

A mérési adatokból nyert információk hibáinak csökkentése általunk ismételten generált többlethibáknak a mérési adatokra történő szuperponálásával¹

STEINER FERENC²

Kellően nagyszámú többlethiba-sorozatot szuperponálva a mérési adatrendszerre, az inverziók során kapott modellparaméter-adatok hiába nagyobb hibájúak az egyes esetekben, mint a szokásos inverziónál, mediánjuk (vagy csoportok mediánjainak mediánja) szignifikánsan pontosabb értéket szolgáltathat a modellparaméterekre, mint az eredeti inverzió (a hibák körülbelül feleződnek).

Az alapgondolat gyakorlati alkalmazhatóságát a dolgozat mikrogravitációs példán mutatja be. Kiderült, hogy ez a „többlethiba-módszer” jól használható modellparamétereket tud szolgáltatni akkor is, ha a szokásos inverzió használhatatlan eredményekre vezet. A szerző mind ez ideig ugyan nem tette vizsgálat tárgyává, de előzetes véleménye szerint a többlethiba-módszer sikeres alkalmazásához nem szükséges a „természetes” és „mesterséges” hibák típusának az egyezése.

F. STEINER: The decrease of the errors of information exhausted from measuring data by using the “method of surplus errors”

Although the errors of the model parameters increase if the primary measuring errors are artificially increased by superposition of surplus errors to the measured values, if this surplus error superposition, however, is carried out many times enough, the medians of the inversion-determined model parameters can be more (or even much more) accurate than the model parameter got by one single inversion of the originally measured data.

The practical application of this fundamental conception is shown in the present paper on a microgravimetric example. It turned out that this „surplus-error method” can give well applicable model parameters even if some model parameter values are fully unusable if these values are calculated directly on the basis of the original measuring data but using only one single inversion.

1. Bevezetés

Közhelynek számít az, hogy a számítástechnika rohamosan, immár több mint fél évszázada exponenciálisan (vagy annál is nagyobb mérvűen) fejlődik. Nem érdektelen ennek tükrében a geofizikában gyakran alkalmazott (és általában egyre bonyolultabbá váló) algoritmusokat az alábbi három osztályba sorolni.

- A) A direkt feladatok némelyike a gazdaságos számíthatóság határán mozog (a „forward problem” megfogalmazása olyan, hogy az algoritmus számítógépes megoldásának gépidőigénye még éppen nem lép túl egy ésszerű határt).
- B) Az inverzió egészen más jellegű problematikájával kapcsolatban most ennek csak az azonos bonyolultsági fokú direkt feladatoknál szükséges gépidő természetesen szignifikánsan nagyobb volta emelendő ki. Egy adott rövid időszak számítástechnikai szintje a bonyolultsági fok egy ésszerű limitjét ugyanúgy meghatározza, mint az A) esetben.
- C) Ha egy adott típusú inverzióknak egyszeri lefuttatása valamikor még éppen ésszerű gépidőigényűnek minősült (azaz B) kategóriájú volt), az néhány év (esetleg egy évtized) elteltével akár több százszor is lefutathatóvá válik, ugyanannyi gépidőt véve igénybe.

2. A többlethiba-módszer alapgondolata

Tegyük fel, hogy adekvát modell paramétereit keressük az inverzió során, mégpedig az algoritmust illetően ez az eset számítástechnikai szempontból az 1. pont C)-vel jelölt csoportjába tartozik — így gyakorlatilag megvalósítható az alábbi alapgondolat:

Kellően nagyszámú többlethiba-sorozatot szuperponálva a mérési adatrendszerre, az inverziók során kapott modellparaméter-adatok hiába nagyobb hibájúak az egyes esetekben, mint a szokásos inverziónál, mediánjuk (vagy csoportok mediánjainak a mediánja) szignifikánsan pontosabb értéket szolgáltathat a modellparaméterekre, mint az eredeti inverzió.

A jelen dolgozat szerzője bevallja, hogy ezt a 2002 januárjában eszébe jutott, egyszerűségében is merésznek tartott alapgondolatot először még a Miskolci Egyetem Geofizikai Tanszékén működő geostatistikai team tagjával is csak óvatosan közölte. Végül azonban nemcsak bátorítást kaptam, hanem HAJAGOS Béla kollégám számításai révén az alapgondolat Monte-Carlo-bizonyítását is, amiért hálás köszönetemet fejezem ki. Mivel a munka folytatását tervezzük, egy, a jelenleginél bővebb változatot kétszerzős: Hajagos–Steiner-cikk gyanánt szeretnénk az Acta Geodaetica et Geophysica Acad. Sci. Hung. című szaklapban megjelentetni, esetleg „War against error using the method of surplus errors” címmel.

¹ Beérkezett: 2002. április 23-án

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

3. Gyakorlati példa az alap gondolat alkalmazására

Helytakarékossági célból az alapelvet a STEINER (ed.) [1997] monográfia 254. oldalán már ábrán is bemutatott 2-D esetre tesszük vizsgálat tárgyává. Két vízszintes, csekély mélységű körszelvényű vágat helyét (tengelyeik felszíni vetületének egy önkényesen felvett — és a mérések során is origónak tekintett — ponttól mért t_1 és t_2 vízszintes távolságait), m_1 -gyel és m_2 -vel jelölt mélységüket, valamint R_1 és R_2 sugaraikat kívánjuk mikrogravitációs mérésekkel meghatározni egy olyan területen, ahol már régen felhagytak a bányászkozással, a közensűrűség $2,6 \text{ t/m}^3$ és a talajvízszint a vágatok alatt van. A méréseket az origótól indulva méterenként végezzük 18 m távolsáig. Így 19 db graviméteres adatot mérhetünk az origótól y_i távolságokban. Ha hiba nélkül lehetne mérni, JUNG [1961] szerint a következő g_i adatrendszert nyernénk:

