Globális optimalizálási módszerek továbbfejlesztése, tesztelése és alkalmazása atomklaszter feladatokra

doktori értekezés

Vinkó Tamás

Témavezető: Dr. Csendes Tibor

Szegedi Tudományegyetem

Szeged, 2006

Tartalomjegyzék

Előszó **v**

1. Bevezetés **1**
1.1. A vizsgált feladatok általános alakjai 1
1.2. A globális optimalizáló módszerek osztályozása 3

2. Az intervallumos globális optimalizálási módszerek gyorsítása **5**
2.1. Intervallum-aritmetika 5
2.1.1. Műveletek intervallumokkal 6
2.1.2. Intervallumos befoglaló függvények 7
2.1.3. Az intervallumos befoglaló függvények néhány tulajdonsága . 9
2.2. A korlátozás és szétválasztás típusú algoritmus 9
2.3. Középponti formulák 12
2.3.1. Optimális középponti formula 12
2.3.2. Lineáris határvonal formula 13
2.4. Kite befoglaló függvény – egydimenziós eset 15
2.4.1. Optimális kifejtési pont 17
2.4.2. A kite befoglalás tulajdonságai 21
2.4.3. Metszés 22
2.4.4. Kiterjesztett kite algoritmus 25
2.4.5. Numerikus eredmények 26
2.5. Kite befoglaló függvény – többdimenziós eset 33
2.5.1. A kite befoglalás komponensenkénti kiterjesztése 33
2.5.2. Komponensenkénti metszés magasabb dimenzióban 35
2.5.3. A javasolt algoritmus 38
2.5.4. Numerikus eredmények 39

3. Egy módszertan globális optimalizáló programok összehasonlítására **45**
3.1. Előkészületek 46
3.1.1. Tesztfeladatok 46
3.1.2. Időzítés 47
3.1.3. Egységes input 48
3.1.4. Teljesítmény kritériumok 51
3.1.5. Legjobb függvényértékek előállítása, vizsgálata 52

3.2. Jelölések a táblázatokban 53
3.2.1. Összefoglaló statisztikák 53
3.2.2. Feladatok osztályozása nehézség szerint 54
3.2.3. Részletező táblázatban használt jelölések 54
3.2.4. Futási idők összehasonlítása 55
3.2.5. Megbízhatósági analízis 56
3.2.6. A teszteredmények összefoglalása 58

4. Atomklaszter feladatok 61
4.1. Alapfogalmak 61
4.1.1. Vizsgálandó tulajdonságok 62
4.1.2. Eredmények használhatósága 62
4.1.3. Korábbi eredmények 63
4.1.4. Jelölések 64
4.1.5. Feltételek a párpotenciál függvényre 65
4.2. Méretfüggő korlátok 65
4.3. Méretfüggetlen korlátok 67
4.3.1. Első változat 67
4.3.2. Továbbfejlesztett változat 73
4.4. Lennard-Jones klaszterek 79
4.4.1. Méretfüggő korlát a minimális atompár távolságra 80
4.4.2. Méretfüggetlen alsó korlátok a minimális atompár távolságra . 81
4.4.3. Lineáris alsó korlát az optimum értékére 84
4.4.4. Statisztikák empírikus adatokból 84
4.5. Morse klaszterek 88
4.5.1. Méretfüggő alsó korlát a minimális atompár távolságra 88
4.5.2. Méretfüggetlen alsó korlát a minimális atompár távolságra . . 89
4.5.3. Lineáris alsó korlát az optimum értékére 91
4.6. Konklúzió és további feladatok 91

Összefoglalás 93

Summary 95

Irodalomjegyzék 97

Köszönettel tartozom témavezetőmnek, Csendes Tibornak, akitől a tudomány művelésének alapvető módszereit tanultam; Arnold Neumaiernek az inspirációért és mert tanítványává fogadott; Waltraud Huyernek, Jean-Loius Lagouanelle-nek, Dietmar Ratznak és Oleg Shcherbina-nak a közös munkáért; valamint Csirik Jánosnak, amiért a szegedi Mesterséges Intelligencia Kutatócsoportnál lehetőséget biztosít tudományos munkámhoz.

Köszönet illeti Tóth Boglárkát, Gazdag Zsoltot és Szörényi Balázst, akik munkámhoz hasznos ötleteikkel járultak hozzá.

Szüleimnek és feleségem szüleinek a végtelen szeretetért és bizalomért vagyok hálás.

Az értekezés elkészítésében feleségem, Anita, és gyermekeink, Marci és Barnus türelmükkel, megértésükkel és szeretetükkel támogattak. Eredményeimet Nekik ajánlom.

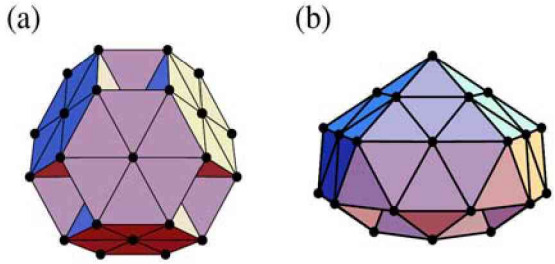
Előszó

„Ez a módszer azonban akkora éberséget és lelkierőt követelt,
hogy sokakat rabul ejtett egy képzeletbeli valóság,
a maguk agyának szüleménye, amelyből kevesebb gyakorlati hasznot,
de több vigaszt merítettek."

Gabriel García Márquez: Száz év magány (részlet)

Jelen értekezés témája a globális optimalizálás, a feladatunk az összes lehetséges megoldás közül megadni mindazokat, amelyek a legjobb eredményt szolgáltatják. Matematikai értelemben ez azt jelenti, hogy megadott feltételek mellett keressük meg a célfüggvény összes globális szélsőértékét (a feladattól függően minimumát vagy maximumát). A témakör matematikai háttere több, mint száz éves múltra tekint vissza. A digitális számítógépek megjelenésével és rendkívül gyors technikai fejlődésével egyidőben az optimalizálás gyakorlati jelentősége is megnövekedett. A jelenleg elérhető és ténylegesen futtatható (globális) optimalizáló módszerek száma több tucat. Az értekezésben ezek közül csak az ú.n. teljes megoldókkal foglalkozunk: ide azokat az eljárásokat soroljuk, amelyek biztosan megtalálják a globális szélsőértékeket, amennyiben egzakt számítást és végtelen hosszú futási időt feltételezünk. Itt ha a célunk az, hogy a globális megoldás egy előírt közelítését találjuk meg, akkor az eljárás garantáltan végezni fog véges határidőn belül. Ezen módszertípuson belül értelmezhetjük a szigorúan teljes keresők fogalmát, ahol a globális optimum megkeresése mellett matematikai szigorúsággal állíthatjuk a kapott megoldás globalitását (még véges pontosságú –tehát kerekítési hibákkal terhelt– aritmetika esetén is).

Az optimalizálási feladatoknál a globális megoldás megkeresése gyakran döntő fontosságú lehet. Példaként említhetjük a kémiai számításokban felmerülő potenciálfüggvény optimális értékének és helyének meghatározására vonatkozó feladatot, amelynek megoldása csak akkor jelent tényleges megoldást, ha az a globális minimumot írja le. Jelen értekezésben foglalkozunk majd ezen témakörhöz tartozó feladatokkal. Szemléltetésképpen tekintsük az 1. ábrát, ahol a 38 atomból álló ún. Lennard-Jones energiafüggvény globális minimumhoz (pontosabban globálisnak sejtett minimumához) tartozó konfigurációját ((a) ábra) és egy szerkezetében teljesen különböző lokális optimumot ((b) ábra) látunk, amelyek értékben igen közel állnak



1. ábra. A 38 atomból álló Lennard-Jones feladat két lehetséges megoldása: (a) globális optimumhoz, és egy (b) lokális (nem globális) optimumhoz tartozó.

egymáshoz. Rögtön láthatjuk, hogy a lokális minimum meghatározásával a keresett globális megoldástól még meglehetősen távol vagyunk.

Egy másik szemléletes példa a robotikából származik. A Lee & Mavroidis [34] cikkben tárgyalt feladat egy egyszerű robotkar lehetséges állapotainak megállapítására vonatkozik. Könnyű látni, hogy lokális megoldásnak itt sem vesszük hasznát. Bár a feladat formalizálás után egy alacsony fokszámú polinomrendszerből áll, a szerzők egy 64 darab processzort tartalmazó rendszerrel 70 óráig számolták, míg az összes globális megoldást megtalálták.

További motivációs példákat a globális optimalizálás fontosságára a Neumaier [47] összefoglaló cikkben olvashatunk.

A globális megoldás megkeresésére olyan módszerek kifejlesztése érdekes számunkra, amely számítógépen megvalósítható. Ebben az aspektusban viszont fontos a megbízhatóság kérdése. Már a bevezető jellegű numerikus matematika kurzusok is a hibaszámítás és a számítógéppel, lebegőpontos műveletekkel elvégzett számításokban előforduló (gyakran végzetes kimenetelű) hibalehetőségek tárgyalásával kezdődnek. Rendkívül fontos tehát, hogy számításaink eredménye olyan legyen, amelyre tudunk biztosítékot adni, a globális minimumhelyet, illetve -értéket a kívánt tolerancia megsértése nélkül szolgáltatni tudjuk. Ez a témakör a megbízható számítások területe. Ezen belül a globális optimalizálási eljárások az ún. korlátozás és szétválasztás (B&B) módszerén és az intervallum matematikán alapszanak.

A B&B módszer lényege, hogy a keresési teret rekurzív módon részproblémákra osztjuk (ez a szétválasztás), és az egyes részfeladatokon alsó és felső korlátokat állítunk a célfüggvény lehetséges értékeire (ez a korlátozás), melyek segítségével előbb vagy utóbb eliminálhatjuk azokat a részeket, amelyek nem adnak jobb megoldást, mint az addig ismert legjobb. Ezt az eljárást kombinálhatjuk az intervallum matematika eszköztárával, amely természetes módon szolgáltatja a megfelelő alsó és felső korlátokat az egyes részfeladatokra, valamint igen kifinomult technikákat azon részek elvetésére, amelyek garantáltan nem tartalmaznak globális minimumhelyet.

A globális optimalizálás matematikai eszközökkel megalapozott módszereit számí-

tógépes környezetben kívánjuk felhasználni. Ha már van egy kész programunk, akkor fontos lehet meggyőződni arról, hogy az valóban helyesen működik-e, mennyire megbízható. Általában kiváncsiak vagyunk arra is, hogy egy adott megoldó módszer más (hasonló) módszerekhez képest mennyire hatékony – és itt elsősorban a feladatmegoldás gyorsaságát tekintjük mérvadónak.

Bizonyos optimalizálási feladattípusok esetén pedig nem elég, hogy általánosságban jól, gyorsan és megbízhatóan működő eljárásaink vannak. Gyakran előfordul, hogy az általános globális optimalizáló módszerekkel nem tudjuk megoldani az adott problémát (tipikus eset erre például a már említett potenciálfüggvény optimalizálás). Ilyenkor általában az egyedüli célravezető út az, ha kihasználjuk az adott feladat néhány sajátos tulajdonságát. Az atomklaszter feladatoknál például tudjuk, hogy az optimális szerkezetben az atomok nem lehetnek túl közel egymáshoz, illetve túl távol sem egymástól.

Az értekezés 4 fő fejezetre oszlik. Az 1. fejezetben a továbbiakhoz szükséges alapfogalmakat és tételeket vezetjük be, illetve ismertetjük.

A 2. fejezetben intervallumos globális optimalizálási algoritmusok továbbfejlesztésével foglalkozunk. Egy új befoglalófüggvény elméleti és numerikus vizsgálatát végezzük el. Először az egydimenziós változatra megmutatjuk, hogy a javasolt befoglaló függvény mindig jobb eredményt ad, mint az ötlet alapját képező másik két módszer. Bebizonyítjuk a felhasználáshoz szükséges tulajdonságok meglétét (befoglalási monotonitás, négyzetes konvergencia sebesség, és egy rendkívül hasznos metszési tulajdonság); valamint numerikus vizsgálatokkal kimutatjuk, hogy a klasszikus intervallumos korlátozás és szétválasztás típusú optimalizáló algoritmusba történő implementálása milyen hatékonyság-növekedést eredményez. Ugyanezen részhez tartozik még a módszer egy lehetséges többdimenziós kiterjesztésének vizsgálata is. Ebben az esetben is megmutatjuk, hogy a javasolt új technikából egy hatékony gyorsító módszer származtatható, amely (numerikus vizsgálatokkal igazolt módon) teljesítmény-növekedéshez vezet.

A 3. fejezet az ún. teljes globális optimalizálók tesztelésének és összehasonlításának módszertanával foglalkozik. A munka jelentőségét mutatja, hogy ez volt az első eset, amikor különböző korlátozásos globális optimalizálási és feltétel kielégítési feladatokat megoldó programok összehasonlítása megvalósult egyrészt szisztematikus alapon, másrészt egy olyan teszthalmazon, amely megengedi statisztikusan szignifikáns következtetések levonását. Az ismertetett módszertan tehát arra vállakozik, hogy algoritmikus úton olyan keretet adjon, amely számítógépen implementálható, és lényegében emberi beavatkozás nélkül elvégezzen egy olyan lépéssorozatot, amelynek a végén emberi feldolgozásra alkalmas értelmes kimutatásokat kapunk a tesztelt programok gyorsaságára, helyességére és megbízhatóságára vonatkozóan.

A 4. fejezet témája pedig atomklaszterek szerkezetének vizsgálata optimalizálási keretben. Célunk az volt, hogy minél jobb (méretfüggő és méretfüggetlen) alsó

korlátot adjunk az optimális konfigurációban előforduló minimális atompár távolságra. Ilyen információ birtokában az optimum megkeresésére szolgáló eljárások hatékonysága növelhető, valamint az optimum értékére lineáris alsó korlát adható (az eredményekből explicit módon számolható is ez a korlát).

Az értekezést magyar és angol nyelvű összefoglaló, valamint az irodalomjegyzék zárja.

Az értekezésben az egyes fogalmak első előfordulását dőlt betűtípussal emeljük ki, ez szolgál tehát a definíciók megadására.

Az állítások, tételek és következmények tekintetében minden esetben megadjuk annak forrását. Bizonyítást csak abban az esetben közlünk, ha az teljes egészében saját eredmény (és ha az értekezés alapját képező cikkekben az ugyancsak megtalálható).

1. fejezet

Bevezetés

Ebben a fejezetben bevezetjük a vizsgálandó feladatok általános alakját, valamint megmutatjuk, hogy az értekezés tárgyát képző módszerek milyen módon osztályozhatók. A további (konkrét eredményeket tárgyaló) fejezetek önálló egységet képeznek, ezért az ott felhasznált fogalmak és eredményeket is ott vezetjük be, illetve közöljük.

Az értekezésben $\mathbb{R}$ jelöli a valós számok, $\mathbb{R}^n$ pedig a n-dimenziós valós vektorok halmazát.

1.1. A vizsgált feladatok általános alakjai

Feltétel nélküli globális optimalizálási feladaton a

$$\min_{x \in S} f(x) \tag{1.1}$$

alakú feladatot értjük, ahol az $f : \mathbb{R}^n \to \mathbb{R}$ függvényt *célfüggvénynek*, az $S \subseteq \mathbb{R}^n$ tartományt pedig a *keresési tartománynak* nevezzük. Az (1.1) feladatot szokás még a *keresési tartomány korlátaival adott* (bound constrained) optimalizálási feladatnak is nevezni, abban az esetben, ha S alsó és felső korlátaival megadott intervallum. Jelen értekezés 2. fejezetében (1.1) alakú feladatok vizsgálatával foglalkozunk.

Megjegyzés. Fontos megkülönböztetnünk az $n = 1$ esetet, az egyváltozós globális optimalizálási problémát. A többváltozós esethez képest ez egyszerűbb probléma, hiszen a „dimenzionalitás átka" itt nincs jelen. Általában az is igaz, hogy az egydimenziós esetekre kifejlesztett technikák, módszerek, elméletek nem minden esetben vihetők át természetes módon magasabb dimenzióba. Mindazonáltal számos alkalmazási területe van az egydimenziós globális optimalizálásnak (lásd például a Casado *et al.* [7] cikkben megadott hivatkozásokat). Ha az általános esetet tekintjük, akkor az (1.1) feladat NP-nehéz.

Korlátozó feltételekkel megadott globális optimalizálási feladaton a

$$\begin{array}{rl} \min & f(x) \\ \text{úgy, hogy} & g_i(x) \leq \ 0 \quad (i = 1, \ldots, l) \\ & x \in S \end{array} \tag{1.2}$$

alakú feladatot értjük, ahol minden $i \in \{1, \ldots, l\}$ indexre $g_i : \mathbb{R}^n \to \mathbb{R}$ (*korlátozó feltétel*). Az értekezésben (1.2) alakú feladatokkal csak közvetett módon foglalkozunk, a 3. fejezetben adunk egy módszertant az ilyen típusú feladatok megoldására kifejlesztett programok tesztelésére.

Feltétel kielégítési feladatról akkor beszélünk, ha az (1.2) alakú feladatban nincs célfüggvényünk, csak korlátozó feltételek egy rendszere.

Megjegyzés. Vegyük észre, hogy az (1.2) alakú megfogalmazásban benne van az (1.1) alakú és a feltétel kielégítési feladat megfogalmazása is, tehát ha globális optimalizálási feladatról beszélünk, akkor mindig gondolhatunk az (1.2)-re.

Az (1.2) és a feltétel kielégítési feladatban a feltételeket kielégítő pontok halmazát *lehetséges megoldásoknak* nevezzük. Azon pontokat pedig, amelyek nem teljesítik a megadott feltételeket *nem lehetséges megoldásoknak* nevezzük[1].

Azt mondjuk, hogy egy probléma *nem kielégíthető*, ha a feltételrendszere olyan, hogy nincs hozzá lehetséges megoldás.

Megjegyzés. A feltétel kielégítési feladatoknál minden lehetséges megoldás egyben globális megoldás is.

A globális optimum értékét f^*, az ehhez tartozó globális minimumpontot (amennyiben egy van) pedig x^* jelöli.

Példa. Legyen adott atomok n elemű d dimenziós halmazában (klaszterében) az atomok egymásra hatását leíró potenciál függvény. Keressük meg a minimális energiához tartozó optimális szerkezetet. Ez ebben a formában egy globális optimalizálási feladat az nd-dimenziós Euklidészi térben. Amennyiben a feladat leírását kiegészítjük például olyan korlátozó feltételekkel, amelyek kizárják a forgatási és tükrözési szimmetriákat, akkor (1.2) alakú feladatot kapunk. Ha pedig adott egy feltételezett minimális energiaszint és azt kell megmutatnunk, hogy ennél az energiaszintnél nem érhető el alacsonyabb, akkor feltétel kielégítési feladatot kapunk.

[1] Használatos még a *fízibilis* és *infízibilis* pontok szóhasználat is.

1.2. A globális optimalizáló módszerek osztályozása

A NEUMAIER [47] által javasolt felosztás szerint az (1.2) alakú feladatok megoldására szolgáló módszerek a következőképpen osztályozhatók.

A *nemteljes* módszerek heurisztikán alapuló eljárások. Itt nincs biztosítékunk arra, hogy egy lokális megoldásba beragadunk-e vagy sem, valamint arról sincs információnk, hogy milyen közel vagyunk a globális minimumhoz. Ezért a megállási feltételek is heurisztikusak.

Az *aszimptotikusan teljes* módszerekre bebizonyítható, hogy korlátlan futási időt feltételezve egy valószínűséggel megtalálják a globális minimumot (egy előírt tolerancia mellett). A megállási feltétel azonban itt is heurisztikus, hiszen az ide tartozó módszerek nem tudják, hogy a globális megoldást találták-e meg.

A *teljes* módszerek pontos aritmetikát feltételezve megjósolható időkorláton belül garantáltan megtalálják a globális optimumot (valamilyen toleranciával). Itt a megjósolhatóság azt jelenti, hogy van valamilyen információnk a problémával kapcsolatban (például Lipschitz konstans vagy más globális jellegű információ), amivel a konvergencia sebességet becsülhetjük.

A *szigorúan megbízható* (rigorous) módszerek olyan teljes módszerek, amelyek még kerekítési hibák megléte esetén is garantáltan megtalálják a globális optimumot (valamilyen toleranciával).

Az értekezés 2. fejezetében szigorúan megbízható módszerek továbbfejlesztésével foglalkozunk, míg a 3. fejezetben ismertetett módszertan teljes keresők tesztelésére és összehasonlítására ad eljárást.

2. fejezet

Az intervallumos globális optimalizálási módszerek gyorsítása

Ebben a fejezetben a valós számokat kisbetűvel, az intervallumokat pedig nagybetűvel jelöljük.

2.1. Intervallum-aritmetika

Az X *intervallumot* az alsó és felső korlátja között lévő pontok (nem üres) halmazával definiáljuk:

$$X = [\underline{X}, \overline{X}] = \{x \in \mathbb{R} \mid \underline{X} \leq x \leq \overline{X}\},$$

tehát azt mondjuk, hogy egy $x \in \mathbb{R}$ *benne van* az X intervallumban, azaz $x \in X$ akkor és csak akkor, ha $\underline{X} \leq x \leq \overline{X}$. Itt tehát $\underline{X}$ jelöli az *alsó végpontot*, $\overline{X}$ pedig a *felső végpontot*. Az n dimenziós *intervallum vektor* esetén $X = (X_1, \ldots, X_n)^T$ jelöli az $X_k = [\underline{X}_k, \overline{X}_k]$ $(k = 1, \ldots, n)$ *komponenseket*. Az értekezésben mindvégig az intervallum szót fogjuk használni, abban az esetben is, ha többdimenziós esetet tárgyalunk.

Az összes n dimenziós intervallumot tartalmazó halmazt $\mathbb{I}^n$ jelöli. (Szokás még az $\mathbb{IR}^n$ jelölés is, de mi itt csak a valós esettel foglalkozunk, így az $\mathbb{R}$ megkülönböztetést elhagyjuk.) Amennyiben $D \subseteq \mathbb{R}^n$ egy halmaz, akkor $\mathbb{I}(D)$ jelöli az összes olyan X intervallum halmazát, amelyre $X \subseteq D$. Az $X = [x, x]$ *vékony* intervallum (tehát nulla szélességű intervallum) általában az x ponttal van azonosítva. Az X intervallum egy általános pontját x jelöli (általában az x, y, z esetleg $\tilde{x}$ vagy c, d jelöléseket használjuk majd).

Az $X \in \mathbb{I}$ *szélessége* a

$$\operatorname{wid}(X) = \overline{X} - \underline{X} \geq 0,$$

az $X \in \mathbb{I}^n$ szélessége a $\operatorname{wid}(X) = \max_{i=1,\ldots,n} \operatorname{wid}(X_i)$ szerint definiált.

Az $X \in \mathbb{I}^n$ *középpontja* a

$$\operatorname{mid}(X) = \frac{1}{2}(\underline{X} + \overline{X}),$$

kifejezéssel van meghatározva, ahol $\underline{X} = (\underline{X}_1, \ldots, \underline{X}_n)$ és $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_n)$.

Az $X \in \mathbb{I}^n$ *relatív szélessége* pedig a

$$\operatorname{wid}_{\mathrm{rel}}(X) = \frac{\operatorname{wid}(X)}{\max\{1, \min_{x \in X} |x|\}},$$

által definiált.

Korlátos $S \subseteq \mathbb{R}^n$ halmazokra

$$[\![S := [\inf S, \sup S]$$

halmazt az S *intervallum burkának* (intervall hull) nevezzük. Ez tehát a legszűkebb intervallum, amely tartalmazza az S halmazt.

Az *elemi műveletek* halmazát

$$\Omega := \{+, -, \cdot, /\}$$

definiálja. Az *elemi függvények* egy előre megadott Φ halmaz elemei, folytonosak minden olyan zárt intervallumon, amelyen definiáltak[1]. Például a

$$\Phi := \{\sin, \cos, \exp, \ln, \sqrt{\ }, \operatorname{abs}, \arctan, \ldots\}$$

a szokásos elemi függvényeket tartalmazza.

2.1.1. Műveletek intervallumokkal

A valós számokon értelmezett *elemi műveletek intervallumos kiterjesztése* az

$$X \circ Y := \{x \circ y \mid x \in X, y \in Y\} \in \mathbb{I}, \quad \text{ahol} \quad \circ \in \Omega \tag{2.1}$$

definíció alapján történik. A definícióból látható, hogy a megfelelő eredmény intervallumot a két intervallumból szóba jöhető összes elemre (valós számra) elvégezett művelet adja. Ez tehát végtelen sok művelet elvégzését jelentené. Könnyen látható

[1] Használatos még a *standard függvények* elnevezés is; ezt Kearfott [30] úgy definiálja, hogy azon függvények halmaza, amelyek a FORTRAN-77 nyelvben adottak.

azonban, hogy az alapműveletek folytonossága miatt a (2.1) képlettel adott műveletek valójában könnyen számítók:

$$\begin{aligned} X+Y &= [\underline{X}+\underline{Y}, \overline{X}+\overline{Y}], \\ X-Y &= [\underline{X}-\overline{Y}, \overline{X}-\underline{Y}], \\ XY &= [\min\{\underline{X}\underline{Y}, \underline{X}\overline{Y}, \overline{X}\underline{Y}, \overline{X}\overline{Y}\}, \max\{\underline{X}\underline{Y}, \underline{X}\overline{Y}, \overline{X}\underline{Y}, \overline{X}\overline{Y}\}], \\ X/Y &= X \cdot [1/\overline{Y}, 1/\underline{Y}], \quad \text{ha } 0 \notin Y. \end{aligned}$$

Valós függvények intervallumos kiterjesztése is hasonlóképpen történik. A $\varphi \in \Phi$ elemi függvényre

$$\varphi(X) := \square\{\varphi(x) \mid x \in X\},$$

ahol a jobb oldal definiált. Az $X\omega Y$ *reláció* (ahol $\omega \in \{=, <, \leq, >, \geq\}$) az X és Y intervallumok között akkor és csak akkor teljesül, ha $x\omega y$ teljesül minden $x \in X$ és $y \in Y$ elemre.

Megjegyzés. Fontos megjegyeznünk, hogy amennyiben véges pontosságú aritmetika áll rendelkezésünkre (és pontosan ez az eset áll fent amennyiben az intervallum-aritmetika számítógépes megvalósítását használjuk), akkor az intervallumos műveletek elvégzésekor kifelé kerekítést kell végrehajtani (lásd Kearfott [30] 147. oldal, illetve Neumaier [45] 8. oldal). Az értekezésben az egyes numerikus megvalósításoknál az így kapott *gépi intervallum-aritmetikát* használjuk. Szokásos erre külön jelölésrendszert bevezetni (a műveletekre), amitől a tézisben eltekintünk: elméleti megfontolásainkban a valós intervallum-aritmetikát, míg a számítógéppel elvégzett numerikus vizsgálatoknál a gépi aritmetikát használjuk, így egyértelmű, hogy mikor melyik van érvényben.

2.1.2. Intervallumos befoglaló függvények

Azt mondjuk, hogy az $F : \mathbb{I}^n(X) \to \mathbb{I}$ az $f : \mathbb{R}^n \to \mathbb{R}$ egy *intervallumos befoglaló függvénye* az X intervallumon, ha $x \in Y$ esetén $f(x) \in F(Y)$ teljesül minden $Y \in \mathbb{I}^n(X)$ intervallumra.

Az f függvény *értékkészletét* az Y intervallumon $f(Y)$, továbbá $\underline{f}(X)$ az értékkészlet alsó korlátját, valamint $\underline{F}(X)$ és $\overline{F}(X)$ az intervallumos befoglalás alsó- és felső korlátját jelöli.

Könnyű látni, hogy egy tetszőleges valós függvény értékkészletének pontos kiszámítása két globális optimalizálási feladatnak felel meg az X intervallumon. Ebből következik, hogy általános esetben az értékkészlet csak túlbecsléssel adható meg. Megfelelő befoglaló függvény konstruálása ezért az intervallum-aritmetika központi jelentőségű *alapfeladata.*

A legegyszerűbb befoglalást az intervallum-aritmetika automatikusan szolgáltatja. Ehhez tekintsük az $f : \mathbb{R}^n \to \mathbb{R}$ függvényt, mint matematikai kifejezést (tehát a Φ halmaz elemeit és Ω halmaz műveleteit változókkal összekapcsoló kifejezést). Az $F : \mathbb{I}^n \to \mathbb{I}$ függvény által meghatározott $F(X)$ intervallumot az f függvény *természetes intervallumos kiterjesztésének* nevezzük, amelyet úgy kapunk, hogy az f-et megadó kifejezésben minden $i \in 1, \ldots, n$-re az x_i változót X_i-re cseréljük, és minden valós alapműveletet és elemi függvényt az intervallumos megfelelőire cseréljük.

