A COULOMB-SZÓRÓDÁS PARAMÉTERÉNEK BECSLÉSE FOTO-EMULZIÓBAN VÉGZETT MÉRÉSEK ALAPJÁN

JÁNOSSY LAJOS[1] — LEE ANNA — RÓZSA PÁL

Bevezetés

Gyors részecskék energiájának meghatározása a fotoemulzióban a részecskenyom szóródásának mérése alapján történik. Ugyanis, ha ϑ jelöli a részecske s hosszúságú úton (cellahossz) elszenvedett szóródásának, a többszörös Coulomb-szóródásnak síkbeli vetített szögét, akkor e ϑ szög négyzete $\langle\vartheta^2\rangle$ várható értékének és a részecske tömegének ismeretében energiája már meghatározható.

Ebben a dolgozatban a Coulomb-szóródás $\langle\vartheta^2\rangle$ paraméterének — pontosabban az ezzel arányos $\alpha^{(1)}$ jel-paraméternek — a becslésével foglalkozunk. A becslésnél a mérési hibának — a háttér zajának — a hatását is figyelembe vesszük. A probléma fizikai hátterével itt nem foglalkozunk, erre vonatkozóan néhány korábbi dolgozatra utalunk ([1]—[4]). E helyen a feladatot elsősorban matematikai szempontból kívánjuk tárgyalni.

A dolgozat első részében a Coulomb-szóródás $\alpha^{(1)}$ paraméterét és a háttér hatásának (a mérési hibának) $\alpha^{(2)}$ paraméterét becsüljük először a maximum likelihood módszer alapján, majd mért d_i értékeknek egy kvadratikus alakjával. Megmutatjuk, hogy ha az $\alpha^{(\nu)}$ paraméterek kvadratikus becslésétől torzítatlanságot és minimális szórást követelünk meg, akkor ez az *optimális kvadratikus* becslés megegyezik a maximum likelihood módszer alapján kapott becsléssel. Kvadratikus becslésnél a torzítatlanság és minimális szórás követelménye a kvadratikus alak mátrixának elemeire vonatkozóan feltételes szélsőértékfeladathoz vezet. Ennek megoldására — kis cellahossz esetén érvényes — egyszerű aszimptotikus kifejezéseket nyerünk. Az optimális kvadratikus becslés szórását jellemző $q_{\nu\mu}(x)$ görbéket — a becsült paraméterek hányadosának függvényében — az 1. ábra tünteti fel. (Az ábráknál a *b)* és a *c)* jelzés arra utal, hogy az *a)* jelű ábra egy-egy részét a gyakorlati használhatóság követelményeinek megfelelően kinagyítottuk.) Megmutatjuk azt is, hogy az $\alpha^{(1)}$ paraméter becslésének relatív szórása aszimptotikusan a mérési pontok számának nyolcadik gyökével csökken.

Ez az optimális kvadratikus becslés gyakorlati számításra nem alkalmas, mert rendkívül körülményes és hosszadalmas számításokat igényel. A második részben általános módszert adunk gyakorlati számítás szempontjából alkalmas kvadratikus becslésre. A becslés természetesen rosszabb, mint az előbbiekben nyert optimális, azonban az optimális becslés szórásának ismeretében esetenként mérlegelhető, hogy a választott becslés hatásfoka kielégítő-e. A módszer

[1] Központi Fizikai Kutató Intézet.

lényege abban áll, hogy a kvadratikus becslés mátrixát — eltérően az optimális becslés esetétől — alkalmasan választott *alapmátrixok* lineáris kombinációjaként írjuk fel. Íly módon nem a mátrix elemeire, hanem csupán a lineáris kombinációban szereplő együtthatókra nézve jutunk feltételes szélsőértékfeladatra.

Az első két részben foglaltak lényegileg megtalálhatók korábbi [4] dolgozatunkban. A [4] dolgozat harmadik részében a fent említett $\mathbf{G}^{(\lambda)} = [G_{ij}^{(\lambda)}]$ alapmátrixokat a következő *Toeplitz-típusú* mátrixoknak választottuk:

$$(*) \qquad G_{ij}^{(\lambda)} = \frac{1}{2(N-\lambda)}(\delta_{i,j+\lambda} + \delta_{i+\lambda,j}) \quad \lambda = 0, 1, \ldots, l;\ i, j = 1, 2, \ldots, N.$$

Jelen dolgozatunk harmadik részében először annak elméleti indokolását adjuk, hogy az alapmátrixokat miért célszerű Toeplitz-típusúaknak választani. Ezután M. E. Cosyns javaslatára az alapmátrixokat úgy választjuk meg, hogy azok a koordináták *átugrásos* második differenciáinak tiszta négyzetösszegével való becslésnek feleljenek meg. Az alapmátrixoknak ilyenképpen való megválasztása elméleti és gyakorlati szempontból egyaránt jobbnak bizonyul, mint a (*) mátrixoké. Elméleti szempontból azért, mert az ezekkel végzett becslés az optimális kvadratikus becslés hatásfokát jobban megközelíti (vö. [4] dolgozat 3., 2. ábrái és jelen dolgozat 2., 3. ábrái). Gyakorlati szempontból pedig azért, mert a becslés elvégzésénél vegyesszorzatok számítása feleslegessé válik. Az a körülmény, hogy a becsléshez a mérési pontok ordinátáiból megfelelő módon képzett második differenciáknak kizárólag a tiszta négyzetösszegét kell felhasználni, a mérések gépi úton való egyszerű és gyors kiértékelésének lehetőségét könnyíti meg.

Lényeges különbség mutatkozik azon két eset között, amikor a becslést két, illetve három alapmátrixszal végezzük. Két alapmátrix alkalmazása esetén három különböző becslési módszert mutatunk meg aszerint, hogy az alapbeosztáshoz tartozó második differenciák mellett az egy-, két-, illetve háromátugrásos második differenciák négyzetösszegét használjuk fel a becsléshez. Tekintettel arra, hogy ekkor a *feltételi* egyenletek egyértelműen határozzák meg a paraméterek becslését, nagyon jó hatásfokú becslést nem várhatunk. A 2. ábra segítségével azonban eldönthető, hogy melyik módszerrel nyert becslés tekinthető a paraméterek első közelítésének. A három alapmátrixszal végzett becslésnél $\overline{\alpha}^{(1)}$ és $\overline{\alpha}^{(2)}$ az $\overline{x} = \dfrac{\overline{\alpha}^{(2)}}{\overline{\alpha}^{(1)}}$ hányados függvényeként adódik. Ahhoz, hogy az $\alpha^{(1)}$ és $\alpha^{(2)}$ paraméterek kívánt becslését megkapjuk, e függvényekben $\overline{x}$ helyére egy harmadfokú egyenletnek azt a pozitív gyökét kell behelyettesíteni, amelynek első közelítéseként a két alapmátrixszal végzett becslések hányadosa tekinthető. — Az $\alpha^{(1)}$ becslés *relatív* szórását meghatározó $r^{(0j)}(z)$ függvénygörbéket a 3. ábra tünteti fel. Innen látható, hogy az optimális kvadratikus becslés relatív szórásgörbéjétől eltérően e görbéknek jóldefiniált minimumuk van. Ez a minimumhely az $\dfrac{\alpha^{(2)}}{\alpha^{(1)}}$ hányados minden értékéhez a mérések számának egy meghatározott N_0 értékét rendeli hozzá, úgy, hogy e becslés akkor lesz a legjobb, ha a cellahosszt az N_0 értéknek megfelelően választjuk.

Nyilvánvaló, hogy amennyiben az ilyen módon végzett becslés nem közelíti meg kielégítően az optimális kvadratikus becslést, akkor az alapmátrixok más megválasztásával, vagy az alapmátrixok számának növelésével a becslés jósága tovább fokozható.

I. rész

1. Egy emulzión áthaladó nagyenergiájú részecske pályáját — az emulzió rétegvastagságát elhanyagolva — síkmozgás pályájaként tekintjük. A valódi pályának e síkbeli vetületét a részecske nyomának fogjuk nevezni. A síkon úgy vesszük fel az X, Y koordinátarendszert, hogy a részecske nyomának kezdő és végpontja az X tengelyre essék. Mérjük a nyomnak ekvidisztáns $X_i = is$ abszcisszákhoz tartozó Y_i $(i = 0, 1, \ldots, N + 1;\ Y_0 = Y_{N+1} = 0)$ ordinátáit, amelyek a mérési hibák következtében valószínűségi változók (itt s a cellahosszt jelöli). A jelenség természetéből következik, hogy az Y_i valószínűségi változók várható értéke $\langle Y_i \rangle$ lineáris függvény.[2]

A további tárgyalás során az Y_i valószínűségi változók helyett ezeknek második differenciáit, a

$$d_i = Y_{i+1} - 2Y_i + Y_{i-1} \qquad i = 1, 2, \ldots, N$$

új változókat fogjuk tekinteni. Nyilvánvaló, hogy ezeknek várható értéke: $\langle d_i \rangle = 0$ $(i = 1, 2, \ldots, N)$. Feltesszük, hogy a d_i valószínűségi változók együttes eloszlása N-dimenziós normális eloszlás (lásd [2]), tehát ezek $f(\mathbf{d}')$ együttes sűrűségfüggvénye:

$$f(\mathbf{d}') = (2\pi)^{-\frac{N}{2}} (\det \mathbf{M})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\mathbf{d}'\mathbf{M}^{-1}\mathbf{d}\right]. \tag{1.1}$$

Itt $\mathbf{M} = [M_{ij}]$ jelöli a $\mathbf{d}' = (d_1, d_2, \ldots, d_N)$ valószínűségi vektorváltozó kovariancia mátrixát, tehát

$$M_{ij} = \langle d_i d_j \rangle \qquad i, j = 1, 2, \ldots, N.$$

Az $\mathbf{M}$ kovariancia mátrix két részből tevődik össze:

$$\mathbf{M} = \alpha^{(1)}\mathbf{A}^{(1)} + \alpha^{(2)}\mathbf{A}^{(2)}, \tag{1.2}$$

ahol az első tag az ütközés okozta szóródást (vagy Coulomb-szóródást), a második tag pedig a mérés hibája, a „háttér" hatása által okozott szóródást fejezi ki. Az $\mathbf{A}^{(1)} = [A_{ij}^{(1)}]$ és $\mathbf{A}^{(2)} = [A_{ij}^{(2)}]$ egyszerű szerkezetű Toeplitz-típusú mátrix elemeire $A_{ij}^{(\nu)} = A_{|i-j|}^{(\nu)}$ jelöléssel a következő kifejezések adódnak:

$$A_k^{(1)} = \begin{cases} 4 & \text{ha} \quad k = 0 \\ 1 & \text{ha} \quad k = 1 \\ 0 & \text{ha} \quad k \geq 2 \end{cases} \tag{1.3}$$

[2] Itt és a továbbiakban — a fizikában szokásos statisztikai jelölésmódnak megfelelően — $\langle \xi \rangle$ jelöli a ξ valószínűségi változó várható értékét.

$$A_k^{(2)} = \begin{cases} 6 & \text{ha} \quad k = 0 \\ -4 & \text{ha} \quad k = 1 \\ 1 & \text{ha} \quad k = 2 \\ 0 & \text{ha} \quad k \geqq 3 \end{cases} \tag{1.4}$$

Az $\alpha^{(1)}$ és $\alpha^{(2)}$ együttható a Coulomb-szóródás paraméterével, illetve a mérési hiba szórásával arányos (lásd [2]):

$$\alpha^{(1)} = as^3, \quad \alpha^{(2)} = \varkappa_2^2 \quad \text{és} \quad a = \frac{1}{6} \langle \vartheta^2 \rangle, \tag{1.5}$$

ahol $\varkappa_2^2$ az Y_i értékek mérési hibájának szórásnégyzetét jelöli. Az (1.5) összefüggések alapján látható, hogy a részecske energiájának meghatározása szempontjából milyen fontos az $\alpha^{(1)}$ együttható — illetőleg az (1.2) összefüggés miatt az $\alpha^{(1)}$ és $\alpha^{(2)}$ együtthatók — minél jobb becslése.

2. A következőkben az $\alpha^{(1)}$ és $\alpha^{(2)}$ együtthatóknak — a továbbiakban a ,,jel", illetve ,,zaj" paramétereinek — a becslésével foglalkozunk.

Az $\alpha^{(\nu)}$ paraméter becslésére az $\overline{\alpha}^{(\nu)}$ jelölést használjuk. A becslést kétféleképpen végezzük. Először a maximum likelihood módszer alapján, azután pedig a mért d_i differenciáknak egy kvadratikus alakjával. Megmutatjuk majd, hogy ha az utóbbi, kvadratikus becsléstől megköveteljük, hogy *torzítatlan és minimális szórású* legyen, akkor ez az ilyen értelemben *optimális kvadratikus becslés* megegyezik a maximum likelihood becsléssel.

