
# Sumar

✓ **Bibliografie**

✓ ⓿ Organizarea temelor de laborator

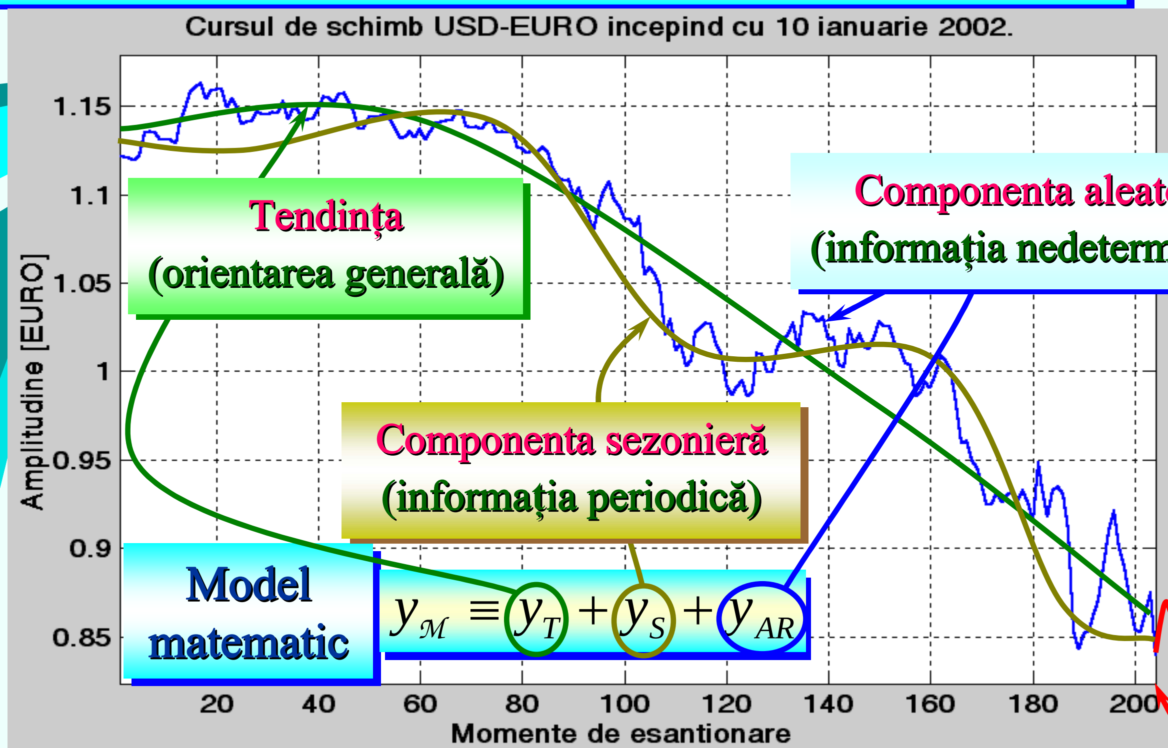
☞ ❽ Modelarea şi predicţia seriilor de timp

- ☞ ❽.❶ Estimarea modelului polinomial al tendinţei
- ❽.❷ Estimarea componentei sezoniere
- ❽.❸ Estimarea componentei nedeterministe (aleatoare)
- ❽.❹ Predicţia seriei de timp

# ❽ Modelarea şi predicţia seriilor de timp

## ☞ Notaţii şi definiţii de bază

**Serie de timp** ⇒ Şir de date înregistrat în urma evoluţiei unui proces, fără a putea cuantifica sau cunoaşte cauzele acelei evoluţii.




