

EGYENLŐTLENSÉGEK NUMERIKUS BIZONYÍTÁSA, I.*

Írta: RUDA MIHÁLY

In memoriam Ottó Varga (1909—1969)

1. Bevezetés

A nagy sebességű és nagy kapacitású elektronikus számológépek megjelenésével elvileg új lehetőségek nyíltak meg a matematika elméleti továbbfejlesztésének területén. Egy ilyen lehetőség matematikai tételek, állítások gépi segítséggel történő bizonyítása. Bizonyításon Hilbert-féle értelemben vett (l. [3]) bizonyítást értünk.

Az utóbbi időben olyan, gépre alkalmazható bizonyítási eljárásokat is kidolgoztak, melyek segítségével nem csak a bizonyítás egyes részfeladatai oldhatók meg, hanem amelyek lehetővé teszik matematikai tételek bizonyításának teljes automatizálását — elvileg.

Bár CHURCH tétele szerint a formalizált matematikán belül nincs univerzális eldöntési eljárás, mégis vannak már olyan bizonyítási módszerek, melyek teljesnek mondhatók azon az alapon, hogy valahányszor egy igaz tétellel kerülünk szembe, mindannyiszor lehetővé teszik egy bizonyítás előállítását. Ilyen, különböző bizonyítási eljárások kidolgozásában értek el eredményeket többek között GELERNTER [2] és NEWELL [5], illetve DAVIS [1], PRAWITZ [6] és ROBINSON [7].

Ezen a területen elért eredmények összefoglalását adja MELTZER egy 1968-as dolgozata [4], melyben különböző tételbizonyítási eljárások szerepelnek.

Ezek a bizonyítási módszerek a vizsgált matematikai részterület formalizálásával, nagyszámú logikai művelet elvégzésén keresztül adják a kérdéses tétel bizonyítását. A kidolgozott módszerek azonban még elég nehézkesek és végrehajtásuk rendkívül időigényes. Ezek a hátrányok éppen abból erednek, hogy a fenti módszerek általános jellegűek.

POGÁNY CSABA hívta fel a szerző figyelmét arra, hogy automatizált bizonyítási eljárások nemcsak általános formában — logikai módszerek segítségével —, hanem speciális matematikai részterületeken is alkalmazhatók, például függvények közti egyenlőtlenségek bizonyításánál. Ugyanakkor felvetette azt a kérdést is, hogy ezen a speciális területen kívül, általában melyek azok a feladattípusok, ahol szintén lehetőség nyílik egzakt bizonyítási módszerek automatizált alkalmazására.

A következőkben néhány általános jellegű megjegyzés után konkrét numerikus módszerek szerepelnek, melyek függvények közti egyenlőtlenségek bizonyítására szolgálnak. Ezek a módszerek, és általában az automatizált bizonyítási módszerek a következő általános elv szerint működnek:

Adott egy véges számú elemből álló H halmaz, az *alaphalmaz*. Ennek a halmaznak az elemeiből egy adott A algoritmus alapján képezünk egy újabb G halmazt

* Ez a dolgozat a szerzőnek „*A számítástechnika alkalmazása új tudományterületeken*” című kollokviumon 1969. június 5-én tartott előadását tartalmazza.

($G = A(H)$). Nevezzük ezt *generált* halmaznak! Ezután a G halmazt, illetve annak elemeit vizsgáljuk valamilyen előre adott szempont szerint.

A bizonyítási eljárást mint egy automatizált folyamatot tekintve elmondható, hogy a H alaphalmaz az automatába bemenő paraméterek (jelek) halmaza, az A algoritmus magának az automatának, vagyis a megfelelően programozott számológépnek felel meg, míg a generált halmaz a kimenő jelekből áll. Ezután a sorra jövő vizsgálat számára a G halmaz már mint bemenő halmaz szerepel. Ennek viszont már csak egyetlen kimenő eleme van, mely jelezheti a tétel igazságát vagy hamisságát.

Ezek a bizonyítási eljárásokon belül még kétféle típust választhatunk szét:

1. Az alaphalmazból kiindulva egy egyértelműen meghatározott G halmazt képezünk egy rögzített algoritmus szerint. Ennek a G halmaznak az előállítása után végezzük el a kitűzött vizsgálatot, melynek eredménye az általunk bizonyítandó tétel igazságát vagy hamisságát mutathatja.

