

## 0 Poznámky k poznámkam

- Okrem úvodu vznikali poznámky v roku 2020, úvod bol doplnený v roku 2022 - je teda možné, že niektoré veci sa nebudú úplne zhodovať s tým, čo budem hovoriť na prednáškach (keďže niekedy sa rozhodnem klásť menší/väčší dôraz na iné veci, prípadne sa na niečo budem pozerat' iným pohľadom, prípadne niečo pridám alebo vynechám).
- Ak je niečo sivou farbou, berte to ako nejakú drobnú poznámku, prípadne snahu niečo viac rozviesť/dovysvetliť. V princípe sa tieto časti dajú preskakovať.
- Ak ste našli nejakú chybu (či už matematickú, gramatickú alebo štylistickú), prípadne máte pocit, že niečo nie je dobre/jasne vysvetlené, dajte mi, prosím, vedieť prostredníctvom formulára na stránke <https://forms.gle/nTox8o6J5FXj74bj9>. Budem veľmi vďačný.

## 1 Úvod

### Čo je to časový rad?

Voľnejšie môžeme povedať, že sú to chronologicky usporiadané dáta týkajúce sa jednej veci (jedného štatistického znaku). Presnejšie je to nejaká nekonečná postupnosť náhodných premenných

$$(\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots),$$

skrátene iba  $\{X_t\}_{t \in \mathbb{Z}}$ .

V praxi pracujeme vždy iba s tzv. konečným pozorovaním  $(x_1, x_2, \dots, x_n)$ , čo predstavuje realizáciu náhodného vektora  $(X_1, X_2, \dots, X_n)$ , kde  $n$  je nejaké prirodzené číslo. V zmysle voľnejšej definície budeme aj konečné pozorovanie  $(x_1, \dots, x_n)$  nazývať časový rad. Iba drobná poznámka k pojmu „realizácia“, ak by náhodou čitateľ nebol s týmto pojmom zžitý - je to nejaké konkrétne uskutočnenie javu, na ktorý sa pozeráme ako na náhodný. Napríklad ak náhodná premenná  $X$  predstavuje číslo na kocke, ktorú sa chystáme hodiť, tak jej realizácia  $x$  bude konkrétne číslo, ktoré nám padlo. Čiže rozdiel medzi veľkým a malým písmenom je, že veľké je náhodná premenná (nevieme konkrétnu hodnotu) a malé je už konkrétne číslo (v prípade kocky napríklad 4). V časových radoch väčšinou pracujeme s číselnými dátami (čiže realizáciou), ale pri niektorých teoretických úvahách je užitočné za tým vidieť aj tie náhodné premenné.

Ak nie je povedané inak (a na tomto predmete zrejme inak povedané nebude), časové úseky medzi susednými dátovými bodmi sú aspoň približne rovnaké. Najbežnejšie budeme pracovať s mesačnými, kvartálnymi, ročnými a niekedy dennými dátami.

### Kde sa s niečím takýmto stretnúť?

Napríklad v

- medicíne (tep/tlak pacienta, denné počty nakazených/úmrtí koronavírusom, ...),
- financiách, ekonomike (vývoj úrokových mier, menových kurzov, tržby v obchode, ...),
- meteorológii, hydrológii (hladina riek, teplota vzduchu, ...),
- demografii

a mnohých iných oblastiach.

## Načo je to dobré?

Dôvodov je mnoho, spomeňme tri hlavné.

1. Predikcia - jednou z najdôležitejších úloh pri analýze časových radov je jeho predikcia, t.j. na základe konečného pozorovania časového radu, ktorý máme k dispozícii nájsť rozumné odhady pre  $X_{n+1}, X_{n+2}, \dots$  (prípadne  $X_0, X_{-1}, \dots$ ). Ak sme schopní predikovať, čo sa bude diať v budúcnosti, tak na základe toho vieme rozumne prispôbiť svoje správanie v prítomnosti. Napríklad ak nám výjde, že cena bitcoinu bude v najbližších dňoch prudko klesať, tak ho dnes nenakúpime. Alebo ak vychádza predikcia, že počty nakazených budú prudko rásť, treba na to adekvátne zareagovať už teraz.
2. Deskripcia - niekedy nás nemusí zaujímať, čo sa bude diať v budúcnosti - chceme iba popísať dané dáta (napríklad ich nejako charakterizovať menším počtom čísel). Do kategórie deskripcia možno zaradiť napríklad tieto veci (na tomto kurze sa budeme venovať iba prvej z nich):
  - Dekompozícia časového radu - v princípe ide o rozklad tvaru

$$X_t = Tr_t + S_t + E_t, \quad (1.1)$$

kde  $Tr_t$  je trendová zložka (predstavuje akýsi dlhodobý priemer),  $S_t$  je sezónna zložka (opisuje periodické zmeny v časovom rade, ktoré sa opakujú každý väčší časový úsek - napríklad pri mesačných dátach môžeme často vidieť, že každý rok je v dátach podobný vzor) a  $E_t$  je reziduálna (alebo náhodná) zložka nemajúca deterministický charakter. Takýto prístup nám vie dať nejaký približný náhľad do dát (napríklad nám povedať, čo sa deje v jednotlivých „sezónach“) a môže slúžiť aj na porovnanie viacerých časových radoch - napríklad majme denné tržby dvoch obchodov - ak v jednom trendová zložka vychádza  $Tr_t^{(1)} = 500 + 5t$  a v druhom  $Tr_t^{(2)} = 400 + 7t$ , tak môžeme usudzovať, že v druhom obchode priemerné tržby rastú rýchlejšie.

- Spektrálna analýza - slúži hlavne pri spracovaní signálu na popísanie toho, ako veľmi sú jednotlivé frekvencie (resp. cykly) v signáli zastúpené. Signál možno chápať ako nejakú spojitú vlnu, avšak reálne (ak ju chceme mať uloženú v počítači), musíme ju nejako diskretizovať, čím dostaneme postupnosť bodov, čiže časový rad (náhodnosť vzniká kvôli šumu v signáli).
- Lineárna regresia - niekedy máme okrem časového radu (resp. jeho konečného pozorovania  $(x_1, \dots, x_n)$ ) aj nejaké potenciálne vysvetľujúce premenné  $z_{t,1}, z_{t,2}, \dots, z_{t,m}$  (pre  $t = 1, 2, \dots, n$ ). Štandardne v takýchto situáciách uvažujeme lineárny model

$$x_t = \beta_0 + \beta_1 z_{t,1} + \beta_2 z_{t,2} + \dots + \beta_m z_{t,m} + \varepsilon_t, \quad (1.2)$$

$t = 1, 2, \dots, n$  kde  $\varepsilon_t$  je nejaká náhodná chyba. Jednou z vecí, ktorá nás môže zaujímať je „vplyv“ jednotlivých vysvetľujúcich premenných na sledovaný časový rad, t.j. aké sú hodnoty jednotlivých biet. Slovo „vplyv“ je v úvodzovkách, nakoľko nemusí ísť o kauzalitu, môže to byť iba korelácia - presnejšie slovo by teda mohlo byť „súvislosť“. Problémom v časových radoch býva, že náhodné chyby  $\varepsilon_1, \dots, \varepsilon_n$  sú často korelované, čím je porušený základný predpoklad „klasickej“ regresie - treba preto použiť iné postupy využívajúce modely pre časové rady.

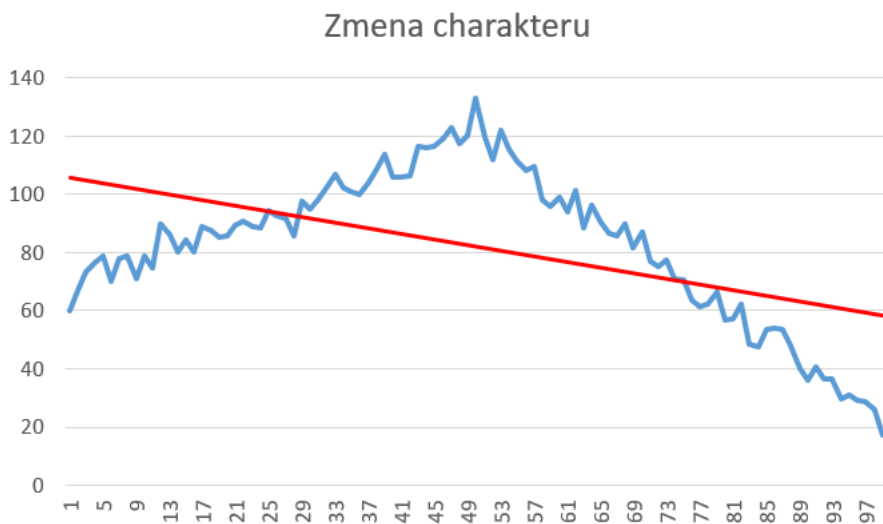
3. Riadenie - niekedy môžeme správanie časového radu ovplyvňovať (napríklad ak chceme nejako ovplyvniť tržby v obchode a máme na to potrebný kapitál, tak môžeme sa napríklad investovať do reklamy, alebo môžeme „investovať“ do zliav na tovar) - v takom prípade môžeme chcieť nájsť optimálnu stratégiu, ktorá nám bude maximalizovať nejakú účelovú funkciu (napríklad očakávaný zisk z tržieb v nasledujúcom mesiaci).

## Na čo si dávať pozor pri práci s časovými radmi?

- Na kalendár - niektoré sledované znaky môžu byť výrazne ovplyvnené kalendárom, napríklad počtom dní v mesiaci (celkové mesačné tržby obchodu budú zrejme vo februári nižšie), počet víkendov v mesiaci, sviatky a prázdniny (najmä pohyblivé). Niekedy je teda vhodné dáta nejakým spôsobom znormovať (napríklad zobrať mesačné tržby predelené počtom pracovných dní, prípadne počtom dní, kedy bol obchod v danom mesiaci otvorený), prípadne zobrať nejaký „vyrovnanejší“ časový úsek (napríklad kvartál - v jednotlivých kvartáloch sú počty dní takmer rovnaké). Niekedy netreba priamo upravovať dáta, stačí si iba dávať pozor pri interpretácii výsledkov.
- Zmena podmienok, respektíve charakteru časového radu - napríklad predikovať počty nakazených počas lockdownu na základe dát z času, kedy lockdown nebol môže byť náročné, až nemožné. V princípe celá predikcia je založená na tom, že ak poznáme minulosť, resp. prítomnosť, tak *ak sa nezmení charakter časového radu*, tak vieme povedať, ako bude zhruba vyzeráť budúcnosť. Toto je aj jeden z dôvodov, prečo robiť iba kratšie predikcie - čím ďalej do budúcnosti by sme išli, tým je väčšia šanca, že sa niečo zmení (ďalší dôvod je ten, že čím ďalej do budúcnosti ideme, tým viac sa prejavia potenciálne nepresnosti odhadov/meraní, ktoré sme na základe súčasných dát spravili).

Podobne aj pri deskripcii môže byť niekedy vhodnejšie rozdeliť časový rad na viac homogénnych úsekov. Napríklad na Obr. 1 je modrou farbou zobrazený časový rad, ktorý sa síce správa lineárne, ale do času 50 podľa nejakého predpisu a od času podľa nejakého iného predpisu. (V čase 50 sa zmení charakter.) Ak by sme to ale ignorovali a použili iba jednu priamku, tak najlepšie (v nejakom zmysle), čo môžeme dostať je priamka zobrazená na Obr. 1 červenou farbou, čo zjavne vystihuje časový rad veľmi zle.

Takisto, ak by sme mali v tomto príklade k dispozícii iba dáta do času 50, tak kvôli zmene charakteru nemáme šancu rozumne predikovať správanie časového radu v časoch 51, 52, ...



Obr. 1: Príklad zmeny charakteru časového radu

- Voľba dátových bodov - niekedy máme možnosť si zvoliť, kedy, prípadne ako často budeme pozorovať hodnoty sledovaného znaku - treba dbať na to, aby to bolo zmysluplné a aj matematicky zvládnuteľné - napríklad pri numerických výpočtoch môže veľký počet meraní výrazne sťažiť výpočet (nebudeme preto napríklad sledovať počet zákazníkov v obchode každú minútu - nakoľko tento počet sa až tak často mení nebude - stačí napríklad celkový počet za deň, prípadne hodinu). Naopak ak chceme nájsť nejaké zákonitosti, tak meraní potrebujeme dostatočne veľa, respektíve často (napríklad ak máme iba ročné dáta, nepovie nám to nič o mesačných fluktuáciách).

## Prístupy k modelovaniu časových radov

Prístupov je veľa, z časových dôvodov sa ale budeme venovať iba nasledovným dvom (zrejme najpoužívanejším):

1. Dekompozičný prístup - ten už sme spomínali vyššie - ide o rozloženie časového radu na trend, sezónnu zložku a šum (reziduálna zložka). Väčšinou sa uvažuje tzv. aditívny model

$$X_t = Tr_t + S_t + E_t,$$

avšak niekedy môže dáta lepšie vystihovať tzv. multiplikatívny model tvaru

$$X_t = Tr_t S_t E_t.$$

V takom prípade je ale bežným postupom zlogaritmovanie dát (resp. celého modelu), čím dostaneme

$$\ln X_t = \ln Tr_t + \ln S_t + \ln E_t,$$

čo je opäť aditívny model - z toho dôvodu sa multiplikatívnemu modelu nebudeme až tak venovať (spomenieme iba zopár metód).

Spomeňme ešte, že v dekompozičnom prístupe sa niekedy uvažuje aj takzvaná cyklická zložka  $C_t$  charakterizovaná striedajúcou fázou rastu a fázou poklesu, ktorá ale nemusí byť úplne pravidelná (napr. v literatúre sa dá stretnúť s pojmom „business cycle“, ktorého dĺžka fluktuuje medzi 5-7 rokmi (aj keď aj tento údaj sa mení v závislosti od literatúri)). Ide o trochu spornú zložku (vzhľadom na nie úplne poriadnu definovateľnosť) a často špecifickú pre nejakú oblasť, takže na tomto kurze sa jej nebudeme venovať.

2. Box-Jenkinsonova metodológia - základným pojmom v tomto prístupe je tzv. biely šum - to je časový rad  $\{W_t\}_{t \in \mathbb{Z}}$  s vlastnosťami

- $E(W_t) = 0$  pre každé  $t \in \mathbb{Z}$ ,
- $Cov(W_t, W_s) = \begin{cases} 0 & \text{ak } t \neq s, \\ \sigma^2 & \text{ak } t = s, \end{cases}$  pre ľubovoľné  $t, s \in \mathbb{Z}$ , kde  $\sigma^2 > 0$  je nejaký parameter.

Pomocou bieleho šumu sa potom vytvárajú tzv. ARMA modely, ktoré majú tvar

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q},$$

kde  $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  sú parametre. Časový rad teda v tomto prípade modelujeme ako nejakú lineárnu kombináciu minulosti ( $X_{t-1}, \dots, X_{t-p}$ ) a bieleho šumu ( $W_s$  pre  $s \in \mathbb{Z}$  predstavuje akýsi náhodný impulz v čase  $s$ ). Na rozdiel od mnohých iných štatistických metód, v ARMA modeloch sa predpokladá (a využíva) závislosť jednotlivých meraní  $X_1, X_2, \dots$ . Výhodou takýchto modelov je ich vysoká flexibilita (vedia sa prispôbiť rôznym, aj „divokým“ tvarom časového radu) a kvalitné predikcie (spravidla lepšie ako pri dekompozičných metódach). Naopak nevýhodou je ich ťažšia interpretovateľnosť (často je ťažké povedať, čo sa skrýva za bielym šumom a za hodnotami jednotlivých parametrov).

## 2 Dekompozičné metódy

Pripomeňme, že pri (aditívnej) dekompozícii predpokladáme, že časový rad sa dá zapísať v tvare

$$X_t = Tr_t + S_t + E_t,$$

pričom našou úlohou je v princípe nájsť nejaké odhady  $\widehat{Tr}_t$  a  $\widehat{S}_t$  tak, aby  $\widehat{Tr}_t + \widehat{S}_t$  bolo (v nejakom zmysle) čo najbližšie k dátam  $x_t$  pre  $t = 1, \dots, n$ . V niektorých metódach majú tieto odhady nejaký

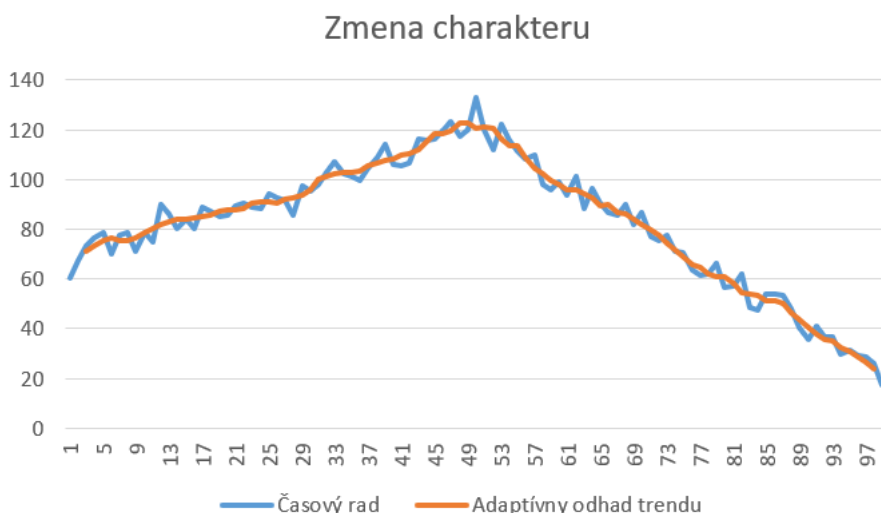
konkrétny predpis, ktorý vieme aplikovať pre ľubovoľné  $t \in \mathbb{Z}$  (nie len pre  $t = 1, \dots, n$ ) - v takom prípade môžeme tieto odhady použiť aj na predikciu - pre nejaké prirodzené číslo  $T > n$  môžeme časový rad v čase  $T$  odhadnúť (predikovať) ako

$$\widehat{X}_T = \widehat{Tr}_T + \widehat{S}_T$$

( $E_T$  odhadneme nulou - nakoľko ide o šum, tak predpokladáme, že sa pohybuje okolo nuly a keďže je to nejaká nesystematická časť, tak 0 je väčšinou najrozumnejšia vec, ktorú môžeme zobrať).

Začneme metódami, ktoré tieto zložky odhadujú postupne, t.j. najprv sa nájde  $\widehat{Tr}_t$  a pomocou neho sa dopočíta  $\widehat{S}_t$ . Ukážeme si spôsoby odhadovania trendu

- matematickými krivkami definovanými menším počtom parametrov (teda úloha odhadovania trendu sa zredukuje na odhad parametrov),
- kľzavými priermi - výhodou tejto metódy je, že je adaptívna (teda vieme odhadnúť aj trendy „škaredších“ tvarov, ktoré by sa jednoduchými funkciami modelovali ťažko - napríklad na Obr. 2 je touto metódou odhadnutý trend dát, ktorý sa ťažko modeloval jednou priamkou - porovnaj s Obr. 1), naopak jej nevýhodou sú predikcie, ktoré sa v tejto metóde robia komplikovane a nie sú až tak dobré (čiže táto metóda je vhodnejšia skôr na deskripciu),



Obr. 2: Príklad použitia adaptívnej metódy na dáta, v ktorých sa mení charakter.

- metódou exponenciálneho vyrovnávania - tiež ide o adaptívnu metódu, avšak s lepšími predikciami.

Neskôr si ukážeme aj metódy, kedy sa trend aj sezónna zložka modelujú naraz.

## 2.1 Modelovanie trendu matematickými krivkami

V tejto časti budeme predpokladať, že trend sa dá zapísať pomocou nejakého malého počtu parametrov, ktoré budeme označovať ako  $\beta_0, \beta_1, \dots, \beta_k$  (prípadne  $\alpha, \beta, \gamma$  - v závislosti od metódy) - trend potom budeme označovať ako  $Tr_t(\beta_0, \beta_1, \dots, \beta_k)$ . Pre predstavu - ak predpokladáme, že trend má tvar priamky, tak sa dá zapísať ako  $Tr_t(\beta_0, \beta_1) = \beta_0 + \beta_1 t$ . Okrem toho krátka poznámka k značeniu - trend je funkcia premennej  $t$  a  $\beta_0, \beta_1, \dots, \beta_k$  sú parametre - „štandardnejšie“ značenie by teda malo byť skôr  $Tr_{\beta_0, \beta_1, \dots, \beta_k}(t)$  - to ale môže spôsobiť, že v indexe bude priveľa písmeniek, čo nepôsobí úplne prehľadne. Navyše v časových radoch je premenná  $t$  bežne v indexe - preto budeme preferovať značenie  $Tr_t(\beta_0, \beta_1, \dots, \beta_k)$ , resp. ak bude zjavné, čo sú parametre trendu, tak iba  $Tr_t$ .

Úloha odhadnutia trendu sa tým pádom zjednoduší na úlohu odhadnutia parametrov - tie spravidla (a na tomto kurze výlučne) hľadáme metódou najmenších štvorcov, t.j. hľadáme také odhady  $\hat{\beta}_0, \dots, \hat{\beta}_k$ , ktoré minimalizujú funkciu  $g$  definovanú ako

$$g(\beta_0, \dots, \beta_k) \equiv \sum_{t=1}^n (Tr_t(\beta_0, \dots, \beta_k) - x_t)^2. \quad (2.1)$$

V niektorých prípadoch túto úlohu vieme vyriešiť analyticky zderivovaním funkcie  $g$ , inokedy budeme hľadať optimálne hodnoty parametrov numerickými metódami.

Výhodou takejto parametrizácie je aj možnosť v niektorých prípadoch porovnať trendy dvoch časových radov - napríklad ak máme dve priamky, tak vieme povedať, ktorá rastie rýchlejšie.

Ďalej spomenieme najbežnejšie používané trendy a spôsoby odhadu parametrov.

### Konštantný „trend“<sup>1</sup>

V takomto prípade predpokladáme, že  $Tr_t(\beta_0) = \beta_0$ . Odhad  $\beta_0$  môžeme nájsť deriváciou funkcie

$$g(\beta_0) = \sum_{t=1}^n (\beta_0 - x_t)^2.$$

Derivácia tejto funkcie je

$$g'(\beta_0) = \sum_{t=1}^n 2(\beta_0 - x_t).$$

Pre odhad  $\hat{\beta}_0$  musí platiť, že  $g'(\hat{\beta}_0) = 0$ , teda

$$\begin{aligned} \sum_{t=1}^n 2(\hat{\beta}_0 - x_t) &= 0 \\ \sum_{t=1}^n \hat{\beta}_0 &= \sum_{t=1}^n x_t \\ n\hat{\beta}_0 &= \sum_{t=1}^n x_t \\ \hat{\beta}_0 &= \bar{x}. \end{aligned}$$

Čiže konštantný trend odhadneme ako aritmetický priemer dát (čo je vcelku zmysluplné).

### Lineárny trend

V tomto prípade predpokladáme, že trend má tvar priamky, t.j.  $Tr_t(\beta_0, \beta_1) = \beta_0 + \beta_1 t$ . Funkcia  $g$  a jej parciálne derivácie majú tvar

$$\begin{aligned} g(\beta_0, \beta_1) &= \sum_{t=1}^n (\beta_0 + \beta_1 t - x_t)^2, \\ \frac{\partial g}{\partial \beta_0}(\beta_0, \beta_1) &= \sum_{t=1}^n 2(\beta_0 + \beta_1 t - x_t), \\ \frac{\partial g}{\partial \beta_1}(\beta_0, \beta_1) &= \sum_{t=1}^n 2(\beta_0 + \beta_1 t - x_t)t. \end{aligned}$$

<sup>1</sup> Ťažko povedať, či konštantu možno nazvať trendom, preto je slovíčko trend v úvodzovkách. V takomto prípade skôr hovoríme, že v dátach trend nie je (a iba fluktuujú okolo nejakej konštantnej hodnoty).

Pre odhady  $\hat{\beta}_0, \hat{\beta}_1$  musí platiť  $\frac{\partial g}{\partial \beta_0}(\hat{\beta}_0, \hat{\beta}_1) = 0$  a zároveň  $\frac{\partial g}{\partial \beta_1}(\hat{\beta}_0, \hat{\beta}_1) = 0$ , čiže v princípe ide o sústavu dvoch rovníc o dvoch neznámych. Po dosadení do prvej rovnice dostávame

$$\sum_{t=1}^n 2(\hat{\beta}_0 + \hat{\beta}_1 - x_t) = 0$$

$$n\hat{\beta}_0 = \sum_{t=1}^n x_t - \hat{\beta}_1 \sum_{t=1}^n t$$

$$\hat{\beta}_0 = \bar{x} - \hat{\beta}_1 \bar{t}$$

( $\bar{t}$  sa dá jednoducho vyjadriť ako  $\frac{n+1}{2}$ , ale nie je to až tak podstatné - môžeme to nechať v tomto tvare). Z druhej rovnice (už aj s dosadením za  $\hat{\beta}_0$ ) dostávame

$$\sum_{t=1}^n 2(\hat{\beta}_0 + \hat{\beta}_1 t - x_t)t = 0$$

$$\sum_{t=1}^n (\bar{x} - \hat{\beta}_1 \bar{t} + \hat{\beta}_1 t - x_t)t = 0$$

$$\bar{x}n\bar{t} - \hat{\beta}_1 \bar{t}^2 n + \hat{\beta}_1 n\bar{t}^2 - \sum_{t=1}^n x_t t = 0$$

$$\hat{\beta}_1 (n\bar{t}^2 - n\bar{t}^2) = \sum_{t=1}^n x_t t - n\bar{x}\bar{t}$$

$$\hat{\beta}_1 = \frac{\frac{1}{n} \sum_{t=1}^n x_t t - \bar{x}\bar{t}}{\bar{t}^2 - \bar{t}^2}.$$

Pozor, medzi  $\bar{t}^2$  a  $\bar{t}^2$  je rozdiel - prvý výraz znamená priemer druhých mocnín hodnôt  $1, 2, 3, \dots, n$  a druhý výraz znamená priemer hodnôt  $1, 2, 3, \dots, n$  umocnený na druhú. Opäť platí, že tieto výrazy sa dajú vyjadriť čisto pomocou  $n$ , ale v praxi to aj tak budeme počítajú v Exceli, kde je takýto zápis možno aj rýchlejší. Hodnotu  $\hat{\beta}_1$  by sme mohli spätne dosadiť do vzorca pre  $\hat{\beta}_0$ , avšak opäť to pri praktickej realizácii nie je potrebné (keďže budeme vyčíslávať hodnotu  $\hat{\beta}_1$  a priamo toto číslo môžeme využiť pri výpočte  $\hat{\beta}_0$ ).

