

Diplomová práca
Spektrálna analýza metódou najmenších štvorcov
a jej aplikácie v geodézii

Autor: Bc. Tomáš Bacigál

Bratislava 2003

Vedúci práce: Ing. Juraj Janák, PhD.

Obsah

Zoznam skratiek

1. <u>Úvod</u>	3
2. <u>Jednorozmerné (časové) rady a metódy ich analýzy</u>	3
3. <u>Fourierova spektrálna analýza</u>	5
4. <u>Od Fourierovej transformácie k metóde najmenších štvorcov</u>	8
5. <u>Základná myšlienka spektrálnej analýzy metódou najmenších štvorcov</u>	10
6. <u>Spektrum, spektrálna analýza a metóda najmenších štvorcov z pohľadu funkcionálnej analýzy</u>	13
7. <u>Štatistické testovanie významnosti spektrálnych vrcholov</u>	20
8. <u>Praktická aplikácia</u>	22
8.1. <u>Popis vstupných a výstupných súborov programu LSSA</u>	23
8.2. <u>Definícia a vyhodnotenie experimentu</u>	27
9. <u>Záver</u>	36

Zoznam literatúry

Zhrnutie

Prílohy

Zoznam skratiek

CV	celkový výkon
DFT	diskrétna Fourierova transformácia
FSA	Fourierova spektrálna analýza
FT	Fourierova transformácia
GPS	Globálny systém na určenie polohy
HVS	Hustota výkonového spektra
MNŠ	metóda najmenších štvorcov
RFT	rýchla Fourierová transformácia
SA	spektrálna analýza
SA-MNŠ	spektrálna analýza metódou najmenších štvorcov
T-MNŠ	transformácia metódou najmenších štvorcov

1. Úvod

Základom každej experimentálnej vedy je meranie, pretože je jediným nástrojom ako kvantitatívne opísať vlastnosti prírodných procesov, ktoré okolo nás vytvárajú reálny svet. Výsledky merania sú závislé od charakteru skúmaných javov i tých, ktoré merania svojim nežiaducim vplyvom znehodnocujú. Ak je proces jednorozmerný, získavame napr. časové rady, ak je funkciou vektora argumentov, tak najčastejšie sa stretáme s priestorovými radmi. Je zrejmé, že rad – súbor údajov vzniká opakovaním zberu údajov za odlišných podmienok, ktoré sú výsledkom dynamiky procesov závislej od hodnôt argumentov. Procesy alebo javy ešte možno odlíšiť na spojité (pohyb družice na dráhe) a diskrétny (hromadenie chýb na bodoch geodetickej siete), alebo na deterministické (explicitne matematicky definované) a stochastické (náhodné). Klasifikácia na posledné dva je do značnej miery subjektívna, pretože odzrkadľuje jednak našu schopnosť a potom aj vôľu určité zložky modelovať, resp. zo súboru meraní filtrovať.

Ďalej sa budeme zberať iba jednorozmernými spojitými procesmi, a ich kvantitatívnym vyjadrením vo forme empirickej funkcie $f(t)$, danej hodnotami v diskretných bodoch argumentu t . Ako som spomínal, argumentom t v geodetickej praxi býva najčastejšie čas, potom hovoríme o súboroch meraní ako o časových radoch, ale môže ním byť kľudne aj vzdialenosť, napr. ak skúmame deformácie na dĺžkovom meradle, a pod.

Vzhľadom na to, že výskumu dynamiky objektov v prírode sa v geodézii venuje čoraz väčšia pozornosť, zamerajme sa na časové rady. V nasledujúcej kapitole si priblížime prístupy, ktoré sa v analýze jednorozmerných radov v súčasnosti uplatňujú. Tieto nám umožnia lepšie pochopiť význam a výhody prezentovanej a ťažiskovej metódy tejto práce, ktorou je spektrálna analýza metódou najmenších štvorcov. Cieľom práce je predstaviť matematický základ, výhody a oblasti aplikácie tejto perspektívnej metódy, ktorá na analýzu geodetických meraní používa v geodézii notoricky známy princíp minimalizácie štvorcov odchýlok, a taktiež jej programovú implementáciu pre súčasné výpočtové platformy – program LSSA, ktorý bol navrhnutý na efektívne využitie počítačov pre vykonanie pomerne náročných algoritmov/operácií s maticami veľkých rozmerov.

2. Jednorozmerné (časové) rady a metódy ich analýzy

Časový rad (č.r.), ako príklad jednorozmerných radov spomínaných v úvode, môžeme chápať ako súbor údajov alebo pozorovaní určitého javu chronologicky usporiadaných v čase, tvorený diskretnými hodnotami v časových okamihoch najlepšie konštantného rozostupu. Požiadavka rovnomerného delenia časových intervalov je limitujúca pre mnoho metód, ktoré nerovnomerné rozloženie hodnôt argumentu t pred samotným spracovaním riešia interpoláciou alebo aproximáciou. Tým sa však znižuje informačná hodnota analyzovaného radu.

Význam analýzy spočíva v **pochopení** mechanizmu generujúceho pozorované údaje a **predpovedi** hodnôt do budúcnosti.

Pred rozborom zložiek (časového) radu je dobré aspoň sčasti poznať jeho charakter a hypoteticky ho rozložiť na známe, či predpokladané komponenty, v závislosti od charakteru procesov pôsobiacich pri meraní. V tomto smere možno v domácej a zahraničnej literatúre nájsť množstvo klasifikácií, ktoré rozkladajú časové rady do zložiek, či už podľa veľkosti vplyvu, štatistických vlastností alebo aj významu pre vyhodnocovateľa.

Tak napr. podľa (Cipra, 1986), vo všeobecnosti každý časový rad pozostáva zo štyroch zložiek. Po prvé z *trendu*, ktorý reprezentuje dlhodobý vývoj radu a môže byť aproximovaný priamkou alebo polynómom vyšších stupňov. Ďalej je to *sezónna* zložka popisujúca periodické zmeny v č.r. s periódou jedného roka prípadne nejakým iným násobkom štandardnej časovej jednotky, a *cyklická* zložka predstavujúca komplikovanejšie vplyvy, ktorých perióda aj amplitúda sa dokonca môže v čase meniť. Poslednou zložkou je *reziduálna*, ktorá ako jediná má náhodný charakter.

Podľa toho istého zdroja, na určenie a elimináciu uvedených zložiek sa dnes v analýze č.r. používajú tieto skupiny metód:

- § *Dekompozícia časových radov*. Kladie dôraz na prácu so systematickými zložkami č.r., jednotlivé pozorovania sa obvykle berú ako navzájom nekorelované.
- § *Box-Jenkinsonova metodológia*. Základný prvok konštrukcie modelu č.r. je reziduálna zložka, ktorá môže byť tvorená korelovanými náhodnými veličinami. Ťažisko tejto skupiny tvorí korelačná analýza.
- § *Lineárne dynamické modely*. Hodnoty určitého č.r. sú vysvetľované pomocou hodnôt ďalších vysvetľujúcich č.r.ádov.
- § *Spektrálna analýza*. Č.r. považuje za zmes sinusových a kosinusových kriviek s rôznymi amplitúdami a frekvenciami. Na rozdiel od predošlých troch metód, ktoré predstavujú analýzu č.r. v časovej doméne, táto je zameraná na doménu spektrálnu.

Uvedené metódy sa podľa klasického prístupu aplikujú v nasledujúcich niekoľkých krokoch. Regresnou analýzou najprv odstránime trend a prípadne aj sezónnu zložku. Ďalej spektrálnou analýzou určíme výrazné frekvencie pre cyklickú zložku a pomocou regresnej analýzy ju izolujeme ako sumu trigonometrických polynómov. Nakoniec sa testami náhodnosti overí, či reziduálny rad tvoria nekorelované veličiny. Ak nie, reziduálny rad sa spracuje Box-Jenkinsonovou metodológiou alebo modernými nelineárnymi modelmi. Ako si povieme neskôr, takéto separátne riešenie prináša pomerne veľké riziká straty informácie.

Iný pohľad na kompozíciu časových radov zohľadňuje prioritu ich analýzy. Podľa [Pagiatakis, 1999], zloženie časových radov možno pripísať dvom veličinám (resp. ich komplexu). Sú nimi *signál*, ktorý je predmetom nášho záujmu, a *šum* (z angl. *noise*), ktorý skúmaný signál v časovom rade skresľuje. Šum môže byť *náhodný* alebo *systematický*. Idealizovaný koncept náhodného šumu je biely šum, ktorý je kompletne nekorelovaný, jeho spektrálna hustota je konštantná (pozri kap.7), a môže a nemusí mať normálne (Gaussovo) rozdelenie pravdepodobnosti. V praxi sa však oveľa častejšie stretáme s náhodným šumom, ktorý na rozdiel od toho bieleho je pásmovo obmedzenou funkciou času.

Systematický šum, ktorého podobu môžeme popísať určitým funkčným vzťahom, je buď periodický (farebný), alebo neperiodický. Neperiodický šum štandardne pozostáva jednak z tzv. *dátových skokov* (z angl. *datum shift*), ktorými sa rozumie vychýlenie alebo náhlu zmenu strednej hodnoty radu v určitých časových intervaloch, a jednak z *trendu* (lineárneho, kvadratického, exponenciálneho a pod.). Neperiodický systematický šum spôsobuje nestacionaritu časového radu, t.j. závislosť štatistických vlastností (stredná hodnota, autokorelačná funkcia, rozptyl a všetky momenty vyšších rádov) na voľbe počiatku epochy.

Nestacionarita časového radu je podobne ako nepravidelný rozostup jeho hodnôt vlastnosťou, s ktorou sa bežné techniky spektrálnej analýzy nevedia priamo vysporiadať, a tak v rámci predspracovania sa z časového radu zvyčajným spôsobom zvykne odstrániť vplyv spomínaného systematického vplyvu. „Zvyčajný spôsob“ znamená separátne odhadnutie veľkostí všetkých známych zložiek v rade a ich odfiltrovanie. Takýto postup však necháva negatívne následky na spektrálnom obraze reziduálneho radu, dochádza k jeho skresleniu.

Jednou z bežných a najznámejších techník analýzy periodických procesov (signálov) v časových radoch je Fourierova spektrálna analýza. Vychádza z matematických princípov Fourierovej transformácie a stručne si ju zhrnieme v nasledujúcej kapitole. Dôvod, prečo ju uvádzam je ten, že Fourierova spektrálna analýza (FSA) je špecifickým prípadom v mnohom univerzálnejšej spektrálnej analýzy pomocou MNS. Veľmi dobre je súvislosť medzi fourierovou a MNS transformáciou jednorozmerných časových radov popísaná v (Craymer, 1998).

3. Fourierova spektrálna analýza

Fourierova transformácia

Je známe, že akúkoľvek spojitú periodickú funkciu $f(t)$ s periodou T možno matematicky vyjadriť vo forme nekonečného radu trigonometrických funkcií

$$f(t) = \frac{1}{2} a_0 + \sum_{i=1}^{\infty} (a_i \cos 2\pi f_i t + b_i \sin 2\pi f_i t) \quad (3.1)$$

nazývaného aj *Fourierov rad* (FR), kde a_i a b_i sú Fourierové koeficienty prislúchajúce frekvencii f_i . Frekvencia f_i sa dá zapísať aj ako násobok fundamentálnej frekvencie f_0 ako $f_i = i f_0$, kde $f_0 = 1/T$. Ak uvážime fakt, že funkcie kosínus a sínus formujú ortogonálnu bázu (pozri kap. 6) nad intervalom $(-T/2, T/2)$, Fourierove koeficienty sú dané

$$a_i = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos 2\pi f_i t dt$$
$$b_i = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin 2\pi f_i t dt \quad (3.2)$$

pre každé $i = 0, 1, \dots, \infty$.

Pre neperiodické funkcie podobný Fourierov rad neexistuje, môžeme však definovať novú funkciu, ktorá bude tvorená opakovanou sekvenciou neperiodickej funkcie na určitom konečnom intervale, povedzme $(-T/2, T/2)$. Novú, periodickú funkciu s periódou T tak už bude možné vyjadriť pomocou FR. Ak $T \rightarrow \infty$, diskrétna sada frekvencií vo FR sa stane spojitou a tým FR prejde na tvar tzv. Fourierovho integrálu (FI)

$$f(t) = \int_{-\infty}^{\infty} (a(f) \cos 2\pi f t + b(f) \sin 2\pi f t) df, \quad \forall t \quad (3.3)$$

kde Fourierove koeficienty $a(f)$ a $b(f)$ sú definované podobne ako v (3.2), ale v inegračnom rozsahu $(-\infty, \infty)$. Samozrejme za podmienky, že funkcia $f(t)$ je absolútne integrovateľná, čo je splnené ak v nekonečne konverguje k nule.

Ok, uvažovali sme periodické a neperiodické funkcie, ale v praxi sa stretávame skôr s náhodnými alebo stochastickými funkciami (procesmi), ktoré nemusia byť nutne periodické, a navyše ani nebývajú absolútne integrovateľné, keďže už z definície stacionarity nekonvergujú v nekonečne k nule. I tento problém však možno obísť jednoduchým orezaním stochastickej funkcie v rozsahu, napríklad od $-T/2$ do $T/2$, podobne ako to bolo pri neperiodickej funkcii, pričom mimo tohto intervalu bude nulová. Tak splníme podmienku absolútnej integrovateľnosti a reprezentáciu stochastickej funkcie môžeme zapísať výrazom (3.3), kde však Fourierove koeficienty sú definované konečným Fourierovým integrálom v hraniciach $(-T/2, T/2)$ namiesto $(-\infty, \infty)$.

Všeobecný zápis FR a FI počíta s komplexným charakterom funkcie $f(t)$. Reálnej zložke je priradený výraz \sin , imaginárnej zložke výraz \cos a použitím Eulerovej formuly $\cos 2pft + j \sin 2pft = e^{j2pft}$, kde $j = \sqrt{-1}$ je imaginárna jednotka, dostaneme FI v tvare

$$f(t) = \int_{-\infty}^{\infty} F(f) e^{j2pft} df \quad (3.4)$$

a takisto

$$F(f) = \begin{cases} \int_{-\infty}^{\infty} f(t) e^{-j2pft} dt & \text{pre neperiodickú funkciu} \\ \int_{-T/2}^{T/2} f(t) e^{-j2pft} dt & \text{pre stochastickú funkciu} \end{cases} \quad (3.5)$$

, kde $F(f)$ je tzv. komplexné *fourierovské spektrum*.

Výraz (3.4), resp. (3.5), predstavuje priamu ($f(t) \rightarrow F(f)$), resp. inverznú ($F(f) \rightarrow f(t)$) Fourierovú transformáciu, a funkcie $f(t)$ a $F(f)$ tvoria tzv. *Fourierov pár*, symbolicky $f(t) \Leftrightarrow F(f)$.

Opäť, v praxi zriedkavo narazíme na spojitý stochastický proces nekonečnej dĺžky, skôr skúmame diskrétny procesy v konkrétnom okamihu, alebo diskrétny vzorkované (digitalizované) spojité procesy konečnej dĺžky. Tie môžu a nemusia byť vždy vzorkované ekvidištane, čo je nanešťastie jednou z podmienok aplikácie FSA. Pre diskrétny a pravidelne vzorkované stochastické procesy $\{f(t), i=0,1,\dots,n-1\}$, ak sumáciou aproximujeme Fourierov integrál (3.5), dostávame diskrétnu Fourierovú transformáciu DFT

$$F(f_k) = \int_{-\infty}^{\infty} f(t) e^{-j2p f_k t} dt \cong Dt \sum_{i=0}^{n-1} f(t_i) e^{-j2p f_k t_i} \quad (3.6)$$

kde n je počet observácií (vzoriek), Dt vzorkovací krok (tiež digitalizačný krok, alebo vzorkovací interval) a f_k je jedna zo sady odhadnuteľných frekvencií

$$f_k = \frac{k}{T} = \frac{k}{nDt} = kf_0, \quad \forall k = -\frac{n}{2} \dots \frac{n}{2} \quad (3.7)$$

kde $T = nDt$ je dĺžka radu údajov, $f_0 = 1/T$ základná (fundamentálna) frekvencia. Tento rad celočíselných násobkov základnej frekvencie budeme jednoducho označovať ako „Fourierove“ frekvencie. Kritická frekvencia

$$f_N = \frac{1}{2Dt} \quad (3.8)$$

sa nazýva *Nyquistova frekvencia* a pri danom digitalizačnom kroku Dt je principiálne nemožné zistiť Fourierovo spektrum pre frekvencie väčšie ako je Nyquistova.

Podľa konvencie, výsledné vzťahy DFT sa udávajú bez člena Dt pred sumáciou. Priama DFT je teda definovaná vzťahom

$$F_k = \sum_{i=0}^{n-1} f(t_i) e^{-j2p f_k t_i}, \quad \forall k \quad (3.9)$$

a inverzná DFT zase

$$f(t_i) = \frac{1}{n} \sum_{k=0}^{n-1} F_k e^{j2p f_k t_i}, \quad \forall i \quad (3.10)$$

Verím, že z doterajších úvah si pozorný čitateľ i bez základných znalostí z oblasti spektrálnej analýzy uvedomil, že algoritmus Fourierovej transformácie aplikovaný napríklad na časový rad predstavuje transformáciu hodnôt empirickej funkcie medzi dvoma doménami – časovou

a frekvenčnou, čo prakticky znamená že môžeme skúmať dynamiku prebiehajúcich prírodných procesov akoby z dvoch odlišných strán, v dvoch celkom rozdielnych priestoroch.

