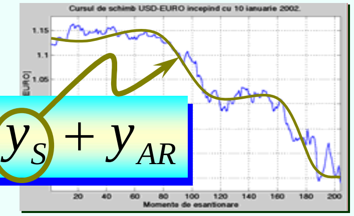
Sumar

- ✓ **Bibliografie**
- ✓ ⓿ Organizarea temelor de laborator
- ✓ ⑧ Modelarea şi predicţia seriilor de timp
  - ✓ ⑧.① 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

❽.❷ Estimarea componentei sezoniere

Componentă sezonieră ⇒ Model determinist asociat fenomenului de repetabilitate relevat de seria de timp.



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

$$y_S = \cdots \cup \left\{ y_{S,1}^{k-1}, y_{S,2}^{k-1}, \cdots, y_{S,P_{k-1}}^{k-1} \right\} \cup \left\{ y_{S,1}^{k}, y_{S,2}^{k}, \cdots, y_{S,P_k}^{\boxed{k}} \right\} \cup \cdots$$

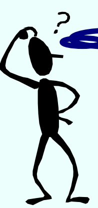
coeficienţi sezonieri (necunoscuţi)

indice de perioadă $k \in \mathbb{Z}$

prelungire prin periodicitate

perioada modelului (necunoscută) $P \in \mathbb{N}^*$

☝ Nu neapărat o armonică elementară.



Cum se poate realiza prelungirea prin periodicitate?

Eşantionare neuniformă

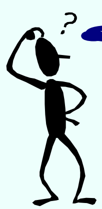
Sunt necesare interpolarea şi re-eşantionarea la momentele indicate de fiecare perioadă.

Eşantionare uniformă

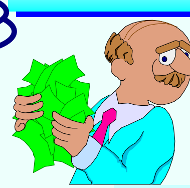
Natural, prin repetarea perioadei principale.

$P = P_k$ $\quad \forall\, k \in \mathbb{Z}$

$y_{S,p} = y_{S,p}^{k-1} = y_{S,p}^{k}$ $\quad \forall\, p \in \overline{1,P}$



Interpolare?



Există numeroase tehnici, dintre care două sunt preferate în acest context.

Tabelă de interpolare

nod de interpolare

momente de eşantionare (nu neapărat uniforme)

eşantioane

t_1	t_2	…	t_n	…	t_N
y_1	y_2	…	y_n	…	y_N

Interpolare liniară

$$y(t) = y_{n-1} + \frac{t - t_{n-1}}{t_n - t_{n-1}} (y_n - y_{n-1}) \quad \forall\, t \in \mathbb{R}$$

Interpolare cu funcţii spline cubice

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Funcţii spline?

Funcţii polinomiale pe porţiuni. Polinoamele au grade egale, dar pot avea coeficienţi diferiţi între perechi diferite de momente de eşantionare.

$$\mathcal{S} \stackrel{def}{\equiv} \sum_{n=1}^{N-1} \mathcal{S}_n$$

Interpolatorul spline cubic → **Sumă de *N*-1 polinoame de grad 3, determinate de tabela de interpolare.**

t_1	t_2	…	t_n	…	t_N
y_1	y_2	…	y_n	…	y_N

$$\mathcal{S}_n(t) \stackrel{def}{=} a_{n,3}t^3 + a_{n,2}t^2 + a_{n,1}t + a_{n,0} \quad \forall\, t \in [t_n, t_{n+1}] \quad \forall\, n \in \overline{1, N-1}$$

coeficienţi necunoscuţi (în număr de 4)

Condiţii generale care pot determina interpolatorul spline cubic

① Suportul fiecărui polinom din sumă este inclus în intervalul compact dintre două momente succesive de eşantionare.

② Interpolatorul trece prin nodurile de interpolare.

③ Derivabilitatea interpolatorului este cel puţin 1 şi cel mult egală cu 3.

