

LATVIJAS UNIVERSITĀTE  
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE  
MATEMĀTIKAS NODAĻA

**DAŽI NEPARAMETRISKĀS STATISTIKAS PIELIETOJUMI  
SPEKTRĀLAJĀ ANALĪZĒ**

KURSA DARBS

Autors: **Agris Vaselāns**

Stud. apl. av07050

Darba vadītājs: doc. Dr.math. Jānis Valeinis

RĪGA 2011

# Saturs

1.	Ievads . . . . .	2
2.	Spektrs un spektrālā blīvuma funkcija . . . . .	3
2.1.	$ARMA(p, q)$ spektri . . . . .	5
2.2.	Sezonālo $ARIMA(p, d, q) \times (P, D, Q)_s$ spektri . . . . .	6
3.	Periodogramma . . . . .	7
4.	Periodogrammas gludināšana . . . . .	10
4.1.	Kodolu gludinātā periodogramma . . . . .	10
4.2.	Logaritmiskā vai lineārā skala periodogrammas attēlošanā . . . . .	15
4.3.	Punktveida ticamības intervāli . . . . .	16
4.4.	Taperošana . . . . .	17
5.	Neparametriskās regresijas metodes periodogrammas gludināšanā . . . . .	19
6.	Trokšņu spektri un analīze . . . . .	24
7.	Dispersiju analīze un diskriminantu analīze frekvenču skalā . . . . .	28
8.	Secinājumi . . . . .	35
	<b>Izmantotā literatūra un avoti</b>	<b>36</b>
<b>A</b>	<b>Izveidoto programmu kods</b>	<b>37</b>
A1.	Periodogrammas gludināšanas programmas kods . . . . .	37
A2.	Kosinusa zvans . . . . .	38
A3.	Sunspot datu analīzes un ticamības intervālu programmas kods . . . . .	38
A4.	ARMA procesu teorētisko spektru programmas kods . . . . .	39
A5.	Brauna kustības pārbaudes programmas kods . . . . .	39
A6.	Vidējo vērtību pārbaudes programmas kods . . . . .	39
A7.	Spektrālo matricu AR(2) procesiem salīdzināšanas programmas kods . . . . .	41
A8.	Diskriminanta analīzes programmas kods . . . . .	43

# 1. Ievads

Lai gan spektrālās analīzes pirmsākumi meklējami jau 19. gadsimta beigās, kad Šusters (Schuster) 1898. gadā ieviesa periodogrammas jēdzienu, kas aprakstīts viņa 1906. gada darbā izpētot periodiskumu datos par saules aktivitāti [1, 181. lpp.], tā nav zaudējusi savu aktualitāti arī mūsdienās, jo pēdējos gados arvien vairāk tiek izmantotas dažādas skaitlošanas un statistiskās piejas stohastisku procesu analīzei, un aktuālāka kļuvusi arī spektrālās analīzes problemātika. Spektrālā analīze galvenokārt balstās uz Furjē analīzi (diskrēto Furjē transformāciju realizāciju analīzē) un dod iespēju analizēt stohastiskus procesus nevis laika apgabalā (domēnā), bet gan frekvenču apgabalā. Tieki apskatīts procesa cikliskums un tādejādi iegūta jauna pieeja procesa analīzē izmantojot frekvenču jēdzienu ar mērķi atklāt procesā apslēptu ciklu. Spektrālā analīze ir svarīga fizikā trokšņu spektru analīzē, ekonometrijā ekonomisko procesu ciklu izpētē, inženierzinātnēs signālu pārraidīšanas analīzē, kā arī astronomijā un citās nozarēs. Procesu analīzei, izmantojot frekvenču apgabalu, literatūrā minēti ļoti dažādi piemēri: seismiologisko procesu salīdzināšana un analīze, automašīnu motora darba analīze signālam ar troksni (izmantojot spektru salīdzināšanu bojātam un veselam motoram) kā arī tiek minēts, ka spektru analīzi var izmantot dažādu ekonomisku procesu analīzē un fizikā, pētot Brauna trokšņus.

Šim kursa darbam tik izvirzīti šādi mērķi

1. Iepazīties ar stohastisku procesu spektrālo analīzes problemātiku un tās analīzi ar neparametriskās statistikas metodēm.
2. Aprakstīt spektra gludināšanas metodes.
3. Apskatīt, kā veikt spektru salīdzināšanu.

Šajā kursa darbā apskatīti spektrālās analīzes pamatjēdzieni un dots ieskats neparametriskās statistikas pielietojumiem spektrālajā analīzē. Pirmajā daļā tiek apskatīts spektrālās blīvuma funkcijas  $f(\omega)$  jēdziens, kas cieši saistīts ar jau daudz plašāk zināmo autokovariāciju funkcijas jedzienu. Tad seko neliels ieskats ARMA procesu spektru apskatē. Otrā kursa darba daļa veltīta periodogrammas kā spektrālās blīvuma funkcijas novērtējuma izpētē: periodogrammas gludināšana ar dažādām metodēm, skalas izvēle, u.c. Trešājā daļā dots neliels ieskats neparametriskās regresijas metodei periodogrammas gludināšanā. Tālāk dots ieskats trokšņu spektru analīzē. Piektajā sadaļā aprakstīta periodogrammas izmantošanai procesu klasifikācijā.

## 2. Spektrs un spektrālā blīvuma funkcija

Lai ievestu spektra un spektrālās blīvuma funkcijas jēdzienu, apskatīsim dažus jēdzienus un teorēmas.

**Definīcija 1.** [1, 20.lpp.] Par gadījuma procesa  $\{x_t, t \in T\}$  kovariāciju funkciju sauc

$$\gamma_x(s, t) = E[x_s - \mu_s][x_t - \mu_t],$$

kur  $\mu_s$  ir procesa vidējās vērtības funkcija  $\mu_s = \int_{-\infty}^{\infty} xf_s(x)dx$  un  $f_s(x) = \partial F_s(x)/\partial x$  ir procesa sadalījuma funkcijas atvasinājums jeb blīvuma funkcija. Ja nav šaubu, kuru procesu apskata  $\gamma_x$  vietā lietosim  $\gamma$ .

**Definīcija 2.** [1, 24.lpp.] Gadījuma procesu  $\{x_t, t \in T\}$  sauc par stacionāru vājā nozīmē, ja

- $E x_t = c$ , kur  $c$  ir konstante, un nav atkarīga no parametra  $t$ ,
- kovariāciju funkcija  $\gamma_x(s, t)$  ir atkarīga tikai no laika atstarpes  $|s - t|$ .

Turpmāk ar terminu *stacionārs* sapratīsim, ka process ir stacionārs vājā nozīmē. Tad aplūkojot satcionāra gadījuma procesa kovariāciju funkciju un ņemot  $s = t + h$ , iegūsim, ka  $\gamma(t + h, t) = \gamma(h, 0)$ , kas noved pie autokovariāciju funkcijas jēdziena.

**Definīcija 3.** [1, 25.lpp.] Par stacionāra procesa autokovariāciju funkciju sauc

$$\gamma(h) = E[x_{t+h} - \mu][x_t - \mu].$$

**Teorēma 1.** Ja  $\gamma(h)$  ir stacionāra procesa  $\{x_t\}$  autokovariāciju funkcija, turklāt

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,$$

tad

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} f(\omega) d\omega, \quad h = 0, \pm 1, \pm 2, \dots$$

un procesa  $\{x_t\}$  spektrālā blīvuma funkcija ir

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h}, \quad -1/2 \leq \omega \leq 1/2.$$

*Pierādījums.* Teorēmas pierādījums atrodams, piemēram, grāmatā [1, 537-538.lpp]. Tur sākumā tiek izmantota vājā nozīmē stacionāra (iespējams, kompleksvērtīga) procesa  $x_t$  ar nulles videjo vērtību autokovariāciju funkcijas  $\gamma(h) = E(x_{t+h}, x_t^*)$  reprezentācija. Tad tiek pamatos, ka autokovariāciju funkcija ir Ermita nenegatīvi definēta, un tad dota jebkuras nenegatīvas funkcijas  $\gamma(h)$  reprezentācija.

Autokovariāciju funkcija ir Ermita nenegatīva, jo jebkurai komplekso skaitļu kopai  $a_t, t = 0, \pm 1, \pm 2, \dots$ , kas ir galīga, varam pierakstīt [1, 534.lpp]

$$E \left| \sum_{s=1}^n a_s^* x_s \right|^2 = \sum_{s=1}^n \sum_{t=1}^n a_s^* \gamma(s-t) a_t \geq 0.$$

Tālāk teorēmās dotā reprezentācija balstīta uz nenegatīvi definētām funkcijām un spektrālo sadalījuma funkciju  $F(\omega)$ , kas ir monotonu nedilstoša, nepārtraukta no labās pusēs, turklāt  $F(-1/2) = 0$  un  $F(1/2) = \sigma^2 = \gamma(0)$ , kad attiecīgi  $\omega = -1/2$  un  $\omega = 1/2$ .

**Teorēma 2.** [1, 534-535.lpp] *Funkcija  $\gamma(h)$ ,  $h = 0, \pm 1, \pm 2, \dots$  ir Ermita nenegatīvi definēta tad un tikai tad, ja tā izsakāma formā:*

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} F(\omega).$$

□

Tātad spektrālā blīvuma funkcija zināmā mērā raksturo procesa dispersiju

**Piezīme 3.** Samērā bieži spektrālo blīvuma funkciju definē nedaudz savādāk, izmantojot frekvences  $\lambda = 2\pi\omega$ , kur cikli tiek mēriti  $\pi$  radiānos. Šādi definē, piemēram, [2, 192-193.lpp], arī populārajā Hamiltona grāmatā [3] ir šāda pīeja.

**Spektrālās blīvuma funkcijas īpašības.** Spektrālā blīvuma funkcija ir analogiska sadalījuma (varbūtību) blīvuma funkcijai:

- $f(\omega) \geq 0, \forall \omega;$
- $f(\omega) = f(-\omega);$
- $f(\omega + 1) = f(\omega)$  (periods 1).

Precizēsim kā spektrālā blīvuma funkcija zināmā mērā raksturo procesa dispersiju, proti, izvēloties  $h = 0$  autokovariāciju funkcijā, zināms, ka iegūstam procesa dispersiju, tad

$$\gamma(0) = D(x_t) = \int_{-1/2}^{1/2} f(\omega) d\omega.$$

Tātad integrējot spektrālo blīvumu funkciju pa visām frekvencēm, iegūsim procesa dispersiju un, ja izpildīti visi teorēmas nosacijumi, tad autokovariāciju un spektrālā blīvuma funkcijas satur vienu un to pašu informāciju. Tikai autokovariāciju funkcija informāciju sniedz, izmantojot *laga* jēdzienu, bet spektrālā blīvuma funkcija - frekvences (cikli novērotā procesa laika atstatumā (cikli/gadā; cikli/mēnesī utt.))

## 2.1. $ARMA(p, q)$ spektri

Atgādināsim  $ARMA(p, q)$  procesa definīciju, lai precizētu apzīmējumus un vēlāk dotu  $ARMA(p, q)$  teorētiskā spektra formulu.

**Definīcija 4.** Process (laikrinda)  $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$  ir  $ARMA(p, q)$ , ja tas ir stacionārs un

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

kur  $\theta_q \neq 0$ ,  $\phi_p \neq 0$  un  $\sigma_\varepsilon^2 > 0$ , un  $\varepsilon_t$  ir baltā trokšņa process ar  $E\varepsilon_t = 0$ ,  $E\varepsilon_t^2 = \sigma_\varepsilon^2$ .

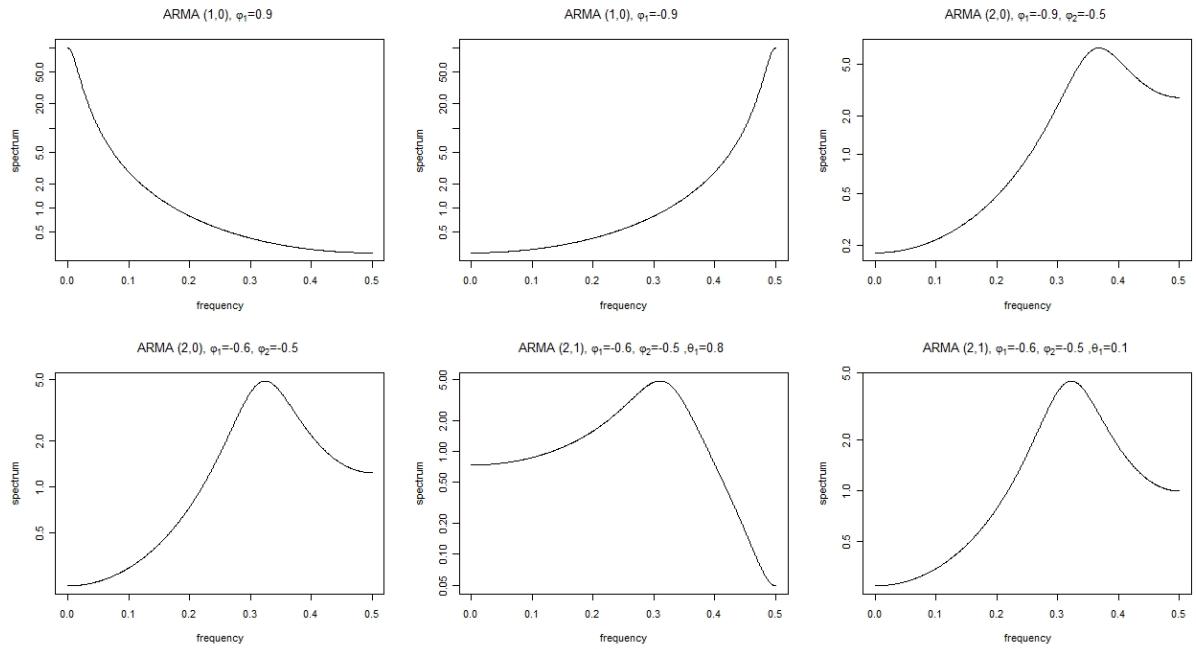
**Apgalvojums 4.** [1, 229.lpp]

$ARMA(p, q)$  teorētiskais spektrs ir pierakstāms formā

$$f(\omega) = \sigma_\varepsilon^2 \frac{|1 + \sum_{k=1}^q \theta_k e^{-2\pi i \omega k}|}{|1 - \sum_{m=1}^p \phi_m e^{-2\pi i \omega m}|}.$$

Tātad jebkuram  $ARMA(p, q)$  procesam var precīzi aprakstīt un uzzīmēt tā teorētisko spektru. Apskatīsim dažus piemērus.

Lai gan jebkuram procesam zināms tā teorētiskais spektrs, praktiski tos izmantot ir visai grūti, jo līdzīgiem spektriem var atbilst samērā dažādi procesi. Turklāt vēlāk veicot simulācijas redzēsim, ka arī katrā realizācijā procesu spektri ir nedaudz atšķirīgi un tādēļ izmantot procesa spektru, lai noteiktu  $ARMA(p, q)$  kārtas  $p$  un  $q$  un vēl jo vairāk parametru novērtējumus būs pietiekoši apgrūtinoši. Kā minēts arī literatūrā, spektrus drīzāk var izmantot procesa analīzes beigās, kad jāizšķiras starp dažiem vispemērotāka-



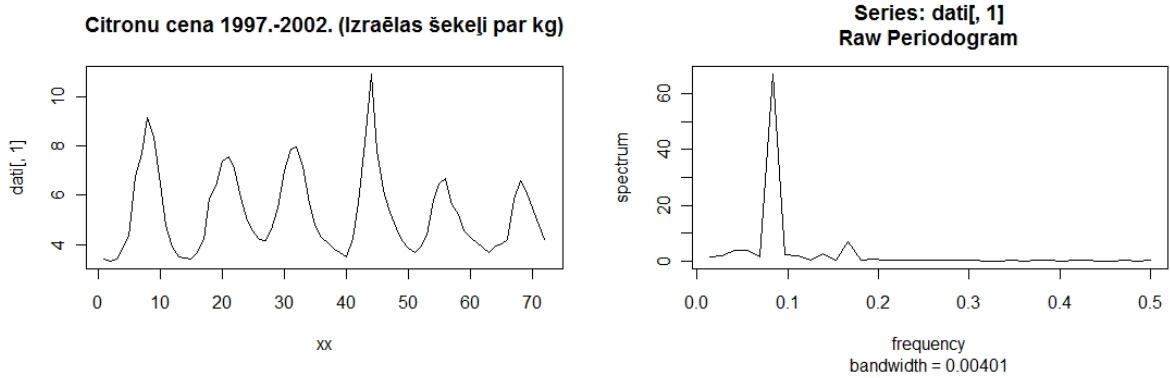
1. att. Dažādi  $ARMA(p, q)$  spektri.

jiem modeļiem vai interesē pētāmā procesa spektrs. Grāmatā [4, 363-368.lpp] sākumā noteikta procesa kārta ar AIC kritēriju un tad analizēti gludinātie spektri.