V spektrálnej oblasti je nositeľom informácie o procesoch funkcia, ktorá každej frekvencii z množiny definovaných frekvencií priradí hodnotu jej významnosti, resp. prislúchajúcej periodickej zložky v časovom rade. Vo všeobecnosti jej hovoríme spektrum, je základom spektrálnej analýzy a v nasledujúcom texte si o nej povieme viac.

Spektrum

Aproximácia empirických funkcií v zmysle Fourierových radov má jednu zaujímavú fyzikálnu interpretáciu spojenú s termínom výkon (z angl. *power*, ktorý značí aj silu, energiu, či účinnosť). Uvažujme absolútne integrovateľnú neperiodickú funkciu $f(t)$, potom tzv. celkový výkon (total power) CV funkcie $f(t)$ je definovaný vzťahom

$$CV = \int_{-\infty}^{\infty} f(t)^2 dt \quad (3.11)$$

čo je ekvivalentom vzťahu

$$CV = \int_{-\infty}^{\infty} |F(f)|^2 df \quad (3.12)$$

kde namiesto originálnej funkcie integrujeme jej Fourierovú transformáciu. Oba predošlé výrazy sú známe pod názvom Parsevalov vzťah. Aby sme lepšie pochopili pojem „celkový výkon“, poznamenajme, že je taktiež rovný n -násobku variancie (rozptylu) σ^2 . Každý člen $|F(f)|^2 df$ reprezentuje príspevok (podiel, účasť) do celkového výkonu v $f(t)$, tvoreného zložkami o frekvenciách z intervalu $(f, f+df)$. Potom tzv. hustota výkonového spektra HVS je definovaná

$$s(f) = |F(f)|^2 \quad (3.13)$$

a grafické znázornenie $s(f)$ podľa frekvencie f sa tiež nazýva výkonové spektrum, prípadne len spektrum.

Pre neperiodické a stochastické funkcie, orezané do intervalu $(-T/2, T/2)$, je podobne ako pre periodické funkcie výhodnejšie definovať celkový výkon za časovú jednotku cez konečný interval, teda

$$\frac{CV\left(-\frac{T}{2}, \frac{T}{2}\right)}{T} = \frac{1}{T} \int_{-T/2}^{T/2} f(t)^2 dt = \frac{1}{T} \int_{-\infty}^{\infty} |F(f)|^2 df = \int_{-\infty}^{\infty} s(f) df \quad (3.14)$$

kde $s(f) = 1/T |F(f)|^2$. Toto spektrum je funkciou kladných i záporných frekvencií, čo je pre nás v podstate zbytočné, pretože medzi nimi zvyčajne nerozlišujeme. Navyše, ak $f(t)$ je reálna, jej FT je párna (tzn. $F(f) = F(-f)$) a tak si môžeme vyjadriť spektrum ako funkciu kladných frekvencií a nazvať ho *jednostrannou funkciou spektrálnej hustoty* $s(f) = 2|F(f)|^2, \forall f \in \langle 0, \infty \rangle$. Ak uvažujeme digitalizované procesy, kedy je spektrum definované pre diskretnú množinu „Fourierových“ frekvencií, diskretná forma Parsevalovho vzťahu dostáva tvar

$$CV = \sum_{k=0}^{n-1} f(t_k)^2 = \frac{1}{n} \sum_{k=0}^{n-1} |F(f_k)|^2 = \sum_{k=0}^{n-1} s(f_k) \quad (3.15)$$

a individuálne hodnoty spektra $s(f_k)$ pre HVS

$$s(f_k) = \frac{1}{n} |F(f_k)|^2 \quad (3.16)$$

ktoré je vhodné normalizovať, aby vyjadrovali percentuálny podiel zložky s frekvenciou f_k na celkovom výkone, alebo rozptyle, vzťahom

$$\mathcal{P}(f_k) = \frac{s(f_k)}{\sum_{k=0}^{n-1} s(f_k)} = \frac{s(f_k)}{\sum_{k=0}^{n-1} f(t_k)^2} = \frac{|F(f_k)|^2}{n^2 S^2}. \quad (3.17)$$

Pre reálne funkcie je záporná časť frekvencií spektra zrkadlovým obrazom tej kladnej, preto uvažujúc iba Fourierové frekvencie $f_k = 0, \dots, 1/(2Dt)$ dostaneme jednostranné spektrum. Spektrum počítané na tomto princípe sa všeobecne označuje ako *periodogram* a ako si neskôr ukážeme, formuje základ spektra MNŠ.

V minulosti bola náročnosť výpočtu DFT pre väčšie množstvo údajov limitujúcim faktorom jej použitia, preto bola vyvinutá tzv. *rýchla fourierova transformácia* RFT, ktorá dramaticky znížila počet nutných operácií na vykonanie DFT a revolucionalizovala tak jej použitie. RFT však rovnako ako DFT vyžaduje konštantný digitalizačný krok bez prerušení („diery“ v rade) a navyše dĺžka radu musí byť celočíselná mocnina dvoch. Nesplnenie druhej podmienky sa rieši alebo useknutím alebo doplnením núl na koniec radu. Takisto obmedzenie sa na sadu „Fourierových“ frekvencií je zdedenou nevýhodou RFT, pretože za prítomnosti iných významných frekvencií ako štandardných dochádza k fenoménu známemu ako *únik spektra* (z angl. *spectral leakage*), ktorý degraduje výsledky transformácie. Čiastočným riešením, ak sa zameriame iba na určitú oblasť spektra, môže byť použitie tzv. filtrov. Viac informácií o vlastnostiach a aplikáciách Fourierovej transformácie a odkazy na iné zdroje môže čitateľ nájsť v špecializovanej literatúre, napr. (Červený, 1977).

4. Od Fourierovej transformácie k metóde najmenších štvorcov

Na úvod tejto kapitoly uvediem zápis uvedených vzťahov pre DFT v maticovom tvare, ktorý nám v ďalších úvahách pomôže pochopiť súvislosť s transformáciou MNŠ, ako ju uvádza (Craymer, 1998).

DFT

Priama Fourierova transformácia zo vzťahu (3.9) sa dá vyjadriť ako

$$\mathbf{F}^c = \mathbf{A}^{cT} \mathbf{f}, \quad (4.1)$$

kde

$$\mathbf{F}^c = \begin{bmatrix} F_0 \\ F_1 \\ \mathbf{M} \\ F_{n-1} \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} f(t_0) \\ f(t_1) \\ \mathbf{M} \\ f(t_{n-1}) \end{bmatrix}, \quad (4.2)$$

$$\mathbf{A}^c = \begin{bmatrix} \mathbf{A}^c_{f_0} & \mathbf{A}^c_{f_1} & \mathbf{K} & \mathbf{A}^c_{f_{n-1}} \end{bmatrix} = \begin{bmatrix} e^{j2p f_0 t_0} & e^{j2p f_1 t_0} & \mathbf{L} & e^{j2p f_{n-1} t_0} \\ e^{j2p f_0 t_1} & e^{j2p f_1 t_1} & \mathbf{L} & e^{j2p f_{n-1} t_1} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ e^{j2p f_0 t_{n-1}} & e^{j2p f_1 t_{n-1}} & \mathbf{K} & e^{j2p f_{n-1} t_{n-1}} \end{bmatrix}. \quad (4.3)$$

Horný index „c“ označuje komplexnú maticu a transpozícia matice \mathbf{A}^c , ktorá obsahuje komplexne združené členy, bude

$$\mathbf{A}^{cT} = \begin{bmatrix} \mathbf{A}^c_{f_0}{}^T \\ \mathbf{A}^c_{f_1}{}^T \\ \mathbf{M} \\ \mathbf{A}^c_{f_{n-1}} \end{bmatrix} = \begin{bmatrix} e^{-j2p f_0 t_0} & e^{-j2p f_0 t_1} & \mathbf{L} & e^{-j2p f_0 t_{n-1}} \\ e^{-j2p f_1 t_0} & e^{-j2p f_1 t_1} & \mathbf{L} & e^{-j2p f_1 t_{n-1}} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ e^{-j2p f_{n-1} t_0} & e^{-j2p f_{n-1} t_1} & \mathbf{K} & e^{-j2p f_{n-1} t_{n-1}} \end{bmatrix}. \quad (4.4)$$

Inverznú DFT podľa (3.10) zapíšme

$$\mathbf{f} = \frac{1}{n} \mathbf{A}^c \mathbf{F}^c. \quad (4.5)$$

Pre praktické aplikácie napr. v geodézii a geofyzikálnych vedách, obmedzíme sa na údaje z oboru reálnych (nie komplexných) čísel a kladné frekvencie. Nahradením komplexných výrazov zodpovedajúcimi reálnymi výrazmi v trigonometrickej forme podľa Eulerovej formuly vo vzťahu pre priamu resp. inverznú DFT

$$F_k = \sum_{i=0}^{n-1} f(t_i) (\cos 2p f_k t_i - j \sin 2p f_k t_i), \quad \forall k = 0, 1, \dots, n-1, \quad (4.6)$$

resp.

$$f(t_i) = F_0 (\cos 2p f_0 t_i) + \sum_{k=1}^{n-1} F_k (\cos 2p f_k t_i + j \sin 2p f_k t_i), \quad \forall i = 0, 1, \dots, n-1 \quad (4.7)$$

budú komponenty maticového zápisu DFT v (4.4) resp. (4.5) zbavené indexu „c“ a pre f_k vyjadrené nasledovne:

$$\mathbf{F}_k = \begin{bmatrix} \text{Re}(F_k) \\ \text{Im}(F_k) \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{n-1} f(t_i) \cos 2p f_k t_i \\ \sum_{i=0}^{n-1} f(t_i) \sin 2p f_k t_i \end{bmatrix} \quad (4.8)$$

$$\mathbf{A}_{f_k} = \begin{bmatrix} \cos 2p f_k t_0 & \sin 2p f_k t_0 \\ \cos 2p f_k t_1 & \sin 2p f_k t_1 \\ \mathbf{M} & \mathbf{M} \\ \cos 2p f_k t_{n-1} & \sin 2p f_k t_{n-1} \end{bmatrix}. \quad (4.9)$$

Pre nulovú frekvenciu je $\text{Im}(f_0)=0$ a všetky sinusy v \mathbf{A}_{f_0} takisto nulové. Všimnime si, že zložky komplexného spektra F_k predstavujú základ fourierových koeficientov a_k a b_k pre DFT.

Vektor \mathbf{F} a matica \mathbf{A} teda vyzerajú

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_0 \\ \mathbf{F}_1 \\ \mathbf{M} \\ \mathbf{F}_{n-1} \end{bmatrix} = \begin{bmatrix} \text{Re}(F_0) \\ \text{Re}(F_1) \\ \text{Im}(F_1) \\ \mathbf{M} \\ \text{Re}(F_{n-1}) \\ \text{Im}(F_{n-1}) \end{bmatrix} \quad (4.10)$$

$$A = \begin{bmatrix} \cos 2p f_0 t_0 & \cos 2p f_1 t_0 & \sin 2p f_1 t_0 & \mathbf{K} & \cos 2p f_{n-1} t_0 & \sin 2p f_{n-1} t_0 \\ \cos 2p f_0 t_1 & \cos 2p f_1 t_1 & \sin 2p f_1 t_1 & \mathbf{K} & \cos 2p f_{n-1} t_1 & \sin 2p f_{n-1} t_1 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} & \mathbf{M} \\ \cos 2p f_0 t_{n-1} & \cos 2p f_1 t_{n-1} & \sin 2p f_1 t_{n-1} & \mathbf{K} & \cos 2p f_{n-1} t_{n-1} & \sin 2p f_{n-1} t_{n-1} \end{bmatrix}.$$

Transformácia MNŠ

Z predošlých maticových zápisov DFT možno odvodiť všeobecnejšiu formu transformácie, a to transformáciu metódou najmenších štvorcov (T-MNŠ). Stačí si uvedomiť, že DFT a jej inverzia sú ekvivalentom aproximácie (al.interpolácie) funkčných vzťahov technikou MNŠ použitím trigonometrických funkcií (t.j. Fourierových radov) ako bazových funkcií (bližšie (Vaníček a Krakiwsky, 1986), alebo aj kap.6).

Konkrétne, aproximácia vektoru meraní f v zmysle Fourierových radov môže byť zapísaná v maticovom tvare ako

$$f = Ax, \quad (4.11)$$

kde

$$x = \begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ \mathbf{M} \\ a_{n-1} \\ b_{n-1} \end{bmatrix} \quad (4.12)$$

je vektor odhadovaných Fourierových koeficientov a matica A reprezentuje bázu trigonometrických funkcií ako je definovaná v (4.10). Na odhad Fourierových koeficientov x sa využije minimalizačné kritérium MNŠ (Vaníček a Krakiwsky, 1986, str. 204-207), riešením je

$$\hat{x} = N^{-1} A^T f, \quad (4.13)$$

kde $N = A^T A$ je matica koeficientov normálnych rovníc. Ak si uvedomíme, že $F = A^T f$ je (simultánna) priama DFT, môžeme (4.13) prepísať na

$$\hat{x} = N^{-1} F \quad (4.14)$$

a (4.11) na

$$\hat{f} = AN^{-1} F, \quad (4.15)$$

čo predstavuje (simultánnu) priamu a inverznú T-MNŠ.

To, že konvenčná Fourierova transformácia je iba špeciálnym prípadom uvedených definícií T-MNŠ pre váhovo rovnocenné a ekvidištančne krokované údaje, je na prvý pohľad zjavné z porovnania oboch priamych transformácií. Z maticovej formy oboch inverzných transformácií to už také očividné nie je, stačí však iba vyšetriť prvky N^{-1} ($N = \text{diag}(n, n/2, n/2, \dots)$), pozri (Craymer, 1998, na str.48).

Popri schopnosti spracovať aj nerovnomerne delené rady, vyznačuje sa T-MNŠ ešte ďalšou významnou vlastnosťou, a síce, že nie sme obmedzení iba na sadu „Fourierových“ frekvencií $f_k = k/(nDt)$ pre $k=0,1,\dots,n-1$, ale k dispozícii máme akékoľvek frekvencie z intervalu $(0, f_N)$. Najviac však iba $n/2$ frekvencií (n Fourierových koeficientov) možno súbežne odhadnúť z n observácií. Navyše, ak výberom určitých frekvencií spôsobíme takmer lineárnu závislosť medzi dotknutými Fourierovými komponentmi (trig. funkciami), matica N bude zle podmienená (blízko singularná). Tomu sa vyhneme, ak buď zvolíme inú množinu frekvencií (napr. rovnomerne krokované), alebo jednoducho zanedbáme korelácie v N (t.j.

mimodiagonálne prvky) a odhad inverznej transformácie vykonáme oddelene pre individuálne frekvencie.

Spektrum

Zapíšme celkový výkon z (3.15) maticami

$$CV = \mathbf{f}^T \mathbf{f} = \frac{1}{n} \mathbf{F}^T \mathbf{F}, \quad (4.16)$$

kde

$$\mathbf{F}^T \mathbf{F} = \sum_{k=0}^{n-1} \mathbf{F}_k^T \mathbf{F}_k = \sum_{k=0}^{n-1} |F(f_k)|^2 \quad (4.17)$$

a jednotlivé spektrálne hodnoty z (3.16) budú

$$s(f_k) = \frac{1}{n} \mathbf{F}_k^T \mathbf{F}_k = \frac{1}{n} |F(f_k)|^2, \quad \forall k = 0, 1, \dots, n-1 \quad (4.18)$$

(jednostranná HVS je 2-násobkom obojstrannej ak $k=1 \dots (n/2)-1$).

Ekvivalentom predošlých vzťahov v zmysle transformácie MNŠ sú rovnice

$$CV = \mathbf{f}^T \mathbf{f} = \mathbf{F}^T \mathbf{N}^{-1} \mathbf{F} \quad (4.19)$$

pre celkový výkon a

$$s(f_k) = \mathbf{F}_k^T \mathbf{N}_k^{-1} \mathbf{F}_k \quad (4.20)$$

pre jednostrannú HVS, kde \mathbf{N}_k je k -ty diagonálny blok matice \mathbf{N} prislúchajúci frekvencii f_k . Vo všeobecnosti, CV nemožno vyjadriť sumou individuálnych príspevkov rozdielnych frekvenčných zložiek. Problém opäť spôsobuje nerovnomerné delenie údajov, kedy \mathbf{N} netvorí diagonálnu maticu pretože Fourierove komponenty (trig. funkcie) nie sú vzájomne ortogonálne. Riešením je vyhodnotenie každej frekvencie zvlášť, nezávisle na ostatných, čo je rovné zanedbaniu lineárnych závislosti medzi rozdielnymi frekvenčnými zložkami v \mathbf{N} , a definovaniu spektra na základe nezávislosti príspevkov každej frekvenčnej zložky ku CV .

Normalizované spektrum $s(f_k)$ je dané

$$\%f_k = \frac{s(f_k)}{\mathbf{f}^T \mathbf{f}} = \frac{\mathbf{F}_k^T \mathbf{N}_k^{-1} \mathbf{F}_k}{\mathbf{f}^T \mathbf{f}} \quad (4.21)$$

a reprezentuje *percento variancie* súboru údajov nezávisle pre každú spektrálnu zložku (ľubovolnej nezápomej frekvencie).

Spektrum MNŠ v základnej forme bolo prvý krát odvodené a publikované v (Vaníček; 1969a, 1971) a jeho základnú ideu by som rád načrtol v nasledujúcej kapitole. Autor tu spomína i vzťah k Fourierovej spektrálnej analýze, súvis s vyššie uvedenými vzťahmi však bude zrejmy najmä v kap.6.