① $Supp(\mathcal{S}_n) \subseteq [t_n, t_{n+1}] \quad \forall\, n \in \overline{1, N-1}$

② $\mathcal{S}(t_k) = y_k \quad \forall\, k \in \overline{1, N}$ ⟺ $\begin{cases} \mathcal{S}_k(t_k) = y_k \\ \mathcal{S}_k(t_{k+1}) = y_{k+1} \end{cases} \quad \forall\, k \in \overline{1, N-1}$

③ $\mathcal{S}_k^{(i)}(t_{k+1}) = \mathcal{S}_{k+1}^{(i)}(t_{k+1}) \quad \forall\, i \in \overline{1,3} \quad \forall\, k \in \overline{2, N-1}$

Polinoamele pot să coincidă pînă la derivata a 3-a (inclusiv) în nodurile interne.

- $N = 2$ → 2 ecuaţii, 4 necunoscute
- $N = 3$ → 7 ecuaţii, 8 necunoscute
- $N = 4$ → 12 ecuaţii, 12 necunoscute
- $N > 4$ → mai multe ecuaţii decît necunoscute

☝ Ar trebui estimate derivatele de ordin 1.

☝ Sistem liniar cu *5N-8* ecuaţii şi *4N-4* necunoscute.

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

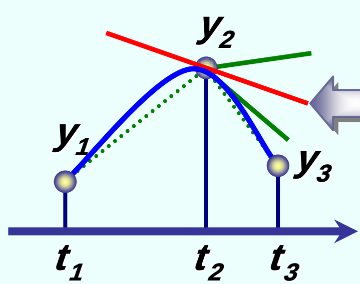
Cum se determină practic interpolatorul spline cubic?

Numărul de ecuaţii trebuie să fie egal cu numărul de necunoscute, iar derivatele de ordinul întîi trebuie estimate. Se renunţă la egalităţile derivatelor de ordin superior.

$N = 2$ ⇒ 2 ecuaţii, 4 necunoscute ⇒ Pentru o tabelă cu două noduri este de preferat o interpolare liniară.

$N = 3$ ⇒ 8 ecuaţii, 8 necunoscute ⇒ Trebuie estimate derivatele de ordinul întîi în toate nodurile de interpolare.

Principiul general



(pentru nodurile interne)

$$y_{2,1} \overset{def}{=} \operatorname{tg}\left[w_1 \operatorname{arctg}\left(\frac{y_2 - y_1}{t_2 - t_1}\right) + w_2 \operatorname{arctg}\left(\frac{y_3 - y_2}{t_3 - t_2}\right)\right]$$

unghiuri obţinute prin interpolare liniară

ponderi determinate de duratele de eşantionare

$$w_1 = \frac{t_3 - t_2}{t_3 - t_1} \qquad w_2 = \frac{t_2 - t_1}{t_3 - t_1}$$

☝ Ponderile trebuie să fie invers proporţionale cu duratele de eşantionare.

- Pentru nodurile marginale, se pot considera derivatele laterale.

$$y_{1,1} \overset{def}{=} \frac{y_2 - y_1}{t_2 - t_1} \qquad y_{3,1} \overset{def}{=} \frac{y_3 - y_2}{t_3 - t_2}$$

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Aşadar $N=3$ ⇨ 8 ecuaţii, 8 necunoscute

$$\begin{cases}
a_{1,3}t_1^3+a_{1,2}t_1^2+a_{1,1}t_1+a_{1,0}=y_1 & \text{1 stînga}\\
a_{1,3}t_2^3+a_{1,2}t_2^2+a_{1,1}t_2+a_{1,0}=y_2 & \text{1 dreapta}\\
a_{2,3}t_2^3+a_{2,2}t_2^2+a_{2,1}t_2+a_{2,0}=y_2 & \text{2 stînga}\\
a_{2,3}t_3^3+a_{2,2}t_3^2+a_{2,1}t_3+a_{2,0}=y_3 & \text{2 dreapta}\\
3a_{1,3}t_1^2+2a_{1,2}t_1+a_{1,1}=y_{1,1} & \text{1 derivată stînga}\\
3a_{1,3}t_2^2+2a_{1,2}t_2+a_{1,1}=y_{2,1} & \text{1 derivată dreapta}\\
3a_{2,3}t_2^2+2a_{2,2}t_2+a_{2,1}=y_{2,1} & \text{2 derivată stînga}\\
3a_{2,3}t_3^2+2a_{2,2}t_3+a_{2,1}=y_{3,1} & \text{2 derivată dreapta}
\end{cases}$$

