MTA Számítástechnikai és Automatizálási Kutató Intézete, Közlemények 10.

GAUSS FOLYAMATOK EGY STATISZTIKAI PROBLÉMÁJÁRÓL

Pham Ngoc Phuc

1. BEVEZETÉS

Vizsgáljuk az

(1) $x(j) = \xi(j) + \Theta \qquad j = 1, 2, \ldots, n$

folyamatot, ahol a $\xi(1), \ldots, \xi(n)$ valószínűségi változók (általában nem függetlenek) eleget tesznek az $E\xi(j) = 0$, $E\xi^2(j) < \infty$ feltételeknek.

Legyen a Θ paraméter $(\Theta \epsilon R^1)$ legjobb lineáris torzítatlan becslése az

$$\hat{l} = \sum_{j=1}^{n} c_j^0 x(j) \quad \left(\sum_{j=1}^{n} c_j^0 = 1 \right) \quad \text{statisztika.}$$

Felmerül a kérdés, hogy az $\hat{l}$ becslés megengedhetősége a Θ torzítatlan becslései osztályában jellemző-e Gauss-folyamatokra?

A.M. Kagan, Ju.V. Linnik, C.R. Rao a [2] könyvben vizsgálják ezt a problémát abban az esetben, amikor $\xi(j)$ elsőrendű autoregressziós folyamat.

Az eredmény a következő (lásd [2], 341.o.):

Tétel. *Legyenek az* $x(1), \ldots, x(n)$, $n \geqslant 3$ *megfigyelések* (1) *alakúak, és* $\xi(j)$ *a következő elsőrendű autoregressziós folyamat*:

(2) $$\xi(1) = \epsilon(1)$$
$$\xi(j) = \lambda\xi(j-1) + \epsilon(j) \qquad j = 2, \ldots, n,$$

ahol az $\epsilon(1), \ldots, \epsilon(n)$ *valószinűségi változók függetlenek,* $F_j(x)$ *eloszlással (az* $F_j(x)$ *eloszlások nem szükségképpen azonosak),* $E\epsilon(j) = 0$, $E\epsilon^2(j) = \sigma_j^2$, $0 < \sigma_j^2 < \infty$.

Ha $\lambda \neq 1$, *akkor az* $\hat{l} = \sum_{j=1}^{n} c_j^0 x(j)$ *becslés megengedhetősége a* Θ *torzitatlan becslései osztályában jellemző tulajdonsága az* $\epsilon(1), \ldots, \epsilon(n)$ *Gauss-féle valószinűségi változóknak, egyben az* $x(j)$ *Gauss-folyamatnak.*

A következőkben a $\xi(j)$ p-edrendű autoregressziós folyamat esetén vizsgáljuk a problémát. Megmutatjuk, hogy bizonyos egyszerű feltételek mellett az elsőrendű autoregressziós sémára vonatkozó eredmény általánosítható.

A probléma felvetése Arató Mátyástól ered. Munkám során segítséget kaptam a Számítástechnikai és Automatizálási Kutató Intézet Valószínűségszámítási és statisztikai osztálya munkatársaitól. Ezúton fejezem ki köszönetemet nekik.

2. Legyen $\xi(t)$ stacionárius folyamat, és elégítse ki a

(3) $$\xi(t) + a_1 \xi(t-1) + \ldots + a_p \xi(t-p) = \epsilon(t)$$

sztochasztikus differenciálegyenletet, ahol $\epsilon(t)$ $(E\epsilon(t) = 0,\ E\epsilon^2(t) = \sigma_\epsilon^2)$ azonos eloszlású független sorozat, melyre $\epsilon(t)$ független $\xi(t-1), \xi(t-2), \ldots$, -től is.

Feltesszük, hogy $x(t)$ Gauss-folyamat. $n \geqslant 2p+1$ esetén $x(1), \ldots, x(n)$ együttes sűrűségfüggvényét a következőképpen határozhatjuk meg.