## 2.2. Sezonālo $ARIMA(p, d, q) \times (P, D, Q)_s$ spektri

Sezonālo ARMA procesu spektru piemēri doti grāmatā [4, 339-340.lpp]. Sezonālitāte grafiski izpaužas kā *pīķa* parauga atkārtojumi ar mazāku svārstību amplitūdu ik pēc  $1/k$  frekvencēm, kur  $k$  apzīmē sezonalitātes kārtu. Šajā grāmatā labi ilustrēti gadījumi  $ARIMA(1, 0, 0) \times (1, 0, 0)_{12}$  un  $ARIMA(0, 0, 1) \times (0, 0, 1)_{12}$ .

Savukārt šeit apskatīsim reālu piemēru par citronu cenām (dati peejami <http://faculty.arts.ubc.ca/ediewert/concepts3.pdf> [atsauce 23.01.2011.]). (Var atzīmēt, ka šim piemēram samērā labi atbilst modelis  $ARIMA(2, 0, 0) \times (2, 0, 0)_{12}$ .) Kā redzams, tad periodogrammā pie vērtības  $\omega \approx 0.82$  ir liels *pīķis*, kas tieši norāda uz datu periodiskumu, proti, 12 mēnešiem (jo  $1/12 \approx 0.83$ ), savukārt nākošais leciens ir šī lielā *pīķa atbalsošanās*.



2. att. Citronu cenu laikrindas attēls un periodogrammai.

### 3. Periodogramma

**Definīcija 5.** [1, 187.lpp] Ja doti dati  $x_1, x_2, \dots, x_n$ , tad par diskrēto Furjē transformāciju (DFT) sauc

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^n x_t e^{-2\pi i \omega_j t},$$

kur  $j = 0, 1, \dots, n-1$  un frekvences  $\omega_j = j/n$  tiek sauktas par Furjē jeb fundamentālajām frekvencēm.

**Definīcija 6.** [1, 188.lpp] Ja doti dati  $x_1, x_2, \dots, x_n$ , tad par periodogrammu sauc

$$I(\omega_j) = |d(\omega_j)|^2,$$

kur  $j = 0, 1, \dots, n-1$ .

Šeit jāpiezīmē, ka  $I(0) = n\bar{x}^2$ , kur  $\bar{x}$  ir izlases vidējā vērtība. Bet, ja  $j \neq 0$ , tad  $I(\omega_j) = \sum_{h=-n+1}^{n-1} \hat{\gamma}(h) e^{2\pi i \omega_j h}$ , kur  $\hat{\gamma}(h)$  ir izlases autokovariāciju funkcija

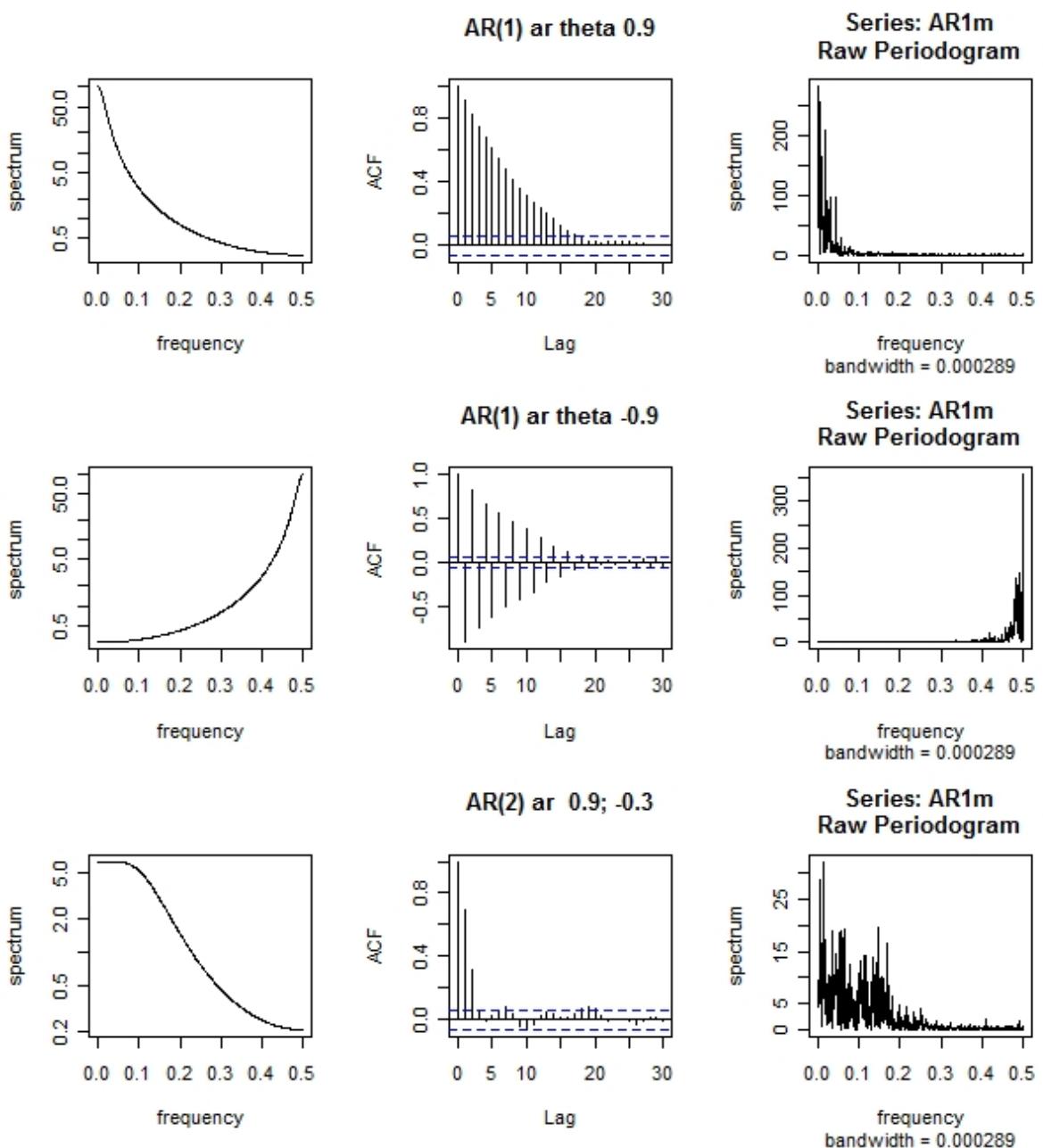
$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}).$$

**Piezīme 5.** Līdzīgi kā definējot spektrālo blīvuma funkciju, definīcijā var izmantot frekvences  $\pi$  radiānos  $\lambda = 2\pi\omega$ , tā arī, definējot periodogrammu, pastāv šīs pašas atšķirības. Tātad nepieciešams izvēlēties vienu no pieejām un saskaņot šīs definīcijas. (Piemēram, grāmatā [2, 185.lpp] dota gadījuma procesa periodogramma izmantojot tieši frekvences  $\pi$  radiānos.)

Var teikt, ka periodogramma sadala procesa dispersiju pa frekvencēm, bet periodogrammai kā spektrālās blīvuma funkcijas novērtējumam ir 2 būtiski trūkumi. Pirmais ir tas, ka aplūkoto frekvenču skaits ir mainīgs un ierobežots, jo tas ir atkarīgs no izlases novēro-

jumu skaita. Un otrs trūkums ir tas, ka periodogramma neklūst gludāka (netiecas uz teorētisko spektru), ja palielina novērojumu skaitu, t.i., periodogramma nav būtisks spektrālās blīvuma funkcijas novērtējums [2, 175.lpp], [4, 342.lpp]. Tādēļ svarīgi veikt periodogrammas gludināšanu ar dažādām metodēm, kas arī sagādā vislielākās grūtības spektrālajā analīzē. Daži gludināšanas veidi tiks apskatīti vēlāk.

Vienkāršākajos gadījumos var aplūkot procesa autokorelāciju funkciju un spektru, tādejādi mēģinot apskatīt, vai pastāv kāda saistība, jo, kā jau iepriekš minēts, pēc būtības autokovariāciju funkcija un spektrālā blīvuma funkcija satur vienu un to pašu informāciju. Līdz ar to arī autokorelāciju funkcijai varētu aplūkot saistību ar spektru.



3. att. Dažādi  $AR(p)$  spektri.

Bet, kā redzams, praktiski nekādas saistības nav, jo autorelāciju funkcija izmanto *lagus*, bet periodogramma (kā spektrālās blīvuma funkcijas novērtējums) informāciju izsaka izmantojot frekvences.

## 4. Periodogrammas gludināšana

Kā jau iepriekš noskaidrojām, visbūtiskāk ir noskaidrot ar kādām metodēm veikt periodogrammas gludināšanu. Pastāv divas būtiski atšķirīgas metodes. Visbiežāk tiek apskatīta tā sauktā ”kodolu gludināšana”, kur, līdzīgi kā kodolu gludināšanā, meklējot sadalījuma blīvuma funkciju, tiek izvēlēts kāds noteikts joslas platums un veikta periodogrammas ”vidējošana” attiecīgajā frekvencē dažādi izsvarojot vai neizsvarojot apkārtējo frekvenču ietekmi. Tas tiek darīts, izvēloties apkārtējo frekvenču skaitu  $m$ , kas arī sagādā vislielākās grūtības šajā metodē, jo nav viennozīmīgas metodes kā noteikt parametru  $m$ . Otra fundamentālā pieeja ir periodogrammas gludināšana izmantojot neparametisko regresiju. Tad savukārt, kā tas ir vienmēr, vissvarīgāk noteikt parametru  $h$ , t.i., joslas platumu.

### 4.1. Kodolu gludinātā periodogramma

Aprakstīsim visbiežāk sastopamo periodogrammas gludināšanas veidu - kodolu gludināšanu. Šeit tiek izvēlēta frekvenču josla  $\mathcal{B}$  no  $L \ll n$  sekojošām fundamentālām frekvencēm, kas centrētas ap  $\omega_j = j/n$ , kas tuva interesējošai frekvencei  $\omega$ :

$$\mathcal{B} = \{\omega : \omega_j - m/n \leq \omega \leq \omega_j + m/n\},$$

kur  $L = 2m + 1$ . Lielumu  $\mathcal{B}_\omega = L/n$  sauc par joslas platumu. [1, 197. lpp]

**Definīcija 7.** [1, 197. lpp] Ja  $\mathcal{B}$  ir frekvenču josla no  $L \ll n$  sekojošām fundamentālām frekvencēm, kas centrētas ap  $\omega_j = j/n$ , tad par vidējoto (angliski *averaged*) periodogrammu sauc

$$\bar{f}(\omega) = \frac{1}{L} \sum_{k=-m}^m I(\omega_j + k/n).$$

Lai iegūtu precīzāku spektra novērtējumu, periodogrammu var gludināt izmantojot dažādus svarus, tādēļ definēsim svērto periodogrammu.

**Definīcija 8.** [1, 203. lpp] Par svērto periodogrammu sauc

$$\hat{f}(\omega) = \sum_{k=-m}^m h_k I(\omega_j + k/n),$$

kur svari  $h_k$  apmierina nosacījumus  $\forall k : h_{-k} = h_k$  un  $\sum_{k=-m}^m h_k = 1$ .

Svaru  $h_k$  izvēlei tiek ieteikti Daniela, Fejer, Dirihlē un modifīcētais Daniela kodols. Šajā gadījumā parametrs  $L$  jāaizvieto ar  $L_h = (\sum_{k=-m}^m h_k^2)^{-1}$ , tad joslas platums  $\mathcal{B}$  svērtajai periodogrammai būs  $\mathcal{B} = L_h/n$ .

### Parametra $m$ izvēle

Tātad svarīga ir parametra  $L$  vai  $m$  (angliski tiek arī saukts par *window closing*, jo, samazinot  $m$ , samazinās joslas platums) izvēle. Neliels ieskats par parametru  $m$  tiek dots grāmatā [4, 355. lpp]. Dženkins (*Jenkins*) un Vatts (*Watts*) (1968) esot ieteikuši mēģināt ar trim dažādām  $m$  vērtībām un tad izvēlēties atbilstošāko. Mazākā vērtība dos priekšstatu par to, kur atrodas lielākie pīķi, bet periodogrammā būs daudz pīķu, un tā būs samērā negluda. Liela parametra vērtība var veidot līkni, kas būs *pārgludināta*. Kompromisu došot trešā vērtība. Savukārt Čatfīlds (*Chatfield*) (2004) iesakot ņemt  $m = \sqrt{n}$ . Tādēļ, vēl jāaplūko tikai  $m = 2\sqrt{n}$  un  $m = 0.5\sqrt{n}$ , kas dos priekšstatu par spektru. Un piedevām tiek minēts spilgts citāts, kas raksturo šo problemātiku: "Experience is the real teacher and cannot be got from a book." ( *burtiski* tulkojot no angļu valodas "Pieredze māca rīkoties un tā nav atrodama/izlasāma grāmatās" )

Tātad varam apkopot ieteikumus:

- izvēlēties  $m = \sqrt{n}$ ;
- vēl var aplūkot  $m = 2\sqrt{n}$  ,  $m = 0.5\sqrt{n}$ .

### Gludināšanas kodoli

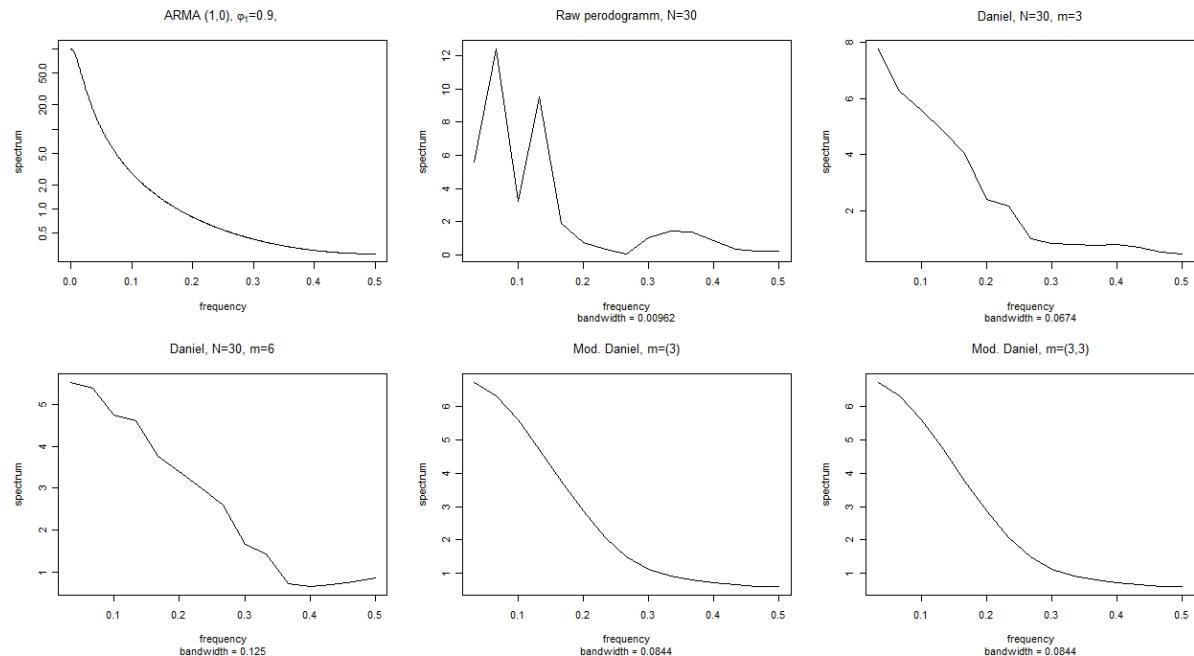
Kā jau ierasts statistikā kodolu gludināšanā paša kodola izvēle nav tik būtiska kā joslas platura (šajā gadījumā parametra  $m$ ) izvēle. Literatūrā visbiežāk tiek minēti Daniela un modifīcētais daniela kodols. Daniela kodols ir visvienkārsākais, tas visām frekvencēm no joslas piekārto vienādus svarus, proti, tiek piekārtoti svari  $1/L$ , kur  $L = 2m + 1$ . Savukārt modifīcētais Daniela kodols pirmajam un pēdējam elementam piešķir pusi no pārējiem svariem, t.i., pirmais un pēdējais svara koeficienti ir uz pusi mazāki nekā pārējie, kas ir vienādi savā starpā. Protams arī modifīcēta Daniela kodola svaru summa ir 1.