A

$$\frac{\partial \log f(\mathbf{d}')}{\partial \alpha^{(\nu)}} = 0 \qquad \alpha^{(\nu)} = \overline{\alpha}^{(\nu)}; \quad \nu = 1, 2$$

maximum likelihood egyenletek a kijelölt parciális deriválás elvégzése után a

$$-\frac{1}{2} \frac{\partial \log (\det \mathbf{M})}{\partial \alpha^{(\nu)}} - \frac{1}{2} \mathbf{d}' \frac{\partial \mathbf{M}^{-1}}{\partial \alpha^{(\nu)}} \mathbf{d} = 0 \qquad \alpha^{(\nu)} = \overline{\alpha}^{(\nu)}; \quad \nu = 1, 2 \tag{1.6}$$

egyenletrendszerre vezetnek. Ebből kell az $\overline{\alpha}^{(\nu)}$ becsléseket a mért d_i differenciák függvényében kifejeznünk. Deriváljuk e célból az

$$\mathbf{M}^{-1} \mathbf{M} = \mathbf{E}$$

összefüggést $\alpha^{(\nu)}$ szerint. A

$$\frac{\partial \mathbf{M}}{\partial \alpha^{(\nu)}} = \mathbf{A}^{(\nu)} \qquad \nu = 1, 2$$

összefüggés alapján

$$\frac{\partial \mathbf{M}^{-1}}{\partial \alpha^{(\nu)}} \mathbf{M} + \mathbf{M}^{-1} \mathbf{A}^{(\nu)} = 0 \qquad \nu = 1, 2,$$

így az $\mathbf{M}^{-1}$ mátrix $\alpha^{(\nu)}$ szerinti parciális deriváltjára a

$$\frac{\partial \mathbf{M}^{-1}}{\partial \alpha^{(\nu)}} = - \mathbf{M}^{-1} \mathbf{A}^{(\nu)} \mathbf{M}^{-1} \qquad \nu = 1, 2 \tag{1.7}$$

kifejezés adódik. Ha most figyelembe vesszük, hogy

$$(1.8)\qquad \left\langle \frac{\partial \log f(\mathbf{d}')}{\partial \alpha^{(\nu)}} \right\rangle = 0 \qquad \nu = 1, 2,$$

akkor (1.7) és (1.8) alapján a

$$(1.9)\qquad \frac{\partial \log (\det \mathbf{M})}{\partial \alpha^{(\nu)}} = \operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}) \qquad \nu = 1, 2$$

összefüggésre jutunk. Behelyettesítve az (1.7) és (1.9) kifejezéseket, az (1.6) egyenletrendszer

$$(1.10)\qquad \frac{1}{2}\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}) - \frac{1}{2}\mathbf{d}'(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}\mathbf{M}^{-1})\,\mathbf{d} = 0 \qquad \alpha^{(\nu)} = \overline{\alpha}'^{(\nu)};\ \nu = 1,2$$

alakban írható. Ha a sűrűségfüggvény logaritmusát most $\alpha^{(\mu)}$ szerint deriváljuk, akkor a

$$\frac{\partial^2 \log (\det \mathbf{M})}{\partial \alpha^{(\mu)}\, \partial \alpha^{(\nu)}} = -\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\mu)}\mathbf{M}^{-1}\mathbf{A}^{(\nu)}),$$

illetve a

$$\left\langle \mathbf{d}' \frac{\partial^2 \mathbf{M}^{-1}}{\partial \alpha^{(\mu)}\, \partial \alpha^{(\nu)}} \mathbf{d} \right\rangle = \operatorname{spur}\left(\frac{\partial^2 \mathbf{M}^{-1}}{\partial \alpha^{(\mu)}\, \partial \alpha^{(\nu)}} \mathbf{M} \right) = 2\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\mu)}\mathbf{M}^{-1}\mathbf{A}^{(\nu)})$$

összefüggések alapján a $\log f(\mathbf{d}')$ függvény második parciális deriváltjainak várható értékére a

$$\left\langle \frac{\partial^2 \log f(\mathbf{d}')}{\partial \alpha^{(\mu)}\, \partial \alpha^{(\nu)}} \right\rangle = -\frac{1}{2}\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\mu)}\mathbf{M}^{-1}\mathbf{A}^{(\nu)}) \qquad \mu, \nu = 1, 2$$

kifejezéseket kapjuk. Vezessük be ezekre a $-Q^*_{\mu\nu}$ jelölést:

$$(1.11)\qquad Q^*_{\mu\nu} = -\left\langle \frac{\partial \log f(\mathbf{d}')}{\partial \alpha^{(\mu)}\, \partial \alpha^{(\nu)}} \right\rangle = \frac{1}{2}\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\mu)}\mathbf{M}^{-1}\mathbf{A}^{(\nu)}) \qquad \mu, \nu = 1, 2.$$

Könnyen belátható, hogy a $Q^*_{\mu\nu}$ kifejezésekre érvényesek az alábbi összefüggések:

$$Q^*_{\mu\nu} = Q^*_{\nu\mu} \qquad \mu, \nu = 1, 2,$$

$$(1.12)\qquad \alpha^{(1)} Q^*_{1\nu} + \alpha^{(2)} Q^*_{2\nu} = \frac{1}{2}\operatorname{spur}(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}) \qquad \nu = 1, 2,$$

$$\alpha^{(1)2} Q^*_{11} + 2\,\alpha^{(1)}\alpha^{(2)} Q^*_{12} + \alpha^{(2)2} Q^*_{22} = \frac{1}{2}\operatorname{spur}\mathbf{E} = \frac{N}{2}.$$

Az (1.12) összefüggések alapján a megoldandó (1.10) egyenletrendszer végül is

$$\alpha^{(1)} Q^*_{1\nu} + \alpha^{(2)} Q^*_{2\nu} = \frac{1}{2}\mathbf{d}'(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}\mathbf{M}^{-1})\,\mathbf{d} \qquad \alpha^{(\nu)} = \overline{\alpha}^{(\nu)};\ \nu = 1, 2$$

alakban írható. Ebből pedig az $\overline{\alpha}^{(\nu)}$ becslések már kifejezhetők:

$$\overline{\alpha}^{(\nu)} = \frac{1}{2}\mathbf{d}'(Q_{\nu 1}\,\mathbf{M}^{-1}\,\mathbf{A}^{(1)}\,\mathbf{M}^{-1} + Q_{\nu 2}\,\mathbf{M}^{-1}\,\mathbf{A}^{(2)}\,\mathbf{M}^{-1})\,\mathbf{d} \qquad \nu = 1, 2, \tag{1.13}$$

ahol $\mathbf{Q} = [Q_{\nu\mu}]$ jelöli a $Q^*_{\nu\mu}$ elemekből alkotott mátrix inverzét.

Határozzuk meg végül az $\overline{\alpha}^{(\nu)}$ becslések, mint valószínűségi változók kovariancia mátrixát. Mivel a becslések várható értéke

$$\langle\overline{\alpha}^{(\nu)}\rangle = \frac{1}{2}\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)}) \qquad \nu = 1, 2,$$

a kovariancia mátrix elemeire

$$\langle\delta\overline{\alpha}^{(\nu)}\,\delta\overline{\alpha}^{(\mu)}\rangle = \frac{1}{4}\left\langle \mathbf{d}'\left(\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)}\,\mathbf{M}^{-1}\right)\mathbf{d}\mathbf{d}'\left(\sum_{\lambda=1}^{2} Q_{\mu\lambda}\,\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)}\,\mathbf{M}^{-1}\right)\mathbf{d}\right\rangle -$$

$$- \frac{1}{4}\left[\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)})\right]\left[\sum_{\lambda=1}^{2} Q_{\mu\lambda}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)})\right] =$$

$$= \frac{1}{4}\sum_{\varkappa,\lambda=1}^{2} Q_{\nu\varkappa}\,Q_{\mu\lambda}\sum_{i,j=1}^{N}\sum_{l,m=1}^{N}\sum_{p,q=1}^{N}\sum_{r,s=1}^{N} M^*_{ip}\,A^{(\varkappa)}_{pq}\,M^*_{qj}\,M^*_{lr}\,A^{(\lambda)}_{rs}\,M^*_{sm}\,\langle d_i\,d_j\,d_l\,d_m\rangle -$$

$$- \frac{1}{4}\left[\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)})\right]\left[\sum_{\lambda=1}^{2} Q_{\mu\lambda}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)})\right]$$

adódik. ($M^*_{\alpha\beta}$ jelöli az $\mathbf{M}^{-1}$ mátrix elemeit.) Az itt szereplő $\langle d_i\,d_j\,d_l\,d_m\rangle$ várható értékeket legkönnyebben a $\mathbf{d}' = (d_1, d_2, \ldots, d_N)$ vektorváltozó karakterisztikus függvényéből határozhatjuk meg. A $\mathbf{d}'$ valószínűségi vektorváltozó karakterisztikus függvénye

$$\varphi_{\mathbf{d}'}(t_1, t_2, \ldots, t_N) = \exp\left|-\frac{1}{2}\sum_{k,n=1}^{N} M_{kn}\,t_k\,t_n\right|,$$

így a keresett várható értékekre a karakterisztikus függvény négyszeri parciális deriválásával a

$$\langle d_i\,d_j\,d_l\,d_m\rangle = \left.\frac{\partial\varphi_{\mathbf{d}'}(t_1, t_2, \ldots, t_N)}{\partial t_i\,\partial t_j\,\partial t_l\,\partial t_m}\right|_{t_1=t_2=\ldots=t_N=0} = \tag{1.14}$$

$$= M_{ij}\,M_{lm} + M_{il}\,M_{jm} + M_{jl}\,M_{im}$$

kifejezést kapjuk. Ennek alapján

$$\sum_{\varkappa,\lambda=1}^{2} Q_{\nu\varkappa}\,Q_{\mu\lambda}\sum_{i,j=1}^{N}\sum_{l,m=1}^{N}\sum_{p,q=1}^{N}\sum_{r,s=1}^{N} M^*_{ip}\,A^{(\varkappa)}_{pq}\,M^*_{qj}\,M^*_{lr}\,A^{(\lambda)}_{rs}\,M^*_{sm}\,\langle d_i\,d_j\,d_l\,d_m\rangle =$$

$$= \sum_{\varkappa,\lambda=1}^{2} Q_{\nu\varkappa}\,Q_{\mu\lambda}\,[\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)})\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)}) + 2\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)}\,\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)})] =$$

$$= \sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\varkappa)})\sum_{\lambda=1}^{2} Q_{\mu\lambda}\,\mathrm{spur}\,(\mathbf{M}^{-1}\,\mathbf{A}^{(\lambda)}) + 4\sum_{\varkappa=1}^{2}\sum_{\lambda=1}^{2} Q_{\nu\varkappa}\,Q_{\mu\lambda}\,Q^*_{\varkappa\lambda}\,.$$

Figyelembe véve még, hogy $\sum_{\varkappa=1}^{2}\sum_{\lambda=1}^{2} Q_{\nu\varkappa} Q_{\mu\lambda} Q^*_{\varkappa\lambda} = Q_{\nu\mu}$, a becslések kovariancia mátrixának elemeire végül is

$$\langle\delta\overline{\alpha}^{(\nu)}\delta\overline{\alpha}^{(\mu)}\rangle = Q_{\nu\mu} \qquad \nu,\mu = 1,2 \tag{1.15}$$

adódik. Ezzel a maximum likelihood módszer alapján megadtuk az $\alpha^{(\nu)}$ paraméterek becslését és e becslések kovariancia mátrixát. Érdekes megjegyezni, hogy az (1.11) és (1.15) összefüggések alapján a $\log f(\mathbf{d}')$ függvény Hesse-féle mátrixának várható értéke és a becslések kovariancia mátrixa között — egy (—1)-es faktortól eltekintve — reciprocitási kapcsolat áll fenn.

3. Most rátérünk annak a másik módszernek az ismertetésére, amellyel a paraméterek becslését végezni kívánjuk. Mint láttuk, az (1.13) összefüggés az $\alpha^{(\nu)}$ paraméter becslését a mért d_i differenciák egy kvadratikus alakjaként adja meg. Bár az (1.13) alatti kvadratikus alak mátrixa maga is függ még a becsült paraméterektől, mégis célszerűnek látszik, hogy a mért d_i differenciáknak egy kvadratikus alakjából kiindulva próbáljunk az $\alpha^{(\nu)}$ paraméterekre torzítatlan, jó hatásfokú becslést adni.

Legyen tehát

$$\overline{\alpha}^{(\nu)} = \mathbf{d}'\,\mathbf{C}^{(\nu)}\,\mathbf{d} \qquad \nu = 1,2 \tag{1.16}$$

az $\alpha^{(\nu)}$ paraméternek egy kvadratikus becslése, ahol $\mathbf{C}^{(\nu)}$ a kvadratikus becslés mátrixa.

Mivel megkívánjuk, hogy a becslés torzítatlan legyen, tehát

$$\langle\overline{\alpha}^{(\nu)}\rangle = \alpha^{(\nu)} \qquad \nu = 1,2$$

(eljesüljön, így a $\mathbf{C}^{(\nu)}$ mátrixnak ki kell elégítenie a

$$\mathrm{spur}\,(\mathbf{C}^{(\nu)}\mathbf{A}^{(\mu)}) = \delta_{\nu\mu} \qquad \nu,\mu = 1,2 \tag{1.17}$$

feltételeket, ahol $\delta_{\nu\mu}$ a Kronecker-szimbolumot jelenti:

$$\delta_{\nu\mu} = \begin{cases} 1 & \text{ha} \quad \nu = \mu \\ 0 & \text{ha} \quad \nu \neq \mu . \end{cases}$$

Ahhoz, hogy a becslések hatásfokáról beszélhessünk, fel kell írnunk az $\overline{\alpha}^{(1)}$, $\overline{\alpha}^{(2)}$ valószínűségi változók $\mathbf{Q} = [Q_{\nu\mu}]$ kovariancia mátrixának elemeit. Figyelembe véve a $\langle d_i\, d_j\, d_l\, d_m\rangle$ várható értékek (1.14) alatti kifejezéseit, a kovariancia mátrix elemeire

$$Q_{\nu\mu} = \langle\delta\overline{\alpha}^{(\nu)}\delta\overline{\alpha}^{(\mu)}\rangle = \sum_{i,j=1}^{N}\sum_{l,m=1}^{N} C^{(\nu)}_{ij} C^{(\mu)}_{lm}\,[\langle d_i\, d_j\, d_l\, d_m\rangle - M_{ij}\, M_{lm}] = \tag{1.18}$$

$$= 2\,\mathrm{spur}\,(\mathbf{C}^{(\nu)}\,\mathbf{M}\mathbf{C}^{(\mu)}\,\mathbf{M}) \qquad \nu,\mu = 1,2$$

adódik. Jó hatásfokú becsléseket akkor kapunk, ha a $\mathbf{C}^{(\nu)}$ mátrixokat úgy választjuk meg, hogy a $Q_{\nu\nu}$ szórásnégyzeteket minimalizálja. Így tehát a következő szélsőérték-feladatra jutottunk: Meghatározandó

$$\min_{(\mathbf{C}^{(\nu)})} Q_{\nu\nu} \qquad \nu = 1,2 \tag{1.19}$$

és azok a $\mathbf{C}^{(\nu)}$ mátrixok, amelyek a minimumot szolgáltatják, ahol a $\mathbf{C}^{(\nu)}$ befutja azoknak az N-edrendű szimmetrikus mátrixoknak az osztályát, amelyek az (1.17) feltételi egyenleteket kielégítik. Ezt a feltételes szélsőértékfeladatot a Lagrange-féle multiplikátorok segítségével oldjuk meg. A minimumot szolgáltató $\mathbf{C}^{(\nu)}$ mátrixok meghatározására így a

$$4\,\mathbf{M}\mathbf{C}^{(\nu)}\mathbf{M} - \sum_{\varkappa=1}^{2} L_{\nu\varkappa}\mathbf{A}^{(\varkappa)} = 0 \qquad \nu = 1, 2 \tag{1.20}$$

egyenleteket kapjuk, ahol $L_{\nu\varkappa}$ jelöli a Lagrange-féle multiplikátorokat.

