koeficientu τ*

Spočteme v dané řadě počet v takových dvojic y_s a y_t , že $y_s < y_t$ pro $s < t$. Test je založen na Kendallově koeficientu pořadové korelace τ , definovaném jako

$$\tau = \frac{4v}{n(n-1)} - 1.$$

Tento koeficient, který byl původně zaveden pro ohodnocení závislosti mezi různými uspořádáními daných hodnot, může nabývat hodnot pouze z intervalu $\langle -1, 1 \rangle$ a má za platnosti hypotézy H_0 nulovou střední hodnotu a rozptyl

$$\text{var}(\tau) = \frac{2(2n+5)}{9n(n-1)}.$$

Hypotézu H_0 zamítáme, když

$$\frac{|\tau|}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}} > u_{1-\alpha/2}$$

d) *Test založený na Spearmanově koeficientu ρ*

Nechť q_1, \dots, q_n označuje pořadí hodnot dané řady. Spearmanův koeficient pořadové korelace ρ budeme používat ve tvaru

$$\rho = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^n (i - q_i)^2$$

Kritické hodnoty $r_S(p)$ pro testování hypotézy H_0 takové, že $P(|\rho| > r_S(p)) \leq p$ jsou tabelovány např. v učebnici [Anděl, J. *Matematická statistika*. Praha. SNTL/Alfa 1978] pro $n \leq 30$. Pro větší n hypotézu H_0 zamítáme, když

$$|q| \sqrt{n-1} \geq u_{1-\alpha/2}$$

e) *Mediánový test*

Pro tento test je nutné zkonstruovat výběrový medián M daných pozorování. V grafickém záznamu řady hledáme přímkou rovnoběžnou s časovou osou, která má tu vlastnost, že nad ní i pod ní leží stejný počet pozorování. Vyloučíme všechna pozorování ležící na zkonstruované přímce a ostatní pozorování sdružíme do skupin tak, že všechna pozorování v každé takové skupině leží buď nad danou přímkou, nebo pod ní. Označme u počet těchto skupin a m celkový počet pozorování pod přímkou. Odpovídající kritické hodnoty pro test založený na počtu u jsou uvedeny v [Likeš, J. – Jaga, J. *Základní statistické tabulky*. Praha. SNTL 1978] Hypotézu H_0 zamítáme, když

$$\frac{|u - (m+1)|}{\sqrt{m(m-1)/(2m-1)}} \geq u_{1-\alpha/2}$$

Výběr modelu

V některých případech se může stát, že se identifikují dva či více přijatelných modelů. Potom vzniká otázka, který z nich zvolit.

Existují dva přístupy pro výběr modelu. První z nich je založen na dalším zkoumání a porovnání odhadnutých reziduí. Dává se přednost modelu s nejmenší hodnotou odhadu reziduálního rozptylu a s menším počtem parametrů. Z tohoto principu vychází několik následujících kritérií.

1. Kritérium AIC (Akaike Information Criterion) zavedl Akaike (1973). Je definováno jako

$$AIC(M) = T \ln \hat{\sigma}_u^2 + 2M,$$

kde T je počet pozorování po diferenciaci, M je počet parametrů. Vybírá se ten model, který minimalizuje $AIC(M)$

2. Kritérium BIC (Bayesian extension of Information Criterion) zavedl Akaike z důvodu, že AIC nadhodnocovalo řád autoregrese. Toto kritérium má formu

$$BIC(M) = T \ln \hat{\sigma}_a^2 - (T - M) \ln \left(1 - \frac{M}{T} \right) + M \ln T + M \ln \left[\frac{\left(\frac{\hat{\sigma}_X^2}{\hat{\sigma}_a^2} - 1 \right)}{M} \right]$$

kde $\hat{\sigma}_X^2$ je odhad rozptylu časové řady X . Volí se ten model, který minimalizuje $BIC(M)$

3. Kritérium SBC (Schwartz's Bayesian Criterion) bylo navrženo Schwartzem (1978) a má formu

$$SBC(M) = T \ln \hat{\sigma}_u^2 + M \ln T$$

Opět se vybere ten model, který minimalizuje $SBC(M)$

Druhý způsob volby vhodného modelu spočívá v porovnání přesnosti předpovědí, které jednotlivé modely poskytují. Časová řada se rozdělí na dvě části. Modely jsou odhadnuty na základě první části, druhá část je pak použita pro měření přesnosti předpovědí „ex post“

3.6 Konstrukce předpovědí

3.6.1 Výpočet předpovědí

Výpočet předpovědí se provádí rekurzivně na základě odhadů parametrů daného modelu. Nejprve se vypočítá předpověď na jeden krok dopředu. Tato předpověď se využije pro výpočet předpovědí na dva kroky dopředu atd. Model (3.3) lze zapsat také jako

$$X_t^* = \phi_1 X_{t-1}^* + \dots + \phi_p X_{t-p}^* + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

kde $X_t^* = (1-B)^d X_t$

Aby bylo možné vypočítat předpověď $X_n^*(h)$, je třeba vypočítat předpovědi $X_n^*(1)$, $X_n^*(2)$, ..., $X_n^*(h-1)$.

Obecně předpověď $X_n^*(h)$ lze zapsat jako

$$X_n^*(h) = \hat{\phi}_1 \hat{X}_n^*(h-1) + \dots + \hat{\phi}_h \hat{X}_n^* + \dots + \hat{\phi}_p \hat{X}_{n-p+h}^* - \hat{\theta}_h \hat{a}_n - \dots - \hat{\theta}_q \hat{a}_{n-q+h}$$

Jestliže $h > p$ a $h > q$, potom bude tato předpověď

$$X_n^*(h) = \hat{\phi}_1 \hat{X}_n^*(h-1) + \dots + \hat{\phi}_p \hat{X}_n^*(h-p)$$

Z tohoto zápisu je patrná „paměť“ stacionárního procesu AR a invertibilního procesu MA. Zatímco „paměť“ procesu AR se ztrácí postupně, jak se zvyšuje horizont předpovědi h , a předpovědi se blíží ke střední hodnotě daného procesu AR, tj. nule (pokud není zařazen volný parametr θ_0), „paměť“ procesu MA končí po horizontu $h = q$, resp. předpověď konstruovaná pro $h > q$ má vždy hodnotu nula (resp. hodnotu volného parametru θ_0 , pokud je zařazen do modelu). O stacionárních procesech se někdy hovoří také jako o procesech s krátkou „pamětí“.

Jestliže se modeluje řada pomocí nestacionárního procesu, potom je předpověď tvořena jako kumulativní součet předpovědi získaných výše uvedeným způsobem. Takže, když $d = 1$, potom předpověď s horizontem h se počítá jako

$$X_n(h) = X_n + \hat{X}_n^*(1) + \hat{X}_n^*(2) + \dots + \hat{X}_n^*(h)$$

Jestliže $d = 2$, potom předpověď $\hat{X}_n(h)$ se počítá jako

$$\begin{aligned} \hat{X}_n(h) = X_n + [(1-B)X_n + \hat{X}_n^*(1)] + [(1-B)X_n + \hat{X}_n^*(1) + \hat{X}_n^*(2)] + \dots + \\ + [(1-B)X_n + \hat{X}_n^*(1) + \dots + \hat{X}_n^*(h)] \end{aligned}$$

Vzhledem k tomu, že se jednotlivé předpovědi nestacionárních procesů kumulují, informace v nich obsažené se s rostoucím horizontem neztrácejí. V tomto smyslu mají nestacionární procesy dlouhou „paměť“.

Interval spolehlivosti pro předpověď $\hat{X}_n(h)$ je počítán jako

$$C_n = \hat{X}_n(h) \pm c \sqrt{\sum_{j=0}^{h-1} \hat{\psi}_j^2} \hat{\sigma}_a$$

kde např. pro 95% interval spolehlivosti $c = 2$. Je logické, že pro zvyšující se horizont předpovědi h se interval spolehlivosti rozšiřuje.

3.7 Literatura

- Akaike, H. *Information theory and an extension of the maximum likelihood principle*. 2nd Internation Symposium on Information Theory, Budapest: 1973. Akademiai Kiado. s. 267-281
- Anděl, J. *Statistická analýza časových řad*. Praha: SNTL. 1976
- Arlt, J. *Moderní metody modelování ekonomických časových řad*. 1. vyd. Praha: Grada Publishing. 1998. 312 s. ISBN 80-7169-539-4
- Arlt, J. – Arltová, M. *Příklady z analýzy ekonomických časových řad*. Praha: VŠE 1997.
- Box, G. E. P. – Jenkins, G. M. *Time Series Analysis: Forecasting and Control*, rev. ed., San Francisco: Holden Day. 1976
- Cipra, T. *Analýza časových řad s aplikacemi v ekonomii*. Praha: SNTL. 1986.
- Kozák, J. – Hindls, R. – Arlt, J. *Úvod do analýzy ekonomických časových řad*. Praha: VŠE 1994. ISBN 80-7079-760-6.
- Said, E. S. – Dickey, D. A. *Testing for unit roots in autoregressive-moving average models of unknown order*. Biometrika: 1984, 71, s. 599-607
- Schwartz, G. *Estimating the dimension of a model*. Ann. Statist., s. 461-464. 1978
- Stuchlý, J. *Statistika II*. Praha: VŠE 1999.

4 Wavelet transformace ve zpracování signálů

V druhé kapitole následuje popis teorie waveletů, kde se soustředím především na dekompozici signálu, tzv. MR-analýzu (z angl. multiresolution analysis - MRA), což lze volně přeložit jako analýza o více úrovních rozlišení. Tato metoda je výpočetně velmi atraktivní, neboť umožňuje jednoduchou aplikaci vhodné dvojice lineárních filtrů postupně snižovat nebo zvyšovat rozlišovací schopnost waveletové aproximace. Wavelet transformaci lze využít pro analýzu stacionárních i nestacionárních signálů.

Wavelet transformace představuje moderní nástroj pro analýzu, dekompozici a rekonstrukci signálů a obrazů. Její předností v porovnání s krátkou diskrétní Fourierovou transformací je zejména možnost volby funkcí pro analýzu v závislosti na řešeném problému a dále proměnná rozlišovací schopnost pro dílčí frekvenční složky signálu. Práce prezentuje základní vztahy pro implementaci diskrétní Wavelet transformace a zabývá se jejími některými typickými aplikacemi.

4.1 Úvod

Wavelet transformace (WT) představuje poměrně nový matematický prostředek pro analýzu signálů pomocí Wavelet funkcí s aplikací na široké spektrum reálných signálů, které zahrnuje technologické časové řady, biomedicínské signály a obrazy, družicové snímky, ekonomická data i lidskou řeč. V řadě případů je problémem efektivní kódování, komprese, potlačování rušivých složek, modelování a detekce dílčích složek signálu. Wavelet transformace představuje v řadě těchto oblastí moderní a pružný nástroj, který lze modifikovat s ohledem na řešený problém.

Prostředky Wavelet transformace se ve svých principech opírají o práce Josepha Fouriera, který v 19. století položil základy frekvenční analýzy. Tyto principy však byly zásadně modifikovány zejména v posledních 20 letech francouzskými matematiky Y. Meyerem, S. Malattem a I. Daubechies. Důvodů pro současný velice rychlý rozvoj Wavelet funkcí je celá řada. Základním z nich je nutnost analýzy nestacionárních signálů, pro které je použitelná Fourierova transformace s posuvným okénkem konečné délky. Wavelet funkce představují v této oblasti nový nástroj s možností analýzy signálů na různých rozlišovacích úrovních v závislosti na zpracovávaném frekvenčním pásmu. Jejich význam je i v tom, že

umožňují velice efektivně využít pro analýzu signálů a obrazů i jiné než harmonické funkce, a to zejména pro analýzu fraktálů a přechodových složek signálů.

Teoretický základ Wavelet transformace byl studován v mnoha pracích a knihách. Mezi důležité výsledky těchto studií patří analýza a konstrukce wavelet funkcí v analytické i rekurentní formě spolu s popisem jejich vlastností. Wavelet transformace přitom představuje alternativní přístup k diskrétní Fourierově transformaci (STFT) pro analýzu nestacionárních signálů a detekci bodů, pro které signál mění vlastnosti. Hlavní výhoda WT spočívá v její schopnosti analýzy signálů s proměnným časově-frekvenčním rozlišením.

Mezi aplikace Wavelet transformace patří detekce složek signálu a dekompozice signálu, komprese dat a potlačování šumu. Problémy, které jsou úzce spjaty s tímto tématem, zahrnují klasifikaci a predikci signálu, statistické zpracování časových řad, analýzu geofyzikálních signálů a analýzu obrazů.

4.2 Matematický základ

4.2.1 Fourierova a waveletová řada

Označení:

R ... množina všech reálných čísel,

Z ... množina všech celých čísel,

C ... množina všech komplexních čísel

δ_{ij} ... Kroneckerův symbol (= 1 pro $i = j$ a 0 jinak)

vyraz := s ... označení výrazu symbolem s

Budeme pracovat obecně s komplexními funkcemi (periodickými i neperiodickými) definovanými na celé reálné ose, např. $x := x(t): R \rightarrow C$. Většinou se bude jednat o funkce z následujících funkcionálních prostorů, kde $J \subset R$ je interval periody v případě periodických funkcí nebo $J = R$ v případě neperiodických funkcí:

$$L_p(J) := \{x \mid x \text{ měřitelná a } \int_J |x(t)|^p dt < \infty\}, 1 \leq p < \infty.$$

Norma v $L_p(J)$: $\|x\|_p = \left(\int_J |x(t)|^p dt \right)^{\frac{1}{p}}$ pro $1 \leq p < \infty$

$L_p(J)$, $1 \leq p \leq \infty$, je **Banachův prostor** nad C .

$$L_p := L_p(\mathbb{R})$$

$L_2(J)$ je dokonce Hilbertův prostor, kde pro $x, y \in L_2(J)$:

$$\langle x, y \rangle = \int_J x(t) \overline{y(t)} dt \text{ je skalární součin, přičemž}$$

$$\langle x, x \rangle = \|x\|_2^2 \quad (\|x\|_2^2 = \|x\|_{L_2}^2) \text{ a}$$

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0 \text{ určuje ortogonalitu } x \text{ a } y.$$

Interpretujeme-li funkci x jako signál (spojité měření nebo pozorování nějakých dat), pak hodnota $\|x\|_2^2$ je energie signálu, tj. $L_2(J)$ představuje třídu všech signálů o konečné energii na intervalu J .

Z matematického pohledu představuje waveletová analýza přímou analogii klasické okenní Fourierovy analýzy T -periodických funkcí v $L_2([a, b])$, $b - a = T$. Fourierově řadě, která představuje vyjádření T -periodické funkce ve spočetné ortonormální bázi (indexované celými čísly) $\{\omega(jt)\}_{j \in \mathbb{Z}}$ harmonických kmitů odvozených z komplexní sinusové vlny $\omega(t) = e^{i2\pi t/T} / \sqrt{T} = (\cos 2\pi t/T + i \sin 2\pi t/T) / \sqrt{T}$ pouhou změnou měřítka (frekvence) v závislosti na j , odpovídá **waveletová řada** jakožto rozvoj funkce ve vhodné spočetné bázi opět určené jedinou funkcí $\psi(t)$ nazývanou **mateřský wavelet (mother wavelet)**. Protože $\omega(t)$ je funkce mající konečnou energii pouze na ohraničeném intervalu ($\int_a^b |\omega(t)|^2 dt < \infty$), je třeba hledat $\psi \neq \omega$, která má konečnou energii na celé reálné ose. Taková funkce však musí vymizet k nule pro $t \rightarrow \pm\infty$ a dokonce se ukazuje, že za dodatečného předpokladu integrovatelnosti (tj. $x \in L_1 \cap L_2$) musí měnit i znaménko, tj. ψ musí být tlumený kmit neboli **vlnka (= wavelet)**. V důsledku tohoto tlumení, které může být dokonalé v tom smyslu, že ψ je nenulová pouze na ohraničeném intervalu (říkáme, že má kompaktní nosič), nestačí generovat bázi pouhou změnou měřítka mateřského waveletu, ale navíc je nutné jej i posouvat. Waveletová báze se pak dostane jako spočetný dvojnásobně celými čísly indexovaný systém $\{\psi_{j,k}(t)\}_{j,k \in \mathbb{Z}}$, kde $\psi_{j,k}(t) = 2^{j/2} \psi\left(\frac{t - k2^{-j}}{2^{-j}}\right)$ a $2^{j/2}$ je normalizační konstanta zajišťující konstantní (jednotkovou) normu $\|\psi_{j,k}\|_2 = \|\psi\|_2$.

Označme $E = \{\psi_{j,k}(t)\}_{j,k \in \mathbb{Z}}$. Mateřský wavelet ψ se nazývá **ortogonální**, jestliže E je ortonormální báze v L_2 : $\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \delta_{k,m}$, kde $\delta_{i,j}$ je Kroneckerův symbol.

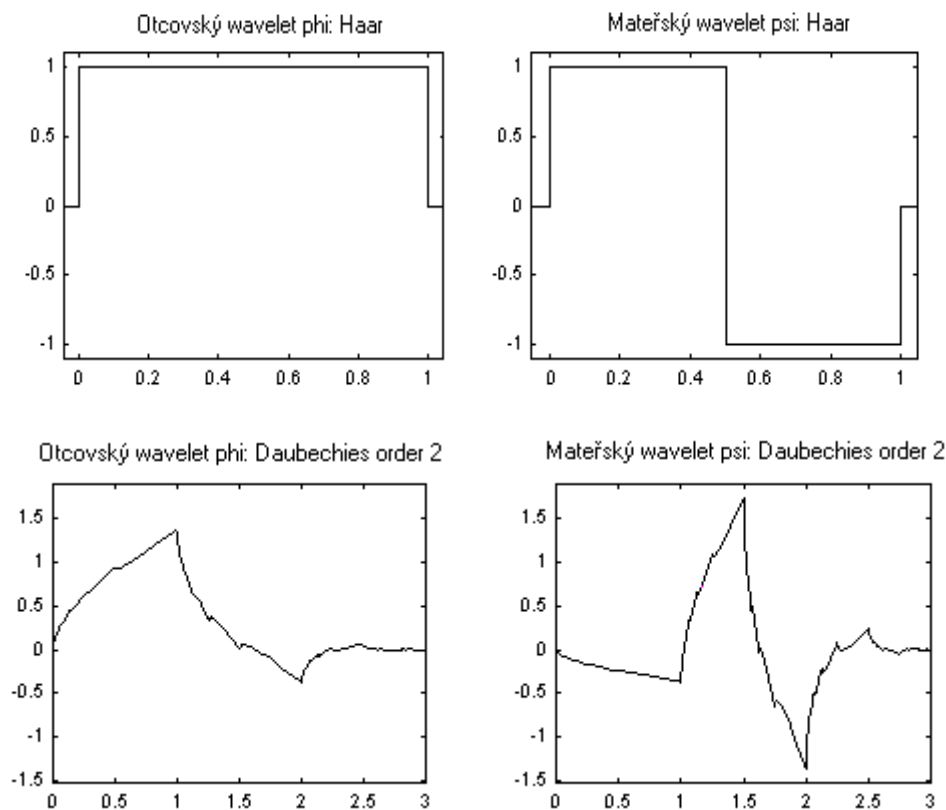
Wavelet řada, za předpokladu použití ortogonálních waveletů, je tedy zobecněním Fourierovy řady s tím rozdílem, že má dva parametry, tudíž je dvourozměrná.

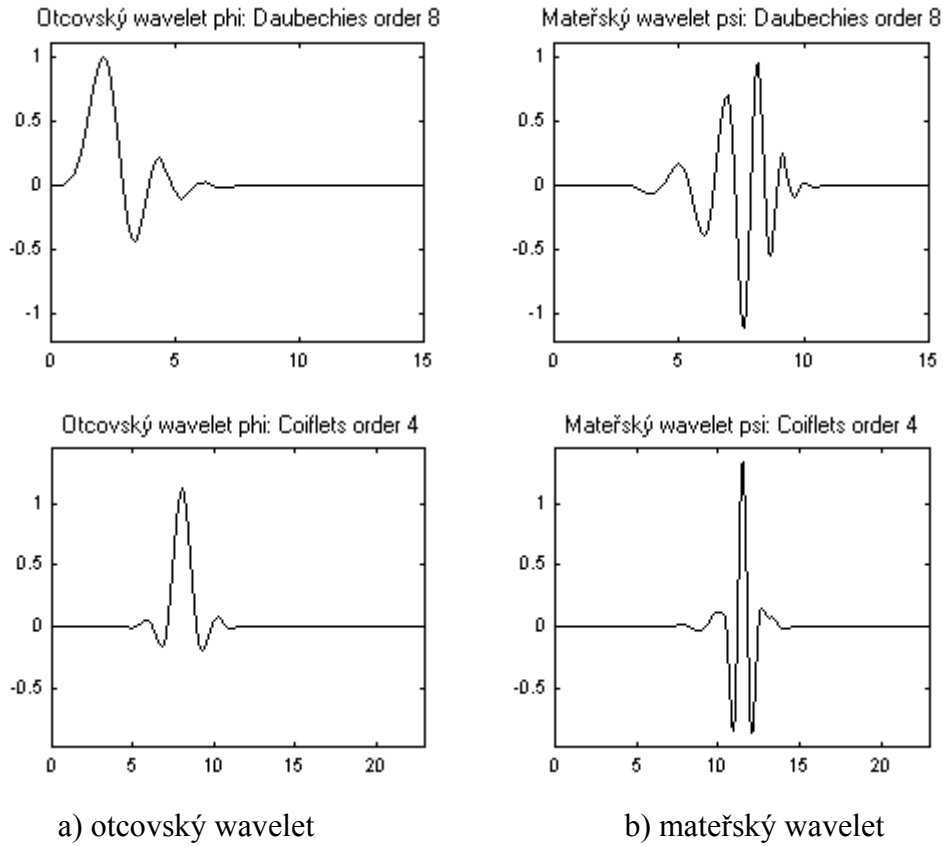
Mateřský wavelet ψ se nazývá **waveletem s kompaktním nosičem (compactly supported wavelet)**, jestliže $\{t \mid \psi(t) \neq 0\}$ je ohraničená množina.