Előfordulhat azonban, hogy az eljárás véghezvitele után sem tudjuk a tétel igazságát vagy hamisságát kimutatni.

2. Az alaphalmazból generált halmaz minden egyes elemének előállítása után elvégezzük a vizsgálatot és az újabb elem előállítása a vizsgálat eredményének függvényében történik.

Ebben az esetben nincs biztosítva az algoritmus véges volta.

Tekintsünk ezután egy konkrét példát! A kiindulási H halmaz lehet például egy axiómarendszer axiómáinak, illetve azokból levezethető néhány tételnek a halmaza. Ezen halmaz elemeiből — melyeket megfelelő módon formalizáltunk — logikai műveletek segítségével újabb kijelentéseket képezünk, melyek a G halmazt fogják alkotni. Ezeknek a kijelentéseknek a halmazán belül keressük aztán azt a kijelentést, amelyet előre megadtunk és bizonyítani szeretnénk.

Természetesen nemcsak kijelentéseket, hanem tetszőleges más objektumokat — például számértékeket — is vizsgálhatunk ilyen módon. Általában a kimondott (vagyis a bizonyítandó) tétel éppen ezeknek az objektumoknak valamilyen tulajdonságával kapcsolatos. Az előző esetben ez a tulajdonság egyszerűen egy előre adott objektummal (a bizonyítandó tétellel) való azonosság: ha a generált elemek között megtaláljuk a bizonyítandó kijelentéssel azonos állítást, akkor kijelentésünk igaz, ha az ellentettjét, akkor hamis.

Előfordulhat, hogy nem az egyes elemeket, hanem az egész generált halmazt vizsgáljuk, például olyan szempontból, hogy hány elemből áll.

Ezek után már rátérünk a függvények közti egyenlőtlenségek bizonyításánál használható eljárások vizsgálatára.

2. Bizonyítási eljárások szerkezete egyenlőtlenségek vizsgálatánál

Vizsgálunk tehát egy $f(x) > g(x)$ $x \in M$ egyenlőtlenséget. (Az $f(x)$ és $g(x)$ egy tetszőleges korlátos és zárt M halmazon értelmezett valós függvény.)

A H alaphalmazba az M értelmezési tartományt meghatározó elemek, a függvények helyettesítési értékeinek kiszámításához szükséges paraméterek, a számítás pontosságát jellemző ε érték, valamint még néhány, a függvényekre jellemző adat — például a deriváltakra adható korlátok — tartozik. Ezeknek az elemeknek a segítségével a megfelelő módon — véges sok (n) részre — felosztott értelmezési tartomány minden egyes részén egy alsó becslést adunk a nagyobbban feltételezett $f(x)$

függvényre, illetve egy felső becslést a kisebbnek feltételezett $g(x)$ függvényre. Az így nyert függvények alkotják a G generált halmazt:

Ha $\bigcup_{i=1}^n M_i \supseteq M$ és $f_i(x) < f(x)$, $g_i(x) > g(x)$ ($x \in M_i$), akkor legyen

$$G = \{f_i(x), g_i(x); i = 1, 2, \dots, n\}.$$

A vizsgálat abból áll, hogy megnézzük, minden i -re igaz-e az $f_i(x) > g_i(x)$ $x \in M_i$ reláció.

Ha valóban minden i -re teljesül az egyenlőtlenség, akkor a tétel, azaz a bizonyítandó egyenlőtlenség is igaz, ha nem, akkor még nem feltétlenül hamis az állítás (vagyis az egyenlőtlenség), hiszen itt csak becslésekről van szó.

Ilyen bizonyítási eljárásokon belül az $f_i(x)$ és $g_i(x)$ függvények meghatározásakor a következő konkrét becslések alkalmazhatók.

3. Néhány becslési eljárás

Elsőként egy $[a, b]$ zárt intervallumon differenciálható $f(x)$ és $g(x)$ függvény közti

$$f(x) > g(x) \quad x \in [a, b]$$

egyenlőtlenség bizonyításának módszerével foglalkozunk.

Ha ismerjük az $[a, b]$ intervallumon a két függvény deriváltjára adható korlátokat, melyek:

$$K_1 > f'(x) > k_1, \quad K_2 > g'(x) > k_2, \quad (x \in [a, b]),$$

akkor kiszámítva egy x_i pontban ($x_i \in [a, b]$) az $f(x)$ és $g(x)$ függvények értékeit, lineáris függvényekkel becsülhetjük a két függvényt. (Ezek a lineáris függvények lesznek a G halmaz elemei.)