## Kvadratický trend

Ďalej budeme predpokladať, že trend má tvar  $Tr_t(\beta_0, \beta_1, \beta_2) = \beta_0 + \beta_1 t + \beta_2 t^2$ . Opäť by sme mohli využiť parciálne derivácie, avšak jednoduchšie je ísť na odhad parametrov maticovo - konkrétne funkciu  $g$  môžeme zapísať ako

$$g(\beta) = (X - F\beta)^\top (X - F\beta),$$

kde

- $\beta$  je vektor parametrov, čiže  $\beta = (\beta_0, \beta_1, \beta_2)^\top$ ,
- $X$  je vektor meraní, čiže  $X = (x_1, \dots, x_n)^\top$  a
- $F$  je matica tvaru

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ \vdots & \vdots & \vdots \\ 1 & n & n^2 \end{pmatrix}.$$

Ak by sme si rozpísali výraz  $X - F\beta$ , dostali by sme  $n$ -rozmerný vektor, ktorý má na  $t$ -tom mieste hodnotu  $x_t - (\beta_0 + \beta_1 t + \beta_2 t^2)$ , takže vo funkcii  $g$  naozaj dostaneme výraz

$$\sum_{t=1}^n (x_t - (\beta_0 + \beta_1 t + \beta_2 t^2))^2.$$

Tým, že umocňujeme na druhú, tak je jedno, či v sume ide najprv  $x_t$ , alebo hodnota trendu. To, že naozaj výjde súčet štvorcov je jednoduchá vlastnosť násobenia vektorov - napr. ak máme nejaký

číselný vektor  $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ , tak platí

$$a^\top a = (a_1 \quad a_2 \quad a_3) \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1^2 + a_2^2 + a_3^2.$$

Dá sa ukázať, že v takomto prípade funkcia  $g$  nadobúda minimum v bode

$$\hat{\beta} = (F^\top F)^{-1} F^\top X.$$

Zrejme bolo na predmete Pravdepodobnosť a štatistika, prípadne Numerických metódach, budeme o tom hovoriť aj na Matematickej štatistike. Nejaká intuícia za tým je, že  $g(\beta)$  sa dá rozpísať ako

$$g(\beta) = (X - F\beta)^\top (X - F\beta) = X^\top X - 2\beta^\top F^\top X + \beta^\top F^\top F\beta$$

a následne vieme použiť akési maticové derivovanie (funguje veľmi podobne ako pri normálnom derivovaní, len treba dávať pozor na rozmery matíc, resp. vektorov) - ak máme v sčítanci nemáme betu, tak sa zderivuje na nulu, ak tam máme jednu betu (ako nejaký činiteľ), tak odstránime iba betu a ak máme v sčítanci dve bety (tiež vo forme súčinu), tak tam necháme iba jednu a vynásobíme dvomi - v tomto prípade

$$\frac{\partial g}{\partial \beta}(\beta) = -2F^\top X + 2F^\top F\beta.$$

Pre odhad  $\hat{\beta}$  teda musí platiť

$$-2F^\top X + 2F^\top F\hat{\beta} = 0,$$

z čoho po úprave naozaj dostávame

$$\hat{\beta} = (F^\top F)^{-1} F^\top X.$$

## Polynomický trend

Všeobecne môžeme uvažovať polynomický trend, čiže  $Tr_t(\beta_0, \beta_1, \dots, \beta_k) = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_k t^k$ , pričom všetko funguje analogicky ako v kvadratickom trende, jedine vektor parametrov  $\beta$  je o niečo „dlhší“ a matica  $F$  je o niečo „širšia“, konkrétne v tomto prípade má tvar

$$F = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \dots & n^k \end{pmatrix}.$$

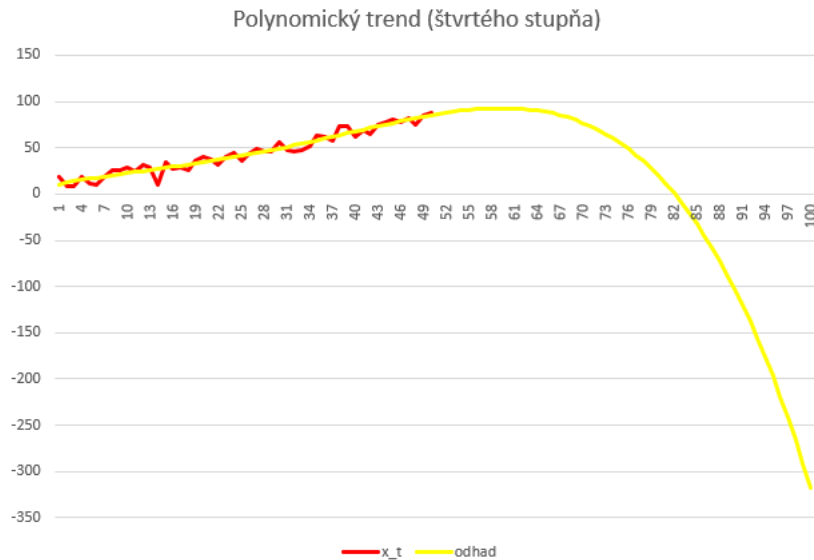
Odhad parametrov výjde rovnako ako v predchádzajúcom, čiže

$$\hat{\beta} = (F^\top F)^{-1} F^\top X.$$

Môžeme si všimnúť, že tento prípad zahŕňa všetky doteraz spomenuté. V praxi sa najbežnejšie stretne s lineárnym, prípadne kvadratickým trendom - vyššie stupne polynómu sú zriedkavé. Platí síce, že čím vyšší stupeň polynómu, tým lepší bude „fit“ (to ako presne krivka prechádza dátovými bodmi), avšak často je to nežiadúce, lebo



- od trendu ani principiálne nechceme, aby úplne presne opisoval dátové body  $x_1, \dots, x_n$  - chceme iba akési „priemerné správanie“ časového radu,
- polynóm vysokého stupňa má tú zľú vlastnosť, že veľmi rýchlo „uletí“ do plus alebo mínus nekonečna - často sa teda stane, že v tejto situácii bude síce model pre dátové body ako tak v poriadku, ale bude dávať veľmi nereálne predikcie. Napríklad na Obr. 3 sú zobrazené dáta, ktoré by sa dali modelovať lineárnym trendom - ak ich budeme modelovať polynómom štvrtého stupňa, tak síce je fit dobrý, podobne aj krátkodobé predikcie, avšak časom tento polynóm veľmi rýchlo začne klesať do  $-\infty$ .



Obr. 3: Príklad predikcií získaných použitím polynomickeho trendu (štvrtého stupňa).

### Predikčné intervaly v polynomickeho trende

Ak by náhodou v časovom rade nebola sezónna zložka a ak by platilo, že  $E_1, E_2, \dots$  sú nezávislé a všetky majú rozdelenie  $N(0, \sigma^2)$ , tak pre „budúcnosť“ časového radu  $X_T$ , ( $T > n$ ) možno konštruovať 95% predikčné intervaly (interval, do ktorého na 95% „padne  $X_T$ “) - tie majú tvar

$$\left( \hat{X}_T - t_{n-(k+1)}(97.5\%)sf_T, \hat{X}_T + t_{n-(k+1)}(97.5\%)sf_T \right),$$

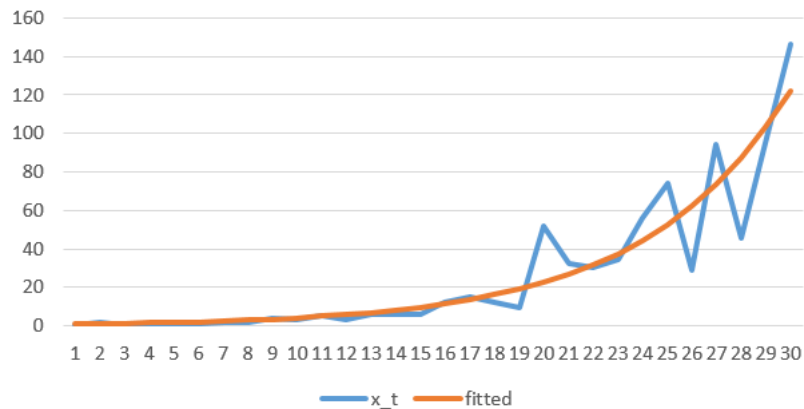
kde

- $\hat{X}_T$  je predikcia časového radu v čase  $T$  (získame ako  $\hat{\beta}_0 + \hat{\beta}_1 T + \dots + \hat{\beta}_k T^k$ ),
- $s = \sqrt{\frac{(X - F\hat{\beta})^\top (X - F\hat{\beta})}{n - (k + 1)}}$ ,
- $f_T = \sqrt{1 + (1, T, T^2, \dots, T^k)(F^\top F)^{-1}(1, T, T^2, \dots, T^k)^\top}$ ,
- $t_{n-(k+1)}(97.5\%)$  je 97.5% kvantil  $t$  rozdelenia s  $n - (k + 1)$  stupňami voľnosti (v Exceli dostaneme ako T.INV(0.975,  $n - (k + 1)$ )).

Je možné použiť aj iné percento ako 95% (v takom prípade potom stačí zmeniť kvantil  $t$  rozdelenia), avšak toto je zrejme najpoužívanejšie. Ak by sme chceli  $\alpha$  percentný predikčný interval, miesto 97.5% kvantilu  $t$  rozdelenia použijeme  $100\% - \frac{100\% - \alpha}{2}$  percentný kvantil. Ďalej je potrebné poznamenať, že predpoklad o nezávislosti  $E_1, E_2, \dots$ , je v časových radoch skôr výnimkou ako pravidlom, preto je pred prípadnou konštrukciou takýchto intervalov vhodné sa presvedčiť, že predpoklad o nezávislosti a rovnako aj o normalite je splnený, inak môžu mať výrazne nižšiu spoľahlivosť, ako si myslíme.

## Exponenciálny trend

V tomto prípade predpokladáme, že  $Tr_t(\alpha, \beta) = \alpha\beta^t$ , kde  $\beta > 0$ . Ak  $\alpha > 0$ , tak  $\beta > 1$  znamená exponenciálny nárast a  $\beta \in (0, 1)$  znamená exponenciálny pokles ( $\beta = 1$  znamená konštantný trend). Odhady parametrov  $\alpha, \beta$  vieme nájsť opäť metódou najmenších štvorcov - v tomto prípade ale nemáme explicitné vyjadrenie pre  $\hat{\alpha}$  a  $\hat{\beta}$  - odhady je potrebné nájsť numericky (to sa dá aj v Exceli, návod bude na konci tejto časti).



Obr. 4: Exponenciálny trend - parametre odhadnuté numericky MNŠ.

Pri dátach exponenciálneho tvaru môže dobre fungovať aj multiplikatívny model - v takom prípade po zlogaritmovaní dostávame

$$\ln Tr_t(\alpha, \beta) = \ln \alpha + t \ln \beta, \quad (2.2)$$

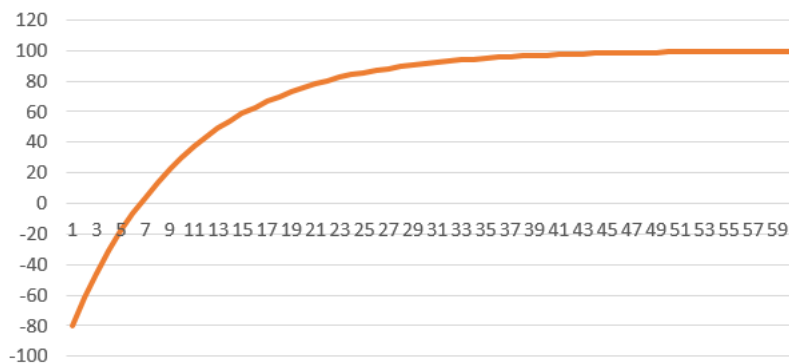
čo je už lineárny trend s parametrami  $\ln \alpha$  a  $\ln \beta$ . Odhady  $\widehat{\ln \alpha}$  a  $\widehat{\ln \beta}$  vieme nájsť rovnako ako pri lineárnom trende, teda minimalizáciou výrazu

$$\sum_{t=1}^n (\ln x_t - \ln \alpha - t \ln \beta)^2. \quad (2.3)$$

Z odhadov  $\widehat{\ln \alpha}$  a  $\widehat{\ln \beta}$  potom vieme vypočítať odhady  $\tilde{\alpha} = e^{\widehat{\ln \alpha}}$  a  $\tilde{\beta} = e^{\widehat{\ln \beta}}$ . Pozor,  $\tilde{\alpha} \neq \hat{\alpha}$  a  $\tilde{\beta} \neq \hat{\beta}$  - ide o rôzne odhady (ale obidva v nejakom zmysle rozumné).

## Modifikovaný exponenciálny trend

V tomto prípade má trend tvar  $Tr_t(\alpha, \beta, \gamma) = \gamma + \alpha\beta^t$ ,  $\beta > 0$ , čiže môžeme exponenciálnu krivku posúvať v smere  $y$ -ovej osi. Odhady parametrov  $\alpha, \beta, \gamma$  hľadáme numericky metódou najmenších



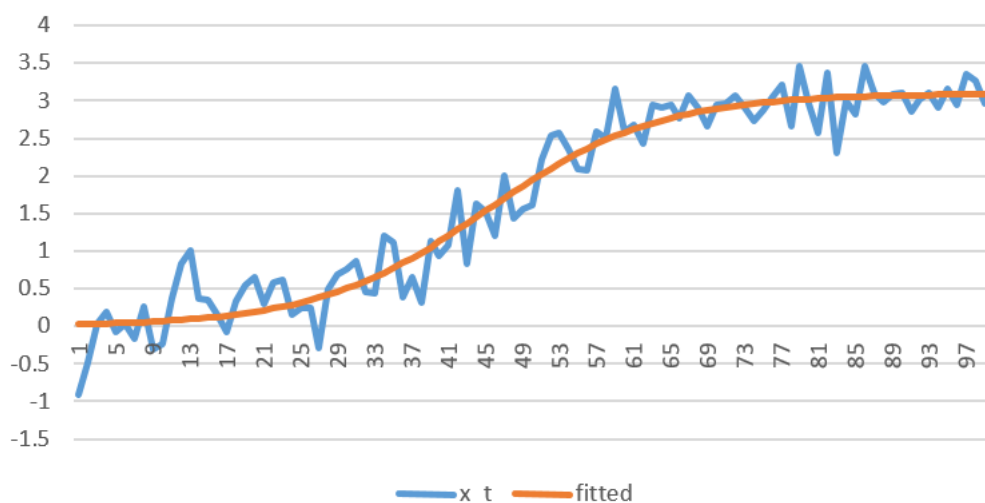
Obr. 5: Príklad modifikovaného exponenciálneho trendu pre  $\alpha < 0, \beta \in (0, 1)$  a  $\gamma > 0$ .

štvorcov. V trende máme teraz súčet, takže zlogaritmovanie by nám nepomohlo (na odhady existujú aj „kalkulačkové postupy“, ale numerické riešenie je v dnešnej dobe rýchlejšie aj presnejšie).

Keďže máme viac parametrov, tak táto modifikovaná verzia bude vždy presnejšia - v praxi preto môžeme pri exponenciálne vyzerajúcich dátach vždy vyskúšať túto (aj pri podozrení, že to je iba obyčajný exponenciálny trend) a „v najhoršom“ nám  $\hat{\gamma}$  výjde 0 (resp. blízko 0).

## Logistický trend

Logistický trend je daný vzorcom  $Tr_t(\alpha, \beta, \gamma) = \frac{\gamma}{1 + \alpha\beta^t}$ ,  $\alpha, \beta, \gamma > 0$ . Z predpisu vyplýva, že ak  $\beta \in (0, 1)$ , tak  $Tr_t \rightarrow 0$  pre  $t \rightarrow -\infty$  a  $Tr_t \rightarrow \gamma$  pre  $t \rightarrow \infty$  (ak  $\beta > 1$ , tak naopak). Krivka má „esíkovitý“ tvar a je symetrická okolo inflexného bodu  $-\frac{\ln \alpha}{\ln \beta}$  (viď Obr. 6) - kto by si náhodou nepamätal, tak inflexný bod je taký bod, v ktorom sa mení funkcia z konvexnej na konkávnu (resp. naopak). Odhady parametrov metódou najmenších štvorcov hľadáme numericky.



Obr. 6: Odhad logistického trendu.

## Gompertzova krivka

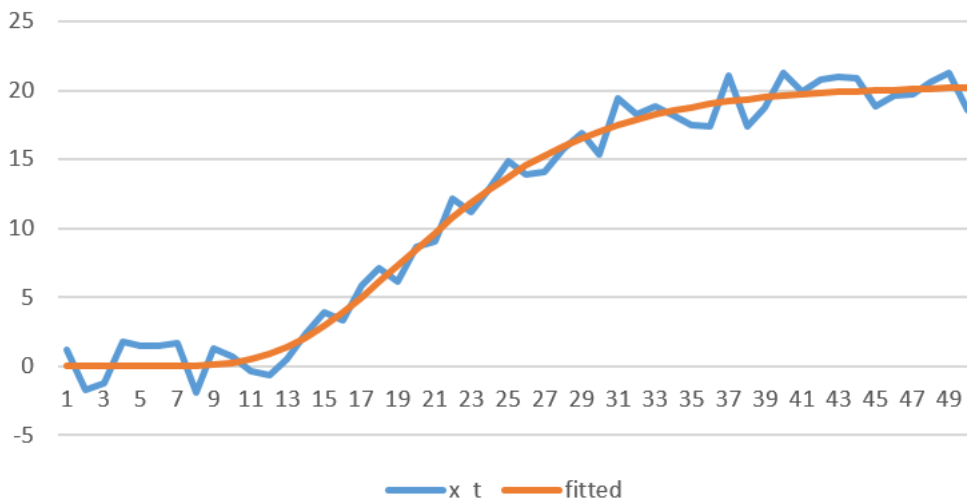
Ako poslednú spomenieme Gompertzovu krivku - tá má predpis  $Tr_t(\alpha, \beta, \gamma) = e^{\gamma + \alpha\beta^t}$ ,  $\beta > 0, \alpha < 0$ . Tvarovo vie byť takáto krivka veľmi podobná logistickému trendu, hlavný rozdiel je, že Gompertzova krivka nie je symetrická okolo inflexného bodu  $-\frac{\ln(-\alpha)}{\ln \beta}$  (viď Obr. 7 - trochu to tam vidno). Odhady parametrov hľadáme numericky.

## Voľba krivky

Najdôležitejšie je vykresliť si dáta a sledovať aký majú tvar - z toho sa dá často usúdiť, aká krivka by mohla byť vhodná. V niektorých prípadoch nám niekto túto informáciu dá, prípadne vieme typ krivky uhádnuť z povahy dát (niektoré deje sa prirodzene správajú exponenciálne alebo napr. v populačnej biológii je zase celkom bežná Gompertzova krivka a podobne). V prípade, že si nie sme istý, treba vyskúšať viac kriviek a pozrieť sa, ktorá lepšie popisuje dáta - napr. ktorá z nich má menšiu sumu štvorcov rezíduí

$$\sum_{t=1}^n (x_t - \widehat{Tr}_t)^2, \quad (2.4)$$

prípadne (ak to je zjavné) z obrázka. V učebnici možno na strane 40 nájsť zopár informatívnych testov pre voľbu krivky - skúsil som ich na generovaných dátach a pravdupovediac veľmi nesedeli, takže skôr neodporúčam.

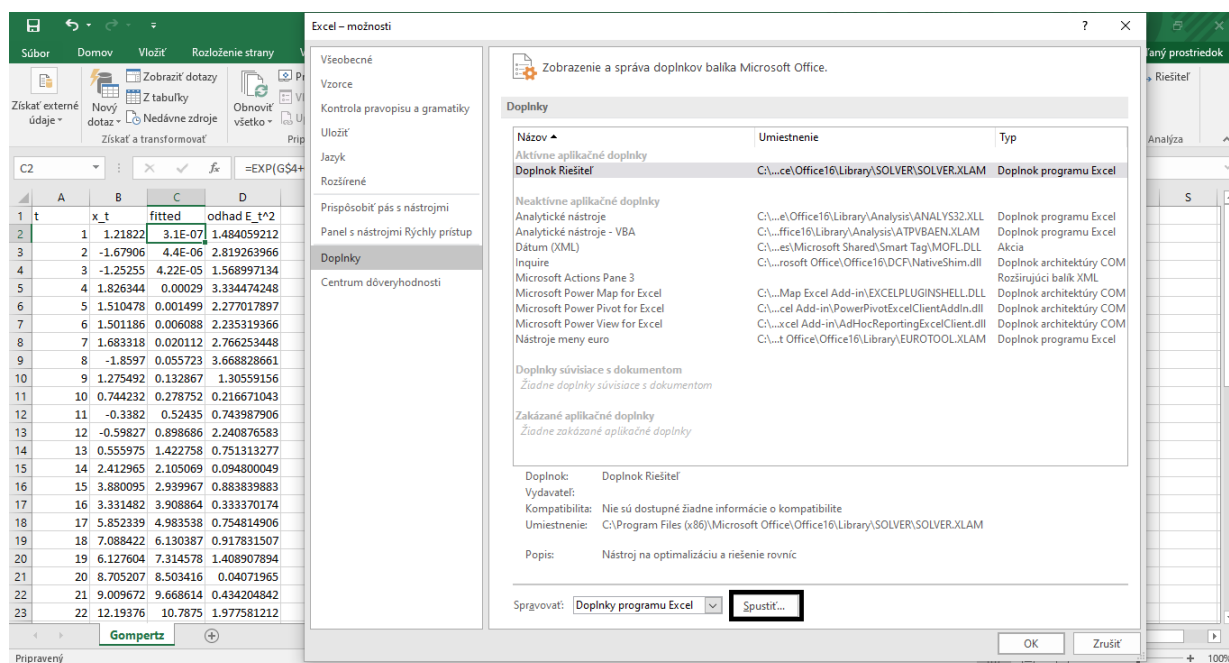


Obr. 7: Odhad Gompertzovej krivky.

## Numerické hľadanie parametrov metódou najmenších štvorcov v Exceli

V prvom rade si treba spustiť doplnok „Riešiteľ“ (Solver). To spravíme nasledovne:

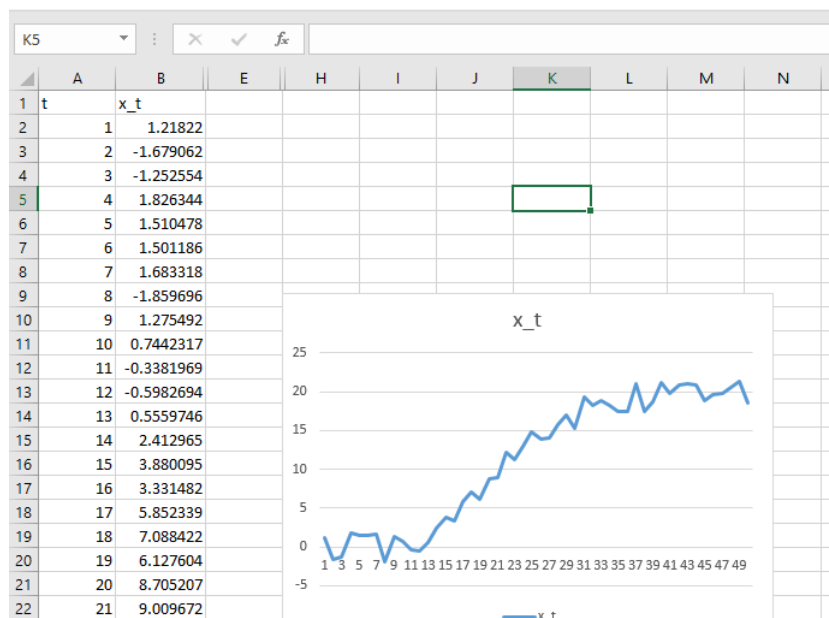
- Pôjdeme na Súbor → Možnosti → Doplnky (File → Options → Add-ins).
- Dole je „Spravovať“ (Manage)- tam vyberieme „Doplnky programu Excel“ (Excel Add-ins) a stlačíme Spustiť (Go).
- Zaklikneme „Doplnok Riešiteľ“ (Solver Add-in) a klikneme OK.



Obr. 8: Spúšťanie Excelovského doplnku „Riešiteľ“.

Samotné hľadanie si ukážeme na príklade - povedzme, že máme dáta ako na Obr. 9 a rozhodneme sa, že chceme cez tieto dáta preložiť nejakú Gompertzovu krivku - musíme teda nájsť vhodné parametre  $\alpha, \beta, \gamma$ . Na úvod treba spraviť v Exceli zopár prípravných krokov:

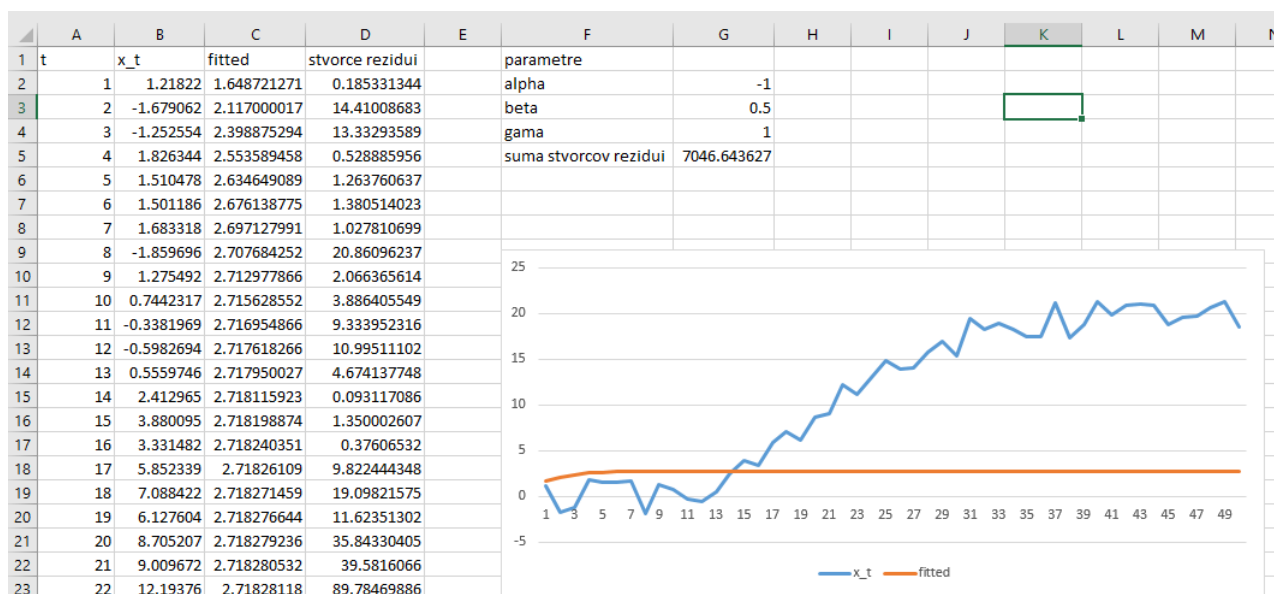
- Nastavíme si hodnoty nejakých štartovacích parametrov, napr.  $\alpha_0 = -1, \beta_0 = 0.5, \gamma_0 = 1$ . Tieto čísla zadáme napr. do buniek G2, G3 a G4 (viď Obr. 10). Čím bližšie sú štartovacie hodnoty k optimálnym parametrom, tým lepšie, ale v takejto jednoduchšej úlohe to Excel



Obr. 9

nšťastie zvládne aj s nepresnými hodnotami (aké sme dali my, vid' oranžová čiara v grafe na Obr. 10). Štartovacie parametre by mali spĺňať podmienky pre túto krivku, čiže  $\alpha_0 < 0$  a  $\beta_0 > 0$ .