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

Tételezzük fel, hogy 6 db modellparaméterünk pontos értékei a következők: $R_1=1,5 \text{ m}$; $m_1=7,5 \text{ m}$; $t_1=5 \text{ m}$; $R_2=1,5 \text{ m}$; $m_2=6,5 \text{ m}$; $t_2=13 \text{ m}$. A 19 db pontos g_i érték az origótól indulva az (1)-ből már egyszerűen számítható, amelyek rendre a következők: $-30,2$; $-34,0$; $-37,9$; $-41,7$; $-45,0$; $-47,7$; $-49,6$; $-50,9$; $-51,9$; $-52,8$; $-53,7$; $-54,4$; $-54,3$; $-53,0$; $-50,2$; $-46,2$; $-41,5$; $-36,5$; $-31,9 \mu\text{Gal}$. Méréskor persze ezeket az értékeket csak akarunktól független, első pillanatban ismeretlen nagyságú hibával terhelve tudjuk meghatározni. DUTTER [1986/87] szerint a geostatistikai hibatípus a következő sűrűségfüggvényű standard esetben:

$$f_{st}(x) = 0,75 \cdot (1+x^2)^{-5/2}, \quad (2)$$

(ez a STEINER [1990] 30. oldalán definiált $f_d(x)$ szupermodell $a=5$ -höz tartozó típusa,) aminek a következő eloszlásfüggvény felel meg (ld. STEINER [1990] 50. oldalán az (1-44) formulát):

$$F_{st}(x) = \frac{1}{2} + \frac{x}{2 \cdot \sqrt{1+x^2}} + \frac{x}{4 \cdot (\sqrt{1+x^2})^3}. \quad (3)$$

Az $F_{st}(x)$ függvény segítségével könnyen generálhatunk geostatistikus típusú véletlen számokat, s mivel ezek hibaértelműek lesznek, a helyparaméter zérus értéke miatt csak az S skálaparamétert kell figyelembe vennünk a standard (azaz az $S=1$ esetre vonatkozó) (3) formula általánosításakor, mégpedig egyszerűen úgy, hogy x helyére mindenütt x/S kerül.

Kérdés, mekkorára célszerű S -et választanunk. Ismert, hogy a mikrogravitációs mérések hibája legalább 2–3 μGal . Ha a hiba alatt a Q interszextilis félterjedelmet értjük, amely a geostatistikus eloszlásra standard esetben 0,55 (ld. STEINER [1990] 52. oldalán az $a=5$ oszlopot), az S -et legalább 4-nek kell választanunk. Gyanítható azonban, hogy a szakirodalom pontos definíció nélküli „hiba” alatt szórást ért, amelynek a standard alakú geostatistikus eloszlásra $1/\sqrt{2} = 0,707$ az értéke, így $S=4$ esetén 2,8 μGal adódik az így értelmezett hibára. Mivel a szakirodalom gyakrabban idézi a 3 μGal maximálisan elérhető pontosságot, a mérési hibák generálása mindig $S=4$ -gyel történt

(és persze geostatistikus típusú volt). 19-szer generálva így véletlen számokat, „mérési adatrendszerünk” az I. Táblázat szerintieknek adódtak (ha szükséges, megkülönböztetésül $g_i^{(m)}$ -mel jelöljük ezeket az adatokat).

A továbbiakban valódi mérési eredményekként kezeljük a 19 db adatot (természetesen akármelyik adatoszlopról legyen is szó.) Így első lépésként inverziót hajtunk végre a P -norma minimumhelyének meghatározásával, vagy ami ugyanaz: a legkisebb szorzatok elvének teljesítésével (ld. STEINER (ed.) [1997] 367. oldalát): megkeressük azokat az R_1, m_1, t_1, R_2, m_2 és t_2 értékeket, amelyek kielégítik a

$$\prod_{i=0}^{18} (4\varepsilon_k^2 + X_i^2) = \text{minimum} \quad (4)$$

követelményt, ahol az X_i az eltérést jelenti az (1) szerint számított és a $g_i^{(m)}$ mért érték között, miközben az ε_k dihézió egyidejűleg eleget tesz a következő egyenletnek:

$$\varepsilon_k^2 = 3 \cdot \frac{\sum_{i=0}^{18} \frac{X_i^2}{[X_i^2 + \varepsilon_k^2]^2}}{\sum_{i=0}^{18} \frac{1}{[X_i^2 + \varepsilon_k^2]^2}}. \quad (5)$$

A 6 db modellparaméter és az ε_k dihézió nyilván kettős iteráció eredményeképpen adódik, felváltva elégítve ki (4)-et és (5)-öt. — Az ε_k -ra adódó értékeket az I. táblázat utolsó sora tünteti fel.

A fentiek elvégzése után még csak a P -norma minimumhelyének vagyunk a birtokában: kezünkben van az egyetlen, de első inverzió eredménye, azaz 6 db modellparaméter ahhoz a k -adik mérési adatsorhoz, amelyet kiválasztottunk. Még eddig egyetlen szó sem esett tulajdonképpeni célunkról: a többlethiba-módszer alkalmazásáról.

