

Implementarea relațiilor (4.4) și (4.5) în forma originală conduce de regulă la erori importante, matricea \mathbf{R}_{N_y} fiind dezechilibrată numeric. Datorită inegalității (4.1), se poate scrie că suma generică a matricii verifică următoarea proprietate numerică:

$$\sum_{n=1}^{N_y} t_n^{i+j} \geq \sum_{n=1}^{N_y} n^{i+j} T_s^{i+j} \sim N_y^{i+j+1} T_s^{i+j+1}, \quad (4.6)$$

ceea ce implică faptul că elementele de pe diagonala matricii au ordine de mărime extrem de diferite:

$$N_y T_s, N_y^3 T_s^3, N_y^5 T_s^5, \dots, N_y^{2p+1} T_s^{2p+1} \quad (4.7)$$

și inversarea conduce la valori numerice extrem de dezechilibrate. Pentru a corecta acest fenomen, matricii \mathbf{R}_{N_y} i se aplică un operator diagonal de *balansare* (adică de echilibrare numerică) de forma:

$$\mathbf{B}_M \stackrel{\text{def}}{=} \text{diag} \left[\frac{1}{\sqrt{MT_s}}, \frac{1}{MT_s \sqrt{MT_s}}, \dots, \frac{1}{M^p T_s^p \sqrt{MT_s}} \right], \quad (4.8)$$

unde $M \stackrel{\text{def}}{=} \lfloor T_{\max} / T_s \rfloor$. Evident, în cazul eșantionării uniforme, $M = N_y$.

Cu definiția (4.8), ecuația (4.4) se poate exprima echivalent astfel:

$$\hat{\boldsymbol{\theta}}_{N_y} = \mathbf{B}_M \left(\mathbf{B}_M \mathbf{R}_{N_y} \mathbf{B}_M \right)^{-1} \mathbf{B}_M \mathbf{r}_{N_y}. \quad (4.9)$$

În (4.9), inversarea matricii dintre paranteze se poate efectua acum cu precizie. Se observă de asemenea că matricea de balansare nu se inversează explicit niciodată. În pofida echilibrării numerice, erorile de calcul continuă să existe, dar ele devin inacceptabile pentru grade ale polinomului tendință care depășesc valoarea 11, după cum au demonstrat simulările. O altă posibilitate de a evita efectele numerice neplăcute constă în utilizarea unei implementări speciale (dar mai complicate) a MCMMP, bazată pe descompunerea Q-R [SCS05]. Aceasta permite operarea cu grade superioare lui 11, fără erori numerice importante. Însă, în cazul predicției, gradul tendinței nu trebuie să fie exagerat de mare, deoarece polinomul tinde să înglobeze în modelul său și o parte din zgomotul care corupe seria de timp. Acest fenomen conduce la performanțe foarte scăzute de predicție. În consecință, gradul tendinței va fi variat doar în gama $\overline{0, P_{\max}}$, cu $P_{\max} = 10$.

O altă proprietate interesantă utilă implementării metodei de estimare parametrică este recurența verificată de matricile \mathbf{R}_{N_y} și vectorii \mathbf{r}_{N_y} pentru diferite grade ale polinomului tendință. Astfel, dacă \mathbf{R}_{N_y} este renotată cu $\mathbf{R}_{N_y, p}$, iar \mathbf{r}_{N_y} cu $\mathbf{r}_{N_y, p}$ (pentru a pune în evidență gradul polinomului), atunci se constată cu ușurință că:

$$\mathbf{R}_{N_y, p} = \begin{bmatrix} & & & \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^p \\ & & & \vdots \\ & & \mathbf{R}_{N_y, p-1} & \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^{2p-1} \\ \hline \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^p & \dots & \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^{2p-1} & \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^{2p} \end{bmatrix} \quad \text{și} \quad \mathbf{r}_{N_y, p} = \begin{bmatrix} \mathbf{r}_{N_y, p-1} \\ \hline \frac{1}{N_y} \sum_{n=1}^{N_y} t_n^p y(t_n) \end{bmatrix}, \quad (4.10)$$

matricile $\mathbf{R}_{N_y, p}$ fiind, în plus, simetrice. Recurențele (4.10) arată de fapt că efortul de calcul depus pentru a evalua matricea inversabilă și vectorul liber pentru un anumit grad p poate fi conservat la evaluarea acestora pentru gradul următor, $p+1$.

Alegerea gradului polinomului tendință se poate efectua apelînd la anumite criterii structurale, bazate pe eroarea pătratică (4.3). Acestea sunt descrise pe larg în [SMS04] și [SCS05]. Din cauza caracterului lor cvasi-empiric, pentru predicție, însă, aceste criterii nu sunt recomandate. Mult mai util în alegerea gradului tendinței este criteriul introdus de noi, numit *calitatea predicției* și definit în finalul secțiunii.

De subliniat că nu se pot trasa linii de demarcație nete între cele 3 componente ale unei serii de timp. Acest lucru este ilustrat din plin de polinomul tendinței. Dacă gradul său este prea mare, componenta sezonieră tinde să fie parțial înglobată în modelul tendinței. Dacă gradul acesteia crește și mai mult, atunci și zgomotele care afectează datele tind să fie parțial modelate prin tendință. Un grad prea mic conduce la o modelare grosieră a tendinței, o parte din informația ei fiind preluată de celelalte 2 componente. Este deci preferabil ca modelul să fie parsimonios, adică tendința să aibă un grad mic, dar suficient, pentru a discrimina cu o bună acuratețe tendința și celelalte 2 componente.