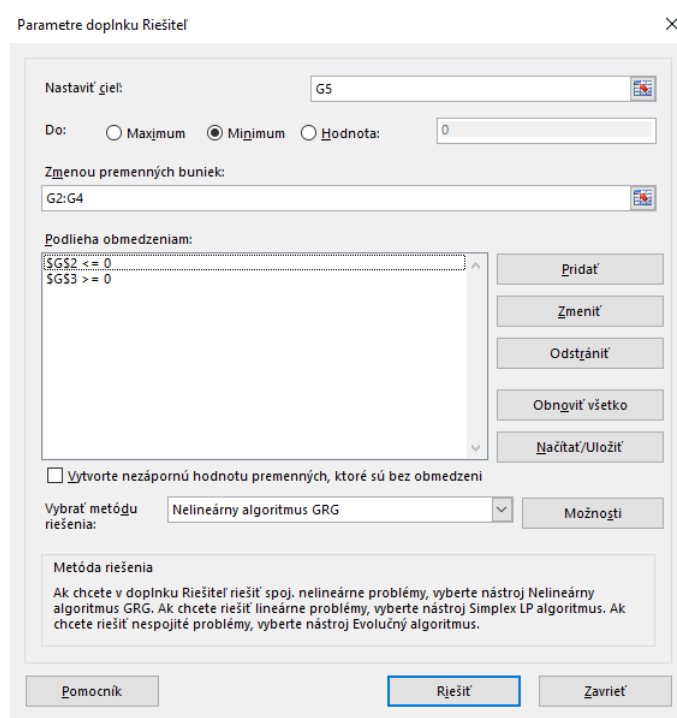
- Vypočítame aké by boli fitované hodnoty pri takýchto parametroch - t.j. vypočítame funkciu  $Tr_t(\alpha_0, \beta_0, \gamma_0)$  pre  $t = 1, \dots, n$  - v našom príklade sme si tieto hodnoty dali do stĺpca C. Ak chceme, môžeme funkciu  $Tr_t(\alpha_0, \beta_0, \gamma_0)$  pridať aj do grafu.
- Vypočítame štvorce rezíduí pri takýchto parametroch, t.j. hodnoty  $(x_t - Tr_t(\alpha_0, \beta_0, \gamma_0))^2$  pre  $t = 1, \dots, n$  - v našom príklade si ich uložíme do stĺpca D.
- Vypočítame sumu štvorcov rezíduí pri týchto parametroch, t.j.  $\sum_{t=1}^n (x_t - Tr_t(\alpha_0, \beta_0, \gamma_0))^2$  a uložíme si ju napr. do políčka G5. Všetky tieto výpočty musia byť robené („previazané“) pomocou Excelovských vzorcov, čiže ak by sme teraz zmenili napr. hodnotu  $\beta$  (bunka G3), mali by sa nám zmeniť aj hodnoty v stĺpcoch C, D a hodnota v bunke G5



Obr. 10

Teraz môžeme v Exceli prejsť na Údaje → Riešiteľ (Data → Solver). Tam nastavíme (vid' Obr. 11):

- V „Nastaviť cieľ” (Set Objective) zadáme bunku G5 (tú chceme minimalizovať).
- V „do” (to) nastavíme Minimum (lebo minimalizujeme).
- V „Zmenou premenných buniek” (By Changing Variable Cells) zadáme G2:G4 (tam sú naše parametre).
- V „Podlieha obmedzeniam” (Subject to the Constraints) pridáme  $G2 \leq 0$  (lebo  $\alpha < 0$ ) a  $G3 \geq 0$  (lebo  $\beta > 0$ ) - ostré nerovnosti sa tam nedajú pridať, preto je tam aj rovnosť.
- „Vytvorte nezápornú hodnotu premenných, ktoré sú bez obmedzení” (Make Unconstrained Variable Non-Negative) odklikneme (to nechceme).
- Vo „Vybrať metódu riešenia” (Select a Solving Method) zvolíme Nelineárny algoritmus GRG (GRG nonlinear).
- Spustíme „Riešiť” (Solve).



Obr. 11

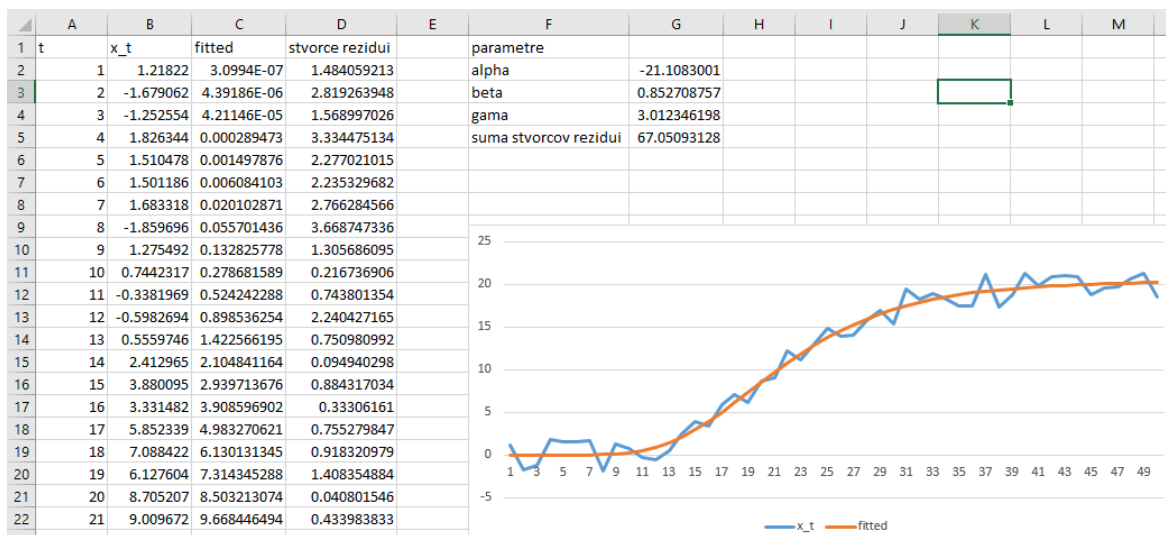
Excel nám oznámi, či našiel riešenie (v takýchto úlohách väčšinou nie je problém), ktoré môžeme skontrolovať (napr. z grafu). Ak by sme s ním náhodou neboli spokojný, môžeme to skúsiť zopakovať s inými štartovacími parametrami. V našom prípade to našlo riešenia  $\hat{\alpha} \approx -21.11$ ,  $\hat{\beta} \approx 0.85$  a  $\hat{\gamma} \approx 3.01$ , čo podľa grafu vyzerá byť v poriadku (viď Obr. 12).

### 3 Metóda kľzavých priemerov

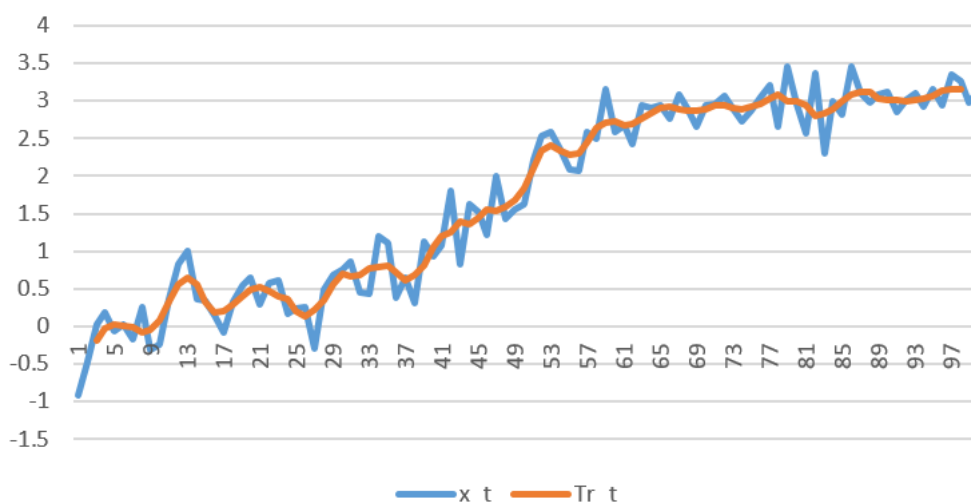
V tejto metóde odhadujeme trend váženým priemerom okolitých hodnôt, napríklad

$$\widehat{Tr}_t = \frac{1}{8}(x_{t-2} + 2x_{t-1} + 2x_t + 2x_{t+1} + x_{t+2}), \quad (3.1)$$

pre  $t = 3, 4, \dots, n - 2$  (viď Obr. 13). Zjednodušene môžeme pravú stranu výrazu (3.1) zapísať ako  $\frac{1}{8}(1, 2, 2, 2, 1)x_t$ . Táto metóda patrí medzi tzv. adaptívne metódy, t.j. vie sa „prispôbiť” zmenám v trende (ktoré by sme napríklad nedokázali „pokryť”, ak by sme sa obmedzili iba na jednu z kriviek spomenutých v predchádzajúcej časti). V nasledujúcich častiach sa zameriame na nasledovné otázky:



Obr. 12



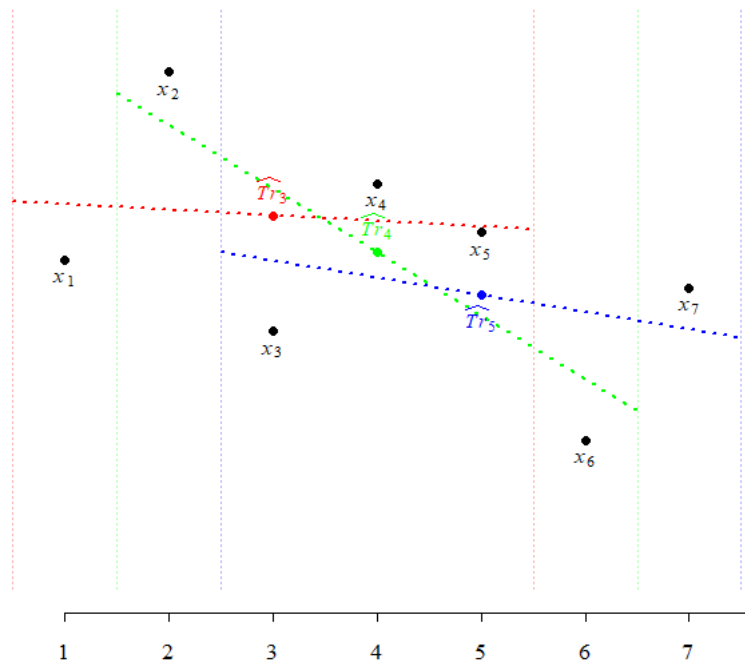
Obr. 13: Odhad trendu klzavým priemerom  $\frac{1}{81}(1, 2, 2, 2, 1)$ .

- Aká je za tým idea?
- Ako rozumne voliť váhy vo váženom priemere?
- Ako riešiť okrajové body (napr. v (3.1)  $\widehat{Tr}_t$  v časoch  $t = 1, 2, n - 1, n - 2$ ), prípadne predikcie trendu  $\widehat{Tr}_{n+1}, \widehat{Tr}_{n+2}, \dots$ ?
- Ako voliť tzv. „šírku okna“? (Koľko okolitých bodov brať?)

### Aká je za tým idea?

Idea je taká, že trend budeme modelovať priamkou/polynómom lokálne - napríklad pre prvých päť bodov  $x_1, x_2, x_3, x_4, x_5$  nájdeme priamku  $\beta_0(3) + t\beta_1(3)$  (trojka v zátvorke je na počesť toho, že berieme body v okolí bodu  $x_3$ ), ktorá je k týmto bodom najbližšie ( $\beta_0(3), \beta_1(3)$  nájdeme metódou najmenších štvorcov) a  $Tr_3$  odhadneme ako  $\beta_0(3) + 3\beta_1(3)$ . Ďalej sa posunieme o jeden časový úsek „doprava“ - bodom  $x_2, x_3, x_4, x_5, x_6$  nájdeme priamku  $\beta_0(4) + t\beta_1(4)$ , ktorá je k nim najbližšie a  $Tr_4$  odhadneme ako  $\beta_0(4) + 4\beta_1(4)$ . Analogicky sa posúvame pozdĺž celých dát (viď Obr. 14).

Konkrétny výpočet si ukážeme pre polynóm tretieho stupňa v prípade, že šírka okna je 5. Majme teda nejakú päťicu bodov  $x_{t-2}, x_{t-1}, x_t, x_{t+1}, x_{t+2}$  - chceme nájsť polynóm tretieho stupňa, ktorý je k týmto bodom najbližšie a hodnotu  $Tr_t$  potom odhadneme ako funkčnú hodnotu tohto polynómu



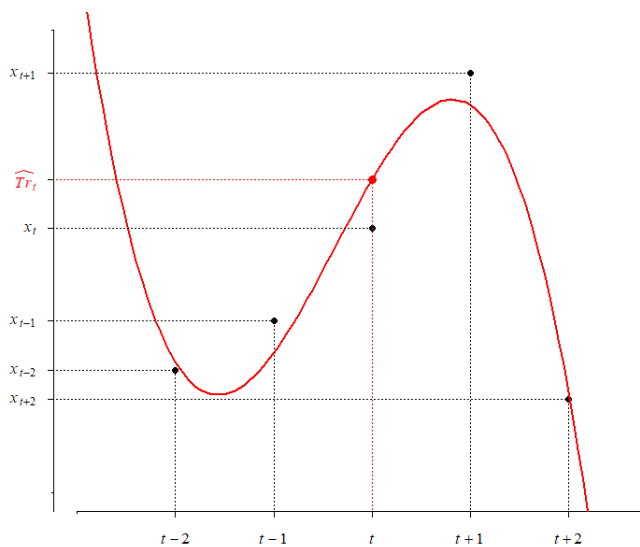
Obr. 14

v bode  $t$  (viď Obr. 15). Užitočná vec, ktorú si môžeme uvedomiť je, že z hľadiska odhadu hodnoty  $Tr_t$  sa nič nestane, keď si celú situáciu „prenesieme“ v smere  $x$ -ovej osi tak, aby bola symetricky okolo bodu 0 (viď Obr. 16) - toto nám značne zjednoduší výpočty. Hľadáme teda funkciu  $f_t$  ( $t$  v indexe je preto, lebo riešime  $Tr_t$ ) tvaru

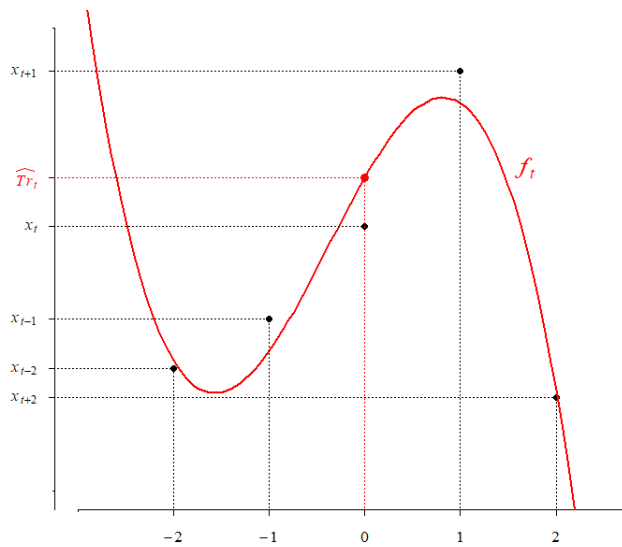
$$f_t(\tau) = \beta_0 + \beta_1\tau + \beta_2\tau^2 + \beta_3\tau^3, \quad (3.2)$$

ktorá bude v bodoch  $-2, -1, 0, 1, 2$  čo najbližšie k hodnotám  $x_{t-2}, x_{t-1}, x_t, x_{t+1}, x_{t+2}$  (patrilo by sa písať  $\beta_0(t), \dots, \beta_3(t)$ , ale pre jednoduchosť to budeme vynechávať). Reálne nás potom bude zaujímať iba  $f_t(0) = \beta_0$  (viď Obr. 16) - to je jedna z prvých výhod toho, že sme si dáta posunuli v smere  $x$ -ovej osi. Parametre  $\beta_0, \beta_1, \beta_2$  a  $\beta_3$  hľadáme metódou najmenších štvorcov, teda minimalizáciou výrazu

$$g(\beta_0, \beta_1, \beta_2, \beta_3) = \sum_{\tau=-2}^2 (x_{t+\tau} - \beta_0 - \beta_1\tau - \beta_2\tau^2 - \beta_3\tau^3)^2 \quad (3.3)$$



Obr. 15



Obr. 16



Deriváciou podľa  $\beta_0$  dostaneme

$$\begin{aligned}\frac{\partial g}{\partial \beta_0}(\beta_0, \beta_1, \beta_2, \beta_3) &= \sum_{\tau=-2}^2 2(x_{t+\tau} - \beta_0 - \beta_1\tau - \beta_2\tau^2 - \beta_3\tau^3) \cdot (-1) = \\ &= -2 \sum_{\tau=-2}^2 x_{t+\tau} + 2\beta_0 \sum_{\tau=-2}^2 1 + 2\beta_1 \sum_{\tau=-2}^2 \tau + 2\beta_2 \sum_{\tau=-2}^2 \tau^2 + 2\beta_3 \sum_{\tau=-2}^2 \tau^3 = \\ &= -2 \sum_{\tau=-2}^2 x_{t+\tau} + 10\beta_0 + 20\beta_2,\end{aligned}$$

kde sme využili, že  $\sum_{\tau=-2}^2 \tau^k = 0$  pre nepárne  $k$  (a  $\sum_{\tau=-2}^2 \tau^k$  pre párne  $k$  vieme ľahko vyčísliť) - to je ďalšia výhoda toho, že sme si dáta posunuli v smere  $x$ -ovej osi tak, aby boli symetricky okolo 0. Podobne sa nám zjednodušia aj ďalšie derivácie

$$\begin{aligned}\frac{\partial g}{\partial \beta_1}(\beta_0, \beta_1, \beta_2, \beta_3) &= -2 \sum_{\tau=-2}^2 \tau x_{t+\tau} + 20\beta_1 + 68\beta_3 \\ \frac{\partial g}{\partial \beta_2}(\beta_0, \beta_1, \beta_2, \beta_3) &= -2 \sum_{\tau=-2}^2 \tau^2 x_{t+\tau} + 20\beta_0 + 68\beta_2 \\ \frac{\partial g}{\partial \beta_3}(\beta_0, \beta_1, \beta_2, \beta_3) &= -2 \sum_{\tau=-2}^2 \tau^3 x_{t+\tau} + 68\beta_1 + 260\beta_3.\end{aligned}$$

Keďže nás zaujíma iba  $\beta_0$ , stačí zobrať iba  $\frac{\partial g}{\partial \beta_0}$  a  $\frac{\partial g}{\partial \beta_2}$  a položiť ich nule, z čoho dostaneme dve rovnice o dvoch neznámych

$$\begin{aligned}5\beta_0 + 10\beta_2 &= \sum_{\tau=-2}^2 x_{t+\tau} & / \cdot 17 \\ 10\beta_0 + 34\beta_2 &= \sum_{\tau=-2}^2 \tau^2 x_{t+\tau} & / \cdot (-5) \\ \hline 85\beta_0 + 170\beta_2 &= 17 \sum_{\tau=-2}^2 x_{t+\tau} \\ -50\beta_0 - 170\beta_2 &= -5 \sum_{\tau=-2}^2 \tau^2 x_{t+\tau}.\end{aligned}$$

Sčítaním týchto dvoch rovníc dostaneme

$$\begin{aligned}35\beta_0 &= 17(x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}) - 5(4x_{t-2} + x_{t-1} + x_{t+1} + 4x_{t+2}) = \\ &= -3x_{t-2} + 12x_{t-1} + 17x_t + 12x_{t+1} - 3x_{t+2},\end{aligned}$$

čiže

$$\widehat{Tr}_t = \beta_0 = \frac{1}{35}(-3, 12, 17, 12, -3)x_t. \quad (3.4)$$

Takúto lineárnu kombináciu okolitých hodnôt by sme dostali pri polynóme ľubovoľného stupňa a pri ľubovoľnej šírke okna - jediné, čo sa bude meniť, sú váhy, ktoré priradíme jednotlivým okolitým hodnotám (odvodenie by bolo rovnaké a pri takomto postupe nám nemá ako vzniknúť niečo iné ako lineárna kombinácia).

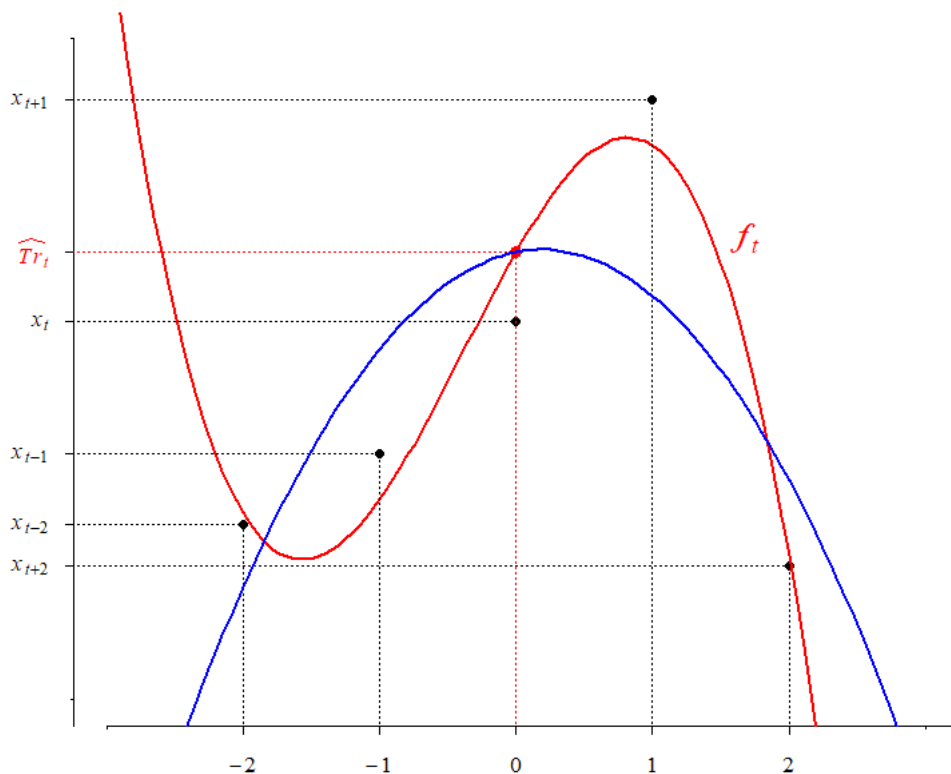
Platí, že

1. súčet váh bude vždy 1, pretože váhy  $w_{-m}, \dots, w_m$  nezávisia od hodnôt  $x_{t-m}, \dots, x_{t+m}$  ( $m$  je nejaké prirodzené číslo určujúce šírku okna) a ak by sa stalo, že všetky  $x_{t-m}, \dots, x_{t+m}$  sú rovné nejakej konštante  $x^*$  tak pri polynóme stupňa  $d$  ( $d$  je ľubovoľné nezáporné celé číslo) nám musí výjsť  $\widehat{Tr}_t = \beta_0 = x^*$  a  $\beta_1 = \dots = \beta_d = 0$  (lebo takýto „polynóm“ by dokonale prechádzal všetkými bodmi) - platilo by teda

$$x^* = \beta_0 = \sum_{\tau=-m}^m w_{\tau} x_{t+\tau} = \sum_{\tau=-m}^m w_{\tau} x^* = x^* \sum_{\tau=-m}^m w_{\tau}, \quad (3.5)$$

z čoho máme  $\sum_{\tau=-m}^m w_{\tau} = 1$ ;

2. váhy su symetrické okolo stredy, pretože nám v odvodení (pri ľubovoľnom  $m$  a  $d$ ) budú  $x_{t-m}, \dots, x_{t+m}$  vždy vystupovať iba v sumách  $\sum_{\tau=-m}^m \tau^k x_{t+\tau}$ , kde  $k$  je párne číslo, t.j. pri  $x_{t-1}$  bude vo všetkých úpravách vždy rovnaký koeficient ako pri  $x_{t+1}$ , podobne pre  $x_{t-2}$  a  $x_{t+2}$ , atď.;
3. ak stupeň polynómu  $d$  je nepárne číslo, tak rovnaké váhy by sme dostali aj ak by sme miesto neho zobrali iba polynóm stupňa  $d-1$  (viď Obr. 17) - napr. ak by v našom príklade vo funkcii  $g(\beta_0, \beta_1, \beta_2, \beta_3)$  chýbal posledný člen  $\beta_3 \tau^3$  (čiže by sme hľadali iba polynóm druhého stupňa), tak na  $\frac{\partial g}{\partial \beta_0}$  a  $\frac{\partial g}{\partial \beta_2}$  (z ktorých sme počítali  $\beta_0$ ) by sa nič nezmenilo; analogicky by to fungovalo aj vo všeobecnosti („geometricky“ to pritom vôbec nie je intuitívne).



Obr. 17: Najbližšia možná (MNŠ) parabola a kubický polynóm.

V Tabuľke 1 sú uvedené váhy pre rôzne stupne polynómu a rôzne šírky okna. Keďže váhy sú symetrické, väčšinou je uvedená iba polovica týchto váh. Situácie, kedy je stupeň polynómu väčší ako šírka okna, nemajú veľmi zmysel, preto nie sú uvedené. Prívetký zmysel nemajú ani situácie, kedy je stupeň polynómu rovnaký ako šírka okna, pretože  $d$  bodov vieme vždy interpolovať polynómom  $d$ -teho stupňa, takže v takom prípade výjde  $\widehat{Tr}_t = x_t$ , čo nie je zrovna účelom trendu. Za povšimnutie stojí aj situácia  $d = 0$ , resp.  $d = 1$ , kedy vychádza obyčajný aritmetický priemer - tomu sa hovorí **jednoduchý kľzavý priemer**.