A becslési eljárás a következő egyenlőtlenségeken alapszik:

Ha $f(x) > g(x)$ és $\Delta x > 0$ ($x, x + \Delta x \in [a, b]$) és ha

$$(1) \quad \Delta x < \frac{f(x) - g(x)}{K_2 - k_1},$$

akkor $f(x + \vartheta \Delta x) > g(x + \vartheta \Delta x)$ is igaz, ahol $0 \leq \vartheta \leq 1$ tetszőleges rögzített érték (1. ábra).

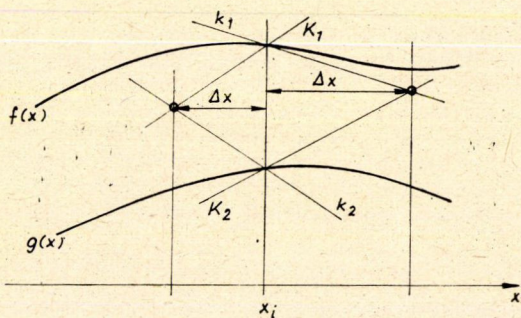
Hasonlóan, ha $f(x) > g(x)$ és $\Delta x < 0$ ($x, x + \Delta x \in [a, b]$) és ha

$$(2) \quad \Delta x > \frac{f(x) - g(x)}{k_2 - K_1},$$

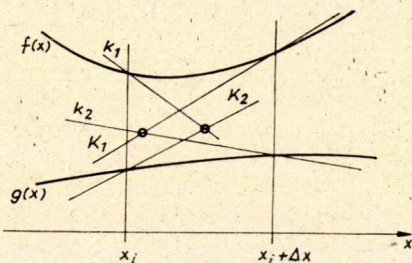
akkor $f(x + \vartheta \Delta x) > g(x + \vartheta \Delta x)$, ahol $0 \leq \vartheta \leq 1$ (1. ábra).

Végül, ha $0 < \Delta x$ olyan, hogy

$$(3) \quad \frac{f(x) - g(x)}{K_2 - k_1} + \frac{f(x + \Delta x) - g(x + \Delta x)}{K_1 - k_2} > \Delta x,$$



1. ábra



2. ábra

ahol $x, x + \Delta x \in [a, b]$, $f(x) > g(x)$ és $f(x + \Delta x) > g(x + \Delta x)$, akkor $f(x + \vartheta \Delta x) > g(x + \vartheta \Delta x)$, $0 \leq \vartheta \leq 1$ (2. ábra).

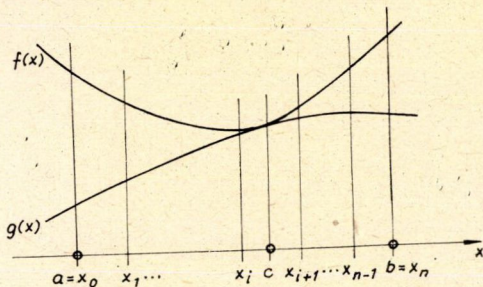
A bizonyítási eljárás — a bevezetésben adott — kétféle módon történhet:

1. n egyenlő részre osztjuk az $[a, b]$ intervallumot x_i osztópontok segítségével, és megvizsgáljuk, hogy a $\Delta x = x_{i+1} - x_i$ értékek minden i -re kielégítik-e az (1), (2) vagy (3) egyenlőtlenségek valamelyikét. Ha igen, akkor a bizonyítandó egyenlőtlenség is igaz a teljes $[a, b]$ intervallumon.