Nejjednodušším a historicky nejstarším příkladem ortogonálního waveletu je tzv. **Haarova funkce** ψ_H definovaná předpisem

$$\psi_H(t) = \begin{cases} 1 & \text{pro } 0 \leq t < 0,5 \\ -1 & \text{pro } 0,5 \leq t < 1 \\ 0 & \text{jinak} \end{cases}$$

V nedávné době bylo nalezeno mnoho dalších ortogonálních waveletů, ukázky některých z nich i Haarova waveletu jsou na obr. 7.





Obr. 7 Ukázky waveletů

Každá funkce $x \in L_2([a, b])$, $b - a = T$ je součtem své Fourierovy řady:

$$x(t) = \sum_{j=-\infty}^{\infty} x_j \frac{1}{\sqrt{T}} e^{i2\pi \frac{j}{T} t} = \sum_{j=-\infty}^{\infty} \langle x, e_j \rangle \frac{1}{\sqrt{T}} e^{i2\pi \frac{j}{T} t} = \sum_{j=-\infty}^{\infty} c_j e^{i2\pi \frac{j}{T} t} \quad (4.1)$$

kde

$$c_j := c_j(x) = \frac{1}{\sqrt{T}} \underbrace{\langle x, e_j \rangle}_{x_j} = \frac{1}{\sqrt{T}} \int_a^b x(t) \frac{1}{\sqrt{T}} e^{-i2\pi \frac{j}{T} t} dt = \frac{1}{T} \int_a^b x(t) e^{-i2\pi \frac{j}{T} t} dt \quad (4.2)$$

je tzv. j -tý **Fourierův koeficient** a $\{c_j(x)\}_{j \in \mathbb{Z}}$ je tzv. **fourierovské spektrum** funkce $x(t) \in L_2([a, b])$.

Podobně každá funkce $x(t) \in L_2(\mathbb{R})$ je součtem své **waveletové řady**

$$x(t) = \sum_{j,k=-\infty}^{\infty} c_{j,k} \psi_{j,k}^*(t) \quad (4.3)$$

kde

$$c_{j,k} := c_{j,k}(x) = \langle x, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} x(t) \overline{\psi_{j,k}(t)} dt = 2^{j/2} \int_{-\infty}^{\infty} x(t) \overline{\psi\left(\frac{t-k2^{-j}}{2^{-j}}\right)} dt \quad (4.4)$$

je tzv. (j, k) -tý **waveletový koeficient** a $\{c_{j,k}\}_{j,k \in \mathbb{Z}}$ je tzv. **waveletové spektrum** funkce $x(t) \in L_2(\mathbb{R})$.

Parsevalova identita

- **pro Fourierovu řadu:**

$$\underbrace{\frac{1}{T} \int_a^b |x(t)|^2 dt}_{\text{střední výkon}} = \sum_{j=-\infty}^{\infty} c_j(x) \overline{c_j(x)} = \sum_{j=-\infty}^{\infty} |c_j(x)|^2$$

- **pro waveletovou řadu:**

$$\underbrace{\int_{-\infty}^{\infty} |x(t)|^2 dt}_{\text{energie}} = \sum_{j,k=-\infty}^{\infty} |c_{j,k}(x)|^2$$

Při zpracování konkrétních diskretních dat vede numerický výpočet integrálu v (4.2) na známý operátor diskretní Fourierovy transformace (DFT), jehož inverze (IDFT) koresponduje s (4.1), kde nekonečnou řadu nahradíme pouze jejím částečným součtem. Podobně diskretizací (4.4) a (4.3) dostaneme příslušné operátory diskretní waveletové transformace a její inverze. Pro jejich výpočet existují rychlé algoritmy (FWT = Fast Wavelet Transform) a rychlá Fourierova transformace (FFT = Fast Fourier Transform).

4.2.2 Časově-frekvenční spektrální charakteristika

V klasické fourierovské bázi je každá bazová funkce sinusový kmit, který ovlivňuje průběh $x(t)$ stejnoměrně na celé reálné ose a má tedy **globální** účinek. Je to tedy funkce, která **není lokalizována v čase**, ale naopak je **ideálně lokalizována ve frekvenci** (její spektrum je jednobodové). Výpočet každého jednotlivého spektrálního koeficientu podle (4.2) tedy vyžaduje úplnou znalost funkce (signálu) $x(t)$ v minulosti i budoucnosti. Nevýhoda této reprezentace se nejvýrazněji projeví u signálu, jehož dynamika (tj. frekvenční charakteristika) se s časem výrazně mění. Takový signál si vynucuje silné zastoupení velkého množství vysokofrekvenčních komponent (např. skokové změny se vymodelují až v limitě) neboli

zabírá **široké frekvenční pásmo**. Prakticky nepříznivým důsledkem je velký objem dat spektra $\{c_j(x)\}_{j \in \mathbb{Z}}$ potřebný k vyhovujícímu popisu $x(t)$.

Vzniká tak požadavek vyšetřovat **lokální frekvenční charakteristiky** signálu postupně v čase. Standardní postup využívá tzv. **okenní Fourierovu transformaci**, které postupně „klouže“ po datech (kontinuálně nebo v diskrétních krocích) a vybírá tak z dat pro frekvenční analýzu pouze lokální úsek:

$$(F^{win} x)(t, f) = \int_{-\infty}^{\infty} x(t) \overline{g(t-t')} e^{-i2\pi ft} dt \quad (4.5)$$

kde $g(\cdot)$ je váhové okno se středem v 0.

Protože frekvence je nepřímo úměrná délce cyklu, tak pro zachycení vysokofrekvenční informace vystačíme s kratším intervalem, zatímco pro zachycení nízkofrekvenční informace je naopak potřeba interval delší. Jinými slovy, potřebujeme mít flexibilní časové okno, které se automaticky zužuje, je-li střední frekvence jeho spektra vyšší a rozšiřuje, je-li tato frekvence nižší. Tuto vlastnost mají právě $\psi_{j,k}$. Pro rostoucí j se $\psi_{j,k}$ zužuje (šířka = $1/2^j$). Se zužováním je třeba rovněž zjemňovat posuv $k / 2^j$, aby nevznikly časové díry nepokryté vysokofrekvenční informací. Aby se $\psi_{j,k}$ mohly používat pro časově-frekvenční analýzu ve výše uvedeném smyslu, musí být tedy dobře lokalizovány současně v čase i ve frekvenci. Což jsou ovšem protichůdné požadavky, kterým lze rozumně vyhovět pouze v případě, že $\psi(t)$ se rychle tlumí (konvergují k nule) pro $t, f \rightarrow \pm\infty$. Navíc musíme pro ψ a následně i pro $\psi_{j,k}$ být schopni určit jejich střed a šířku. Taková funkce ψ se pak nazývá **váhovou funkcí časového, resp. frekvenčního okna**.

Definice: Funkce $0 \neq \omega(t) \in L_2$ se nazývá **váhová funkce okna** (stručně **okno**), jestliže rovněž $t\omega(t) \in L_2$. Střed t^* a poloměr Δ_ω okna ω jsou pak definovány vztahy

$$t^* = \frac{1}{\|\omega\|_2^2} \int_{-\infty}^{\infty} t |\omega(t)|^2 dt \quad (4.6)$$

a

$$\Delta_\omega = \frac{1}{\|\omega\|_2} \left\{ \int_{-\infty}^{\infty} (t-t^*) |\omega(t)|^2 dt \right\}^{\frac{1}{2}}$$

Číslo $2\Delta_\omega$ se nazývá **šířkou okna** ω .

Ve statistické terminologii můžeme t^* volně interpretovat jako „střední hodnotu“ a Δ_ω jako „směrodatnou odchylku“ rozložení výkonu $|\omega(t)|^2$ funkce signálu $\omega(t)$.

4.2.3 Waveletové vyhlazování (filtrace)

Uvažujme aditivní model

$$y(t) = x(t) + e(t); t \in R,$$

kde $y(t)$ jsou pozorované hodnoty, $x(t) \in L_2$, je neznámá odhadovaná funkce a $e(t)$ je bílý šum.

Vzhledem k linearitě (4.4) dostáváme

$$c_{j,k}(y) = c_{j,k}(x) + c_{j,k}(e)$$

kde $c_{j,k}(y)$, $c_{j,k}(x)$ a $c_{j,k}(e)$ jsou waveletové koeficienty pořadě ve waveletových řadách pro $y(t)$, $x(t)$ a $e(t)$. Cílem waveletového vyhlazování je nalezení vhodného modifikačního předpisu $\mu(\cdot)$ takového, že $\mu(c_{j,k}(y)) \approx c_{j,k}(x)$ je dobrým odhadem $c_{j,k}(x)$. Tento přístup představuje opět analogii s fourierovskou filtrací, kde modifikujeme Fourierovy koeficienty. Pro pevné k je příspěvek každého koeficientu $c_{j,k}(x)$ pouze lokální, takže waveletová reprezentace dovoluje tímto způsobem konstruovat **lokálně adaptivní filtry**.

Běžně se používají čtyři modifikační techniky pro waveletové koeficienty. První z nich operuje na datech (koeficientech) lineárně, zbývající tři nelineárně.

1. Positive scaling

$$\hat{c}_{j,k}^{pos} = \mu_{j,k} c_{j,k}(y), \quad \mu_{j,k} \geq 0,$$

kteřá je přímým zobecněním výše zmíněné přenosové charakteristiky.

2. Tvrdé prahování (hard thresholding)

Všechny waveletové koeficienty pod jistou prahovou hodnotou $\lambda > 0$ se vynulují a ostatní se ponechají beze změny:

$$\hat{c}_{j,k}^{hard} = \begin{cases} 0 & \text{pro } |c_{j,k}| < \lambda \\ c_{j,k} & \text{pro } |c_{j,k}| \geq \lambda \end{cases}$$

3. Měkké prahování (soft thresholding)

Velikosti všech waveletových koeficientů se sníží o prahovou hodnotu $\lambda > 0$

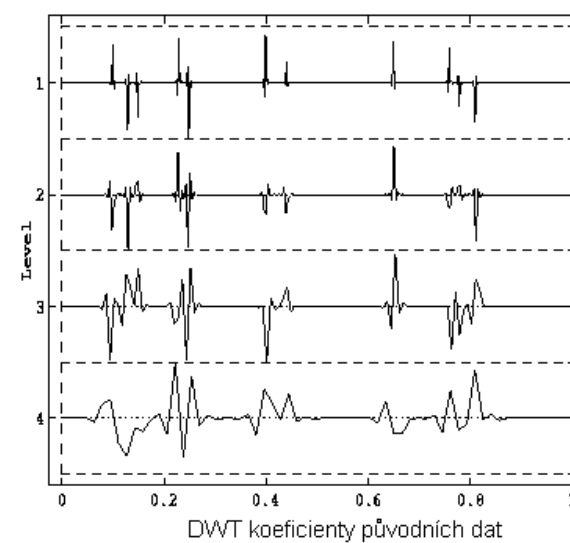
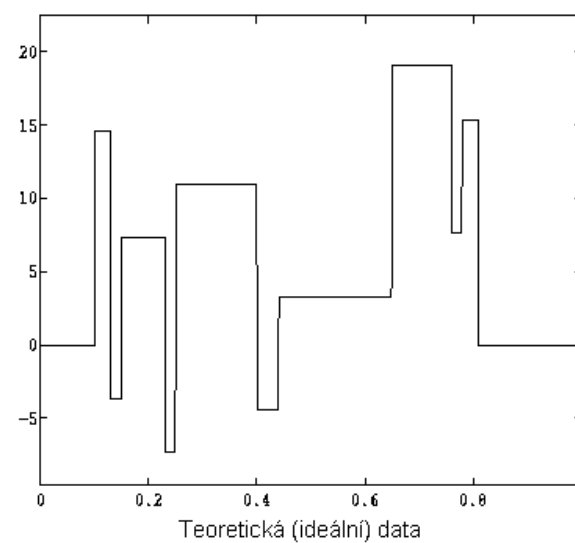
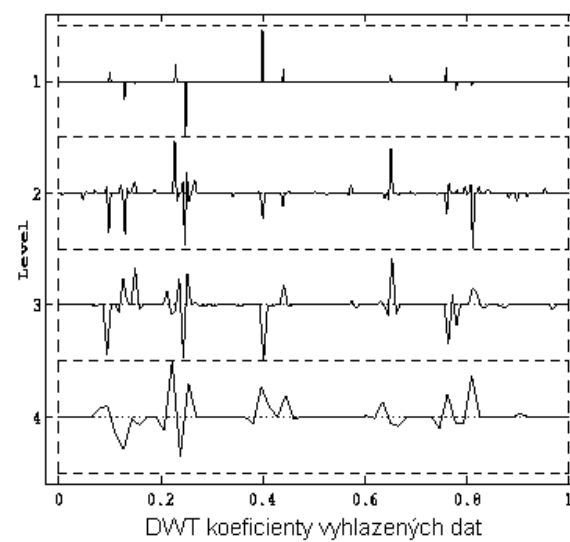
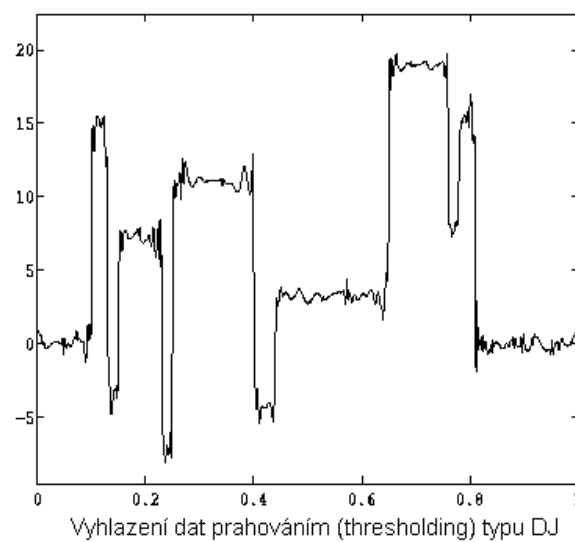
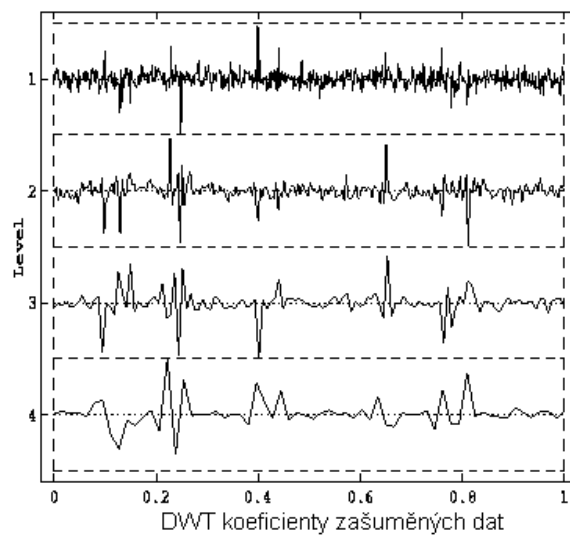
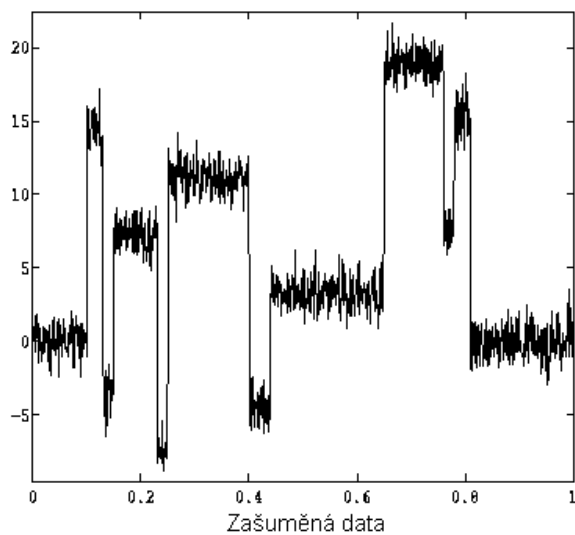
$$\hat{c}_{j,k}^{soft} = \text{sign}(c_{j,k}) \max(0, |c_{j,k}| - \lambda)$$

4. Kvantilové prahování (quantile thresholding)

Podobně jako 2, místo λ se však použije kvantil z množiny všech waveletových koeficientů, tj. např. se vynuluje 30 % nejmenších waveletových koeficientů.

Donoho a Johnstone navrhli metodu pro v jistém smyslu optimální volbu prahové hodnoty λ , která je buď univerzální $\lambda = \sigma \sqrt{2 \ln n}$, kde n je počet dat, nebo specifická pro každou úroveň měřítka j ($\lambda = \lambda_j$). Tyto a jiné podobné metody se staly známými pod anglickými názvy **wavelet shrinkage** nebo **wavelet de-noising**.

Na obr. 8 je ukázka waveletového vyhlazení signálu s nespojitými skoky kontaminovaného simulovaným šumem. V levém sloupci jsou odshora dolů uvedeny pořadě data zašuměná, vyhlazená a originální bez šumu. V pravém sloupci jsou grafy odpovídajících wavelet koeficientů, kde k je vyneseno na vodorovné ose a j na svislé ose (číslo 1 odpovídá maximální úrovni). Pro vyhlazení bylo použito Donoho-Johnstonova měkkého prahování.



Ob. 8 Ukázka vyhlazení měkkým práhovááním

4.3 MR-analýza

MR-analýza, neboli analýza o více úrovních rozlišení (z angl. multiresolution analysis), umožňuje jednoduchou aplikací vhodné dvojice lineárních filtrů postupně snižovat nebo zvyšovat rozlišovací schopnost waveletové aproximace.

Je-li X Banachův prostor nad C s normou $\|\cdot\|_X$ a $G \subset X$, pak $\tilde{\lambda}(G)$ značí **lineární obal** množiny G a \overline{G} její uzávěr v X . Jestliže $X = \overline{\tilde{\lambda}(G)}$, budeme říkat, že G **generuje** X .

Nechť ψ je mateřský wavelet. Pro každé $j \in Z$ uvažujme (uzavřené) podprostory v L_2 generované těmito částečnými součty waveletové řady:

$$W_j = \overline{\{\psi_{j,k}\}_{k \in Z}} = \{g_j(t) \mid g_j(t) = \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(t)\} \quad (4.7)$$

a

$$V_j = \overline{\{\psi_{i,k}\}_{i,k \in Z, i < j}} = \{x_j(t) \mid x_j(t) = \sum_{i=-\infty}^{j-1} \sum_{k=-\infty}^{\infty} c_{i,k} \psi_{i,k}(t)\} \quad (4.8)$$

Protože waveletový rozvoj (4.3) libovolné funkce $x \in L_2(R)$ lze přepsat jako

$$x(t) = \sum_{j=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(t) \right) = \sum_{j=-\infty}^{\infty} g_j(t), \text{ kde } g_j \in W_j \quad (4.9)$$

přičemž $g_j \in W_j$ jsou zřejmě určeny jednoznačně, můžeme celý prostor L_2 psát jako přímý součet ($\dot{+}$) podprostorů W_j

$$L_2 = \sum_{j \in Z} W_j = \dots \dot{+} W_{-1} \dot{+} W_0 \dot{+} W_1 \dot{+} \dots \quad (4.10)$$

Je-li mateřský wavelet ψ ortogonální, pak $W_i \perp W_j$ (podprostory jsou k sobě kolmé) pro $i \neq j$ a přímý součet (4.10) přejde v ortogonální součet (\oplus)

$$L_2 = \oplus_{j \in Z} W_j = \dots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \dots$$

Pro každé $j \in Z$ platí

$$V_j = \overline{\left(\bigcup_{i=-\infty}^{j-1} W_i \right)} = \sum_{i=-\infty}^{j-1} W_i = \dots \dot{+} W_{j-2} \dot{+} W_{j-1} \quad (4.11)$$

a systém podprostorů $\{V_j\} := \{V_j\}_{j \in Z}$ má následující vlastnosti:

1° $\dots V_{-1} \subset V_0 \subset V_1 \dots$ neboli $V_{j-1} \subset V_j, j \in Z$

$$2^\circ \overline{\lambda\left(\bigcup_{j \in Z} V_j\right)} = L_2$$

$$3^\circ \bigcap_{j \in Z} V_j = \{0\}$$

$$4^\circ V_{j+1} = V_j \dot{+} W_j, j \in Z$$

$$5^\circ x(t) \in V_j \Leftrightarrow x(2t) \in V_{j+1}, j \in Z \quad [\Rightarrow x(t) \in V_0 \Leftrightarrow x(2^j t) \in V_j, j \in Z]$$

Definice. Funkce $\phi(t) \in L_2$ generuje MR-analýzu $\{V_j\}$ v L_2 , jestliže generuje posloupnost podprostorů $\{V_j\}$ s vlastnostmi 1°, 2°, 3° a 5° předpisem

$$V_j = \overline{\lambda(\{\phi_{j,k}\}_{k \in Z})}, \text{ kde } \phi_{j,k} = 2^{j/2} \phi(2^j t - k) = 2^{j/2} \phi\left(\frac{t - k2^{-j}}{2^{-j}}\right), \quad (4.12)$$

přičemž $E_0 = \{\phi_{0,k}\}_{k \in Z}$ je Rieszovou bází podprostoru V_0 . Pak také $E_j = \{\phi_{j,k}\}_{k \in Z}$ jsou Rieszovy báze V_j pro každé $j \in Z$. Funkci $\phi(t)$ nazýváme funkcí měřítka (**scaling function**) pro MR-analýzu $\{V_j\}$. Podobně jako pro ψ platí, $\|\phi\|_2 = \|\phi_{j,k}\|_2 = 1$ pro každé $j, k \in Z$.