5. Základna myšlienka spektrálnej analýzy metódou najmenších štvorcov

V tejto kapitole si priblížime základné princípy metódy spektrálnej analýzy MNŠ (SA-MNŠ, alebo iba SA) ako ich odvodil jej autor P.Vaníček vo svojej práci z roku 1969, resp.1971. Na jeho výskum nadviazali mnohí autori, ktorí túto perspektívnu techniku v splupráci s autorom rozpracovali do pozoruhodných detailov, a tie publikovali na medzinárodných odborných seminároch a v špecializovaných časopisoch (pozri zoznam literatúry na konci DP).

Klasické techniky spektrálnej analýzy predpokladajú stacionárny charakter časových radov, to znamená, že ich stredná hodnota nie je závislá na počiatočnom okamihu. To však samozrejme

v geodézii a geofyzikálnych vedách splnené nebýva, najmä pre prítomnosť sekulárnych trendov, dlhoperiodických vplyvov a pod. Aj keď určité techniky dokážu problém nestacionarity obísť, ich výsledky v prípade neznalosti analyzovanej empirickej funkcie strácajú svoju silu. SA-MNŠ rieši problémy s nestacionaritou a navyše s nepravidelným delením časových radov z odlišného uhla.

V nasledujúcom texte budeme používať označenia

$$\langle F, G \rangle = \sum_{t \in N_n} (F(t)G(t)) \quad (5.1)$$

pre skalárny súčin dvoch ľubovoľných funkcií F a G definovaných na $N_n \equiv \{t_1, t_2, \dots, t_n\}$,

$$r(F) = \langle F, F \rangle \quad (5.2)$$

pre normu funkcie F , a

$$r(F, G) = r(F - G) \quad (5.3)$$

pre strednú kvadratickú vzdialenosť medzi F a G .

Majme časový rad observácií $F(t)$, $t \in N_n$ o ktorom vieme iba to, že obsahuje konštantnú zložku, závislú od zvoleného časového intervalu. Budeme predpokladať, že spektrum funkcie F je nanajvyš v lineárnej závislosti so spektrom funkcie $F+C$ pre ľubovoľnú konštantu C . Druhou požiadavkou je, aby hodnota spektra predstavovala príspevok príslušnej frekvencie do celkovej variancie funkcie F . Ďalej definujeme

$$r(F, G) = \sum_{t \in N_n} (F(t) - T(t))^2, \quad (5.4)$$

kde

$$T(t) = c_1 + c_2 \cos wt + c_3 \sin wt. \quad (5.5)$$

$r(F, T)$ vyjadruje určitú lineárnu transformáciu variancie reziduí funkcie F po odstránení trigonometrického členu T . Potom

$$s(F, T) = r(F) - r(F, T) \quad (5.6)$$

vyjadruje príspevok frekvencie w do $r(F)$. Pretože sme zatiaľ neurčili, o aký príspevok má ísť, $s(F, T)$ nebude jediné. Prírodzene, najvhodnejší bude maximálny príspevok frekvencie w , čo nás okamžite vedie k minimalizácii $r(F, T)$. Minimum je dané pre koeficienty c_1, c_2, c_3 zo sústavy normálnych rovníc, a tak v konečnom dôsledku $s(F, T)$ bude funkciou iba premennej w . $s(w)$ predstavuje maximálny príspevok frekvencie w do $r(F)$ a po odstránení strednej hodnoty z F nadobudne zjednodušený tvar

$$s(w) = p(w) \left(\sum_{t \in N_n} F(t) \cos wt \right)^2 + r(w) \left(\sum_{t \in N_n} F(t) \sin wt \right)^2 \quad (5.7)$$

čo nie je nič iné ako príspevok w do n -násobku variancie F . Všimnime si podobnosť s Fourierovou transformáciou

$$R(w) = \frac{2}{n} \left(\left(\sum_{t \in N_n} F(t) \cos wt \right)^2 + \left(\sum_{t \in N_n} F(t) \sin wt \right)^2 \right) \quad (5.8)$$

(jednostranná hustota výkonového spektra). A keďže oboma funkciami možno priamo odhaľovať frekvencie (a tak i periódy) v časových radoch, zvyknú sa označovať pojmom *periodogram*. Koeficienty $p(w)$ a $r(w)$ v (5.7)

$$p(w) = \frac{1}{\sum_{t \in N_n} \cos^2 wt - \frac{1}{n} \left(\sum_{t \in N_n} \cos wt \right)^2} \quad r(w) = \frac{1}{\sum_{t \in N_n} \sin^2 wt} \quad (5.9)$$

považujme za korekčné členy zavedené do Fourierovej transformácie, a zodpovedné za spomínané vlastnosti periodogramu MNŠ.

Ak F je jednoduchá sinusoida frekvencie w , $s(w) = r(F)$ je jediné maximum. To nám dáva prísluš viacmenej správneho obrazu spektra funkcie, ak v určitej frekvenčnej oblasti obsahuje jediné dominantné periodické komponentum, dokonca i v oblasti nízkych frekvencií. Toto tvrdenie o Fourierovej transformácii všeobecne neplatí.

Uvedené poznatky sú iba jednoduchým úvodom do problematiky, ktorú v nasledujúcich kapitolách rozšírime jednak o ďalšie súvislosti s minimalizačným princípom MNŠ a jeho interpretáciou v jazyku funkcionálnej analýzy, a jednak o zahrnutie vplyvov, ktoré sa v analyzovaných časových radoch podieľajú na ich nestacionarite, a s ktorými sa iné techniky SA nevedia tak elegantne vyrovnáť.

6. Spektrum, spektrálna analýza a metóda najmenších štvorcov z pohľadu funkcionálnej analýzy

Doteraz sme si predstavili spektrálnu analýzu MNŠ z dvoch rôznych uhlov pohľadu, ako rozšírenie známej Fourierovej SA, a ako jednoduchý koncept reprezentácie minimalizačného procesu vo vektorovom priestore. V tejto kapitole sa pokúsím podať rozšíreý pohľad na veľmi známu a aj v geodézii obľúbenú MNŠ, využitím pojmov funkcionálnej analýzy ako vektorový priestor, báza, ortogonálna projekcia a pod. (podľa (Wells, 1985)).

Každý geodet určite pozná dva najznámejšie postupy spracovania meraní, ktoré používajú rovnaký algoritmus, ale každý na iný účel. Ide o *lineárny odhad parametrov MNŠ* a *aproximáciu empirických funkcií MNŠ* (názvy sú orientačné, inak môžeme hovoriť aj o vyrovnaní nepriamych meraní, resp. priamych). Algoritmus je známy vo svojom maticovom zápise

$$\hat{x} = (A^T P A)^{-1} A^T P f \quad (6.1)$$

$$\hat{v} = f - A \hat{x}, \quad (6.2)$$

kde, pri lineárnom odhade parametrov

f je vektor observácií,

A matica plánu, ktorá modeluje fyzikálny vzťah medzi meraniami a vektorom neznámych parametrov x z observačných rovníc $f = Ax$,

P váhová matica, ktorá je inverziou kovariančnej matice C_f meraní,

a úlohou je získať odhad určitých fyzikálnych parametrov x , na základe observácií f . Teda chceme \hat{x} (plus ich kovariančnú maticu C_x).

Pri aproximácii

f známy vektor, ktorý má byť aproximovaný,

A matica obsahujúca niekoľko stĺpcových vektorov $A = [A_1, A_2 \dots A_m]$, nazývaných bázové funkcie. $\text{Rows}(A) = \text{dim}(f)$.

P je váhová matica bez štatistického významu, často rovná jednotkovej matici.

Úlohou je nájsť funkciu \mathbf{p} , ktorá najlepšie aproximuje \mathbf{f} v zmysle minimalizácie štvorcov reziduí $\mathbf{v}=\mathbf{f}-\mathbf{p}$, pričom koeficienty $\hat{\mathbf{x}}$ sú už menej dôležité. Obe funkcie možno zapísať i v tvare

$$\mathbf{f} = \sum_{i=1}^m x_i \mathbf{A}_i, \quad (6.3)$$

$$\mathbf{p} = \sum_{i=1}^m \hat{x}_i \mathbf{A}_i. \quad (6.4)$$

V prípade spektrálnej analýzy MNŠ, ktorá je aplikáciou techniky aproximácie empirických funkcií, poznáme $\mathbf{f}(\{f_i, t_i\}$ je časový rad) a definujeme

$$\mathbf{A}_1 = \cos w_j t, \quad (6.5)$$

$$\mathbf{A}_2 = \sin w_j t.$$

Pre každú frekvenciu w_j počítame

$$\mathbf{p} = \hat{x}_1 \cos w_j t + \hat{x}_2 \sin w_j t, \quad (6.6)$$

kde $\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$ sa určí z (6.1).

Stupeň aproximácie \mathbf{f} funkciou \mathbf{p} môžeme vo všeobecnosti vyjadriť pomerom

$$s = \frac{\text{dĺžka ortogonálnej projekcie } \mathbf{p} \text{ na } \mathbf{f}}{\text{dĺžka } \mathbf{f}} = \frac{\mathbf{f}^T \mathbf{p}}{\mathbf{f}^T \mathbf{f}}, \quad (6.7)$$

pričom ak je aproximácia dokonalá ($\mathbf{v} = \mathbf{0}$), $s = 1$. Ak však $\hat{\mathbf{x}} = \mathbf{0}$, \mathbf{f} je kolmá na \mathbf{A} , $\mathbf{p} = \mathbf{0}$ a $s = 0$ (pozri vysvetlenie nižšie). Ak zavedieme označenie $\mathbf{v}(c) = \{\mathbf{v}(c)_i; t_i\}$ pre ľubovoľnú vektorovú funkciu \mathbf{v} parametra c , môžeme pre \mathbf{p} zo (6.6) napísať $\mathbf{p} = \mathbf{p}(w_j)$ a pomer v v (6.7) bude vyjadrovať hľadanú hodnotu spektra $s(w_j)$ pre frekvenciu w_j

$$s(w_j) = \frac{\mathbf{f}^T \mathbf{p}(w_j)}{\mathbf{f}^T \mathbf{f}} \quad (6.8)$$

V konečnom dôsledku tak k výpočtu spektra MNŠ $s(w_j) = \{s_j, w_j\}$, $j=1,2,\dots,k$, treba k -krát zopakovať aproximačný postup MNŠ, zakaždým vypočítať $\mathbf{p}(w_j)$ pre odlišné frekvencie w_j .

Určenie spektra originálneho, úplného časového radu je jednou z hlavných oblastí aplikácie SA. Druhou je odstránenie známych zložiek (ale neznámej veľkosti) z časových radov a výpočet spektra reziduálneho radu. To je trochu komplikovanejšie, preto si poďme problém najprv definovať v jazyku funkcionálnej analýzy.

Funkcionálna analýza a spektrum MNŠ

Funkcionálna analýza je analýza funkcionálov. Funkcionál je skalárna funkcia vektorových veličín, nás budú zaujímať najmä tri funkcionály: skalárny súčin, norma a vzdialenosť. Na začiatok, definujme si niektoré konečné priestory.

Vektorový priestor L rozmeru $n = \dim(L)$ je priestor všetkých možných n -tíc (l_1, l_2, \dots, l_n) reálnych čísiel. Platí, že lineárna kombinácia ľubovoľných prvkov patrí do toho istého priestoru, t.j ak $\forall i: \mathbf{a}_i \in L$ a $a \in \mathbb{R}$ (skaláry), potom

$$\mathbf{b} = \sum_i a_i \mathbf{a}_i \quad (6.9)$$

je tiež z L .

Hilbertov priestor H je vektorový priestor, v ktorom je definovaný skalárny súčin dvoch vektorov. Ak $\mathbf{a}, \mathbf{b} \in H$, označme ich skalárny súčin ako

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b} = \sum_i a_i b_i, \quad (6.10)$$

definujeme normu (dĺžku, magnitúdu) vektora $\mathbf{a} \in H$

$$\|\mathbf{a}\| = (\langle \mathbf{a}, \mathbf{a} \rangle)^{1/2} \quad (6.11)$$

a vzdialenosť medzi $\mathbf{a}, \mathbf{b} \in H$

$$d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\| = [\langle (\mathbf{a} - \mathbf{b}), (\mathbf{a} - \mathbf{b}) \rangle]^{1/2}. \quad (6.12)$$

n -tica vektorov $\mathbf{a}_i \in L$ je lineárne nezávislá ak

$$\sum_{i=1}^n a_i \mathbf{a}_i = 0, \quad \forall i: a_i \in R \quad (6.13)$$

je splnená vtedy, a iba vtedy, keď $\forall i: a_i = 0$. To znamená, že žiadnu z \mathbf{a}_i nemožno vyjadriť ako lineárnu kombináciu ostatných. Potom množina všetkých vektorov

$$\mathbf{b}_j = \sum_i a_i \mathbf{a}_i \quad (6.14)$$

(vytvorená použitím všetkých možných kombinácií skalárnych hodnôt a_i) formuje oblasť (z angl. *manifold*) S priestoru L , a $\{\mathbf{a}_i, \forall i\}$ sa nazýva *báza*, ktorá generuje S . Počet n vektorov v $\{\mathbf{a}_i, \forall i\}$ je zároveň rozmerom oblasti S .

Dva vektory \mathbf{a}, \mathbf{b} sú ortogonálne, ak platí $\langle \mathbf{a}, \mathbf{b} \rangle = 0$. V dvojrozmernom priestore sa pretínajú pod pravým uhlom. Ak sa pretínajú pod všeobecným uhlom q , spustíme kolmicu napr. z \mathbf{a} na \mathbf{b} , ktorá na ňom vymedzí vektor \mathbf{y} . Vektor \mathbf{y} nazývame ortogonálnou projekciou vektora \mathbf{a} na \mathbf{b} . Keďže skalárny súčin v rovinnej reprezentácii je

$$\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\| \|\mathbf{b}\| \cos q, \quad (6.15)$$

z neho

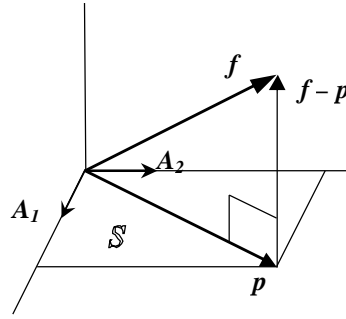
$$\|\mathbf{y}\| = \|\mathbf{a}\| \cos q = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{b}\|} \quad (6.16)$$

$$\mathbf{y} = \|\mathbf{y}\| \frac{\mathbf{b}}{\|\mathbf{b}\|} = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} \mathbf{b}$$

platia aj v celom Hilbertovom priestore.

Jednoduché projekčné pravidlo hovorí: Najmenšia vzdialenosť medzi bodom a rovinou je kolmica spustená z bodu na rovinu. Preformulujme ho do kontextu predošlých úvah nahradením niektorých výrazov. Vzdialenosťou rozumieme normu, bod reprezentuje (vektorový) element Hilbertového priestoru, a rovina zase oblasť S Hilbertovho priestoru. Majme teda $f \in H$ (bod) a $S \subset H$ (rovinu), potom zo všetkých elementov $s \in S$ existuje jeden $p \in S$ taký, že $d(f, p) \leq d(f, s)$ (najmenšia vzdialenosť). Tento element p je daný ortogonálnou projekciou f na S , t.j. $(f-p) \perp S$.

Jednoduchá ilustrácia geometrických vzťahov medzi f (časový rad), p (aproximujúca funkcia) a množinou $\{\mathbf{a}_i, \forall i\}$ generujúcou oblasť S v prípade trojrozmernej f a dvojrozmernej S je na obr. 6.1.



Obr. 6.1: Schéma projekčného teoremu.

Lubovolné $s \in S$ môžeme vyjadriť vzt'ahom

$$s = \sum_i x_i A_i, \quad (6.17)$$

potom elementu p bude náležať iba jedna konkrétna n -tica skalárov $\{\hat{x}_i, \forall i\}$ (pozri (6.4)). Podmienku $(f-p) \perp S \equiv (f-p) \perp A_j; \forall i$, ktorú podľa projekčného pravidla musíme splniť, zapíšme v tvare skalárneho súčinu

$$\forall j: \left\langle \left(f - \sum_i \hat{x}_i A_i \right), A_j \right\rangle = 0. \quad (6.18)$$

Po úprave

$$\sum_i \hat{x}_i \langle A_i, A_j \rangle = \langle f, A_j \rangle, \quad j = 1, 2, \dots, m \quad (6.19)$$

a prepísaní do maticového tvaru, keď

$$N = \begin{bmatrix} \langle A_1, A_1 \rangle & \langle A_2, A_1 \rangle & \mathbf{K} & \langle A_m, A_1 \rangle \\ \langle A_1, A_2 \rangle & \langle A_2, A_2 \rangle & \mathbf{K} & \langle A_m, A_2 \rangle \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ \langle A_1, A_m \rangle & \langle A_2, A_m \rangle & & \langle A_m, A_m \rangle \end{bmatrix} \quad u = \begin{bmatrix} \langle f, A_1 \rangle \\ \langle f, A_2 \rangle \\ \mathbf{M} \\ \langle f, A_m \rangle \end{bmatrix} \quad \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \mathbf{M} \\ \hat{x}_m \end{bmatrix},$$

dostávame známe *normálne rovnice* ($f-p$ je normálny (=kolmý) na S)

$$N\hat{x} = u. \quad (6.20)$$

Keďže $A = [A_1 \ A_2 \ \dots \ A_m]$,

$$N = A^T A$$

a $u = A^T f$,

z normálnych rovníc dostaneme odhad (neznámych) parametrov (porov. (6.1), ak $P=I$)

$$\hat{x} = N^{-1} u = \left(A^T A \right)^{-1} A^T f, \quad (6.21)$$

aproximujúcu funkciu

$$p = A\hat{x} \quad (6.22)$$

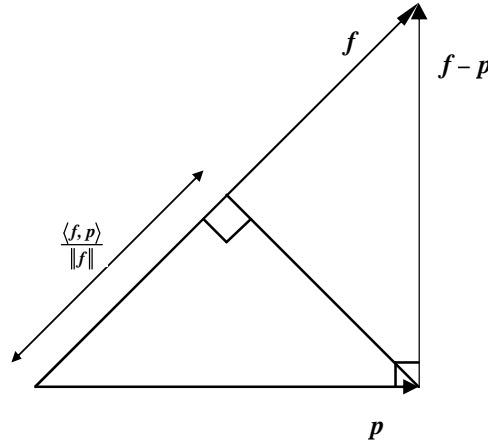
a vektor reziduí (v kontexte vyrovnávacieho počtu sú zápornými hodnotami opráv meraní)

$$v = f - p = f - A\hat{x} = f - A \left(A^T A \right)^{-1} A^T f \quad (6.23)$$

Poznamenajme, že v je kolmý na p ($v \perp p$) tak, ako na každý vektor oblasti S , teda i na stĺpcové vektory A , ktoré generujú S a ktorých lineárnou kombináciou je vektor p . Projekčný

teorém (obr. 6.1) teda rozkladá vektor f na dve kolmé zložky. Jedna je ortogonálnou projekciou f na S , druhá kolmicou „spustenou“ z f na S .