1. Tétel. (Moore [44]) *A természetes intervallum kiterjesztés befoglaló függvény. Amennyiben a tekintett f kifejezésben minden változó pontosan egyszer fordul csak elő, akkor a befoglalás pontos lesz.*

Amennyiben az f képletében egy változó többször is előfordul, akkor általában túlbecsléssel kapjuk meg az értékkészlet befoglalását. Ezt a jelenséget *függőségi problémának* (dependency problem) nevezzük. Megjegyezzük, hogy algebrai átalakításokkal sokat lehet tenni a függőségi problémákból adódó túlbecslések csökkentésére. Másrészt innen az is látszik, hogy ugyanazon intervallumon értelmezett, matematikailag ekvivalens kifejezések intervallumos kiterjesztésével kapott befoglalásai különbözőek lehetnek. Ennek a jelenségnek azonban hasznát is vehetjük: az intervallumos globális optimalizáló algoritmusok tesztelésére használhatunk olyan célfüggvényeket, amelyek a hagyományos (nem teljes) módszerek (lásd 1.2. alfejezet) számára gyorsan megoldhatók, míg az intervallumos módszerek sok munka árán végeznek csak megoldásukkal.

A továbbiakban f' az f függvény deriváltját (többváltozós esetben a gradiens vektort), F' pedig az f' egy intervallumos befoglalását jelöli.

Amennyiben a szóban forgó függvény folytonosan differenciálható, alkalmazhatjuk a *középponti formulákat.* A módszert az analízisből jól ismert középérték tételből származtatjuk. Nevezetesen, $f(x) = f(c) + f'(\xi)(x - c)$ teljesül $c, x \in Y$ és $\xi \in [\min\{c, x\}, \max\{c, x\}]$ esetén, ezért

$$f(x) \in F_{CF}(Y, c) := f(c) + F'(Y)(Y - c). \tag{2.2}$$

Itt az f függvényt minden $x \in Y$ értékre kiterjesztettük, hiszen $F'(Y)$ az f deriváltjának intervallumos befoglalása az Y intervallumon.

A c kifejtési pontot leggyakrabban az Y intervallum közepének választják. A 2.3.1. alfejezetben azonban látni fogjuk, hogy ez a kifejtési pont választható úgy is, hogy a középponti formula által elérhető lehető legjobb befoglalást kapjuk. Megjegyezzük továbbá, hogy (2.2) kiszámítható intervallumos lejtő aritmetikával is (Neumaier [45], Ratz [55]), amely gyakran az $f(Y)$ jobb befoglalását eredményezi.

2.1.3. Az intervallumos befoglaló függvények néhány tulajdonsága

Azt mondjuk, hogy az f függvény egy F befoglalása *izoton* (vagy *befoglalásra nézve monoton*) tulajdonságú X felett, ha minden $Y \subseteq Z$ $(Y, Z \in \mathbb{I}^n(X))$ esetén $F(Y) \subseteq F(Z)$ teljesül.

Az intervallumos alapműveletek és az alapfüggvények intervallumos kiterjesztései izoton tulajdonságú. Ebből indukcióval következik, hogy a természetes intervallumos kiterjesztés is izoton. Amennyiben a (2.2) képletben a $c = \mathrm{mid}\,(Y)$ választást használjuk, akkor az így kapott középponti formula is izoton tulajdonságú lesz (különben nem mindig).

Azt mondjuk, hogy az F befoglaló függvény *α-konvergens* az X intervallum felett, ha minden $Y \in \mathbb{I}(X)$ intervallumra $\mathrm{wid}\,(F(Y)) - \mathrm{wid}\,(f(Y)) \leq k(\mathrm{wid}\,(Y))^\alpha$ teljesül, ahol α és k pozitív konstansok.

Az $\alpha = 1$ esetet *lineáris-*, az $\alpha = 2$ esetet pedig *kvadratikus konvergenciának* nevezzük. A definíció alapján nagyobb konvergenciarendű befoglaló függvény esetén keskeny intervallumokra a befoglalás jobb lesz.

Az $F : \mathbb{I}^n \to \mathbb{I}$ függvényt *Lipschitz-folytonosnak* nevezzük az $X \in \mathbb{I}^n$ intervallumon, ha létezik olyan $k \in \mathbb{R}$, hogy $\mathrm{wid}\,(F(Y)) \leq k\mathrm{wid}\,(Y)$ teljesül minden $Y \subseteq X$ intervallumra.

2. Tétel. (Ratschek & Rokne [52]) *A természetes intervallum kiterjesztés lineárisan konvergens. Ha $c = \mathrm{mid}\,(X)$ és F' komponensei Lipschitz-folytonosak, akkor $F_{CF}(X, c)$ kvadratikusan konvergens.*

Általánosan elfogadott szabály, hogy ha az intervallum szélessége nagyobb, mint 1, akkor a természetes intervallumos kiterjesztést érdemes használni, ellenkező esetben viszont a középponti formulát.

A különféle befoglaló függvények konvergencia rendjének empírikus úton történő meghatározásáról a Tóth & Csendes [63] cikkben olvashatunk. A szerzők javaslatot tesznek arra, hogy az intervallum szélességét tekintve melyik befoglalást érdemes használni.

2.2. A korlátozás és szétválasztás típusú algoritmus

Teljes globális keresés elvégzésére általában a korlátozás és szétválasztás (branch-and-bound, továbbiakban B&B) módszere a használatos. Az ötlet lényege, hogy

rekurzív módon osszuk fel a keresési teret (szétválasztás) és ezeken az altereken alsó korlátokat állítva a célfüggvényre (korlátozás) eliminájuk azokat a részeket, amelyekről tudjuk, hogy nem vezetnek az eddig ismert legjobb megoldásnál jobbhoz. Az algoritmus legrosszabb esetben exponenciális futásigényű; bár az esetek többségében a keresés során a résztartományok jelentős részét el tudjuk vetni: például ha az aktuálisan vizsgált résztartományon a függvény alsó korlátja nagyobb, mint a monoton csökkenő felső korlát, akkor tudjuk, hogy a tekintett résztartomány nem tartalmazhatja a globális minimumot.

A B&B ötlet természetes módon alkalmazható az intervallum-aritmetikával együtt, hiszen ez utóbbi automatikusan ad korlátokat a vizsgált célfüggvényre. A megvalósítás Moore nevéhez fűződik (MOORE [44]), amely módszert aztán SKELBOE [60] módosított úgy, hogy az ténylegesen is egy jól használható eljárássá vált. Az *intervallum-aritmetikán alapuló, korlátozás és szétválasztás elvén működő* algoritmus általános alakja a következő.

1. lépés Legyen X a kezdő intervallum, $\mathcal{L}$ a munkalista, $\mathcal{Q}$ pedig az eredménylista. Számítsuk ki az $F(X)$ befoglalást, legyenek $\mathcal{L} := \{(X, \underline{F}(X))\}$, $\mathcal{Q} := \{\}$ és állítsuk be az f^* értékre vonatkozó garantált felső korlátot: $\tilde{f} = \overline{F}(c)$, $(c \in X)$.

2. lépés Mindaddig, amíg $\mathcal{L}$ nem üres, hajtsuk végre a következő lépéseket.

3. lépés Vegyünk le egy $(Y, \underline{F}(Y))$ elemet az $\mathcal{L}$ listáról. Osszuk fel az Y intervallumot $U_1 \cup U_2 \cup \ldots \cup U_k = Y$ részintervallumra $(k > 1)$ úgy, hogy $\mathrm{int}(U_1) \cap \ldots \cap \mathrm{int}(U_k) = \emptyset$ teljesüljön, ahol 'int' az U intervallum belsejét jelöli.

4. lépés Minden $i = 1, \ldots, k$-ra számítsuk ki az $F(U_i)$ befoglalásokat, alkalmazzunk gyorsító teszteket az U_i vagy annak bizonyos részeinek eliminálására majd frissítsük az $\tilde{f}$ értékét, ha lehetséges.

5. lépés Minden $i = 1, \ldots, k$-ra, amennyiben bizonyos feltételek teljesülnek, legyen $\mathcal{Q} = \mathcal{Q} + (U_i, \underline{F}(U_i))$ különben pedig legyen $\mathcal{L} = \mathcal{L} + (U_i, \underline{F}(U_i))$. Menjünk a 2. lépésre.

Az alábbiakban a fenti intervallumos B&B algoritmus néhány fontos részletét tárgyaljuk.

Értékkészlet befoglalás

Mint láttuk a 2.1.2. alfejezetben, az intervallum-aritmetika lényegében automatikusan szolgáltatja a szóban forgó függvény értékkészletének befoglalását. A garantált megbízhatóságú globális optimalizálásban az aktuálisan vizsgált intervallumon a

célfüggvény értékkészletének alsó korlátjára van csak szükségünk (a globális minimumra vonatkozó felső korlátot globális információként használva monoton csökkentjük). A módszereink megvalósításában a természetes intervallumos kiterjesztést és a középponti formákat használjuk, mint alap eszközöket. Jelen értekezés egyik eredményeként újabb befoglalási módszereket adunk. Ezen módszerek és a középponti alakok a deriváltfüggvény befoglalásait is felhasználják – ezt az automatikus deriválással számítjuk (lásd például a KEARFOTT [30] és CSENDES [12] cikkeket). A számítógépes implementációkban a C-XSC programcsomag HAMMER *et al.* [25] által adott előrefele történő (tehát forward mode) automatikus differenciálást használjuk.

Gyorsító eljárások

Az intervallumos B&B eljárás 4. lépésében láttuk, hogy alkalmazhatunk olyan eljárásokat, amelyek a keresési tér azon részeit eliminálják, amelyek garantáltan nem tartalmaznak globális minimumot. Részletes ismertetés nélkül: az implementált algoritmusban a kivágási tesztet, monotonitási tesztet, konkavitási tesztet és az intervallumos Newton-lépést használjuk (bővebben lásd HANSEN [27]).

Felosztási irányok, megállási feltétel, konvergencia

Az algoritmus 3. lépésében az Y intervallumot felosztjuk. A felosztás lehet két részintervallumra (bisection) vagy több részintervallumra (multisection) történő felosztás. Az ide vonatkozó elméleti és numerikus vizsgálatokat a CSALLNER *et al.* [11] és MARKÓT *et al.* [40] cikkek tartalmazzák. Az értekezésben vizsgált algoritmusokban biszekciót alkalmazunk.

A felosztás irányának megválasztása is teljesítmény-változáshoz vezethet. Ilyen irányú vizsgálatokat a CSENDES & RATZ [13] cikk tartalmaz.

Az 5. lépésben alkalmazhatunk különféle megállási feltételeket, amelyek befolyással vannak az algoritmus futási idejére és a megoldás minőségére is. Általában az aktuálisan vizsgált intervallum szélességét, illetve a befoglaló függvény szélességét szokás alapul venni, ezek egyikének (vagy mindkettőnek egyszerre) kell kisebbnek lennie egy-egy előírt tolerancia értéknél.

Az algoritmus konvergenciáját úgy szokás vizsgálni, hogy az 5. lépésben a megállási feltételt kikapcsoljuk, azaz feltesszük, hogy sohasem teljesül. Bizonyítandó ilyenkor, hogy a részintervallumok sorozatán vett értékkészlet befoglalások alsó értéke a globális minimum értékéhez tart. Az értekezésben ilyen típusú vizsgálatokkal nem foglalkozunk, a 2. fejezetben megvalósított eljárásokat egy olyan módszer módosításával készítettük el, amelyek teljesítik a konvergenciát, a módosítások pedig nem befolyásolják azt.

2.3. Középponti formulák

A fejezet hátralevő részében az f célfüggvényről feltesszük, hogy folytonosan differenciálható.

Mint azt láttuk, amennyiben a célfüggvényről elsőrendű információ is rendelkezésre áll (például derivált), akkor a (2.2) formulával javíthatunk az értékkészlet befoglalás szélességén. Mivel a kifejtési pont nincs rögzítve, ezért felmerül a kérdés, hogy annak megválasztása mennyire befolyásolja a befoglalás jóságát. A következő részben a kifejtési pont megválasztásának lehetőségeit tárgyaljuk.

Megjegyzés. Az egyszerűbb jelölés kedvéért az aktuálisan vizsgált egydimenziós Y intervallum végpontjait a és b jelöli, tehát $Y = [a, b]$, valamint a gradiens (vektor) elemeit $[\ell_i, u_i]$ $i = 1, \ldots, n$, és egydimenziós esetben az alsó indexeket elhagyjuk.

A továbbiakban feltesszük, hogy minden $i = 1, \ldots, n$ indexre $\ell_i < 0 < u_i$ teljesül. Ha valamely i-re $u_i \leq 0$ vagy $\ell_i \geq 0$, akkor f monoton, tehát az értékkészlet egyszerűen számítható.

2.3.1. Optimális középponti formula

Először az egydimenziós esetet tekintjük. A (2.2) képlet által adódó $F_{CF}(Y, c)$ alsó korlátját vizsgáljuk. A 2.1. ábrán láthatjuk, hogy minden $c \in [a, b]$-re a $(c, f(c))$ pont és az ℓ és u meredekségek által definiált két egyenes alsó korlátot ad f-re az Y intervallumon:

$$\min\{y_p(c), y_q(c)\} \leq \inf_{x \in Y} f(x),$$

ahol

$$y_p(c) := f(c) + u(a - c) \quad \text{és} \quad y_q(c) := f(c) + \ell(b - c).$$

Ebből az összefüggésből az alsó korlátra vonatkozó optimális c meghatározható. Baumann [2] bebizonyította, hogy c-re a legjobb választás akkor adódik, amikor $y_p(c) = y_q(c)$ teljesül, azaz a $c^- \in Y = [a, b]$ pont maximalizálja a $\min\{y_p(c), y_q(c)\}$ értékét. A következő tétel a megfelelő képleteket adja.

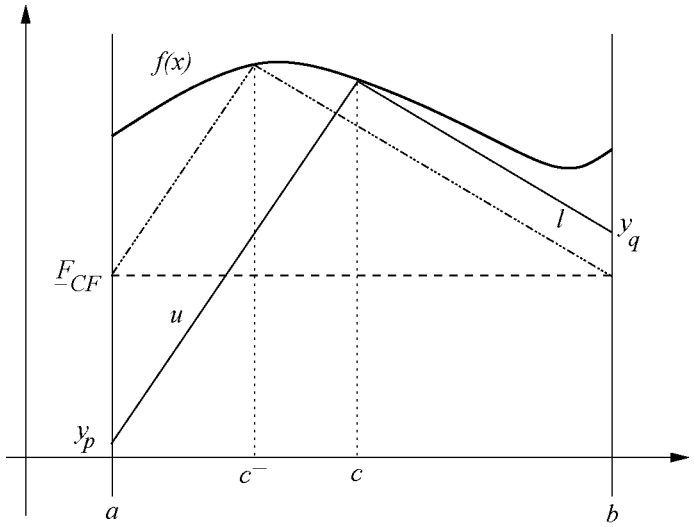
3. Tétel. (Baumann [2]) *A középponti formulában az optimális kifejtési pont és az ehhez tartozó alsó korlát a*

$$c^- = \frac{au - b\ell}{u - \ell}$$

és

$$\underline{F}_{CF}(Y, c^-) = f(c^-) + (b - a)\frac{\ell u}{u - \ell} \tag{2.3}$$

képletekkel adott.



2.1. ábra. Az aktuális intervallum középpontjára (egyenes vonalakkal) és az optimális alappontra (szaggatott vonalakkal) kifejtett középponti formula.

Megjegyezzük, hogy c^- értéke független az f értékeitől. Továbbá az intervallumos globális optimalizáló algoritmusban az ℓ és az u értékeket általában ettől függetlenül is kiszámítjuk, mert ezeket a monotonitási tesztben is fel tudjuk használni. Így a Baumann középponti formula nem kíván extra függvény- vagy gradiens hívást.

Hasonló meggondolással a felső korlátot optimalizáló c^+ pont is megkapható (észrevétel: ez a c^+ pont a c^- szimmetrikus párja a $\operatorname{mid}(Y)$ pontra nézve). Ezért ha a középponti formulák által kiszámítható legjobb befoglalást akarjuk megkapni, mindkét formulát használnunk kell, amely növeli a számítási igényt. A globális optimalizáló eljárásban azonban általában csak az alsó korlátot számoljuk.

A Baumann középponti formula többdimenziós kiterjesztése szintén megtalálható a Baumann [2] cikkben. Ez az általánosítás viszonylag egyszerűen adódik.

2.3.2. Lineáris határvonal formula

Először itt is szintén az egyváltozós esetet vizsgáljuk. Amikor a középponti formulát az intervallum alsó- és felső végpontjára egyidejűleg alkalmazzuk, akkor a *lineáris határvonal formulát* (linear boundary value form, a továbbiakban lbvf) kapjuk (Neumaier [45]). Ezt az esetet a 2.2. ábra szemlélteti. Az $y = f(a)+\ell(x-a)$ és $y = f(b)+u(x-b)$ egyenesek (x_s, y_s) metszéspontjának kiszámításával megkapjuk az alsó korlát előállítására vonatkozó képleteket. Ezt állítja a következő tétel.

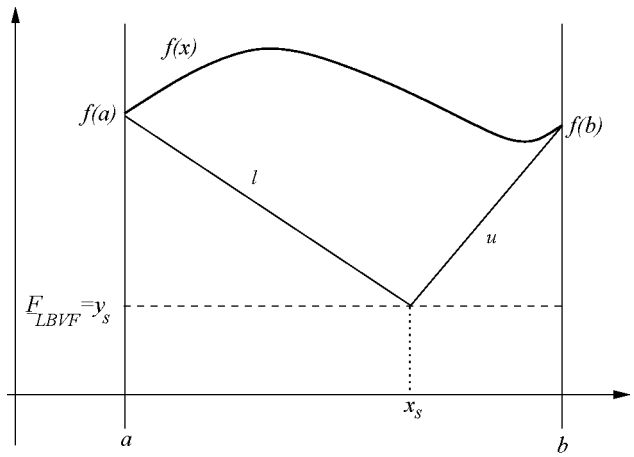
4. Tétel. (Neumaier [45]) *Az $(a, f(a))$ és $(b, f(b))$ pontok, valamint az ezekhez tartozó ℓ és u meredekségek által definiált egyenesek alsó korlátot adnak f-re:*

$$x_s = \frac{f(a) - f(b)}{u - \ell} + \frac{bu - a\ell}{u - \ell}, \tag{2.4}$$

$$\underline{F}_{LBVF}(Y) = y_s \quad = \quad \frac{uf(a) - \ell f(b)}{u - \ell} + (b - a)\frac{\ell u}{u - \ell}, \tag{2.5}$$

amelyet az lbvf alsó korlátjának nevezünk.

Világos, hogy az $\underline{F}_{LBVF}(Y) \leq \underline{f}(Y)$ egyenlőtlenség mindig teljesül, hiszen az $y = f(a) + \ell(x - a)$ és $y = f(b) + u(x - b)$ egyenesek az f függvény alatt vannak az $[a, b]$ intervallumon és soha nem metszik azt (a végpontokat kivéve).



2.2. ábra. Az lbvf befoglalás alsó korlátjának geometriai értelmezése.

Ezekből az eredményekből a következő kérdés adódik: melyik eljárás szolgáltat jobb alsó korlátot f értékkészletére? Egy egyszerű észrevétel az, hogy a (2.3) és (2.5) képletek meghatározása az $f(c^-)$ és az $(uf(a) - \ell f(b))/(u - \ell)$ kifejezésekben különböznek. Ez adja a következő állítást.

1. Állítás. [65] *Az* $\underline{F}_{CF}(Y, c^-) \leq \underline{F}_{LBVF}(Y)$ *egyenlőtlenség akkor és csak akkor teljesül, ha* $f(c^-) \leq \frac{uf(a)-\ell f(b)}{u-\ell}$.

Mint azt láthatjuk, az lbvf néha jobb eredményt ad, mint a Baumann forma. Az 1. Állítás azt mondja, hogy ez teljesül, ha például f konvex az adott intervallumon.

Jegyezzük meg, hogy a (2.4) és (2.5) képletekben minden érték rögzített, nincs lehetőség optimalitási vizsgálatokra. Az $\underline{F}_{LBVF}$ kiszámítása több információt igényel, hiszen szükségünk van az $f(a)$ és $f(b)$ értékekre; ez magasabb műveletigényhez vezethet az optimalizálási eljárásban. Az Y végpontjaiban vett függvényértékeket azonban felhasználhatjuk később is, amikor Y részintervallumait vizsgáljuk. Az f felső korlátjára vonatkozó formula hasonló (2.5)-hez és az alsó korlátokhoz már kiszámított értékeket ($\ell, u, f(a)$ és $f(b)$) tartalmazza.

A többváltozós esetre vonatkozó elméleti és numerikus vizsgálatokat a Messine & Lagouanelle [41] cikkben találjuk.

2.4. Kite befoglaló függvény – egydimenziós eset

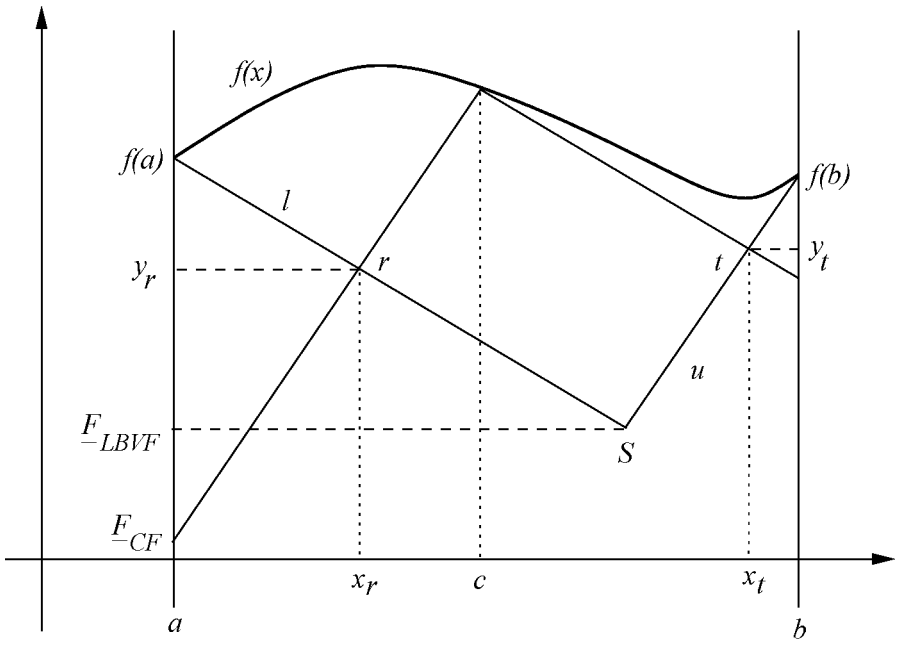
Származik-e valami előnyünk a fentebb tárgyalt két módszer együttes használatából? A választ a 2.3. ábra adja, amiből levezethető, hogy a szimultán használat nem rosszabb (és általában határozottan jobb) eredményt ad a célfüggvény befoglalására. Ezért definiáljuk az $\underline{F}_K(Y,c) := \min\{y_r(c), y_t(c)\}$ függvényt, ahol

$$y_r(c) := \frac{uf(a) - \ell f(c) + \ell u(c-a)}{u-\ell}, \tag{2.6}$$

és

$$y_t(c) := \frac{uf(c) - \ell f(b) + \ell u(b-c)}{u-\ell}. \tag{2.7}$$

Az $\underline{F}_K(Y,c)$ értéket a *kite befoglalás alsó korlátjának* nevezzük.



2.3. ábra. A középponti formula (kifejtési pontként az aktuális intervallum középpontját használva) és az lbvf szimultán használata.

5. Tétel. [65] *A*

$$\max\{\underline{F}_{LBVF}(Y), \underline{F}_{CF}(Y,c)\} \le \underline{F}_K(Y,c) \le \underline{f}(Y)$$

egyenlőtlenségek teljesülnek.

Bizonyítás. Legyen az r pont az $y = f(a)+\ell(x-a)$ és $y = f(c)+u(x-c)$ egyenesek metszéspontja:

$$x_r(c) = \frac{f(a) - f(c) + uc - \ell a}{u-\ell}, \tag{2.8}$$

és $y_r(c)$ fentebb definiált a (2.6) képletben. A t pont az $y = f(b) + u(x - b)$ és $y = f(c) + \ell(x - c)$ egyenesek metszéspontja:

$$x_t(c) = \frac{f(c) - f(b) + ub - \ell c}{u - \ell}, \tag{2.9}$$

és $y_t(c)$ fentebb definiált a (2.7) képlettel.

A következőkben négy esetet kell megvizsgálnunk.

(i) Tegyük fel, hogy $\underline{F}_{CF}(Y, c) \leq \underline{F}_{LBVF}(Y)$ teljesül. Meg kell mutatnunk, hogy $\underline{F}_{LBVF}(Y) \leq y_r(c)$ is igaz, azaz

$$\begin{aligned}
\frac{uf(a) - \ell f(b) + \ell u(b - a)}{u - \ell} &\leq^? \frac{uf(a) - \ell f(c) + \ell u(c - a)}{u - \ell} \\
-\ell f(b) + \ell ub &\leq^? -\ell f(c) + \ell uc \\
\ell(f(c) - f(b)) &\leq^? -\ell u(b - c) \\
f(c) - f(b) &\geq^? u(c - b) \\
\frac{f(c) - f(b)}{c - b} &\leq u.
\end{aligned}$$

Az utolsó egyenlőtlenség mindig teljesül, hiszen a baloldalon a $(c, f(c))$ és $(b, f(b))$ pontok által meghatározott egyenes meredeksége áll, míg a jobboldalon szereplő u az $f'(x)$ felső korlátja az $[a, b]$ intervallumon.

(ii) Most megmutatjuk, hogy ha $\underline{F}_{CF}(Y, c) \leq \underline{F}_{LBVF}(Y)$ akkor $\underline{F}_{LBVF}(Y) \leq y_t$:

$$\begin{aligned}
\frac{uf(a) - \ell f(b) + \ell u(b - a)}{u - \ell} &\leq^? \frac{uf(c) - \ell f(b) + \ell u(b - c)}{u - \ell} \\
uf(a) - \ell ua &\leq^? uf(c) - \ell uc \\
f(a) - f(c) &\leq^? \ell(a - c) \\
\frac{f(c) - f(a)}{c - a} &\geq \ell.
\end{aligned}$$

Az utolsó egyenlőtlenség mindig teljesül, mivel a baloldalán a $(c, f(c))$ és $(a, f(a))$ pontok által definiált egyenes meredeksége áll, a jobboldalán pedig ℓ, ami egy alsó korlátja $f'(x)$-nek az $[a, b]$ intervallumon.

(iii) Tegyük fel most, hogy $\underline{F}_{LBVF}(Y) \leq \underline{F}_{CF}(Y, c)$. Először megnézzük, hogy $\underline{F}_{CF}(Y, c) \leq y_r$ teljesül-e. Ennek bizonyítása az (i) eset bizonyításával analóg, meg kell mutatni, hogy $f(c) + u(a - c) \leq y_r(c)$. Ez hasonló okok miatt teljesül, mint azt az (i) pontban láttuk.

(iv) Végül azt nézzük meg, hogy ha $\underline{F}_{LBVF}(Y) \leq \underline{F}_{CF}(Y, c)$ akkor $\underline{F}_{CF}(Y, c) \leq y_t(c)$ is igaz. Ennek az esetnek a bizonyítása pedig a (ii) esethez hasonló. Belátható, hogy $f(c) + \ell(b - c) \leq y_t(c)$ teljesül hasonló okok miatt, mint az (ii) esetben.

A fenti négy eset megvizsgálásával beláttuk, hogy $\max\{\underline{F}_{LBVF}, \underline{F}_{CF}\} \leq \underline{F}_K(Y, c)$.

Hátra van még annak a bizonyítása, hogy $\underline{F}_K(Y, c) \leq \underline{f}(Y)$ is teljesül. Tekintsük az $Y_1 = [a, c]$ és $Y_2 = [c, b]$ intervallumokat, ahol $c \in [a, b]$. Az y_r és y_t értékek rendre az f függvény Y_1 és Y_2 intervallumokon vett két lbvf által adódó alsó korlátjai. A 4. Tételből tudjuk, hogy $y_r \leq \underline{f}(Y_1)$ és $y_t \leq \underline{f}(Y_2)$ mindig teljesülnek. Következésképpen az $y_K = \min\{y_r, y_t\} \leq \underline{f}(X)$ egyenlőtlenség is áll, amit bizonyítani kellett. □

Megjegyzés. Ha $c_1 \neq c_2$, akkor az $\underline{F}_{CF}(Y, c_1) \leq \underline{F}_K(Y, c_2)$ egyenlőtlenség nem feltétlen teljesül minden esetben. Példának vehetjük azt az esetet, amikor $c_2 = a$ vagy $c_2 = b$, mivel ekkor $\underline{F}_K(Y, c_2) = \underline{F}_{LBVF}(Y)$ és ha $c_1 = c^-$, akkor az 1. Állítás szerint $\underline{F}_{CF}(Y, c^-)$ lehet nagyobb, mint $\underline{F}_{LBVF}(Y)$.

