

A geofizikai inverzióval nyert paraméterhibák meghatározása, ha az inverziót egyrészt tetszőleges mérési hibatípus esetére, másrészt a reziduálok tetszőleges normája abszolút minimumhelyének globális optimalizációval történő meghatározásával hajtjuk végre¹

A módszer részletes bemutatása gravitációs környezetgeofizikai modellre, a mérések Tarantola-hibájára és a P_C -norma alkalmazására történik

STEINER FERENC², HAJAGOS BÉLA²

Az univerzális gravitációs állandó különböző munkacsoportok által meghatározott értékeinek különbsége egy egész nagyságrenddel is nagyobb lehet, mint az egyes szerzők által a saját mérési eredményükre becsült hibák. Különös élességgel veti fel ez a körülmény azt, hogy az inverzió hibaviszonyainak vizsgálatát megelőzően mindenekelőtt olyan hibajellemző alkalmazásában kell megállapodnunk, amely hibatípustól függetlenül megbízhatóan alkalmazható. Az elővezetett 12 lehetőség közül a Q interszextilis félterjedelem több szempont szerint is megfelelőnek mutatkozik (és ráadásul a Gauss-hibatípus esetén értéke csaknem azonos a szórás értékével, így a megszokott klasszikus statisztikához ezzel a köldökszínórral csatlakozik).

A tetszőleges reziduál-norma globális minimalizálásával, tetszőleges mérési hibatípusú környezetben végrehajtott geofizikai inverzióknak az eredményei (pl. gravitációs modell geometriai paraméter értékei) nyilván hibával terheltek, azonban az adott premisszák esetére jelenleg a matematika még nem szolgáltat olyan elegáns lehetőséget a modellparaméter-hibák meghatározására, mint az L_2 -normára épülő klasszikus statisztika nyújt a szokásosan elfogadott (de ritkán teljesülő) premisszák mellett. A dolgozat szerzői jól definiált Monte Carlo-módszert javasolnak a meghatározott paraméter értékek hibáinak a meghatározására, a fenti értelemben felfogott általános geofizikai inverzió esetére.

F. STEINER, B. HAJAGOS: Determination of the parameter errors (demonstrated on a gravimetric example) if the geophysical inversion is carried out as the global minimization of arbitrary norms (demonstrated by the P_C norm)

It seems to be unavoidable to accept leastways in the geosciences a new error-characteristic instead of the scatter (standard deviation) to exclude e.g. in the practice of inversion such crass contradictions (see Fig. 1) which are at least partly due to the error determination method of the classical statistics. From the presented 12 possibilities the semi-intersextile range (Q) was chosen and used by the authors for characterize the errors of model parameters gained by general statistical inversion methods. As for the errors of the gravimetrically measured g_i values Tarantola type (i.e., Cauchy type) was supposed (Tarantola: "the Cauchy function"... "seems to be adequate for modelling suspected outliers by an unknown amount"), the global minimization of the P_C norm of the X_i residuals was applied as inversion method. The nowadays not yet reachable measuring accuracy figuring in the paper has no influence to the conclusions but the authors strived after such presentation of the proposed method and of the results which are as simple as possible, see e.g. the extreme simplicity of expression (11) which is to be minimized in such circumstances. Unfortunately, for arbitrary type of measuring errors and for arbitrary residual norms as yet not any mathematical method is available (to make clear the difficulties: even the error distribution types of the estimates for different model parameters figuring in a single inversion can differ from each other significantly). The Monte Carlo method proposed by the authors for the error determination of inversion results is generally applicable.

¹ Elhangzott a MGE Általános Geofizikai Szakosztálya Tudományos Anktáján, Miskolcon 1995. december 12-én. (A szerzők a tisztelt olvasó kollégák megértését kéri azért, hogy a dolgozaton néhol túlságosan érezhető az előadásszöveg jelleg, mivel nem tettünk elegendő lépést annak érdekében, hogy a dolgozatok — pardon: dolgozataink — szövegezésének egyébként megszokott szűrkeségi szintjét elérjük.)

² Miskolci Egyetem Geofizikai Tanszék, H-3515 Miskolc, Egyetemváros

A gravitációs körökben közismert BARTA és HARJÓSY [1985, 1986] cikk egyik elgondolkodtató ábráján ugyanarra az univerzális állandóra (nevezetesen az általános tömegvonzás konstansára) három szerzőtől látunk eredményeket, hibaintervallumokkal együtt. A három hibaintervallumnak egyetlen közös pontja sincs, amiből megrendülten értesülünk, statisztikai szemszögből értékelve a szituációt, hogy a konvencionális hibaszámítást elfogadva teljesen abszurd helyzetekkel is szembesülhetünk. VARGA Péter igazgató úr jóvoltából, aki egyik szakmai konzultációnk során újabb mérési eredmények leleghelyére, CHEN és COOK [1993]-ra hívta fel figyelmet, módomban van most egy analóg, és még részletesebb ábra bemutatására, négy szerző eredményei alapján, akik ugyancsak az általános tömegvonzás állandójának meghatározását tűzték ki célul (1. ábra). Az 1. ábra alapján könnyű megállapítani,

Az 1. ábrán eredményeikkel szereplő kutatók vagy csapatok nyilván levezetett mennyiségként számították az univerzális gravitációs állandó értékét. Ennek leg egyszerűbb esetében az y levezetett mennyiséget a közvetlenül mért $x_1, x_2, \dots, x_j, \dots, x_r$ mennyiségekből valamely $g(x_1, \dots, x_r)$ függvény segítségével számíthatjuk, azaz

$$y = g(x_1, x_2, \dots, x_j, \dots, x_r). \quad (1)$$

Ha az összes x_j -t Gauss-típusú (ún. normáeloszlású) hiba terheli, akkor a jól ismert formulával az y leszámaztatott mennyiség σ_y szórását a következő formula szolgáltatja:

$$\sigma_y = \sqrt{\sum_{j=1}^r \left(\frac{\partial g}{\partial x_j} \right)^2 \cdot \sigma_{x_j}^2}, \quad (2)$$

ahol a parciális deriváltak értékei természetesen az x_j -k meghatározott értékeivel számítandók, a σ_{x_j} pedig az x_j szórása.

Tekintettel arra, hogy a Gauss-eloszlásra vonatkozóan a q interkvantilis félterjedelem a $q=0,6745 \cdot \sigma$ összefüggés szerinti érték, a q_y számítása nyilván a következő összefüggés alapján történik:

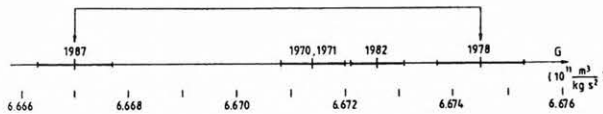
$$q_y = \sqrt{\sum_{j=1}^r \left(\frac{\partial g}{\partial x_j} \right)^2 \cdot q_{x_j}^2}. \quad (3)$$

Itt q helyett bármely interkvantilis félterjedelemre, pl. a Q interszextilis félterjedelemre is áttérhetünk volna; utóbbi esetben a Gauss-eloszlásra érvényes $Q=0,9674 \cdot \sigma$ összefüggést idéztük volna. (A q -ra és Q -ra vonatkozóan emlékeztetőül ld. a 2. ábrát.)

A (3) kifejezés általánosítható, mégpedig arra az esetre, amikor az x_j -k hibái nem Gauss-típusúak, hanem valamely más szimmetrikus, ún. stabilis eloszlást követnek, azaz sűrűségfüggvényük az

$$f_\alpha(x) = \frac{1}{\pi} \int_0^\infty \exp(-t^\alpha/\alpha) \cdot \cos(tx) dt \quad (0 < \alpha \leq 2) \quad (4)$$

formulával írható le standard esetben [ld. pl. STEINER 1990]. (Bizonyára emlékezünk arra, hogy a stabilitás fogalma a következőt jelenti: valamely α -val jellemzett típusú valószínűségi változókat összegezve az eredmény is ugyanazzal az α típusparaméterrel jellemzett stabilis eloszlás lesz.) Ezek után lássuk a (3) tetszőleges szimmetrikus stabilis eloszlásra érvényes, általános alakját. Íme:



1. ábra. Ugyanazon mennyiségre végzett mérések eredményei között egy nagyságrenddel is nagyobb lehet az eltérés, mint amilyen hibát a szerzők (feltehetően a klasszikus statisztika formuláival számolva) megadtak saját adatuk bizonytalanságának a jellemzésére

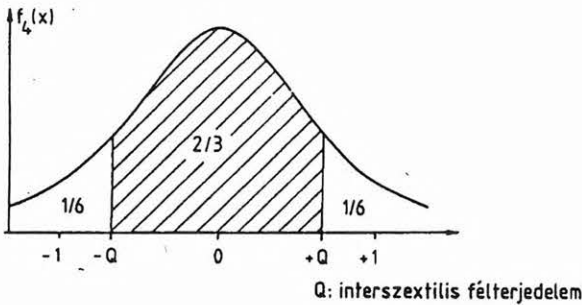
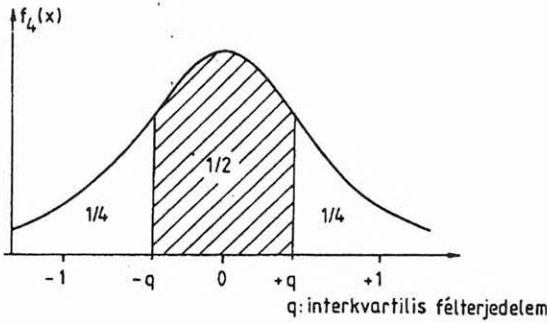
Fig. 1. Real differences can be by an order larger than the estimated errors. This contradiction can be (at least partly) due to the mechanical application of the results of the classical statistics

hogy ugyanarra az univerzális állandóra meghatározott értékek különbségei egy egész nagyságrenddel is nagyobbak lehetnek, mint a szerzők által jóhiszeműen, a klasszikus statisztika alapján megadott, azaz vélhetően empirikus szórásként számított hibaértékek. Azt hiszem, egyetértünk, hogy BECKETT és IONESCO drámáin, az abszurdítás mértékét illetően, ez az 1. ábra messze túlsz.

Az okok egyikére STEINER et al. [1996] már rámutatott, hogy ti. nagyobb valószínűséggel kaphatunk a σ szórás valódi értékére szignifikáns mértékben optimista becslést (azaz a valóságosnál jelentősen kisebb értéket, mint amilyen valószínűséggel annál nagyobb), ha a közvetlenül mért mennyiségek szokásosan korrigált empirikus szórását számítjuk a σ szórás becsléseként. Legyen szabad a következőkben egy másik lehetséges, szintén statisztikai jellegű okra felhívnom a figyelmet.

$$q_y = \left[\sum_{j=1}^r \left| \frac{\partial g}{\partial x_j} \right|^\alpha \cdot q_{x_j}^\alpha \right]^{1/\alpha} \quad (5)$$

(ld. újra az imént idézett könyvet). Az $\alpha=2$ -vel a (3) formulaként közölt (Gauss-típusú hibára vonatkozó) kifejezést kapjuk valóban vissza, azaz az (5) valóban a (3) általánosítása.