Vēl bieži kodolu modifikācija notiek vienārši atkārtojot kodolu gludināšanu periodogrammai. Matemātiski tas nozīmē, ka tiek ņemta kodolu konvolūcija. Grāmatā [4, 354. lpp] attēlots kā izmainās svara koeficienti atkārtojot izsvarošanu ar modifīcēto Daniela kodolu pie  $m = 3$ . Atkārtojot kodolu otru reizi svari veido trijstūra formas grafiku, bet jau pie trešās atkāršanas, tā forma jau atgādina normālo sadalījumu.

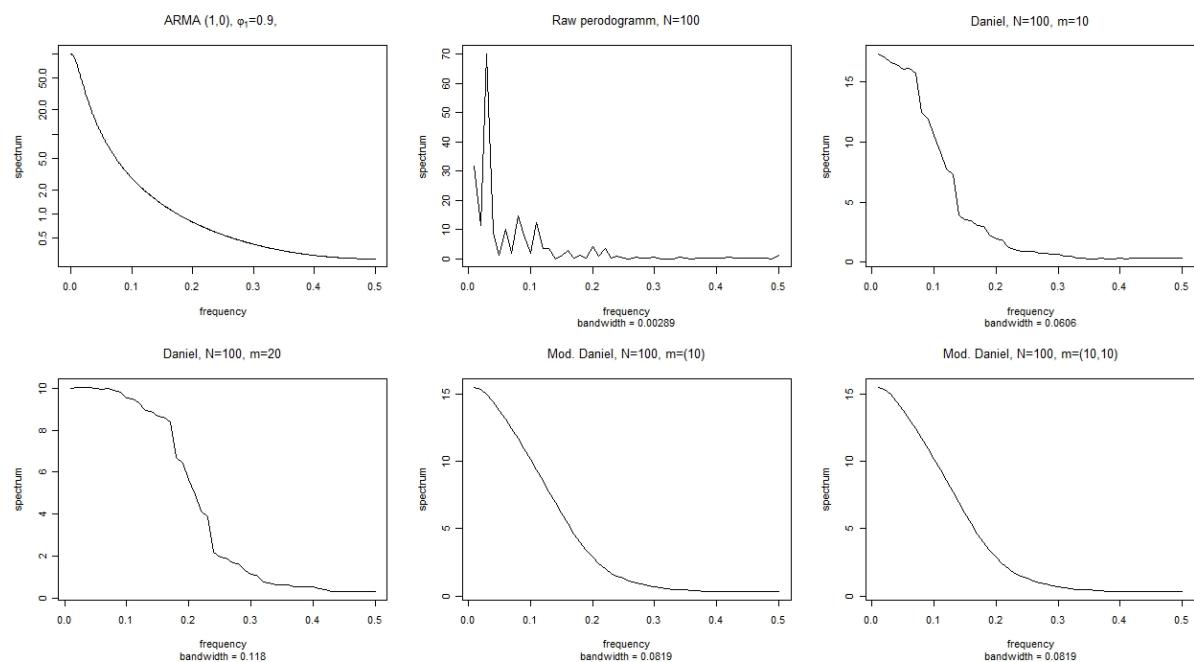
Vēl literatūrā atrodami arī Fejer un Dirihielē kodoli (skat., piemēram, [1, 207.-215. lpp])

### Kodolu gludināto periodogrammas (gludinātie spektra novērtējumi)

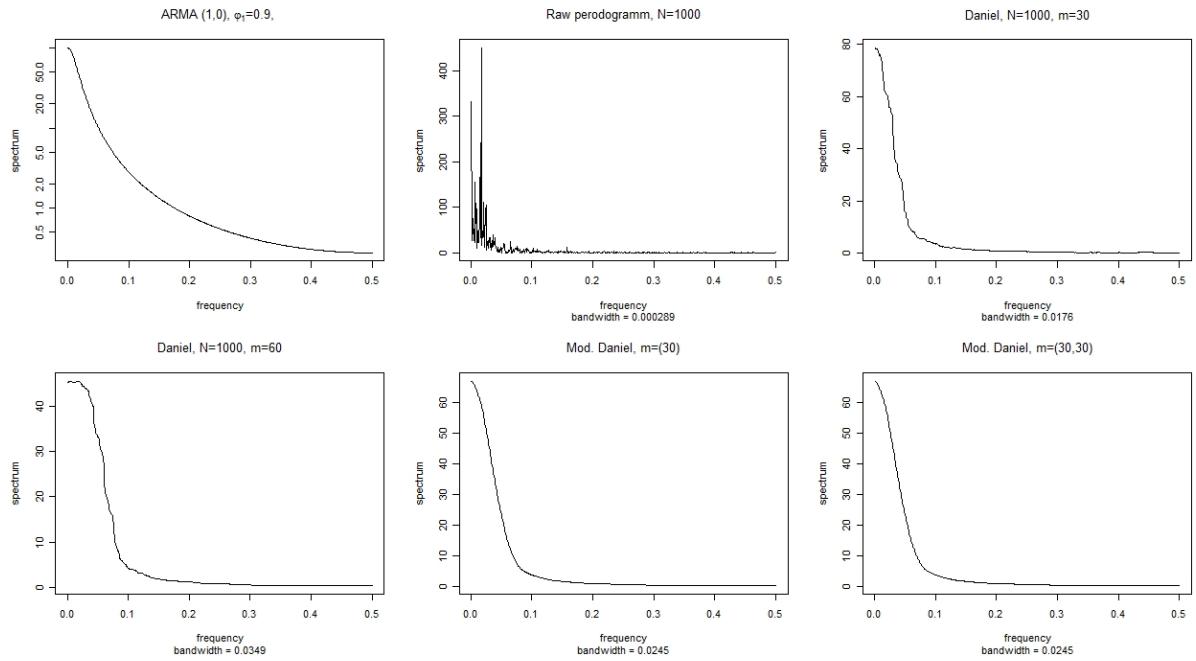
Tagad nedaudz ilustrēsim kodolu gludināšanas metodi ar dažiem piemēriem. Sākumā apskatīsim  $ARMA(p, q)$  procesu simulācijas, līdzīgi kā apskatījām to teorētiskos spektrus jau iepriekš.



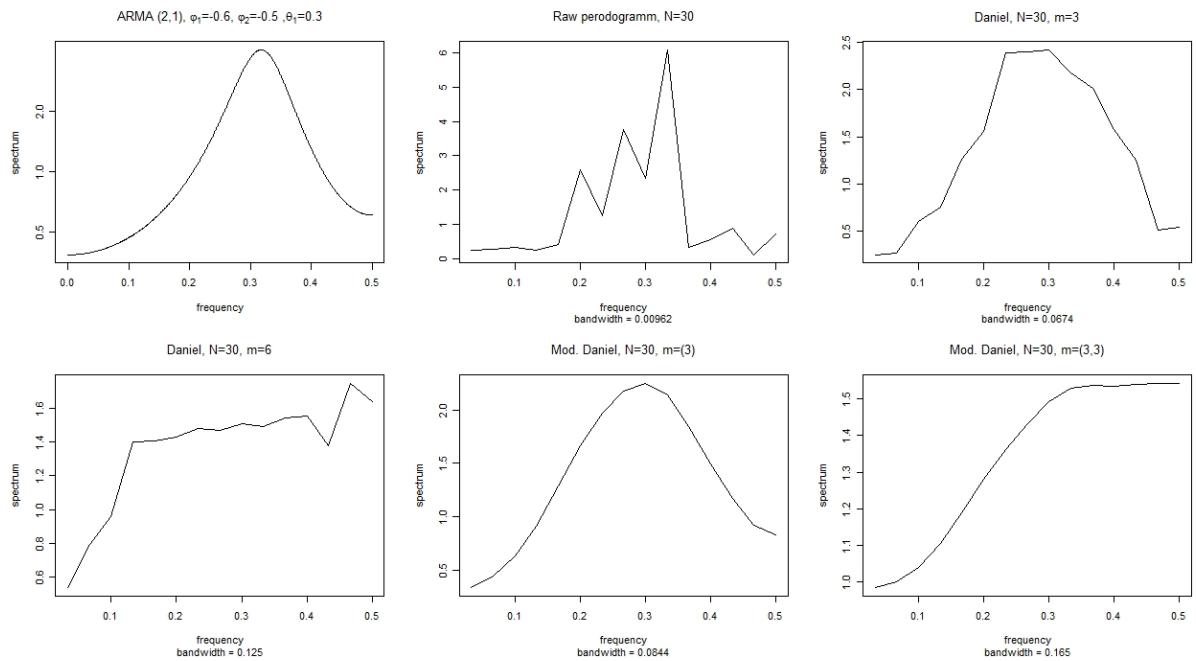
4. att.:  $AR(1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 30$ .



5. att.:  $AR(1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 100$ .

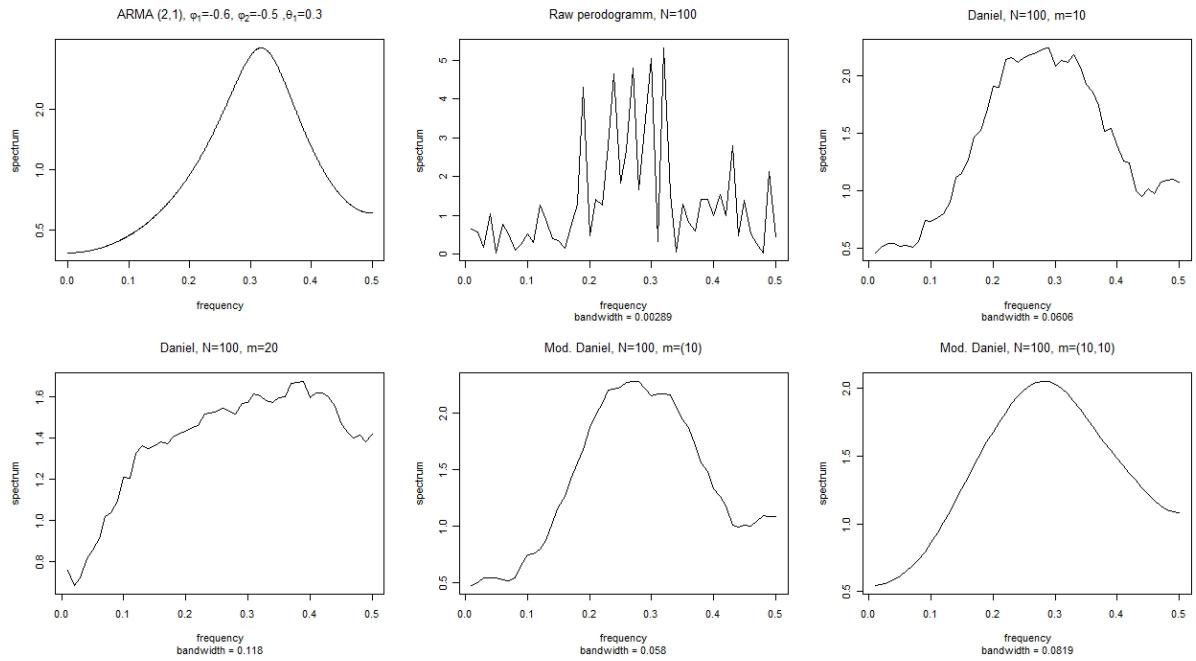


6. att.:  $AR(1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 1000$ .

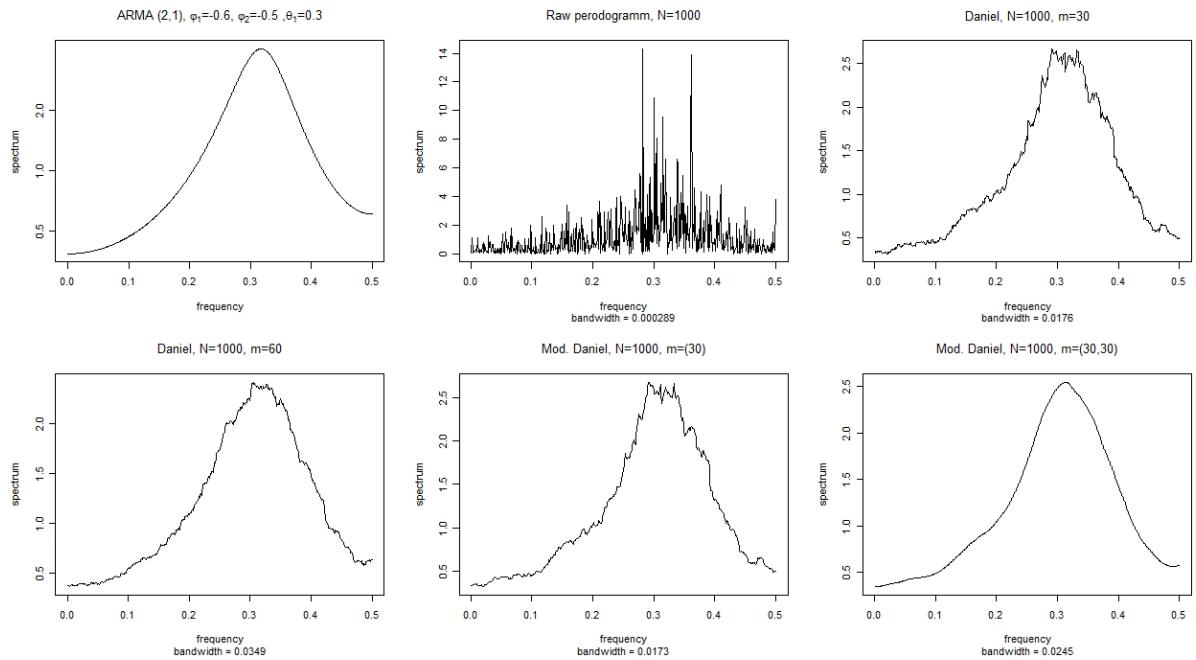


7. att.:  $ARMA(2,1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 30$ .

Kā redzams, negludinātā periodogramma, protams, ir samērā slikts spektra novērtējums, bet jau vienkāršākā gludināšana ar Daniela kodolu un ietekto  $m \approx \sqrt{N}$  dod priekšstatu par spektra formu. Izmantojot modificēto Daniela kodolu vienkāršākajā  $AR(1)$  procesa gadījumā jau iegūst samērā labu spektra novērtējumu. Savukārt jau nedaudz komplikētāks piemērs ar  $ARMA(2,1)$  ir jau interesantāks. Pie izlases apjoma  $N = 30$  Daniela kodols ar  $m \approx \sqrt{N}$  dod kādu priekšstatu par spektra formu, bet otrs ieteiktais variants ar



8. att.:  $ARMA(2,1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 100$ .

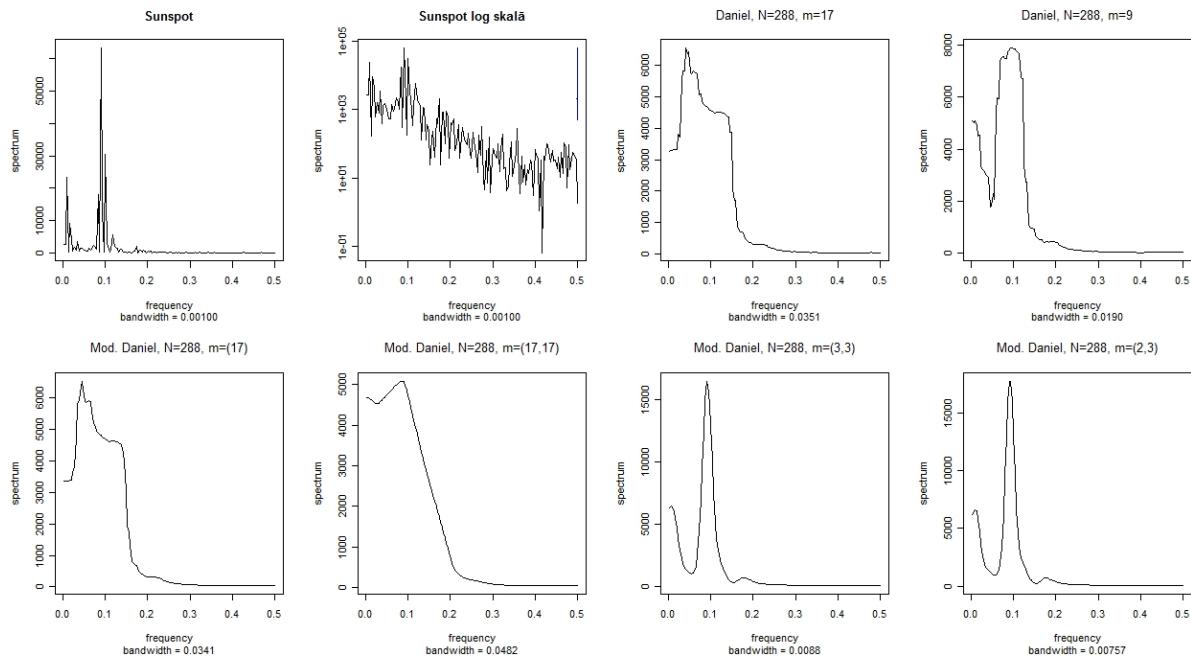


9. att.:  $ARMA(2,1)$  teorētiskais spektrs un simulāciju gludināšana pie izlases apjoma  $N = 1000$ .