Nem volt felesleges azonban áttekinteni ezt az első lépést kicsit részletesebben ahhoz, hogy a 2. pontban definiált új módszert most hogyan célszerű alkalmaznunk. Mivel STEINER [1990] 106. oldalán a geostatistikus eloszlás ($a=5$) standard esetére $\varepsilon=0,4818$ nagyságú dihézióértéket olvashatunk ki, a primer hibageneráláskor alkalmazott $S=4$ miatt az először mérési hibaként szuperponált véletlen számok dihéziója mindegyik k sorozatra $0,4818 \times 4 = 1,9272$ volt. Naivitás lenne azonban azt várnunk, hogy mindössze 19 adatból ε_k -ként ezt az értéket pontosan vissza is kapjuk — mivel azonban terepi mérés inverziójakor csak 19 adat (az I. táblázat valamelyik oszlopa) áll csak rendelkezésünkre, kénytelenek vagyunk az ehhez tartozó ε_k értéket dihézióként elfogadni. (A P -norma robusztussága miatt szerencsére még viszonylag nagy dihézióeltérések sem befolyásolják szignifikánsan a paraméterter-beli minimumhelyet.) De miért is fontos ez az ε_k érték? Hogy a többlethiba se túl kicsiny, se túl nagy ne legyen, célszerű ugyanolyan ε_k -vel generált, geostatistikus típusú „mesterséges” hibát az I. táblázatbeli mérési adatokra szuperponálni. Mivel ekkor a két azonos hiba összegének skálaparamétere másfélszeres lesz, végül a többlethibás adatokra elegendő egyszeres iterációval a következő feltételt teljesíteni:

$$\prod_{i=0}^{18} (9\varepsilon_k^2 + X_i^2) = \text{minimum} . \quad (6)$$

A (6)-ban azért kapja ε_k a kilences faktort, mert a geostatistikus típusnak ($\alpha=5$ -nek) a szimmetrikus stabilis eloszlástípusok $f_{\alpha}(x)$ családjában (ld. STEINER [1990] 34. oldal) a legjobban az $\alpha=1,677$ típusparaméterű felel meg [HAJAGOS, STEINER 1997], — márpedig ha két azonos α típusú és azonos skálaparaméterű valószínűségi változót

összeadunk, annak (az ugyancsak α típusú) eloszlásnak $2^{1/\alpha}$, azaz jelenleg $2^{1/1,677} = 3/2$ lesz a skálaparamétere, így most (a többlethiba hozzáadása után) a dihézió az ε_k $3/2$ -szerese lesz, s mivel a P -normában (és a hozzátartozó szorzatminimum elvben) a dihézió kétszeresének a négyzete szerepel, 4-gyel egyszerűsíthetünk és végül csak $3^2=9$ marad meg ε_k szorzójaként.

i	k	1	2	3	4	5	6	7	8	9	10
0		37,78	31,34	27,87	30,16	32,58	33,27	31,57	29,18	28,70	31,87
1		31,95	33,12	31,40	37,11	33,70	29,66	29,00	33,25	35,73	33,82
2		38,12	39,05	40,31	39,29	36,54	43,20	41,61	39,44	36,81	39,56
3		42,11	40,86	39,69	39,56	59,59	40,90	46,08	45,65	41,17	43,23
4		43,77	46,00	47,31	47,34	45,65	46,97	45,29	42,43	39,99	45,89
5		55,02	44,41	45,11	47,30	46,46	44,60	48,29	47,56	49,86	49,08
6		47,23	50,62	52,30	49,39	49,08	46,05	50,53	49,48	49,36	49,22
7		47,23	46,38	52,49	51,89	45,12	54,90	51,43	48,15	52,04	51,99
8		50,72	51,46	53,34	51,11	50,43	49,96	47,66	51,13	54,11	54,51
9		52,08	53,71	53,00	54,54	53,40	55,85	50,30	53,73	54,97	52,97
10		55,09	52,48	55,38	53,96	55,66	53,23	55,82	53,28	50,84	55,40
11		57,89	55,88	53,80	53,95	53,14	49,19	54,09	53,29	52,72	55,02
12		48,37	50,57	54,28	55,05	55,35	53,96	52,53	56,20	55,55	51,99
13		55,77	50,45	55,07	53,05	50,76	54,80	51,46	51,59	51,90	52,19
14		51,23	51,33	48,58	48,86	51,29	51,62	54,67	52,04	50,57	46,42
15		54,48	40,79	45,19	49,40	46,91	50,36	47,53	50,57	43,67	47,81
16		40,75	41,26	42,32	43,41	43,12	36,62	41,45	41,26	45,54	41,21
17		37,09	36,45	36,97	35,43	39,52	38,83	39,35	35,62	30,35	38,35
18		35,99	32,02	28,72	30,13	35,18	29,30	29,93	27,06	33,25	33,09
ε_k		1,39	1,96	1,70	0,91	1,23	2,79	2,06	2,02	2,55	1,30

I. táblázat. A vizsgálatok céljára a következő 2-D-modell szolgál: két vízszintes üres hengernek mint egy felhagyott bányá vágatainak mélységei és egyéb paraméterei határozandók meg mikrograviméteres mérések eredményeinek P -norma bázisú inverziójával. A fentiekhez szimulált, e táblázatban bemutatott „mérési” eredményeket $S=4$ skálaparaméterű statisztikus hiba többszöri generálásával és az egzakt adatokra való szuperponálásával nyertük, hogy a mikrogravimetriában reálisnak számító hibával terhelt adatrendszerek lehessenek vizsgálataink alapjai. Az összes adat negatív (ezért előjeleket nem tüntettünk fel) és μGal -ban értendő. Az utolsó sorban a megfelelő (k -val definiált) mérési adatrendszer egyetlen, P -norma alapú inverziójakor adódó dihézióértékek szerepelnek; egyaránt 5–5 ε_k érték nagyobb, ill. kisebb a dihézió elméleti értékénél (1,9272-nél)

Table I. Different data sets of simulated microgravimetric measurements consisting of 19 data each above the chosen 2-D model: two horizontal cylinder considered as two galleries in an old mine. The random numbers for errors were generated according to the statistical type using $S=4$ as parameter of scale to get errors in the scale of the real microgravimetric measurements. In the last row are given the corresponding ε_k values of the dihesion gotten as results of the inversion carried out on the basis of the P -norm; five ε_k values are greater, the other five ones are less than the theoretical value of the dihesion (=1.9272). All gravimetric values are given in μGal -s and are negative (therefore signs are not given)