2.4.1. Optimális kifejtési pont

A fenti eredményeink azt mutatják, hogy az lbvf és a középponti formula együttes használatával kapott alsó korlát legalább olyan jó, mint a kettő közül a jobbik. Most – ugyanúgy, mint azt vizsgáltuk a középponti formulánál – azt vizsgáljuk meg, hogy van-e lehetőség a felhasznált középponti formula középpontjának optimális megválasztására. Ez a c^* pont tehát olyan, hogy

$$\underline{F}_K(Y, c^*) = \max_{c \in [a,b]} \underline{F}_K(Y, c) = \max_{c \in [a,b]} \min\{y_r(c), y_t(c)\}. \tag{2.10}$$

A következő tételben a kite optimális középpontjára vonatkozó megállapításainkat mondjuk ki.

6. Tétel. [65] *A következők teljesülnek.*

1. *Létezik egy egyértelmű $c^* \in [a, b]$ pont, amelyre $y_r(c^*) = y_t(c^*)$ teljesül, és*
2. *c^* a maximumhelye a $\underline{F}_K(Y, c)$ függvénynek a c-re vonatkozóan.*

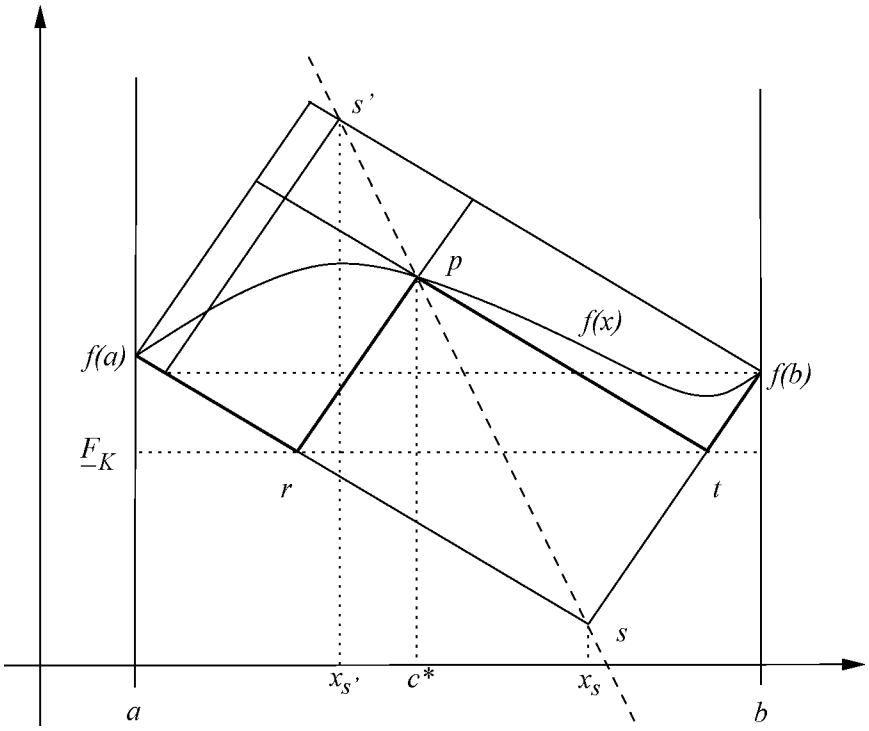
Bizonyítás. 1. Megvizsgáljuk a $\Delta := y_t - y_r$ különbséget. Deriválást alkalmazva azt kapjuk, hogy

$$y_r'(c) = \frac{-\ell f'(c)}{u - \ell} + \frac{\ell u}{u - \ell} \leq 0$$

és

$$y_t'(c) = \frac{u f'(c)}{u - \ell} - \frac{\ell u}{u - \ell} \geq 0$$

teljesülnek minden $c \in [a, b]$ pontra, ami azt jelenti, hogy y_r monoton csökken, y_t pedig monoton növekszik. Kihasználva, hogy $\ell < 0 < u$, ebből az következik, hogy



2.4. ábra. A kite kifejtési pontjának optimális választása.

$\Delta'(c) > 0$ minden $c \in [a, b]$ pontra. Így Δ szigorúan növekedő. Könnyű látni, hogy $\Delta(a) \leq 0$ és $\Delta(b) \geq 0$, ezért Δ-nak pontosan egy zérushelye van, a c^* pont az $[a, b]$ intervallumban, azaz

$$\frac{uf(a) - \ell f(c^*) + (c^* - a)\ell u}{u - \ell} = \frac{uf(c^*) - \ell f(b) + (b - c^*)\ell u}{u - \ell}. \tag{2.11}$$

A (2.11) egyenlőségből azt kapjuk, hogy

$$c^* = \frac{f(c^*) - f(a)}{2\ell} + \frac{f(c^*) - f(b)}{2u} + \frac{a + b}{2},$$

ami azt jelenti, hogy c^* az egyetlen fixpontja egy $\alpha f + \beta$ alakú függvénynek, ahol

$$\alpha = \frac{\ell + u}{2\ell u}, \qquad \beta = \frac{\ell u(a + b) - uf(a) - \ell f(b)}{2\ell u}. \tag{2.12}$$

Ezzel a tétel első állítását bebizonyítottuk.

2. Láttuk, hogy az y_r függvény monoton csökkenő, míg az y_t függvény monoton növekedő. Vegyük a $d \neq c^*$ pontot. Ha $d < c^*$, akkor

$$y_r(d) \geq y_r(c^*) = y_t(c^*) \geq y_t(d),$$

ahol valamelyik egyenlőtlenség szigorú, mivel $d \neq c^*$. Ezért

$$\underline{F}_K(Y, d) = \min\{y_r(d), y_t(d)\} = y_t(d) \leq y_t(c^*) = y_r(c^*) = \underline{F}_K(Y, c^*)$$

teljesül. Ha $d > c^*$, akkor

$$y_r(d) \leq y_r(c^*) = y_t(c^*) \leq y_t(d),$$

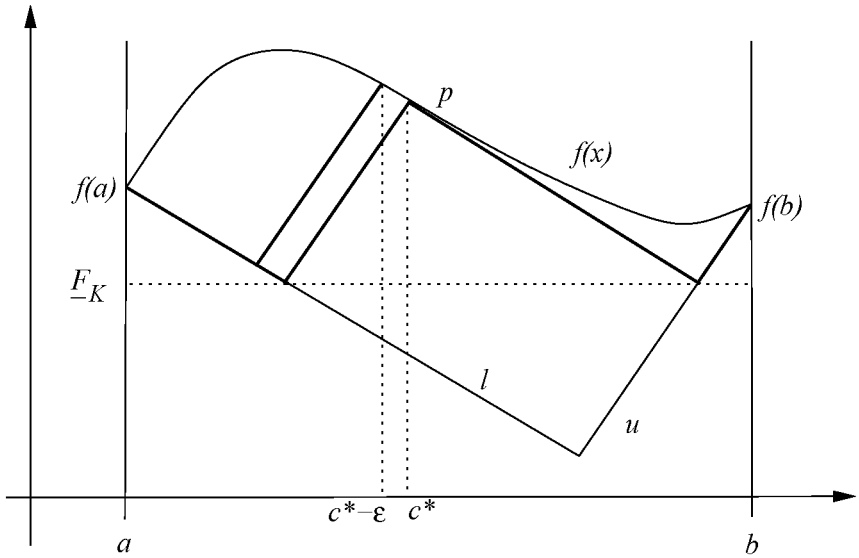
ahol az egyik egyenlőtlenség megintcsak szigorú, mert $d \neq c^*$. Ezért

$$\underline{F}_K(Y, d) = \min\{y_r(d), y_t(d)\} = y_r(d) \leq y_t(c^*) = y_r(c^*) = \underline{F}_K(Y, c^*).$$

Most mivel minden $d \neq c^*$ pontra az $\underline{F}_K(Y, d) \leq \underline{F}_K(Y, c^*)$ egyenlőtlenség is áll, ezért $\underline{F}_K$ maximális értéke a c^* pontban vétetik fel.

Egy a fentieket jól szemléltető példát a 2.4. ábrán láthatunk. □

Előfordulhat, hogy az $\underline{F}_K(Y, \cdot)$ függvénynek több maximumhelye is van. Amennyiben $f'(c^*) = \ell$ vagy $f'(c^*) = u$ és $f'(d) = \ell$ vagy $f'(d) = u$ teljesül minden $d \in [c^* - \varepsilon, c^* + \delta]$, $(\varepsilon, \delta > 0)$ értékre, akkor az $\underline{F}_K(Y, \cdot)$ függvénynek megszámlálhatatlan végtelen sok maximimhelye van a $[c^* - \varepsilon, c^* + \delta]$ intervallumban. Egy egyszerű példát látunk erre az esetre a 2.5. ábrán, ahol az $\underline{F}_K(Y, \cdot)$ függvénynek végtelen sok maximumhelye van a $[c^* - \varepsilon, c^*]$ intervallumon.



2.5. ábra. Az $\underline{F}_K(Y, \cdot)$ függvénynek végtelen sok maximumpontja is lehet.

Megjegyzés. Ha $f'(Y)$ az (ℓ, u) nyílt intervallumban van, akkor pontosan egy optimális pont van. Amennyiben a gépi megvalósítást vizsgáljuk, ez az eset általában teljesül is, hiszen az ℓ és u értékeket kifelé kerekítést használó intervallum aritmetikával számítjuk ki.

1. Következmény. *Az optimális kite befoglaló függvény mindig legalább olyan jó befogalást ad, mint a Baumann középponti formula, azaz*

$$\underline{F}_{CF}(Y, c^-) \leq \underline{F}_K(Y, c^*).$$

Bizonyítás. A 5. Tétel alapján az $\underline{F}_{CF}(Y, c^-) \leq \underline{F}_K(Y, c^-)$ egyenlőtlenség igaz. Láttuk, hogy $\underline{F}_K(Y, c) \leq \underline{F}_K(Y, c^*)$ is igaz, speciálisan $c = c^-$-re is. □

Az aktuális lépésben rendelkezésre álló információk alapján a c^* pont fixpont iterációval meghatározható. Ehhez tekintsük az (2.4)-ben definiált x_s pont és az

$$x_{s'} = \begin{cases} \dfrac{u(ua - \ell b) - \ell(f(a) - f(b))}{u(u - \ell)} & \text{ha } f(a) \leq f(b), \text{ és} \\ \dfrac{\ell(\ell b - ua) - u(f(b) - f(a))}{\ell(\ell - u)} & \text{ha } f(a) \geq f(b) \end{cases}$$

pont által definiált intervallumot. Világos, hogy c^* benne van ebben az intervallumban, hiszen a $(c^*, f(c^*))$ pont az f függvény grafikonjának és az $s = (x_s, f(x_s))$ és $s' = (x_{s'}, f(x_{s'}))$ pontok által megadott egyenes metszéspontja (lásd 2.4. ábra).

Gyorsabb konvergencia érdekében intervallumos Newton módszert is alkalmazhatunk az

$$\alpha f(c) + \beta - c = 0 \tag{2.13}$$

egyenletre, ahol α és β a (2.12)-ben definiáltak. Bár ebben az esetben az $f'(c)$ intervallumos kiértékelésére szintén szükségünk van. Alkalmazhatunk viszont kvázi Newton módszert is a (2.13) egyenleten az előzőleg kiszámított derviált befoglalást, mint konstanst használva. Mindkét módszer esetén általában egyetlen lépés elegendő ahhoz, hogy az optimális pont egy megfelelően jó közelítését kapjuk. Ugyanakkor tudjuk a 5. Tételből, hogy a kite befoglalás mindig legalább olyan jó alsó korlátot ad, mint a másik két módszer, ezért az intervallumos globális optimalizálási eljárásban nincs szükségünk az optimális pont nagy pontosságú meghatározására, lényegében bármilyen $c \in Y$ megfelelő, főleg, ha az az $[x_s, x_{s'}]$ intervallumból van.

A c^* pont c közelítését és a $\underline{F}_K(Y, c)$ befoglalását kiszámító eljárást *kite algoritmusnak* nevezzük. Mint azt később látjuk majd ezt az eljárást könnyen beépíthetjük az intervallumos globális optimalizáló módszerbe, valamint használhatjuk majd mint gyorsító eljárást is.

Ha felső korlátot szeretnénk meghatározni, akkor az ennek megfelelő c' középpont kiszámítható az $y_{r'} = y_{t'}$ egyenlőségből, ahol az r' pont az $y = f(a) + u(x - a)$ és $y = f(c) + \ell(x - c)$ egyenesek metszéspontja, a t' pont pedig az $y = f(c) + u(x - c)$ és $y = f(b) + \ell(x - b)$ egyenesek metszéspontja. Ebből kapjuk a megfelelő formulákat:

$$x_{r'} = \frac{f(c) - f(a) + ua - \ell c}{u - \ell}, \qquad y_{r'} = \frac{uf(c) - \ell f(a) + (a - c)\ell u}{u - \ell},$$

$$x_{t'} = \frac{f(b) - f(c) - \ell b + uc}{u - \ell}, \qquad y_{t'} = \frac{uf(b) - \ell f(c) + (c - b)\ell u}{u - \ell}.$$

Ezeket együtt használva a (2.6), (2.7), (2.8) és (2.9) képletekkel kapjuk az $f(Y)$ alsó- és felső korlátjait. Könnyű látni, hogy a felső korlátra is érvényes a 5. Tételben

megfogalmazott állítás: legalább olyan jó, mint a középponti formula vagy az lbvf által adódó felső korlát. Továbbá az 6. Tétel is átvihető a megfelelő módosításokkal a felső korlátra vonatkozó számításainkra.

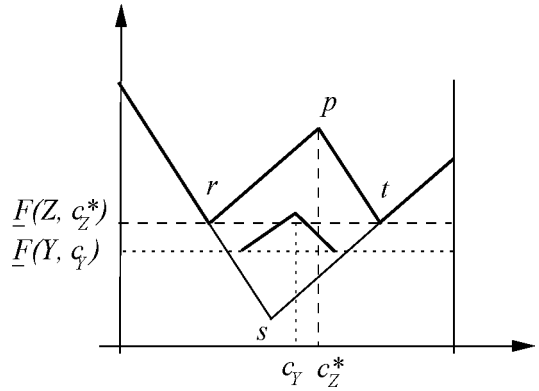
Megjegyzés. LAGOUANELLE & SOURBY [32] javaslata alapján a kite befoglalás általánosítható úgy, hogy ne csak egy c^* kifejtési pontra támaszkodjon, hanem az aktuális intervallumon belül válasszunk p darabot ezekből (így kapjuk a p-kite befoglalást). Ez a stratégia arra is jó, hogy az intervallum felosztás ne felezés legyen, hanem több részre osztás (multisection). A cikkben numerikus eredmények nem találhatók, így kérdéses, hogy a javasolt eljárás milyen befolyással van a hatékonyságra, amennyiben azt globális optimalizálási algoritmusban használjuk.

2.4.2. A kite befoglalás tulajdonságai

Ebben a részben a kite befoglaló függvény néhány –az intervallumos módszerek számára fontos– tulajdonságát tárgyaljuk.

7. Tétel. [65] *Tegyük fel, hogy az F' befoglalás izoton és legyen az f befoglalását adó F függvény a kite algoritmussal adott, azaz $F(Y) = [\underline{F}_K(Y, c^*), \overline{F}_K(Y, c')]$ minden $Y \in \mathbb{I}(X)$ intervallumra. Akkor az F befoglalás izoton.*

Bizonyítás. Legyen $Y \subset Z = [a, b]$ adott és c_Z^* a kite maximum helye a Z intervallumon. Először megmutatjuk azt, hogy $\underline{F}_K(Y, c) \geq \underline{F}_K(Z, c_Z^*)$ igaz minden $c \in Y$ pontra. Legyen $F'(Z) = [\ell, u]$ és $F'(Y) = [\ell', u']$. Ha $\ell' \geq 0$ teljesül, akkor legyen $\underline{F}_K(Y, c) := [f(\underline{Y}), f(\overline{Y})]$ minden $c \in Y$ pontra, vagy ha $u' \leq 0$ teljesül, akkor pedig legyen $\underline{F}_K(Y, c) := [f(\overline{Y}), f(\underline{Y})]$ minden $c \in Y$ pontra. Mindkét esetben $\underline{F}_K(Y, c) \geq \underline{F}_K(Z, c_Z^*)$ igaz, hiszen az $f(\underline{Y})$ és $f(\overline{Y})$ értékek nem lehetnek az $y = f(a) + \ell(x - a)$ és $y = f(b) + u(x - b)$ egyenesek alatt.



2.6. ábra. Az ábrán megpróbálunk konstruálni olyan f függvényt, amely nem engedi meg a kite izotonitását.

Az $\ell' < 0 < u'$ esetre egy indirekt bizonyítást adunk. Az ötletet a 2.6. ábra adja: legyenek az ℓ és u értékek adottak, azaz az $y = f(a) + \ell(x-a)$ és $y = f(b) + u(x-b)$ egyenesek rögzítettek. Megpróbálunk olyan f függvényt konstruálni, amelynek befoglalása megsérti az izotonitást. A 2.6 ábrán $\underline{F}(Z, c_Z^*)$-gal jelzett szaggatott vonal a Z intervallumon a kite-ot maximalizáló c_Z^* pont által adott befoglalás alsó korlátja. Olyan $Y \subset Z$ intervallumot konstruálunk, amelyben a kite c_Y középpontja olyan $(c_Y, f(c_Y))$ pont, amely eredményeképpen izotonitást sértő befoglalást kapunk. Könnyű látni, hogy ilyen pont csak a p, t, r és s pontok által meghatározott paralelogrammában létezhet, mivel csak az ottani pontok adhatnak alacsonyabb $\underline{F}$ értéket. Ez viszont ellentmondásra vezet, mivel a $(c_Y, f(c_Y))$ és $(c_Z^*, f(c_Z^*))$ pontok által meghatározott egyenes meredeksége nincs benne az $[\ell, u]$ intervallumban. Következésképpen nincs olyan c_Y pont, amelyre $\underline{F}_K(c_Y) < \underline{F}_K(c_Z^*)$ teljesülne.

Az $\overline{F}_K(c_Y) \leq \overline{F}_K(c'_Z)$ eset bizonyítása a fentiekkel analóg. Itt a $c'_Z \in Z$ a kite felső korlátját minimalizáló pont, míg $c_Y \in Y$. □

8. Tétel. [65] *Ha a derivált befoglalása Lipschitz-folytonos, akkor a kite algoritmus által adott befoglalás α-konvergens, ahol $\alpha \geq 2$.*

Bizonyítás. A 5. Tételből tudjuk, hogy a kite algoritmus legalább olyan jó, mint a középponti formula. Tudjuk továbbá, hogy a középponti formula négyzetesen konvergens, ha $F'(X)$ Lipschitz-folytonos (Krawczyk & Nickel [31]). Következésképpen a kite algoritmus által adott befoglalás is legalább négyzetesen konvergens. Továbbá legalább akkora α érték érvényes F_K-ra, mint F_{CF}-re. □

2.4.3. Metszés

Mint a bevezetőben említettük, az intervallumos globális optimalizálásban számos gyorsító eljárás létezik. Ezen ún. tesztek lényege, hogy a keresési tér minél nagyobb olyan részeit eltávolítsák, amelyek garantáltan nem tartalmaznak globális minimum pontot. A Ratz [55] cikkben egy lejtő aritmetikán alapuló metszési (pruning) technikáról olvashatunk. Hasonló eljárás dolgozható ki a kite befoglaló függvényre is. Ez a jelen alfejezet témája.

9. Tétel. [65] *Legyen $Y = [a, b] \subseteq X$ az aktuálisan vizsgált intervallum, $c^* \in [a, b]$ egy maximumhelye az $\underline{F}_K(Y, \cdot)$ függvénynek, továbbá $\tilde{f}$ egy garantált felső korlát az f globális minimumára. Definiáljuk a következő értékeket:*

$$p = a + \frac{\tilde{f} - f(a)}{\ell}, \qquad q = c^* + \frac{\tilde{f} - f(c^*)}{u},$$
$$r = c^* + \frac{\tilde{f} - f(c^*)}{\ell}, \qquad s = b + \frac{\tilde{f} - f(b)}{u}.$$

Ha $\ell < 0 < u$, akkor a kite algoritmusban használhatjuk a következő kivágási technikákat:

(a) *Ha $\tilde{f} < \min\{f(a), f(b), f(c^*)\}$, akkor $[p, q] \cup [r, s]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

(b) *Ha $f(b) \leq \tilde{f} < \min\{f(a), f(c^*)\}$, akkor $[p, q] \cup [r, b]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

(c) *Ha $f(a) \leq \tilde{f} < \min\{f(b), f(c^*)\}$, akkor $[a, q] \cup [r, s]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

(d) *Ha $f(c^*) \leq \tilde{f} < \min\{f(a), f(b)\}$, akkor $[p, s]$ tartalmazza az összes Y-ban lévő globális minimumpontot..*

(e) *Ha $\max\{f(b), f(c^*)\} \leq \tilde{f} < f(a)$, akkor $[p, b]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

(f) *Ha $\max\{f(a), f(c^*)\} \leq \tilde{f} < f(b)$, akkor $[a, s]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

(g) *Ha $\max\{f(a), f(b)\} \leq \tilde{f} < f(c^*)$, akkor $[a, q] \cup [r, b]$ tartalmazza az összes Y-ban lévő globális minimumpontot.*

Bizonyítás. (a) Legyen $z \in [a, b]$ úgy, hogy $f(z) = \min_{x \in [a,b]} f(x)$ és $\tilde{f} \geq f(z)$ (lásd a 2.7. ábrát). Meg kell mutatnunk, hogy

$$a + \frac{\tilde{f} - f(a)}{\ell} \leq z. \tag{2.14}$$

Tudjuk, hogy minden $x \in [a, b]$ pontra az $f(a) + \ell(x - a) \leq f(x)$ egyenlőtlenség teljesül. Ha $x = z$, akkor $\ell(z - a) \leq f(z) - f(a)$, ami ekvivalens az

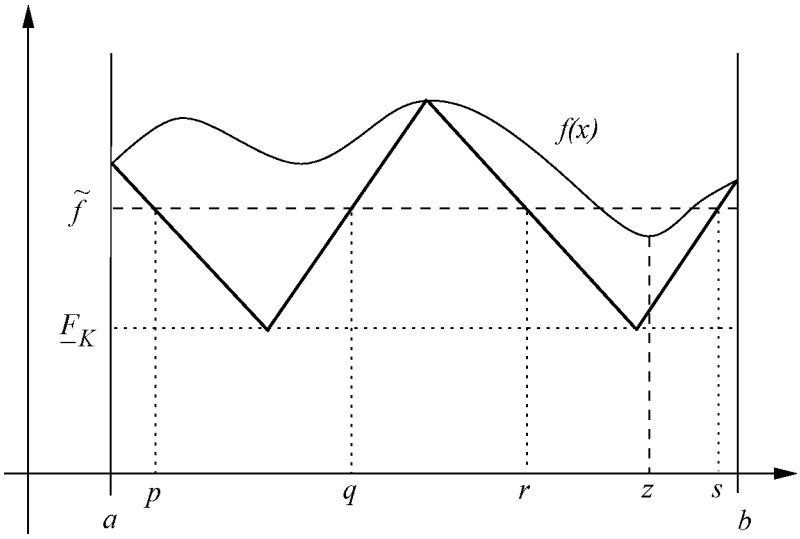
$$\ell \leq \frac{f(z) - f(a)}{z - a} \leq \frac{\tilde{f} - f(a)}{z - a}, \tag{2.15}$$

relációval, amennyiben $z \neq a$. Ha $z = a$, akkor (2.14) teljesül, hiszen $\tilde{f} < f(a)$. A (2.15) képletből (mivel $z > a$) kapjuk, hogy

$$\ell z \leq \ell a + \tilde{f} - f(a)$$

ami (2.14) bizonyítását adja, mert $\ell < 0$. Annak bizonyításához, hogy

$$z \leq c^* + \frac{\tilde{f} - f(c^*)}{u} \tag{2.16}$$



2.7. ábra. A kite metszési tulajdonsága. Az ábrán a 9. Tétel (a) esetét látjuk: az $[a,p),(q,r)$ és $(s,b]$ intervallumokat törölhetjük, azok garantáltan nem tartalmaznak globális minimumpontot.

is teljesül, használjuk az $f(c^*) + u(x - c^*) \leq f(x)$ egyenlőtlenséget, amely igaz minden $x \in [a, c^*]$, $c^* \in [a, b]$ pontra. Ha $x = z$, akkor

$$u \leq \frac{f(z) - f(c^*)}{z - c^*} \leq \frac{\tilde{f} - f(c^*)}{z - c^*},$$

teljesül, amennyiben $z \neq c^*$. Ebből az egyenlőtlenségből kapjuk az

$$uz \leq uc^* + \tilde{f} - f(c^*),$$

összefüggést, ami bizonyítja a (2.16) relációt, mert $u > 0$. A $z = c^*$ eset nem lehetséges, hiszen feltettük, hogy $\tilde{f} < f(c^*)$ teljesül, továbbá $f(z) \leq \tilde{f}$.
Az $x^* \in [r, s]$ esetre vonatkozó rész bizonyítása a fentiekhez hasonló meggondolásokkal elvégezhető.

(b) – (g) Ezeket az eseteket az (a) esettel megegyező módon bizonyíthatjuk. □

Vegyük észre, hogy a kivágás használatához nincs szükségünk további információra, minden értéket, amit a 9. Tételben használunk már előzetesen kiszámoltunk a kite befoglaló függvény előállításához. A metszési tulajdonság hatékonyságát numerikus tesztekkel támasztjuk majd alá.

Példa. A fenti meggondolásokat egy egyszerű példán szemléltetjük. Legyen $f(x) = x^2 - x, X = [0, 0.75]$. A globális minimum $f^* = -0.25$, a minimumhely pedig $x^* = 0.5$. Automatikus differenciálással (vagy „kézzel” számolva) kapjuk, hogy

$F'(X) = [-1, 0.5]$. Használva a fenti formulákat adódnak a következő $F(X)$-re vonatkozó alsó korlátok:

$$\begin{aligned} \underline{F}_{CF}(c^-) &= -0.5 \\ \underline{F}_{LBVF} &= -0.375 \\ \underline{F}_K(c^*) &= -0.31066, \end{aligned}$$

ahol c^* megközelítőleg 0.43934. Láthatjuk, hogy az optimális c^* pontot használó kite algoritmus adja a legjobb alsó korlátot. Használva a kivágási technikát, a $X_1 = [0, 0.25)$ és $X_2 = (0.63, 0.75]$ intervallumok eldobhatók, csak az $X' = [0.25, 0.63]$ részintervallum tartalmazhat globális minimumhelyet.

Megjegyzés. Érdekes észrevétel, hogy a metszés annál hatékonyabb, minél távolabb van a kite kifejtési pontjához tartozó függvényérték geometriai értelemben az $\tilde{f}$ vonaltól. Ezért lényegében a minél jobb metszési hatékonyság érdekében egy lokális maximalizálást kellene végrehajtanunk. Ez azonban jelentősen megnövelné a költségeket és a kite befoglalás sem lenne (általában) optimális, ezért az ötlet alkalmazását elvetjük.

2.4.4. Kiterjesztett kite algoritmus

Most a kite algoritmus kiterjesztését részletezzük, amelyet aztán beépíthetünk az intervallumos globális optimalizáló eljárásba. Láttuk, hogy a kite befoglaló függvény és a metszési teszt használatához elsőrendű derivált információra van szükség.

i. lépés Számítsuk ki az $f(a), f(b)$, és $F'(X) = [\ell, u]$ értékeket.

ii. lépés Ha $\ell < 0 < u$, akkor határozzuk meg a c^* pontot (vagy annak egy (jó) közelítését). Értékeljük ki az $f(c^*)$ kifejezést és számítsuk ki $\underline{F}_K(c^*)$ értékét.

iii. lépés Ha $\tilde{f} > \min\{\overline{F}(c^*), \overline{F}(a), \overline{F}(b)\}$, akkor frissítsük $\tilde{f}$-t. Alkalmazzuk a kivágási tesztet az $\overline{F}(c^*)$ segítségével.

iv. lépés Alkalmazzuk a metszés eljárást a 9. Tétel alapján.

Ezen algoritmus a B&B algoritmus 4. lépésébe illeszthető be.