În general $N\geq 3$ ⇨ 4N-4 ecuaţii, 4N-4 necunoscute

$$\begin{cases}
a_{1,3}t_1^3+a_{1,2}t_1^2+a_{1,1}t_1+a_{1,0}=y_1\\
a_{1,3}t_2^3+a_{1,2}t_2^2+a_{1,1}t_2+a_{1,0}=y_2\\
\vdots\\
a_{N-1,3}t_{N-1}^3+a_{N-1,2}t_{N-1}^2+a_{N-1,1}t_{N-1}+a_{N-1,0}=y_{N-1}\\
a_{N-1,3}t_N^3+a_{N-1,2}t_N^2+a_{N-1,1}t_N+a_{N-1,0}=y_N\\
3a_{1,3}t_1^2+2a_{1,2}t_1+a_{1,1}=y_{1,1}\\
3a_{1,3}t_2^2+2a_{1,2}t_2+a_{1,1}=y_{2,1}\\
\vdots\\
3a_{N-1,3}t_{N-1}^2+2a_{N-1,2}t_{N-1}+a_{N-1,1}=y_{N-1,1}\\
3a_{N-1,3}t_N^2+2a_{N-1,2}t_N+a_{N-1,1}=y_{N,1}
\end{cases}$$

Matricea sistemului nu are dimensiune prea mare?

Dacă se încearcă rezolvarea sistemului global, da. Însă fiecare polinom este unic determinat de 4 ecuaţii, astfel că ar putea fi construit separat.

$$a_{n,3}t_n^3+a_{n,2}t_n^2+a_{n,1}t_n+a_{n,0}=y_n$$
$$a_{n,3}t_{n+1}^3+a_{n,2}t_{n+1}^2+a_{n,1}t_{n+1}+a_{n,0}=y_{n+1}$$
$$3a_{n,3}t_n^2+2a_{n,2}t_n+a_{n,1}=y_{n,1}$$
$$3a_{n,3}t_{n+1}^2+2a_{n,2}t_{n+1}+a_{n,1}=y_{n+1,1}$$

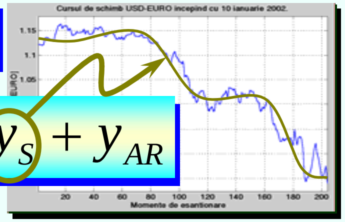
$\forall\, n\in\overline{1,N-1}$

☝ Matricea sistemului liniar global este, oricum, bloc diagonală.

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Revenim la



$$y_M \equiv y_T + y_S + y_{AR}$$

Cum se pot determina perioada şi coeficienţii sezonieri?

Există două tehnici de bază.

În timp: Wittacker-Robinson

Recomandată pentru determinarea coeficienţilor sezonieri.

În frecvenţă: periodograma Schuster

Recomandată pentru determinarea perioadei.

☝ Greoaie şi subiectivă. Nu va fi implementată.

Metoda Wittacker-Robinson

- Se poate aplica numai seriilor de timp uniform eşantionate.

☝ Seria staţionarizată neuniform eşantionată trebuie întâi interpolată şi apoi re-eşantionată uniform.

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

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

interpolare liniară / spline

$$y_{\text{sta}} : [1, T_{\max} = t_N] \to \mathbb{R}$$

(funcţie continuă)

Re-eşantionare

$$T_s \overset{def}{=} \min_{n \in \overline{1,N-1}} \{t_{n+1} - t_n\}$$

(pseudo-perioadă de eşantionare)

☝ Eeşantionare uniformă

$$M = N$$

$$\mathcal{D}' = \{y_{\text{sta}}[m] \overset{def}{=} y_{\text{sta}}(m\,T_s)\}_{n=\overline{1,M}}$$

$$M \overset{def}{=} \left\lfloor \frac{T_{\max}}{T_s} + \frac{1}{2} \right\rfloor$$

Principiul metodei

- Pentru fiecare perioadă posibilă ***P***, se rearanjează seria de timp într-o matrice cu ***P*** coloane şi se evaluează media fiecărei coloane.

