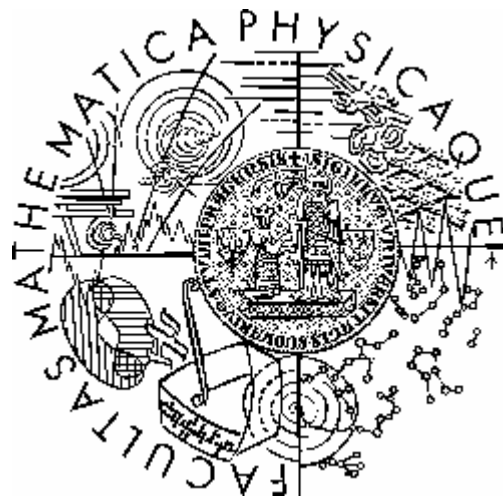


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Tomáš Hanzák

Dekompoziční metody pro časové řady
s nepravidelně pozorovanými hodnotami

Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: Prof. RNDr. Tomáš Cipra, DrSc.

Studijní program: Matematika

Studijní obor: Matematická statistika, pravděpodobnost a ekonometrie

Studijní plán: Ekonometrie

Upřímně děkuji vedoucímu své diplomové práce Prof. Tomáši Ciprovi za výběr a udělení tématu práce, poskytnutí potřebné literatury, jeho cenné rady a připomínky a především za laskavou ochotu při vzájemné spolupráci.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 12.4. 2007

Tomáš Hanzák

Obsah

Abstrakt / Abstract	4
1 Úvod	5
2 Základní pojmy a metody	7
2.1 Nepravidelné časové řady.....	7
2.2 Dekompoziční a rekurentní metody pro časové řady.....	9
2.3 Exponenciální vyrovnávání	11
3 Jednoduché exponenciální vyrovnávání	14
3.1 Jednoduché exponenciální vyrovnávání pro pravidelné časové řady	14
3.2 Wrightova modifikace pro nepravidelné řady	17
3.3 Nepravidelně pozorovaný ARIMA(0, 1, 1) proces.....	20
4 Holtova metoda	24
4.1 Holtova metoda pro pravidelné časové řady	24
4.2 Wrightova modifikace pro nepravidelné časové řady.....	30
4.3 Holt-Wintersova metoda pro řady s chybějícími pozorováními.....	34
5 Exponenciální vyrovnávání řádu m	38
5.1 Exponenciální vyrovnávání řádu m pro pravidelné časové řady	38
5.2 Exponenciální vyrovnávání řádu m pro nepravidelné časové řady	42
5.3 DLS odhad polynomického trendu stupně m	51
6 Některé výpočetní aspekty metod	57
6.1 Odhad parametrů metodou maximální věrohodnosti.....	57
6.2 Míry přesnosti a adekvátnosti předpovědních metod	59
6.3 Transformace časových řad	62
6.4 Praktické problémy a zkušenosti	66
7 Softwarová realizace	68
7.1 Program DMITS	68
7.2 Numerické příklady.....	72
8 Závěr	78
Literatura	79

Abstrakt

Název práce: Dekompoziční metody pro časové řady s nepravidelně pozorovanými hodnotami

Autor: Tomáš Hanzák

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Prof. RNDr. Tomáš Cipra, DrSc.

e-mail vedoucího: cipra@karlin.mff.cuni.cz

Abstrakt: Práce se věnuje zobecněním klasických metod typu exponenciálního vyrovnávání pro jednorozměrné časové řady s nepravidelně pozorovanými hodnotami. Prezentována jsou zobecnění jednoduchého exponenciálního vyrovnávání, Holtovy a Holt-Wintersovy metody a dvojitého exponenciálního vyrovnávání pro nepravidelné časové řady, která byla v minulosti vyvinuta. Je navržena metoda alternativní k Wrightově modifikaci jednoduchého exponenciálního vyrovnávání pro nepravidelné řady, založená na příslušném ARIMA procesu. Odvozeno je exponenciální vyrovnávání řádu m pro nepravidelné časové řady, které je zobecněním jednoduchého a dvojitého exponenciálního vyrovnávání. Podobná metoda, založená na DLS (*discounted least squares*) odhadu polynomického trendu stupně m , je též odvozena. Ve všech případech je zachován rekurentní charakter původních metod a tak i jejich implementační a výpočetní nenáročnost. Součástí diplomové práce je program, v němž je dostupná většina zde prezentovaných metod. Je též uvedeno několik numerických příkladů jejich použití.

Klíčová slova: časové řady, exponenciální vyrovnávání řádu m , Holtova metoda, jednoduché exponenciální vyrovnávání, nepravidelná pozorování.

Abstract

Title: Decomposition methods for time series with irregular observations

Author: Tomáš Hanzák

Department: Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Tomáš Cipra, DrSc.

Supervisor's e-mail address: cipra@karlin.mff.cuni.cz

Abstract: This work deals with extensions of classical exponential smoothing type methods for univariate time series with irregular observations. Extensions of simple exponential smoothing, Holt method, Holt-Winters method and double exponential smoothing which have been developed in past are presented. An alternative method to Wright's modification of simple exponential smoothing for irregular data, based on the corresponding ARIMA process, is suggested. Exponential smoothing of order m for irregular data as a generalization of simple and double exponential smoothing is derived. A similar method using a DLS (*discounted least squares*) estimation of polynomial trend of order m is derived as well. In all cases the recursive character of these methods is preserved making them easy to implement and high computationally effective. A program in which most of the methods presented here are available is a part of the work. Some numerical examples of their application are also included.

Keywords: exponential smoothing of order m , Holt method, irregular observations, simple exponential smoothing, time series.

1 Úvod

Schopnost činit kvalitní předpovědi o budoucím vývoji sledovaných jevů a veličin je bezesporu klíčovým faktorem pro naše současná rozhodnutí. Jedním z nástrojů pro získání takových předpovědí je statistická disciplína označovaná souhrnně jako *analýza časových řad*. Ze známých hodnot sledované veličiny z minulosti až do současnosti, které tvoří tzv. *časovou řadu*, se snažíme odhadnout její budoucí vývoj. I přes nepopíratelné limity uvedeného přístupu jsou takto získané předpovědi často uspokojivě přesné. Navíc jsou metody předpovídání v časových řadách podrobně zpracované a mnohé z nich jsou dostupné v příslušném softwaru.

Drtivá většina metod analýzy časových řad je navrhovaná pro práci s pravidelnými časovými řadami, tedy takovými, jejichž sousední pozorování mají od sebe konstantní časovou vzdálenost. Ještě poměrně dost pozornosti bylo věnováno zvládnutí problému tzv. *chybějících pozorování* v jinak pravidelných časových řadách. Podstatně méně prostoru ale doposud náleželo metodám, které by dokázaly zacházet s obecně nepravidelnými časovými řadami. I s nimi se přitom v praxi můžeme setkat překvapivě často.

Tato diplomová práce se věnuje právě metodám pro vyrovnávání a předpovídání v časových řadách s nepravidelně pozorovanými hodnotami. Jejím cílem je podat přehled o existujících metodách, navrhnout jejich případná vylepšení či k nim alternativní metody. Značná pozornost je věnována praktickým problémům souvisejícím s použitím těchto metod, například konstrukci předpovědních intervalů.

Práce se soustředí výhradně na metody typu *exponenciálního vyrovnávání* a jejich zobecnění pro nepravidelné časové řady. U všech těchto zobecnění jsou zachovány důležité vlastnosti původních metod, pro které jsou v praxi ceněny. I nadále jde o metody *dekompoziční* a *adaptivní*, z praktického hlediska je významné zachování jejich *rekurentního charakteru*, a tak i implementační a výpočetní nenáročnosti.

Tato práce se omezuje na jednorozměrné časové řady, i když v odborné literatuře se můžeme setkat i s jednoduchými zobecněními metod exponenciálního vyrovnávání na vícerozměrné časové řady, viz. např. Cipra (1989). Stejně tak další modifikace a zobecnění těchto klasických metod, vyvinuté pro různé speciální situace, zde nebudou zmíněny a dále zobecňovány pro nepravidelné časové řady, přestože by to bylo možné. Až na jednu výjimku se budeme soustředit na metody pro nesezónní časové řady.

V práci jsou prezentována zobecnění jednoduchého exponenciálního vyrovnávání (odst. 3.2), Holtovy metody včetně její verze s tlumeným lineárním trendem (odst. 4.2), Holt-Wintersovy metody (odst. 4.3) a dvojitého exponenciálního vyrovnávání (odst. 5.2) pro nepravidelné časové řady, které byly v uplynulých letech publikovány v odborné literatuře, viz. Wright (1986), Cipra a kol. (1995) a Cipra (2006).

V odstavci 3.3 je navržena metoda alternativní k Wrightovu exponenciálnímu vyrovnávání pro nepravidelné řady, odvozená na základě předpokladu, že zkoumaná

časová řada je nepravidelně pozorovaný ARIMA(0, 1, 1) proces. V odstavci 5.2 je odvozeno exponenciální vyrovnávání řádu m pro nepravidelné časové řady, které je zobecněním dvojitého exponenciálního vyrovnávání (případu $m = 1$) pro nepravidelné časové řady z článku Cipra (2006). Vedle toho je odvozena i podobná metoda založená na DLS (*discounted least squares*) odhadu polynomiálního trendu stupně m (odst. 5.3), která s exponenciálním vyrovnáváním řádu m splývá pouze ve verzi pro pravidelné časové řady.

Pozornost je věnována výpočtu počátečních hodnot rekurentních metod, pro některé existující metody jsou navrženy drobné modifikace používaných postupů. Pro jednoduché exponenciální vyrovnávání a Holtovu metodu jsou navrženy vzorce pro rozptyly chyb předpovědí s delším časovým horizontem, založené na předpokladu, že jedнокrokové předpovědní chyby tvoří bílý šum. Při výpočtu počátečních hodnot a rozptylů předpovědních chyb je využíván vztah mezi dvojitým exponenciálním vyrovnáváním a Holtovou metodou, který existuje v případě pravidelné časové řady.

Je navržen maximálně věrohodný odhad parametrů předpovědní metody (odst. 6.1), který je za předpokladu normality zobecněním klasické minimalizace MSE kritéria. Pro předpovídání v nepravidelné časové řadě je zaveden pojem *normalizovaných předpovědních chyb*, které jsou mimo jiné využívány při testech adekvátnosti použití dané předpovědní metody (odst. 6.2). Pro vyhodnocení efektivity předpovědních metod jsou navrženy ukazatele obdobné koeficientu determinace v lineární regresi (odst. 6.2). V odstavci 6.3 jsou kromě logaritmické transformace navrženy i některé další transformace časových řad. Odstavec 6.4 shrnuje autorovy praktické zkušenosti s aplikací prezentovaných metod.

Součástí diplomové práce je program DMITS (zkratka pro *Decomposition Methods for Irregular Time Series*), v němž jsou dostupné všechny metody v této práci uvedené či odvozené, s výjimkou Holt-Wintersovy metody. Implementace metod zahrnuje výpočet počátečních hodnot, optimální volbu parametrů, výpočet bodových a intervalových předpovědí a vyhodnocení přesnosti a adekvátnosti použití dané předpovědní metody (viz. odst. 7.1).

V odstavci 7.2 jsou uvedeny dva konkrétní numerické příklady použití prezentovaných metod. Dále je zde pomocí většího množství simulovaných časových řad porovnávána přesnost Wrightova jednoduchého exponenciálního vyrovnávání a alternativní metody navržené v této práci.

2 Základní pojmy a metody

V této rozsáhlejší úvodní kapitole jsou vyloženy některé základní pojmy a typy metod, kterých se text diplomové práce týká. Odstavec 2.1 detailněji pojednává o pravidelných a nepravidelných časových řadách a uvádí příklady, kdy se v praxi můžeme setkat s nepravidelnými časovými řadami. V odstavci 2.2 je velmi stručně popsán obecný princip dekompozice časových řad, do jehož rámce spadají i všechny zde probírané metody. Ze stejného důvodu je zde obecná zmínka o rekurentních metodách. Odstavec 2.3 se věnuje základní myšlence a historii metod typu exponenciálního vyrovnávání, jimiž je diplomová práce věnována.

2.1 Nepravidelné časové řady

Klasická (pravidelná) časová řada je soubor hodnot (pozorování, měření) jisté veličiny v pravidelně rozmístěných časových okamžicích. Může jít o hodnoty nějaké spojité veličiny na předem zvolené pravidelné časové mřížce, o agregaci nějaké aktivity v rámci stejně dlouhých a pravidelně rozmístěných časových intervalů nebo o pozorování nějakého pravidelně se opakujícího jevu. Jako ilustrační příklady těchto tří možností můžeme uvést např. teploty měřené na daném místě každý den vždy přesně v poledne, počty zákonů schválených PSP ČR v jednotlivých letech a denní řadu počtu diváků hlavních zpráv vybrané televizní stanice.

Reálné situace, z nichž vznikají konkrétní časové řady, jsou tedy velmi rozmanité nejen svou obsahovou stránkou ale také smyslem příslušné časové osy. Důležitý je zde předpoklad, že časové okamžiky, ke kterým jednotlivá pozorování vztahujeme, jsou (z nějakého rozhodujícího hlediska) pravidelně rozmístěné. Jak bude snad vidět později, tento předpoklad není vždy tak jednoznačný, jak by se mohlo zdát.

Rozhodneme-li se danou časovou řadu považovat za řadu pravidelnou, značíme její hodnoty obvykle jako

$$y_1, y_2, \dots, y_n, \quad (2.1.1)$$

kde y_t je pozorování řady y v čase t , $t=1, 2, \dots, n$ a přirozené číslo n udává délku dotyčné časové řady. Drtivá většina prací z oblasti analýzy časových řad se zabývá modely a metodami použitelnými právě pro takovéto (pravidelné) časové řady.

V praxi se skutečně ve velké míře setkáváme s pravidelně pozorovanými časovými řadami. Většina institucí publikuje své statistiky pravidelně (měsíčně, čtvrtletně, ročně apod.), pokud máme my sami možnost organizovat naše vlastní měření, zvolíme pravidelné časové intervaly, leda by tomu bránily podstatné objektivní důvody.

Přesto existují situace, kdy máme k dispozici časovou řadu složenou z hodnot pozorovaných v nepravidelně rozmístěných časových okamžicích. Označíme-li tyto okamžiky jako

$$t_1, t_2, \dots, t_n, \quad (2.1.2)$$

pak hodnoty dané časové řady budeme značit podobně jako v (2.1.1)

$$y_{t_1}, y_{t_2}, \dots, y_{t_n}. \quad (2.1.3)$$

Přirozeně je požadováno, aby platilo $t_1 < t_2 < \dots < t_n$. Mluvíme o časové řadě s *nepravidelně pozorovanými hodnotami* či stručněji o *nepravidelné časové řadě*. Pravidelnou časovou řadu je možné chápat jako speciální případ nepravidelné časové řady, z opačného pohledu lze říci, že nepravidelné časové řady jsou zobecněním těch pravidelných.

Je samozřejmě možné mírnou nepravidelnost v pozorování zkoumané časové řady zanedbat. Jsou-li rozdíly $t_{i+1} - t_i$ „téměř konstantní“, nelze očekávat, že by použití metody pro pravidelné řady vedlo k zásadně chybným výsledkům. Tato situace může vzniknout např. kvůli nepravidelnostem v používaném kalendáři (různá délka měsíců a let).

Jakýmsi mezistupněm jsou tzv. *časové řady s chybějícími pozorováními*, které vznikají z pravidelné řady (2.1.1) pomyslným vypuštěním některých jednotlivých pozorování či jejich celých úseků. Nutným požadavkem, abychom mohli mluvit o časové řadě s chybějícími pozorováními, je tedy to, aby všechny rozdíly $t_{i+1} - t_i$ byly násobky nějakého základního časového kroku Δt a pokud možno se mu často rovnaly. Existují modely a metody, které jsou použitelné na řady s chybějícími pozorováními, ale nikoli tak už na řady zcela obecně nepravidelné. To platí obzvláště pro sezónní časové řady. Jsou-li mezery v datech krátké a řídké, můžeme se pokusit doplnit chybějící hodnoty, ať už expertními odhady nebo nějakou formou interpolace, a na doplněnou řadu použít již některou z metod pro pravidelné časové řady. Jak již bylo řečeno v úplném úvodu, v této práci se soustředíme na metody aplikovatelné na zcela obecné nepravidelné časové řady.

Četné příklady časových řad s nepravidelně pozorovanými hodnotami lze nalézt v článku Wright (1986). Např. pokud v průběhu času dojde ke změně frekvence, se kterou naše pozorování provádíme, je celkovým výsledkem nepravidelná časová řada. Běžně se stává, že statistické úřady a jiné instituce zvyšují publikační frekvenci některých veličin, např. z roční na kvartální. V souhrnných přehledech je často pro úsporu místa volena pro starší období menší frekvence než pro období nedávné.

Jindy jsou nepravidelnosti v našich pozorováních způsobené objektivní nemožností měřit danou veličinu pravidelně, případně vznikají neplánovaně v důsledku výpadku měřících přístrojů nebo dokonce pozdější ztráty pozorovaných hodnot.

Někdy je časová nepravidelnost vnitřně obsažena už v samotném sledovaném jevu. Je možné např. uvažovat časovou řadu počtu zaznamenaných výskytů jisté choroby na daném území, přičemž okamžikem pozorování je vždy výskyt nového případu. Nebo

můžeme chtít předpovědět, jaká bude hodnota mužského světového rekordu v běhu na jednu míli v roce 2020, a to z časové řady historických hodnot tohoto rekordu, kde okamžikem pozorování je vždy datum vytvoření rekordu (viz. odst. 7.2).

Na závěr uveďme jeden příklad ilustrující, že hranice mezi pravidelnými a nepravidelnými časovými řadami nemusí být vždy tak ostrá. Uvažujme řadu denních uzavíracích cen vybraného akciového titulu na burze cenných papírů, kde se však obchoduje jen ve všední dny. Vzniklá časová řada vypadá jako typická řada s chybějícími pozorováními (víkendy a svátky), ovšem tyto chybějící hodnoty nebyly ani nikdy realizovány. Otázka zní, co je z hlediska vývoje ceny akcie důležitější. Zda fakt, že pátek a pondělí jsou od sebe vzdáleny 72 hodin, nebo že jde o dva po sobě jdoucí obchodní dny stejně jako třeba úterý a středa.

2.2 Dekompoziční a rekurentní metody pro časové řady

Dekompozice je jednou ze základních a široce rozšířených metod pro modelování, vyrovnávání a předpovídání časových řad. Jde o princip velice obecný, pod jehož hlavičku lze zařadit velké množství celých tříd metod. Základní myšlenkou je danou časovou řadu rozložit na několik složek s charakteristickými vlastnostmi a dále tohoto rozkladu využít při řešení jednotlivých praktických úloh analýzy časových řad.

Mějme časovou řadu y_1, y_2, \dots, y_n a uvažujme rozklad jejích hodnot

$$y_t = T_t + I_t + \varepsilon_t . \quad (2.2.1)$$

Zde T_t je hodnota *trendu*, I_t hodnota *sezónní složky* a ε_t hodnota *náhodné (nepravidelné) složky* časové řady y v čase t . Trend T_t , $t = 1, 2, \dots, n$ by měl vykazovat určitou míru hladkosti (v porovnání s originální řadou) a absenci periodicity. Sezónní složka I_t , $t = 1, 2, \dots, n$ by měla vykazovat periodicitu s pevnou periodou $p \geq 2$ a měla by být centrována kolem 0. Náhodná složka ε_t , $t = 1, 2, \dots, n$ by ideálně měla tvořit bílý šum. Trend a sezónní složku nazýváme *systematickými* složkami.

Existuje velké množství modifikací, rozšíření či zjednodušení tohoto základního schématu. Místo *aditivní* dekompozice (2.2.1) lze uvažovat *multiplikativní* dekompozici

$$y_t = T_t \cdot I_t \cdot \varepsilon_t , \quad (2.2.2)$$

kde složky I_t a ε_t jsou centrovány kolem 1. Zlogaritmováním (2.2.2) získáme aditivní rozklad řady $\log y_t$. Jiné dekompozice v sobě kombinují vlastnosti obou uvedených.

Dekompoziční schémata se mohou lišit i počtem a charakterem složek, na které je řada rozkládána. Tak pro nesezónní časové řady bude rozklad postrádat sezónní složku I_t . Jiné časové řady mohou naopak vykazovat více různých periodicit, příkladem buď denní a týdenní periodicitu spotřeby elektrické energie, viz. Taylor (2003). Někdy se uvažuje tzv. *cyklická složka* C_t , která vyjadřuje cyklické kolísání řady kolem jejího

dlouhodobého trendu. Na rozdíl od sezónní složky nemá však toto kolísání periodu pevné délky. Jako příklad vezmeme kolísání HDP v rámci ekonomického cyklu.

Konkrétní dekompoziční metody se liší také tím, jakým způsobem je rozklad řady na jednotlivé složky prakticky proveden. Jedny se snaží v časové řadě identifikovat trend a sezónní složku ve tvaru neměnném v průběhu času. Sem patří především použití regrese k nalezení trendu v podobě různých matematických křivek a sezónní složky tvořené goniometrickými funkcemi času nebo sezónními indexy. Parametry regresních modelů jsou tu odhadovány jako konstantní v rámci celé časové řady.

Tzv. *adaptivní* metody naopak připouštějí, že tvar trendu a sezónní složky se mohou v průběhu času měnit. Odhady příslušných parametrů jsou prováděny lokálně, hovoříme o *lokálním trendu*. Mezi adaptivní metody patří nejrůznější způsoby tzv. *vyrovnávání* časových řad, mezi nejznámější patří různé typy *klouzavých průměrů*. Jejich speciálním případem je i *jednoduché exponenciální vyrovnávání*, na jehož základě je vyvinuta celá škála adaptivních metod pro různé typy trendů a sezónností. Zobecnění těchto metod na nepravidelné časové řady jsou tématem této práce.

Předpovídání neznámých budoucích hodnot časové řady pomocí dekompoziční metody probíhá následujícím způsobem: nejprve je provedena samotná dekompozice řady a následně jsou takto získané systematické složky vhodným způsobem extrapolovány do budoucnosti pro získání příslušných bodových předpovědí.

Z nepřeberného množství vyrovnávacích a předpovědních metod pro časové řady se v praxi těší velké oblibě především tzv. *rekurentní metody*. Jejich výhodou je snadná softwarová implementace a následná výpočetní nenáročnost. Předpokládejme, že jsme již pozorovali hodnoty y_1, y_2, \dots, y_n a v rámci algoritmu metody jsme z nich napočítali hodnoty k statistik S_n^1, \dots, S_n^k . Samotné hodnoty y_1, y_2, \dots, y_n již v tuto chvíli v paměti nutně neuchováme. Číslo k je přitom relativně malé (může být i $k = 1$) a především pevné pro danou metodu (speciálně nezávisí na počtu pozorování n).

Vyrovnanou hodnotu v čase n získáme jako funkci statistik S :

$$\hat{y}_n = Y(S_n^1, \dots, S_n^k) . \quad (2.2.3)$$

Bodovou předpověď z času n o h časových jednotek vpřed pak obdobně jako

$$\hat{y}_{n+h}(n) = F(S_n^1, \dots, S_n^k; h) . \quad (2.2.4)$$

Jakmile pozorujeme novou hodnotu y_{n+1} , provedeme přepočítání hodnot našich k statistik pomocí rekurentního vzorce

$$(S_{n+1}^1, \dots, S_{n+1}^k) = S(S_n^1, \dots, S_n^k; y_{n+1}) , \quad (2.2.5)$$

kde S je nějaká pevná k -složková vektorová funkce. Samotnou hodnotu y_{n+1} můžeme poté zapomenout a celý postup se opakuje. V případě, že bychom pracovali s nepravidelnou časovou řadou, tak kromě nového pozorování y_{n+1} vstoupí do vztahu (2.2.5) ještě hodnota časového kroku $t_{n+1} - t_n$. Funkce Y , F a S , které určují danou rekurentní metodu, jsou často velice jednoduché a názorné.

Pro praktické použití rekurentní metody je nutné nejprve nějakým způsobem získat počáteční hodnoty S_0^1, \dots, S_0^k . Ty se určí většinou z několika prvních pozorování řady, která již máme k dispozici. Pro danou rekurentní metodu zpravidla existuje více alternativních způsobů výpočtu počátečních hodnot S_0^1, \dots, S_0^k .

Velmi obecnou rekurentní metodou je např. tzv. *Kalmanův filtr* použitý na časovou řadu modelovanou tzv. *dynamickým lineárním modelem*, viz. Brockwell a Davis (2002), kap. 8. Patrně ve všech ohledech nejrozšířenějšími rekurentními metodami pro časové řady jsou metody typu *exponenciálního vyrovnávání*, jejichž zobecněním pro nepravidelné časové řady je věnována tato práce.

2.3 Exponenciální vyrovnávání

Metody dnes souhrnně nazývané jako *exponenciální vyrovnávání* byly vyvinuty na konci 50. let k předpovídání budoucích objemů prodeje zboží za účelem optimálního řízení jejich výroby a skladování. Myšlenku použít konceptu exponenciálního vážení k odhadu nejen úrovně časové řady, ale i jejího trendu a sezónní složky, což pak umožní předpovídat její budoucí hodnoty, publikoval jako první Američan Charles C. Holt v roce 1957 ve svém memorandu pro Office of Naval Research. Výsledné předpovědní metody včetně dobového aplikačního kontextu lze nalézt v článku Winters (1960).

Jejich nejjednodušší variantou je tzv. *jednoduché exponenciální vyrovnávání* vhodné pro nesezónní řady s lokálně konstantním trendem. *Holtova metoda* je vhodná pro nesezónní časové řady s lokálně lineárním trendem, *Holt-Wintersova metoda* navíc připouští aditivní nebo multiplikatívni sezónnost. V průběhu let se objevily další odvozené varianty jako např. Holtova metoda s *exponenciálním* či tzv. *tlumeným lineárním* trendem. Podrobný přehled lze nalézt v článku Gardner (1985).

Přes velkou různorodost mají všechny varianty exponenciálního vyrovnávání důležité společné rysy. Předně jsou založeny na stejné základní myšlence, kterou je vážení s vahami exponenciálně klesajícími směrem do minulosti. Odtud pak pramení dvě v praxi ceněné vlastnosti exponenciálního vyrovnávání: jeho rekurentní a adaptivní charakter. Základní myšlenku i uvedené vlastnosti exponenciálního vyrovnávání si budeme ilustrovat na jednoduchém exponenciálním vyrovnávání, které je základem všech dalších variant.

Dejme tomu, že jsme již pozorovali hodnoty y_1, y_2, \dots, y_n časové řady a naším úkolem je sestavit předpověď $\hat{y}_{n+1}(n)$ neznámé budoucí hodnoty y_{n+1} (z času n). Existují dva velmi jednoduché způsoby konstrukce této předpovědi.

První z nich je vzít jako předpověď neznámé budoucí hodnoty aritmetický průměr všech doposud pozorovaných hodnot dotyčné časové řady, tedy

$$\hat{y}_{n+1}(n) = \frac{1}{n}(y_1 + y_2 + \dots + y_n) . \quad (2.3.1)$$

Tato předpověď je vhodná, pokud hodnoty řady y náhodně kolísají kolem jisté úrovně, která se v čase nemění (*model konstantní úrovně*).

Druhou možností je vzít za předpověď neznámé hodnoty y_{n+1} předcházející pozorovanou hodnotu y_n , tedy

$$\hat{y}_{n+1}(n) = y_n . \quad (2.3.2)$$

Tato předpověď je naopak vhodná, pokud hodnoty řady y vznikají náhodným odchýlením se od předcházející hodnoty (*model náhodné procházky*).

V obou případech předpovídáme hodnotu y_{n+1} pomocí *vyrovnané hodnoty* \hat{y}_n , která představuje odhad úrovně řady y v čase n . Přitom \hat{y}_n je váženým průměrem pozorovaných hodnot y_1, y_2, \dots, y_n , kde v prvním případě jsou váhy rovnoměrné, ve druhém koncentrované na nejaktuálnější pozorování.

Pro model konstantní úrovně získáme snadno rekurentní formuli

$$\hat{y}_{n+1} = \left(1 - \frac{1}{n+1}\right) \cdot \hat{y}_n + \frac{1}{n+1} \cdot y_{n+1} \quad (2.3.3)$$

a pro model náhodné procházky podobně

$$\hat{y}_{n+1} = (1-1) \cdot \hat{y}_n + 1 \cdot y_{n+1} . \quad (2.3.4)$$

Nová vyrovnaná hodnota \hat{y}_{n+1} tedy vzniká jako konvexní lineární kombinace předchozí vyrovnané hodnoty \hat{y}_n a naposledy pozorované hodnoty y_{n+1} . V prvním případě je váha soustředěna na hodnotu \hat{y}_n (předpokládáme-li $n \gg 0$), ve druhém je celá váha soustředěna na nejnovější pozorování y_{n+1} .

Označíme-li chybu předpovědi $\hat{y}_{n+1}(n)$ jako e_{n+1} , tedy

$$e_{n+1} = y_{n+1} - \hat{y}_{n+1}(n) = y_{n+1} - \hat{y}_n , \quad (2.3.5)$$

můžeme rovnosti (2.3.3) a (2.3.4) přepsat na tzv. *chybový tvar* :

$$\hat{y}_{n+1} = \hat{y}_n + \frac{1}{n+1} \cdot e_{n+1} \quad (2.3.6)$$

pro model konstantní úrovně a

$$\hat{y}_{n+1} = \hat{y}_n + 1 \cdot e_{n+1} \quad (2.3.7)$$

pro model náhodné procházky. Obě předpovědní metody se tedy liší tím, do jaké míry absorbuje nová vyrovnaná hodnota \hat{y}_{n+1} poslední zaznamenanou předpovědní chybu e_{n+1} . Zatímco v prvním případě je tato chyba považována pouze za důsledek náhodné odchylky od současné úrovně řady a jako taková není téměř absorbována, ve druhém

případě je předpovědní chyba brána jako příznak změny úrovně řady a jako taková je absorbována v plné míře.

Jak bylo řečeno, obě popsané předpovědní metody jsou vhodné pouze pro velmi úzce definované třídy časových řad, které se v praxi vyskytují zřídka. Mnohem realističtější předpovědní metodu ale můžeme získat jako jejich kompromis. Vyrovnanou hodnotu \hat{y}_n budeme opět počítat jako vážený průměr pozorovaných hodnot y_1, y_2, \dots, y_n , váhy ale tentokrát zvolíme exponenciálně klesající směrem do minulosti s *diskontním faktorem* $\beta \in (0, 1)$. Tedy

$$\hat{y}_n = \frac{y_n + \beta y_{n-1} + \beta^2 y_{n-2} + \dots + \beta^{n-1} y_1}{1 + \beta + \beta^2 + \dots + \beta^{n-1}} = \frac{1 - \beta}{1 - \beta^n} \cdot \sum_{t=1}^n y_t \beta^{n-t}. \quad (2.3.8)$$

Tyto váhy jsou rozumným kompromisem mezi rovnoměrným rozložením vah a vahou soustředěnou jen na poslední pozorování. Navíc díky jejich speciálnímu tvaru je stále možné počítat vyrovnané hodnoty řady rekurentně. Formule obdobná těm v (2.3.3) a (2.3.4) má nyní tvar

$$\hat{y}_{n+1} = (1 - \alpha_n) \cdot \hat{y}_n + \alpha_n \cdot y_{n+1}, \quad (2.3.9)$$

kde jsme označili $\alpha_n = \frac{1 - \beta}{1 - \beta^{n+1}}$. Její příslušný chybový tvar je

$$\hat{y}_{n+1} = \hat{y}_n + \alpha_n \cdot e_{n+1}. \quad (2.3.10)$$

Pro $n \gg 0$ platí $\alpha_n = \frac{1 - \beta}{1 - \beta^{n+1}} \approx 1 - \beta$ a označíme-li $\alpha = 1 - \beta$, můžeme vztahy (2.3.9)

a (2.3.10) psát přibližně jako

$$\hat{y}_{n+1} = (1 - \alpha) \cdot \hat{y}_n + \alpha \cdot y_{n+1}, \quad (2.3.11)$$

$$\hat{y}_{n+1} = \hat{y}_n + \alpha \cdot e_{n+1}. \quad (2.3.12)$$

Volbou tzv. *vyrovnávací konstanty* $\alpha = 1 - \beta$ určíme rozložení vah v příslušné konvexní lineární kombinaci. Předpovědní chyba e_{n+1} je chápána z části jako důsledek náhodné odchylky od současné úrovně řady a z části jako příznak změny úrovně řady. Parametr $\alpha \in (0, 1)$ určuje míru její absorpce.