După construcția modelului tendinței, valorile simulate ale acestuia se scad din valorile seriei de timp inițiale, rezultatul fiind o serie de timp *staționarizată*:

$$y_{\text{sta}}(t_n) \stackrel{\text{def}}{=} y(t_n) - y_T(t_n) = y(t_n) - a_0 - a_1 t_n - \dots - a_p t_n^p, \quad \forall n \in \overline{1, N}. \quad (4.11)$$

4.3. Estimarea componentei sezoniere

Seria de timp staționarizată constituie punctul de plecare pentru determinarea următoarei componente, cea sezonieră. Aceasta exprimă fenomenul de repetabilitate din evoluția procesului care a furnizat datele măsurate. Ea este modelată cu ajutorul a P valori succesive numite *coeficienți sezonieri*, care sunt replicați prin periodicitate pe durata orizontului de măsură. Noul semnal discret obținut, y_s , este periodic, de perioadă PT_s , unde $P \in \mathbb{N}^*$, iar T_s este perioada de eșantionare stabilită conform convențiilor de la primul paragraf al secțiunii.

Prelungirea prin periodicitate a coeficienților sezonieri se efectuează simplu în cazul eșantionării uniforme. În cazul eșantionării neuniforme, după determinarea coeficienților sezonieri, este necesară o interpolare înaintea prelungerii prin periodicitate. Interpolarea poate fi polinomială (de exemplu, liniară sau cu polinoame Lagrange) sau cu funcții spline cubice (de preferat).

În mod convențional, coeficienții sezonieri (parametrii necunoscuți ai modelului) sunt notați prin $y_{s,1}, \dots, y_{s,P}$, în timp ce indicele structural este numărul P (perioada în timp normalizat) – de asemenea necunoscut.

Determinarea componentei sezoniere se bazează pe două abordări (în care intervine MCMMP): una temporală (*Metoda Wittacker-Robinson*) și alta frecvențială (*Metoda periodogramei Schuster*). Le vom descrie pe scurt pe ambele, în cele ce urmează.

4.3.1. Metoda Wittacher-Robinson (în timp)

Coeficienții sezonieri se pot obține folosind media temporală a unor submulțimi de date consecutive staționarizate.

Dacă seria de timp este eșantionată neuniform, atunci seria staționarizată poate fi interpolată și apoi re-eșantionată uniform, la momente de timp de tipul mT_s , unde $m \in \overline{1, M_y}$, cu $M_y > N_y$. Perioada de eșantionare T_s , dacă nu este precizată, va fi aleasă

egală cu durata minimă dintre momentele de eșantionare adiacente. Pentru ușurința exprimării, se poate considera interpolarea liniară. Astfel, $y_{sta,1}$ – versiunea interpolată liniar a seriei staționarizate – are următoarea exprimare:

$$y_{sta,1}(t) \stackrel{def}{=} y_{sta}(t_{n-1}) + \frac{t - t_{n-1}}{t_n - t_{n-1}} [y_{sta}(t_n) - y_{sta}(t_{n-1})], \quad \forall t \in [t_{n-1}, t_n], \quad \forall n \in \overline{2, N_y}. \quad (4.12)$$

Prin re-eșantionarea semnalului (4.12), se obține setul de date $\{y_{sta,1}(mT_s)\}_{m \in \overline{1, M_y}}$, unde $M_y = \lfloor T_{max} / T_s + 0.5 \rfloor$ (rotunjire la cel mai apropiat întreg).

În cazul eșantionării uniforme, $M_y = N_y$ și $y_{sta,1}(mT_s) = y_{sta}[m]$, $\forall m \in \overline{1, M_y}$.

Pentru fiecare $P \in \overline{2, \lfloor M_y / 2 \rfloor}$, setul de date $\{y_{sta,1}(mT_s)\}_{m \in \overline{1, M_y}}$ este segmentat într-un număr de $Q = \lfloor M_y / P \rfloor$ seturi de date consecutive (numite și *cadre* sau *segmente de semnal*), aranjate într-o matrice, pe linii:

$$\mathbf{Y}_{sta} \stackrel{def}{=} \begin{bmatrix} y_{sta,1}(T_s) & y_{sta,1}(2T_s) & \cdots & y_{sta,1}(PT_s) \\ y_{sta,1}((P+1)T_s) & y_{sta,1}((P+2)T_s) & \cdots & y_{sta,1}(2PT_s) \\ \vdots & \vdots & \ddots & \vdots \\ y_{sta,1}(((Q-1)P+1)T_s) & y_{sta,1}(((Q-1)P+2)T_s) & \cdots & y_{sta,1}(PQT_s) \end{bmatrix}. \quad (4.13)$$

Ultimele date (care nu pot constitui un segment complet) se pierd.