Figyelembe véve, hogy a $\xi(1), \ldots, \xi(p)$ változók függetlenek az $\epsilon(p+1), \ldots, \epsilon(n)$ változóktól, kapjuk, hogy

(4) $$p_{\xi(1),\ldots,\xi(p),\epsilon(p+1),\ldots,\epsilon(n)}(u_1, \ldots, u_p, z_{p+1}, \ldots, z_n) = (2\pi)^{-n/2} |R_p|^{-1/2} \sigma_\epsilon^{-(n-p)} \cdot$$
$$\cdot \exp\left[-\frac{1}{2}\left(\underline{U}_p R_p^{-1} \underline{U}_p^* + \frac{1}{\sigma_\epsilon^2} \sum_{p+1}^{n} z_i^2\right)\right]$$

ahol R_p jelöli a $\xi(1), \ldots, \xi(p)$ változók kovarianciamátrixát, R_p^{-1} ennek inverze,

$$\underline{U}_p = (u_1, \ldots, u_p), \quad \underline{U}_p^* = \begin{pmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_p \end{pmatrix}$$

Felhasználva a

$$\xi(1) = \xi(1)$$
$$\ldots\ldots\ldots\ldots$$
(5) $$\xi(p) = \xi(p)$$
$$\xi(p+1) + a_1 \xi(p) + \ldots + a_p \xi(1) = \epsilon(p+1)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\xi(n) + a_1 \xi(n-1) + \ldots + a_p \xi(n-p) = \epsilon(n)$$

leképezést, melynek determinánsa 1, (4)-ből belátható, hogy

(6) $$p_{\xi(1),\ldots,\xi(n)}(u_1, \ldots, u_n) = (2\pi)^{-n/2} |R_p|^{-1/2} \sigma_\epsilon^{-(n-p)} \exp\left\{-\frac{1}{2}\left[\underline{U}_p R_p^{-1} \underline{U}_p^* + \right.\right.$$
$$\left.\left. + \frac{1}{\sigma_\epsilon^2} \sum_{i=p+1}^{n} (u_i + a_1 u_{i-1} + \ldots + a_p u_{i-p})^2\right]\right\}$$

Innen $x(1), \ldots, x(n)$ együttes sűrűségfüggvénye a következő:

(7) $$p_{x(1),\ldots,x(n)}(x_1, \ldots, x_n; \Theta) = (2\pi)^{-n/2} |R_p|^{-1/2} \sigma_\epsilon^{-(n-p)} \exp -\frac{1}{2}\left\{(\underline{X}_p - \underline{\Theta}) R_p^{-1} (\underline{X}_p - \underline{\Theta})^* + \right.$$
$$\left. + \frac{1}{\sigma_\epsilon^2} \sum_{p+1}^{n} \left[(x_i - \Theta) + a_1 (x_{i-1} - \Theta) + \ldots + a_p (x_{i-p} - \Theta)\right]^2\right\}$$

ahol

$$\underline{X}_p - \underline{\Theta} = (x_1 - \Theta, \ldots, x_p - \Theta), \qquad (\underline{X}_p - \underline{\Theta})^* = \begin{pmatrix} x_1 - \Theta \\ \cdot \\ \cdot \\ \cdot \\ x_p - \Theta \end{pmatrix}$$

Ismeretes (lásd [1]), hogy $(a_0 = 1)$

$$R_p^{-1} = \frac{1}{\sigma_\epsilon^2} \begin{pmatrix} a_0^2 & a_0a_1 & a_0a_2 \ldots a_0a_{p-1} \\ a_0a_1 & a_0^2 + a_1^2 & a_0a_1 + a_1a_2 \ldots a_0a_{p-2} + a_1a_{p-1} \\ a_0a_2 & a_0a_1 + a_1a_2 & a_0^2 + a_1^2 + a_2^2 \ldots \\ \cdots & \cdots & \cdots \\ a_0a_{p-1} & a_0a_{p-2} + a_1a_{p-1} & \ldots a_0^2 + a_1^2 + \ldots + a_{p-1}^2 \end{pmatrix} \tag{8}$$

Innen

$$(\underline{X}_p - \underline{\Theta})R_p^{-1}(\underline{X}_p - \Theta)^* = \frac{1}{\sigma_\epsilon^2}\Big[a_0^2(x_1 - \Theta)^2 + a_0a_1(x_1 - \Theta)(x_2 - \Theta) + \ldots + a_0a_{p-1} \cdot$$

$$\cdot (x_1 - \Theta)(x_p - \Theta) + a_0a_1(x_2 - \Theta)(x_1 - \Theta) + (a_0^2 + a_1^2)(x_2 - \Theta)^2 +$$

$$+ \ldots + (a_0a_{p-2} + a_1a_{p-1})(x_2 - \Theta)(x_p - \Theta) + \ldots +$$

$$+ (a_0^2 + a_1^2 + \ldots + a_{p-1}^2)(x_p - \Theta)^2\Big]$$

$$= -\frac{2\Theta}{\sigma_\epsilon^2}\Big[x_1(a_0^2 + a_0a_1 + \ldots + a_0a_{p-1}) + x_2(a_0a_1 + a_0^2 + a_1^2 +$$