$m \approx 2\sqrt{N}$  jau ir *pārgludināts*. Tāpat līdzīgi modifiķētā Daniela kodola gadījumā pie  $m \approx \sqrt{m}$  iegūst jau samērā ļoti labu rezultātu, bet pie  $m \approx 2\sqrt{N}$  atkal periodogramma ir *pārgludināta*. Savukārt  $ARMA(2,1)$  simulācijai pie izlases apjoma  $N = 100$  tiesi atkārtoti pielietots modifiķētais Daniela kodols dod vislabāko rezultātu, līdzīgi arī tas ir, ja izlases apjomu palielina līdz  $N = 1000$ . Interesanti, ka simulētajam procesam pie  $N = 1000$

negludinātā periodogramma vispār nedod priekšstatu par procesa spektru. Patiesībā tas jau bija sagaidāms, jo, kā jau iepriekš atzīmējām, periodogramma nav būtisks spektrālās blīvuma funkcijas novērtējums.

Tagad apskatīsim nedaudz interesantāku piemēru par saules aktivitātes datiem, kas bieži tiek izmantoti laikrindu analīzē (datus var iegūt, piemēram, <http://www.napscience.com/astro/sunspots/sunspotdata.htm>, nēmot datus par 1700.-1987. gadu).



10. att.: Sunspot periodogramma lineārajā un logaritmiskajā skalā, periodogrammas lineārajā skalā gludināšana.

Kā redzams, saules aktivitātes datiem piemērotāka ir lineārā skala (skat. *Logaritmiskā vai lineārā skala periodogrammas attēlošanā*). Izmantojot ieteikumu, ka  $m \approx \sqrt{N}$  vai  $m \approx 0.5\sqrt{N}$  un tad veicot periodogrammas gludināšanu, var ievērot, ka parametrs  $m$  ir izvēlēts pārāk liels. Šis piemērs labi ilustrē jau iepriekš minēto, ka parametra  $m$  noteikšana ir samērā nenoteikta un ka tā jābalsta arī uz iepriekšējo pētnieka pieredzi un *apriorajām* zināšanām par procesu. Var redzēt, ka, pielietojot modificēto Daniela kodolu ar  $m = 3$  atkārtoti (divas reizes), iegūtais tuvinājums ir samērā labi pielāgots datiem, arī gludinot ar modificēto Daniela kodolu ar  $m = 2$  un atkārtoti ar  $m = 3$ , iegūtais rezultāts nav slikts.

## 4.2. Logaritmiskā vai lineārā skala periodogrammas attēlošanā

Apskatot periodogrammu un analizējot procesa spektru samērā bieži vēl tiek apskatītas divas pieejas, kā attēlo periodogrammu, proti, dažkārt periodogrammas vērtības apskata logaritmiskajā skālā. Protams nav grūti aplūkot, piemēram, gludināto periodogrammu

abās - gan lineārajā, gan logaritmiskajā skalā-, bet tomēr būtu būtiski zināt kādu *aprioro* informāciju par procesu. Visbiežāk gan ir tā, ka vienā no izvēlētajām skalām gludinātā periodogramma samērā labi raksturos procesu, bet, izvēloties nepiemērotu skalu, attēlojums būs pārāk *vienveidīgs*, proti, nebūs frekvences, kurās spektra novērtējums būtiski atšķirsies no spektra novērtējuma citā frekvencē. Apskatīsim logaritmiskās (*log*) skalas īpašības. *Log* skala

- lai aplūkotu lielas un mazas amplitūdas signālus vienlaicīgi, proti, log skalā katrai frekvences amplitūdai ir vienāds ”nozīmīgums”;
- lietojot logaritmisko skalu mazi signāli ir vieglāk pamanāmi;
- ja ir kāda liela frekvence, kas ”traucē” pamanīt citas, kas arī ir būtiskas;
- biežāk izmantojama fizikālu eksperimentu analizēšanai.

Savukārt lineārā skala labi parāda frekvenču amplitūdu atšķirības un biežāk izmantojama ekonomisku utml. datu analīzē. Lineārājā skalā arī labi redzams, kurā frekvencē ir izteikts *pīķis*, t.i., kur novērojams kāds būtisks cikls.

### 4.3. Punktveida ticamības intervāli

Zinot periodogrammas īpašības, var konstruēt periodogrammas punktveida ticamības intervālus. Ja izpildās iepriekšminētie nosacījumi par svariem  $h_k$ , turklāt  $m \rightarrow \infty$ , kad  $n \rightarrow \infty$  un  $m/n \rightarrow 0$ , tad var pierādīt (sīkāk skat.[1, 197.-198.lpp, 543.lpp, teorēma C.4] vai [4, 356. lpp.] , ka

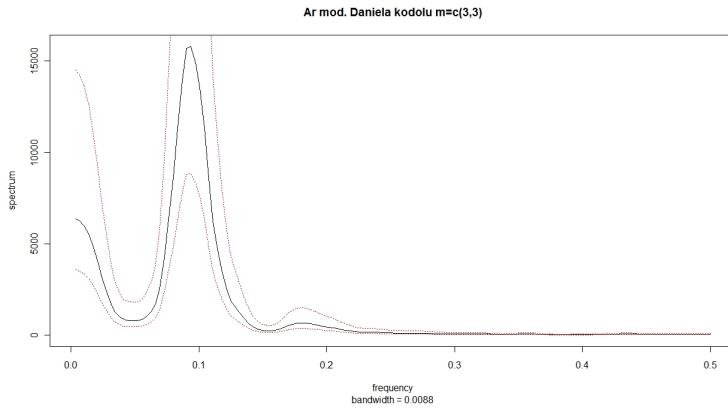
$$\frac{2L_h \hat{f}(\omega)}{f(\omega)} \sim \chi^2_{2L_h}.$$

Tātad ticamības intervāls

$$\frac{2L_h \hat{f}(\omega)}{\chi^2_{2L_h}(1 - \alpha/2)} \leq f(\omega) \leq \frac{2L_h \hat{f}(\omega)}{\chi^2_{2L_h}(\alpha/2)}.$$

Šo ilustrēsim tikai ar vienu piemēru. Atkal izmantosim jau apskatītos datus par saules aktivitāti.

Šeit attēlota jau gludinātā periodogramma ar atbilstošiem punktveida ticamības intervāliem. Kā var redzēt galvenās neskaidrības ir frekvencēs, kur periodogramma pieņem lielas vērtības, proti, tur šie intervāli ir samērā plaši.



11. att.: Sunspot periodogrammas lineārajā skalā divkārša gludināšana ar modificēto Daniela kodolu ar  $m = 3$ .

#### 4.4. Taperošana

Tā kā spektra novērtējumā periodogrammas definēšanai izmatojām diskrēto Furjē transformāciju, tad spektra novērtējumā diezgan būtiskas novirzes dod novērotā procesa sākuma un beigu vērtības. Lai mazinātu šo datu ietekmi, lieto t.s. ”taperošanu”.

Aizvieto novērotos datus  $x_t$  ar  $y_t = h_t x_t$ ,  $t = 1, 2, \dots, n$  un tad pielieto DFT:

$$d_y(\omega_j) = n^{-1/2} \sum_{t=1}^n h_t x_t e^{-2\pi i \omega_j t}.$$

Un tad definē šiem datiem periodogrammu  $I_y(\omega_j) = |d(\omega_j)|^2$ . Būtībā tiek samazināta novērotā procesa sākuma un beigu vērtību ietekme. Tas tiek darīts, jo nav zināms vai process novērots tieši atbilstoši tā ciklu sākumam un beigām, proti, vai novērojumu sākums sakrīt arī ar procesa ciklu sākumu un novērojumi pārtraukti arī noslēdzoties ciklam. Tā kā reāli veicot novērojumus, visticamāk, par procesu tas nav zināms, var gadīties, ka dati fiksēti, kad process atrodas jau kādā cikla fāzē. Tāpēc piemēro sākumu un beigu datu, visbiežāk tieši 10%, datu ietekmes mazināšanu. Piemēram, 10% no datiem sākumā un 10% datiem beigās tiek piekārtoti svari  $h_t$  tā, ka, jo tuvāk datu sākumam/beigām ir novērojums, jo mazāks svara koeficients tam tiek piešķirts. Kā piemērs bieži tiek minēts ([1, 207.lpp] un [4, 360.lpp]) kosinusa zvans (angliski *cosine bell*) vai nošķeltais kosinusa zvans (angliski *split cosine bell*). Kosinusa zvana svari tiek uzdoti pēc likuma

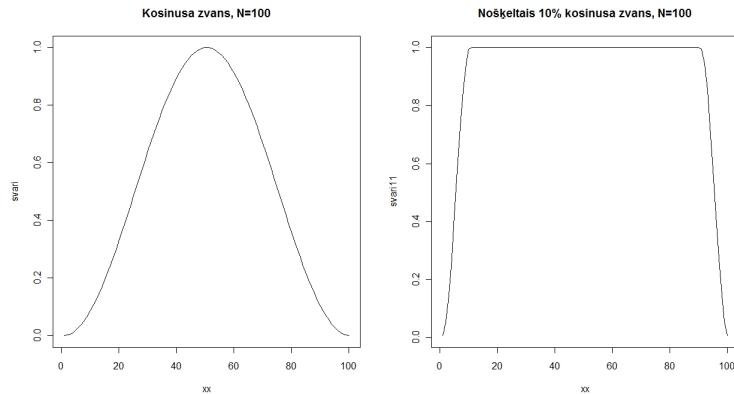
$$h_t = \frac{1}{2} \left( 1 - \cos \left[ \frac{2\pi(t - 0.5)}{n} \right] \right),$$

bet nošķeltais kosinusa zvans pec formulas

$$h_t = \begin{cases} \frac{1}{2} \left( 1 - \cos \left[ \frac{\pi(t-0.5)}{m} \right] \right), & ja 1 \leq t \leq m \\ 1, & ja m+1 \leq t \leq n-m \\ \frac{1}{2} \left( 1 - \cos \left[ \frac{\pi(n-t+0.5)}{m} \right] \right), & ja n-m+1 \leq t \leq n, \end{cases}$$

kurš tiek saukts par  $100p\%$  kosinusa zvana taperošanu, kur  $p = 2m/n$ . Praksē visbiežāk izmanto 10% vai 20% datu taperošanu.

**Piezīme 6.** *Datorprogrammu paketē R automātiskā periodogrammas attēlošanas komanda `spec.pgram()` veic 10% datu taperošanu pēc noklusējuma, kaut gan paziņojumā virs periodogrammas tiek norādīts, ka tā ir negludinātā periodogramma (angļiski Raw periodogram)!*



12. att.: Kosinusa zvana svaru funkcija un 10% nošķeltā kosinusa zvana svaru funkcija datiem ar apjomu  $N = 100$ .

## 5. Neparametriskās regresijas metodes periodogrammas gludināšanā

Fana un Kreutzberges publikācijā [5, ] tiek minēts, ka ir trīs veidi, ko pielieto periodogrammas gludināšanā, turklāt turpat minēts, ka līdz šim tā arī nav noteikts, kur pieeja ir labākā, jo tās ir grūti salīdzināt dažādo gludināšanas procedūru dēļ. Šajā publikācijā viņi paši apskata metodi, kas izmanto Whittle ticamības funkciju spektrālajai blīvuma funkcijai. Kā pirmā pieeja, kas arī tika apskatīta iepriekš, tiek minēta kodolu gludināšana. Kā otrā pieeja tiek minēta logaritmiskās periodogrammas gludināšana ar mazāko kvadrātu metodi. Kā piemērs minēta Wahba (1980), kur tiek izmantoti splaini. Trešā pieeja balstoties uz Whittle (1962) ticamības funkcijas aproksimāciju periodogrammai. Kā arī minēts, ka lielā daļā literatūras tiek izmantota parametriskā pieeja spektra analīzei izmantojot  $ARMA(p, q)$  spektrus. Tad sākumā tiek piemeklēts atbilstošs  $ARMA(p, q)$  process, visbiežāk izmantojot Beijesa informācijas kritēriju (BIC). Bet šādas, parametriskās, pieejas galvenais trūkums ir tas, ka ne visi procesi pakļaujas labai  $ARMA(p, q)$  aproksimācijai.

Viens no veidiem kā gludināt periodogrammu varētu būt izmantojot lokālās regresijas polinomu, bet šeit samērā lielas grūtības sagādā gludinošā parametra noteikšana. Datorprogrammu paketē R pieejama iebūvētā funkcija `locfit`, ko apskatīsim nedaudz vēlāk.

### Lokālā regresija [6, 94.-96. lpp.]

Pieņemsim, ka  $(X_1, Y_1), (X_2, Y_2, \dots, (X_n, Y_n)$  ir neatkarīgi pa pāriem iegūti novērojumi. Sakaribu starp rezultējošo jeb atbildes mainīgo  $Y$  un  $X$  var izteikt ar vienādojumiem

$$Y_i = m(X_i) + \epsilon_i, \quad \mathbb{E}(\epsilon_i) = 0, \quad i = 1, 2, \dots, n,$$

kur  $m$  ir regresijas funkcija.  $m(x)$  var interpretēt kā  $Y$  vidējo vērtību, ja dota vērtība  $X = x$ , tas ir,

$$m(x) = \mathbb{E}(Y|X = x).$$

Aplūkosim nosacītās vidējās vērtības funkcijas  $m(\cdot)$  izvirzījumu. Teilora rindā patvalīgam  $t$  kādā punkta  $x$  apkārtnē

$$m(t) \approx m(x) + m'(x)(t - x) + \dots + m^{(p)}(x)(t - x)^p \frac{1}{p!}.$$

Šis var kalpot par iemeslu, lai aplūkotu lokālo lineāro regresiju (meklēt polinomiālu aproksimāciju punkta  $x$  apkātnē). Ideja par lokālo lineāro regresiju punkta  $x$  apkātnē tiek realizēta izmantojot tā sauktos *kodola svarus* minimizācijas problēmā

$$\min_{\beta} \sum_{i=1}^n \{Y_i - \beta_0 - \beta_1(X_i - x) - \dots - \beta_p(X_i - x)^p\}^2 K_h(x - X_i),$$

kur  $\beta$  ir koeficientu vektors  $(\beta_0, \beta_1, \dots, \beta_p)^T$ . Tātad tiek izmantota svērtā mazāko kvadrātu metode ar svariem  $K_h(x - X_i)$ . Izmantosim apzīmējumus

$$X = \begin{pmatrix} 1 & (X_1 - x) & (X_1 - x)^2 & \dots & (X_1 - x)^p \\ 1 & (X_2 - x) & (X_2 - x)^2 & \dots & (X_2 - x)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (X_n - x) & (X_n - x)^2 & \dots & (X_n - x)^p \end{pmatrix}; \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

$$W = \begin{pmatrix} K_h(x - X_1) & 0 & 0 & \dots & 0 \\ 0 & K_h(x - X_2) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & K_h(x - X_n) \end{pmatrix}.$$

Tad, izmantojot tradicionālo svērto vidējo kvadrātu metodes formulu, iegūsim parametru vektora  $\beta$  novērtējumu

$$\hat{\beta} = (X^T W X)^{-1} X^T W Y.$$

Jāuzsver, ka atšķirībā no novērtējuma, kas iegūts ar parametriskās regresijas palīdzību, šis novērtējums ir atkarīgs no mainīgā  $x$ . Tātad šī patiešām ir lokālā regresija punktā  $x$ . Apzīmēsim vektora  $\hat{\beta}$  komponentes ar  $\hat{\beta}_0(x), \hat{\beta}_1(x), \dots, \hat{\beta}_p(x)$ . Tātad lokālais regresijas funkcijas  $m$  novērtējums ir

$$\widehat{m}_{p,h}(x) = \hat{\beta}_0(x).$$

Tātad  $m(x) \approx \beta_0(x)$ . Vēl tikai atzīmēsim, ka šis novērtējums ir atkarīgs no parametra  $h$ .