		Stupeň polynómu		
		$d = 0, d = 1$	$d = 2, d = 3$	$d = 4, d = 5$
Šírka okna	1	(1)	-	-
	3	$\frac{1}{3}(1, 1, 1)$	(0,1,0)	-
	5	$\frac{1}{5}(1, 1, 1, 1, 1)$	$\frac{1}{35}(-3, 12, 17, \dots)$	(0,0,1,0,0)
	7	$\frac{1}{7}(1, 1, 1, 1, \dots)$	$\frac{1}{21}(-2, 3, 6, 7, \dots)$	$\frac{1}{231}(5, -30, 75, 131, \dots)$
	9	$\frac{1}{9}(1, 1, 1, 1, 1, \dots)$	$\frac{1}{231}(-21, 14, 39, 54, 59, \dots)$	$\frac{1}{429}(15, -55, 30, 135, 179, \dots)$

Tabuľka 1

## Okrajové body a predikcie trendu

Keď už vieme, že za kľzavými priemerami sa „skrýva“ lokálna aproximácia polynómom, túto znalosť môžeme využiť aj pri riešení okrajových bodov, resp. predikcií. Uvažujme opäť náš príklad (polynóm tretieho stupňa, šírka okna 5) a predpokladajme, že chceme nájsť odhad pre  $Tr_{n-1}$  (už nemôžeme použiť

$$\frac{1}{35}(-3, 12, 17, 12, -3)x_{n-1} = \frac{1}{35}(-3x_{n-3} + 12x_{n-2} + 17x_{n-1} + 12x_n - 3x_{n+1}),$$

pretože hodnotu  $x_{n+1}$  už nemáme k dispozícii). Miesto toho ale môžeme použiť posledný polynóm  $f_{n-2}(\tau)$  - pripomeňme, že to je taký, ktorý je v bodoch  $-2, -1, 0, 1, 2$  najbližšie bodom  $x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n$ . Na odhad  $Tr_{n-2}$  sme použili hodnotu tohto polynómu v bode 0, t.j.  $f_{n-2}(0) = \beta_0(n-2)$  a na odhad  $Tr_{n-1}$  môžeme podobne použiť

$$\widehat{Tr}_{n-1} = f_{n-2}(1) = \beta_0(n-2) + \beta_1(n-2) + \beta_2(n-2) + \beta_3(n-2) \quad (3.6)$$

a analogicky

$$\widehat{Tr}_n = f_{n-2}(2) = \beta_0(n-2) + 2\beta_1(n-2) + 4\beta_2(n-2) + 8\beta_3(n-2). \quad (3.7)$$

Rovnako tento polynóm môžeme využiť aj na konštrukciu predpovede trendu

$$\widehat{Tr}_{n+T} = f_{n-2}(2+T) = \beta_0(n-2) + (2+T)\beta_1(n-2) + (2+T)^2\beta_2(n-2) + (2+T)^3\beta_3(n-2), \quad (3.8)$$

pre  $T > 0$ . Treba ale poznamenať, že takáto metóda je vhodná iba na krátkodobé predpovede. Podobne by sme vedeli vypočítať aj  $\widehat{Tr}_1$  a  $\widehat{Tr}_2$  pomocou  $f_3$ .

Vidíme, že tentokrát nám nestačí iba  $\beta_0$ , potrebujeme vypočítať aj  $\beta_1, \beta_2, \beta_3$ . Stále však platí, že výsledkom bude vážený priemer hodnôt  $x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n$  - už ale nebude symetrický. Napríklad v našom príklade by vyšlo

$$\begin{aligned} \widehat{Tr}_{n-1} &= f_{n-2}(1) = \frac{1}{35}(2, -8, 12, 27, 2)x_{n-2}, \\ \widehat{Tr}_n &= f_{n-2}(2) = \frac{1}{70}(-1, 4, -6, 4, 69)x_{n-2}, \\ \widehat{Tr}_{n+1} &= f_{n-2}(3) = \frac{1}{5}(-4, 11, -4, -14, 16)x_{n-2}. \end{aligned}$$

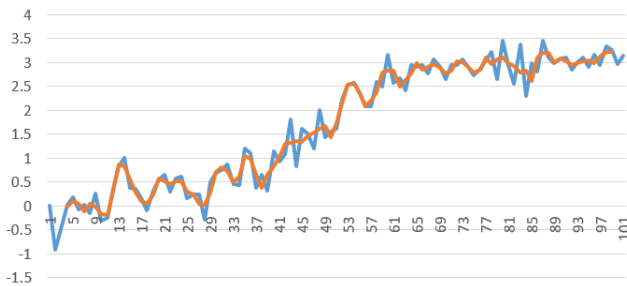
Takisto už neplatí, že pre polynóm nepárneho stupňa  $d$  výjdu rovnaké váhy ako pre polynóm stupňa  $d-1$ .

V prípade jednoduchých kľzavých priemerov (ak bolo  $d=0$ ), výjdu okrajové body aj predikcie stále ako obyčajné aritmetické priemery posledných hodnôt. V prípade  $d > 0$  sú váhy vo váženom priemere tabelizované pre rôzne  $d$  a rôzne šírky okna. Keďže ale výpočet týchto váh bude súčasťou domácej úlohy, tabuľky takéhoto typu neuvedieme.

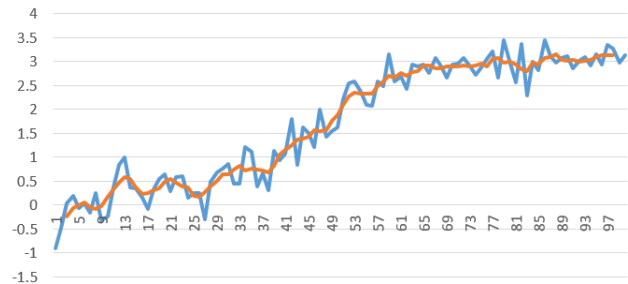
## Voľba stupňa polynómu a šírky okna

Síce existuje zopár objektívnych metód (ktoré sú ale nad rámec tohto kurzu), voľba týchto parametrov je stále do istej miery veľmi subjektívna. Často treba skúsiť viac možností a vybrať si takú, ktorá sa nám „pozdáva najviac“. Uvedieme aspoň zopár zásad, ktoré je dobré dodržiavať.

- Trend má zachytávať dlhodobý priemerný vývoj - nemusí teda čo najpresnejšie opisovať dáta, skôr je lepší menej „rozlietaný“ odhad trendu. Napríklad z Obr. 18 a Obr. 19 by sme preferovali skôr Obr. 19.

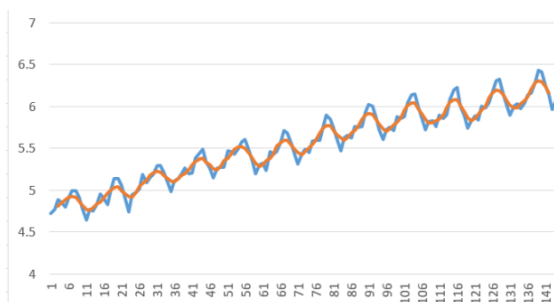


Obr. 18

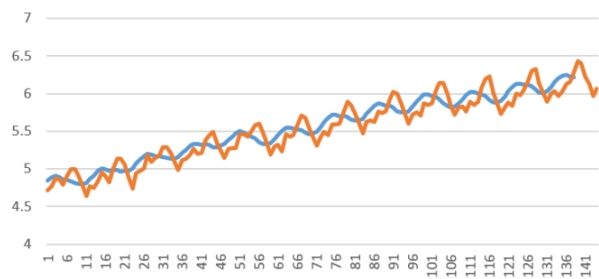


Obr. 19

- Hoci sú jednoduché kľzavé priemery jednoduché, často môžu byť postačujúce - stačí poskúšať rôzne šírky okna. Polynómy vyšších stupňov môžu byť fajn, ak chceme hladší trend, na druhej strane môžu niekedy spôsobiť prílišné „kopírovanie“ dát (ako na Obr. 18).
- Ak je v dátach sezónna zložka, kľzavý priemer by ju mal ignorovať - sezónna zložka sa rieši samostatne (nie je „úlohou“ trendu riešiť sezónne výkyvy). Napríklad by sa nemalo diať to, čo je na Obr. 20 (oranžovou je jednoduchý kľzavý priemer šírky 5, t.j.  $\frac{1}{5}(1, 1, 1, 1, 1)$ ) a rozhodne nie to, čo je na Obr. 21 (oranžovou je jednoduchý kľzavý priemer šírky 7) - tam to osciluje a ešte aj v „opačných časoch“ ako dáta.

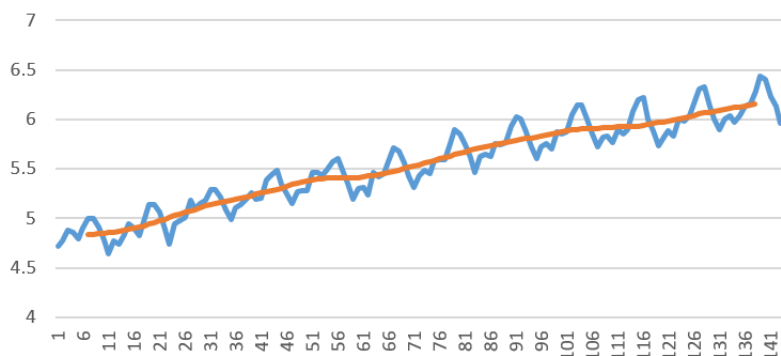


Obr. 20



Obr. 21

- Ak máme dáta zo sezónnou zložkou, pričom dĺžka periódy  $p$  je nepárna, odporúča sa použiť jednoduchý kľzavý priemer  $\frac{1}{p}(1, 1, \dots, 1)$  - to nám vie dobre pokryť priemerné správanie časového radu (lebo berieme do úvahy všetky časové obdobia v danej perióde). Ak je dĺžka periódy  $p$  párna (väčšinou sa stretávame skôr s takouto situáciou - napríklad pri mesačných alebo kvartálnych dátach), odporúča sa použiť takzvaný centrovaný kľzavý priemer - ten má tvar  $\frac{1}{2p}(1, 2, 2, \dots, 2, 1)$  (šírka je  $p+1$ ). Idea za tým je taká, že ak by sme mali napríklad  $p = 4$ , tak ak by sme uvažovali niečo typu  $\frac{1}{4}(x_{t-2} + x_{t-1} + x_t + x_{t+1})$  (aby sme rovnomerne zobrali do úvahy každé obdobie z danej periódy), tak toto by odhadovalo hodnotu  $Tr_{t-0.5}$  (stred medzi týmito bodmi), čo ale nechme. Ak však zoberieme aj  $\frac{1}{4}(x_{t-1} + x_t + x_{t+1} + x_{t+2})$  (čo zodpovedá



Obr. 22: Centrováný kľzavý priemer aplikovaný na mesačné dáta

odhadu  $Tr_{t+0.5}$ ), na odhad hodnoty  $Tr_t$  môžeme tieto dve hodnoty spriemerovať, t.j.

$$\begin{aligned}\widehat{Tr}_t &= \frac{1}{2} \left( \widehat{Tr}_{t-0.5} + \widehat{Tr}_{t+0.5} \right) = \\ &= \frac{1}{2} \left( \frac{1}{4}(x_{t-2} + x_{t-1} + x_t + x_{t+1}) + \frac{1}{4}(x_{t-1} + x_t + x_{t+1} + x_{t+2}) \right) = \\ &= \frac{1}{8}(x_{t-2} + 2x_{t-1} + 2x_t + 2x_{t+1} + x_{t+2}) = \frac{1}{8}(1, 2, 2, 2, 1)x_t,\end{aligned}$$

čo je práve centrováný kľzavý priemer (šírky 5). Na Obr. 22 je príklad centrovaného kľzavého priemeru šírky 13, čiže  $\frac{1}{24}(1, 2, 2, \dots, 2, 1)$ .

## 4 Exponenciálne vyrovnávanie

Na úvod poznámka k terminológii - namiesto „vyrovnávanie“ sa používajú aj pojmy „zhladzovanie“, prípadne „vyhladzovanie“. Všetky tieto tri pojmy sa zvyknú používať v súvislosti s hľadaním trendu, keďže tieto veci s pôvodným časovým radom trend robí (viď skoro všetky obrázky).

V tejto časti budeme opäť predpokladať, že časový rad nemá sezónnu zložku, t.j. platí

$$X_t = Tr_t + E_t \quad (4.1)$$

(v prípade, že by sezónnu zložku mal, používa sa veľmi podobná metóda - o tom trochu neskôr).

Idea exponenciálneho vyrovnávania je podobná ako pri metóde kľzavých priemerov, t.j. trend v istom zmysle lokálne modelujeme polynómom stupňa  $d$ . V tomto prípade ale odhady parametrov tohto polynómu nehľadáme pomocou okolitých hodnôt, ale pomocou všetkých predchádzajúcich pozorovaných hodnôt. Okrem toho sa pri tejto metóde modifikuje metóda najmenších štvorcov takým spôsobom, že jednotlivým „štvorcov“ sa priradia váhy, ktoré sa „smerom do minulosti“ exponenciálne znižujú (odtiaľ názov metódy) - čiže čím staršie pozorovanie, tým má menší vplyv.

V praxi sa ako stupeň polynómu používajú najmä  $d = 0$  (tzv. jednoduché exponenciálne vyrovnávanie),  $d = 1$  (dvojité exponenciálne vyrovnávanie) a niekedy  $d = 2$  (trojité exponenciálne vyrovnávanie) - my si povieme o jednoduchom a dvojitom exponenciálnom vyrovnávaní.

### Jednoduché exponenciálne vyrovnávanie

V tomto prípade predpokladáme, že trend je približne konštantna, t.j.  $Tr_t \approx \beta_0$  - „približne“ v takom zmysle, že s meniacim sa časom sa môže táto konštantna trochu meniť. V zmysle úvodu v čase  $t$  odhadujeme  $\beta_0$  (tento odhad budeme označovať  $\widehat{\beta}_0(t)$ ) tak, aby sme minimalizovali funkciu

$$g_t(\beta_0) \equiv (x_t - \beta_0)^2 + \alpha(x_{t-1} - \beta_0)^2 + \alpha^2(x_{t-2} - \beta_0)^2 + \dots = \sum_{j=0}^{\infty} \alpha^j (x_{t-j} - \beta_0)^2, \quad (4.2)$$

kde  $\alpha \in (0, 1)$  (nazýva sa aj vyrovnávací konštanta - o jej voľbe o chvíľu). Pozorný čitateľ si iste všimol, že sme použili naozaj všetky predchádzajúce hodnoty - dokonca aj tie, čo reálne nemáme (t.j.  $x_0, x_{-1}, x_{-2}, \dots$ ) - v teoretických úvahách je ale „krajšie“ (a jednoduchšie) pracovať aj s týmito hodnotami a to, že ich nemáme riešiť až po odvodení. Derivácia funkcie  $g_t$  má tvar

$$g'_t(\beta_0) = \sum_{j=0}^{\infty} 2\alpha^j (x_{t-j} - \beta_0) \quad (4.3)$$

a ak má  $\widehat{\beta}_0(t)$  minimalizovať funkciu  $g_t$ , musí platiť  $g'_t(\widehat{\beta}_0(t)) = 0$ , čiže

$$\begin{aligned} 2 \sum_{j=0}^{\infty} \alpha^j (x_{t-j} - \widehat{\beta}_0(t)) &= 0 \\ \sum_{j=0}^{\infty} \alpha^j x_{t-j} &= \widehat{\beta}_0(t) \sum_{j=0}^{\infty} \alpha^j = \widehat{\beta}_0(t) \frac{1}{1-\alpha} \\ \widehat{\beta}_0(t) &= \sum_{j=0}^{\infty} (1-\alpha)\alpha^j x_{t-j}. \end{aligned}$$

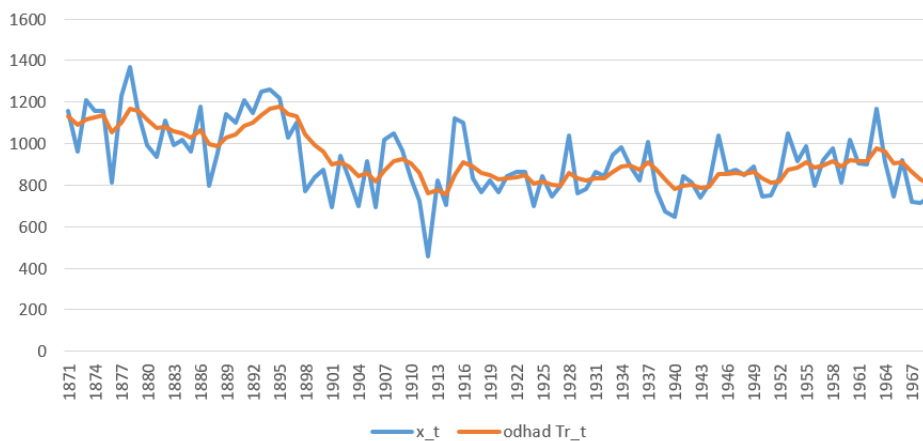
Platí (lebo geometrický rad), že

$$\sum_{j=0}^{\infty} (1-\alpha)\alpha^j = 1, \quad (4.4)$$

takže  $\widehat{\beta}_0(t)$  je vážený priemer hodnôt  $x_t, x_{t-1}, \dots$ . Môžeme si všimnúť, že

$$\begin{aligned} \widehat{\beta}_0(t) &= \sum_{j=0}^{\infty} (1-\alpha)\alpha^j x_{t-j} = (1-\alpha)x_t + \sum_{j=1}^{\infty} (1-\alpha)\alpha^j x_{t-j} = |\text{substitúcia } j = k+1| \\ &= (1-\alpha)x_t + \alpha \sum_{k=0}^{\infty} (1-\alpha)\alpha^k x_{t-1-k} = (1-\alpha)x_t + \alpha \widehat{\beta}_0(t-1), \end{aligned}$$

takže  $\widehat{Tr}_t = \widehat{\beta}_0(t)$  môžeme veľmi jednoducho počítať rekurentne - jediné čo nám chýba je nejaké začiatkové  $\widehat{\beta}_0(1)$ . Na jeho výpočet by sme ale potrebovali aj hodnoty  $x_0, x_{-1}, \dots$ , ktoré ale nemáme. V praxi sa to rieši tak, že za  $\widehat{\beta}_0(1)$  sa položí  $x_1$ , prípadne priemer niekoľkých prvých hodnôt  $\frac{x_1 + \dots + x_m}{m}$  (to ako zvoliť  $m$  nie je vždy jasné, takže môžeme ostať pri  $\widehat{\beta}_0(1) \equiv x_1$ ).



Obr. 23: Jednoduché exponenciálne vyrovnávanie aplikované na ročný prietok rieky Níl mestom Asuán v rokoch 1871-1970.

## Predikcie a voľba $\alpha$

Ak máme k dispozícii  $n$  pozorovaní a chceme urobiť predikciu trendu v čase  $n + h, h > 0$  (toto označíme ako  $\widehat{Tr}_{n+h}(n)$ ), tak vezmeme „najaktuálnejší“ odhad  $\beta_0$ , ktorý máme k dispozícii, čiže

$$\widehat{Tr}_{n+h}(n) = \widehat{\beta}_0(n). \quad (4.5)$$

Keďže predpokladáme model (4.1), tak predikcia trendu je zároveň aj predikcia samotného časového radu (predikciu časového radu v čase  $n + h$  vytvorenú pomocou pozorovaní  $x_1, \dots, x_n$  budeme označovať  $\widehat{x}_{n+h}(n)$ ). V čase  $t = 2, \dots, n$  vieme spočítať aj chybu predikcie (označíme  $\varepsilon_t$ ) ako

$$\varepsilon_t = x_t - \widehat{x}_t(t-1) = x_t - \widehat{Tr}_t(t-1) = x_t - \widehat{\beta}_0(t-1) \quad (4.6)$$

( $\widehat{x}_t(t-1)$  (resp.  $\widehat{Tr}_t(t-1)$ ) je odhad časového radu (resp. trendu) v čase  $t$  pomocou predchádzajúcich hodnôt, čiže pomocou  $x_1, \dots, x_{t-1}$  (tvárimo sa, že  $x_t$  nemáme k dispozícii) - tento odhad vieme porovnať so skutočnou nameranou hodnotou  $x_t$  a to, ako sme „netrafili“ je  $\varepsilon_t$ ).

Parameter  $\alpha$  volíme potom tak, aby súčet štvorcov týchto chýb predikcie (čiže  $\sum_{t=2}^n \varepsilon_t^2$ ) bol čo najmenší - v Exceli na to opäť môžeme použiť Solver.

## Dvojité exponenciálne vyrovňovanie

V tomto prípade predpokladáme, že trend je približne priamka, t.j.  $Tr_t \approx \beta_0 + \beta_1 t$  - „približne“ v rovnakom zmysle ako pri jednoduchom exponenciálnom vyrovňovaní, čiže  $\beta_0$  a  $\beta_1$  sa s meniacim časom môžu trochu meniť. Inak je idea úplne rovnaká, čiže

- odhady  $\widehat{\beta}_0(t)$  a  $\widehat{\beta}_1(t)$  v čase  $t$  získame minimalizáciou funkcie

$$g_t(\beta_0, \beta_1) = \sum_{j=0}^{\infty} \alpha^j [x_{t-j} - (\beta_0 + \beta_1 \cdot (t-j))]^2, \quad (4.7)$$

- trend v čase  $t$  potom odhadneme ako

$$\widehat{Tr}_t = \widehat{\beta}_0(t) + \widehat{\beta}_1(t) \cdot t. \quad (4.8)$$

- ak máme  $n$  meraní, tak predikciu v čase  $n + h, h > 0$  vieme získať ako

$$\widehat{x}_{n+h}(n) = \widehat{Tr}_{n+h}(n) = \widehat{\beta}_0(n) + \widehat{\beta}_1(n) \cdot (n+h), \quad (4.9)$$

- chybu predikcie v čase  $t = 2, \dots, n$  vypočítame ako

$$\varepsilon_t = x_t - \widehat{x}_t(t-1) = x_t - \widehat{Tr}_t(t-1) = x_t - \left( \widehat{\beta}_0(t-1) + \widehat{\beta}_1(t-1) \cdot t \right), \quad (4.10)$$

- parameter  $\alpha$  zvolíme tak, aby sme minimalizovali  $\sum_{t=2}^n \varepsilon_t^2$ .

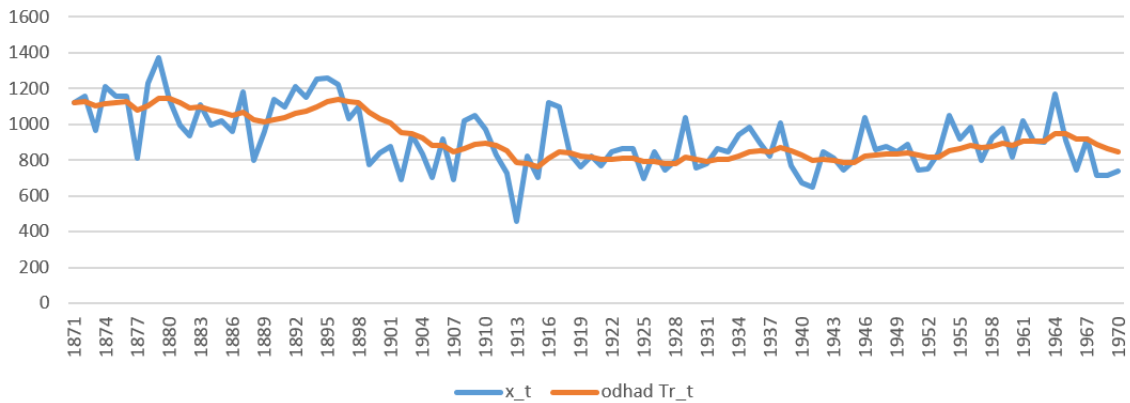
## Výpočet v praxi

Toto síce vyzerá pomerne zložito, avšak v praxi si stačí vypočítať pomocné štatistiky - takzvanú

- jednoduchú vyrovňovaciu štatistiku  $S_t$  danú rekurentne ako

$$S_t \equiv (1 - \alpha)x_t + \alpha S_{t-1} \quad (4.11)$$

pre  $t = 2, 3, \dots$  a  $S_1 = x_1$  (sú aj iné možnosti ako zvoliť  $S_1$ , my ostaňme pri tomto),



Obr. 24: Dvojité exponenciálne vyrovňovanie aplikované na ročný prietok rieky Níl mestom Asuán v rokoch 1871-1970.

- dvojitú vyrovňovaciu štatistiku  $S_t^{[2]}$  danú rekurentne ako

$$S_t^{[2]} = (1 - \alpha)S_t + \alpha S_{t-1}^{[2]} \quad (4.12)$$

pre  $t = 2, 3, \dots$  a  $S_1^{[2]} = x_1$  (opäť sú aj iné možnosti ako  $x_1$ ).

Pomocou nich vieme potom jednoducho vypočítať (odvodenie, že to platí, jednoduché nie je)

- $\widehat{Tr}_t$  pre  $t = 1, \dots, n$  ako  $\widehat{Tr}_t = 2S_t - S_t^{[2]}$ ,
- $\widehat{Tr}_{n+h}(n)$  pre  $h > 0$  ako

$$\widehat{Tr}_{n+h}(n) = 2S_n - S_n^{[2]} + \frac{1 - \alpha}{\alpha}(S_n - S_n^{[2]}) \cdot h \quad (4.13)$$

(špeciálne na výpočet  $\varepsilon_t$  potrebujeme  $\widehat{Tr}_t(t-1) = 2S_{t-1} - S_{t-1}^{[2]} + \frac{1-\alpha}{\alpha}(S_{t-1} - S_{t-1}^{[2]})$ ).

**Poznámka.** Trojité exponenciálne vyrovňovanie by fungovalo veľmi podobne - tam by sme predpokladali  $Tr_t \approx \beta_0 + \beta_1 t + \beta_2 t^2$ . Opäť sa dajú využiť pomocné, rekurentne definované štatistiky (takže to vieme bez problémov „zrátat“ aj v Exceli), avšak tie vzorce už sú dosť komplikované - koho by zaujímali, dajú sa nájsť v učebnici. V praxi ale väčšinou stačí dvojité exponenciálne vyrovňovanie a to aj na dáta, ktoré nemajú „priamkový tvar“ - tým, že v čase trochu meníme  $\beta_0$  a  $\beta_1$ , tak táto metóda sa vie adaptovať aj na komplikovanejšie krivky.

