*Dinamikus rendszerek paramétereinek becslése*

# *A maximum likelihood (ML) becslés*
# *A becsült paraméterek kovariancia mátrixa*

**Hangos Katalin**

PE Villamosmérnöki és Információs Rendszerek Tanszék

*ISMÉTLÉS*

# Együttes valószínűségi sűrűségfüggvény

A $\xi_1, ..., \xi_n$ skalár értékű valószínűségi változók együttes valószínűségi sűrűségfüggvénye egy $n$ változós nem-negatív függvény $f(x_1, ..., x_n)$, amelyre

$$P[a_1 \leq x_1 \leq b_1,\ ...,\ a_n \leq x_n \leq b_n] = \int_{a_1}^{b_1} ... \int_{a_n}^{b_n} f(x_1, ..., x_n) dx_1 ... dx_n$$

Marginális valószínűségi sűrűségfüggvény (a 2 változós esetben)

$$f_{\xi_1}(x_1) = \int_{-\infty}^{\infty} f_{\xi_1,\xi_2}(x_1, x_2) dx_2$$

**Függetlenség** (a 2 változós esetben)

$$f_{\xi_1,\xi_2}(x_1, x_2) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2)$$

## Kovariancia mártixok

A $\xi$ valószínűségi változó vektor értékű, azaz

$$\xi \;:\; \xi(\omega), \quad \omega \in \Omega, \quad \xi(\omega) \in \mathbb{R}^{\mu} \tag{1}$$

A változó $m \in R^{\mu}$ várható értéke egy valós vektor.
Varianciája egy valós elemű négyzetes mátrix, az ún. *kovariancia mátrix*:

$$COV\{\xi\} = E\{(\xi - E\{\xi\})(\xi - E\{\xi\})^T\}$$

Két skalár értékű valószínűségi változó $\xi_i$ és $\xi_j$ *kovarianciája*:

$$COV\{\xi_i, \xi_j\} = E\{(\xi_i - E\{\xi_i\})(\xi_j - E\{\xi_j\})\}$$

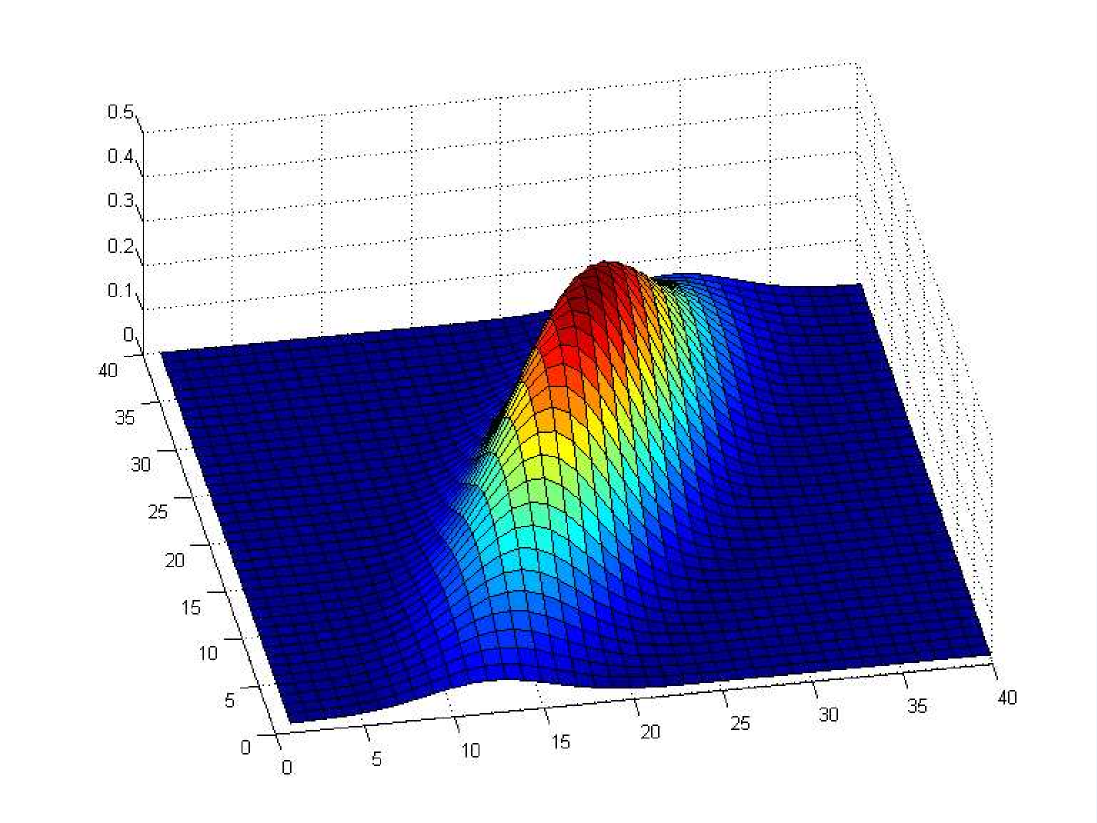
A *kovariancia mátrixok pozitív definit szimmetrikus mátrixok*, azaz