**Eşantionare uniformă**

$$\mathcal{D} = \{y[n] = y(nT_s)\}_{n=\overline{1,N}}$$

perioadă de eşantionare

**Eşantionare neuniformă**

$$\mathcal{D} = \{y(t_n)\}_{n=\overline{1,N}}$$

$$t_1 < t_2 < \cdots < t_n < \cdots < t_N$$

momente de eşantionare

Tendinţa (orientarea generală)

Componenta aleatoare (informaţia nedeterministă)

Componenta sezonieră (informaţia periodică)

**Model matematic**

$$y_{\mathcal{M}} \equiv y_T + y_S + y_{AR}$$

☝ Orizontul de măsură este limitat. $T_{max}$

?

✍ Convenţie: perioada de eşantionare este egală cu minimul duratei dintre oricare două momente de eşantionare succesive. ⇒

$$t_n \geq nT_s \quad \forall\, t \in \mathbb{N}^*$$

**Obiectivul modelării** ⇒ Predicţia/Prognoza seriei de timp.

# ❽ Modelarea şi predicţia seriilor de timp

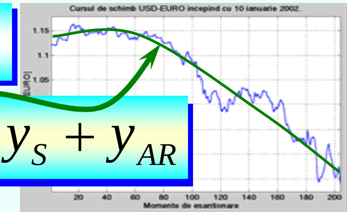

## ❽.❶ Estimarea modelului polinomial al tendinţei

**Tendinţă** ➔ Model determinist asociat orientării generale a seriei de timp.

$$y_T(t) = a_0 + a_1 t + \cdots + a_p t^{\boxed{p}} \quad \forall t \in \mathbb{R}$$

☝ De tip polinomial.

$$y_{\mathcal{M}} \equiv y_T + y_S + y_{AR}$$




parametri (necunoscuţi)

ordinul modelului (necunoscut)

**Eşantionare uniformă**

$$y_T(nT_s) = a_0 + a_1 nT_s + \cdots + a_p n^p T_s^p \quad \forall n \in \mathbb{Z}$$

← dacă se cunoaşte perioada de eşantionare

$$y_T(nT_s) = a_0 + a_1 n + \cdots + a_p n^p \quad \forall n \in \mathbb{Z}$$

← dacă nu se cunoaşte perioada de eşantionare sau este considerată unitară

$$\boldsymbol{\theta} \overset{def}{=} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} \in \mathbb{R}^{\boxed{p+1}}$$

vectorul parametrilor necunoscuţi ⤴

indicele stuctural necunoscut

**Eşantionare neuniformă**

$$y_T(nT_s) = a_0 + a_1 t_n + \cdots + a_p t_n^p \quad \forall n \in \mathbb{Z}$$

← pentru orice colecţie de momente de eşantionare

$$\cdots < t_{-n} < \cdots < t_{-1} < t_0 < t_1 < \cdots < t_n \cdots$$

☝ inclusiv uniforme

Cum poate fi determinat modelul?

Folosind Metoda Celor Mai Mici Pătrate (MCMMP).

$$\hat{\boldsymbol{\theta}}_N = \arg\min_{\boldsymbol{\theta} \in \mathcal{S}} \mathcal{V}_N(\boldsymbol{\theta})$$

domeniul de stabilitate

$$\mathcal{V}_N(\boldsymbol{\theta}) \overset{def}{=} \sum_{n=1}^{N} \left( y(t_n) - y_T(t_n) \right)^2 \quad \forall \boldsymbol{\theta} \in \mathcal{S}$$

Curs IS ➔


# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei

**MCMMP**

Estimaţie consistentă

$$\hat{\boldsymbol{\theta}}_N = \mathbf{R}_N^{-1}\,\mathbf{r}_N$$

$$\mathbf{r}_N \stackrel{def}{=} \left[\frac{1}{N}\sum_{n=1}^{N} t_n^i\, y(t_n)\right]_{i\in\overline{0,p}}$$

vector de covarianţe

$$\mathbf{R}_N \stackrel{def}{=} \left[\frac{1}{N}\sum_{n=1}^{N} t_n^{i+j}\right]_{i,j\in\overline{0,p}}$$