A részletezett algoritmusban láthatjuk, hogy a ii. lépés költséges is lehet – attól függően, hogy milyen módszert használunk a c^* meghatározására. A konkrét megvalósításban itt az $x_{s'}$ és x_s pontok által meghatározott intervallum középpontját vettük közelítésnek. Tapasztalataink szerint ez az olcsó becslés megfelelő.

Figyelembe véve a fenti meggondolásokat, a következő állítással zárjuk elméleti vizsgálódásainkat.

2. Következmény. [65] *A javasolt algoritmust használva soha nem veszíthetünk a kiindulási X intervallumban lévő globális minimum pontokból. A metszési lépésben ha egy adott Y intervallumra üres intervallumot kapunk, akkor f-nek nincs (az X intervallumra nézve) globális minimumhelye Y-ban.*

2.4.5. Numerikus eredmények

Ebben a szakaszban a fenti elméleti eredmények numerikus igazolását mutatjuk be. Láttuk, hogy az optimális kite befoglaló függvény mindig jobb befoglalást ad, mint a középponti alak vagy az lbvf. A B&B algoritmusban történő alkalmazása hatékonyság szempontjából azonban kérdéses lehet, hiszen a kite előállításához több függvénykiértékelésre van szükségünk. A szakasz célja tehát, hogy igazoljuk, a kite befoglaló függvény használata az intervallumos globális optimalizálási algoritmusban hatékonyság növekedést eredményez.

Összesen 40 darab egyváltozós standard tesztfüggvényt vizsgáltunk meg (ezek leírását lásd Casado *et al.* [7]). A számításokat egy duál processzoros Pentium-II gépen (233 MHz, 256 Mbyte), Linux operációs rendszerben végeztük el. Programozási környezetként a C++ Toolbox for Verified Computing [26] és a C-XSC [25] programcsomagokat használtuk.

A megvalósításban fontos volt, hogy a kite módszer több információt igényel, mint a hagyományos módszerek. Hogy csökkentsük a redundás számítások mennyiségét a következő megfontolásokat tettük:

- Az optimális c^* pont értéke jól közelíthető az x_s és $x_{s'}$ pontok által adott intervallum közepével (ahelyett, hogy intervallumos Newton módszert alkalmaznánk a (2.13) egyenletre). Ezzel a technikával a számítások mennyisége csökkenthető, az $F(c)$ és $F'(c)$ intervallumokat nem kell kiszámítanunk. Az így megadott kite középpont jó közelítése az optimálisnak, ha a vizsgált részintervallum már elegendően keskeny.

- A metszési lépést csak az egyes iterációk végén végezzük el, mivel az megváltoztathatja az aktuális részintervallum végpontjait. Ha nem így teszünk, a végpontokban vett függvényértékeket esetleg újra ki kell számolnunk.

A következőkben két megvalósításra kapott numerikus eredményeket ismertetünk és elemezünk. Az első esetben csak gradiens információt használtunk, míg a másik esetben a másodrendű derivált befoglalását is használhatjuk. Mindkét változat esetén megállási feltételként a

$$\mathrm{wid}_{\mathrm{rel}}([\underline{F}_K, \tilde{f}]) \leq 10^{-12} \quad \text{vagy} \quad \mathrm{wid}_{\mathrm{rel}}(Y) \leq 10^{-12}$$

relációkat használtuk. Az összehasonlításban a [26] könyvben megadott és megvalósított eljárást használtuk; ez középponti formulát használ befoglaló függvényként. A hatékonyság mérésére a következő mutatókat vizsgáltuk: függvénykiértékelések száma + 2×(deriváltkiértékelések száma) + 3×(másodrendű deriváltkiértékelések száma). Vizsgálataink azt mutatják, hogy ez egy korrekt súlyozás a teljes műveletigény leírására.

Megjegyezzük továbbá, hogy a teljes CPU idő a tesztfeladatsor megoldására kevesebb volt, mint egyetlen másodperc, ezért annak feltüntetésétől és az ezen alapuló összehasonlítástól eltekintünk.

Elsőrendű algoritmus

Az algoritmus variánsok a kivágási- és monotonitási teszteket tartalmazták és a középponti formát, illetve a kite algoritmust használták kivágással és anélkül. Ezek a változatok tehát csak elsőrendű információt használtak. A numerikus eredményeket a 2.1. táblázat tartalmazza.

Mindhárom változat az összes tesztfeladatot sikeresen megoldotta. Minden tesztfüggvényre a függvénykiértékelések számát, a deriváltkiértékelések számát, a biszekciók számát és a felhasznált maximális listahosszat tüntettük fel. Ezek a mutatók mindhárom változatnál szerepelnek. A táblázat végén a megfelelő mutatók összegei, illetve az új módszer(ek)nek a hagyományos eljáráshoz viszonyított százalékos összevetése szerepel.

A függvénykiértékelések számának összege 15706 volt a tradícionális változat esetén, míg a kite módszerekre 8416 és 11710 attól függően, hogy használtuk-e a kivágást vagy nem. Ezek rendre 46%-os illetve 26%-os hetékonyság javulást jelentenek. Az itt tapasztalható javulást jórészt a nehezebb tesztfeladatokon értük el, speciálisan az utolsó két függvényre a számítások 38%-os és 37%-os csökkenését tapasztaltuk.

A deriváltkiértékelések száma 9646 volt a hagyományos módszerre, míg a kite algorimus használatával rendre 3406 és 5536 volt a kivágással és nélküle. Ez arányaiban 65%-os, illetve 43%-os javítást jelent.

A hatékonysági mutató (ami a teljes műveletigényt jelenti) 34998 a régi módszerre, a kite módszerre a kivágás használatával 15228, annak használata nélkül pedig 22728. Ez 56%-os, illetve 35%-os hatékonyság növekedést jelent. Megállapíthatjuk tehát, hogy a javasolt kite módszer használata jelentősen felgyorsítja az optimalizáló eljárás sebességét.

Az alkalmazott intervallum-felezések száma 3.114 volt a hagyományos módszernél, 1683 és 2728 az új eljárásnál a metszéssel, illetve metszés nélkül, ami 46%-es, illetve 13% hatékonyság-növekedést jelent. Ez a mutató a hatékonysági mutatóval együtt jelzi, hogy a kite módszernél a metszés műveletet érdemes használni.

2.1. táblázat. Elsőrendű algoritmus numerikus eredményei.

Fel-	F kiért. száma			D kiért. száma			biszekciók száma			max. lista hossz		
adat	cf	k+pr	k	cf	k+pr	k	cf	k+pr	k	cf	k+pr	k
1	84	60	93	53	25	44	25	12	21	4	5	3
2	88	77	99	55	31	46	26	15	22	5	6	5
3	98	93	100	55	39	48	25	19	23	3	5	2
4	96	95	119	59	35	56	27	17	27	4	7	4
5	109	100	141	69	45	68	33	22	33	3	3	2
6	88	69	110	55	27	52	25	13	25	3	4	3
7	79	57	95	51	25	46	23	12	22	2	2	2
8	92	80	115	59	37	56	27	18	27	2	6	2
9	94	84	118	61	37	58	28	18	28	2	4	2
10	89	76	113	57	33	54	27	16	26	2	3	2
11	83	70	101	53	29	48	24	14	23	3	2	1
12	83	68	103	53	29	50	24	14	24	2	3	2
13	91	71	111	59	33	54	27	16	26	3	4	3
14	118	95	138	77	41	66	36	20	32	2	3	2
15	107	85	132	69	33	66	32	16	32	6	13	6
16	113	107	138	73	49	70	35	24	34	8	12	8
17	109	107	128	71	47	62	33	23	30	2	6	3
18	153	105	117	99	49	56	47	24	27	4	4	3
19	95	72	93	59	31	44	27	15	21	4	4	3
20	82	52	79	53	23	38	24	11	18	1	3	1
21	83	64	79	53	29	38	25	14	18	2	2	1
22	145	125	147	93	57	68	43	28	33	4	6	3
23	161	144	167	103	63	78	47	31	38	3	5	3
24	158	136	166	103	65	76	47	32	37	3	6	3
25	155	128	192	101	59	94	47	29	46	4	6	3
26	223	110	202	145	51	98	69	25	48	4	5	3
27	179	143	220	117	69	108	55	34	53	4	6	4
28	229	122	226	149	55	110	69	27	54	4	5	3
29	215	156	209	139	67	92	63	33	45	4	5	4
30	310	212	302	203	101	148	93	50	73	4	8	4
31	88	75	99	57	33	48	26	16	23	2	3	2
32	602	251	386	395	119	188	186	59	93	8	15	8
33	345	272	401	225	131	198	107	65	98	17	18	15
34	292	216	242	189	101	102	86	50	50	8	7	5
35	88	74	87	57	33	42	26	16	20	1	2	1
36	762	559	529	383	197	212	177	98	105	14	14	9
37	352	289	291	217	117	122	101	58	60	10	13	7
38	1026	567	530	675	253	212	201	126	105	12	14	4
39	3965	1519	1409	2329	511	564	436	255	281	109	31	19
40	4377	1631	3583	2673	597	1856	635	298	927	97	40	77
Σ	15706	8416	11710	9646	3406	5536	3114	1683	2728	379	310	237
		54%	74%		35%	57%		54%	87%		82%	63%

A maximális listahosszak összege 379 a hagyományos módszerre, 310 és 237 az új eljárásra a metszéssel és nélküle. Ez 18%-os és 37%-os javulást jelent.

Ezekből az eredményekből láthatjuk, hogy amennyiben az f célfüggvény deriváltjának befoglalása rendelkezésre áll, akkor érdemes használni a kite módszert. Numerikus vizsgálataink azt mutatják, hogy ekkor a vizsgált tesztfeladatok kisebb számítási ráfordítással oldhatók meg.

Másodrendű algoritmus

Ebben az alfejezetben a másodrendű deriváltat is használó algoritmusra mutatunk numerikus eredményeket. Az eljárásban a kivágási tesztet, monotonitási tesztet, konkavitási tesztet és az intervallumos Newton lépést használtunk. Futtatási eredményeinket a 2.2. táblázat tartalmazza, ahol a mutatók ismét a függvénykiértékelések száma, deriváltkiértékelések száma, másodrendű derivált kiértékelések száma, az alkalmazott intervallum felezések száma és a maximális listahossz.

Ezek az algoritmus változatok sokkal kifinomultabbak, minden szokásos gyorsító eljárást tartalmaznak, ezért nem számíthatunk nagy mértékű teljesítmény növekedésre.

A megvalósításban az aktuális intervallum felezése után egy természetes intervallumos kiértékelést alkalmaztunk. A monotonitási teszt után kivágási tesztet, konkavitási tesztet és egy Newton lépést hajtottunk végre. A kite befoglalást a metszéssel együtt csak a Newton lépés által visszaadott részintervallumokra alkalmaztuk. Vizsgálataink azt mutatják, hogy a (d), (e) és (f) kivágási lépéseket érdemes használni, hiszen ezek csak egy részintervallumot adnak eredményül, ami jobban alkalmazkodik ehhez az algoritmus változathoz. A gyorsító eljárások ilyen módon való használata a számítási költségek csökkentéséhez vezetett.

A másodrendű algoritmusokra a függvénykiértékelések száma 4029 volt a hagyományos esetben, míg 3101 és 3401 az új módszernél metszéssel és nélküle. Ez 23%-os, illetve 16%-os hatékonyság növekedést jelent.

A deriváltkiértékelések száma 2747 volt a hagyományos módszer esetén, 1659 és 1899 az új módszerrel metszéssel és nélküle. Ez 40%-os és 31%-os növekedést jelent a hatékonyságban.

A másodrendű derivált kiértékelések száma összesen 612 volt a régi módszernél, míg 638 illetve 733 az új módszerrel metszéssel és nélküle. Ez azt mutatja, hogy a kite módszer hatékonysága ebből a szempontból romlott. Viszont a tesztfüggvényekre a másodrendű derivált kiértékelések száma csekély a függvény- és deriváltkiértékelésekhez képest, ezért ez a mutató nem rontja le nagyon a hatékonyságot.

A teljesítmény mutató értéke 11359 a hagyományos algoritmusra, 8333 és 9398 a kite módszerrel metszést alkalmazva, illetve metszés nélkül. Ez 27%-os, illetve 17%-

2.2. táblázat. Másodrendű algoritmus numerikus eredményei.

Fel-	F kiért. száma			D kiért. száma			H kiért. száma			biszekciók			max. lista hossz		
adat	cf	k+pr	k	cf	k+pr	k	cf	k+pr	k	cf	k+p	k	cf	k+pr	k
1	60	38	48	40	18	25	8	6	9	6	3	4	4	3	3
2	74	49	56	48	23	30	10	10	11	6	5	5	4	4	4
3	66	49	52	45	26	29	11	12	13	7	6	6	1	1	1
4	67	68	71	44	30	33	9	10	11	6	5	5	8	4	4
5	95	74	84	67	42	49	13	16	19	10	8	9	3	2	2
6	58	38	52	38	18	29	8	8	11	5	4	5	4	1	1
7	43	35	38	29	18	21	6	8	9	5	4	4	2	2	2
8	51	35	39	34	18	22	7	8	9	5	4	4	2	1	1
9	52	43	46	35	23	26	7	10	11	6	5	5	2	2	2
10	55	42	53	37	19	27	8	8	11	6	4	5	2	3	3
11	44	35	38	29	18	21	6	8	9	5	4	4	2	1	1
12	45	47	43	30	21	20	6	8	7	5	4	3	3	1	1
13	59	50	53	40	27	30	8	12	13	7	6	6	3	3	3
14	57	46	60	40	26	34	8	10	15	7	5	7	2	2	2
15	106	84	94	69	43	50	14	14	17	9	7	8	8	6	6
16	118	100	103	77	53	56	15	16	17	11	8	8	8	5	5
17	54	43	46	37	23	26	9	10	11	2	5	5	2	2	2
18	84	51	70	58	25	38	13	10	16	9	5	7	3	2	2
19	74	40	50	48	20	27	10	8	11	6	4	5	5	2	2
20	43	36	46	30	19	26	6	8	11	5	4	5	1	1	1
21	50	36	39	34	19	22	7	8	9	5	4	4	2	1	1
22	77	58	61	54	35	38	13	12	13	7	6	6	1	1	1
23	92	75	76	62	40	44	15	14	15	8	7	7	6	4	4
24	73	55	58	50	32	35	12	12	13	7	6	6	1	1	1
25	92	66	80	64	34	45	14	14	18	10	7	8	2	2	2
26	113	92	99	80	48	55	17	20	23	13	10	10	3	3	3
27	92	62	82	64	33	47	15	14	20	10	7	9	3	3	3
28	103	86	93	72	45	52	16	20	23	12	10	10	3	3	3
29	53	43	46	35	23	26	5	8	9	5	4	4	2	1	1
30	144	121	130	103	65	74	22	28	32	17	14	14	5	4	4
31	51	44	47	36	24	27	6	10	11	6	5	5	2	1	1
32	275	158	174	195	84	100	43	32	40	31	16	16	8	7	7
33	362	229	243	243	128	139	48	56	61	33	28	29	17	16	16
34	116	90	96	81	52	55	19	18	19	12	9	9	5	3	3
35	51	44	54	36	24	31	7	10	13	6	5	6	1	1	1
36	207	191	186	140	108	106	32	36	35	20	18	17	9	7	7
37	120	102	105	79	49	52	18	16	17	12	8	8	5	3	3
38	162	118	143	111	71	84	30	26	32	15	13	15	6	4	4
39	273	218	238	188	123	137	50	44	50	25	22	24	8	9	9
40	218	210	209	145	112	111	36	40	39	20	20	19	10	9	9
Σ	4029	3101	3401	2747	1659	1899	612	638	733	406	319	336	168	131	131
		77%	84%		60%	69%		104%	120%		79%	83%		78%	78%

os hatékonyság növekedést jelent. Itt megint azt láthatjuk, hogy a kite módszer a metszést használva éri el a nagyobb hatékonyságot.

Az alkalmazott intervallum felezések száma 406 volt az eredeti módszernél, míg 319 és 336 az új módszernél a metszéssel és anélkül. Ez 21%-os, illetve 17%-os javítást jelent.

A felhasznált maximális listaelemek összege 168 a régi módszer esetén, míg 131 az új eljárásban, ami 22%-os javulást eredményezett.

Összegezve az eredményeket láthatjuk, hogy a másodrendű algoritmus alkalmazásakor is jobb teljesítményt érhetünk el. Bár a hatékonyság nem javult olyan mértékben, mint az elsőrendű algoritmus esetében, az új befoglaló függvény használata így is javasolt.

Összehasonlítás más módszerekkel

Ebben az alfejezetben az egydimenziós kite módszert hasonlítjuk össze két hasonló eljárással, amelyek a célfüggvényről elsőrendű információt használnak, illetve alkalmazzák a metszés technika megfelelő változatát is.

A kite módszer kidolgozásával nagyjából azonos időben megjelent cikkben Casado *et al.* [7] közöl elméleti és numerikus eredményeket, amelyek az alap B&B módszer gyorsítását érték el. Módszerük lényegében az lbvf formula alkalmazása egy metszési technikával.

Egy nemrégiben megjelent cikkben Sotiropoulos & Grapsa [61] a Baumann középponti formulára fejlesztett ki kifinomult metszési technikát. A cikkben elméleti és numerikus eredmények egyaránt azt mutatják, hogy módszerük hatékonyabb az eddig ismerteknél. (Ez természetesen nem a befoglalás jóságára vonatkozik – hiszen láttuk, hogy a kite befoglaló függvény sohasem rosszabb, mint a Baumann középponti formula –, hanem a B&B algoritmusba történő alkalmazás hatékonyságára.) Az ott közölt adatok előállításához viszont a kite eljárásban nem alkalmazták a metszési technikát, ami nélkül (mint azt láttuk) általában rosszabb eredményeket kapunk. Másrészt a kite befoglalást és metszést használó B&B módszer tovább javítható. A következő algoritmus erre tesz javaslatot.

A lépés. Legyen X a kezdő intervallum, $\mathcal{L}$ a munkalista, $\mathcal{Q}$ pedig az eredménylista. Alkalmazzuk a kite befoglalást az X intervallumon. Legyenek $\mathcal{Q} := \{\}$ és $\mathcal{L} := \{(X, c^*, f(a), f(b), f(c^*), \underline{F'}(X), \overline{F'}(X), \underline{F}(X))\}$, és állítsuk be az f^* értékre vonatkozó garantált felső korlátot: $\tilde{f} = \overline{F}(c^*)$.

B lépés. Mindaddig, amíg $\mathcal{L}$ nem üres, hajtsuk végre a következő lépéseket.

C lépés. Vegyünk le az $\mathcal{L}$ lista legelső elemét. A rendelkezésre álló értékek alapján alkalmazzuk a kite metszési technikát.

D lépés. Legyenek U_1 és U_2 a metszés által kapott részintervallumok (ahol lehet, hogy U_2 üres). Amennyiben a metszés sikertelen volt, vágjuk ketté az aktuálisan vizsgált intervallumot az optimális c^* pontban (ekkor ezek lesznek az U_1 és U_2 részintervallumok.

E lépés. Az $i = 1, 2$ indexre számítsuk ki az U_i intervallumon a derivált befoglalását, alkalmazzunk monotonitási tesztet és a kite befoglalást. Aktualizáljuk az $\tilde{f}$ értékét.

F lépés. Ha U_i-re ($i = 1, 2$) a megállási feltételek teljesülnek, akkor $\mathcal{Q} := \mathcal{Q} + \{U_i, \underline{F}_K(U_i)\}$, különben pedig $\mathcal{L} := \mathcal{L} + \{(U_i, c^*_{U_i}, f(\underline{U_i}), f(\overline{U_i}), f(c^*_{U_i}), \underline{F'}(U_i), \overline{F'}(U_i), \underline{F}(X))\}$ (ahol $c^*_{U_i}$ az U_i részintervallumon vett optimális kite kifejtési pontot jelenti), és menjünk vissza a B lépésre.

Az itt ismertetett eljárás hatékonyabb, mint amit a 2.4.4. alfejezetben alkalmaztunk, hiszen lehetővé teszi a már egyszer kiszámolt függvényértékek újbóli felhasználását, amivel az eljárás összköltsége csökkenthető.

A 2.3. táblázat az [61] cikkből vett adatokat tartalmazza, a kite módszerrel ott kapott eredmények kivételével, ahelyett az imént ismertetett változat implementálásával kapott eredményeket közöljük. A táblázatban csak a 40 tesztfeladat megoldásakor összesen felhasznált függvény- és deriváltkiértékelések számát, az aktuális intervallumon alkalmazott kettévágások számát és a felhasznált leghosszabb lista elemszámát tüntettük fel. A B módszer az alap B&B eljárást jelöli, amely a Baumann középponti formulát használja, C a Casado *et al.* [7] cikkben leírt befoglalófüggvényt és metszést, K a kite módszert az imént ismertetett algoritmussal, végül SG pedig a Sotiropoulos & Grapsa [61] cikkben közölt módszert.

Az összehasonlított módszerek mindegyike használta a kivágási tesztet és a monotonitási tesztet. A megállási feltétel a $\mathrm{wid}_{\mathrm{rel}}(Y) \leq 10^{-8}$ volt.

2.3. táblázat. Az elsőrendű kite módszer összehasonlítása más hasonló módszerekkel.

F kiért. száma				D kiért. száma				biszekciók száma				max. lista hossz			
B	C	K	SG	B	C	K	SG	B	C	K	SG	B	C	K	SG
7124	10351	5519	4487	4068	3430	1920	2509	2014	600	229	260	220	311	161	199
	102%	77%	63%		84%	47%	62%		30%	9%	13%		141%	73%	90%

Ha kiszámoljuk a teljesítmény mutatókat a C, K és SG módszerekre rendre a 17211 (113%), 9359 (61%) és 9505 (62%) értékeket kapjuk a Baumann középponti formát használó alap algoritmushoz képest, ami szerint a kite módszer megfelelő algoritmikus környezetben legalább olyan hatékony, mint az [61] által javasolt technika. Láthatjuk továbbá, hogy az alkalmazott intervallum felezések és a tárméret is a kite esetében a legígéretesebb.

2.5. Kite befoglaló függvény – többdimenziós eset

Jelen fejezet az imént bemutatott kite befoglaló függvény egy lehetséges magasabb dimenziós kiterjesztését tárgyalja a VINKÓ & RATZ [68] cikk alapján.

A fejezet hátralevő részében az $f'(y)$ gradiensvektor egy befoglalását $F'(Y)$ jelzi, míg ezen vektor i-edik komponensére az $F'_i(Y) = [\ell_i, u_i]$ jelölést használjuk a könnyebb olvashatóság kedvéért. Feltesszük továbbá, hogy minden $i = 1, \ldots, n$ indexre $\ell_i u_i < 0$ teljesül.

2.5.1. A kite befoglalás komponensenkénti kiterjesztése

Mint azt láttuk, a kite befoglaló függvény az lbvf és a középponti forma egyszerű szimultán használatából vezethető le a kifejtési pont megfelelő megválasztásával. Magasabb dimenziókra az lbvf kiterjesztését MESSINE & LAGOUANELLE [41] tárgyalja. A középponti formák természetes módon vihetők át a magasabb dimenziós térre. Ezen két módszer szimultán használata viszont nagyon komplikált, nehezen kivitelezhető és optimalizáló eljárásokba való használata –a magas műveletigény miatt– egyáltalán nem javasolt. A továbbiakban egy hatékony és könnyen implementálható kiterjesztést tárgyalunk.

RATZ [54] munkájában a lejtő aritmetikán alapuló középponti formák és azok metszési eljárásáról találhatunk értekezést, ahol a szerző egy komponensenkénti kiterjesztést javasol. A kite kiterjesztése ezen az ötleten alapszik.

Legyen adott az $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ függvény és az $Y = Y_1 \times \ldots \times Y_n \subseteq D$ intervallum. Definiáljuk a $g_i : Y_i \subseteq \mathbb{R} \to \mathbb{I}$ ($i \in \{1, \ldots, n\}$) függvényt úgy, hogy

$$g_i(w) := f(Y_1, \ldots Y_{i-1}, w, Y_{i+1}, \ldots, Y_n), \quad w \in Y_i.$$

Az egydimenziós intervallumos függvények ilyen használatával az egydimenziós kite befoglalást is használhatjuk. Ha adottak a $V \supseteq g_i(\underline{Y}_i), W \supseteq g_i(\overline{Y_i})$ és $Z \supseteq g_i(c_i)$ ($c_i \in Y_i$) befoglalások, akkor a *komponensenkénti kite befoglalás* konstruálható a *komponensenkénti középponti forma*:

$$\underline{F}_{CF}(Y, c, i) = \underline{Z} + F'_i(Y)(Y_i - c_i), \quad (c_i \in Y_i), \tag{2.17}$$

és a *komponensenkénti lbvf*:

$$\underline{F}_{LBVF}(Y, i) = \frac{u_i \underline{V} - \ell_i \underline{W}}{u_i - \ell_i} + (\overline{Y_i} - \underline{Y}_i)\frac{\ell_i u_i}{u_i - \ell_i} \tag{2.18}$$

együttes használatával. Ez a következő eredményre vezet.

10. Tétel. [68] *Legyen* $\underline{F}_K(Y,c,i) = \min\{y_r(c,i), y_t(c,i)\}$, *ahol* $c \in Y$, *valamint*

$$\begin{aligned} y_r(c,i) &= \frac{u_i\underline{V} - \ell_i\underline{Z} + u_i\ell_i(c_i - \underline{Y}_i)}{u_i - \ell_i}, \\ y_t(c,i) &= \frac{u_i\underline{Z} - \ell_i\underline{W} + u_i\ell_i(\overline{Y_i} - c_i)}{u_i - \ell_i}, \end{aligned}$$

ahol $Z \supseteq g_i(c_i)$, $V \supseteq g_i(\underline{Y}_i)$ *és* $W \supseteq g_i(\overline{Y_i})$, *és* $i = 1, \ldots, n$. *Akkor*

$$\max\{\underline{F}_{LBVF}(Y,i), \underline{F}_{CF}(Y,c,i)\} \leq \underline{F}_K(Y,c,i) \leq \underline{f}(Y) \tag{2.19}$$

teljesül minden $i = 1, \ldots n$*-re.*

Bizonyítás. Alkalmazzuk a 5. Tétel bizonyítását minden $i = 1, \ldots, n$-re. □

Az 10. Tétel tehát azt mondja ki, hogy a komponensenkénti kite módszer nem roszszabb, mint a komponensenkénti középponti formula vagy mint a komponensenkénti lbvf (az Y intervallum ugyanazon Y_i irányára nézve).

Csakúgy, mint az egydimenziós esetben, a c paraméter a (2.19) egyenlőtlenségben itt is választható optimálisan. Keressük tehát azt a c^* pontot, amelyre

$$\underline{F}_K(Y,c^*,i) = \max_{c \in Y} \underline{F}_K(Y,c,i) = \max_{c \in Y} \min\{y_R(c,i), y_T(c,i)\}. \tag{2.20}$$

Az optimális c^* meghatározásához minden koordináta irányra használhatjuk az 6. Tételt a 2.4.1. fejezetből.

11. Tétel. [68] *Minden* $i = 1, \ldots, n$*-re a következők teljesülnek.*

1. *Létezik egyértelmű* $c^* \in Y$ *pont, amelyre* $y_R(c^*,i) = y_T(c^*,i)$ *teljesül, és*
2. c^* *a maximumhelye a* $\underline{F}_K(Y,c,i)$ *függvénynek.*

Bizonyítás. Alkalmazzuk a 6. Tétel bizonyítását minden $i = 1, \ldots, n$-re. □

Jegyezzük meg, hogy itt c^* értéke egyaránt függ az Y intervallumtól és az i iránytól, azaz ha $i \neq j$, akkor a $c^*(Y,i)$ értéke általában nem egyezik meg a $c^*(Y,j)$ értékével. A 2.4.1. fejezetből tudjuk, hogy a $c^*(Y,i)$ pont nem feltétlenül egyértelmű, továbbá, hogy egy $\alpha_i g_i(z) + \beta_i = 0 \quad (\alpha_i, \beta_i, z \in \mathbb{R})$ alakú nemlineáris egyenlet megoldásaként határozható meg. A számítógépes megvalósításban általában nem számoljuk ki $c^*(Y,i)$ pontos értékét, annak csak egy közelítését használjuk, hasonló technikával, mint az egydimenziós esetben (lásd a 2.4.4. alfejezet).