Világos, hogy a (6) minimumhelyét — a nagyobb össz-hiba miatt — pontatlanabban kapjuk meg. Többlethibát azonban egyetlen mérési adatrendszerhez tetszőlegesen sokszor generálhatunk, így kellően sok így kapott modellparaméter mediánja már pontosabb értéket szolgáltathat az egyszeri iterációval kapottnál. Mivel vizsgált modellünk egyszeres inverziójának gépidejénél a ME geofizikai teamjének rendelkezésére álló számítástechnika akár három nagyságrenddel nagyobb gépidővel is még ésszerűen dolgozhat, azonos k -hoz 21-szer szuperponáltunk többlethibát, az így adódó 6 modellparaméter mindegyikére meghatároztuk a mediánt — és ezt az egész procedúrát 21-szer megismételve, végül a mediánok mediánjait fogadtuk el mint a többlethiba-módszer ered-

ményeit. Ez utóbbiakat, felülvonás alkalmazásával, $\bar{p}_{j,k}$ -val jelöltük, megkülönböztetésül az egyszeri inverzióval kapott $p_{j,k}$ paraméterekről. A p_j -k a helyes paraméterértékek, amelyekre terepi mérés esetén persze nem ismerünk, mostani vizsgálatainkban azonban célszerű felhasználnunk a többlethiba-módszer előnyös alkalmazhatóságát bemutató *II. táblázatban*. Itt mindegyik k -hoz és ezen belül mind a hat paraméterhez részletesen adottak a többlethiba-módszer alkalmazásával kapott, hibacsökkenést jellemző számértékek. Röviden úgy foglalhatjuk össze eredményeinket, hogy a dolgozat által javasolt módszerrel nagy valószínűséggel érhetjük el a paraméterhibák szignifikáns csökkenését.

k	p_j	$p_{j,k}$	$\overline{p_{j,k}}$	$ p_{j,k} - p_j $	$ \overline{p_{j,k}} - p_j $	$\frac{ \overline{p_{j,k}} - p_j }{ p_{j,k} - p_j } \cdot 100$
1	R_1	2,42	1,46	0,92	0,04	4,59%
	m_1	12,3	7,94	4,76	0,44	9,24%
	t_1	9,32	4,12	4,32	0,88	20,4%
	R_2	0,48	1,64	1,02	0,14	13,5%
	m_2	0,08	6,79	6,42	0,29	4,55%
	t_2	14,6	12,6	1,59	0,40	25,2%
2	R_1	1,00	1,46	0,50	0,04	8,24%
	m_1	5,95	8,02	1,55	0,52	33,7%
	t_1	3,26	4,66	1,74	0,34	19,5%
	R_2	1,91	1,58	0,41	0,08	20,5%
	m_2	8,48	7,47	1,98	0,97	49,0%
	t_2	11,6	12,6	1,43	0,45	31,3%
3	R_1	1,49	1,49	0,01	0,01	168%
	m_1	6,65	6,83	0,85	0,67	78,9%
	t_1	5,56	5,45	0,56	0,45	79,4%
	R_2	1,43	1,45	0,07	0,05	73,8%
	m_2	6,00	6,05	0,50	0,45	91,4%
	t_2	13,2	13,1	0,20	0,06	29,9%
4	R_1	1,42	1,58	0,08	0,08	98,0%
	m_1	6,96	7,94	0,54	0,44	81,3%
	t_1	4,42	5,06	0,58	0,06	11,1%
	R_2	1,55	1,42	0,05	0,08	158,0%
	m_2	6,46	5,97	0,04	0,53	1380%
	t_2	12,7	13,0	0,35	0,04	12,6%
5	R_1	0,55	1,38	0,95	0,12	12,1%
	m_1	0,05	8,18	7,45	0,68	9,15%
	t_1	3,28	4,32	1,72	0,68	39,6%
	R_2	2,40	1,73	0,90	0,23	25,2%
	m_2	11,5	8,01	4,98	1,51	30,3%
	t_2	10,1	12,6	2,90	0,36	12,4%
6	R_1	2,21	1,60	0,71	0,10	13,5%
	m_1	10,7	8,12	3,24	0,62	19,0%
	t_1	8,02	4,87	3,02	0,13	4,44%
	R_2	0,59	1,40	0,91	0,10	11,3%
	m_2	49	5,67	4,01	0,83	20,6%
	t_2	13,9	13,3	0,93	0,26	28,0%
7	R_1	1,39	1,50	0,11	0,00	2,84%
	m_1	6,24	6,71	1,26	0,79	63,0%
	t_1	4,13	4,59	0,87	0,41	47,3%
	R_2	1,57	1,46	0,07	0,04	56,1%
	m_2	6,33	5,92	0,17	0,59	336,0%
	t_2	13,1	13,3	0,09	0,32	373,0%
8	R_1	1,62	1,51	0,12	0,01	9,9%
	m_1	7,55	7,13	0,05	0,37	766,%
	t_1	5,29	4,87	0,29	0,13	44,2%
	R_2	1,31	1,40	0,19	0,10	50,1%
	m_2	5,08	5,31	1,42	1,19	83,8%
	t_2	13,3	13,1	0,32	0,15	45,7%
9	R_1	2,02	1,53	0,52	0,03	6,22%
	m_1	9,39	7,51	1,89	0,01	0,39%
	t_1	7,59	5,60	2,59	0,60	23,2%
	R_2	0,88	1,47	0,62	0,03	5,29%
	m_2	4,66	6,63	1,84	0,13	6,98%
	t_2	14,3	13,2	1,33	0,22	16,5%
10	R_1	1,21	1,52	0,29	0,02	8,20%
	m_1	6,97	8,08	0,53	0,58	109%
	t_1	4,16	5,18	0,84	0,18	21,1%
	R_2	1,85	1,60	0,35	0,10	28,0%
	m_2	8,54	7,94	2,04	1,44	70,6%
	t_2	12,0	12,7	1,02	0,33	32,3%