Teraz by sme sa mali prepracovať k niečomu, na čo neskôr môžeme odkázať hodnotu spektra. Vykonajme druhú, podstatne jednoduchšiu projekciu p späť na f (obr.6.2).



Obr. 6.2: Druhá projekcia v spektrálnej analýze.

Z (6.16), pomer dĺžky tejto ortogonálnej projekcie $\frac{\langle f, p \rangle}{\|f\|}$ k dĺžke f je rovný

$$\frac{\langle f, p \rangle}{\langle f, f \rangle} = \frac{f^T p}{f^T f} \quad (6.24)$$

a predstavuje mieru, akou sa p podiela na f , a tiež nám hovorí čosi o tom, koľko z f je „obsiahnuté“ v S . Čím f leží „bližšie“ k S , tým viac sa spomínaný pomer blíži k 1, naopak v prípade $f \perp S$ je rovný nule.

Aplikujme predošlé poznatky na SA. Pre každú spektrálnu frekvenciu w_j , $j=1,2,\dots,k$, je prostredníctvom $A = [\cos w_j t \quad \sin w_j t]$, definovaná iná oblasť S , preto aj ortogonálna projekcia $p(w_j)$ sa bude s frekvenciou meniť.

Pomer z (6.24), pre spomínané vlastnosti, bude hľadanou hodnotou spektra (resp. spektr.hodnotou) MNŠ funkcie f pre frekvenciu w_j (rovn. 6.8)

$$s(w_j) = \frac{f^T p(w_j)}{f^T f}.$$

Spektrum MNŠ funkcie f je množina spektrálnych hodnôt pre všetky (vybrané) frekvencie w_j ,

$$s(w) = \{s(w_j), j=1,2,\dots,k\}.$$

Týmto sme si definovali spektrum, hodnotu spektra a princípy SA pri spracovaní časových radov, ak neobsahujú žiadne rušivé zložky (šum), ktoré by nepriaznivo ovplyvnili spektrálny obraz. Ak sa takéto komponenty v rade nachádzajú, čo je v praxi bežné, ich eliminácia znamená zjavné nebezpečenstvo, pokiaľ nepoznáme ich veľkosť. Bez ohľadu na techniku, ktorú by sme použili, je nemožné určiť presné hodnoty x , pokiaľ nie je známe kompletne zloženie f . Ak by bolo, nemuseli by sme hľadať skryté periódy. Určovanie x však neprináša iba chybné hodnoty parametrov, môže v reziduálnom rade zanechať i falošné frekvencie. Na druhej strane, je rovnako škodlivé nechať tieto komponenty bez povšimnutia, keďže dokážu spektrum skresliť v porovnateľnej miere (Vaníček, 1971).

Z daných dôvodov je rozumné navrhnúť vzťah pre spektrum, ktorý by zahrnul vplyv systematického šumu bez nutnosti určenia \mathbf{x} . Toto užitočné rozšírenie SA dosiahneme relatívne ľahko a opäť s využitím jednoduchosti geometrickej interpretácie.

SA-MNŠ ak sú známe niektoré zložky časového radu

Časový rad zvyčajne obsahuje okrem signálu, ktorého spektrum nás zaujíma, aj náhodný a systematický šum. O vplyve náhodného sa zmienime v kap.7, systematický nás bude zaujímať práve teraz. Či už periodický alebo ten, spôsobujúci nestacionaritu radu, o oboch musíme mať aspoň akú-takú informáciu, aby sme ich vplyv mohli modelovať. Potom všetky zložky, podielajúce sa na systematickom šume, ktorý chceme z SA vylúčiť, nazvime známe komponenty a každý bude reprezentovaný svojou báзовou funkciou A_i .

Uvažujme najbežnejší prípad, že poznáme báзовé funkcie $A_i(t)$ šumu, ale nepoznáme veľkosť, akou sa na zložení časového radu podielajú, teda

$$f(t) = \sum_{i=1}^u x_i A_i(t), \quad (6.25)$$

kde práve x_i sú neznáme parametre zodpovedné za rozmer modelovanej zložky, u je počet známych komponentov. Rozdelme A na známe komponenty \hat{A} a spektrálne funkcie $\cos w_j t$, $\sin w_j t$:

$$A = \begin{bmatrix} \hat{A}_1 & \hat{A}_2 & \mathbf{K} & \hat{A}_u & \cos w_j t & \sin w_j t \end{bmatrix}. \quad (6.26)$$

Dekompozíciou $f(t)$ dostaneme jej ortogonálnu projekciu $\hat{p} \in \{\hat{A}\mathbf{x}\}$ a reziduálny časový rad

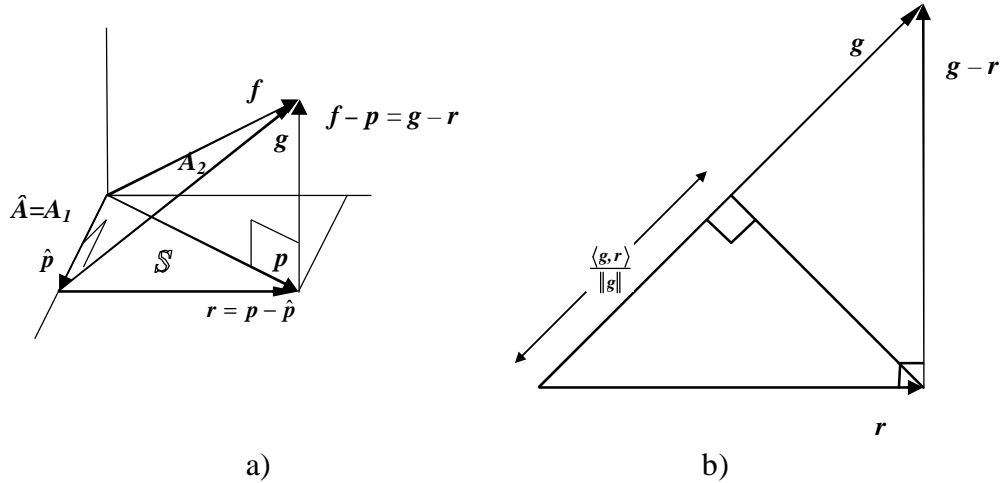
$$g(t) = f(t) - \hat{p}(t), \quad (6.27)$$

ktorého spektrum ideme počítať. Následnou kolmou projekciou g na oblasť (manifold) $S(A)$ dostávame projekciu $r = p - \hat{p}$ a reziduá $v = g - r = f - p$. Všimnime si, že, či už premietame f na $S(A)$ priamo, alebo najprv na $S(\hat{A})$ a potom $f - \hat{p}$ na $S(A)$, dostaneme rovnaké reziduá. Nakoniec, našu spektrálnu hodnotu získame kolmou projekciou r na g ako pomer dĺžky tejto projekcie ku dĺžke g .

Zhrňme doterajšie poznatky:

<u>technika</u>	\approx	<u>projekcia</u>	\rightarrow	<u>výsledok</u>
a) Klasické vyrovnanie MNŠ (aproximácia aj odhad parametrov)		f na $S(A)$		p, \hat{x}
b) SA-MNŠ bez známych komponentov		1. f na $S(A)$ 2. p na f		p $s(w)$
c) SA-MNŠ so známymi komponentami		1. f na $S(\hat{A})$ 2. g na $S(A)$ 3. r na g		g r $s(w)$

Aby sme si uvedené ortogonálne projekcie mohli názorne ilustrovať (obr.6.3), musíme (zjavne nerealisticky) pre A vyhradiť iba dva rozmery, pričom nech $\hat{A}=A_1$ je jednorozmerné a rovnako tak reprezentant spektrálnych funkcií A_2 . Ak sa cez toto obmedzenie dokážeme preniesť, na obr. 6.3a) môžeme pekne vidieť prvú a druhú projekciu z prípadu c), obr.6.3b) dáva prehľad o druhej a tretej projekcii.



Obr.6.3: Geometrická interpretácia SA-MNŠ so známymi komponentmi

Analogicky ku (6.24), spektrálna hodnota bude

$$s(w_j) = \frac{\mathbf{g}^T \mathbf{r}(w_j)}{\mathbf{g}^T \mathbf{g}}, \quad (6.28)$$

kde $\mathbf{g} = \mathbf{f} - \hat{\mathbf{p}}$

a $\mathbf{r} = \mathbf{p} - \hat{\mathbf{p}}$.

Zo vzťahu (6.23)

$$\mathbf{g} = \mathbf{f} - \hat{\mathbf{A}}(\hat{\mathbf{A}}^T \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^T \mathbf{f} \quad (6.29)$$

a jej kvadratická norma

$$\mathbf{g}^T \mathbf{g} = \mathbf{f}^T (\mathbf{I} - \hat{\mathbf{A}}(\hat{\mathbf{A}}^T \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^T) \mathbf{f}$$

Pretože (w_j) je projekciou \mathbf{g} na $S(\hat{\mathbf{A}})$

$$\mathbf{r} = \mathbf{A}^T (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{g} \quad (6.30)$$

bude čitateľ v (6.28) rovný

$$\mathbf{g}^T \mathbf{r} = \mathbf{g}^T \mathbf{A}^T (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{g}. \quad (6.31)$$

$\mathbf{g}^T \hat{\mathbf{A}} = \mathbf{g}^T [\hat{\mathbf{A}} \quad \cos w_j t \quad \sin w_j t]$, zároveň však vieme, že \mathbf{g} je kolmé na $S(\hat{\mathbf{A}})$, t.j. $\mathbf{g}^T \hat{\mathbf{A}} = \mathbf{0}$, preto $\mathbf{g}^T \hat{\mathbf{A}} = [\mathbf{0} \quad \mathbf{0} \quad \dots \quad \mathbf{0} \quad \mathbf{g}^T \cos w_j t \quad \mathbf{g}^T \sin w_j t]$, a z toho dôvodu stačí z matice $\mathbf{N} = (\mathbf{A}^T \mathbf{A})^{-1}$ počítať iba posledný 2×2 diagonálny blok.

Podrobnosti o ďalšom postupe, ktorým sa získa výpočtovo menej náročná formula pre hodnotu spektra, sú v (Wells, 1985, str.19-20) alebo (Vaníček, 1971, str.17-19). Vychádza zo štruktúry matice $\mathbf{A}^T \mathbf{A}$

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} \hat{\mathbf{K}} & \mathbf{a} & \mathbf{b} \\ \mathbf{a}^T & CC & CS \\ \mathbf{b}^T & CS & SS \end{bmatrix} \approx \begin{bmatrix} u \times u & u \times 1 & u \times 1 \\ 1 \times u & 1 \times 1 & 1 \times 1 \\ 1 \times u & 1 \times 1 & 1 \times 1 \end{bmatrix},$$

kde u je počet známych komponentov, ďalej

$$\begin{aligned} \hat{\mathbf{K}} &= \hat{\mathbf{A}}^T \hat{\mathbf{A}} & \mathbf{a} &= \mathbf{A}_j^T \mathbf{A}_{u+1} & \mathbf{b} &= \mathbf{A}_j^T \mathbf{A}_{u+2} \\ CC &= \mathbf{A}_{u+1}^T \mathbf{A}_{u+1} & CS &= \mathbf{A}_{u+1}^T \mathbf{A}_{u+2} & SS &= \mathbf{A}_{u+2}^T \mathbf{A}_{u+2} \end{aligned}$$

a $j=1,2,\dots,u$. Posledný 2×2 diagonálny blok matice N normálnych rovníc je nakoniec

$$\frac{1}{DET} \begin{bmatrix} SS - V & -(CS - W) \\ -(CS - W) & CC - U \end{bmatrix},$$

kde

$$U = \mathbf{a}^T \hat{\mathbf{K}}^{-1} \mathbf{a} \quad W = \mathbf{a}^T \hat{\mathbf{K}}^{-1} \mathbf{b} \quad V = \mathbf{b}^T \hat{\mathbf{K}}^{-1} \mathbf{b}$$

$$DET = (CC - U)(SS - V) - (CS - W)^2.$$

Označme ešte

$$\mathbf{x} = \mathbf{g}^T \mathbf{A}_{u+1} \quad \mathbf{h} = \mathbf{g}^T \mathbf{A}_{u+2}$$

potom algoritmus výpočtu spektra funkcie f je

$$s(\omega) = \frac{(SS - V)\mathbf{x}^2 - 2(CS - W)\mathbf{x}\mathbf{h} + (CC - U)\mathbf{h}^2}{(\mathbf{g}^T \mathbf{g})DET} \quad (6.32)$$

Ak v časovom rade dopredu žiadne známe komponenty neuvažujeme, $U=V=W=0$ a $\mathbf{g}=\mathbf{f}$, ak je navyše konštantne delený a symetrický voči počiatku, $CS=0$ a spektrum bude $s(\omega) = (1/f^T f)[(\mathbf{x}^2/CC) + (\mathbf{h}^2/SS)]$.

Praktický postup pri analýze časových radov so známymi komponentmi ZK, ale i bez nich, spočíva v iteratívnom výpočte spektra. To znamená, že v prvom kroku (pri uvážení všetkých ZK) sa vypočíta (6.32) spektrálny obraz, ktorý pozostáva z „vrcholov a údolí“. Najväčšie vrcholy odhaľujú frekvencie (periodických) signálov s najväčším podielom na zložení časového radu. Akonáhle je takýto signál odhalený, z hľadiska ďalšieho spracovania je zaradený do systematického šumu, a v ďalšom kroku figuruje medzi ZK. Takýmto spôsobom možno nájsť aj signály, ktoré sú akoby v tieni významnejších „kolegov“, čo sa prejavuje najmä pri blízkych frekvenciách, alebo zvýrazňovaním nižších frekvencií. Viac sa dozvieme v kap.8.

7. Štatistické testovanie významnosti spektrálnych vrcholov

Doteraz sme si ukázali, ako určiť spektrum, ale bez toho, aby sme vedeli i posúdiť významnosť jeho vrcholov. Toto je veľmi dôležité, pretože iba štatisticky významné vrcholy má zmysel v jednotlivých iteráciách analýzy potlačiť, a tak z časového radu vyčistiť frekvencie reálneho významu. V tejto kapitole si preto povieme, akým štatistickým testom posúdime významnosť spektrálnych hodnôt vo vrcholoch. Podrobnosti možno získať v (Pagiatakis, 1999) alebo (Steeves, 1981)

Prvý krát do úvah zahmieme kovariančnú maticu \mathbf{C}_f časového radu \mathbf{f} , potom vzťah pre výpočet hodnoty spektra bude

$$s(\omega_j) = \frac{\mathbf{f}^T \mathbf{C}_f^{-1} \mathbf{p}(\omega_j)}{\mathbf{f}^T \mathbf{C}_f^{-1} \mathbf{f}}, \quad j = 1, 2, \dots, k. \quad (7.1)$$

Označme reziduá

$$\mathbf{r} = \mathbf{f} - \mathbf{p} = (\mathbf{I} - \mathbf{J})\mathbf{f}, \quad (7.2)$$

kde \mathbf{I} je jednotková matica a $\mathbf{J} = \mathbf{A}(\mathbf{A}^T \mathbf{C}_f^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_f^{-1}$. Po viacerých úpravách dostávame vzťah pre spektrum

$$s = \left(1 + \frac{\mathbf{f}^T \mathbf{C}_f^{-1} (\mathbf{I} - \mathbf{J}) \mathbf{f}}{\mathbf{f}^T \mathbf{C}_f^{-1} \mathbf{J} \mathbf{f}} \right)^{-1} = \left(1 + \frac{Q_n}{Q_s} \right)^{-1} \quad (7.3)$$

ako funkciu stochasticky nezávislých kvadratických foriem Q_n a Q_s . Q_s reprezentuje vplyv signálu ($Q_s=f^T C_f^{-1} p$), na druhej strane Q_n vplyv šumu (noise), ktorý (a to je dočasný predpoklad) v f tvorí zvyšok ($Q_n=f^T C_f^{-1} r$). Zároveň treba poznamenať, že pri SA uvažujeme iba náhodný a prípadne periodický šum, ktorý čaká na odhalenie, a známe komponenty systematického šumu sú zahnuté v A . Signálom tak označujeme iba tie periodické zložky, ktoré práve detekujeme.