A (2.19) egyenlőtlenségből, vagy méginkább a (2.20) egyenletből $f(X)$ garantált alsó becslése adható: $\max_{1\le i\le n} \underline{F}_K(Y,c,i)$ értéke mindig kisebb vagy egyenlő $\underline{f}(Y)$-nél. Egyszerű azonban látni, hogy a középponti formula, vagy a többdimenziós lbvf általában jobb (nagyobb) alsó korlátot ad a célfüggvényünk értékkészletére. Továbbá az $f(Y)$ értékének befoglalása a komponensenkénti kite módszerrel $3n$ függvényhívást (minden irányra 2 kiértékelés a végpontokban és 1 kiértékelés a középpontban) és egy gradiens kiértékelést igényel. Ezért az itt bemutatott módszer használata önmagában nem javasolt globális optimalizáló módszerekben. Ez az oka annak, amiért a módszert inkább egy gyorsító technika kidolgozására és megvalósítására használjuk. Ez a következő szakasz tartalma.

2.5.2. Komponensenkénti metszés magasabb dimenzióban

A komponensenkénti kite módszer használatához kiszámolt értékek segítségével egy metszési (pruning) technikát dolgozhatunk ki. A következő tétel az ehhez szükséges formulákat ismerteti.

12. Tétel. [68] *Legyen $Y \subseteq X \subseteq \mathbb{I}^n$ az aktuálisan vizsgált részintervallum, $c \in Y$, $F'(Y)$ az $f(y)$ gradiensének egy befoglalása és $\tilde{f}$ pedig az aktuális (garantált) felső korlát a globális minimum értékére. Legyen $\mathcal{Y}^*$ az f függvény Y intervallumba eső X intervallumra vonatkozó globális minimumhelyeinek halmaza. Ha $Z \supseteq g_i(c_i)$, $V \supseteq g_i(\underline{Y}_i)$ és $W \supseteq g_i(\overline{Y_i})$,*

$$p_i = \underline{Y}_i + \frac{\tilde{f} - \underline{V}}{\ell_i}, \qquad q_i = c_i + \frac{\tilde{f} - \underline{Z}}{u_i},$$
$$r_i = c_i + \frac{\tilde{f} - \underline{Z}}{\ell_i}, \qquad s_i = \overline{Y_i} + \frac{\tilde{f} - \underline{W}}{u_i},$$

akkor minden $i \in \{1,\dots,n\}$ indexre a következő állítások teljesülnek.

(a) Ha $\tilde{f} < \min\{\underline{V}, \underline{W}, \underline{Z}\}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [p_i, q_i] \times Y_{i+1} \times \ldots \times Y_n \cup Y_1 \times \ldots \times Y_{i-1} \times [r_i, s_i] \times Y_{i+1} \times \ldots \times Y_n$.

(b) Ha $\underline{W} \le \tilde{f} < \min\{\underline{V}, \underline{Z}\}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [p_i, q_i] \times Y_{i+1} \times \ldots \times Y_n \cup Y_1 \times \ldots \times Y_{i-1} \times [r_i, \overline{Y_i}] \times Y_{i+1} \times \ldots \times Y_n$.

(c) Ha $\underline{V} \le \tilde{f} < \min\{\underline{Z}, \underline{W}\}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [\underline{Y}_i, q_i] \times Y_{i+1} \times \ldots \times Y_n \cup Y_1 \times \ldots \times Y_{i-1} \times [r_i, s_i] \times Y_{i+1} \times \ldots \times Y_n$.

(d) Ha $\underline{Z} \le \tilde{f} < \min\{\underline{V}, \underline{W}\}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [p_i, s_i] \times Y_{i+1} \times \ldots \times Y_n$.

(e) Ha $\max\{\underline{W}, \underline{Z}\} \le \tilde{f} < \underline{V}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [p_i, \overline{Y_i}] \times Y_{i+1} \times \ldots \times Y_n$.

(f) Ha $\max\{\underline{V}, \underline{Z}\} \le \tilde{f} < \underline{W}$, akkor $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [\underline{Y}_i, s_i] \times Y_{i+1} \times \ldots \times Y_n$.

(g) *Ha* $\max\{\underline{V}, \underline{W}\} \leq \tilde{f} < \underline{Z}$, *akkor* $\mathcal{Y}^* \subseteq Y_1 \times \ldots \times Y_{i-1} \times [\underline{Y}_i, q_i] \times Y_{i+1} \times \ldots \times Y_n \cup Y_1 \times \ldots \times Y_{i-1} \times [r_i, \overline{Y_i}] \times Y_{i+1} \times \ldots \times Y_n$.

Bizonyítás. Az (a) esetet bizonyítjuk, a (b)–(g) esetek bizonyítása teljesen hasonlóan megy. Legyen $x^* \in Y \subseteq X$ egy globális minimum és legyen $i \in \{1, \ldots, n\}$ tetszőleges, de rögzített index. Először megmutatjuk, hogy $p_i \leq x_i^*$ teljesül. Mivel feltételeztük, hogy $\tilde{f} < \underline{V}$ és $\underline{V} \leq f(x^*)$, ezért $x_i^* \neq \underline{Y}_i$. Az

$$\ell_i \leq \frac{f(x^*) - \underline{V}}{x_i^* - \underline{Y}_i} \leq \frac{\tilde{f} - \underline{V}}{x_i^* - \underline{Y}_i}$$

egyenlőtlenségekből $(x_i^* - \underline{Y}_i)\ell_i \leq \tilde{f} - \underline{V}$ adódik. Ezért

$$x^* \geq \frac{\tilde{f} - \underline{V}}{\ell_i} + \underline{Y}_i = p_i$$

teljesül, hiszen feltettük, hogy $\ell_i < 0$.

Ahhoz, hogy megmutassuk, x_i^* nincs benne a (q_i, r_i) nyílt intervallumban, először tegyük fel, hogy $x_i^* < c_i$. Akkor felhasználva, hogy

$$u_i \geq \frac{f(x^*) - \underline{Z}}{x_i^* - c_i}$$

következik $u_i(x_i^* - c_i) \leq f(x^*) - \underline{z} \leq \tilde{f} - \underline{Z}$. Ebből az egyenlőtlenségből

$$x_i^* \leq c_i + \frac{\tilde{f} - \underline{Z}}{u_i} = q_i$$

adódik. Most tegyük fel, hogy $x_i^* > c_i$. Akkor

$$\ell_i \leq \frac{f(x^*) - \underline{Z}}{x_i^* - c_i} \leq \frac{\tilde{f} - \underline{Z}}{x_i^* - c_i}$$

implikálja az $(x_i^* - c_i)\ell_i \leq \tilde{f} - \underline{Z}$ egyenlőtlenséget. Ekkor pedig

$$x_i^* \geq c_i + \frac{\tilde{f} - \underline{Z}}{\ell_i}$$

áll, mivel feltettük, hogy $x_i^* - c_i > 0$ és $\ell_i < 0$. Az $x_i^* = c_i$ eset lehetetlen, mert feltettük, hogy $\underline{Z} = \underline{g_i}(c_i) > \tilde{f}$ és $\tilde{f} \geq f(x^*)$.

Végül az

$$u_i \geq \frac{f(x^*) - \underline{W}}{x_i^* - \overline{Y_i}} \geq \frac{\tilde{f} - \underline{W}}{x_i^* - \overline{Y_i}}$$

egyenlőtlenségből $(x_i^* - \overline{Y_i})u_i \leq \tilde{f} - \underline{W}$ következik. Ekkor

$$x_i^* \leq \overline{Y_i} + \frac{\tilde{f} - \underline{W}}{u_i}$$

is teljesül, mivel $u_i > 0$ egy korábbi feltételből. Ezzel befejeztük az (a) eset bizonyítását. □

3. Következmény. [68] *Minden esetben, amikor* $Y_i = [\underline{Y_i}, \overline{Y_i}]$ *és* $((p_i > \underline{Y_i}) \wedge (s_i < \underline{Y_i}))$ *vagy* $((q_i < \underline{Y_i}) \wedge (r_i > \overline{Y_i}))$ *teljesül, akkor a teljes* Y *részintervallum kidobható: nem tartalmazhat globális minimimpontot.*

Mielőtt rátérnénk a fentiek alapján a javasolt algoritmus ismertetésére, mutatunk egy példát, ami segít megérteni a fenti gondolatmenetet.

Példa. Tekintsük az $f(x_1, x_2) = x_1^2 + x_2^2 + x_1$ függvényt az $X = X_1 \times X_2 = [-1, 0.5] \times [-0.5, 1]$ tartományon. Akkor $g_1(c_1, X_2) = c_1^2 + X_2^2 + c_1$ és $g_2(X_1, c_2) = X_1^2 + c_2 + X_1$, ahol $c_1 \in X_1$ és $c_2 \in X_2$. Automatikus deriválással (vagy „kézzel" számolva) kapjuk a derivált befoglalását, ami $F' = [-1, 2] \times [-1, 2]$.

A következő értékek az intervallum aritmetika használatával kaphatók:

$$\begin{aligned} &g_1(c_1, X_2) = [-0.1875, -0.1875], && g_2(X_1, c_2) = [-0.9375, 1.5625], \\ &g_1(\underline{X_1}, X_2) = [0, 1], && g_1(\overline{X_1}, X_2) = [0.75, 1.75], \\ &g_2(\overline{X_1}, \underline{X_2}) = [-0.75, 1.75], && g_2(X_1, \overline{X_2}) = [0, 2.5], \end{aligned}$$

ahol $(c_1, c_2) = \mathrm{mid}\,(X)$. Így előállíthatjuk az $f(x_1, x_2)$ komponensenkénti befoglalásait. Először a komponensenkénti középponti formulával azt kapjuk, hogy

$$\underline{F}_{CF}(X, c, 1) = -0.9375,$$
$$\underline{F}_{CF}(X, c, 2) = -2.4375,$$

míg az lbvf az

$$\underline{F}_{LBVF}(X, 1) = -0.75,$$
$$\underline{F}_{LBVF}(X, 2) = -1.5,$$

értéket adja; végül a kite befoglalással az

$$\underline{F}_K(X, \tilde{c}, 1) = -0.5390625,$$
$$\underline{F}_K(X, \tilde{c}, 2) = -1.2890625$$

alsó korlátokat kapjuk. Itt a $\tilde{c} \in \mathbb{R}^2$ pont az optimális kite kifejtési pontjának egy közelítése.

A fenti befoglalásokkal kapott alsó korlát tehát az $\underline{F}(X) = \max\{\underline{F}_{CF}(X, c, 1), \underline{F}_{CF}(X, c, 2), \underline{F}_{LBVF}(X, 1), \underline{F}_{LBVF}(X, 2), \underline{F}_K(X, \tilde{c}, 1), \underline{F}_K(X, \tilde{c}, 2)\} = -0.5390625$ érték.

2.5.3. A javasolt algoritmus

Most a fenti eredmények alapján egy új B&B alapú globális optimalizálási eljárás algoritmikus leírását adjuk.

A lépés. Legyen X a kiindulási intervallum. Számítsuk ki az $F(X)$ és $F'(X)$ értékeket. Végezzük el az $\mathcal{L} = \{(X, \underline{F}(X), F'(X))\}$, $\mathcal{Q} = \{\}$, és $\tilde{f} = F(c)$ (garantált felső korlát a globális minimim értékére) inicializáló műveleteket.

B lépés. Mindaddig, amíg $\mathcal{L}$ nem üres, hajtsuk végre az alábbi lépéseket.

C lépés. Vegyük le az $(Y, \underline{F}(Y), F'(Y))$ hármast az $\mathcal{L}$ listáról, majd Y minden koordináta irányára csináljuk a következőket.

C.1 lépés. Számítsuk ki a komponensenkénti kite befoglalást az i-edik koordinátára.

C.2 lépés. Alkalmazzuk a metszés módszert az i-edik koordinátára.

D lépés. A metszés által keletkezett U_i $(i = 1 \ldots m \leq n+1)$ részintervallum(ok)ra hajtsuk végre a következőket.

D.1 lépés. Számítsuk ki az $F(U_i)$ és $F'(U_i)$ értékeket. Alkalmazzuk a monotonitási- és a középponti tesztet.

D.2 lépés. Számítsuk ki a középponti formulát (és frissítsük $\tilde{f}$ értékét, ha lehetséges).

D.3 lépés. Ha a megállási feltétel teljesül az aktuális intervallumra, akkor tegyük fel a $\mathcal{Q}$ listára, különben tegyük rá (az $\underline{F}(U_i), F'(U_i)$ értékekkel együtt) az $\mathcal{L}$ listára.

E lépés. Menjünk vissza a B lépésre.

Mindenekelőtt hangsúlyozzuk, hogy ez az algoritmus a komponensenkénti kite módszert mint metszési lépést használja (azaz gyorsítóként) és nem (csak) mint befoglaló függvényt. Hogy (általában) jobb befoglalást kapjunk a célfüggvényre, az aktuális intervallumon mindig használjuk a középponti formulát (lásd D.2 lépés). Erre azért van szükség, mert a középponti formula általában jobb alsó korlátot ad, mint a komponensenkénti kite módszer. Mindazonáltal a C lépésben a kite kiszámításához szükséges információ felhasználható az $\tilde{f}$ értékének csökkentésére. Másrészt a vizsgált részintervallum elvethető, ha az $\tilde{f} < \underline{F}_K(Y, c, i)$ egyenlőtlenség (azaz egy értékkészlet teszt) teljesül.

A C.2 lépésben használhatjuk a Ratz [53] által bevezetett speciális vágási technikát. Ez a következő séma szerint működik:

1. Legyenek $V, W \subseteq Y_i$ a metszés lépés által produkált részintervallumok.
2. Ha $W = V = \emptyset$, akkor megállunk (nincs megoldás Y-ban).
3. Ha $V \neq \emptyset$, akkor legyen $Y_i := V$ és tároljuk el Y-t.
4. Legyen $Y_i := W$ és folytassuk a következő i-vel.

Ezt a módszert alkalmazva a metszés eljárás legfeljebb $n+1$ részintervallumot produkál (ahogyan ezt jeleztük az algoritmus D lépésének leírásában). Ha egy iterációs lépésben a metszés eredményes volt (tehát tudtunk csökkenteni az aktuális intervallum méretén), akkor a derivált befoglalását nem számítjuk ki a következő iterációban (amely tehát az előző lépésben lecsökkentett méretű intervallummal dolgozik). Megjegyezzük továbbá, hogy a C lépés egy egyszerű kettévágást hajt végre amennyiben a metszés sikertelen volt.

További észrevétel, hogy Y_i kiszámítása tetszőleges indexezés szerint történhet – tehát nem szükséges rögzített $i = 1, \ldots, n$ sorrend. Használhatunk egy rendezett index vektort, amely a komponensek egy meghatározott sorrendjét tartalmazza. Vizsgálatainkban a intervallumos felosztási eljárásokból ismert A, B, C és D sorbarendezési technikákat alkalmaztuk (részletes leírást lásd Csendes & Ratz [13]). Numerikus eredményeink szerint a C szabály tűnik a legkedvezőbnek. Ez a

$$D(i) = \operatorname{wid}\left(F_i'(Y)(Y_i - \operatorname{mid}(Y_i))\right), \tag{2.21}$$

érdem-függvény maximalizálásán alapszik. Az új $t = (t_1, \ldots, t_n)$ index vektor, ahol $t_k \in \{1, \ldots, n\}$ és $t_i \neq t_j$ ha $i \neq j$ kielégíti a $D(t_k) \geq D(t_{k+1})$ egyenlőtlenséget minden $k = 1, \ldots, n-1$ indexre.

Figyelembe véve a fenti meggondolásokat, a következő állítással zárjuk elméleti vizsgálódásainkat.

4. Következmény. *A javasolt algoritmust használva soha nem veszíthetünk el a kiindulási X intervallumban lévő globális minimum pontokat. Továbbá a metszési lépésben ha egy adott Y intervallumra az m értéke 0, akkor f-nek nincs (az X intervallumra véve) globális minimumhelye Y-ban.*

2.5.4. Numerikus eredmények

Ez a szakasz a fentiekben ismertetett komponensenkénti kite befoglalás és a hozzá kidolgozott metszés eljárás intervallumos globális optimalizálási algoritmusba történt implementálásával és tesztelésével kapott numerikus eredmények diszkussziót tartalmazza.

A tesztelés célja, hogy kimutassuk az új gyorsító eljárás hatékonyságát (a hagyományos algoritmussal szemben), megvizsgáljuk a viselkedését. Az implementációt

egy 1 GHz-es Pentium III gépen, Linux operációs rendszer alatt a C++ Toolbox for Verified Computing [26] környezetben végeztük. Az összehasonlításban hagyományos algoritmusnak a 2.5.3. alfejezetben ismertetett algoritmust használtuk a következő módosításokkal:

- a C lépést nem hajtottuk végre,
- a D lépésben az m értékét mindig 2-re állítottuk (tehát biszekciót alkalmaztunk).

A vizsgálatokban az irodalomban jól ismert és gyakran használt 40 darab standard tesztfüggvényt használtuk. Megállási feltételként az aktuális intervallum relatív szélességének maximális nagyságaként 10^{-6} értéket követeltünk meg (kivéve a GP, Sch27, Sch214, G7, R5, R6, R7, R8 és EX2 feladatokra, ahol ez az érték 10^{-2} volt.)

Numerikus eredményeink azt mutatták, hogy a (2.20) formulát használva a komponensenkénti kite befoglalás kiszámításakor kapott $g_i(\underline{Y}_i), g_i(c_i), g_i(\overline{Y_i})$ intervallumok nagyon szélesek lehetnek. Ilyenkor a nagymértékű túlbecslés miatt a metszés lépés nem használható sikeresen. Ezért amennyiben a $g_i(\underline{Y}_i), g_i(c_i), g_i(\overline{Y_i})$ intervallumok valamelyike szélesebb, mint egy meghatározott heurisztikus paraméter, akkor az algoritmus kihagyja a metszés lépést (azaz a C lépést) és egy intervallum felezést hajt végre. Megvalósításunkban a $\max\{D(t_1), 100\}$ paramétert használtuk erre a célra, ahol a D érdem-függvényt a (2.21) képletben definiáltuk. Ezt a módosítást alkalmazva a számítási költségek csökkenthetők.

Mindkét algoritmus sikeresen megoldotta az összes tesztfeladatot. A numerikus eredményeket a 2.4. és a 2.5. táblázatok tartalmazzák. A megadott hatékonysági mutatók:

- függvénykiértékelések száma,
- deriváltkiértékelések száma,
- maximális listaméret,
- és a feladat megoldására igénybe vett CPU idő.

Az utolsó előtti sorban Σ jelzi a megadott hatékonysági mutatók összegzett értékét. Az utolsó sor megfelelő oszlopai az átlagok átlagát (AoP) tartalmazzák.

Összefoglalva az eredményeket láthatjuk, hogy a függvénykiértékelések száma 24 tesztfüggvény esetében nagyobb volt az új módszer esetében. Az eredmény nem meglepő: a komponensenkénti kite kiszámításához az adott részintervallum szélein vett függvényértékekre is szükségünk van.

A derivált-kiértékelések száma majdnem minden esetben kevesebb az új módszer esetén. Az algoritmus felépítéséből következik, hogy valójában ez az érték szoros

2.4. táblázat. Numerikus eredmények többdimenziós kite-ot használó algoritmusra.

Feladat		Függvénykiértékelések száma			Deriváltkiértékelések száma		
neve	dim.	régi	új	%	régi	új	%
S5	4	281	450	160	179	177	98
S7	4	291	478	164	183	184	100
S10	4	291	478	164	183	184	100
H3	3	1338	1232	92	889	683	76
H6	6	3654	4705	128	2399	1767	73
GP	2	15991	24043	150	8653	7696	88
SHCB	2	1366	1896	138	859	750	87
THCB	2	874	848	97	563	291	51
BR	2	1278	769	60	831	297	35
RB	2	460	559	121	283	252	89
RB5	5	2582	2775	107	1601	1445	90
L3	2	2522	2481	98	1629	671	41
L5	2	587	933	158	385	285	74
L8	3	237	282	118	153	150	98
L9	4	315	369	117	203	200	98
L10	5	393	453	115	253	251	99
L11	8	627	709	113	403	401	99
L12	10	783	878	112	503	501	99
L13	2	162	229	141	103	96	93
L14	3	243	329	135	153	145	94
L15	4	323	436	134	203	194	95
L16	5	388	514	132	243	238	97
L18	7	542	708	130	339	334	98
Sch21	2	2004	2125	106	1249	868	69
Sch31	3	253	357	141	153	159	103
Sch25	2	649	736	113	415	323	77
Sch27	3	708262	28505	4	472269	15726	3
Sch214	4	15771	11692	74	10317	6399	62
Sch218	2	2022	2393	118	1215	1140	93
Sch32	3	866	863	99	545	411	75
Sch37	5	8830	8766	99	5887	5823	98
Sch37	10	559102	557054	99	372735	370687	99
G5	5	14590	1741	11	9727	705	7
G7	7	43774	11578	26	29183	2855	9
R4	2	2454	1390	56	1615	633	39
R5	3	33386	14727	44	22251	8893	39
R6	5	52558	31543	60	35023	19881	56
R7	7	71730	44337	61	47795	28349	59
R8	9	90902	79971	87	60567	51541	85
EX2	5	425349	690379	162	279673	213027	76
Σ		2068030	1534711	74	1371812	744612	54
AoP				106			76

2.5. táblázat. Numerikus eredmények többdimenziós kite-ot használó algoritmusra.

Feladat		Maximális listahossz			Felhasznált CPU idő		
neve	dim.	régi	új	%	régi	új	%
S5	4	9	9	100	0,36	0,55	152
S7	4	12	12	100	0,50	0,79	158
S10	4	12	12	100	0,71	1,12	157
H3	3	21	13	61	1,07	0,99	92
H6	6	118	79	66	9,01	10,87	120
GP	2	798	761	95	8,32	10,76	129
SHCB	2	60	57	95	0,31	0,38	122
THCB	2	24	17	70	0,15	0,13	86
BR	2	17	10	58	0,26	0,14	53
RB	2	11	12	109	0,07	0,07	100
RB5	5	58	58	100	1,93	1,93	100
L3	2	119	98	82	1,40	1,22	87
L5	2	29	37	127	0,36	0,53	147
L8	3	9	9	100	0,13	0,15	115
L9	4	12	12	100	0,27	0,30	111
L10	5	15	15	100	0,48	0,54	112
L11	8	24	24	100	1,90	2,09	109
L12	10	30	30	100	3,71	4,15	111
L13	2	7	6	85	0,04	0,05	125
L14	3	10	11	110	0,10	0,13	130
L15	4	13	14	107	0,21	0,27	128
L16	5	16	12	75	0,37	0,46	124
L18	7	22	16	72	0,91	1,13	124
Sch21	2	36	31	86	0,47	0,45	95
Sch31	3	3	5	166	0,08	0,11	137
Sch25	2	8	8	100	0,10	0,10	100
Sch27	3	45364	1901	4	9597,84	53,44	0
Sch214	4	382	355	92	6,92	4,47	64
Sch218	2	18	18	100	0,26	0,28	107
Sch32	3	13	10	76	0,21	0,20	95
Sch37	5	32	32	100	5,84	5,67	97
Sch37	10	1024	1024	100	2070,49	2026,59	97
G5	5	32	32	100	11,28	1,25	11
G7	7	128	128	100	64,03	13,45	21
R4	2	72	32	44	0,41	0,20	48
R5	3	1024	512	50	28,48	10,16	35
R6	5	1024	768	75	91,02	48,48	53
R7	7	1024	768	75	203,42	112,16	55
R8	9	1024	896	87	409,26	328,78	80
EX2	5	13236	12007	90	2428,82	2556,78	105
Σ		65890	19851	30	14951,50	5270,99	35
AoP				89			97

összefüggésben van a végrehajtáshoz szükséges iterációs lépések számával. Ebből arra következtethetünk, hogy az új módszer egy másik utat jár be a B&B fában a feladatok megoldása során.

A felhasznált tárméret az új módszer esetén kisebb volt, kevesebb részintervallumot helyezett el a még szóbajöhető intervallumok listájára.

Az új algoritmus által felhasznált teljes CPU idő 35%-a volt a hagyományos eljárás lefutásához szükséges időnek. Ha azonban kiszámítjuk az egyes feladatokra kapott százalékok átlagát, mindössze 3%-os növekedést kapunk. Ebből a két mutatóból azt a konklúziót vonhatjuk le, hogy az új módszer jobban működik a nehezebben megoldható feladatokon.

További megállapításunk, hogy az új módszer rosszabbul működik a Shekel függvényekre (S5, S7, S10). Másrészről a Ratz függvényekre (R4 – R8) sokkal jobban teljesít. A legnagyobb teljesítmény növekedést a Schwefel-27 (Sch27) és a Griewank (G5, G7) feladatokon értük el.

Összefoglalva a numerikus eredményeket megállapíthatjuk, hogy a metszési technikát alkalmazó algoritmus a fenti tesztfeladatsoron bizonyítottan jobb eredményt produkált. A teljesítmény növekedés ráadásul a nehezebb feladatok esetén volt nagyobb.

Összehasonlítás más rendszerekkel

Ugyanúgy, mint az egydimenziós esetre, a magasabb dimenzióra is létezik alternatíva, például a MIAG rendszer (Martínez *et al.* [49]) és az AMIGO (Martínez *et al.* [50]). Ezek lényegében a 2.3.2. szakaszban említett ötlet (lásd Casado *et al.* [7]) többdimenziós komponensenkénti kiterjesztése néhány egyéb szofisztikált gyorsító technikával. Ebben az alfejezetben empírikus összevetést végzünk a kite módszer, a MIAG és az AMIGO között.

Az összehasonlításnál a megállási feltételként a $\operatorname{wid}(X) < \varepsilon$ teljesülését vizsgáljuk (míg az előző alfejezetben a vizsgált részintervallumok relatív szélességét vizsgáltuk a megállási feltételben). Ennek az a magyarázata, hogy az összehasonlítás alapjául vett cikkekben ez volt az alkalmazott megállási feltétel.

Mint azt a már közölt numerikus eredményekből láttuk, a könnyebb feladatokon a kite módszer nem hozott teljesítmény javítást, sőt, ennek épp az ellenkezőjét tapasztaltuk. Ugyanez a helyzet a MIAG és az AMIGO esetében is. Ezért a numerikus tesztjeinkben csak azokat a tesztfeladatokat vettük figyelembe, amelyek mindhárom módszer számára nehezebben megoldhatónak bizonyultak, valamint mindhárom módszernél rendelkezésre állnak a futtatási adatok.

A futtatási eredményeket a 2.6. táblázat tartalmazza. Mivel a [49] és [50] cikkekben teljesítmény mutatóként a függvénykiértékelések száma + n(deriváltkiértékelések

2.6. táblázat. A többdimenziós kite algoritmus összehasonlítása a MIAG és az AMIGO módszerekkel.

Probléma	ε értéke	n-kite	AMIGO	%	MIAG	%
Schw12	$1e-8$	42602	22341	191	22963	186
GP	$1e-8$	51405	30493	169	30128	171
H6	$1e-8$	13959	12998	107	13020	107
HM4	$1e-8$	28211	28726	98	59870	47
Sch214	$1e-5$	168711	139335	121	595993	28
R5	$1e-3$	92937	364215	26	331049	28
R6	$1e-3$	308491	502237	61	468513	66
Schw210	$1e-2$	380544	520749	73	496155	77
G10	$1e-2$	1770934	2436103	73	3869704	46
RB10	$1e-2$	1163248	1524310	76	2045727	57
EX2	$1e-2$	849652	261241	325	256975	331
R8	$1e-2$	137432	75231	183	75231	183
Σ		5008126	5917979	85	8265328	65
AoP				125		110

száma) van feltüntetve, ezért a kite módszerre is ezt a mutatót tüntettük fel az egyes tesztfeladatoknál.

A táblázatból láthatjuk, hogy meglehetősen vegyes képet kapunk az egyes módszerek hatékonyságáról. Megállapíthatjuk, hogy a kite módszer összességében jobban teljesített, ezt mutatják a Σ sorban szereplő mutatók. A hatékonyság javulás itt 15%, illetve 35% lett rendre az AMIGO-hoz és a MIAG-hoz viszonyítva. Ha azonban az átlagok átlagát számoljuk, akkor rosszabb eredményt kapunk. Így az AMIGO 25%-kal, míg a MIAG 10%-kal volt gyorsabb a kite módszernél.

Megfigyelhetjük, hogy a kite módszer az EX2 feladaton teljesített a legrosszabbul, ez okozza az átlagos hatékonyságának csökkenését. Viszont az R5 tesztfüggvényre a másik két eljárásnál sokkal gyorsabban szolgáltatott eredményt.

Amennyiben a dimenziószámot vesszük figyelembe, akkor általában a kite módszer hatékonyabb volt (ez legjobban a G10 tesztfüggvényre igaz). Az R8 feladat esetében viszont lassabb volt, mint a másik két eljárás.

Módszerünk teljesítményén körültekintőbb implementáció javíthat (például a rész-intervallum-széleken már kiszámított függvényértékek eltárolása, stb).