II. táblázat. A többlethiba-módszer előnyei. A p_j ($j=1,2,\dots,6$) modellparaméterek valódi értékei a következők: $R_1=1,5$ m; $m_1=7,5$ m; $t_1=5$ m; $R_2=1,5$ m; $m_2=6,5$ m és $t_2=13$ m. A táblázat első oszlopának k értékei az I. táblázatra (folytatás a következő oldalon)

utalnak, hogy t_i éppen melyik mérési adatrendszer a vizsgálat alapja. A második oszlop definiálja, hogy a táblázat adott sorában p_j alatt melyik modellparaméter értendő. A 3. oszlop $p_{j,k}$ értékei egyetlen inverzió eredményeit jelentik, ha a k -edik mérési adatoszlopot választottuk az I. táblázatból. A többlethiba-módszer eredményeit a negyedik oszlop tartalmazza $\bar{p}_{j,k}$ -val jelölve; ezek abszolút értékben értett eltéréseit a valódi értékektől a hatodik oszlop adja meg. Az egyetlen inverzió eredményeinek analóg különbségei az ötödik oszlopban adóttak. Utóbbiakhoz viszonyítva a 6. oszlop különbségeit, a százalékokban kifejezett és a 7. oszlopban számszerűleg megadott arányok kvantitatív tájékoztatást nyújtanak a dolgozatban javasolt módszernek az esetek túlnyomó többségében teljesülő előnyeiről. (A többszáz %-os esetek mindegyikében néhány ezrelékes, szinte elhanyagolható volt az első iterációs lépésben kapott modellparaméter-érték eltérése a valódi értéktől, míg a többlethiba-módszer 8% körüli, tehát nem irreálisan nagy hibára vezetett. A $k=3$ sor R_1 -hez tartozó 168%-os aránya az előző két, mm-ekben mérhető és a gépben így szereplő értékek hányadosa; mivel a táblázat csak cm-ekig tüntet fel értékeket, mindkét esetben a kerekített 0,01 érték szerepel)

Table II. This Table shows the advantages occurring in the overwhelming majority of cases using the method of surplus errors. The true values of the p_j ($j=1,2,\dots,6$) model parameters are the followings: $R_1=1.5$ m; $m_1=7.5$ m; $t_1=5$ m; $R_2=1.5$ m; $m_2=6.5$ m and $t_2=13$ m. The k value in the first column refers to Table I: which data set was actually chosen for the investigations. The second column gives the actual meaning of the p_j model parameter in the row in question. The $p_{j,k}$ values in the third column mean the results of one single inversion, carried out on the basis of the measured values given in the k^{th} column in Table I. The results of the method of the surplus errors are denoted by $\bar{p}_{j,k}$ and given in the 4th column. The absolute values of the difference $p_{j,k}$ minus the true value of p_j are in the 5th column; the analogous differences for $\bar{p}_{j,k}$ are given in the 6th column. The last column presents (in percent) the value of the decrease of the error of the model parameters if the surplus error method is used instead of one single inversion. (In the cases if some hundred % figure in the last column, the error of the result of the first iteration is only some per thousand whilst the error of the proposed method resulted in the acceptable values of about 8%. The ratio of 168% figuring for R_1 in row 3 of the table is the ratio of small values; after rounding them to cm-s, both have the value of 1 cm)

Nem maradhat megemlítés nélkül, hogy a fent idézett két ismétlési számot (a többlethiba-generálások, valamint a teljes procedúra ismétléseinek a számát) más irányú, de rokon gyakorlati tapasztalataink alapján adtuk meg. A későbbi vizsgálatok egyik célja éppen az lehet, hogy a mediánok mediánjának képzésekor milyen ismétlési számpár alkalmazása tekinthető optimálisnak. — Az olvasó joggal vethetné fel azt a kérdést, hogy ha egyetlen lépéssorozatban szuperponálnánk 400–500-szor többlethibát, az inverziókkal nyert modellparaméterek mediánja nem adódna-e hasonlóan pontosnak, mint a 21 darab, szintén 21 elemű paramétercsoportból meghatározott mediánoknak a $\bar{p}_{j,k}$ mediánja. Ilyen vizsgálat nem történt ugyan, de valószínűleg igenlő lenne a válasz a fenti kérdésre, annál is inkább, mert a paraméterek külön-külön történő mediánképzésével mindkét esetben feltételezhetően ugyanúgy csökken az egyes paraméterpárok közötti kapcsolat vélhetően káros hatása. Ha azonban a „21-szer 21-es” szisztémánál maradunk, a 21-es csoportok mediánjainak nemcsak a mediánja képezhető, hanem a nagyság szerinti rendezés után a negyedik elem a Q_a alsó szextilist, a 18-ik elem a Q_f felső szextilist szolgáltatja, amelyből nyilvánvalóan képezhetjük a mediánokra az interszextilis félterjedelmet $Q=0,5 \cdot (Q_f - Q_a)$ szerint, azaz kvantitatív adattal tudjuk jellemezni a $\bar{p}_{j,k}$ végeredményeink hibáját, utóbbi alatt a statisztikus ingadozások mértékét értve. Indokolt tehát a „ Q_a ; medián; Q_f ” értékhármassok számértékeinek áttekinthető megadása is, a hatféle modellparaméterre és mind a tíz, az I. táblázatban megadott mérési adatrendszerre (III. táblázat).