Tvrzení. Jestliže $\phi \in L_2$ generuje MR-analýzu $\{V_j\}$, pak

$$6^\circ x(t) \in V_j \Leftrightarrow x(t + 2^{-j}) \in V_j, j \in Z$$

Nemusí však existovat funkce ψ , která tyto W_j generuje vztahem (4.7). V důsledku toho se při definici MR-analýzy a funkce měřítka někdy místo 1°, 2°, 3°, 5° předpokládá platnost 1°, 2°, 5°, 6°, takže potom MR-analýza $\{V_j\}$ v L_2 má všechny vlastnosti 1° - 6°. Zde budeme však vždy předpokládat navíc (4.7). V tomto případě se někdy funkce měřítka ϕ nazývá **otcovský wavelet (father wavelet)**, neboť s mateřským waveletem ψ tvoří přirozenou dvojici. Říkáme potom, že dvojice (ϕ, ψ) generuje MR-analýzu $\{V_j\}$. Ukázky dvojic (ϕ, ψ) jsou na obr. 3, kde mateřský wavelet je ve sloupci b) a odpovídající otcovský wavelet ve sloupci a).

Necht' (ϕ, ψ) generují MR-analýzu $\{V_j\}$ v L_2 . Podle (4.8) a (4.9) je $x_j(t) = \sum_{i=-\infty}^{j-1} g_i(t) \approx \approx x(t) = \sum_{i=-\infty}^{\infty} g_i(t)$ pro každou funkci $x \in L_2$. Přitom $x_j \in V_j \Rightarrow x_j \xrightarrow{L_2} x$ pro $j \rightarrow \infty$.

Tedy $x_j \in V_j$ aproximuje x až do $(j-1)$ -té úrovně rozlišení.

Dostáváme následující vztahy:

$$x_j \in V_j \Rightarrow x_j(t) \stackrel{(2.12)}{=} \sum_{k=-\infty}^{\infty} c_k^j \phi(2^j t - k)$$

$$c^j := \{c_k^j\}_{k \in Z} \in \ell_2 \quad \forall j \in Z \quad (4.13)$$

$$g_j \in W_j \Rightarrow g_j(t) \stackrel{(2.7)}{=} \sum_{k=-\infty}^{\infty} d_k^j \psi(2^j t - k) \quad (4.14)$$

$$d^j := \{d_k^j\}_{k \in Z} \in \ell_2 \quad \forall j \in Z$$

$$4^\circ \Rightarrow x_j(t) = x_{j-1}(t) + g_{j-1}(t) \quad \forall j \in Z \quad (4.15)$$

kde posloupnosti souřadnic c^j a d^j jsou jednoznačně určeny, neboť

$$\underbrace{\{\phi(2^j t - k)\}_{k \in Z}}_{2^{-j/2} \phi_{j,k}} \text{ a } \underbrace{\{\psi(2^j t - k)\}_{k \in Z}}_{2^{-j/2} \psi_{j,k}} \text{ tvoří Rieszovu bázi pořadě ve } V_j \text{ a } W_j.$$

Dekompozice v MR-analýze (výpočet c^{j-1} a d^{j-1} pomocí c^j)

Při dekompozici snižujeme úroveň rozlišení o jedna. K tomu bude potřeba vyjádřit $\phi(2^j t - k) \in V_j$ z (4.13) ve tvaru (4.15).

Tvrzení: Existují posloupnosti $a := \{a_n\}_{n \in Z}$, $b := \{b_n\}_{n \in Z} \in \ell_2$ takové, že platí

$$\underbrace{\phi(2t - l)}_{\in V_1} = \sum_{k=-\infty}^{\infty} \left[\underbrace{a_{2k-l} \phi(t - k)}_{\in V_0} + \underbrace{b_{2k-l} \psi(t - k)}_{\in W_0} \right] \quad \forall l \in Z, t \in R \quad (4.16)$$

Důsledek. Platí tzv. **dekompoziční vztah (decomposition relation):**

$$\underbrace{\phi(2^j t - l)}_{\in V_j} = \sum_{k=-\infty}^{\infty} \left[\underbrace{a_{2k-l} \phi(2^{j-1} t - k)}_{\in V_{j-1}} + \underbrace{b_{2k-l} \psi(2^{j-1} t - k)}_{\in W_{j-1}} \right] \quad \forall j, l \in Z, t \in R \quad (4.17)$$

Tvrzení (algoritmus dekompozice: výpočet c^{j-1} a d^{j-1} z c^j)

Pro c^j a d^j z (4.13) a (4.14) platí následující rekurentní vztah:

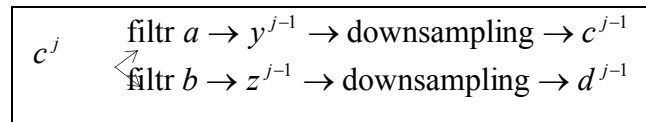
$$c_k^{j-1} = \sum_{l=-\infty}^{\infty} a_{2k-l} c_l^j \stackrel{2k-l=n}{=} \sum_{n=-\infty}^{\infty} a_n c_{2k-n}^j \quad (4.18)$$

$$d_k^{j-1} = \sum_{l=-\infty}^{\infty} b_{2k-l} c_l^j = \sum_{n=-\infty}^{\infty} b_n c_{2k-n}^j \quad (4.19)$$

kde $a = \{a_n\}_{n \in \mathbb{Z}}$, $b = \{b_n\}_{n \in \mathbb{Z}}$ jsou posloupnosti z dekompozičního vztahu (4.17)

Označíme-li $y^{j-1} = \{y_m^{j-1}\}_{m \in \mathbb{Z}}$ a $z^{j-1} = \{z_m^{j-1}\}_{m \in \mathbb{Z}}$ pro každé $j \in \mathbb{Z}$, pak (4.18) a (4.19) definují dva lineární konvoluční filtry pořadě s impulzivními odezvami a a b pro výpočet $y^{j-1} = a * c^j$ a $z^{j-1} = b * c^j$, kde $*$ je symbol operátoru lineární konvoluce. Po vynechání hodnot s lichými indexy (tzv. downsampling) dostáváme $c_k^{j-1} = y_{2k}^{j-1}$ a $d_k^{j-1} = z_{2k}^{j-1}$.

Schématicky tedy můžeme algoritmus dekompozice znázornit takto



Rekonstrukce v MR-analýze (výpočet c^j pomocí c^{j-1} a d^{j-1})

Při rekonstrukci naopak zvyšujeme úroveň rozlišení o jedna. K tomu je třeba vyjádřit $\phi(2^{j-1}t-k)$ a $\psi(2^{j-1}t-k)$ z (4.13) a (4.14) pomocí $\phi(2^j t-k)$

Tvrzení. Existují posloupnosti $p := \{p_n\}_{n \in \mathbb{Z}}$, $q := \{q_n\}_{n \in \mathbb{Z}} \in \ell_2$ takové, že platí

$$\phi(t) = \sum_{n=-\infty}^{\infty} p_n \phi(2t-n) \quad (4.20)$$

$$\psi(t) = \sum_{n=-\infty}^{\infty} q_n \phi(2t-n) \quad (4.21)$$

Důsledek. Platí tzv. **vztahy dvou měřítek (two-scale relations):**

$$\phi(2^{j-1}t-l) = \sum_{k=-\infty}^{\infty} p_{k-2l} \phi(2^j t-k) \quad (4.22)$$

$$\psi(2^{j-1}t-l) = \sum_{k=-\infty}^{\infty} q_{k-2l} \phi(2^j t-k) \quad (4.23)$$

Tvrzení: (Algoritmus rekonstrukce: výpočet c^j z c^{j-1} a d^{j-1}).

Pro c^j a d^j z (4.13) a (4.14) platí následující rekurentní vztah:

$$c_k^j = \sum_{l=-\infty}^{\infty} [p_{k-2l} c_l^{j-1} + q_{k-2l} d_l^{j-1}] = \sum_{l=-\infty}^{\infty} p_{k-2l} c_l^{j-1} + \sum_{l=-\infty}^{\infty} q_{k-2l} d_l^{j-1} \quad (4.24)$$

kde $p = \{p_n\}_{n \in \mathbb{Z}}$ a $q = \{q_n\}_{n \in \mathbb{Z}}$ jsou posloupnosti ze vztahů dvou měřítek (4.22) a (4.23).

Označíme-li $y^{j-1} = \{y_m^{j-1}\}_{m \in \mathbb{Z}}$ a $z^{j-1} = \{z_m^{j-1}\}_{m \in \mathbb{Z}}$ pro každé $j \in \mathbb{Z}$, kde pro každé $l \in \mathbb{Z}$

klademe

$$\left. \begin{array}{l} y_{2l}^{j-1} = c_l^{j-1} \quad z_{2l}^{j-1} = d_l^{j-1} \\ y_{2l+1}^{j-1} = 0 \quad z_{2l+1}^{j-1} = 0 \end{array} \right\} \text{tzv. upsampling}$$

pak vztah (4.24) určuje dva konvoluční filtry po řadě s impulzivními odezvami p a q pro výpočet $c^j = p * y^{j-1} + q * z^{j-1}$.

Schematicky tedy můžeme algoritmus rekonstrukce znázornit takto.

$$\boxed{c^j \quad \begin{array}{l} + \leftarrow \text{filtr } p \leftarrow y^{j-1} \leftarrow \text{upsampling} \leftarrow c^{j-1} \\ \leftarrow \text{filtr } q \leftarrow z^{j-1} \leftarrow \text{upsampling} \leftarrow d^{j-1} \end{array}}$$

4.3.1 Dekompozice a rekonstrukce signálu

Velice efektivním a často používaným způsobem výpočtu Wavelet transformace je užití Mallatovy pyramidální struktury pro určení příslušných koeficientů pro daný sloupcový vektor $\{x(n)\}_{n=0}^{N-1}$. Z hlediska metod zpracování signálů se jedná o užití nízkofrekvenčního filtru pro mezní frekvenci v polovině příslušného frekvenčního pásma pro funkci měřítka (a určení aproximativních složek signálu) a dále o aplikaci vysokofrekvenčního (Wavelet) filtru $\{(n)\}_{n=0}^{L-1}$ pro stanovení detailních složek signálu. Tyto analyzující posloupnosti jsou užity v konvoluci s daným signálem $\{x(n)\}$ a příslušný výsledek je dále podvzorkován faktorem 2. Podobné dekompoziční schéma je použitelné i v případě analýzy obrazů s tím, že příslušná dekompozice se provádí nejprve po řádcích a posléze po sloupcích vždy pro

nízkofrekvenční a vysokofrekvenční filtr s následným podvzorkováním. Výsledkem jednoho kroku dekompozice jsou tedy v případě jednorozměrných signálů dvě funkce a v případě obrazů jsou to čtyři dílčí obrazové složky.

4.4 Aplikace

4.4.1 Detekce nelinearit a segmentace časových řad

Pro potřeby zpracování nestacionárních signálů je nutné provést detekci změn a nalézt hranice jednotlivých segmentů. Nalezené segmenty je možno samostatně zpracovávat, klasifikovat, provádět predikci apod. V současné době se ukazuje Wavelet analýza jako velice výhodná metoda pro segmentaci. Jak je popsáno výše, je založena na aplikaci nízkofrekvenčních a vysokofrekvenčních filtrů na daný signál. Výsledkem jsou dvě nové posloupnosti, které jsou označovány jako aproximační a detailní složka. Tyto názvy vychází z předpokladu, že každý signál má důležité informace uloženy v nízkých frekvencích, které postačí k aproximaci daného signálu, zatímco vysoké frekvence nesou informace pouze zpřesňující, dokreslující, tedy detaily daného signálu.

Rozklad pomocí wavelet analýzy do 1. rozlišovací hladiny je možno použít pouze u testovacích signálů, které neobsahují příliš mnoho frekvenčních složek. U technických, biomedicínckých, elektrotechnických a dalších signálů je potřeba celý postup opakovat. Celý postup odpovídá dekompozici podle Mallatova pyramidálního schématu. Pro dekompozici do druhé úrovně použijeme aproximační složku z úrovně první. Dostaneme tak novou detailní a aproximační složku. To znamená, že po dvou krocích máme signál rozložen do dvou detailních a jedné aproximační složky.

4.5 Informace o software

Existuje několik komerčních i nekomerčních podpurných softwarových balíků pro práci s wavelety, z nichž bych zmínil tyto:

WAVELET TOOLBOX: Firemní komerční knihovna (toolbox) pro numerický výpočetní systém MATLAB firmy MathWorks, Inc. (USA). V České republice jej spolu se systémem MATLAB a dalšími firemními toolboxy distribuuje firma HUMUSOFT, s. r. o. Praha.

WAVBOX 4: Komerční nefiremní toolbox pro MATLAB, jehož starší verze jsou nekomerční a volně dostupné přes FTP. Autorem je Carl Taswell, Stanford University, USA.

DPWT Toolbox, tools for working with the discrete periodic wavelet transform (DPWT).
N.H. Getz, "A discrete periodic wavelet transform", UCB/ERL M92/138,
Electronics Research Laboratory, University of California, Berkeley,
December 1992.
<http://www.InversionInc.COM/wavelet.html>

MULTISCALE METHODS
STATISTICS IN TIME/FREQUENCY AND TIME/SCALE DOMAINS
B. Vidakovic, Statistical Modeling By Wavelets, Wiley, 1999 [ISBN 0-471-29365-2]
<http://www.isds.duke.edu/~brani/TFM.html>

TIME FREQUENCY TOOLBOX, Version 1.0 January 1996
Copyright (c) 1994-96 by CNRS (France) - RICE University (USA)
<http://www-isis.enst.fr/Applications/tftb/iutsn.univ-nantes.fr/auger/tftbftp.html>

WAVELAB 802

Library of Matlab routines for wavelet analysis, wavelet packet analysis, cosine-packet analysis and matching pursuit. The library is available free of charge over the Internet. Versions are provided for Macintosh, UNIX and Windows machines. WaveLab has been used in teaching courses in adapted wavelet analysis at Stanford and at Berkeley.

TSA (Time Series Analysis) Toolbox 2.40

E-Mail: a.schloegl@ieee.org

WWW: <http://www-dpml.tu-graz.ac.at/~schloegl/matlab/tsa>

WAVEKIT toolbox for Matlab by Harri Ojanen

(C) 1998 Harri Ojanen

The documentation is available on-line at

<http://www.math.rutgers.edu/~ojanen/wavekit/>

4.6 Závěr

Wavelet transformace představuje moderní matematický aparát s rozsáhlými aplikacemi v řadě oborů. Význam využití Wavelet funkcí spočívá přitom zejména v možnosti jejich výběru podle chování dané časové řady nebo obrazu a v možnostech různého způsobu dekompozice, úpravy a rekonstrukce signálu. Z těchto důvodů je tato tematika předmětem širokého zájmu matematiků i inženýrů v oblasti teoretického popisu Wavelet funkcí, jejich vlastností a dále v oblasti jejich využití. S tím souvisí i řada diskusních skupin na Internetu, z nichž mezi nejaktivnější patří skupina, kterou lze nalézt na adrese <http://www.wavelet.org>, a která pravidelně informuje o konferencích a publikacích zaměřených na Wavelet transformaci.

Pokud si na závěr položíme otázku, zda se při analýze dat vyplatí používat wavelety či nikoli, pak odpověď samozřejmě závisí na typu zpracovávaných dat. Vzhledem k mnoha dobrým vlastnostem představují však rozhodně aparát, který není vhodné ignorovat.

Ve prospěch waveletů hovoří především myšlenková jednoduchost blízká zaběhaným fourierovským přístupům. Oproti nim však nabízí větší pružnost danou lokálním charakterem příspěvků jednotlivých waveletových koeficientů ve waveletové řadě. Například při waveletové filtraci tedy postupujeme obdobně jako při fourierovské filtraci, ovšem s tím podstatným rozdílem, že při redukci $c_{j,k}$ je vliv shlazení pouze lokální v závislosti na zvoleném k a nikoliv globální. Waveletový filtr tedy poskytuje významnou novou kvalitu v tom, že se může adaptovat na lokální charakter dat. Tam kde data vykazují vyšší dynamiku, tj. nesou užitečnou informaci ve vyšších frekvenčních pásmech, můžeme míru redukce vysokofrekvenčních komponent snížit (neboli snížit prahovou hodnotu signifikantnosti) a naopak postupovat v místech, kde je změna jen pozvolná. Tím se odstraňuje hlavní nevýhoda fourierovské filtrace, která v takovém případě má tendenci dynamický úsek přehlédnout. Typickým příkladem může být rozvlnění v okolí nespojitých změn (Gibbsův jev). Toto je ovšem spíše problém volby správné strategie vyhlazení, než samotného principu waveletové filtrace.

Dalším nezanedbatelným argumentem ve prospěch waveletů je také výpočetní efektivita. Jestliže pro výpočet DFT existují rychlé algoritmy FFT s výpočtovou složitostí řádu $O(n \log_2(n))$, pak v případě FWT složitost mnohdy klesá dokonce až na $O(n)$.

Jako žádná jiná metoda, nejsou ovšem ani wavelety universální všelék pro všechny typy úloh, ale v každém případě významně obohacují repertoár dosud běžně používaných technik.

4.7 Literatura

- Akansu, A. N. – Haddad, R. A. *Multiresolution Signal Decomposition*. Academic Press, Inc., Boston, USA., 1992.
- Al-Adnani, A. M. S. – Nandi, A. K. – Chapman, R. – MacDougall, S. *Multiresolution Methods for Singularity Detection*. In *Signal Analysis and Prediction I*, pages 109-112. ICT Press, 1997.
- Arino, M. A. *Time Series Forecasts Via Wavelets*. IESE Universidad de Navarra. Barcelona 1995.
- Arino, M. A. – Vidakovic, B. *On wavelet scalograms and their applications in economic time series*. Discussion Paper 95-21, ISDS, Duke University. 1995.
- Crouse, M. S. – Nowak, R. D. – Baraniuk, R. G. *Wavelet-Based Statistical Signal Processing Using Hidden Markov Models*. IEEE Trans. Signal Processing, Special Issue on Theory and Application of Filter Banks and Wavelet Transforms, 46(4):886-902, April 1998.
- Chui, K. *An Introduction to Wavelets*. Academia Press. 1992.
- Daubechies, I. *The Wavelet Transform, Time-Frequency Localization and Signal Analysis*. IEEE Trans. on Inform. Theory, 36:961, September 1990.
- Daubechies, I. *Ten Lectures on Wavelets*. SIAM, CBMS-NST Conference Series 61.
- Donoho, Bruce, D. – Gao, Hong-Ye. *Wavelet Analysis*. IEEE Spectrum, 33(10):26-35, October 1996.
- Etter, W. *Restoration of Discrete-Time Signal Segments by Interpolation Based on the Left-Sided and Right-Sided Autoregressive Parameters*. IEEE Trans. on Signal Processing, 44(5):1124-1135, May 1996.
- Hayes, M. H.. *Statistical Digital Signal Processing and Modeling*. John Wiley and Sons, Inc., 1996.
- Haykin, S. *Neural Networks*. IEEE Press, New York, 1994.
- Holben, B. – Vermote, E. – Kaufman, Y. J. - Tanrie, D. – Kalb, V. *Aerosol Retrieval over Land from AVHRR Data-Application for Atmospheric Correction*. IEEE Trans. on Geoscience and Remote Sensing, 30(2):212-222, March 1992.
- Chan, Y. T. *Wavelet Basics*. Kluwer Academic Publishers, Boston, 1995.
- Kay, S. M. *Fundamentals of Statistical Signal Processing*. Prentice Hall, 1993.

- Kingsbury, N. G. – Magarey, J. *Wavelet Transforms in Image Processing*. Birkhauser Boston, 1998.
- Lim, J. S. *Two-Dimensional Signal and Image Processing*. Prentice Hall, 1990.
- Louis, A. K. – Mass, P. – Rieder, A. *Wavelets: Theory and Applications*. John Wiley & Sons, Chichester, U. K., 1997.
- Magarey, J. – Kingsbury, N. G. *Motion Estimation Using a Complex-Valued Wavelet Transform*. IEEE Trans. Signal Processing, Special Issue on Theory and Application of Filter Banks and Wavelet Transforms, April 1998.
- Misiti, M. – Misiti, Y. – Oppenheim, G. – Poggi, J. M. *Wavelet Toolbox*. The MathWorks, Inc., Natick, Massachusetts 01760, March 1996.
- Newland, D. E. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*. Longman Scientific & Technical, Essex, U. K., third edition, 1994.
- Polikar, R. *The Wavelet Tutorial*. Iowa State University. 1999.
<http://www.public.iastate.edu/~rpolikar/WAVELETS>.
- Procházka, A. – Sláma, M. – Pelikán, E. *Bayesian Estimator Use in Signal Processing*. Neural Network World, 6(2):209{213, 1996.
- Procházka, A. – Štorek, M.. *Wavelet Transform Use for Signal Classification by Self-Organizing Neural Networks*. In Fourth International Conference on Artificial Neural Networks ANN-95. IEE, Cambridge, England, 1995.
- Rayner, P. J. W. – Fitzgerald, W. J. *The Bayesian Approach to Signal Modelling and Classification*. In Signal Analysis and Prediction I. ICT Press, 1997.
- Satorie, M. *Metody minimalizace dynamických chyb měření teplotních profilů zemské kůry*. PhD thesis, ČVUT, Fakulta elektrotechnická, 1998.
- Simhadri, K. K. – Iyengar, S. S. – Holyer, R. J. – Lybanon, M. – Zachary, J. M.. *Wavelet-Based Feature Extraction from Oceanographic Images*. IEEE Trans. on Geoscience and Remote Sensing, 36(3):767-778, May 1998.
- Strang, G. – Nguyen, T. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.
- Veselý, V. *Wavelety a jejich použití při filtraci dat*. In ROBUST'1996. Praha 1997, s. 241-272.
- Vetterli, M. *Wavelets and Filter Banks: Theory and Design*. IEEE Trans. Signal Processing, 40(9):2207-2232, September 1992

Případové studie

Analýza ekonomických dat - inflace

V současné době jsou často požadovány údaje o míře inflace, a to nejen pro potřeby ekonomiky, ale také z řad podnikatelů a veřejnosti. Nezanedbatelnou informací jsou také prognózy týkající se dalšího vývoje inflace v následujícím období. Je známo, že odhady míry inflace, které zveřejňují různí odborníci, se liší až o 4%. Cílem článku je ukázat, jak různé statistické modely vývoje inflace vedou k různým prognózám míry inflace.

