Metoda lui Bayes (continuare)

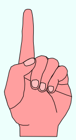
- Se poate arăta (deşi dificil) că, în anumite condiții, maximizarea probabilității din cadrul problemei lui Bayes conduce la minimizarea matricii de auto-covarianță a erorii de estimare.

Cum poate fi rezolvată problema lui Bayes?

Folosind o proprietate elementară de Teoria Probabilităților.

$$p(\theta,\mathcal{D}) = p(\mathcal{D}\,|\,\theta)\,p(\theta) = p(\theta\,|\,\mathcal{D})\,p(\mathcal{D})$$

$$\hat{\theta}(\mathcal{D}) = \underset{\theta\in\mathcal{S}\subseteq\mathbb{R}^{n\theta}}{\operatorname{argmax}}\, p(\theta\,|\,\mathcal{D}) = \underset{\theta\in\mathcal{S}\subseteq\mathbb{R}^{n\theta}}{\operatorname{argmax}} \frac{p(\mathcal{D}\,|\,\theta)\,p(\theta)}{p(\mathcal{D})} = \underset{\theta\in\mathcal{S}\subseteq\mathbb{R}^{n\theta}}{\operatorname{argmax}}\, p(\mathcal{D}\,|\,\theta)\,p(\theta)$$



Noua problemă se poate rezolva practic după o strategie recursivă.

☝ Probabilitatea de a obține setul măsurat de date nu depinde de parametrii care urmează a fi estimați.

- Dacă ambele densități de probabilitate din produsul ce trebuie maximizat sunt cunoscute, se poate utiliza o tehnică de optimizare (de exemplu, de tipul MNR).

  ☹ Cunoaşterea acestora este însă dificilă, dacă nu imposibilă.

  ☺ Este totuşi posibilă cunoaşterea densităților de probabilitate: $p(\mathcal{D}\,|\,\theta)$ & $p(\mathcal{D})$

Proprietatea de Teoria Probabilităților poate fi utilizată şi pentru a estima recursiv $p(\theta)$, plecînd de de la o inițializare.

Metoda lui Bayes (continuare)

Algoritmul lui Bayes

Date de intrare

- $\mathcal{D} = \{(u[n], y[n])\}_{n\in\overline{1,N}}$ (date intrare-ieşire măsurate)
- $p(\mathcal{D})$, $p(\mathcal{D}\,|\,\boldsymbol{\theta})$ (expresile densităţilor de probabilitate)
- $\eta > 0$ (prag de precizie)

Iniţializare

- $p_0(\boldsymbol{\theta})$ ➔ uniformă, dacă nu se poate stabili altfel $p_0(\boldsymbol{\theta}) \equiv \text{const.}$
- $\hat{\boldsymbol{\theta}}_0$ ➔ ales arbitrar în domeniul de variaţie
- $k = 0$ ➔ indicele iterativ iniţial

Buclă iterativă

① Se reactualizează densitatea de probabilitate a parametrilor necunoscuţi:

$$p_{k+1}(\boldsymbol{\theta}) \overset{def}{=} p_{k+1}(\boldsymbol{\theta}\,|\,\mathcal{D}) = \frac{p(\mathcal{D}\,|\,\boldsymbol{\theta})\,p_k(\boldsymbol{\theta})}{p(\mathcal{D})}$$

Vectorul parametrilor trebuie să fie din ce în ce mai bine adaptat la datele achiziţionate.

② Se evaluează punctul de maxim: $\hat{\boldsymbol{\theta}}_{k+1} = \underset{\boldsymbol{\theta}\in\mathcal{S}\subseteq\mathbb{R}^{n\theta}}{\operatorname{argmax}}\, p_{k+1}(\boldsymbol{\theta}\,|\,\mathcal{D})$

☝ Eventual, folosind o tehnică de optimizare.