Získali jsme tedy celou škálu adaptivních rekurentních předpovědních metod. pro $\alpha \approx 0$ se blížíme modelu konstantní úrovně, pro $\alpha \approx 1$ naopak modelu náhodné procházky. Základními vzorci jednoduchého exponenciálního vyrovnávání jsou tedy předpovědní formule $\hat{y}_{n+1}(n) = \hat{y}_n$ a rekurentní formule (2.3.11), jejíž tvar je charakteristický pro všechny metody exponenciálního vyrovnávání.

3 Jednoduché exponenciální vyrovnávání

Tato kapitola bude věnována jednoduchému exponenciálnímu vyrovnávání a jeho zobecněním pro nepravidelné časové řady. V odstavci 3.1 bude nejprve prezentována klasická verze této metody pro pravidelné časové řady a její spojitost s modelem ARIMA(0, 1, 1). V odstavci 3.2 bude uvedeno Wrightovo zobecnění jednoduchého exponenciálního vyrovnávání pro případ časových řad s nepravidelně pozorovanými hodnotami. V odstavci 3.3 odvodíme alternativní metodu vycházející z toho, že zkoumaná časová řada je nepravidelně pozorovaný ARIMA(0, 1, 1) proces.

3.1 Jednoduché exponenciální vyrovnávání pro pravidelné časové řady

Jednoduché exponenciální vyrovnávání je nejjednodušší a nejznámější metodou exponenciálního vyrovnávání. Vhodné je pro použití na nesezónní časové řady bez zřetelného rostoucího nebo klesajícího trendu. Někdy říkáme, že řada by měla mít *lokálně konstantní* trend. Při použití na časovou řadu, která nemá výše popsané vlastnosti, lze očekávat velmi špatné výsledky. Pro řady s trendem či sezónností jsou určeny složitější metody exponenciálního vyrovnávání (viz. kap. 4 a 5).

Myšlenka jednoduchého exponenciálního vyrovnávání byla již nastíněna v odstavci 2.3. Pracujeme-li s časovou řadou $\dots y_{n-1}, y_n, y_{n+1} \dots$, můžeme základní vzorce metody zapsat takto:

$$S_{n+1} = (1 - \alpha) \cdot S_n + \alpha \cdot y_{n+1} , \quad (3.1.1)$$

$$\hat{y}_n = S_n , \quad (3.1.2)$$

$$\hat{y}_{n+\tau}(n) = S_n , \quad \tau > 0 . \quad (3.1.3)$$

Parametr $\alpha \in (0, 1)$ je tzv. *vyrovnávací konstanta*, hodnotě S_n říkáme (*jednoduchá*) *vyrovnávací statistika*. Bodová předpověď z času n je rovna S_n nezávisle na horizontu předpovědi $\tau > 0$, tedy tyto předpovědi tvoří horizontální přímku. Vzorec (3.1.1) lze zapsat v jeho *chybovém tvaru* jako

$$S_{n+1} = S_n + \alpha \cdot e_{n+1} , \quad (3.1.4)$$

kde $e_{n+1} = y_{n+1} - S_n$ je chyba jednokrokové předpovědi z času n .

Pro praktické použití metody na řadu $y_1, y_2, \dots, y_n, \dots$ je nutné nějak zvolit počáteční hodnotu S_0 , která je nezbytná k nastartování rekurentního výpočtu podle vzorce (3.1.1). S_0 se často volí jednoduše rovno y_1 nebo průměru několika prvních

pozorování řady y . Za nejlepší volbu je ale považována tzv. *zpětná předpověď*, kdy se S_0 bere jako vyrovnaná hodnota v čase 1 získaná použitím stejné předpovědní metody na časově převrácenou řadu y , viz. Gardner (1985). Tomuto postupu víceméně odpovídá volba S_0 ve tvaru váženého průměru několika prvních členů řady, kde váhy klesají směrem *do budoucnosti* s diskontním faktorem $\beta = 1 - \alpha$.

Hodnota vyrovnávací konstanty $\alpha \in (0, 1)$ se volí buď expertně nebo minimalizací jistého kritéria vyjadřujícího nepřesnost prováděných předpovědí v nějakém úseku řady y . Tato minimalizace přes $\alpha \in (0, 1)$ se provádí nějakým numerickým algoritmem, který vyžaduje vyčíslit dotyčné předpovědi v řadě y pro větší množství různých hodnot α . Jiný způsob volby parametru α bude popsán na konci tohoto odstavce. Volbě parametrů předpovědních metod je věnována odstavce 6.1 a částečně i odstavce 6.2 a 6.4.

Jednoduché exponenciální vyrovnávání je typická *ad hoc* metoda, která není explicitně založena na nějakém pravděpodobnostním modelu pro zkoumanou řadu y . Zdánlivé ospravedlnění může poskytnout následující skutečnost, viz. Chatfield (2000), str. 96: Uvažujme časovou řadu $\dots y_{n-2}, y_{n-1}, y_n$ s nekonečnou minulostí a její model

$$y_t = \mu + \varepsilon_t, \quad t = \dots n-2, n-1, n, \quad (3.1.5)$$

kde μ je neznámý parametr a ε_t je náhodná složka. Odhadem parametru μ metodou DLS (*discounted least squares*) s diskontním faktorem $\beta = 1 - \alpha$, tedy řešením úlohy

$$\arg \min_{\mu} \left\{ \sum_{j=-\infty}^n \beta^{n-j} (y_j - \mu)^2 \right\}, \quad (3.1.6)$$

je hodnota $S_n = (1 - \beta) \cdot \sum_{j=-\infty}^n \beta^{n-j} y_j$, která vyhovuje rekurentní formuli (3.1.1).

Jednoduché exponenciální vyrovnávání použité na časovou řadu s konečnou minulostí je tedy aproximací odhadu úrovně μ metodou DLS. To ovšem neznamená, že bychom snad skutečně věřili v platnost modelu (3.1.5) nebo že bychom měli nějaký jasný důvod, proč k odhadu parametru μ použít zrovna metodu DLS. Jednoduché exponenciální vyrovnávání tedy pro tuto chvíli stále zůstává *ad hoc* metodou.

Důsledkem toho, že jednoduché exponenciální vyrovnávání nemusí být založeno na žádném pravděpodobnostním modelu, je také fakt, že nemáme žádné vodítko pro určení rozptylů předpovědních chyb a tedy ani pro určení mezí předpovědních intervalů. Řešením může být využití vztahu jednoduchého exponenciálního vyrovnávání k modelu ARIMA(0, 1, 1), který bude osvětlen na následujících řádcích.

Uvažujme první diferenci $\Delta y_n = y_n - y_{n-1}$ řady y . S využitím vztahu (3.1.4) a definice předpovědní chyby e_n získáme snadno

$$\Delta y_n = y_n - y_{n-1} = (S_{n-1} + e_n) - (S_{n-2} + e_{n-1}) = e_n + (\alpha - 1) \cdot e_{n-1}. \quad (3.1.7)$$

Rozumnou matematickou formulací toho, že metoda jednoduchého exponenciálního vyrovnávání s vyrovnávací konstantou α je vhodná pro řadu y , je pokládat náhodné veličiny $\{e_n, n \in \mathbf{Z}\}$ za bílý šum s rozptylem $\sigma^2 > 0$. Tedy předpokládat, že platí $E(e_n) = 0$, $\text{var}(e_n) = \sigma^2$ a $\text{cov}(e_n, e_m) = 0$ pro všechna $m \neq n$.

Vztah (3.1.7) pak znamená, že řada $\{\Delta y_n, n \in \mathbf{Z}\}$ se řídí modelem MA(1) generovaným bílým šumem $\{e_n, n \in \mathbf{Z}\}$, tedy původní řada $\{y_n, n \in \mathbf{Z}\}$ se řídí modelem ARIMA(0, 1, 1). Lze naopak snadno ukázat, že pro řadu řídicí se modelem ARIMA(0, 1, 1) je jednoduché exponenciální vyrovnávání optimální předpovědní metodou z hlediska minimalizace střední čtvercové chyby jednokrokové předpovědi, viz. např. Chatfield (2000), str. 90.

Při použití jednoduchého exponenciálního vyrovnávání na časovou řadu y se tedy jeví jako rozumné považovat ji za realizaci pravděpodobnostního modelu ARIMA(0, 1, 1). Tento předpoklad nám pak umožní odvodit přesné předpovědní rozptyly pro libovolný krok $\tau = 1, 2, \dots$ a při vyslovení předpokladu o typu rozdělení procesu $\{e_n, n \in \mathbf{Z}\}$ i přesné meze předpovědních intervalů s danou spolehlivostí $1 - \theta$.

Uvažujme předpověď o $\tau = 1, 2, \dots$ kroků vpřed $\hat{y}_{n+\tau}(n) = S_n$. Pro chybu této předpovědi (zatím nezávisle na předpokladu o ARIMA procesu) platí

$$\begin{aligned} e_{n+\tau}(n) &= y_{n+\tau} - \hat{y}_{n+\tau}(n) = y_{n+\tau} - S_n = \\ &= (S_n + \alpha \cdot e_{n+1} + \alpha \cdot e_{n+2} + \dots + \alpha \cdot e_{n+\tau-1} + e_{n+\tau}) - S_n = \\ &= \alpha \cdot (e_{n+1} + e_{n+2} + \dots + e_{n+\tau-1}) + e_{n+\tau}. \end{aligned} \quad (3.1.8)$$

Nyní již s využitím předpokladu o kovarianční struktuře procesu $\{e_n, n \in \mathbf{Z}\}$ dostaneme

$$\text{var}[e_{n+\tau}(n)] = \text{var}[\alpha \cdot (e_{n+1} + e_{n+2} + \dots + e_{n+\tau-1}) + e_{n+\tau}] = \sigma^2[\alpha^2(\tau-1) + 1]. \quad (3.1.9)$$

Pro $\tau = 1$ je tento rozptyl přirozeně roven σ^2 , obecně je pak přímo úměrný tomuto parametru. Dále dle očekávání $\text{var}[e_{n+\tau}(n)]$ roste (a to lineárně) s rostoucím horizontem předpovědi τ . Také větší hodnota parametru α , která znamená rychlejší změny úrovně řady y , se projevuje většími předpovědními rozptyly.

Pokud budeme navíc předpokládat, že $e_n \sim N(0, \sigma^2)$, pak

$$e_{n+\tau}(n) \sim N(0, \sigma^2 \cdot [\alpha^2(\tau-1) + 1]) \quad (3.1.10)$$

a příslušný předpovědní interval se spolehlivostí $1 - \theta$ má meze

$$S_n \pm u_{1-\theta/2} \cdot \sigma \sqrt{\alpha^2(\tau-1) + 1}, \quad (3.1.11)$$

kde $u_{1-\theta/2}$ je $(1 - \theta/2)$ -kvantil standardního normálního rozdělení $N(0, 1)$.

Všechny zde prováděné úvahy vycházejí z toho, že známe skutečné hodnoty parametrů α a σ^2 . V praxi je samozřejmě pouze odhadujeme z pozorovaných hodnot řady y . Vyrovnávací konstantu α určíme např. minimalizací *střední čtvercové chyby*

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (3.1.12)$$

(*mean square error*) a parametr σ^2 odhadneme zde dosaženým minimem, tedy výběrovým rozptylem jedнокrokových předpovědních chyb.

Alternativním přístupem je odhadnout parametr α na základě výběrové autokorelace řady $\{\Delta y_n, n \in \mathbf{Z}\}$, což odpovídá tzv. *momentovému odhadu* parametru v modelu MA(1) této řady, viz. Prášková (2004), str. 143-145. Platí

$$\rho_1 = \text{cor}(\Delta y_n, \Delta y_{n+1}) = \frac{\alpha - 1}{1 + (\alpha - 1)^2} \quad (3.1.13)$$

a pokud uvažujeme $\alpha \in (0, 1)$, pak $\rho_1 \in (-1/2, 0)$. Můžeme naopak vyjádřit

$$\alpha = 1 + \frac{1 - \sqrt{1 - 4\rho_1^2}}{2\rho_1} . \quad (3.1.14)$$

Odhad $\hat{\alpha}$ získáme dosazením výběrové korelace r_1 na místo teoretické hodnoty ρ_1 :

$$\hat{\alpha} = 1 + \frac{1 - \sqrt{1 - 4r_1^2}}{2r_1} . \quad (3.1.15)$$

Hodnota $\hat{\alpha}$ je rostoucí funkcí r_1 , pro $r_1 \rightarrow 0$ je $\alpha \rightarrow 1$ a naopak pro $r_1 \rightarrow -1/2$ je $\alpha \rightarrow 0$. Hodnota r_1 mimo interval $(-1/2, 0)$ značí nevhodnost použití jednoduchého exponenciálního vyrovnávání na danou časovou řadu.

Odhad $\hat{\alpha}$ je pochopitelně tím přesnější, čím přesnější je odhad r_1 a tedy čím více pozorování řady y máme k dispozici. Doporučuje se přitom alespoň 50 pozorování, viz. Prášková (2004), str. 129. Odhad $\hat{\alpha}$ se však obecně nechová dobře pro α blízká 0, jelikož funkce $\hat{\alpha}(r_1)$ má pro $r_1 \rightarrow -1/2$ (a tedy $\hat{\alpha} \rightarrow 0$) derivaci blížící se $+\infty$ a hodnota $\hat{\alpha}$ v blízkosti 0 tak velmi silně reaguje i na nepatrné změny hodnoty r_1 . Hodnotu odhadu $\hat{\alpha}$ lze použít jak přímo, tak jako počáteční hodnotu pro numerický algoritmus hledající optimální hodnotu α .

3.2 Wrightova modifikace pro nepravidelné řady

V tomto odstavci bude prezentováno jednoduché exponenciální vyrovnávání pro nepravidelně pozorované časové řady, viz. Wright (1986). Jde o přímočaré *ad hoc* zobecnění stejné metody pro pravidelné časové řady.

V případě pravidelné časové řady s nekonečnou minulostí je vyrovnávací statistika S_n rovna exponenciálně váženému průměru hodnot řady y pozorovaných do času n :

$$S_n = (1 - \beta) \cdot \sum_{j=-\infty}^n \beta^{n-j} y_j , \quad (3.2.1)$$

kde $\beta \in (0, 1)$ je použitý diskontní faktor. Úplně stejným způsobem můžeme postupovat i v případě časové řady $\dots y_{t_{n-2}}, y_{t_{n-1}}, y_{t_n}, \dots$ pozorované v nepravidelných časových okamžicích $\dots < t_{n-2} < t_{n-1} < t_n < \dots$. Ve vzorci (3.2.1) jednoduše zohledníme nepravidelnou časovou strukturu řady y a položíme

$$S_{t_n} = \frac{1}{\sum_{j=-\infty}^n \beta^{t_n-t_j}} \cdot \sum_{j=-\infty}^n \beta^{t_n-t_j} y_{t_j} . \quad (3.2.2)$$

Předpokládáme přitom, že posloupnost $\dots < t_{n-2} < t_{n-1} < t_n < \dots$ má takovou podobu, že všechny příslušné nekonečné řady konvergují. Označíme-li

$$\alpha_{t_n} = \left(\sum_{j=-\infty}^n \beta^{t_n-t_j} \right)^{-1} , \quad (3.2.3)$$

můžeme psát $S_{t_n} = \alpha_{t_n} \cdot \sum_{j=-\infty}^n \beta^{t_n-t_j} y_{t_j}$. Jednoduchými úpravami dostaneme

$$S_{t_{n+1}} = \alpha_{t_{n+1}} \cdot \sum_{j=-\infty}^{n+1} \beta^{t_{n+1}-t_j} y_{t_j} = \frac{\alpha_{t_{n+1}}}{\alpha_{t_n}} \cdot \beta^{t_{n+1}-t_n} \cdot S_{t_n} + \alpha_{t_{n+1}} \cdot y_{t_{n+1}} , \quad (3.2.4)$$

$$\alpha_{t_{n+1}} = \left(\sum_{j=-\infty}^{n+1} \beta^{t_{n+1}-t_j} \right)^{-1} = \frac{1}{1 + \beta^{t_{n+1}-t_n} \cdot \sum_{j=-\infty}^n \beta^{t_n-t_j}} = \frac{\alpha_{t_n}}{\alpha_{t_n} + \beta^{t_{n+1}-t_n}} . \quad (3.2.5)$$

Odtud můžeme (3.2.4) přepsat jako

$$S_{t_{n+1}} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n} + \alpha_{t_{n+1}} \cdot y_{t_{n+1}} , \quad (3.2.6)$$

případně v chybovém tvaru jako

$$S_{t_{n+1}} = S_{t_n} + \alpha_{t_{n+1}} \cdot e_{t_{n+1}} . \quad (3.2.7)$$

Toto jsou rekurentní formule obdobné těm ve (3.1.1) a (3.1.4). Pouze místo pevné vyrovnávací konstanty $\alpha = 1 - \beta$ se zde vyskytuje hodnota $\alpha_{t_{n+1}}$, která je přepočítávána rekurentním vztahem (3.2.5). Stejně jako v pravidelném případě budeme brát $\hat{y}_{t_n} = S_{t_n}$ a $\hat{y}_{t_n+\tau}(t_n) = S_{t_n}$, kde $\tau > 0$ může nyní nabývat i neceločíselných hodnot.

Narozdíl od jednoduchého exponenciálního vyrovnávání pro pravidelné časové řady tedy musíme kromě vyrovnávací statistiky S_{t_n} přepočítávat ještě proměnný vyrovnávací koeficient α_{t_n} . Jeho význam ve vzorcích (3.2.6) a (3.2.7) je zřejmý a stejný jako v pravidelném případě. Podívejme se podrobněji na rekurentní

vzorec (3.2.5), podle kterého je počítán. Rostoucí závislost $\alpha_{t_{n+1}}$ na časovém kroku $t_{n+1} - t_n$ znamená, že je-li aktuální vyrovnaná hodnota S_{t_n} již značně dále v minulosti od nového pozorování $y_{t_{n+1}}$, je větší váha kladena na nově pozorovanou hodnotu. Nová hodnota vyrovnávacího koeficientu $\alpha_{t_{n+1}}$ závisí rostoucím způsobem ještě na minulé hodnotě α_{t_n} , která v sobě obsahuje informaci o časové struktuře řady y od času n do minulosti.

Máme-li k dispozici pozorování řady y počínaje časem t_1 a chceme-li nastartovat rekurentní výpočty podle vzorců (3.2.5) a (3.2.6) pro $n = 0, 1, 2, \dots$, musíme nejprve určit počáteční hodnoty t_0 , S_{t_0} a α_{t_0} . Hodnotu S_{t_0} se opět doporučuje volit jako průměr několika prvních pozorování řady y , případně pomocí metody zpětné předpovědi. Dále vezmeme $t_0 = t_1 - q$, kde $q > 0$ je střední časová vzdálenost mezi dvěma sousedními pozorováními řady y . Hodnotu α_{t_0} doporučuje Wright vzít rovnu $1 - \beta^q = 1 - (1 - \alpha)^q$. To odpovídá předpokladu, že fiktivní pozorování řady y od času t_1 do minulosti jsou pozorována s pevným časovým rozestupem q . Tedy

$$\alpha_{t_0} = \left(\sum_{j=0}^{\infty} \beta^q \right)^{-1} = 1 - \beta^q . \quad (3.2.8)$$

Hodnota $\alpha_{t_0} = 1 - \beta^q$ vyhovuje rovnici

$$a = \frac{a}{a + \beta^q}$$

s neznámou a , tedy jde o pevný bod rekurence (3.2.5) s $t_{n+1} - t_n = q$.

Je snadné ověřit, že Wrightovo jednoduché exponenciální vyrovnávání je při použití na pravidelnou časovou řadu totožné s klasickým jednoduchým exponenciálním vyrovnáváním. Pokud jde o předpovědní rozptyly Wrightovy metody, nemůžeme se již přímo opřít o příslušný ARIMA(0, 1, 1) proces jako v případě pravidelné časové řady. Je však možné přijmout takto odvozené předpovědní rozptyly, tedy předpokládat opět

$$\text{var}[e_{t_n+\tau}(t_n)] = \sigma^2 [\alpha^2(\tau - 1) + 1] \quad (3.2.9)$$

pro $\tau \in (0, \infty)$, kde $\sigma^2 > 0$ je parametr určující chybový rozptyl předpovědi o jeden krok vpřed. Je důležité si všimnout, že $\text{var}[e_{t_n+\tau}(t_n)]$ počítaný podle vzorce (3.2.9) je kladný i pro $\tau \in (0, 1)$.

Nevýhodou vzorce (3.2.9) je, že nezohledňuje časovou strukturu řady y do času t_n , vyjádřenou například hodnotou α_{t_n} . Jako rozumné by se jevilo předpokládat, že předpovědní rozptyl $\text{var}[e_{t_n+\tau}(t_n)]$ roste s rostoucí hodnotou α_{t_n} , která značí větší časové rozestupy mezi pozorováními řady y v období před časem t_n . Jednou z možných modifikací vzorce (3.2.9) v tomto směru je například

$$\text{var}[e_{t_n+\tau}(t_n)] = \sigma^2 [\alpha^2(\tau-1) + 1 + (1-\alpha)(\alpha_{t_n} - \alpha)^+], \quad (3.2.10)$$

kde $x^+ = \max(x, 0)$ je kladná část čísla x . Vzorec (3.2.10) splývá se vzorcem (3.2.9) pro $\alpha_{t_n} \leq \alpha$, pro $\alpha_{t_n} \geq \alpha$ je $\text{var}[e_{t_n+\tau}(t_n)]$ lineárně rostoucí funkcí α_{t_n} .

3.3 Nepravidelně pozorovaný ARIMA(0, 1, 1) proces

V tomto odstavci odvodíme alternativní metodu k Wrightovu jednoduchému exponenciálnímu vyrovnávání pro nepravidelné časové řady. Jak bylo řečeno v odstavci 2.3, jednoduché exponenciální vyrovnávání je optimální předpovědní metodou pro časové řady řídicí se modelem ARIMA(0, 1, 1). Vyjdeme tedy z toho, že zkoumaná nepravidelná časová řada je nepravidelně pozorovaným ARIMA(0, 1, 1) procesem a na základě tohoto předpokladu pro ni odvodíme optimální předpovědní metodu. Přestože tato metoda bude tak odvozena pouze pro časové řady s chybějícími pozorováními, půjde ji v praxi použít na libovolnou nepravidelnou časovou řadu.

Nechť $\{y_t, t \in \mathbf{Z}\}$ je časová řada řídicí se modelem ARIMA(0, 1, 1). Řada jejích prvních diferencí $\{\Delta y_t, t \in \mathbf{Z}\}$ se tedy řídí modelem MA(1), což můžeme zapsat například jako

$$\Delta y_t = y_t - y_{t-1} = e_t + (\alpha - 1) \cdot e_{t-1}, \quad (3.3.1)$$

kde $\{e_t, t \in \mathbf{Z}\}$ je bílý šum s rozptylem $\sigma^2 > 0$. Předpokládejme, že $\alpha \in (0, 1)$. Položíme-li pro $t \in \mathbf{Z}$

$$S_t = y_t - (1 - \alpha) \cdot e_t, \quad (3.3.2)$$

můžeme model řady $\{y_t, t \in \mathbf{Z}\}$ zapsat jako

$$y_{t+1} = S_t + e_{t+1}, \quad (3.3.3)$$

$$S_{t+1} = (1 - \alpha) \cdot S_t + \alpha \cdot y_{t+1}. \quad (3.3.4)$$

Z tohoto zápisu je patrná souvislost modelu řady $\{y_t, t \in \mathbf{Z}\}$ s jednoduchým exponenciálním vyrovnáváním s vyrovnávací konstantou α .

Uvažujme nyní rostoucí posloupnost $\{t_j, j \in \mathbf{Z}\}$ představující časovou mřížku, na níž pozorujeme hodnoty řady $\{y_t, t \in \mathbf{Z}\}$. Pozorujeme tedy pouze hodnoty $y_{t_j}, j \in \mathbf{Z}$, zatímco hodnoty y_t pro $t \notin \{t_j, j \in \mathbf{Z}\}$ jsou pro nás nepozorovatelné. Zajímat se budeme o takto vzniklou časovou řadu $\{y_{t_j}, j \in \mathbf{Z}\}$ s chybějícími pozorováními. Konkrétně pro ni budeme hledat optimální předpovědní metodu z hlediska minimalizace rozptylu předpovědní chyby.

Nechť jsme již pozorovali hodnoty y_{t_j} pro $j \leq n$ a na jejich základě získali náhodnou veličinu \tilde{S}_{t_n} představující předpověď neznámé hodnoty S_{t_n} . Je realistické předpokládat, že náhodná veličina $y_{t_n} - \tilde{S}_{t_n}$ má konečný rozptyl a je nekorelovaná s náhodnými veličinami e_t pro $t > t_n$. Stejně vlastnosti pak podle (3.3.2) bude mít i náhodná veličina $S_{t_n} - \tilde{S}_{t_n}$. Označme

$$v_{t_n} = \text{var}(S_{t_n} - \tilde{S}_{t_n}) / \sigma^2 < \infty . \quad (3.3.5)$$

Hledejme nyní předpověď hodnoty $S_{t_{n+1}}$ ve tvaru

$$\tilde{S}_{t_{n+1}} = (1-a) \cdot \tilde{S}_{t_n} + a \cdot y_{t_{n+1}} , \quad (3.3.6)$$

kde $y_{t_{n+1}}$ je nově pozorovaná hodnota řady y . Parametr $a \in \mathbf{R}$ budeme volit s cílem minimalizovat rozptyl

$$\text{var}(S_{t_{n+1}} - \tilde{S}_{t_{n+1}}) = \sigma^2 \cdot v_{t_{n+1}} . \quad (3.3.7)$$

Ze vzorců (3.3.1) a (3.3.2) snadno odvodíme vztah

$$S_{t_{n+1}} = S_{t_n} + \alpha \cdot (e_{t_{n+1}} + e_{t_{n+2}} + \dots + e_{t_{n+1-1}} + e_{t_{n+1}}) \quad (3.3.8)$$

a odtud dosazením do (3.3.6) dostaneme

$$\tilde{S}_{t_{n+1}} = (1-a) \cdot \tilde{S}_{t_n} + a \cdot [S_{t_n} + \alpha \cdot (e_{t_{n+1}} + e_{t_{n+2}} + \dots + e_{t_{n+1-1}}) + e_{t_{n+1}}] . \quad (3.3.9)$$

Odečtením rovností (3.3.8) a (3.3.9) obdržíme

$$S_{t_{n+1}} - \tilde{S}_{t_{n+1}} = (1-a)(S_{t_n} - \tilde{S}_{t_n}) + \alpha(1-a)(e_{t_{n+1}} + e_{t_{n+2}} + \dots + e_{t_{n+1-1}}) + (\alpha - a)e_{t_{n+1}} . \quad (3.3.10)$$

Z nekorelovanosti náhodných veličin $S_{t_n} - \tilde{S}_{t_n}$ a e_t , $t > t_n$ dostáváme

$$\text{var}(S_{t_{n+1}} - \tilde{S}_{t_{n+1}}) = \sigma^2 [(1-a)^2 v_{t_n} + \alpha^2 (1-a)^2 (t_{n+1} - t_n - 1) + (\alpha - a)^2] . \quad (3.3.11)$$

Řešíme tedy úlohu

$$\min_{a \in \mathbf{R}} \{ (1-a)^2 v_{t_n} + \alpha^2 (1-a)^2 (t_{n+1} - t_n - 1) + (\alpha - a)^2 \} . \quad (3.3.12)$$

Jde o minimalizaci konvexní kvadratické funkce proměnné a , bod minima \hat{a} tedy nalezneme velice snadno, konkrétně

$$\hat{a} = \frac{v_{t_n} + \alpha^2 (t_{n+1} - t_n - 1) + \alpha}{v_{t_n} + \alpha^2 (t_{n+1} - t_n - 1) + 1} . \quad (3.3.13)$$

Dosažené minimum má přitom hodnotu

$$v_{t_{n+1}} = \frac{\text{var}(S_{t_{n+1}} - \tilde{S}_{t_{n+1}})}{\sigma^2} = (1 - \hat{a})^2 [v_{t_n} + \alpha^2(t_{n+1} - t_n - 1)] + (\hat{\alpha} - a)^2 . \quad (3.3.14)$$

Všimněme si, že platí $\hat{a} \in (0, 1)$ a vzorec

$$\tilde{S}_{t_{n+1}} = (1 - \hat{a}) \cdot \tilde{S}_{t_n} + \hat{a} \cdot y_{t_{n+1}} \quad (3.3.15)$$

tedy počítá předpověď $\tilde{S}_{t_{n+1}}$ jako konvexní lineární kombinaci stávající předpovědi \tilde{S}_{t_n} a nově pozorované hodnoty $y_{t_{n+1}}$.

Ze vzorce (3.3.13) plyne, že \hat{a} je rostoucí funkcí v_{t_n} , α i $t_{n+1} - t_n$, což odpovídá naší intuitivní představě. Velká hodnota v_{t_n} znamená, že \tilde{S}_{t_n} není kvalitní předpovědí skutečné hodnoty S_{t_n} , a tak je ve vztahu (3.3.15) větší váha kladena na nové pozorování $y_{t_{n+1}}$. Podobně větší časový krok $t_{n+1} - t_n$ znamená, že hodnota S_{t_n} , jejíž je \tilde{S}_{t_n} předpovědí, je již dále v minulosti od nového pozorování $y_{t_{n+1}}$. Parametr α pak představuje vyrovnávací konstantu při použití jednoduchého exponenciálního vyrovnávání na řadu $\{y_t, t \in \mathbf{Z}\}$, jeho vztah k \hat{a} tudíž také nepřekvapí. Pro $\alpha \rightarrow 1$, $t_{n+1} - t_n \rightarrow \infty$ či $v_{t_n} \rightarrow \infty$ je $\hat{a} \rightarrow 1$. Je-li $v_{t_n} = 0$ a $t_{n+1} - t_n = 1$, což odpovídá pravidelné časové řadě $\{y_{t_j}, j \in \mathbf{Z}\}$ s $t_j = j$, pak $\hat{a} = \alpha$.

Odvozená metoda se skládá ze vzorce (3.3.13) pro výpočet optimálního vyrovnávacího koeficientu \hat{a} v daném kroku, rekurentního vzorce (3.3.15) pro aktualizaci vyrovnané hodnoty \tilde{S} a rekurentního vzorce (3.3.14) pro aktualizaci rozptylového faktoru v . Předpovědi z času t_n o $\tau > 0$ časových jednotek vpřed budoucí neznámé hodnoty $y_{t_n+\tau}$ je vyrovnaná hodnota $\hat{y}_{t_n} = \tilde{S}_{t_n}$ podobně jako u jednoduchého exponenciálního vyrovnávání. Pro rozptyl chyby $e_{t_n+\tau}(t_n) = y_{t_n+\tau} - \tilde{S}_{t_n}$ této předpovědi platí

$$\begin{aligned} \text{var}[e_{t_n+\tau}(t_n)] &= \text{var}[S_{t_n} - \tilde{S}_{t_n} + \alpha \cdot (e_{t_{n+1}} + e_{t_{n+2}} + \dots + e_{t_{n+\tau-1}}) + e_{t_n+\tau}] = \\ &= \sigma^2 [v_{t_n} + \alpha^2(\tau - 1) + 1] . \end{aligned} \quad (3.3.16)$$

vzhledem k nekorelovanosti veličin $S_{t_n} - \tilde{S}_{t_n}$ a e_t pro $t > t_n$. Je vidět, že tento rozptyl je minimální právě tehdy, když je minimální hodnota v_{t_n} . Tedy odvozená hodnota \hat{a} je optimální vyrovnávací konstantou i z hlediska minimalizace předpovědní chyby.

Pokud budeme předpokládat, že $e_t \sim N(0, \sigma^2)$ a $S_{t_n} - \tilde{S}_{t_n} \sim N(0, \sigma^2 v_{t_n})$, pak

$$e_{n+\tau}(n) \sim N(0, \sigma^2 [v_{t_n} + \alpha^2(\tau - 1) + 1]) \quad (3.3.17)$$

a příslušný předpovědní interval se spolehlivostí $1 - \theta$ má meze

$$\tilde{S}_{t_n} \pm u_{1-\theta/2} \cdot \sigma \sqrt{v_{t_n} + \alpha^2(\tau - 1) + 1} . \quad (3.3.18)$$

Všimněme si, že jakmile je náhodná veličina $S_{t_m} - \tilde{S}_{t_m}$ nekorelovaná s veličinami e_t pro $t > t_n$, pak již vzhledem ke vzorci (3.3.10) a nekorelovanosti $\{e_t, t \in \mathbf{Z}\}$ platí stejný fakt i pro každé $m \geq n$. Pokud je $e_t \sim N(0, \sigma^2)$, pak podobně z normality náhodné veličiny $S_{t_m} - \tilde{S}_{t_m}$ již plyne normalita $S_{t_m} - \tilde{S}_{t_m}$ pro všechna $m \geq n$.