$$z^T COV\{\xi\} z \geq 0 \quad , \quad \forall z \in \mathbb{R}^{\mu}$$

# Két dimenziós normális eloszlás

Valószínűségi sűrűségfüggvény:

$$f(x_1, x_2) = \frac{1}{2\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left(\frac{(x_1-m_1)^2}{\sigma_1^2} - 2r\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right)}$$



# A predikciós hiba minimalizálása

A paraméterbecslési módszer: $D^N \longrightarrow \hat{\theta}_N$
**Feladatkitűzés**: Adott

- mért értékek: $D[1, N] = D^N = \{(y(k), u(k)) \mid k = 1, \ldots N\}$
- prediktív parametrizált modell: $\hat{y}(k|\theta) = g(k, D[1, k-1]; \theta)$
  predikciós hiba sorozat (diszkrét idejű jel): $\varepsilon(k, \theta) = y(k) - \hat{y}(k|\theta) \;\;, \;\; k = 1, \ldots, N$
- norma a predikciós hibán: $V_N(\theta, D^N) = \frac{1}{N} \sum_{k=1}^{N} \ell(\varepsilon(k, \theta))$ ahol $\ell(.)$ egy pozitív skalár értékű függvény, leggyakrabban: $\ell(\varepsilon) = \frac{1}{2}\varepsilon^2$

*Az ismert $D^N$ mért érték sorozatból és a $\theta$ paraméter vektor értékéből kiszámíthatjuk a $V_N(\theta, D^N)$ norma értéket. A $k = N$ időpillanatban válasszuk meg a becsült $\hat{\theta}_N$ paraméter vektort úgy, hogy*

$$\hat{\theta}_N = \hat{\theta}_N(D^N) = \arg\min_{\theta} V_N(\theta, D^N)$$

# Példa: SISO ARX modellek

**ALAPESET:** Az általános alakú input-output modellben a mozgóátlag tag nulla, azaz *a kimeneti rendszerzaj fehér*

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k) + e(k)$$

A modell prediktív alakja:

$$\hat{y}(k|\theta) = -a_1 y(k-1) - a_2 y(k-2) \cdots - a_n y(k-n) + b_0 u(k) + \cdots + b_m u(k-m)$$

A paramétervektor:

$$\theta = [-a_1 \ \ -a_2 \ \ \ldots \ \ -a_n \ \ b_0 \ b_1 \ \ \ldots \ \ b_m]^T$$

A predikciós hiba (**fehérzaj!**):

$$\varepsilon(k) = \hat{y}(k|\theta) - y(k) = e(k)$$

# LKN becslés paraméterekben lineáris esetben

A predikciós hiba:

$$\varepsilon(k,\theta) = y(k) - \theta^T \varphi(k)$$

A minimalizálandó kritériumfüggvény:

$$V_N(\theta, D^N) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{2} \left[ y(k) - \theta^T \varphi(k) \right]^2$$

aminek minimalizálása $\theta$ függvényében adja a LKN-becslést:

$$\hat{\theta}_{LKN} = \left[ \frac{1}{N} \sum_{k=1}^{N} \varphi(k) \varphi^T(k) \right]^{-1} \frac{1}{N} \sum_{k=1}^{N} \varphi(k) y(k)$$

# *A MAXIMUM LIKELIHOOD (ML) BECSLÉS*

# Valószínűségi modellen alapuló becslések

Cél: egy ismeretlen $\theta$ paraméter becslése
Adottak:

- A paraméterről információt hordozó $S(y^N) = \{y(1), \ldots, y(N)\}$ megfigyelések (skalár értékű valószínűségi változók) *nem feltétlenül függetlenek, és azonos eloszlásúak*
- Ezek együttes valószínűségi sűrűségfüggvénye

$$f(\theta; x_1, \ldots, x_N) = f_y(\theta; x^N) \quad , \quad f_y \ : \ \mathbb{R}^d \times \mathbb{R}^N \ \longrightarrow \ \mathbb{R}$$

  amely függ az ismeretlen $\theta \in \mathbb{R}^d$ paramétertől,
- A paraméter becslésére konstruált *becslő*: $\hat{\theta}(y^N)$ , $\hat{\theta} \ : \ \mathbb{R}^N \ \longrightarrow \ \mathbb{R}^d$ egy $N$-változós $d$ dimenziós vektor értékű függvény.

A $y_*^N$ konkrét megfigyelt értékeket, amelyek determinisztikusak a fenti becslő formulába behelyettesítve kapjuk a paraméter $\hat{\theta}_* = \hat{\theta}(y_*^N)$ becslését.

# A maximum likelihood becslés elve

Az úgynevezett **maximum likelihood becslő** *a megfigyelt esemény valószínűségét maximalizálja*.
A megfigyelt mitaelemek $f_y$ együttes valószínűségi sűrűségfüggvénye a konkrét megfigyelt $y_*^N$ értékek behelyettesítése után az ismeretlen $\theta$ paraméter egy determinisztikus függvénye, amelyet

$$\ell(\theta; y_*^N) = f_y(\theta; y_*^N)$$

*likelihood függvény*nek nevezünk.
A *maximum likelihood becslés* az a

$$\hat{\theta}_{ML}(y_*^N) = \arg\max_\theta f_y(\theta; y_*^N)$$

érték, amely az együttes valószínűségi sűrűségfüggvényt maximalizálja.

# A ML becslés normális eloszlás esetén

Normális eloszlású val. változók esetén az együttes eloszlás is normális lesz, ezért a gyakorlatban a fenti maximalizálás helyett a

$$\arg\max_{\theta} \ell(\theta; y_*^N) \equiv \arg\min_{\theta} \left[ -\ln \left[ \ell(\theta; y_*^N) \right] \right]$$

minimalizálást szoktuk elvégezni úgy, hogy

$$\frac{d \ln(\ell)}{d\theta} = 0$$

legyen ahol $ln(.)$ a természetes alapú logaritmusfüggvény.

## A ML becslés független és azonos eloszlású megfigyelésekre

Ha az $S(y^N)$ mintában szereplő mintaelemek független és azonos eloszlású valószínűségi változók, akkor együttes valószínűségi sűrűségfüggvényük a sűrűségfüggvények szorzata:

$$\ell(\theta; y_*^N) = \prod_{i=1}^{N} f_{y(i)}(\theta; y_{*_i}) = \prod_{i=1}^{N} f_y(\theta; y_{*_i})$$

Ekkor a *log-likelihood függvény* összeg alakú:

$$\ln[\ell(\theta; y_*^N)] = \sum_{i=1}^{N} \ln f_y(\theta; y_{*_i})$$

# Példa: Egy egyszerű maximum likelihood becslés 1.

**Minta**: független skalár értékű normális eloszlású, azonos, de *ismeretlen* $\theta_0$ várható értékű és mintaelemenként különböző $\lambda_i$ szórásnégyzetű $y(i)$ valószínűségi változók:

$$y(i) \quad \sim \quad \mathbb{N}(\theta_0, \lambda_i)$$

Az ismeretlen várható érték becslése *mintaközép*pel:

$$\hat{\theta}_{SM}(y^N) = \frac{1}{N} \sum_{i=1}^{N} y(i)$$

A mintaelemek független normális eloszlású val. változók a

$$f_{y(i)}(x_i) = \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{(x_i - \theta)^2}{2\lambda_i}}$$

# Példa: Egy egyszerű maximum likelihood becslés 2.

Az együttes sűrűségfüggvény:

$$f_y(\theta; x^N) = \prod_{i=1}^{N} \left( \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{(x_i-\theta)^2}{2\lambda_i}} \right)$$

A log-likelihood függvényt minimalizáljuk $\theta$ függvényében:

$$\hat{\theta}_{ML}(y^N) = \arg\min_{\theta} \ln f_y(\theta; y^N) = \arg\min_{\theta} \{ -\frac{N}{2} \ln 2\pi - \sum_{i=1}^{N} \frac{1}{2} \ln \lambda_i + \frac{1}{2} \sum_{i=1}^{N} \frac{(y(i)-\theta)^2}{\lambda_i} \}$$

A ML becslés:

$$\hat{\theta}_{ML}(y^N) = \frac{1}{\sum_{i=1}^{N} (1/\lambda_i)} \sum_{i=1}^{N} \frac{y(i)}{\lambda_i}$$

A maximum likelihood becslés figyelembe veszi a mintaelemek szórására vonatkozó információt is, és egy súlyozott mintaközepet használ becslőként.