$$+ \ldots + a_0a_{p-2} + a_1a_{p-1}) + \ldots + x_p(a_0a_{p-1} + a_0a_{p-2} +$$

$$+ a_1a_{p-1} + \ldots + a_0^2 + a_1^2 + \ldots + a_{p-1}^2)\Big] + f_1(x_1, \ldots, x_n) +$$

$$+ g_1(\Theta)$$

ahol $f_1(x_1, \ldots, x_n)$ és $g_1(\Theta)$ értelmezése leolvasható.

Ily módon,

$$(\underline{X}_p - \underline{\Theta})R_p^{-1}(\underline{X}_p - \underline{\Theta})^* = -\frac{2\Theta}{\sigma_\epsilon^2}(\alpha_1x_1 + \alpha_2x_2 + \ldots + \alpha_px_p) + f_1(x_1, \ldots, x_n) + g_1(\Theta) \tag{9}$$

ahol

$$\alpha_1 = a_0^2 + a_0 a_1 + \ldots + a_0 a_{p-1}$$

$$\alpha_2 = a_0 a_1 + (a_0^2 + a_1^2) + \ldots + (a_0 a_{p-2} + a_1 a_{p-1}) \tag{10}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\alpha_p = a_0 a_{p-1} + (a_0 a_{p-2} + a_1 a_{p-1}) + \ldots + (a_0^2 + a_1^2 + \ldots + a_{p-1}^2)$$

Másrészt

$$\sum_{i=p+1}^{n}\left[(x_i - \Theta) + a_1(x_{i-1} - \Theta) + \ldots + a_n(x_{i-p} - \Theta)\right]^2 = \sum_{i=p+1}^{n}\left[(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p}) - \Theta \cdot (1 + a_1 + \ldots + a_p)\right]^2 \tag{11}$$

$$= \sum_{i=p+1}^{n}(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p})^2 - 2\Theta(1 + a_1 + \ldots + a_p)\sum_{i=p+1}^{n}(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p}) + \Theta^2 (n-p)(1 + a_1 + \ldots + a_p)^2$$

$$= -2\Theta(1 + a_1 + \ldots + a_p)\sum_{i=p+1}^{n}(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p}) + f_2(x_1, \ldots, x_n) + g_2(\Theta)$$

(9) és (11) behelyettesítésével $x(1), \ldots, x(n)$ (7) alatti együttes sűrűségfüggvényét a következő alakban írhatjuk:

$$p_{x(1),\ldots,x(n)}(x_1, \ldots, x_n; \Theta) = F(x_1, \ldots, x_n)G(\Theta)\exp\frac{\Theta}{\sigma_\epsilon^2}\left\{\alpha_1 x_1 + \ldots + \alpha_p x_p + (1 + a_1 + \ldots + a_p)\sum_{i=p+1}^{n}(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p})\right\} \tag{12}$$

Tekintsük az

$$l = \alpha_1 x_1 + \ldots + \alpha_p x_p + (1 + a_1 + \ldots + a_p)\sum_{i=p+1}^{n}(x_i + a_1 x_{i-1} + \ldots + a_p x_{i-p}) \tag{13}$$

$$= \alpha_1 x_1 + \ldots + \alpha_p x_p + (1 + a_1 + \ldots + a_p)\left[\sum_{j=1}^{p}(a_p + \ldots + a_{p-j+1})x_j + \sum_{j=p+1}^{n-p}(1 + a_1 + \ldots + a_p)x_j + \sum_{j=n-p+1}^{n}(1 + a_1 + \ldots + a_{n-j})x_j\right]$$

lineáris statisztikát.

Látható, hogy

$$l = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$$

alakú, ahol

$$\begin{aligned} &\lambda_j = \alpha_j + (1 + a_1 + \ldots + a_p)(a_{p-j+1} + \ldots + a_p) && \text{ha } j = 1, \ldots, p \\ (14) \quad &\lambda_j = (1 + a_1 + \ldots + a_p)^2 && \text{ha } p + 1 \leqslant j \leqslant n - p \\ &\lambda_j = (1 + a_1 + \ldots + a_p)(1 + a_1 + \ldots + a_{n-j}) && \text{ha } n - p + 1 \leqslant j \leqslant n\,. \end{aligned}$$

(12)-ből nyilvánvaló, hogy az l statisztika elégséges az $x(1), \ldots, x(n)$ együttes sűrűségfüggvényeinek összességére nézve. Továbbá könnyű megmutatni, hogy az l lineáris statisztika egyben teljes is.