A feladat tehát az (1.17) és (1.20) egyenletekből a $\mathbf{C}^{(\nu)}$ $(\nu = 1,2)$ mátrixok és az $L_{\nu\varkappa}$ $(\nu, \varkappa = 1, 2)$ multiplikátorok kiszámítása. Könnyen belátható, hogy az (1.17) és (1.20) egyenletek az ismeretleneket egyértelműen meghatározzák. Mielőtt az egyenletrendszer megoldására rátérnénk, megmutatjuk, hogy a Lagrange-féle multiplikátorok a kovariancia mátrix elemeinek kétszeresei:

$$L_{\nu\mu} = 2\,Q_{\nu\mu} \qquad \nu, \mu = 1, 2\,. \tag{1.21}$$

Valóban, ha az (1.20) egyenletet balról a $\mathbf{C}^{(\mu)}$ mátrixszal megszorozzuk és képezzük az így nyert mátrixok spurját, akkor az (1.17) és (1.18) összefüggés felhasználásával az (1.21) összefüggés adódik.

Ezek után a $\mathbf{C}^{(\nu)}$ és a $\mathbf{Q}$ mátrixok kiszámítása céljából a következőképpen járunk el. Helyettesítsük a Lagrange-féle multiplikátoroknak az (1.21) kifejezését az (1.20) egyenletbe és szorozzuk meg az egyenletet mindkét oldalról az $\mathbf{M}^{-1}$ mátrixszal. Így a $\mathbf{C}^{(\nu)}$ mátrixra

$$\mathbf{C}^{(\nu)} = \frac{1}{2}\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathbf{M}^{-1}\mathbf{A}^{(\varkappa)}\mathbf{M}^{-1} \tag{1.22}$$

adódik. Ebből az összefüggésből az (1.17) feltételi egyenletek felhasználásával kifejezhetők a $Q_{\nu\varkappa}$ szórások. Szorozzuk meg ugyanis az (1.22) egyenletet jobbról az $\mathbf{A}^{(\mu)}$ mátrixszal és vegyük az így adódó mátrixok spurját. Figyelembe véve az (1.17) feltételi egyenleteket, ezzel az

$$\frac{1}{2}\,\mathrm{spur}\sum_{\varkappa=1}^{2} Q_{\nu\varkappa}\,\mathbf{M}^{-1}\mathbf{A}^{(\varkappa)}\,\mathbf{M}^{-1}\mathbf{A}^{(\mu)} = \delta_{\nu\mu} \qquad \nu, \mu = 1, 2 \tag{1.23}$$

összefüggésre jutunk. Innen már közvetlenül következik, hogy ha $Q^{*}_{\nu\mu}$ $(\nu, \mu = 1,2)$ jelöli a $\mathbf{Q} = [Q_{\nu\mu}]$ kovariancia mátrix inverzének az elemeit, akkor ezekre a

$$Q^{*}_{\nu\mu} = \frac{1}{2}\,\mathrm{spur}\,(\mathbf{M}^{-1}\mathbf{A}^{(\nu)}\,\mathbf{M}^{-1}\mathbf{A}^{(\mu)}) \qquad \nu, \mu = 1, 2 \tag{1.24}$$

kifejezést kapjuk. Figyelembe véve, hogy az (1.24) jobboldalán álló kifejezés N-nek monton növekvő függvénye, könnyű belátni, hogy a becslések $\mathbf{Q}$ kovariancia mátrixa a méréspontok számának növelésével 0-hoz tart, tehát *a becslés konzisztens.*

Ezzel az optimális kvadratikus becslés mátrixait és kovariancia mátrixát meghatároztuk.[3]

[3] E mátrixok ugyan még tartalmazzák a becsülendő paramétereket, azonban amint a III. részben látni fogjuk, a kapott összefüggéseink alkalmasak azok meghatározására.

Összevetve az (1.16) és (1.22) összefüggéseket az (1.13) összefüggéssel, illetve az (1.24) összefüggést az (1.15) alattival, látható, hogy az *optimális kvadratikus becslés* valóban *megegyezik a maximum likelihood becsléssel*, amint állítottuk.

A kapott (1.22) és (1.24) képletek azonban $N \gg 1$ esetén, — vagyis amikor a mérési pontokat sűrítjük, — gyakorlati számításra nem alkalmasak. Célszerűnek látszik, hogy ilyen esetre érvényes, egyszerűbb és könnyen kezelhető aszimptotikus formulákat vezessünk le. A következőkben ezzel a feladattal foglalkozunk.

4. Az (1.2) és (1.24) összefüggésből látható, hogy a $Q^*_{\nu\mu}$ elemek az $\alpha^{(1)}$ és $\alpha^{(2)}$ paraméterek homogén kifejezései. Ezért célszerű, ha a paraméterek, illetve becsléseik hányadosára egy új változót vezetünk be és a további tárgyalás során valamennyi előforduló kifejezést e változó függvényében írunk fel. Jelölje x illetve $\bar{x}$ ezt a változót, tehát

$$x = \frac{\alpha^{(2)}}{\alpha^{(1)}} \quad \text{és} \quad \bar{x} = \frac{\bar{\alpha}^{(2)}}{\bar{\alpha}^{(1)}}. \tag{1.25}$$

Vezessük be a

$$\mathbf{P} = \frac{1}{\alpha^{(1)}} \mathbf{M} = \mathbf{A}^{(1)} + x\,\mathbf{A}^{(2)} \tag{1.26}$$

mátrixot és a

$$Q^*_{\nu\mu} = \frac{N}{2\,\alpha^{(1)2}}\, q^*_{\nu\mu}(x) \qquad \nu, \mu = 1, 2 \tag{1.27}$$

összefüggés segítségével definiált $q^*_{\nu\mu}(x)$ függvényeket. Ezekre az (1.2) és (1.26) összefüggések alapján a következő relációk érvényesek:

$$q^*_{\nu\mu}(x) = q^*_{\mu\nu}(x) \qquad \nu, \mu = 1, 2\,,$$

$$q^*_{1\nu}(x) + x q^*_{2\nu}(x) = \frac{1}{N} \operatorname{spur}(\mathbf{P}^{-1}\mathbf{A}^{(\nu)}) \qquad \nu = 1, 2\,,$$

$$q^*_{11}(x) + 2\,x q^*_{12}(x) + x^2\, q^*_{22}(x) = 1\,.$$

Az (1.24) összefüggés helyett

$$q^*_{\nu\mu}(x) = \frac{1}{N} \operatorname{spur}(\mathbf{P}^{-1}\mathbf{A}^{(\nu)}\mathbf{P}^{-1}\mathbf{A}^{(\mu)}) \qquad \nu, \mu = 1, 2 \tag{1.28}$$

írható. Az (1.27) összefüggésből látszik, hogy az így nyert $q^*_{\nu\mu}(x)$ elemekből alkotott mátrix inverzének $q_{\nu\mu}(x)$ elemei és a keresett $\mathbf{Q}$ kovariancia mátrix elemei között a következő egyszerű összefüggés áll fenn:

$$Q_{\nu\mu} = \frac{2\,\alpha^{(1)2}}{N}\, q_{\nu\mu}(x) \qquad \nu, \mu = 1, 2. \tag{1.29}$$

Feladatunk tehát abban áll, hogy a $q^*_{\nu\mu}(x)$ (1.28) kifejezésekre, majd az inverz mátrix $q_{\nu\mu}(x)$ elemeire $N \gg 1$ esetén érvényes aszimptotikus formulákat nyerjünk. Ezt az a körülmény teszi lehetővé, hogy az (1.28) összefüggés jobb oldalán álló $\mathbf{P}$ és $\mathbf{A}^{(\nu)}$ mátrixok — egy-egy sarokelemtől eltekintve — a $\mathbf{H} = [H_{ij}]$

$$H_{ij} = \delta_{1,|i-j|} \qquad i, j = 1, 2, \ldots, N \tag{1.30}$$

mátrix polinomjaiként írhatók fel. A $\mathbf{H}$ mátrixot a továbbiakban *elemi Toeplitz-féle* mátrixnak fogjuk nevezni. Az (1.28) összefüggések szerint tehát a $q^*_{\nu\mu}(x)$ kifejezések meghatározásához a $\mathbf{H}$ mátrix bizonyos racionális tört kifejezéseinek a spurját kell képezni. (A sarokelemek módosításával ugyanis elérhető, hogy az (1.28) összefüggés jobb oldalán álló mátrixok felcserélhetők legyenek és így a négy tényezős mátrix-szorzat a $\mathbf{H}$ mátrixnak egy racionális törtkifejezéseként írható fel.)

A mondottak szerint tehát az $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ és a $\mathbf{P}$ mátrix, az (1.3), (1.4) és (1.26) definíciójuk alapján a $\mathbf{H}$ mátrix segítségével a következőképpen fejezhető ki:

$$\begin{aligned} \mathbf{A}^{(1)} &= 4\,\mathbf{E} + \mathbf{H} \\ \mathbf{A}^{(2)} &= 4\,\mathbf{E} - 4\,\mathbf{H} + \mathbf{H}^2 + \mathbf{H}_0 \\ \mathbf{P} &= 4\,(1 + x)\,\mathbf{E} + (1 - 4x)\,\mathbf{H} + x\,\mathbf{H}^2 + x\,\mathbf{H}_0 \,. \end{aligned} \tag{1.31}$$

Az itt szereplő $\mathbf{H}_0$ mátrixnak csak két zérustól különböző eleme van (a bal felső és a jobb alsó sarokelem). Megmutatható, hogy ha elhanyagoljuk a $\mathbf{H}_0$ mátrixot, akkor a végeredményben elkövetett hiba $\frac{1}{N}$ nagyságrendű, ezért a továbbiakban a $\mathbf{H}_0$ mátrixot elhagyjuk.

Vezessük be az (1.28) összefüggésben szereplő mátrix-szorzatra az

$$\mathbf{R}^{(\nu\mu)} = \mathbf{P}^{-1}\,\mathbf{A}^{(\nu)}\,\mathbf{P}^{-1}\,\mathbf{A}^{(\mu)} \qquad \nu, \mu = 1, 2 \tag{1.32}$$

jelölést. Az (1.31) összefüggések felhasználásával tehát (elhagyva a $\mathbf{H}_0$ mátrixot),

$$\begin{aligned} \mathbf{R}^{(11)} &= \left[\frac{4\,\mathbf{E} + \mathbf{H}}{4\,(1 + x)\,\mathbf{E} + (1 - 4x)\,\mathbf{H} + x\,\mathbf{H}^2}\right]^2 \\ \mathbf{R}^{(12)} = \mathbf{R}^{(21)} &= \frac{(4\,\mathbf{E} + \mathbf{H})\,(4\,\mathbf{E} - 4\,\mathbf{H} + \mathbf{H}^2)}{[4\,(1 + x)\,\mathbf{E} + (1 - 4x)\,\mathbf{H} + x\,\mathbf{H}^2]^2} \\ \mathbf{R}^{(22)} &= \left[\frac{4\,\mathbf{E} - 4\,\mathbf{H} + \mathbf{H}^2}{4\,(1 + x)\,\mathbf{E} + (1 - 4x)\,\mathbf{H} + x\,\mathbf{H}^2}\right]^2 \end{aligned} \tag{1.33}$$

adódik. Vegyük most tekintetbe, hogy egy mátrix valamely racionális kifejezésének a sajátértékei megegyeznek a mátrix sajátértékeinek e racionális kifejezéseivel, továbbá, hogy egy mátrix spurját megadja a sajátértékeinek az összege. Ennek alapján, ha az $\mathbf{R}^{(\nu\mu)}$ mátrix sajátértékeit $\varrho_k^{(\nu\mu)}$ jelöli, az (1.33) összefüggések szerint pedig az $\mathbf{R}^{(\nu\mu)}$ mátrixokat $\mathbf{H}$ függvényében az

$$\mathbf{R}^{(\nu\mu)} = f_{\nu\mu}(\mathbf{H}) \qquad \nu, \mu = 1, 2 \tag{1.34}$$

függvénykapcsolat határozza meg, akkor a keresett $q^*_{\nu\mu}(x)$ függvényekre

$$q^*_{\nu\mu}(x) = \frac{1}{N}\sum_{k=1}^{N} \varrho_k^{(\nu\mu)} = \frac{1}{N}\sum_{k=1}^{N} f_{\nu\mu}(\eta_k) \qquad \nu, \mu = 1, 2 \tag{1.35}$$

adódik, ahol

$$\eta_k = 2\cos\frac{k\pi}{N + 1} \qquad k = 1, 2, \ldots, N$$

a $\mathbf{H}$ elemi Toeplitz-féle mátrix sajátértékeit jelenti (lásd pl. [5] 155—156 o.).

Az (1.35) összefüggésben szereplő összeg határozott integrál közelítő összegeként tekinthető, ezért ha $N \gg 1$, akkor az összeg az integrállal helyettesíthető:

$$q^*_{\nu\mu}(x) = \frac{1}{\pi} \int_0^{\pi} f_{\nu\mu}(2\cos z)\, dz\,.$$

Figyelembe véve az (1.33) és (1.32) összefüggéseket, innen

$$q^*_{11}(x) = \frac{1}{\pi} \int_0^{\pi} \left[\frac{1 + \frac{1}{2}\cos z}{1 + x + \left(\frac{1}{2} - 2x\right)\cos z + x\cos^2 z} \right]^2 dz,$$

$$q^*_{12}(x) = \frac{1}{\pi} \int_0^{\pi} \frac{\left(1 + \frac{1}{2}\cos z\right)(1 - \cos z)^2}{\left[1 + x + \left(\frac{1}{2} - 2x\right)\cos z + x\cos^2 z\right]^2}\, dz,$$

$$q^*_{22}(x) = \frac{1}{\pi} \int_0^{\pi} \left[\frac{(1 - \cos z)^2}{1 + x + \left(\frac{1}{2} - 2x\right)\cos z + x\cos^2 z} \right]^2 dz$$

adódik.

Ha ezeket az integrálokat kiszámítjuk, és az így nyert $[q^*_{\nu\mu}(x)]$ másodrendű mátrixot invertáljuk, akkor az inverz mátrix $q_{\nu\mu}(x)$ elemeire az alábbi kifejezéseket kapjuk:

$$q_{11}(x) = \frac{1}{K}\{32v^5 + 117v^4 + 126v^3 + 13v^2 - 72v - 48\},$$

$$q_{12}(x) = -\frac{1}{24K}\{(4v^3 + 20v^2 + 40v + 32)v^3\sqrt{6(v+2)} +$$

$$+ (15v^6 + 90v^5 + 222v^4 + 114v^3 - 117v^2 - 216v - 144)\}, \tag{1.36}$$

$$q_{22}(x) = \frac{1}{576K}\{(12v^5 + 60v^4 + 100v^3 + 92v^2 + 120v + 96)v^3\sqrt{6(v+2)} +$$

$$+ (45v^8 + 270v^7 + 591v^6 + 612v^5 + 423v^4 + 270v^3 - 171v^2 - 648v - 432)\},$$

ahol

$$K = (12v^3 + 30v^2 + 22v + 8)v\sqrt{6(v+2)} - (3v^4 + 18v^3 + 51v^2 + 72v + 48)$$

és

$$v = \sqrt{3 + 24x}\,.$$

A $q_{\nu\mu}(x)$ függvények az 1. ábrán láthatók.