$\left\|\hat{\boldsymbol{\theta}}_{k+1} - \hat{\boldsymbol{\theta}}_k\right\| < \eta$

NU: $k \leftarrow k+1$ (se revine la ①)

DA:

Date de ieşire

- $\hat{\boldsymbol{\theta}}_{k+1}$ ➔ aproximaţia de precizie η
- $k+1$ ➔ numărul de iteraţii efectuate

❹ Metode de identificare şi validare

❹.❼ Metode bazate pe Teoria Estimaţiei

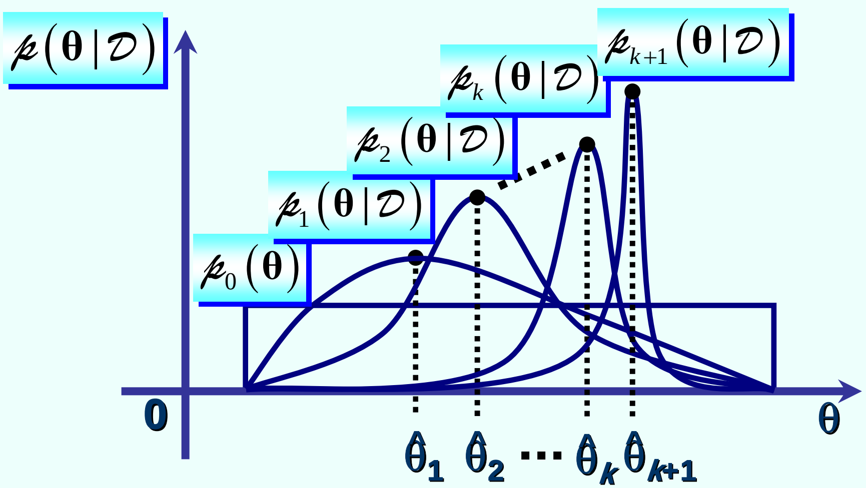
Metoda lui Bayes (continuare)

$$p_{k+1}(\boldsymbol{\theta}) \overset{def}{=} p_{k+1}(\boldsymbol{\theta}\,|\,\mathcal{D}) = \frac{p(\mathcal{D}\,|\,\boldsymbol{\theta})\,p_k(\boldsymbol{\theta})}{p(\mathcal{D})} \qquad \forall k \in \mathbb{N}^*$$

Pe măsură ce procesul iterativ avansează, corelaţia dintre parametrii estimaţi şi datele măsurate **trebuie să devină tot mai puternică**.

☝ **În acelaşi timp, densitatea de probabilitate se grupează tot mai mult în jurul unei valori maxime.**

Apariţia necondiţionată a parametrilor este tot mai puţin probabilă, ei fiind din ce în ce **mai puternic condiţionaţi de datele măsurate**.



- Densitatea de probabilitate a datelor achiziţionate **nu depinde de valorile parametrilor necunoscuţi**, astfel că relaţia recursivă se poate simplifica:

$$p_{k+1}(\boldsymbol{\theta}) \overset{def}{=} p_{k+1}(\boldsymbol{\theta}\,|\,\mathcal{D}) = p(\mathcal{D}\,|\,\boldsymbol{\theta})\,p_k(\boldsymbol{\theta}) \qquad \forall k \in \mathbb{N}^*$$

- Se poate arăta că acest proces iterativ este **convergent** către soluţia problemei lui Bayes.

☹ Algoritmul lui Bayes **necesită cunoaşterea prealabilă** cel puţin a densităţii de probabilitate a setului de date $\mathcal{D}$, condiţionate de vectorul θ, lucru care s-ar putea dovedi **dificil**.

☹ **Viteza de convergenţă** a algoritmului este **în general scăzută**, dacă se pleacă de la o **iniţializare uniformă**.

☺ Cu toate acestea, **algoritmul lui Bayes este unul dintre puţinii algoritmi implementabili** din cadrul **TE**.

Metoda Verosimilităţii Maxime (MVM)

- Metodă înrudită cu MB, care propune **rezolvarea problemei de identificare în formă completă**, fără a apela la un algoritm recursiv, **pentru anumite tipuri de modelel de identificare**.