2. ábra. A q interkvartilis, ill. a Q intersextilis félterjedelmek definíciói (a törtszámok valószínűségek)

Fig. 2. The definition of the semi-interquartile (q) and semi-intersextile (Q) ranges

Ritkán hivatkozunk tekintélyekre, de most (és a dolgozat további részeiben is) fogadjuk el az inverzió egyik pápájának, TARANTOLÁnak a felfogását arra vonatkozóan, hogy Cauchy-eloszlású hibát feltételezve egyúttal már az outlier-előfordulásokat is modelleztük. Legyen szabad ezt TARANTOLA [1987] 303. oldaláról szó szerint is idéznem: „the Cauchy-function...seems to be adequate for modeling suspected outliers by an unknown amount”. A leszámított hiba szemszögéből persze édes mindegy, hogy az x_j -k hibái milyen effektusok eredőjeként adódtak Cauchy típusúaknak. Ismeretes másrészt, hogy az $\alpha=1$ típusparaméterhez tartozó stabilis eloszlás azonos a Cauchy-eloszlással, amikor is az általános (5) formula a következő, igen egyszerű alakot ölti:

$$q_y = \sum_{j=1}^r \left| \frac{\partial g}{\partial x_j} \right| \cdot q_{x_j} \quad (6)$$

Jó lenne tudni, hogy azonos q_{x_j} primer hibákhoz q_y -ra ez az utóbbi, vagy a klasszikus (3) szolgáltat-e nagyobb eredményeket. Egy teljesen általános vizsgálat bonyodalma elkerülendő, radikálisan egyszerűsítő feltevessel élünk: feltesszük, hogy mindegyik x_j hibája azonos mértékben járul hozzá a q_y hibájához, azaz fennáll, hogy

$$\left| \frac{\partial g}{\partial x_1} \right| q_{x_1} = \left| \frac{\partial g}{\partial x_2} \right| q_{x_2} = \dots = \left| \frac{\partial g}{\partial x_j} \right| q_{x_j} = \dots = \left| \frac{\partial g}{\partial x_r} \right| q_{x_r} \quad (7)$$

Ekkor (6) egyszerűen

$$q_y = r \cdot \left| \frac{\partial g}{\partial x_1} \right| q_{x_1} \quad (8)$$

alakban, a klasszikus (3) formula pedig így írható:

$$q_y = \sqrt{r} \cdot \left| \frac{\partial g}{\partial x_1} \right| q_{x_1} \quad (9)$$

A két formula összehasonlításából az következik tehát, hogy a leszámított mennyiségnek a Tarantola-feltevésnek megfelelő (8) formulából számítható, *reális* q_y hibaértékéhez képest a klasszikus (9) formula négy közvetlenül mért mennyiség esetén csak feleakkora, kilenc közvetlenül mért mennyiség esetén csak harmadakkora értéket szolgáltat. A Tarantola-feltevés negligálása, azaz a klasszikus (2) formula használata a leszámított mennyiség hibájának meghatározására ahhoz vezethet tehát, hogy a valószínűs hibánál 50%-kal, 67%-kal vagy akár ennél is nagyobb mértékben, — mindenesetre, ha a (7) szerinti egyszerűsítés jó közelítéssel teljesül —, általában $100 \cdot \sqrt{r}$ %-kal csökkentett hibaértéket kaphatunk eredményül még akkor is, ha a (2)-ben a közvetlenül mért mennyiség σ_x hibájának valódi értéke szerepel. (Persze nem felejtettük közben el, hogy ezekre is többé-kevésbé, de esetenként irreálisan optimista becsléseket alkalmazunk a klasszikus statisztika gyakorlatában, STEINER et al. [1996] szerint.)

A jelen dolgozat szerzői nem gondolják, hogy az 1. ábrán nyilvánvalóvá tett ellentmondások egészükben statisztikai eredetűek, de már maga az is megdöbbentő, hogy a konvencionális módon nem kezelhető hibatípusokra a klasszikus statisztikát mechanikusan alkalmazva *elvileg lehetséges az, hogy ezek az ellentmondások kizárólag statisztikai okokra legyenek visszavezethetők.*

Mindezeket átgondolva (és esetleg még egyszer egy pillantást vetve a hibát illetően félelmetes mértékű bizonytalanságokat tükröző 1. ábrára) talán kimondhatjuk, hogy ha a szórás helyett újabb hibajellemzőt keresünk, *megengedhető, hogy az már a*

definícióját tekintve is valamennyire, mondjuk 15–16%-os mértékben bizonytalan legyen. Különbösen is, az inverzióval meghatározandó mennyiségek hibáit általános esetben aligha várhatjuk lényegesen kisebbnek, mint mondjuk 5%-osnak, s így ha a hiba hibája 20%-os lenne is, ez maximálisan 1%-ot jelent a hiba meghatározott értékében.

A tisztelt olvasó már kezdhet jogosan aggódni, hogy a dolgozat végül révbe ér-e abban az értelemben, hogy valamilyen gravitációs inverzió hibaviszonyainak bemutatásáig sikerül-e eljutni. De ezzel a kérdéssel nem foglalkozhatunk addig, míg valamely javaslattal nem élünk és meg nem állapodunk az alkalmazandó hibadefiníciót illetően, amely tehát kielégítően működik a Tarantola-féle hibamodell esetében is, és nem mutatja a szórásnak, általában a klasszikus statisztikának az imént egy kicsit közelebb-ről is megismert, hibaszámításra vonatkozó veszélyeit.

A hibadefiníciók kiválasztásához bemutatok egy táblázatot 13-féle hibát bemutatva (mint egy étlapot ld. az I. táblázat felső részét, amelyet HAJAGOS és STEINER [1993] dolgozatából vettünk át), amelyek öt jól ismert hibaeloszlás-típus standard eseteit jellemzik. A táblázat alsó része megadja ennek az ötféle sűrűségfüggvénynek az analitikus alakját is.

Időhiány miatt lehetetlen most 13-féle hibával foglalkoznunk. Az I. táblázat utolsó két oszlopában az interkvartilis (q) és interszextilis (Q) féltérjedelmek állnak (ezek — főleg az utóbbi — definíciójára vonatkozóan legyen szabad újra a 2. ábrára utalnunk).

Az ε dihézióának a táblázat utolsó sorában másodikként feltüntetett egyenletnek kell eleget tennie. Itt X_i a teljesen általános értelemben vett eltérés vagy másképpen fogalmazva: reziduál (mért mínusz számított érték, akárhány ismeretlen paraméter szerepeljen is a számított érték meghatározási algoritmusában). Ugyanezekkel az X_i -ekkel definiált (ugyanezen sor elején) a P_k -norma, amelyet $k=3$ esetén P_J -vel, $k=2$ esetén index nélküli P -vel, $k=1$ -et választva P_C -vel, végül $k=1/2$ -nél $P_{1/2}$ -vel jelölünk (legyen szabad ezúttal mellőznünk a jelölések indokolását). Ugyanezeket az indexeket kapják a P_k -normák minimális értékei is, amelyeket határozatlanság-jellemzők lévén, U_J -vel, U -val, U_C -vel és $U_{1/2}$ -vel jelölünk. A határozatlanságot jelentő „uncertainty” szó első betűjéből származnak a határozatlanságok U -jelölései, valamint az U_j^* , U^* , U_C^* és $U_{1/2}^*$ jelölések is, amelyek az itt nem definiált P_k^* -normák minimális értékei (közelebbit ezekről HAJAGOS és STEINER [1993] közöl), és így szintén határozatlansági jellemzők.

Mivel leggyakrabban a P norma minimalizálásával dolgoznak e dolgozat szerzői az ismeretlen regressziós vagy modellparaméterek meghatározásakor, szorítkozzunk most a 8-féle U_k , ill. U_k^* közül csak egyetlenegyre: az U -ra.

A II. táblázat jelentősen csökkentti ugyan a hibák számát (az I. táblázatban szereplő 13 helyett 6-ra): a d átlagos eltérést, a σ szórást, az U határozatlanságot, mint rendre az L_1 -norma, az L_2 -norma, valamint a P -norma minimális értékeit hagytuk csak meg a hibák első csoportjából. (Az utóbbi esetben, azaz U -nál, teljesülnie kell az ε dihézióra az I. táblázat utolsó sorában szereplő feltételnek; ugyanitt szerepeltettük a P_k -norma formuláját is.)

A hibák számának csökkenésével ellentétben a hibaeloszlások típusválasztéka a II. táblázatban jelentősen megnövekedett: háromféle szupermodell típusai szerepelnek a táblázatban; a T hely-, ill. S skálaparaméterre vonatkozóan standard alakban e szupermodellek sűrűségfüggvényeinek általános analitikus alakjait is feltünteteti a II. táblázat.

A II. táblázat a típusparaméterek (az a , p vagy α) bizonyos értékeire adja meg az immár csak hatféle hibát. Kérdés, hogyan értékeljük ezt az így is tengernyi adatot?

Legjobb a kályhától elindulni: a jól ismert és megszokott, (már-már a vérünkben levő) Gauss-eloszlás 1 értékű szórásától a legkevesebb eltérésre vezető hibadefiníciót fogadjuk el. Látjuk, hogy az 1-es értékhez a Gauss-típusnál legközelebb a Q interszextilis féltérjedelem értéke áll, mindössze 3,26%-os eltéréssel. Ha tehát a továbbiakban a Q interszextilis féltérjedelmet tekintjük hibajellemzőnek, akkor ez a Gauss-eloszlásnál majdnem pontosan ugyanarra a hibaértékre vezet, mintha a szórást számolnánk (persze utóbbi esetben pontosan, azaz extrém nagy n mintaelemszám és outlier-mentesség esetén).

A táblázat utolsó oszlopai a nagy Q -ra vonatkoztatott hibaarányokat adják meg d -re, σ -ra, U -ra és e -ra. Emlékezve arra a megállapodásunkra, hogy 15–16%-os eltéréseket hibák definíciói esetében még megengedhetők tartunk, érdemes áttekinteni az arányok oszlopait, hogy ez a feltétel milyen típusstartományokra teljesül. Ami az U/Q arányt illeti, csak a két túszerűen hegyes, Eiffel-torony-szerű f_p -eloszlásnál nem teljesül ez a feltétel — ld. az U/Q -arányok értékeire bekeretezéssel végrehajtott (és csak két adatot elhagyó) markírozást a II. táblázatnak ebben az oszlopában. A kimaradt hibamodellekkel (azaz 1-nél kisebb p -kkel) azonban aligha dolgozunk a földtudományokban. Konklúziónk végül az, hogy a Q interszextilis féltérjedelem és a

a hiba eloszlás-típusa	Elméleti hibaértékek néhány eloszlástípusra												
	normaminimumok										egyéb jellemzők		
	d	σ	U_J	U	U_C	U_k	U_J^*	U^*	U_C^*	U_k^*	ε	q	Q
Laplace	1.0000	1.4142	0.9008	0.9790	1.2419	1.7975	1.2528	1.0974	0.9302	0.9662	0.8072	0.6931	1.0986
Cauchy	∞	∞	1.3333	1.5000	2.0000	2.9853	3.2278	2.2401	1.4641	1.3382	1.0000	1.0000	1.7320
geostatistical	0.5000	0.7071	0.5223	0.5585	0.6875	0.9739	0.5444	0.5005	0.4698	0.5327	0.4818	0.3704	0.5496
Jeffreys	0.3125	0.4083	0.3575	0.3796	0.4615	0.6480	0.3081	0.2981	0.3015	0.3577	0.3347	0.2498	0.3640
Gaussian	0.7979	1.0000	0.9784	1.0327	1.2421	1.7303	0.7141	0.7170	0.7740	0.9587	0.9254	0.6745	0.9672