Az (1.26) és az (1.18) összefüggés alapján meghatározható az $\overline{\alpha}^{(\nu)}$ valószínűségi változók relatív szórása.

Tekintettel arra, hogy a feladat tulajdonképpen az $\alpha^{(1)}$ jel-paraméter minél jobb becslése, ezért a továbbiakban elsősorban $\alpha^{(1)}$ becslésének relatív szórását fogjuk vizsgálni. Jelölje ezt σ, akkor

$$\sigma^2 = \frac{\langle \delta(\overline{\alpha}^{(1)2} \rangle}{\alpha^{(1)2}} = \frac{2}{N}\, q_{11}(x)\,, \tag{1.37}$$

ahol

$$q_{11}(x) = \frac{q_{22}^*(x)}{q_{11}^*(x)\, q_{22}^*(x) - q_{12}^{*2}(x)}\,.$$

Gyakorlati alkalmazásnál az (1.37) egyenlet jobboldali kifejezésében $\alpha^{(1)}$ és x valódi értéke helyére a becsléssel kapott $\overline{\alpha}^{(1)}$ és $\overline{x}$ értékeket kell behelyettesítenünk.

5. Az alábbiakban azzal foglalkozunk, hogy egyrészt „zaj" jelenléte, másrészt a mérések számának a növelése hogyan befolyásolja az $\alpha^{(1)}$ jel-paraméter becslését.

Abban az esetben, ha tudjuk, hogy a mérési hiba *elhanyagolható*, akkor az $\alpha^{(1)}$ jel-paraméter becslésének relatív szórását a következő kifejezés adja:

$$\frac{\langle(\delta\overline{\alpha}^{(1)})^2\rangle^{\frac{1}{2}}}{\alpha^{(1)}} = \left(\frac{2}{N}\right)^{\frac{1}{2}}. \tag{1.38}$$

Összehasonlítva az (1.37) és (1.38) kifejezést, látható, hogy a $q_{11}(x)$ függvény annak mértékéül szolgál, hogy zaj jelenléte mennyire rontja az elérhető pontosságot.

Érdekes megjegyezni, hogy

$$q_{11}(0) = 1{,}7561,$$

azaz a háttér figyelembevételével számoló két-paraméteres módszer segítségével $\alpha^{(1)}$ kisebb pontossággal becsülhető, mint az egyparaméteres módszerrel, még abban az esetben is, amikor a kétparaméteres módszerrel végzett becslésnél $\overline{\alpha}^{(2)} \sim 0$ adódik. Tehát zaj létezésének a puszta feltételezése is számottevően befolyásolja a jel-paraméter becslésének a pontosságát, még abban az esetben is, ha a zaj elhanyagolhatóan kicsi.

Az (1.37) egyenlet abban az esetben adja meg az $\overline{\alpha}^{(1)}$ becslés relatív szórását, ha az $\alpha^{(1)}$ és $\alpha^{(2)}$ paraméterekre egyidejűleg végezzük el a becslést a mért d_i differenciákkal.

Ha az $\alpha^{(2)}$ zaj-paraméter más nyomokon végzett független mérésekből becsülhető, akkor a jel-paraméter becslésének hatásfoka valamivel javítható. Jelölje ebben az esetben σ_1 az $\overline{\alpha}^{(1)}$ relatív szórását, akkor nyilvánvalóan

$$\sigma_1^2 = \frac{2}{N\, q_{11}^*(x)}\,.$$

Annak mértékéül, hogy a becslés hatásfokában milyen veszteséget okoz, ha a zaj-paramétert nem tudjuk előzetesen más független mérésekből becsülni, a σ^2 és σ_1^2 relatív szórásnégyzetek hányadosa szolgál:

$$p(x) = \frac{\sigma^2}{\sigma_1^2} = q_{11}(x)\, q_{11}^*(x) = \frac{1}{1 - \dfrac{q_{12}^{*2}(x)}{q_{11}^*(x)\, q_{22}^*(x)}}.$$

A $p(x)$ függvény menetét az 1. ábra mutatja.

Most vizsgáljuk meg azt, hogy a becslés pontossága hogyan növekszik a mérési pontok sűrítésével. Tekintsünk egy nyomot, amelynek az X tengelyen való vetülete L hosszúságú. Ha $N + 2$ mérést végzünk, akkor a cellahossz:

$$s = \frac{L}{N+1}.$$

Bevezetve a

$$\gamma = \frac{\varkappa_2^2}{a L}$$

jelölést, az (1.5) és (1.25) összefüggésből

$$x = \frac{\alpha^{(2)}}{\alpha^{(1)}} = (N+1)^3 \gamma$$

következik. Ezzel a relatív szórásnégyzet (1.37) képlete a

$$\sigma^2 = \frac{2\, q_{11}([N+1]^3 \gamma)}{N+1}$$

alakban, illetve

$$z = (N+1)\, \gamma^{\frac{1}{3}}$$

transzformációval

(1.39) $$\sigma^2 = 2 \gamma^{\frac{1}{3}} \left(1 + \frac{1}{N}\right) r(z)$$

alakban írható, ahol

(1.40) $$r(z) = \frac{q_{11}(z^3)}{z}.$$

Az $r(z)$ függvényt a 3. ábra tünteti fel. Az $r(z)$ függvény monoton csökkenő, amiből következik, hogy $\overline{\alpha}^{(1)}$ relatív szórása a mérési pontok sűrítésével monoton csökken. Ez az eredmény várható is, hiszen minél több koordinátát mérünk, annál több információhoz jutunk, tehát annál jobb becslést tudunk adni.

A $q_{11}(x)$ függvény (1.36) alatti kifejezéséből következik, hogy nagy x értékekre

$$q_{11}(x) \sim x^{\frac{1}{4}},$$

ahonnan az (1.40) összefüggés felhasználásával nagy N értékekre

$$\sigma \sim N^{-\frac{1}{8}}$$

adódik. Ez azt jelenti, hogy $\alpha^{(1)}$ becslésének a pontossága mindössze a mérési pontok számának *nyolcadik gyökével* nő, azaz a növekedés olyan lassúvá válik, hogy bizonyos határon túl nem érdemes a mérési pontokat sűríteni.

II. rész

Az első részben foglaltak alapján látható, hogy a $\mathbf{Q}$ kovariancia mátrix ismeretében meghatározhatók az optimális kvadratikus becslés $\mathbf{C}^{(\nu)}$ mátrixai (lásd [1.16] és [1.22]). Gyakorlatilag azonban ezeknek a $\mathbf{C}^{(\nu)}$ mátrixoknak a meghatározása igen fáradságos és hosszadalmas numerikus munkát igényelne. Ezért kívánatosnak látszik az első részben ismertetett optimális becslés helyett olyan *közelítő eljárás* kidolgozása, amelynél a megfelelő kvadratikus alakok mátrixai egyszerű szerkezetűek, tehát a jel- és zaj-paraméter becslése a mért d_i differenciákból könnyen számítható, $\bar{\alpha}^{(1)}$ relatív szórása pedig nem növekszik számottevően. Ezért a következőkben az optimális kvadratikus becsléshez tartozó $\mathbf{C}^{(\nu)}$ mátrixok helyett egyszerű szerkezetű, könnyen kezelhető mátrixok segítségével adunk az $\alpha^{(\nu)}$ paraméterekre torzítatlan becslést. Az optimális becsléstől való megkülönböztetésül jelölje $\alpha^{(1)}$ és $\alpha^{(2)}$ ezen becslését $\bar{\bar{\alpha}}^{(1)}$ és $\bar{\bar{\alpha}}^{(2)}$. Nyilván a becsléshez tartozó relatív szórások az optimális kvadratikus becsléshez tartozó relatív szórásoknál nagyobbak lesznek, azonban, mivel az ,,optimális relatív szórás'' ismert, minden egyes esetben eldönthető, hogy kielégítő-e a választott becslés pontossága.

Tekintsük az $\alpha^{(\nu)}$ paraméter

$$\bar{\bar{\alpha}}^{(\nu)} = \mathbf{d}'\mathbf{V}^{(\nu)}\mathbf{d} \qquad \nu = 1,2 \tag{2.1}$$

kvadratikus becslését. A $\mathbf{V}^{(\nu)}$ mátrixokat úgy választjuk, hogy torzítatlan becsléseket kapjunk és a becslések relatív szórása lehetőleg kevéssé térjen el az ,,optimális relatív szórástól''. Vegyünk fel e célból egyelőre tetszőleges $\mathbf{G}^{(\lambda)}$ $(\lambda = 0, 1, \ldots, l)$ alapmátrixokat és a $\mathbf{V}^{(\nu)}$ mátrixokat tekintsük ezek lineáris. kombinációinak:

$$\mathbf{V}^{(\nu)} = \sum_{\lambda=0}^{l} B_{\lambda\nu}\mathbf{G}^{(\lambda)} \qquad \nu = 1, 2 \tag{2.2}$$

Az itt szereplő $B_{\lambda\nu}$ $(\lambda = 0, 1, \ldots, l;\ \nu = 1,2)$ együtthatókat kell tehát úgy meghatározni, hogy a $\mathbf{V}^{(\nu)}$ mátrixokra a fenti feltételek teljesüljenek. Az első rész (1.17), (1.18) és (1.19) képleteinek megfelelően most a következő feladatra jutunk:

A

$$\mathrm{spur}\,(\mathbf{V}^{(\nu)}\mathbf{A}^{(\mu)}) = \delta_{\nu\mu} \qquad \nu, \mu = 1, 2 \tag{2.3}$$

feltételek mellett keressük a $Q_{\nu\nu}$ kifejezésnek a $B_{\lambda\nu}$ értékek szerint vett minimumát, ahol

$$Q_{\nu\mu} = \langle\delta\bar{\bar{\alpha}}^{(\nu)}\delta\bar{\bar{\alpha}}^{(\mu)}\rangle = 2\,\mathrm{spur}\,(\mathbf{V}^{(\nu)}\mathbf{M}\mathbf{V}^{(\mu)}\mathbf{M}) \qquad \nu, \mu = 1, 2\,. \tag{2.4}$$

Bevezetve a

$$T_{\lambda\mu} = \mathrm{spur}\,(\mathbf{G}^{(\lambda)}\mathbf{A}^{(\mu)}) \qquad \lambda = 0, 1, \ldots, l\,;\ \mu = 1, 2 \tag{2.5}$$

jelöléseket, esetünkben a (2.3) feltételek a

$$\sum_{\lambda=0}^{l} B_{\lambda\nu}T_{\lambda\mu} = \delta_{\nu\mu} \qquad \nu, \mu = 1, 2 \tag{2.6}$$

összefüggésekre vezetnek. A feltételes szélsőérték-feladat a minimalizáló $B_{\lambda\nu}$ értékek meghatározására a

(2.7) $$4 \sum_{\mu=0}^{l} \operatorname{spur} (\mathbf{G}^{(\lambda)} \mathbf{M} \mathbf{G}^{(\mu)} \mathbf{M}) B_{\mu\nu} - \sum_{\varkappa=1}^{2} T_{\lambda\varkappa} L_{\varkappa\nu} = 0,$$
$$\lambda = 0, 1, \ldots, l\,; \quad \nu = 1, 2$$

egyenletet adja, ahol $L_{\varkappa\nu}$ a Lagrange-multiplikátorokat jelöli.

Jelölje $\mathbf{T}$ a $T_{\lambda\nu}$ és $\mathbf{B}$ a $B_{\lambda\nu}$ ($\lambda = 0, 1, \ldots l$; $\nu = 1, 2$) elemekből álló mátrixot, vezessük be továbbá az $\mathbf{S} = [S_{\lambda\mu}]$ kvadratikus mátrixot, amelynek elemeit az

(2.8) $$S_{\lambda\mu} = \operatorname{spur} (\mathbf{G}^{(\lambda)} \mathbf{P} \mathbf{G}^{(\mu)} \mathbf{P}) \qquad \lambda, \mu = 0, 1, \ldots, l$$

kifejezések adják (a $\mathbf{P}$ mátrixot az (1.26) összefüggés definiálja).

A most bevezetett jelöléssel a (2.6) feltételek a

(2.9) $$\mathbf{B}'\mathbf{T} = \mathbf{E}$$

mátrixegyenlet alakjában írhatók, ahol $\mathbf{B}'$ a $\mathbf{B}$ mátrix transzponáltját, $\mathbf{E}$ pedig a másodrendű egységmátrixot jelöli.

A (2.8) alatt bevezetett $\mathbf{S}$ mátrix segítségével a $\mathbf{Q}$ kovariancia mátrixra a (2.4) helyett a

(2.10) $$\mathbf{Q} = 2\,\alpha^{(1)^2}\, \mathbf{B}'\mathbf{S}\mathbf{B}$$

összefüggés írható, a (2.7) minimumfeltétel pedig a

(2.11) $$4\,\alpha^{(1)^2}\, \mathbf{S}\mathbf{B} - \mathbf{T}\mathbf{L} = 0$$

mátrixalakot ölti, ahol $\mathbf{L}$ az $L_{\varkappa\nu}$ elemekből álló mátrix. Szorozzuk balról a (2.11) egyenletet a $\mathbf{B}'$ mátrixszal, ekkor a (2.9) összefüggés alapján az $\mathbf{L}$ mátrixra az

(2.12) $$\mathbf{L} = 4\,\alpha^{(1)^2}\, \mathbf{B}'\mathbf{S}\mathbf{B}$$

kifejezést kapjuk. Visszahelyettesítve ezt a (2.11) egyenletbe, az

(2.13) $$\mathbf{S}\mathbf{B} = \mathbf{T}\mathbf{B}'\mathbf{S}\mathbf{B}$$

összefüggésre jutunk. A (2.13) egyenlet már meghatározza a $\mathbf{B}$ mátrixot. Szorozzuk ugyanis (2.13) mindkét oldalát balról az $\mathbf{S}^{-1}$ mátrixszal (a $\mathbf{G}^{(\lambda)}$ mátrixokat úgy kell megválasztani, hogy $\mathbf{S}$ inverze létezzék), akkor:

(2.14) $$\mathbf{B} = \mathbf{S}^{-1}\mathbf{T}\mathbf{B}'\mathbf{S}\mathbf{B}.$$

A kapott (2.14) összefüggést ismét szorozzuk meg balról a $\mathbf{T}'$ mátrixszal (a $\mathbf{T}$ transzponáltjával), akkor (2.9) alapján az

$$\mathbf{E} = (\mathbf{T}'\mathbf{S}^{-1}\mathbf{T})\,(\mathbf{B}'\mathbf{S}\mathbf{B})$$

összefüggésre jutunk. Ha a $\mathbf{G}^{(\lambda)}$ mátrixokat úgy választottuk, hogy a $\mathbf{T}$ oszlopai lineárisan függetlenek legyenek, akkor a másodrendű $\mathbf{T}'\mathbf{S}^{-1}\mathbf{T}$ mátrix invertálható és így

(2.15) $$\mathbf{B}'\mathbf{S}\mathbf{B} = (\mathbf{T}'\mathbf{S}^{-1}\mathbf{T})^{-1}$$

adódik. Behelyettesítve ezt most már a (2.14) ill. (2.10) összefüggésbe, a **B** ill. **Q** mátrixra végül is a

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{T}(\mathbf{T}'\mathbf{S}^{-1}\mathbf{T})^{-1}, \tag{2.16}$$

$$\mathbf{Q} = 2\,\alpha^{(1)2}(\mathbf{T}'\mathbf{S}^{-1}\mathbf{T})^{-1} \tag{2.17}$$

kifejezést kapjuk. Ilyenformán a becslés szórása, a **B** mátrix explicit ismerete nélkül is, csupán a **T** és **S** mátrixokkal kifejezhető.