Ipoteză simplificatoare

Indiferent de procesele furnizoare de date avînd structură fixată şi domeniu de stabilitate comun, vectorul parametrilor adevăraţi poate lua orice valoare din acest domeniu, cu aceeaşi probabilitate.

$$p(\boldsymbol{\theta}) \equiv \text{const.}$$

Vectorul parametrilor necunoscuţi θ este **echiprobabil** pe domeniul de stabilitate.

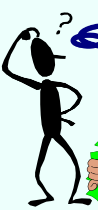
Problema verosimilităţii

$$\hat{\boldsymbol{\theta}}(\mathcal{D}) = \underset{\boldsymbol{\theta}\in\mathcal{S}\subseteq\mathbb{R}^{n\theta}}{\operatorname{argmax}} \boxed{p(\mathcal{D}\,|\,\boldsymbol{\theta})}$$

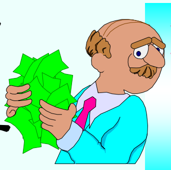
Verosimilitate

☝ Datele măsurate nu pot fi obţinute prin simulare decît dacă vectorul parametrilor modelului de identificare este **cel mai plauzibil sau verosimil**, adică produce o frecvenţă maximă de obţinere a setului de date măsurate.

Vectorii parametrilor care nu verifică această proprietate sunt declaraţi **mai puţin verosimili** sau chiar **neverosimili**.



Cum poate fi rezolvată problema verosimilităţii maxime?



În formă completă, dacă se folosesc **modele de regresie liniară**.

Metoda Verosimilităţii Maxime (continuare)

Exemplu Media şi varianţa – estimaţii de verosimilitate maximă

$\mathcal{P}(\theta^*)$

$y[n] = \theta^* + v[n]$ $\quad \forall n \in \mathbb{N}^*$ ← zgomot alb normal distribuit

☝ Aceeaşi dispersie necunoscută.

$$v \in za\left(0,(\lambda^*)^2\right) \cap \mathcal{N}\left(0,(\lambda^*)^2\right)$$

$$p(v[n]) \overset{def}{=} \frac{1}{\sqrt{2\pi}\lambda^*}\exp\left[-\frac{v^2[n]}{2(\lambda^*)^2}\right] \quad \forall n \in \mathbb{N}^*$$

← zgomotul afectează direct datele măsurate

$$y \in za\left(\theta^*,\sigma_y^2\right) \cap \mathcal{N}\left(\theta^*,(\lambda^*)^2\right)$$

Exerciţiu

$$\sigma_y^2 = (\theta^*)^2 + (\lambda^*)^2$$

Se măsoară valorile unui parametru în mod direct.

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

$$\hat{\theta}(\mathcal{D}) = \underset{\theta\in\mathbb{R}}{\operatorname{argmax}}\, p\left(\mathcal{D}\,|\,\theta,\lambda^2\right) = ?$$

$$\hat{\lambda}^2(\mathcal{D}) = \underset{\lambda^2\in\mathbb{R}_+^*}{\operatorname{argmax}}\, p\left(\mathcal{D}\,|\,\theta,\lambda^2\right) = ?$$

$$p(y[n]) \overset{def}{=} \frac{1}{\sqrt{2\pi}\lambda^*}\exp\left[-\frac{\left(y[n]-\theta^*\right)^2}{2(\lambda^*)^2}\right] \quad \forall n \in \mathbb{N}^*$$

$\mathcal{M}(\theta)$

$$y[n] = \theta + \varepsilon\left[n,\theta,\lambda^2\right] \quad \forall n \in \mathbb{N}^*$$

$$\varepsilon\left[n,\theta^*,(\lambda^*)^2\right] = v[n]$$

(modelul de identificare corespunzător)

$$p\left(y[n]\,|\,\theta,\lambda^2\right) \overset{def}{=} \frac{1}{\sqrt{2\pi}\lambda}\exp\left[-\frac{\left(y[n]-\theta\right)^2}{2\lambda^2}\right] \quad \forall n \in \mathbb{N}^*$$

☝ Verosimilitatea fiecărui eşantion.