Úvod

Inflace je pojem, se kterým se setkáváme takřka denně, většinou bez přesného vymezení. Přijímáme pouze konstatování, že inflace vzrostla o tolik a tolik procent, aniž bychom podrobněji zkoumali metody výpočtu a vypovídající schopnosti použitých charakteristik.

Inflace je definována jako projev ekonomické nerovnováhy, jehož vnějším znakem je růst cenové hladiny - viz [6].

Cenová hladina (P) představuje průměrnou úroveň cen určitého souboru statků v běžném období (ceny p_t) ve srovnání s cenami určitého vybraného základního období (p_0). Cenové hladině v základním období je přiřazen index 100,00 (tedy $P_0 = 100,00$).

Vývoj cenové hladiny může být vyjadřován uvedením indexů z jednotlivých období (P_0, P_1, P_2, \dots), častěji se však používá tempo jejich růstu. Jde tedy o tempo růstu cenové hladiny neboli o **míru inflace**.

Výraz

$$\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}} 100 \quad (1)$$

vyjadřuje, o kolik procent vzrostla cenová hladina, vztahující se k období t , ve srovnání s cenovou hladinou, vztahující se k období $t-1$.

Index spotřebitelských cen

Nejčastěji používanými indexy k vyjádření obecné cenové hladiny jsou:

- **index spotřebitelských cen ISC**, který měří změnu hladiny cen tzv. spotřebního koše, tj. vybraného souboru reprezentativního zboží a služeb.
- **index cen výrobců ICV**, který měří změnu cenové hladiny vybraných výrobků při prvním prodeji, tedy při prodeji výrobcem, a je využíván zejména v podnikové sféře. Všeobecně vývoj ICV signalizuje nadcházející změny ISC.
- **cenový deflátor hrubého domácího produktu, deflátor HDP**, který měří změnu cenové hladiny hrubého domácího produktu jako celku.

Ve světě i u nás je **nejfrekventovanější mírou inflace index spotřebitelských cen**, který zachycuje míru změny cenové hladiny na základě tržních cen vybraných reprezentantů, tj. zboží a služeb, za které skutečně nakupuje obyvatelstvo v konečné fázi ve vybraných prodejnách a provozovnách.

Index spotřebitelských cen se počítá podle Laspeyresova vzorce, tj. s váhami základního období. **Laspeyresův agregátní index** má tvar

$$P = \frac{\sum_{i \in I} \frac{P_{1i}}{P_{0i}} w_{si}}{\sum_{i \in I} w_{si}} \cdot 100 = \frac{\sum_{i \in I} \frac{P_{1i}}{P_{0i}} p_{0i} q_{0i}}{\sum_{i \in I} p_{0i} q_{0i}} \cdot 100, \quad (2)$$

kde

p_{1i} je cena i-tého reprezentanta v období sledovaném,

p_{0i} je cena i-tého reprezentanta v období základním,

q_{0i} je množství i-tého reprezentanta v období základním,

$w_{si} = p_{0i} q_{0i}$ je stálá váha (strukturní ukazatel hodnoty) i-tého reprezentanta.

Vývoj spotřebitelských cen se sleduje na **spotřebním koši**, který tvoří soubor vybraných druhů zboží a služeb placených obyvatelstvem. Celkový počet cenových reprezentantů je v současné době $I = 755$ konkrétních výrobků a služeb, které jsou rozděleny do 10 relativně homogenních tříd (Potraviny, nápoje, tabák: Odívání: Bydlení: Zařízení a provoz domácnosti: Zdravotnictví: Doprava: Volný čas: Vzdělání: Veřejné stravování a ubytování: Ostatní zboží a služby). Podrobně je spotřební koš popsán v publikaci Českého statistického úřadu „Indexy spotřebitelských cen (životních nákladů), revize 1994, která byla vydána v roce 1995 v edici Česká statistika.

Spotřební koš a tedy i váhy jednotlivých reprezentantů jsou po určitou dobu (např. 5 let) fixní - jejich statistické zjišťování je totiž velmi náročné. Tyto stálé váhy jsou hlavní slabinou indexů Laspeyresova typu (nadhodnocují skutečné tempo inflace), protože spotřební zvyklosti obyvatelstva se časem mění. Dalšími problémy jsou zohlednění postupné změny kvality statků obsažených v koši a zohlednění nových statků, v období konstrukce koše do něj nezařazených. Po roce 2000 se uvažuje o tom, že se budou váhy v indexní formuli (2) měnit každoročně.

Index spotřebitelských cen vyjadřuje **vývoj životních nákladů průměrné domácnosti**. Pro vyjádření vývoje těchto nákladů v různých skupinách (např. zaměstnanci, zemědělci, důchodci) jsou sestavovány speciální indexy životních nákladů.

Přes uvedené nedostatky nemáme v současné době za index spotřebitelských cen (2) vhodnou náhradu, která by snadno a hodnověrně odrážela pohyby inflace - viz [9].

Časová řada a její vyrovnávání

Příčinami, formami a důsledky inflace se nebudeme zabývat. Na základě analýzy údajů z let 1985 až 1997 - viz tabulku 1 - se pokusíme o prognózu míry inflace pro rok 1998.

Inflační skok v roce 1991 byl způsoben liberalizací převážné míry cen zboží a služeb, v roce 1993 pak reformou daňové soustavy. Pro rok 1998 zavedlo Ministerstvo financí nejprve odhad míry inflace na 9,8%. Pod tlakem nepříznivého vývoje inflace v prvních měsících letošního roku bylo Ministerstvo financí nuceno revidovat svou optimistickou prognózu průměrné míry inflace pro rok 1998 na 11,0 - 11,7%. Odhad Českého statistického úřadu i nezávislých ekonomů byl však vyšší (13,0%).

Hodnoty meziroční míry inflace, uvedené v tabulce 1, tvoří časovou řadu. **Časová řada** hodnot y_t , $t=1,2,\dots,n$, může obsahovat následující složky: trend T_t , periodické kolísání P_t a náhodnou složku ε_t . Některé složky (např. sezónnost) mohou v časové řadě chybět.

Tab.1 Dlouhodobý vývoj inflace v letech 1960 – 1997

Pramen: Statistická ročenka ČR 1997

Rok	ISC v %	Míra inflace, nárůst v %
1970	100,0	2,2
1980	111,2	2,9
1985	122,9	2,3
1986	123,5	0,5

1987	123,6	0,1
1988	125,8	0,2
1989	125,5	1,4
1990	137,7	9,7
1991	215,6	56,6
1992	239,5	11,1
1993	289,3	20,8
1994	318,2	10,0
1995	347,2	9,1
1996	377,6	8,8
1997	409,7	8,5

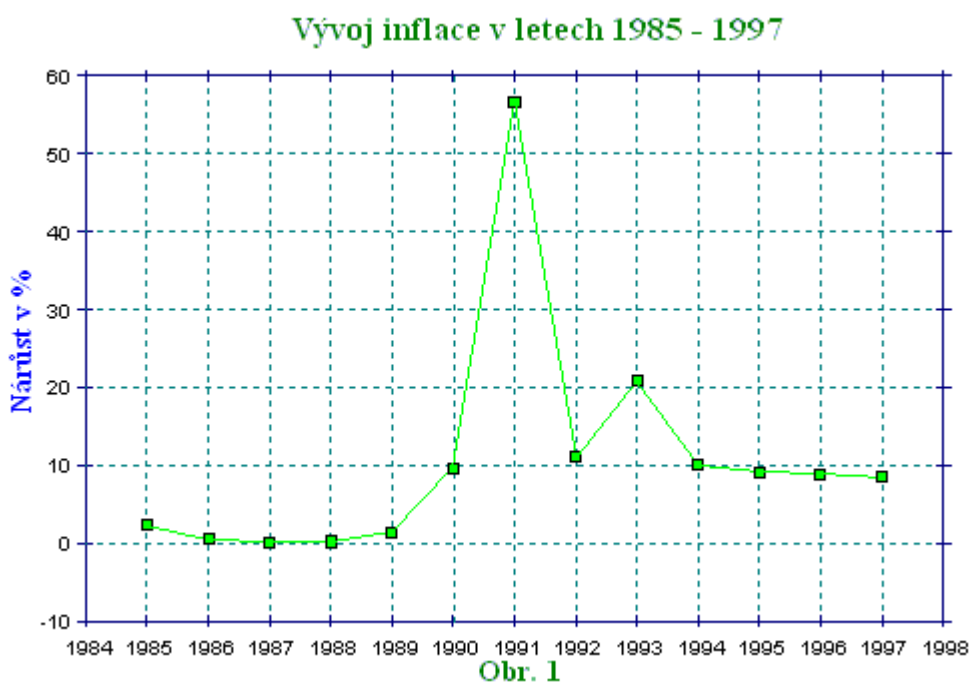
Obvykle se předpokládá, že analyzovaná časová řada má tvar

$$y_t = T_t + \varepsilon_t. \quad (3)$$

Při hledání nejvhodnějšího typu trendu nejčastěji hledáme spojitou funkci času

$$\hat{T}_t = f(t), \quad t = 1, 2, \dots, n., \quad (4)$$

kde \hat{T}_t je odhad trendu.

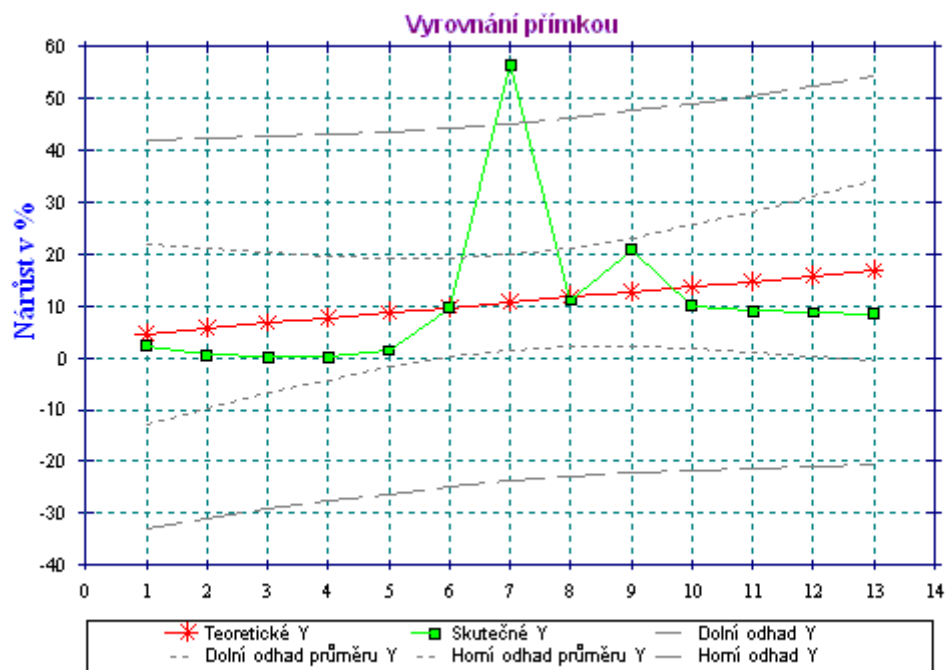


Jde-li o funkce lineární v parametrech (přímka, jednoduchá parabola, případně polynom vyššího stupně) - viz obr. 2 a obr. 3, využíváme pro výpočet odhadů parametrů známou **metodu regresní analýzy**, metodu nejmenších čtverců, která požaduje, aby reziduální součet čtverců byl minimální, tj.

$$S_R = \sum_{t=1}^n (y_t - \hat{T}_t)^2 = \min. \quad (5)$$

Známe-li trend časové řady, čili známe hlavní tendenci ve vývoji analyzovaného ukazatele v čase, můžeme modelovat i další vývoj trendu v budoucnosti, ovšem za předpokladu, že se jeho charakter v podstatě nezmění.

Míra inflace v letech 1985 - 1997



Obr. 2

Vzhledem k tomu, že k vyjádření trendu se používají i některé složitější funkce, které nejsou lineární vzhledem k parametrům (exponenciální křivka, různé S-křivky atd.)- viz obr. 4 a obr. 5, používáme k odhadu parametrů těchto křivek iterační metody nelineární regrese, např. Marquardtovy metody.

Právě uvedené postupy vycházejí z předpokladu, že v průběhu sledované doby se parametry modelu časové řady nemění a že tyto parametry lze v rámci celé časové řady odhadnout „naráz“ na základě tzv. **analytického vyrovnávání**. Metody analytického vyrovnávání tedy dávají všem pozorováním stejné váhy, což je pro předvídaní budoucího vývoje předpoklad velmi omezující. Předpověď do budoucna bezpochyby více závisí na datech nedávných a kromě toho odlišný vývoj ve vzdálené minulosti může velmi zkreslit předpovědi do budoucna.

Realističtější přístup k modelování časových řad poskytují **adaptivní metody**. Adaptivní metody vycházejí z předpokladu, že pro konstrukci extrapolační prognózy budoucího vývoje jsou nejcennější nejnovější pozorování časové řady. Adaptivní metody tedy berou v úvahu „stárnutí“ informací. Nejrozšířenější adaptivní metodou je **metoda exponenciálního vyrovnávání**. Její princip zohledňuje stáří dat jejich vážením, kdy váha každého pozorování je nepřímo úměrná jeho věku.

Při exponenciálním vyrovnávání předpokládáme, že v časovém okamžiku t , který představuje pozorování v přítomném čase, máme k dispozici řadu hodnot y_{t-k} , kde jednotlivá $k = 0, 1, \dots, t-1$ interpretujeme jako „stáří“ pozorování, čili vztah (3) píšeme ve tvaru

$$y_{t-k} = T_{t-k} + \varepsilon_{t-k}. \quad (6)$$

Trendovou složku T_{t-k} ve vztahu (6) předpokládáme ve tvaru

$$T_{t-k} = a_0 - a_1 k + a_2 k^2 + \dots + (-1)^k a_k k^k, \quad k = 0, 1, \dots, t-1, \quad (7)$$

kde k je časová proměnná, kterou lze chápat jako „věk“ pozorování z hlediska časového okamžiku t .

Odhad parametrů trendové složky (4) počítáme váženou metodou nejmenších čtverců tj. minimalizací výrazu

$$\sum_{k=0}^{n-1} (y_{t-k} - \hat{T}_{t-k})^2 w_k = \sum_{k=0}^{n-1} (y_{t-k} - \hat{T}_{t-k})^2 \alpha^k = \min. \quad (8)$$

Váha

$$w_k = \alpha^k, \quad 0 < \alpha < 1, \quad k = 0, 1, \dots, t-1, \quad (9)$$

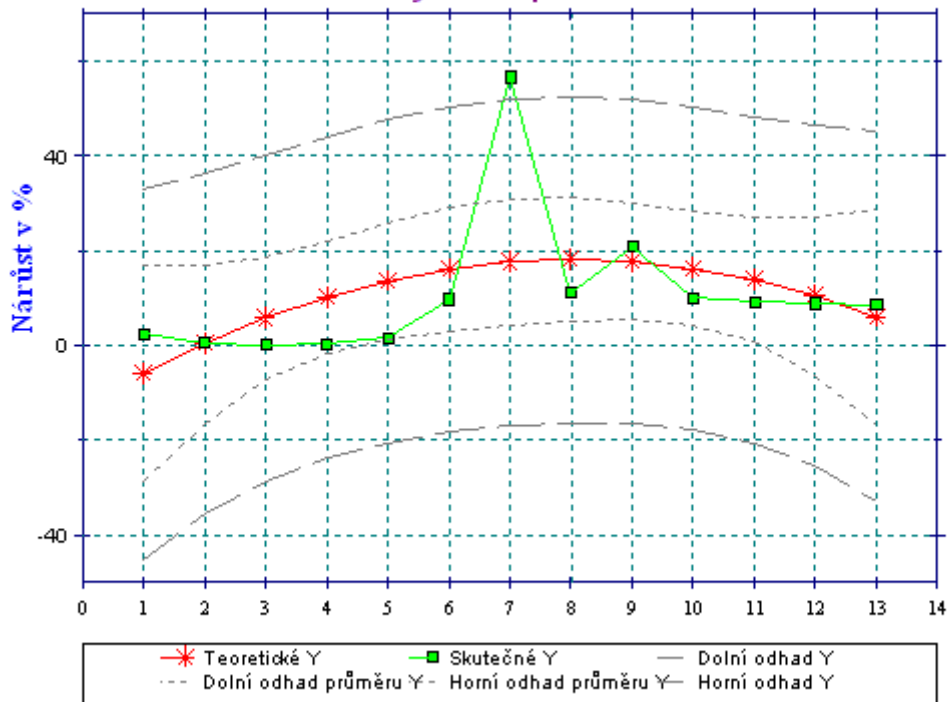
kde α se nazývá vyrovnávací konstanta.

Váha w_k je klesající exponenciální funkcí „stáří“ pozorování k a vyrovnávání časových řad na uvedeném principu se proto nazývá exponenciální vyrovnávání. Tyto metody se rozlišují podle typu předpokládaného trendu (jsou propracovány metody vyrovnání pro

časové řady s trendem konstantním, přímkovým či kvadratickým), kdy se postupně hovoří o jednoduchém, dvojitém či trojitém vyrovnávání a dále podle přítomnosti či nepřítomnosti sezónní složky