Tagad aplūkosim lokālo polinomiālo regresiju dažādiem  $p$

- ja  $p=0$ , tad iegūstam *lokālo konstantes* jeb Nadaraja-Watsona novērtējumu

$$\widehat{m}_{0,h}(x) = \widehat{m}_h(x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)};$$

- pie  $p=1$  iegūstam lineāro lokālās rregresijas novērtējumu. Sākumā apzīmēsim

$$S_{h,j}(x) = \sum_{i=1}^n K_h(x - X_i)(X_i - x)^j,$$

$$T_{h,j}(x) = \sum_{i=1}^n K_h(x - X_i)(X_i - x)^j Y_i,$$

tad iegūsim, ka

$$\widehat{\beta}(x) = \begin{pmatrix} S_{h,0}(x) & S_{h,1}(x) \\ S_{h,1}(x) & S_{h,2}(x) \end{pmatrix} \begin{pmatrix} T_{h,0}(x) \\ T_{h,1}(x) \end{pmatrix}.$$

Līdz ar to esam ieguvuši lokālo lineāro regresijas novērtējumu

$$\widehat{m}_{1,h}(x) = \widehat{\beta}_0(x) = \frac{T_{h,0}(x)S_{h,2}(x) - T_{h,1}(x)S_{h,1}(x)}{S_{h,0}(x)S_{h,2}(x) - S_{h,1}^2(x)}.$$

### Lokālās regresijas ticamības joslas

Lai aplūkotu lokālās regresijas ticamības joslas, ievērosim, ka lokālās regresijas novērtējums ir mazāko kvadrātu metodes atrisinājums un ir lineārs novērtējums. Tad definēsim svaru diagrammas vektoru.

**Apgalvojums 7.** [7, 27. lpp.] Katram  $x$  eksistē svaru diagrammas vektors  $l(x) = (l_1(x), \dots, l_n(x))$ , ka

$$\widehat{m}_{p,h}(x) = \sum_{i=1}^n l_i(x)Y_i.$$

**Apgalvojums 8.** [7, 34. lpp.] Svaru diagrammas vektora koeficientus var aprēķināt pēc formulas

$$l(x)^T = e_1^T (X^T W X)^{-1} X^T W,$$

kur ar  $e_1$  apzīmēts vienības vektors  $e_1 = (1, 0, \dots, 0)^T$ .

Tagad samērā viegli aprēķināt matemātisko cerību un dispersiju.

**Apgalvojums 9.** [7, 28. lpp.] Lokālās regresijas novērtējuma  $\widehat{m}_{p,h}(x)$  matemātiskā cerība un dispersija ir izsakāma ar svaru diagrammas vektoru šādā formā

$$E(\widehat{m}_{p,h}(x)) = \sum_{i=1}^n l_i(x) \widehat{m}_{p,h}(x_i)$$

$$D(\widehat{m_{p,h}}(x)) = \sigma^2 \sum_{i=1}^n l_i(x)^2 =: \sigma^2 \|l(x)\|^2$$

Šeit pieņemts, ka  $Y_i$  ir neatkarīgi un ar konstantu dispersiju  $\sigma^2 < \infty$ .

**Apgalvojums 10.** *Lokālās regresijas novērtējuma dispersija aprēķināma pēc formulas*

$$D(\widehat{m_{p,h}}(x)) = \sigma^2 \|l(x)\|^2 = \sigma^2 e_1^T (X^T W X)^{-1} (X^T W^2 X) (X^T W X)^{-1} e_1.$$

Tagad varam konstruēt ticamības joslas. Apzīmēsim ar  $\bar{m}(x) = E(\widehat{m_{p,h}}(x))$ .

**Apgalvojums 11.** [8, 91.-92. lpp.]  $\bar{m}(x)$  ticamības joslas ir formā

$$I(x) = [\widehat{m_{p,h}}(x) - c\widehat{\sigma}\|l(x)\|, \widehat{m_{p,h}}(x) + c\widehat{\sigma}\|l(x)\|], x \in [a, b]$$

kur  $c$  ir konstante, kas novērtēta no vienādojuma

$$2(1 - \Phi(c)) + \frac{k_0}{\pi} e^{-\frac{c^2}{2}} = \alpha,$$

kur savukārt  $\alpha$  ir uzdotois drošības līmenis un

$$k_0 = \int_a^b \|T'(x)\| dx,$$

kur  $T'(x) = (T'_1(x), \dots, T'_n(x))$ ,  $T'_i(x) = \partial T_i(x) / \partial x$  un  $T_i(x) = l_i(x) / \|l(x)\|$ .

**Teorēma 12.** [8, 85.-86. lpp.] Ja  $\widehat{m_{p,h}}(x)$  ir lineāra regresija, tad

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_i - \widehat{m_{p,h}}(x_i))^2}{n - 2\nu + \tilde{\nu}},$$

kur

$$\nu = \text{tr}(L), \quad \tilde{\nu} = \text{tr}(L^T L) = \sum_{i=1}^n \|l(x_i)\|^2,$$

kur matricas  $L$  elementi ir  $L_{i,j} = l_j(x_i)$ . Turklāt, ja  $m(x)$  ir pietiekami gluda, t.i.,  $\nu = o(n)$  un  $\tilde{\nu} = o(n)$ , tad  $\widehat{\sigma}^2$  ir būtisks novērtējums.

**Piezīme 13.** Ja atlikumu dispersija ir heteroskadiska, tad ticamības joslu konstruēšanā nepieciešams izmantot dispersijas novērtējumu  $s(x)$ , ko šeit neaplūkosim, bet var apskatīt, piemēram, [8, 92.lpp.].

**Piezīme 14.** Datorprogrammu paketē R lokālo regresiju var konstruēt ar paketes `locfit()` palīdzību. Piemēram (`locfit(y~x)`) nozīmē, ka tiek veikta  $y$  lokālā regresija pēc  $x$ . Pēc noklusēšanas tiek izmantota kvadrātiskā regresija, bet, lai veiktu lokālo lineāro regresiju, jānorāda (`locfit(y~x)`), `deg=1` .

**Piezīme 15.** Datorprogrammu paketē R lokālo regresijas ticamības joslas var konstruēt arī ar paketes `locfit()` palīdzību, t.i., jānorāda `fit<-(locfit(y~x))` un tad, attēlojot grafiku ar komandas `plot` palīdzību, norāda `plot(fit, band= "global")`, kur `global` nozīmē, ka atlikumi ir ar konstantu dispersiju. Ja tiek norādīts `band= "local"`, tad tiek izdarīts pieņēmums, ka atlikumu dispersija ir mainīga, t.i.,  $\sigma^2 = \sigma^2(x)$ .

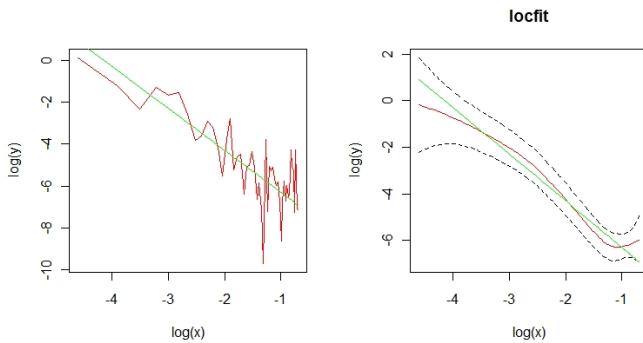
## 6. Trokšņu spektri un analīze

Viens no veidiem kā raksturot dažāda tipa (balto, rozā un sarkano jeb Brauna) trokšņus ir aplūkot to spektrus. Teorētiskais trokšņu spektrs ir  $1/\omega^\beta$ ,  $\beta \geq 0$ .

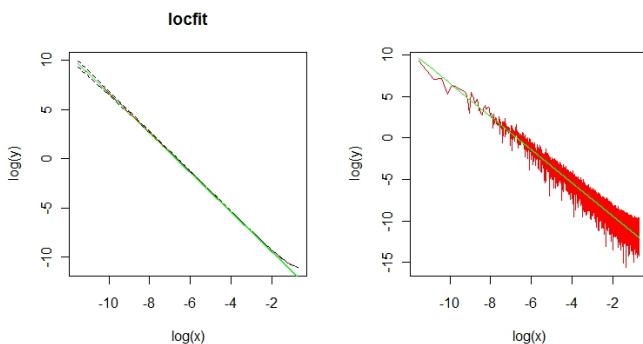
- ja  $\beta = 0$  tad process ir baltais troksnis;
- $\beta = 2$  tad procesam atbilst Brauna troksnis jeb kustība (arī sarkanais troksnis);
- $0 < \beta < 2$ , tad procees ir rozā troksnis.

Skaidrs, ka attēlojot spektru log-log skalā, taisnes virziena koeficients norādīs  $\beta$ .

Tātad viens no veidiem, kā pārbaudīt, vai process nav Brauna kustība, ir konstruēt periodogrammas lokālās regresijas novērtējumu un tad uzzīmējot vienlaicīgās ticamības joslas teorētiskajam spektram. Tātad veikt grafisko pārbaudi. Vienīgi šeit jāpiezīmē, ka Brauna kustība nav stacionārs process, jo tā dispersija laikā aug, tādēļ nepieciešams savādāk definēt šāda procesa spektru.



13. att. Brauna kustības pārbaude,  $N = 100$



14. att. Brauna kustības pārbaude,  $N = 10000$

Tātad, ja veicam šādu pārbaudi un izmantojam paketes R iebūvēto komandu `locfit`, tad tur varam arī konstruēt vienlaicīgās ticamības joslas lokālajai regresijai pār periodogrammu un pārbaudīt, vai log-log skalā taisni ar attiecīgo virziena koeficientu var

ievietot starp vienlaicīgajām ticamības joslām. Kā redzams, simulētā Brauna kustība apmierina šos nosacījumus, kas arī bija sagaidāms.

### Brauna kustības spektrs.

Iepriekš apskatījām vienkāršāko procesa pārbaudi par tā atbilstību Brauna kustībai, kur izmantojām zināmu rezultātu, ka Brauna kustības spektrs (literatūrā bieži tiek saukts arī par jaudas jeb spēka spektru (angliski - *power spectrum*) ir proporcionāls  $\omega^{-2}$ . Šis rezultāts diezgan bieži literatūrā tiek dots, bet arvien atsaucoties uz kāda citu autoru darbu, bet tālāk to neapspriežot, piemēram, nesen izdotajā grāmatā [9, 328.lpp], kur tiesi dota atsuace uz grāmatu [10], kurā dotos rezultātus tagad arī apskatīsim.

**Teorēma 16.** [10, 46. lpp.] *Pienemsim, ka funkcija  $f(t)$  (vispārīgā gadījumā kopleksvērtīga funkcija) apmierina nosacījumu  $\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty$ , tad to var izteikt*

$$f(t) = \int_{-\infty}^{\infty} \phi(\nu) e^{2\pi i \nu t} d\nu,$$

*kur Furjē koeficienti*

$$\phi(\nu) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt.$$

Viegli redzēt, ka  $\phi(-\nu) = \phi(\nu)^*$ , kur \* apzīmē kompleksi saistīto skaitli. Kā arī varam ievērot, ka  $f'(t) = 2\pi i \nu f(t)$ . Tālāk parādīts, kā izmantojot šādā veida pāreju un savādāk definē baltā trokšņa procesu un Brauna kustību. Sākumā atgādināsim Dīraka delta funkcijas definīciju.

**Definīcija 9.** [10, 45. lpp.] Aplūkosim funkciju klasi  $\delta_\varepsilon(t, t')$ , kas atkarīga no parametra  $\varepsilon$ . Ja katrai nepārtrauktai funkcijai  $\psi(t)$

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_\varepsilon(t, t') \psi(t') dt' = \psi(t),$$

tad

$$\delta(t - t') := \lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(t, t')$$

sauc par Dīraka delta funkciju.

**Definīcija 10.** [10, 47. lpp.] Par baltā trokšņa procesu sauc stohastisku procesu, kur gan reālās, gan imaginārās daļas Furjē koeficienti  $\phi(\nu)$  ir ar Gausa sadalījumu un

$$E\phi(\nu) = 0, \quad E\phi(\nu)\phi(\nu')^* = c\delta(\nu - \nu'),$$

kur  $c$  ir konstante. Brauna kustību raksturojot ar Furjē koeficientiem, tiek izmantots, ka tā ir integrēts baltais troksnis un fakts, ko atzīmējām iepriekš, ka Furjē koeficientu integrēšana ekvivalenta to dalīšanai ar  $2\pi\nu i$ . Līdz ar to

$$E\phi(\nu)\phi(\nu')^* = \frac{c}{(2\pi\nu i)(2\pi\nu'i)}\delta(\nu - \nu') = \frac{c}{(2\pi\nu)^2}\delta(\nu - \nu')$$

**Definīcija 11.** [10, 48. lpp.] Par Brauna kustību sauc stohastisku procesu, kur gan reālās, gan imaginārās daļas Furjē koeficienti  $\phi(\nu)$  ir ar Gausa sadalījumu un

$$E\phi(\nu) = 0, \quad E\phi(\nu)\phi(\nu')^* \sim \nu^{-2}\delta(\nu - \nu').$$

Stingri runājot šī definīcija precīzi nesakrīt ar parasti pieņemto, kur  $f(0) = 0$  un  $f(t)$  nav definēta, ja  $t < 0$ . Šeit funkcija definēta visiem  $t$ , bet tā saglabā visas būtiskās Brauna kustības klasiskās definīcijas īpašības.

Kā redzams, tad baltā trokšņa procesa un Brauna kustības Furjē koeficienti nav korlēti, ja vien  $\nu \neq \nu'$ . Šo īpašību aprakstīsim nedaudz savādāk. Izmantosim funkciju  $P(\nu)$

$$E\phi(\nu)\phi(\nu')^* = P(\nu)\delta(\nu - \nu'),$$

kur  $P(\nu)$  ir konstante baltā trokšņa procesam un izturas kā  $\nu^{-2}$  Brauna kustības gadījumā.

**Definīcija 12.** Funkciju  $P(\nu)$  sauksim par procesa spektru, ja tā apmierina iepriekšminētos nosacījumus.

Šāda pieeja dod iespēju definēt  $P(\nu)$  citu procesu gadījumā ar patvaļīgām pakāpes funkcijām.

**Definīcija 13.** [10, 49. lpp.] Par Frakcionālo Brauna kustību sauc stohastisku procesu, kur gan reālās, gan imaginārās daļas Furjē koeficienti  $\phi(\nu)$  ir ar Gausa sadalījumu un

$$E\phi(\nu) = 0, \quad E\phi(\nu)\phi(\nu')^* = P(\nu)\delta(\nu - \nu'),$$

kur

$$P(\nu) \sim |\nu|^{-\beta}.$$

Pakāpi  $\beta$  sauc par *procesa spektrālo pakāpi* jeb *spektrālo eksponenti*.

Tālāk literatūrā [10] tiek piedāvāts apskatīt procesa raksturojošo izvirzījumu galīgas

summas gadījumā

$$f(t) = \int_{-\infty}^{\infty} \phi(\nu) e^{2\pi i \nu t} d\nu = \sum_{j=-n}^n \phi_j(\nu) e^{2\pi i \nu_j t},$$

kur  $\nu_j = j/2n\delta t$ , kur  $\delta t$  atkarīgs no  $f(t)$  un nosaka augstākās frekvences  $\nu_n$  pusperiodu.