III. rész

1. A következőkben azzal foglalkozunk, hogy hogyan célszerű a $\mathbf{G}^{(\lambda)}(\lambda = 0, 1, \ldots, l)$ alapmátrixokat úgy megválasztani, hogy ezeknek egy-egy alkalmas lineáris kombinációja az optimális kvadratikus becsléshez tartozó (1.22) alatti $\mathbf{C}^{(\nu)}$ mátrixokat lehetőleg jól közelítse. A közelítés mértékéül természetszerűen a választott becslés relatív szórásának az ,,optimális relatív szórás''-tól való (százalékos) eltérése szolgál. Először tehát azt vizsgáljuk meg, hogy mi jellemzi a $\mathbf{C}^{(\nu)}$ mátrixok szerkezetét. Meg fogjuk mutatni, hogy — nagy rendszám esetén, azaz ha $N \gg 1$ — e mátrixok ,,*kvázi Toeplitz-típusúak*'', vagyis a mátrixok bal felső és jobb alsó sarkának bizonyos környezetétől eltekintve az elemek a főátló, valamint az ezzel párhuzamos ferde sorok mentén *közel* állandóak. Könnyen belátható, hogy a kvázi Toeplitz-típusú mátrixok szorzata is kvázi Toeplitz-típusú, ezért a $\mathbf{C}^{(\nu)}$ mátrixok (1.22) alatti kifejezését, valamint az (1.26), (1.3) és (1.4) jelöléseket figyelembe véve, elegendő azt megmutatni, hogy a **P** mátrix inverze minden pozitív x-re kvázi Toeplitz-típusú. Bontsuk fel ezért a **P** mátrixot a **H** mátrixban lineáris tényezők szorzatára:

$$\mathbf{P} = x\,(\mathbf{H} - z_1\mathbf{E})\,(\mathbf{H} - z_2\mathbf{E}), \tag{3.1}$$

ahol

$$z_{1,2} = \begin{cases} \dfrac{4x - 1 \pm \sqrt{1 - 24x}}{2x} & \text{ha} \quad x < \dfrac{1}{24} \\[2ex] -10 & \text{ha} \quad x = \dfrac{1}{24} \\[2ex] \dfrac{4x - 1 \pm \sqrt{24x - 1}}{2x} & \text{ha} \quad x > \dfrac{1}{24}. \end{cases} \tag{3.2}$$

A **P** mátrix inverzét $x \neq \dfrac{1}{24}$ esetén a

$$[(\mathbf{H} - z_1\mathbf{E})\,(\mathbf{H} - z_2\mathbf{E})]^{-1} \equiv \frac{1}{z_1 - z_2}[(\mathbf{H} - z_1\mathbf{E})^{-1} - (\mathbf{H} - z_2\mathbf{E})^{-1}] \tag{3.3}$$

azonosság felhasználásával számítjuk, $x = \dfrac{1}{24}$ esetén pedig

$$\mathbf{P}^{-1} = 24\,(\mathbf{H} + 10\,\mathbf{E})^{-2} \tag{3.4}$$

a mátrix inverze. Vizsgáljuk először azt az esetet, amikor $z_{1,2}$ valós. Mivel csak pozitív x értékek jönnek tekintetbe, ekkor $0 < x \leqq \frac{1}{24}$ és $z_{1,2} < -4$. Ismeretes [6], hogy az $\mathbf{R} = (\mathbf{H} - 2z\,\mathbf{E})^{-1}$ $(n-1)$-edrendű reciprokmátrix elemei a

$$z = -2\cos\Theta \tag{3.5}$$

transzformációval

$$R_{kl} = \begin{cases} (-1)^{k+l} \dfrac{\sin k\Theta \sin(n-l)\Theta}{\sin\Theta \sin n\Theta} & \text{ha } k \leqq l \\[2ex] (-1)^{k+1} \dfrac{\sin l\Theta \sin(n-k)\Theta}{\sin\Theta \sin n\Theta} & \text{ha } k \geqq l \end{cases} \tag{3.6}$$

alakban írhatók. Innen látható, hogy amennyiben z olyan valós szám, amelyre $z < -2$, akkor Θ tiszta imaginárius: $\Theta = i\,\Psi$, a trigonometrikus függvények helyébe tehát hiperbolikus függvények lépnek. Nagy rendszámú mátrixok esetén a reciprokmátrix elemeire tehát

$$R_{kl} \sim \begin{cases} (-1)^{k+l} \dfrac{\operatorname{sh} k\Psi}{\operatorname{sh}\Psi} e^{-l\Psi} & \text{ha } k \leqq l \\[2ex] (-1)^{k+l} \dfrac{\operatorname{sh} l\Psi}{\operatorname{sh}\Psi} e^{-k\Psi} & \text{ha } k \geqq l \end{cases} \tag{3.7}$$

adódik. Ha a Ψ értékétől függően a k és l indexeket olyan nagyra választjuk, hogy $k\Psi$ és $l\Psi$ hiperbolikus függvényei exponenciális függvénnyel jól közelíthetők, azaz ha a mátrix bal felső és jobb alsó sarkának bizonyos környezetében levő elemektől eltekintünk, akkor

$$R_{kl} \sim (-1)^{|k-l|} \frac{e^{-|k-l|\Psi}}{2\operatorname{sh}\Psi} \tag{3.8}$$

írható. Ez azt jelenti, hogy nagy rendszám esetén az $\mathbf{R}$ reciprokmátrix *kvázi Toeplitz-típusú.* Mivel esetünkben az (3.3) azonosságban szereplő z_1 és z_2 teljesíti a kívánt feltételt (ugyanis $z_{1,2} < -4$), tehát valós z_1, z_2 mellett $\mathbf{P}$ inverze is kvázi Toeplitz-típusú lesz. Felhasználva, hogy két kvázi Toeplitz-típusú mátrix szorzata is kvázi Toeplitz-típusú, a (3.4) összefüggésből látható, hogy a $\mathbf{P}^{-1}$ reciprokmátrix az $x = \frac{1}{24}$ határesetben is *kvázi Toeplitz-típusú.*

A következőkben azt az esetet kell megvizsgálni, amikor z_1 és z_2 konjugált komplexek. A

$$z_1 = a + ib, \quad z_2 = \bar{z}_1$$

jelöléssel a (3.3) azonosság

$$[(\mathbf{H} - z_1\mathbf{E})(\mathbf{H} - \bar{z}_1\mathbf{E})]^{-1} \equiv \frac{\operatorname{Im}(\mathbf{H} - z_1\mathbf{E})^{-1}}{\operatorname{Im} z_1} \tag{3.9}$$

alakban írható. Ahhoz tehát, hogy a $\mathbf{P}^{-1}$ reciprokmátrixról ebben az esetben is belássuk, hogy az kvázi Toeplitz-típusú, elegendő megmutatni, hogy az

Im $(\mathbf{H}-z_1\mathbf{E})^{-1}$ mátrix ilyen. A $(\mathbf{H}-z_1\mathbf{E})^{-1}$ reciprokmátrix elemeit a (3.5) transzformáció alkalmazásával ismét a (3.6) képlet szolgáltatja, most azonban Θ komplex:

$$\Theta = \Phi + i\Psi .$$

Behelyettesítve a (3.6) kifejezésekbe, a reciprokmátrix elemeinek imaginárius része, Im R_{kl} felírható Φ trigonometrikus és Ψ hiperbolikus függvényeinek a segítségével. Figyelembe véve, hogy nagy argumentumú hiperbolikus függvények exponenciális függvénnyel jól közelíthetők, $n \gg 1$ esetén

$$\begin{aligned} \operatorname{Im} R_{kl} &\sim (-1)^{k+l} \frac{e^{-l\Psi}}{\sin^2\Phi \operatorname{ch}^2\Psi + \cos^2\Phi \operatorname{sh}^2\Psi} \times \\ &\times \{\sin\Phi \operatorname{ch}\Psi (\cos l\Phi \cos k\Phi \operatorname{sh} k\Psi + \sin l\Phi \sin k\Phi \operatorname{ch} k\Psi) + \\ &+ \cos\Phi \operatorname{sh}\Psi (\sin l\Phi \cos k\Phi \operatorname{sh} k\Psi - \cos l\Phi \sin k\Phi \operatorname{ch} k\Psi)\} \end{aligned} \tag{3.10}$$

adódik, ha $k \leq l$ ($k \geq l$ esetén k és l szerepet cserél). Ha Ψ értékétől függően a k és l indexeket ismét elég nagynak választjuk, akkor

$$\operatorname{Im} R_{kl} \sim \frac{(-1)^{|l-k|}}{2} \frac{\sin\Phi \cos|l-k|\Phi \operatorname{ch}\Psi + \cos\Phi \sin|l-k|\Phi \operatorname{sh}\Psi}{\sin^2\Phi \operatorname{ch}^2\Psi + \cos^2\Phi \operatorname{sh}^2\Psi} e^{-|l-k|\Psi}$$

adódik. Ebből következik, hogy a $\mathbf{P}^{-1}$ reciprokmátrix komplex $z_{1,2}$ esetén, azaz $x > \frac{1}{24}$ esetén is *kvázi Toeplitz-típusú*. Ezzel megmutattuk, hogy az optimális becsléshez tartozó $\mathbf{C}^{(\nu)}$ mátrixok kvázi Toeplitz-típusúak.

2. A fentiekből önként kínálkozik az az elgondolás, hogy a $\mathbf{G}^{(\lambda)}$ alapmátrixokat Toeplitz-típusúaknak válasszuk. E Toeplitz-típusú alapmátrixok megválasztásánál arra törekszünk, hogy a mérési adatok kiértékelése lehetőleg egyszerű számítások segítségével legyen elvégezhető. Mivel az $\alpha^{(\nu)}$ paramétereket az Y_i mérési adatok második differenciáinak kvadratikus alakjával becsüljük, célszerű a $\mathbf{G}^{(\lambda)}$ alapmátrixokat úgy megválasztani, hogy azok a mérési adatokból különböző módon számított második differenciák *tiszta négyzetösszegével* való becslésnek feleljenek meg:

$$\mathbf{d}'\mathbf{G}^{(\lambda)}\mathbf{d} = \mathbf{d}^{(j\lambda)\prime}\mathbf{d}^{(j\lambda)} \qquad \mathbf{d}^{(j)\prime} = (d_1^{(j)}, d_2^{(j)}, \ldots, d_N^{(j)}), \tag{3.11}$$

ahol $d_i^{(j)}$ $(i = 1, 2, \ldots, N)$ a mérési adatokból, tetszőleges, de előre meghatározott módon számított második differenciákat jelöli. Elméleti és gyakorlati szempontból egyaránt indokolt, hogy a $d_i^{(j)}$ második differenciákat a mérési adatokból a

$$\begin{aligned} &d_i^{(j)} = Y_{i+2^j} - 2Y_i + Y_{i-2^j} \\ &i = 1, 2, \ldots, N; \quad Y_{-k} = Y_{N+1+k} = 0 \quad (k = 1, 2, \ldots, 2^j) \end{aligned} \tag{3.12}$$

képlet alapján számítsuk. Mint látható, e második differenciákat bizonyos mérési adatok átugrásával képezzük, ezért a (3.12) képlett szerint számított második differenciákat *j-átugrásos második differenciáknak* fogjuk nevezni. Nyilvánvalóan $j = 0$ esetén ezek a közönséges második differenciákat adják.

Abból a célból, hogy a (3.12) j-átugrásos második differenciákhoz tartozó (3.11) képlettel definiált $\mathbf{G}$ alapmátrixot meghatározzuk, fel kell írnunk azt a lineáris transzformációt, amely $\mathbf{d}$ és $\mathbf{d}^{(j)}$ között fennáll.