Čo je dôležité, pomer Q_s/Q_n (S/N) nám umožní vypočítať hustotu výkonového spektra HVS v decibeloch (dB), ak vyjadríme S/N ako funkciu spektra s , zlogaritmujeme a vynásobíme desiatimi:

$$HVS = 10 \log_{10} \left(\frac{s}{1-s} \right) \quad (7.4)$$

V (Steeves, 1981) je významnosť vrcholov odvodená na základe predpokladu, že časový rad obsahuje nekorelovaný (biely) Gaussov šum. V našom ďalšom texte budeme podľa (Pagiatakis, 1999) predpokladať, že časový rad f vznikol zastúpením náhodných premenných s viacrozmerým normálnym rozdelením pravdepodobnosti. Odvodenie funkcie rozdelenia pravdepodobnosti FRP spektra MNŠ pritom nie je limitované iba na trigonometrické bázové funkcie a tak nasledujúce štatistické testy môžu byť aplikované i na lineárnu regresnú analýzu. Vo všeobecnosti, počítame s plnou vstupnou kovariančnou maticou C_f a dvoma možnými prípadmi: Jej „mierka“ alebo variančný faktor S_0^2 je buď vopred známy, alebo je vopred neznámy, a potom sa ráta jeho odhad \hat{S}_0^2 klasickým spôsobom z reziduí.

Zoberme si teraz prvý prípad, ak je variančný faktor známy.

Majme rad f s viacrozmerým normálnym rozdelením tvoriacich náhodných premenných, a každé z Q_s a Q_n má chí-kvadrát rozdelenie (χ^2) s m a n stupňami voľnosti. Keďže sú štatisticky nezávislé, ich pomer má Snedecorovo-Fisherovo rozdelenie pravdepodobnosti (F).

Predstavme si, že sme po určitom počte pokusov s rôznymi bázovými funkciami stanovili typický priebeh časového radu a formu systematického šumu, a tak definovali sústavu bázových vektorov o rozmere u . Teraz hľadáme spektrálny obsah zvyšku časového radu postupným aplikovaním bázových funkcií $[\cos w_i t \quad \cos w_i t]$ pre všetky frekvencie z oblasti záujmu W . Kvadratická forma reziduí Q_n bude mať $n=n-u-2$ stupňov voľnosti (sv), kde „2“ označuje ďalšie dve neznáme pre sinusové a kosínusové koeficienty periodickej zložky (signálu). m je dĺžka časového radu f . Ak $Q=Q_s+Q_n=f^T C_f^{-1} f$ a Q_s , Q_n sú štatisticky nezávislé, potom $sv(Q)=sv(Q_s)+sv(Q_n)=n-u$. Preto $m=2$, $n=n-u-2$, a $(2/n)Q_n/Q_s \sim F_{n,2}$.

Je potrebné vysloviť nulovú hypotézu H_0 : *Vstupný časový rad má mnohorozmerné normálne rozdelenie $n(\theta, C_f)$* . H_0 tvrdí, že ak $Q_n/Q_s \geq (2/n)F_{n,2,a}$, časový rad obsahuje štatisticky bezvýznamný šum. Alternatívna hypotéza je H_1 : $Q_n/Q_s < (2/n)F_{n,2,a}$. To, že sa použije dolný koniec F dáva zmysel, vysoký pomer Q_n/Q_s (horný koniec) implikuje nižšiu významnosť signálu voči šumu, čo ho robí neodhaliteľným. Vzhľadom k (7.3) štatisticky významné spektrálne vrcholy musia spĺňať nerovnosť

$$s(w_i) \geq \left(1 + \frac{n}{2} F_{n,2,a} \right)^{-1}, \quad (7.5)$$

z ktorej vidieť, že spektrum MNŠ má β (beta) rozdelenie pravdepodobnosti s parametrami $g=1$ a $d=(n-u-2)/2$. Stredná hodnota β -rozdelenia tak je $m=g/(g+d) = 2/(n-u)$, čo je očakávaná spektrálna hodnota šumu (identická s tou, ktorú pre biely šum odvodil Vaníček(1971)). Môžeme sa domnievať, že šum je nekorelovaný. Reakcia spektra MNŠ je konštantná.

Aby sme neboli pri počítaní kritickej hodnoty odkázaní na štatistické tabuľky, je užitočné použiť rovnosť

$$F_{n,2,a} = F_{2,n,1-a}^{-1} = \frac{2}{n} (a^{-2/n} - 1)^{-1}, \quad (7.6)$$

ktorá nám dosadením do (7.5) pri hladine významnosti α dáva *kritickú hodnotu*

$$c_a = \left(1 + \frac{1}{a^{-2/n} - 1} \right)^{-1}, \quad n = n - u - 2 \quad (7.7)$$

(u je počet odhadovaných neznámych parametrov známych komponentov). Zapíšme ešte tzv. rozhodovaciu funkciu

$$s(w_i) \begin{cases} \leq c_a \forall \omega_i \in \Omega; H_0 \text{ nezamietame, t.j. } f \text{ obsahuje štatisticky bezvýznamný šum} \\ > c_a \text{ pre aspoň jedno } \omega_i \in \Omega; H_0 \text{ zamietame, } f \text{ obsahuje štat. významnú zložku}(y) \end{cases}$$

Často však v experimentálnych vedách variančný faktor σ_0^2 nie je známy, prvky kovariančnej matice teda nemajú rozmer. Musíme použiť odhadnutú kovariančnú maticu, danú súčinom inverzie matice normálnych rovníc a a posteriorneho variančného faktora \hat{S}_0^2 . Keďže \hat{S}_0^2 je iba odhadom (náhodnou premennou) a $f \sim n(\mathbf{0}, \hat{C}_f = \hat{S}_0^2 N^-)$ (N^- je G-inverzia, N nemusí byť nutne regulárna), Q_n ani Q_s nebudú mať χ^2 -rozdelenie. Po sérii úvah prideme k nasledovnému:

- $p = Q_n / Q_s$ má hypergeometrické rozdelenie pravdepodobnosti dané funkciou

$$h_n(p) = \frac{\Gamma(n)}{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2})} p^{(n-2)/2} \int_{-\infty}^{\infty} \frac{p^{n/2}}{(1+pq)^n (1+\frac{2}{n}q)^{(n+2)/2}} dq, \quad \text{kde tzv. gamafunkcia } \Gamma \text{ je}$$

$$\text{definovaná vzťahom (Kubáček a Pázman, 1979) } \Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx, \quad n > 0.$$

- kritická hodnota percenta variancie c_a , ponad ktorú sú spektrálne vrcholy významné, sa určí z

$$c_a = \left(1 + \frac{n}{2} H_{n,a} \right)^{-1}, \quad n = n - u - 2 \quad (7.8)$$

kde $H_{v,\alpha}$ je pravdepodobnostný integrál z $h_v(p)$;

- c_α konverguje ku kritickej hodnote určenej z (7.7), keď rad obsahuje viac ako 150 bodov;
- Pre vyhodnotenie testu použijeme predošlú rozhodovaciu funkciu.

Vidíme, že rozdiel medzi uvedenými prípadmi, S_0^2 známe alebo neznáme, je výrazný pre kratšie rady (do 150 bodov), tým viac, ak zmenšujeme hladinu významnosti. Je dobré uvedomiť si z toho vyplývajúci fakt, že ak nepoznáme presnú informáciu o presnosti vstupných údajov, úrovně významnosti ležia vyššie a my musíme striktnejšie posudzovať a identifikovať signifikantné vrcholy.

8. Praktická aplikácia

Spektrálna analýza metódou najmenších čtvorcov je nepochybne silný nástroj na odhaľovanie skrytých periodických procesov, ktoré pôsobia cez merané veličiny. Matematický aparát SA, ako bolo vidieť z predošlých kapitol, však vôbec nie je jednoduchý z hľadiska množstva operácií, ktorých vykonanie si vyžaduje. Náročnosť najviac súvisí s rozsahom časových radov a ich delení, potom aj s výberom rozsahu frekvenčnej oblasti, počtom a matematickou

formou odhadovaných procesov (známych komponentov). Tu klasické kalkulačky a vyhodnocovatelia strácajú zmysel, a všetky automatizovateľné procedúry je vhodné, ba nevyhnutné implementovať do programových algoritmov počítačov. Jednou, a zdá sa i jedinou implementáciou SA-MNŠ svojho druhu (a rozsahu) v súčasnosti je program LSSA (z angl. Least Square Spectral Analysis), ktorý bol vyvinutý na kanadskej akademickej pôde pre výskumné účely v oblasti geodynamiky a ďalších geofyzikálnych vied. Primárne existuje vo forme zdrojového kódu jazyka Fortran, pre naše experimentálne účely bol skompilovaný na vykonateľný súbor LSSA.exe pre platformu MS DOS/Windows. Jeho komplexnosť presahuje možnosti tejto práce popísať všetky matematické postupy a špecifikácie, ku ktorým SA vo svojom vývoji dospela a ktoré boli zavedené do programu. Pozrieme sa naň ale z užívateľského hľadiska. Aké informácie a v akej forme dokáže spracovať, a aké poskytnúť. Nakoniec pomocou LSSA zrealizujeme analýzu niekoľkých časových radov, empirických aj analytických, tak aby boli možnosti jeho využitia v geodetickej praxi názornejšie.

8.1 Popis vstupných a výstupných súborov programu LSSA

Vstupné súbory

Program vyžaduje prítomnosť dvoch ďalších súborov v tom istom adresári: lssa.in a samotný časový rad.

LSSA.IN

Obsahuje všetky užívateľom definované parametre pre analýzu časového radu. Sú rozdelené do 8 blokov, prvý riadok každého bloku popisuje jeho význam. Vstup parametrov je citlivý na dodržanie prideleného formátu, preto je nevyhnutné všimnúť si údaj o formátovej špecifikácii (I-integer-prirodzené č., F-float-reálne č., A-alfanumerické znaky) v každom relevantnom riadku. Dĺžka sekvencie znakov je počítaná od začiatku riadku, aj medzera je znak. Nasleduje detailný popis súboru, po blokoch. V rámečkoch je uvedený originálny obsah (testovacieho) súboru.

Blok 1

INPUT FILE CHARACTERISTICS	
TEST-SER.DAT	NAME OF INPUT FILE FMT:A12
64	NUMBER OF DATA POINTS (I6)
Hour	UNITS OF TIME (EX. HOUR, MIN, SEC, etc.) (FMT: A4)
Volt	UNITS OF TIME SERIES (EX. mGal, m, mbar, etc)(FMT: A4)

Údaje o vstupnom súbore, ktorý obsahuje časový rad.

Riadok 2: Názov súboru vo formáte A12, t.j. 8 znakov pre názov a 3 pre príponu

Riadok 3: Dĺžka radu (počet časových bodov): 6 ciferné prirodzené číslo.

Riadok 4: Jednotka času v rade, napr. rok, hodina, minúta či sekunda: 4znaky.

Riadok 5: Jednotka hodnôt radu, napr. mm, mGal, °C a pod: 4 znaky. Názvy obidvoch jednotiek figurujú vo výstupnom súbore.

Blok 2

RANDOM CONSTANT (DATUM SHIFTS)	
4	NUMBER OF DATUM SHIFTS (I6)
4.0000	DATUM BIASES (TIMES) (F20.12)
13.0000	
56.0000	
128.0000	

Náhodné skoky (zdvihy, z angl. datum shifts, alebo offsets) údajov. Sú konštantné pre určité časové intervaly, ktoré užívateľ definuje ich začiatkom. Je vhodné uviesť všetky intervaly, pri ktorých je podozrenie na zmenu tohto konštantného člena, pretože dokáže výrazne skresliť spektrálny obraz aj parametre ostatných trendov.

- R2: Počet predpokladaných skokov, minimálne však jeden, ktorý zodpovedá začiatku radu.
 R3,4...: Časové okamihy, kedy došlo k skoku. Prvý ofset je vždy definovaný v prvom okamihu časového radu. Ak je zadaných viac hodnôt ako deklarovaných v druhom riadku, prebytočné budú ignorované. Formát: reálne číslo, 7 cifier pred a 12 za desatinnou čiarkou.

Blok 3

TRENDS	
1	RANDOM RAMP (LINEAR TREND) (1=ON, 0=OFF) (I6)
0	EXPONENTIAL TREND ($a \cdot \exp(-b \cdot t)$) (1=ON, 0=OFF)
1.000000	DECAY CONSTANT (=b above) (F15.6)
0	POLYNOMIAL SECOND ORDER TERM (1=ON, 0=OFF) (I6)
0	POLYNOMIAL THIRD ORDER TERM (1=ON, 0=OFF) (I6)
0	POLYNOMIAL FORTH ORDER TERM (1=ON, 0=OFF) (I6)
0	POLYNOMIAL FIFTH ORDER TERM (1=ON, 0=OFF) (I6)

Všeobecný tvar polynómu, ktorý predstavuje trend (max. do 5 rádu), je daný vzťahom $y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$. Kombináciu členov možno dosiahnuť zapnutím (vypnutím) výpočtu koeficientov v príslušných riadkoch.

Trend.

- R2: Prepínač rozhodujúci o tom, či bude časový rad aproximovaný polynómom 1 rádu (lineárny trend, priamka). 1 znamená zapnuté, 0 vypnuté. Ak je zapnutý, odhaduje sa koeficient a_1 , predstavujúci sklon priamky.
 R3: Prepínač exponenciálneho trendu: $y = a \cdot \exp(-b \cdot t)$. Koeficient b sa zadáva v R4.
 R4: Reálna konštanta b v exponenciálnej funkcii (R3).
 R5-8: Prepínač zavedenia členov ďalších rádov polynómu do vyrovnania MNŠ. Výpočet koeficientov a_2, a_3, a_4 , resp. a_5 .

Blok 4

PROCESSES		ALL FORMATS:I6
0	FIRST ORDER TERM OF AR OR RANDOM WALK (1=ON, 0=OFF)	
0	SECOND ORDER TERM OF AR (1=ON, 0=OFF)	
0	THIRD ORDER TERM OF AR (1=ON, 0=OFF)	
0	FORTH ORDER TERM OF AR (1=ON, 0=OFF)	
0	FIFTH ORDER TERM OF AR (1=ON, 0=OFF)	

Všeobecný tvar autoregresného procesu AR(p) je $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t$, kde ϕ sú parametre náhodného procesu a ϵ_t biely šum.

Náhodné procesy. Opäť: 1-zapnutý výpočet, 0-vypnutý.

- R2: AR(1). Vykoná sa výpočet koeficientu a štatistický test významnosti tohto náhodného procesu. Navyše sa otestuje, či je koeficient významne odlišný od 1. Ak je rovný jednej, ide o tzv. „náhodnú prechádzku“ (z angl. random walk).
 R3-6: Zavedenie druhého až piateho člena AR.

Blok 5

PERIODIC SIGNALS

9	NUMBER OF FORCED FREQUENCIES (I6)
28.1783	NAME 1 FORCED PERIODS AND NAMES (F20.12,1X,A7)
14.0557	NAME 2
10.0760	NAME 3
8.1954	NAME 4
7.5816	NAME 5

Periodické signály. Podobne ako trend, aj určité periodické komponenty môžu byť v rade potlačené, aby tak vynikli slabšie, ale ešte štatisticky významné spektrálne vrcholy. Výsledkom vyrovnaní bude ich amplitúda a fáza, spolu so štatistickým testom ich významnosti.

R2: Počet známych periodických komponentov.

R3,4...:Periódá prvej a ďalších komponentov v časovej jednotke radu, spolu s orientačným názvom. Ak je zadaných viac hodnôt ako deklarovaných v druhom riadku, prebytočné budú ignorované. Ak však menej, program sa predčasne ukončí.

Blok 6

CHARACTERISTICS OF SERIES	
1	A-PRIORI VAR. FACTOR (0 = UNKNOWN, 1 KNOWN) (I6)
0	WEIGHTS = 1 (1=YES, 0=USE GIVEN WEIGHTS) (I6)
1.0000000	SCALE FACTOR FOR THE STANDARD DEVIATIONS (F15.7)

Charakteristika radu. Doplní informáciu o kovariančnej matici, zadanej výhradne v diagonálnej forme vo vstupnom súbore hodnôt radu.

R2: 1, ak je a-priori variančný factor známy.

R3: 1, ak chceme váhy meraní považovať za rovnaké a rovné jednej. V opačnom prípade (0) sa kovariančná matica vypočíta inverziou variancií hodnôt časového radu. Štandardné odchýlky musia byť uvedené v treťom stĺpci vstupného súboru.

R4: Používateľ tu môže špecifikovať pomerný rozmer – variančný faktor meraní.

Blok 7

SPECTRUM CHARACTERISTICS	
2	IB: NUMBER OF SPECTRAL BANDS (I6)
250	NW: NUMBER OF SPECTRAL VALUES IN BAND (I6)
64.0000	2.0000 PL, PS: LARGEST, SMALEST PERIOD IN BAND
18.0000	15.0000

Charakteristika počtu a rozsahu frekvenčných pásiem, ktoré sú v záujme používateľa, spolu s jednotným počtom spektrálnych hodnôt v každom pásme.

R2: Počet spektrálnych pásiem (max. 50).

R3: Počet spektrálnych hodnôt v každom pásme(max. 5000). Čím väčšia hodnota, tým hladší je výsledný spektrálny obraz, no väčšie nároky na výpočet. Optimum sa stanoví experimentálne. Formát: F12.4,1X,F12.4 .

R4,5...:Najväčšia a najmenšia periódá intervalu (pásma) v časovej jednotke radu.

Blok 8

STATISTICS	
0.01	CRITICAL LEVEL FOR DETECTING SIGNIFICANT PEAKS (F10.2)
0.01	LEVEL OF SIGNIFICANCE FOR STATISTICAL TESTING (F10.2)

Štatistické parametre. Kritické hladiny (hl. významnosti) pre štatistické testovanie, zvyčajne $\alpha=0.05$, alebo $\alpha=0.01$.