Turklāt jaunie koeficienti apmierina

$$E\phi_j = 0, \quad E\phi_j\phi_k^* = \begin{cases} c|v_j|^{-\beta}, & j = k \\ 0, & j \neq k \end{cases},$$

turklāt  $\phi_0 := 0$ . Tālāk apskatīsim laika momentus  $t_k$ , un tad

$$f(t_k) = \sum_{j=-n}^n \phi_j e^{2\pi i \frac{jk}{2n}}.$$

Tālāk tiek apskatīta procesu analīze parametra  $\beta$  noteikšanai, ko veicām jau iepriekš. Tas tiek darīts tā, ka tiek noskaidroti koeficienti  $\phi_j$ , kas izteikti kā inversā funkcija no iepriekš apskatītās formulas

$$\phi_j = \frac{1}{2n} \sum_{k=0}^{2n} f(t_k) e^{-2\pi i \frac{jk}{2n}},$$

kas tieši ir diskrētā Furjē transformācija, kas bija apskatīta jau iepriekš (ar precizitāti līdz apzīmējumiem). Tad aplūkojot grafiku  $|\phi_j|^2$  pret  $\nu_j$  iegūsim tieši periodogrammu. Tātad veiktā pārbaude bija korekta un esam pamatojuši, ka šim grafikam (periodogrammai) jābūt proporcionālai  $|v_j|^{-\beta}$ . Šajā grāmatā gan arī uzsvērts, ka šāda koeficiente  $\beta$  pārbaude lielām izlasēm ir pārāk neprecīza un labāk lietot variogrammas analīzi, ar ko sīkāk var iepazīties [10, 59.-64. lpp.].

Kā alternatīvu pieeju trokšņu analīzē var minēt pāreju uz frakcionālajiem Gausa trokšņiem, kas tiek definēti kā frakcionālās Brauna kustības pieaugumi. Turklāt šādā gadījumā process ir stacionārs un tam var definēt autokovariāciju funkciju. Tālāk tiek piedāvāti dažādi paņēmieni frakcionālās Brauna kustības parametra  $H$  noteikšanai. Kā jau minēts, šo pieeju neapskatīsim. Sīkāk skatīt, piemēram, [11]

## 7. Dispersiju analīze un diskriminantu analīze frekvenču skalā

Viens no interesantākajiem momentiem spektrālajā analīzē ir spektru salīdzināšana un kā to veikt. Šis temats ir jo aktuālaks tāpēc, ka tam ir piemērojami vairāki praktiski pielietojumi. Literatūrā [1, 412.-416.lpp.] kā ievadpiemērs spektru klasifikācijai tiek aprakstīta smadzeņu signālu reakcija uz dažādiem regulāriem kairinājumiem pirms anestēzijas un pēc, kā arī piemērs par sprādzienu un zemestrīču datiem, kur tiek analizēti to spektri un, klasificējot tos grupās, apskata jaunu gadījumu *Novoya Zemlya*, lai noskaidrotu kurai grupai tas pieder. Piemērs par sprādzieniem un zemestrīcēm kalpo kā viens no galvenajiem piemēriem šajā grāmatas sadaļā par spektru klasifikāciju un sava veida salīdzināšanu.

Kā viena no galvenajām metodēm procesu salīdzināšanai tiek apskatīta spektru dispersiju analīze. Sākumā definēsim spektru (spektrālo matricu) stacionāram daudzdimensiju procesam.

**Definīcija 14.** [1, 215.lpp.] Ja  $x_t$  un  $y_t$  ir stacionāras laikrindas un

$$\gamma_{xy}(h) = E[(x_{t+h} - \mu_x)(y_t - \mu_y)]$$

ir jauktā kovariāciju funkciju ar reprezentāciju

$$\gamma_{xy}(h) = \int_{-1/2}^{1/2} f_{xy}(\omega) e^{2\pi i \omega h} d\omega, \quad h = 0, \pm 1, \pm 2, \dots,$$

kur jaukto spektru definē kā Furjē transformāciju

$$f_{xy}(\omega) = \sum_{h=-\infty}^{\infty} \omega_{xy}(h) e^{-2\pi \omega h}, \quad -1/2 \leq \omega \leq 1/2,$$

pieņemot, ka jauktā kovariāciju funkcija, līdzīgi kā autokovariāciju funkcija ir absolūti summējama.

**Definīcija 15.** [1, 217.lpp.] Ja  $p \times p$  dimensiju autokovariāciju funkcija

$$\Gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)^T],$$

kur  $x_t = (x_{t1}, x_{t2}, \dots, x_{tp})^T$  ir  $p$ -dimensionāla laikrinda, kuras elementi apmierina

$$\sum_{h=-\infty}^{\infty} |\gamma_{jk}(h)| < \infty, \forall j, k = 1, \dots, p,$$

tad  $\Gamma(h)$  var izteikt

$$\Gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} f(\omega) d\omega, h = 0, \pm 1, \pm 2, \dots,$$

kas ir kā inversā transformācija **spektrālajai matricai**,  $f(\omega) = \{f_{jk}(\omega)\}$ ,  $j, k = 1, \dots, p$ , kur matrica sastāv no iepriekš definētajām jauktajām spektrālajām komponentēm. Tad matricu  $f(\omega)$  var izteikt

$$f(\omega) = \sum_{-\infty}^{\infty} \Gamma(h) e^{-2\pi i \omega h}, -1/2 \leq \omega \leq 1/2.$$

### Spektra vidējo vienādību (frekvencēs) [1, 438.-450.lpp.]

Bieži vien analizējot reālus datus pietiek ar vidējo vērtību vienādību pārbaudi, kur šoreiz varētu izmantot laikrindu

$$y_{ijt} = \mu_t + \alpha_i t + \nu_{ijt},$$

kur  $i = 1, \dots, I$  apzīmē grupau skaitu,  $j = 1, \dots, N_i$  grupas  $i$  novērojumu (laikrindu) skaitu. Šeit  $\mu_t$  apzīmē visu novērojumu kopējo vidējo vērtību,  $\alpha_{it}$  -  $i$ -tās grupas efektu laika momentā  $t$ , turklāt pieņemam, ka  $\sum_i \alpha_{it} = 0$ . Pie pieņēmuma (hipotēzes  $H_0$ ), ka visu grupu vidējās vērtības laika momentā  $t$  ir vienādas, modeli var pierrakstīt

$$y_{ijt} = \mu_t + \nu_{ijt}.$$

Tagad apskatīsim  $p$ -dimensionālas laikrindas, kur  $y_{ijt} = (y_{ijt1}, y_{ijt2}, \dots, y_{ijtp})^T$ , kas pieder  $i$ -tajai grupai ar  $j = 1, \dots, N_i$  novērojumiem tajā, turklāt visiem grupas vidējā vērtība ir  $\mu_{it}$  un stacionāra autokovariāciju matrica  $\Gamma_i(h)$ . Tad grupu vidējo vērtību vektora diskrēto Furjē transformāciju apzīmēsim ar  $Y_i(\omega_k)$  un  $p \times p$  spektrālās matricas ar  $\hat{f}_i(\omega_k)$ ,  $i = 1, \dots, I$ . Un sastādam klasiskās dispersiju analīzes formulas

$$SPR(\omega_k) = \sum_{i=1}^I \sum_{j=1}^{N_i} (Y_{i.}(\omega_k) - Y_{..}(\omega_k))(Y_{i.}(\omega_k) - Y_{..}(\omega_k))^*,$$

kur \* apzīmē kompleksi saistīto un transponēto jeb Ermita saistīto matricu, bet . apzīmē vidējo vērtību pa attiecīgo indeksu. Līdzīgi

$$SPE(\omega_k) = \sum_{i=1}^I \sum_{j=1}^{N_i} (Y_{ij}(\omega_k) - Y_{i\cdot}(\omega_k))(Y_{ij}(\omega_k) - Y_{i\cdot}(\omega_k))^*.$$

Hipotēzes pārbaudišanai izmanto šādu statistiku

$$\Lambda(\omega) = \frac{|SPE(\omega_k)|}{|SPE(\omega_k) + SPR(\omega_k)|}.$$

Khatri un Hannan ir devuši statistikas asimptotisko sadalījumu

$$\chi^2_{2(I-1)p} = -2(\sum_i N_i - I - p - 1) \ln \Lambda(\omega_k),$$

proti, hī-kvadrāta sadalījums ar  $2(I - 1)p$  brīvības pakāpēm, ja patiesa hipotēze par grupu vidējo vērtību vienādību. Gadījumā  $I = 2$  šī statistika reducējas uz Hottelinga  $T^2$  statistiku, kā to parādījis Giri,

$$T^2 = \frac{N_1 N_2}{(N_1 + N_2)} [Y_{1\cdot}(\omega_k) - Y_{2\cdot}(\omega_k)]^* \hat{f}_\nu^{-1}(\omega_k) [Y_{1\cdot}(\omega_k) - Y_{2\cdot}(\omega_k)],$$

kur

$$\hat{f}_\nu(\omega_k) = \frac{SPE(\omega_k)}{\sum_i N_i - I}$$

ir tā sauktā kopējā (angliski *pooled*) kļūda ar  $I = 2$ . Tātad esam ieguvuši

$$F_{2p, 2(N_1 + N_2 - p - 1)} = \frac{(N_1 + N_2 - 2)p}{(N_1 + N_2 - p - 1)} T^2.$$

Gadījumā, ja tiek aplūkotas viendimensionālas laikrindas, t.i.,  $p = 1$ , iegūsim  $t$ -testa statistiku.

Tāpat lieliu interesu izraisa spektrālo matricu vienādības pārbaude, jo tā kā pieņēmums tiek izmantota vidējo vērtību vienādības pārbaudē. Kā arī šāds pieņēmums tiek izdarīts diskriminantu analizē, kuras neliels piemērs tiks apskatīts vēlāk. Tieka salīdzinātas atsevišķo grupu spektrālās matricas

$$\hat{f}_i(\omega_k) = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (Y_{ij}(\omega_k) - Y_{i\cdot}(\omega_k))(Y_{ij}(\omega_k) - Y_{i\cdot}(\omega_k))^*.$$

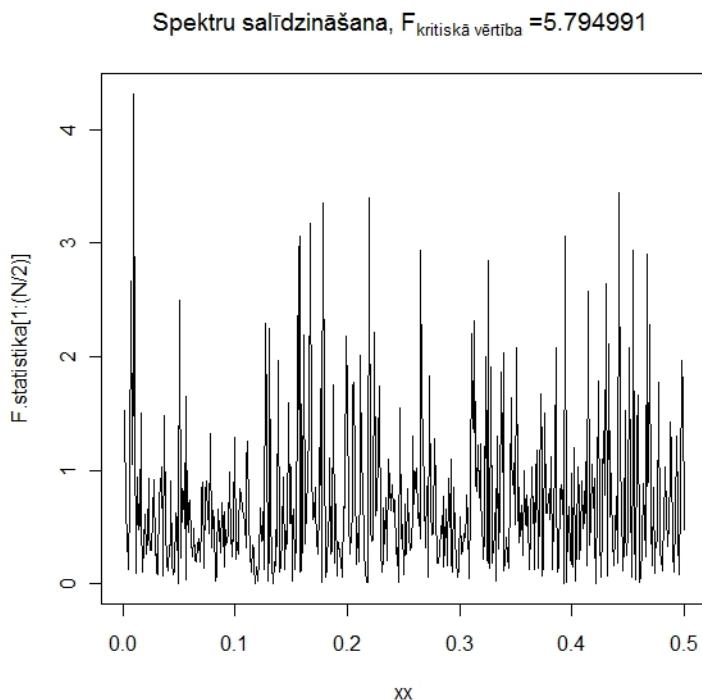
Divu viendimensionālu laikrindu gadījumā testa statistika ir

$$F_{2LM_1,2LM_2} = \frac{\hat{f}_1(\omega)}{\hat{f}_2(\omega)},$$

kur  $M_i = N_i - 1$  un  $L = 2m + 1$ .

## Piemērs

Grāmatā [1] aprakstīts piemērs par zemestrīcēm un sprādzieniem. Šis piemērs tālak analizēts divdimensionālu laikrindu gadījumā. Šajā kursa darbā šos pašus datus analizēsim kā viendimensionālas laikrindas. Sākumā aplūkosim pārbaudi par vidējo vienādību. Atbilstoši sastādam testa statistiku un veicam pārbaudi visām frekvencēm. Šeit jāatzīmē, ka šos testus ērti aplūkot grafiski, jo veicam pārbaudi katrā frekvencē. Ja  $F$ -statistikas vērtība pārsniedz kritisko vērtību kaut vienā frekvencē, hipotēzi par vidējo vērtību vienādību noraidam. Šiem datiem (dati *eqexp.dat* pieejami grāmatas mājaslapā) aplūkojot tos kā viendimensionālas laikrinadas iegūst grafiku, pēc kura redzams, ka vidējo vērtību vienādību nevar noraidīt.



15. att.: Sprādzienu un eksploziju grupu ( $i = 2$ ,  $j = 16$ ,  $p = 1$ ) vidējo vērtību vienādības pārbaude .

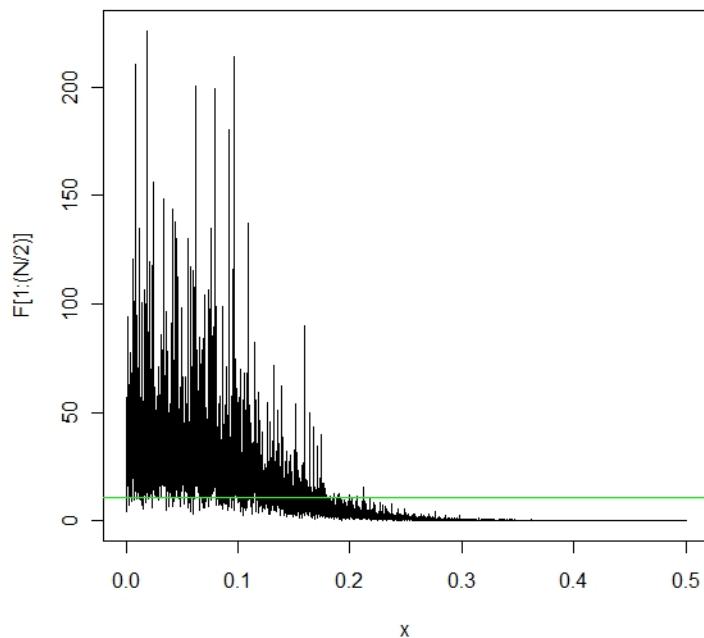
## Piemērs

Apskatīsim vēl vienu piemēru, kur salīdzināsim divu dažādu  $AR(2)$  procesu spektrālās matricas, izmantojot jau aprakstīto formulu

$$F_{2LM_1,2LM_2} = \frac{\hat{f}_1(\omega)}{\hat{f}_2(\omega)},$$

kur  $M_i = N_i - 1$  un  $L = 2m + 1$ . Šajā gadījumā ġenerēti 6  $AR(2)$  procesi ar parametriem  $\phi_1 = 0.9, \phi_2 = -0.3$  un 6  $AR(2)$  procesi ar parametriem  $\phi_1 = -0.9, \phi_2 = -0.3$ . Izveidotas to spektrālās matricas un aprēķināta testa statistikas vērtība katrā frekvencē. Tad, attēlojot grafiski, var redzēt, ka spektrālo matricu vienādība tiek noraidīta, jo testa statistikas vērtības pārniedz kritisko vērtību 10.96702.

**Spektru salīdzināšana**



16. att. Divu procesu grupu spektrālo matricu vienādības pārbaude .