Metoda Verosimilităţii Maxime (continuare)

Exemplu Media şi varianţa – estimaţii de verosimilitate maximă (continuare)

$$\wp\left(y[n]\,|\,\theta,\lambda^2\right) \stackrel{def}{=} \frac{1}{\sqrt{2\pi\lambda}}\exp\left[-\frac{\left(y[n]-\theta\right)^2}{2\lambda^2}\right] \quad \forall n\in\mathbb{N}^*$$

$$\wp\left(\mathcal{D}\,|\,\theta,\lambda^2\right) = \wp\left(y[1],\ldots,y[N]\,|\,\theta,\lambda^2\right) = \prod_{n=1}^{N}\wp\left(y[n]\,|\,\theta,\lambda^2\right) \quad \forall n\in\mathbb{N}^*$$

☝ **Propoziţia 2** (orice proces neautocorelat şi Gaussian este independent)

$$\wp\left(\mathcal{D}\,|\,\theta,\lambda^2\right) = \frac{1}{\left(\sqrt{2\pi\lambda}\right)^N}\prod_{n=1}^{N}\exp\left[-\frac{\left(y[n]-\theta\right)^2}{2\lambda^2}\right] = \boxed{\frac{1}{\left(\sqrt{2\pi\lambda}\right)^N}\exp\left[-\frac{1}{2\lambda^2}\sum_{n=1}^{N}\left(y[n]-\theta\right)^2\right]} \quad \forall\left(\theta,\lambda^2\right)\in\mathbb{R}\times\mathbb{R}_+^*$$

Aşadar Pentru a rezolva problema verosimilităţii, trebuie determinat punctul de maxim al unei funcţii Gaussiene generalizate.

Punctul de maxim este o rădăcină a gradientului. ☝ Indefinit derivabilă în raport cu oricare dintre cei 2 parametri.

☹ Evaluarea gradientului este complicată.

Forma funcţiei sugerează că extremele ei coincid cu cele ale logaritmului natural aplicat acesteia.

$$\left(\hat{\theta},\hat{\lambda}^2\right)(\mathcal{D}) = \underset{\left(\theta,\lambda^2\right)\in\mathbb{R}\times\mathbb{R}_+^*}{\operatorname{argmax}}\left\{\boxed{\ln\left[\wp\left(\mathcal{D}\,|\,\theta,\lambda^2\right)\right]}\right\}$$

$$F(\theta,\lambda^2)$$

$$F(\theta,\lambda^2) \stackrel{def}{=} -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\lambda^2 - \frac{1}{2\lambda^2}\sum_{n=1}^{N}\left(y[n]-\theta\right)^2 \quad \forall\left(\theta,\lambda^2\right)\in\mathbb{R}\times\mathbb{R}_+^*$$

☺ Mai simpu de derivat.

Metoda Verosimilităţii Maxime (continuare)

Exemplu — Media şi varianţa – estimaţii de verosimilitate maximă (continuare)

$$F(\theta,\lambda^2) \stackrel{def}{=} -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\lambda^2 - \frac{1}{2\lambda^2}\sum_{n=1}^{N}\left(y[n]-\theta\right)^2$$

$$\forall(\theta,\lambda^2)\in\mathbb{R}\times\mathbb{R}_+^*$$

$$\begin{cases} \dfrac{\partial}{\partial\theta}F(\theta,\lambda^2) = \dfrac{1}{\lambda^2}\displaystyle\sum_{n=1}^{N}\left(y[n]-\theta\right) = 0 \\ \dfrac{\partial}{\partial\lambda^2}F(\theta,\lambda^2) = -\dfrac{N}{2\lambda^2} + \dfrac{1}{2\lambda^4}\displaystyle\sum_{n=1}^{N}\left(y[n]-\theta\right)^2 = 0 \end{cases}
\Leftrightarrow
\begin{cases} \hat{\theta}(\mathcal{D}) = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} y[n] \\ \hat{\lambda}^2(\mathcal{D}) = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N}\left(y[n]-\hat{\theta}(\mathcal{D})\right)^2 \end{cases}$$

(parametru)

(1/precizie)

☝ Soluţia este unică.