# Prediktív modellek paramétereinek ML becslése 1.

**A teljes valószinűségi modell**

$$\mathbb{M}(\theta): \qquad \hat{y}(k \mid \theta) = g(k, D^{k-1}; \theta)$$

$\varepsilon(k,\theta) = y(k) - \hat{y}(k \mid \theta)$ függetlenek,

adott $\mathbf{f}_\varepsilon(\mathbf{x}; \mathbf{k}; \theta)$ sűrűségfüggvénnyel

**A likelihood függvény és a maximum likelihood becslés**

A $\varepsilon(k,\theta)$ hibák függetlenek, így az együttes sűrűségfüggvény:

$$\hat{f}_y(\theta; y^N) = \prod_{i=1}^{N} f_\varepsilon(y(i) - g(i, D^{i-1}; \theta), i; \theta) = \prod_{i=1}^{N} f_\varepsilon(\varepsilon(i,\theta), i; \theta)$$

a log-likelihood függvény pedig

$$\ell(\varepsilon, \theta, k) = -\ln f_\varepsilon(\varepsilon, k; \theta)$$

# Prediktív modellek paramétereinek ML becslése 2.

A ML becslés:

$$\hat{\theta}_{ML}(y^N) = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} -\ell(\varepsilon(i,\theta), i; \theta)$$

A LKN becslés:

$$\hat{\theta}_{ML}(y^N) = \arg \min_{\theta \in D_M} V_N(\theta, y^N) \quad , \quad V_N(\theta, D^N) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{2} \varepsilon(k,\theta)^2$$

Látható, hogy a ML becslés a LKN becslés abban a speciális esetben, ha a veszteségfüggvényt a

$$V_N(\theta, y^N) = \frac{1}{N} \sum_{i=1}^{N} - \ln f_{\varepsilon}(\varepsilon(i,\theta), i; \theta)$$

alakúra választjuk.

# Példa: ARX modell paramétereinek ML becslése 1.

$$\hat{y}_t = ay_{k-1} + bu_{k-1}$$

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\varepsilon(k,\theta) = y_k - ay_{k-1} - bu_{k-1} := e(k) \quad \text{függetlenek}$$

$$f_e(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad \text{normális eloszlásúak}$$

Ekkor a mintaelemek együttes eloszlásfüggvénye:

$$\overline{f}_y(\varepsilon) = \prod_{k=1}^{N} f_e(y_k - ay_{k-1} + bu_{k-1}) = \prod_{k=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left( \frac{-(y_k - ay_{k-1} - bu_{k-1})^2}{2\sigma^2} \right)$$

# Példa: ARX modell paramétereinek ML becslése 2.

A log-likelihood függvény:

$$\ln \overline{f}_y(\theta; y^N) = \sum_{k=1}^{N} - \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \ln\left(\exp\left(-\frac{(y_k - ay_{k-1} - bu_{k-1})^2}{2\sigma^2}\right)\right)$$

amely a paraméterek kvadratikus függvénye. Ennek egyértelmű minimuma:

$$\frac{d}{d\theta} \ln \overline{f}_y(\theta, y^N) = \begin{bmatrix} \sum_{k=1}^{N} 2y_{k-1}\frac{y_k - ay_{k-1} - bu_{k-1}}{2\sigma^2} \\ \sum_{k=1}^{N} 2u_{k-1}\frac{y_k - ay_{k-1} - bu_{k-1}}{2\sigma^2} \end{bmatrix} = 0$$

Az $a$ és $b$ paraméterek ML becslését az alábbi lineáris egyenletrendszer megoldása adja:

$$\sum_{k=1}^{N} 2y_{k-1}\frac{y_k - ay_{k-1} - bu_{k-1}}{2\sigma^2} = 0$$

$$\sum_{k=1}^{N} 2u_{k-1}\frac{y_k - ay_{k-1} - bu_{k-1}}{2\sigma^2} = 0$$

# *A BECSÜLT PARAMÉTEREK KOVARIANCIA MÁTRIXA: A CRAMÉR-RAO EGYENLŐTLENSÉG ÉS A FISHER INFORMÁCIÓS MÁTRIX*

# Cramér-Rao egyenlőtlenség, Fisher információs mátrix

A becsült $\hat{\theta}$ paraméterek "jósága" (torzítatlan becslés!)