Könnyen belátható, hogy

$$\mathbf{d}^{(j)} = \mathbf{K}^{(j)}\,\mathbf{d}\,,$$

ahol a $\mathbf{K}^{(j)}$ mátrix elemei:

$$K_{ik}^{(j)} = \begin{cases} 2^j - |i-k| & \text{ha} \quad 2^j \geqq |i-k| \\ 0 & \text{ha} \quad 2^j \leqq |i-k|\,. \end{cases} \tag{3.13}$$

Ezzel a (3.11) összefüggés alapján

$$\mathbf{d}'\,\mathbf{G}^{(\lambda)}\,\mathbf{d} = \mathbf{d}'\,(\mathbf{K}^{(j\lambda)})^2\,\mathbf{d}$$

adódik, azaz a j-átugrásos második differenciáknak megfelelő $\mathbf{G}$ alapmátrix

$$\mathbf{G} = (\mathbf{K}^{(j)})^2 \tag{3.14}$$

lesz. A (3.14) összefüggésből látható, hogy a $\mathbf{G}^{(\lambda)}$ mátrix a Toeplitz-típusú $\mathbf{K}^{(j\lambda)}$ mátrix négyzete, tehát *majdnem Toeplitz-típusú.*[4] Ebből következik, hogy felírható az elemi Toeplitz-típusú $\mathbf{H}$ mátrix egy $2(2^{j\lambda}-1)$-edfokú polinomjának és egy „majdnem zérus" $\mathbf{G}_0^{(\lambda)}$ mátrixnak az összegeként, amely tehát csak a bal felső és jobb alsó sarkának bizonyos környezetében tartalmaz zérustól különböző elemeket:

$$\mathbf{G}^{(\lambda)} = P_{2(2^{j\lambda}-1)}(\mathbf{H}) + \mathbf{G}_0^{(\lambda)}. \tag{3.15}$$

A $P_{2(2^{j\lambda}-1)}(\mathbf{H})$ polinomok $j_\lambda = 0, 1, 2, 3$ esetén az alábbiak:

$$\begin{aligned} j_\lambda &= 0 & P_0(\mathbf{H}) &= \mathbf{E} \\ j_\lambda &= 1 & P_2(\mathbf{H}) &= (\mathbf{H} + 2\,\mathbf{E})^2 \\ j_\lambda &= 2 & P_6(\mathbf{H}) &= (\mathbf{H}^3 + 2\,\mathbf{H}^2)^2 \\ j_\lambda &= 3 & P_{14}(\mathbf{H}) &= (\mathbf{H}^7 + 2\,\mathbf{H}^6 - 4\,\mathbf{H}^5 - 8\,\mathbf{H}^4 + 4\,\mathbf{H}^3 + 8\,\mathbf{H}^2)^2. \end{aligned} \tag{3.16}$$

Ezzel tehát meghatároztuk a j-átugrásos második differenciákhoz tartozó alapmátrixokat.

A következőkben olyan becslésekkel foglalkozunk, amikor két, illetőleg három $\mathbf{G}^{(\lambda)}$ alapmátrixot választunk. Abban az esetben, amikor két alapmátrixot választunk, a kvadratikus becslés négy $B_{\lambda\nu}$ paraméterét egyértelműen meghatározzák a becslés torzítatlanságának (2.6) alatti feltételi egyenletei, így ebben az esetben a szórásnégyzetek minimalizálásáról nem beszélhetünk. Ekkor nyilván nem várható, hogy a relatív szórás az „optimális relatív szórás"-t jól közelítse, de mindenesetre érdekes azt vizsgálni, hogy az alapmátrixok különböző megválasztása miként befolyásolja a relatív szórást.

A második részben foglaltak szerint a kvadratikus becslés $B_{\lambda\nu}$ paramétereit és a becslés $\mathbf{Q}$ kovariancia mátrixát a (2.16) és (2.17) képletek alapján

[4] *Majdnem Toeplitz-típusúnak* akkor nevezünk egy mátrixot, ha a jobb felső és bal alsó sarkának bizonyos környezetétől eltekintve a fődiagonálisban, illetve a fődiagonálissal párhuzamos sorokban álló elemek állandók. Belátható, hogy két majdnem Toeplitz-típusú mátrix szorzata is majdnem Toeplitz-típusú.

számíthatjuk. Két alapmátrix választása esetén az ezekben a képletekben előforduló mátrixok *kvadratikus* mátrixok lesznek, tehát az előírt invertálás tényezőnként elvégezhető. Így az alábbi egyszerű összefüggéseket nyerjük:

$$\mathbf{B} = \mathbf{T}'^{-1} \tag{3.17}$$

$$\mathbf{Q} = 2\,\alpha^{(1)^2}\,\mathbf{T}^{-1}\,\mathbf{S}\mathbf{T}'^{-1}\,. \tag{3.18}$$

A következő három esetet fogjuk vizsgálni:

1^0 $\qquad j_0 = 0\,,\quad j_1 = 1\,,$

azaz a becslést a közönséges és az egyátugrásos második differenciák tiszta négyzetösszegével végezzük;

2^0 $\qquad j_0 = 0\,,\quad j_1 = 2\,,$

azaz a becslést a közönséges és a kétátugrásos második differenciák tiszta négyzetösszegével végezzük;

3^0 $\qquad j_0 = 0\,,\quad j_1 = 3\,,$

azaz a becslést a közönséges és a háromátugrásos második differenciák tiszta négyzetösszegével végezzük.

Az egyes esetekben a $\mathbf{T}$, illetve $\mathbf{S}$ mátrixok elemeit a (2.5) és (2.8) összefüggésekből számítjuk. Figyelembe véve a $\mathbf{G}^{(\lambda)}$ alapmátrixok (3.15) alatti előállítását, valamint az $\mathbf{A}^{(\nu)}$ és $\mathbf{P}$ mátrixok (1.3), (1.4), (1.26) alatti kifejezéseit, megállapítható, hogy mind a $\mathbf{T}$ mind az $\mathbf{S}$ mátrix elemei olyan mátrixok spurjaiból számíthatók, amelyek a bal felső és jobb alsó sarkok bizonyos környezetétől eltekintve, az elemi Toeplitz-típusú mátrix polinomjaiként írhatók fel. E polinomok majdnem Toeplitz-típusú mátrixok.

Tekintetbe véve, hogy a $\mathbf{H}$ mátrix páratlan hatványainak diagonálelemei zérussal egyenlők, páros hatványának diagonálelemei pedig, a sarkok bizonyos környezetétől eltekintve, a következők:

$$(H^{2k})_{ii} = \binom{2k}{k} \qquad k=0,1,2,\ldots,$$

tehát az elemi Toeplitz-típusú mátrix tetszőleges polinomjának a spurját a

$$\operatorname{spur}\left(\sum_{j=0}^{n} a_j\,\mathbf{H}^j\right) = n \sum_{k=0}^{\left[\frac{n}{2}\right]} a_{2k}\binom{2k}{k} + O(1)$$

képlet szolgáltatja. Könnyen belátható, hogy az a hiba, amelyet akkor követünk el, ha a keresett spurokat a sarkok bizonyos környezetében levő elemek módosításával Toeplitz-típusúvá kiegészített mátrixokból számítjuk, $\frac{1}{n}$ nagyságrendű lesz. Mivel tárgyalásaink során feltettük, hogy $N \gg 1$, az eddigiekhez hasonlóan az alábbiakban közölt eredményeket is aszimptotikusaknak kell tekintenünk.

A fentiek szerint az említett három esetben a következő eredményeket nyerjük:

1° Az alapmátrixok:

$$\mathbf{G}^{(0)} = \mathbf{E}\,, \quad \mathbf{G}^{(1)} = \mathbf{H}^2 + 4\,\mathbf{H} + 4\,\mathbf{E}\,.$$

A **T** mátrix elemei:

$$T_{01} = \operatorname{spur}(\mathbf{G}^{(0)}\mathbf{A}^{(1)}) = \operatorname{spur}(\mathbf{H} + 4\,\mathbf{E}) = 4\,N\,,$$

$$T_{02} = \operatorname{spur}(\mathbf{G}^{(0)}\mathbf{A}^{(2)}) = \operatorname{spur}(\mathbf{H}^2 - 4\,\mathbf{H} + 4\,\mathbf{E}) = 6\,N\,,$$

$$T_{11} = \operatorname{spur}(\mathbf{G}^{(1)}\mathbf{A}^{(1)}) = \operatorname{spur}(\mathbf{H}^3 + 8\,\mathbf{H}^2 + 20\,\mathbf{H} + 16\,\mathbf{E}) = 32\,N\,,$$

$$T_{12} = \operatorname{spur}(\mathbf{G}^{(1)}\mathbf{A}^{(2)}) = \operatorname{spur}(\mathbf{H}^4 - 8\,\mathbf{H}^2 + 16\,\mathbf{E}) = 6\,N\,,$$

tehát

$$\mathbf{T} = N\begin{bmatrix} 4 & 6 \\ 32 & 6 \end{bmatrix}.$$

Az **S** mátrix elemei:

$$\begin{aligned} S_{00} &= \operatorname{spur}(\mathbf{G}^{(0)}\mathbf{P}\mathbf{G}^{(0)}\mathbf{P}) = \operatorname{spur}\{[x\,\mathbf{H}^2 + (1 - 4\,x)\,\mathbf{H} + 4(1 + x)\,\mathbf{E}]^2\} = \\ &= (70\,x^2 + 32\,x + 18)\,N\,, \\ S_{01} &= S_{10} = \operatorname{spur}(\mathbf{G}^{(0)}\mathbf{P}\mathbf{G}^{(1)}\mathbf{P}) = \\ &\qquad = \operatorname{spur}\{[\mathbf{H}^2 + 4\,\mathbf{H} + 4\,\mathbf{E}]\,[x\,\mathbf{H}^2 + (1 - 4\,x)\,\mathbf{H} + 4(1 + x)\,\mathbf{E}]^2\} = \\ &\qquad = (28\,x^2 + 48\,x + 174)\,N, \\ S_{11} &= \operatorname{spur}(\mathbf{G}^{(1)}\mathbf{P}\mathbf{G}^{(1)}\mathbf{P}) = \\ &= \operatorname{spur}\{[\mathbf{H}^2 + 4\,\mathbf{H} + 4\,\mathbf{E}]^2\,[x\,\mathbf{H}^2 + (1 - 4\,x)\,\mathbf{H} + 4(1 + x)\,\mathbf{E}]^2\} = \\ &= (70\,x^2 + 256\,x + 2\,212)\,N\,, \end{aligned}$$

tehát

$$\mathbf{S} = N\begin{bmatrix} 70\,x^2 + 32\,x + 18 & 28\,x^2 + 48\,x + 174 \\ 28\,x^2 + 48\,x + 174 & 70\,x^2 + 256\,x + 2212 \end{bmatrix}.$$

A (3.17) és (3.18) képletekből a becslés $B_{\lambda\nu}$ paramétereit tartalmazó **B** mátrix és a **Q** kovariancia mátrix most már közvetlenül számítható:

$$\mathbf{B} = \frac{1}{N}\,\frac{1}{84}\begin{bmatrix} -3 & 16 \\ 3 & -2 \end{bmatrix}, \tag{3.19}$$

$$\mathbf{Q} = \frac{2\,\alpha^{(1)2}}{N}\,\frac{1}{3528}\begin{bmatrix} 378\,x^2 + 864\,x + 8469 & -(1134\,x^2 + 240\,x + 2370) \\ -(1134\,x^2 + 240\,x + 2370) & 8204\,x^2 + 3072\,x + 1160 \end{bmatrix}. \tag{3.20}$$

Vezessük be a j-átugrásos második differenciák négyzeteinek átlagára a

$$\bar{d}^{(j)} = \frac{1}{N}\sum_{i=1}^{N} d_i^{(j)2} \tag{3.21}$$

jelölést, akkor $\alpha^{(\nu)}$ becslésére (3.19) alapján az

$$\bar{\bar{\alpha}}^{(1)} = \frac{1}{28}(-\tilde{d}^{(0)} + \tilde{d}^{(1)})$$

(3.22)

$$\bar{\bar{\alpha}}^{(2)} = \frac{1}{42}(8\,\tilde{d}^{(0)} - \tilde{d}^{(1)})$$

kifejezéseket nyerjük. Az $\bar{\bar{\alpha}}^{(1)}$ becslés relatív szórásnégyzetét (3.20) szerint a

$$q_{11}^{(01)}(x) = \frac{1}{392}(42\,x^2 + 96\,x + 941) \tag{3.23}$$

függvény, továbbá I. részben bevezetett (1.39), (1.40) jelölésekkel, az

$$r^{(01)}(z) = \frac{q_{11}^{(01)}(z^3)}{z} = \frac{1}{392}\left(42\,z^5 + 96\,z^2 + 941\,\frac{1}{z}\right) \qquad (z^3 = x) \tag{3.24}$$

függvény segítségével számíthatjuk, ha x helyére az

$$\bar{\bar{x}} = \frac{\bar{\bar{\alpha}}^{(2)}}{\bar{\bar{\alpha}}^{(1)}} = \frac{2}{3}\,\frac{8\,\tilde{d}^{(0)} + \tilde{d}^{(1)}}{-\tilde{d}^{(0)} + \tilde{d}^{(1)}} \tag{3.25}$$

értéket helyettesítjük be.

Hasonló módon határozható meg a **B** és a **Q** mátrix a másik két esetben. A számítások mellőzésével csupán a végeredményeket közöljük.

2° Az alapmátrixok:

$$\mathbf{G}^{(0)} = \mathbf{E}\,, \quad \mathbf{G}^{(1)} = \mathbf{H}^6 + 4\,\mathbf{H}^5 + 4\,\mathbf{H}^4\,.$$

A becslés paramétereit tartalmazó **B** mátrix:

$$\mathbf{B} = \frac{1}{N}\,\frac{1}{756}\begin{bmatrix} -3 & 128 \\ 3 & -2 \end{bmatrix}.$$

A kovariancia mátrix:

$$\mathbf{Q} = \frac{2\,\alpha^{(1)^2}}{N}\,\frac{1}{285768}\times$$

$$\times\begin{bmatrix} 306\,x^2 + 8928\,x + 1259055 & -(6630\,x^2 + 2928\,x + 560406) \\ -(6630\,x^2 + 2928\,x + 560406) & 564364\,x^2 + 253952\,x + 330512 \end{bmatrix}.$$

Innen $\alpha^{(\nu)}$ becslésére az

$$\bar{\bar{\alpha}}^{(1)} = \frac{1}{252}(-\tilde{d}^{(0)} + \tilde{d}^{(2)})$$

(3.26)

$$\bar{\bar{\alpha}}^{(2)} = \frac{1}{378}(64\,\tilde{d}^{(0)} - \tilde{d}^{(2)})$$

kifejezéseket kapjuk. Az $\bar{\bar{\alpha}}^{(1)}$ becslés relatív szórásnégyzetét most a

$$q_{11}^{(02)}(x) = \frac{1}{31752}(34x^2 + 992x + 139895), \tag{3.27}$$

illetve az

$$r^{(02)}(z) = \frac{q_{11}^{(02)}(z^3)}{z} = \frac{1}{31752}\left(34z^5 + 992z^2 + 139895\frac{1}{z}\right) \qquad (z^3 = x) \tag{3.28}$$

függvény segítségével számíthatjuk, ha x helyére most a

$$\bar{\bar{x}} = \frac{\bar{\bar{\alpha}}^{(2)}}{\bar{\bar{\alpha}}^{(1)}} = \frac{2}{3}\,\frac{64\tilde{d}^{(0)} - \tilde{d}^{(2)}}{-\tilde{d}^{(0)} + \tilde{d}^{(2)}} \tag{3.29}$$

értéket helyettesítjük.