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Metoda Wittacker-Robinson (continuare)

Aşadar $\mathcal{D}' = \{y_{\text{sta}}[m]\}_{n=\overline{1,M}}$ rearanjare $\mathbf{Y}_{\text{sta}} \in \mathbb{R}^{L \times P}$ $L \overset{def}{=} \left\lfloor \frac{M}{P} \right\rfloor$

indice temporal p

indice de segment k

$y_{\text{sta}}[1]$	$y_{\text{sta}}[2]$	...	$y_{\text{sta}}[p]$	...	$y_{\text{sta}}[P]$
$y_{\text{sta}}[P+1]$	$y_{\text{sta}}[P+2]$	...	$y_{\text{sta}}[P+p]$	...	$y_{\text{sta}}[2P]$
...	...	...	...	...	...
$y_{\text{sta}}[kP+1]$	$y_{\text{sta}}[kP+2]$	...	$y_{\text{sta}}[kP+p]$	...	$y_{\text{sta}}[(k+1)P]$
...	...				...
$y_{\text{sta}}[(L-1)P+1]$	$y_{\text{sta}}[(L-1)P+2]$	...	$y_{\text{sta}}[(L-1)P+p]$	...	$y_{\text{sta}}[LP]$

☝ Segment de semnal (pe o perioadă).

mediere pe coloane

$y_{S,1}$ $y_{S,2}$ ... $y_{S,p} \overset{def}{=} \dfrac{1}{L}\sum_{k=0}^{L-1} y[kP+p]$ ... $y_{S,P}$

Coeficienţi sezonieri pe durata unei perioade (pentru o perioadă precizată)

☝ În matricea de semnal $\mathbf{Y}_{\text{sta}}$ ultimele date staţionarizate pot lipsi.

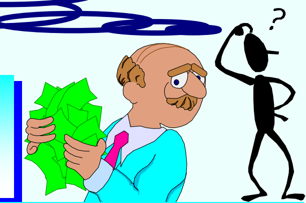
Evident Matricea de semnal trebuie să conţină minim două linii şi două coloane. $P \in \overline{2, \left\lfloor \frac{M}{2} \right\rfloor}$

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Metoda Wittacker-Robinson (continuare)

Cum se poate alege o perioadă adecvată?

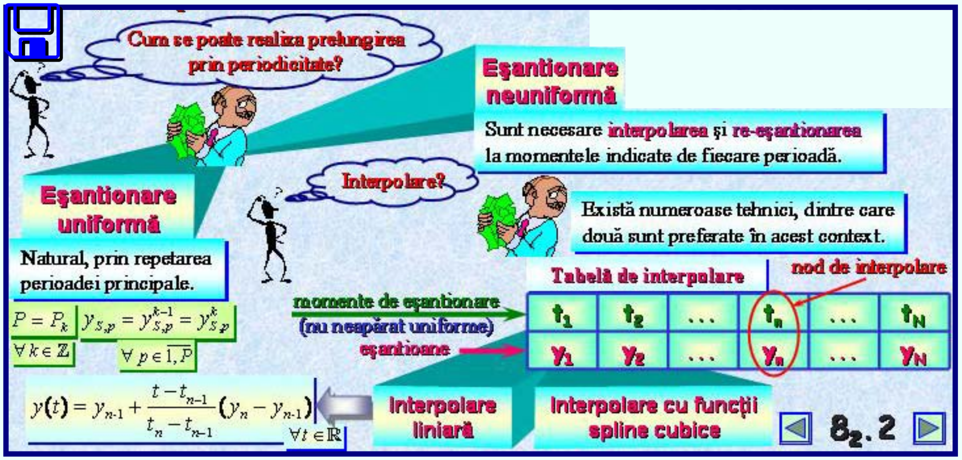


$$\mathcal{V}[P] \stackrel{def}{=} \sum_{n=1}^{N}\left[y_{sta}(t_n) - y_S^P(t_n)\right]^2 \qquad \forall P \in \overline{2, \left\lfloor \frac{M}{2} \right\rfloor}$$