$$P = E\left[\hat{\theta}(y^N) - \theta_0\right]\left[\hat{\theta}(y^N) - \theta_0\right]^T$$

kovariancia mátrix (valamilyen normában mért) nagysága

- Cramér-Rao egyenlőtlenség a kovariancia mátrixra
- Fisher információs mátrix: alsó korlát értéke

# Cramér-Rao egyenlőtlenség

**1. Tétel (Cramér-Rao).** *Legyen a $\hat{\theta}(y^N)$ a $\theta$ paraméter egy torzítatlan becslése úgy, hogy $E[\hat{\theta}(y^N)] = \theta_0$ a paraméter valódi értéke, és az $y^N$ minta elemeinek együttes valószínűségi sűrűség függvénye $f_y(\theta_0; y^N)$. Ekkor a $\hat{\theta}(y^N)$ becslés kovariancia mátrixára az alábbi egyenlőtlenség érvényes:*

$$E\left[\hat{\theta}(y^N) - \theta_0\right]\left[\hat{\theta}(y^N) - \theta_0\right]^T \geq M^{-1}$$

*ahol $M$ az úgynevezett **Fisher információs mátrix**:*

$$M = E\left[\frac{d}{d\theta}\ln f_y(\theta; y^N)\right]\left[\frac{d}{d\theta}\ln f_y(\theta; y^N)\right]^T\Bigg|_{\theta=\theta_0} =$$

$$= -E\left[\frac{d^2}{d\theta^2}\ln f_y(\theta; y^N)\right]\Bigg|_{\theta=\theta_0}$$

# ARX modell példa: Fisher információs mátrix 1

$$\hat{y}_t = ay_{k-1} + bu_{k-1}$$

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\varepsilon(k,\theta) = y_k - ay_{k-1} - bu_{k-1} := e(k) \text{ függetlenek}$$

$$f_e(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \text{ normális eloszlásúak}$$

példában szereplő egyszerű ARX rendszert a $\theta_0 = [a_0 \;\; b_0]^T$ valódi paramétervektorral. A becslés Fischer információs mátrixa:

$$M = E \begin{bmatrix} m_{aa} & m_{ab} \\ m_{ab} & m_{bb} \end{bmatrix}$$

# ARX modell példa: Fisher információs mátrix 2

A Fisher információs mátrix elemei

$$m_{aa} = (\sum_{k=1}^{N} 2y_{k-1} \frac{y_k - a_0 y_{k-1} - b_0 u_{k-1}}{2\sigma^2})^2$$

$$m_{ab} = (\sum_{k=1}^{N} 2y_{k-1} \frac{y_k - a_0 y_{k-1} - b_0 u_{k-1}}{2\sigma^2})(\sum_{k=1}^{N} 2u_{k-1} \frac{y_k - a_0 y_{k-1} - b_0 u_{k-1}}{2\sigma^2})$$

$$m_{bb} = (\sum_{k=1}^{N} 2u_{k-1} \frac{y_k - a_0 y_{k-1} - b_0 u_{k-1}}{2\sigma^2})^2$$

Az $m_{aa}, m_{ab}, m_{bb}$ elemekben az $\{y(k)\}_{k=1}^{N}$ és az $\{u(k)\}_{k=1}^{N}$ sztochasztikus folyamatok negyedrendű auto- és keresztkorrelációs együtthatói állnak.

# A ML becslés aszimptotikus tulajdonságai 1

**2. Tétel (Wald-Cramér).** *Tegyük fel, hogy az* $\{y(i)\}_{i=1}^{N}$ *függetlenek és azonos eloszlásúak úgy, hogy együttes sűrűségfüggvényük*

$$f_y(\theta; x_1, \ldots, x_N) = \prod_{i=1}^{N} f_{y(i)}(\theta; x_i)$$

*és az* $y^N$ *mintaelemek együtes sűrűségfüggvénye az* $f_y(\theta_0; x^N)$ *valamely* $\theta_0$ *valódi paraméter vektorral. Ekkor a* $\hat{\theta}_{ML}(y^N)$ *ML becslés 1 valószínűséggel* $\theta_0$*-hoz tart, ha* $N$ *tart a végtelenhez, és a*

$$\sqrt{N}[\hat{\theta}_{ML}(y^N) - \theta_0]$$

*eloszlása egy normális, 0 várható értékű és* $M^{-1}$ *kovariancia mártixú eloszláshoz tart.*