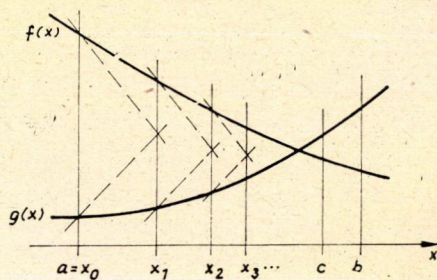
2. Az x osztópontokat az (1) vagy a (2) egyenlőtlenséget az x_{i-1} pontban kielégítő Δx érték segítségével adjuk meg úgy, hogy $\Delta x = x_i - x_{i-1}$, ahol $x_1 = a$ vagy $x_1 = b$. Tulajdonképpen az eredeti intervallum valamely végpontjából kiindulva a lehető legritkábban helyezzük el az x_i osztópontokat úgy, hogy a köztük levő Δx távolságok az (1) vagy a (2) egyenlőtlenséget még kielégítsék. Természetesen kiindulhatunk egyidejűleg az $[a, b]$ intervallum mindkét végpontjából is (ekkor mondjuk $x_1 = a$, $x_2 = b$ és $x_{i+2} = x_i + \Delta x$). Ha véges sok lépésben le tudjuk fedni ilyen $[x_i, x_{i+1}]$ (illetve $[x_i, x_{i+2}]$) intervallumokkal az $[a, b]$ intervallumot, akkor szintén igaz az $f(x) > g(x)$ ($x \in [a, b]$) egyenlőtlenség.

Ha $f(x) > g(x)$ ($x \in [a, b]$) (és az 1. változatnál az előre adott felosztás elég sűrű), akkor ezekkel a becslésekkel véges sok lépésben kimutatható az egyenlőtlenség. Ha van olyan c pont $c \in [a, b]$, melyben $f(c) \leq g(c)$, akkor az 1. módszernél mindig lesz olyan részintervallum, melyben nem teljesül az (1), (2) és (3) egyenlőtlenségek egyike sem (3. ábra), míg a 2. módszernél az x_i pontoknak egy konvergens sorozatához jutunk (4. ábra).

Ha olyan c pont is létezik, melyben $g(x) > f(x)$, akkor az 1. típusú eljárásoknál,



3. ábra



4. ábra

ha az intervallum felosztása elég finom, van olyan x_i osztópont, hogy $f(x_i) < g(x_i)$. Így kimutatható a bizonyítandó $f(x) > g(x)$ $x \in [a, b]$ egyenlőtlenség hamissága.

Természetesen gyakorlatilag nem lehetséges tetszőlegesen közeli függvényértékek közti egyenlőtlenség kimutatása, mivel a becsléseknél figyelembe kell venni a számítási pontosság korlátozottságából adódó hibát. Általában a számítás pontosságának kérdése központi szerepet játszik eljárásainknál (hiszen numerikus módszerekről van szó).

Az eddig tárgyalt két eljárástípus közül az elsőnek előnye, hogy csak az $f(x)$ és $g(x)$ függvényértékek különbségét kell kiszámítani az előre adott x_i pontokban, és ezt az értéket kell például az $(x_{i+1} - x_i)(K_2 - k_1)$ értékhez hasonlítani (ha az (1) egyenlőtlenséget használjuk). A második esetben a függvényértékek különbségén kívül minden alkalommal (minden i -re) a következő osztópont helyét is meg kell határozni, viszont a deriváltakra előírt korlátok adott értéke mellett az osztópontok a legtrikábban helyezkednek el, azaz eljárásunk a lehető legkevesebb lépésből áll. Az első módszernél még az is előfordulhat, hogy ha nem elég sűrűn helyezük el az osztópontokat, akkor az eljárást egy finomabb felosztás mellett meg kell ismételni. Ez a második módszernél nem fordulhat elő.

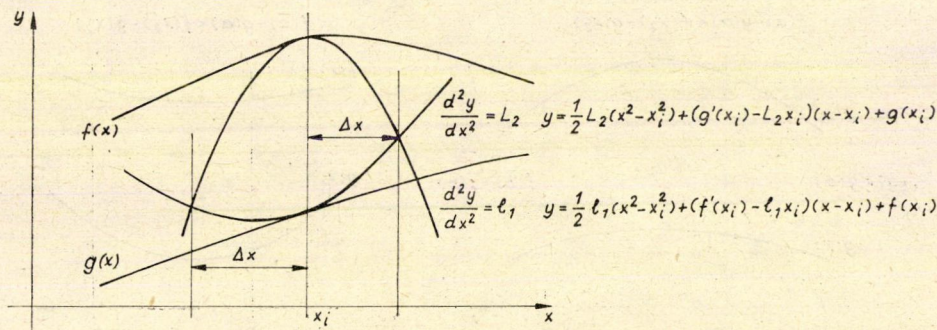
Mindkét eljárás hátránya, hogy ha a deriváltakra adott korlátok között nagy a különbség (5. ábra), akkor az (1), (2) és (3) egyenlőtlenségek által adott Δx értékek igen kicsik, és így az eljárás lassú, vagy a számítási pontosság korlátozottsága miatt leáll.