Medie

Varianţă

+

$$\hat{v}[n] = y[n]-\hat{\theta}(\mathcal{D}) = y[n] - \frac{1}{N}\sum_{n=1}^{N} y[n]$$

(zgomot perturbator)

În concluzie

Media şi varianţa sunt cele mai verosimile (plauzibile) estimaţii.

☝ Estimaţiile sunt consistente, datorită IE.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} y[n] = E\{y[n]\} = E\left\{\theta^* + v[n]\right\} = \theta^*$$

$$v \in za\left(0,\left(\lambda^*\right)^2\right)$$

- Nu întîmplător prima caracterizare statistică a numeroase procese nedeterministe, chiar dacă este grosieră, constă în evaluarea mediilor şi varianţelor acestora.
- **Exemplu**: media generală a unui student, calculată după 3 ani de studii, are şanse mici să varieze semnificativ după 4 ani de studii, dacă notele sale sunt grupate în jurul acesteia şi şanse mari să varieze sensibil în finalul celor 4 ani, dacă notele sale sunt larg dispersate în jurul acesteia.

❹ Metode de identificare şi validare

❹.❼ Metode bazate pe Teoria Estimaţiei

Metoda Verosimilităţii Maxime (continuare)

Dar în cazul modelelor de regresie liniară?

Se poate demonstra un rezultat remarcabil.

Teorema 6 (MVI şi MCMMP)

În contextul modelelor de regresie liniară, următoarele două ipoteze se consideră verificate:

a. matricea R (sau R_N) nu este neapărat deterministă, dar este (strict) pozitiv definită (adică inversabilă) pentru toate dimensiunile orizontului de măsură suficient de mari;

b. perturbaţia v aparţine clasei $za\left(0,\left(\lambda^*\right)^2\right)\cap\mathcal{N}\left(0,\left(\lambda^*\right)^2\right)$

(adică este un zgomot alb de medie nulă şi dispersie necunoscută , cu distribuţie Gaussiană centrată în zero, de varianţă egală tot cu dispersia necunoscută).

Atunci estimaţiile vectorului parametrilor adevăraţi şi dispersiei zgomotului alb obţinute aplicînd MCMMP sunt de verosimilitate maximă.

Demonstraţie → Exerciţiu

- Chiar dacă proprietatea de consistenţă este dificil de verificat în contextul Teoremei 6 (matricea regresorilor nu mai este strict deterministă), cel puţin rămîne proprietatea de verosimilitate maximă.

☺ Probabilitatea de a obţine setul de date măsurate prin simularea cu un model identificat folosind MCMMP este maximă.

❹ Metode de identificare şi validare

❹.❽ Identificarea şi predicţia proceselor auto-regresive

Contextul de lucru

☝ Dispersie necunoscută.

$e \in za\left(0,\lambda^2\right)$

☝ Model de tip AR.

$\mathcal{P}(\theta^*)$

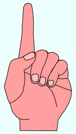
$A^*(q^{-1})y[n] = e[n]$ $\quad \forall n \in \mathbb{N}^*$ $\iff$ $y[n] = \boldsymbol{\varphi}^T[n]\boldsymbol{\theta}^* + e[n]$ $\quad \forall n \in \mathbb{N}^*$

$\mathcal{M}(\theta)$

$A(q^{-1})y[n] = \varepsilon[n,\boldsymbol{\theta}]$ $\quad \forall n \in \mathbb{N}^*$ $\iff$ $y[n] = \boldsymbol{\varphi}^T[n]\boldsymbol{\theta} + \varepsilon[n,\boldsymbol{\theta}]$ $\quad \forall n \in \mathbb{N}^*$

$\varepsilon\left[n,\theta^*\right] = e[n]$ $\quad \forall n \in \mathbb{N}^*$

$$\left[\begin{array}{l} \boldsymbol{\varphi}^T[n] \overset{def}{=} \left[-y[n-1] \quad \cdots \quad -y[n-na]\right] \\ \boldsymbol{\theta}^T \overset{def}{=} \left[a_1 \quad \cdots \quad a_{na}\right] \end{array}\right.$$



Vectorul regresorilor nu este determinist,
dar este măsurabil.