4. A többlethiba-módszer hatásosságának bemutatása a mérési adatok terében

4.1. A 21x21-es többlethiba-módszer eredmény-paramétereivel számított, valamint a valódi modellhez tartozó graviméteres adatok eltérése

A 3. pontbeli, paramétertérben megadott eredmények legegyszerűbb, ha nem is legtanulságosabb átfogalmazása a mérési adatok terére, ha kiszámítjuk a fenti (4.1) címben definiált különbségeket μGal -ban, és ezeket az értékeket mind a tíz esetre ($k=1,2,\dots,10$) és esetenként mind a 19 mérési pontra ($i=0,1,2,\dots,18$) megadjuk (ld. a IV. Táblázatot, ahol a különbségi értékek μGal -ban adóttak). Megállapíthatjuk, Q_f hogy a 21x21-es többlethibamódszer eredményparamétereivel számított graviméteres adatok a hibamentes mérési adatoktól csak a mikrogravitációs mérések reális (legalább 2–3 μGal -nyi) mérési hibájának legtöbbször néhányszor 10%-nyi részével térnek csak el.

4.2. Milyen mérvű mérési hibacsökkentés volna szükséges ahhoz, hogy egyetlen inverzió a 21x21-es többlethiba-módszernek a 3. pontban már megismert paraméter-pontosságához hasonlót szolgáltatasson?

Gondolatmenetünket indítsuk a mérési adatokból kiindulva, ld. pl. az I. ábrán a 19 db mérési pontot, amely a $k=9$ esetnek felel meg (ld. az I. táblázatot). Végezzünk kiegyenlítést az (1) formula analitikus alakját használva, hogy a mérési pontoknak jól megfelelő sima görbét kapjunk. Az analitikus formulában szereplő 6 paraméteradat valamely statisztikai norma alapul választása után (esetünkben ez a P -norma) azon 6 értéként kapjuk, amelyekkel definiált folytonos görbe mérési pontokbeli értékeinek és a mérésről kapott adatoknak az eltéréseire minimális a választott eltérés-norma (itt a P -norma). Így az 1. ábrán vékony vonallal ábrázolt folytonos görbét kapjuk, ami

„ránézésre” is jól egyenlíti ki a mérési pontokat, azaz az egyetlen minimumhely-kereséssel kapott 6 érték, mint kiegyenlítési paramétervektor, kielégítően „vizsgázott” a mérési adatok terében.

	$R_1=1,5$	$m_1=7,5$	$t_1=5$	$R_2=1,5$	$m_2=6,5$	$t_2=13$
$k=1$	1,42	7,70	3,97	1,60	6,66	12,5
	1,46	7,94	4,12	1,64	6,79	12,6
	1,47	8,21	4,31	1,65	6,95	12,7
$k=2$	1,43	7,77	4,40	1,56	7,27	12,6
	1,46	8,02	4,66	1,58	7,47	12,6
	1,47	8,22	4,80	1,63	7,61	12,6
$k=3$	1,47	6,60	5,35	1,43	5,87	13,0
	1,49	6,83	5,45	1,45	6,05	13,1
	1,50	6,87	5,55	1,47	6,26	13,2
$k=4$	1,54	7,75	4,87	1,35	5,87	12,9
	1,58	7,94	5,06	1,42	5,97	13,0
	1,65	8,10	5,45	1,46	6,17	13,2
$k=5$	1,33	8,02	4,06	1,70	7,85	12,6
	1,38	8,18	4,32	1,73	8,01	12,6
	1,42	8,38	4,44	1,78	8,24	12,7
$k=6$	1,57	7,93	4,74	1,36	5,37	13,2
	1,60	8,12	4,87	1,40	5,67	13,3
	1,64	8,31	5,14	1,42	5,93	13,4
$k=7$	1,49	6,54	4,41	1,45	5,79	13,2
	1,50	6,71	4,59	1,46	5,92	13,3
	1,52	6,97	4,66	1,49	6,02	13,4
$k=8$	1,50	6,94	4,79	1,38	5,20	13,1
	1,51	7,13	4,87	1,40	5,31	13,1
	1,54	7,41	4,95	1,43	5,53	13,3
$k=9$	1,52	7,40	5,50	1,45	6,52	13,1
	1,53	7,51	5,60	1,47	6,63	13,2
	1,56	7,71	5,85	1,48	6,84	13,3
$k=10$	1,49	7,91	5,01	1,57	7,77	12,6
	1,52	8,08	5,18	1,60	7,94	12,7
	1,54	8,27	5,28	1,64	8,20	12,9

III. táblázat. Mind a hat modellparaméterhez és mind a tíz (az I. táblázatban megadott) mérési adatrendszerhez adottak a 21-szeri többlethiba generaláskor kapott paraméter-mediánok alsó és felső szextilisei (mint az adathármasok első és harmadik értékei); a második érték e mediánok mediánja. Ezek az értékek a többlethiba-módszer jellemzésére pontosnak tekinthetők, hiszen az egész procedúra szintén 21-szeres ismétlésének eredményei alapján történt a meghatározásuk

Table III. Three values are given for all six model parameters and for all ten measuring data systems given in Table I: the first one means the lower, the third one the upper sextile value of the medians, the second one the median of the medians superposing 21 times surplus errors. Repeated this latter procedure 21 times, too, these three values can be determined accurately enough