Vizsgáljuk most az $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ statisztikát.

Mivel $\hat{l}$ a legjobb lineáris torzítatlan becslése Θ-nak, és a vizsgálandó esetben létezik az l elégséges statisztika, akkor $\hat{l}$ szükségképpen csak l-től függ, másszóval $\hat{l}$ függvénye l-nek.

Valójában, ellenkező esetben $\hat{l}$ nem volna függvénye l-nek. Vizsgáljuk a $g(l) = E(\hat{l} \mid l)$ függvényt. Minthogy $x_1, \ldots, x_n$ Gauss-eloszlásúak, ennélfogva $\hat{l}$ és l szintén Gauss-eloszlású. Igy

$$g(l) = E(\hat{l} \mid l) = \alpha l + \beta$$

alakú, azaz $g(l)$ lineáris függvénye l-nek, ebből következik, hogy $E(\hat{l} \mid l)$ lineáris függvénye $x_1, \ldots, x_n$-nek.

A Blackwell-Kolmogorov-Rao-tétel szerint $E(\hat{l} \mid l)$ torzítatlan becslése Θ-nak (így $E(\hat{l} \mid l)$ lineáris torzítatlan becslése Θ-nak), és

$$(15) \qquad E_\Theta (\hat{l} - \Theta)^2 \geqslant E_\Theta \left[E(\hat{l} \mid l) - \Theta\right]^2 \qquad \text{minden } \Theta \epsilon R^1\text{-ra.}$$

A feltevés szerint $\hat{l}$ legjobb lineáris torzítatlan becslése Θ-nak, ezért a (15) relációban szükségképpen teljesül az egyenlőség:

$$(16) \qquad E_\Theta (\hat{l} - \Theta)^2 = E_\Theta \left[E(\hat{l} \mid l) - \Theta\right]^2 \qquad \text{minden } \Theta \epsilon R^1\text{-ra.}$$

Ha visszaemlékszünk a Blackwell-Kolmogorov-Rao-tétel bizonyítására (lásd pl. [3]), (16) akkor és csak akkor teljesül, ha

$$\hat{l} = E(\hat{l} \mid l) = g(l) \qquad \text{m.m. } P_\Theta,\ \Theta \epsilon R^1,$$

ami ellentmond a feltevésünknek, hogy $\hat{l}$ nem függ l-től. Ezzel igazoltuk állításunkat.

Az l statisztika teljessége miatt $\hat{l}$ Θ-nak egyetlen olyan torzítatlan becslése, amely függvénye l-nek, így $\hat{l}$ legjobb torzítatlan becslés Θ-ra.

A fenti eredményeket összefoglalva a következő állítást kapjuk:

1. Tétel. *Ha $x(j)$ Gauss-folyamat és $x(j) = \xi(j) + \Theta$ alakú, ahol $\xi(j)$ stacionárius folyamat és kielégiti a (3) sztochasztikus differenciálegyenletet, akkor az*

$$\hat{l} = \sum_{j=1}^{n} c_j^0 x_j \qquad (n \geqslant 2p + 1)$$

legjobb lineáris torzitatlan becslés megengedhető – sőt optimális – a Θ torzitatlan becsléseinek osztályában.

3. Tegyük most fel, hogy a Θ paraméter legjobb lineáris torzítatlan becslése $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ megengedhető a Θ torzítatlan becslései osztályában.

Legyen $\underline{y} = (x_2 - x_1, \ldots, x_n - x_1)$

Bevezetjük a

(17) $\quad \hat{\Theta} = \hat{l} - E_0(\hat{l} \mid \underline{y})$

Pitman-féle becslést.

A $\xi(t)$ elsőrendű autoregressziós folyamat esetén teljesülnek az alábbi 1. és 2. lemmák (lásd [2]), amelyek érvényben maradnak a $\xi(t)$ p-edrendű autoregressziós folyamatra is.