- Modelul autoregresiv (AR) conduce la una dintre cele mai atractive şi simple abordări pentru identificarea surselor de perturbaţii.
- În pofida preciziei sale limitate, modelul AR este adesea preferat în aplicaţii, în special pentru metoda extrem de rapidă şi eficientă de estimare a parametrilor săi.
- Modelul concurent al acestuia în aplicaţii este ARMA, care, deşi necesită utilizarea MCMMPE sau MMEP pentru estimarea parametrilor săi, este tot mai des utilizat, datorită creşterii sensibile a performanţelor tehnicii de calcul automat.

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

date măsurate

$$\left\{\begin{array}{l} \hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^{na}}{\operatorname{argmin}} \mathcal{V}(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}\in\mathbb{R}^{na}}{\operatorname{argmin}} \sum_{n=1}^{N}\left(y[n] - \boldsymbol{\varphi}^T[n]\boldsymbol{\theta}\right)^2 = ? \\ \hat{\lambda}^2 = \dfrac{1}{\gamma_N}\mathcal{V}\left(\hat{\boldsymbol{\theta}}\right) \end{array}\right.$$

$\gamma_N \in \{N-na, N\}$

❹ Metode de identificare şi validare

❹.❽ Identificarea şi predicţia proceselor autoregresive

Performanţele MCMMP în cazul modelului AR

$$\varphi^T[n] \overset{def}{=} \left[-y[n-1] \quad \cdots \quad -y[n-na]\right]$$

$$\theta^T \overset{def}{=} \left[a_1 \quad \cdots \quad a_{na}\right]$$

$$\Rightarrow \begin{cases} \hat{\theta} = \mathbf{R}^{-1}\mathbf{r} \\ \hat{\lambda}^2 = \dfrac{1}{\gamma_N}\mathcal{V}(\hat{\theta}) \end{cases} \qquad \gamma_N \in \{N-na, N\}$$

$$\mathbf{R} \overset{def}{=} \sum_{n=1}^{N} \varphi[n]\varphi^T[n] = \left[\sum_{n=\max\{i,j\}+1}^{N} y[n-i]y[n-j]\right]_{i,j\in\overline{1,na}}$$

$$\mathbf{r} \overset{def}{=} \sum_{n=1}^{N} \varphi[n]y[n] = -\left[\sum_{n=i+1}^{N} y[n]y[n-i]\right]_{i\in\overline{1,na}}$$

☝ Simetrică, dar nu neapărat de tip Toeplitz (elementele de pe diagonala principală sunt diferite).

Pentru inversare, se foloseşte Algoritmul (clasic al) lui Gauss.

Exerciţiu

- Estimaţiile oferite de MCMMP sunt consistente, chiar dacă vectorul regresorilor nu este perfect determinist.

Efortul de calcul

	Înmulţiri	Adunări
Construcţia matricii **R**	$\left[\dfrac{na(na+1)(3N-2na-1)}{6}\right]_\times$	$\left[\dfrac{na(na+1)(3N-2na-4)}{6}\right]_+$
Construcţia vectorului **r**	$\left[N-na\right]_\times$	$\left[N-na-1\right]_+$
Inversarea matricii **R**	$\left[\dfrac{na(3na^2-3na+2)}{2}\right]_\times$	$\left[\dfrac{3na(na-1)^2}{2}\right]_+$
Estimarea parametrilor ($\mathbf{R}^{-1}\mathbf{r}$)	$\left[na^2\right]_\times$	$\left[na(na-1)\right]_+$

Total

$$O_{MCMMP} \sim \left[\frac{na\left[7na^2+3na(N-2)+3N+5\right]}{6}\right]_\times + \left[\frac{na\left[7na^2+3na(N-6)+3N-1\right]}{6}\right]_+$$

❹ Metode de identificare şi validare

❹.❽ Identificarea şi predicţia proceselor autoregresive

☹ Efortul de calcul al MCMMP cu modelul AR este susţinut, deoarece aît numărul datelor măsurate cît şi numărul de parametri au valori mari (pentru a asigura o precizie suficientă).