Máme-li k dispozici pozorování řady y počínaje časem t_1 a chceme-li nastartovat rekurentní výpočty podle vzorců (3.3.13) až (3.3.15) pro $n = 0, 1, 2, \dots$, musíme nejprve určit počáteční hodnoty t_0 , \tilde{S}_{t_0} a v_{t_0} . Stejně jako Wright označme jako q průměrný časový rozestup mezi sousedními pozorováními řady $\{y_{t_j}, j \in \mathbf{Z}\}$ a položme $t_0 = t_1 - q$. Hodnotu \tilde{S}_{t_0} určíme opět jako vážený průměr několika prvních pozorování řady y , která máme k dispozici, s tím, že váhy budou klesat do budoucnosti s diskontním faktorem $\beta = 1 - \alpha$. Hodnotu v_{t_0} zvolíme jako pevný bod rekurence (3.3.14) pro $t_{n+1} - t_n = q$, tedy jako složku v řešení soustavy rovnic

$$\begin{aligned} a &= \frac{v + \alpha^2(q-1) + \alpha}{v + \alpha^2(q-1) + 1}, \\ v &= (1-a)^2[v + \alpha^2(q-1)] + (\alpha - a)^2. \end{aligned} \quad (3.3.19)$$

s neznámými v a a . Po formálních algebraických úpravách takto získáme

$$v_{t_0} = \tilde{a} = \frac{(1-\tilde{a})^2 \alpha^2 (q-1) + (\tilde{a} - \alpha)^2}{\tilde{a}(2-\tilde{a})}, \quad (3.3.20)$$

kde hodnota $\tilde{a} \in (0, 1)$ je určena vzorcem

$$\tilde{a} = \frac{\alpha^2 q - \sqrt{\alpha^4 q^2 + 4(1-\alpha)\alpha^2 q}}{2(\alpha-1)}. \quad (3.3.21)$$

Výše odvozená metoda pro vyrovnávání a předpovídání v časových řadách s chybějícími pozorováními má podobný charakter jako Wrightovo jednoduché exponenciální vyrovnávání. I zde je pomocí rekurentní formule typického tvaru, viz. vzorec (3.3.15), přepočítávána vyrovnaná hodnota řady. Přitom vyrovnávací koeficient \hat{a} se mění krok od kroku a za tímto účelem je nutné rekurentně přepočítávat vedle vyrovnané hodnoty \tilde{S}_{t_j} ještě statistiku v_{t_j} .

Přesto, že tato metoda byla odvozena pouze pro časové řady s chybějícími pozorováními, lze jí v praxi stejně tak dobře použít na jakoukoli časovou řadu s nepravidelně pozorovanými hodnotami. Fakt, že v tomto případě již časový krok $t_{n+1} - t_n$ nebude obecně přirozené číslo, nijak neznemožňuje použití vzorců (3.3.13) až (3.3.15).

4 Holtova metoda

Tato kapitola bude věnována Holtově metodě a jejímu zobecnění pro nepravidelné časové řady. V odstavci 4.1 bude nejprve prezentována klasická verze této metody pro pravidelné časové řady a její vztah k modelu ARIMA(0, 2, 2). Odtud budou odvozeny předpovědní rozptyly a předpovědní intervaly pro tuto metodu. Zmíněna bude též Holtova metoda s exponenciálním trendem a Holtova metoda s tlumeným lineárním trendem. V odstavci 4.2 bude popsáno Wrightovo zobecnění Holtovy metody pro nepravidelné časové řady, včetně analogických zobecnění Holtovy metody s exponenciálním a tlumeným lineárním trendem. Odstavec 4.3 bude věnována Holt-Wintersově metodě a jejímu zobecnění pro časové řady s chybějícími pozorováními.

4.1 Holtova metoda pro pravidelné časové řady

Holtova metoda je další velice známou a používanou variantou exponenciálního vyrovnávání. Vhodná je pro aplikaci na nesezónní časové řady s *lokálně lineárním* trendem. V praxi lze Holtovu metodu nebo některou její modifikaci úspěšně použít na většinu nesezónních časových řad. Při použití na sezónní časovou řadu nelze však očekávat dobré výsledky, jelikož nebude brán zřetel na přítomnost sezónnosti. Pro sezónní časové řady je určena Holt-Wintersova metoda, viz. odstavec 4.3.

Holtova metoda používá rekurentní formule vycházející z myšlenky jednoduchého exponenciálního vyrovnávání k odhadu nejen úrovně, ale i směrnice trendu zkoumané časové řady. Její vzorce jsou stále velice názorné a výpočetně nenáročné.

Vycházíme z toho, že zkoumaná časová řada $\dots y_{n-1}, y_n, y_{n+1}, \dots$ vykazuje lineární trend, jehož směrnice se ale může v čase pozvolna měnit. Stav časové řady y v daném okamžiku n je tak charakterizován jednak její *úrovní* S_n a také *směrnici trendu* T_n . Úroveň S_n představuje vyrovnanou hodnotu řady y v čase n , tedy

$$\hat{y}_n = S_n . \quad (4.1.1)$$

Směrnice trendu T_n udává aktuální směrnici lineárního trendu řady y v čase n , tedy očekávanou změnu její úrovně vztahenou k jedné časové jednotce. Předpověď o $\tau > 0$ časových jednotek vpřed budoucí neznámé hodnoty $y_{n+\tau}$ z času n má tak následující intuitivní tvar:

$$\hat{y}_{n+\tau}(n) = S_n + \tau \cdot T_n . \quad (4.1.2)$$

Bodové předpovědi z času n pro různé předpovědní horizonty $\tau > 0$ tedy tvoří přímkou procházející vyrovnanou hodnotou $\hat{y}_n = S_n$ a mající za směrnici hodnotu T_n směrnice trendu řady y v čase n .

Jádrem Holtovy metody jsou rekurentní formule, pomocí nichž ze stávajících hodnot S_n a T_n a nově pozorované hodnoty y_{n+1} získáme úroveň S_{n+1} a směrnici trendu T_{n+1} řady y v následujícím čase $n+1$. Naší předpovědi hodnoty y_{n+1} z času n je výraz $S_n + T_n$, jak plyne ze vzorce (4.1.2) pro $\tau = 1$. Tedy je přirozené očekávat $S_{n+1} \approx S_n + T_n$. Na druhou stranu jak úroveň tak směrnice trendu řady y se může v čase měnit, signálem této změny pro nás může být pozorovaná hodnota y_{n+1} . Jelikož též $S_{n+1} \approx y_{n+1}$, jeví se rozumným volit hodnotu S_{n+1} jako konvexní lineární kombinaci hodnot $S_n + T_n$ a y_{n+1} . Zvolme tedy pevně číslo $\alpha \in (0, 1)$ a poloźme

$$S_{n+1} = (1 - \alpha)(S_n + T_n) + \alpha \cdot y_{n+1} . \quad (4.1.3)$$

Hodnota α je *vyrovnávací konstanta pro úroveň*. Vzorec (4.1.3) je obdobou vzorce (3.1.1) z jednoduchého exponenciálního vyrovnávání. Rozdíl je pouze v jiném tvaru předpovědi $\hat{y}_{n+1}(n)$, která je zde kombinována s nově pozorovanou hodnotou y_{n+1} . Také vzorec (4.1.3) lze přepsat do jeho chybového tvaru

$$S_{n+1} = S_n + T_n + \alpha \cdot e_{n+1} , \quad (4.1.4)$$

kde $e_{n+1} = y_{n+1} - \hat{y}_{n+1}(n)$ je chyba jednokrokové předpovědi.

Podobnou úvahou dojdeme i ke vzorci pro aktualizaci směrnice T trendu řady y . O té předpokládáme, že má lokálně *lineární* trend, lze tedy očekávat $T_{n+1} \approx T_n$. Taktéž však má smysl očekávat $T_{n+1} \approx S_{n+1} - S_n$. Zvolme tedy pevně číslo $\gamma \in (0, 1)$ a poloźme

$$T_{n+1} = (1 - \gamma) \cdot T_n + \gamma \cdot (S_{n+1} - S_n) . \quad (4.1.5)$$

Hodnota γ je *vyrovnávací konstanta pro směrnici trendu*, přitom může být obecně $\gamma \neq \alpha$. I vzorec (4.1.5) lze přepsat do jeho chybového tvaru:

$$T_{n+1} = T_n + \gamma\alpha \cdot e_{n+1} . \quad (4.1.6)$$

Vzorce (4.1.3) a (4.1.5) jsou podstatou Holtovy metody.

Chybové tvary (4.1.4) a (4.1.6) rekurentních vzorců (4.1.3) a (4.1.5) jsou opět velice názorné. Parametr $\alpha \in (0, 1)$ udává, jak moc je předpovědní chyba e_{n+1} zahrnuta do nové hodnoty úrovně S_{n+1} . Hodnota $\gamma\alpha$ pak udává, do jaké míry je chyba e_{n+1} zahrnuta do nové hodnoty T_{n+1} směrnice trendu. Jinak řečeno, volbou hodnot parametrů α a γ se řídí rozložení vah ve vzorcích (4.1.3) a (4.1.5) mezi stávající hodnotu (či předpověď ze stávajících hodnot) a hodnotu založenou na novém pozorování y_{n+1} .

Volba vyrovnávacích konstant $\alpha, \gamma \in (0, 1)$ se v praxi provádí stejně jako volba $\alpha \in (0, 1)$ u jednoduchého exponenciálního vyrovnávání, tedy buď expertně nebo

minimalizací např. MSE přes jistý úsek zkoumané časové řady. Numerická minimalizace MSE či jiného kritéria je zde pochopitelně výpočetně náročnější, protože probíhá na celém jednotkovém čtverci $(0, 1) \times (0, 1)$.

Pro praktické použití metody na řadu $y_1, y_2, \dots, y_n, \dots$ je nutné nějak zvolit počáteční hodnoty S_0 a T_0 , které jsou nezbytné k nastartování rekurentního výpočtu podle vzorců (4.1.3) a (4.1.5). Nejjednodušší a v praxi často používanou volbou je

$$T_0 = y_2 - y_1, \quad (4.1.7)$$

$$S_0 = y_1 - T_0 = 2y_1 - y_2. \quad (4.1.8)$$

Jinou možností je vzít $T_0 = \hat{a}$ a $S_0 = \hat{b}$, kde \hat{a} a \hat{b} jsou regresní odhady parametrů a a b v modelu $y_t \approx a \cdot t + b$ pro několik prvních pozorování řady y . Obvykle se zde používají klasické OLS (*ordinary least squares*) odhady. Myšlenka zpětné předpovědi však lépe odpovídá použití časově invertovaných DLS odhadů, kde váhy klesají směrem *do budoucnosti* s diskontním faktorem $\beta = 1 - \sqrt{\alpha \cdot \gamma}$. Odůvodnění volby diskontního faktoru spočívá ve vztahu Holtovy metody k dvojitému exponenciálnímu vyrovnávání, viz. vzorec (5.1.14) v odstavci 5.1.

Holtova metoda je typickou *ad hoc* vyrovnávací a předpovědní metodou. Hlavní ospravedlnění pramení z jejího úspěšného fungování v praxi. Při velice názorné a výpočetně nenáročné podobě mají získané předpovědi uspokojivou přesnost. Podobně jako u jednoduchého exponenciálního vyrovnávání byly až dodatečně dohledány pravděpodobnostní modely časových řad, pro něž je Holtova metoda optimální předpovědní metodou. V následujících odstavcích ukážeme souvislost Holtovy metody s modelem ARIMA(0, 2, 2) a na základě této souvislosti odvodíme pro Holtovu metodu rozptyly předpovědních chyb a meze předpovědních intervalů.

Podobně jako v případě jednoduchého exponenciálního vyrovnávání budeme předpokládat, že jednokrokové předpovědní chyby $\{e_n, n \in \mathbf{Z}\}$ Holtovy metody, definované jako $e_n = y_n - \hat{y}_n(n-1)$, tvoří bílý šum s rozptylem $\sigma^2 > 0$. Vyjdeme ze vzorců

$$S_{n+1} = S_n + T_n + \alpha \cdot e_{n+1}, \quad (4.1.9)$$

$$T_{n+1} = T_n + \gamma \alpha \cdot e_{n+1}, \quad (4.1.10)$$

$$y_{n+1} = S_n + T_n + e_{n+1} \quad (4.1.11)$$

a ukážeme, že řada $\{y_n, n \in \mathbf{Z}\}$ se řídí modelem ARIMA(0, 2, 2), tedy řada $\{\Delta^2 y_n, n \in \mathbf{Z}\}$ jejích druhých diferencí se řídí modelem MA(2). Počítejme

$$\begin{aligned} \Delta y_n &= y_n - y_{n-1} = (S_{n-1} + T_{n-1} + e_n) - (S_{n-2} + T_{n-2} + e_{n-1}) = \\ &= (S_{n-1} - S_{n-2}) + (T_{n-1} - T_{n-2}) + e_n - e_{n-1} = \\ &= T_{n-2} + \alpha \cdot e_{n-1} + \gamma \alpha \cdot e_{n-1} + e_n - e_{n-1} = T_{n-2} + (\alpha + \gamma \alpha - 1)e_{n-1} + e_n. \end{aligned} \quad (4.1.12)$$

Dalším diferencováním obdržíme konečně

$$\begin{aligned}
\Delta^2 y_n &= \Delta y_n - \Delta y_{n-1} = \\
&= [T_{n-2} + (\alpha + \gamma\alpha - 1)e_{n-1} + e_n] - [T_{n-3} + (\alpha + \gamma\alpha - 1)e_{n-2} + e_{n-1}] = \\
&= e_n + (\alpha + \gamma\alpha - 2) \cdot e_{n-1} + (1 - \alpha) \cdot e_{n-2} .
\end{aligned} \tag{4.1.13}$$

Tedy řada $\{\Delta^2 y_n, n \in \mathbf{Z}\}$ se skutečně řídí modelem MA(2), kde jeho parametry jsou

$$\begin{aligned}
v_1 &= \alpha + \gamma\alpha - 2 , \\
v_2 &= 1 - \alpha .
\end{aligned} \tag{4.1.14}$$

Ty můžeme odhadnout např. momentovým odhadem, založeným na hodnotách r_1 a r_2 první a druhé výběrové autokorelace řady $\{\Delta^2 y_n, n \in \mathbf{Z}\}$. Dosazením těchto odhadů \hat{v}_1 a \hat{v}_2 do vztahů (4.1.14) a vyjádřením α a γ získáme odhady $\hat{\alpha}$ a $\hat{\gamma}$ vyrovnávacích konstant pro Holtovu metodu. Tyto odhady nejsou však ani při velkém počtu pozorování řady y příliš přesné.

Je známo, viz. Chatfield (2000), str. 99 nebo Brockwell a Davis (2002), str. 325, že Holtova metoda je optimální předpovědní metodou pro výše specifikovaný ARIMA(0, 2, 2) proces. Je tedy možné při použití Holtovy metody předpokládat, že zkoumaná časová řada se skutečně řídí tímto modelem. Za předpokladu, že známe skutečné hodnoty parametrů α a γ , můžeme snadno odvodit rozptyly předpovědních chyb a za předpokladu normality i meze předpovědních intervalů pro Holtovu metodu.

Zajímáme se o rozptyl chyby $e_{n+\tau}(n) = y_{n+\tau} - \hat{y}_{n+\tau}(n)$ předpovědi $\hat{y}_{n+\tau}(n) = S_n + \tau \cdot T_n$ z času n o $\tau > 0$ kroků vpřed. Opakovaným použitím vzorců (4.1.4) a (4.1.6) dostaneme

$$y_{n+\tau} = S_n + \tau \cdot T_n + \alpha \cdot \sum_{i=1}^{\tau-1} e_{n+i} [\gamma(\tau - i) + 1] + e_{n+\tau} . \tag{4.1.15}$$

Odtud pak

$$e_{n+\tau}(n) = \alpha \cdot \sum_{i=1}^{\tau-1} e_{n+i} [\gamma(\tau - i) + 1] + e_{n+\tau} \tag{4.1.16}$$

a z předpokládané nekorelovanosti chyb $\{e_n, n \in \mathbf{Z}\}$ nakonec

$$\text{var}[e_{n+\tau}(n)] = \sigma^2 \cdot \left\{ \alpha^2 \cdot \sum_{i=1}^{\tau-1} [\gamma(\tau - i) + 1]^2 + 1 \right\} . \tag{4.1.17}$$

Jednoduchými algebraickými úpravami tohoto výrazu za použití známých vzorců

$$\sum_{i=1}^N i = \frac{N(N+1)}{2} \quad \text{a} \quad \sum_{i=1}^N i^2 = \frac{N(N+1)(2N+1)}{6} , \tag{4.1.18}$$

viz. Bartsch (2000), str. 168 a 169, dospějeme k následujícímu vzorci pro rozptyl předpovědní chyby:

$$\text{var}[e_{n+\tau}(n)] = \sigma^2 \left\{ \alpha^2 (\tau - 1) \left[1 + \gamma\tau + \gamma^2 \frac{\tau(2\tau - 1)}{6} \right] + 1 \right\}. \quad (4.1.19)$$

Tento rozptyl je úměrný parametru σ^2 , pro $\tau = 1$ je roven σ^2 . Jde o rostoucí funkci předpovědního horizontu τ i obou vyrovnávacích konstant α a γ . Pro $\tau \rightarrow \infty$ je

$$\text{var}[e_{n+\tau}(n)] \approx \sigma^2 \cdot \frac{\alpha^2 \gamma^2}{3} \cdot \tau^3. \quad (4.1.20)$$

Doposud jsme se zabývali původní Holtovou metodou pro lokálně lineární trend, jak ji Holt navrhl v roce 1957. Existuje však několik modifikací Holtovy metody, které využívají stejné struktury základních vzorců, ale jsou uzpůsobeny pro jiné než lineární trendy. Jednou z takových modifikací je Holtova metoda s tzv. *tlumeným lineárním trendem*, viz. Gardner a McKenzie (1985). Ta je motivována empirickým poznatkem o tom, že současný trend v časové řadě většinou nepřetrvává dlouhodobě, ale má tendenci se spíše tlumit. Z hlediska přesnosti dlouhodobějších předpovědí se pak jeví výhodné tento poznatek zohlednit v použité předpovědní metodě.

Bodové předpovědi klasické Holtovy metody s lineárním trendem tvoří přímku, tedy předpokládá se konstantní nárůst úrovně řady za časovou jednotku. Oproti tomu *tlumený lineární trend* je takový, kde přírůstek úrovně řady za časovou jednotku v průběhu času geometricky klesá s pevným kvocientem $\varphi \in (0, 1)$. Tomuto číslu se říká *tlumicí konstanta* a určuje rychlost, s jakou se trend tlumí: nižší resp. vyšší hodnota φ znamená rychlejší resp. pomalejší tlumení trendu. Příklad $\varphi = 1$ by znamenal žádné tlumení, tedy klasický lineární trend.

Předpovědní formule Holtovy metody s tlumeným lineárním trendem má tvar

$$\hat{y}_{n+\tau}(n) = S_n + \varphi \cdot T_n + \varphi^2 \cdot T_n + \dots + \varphi^\tau \cdot T_n. \quad (4.1.21)$$

Rekurentní vzorce pro výpočet úrovně S a směrnice trendu T řady y jsou nyní

$$S_{n+1} = (1 - \alpha)(S_n + \varphi \cdot T_n) + \alpha \cdot y_{n+1}, \quad (4.1.22)$$

$$T_{n+1} = (1 - \gamma) \cdot \varphi \cdot T_n + \gamma \cdot (S_{n+1} - S_n) \quad (4.1.23)$$

a jejich chybové tvary mají názornou podobu

$$S_{n+1} = S_n + \varphi \cdot T_n + \alpha \cdot e_{n+1}, \quad (4.1.24)$$

$$T_{n+1} = \varphi \cdot T_n + \alpha \gamma \cdot e_{n+1}. \quad (4.1.25)$$

Dodatečný parametr $\varphi \in (0, 1)$ se volí buď expertně (většinou blízko 1) nebo stejně jako vyrovnávací konstanty minimalizací jisté míry nepřesnosti předpovědí. Počáteční hodnoty S_0 a T_0 se obvykle volí stejně jako u klasické Holtovy metody, nezávisle na hodnotě φ . Vzorce (4.1.7) a (4.1.8) by bylo možné upravit do podoby

$$T_0 = \frac{y_2 - y_1}{\varphi^2}, \quad (4.1.26)$$

$$S_0 = y_1 - \varphi \cdot T_0 = y_1 - \frac{y_2 - y_1}{\varphi}. \quad (4.1.27)$$

V případě použití regresních odhadů \hat{a} a \hat{b} jako u klasické Holtovy metody by bylo možné původně lineární trend $y_t \approx a \cdot t + b$ nahradit tlumeným lineárním trendem $y_t \approx a \cdot g(t) + b$, kde $g(t) = \varphi + \varphi^2 + \dots + \varphi^t$.

Za předpokladu, že předpovědní chyby $e_n = y_n - \hat{y}_n(n-1)$ v Holtově metodě s tlumeným lineárním trendem tvoří bílý šum s rozptylem $\sigma^2 > 0$, lze snadnými algebraickými úpravami odvodit, že zkoumaná časová řada $\{y_n, n \in \mathbf{Z}\}$ se řídí modelem ARIMA(1, 1, 2), kde tlumicí konstanta φ představuje autoregresní parametr. Přesně

$$\Delta y_n = \varphi \cdot \Delta y_{n-1} + \varphi(1 - \alpha)e_{n-2} + (\alpha + \varphi\alpha\gamma - 2)e_{n-1} + e_n. \quad (4.1.28)$$

Pokud jde o rozptyly předpovědních chyb, lze postupovat obdobně jako v případě klasické Holtovy metody:

$$\hat{y}_{n+\tau} = S_n + g(\tau) \cdot T_n, \quad (4.1.29)$$

$$y_{n+\tau} = S_n + g(\tau) \cdot T_n + \alpha \cdot \sum_{i=1}^{\tau-1} e_{n+i} [\gamma \cdot g(\tau - i) + 1] + e_{n+\tau}, \quad (4.1.30)$$

kde $g(t) = \varphi + \varphi^2 + \dots + \varphi^t$. Odečtením těchto dvou rovností získáme

$$e_{n+\tau}(n) = \alpha \sum_{i=1}^{\tau-1} e_{n+i} [\gamma \cdot g(\tau - i) + 1] + e_{n+\tau} \quad (4.1.31)$$

a z předpokládané nekorelovanosti chyb $\{e_n, n \in \mathbf{Z}\}$ nakonec

$$\text{var}[e_{n+\tau}(n)] = \sigma^2 \cdot \left\{ \alpha^2 \cdot \sum_{i=1}^{\tau-1} [\gamma \cdot g(i) + 1]^2 + 1 \right\}. \quad (4.1.32)$$

Další algebraické úpravy využívající vzorce

$$g(t) = \varphi + \varphi^2 + \dots + \varphi^t = \varphi \frac{1 - \varphi^t}{1 - \varphi} \quad (4.1.33)$$

vedou již na nepříliš přehledný vzorec, konkrétně

$$\text{var}[e_{n+\tau}(n)] = \sigma^2 \cdot \left\{ \alpha^2 \cdot \left(\tau - 1 + 2\gamma \cdot \sum_{i=1}^{\tau-1} g(i) + \gamma^2 \cdot \sum_{i=1}^{\tau-1} g^2(i) \right) + 1 \right\}, \quad (4.1.34)$$

kde

$$\sum_{i=1}^{\tau-1} g(i) = \frac{\varphi}{1-\varphi} \left(\tau - 1 - \frac{1-\varphi^{\tau-1}}{1-\varphi} \right), \quad (4.1.35)$$

$$\sum_{i=1}^{\tau-1} g^2(i) = \frac{\varphi^2}{(1-\varphi)^2} \left(\tau - 1 - 2\varphi \frac{1-\varphi^{\tau-1}}{1-\varphi} + \frac{1-\varphi^{2\tau-2}}{1-\varphi^2} \right). \quad (4.1.36)$$

Jinou modifikací Holtovy metody je tzv. Holtova metoda s *exponenciálním trendem*. Myšlenka i struktura vzorců metody je zde shodná jako v případě klasické Holtovy metody s lineárním trendem. Jediný rozdíl je, že se předpokládá exponenciální namísto lineárního trendu. Uvažuje se tedy situace, kdy úroveň zkoumané časové řady vzroste (klesne) každou časovou jednotku jistým násobkem. Toto je v praxi velmi častý jev ať už v ekonomických časových řadách, nebo v těch pocházejících z přírodních věd. Exponenciální trend má smysl pouze u časových řad s kladnými pozorováními.

Předpovědní vzorec, obdoba vzorců (4.1.2) a (4.1.21), vypadá nyní takto

$$\hat{y}_{n+\tau}(n) = S_n \cdot T_n^\tau \quad (4.1.37)$$

a příslušné rekurentní formule mají tvar

$$S_{n+1} = (1-\alpha)(S_n \cdot T_n) + \alpha \cdot y_{n+1}, \quad (4.1.38)$$

$$T_{n+1} = (1-\alpha)T_n + \alpha \cdot S_{n+1}/S_n. \quad (4.1.39)$$

Počáteční hodnoty S_0 a T_0 se volí obdobně jako v případě klasické Holtovy metody:

$$T_0 = y_2/y_1 \quad \text{a} \quad S_0 = y_1/T_0 = y_1^2/y_2 \quad (4.1.40)$$

nebo jako $T_0 = \hat{a}$ a $S_0 = \hat{b}$, kde \hat{a} a \hat{b} jsou odhady parametrů a a b v regresním modelu $y_t \approx a \cdot b^t$. Ty se běžně volí jako exponenciální funkce OLS odhadů v logaritmovaném modelu $\log y_t \approx \log a + t \cdot \log b$. Pokud zde namísto OLS odhadů budeme používat časově invertované DLS odhady, diskontní faktor, snižující váhy směrem *do budoucnosti*, budeme opět brát jako $\beta = 1 - \sqrt{\alpha \cdot \gamma}$.

4.2 Wrightova modifikace pro nepravidelné časové řady

V tomto odstavci bude prezentováno Wrightovo zobecnění Holtovy metody pro nepravidelné časové řady, viz. Wright (1986). Stejně jako v případě jednoduchého exponenciálního vyrovnávání od stejného autora půjde o přímočarou adaptaci vzorců Holtovy metody na časovou nepravidelnost ve zkoumané časové řadě a opět při tom bude zachována názornost a výpočetní nenáročnost původní metody. Na závěr ještě zmíníme analogická zobecnění Holtovy metody s tlumeným lineárním a exponenciálním trendem pro nepravidelné časové řady, viz. Cipra (2006).

Mějme nepravidelnou časovou řadu $\dots y_{t_{n-1}}, y_{t_n}, y_{t_{n+1}}, \dots$ s lokálně lineárním trendem. Uvažujme opět úroveň S_{t_n} a směrnici trendu T_{t_n} řady y v čase t_n . Předpověď z času t_n o $\tau > 0$ časových jednotek vpřed bude tedy

$$\hat{y}_{t_n+\tau}(t_n) = S_{t_n} + \tau \cdot T_{t_n} . \quad (4.2.1)$$

Rekurentní vzorce pro výpočet nových hodnot $S_{t_{n+1}}$ a $T_{t_{n+1}}$ musí nově zohledňovat délku časového kroku $t_{n+1} - t_n$ mezi příslušnými dvěma pozorováními časové řady y v tom smyslu, že nyní

$$S_{t_{n+1}} \approx S_{t_n} + (t_{n+1} - t_n)T_{t_n} \quad \text{a} \quad T_{t_{n+1}} \approx \frac{S_{t_{n+1}} - S_{t_n}}{t_{n+1} - t_n} . \quad (4.2.2)$$

Opět však současně $S_{t_{n+1}} \approx y_{t_{n+1}}$ a $T_{t_{n+1}} \approx T_{t_n}$, což vede ke konvexním lineárním kombinacím podobným těm ve vzorcích (4.1.3) a (4.1.5). Uvažujme opět dvě různé vyrovnávací konstanty $\alpha, \gamma \in (0, 1)$, jednu pro vyrovnávání úrovně S a druhou pro vyrovnávání směrnice trendu T . Bude zde ovšem aplikována myšlenka proměnných vyrovnávacích koeficientů podobně jako u Wrightova jednoduchého exponenciálního vyrovnávání:

$$S_{t_{n+1}} = (1 - \alpha_{t_{n+1}}) \cdot [S_{t_n} + (t_{n+1} - t_n)T_{t_n}] + \alpha_{t_{n+1}} \cdot y_{t_{n+1}} , \quad (4.2.3)$$

$$T_{t_{n+1}} = (1 - \gamma_{t_{n+1}}) \cdot T_{t_n} + \gamma_{t_{n+1}} \cdot \frac{S_{t_{n+1}} - S_{t_n}}{t_{n+1} - t_n} , \quad (4.2.4)$$

kde vyrovnávací koeficienty $\alpha_{t_{n+1}}$ a $\gamma_{t_{n+1}}$ jsou přepočítávány rekurentními vzorci

$$\alpha_{t_{n+1}} = \frac{\alpha_{t_n}}{\alpha_{t_n} + (1 - \alpha)^{t_{n+1} - t_n}} \quad \text{a} \quad \gamma_{t_{n+1}} = \frac{\gamma_{t_n}}{\gamma_{t_n} + (1 - \gamma)^{t_{n+1} - t_n}} . \quad (4.2.5)$$

Na rozdíl od Holtovy metody pro pravidelné časové řady tedy musíme kromě úrovně S_{t_n} a směrnice trendu T_{t_n} přepočítávat ještě dva proměnné vyrovnávací koeficienty α_{t_n} a γ_{t_n} . Jejich význam ve vzorcích (4.2.3) a (4.2.4) je zřejmý a stejný jako v pravidelném případě. Tvar vzorců (4.2.5), podle nichž se vyrovnávací koeficienty rekurentně přepočítávají, je stejný jako v případě jednoduchého exponenciálního vyrovnávání, viz. vzorec (3.2.5), a nevyžaduje tudíž žádný komentář.

Podívejme se ještě na vzorec (4.2.4), kde se ve jmenovateli vyskytuje délka časového kroku $t_{n+1} - t_n$. Pokud jsou si dvě sousední pozorování časové řady velice blízká v čase, ale nikoli z hlediska svých hodnot, může zlomek $(S_{t_{n+1}} - S_{t_n}) / (t_{n+1} - t_n)$ v tomto vzorci nabývat hodnot řádově vyšších než předchozí směrnice T_{t_n} . Takto pak může dojít k nežádoucímu vychýlení nové hodnoty $T_{t_{n+1}}$, což má fatální důsledky na následující předpovědi. Míra tohoto efektu závisí jak na velikosti směrnice

$(y_{t_{n+1}} - y_{t_n}) / (t_{n+1} - t_n)$, tak na předchozí časové struktuře řady a na hodnotách vyrovnávacích konstant.

Tento problém se pochopitelně netýká pravidelných časových řad či řad s chybějícími hodnotami, kde nikdy nemůže nastat $t_{n+1} - t_n \ll q$. Taktéž se nebude vyskytovat u metod pro řady s lokálně lineárním trendem z kapitol 5.2 a 5.3. Jako řešení se nabízí buď jedno ze dvou velice blízkých pozorování vyřadit nebo je sloučit do jednoho nového pozorování s časem t a hodnotou y zkonstruovanými jako aritmetický průměr příslušných dvou hodnot.

Volba počátečních hodnot t_0 , S_{t_0} , T_{t_0} , α_{t_0} a γ_{t_0} se provádí podobným způsobem jako u klasické Holtovy metody resp. u jednoduchého exponenciálního vyrovnávání pro nepravidelné časové řady. Opět označme jako q střední časovou vzdálenost mezi dvěma sousedními pozorováními řady y a vezměme $t_0 = t_1 - q$. Hodnoty S_{t_0} a T_{t_0} můžeme buď zvolit jako

$$T_{t_0} = \frac{y_{t_2} - y_{t_1}}{t_2 - t_1} \quad \text{a} \quad S_{t_0} = y_{t_1} - q \cdot T_{t_0} \quad (4.2.6)$$

nebo stejně jako v pravidelném případě rovny příslušným regresním odhadům parametrů lineárního trendu v počátečním úseku řady y . Pokud k odhadu těchto parametrů volíme DLS metodu, měly by váhy (klesající do budoucnosti) zohlednit nepravidelnou časovou strukturu řady y . Diskontní faktor volíme opět $\beta = 1 - \sqrt{\alpha \cdot \gamma}$.