1. Lemma. *A Θ paraméter (17) becslése torzitatlan és*

$$E_\Theta(\hat{\Theta} - \Theta)^2 \leqslant E_\Theta(\hat{l} - \Theta)^2,$$

az egyenlőség minden $\Theta \epsilon R^1$-ra akkor és csak akkor áll fenn, ha

$$E_0(\hat{l} \mid \underline{y}) = 0.$$

Az 1. lemmából következik, hogy az $\hat{l}$ becslés akkor és csak akkor megengedhető a Θ torzítatlan becslései osztályában ha

$$E_0(\hat{l} \mid \underline{y}) = 0.$$

Legyenek

$$f(t) = E\exp it\epsilon_j, \qquad g(t) = \frac{f'(t)}{f(t)}$$

$$\Phi(t_1, \ldots, t_n) = E\exp i\sum_{j=1}^{n} t_j\xi_j = E_0 \exp i\sum_{j=1}^{n} t_j x_j$$

$$\varphi(t_1, \ldots, t_n) = \log \Phi(t_1, \ldots, t_n)$$

ahol $t_1, \ldots, t_n$ valós számok, a $g(t)$ függvény a $t = 0$ pont valamilyen környezetében, $\varphi(t_1, \ldots, t_n)$ pedig a $(0, \ldots, 0)$ pont egy környezetében van értelmezve.

2. Lemma. *Tegyük fel, hogy* $E_0(\hat{l} \mid \underline{y}) = 0$ *akkor a* $(0, \ldots, 0)$ *pont valamilyen környezetében*

$$\sum_{j=1}^{n} c_j^0 \frac{\partial\varphi(\tau_1, \ldots, \tau_n)}{\partial\tau_j} = 0, \quad \textit{ha} \quad \sum_{j=1}^{n} \tau_j = 0. \tag{18}$$

Legyen $\xi(k)$ a következő p-edrendű autoregressziós folyamat:

$$\xi(1) = \epsilon(1)$$

$$\xi(2) + a_1\xi(1) = \epsilon(2)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\xi(p) + a_1\xi(p-1) + \ldots + a_{p-1}\xi(1) = \epsilon(p) \tag{19}$$

$$\xi(K) + a_1\xi(K-1) + \ldots + a_p\xi(K-p) = \epsilon(K), \quad \text{ha} \quad K \geqslant p+1$$

ahol $\epsilon(1)$, $\epsilon(2), \ldots$ $(E\epsilon(k) = 0$, $E\epsilon^2(k) = \sigma_\epsilon^2)$ független, azonos eloszlású sorozat. Irjuk fel a (19) egyenleteket $k = n$, $n-1, \ldots, 2, 1$ estén, azután szorozzuk meg azokat rendre a $b_0 (b_0 = 1)$, $b_1, \ldots, b_{n-2}, b_{n-1}$ számokkal, majd adjuk össze őket. A kapott összefüggésben $\xi(n-1), \ldots, \xi(2), \xi(1)$ akkor és csak akkor nem szerepel, ha

$$a_0b_1 + a_1b_0 = 0$$

$$a_0b_2 + a_1b_1 + a_2b_0 = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_0b_p + a_1b_{p-1} + \ldots + a_pb_0 = 0 \tag{20}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_0b_{n-1} + a_1b_{n-2} + \ldots + a_pb_{n-(p+1)} + a_{p+1}b_{n-(p+2)} + \ldots + a_{n-1}b_0 = 0$$

ahol $a_0 = b_0 = 1$, és $a_k = 0$ ha $k \geqslant p + 1$.

(20) teljesülése esetén kapjuk, hogy

$$\xi(n) = \epsilon(n) + b_1 \epsilon(n-1) + \ldots + b_{n-1}\epsilon(1) = \sum_{K=0}^{n-1} b_k \epsilon(n-k). \tag{21}$$

Ily módon a $\xi(n)$ sorozat előállítható a (21) alakban, ahol a b_k együtthatók (20)-ból határozhatók meg.

3. Lemma. *A (19) $\xi(t)$ p-edrendű autoregressziós folyamat esetén*

$$\Phi(t_1, \ldots, t_n) = f(t_1 + b_1 t_2 + \ldots + b_{n-1} t_n) \ldots f(t_{n-1} + b_1 t_n) \cdot f(t_n). \tag{22}$$