3. fejezet

Egy módszertan globális optimalizáló programok összehasonlítására

Ebben a fejezetben egy olyan algoritmikus eljárást ismertetünk, amely a teljes globális optimalizáló programok tesztelésére és összehasonlítására szolgál[1]. A keretrendszer összeállítása a COCONUT projekt [8] egyik vállalt célja volt, azon belül valósult meg.

A módszer fontosságát azzal tudjuk alátámasztani, hogy jelenleg kb. egy tucat teljes globális optimalizáló szoftver létezik (kommerciális és public domain), amelyek szerzőik szerint gyorsan, helyesen és megbízhatóan oldják meg az optimalizálási feladatokat. Ezen tulajdonságok megléte csak akkor nyer értelmet, ha van viszonyítási alapunk. Az eredmény jelentőségét mutatja továbbá, hogy ez volt az első eset, amikor különböző korlátozásos globális optimalizálási és feltétel kielégítési feladatokat megoldó programok összehasonlítása megvalósult egyrészt szisztematikus alapokon, valamint olyan teszthalmazon, amely megengedi statisztikusan szignifikáns következtetések levonását.

Eredményeinket a Neumaier *et al.* [48] cikk közli. Míg a cikk főleg a konkrét tesztelési eredményeket tartalmazza, jelen értekezésben a munka alapját képező módszertant is részletesen ismertetjük.

Az itt ismertetett módszertan tehát arra vállalkozik, hogy algoritmikus úton olyan keretet adjon, amely számítógépen implementálható, és lényegében emberi beavatkozás nélkül elvégezzen egy olyan lépés sorozatot, amelynek a végén emberi feldolgozásra alkalmas és értelmes kimutatásokat kapjunk a tesztelt programok gyorsaságára, helyességére és megbízhatóságára vonatkozóan.

[1] Mindazonáltal a nem teljes optimalizálók tesztelése és összevetése is lehetséges a környezeten belül.

Az irodalomban számos olyan eredményt találunk, amelyek lokális (lásd például BARR *et al.* [1], CROWDER *el al.* [10] és DOLAN & MORÉ [15] cikkeket) vagy nemteljes (például JANKA [29] és MONGEAU *el al.* [43]) optimalizálók tesztelését vették célba (további hivatkozások és eredmények tekintetében MITTELMANN [42] összefoglaló web oldala ad eligazítást). Teljes optimalizálók részletes tesztelésére azonban csak a NEUMAIER *el al.* [48] cikkben, illetve a COCONUT projekt keretében találunk eredményeket.

3.1. Előkészületek

Előkészületként tesztfeladatokat kell gyűjtenünk, időzítéssel kell foglalkoznunk, egységesíteni kell az inputot, le kell rögzítenünk, hogy milyen teljesítmény-kritériumokat követelünk, valamint rendelkeznünk kell egy listával, ami az egyes tesztfeladatok legjobb megoldásait tartalmazza. A következőkben ezeket tárgyaljuk részletesen.

3.1.1. Tesztfeladatok

A tesztelés első lépése, hogy tesztfeladatokkal rendelkezzünk. A COCONUT projekt keretében összeállított tesztfeladat gyűjtemény összesen 1322 optimalizálási feladatból áll (ez a COCONUT Benchmarking Set). Az itt ismertetett módszertan megvalósításában ezen tesztfeladatsor egy részén futtattuk a vizsgált megoldókat (kihagytuk a legnagyobb méretű feladatokat, ahol a változók száma nagyobb volt, mint 1000).

A feladatokat 3 fő könyvtárra osztottuk, ezek a származásukra és jellegükre utaló osztályozások:

- Library1 = Global Library (GAMS World, [21])
- Library2 = CUTE (globális- és lokális feladatok, [24])
- Library3 = EPFL (feltétel kielégítési feladatok, [59])

Az egyes könyvtárakon belül méret szerint (a feladatokban előforduló változók száma) csoportosítottuk a feladatokat: size1 ($n \leq 10$), size2 ($10 < n \leq 100$), size3 ($100 < n \leq 1000$).

Megjegyzés. Egy korai verzióban a tiny, small és large elnevezéseket használtuk, ami azért lehet félrevezető, mert az optimalizálás világán belül is más-más értelmezést kap például a „small” elnevezés a különböző felhasználási területeket tekintve.

Példa. Jelölésüket a `lib2s1` = Library2 size1 példával illusztráljuk.

3.1. táblázat. A tesztelés során felhasznált számítógépek adatai.

Szgép	CPU típus	OS	CPU/MHz	BogoMips	STU/sec	Linpack
Lisa	AMD Athlon XP2000+	Linux	1678.86	3348.88	50	7.42
Hektor	AMD Athlon XP1800+	Linux	1544.51	3080.19	53	6.66
Zenon	AMD Family 6 Model 4	Windows NT 4.0	1001	—	74	46.78
Theseus	Pentium III	Linux	1000.07	1992.29	130	4.12
Bagend	AMD Athlon MP2000+	Linux	1666.72	3329.22	36	5.68

3.1.2. Időzítés

Fontos szempont az időzítés kérdése. A probléma abból adódik, hogy a különböző megoldók futási idejét szeretnénk összehasonlítani és ez alapján (is) rangsorolni őket. Egy nagyméretű tesztfeladatsoron végrehajtott komplett tesztelést nem feltétlenül egyetlen számítógépen végezzük, hanem több, esetleg különböző kapacitású és sebességű gépen. Arra, hogy hogyan mérhetjük a felhasznált időt, számos javaslat született, például:

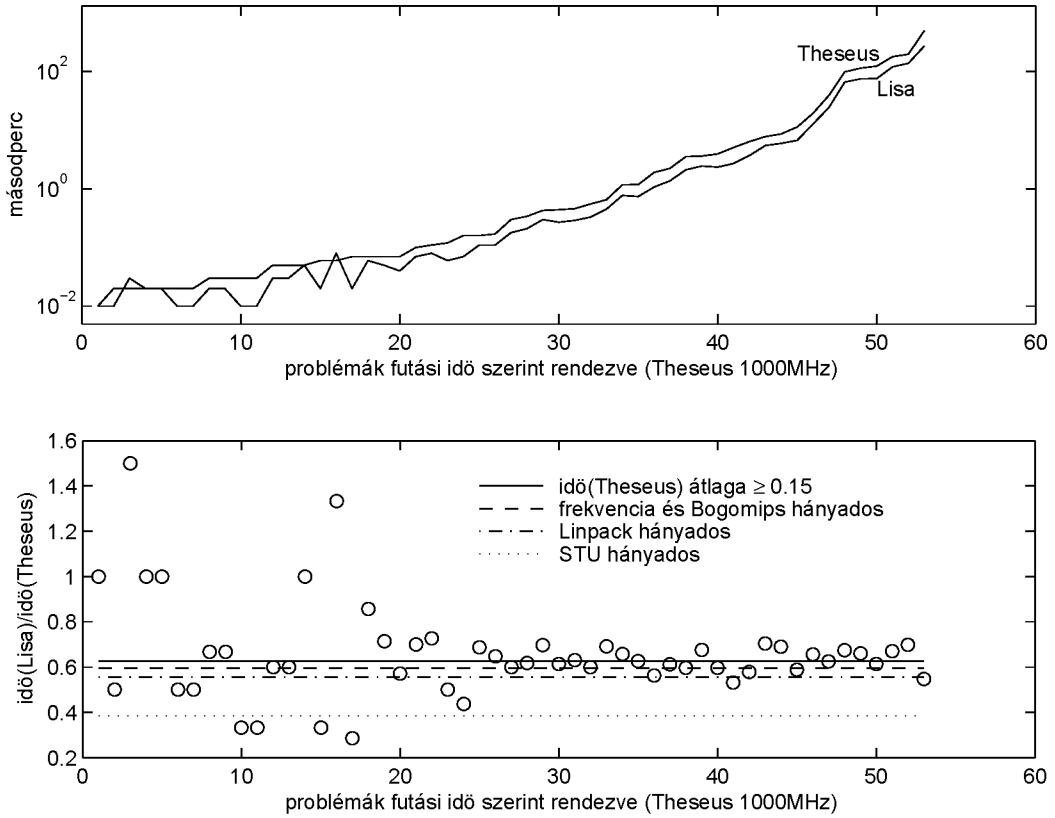
- a processzor órajel frekvenciája (MHz),
- a standard időegység (Standard Time Unit, lásd Dixon & Szegő [14] és Shcherbina *el al.* [59]),
- a Linpack [35] csomag által javasolt Java alapú időmérő lefuttatása,
- vagy a BogoMips [5], amely a Linux operációs rendszereken a CPU teljesítményét meghatározó mértékegység.

3.2. táblázat. A leggyorsabb és a leglassabb gépek egymáshoz viszonyított teljesítményeinek összehasonlítása.

	Lisa	Theseus	hányados	inverz hányados
CPU frekvencia	1678.86	1000.07	1.68	0.60
Bogomips	3348.88	1992.29	1.68	0.59
STU	50.00	130.00	0.38	2.60
Linpack	7.42	4.12	1.80	0.56

A 3.1. táblázat a tesztelés során felhasznált számítógépeken lefuttatott sebességmérők eredményeit mutatja. A 3.1.2. táblázat pedig a Lisa és a Theseus gépek teljesítmény mutatóinak hányadosait tartalmazza. Ez alapján a CPU frekvencia és a BogoMips mérőszámok tűnnek a legjobb választásnak.

Annak eldöntésére, hogy a CPU frekvencia (amelynek mutatószáma minden gépre könnyen megmondható) valóban megbízható mutató, lefuttattuk a BARON optimalizáló programot a `lib1s1` könyvtárra a Theseus és a Lisa gépeken. A kapott eredményt a 3.1. ábra mutatja. Az ábrából kiderül, hogy nagyon rövid futásidő esetén az összehasonlítás nehéz, ezért a t időeredményeket másodpercben egy tizedesjegyre adtuk meg, ha $t < 10$ és a legközelebbi egészre kerekítve, ha $t \geq 10$. A nagyon pici időket (ahol a kerekítés miatt 0 jött ki a felhasznált időre) egységesen 0.05-re állítottuk.



3.1. ábra. A BARON futási ideje a `lib1s1` könyvtárra.

3.1.3. Egységes input

Következő fontos szempont, hogy a tesztfeladatok olyan formátumban legyen elérhetők, amely implicit vagy explicit módón feldolgozhatók a tesztelésben szereplő

programok által. Ennek a kérdésnek a megoldására az AMPL (Fourer *et al.* [17]) nevű, matematikai programozási feladatok leírására alkalmas modellezési nyelvet választhatjuk, mint kiindulási formátumot. Mint azt látni fogjuk később, ehhez mellékelünk majd olyan konvertereket, amelyek az AMPL formátumból előállítják a megfelelő input formátumot. A COCONUT Benchmarking Set tehát AMPL formátumban tartalmazza a tesztfeladatokat.

Amennyiben a tesztfeladat eredetileg maximalizálási feladat volt, akkor a célfüggvényt beszoroztuk -1-gyel.

Az AMPL modellezési nyelv mellett egy másik, alapjaiban véve teljesen különböző filozófián alapuló input formátum is részét képezi módszertanunknak, a DAG (directed acyclic graph, irányított körmentes gráf). A formátum részletes leírását és az optimalizálás szempontjából fontos és hasznos tulajdonságok tárgyalását a Schichl & Neumaier [58] cikk tartalmazza. Számunkra jelen pillanatban azért fontos a DAG formátum, mert egy közbülső formátumot képez az AMPL és más input formátumok között. Nevezetesen, mint azt említettük, a COCONUT környezet számos olyan konvertert biztosít, amely az AMPL nyelven leírt optimalizálási feladatokat átírja valamilyen más nyelvre.

Konverterek

A COCONUT környezetben jelenleg elérhetők konverterek listáját a 3.3. táblázat tartalmazza.

3.3. táblázat. A COCONUT környezetben elérhető input konverterek listája és funkciója.

név	funkció
`ampl2dag`	AMPL formátumból DAG formátumra
`dag_simplify`	DAG formátumot egyszerűsíti
`dag2gams`	DAG formátumból GAMS formátumra
`dag2lgo`	DAG formátumból Windows LGO formátumra
`dag2c`	DAG formátumból C nyelvre
`dag2globsol`	DAG formátumról GlobSol input formátumra
`c2dag`	speciális C++ formátumról DAG formátumra

Láthatjuk, hogy ezek a konverterek lehetővé teszik, hogy a széles körben használt input formátumok mind elérhetők legyenek az AMPL teszthalmazból kiindulva.

A konverterek helyessége

Az egyes feladatok korrekt megoldásához és az egyes megoldók megbízható összehasonlításához fontos biztosítanunk, hogy a konverterek működése helyes legyen.

Ennek egy lehetséges tesztelése a következőképpen mehet:

1. Az AMPL tesztfeladatokból csináljunk más formátumokat.
2. Ezeket oldjuk meg a különböző optimalizáló programokkal.
3. Vessük össze a kapott eredményeket.

Jegyezzük meg azonban, hogy a 3. lépésben a kapott eredmények esetleges eltérését okozhatják a megoldó programokban előforduló hibák is. Fontos továbbá, hogy ez a módszer csak szükséges feltételt biztosít a konverterek helyességére.

A következőkben leírjuk, hogy hogyan történt egy konkrét tesztelés a fenti konverterek helyességének ellenőrzésére.

Első lépés: ellenőrzés GAMS rendszerrel. Először a `lib1s1` feladatkönyvtárra alkalmaztuk a következő konvertálás-sorozatot:

$$\text{GAMS} \to \text{AMPL} \to \text{DAG} \to \text{GAMS},$$

valamint a `lib2s1` probléma könyvtárra az

$$\text{AMPL} \to \text{DAG} \to \text{GAMS} \to \text{AMPL}$$

konverzió-sorozatot. Így tehát adott ezen könyvtáraknak két-két változata, mindegyik GAMS [19] formátumban. Ekkor futtatuk a GAMS rendszert a BARON programmal ezeken a könyvtárakon és összehasonlítottuk a kapott eredményeket. Ha valamelyik feladatra a két verzió megoldása különbséget mutatott, akkor megvizsgáltuk a feladat különböző formátumait (ezt kézzel kell elvégezni, de mivel ezek alacsony dimenziós feladatok, így ez nem jelent nagy problémát).

Második lépés: Ellenőrzés különböző megoldó programokkal. A konverterek helyességét úgy is érdemes megvizsgálni, hogy a különböző megoldó programokat lefuttatjuk egy kis feladathalmazra és összevetjük a kapott eredményeket az eltérő megoldásokra koncentrálva. Elég csak

- azokat az eseteket megvizsgálni, ahol valamelyik megoldó program egy eredményről optimalitást állít és az nem tűnik helyesnek;
- illetve azokat az eseteket, ahol furcsa eredményeket kapunk, például ahol a BARON a többi program által kapott eredménytől lényeges különböző megoldást ad.

Érdemes megjegyezni, hogy itt a második lépésben találkozhatunk olyan esettel is (konkrétan találtunk is ilyet), amikor az eredmények különbözőségét nem a konverter hibája okozza, hanem maga a megoldó program. Az ilyen típusú ellenőrzés elvégzése tehát a fejlesztők munkáját is segítik.

3.1.4. Teljesítmény kritériumok

Szinte minden megoldóprogram rendelkezik konfigurációs lehetőségekkel. Amennyiben adott egy globális optimalizálási feladat, amelynek ismerjük valamilyen tulajdonságait, akkor esetleg végezhetünk olyan beállításokat a keresőprogramban, amelyek gyorsabb feladat megoldáshoz vezetnek. Más azonban a helyzet akkor, ha a megoldók tesztelésére, összehasonlítására van szó. Ezért minden megoldóra a gyártó által javasolt alapértelmezett beállításokat használtuk. Ez alapvetően peszszimista eredményekhez vezet(het), de egy ilyen méretű tesztsorozaton ez az egyedüli járható út.

Időkorlátok

Amennyiben a tesztelést különböző teljesítményű gépeken végezzük, akkor az időkorlátokat normalizálni kell úgy, hogy az eredmények végül összevethetők legyenek. Az egyes méret szerinti osztályozásra különböző (lényegében tetszés szerinti) időkorlátokat kell rögzítenünk. Ezen egységek alapján egy konkrét gépre a

$$\frac{\text{korlát} \times 1000}{\text{CPU MHz}}$$

képlet alapján állítottuk be az időkorlátot.

Példa. A teszteléskor a használt időkorlátok: 180, 900, 1800 másodperc rendre a size1, size2 és size3 mértekre. Ez alapján egy 1666MHz-es számítógépen a fenti képlet 108 másodpercet ad a size1 feladatokra.

Kategóriák a kimenet osztályozására

A jelenleg elérhető globális optimalizáló programok egyik leggyengébb része a kapott megoldások osztályozásának megbízhatósága. A 3.1.4. táblázat tartalmazza azokat a jelöléseket, amelyeket a programok kimenetének egységes osztályozására javasoltunk. Ez alapján még könnyebb az egyes megoldók sikeresség szempontjából történő összehasonlítása.

A „feloldatlan” osztályozás tartalmazhat olyan eseteket is, amikor a megoldó fízibilis, de nem optimális megoldást talált, és a futást befejezte még azelőtt, mielőtt meg tudta volna vizsgálni, hogy lokális vagy globális optimumot talált-e.

3.4. táblázat. Kategóriák a kimenetek osztályozására.

Jel	Jelentés
X	a feladatot nem fogadta el a megoldó
I	a feladatot infízibilisnek nyilvánította a megoldó
G	a megoldást globálisnak nyilvánította a megoldó
L	a megoldást lokálisnak (esetleg globális) nyilvánította a megoldó
U	feloldatlan (nem talált megoldást vagy hibaüzenet)
T	időkorlát túlhaladva

3.1.5. Legjobb függvényértékek előállítása, vizsgálata

A következő fontos kérdés, hogy honnan tudjuk megállapítani, hogy egy megoldó program a globális megoldást találta-e meg, illetve, hogy a 3.1.4. táblázat alapján kiadott sikerességi mutató helyes-e? Világos, hogy ehhez először össze kell állítani egy listát, amely minden egyes vizsgált feladatra tartalmazza a globális optimum értékét (illetve, ha ilyet nem találunk, akkor a lehető legjobb megoldást). Ehhez a következő lépéseket alkalmaztuk.

1. Először minden megoldó kimenetét egységesítettük (ezek lesznek a `.res` fájlok, lásd később).
2. Az egységes kimeneteket fízibilitási tesztnek vetettük alá. Ez a COCONUT környezetben található `solcheck` programmal történt. Egy pontot fízibilisnek tekintünk, ha kielégít minden $c(x) \in [\underline{c}, \overline{c}]$ feltételt egy rögzített `tol` abszolút hiba mellett azokra a korlátokra, amelyek felső korlátjainak abszolút értéke kisebb, mint 1, és `tol` nagyságú relatív hiba mellett a többi korlátra. Az egyenlőség feltételek vizsgálatát hasonlóan végeztük el a $\underline{c} = \overline{c}$ használatával.

Végül a fízibilis megoldások közül (amiből egy-egy feladat esetén több is lehet) kiválasztottuk a legkisebb függvényértékűt.

Legjobb megoldások listája

Ezek után jöhet a legjobb megoldások listájának (hitlist) összeállítása. Ez úgy történik, hogy minden könyvtárra vesszük az összes futtatási értéket és ezek közül minden feladatra kiválasztjuk a legkisebb függvényértékkel rendelkező megoldást, illetve ha több ilyen is van, akkor azok közül azt, amelyiknek a maximális fízibilitása a legkisebb; ez lesz a globális optimum.

Ha nincs fízibilis megoldás, akkor a korlátok túllépésében a lehető legkisebb eltérésű megoldást választottuk, de megjelölve azt infízibilisként.

A hitlist tehát a következő oszlopokat tartalmazza:

- feladat neve,
- feladat mérete (változók száma és korlátozó feltételek száma),
- az optimális pontot tartalmazó `.res` fájl elérhetősége
- maximális fízibilitás,
- és a globális minimum értéke.

Az eredményül kapott legjobb megoldások listája elérhető a COCONUT Benchmark [9] internetes oldalról.

3.2. Jelölések a táblázatokban

A tesztkörnyezet számos táblázatba rendezett kimutatást készít a megoldó programok minőségi viselkedéséről. Ezeket a táblázatokat ismertetjük ebben a szakaszban.

3.2.1. Összefoglaló statisztikák

Az összefoglaló statisztikákat tartalmazó táblázatokban használt jelöléseket a 3.5. táblázat tartalmazza.

3.5. táblázat. Összefoglaló táblázatokban használt jelölések.

Oszlop	Jelentés
library	könyvtár leírása
all	könyvtár/méret
accepted	a kereső által elfogadott feladatok száma
+G	feladatok száma, ahol megtalálta a globális optimumot
G!	feladatok száma, ahol a globális optimumot a globalitás állításával együtt helyesen megtalálta
G?	feladatok száma, ahol a globalitás állítása megvolt, de az igazi globális megoldás valójában jobb, vagy a globális megoldásnak kinevezett pont valójában infízibilis
I?	feladatok száma, ahol a feladat infízibilisnek lett mondva, bár fízibilis megoldás is létezik.

3.6. táblázat. Példa összefoglaló táblázatra a BARON eredményeivel.

BARON7.2/GAMS summary statistics						
library	all	accepted	+G	G!	G?	I?
`lib1s1`	91	88	88	64	0	0
`lib1s2`	80	77	71	46	3	0
`lib1s3`	41	33	23	5	1	0
`lib2s1`	324	296	254	206	11	0
`lib2s2`	99	89	82	48	2	0
`lib2s3`	95	87	51	25	6	0
`lib3s1`	217	195	182	180	3	3
`lib3s2`	69	63	57	57	2	1
`lib3s3`	22	20	14	13	1	0

Példa. A 3.2.1. táblázatban a BARON megoldóra kapott eredményeinkkel demonstráljuk az összefoglaló statisztikai táblát. Láthatjuk, hogy a kimutatásokat az egyes feladat osztályokra külön sorokban kapjuk, ami nagyban megkönnyíti az eredmények értékelését.

Megjegyzés. A teszteléshez elkészített környezet, amely majdnem teljes egészében automaziálja az eredmények feldolgozását, angol nyelvű táblázatokat ad végeredményül. A 3.2.1. táblázat erre egy példa, ahol meghagytuk az angol szövegeket.

3.2.2. Feladatok osztályozása nehézség szerint

Ha egy megoldó talált egy globális minimumot (anélkül, hogy tudná a globalitást), akkor a globalitás ellenőrzése abból áll, hogy megállapítsa: vajon tényleg nincs a talált pontnál jobb. Ez a legidőigényesebb része egy teljes keresésnek. Másrészt látjuk, hogy a globális minimum megtalálása anélkül, hogy tudnánk a globalitását, lényegében lokális minimumkeresés. Ezért két osztályba soroljuk a feladatokat:

"könnyen lokalizálható feladat" ahol a lokális megoldó program (esetünkben a MINOS) talált globális optimumhoz tartozó fízibilis pontot;

"nehezen lokalizálható feladat" minden más eset, ahol a lokális kereső (MINOS) sikertelen volt.

3.2.3. Részletező táblázatban használt jelölések

Az előző alfejezetben ismertetett könnyen/nehezen lokalizálható feladat fogalmát felhasználva részletező táblázatokat készítünk az egyes problémakönyvtárakról. Ezek-

3.7. táblázat. A részletező táblázatban használatos jelölések.

Oszlop	leírás
wrong	rossz állítások száma, azaz a G? és I? esetek összege az összefoglaló statisztikai táblázatból
+G	hányszor volt a megoldás valóban globális
−G	hányszor volt a megoldás valójában nem globális
I	hány feladat volt valójában infízibilis

ben a táblázatokban használt jelöléseket mutatja a 3.2.2. táblázat.

Példa. A 3.2.3. táblázat mutat egy példát a részletező táblázatra, ez a BARON eredményeit tartalmazza a `lib1s1` problémakönyvtárra. A táblázatból kiolvasható, hogy 91 feladatot tartalmazott a `lib1s1` probléma könyvtár. Ebből a BARON számára 64 volt könnyen, 27 pedig nehezen lokalizálható. Összesen 64 esetben állította (helyesen) a globalitást, 15 esetben, hogy lokális megoldást talált, ezek azonban valójában globális megoldások voltak (8 könnyen, 7 pedig nehezen lokalizálható). A rendelkezésre álló idő 9 esetben letelt, mire végzett volna a teljes keresésse (az LT sor), ezért lokális megoldásként adta meg ezeket az eredményeket; a táblázatból viszont látható, hogy ezek valójában globális optimumok voltak. Végül 3 esetben nem tudta elfogadni a feladatot (ezek trigonometrikus függvényeket tartalmaztak, amelyeket a BARON nem tud kezelni).

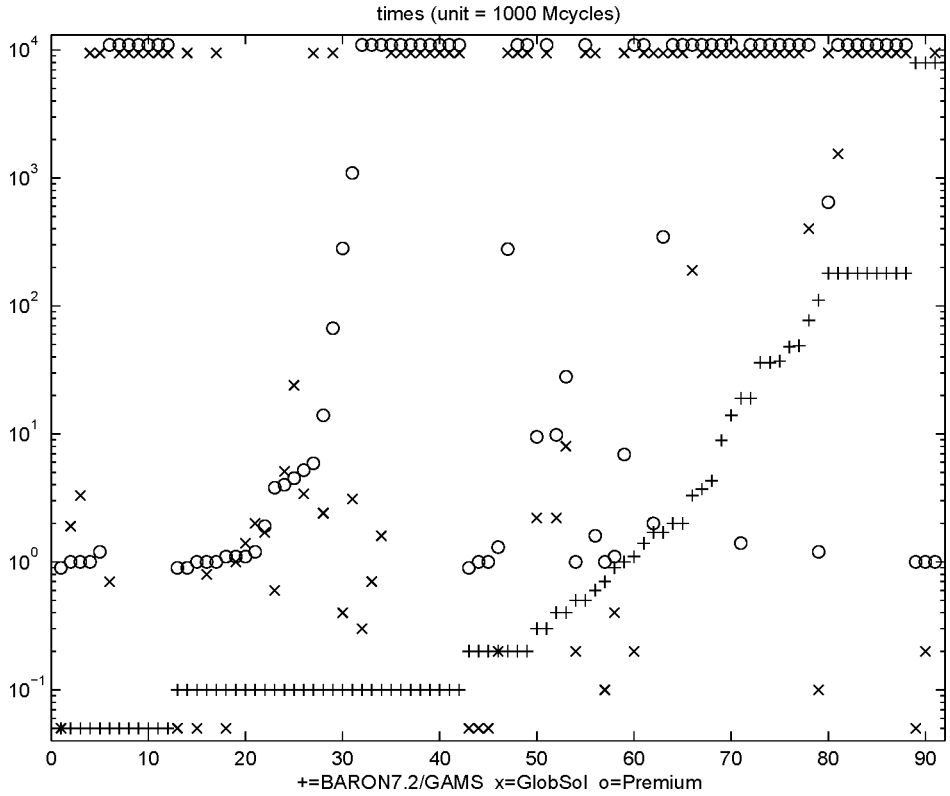
3.8. táblázat. Példa a részletező táblázatra a BARON eredményeivel.

BARON7.2/GAMS on `lib1s1`								
status	all	wrong	easy location			hard location		
			+G	−G	I	+G	−G	I
all	91	0	62	2	0	26	1	0
G	64	0	50	0	0	14	0	0
L	15	0	8	0	0	7	0	0
LT	9	0	4	0	0	5	0	0
X	3	0	0	2	0	0	1	0

3.2.4. Futási idők összehasonlítása

Amennyiben az egyes megoldókat a feladatok megoldására felhasznált futási idejük alapján szeretnénk összevetni, akkor egy erre alkalmas ábra sokat segíthet az adatok értelmezésében. A 3.2. ábrán látunk erre példát: a BARON 7.2, a GlobSol [23] és a Premium Solver [18] optimalizálók eredményeit hasonlítjuk össze a `lib1s1`

feladatsoron. Itt a BARON által adott időeredmények alapján vannak sorbarakva a felhasznált idők. Fontos azonban megjegyeznünk, hogy az esetleges konverziókra (például a `dag2globsol` futtatására) felhasznált idők itt nincsennek feltüntetve. A 0.05 időegység alatti idők az ábra legalján vannak. Azokhoz a feladatokhoz, amelyekre a megoldó nem találta meg a globális megoldást egy fiktív időegységet rendelünk, ami a rögzített időkorláton felül van – ezek az esetek vannak az ábra tetején.



3.2. ábra. Példa a futási idők összehasonlítására: BARON, GlobSol és Premium Solver optimalizálók a `lib1s1` feladatsoron.