Obiectiv: Reducerea efortului de calcul prin metode alternative de identificare.

- Metoda Yule-Walker-Wiener
- Algoritmul Levinson-Durbin

Metoda Yule-Walker-Wiener (MYWW)

Ideea lui G.U. Yule & G. Walker (anii'30) ⇒ Estimarea parametrilor necunoscuţi şi a dispersiei zgomotului alb se poate realiza apelînd la secvenţa de auto-covarianţă a ieşirii.

- Se aplică două operaţii succesive asupra ecuaţiei procesului:

$$E \;②\; \Big|\; y[n-k]\times \;①\; \Big|\; A^*(q^{-1})y[n]=e[n] \qquad \forall k\in\mathbb{N},\ \forall n\in\mathbb{N}^*$$

$$r_y[k]+a_1^* r_y[k-1]+\cdots+a_{na}^* r_y[k-na]=r_{e,y}[k] \qquad \forall k\in\mathbb{N} \quad ?$$

ecuaţie autoregresivă deterministă

Împărţire infinită.

$$y[n]=e[n]+\sum_{m\geq 1}\alpha_m^* e[n-m] \;\;\forall n\in\mathbb{N}^* \iff y[n]=\frac{1}{A^*(q^{-1})}e[n] \;\;\forall n\in\mathbb{N}^* \iff A^*(q^{-1})y[n]=e[n] \;\;\forall n\in\mathbb{N}^*$$

$$r_{e,y}[k]\overset{def}{=}E\{e[n]y[n-k]\}=\lambda^2\delta_0[k]+\lambda^2\sum_{m\geq 1}\alpha_m^*\delta_0[k+m]=\lambda^2\delta_0[k] \qquad \forall k\in\mathbb{N}$$

❹ Metode de identificare şi validare

❹.❽ Identificarea şi predicţia proceselor autoregresive

Metoda Yule-Walker-Wiener (continuare)

$$r_{e,y}[k] \overset{def}{=} E\{e[n]y[n-k]\} = \lambda^2\delta_0[k] + \lambda^2\sum_{m\geq 1}\alpha_m^*\delta_0[k+m] = \lambda^2\delta_0[k] \quad \forall k \in \mathbb{N}$$

☝ Ieşirea nu este corelată cu valori viitoare ale zgomotului.

Ecuaţiile Yule-Walker-Wiener

$$r_y[k] + a_1^* r_y[k-1] + \cdots + a_{na}^* r_y[k-na] = \lambda^2\delta_0[k] \quad \forall k \in \mathbb{N}$$

(similare ecuaţiilor Wiener-Hopf)

☹ Sistemul este **inoperant**, deoarece are un **număr infinit de ecuaţii** şi apelează la **secvenţa ideală de auto-covarianţă a ieşirii**.

Se renunţă la sistemul ideal (care ar conduce la determinarea exactă a parametrilor adevăraţi) şi se înlocuieşte cu un sistem practic, compatibil.

$$\begin{cases} r_y[0] + a_1^* r_y[1] + \cdots + a_{na}^* r_y[na] - \lambda^2 = 0 \\ r_y[1] + a_1^* r_y[0] + \cdots + a_{na}^* r_y[na-1] = 0 \\ r_y[2] + a_1^* r_y[1] + \cdots + a_{na}^* r_y[na-2] = 0 \\ \vdots \\ r_y[na] + a_1^* r_y[na-1] + \cdots + a_{na}^* r_y[0] = 0 \end{cases} \Leftrightarrow \begin{cases} a_1^* r_y[1] + \cdots + a_{na}^* r_y[na] - \lambda^2 = -r_y[0] \\ \underbrace{\begin{bmatrix} r_y[0] & r_y[1] & \cdots & r_y[na-1] \\ r_y[1] & r_y[0] & \cdots & r_y[na-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_y[na-1] & r_y[na-2] & \cdots & r_y[0] \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_{na}^* \end{bmatrix} = -\overbrace{\begin{bmatrix} r_y[1] \\ r_y[2] \\ \vdots \\ r_y[na] \end{bmatrix}}^{\mathbf{r}} \end{cases}$$