Hodnoty α_{t_0} a γ_{t_0} doporučuje Wright vzít rovny $1 - (1 - \alpha)^q$ a $1 - (1 - \gamma)^q$, což opět odpovídá předpokladu, že fiktivní pozorování řady y od času t_1 do minulosti jsou pozorována s pevným časovým rozestupem q . Dodejme, že takto volené hodnoty α_{t_0} a γ_{t_0} jsou opět pevnými body příslušných rekurencí s $t_{n+1} - t_n = q$.

Je-li $\alpha = \gamma$, pak též $\alpha_{t_n} = \gamma_{t_n}$ pro všechna n a paměťové i výpočetní nároky metody se tak ještě nepatrně sníží. Obecně je pochopitelně $\alpha \neq \gamma$ a praxe ukazuje, že omezením se na případ $\alpha = \gamma$ ztrácí metoda na flexibilitě při jejím používání. Konkrétní hodnoty vyrovnávacích konstant $\alpha, \gamma \in (0, 1)$ se volí stejně jako v pravidelném případě.

Pokud jde o předpovědní rozptyly Wrightovy modifikace Holtovy metody, nemůžeme se již přímo opřít o příslušný ARIMA(0, 2, 2) proces jako v případě pravidelné časové řady. Je však možné přijmout takto odvozené předpovědní rozptyly, tedy předpokládat opět

$$\text{var}[e_{t_n+\tau}(t_n)] = \sigma^2 \left\{ \alpha^2 (\tau - 1) \left[1 + \gamma \tau + \gamma^2 \frac{\tau(2\tau - 1)}{6} \right] + 1 \right\} \quad (4.2.7)$$

pro $\tau \in (0, \infty)$, kde $\sigma^2 > 0$ je parametr určující chybový rozptyl předpovědi o jednu časovou jednotku vpřed. Je důležité si všimnout, že $\text{var}[e_{t_n+\tau}(t_n)]$ počítané podle vzorce (4.2.7) je kladný pro všechna $\alpha, \gamma \in (0, 1)$ a $\tau \in (0, \infty)$.

Vzorec (4.2.7) můžeme opět modifikovat tak, aby zohledňoval časovou strukturu řady y do času t_n :

$$\text{var}[e_{t_n+\tau}(n)] = \sigma^2 \left\{ \alpha^2(\tau-1) \left[1 + \gamma\tau + \gamma^2 \frac{\tau(2\tau-1)}{6} \right] + \right. \\ \left. + (\alpha_{t_n} - \alpha)^+(1+\alpha) + \alpha^2\tau^2(\gamma_{t_n} - \gamma)^+(1+\gamma) + 1 \right\}, \quad (4.2.8)$$

kde $x^+ = \max(x, 0)$ je kladná část čísla x . Vzorec (4.2.8) splývá se vzorcem (4.2.7) pro $\alpha_{t_n} \leq \alpha$ a současně $\gamma_{t_n} \leq \gamma$. Pro $\alpha_{t_n} \geq \alpha$ je $\text{var}[e_{t_n+\tau}(t_n)]$ lineárně rostoucí funkcí α_{t_n} , podobně pro γ_{t_n} . Meze předpovědních intervalů za předpokladu normality jsou zřejmé.

Na závěr tohoto odstavce uvedme analogická zobecnění Holtovy metody s tlumeným lineárním a exponenciálním trendem pro nepravidelné časové řady, viz. Cipra (2006). Půjde o výsledek stejného přístupu, jaký použil Wright na klasickou Holtovu metodu s lineárním trendem. Pro Holtovu metodu s tlumeným lineárním trendem dojdeme k

$$S_{t_{n+1}} = (1 - \alpha_{t_n}) \cdot \left(S_{t_n} + T_{t_n} \cdot \varphi \frac{1 - \varphi^{t_{n+1} - t_n}}{1 - \varphi} \right) + \alpha_{t_n} \cdot y_{t_{n+1}}, \quad (4.2.9)$$

$$T_{t_{n+1}} = (1 - \gamma_{t_n}) \cdot \varphi^{t_{n+1} - t_n} \cdot T_{t_n} + \gamma_{t_n} \cdot \frac{(S_{t_{n+1}} - S_{t_n})(1 - \varphi)}{1 - \varphi^{t_{n+1} - t_n}} \cdot \varphi^{t_{n+1} - t_n - 1}, \quad (4.2.10)$$

$$\hat{y}_{t_n+\tau}(t_n) = S_{t_n} + \varphi \frac{1 - \varphi^\tau}{1 - \varphi} \cdot T_{t_n}^\tau, \quad \tau > 0. \quad (4.2.11)$$

Byl přitom použit vzorec $\varphi + \varphi^2 + \dots + \varphi^\tau = \varphi \frac{1 - \varphi^\tau}{1 - \varphi}$, což umožňuje použít tuto metodu i na časové řady, kde délka časového kroku $t_{n+1} - t_n$ mezi sousedními pozorováními není přirozené číslo.

Pro Holtovu metodu s exponenciálním trendem dospějeme ke vzorcům

$$S_{t_{n+1}} = (1 - \alpha_{t_n}) \cdot S_{t_n} \cdot T_{t_n}^{t_{n+1} - t_n} + \alpha_{t_n} \cdot y_{t_{n+1}}, \quad (4.2.12)$$

$$T_{t_{n+1}} = (1 - \gamma_{t_n}) \cdot T_{t_n} + \gamma_{t_n} \cdot \left(\frac{S_{t_{n+1}}}{S_{t_n}} \right)^{\frac{1}{t_{n+1} - t_n}}, \quad (4.2.13)$$

$$\hat{y}_{t_n+\tau}(t_n) = S_{t_n} \cdot T_{t_n}^\tau, \quad \tau > 0. \quad (4.2.14)$$

Při volbě počátečních hodnot t_0 , S_{t_0} , T_{t_0} , α_{t_0} a γ_{t_0} se v obou případech postupuje analogicky jako v případě dotyčných metod pro pravidelné časové řady resp. jako při Wrightově modifikaci Holtovy metody s lineárním trendem.

Pokud bychom chtěli konstruovat předpovědní intervaly pro Holtovu metodu s tlumeným lineárním trendem, můžeme opět převzít rozptyly předpovědních chyb odvozené pro odpovídající metodu pro pravidelné časové řady:

$$\text{var}[e_{t_n+\tau}(t_n)] = \sigma^2 \cdot \{ \alpha^2 \cdot (\tau - 1 + 2\gamma \cdot G_1(\tau) + \gamma^2 \cdot G_2(\tau)) + 1 \}, \quad (4.2.15)$$

kde

$$G_1(\tau) = \frac{\varphi}{1-\varphi} \left(\tau - 1 - \frac{1-\varphi^{\tau-1}}{1-\varphi} \right), \quad (4.2.16)$$

$$G_2(\tau) = \frac{\varphi^2}{(1-\varphi)^2} \left(\tau - 1 - 2\varphi \frac{1-\varphi^{\tau-1}}{1-\varphi} + \frac{1-\varphi^{2\tau-2}}{1-\varphi^2} \right). \quad (4.2.17)$$

Tyto vzorce opět nezohledňují časovou strukturu řady y do času t_n . Modifikace v tomto směru, analogická vzorcům (4.2.8), by mohla vypadat takto:

$$\text{var}[e_{t_n+\tau}(t_n)] = \sigma^2 \left[\alpha^2 (\tau - 1 + 2\gamma \cdot G_1(\tau) + \gamma^2 \cdot G_2(\tau)) + (\alpha_{t_n} - \alpha)^+ (1 + \alpha) + \alpha^2 g^2(\tau) (\gamma_{t_n} - \gamma)^+ (1 + \gamma) + 1 \right], \quad (4.2.18)$$

kde $G_1(\tau)$ a $G_2(\tau)$ mají stejný význam jako dříve.

4.3 Holt-Wintersova metoda pro řady s chybějícími pozorováními

Použitelnost Holtovy metody a jejích různých modifikací v praxi je limitována především tím, že jde o metody určené pro nesezónní časové řady. Většina reálných časových řad však vykazuje sezónnost, nejčastěji roční, týdenní nebo denní.

Rozšíření Holtovy metody pro sezónní časové řady je známo jako Holt-Wintersova metoda a pochází ze stejné doby jako samotná Holtova metoda, viz. původní článek Winters (1960). Sezónnost je zde modelována pomocí sezónních indexů, a to buď v aditivní nebo multiplikatívni podobě. Stejně jako v případě Holtovy metody lze uvažovat Holt-Wintersovu metodu s tlumeným lineárním trendem, viz. Gardner a McKenzie (1989), či s exponenciálním trendem. Úplný přehled všech možných kombinací lze nalézt v článku Gardner (1986).

V článku Cipra a kol. (1995) byla navržena modifikace Holt-Wintersovy metody pro časové řady s chybějícími pozorováními, a to aplikací stejného postupu, jaký zvolil Wright na jednoduché exponenciální vyrovnávání a Holtovu metodu, viz. Wright (1986) či odstavce 3.2 a 4.2 této práce.

V tomto odstavci nejprve osvětlíme podstatu Holt-Wintersovy metody na její verzi pro pravidelné časové řady a poté bude prezentováno zmíněné zobecnění této metody pro časové řady s chybějícími pozorováními. V obou případech bude k ilustraci použita varianta s aditivní sezónností, analogické vzorce stejných metod s multiplikatívni sezónností budou vždy uvedeny dodatečně. V obou případech též půjde o klasické verze Holt-Wintersovy metody, tedy s lokálně lineárním trendem.

Uvažujme pravidelnou časovou řadu $\dots y_{n-1}, y_n, y_{n+1}, \dots$ s lokálně lineárním trendem a sezónností s periodou $p \geq 2$. Holt-Wintersova metoda kromě *sezóně očištěné*

úrovně S_n a směrnice T_n lokálně lineárního trendu uvažuje ještě tzv. *sezónní index* I_n , vyjadřující vliv sezónnosti v čase n . To odpovídá následující aditivní dekompozici hodnoty y_n řady y v čase n :

$$y_n = S_n + I_n + \varepsilon_n , \quad (4.3.1)$$

kde ε_n je náhodná složka. Pro vyrovnanou hodnotu řady y v čase n platí

$$\hat{y}_n = S_n + I_n . \quad (4.3.2)$$

Předpověď o $\tau = 1, 2, 3, \dots$ kroků vpřed budoucí neznámé hodnoty $y_{n+\tau}$ z času n je

$$\hat{y}_{n+\tau}(n) = S_n + \tau \cdot T_n + I_{n-p+1+(\tau-1)\text{mod } p} , \quad (4.3.3)$$

kde $m \text{ mod } p$ je zbytek celého čísla m po dělení p . Tento poněkud formálně složitý vzorec v sobě přirozeným způsobem kombinuje lineární trend a aditivní sezónnost pomocí sezónních indexů s periodou p .

Na základě podobných úvah jako v případě Holtovy metody dojdeme k následujícím rekurentním vzorcům pro výpočet hodnot S_{n+1} , T_{n+1} a I_{n+1} :

$$S_{n+1} = (1 - \alpha)(S_n + T_n) + \alpha(y_{n+1} - I_{n+1-p}) , \quad (4.3.4)$$

$$T_{n+1} = (1 - \gamma)T_n + \gamma(S_{n+1} - S_n) , \quad (4.3.5)$$

$$I_{n+1} = (1 - \delta)I_{n+1-p} + \delta(y_{n+1} - S_{n+1}) , \quad (4.3.6)$$

kde $\alpha, \gamma, \delta \in (0, 1)$ jsou postupně vyrovnávací konstanta pro úroveň, směrnici trendu a sezónní index. Také vzorce (4.3.4) až (4.3.6) lze přepsat do jejich chybových tvarů:

$$S_{n+1} = S_n + T_n + \alpha \cdot e_{n+1} , \quad (4.3.7)$$

$$T_{n+1} = T_n + \gamma\alpha \cdot e_{n+1} , \quad (4.3.8)$$

$$I_{n+1} = I_{n+1-p} + \delta(1 - \alpha) \cdot e_{n+1} , \quad (4.3.9)$$

kde $e_{n+1} = y_{n+1} - \hat{y}_{n+1}(n)$ je chyba jednokrokové předpovědi. Volba vyrovnávacích konstant $\alpha, \gamma, \delta \in (0, 1)$ se provádí stejně jako u Holtovy metody.

Pro praktické použití metody na řadu $y_1, y_2, \dots, y_n, \dots$ je nutné nějak zvolit počáteční hodnoty S_0 , T_0 a I_t pro $t = -p+1, -p+2, \dots, -1, 0$. Pokud máme pro nastavení těchto hodnot k dispozici k kompletních period, $k \geq 2$, můžeme provést volbu počátečních hodnot podle následujících intuitivních vzorců:

$$T_0 = \frac{\bar{y}_k - \bar{y}_1}{(k-1)p} , \quad S_0 = \bar{y}_1 - \frac{p+1}{2} T_0 , \quad (4.3.10)$$

$$I_t = \sum_{i=1}^k \left[y_{t+i,p} - \bar{y}_i - T_0 \left(t + \frac{p-1}{2} \right) \right] , \quad t = -p+1, \dots, -1, 0 , \quad (4.3.11)$$

kde \bar{y}_i je aritmetický průměr pozorování i -té periody, $i = 1, 2, \dots, k$. Jinou možností je použít k získání počátečních hodnot odhady příslušných parametrů v modelu lineární regrese se sezónními indexy.

V případě Holt-Wintersovy metody s multiplikativní sezónností stačí vzorce (4.3.2), (4.3.3), (4.3.4) a (4.3.6) nahradit takto

$$\hat{y}_{n+\tau}(n) = S_n \cdot I_n, \quad (4.3.12)$$

$$\hat{y}_{n+\tau}(n) = S_n + \tau \cdot T_n \cdot I_{n-p+1+(\tau-1)\text{mod } p}, \quad (4.3.13)$$

$$S_{n+1} = (1 - \alpha)(S_n + T_n) + \alpha(y_{n+1}/I_{n+1-p}), \quad (4.3.14)$$

$$I_{n+1} = (1 - \delta)I_{n+1-p} + \delta(y_{n+1}/S_{n+1}). \quad (4.3.15)$$

Nyní popíšeme zobecnění Holt-Wintersovy metody pro případ časové řady s chybějícími pozorováními, jak jej navrhli Cipra a kol. (1995). Zatímco Wrightovo zobecnění Holtovy metody lze použít na jakoukoli nepravidelnou časovou řadu, v případě Holt-Wintersovy metody se skutečně omezuje jen na řady s chybějícími pozorováními. Důvodem tohoto omezení je modelování sezónnosti pomocí sezónních indexů, které přináší nutnost jednoznačně zařadit každé pozorování zkoumané časové řady do právě jednoho z p časových úseků periody délky p . Např. při měsíčních pozorováních v rámci roční periody ($p = 12$) je nutné, aby každé pozorování odpovídalo buď lednu, únoru, ... či prosinci.

Uvažujme sezónní časovou řadu $\dots y_{t_{n-1}}, y_{t_n}, y_{t_{n+1}}, \dots$ s periodou p , kde časové okamžiky $\dots < t_{n-1} < t_n < t_{n+1} < \dots$ jsou celá čísla a představují tak výběr z pravidelné časové mřížky. Podobně jako v pravidelném případě označme symboly S_t , T_t a I_t postupně sezónně očištěnou úroveň, směrnici trendu a sezónní index řady y v čase t a k nim uvažujme tři příslušné vyrovnávací konstanty $\alpha, \gamma, \delta \in (0, 1)$. Pro jednoduchost následujících zápisů ještě zavedme pro $t, n \in \mathbf{Z}$ následující značení:

$$t * n = \max \{ t_k : k \leq n \wedge p \mid (t - t_k) \}, \quad (4.3.16)$$

kde zápis $p \mid d$ znamená, že číslo d je dělitelné číslem p .

Analogie vzorců (4.3.3) až (4.3.6) vypadají nyní takto:

$$\hat{y}_{t_n+\tau}(t_n) = S_{t_n} + \tau \cdot T_{t_n} + I_{(t_n+\tau)*n}, \quad (4.3.17)$$

$$S_{t_{n+1}} = (1 - \alpha_{t_{n+1}})[S_{t_n} + T_{t_n}(t_{n+1} - t_n)] + \alpha_{t_{n+1}}[y_{t_{n+1}} - I_{(t_{n+1})*n}], \quad (4.3.18)$$

$$T_{t_{n+1}} = (1 - \gamma_{t_{n+1}})T_{t_n} + \gamma_{t_{n+1}} \frac{S_{t_{n+1}} - S_{t_n}}{t_{n+1} - t_n}, \quad (4.3.19)$$

$$I_{t_{n+1}} = (1 - \delta_{t_{n+1}})I_{(t_{n+1})*n} + \delta_{t_{n+1}}(y_{t_{n+1}} - S_{t_{n+1}}). \quad (4.3.20)$$

Proměnné vyrovnávací koeficienty $\alpha_{t_n}, \gamma_{t_n}, \delta_{t_n}$ se přepočítávají pomocí stejných rekurentních vzorců jako v případě Wrightovy modifikace Holtovy metody, viz vzorce (4.2.5), tedy:

$$\alpha_{t_{n+1}} = \frac{\alpha_{t_n}}{\alpha_{t_n} + (1 - \alpha)^{t_{n+1} - t_n}}, \quad \gamma_{t_{n+1}} = \frac{\gamma_{t_n}}{\gamma_{t_n} + (1 - \gamma)^{t_{n+1} - t_n}}, \quad (4.3.21)$$

$$\delta_{t_{n+1}} = \frac{\delta_{(t_n+1)^*n}}{\delta_{(t_n+1)^*n} + (1 - \delta)^{\frac{t_{n+1} - (t_n+1)^*n}{p}}}. \quad (4.3.22)$$

Ve vzorci pro $\delta_{t_{n+1}}$ se exponent dělí délkou periody p , aby vyrovnávací konstanta δ odpovídala vyrovnávací konstantě z metody pro pravidelné časové řady. Je důležité si uvědomit, že stejně jako je třeba v paměti uchovávat p hodnot sezónních indexů I_{t_n} , je třeba držet v paměti i posledních p hodnot vyrovnávacího koeficientu δ_{t_n} .

Pokud jde o volbu počátečních hodnot, postupujeme analogicky jako v případě Wrightovy modifikace Holtovy metody resp. jako v případě Holt-Wintersovy metody pro pravidelné řady. Tedy

$$t_k = t_1 - k \cdot q, \quad \delta_{t_k} = 1 - (1 - \delta)^{Q_k}, \quad k = -p + 1, \dots, -1, 0, \quad (4.3.23)$$

$$\alpha_{t_0} = 1 - (1 - \alpha)^q \quad \text{a} \quad \gamma_{t_0} = 1 - (1 - \gamma)^q, \quad (4.3.24)$$

kde q je průměrná časová vzdálenost sousedních pozorování řady y a Q_k je průměrný počet period délky p mezi dvěma pozorováními ze stejného období odpovídajícího času t_k , $k = -p + 1, \dots, -1, 0$. Je možné uvažovat realistické zjednodušení $Q_{-p+1} = \dots = Q_{-1} = Q_0 = q$. Volba hodnot S_{t_0} , T_{t_0} a I_{t_k} , $k = -p + 1, \dots, -1, 0$ lze provést obdobně jako v pravidelném případě.

V případě multiplikativní sezónnosti přejdou vzorce (4.3.17), (4.3.18) a (4.3.20) na

$$\hat{y}_{t_n+\tau}(t_n) = S_{t_n} + \tau \cdot T_{t_n} \cdot I_{(t_n+\tau)^*n}, \quad (4.3.25)$$

$$S_{t_{n+1}} = (1 - \alpha_{t_{n+1}})[S_{t_n} + T_{t_n}(t_{n+1} - t_n)] + \alpha_{t_{n+1}} \frac{y_{t_{n+1}}}{I_{(t_{n+1})^*n}}, \quad (4.3.26)$$

$$I_{t_{n+1}} = (1 - \delta_{t_{n+1}})I_{(t_{n+1})^*n} + \delta_{t_{n+1}} \frac{y_{t_{n+1}}}{S_{t_{n+1}}}. \quad (4.3.27)$$

5 Exponenciální vyrovnávání řádu m

V této kapitole se budeme zabývat exponenciálním vyrovnáváním řádu m . Jde o adaptivní a rekurentní způsob odhadu parametrů polynomického trendu stupně m v nesezónní časové řady s využitím pouze jedné vyrovnávací konstanty α . Jednoduché exponenciální vyrovnávání je speciálním případem této metody pro $m = 0$. Případ $m = 1$, tedy metoda odhadující lineární trend, se nazývá *dvojitě exponenciální vyrovnávání* či *Brownova metoda*. Metodě odhadující kvadratický trend ($m = 2$) se říká *trojitě exponenciální vyrovnávání* atd.

Odstavec 5.1 bude věnována exponenciálnímu vyrovnáváním řádu m pro pravidelné časové řady, které vychází z DLS odhadu parametrů polynomického trendu stupně m . Metoda bude nejprve ilustrována na příkladu dvojitě exponenciálního vyrovnávání, poté bude objasněno její fungování pro obecný řád m .

Cipra (2006) odvodil dvojitě exponenciální vyrovnávání pro nepravidelné časové řady. Stejným způsobem bude v odstavci 5.2 odvozeno exponenciální vyrovnávání obecného řádu m pro případ nepravidelné časové řady. V odstavci 5.3 pak bude vedle toho odvozena i podobná metoda založená na DLS odhadu polynomiálního trendu stupně m . Tato metoda je s exponenciálním vyrovnáváním řádu m ekvivalentní pouze ve verzi pro pravidelné časové řady.

5.1 Exponenciální vyrovnávání řádu m pro pravidelné časové řady

V odstavci 3.1 bylo zmíněno, že jednoduché exponenciální vyrovnávání pro pravidelné časové řady s vyrovnávací konstantou $\alpha \in (0, 1)$ odpovídá odhadu konstantního trendu metodou DLS s diskontním faktorem $\beta = 1 - \alpha$. Podobně dvojitě exponenciální vyrovnávání pro pravidelné časové řady odpovídá DLS odhadu lineárního trendu atd. V tomto odstavci budeme věc ilustrovat právě na dvojitě exponenciálním vyrovnáváním (Brownova metoda), které je fakticky jistým speciálním případem Holtovy metody s lineárním trendem.

Uvažujme pravidelnou časovou řadu $\dots, y_{n-2}, y_{n-1}, y_n, y_{n+1}, \dots$ s nekonečnou minulostí. Uvažujme dále model

$$y_{n-j} = b_0 - b_1 j + \varepsilon_{n-j}, \quad j = 0, 1, 2 \dots \quad (5.1.1)$$

pro pozorování řady y do času n včetně. Zde přítomné náhodné složky ε_t nechť mají nulovou střední hodnotu. Odhadněme nyní neznámé parametry b_0 a b_1 tohoto modelu

metodou DLS s diskontním faktorem $\beta = 1 - \alpha$, kde $\alpha \in (0, 1)$ je pevně zvolená vyrovnávací konstanta. Řešíme tedy úlohu

$$\arg \min_{b_0, b_1} \left\{ \sum_{j=0}^{\infty} \beta^j (y_{n-j} - b_0 + b_1 j)^2 \right\}. \quad (5.1.2)$$

Derivováním minimalizovaného výrazu podle b_0 a b_1 a položením těchto derivací rovných nule získáme soustavu dvou normálních rovnic pro neznámé b_0 a b_1 :

$$\begin{aligned} \sum_{j=0}^{\infty} \beta^j (y_{n-j} - b_0 + b_1 j) &= 0, \\ \sum_{j=0}^{\infty} j \beta^j (y_{n-j} - b_0 + b_1 j) &= 0. \end{aligned} \quad (5.1.3)$$

S použitím součtových vzorců $\sum_{j=0}^{\infty} j \beta^j = \frac{\beta}{(1-\beta)^2}$ a $\sum_{j=0}^{\infty} j^2 \beta^j = \frac{\beta(1+\beta)}{(1-\beta)^3}$ a po drobných algebraických úpravách můžeme tuto soustavu přepsat do tvaru

$$\begin{aligned} b_0 - \frac{\beta}{1-\beta} b_1 &= (1-\beta) \sum_{j=0}^{\infty} y_{n-j} \beta^j, \\ b_0 - \frac{\beta(1+\beta)}{1-\beta} b_1 &= (1-\beta)^2 \sum_{j=0}^{\infty} y_{n-j} j \beta^j. \end{aligned} \quad (5.1.4)$$

Zavedme nyní tzv. *jednoduchou vyrovnávací statistiku*

$$S_n = S_n^{[1]} = (1-\beta) \cdot \sum_{j=0}^{\infty} y_{n-j} \beta^j \quad (5.1.5)$$

a podobně tzv. *dvojitou vyrovnávací statistiku*

$$S_n^{[2]} = (1-\beta) \cdot \sum_{j=0}^{\infty} S_{n-j} \beta^j. \quad (5.1.6)$$

Hodnota S_n je tedy váženým průměrem hodnot řady y v časech počínaje n směrem do minulosti, kde váhy klesají exponenciálně s diskontním faktorem β . Podobně $S_n^{[2]}$ je stejným váženým průměrem hodnot $\dots, S_{n-2}, S_{n-1}, S_n$. Platí rekurentní formule známé již z jednoduchého exponenciálního vyrovnávání, viz. např. (3.1.1):

$$S_{n+1} = (1-\alpha) \cdot S_n + \alpha \cdot y_{n+1}, \quad (5.1.7)$$

$$S_{n+1}^{[2]} = (1-\alpha) \cdot S_n^{[2]} + \alpha \cdot S_{n+1}. \quad (5.1.8)$$

Pokusme se nyní vyjádřit hodnotu $S_n^{[2]}$ přímo pomocí hodnot řady y :

$$\begin{aligned}
S_n^{[2]} &= (1 - \beta) \cdot \sum_{j=0}^{\infty} S_{n-j} \beta^j = (1 - \beta)^2 \cdot \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{n-(i+j)} \beta^{i+j} = \\
&= (1 - \beta)^2 \cdot \sum_{k=0}^{\infty} (k + 1) y_{n-k} \beta^k .
\end{aligned} \tag{5.1.9}$$

Soustavu rovnic (5.1.4) můžeme nyní přepsat do tvaru

$$\begin{aligned}
b_0 - \frac{\beta}{1 - \beta} b_1 &= S_n , \\
b_0 - \frac{\beta(1 + \beta)}{1 - \beta} b_1 &= S_n^{[2]} - S_n .
\end{aligned} \tag{5.1.10}$$

Odtud snadno dostaneme její řešení jako

$$\hat{b}_0(n) = 2S_n - S_n^{[2]} , \quad \hat{b}_1(n) = \frac{\alpha}{1 - \alpha} (S_n - S_n^{[2]}) . \tag{5.1.11}$$

Vyrovnaná hodnota řady y a předpověď o $\tau > 0$ kroků vpřed z času n jsou přirozeně

$$\hat{y}_n = \hat{b}_0(n) = 2S_n - S_n^{[2]} , \tag{5.1.12}$$

$$\hat{y}_{n+\tau}(n) = \hat{b}_0(n) + \hat{b}_1(n) \cdot \tau = \left(2 + \frac{\alpha\tau}{1 - \alpha} \right) S_n - \left(1 + \tau \frac{\alpha\tau}{1 - \alpha} \right) S_n^{[2]} . \tag{5.1.13}$$

Chceme-li použít tuto rekurentní metodu na časovou řadu y_1, y_2, y_3, \dots , musíme nějakým způsobem získat počáteční hodnoty S_0 a $S_0^{[2]}$. Ty se vyjádří z rovnic (5.1.10), kam se za $\hat{b}_0(0)$ a $\hat{b}_1(0)$ dosadí odhady parametrů v modelu $y_j = b_0 + b_1 j + \varepsilon_j$ získané z několika prvních pozorování řady y . Běžně se zde používají OLS odhady. Alternativou jsou již několikrát zmíněné invertované DLS odhady, tentokrát přirozeně s diskontním faktorem β . Hodnota vyrovnávací konstanty $\alpha \in (0, 1)$ se volí stejným způsobem jako u jednoduchého exponenciálního vyrovnávání.

Jednoduchými algebraickými úpravami lze ukázat, že Brownova metoda s vyrovnávací konstantou α je ekvivalentní Holtově metodě s lineárním trendem s vyrovnávacími konstantami α_H a γ_H určenými vzorci

$$\alpha_H = \alpha(2 - \alpha) \quad \text{a} \quad \gamma_H = \frac{\alpha}{2 - \alpha} . \tag{5.1.14}$$

Všimněme si, že $\alpha = \sqrt{\alpha_H \cdot \gamma_H}$, tedy vyrovnávací konstanta Brownovy metody je geometrickým průměrem vyrovnávacích konstant ekvivalentní Holtovy metody.

Holtova metoda nám díky dvou různým a spolu nijak nesvázaným vyrovnávacím konstantám poskytuje větší flexibilitu, a tím i lepší praktické výsledky. Z toho důvodu je jí v praxi před Brownovou metodou dáována přednost. Situace může být však méně jednoznačná v případě nepravidelných časových řad (viz. odst. 7.2).

Při konstrukci předpovědních intervalů pro Brownovu metodu můžeme s výhodou využít faktu, že jde o speciální případ Holtovy metody. To nás vede ke vzorci

$$\text{var}[e_{n+\tau}(n)] = \sigma^2 \left\{ \alpha_H^2 (\tau - 1) \left[1 + \gamma_H \cdot \tau + \gamma_H^2 \frac{\tau(2\tau - 1)}{6} \right] + 1 \right\} \quad (5.1.15)$$

pro rozptyl chyby předpovědi o $\tau > 0$ kroků vpřed, viz. vzorec (4.1.19) pro Holtovu metodu. Hodnoty α_H a γ_H jsou získány z vyrovnávací konstanty α podle vzorců (5.1.14). Předpovědní interval za předpokladu normality dostaneme již snadno.

Exponenciální vyrovnávání obecného řádu m pro pravidelné časové řady odvodíme analogicky jako dvojité exponenciální vyrovnávání v předchozích odstavcích. Uvažujeme polynomický trend stupně m centrováný v čase n , tedy

$$y_{n-j} = b_0 + b_1(-j) + b_2(-j)^2 + \dots + b_m(-j)^m + \varepsilon_{n-j}, \quad j = 0, 1, 2, \dots \quad (5.1.16)$$

Jeho parametry b_0, b_1, \dots, b_m opět odhadujeme metodou DLS s diskontním faktorem $\beta = 1 - \alpha$.

Pro úpravu vzniklé soustavy normálních rovnic je třeba určit součty nekonečných řad $\sum_{j=0}^{\infty} j^k \beta^j$, $k = 0, 1, \dots, m$, což však není problém, viz. rekurentní výpočet (5.3.18) v odstavci 5.3. Dále se v této soustavě normálních rovnic vyskytují nekonečné řady $\sum_{j=0}^{\infty} j^k y_{n-j} \beta^j$ pro $k = 0, 1, \dots, m$. Uvažujme nyní prvních $m+1$ vyrovnávacích statistik $S_n^{[1]}, S_n^{[2]}, \dots, S_n^{[m+1]}$, které jsou definovány analogicky jako $S_n^{[1]}$ a $S_n^{[2]}$, tj.

$$S_n^{[k+1]} = (1 - \beta) \cdot \sum_{j=0}^{\infty} S_{n-j}^{[k]} \beta^j, \quad k = 1, \dots, m. \quad (5.1.17)$$

Přirozeně platí rekurentní formule

$$S_{n+1}^{[k+1]} = (1 - \alpha) \cdot S_n^{[k+1]} + \alpha \cdot S_{n+1}^{[k]}, \quad k = 1, \dots, m. \quad (5.1.18)$$

Indukcí lze dokázat, že pro $k \geq 1$ platí také vyjádření

$$S_n^{[k]} = (1 - \beta)^k \cdot \sum_{j=0}^{\infty} C_{k-1}(j) y_{n-j} \beta^j, \quad (5.1.19)$$

kde $C_k(j) = \binom{k+j}{k}$ lze chápat jako polynom stupně k proměnné j . Polynomy

C_0, C_1, \dots, C_k tvoří bázi vektorového prostoru všech reálných polynomů stupně nejvýše k . Tudíž pro každé $k \geq 1$ existují reálné koeficienty $\pi_0^k, \pi_1^k, \dots, \pi_k^k$ takové, že

$$j^k \equiv \sum_{i=0}^k \pi_i^k C_i(j) \text{ a odtud}$$

$$\sum_{j=0}^{\infty} j^k y_{n-j} \beta^j = \sum_{i=0}^k \pi_i^k \frac{S_n^{[i+1]}}{(1-\beta)^{i+1}}. \quad (5.1.20)$$

Tedy výrazy $\sum_{j=0}^{\infty} j^k y_{n-j} \beta^j$ vyskytující se pro $k = 0, 1, \dots, m$ v soustavě normálních rovnic pro neznámé parametry b_0, b_1, \dots, b_m lze vyjádřit jako lineární kombinace prvních $m+1$ vyrovnávacích statistik $S_n^{[1]}, S_n^{[2]}, \dots, S_n^{[m+1]}$. Totéž pak platí i pro odhady $\hat{b}_0(n), \hat{b}_1(n), \dots, \hat{b}_m(n)$ parametrů b_0, b_1, \dots, b_m .