Hibaeloszlástípusok sűrűségfüggvényei
standard esetben

Laplace	$\frac{1}{2} e^{- x }$
Cauchy	$\frac{1}{\pi} \frac{1}{1+x^2}$
geostatistical	$\frac{3}{4} \cdot \frac{1}{(1+x^2)^2 \cdot \sqrt{1+x^2}}$
Jeffreys	$\frac{35}{32} \cdot \frac{1}{(1+x^2)^4 \cdot \sqrt{1+x^2}}$
Gaussian	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$$P_k = \varepsilon \cdot \left\{ \prod_{j=1}^n [1 + (X_j/k\varepsilon)^2] \right\}^{1/2n}; \quad \varepsilon^2 = 3 \cdot \sum_{i=1}^n \frac{X_i^2}{[\varepsilon^2 + X_i^2]^2} / \sum_{i=1}^n \frac{1}{[\varepsilon^2 + X_i^2]^2}$$

I. táblázat
Table I

P -norma minimális értéke egyaránt javasolható hibajellemzőként. Az utolsó előtti oszlop bekeretezett részeinek bizonyága szerint ez a Q -val való „felcserélhetőség” (az elfogadott bizonytalansági intervallumon belül), ismételjük, az $f_a(x)$ supermodell összes feltüntetett típusára teljesül; az L_1 minimális

értékére: d -re azonban a d/Q arányok az $1 \pm 16\%$ feltételt ugyanazoknak az $f_a(x)$ -típusoknak csak kb. a felére teljesítik.

Hogy elfogultsággal ne legyünk vádolhatók (hiszen az U határozatlanságnak, mint hibajellemzőnek a használatát a miskolci Geofizikai Tanszék statisztici-

$$f_a(x) = \Gamma(a/2) \left\{ \sqrt{\pi} \Gamma[(a-1)/2] \right\}^{-1} \cdot (1+x^2)^{-a/2} \quad (a > 1)$$

Az $f_a(x)$ szupermodell hibajellemzői

a	$\frac{1}{a-1}$	q	Q	d	σ	U	ε	d/Q	σ/Q	U/Q	ε/Q
2	1.0000	1.0000	1.7320	∞	∞	1.5000	1.0000	∞	∞	0.8660	0.5774
2.2	0.8333	0.8523	1.4170	3.6043	∞	1.2693	0.9117	2.5437	∞	0.8958	0.6434
2.4	0.7143	0.7518	1.2166	1.9954	∞	1.1174	0.8418	1.9954	∞	0.9184	0.6919
2.8	0.5556	0.6221	0.9742	1.1711	∞	0.9258	0.7374	1.2021	∞	0.9503	0.7569
3	0.5000	0.5774	0.8944	1.0000	∞	0.8605	0.6974	1.1181	∞	0.9621	0.7797
3.2	0.4546	0.5407	0.8306	0.8831	2.2361	0.8070	0.6627	1.0632	2.6921	0.9716	0.7978
4	0.3333	0.4416	0.6642	0.6366	1	0.6628	0.5616	0.9584	1.5055	0.9979	0.8455
5	0.2500	0.3703	0.5497	0.5	0.7071	0.5585	0.4818	0.9096	1.2864	1.0161	0.8765
6	0.2000	0.3250	0.4787	0.4244	0.5773	0.4916	0.4281	0.8866	1.2060	1.0269	0.8943
8	0.1429	0.2688	0.3927	0.3395	0.4472	0.4081	0.3588	0.8646	1.1388	1.0392	0.9136
10	0.1111	0.2342	0.3408	0.2910	0.3780	0.3564	0.3149	0.8540	1.1092	1.0459	0.9241
40	0.0256	0.1090	0.1568	0.1303	0.1644	0.1669	0.1492	0.8305	1.0481	1.0641	0.9512
standard Gauss ($a \rightarrow \infty$)	0	0.6745	0.9674	0.7979	1	1.0327	0.9254	0.8248	1.0337	1.0675	0.9565

$$f_p(x) = p^{1-1/p} \cdot \{2 \cdot \Gamma(1/p)\}^{-1} \cdot \exp(-|x|^p/p)$$

Az $f_p(x)$ szupermodell hibajellemzői

p	$1/p$	q	Q	d	σ	U	ε	d/Q	σ/Q	U/Q	ε/Q
0.5	2	0.7042	1.3102	1.5000	2.7386	0.7774	0.4780	1.1449	2.0902	0.5933	0.3648
0.7	1.4286	0.6996	1.1939	1.1933	1.6591	0.8994	0.6694	0.9994	1.3896	0.7533	0.5607
1.0	1	0.6932	1.0986	1.0000	1.4142	0.9785	0.8065	0.9102	1.2873	0.8907	0.7341
1.5	0.6667	0.6833	1.0150	0.8641	1.1261	1.0237	0.8960	0.8514	1.1094	1.0086	0.8828
2	0.5	0.6745	0.9674	0.7979	1	1.0327	0.9254	0.8248	1.0337	1.0675	0.9566
3	0.3333	0.6593	0.9124	0.7290	0.8812	1.0221	0.9334	0.7990	0.9658	1.1202	1.0230
5	0.2	0.6356	0.8575	0.6666	0.7858	0.9861	0.9109	0.7774	0.9164	1.1500	1.0623
10	0.1	0.5989	0.7992	0.6075	0.7060	0.9256	0.8600	0.7601	0.8834	1.1582	1.0761
∞	0	0.5000	0.6667	0.5000	0.5774	0.7727	0.7198	0.7500	0.8661	1.1591	1.0798

$$f_\alpha(x) = \frac{1}{\pi} \int_0^\infty \exp(-t^\alpha/\alpha) \cdot \cos(tx) dt \quad (0 < \alpha \leq 2)$$

Az $f_\alpha(x)$ szupermodell néhány hibajellemzője

α	q	Q	σ	U	ε	σ/Q	U/Q	ε/Q
2	0.6745	0.9674	1	1.033	0.9254	1.0337	1.0678	0.9566
1.8	0.6924	1.0044	∞	1.057	0.9266	∞	1.0524	0.9225
1.5	0.7394	1.1051	∞	1.124	0.9353	∞	1.0171	0.8463
1.2	0.8432	1.3443	∞	1.273	0.9610	∞	0.9470	0.7149
1	1.0000	1.7320	∞	1.500	1.0000	∞	0.8661	0.5774

II. táblázat

Table II

kai teamje javasolta először), a továbbiakban a nagy Q -kkal dolgozunk, feltüntetve azonban a megszokottabb kis q -kat is. (Mondhatnánk, bizonyos értelemben megfelelve a valóságnak, hogy ez utóbbit egyrészt a BESSEL iránti tiszteletből tesszük, aki még egyszerűen „valószínű hibának” nevezte ezt a hibajellemzőt; lényegesen kacifántosabban, de ma már egységesen „interkvartilis félterjedelemnek” — „semiinterquartile range”-nek — nevezik ezt a kis q -t.) Persze lehetne úgy is gondolkodni — de most nem tesszük ezt, — hogy a σ szórás az L_2 -norma minimális értéke volt, s ezért konzekvensebben járnánk el, ha az új hibajellemzőt szintén normaminimumként definiálnánk; emlékezünk, hiszen az imént említettem, hogy az U hibajellemző, (amit határozatlanságnak neveztünk el,) a P -norma minimális értéke. Még egyszer rápillantva azonban a II. táblázatban a bekeretezéssel kiemelt U/Q értékekre, csak megismételhetjük, hogy gyakorlatilag (azaz ± 15 – 16% -on belül) nincs különbség a két definícióval adódó értékek között.

Végre eljutottunk tehát odáig, hogy a gravitációs inverzióval, pontosabban az inverzió hibaviszonyaival foglalkozunk. Az eredményeket a következő, környezetgeofizikai jellegű modellre mutatjuk be: talajvízszint fölötti, $2,6 \text{ t/m}^3$ sűrűségű kőzetbe kör keresztmetszettel vajt vízszintes vágatpárra vonatkozóan a mélységeket, a vágatsugarakat és a vízszintes elhelyezkedést, azaz 6 modellparamétert kell graviméteres mérések eredményeiből kiindulva meghatározni. A 3. ábra alsó fele a valóságos viszonyokat mutatja be, az ennek megfelelő, hibátlan g -szelvényt a 3. ábra felső részén sima görbe vonalként láthatjuk. Utóbbi szelvény értékeit a mérési pontokban nullkörök jelzik (a 3. ábra ezeket számértékeként is feltünteti). Biztos vagyok benne, hogy olvasóim közül többen emlékezetből is felírnák a vízszintes hengerre vonatkozó g -hatásnak a formuláját, amelyből modellünkre a következő egyszerű képlet származtatható:

$$g_i = -41,9 \cdot 2,6 \cdot \left[\frac{R_1^2 \cdot m_1}{m_1^2 + (y_i - t_1)^2} + \frac{R_2 \cdot m_2}{m_2^2 + (y_i - t_2)^2} \right] \mu\text{Gal} \quad i = y_i = 0; 1; 2; \dots; 17; 18. \quad (10)$$

(Az y_i ugyanúgy méterben értendő, mint a modellparaméterek számértékei.)

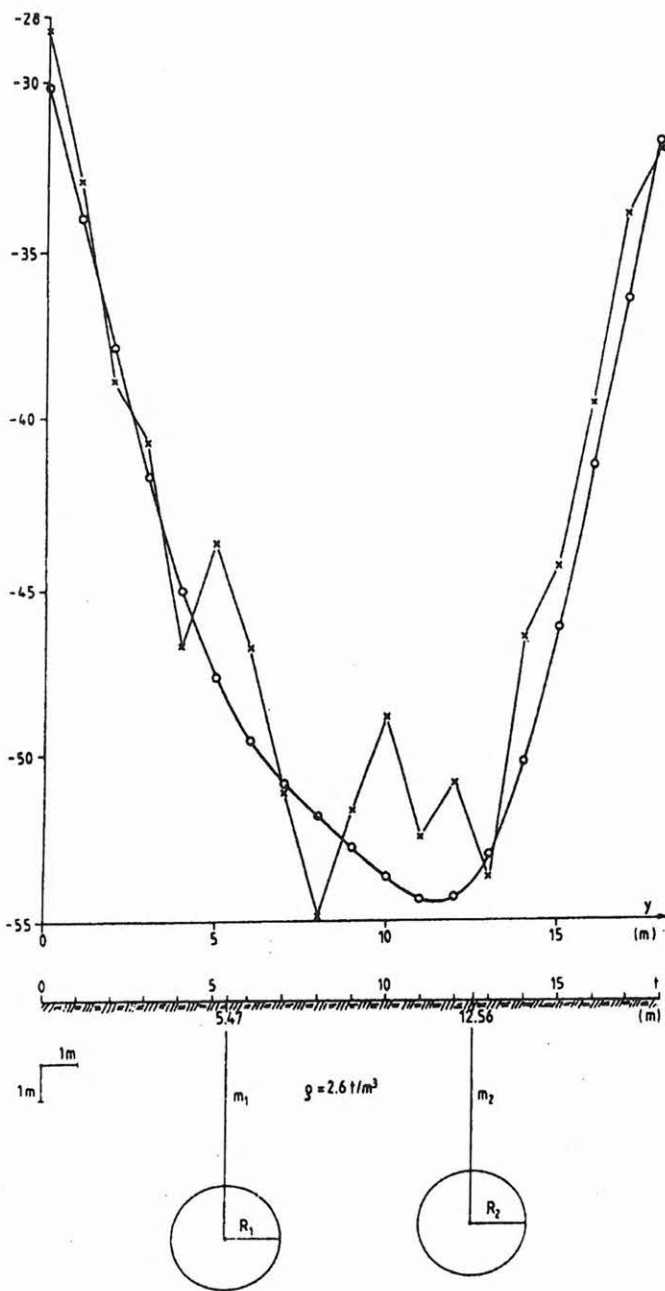
A formula egyszerű ugyan, de az ismeretlen paraméterekben nem lineáris. Ez a körülmény az inverzió végrehajtását, azaz a modellparaméterek meghatározását megnehezítő körülménynek számított még nem is olyan régen. A globális (azaz abszolút) minimumot egyre inkább gazdaságosan szol-

gáltató algoritmusok, ill. programok megjelenésével azonban az inverzió feladata ún. „optimalizációs” feladattá egyszerűsödött; döntenünk csak arra vonatkozóan kell ilyenkor, hogy a reziduálok melyik normáját akarjuk minimalizálni.