3.2.5. Megbízhatósági analízis

Az eddigi statisztikákból készíthetünk egy megbízhatósági analízist, amely a következőket tartalmazza:

- az elfogadott feladatok közül az esetek hány százalékában találta meg a globális minimumot (itt nem kell, hogy a megoldó állítsa is a globalitást),
- az elfogadott feladatok közül az esetek hány százalékában volt helyes a globalitás megállapítása,

- az esetek hány százalékában állította helytelenül a globalitást,
- végül az elfogadott és fízibilis feladatok közül az esetek hány százalékában történt meg az, hogy a megoldó infízibilitást állított.

Mindezekből látható, hogy egy megoldó program akkor ígéretes, ha az első két kategóriában minél nagyobb százalékos teljesítményt hoz, míg az utolsó két kategóriában lehetőleg minél kisebbet.

Példa. Megbízhatósági analízis példát a BARON-ra mutatunk a 3.9. táblázatban. Láthatjuk, hogy a BARON azon feladatokon, amelyeken lehetett futtatni, az esetek 86%-ában találta meg a globális optimumot. A legjobban a size2 méretű feladatokon teljesített (92%), míg a legnagyobb méretű feladatok valóban nehéznek bizonyultak, itt csak 62% volt a sikeresség a globális optimum megtalálásában. Az összes elfogadott feladatot tekintve az esetek 62%-ában állította helyesen, hogy globális optimumot talált. Nagyon kevés esetben állította helytelenül a globalitást (mindössze 4%), míg az infízibilitás helytelen állítása is igen csekély számban fordult elő.

3.9. táblázat. Példa a BARON megbízhatósági analízisére.

Reliability analysis for BARON 7.2	
	global minimum found/accepted
size 1	$524/579 \approx 91\%$
size 2	$210/229 \approx 92\%$
size 3	$88/140 \approx 63\%$
all	$821/950 \approx 86\%$
	correctly claimed global/accepted
size 1	$450/579 \approx 78\%$
size 2	$151/229 \approx 66\%$
size 3	$43/140 \approx 31\%$
all	$644/950 \approx 68\%$
	wrongly claimed global/claimed global
size 1	$14/464 \approx 3\%$
size 2	$7/158 \approx 4\%$
size 3	$8/51 \approx 16\%$
all	$29/675 \approx 4\%$
	claimed infeasible/accepted and feasible
size 1	$3/571 \approx 1\%$
size 2	$1/222 \approx 0\%$
size 3	$0/128 = 0\%$
all	$4/921 \approx 0.4\%$

3.2.6. A teszteredmények összefoglalása

A COCONUT projekt keretén belül a BARON/GAMS (7.2-es verzió) [62], COCOS (2004. szeptember 20-án kiadott béta teszt verzió), GlobSol (2004. szeptember 11-én kiadott verzió) [23], ICOS (2004. március 29-én kiadott béta teszt verzió) [33], LGO/GAMS [51], LINGO 9.0 [28], OQNLP/GAMS [20], Premium Solver 5 [18] és a MINOS/GAMS lokális kereső összehasonlítását végeztük el. Az eredmények rövid összefoglalása a következő.

A tesztelt programok közül a BARON a leggyorsabb és legrobosztusabb. Tőle nem sokkal marad le az OQNLP. Az elérhető megoldók közül egyik sem teljesen megbízható, egyetlen kivétellel: feltétel kielégítési feladatok megoldására szolgáló ICOS, ami bár lassabb, mint a BARON, kiváló megbízhatósági jellemzőkkel rendelkezik (amikor be tudja fejezni a keresést a rendelkezésre álló időkorláton belül). A BARON a 100 változónál kevesebb változót tartalmazó tesztfeladatsorokat 90%-os sikerrel oldotta meg, míg az ennél nagyobb feladatoknak valamivel több, mint kétharmadát.

A sztochasztikus megoldók közül az OQNLP volt a legjobb. Hátránya a BARON-nal szemben, hogy lassabb és nem tud információval szolgálni arról, hogy a keresés teljes volt-e. A 100 változónál nagyobb feladatok 72%-át oldotta meg (a megadott időkorláton belül).

A GlobSol és a Premium Solver esetében csak a `lib1s1` könyvtárra végeztük el a tesztelést. Mivel ezek a programok a szigorúan megbízható kategóriába tartoznak, ezért nem meglepő, hogy például a BARON-hoz képest lényegesen lassabban dolgoznak. Ennek ellenére találtunk olyan eseteket, amikor rosszul határoztak meg megoldásokat, ami implementálási hibákra vall.

Az ICOS, amely szintén a szigorúan megbízható kategóriába tartozik, csak feltétel kielégítési feladatok megoldására képes, ezért csak a Library 3 könyvtárra teszteltük. Az ICOS volt az egyetlen program, amelynél egyszer sem fordult elő, hogy hamisan állította volna a globalitást.

Végül néhány pontban összefoglaljuk a tesztelés során kialakult tapasztalatainkat:

- A GAMS rendszer LGO és az OQNLP megoldói nagyon óvatosak, sohasem állítanak globalitást. Másrészről, ugyancsak a GAMS rendszerben a MINOS néha globalitást állít egy feladat megoldása végén. Ez annak köszönhető, hogy néhány feladat esetén észreveheti, hogy lényegében lineáris feladatról van szó, ahol a lokális megoldás egyben globális is. (A G? eseteket a megoldások pontatlan közelítése okozta.)

- Néhány esetben az optimalizálók infízibilitást észleltek, annak ellenére, hogy a kapott megoldás a `solcheck` szerint fízibilis volt.

- Számos esetben tapasztaltuk (legtöbbször a LINGO esetén), hogy egy minimumpont közelítő értékének megtalálásakor a globalitás állítása hamis volt azért, mert a korlátozó feltételek nem voltak kielégítve az előírt toleranciával.
- A tesztelt programok általában nem vették észre, ha a célfüggvény konstans volt (tehát lényegében feltétel kielégítési feladatot kaptak).

A teszteredmények közzététele a fejlesztők számára is hasznos volt. A BARON és az ICOS szerzői az eredmények ismeretében javítani tudtak a megoldóprogramjaik hatékonyságán és megbízhatóságán.

Az ismertetett módszertan egy lehetséges alternatívája a Dolan & Moré [15] által kidolgozott teljesítmény profil (performance profile) lehet. Ennek alkalmazása és értelmezése azonban (teljes) globális optimalizálókra óvatosságot igényel, mivel ezek a programok ugyan a globális optimumot gyakran hamar megtalálják, viszont jelentős időt töltenek azzal, hogy kiderítsék van-e másik megoldás is.

4. fejezet

Atomklaszter feladatok

A globális optimalizálás számára az egyik nagy kihívást jelentő feladat az atomklaszterek szerkezetének meghatározása. Az ún. computational chemistry tudományterületnek ez csak egy apró része, annak számos egyéb szép, matematikai szempontból is érdekes feladata létezik és megoldásra vár (bővebben lásd NEUMAIER [46]).

Ebben a fejezetben bizonyos tulajdonságoknak elegettevő atomklaszterek optimális szerkezetének vizsgálatával foglalkozunk. A közölt eredményeket a VINKÓ [64] és a VINKÓ & NEUMAIER [67] cikkek tartalmazzák.

4.1. Alapfogalmak

Tekintsünk n darab atomot. Az i-edik atom pozícióját az $x_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ és $d = 2, 3, \ldots$ jelöli, így egy *atomot* tekinthetünk úgy, mint a (d-dimenziós) Euklideszi-tér egy pontja[1].

Az $x = (x_1, \ldots, x_n) \in \mathbb{R}^{dn}$ atomklaszter *energiáját* az atomok közötti interakciók határozzák meg. Matematikailag ezt egy

$$E(x) = \sum_{i<j} v(x_i, x_j) + \sum_{i<j<k} v(x_i, x_j, x_k) + \ldots \tag{4.1}$$

alakú függvény írja le; ez tehát egy $E : \mathbb{R}^{dn} \to \mathbb{R}$ alakú függvény. Az atomok optimális elhelyezkedésének (ez a minimális energiaszint) meghatározása az E függvény globális minimalizálásával ekvivalens.

Mint látható, az E függvény általában számos tagot tartalmaz, ezek a tagok az atomok között különféle kölcsönhatások leírására szolgálnak. A kémiai és fizikai

[1] Általában $d = 3$; itt pusztán arról van szó, hogy a vizsgált modellek tetszőleges dimenzióra definiálhatók.

szimulációknál fontos, hogy megfelelő típusú modellt találjunk a vizsgált rendszer leírására. A (4.1) képlet elméletileg tetszőlegesen bonyolult is lehet, ez azonban rendkívül nehézzé teszi a tényleges modellezést. Konkrét vizsgálatokban a (4.1) jobboldalának általában csak az első két tagját tekintik. Ekkor azonban még mindig kérdés, hogy a szóban forgó függvények milyen feltételeknek tegyenek eleget, illetve, hogy ezek mennyire tükrözik a valóságot. Ezért legtöbbször csak az első tagot, az ú.n. *párpotenciált* vesszük figyelembe. A tapasztalatok azt bizonyítják, hogy ezek is jól közelítik a valóságot, másrészről a modellezés közben még így is temérdek számítást kell elvégezni.

Jelen értekezésben mi is csak olyan energiafüggvény vizsgálatával foglalkozunk, amelyek csak párpotenciál függvényt tartalmaznak, tehát

$$E(x_1, \ldots, x_n) = \sum_{i<j} v(r_{ij}) \tag{4.2}$$

alakúak, ahol $r_{ij} := \|x_i - x_j\|$. A bemutatott vizsgálataink és módszereink viszont általánosak abban az értelemben, hogy nem rögzítjük le a párpotenciál függvényt, hanem megadunk egy feltételrendszert, amelynek eleget tevő párpotenciált tartalmazó energiafüggvény bizonyos tulajdonságai meghatározhatók.

4.1.1. Vizsgálandó tulajdonságok

A továbbiakban a (4.2) függvény által leírt atomklaszterek optimális szerkezetének tulajdonságait fogjuk megvizsgálni:

(a) Milyen alsó korlátot adhatunk az atompárok közötti minimális távolságra az előforduló atomok számától függetlenül?

(b) Amennyiben figyelembe vesszük az atomok számát milyen alsó korlátot adhatunk meg az atompárok közötti minimális távolságra?

(c) Milyen (lehetőleg lineáris) alsó- és felső korlátot adhatunk a célfüggvény optimális értékére?

4.1.2. Eredmények használhatósága

Fontos kérdés, hogy az imént felsorolt tulajdonságok ismeretében hogyan javíthatunk az atomklaszter feladatok megoldására kidolgozott globális optimalizálási módszerek hatékonyságán.

Az atompárok közötti minimális- és maximális távolság ismerete

- a B&B módszerben alkalmazható gyorsító eljárásként;

- felhasználható nemteljes vagy aszimptotikusan teljes keresők esetében a kiindulási pont előállítására. Erre a LOCATELLI & SCHOEN [36] cikkben találunk példát, ahol éppen méretszám független minimális távolságot használták fel a hatékonyság növelése céljából;

- illetve, mint azt XUE [71] bizonyította, ilyen jellegű információ birtokában hatékony adatstruktúrát konstruálhatunk a potenciálfüggvény értékének kiszámítására. Meglepő eredmény, hogy a (4.2) potenciálfüggvény értéke $\mathcal{O}(n)$ időben számítható (míg a naív eljárás $\mathcal{O}(n^2)$ időigényű).

Az optimális szerkezethez tartozó globális minimumra adott alsó- és felső korlátok felhasználhatók a B&B módszerben mint kivágási értékek.

4.1.3. Korábbi eredmények

Atomklaszterek tulajdonságainak elméleti vizsgálatával számos fizikai tárgyú cikkben és könyvben találkozhatunk. Ezek az eredmények javarészt a 70-es évekből származnak és jellemző rájuk, hogy bár bonyolult matematikai apparátust alkalmaznak, explicit, (optimalizálási módszerekhez) jól használható eredményeket mégsem tartalmaznak. Ezekben a cikkekben találkozhatunk először a *stabilitás* fogalmával: egy potenciált stabilisnak nevezünk, ha az optimális konfigurációjára létezik az atomok számában lineáris alsó korlát (lásd RUELLE [56] ide vontakozó alapkönyvében, illetve jelen tézis 4.1.1. fejezetének (c) pontja). Fontos jellemzője még ezen eredményeknek, hogy a párpotenciált általában nem rögzítik le, hanem néhány általános tulajdonságnak eleget tevő potenciálfüggvényt vizsgálnak, és arra állapítanak meg tulajdonságokat.

Számítástudományi szempontból az első eredményeket XUE *et al.* [72] közli. Megmutatták, hogy a Lennard-Jones potenciál (lásd később) lineárisan korlátos (ez egyebként Ruelle munkásságából már ismert volt, bár Xue és munkatársai feltehetőleg nem ismerték ezeket az eredményeket), valamint az optimális konfigurációra vonatkozólag explicit alsó korlátot adtak az atomszámtól független minimális atompár távolságra. MARANAS & FLOUDAS [39] az optimális Lennard-Jones klaszterben előforduló minimális atompár távolságra ad méretfüggő alsó korlátot – ezek az értékek kisméretű konfigurációra nagyon jó eredményt adnak, nagyobb méretek esetén viszont használhatatlanok. XUE [70] megadta az első, gyakorlati szempontból is releváns méretfüggetlen minimális távolságot. Egy nemrégiben megjelent cikkben BLANC [3] javított ezen a korláton. A Morse klaszterekre (lásd 4.5. fejezet) LOCATELLI & SCHOEN [37] ad méretszám független minimális atompár távolságot. Ez az eredmény azért érdekes, mert a Morse klaszterben szereplő párpotenciál megengedi azt az esetet is, hogy két (vagy több) atom a tér ugyanazon pontjában helyezkedjék el – és ezen tulajdonságukban különböznek a Lennard-Jones párpotenciáltól, ahol ez nem megengedett.

A következőkben a VINKÓ [64] és VINKÓ & NEUMAIER [67] cikkekben elért saját eredményeinket közöljük. A konkrét példák esetén megmutatjuk, hogy az általunk javasolt módszerek használatával az imént hivatkozott korábbi eredményekhez képest milyen javításokat érhetünk el.

4.1.4. Jelölések

A fejezet további részében a következő jelöléseket használjuk. Az E függvény globális minimumhelye az az $x^* \in \mathbb{R}^{3n}$ konfiguráció, amelyre

$$E(x^*) = \min_{x \in \mathbb{R}^{3n}} E(x). \tag{4.3}$$

A globális miniumumot

$$E^* = E(x^*)$$

jelöli. Legyen r_{ij} az x_i^* és x_j^* $(i, j = 1, \ldots, n)$ pontok közötti Euklidészi távolság. Az *i címkéjű atomhoz tartozó potenciális energiát* a

$$E_i(x) = \sum_{i \neq j} v(\|x_i - x_j\|) \quad (i = 1, \ldots, n)$$

egyenlet szerint definiáljuk, valamint $E_i^* = E_i(x^*)$. Nyilvánvaló, hogy

$$E(x) = \frac{1}{2} \sum_{i=1}^{n} E_i(x). \tag{4.4}$$

Az optimális struktúrában

$$r_{\min} := \min_{i,j} r_{ij} \quad (i, j = 1, \ldots, n)$$

a *minimális atompár távolság.* A minimális távolság egy alsó korlátját q-val fogjuk jelölni; célunk tehát hogy találjunk lehetőleg minél jobb

$$q \leq r_{\min}$$

alsó becslést.

Amennyiben a v párpotenciál függvénynek létezik pozitív zérushelye, azt t-vel jelöljük.

Az általánosság elvesztése nélkül feltesszük, hogy $x_1 = 0$ és $0 = r_1 < r_2 \leq \ldots \leq r_n$, ahol

$$r_j = \|x_j - x_1\| = \|x_j\| \quad (j = 1, \ldots, n)$$

szerint definiált. A továbbiakban (hacsak külön nem hangsúlyozzuk) csak az $n > 2$ esetet vizsgáljuk.

4.1.5. Feltételek a párpotenciál függvényre

A párpotenciál függvényre az elméleti eredményekben a következő feltételrendszer teljesülését feltételezzük.

(P1) v folytonos.

(P2) Egyértelműen létezik egy nemnegatív s úgy, hogy $v(s) < 0$ és ez az egyetlen globális minimumpontja v-nek.

(P3) $v(r) \to 0$ $(r \to \infty)$.

(P4) $v(r)$ monoton csökkenő ha $r < s$ és monoton növekedő, ha $r \geq s$.

A (P1)–(P4) feltételrendszer meglehetősen általános, az általában használt párpotenciál függvények csak bizonyos megszorításokkal teljesítik ezt. Amennyiben ilyen megszorításra van szükségünk, azt mindig jelezni fogjuk az adott elméleti eredmény tárgyalásakor (látni fogjuk, hogy a méretfüggetlen eredmények ismertetésekor lesz ilyenre szükségünk).

4.2. Méretfüggő korlátok

Ebben az alfejezetben a v függvényről feltesszük, hogy teljesíti a (P1)–(P4) tulajdonságokat.

Az első lemma a Maranas & Floudas [39] által a Lennard-Jones klaszterekre (lásd 4.4. alfejezet) talált alsó- és felső korlátok általánosítását adja.

1. Lemma. [67] *Az optimális atomklaszterre érvényesek a*

$$-\frac{n(n-1)}{2}|v(s)| \leq E^*(n) \leq -d(n-d+1)|v(s)| \tag{4.5}$$

korlátok.

Bizonyítás. Mivel $v(r_{ij}) - v(s) \geq 0$ teljesül, ezért az alsó korlát bizonyítása:

$$\begin{aligned} E^*(n) &= \sum_{i<j} (v(r_{ij}) - v(s) + v(s)) \\ &= \sum_{i<j} (v(r_{ij}) - v(s)) + \sum_{i<j} v(s) \\ &\geq -\frac{n(n-1)}{2}|v(s)|. \end{aligned}$$

Ha tekintünk egy olyan n darab atomból álló klasztert, amelyben $n-d$ atom olyan pozícióban van, hogy mindegyikük d darab másikat "érint"; kezdve d darab atommal úgy, hogy a köztük lévő távolság pontosan s (azaz egy s hosszú szakasz 2 dimenzióban, egy egyenlő oldalú háromszög 3 dimenzióban, stb.), akkor

$$E^*(n) \leq -d|v(s)| - d(n-d)|v(s)| + M \leq -d(n-d+1)|v(s)|$$

teljesül, – ahol M nem pozitív tag – amely egy felső korlátot ad az optimális szerkezetre. □

Az 1. Lemma egy lineáris felső korlátot ad az optimum értékére. Ez tehát egy válasz a 4.1.1. szakasz (c) pontjának egyik kérdésére. Megjegyezzük továbbá, hogy ez a jelenleg ismert legjobb felső korlát.

2. Lemma. [67] *Az optimális konfigurációban az i atomhoz tartozó potenciál korlátozható a*

$$-(n-1)|v(s)| \leq E_i^*(n) < -e_d|v(s)|, \tag{4.6}$$

értékekkel, ahol $e_d = 1$.

Bizonyítás. A felső korlát bizonyításához legyen $k = n$ ha $i \neq n$ és $k = n-1$ ha $i = n$, és definiáljuk a $z = (z_1, \ldots, z_n)$ konfigurációt úgy, hogy legyen $z_j = x_j^*$ minden $j \neq i$ indexre, valamint legyenek $\|z_i - z_k\| = s$ és $\|z_i - z_l\| \geq s$ minden $l \neq i$ indexre. Akkor helyezzük el a z_i atomot az origó és a z_k atom által meghatározott egyenesen úgy, hogy a z_i rendelkezzen a legnagyobb r_j értékkel. Ekkor $E_i(z) < -|v(s)|$. A z konstrukciójából adódóan

$$E^* - E_i^* \quad = \quad E(z) - E_i(z).$$

Tehát $E_i(z) < -|v(s)|$ és

$$E^* - E_i^* = E(z) - E_i(z) > E(z) + |v(s)|,$$

amiből $E_i^* < -|v(s)|$.

Az alsó korlát a (P4) tulajdonságból és az E_i^* definíciójából jön, nevezetesen E_i^* pontosan $n-1$ tag összege, ahol minden tagnak $v(s)$ alsó korlátja. □

Megjegyzés. Vegyük észre, hogy a (4.6) képletben a felső korlát valójában méret- és dimenziófüggetlen korlát. Létezik továbbá egy sejtés, hogy $e_d = d$ is teljesül, de ennek a bizonyítása egyelőre nyitott. A következőkben a $-e_d|v(s)|$ kifejezést írjuk az E_i^* felső korlátjaként.

3. Lemma. [67] *Ha $n > 2 + e_d$, akkor az optimális konfigurációban a minimális atompár távolságra teljesül a*

$$q(n) := w\Big((n-2-e_d)|v(s)|\Big) \leq r_{\min} \tag{4.7}$$

egyenlőtlenség, ahol w a v inverz függvénye, amely a

$$w(x) = \begin{cases} r & \text{akkor és csak akkor, ha } x = v(r) \text{ és } r \geq s, \\ 0 & \text{különben} \end{cases}$$

definícióval adott.

Bizonyítás. A 2. Lemmát használva a következő levezetést alkalmazhatjuk:

$$\begin{aligned} -e_d|v(s)| &\geq E_1^* = \sum_{j=2}^{n} v(r_j) \\ &= \sum_{j=3}^{n} v(r_j) + v(r_2) \\ &\geq -(n-2)|v(s)| + v(r_2). \end{aligned}$$

Az egyenlőtlenséget átrendezve kapjuk, hogy $v(r_2) \leq (n-2-e_d)|v(s)|$, amely a (4.7) összefüggést eredményezi. □

4.3. Méretfüggetlen korlátok

4. Lemma. [67] *Az optimális konfigurációban a minimális atompár távolság mindig kisebb vagy egyenlő, mint a párpotenciál függvény minimumpontja, azaz $r_{\min} \leq s$ teljesül.*

Bizonyítás. Tegyük fel, hogy az optimális konfigurációban $r_{\min} > s$. Tudjuk, hogy a v függvény növekedő, ha $r \geq s$. Ezért ha alkalmazunk egy skálázást, amely minden távolságot lecsökkent úgy, hogy $r_{\min} = s$ is teljesüljön, akkor a konfiguráció teljes energiáját is csökkentenénk. Ezért $r_{\min} \leq s$. □

4.3.1. Első változat

A következőkben a Vinkó [64] cikk eredményeit közöljük, amely Xue [70] és Blanc [3] eredményeit általánosítja, illetve javítja tovább.

Módszerünk a következő. Feltesszük, hogy a vizsgált konfigurációban a minimális atompár távolság pontosan q. Először egy felső korlátot adunk az E_i^* $(i = 1, \ldots, n)$ értékekre. Tegyük fel, hogy $p \in \mathbb{R}_+$ egy olyan paraméter, hogy

$$pq \geq s \tag{4.8}$$

teljesül. Ekkor használjuk a

$$E_1^* = \sum_{q \leq r_j < pq} v(r_j) + \sum_{r_j \geq pq} v(r_j) \tag{4.9}$$

felosztást és alsó korlátokat adunk erre a két tagra. Megfelelően megválasztott paraméterekkel megmutatjuk, hogy ha a minimális atompár távolság túl kicsi, akkor ellentmondásra jutunk az E_1-re adott felső korláttal.

Feltételek a párpotenciálra

A módszer használhatóságához szükségünk van a (P1)–(P4) feltételrendszer szigorítására. Nevezetesen felteszük, hogy a v függvényre teljesül (P1), (P2), továbbá

(P3') Ha $r \leq s$, akkor v szigorúan monoton csökkenő és $v(r) \geq r^{-4}$.

(P4') Ha $r > s$, akkor v szigorúan monoton növekvő és $v(r) \geq -r^{-4}$.

A (P3') és (P4') tulajdonságok megkövetelése az alkalmazott gömbpakolási módszerrel van összefüggésben. Itt írhatunk Cr^{-3} alakú korlátot is, azonban a C konstans *a priori* meghatározása meglehetősen bonyolult.

Vegyük észre továbbá, hogy a (P1), (P2) és (P3') tulajdonságokból következik, hogy a v függvénynek van $t < s$ zérushelye.

A felhasznált korlátok

5. Lemma. [64] *Ha $\frac{r}{2} < a < b$, akkor az $\mathcal{J}_{ab} = \{j \mid a \leq r_j < b\}$ indexhalmaz méretére érvényes az*

$$|\mathcal{J}_{ab}| \leq \left(\frac{2b}{r} + 1\right)^d - \left(\frac{2a}{r} - 1\right)^d$$

korlát.

Bizonyítás. Feltételezhetjük, hogy a vizsgált konfigurációban szereplő atomok $r/2$ sugarú gömbök. A $\mathcal{J}_{ab}$ halmaz mérete nem haladhatja meg azon $r/2$ sugarú gömbök

számát, amelyeket az origó középpontú $b + r/2$ sugarú gömb tartalmaz. Térfogat összehasonlítással ebből az

$$|\mathcal{J}_{ab}| \leq \left(\frac{b+\frac{r}{2}}{\frac{r}{2}}\right)^d$$

felső korlát adódik. Másrészt, mivel $r_j \geq a$ teljesül, ezért kidobhatjuk az összes olyan $r/2$ sugarú gömböt, amely az origó középpontú $a - r/2$ sugarú gömbben van. Ezen elemek száma szintén térfogat összehasonlítással felülről becsülhető, ahogyan azt a lemma állítja. □

6. Lemma. [64] *Ha $pq \geq s$, akkor a (4.9) első tagja alulról korlátozható a*

$$\sum_{q \leq r_j < pq} v(r_j) \geq v(q) - \left((2p+1)^d - 1)\right) |v(s)| \tag{4.10}$$

egyenlőtlenséggel.

Bizonyítás. Tegyük fel, hogy $r_2 = r_3 = \ldots = r_{m+1} = q$ (azaz létezik $m \geq 1$ darab q-val egyenlő távolság). Mivel ezek pozitív értékeket adnak a potenciál értékében, ezért egy kivételével (amiről feltettük, hogy létezik) mindegyiket elhagyhatjuk. Ekkor

$$\sum_{q \leq r_j < pq} v(r_j) \geq v(q) + \sum_{q < r_j < pq} v(r_j) \tag{4.11}$$

teljesül. Továbbá az 5. Lemma és a párpotenciál monotonitása miatt kapjuk, hogy

$$\begin{aligned} v(q) + \sum_{q < r_j < pq} v(r_j) &\geq v(q) - \left(\left(\frac{2pq+q}{q}\right)^d - \left(\frac{2q-q}{q}\right)^d\right) |v(s)| \\ &= v(q) - \left((2p+1)^d - 1)\right) |v(s)|. \end{aligned}$$

□

7. Lemma. [64] *Legyen $s \leq pq = R_0 < R_1 < R_2 < \ldots$ egy végtelen, szigorúan növekedő sorozat, és definiáljuk az $\mathcal{I}_k = \{j \mid 2 \leq j \leq n, R_k \leq r_j < R_{k+1}\}$ ($k = 0, 1, 2, \ldots$) indexhalmazt. Ha $pq \geq s$, akkor a (4.9) második tagja alulról becsülhető a*

$$\sum_{r_j \geq pq} v(r_j) \geq \frac{1}{q^d} \sum_{k=0}^{\infty} v(R_k) \left((2R_{k+1}+q)^d - (2R_k - q)^d\right) \tag{4.12}$$

egyenlőtlenséggel.

Bizonyítás. Használhatjuk ismét a v függvény (P4') által biztosított monotonitási tulajdonságát és a 5. Lemmát az $\mathcal{I}_k$ indexhalmazra:

$$\begin{aligned}\sum_{r_j \geq pq} v(r_j) &= \sum_{k=0}^{\infty} \sum_{r_j \in \mathcal{I}_k} v(r_j) \\ &\geq \sum_{k=0}^{\infty} \sum_{r_j \in \mathcal{I}_k} v(R_k) \\ &\geq \frac{1}{q^d} \sum_{k=0}^{\infty} v(R_k) \left((2R_{k+1} + q)^d - (2R_k - q)^d \right),\end{aligned}$$

amely a bizonyítást adja. □

Minimális atompár távolság

A fenti lemmákat használva egy általános módszer adható az optimális szerkezetben előforduló minimális atompár távolság egy alsó korlátjának meghatározására. Idézzük fel, hogy t és s a párpotenciál zérus- és minimumhelye. A 7. Lemmában egy végtelen R_k sorozatot használtunk, amely egy végtelen, egymásba ágyazott gömbsorozatot reprezentál. Ehelyett a sorozat helyett azonban használhatunk $R : \mathbb{R}_+ \times \mathbb{N}_0 \to \mathbb{R}_+$ alakú függvényeket is, amelyek az

$$R(Q,k) < R(Q,k+1) \qquad \text{és} \qquad R(Q,0) = c$$

tulajdonságokkal rendelkeznek, ahol $c \in \mathbb{R}_+$ egy konstans (a 7. Lemmában ez a konstans pq, a végtelen sorozat kezdőpontja). A rövidség kedvéért az R_k^Q jelölést fogjuk használni az $R(Q,k)$ függvényre. Használjuk még továbbá a

$$U_c^Q := \{R_k^Q \mid R_k^Q < R_{k+1}^Q \text{ és } R_k^Q = c \text{ és } k = 0, 1, \ldots\}$$

jelölést is. Definiáljuk most a

$$F(q,p) := v(q) - \left((2p+1)^3 - 1\right) |v(s)|, \tag{4.13}$$

$$S(q,p,R) := \frac{1}{q^d} \sum_{k=0}^{\infty} v(R_k^Q) \left(\left(2R_{k+1}^Q + q\right)^d - \left(2R_k^Q - q\right)^d \right), \tag{4.14}$$

$$G(q,p,R) := F(q,p) + S(q,p,R) \tag{4.15}$$

függvényeket. Ezeket a jelöléseket és a 6. és 7. Lemmákat használva az

$$\begin{aligned}E_1^* &= \sum_{q \leq r_j < pq} v(r_j) + \sum_{r_j \geq pq} v(r_j) \\ &\geq G(q,p,R)\end{aligned} \tag{4.16}$$

alsó korlát adódik, ahol $p \in \mathbb{R}_+$ úgy, hogy $pq \geq s$ és $R \in U_{pq}^Q$.