## 5 Modelovanie sezonality

V tejto časti budeme predpokladať model

$$X_t = Tr_t + S_t + E_t \quad (5.1)$$

a budeme sa venovať sezónnej zložke  $S_t$ . Pripomeňme, že táto zložka opisuje periodické zmeny opakujúce sa každý väčší časový úsek. Pod „väčším časovým úsekom“ si môžeme predstaviť napríklad kalendárny rok (v praxi máme často mesačné alebo kvartálne dáta - vtedy sa môžeme stretnúť s takýmto typom sezonality), prípadne deň (ak máme hodinové dáta) alebo týždeň (pri denných dátach). Tým, že tieto zmeny sa pravidelne opakujú každých  $p$  kratších časových období (napr. každých  $p = 12$  mesiacov,  $p = 7$  a pod.), štandardne pre sezónnu zložku predpokladáme  $S_t = S_{t+p}$  pre každé  $t$  (avšak existujú aj adaptívne metódy, kde sa sezónna zložka môže „vyvíjať“, takže platí iba  $S_t \approx S_{t+p}$  - jednu si spomenieme). Na  $S_t$  sa môžeme pozeráť aj ako na očakávanú odchýlku od



trendu (typickú pre obdobie zodpovedajúce času  $t$  - napr. v soboty môžu byť tržby v potravinách v priemere vyššie ako v iných dňoch, naopak v nedele nižšie a pod.), preto sa pridáva aj predpoklad

$$\sum_{j=1}^p S_j = 0 \quad (5.2)$$

(od odchýliek sa takáto vlastnosť očakáva - inak by nám sezónne prvky „vychyľovali trend“). Dĺžku periódy (sezóny)  $p$  spravidla poznáme z typu dát, ale aj keby sme ju nepoznali, tak existujú metódy, ako toto  $p$  nájsť. To, či sezónnu zložku treba vôbec uvažovať sa dá vo väčšine prípadov vidieť priamo z obrázku - pri tom zatiaľ zostaneme a v druhej polovici semestra si povieme aj o sofistikovanejšej technike (z ktorej sa dá dobre zistiť aj  $p$ , ak by sme ho náhodou nepoznali). V ďalších častiach si povieme o niekoľkých prístupoch k odhadovaniu  $S_t$ .

## Jednoduchý prístup

Predpokladajme najprv, že tentokrát nemáme trendovú zložku, t.j. máme iba model

$$X_t = S_t + E_t. \quad (5.3)$$

Zároveň predpokladáme, že pre každé  $t$  platí  $S_t = S_{t+p}$ , čiže napr.  $S_1 = S_{1+p} = S_{1+2p} = \dots$ . Dosadením do modelu (5.3) by sme dostali (dosadzujeme do každej  $p$ -tej „rovnice“)

$$\begin{aligned} X_1 &= S_1 + E_1 \\ X_{1+p} &= S_1 + E_{1+p} \\ X_{1+2p} &= S_1 + E_{1+2p} \\ &\vdots \\ X_{1+kp} &= S_1 + E_{1+kp}, \end{aligned}$$

kde  $k$  je najväčšie prirodzené číslo, pre ktoré  $1 + kp \leq n$  (čiže berieme tie riadky, z ktorých máme dáta - napr. ak by sme mali mesačné dáta, tak zoberieme všetky januárové). Rozumným odhadom  $S_1$  by v takomto prípade bol priemer hodnôt  $x_1, x_{1+p}, \dots, x_{1+kp}$  (viď druhá prednáška - odhadovanie konštantného trendu), čiže

$$\tilde{S}_1 = \frac{x_1 + x_{1+p} + x_{1+2p} + \dots + x_{1+kp}}{k + 1}. \quad (5.4)$$

Analogicky by sme vedeli odhadnúť aj  $S_2$  ako

$$\tilde{S}_2 = \frac{x_2 + x_{2+p} + x_{2+2p} + \dots + x_{2+kp}}{k + 1} \quad (5.5)$$

a rovnako aj  $\tilde{S}_3, \dots, \tilde{S}_p$ .

V praxi ale trend spravidla máme - vtedy postupujeme tak, že tento trend najprv odhadneme nejakou z vyššie spomenutých metód (konkrétna krivka alebo kĺzavé priemery) a potom namiesto hodnôt  $x_t$  pracujeme s hodnotami  $x_t - \widehat{Tr}_t$  (v angličtine sa na takéto hodnoty používa výraz „detrended data“). Výraz (5.4) by sme teda nahradili výrazom

$$\tilde{S}_1 = \frac{(x_1 - \widehat{Tr}_1) + (x_{1+p} - \widehat{Tr}_{1+p}) + (x_{1+2p} - \widehat{Tr}_{1+2p}) + \dots + (x_{1+kp} - \widehat{Tr}_{1+kp})}{k + 1}, \quad (5.6)$$

podobne aj  $\tilde{S}_2, \dots, \tilde{S}_p$ .

Aby naše odhady hodnôt  $S_1, S_2, \dots, S_p$  boli v súlade s podmienkou (5.2), nakoniec sa ešte položí

$$\widehat{S}_t = \tilde{S}_t - \frac{1}{p}(\tilde{S}_1 + \dots + \tilde{S}_p) = \tilde{S}_t - \frac{1}{p} \sum_{i=1}^p \tilde{S}_i. \quad (5.7)$$

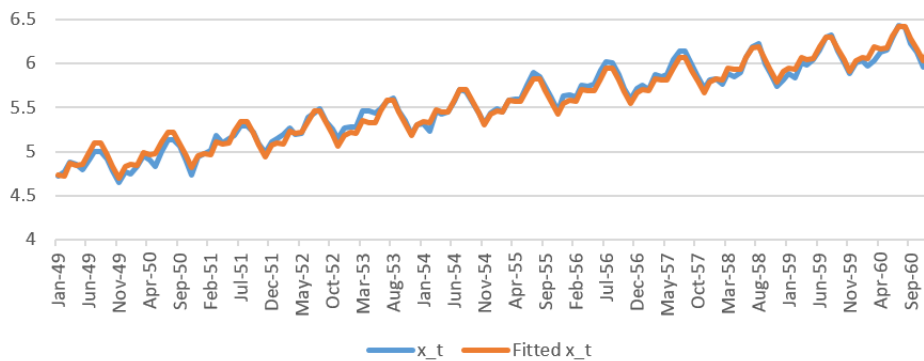
Pre takéto odhady už platí

$$\sum_{j=1}^p \hat{S}_j = \sum_{j=1}^p \left( \tilde{S}_j - \frac{1}{p}(\tilde{S}_1 + \dots + \tilde{S}_p) \right) = \sum_{j=1}^p \tilde{S}_j - p \frac{1}{p}(\tilde{S}_1 + \dots + \tilde{S}_p) = 0. \quad (5.8)$$

Ak sme mali dobrý odhad trendu, tak „korekcia“  $\frac{1}{p} \sum_{i=1}^p \tilde{S}_i$  (ktorú odčítavame od  $\tilde{S}_t$ ) bude iba malá.

Niekedy sa odporúča potom zobrať tzv. „sezónne očistené dáta“  $x_t - \hat{S}_t$  a pre tieto znova odhadnúť trend - môžeme na to použiť aj inú metódu akú sme použili v prvom kroku. Aby sme všetko zhrnuli, tak celý postup by bol nasledovný.

1. Z  $x_1, \dots, x_n$  odhadneme  $\hat{T}r_t$  (nejakou z vyššie spomenutých metód).
2. Vypočítame hodnoty  $x_1 - \hat{T}r_1, \dots, x_n - \hat{T}r_n$ .
3. Pomocou týchto hodnôt nájdeme  $\tilde{S}_1$  (podľa vzorca (5.6)), podobne  $\tilde{S}_2, \dots, \tilde{S}_p$  (z nich už vieme aj  $\tilde{S}_{p+1}, \dots, \tilde{S}_n$ , lebo predpokladáme, že  $S_t = S_{t+p}$ , takže tieto hodnoty sa budú opakovať).
4. Znормujeme odhady  $\tilde{S}_1, \dots, \tilde{S}_p$  tak, aby bol súčet 0, čiže položíme  $\hat{S}_t = \tilde{S}_t - \frac{1}{p} \sum_{i=1}^p \tilde{S}_i$  - tým dostaneme naše odhady sezónnej zložky.
5. Ak chceme, môžeme vypočítať hodnoty  $x_1 - \hat{S}_1, \dots, x_n - \hat{S}_n$  a týmto nájsť nejaký nový odhad trendu  $\hat{T}r_t^{(2)}$  (môže nám to zlepšiť model, t.j. bude lepšie fitovať dáta). V tomto kroku môžeme na odhad trendu použiť inú metódu ako v bode 1 - napr. nie až tak zriedkavá metóda je v prvom kroku použiť klzavé priemery (ktoré dobre odhadujú trend, ale nie sú vôbec vhodné na predikcie) a v tomto kroku použiť nejakú krivku (ktorá už na predikcie vhodná je).
6. Občas sa odporúča aj znova prepočítať sezónnu zložku, t.j. zobrať dáta  $x_t - \hat{T}r_t^{(2)}$  a pokračovať bodom 3 - toto sa dá teoreticky opakovať veľa krát. Pre účely tohto predmetu nám ale bude postačovať prvých päť (prípadne iba prvé štyri) kroky.



Obr. 25: Príklad jednoduchého prístupu aplikovaný na počet (v tisíckach) letecky prepravených pasažierov (zlogaritmované mesačné dáta).

## Regresný prístup

Táto metóda sa používa v situáciach, kedy v dátach uvažujeme polynomický trend. Pre jednoduchosť budeme v tejto časti predpokladať, že

- stupeň polynómu je 1 (teda trend je priamka),
- máme kvartálne dáta, čiže dĺžka periódy  $p$  je 4,

- čas  $t = 1$  zodpovedá prvému kvartálu a čas  $t = n$  zodpovedá štvrtému kvartálu

(vyšší stupeň polynómu a iné  $p$  a iné „rozloženie dát“ by sa spravilo úplne takým istým spôsobom - to len aby sme sa „neupísmenovali“). Na odhad sezónnej zložky nám teda stačí nájsť odhady  $S_1, S_2, S_3, S_4$  - ostatné už dopočítame, lebo  $S_t = S_{t+4}$ . V takejto situácii vieme písať

$$X_t = Tr_t + S_t + E_t = \beta_0 + \beta_1 t + \alpha_1 z_{1t} + \alpha_2 z_{2t} + \alpha_3 z_{3t} + \alpha_4 z_{4t} + E_t, \quad (5.9)$$

kde sezónnu zložku sme si zapísali ako

$$S_t = \alpha_1 z_{1t} + \alpha_2 z_{2t} + \alpha_3 z_{3t} + \alpha_4 z_{4t}, \quad (5.10)$$

kde  $z_{1t}, z_{2t}, z_{3t}, z_{4t}$  sú tzv. „dummy variables“ (niekedy sa to prekladá ako „kvalitatívne premenné“), ktoré sú pre  $i = 1, 2, 3, 4$  definované ako

$$z_{it} = \begin{cases} 1 & \text{ak okamih } t \text{ zodpovedá } i\text{-temu kvartálu,} \\ 0 & \text{inak} \end{cases} \quad (5.11)$$

(t.j. napr. pre  $t = 1$  - čo predpokladáme, že zodpovedá prvému kvartálu - máme  $z_{1t} = 1$  a  $z_{2t} = z_{3t} = z_{4t} = 0$ ). Z toho potom triviálne vyplýva  $S_1 = \alpha_1, S_2 = \alpha_2, S_3 = \alpha_3, S_4 = \alpha_4$ . Keďže však predpokladáme aj  $S_1 + S_2 + S_3 + S_4 = 0$ , tak z  $\alpha_2, \alpha_3, \alpha_4$  vieme vyjadriť hodnotu  $\alpha_1 = -\alpha_2 - \alpha_3 - \alpha_4$  - model (5.9) je teda zbytočne preparametrizovaný a vieme ho zjednodušiť na

$$\begin{aligned} X_t &= \beta_0 + \beta_1 t + (-\alpha_2 - \alpha_3 - \alpha_4)z_{1t} + \alpha_2 z_{2t} + \alpha_3 z_{3t} + \alpha_4 z_{4t} + E_t = \\ &= \beta_0 + \beta_1 t + \alpha_2(z_{2t} - z_{1t}) + \alpha_3(z_{3t} - z_{1t}) + \alpha_4(z_{4t} - z_{1t}) + E_t. \end{aligned}$$

Maticovo toto vieme zapísať ako

$$\underbrace{\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ \vdots \\ X_n \end{pmatrix}}_{\mathbf{X}} = \underbrace{\begin{pmatrix} 1 & 1 & -1 & -1 & -1 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 \\ 1 & 4 & 0 & 0 & 1 \\ 1 & 5 & -1 & -1 & -1 \\ & & \vdots & & \\ 1 & n & 0 & 0 & 1 \end{pmatrix}}_F \begin{pmatrix} \beta_0 \\ \beta_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} + \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \\ \vdots \\ E_n \end{pmatrix}. \quad (5.12)$$

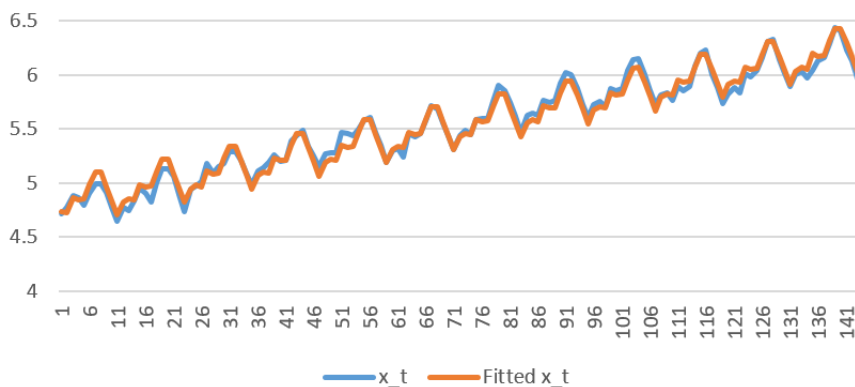
To je obyčajný lineárny model, čiže odhad vektora parametrov  $(\beta_0, \beta_1, \alpha_2, \alpha_3, \alpha_4)^T$  metódou najmenších štvorcov nájdeme ako  $(F^T F)^{-1} F^T \mathbf{x}$  ( $\mathbf{x}$  je realizácia náhodného vektora  $\mathbf{X}$ , ľudskou rečou naše dáta). Z tohto potom vieme dopočítať aj odhad pre  $\alpha_1$  ( $\hat{\alpha}_1 = -\hat{\alpha}_2 - \hat{\alpha}_3 - \hat{\alpha}_4$ ).

Týmto spôsobom sme naraz dostali odhad trendu aj sezonality. Okrem toho sa - za predpokladu, že  $E_1, E_2, \dots$  sú nezávislé s rozdelením  $N(0, \sigma^2)$  - dajú v takomto modeli konštruovať aj predikčné intervaly (to ako konštruovať predikčné intervaly vo všeobecnom lineárnom modeli budete mať zrejme na predmete Matematická štatistika alebo Ekonometria, takže sa tomu nebudeme hlbšie venovať).

## Holt-Wintersova metóda<sup>2</sup>

Táto metóda je zovšeobecnením dvojitého exponenciálneho vyrovnávania - opäť predpokladáme, že  $Tr_t \approx \beta_0 + \beta_1 t$  ( $\beta_0$  a  $\beta_1$  sa môžu v čase trochu meniť) a tentoraz pridávame aj sezónne koeficienty, ktoré sa tiež môžu v čase vyvíjať, t.j. platí iba  $S_t \approx S_{t+p}$  (predtým sme mali rovnosť). Korektné odvodenie by bolo nad rámec tohto kurzu, takže uvedieme iba „kuchársky prístup“, t.j. „recept“ ako v čase  $t$  odhadovať  $Tr_t, S_t$  a ako konštruovať predpovede a predikčné chyby. Metóda má tri parametre -  $\alpha, \beta, \gamma \in [0, 1]$  - tie hľadáme rovnako ako pri exponenciálnom vyrovnávaní, t.j. minimalizovaním sumy štvorcov predikčných chýb. Pre jednoduchosť opäť predpokladajme, že dĺžka periódy (sezóny) je  $p = 4$  (zovšeobecnenie na iné  $p$  je triviálne). Postup je nasledovný:

<sup>2</sup>V niektorých zdrojoch iba Wintersova metóda.



Obr. 26: Príklad regresného prístupu aplikovaný na počet (v tisíckach) letecky prepravených pasažierov (zlogaritmované mesačné dáta).

1. Zvolíme štartovacie parametre  $a_0, b_0, \hat{S}_{-3}, \hat{S}_{-2}, \hat{S}_{-1}, \hat{S}_0$  - tie získame použitím „regresného prístupu“ z predchádzajúcej podkapitoly na prvých  $2p$  (pri kvartálnych dátach 8) dát<sup>3</sup> ( $a_0$  zodpovedá odhadu  $\beta_0$  z „regresného prístupu“,  $b_0$  zodpovedá  $\beta_1$ ,  $\hat{S}_{-3}$  zodpovedá  $\alpha_1, \dots, \hat{S}_0$  zodpovedá  $\alpha_4$ ). Ak by sme mali napríklad mesačné dáta, tak potrebujeme  $a_0, b_0, S_0, \dots, S_{-11}$ .
2. Pre  $t = 1, 2, \dots, n$  prepočítame parametre nasledovne<sup>4</sup>

$$\begin{aligned} a_t &= (1 - \alpha)(x_t - \hat{S}_{t-4}) + \alpha(a_{t-1} + b_{t-1}), \\ b_t &= (1 - \beta)(a_t - a_{t-1}) + \beta b_{t-1}, \\ \hat{S}_t &= (1 - \gamma)(x_t - a_t) + \gamma \hat{S}_{t-4}. \end{aligned}$$

3. Odhad trendu v čase  $t$  je  $\widehat{Tr}_t = a_t$ , „fitovanú hodnotu“ z toho získame ako  $\hat{x}_t = \widehat{Tr}_t + \hat{S}_t$ .
4. Ak berieme do úvahy všetkých  $n$  meraní, tak predikciu v čase  $n + h, h > 0$  získame ako

$$\hat{x}_{n+h}(n) = a_n + b_n h + \hat{S}_{n+h-4k}, \quad (5.13)$$

kde  $k$  najmenšie možné prirodzené číslo tak, aby platilo  $n + h - 4k \leq n$  (ľudskou rečou zoberiem posledný odhad sezónnej zložky, ktorý máme a ktorý zodpovedá časovému okamihu  $n + h$  (rovnakému kvartálu ako je v čase  $n + h$ ) - toto ešte asi stále nemusí pripomínať ľudskú reč, takže skúsme príklad - ak by sme mali štvrťročné dáta od prvého kvartálu roku 2001 ( $t = 1$ ) po štvrtý kvartál roku 2010 (t.j.  $t = n = 40$ ) a chceli by sme napr. odhad pre tretí kvartál roku 2012 (t.j.  $h = 7$ ), tak by sme zobrali  $\hat{x}_{40+7}(40) = a_{40} + 7 \cdot b_{40} + \hat{S}_{39}$  ( $39 = 47 - 2 \cdot 4$ ), lebo 39 je posledný odhad sezónnej zložky, ktorý máme a ktorý zodpovedá tretiemu kvartálu).

5. Chybu predikcie  $\varepsilon_t$  pre  $t = 1, 2, \dots, n$  vypočítame ako

$$\varepsilon_t = x_t - \hat{x}_t(t-1), \quad (5.14)$$

kde  $\hat{x}_t(t-1) = a_{t-1} + b_{t-1} + \hat{S}_{t-4}$  (to isté ako keby sme vo vzorci (5.13) zobrali  $n = t-1$  a  $h = 1$ , čiže tvárime sa, že máme k dispozícii iba body  $x_1, \dots, x_{t-1}$  a na základe nich odhadujeme  $x_t$ ). Hodnotu  $\hat{x}_t(t-1)$  možno považovať za trochu „férovejšiu“ fitovanú hodnotu, nakoľko  $\hat{x}_t$  „využíva“ aj hodnotu  $x_t$  (vid' prepočet parametrov a výpočet  $\hat{x}_t$ ) - štandardne sa pri tejto metóde teda do grafu zobrazuje pôvodný časový rad spoločne s časovým radom  $\hat{x}_t(t-1)$ .

<sup>3</sup>V predchádzajúcej verzii poznámok to bolo uvedené trochu inak (regresný prístup na všetky dáta) - ono je to v princípe jedno, na výsledné predikcie to nebude mať až taký vplyv. Aby to bolo ale v súlade s prednáškou, budeme používať tento postup.

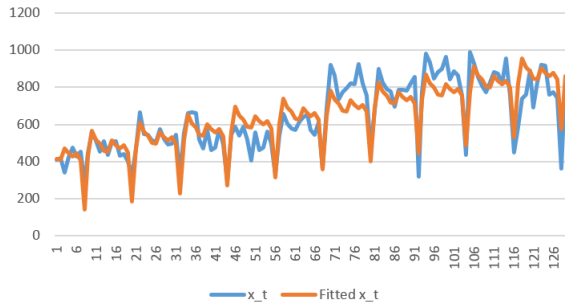
<sup>4</sup>Opäť som to trochu zmenil oproti pôvodnej verzii, aby to bolo v súlade s prednáškou - najpr ide  $(1 - \alpha)$ , až potom  $\alpha$ , podobne  $\beta, \gamma$ . Tu už je to z hľadiska čísel úplne jedno (všetky relevantné hodnoty výjdu rovnako), ale štandardnejšie značenie je toto (aktuálna verzia), preto táto drobná zmena.

6. Parametre  $\alpha, \beta, \gamma$  zvolíme tak, aby

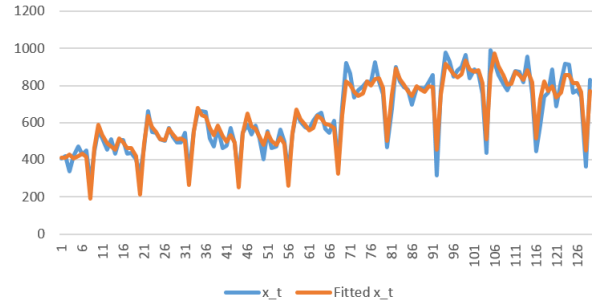
$$\sum_{t=1}^n \varepsilon_t^2 \quad (5.15)$$

bolo čo najmenšie.

Táto metóda patrí medzi adaptívne metódy, takže sa vie lepšie „vysporiadať“ so zmenami charakteru dát - vid' napr. Obr. 27 a Obr. 28, kde sme na tie isté dáta (modrou farbou) aplikovali regresný prístup a Holt-Wintersovu metódu (fitované hodnoty sú oranžovou).



Obr. 27: Regresný prístup.



Obr. 28: Holt Wintersova metóda.

Drobná poznámka k prepočtu parametrov  $a_t, b_t$  a  $\hat{S}_t$  - predstava je taká, že trend má v čase  $t-1$  tvar priamky prechádzajúcej bodom  $a_{t-1}$  so sklonom  $b_{t-1}$ . Očakávaná hodnota trendu v čase  $t$  by potom bola  $a_{t-1} + b_{t-1}$  (tak funguje priamka) - tú hodnotu potom „vyvážime“ hodnotou  $x_t - \hat{S}_{t-p}$ , čo je nejaký náš odhad skutočnej hodnoty trendu v čase  $t$ , lebo  $x_t$  sme „očistili“ sezónneho vplyvu (jeho odhadu) - z toho vzorec pre  $a_t$ . Podobne ak už máme  $a_t$ , tak „skutočný“ sklon trendu v čase  $t$  je  $a_t - a_{t-1}$  (tak funguje priamka) - ten opäť „vyvážime“ predchádzajúcim sklonom  $b_{t-1}$ , čím dostaneme hodnotu  $b_t$ . Nakoniec  $\hat{S}_t$  dostaneme opäť ako vážený priemer „očakávanej“ hodnoty, čiže  $\hat{S}_{t-p}$  (to by bolo vtedy, ak by sa so sezónnym vplyvom nič nedialo) a „skutočnej“ hodnoty  $x_t - a_t$  ( $x_t$  „očistené“ od trendu  $a_t$ ). Čiže funguje to ako jednoduché exponenciálne vyrovnávanie, kde sme tiež brali vážený priemer očakávanej a skutočnej hodnoty.

### Stručne k multiplikatívnym modelom

Doteraz sme predpokladali, že platí aditívny model

$$X_t = Tr_t + S_t + E_t. \quad (5.16)$$

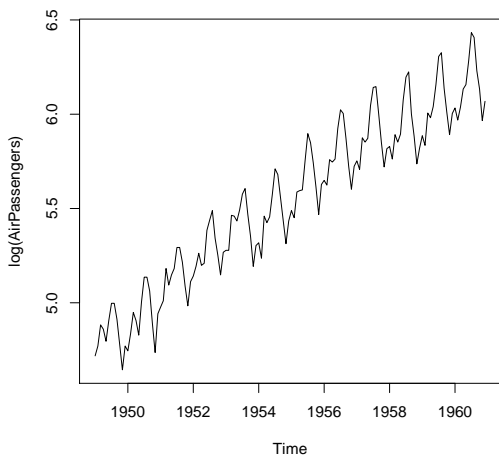
Ako sme ale aj v úvode spomínali, niekedy dáta lepšie popisuje takzvaný multiplikatívny model

$$X_t = Tr_t \cdot S_t \cdot E_t. \quad (5.17)$$

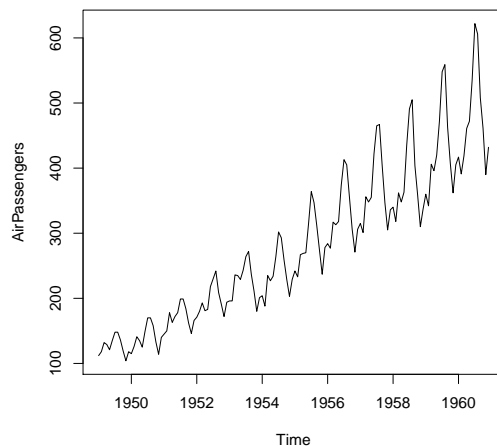
Dá sa to rozoznať napríklad tak, že „sezónne výkyvy“ nie sú približne konštantné, ale výraznejšie sa menia - porovnaj napr. Obr. 29 a Obr. 30 - v prvom z nich sú tieto odchýlky približne konštantné, v druhom zjavne narastajú. V takejto situácii je najjednoduchším (a často postačujúcim) riešením dáta zlogaritmovať - tým dostaneme opäť aditívny model (len s inými dátami)

$$\ln X_t = \ln Tr_t + \ln S_t + \ln E_t. \quad (5.18)$$

V prípade, že takúto operáciu spraviť nechceme (alebo nevieme kvôli záporným číslam), existujú multiplikatívne alternatívy k vyššie spomenutým metódam. Stručne spomenieme aspoň „jednoduchý prístup“ a Holt-Wintersovu metódu.