Nyár elején Miskolcon volt a középiskolai fizikatanárok országos ankétja, amikor is az inverzió lényegét ezzel a modellel mutattam be, $3 \mu\text{Gal}$ szórású Gauss-hibát szuperponálva a hibátlan g_i adatokra. Ekkor természetesen a reziduálok L_2 -normáját minimalizáltuk; az eredményül kapott hat modellparaméter-érték eltérése a valódi értékektől 6; 6; 9; 2; 2, ill. 3% -os volt.

Az alább bemutatandó vizsgálati eredmények az OTKA által támogatott tudományos kutatás keretében születtek, amikor is illik a *jövő* (persze a nem túl távoli jövő) várható mérési és számítástechnikai lehetőségeiből kiindulni. Jól tudjuk, hogy a mikrograviméter-mérések Q hibáját $3 \mu\text{Gal}$ -nál kisebb értékre *jelenleg* gyakorlatilag nem lehet lecsökkenteni, de *nem tartjuk irreálisan távolinak azt az időt*, amikor további műszerfejlesztésekkel (gondoljunk pl. a Scintrex-mikrograviméterbe épített automatikus mérésismétlőre, az eredmény sorozat azonnali értékelőjére, mely utóbbit pontosabb eredményre vezető algoritmussal lehetne működtetni akár már ma is, ld. HAJAGOS és STEINER [1992]), valamint a mérés-kivitelezés és a *korrekciószámítás további tökéletesítésével* $\sqrt{3} \mu\text{Gal}$ -ra is le lehet majd szorítani a mikrograviméter-mérések hibáját. Vizsgálatainkat tehát a hiba $Q = \sqrt{3} \mu\text{Gal}$ nagyságának a feltételezésével végeztük, de ez nem érinti következtetéseinket: próbaképpen $Q = 2 \cdot \sqrt{3} \mu\text{Gal}$ (tehát $3 \mu\text{Gal}$ -nál nagyobb) hibával terhelt mérési adatokra is végeztünk teljesen problémamentesen inverzió sorozatokat azzal az egyáltalában nem meglepő eredménnyel, hogy a paraméter-hibák is megkétszereződtek.

A Monte Carlo-vizsgálatok túlnyomó többségében ezt a $Q = \sqrt{3} \mu\text{Gal}$ nagyságú hibát Tarantola-hibaként szuperponáltuk a valóságos ható 19 pontban mért hatásának tetszőleges pontossággal ismert értékeire. Ennek egyik realizációját a 3. ábra tört vonalú szelvénye mutatja be. A példát — a könnyebb ábrázolhatóság kedvéért — a viszonylag szelídebb esetek közül választottuk, hiszen gyakran léptek fel a 19 pont közül 1, 2, 3 (vagy még több) pontban többször $10 \mu\text{Gal}$ abszolút hibával terhelt outlierok. Persze látjuk, hogy még ebben a szelídek minősített esetben sem ritkák a teljes g -változási tartomány 10% -ánál is nagyobb eltérések: történetesen a pontok 37% -ában tapasztaljuk ezt az ábrázolt példa esetében.



i	g_i
0	-30.2
1	-34.0
2	-37.9
3	-41.7
4	-45.0
5	-47.7
6	-49.6
7	-50.9
8	-51.9
9	-52.8
10	-53.7
11	-54.4
12	-54.3
13	-53.0
14	-50.2
15	-46.2
16	-41.5
17	-36.5
18	-31.9

$R_1 = 1.5$ m
$m_1 = 7.5$ m
$t_1 = 5$ m
$R_2 = 1.5$ m
$m_2 = 6.5$ m
$t_2 = 13$ m

3. ábra. Környezetgeofizikai jellegű gravitációs 2D-modell rajza (balra alul) a pontos modellparaméter-értékekkel együtt (alul jobbra). Felül ennek a modellnek a hibátlan g_i -hatásait jobbra számértékekkel, balra (sima vonallal összekötött) nullköröként adtuk meg. A tört vonalú szelvényt a Tarantola-hiba egy realizációjával kaptuk

Fig. 3. Gravity 2D-model with given geometrical parameters. The error-free values are given (as g_i -values as well as by a smooth curve). The zigzag line represents one single (and moderate) realisation of the measured profile if Tarantola type errors occur

Egy rövid kitérő: a mérési adatok Q_g hibáját miért választottuk ausgetippelt $Q = \sqrt{3} \mu\text{Gal}$ -nak? Nos, ez a nagyságú Tarantola-hibamodel azonos a Cauchy-típusú eloszlás standard változatával, amikor is a q interkvartilis félterjedelem és az ε dühöz is pontosan 1-gyel egyenlő, hozzájárulva a viszonyok valamivel jobb technikai áttekinthetőségéhez.

A továbbiakban ezt a Q_g hibát *a priori ismertnek* tekintjük, hiszen graviméterek esetében valóban nem irreális feltételezni a hiba előzetes ismeretét. Ha valamilyen okból ez a feltétel mégsem teljesülne, az inverzióval kapott paraméterekkel számított értékeket célszerű levonni a mért értékekből, és az így kapott reziduál adatrendszer empirikus Q értékét fogjuk Q_g -ként elfogadni. Mondhatná valaki, hogy

ez utóbbi eljárás pontossága megkérdőjelezhető. Ezzel a problémával azonban a P - és P_C -norma robusztussága miatt teljesen felesleges foglalkozni.

A $\sqrt{3}$ -választás formai egyszerűsítést tesz lehetővé a minimalizálandó formulát illetően: végül tehát csak a

$$\prod_{i=0}^{18} (1+X_i^2) \quad (11)$$

kifejezés minimumhelyét kell megkeresnünk.

Hogy azonos hibaviszonyok esetén hogyan viselkednek az egyes normák, ezzel a problémával SZŰCS [1995b] foglalkozott, nem kevesebb, mint 10 normára vonatkozóan. Nem tudom azonban megállni, hogy egyetlen ábrát (ld. 4. ábra) ne szenteljek a P_C - és L_2 -norma viselkedésének összehasonlítására a vizsgált gravitációs modellnél.

A megállapodásunk szerinti ($\sqrt{3}$ μGal -os Tarantola-) hibát szuperponáltuk 25-féleképpen a hibamentes g_i értékekre, és mind a 25 esetre elvégeztük a P_C -inverzió mellett az L_2 -inverziót is. A hat modellparaméter mindegyikére tehát 25-25 db értéket kaptunk, így az inverzióval kapott R_1, m_1, t_1, R_2, m_2 és t_2 értékek Q -hibáit mindkét normát illetően megtudtuk határozni, így a Q_{L_2}/Q_P arány is képezhető volt. Persze 25 adat még elég kevés ahhoz, hogy megbízható értékeket kapjunk, ezért az egész procedúrát 5-ször megismételtük; az így adódó 5 db Q_{L_2}/Q_P arányt mind a hat modellparaméterre a 4. ábra felső felén ábrázoltuk, mediánjaikat hosszabb vonallal. E mediánok legkisebb értéke 1,8 volt, a legnagyobbja 2,76, és amint azt az utolsó sorban látjuk, a mediánok mediánja kerekén 2,4-nek vehető. Konklúzióink tehát az, hogy Tarantola-hibák esetén az L_2 -inverzióval nyert paraméterek hibái lényegesen (aktuális Monte Carlo-vizsgálatunk esetében 2,4-szeresen) nagyobbak a P_C -inverzióval nyert paraméterértékekhez viszonyítva.

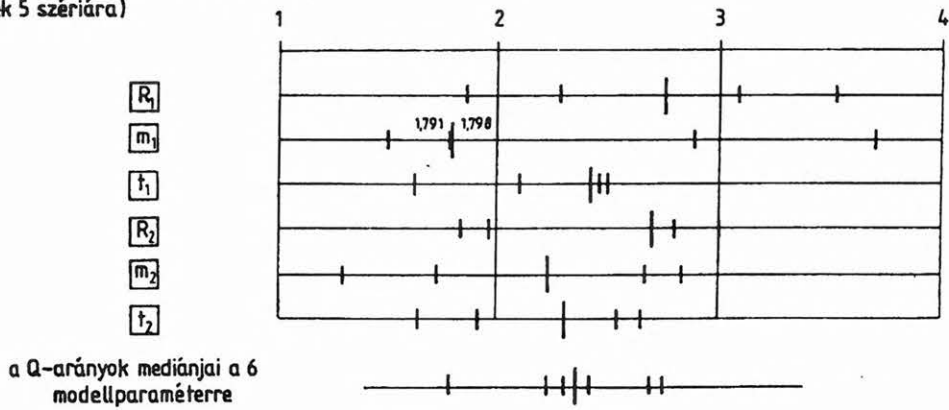
Mondhatná valaki, hogy ebben semmi meglepő nincs, hiszen a Tarantola-hibákhoz adekvát módon a P_C -inverzió tartozik. Fordítsuk meg ezért a szituációt: szuperponáljunk ugyancsak $Q_g = \sqrt{3}$ nagyságú, de Gauss-típusú hibákat a hibamentes g_i értékekre, egyebekben pedig szóról szóra ugyanazt hajtsuk végre, amit az előbbiekben tettünk: 5-ször megismételve a 2×25 db inverziót, a hat modellparaméterre vonatkozóan ugyanúgy a Q_{L_2}/Q_P hibaarányokat hordjuk fel (ld. a 4. ábra alsó felét), végül pedig (az utolsó sorban) a mediánokat. A mediánok mediánja most meglepően közel adódik 1-hez (pontos értéke 0,91), így Gauss-típusú hiba esetén gyakorlatilag mindegy, hogy L_2 - vagy P_C -in-

verziót alkalmazunk-e. Ez az eredményünk már aligha nevezhető triviálisnak, hiszen ez úgy is interpretálható, hogy az L_2 -normát még Gauss-típusú mérési hibák esetén sem kötelező alkalmazni, esetleg csak akkor, ha az outlier-mentesség és a kristálytisza Gauss-típus egyidejűleg abszolúte garantált. A P_C -norma alkalmazása pedig ez utóbbi, szinte sohasem előforduló esetre sem növeli 10%-nál nagyobb mértékben az L_2 alkalmazásakor jelentkező paraméterhibát.