Calculînd media fiecărei coloane a matricii \mathbf{Y}_{sta} , se obțin coeficienții sezonieri:

$$y_{S,p} \stackrel{def}{=} \frac{1}{Q} \sum_{k=0}^{Q-1} y_{sta,1}((kP+p)T_s), \quad \forall p \in \overline{1, P}. \quad (4.14)$$

Urmează interpolarea coeficienților sezonieri în cazul eșantionării neuniforme. Astfel, componenta sezonieră este exprimată pe o perioadă de un semnal continuu cum este și cel de mai jos, obținut prin interpolare liniară:

$$\tilde{y}_{S,1}(t) \stackrel{def}{=} y_{S,p-1} + \frac{t - (p-1)T_s}{T_s} (y_{S,p} - y_{S,p-1}), \quad \forall t \in [(p-1)T_s, pT_s], \quad \forall p \in \overline{1, P}, \quad (4.15)$$

unde, datorită periodicității, $y_{S,0} = y_{S,P}$. Interpolarea cu funcții spline cubice poate fi de asemenea utilizată (fiind mai precisă). Pe întregul orizont de măsură, componenta sezonieră $y_S(\bullet)$ este exprimată în timp continuu prin replicarea consecutivă a semnalului (4.15) de un număr corespunzător de ori. În timp discret, semnalul sezonier continuu trebuie re-eșantionat la aceleași momente de timp ca și seria de timp originală, obținîndu-se valorile: $y_S(t_n)$, $\forall n \in \overline{1, N_y}$.

De notat că interpolarea induce distorsiuni ale semnalului periodic rezultat. În cazul interpolării liniare, distorsiunile pot fi mai importante decît pentru interpolarea cu funcții spline cubice. Operațiile de interpolare și re-eșantionare nu mai sunt necesare în cazul în care seria de timp a fost eșantionată uniform.

Alegerea unei componente sezoniere adecvate se efectuează baleind gama de perioade posibile între 2 și $\lfloor M_y / 2 \rfloor$. Fie y_S^P componenta sezonieră discretă de perioadă

P (eventual eșantionată neuniform). Dintre toate componentele sezoniere $\{y_S^P\}_{P \in \overline{2, \lfloor M_y / 2 \rfloor}}$ se va alege aceea care conduce la o eroare pătratică minimă față de seria staționarizată. Mai precis, perioada optimă rezultă prin rezolvarea următoarei probleme de minimizare a erorii pătratice dintre model și procesul furnizor de date:

$$P_{0,t} = \arg \min_{P \in 2, \left\lfloor \frac{M_y}{2} \right\rfloor} \mathcal{V}[P], \text{ unde } \mathcal{V}[P] = \sum_{n=1}^{N_y} \left[y_{\text{sta}}(t_n) - y_S^P(t_n) \right]^2, \dots \forall P \in 2, \left\lfloor \frac{M_y}{2} \right\rfloor. \quad (4.16)$$

Deoarece criteriul \mathcal{V} este discret, minimizarea acestuia se poate efectua apelînd la o procedură de căutare exhaustivă, cu condiția ca numărul total al perioadelor posibile, $\left\lfloor \frac{M_y}{2} \right\rfloor$, să fie suficient de mic (de ordinul zecilor de mii, cel mult). Dacă acest număr este prea mare (sute de mii, milioane, etc.), determinarea minimului se poate realiza folosind algoritmi de căutare evoluționiști (algoritmul de analizare, algoritmi genetici, algoritmi de ascensiune montană, etc.) [RuNo95], [MiM95].

Atunci cînd numărul perioadelor posibile este mic, alegerea perioadei optime se poate realiza și pe cale grafică, după trasarea variației criteriului \mathcal{V} , ca în Figura 4.2.

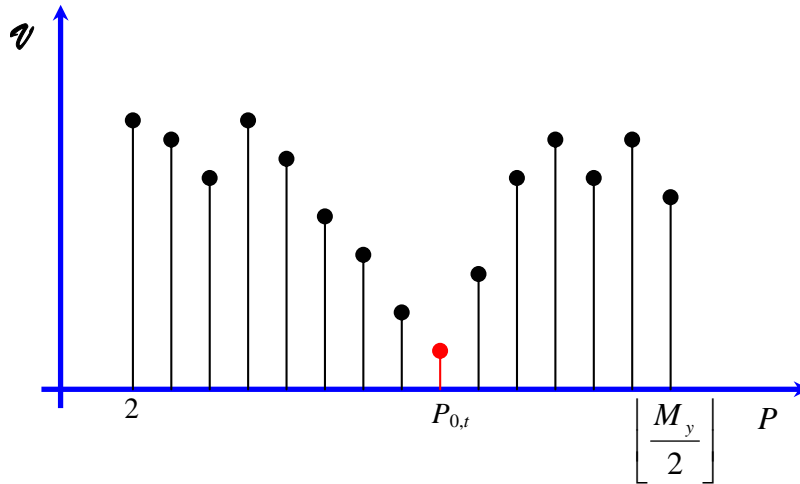


Figura 4.2. Determinarea perioadei optime cu Metoda Wittacker-Robinson.

Graficul lui \mathcal{V} poate pune în evidență mai multe minime locale situate la perioade multiple ale unei perioade date. Acestea indică de fapt doar perioada de bază, deoarece un semnal periodic de perioadă P este periodic și de perioadă nP , cu $n \geq 2$.