Obr. 29



Obr. 30

### Jednoduchý prístup v multiplikatívnom modeli

Oproti postupu v aditívnom modeli sú iba dva rozdiely:

1. Namiesto členov  $x_1 - \widehat{Tr}_1, x_{1+p} - \widehat{Tr}_{1+p} \dots$  vo výraze (5.6) berieme prirodzene namiesto rozdielov podiely, čiže  $\frac{x_1}{\widehat{Tr}_1}, \frac{x_{1+p}}{\widehat{Tr}_{1+p}} \dots \widetilde{S}_1$  by sme teda odhadovali ako

$$\widetilde{S}_1 = \frac{\frac{x_1}{\widehat{Tr}_1} + \frac{x_{1+p}}{\widehat{Tr}_{1+p}} + \dots + \frac{x_{1+kp}}{\widehat{Tr}_{1+kp}}}{k+1} \quad (5.19)$$

( $k$  je podľa počtu dát),  $\widetilde{S}_2, \dots, \widetilde{S}_p$  analogicky.

2. O  $S_1, \dots, S_p$  teraz predpokladáme, že sú v priemere 1, takže normovanie prebieha podľa predpisu

$$\widehat{S}_t = \frac{p\widetilde{S}_t}{\sum_{i=1}^p \widetilde{S}_i} \quad (5.20)$$

(môžete sa presvedčiť, že priemer bude naozaj 1).

### Multiplikatívna Holt-Wintersova metóda

V multiplikatívnej Holt-Wintersovej metóde postupujeme nasledovne.

1. Zvolíme štartovacie parametre  $a_0, b_0$  (napríklad tak, že nejakému odhadu trendu „nafitujeme“ metódou najmenších štvorcov priamku  $a_0 + b_0 t$ ) a  $\widehat{S}_{1-p}, \dots, \widehat{S}_0$  (napríklad použitím jednoduchého prístupu v multiplikatívnom modeli).
2. Parametre prepočítame podľa predpisov

$$\begin{aligned} a_t &= (1 - \alpha) \frac{x_t}{\widehat{S}_{t-p}} + \alpha(a_{t-1} + b_{t-1}) \\ b_t &= (1 - \beta)(a_t - a_{t-1}) + \beta b_{t-1} \\ \widehat{S}_t &= (1 - \gamma) \frac{x_t}{a_t} + \gamma \widehat{S}_{t-p}, \end{aligned}$$

3. Odhad trendu v čase  $t$  je  $\widehat{Tr}_t = a_t$ , „fitovanú hodnotu“ z toho získame ako  $\widehat{x}_t = \widehat{Tr}_t \widehat{S}_t$ .

4. Ak berieme do úvahy všetkých  $n$  meraní, tak predikciu v čase  $n + h, h > 0$  získame ako

$$\widehat{x}_{n+h}(n) = (a_n + b_n h) \widehat{S}_{n+h-pk}, \quad (5.21)$$

kde  $k$  je rovnaké ako v aditívnej verzii.

5. Chybu predikcie  $\varepsilon_t$  pre  $t = 1, 2, \dots, n$  vypočítame ako

$$\varepsilon_t = x_t - \widehat{x}_t(t-1), \quad (5.22)$$

kde  $\widehat{x}_t(t-1) = (a_{t-1} + b_{t-1}) \widehat{S}_{t-p}$ .

6. Parametre  $\alpha, \beta, \gamma$  zvolíme tak, aby

$$\sum_{t=1}^n \varepsilon_t^2 \quad (5.23)$$

bolo čo najmenšie.

## 6 Testy náhodnosti

Doteraz sme pracovali s časovými radmi, v ktorých bola zjavná nejaká systematická časť (trend a sezonalita), takže z  $(X_1, \dots, X_n)$  sme s istou mierou spoľahlivosti vedeli niečo povedať o nasledujúcich/predchádzajúcich hodnotách. Niekedy ale v dátach  $(x_1, \dots, x_n)$  nič systematické nevidíme a o tejto „absencii systematickosti“ by sme sa potrebovali presvedčiť nejakým objektívnym spôsobom. Jednou z možností je otestovať hypotézu o tom, že  $x_1, \dots, x_n$  sú realizácie nejakých nezávislých, rovnako rozdelených náhodných premenných  $X_1, \dots, X_n$ . Ak je táto hypotéza pravdivá, tak naozaj nemožno hovoriť o niečom systematickom (ak  $X_{n+1}$  nezávisí od  $X_n, \dots, X_1$ , tak vedomosť o realizáciách  $X_1, \dots, X_n$  nám nepovie nič zaujímavé o  $X_{n+1}$ ). V tejto časti si teda poviem o niektorých takýchto štatistických testoch.

Tie možno využiť aj v prípade, ak chceme skontrolovať, či už sme naozaj „odhalili“ celú systematickú časť časového radu. Napríklad ak sme v aditívnom modeli našli odhady  $\widehat{T}r_t$  a  $\widehat{S}_t$ , môžeme týmito testami odhaliť, či je v reziduálnej zložke  $E_t$  ešte niečo systematické (čím by sme potenciálne mohli vylepšiť naše predikcie), alebo sú to nejaké nezávislé (alebo aspoň nekorelované - o tom trochu neskôr) veličiny. V tomto prípade, ale treba tieto testy brať s miernou rezervou, lebo k dispozícii nemáme samotné realizácie  $E_t$ , ale iba ich odhady (tie prirodzene dostaneme ako  $x_t - \widehat{T}r_t - \widehat{S}_t$ ), ktoré už sú „zaťažené“ nejakou chybou (lebo odhady  $\widehat{T}r_t$  a  $\widehat{S}$  skoro iste nebudú úplne presné).

### Štatistické testovanie hypotéz

Predtým než uvedieme konkrétne testy, stručne pripomeňme, ako fungujú štatistické testy vo všeobecnosti - kto v tom má jasno, môže pokojne túto časť preskočiť.

Každý test má nejakú nulovú hypotézu (označujeme  $H_0$ ), ktorú testujeme a nejakú testovaciu štatistiku - to je náhodná premenná, ktorá má za platnosti  $H_0$  známe rozdelenie. Z dát, ktoré máme k dispozícii, vieme vypočítať jednu realizáciu tejto testovacej štatistiky - to je nejaké konkrétne číslo. Ak je toto číslo „nepravdepodobnou“<sup>5</sup> realizáciou rozdelenia testovacej štatistiky za platnosti  $H_0$ , tak hypotézu zamietneme.

Lepšie si to asi predstavíme na príklade - majme mincu, pričom chceme testovať

$$H_0 : \text{znak padá s rovnakou pravdepodobnosťou ako hlava, čiže s pravdepodobnosťou } \frac{1}{2},$$

pričom mincu môžeme hodiť desaťkrát. Rozumná testovacia štatistika by v tomto prípade bola napríklad náhodná premenná  $T$  daná predpisom  $T = \text{počet padnutých znakov z 10 hodov}$ .

<sup>5</sup>Toto je povedané trochu voľnejšie, pre hrubú predstavu to stačí.

Z PaŠ(2) vieme, že ak platí  $H_0$  (čiže znak padá s pravdepodobnosťou  $\frac{1}{2}$ ), tak táto náhodná premenná má binomické rozdelenie s parametrami 10 a  $\frac{1}{2}$  ( $T \sim Bin(10, \frac{1}{2})$ ). Nepravdepodobné realizácie takejto náhodnej premennej sú 0,1,9,10 - pre takéto hodnoty totiž platí

$$P(T = 0) + P(T = 1) + P(T = 9) + P(T = 10) \approx 0.02148.$$

Ak by nám teda z 10 hodov padlo napríklad 9 znakov, tak už veľmi nebudeme veriť, že je to realizácia binomického rozdelenia s parametrami 10 a  $\frac{1}{2}$ . Inými slovami, nebudeme veriť, že  $H_0$  platí, takže ju zamietneme. Môžeme sa, samozrejme, pomýliť - aj pri vyváženej minci sa nám môže stať, že nám padne deväťkrát znak - je to ale nepravdepodobné. Množina „nepravdepodobných hodnôt“ (kedy už hypotézu zamietame - v našom prípade  $\{0, 1, 9, 10\}$ ) sa nazýva kritická oblasť alebo oblasť zamietnutia. Spravidla sa volí tak, že sú v nej okrajové body, pričom chceme, aby zároveň platilo

$$P(T \in \text{kritická oblasť}) \leq 0.05 = 5\%^6.$$

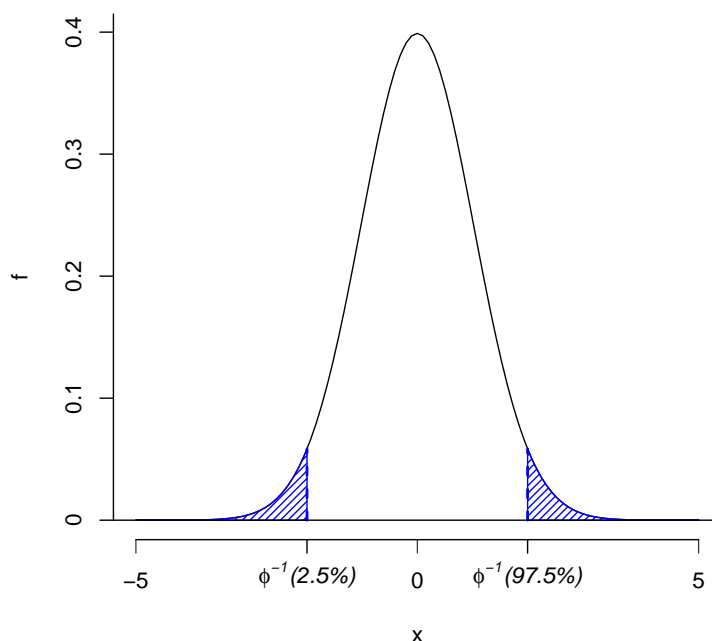
Ak je  $T$  spojitá náhodná premenná, tak vieme dosiahnuť presne 5%, pri diskrétnych to často nejde - napríklad ak by sme v našom príklade pridali do kritickej oblasti 2 alebo 8, tak už dostaneme

$$P(T \in \{0, 1, 2, 9, 10\}) \approx 0.0654.$$

Vo veľa testoch (aj v tých, ktoré budú v tejto kapitole) má testovacia štatistika  $T$  za platnosti  $H_0$  asymptoticky (pre veľké  $n$ ) štandardizované normálne rozdelenie  $N(0, 1)$ . V takom prípade má kritická oblasť tvar

$$\mathbb{R} \setminus (\phi^{-1}(2.5\%), \phi^{-1}(97.5\%))$$

(viď Obr. 31), kde  $\phi^{-1}$  je kvantilová funkcia (inverzia distribučnej funkcie) štandardizovaného normálneho rozdelenia. V Exceli vieme  $\phi^{-1}(x)$  vypočítať pomocou funkcie NORM.S.INV(x). Na-



Obr. 31: Kritická oblasť v prípade, že testovacia štatistika má  $N(0, 1)$  rozdelenie. Každá z modrých oblastí má plochu 2.5%, dokopy, teda 5%.

<sup>6</sup>5% je nejaký zaužívaný štandard, niekedy sa používa aj 1% alebo 10%, prípadne niečo iné. V tomto kurze si ale vystačíme s hodnotou 5%.



miesto kvantilov  $\phi^{-1}(2.5\%)$  a  $\phi^{-1}(97.5\%)$  sa občas píše takzvané kritické hodnoty  $u(97.5\%)$  a  $u(2.5\%)$  (tak to bolo zrejme aj na PaŠ(2)) - je to viac menej to isté, len symetricky otočené, t.j. platí vzťah  $\phi^{-1}(x) = u(100\% - x)$ . Z obrázku tiež môžeme vidieť, že hustota  $N(0, 1)$  je symetrická okolo nuly, takže v praxi by sme hypotézu zamietli v prípade, že  $|T| > \phi^{-1}(97.5\%) \approx 1.96$  (nezaškodí si túto hodnotu zapamätať, človek ju nemusí potom stále hľadať).

## Testy nezávislosti

V tejto časti teda spomenieme niektoré testy hypotézy

$H_0 : x_1, \dots, x_n$  sú realizáciou nezávislých, rovnako rozdelených náhodných premenných.

V praxi nezaškodí vyskúšať týchto testov viac, pretože testy majú rôznu silu (môže sa stať, že v niektorom teste hypotézu nezamietneme aj vtedy, keď neplatí). Ak aspoň v jednom z testov hypotézu zamietneme, je dosť možné, že je nepravdivá a nemali by sme s ňou rátať.

1. **Test založený na znamienkách diferencii.** Testovacia štatistika tohto testu je

$$T = \frac{k - \frac{n-1}{2}}{\sqrt{\frac{n+1}{12}}},$$

kde  $k$  je počet bodov rastu - takých bodov  $x_i$ , že  $x_{i+1} > x_i$ . Formálne vieme  $k$  zapísať ako  $k = \sum_{i=1}^{n-1} V_i$ , kde  $V_i = 1$  ak  $x_i < x_{i+1}$ , inak je  $V_i = 0$ . V prípade, že je nejaká dvojica susedných bodov rovnaká, t.j.  $x_i = x_{i+1}$ , tak jeden z nich treba v tomto teste vyškrtnúť (a potom pracovať aj so zníženým  $n$ ). Testovacia štatistika má za platnosti  $H_0$  asymptoticky normálne rozdelenie, takže (v duchu predchádzajúcej časti)  $H_0$  zamietame ak  $|T| > 1.96$ .

Idea tohto testu je taká, že ak sú  $X_1, \dots, X_n$  nezávislé a rovnako rozdelené (čiže ak platí  $H_0$ ), tak v realizáciach  $x_1, \dots, x_n$  bude počet bodov rastu tvoriť približne polovicu bodov, t.j.  $\frac{n-1}{2}$  (o  $x_n$  nevieme povedať, či je to bod rastu, alebo nie - preto je v čitateli  $n - 1$ ).  $H_0$  teda zamietneme vtedy, ak je počet bodov rastu „výrazne väčší“ (resp. „výrazne menší“) ako očakávaná hodnota  $\frac{n-1}{2}$  (inými slovami - ak sú v dátach nejaké rastúce, resp. klesajúce tendencie). To, čo je „výrazne väčší“ (resp. „výrazne menší“) počet nám nejakým štatisticky presným spôsobom kvantifikuje práve testovacia štatistika  $T$  (napríklad ak je počet bodov rastu  $k$  „výrazne väčší“ ako očakávaná hodnota  $\frac{n-1}{2}$ , tak  $k - \frac{n-1}{2}$  bude tiež relatívne veľké číslo - po nejakom znormovaní hodnotou  $\sqrt{\frac{n+1}{12}}$  väčšie ako 1.96).

2. **Test založený na bodoch zvratu.** Testovacia štatistika tohto testu je

$$T = \frac{r - \frac{2(n-2)}{3}}{\sqrt{\frac{16n-29}{90}}},$$

kde  $r$  je počet bodov zvratu - takých bodov  $x_i$ , že  $x_{i-1} < x_i > x_{i+1}$  alebo  $x_{i-1} > x_i < x_{i+1}$ . Opäť platí, že ak je nejaká dvojica susedných bodov rovnaká ( $x_i = x_{i+1}$  pre nejaké  $i$ ), tak jeden z bodov treba pred vykonaním tohto testu vyškrtnúť. Rovnako ako v predchádzajúcom teste má  $T$  za platnosti  $H_0$  asymptoticky štandardizované normálne rozdelenie.

Idea je analogická ako v predchádzajúcom teste - dá sa ukázať, že v postupnosti nezávislých náhodných premenných  $X_1, X_2, \dots, X_n$  je očakávaný počet bodov zvratu  $\frac{2(n-2)}{3}$ .  $H_0$  teda zamietneme, ak sa skutočný počet bodov zvratu  $r$  príliš odlišuje od tohto očakávaného počtu.

3. **Test založený na Kendallovom  $\tau^7$ .** Testovacia štatistika tohto testu je

$$T = \frac{\tau}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}},$$

<sup>7</sup>S tým sa pravdepodobne ešte stretnete v iných súvislostiach

kde  $\tau = \frac{4v}{n(n-1)} - 1$  a  $v$  je počet takých dvojíc  $x_i$  a  $x_j$ , že  $x_i < x_j$  (ide sa cez všetky možné  $i$  a  $j$ ), kde  $i < j$ . Za platnosti  $H_0$  má testovacia štatistika  $T$  opäť asymptoticky štandardizované normálne rozdelenie. Idea je opäť analogická ako v predchádzajúcich dvoch testoch - pri nezávislých, rovnako rozdelených náhodných premenných by sa hodnota  $v$  pohybovala okolo hodnoty  $\frac{n(n-1)}{4}$  (skúste si to premyslieť) a ak bude  $v$  príliš ďaleko od tejto očakávanej hodnoty, tak  $H_0$  zamietneme (avšak môžeme si všimnúť, že teraz sa to čo je „ďaleko“ nemeria rozdielom, ale podielom, od ktorého potom odčítame jednotku).

Môžeme si všimnúť, že vo všetkých troch prípadoch používame asymptotické rozdelenie, čiže potrebujeme veľké  $n$ . V týchto prípadoch sa odporúča  $n \geq 100$ . Ak toľko dát nemáme, tak sú známe aj presné rozdelenia (nie asymptotické), ktorých kvantily sú tabelizované (viď učebnica, kapitola 10). Druhá možnosť je aj tak použiť normálne rozdelenie a zmieriť sa s tým, že výsledky môžu byť trochu nepresné.

Testov nezávislosti je samozrejme viac, niektoré ďalšie možno nájsť v učebnici v kapitole 10.

## 7 Box-Jenkinsova metodológia - ARMA modely

Doteraz sme sa takmer výlučne sústredili na deterministickú (nenáhodnú) časť časového radu (trend a sezonalitu) a reziduálnej (náhodnej) zložke  $E_t$  sme nevenovali príliš veľkú pozornosť. V niektorých metódach (napríklad v exponenciálnom vyrovnávaní a Holt-Wintersovej metóde) ani nezohrávala príliš veľkú úlohu, v iných (napríklad pri konštrukcii predikčných intervalov) sme potrebovali predpoklad, že  $E_t$  sú nezávislé, normálne rozdelené, s nulovou strednou hodnotou a konštantnou disperziou. To je dosť silný predpoklad (v časových radoch väčšinou nereálny), avšak na to, aby metódy (aspoň asymptoticky) fungovali, spravidla stačí, aby bol  $E_t$  biely šum. Pripomeňme, že biely šum je postupnosť nekorelovaných náhodných premenných (t.j. ľubovoľné dve rôzne náhodné premenné, ktoré z tejto postupnosti vyberieme, majú nulovú koreláciu) s nulovou strednou hodnotou a konštantnou disperziou.

V praxi ale dosť často nie je splnený ani predpoklad o bielom šume, pretože v  $E_t$  môže byť „schovaná“ nejaká aspoň sčasti systematická štruktúra, ktorú ale nevieme identifikovať ani ako trend, ani ako sezónnu zložku. To má zopár dôsledkov:

1. Predikcie, ktoré sme doteraz používali ( $\hat{X}_T = \hat{T}r_T + \hat{S}_T$  pre nejaké „budúce“  $T > n$ ) sa potenciálne dajú ešte vylepšiť. Ak by sme napríklad vedeli, že pre ľubovoľné  $t \in \mathbb{Z}$  platí  $E_t = 0.8E_{t-1} + 0.4E_{t-2} + W_t$ , kde  $W_t$  je biely šum, tak napríklad v  $n + 1$  by sme namiesto odhadu  $\hat{X}_{n+1} = \hat{T}r_{n+1} + \hat{S}_{n+1}$  mohli zobrať

$$\hat{X}_{n+1} = \hat{T}r_{n+1} + \hat{S}_{n+1} + \hat{E}_{n+1} = \hat{T}r_{n+1} + \hat{S}_{n+1} + 0.8\hat{E}_n + 0.4\hat{E}_{n-1}.$$

2. Niektoré odhady (napríklad MNS odhady polynomickeho trendu) môžu byť vychýlené (ak by bol  $E_t$  biely šum, tak MNS odhady polynomickeho trendu by boli nevychýlené a v istom zmysle najlepšie možné), takže môže byť vhodnejšie použiť iný postup (časom si ukážeme aký).

Box-Jenkinsova metodológia, ktorej sa v tejto časti semestra budeme venovať, sa zameriava na to, ako túto systematickú štruktúru v náhodnej zložke vhodne modelovať pomocou takzvaných ARMA modelov. Zároveň na (väčšiu) chvíľu zabudneme na trend a sezonalitu - na časový rad

$$(\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots) \tag{7.1}$$

sa budeme pozerať čisto ako na postupnosť náhodných premenných (tak sme v konečnom dôsledku časový rad definovali), ktoré spravidla nemajú nejakú konkrétnu deterministickú časť. Výraz (7.1) budeme zapisovať skrátene ako  $\{X_t\}_{t \in \mathbb{Z}}$ .

ARMA modely (autoregressive-moving-average model) majú veľmi jednoduchý princíp (a tvar) - zoberie sa biely šum  $\{W_t\}_{t \in \mathbb{Z}}$ , ktorý predstavuje nejaké náhodné impulzy v daných časových

okamihoch (v literatúre sa namiesto výrazu „impulzy” používa aj „šoky” alebo „inovácie”) a hodnotu časového radu v čase  $t$  potom modelujeme ako lineárnu kombináciu predchádzajúcich hodnôt časového radu a náhodných impulzov - napríklad niečo typu

$$X_t = 2 + 0.3X_{t-1} + 0.4X_{t-2} - 0.4W_{t-1} + W_t \quad (7.2)$$

- v takomto príklade by hodnota časového radu v čase  $t$  závisela od predchádzajúcich dvoch hodnôt ( $X_{t-1}, X_{t-2}$ ) a náhodného impulzu v danom roku ( $W_t$ ) a predchádzajúcom roku ( $W_{t-1}$ ) - toto je takzvaný ARMA(2,1) model (2 a 1, lebo „najstaršie” členy sú  $X_{t-2}$  a  $W_{t-1}$ ). Úplne všeobecný ARMA( $p, q$ ) model má tvar

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}, \quad (7.3)$$

kde  $p, q \in \mathbb{N}$  a  $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q \in \mathbb{R}$  (a ešte nejaké iné podmienky - o všetkom ešte budeme hovoriť podrobnejšie, teraz len stručne pre predstavu). Modely takéhoto typu sú veľmi flexibilné (dobré sa dokážu prispôbiť aj na zložité tvary časového radu) a spravidla dávajú kvalitné predikcie. Nevýhodou býva horšia interpretovateľnosť - napríklad v modeli (7.2) sa dosť ťažko zdôvodňuje (a predstavuje), aký je na pravej strane význam 0.4-násobku náhodného impulzu z minulého roku.

Celý proces „výstavby” modelu sa dá zhrnúť do troch bodov.

1. Identifikácia modelu - v tomto bode sa na základe niektorých charakteristík snažíme určiť  $p$  a  $q$ , t.j. ako vôbec bude vyzeráť konkrétny model, ktorý pôjdeme dátam „napasovať”.
2. Odhad parametrov - keď už máme vybrané konkrétne  $p$  a  $q$ , v tomto bode odhadneme hodnoty neznámych parametrov  $c, \phi_1, \dots, \phi_p$  a  $\theta_1, \dots, \theta_q$  tak, aby model v čo najlepšej zhode s dátami (aby bol čo najlepší „fit”) - to je pomerne zložitá vec, bude to za nás robiť počítač (pričom ani pre počítač to nie je úplne najľahšie).
3. Kontrola (diagnostika) modelu - keď už máme odhadnuté parametre, tak spätne skontrolujeme, či je náš model v poriadku - či spĺňa predpoklady (napríklad ten, že  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum), prípadne či sa nedá zjednodušiť (použiť model s menším počtom parametrov) - na to sú štatistické testy, o ktorých si povieme. Ak model úplne v poriadku nie je, vraciame sa na bod 1 (a skúsime iné  $p$  a  $q$ ).

O všetkých troch bodoch, rovnako aj o základných vlastnostiach ARMA modelov si podrobnejšie povieme v tejto časti - na úvod ale najprv začneme s definíciami niektorých základných pojmov v časových radoch (úplne všeobecných) a ich vlastnosťami.

## Základné pojmy a ich vlastnosti

**Definícia.** Stredná hodnota časového radu  $\{X_t\}_{t \in \mathbb{Z}}$  je funkcia  $\mu : \mathbb{Z} \rightarrow \mathbb{R}$  daná predpisom

$$\mu(t) = E(X_t). \quad (7.4)$$

Nameište  $\mu(t)$  budeme pre jednoduchosť písať iba  $\mu_t$  (ale treba za tým stále vidieť funkciu premennej  $t$ ).

**Definícia.** Funkcia  $\gamma : \mathbb{Z}^2 \rightarrow \mathbb{R}$  definovaná predpisom

$$\gamma(s, t) = Cov(X_s, X_t) = E[(X_s - E(X_s))(X_t - E(X_t))] = E(X_s X_t) - E(X_s)E(X_t) \quad (7.5)$$

sa nazýva autokovariančná (prípadne iba kovariančná) funkcia časového radu  $\{X_t\}_{t \in \mathbb{Z}}$ .

Pripomeňme z PaŠ(1) niektoré vlastnosti kovariancie - pre ľubovoľné náhodné premenné  $X, Y, Z$  platí

1.  $Cov(X, X) = D(X)$  (disperzia),

2.  $Cov(X, Y) = Cov(Y, X)$ ,
3.  $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$
4.  $Cov(aX + b, cY + d) = acCov(X, Y)$  pre ľubovoľné  $a, b, c, d \in \mathbb{R}$ ,
5.  $|Cov(X, Y)| \leq \sqrt{D(X)D(Y)}$  (Cauchy-Schwarzova nerovnosť pre náhodné premenné).

Z tohto vieme ľahko odvodiť nasledujúce vlastnosti autokovariančnej funkcie:

1.  $\gamma(s, t) = \gamma(t, s)$  pre ľubovoľné  $s, t \in \mathbb{Z}$ ,
2.  $\gamma(t, t) = D(X_t) \geq 0$ ,
3.  $|\gamma(s, t)| \leq \sqrt{\gamma(s, s)\gamma(t, t)}$  pre ľubovoľné  $s, t \in \mathbb{Z}$ .

**Definícia.** Autokorelačná funkcia časového radu  $\{X_t\}_{t \in \mathbb{Z}}$  je funkcia  $\rho : \mathbb{Z}^2 \rightarrow [-1, 1]$  daná predpisom

$$\rho(s, t) = \rho_{X_s, X_t} = \frac{Cov(X_s, X_t)}{\sqrt{D(X_s)D(X_t)}}, \quad (7.6)$$

čiže je to korelácia medzi náhodnými premennými  $X_s$  a  $X_t$ . Pre autokorelačnú funkciu sa veľmi často používa skratka ACF (autocorrelation function).

Zjavne platí

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} \quad (7.7)$$

(priamo z definície) a to, že ACF nadobúda hodnoty iba z intervalu  $[-1, 1]$  vyplýva z vlastnosti 3 autokovariančnej funkcie.