## Diskriminantu analīze frkvenču skalā

Šajā sadaļā tikai dosim pavisam nelielu ieskatu, kā pielietot klasisko diskriminantu analīzi spektriem. Apzīmēsim ar  $X(\omega_k)$  datu diskrētās Furjē transformācijas un pieņemsim, ka to robežsadaliņums ir normālais sadaliņums ar vidējām vērtībām  $M_j(\omega_k)$  un spektrālajām matricām  $f_j(\omega_k)$ , kas pieder  $g$  grupām (populācijām)  $\Pi_1, \Pi_2, \dots, \Pi_g$ . Tātad galvenā problēma ir saklasificēt datus  $g$  apakšgrupās pēc kādas piemērotas pazīmes. Šo problēmu var formalizēt, pieņemot, ka  $X(\omega_k)$  būs ar sadaliņumu  $p_i(X)$ , ja vien dati ir no populācijas  $\Pi_i$ ,  $i = 1, \dots, g$ . Tātad jāsadala apgabalu  $g$  reģionos  $R_1, \dots, R_g$ . Kļūdainas klasifikācijas varbūtība ir

$$P(j|i) = \int_{R_j} p_i(X)d(X).$$

Kopējā kļūdainas klasifikācijas varbūtība sevī ietver arī apriorās varbūtības  $\pi_1, \dots, \pi_g$ , kas raksturo jebkura notikuma varbūtību piederēt katrai no  $g$  grupām. Piemēram, notikumu no grupas  $i$  klasificē kļūdaini grupā  $j$  ar varbūtību  $\pi_i P(j|i)$ . Tātad kopējā kļūdaino klasifikāciju varbūtība ir

$$P_e = \sum_{i=1}^g \pi_i \sum_{i \neq j} P(j|i).$$

Šo kļūdu minimizē nosacījums

$$\frac{p_i(X)}{p_j(X)} > \frac{\pi_j}{\pi_i}$$

visiem  $i \neq j$ . Tātad interesē nosacītā varbūtība

$$P(\Pi_i|X) = \frac{\pi_i p_i(X)}{\sum_j \pi_j(X) p_j(X)},$$

proti notikums tiek klasificēts pie grupas, kurai šī varbūtība ir vislielākā. Ja  $X$  ir ar normālo sadaliņumu pie kā mūsu gadījumā ar vidējo vērtību  $M_j(\omega_k)$  un spektrālajām matricām  $f_j(\omega_k)$ , ja notikums no  $\Pi_j$ , tad aplūko tā sadaliņuma blīvuma funkcijas  $p_j(X)$ . Pēc tam logaritmējot var iegūt atbilstoši funkcijas  $g_j$  [1, 457. lpp.], proti,

$$\begin{aligned} g_j(X) = & \ln \pi_j - \sum_{0 < \omega_k < 1/2} [\ln |f_j(\omega_k)| + X^*(\omega_k) f_j^{-1}(\omega_k) X(\omega_k) \\ & - 2M_j^*(\omega_k) f_j^{-1}(\omega_k) X(\omega_k) + M_j^*(\omega_k) f_j^{-1}(\omega_k) M_j(\omega_k)], \end{aligned}$$

kur tiek summēts pa frekvencēm  $\omega_k$ , kurās  $f(\omega_k) \neq 0$ . Ja tiek izdarīts pieņēmums par vidējo vērtību vienādību (veikta pārbaude), tad varbūtību notikumam piederēt kādai

grupai var vienkāršot [1, 457. lpp.]

$$\ln p_j(X) = \sum_{0 < \omega_k < 1/2} [-\ln|f_j(\omega_k)| - X^*(\omega_k)f_j^{-1}(\omega_k)X(\omega_k)] .$$

Šajā gadījumā notikumu  $X$  klasificē pie grupas  $i$ , ja

$$\ln p_i(X) > \ln p_j(X).$$

Tad divu grupu gadījumā var sastādīt atšķirības mēru [1, 458. lpp.]

$$I(f_1; f_2) = \sum_{0 < \omega_k < 1/2} \left[ \text{tr}\{f_1(\omega_k)f_2^{-1}(\omega_k)\} - \ln \frac{|f_1(\omega_k)|}{|f_2(\omega_k)|} - p \right] ,$$

kur ar  $\text{tr}\{\cdot\}$  apzīmēta matricas pēda un  $p$  nosaka laikrindas dimensiju.

### Piemērs.

Grāmatā [1] aprakstīts piemērs par zemestrīcēm un sprādzieniem. Šis piemērs tiek tālāk analizēts arī diskriminantu analīzē divdimensionālu laikrindu gadījumā. Šajā kursa darbā atkal šos pašus datus analizēsim kā viendimensionālas laikrindas. Apskatot kritēriju  $I(f_1; f_2)$  lineārajā diskriminantu analīzē pie pieņēmua, ka grupu vidējās vērtības ir vienādas. Iegūtās vērtības apkopotas tabulā. Pozitīva vērtība norāda par klasifikāciju pie pirmās grupas, t.i., zemestrīcēm, attiecīgi negatīva par piederību sprādzienu grupai.

1. tabula Atšķirības mēra vērtība

Nr.	1	2	3	4	5	6	7	8
Zemestrīces (eq)	-0.51973	3.77204	-0.52363	-0.57132	-0.56735	8.95826	10.25207	-0.55733
Sprādzieni (exp)	-0.43851	-0.44826	-0.37423	-1.04176	-1.58136	-0.56848	-0.87415	-0.51279

## 8. Secinājumi

Šajā kursa darbā tika apskatīti spektrālās analīzes pamatjēdzieni un vienkāršākie lietojumi, un to teorētiskie apraksti. Katrā kursa darba sadaļā mēģināts ilustrēt attiecīgo tēmu ar kādu piemēru un sastādīts atbilstošs programmas R kods, kas ietverts kursa darba pielikumā. Piemēru analīze galvenokārt veikt ilustratīvos nolūkos, tādēļ to rezultātiem nav tik būtiskas nozīmes. Tā kā spektrālā analīze ir samērā plaša matemātikas (kā arī fizikas un finanšu matemātikas) disciplīna, tad būtu nepieciešams vairāk laika kā arī lielāka specializācija kādā no spektrālās analīzes apakštēmām, lai to izklāstītu pilnīgāk un aprakstītu konkrētākus un plašākus spektrālās analīzes pielietojumus. Turklat, tā kā spektrālās analīzes pamatā ir Furjē analīze (un diskrētā Furjē transformācija), teorētisko formulēju mu izpratne prasa samērā daudz laika un labas priekšzināšanas. Jāatzīmē, ka, pētot šo tēmu, radusies pārliecība, ka konkrētas realizācijas analīzei lieti noderēs šāds ieskats pamatjautājumos, un tad, konkrētāk specializējoties kādā no tēmām, varēs precīzāk un ātrāk risināt konkrēto uzdevumu.

# Izmantotā literatūra un avoti

- [1] Stoffer D.S Shumway R.H. *Time Series Analysis and Its Applications with R Examples 2nd edition.* Springer Science+Business Media, LLC, New York, 2006.
- [2] Metcalfe A.V. Cowpertwait P.S.P. *Time series analysis.* Princeton University Press, New Jersey, 1994.
- [3] Hamilton J.D. *Introductory Time Series with R.* Springer Science+Business Media, LLC, New York, 2009.
- [4] Chan K.S. Cryer J.D. *Time Series Analysis with Applications in R 2nd edition.* Springer Science+Business Media, LLC, New York, 2008.
- [5] Kreuzberger E. Fan J. Automatic local smoothing for spectral density estimation. *Scandinavian Journal of Statistics*, 25:359–369, June 1998.
- [6] Sperlich S. Werwatz A. Hardle W., Muller M. *Nonparametric and semiparametric models.* Springer-Verlag Berlin Heidelberg, Heidelberg, 2004.
- [7] Loader C. *Local Regression and Likelihood.* Springer-Verlag New York Inc., New York, 1999.
- [8] Wasserman L. *All of Nonparametric Statistics.* Springer Science + Business Media, LLC, New York, 2006.
- [9] Aschwanden M. *SELF-ORGANIZED CRITICALITY IN ASTROPHYSICS: The Statistics of Nonlinear Processes in the Universe.* Springer-Verlag Berlin Heidelberg, Heidelberg, 2011.
- [10] Hergarten S. *SELF-ORGANIZED CRITICALITY IN EARTH SYSTEMS.* Springer-Verlag Berlin Heidelberg, Heidelberg, 2002.
- [11] Dieker T. Simulation of fractional brownian motion. Master's thesis, University of Twente Department of Mathematical Sciences, Amsterdam, 2002.

# A Izveidoto programmu kods

## A1. Periodogrammas gludināšanas programmas kods

```
###  
###kodolu gludinata periodogramma kursa darba piemēriem  
###iebūvētās f-jas  
source("itall2.R")  
##koeficienti ar + zīmi visiem : gan AR gan MA kā kursa darbā  
par(mfrow=c(2,3))  
  
theta<- c(-0.6,-0.5)  
N<-100  
AR1m<-arima.sim(n = N, list(ar = c(theta), ma=0.3 ))  
##  
  
spec.arma(ar=c(-.6,-.5), ma=0.3, Nosaukums=expression  
("ARMA (2,1), "*phi1[1]*"=-0.6, "*phi1[2]*"=-0.5 , "*theta[1]*"=0.3"))  
spec.pgram(AR1m,taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Raw periodogramm, N=100"))  
spec.pgram(AR1m,kernel("daniell",10), taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Daniel, N=100, m=10"))  
spec.pgram(AR1m,kernel("daniell",20), taper=0, fast=FALSE, detrend=FALSE, log="no"  
,main=expression("Daniel, N=100, m=20"))  
spec.pgram(AR1m,kernel("modified.daniell",c(10)), taper=0, fast=FALSE, detrend=FALSE,  
log="no",main=expression("Mod. Daniel, N=100, m=(10)"))  
spec.pgram(AR1m,kernel("modified.daniell",c(10,10)), taper=0, fast=FALSE, detrend=FALSE,  
log="no",main=expression("Mod. Daniel, N=100, m=(10,10)"))  
  
theta<- c(-0.6,-0.5)  
N<-1000  
AR1m<-arima.sim(n = N, list(ar = c(theta), ma=0.3 ))  
##  
  
spec.arma(ar=c(-.6,-.5), ma=0.3, Nosaukums=expression("ARMA (2,1), "*phi1[1]*"=-0.6, "*phi1[2]*"=-0.5  
,"*theta[1]*"=0.3))  
spec.pgram(AR1m,taper=0, fast=FALSE, detrend=FALSE, log="no",main=expression("Raw periodogramm, N=1000"))  
  
spec.pgram(AR1m,kernel("daniell",30), taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Daniel, N=1000, m=30"))  
spec.pgram(AR1m,kernel("daniell",60), taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Daniel, N=1000, m=60"))  
spec.pgram(AR1m,kernel("modified.daniell",c(30)), taper=0, fast=FALSE, detrend=FALSE,  
log="no",main=expression("Mod. Daniel, N=1000, m=(30)"))  
spec.pgram(AR1m,kernel("modified.daniell",c(30,30)), taper=0, fast=FALSE, detrend=FALSE,  
log="no",main=expression("Mod. Daniel, N=1000, m=(30,30)"))  
  
theta<- c(-0.6,-0.5)  
N<-30  
AR1m<-arima.sim(n = N, list(ar = c(theta), ma=0.3 ))  
##  
  
spec.arma(ar=c(-.6,-.5), ma=0.3, Nosaukums=expression("ARMA (2,1), "*phi1[1]*"=-0.6, "*phi1[2]*"=-0.5  
,"*theta[1]*"=0.3))  
spec.pgram(AR1m,taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Raw periodogramm, N=30"))  
spec.pgram(AR1m,kernel("daniell",3), taper=0, fast=FALSE, detrend=FALSE, log="no"  
,main=expression("Daniel, N=30, m=3"))  
spec.pgram(AR1m,kernel("daniell",6), taper=0, fast=FALSE, detrend=FALSE, log="no",  
main=expression("Daniel, N=30, m=6"))
```

```

spec.pgram(AR1m,kernel("modified.daniell",c(3,3)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=30, m=(3)"))
spec.pgram(AR1m,kernel("modified.daniell",c(6,6)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=30, m=(3,3)"))

spec.pgram(dati[2], taper=0, fast=FALSE, detrend=FALSE, log="no",main="Sunspot") #Istā automātiskā

```

## A2. Kosinusa zvans

```

###cosine bell attēlam

xx<-seq(1,100,1)
svari<- (1/2)*(1-cos((2*pi*(xx-(1/2)))/100))

svari1<-(1/2)*(1-cos((pi*(xx[1:10]-(1/2))/10)))
svari2<-seq(1,1,length=80)
svari3<-(1/2)*(1-cos((pi*(100-xx[91:100]+(1/2))/10)))
svari11<-c(svari1,svari2,svari3)
par(mfrow=c(1,2))
plot(xx,svari,type="l", main= "Kosinusa zvans, N=100")
plot(xx,svari11,type="l", main= "Nošķeltais 10% kosinusa zvans, N=100")

```

## A3. Sunspot datu analīzes un ticamības intervālu programmas kods

```

### Sunspot datu periodogramma salīdzināšanai ar matlab periodogrammu

####iebūvētās f-jas
source("itall.R")
##
par(mfrow=c(2,4))
dati<-read.table(file="sunspot.txt",header=F)
sqrt(288)

spec.pgram(dati[2], taper=0, fast=FALSE, detrend=FALSE, log="no",main="Sunspot") #Istā automātiskā
spec.pgram(dati[2], taper=0, fast=FALSE, detrend=FALSE, log="yes",main="Sunspot log skalā") #Istā automātiskā
spec.pgram(dati[2],kernel("daniell",17), taper=0, fast=FALSE, detrend=FALSE, log="no",
main=expression("Daniel, N=288, m=17"))
spec.pgram(dati[2],kernel("daniell",9), taper=0, fast=FALSE, detrend=FALSE, log="no"
,main=expression("Daniel, N=288, m=9"))
spec.pgram(dati[2],kernel("modified.daniell",c(17)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=288, m=(17)"))
spec.pgram(dati[2],kernel("modified.daniell",c(17,17)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=288, m=(17,17)"))
spec.pgram(dati[2],kernel("modified.daniell",c(3,3)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=288, m=(3,3)"))
spec.pgram(dati[2],kernel("modified.daniell",c(2,3)), taper=0, fast=FALSE, detrend=FALSE,
log="no",main=expression("Mod. Daniel, N=288, m=(2,3)"))

####tic. intervali
par(mfrow=c(1,1))
gludinata.2<-spec.pgram(dati[2],kernel("modified.daniell",c(3,3)),log="no",
main="Ar mod. Daniela kodolu m=c(3,3)") #ar kodolu
df<-gludinata.2$df
##Hī kvadrāta kritiskās vērtības
U<-qchisq(0.025,df)
L<-qchisq(0.975,df)
u.bound<-df*gludinata.2$spec/U
l.bound<-df*gludinata.2$spec/L
lines(gludinata.2$freq,u.bound,lty="dotted", col="red")
lines(gludinata.2$freq,l.bound,lty="dotted", col="red")

##### tic. joslas
x<-spec.pgram(dati[2], taper=0, fast=FALSE, detrend=FALSE, log="no",main="Sunspot")$freq

```

```

y<-spec.pgram(dati[2], taper=0, fast=FALSE, detrend=FALSE, log="no", main="Sunspot")$spec

library(KernSmooth)
h<-dpill(x,y) #svariķa h izvēle
lines(ksmooth(x, y, kernel = c("normal"), bandwidth =h), col="red", type="lty")
plot(locfit(y~x), band="global", col="green")

```

## A4. ARMA procesu teorētisko spektru programmas kods

```

####iebūvētās f-jas
source("itall12.R")
##koeficienti ar + zīmi visiem : gan AR gan MA kā kursa darbā
par(mfrow=c(2,3))

spec.arma(ar=c(-.6,-0.5), ma=0.1,
Nosaukums=expression("ARMA (2,1), \"*phi1[1]*\"=-0.6, \"*phi1[2]*\"=-0.5 , \"*theta[1]*\"=0.1 "))

```

## A5. Brauna kustības pārbaudes programmas kods

```

library(locfit)
library(sde)
source("itall.R")

par(mfrow=c(2,2))

Brauna.kustiba<-BM(N=10000)[1:10000]
plot(Brauna.kustiba)
##īstās (bez pārveidojumiem) :
spec.pgram(Brauna.kustiba, taper=0, fast=FALSE, detrend=FALSE, log="yes")
# log skalā īstās:
spec.pgram(Brauna.kustiba,taper=0, fast=FALSE, detrend=FALSE, log="yes")
x<-spec.pgram(Brauna.kustiba,taper=0, fast=FALSE, detrend=FALSE, log="yes")$freq
y<-spec.pgram(Brauna.kustiba,taper=0, fast=FALSE, detrend=FALSE, log="yes")$spec
##x<-x[-1]; y<-y[-1]
plot(locfit(log(y)^log(x)), band="global", col="red",main="locfit")

plot(log(x),log(y),type="l", col="red")
lines(log(x),log(0.0000035/(x^2)), col="green" )

#####

```

## A6. Vidējo vērtību pārbaudes programmas kods

```

####iebūvētās f-jas
source("itall.R")
##
nser=17;
# Series Dimension
np=2;
# Series Length
n=1024
# Number of frequencies out to the folding frequency
n2=n/2+1

x = matrix(scan("eqexp.dat"), nrow=1024 ,ncol=nser*np)

x;
x1<-x[,1:16]
x2<-x[,16:34]

spec.pgram(x[,25],log="no")
plot(1:1024,x[,28],type="l")