☝ (*na*+1) ecuaţii.

Sistemul Yule-Walker-Wiener

❹ Metode de identificare şi validare

❹.❽ Identificarea şi predicţia proceselor autoregresive

Metoda Yule-Walker-Wiener (continuare)

$$\begin{cases} a_1^* r_y[1] + \cdots + a_{na}^* r_y[na] - \lambda^2 = -r_y[0] \\ \underbrace{\begin{bmatrix} r_y[0] & r_y[1] & \cdots & r_y[na-1] \\ r_y[1] & r_y[0] & \cdots & r_y[na-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_y[na-1] & r_y[na-2] & \cdots & r_y[0] \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_{na}^* \end{bmatrix} = -\overbrace{\begin{bmatrix} r_y[1] \\ r_y[2] \\ \vdots \\ r_y[na] \end{bmatrix}}^{\mathbf{r}} \end{cases}$$

Soluţia teoretică

$$\begin{cases} \boldsymbol{\theta}^* = \mathbf{R}^{-1}\mathbf{r} \\ \lambda^2 = r_y[0] - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \end{cases}$$

Soluţia practică

$$\begin{cases} \hat{\boldsymbol{\theta}}_N = \mathbf{R}_N^{-1}\mathbf{r}_N \\ \hat{\lambda}_N^2 = r_y^N[0] - \mathbf{r}_N^T \mathbf{R}_N^{-1} \mathbf{r}_N \end{cases}$$

$$\mathbf{R}_N \overset{def}{=} \begin{bmatrix} r_y^N[0] & r_y^N[1] & \cdots & r_y^N[na-1] \\ r_y^N[1] & r_y^N[0] & \cdots & r_y^N[na-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_y^N[na-1] & r_y^N[na-2] & \cdots & r_y^N[0] \end{bmatrix}$$

☝ Simetrică şi de tip Toeplitz.

$$\mathbf{r}_N \overset{def}{=} -\begin{bmatrix} r_y^N[1] \\ r_y^N[2] \\ \vdots \\ r_y^N[na] \end{bmatrix}$$

Vectorul estimat al parametrilor este de regulă bine definit, dar estimaţia dispersiei zgomotului alb poate fi negativă, din cauza erorilor numerice.

Efortul de calcul

	Înmulţiri	Adunări
Construcţia matricii $\mathbf{R}_N$	$\left[\dfrac{na(2N-na+1)}{2}\right]_\times$	$\left[\dfrac{na(2N-na-1)}{2}\right]_+$
Construcţia vectorului $\mathbf{r}_N$	$[N-na]_\times$	$[N-na-1]_+$
Inversarea matricii $\mathbf{R}_N$	$\left[\dfrac{na^2(3na-1)}{2}\right]_\times$	$\left[\dfrac{na(na-1)(3na-1)}{2}\right]_+$
Estimarea parametrilor ($\mathbf{R}_N^{-1}\mathbf{r}_N$)	$[na^2]_\times$	$[na(na-1)]_+$

Total

$$O_{MYWW} \sim \left[\frac{na(3na^2+2N+1)}{2}\right]_\times + \left[\frac{na(3na^2-3na+2N-2)}{2}\right]_+$$

☝ Efortul de calcul a scăzut.

Exemplu

$N = 1000$ $na = 10$

$O_{MCMMP} \sim [56075]_\times + [55865]_+$

$O_{MYWW} \sim [11505]_\times + [11340]_+$