R2: Hladina významnosti pre odhalenie významných vrcholov v spektre.

R3: Hladina významnosti pre testovanie významnosti odhadovaných parametrov.

Vstupný súbor observácií

Je to súbor obsahujúci analyzovaný časový rad, ktorého názov, dĺžka a údaje o fyzikálnych jednotkách sú uvedené v prvom bloku súboru `lssa.in`.

Prvý riadok je vyhradený pre názov projektu (max. 40 riadkov), nanajvýš musí zostať prázdny. Časový rad začína druhým riadkom. Prvý stĺpec je **čas**, druhý **hodnota** radu a tretí **štandardná odchýlka** hodnoty radu. Štandardná odchýlka je v rovnakých jednotkách ako hodnota radu, a ak nie je známa, možno ju nahradiť jedničkou alebo akýmkoľvek zmysluplným číslom.

- Formát vstupných údajov nie je obmedzený ako v `lssa.in`.
- Je možné spracovať až 100 000 hodnôt, samozrejme v závislosti na počítačovom systéme.
- Časový rad môže a nemusí byť ekvidištančne krokovaný, bez potreby uvedenia tejto skutočnosti.
- Rad musí byť usporiadaný chronologicky, prográdne alebo retrográdne, a nesmie obsahovať dva rovnaké okamihy. Porušenie tejto podmienky povedie k ukončeniu programu a chybovému hláseniu v súbore `lssa.out`.

Výstupné súbory

Program po úspešnom výpočte vytvorí hlavný výstupný súbor, `lssa.out`, a tri doplnkové, `spectrum.dat`, `histogram.dat` a `rezidual.dat`.

LSSA.OUT

V tomto súbore sú výsledky analýzy spolu so štatistickými testami. Spektrum je vyjadrené v troch rozdielnych tvaroch: Percento variancie (spektrum MNŠ), hustota výkonového spektra v dB (decibel), a hustota výkonového spektra v_j^2/f , kde j je jednotka hodnôt radu, f (frekvencia) predstavuje počet cyklov za časovú jednotku ($f=1/T$, T je perióda).

Spektrum je zapísané v 6 stĺpcoch nasledovne:

1. *Perióda*.
2. *Frekvencia* (prevrátená hodnota periódy).
3. *Spolahľivosť (Fidelity)*: Je udávaná v tej istej časovej jednotke ako okamihy v rade, a indikuje polohu významného vrcholu. Jej hodnota prezrádza, či tento vrchol môže byť rušený (z angl. resolved from) vedľajšími významnými vrcholmi. Napríklad, ak vedľajší vrchol leží v rozsahu \pm fidelity, potom vrchol nemôže byť rozptýlený. V opačnom prípade je rozptýliteľný. Kritérium: Dva vrcholy sa môžu vyrušiť v danom časovom rade, ak v jeho priebehu dospejú k rozdielu fáz 180° .
4. *Spektrum MNŠ v percentách variancie (%)*.
5. *Hustota výkonového spektra (power spectral density) v dB*. Je ekvivalentná tej z FFT, ak sú časové body rovnomerne rozložené.
6. *Hustota výkonového spektra v jednotkách štvorcových na frekvenciu*. Opäť, táto spektrálna hustota (normalizovaná frekvenciou) je ekvivalentná tej z FFT, keď sú časové body ekvidištančne rozmiestnené.
7. Nakoniec je spektrum znázornené hviezdikami pre rýchlu identifikáciu vrcholov.

Nasledujúce súbory sú vhodné na jednoduché grafické znázornenie.

RESIDUAL.DAT

ASCII súbor so 4 stĺpcami:

- a. Čas.
- b. Vstupný časový rad.
- c. Normalizovaný reziduálny časový rad (po odstránení trendov, periód, náhodných procesov a pod.).
- d. Štandardná odchýlka reziduí.

SPECTRUM.DAT

ASCII súbor obsahujúci prvých 6 stĺpcov spektra zo súboru `lssa.out`, v rovnakom poradí.

HISTOGRAM.DAT

Histogram normalizovaných reziduí.

8.2 Definícia a vyhodnotenie experimentu

V tejto kapitole sa pozrieme na praktickú aplikáciu SA-MNŠ. Ako zdroj meraných údajov mi poslúžili časové rady dvojnásobných denných observácií polohy permanentných staníc GPS. Poloha permanentnej stanice sa primárne získava v pravouhlých geocentrických súradniciach, tie sa však pre ľahšiu interpretáciu transformujú do lokálneho horizontálneho systému (n, e, v ~ smer osí: severný, východný, vertikálny). V mojej práci som analyzoval zmeny polohy v jednotlivých súradniciach n, e, v (samostatne) a to permanentných staníc Hafelekar (HFLK) a Boroviec (BOR1), ktoré sú spolu s ďalšími asi tridsiatimi stanicami zapojené do projektu CERGOP (The Central Europe Regional Geodynamics Project) pre dlhodobé GPS monitorovanie tektonických procesov v regióne Strednej Európy, a tvoria sieť, ktorá je vzťahnutá k strednej hodnote polôh vybraných staníc a umožňuje sledovať i veľmi jemné relatívne variácie súradníc. Analýzou týchto zmien možno odhaliť nielen lineárne zmeny polohy (rýchlosť), ale aj pôsobenie množstva periodických javov. To je oblasť, v ktorej sa môže SA-MNŠ s výhodou aplikovať.

Na analýzu časových radov som použil program LSSA, pretože je veľmi prepracovaný a efektívne aplikuje matematický aparát SA-MNŠ. Pred samotným spracovaním GPS meraní som vykonal niekoľko testov na analytických časových radoch s vopred presne definovanými parametrami tvoriacich komponentov.

Syntetické časové rady

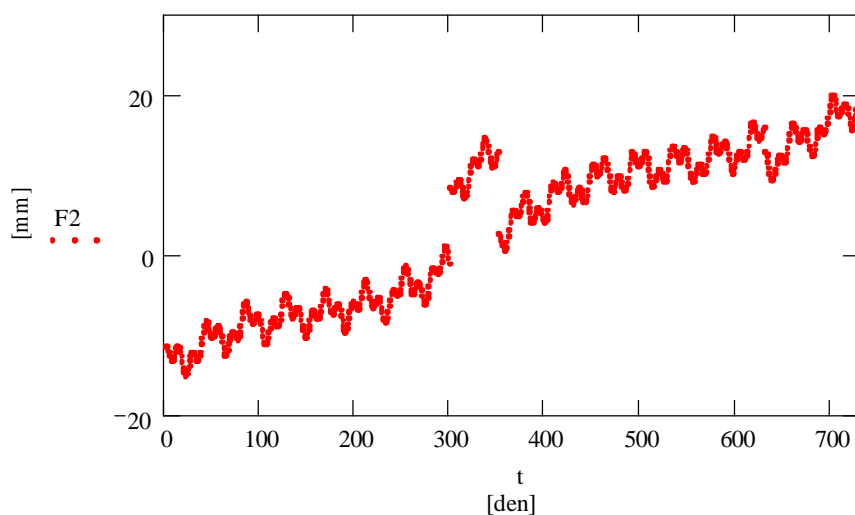
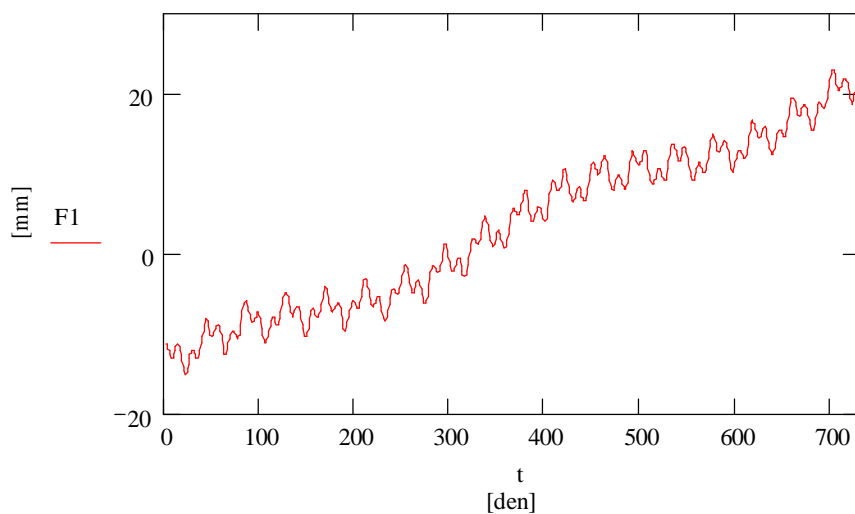
Pri definovaní skladby umelých testovacích funkcií som vychádzal zo skladby skutočných empirických funkcií. Zvolil som 4 najčastejšie zložky: dátový skok, lineárny trend, periodickú zložku a náhodný šum, a ich kombináciou vytvoril 5 syntetických funkcií podľa tabuľky ?1.

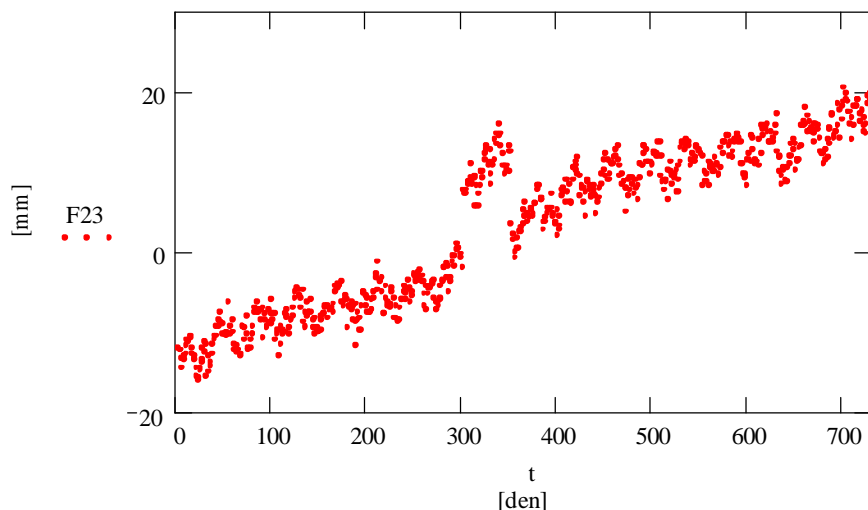
Tab.8:1 Definícia syntetických funkcií

Zložky:	Konštantný člen (dátový skok)		Lineárny trend (sklon)	Periodický člen			Náhodný šum ~ n(m,s)
	parameter	interval	parameter	perióda	amplitúda	fáza	št. odchýlka
Funkcia	a [mm]	I [deň]	b [$\mu\text{m}/\text{deň}$]	T [deň]	r [mm]	ϕ [deg]	σ [mm]
F1	-15.0	1 - 730	0.05	365.0	2.00	63.4	---
F3				41.0	1.70	45.0	1.0
F2	- 15.0	1 - 300	0.05	365.0	2.00	63.4	---
	- 5.0	301 - 350					1.0
F23	- 15.0	351 - 630					1.0
F24	- 18.0	631 - 730					2.0

Tri z nich sú znázornené i graficky na obr.8.1.

Ďalej som uvažoval prípad, ak v súbore meraní chýbajú hodnoty meranej veličiny v určitých časových intervaloch. Tak som vytvoril 7 nových časových radov, ktoré sú dané diskretnými hodnotami funkcií F1 a F3, ale na množine navzájom nenadväzujúcich intervalov podľa tabuľky 8.2 (pozri obr.8.2).

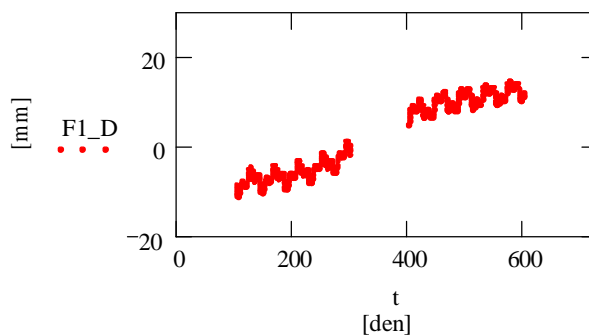
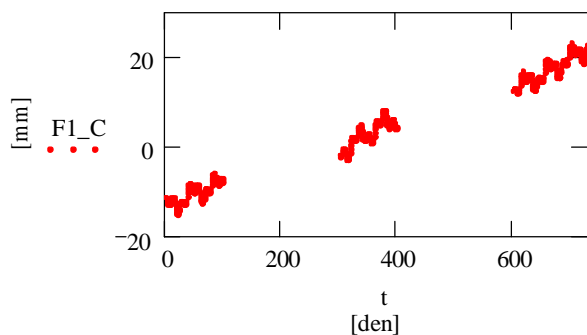
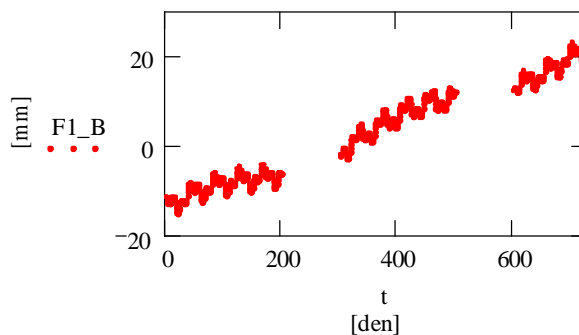
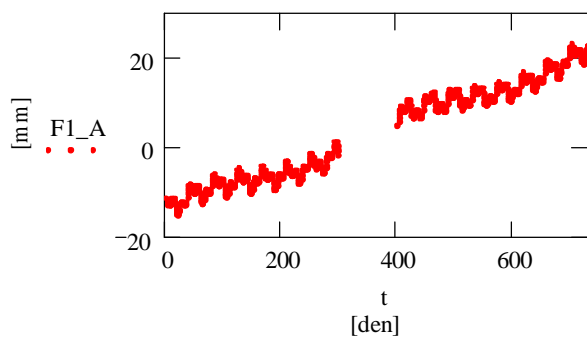




Obr.č.8.1: Synteticky pripravené časové rady.

Tab.8.2: Definovanie radov s intervalmi chýbajúcich údajov.

Intervaly diskretných hodnôt [deň]	Pôvodná funkcia:	F1	F3	počet diskretných hodnôt [deň]
$\langle 1, 300 \rangle \cup \langle 401, 730 \rangle$		F1_A	F3_A	630
$\langle 1, 200 \rangle \cup \langle 301, 500 \rangle \cup \langle 601, 730 \rangle$		F1_B	F3_B	530
$\langle 1, 100 \rangle \cup \langle 301, 400 \rangle \cup \langle 601, 730 \rangle$		F1_C	F3_C	330
$\langle 101, 300 \rangle \cup \langle 401, 600 \rangle$		F1_D		400

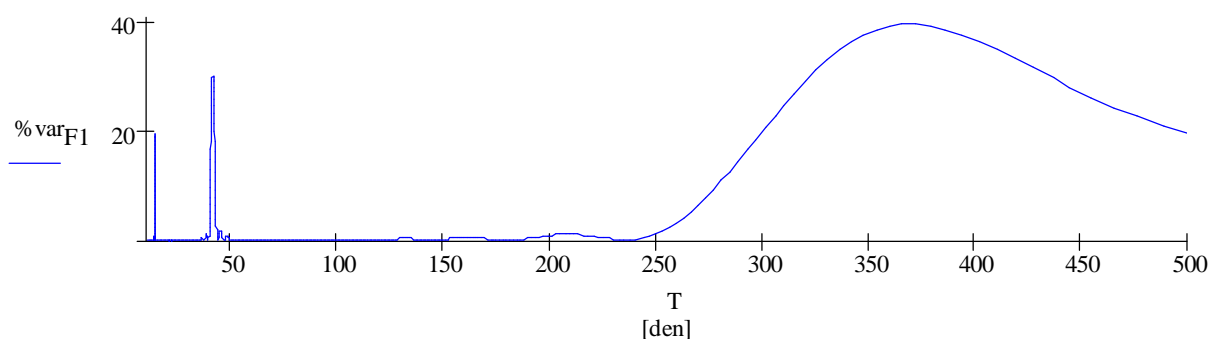


. Obr.8.2: Syntetické časové rady s intervalmi chýbajúcich údajov.

Uvedená vzorka časových radov vo všeobecnosti pomerne dobre vystihuje vplyvy, ktorých neuváženie v praxi spôsobuje väčšie či menšie skreslenie výsledných parametrov modelovaných procesov. Nás zaujíma, ako a do akej miery sa podpisujú pod celkový spektrálny obraz časového radu a hodnoty odhadnutých parametrov periodických zložiek: *amplitúdy* r a *fázy* f . Tieto parametre sme doteraz neuvádzali, namiesto nich vo všeobecnom tvare trigonometrického polynómu vystupovali parametre a , b (prípadne c_1 , c_2). Použitím r a f dostaneme $T = a \cos(\frac{2\pi}{T}t) + b \sin(\frac{2\pi}{T}t) = r \cos(\frac{2\pi}{T}t - f)$, kde T je perióda modelovanej zložky. Okrem toho platí $r = \sqrt{a^2 + b^2}$ a $f = \arctg \frac{b}{a}$.

Testovacie časové rady som najprv podrobil spektrálnej analýze na zistenie významných periód a ich percentuálnej účasti na celkovom rozptyle radu. Výsledky sú v tabuľke 8.3.