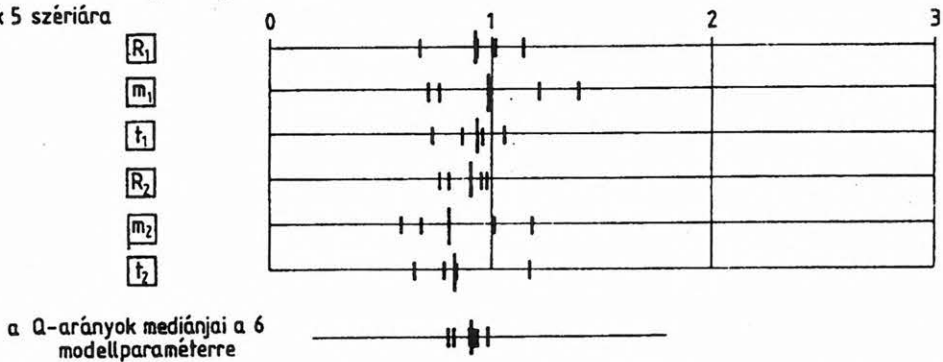
A 4. ábra csak hibaarányokat tüntet fel, az egyes modellparaméterekre vonatkozó Q_p -k abszolút értékeit nem. Persze ami késik, az nem múlik, de legyen szabad először felvetnünk azt a kérdést, hogy az előzőekben fontos praktikus előnyökkel bírónak megismert P_C -inverzió esetén lehet-e valamilyen eredményünk arra, hogy ezeket a paraméterhibákat az L_2 -inverzióhoz hasonló módon, kellő matematikai megalapozással, elegánsan számíthassuk. Sajnos, nemleges a válaszuk. Megvizsgáltuk ugyanis a Tarantola-hiba 200-féle realizációjával azt, hogy a P_C -inverzióval nyert paraméterek milyen típusú hibaeloszlást mutatnak az $f_a(x)$ supermodellen belül (ilyen vizsgálatokhoz 25 adat nyilván kevés lenne). A vizsgálatokat megismételtük Gauss-típusú hibák 200-féle szuperponálásával is. Az 5. ábrán a $t=1/(a-1)$ típusparaméter-tengelyen hordtuk fel az adatokat azzal a meglepő eredménnyel, hogy a mélység- és vízszintes távolság-adatok (m_1, m_2, t_1 és t_2) típusa a $t=0,25$ -tel jellemzett *geostatistikus típus*hoz áll közel, az R_1 és R_2 hengersugár-adatok pedig jó közelítéssel *Gauss-eloszlásúaknak* ($t=0$) vehetők. (A típus-meghatározások STEINER és HAJAGOS [1995] szerint történtek.) Ebből az önmagában is érdekes eredményből — amelyre nézve érdemes volna még további vizsgálatokat is végezni, — most csak azt a következtetést vonjuk le, hogy egy elméleti megközelítésben jelentősen eltérő típusokat egyidőben kellene kezelni tudni, s ezért vagyunk kénytelenek arra következtetni, hogy a közeljövőben elegáns matematikai eredményt a P_C -inverzióval nyert paraméterek hibáinak meghatározására aligha várhatunk. (Bár ne lenne igazam.)

Adós vagyok még a P_C -inverzióval nyert paraméterek hibaértékeinek a megadásával. Ha a Q -arányok (ld. 4. ábra) elfogadható pontosságú meghatározására elegendőek voltak 25-elemű paraméterhalmazok (25-féleképpen szuperponálva Tarantola-hibát a hibamentes értékekre és L_2 -, valamint P_C -inverziót végrehajtva), akkor ez bizonyára itt is elegendő lesz; esetleg a procedúrát nem 5-ször, hanem, mondjuk, 10-szer fogjuk megismételni.

Q_{L_2}/Q_p -arányok a modellparaméterre
 $\sqrt{3}$ -as Q -jú Tarantola-hiba esetén
 (25-szörös ismétlés alapján meghatározott
 értékek 5 szériára)

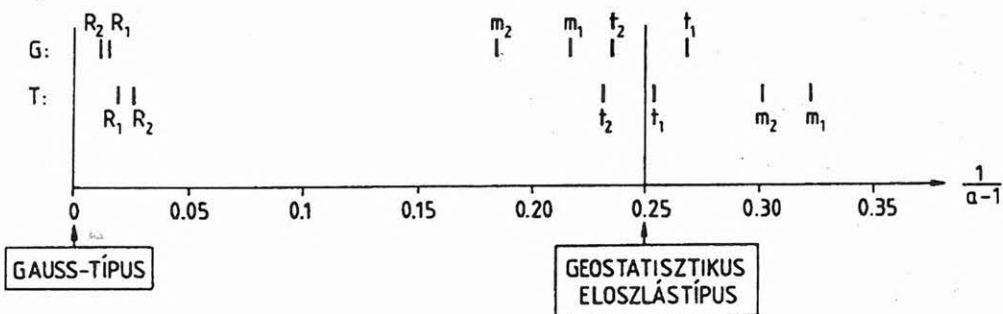


Q_{L_2}/Q_p -arányok a modellparaméterre
 $\sqrt{3}$ -as Q -jú Gauss-hiba esetén
 (25-szörös ismétlés alapján meghatározott
 értékek 5 szériára)



4. ábra. Tarantola-hibájú g -adatrendszer inverziója az L_2 -normával kb. két és félszer nagyobb paraméterhibákra vezet, mint a Tarantola-hibának megfelelő P_C -normával végrehajtott inverzió, az utóbbi normával azonban Gauss-típusú mérési hibák esetében is az optimálisnál alig nagyobbak a modellparaméterek hibái

Fig. 4. The errors of the model parameter values gained by the L_2 -inversion are more than twice as great than by using P_C -inversion, if the measured values show Tarantola type error. In the contrary, both inversions lead to very near lying parameter error values, if the measured data have Gaussian error



5. ábra. A 3. ábra gravitációs modelljénél a geometriai paraméterekre vonatkozó becslések eloszlástípusa a Gauss-féléhez közelinek adódott a hengersizűgerek (R_1 és R_2) esetén, míg a mélység- (m_1 és m_2) és távolságparaméterek (t_1 és t_2) becsléseinek eloszlástípusa a geostatistikus eloszlástípushoz ($a=5$) adódott közelinek, attól függetlenül, hogy a méréseket Gauss-féle, vagy Tarantola-típusú hiba terhelte-e

Fig. 5. Independently of the type of the errors of the measured g -data, the estimates of two model parameters (R_1 and R_2) show nearly Gaussian type, in the contrary, the estimates for the other four model parameters (m_1 , m_2 , t_1 and t_2) are approximately geostatistically distributed

A 6a., 6b., 6c., 6d., 6e., és 6f. ábrák vonalpárjai rendre az R_1 , R_2 , m_1 , m_2 , t_1 és t_2 modellparaméterekre vastag vonallal az interszextilis intervallumot jelzik, ahol a mediánt és a kvartilisek értékeit is bejelöltük. Meghosszabítottuk továbbá — vékony vonallal — a hibákat jellemző, $2Q$ hosszúságú intervallumot az aktuálisan minimálisnak, ill. maximálisnak adódó paraméterértékig.

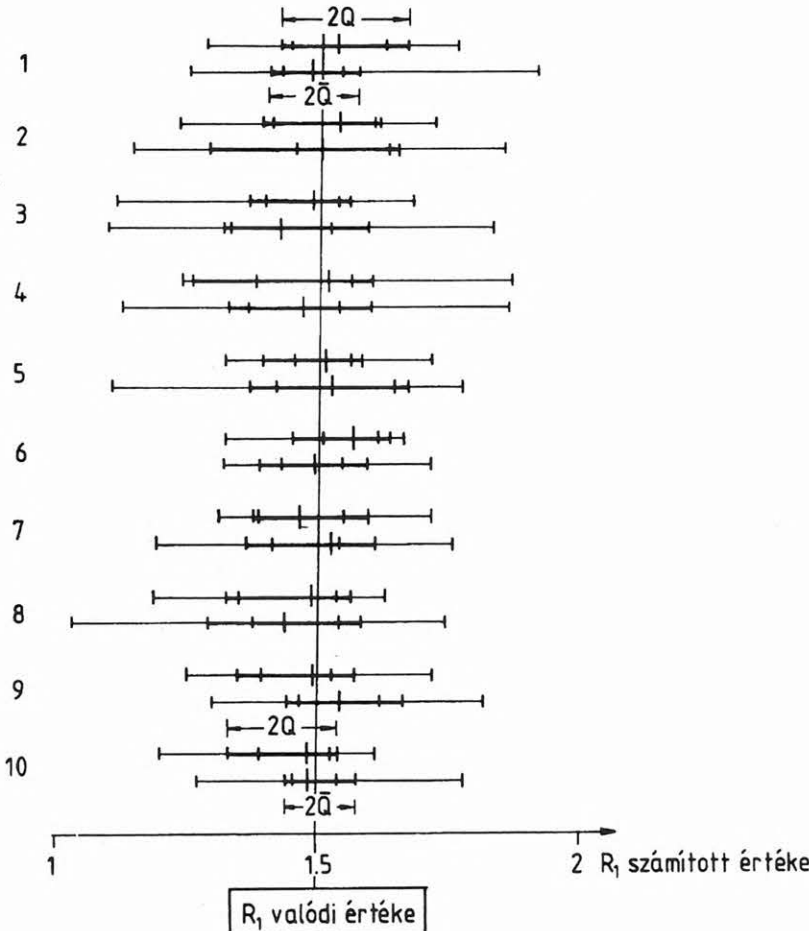
A vonalpárok közül a felső mind a hat paraméterre annak az esetnek felelt meg, amikor 25-féleképpen szuperponáltunk Tarantola-hibát a hibamentes értékekre. A tízszeres ismétléssel 9 cm-től 17 cm-ig terjedő 10 db Q hibaérték adódott az R_1 hibájára, amelynek mediánja 10,5 cm-nek, azaz 7%-osnak adódott. Ugyanezt a procedúrát végrehajtva R_2 -re 6,3%-os a meghatározott hiba; m_1 -re 8,5%-os, m_2 -re 6,9%-os hibát kaptunk. A t_1 -re és t_2 -re nyilván nincs értelme, hogy százalékosan adjuk meg a hibát, hiszen a t_1 és t_2 távolságot egy önkényesen felvett origótól számítottuk. A vízszintes távolságok Q_1 , ill. Q_2 abszolút hibaértékei meglepően kicsinyeknek: 39 cm-nek, ill. 25 cm-nek adódtak. A vázolt eljárás után szinte pillanatokon belül meg lehet adni a becslést a hibák (szintén Q értelemben felfo-

gott) hibájára is: ez az érték R_1 -re, R_2 -re, m_1 -re, m_2 -re, t_1 -re és t_2 -re rendre 14%-os, 21%-os, 16%-os, 17%-os, 27%-os és 33%-os értékűnek adódott.