Míra inflace v letech 1985 - 1997

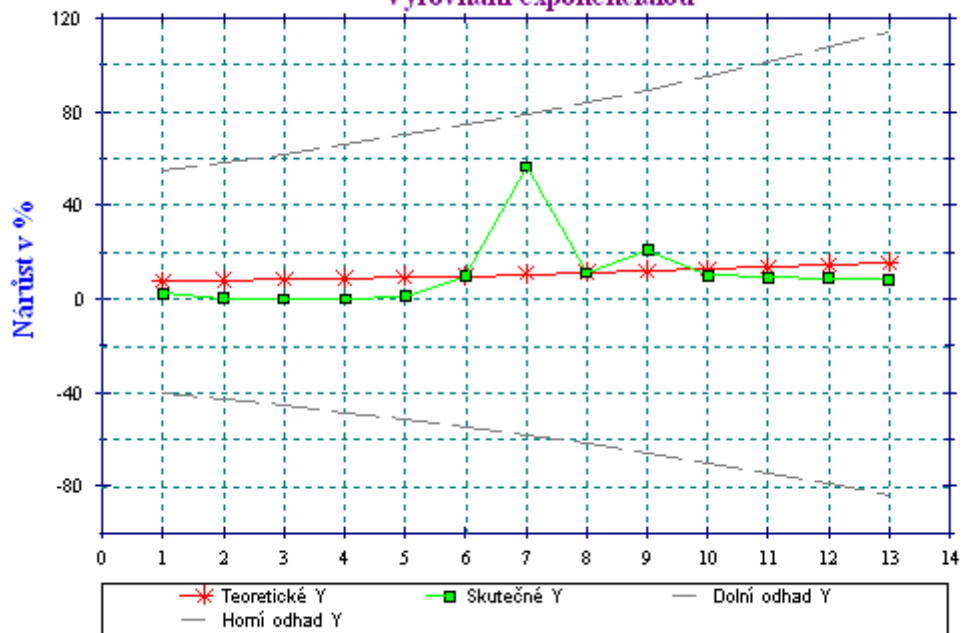
Vyrovnání parabolou



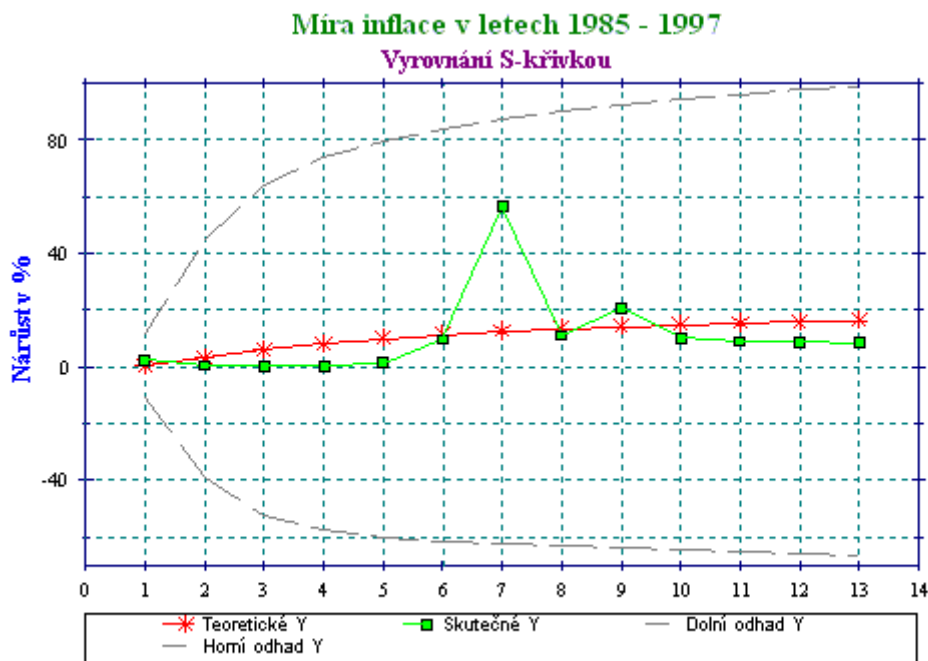
Obr. 3

Míra inflace v letech 1985 - 1997

Vyrovnání exponenciálou



Obr. 4



Obr. 5

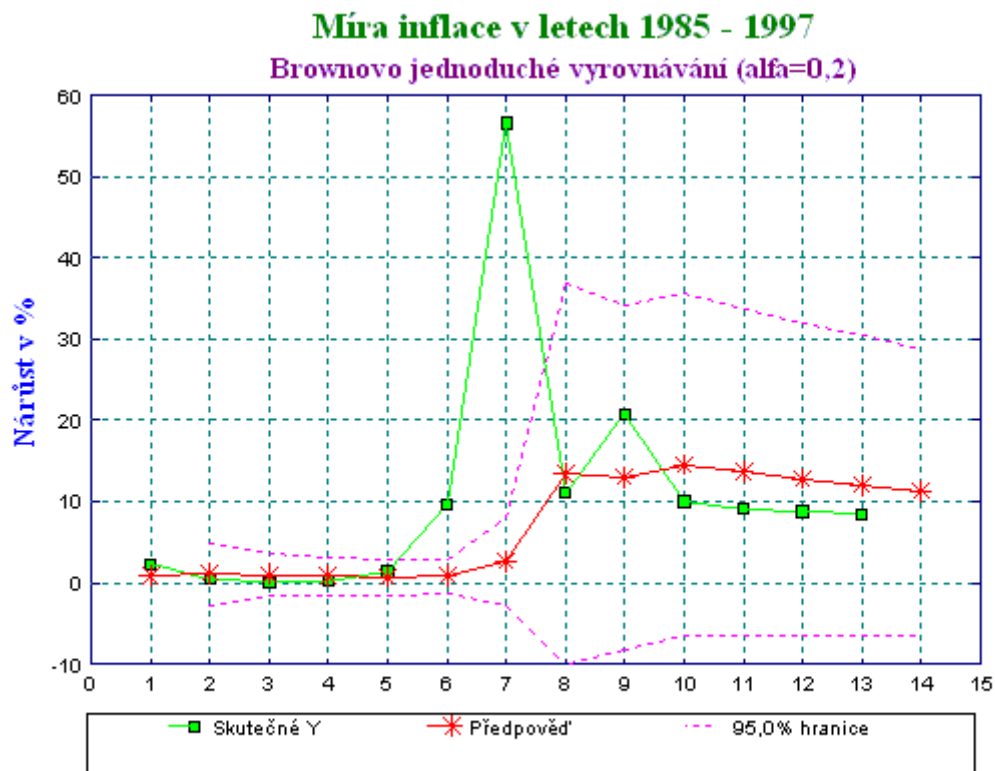
Jednotlivé metody exponenciálního vyrovňování se liší počtem vyrovňovacích konstant, který se pohybuje od jedné do tří. V literatuře a v nabídce příslušných programů pro statistickou analýzu časových řad nalezneme tyto metody často pod jmény jejich autorů jako Brownovo, Holtovo a Wintersovo vyrovňování. Podrobnější informace o těchto metodách nalezne zájemce v [7].

Statistický program Statistica, pomocí kterého bylo provedeno vyrovňování časové řady hodnot roční míry inflace, nabízí pro vyrovňování a předpovědi jednoparametrické (α) Brownovo jednoduché vyrovňování ($\hat{T}_{t-k} = a_0$), dvouparametrické (α, β) Holtovo lineární vyrovňování a tříparametrické (α, β, γ) Wintersovo (aditivní a multiplikativní) vyrovňování pro sezónní časové řady.

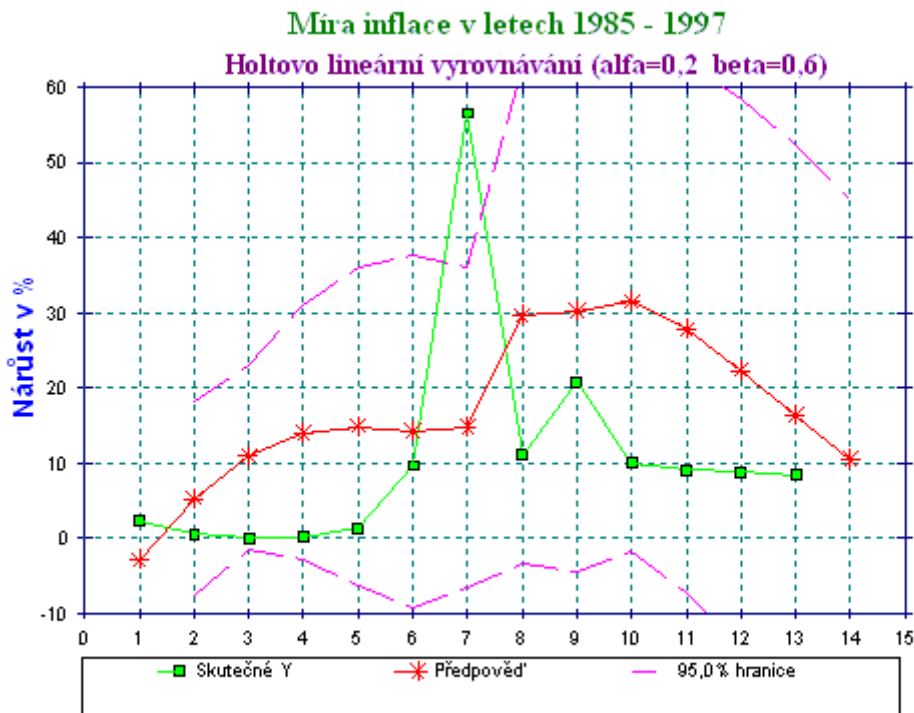
Rekurentní vzorec **Brownova jednoduchého vyrovňování** počítá vyrovnanou hodnotu \hat{T}_t v čase t pomocí vztahu

$$\hat{T}_t = \alpha y_t + (1 - \alpha) \hat{T}_{t-1}, \quad (10)$$

kde y_t je skutečná hodnota řady, \hat{T}_{t-1} značí vyrovnanou hodnotu v čase t-1, α je vyrovnávací konstanta. Počáteční hodnota \hat{T}_0 je průměr z první čtvrtiny dat zkoumané časové řady.



Obr. 6



Obr. 7

Holtovo lineární vyrovnávání předpokládá vyrovnanou hodnotu v čase t ve tvaru součtu

$$\hat{T}_t = m_t + b_t, \quad (11)$$

kde

$$m_t = \alpha y_t + (1 - \alpha)(m_{t-1} + b_{t-1}) \text{ je hodnota v čase } t, \quad (12)$$

α je vyrovnávací konstanta, $\alpha \in (0,1)$,

$$b_t = \beta(m_t - m_{t-1}) + (1 - \beta)b_{t-1} \text{ je trend v čase } t, \quad (13)$$

β je trendová vyrovnávací konstanta, $\beta \in (0,1)$.

Počáteční hodnoty m_0 a b_0 se počítají přímkovou regresí z první poloviny hodnot analyzované řady.

Protože analyzovaná časová řada obsahuje hodnoty roční míry inflace, tudíž se jedná o řadu bez sezónní složky, nebudeme se zabývat Wintersovým exponenciálním vyrovnáváním.

Výpočty související s exponenciálním vyrovnáváním se dnes běžně provádějí, jak už bylo řečeno, pomocí statistického softwaru. Největším problémem je volba vyrovnávacích konstant. Např. v případě Brownova jednoduchého vyrovnávání, kdy uvažujeme dosti hrubé třídění vyrovnávací konstanty α po 0,1 v rozmezí $(0,1)$, představuje výpočet 10 modelů, v případě Holtova vyrovnávání již 100 a v případě Wintersova (tři konstanty α, β, γ)

vyrovnávání již 1000 možných kombinací hodnot vyrovnávacích konstant. Existují statistické programy, které přímo nabízejí optimální hodnoty vyrovnávacích konstant.

Kritéria pro volbu modelu časové řady

V současné době existuje široká škála metod a postupů použitelných při analýze a prognóze časových řad ukazatelů a jejich počet se neustále zvyšuje. Kritéria a přístupy k volbě modelu časové řady je možno rozdělit do dvou skupin:

* **věcně ekonomická kritéria**, která vycházejí z požadavku, aby model byl zvolen v souladu s věcnou analýzou sledovaného ekonomického jevu. Zařazujeme sem znalosti, které má specialista příslušného oboru k dispozici a které často nejsou ani kvantifikovatelné. Jedná se o určité představy o charakteru vývoje příslušného ukazatele, na jehož základě můžeme zvolit vhodnou trendovou křivku: např. přímku v případě předpokladu rovnoměrného růstu nebo poklesu, parabolu při očekávání zpomalení přírůstků. Pokud dochází ke změnám trendu, nepravidelностям a výkyvům v průběhu časové řady, volíme některou z adaptivních metod.

* **statistická kritéria**, která využívají prostředků popisné a matematické statistiky. Tato kritéria umožňují navzájem porovnat modely různého typu - např. vyrovnání řady analytickou trendovou funkcí, adaptivní metodou apod. Statistická kritéria dělíme na kritéria interpolační a extrapolaci v závislosti na tom, zda vybíráme model především pro analýzu nebo hlavně pro extrapolaci.

Interpolační kritéria vycházejí z chování časové řady ve sledovaném období, tj. ze zjištěných dat. Pomocí těchto kritérií volíme model, který nejlépe vystihne průběh a chování řady v minulosti.

Extrapolaci kritéria naproti tomu slouží k výběru modelu, který nejlépe simuluje chování řady do budoucnosti.

Rozlišujeme přitom dvě skupiny kritérií: *souhrnná* kritéria, která popisují celý model jedním číslem a kritéria *strukturní*, popisující pouze určitou vlastnost modelu.

Souhrnná interpolační kritéria využívají metod regresní analýzy. Nejčastěji se za kritérium volí součet čtverců odchylek skutečných hodnot časové řady a teoretických hodnot, tzv. *reziduální součet čtverců* (5).

Za nejlepší model je zpravidla považován ten, který dává nejmenší S_R . Riziko aplikace těchto kritérií spočívá v tom, že při použití trendových křivek polynomů vyšších řádů se S_R snižuje, aniž bychom měli záruku, že daná křivka vystihuje celkovou tendenci řady.

Jiným používaným kritériem je *index determinace* R^2 ,

$$R^2 = \frac{\sum_{t=1}^n (\hat{T}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2}, \quad (14)$$

kde \bar{y} je celkový průměr časové řady.

Za vhodnější považujeme model, který má větší $R^2 \in (0,1)$. I toto kritérium má rovněž své nedostatky, protože s růstem počtu parametrů trendové funkce roste R^2 , což je v rozporu s principem parsimonie (ekonomie parametrů). Tento nedostatek řeší upravená hodnota indexu determinace

$$R_{kor}^2 = 1 - (1 - R^2) \frac{n-1}{n-p}, \quad (15)$$

kde n je počet pozorování a p je počet parametrů (bez absolutního členu) trendové funkce.

Strukturní kritéria popisují dílčí vlastnosti modelů a mají nejčastěji **interpolační charakter**. Jde o kritéria používaná v regresní analýze. Nejdůležitější jsou testy významnosti parametrů a testy náhodnosti reziduí.

Durbin-Watsonův test autokorelace je nejpoužívanější test, jímž ověřujeme nezávislost náhodných poruch (reziduí) v modelu. Používá testovací kritérium

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}. \quad (16)$$

kde reziduum $e_t = y_t - \hat{T}_t$.

Kritérium (16) nabývá hodnot z intervalu (0,4). V případě, že hodnota DW se pohybuje kolem 2, nelze zamítnout hypotézu o nezávislosti náhodných poruch. Blíží-li se hodnota DW 0 nebo 4, jsou rezidua závislá. Pokud test (16) ukáže závislost reziduí, většinou to znamená, že model trendu není vhodný.

Vhodnost modelu pro popis průběhu řady v minulosti ale ještě není zárukou, že tento model bude poskytovat dobré extrapolační předpovědi. Na základě rozsáhlých srovnávacích

studií bylo dokázáno, že jenom 50 - 60 % modelů kvalitních při popisu minulosti dalo také kvalitní krátkodobé prognózy.

Souhrnná extrapolační kritéria umožňují posoudit kvalitu extrapolace daným modelem. Nejčastěji používaný postup je založen na *simulaci*.

Jako míra kvality modelu jsou používány různé koeficienty nesouladu, z nichž nejznámější je *Theilův koeficient nesouladu*

$$T^2 = \frac{\sum_{j=1}^m (y_{n-m+j} - \hat{P}_j)^2}{\sum_{j=1}^m y_{n-m+j}^2}, \quad (17)$$

kde

(n-m) je délka zkrácené časové řady, $j=1,2,\dots,m$,

\hat{P}_j je předpověď na j-období dopředu modelem vypočteným z (n-m) hodnot.

Při srovnání více modelů preferujeme model, který má menší koeficient nesouladu. Pro krátkodobé předpovědi, na 1 až 3 období dopředu, se osvědčil velmi jednoduchý postup, kdy časová řada je zkrácena o poslední (známé) pozorování.

Z (n-1) hodnot vypočítáme „pseudopředpověď“ (\hat{y}_{13}) a porovnáme se známou hodnotou tohoto pozorování (y_{13}) - viz tabulku 2. Model, který poskytl nejlepší pseudopředpověď, použijeme pro extrapolaci.

Tab. 2 Hodnocení statistických modelů

Model trendu	S _R	R ²	DW	y ₁₃	\hat{y}_{13}	Rok 1998
přímka 3,6115 + 1,0126 t	2519	0,0690	1,8086	16,7758	17,2955	17,7879
parabola -13,4776+7,8483t-0,4883 t ²	2041	0,2454	2,2324	6,0341	-3,3233	0,6918
exponenciála 6,8666*1,0632 ^t	2583	0,0453	1,7652	15,2334	15,6893	16,1937
S-křivka 22,1074*exp(-3,9981/t)	2349	0,1318	1,9358	16,2476	16,5861	16,6155
Brownovo jednoduché vyrovnávání, $\alpha = 0,2$	3134	-	-	11,9248	11,2034	11,2398
Holtovo lin. vyrovnávání $\alpha = 0,2 \beta = 0,6$	3792	-	-	16,3362	11,7968	10,5377

Statistických metod, které lze s úspěchem aplikovat v oblasti extrapolace časových řad, existuje velké množství. Přitom početní složitost přestává hrát roli neboť vybavenost

pracovišť počítači a statistickým softwarem se stává samozřejmostí. Volba vhodné metody pro analýzu a extrapolaci však zůstává stále subjektivní záležitostí pracovníka provádějícího extrapolace, neboť sebelepší manuál nemůže pracovníkovi s menší znalostí metod a s menšími zkušenostmi z praktické prognostické činnosti poradit nejvhodnější metodu pro danou analýzu.

Závěr

Rozhodnutí, který z možných modelů vybrat pro analýzu a pro krátkodobou extrapolaci časové řady, je i při existenci velkého množství kritérií vždy značně subjektivní záležitostí.

Volba vhodného modelu je konfrontací výsledků statistických testů a vypočtených charakteristik především s věcnou analýzou časové řady. Volba vhodného modelu se tak stává úlohou vícekritériálního rozhodování, kde váhy jednotlivých kritérií závisejí na znalostech a zkušenostech toho, kdo model tvoří. V období výrazných ekonomických změn je věcná analýza hlavním kritériem při rozhodování mezi různými modely.

Na základě srovnání různých statistických kritérií uvedených v tabulce 2 se jeví jako nejlepší model popisu trendu míry inflace v letech 1985 až 1997 tzv. S-křivka. Tyto křivky, nazvané podle svého typického průběhu, popisují tzv. logistický trend, kdy hodnota příslušného ukazatele nejprve pomalu vzrůstá, poté dojde k jejímu strmějšímu růstu a posléze je tento růst zpomalen. Vhodnost modelu pro popis průběhu řady v minulosti ale ještě není zárukou, že tento model bude poskytovat dobré extrapolační předpovědi. Na základě rozsáhlých srovnávacích studií bylo dokázáno, že jenom 50 - 60 % modelů kvalitních při popisu minulosti dalo také kvalitní krátkodobé prognózy. Realističtější přístup k modelování časových řad poskytují adaptivní metody, které vycházejí z předpokladu, že pro konstrukci extrapolační prognózy budoucího vývoje jsou nejcennější nejnovější pozorování časové řady.

Různé metody tedy poskytují různé extrapolační předpovědi pro tutéž časovou řadu. Proto různí autoři navrhnou různé metody kombinování předpovědí. Nejjednodušší způsob kombinování je výpočet aritmetického průměru předpovědí poskytnutých různými metodami. Z empirické studie [3] vyplynulo, že u krátkých řad se jeví vhodné kombinovat předpovědi z metod exponenciálního vyrovnání s předpověďmi z analytických metod. V našem případě byl vypočten vážený aritmetický průměr z předpovědi S-křivkou (váha = 0,6 vzhledem ke zmíněné 60% úspěšnosti analytických modelů) a předpovědi z Holtova lineárního vyrovnání (váha = 1).

Kombinací předpovědí, tj. výpočtem váženého aritmetického průměru, dostáváme **pro rok 1998 odhad míry inflace 10,3 %**. Odhad vychází z předpokladu, že charakter vývoje ekonomiky v roce 1998 se oproti minulým rokům v podstatě nezmění. Tento předpoklad však podle posledních informací nebude splněn. V ekonomice je absence inflačních tlaků. Kromě restriktivní politiky ČNB hraje velkou roli propad domácí poptávky, klesající reálné mzdy, silná koruna a celosvětově nízké ceny surovin. Meziroční růst spotřebních cen, který v prvním čtvrtletí vystoupil na 13,3%, se v dalších čtvrtletích zpomaloval na 12,7%, resp. na 9,5%. V prosinci se očekává zmírnění celoročního růstu cen na 8%.

Jak tyto prognózy ovlivní deregulace cen energií a nájemného, stagnace cen spotřebního zboží, kurz koruny, měnové restrikce centrální banky atd., ukáží až statistická čísla za rok 1998.

Literatura:

- [1] BLATNÁ, D.: Kritéria výběru vhodného modelu ekonomické časové řady. Statistika č.5. Český statistický úřad, Praha 1994.
- [2] BLATNÁ, D.: Kombinované předpovědi. Statistika č.7. Český statistický úřad, Praha 1994.
- [3] BLATNÁ, D. : K pokusu vytvořit praktickou pomůcku volby metody pro extrapolaci časových řad. Statistika č.6. Český statistický úřad, Praha 1995.
- [4] CIPRA, T.: Analýza časových řad s aplikacemi v ekonomii. SNTL/Alfa, Praha 1986.
- [5] HINDLS, R. - KAŇOKOVÁ, J. - NOVÁK, I.: Metody statistické analýzy pro ekonomy. Management Press, Praha 1997.
- [6] HOLÍSEK, M.: Makroekonomie pro bakalářské studium. Melandrium, Slaný 1998.
- [7] KOZÁK, J. - HINDLS, R. - ARLT, J.: Úvod do analýzy ekonomických časových řad. VŠE, Praha 1994.
- [8] SEGER, J.- HINDLS, R.: Statistické metody v tržním hospodářství. Victoria publishing, Praha 1995.
- [9] STEINDL, Ch.: Máme dobrou náhradu za index spotřebitelských cen? Statistika č.11. Český statistický úřad, Praha 1997.
- [10] Statistická ročenka České republiky 1997

Kinematická analýza

Úvod

Zpracování obrazu může být prováděno za různým účelem (analýza povrchu, rozpoznávání obrazců, určování velikosti atd.). V našem případě se budeme zabývat kinematickou analýzou. Výsledná data (prostorové souřadnice) mohou být použita pro pozdější zpracování v programech zabývajících se designem a animací nebo mohou být podrobena dalším matematickým výpočtům.

Na sportovištích již není neobvyklá přítomnost kamer různých kvalit, různých značek a různého počtu. Diváci, rodiče, trenéři nebo pověření členové realizačního týmu provádějí videozáznam většinou za účelem archivace a mnohdy i pro získání zpětné vazby z natočeného materiálu. S rostoucí oblibou audiovizuální a výpočetní techniky rostou i požadavky sportovců a jejich týmů po biomechanické analýze za účelem vyhodnocení a vylepšení technického provedení pohybu. Vzdělávací instituce (CASRI Praha, tělovýchovné a sportovní fakulty), které jsou úzce napojeny na sport, tento trend zachycují a nabízejí závodníkům různé možnosti. Kinematická 3D analýza přispívá k přesnému a podrobnému posouzení techniky s možností odhalení slabých i silných stránek v technickém provedení, s následným rozbohem poukázat na klíčové faktory ovlivňující konečný výkon.