Dacă seria staționarizată nu posedă componentă sezonieră, graficul criteriului \mathcal{V} este fie aproape constant, fie extrem de oscilant cu numeroase minime locale situate aproape peste tot de-a lungul abscisei graficului.

4.3.2. Metoda periodogramei Schuster (în frecvență)

Potrivit Teoriei lui Joseph Fourier [OpSc85], [PrMa96], componenta sezonieră fiind un semnal periodic, poate fi aproximată punctual cu o sumă finită de armonice elementare:

$$y_S(t_n) \cong \sum_{m=1}^{M_y} \left[a_m \sin(\omega_m t_n) + b_m \cos(\omega_m t_n) \right], \quad \forall n \in \overline{1, N_y}. \quad (4.17)$$

Numărul maxim al armonicelor, M_y , poate fi ales în așa fel încît să fie inferior lui $\left\lfloor T_{\max} / (2T_s) \right\rfloor$ (adică $\left\lfloor N_y / 2 \right\rfloor$ în cazul eșantionării uniforme). Acest număr poate fi ales și mai mare, dar, dincolo de $\left\lfloor T_{\max} / (2T_s) \right\rfloor$, armonicile au puteri spectrale nesemnificative în raport cu puterile armonicelor anterioare, datorită Teoremei de eșantionare Kotelnikov-Shannon-Nyquist [StD96].

Pulsățiile armonicelor sunt alese să fie echidistante (chiar și în cazul eșantionării neuniforme). Se poate arăta că dacă momentele de eșantionare sunt multipli raționali ai perioadei de eșantionare, atunci o reprezentare Fourier corectă se obține alegînd următorul set de pulsății [StD99]:

$$\omega_m \stackrel{\text{def}}{=} \frac{mM_0\pi}{M_y T_s}, \quad \forall m \in \overline{1, M_y}, \quad (4.18)$$

unde M_0 este restul împărțirii celui mai mic multiplu comun (cmmmc) al numitorilor numerelor raționale $\{t_n/T_s\}_{n \in \overline{1, N_y}}$ la N_y . În cazul eșantionării uniforme, $M_0 = 1$. Pulsația normalizată de eșantionare $\omega_s = \pi/M_y$ controlează rezoluția în frecvență, care trebuie să fie cât mai bună, deoarece, în acest fel, precizia de determinare a perioadei optime a componentei sezoniere este mai mare. Rezoluția crește odată cu scăderea lui ω_s , adică odată cu mărirea numărului de date achiziționate, N_y . Condiția ca momentele de eșantionare să fie multipli raționali ai perioadei de eșantionare nu este restrictivă, deoarece, în practică, se operează numai cu numere raționale.

Pentru a estima parametri necunoscuți ai modelului sezonier (4.17), adică $\{a_m\}_{m \in \overline{1, M_y}}$ și $\{b_m\}_{m \in \overline{0, M_y}}$ (numiți și *coeficienți Fourier*), trebuie rezolvată o problemă de minimizare a erorii pătratice dintre model și proces:

$$\hat{\theta}_{M_y} = \arg \min_{\theta \in \mathbb{R}^{2M_y}} \mathcal{V}_{M_y}(\theta), \text{ unde: } \mathcal{V}_{M_y}(\theta) \stackrel{\text{def}}{=} \sum_{n=1}^{N_y} [y_{\text{sta}}(t_n) - y_s(t_n)]^2, \quad \forall \theta \in \mathbb{R}^{2M_y}, \quad (4.19)$$

iar θ este vectorul coeficienților Fourier. De notat că, deoarece seria de timp a fost staționarizată, coeficientul Fourier b_0 (care ar fi trebuit să apară în expresia (4.17)) are valoarea nulă, fiind proporțional cu media datelor staționarizate. Din acest motiv el lipsește din modelul Fourier (4.17). Rezolvarea problemei (4.19) se poate efectua prin metoda clasică (anularea gradientului criteriului \mathcal{V}_{M_y}). Rezultă următoarele estimări ale coeficienților Fourier:

$$\hat{\theta}_{M_y} = \mathbf{R}_{M_y}^{-1} \mathbf{r}_{M_y}, \quad \text{unde:} \quad (4.20)$$

$$\mathbf{R}_{M_y} = \begin{bmatrix} \left[\sum_{n=1}^{N_y} \sin(\omega_i t_n) \sin(\omega_j t_n) \right]_{i,j \in \overline{1, M_y}} & \left[\sum_{n=1}^{N_y} \sin(\omega_i t_n) \cos(\omega_j t_n) \right]_{i,j \in \overline{1, M_y}} \\ \left[\sum_{n=1}^{N_y} \sin(\omega_j t_n) \cos(\omega_i t_n) \right]_{i,j \in \overline{1, M_y}} & \left[\sum_{n=1}^{N_y} \cos(\omega_i t_n) \cos(\omega_j t_n) \right]_{i,j \in \overline{1, M_y}} \end{bmatrix}; \quad (4.21)$$

$$\mathbf{r}_{M_y} = \begin{bmatrix} \left[\sum_{n=1}^{N_y} y_{\text{sta}}(t_n) \sin(\omega_i t_n) \right]_{i \in \overline{1, M_y}} \\ \left[\sum_{n=1}^{N_y} y_{\text{sta}}(t_n) \cos(\omega_i t_n) \right]_{i \in \overline{1, M_y}} \end{bmatrix}. \quad (4.22)$$