```

```

N<-1024
xx<-spec.pgram(x[,1],kernel("modified.daniell",c(round(sqrt(N)),round(sqrt(N)))), log="no")$freq
yy<-spec.pgram(x[,1],kernel("modified.daniell",c(round(sqrt(N)),round(sqrt(N)))), log="no")$spec

plot(xx,yy,col="red",type="l")
#####
#####ACF

xxx<-spec.pgram(x[,1],kernel("modified.daniell",c(sqrt(N),sqrt(N))), log="no")$spec

acf(xxx,lag.max=3000)

#####
#####
par(mfrow=c(1,1))
##N_i=16 (N=100 vai 1000, vai 10 000)
Ni<-16
##DFT no x[,j] j=1,2,...,16 apzimee ar y1

y1<-matrix(x[,1:16],nrow = 1024, ncol=16, byrow=FALSE)
for (i in 1:16)
{
y1[,i]<-fft(x[,i])/(sqrt(N))
}

#### vidējo vērtību vektoru DFT Y_i.
aa<-c((1:16)/(1:16))
sum(aa)
x.matri1<-matrix(x[,1:16],nrow = 1024, ncol=16, byrow=FALSE)
vid<-(aa%*%t(x.matri1))/16
Y_1.<-fft(vid)/sqrt(N)

## SPE## DFT matrica
## DFT matrica i=1
matr_1<-(y1)
matr_2<-cbind( c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),
c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.) )

matr_Y<-matr_1-matr_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvencees:
rezul_matr<-(matr_Y*Conj(matr_Y))
SPE.1<-c()
for (i in 1:N)
{
SPE.1[i]<-( sum(t(rezul_matr[i,])) ) #neists reizinaajums, bet ta vajag
}
SPE.1<-abs(SPE.1)
#plot(x,f_1_2[1:(N/2)])
#lines(x,f_1_2[1:(N/2)])

#####
#####otram tas pats - iegūst otru spektrālo matricu:
##DFT no x[,j] j=1,2,...,16 apzimee ar y2

y2<-matrix(x[,16:32],nrow = 1024, ncol=16, byrow=FALSE)
for (i in 1:16)
{
y2[,i]<-fft(x[,i+15])/(sqrt(N))
}

#### vidējo vērtību vektoru DFT Y_i.
aa<-c((1:16)/(1:16))
sum(aa)
x.matri2<-matrix(x[,16:32],nrow = 1024, ncol=16, byrow=FALSE)
vid<-(aa%*%t(x.matri2))/16
Y_2.<-fft(vid)/sqrt(N)

## SPE## DFT matrica

```

```

## DFT matrica i=1
matr_1<-c(y2)
matr_2<-cbind( c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),
c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),
c(Y_2.),c(Y_2.) )

matr_Y<-matr_1-matr_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvences:
rezul_matr<-(matr_Y*Conj(matr_Y))
SPE.2<-c()
for (i in 1:N)
{
SPE.2[i]<-( sum(t(rezul_matr[i,])) ) #neists reizinaajums, bet ta vajag
}
SPE.2<-abs(SPE.1)
#plot(x,f_1_2[1:(N/2)])
#lines(x,f_1_2[1:(N/2)])

SPE<-SPE.1+SPE.2

#####
SPE

#####tests:
T_2<- ((Ni*Ni)/(Ni+Ni))*( (SPE/(Ni+Ni-2))^( -1) ) * abs(Y_1.-Y_2.)^2
T_2<-abs(T_2)
p<-1
F.statistika<-((Ni+Ni-2)*p) / (Ni+Ni-p-1)*(T_2)
##uzīmē statistiku:
plot(xx,F.statistika[1:(N/2)], main="Spektru salīdzināšana ", type="l")

alpha<-0.01
F_stat_kr_value<-qf(1-alpha/2,2*p,2*(Ni+Ni-p-1))
F_stat_kr_value;
abline(F_stat_kr_value,0,col="green")
plot(xx,F.statistika[1:(N/2)],
main=expression("Spektru salīdzināšana, " * F[kritiskā] * " *vērtība]" = 5.794991 ), type="l")

##vajadzētu būt, ka ir ok. bet es nēmu kā vienvarianta (vienvērtīgas laikrindas)

```

## A7. Spektrālo matricu AR(2) procesiem salīdzināšanas programmas kods

```

#####
#####

###
AR1<-arima.sim(n = N, list(ar = theta, ma=0 ))
AR2<-arima.sim(n = N, list(ar = c(theta), ma=0 ))
AR3<-arima.sim(n = N, list(ar = c(theta), ma=0 ))
AR4<-arima.sim(n = N, list(ar = c(theta), ma=0 ))
AR5<-arima.sim(n = N, list(ar = c(theta), ma=0 ))
AR6<-arima.sim(n = N, list(ar = c(theta), ma=0 ))

##N_i=6 (N=100 vai 1000, vai 10 000)
Ni<-6
##DFT no AR_j j=1,2,...,6 apzimee ar Y_i
Y_11<-fft(AR1)/sqrt(N)
Y_12<-fft(AR2)/sqrt(N)
Y_13<-fft(AR3)/sqrt(N)
Y_14<-fft(AR4)/sqrt(N)
Y_15<-fft(AR5)/sqrt(N)
Y_16<-fft(AR6)/sqrt(N)

### vidējo vērtību vektoru DFT Y_i.
vid_AR<-(AR1+AR2+AR3+AR4+AR5+AR6)/6
Y_1.<-fft(vid_AR)/sqrt(N)

```

```

## spekraalaas matricas f_i (šoreiz skaitli)
## DFT matrica i=1
matrica_1<-matrix(c(Y_11,Y_12,Y_13,Y_14,Y_15,Y_16),nrow=Ni,ncol=N,byrow=T)
matrica_2<-matrix(c(Y_1.,Y_1.,Y_1.,Y_1.,Y_1.,Y_1.),nrow=Ni,ncol=N,byrow=T)
matr_Y<-matrica_1-matrica_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvencees:
rezul_matr<-(matr_Y*Conj(matr_Y))
f_1<-c()
for (i in 1:N)
{
f_1[i]<-(1/(Ni-1))*( sum(t(rezul_matr[,i])) ) #neists reizinaajums, bet ta vajag
}
f_1_2<-abs(f_1)
plot(x,f_1_2[1:(N/2)])
lines(x,f_1_2[1:(N/2)])

#####
#####otram tas pats - ģenerē otru, iegūst otru spektrālo matricu:

theta2<-c( -0.9,-.3)
##
AR1<-arima.sim(n = N, list(ar = theta2, ma=0 ))
AR2<-arima.sim(n = N, list(ar = c(theta2), ma=0 ))
AR3<-arima.sim(n = N, list(ar = c(theta2), ma=0 ))
AR4<-arima.sim(n = N, list(ar = c(theta2), ma=0 ))
AR5<-arima.sim(n = N, list(ar = c(theta2), ma=0 ))
AR6<-arima.sim(n = N, list(ar = c(theta2), ma=0 ))

##N_i=6 (N=100 vai 1000, vai 10 000)
Ni<-6
##DFT no AR_j j=1,2,...,6 apzimee ar Y_i
Y_11<-fft(AR1)/sqrt(N)
Y_12<-fft(AR2)/sqrt(N)
Y_13<-fft(AR3)/sqrt(N)
Y_14<-fft(AR4)/sqrt(N)
Y_15<-fft(AR5)/sqrt(N)
Y_16<-fft(AR6)/sqrt(N)

### vidējo vērtību vektoru DFT Y_i.
vid_AR<-(AR1+AR2+AR3+AR4+AR5+AR6)/6
Y_1.<-fft(vid_AR)/sqrt(N)

## spekraalaas matricas f_i (šoreiz skaitli)
## DFT matrica i=1
matrica_1<-matrix(c(Y_11,Y_12,Y_13,Y_14,Y_15,Y_16),nrow=Ni,ncol=N,byrow=T)
matrica_2<-matrix(c(Y_1.,Y_1.,Y_1.,Y_1.,Y_1.,Y_1.),nrow=Ni,ncol=N,byrow=T)
matr_Y<-matrica_1-matrica_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvencees:
rezul_matr<-(matr_Y*Conj(matr_Y))
f_2<-c()
for (i in 1:N)
{
f_2[i]<-(1/(Ni-1))*( sum(t(rezul_matr[,i])) ) #neists reizinaajums, bet ta vajag
}
f_2_2<-abs(f_2)
plot(x,f_2_2[1:(N/2)])
lines(x,f_2_2[1:(N/2)])

#####
#####

#####teorētiskais:
###f_2<-spec.arma(ar=theta_teo, ma=0,n.freq=N)$spec
###F<-f_1_2/f_2

F<-f_1_2/f_2_2 ## (7.104)

```

```

plot(x,F[1:(N/2)], main="Spektru salīdzināšana ", type="l")
##lines(x,F[1:(N/2)])
alpha<-0.01
F_stat_kr_value<-qf(1-alpha,Ni-1,Ni-1)
F_stat_kr_value;
abline(F_stat_kr_value,0,col="green")

```

## A8. Diskriminanta analīzes programmas kods

```

####iebūvētās f-jas
source("itall.R")
##
nser=17;
# Series Dimension
np=2;
# Series Length
n=1024
# Number of frequencies out to the folding frequency
n2=n/2+1

x = matrix(scan("eqexp.dat"),nrow=1024 ,ncol=nser*np)

x;

#x1<-x[,1:16]
#x2<-x[,16:34]
x1<-cbind(x[,1], x[,3], x[,5], x[,7], x[,9], x[,11], x[,13], x[,15]) #p komponentes
x2<-cbind(x[,2], x[,4], x[,6], x[,8], x[,10], x[,12], x[,14], x[,16]) #s komponentes
ex1<-cbind(x[,17], x[,19], x[,21], x[,23], x[,25], x[,27], x[,29], x[,31]) #p komponentes
ex2<-cbind(x[,18], x[,20], x[,22], x[,24], x[,26], x[,28], x[,30], x[,32]) #p komponentes

NZ1<-cbind(x[,33])
NZ2<-cbind(x[,34])

spec.pgram(x[,25],log="no")
par(mfrow=c(1,2))
plot(1:1024,NZ1,type="l")
plot(1:1024,NZ2,type="l")
N<-1024
xx<-spec.pgram(ex1[,5],kernel="modified.daniell",c(round(sqrt(N)),round(sqrt(N)))), log="no")$freq
yy<-spec.pgram(ex2[,5],kernel("modified.daniell",c(round(sqrt(N)),round(sqrt(N)))), log="no")$spec

plot(xx,yy,col="red",type="l")
#####
#####ACF

xxx<-spec.pgram(x[,1],kernel("modified.daniell",c(sqrt(N),sqrt(N))), log="no")$spec

acf(xxx,lag.max=3000)

#####
#####

##N_i=8
Ni<-8
##DFT no x[,j] j=1,2,...,8 apzimee ar y1

y1<-matrix(x[,1:8],nrow = 1024, ncol=8, byrow=FALSE)
for (i in 1:8)
{
y1[,i]<-fft(x1[,i])/(sqrt(N))
}

### vidējo vērtību vektoru DFT Y_i.
aa<-c((1:8)/(1:8))
sum(aa)
x.matri1<-matrix(x1[,1:8],nrow = 1024, ncol=8, byrow=FALSE)
vid<-(aa%*%t(x.matri1))/8
Y_1.<-fft(vid)/sqrt(N)

```

```

##  SPE## DFT matrica
## DFT matrica i=1
matr_1<-(y1)
matr_2<-cbind( c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),c(Y_1.),
c(Y_1.),c(Y_1.) )

matr_Y<-matr_1-matr_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvences:
rezul_matr<-(matr_Y*Conj(matr_Y))
SPE.1<-c()
for (i in 1:N)
{
SPE.1[i]<-( sum(t(rezul_matr[i,])) ) #neists reizinaajums, bet ta vajag
}
SPE.1<-abs(SPE.1)
#plot(x,f_1_2[1:(N/2)])
#lines(x,f_1_2[1:(N/2)])

#####
#####otram tas pats - iegūst otru spektrālo matricu:

##DFT no x[,j] j=1,2,...,8 apzimee ar y2

y2<-matrix(ex1[,1:8],nrow = 1024, ncol=8, byrow=FALSE)
for (i in 1:8)
{
y2[,i]<-fft(ex1[,i])/(sqrt(N))
}

#### vidējo vērtību vektoru DFT Y_i.
aa<-c((1:8)/(1:8))
sum(aa)
x.matr2<-matrix(ex1[,1:8],nrow = 1024, ncol=8, byrow=FALSE)
vid<-(aa%*%t(x.matr2))/8
Y_2.<-fft(vid)/sqrt(N)

##  SPE## DFT matrica
## DFT matrica i=1
matr_1<-(y2)
matr_2<-cbind( c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),c(Y_2.),
c(Y_2.),c(Y_2.) )

matr_Y<-matr_1-matr_2
##spektr. matr. novērtējums: f_i ar jumtiņu
#frekvences:
rezul_matr<-(matr_Y*Conj(matr_Y))
SPE.2<-c()
for (i in 1:N)
{
SPE.2[i]<-( sum(t(rezul_matr[i,])) ) #neists reizinaajums, bet ta vajag
}
SPE.2<-abs(SPE.2)
#plot(x,f_1_2[1:(N/2)])
#lines(x,f_1_2[1:(N/2)])

SPE<-SPE.1+SPE.2

#####
SPE

#####
tests:

T_2<- ((Ni*Ni)/(Ni+Ni))*( (SPE/(Ni+Ni-2))^-1 ) * abs(Y_1.-Y_2.)^2
T_2<-abs(T_2)
p<-1
F.statistika<-( ((Ni+Ni-2)*p) / (Ni+Ni-p-1) )*(T_2)
##uzzīmē statistiku:
plot(xx,F.statistika[1:(N/2)], main="Spektru salīdzināšana ", type="l")

alpha<-0.01

```

```

F_stat_kr_value<-qf(1-alpha/2,2*p,2*(Ni+Ni-p-1))
F_stat_kr_value;
abline(F_stat_kr_value,0,col="green")

##vajadzētu būt, ka ir ok. bet es nēmu kā vienvarianta (vienvērtīgas laikrindas)

#####
######
#spektrālās metricas

f_1<-SPE.1/(Ni-1)
f_2<-SPE.2/(Ni-1)

# aprekinatas grupesanas ln varbutibu vertibas eq datiem pēc (7.122)
Rezultati_1<-matrix(seq(0,0,length=8),nrow = 8, ncol=2, byrow=FALSE)
for (i in 1:8){
ln.p_1<- -log(abs(f_1)) - ( (abs(fft(x1[,i])/sqrt(1024)))^2 )/ f_1 )
Ln.p_1<-sum(ln.p_1)
Ln.p_1;
ln.p_2<- -log(abs(f_2)) - ( (abs(fft(x1[,i])/sqrt(1024)))^2 )/ f_2 )
Ln.p_2<-sum(ln.p_2)
Ln.p_2;
Rezultati_1[i,]<-rbind(Ln.p_1,Ln.p_2)
}

Rezultati_1;

Grupesana_eq<-matrix(seq(0,0,length=8),nrow = 8, ncol=2, byrow=FALSE)
Grupesana_eq<-Rezultati_1[,1]>Rezultati_1[,2]
Grupesana_eq;

# aprekinatas grupesanas ln varbutibu vertibas exp datiem pēc (7.122)
Rezultati_2<-matrix(seq(0,0,length=8),nrow = 8, ncol=2, byrow=FALSE)
for (i in 1:8){
ln.p_1<- -log(abs(f_1)) - ( (abs(fft(ex1[,i])/sqrt(1024)))^2 )/ f_1 )
Ln.p_1<-sum(ln.p_1)
Ln.p_1;
ln.p_2<- -log(abs(f_2)) - ( (abs(fft(ex1[,i])/sqrt(1024)))^2 )/ f_2 )
Ln.p_2<-sum(ln.p_2)
Ln.p_2;
Rezultati_2[i,]<-rbind(Ln.p_1,Ln.p_2)
}

Rezultati_2;

Grupesana_exp<-matrix(seq(0,0,length=8),nrow = 8, ncol=2, byrow=FALSE)
Grupesana_exp<-Rezultati_2[,1]<Rezultati_2[,2]
Grupesana_exp;

# aprekinatas grupesanas I varbutibu vertibas exp datiem pēc (7.126)
I_eq<-(Rezultati_1[,1]-Rezultati_1[,2])/1024
I_eq
I_exp<-(Rezultati_2[,1]-Rezultati_2[,2])/1024
I_exp

## tapēc, ka nēmu atsevišķi p komponenti.

```