Vyrovnaná hodnota řady y a předpověď o $\tau > 0$ kroků vpřed z času n jsou

$$\hat{y}_n = \hat{b}_0(n), \quad (5.1.21)$$

$$\hat{y}_{n+\tau}(n) = \hat{b}_0(n) + \hat{b}_1(n) \cdot \tau + \hat{b}_2(n) \cdot \tau^2 + \dots + \hat{b}_1(n) \cdot \tau^m. \quad (5.1.22)$$

Počáteční hodnoty $S_0^{[1]}, S_0^{[2]}, \dots, S_0^{[m+1]}$ se mohou volit obdobně jako u dvojitého exponenciálního vyrovnávání.

5.2 Exponenciální vyrovnávání řádu m pro nepravidelné časové řady

V předchozím odstavci bylo ilustrováno exponenciální vyrovnávání řádu m pro pravidelné časové řady. Východiskem zde byl odhad parametrů polynomického trendu stupně m metodou DLS. Při teoretickém předpokladu nekonečné minulosti zkoumané časové řady bylo možné tyto odhady parametrů vyjádřit jako lineární kombinace (s koeficienty neměnnými v čase) prvních $m+1$ vyrovnávacích statistik $S_n^{[1]}, S_n^{[2]}, \dots, S_n^{[m+1]}$. Ty jdou navíc počítat pomocí rekurentních vzorců (5.1.18). Výsledkem je tedy výpočetně velice efektivní adaptivní metoda pro pravidelnou časovou řadu vykazující lokálně polynomický trend stupně m .

Budeme-li chtít odvodit podobnou metodu pro nepravidelné časové řady, narazíme hned na několik obtíží. Předně již nebudeme moci využít vztahu podobného tomu v (5.1.20), či-li nebudeme mít k dispozici vztah mezi vyrovnávacími statistikami S a součty nekonečných řad zahrnujících hodnoty řady y vyskytujícími se v soustavě normálních rovnic. Tedy zatímco v případě pravidelné časové řady vedlo použití DLS odhadu na metodu pracující s vyrovnávacími statistikami S , v případě nepravidelné časové řady již tomu tak nebude. Zde se nabízejí obecně dvě různé metody pro adaptivní odhad polynomického trendu - metoda pracující s vyrovnávacími statistikami a metoda využívající DLS odhadu parametrů. *Exponenciálním vyrovnáváním* budeme nazývat první z těchto metod, která bude také odvozena v tomto odstavci. Postup bude analogický tomu, jaký použil Cipra (2006) k odvození dvojitého exponenciálního vyrovnávání pro nepravidelné časové řady.

V případě pravidelně pozorovaných hodnot byla nutnou podmínkou pro odvození exponenciálního vyrovnávání (aspoň v té podobě jako v odstavci 5.1) nekonečná

minulost dané časové řady. Pracujeme-li s nepravidelnou časovou řadou, nehraje již otázka nekonečnosti či konečnosti její minulosti žádnou roli. Nepravidelnost jejich pozorování beztak neumožňuje dospět k jednoduchým vzorcům z předchozího odstavce.

Zde uvažujme stejně jako Cipra (2006) nepravidelnou časovou řadu $y_{t_1}, y_{t_2}, \dots, y_{t_n}, y_{t_{n+1}}, \dots$ pozorovanou v časech $t_1 < t_2 < \dots < t_n < t_{n+1} < \dots$, tedy případ s konečnou minulostí. Nechť $m \in \mathbf{N}_0$ je řád exponenciálního vyrovnávání, tedy stupeň uvažovaného polynomického trendu. Dále nechť $n \geq m + 1$ a uvažujme model

$$y_{t_j} = b_0 + b_1(t_n - t_j) + b_2(t_n - t_j)^2 + \dots + b_m(t_n - t_j)^m + \varepsilon_{t_j}, \quad j = 1, 2, 3, \dots, \quad (5.2.1)$$

kde $b_0, b_1, \dots, b_m \in \mathbf{R}$ jsou neznámé parametry a náhodné složky ε_{t_j} mají nulovou střední hodnotu. O jejich kovarianční struktuře nečiníme žádné předpoklady. Uvažujme daný *diskontní faktor* $\beta \in (0, 1)$ a k němu *vyrovnávací konstantu* $\alpha = 1 - \beta$.

V následujících odstavcích sestrojíme nestranné odhady $\hat{b}_0(t_n), \hat{b}_1(t_n), \dots, \hat{b}_m(t_n)$ neznámých parametrů b_0, b_1, \dots, b_m založené na prvních n pozorování řady y . Za tímto účelem budeme pro $j = 1, 2, 3, \dots$ a $p = 1, 2, \dots, m + 1$ definovat *vyrovnávací statistiky* $S_{t_j}^{[p]}$ následujícím způsobem:

$$S_{t_j}^{[1]} = \alpha_{t_j} \cdot \sum_{i=1}^j y_{t_i} \beta^{t_j - t_i}, \quad (5.2.2)$$

$$S_{t_j}^{[p+1]} = \alpha_{t_j} \cdot \sum_{i=1}^j S_{t_i}^{[p]} \beta^{t_j - t_i}, \quad p = 1, 2, \dots, m, \quad (5.2.3)$$

kde

$$\alpha_{t_j} = \left(\sum_{i=1}^j \beta^{t_j - t_i} \right)^{-1}, \quad j = 1, 2, 3, \dots. \quad (5.2.4)$$

Dále pro $k = 0, 1, 2, \dots, m$ a $j = 1, 2, 3, \dots$ označme

$${}^k T_{t_j}^{[1]}(t_n) = \alpha_{t_j} \cdot \sum_{i=1}^j (t_n - t_j)^k \beta^{t_j - t_i}, \quad (5.2.5)$$

$${}^k T_{t_j}^{[p+1]}(t_n) = \alpha_{t_j} \cdot \sum_{i=1}^j {}^k T_{t_i}^{[p]}(t_n) \beta^{t_j - t_i}, \quad p = 1, 2, \dots, m. \quad (5.2.6)$$

Zřejmě ${}^0 T_{t_j}^{[p]}(t_n) = 1$ pro všechna j, p a n (symbol ${}^0 T_{t_j}^{[p]}(t_n)$ zavádíme jen pro zjednodušení dalších zápisů). Pro zkrácení zápisu ještě označme ${}^k T_{t_n}^{[p]} = {}^k T_{t_n}^{[p]}(t_n)$.

Nyní se podívejme na náš model (5.2.1). Předně pro $j = 1, 2, 3, \dots$ platí

$$E(y_{t_j}) = b_0 + b_1(t_n - t_j) + b_2(t_n - t_j)^2 + \dots + b_m(t_n - t_j)^m. \quad (5.2.7)$$

Aplikací lineárního vyrovnávacího operátoru p -tého řádu na tuto rovnost pak můžeme s využitím zavedeného značení vyjádřit střední hodnotu vyrovnávacích statistik jako

$$E(S_{t_n}^{[p]}) = b_0 + b_1 \cdot {}^1T_{t_n}^{[p]} + b_2 \cdot {}^2T_{t_n}^{[p]} + \dots + b_m \cdot {}^mT_{t_n}^{[p]}, \quad p = 1, 2, \dots, m+1. \quad (5.2.8)$$

Toto je soustava $m+1$ lineárních rovnic pro $m+1$ neznámých parametrů b_0, b_1, \dots, b_m . Ty jsou jako řešení této soustavy lineární funkcí vektoru jejích pravých stran (v uvedeném zápisu jde o levé strany). Nahradíme-li zde hodnoty $E(S_{t_n}^{[p]})$ hodnotami $S_{t_n}^{[p]}$ samotnými, obdržíme nestranné odhady $\hat{b}_0(t_n), \hat{b}_1(t_n), \dots, \hat{b}_m(t_n)$ parametrů b_0, b_1, \dots, b_m . Jejich nestrannost plyne z linearit y střední hodnoty. Tedy naše odhady parametrů polynomickeho trendu v čase n získáme řešením soustavy rovnic

$$b_0 + b_1 \cdot {}^1T_{t_n}^{[p]} + b_2 \cdot {}^2T_{t_n}^{[p]} + \dots + b_m \cdot {}^mT_{t_n}^{[p]} = S_{t_n}^{[p]}, \quad p = 1, 2, \dots, m+1. \quad (5.2.9)$$

Vyrovnanou hodnotu v čase n a předpověď o $\tau > 0$ časových jednotek vpřed získáme pak jako

$$\hat{y}_{t_n} = \hat{b}_0(t_n), \quad (5.2.10)$$

$$\hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) + \hat{b}_1(t_n) \cdot (-\tau) + \hat{b}_2(t_n) \cdot (-\tau)^2 + \dots + \hat{b}_m(t_n) \cdot (-\tau)^m. \quad (5.2.11)$$

Jde přitom o nestranné předpovědi, totiž

$$E(\hat{y}_{t_n}) = E(y_{t_n}) \quad \text{a} \quad E[\hat{y}_{t_n+\tau}(t_n)] = E(y_{t_n+\tau}). \quad (5.2.12)$$

Získáme-li nové pozorování $y_{t_{n+1}}$, posuneme se z času t_n do času t_{n+1} a budeme odhadovat parametry b_0, b_1, \dots, b_m v aktualizovaném modelu

$$y_{t_j} = b_0 + b_1(t_{n+1} - t_j) + b_2(t_{n+1} - t_j)^2 + \dots + b_m(t_{n+1} - t_j)^m + \varepsilon_{t_j}, \quad j = 1, 2, 3, \dots. \quad (5.2.13)$$

Tyto odhady $\hat{b}_0(t_{n+1}), \hat{b}_1(t_{n+1}), \dots, \hat{b}_m(t_{n+1})$ získáme řešením soustavy (5.2.9) posunuté do času t_{n+1} , tedy soustavy

$$b_0 + b_1 \cdot {}^1T_{t_{n+1}}^{[p]} + b_2 \cdot {}^2T_{t_{n+1}}^{[p]} + \dots + b_m \cdot {}^mT_{t_{n+1}}^{[p]} = S_{t_{n+1}}^{[p]}, \quad p = 1, 2, \dots, m+1. \quad (5.2.14)$$

V následujících odstavcích odvodíme rekurentní formule, které nám umožní získat koeficienty nové soustavy (5.2.14) pomocí koeficientů původní soustavy (5.2.9).

Předně pro vyrovnávací statistiky $S_{t_{n+1}}^{[p]}$ figurující na pravé straně soustavy (5.2.14) platí následující rekurentní vzorce, známé již z Wrightovy modifikace jednoduchého exponenciálního vyrovnávání:

$$S_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[1]} + \alpha_{t_{n+1}} \cdot y_{t_{n+1}}, \quad (5.2.15)$$

$$S_{t_{n+1}}^{[p+1]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[p+1]} + \alpha_{t_{n+1}} \cdot S_{t_{n+1}}^{[p]}, \quad p = 1, 2, \dots, m. \quad (5.2.16)$$

Vyrovnávací koeficient $\alpha_{t_{n+1}}$ získáme rovněž pomocí známého rekurentního vzorce

$$\alpha_{t_{n+1}} = \frac{\alpha_{t_n}}{\alpha_{t_n} + \beta^{t_{n+1}-t_n}}. \quad (5.2.17)$$

Rekurentní výpočet koeficientů na levé straně soustavy (5.2.14) bude o něco komplikovanější. Nejprve vyjádříme hodnoty ${}^k T_{t_n}^{[p]}(t_{n+1})$ pomocí hodnot ${}^k T_{t_n}^{[p]} = {}^k T_{t_n}^{[p]}(t_n)$. K tomu využijeme následující rovnosti vyplývající z binomické věty:

$$(t_{n+1} - t_j)^k = [(t_n - t_j) + (t_{n+1} - t_n)]^k = \sum_{i=0}^k \binom{k}{i} (t_{n+1} - t_n)^{k-i} (t_n - t_j)^i. \quad (5.2.18)$$

Na tuto rovnost budeme aplikovat příslušný vyrovnávací operátor a takto obdržíme

$${}^k T_{t_n}^{[p]}(t_{n+1}) = \sum_{i=0}^k \left[\binom{k}{i} (t_{n+1} - t_n)^{k-i} \cdot {}^i T_{t_n}^{[p]} \right]. \quad (5.2.19)$$

Nyní využijeme rekurentní vztahy obdobné těm v (5.2.15) a (5.2.16) k tomu, abychom z hodnot ${}^k T_{t_n}^{[p]}(t_{n+1})$ získali hodnoty ${}^k T_{t_{n+1}}^{[p]} = {}^k T_{t_{n+1}}^{[p]}(t_{n+1})$ vyskytující se již přímo v soustavě (5.2.14):

$${}^k T_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot {}^k T_{t_n}^{[1]}(t_{n+1}), \quad (5.2.20)$$

$${}^k T_{t_{n+1}}^{[p+1]} = (1 - \alpha_{t_{n+1}}) \cdot {}^k T_{t_n}^{[p+1]}(t_{n+1}) + \alpha_{t_{n+1}} \cdot {}^k T_{t_{n+1}}^{[p]}, \quad p = 1, 2, \dots, m. \quad (5.2.21)$$

Praktická realizace popsané vyrovnávací a předpovědní metody vypadá tedy následujícím způsobem:

- V čase t_n máme k dispozici hodnoty α_{t_n} , $S_{t_n}^{[p]}$ a ${}^k T_{t_n}^{[p]}$ pro $k = 0, 1, 2, \dots, m$, $p = 1, 2, \dots, m + 1$.
- Řešením soustavy (5.2.9) získáme odhady $\hat{b}_0(t_n), \hat{b}_1(t_n), \dots, \hat{b}_m(t_n)$.
- Určíme vyrovnanou hodnotu a předpovědi z času t_n podle vzorců (5.2.10) a (5.2.11).
- Získáme nové pozorování $y_{t_{n+1}}$ v čase $t_{n+1} > t_n$.
- Pomocí rekurentních vzorců (5.2.15) až (5.2.17) a (5.2.19) až (5.2.21) vypočteme postupně z hodnot α_{t_n} , $S_{t_n}^{[p]}$ a ${}^k T_{t_n}^{[p]}$ hodnoty $\alpha_{t_{n+1}}$, $S_{t_{n+1}}^{[p]}$ a ${}^k T_{t_{n+1}}^{[p]}$.
- Položíme $n := n + 1$ a vrátíme se k prvnímu bodu.

Hlavní rozdíl oproti stejné metodě pro pravidelné časové řady je tedy ten, že kromě $m + 1$ vyrovnávacích statistik $S_{t_n}^{[p]}$ musíme v každém kroku rekurentně přepočítávat ještě vyrovnávací koeficient α_{t_n} a koeficienty ${}^k T_{t_n}^{[p]}$ levé strany soustavy (5.2.9). Jejich proměnlivost navíc znamená, že v každém kroku musíme znovu řešit novou soustavu $m + 1$ lineárních rovnic o $m + 1$ neznámých. V případě pravidelné řady s nekonečnou

minulostí jsou zřejmě jak α_{t_n} , tak hlavně ${}^k T_{t_n}^{[p]}$ konstantní. To především značně usnadní řešení uvažované soustavy rovnic, jelikož její levá strana se nemění.

Výpočetní složitost popsaného exponenciálního vyrovnávání pro nepravidelné časové řady pochopitelně závisí především na jeho řádu m . Předně rekurentně přepočítáváme celkem $m^2 + 2m + 2$ hodnot. Navíc v rekurentním vzorci (5.2.19) je nová hodnota ${}^k T_{t_n}^{[p]}(t_{n+1})$ počítána z $k + 1$ hodnot ${}^i T_{t_n}^{[p]}$. To způsobuje, že celkový počet aritmetických operací prováděných při všech rekurentních přepočtech v každém kroku metody je pro $m \rightarrow \infty$ asymptoticky úměrný dokonce m^3 . Zapomenout nesmíme ani na následné řešení soustavy $m + 1$ lineárních rovnic, jehož náročnost pochopitelně také roste s rostoucím řádem m .

Pro $m = 0, 1, 2$, tedy pro jednoduché, dvojité a trojitě exponenciální vyrovnávání, jsou však výsledné vzorce stále ještě poměrně jednoduché. Jejich přehledný zápis bude proveden na konci tohoto odstavce. Exponenciální vyrovnávání s řádem $m \geq 3$ (tedy počínaje kubickým trendem) je již v praxi zřídka využitelné, většinou vystačíme s $m = 1$ (lineární trend) či případně s $m = 2$ (kvadratický trend).

Zabývejme se nyní ještě otázkou nastartování rekurentního výpočtu podle vzorců (5.2.15) až (5.2.17) a (5.2.19) až (5.2.21). Připomeňme, že metoda byla explicitně odvozena pro časovou řadu s konečnou minulostí (v případě nekonečné minulosti by ovšem všechny rekurentní vzorce byly naprosto totožné). Není tedy nic, co by nám bránilo podle vzorců (5.2.2) až (5.2.6) určit hodnoty α_{t_1} , $S_{t_1}^{[p]}$ a ${}^k T_{t_1}^{[p]}$. Konkrétně bude

$$\alpha_{t_1} = 1, \quad S_{t_1}^{[p]} = y_{t_1}, \quad {}^k T_{t_1}^{[p]} = 0 \quad \text{a} \quad {}^0 T_{t_1}^{[p]} = 1 \quad (5.2.22)$$

pro $p = 1, 2, \dots, m + 1$ a $k = 1, 2, \dots, m$. Dále již můžeme pokračovat rekurentním přepočtem z času t_1 do času t_2 atd. Máme-li k dispozici prvních n pozorování řady y , kde $n \geq m + 1$, můžeme takto rekurentně napočítat pomocné hodnoty až do času t_n . Zde již můžeme generovat vyrovnanou hodnotu předpovědi v řadě y . Podmínka $n \geq m + 1$ je zde důležitá proto, aby měla soustava (5.2.9) právě jedno řešení.

Přesto, že s volbou počátečních hodnot pro nastartování rekurentního výpočtu zde není teoreticky žádný problém, ukážeme si možnou volbu počátečních hodnot t_0 , α_{t_0} , $S_{t_0}^{[p]}$ a ${}^k T_{t_0}^{[p]}$, která nám umožní konstruovat vyrovnané hodnoty a předpovědi již z času t_0 , stejně jako u všech předchozích metod popsaných v této diplomové práci. Při volbě těchto počátečních hodnot budeme postupovat analogicky jako Cipra (2006) v případě dvojitého exponenciálního vyrovnávání.

Předně opět označme jako q průměrnou časovou vzdálenost mezi sousedními pozorováními řady y a položíme

$$t_0 = t_1 - q. \quad (5.2.23)$$

$$\alpha_{t_0} = 1 - (1 - \alpha)^q. \quad (5.2.24)$$

Logicky bereme ${}^0T_{t_0}^{[p]} = 1$. Další hodnoty ${}^kT_{t_0}^{[p]}$ pro $k = 1, 2, \dots, m$ a $p = 1, 2, \dots, m + 1$ získáme na základě předpokladu, že nekonečná posloupnost fiktivních pozorování řady y od času t_0 počínaje směrem do minulosti byla pozorovaná s pravidelnými rozestupy délky q . Hodnoty ${}^kT_{t_0}^{[p]}$ by tak bylo možné určit explicitně jako součty nekonečných řad, ve které se promění vzorce (5.2.5) a (5.2.6). To je však technicky zbytečně náročné. Stejně hodnoty ${}^kT_{t_0}^{[p]}$ můžeme mnohem snáze získat jako pevné body rekurentního přepočtu dle vzorců (5.2.19) až (5.2.21) s $\alpha_{t_n} = \alpha_{t_0}$ a $t_{n+1} - t_n = q$. V praxi tento postup ukážeme pro konkrétní řady $m = 1, 2$, odkud snad bude zřejmé i to, že je stejně tak použitelný i pro řady vyšší.

Nakonec hodnoty $S_{t_0}^{[p]}$ pro $p = 1, 2, \dots, m + 1$ určíme jako

$$S_{t_0}^{[p]} = \hat{b}_0(t_0) + \hat{b}_1(t_0) \cdot {}^1T_{t_0}^{[p]} + \hat{b}_2(t_0) \cdot {}^2T_{t_0}^{[p]} + \dots + \hat{b}_m(t_0) \cdot {}^mT_{t_0}^{[p]}, \quad (5.2.25)$$

kde $\hat{b}_0(t_0), \hat{b}_1(t_0), \dots, \hat{b}_m(t_0)$ jsou odhady parametrů b_0, b_1, \dots, b_m v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + b_2(t_0 - t_j)^2 + \dots + b_m(t_0 - t_j)^m + \varepsilon_{t_j} \quad (5.2.26)$$

založené na několika prvních (minimálně $m + 1$) pozorováních řady y , získané například metodou DLS s diskontním faktorem β .

Nyní sepišme přehledně vzorce odvozeného exponenciálního vyrovnávání pro $m = 1$ a $m = 2$, tedy pro dvojité a trojitě exponenciální vyrovnávání ($m = 0$ je jednoduché exponenciální vyrovnávání z odst. 3.2 a pro úsporu místa se jím zde zabývat nebudeme). Pro tyto případy také explicitně uvedeme vzorce pro výpočet počátečních hodnot ${}^kT_{t_0}^{[p]}$ a pokusíme se o *ad hoc* konstrukci předpovědních intervalů. Vzorce (5.2.17), (5.2.23) a (5.2.24), které jsou společné všem metodám bez ohledu na řád m , již nebudeme pro úsporu místa uvádět.

Dvojitě exponenciální vyrovnávání

Případ $m = 1$ pro řady s lokálně lineárním trendem povede ke stejné metodě, jakou je Ciprovo dvojitě exponenciální vyrovnávání pro nepravidelné časové řady, viz. Cipra (2006). Přepočítávat je nyní třeba vyrovnávací koeficient α_{t_n} , první dvě vyrovnávací statistiky $S_{t_n}^{[1]}$ a $S_{t_n}^{[2]}$ a ještě hodnoty ${}^1T_{t_n}^{[1]}$ a ${}^1T_{t_n}^{[2]}$ (index 1 v levém horním rohu budeme pro jednoduchost vynechávat). Příslušné rekurentní vzorce (5.2.15), (5.2.16) a (5.2.19) až (5.2.21) zde vypadají takto:

$$S_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[1]} + \alpha_{t_{n+1}} \cdot y_{t_{n+1}}, \quad (5.2.27)$$

$$S_{t_{n+1}}^{[2]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[2]} + \alpha_{t_{n+1}} \cdot S_{t_{n+1}}^{[1]}, \quad (5.2.28)$$

$$T_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot (T_{t_n}^{[1]} + t_{n+1} - t_n), \quad (5.2.29)$$

$$T_{t_{n+1}}^{[2]} = (1 - \alpha_{t_{n+1}}) \cdot (T_{t_n}^{[2]} + t_{n+1} - t_n) + \alpha_{t_{n+1}} \cdot T_{t_n}^{[1]}. \quad (5.2.30)$$

Soustava (5.2.9) má tvar

$$\begin{aligned}\hat{b}_0(t_n) + \hat{b}_1(t_n) \cdot T_{t_n}^{[1]} &= S_{t_n}^{[1]}, \\ \hat{b}_0(t_n) + \hat{b}_1(t_n) \cdot T_{t_n}^{[2]} &= S_{t_n}^{[2]}.\end{aligned}\quad (5.2.31)$$

s řešením

$$\begin{aligned}\hat{b}_0(t_n) &= \frac{S_{t_n}^{[1]} \cdot T_{t_n}^{[2]} - S_{t_n}^{[2]} \cdot T_{t_n}^{[1]}}{T_{t_n}^{[2]} - T_{t_n}^{[1]}}, \\ \hat{b}_1(t_n) &= \frac{S_{t_n}^{[2]} - S_{t_n}^{[1]}}{T_{t_n}^{[2]} - T_{t_n}^{[1]}}.\end{aligned}\quad (5.2.32)$$

Pro vyrovnanou hodnotu a předpověď o $\tau > 0$ časových jednotek vpřed platí

$$\hat{y}_{t_n+\tau} = \hat{b}_0(t_n) \quad \text{a} \quad \hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) - \tau \cdot \hat{b}_1(t_n). \quad (5.2.33)$$

Počáteční hodnoty $T_{t_0}^{[1]}$ a $T_{t_0}^{[2]}$ získáme jako řešení soustavy, kterou obdržíme ze vzorců (5.2.29) a (5.2.30) dosazením $\alpha_{t_n} = \alpha_{t_0}$ a $t_{n+1} - t_n = q$:

$$\begin{aligned}T_{t_0}^{[1]} &= (1 - \alpha_{t_0}) \cdot (T_{t_0}^{[1]} + q), \\ T_{t_0}^{[2]} &= (1 - \alpha_{t_0}) \cdot (T_{t_0}^{[2]} + q) + \alpha_{t_0} \cdot T_{t_0}^{[1]}.\end{aligned}\quad (5.2.34)$$

Odtud a s použitím vzorce (5.2.24) již velice snadno získáme

$$T_{t_0}^{[1]} = \frac{q\beta^q}{1 - \beta^q} \quad \text{a} \quad T_{t_0}^{[2]} = \frac{2q\beta^q}{1 - \beta^q}. \quad (5.2.35)$$

Počáteční hodnoty $S_{t_0}^{[1]}$ a $S_{t_0}^{[2]}$ dostaneme jako

$$\begin{aligned}S_{t_0}^{[1]} &= \hat{b}_0(t_0) + \hat{b}_1(t_0) \cdot T_{t_0}^{[1]}, \\ S_{t_0}^{[2]} &= \hat{b}_0(t_0) + \hat{b}_1(t_0) \cdot T_{t_0}^{[2]}.\end{aligned}\quad (5.2.36)$$

kde $\hat{b}_0(t_0)$, $\hat{b}_1(t_0)$ jsou odhady parametrů b_0 , b_1 v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + \varepsilon_{t_j}. \quad (5.2.37)$$

Cipra (2006) formuluje dvojité exponenciální vyrovnávání pro nepravidelné časové řady pomocí vyrovnávacího koeficientu α_{t_n} , vyrovnávacích statistik S_{t_n} a $S_{t_n}^{[2]}$ a dvou pomocných statistik w_{t_n} a z_{t_n} . První tři symboly mají shodný význam jako v tomto odstavci, dále pak platí vztahy

$$T_{t_n}^{[1]} = \frac{\alpha_{t_n}}{w_{t_n}} \quad \text{a} \quad T_{t_n}^{[2]} = \frac{\alpha_{t_n}}{z_{t_n}} + \frac{\alpha_{t_n}}{w_{t_n}} . \quad (5.2.38)$$

Zabývejme se nyní otázkou předpovědních intervalů pro dvojitě exponenciální vyrovnávání pro nepravidelné časové řady. Jak bylo řečeno v předchozím odstavci, dvojitě exponenciální vyrovnávání pro pravidelné časové řady (tj. Brownova metoda) je jistým speciálním případem Holtovy metody s lineárním trendem. Na základě vzorců (4.2.8) a (5.1.15) bychom tedy mohli uvažovat např. následující vzorec pro rozptyl chyby předpovědi o $\tau > 0$ časových jednotek vpřed:

$$\text{var}[e_{t_n+\tau}(n)] = \sigma^2 \left\{ \alpha_H^2 (\tau - 1) \left[1 + \gamma_H \cdot \tau + \gamma_H^2 \frac{\tau(2\tau - 1)}{6} \right] + 1 \right. \\ \left. + [\alpha_{H t_n} - \alpha_H]^+ (1 + \alpha_H) + \alpha_H^2 \cdot \tau^2 [\gamma_{H t_n} - \gamma_H]^+ (1 + \gamma_H) \right\} , \quad (5.2.39)$$

kde hodnoty α_H , γ_H , $\alpha_{H t_n}$ a $\gamma_{H t_n}$ jsou dány vzorcí

$$\alpha_H = \alpha(2 - \alpha) , \quad \gamma_H = \frac{\alpha}{2 - \alpha} , \quad \alpha_{H t_n} = \alpha_{t_n}(2 - \alpha_{t_n}) \quad \text{a} \quad \gamma_{H t_n} = \frac{\alpha_{t_n}}{2 - \alpha_{t_n}} . \quad (5.2.40)$$

Trojité exponenciální vyrovnávání

Případ $m = 2$ představuje trojitě exponenciální vyrovnávání pro řady s lokálně kvadratickým trendem. Přepočítávat je nyní třeba vyrovnávací koeficient α_{t_n} , první tři vyrovnávací statistiky $S_{t_n}^{[1]}$, $S_{t_n}^{[2]}$ a $S_{t_n}^{[3]}$ a hodnoty ${}^1T_{t_n}^{[1]}$, ${}^1T_{t_n}^{[2]}$, ${}^1T_{t_n}^{[3]}$, ${}^2T_{t_n}^{[1]}$, ${}^2T_{t_n}^{[2]}$ a ${}^2T_{t_n}^{[3]}$.