Druhým testom bolo simultánne vyrovnanie radu trigonometrickými polynómami so známymi (definovanými) periódami 365, 41 a 14 dní, určením ich parametrov (tab.8.4). Pripomínam, že v prvom aj druhom prípade sa z pochopiteľných dôvodov súčasne uvažoval vplyv aj ostatných známych komponentov radu (dátové skoky, trend). Pre názornosť, obr.8.3 ukazuje jednoduchý spektrálny obraz funkcie F1.



Obr.8.3: Spektrum funkcie F1.

Časové rady GPS observácií

Na rozdiel od predošlých jednoduchých funkcií, časový rad z meraní GPS na permanentných staniách je výsledkom pôsobenia obrovského komplexu geodynamických, atmosférických a ďalších fyzikálnych procesov. Dynamika polohy sa mení z bodu na bod, pre každý zo smerov má charakteristický priebeh a nie vždy na prvý pohľad vykazuje prítomnosť periodických procesov. Najvýraznejšie a najčastejšie bývajú ročné a polročné (teda sezónne) vplyvy.

Nasledujúci experiment sa nesie v duchu toho predošlého, tentokrát však skutočné periódny nie sú známe. Hoci spektrálna analýza z periodogramu odhalí niekoľko významných frekvencií, môže ísť o tzv.falošné frekvencie a ich potlačenie sa značná časť spektra znehodnotí. Pri interpretácii významnosti spektrálnych vrcholov treba totiž predovšetkým pamätať na ich previazanie s viac, alebo menej známym fyzikálnym javom. Problém interpretácie presahuje náplň tejto práce, niekedy i oblasť pôsobnosti jedného vedného odboru, preto za reálne boli ďalej uvažované iba sezónne vplyvy.

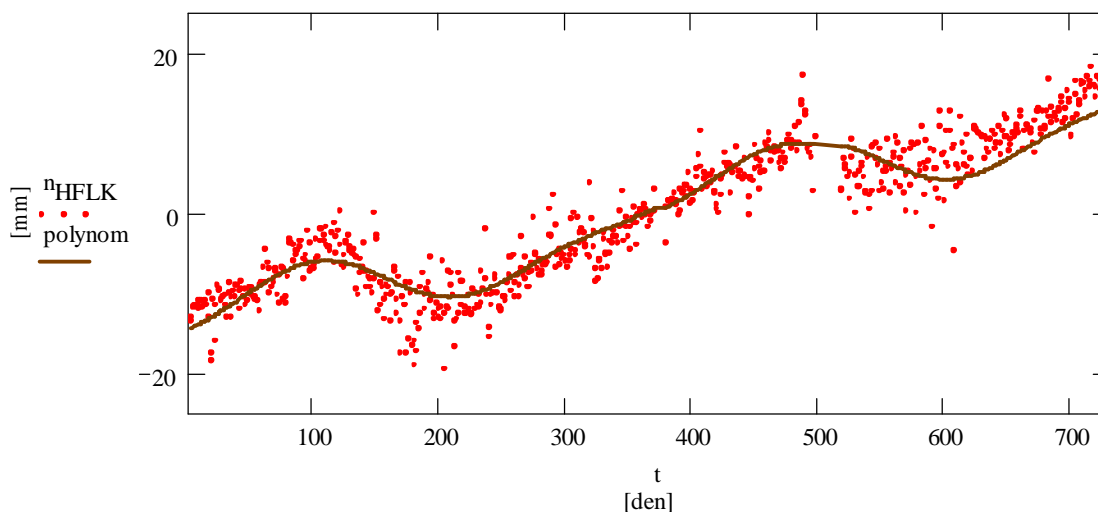
Experiment pre HFLK som navrhol v štyroch krokoch. Prvým je získanie spektra popri vyrovnaní konštantného člena a lineárneho trendu. V druhom je potlačená už aj polročná perióda, v treťom namiesto nej ročná, a v štvrtom kroku sú do deterministického modelu

zahrnuté obidve. Pre BOR1 sú vykonané iba prvý a posledný krok, nakoľko ostatné dva mali iba informatívny charakter. Konkrétne ma zaujímalo, ako ovplyvní eliminácia jednej sezónnej zložky polohu spektrálneho vrchola druhej zložky.

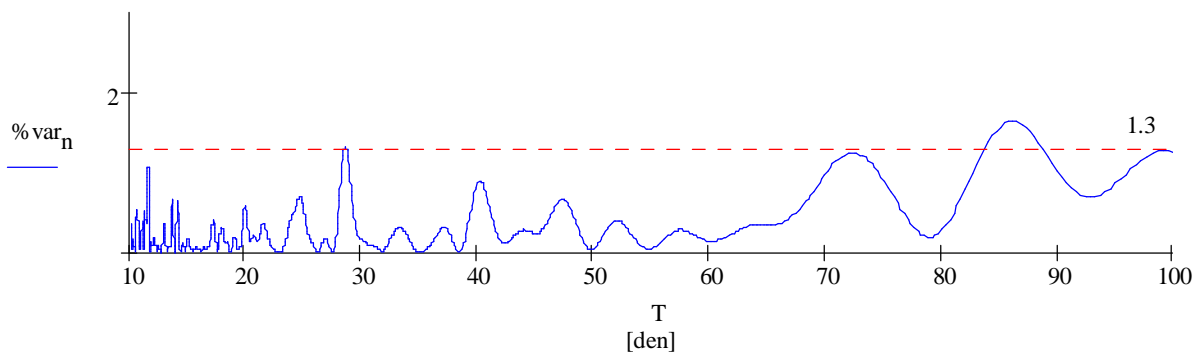
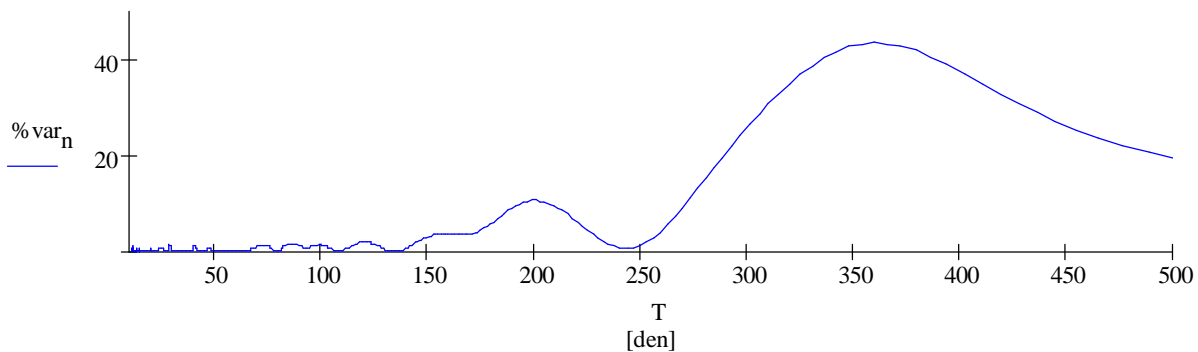
Časové rady zo stanice HFLK obsahujú intervaly, v ktorých merania neboli vykonané alebo zahrnuté medzi ostatné pravdepodobne z technických príčin (manipulácia s anténou a pod.). Hoci boli časové body neskôr dointerpolované, nezahrnul som ich do vstupného súboru, pretože by pôsobili viac skresľujúco ako samotné medzery. Interpolácia má význam, ak sa použije klasická technika analýzy radu, ktorá nepravidelnosť rozostupu časových bodov nedokáže akceptovať.

Výsledok experimentu je v tab.č.8.5. V stĺpcoch pre spektrálnu analýzu sú uvedené najprv periódy, ktoré približne súhlasia s jednou zo sezónnych zložiek, a potom aj tie periódy, ktoré som subjektívne vybral za zaujímavé spomedzi všetkých „štatisticky významných“ spektrálnych vrcholov. Kritérium štatistickej významnosti bolo $c_{\alpha}=1.3\%$ variancie pre hladinu významnosti $\alpha=0.01\%$.

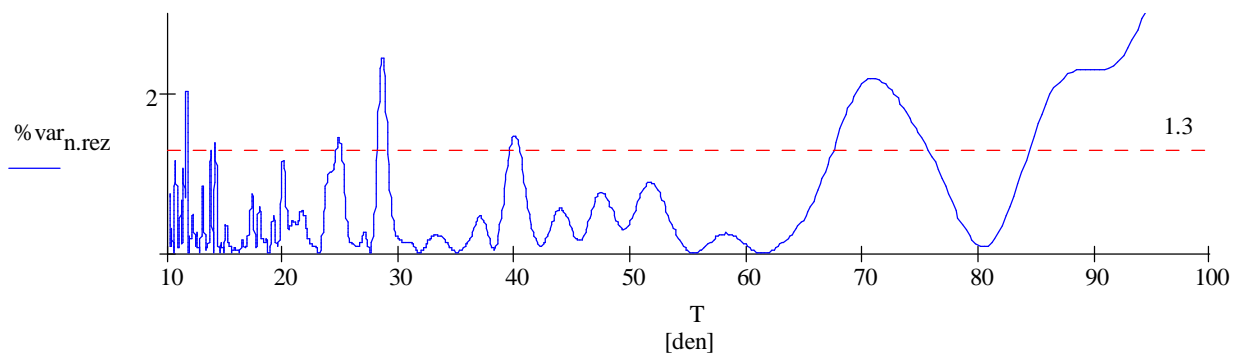
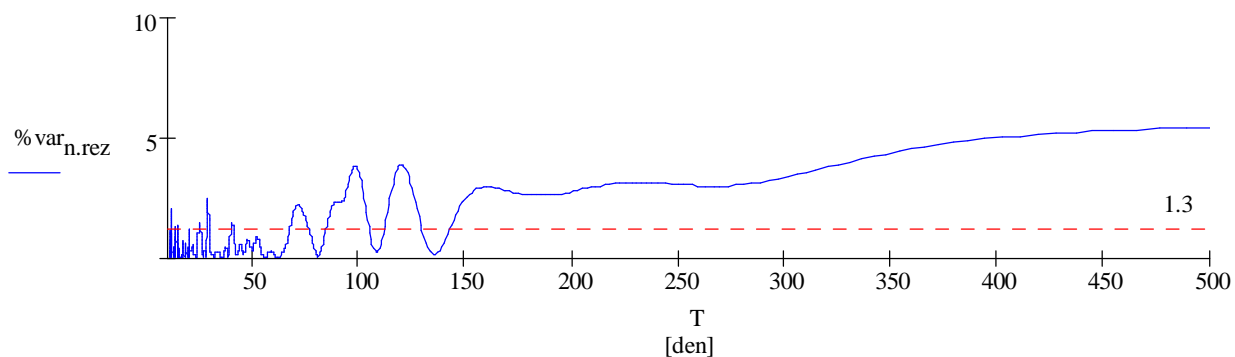
Pozrime sa na analýzu jedného radu, napr. pre súradnicu n HFLK, z trochu názornejšieho hľadiska. Originálny rad i aproximujúci polynóm, ktorý obsahuje obidva periodické členy, sú znázornené na obr.8.4. Parametre polynómu sú z experimentu č.4 (Tab.č.8.5). Ďalej na obr.8.5 je zobrazené spektrum radu ešte pred potlačením ktorejkoľvek periodickej zložky (exp.č.1) a na obr.8.6 po ich eliminácii.



Obr.8.4: Časový rad n HFLK a aproximujúci polynóm s ročnou aj polročnou periódou



Obr.8.5: Spektrum časového radu n HFLK ak neboli eliminované sezónne zložky.



Obr.8.6: Spektrum časového radu n HFLK ak boli eliminované sezónne zložky (ročná aj polročná).

Výsledné tabuľky

Tab.8.3: Detekcia periód v syntetických funkciách.

Funkcia	odhad parametrov známych komponentov									spektrálna analýza			
	konšt. člen		lineárny trend		periodický člen					významné periódy			
	a	σ_a	b	σ_b	T	r	σ_r	ϕ	σ_ϕ	predpokladané		iné	
	[mm]		[$\mu\text{m}/\text{deň}$]		[deň]	[mm]		[deg]		[deň]	% var.	[deň]	% var.
F1	-14.04	0.15	47.3	0.4						395.0	39.7		
										41.0	30.1		
										14.0	19.5		
F3	-14.11	0.18	47.0	0.4						395.0	30.6		
										41.0	23.8		
										14.0	13.7		
F2	-12.67	0.17	35.3	0.9						365.0	21.1		
	-0.82	0.39								41.0	42.9		
	-7.57	0.45								14.0	26.2		
	-8.90	0.64											
F23	-12.77	0.21	35.1	1.1						365.0	16.2		
	-0.88	0.47								41.0	32.3		
	-7.60	0.55								14.0	18.1		
	-9.11	0.77											
F24	-12.55	0.25	34.9	1.3						365.0	8.9		
	-0.15	0.56								41.0	24.2		
	-7.30	0.65								14.0	10.6		
	-8.79	0.92											
F1_A	-14.03	0.16	47.2	0.4						371.7	41.1		
										41.0	28.8		
										14.0	19.0		
F1_B	-13.5	0.16	47.0	0.4						365.0	32.9		
										41.1	34.1		
										14.0	19.1		
F1_C	-13.0	0.18	45.5	0.4						353.3	21.4	280.1	20.0
										41.1	35.6	162.1	20.1
										14.0	23.5	47.5	28.5
												36.3	21.7
F1_D	-15.5	0.26	50.8	0.7						362.3	42.8		
										41.0	29.9		
										14.0	20.1		

Tab.8.4: Odhad parametrov všetkých známych zložiek syntetických funkcií.

funkcia	odhad parametrov známych komponentov									spektrálna analýza			
	konšt. člen		lineárny trend		periodický člen					významné periódy			
	a	σ_a	b	σ_b	T	r	σ_r	ϕ	σ_ϕ	predpokladané		iné	
	[mm]		[$\mu\text{m}/\text{deň}$]		[deň]	[mm]		[deg]		[deň]	% var.	[deň]	% var.
F1	-14.94	0.05	49.8	0.1	395.0	2.00	0.04	62.2	2.0				
					41.0	1.65	0.03	48.7	2.0				
					14.0	1.30	0.03	21.5	2.0				
F3	-15.08	0.11	49.7	0.3	395.0	2.10	0.08	67.4	4.2				
					41.0	1.74	0.07	50.3	4.2				
					14.0	1.29	0.07	17.6	4.2				
F2	-14.91	0.10	49.6	0.6	365.0	1.99	0.06	61.9	2.7				
	-4.97	0.22			41.0	1.65	0.03	48.8	2.0				
	-14.86	0.30			14.0	1.30	0.03	21.4	2.0				
	-17.84	0.40											
F23	-15.07	0.19	49.4	1.2	365.0	2.08	0.12	68.8	5.4				
	-4.80	0.44			41.0	1.74	0.07	50.2	3.9				
	-14.98	0.59			14.0	1.30	0.07	18.2	3.9				
	-17.82	0.81											
F24	-14.60	0.29	47.8	1.8	365.0	1.84	0.19	65.0	8.2				
	-3.77	0.67			41.0	1.80	0.10	47.2	6.0				
	-13.92	0.91			14.0	1.18	0.10	22.5	6.0				
	-16.73	1.24											
F1_A	-14.92	0.06	49.7	0.1	365.0	2.01	0.04	61.5	2.5				
					41.0	1.64	0.04	48.6	2.3				
					14.0	1.30	0.04	21.5	2.3				
F1_B	-14.90	0.07	49.6	0.2	365.0	2.00	0.06	60.8	2.9				
					41.0	1.64	0.05	48.2	2.7				
					14.0	1.29	0.05	21.5	2.7				
F1_C	-14.9	0.11	49.8	0.2	365.0	2.02	0.09	62.8	5.7				
					41.0	1.66	0.06	46.7	3.2				
					14.0	1.28	0.06	24.5	3.2				
F1_D	-14.98	0.08	49.8	0.2	365.0	1.99	0.04	62.5	3.4				
					41.0	1.64	0.04	50.5	2.5				
					14.0	1.32	0.04	19.0	2.4				
F3_A	-15.13	0.12	49.8	0.3	365.0	2.11	0.08	68.7	5.1				
					41.0	1.76	0.08	49.6	4.7				
					14.0	1.25	0.08	16.0	4.7				
F3_B	-15.05	0.15	49.5	0.4	365.0	2.09	0.12	65.2	5.7				
					41.0	1.72	0.09	49.7	5.4				
					14.0	1.32	0.09	13.2	5.4				
F3_C	-15.44	0.26	49.7	0.6	365.0	2.28	0.21	53.5	13.9				
					41.0	1.67	0.14	48.9	7.8				
					14.0	1.43	0.13	11.0	7.8				

Tab.8.5: Spektrálna analýza a odhad parametrov sezónnych zložiek v časových radoch GPS meraní.