Jól tudjuk azonban, hogy ennek a hat adatnak fizikai, ha tetszik, modellgeometriai értelme van: két hengerkeresztmetszetű vágat sugarait, valamint tengelyük mélységeit és e tengelyek felszíni vetületeinek origótól mért távolságait jelentik. Az 1. ábrán a (kiegyenlítéskor kielégítő eredményt adó) 6 paraméterértékhez vékony vonalú körök adják meg a hozzájuk tartozó vágat-metszeteket. Az eredmény annyira rossz, hogy már félrevezetőnek sem nevezhető. (A jobb felső vágatot, kis keresztmetszete miatt, leginkább Hófehérke törpéi tudták volna valaha kihajtani...) Láttuk viszont, hogy a 21x21-es többlethiba-módszer ekkor is a valódi paraméterértékekhez közeliekhez vezet (ld. a II. táblázatban a megfelelő különbségi értékeket $k=9$ esetére); az első ábra vastag vonallal rajzolt körei egyrészt valóban közel vannak a valódi, azaz az $R_1=1,5$ m; $m_2=7,5$ m; $t_1=5$ m; $R_2=1,5$ m; $m_2=6,5$ m; $t_2=13$ m modellparaméterekkel definiált esethez, — másrészt a vastag körök szerinti üres vágatok (1) szerint számolt graviméteres hatás-

görbéje, amint azt a vastag vonalú görbe mutatja az 1. ábrán, vizuálisan megítélve hasonlóan jól egyenlíti ki a pontokkal jelölt mérési eredményeket, mint az egyetlen inverzióknak megfelelő, vékony vonalú görbe. — A 4.2. alpont címében feltett kérdést, legalábbis $k=9$ -re, most az 1. ábrára utalva fogalmazhatjuk át: *a mérési hibák milyen mérvű csökkentése vezetne (a vékony vonalú körök helyett) egyetlen inverzióval is a valóságos határhoz oly mértékben közeli eredményhez, mint az 1. ábra vastag vonallal rajzolt körei, amelyeket a 21x21-es többlethiba-módszer szolgáltatott?*

A „közelséget”, általában a valódi esettől való távolságot persze célszerű volna kvantitatíve is definiálni, hogy áttekinthetőbben lehessen a fenti kérdést megválaszolni. Az inverzió irodalma szerencsére használja a modell-távolság fogalmát, s így például GYULAI [2000]-rel (ld. 108. oldal) összhangban vagyunk, ha az akármilyen hibájú mérési adatokból az éppen alkalmazott algoritmussal adódó számí-

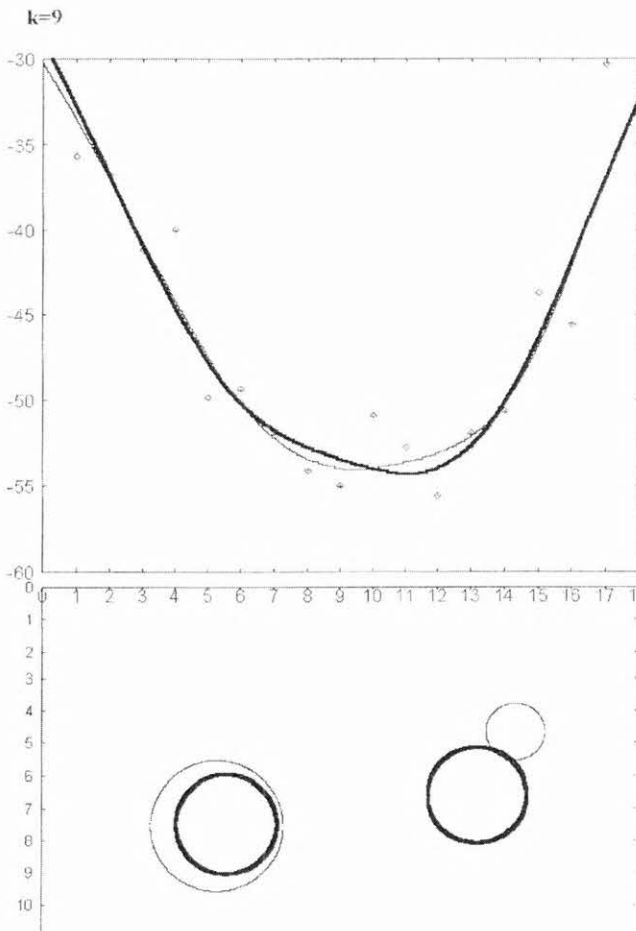
tott modellparaméter-értékeket $p_j^{számított}$ -tal jelölve, a valódi hatótól a paraméter-térben értelmezett (és Δ -val jelölt) távolságot a következőképpen definiáljuk:

$$\Delta = \sqrt{\frac{1}{J} \sum_{j=1}^J \left(\frac{p_j^{számított} - p_j^{valódi}}{p_j^{valódi}} \right)^2} \quad (7)$$

i	k	1	2	3	4	5	6	7	8	9	10
0		-2,488	-1,022	1,945	-0,491	-1,396	-0,793	-1,113	0,707	1,181	-1,905
1		-2,240	-0,724	1,646	-0,334	-0,932	-0,492	-1,676	0,563	1,197	-1,791
2		-1,803	-0,334	1,137	-0,128	-0,341	-0,068	-2,191	0,432	1,092	-1,667
3		-1,224	0,096	0,439	0,098	0,295	0,448	-2,488	0,377	0,828	-1,583
4		-0,636	0,477	-0,345	0,303	0,834	0,989	-2,387	0,468	0,405	-1,598
5		-0,227	0,710	-1,028	0,443	1,124	1,468	-1,805	0,740	-0,107	-1,736
6		-0,164	0,748	-1,429	0,496	1,079	1,812	-0,840	1,148	-0,575	-1,951
7		-0,505	0,638	-1,477	0,468	0,751	1,980	0,251	1,551	-0,859	-2,115
8		-1,166	0,518	-1,259	0,367	0,318	1,946	1,173	1,749	-0,878	-2,058
9		-1,955	0,559	-0,953	0,190	0,004	1,673	1,699	1,534	-0,647	-1,640
10		-2,643	0,881	-0,725	-0,073	-0,034	1,110	1,708	0,775	-0,267	-0,830
11		-3,040	1,471	-0,643	-0,399	0,199	0,262	1,209	-0,467	0,116	0,238
12		-3,076	2,143	-0,641	-0,688	0,501	-0,708	0,379	-1,802	0,358	1,261
13		-2,832	2,612	-0,567	-0,791	0,579	-1,467	-0,449	-2,600	0,381	1,883
14		-2,479	2,664	-0,311	-0,621	0,239	-1,709	-0,935	-2,420	0,212	1,893
15		-2,165	2,283	0,106	-0,241	-0,482	-1,406	-0,947	-1,398	-0,049	1,340
16		-1,945	1,630	0,561	0,193	-1,374	-0,795	-0,604	-0,071	-0,294	0,476
17		-1,801	0,911	0,940	0,548	-2,206	-0,160	-0,127	1,088	-0,466	-0,425
18		-1,695	0,271	1,193	0,774	-2,837	0,335	0,315	1,880	-0,557	-1,185