Pripomeňme, že korelácia je akási miera *lineárnej* závislosti medzi náhodnými premennými - zjednodušene si to môžeme predstaviť tak, že ak je korelácia medzi  $X_s$  a  $X_t$  kladná, tak čím väčšie  $X_s$ , tým *pravdepodobne* väčšie  $X_t$  a naopak, ak je záporná, tak čím väčšie  $X_s$ , tým *pravdepodobne* menšie  $X_t$ . Ak je korelácia nulová, tak niečo takéto povedať nevieme. Ak by korelácia bola presne 1 alebo  $-1$ , tak so stopercentnou pravdepodobnosťou sú  $X_s$  a  $X_t$  lineárne závislé, čiže s pravdepodobnosťou 1 platí  $X_t = aX_s + b$  pre nejaké reálne čísla  $a, b$ , kde  $a > 0$  ak  $\rho(s, t) = 1$  a  $a < 0$  ak  $\rho(s, t) = -1$ . Keďže v ARMA modeloch je  $X_t$  vyjadrené aj ako *lineárna* kombinácia predchádzajúcich hodnôt  $X_{t-1}, X_{t-2}, \dots$ , uvidíme, že ACF bude zohrávať dôležitú úlohu pri identifikovaní modelu.

**Definícia.** Časový rad  $\{X_t\}_{t \in \mathbb{Z}}$  je (slabo) stacionárny<sup>8</sup>, ak sú splnené nasledovné podmienky:

1. Stredná hodnota  $\mu_t$  nezávisí od  $t$ , t.j. môžeme písať  $\mu_t = \mu$ , čiže je to konštantná funkcia premennej  $t$ ,
2. pre ľubovoľné  $t, h \in \mathbb{Z}$  platí, že funkcia  $\gamma(t, t + h)$  závisí iba od  $h$  bez ohľadu na to, aké je  $t$  - čiže iba od toho, ako sú  $X_t$  a  $X_{t+h}$  časovo vzdialené, nie od konkrétnych okamihov  $t$  a  $t + h$ .

To, že časový rad je stacionárny zjednodušene znamená to, že jeho hodnoty nám časom „neuletia“, ale stabilne sa pohybujú okolo nejakej pevnej hodnoty  $\mu$  bez výraznejších výkyvov.

<sup>8</sup>Existuje aj pojem silnej stacionarity, avšak slabá stacionarita sa používa bežnejšie, takže slovíčko „slabo“ sa pri slabostacionárnych časových radoch spravidla vynecháva.

## Príklady časových radov<sup>9</sup> časových radov (stacionárnych aj nestacionárnych)

Nech  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum s disperziou  $\sigma^2$ . Uvažujme nasledovné časové rady a pozrime sa, či sú stacionárne, alebo nie.

a)  $X_t = W_t$  - platí

$$\begin{aligned} \mu_t &= E(X_t) = E(W_t) = 0 \text{ vlastnosť bieleho šumu,} \\ \gamma(t, t+h) &= Cov(X_t, X_{t+h}) = Cov(W_t, W_{t+h}) = \\ &= \begin{cases} 0 & \text{ak } t \neq t+h \text{ (lebo je to biely šum - postupnosť nekorelovaných} \\ & \text{náhodných premenných)} \Leftrightarrow h \neq 0, \\ \sigma^2 & \text{ak } h = 0. \end{cases} \end{aligned}$$

Vidíme, že  $\mu_t$  je konštantné a  $\gamma(t, t+h)$  závisí iba od  $h$ , takže tento časový rad je stacionárny.

b)  $X_t = a + bt + W_t$ , kde  $a, b \in \mathbb{R}$  (lineárny trend, len s inými písmenkami) - platí

$$\begin{aligned} \mu_t &= E(X_t) = E(a + bt + W_t) = a + bt + E(W_t) = a + bt, \\ \gamma(t, t+h) &= Cov(a + bt + W_t, a + b(t+h) + W_{t+h}) = Cov(W_t, W_{t+h}) = \begin{cases} 0 & \text{ak } h \neq 0, \\ \sigma^2 & \text{ak } h = 0. \end{cases} \end{aligned}$$

V tomto prípade síce  $\gamma(t, t+h)$  závisí iba od  $h$ , avšak stredná hodnota nie je konštantná, takže toto nie je stacionárny časový rad.

c)  $X_t = c + W_t + \theta W_{t-1}$ , kde  $c, \theta \in \mathbb{R}$  - platí

$$\begin{aligned} \mu_t &= E(X_t) = E(c + W_t + \theta W_{t-1}) = c + E(W_t) + \theta E(W_{t-1}) = c, \\ \gamma(t, t+h) &= Cov(c + W_t + \theta W_{t-1}, c + W_{t+h} + \theta W_{t+h-1}) = \\ &= Cov(W_t, W_{t+h}) + \theta Cov(W_t, W_{t+h-1}) + \theta Cov(W_{t-1}, W_{t+h}) + \theta^2 Cov(W_{t-1}, W_{t+h-1}) = \\ &= \begin{cases} 0 & \text{ak } |h| \geq 2, \\ \theta \sigma^2 & \text{ak } |h| = 1, \\ \sigma^2 + \theta^2 \sigma^2 & \text{ak } h = 0. \end{cases} \end{aligned}$$

Stredná hodnota je konštantná a  $\gamma(t, t+h)$  závisí iba od  $h$ , takže v tento časový rad je stacionárny.

d)  $X_t = W_1 + W_2 + \dots + W_t$  - máme

$$\begin{aligned} \mu_t &= E(W_1 + W_2 + \dots + W_t) = E(W_1) + E(W_2) + \dots + E(W_t) = 0 \\ \gamma(t, t+h) &= Cov(W_1 + \dots + W_t, W_1 + \dots + W_{t+h}) = \sum_{i=1}^t \sum_{j=1}^{t+h} Cov(W_i, W_j) = \\ &= |Cov(W_i, W_j) \neq 0 \text{ iba ak } i = j, \text{ lebo } \{W_t\}_{t \in \mathbb{Z}} \text{ je biely šum}| = \\ &= \sum_{i=1}^{\min(t, t+h)} Cov(W_i, W_i) = \begin{cases} t\sigma^2 & \text{ak } h \geq 0, \\ (t+h)\sigma^2 & \text{ak } h < 0. \end{cases} \end{aligned}$$

Vidíme, že  $\gamma(t, t+h)$  závisí aj od  $t$ , takže tento časový rad nie je stacionárny.

V praxi vieme veľkú väčšinu dát - po nejakých rozumných transformáciach - modelovať ako stacionárne časové rady - budeme sa teda zaoberať iba nimi. To znie možno trochu prekvapivo, keďže

<sup>9</sup>Ak sme si pod časovými radmi doteraz predstavovali skôr dáta, v tejto časti budeme myslieť skôr teoretické modely, ktoré sa na dáta dajú potenciálne napasovať.

napríklad ani obyčajný lineárny trend (príklad b)) nie je stacionárny časový rad - dôležité je ale to „po nejakých rozumných transformáciach” - napríklad ak by sme v príklade b) (rovnako aj v príklade d)) namiesto  $X_t$  zobrali tazkvanú prvú diferenciu  $X_t - X_{t-1}$ , tak časový rad  $Y_t = X_t - X_{t-1}$  už stacionárny bude (môžete sa o tom presvedčiť, je to veľmi ľahké). Čiže aj v praxi, keď v nejakých dátach  $x_1, \dots, x_n$  uvidíme zjavný lineárny trend, tak stacionárnym modelom nebudeme modelovať priamo dáta  $x_1, \dots, x_n$ , ale ich diferencie  $x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1}$ . V stacionárnych časových radoch sa nám veľa vecí zjednoduší, konkrétne

- pre ľubovoľné  $t, h \in \mathbb{Z}$  závisí  $\gamma(t, t+h)$  iba od  $h$ , teda môžeme (a aj budeme) písať iba  $\gamma(h)$  (to nám vyjadruje, aká je kovariancia náhodných premenných vzdialených od seba o  $h$  časových úsekov),
- $D(X_t) = \gamma(t, t) = \gamma(t, t+0) = \gamma(0)$  pre ľubovoľné  $t$ , stacionárny časový rad má teda okrem konštantnej strednej hodnoty aj konštantnú disperziu,
- vlastnosti autokovariančnej funkcie môžeme zapísať ako
  1.  $\gamma(h) = \gamma(-h)$  pre ľubovoľné  $h \in \mathbb{Z}$  - zaujímať nás teda budú iba nezáporné hodnoty  $h$  - zvyšné si vieme dopočítať,
  2.  $\gamma(0) \geq 0$ ,
  3.  $|\gamma(h)| \leq \gamma(0)$  pre ľubovoľné  $h \in \mathbb{Z}$ ,
- aj pre ACF platí, že  $\rho(t, t+h)$  je iba funkciou  $h$ , pretože

$$\rho(t, t+h) = \frac{\gamma(t, t+h)}{\sqrt{\gamma(t, t)\gamma(t+h, t+h)}} = \frac{\gamma(h)}{\sqrt{\gamma(0)\gamma(0)}} = \frac{\gamma(h)}{\gamma(0)}$$

- rovnako teda namiesto  $\rho(t, t+h)$  budeme písať iba  $\rho(h)$ .

## Odhady v stacionárnych časových radoch

V praxi hodnoty konštantnej strednej hodnoty  $\mu$  a autokovariančnej funkcie  $\gamma(h)$  nepoznáme - potrebujeme ich odhadnúť z dát  $x_1, \dots, x_n$ . Na odhad strednej hodnoty sa prirodzene používa výberový priemer, t.j.

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (7.8)$$

Za odhad  $\gamma(h)$  sa berie

$$\hat{\gamma}(h) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}) & \text{pre } h \in \{0, \dots, (n-1)\}, \\ 0 & \text{pre } h \geq n, \end{cases} \quad (7.9)$$

čo je v podstate výberová kovariancia z  $(X_1, X_{1+h}), (X_2, X_{2+h}), \dots, (X_{n-h}, X_n)$  - je to niečo ako

$$E[(X_{t+h} - E(X_{t+h}))(X_t - E(X_t))]$$

(porovnaj so (7.5)), len namiesto stredných hodnôt máme priemery. Keďže najväčšia časová vzdialenosť, ktorú z dát máme k dispozícii je  $n-1$  (vzdialenosť medzi  $x_1$  a  $x_n$ ), pre väčšie hodnoty kladíme autokovariančnú funkciu 0. Pozorný čitateľ si zrejme všimol, že na to, aby to bol priemer, pred sumou by malo byť  $\frac{1}{n-h}$  namiesto  $\frac{1}{n}$  - to je, samozrejme, pravda, avšak ak by sme použili  $\frac{1}{n-h}$ , mohli by nám z matematického hľadiska výjsť aj nezmysly (záporný odhad disperzie), čo sa pri  $\frac{1}{n}$  nestane, preto sa berie  $\frac{1}{n}$  - z hľadiska presnosti odhadu to príliš veľký rozdiel nespraví. Za odhad autokorelačnej funkcie sa berie iba jednoducho

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}. \quad (7.10)$$

Ako už aj bolo spomínané vyššie, autokorelačná funkcia môže byť vhodným nástrojom na identifikáciu modelu - jednoducho si to môžeme ilustrovať na príklade c) (o stranu vyššie), kde sme uvažovali časový rad tvaru

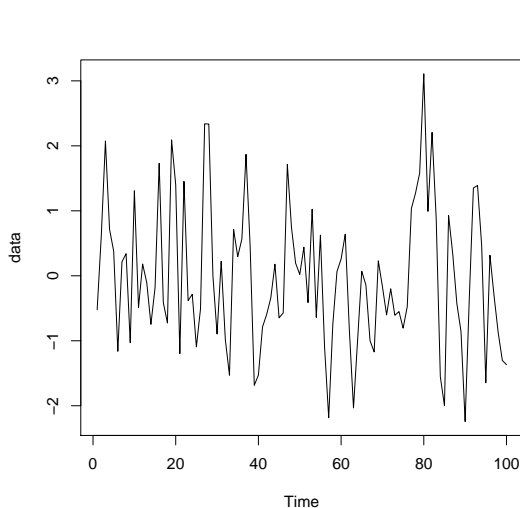
$$X_t = c + W_t + \theta W_{t-1},$$

čo je vlastne ARMA(0,1) model. Pre tento časový rad sme spočítali autokovariančnú funkciu, z ktorej ľahko vieme vyjadriť autokorelačnú funkciu

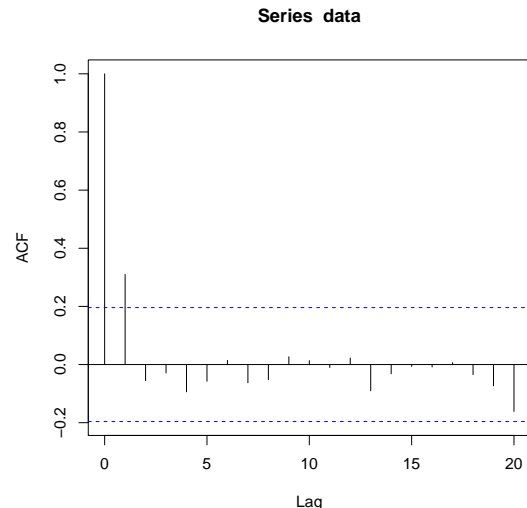
$$\rho(h) = \begin{cases} 0 & \text{ak } h \geq 2, \\ \frac{\theta\sigma^2}{\sigma^2 + \theta^2\sigma^2} = \frac{\theta}{1+\theta^2} & \text{ak } h = 1, \\ 1 & \text{ak } h = 0. \end{cases}$$

Takýto tvar autokorelačnej funkcie (nenulový práve pre  $h = 0$  a  $h = 1$  a inde nulový) je špecifický práve pre ARMA(0,1) model (žiadny iný ARMA( $p, q$ ) nemá autokorelačnú funkciu takéhoto tvaru) - ak by sme teda narazili na dáta, pre ktoré platí  $\hat{\rho}(0) = 1$  (to v skutočnosti triviálne platí vždy),  $\hat{\rho}(1) \neq 0$  (pričom pod „ $\neq 0$ “ teraz rozumieme, že to je od 0 dosť ďaleko) a  $\hat{\rho}(h) \approx 0$  pre  $h \geq 2$  (tým, že ide iba o *odhad* autokorelačnej funkcie, tak presnú nulu nedostaneme takmer nikdy, preto je tam iba približne), tak ARMA(0,1) model by mohol byť vhodný kandidát pre takéto dáta. Môže sa samozrejme stať, že netrafíme - tým, že z dát máme iba odhady  $\hat{\rho}$ , ktoré sú do istej miery nepresné, môže sa stať, že napríklad v skutočnosti platí  $\rho(2) > 0$ , ale mi to vďaka nepresnosti odhadu nesprávne vyhodnotíme ako nulové.

Na rozlíšenie toho, či je  $\hat{\rho}(h) \approx 0$ , alebo  $\hat{\rho}(h) \neq 0$  sa berie hraničná hodnota  $\pm \frac{2}{\sqrt{n}}$  - ak je  $\hat{\rho}(h)$  v týchto hraniciach, tak ho považujeme za blízke nule a ak je mimo týchto hraníc, tak ho berieme za rôzne od nuly. Netreba byť ale na tieto hranice prehnane upätý, skôr ich brať orientačne. Za tými hranicami sa v skutočnosti skrýva približný 95%-ný interval spoľahlivosti pre  $\hat{\rho}(h)$  v bielom šume,  $h = 1, 2, \dots$ . V bielom šume je skutočné  $\rho(h)$  rovné nule a jeho odhad  $\hat{\rho}(h)$  na 95% „padne“ približne do tohto intervalu. V štatistickom softvéri R (a aj v iných štatistických softvéroch) sa spolu s autokorelačnou funkciou automaticky vykresľuje aj táto hranica (viď modrá čiara na Obr. 33).



Obr. 32: Príklad realizácie ARMA(0,1) modelu.



Obr. 33: Autokorelačná funkcia realizácie ARMA(0,1) modelu.

Ďalším dôležitým nástrojom na identifikáciu modelu je v stacionárnych časových radoch takzvaná **parciálna autokorelačná funkcia** (skrátene ju budeme označovať PACF)  $\tau : \mathbb{N} \rightarrow [-1, 1]$ . Tá je daná predpisom

$$\tau(k) = \rho_{X_{t+k} - P_{t,k}(X_{t+k}), X_t - P_{t,k}(X_t)}, \quad (7.11)$$

kde pre ľubovoľnú náhodnú premennú  $Y$  je  $P_{t,k}(Y)$  najlepší lineárny odhad pomocou náhodných premenných  $X_{t+1}, \dots, X_{t+k-1}$ . Pod najlepším lineárnym odhadom sa myslí lineárna kombinácia

$\hat{Y} = \alpha_0 + \alpha_1 X_{t+1} + \dots + \alpha_{k-1} X_{t+k-1}$ , ktorá minimalizuje hodnotu  $E\left(\left(Y - \hat{Y}\right)^2\right)$ . Voľnejšie môžeme povedať, že  $\tau(k)$  je korelácia medzi náhodnými premennými vzdialenými o  $k$  časových úsekov (t.j.  $X_{t+k}$  a  $X_t$ ) „očistených“ od (lineárneho) vplyvu náhodných premenných, ktoré sú medzi nimi (t.j.  $X_{t+1}, \dots, X_{t+k-1}$ ). Hoci v zápise používame aj  $t$ , táto funkcia v stacionárnych časových radoch od  $t$  nezávisí (bola by rovnaká pre ľubovoľné  $t$ ). Platí, že ju vieme vypočítať pomocou ACF ako

$$\tau(k) = \frac{\det(P_k^*)}{\det(P_k)}, \quad (7.12)$$

kde

$$P_k = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(k-2) & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(k-1) & \rho(k-2) \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \dots & \rho(1) & 1 \end{pmatrix} \quad (7.13)$$

( $i,j$ -ty prvok tejto matice je  $\rho(|i-j|)$ , pretože  $\rho(0) = 1$ ) a  $P_k^*$  je taká istá ako  $P_k$ , len posledný stĺpec nahradíme vektorom  $(\rho(1), \rho(2), \dots, \rho(k))^T$ . Ak chceme PACF odhadnúť z dát, tak použijeme rovnaký vzorec, len v maticiach  $P_k$  a  $P_k^*$  namiesto  $\rho$  použijeme  $\hat{\rho}$ .

V ďalších častiach si povieme o niektorých špecifických ARMA modeloch a o všeobecnom ARMA modeli, pričom pri každom modeli si okrem iného povieme

- pre aké hodnoty parametrov je model stacionárny,
- aká je stredná hodnota daného modelu,
- ako vyzerá ACF,
- ako vyzerá PACF.

Na ukázanie niektorých týchto vecí ale ešte budeme potrebovať dva dôležité pojmy - operátor spätného posunu a lineárny proces.

**Operátor spätného posunu**  $B$  je operátor, ktorý časový rad posunie o jeden časový úsek dozadu, čiže  $BX_t = X_{t-1}$  pre ľubovoľné  $t$ . Vieme ho aplikovať aj viackrát - napríklad  $B^2X_t = BBX_t = BX_{t-1} = X_{t-2}$  a analogicky  $B^jX_t = X_{t-j}$  pre  $j = 0, 1, 2, \dots$ . Takisto vieme s týmto operátorom pracovať ako s polynómom, napríklad  $(B^2 + 4B + 4)X_t = X_{t-2} + 4X_{t-1} + 4X_t$ , pričom roznásobovanie funguje rovnako ako pri klasických polynómoch, čiže vyššie uvedený výraz vieme zapísať aj ako

$$\begin{aligned} (B+2)^2X_t &= (B+2)(B+2)X_t = (B+2)(X_{t-1} + 2X_t) = (B+2)X_{t-1} + (B+2)2X_t = \\ &= (X_{t-2} + 2X_{t-1}) + (2X_{t-1} + 4X_t) = X_{t-2} + 4X_{t-1} + 4X_t, \end{aligned}$$

čo je naozaj to isté ako  $(B^2 + 4B + 4)X_t$ .

**Lineárny proces** je časový rad tvaru

$$X_t = \mu + \psi_0 W_t + \psi_1 W_{t-1} + \psi_2 W_{t-2} + \dots = \mu + \sum_{i=0}^{\infty} \psi_i W_{t-i}, \quad (7.14)$$

kde  $\{W_t\}$  je biely šum s disperziou  $\sigma^2$  a  $\mu, \psi_0, \psi_1, \dots \in \mathbb{R}$  sú parametre. Dá sa ukázať, že postačujúcou podmienkou na stacionaritu takéhoto časového radu je

$$\psi_0^2 + \psi_1^2 + \psi_2^2 + \dots = \sum_{i=0}^{\infty} \psi_i^2 < \infty. \quad (7.15)$$

V takom prípade platí aj to, že lineárny proces je „slušný“ v zmysle, že pre ľubovoľné  $t$  je nulová pravdepodobnosť toho, že  $X_t$  bude  $\infty$  alebo  $-\infty$ . Prakticky sa takýto model nedá použiť úplne priamo (keďže má nekonečne veľa členov), avšak má teoretický význam, pretože:



1. Ak pre nejaký model ukážeme, že sa dá zapísať ako (7.14), kde  $\psi_0^2 + \psi_1^2 + \dots < \infty$ , tak z toho už vyplýva, že to je stacionárny model (a nemusíme overovať podmienky stacionarity, ktoré nemusia ísť vždy ľahko).
2. Dá sa ukázať, že úplne každý náhodný stacionárny časový rad<sup>10</sup> sa dá zapísať v tvare (7.14) - tomu sa hovorí Woldova reprezentácia stacionárneho časového radu. V praxi to znamená, že ak máme dáta, ktoré sú realizáciou stacionárneho časového radu (väčšinou rozumný predpoklad), tak model (7.14) s podmienkou (7.15) by bol teoreticky použiteľný model pre tieto dáta (len by sme museli nájsť jeho parametre), čo sa spravidla dá dobre aproximovať iba nejakým konečným počtom členov, čiže

$$X_t \approx \mu + \sum_{i=0}^k \psi_i W_{t-i}$$

pre nejaké  $k \in \mathbb{N}$ , čo je po predelení hodnotou  $\psi_0$  obyčajny ARMA(0,  $k$ ) model.

## Autoregresný proces

Časový rad  $\{X_t\}_{t \in \mathbb{Z}}$  sa nazýva autoregresný proces rádu  $p$ , ak platí

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t, \quad (7.16)$$

kde  $c, \phi_1, \dots, \phi_p \in \mathbb{R}$  sú parametre,  $\phi_p \neq 0$  a  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum s disperziou  $\sigma^2$ , pričom  $W_t$  a  $X_s$  sú nezávislé pre  $s < t$ . Môžeme si všimnúť, že to je špeciálny prípad ARMA( $p, q$ ) modelu (7.3), kde  $q = 0$  - skrátene teda takýto časový rad označujeme ako AR( $p$ ). Platí, že AR( $p$ ) je stacionárny práve vtedy, keď sú korene polynómu (aj komplexné)

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p \quad (7.17)$$

v absolútnej hodnote väčšie ako jedna (niekedy sa to formuluje tak, že korene sú v komplexnej rovine mimo jednotkového kruhu).

Polynóm (7.17) sa nazýva autoregresný polynóm a má úzky súvis s predpisom (7.16) - ten totiž môžeme pomocou operátora spätného posunu zapísať aj ako

$$\begin{aligned} X_t &= c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t \\ X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} &= c + W_t \\ X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \dots - \phi_p B^p X_t &= c + W_t \\ (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= c + W_t \\ \Phi(B) X_t &= c + W_t. \end{aligned}$$

Dokázať tvrdenie o stacionarite je pomerne zložité, ilustrujeme si to aspoň na AR(1) modeli s  $c = 0$ , čiže ak platí

$$X_t = \phi_1 X_{t-1} + W_t.$$

V takom prípade má autoregresný polynóm tvar  $\Phi(z) = 1 - \phi_1 z$ , čiže má iba jeden koreň, a to  $\frac{1}{\phi_1}$ . AR(1) je teda stacionárny  $\Leftrightarrow \left| \frac{1}{\phi_1} \right| > 1 \Leftrightarrow |\phi_1| < 1$ . O tom sa dá aj ľahko presvedčiť - ak by totiž platilo, že  $|\phi_1| \geq 1$ , tak

$$D(X_t) = D(\phi_1 X_{t-1} + W_t) = |\text{nezávislosť } X_{t-1} \text{ a } W_t| = \underbrace{\phi_1^2}_{\geq 1} D(X_{t-1}) + \underbrace{D(W_t)}_{> 0} > D(X_{t-1}),$$

<sup>10</sup> „Náhodný“ preto, lebo existujú totiž aj v istom zmysle „deterministické“ časové rady - napríklad ak by sme zobrali  $X_t = A \cos(t)$ , kde  $A$  je náhodná premenná, tak síce tam je prvok náhodnosti (kvôli  $A$ ), ale ak poznáme hodnotu v jednom časovom okamihu, tak  $A$  už si vieme vyjadriť, takže by sme poznali všetky hodnoty. To je ale skôr teoretická záležitosť, v praxi sa stretávame takmer výlučne s náhodnými časovými radmi.

čo sa ale v stacionárnom časovom rade nemôže stať, lebo tam sú všetky disperzie rovnaké. Naopak ak  $|\phi_1| < 1$ , tak  $X_t$  vieme zapísať ako

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + W_t = \phi_1 \overbrace{(\phi_1 X_{t-2} + W_{t-1})}^{X_{t-1}} + W_t = \phi_1^2 X_{t-2} + \phi_1 W_{t-1} + W_t = \\ &= \phi_1^2 \overbrace{(\phi_1 X_{t-3} + W_{t-2})}^{X_{t-2}} + \phi_1 W_{t-1} + W_t = \phi_1^3 X_{t-3} + \phi_1^2 W_{t-2} + \phi_1 W_{t-1} + W_t = \dots \\ &= \phi_1^k X_{t-k} + \phi_1^{k-1} W_{t-(k-1)} + \dots + \phi_1 W_{t-1} + W_t = \phi_1^k X_{t-k} + \sum_{j=0}^{k-1} \phi_1^j W_{t-j}. \end{aligned}$$

To platí pre ľubovoľné  $k$ , takže ak  $k \rightarrow \infty$ , tak  $\phi_1^k \rightarrow 0$  (lebo  $|\phi_1| < 1$ ), takže dostaneme

$$X_t = \sum_{j=0}^{\infty} \phi_1^j W_{t-j},$$

čo je obyčajný lineárny proces s  $\psi_j = \phi_1^j$  pre  $j = 0, 1, \dots$ , pričom

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{\infty} \phi_1^{2j} = |\text{geometrický rad s kvocientom } \phi_1^2 < 1| = \frac{1}{1 - \phi_1^2} < \infty,$$

čiže je to stacionárny časový rad.