A lineáris függvényekkel való becslésen kívül más módszereket is használhatunk. Általában egy adott x_i pontbeli függvényérték (esetleg még más jellemző érték) kiszámításával adunk becslést a függvényekre, ismerve változásuk maximális mértékét.

Például, ha egy x_i pontban a függvényértékeken kívül a deriváltak értékeit is kiszámítjuk, és ha ismerjük a második deriváltakra adható korlátokat:

$$L_1 > f''(x) > l_1, \quad L_2 > g''(x) > l_2, \quad (x \in [a, b]),$$

akkor parabolákkal helyettesíthetjük a függvényeket (6. ábra).



6. ábra

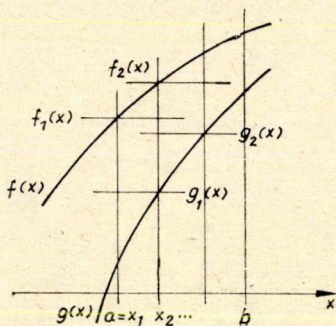
Megjegyezhető, hogy a becslésnél csak az L_2 és l_1 értékek ismerete szükséges. Ezután ugyanazok a módszerek használhatók, mint a lineáris függvényekkel való becslésnél, csak most nem egyeneseknek, hanem a paraboláknak a metszéspontjait kell meghatározni. A metszéspontoknak az $x = x_i$ egyenestől mért távolsága lesz a megfelelő Δx maximális értéke.

A parabolákkal való becslésnél hátrányos, ha a második deriváltakra adott korlátok távol vannak egymástól, vagyis ha az $L_2 - l_1$ abszolút értéke túl nagy.

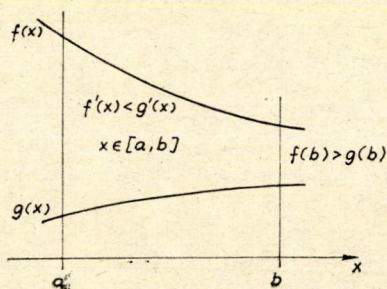
Az eddig tárgyalt becslési módszereken kívül természetesen más módszerek is adhatók. Az alkalmazott módszert mindig a szereplő függvények típusának (alakjának) megfelelően választhatjuk meg. Előfordulhat, hogy egy becslési eljárás belül egyszerre több különböző módszert is alkalmazunk a különböző részintervallumokon belül.

Speciális függvények esetében speciális módszereket is használhatunk.

Például monoton növekvő (vagy csökkenő) függvények esetén konstans függvényekkel is becsülhetünk: $f_i(x) = f(x_i)$, $g_i(x) = g(x_{i+1})$ (7. ábra).



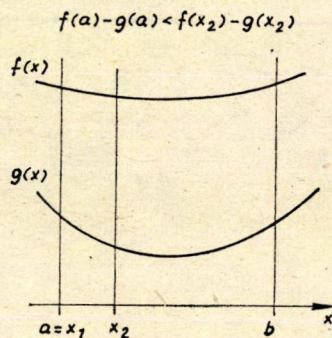
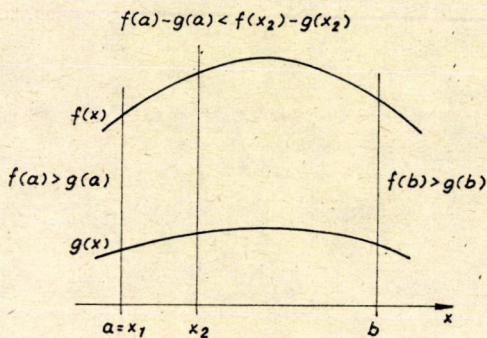
7. ábra



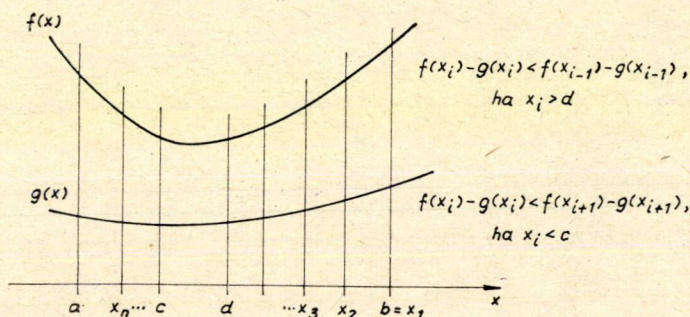
8. ábra

A deriváltak közt fennálló egyenlőtlenség ismeretében (például $g'(x) > f'(x)$, ha $x \in [a, b]$) csupán az intervallum egyik végpontjában kell ellenőrizni az egyenlőtlenség teljesülését (8. ábra).