Zpracování obrazu

Ve srovnání s většinou ostatních metod měření má analýza obrazu tu výhodu, že nemá přímý negativní dopad. To znamená, že stanovení kvantitativních rozměrů prostřednictvím měřicího systému nemá žádný dopad na chování měřeného objektu. A to protože samotné měření není prováděno na konkrétním objektu, ale na jeho obrazu.

Pro nejjednodušší měřicí techniku představuje tento fakt jednu nevýhodu: trojrozměrný objekt je zobrazen ve dvou dimenzích. Tato nevýhoda je akceptovatelná, jestliže máme zájem pouze o dvě dimenze (2D analýza), např. pro určení nejvyššího skoku, rozběhové rychlosti při skoku dalekém nebo odrazový úhel. Při nahrávání těchto pohybů je důležité, aby tyto pohyby byly kompletně popsány v jedné rovině. Abychom se vyhnuli chybám plynoucím z toho, že se určité části těla pohybují mimo rovinu pohybu, kamera by měla být umístěna dostatečně

daleko od této roviny pohybu. Fyzikální rozměry zaznamenané tímto měřením jsou v první řadě kinematografickými rozměry (vzdálenost, čas, rychlost, zrychlení, úhly).

Problémy související s analýzou obrazu

Poté co byl pohyb nahrán, může být provedena analýza záběru. Abychom ji mohli provést, musí být určeny body na těle nebo body, které jsou určitým způsobem důležité pro vykonání pohybu. Použitými body na těle jsou většinou průsečíky kloubních os, nebo jejich středy. Při tomto určování můžeme narazit na tři hlavní zdroje chyb:

- osy kloubů nemohou být jasně definovány
- průsečíky os nelze na záběru jasně rozlišit
- průsečíky jsou skryty za ostatními částmi těla a na záběru nejsou viditelné

Řešení

- pouze precizní znalost anatomie může minimalizovat tuto chybu
- průsečíky lze označit jasně kontrastní barvou
- střed kloubů musí být interpolován, popřípadě odhadnut

Chyby a tolerance chyb

- chyby v určování časového rozpětí mezi jednotlivými snímky záznamu
- chyby v určování pozice měřených bodů
- kumulativní chyby, které nastanou, když k výpočtům použijí nesprávné hodnoty, např. $\text{rychlost} = \text{vzdálenost} / \text{čas}$, přičemž naměřené hodnoty vzdálenosti i času jsou obě nepřesné.

Rozsah těchto chyb může být vyjádřen jako matematická funkce citlivosti použitého filmu, přesnosti snímací metody, přesnosti určení ohniskových bodů při měření, chyb vzniklých při zaznamenávání času atd. Různorodost těchto faktorů ukazuje, jak komplikované mohou tyto výpočty být.

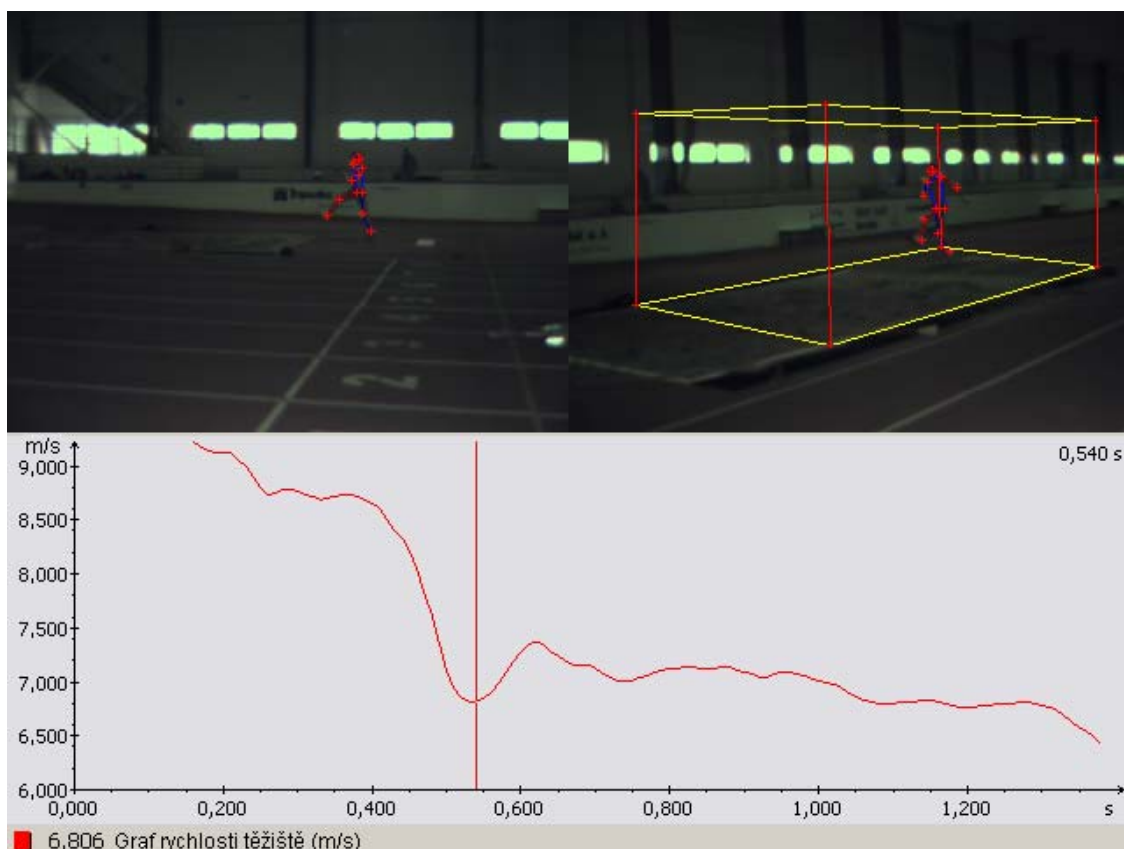
V praxi je dostačující, že tolerance chyb jsou zjištěny s odvoláním na známé vnější hodnoty. Jestliže je například známá hodnota vzdálenosti mezi vrchním hlezenním kloubem a kolenním kloubem, potom musíme dospět ke stejné hodnotě i po sejmutí obrazu a provedení výpočtů (Janura, 2004).

Zobrazení dat

Sledovat lze jednotlivý bod, spojnice bodů a těžiště. Je možné zvýraznit tyto spojnice a sledovat je během pohybu. Například spojnice mezi kyčlí a kolenem může být v průběhu určité fáze pohybu vyobrazena v jiné barvě.

Existují různé typy těžiště pro různé pohybové sekvence. Pro každý model je požadován určitý počet bodů. To znamená, že body specifikace musí být nejprve přiřazeny k bodům daného modelu. Určování těžiště je matematickým odhadem a je založeno na zkušenostech a naměřených hodnotách. Přesné parametry pro výpočet těžiště jsou pro každého člověka rozdílné, takže s použitím jednoho modelu pro různé typy lidí (muži/ženy, dospělí/děti nebo sprinteři/vytrvalci) by se mělo zacházet opatrně. Je možné chybu minimalizovat pomocí software doplňku, jež umožňuje získání parametrů určité osoby na základě individuálních měření (váha, výška, měření hrudního koše, šíře zad, délka nohy atd.)

Následné zobrazení (obr. 1) modelovaných dat v libovolné ose x, y, z 3-rozměrného prostoru, spolu se sledováním jednotlivých charakteristik – vzdálenosti, rychlosti, zrychlení, úhly se sledováním vlastní provedení sportovního výkonu dává do rukou trenéra velmi mocný nástroj na posouzení individuální technické vyspělosti sportovce (Sebera, 2006).



Obr. 1 Rychlost těžiště v průběhu skoku s vizualizací

Výsledek kinematické analýzy

- 3D model pohybu s možností náhledů a podhledů
- individuální biomechanická charakteristika skokana
- možnost srovnání dvou špičkových skokanů, hledání jejich silných a slabých stránek
- možnost duálního porovnání parametrů výkonu, např. před zraněním a po zranění
- hledání silných a slabých stránek vlastního výkonu
- kinogram

Vyhodnocování dat

Nepřeberné množství dat z 3D kinematických analýz je na první pohled pro výzkumníka velmi potěšující, neboť tuší, že má před sebou velmi podrobnou podobu reálného pohybu a předpokládá, že tam najde přesné odpovědi na otázky, které si před vlastním výzkumem položil. Zejména zpočátku, kdy chybí výzkumníkovi zkušenosti s vyhodnocováním, se prvotní nadšení brzy promění ve skepsi, neboť pouhá orientace v množství naměřených dat je nelehká. Natož pak provádět jednotlivé dílčí kroky vedoucí k naplnění cílů výzkumu. Pro představu, pokud snímáme chůzi frekvencí 500 snímků za sekundu a natočíme a vyhodnotíme 4 kroky, kdy každý trvá průměrně 0,55 s, tak máme k dispozici 1100 údajů a snímáme-li 13 bodů (pro vytvoření 3D modelu) na lidském těle, jde pak o 14300 údajů. Experiment se může několikrát opakovat a počet dat se tak násobně zvyšuje.

Modelování časových řad pomocí waveletů

Wavelet transformace (WT) představuje nový matematický prostředek pro analýzu signálů pomocí Wavelet funkcí s aplikací na široké spektrum reálných signálů, které zahrnuje technologické časové řady, biomedicínské signály a obrazy, družicové snímky, ekonomická data i lidskou řeč. V řadě případů je problémem efektivní kódování, komprese, potlačování rušivých složek, modelování a detekce dílčích složek signálu. Mezi důležité výsledky těchto studií patří analýza a konstrukce wavelet funkcí v analytické i rekurentní formě spolu s popisem jejich vlastností. Wavelet transformace přitom představuje alternativní přístup k diskretní Fourierově transformaci (FFT) pro analýzu nestacionárních signálů a detekci

bodů, pro které signál mění vlastnosti. Hlavní výhoda WT spočívá v její schopnosti analýzy signálů s proměnným časově-frekvenčním rozlišením.

Multirozklad – Multiresolution analysis (MR)

Velice efektivním a často používaným způsobem výpočtu Wavelet transformace je užití Mallatovy pyramidální struktury pro určení příslušných koeficientů pro daný vektor. Z hlediska metod zpracování signálů se jedná o užití nízkofrekvenčního filtru pro mezní frekvenci v polovině příslušného frekvenčního pásma pro funkci měřítka (a určení aproximativních složek signálu) a dále o aplikaci vysokofrekvenčního (Wavelet) filtru pro stanovení detailních složek signálu. Tyto analyzující posloupnosti jsou užity v konvoluci s daným signálem a příslušný výsledek je dále podvzorkován faktorem 2. Výsledkem jednoho kroku dekompozice jsou v případě jednorozměrných signálů dvě funkce (Veselý, 1996).

Pro potřeby zpracování nestacionárních signálů je nutné provést detekci změn a nalézt hranice jednotlivých segmentů. Nalezené segmenty je možno samostatně zpracovávat, klasifikovat, provádět predikci apod. V současné době se ukazuje Wavelet analýza jako velice výhodná metoda pro segmentaci. Jak je popsáno výše, je založena na aplikaci nízké a vysokofrekvenčních filtrů na daný signál. Výsledkem jsou dvě nové posloupnosti, které jsou označovány jako aproximační a detailní složka. Tyto názvy vychází z předpokladu, že každý signál má důležité informace uloženy v nízkých frekvencích, které postačí k aproximaci daného signálu, zatímco vysoké frekvence nesou informace pouze zpřesňující, dokreslující, tedy detaily daného signálu.

Rozklad pomocí wavelet analýzy do 1. rozlišovací hladiny je možno použít pouze u testovacích signálů, které neobsahují příliš mnoho frekvenčních složek. U technických, biomedicínckých, elektrotechnických a dalších signálů je potřeba celý postup opakovat. Celý postup odpovídá dekompozici podle Mallatova pyramidálního schématu. Pro dekompozici do druhé úrovně použijeme aproximační složku z úrovně první. Dostaneme tak novou detailní a aproximační složku. To znamená, že po dvou krocích máme signál rozložen do dvou detailních a jedné aproximační složky.

Srovnání s fourierovskou transformací

Při rozkladu časové řady (obr. 5, řada „s“) lze stěží použít standardních statistických postupů, jakými jsou např. regresní analýza, neboť vstupní data nemají funkční předpis. Klasická statistika tady končí a musí přijít teorie zpracování signálu.

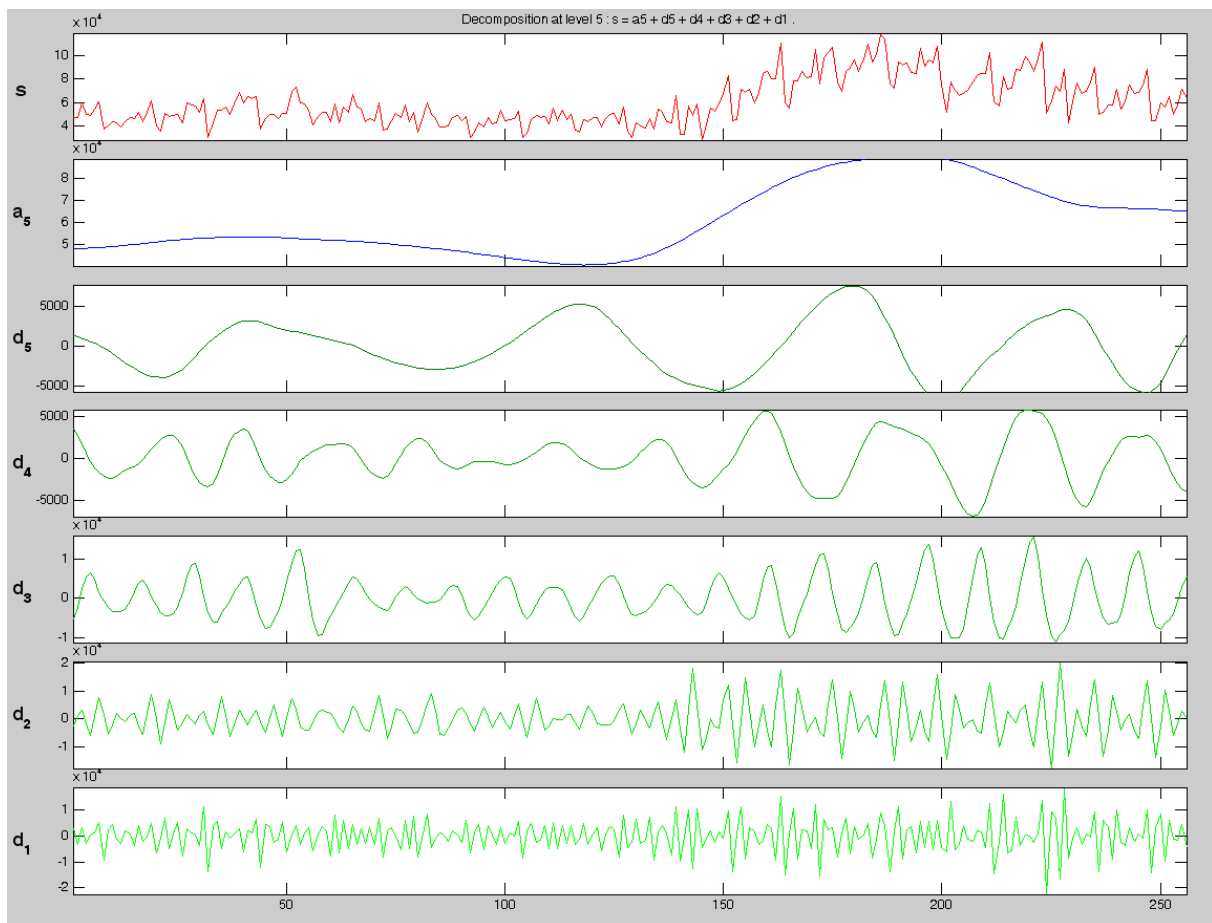
Wavelet transformace se zjednodušeně vysvětlí tak, že procházíme po celém signálu a hledáme shodu mezi mateřskou funkcí (obr. 3) a originálním signálem. Mateřská funkce je deformovatelná ve 2 parametrech, které stanoví shodu (wavelet koeficienty). Obrovskou výhodou oproti FFT je skutečnost, že pokud změněme např. poslední část signálu, pak se nemění předchozí wavelet koeficienty. Na rozdíl od FFT, kde se změna projeví globálně ve všech parametrech. FFT vypovídá pouze o přítomnosti frekvencí v signále, neříká však nic o jejich umístění v čase.

Pomocí wavelet báze Daubechies řádu 8 (jedna z tzv. mateřských funkcí) jsme provedli MR-analýzu na vstupních datech (s, záznam rychlosti plavce), jejímž výsledkem je 6 komponent. Aproximační složka a_1 a 5 detailních složek d_1 až d_5 (obr. 5), kde platí $s = a_5 + d_1 + d_2 + d_3 + d_4 + d_5$. Z původních dat tak získáváme trendovou křivku (a_5), společně s dalšími aproximacemi, které už posléze dovolují „rozumnou“ interpretaci, ať už při srovnávání dvou respondentů nebo při hledání změn u jednoho jedince (např. před a po zranění). Pro výpočty a vizualizace byl použit systém Matlab (<http://www.mathworks.com/products/matlab/index.html>).

Ve prospěch waveletů hovoří především myšlenková jednoduchost blízká zaběhaným fourierovským přístupům. Oproti nim však nabízí větší pružnost danou lokálním charakterem příspěvků jednotlivých waveletových koeficientů ve waveletové řadě. Například při waveletové filtraci tedy postupujeme obdobně jako při fourierovské filtraci, ovšem s tím podstatným rozdílem, že při redukci wavelet koeficientů je vliv shlázení pouze lokální v závislosti na zvolené úrovni a nikoliv globální. Waveletový filtr tedy poskytuje významnou novou kvalitu v tom, že se může adaptovat na lokální charakter dat. Tam kde data vykazují vyšší dynamiku, tj. nesou užitečnou informaci ve vyšších frekvenčních pásmech, můžeme míru redukce vysokofrekvenčních komponent snížit (neboli snížit prahovou hodnotu signifikantnosti) a naopak postupovat v místech, kde je změna jen pozvolná. Tím se odstraňuje hlavní nevýhoda fourierovské filtrace, která v takovém případě má tendenci dynamický úsek přehlédnout. Typickým příkladem může být rozvlnění v okolí nespojitých změn

(Gibbsův jev). Toto je ovšem spíše problém volby správné strategie vyhlazení, než samotného principu waveletové filtrace.

Jako žádná jiná metoda, nejsou ovšem ani wavelety universální všelék pro všechny typy úloh, ale v každém případě významně obohacují repertoár dosud běžně používaných technik.



Obr. 5 Původní data a multirozklad původních dat

Literatura

- Hebák, P. - Hustopecký, J. - Pecáková, I. - Plašil, M. - Průša, M. - Řezanková, H. - Svobodová, A. - Vlach, P. *Vícerozměrné statistické metody 3*. Praha: Informatorium, 2007. ISBN 80-7333-039-3.
- Janura, M. – Zahálka, F. *Kinematická analýza pohybu člověka*. Olomouc: 2004 1. vyd. ISBN 80-244-0930-5
- Sebera, M. *Využití multimediálních prostředků v práci trenéra atletiky*. Brno: FSpS MU. 2006. 1. trenérské třídy atletiky. Závěrečná práce.

- Sebera, M. - Seberová, H. Časové řady a wavelety. Vyškov: VVŠ PV Vyškov, 2002. ISBN 80-7231-090-9, s. 352-356.
- Veselý, V. *Wavelety a jejich použití při filtraci dat*. In ROBUST'1996. Praha 1997, s. 241-272.

Distance-Time Dependencies of Spatial Coordinates for Gait Recognition

Martin Sebera, Jan Sedmidubský*, Martin Zvonař

Masaryk University - Faculty of Sports Studies, *Faculty of Informatics

Gait recognition; MUFIN; kinematic analysis; SIMI

Purpose

This paper presents a system for identifying people by their gait pattern. We take into account different recognizable features which can be obtained by walking with help of the SIMI Motion system.

In recent years we have witnessed a trend where we are increasingly dependent on computers, especially in the form of information storage and processing. The same is true for the biometric data, for automatic recognition of people based on their anatomical characteristics (e.g. face, fingerprint, iris, retina) and behaviour (e.g., signature, gait) [1]. From these options we focus on gait, which is some way of expressing the human uniqueness. What is more, there is considerable evidence that gait is unique in its ability to determine one's identity [2, 3]. We can find several studies in areas such as biomechanics, mathematics and psychology. Walking is ultimately attractive as a biometric identifier, because it does not affect the measurement and recognition of the person under surveillance and can be detected and measured in low resolution.

Gafurov [4] described an overview of biometrics systems of walking and general procedures such as vision-based approach, the floor-based approach and a portable sensor-based approach. Vision-Based Gait Recognition focuses on identifying individuals with different features extracted from video sequences of walking people [5]. Compared to other biometric features such as fingerprints, the vision-based gait recognition has the advantage of being inconspicuous. Movement is probably the only perceivable biometric feature from distance with a low resolution in comparison with systems such as facial recognition or iris. Nixon et al. presented an extensive overview of different identification methods based on human gait [6]. Most current approaches work on the basis of analysis of silhouette images in

a sequence of characters walking man, e.g. trajectories, an estimate of the pace and stride length, relative body parameters or changing angles of trajectory extracted from markers arranged on the human body. Another way is searching of relationships and estimates positions, velocity and acceleration.