3° Az alapmátrixok:

$$\mathbf{G}^{(0)} = \mathbf{E}, \quad \mathbf{G}^{(1)} = \mathbf{H}^{14} + 4\mathbf{H}^{13} - 4\mathbf{H}^{12} - 32\mathbf{H}^{11} - 8\mathbf{H}^{10} + 96\mathbf{H}^{9} +$$
$$+ 64\mathbf{H}^{8} - 128\mathbf{H}^{7} - 112\mathbf{H}^{6} + 64\mathbf{H}^{5} + 64\mathbf{H}^{4}.$$

A $\mathbf{B}$ mátrixra

$$\mathbf{B} = \frac{1}{N}\,\frac{1}{6132}\begin{bmatrix} -3 & 1024 \\ 3 & -2 \end{bmatrix}$$

adódik. A kovariancia mátrix pedig

$$\mathbf{Q} = \frac{2\,\alpha^{(1)^2}}{N}\,\frac{1}{18800712}\times$$
$$\times\begin{bmatrix} 306x^2+73440x+162738963 & -(52326x^2+24432x+87981102) \\ -(52326x^2+24432x+87981102 & 36626572x^2+16711680x+54380064 \end{bmatrix}$$

lesz. Innen $\alpha^{(\nu)}$ becslésére

$$\begin{aligned} \bar{\bar{\alpha}}^{(1)} &= \frac{1}{2044}(-\tilde{d}^{(0)} + \tilde{d}^{(3)}) \\ \bar{\bar{\alpha}}^{(2)} &= \frac{1}{3066}(512\tilde{d}^{(0)} - \tilde{d}^{(3)}) \end{aligned} \tag{3.30}$$

kifejezéseket kapjuk. Az $\bar{\bar{\alpha}}^{(1)}$ becslés relatív szórásnégyzetét itt a

$$q_{11}^{(03)}(x) = \frac{1}{2088968}(34x^2 + 8160x + 18082107), \tag{3.31}$$

illetve az

$$r^{(03)}(z) = \frac{q_{11}^{(03)}(z^3)}{z} = \frac{1}{2088968}\left(34z^5 + 8160z^2 + 18082107\frac{1}{z}\right), \qquad (z^3 = x) \tag{3.32}$$

függvény segítségével számíthatjuk, ha x helyére az

$$\bar{\bar{x}} = \frac{\bar{\bar{\alpha}}^{(2)}}{\bar{\bar{\alpha}}^{(1)}} = \frac{2}{3}\,\frac{512\tilde{d}^{(0)} - \tilde{d}^{(3)}}{-\tilde{d}^{(0)} + \tilde{d}^{(3)}}$$

értéket helyettesítjük.

Az $\bar{\bar{\alpha}}^{(1)}$ becslés relatív szórását jellemző $q_{11}^{(0j)}(x)$ függvények görbéit, amelyek az 1°, 2° és 3° esetnek megfelelően a (3.23), (3.27) illetve (3.31) egyenletű parabolák, a 2. ábrán tüntettük fel. Az ábra egyszersmind az optimális kvadratikus becsléshez tartozó $q_{11}(x)$ függvény görbéjét is feltünteti. Az ábrából látható, hogy az 1° és 2° becsléseknek megfelelő parabolák metszéspontjának $x^{(1,2)} = 3{,}4560$ abszcisszája, illetve a 2° és 3° becslésnek megfelelő parabolák metszéspontjának $x^{(2,3)} = 51{,}8336$ abszcisszája az x tengelyt három intervallumra osztja. Nyilvánvaló, hogy $x < x^{(1,2)}$ esetén az 1°, $x^{(1,2)} < x < x^{(2,3)}$ esetén a 2°, $x > x^{(2,3)}$ esetén pedig a 3° becslés közelíti meg legjobban az optimális kvadratikus becslést. Nevezzük ezeket az intervallumokat az 1°, 2°, illetve 3° becslés számára ,,jó'' intervallumnak. Ebből következik, hogy amennyiben az 1°, 2°, illetve 3° becsléssel számított (3.25), (3.29) illetve (3.33) $\bar{\bar{x}}$ érték az illető becslés számára jó intervallumba esik, akkor az $\alpha^{(1)}$ és $\alpha^{(2)}$ paraméterek becslésére kapott $\bar{\bar{\alpha}}^{(1)}$ és $\bar{\bar{\alpha}}^{(2)}$ értékeket mint ,,jó'' becsléseket elfogadhatjuk, ha azonban $\bar{\bar{x}}$ egy másik becslés számára jó intervallumba esik, akkor a becslést ennek megfelelően kell megismételni.

Az $\bar{\bar{\alpha}}^{(1)}$ becslés *relatív* szórását változó N esetén jellemző, — az 1°, 2° és 3° esetnek megfelelő (3.24), (3.28) illetve (3.32) egyenletű — $r^{(0j)}(z)$ függvények görbéit a 3. ábra tünteti fel. Ugyanitt ábrázoltuk az optimális kvadratikus becsléshez tartozó $r(z)$ függvényt is. Az ábrából látható, hogy az 1°, 2° és 3° becslésekhez olyan $r^{(0j)}(z)$ függvények tartoznak, amelyeknek jól definiált pozitív $z_0^{(0j)}$ helyeken minimumuk van. Tekintettel arra, hogy z arányos a mérések számával, $N+1$-gyel, ez annyit jelent, hogy — az optimális kvadratikus becsléstől eltérően — az 1°, 2° illetve 3° becslés esetén a mérési pontok számának növelésével $\bar{\bar{\alpha}}^{(1)}$ relatív szórása csak egy bizonyos határig csökken, azon túl ismét növekszik. A $z_0^{(0j)}$ minimumhelyek az egyes becsléseknél a következők:

$$1^\circ \quad z_0^{(01)} = 1{,}1956\,,$$

$$2^\circ \quad z_0^{(02)} = 2{,}8618\,,$$

$$3^\circ \quad z_0^{(03)} = 6{,}5550\,.$$

A 2. ábrából látható, hogy az $\bar{\bar{\alpha}}^{(1)}$ becslés szórását jellemző $q_{11}^{(0j)}(x)$ függvény általában nem közelíti meg elég jól az optimális kvadratikus becsléshez tartozó $q_{11}(x)$ függvényt. A közelítés mértékéül nyilvánvalóan a $q_{11}^{(0j)}(x)$ függvénynek az optimális kvadratikus becsléshez tartozó $q_{11}(x)$ függvénytől való százalékos eltérése, azaz

$$h^{(0j)}(x) = \frac{q_{11}^{(0j)}(x) - q_{11}(x)}{q_{11}(x)}\, 100\ \%$$

szolgál. A $h^{(0j)}(x)$ függvényeket szakaszosan, — csak a megfelelő jósági intervallumokban — ábrázolva a 7. ábrán tüntettük fel.

3. A következőkben azzal az esettel foglalkozunk, amikor a $\mathbf{G}^{(\lambda)}$ alapmátrixokat a közönséges, az egyátugrásos és a kétátugrásos második differenciák tiszta négyzetösszegének megfelelően választjuk. A (3.15) és (3.16) képletek felhasználásával tehát, $j_\lambda = \lambda$ ($\lambda = 0, 1, 2$) választással az alapmátrixok a következők:

$$(3.34)\quad \mathbf{G}^{(0)} = \mathbf{E}\,, \quad \mathbf{G}^{(1)} = \mathbf{H}^2 + 4\,\mathbf{H} + 4\,\mathbf{E}\,, \quad \mathbf{G}^{(2)} = \mathbf{H}^6 + 4\,\mathbf{H}^5 + 4\,\mathbf{H}^4\,.$$

A (2.5) illetve (2.8) összefüggéssel definiált **T** és **S** mátrixra ebben az esetben

$$\mathbf{T} = N\begin{bmatrix} 4 & 6 \\ 32 & 6 \\ 256 & 6 \end{bmatrix} \tag{3.35}$$

$$\mathbf{S} = N\begin{bmatrix} 70x^2+32x+18 & 29x^2+48x+174 & 36x^2+48x+1494 \\ 28x^2+48x+174 & 70x^2+256x+2212 & 28x^2+384x+22188 \\ 36x^2+48x+1494 & 28x^2+384x+22188 & 70x^2+2048x+282760 \end{bmatrix} \tag{3.36}$$

adódik.

A **Q** kovariancia mátrix és a becslés $B_{\lambda\nu}$ paramétereiből alkotott **B** mátrix meghatározásához, amint az a (2.16), (2.17) képletekből látható, a háromtényezős $\mathbf{T}'\mathbf{S}^{-1}\mathbf{T}$ mátrix-szorzat invertálását kell elvégezni. Tekintettel arra, hogy a **T** mátrix ebben az esetben *nem kvadratikus*, az invertálást nem lehet tényezőnként elvégezni, indokolt azonban az a törekvés, hogy a keresett reciprokmátrixot ismét közvetlenül az **S** mátrix segítségével határozzuk meg. Esetünkben ez márcsak azért is igen célszerű lenne, mert az **S** mátrix elemei polinomok, ami az invertálást nehézkessé teszi, a bonyolult végeredményből pedig nem volna könnyű felismerni az egyszerűsítés lehetőségét.

A $\mathbf{T}'\mathbf{S}^{-1}\mathbf{T}$ mátrix invertálásának leglényegesebb gondolata abban áll, hogy a téglalap alakú **T** mátrixot egy harmadik oszloppal nemszinguláris harmadrendű mátrixszá egészítjük ki. Ennek lehetőségét az biztosítja, hogy a **T** mátrix két oszlopa lineárisan független. Jelölje a kiegészített **T** mátrixot $\hat{\mathbf{T}}$. Nyilvánvaló, hogy $\mathbf{T}'\mathbf{S}^{-1}\mathbf{T}$ a kiegészített $\hat{\mathbf{T}}$ mátrixszal számított $\mathbf{Z}=\hat{\mathbf{T}}'\mathbf{S}^{-1}\hat{\mathbf{T}}$ mátrixnak az a minormátrixa, amelyet a harmadik sor és oszlop elhagyásával nyerünk. A feladatot tehát a **Z** mátrix egy minor-mátrixának az invertálására vezettük vissza. Ismeretes azonban (lásd pl. [6]), hogy egy minormátrix inverze az adott mátrix inverzének ismeretében úgy számítható, hogy az adott mátrix ismert inverzéből egy alkalmasan választott diádot levonunk, s a kapott mátrixnak a diád levonása által triviálisan csupa zérust tartalmazó sorát és oszlopát elhagyjuk.[5]

Esetünkben a

$$\mathbf{Z}^{-1} = \hat{\mathbf{T}}^{-1}\mathbf{S}\hat{\mathbf{T}}'^{-1}$$

mátrixból kell levonni ennek harmadik oszlopából s harmadik sorából alkotott, és a mátrix (zérustól különböző) jobb alsó sarokelemének reciprokával szorzott diádot.

Ha **u** jelöli a kiegészített $\hat{\mathbf{T}}'$ mátrix inverzének harmadik oszlopát, akkor ez a levonandó diád

$$\hat{\mathbf{T}}^{-1}\mathbf{Su}(\mathbf{u}'\mathbf{Su})^{-1}\mathbf{u}'\mathbf{S}\hat{\mathbf{T}}'^{-1}$$

[5] Megjegyezzük, hogy ez az eljárás tetszőleges rendszámú mátrixok esetén alkalmazható, az egyetlen feltétel természetesen az, hogy a **T** mátrix oszlopai lineárisan független vektorok legyenek. Ha **S** n-edrendű, **T** pedig r oszlopos mátrix, és **S** elemei polinomok, akkor a fenti három tényezős mátrix-szorzat invertálásához, polinomot tartalmazó n-edrendű és r-edrendű mátrixok invertálása helyett csupán egyetlen, $(n-r)$-edrendű (polinomot tartalmazó) mátrix invertálását kell elvégezni.

tehát a $\mathbf{Z}$ mátrixból a harmadik sor és oszlop elhagyásával adódó minormátrix keresett inverzét a

$$(3.37) \qquad \hat{\mathbf{T}}^{-1}\mathbf{S}\hat{\mathbf{T}}'^{-1} - \hat{\mathbf{T}}^{-1}\mathbf{Su}(\mathbf{u}'\mathbf{Su})^{-1}\mathbf{u}'\mathbf{S}\hat{\mathbf{T}}'^{-1}$$

mátrixnak a csupa zérust tartalmazó, harmadik sor és oszlop elhagyásával kapott minormátrixa adja. Kiemelve balra a $\hat{\mathbf{T}}^{-1}$, jobbra a $\hat{\mathbf{T}}'^{-1}$ mátrixokat, a $\hat{\mathbf{T}}'^{-1}$ mátrix első két oszlopából alkotott minormátrixot pedig $\mathbf{U}$-val jelölve, a keresett reciprokmátrixra a

$$(3.38) \qquad (\mathbf{T}'\mathbf{S}^{-1}\mathbf{T})^{-1} = \mathbf{U}'\{\mathbf{S} - \mathbf{Su}(\mathbf{u}'\mathbf{Su})^{-1}\mathbf{u}'\mathbf{S}\}\,\mathbf{U}$$

kifejezés adódik. Ezzel tehát elértük azt, hogy a keresett reciprokmátrix meghatározásához nem szükséges a kellemetlen polinom-elemeket tartalmazó $\mathbf{S}$ mátrixot invertálni, csupán az $\mathbf{u}'\mathbf{Su}$ egyetlen elemet alkotó kifejezést. Esetünkben, amikor az $\mathbf{S}$ mátrix elemei másodfokú polinomok, látható, hogy a keresett reciprokmátrix elemei negyedfokú és másodfokú polinomok hányadosai lesznek. (Ha a $\mathbf{T}'\mathbf{S}^{-1}\mathbf{T}$ szorzatot közvetlenül invertáltuk volna, akkor ezen elemek egy tizedfokú és egy nyolcadfokú polinom hányadosaként adódtak volna, amelyek egyszerűsíthetősége nem nyilvánvaló.) A (3.37) képlet felhasználásával, figyelembevéve, hogy e mátrix harmadik sora és oszlopa csupa zérus elemet tartalmaz, a $\mathbf{B}$ mátrixra

$$(3.39) \qquad \mathbf{B} = \mathbf{U} - \mathbf{u}(\mathbf{u}'\mathbf{Su})^{-1}\mathbf{u}'\mathbf{SU}$$

adódik.