Příslušné rekurentní vzorce (5.2.15), (5.2.16) a (5.2.19) až (5.2.21) zde vypadají takto:

$$S_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[1]} + \alpha_{t_{n+1}} \cdot y_{t_{n+1}} , \quad (5.2.41)$$

$$S_{t_{n+1}}^{[2]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[2]} + \alpha_{t_{n+1}} \cdot S_{t_{n+1}}^{[1]} , \quad (5.2.42)$$

$$S_{t_{n+1}}^{[3]} = (1 - \alpha_{t_{n+1}}) \cdot S_{t_n}^{[3]} + \alpha_{t_{n+1}} \cdot S_{t_{n+1}}^{[2]} , \quad (5.2.43)$$

$${}^1T_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot ({}^1T_{t_n}^{[1]} + t_{n+1} - t_n) , \quad (5.2.44)$$

$${}^1T_{t_{n+1}}^{[2]} = (1 - \alpha_{t_{n+1}}) \cdot ({}^1T_{t_n}^{[2]} + t_{n+1} - t_n) + \alpha_{t_{n+1}} \cdot {}^1T_{t_{n+1}}^{[1]} , \quad (5.2.45)$$

$${}^1T_{t_{n+1}}^{[3]} = (1 - \alpha_{t_{n+1}}) \cdot ({}^1T_{t_n}^{[3]} + t_{n+1} - t_n) + \alpha_{t_{n+1}} \cdot {}^1T_{t_{n+1}}^{[2]} , \quad (5.2.46)$$

$${}^2T_{t_{n+1}}^{[1]} = (1 - \alpha_{t_{n+1}}) \cdot \left[{}^2T_{t_n}^{[1]} + 2(t_{n+1} - t_n) \cdot {}^1T_{t_{n+1}}^{[1]} + (t_{n+1} - t_n)^2 \right] , \quad (5.2.47)$$

$${}^2T_{t_{n+1}}^{[2]} = (1 - \alpha_{t_{n+1}}) \cdot \left[{}^2T_{t_n}^{[2]} + 2(t_{n+1} - t_n) \cdot {}^1T_{t_{n+1}}^{[2]} + (t_{n+1} - t_n)^2 \right] + \alpha_{t_{n+1}} \cdot {}^2T_{t_{n+1}}^{[1]} , \quad (5.2.48)$$

$${}^2T_{t_{n+1}}^{[3]} = (1 - \alpha_{t_{n+1}}) \cdot \left[{}^2T_{t_n}^{[3]} + 2(t_{n+1} - t_n) \cdot {}^1T_{t_{n+1}}^{[3]} + (t_{n+1} - t_n)^2 \right] + \alpha_{t_{n+1}} \cdot {}^2T_{t_{n+1}}^{[2]} . \quad (5.2.49)$$

Soustava (5.2.9) má nyní tvar

$$\hat{b}_0(t_n) + \hat{b}_1(t_n) \cdot {}^1T_{t_n}^{[p]} + \hat{b}_2(t_n) \cdot {}^2T_{t_n}^{[p]} = S_{t_n}^{[p]} , \quad p = 1, 2, 3 . \quad (5.2.50)$$

Explicitní vzorce pro její řešení $\hat{b}_0(t_n), \hat{b}_1(t_n), \hat{b}_2(t_n)$ lze získat např. Cramerovým pravidlem, viz. Bartsch (2000), str. 207. Pro vyrovnanou hodnotu a předpověď o $\tau > 0$ časových jednotek vpřed platí

$$\hat{y}_{t_n+\tau} = \hat{b}_0(t_n) \quad \text{a} \quad \hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) - \tau \cdot \hat{b}_1(t_n) + \tau^2 \cdot \hat{b}_2(t_n) . \quad (5.2.51)$$

Počátečních hodnoty ${}^1T_{t_0}^{[1]}$, ${}^1T_{t_0}^{[2]}$, ${}^1T_{t_0}^{[3]}$, ${}^2T_{t_0}^{[1]}$, ${}^2T_{t_0}^{[2]}$ a ${}^2T_{t_0}^{[3]}$ získáme jako řešení soustavy vzniklé ze vzorců (5.2.44) až (5.2.49) po dosazení $\alpha_{t_n} = \alpha_{t_0}$ a $t_{n+1} - t_n = q$:

$$\begin{aligned} {}^1T_{t_0}^{[1]} &= (1 - \alpha_{t_0}) \cdot ({}^1T_{t_0}^{[1]} + q) , \\ {}^1T_{t_0}^{[2]} &= (1 - \alpha_{t_0}) \cdot ({}^1T_{t_0}^{[2]} + q) + \alpha_{t_0} \cdot {}^1T_{t_0}^{[1]} , \\ {}^1T_{t_0}^{[3]} &= (1 - \alpha_{t_0}) \cdot ({}^1T_{t_0}^{[3]} + q) + \alpha_{t_0} \cdot {}^1T_{t_0}^{[2]} , \\ {}^2T_{t_0}^{[1]} &= (1 - \alpha_{t_0}) \cdot [{}^2T_{t_0}^{[1]} + 2q \cdot {}^1T_{t_0}^{[1]} + q^2] , \\ {}^2T_{t_0}^{[2]} &= (1 - \alpha_{t_0}) \cdot [{}^2T_{t_0}^{[2]} + 2q \cdot {}^1T_{t_0}^{[2]} + q^2] + \alpha_{t_0} \cdot {}^2T_{t_0}^{[1]} , \\ {}^2T_{t_0}^{[3]} &= (1 - \alpha_{t_0}) \cdot [{}^2T_{t_0}^{[3]} + 2q \cdot {}^1T_{t_0}^{[3]} + q^2] + \alpha_{t_0} \cdot {}^2T_{t_0}^{[2]} . \end{aligned} \quad (5.2.52)$$

Odtud a s použitím vzorce (5.2.24) již velice snadno získáme

$${}^1T_{t_0}^{[1]} = \frac{q\beta^q}{1 - \beta^q} , \quad {}^1T_{t_0}^{[2]} = \frac{2q\beta^q}{1 - \beta^q} , \quad {}^1T_{t_0}^{[3]} = \frac{3q\beta^q}{1 - \beta^q} , \quad (5.2.53)$$

$${}^2T_{t_0}^{[1]} = \frac{q^2\beta^q(1 + \beta^q)}{(1 - \beta^q)^2} , \quad {}^2T_{t_0}^{[2]} = \frac{q^2\beta^q(2 + 4\beta^q)}{(1 - \beta^q)^2} \quad \text{a} \quad {}^2T_{t_0}^{[3]} = \frac{q^2\beta^q(3 + 9\beta^q)}{(1 - \beta^q)^2} . \quad (5.2.54)$$

Všimněme si, že soustava rovnic (5.2.52) má speciální tvar, díky němuž je její řešení velice snadné. Počátečních hodnoty $S_{t_0}^{[1]}$, $S_{t_0}^{[2]}$ a $S_{t_0}^{[3]}$ dostaneme jako

$$S_{t_0}^{[p]} = \hat{b}_0(t_0) + \hat{b}_1(t_0) \cdot {}^1T_{t_0}^{[p]} + \hat{b}_2(t_0) \cdot {}^2T_{t_0}^{[p]} , \quad p = 1, 2, 3 , \quad (5.2.55)$$

kde $\hat{b}_0(t_0), \hat{b}_1(t_0), \hat{b}_2(t_0)$ jsou odhady parametrů b_0, b_1, b_2 v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + b_2(t_0 - t_j)^2 + \varepsilon_{t_j} . \quad (5.2.56)$$

Pokud jde o rozptyly předpovědních chyb, nemáme se zde již o co opřít. Možný *ad hoc* vzorec pro rozptyl předpovědní chyby o $\tau > 0$ časových jednotek vpřed by mohl být např.

$$\text{var}[e_{n+\tau}(n)] = \sigma^2[\alpha^2 \cdot (\tau^4 - 1) + 1] . \quad (5.2.57)$$

Je volen tak, aby měl podobné vlastnosti jako vzorec (3.2.9) pro jednoduché exponenciální vyrovnávání: pro $\tau = 1$ je uvažovaný rozptyl roven parametru $\sigma^2 > 0$, pro $\tau = 0$ je roven $\sigma^2(1 - \alpha^2)$. Jde o rostoucí funkci proměnné τ a pro $\tau \geq 1$

i o rostoucí funkci proměnné α . Fakt, že pro $\tau \rightarrow \infty$ je $\text{var}[e_{n+\tau}(n)] \approx \sigma^2 \alpha^2 \tau^4$, by měl odrážet vyšší stupeň uvažovaného polynomického trendu.

5.3 DLS odhad polynomického trendu stupně m

V tomto odstavci si ukážeme vyrovnávací a předpovědní metodu pro nepravidelné časové řady s lokálně polynomickým trendem stupně m , založenou na DLS odhadu parametrů tohoto trendu. Jde o druhou z možností, zmíněných v úvodu odstavce 5.2, jak zobecnit exponenciální vyrovnávání řádu m z případu pravidelné na případ nepravidelné časové řady. První takové zobecnění, pracující s vyrovnávacími statistikami S , bylo odvozeno ve zmíněném odstavci 5.2.

Metoda používající DLS odhadu parametrů polynomického trendu bude mít s touto metodou v mnohém podobné vlastnosti. Půjde také o adaptivní metodu, pracující s jednou vyrovnávací konstantou α , resp. s příslušným diskontním faktorem $\beta = 1 - \alpha$. Taktéž se bude jednat o metodu rekurentní a také bude třeba v každém kroku znovu řešit soustavu $m+1$ rovnic pro $m+1$ neznámých parametrů polynomického trendu. Zmíněné dvě metody jsou ekvivalentní v případě pravidelné časové řady (pak jde o klasické exponenciální vyrovnávání) a v případě řádu $m=0$ (lokálně konstantní trend), kdy jsou obě metody totožné s Wrightovým zobecněním jednoduchého exponenciálního vyrovnávání (odst. 3.2). V ostatních případech dávají obě metody v praxi více či méně podobné výsledky.

Začneme stejně jako v odstavci 5.2. Tedy mějme nepravidelnou časovou řadu $y_{t_1}, y_{t_2}, \dots, y_{t_n}, y_{t_{n+1}}, \dots$ pozorovanou v časech $t_1 < t_2 < \dots < t_n < t_{n+1} < \dots$. Necht' $m \in \mathbf{N}$ je stupeň uvažovaného polynomického trendu. Dále buď $n \geq m+1$ a uvažujme model

$$y_{t_j} = b_0 + b_1(t_n - t_j) + b_2(t_n - t_j)^2 + \dots + b_m(t_n - t_j)^m + \varepsilon_{t_j}, \quad j = 1, 2, 3, \dots, \quad (5.3.1)$$

kde $b_0, b_1, \dots, b_m \in \mathbf{R}$ jsou neznámé parametry a náhodné složky ε_{t_j} mají nulovou střední hodnotu. O jejich kovarianční struktuře nečiníme žádné předpoklady. Uvažujme daný *diskontní faktor* $\beta \in (0, 1)$ a k němu *vyrovnávací konstantu* $\alpha = 1 - \beta$.

Neznámé parametry b_0, b_1, \dots, b_m modelu (5.3.1) odhadneme z prvních n pozorování řady y metodou DLS s diskontním faktorem β . Příslušná soustava normálních rovnic pro odhady $\hat{b}_0(t_n), \hat{b}_1(t_n), \dots, \hat{b}_m(t_n)$ má tvar

$$b_0 \cdot T_{t_n}^{(k)} + b_1 \cdot T_{t_n}^{(k+1)} + \dots + b_m \cdot T_{t_n}^{(k+m)} = Y_{t_n}^{(k)}, \quad k = 0, 1, 2, \dots, m, \quad (5.3.2)$$

kde jsme označili

$$T_{t_n}^{(k)} = \sum_{j=1}^n \beta^{t_n-t_j} (t_n - t_j)^k, \quad k = 0, 1, 2, \dots, 2m, \quad (5.3.3)$$

$$Y_{t_n}^{(k)} = \sum_{j=1}^n \beta^{t_n-t_j} y_{t_j} (t_n - t_j)^k, \quad k = 0, 1, 2, \dots, m. \quad (5.3.4)$$

Za předpokladů $n \geq m+1$ a $t_1 < t_2 < \dots < t_n$ má tato soustava právě jedno řešení. Vyrovnanou hodnotu řady y v čase n a předpověď z času n o $\tau > 0$ časových jednotek vpřed získáme pak opět jako

$$\hat{y}_{t_n} = \hat{b}_0(t_n), \quad (5.3.5)$$

$$\hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) + \hat{b}_1(t_n) \cdot (-\tau) + \hat{b}_2(t_n) \cdot (-\tau)^2 + \dots + \hat{b}_m(t_n) \cdot (-\tau)^m. \quad (5.3.6)$$

Protože DLS odhady $\hat{b}_0(t_n), \hat{b}_1(t_n), \dots, \hat{b}_m(t_n)$ parametrů b_0, b_1, \dots, b_m jsou nestranné, jsou nestranné i výše uvedená vyrovnaná hodnota a předpovědi, tedy platí

$$E(\hat{y}_{t_n}) = E(y_{t_n}) \quad \text{a} \quad E[\hat{y}_{t_n+\tau}(t_n)] = E(y_{t_n+\tau}). \quad (5.3.7)$$

Získáme-li nové pozorování $y_{t_{n+1}}$, posuneme se z času t_n do času t_{n+1} a budeme odhadovat stejnou DLS metodou parametry b_0, b_1, \dots, b_m v aktualizovaném modelu

$$y_{t_j} = b_0 + b_1(t_{n+1} - t_j) + b_2(t_{n+1} - t_j)^2 + \dots + b_m(t_{n+1} - t_j)^m + \varepsilon_{t_j}, \quad j = 1, 2, 3, \dots. \quad (5.3.8)$$

Tyto odhady $\hat{b}_0(t_{n+1}), \hat{b}_1(t_{n+1}), \dots, \hat{b}_m(t_{n+1})$ získáme řešením soustavy (5.3.2) posunuté do času t_{n+1} , tedy soustavy

$$b_0 T_{t_{n+1}}^{(k)} + b_1 T_{t_{n+1}}^{(k+1)} + \dots + b_m T_{t_{n+1}}^{(k+m)} = Y_{t_{n+1}}^{(k)}, \quad k = 0, 1, 2, \dots, m. \quad (5.3.9)$$

V následujících odstavcích odvodíme rekurentní formule, které nám umožní získat koeficienty nové soustavy (5.3.9) pomocí koeficientů původní soustavy (5.3.2).

Nejprve pro dané $k = 1, 2, \dots, 2m$ počítejme:

$$\begin{aligned} T_{t_{n+1}}^{(k)} &= \sum_{j=1}^{n+1} \beta^{t_{n+1}-t_j} (t_{n+1} - t_j)^k = \sum_{j=1}^n \beta^{t_{n+1}-t_j} [(t_n - t_j) + (t_{n+1} - t_n)]^k = \\ &= \sum_{j=1}^n \left[\beta^{t_{n+1}-t_j} \sum_{i=0}^k \binom{k}{i} (t_{n+1} - t_n)^{k-i} (t_n - t_j)^i \right] = \\ &= \beta^{t_{n+1}-t_n} \sum_{i=0}^k \left[\binom{k}{i} (t_{n+1} - t_n)^{k-i} \cdot \sum_{j=1}^n \beta^{t_n-t_j} (t_n - t_j)^i \right] = \\ &= \beta^{t_{n+1}-t_n} \sum_{i=0}^k \binom{k}{i} (t_{n+1} - t_n)^{k-i} T_{t_n}^{(i)}. \end{aligned} \quad (5.3.10)$$

Navíc platí

$$T_{t_{n+1}}^{(0)} = \sum_{j=1}^{n+1} \beta^{t_{n+1}-t_j} = 1 + \beta^{t_{n+1}-t_n} \sum_{j=1}^n \beta^{t_n-t_j} = 1 + \beta^{t_{n+1}-t_n} \cdot T_{t_n}^{(0)}. \quad (5.3.11)$$

Úplně stejně odvodíme rekurentní vzorce pro výpočet $Y_{t_{n+1}}^{(k)}$. Pro $k = 1, 2, \dots, m$ platí

$$\begin{aligned} Y_{t_{n+1}}^{(k)} &= \sum_{j=1}^{n+1} \beta^{t_{n+1}-t_j} y_{t_j} (t_{n+1} - t_j)^k = \\ &= \beta^{t_{n+1}-t_n} \sum_{i=0}^k \left[\binom{k}{i} (t_{n+1} - t_n)^{k-i} \cdot \sum_{j=1}^n \beta^{t_n-t_j} y_{t_j} (t_n - t_j)^i \right] = \\ &= \beta^{t_{n+1}-t_n} \sum_{i=0}^k \binom{k}{i} (t_{n+1} - t_n)^{k-i} Y_{t_n}^{(i)} \end{aligned} \quad (5.3.12)$$

a k tomu

$$Y_{t_{n+1}}^{(0)} = \sum_{j=1}^{n+1} y_{t_j} \beta^{t_{n+1}-t_j} = y_{t_{n+1}} + \beta^{t_{n+1}-t_n} \sum_{j=1}^n y_{t_j} \beta^{t_n-t_j} = y_{t_{n+1}} + \beta^{t_{n+1}-t_n} \cdot Y_{t_n}^{(0)}. \quad (5.3.13)$$

Schéma praktického použití této metody je stejné jako v případě exponenciálního vyrovnávání pro nepravidelné časové řady z odstavce 5.2. Výpočetní složitost metody pochopitelně opět závisí především na číslu m . Rekurentně zde přepočítáváme celkem $3m + 2$ hodnot. Vzhledem ke tvaru vzorců (5.3.10) a (5.3.12) je celkový počet aritmetických operací prováděných při všech rekurentních přepočtech v jednom kroku metody pro $m \rightarrow \infty$ asymptoticky úměrný m^2 . Pro $m \leq 1$ mají obě metody stejnou výpočetní složitost, pro $m \geq 2$ vychází z tohoto srovnání lépe metoda založená na DLS odhadu.

Ta byla odvozena pro nepravidelnou časovou řadu s konečnou minulostí, ovšem stejně tak by mohla být odvozena pro nepravidelnou časovou řadu s nekonečnou minulostí - výsledné rekurentní vzorce by byly naprosto totožné. Pokud jde o nastartování výpočtu podle rekurentních vzorců (5.3.10) až (5.3.13), nic nám nebrání podle vzorců (5.3.3) a (5.3.4) určit počáteční hodnoty $T_{t_1}^{(k)}$ a $Y_{t_1}^{(k)}$ a dále již pokračovat rekurentním přepočtem z času t_1 do času t_2 atd. Konkrétně bude

$$T_{t_1}^{(0)} = 1, \quad T_{t_1}^{(k)} = 0, \quad Y_{t_1}^{(0)} = y_{t_1} \quad \text{a} \quad Y_{t_1}^{(k)} = 0 \quad (5.3.14)$$

pro $k \geq 1$. Máme-li k dispozici prvních n pozorování řady y , kde $n \geq m + 1$, můžeme takto rekurentně napočítat pomocné hodnoty až do času t_n . Zde již můžeme generovat vyrovnanou hodnotu předpovědi v řadě y . Podmínka $n \geq m + 1$ je zde důležitá proto, aby měla soustava (5.3.2) právě jedno řešení.

Stejně jako v odstavci 5.2 si ukážeme možnou volbu počátečních hodnot t_0 , $T_{t_0}^{(k)}$ a $Y_{t_0}^{(k)}$, která nám umožní konstruovat vyrovnané hodnoty a předpovědi již z času t_0 . Postup bude analogický. Položme opět

$$t_0 = t_1 - q, \quad (5.3.15)$$

kde q je průměrná časová vzdálenost mezi sousedními pozorováními řady y . Hodnoty $T_{t_0}^{(k)}$ pro $k = 0, 1, 2, \dots, 2m$ získáme na základě předpokladu, že nekonečná posloupnost fiktivních pozorování řady y od času t_0 počínaje směrem do minulosti byla pozorovaná s pravidelnými časovými rozestupy délky q . Tedy vezmeme

$$T_{t_0}^{(k)} = \sum_{j=0}^{\infty} (jq)^k \beta^{jq}, \quad k = 0, 1, 2, \dots, 2m. \quad (5.3.16)$$

Označíme-li $T_k(x) = \sum_{j=0}^{\infty} j^k x^j$ pro $|x| < 1$ a $k \geq 0$, můžeme psát

$$T_{t_0}^{(k)} = q^k \cdot T_k(\beta^q), \quad (5.3.17)$$

Hodnoty $T_k(x)$ můžeme počítat rekurentně. Pro $k \geq 0$ platí

$$\begin{aligned} T_{k+1}(x) &= \sum_{j=0}^{\infty} j^{k+1} x^j = \sum_{l=0}^{\infty} \left\{ [(l+1)^{k+1} - l^{k+1}] \sum_{i=0}^{\infty} x^{i+l+1} \right\} = \\ &= \frac{x}{1-x} \sum_{l=0}^{\infty} [(l+1)^{k+1} - l^{k+1}] x^l = \frac{x}{1-x} \sum_{l=0}^{\infty} \left[\sum_{i=0}^k \binom{k}{i} l^i \cdot x^l \right] = \\ &= \frac{x}{1-x} \sum_{i=0}^k \left[\binom{k}{i} \sum_{l=0}^{\infty} l^i x^l \right] = \frac{x}{1-x} \sum_{i=0}^k \binom{k}{i} T_i(x). \end{aligned} \quad (5.3.18)$$

K tomu je

$$T_0(x) = \sum_{j=0}^{\infty} x^j = \frac{1}{1-x}. \quad (5.3.19)$$

Počáteční hodnoty $Y_{t_0}^{(k)}$ pro $k = 0, 1, 2, \dots, m$ určíme jako

$$Y_{t_0}^{(k)} = \hat{b}_0(t_0) \cdot T_{t_0}^{(k)} + \hat{b}_1(t_0) \cdot T_{t_0}^{(k+1)} + \dots + \hat{b}_m(t_0) \cdot T_{t_0}^{(k+m)}, \quad (5.3.20)$$

kde $\hat{b}_0(t_0), \hat{b}_1(t_0), \dots, \hat{b}_m(t_0)$ jsou odhady parametrů b_0, b_1, \dots, b_m v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + b_2(t_0 - t_j)^2 + \dots + b_m(t_0 - t_j)^m + \varepsilon_{t_j} \quad (5.3.21)$$

založené na několika prvních (minimálně $m+1$) pozorováních řady y , získané např. metodou DLS s diskontním faktorem β .

Je zřejmé, že případ $m=0$ popisované metody je ekvivalentní Wrightovu jednoduchému exponenciálnímu vyrovnávání pro nepravidelné časové řady a nebudeme se jím tedy již zabývat. Sepíšeme ale přehledně vzorce popisované metody pro $m=1, 2$, tedy pro lokálně lineární a lokálně kvadratický trend. Případ $m=1$

povede na metodu podobnou (nikoli však totožnou) dvojitému exponenciálnímu vyrovnávání z odstavce 5.2, podobně pro $m = 2$ a trojitě exponenciální vyrovnávání. Pro $m = 1, 2$ můžeme v praxi použít stejné předpovědní intervaly jako u exponenciálního vyrovnávání příslušného řádu v odstavci 5.2.

Lokálně lineární trend

V případě $m = 1$ přepočítáváme rekurentně hodnoty $T_{t_n}^{(0)}$, $T_{t_n}^{(1)}$, $T_{t_n}^{(2)}$, $Y_{t_n}^{(0)}$ a $Y_{t_n}^{(1)}$ podle vzorců

$$T_{t_{n+1}}^{(0)} = 1 + \beta^{t_{n+1}-t_n} \cdot T_{t_n}^{(0)}, \quad (5.3.22)$$

$$T_{t_{n+1}}^{(1)} = \beta^{t_{n+1}-t_n} [T_{t_n}^{(1)} + T_{t_n}^{(0)}(t_{n+1} - t_n)], \quad (5.3.23)$$

$$T_{t_{n+1}}^{(2)} = \beta^{t_{n+1}-t_n} [T_{t_n}^{(2)} + 2T_{t_n}^{(1)}(t_{n+1} - t_n) + T_{t_n}^{(0)}(t_{n+1} - t_n)^2], \quad (5.3.24)$$

$$Y_{t_{n+1}}^{(0)} = y_{t_{n+1}} + \beta^{t_{n+1}-t_n} \cdot Y_{t_n}^{(0)}, \quad (5.3.25)$$

$$Y_{t_{n+1}}^{(1)} = \beta^{t_{n+1}-t_n} [Y_{t_n}^{(1)} + Y_{t_n}^{(0)}(t_{n+1} - t_n)]. \quad (5.3.26)$$

Soustava (5.3.2) má tvar

$$\begin{aligned} b_0 \cdot T_{t_n}^{(0)} + b_1 \cdot T_{t_n}^{(1)} &= Y_{t_n}^{(0)}, \\ b_0 \cdot T_{t_n}^{(1)} + b_1 \cdot T_{t_n}^{(2)} &= Y_{t_n}^{(1)} \end{aligned} \quad (5.3.27)$$

s řešením

$$\hat{b}_0(t_n) = \frac{Y_{t_n}^{(0)}T_{t_n}^{(2)} - Y_{t_n}^{(1)}T_{t_n}^{(1)}}{T_{t_n}^{(0)}T_{t_n}^{(2)} - (T_{t_n}^{(1)})^2}, \quad \hat{b}_1(t_n) = \frac{Y_{t_n}^{(1)}T_{t_n}^{(0)} - Y_{t_n}^{(0)}T_{t_n}^{(1)}}{T_{t_n}^{(0)}T_{t_n}^{(2)} - (T_{t_n}^{(1)})^2}. \quad (5.3.28)$$

Pro vyrovnanou hodnotu v čase n a předpověď o $\tau > 0$ časových jednotek vpřed platí

$$\hat{y}_{t_n} = \hat{b}_0(t_n), \quad \hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) - \tau \cdot \hat{b}_1(t_n). \quad (5.3.29)$$

Počáteční hodnoty jsou

$$T_{t_0}^{(0)} = \frac{1}{1 - \beta^q}, \quad T_{t_0}^{(1)} = \frac{q\beta^q}{(1 - \beta^q)^2}, \quad T_{t_0}^{(2)} = \frac{q^2\beta^q(1 + \beta^q)}{(1 - \beta^q)^3}, \quad (5.3.30)$$

$$\begin{aligned} Y_{t_0}^{(0)} &= \hat{b}_0(t_0) \cdot T_{t_0}^{(0)} + \hat{b}_1(t_0) \cdot T_{t_0}^{(1)}, \\ Y_{t_0}^{(1)} &= \hat{b}_0(t_0) \cdot T_{t_0}^{(1)} + \hat{b}_1(t_0) \cdot T_{t_0}^{(2)}, \end{aligned} \quad (5.3.31)$$

kde $\hat{b}_0(t_0)$, $\hat{b}_1(t_0)$ jsou odhady parametrů b_0 , b_1 v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + \varepsilon_{t_j}. \quad (5.3.32)$$

Lokálně kvadratický trend

V případě $m = 2$ přepočítáváme rekurentně hodnoty $T_{t_n}^{(0)}$ až $T_{t_n}^{(4)}$ a $Y_{t_n}^{(0)}$ až $Y_{t_n}^{(2)}$:

$$T_{t_{n+1}}^{(0)} = 1 + \beta^{t_{n+1}-t_n} \cdot T_{t_n}^{(0)} , \quad (5.3.33)$$

$$T_{t_{n+1}}^{(1)} = \beta^{t_{n+1}-t_n} [T_{t_n}^{(1)} + T_{t_n}^{(0)}(t_{n+1} - t_n)] , \quad (5.3.34)$$

$$T_{t_{n+1}}^{(2)} = \beta^{t_{n+1}-t_n} [T_{t_n}^{(2)} + 2T_{t_n}^{(1)}(t_{n+1} - t_n) + T_{t_n}^{(0)}(t_{n+1} - t_n)^2] , \quad (5.3.35)$$

$$T_{t_{n+1}}^{(3)} = \beta^{t_{n+1}-t_n} [T_{t_n}^{(3)} + 3T_{t_n}^{(2)}(t_{n+1} - t_n) + 3T_{t_n}^{(1)}(t_{n+1} - t_n)^2 + T_{t_n}^{(0)}(t_{n+1} - t_n)^3] , \quad (5.3.36)$$

$$T_{t_{n+1}}^{(4)} = \beta^{t_{n+1}-t_n} \left[T_{t_n}^{(4)} + 4T_{t_n}^{(3)}(t_{n+1} - t_n) + 6T_{t_n}^{(2)}(t_{n+1} - t_n)^2 + \right. \\ \left. + 4T_{t_n}^{(1)}(t_{n+1} - t_n)^3 + T_{t_n}^{(0)}(t_{n+1} - t_n)^4 \right] , \quad (5.3.37)$$

$$Y_{t_{n+1}}^{(0)} = y_{t_{n+1}} + \beta^{t_{n+1}-t_n} \cdot Y_{t_n}^{(0)} , \quad (5.3.38)$$

$$Y_{t_{n+1}}^{(1)} = \beta^{t_{n+1}-t_n} [Y_{t_n}^{(1)} + Y_{t_n}^{(0)}(t_{n+1} - t_n)] , \quad (5.3.39)$$

$$Y_{t_{n+1}}^{(2)} = \beta^{t_{n+1}-t_n} [Y_{t_n}^{(2)} + 2Y_{t_n}^{(1)}(t_{n+1} - t_n) + Y_{t_n}^{(0)}(t_{n+1} - t_n)^2] . \quad (5.3.40)$$

Soustava (5.3.2) má tvar

$$b_0 \cdot T_{t_n}^{(k)} + b_1 \cdot T_{t_n}^{(k+1)} + b_2 \cdot T_{t_n}^{(k+2)} = Y_{t_n}^{(k)} , \quad k = 0, 1, 2 . \quad (5.3.41)$$

Explicitní vzorce pro její řešení $\hat{b}_0(t_n), \hat{b}_1(t_n), \hat{b}_2(t_n)$ lze získat např. Cramerovým pravidlem, viz. Bartsch (2000), str. 207. Pro vyrovnané hodnoty a předpovědi platí

$$\hat{y}_{t_n+\tau} = \hat{b}_0(t_n) \quad \text{a} \quad \hat{y}_{t_n+\tau}(t_n) = \hat{b}_0(t_n) - \tau \cdot \hat{b}_1(t_n) + \tau^2 \cdot \hat{b}_2(t_n) . \quad (5.3.42)$$

Počáteční hodnoty jsou

$$T_{t_0}^{(0)} = \frac{1}{1 - \beta^q} , \quad T_{t_0}^{(1)} = \frac{q\beta^q}{(1 - \beta^q)^2} , \quad T_{t_0}^{(2)} = \frac{q^2\beta^q(1 + \beta^q)}{(1 - \beta^q)^3} , \quad (5.3.43)$$

$$T_{t_0}^{(3)} = \frac{q^3\beta^q(1 + 4\beta^q + \beta^{2q})}{(1 - \beta^q)^4} , \quad T_{t_0}^{(4)} = \frac{q^4\beta^q(1 + 11\beta^q + 11\beta^{2q} + \beta^{3q})}{(1 - \beta^q)^5} , \quad (5.3.44)$$

$$Y_{t_0}^{(k)} = \hat{b}_0(t_0) \cdot T_{t_0}^{(k)} + \hat{b}_1(t_0) \cdot T_{t_0}^{(k+1)} + \hat{b}_2(t_0) \cdot T_{t_0}^{(k+2)} , \quad k = 0, 1, 2 , \quad (5.3.45)$$

kde $\hat{b}_0(t_0), \hat{b}_1(t_0), \hat{b}_2(t_0)$ jsou odhady parametrů b_0, b_1, b_2 v modelu

$$y_{t_j} = b_0 + b_1(t_0 - t_j) + b_2(t_0 - t_j)^2 + \varepsilon_{t_j} . \quad (5.3.46)$$

6 Některé výpočetní aspekty metod

V této kapitole se budeme zabývat problémy spojenými s praktickou aplikací vyrovnávacích a předpovědních metod prezentovaných v této diplomové práci. V odstavci 6.1 popíšeme volbu vyrovnávacích konstant a jiných parametrů metodou maximální věrohodnosti. V odstavci 6.2 se budeme zabývat různými mírami přesnosti a adekvátnosti použití zvolené předpovědní metody na konkrétní časovou řadu. Odstavec 6.3 bude věnována použití transformací na časové řady. Nakonec do odstavce 6.4 budou soustředěny některé autorovi zkušenosti s používáním dotyčných metod. Diskutována zde bude též otázka volby časového měřítka.

6.1 Odhad parametrů metodou maximální věrohodnosti

Všechny předpovědní metody vyskytující se v této diplomové práci mají jeden či více číselných parametrů. Jsou jimi vyrovnávací konstanty α , γ či δ , případně tlumící konstanta φ . Po otázce výběru nejvhodnější předpovědní metody tak vyvstává neméně důležitá otázka výběru nejvhodnějších hodnot jejích parametrů. Klasická volba parametrů minimalizací např. MSE kritéria má v případě nepravidelné časové řady tu nevýhodu, že nezohledňuje různě dlouhé předpovědní kroky $t_j - t_{j-1}$. Pokud nám jde o nalezení „skutečných“ hodnot parametrů, které budou optimální z hlediska předpovídání budoucích pozorování, má informace o časové struktuře dosavadních pozorování také svůj přínos.