Matricea simetrică \mathbf{R}_{M_y} din definiția (4.21) devine diagonală în cazul în care seria de timp este uniform eșantionată. Aceasta conduce la exprimarea coeficienților Fourier în formă completă:

$$\begin{cases} \hat{a}_m = \frac{2}{N_y} \sum_{n=1}^{N_y} y_{\text{sta}}(t_n) \sin(\omega_m t_n) \\ \hat{b}_m = \frac{2}{N_y} \sum_{n=1}^{N_y} y_{\text{sta}}(t_n) \cos(\omega_m t_n) \end{cases}, \quad \forall m \in \overline{1, M_y}. \quad (4.23)$$

După estimarea coeficienților Fourier, se poate trasa graficul *periodogramei Schuster*, care este definită prin:

$$\mathcal{P}[m] \stackrel{\text{def}}{=} \sqrt{\hat{a}_m^2 + \hat{b}_m^2}, \quad \forall m \in \overline{1, M_y}. \quad (4.24)$$

Valorile periodogramei aproximează puterile spectrale ale armonicelor din componența seriei de timp staționarizate. Armonica dominantă se poate determina prin evaluarea punctului de maxim al periodogramei, adică prin rezolvarea următoarei probleme de optimizare:

$$m_0 = \arg \max_{m \in \overline{1, M_y}} \mathcal{P}[m]. \quad (4.25)$$

Indicele pulsației optime, m_0 , conduce direct la perioada optimă a componentei sezoniere în timp continuu:

$$T_0 = \frac{2\pi}{\omega_{m_0}} = \frac{2M_y T_s}{m_0 M_0}. \quad (4.26)$$

În timp discret, perioada optimă se obține prin rotunjire la cel mai apropiat întreg:

$$P_{0,f} = \left\lfloor \frac{T_0}{T_s} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{2M_y}{m_0 M_0} + \frac{1}{2} \right\rfloor. \quad (4.27)$$

Ca și în cazul abordării anterioare (în timp), rezolvarea problemei (4.25) se poate realiza practic prin căutarea exhaustivă a maximului său direct de pe graficul periodogramei (ca în Figura 4.3). De altfel, graficul periodogramei conduce la aceleași concluzii ca și cel al erorii pătratice din cazul metodei anterioare.

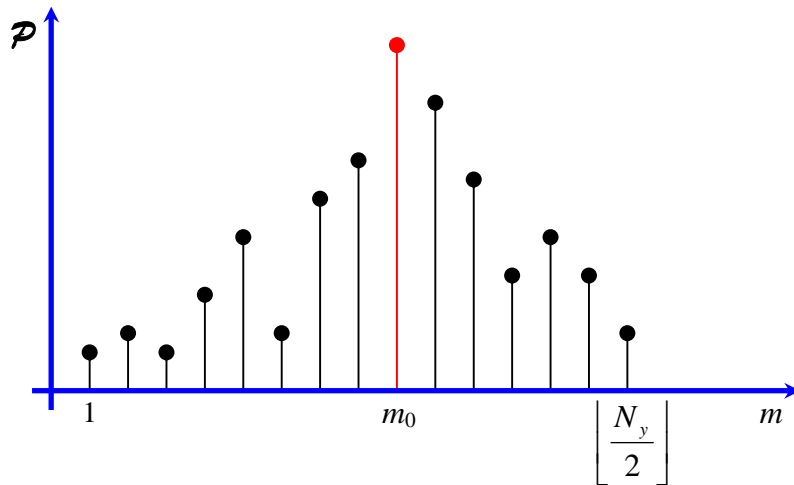


Figura 4.3. Determinarea perioadei optime cu Metoda periodogramei Schuster.

Abordarea în frecvență completează demersul anterior, bazat pe estimarea componentei sezoniere direct în domeniul timpului, în sensul că, pentru a decide perioada optimă a componentei sezoniere, trebuie comparate perioadele oferite de ambele metode.

Astfel:

- Dacă graficul criteriului \mathcal{V} și cel al periodogramei \mathcal{P} sunt grupate într-o bandă de $\pm 10\%$ în jurul mediilor lor, seria de timp nu posedă componentă sezonieră.
- Dacă $P_{0,t} = P_{0,f}$ (caz destul de rar), atunci perioada în timp discret este $P_0 = P_{0,t} = P_{0,f}$.

- Dacă $P_{0,t} \cong P_{0,f}$ (adică dacă $P_{0,t}$ și $P_{0,f}$ diferă cu o valoare mică în raport cu valorile lor, se preferă Metoda Wittacker-Robinson (deoarece introduce mai puține erori de calcul). Așadar, în acest caz: $P_0 = P_{0,t}$.
- Dacă una din cele două perioade o divide pe cealaltă, se va alege perioada optimă în mod natural: $P_0 = \min\{P_{0,t}, P_{0,f}\}$.
- Dacă perioadele $P_{0,t}$ și $P_{0,f}$ sunt total diferite, se vor lua în calcul și alte valori rezultate din extremele locale ale criteriului \mathcal{V} și periodogramei \mathcal{P} . Se va alege o pereche de valori apropiate ale perioadei (dacă este posibil) și se va selecta perioada oferită de minimul local corespunzător al criteriului \mathcal{V} . Dacă nu se poate stabili nici o corespondență între valorile posibile ale perioadelor rezultate în urma celor 2 abordări, se poate considera că seria de timp nu posedă componentă sezonieră.