Esetünkben a (3.35) mátrixot a következőképpen egészítjük ki nemszinguláris $\hat{\mathbf{T}}$ mátrixszá:

$$\hat{\mathbf{T}} = N\begin{bmatrix} 4 & 6 & 0 \\ 32 & 6 & 0 \\ 256 & 6 & 1 \end{bmatrix}.$$

Innen

$$\hat{\mathbf{T}}'^{-1} = \frac{1}{84N}\begin{bmatrix} -3 & 16 & 672 \\ 3 & -2 & -756 \\ 0 & 0 & 84 \end{bmatrix}.$$

Tehát a képleteinkben szereplő $\mathbf{U}$ és $\mathbf{u}$ mátrixokra

$$\mathbf{U} = \frac{1}{84N}\begin{bmatrix} -3 & 16 \\ 3 & -2 \\ 0 & 0 \end{bmatrix}; \quad \mathbf{u} = \frac{1}{84N}\begin{bmatrix} 672 \\ -756 \\ 34 \end{bmatrix} = \frac{1}{N}\begin{bmatrix} 8 \\ -9 \\ 1 \end{bmatrix}$$

adódik. Behelyettesítve a (3.38) és (3.39) képletekbe, végül a következő eredményre jutunk:

$$\mathbf{B} = \frac{1}{N}\begin{bmatrix} \dfrac{-(121x^2+128x+22837)}{28(1565x^2+2944x+15637)} & \dfrac{2(1565x^2+5632x+32822)}{21(1565x^2+2944x+15637)} \\ \dfrac{-(119x^2+448x-47474)}{56(1565x^2+2944x+15637)} & \dfrac{(1565x^2-512x-6458)}{12(1565x^2+2944x+15637)} \\ \dfrac{(361x^2+704x-1800)}{56(1565x^2+2944x+15637)} & \dfrac{-(1565x^2+256x-1548)}{84(1565x^2+2944x+15637)} \end{bmatrix} \tag{3.40}$$

és

$$\mathbf{Q} = \frac{2\alpha^{(1)^2}}{N} \times \tag{3.41}$$

$$\times \begin{bmatrix} \dfrac{1139x^4+39488x^3+5628070x^2+11077312x+26188834}{784(1565x^2+2944x+15637)} & q_{12}^{(012)}(x) \\ q_{12}^{(012)}(x) & q_{22}^{(012)}(x) \end{bmatrix}.$$

(A mátrix $q_{12}^{(012)}(x)$ és $q_{22}^{(012)}(x)$ elemeit, — amelyek szintén racionális törtfüggvényei az x-nek —, ezúttal nem írtuk ki.) A (3.40) képletből látható, hogy a becslés paramétereire nem állandó értékeket, hanem függvényeket kaptunk. Ezeket a 4. és 5. ábrán tüntettük fel. A független változó az (1.25) összefüggés szerint éppen azoknak a paramétereknek a hányadosa, amelyeket becsülni akarunk. A keresett becslésekhez most a következő meggondolással juthatunk. Az $\bar{\bar{\alpha}}^{(\nu)}$ becsléseket először felírjuk, mint a j-átugrásos második differenciák négyzetátlagának a $B_{\lambda\nu}(x)$ becslési paraméterekkel képzett lineáris kombinációját:

$$\bar{\bar{\alpha}}^{(1)}(x) = \sum_{\lambda=0}^{2} B_{\lambda 1}(x)\, \tilde{d}^{(\lambda)} \tag{3.42}$$

$$\bar{\bar{\alpha}}^{(2)}(x) = \sum_{\lambda=0}^{2} B_{\lambda 2}(x)\, \tilde{d}^{(\lambda)}. \tag{3.43}$$

Ha képezzük a két becslés hányadosát, és x helyére az $\bar{\bar{x}} = \dfrac{\bar{\bar{\alpha}}^{(2)}}{\bar{\bar{\alpha}}^{(1)}}$ kifejezést írjuk, akkor nyilvánvaló, hogy a hányadosnak az az értéke, amelyik $\bar{\bar{x}}$-szel megegyezik, szolgáltatja azt az $\bar{\bar{x}}_0$ értéket, amelyből a (3.42) és (3.43) képletek segítségével a keresett becslések kiszámíthatók. Tehát a

$$\bar{\bar{x}} = \frac{\sum_{\lambda=0}^{2} B_{\lambda 2}(\bar{\bar{x}})\, \tilde{d}^{(\lambda)}}{\sum_{\lambda=0}^{2} B_{\lambda 1}(\bar{\bar{x}})\, \tilde{d}^{(\lambda)}} \tag{3.44}$$

egyenlet $\bar{\bar{x}}_0$ gyökével az $\bar{\bar{\alpha}}^{(\nu)}$ paraméterek becslésére

$$\bar{\bar{\alpha}}^{(\nu)} = \sum_{\lambda=0}^{2} B_{\lambda\nu}(\bar{\bar{x}}_0)\, \tilde{d}^{(\lambda)} \qquad \nu = 1,2$$

adódik. Az $\bar{\bar{\alpha}}^{(1)}$ becslés relatív szórásnégyzete a (3.41) kifejezés szerint a

$$(3.45)\quad q_{11}^{(012)}(x) = \frac{1139\,x^4 + 39488\,x^3 + 5628070\,x^2 + 11077312\,x + 26188834}{784\,(1565\,x^2 + 2944\,x + 15637)},$$

illetve az

$$(3.46)\quad r^{(012)}(z) = \frac{q_{11}^{(012)}(z^3)}{z} =$$

$$= \frac{1139\,z^{12} + 39488\,z^9 + 5628070\,z^6 + 11077312\,z^3 + 26188834}{784\,(1565\,z^7 + 2944\,z^4 + 15637\,z)} \qquad (z^3 = x)$$

függvény segítségével számítható, ha x helyére az $\bar{\bar{x}}_0$ értéket helyettesítjük be. A (3.45) függvényt a 2. ábra tünteti fel. Látható az ábrából, hogy $x < 54{,}3766$ esetén e függvény görbéje jobban közelíti az optimális kvadratikus becsléshez tartozó $q_{11}(x)$ függvény görbéjét, mint a két alapmátrixszal számított becsléshez tartozó bármelyik $q_{11}^{(0j)}(x)$ függvény, $x > 54{,}3766$ esetén azonban a $q_{11}^{(03)}(x)$ függvény görbéje közelíti jobban a $q_{11}(x)$ függvény görbéjét. Abban a — gyakorlatilag kevésbé érdekes — esetben tehát, amikor a háttér zaja a jelhez képest igen nagy, akkor a két alapmátrixszal számított 3° becslés jobb a három alapmátrixszal számított becslésnél.

A közelítés mértékéül nyilvánvalóan a $q_{11}^{(012)}(x)$ függvénynek az optimális kvadratikus becsléshez tartozó $q_{11}(x)$ függvénytől való százalékos eltérése, azaz

$$h^{(012)}(x) = \frac{q_{11}^{(012)}(x) - q_{11}(x)}{q_{11}(x)}\,100\%$$

szolgál. (A $h^{(012)}(x)$ függvény görbéje a 7. ábrán látható.)

A (3.46) függvény görbéjét a 3. ábrán tüntettük fel. Innen látható, hogy az $r^{(012)}(z)$ függvénynek is van minimuma, mégpedig a

$$(3.47)\quad z_0^{(012)} = 2{,}9088$$

helyen. Az $r^{(012)}(z)$ függvény minimuma itt

$$r_{\min}^{(012)} = 1{,}9954\,.$$

Ez azt jelenti, hogy a három alapmátrixszal végzett becslésnél a legkisebb relatív szórás a (3.47) kifejezésnek megfelelő N_0 mérés-szám esetén érhető el. A 3. ábrából az is látható, hogy a $z_0^{(012)}$ minimumhely a két alapmátrixszal végzett 1° és 2° becslésnek megfelelő $r^{(0j)}(z)$ függvények $z_0^{(0j)}$ minimumhelyénél nagyobb, az $r^{(012)}(z)$ függvény $r_{\min}^{(012)}$ minimuma pedig e két $r^{(0j)}(z)$ függvény minimumánál kisebb. A 3° becslésnek megfelelő $r^{(03)}(z)$ függvény minimumhelye azonban nagyobb, a minimuma pedig kisebb, mint az $r^{(012)}(z)$ függvényé, ami azt jelenti, hogy megfelelő számú mérés esetén a 3° becsléssel kisebb relatív szórás érhető el, mint a három alapmátrixszal végzett becsléssel.

A három alapmátrixszal végzett becslés legfőbb nehézsége a (3.44) egyenlet numerikus megoldásából áll. A (3.44) egyenletet az egyszerűbb

$$(3.48)\quad \sum_{\lambda=0}^{2} \tilde{d}^{(\lambda)}\, g_\lambda(x) = 0$$

alakban írjuk, ahol

$$g_0(x) = -(0{,}0726\,x^3 + 2{,}5808\,x^2 + 22{,}7134\,x + 52{,}5152),$$

$$g_1(x) = -0{,}0357\,x^3 - 2{,}3254\,x^2 + 14{,}9590\,x + 9{,}0412,$$

$$g_2(x) = 0{,}1083\,x^3 + 0{,}5242\,x^2 - 0{,}4888\,x - 0{,}3096.$$

(A függvények görbéi a 6. ábrán láthatók.) E harmadfokú egyenlet keresett gyökének első közelítéseként a (3.25), (3.29) vagy (3.33) képletből számított $\bar{\bar{x}}$ értéket tekinthetjük, amelyből kiindulva pl. a Newton-módszerrel néhány iterációs lépésben eljuthatunk a keresett gyökhöz.

A 2. ábrán vázolt függvények vizsgálata azt mutatja, hogy három alapmátrix esetén a gyakorlatban szóba jöhető $\frac{\alpha^{(2)}}{\alpha^{(1)}}$ értékek esetén az $\alpha^{(1)}$ jel-paraméter becslésének szórása kielégítő pontossággal közelíti az optimális kvadratikus becslés szórását. Amennyiben a közelítés mértéke nem volna kielégítő, a kívánt pontosság mértéke az alapmátrixok megfelelő más megválasztásával, számuk növelésével azonban mindenesetre tovább fokozható.

(Beérkezett: 1961. XII. 1.)

IRODALOM

[1] JÁNOSSY, L.: „On the determination of the energy of a particle from its track in an emulsion". *Acta Physica Acad. Sci. Hung.* **7** (1957) 385—401.
[2] JÁNOSSY, L.: „On the simultaneous distribution of the sagittas of a track in emulsion in the case of background noise". *Acta Physica Acad. Sci. Hung.* **12** (1960) 139—150.
[3] SOLNTSEFF, N.: „On the theory of scattering measurements in nuclear emulsion". *Nuclear Physics* **6** (1958) 222—251.
[4] JÁNOSSY, L.—RÓZSA, P.: „Maximum likelihood determination of the scattering constant of an emulsion track in the presence of noise". *Il Nuovo Cimento* Serie X, Vol. **20** (1961) 817—835.
[5] PASCAL, E.: *Die Determinanten.* Teubner, Leipzig, 1900.
[6] EGERVÁRY, E.: „Über eine Methode zur numerischen Lösung der Poissonschen Differenzengleichung für beliebige Gebiete". *Acta Mathematica Acad. Sci. Hung.* **11** (1960) 341—361.

ОЦЕНКА ПАРАМЕТРА РАССЕИВАНИЯ КУЛОНА НА ОСНОВАНИИ ИЗМЕРЕНИЙ, ВЫПОЛНЕННЫХ В ФОТОЭМУЛЬСИИ

L. JÁNOSSY — A. LEE — P. RÓZSA

Резюме

В настоящей статье речь идет об оценке параметра рассеивания Кулона частиц с большой энергией на основании измерений в фотоэмульсии, с учетом ошибок измерений. В I. части авторы дают оценку параметра рассеивания при помощи метода наибольшего правоподобия.

Из-за дальнейших приложений авторы дадут оценку с помошью такой квадратичной формы вторых разносрей, которая является несмещенной оценкой с минимальной дисперсией. Они называют эту оценку *оптимально* *квадратичной* и доказывают, что она совпадает с оценкой наибольшего правдоподобия. Для относительного среднего квдратического отклонения оценки дается асимптотически явные формулы при маленькой длины ячейки; таким образом, становится возможным определение оптимально квадратичной оценки параметра рассеивания в случае произвольного отношения сигнала шума. Они дают в явном виде матрицы, принадлежащие к оптимально квадратичной оценке; численное определение их элементов, однако, очень длительно, поэтому необходимо отыскать подходящий метод приближения. Так как относительно среднее квадратическое отклонение оптимальной оценки параметра рассеивания известно, то можно определить, насколько ухудшается точность оценки в случае применения какого-либо метода приближения, и так в каждом отдельном случае можно взвесить, подходит ли примененная оценка или же необходимо улучшить.

Во II-ой части авторы выработали общий метод квадратичной оценки, подходящий для практического применения. Сущность метода состоит в том, что они выражают матрицу квадратичной оценки как линейную комбинацию подходящим образом выбранных *основных матриц*. Наконец в части III, с помощью общих рассуждений второй части, авторы рекомендуют, для оценки параметра рассеяния конкретный метод, который является годным для вычислений. По предложению М. Козинса (Cosyns) упомянутые основные матрицы выбираются таким образом, чтобы они соответствовали оценке, так называемой чистой суммы квадратов *перепрыгающих* вторых разностей. В явном виде рассматриваются оценки при двух и трех основных матриц.

ON THE ESTIMATION OF THE SCATTERING CONSTANT OF AN EMULSION TRACK IN THE PRESENCE OF NOISE

by

L. JÁNOSSY — A. LEE — P. RÓZSA

Abstract

The present paper deals with the estimation of the scattering constant of high energy particles passing through an emulsion taking the background noise into account. In Part I an estimation of the scattering constant is given based on the maximum likelihood method. In view of further application another estimation is given in the form of quadratic expression in the sagittas which is unbiased and the variance of which is minimum. This is being called *optimum quadratic* estimation. This optimum quadratic estimation is proved to be the same as that gained by the maximum likelihood method. For the relative standard deviations of the estimations of the scattering parameters asymptotically exact explicit solutions are given in the case of a large number of measurements. Thus it becomes possible to determine the optimum quadratic estimation of the scattering constant in the case of an arbitrary noise-to-signal ratio. The matrices of the optimum quadratic estimation are given in explicit form, their numerical computation is however extremely cumbersome so that an approximate method has to be chosen. As the relative standard deviation of

the best estimation of the scattering parameter is known, the loss of accuracy when using an approximate method can be determined, and thus the suitability of the approximation used can in every case be decided and improved upon if necessary.

In Part II a general method for the quadratic estimation convenient for practical use is worked out. The method essentially consists in expressing the matrix of the quadratic estimation as a linear combination of suitably chosen *basic matrices*. Finally in Part III. making use of the general principle developed in Part II a concrete method for the estimation of the scattering constant convenient for practical computation is proposed. Following a suggestion by M. Cosyns the above-mentioned basic matrices were chosen to correspond to the estimation using quadratic sums of the so-called *several-step* sagittas. Estimations using two and three resp. basic matrices are treated explicitly.