Bizonyítás. Valóban (21) szerint $\xi(k) = b_{k-1}\epsilon(1) + b_{k-2}\epsilon(2) + \ldots + \epsilon(k)$, $k = 1,2,\ldots,n$,

ekkor,

$$\begin{aligned}
\Phi(t_1, \ldots, t_n) &= E \exp i \sum_{j=1}^{n} t_j \xi(j) \\
&= E \exp i \big[t_1 \epsilon_1 + t_2(b_1 \epsilon_1 + \epsilon_2) + \ldots + t_n(b_{n-1}\epsilon_1 + \ldots + b_1 \epsilon_{n-1} + \\
&\qquad + \epsilon_n)\big] \\
&= E \exp i \big[(t_1 + b_1 t_2 + \ldots + b_{n-1} t_n)\epsilon_1 + \ldots + (t_{n-1} + b_1 t_n)\epsilon_{n-1} + \\
&\qquad + t_n \epsilon_n\big] \\
&= E e^{i(t_1 + b_1 t_2 + \ldots + b_{n-1} t_n)\epsilon_1} \cdot E e^{i(t_{n-1} + b_1 t_n)\epsilon_{n-1}} \cdot E e^{i t_n \epsilon_n} \\
&= f(t_1 + b_1 t_2 + \ldots + b_{n-1} t_n) \ldots f(t_{n-1} + b_1 t_n) \cdot f(t_n).
\end{aligned}$$

A 3. lemma alapján

$$\begin{aligned}
\varphi(t_1, \ldots, t_n) = \log \Phi(t_1, \ldots, t_n) &= \log f(t_1 + b_1 t_2 + \ldots + b_{n-1} t_n) + \ldots + \\
&+ \log f(t_{n-1} + b_1 t_n) + \log f(t_n)
\end{aligned}$$

Innen könnyen belátható, hogy a p-edrendű autoregressziós folyamat esetén a (18) összefüggés felírható a következő alakban:

$$c_n^0\left\{\frac{f'(t_n)}{f(t_n)} + b_1\frac{f'(t_{n-1} + b_1 t_n)}{f(t_{n-1} + b_1 t_n)} + \ldots + b_{n-1}\frac{f'(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)}{f(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)}\right\} +$$

$$+ c_{n-1}^0\left\{\frac{f'(t_{n-1} + b_1 t_n)}{f(t_{n-1} + b_1 t_n)} + b_1\frac{f'(t_{n-2} + b_1 t_{n-1} + b_2 t_n)}{f(t_{n-2} + b_1 t_{n-1} + b_2 t_n)} + \ldots +\right.$$

$$\left.+ b_{n-2}\frac{f'(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)}{f(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)}\right\} + \ldots + c_1^0\frac{f'(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)}{f(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n)} = 0$$

vagy,

$$(23)\qquad c_n^0 g(t_n) + (b_1 c_n^0 + c_{n-1}^0)g(t_{n-1} + b_1 t_n) + \ldots + (b_{n-1}c_n^0 + b_{n-2}c_{n-1}^0 + \ldots + c_1^0)\cdot$$

$$\cdot\, g(t_1 + b_1 t_2 + \ldots + b_{n-1}t_n) = 0,$$

$$\text{ha} \quad \sum_{j=1}^{n} t_j = 0$$

Legyenek

$$\gamma_n = c_n^0$$

$$\gamma_{n-1} = b_1 c_n^0 + c_{n-1}^0$$

$$(24)\qquad \gamma_{n-2} = b_2 c_n^0 + b_1 c_{n-1}^0 + c_{n-2}^0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\gamma_1 = b_{n-1}c_n^0 + b_{n-2}c_{n-1}^0 + \ldots + c_1^0$$

Vezessük be a

$$z_n = t_n$$

$$z_{n-1} = b_1 t_n + t_{n-1}$$

$$(25)\qquad z_{n-2} = b_2 t_n + b_1 t_{n-1} + t_{n-2}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$z_1 = b_{n-1}t_n + b_{n-2}t_{n-1} + \ldots + t_1$$

új változókat, akkor (23) a

$$(26)\qquad \sum_{j=1}^{n}\gamma_j g(z_j) = 0$$

alakra hozható.