Alegerea perioadei optime a componentei sezoniere constituie punctul cel mai delicat al modelării seriilor de timp. În afara celor 2 abordări, se pot utiliza în acest scop și o serie de informații apriorice legate de procesul furnizor de date, dacă sunt cunoscute. De exemplu, seria de timp a ratei șomajului înregistrat lunar într-o anumită țară va avea probabil o perioadă de 12 luni (considerând că perioada de eșantionare este de o lună).

Odată ce perioada componentei sezoniere a fost stabilită, coeficienții sezonieri se evaluează după procedeul descris în cadrul metodei Wittacker-Robinson (definițiile (4.13)-(4.14), cu P_0 în loc de P și numărul de cadre K_y evaluat corespunzător).

Dacă există, semnalul periodic asociat, y_s , se obține apoi tot în maniera descrisă în cadrul Metodei Wittacker-Robinson (prelungire prin periodicitate și, eventual, re-eșantionare). Dacă seria de date nu posedă componentă sezonieră, aceasta este asimilată cu un semnal nul. Prin scăderea componentei sezoniere din seria de date staționarizată se obține semnalul perturbator al datelor (un set de date reziduale, de fapt):

$$v(t_n) \stackrel{\text{def}}{=} y_{\text{sta}}(t_n) - y_s(t_n), \quad \forall n \in \overline{1, N}. \quad (4.28)$$

Avînd în vedere gradul de subiectivism care însoțește determinarea modelului componentei sezoniere, este necesară totuși o ultimă validare. Astfel, dacă energia semnalului perturbator (4.28) are valori semnificative în raport cu energia semnalului original centrat pe medie \mathcal{E}_y^0 , atunci componenta sezonieră nu a fost corect determinată. Energia \mathcal{E}_y^0 este, prin definiție, următoarea:

$$\mathcal{E}_y^0 \stackrel{\text{def}}{=} \sum_{n=1}^{N_y} [y(t_n) - \hat{a}_0]^2, \quad (4.29)$$

unde \hat{a}_0 este coeficientul liber al polinomului tendință (media seriei de date). Pentru mai multă claritate, se poate stabili un prag de validare $\xi \in (0, 1]$, care va activa sau inhiba componenta sezonieră. Alegerea pragului de validare nu constituie decît o manieră de cuantificare a subiectivismului în stabilirea perioadei optime. Odată precizat, pragul de validare este utilizat pentru a testa verificarea inegalității:

$$\|v\|^2 \leq \xi \mathcal{E}_y^0, \quad (4.30)$$

unde $\|v\|^2 = \mathcal{E}_v$ este energia semnalului perturbator. Dacă inegalitatea (4.30) se verifică, atunci perioada componentei sezoniere a fost corect selectată, altfel, probabil, seria de timp nu posedă componentă sezonieră. În acest caz, se va stabili direct că semnalul perturbator este chiar seria de timp staționarizată:

$$\stackrel{def}{v} \equiv y_{sta} . \quad (4.31)$$

Valorile tipice ale pragului de validare sunt cuprinse între 0.25 și 0.35. El depinde însă de raportul semnal-zgomot (SNR) al seriei de timp. Dacă acesta este inferior valorii de 6 dB (adică dacă energia zgomotului perturbator este mai mare decât un sfert din energia semnalului), atunci performanțele de predicție ale modelelor clasice devin din ce în ce mai slabe (pe măsură ce SNR scade).

4.4. Estimarea componentei aleatoare

Dacă tendința și componenta sezonieră au fost corect determinate, setul de date $\{v(t_n)\}_{n \in \overline{1, N_y}}$, obținut după extragerea lor din seria de timp, are caracteristicile unui zgomot colorat rezidual. Filtrul de zgomot poate fi de tip FIR (cu răspuns finit la impuls) sau IIR (cu răspuns infinit la impuls). În primul caz, se poate opera cu un model de identificare de tip *medie alunecătoare* (MA). În al doilea caz, este utilizat unul dintre modelele AR sau ARMA [SCS05]. Filtrele de tip FIR sunt de asemenea utilizate pentru a aproxima filtre de tip IIR, dar funcțiile pondere au, în general, o lungime ridicată. Vom adopta în continuare filtrul de tip IIR. Deoarece în abordarea care urmează maniera de eșantionare nu are importanță, valorile zgomotului sunt renotate, pentru ușurința exprimării, prin $\{v[n]\}_{n \in \overline{1, N_y}}$. Vor fi descrise separat cele două modele de filtre IIR care pot fi utilizate în vederea predicției.