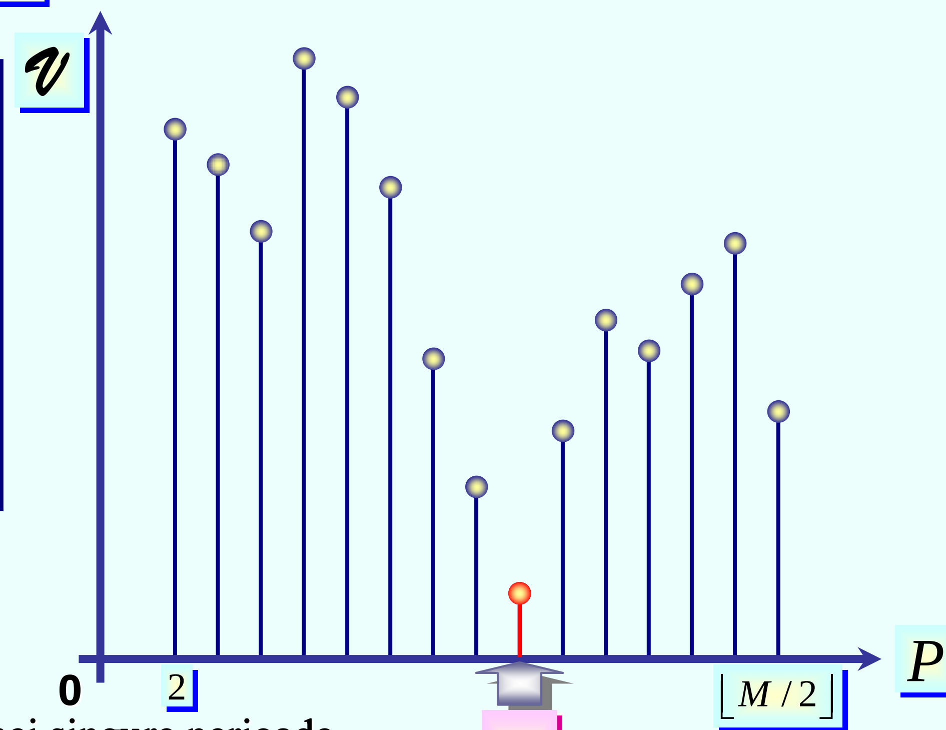
Folosind un criteriu de adecvanţă bazat pe eroarea pătratică.

Minimul acestui criteriu indică perioada adecvată.

Componenta sezonieră, de perioadă ***P***, obţinută în urma prelungirii prin periodicitate.



☝ În cazul interpolării spline, derivatele laterale din nodurile marginale se pot evalua ca şi cele din nodurile interioare, folosind periodicitatea.

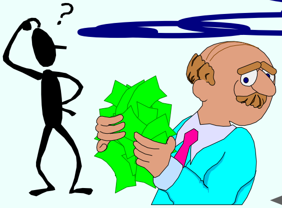


- Minimele locale pot indica divizori sau multipli ai unei singure perioade.
- Prea multe minime locale necorelate pot indica absenţa componentei sezoniere.

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Dar în cazul serilor de timp care nu au comportament sezonier?



Comportamentul sezonier poate fi detectat cu ajutorul următorul criteriu empiric:

Dacă $\mathcal{V}[P_0] < \frac{1}{4}\mathcal{E}[y_{\text{sta}}]$ **atunci seria de timp posedă componentă sezonieră.** → P_0 perioada estimată

energia seriei staţionarizate

$$\mathcal{E}[y_{\text{sta}}] \overset{def}{=} \sum_{n=1}^{N} y_{\text{sta}}^2(t_n)$$

Altfel $P_0 = 0$ & $y_S^{P_0} \equiv 0$

Care este tema de laborator?