matrice simetrică

☝ poate fi construită recursiv

$$\mathbf{R}_{N,p} = \left[\begin{array}{c|c} \mathbf{R}_{N,p-1} & \begin{array}{c} \frac{1}{N}\sum_{n=1}^{N} t_n^{p} \\ \vdots \\ \frac{1}{N}\sum_{n=1}^{N} t_n^{2p-1} \end{array} \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n^{p} \quad \cdots \quad \frac{1}{N}\sum_{n=1}^{N} t_n^{2p-1} & \frac{1}{N}\sum_{n=1}^{N} t_n^{2p} \end{array}\right]$$

$$\mathbf{R}_{N,p-1} = \left[\begin{array}{c|c} \mathbf{R}_{N,1} & \begin{array}{c} \frac{1}{N}\sum_{n=1}^{N} t_n^{p-1} \\ \vdots \\ \frac{1}{N}\sum_{n=1}^{N} t_n^{2p-3} \end{array} \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n^{p-1} \quad \cdots \quad \frac{1}{N}\sum_{n=1}^{N} t_n^{2p-3} & \frac{1}{N}\sum_{n=1}^{N} t_n^{2p-2} \end{array}\right]$$

$$\mathbf{R}_{N,1} = \left[\begin{array}{c|c} \mathbf{R}_{N,0} = 1 & \frac{1}{N}\sum_{n=1}^{N} t_n \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n & \frac{1}{N}\sum_{n=1}^{N} t_n^{2} \end{array}\right]$$

☝ poate fi construit recursiv

$$\mathbf{r}_{N,p} = \left[\begin{array}{c} \mathbf{r}_{N,p-1} \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n^{p}\, y(t_n) \end{array}\right]$$

$$\mathbf{r}_{N,p-1} = \left[\begin{array}{c} \mathbf{r}_{N,0} \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n\, y(t_n) \\ \vdots \\ \hline \frac{1}{N}\sum_{n=1}^{N} t_n^{p-1}\, y(t_n) \end{array}\right]$$

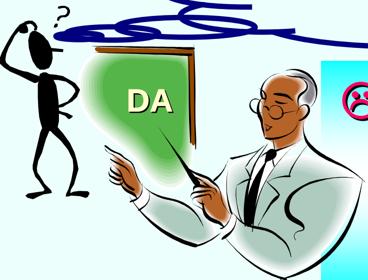
$$\mathbf{r}_{N,0} = \frac{1}{N}\sum_{n=1}^{N} y(t_n)$$

**1/Precizie**

$$\hat{\lambda}^2_{N,p} = \frac{1}{N-p}\,\mathcal{V}_N\!\left(\hat{\boldsymbol{\theta}}_{N,p}\right)$$


# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei




☹ Cu toate acestea, matricea este **puternic dezechilibrată numeric** (avînd un **număr de condiţionare numerică extrem de mare**).

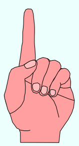
**raportul dintre valoarea proprie de modul maxim şi cea de modul minim** $\nu\left(\mathbf{R}_{N,p}\right)$

**Exemplu**

Momente de eşantionare uniforme, de la 1 la 100 şi polinom de gradul 2.

$$\mathbf{R}_{N,2} = \begin{bmatrix} 1 & \frac{N+1}{2} & \frac{1}{N}\sum_{n=1}^{N} n^2 \\ \frac{N+1}{2} & \frac{1}{N}\sum_{n=1}^{N} n^2 & \frac{1}{N}\sum_{n=1}^{N} n^3 \\ \frac{1}{N}\sum_{n=1}^{N} n^2 & \frac{1}{N}\sum_{n=1}^{N} n^3 & \frac{1}{N}\sum_{n=1}^{N} n^4 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 50.5 & 3383.5 \\ 50.5 & 3383.5 & 255025 \\ 3383.5 & 255025 & 20503333.3 \end{bmatrix}$$



Se observă **diferenţa sensibilă** de ordin de mărime dintre elementele matricii.