Vagy ha tudjuk, hogy mindkét függvény deriváltja egyidejűleg monoton fogy, vagy monoton növekszik, akkor a bizonyítás az egyenlőtlenség véges sok pontban



9. ábra



10. ábra

való ellenőrzésére (ha $f''(x) < g''(x)$, l. 9. ábra), vagy véges sok pontban való számítás mellett egy $[c, d] \subset [a, b]$ részintervallumon való vizsgálatra redukálható (ha $f''(x) > g''(x)$, l. 10. ábra).

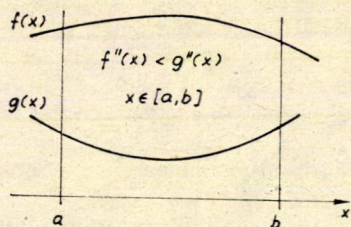
Ha például tudjuk, hogy $g''(x) > f''(x)$ $x \in [a, b]$ (mondjuk $g''(x) > 0$ és $f''(x) < 0$), akkor elegendő az intervallum két végpontjában ellenőrizni az egyenlőtlenséget (11. ábra).

Ezen speciális módszerek alkalmazhatóságának szükséges feltételei — például, ha ismerjük a deriváltak közötti egyenlőtlenséget — teljesülésének kimutatására szintén alkalmazhatók gépi (numerikus) módszerek.

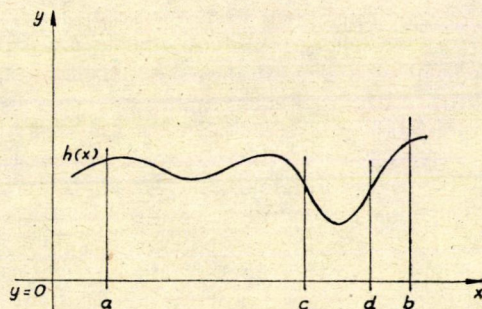
Egy, az eddigiektől eltérő típusú bizonyítási eljárás adható meg azon az alapon, hogy az $f(x) > g(x)$ $x \in [a, b]$ egyenlőtlenség ekvivalens az $f(x) - g(x) > 0$ relációval, és így az egyenlőtlenség bizonyítása visszavezethető a $h(x) = f(x) - g(x)$ függvény abszolút minimumának meghatározására (ilyen szélsőérték-keresési eljárások már ismeretesek). Közelebb visz a megoldáshoz, ha ismerjük azt a $[c, d]$ részintervallumot (illetve általában azt a részhalmazt), melyben a $h(x)$ abszolút minimuma van, hiszen ilyenkor már csak ezen a részhalmazon kell vizsgálni az egyenlőtlenséget (12. ábra).

Természetesen ez a módszer nemcsak egyváltozós függvények esetén alkalmazható.

Általában többváltozós függvényeknél az egyváltozós függvényekhez hasonló módon járhatunk el.



11. ábra



12. ábra

4. Többváltozós függvények esete

Példaként tekintsünk egy $f(x, y)$ és egy $g(x, y)$ $a \leq x \leq b$, $c \leq y \leq d$ függvényt. Bizonyítandó a $g(x, y) < f(x, y)$ egyenlőtlenség a fenti részintervallumon. Ilyenkor, ha a parciális deriváltakra korlátokat tudunk adni:

$$K_1 > \frac{\partial f}{\partial x} > k_1,$$

$$K_2 > \frac{\partial f}{\partial y} > k_2,$$

$$L_1 > \frac{\partial g}{\partial x} > l_1,$$

$$L_2 > \frac{\partial g}{\partial y} > l_2,$$

akkor, mint az egyváltozós esetben is, lineáris függvényekkel, azaz síkokkal helyettesíthetjük a vizsgált függvényeket. Ezek a síkok a G halmaz elemei, és a megfelelő $x_i \leq x \leq x_i + \Delta x$, $y_j \leq y \leq y_j + \Delta y$ részintervallumokon ezek között a lineáris függvények között mutatjuk ki az egyenlőtlenséget.