Pre strednú hodnotu  $\mu_t$  v stacionárnom AR( $p$ ) platí

$$\begin{aligned} \mu_t = E(X_t) &= E(c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + W_t) = c + \phi_1 E(X_{t-1}) + \dots + \phi_p E(X_{t-p}) + E(W_t) = \\ &= c + \phi_1 \mu_{t-1} + \dots + \phi_p \mu_{t-p}. \end{aligned}$$

Keďže ale v stacionárnom časovom  $\mu_t$  nezávisí od  $t$ , tak  $\mu_t = \mu$  pre ľubovoľné  $t$  a teda ľahko vyjadríme, že

$$\mu = \frac{c}{1 - \phi_1 - \dots - \phi_p}.$$

Autokorelačnú funkciu  $\rho$  v stacionárnom AR( $p$ ) modeli počítame nasledovne:

- Pre  $h = 1, 2, \dots, p$  získame  $\rho(h)$  riešením takzvanej sústavy Yule-Wolkerových rovníc, ktorá je tvaru

$$\begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(p-2) & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(p-1) & \rho(p-2) \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \dots & \rho(1) & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{pmatrix},$$

čiže  $p$  rovníc o  $p$  neznámych (môžeme si všimnúť, že naľavo je matica  $P_p$ , vid' (7.13)). Pozor je tu trochu „zrada“, lebo neznáme sú tu  $\rho(1), \dots, \rho(p)$  a nie  $\phi_1, \dots, \phi_p$  (väčšinou sa sústavy zapisujú tak, že v matici a na pravej strane sú známe hodnoty, nie naopak).

- Pre  $h > p$  vypočítame  $\rho(h)$  rekurentne ako

$$\rho(h) = \phi_1 \rho(h-1) + \phi_2 \rho(h-2) + \dots + \phi_p \rho(h-p).$$

Ak by niekoho zaujímalo, ako sme sa k týmto predpisom dostali, tak si to ilustrujeme na AR(2) modeli s  $c = 0$ , čiže ak platí  $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$ . Keďže  $c = 0$ , tak aj pre strednú hodnotu platí  $\mu = E(X_t) = 0$ , takže pre autokovariančnú funkciu  $\gamma$  platí

$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)E(X_{t+h}) = E(X_t X_{t+h}).$$

Zoberme teraz predpis modelu

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t,$$

prenásobme obidve strany náhodnou premennou  $X_{t-1}$  a na obe strany aplikujme strednú hodnotu - dostaneme

$$\begin{aligned} E(X_t X_{t-1}) &= \phi_1 E(X_{t-1} X_{t-1}) + \phi_2 E(X_{t-2} X_{t-1}) + E(W_t X_{t-1}), \text{ čiže} \\ \gamma(1) &= \phi_1 \gamma(0) + \phi_2 \gamma(1), \end{aligned}$$

lebo  $W_t$  a  $X_{t-1}$  sú nezávislé, takže  $E(W_t X_{t-1}) = E(W_t)E(X_{t-1}) = 0$ . Ak to isté spravíme aj s  $X_{t-2}$ , dostaneme

$$\begin{aligned} E(X_t X_{t-2}) &= \phi_1 E(X_{t-1} X_{t-2}) + \phi_2 E(X_{t-2} X_{t-2}) + E(W_t X_{t-2}), \text{ čiže} \\ \gamma(2) &= \phi_1 \gamma(1) + \phi_2 \gamma(0). \end{aligned}$$

Ak teraz rovnice

$$\begin{aligned} \gamma(1) &= \phi_1 \gamma(0) + \phi_2 \gamma(1) \\ \gamma(2) &= \phi_1 \gamma(1) + \phi_2 \gamma(0) \end{aligned}$$

predelíme číslom  $\gamma(0)$ , dostaneme práve sústavu Yule-Wolkerových rovníc

$$\begin{aligned} \rho(1) &= \phi_1 + \phi_2 \rho(1) \\ \rho(2) &= \phi_1 \rho(1) + \phi_2. \end{aligned}$$

Analogickú „procedúru“ vieme spraviť aj s  $X_{t-h}$  pre ľubovoľné  $h > 2$ , čiže

$$\begin{aligned} E(X_t X_{t-h}) &= \phi_1 E(X_{t-1} X_{t-h}) + \phi_2 E(X_{t-2} X_{t-h}) + E(W_t X_{t-h}) \\ \gamma(h) &= \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) \quad / \frac{1}{\gamma(0)} \\ \rho(h) &= \phi_1 \rho(h-1) + \phi_2 \rho(h-2), \end{aligned}$$

čím dostaneme rekurentný predpis pre  $h > 2$ .

Pre ACF vo všeobecnosti platí, že je nenulová, ale k nule konverguje.

Funkciu PACF by sme počítali pomocou predpisu (7.12) - vo všeobecnosti nemá nejaké „pekné“ vyjadrenie, ale dá sa ukázať, že  $\tau(k) = 0$  pre každé  $k > p$ . Uvažujme napríklad  $k = p + 1$  - ak si spomenieme na „voľnejšiu“ definíciu PACF, tak  $\tau(p + 1)$  bola korelácia medzi náhodnými premennými vzdialenými o  $p + 1$  časových okamihov, napríklad  $X_t$  a  $X_{t-p+1}$  „očistenými“ od lineárneho vplyvu premenných medzi nimi, t.j.  $X_{t-1}, \dots, X_{t-p}$ . Keď ale  $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + W_t$  „očistíme“ od lineárneho vplyvu  $X_{t-1}, \dots, X_{t-p}$ , ostane nám iba  $W_t$ , čo je z definície AR( $p$ ) nezávislé od  $X_{t-1}, X_{t-2}, \dots$ , takže korelácia  $\tau(p+1)$  bude naozaj nulová. To je veľmi užitočné pri identifikácii modelu - ak sa totiž stretne s dátami, v ktorých odhad ACF je nenulový, ale k nule sa blíži a odhad PACF je od nejakého okamihu  $p + 1$  blízko nuly, tak AR( $p$ ) môže byť dobrý nápad. Napríklad na Obr. 35 ide ACF (horný graf) k nule dosť pomaly, avšak PACF (spodný graf) je od  $k = 2$  už pod hraničnou modrou čiarou (orientačné hranice sa berú rovnaké ako pri ACF). AR(1) model by teda mohol byť dobrý nápad pre takéto dáta.

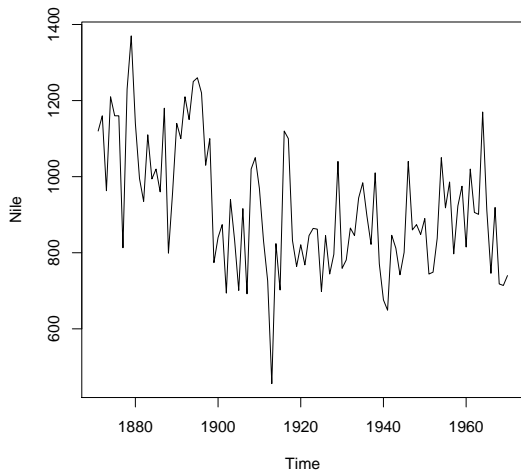
Predikcie v AR( $p$ ) modeli sú veľmi intuitívne - ak nájdeme odhady  $\hat{c}, \hat{\phi}_1, \dots, \hat{\phi}_p$  parametrov  $c, \phi_1, \dots, \phi_p$ , tak  $X_{n+1}$  odhadneme ako

$$\hat{X}_{n+1} = \hat{c} + \hat{\phi}_1 x_n + \dots + \hat{\phi}_p x_{n+1-p},$$

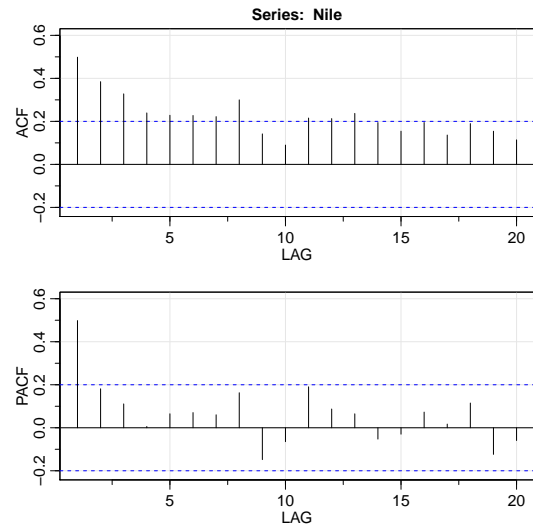
čiže iba použijeme model s odhadnutými parametrami, pričom  $W_{n+1}$  odhadneme jeho strednou hodnotou, čiže 0. Podobne  $X_{n+2}$  by sme odhadli ako

$$\hat{X}_{n+2} = \hat{c} + \hat{\phi}_1 \hat{X}_{n+1} + \hat{\phi}_2 x_n + \dots + \hat{\phi}_p x_{n+2-p},$$

analogicky aj nasledujúce hodnoty.



Obr. 34: Ročný prietok rieky Níl mestom Asuán v rokoch 1871-1970.



Obr. 35: Odhad funkcií ACF a PACF dát z Obr. 34.

## Moving average<sup>11</sup> proces

Časový rad  $\{X_t\}_{t \in \mathbb{Z}}$  sa nazýva moving average rádu  $q$ , ak platí

$$X_t = \mu + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}, \quad (7.18)$$

kde  $\mu, \theta_1, \dots, \theta_q \in \mathbb{R}$  sú parametre,  $\theta_q \neq 0$  a  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum s disperziou  $\sigma^2$ . Môžeme si všimnúť, že to je špeciálny prípad ARMA( $p, q$ ) modelu (7.3), kde  $p = 0$  - skrátene takýto časový rad onačujeme ako MA( $q$ ). Platí, že MA( $q$ ) je vždy stacionárny, lebo je to iba špeciálny prípad lineárneho procesu, v ktorom  $\psi_0 = 1, \psi_1 = \theta_1, \psi_2 = \theta_2, \dots, \psi_q = \theta_q$  a  $\psi_{q+1} = \psi_{q+2} = \dots = 0$ , čiže podmienka (7.15) triviálne platí.

Pre strednú hodnotu platí

$$E(X_t) = E(\mu + W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q}) = \mu + E(W_t) + \theta_1 E(W_{t-1}) + \dots + \theta_q E(W_{t-q}) = \mu.$$

Autokorelačnú funkciu  $\rho(h)$  pre  $h = 0, 1, 2, \dots$  vieme vypočítať ako  $\frac{\gamma(h)}{\gamma(0)}$ , kde

$$\begin{aligned} \gamma(h) &= Cov(X_t, X_{t+h}) = |\text{položme } \theta_0 \equiv 1| = Cov\left(\mu + \sum_{i=0}^q \theta_i W_{t-i}, \mu + \sum_{j=0}^q \theta_j W_{t+h-j}\right) = \\ &= \sum_{i=0}^q \sum_{j=0}^q \theta_i \theta_j Cov(W_{t-i}, W_{t+h-j}) = (\star). \end{aligned}$$

Keďže  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum, tak

$$Cov(W_{t-i}, W_{t+h-j}) = \begin{cases} 0 & \text{ak } t-i \neq t+h-j, \\ \sigma^2 & \text{ak } t-i = t+h-j \Leftrightarrow j = h+i, \end{cases}$$

čiže môžeme písať

$$(\star) = \begin{cases} \sum_{i=0}^{q-h} \theta_i \theta_{i+h} \sigma^2 & \text{pre } h \leq q, \\ 0 & \text{pre } h > q. \end{cases}$$

Keďže  $\gamma(h) = 0$  pre  $h > q$ , tak aj pre ACF platí, že  $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = 0$  pre  $h > q$ .

<sup>11</sup>Toto sa prekladá ako kĺzavý priemer, avšak nemá to nič spoločné s kĺzavým priemerom z tretej časti - aby sme si to nemýlili, tak toto necháme v angličtine.

PACF by sme počítali podľa (7.12) - opäť nemá „pekné“ vyjadrenie, avšak tentokrát pre ňu platí, že je nenulová, ale k nule konverguje. Analogicky ako pri AR modeloch, aj tu vieme ACF a PACF využiť na identifikáciu - ak narazíme na dáta, v ktorých je odhad  $\hat{\tau}$  nenulový, ale k nule sa blíži a ak pre nejaké  $q$  pre odhad ACF platí  $\hat{\rho}(h) \approx 0$  pre  $h > q$ , tak  $MA(q)$  môže byť vhodný kandidát na model pre tieto dáta (viď napr. Obr. 33).

Čo sa týka predikcií, tak v  $MA(q)$  modeloch je to trochu komplikovanejšie - uvažujme napríklad, že na dáta  $x_1, \dots, x_n$  pasuje nejaký  $MA(1)$  model s nulovou strednou hodnotou, t.j.

$$X_t = W_t + \theta_1 W_{t-1} \quad (7.19)$$

pre  $t \in \mathbb{Z}$ . Ak by sme chceli napríklad predikciu v čase  $n+1$ , tak pre ňu platí

$$X_{n+1} = W_{n+1} + \theta_1 W_n. \quad (7.20)$$

$W_{n+1}$  je nejaký budúci náhodný impulz, o ktorom nemáme žiadnu informáciu - odhadujeme ho teda strednou hodnotou (teda nulou). Náhodný impulz  $W_n$  ale už nastal (bolo to nejaké konkrétne číslo), ale pre nás bol nepozorovateľný (toto číslo nepoznáme - poznáme iba  $x_1, \dots, x_n$ ). Z dát  $x_1, \dots, x_n$  sa ale za istých podmienok o tomto  $W_n$  vieme niečo dozvedieť. Ak by pre parameter  $\theta_1$  platilo  $|\theta_1| < 1$ , tak (7.19) vieme vyjadriť ako

$$\begin{aligned} W_t &= X_t - \theta_1 W_{t-1} = X_t - \theta_1 \overbrace{(X_{t-1} - \theta_1 W_{t-2})}^{W_{t-1}} = X_t - \theta_1 X_{t-1} + \theta_1^2 W_{t-2} = \\ &= X_t - \theta_1 X_{t-1} + \theta_1^2 \overbrace{(X_{t-2} - \theta_1 W_{t-3})}^{W_{t-2}} = X_t - \theta_1 X_{t-1} + \theta_1^2 X_{t-2} - \theta_1^3 W_{t-3} = \dots \\ &= X_t + \sum_{j=1}^{\infty} (-\theta_1)^j X_{t-j} \quad (\text{analogicky ako pri AR(1)}). \end{aligned}$$

Po prehodení sumy na druhú stranu máme

$$X_t = - \sum_{j=1}^{\infty} (-\theta_1)^j X_{t-j} + W_t, \quad (7.21)$$

čo je akýsi  $AR(\infty)$  model s parametrami  $\phi_1 = \theta_1, \phi_2 = -\theta_1^2, \phi_3 = \theta_1^3, \dots$ .  $X_{n+1}$  by sme teda vedeli odhadnúť ako

$$\hat{X}_{n+1} = \theta_1 x_n - \theta_1^2 x_{n-1} + \theta_1^3 x_{n-2} - \dots, \quad (7.22)$$

avšak  $x_0, x_{-1}, x_{-2}, \dots$  neponáme, takže sa používa aproximácia

$$\hat{X}_{n+1} \approx \theta_1 x_n - \theta_1^2 x_{n-1} + \theta_1^3 x_{n-2} - \dots + (-1)^{n+1} \theta_1^n x_1. \quad (7.23)$$

Keďže  $|\theta_1| < 1$ , tak čím väčšie  $n$  (alebo čím menšie  $|\theta_1|$ ), tým je táto aproximácia presnejšia (veľkosť členov geometricky klesá s kvocientom  $-\theta_1$ ). Takýto prepis na  $AR(\infty)$  model vieme za istých podmienok spraviť aj pre vyššie rády  $q$ . Ak niečo také vieme, tak hovoríme, že časový rad  $MA(q)$  je invertovateľný.

Platí, že  $MA(q)$  je invertovateľný práve vtedy, keď korene polynómu (aj komplexné)

$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q \quad (7.24)$$

sú v absolútnej hodnote väčšie ako 1. Polynóm  $\Theta(z)$  sa nazýva moving average polynóm a opäť úzko súvisí s  $MA$  modelom (7.18), pretože ho pomocou operátora spätného posunu vieme zapísať ako

$$\begin{aligned} X_t &= \mu + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q} = \mu + W_t + \theta_1 B W_t + \theta_2 B^2 W_t + \dots + \theta_q B^q W_t = \\ &= \mu + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) W_t = \mu + \Theta(B) W_t. \end{aligned}$$

Ak máme teda invertovateľný  $MA(q)$  časový rad (invertovateľnosť spravidla vyžadujeme), tak pri predikciách postupujeme analogicky ako pri  $MA(1)$  - časový rad  $MA(q)$  prepíšeme na  $AR(\infty)$ , zanedbáme posledné členy (obsahujúce  $X_0, X_{-1}, X_{-2}, \dots$ ) a použijeme predikcie ako pri autoregresnom procese.

## Autoregresný-moving average proces

Časový rad  $\{X_t\}_{t \in \mathbb{Z}}$  sa nazýva autoregresný-moving average proces rádu  $(p, q)$ , ak platí

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}, \quad (7.25)$$

kde  $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q \in \mathbb{R}$  sú parametre,  $\phi_p \neq 0, \theta_q \neq 0$ ,  $\{W_t\}_{t \in \mathbb{Z}}$  je biely šum s disperziou  $\sigma^2$  a  $W_t$  a  $X_s$  sú nezávislé pre  $s < t$ . Skrátene ho označujeme ARMA( $p, q$ ). Rovnosť (7.25) vieme pomocou operátorov spätného posunu zapísať ako

$$\begin{aligned} X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} &= c + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}, \\ X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \dots - \phi_p B^p X_t &= c + W_t + \theta_1 B W_t + \theta_2 B^2 W_t + \dots + \theta_q B^q W_t \\ (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= c + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) W_t \\ \Phi(B) X_t &= c + \Theta(B) W_t, \end{aligned}$$

kde  $\Phi$  je autoregresný polynóm (rovnaký ako pri AR( $p$ )) a  $\Theta$  je moving average polynóm (rovnaký ako pri MA( $q$ )). Platí, že

1. model (7.25) je stacionárny práve vtedy, keď sú korene polynómu (aj komplexné)  $\Phi(z)$  v absolútnej hodnote väčšie ako jedna,
2. model (7.25) je invertovateľný (zapísateľný ako AR( $\infty$ )) práve vtedy, keď sú korene polynómu (aj komplexné)  $\Theta(z)$  v absolútnej hodnote väčšie ako jedna,

čiže rovnaké pravidlá ako v samostatnom AR, resp. MA modeli. Okrem stacionarity a invertovateľnosti od parametrov ARMA( $p, q$ ) modelu vyžadujeme, aby polynómy  $\Phi$  a  $\Theta$  nemali rovnaké korene - v takom prípade by totiž rovnosť (7.25) spĺňal aj nejaký ARMA model rádu  $(p-1, q-1)$  (čiže už by to nebol „čistokrvný“ ARMA( $p, q$ ) model). Ilustrujme si to na ARMA(1,1) modeli s  $c = 0$  - v takom prípade platí

$$X_t = \phi_1 X_{t-1} + W_t + \theta_1 W_{t-1}, \quad (7.26)$$

čiže AR a MA polynómy sú tvaru  $\Phi(z) = 1 - \phi_1 z$  a  $\Theta(z) = 1 + \theta_1 z$ . Tieto polynómy majú rovnaký koreň práve vtedy, keď  $\phi_1 = -\theta_1$  - v takom prípade pre model (7.26) platí

$$X_t = -\theta_1 X_{t-1} + W_t + \theta_1 W_{t-1}.$$

Takúto rovnosť ale spĺňa aj ARMA(0,0) model  $X_t = W_t$  (čiže biely šum), pretože ak  $X_t = W_t$  pre každé  $t$ , tak máme

$$\begin{aligned} X_t &= W_t \\ \theta_1 X_{t-1} &= \theta_1 W_{t-1} \end{aligned}$$

---


$$\begin{aligned} X_t + \theta_1 X_{t-1} &= W_t + \theta_1 W_{t-1} \\ X_t &= -\theta_1 X_{t-1} + W_t + \theta_1 W_{t-1}. \end{aligned}$$

Pre strednú hodnotu stacionárneho ARMA( $p, q$ ) platí

$$E(X_t) = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p},$$

počítala by sa úplne rovnako ako pri AR( $p$ ).

Funkcie ACF a PACF vo všeobecnom stacionárnom, invertovateľnom ARMA( $p, q$ ) nemajú „pekné“ vyjadrenie, avšak platí pre ne, že sú obe nenulové, ale k nule konvergujú. V tomto prípade sa teda z odhadov  $\hat{\rho}$  a  $\hat{\tau}$  nedá úplne priamo identifikovať rád  $(p, q)$  - o tom, ako tento rád v praxi hľadať, si povieme na online hodine. Väčšinou sa snažíme nájsť model, kde je buď  $p$  alebo  $q$  rovné

nule (v takom prípade vieme  $q$ , resp.  $p$  nájsť z ACF, resp. PACF) - ono to spravidla aj ide, avšak niekedy je potom model zbytočne preparametrizovaný - ak dáta napríklad dobre modeluje AR(8) a aj ARMA(1,1), spravidla je rozumnejšie si vybrať model s menším počtom parametrov (čiže ARMA(1,1)).

Predikcie v ARMA( $p, q$ ) fungujú rovnako ako v MA( $q$ ) - model zapíšeme v tvare AR( $\infty$ ) modelu, zanedbáme posledné členy (obsahujúce  $X_0, X_{-1}, X_{-2}, \dots$ ) a predikujeme rovnako ako pri autoregresnom procese. Takýmto predikciám sa hovorí „useknuté prediktory“ (truncated predictors), existujú aj trochu iné prístupy k predikovaniu, nám stačí vedieť o týchto.

## ARIMA modely

Doteraz sme sa sústredili hlavne na stacionárne modely pre časové rady. V praxi sa ale často stretáme s dátami, ktoré zjavne nie sú stacionárne, a teda stacionárne modely sa priamo použiť nedajú. Typickým (a najviditeľnejším) prejavom nestacionarity je prítomnosť nejakého nekonštantného trendu (v stacionárnych modeloch sa o trende (=strednej hodnote) predpokladá, že je konštantný). V Box-Jenkinsovej metodológii sa takéto situácie riešia takzvaným diferencovaním dát, t.j. nodelujeme priamo dáta  $x_1, x_2, \dots, x_n$ , ale ich diferencie  $x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1}$ . Ak potom nájdeme potom predikciu pre diferenciu  $x_{n+1} - x_n$ , tak pomocou nej a  $x_n$  ľahko dostaneme predikciu pre  $x_{n+1}$ . Ak by napríklad dáta pochádzali z modelu s lineárnym trendom

$$X_t = at + b + E_t,$$

kde  $\{E_t\}_{t \in \mathbb{Z}}$  je nejaký stacionárny časový rad, tak pre diferenciu  $X_t - X_{t-1}$  (označuje sa  $\nabla X_t$ ) už platí

$$\begin{aligned} \nabla X_t &= X_t - X_{t-1} = (at + b + E_t) - (a(t-1) + b + E_{t-1}) = at + b + E_t - at + a - b - E_{t-1} = \\ &= a + \nabla E_t, \end{aligned}$$

o čom sa dá ľahko ukázať, že už je to stacionárny časový rad (a teda pre diferencované dáta sa môžeme pokúsiť nájsť nejaký stacionárny ARMA model, ktorý by ich dobre opísal). Podobne ak by dáta pochádzali z modelu s kvadratickým trendom

$$X_t = at^2 + bt + c + E_t,$$

tak po druhej diferencii (diferencia z diferencie, označuje sa  $\nabla^2 X_t$ ), dostaneme

$$\begin{aligned} \nabla^2 X_t &= \nabla(X_t - X_{t-1}) = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} = \\ &= (at^2 + bt + c + E_t) - 2(a(t-1)^2 + b(t-1) + c + E_{t-1}) + (a(t-2)^2 + b(t-2) + c + E_{t-2}) = \\ &= (at^2 + bt + c + E_t) - 2(at^2 - 2at + a + bt - b + c + E_{t-1}) + (at^2 - 4at + 4a + bt - 2b + c + E_{t-2}) = \\ &= at^2 + bt + c + E_t - 2at^2 + 4at - 2a - 2bt + 2b - 2c - 2E_{t-1} + at^2 - 4at + 4a + bt - 2b + c + E_{t-2} = \\ &= 2a + \nabla^2 E_t, \end{aligned}$$

čo už je opäť stacionárny časový rad.

Indukciou sa dá relatívne ľahko ukázať, že ak by bol trend polynóm stupňa  $d$ , tak  $d$ -tou diferenciou  $\nabla^d X_t$  sa tohto polynómu zbavíme (z polynomického trendu ostane iba konštanta). V praxi spravidla stačí  $d \leq 2$  (väčšinu kriviek vieme aspoň lokálne aproximovať priamkou/parabolou, čo stačí), pričom aj  $d = 2$  je dosť zriedkavé.

**Autoregresný integrovaný moving average model** rádu  $(p, d, q)$  (skrátene ARIMA( $p, d, q$ )) je potom iba ARMA( $p, q$ ) model aplikovaný na  $d$ -krát diferencované dáta. Pomocou operátora vieme písať, že časový rad  $\{X_t\}_{t \in \mathbb{Z}}$  spĺňa

$$\Phi(B)\nabla^d X_t = c + \Theta(B)W_t,$$

respektíve

$$\Phi(B)(1 - B)^d X_t = c + \Theta(B)W_t$$

(lebo diferencia  $\nabla$  je to isté, čo  $(1 - B)$ ).

To, či je potrebné dáta diferencovať sa dá vidieť

- z obrázku - ak vidíme trend, resp. ak dáta nekolíšu okolo nejakej konštantnej hodnoty, tak je spravidla potrebné diferencovať,
- z ACF - ak táto funkcia ide k nule veľmi pomaly, tak tiež je to väčšinou znak toho, že je potrebná diferenciacia.

Toto samozrejme nie sú úplne najexaktnejšie pravidlá - treba ich brať skôr orientačne - ak náhodou spravíme diferenciu a nemali sme (respektíve nespravíme a mali sme), tak sa to ukáže aj v ďalších fázach výstavby modelu - ukážeme si na online hodine. Niekedy sa môže stať, že vieme relatívne rozumne modelovať aj diferencované a aj nediferencované dáta - správny model nemusí byť úplne jednoznačná záležitosť. V prípade, že by sme mali „podozrenie“ na potrebu viacerých diferencií (napríklad máme pocit, že v dátach je kvadratický trend), aj tak je dobré robiť diferencie postupne, t.j. vyrobíme si dáta  $y_t = x_t - x_{t-1}$  (jedna diferenciacia), vykreslíme ich spoločne s ACF a až potom zvážime, či naozaj treba ďalšiu diferenciu.