# A ML becslés aszimptotikus tulajdonságai 2

Ugyanilyen tétel áll fenn abban az esetben is, ha dinamikus rendszerek paramétereinek becslésére alkalmazzuk a maximum likelihood becslést.
Így tehát *a maximum likelihood becslés - elméleti szempontból legalábbis - a lehetséges legjobb becslés.*

# *DINAMIKUS RENDSZEREK PARAMÉTERBECSLÉSÉNEK KOVARIANCIA MÁTRIXA*

# A Fisher információs mátrix kiszámítása

Adott

- a teljes valószínűségi modell

$$\mathbb{M}(\theta): \qquad \hat{y}(k \mid \theta) = g(k, D^{k-1}; \theta)$$

$$\varepsilon(k,\theta) = y(k) - \hat{y}(k \mid \theta) \ \text{ függetlenek,}$$

$$\text{adott } \ f_\varepsilon(x;k;\theta) \ \text{ sűrűségfüggvénnyel}$$

- a szorzat alakú likelihood függvény

$$\overline{f}_y(\theta; y^N) = \prod_{i=1}^{N} f_\varepsilon(y(i) - g(i, D^{i-1}; \theta), i; \theta) = \prod_{i=1}^{N} f_\varepsilon(\varepsilon(i,\theta), i; \theta)$$

# 1. a log-likelihood függvény egyszerűbb alakja

$$\ell_0(\varepsilon) = \ln f_\varepsilon(\varepsilon)$$

$$\ln \overline{f}_y(\theta, y^N) = \sum_{i=1}^{N} \ln f_\varepsilon(\varepsilon(i,\theta)) = \sum_{i=1}^{N} -\ell_0(\varepsilon(i,\theta))$$

$$\frac{d}{d\theta} \ln \overline{f}_y(\theta, y^N) = \sum_{i=1}^{N} -\ell_0'(\varepsilon(i,\theta)) \cdot \underbrace{\frac{d}{d\theta}\varepsilon(i,\theta) \cdot (-1)}_{\Psi(i,\theta)}$$

## 2. behelyettesítés a Fisher információs mátrixba

$$E\left[\frac{d}{d\theta}\ln \overline{f}_y(\theta;y^N)\right]\left[\frac{d}{d\theta}\ln \overline{f}_y(\theta;y^N)\right]^T$$

Igy a várható érték:

$$\begin{aligned} M_N &= E\sum_{i=1}^{N}\sum_{j=1}^{N}\ell_0'(\varepsilon(i,\theta))\ell_0'(\varepsilon(j,\theta))\Psi(i,\theta_0)\Psi^T(j,\theta_0) \\ &= \sum_{i=1}^{N}E[\ell_0'(\varepsilon(i,\theta))]^2 E[\Psi(i,\theta_0)\Psi^T(i,\theta_0)] \quad , \quad \Psi(i,\theta_0) = -\frac{d}{d\theta}\varepsilon(i,\theta)\Big|_{\theta=\theta} \end{aligned}$$

figyelembe véve, hogy $\varepsilon(i,\theta)$ független $\varepsilon(j,\theta)$-től $i \neq j$-re, mivel a predikciós hiba fehérzaj sorozatot alkot.

# 3. log-likelihood függvény $\ell_0'(x)$ deriváltja

$$\ell_0'(x) = [\ln f_\varepsilon(x)]' = \frac{f_\varepsilon'(x)}{f_\varepsilon(x)}$$

Ebből

$$E[\ell_0'(\varepsilon(i,\theta))]^2 = \frac{1}{\kappa_0}$$

ahol $\kappa_0$ egy állandó.
Ha speciálisan a $\varepsilon(i,\theta)$ hibák normális eloszlásúak $\lambda_0$ szórásnégyzettel, akkor $\kappa_0 = \lambda_0$.

# 4. a Fisher információs mátrix egyszerű alakja

$$M_N = \frac{1}{\kappa_0} \sum_{i=1}^{N} E[\Psi(i, \theta_0) \Psi^T(i, \theta_0)]$$

A Cramér-Rao egyenlőtlenség segítségével kapjuk végül a dinamikus modellek paraméterbecslésének jóságára vonatkozó alapvető eredményt.