Például, ha $(L_1 - k_1)\Delta x + (L_2 - k_2)\Delta y < f(x_i, y_j) - g(x_i, y_j)$, akkor az egész $x_i \leq x \leq x_i + \Delta x$, $y_j \leq y \leq y_j + \Delta y$ intervallumon igaz az $f(x, y) > g(x, y)$ egyenlőtlenség.

Az egyváltozós esethez hasonlóan — az (1), (2) és (3) egyenlőtlenségekkel analóg képletek alapján — itt is különböző megoldások lehetségesek, és nemcsak a kétváltozós, hanem általában többváltozós esetben is.

Ugyancsak az egyváltozós esettel analóg módon, a lineáris függvényekkel való becsléseken kívül egyéb más módszerek is alkalmazásra kerülhetnek.

Befejezőként még megemlítnék néhány lényeges kérdést.

5. Általános megjegyzések, problémák

Az eddigiekben leírt, illetve megemlített módszerek egyáltalán nem merítik ki az összes lehetséges eljárástípust, melyeket egyenlőtlenségek bizonyításánál lehet használni. A deriváltakon kívül például a függvények más jellemzői is felhasználhatók a becsléseknél. Vannak esetek, amikor egészen más, az eddigiektől elvileg is különböző módszerek alkalmazása is eredményes.

A felhasznált eljárás eredményessége természetesen függ attól, hogy milyen függvényekre alkalmazzuk. Az eljárás kiválasztásánál ezenkívül még a következő probléma is felmerül:

Ha pontosabb becsléseket használunk, akkor az eljárás általában eredményesebb (például kevesebb lépésből áll, közelebbi függvények közti egyenlőtlenség is kimutatható), de ugyanakkor ez bonyolultabb számítást igényel, és bonyolultság miatt a számítási pontosság is romlik, rontva a becslés pontosságát. Hasonlóképpen

a számítási pontosság növelése (például a számológép dupla pontosságú műveletek használatával) hatásosabbá teszi az eljárásokat, de ugyanakkor bonyolultabbá is.

Bonyolult problémát okoz a számítás pontosságának meghatározása, illetve figyelembevétele a becsléseknél, hiszen a számítás hibája sok különböző tényezőtől függ: például maguktól a vizsgált függvényektől, a felhasznált számítási módszertől, az adott számológép szerkezetétől, sőt attól is, hogy az értelmezési tartomány mely pontjában végezzük a vizsgálatot.

Ezek után összefoglalásként elmondhatjuk a következőket.

Módszereink nem túl közeli $f(x)$, illetve $g(x)$ függvények közti egyenlőtlenség bizonyítását gyors és egyszerű úton teszik lehetővé. Eljárásaink hátránya azonban, hogy esetenként csak két konkrét függvény vizsgálata lehetséges, és ez is csak egy korlátos értelmezési tartományon. Az utóbbi probléma bizonyos esetekben megfelelő transzformációk alkalmazásával vagy a függvények esetleges periodikusságának kihasználásával kiküszöbölhető.

IRODALOM

- [1] DAVIS, M.: Eliminating the irrelevant from mechanical proofs, *Proceedings of Symposia in Applied Mathematics*, American Mathematical Society (1963).
- [2] GELERNTER, H.: Realization of a geometry theorem proving machine, *Proc. Inter. Conf. on Information Processing*, Paris (1959).
- [3] KALMÁR L.: A Hilbert-féle bizonyításelmélet célkitűzései, módszerei és eredményei, *Matematikai és Fizikai Lapok* **48**, (1941).
- [4] MELTZER, B.: Some recent developments in complete strategies for theorem-proving by computer, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **14**, (1968), 5.
- [5] NEWELL, A.—J. C. SHAW—H. A. SIMON: Report on a general problem-solving program, *Proc. Inter. Conf. on Information Processing*, Paris (1959).
- [6] PRAWITZ, D.: An improved proof procedure, *Theoria* **26**.
- [7] ROBINSON, J. A.: A machine-oriented logic based on the resolution principle, *Journal Ass. Comp. Mach.* **12**.

(Beérkezett: 1969. VI. 25.)