4.4.1. Modelul auto-regresiv (AR)

Modelul AR este unul dintre primele modele de identificare utilizate în aplicații, în special datorită simplității și a posibilității de a estima parametrii în manieră recursivă. Ecuația modelului AR este următoarea:

$$\begin{cases} v[n] + a_1 v[n-1] + \dots + a_{na} v[n-na] = e[n] \\ E\{e[n]e[m]\} = \lambda^2 \delta_0[n-m] \end{cases}, \quad \forall n \in \overline{1, N_y}, \quad (4.32)$$

unde parametrii necunoscuți sunt coeficienții $\{a_i\}_{i \in \overline{1, na}}$ (asamblați într-un vector $\theta_{na} \in \mathbb{R}^{na}$) și dispersia zgomotului alb e este notată prin λ^2 . Prin E s-a notat operatorul de mediere statistică. Indicele structural al modelului (necunoscut și el) este $na \geq 1$. În general, indicele structural maxim nu depășește valoarea $Na = 100$ pentru majoritatea seriilor de timp. Termenul din stînga al primei ecuații din definiția (4.32) este partea *auto-regresivă* (AR) a modelului (formată din valori *regresate în timp* ale zgomotului colorat). Modelul efectiv al componentei aleatoare asociate seriei de timp se exprimă prin:

$$y_{AR}[n] = v[n] - e[n], \quad \forall n \in \overline{1, N_y} \quad (4.33)$$

(din zgomotul colorat se elimină zgomotul alb).

Definiția (4.32) poate fi exprimată în forma *de filtrare*:

$$v[n] = \frac{1}{A(q^{-1})} e[n], \quad \forall n \in \overline{1, N_y}, \quad (4.34)$$

unde q^{-1} este operatorul de întârziere cu un pas [SCS05], iar A este un polinom:

$$A(q^{-1}) \stackrel{def}{=} 1 + a_1 q^{-1} + \dots + a_{na} q^{-na}. \quad (4.35)$$

Operatorul de întârziere este un instrument extrem de utilizat pentru a exprima modele de identificare. El corespunde variabilei z^{-1} în raport cu care se exprimă Transformata Z (TZ)

a unui semnal discret. Corespondența provine dintr-o proprietate remarcabilă de TZ: *Teorema întârzierii* (TZ a unui semnal întârziat cu un pas este egală cu TZ a semnalului inițial, multiplicată de variabila z^{-1}). Astfel, TZ a setului de coeficienți $\{a_i\}_{i \in \overline{0, na}}$ (unde $a_0 = 1$) are chiar expresia polinomului (4.35), în care s-a înlocuit q^{-1} cu z^{-1} :

$$A(z^{-1}) \stackrel{\text{def}}{=} 1 + a_1 z^{-1} + \dots + a_{na} z^{-na}, \quad (4.36)$$

unde variabila complexă z aparține unei zone de convergență de tip coroană circulară centrate în origine. (Coroana este unic determinată de setul de coeficienți.) În consecință, revenind la expresia (4.33), valorile componente aleatoare y_{AR} se obțin prin filtrarea zgomotului alb e cu un sistem a cărui funcție de transfer are expresia:

$$H_{AR}(z) = \frac{\stackrel{\text{def}}{1} - A(z^{-1})}{A(z^{-1})} \stackrel{\text{def}}{=} \frac{-a_1 z^{-1} - a_2 z^{-2} - \dots - a_{na} z^{-na}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{na} z^{-na}}. \quad (4.36)$$

Acesta este un filtru de tip IIR, ai cărui poli determină în mod unic rădăcinile.

Determinarea modelului AR (4.32) se bazează pe MCMMP. Astfel, pentru fiecare index structural $na \in \overline{1, Na}$, parametri necunoscuți estimați din zgomotul colorat sunt următorii:

$$\hat{\theta}_{N_y, na} = \mathbf{R}_{N_y, na}^{-1} \mathbf{r}_{N_y, na}, \quad \hat{\lambda}_{N_y, na}^2 = \frac{1}{N_y} \sum_{n=1}^{N_y} \left(v[n] - \boldsymbol{\varphi}_{na}^T[n] \hat{\theta}_{N_y, na} \right)^2, \quad (4.37)$$

unde:

$$\boldsymbol{\varphi}_{na}^T[n] \stackrel{\text{def}}{=} [-v[n-1] \quad \dots \quad -v[n-na]]; \quad (4.38)$$

$$\mathbf{R}_{N_y, na} \stackrel{\text{def}}{=} \frac{1}{N_y} \sum_{n=1}^{N_y} \boldsymbol{\varphi}_{na}[n] \boldsymbol{\varphi}_{na}^T[n] = \left[\frac{1}{N_y} \sum_{n=1}^{N_y} v[n-i] v[n-j] \right]_{i, j \in \overline{1, na}}; \quad (4.39)$$

$$\mathbf{r}_{N_y, na} \stackrel{\text{def}}{=} \frac{1}{N_y} \sum_{n=1}^{N_y} \boldsymbol{\varphi}_{na}[n] v[n] = \left[-\frac{1}{N_y} \sum_{n=1}^{N_y} v[n] v[n-i] \right]_{i \in \overline{1, na}}. \quad (4.40)$$

Implementarea relațiilor (4.37)-(4.40) se poate realiza mai eficient dacă se observă că matricea $\mathbf{R}_{N_y, na}$ este (aproximativ) Toeplitz simetrică (adică este generată de prima linie sau coloană), iar majoritatea elementelor vectorului $\mathbf{r}_{N_y, na}$ se regăsesc printre elementele matricii $\mathbf{R}_{N_y, na}$. Pentru o precizie mai mare, diagonala principală a matricii ar trebui să fie constantă și egală cu primul element din definiția (4.39). În acest caz, matricea devine chiar de tip Toeplitz. Relații de recurență similare celor de la (4.10) pot fi de asemenea puse în evidență cu ușurință.