És ezzel le is zárhatnám a paraméterek hibáira vonatkozó ismertetésemet, ha megelégednék elméleti tájékozódással, ill. tájékoztatással. Fülemben azonban a gyakorlati szakember felettébb jogos kérdése: uram, ez mind nagyon szép, de nekem nem 25 (vagy pláne 10×25), hanem csak egyetlenegy g_i adathalmazom van 19 elemmel, és ebből mind a hat paraméterre csak egyetlenegy értéket tudok meghatározni; ennek a mindössze hat paraméteradatnak a hibájáról szeretnék tájékozódni: mi ilyenkor a teendőm?

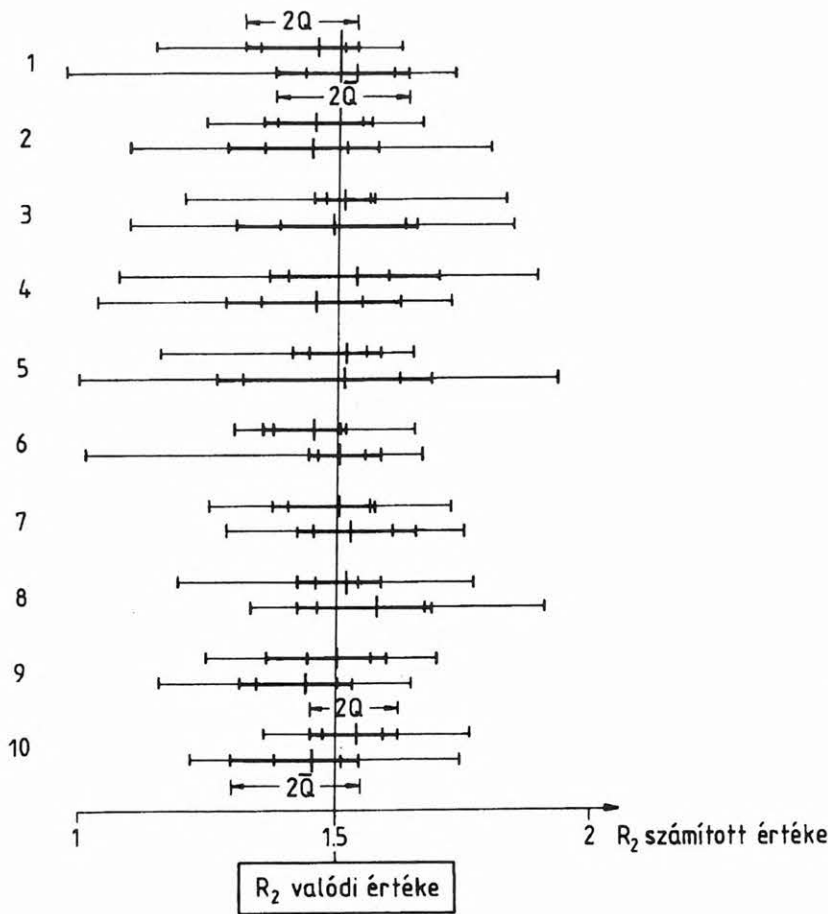
Indoklás nélkül ugyan, de válaszként azonnal megadom a receptet a feltételezett kérdezőnek: a mért (tehát hibával terhelt) adatokra szíveskedjék, mondjuk 25-ször, az eredeti hibával azonos, újabb hibákat szuperponálni, (ez idáig analóg az Arany János-féle szőlősgazda jégveréskor alkalmazott módszerével, amikor doronggal paskolván szőlejét, így kiáltott: „én uram-isten! csak rajta! hadd lám: mire megyünk ketten!”), — és az inverziót (az elsőn túl) még 25-ször elvégezve szíveskedjék megrajzolni

mindegyik modellparaméterre az interszextilis intervallumot. Az így kapott intervallumhossz feleként adódó \overline{Q}_{p_j} kétharmadával, mint felső vonás nélküli Q_{p_j} -vel kellő biztonsággal jellemezhető az első, egyetlen inverzió során kapott modellparamétereinek hibáit.



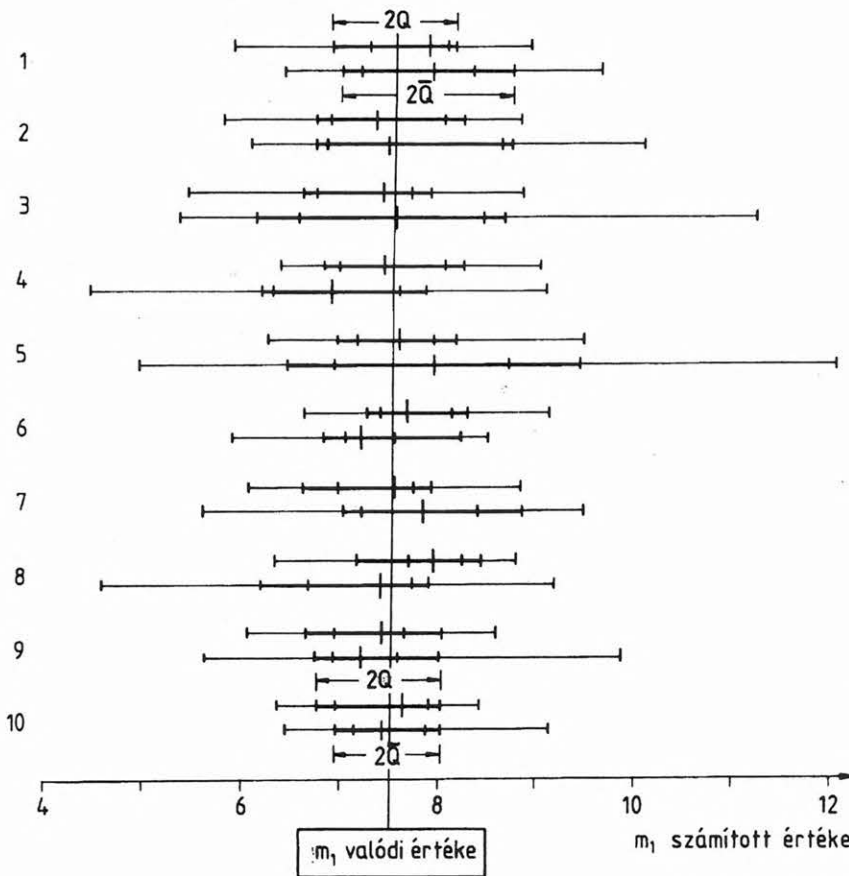
6a. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való szuperponálása után az R_1 becsült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (\overline{Q}_{p_1}) jellemzi akkor, ha a szuperpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok szuperponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait \overline{Q}_{p_1} , a vonalpárok közül az alsón vastagóh rajzolt vonalszakasz félhossza jellemzi

Fig. 6a. Figure to the proposed Monte Carlo method for the estimation of Q_{R_1}



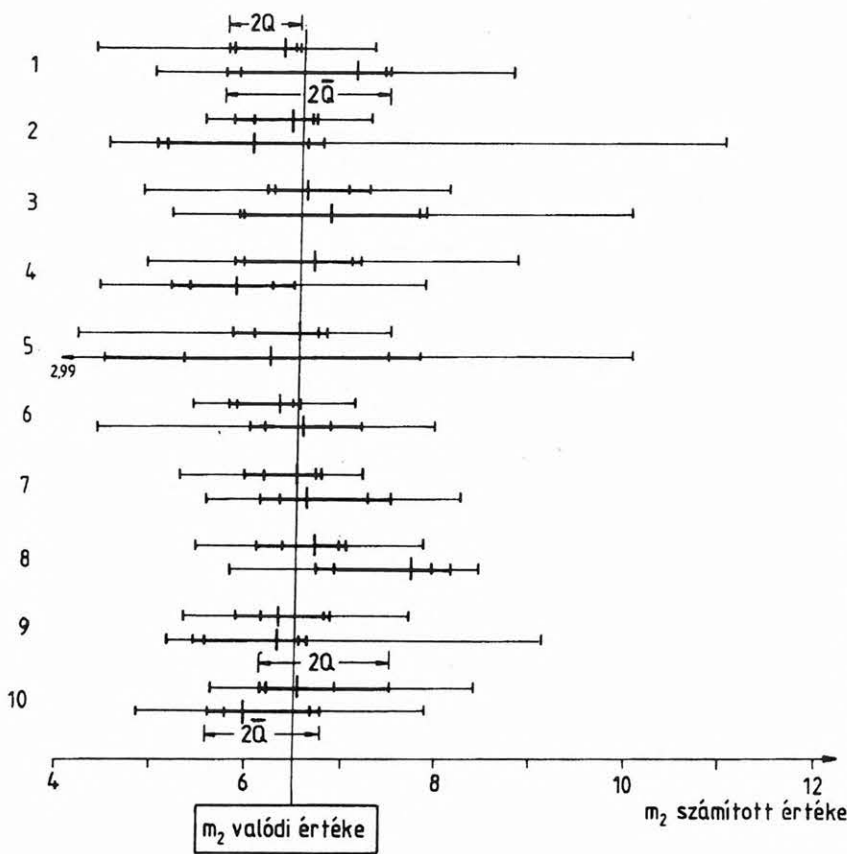
6b. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való szuperponálása után az R_2 becült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (Q_{p_i}) jellemzi akkor, ha a szuperpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok szuperponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait Q_{p_i} , a vonalpárok közül az alsón vastagon rajzolt vonalszakasz félhossza jellemzi

Fig. 6b. Figure to the proposed Monte Carlo method for the estimation of Q_{R_2}



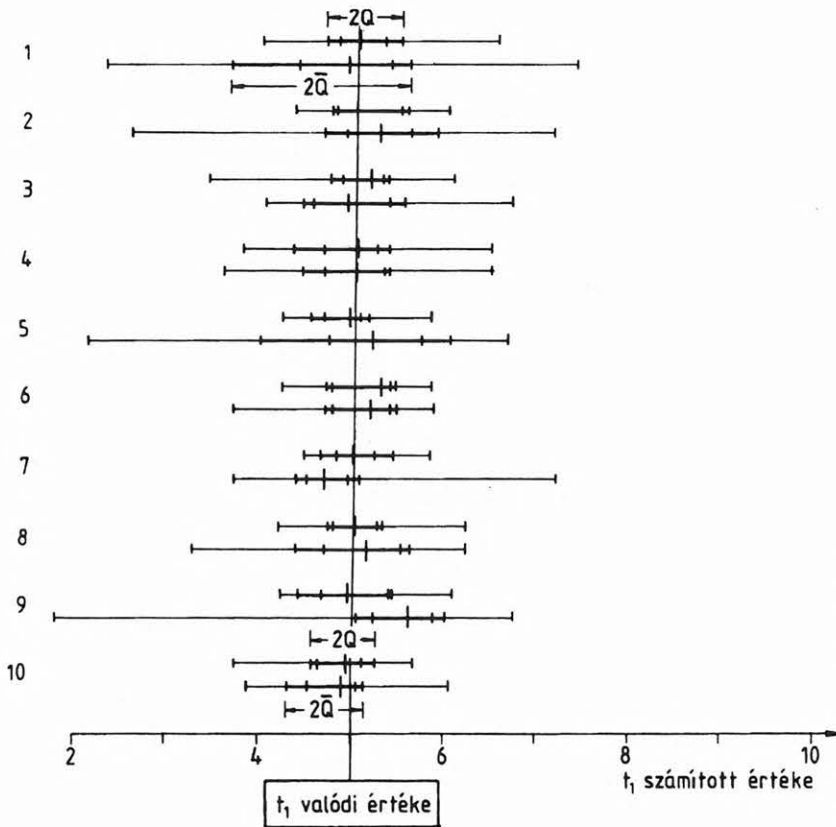
6c. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való szuperponálása után az m_1 becült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (Q_{p_i}) jellemzi akkor, ha a szuperpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok szuperponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait Q_{p_i} , a vonalpárok közül az alsón vastagon rajzolt vonalszakasz félhossza jellemzi