perm. stanica GPS	súradnica	experiment. č.	odhad parametrov známych komponentov								spektrálna analýza				
			konšt. člen		lineárny trend		periodický člen				významné periódy				
			a	σ_a	b	σ_b	T	r	σ_r	ϕ	σ_ϕ	blízke 1r, 1/2r		iné	
			[mm]		[$\mu\text{m}/\text{deň}$]		[deň]	[mm]		[deg]		[deň]	% var.	[deň]	% var.
HFLK	n	1	-13.20	0.29	35.8	0.7						359.7 200.4	43.5 10.4	28.6	1.3
		2	-14.97	0.24	40.8	0.6	365.0	3.94	0.17	69.7	9.4	184.2	10.2	28.5	2.3
		3	-13.00	0.22	35.3	0.7	183.0	1.47	0.20	216.2	11.9	365.0	45.3	28.6	1.3
		4	-14.79	0.23	40.3	0.6	365.0 183.0	3.90 1.34	0.16 0.15	69.9 209.5	8.9 8.8			28.5	2.4
	e	1	-21.29	0.22	58.2	0.5						(365.0) 200.4	13.4 2.5	24.7	2.1
		2	-22.08	0.22	60.5	0.5	365.0	1.67	0.16	84.4	8.5	200.4	4.1	24.7	2.3
		3	-21.40	0.22	58.5	0.5	183.0	0.59	0.16	122.7	8.9	(365.0)	14.1	24.7	2.2
		4	-22.23	0.22	60.9	0.5	365.0 183.0	1.70 0.65	0.16 0.15	83.8 118.3	8.3 8.3	184.2	2.5	24.7	2.5
	v	1	0.71	0.55	-1.3	1.3						347.2 166.1	28.4 8.8	85.8	4.7
		2	1.18	0.51	-2.9	1.2	365.0	5.56	0.34	351.3	21.2	170.4	5.6	86.1	4.9
		3	1.09	0.55	-2.2	1.3	183.0	1.94	0.40	304.2	22.4	341.3 189.4	29.5 14.2	85.8	4.2
		4	1.54	0.51	-3.8	1.2	365.0 183.0	5.52 1.80	0.34 0.34	350.6 307.5	20.8 19.1	190.8	4.1	86.1 47.2	4.4 3.2
BOR1	n	1	-14.12	0.22	38.5	0.5						371.7 168.9	25.1 1.8	14.7	1.4
		4	-13.16	0.21	36.0	0.5	365.0 183.0	2.23 0.53	0.14 0.14	296.1 340.7	8.0 7.8			14.7	1.8
	e	1	-21.78	0.25	60.0	0.6						180.8	4.2	309.6 29.5	11.6 1.9
		4	-21.23	0.26	58.5	0.6	365.0 183.0	1.47 0.86	0.18 0.17	272.8 105.3	9.8 9.7			314.5 29.5	8.8 2.3
	v	1	0.58	0.54	-12.2	13.2						173.3	11.3	418.4 112.2 91.2	7.6 12.2 7.8
		4	0.21	0.55	-6.1	13.6	365.0 183.0	2.63 3.28	0.36 0.36	146.3 186.0	21.6 20.4			156.0 112.9 90.3	3.5 9.9 8.9

9. Záver

Metóda najmenších štvorcov (MNŠ) je v geodézii dobre známa, bežne sa používa pri vyhodnotení nadbytočného počtu meraní, napr. ak potrebujeme získať polohu bodu z určených geometrických veličín geodetickej siete. Oblasť záujmu geodézie však zahŕňa aj monitorovanie zložitejších charakteristík vonkajšieho sveta. Sú nimi napr. určovanie variácií polohy bodov geodetickej siete, pohybu pólu, alebo sledovanie dynamiky tiažového poľa Zeme. Veľa z prírodných procesov, ktoré stoja za zmenami fyzikálnych a geometrických veličín, má periodický charakter. Ich analýzou sa zaoberá niekoľko techník, najrozšírenejšou je Fourierova spektrálna analýza (FSA). V tejto práci som sa pokúsil priblížiť ďalšiu alternatívu SA, ktorá s výhodou využíva minimalizačný princíp MNŠ. Jej vlastnosti sú v mnohých smeroch presvedčivejším argumentom nasadenia v praxi, ide predovšetkým o možnosť spracovania súboru meraní s nekonštantným rozstupom, schopnosť integrovať plnú kovariančnú maticu, vypočítať spektrum v akomkoľvek rozsahu a rozlíšení, a v neposlednom rade aj spracovať údajové rady s ohľadom na prítomnosť neperiodických zložiek (rôzne formy trendu, dátové skoky). Na druhej strane, daňou za spomínané vlastnosti je vyššia výpočtová náročnosť v porovnaní s rýchlou Fourierovou transformáciou RFT. Tento nedostatok sa s rýchlym rozvojom výpočtovej techniky stáva v súčasnosti nepodstatným.

V práci som sa venoval predovšetkým teoretickej podstate SA-MNŠ, pretože doteraz nebola publikovaná žiadna práca v slovenskom alebo českom jazyku, ktorá by rozoberala túto problematiku do uspokojivých detailov. Všetky dostupné informácie bolo potrebné vyhľadať v cudzojazyčnej literatúre a pospájať do ucelenej formy. Snažil som sa priblížiť postavenie SA-MNŠ medzi ostatnými technikami spracovania údajových radov, a najmä súvislosť s používanějšíou FFT. Veľkú časť práce som zameral na priblíženie princípu MNŠ v termínoch funkcionálnej analýzy. Verím, že som spôsob aplikácie MNŠ v spektrálnej analýze predstavil dotatočne názorne i pre menej skúseného čitateľa. Nakoniec bolo v teoretickej časti nevyhnutné uviesť celý problém i zo štatistického hľadiska.

Praktickému experimentu som aj vzhľadom na náročnosť teórie venoval menej priestoru, zameral som sa iba na základné testy pomocou programu LSSA, ktorý predstavuje veľmi silný a komplexný aplikačný nástroj spektrálnej analýzy MNŠ. Bohužiaľ, verzia 5.02, ktorú som mal k dispozícii vykazuje veľkú dávku nestability, ktorá ma v experimentálnych snahách výrazne brzdila. Popis práce s programom a výsledky testov som čo najprehľadnejšie zhrnul v poslednej kapitole. Testy som sa v rámci možností snažil koncipovať jednoduchou formou, aby ich výsledky nepotrebovali rozsiahlejší komentár. Testoval som vplyv rôzneho zloženia časových radov na výsledný spektrálny obraz i odhadované parametre zložiek so známou funkčnou formou.

Problematika si vyžaduje rozsiahlejšie experimenty. Je vidieť, že SA-MNŠ môže dať veľmi cenné informácie o (časových) radoch geodetických pozorovaní, a to v sľubnej kvalite. Podľa všetkého ide o perspektívnu metódu, ktorá má šancu rozšíriť zoznam úspešne používaných techník MNŠ o aplikáciu v spektrálnej oblasti.

Zoznam literatúry

- Cipra, T.: Analýza časových řad s aplikacemi v ekonomii. SNTL, Praha 1986.
- Craymer, M.R.: The least-squares spectrum, its inverse transform and autocorrelation function: Theory and some applications in Geodesy. Graduate Department of Civil Engineering, University of Toronto, 1998.
- Červený, V.: Fourierova spektrální analýza. MFf UK, Praha 1977.
- Kubáček a Pázman: Štatistické metódy v meraní, VEDA Vydavateľstvo SAV, Bratislava 1979.
- Pagiatakis, S.D.: Stochastic significance of peaks in the least-squares spectrum. Journal of Geodesy, 1999: s.67-78.
- Steeves, R.R.: A statistical test for significance of peaks in the least-squares spectrum. Collected Papers, Geodetic Surveys Department of Energy, Mines and Resources, Surveys and Mapping Branch, Ottawa 1981a: s.149-166.
- Vaníček, P.: Approximate spectral analysis by least-squares fit. Astrophysics and Space Science 4, 1969a: s.387-391.
- Vaníček, P.: Further development and properties of the spectral analysis by least-squares. Astrophysics and Space Science 12, 1967: s.10-33.
- Vaníček, P. – Krakiwsky, E.J.: Geodesy: The Concepts. 2nd edn, North Holland, Amsterdam 1986.
- Wells, D.E. – Vaníček, P. – Pagiatakis, S.D.: Least squares spectral analysis revisited. Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton 1985.

Zhrnutie

Práca podáva detailný popis teoretického základu spektrálnej analýzy metódou najmenších štvorcov (SA-MNŠ). Najprv je daný všeobecný pohľad do problematiky spracovania jednorozmerných, časových radov a klasifikácia existujúcich metód. Nasledne je v nevyhnutnej miere popísaná v súčasnosti najpoužívanejšia technika analýzy časových radov v spektrálnej oblasti – Fourierova spektrálna analýza a jej vzťah k univerzálnejšej SA-MNŠ. Tá je v skratke predstavená základnou myšlienkou jej autora P.Vaníčka a ďalej podrobne preberaná použitím vyjadrovacích prostriedkov funkcionálnej analýzy. Možno vidieť dve hlavné aplikácie: určenie spektra a odhad parametrov funkčnej formy systematických zložiek časových radov, pričom sa odporúča ich simultánne vykonanie. Záver teoretickej časti sa venuje štatistickému testovaniu významnosti spektrálnych vrcholov v závislosti od toho, či je vstupná presnosť známa. Experimentálna časť je zameraná na testovanie softvérovej implementácie SA-MNŠ na syntetických radoch, ďalej na určenie spektra a odhad parametrov systematických zložiek empirických časových radov z GPS observácií na dvoch permanentných staniciach (HFLK, BOR1).

Práca si dáva za cieľ vyzdvihnúť významné vlastnosti SA-MNŠ ako je spracovanie nestacionárnych časových radov s nepravidelným delením časových bodov, integrovanie plnej kovariančnej matice a výpočet ľubovoľných frekvencií.

Summary

The thesis gives a detailed description of theoretical basis of least-squares spectral analysis (LSSA). At first, a common view into one-dimensional data series processing is given, as well as an existing procedures classification. Then, a necessary description of frequent analysis, used for time series to be analyzed in spectral domain, called Fourier spectral analysis is given better to understand it's relation to LSSA. There is a short review of basic ideas, as presented by it's author P. Vaníček, and consequently a minute description using the language of functional analysis. Two major applications are mentioned. First one deals with a problem of computing the spectrum of a complet time series, the second one is to remove some constituents before computing the spectrum of residuals, but in one run. Lastly, a derivation of probability density function (underlining the least-squares spectrum) and fomulation of criteria for statistical significance of the least-squares spectral peaks is performed.

Experimantal section is oriented to test the powerful software implementation, named LSSA, on artificial time series, further to perform spectrum determination and systematic parameter estimation of observed time series. These originate from observing the position on GPS permanent stations (HFLK, BOR1) involved in CERGOP project.

This thesis wants to highlight significant properties of LSSA, e.g. non-stationary and unequally spaced data series processing, fully populated covariance matrix integrating and possibility to compute any real frequency.

Prílohy

Príloha A **Výpis súboru lssa.in**

Priloha B **Výpis súboru lssa.out**

Príloha A

Výpis súboru lssa.in

```
=====
SURADNICA N BODU HFLK
HFLK1.DAT          NAME OF INPUT FILE FMT:A12
  710              NUMBER OF DATA POINTS (I6)
Day               UNITS OF TIME (EX. HOUR, MIN, SEC, etc.) (FMT: A4)
mm               UNITS OF TIME SERIES (EX. mGal, m, mbar, etc)(FMT: A4)
=====
RANDOM CONSTANT (DATUM SHIFTS)
  1               NUMBER OF DATUM SHIFTS (I6)
  1.5000         DATUM BIASES (TIMES) (F20.12)
=====
TRENDS
  1              RANDOM RAMP (LINEAR TREND) (1=ON, 0=OFF) (I6)
  0              EXPONENTIAL TREND (a*exp(-b*t))(1=ON, 0=OFF)
  1.000000      DECAY CONSTANT (=b above) (F15.6)
  0              POLYNOMIAL SECOND ORDER TERM (1=ON, 0=OFF) (I6)
  0              POLYNOMIAL THIRD ORDER TERM (1=ON, 0=OFF) (I6)
  0              POLYNOMIAL FORTH ORDER TERM (1=ON, 0=OFF) (I6)
  0              POLYNOMIAL FIFTH ORDER TERM (1=ON, 0=OFF) (I6)
=====
PROCESSES          ALL FORMATS:I6
  0              FIRST ORDER TERM OF AR OR RANDOM WALK (1=ON, 0=OFF)
  0              SECOND ORDER TERM OF AR (1=ON, 0=OFF)
  0              THIRD ORDER TERM OF AR (1=ON, 0=OFF)
  0              FORTH ORDER TERM OF AR (1=ON, 0=OFF)
  0              FIFTH ORDER TERM OF AR (1=ON, 0=OFF)
=====
PERIODIC SIGNALS
  2              NUMBER OF FORCED FREQUENCIES (I6)
  365.0          ROCNA   FORCED PERIODS AND NAMES (F20.12,1X,A7)
  183.0          POLROC
=====
CHARACTERISTICS OF SERIES
  0              A-PRIORI VAR. FACTOR (0 = UNKNOWN, 1 KNOWN) (I6)
  0              WEIGHTS = 1 (1=YES, 0=USE GIVEN WEIGHTS) (I6)
  1.0000000     SCALE FACTOR FOR THE STANDARD DEVIATIONS (F15.7)
=====
SPECTRUM CHARACTERISTICS
  1              IB: NUMBER OF SPECTRAL BANDS (I6)
  2000          NW: NUMBER OF SPECTRAL VALUES IN BAND (I6)
  500.0000     10.0000 PL, PS: LARGEST, SMALEST PERIOD IN BAND (F12.4,1X,F12.4)
=====
STATISTICS
  0.01         CRITICAL LEVEL FOR DETECTING SIGNIFICANT PEAKS (F10.2)
  0.01         LEVEL OF SIGNIFICANCE FOR STATISTICAL TESTING (F10.2)
```

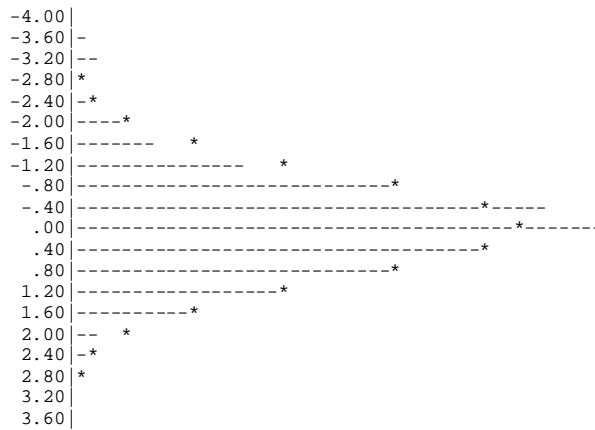
Príloha B1

Výpis súboru lssa.out.

LEAST SQUARES SPECTRAL ANALYSIS
Version 5.02

PROJECT: dipl n HFLK

HISTOGRAM OF NORMALISED RESIDUALS



CHARACTERISTICS OF DISTRIBUTION OF RESIDUALS

1st MOMENT ABOUT ZERO (MEAN) = .000 (Must be equal to 0.0)
2nd MOMENT ABOUT MEAN (VARIANCE) = .994 (Must be equal to 1.0)
3rd MOMENT ABOUT MEAN (SKEWNESS) = -.344 (Must be equal to 0.0)
4th MOMENT ABOUT MEAN (KURTOSIS) = 4.768 (Must be equal to 3.0)

CHI-SQUARE GOODNESS-OF-FIT TEST

Ho HYPOTHESIS : Normalised residuals derived from
a population with a Normal PDF
SIGNIFICANCE LEVEL : .010
DEGREES OF FREEDOM : 18
CHI-SQUARE STATISTIC : .343 (TEST PASSES)

CHI-SQUARE TEST ON THE VARIANCE FACTOR

Ho HYPOTHESIS : a-priori variance factor is EQUAL
to a-posteriori variance factor
SIGNIFICANCE LEVEL : .010
A-PRIORI VARIANCE FACTOR : UNKNOWN
DEGREES OF FREEDOM : 703
RATIO (a-post/a-priori) : *****
CHI SQUARE TEST : FAILS

ASSESSMENT OF THE DETERMINED PARAMETERS

(F-DISTRIBUTION)

Ho HYPOTHESIS : Vector of determined parameters
is different from zero
SIGNIFICANCE LEVEL : .010
A-PRIORI VARIANCE FACTOR : UNKNOWN
DEGREES OF FREEDOM : 703
QUADRATIC NORM OF PARAMETERS : .2202D+01 mm ^2
STATISTICAL TEST : PASSES

Príloha B2

SPECTRAL BAND 1
=====

- SPECTRAL VALUES BETWEEN 500.000000 10.000000:2000

- MEAN SPECTRAL VALUE FOR WHITE NOISE: .3

- CRITICAL %VAR AT 99.0% CONF LEVEL FOR DETECTING SIGNIF PEAKS IN SPECTRUM: 1.3

- AVERAGE SAMPLING INTERVAL: 1.03 Day

PERIOD (Day)	FREQUENCY (CYC/Day)	FIDELITY	%VAR	SNR (dB)	POWER DENSITY (mm ^2/freq)	
*****	.00200		5.4119	-12.42483	.231692D+15	*
*****	.00205		5.3888	-12.44453	.225125D+15	*
*****	.00210		5.3625	-12.46698	.218731D+15	*
...						
...						
...						
...						
...						
...						
10.00981453	.09990		.7270	-21.35325	.593641D+12*	
10.00490486	.09995		.7155	-21.42245	.583971D+12*	
10.00000000	.10000		.7023	-21.50413	.572810D+12*	

CHARACTERISTICS OF RESIDUAL SERIES

RES. SERIES QUADRATIC NORM: .567729D+16 mm ^2

WEIGHTED RMS OF FIT: 2.827754 mm

REDUCED CHI-SQUARE: *****

SOLUTION FOR KNOWN CONSTITUENTS

Significance of parameters based on: SIGMA*EXPANSION FACTOR

Expansion factor = 3.156

DESCRIPTION	NAME	No	VALUE (Day)	AMPLITUDE (mm)	SIGMA (mm)	PHASE (DEG)	SIGMA (DEG)	SIGNIF 99.0%
Datum shift		1	.1500000D+01	-.14794636D+02	.22884279D+00			YES
linear trend				.40346570D-01	.56080023D-03			YES
Periodic constituent	ROCNA	1	.3650000D+03	.38985530D+01	.16187907D+00	69.8777	8.9031	YES
Periodic constituent	POLROCN	2	.1830000D+03	.13447947D+01	.15018993D+00	209.5637	8.8319	YES