Implementarea algoritmului de estimare a coeficienţilor sezonieri, bazat pe Metoda Wittacker-Robinson.

☞ Paşii principali ai algoritmului de modelare

Date de intrare $\mathcal{D}_{\text{sta}}$ (seria de timp staţionarizată)

Iniţializare
- Determinarea lungimii seriei de timp staţionarizate. → N
- Evaluarea perioadei de eşantionare. → T_s ☜ Unitară, pentru seriile eşantionate uniform.

❽.❷ Estimarea componentei sezoniere

☞ Paşii principali ai algoritmului de modelare (continuare)

Se interpolează şi re-eşantionează seria de timp staţionarizată (dacă este cazul). ☝ Cu două opţiuni: liniar, spline.

Se evaluează numărul de eşantioane. M ☝ Eeşantionare uniformă ⇒ $M = N$

Buclă iterativă

Pentru $P \in \overline{2,\lfloor M/2 \rfloor}$

① Se construieşte matricea de semnal cu *P* coloane. $\mathbf{Y}_{\text{sta}}$

② Se estimează valorile coeficienţilor sezonieri. $\left\{ y_{S,p} \right\}_{p \in \overline{1,P}}$ ☝ Acelaşi tip de interpolare ca şi pentru seria staţionarizată.

③ Se interpolează şi re-eşantionează seria coeficienţilor sezonieri (dacă este cazul).

④ Se construieşte componenta sezonieră folosind procedeul prelungirii prin periodicitate. y_S^P

⑤ Se calculează şi se memorează valoarea curentă a criteriului de adecvanţă. $\mathcal{V}[P]$

Se determină punctul de minim al criteriului de adecvanţă. $\left(P_0, \mathcal{V}\left[P_0 \right] \right)$

Dacă $\mathcal{V}\left[P_0 \right] < \frac{1}{4} \mathcal{E}\left[y_{\text{sta}} \right]$ atunci seria de timp **posedă componentă sezonieră**. ⇒ $\begin{cases} P = P_0 \\ y_S \equiv y_S^{P_0} \end{cases}$

Altfel $P = 0$ & $y_S \equiv 0$

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

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

Se evaluează zgomotul colorat. $v(t_n) = y_{\mathrm{sta}}(t_n) - y_S(t_n)$ $\forall n \in \overline{1,N}$

Date de ieşire v y_S P $\mathcal{V}$

Rutină ce trebuie proiectată

```
>> [v,yS,P,V] = seasonal(ysta);
```

- v: vectorul zgomotului colorat
- yS: vectorul componentei sezoniere pe orizontul de măsură
- P: perioada componentei sezoniere
- V: vectorul valorilor criteriului de adecvanţă
- ysta: vectorul de date care conţine seria de timp staţionarizată

Programul de test al rutinei `seasonal` **ISLAB_8B**

☞ **Paşi principali**

Se introduce numărul fişierului care conţine seria de timp staţionarizată. `nts`

Se încarcă seria de timp staţionarizată în spaţiul de lucru. • cu ajutorul funcţei: `eval`

Se apelează rutina `seasonal`.

❽ Modelarea şi predicţia seriilor de timp

❽.❷ Estimarea componentei sezoniere

Programul de test al rutinei `seasonal` **`ISLAB_8B`**

☞ **Paşi principali (continuare)**

Dacă $P \neq 0$:
- Se trasează graficul criteriului de adecvanţă (cu precizarea punctului de minim).
- Se trasează graficele seriei de timp staţionarizate şi al componentei sezoniere (în aceeaşi fereastră, cu precizarea perioadei).

Altfel
- Utilizatorul este informat printr-un mesaj corespunzător.

Se trasează graficul zgomotului colorat (cu precizarea dispersiei pe grafic).

$$\lambda_v^2 \overset{def}{=} \frac{1}{N}\sum_{n=1}^{N} v^2(t_n)$$

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

Mini-programe MATLAB.

`STnts_sta.M`

`nts∈{1,2,3,...,15}`