IV. táblázat. A $\bar{p}_{j,k}$, valamint a helyes p_j modellparaméterekkel számított mikrogravitációs hatások különbségei

Table IV. Differences of microgravimetric effects calculated according to the $\bar{p}_{j,k}$ and to the true p_j model parameters



I. ábra. Az ábrán a pontok az I. táblázat $k=9$ mérési adatoszlopának felelnek meg. Az ezek alapján egyetlen inverzióval nyert hengerkeresztmetszeteket vékony vonallal rajzolt körökkel, a 21×21 -es többlethiba-módszerrel kapottakat vastag vonalú körökkel mutatja be az ábra. Mindkét hatásgörbe, vizuálisan megítélve, hasonlóan jól egyenlíti ki a mérési pontokat

Fig. 1. The points demonstrate the measured data series given in the $k=9^{\text{th}}$ column of Table I. The results of the single step inversion are demonstrated by two thin line circles (as the section of both galleries); the circles drawn by thick line are the results of the method of surplus errors carrying out 21 times 21 superpositions of surplus errors. Both curves of the gravimetric effect seem to fit similarly the original measured points

A jelen dolgozatban persze $J=6$ és a p_j -k jelentésére is emlékezünk: a két vízszintes henger sugarai, tengelyvonalai mélységei és egy rögzített ponttól való vízszintes távolságai a modellparaméterek. A definíció relatív jellege miatt Δ -t célszerű %-okban megadni; az 1. ábra vékony vonalú köreire pl. $\Delta=34,4\%$, míg a vastag vonallal rajzoltakra mindössze $\Delta=5,18\%$. Milyen mérvű mérési hibacsökkentés szükséges ahhoz, hogy az utóbbi Δ értéket kapjunk egyetlen inverzióval?

A $k = 9$ esetre az egyes pontok h_i -vel jelölt primer mérési hibáit nyilván úgy kapjuk, hogy az I. táblázat kilencedik oszlopának minden számértékéből levonjuk az (1)-ből a valódi modellparaméterértékekkel adódó hibamentes értéket; utóbbiak az (1) képletet követő szövegsorokban már kiszámított értékekkel is adóttak. A továbbiakban ezekhez a hibamentes értékekhez $c \cdot h_i$ -ként kapott hibákat adunk hozzá, ahol c kisebb 1-nél, majd elvégezve az egyetlen iterációt, az így adódó modellparamétereket helyettesítjük $p_j^{\text{számított}}$ -ként a (7) szerinti kifejezésbe; az eredményt célszerű $\Delta(c)$ -vel jelölni. Megkeresve azt a c értéket, amely mellett $\Delta(c)$ az 5,18%-os értéket veszi fel (emlékezünk, hogy ez az utóbbi érték a 21x21-es többlethiba-módszerrel kapott eredmény Δ -ja volt), ez a c érték az a mérési hibacsökkentési arány, amelyet keresünk. A számítások a $c=0,199$, azaz kereken a $c=0,2$ értéket eredményezték, s mivel a primer hibageneráláskor igazodtunk a jelenlegi mikrogravitációs mérések legkedvezőbb esetben is 2–3 μGal körüli hibájához, a kapott c érték azt jelenti,

hogy kb. 0,4–0,6 μGal körüli mérési hibák engedhetők csak meg. *A többlethiba-módszer tehát a vizsgált esetben a mért mikrogravitációs értékek olyan mérvű hiba-csökkentésével egyenértékű, amelyet a mikrogravitációs mérés technika feltehetően még a távolabbi jövőben sem fog tudni produkálni.* — Valószínű, hogy számos más geofizikai feladatnál is hasonlóan jó szolgálatot tehet a többlethibák módszere, de ezt természetesen minden esetben külön kell vizsgálat tárgyává tenni.

HIVATKOZÁSOK

- DUTTER R. 1986/87: Mathematische Methoden in der Montangeologie. Vorlesungsnotizen. Manuscript, Leoben
- GYULAI Á. 2000: 2-D földtani szerkezetek interpretációja együttes geoelektromos inverziós módszerekkel. MTA doktori dolgozat. Kézirat. Miskolc
- HAJAGOS B., STEINER F. 1997: Symmetrical stable probability distributions nearest lying to the types of the supermodel $f_a(x)$. Appendix VIII. in STEINER F. (Ed.) 1997: Optimum Methods in Statistics. Akadémiai Kiadó, Budapest
- JUNG K. 1961: Schwerkraftverfahren in der angewandten Geophysik. Akademische Verlagsgesellschaft, Leipzig
- STEINER F. 1990: A geostatistika alapjai. Tankönyvkiadó, Budapest
- STEINER F. (Ed.) 1997: Optimum Methods in Statistics. Akadémiai Kiadó, Budapest