Inversarea matricii $\mathbf{R}_{N_y, na}$ se poate realiza cu ajutorul Algoritmului Levinson-Durbin (ALD) [SCS05], [PrMa96], pe care îl vom prezenta în continuare. (Altfel, această operație poate ridica din nou probleme numerice, ca în cazul modelului tendinței.) ALD permite estimarea recursivă a coeficienților modelului curent AR[na], în funcție de coeficienții modelului de ordin inferior AR[na-1]. Pentru a ilustra aceste relații recursive, coeficienții modelului AR[na] se notează cu $a_{na, i}$, $\forall i \in \overline{1, na}$. De asemenea, se notează cu $\hat{\lambda}_v$ funcția de auto-covarianță aproximativă estimată din datele reziduale:

$$\hat{r}_v[k] \stackrel{def}{=} \frac{1}{N_y - k} \sum_{n=k+1}^{N_y} u[n]u[n-k], \quad \forall k \in \overline{1, \lfloor N_y/4 \rfloor}. \quad (4.41)$$

Procedura Levinson-Durbin este descrisă în tabloul **Algoritmul 4.1**. (Justificarea relațiilor acestuia se găsește, de exemplu, în [SCS05]. Ele au fost deduse prin exploatarea unor proprietăți remarcabile ale matricilor Toeplitz simetrice. De altfel, folosind acest algoritm, se poate proiecta o procedură eficientă de inversare a matricilor Toeplitz simetrice.)

Algoritmul 4.1. Algoritmul Levinson-Durbin.

➤ Date de intrare: seria de date reziduale: $\{v[n]\}_{n \in \overline{1, N_y}}$.

1. Inițializare:
$$\begin{cases} \hat{a}_{1,1} = -\frac{\hat{r}_v[1]}{\hat{r}_v[0]} \\ \hat{\lambda}_1^2 = \hat{r}_v[0] + \hat{a}_{1,1}\hat{r}_v[1] = \hat{r}_v[0](1 - \hat{a}_{1,1}^2) \end{cases}.$$

2. Pentru $na \in \overline{2, Na}$:

$$\begin{cases} \hat{a}_{na,na} = -\frac{1}{\hat{\lambda}_{na-1}^2} (\hat{r}_v[na] + \hat{a}_{na-1,1}\hat{r}_v[na-1] + \dots + \hat{a}_{na-1,na-1}\hat{r}_v[1]) \\ \hat{a}_{na,i} = \hat{a}_{na-1,i} + \hat{a}_{na,na}\hat{a}_{na-1,na-i}, \quad \forall i \in \overline{1, na-1} \\ \hat{\lambda}_{na}^2 = \hat{\lambda}_{na-1}^2 (1 - \hat{a}_{na,na}^2) \end{cases}.$$

➤ Date de ieșire: coeficienții tuturor modelelor AR[na], cu $na \in \{1, 2, \dots, Na\}$.

Se poate arăta că modelul AR determinat cu ajutorul ALD este stabil (rădăcinile polinomului auto-regresiv sunt situate în discul unitar al planului complex) [SCS05], [PrMa96]. De asemenea, coeficienții $\hat{a}_{na,na}$ se mai notează prin \hat{k}_{na} și se numesc *coeficienți de reflexie*. Ei joacă un rol important în estimarea dispersiei zgomotului alb, care îndeplinesc funcția de eroare de predicție cu un pas. Se poate arăta că acești coeficienți sunt teoretic de modul subunitar ($|\hat{k}_{na}| < 1, \forall na \geq 1$), ceea ce arată că estimăția dispersiei zgomotului alb ($\hat{\lambda}_{na}^2$) scade odată cu ordinul modelului. Cu toate acestea, așa cum arată paragraful următor, precizia predicției cu mai mulți pași a modelului AR nu se îmbunătățește în mod necesar odată cu ordinul acestuia.

Alegerea modelului optimal al componentei nedeterminate se realizează fie folosind criteriile structurale amintite în paragraful 4.2. (bazate însă pe variația dispersiei estimate a zgomotului alb $\hat{\lambda}_{na}^2$), fie tot criteriul calității predicției, definit în paragraful 4.6. Prima opțiune nu se recomandă însă, deoarece criteriile structurale nu au o legătură directă cu operația de predicție. Odată ce ordinul optim (na_0) a fost selectat, modelul estimat al componentei aleatoare este dat de următoarea ecuație omogenă recursivă (cu diferențe):

$$\begin{cases} y_{AR}[n] + \hat{a}_{na_0,1}y_{AR}[n-1] + \dots + \hat{a}_{na_0,na_0}y_{AR}[n-na_0] = 0 \\ y_{AR}[1] = v[1]; y_{AR}[2] = v[2]; \dots; y_{AR}[na_0] = v[na_0] \end{cases}, \quad \forall n \in \overline{1, N_y}, \quad (4.42)$$

De notat că zgomotul alb rezidual lipsește din ecuația (4.42), de aceea această ecuație definește doar o estimăție a componentei aleatoare (chiar dacă notația ei se conservă). El va interveni (prin dispersia sa) în faza de predicție.