In [11] a comprehensive survey of video based gait recognition approaches is presented. Generally, gait recognition approaches can be roughly divided into two categories, i.e., model-based approach and model-free approach. Model-free approach aims to extract statistical features from the whole silhouette to identify individuals while model-based approach aims to model gait explicitly. Model-based approach aims to explicitly model human body or motion according to prior knowledge. Usually, each frame of a walking sequence is fitted to the model and the parameters such as trajectories are measured on the model as gait features for recognition. These methods are easy to be understood, however, they tend to be complex and need high computational cost. Model-free approach does not need the prior knowledge of the gait model. It compactly represents gait motion of the walking human based on the gait appearance without considering the underlying structure.

METHODS

In our research we first describe SIMI system for obtaining the 3D coordinates of the gait. Then we use MUFIN system to recognize the individual gait characteristics.

SIMI motion

Simi Motion is a motion analysis software. Its modular design means that a customized system tailored to each user's requirements can be quickly and easily produced. Typical modules which are available are 2D or 3D kinematics (image based motion analysis), inverse dynamics and support for several DV or high-speed video cameras and for EMG, force plates, pressure distribution measuring equipment and other devices. Analog data acquired simultaneously can be automatically synchronized with the video recording. Simi Motion can be used in a wide variety of areas. In science and research it is especially important to us to provide extremely flexible and technically highly sophisticated systems. Simi Motion has been developed for use in biomechanics, sports science, neuroscience, veterinary medicine, materials science, space research and many other fields: in fact, everywhere where motion is an important factor. The 2D or 3D data of selected points (paths, velocities and accelerations) can be calculated from the recorded videos and illustrated in

diagrams synchronized with the video. In analyzing human movement, inverse kinematic calculations based on certain anatomical markers can also be performed to provide the following data: segment centers, joint centers, segment rotations, joint rotations, constraining forces, resistive torques. A large number of mathematical functions such as filtering (moving average, low-pass, high-pass), angle functions (3-point angle, angle to horizontal/vertical etc.), frequency analysis and statistical calculations offer a huge potential for scientific research. All the data can be used for further calculations or exported to other programs.

MUFIN Architecture

From a general point of view, the search problem has three dimensions shown in Figure 1: (1) data and query types, (2) index structures and search algorithms, and (3) infrastructure to run the system on. The MUFIN system [8][9] offers a complex solution that is highly flexible in all three aspects.

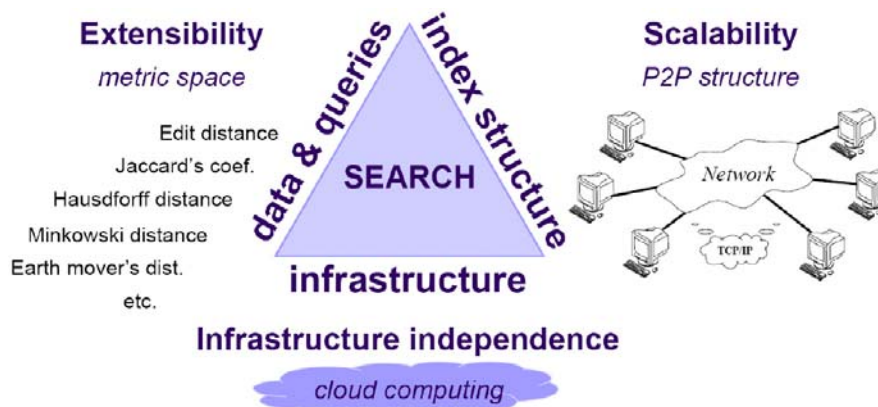


Figure 1: The three aspects of the search problem in the MUFIN system.

The data processed in MUFIN are modelled using the *metric space* approach. It means that the data can be practically anything that has a digital representation if a method to measure their similarity exists. More formally, the mathematical metric space is a pair (D, d) , where D refers to a domain of data objects and d is a function able to compute the distance between any pair of objects from D . It is assumed that the smaller the distance, the closer or more similar the objects are and the zero distance means that the objects are identical. For any three distinct objects $x, y, z \in D$, the distance must satisfy properties of reflexivity, $d(x, x) = 0$, strict positivity, $d(x, y) > 0$ *Chyba! Záložka není definována.*, symmetry, $d(x, y) = d(y, x)$ *Chyba! Záložka není definována.*, and triangle inequality, $d(x, y) \leq d(x, z) + d(z, y)$. There

are many functions satisfying such properties, for example, the commonly used Euclidean distance on 2D or 3D coordinates is a metric function. It is also possible to use a non-metric function – search is not so efficient since all data must be compared to a query to guarantee the precise answer. Data can be searched by a number of similarity queries, including *range* and *k-nearest neighbor* queries. From that point of view, the MUFIN system is highly *extensible* multi-purpose engine that can supply efficient similarity search in various areas.

The MUFIN system achieves the search *scalability* by adopting paradigms of structured peer-to-peer networks: (1) dynamically distributing workload on independent peers for potential parallel query execution, (2) avoiding bottlenecks formed by a single entry point or centralized directories – typical for traditional client-server architectures, and (3) improving fault tolerance by replication of data and adopting multiple search strategies. It is possible by using the concept of virtual processing units (peers) to which the MUFIN system maps the respective parts of the search engine. Due to this open architecture, the system can adapt to virtually any amount of data and also the throughput of the system – the number of simultaneous queries that can be solved without slowing down – is increasing, since queries can be posed from any peer.

Even though the ideas of the system design come from structured peer-to-peer networks, MUFIN is perfectly suitable to run in more controlled environments like computer clusters, GRIDs, or even multi-CPU servers. Actually, implementing a MUFIN searching service on computing clouds (e.g., Amazon EC2) yields a cost-effective fully scalable solution with guaranteed availability. The system hardware abstraction layer allows MUFIN to map its peers onto different hardware architectures. This allows tuning the performance of the system dynamically – if a higher throughput or faster query response is required, the system can utilize additional hardware without changes in the underlying indices.

The implementation of MUFIN builds on a framework, called the Metric Similarity Search Implementation Framework (MESSIF) [1], which is designed to ease the task of building metric-based indexing techniques. Basically, the framework is a collection of modules. The metric space module encapsulates support for metric data objects, the operations module offers a framework for data querying and manipulation methods, and the storage module provides interfaces for creating and maintaining various data storages. The communication module allows exchanging information via the computer network by means of sending and receiving messages. Modules for centralized and distributed index structures encapsulate implemented indexing techniques providing a unified access to searching and

data manipulation regardless of the particular implementation details. Finally, the statistical module allows gathering performance characteristics of all the other modules.

Data Representation

We represent human motion by a structural model of the human body. This model is used for the extraction of gait variables in the form of planar curves derived from graphs indicating the distance of model's components in dependence on time. A collection of these gait variables forms a gait pattern which helps us to distinguish walking persons. Every person is assigned with their gait pattern that can be compared with patterns of other persons by a specific distance function.

The gait pattern is derived from the curves of distance-time dependency graphs for points on the human body that correspond to significant anatomical landmarks. For our human model, we suggest a 3-parameter structure $(P0, P1, P2)$, where Pi represents a 3D coordinate of head ($P0$), left wrist ($P1$), and right hip ($P2$) captured at given time. The gait pattern $G=(C1, C2)$ of each person is then formed by two discrete curves $C1$ and $C2$. The specific value $C1[t]$ represents the Euclidean distance between points $P0$ and $P1$ captured at time t . So the curve $C1$ illustrates how the distance between head and left wrist changes in time as the person walks. The curve $C2$ represents the change of distance between points $P0$ and $P2$.

Two gait patterns G_i and G_j are compared by a specific non-metric function d . This distance function is defined as the following.

$$d(G_i, G_j) = \frac{DTW(G_i.C1, G_j.C1) + DTW(G_i.C2, G_j.C2)}{2}$$

The DTW function stands for the Dynamic Time Warping approach that measures similarity between two discrete curves [10]. This approach does not require the same length of input curves. In other words, the proposed function for comparing two gait patterns is based on similarity of walking styles (curves $C1$) and similarity of physical characteristics (curves $C2$).

Results

We extracted gait patterns from 15 recorded walks that were acquired from 7 different persons. All 15 extracted gait patterns were indexed by the MUFIN system. Individual

patterns started at a different phase (e.g., the pose where the foot span is maximized or minimized) and had a various length. The length of each walk was about 2–3 seconds, which corresponded to 200–350 time moments recorded. In this way, the DTW function compared curves having a variable dimension between 200 and 350.

We created a specialized web user interface for querying gait patterns. The screenshot of this web interface is depicted in Figure 2. We evaluated effectiveness of the proposed recognition function by evaluating similarity queries on each gait pattern recorded. In this way, the MUFIN system evaluated 15 queries. For each query, all the gait patterns were sorted according to their similarity to the query pattern. If the most similar pattern corresponded to the query person, search was successful. The recognition rate over all 15 queries achieved the level of 47% (i.e., 7 queries out of 15 were successful). Such a quite low recognition rate was primarily caused by the following factors: (1) inaccuracy of input 3D coordinates acquired by Simi Motion software, (2) a various length of extracted gait patterns, and (3) a different starting phase of gait patterns. Moreover, effectiveness could be further improved by representing a gait pattern by more than two curves.

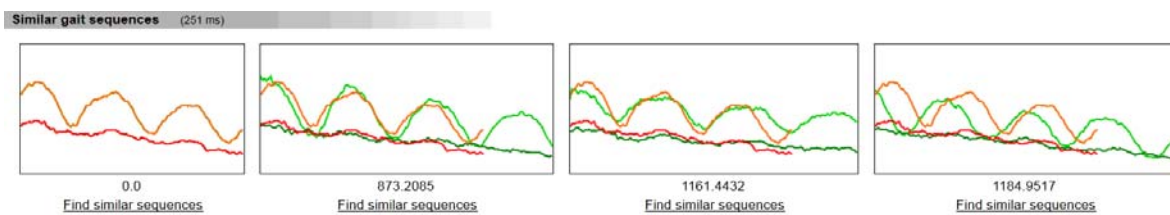


Figure 2: Web user interface demonstrating search for the most similar gait patterns. The left gait pattern (forentik_3_.txt) represents the query and the remaining three ones represent the most similar patterns found (forentik_2_.txt is the most similar pattern to the query pattern). The query pattern is represented by the red color shade (orange – curve C_1 , red – curve C_2) and the specific found pattern is denoted by the green color shade (light green – curve C_1 , dark green – curve C_2).

CONCLUSIONS

We introduced a method of identifying people by their typical aspects of walking. We created a data representation for the system MUFIN. We have also developed a feature metrics for assessing the similarity of two data representations. To improve the gait recognition algorithm it will be necessary to take further steps. In particular, these include:

- To enhance accuracy of extraction of gait patterns.
- To normalize extracted gait patterns in terms of the same length and starting phase.
- To represent a gait pattern by more than two curves.

- To represent a gait pattern by other specification not only from distance-time dependency graphs but also other – velocity, acceleration or pressure distribution measuring.
- To measure the larger number of people in laboratory conditions and also in real situations.

References

- [1] Jain, A.K.; Pankanti, S.; Prabhakar, S.; Hong, L.; Ross, A. Biometrics: A grand challenge. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)*, Cambridge, UK, 23–26 August 2004; Volume 2, pp. 935-942.
- [2] Kale, A.A.; Sundaresan, A.; Rajagopalan, A.N.; Cuntoor, N.P.; Roy-Chowdhury, A.K.; Kruger, V.; Chellappa, R. Identification of humans using gait. *IEEE Trans. In. Image Processing 2004*, 13, 1163-1173.
- [3] Nixon, M.S.; Carter, J.N. Advances in automatic gait recognition. In *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (AFGR 2004)*, Seoul, Korea, 17–19 May 2004; pp. 139-144.
- [4] Gafurov, D. A survey of biometric gait recognition: Approaches, security and challenges. In *Proceedings of the Annual Norwegian Computer Science Conference*, Oslo, Norway, 19–21 November 2007.
- [5] Jaeseok Yun. User Identification Using Gait Patterns on UbiFloorII. In *Sensors 2011*, 11, 2611-2639; ISSN 1424-8220.
- [6] Nixon, M.S.; Tan, T.; Chellappa, R. *Human Identification Based on Gait*, International Series on Biometrics; Springer: Berlin, Germany, 2006.
- [7] Michal Batko, David Novak, and Pavel Zezula. MESSIF: Metric Similarity Search Implementation Framework. In *Digital Libraries: Research and Development, Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer-Verlag. vol. 4877, 10 pages, 2007. ISBN 978-3-540-77087-9.
- [8] Michal Batko, Fabrizio Falchi, Claudio Lucchese, David Novak, Raffaele Perego, Fausto Rabitti, Jan Sedmidubsky, and Pavel Zezula. Building a Web-scale Image Similarity Search System. *Multimedia Tools and Applications*, Springer Netherlands, USA. ISSN 1380-7501, 2010, vol. 47, no. 3, pp. 599–629.
- [9] Michal Batko, Vlastislav Dohnal, David Novak, and Jan Sedmidubsky. MUFIN: A Multi-Feature Indexing Network. In *2nd International Workshop on Similarity Search and Applications*. Los Alamitos, CA 90720-1314: IEEE Computer Society, 2009. ISBN 978-0-7695-3765-8, pp. 158–159. 28.8.2009, Prague.

- [10] C. S. Myers and L. R. Rabiner. A comparative study of several dynamic time-warping algorithms for connected word recognition. *The Bell System Technical Journal*, vol. 60, no. 7, pp. 1389–1409, 1981.
- [11] Ling-Feng Liu, Wei Jia and Yi-Hai Zhu. Survey of Gait Recognition. In *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence Lecture Notes in Computer Science*, 2009, Volume 5755/2009, 652-659, DOI: 10.1007/978-3-642-04020-7_70.

Economical time series prediction via wavelets (wavelet games)

Abstract: We analyze the exchange rate of US dollar to Czech crown and verify how wavelets can help in identifying, estimating and predicting this economical time series. We aim at exploiting how wavelets can discriminate among information at different resolution levels. This paper is organized as follows: In section 2 we present a brief review of wavelets, in section 3 we deal with wavelet-based de-noisation techniques and in section 4 we introduce our experiments with wavelets.

1 INTRODUCTION

In this paper we propose to apply wavelet theory in forecasting economic time series. The method consists of decomposing the series into its long-term trend and its detailed components of the discrete wavelet transform of the series. Each component is then extended to provide a forecast of the series. The method is applied to the data set of the exchange rate of US dollar to Czech crown and the results are compared with the Box-Jenkins forecast of the series.

When the wavelet transform has been carried out, the characteristics of the different transients exist in the coefficients. The statement of the problem then is how to use the characteristics contained in the coefficients so as easily categorize different faults.

2 MULTIREOLUTION WAVELET TRANSFORM

The wavelet transform is one of the latest mathematical tools in the field of signal processing. Its development can be traced to the Fourier transform. The Fourier transform has long been successfully used in the analysis of stationary signals, however, in the case of nonstationary signals, the Fourier transform has difficulties in presenting the time-localization information. Nonstationary signals can be represented as a linear combination of decompositions based on wavelets. The wavelet transform can be used because of its effectiveness in representing transient signals due to its good localization properties in both the time and frequency domains. This makes it possible to simultaneously determine sharp transitions in the spectrum of the signal and in the localization of their occurrence.

A function can be represented with an appropriate basis function. The complex exponential (sines and cosines) functions are the basis functions for the Fourier transform. In the wavelet transform, a set of wavelet functions is chosen as the basis. The set of basis functions is obtained by dilating and shifting a prototype function called the mother wavelet.

One particular useful method to represent wavelets is by using different scales or resolutions. The central part of this is multiresolution analysis.

The wavelet transform [Daubechies, 1993] can be thought of as the inner product of the original signal with the basis function $\Psi(t)$:

$$W_{\Psi} f(a,b) = \int f(t) \cdot \Psi^*(t) dt = \langle f, \Psi_{a,b} \rangle,$$

where

$$\Psi_{a,b} = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (1)$$

In discrete cases, the selection of $a = 2^m$, $b = k 2^m$ ($k, m \in Z$) results in the dyadic wavelet transform.

The definition of the wavelet transform shows that the wavelet analysis is a measure of similarity between the basis functions (wavelets) and the original function. The calculated coefficients indicate how close the function is to the daughter wavelet at the particular scale. If the function has a major contribution at a particular scale, the coefficient computed at this point in the time-scale plane will be a relatively large number.

An invertible transform can be obtained by a simple dyadic translation and dilation scheme, which is based on functions $\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k)$, where $j, k \in Z$. That is, there exist functions \square such that $\{\square_{j,k}; j, k \in Z\}$ forms a complete orthonormal basis of $L^2(R)$, which is the space of the square integrable functions.

For a function $f \in L^2(R)$, there is a wavelet expansion

$$f = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \langle f, \Psi_{j,k} \rangle \Psi_{j,k} \quad (2)$$

where $\langle f, \Psi_{j,k} \rangle = \int f(t) \Psi_{j,k}^*(t) dt$. That is, at each scale J there is a projection of f on some linear space. This space can be alternatively spanned by translates of an accordingly dilated and shifted version of a so-called scaling function \square . Let $\square_{1,k}(t) = 2^{J/2} \square(2^J t - k)$. Now the wavelet expansion of the function f is

$$f = \sum_k \langle f, \Phi_{J,k} \rangle \Phi_{J,k} + \sum_{j=J}^{\infty} \sum_k \langle f, \Psi_{j,k} \rangle \Psi_{j,k} \quad (3)$$

As stated above, the collection of translated scaling function from a fixed resolution scale $\{\Phi_{j,k} ; k \in \mathbb{Z}\}$ forms a linear space, say V_j . The spaces are nested, $\dots \subseteq V_j \subseteq V_{j+1} \subseteq \dots$, in $L^2(\mathbb{R})$. The collections of all these spaces form a *multiresolution* analysis.

Multiresolution analysis is a technique to determine the general wavelet transform. It allows the decomposition of signals into various resolution levels. The data with coarse resolution contain information about lower frequency components and retain the main features of the original signal. The data with finer resolution retain information about the high frequency components.

From equation (3), it is clear that a certain space V_j , can be separated into two component spaces. One is called its subspace V_{j+1} , which has the same main features as V_j . The other one, W_{j+1} , is just the difference of these two spaces:

$$V_j = V_{j+1} + W_{j+1} \quad (4)$$

The basis for the space V_{j+1} will be the collection of translated scaling functions from the fixed resolution level $J+1$, and the basis for the space W_{j+1} is the collection of translated wavelet functions from the particular level J .

The equation (3) and (4) demonstrate that an arbitrary signal can be represented as an approximation part and a detailed part. This process is iterated, with successive approximations being decomposed in turn, so that one signal is broken into many lower resolution components. The original signal can be reconstructed from the sum of the final approximation components and the detail components of all levels. This process is presented in Figure 2. The cAs and cDs are considered as the output of bi-channels of a conjugate quadrature filter. cA is from a low-pass filter and cD is from a high-pass filter. Each stage of the process corresponds to a different scale.

Equation (5) is the fundamental equation for the multiresolution analysis [Daubechies, 1993]:

$$\Phi(t) = 2 \sum_n c_n \Phi(2t - n), \quad (5)$$

from which the mother wavelet is defined as

$$\Psi(t) = 2 \sum_n d_n \Phi(2t - n), \quad (6)$$

where d_n and c_n are quadratic coefficients.

3 THRESHOLDING

Many data operations can now be done by processing the corresponding wavelet coefficients. For instance, one can do data smoothing by thresholding the wavelet coefficients and then returning the thresholded code to the “time domain”. The wavelet coefficients correspond to details. When details are small, they might be omitted without substantially affecting the „general picture“ Thus the idea of thresholding wavelet coefficients is a way of cleaning out „unimportant“ details considered being a noise.

➤ Hard thresholding

The policy for thresholding is keep or kill. The absolute values of all wavelet coefficients are compared to a fixed threshold λ . If the magnitude of the coefficient is less than λ , the coefficient is replaced by zero:

$$d_{jk}^{hard} = \begin{cases} 0, & d_{jk} < \lambda \\ d_{jk}, & d_{jk} > \lambda \end{cases}$$

Hard thresholding is used when one is interested in the shortest possible wavelet code. Long sequences of zeroes that are usually obtained in thresholded wavelet decomposition vector are coded in an efficient way.

➤ Soft thresholding

Soft thresholding shrinks all the coefficients towards to origin. The formula is

$$d_{jk}^{soft} = \text{sign}(d_{jk})(|d_{jk}| - \lambda)$$

➤ Quantile thresholding

Let the rule for thresholding be given as

$$d_{jk}^{quant} = \begin{cases} 0, & d_{jk} < p \\ d_{jk}, & d_{jk} > p \end{cases}$$

where p is a p -quantile of the set of all wavelets coefficients. For example, we might want to replace 30 % of the smallest wavelet coefficients by zero.

➤ Universal thresholding

Donoho and Johnstone (1992) propose the universal threshold $\lambda = \sigma \sqrt{2 \log(n)} / \sqrt{n}$ on data set where n is the sample size, σ is the scale of the noise on a standard deviation scale. Universal thresholding can be hard or soft thresholding with the above defined λ as threshold.

4 OUR EXPERIMENT

Our time series has 2650 items. The data set is listed in Appendix 1. Discrete wavelet transform is applied to the first 2640 (the remaining ten items are used for final comparison) to get the empirical wavelet smooth and detail coefficients. The original series is sum of these components. The wavelet coefficients are shrunken by thresholding (soft, hard, quantile). The ARIMA model and prediction is created in every resolution level. The prediction of the original time series is sum of the predictions of the particular components. The best method we can discover by comparing prediction with original last ten items via value RMSE (root squared mean error).