☝ **Inversarea numerică poate introduce erori semnificative!**

$$\nu\left(\mathbf{R}_{100,2}\right) = 1.9245 \cdot 10^8 \;!$$


Se poate aplica o tehnică de **echilibrare numerică** (a se vedea cursul de *Metode Numerice*).

$$M \overset{def}{=} \left\lceil \frac{T_{\max}}{T_s} \right\rceil$$

**Formula de implementare**

$$\hat{\theta}_N = \mathbf{B}_M \boxed{\left(\mathbf{B}_M \mathbf{R}_N \mathbf{B}_M\right)}^{-1} \mathbf{B}_M \mathbf{r}_N$$

☝ **Echilibrată numeric.**

Se construieşte o **matrice de balansare.**

$$\mathbf{B}_M \overset{def}{=} \operatorname{diag}\left[\frac{1}{\sqrt{MT_s}} \quad \frac{1}{MT_s\sqrt{MT_s}} \quad \cdots \quad \frac{1}{M^pT_s^p\sqrt{MT_s}}\right]$$

# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei

**Exemplu** Momente de eşantionare uniforme, de la 1 la 100 şi polinom de gradul 2.

$$\mathbf{R}_{100,2}=\begin{bmatrix}1 & 50.5 & 3383.5\\ 50.5 & 3383.5 & 255025\\ 3383.5 & 255025 & 20503333.3\end{bmatrix}$$

$$\nu\left(\mathbf{R}_{100,2}\right)=1.9245\cdot10^{8}$$

$$\mathbf{B}_{100}\overset{def}{=}\operatorname{diag}\left[0.1\quad 0.001\quad 0.00001\right]$$

**Balansare**

$$\mathbf{B}_{100}\mathbf{R}_{100,2}\mathbf{B}_{100}=\begin{bmatrix}0.1 & 0.00505 & 0.0033835\\ 0.00505 & 0.0033835 & 0.00255025\\ 0.0033835 & 0.00255025 & 0.00205033333\end{bmatrix}$$

$$\nu\left(\mathbf{B}_{100}\mathbf{R}_{100,2}\mathbf{B}_{100}\right)=5.336\cdot10^{2}$$

Condiționarea numerică poate fi îmbunătățită cu ajutorul unor tehnici de balansare mai sofisticate.

Dar indicele structural? Cum se poate determina?

Se apelează la strategia generală prezentată în cursul de IS.

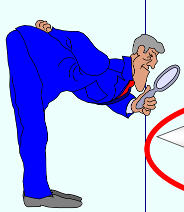
Pentru fiecare structură de model din ce în ce mai bogată ($m\in\{1,2,\ldots,M\}$):
- se determină parametrii modelului ales, $\theta_m$;
- se evaluează precizia modelului ($\mathcal{V}(\theta_m)$).

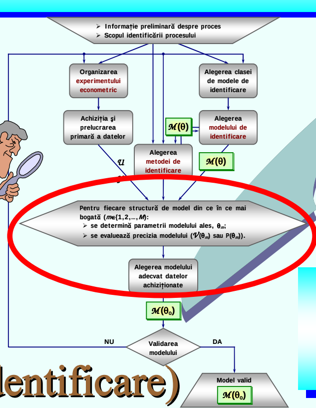
Alegerea modelului adecvat datelor achiziționate

Gradul maxim al modelului.

$$P_{\max}=10$$

Teste (criterii) de adecvanță






Curs IS ➔ (experiment de identificare)

Modele de identificare de acelaşi tip sunt comparate între ele în vederea alegerii celui adecvat.


# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei

Criteriul utilizat în alegerea structurii modelului tendinţei → **Calitatea predicţiei** $\mathrm{PQ}_K[p, P, na]$

Definit în final, pentru toate cele 3 modele componente.


Implementarea algoritmului de estimare a coeficienţilor modelului polinomial al tendinţei, bazat pe MCMMP.