Máme-li pro danou metodu k dispozici vzorec pro rozptyly předpovědních chyb a předpokládáme-li nějaký typ rozdělení těchto chyb, můžeme její parametry odhadnout metodou maximální věrohodnosti. Uvažujme předpovědní metodu s k -rozměrným parametrem $\alpha \in A$, kde $A \subseteq \mathbf{R}^k$ je množina přípustných hodnot. Budeme volit hodnotu $\hat{\alpha}$ tohoto parametru na základě pozorovaných předpovědních chyb $e_{t_1}, e_{t_2}, \dots, e_{t_n}$ v časech t_1, t_2, \dots, t_n . Označme přesněji jako $e_{t_1}(\alpha), e_{t_2}(\alpha), \dots, e_{t_n}(\alpha)$ hodnoty těchto chyb, které budou zaznamenány při použití dané hodnoty parametru $\alpha \in A$. Necht' pro skutečnou hodnotu parametru α platí

$$E[e_{t_j}(\alpha)] = 0 \quad \text{a} \quad \text{var}[e_{t_j}(\alpha)] = \sigma^2 v_{t_j}(\alpha), \quad j = 1, 2, \dots, n, \quad (6.1.1)$$

kde $v_{t_j}(\alpha) > 0$ a $\sigma^2 > 0$ je další neznámý parametr. Abychom mohli dospět ke tvaru věrohodnostní funkce, potřebujeme pracovat s konkrétním rozdělením chyb $e_{t_j}(\alpha)$. Předpokládejme tedy například, že toto rozdělení je normální, tedy podle (6.1.1) platí

$$e_{t_j}(\boldsymbol{\alpha}) \sim N(0, \sigma^2 v_{t_j}(\boldsymbol{\alpha})) . \quad (6.1.2)$$

Bylo by pochopitelně možné na tomto místě uvažovat i jiný typ rozdělení. Dále předpokládejme, že pro skutečnou hodnotu parametru $\boldsymbol{\alpha}$ jsou předpovědní chyby $e_{t_j}(\boldsymbol{\alpha})$ vzájemně nekorelované. Nyní již můžeme napsat příslušnou věrohodnostní funkci:

$$L(\boldsymbol{\alpha}, \sigma^2) = (2\pi\sigma^2)^{\frac{n}{2}} \left[\prod_{j=1}^n v_{t_j}(\boldsymbol{\alpha}) \right]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^n \frac{e_{t_j}^2(\boldsymbol{\alpha})}{v_{t_j}(\boldsymbol{\alpha})} \right\} . \quad (6.1.3)$$

Jejím zlogaritmováním dostaneme logaritmickou věrohodnostní funkci l :

$$l(\boldsymbol{\alpha}, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2} \sum_{j=1}^n \ln[v_{t_j}(\boldsymbol{\alpha})] - \frac{1}{2\sigma^2} \sum_{j=1}^n \frac{e_{t_j}^2(\boldsymbol{\alpha})}{v_{t_j}(\boldsymbol{\alpha})} . \quad (6.1.4)$$

Řešíme tedy úlohu

$$\min_{\boldsymbol{\alpha} \in A, \sigma^2 > 0} \left\{ n \cdot \ln \sigma^2 + \sum_{j=1}^n \ln[v_{t_j}(\boldsymbol{\alpha})] + \frac{1}{\sigma^2} \cdot \sum_{j=1}^n \frac{e_{t_j}^2(\boldsymbol{\alpha})}{v_{t_j}(\boldsymbol{\alpha})} \right\} . \quad (6.1.5)$$

Pro maximálně věrohodné odhady $\hat{\boldsymbol{\alpha}}, \hat{\sigma}^2$ pak platí

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n \frac{e_{t_j}^2(\hat{\boldsymbol{\alpha}})}{v_{t_j}(\hat{\boldsymbol{\alpha}})} , \quad (6.1.6)$$

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha} \in A} \left\{ \ln \sum_{j=1}^n \frac{e_{t_j}^2(\boldsymbol{\alpha})}{v_{t_j}(\boldsymbol{\alpha})} + \frac{1}{n} \sum_{j=1}^n \ln[v_{t_j}(\boldsymbol{\alpha})] \right\} , \quad (6.1.7)$$

srovnej s Prášková (2004), str. 148, 149.

Jak je vidět v (6.1.7), chyby $e_{t_j}(\boldsymbol{\alpha})$ mají v minimalizovaném výrazu váhu nepřímo úměrnou velikosti příslušného rozptylového faktoru $v_{t_j}(\boldsymbol{\alpha})$. Dotyčnou minimalizaci je třeba provést numericky. Ve vzorci (6.1.6) je parametr σ^2 odhadován jako průměr druhých mocnin veličin $\tilde{e}_{t_j} = e_{t_j} / \sqrt{v_{t_j}} = e_{t_j}(\hat{\boldsymbol{\alpha}}) / \sqrt{v_{t_j}(\hat{\boldsymbol{\alpha}})}$, pro které podle (6.1.2) platí

$$\tilde{e}_{t_j} \sim N(0, \sigma^2) . \quad (6.1.8)$$

Říkat jim budeme *normalizované předpovědní chyby* a jelikož (na rozdíl od předpovědních chyb e_{t_j}) tvoří bílý šum, najdou uplatnění v testech adekvátnosti použití dané předpovědní metody, viz. odstavec 6.2.

V případě pravidelné časové řady můžeme bez újmy na obecnosti předpokládat $v_{t_j} \equiv 1$ a vzorce (6.1.6) a (6.1.7) se zjednoduší na

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n e_{t_j}^2(\hat{\alpha}) \quad \text{a} \quad \hat{\alpha} = \arg \min_{\alpha \in A} \sum_{j=1}^n e_{t_j}^2(\alpha), \quad (6.1.9)$$

tedy jde o minimalizaci klasického MSE kritéria.

6.2 Míry přesnosti a adekvátnosti předpovědních metod

Při použití dané předpovědní metody (s danými hodnotami jejích parametrů) na konkrétní časovou řadu nás vždy zajímá, jak přesné předpovědi jsme získali a zda-li bylo použití této metody na dotyčnou časovou řadu adekvátní. Kvantifikovat nějakým způsobem přesnost předpovědí je důležité jednak pro porovnání různých metod použitých na stejnou časovou řadu, tak pro srovnání výsledků mezi různými časovými řadami. Pokud zjistíme, že použití dané předpovědní metody na zkoumanou časovou řadu nebylo adekvátní, znamená to většinou, že existuje jiná předpovědní metoda, pomocí níž můžeme dostat znatelně lepší výsledky.

Jednou z možností jak posoudit přesnost a adekvátnost předpovědní metody je subjektivní vyhodnocení grafu pozorování zkoumané řasové řady a příslušných předpovědí získaných použitím této metody. Z tohoto grafu bývají na první pohled patrné všechny základní problémy předpovědních metod: pomalá adaptace na změnu trendu v časové řadě, zachycení falešných trendů, příliš citlivá reakce na náhodné fluktuace apod. Většinou se hned nabízí i řešení spočívající ve změně vyrovnávacích konstant či jiných parametrů nebo dokonce ve volbě úplně jiné předpovědní metody.

Důležitou roli v této analýze hrají pochopitelně různé kvantitativní míry přesnosti předpovědí a statistické testy adekvátnosti použití metod. Zabývejme se nejprve tím prvním, tedy mírami přesnosti prováděných předpovědí*. Mějme pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$ řady y v časech $t_1 > t_2 > \dots > t_n$ a příslušné bodové předpovědi $\hat{y}_{t_1}(t_0), \hat{y}_{t_2}(t_1), \dots, \hat{y}_{t_n}(t_{n-1})$. Jako obvykle definujeme předpovědní chybu

$$e_{t_j} = y_{t_j} - \hat{y}_{t_j}(t_{j-1}). \quad (6.2.1)$$

Mírou nepřesnosti získaných předpovědí je nějak kvantifikovaná velikost vektoru $(e_{t_1}, e_{t_2}, \dots, e_{t_n})$. Nejpoužívanější je jistě tzv. střední čtvercová chyba (*mean square error*):

$$MSE = \frac{1}{n} \sum_{j=1}^n e_{t_j}^2. \quad (6.2.2)$$

Jde vlastně o odhad rozptylu předpovědních chyb, pokud zanedbáme fakt, že v případě nepravidelné časové řady není rozptyl jednotlivých předpovědních chyb e_{t_j} konstantní, ale závisí přinejmenším na délce příslušného předpovědního kroku $t_j - t_{j-1}$.

* Někdy budeme naopak hovořit o mírách *nepřesnosti* předpovědí, což je ale v důsledku totéž.

Nevýhodou MSE je fakt, že její jednotka je druhou mocninou jednotky, ve které jsou měřeny pozorování řady y . Tento nedostatek se řeší odmocněním MSE, čímž dostáváme RMSE (*root mean square error*), obdobu směrodatné odchylky:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n e_{t_j}^2} . \quad (6.2.3)$$

Statistika RMSE je již vyjádřena ve stejných jednotkách jako hodnoty řady y a lze tedy její hodnotu s hodnotami řady y lépe poměřovat. Podobné vlastnosti jako RMSE má i tzv. střední absolutní chyba (*mean absolute error*):

$$MAE = \frac{1}{n} \sum_{j=1}^n |e_{t_j}| . \quad (6.2.4)$$

Hlavní rozdíl je v tom, že MSE resp. RMSE jsou díky kvadratické ztrátové funkci citlivější na v absolutní hodnotě velké hodnoty e_{t_j} . Vždy platí $RMSE \geq MAE$, pro normálně rozdělené předpovědní chyby je $RMSE \approx 1,25 \cdot MAE$.

MSE ani RMSE však nejsou příliš vhodné pro porovnání přesnosti předpovědi v různých časových řadách, jelikož jejich hodnoty závisí např. na měřítku dotyčných řad. Z tohoto pohledu jsou užitečné bezrozměrné míry nepřesnosti předpovědi. Pro řadu y s kladnými hodnotami, u níž má hodnota 0 rozumný význam, definujeme střední absolutní procentuální chybu (*mean absolute percentage error*):

$$MAPE = \frac{1}{n} \sum_{j=1}^n |e_{t_j} / y_{t_j}| . \quad (6.2.5)$$

Statistika MAPE je bezrozměrná a lze ji vyjádřit v procentech.

Výše definované statistiky kvantifikují pouze míru nepřesnosti resp. přesnosti získaných předpovědí. Neříkají však už nic o tom, do jaké míry je kvalita předpovědi výsledkem efektivní předpovědní metody či jen prostého faktu, že zkoumaná časová řada se chová předvídatelně. Pokud např. řada y kolísá v rozmezí 95 až 105, pak předpovědní metoda s $MAPE = 5\%$ je efektivní jen na první pohled.

Vyjádřit efektivitu předpovědní metody, tedy míru, s jakou přispívá ke kvalitním předpovědím, lze statistikami obdobnými koeficientu determinace v lineární regresi. Klasický koeficient determinace

$$R^2 = 1 - \frac{SSR}{SST} \quad (6.2.6)$$

vyjadřuje, o kolik procent je nižší součet čtvercových odchylek v použitém regresním modelu (SSR) oproti součtu čtvercových odchylek v modelu s pouze konstantním členem (SST).

V případě měření efektivity předpovědní metody stačí zvolit jistou míru nepřesnosti předpovědi M a nějakou triviální předpovědní metodu, s níž budeme naši metodu poměřovat. Koeficient determinace R^2 (*R-Squared*) pak získáme jako

$$R^2 = 1 - \frac{M_1}{M_0}, \quad (6.2.7)$$

kde M_1 resp. M_0 je míra nepřesnosti dosažená posuzovanou resp. referenční metodou. Platí $R^2 \in (-\infty, 0]$, hodnota R^2 blízká 1 znamená efektivní předpovědní metodu, hodnota R^2 blízká 0 či dokonce záporná ukazuje na nízkou efektivitu předpovědní metody.

Za míru M budeme brát konkrétně MSE. Jako referenční předpovědní metody budeme uvažovat tyto čtyři:

- $\hat{y}_{t_j}(t_{j-1}) = \bar{y}$ (*model konstantní úrovně*)
- $\hat{y}_{t_j}(t_{j-1}) = y_{t_{j-1}}$ (*model náhodné procházky*)
- $\hat{y}_{t_j}(t_{j-1}) = \hat{a} + \hat{b} \cdot t_j$ (*globální lineární trend*)
- $\hat{y}_{t_j}(t_{j-1}) = y_{t_{j-1}} + \frac{y_{t_{j-1}} - y_{t_{j-2}}}{t_{j-1} - t_{j-2}}(t_j - t_{j-1})$ (*dvoubodová lineární extrapolace*)

Zde \bar{y} je průměr pozorování řady y a \hat{a}, \hat{b} jsou OLS odhady parametrů lineárního trendu. Na skutečně efektivní předpovědní metodu poukazují pouze kladné hodnoty všech čtyř ukazatelů R^2 .

Čtyři výše vyjmenované referenční předpovědní metody nebyly voleny náhodně. První resp. druhá metoda je mezním případem jednoduchého exponenciálního vyrovnávání pro $\alpha \approx 0$ resp. $\alpha \approx 1$, viz. odstavec 2.3. Třetí a čtvrtá metoda pak ve stejném smyslu představují extrémní polohy metod pro lokálně lineární trend.

Adekvátnost použití dané předpovědní metody na konkrétní časovou řadu budeme posuzovat skrze vlastnosti získaných předpovědních chyb. Přesněji budeme pracovat s normalizovanými předpovědními chybami \tilde{e}_{t_j} (viz. odstavec 6.1), které by v případě adekvátnosti dané předpovědní metody měly tvořit bílý šum. Statistických testů toho, že daná číselná posloupnost je realizací bílého šumu, existuje velké množství, viz. Brockwell a Davis (2002), kap. 1.6. Zde se omezíme na test nulovosti střední hodnoty a test nulovosti první autokorelace posloupnosti $\{\tilde{e}_{t_j}, j = 1, 2, \dots, n\}$.

Testem nulovosti střední hodnoty \tilde{e}_{t_j} zjišťujeme, jestli nejsou naše předpovědi systematicky vychýlené jedním směrem. Za předpokladu, že posloupnost $\{\tilde{e}_{t_j}, j = 1, 2, \dots, n\}$ je nezávislý bílý šum, má statistika

$$T_{bias} = \frac{\sum_{j=1}^n \tilde{e}_{t_j}}{\sqrt{\sum_{j=1}^n \tilde{e}_{t_j}^2}} \quad (6.2.8)$$

podle centrální limitní věty pro $n \rightarrow \infty$ asymptoticky rozdělení $N(0, 1)$, viz. Anděl (2002), věta B.17, str. 340. Příslušná asymptotická p -hodnota našeho testu je tedy

$$p_{bias} = 2 - 2 \cdot \Phi^{-1}(|T_{bias}|), \quad (6.2.9)$$

kde Φ je distribuční funkce rozdělení $N(0, 1)$. Hodnota p_{bias} blízká nule ukazuje, že v našich předpovědích je statisticky významné vychýlení. To nastane např. při použití jednoduchého exponenciálního vyrovnávání na časovou řadu vykazující převažující rostoucí či klesající trend nebo při použití některé metody pro lokálně lineární trend na časovou řadu vykazující převažující konvexní nebo konkávní trend. Řešením je většinou použít metodu pracující s polynomickým trendem vyššího stupně, případně v použití vhodné linearizující transformace na danou časovou řadu, viz. odstavec 6.3.

Za předpokladu, že posloupnost $\{\tilde{e}_{t_j}, j = 1, 2, \dots, n\}$ tvoří nezávislý bílý šum, má první výběrový autokorelační koeficient r_1 této posloupnosti pro velká n přibližně rozdělení $N(0, 1/n)$, viz. Brockwell a Davis (2002), příklad 2.4.2, str. 61. Příslušná asymptotická p -hodnota testu nulovosti první autokorelace posloupnosti normalizovaných předpovědních chyb je tedy

$$p_{corr} = 2 - 2 \cdot \Phi^{-1}(|r_1| \sqrt{n}). \quad (6.2.10)$$

Hodnota p_{corr} blízká nule ukazuje, že předpovědní chyby generované použitou metodou jsou statisticky významně korelované. To může mít několik různých příčin. Pokud např. zvolíme příliš malou resp. velkou hodnotu vyrovnávací konstanty α u jednoduchého exponenciálního vyrovnávání, výsledkem budou významně pozitivně resp. negativně korelované předpovědní chyby. Významná pozitivní korelace předpovědních chyb se objeví také tehdy, pokud na časovou řadu se střídavě konvexním a konkávním trendem použijeme metodu s lokálně lineárním trendem. Jiným zdrojem korelace předpovědních chyb (pozitivní nebo negativní) může být sezónnost časové řady, použijeme-li na ní některou nesezónní metodu.

Pokud konstruujeme předpovědní intervaly na základě předpokladu normality předpovědních chyb, měli bychom též otestovat tento předpoklad. V případě, že se tvar skutečného rozdělení chyb výrazně liší od normálního, jsou proklamované spolehlivosti předpovědních intervalů nepřesné. K testování normality normalizovaných předpovědních chyb \tilde{e}_{t_j} můžeme použít např. d'Agostinovy testy založené na výběrové šikmosti a špičatosti, viz. Anděl (2002), kap. 12.4.

6.3 Transformace časových řad

Použití transformace časové řady v procesu předpovídání jejích hodnot vypadá následovně. Dejme tomu, že pracujeme s časovou řadou $\{y_{t_j}\}$, jejíž hodnoty bychom

rádi předpovídali. Nepoužijeme však některou z předpovědních metod přímo na řadu y , ale na transformovanou řadu $\{z_{t_j}\}$ s hodnotami

$$z_{t_j} = f(y_{t_j}), \quad (6.3.1)$$

kde f je zvolená transformační funkce (nezabýváme se zde transformacemi jako je diferencování apod.). Získáme tak vyrovnané hodnoty \hat{z}_{t_j} a předpovědi $\hat{z}_{t_j+\tau}(t_j)$ v transformované řadě z . Vyrovnané hodnoty a předpovědi v řadě y pak zpětně získáme aplikací inverze funkce f :

$$\hat{y}_{t_j} = f^{-1}(\hat{z}_{t_j}) \quad \text{a} \quad \hat{y}_{t_j+\tau}(t_j) = f^{-1}[\hat{z}_{t_j+\tau}(t_j)]. \quad (6.3.2)$$

Transformaci můžeme použít v případě, že zkoumaná časová řada vykazuje jisté vlastnosti, které brání v úspěšném použití předpovědních metod, jež máme k dispozici. Jestli a jakou transformační funkci f v konkrétním případě použijeme, se tedy vždy rozhodujeme podle vlastností zkoumané časové řady a podle arsenálu předpovědních metod, jež máme k dispozici.

Ony nežádoucí vlastnosti časové řady y mohou být různé. Například exponenciální trend, máme-li k dispozici jen metody pracující s lineárním trendem. Podobně multiplikativní sezónnost, máme-li k dispozici jen metody pracující s aditivní sezónností. Obecně může řada vykazovat nějaký typ trendu, pro nějž nebude adekvátní žádná z uvažovaných předpovědních metod.

Často je amplituda náhodných fluktuací úměrná úrovni časové řady, či jsou tyto fluktuace výrazně asymetrické. Většina předpovědních metod ovšem počítá se symetrickým (nebo přímo normálním) rozdělením předpovědních chyb či jinak definovaných reziduí. Speciálně předpovědní intervaly konstruované na základě předpokladu normality jsou symetrické, což potom neodpovídá realitě.

Zvláštním důvodem pro použití transformace může být fakt, že pozorování řady y mají z podstaty věci omezený obor hodnot. Tak třeba velká většina ekonomických i jiných časových řad má z podstaty věci jen nezáporná pozorování: počet českých individuálních turistů trávící daný měsíc dovolenou ve Španělsku nebo denní uzavírací cena některé akcie na burze. V praxi můžeme většinou uvažovat dokonce jen kladné hodnoty takových časových řad.

Jiné časové řady mohou mít z podstaty věci všechna svá pozorování uvnitř intervalu $[0, 1]$: podíl domácností na území ČR vlastnící barevný televizor nebo procentuální míra nezaměstnanosti na stejném území. Opět můžeme v praxi předpokládat, že hodnoty takových řad budou dokonce uvnitř intervalu $(0, 1)$.

Naproti tomu většina předpovědních metod implicitně připouští hodnoty časové řady z celé reálné přímky a ze stejného oboru potenciálně generuje své předpovědi. Tak např. použitím Holtovy metody bychom snadno mohli získat předpověď, že v roce 2015 bude vlastnit barevný televizor 105 % českých domácností, či že nezaměstnanost na Slovensku bude v roce 2020 na úrovni -1 %. Podobně nesmyslně budou vypadat meze předpovědních intervalů, které budou mimo obor hodnot zkoumané řady.

Použitím vhodné transformace na časovou řadu můžeme výše popsaný problém odstranit, tj. zaručit, že výsledné předpovědi (6.3.2) včetně mezí předpovědních intervalů budou vždy ležet uvnitř oboru hodnot zkoumané časové řady.

Než přistoupíme k popisu konkrétních transformací užitečných v konkrétních situacích, zabývejme se ještě detailněji obecným schématem použití jakékoli transformace. Nejprve odpovíme na otázku, jaké vlastnosti by měla mít transformační funkce f ze vztahu (6.3.1). Jistě by mělo jít o reálnou funkci reálné proměnné, definovanou na oboru hodnot časové řady y . Dále by mělo jít o funkci prostou, aby existovala její inverze, viz. vzorce (6.3.2). Jistě by mělo jít též o funkci spojitou a monotónní, bez újmy na obecnosti předpokládejme že rostoucí. Inverze f^{-1} transformační funkce f by měla být definovaná na celé množině reálných čísel, jelikož předpovědi v transformované řadě z nejsou nijak omezené. Naopak oborem hodnot funkce f musí tedy být množina všech reálných čísel a oborem hodnot funkce f^{-1} obor hodnot původní časové řady y . To řeší výše uváděné problémy s nesmyslnými předpověďmi. Poznamenejme, že pokud bude transformovaná řada z vykazovat lineární trend, pak původní řada y vykazuje trend odpovídající grafu funkce f^{-1} .

Při použití dané předpovědní metody na transformovanou řadu z budeme veškeré předpoklady, které tato metoda vyžaduje, činit o této řadě z . To se týká především příslušných vzorců pro rozptyly předpovědních chyb a předpokladu normality těchto chyb. Také případná volba parametrů metodou maximální věrohodnosti z odstavce 6.1 bude probíhat v rámci transformované řady z , stejně jako testy adekvátnosti popsané v odstavci 6.2. Pokud jde o různé míry nepřesnosti předpovědi jako je MSE apod., ty je lépe vztahovat k původní řadě y , ať už jsou použity jako kritérium při optimální volbě parametrů metody či pouze jako informace o dosažené přesnosti předpovědi. Stejně tak ukazatele R^2 z odstavce 6.2 se budou vztahovat k původní řadě y .

Podobně jako ve vzorcích (6.3.2) budou i meze předpovědních intervalů pro pozorování řady y získány aplikací funkce f^{-1} na meze odpovídajících předpovědních intervalů pro řadu transformovanou řadu z . Nechť $\hat{z}_{t_j+\tau}(t_j)$, $z_{t_j+\tau}^U(t_j)$ a $z_{t_j+\tau}^L(t_j)$ jsou postupně bodová předpověď, horní a dolní mez předpovědního intervalu pro pozorování $z_{t_j+\tau}$ z času t_j . Pak budeme brát

$$\hat{y}_{t_j+\tau}(t_j) = f^{-1}[\hat{z}_{t_j+\tau}(t_j)], \quad (6.3.3)$$

$$y_{t_j+\tau}^U(t_j) = f^{-1}[z_{t_j+\tau}^U(t_j)] \quad \text{a} \quad y_{t_j+\tau}^L(t_j) = f^{-1}[z_{t_j+\tau}^L(t_j)]. \quad (6.3.4)$$

Díky vlastnostem funkce f vyplývají z nerovností

$$z_{t_j+\tau}^L(t_j) \leq \hat{z}_{t_j+\tau}(t_j) \leq z_{t_j+\tau}^U(t_j) \quad (6.3.5)$$

obdobné nerovnosti

$$y_{t_j+\tau}^L(t_j) \leq \hat{y}_{t_j+\tau}(t_j) \leq y_{t_j+\tau}^U(t_j) \quad (6.3.6)$$

pro předpovědi v řadě y . Také platí

$$y_{t_j+\tau}^L \leq y_{t_j+\tau} \leq y_{t_j+\tau}^U \Leftrightarrow z_{t_j+\tau}^L \leq z_{t_j+\tau} \leq z_{t_j+\tau}^U \quad (6.3.7)$$

a tedy předpovědní interval $[y_{t_j+\tau}^L, y_{t_j+\tau}^U]$ má stejnou spolehlivost jako odpovídající předpovědní interval $[z_{t_j+\tau}^L, z_{t_j+\tau}^U]$. Je-li funkce f konvexní resp. konkávní, je bodová předpověď $\hat{y}_{t_j+\tau}$ blíže horní resp. dolní mezi předpovědního intervalu $[y_{t_j+\tau}^L, y_{t_j+\tau}^U]$.

Na závěr tohoto odstavce uveďme čtyři konkrétní transformace. Vždy budou uvedeny předpisy funkcí f a f^{-1} , definiční obor funkce f a stručný komentář k použití dané transformace v praxi.

Logaritmická transformace

$$f(y) = \ln y, \quad f^{-1}(z) = \exp z, \quad D(f) = (0, \infty). \quad (6.3.8)$$

Vhodná je pro převod exponenciálního trendu na lineární, pro převod multiplikační sezónnosti na aditivní, pro odstranění přímé úměry mezi amplitudou náhodných fluktuací a úrovní řady a pro odstranění asymetrie v rozdělení těchto fluktuacích, pokud spočívá v těžším pravém chvostu. Zajišťuje, že hodnoty předpovědi budou kladné.

Odmocninová transformace

$$f(y) = \sqrt{y} - \frac{1}{\sqrt{y}}, \quad f^{-1}(z) = \frac{1}{4} \left(z + \sqrt{z^2 + 4} \right)^2, \quad D(f) = (0, \infty). \quad (6.3.9)$$

Představuje kompromis mezi žádnou transformací a logaritmickou transformací.

Inverzní logistická transformace

$$f(y) = -\ln\left(\frac{1}{y} - 1\right), \quad f^{-1}(z) = \frac{1}{1 + \exp(-z)}, \quad D(f) = (0, 1). \quad (6.3.10)$$

Vhodná je pro použití na časové řady s oborem hodnot $(0, 1)$. Zajišťuje, že i všechny předpovědi budou taktéž náležet do tohoto intervalu. Grafem funkce $f^{-1}(z)$ je tzv. *logistická křivka* symetrická podle svého inflexního bodu $[0, 1/2]$.

Inverzní Gompertzova transformace

$$f(y) = -\ln(-\ln y), \quad f^{-1}(z) = \exp[-\exp(-z)], \quad D(f) = (0, 1). \quad (6.3.11)$$

Má podobné vlastnosti jako předchozí transformace. Grafem funkce $f^{-1}(z)$ je tzv. *Gompertzova křivka* s inflexním bodem $[0, 1/e]$.

6.4 Praktické problémy a zkušenosti

Jedním praktickým problémem při práci s nepravidelnými časovými řadami je volba časového měřítka. Zatímco u pravidelných časových řad je časovým měřítkem automaticky časová vzdálenost mezi sousedními pozorováními, tak v případě nepravidelné časové řady není nikde dáno, mají-li být časy t_j měřeny např. v letech, měsících či týdnech. Jde-li o časovou řadu s chybějícími pozorováními, je přirozené zvolit jako časové měřítko rozestup pomyslné pravidelné řady, z níž naše řada vznikla vypuštěním některých hodnot. Pracujeme-li ovšem s obecně nepravidelnou časovou řadou, může být volba časového měřítka libovolná.

Lze samozřejmě používat metody, které jsou invariantní vůči změně časového měřítka, tedy vůči lineární transformaci vektoru okamžiků pozorování (t_1, t_2, \dots, t_n) . V tom případě nehraje volba časového měřítka roli. Většina metod však tuto vlastnost invariance nemá. Některé jsou sice invariantní, pokud jde o bodové předpovědi, ale již nejsou invariantní pokud jde třeba o intervalové předpovědi. Bylo by možné vždy standardizovat vektor (t_1, t_2, \dots, t_n) např. tak, aby platilo $q = 1$. Tato standardizace by mohla být i součástí implementace vyrovnávací a předpovědní metody a činila by jí tak vlastně invariantní vůči změně časového měřítka.

Význam standardizace časového měřítka pokud jde o metody typu exponenciálního vyrovnávání spočívá také v porovnatelnosti hodnot vyrovnávacích konstant mezi různými časovými řadami, viz. Wright (1986). Tyto metody sice jsou při tvorbě bodových předpovědí invariantní vůči změně časového měřítka, ale pouze za předpokladu adaptace hodnot vyrovnávacích konstant na změnu měřítka. Pokud např. při časové jednotce 1 den používáme $\alpha = 0,1$, pak při časové jednotce 1 týden musíme k získání totožných bodových předpovědí použít hodnotu $\alpha = 1 - (1 - 0,1)^7 \doteq 0,5217$.

Velice důležitá při používání metod typu exponenciálního vyrovnávání je volba těch „správných“ hodnot parametrů, většinou vyrovnávacích konstant. Přestože v dnešní době není po výpočetní stránce problém zvolit tyto hodnoty optimálně vzhledem k jistému kvantifikovanému kritériu (MSE, MAE atd.), není tento způsob výběru parametrů bez rizik.

Pokud máme k dispozici jen málo pozorování zkoumané časové řady, případně není její dosavadní průběh dostatečně reprezentativní vzhledem k možnému budoucímu průběhu, mohou mít optimální hodnoty parametrů nepříjemné vlastnosti. Dost často např. vycházejí hodnoty velice blízké 0 nebo 1.

Průběh optimalizovaného kritéria bývá ovšem v okolí optimální hodnoty velice plochý, takže i znatelná expertní korekce těchto hodnot většinou zhorší hodnotu optimalizovaného kritéria jen nepatrně. Větší dopad mívá změna parametrů metody

na výsledné předpovědní intervaly. Pokud tedy uživatele zajímají, měl by se při výběru parametrů řídit i jejich tvarem.

Optimalizace některého kritéria by tedy měla být jen technickým nástrojem, nikoli nutně konečným arbitrem ohledně volby hodnot parametrů, tím by měl být vždy uživatel sám. Užitečnou pomůckou při této volbě je i graf pozorování zkoumané časové řady spolu s příslušnými předpověďmi.

Asi nezákladnější je volba parametrů u Holtovy metody. V tomto případě je zvláště důležité provádět subjektivní grafickou analýzu získaných předpovědí a především vyzkoušet prakticky více různých kombinací hodnot parametrů. Překvapivě značný vliv na optimální hodnoty parametrů mají u Holtovy metody pozorování ze začátku časové řady, z nichž počítáme počáteční hodnoty statistik S a T . Pokud se hodnota T_0 „strefí“ do převládajícího trendu v časové řadě, často vyjde jakožto optimální $\gamma \approx 0$. Stačí však třeba jen změnit počet pozorování, z nichž jsou tyto počáteční hodnoty počítány, T_0 se již „nestrefí“ do převládajícího trendu a jako optimální hodnota pak vyjde $\gamma \gg 0$.

Problémy s výběrem hodnot parametrů zvolené předpovědní metody nejsou pochopitelně specifickým nepravidelných časových řad. Totéž platí o výběru metody samotné. Zde je však situace přeci jen poněkud rozmanitější než u pravidelných časových řad. Spousta metod, které v případě pravidelné časové řady splývají, dávají totiž pro nepravidelnou časovou řadu více či méně různé výsledky. Paradoxně tedy pro nepravidelné časové řady máme více různých metod, mezi kterými si musíme vybrat. Příkladem buď Wrightova modifikace Holtovy metody, dvojité exponenciální vyrovnávání a metoda založená na DLS odhadu lineárního trendu. Naproti tomu v případě pravidelné časové řady je k dispozici jen Holtova metoda (zbylé dvě splývají a jsou jejím speciálním případem).

7 Softwarová realizace

Součástí diplomové práce je program DMITS, v němž jsou dostupné všechny metody v této práci uvedené či odvozené, s výjimkou Holt-Wintersovy metody. Implementace metod zahrnuje výpočet počátečních hodnot, optimální volbu parametrů, výpočet bodových a intervalových předpovědí a vyhodnocení přesnosti a adekvátnosti použití dané předpovědní metody. Podrobnosti o možnostech programu DMITS přináší odstavec 7.1.

V odstavci 7.2 jsou uvedeny dva konkrétní numerické příklady použití prezentovaných metod. Dále je zde pomocí většího množství simulovaných časových řad porovnávána přesnost Wrightovy modifikace jednoduchého exponenciálního vyrovnávání a alternativní metody navržené v této práci.

7.1 Program DMITS

Součástí diplomové práce je autorem vytvořený software pro předpovídání a vyrovnávání v nepravidelných časových řadách - program DMITS (zkratka pro *Decomposition Methods for Irregular Time Series*). V tomto odstavci budou podrobně popsány výpočetní možnosti tohoto programu. V žádném případě nepůjde o uživatelský manuál k tomuto programu, ten je k dispozici spolu s programem samotným na přiloženém CD.

Jedinou metodou, která se vyskytuje v textu této práce a přitom není naprogramována v programu DMITS, je Holt-Wintersova metoda. Metody pro sezónní nepravidelné časové řady nebyly z důvodů omezeného rozsahu práce ani v centru jejího zájmu. Jde totiž o poněkud složitější problém, jak po stránce teoretické, tak i implementační, a bude předmětem dalšího výzkumu. Program DMITS, stejně jako text práce samotné, naopak poměrně detailně pokrývá metody typu exponenciálního vyrovnávání pro jednorozměrné nesezónní nepravidelné časové řady. Výčet metod je následující:

- Wrightova modifikace jednoduchého exponenciálního vyrovnávání
- Metoda založená na nepravidelně pozorovaném ARIMA(0, 1, 1) procesu
- Wrightova modifikace Holtovy metody, včetně tlumeného trendu
- Dvojitě exponenciální vyrovnávání
- Trojitě exponenciální vyrovnávání
- DLS odhad lineárního trendu
- DLS odhad kvadratického trendu

Vždy se uvažuje fiktivní časový okamžik $t_0 = t_1 - q$ a určují se počáteční hodnoty statistik pro tento čas. Předpovědní intervaly jsou založeny na předpokladu normality předpovědních chyb. Parametr σ^2 je odhadován jako střední čtvercová normalizovaná odchylka. Jsou používány *ad hoc* modifikace předpovědních intervalů k zohlednění časové struktury řady do okamžiku konstrukce předpovědi, které jsou uváděny v textu práce. Počáteční hodnoty jsou vždy získány pomocí DLS odhadu příslušného trendu v počátečním úseku řady, váhy exponenciálně klesají směrem do budoucnosti s uvedeným diskontním faktorem.