The results are shown in Table 1. First column contains information about wavelet function and about level of the multiresolution analysis (MRA). Individual thresholding methods (no thresholding, soft and hard thresholding, 30-quantile and 70-quantile thresholding) are in the first row. The rest of table contains RMSE values. To calculate the RMS (root mean squared) error the individual errors are squared, added together, divided by the number of individual errors, and then square rooted. RMSE gives a single number, which summarizes the overall error.

	thresholding →				
	no	soft	hard	30-quantile	70-quantile
Daubechies: 6, MRA: 1	0,40088	0,40117	0,40418	0,335429	0,293880
Daubechies: 6, MRA: 2	0,42441	0,30096	0,42444	0,274477	0,190078
Symmlets: 16, MRA: 1	0,39171	0,39167	0,39171	0,394008	0,190806
Symmlets: 16, MRA: 2	0,41699	0,30066	0,41739	0,303546	0,192610

Table 1 RMSE values

5 CONCLUSION

How to obtain the best prediction?

What kind of a *wavelet function* to choose: There is not a large difference among wavelets.

What kind of a *threshold rule* to use to obtain acceptable results: 70-quantile thresholding yields the absolutely best result. This has nice “psychological” effect. We replace 70 % of the smallest wavelet coefficients by zero and remaining 30 % coefficients have sufficient and the sharpest information about the whole series. Figure 3 shows last 26 values of the original time series with the best prediction (Daub: 6, MRA: 2, 70-quantile)

Prediction of the original time series via Box-Jenkins without wavelet multiresolution decomposition gives RMS error 0,309063. The improvement is 62,59 % $(0,309063 / 0,190078) - 1$ on confrontation with the best prediction via wavelets.

Using wavelets gives better results with prediction economical time series in comparison with Box-Jenkins ARIMA model. This end however cannot be generalized pursuant to one's example. Another study would create conditions or criteria, for which this method can be use.

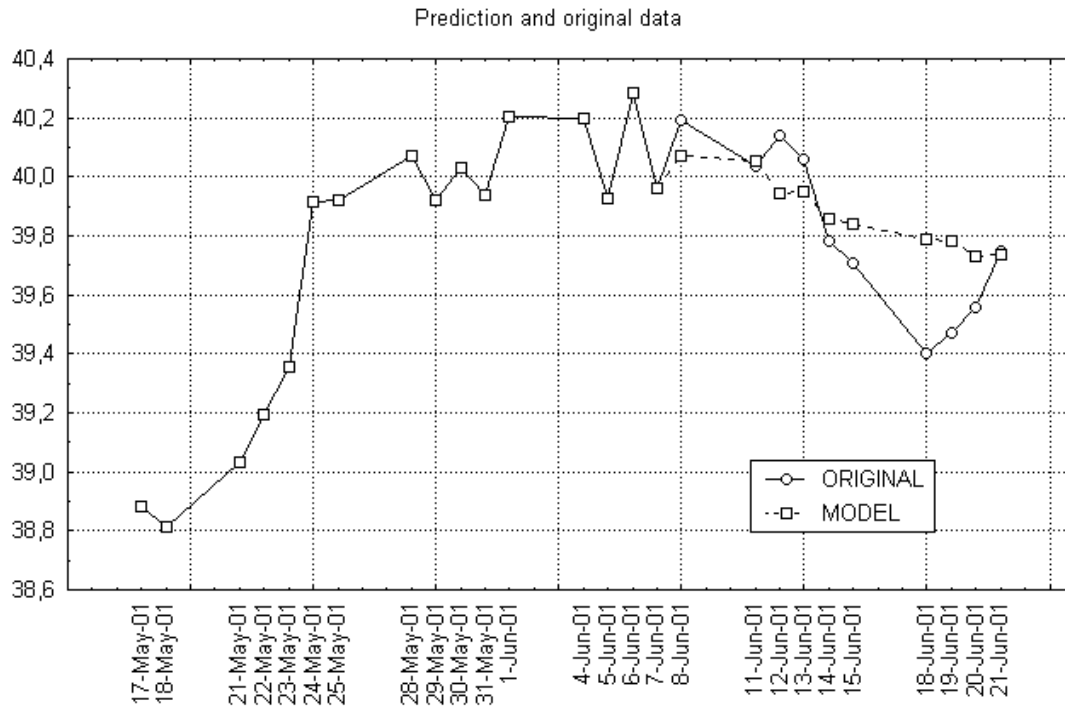


Figure 3 Original time series with the best prediction

REFERENCES

- [1] Daubechies, I. *Ten Lectures on Wavelets*, Kluwer Academic Publishers: 1993.
- [2] Donoho, D. – Johnstone, I. *Adapting to unknown smoothness via wavelet shrinkage*. Technical report 425, Department of Statistics, Stanford University, 1993.
- [3] Box, G.E.P. and Jenkins, G.M. *Time Series Analysis: Forecasting and Control*, rev. ed., San Francisco: Holden Day. 1976.
- [4] Liu, J. – Pillay, P. – Ribeiro, P. Wavelet Analysis of Power Systems Transients Using Scalograms and Multiresolution. In. *Analysis Electric Machines and Power Systems*, 27: 1331–1341, 1999. Taylor & Francis, Inc. 0731-356X

Modelování a prognózování časových řad pomocí waveletů

1 Úvod

Teorie „waveletů“ (vlnek) je relativně novou oblastí aplikované matematiky s počátky sahajícími do let 1982-84. Prostředky Wavelet transformace se ve svých principech opírají o práce Josepha Fouriera, který v 19. století položil základy frekvenční analýzy. Tyto principy byly zásadně modifikovány zejména v posledních 20 letech francouzskými matematiky Y. Meyerem, S. Malattem a I. Daubechies.

Teoretický základ Wavelet transformace byl studován v mnoha pracích a knihách, které vesměs předpokládají u čtenáře dosti vysokou úroveň matematických znalostí, a to především z oblasti funkcionální a klasické Fourierovy analýzy. Případným zájemcům možno doporučit na úvod studia teorie waveletů článek doc. Veselého „*Wavelety a jejich použití při filtraci dat*“. In ROBUST'1996. Praha 1997, s. 241-272, který podává stručný matematický úvod do teorie waveletů a waveletové filtrace .

Mezi důležité výsledky těchto studií patří analýza a konstrukce wavelet funkcí v analytické i rekurentní formě spolu s popisem jejich vlastností. Wavelet transformace přitom představuje alternativní přístup k diskrétní Fourierově transformaci pro analýzu nestacionárních signálů a detekci bodů, pro které signál mění vlastnosti. Hlavní výhoda WT spočívá v její schopnosti analýzy signálů s proměnným časově-frekvenčním rozlišením.

Mezi aplikace Wavelet transformace patří detekce složek signálu a dekompozice signálu, komprese dat a potlačování šumu. Problémy, které jsou úzce spjaty s tímto tématem, zahrnují klasifikaci a predikci signálu, statistické zpracování časových řad, analýzu geofyzikálních signálů a analýzu obrazů.

2 MR - analýza

Velice efektivním a často používaným způsobem výpočtu Wavelet transformace je užití Mallatovy pyramidální struktury pro určení příslušných koeficientů pro daný vektor. Z hlediska metod zpracování signálů se jedná o užití nízkofrekvenčního filtru pro mezní frekvenci v polovině příslušného frekvenčního pásma pro funkci měřítka (a určení aproximativních složek signálu) a dále o aplikaci vysokofrekvenčního (Wavelet) filtru pro stanovení detailních složek signálu. Tyto analyzující posloupnosti jsou užity v konvoluci s daným signálem a příslušný výsledek je dále podvzorkován faktorem 2. Výsledkem jednoho kroku dekompozice jsou v případě jednorozměrných signálů dvě funkce.

Pro potřeby zpracování nestacionárních signálů je nutné provést detekci změn a nalézt hranice jednotlivých segmentů. Nalezené segmenty je možno samostatně zpracovávat, klasifikovat, provádět predikci apod. V současné době se ukazuje Wavelet analýza jako velice výhodná metoda pro segmentaci. Jak je popsáno výše, je založena na aplikaci nízkofrekvenčních a vysokofrekvenčních filtrů na daný signál. Výsledkem jsou dvě nové posloupnosti, které jsou označovány jako aproximační a detailní složka. Tyto názvy vychází z předpokladu, že každý signál má důležité informace uloženy v nízkých frekvencích, které postačí k aproximaci daného signálu, zatímco vysoké frekvence nesou informace pouze zpřesňující, dokreslující, tedy detaily daného signálu.

Rozklad pomocí wavelet analýzy do 1. rozlišovací hladiny je možno použít pouze u testovacích signálů, které neobsahují příliš mnoho frekvenčních složek. U technických, biomedicínckých, elektrotechnických a dalších signálů je potřeba celý postup opakovat. Celý postup odpovídá dekompozici podle Mallatova pyramidálního schématu. Pro dekompozici do

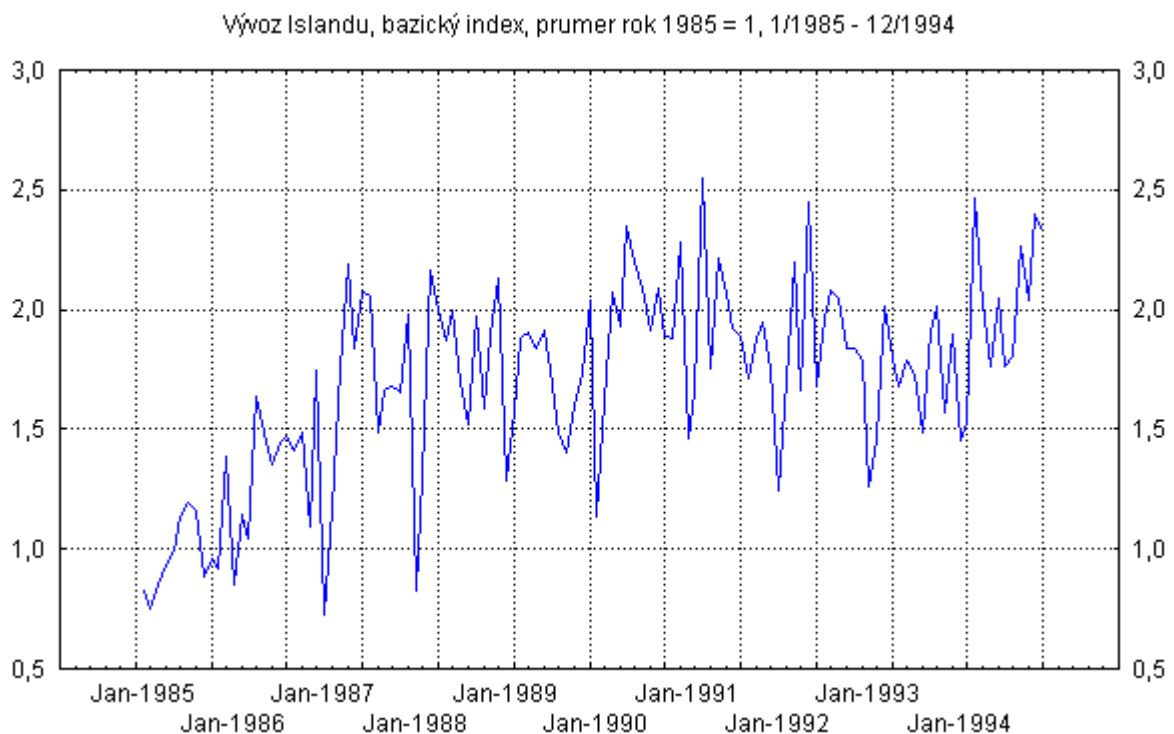
druhé úrovně použijeme aproximační složku z úrovně první. Dostaneme tak novou detailní a aproximační složku. To znamená, že po dvou krocích máme signál rozložen do dvou detailních a jedné aproximační složky.

3 Příklad - měsíční časová řada vývozu Islandu

Máme k dispozici 120 údajů měsíční časové řady vývozu Islandu, kterou máme k dispozici ve formě bazických indexů od ledna roku 1985 do prosince 1994 (základ je průměr roku 1985).

Průběh řady je zachycen na obr. 1. Z obrázku můžeme usuzovat, že časová řada může být nestacionární a obsahovat sezónní složku. Prvních 108 hodnot bude použito k výstavbě modelu, zbylých 12 k posouzení předpovědí.

Obr. 1 Původní časová řada vývozu Islandu

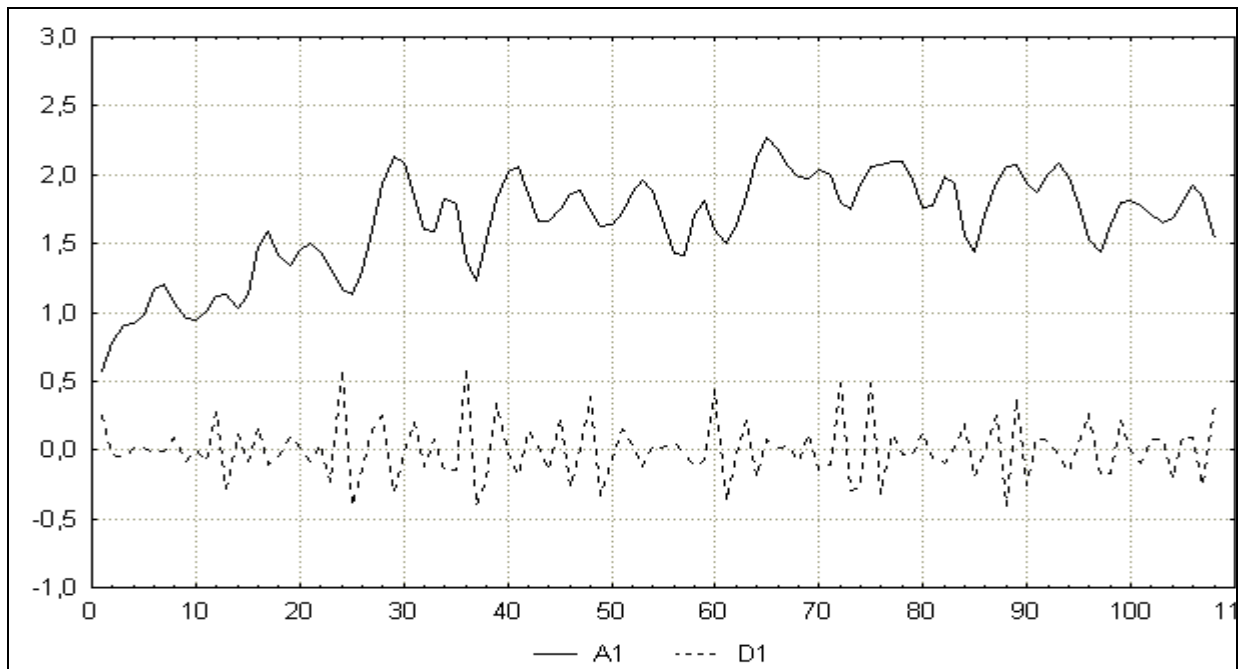


Na základě tohoto ARIMA modelu jsme provedli předpověď 12 hodnot. Srovnáním originálních 12 dat provedeme výpočet hodnoty RMSE (root mean square error). V tomto případě je hodnota RMSE rovna 0,381427. Tuto budeme posléze porovnávat s hodnotou „wavelet“ RMSE modelu.

Pomocí wavelet báze Daubechies řádu 8 jsme provedli MR-analýzu, jejímž výsledkem jsou dvě komponenty. Aproximační složka A1 a detailní složka D1 – obr. 2.

Na těchto dvou složkách jsem provedli Boxovu-Jenkinsonovu metodu modelování časových řad. U zjištěných ARIMA modelu aproximační i detailní složky znovu provedeme predikci 12 hodnot a srovnání s originálními 12 daty. Provedeme výpočet hodnoty RMSE (root mean square error). V tomto případě je hodnota RMSE rovna 0,364865.

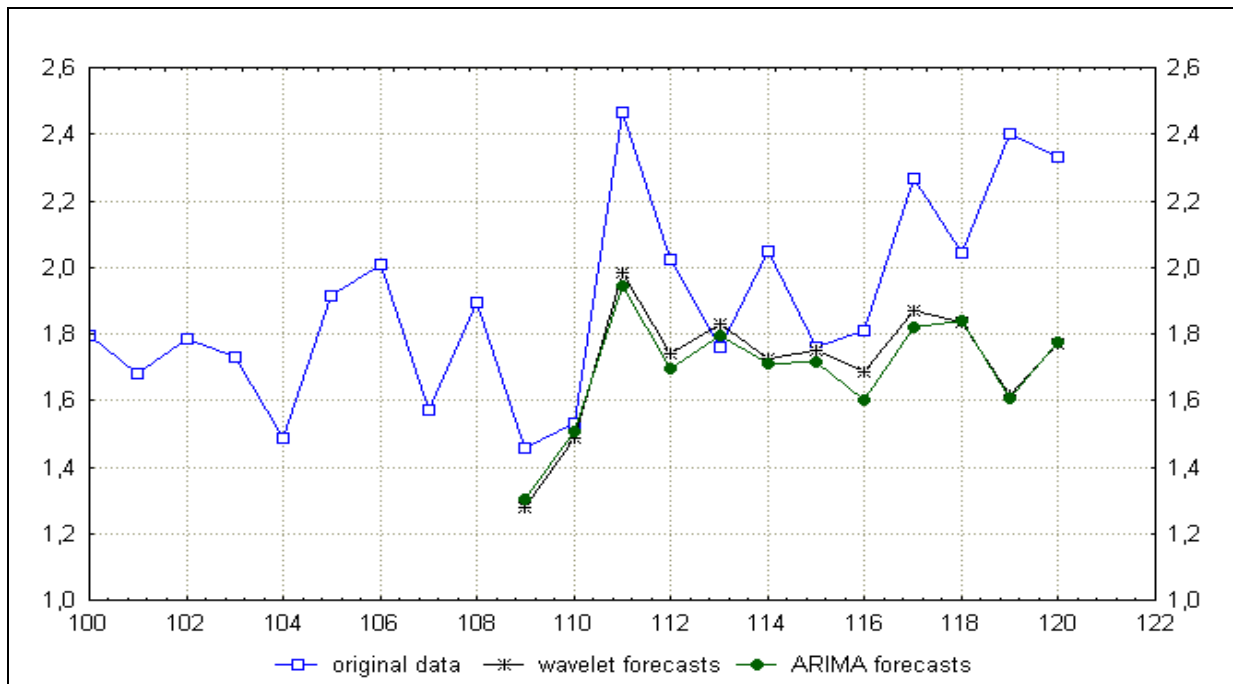
Obr. 2 Detailní a aproximační složka D1, A1



4 Diskuse

Srovnáním hodnot RMSE (root mean squared error) dvou metod bylo zjištěno, že předpověď založená na wavelet teorie přináší zlepšení 4,5% ($0,381427 / 0,364865$) - 1. Porovnáme-li tedy výsledky standardní techniky ARIMA a metody wavelet, lepší výsledky poskytuje wavelet metodologie. Další studie by měly vytvořit podmínky či zjistit kritéria, pro které je vhodné tuto metodou užívat. Určitě by však neměla tato možnost modelování časových řad přehlížena.

Obr. 3 Srovnání původních hodnot, „wavelet“ předpovědí a ARIMA předpovědí



5 Závěr

Wavelet transformace představuje moderní matematický aparát s rozsáhlými aplikacemi v řadě oborů. Význam využití Wavelet funkcí spočívá přitom zejména v možnosti jejich výběru podle chování dané časové řady a v možnostech různého způsobu dekompozice, úpravy a rekonstrukce signálu. Z těchto důvodů je tato tematika předmětem širokého zájmu matematiků i inženýrů v oblasti teoretického popisu Wavelet funkcí, jejich vlastností a dále v oblasti jejich využití.

Ve prospěch waveletů hovoří především myšlenková jednoduchost blízká zaběhaným fourierovským přístupům. Oproti nim však nabízí větší pružnost danou lokálním charakterem příspěvků jednotlivých waveletových koeficientů ve waveletové řadě. Například při waveletové filtraci tedy postupujeme obdobně jako při fourierovské filtraci, ovšem s tím podstatným rozdílem, že při redukci wavelet koeficientů je vliv shlazení pouze lokální v závislosti na zvolené úrovni a nikoliv globální. Waveletový filtr tedy poskytuje významnou novou kvalitu v tom, že se může adaptovat na lokální charakter dat. Tam kde data vykazují vyšší dynamiku, tj. nesou užitečnou informaci ve vyšších frekvenčních pásmech, můžeme míru redukce vysokofrekvenčních komponent snížit (neboli snížit prahovou hodnotu signifikantnosti) a naopak postupovat v místech, kde je změna jen pozvolná. Tím se odstraňuje hlavní nevýhoda fourierovské filtrace, která v takovém případě má tendenci dynamický úsek přehlédnout. Typickým příkladem může být rozvlnění v okolí nespojitých změn (Gibbsův jev). Toto je ovšem spíše problém volby správné strategie vyhlazení, než samotného principu waveletové filtrace.

Jako žádná jiná metoda, nejsou ovšem ani wavelety universální všelék pro všechny typy úloh, ale v každém případě významně obohacují repertoár dosud běžně používaných technik.

Literatura

- [1] Akansu, A. N. – Haddad, R. A. *Multiresolution Signal Decomposition*. Academic Press, Inc., Boston, USA., 1992.
- [2] Arino, M. A. *Time Series Forecasts Via Wavelets*. IESE Universidad de Navarra. Barcelona 1995.
- [3] Chui, K. *An Introduction to Wavelets*. Academia Press. 1992.
- [4] Sebera, Martin - Seberová, Helena. *Economical time series prediction via wavelets*. In Nostradamus 2001. Zlín: Univerzita Tomáše Bati, 2001. 9 s. ISBN 80-7318-030-8.
- [5] Veselý, V. *Wavelety a jejich použití při filtraci dat*. In ROBUST'1996. Praha 1997, s. 241-272.