### ☞ Paşii principali ai algoritmului de modelare

**Date de intrare**
- $\mathcal{D}$ (seria de timp)
- $p$ (gradul polinomului; implicit: $p = 0$)

**Iniţializare**
- Determinarea lungimii seriei de timp. → $N$
- Evaluarea perioadei de eşantionare. → $T_s$ ☜ Unitară, pentru seriile eşantionate uniform.
- Evaluarea numărului de balansare. → $M$ ☜ Egal cu lungimea seriei de timp, pentru seriile eşantionate uniform.

# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei

### ☞ Paşii principali ai algoritmului de modelare (continuare)

Se estimează vectorul parametrilor necunoscuţi prin implementarea MCMMP cu balansare. $\hat{\theta}_{N,p}$

☝ Recursiv!

Se evaluează modelul efectiv al tendinţei pe orizontul de măsură.

$$y_T(t_n) = \hat{a}_0 + \hat{a}_1 t_n + \cdots + \hat{a}_p t_n^p \quad \forall n \in \overline{1,N}$$

Se evaluează seria de timp staţionarizată.

$$y_{sta}(t_n) = y(t_n) - y_T(t_n) \quad \forall n \in \overline{1,N}$$

$$\langle y_{sta} \rangle \overset{def}{=} \frac{1}{N} \sum_{n=1}^{N} y_{sta}(t_n)$$

media datelor staţionarizate

Se corectează modelului tendinţei şi al seriei de timp staţionarizate cu media.

$$y_{sta} \leftarrow y_{sta} - \langle y_{sta} \rangle \qquad \hat{a}_0 \leftarrow \hat{a}_0 + \langle y_{sta} \rangle$$

**Date de ieşire** $y_{sta}$ $y_T$ $\hat{\theta}_{N,p}$

**Rutină ce trebuie proiectată**

```
>> [ysta,yT,theta] = trend(y,p);
```

- ysta: vectorul datelor staţionarizate
- yT: vectorul tendinţei pe orizontul de măsură
- theta: vectorul parametrilor modelului
- y: vectorul de date care conţine seria de timp
- p: gradul polinomului tendinţă

# ❽ Modelarea şi predicţia seriilor de timp

## ❽.❶ Estimarea modelului polinomial al tendinţei

**Programul de test al rutinei `trend`** **`ISLAB_8A`**

Punctaj?

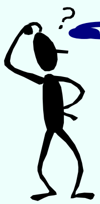
**10 p**

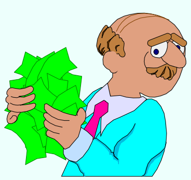
### ☞ Paşi principali

- Se introduce numărul fişierului care conţine seria de timp. `nts`
- Se introduce gradul polinomului tendinţă. `p`
- Se încarcă seria de timp în spaţiul de lucru. • cu ajutorul funcţei: `eval`
- Se apelează rutina `trend`.
- Se trasează graficele seriei de timp şi al tendinţei (în aceeaşi fereastră).
- Se trasează graficul seriei de timp staţionarizate (cu precizarea mediei pe grafic).

☝ Graficele trebuie trasate cu axele corect scalate şi marcate, cu titluri sugestive şi legendă de discriminare (dacă este cazul).

Fişiere de date?





Mini-programe MATLAB.

**`STnts.M`**

**`nts`∈`{1,2,3,...,15}`**

➢ Parametrii încărcaţi în spaţiul de lucru:

| | |
|---|---|
| `y` | 🡐 seria de timp |
| `ntime` | 🡐 suportul seriei de timp (momentele de timp normalizate) |
| `Ts` | 🡐 perioada de eşantionare |
| `unit` | 🡐 unitatea de măsură a perioadei de eşantionare |
| `yunit` | 🡐 unitatea de măsură a valorilor seriei de timp |
| `label` | 🡐 text care arată ce reprezintă seria de timp |