Legyen $\beta_i = -b_i \quad i = 1,2,\ldots,n-1$, (25)-ből következik, hogy

$$
t_n = \nu_0 z_n
$$

$$
t_{n-1} = \nu_1 z_n + \nu_0 z_{n-1}
$$

(27) $\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

$$
t_{n-k} = \nu_k z_n + \nu_{k-1} z_{n-1} + \ldots + \nu_0 z_{n-k}
$$

$$
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
$$

$$
t_1 = \nu_{n-1} z_n + \nu_{n-2} z_{n-1} + \ldots + \nu_0 z_1
$$

ahol

$$
\begin{aligned}
\nu_0 &= 1 \\
\nu_1 &= \beta_1 \\
\nu_2 &= \beta_1^2 + \beta_2 \\
\nu_3 &= \beta_1^3 + 2\beta_1\beta_2 + \beta_3 \\
\nu_4 &= \beta_1^4 + 3\beta_1^2\beta_2 + 2\beta_1\beta_3 + \beta_2^2 + \beta_4
\end{aligned}
$$

és általában

(28) $\quad \nu_{k+1} = \beta_1 \nu_k + \beta_2 \nu_{k-1} + \ldots + \beta_{k+1} \nu_0 .$

Ily módon, a $\sum_{j=1}^{n} t_j = 0$ feltételt átírhatjuk $z_1, z_2, \ldots, z_n$ segítségével az

$$
(1 + \nu_1 + \ldots + \nu_{n-1}) z_n + (1 + \nu_1 + \ldots + \nu_{n-2}) z_{n-1} + \ldots + (1 + \nu_1) z_2 + \\
+ z_1 = 0
$$

alakba.

A fentieket a következőképpen foglalhatjuk össze:

Ha az $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ legjobb lineáris torzítatlan becslés megengedhető a Θ paraméter torzítatlan becslései osztályában, akkor

(29) $\quad \sum_{j=1}^{n} \gamma_j g(z_j) = 0$

ahol $(1 + \nu_1 + \ldots + \nu_{n-1}) z_n + \ldots + (1 + \nu_1) z_2 + z_1 = 0.$

4. Lemma. *Tegyük fel, hogy* $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ *legjobb lineáris torzitatlan becslése* Θ*-nak, akkor*

a c_j^0 *együtthatók a következők:*

$$
(30) \qquad
\begin{aligned}
c_j^0 &= c[\alpha_{p-j+1} + (1 + a_1 + \ldots + a_p)(a_{p-j+1} + \ldots + a_p)] && \text{ha} \quad 1 \leqslant j \leqslant p \\
c_j^0 &= c(1 + a_1 + \ldots + a_p)^2 && \text{ha} \quad p + 1 \leqslant j \leqslant n - p \\
c_j^0 &= c(1 + a_1 + \ldots + a_p)(1 + a_1 + \ldots + a_{n-j}) && \text{ha} \quad n - p + 1 \leqslant j \leqslant n,
\end{aligned}
$$

ahol a c *állandó a* $\sum_{j=1}^{n} c_j^0 = 1$ *reláció segitségével,* $\alpha_j (j = 1,2,\ldots,p)$ *pedig* (10)*-ből határoz-*

ható meg.

Bizonyítás. Mivel a c_j^0, $j = 1,\ldots,n$ konstansok az $x(j)$ folyamatnak csupán első és másodrendű momentumaitól függenek, a c_j^0 értékének kiszámításakor jogunk van $x(j)$ folyamatot Gauss-folyamatnak tekinteni.

Abban az esetben, amikor $x(j) = \xi(j) + \Theta$, ahol $\xi(j)$ (19) alakú p-edrendű autoregressziós folyamat, könnyű megmutatni, hogy az

$$l^* = \lambda_1^* x_1 + \lambda_2^* x_2 + \ldots + \lambda_n^* x_n$$

lineáris statisztika, ahol

$$
\begin{aligned}
\lambda_j^* &= \alpha_{p-j+1} + (1 + a_1 + \ldots + a_p)(a_{p-j+1} + \ldots + a_p) && \text{ha} \quad 1 \leqslant j \leqslant p \\
\lambda_j^* &= (1 + a_1 + \ldots + a_p)^2 && \text{ha} \quad p + 1 \leqslant j \leqslant n - p \\
\lambda_j^* &= (1 + a_1 + \ldots + a_p)(1 + a_1 + \ldots + a_{n-j}) && \text{ha} \quad n - p + 1 \leqslant j \leqslant n
\end{aligned}
$$

elégséges, és egyben teljes is a $p_{x(1),\ldots,x(n)}(\underline{X},\Theta)$ sűrűségfüggvényének összeségére nézve.

Másrészt feltevésünk szerint az $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ statisztika legjobb lineáris torzítatlan becslés

Θ-ra, tehát $\hat{l}$ szükségképpen függvénye l-nek, amiből adódik, hogy c_j^0 és λ_j^*, $j = 1,\ldots,n$ arányosak egymással, tehát fennállnak a (30) összefüggések.