Nyní popíšeme obecné možnosti programu DMITS, nezávislé na konkrétní metodě. Nejprve se budeme zabývat tím, co uživatel musí či může programu zadat. Poté plynule přejdeme na popis výstupu programu.

Zadání časové řady

Programu je zadána posloupnost časových okamžiků t_1, t_2, \dots, t_n a posloupnost příslušných pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$. Musí přitom vždy platit $n \geq 2$, metody s kvadratickým trendem budou kvůli inicializaci vyžadovat $n \geq 3$. Dále nutně $t_1 < t_2 < \dots < t_n$. Rozdíly $t_{j+1} - t_j$ nemusí být celočíselné ani nemusí jít o násobky daného pevného čísla. Není požadováno $t_1 \geq 0$. Posloupnost časových okamžiků t_1, t_2, \dots, t_n není programem nijak upravována, případná změna časového měřítka je věcí uživatele a musí být provedena před zadáním řady programu. Hodnoty y_{t_j} mohou být obecně libovolná reálná čísla, pokud není zvolena transformace s omezeným definičním oborem (viz. níže).

Transformace

Program nabízí před samotnou předpovědní metodou použití čtyř různých transformací na zadanou časovou řadu: logaritmickou, odmocninovou, inverzní logistickou a inverzní Gompertzovu. Pochopitelně je také možné nevyužít žádnou z transformací. V případě prvních dvou typů transformací musí být hodnoty y_{t_j} zadané časové řady vesměs kladné, v případě zbylých dvou pak musí být $y_{t_j} \in (0, 1)$.

Výběr metody

Výběr některé ze seznamu sedmi dostupných předpovědních a vyrovnávacích metod je plně věcí uživatele. Volba konkrétní metody nemá žádný vliv na ostatní volby jako je třeba použití transformace apod. Pochopitelně se metody liší v počtu svých parametrů, které je třeba zvolit, a také minimálním počtem pozorování použitým k výpočtu počátečních hodnot (viz. níže).

Volba hodnot parametrů

Zatímco Holtova metoda má tři parametry (vyrovnávací konstanty α a γ a tlumící konstantu φ), zbylých šest metod má pouze jediný parametr (vyrovnávací konstantu α). U každého parametru je možné buď zadat konkrétní pevnou hodnotu,

příčemž musí platit $\alpha \in (0, 1)$, $\gamma \in (0, 1)$ a $\varphi \in (0, 1]$, nebo nechat program zvolit optimální hodnotu parametru vzhledem k danému kritériu. Optimalizovány tak mohou být jeden, dva nebo i tři parametry současně.

Program nabízí čtyři možná optimalizační kritéria: minimální MSE, MAE či MAPE a maximální věrohodnost. Hodnoty kritérií jsou vždy počítány na základě všech n pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$ a k nim příslušných bodových předpovědí. Pro použití kritéria MAPE musí být hodnoty y_{t_j} zadané časové řady vesměs kladné.

Hledání optimální kombinace parametrů z hlediska zvoleného kritéria je prováděno iteračním numerickým algoritmem. Vždy se všechny parametry až na jeden zafixují a kritérium je optimalizováno přes zbývající parametr. Takto se parametry pravidelně střídají. Optimalizace počítá s konvexní minimalizovanou funkcí, při porušení tohoto předpokladu je možné, že posloupnost přiblížení se bude konvergovat jen k lokálnímu minimu, které bude různé od hledaného globálního minima. Vzhledem k tomu, že se při optimalizaci omezujeme na obor $\alpha, \beta, \gamma \in (0, 1)$, nemusí ani takové globální minimum existovat.

Volba počtu pozorování pro inicializaci

Je třeba zadat počet pozorování ze začátku časové řady, ze kterých budou počítány počáteční hodnoty. U metod s konstantním, lineárním, resp. kvadratickým trendem to musí být alespoň 1, 2, resp. 3 pozorování. Nejvýše je možné inicializaci provést ze všech n pozorování. Počáteční hodnoty jsou také funkcí použitých hodnot parametrů metody. Průměrná časová vzdálenost q mezi dvěma sousedními pozorováními řady je brána jako $q = (t_n - t_1)/(n - 1)$.

Volba parametrů předpovědí

Programu se dále musí zadat parametry budoucích předpovědí: délka předpovědního horizontu, krok předpovědí a spolehlivost předpovědních intervalů. Předpovědní horizont určuje, jak daleko do budoucnosti od času t_n posledního pozorování budou počítány předpovědi. Ty se vykreslí do grafu jako spojitá křivka, jejich hodnoty budou navíc vypsány se zmíněným pevným krokem. Spolehlivost předpovědních intervalů je volena jako 50, 75, 90, 95 nebo 99 %.

Specifikace metody

První sekcí textového výstupu programu jsou informace o zvolené předpovědní metodě. Jde tedy z větší části o přepis toho, jakou transformaci, metodu, způsob volby parametrů apod. uživatel zvolil. Jedinou novou informací jsou použité hodnoty parametrů, pokud byly voleny optimálně.

Vlastnosti časové řady

Jsou vypsány základní informace o zadané časové řadě: počet pozorování n , jejich průměrný časový rozestup q , jejich průměr, rozptyl a směrodatná odchylka.

Přesnost předpovědí

Jsou vypsány dosažené hodnoty měr přesnosti předpovědí: MSE, RMSE, MAE a MAPE. Jsou počítány na základě všech n pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$ a k nim příslušných bodových předpovědí. MAPE není počítáno, pokud nejsou hodnoty y_{t_j} vesměs kladné.

Efektivita předpovědní metody

V další sekci jsou prezentovány dosažené hodnoty čtyř ukazatelů R^2 , popsanych v odstavci 6.2. Jako míra nepřesnosti předpovědí je při jejich výpočtu použita MSE.

Adekvátnost předpovědní metody

Jsou uvedeny výsledky testů předpokladu, že normalizované předpovědní chyby \tilde{e}_{t_j} tvoří bílý šum, viz. odstavce 6.1 a 6.2. Jsou vypsány průměr a průměr čtverců těchto chyb a p -hodnota testu nulovosti jejich střední hodnoty. Dále je vypsána jejich výběrová první autokorelace a p -hodnota testu nekorelovanosti těchto chyb.

Testy normality předpovědních chyb

Pokud $n \geq 9$, jsou vypsány výběrová šikmost a špičatost normalizovaných předpovědních chyb \tilde{e}_{t_j} a p -hodnoty testů normality založené na těchto ukazatelích, viz. odstavec 6.2.

Výkon předpovědních intervalů

Pro všech n pozorování y_{t_j} je určeno, zda-li padla nad, pod, či do vnitřku příslušného předpovědního intervalu konstruovaného z předchozího časového okamžiku t_{j-1} . Tato zjištění jsou pak ve formě tří procentuálních hodnot porovnána s jejich teoretickými protějšky. Poznamenejme, že vzhledem ke způsobu konstrukce předpovědních intervalů, by v průměru mělo vždy stejně pozorování padnout nad jako pod tyto předpovědní intervaly.

Výpis historických předpovědí a vyrovnaných hodnot

Pro $j = 1, 2, \dots, n$ jsou po řádcích vypsány hodnoty $j, t_j, y_{t_j}, \hat{y}_{t_j}$ a $\hat{y}_{t_j}(t_{j-1})$.

Výpis budoucích předpovědí

Dále jsou vypsány bodové předpovědi z času t_n a meze příslušných předpovědních intervalů, vše pro časy $t_n + s, t_n + 2s, t_n + 3s, \dots$, kde $s > 0$ je zadaný krok. Limitem je zadaná délka předpovědního horizontu h .

Graf

Do jednoho grafu jsou vykresleny značky pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$, souvislá křivka bodových předpovědí pro časové rozmezí t_1 až $t_n + h$ (h je délka předpovědního

horizontu) a souvislé křivky mezi předpovědních intervalů pro časové rozmezí t_n až $t_n + h$. Křivka bodových předpovědí je definována hodnotou $\hat{y}_{t_j+\tau}(t_j)$ v bodě $t_j + \tau$, kde $\tau \in [0, t_{j+1} - t_j]$. Dvě různé hodnoty $y_{t_j}(t_{j-1})$ a \hat{y}_{t_j} pro bod t_j , $j = 2, 3, \dots, n$, jsou v grafu spojeny vertikální úsečkou. Takovýto způsob vyobrazení historických předpovědí považuje autor této diplomové práce za správnější a užitečnější než obvyklou lineární interpolaci jednotlivých bodových předpovědí.

7.2 Numerické příklady

Prezentovány budou dva příklady použití předpovědních metod pro nepravidelné časové řady na konkrétních reálných datech. Toto však nebude myšleno tak, že by použití dotyčných metod na tyto časové řady bylo nejlepším možným způsobem konstrukce jejich předpovědí, půjde spíše o pouhou ilustraci. Dále bude použito větší množství simulovaných časových řad k porovnání Wrightovy modifikace jednoduchého exponenciálního vyrovnávání s alternativní metodou navrženou v této práci. Zmíněné příklady byly zpracovány v programu DMITS.

Rekord v běhu na jednu míli

Tento příklad byl inspirován obdobným příkladem z článku Wright (1986). Uvažujme časovou řadu mužských světových rekordů v běhu na jednu míli od roku 1875 do roku 1999. Časovým okamžikem pozorování je rok ustavení nového rekordu, jeho hodnotou dotyčný rekordní čas měřený v sekundách (s přesností na desetiny, od roku 1981 na setiny). Pokud byl rekord překonán vícekrát v jeden kalendářní rok, je brán nejlepší čas z daného roku. Máme k dispozici 31 pozorování této řady s průměrným časovým rozestupem $q = 4.1\bar{3}$.

Řada vykazuje poměrně stálý klesající lineární trend, nabízí se tedy použít na ni některou z metod pro řady s lokálně lineárním trendem. K použití některé transformace v tomto případě nejsou dostatečně pádné důvody. Pro výpočet počátečních hodnot budeme používat vždy prvních 6 pozorování ($t_6 = 1895$). Parametry metod budou voleny optimálně vzhledem k jednomu ze čtyř kritérií. Tabulka 7.2.1 ukazuje výsledky pro různé metody a různá kritéria volby parametrů.

Z tabulky 7.2.1 vyplývá, že nejpřesnější metodou je zde Holtova metoda s tlumeným lineárním trendem, následuje DLS odhad lineárního trendu, dvojité exponenciální vyrovnávání a až na posledním místě se umístila klasická Holtova metoda. Pokud jde o jednotlivá optimalizační kritéria, nejví se MAE a zvláště pak MAPE jako příliš šťastná volba. Jen o málo nižší hodnota těchto ukazatelů je ve většině případů vykoupena znatelně vyšším MSE. Na obrázku 7.2.1 je graf časové řady a předpovědí pomocí Holtovy metody s tlumeným lineárním trendem.

Pochopitelně nelze činit žádné obecné závěry. Zvláště když zjistíme, že jak optimální hodnoty parametrů, tak dosažené míry nepřesnosti předpovědí jsou silně

ovlivněny především pozorováními ze začátku časové řady, potažmo počtem těchto pozorování, z nichž jsou počítány počáteční hodnoty.

Dvojitě exponenciální vyrovnávání

kritérium	α			MSE	MAE	MAPE
MSE	0,08697			2,39833	1,21492	0,4981 %
MAE	0,11327			2,47699	1,19173	0,4879 %
MAPE	0,11992			2,51967	1,19249	0,4879 %
ML	0,11262			2,47328	1,19216	0,4881 %

DLS odhad lineárního trendu

kritérium	α			MSE	MAE	MAPE
MSE	0,10315			2,20806	1,15466	0,4721 %
MAE	0,11056			2,21318	1,14743	0,4686 %
MAPE	0,19082			2,69059	1,15328	0,4690 %
ML	0,11831			2,22874	1,15123	0,4698 %

Holtova metoda s lineárním trendem

kritérium	α	γ		MSE	MAE	MAPE
MSE	0,10030	0,18924		2,62728	1,27793	0,5215 %
MAE	0,15753	0,21684		3,20661	1,27412	0,5205 %
MAPE	0,15854	0,21741		3,21886	1,27437	0,5206 %
ML	0,11353	0,16186		2,65389	1,30267	0,5323 %

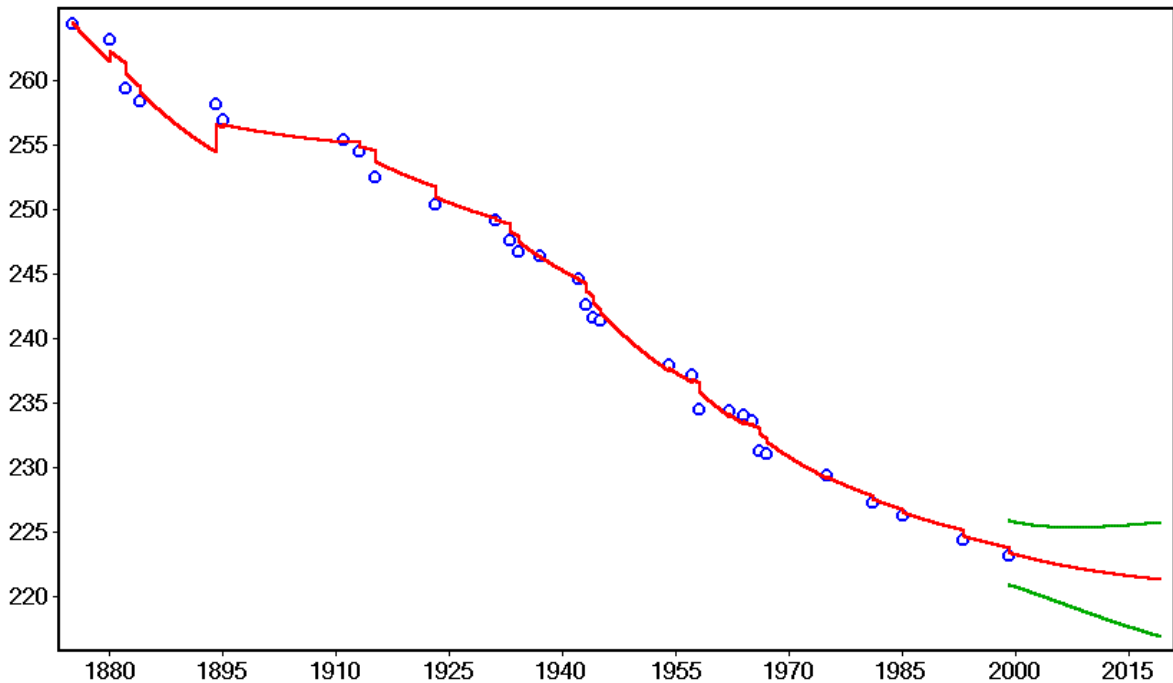
Holtova metoda s tlumeným lineárním trendem

kritérium	α	γ	φ	MSE	MAE	MAPE
MSE	0,13027	0,16599	0,94999	1,64536	0,99688	0,4068 %
MAE	0,16400	0,05751	0,97351	1,81138	0,98316	0,3981 %
MAPE	0,16400	0,05751	0,97351	1,81138	0,98316	0,3981 %
ML	0,11774	0,15534	0,95948	1,65419	1,00280	0,4081 %

Tabulka 7.2.1 Použité parametry a dosažené hodnoty MSE, MAE a MAPE pro čtyři metody v závislosti na optimalizovaném kritériu.

Pokud zapomeneme, že pracujeme s nepravidelnou časovou řadou a budeme jednotlivá pozorování považovat za rovnoměrně rozmístěná v čase, získáme podstatně přesnější předpovědi (půjde ovšem o předpovědi něčeho mírně odlišného). Fakt, že světový rekord není dlouho překonán, totiž ještě neznamená, že následné vylepšení času bude o to výraznější, jak by vyplývalo z předpokladu lineárního trendu v naší časové řadě. Tento příklad je tedy vyložene pouze ilustrativní.

Bylo by možné brát jako pozorování každé zlepšení rekordu a jako časové okamžiky brát přesné datum ustavení rekordu převedené na desetinné číslo vyjadřující rok. Při použití Holtovy metody by však nastal problém popsáný v odstavci 4.2, totiž citlivost metody na časově velmi blízká pozorování. Vzorovým příkladem je rekordní čas z 26. srpna 1981, který byl překonán hned po dvou dnech, a to o 1,07 s.



Obrázek 7.2.1 Světový rekord (s) v běhu na 1 míli v letech 1875 až 1999. Holtova metoda s tlumeným lineárním trendem, parametry voleny minimalizací MSE, předpovědní horizont 20 let, spolehlivost předpovědních intervalů 95 %.

Zastoupení obyvatel nad 60 let

Uvažujme časovou řadu zastoupení věkové kategorie 60+ v obyvatelstvu žijícím na území dnešní ČR. Máme k dispozici 12 pozorování za období 1869 až 1991 pocházejících ze sčítání lidu, které bylo prováděno v průměru zhruba jednou za 11 let (nejkratší interval je 9, nejdelší 20 let). Hodnoty naší časové řady ukazuje tabulka 7.2.2.

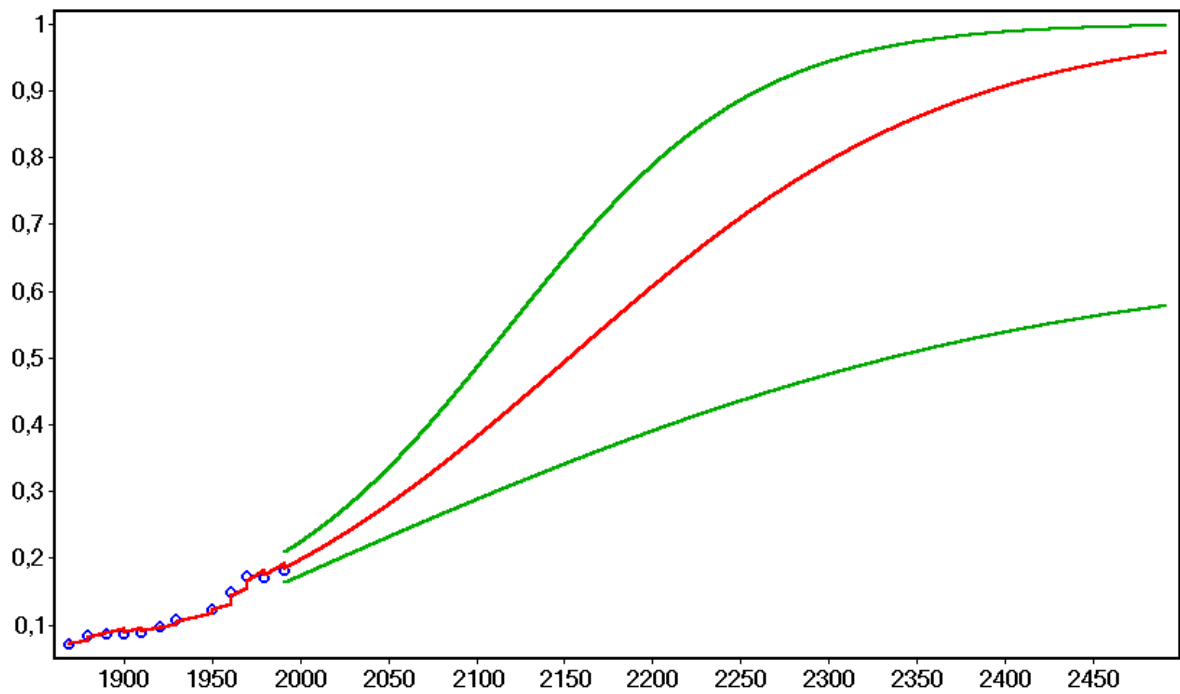
rok	1869	1880	1890	1900	1910	1921	1930	1950	1961	1970	1980	1991
60+	0,071	0,084	0,087	0,087	0,089	0,096	0,107	0,123	0,149	0,172	0,17	0,181

Tabulka 7.2.2 Zastoupení věkové kategorie 60+ v obyvatelstvu žijícím na území ČR podle výsledků sčítání lidu. Zdroj: Český statistický úřad, www.czso.cz.

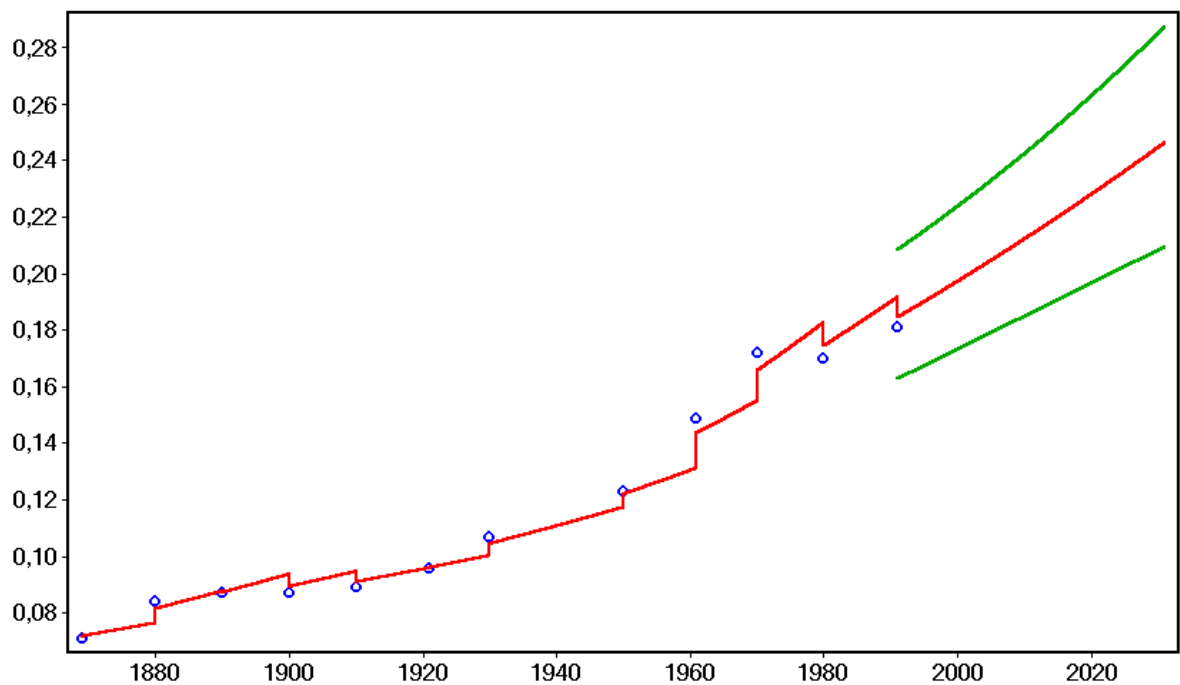
Časová řada vykazuje konvexní rostoucí trend. Můžeme na ni tedy použít metodu s lokálně lineárním trendem spolu s některou z transformací, která by pomohla zachytit konvexní průběh časové řady. Byla zvolena metoda založená na DLS odhadu lineárního trendu spolu s inverzní logistickou transformací. Ta též zaručí, že všechny bodové i intervalové předpovědi padnou do intervalu $(0, 1)$. Počet pozorování použitý pro inicializaci byl 6. Vyrovnávací konstanta α byla volena minimalizací MSE, optimální hodnota činila přibližně 0,05043.

Pro názornější ilustraci vlastností použité transformace byla použita délka předpovědního horizontu 500 let, viz. obrázek 7.2.2. Samozřejmě nejde o seriózní

pokus předpovídat věkové složení obyvatelstva v 25. století. Detail do oblasti pozorování zkoumané časové řady je na obrázku 7.2.3.



Obrázek 7.2.2 Zastoupení věkové kategorie 60+ v obyvatelstvu žijícím na území ČR podle výsledků sčítání lidu z let 1869 až 1991. Předpovědní horizont 500 let.



Obrázek 7.2.3 Zastoupení věkové kategorie 60+ v obyvatelstvu žijícím na území ČR podle výsledků sčítání lidu z let 1869 až 1991. Předpovědní horizont 40 let.

Jednoduché exponenciální vyrovnávání

Zabýval jsem se empirickým srovnáním přesnosti předpovědí získaných Wrightovou modifikací jednoduchého exponenciálního vyrovnávání a zde navrženou metodou založenou na nepravidelném ARIMA(0, 1, 1) procesu. K tomu účelu byly časové řady získávány jako nepravidelný výběr z realizací ARIMA(0, 1, 1) procesu generovaného normálním bílým šumem o rozptylu 5. Počet pozorování výsledné časové řady byl volen 5 000, časový krok mezi dvěma sousedními pozorováními byl volen náhodně z množiny $\{1, 2, \dots, T\}$. Hodnota T byla volena postupně 2, 5 a 10, hodnota vyrovnávací konstanty α použité při generování časových řad postupně 0,1, 0,2, 0,3, 0,4 a 0,5.

Na každou z 15 vzniklých časových řad byly použity obě výše zmíněné předpovědní metody. Počet pozorování pro výpočet počátečních hodnot byl 6, vyrovnávací konstanta α byla volena minimalizací MSE. Tabulka 7.2.3 ukazuje použité optimální hodnoty α a dosažené hodnoty MSE pro každou z 15 časových řad.

		Wrightovo jednoduché exponenciální vyrovnávání		Metoda založená na ARIMA(0, 1, 1) procesu	
skutečné α	N	optimální α	MSE	optimální α	MSE
0,1	2	0,07871	5,10208	0,09574	5,10111
0,1	5	0,05440	5,41406	0,09383	5,41286
0,1	10	0,04246	5,71475	0,09771	5,71173
0,2	2	0,17986	5,11015	0,21687	5,10400
0,2	5	0,11842	5,88946	0,19823	5,88663
0,2	10	0,09156	6,60094	0,20737	6,58273
0,3	2	0,24865	5,41704	0,29839	5,41412
0,3	5	0,18548	6,69379	0,30631	6,69177
0,3	10	0,12365	7,70023	0,27637	7,67747
0,4	2	0,35202	5,54940	0,41465	5,54063
0,4	5	0,25260	7,10261	0,41073	7,07964
0,4	10	0,17889	9,45058	0,38985	9,41525
0,5	2	0,42380	5,72639	0,49597	5,72045
0,5	5	0,31635	7,77051	0,50232	7,74136
0,5	10	0,25545	10,78900	0,51540	10,75790

Tabulka 7.2.3 Použité vyrovnávací konstanty α a dosažené hodnoty MSE pro obě testované metody a 15 různých použitých časových řad.

Ve všech 15 případech vykázala nižší MSE metoda založená na nepravidelném ARIMA(0, 1, 1) procesu. Největší rozdíl v MSE mezi testovanými metodami činil však pouhých 0,38 %, v průměru pak pouhých 0,17 %. Závěr tedy je, že obě metody jsou srovnatelně přesné. Ke stejnému závěru vedlo i použití obou metod na jiné časové řady. Přesto, že jsou procentuální rozdíly v MSE zanedbatelné, je patrná jejich rostoucí závislost na skutečné hodnotě α a na hodnotě N . Na nich pochopitelně stejným

způsobem závisí i dosažené hodnoty MSE v absolutním měřítku, a to plně v souladu se vzorcí pro rozptyl předpovědních chyb z odstavce 3.2.

U metody založené na ARIMA procesu se optimální hodnoty vyrovnávací konstanty α významně neliší od hodnoty použité ke generování řad (v 8 případech je skutečná hodnota větší než použitá, v 7 případech naopak). V absolutní hodnotě největší rozdíl činí 0,02363 v případě časové řady s $\alpha = 0,3$ a $N = 10$. Je-li $N \geq 1$, tak v případě Wrightova jednoduchého exponenciálního vyrovnávání není důvod, proč by zde použité optimální vyrovnávací konstanty měly být blízké těm použitým při generování ARIMA procesu.

8 Závěr

Zadáním této diplomové práce bylo podat přehled o existujících dekompozičních metodách pro vyrovnávání a předpovídání v časových řadách s nepravidelně pozorovanými hodnotami. Práce se měla zaměřit zvláště na metody rekurentního typu.

Vzniklá práce se zabývá výhradně metodami typu exponenciálního vyrovnávání pro jednorozměrné nepravidelné časové řady. Vyjma ilustrativního odstavce o Holt-Wintersově metodě jde o metody pro nesezónní časové řady. Kromě popisu existujících metod byly navrženy některé jejich drobné úpravy, zvláště pokud jde o výpočet počátečních hodnot a výpočet mezi předpovědních intervalů. Některé nové metody byly v práci odvozeny.

Zde navržená alternativa Wrightova jednoduchého exponenciálního vyrovnávání, založená na předpokladu nepravidelně pozorovaného $ARIMA(0, 1, 1)$ procesu, se ukázala být ve svých předpovědích srovnatelně přesná jako původní metoda. V práci jsou prezentovány hned tři metody pro nepravidelné časové řady s lokálně lineárním trendem. Dvojitě exponenciální vyrovnávání a metoda založená na DLS odhadu lineárního trendu již nejsou speciálním případem Holtovy metody, jak je tomu v kontextu pravidelných časových řad. V některých situacích poskytují první dvě metody dokonce lepší výsledky.

Praktické problémy při používání metod typu exponenciálního vyrovnávání pro nepravidelné časové řady jsou stejné jako u jejich klasických verzí. Klíčová je tedy především otázka volby hodnot vyrovnávacích konstant. V práci byla navržena jejich volba metodou maximální věrohodnosti, která u pravidelných časových řad splývá s minimalizací MSE. Užitečnou informaci o efektivitě předpovědní metody představují podle autora ukazatele obdobné koeficientu determinace v lineární regresi.

Jako součást diplomové práce byl vytvořen program DMITS, kde jsou naprogramovány všechny metody zde prezentované (s výjimkou Holt-Wintersovy). Tento program sloužil autorovi práce k testování nově navržených metod a tím k eliminaci případných chyb v jejich odvození. Dále poskytuje uživatelsky přátelský nástroj pro experimentování s jednotlivými metodami a umožňuje tak sbírat užitečné zkušenosti ohledně jejich praktické aplikace.

Autor diplomové práce cítí jistý dluh vůči metodám pro sezónní (nepravidelné) časové řady. Těm zde z důvodu omezeného rozsahu práce nebyl věnován takový prostor, jaký by odpovídal jejich praktické využitelnosti. Jde o obecně složitější problém (minimálně tehdy, pokud modelujeme sezónnost pomocí sezónních indexů) a bude předmětem dalšího výzkumu. Jako problém pocíťoval autor práce též nedostatek veřejně publikovaných nepravidelných časových řad, na nichž by mohla být užitečnost zde popisovaných a odvozovaných metod lépe prokázána.

Literatura

- Anděl, J. (2002): Základy matematické statistiky, MFF UK, Praha.
- Bartsch, H. J. (2000): Matematické vzorce (3. vyd.), Mladá fronta, Praha.
- Brockwell, P. J., Davis, R. A. (2002): Introduction to Time Series and Forecasting (2nd ed.), Springer-Verlag, New York.
- Cipra, T., Trujillo, J., Rubio, A. (1995): Holt-Winters method with missing observations. *Management Science* 41, 174-8.
- Cipra, T. (2006): Exponential smoothing for irregular data. *Applications of Mathematics* 51, 597-604.
- Cipra, T. (1989): Some problems of exponential smoothing. *Aplikace matematiky* 34, 161-9.
- Gardner, E. S. (1985): Exponential smoothing: the state of the art. *Journal of Forecasting* 4, 1-28.
- Gardner, E. S., McKenzie, E. (1985): Forecasting trends in time series. *Management Science* 31, 1237-46.
- Gardner, E. S., McKenzie, E. (1989): Seasonal exponential smoothing with damped trends. *Management Science* 35, 372-6.
- Chatfield, C. (2002): *Time-Series Forecasting*. Chapman & Hall/CRC.
- Prášková, Z. (2004): *Základy náhodných procesů II*, Karolinum, Praha (skripta).
- Taylor, J. W. (2003): Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society* 54, 799-805.
- Winters, P. R. (1960): Forecasting sales by exponentially weighted moving averages, *Management Science* 6, 324-342.
- Wright, D. J. (1986): Forecasting data published at irregular time intervals using extension of Holt's method. *Management Science* 32, 499-510.