Fig. 6c. Figure to the proposed Monte Carlo method for the estimation of Q_{m_1}



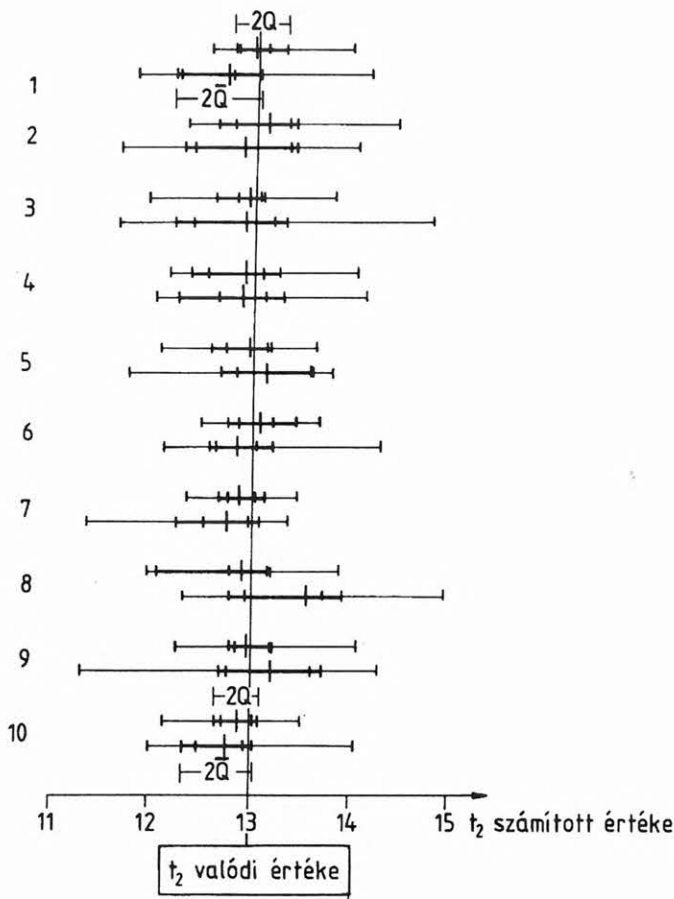
6d. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való superponálása után az m_2 becült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (Q_{p_i}) jellemzi akkor, ha a superpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok superponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait Q_{p_i} , a vonalpárok közül az alsón vastagon rajzolt vonalszakasz félhossza jellemzi

Fig. 6d. Figure to the proposed Monte Carlo method for the estimation of Q_{m_2}



6e. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való superponálása után a t_1 becült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (Q_{p_i}) jellemzi akkor, ha a superpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok superponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait Q_{p_i} , a vonalpárok közül az alsón vastagon rajzolt vonalszakasz félhossza jellemzi

Fig. 6e. Figure to the proposed Monte Carlo method for the estimation of Q_{t_1}



6f. ábra. 25-ször függetlenül generált 19 db Tarantola-hibának a gravitációs adatokra való szuperponálása után a t_2 becslült értékeinek bizonytalanságát a vastag vonallal rajzolt interszextilis intervallum félhossza (Q_{pj}) jellemzi akkor, ha a szuperpozíció hibátlan adatokra történt (ld. a vonalpárok közül a felsőt). Ha már hibás mérési adatokra történt ugyanezen véletlen számok szuperponálása (10 esetre realizálva a primer módon hibás mérési adatokat), akkor az inverzióval kapott eredmények bizonytalanságait Q_{pj} , a vonalpárok közül az alsón vastagon rajzolt vonalszakasz félhossza jellemzi

Fig. 6f. Figure to the proposed Monte Carlo method for the estimation of Q_2

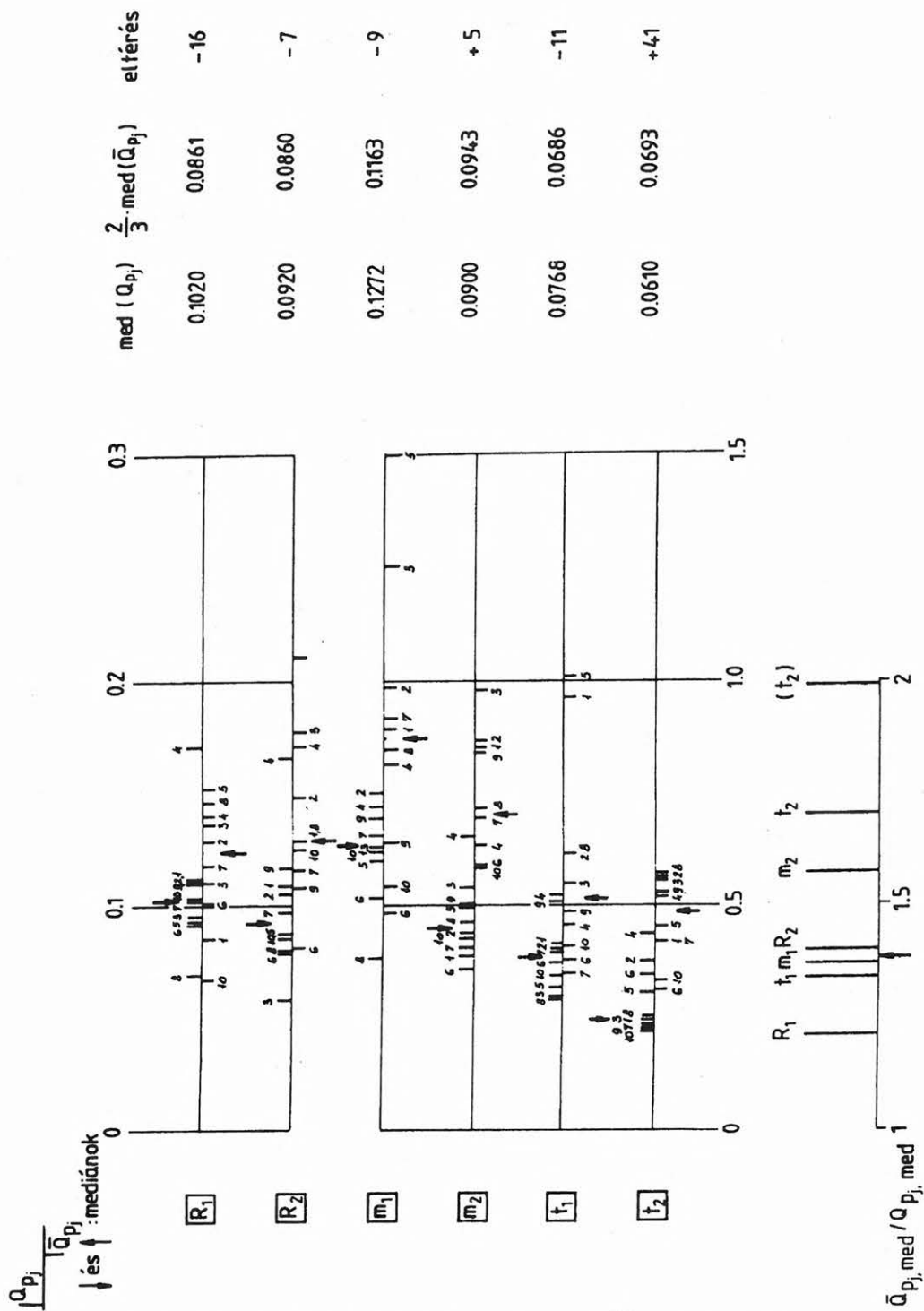
A fenti recepthez bizonyítékul először a 6a., 6b., ..., 6f. ábrákon szereplő vonalpárok közül az alsókat szíveskedjenek tekinteni: ezek úgy készültek, hogy 10 esetben egy-egy $Q_g = \sqrt{3}$ μGal -os Tarantola-hibát szuperponáltunk a hibamentes adatokra, majd ezeket rögzítetten tartva szuperponáltuk ugyanazt a 25-féleképpen realizált Tarantola-hibát, amelyek eredményeképpen a 6a.—6f. ábrák vonalpárjainak felső részén a már megtárgyalt, de a gyakorlat igényeit nem kielégítő módon a Q_{pj} -intervallumokat megszerkeszthettük.

Az összesen 60 vonalpár összehasonlítása ránézésre igazolja azt, hogy az ajánlott Monte Carlo-recept szerint adódó Q_{pj} -kel túlnyomóan felső becslést kapunk az egyetlen inverziókból származó modellparamétereink Q_{pj} hibájára. Ez teljesen érthető, hiszen a \bar{Q}_p/Q_p arányok egyrészt 1-nél biztosan nagyobbak, hiszen a Tarantola-hiba egy másik (igaz, hogy rögzített) Tarantola-hibával van növelve. Másrészt, ha mind a 25 esetben két független Tarantola-hiba összegét szuperponáltuk volna a hibátlan értékekre, akkor — a g_i hiba típusának stabilitásából és $\alpha=1$ -ből következően — pontosan kétszer akkora Tarantola-hiba terhelné mérésünket, amely — ahogyan arról már szó volt, — a paraméterhibákat is megkétszerezné. Ebből az következik, hogy

a \bar{Q}_p/Q_p arányok mindegyike 1 és 2 közé esik, így, kissé nagyvonalúan, 1,5-nek becsüljük ezt az arányt. Az ebből következő 2/3-os értékű korrekciós faktor használhatóságát a 7. ábrán feltüntetett táblázat értékpárjainak összehasonlításával ellenőrizhetjük. A 10 adatpár csak a t_2 modellparaméternél bizonyult kevésnek, a 2×10 adatsor furcsa eloszlását, és ezzel 41%-os felső becslést eredményezve. A többi öt modellparaméternél a dolgozatban javasolt módszerrel számolt és a valódi hibaértékek közötti eltérések nem haladták meg abszolút értékben a 16%-ot.

Itt tulajdonképpen most már valóban vége is lehetne a dolgozatnak, de befejezésül röviden legyen szabad még egyszer visszatérnem a paraméterhibák Monte Carlo-számítások nélküli, matematikai eredményekre támaszkodó elegáns számíthatóságára, ill. egy ilyen eljárás kidolgozásának a reménytelenségére, amely módszer tehát tetszőleges normára és bármilyen mérési hibatípusra alkalmazható lenne.