Térjünk vissza most a (29) relációra, mely szerint

$$\sum_{j=1}^{n} \gamma_j g(z_j) = 0, \quad \text{ha} \quad (1 + \nu_1 + \ldots + \nu_{n-1}) z_n + \ldots + (1 + \nu_1) z_2 + z_1 = 0$$

ahol (24) és 4. lemma alapján $\gamma_1,\ldots,\gamma_n$ kifejezhető $a_1,\ldots,a_p$ segítségével.

Tegyük fel, hogy

(31) $$\frac{\gamma_1}{\gamma_2}(1+\nu_1) > 0.$$

Legyen $z_3 = \ldots = z_n = 0$, $z_2 = t$ szabad változó, és $z_1 = -(1+\nu_1)z_2$.

Mivel $g(z_3) = \ldots = g(z_n) = g(0) = \frac{f'(0)}{f(0)} = 0$, (29)-ből adódik $\gamma_2 g(t) = -\gamma_1 g(z_1)$.

t szerint differenciálva

(32) $$g'(t) = \frac{\gamma_1(1+\nu_1)}{\gamma_2} g'(z_1)$$ -hez

jutunk.

Vegyük figyelembe, hogy $g'(0) = -\sigma_\epsilon^2 < 0$, $g'(t)$ folytonos $t = 0$ környezetében és feltevésünk szerint
$\frac{\gamma_1(1+\nu_1)}{\gamma_2} > 0$, így (32) miatt

$$g'(t) = \text{const.} = -S^2 < 0$$

ha $|t|$ eléggé kicsiny, amiből következik, hogy $f(t) = e^{-\frac{S^2}{2}t^2}$ a $t = 0$ pont valamilyen környezetében. A karakterisztikus függvények tulajdonságaira vonatkozó tétel alapján (lásd pl. [4]) megállapíthatjuk, hogy $\epsilon(j)$ Gauss-eloszlású.

Ily módon igazoltuk a következő állítást:

2. Tétel. *Tegyük fel, hogy* $x(j) = \xi(j) + \Theta$, *ahol* $\xi(j)$ (19) *alakú p-edrendű autoregressziós folyamat és teljesül a*

$$\frac{\gamma_1}{\gamma_2}(1+\nu_1) > 0$$

feltétel. Ha az $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ *legjobb lineáris becslés megengedhető a* Θ *torzítatlan becsléseinek osztályában, akkor az* $x(j)$ *folyamat Gauss-folyamat.*

Irodalom

[1] Arató M., Folytonos állapotú Markov-folyamatok statisztikai vizsgálatáról. IV. A Magy. Tud. Akad. III Osztály Közleményei 15 (1965) 107-124.

[2] А.М.Каган, Ю.В.Линник, С.Р.Рао: Характеризационные задачи математической статистики. Москва 1972.

[3] C.R. Rao, Linear statistical inference and its applications, N.-Y., Wiley, 1965.

[4] Rényi A., Valószínűségszámítás, Budapest, 1968.

Beérkezett: 1973. március 14.

Summary

ON A STATISTICAL PROBLEM OF GAUSSIAN PROCESSES

In the paper the $X(j) = \xi(j) + \Theta$ process is investigated where $\xi(j)$ is a p-order autoregressive process.
It is shown that when $X(j)$ is gaussian, the best linear estimate $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$ $\left(\sum_{j=1}^{n} c_j^0 = 1 \right)$, $n \geq 2p + 1$, is an admissible estimate of the location parameter Θ in the class of all unbiased estimates of Θ.
The inversion of the statement is also proved.
For case $p = 1$ the result is in the book of Kagan, Linnik, Rao [2].

Резюме

ОБ ОДНОЙ СТАТИСТИЧЕСКОЙ ПРОБЛЕМЕ ГАУССОВСКИХ ПРОЦЕССОВ

В данной работе рассматривается процесс $X(j) = \xi(j) + \Theta$, где $\xi(j)$ - авторегрессионный процесс порядка p.

Показано, что если процесс $X(j)$ гауссовский, то наилучшая линейная оценка $\hat{l} = \sum_{j=1}^{n} c_j^0 x_j$, $\sum_{j=1}^{n} c_j^0 = 1$, $n \geq 2p + 1$ является допустимой оценкой параметра θ в классе несмещенных оценок.

В работе рассмотрено также обратное утверждение. Для случая p=1 этот результат рассмотрен в книге Кагана, Линника, Rao [2].