# A becsült paraméterek kovariancia mátrixa

*Dinamikus rendszerek $\theta$ paramétereinek becslésére használt tetszőleges torzítatlan $\hat{\theta}_N$ becslés ($E\hat{\theta}_N = \theta_0$, ahol $\theta_0$ a valódi paraméter) kovariancia mátrixára az alábbi alsó korlát adható:*

$$Cov\ \hat{\theta}_N \geq \kappa_0 \left[ \sum_{i=1}^{N} E\left\{ \Psi(i,\theta_0)\Psi^T(i,\theta_0) \right\} \right]^{-1}$$

$$\Psi(i,\theta_0) = -\frac{d}{d\theta}\varepsilon(i,\theta)\bigg|_{\theta=\theta_0}$$

Normális eloszlású hibák esetén $\kappa_0 = \lambda_0$.

# *ARX RENDSZEREK PARAMÉTERBECSLÉSÉNEK KOVARIANCIA MÁTRIXA*

# A teljes valószinűségi modell

SISO esetben:

$$\mathbb{M}(\theta): \quad \hat{y}(k \mid \theta) = -a_1 y(k-1) \cdots - a_n y(k-n) + b_0 u(k) + \cdots + b_m u(k-m)$$

$\varepsilon(k, \theta) = y(k) - \hat{y}(k \mid \theta)$ függetlenek,

adott $f_\varepsilon(x, \theta)$ sűrűségfüggvénnyel

**A likelihood függvény és a maximum likelihood becslés**

$$\overline{f}_y(\theta; y^N) = \prod_{i=1}^{N} f_\varepsilon(\varepsilon(\theta, i), \theta) \;, \;\; \ell(\varepsilon, \theta, k) = -\ln f_\varepsilon(\varepsilon(\theta, k), \theta) = \ell(\varepsilon(\theta, k), \theta)$$

A paraméterek maximum likelihood becslése:

$$\hat{\theta}_{ML}(y^N) = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} -\ell(\varepsilon(\theta, i), \theta)$$

# Normális eloszlású predikciós hiba esete 1

A predikciós hibák függetlenek és azonos normális eloszlásúak (SISO eset!):

$$f_\varepsilon(x,\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

A **log-likelihood függvény**:

$$\ln \overline{f}_y(\theta; y^N) = \sum_{k=1}^{N} [-\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) +$$

$$+\ln\left(\exp\left(-\frac{(y(k)+a_1y(k-1)+\cdots+a_ny(k-n)-b_0u(k)-\cdots-b_mu(k-m))^2}{2\sigma^2}\right)\right)$$

# Normális eloszlású predikciós hiba esete 2

A **log-likelihood függvény** egyértelmű **minimuma** a paraméterek

$$\theta = [-a_1 \ \ -a_2 \ \ \ldots \ \ -a_n \ \ b_0 \ b_1 \ \ \ldots \ \ b_m]^T$$

szerinti derivált vektor nulla helye:

$$\frac{d}{d\theta} \ln \overline{f}_y(\theta, y^N) =$$

$$\left[ \begin{array}{c} \sum_{k=1}^{N} 2y(k-1) \dfrac{y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) - b_0 u(k) - \cdots + b_m u(k-m)}{2\sigma^2} \\ \cdots \\ \sum_{k=1}^{N} 2u(k) \dfrac{y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) - b_0 u(k) - \cdots - b_m u(k-m)}{2\sigma^2} \\ \cdots \end{array} \right.$$

# Normális eloszlású predikciós hiba esete 3

A LKN módszernél megismert regresszorral

$$\varphi(k) = [y(k-1)\ y(k-2)\ \dots\ y(k-n)\ \ u(k)\ u(k-1)\ \dots\ u(k-m)]^T$$

a **ML becslés**

$$\hat{\theta}_{ML}(y^N) = \hat{\theta}_{LS} = \left[\frac{1}{N}\sum_{k=1}^{N}\varphi(k)\varphi^T(k)\right]^{-1}\frac{1}{N}\sum_{k=1}^{N}\varphi(k)y(k)$$

Ekkor a **Fischer információs mátrix a bemenet és kimenet sztochasztikus folyamatok negyedrendű auto- és keresztkovarianciáit tartalmazza** (ld. egyszerű példa)