Mint láttuk az 5. ábrán, geostatistikai, ill. közel Gauss-eloszlású paraméter-eloszlásokat kaptunk ugyanazon inverzió során. A tisztán geostatistikus típust standard esetben a $Q=0,5496$, ill. a $\sigma = 1/\sqrt{a-3} = 1/\sqrt{2} = 0,7071$ hiba jellemzi; σ/Q tehát 1,29. Emlékezzünk, hogy a Gauss-eloszlásra ez az arány mindössze 1,03 volt, azonos Q -hoz



7. ábra. Előzetes Monte Carlo-igazolás a $Q_{p_j} = 2/3 = \bar{Q}_{p_j}$ összefüggéshez
 Fig. 7. Preliminary Monte Carlo verification to the equation $Q_{p_j} = 2/3 = \bar{Q}_{p_j}$

tehát pusztán típuskülönbségből származóan 60%-os variancia-eltérések tartozhatnak. Talán ez az egyetlen adat is érzékelteti azt, hogy már a paraméterhibák típusainak különbözősége is milyen mértékben teheti az eredmények minősítését bizonytalaná. Ami az elméletet magát illeti, a klasszikus statisztikára támaszkodó, a modellparaméterek hibáira

vonatkozóan jól kidolgozott minősítési módszerek az általánosítása tetszőleges mérési hibatípusok és különböző normák eseteire a fentiek szerint matematikailag e pillanatban szinte megvalósíthatatlannak látszik. Szerencsére azonban a statisztika nem más, mint *matematika + számítástechnika*, és ahogyan azt korábban láttuk egy egyszerű gravitációs inverziós

modell hibaviszonyainak vizsgálatával kapcsolatban, a számítástechnika (a maga hosszadalmas Monte Carlo-vizsgálataival) hozzásegíthet a megoldáshoz bennünket akkor is, amikor az egyébként olyannyira elegáns matematika egyelőre csődöt mond.

Végül, ámbár korántsem utolsósorban, de valószínűleg feleslegesen hívjuk fel tisztelt olvasóink szíves figyelmét arra, hogy a gravitációs modellpéldán e dolgozatban részletesen bemutatott Monte Carlo-eljárás általában, *elvileg bármilyen inverzió esetén alkalmazható a paraméterhibák meghatározására*. A Q_{p_i} / \bar{Q}_{p_i} -arányra általános stabilis eloszlás, mint mérési hibatípus-modell esetén $2/(1+2^{1/\alpha})$ a becsült érték, így a dolgozatban ismertetett Monte Carlo-procedúrával nyert \bar{Q}_{p_j} -ből a

$$Q_{p_i} = \frac{2}{1+2^{1/\alpha}} \bar{Q}_{p_i} \quad (12)$$

formulával becsülhetjük valamely mérési adatrendszer inverziójával nyert geofizikai modellparamétereink hibáit.

Még egy utolsó megjegyzés: ha mérési hibáink típusát az $f_a(x)$ szupermodell egy elemével szeretnénk modellezni, a (12)-vel (némi további közelítéssel ugyan, de) továbbra is dolgozhatunk, mivel nemcsak azt tudjuk már régebről (kvalitatíve ld. pl. a STEINER [1990] 247. oldalán levő ábrát), hogy az $f_a(x)$ eloszlástípusokhoz a Gauss-Cauchy típusstartományban a stabilis eloszlások nagyon hasonlóak, de újabb eredményként kvantitatíve is meg tudjuk adni [HAJAGOS és STEINER 1995], hogy adott α -jú $f_a(x)$ -típushoz milyen α -val jellemzett stabilis eloszlástípus áll a legközelebb. Világos, hogy e közelítések (beleértve a (12) faktorát is) még további, sokoldalú vizsgálatot követelnek, de ezek a számértékek szignifikáns változására semmiképpen nem fognak vezetni.

A szerzők köszönetet mondanak az OTKA támogatásának, amely nélkül a dolgozatban javasolt módszer tesztjeit (pl. több ezer sokváltozós globális minimumhely-meghatározást) nem állt volna módjukban végrehajtani.

HIVATKOZÁSOK

- BARTA GY., HAJÓSY A. 1985, 1986: New procedure for determination of the gravity constant. pp. 5–12. Annales Univ. Sci. Budapestiensis, Sectio Geophysica et Meteorologica, Tomus I-II
- CHEN Y.T., COOK A. 1993: Gravitational experiments in the laboratory. Cambridge University Press, Cambridge, Table 8.1
- HAJAGOS B., STEINER F. 1992: „Built-in” statistics in geophysical instruments. Acta Geod. Geoph. Hung. 27 (2–4), 293–300
- HAJAGOS B., STEINER F. 1993: Determination of fault's position in mines using different norms in inversion. Annales Univ. Sci. Budapestiensis, Sectio Geophysica et Meteorologica 9, 101–146
- HAJAGOS B., STEINER F. 1995: Symmetrical stable probability distributions nearest lying to the types of the supermodel $f_a(x)$. Acta Geod. Geoph. Hung. 30 (2–4), 463–470
- STEINER F. 1990: A geostatistika alapjai. Tankönyvkiadó, Budapest, 363 p.
- STEINER F., HAJAGOS B. 1995: Determination of probability distribution types on ground of samples. Acta Geod. Geoph. Hung. 30 (2–1), 449–461
- STEINER F., HAJAGOS B., HURSÁN G. 1996: MFV-corrected variances. Geophysical Transactions (in press)
- SZÜCS P. 1995a: Theoretical and practical consequences of the global optimization methods. Acta Geod. Geoph. Hung. 30, 301–312
- SZÜCS P. 1995b: Áttekintés néhány olyan eltérésnormáról, amelyek minimalizálása inverziós algoritmus kidolgozását teszi lehetővé. Előadás. Elhangzott az MGE Általános Geofizikai Szakosztálya által rendezett, Geofizikai inverzió c. Országos Tudományos Anketon. Miskolc-Tapolca, 1995. december 12.
- TARANTOLA A. 1987: Inverse Problem Theory. Elsevier, Amsterdam, 613 p.

KÖZKELETŰ TÉVEDÉSEK A ROBUSZTUS STATISZTIKA ELMÉLETÉBEN ÉS GYAKORLATÁBAN

Steiner Ferenc

A statisztikai algoritmus hatásfokának anyaeloszlástípustól való jelentős függetlensége, azaz robusztussága, azonos az outlierekre (durva hibájú adatokra) való érzéketlenséggel, azaz

1 ROBUSZTUSSÁG = REZISZTENCIA.

A rezisztencia pusztá tényének ellenőrzése (akár csak néhány tetszőleges járulékos adat szerepeltetése a mérések valódi értékei mellett, és az így kapott inverzió eredményének egybevetése az outliermentes inverzió eredményeivel) tehát elegendő a robusztusság fokának a megítéléséhez.

A P_k - és P_k^* -normák minimalizálásán alapuló algoritmusok, közös nevükön MFV-eljárások a legkisebb négyzetes elv (azaz az L_2 -minimalizálás) speciális iteratív súlyozással megvalósított variánsai, azaz

2 MFV-ELJÁRÁSOK = SPECIÁLIS IRLS-MÓDSZEREK.

(IRLS: iteratively reweighted least squares, MFV: most frequent value).

3 MFV-ELJÁRÁSOK = SPECIÁLIS ML-MÓDSZEREK

(ML: maximum likelihood)

4 A SKÁLAPARAMÉTER-FORMULA MEGVÁLASZTÁSA CSEKÉLY FONTOSSÁGÚ A HELYPARAMÉTER TULAJDONSÁGAINAK (PL. ROBUSZTUSSÁGÁNAK) A SZEMSZÖGÉBŐL.

5 A REZISZTENCIA VIZSGÁLATOKOR ELHAGYHATÓK A KLASSZIKUS KATASZTRÓFAPONT- (BREAKDOWN BOUND) VIZSGÁLATOK.

(Az utóbbi fogalom definíciójára nézve ld. Andrews et al. immár csaknem negyedszázaddal ezelőtt megjelent, "Robust Estimates of Location" c. könyvét).

MEGJEGYZÉSEK A FENTI TÉVEDÉSEKHEZ

Ad. 1: A robusztusságot ma már kvantitatív mérőszámmal illik jellemezni. Ennek meghatározásában outlierok egyáltalában nem játszanak szerepet, így a robusztusság mértéke a rezisztencia mértékétől teljesen független lehet abban az értelemben, hogy két, A1-gyel és A2-vel jelölt statisztikai algoritmus (hibatípusok valamely várt gyakoriságára vonatkozó) r robusztusságára fennálló $r(A1) > r(A2)$ relációval egyidejűleg (valamely, szintén valószínűnek ítélt outlier-modell

esetén) a rezisztencia mértékére akár éppen a fordított reláció is teljesülhet.

Ad. 2: Az első (kezdetől fogva több ismeretlen mennyiség statisztikai jellegű számítására megfogalmazott) MFV-eljárás 1965-ben norma-extrémumhely keresésként definiáltatott, amelyben semmiféle iteratív súlyozás nem szerepel. A paramétereiktől való lineáris függés speciális esetében valóban sokáig kényelmes volt az MFV-eljárások normaminimum-kereséseit kész least

squares programokkal végrehajtani, de a fent említettek világossá teszik, hogy ezek az utóbbi esetek tekintendők az általános P_k - és P_k^* -normákra épülő statisztikus eljárások speciális eseteinek és nem megfordítva.

Ad. 3: A T helyparaméter (vagy paramétervektor) és az S skálaparaméter meghatározása minden elméletileg megalapozott statisztikai eljárásnál egyetlen algoritmuson belül, szimultán történik. Márpedig kimutatható, hogy ennek skálaparaméter meghatározásra vezető részalgoritmus (a klasszikus Gauss-eloszlás hipotézisét, mint egyetlen kivételt figyelmen kívül hagyva) sohasem vezet azonos formulával megadható feltételre a skálaparamétert illetően, ha egyszer az információvesztéséget minimalizáljuk (azaz MFV-eljárásokkal dolgozunk), máskor pedig a likelihood-függvény S szerinti parciális deriváltjára nézve követeljük szokásosan meg, hogy az zérus legyen. (Ez az utóbbi követelmény egyébként nem is vezethet rezisztens eljárásra, már magából a ML-elvből következően sem). Mivel a (T, S) értékpárt egyetlen algoritmus egyidejűleg szolgáltatja, a 3 tévedésben az egyenlőségjel akkor sem tehető ki, ha az algoritmus T -re vonatkozó

részalgoritmus betűről betűre megegyezik az MFL és az ML esetekre.

Ad. 4: A helyparaméter meghatározás robusztussági mérőszáma lényegesen különbözhet azonos T -meghatározási formula esetén is, ha a skálaparaméter formulák különbözők.

Ad. 5: Szükséges tudni az alkalmazni kívánt módszerekről, hogy azok maximum hány százaléknyi outliert tolerálnak még az eredmények szignifikáns változása nélkül. Kuriózum, de megemlítenő, hogy ez az n_{out}/n arány (n_{out} az outlierek száma, n az összes adatok száma) érdekes módon még a 0,5-öt is meghaladhatja bizonyos P_k , ill. P_k^* normák alkalmazása esetén (egy gravitációs inverzió analízisével kapcsolatban SZÜCS P. mutatott be erre példát az 1995 decemberében Miskolcon tartott „Geofizikai inverzió” témájú országos tudományos ankéton).

Megjegyzés: A tisztelt olvasónak, akit részletek, ill. bizonyítások is érdekelnek, annak a két dolgozatnak az áttanulmányozását ajánlom, amelyek a Geophysical Transactions 38. kötetében (1993-ban) a 193. oldaltól a 230. oldalig jelentek meg (HAJAGOS B. és a szerző tollából).