13. Tétel. [64] *Definiáljuk a* $g_v(q,p,Q) := G(q,p,R)$ *függvényt. Ha* $g_v(q,p,Q) > -\infty$, *akkor a* (4.3) *optimális atomklaszter feladatban a minimális atompár távolság kisebb, vagy egyenlő a*

$$\begin{aligned} \frac{\partial g_v(q,p,Q)}{\partial p} &= 0, && (4.17)\\ \frac{\partial g_v(q,p,Q)}{\partial Q} &= 0, && (4.18)\\ g_v(q,p,Q) + e_d|v(s)| &= 0 && (4.19) \end{aligned}$$

nemlináris egyenletrendszer megoldásában szereplő q *értéknél.*

Bizonyítás. A g_v végessége a (P3') és (P4') tulajdonságok megköveteléséből adódik. Ezek a tulajdonságok garantálják azt is, hogy g_v monoton q-ban a $[0,s]$ intervallumon. Ezért (4.19) rendszernek pontosan egy megoldása van.

A 2. Lemmából tudjuk, hogy $E_1^* < -e_d|v(s)|$. Továbbá a $g_v \leq E_1^*$ a (4.16) miatt teljesül. Most keressük azt a legnagyobb q értéket, amelyre a $g_v < -e_d|v(s)|$ alsó becslés nem teljesül. Ehhez vizsgáljuk a

$$\begin{aligned} &\max \quad q \\ &\text{úgy, hogy} \quad g_v(q,p,Q) \geq -e_d|v(s)| \end{aligned} \tag{4.20}$$

optimalizálási feladatot. Ekkor a (4.17) és (4.18) összefüggések a (4.20) optimalizálási feladat elsőrendű optimalitási feltételei p-re és Q-ra nézve. Végül (4.19) garantálja a lehető legnagyobb q értéket, amelyre $g_v < -e_d|v(s)|$ már nem teljesül. Ekkor tehát (4.3) minimális atompár távolsága legalább q. □

A 13. Tétellel elérhető eredményeket tovább javíthatjuk a következő megfontolások alapján. Ha az R_k sorozat első $m > 1$ tagját a $p_1, \ldots, p_m$ változókkal helyettesítjük, akkor egy $m+2$ változós G függvényt kapunk. Nevezetesen a

$$\begin{aligned} G(q,p_1,\ldots,p_m,R) \;:=\; & F(q,p) + \sum_{i=1}^{m-1} v(p_i q)\left((2p_{i+1}+1)^d - (2p_i-1)^d\right) \\ & + \frac{1}{q^d}\sum_{k=0}^{\infty} v(R_k^Q)\left(\left(2R_{k+1}^Q + q\right)^d - \left(2R_k^Q - q\right)^d\right) \end{aligned}$$

függvényt, ahol $F(q,p)$ a (4.13)-ben definiált, $p_1 q \geq s$, és $R_k^Q \in U_{p_m q}^Q$.

5. Következmény. [64] *Definiáljuk a* $g_v(q,p_1,\ldots,p_m,Q) := G(q,p_1,\ldots,p_m,R)$ *függvényt. Ha* $g_v > -\infty$, *akkor a* (4.3) *optimális atomklaszter feladatban a mini-*

mális atompár távolság nagyobb vagy egyenlő, mint a

$$\begin{aligned}
\frac{\partial g_v(q,p_1,\ldots,p_m,Q)}{\partial p_1} &= 0,\\
&\vdots\\
\frac{\partial g_v(q,p_1,\ldots,p_m,Q)}{\partial p_m} &= 0,\\
\frac{\partial g_v(q,p_1,\ldots,p_m,Q)}{\partial Q} &= 0,\\
g_v(q,p_1,\ldots,p_m,Q)+e_d|v(s)| &= 0
\end{aligned}$$

nemlineáris egyenletrendszer megoldásának q komponense.

Lineáris alsó korlát az optimum értékére

Az előző alfejezet eredményeit használva lineáris alsó korlátot adhatunk az optimális függvény értékére is. Ez a korlát helyes lesz tetszőleges méretű klaszter esetén.

14. Tétel. [64] *Ha q egy olyan alsó korlát a minimális atompár távolságra, amelyet az 5. Következmény felhasználásával kaptunk, akkor létezik olyan B_1 konstans, amelyre*

$$-\frac{B_1}{2}n \le E^*.$$

Továbbá B_1 értéke q értékéből meghatározható.

Bizonyítás. Legyen $i \in \{1,\ldots,n\}$ tetszőleges, de rögzített index. Definiáljuk az $M=[t,pq)$ jobbról nyílt intervallumot, ahol $pq \ge s$. Akkor a

$$\sum_{\substack{j=1\\ j\neq i}}^{n} v(r_{ij}) \ge \sum_{\substack{j=1\\ j\neq i, r_{ij}\in M}}^{n} v(r_{ij}) + \sum_{\substack{j=1\\ j\neq i, r_{ij}\ge pq}}^{n} v(r_{ij})$$

alsó becslés adódik. A 5. Lemmát használva az első tag alulról becsülhető:

$$\sum_{\substack{j=1\\ j\neq i, r_{ij}\in M}}^{n} v(r_{ij}) \ \ge\ -\left((2p+1)^d-\left(\frac{2t-q}{q}\right)^d\right)|v(s)|. \tag{4.21}$$

Az 5. és 7. Lemmákból a második tag is becsülhető alulról a

$$\sum_{\substack{j=1\\ j\neq i, r_{ij}\ge pq}}^{n} v(r_{ij}) \ \ge\ \frac{1}{q^d}\sum_{k=0}^{\infty} v(R_k^Q)\left((2R_{k+1}^Q+q)^d-(2R_k^Q-q)^d\right) \tag{4.22}$$

egyenlőtlenséggel, ahol $R_k^Q \in U_{pr^*}^Q$. Ezen megállapításainkat – ugyanúgy, mint az 5. Következménynél – tovább javíthatjuk több változó használatával a (4.22) képletben. Ez a

$$\begin{aligned}\sum_{\substack{j=1\\ j\neq i}}^{n} v(r_{ij}) \quad &\geq \quad -\left((2p+1)^d - \left(\frac{2t-q}{q}\right)^d\right)|v(s)| + \\ &\quad + \sum_{l=1}^{m-1} v(p_l r^*)\left((2p_{l+1}+1)^d - (2p_l-1)^d\right) + \\ &\quad + \frac{1}{q^d}\sum_{k=0}^{\infty} v(R_k^Q)\left(\left(2R_{k+1}^Q+q\right)^d - \left(2R_k^Q-q\right)^d\right) \\ &=: \quad -B_1\end{aligned}$$

korláthoz vezet, ahol $p_1 q \geq s$ és $R_k^Q \in U_{p_m q}^Q$. Amennyiben g_v véges (lásd 5. Következmény), akkor az 5. Következmény által adott megoldásvektor behelyettesítésével a K végessége is garantált. Végül a (4.4) összefüggés a

$$-\frac{B_1}{2}n \leq E^*$$

lineáris alsó korlátot adja az optimális potenciálfüggvény értékére. □

4.3.2. Továbbfejlesztett változat

A továbbfejlesztett változatot az a tény inspirálta, hogy a fenti módszer nem alkalmazható direkt módon olyan párpotenciálokra, amelyek nem divergálnak az atompár távolságának csökkenésével.

A következőkben ismertetett eredményeket a Vinkó & Neumaier [67] cikk tartalmazza.

Feltételek a párpotenciálra

A módszer alkalmazhatóságához felteszük, hogy v teljesíti a (P1) és (P2) tulajdonságokat, valamint a következőt is.

(P3") Létezik olyan $R \in [0, s]$, amelyre

$$\int_s^\infty \left\lfloor \left(\frac{2r}{R}+1\right)^d \right\rfloor v'(r)dr < \min\left\{v(R)+|v(s)|, \frac{1}{2}v(R)+\frac{3}{2}|v(s)|\right\}.$$

Vegyük észre, hogy ez a tulajdonság automatikusan teljesül, ha v divergál az $r \to 0$ esetben.

Felhasznált korlátok

Az alábbiakban R_k jelöli egy rögzített i indexre az x_i atomtól vett k-adik legkisebb távolságot. Akkor $R_1 = 0$ és

$$R_2 = r_{\min} := \min_{i,j} r_{ij} \quad (i, j = 1, \ldots, n) \tag{4.23}$$

a minimális távolság az optimális konfigurációban. Egy bizonyos atomnak majd az 1 címkét adjuk (ennek meghatározását lásd később) és a többi atomot majd úgy jelöljük, hogy $r_i := r_{1i}$ amelyekre

$$0 = r_1 \leq r_2 \leq \ldots \leq r_n.$$

Megjegyzés. Az első módszernél itt szigorú egyenlőtlenséget tételeztünk fel.

Az E_i^* értékekre vonatkozó méretfüggetlen alsó korlát és a teljes energiára érvényes lineáris alsó korlát megadására a továbbiakban a

$$\Sigma_m := \sum_{k=2}^{m} v(r_k)$$

értékekre keresünk alsó- és felső korlátokat.

Legyen $N_d(r)$ azon diszjunkt nyílt egységgömbök maximális száma, amelyek elhelyezhetők egy r sugarú gömbben. Egyszerű térfogat összehasonlítással az

$$N_d(r) \leq \lfloor r^d \rfloor \tag{4.24}$$

felső korlát adódik, amelyet a továbbiakban használunk. Ezen geometrikus pakolási korlát minden további javítása az itt közölt eredmények javítását vonja maga után.

2. Állítás. [67] *Legyen*

$$K(r) := \min_{m \in \mathbb{N},\ R_m > 0} (m-1) N_d\Big(\frac{2r}{R_m} + 1\Big).$$

Akkor

$$k \leq K(r_k) \quad (k = 1, 2, \ldots), \tag{4.25}$$

és K az r növekvő függvénye. Speciálisan

$$K(r) \leq (m-1)\Big\lfloor \Big(\frac{2r}{R_m} + 1\Big)^d \Big\rfloor \quad (m = 2, 3, \ldots). \tag{4.26}$$

Bizonyítás. Legyen $m \geq 2$ tetszőleges, de rögzített. Rekurzívan válasszunk atomokat, kezdve az 1 címkével ellátottal, és a hozzá legközelebb álló $m-2$ atommal. Ez meghatározza atomok egy $\kappa = \lceil k/(m-1) \rceil$ elemszámú halmazát, amelyek legalább R_m távolságra vannak egymástól. Ezért ezen atomok körüli $R_m/2$ sugarú nyílt gömbök diszjunktak és benne vannak abban a nyílt gömbben, amelynek középpontja az 1 címkével ellátott atom, sugara pedig $r_k + R_m/2 = (2r_k + R_m)/2$. Skálázással kapjuk, hogy

$$\kappa \leq N_d\Big(\frac{2r_k}{R_m} + 1\Big),$$

ezért

$$k \leq (m-1)\kappa \leq (m-1)N_d\Big(\frac{2r_k}{R_m} + 1\Big) \leq (m-1)\Big\lfloor \Big(\frac{2r_k}{R_m} + 1\Big)^d \Big\rfloor,$$

amivel az állítást bebizonyítottuk. □

3. Állítás. [67] *Ha $r_m \leq s$, akkor*

$$\Sigma_m \leq -m|v(s)| + E_1^* + \int_s^\infty K(r)v'(r)dr \tag{4.27}$$

és ha $m \geq 2$ is teljesül, akkor

$$(m-1)v(R_m) + (m+e_d)|v(s)| \leq \int_s^\infty K(r)v'(r)dr. \tag{4.28}$$

Bizonyítás. Legyen először m a legnagyobb egész szám, amelyre $r_m \leq s$. Akkor

$$K(r) \geq K(r_m) \geq m \quad \text{ha } r \geq s,$$

és $r_{m+1} > s$, ezért $v(r_{k+1}) - v(r_k) \geq 0$, amennyiben $k \geq m+1$. Ebből, mivel $r_{n+1} = \infty$, $v(\infty) = 0$, azt kapjuk, hogy

$$\begin{aligned}\sum_{k=m+1}^{n} k(v(r_{k+1}) - v(r_k)) &\leq \sum_{k=m+1}^{n} K(r_k)\int_{r_k}^{r_{k+1}} v'(r)dr \\ &\leq \sum_{k=m+1}^{n}\int_{r_k}^{r_{k+1}} K(r)v'(r)dr = \int_{r_{m+1}}^{\infty} K(r)v'(r)dr.\end{aligned}$$

A bal oldal

$$-mv(r_{m+1}) - \sum_{k=m+1}^{n} v(r_k) \geq -mv(r_{m+1}) - E_1^* + \Sigma_m,$$

és mivel $\int_r^\infty v'(r)dr = -v(r)$, ezért azt kapjuk, hogy

$$\begin{aligned}\Sigma_m &\leq E_1^* + \int_{r_{m+1}}^\infty (K(r)-m)v'(r)dr \leq E_1^* + \int_s^\infty (K(r)-m)v'(r)dr \\ &\leq E_1^* + mv(s) + \int_s^\infty K(r)v'(r)dr.\end{aligned}$$

Ez bizonyítja (4.27) teljesülését m maximálisan megengedhető értékére. Mivel

$$\Sigma_m - mv(s) = \sum_{k=2}^m (v(r_k) - v(s)) - v(s)$$

nemnegatív számok összege, és a bal oldal monoton csökken m-ben, ezért (4.27) teljesül minden kisebb m értékre is.

Speciálisan, ha az m-edik minimális távolságra lévő atomnak az 1 címkét adjuk, $k < m$-re kapjuk a

$$\Sigma_m \geq (m-1)v(R_m)$$

triviális alsó korlátot. Ezt az egyenlőtlenséget (4.27) képlettel és az $E_1^* < -e_d|v(s)|$ becsléssel kombinálva kapjuk, hogy (4.28) is teljesül. □

Minimális atompár távolság

15. Tétel. [67] *Legyen* $[\underline{R}, \overline{R}] \subseteq [0, s]$ *olyan intervallum, amelyre*

$$\int_s^\infty \left\lfloor \left(\frac{2r}{R}+1\right)^d \right\rfloor v'(r)dr \leq v(R) + |v(s)| \quad (R \in [\underline{R}, \overline{R}]), \tag{4.29}$$

és

$$\int_s^\infty \left\lfloor \left(\frac{2r}{\overline{R}}+1\right)^d \right\rfloor v'(r)dr < \min\left\{v(\underline{R}) + |v(s)|, \frac{1}{2}v(\underline{R}) + (1+\frac{e_d}{2})|v(s)|\right\} \tag{4.30}$$

teljesülnek. Akkor az

$$f(q) := v(q) + (2+e_d)|v(s)| - \int_s^\infty \left\lfloor \left(\frac{2r}{q}+1\right)^d \right\rfloor v'(r)dr \tag{4.31}$$

függvény legkisebb q zérushelye benne van az $(\overline{R}, \infty)$ nyílt intervallumban, továbbá $r_{\min} \geq q$.

Vegyük észre, hogy a (P3") tulajdonság implikálja a tétel feltételének teljesíthetőségét (vegyük az $\underline{R} = \overline{R} = R$ esetet).

Bizonyítás. Minden $m \geq 2$ egészre a (4.26) és (4.28) formulákból kapjuk, hogy az $R = R_m$ választással

$$(m-1)v(R) + (m+e_d)|v(s)| \leq \int_s^\infty (m-1)\left\lfloor \left(\frac{2r}{R}+1\right)^d \right\rfloor v'(r)dr,$$

ezért

$$v(R) + |v(s)| < v(R) + \frac{m+e_d}{m-1}|v(s)| \leq \int_s^\infty \left\lfloor \left(\frac{2r}{R}+1\right)^d \right\rfloor v'(r)dr.$$

Ez ellentmond a (4.29) formulának, hacsak az

$$R_m < \underline{R} \quad \text{vagy} \quad R_m > \overline{R}$$

esetek közül valamelyik nem teljesül. Tegyük fel, hogy az első eset igaz valamely $m \geq 2$-re. Legyen m az a legnagyobb egész, amelyre $R_m < \underline{R}$. Akkor $R_{m+1} > \overline{R}$, így

$$K(r) \leq m\left\lfloor \left(\frac{2r}{R_{m+1}}+1\right)^d \right\rfloor < m\left\lfloor \left(\frac{2r}{\overline{R}}+1\right)^d \right\rfloor,$$

és mivel $v(\underline{R}) \leq v(R_m)$ igaz, ezért (4.28) felhasználásával kapjuk, hogy

$$\frac{1}{m}\Big((m-1)v(\underline{R}) + (m+e_d)|v(s)|\Big) \leq \int_s^\infty \left\lfloor \left(\frac{2r}{\overline{R}}+1\right)^d \right\rfloor v'(r)dr.$$

Mivel $m \geq 2$, ezért ez ellentmond a (4.30) formulának. Ezért az első eset nem fordulhat elő. Speciálisan, azt kapjuk, hogy $m = 2$-re

$$r_{\min} = R_2 > \overline{R}.$$

Mivel $m = 2$-re (4.28) implikálja, hogy $f(r_{\min}) \leq 0$, továbbá a (4.29) formulából következik, hogy $f(\overline{R}) > 2|v(s)| > 0$, ezért a középérték tételből kapjuk, hogy f-nek van zérushelye az $(\overline{R}, \infty)$ nyílt intervallumban. Továbbá $r_{\min}$ nem lehet kisebb, mint ez a zérushely. □

Lineáris alsó korlát az optimumra

16. Tétel. [67] *Ha*

$$B_2 := -|v(s)| + \int_s^\infty K(r)v'(r)dr < \infty, \tag{4.32}$$

akkor

$$E_i^* \geq -B_2 \quad \textit{minden } i = 1, \ldots, n \textit{ indexre.} \tag{4.33}$$

Továbbá minden a (4.33) *formulának eleget tévő* B_2 *konstansra*

$$-\frac{B_2}{2}n \leq E^*. \tag{4.34}$$

Bizonyítás. A (4.27) speciális esete az, amikor $m = 1$ a

$$0 = \Sigma_1 \leq -|v(s)| + E_1^* + \int_s^\infty K(r)v'(r)dr = E_1^* + B$$

összefüggést adja, amiből $i = 1$ választással a (4.33) formulát kapjuk. Mivel az 1 címke választása tetszőleges, ezért (4.33) teljesül minden i-re. Végül (4.34) a (4.4) észrevételből következik. □

6. Következmény. *Ha q a minimális atompár távolság egy alsó korlátja, akkor* (4.33) *teljesül a*

$$B_2 := -|v(s)| + \int_s^\infty \left\lfloor \left(\frac{2r}{q} + 1\right)^d \right\rfloor v'(r)dr \tag{4.35}$$

konstanssal.

Bizonyítás. Használjuk a (4.26) formulát az $m = 2$ választással, valamint a (4.23) definíciót. Így B_2 korlátozható, ahogyan a (4.32) formulában definiáltuk:

$$B_2 \leq -|v(s)| + \int_s^\infty \left\lfloor \left(\frac{2r}{r_{\min}} + 1\right)^d \right\rfloor v'(r)dr \leq -|v(s)| + \int_s^\infty \left\lfloor \left(\frac{2r}{q} + 1\right)^d \right\rfloor v'(r)dr.$$

□

Ahogyan azt a fejezet bevezetésében említettük, Ruelle [56] egy potenciál függvényt *stabilisnak* nevez, ha az optimális klaszter energiaszintje alulról korlátozható az atomok számának lineáris függvényével. Egy rövid összefoglalást adunk most ezekről az eredményekről (bővebben lásd [56, 3.2.6. fejezet]). Azt mondjuk, hogy egy folytonos f függvény *pozitív típusú*, ha

$$\sum_{i=1}^n \sum_{j=1}^n f(x_i - x_j) \geq 0 \tag{4.36}$$

teljesül. Általában nem triviális megmutatni, hogy egy párpotenciál pozitív típusú, de Ruelle [56] hivatkozik egy ismert eredményre (Bochner [4]), ami szerint f akkor és csak akkor pozitív típusú, ha f Fourier transzformáltja pozitív.

4. Állítás. (Ruelle [56]) *Ha a v párpotenciális pozitív típusú és $v(0)$ véges, akkor stabilis és*

$$-\frac{n}{2}v(0) \leq E^*. \tag{4.37}$$

Bizonyítás. Valóban,

$$0 \leq \sum_{i=1}^{n} \sum_{j=1}^{n} v(||x_i^* - x_j^*||) = nv(0) + 2 \sum_{i<j} v(||x_i^* - x_j^*||),$$

ezért

$$-\frac{v(0)}{2} n \leq \sum_{i<j} v(||x_i^* - x_j^*||).$$

□

A fenti eredményeket úgy fogalmaztuk meg, hogy csak bizonyos feltételrendszerek teljesülését követeltük meg. Ezen módszerek tekinthetők úgy, mint algoritmusok: a felhasználónak meg kell mondania, hogy mi legyen a párpotenciál függvény és ha az eleget tesz a feltételeknek, akkor a fenti módszerek numerikus számításaival konkrét korlátokat kapunk a minimális atompár távolságra és az optimális energiaszintre.

A fejezet hátralevő részében két jól ismert és az irodalomban legtöbbször hivatkozott párpotenciál függvényre alkalmazzuk az ismertetett eljárásokat.

4.4. Lennard-Jones klaszterek

A fenti módszerek alkalmazását a Lennard-Jones klaszterekkel kezdjük, amelyek kémiai, fizikai és optimalizálási területeken is jelentős modellt képviselnek, mert

- rendkívül egyszerű, de elfogadható matematikai modellt adnak valós fizikai rendszerekre, például alacsony hőmérsékletű ritka gázok (például argon, kripton, xenon) viselkedésére;
- a modell könnyen szimulálható számítógépen, viszont az optimális szerkezetének megállapítása rendkívül nehéznek bizonyul, ezért globális optimalizáló eljárások egyik tesztfeladata is lehet (például az értekezés előszavában említett 38 atomos eset).

A *Lennard-Jones párpotenciál* általános alakja a

$$v_{t,\varepsilon}(r) = 4\varepsilon \left(\left(\frac{t}{r}\right)^{12} - \left(\frac{t}{r}\right)^{6} \right) \tag{4.38}$$

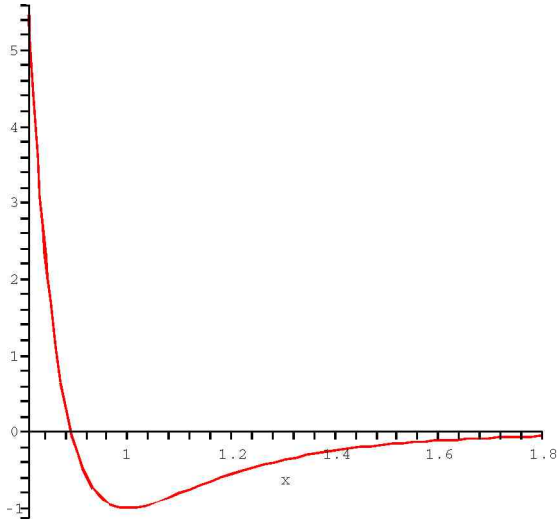
függvénnyel adható meg. A globális optimalizálási irodalomban a (4.38) függvényt az $\varepsilon = t = 1$ és $s = 2^{1/6}$,

$$v_{1,1}(r) = \frac{4}{r^{12}} - \frac{4}{r^6},$$

alakban (reduced unit), vagy az ún. *skálázott* Lennard-Jones potenciál ($\varepsilon = 1$, $t = 2^{-1/6}$, $s = 1$)

$$v_{2^{-1/6},1}(r) = \frac{1}{r^{12}} - \frac{2}{r^6} \tag{4.39}$$

alakban szokás vizsgálni. Ezen utóbbi alakot a 4.1. ábra szemlélteti.



4.1. ábra. A skálázott Lennard-Jones párpotenciál függvény.

A Lennard-Jones potenciál függvény a (4.2) és (4.38) képletek felhasználásával az

$$LJ_{t,\varepsilon}(x) = \sum_{1 \le i < j \le n} v_{\sigma,\varepsilon}(\|x_i - x_j\|). \tag{4.40}$$

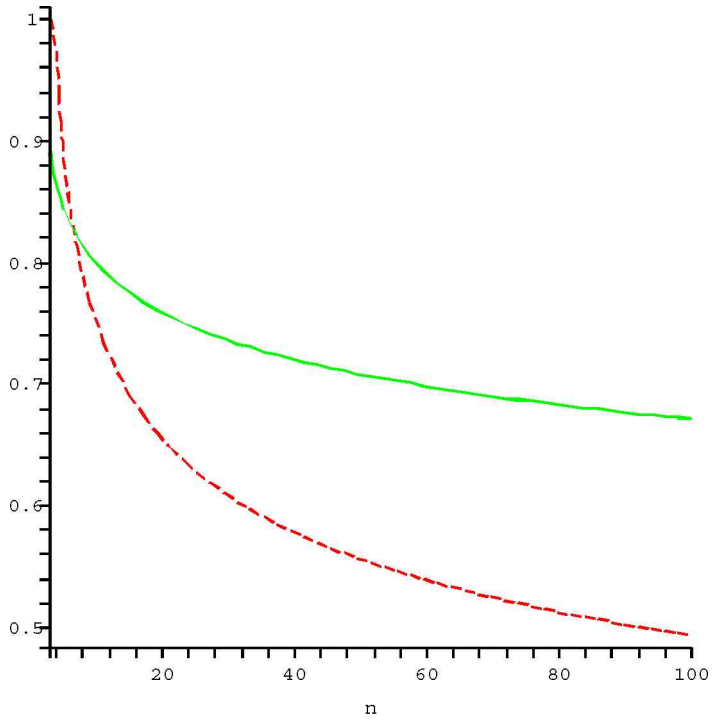
alakban definiálható.

4.4.1. Méretfüggő korlát a minimális atompár távolságra

A 3. Lemmát alkalmazva kapjuk, hogy $v_{t,\varepsilon} \le (n-2-e_d)|v(s)|$. Ebből az egyenlőtlenségből következik, hogy az optimális Lennard-Jones klaszterben ha $n > 2 + e_d$, akkor

$$q(n) = s\Big(\frac{\sqrt{\varepsilon^2 + \varepsilon|v_{t,\varepsilon}(s)|(n-2-e_d)} - \varepsilon}{(n-2-e_d)|v_{t,\varepsilon}(s)|}\Big)^{\frac{1}{6}} \tag{4.41}$$

egy alsó korlát a minimális atompár távolságra.



4.2. ábra. A Maranas és Floudas féle alsó korlát (szaggatott vonal) és a (4.41) által adott alsó korlát összehasonlítása a skálázott Lennard-Jones potenciálra.

A skálázott verzióra a (4.41) korlátot összevetve a Maranas & Floudas [39] által megadott

$$r_{\min} \geq \left(\frac{\sqrt{\frac{1}{2}n^2 - \frac{7}{2}n + 6} - 1}{\frac{1}{2}n^2 - \frac{7}{2}n + 5} \right)^{\frac{1}{6}}$$

korláttal azt kapjuk, hogy $n > 6$ esetén (4.41) jobb becslést ad. Ezt szemlélteti a 4.2. ábra.

4.4.2. Méretfüggetlen alsó korlátok a minimális atompár távolságra

Első változat

Az általános alak és a skálázott változat között a

$$v_{\sigma,\varepsilon}(r) = \varepsilon v_{2^{-1/6},1}(r/s), \tag{4.42}$$

skálázás visz át, tehát a minimális távolságot az s skálázza, míg a potenciál értékét ε. Ezért a skálázott verzióra adjuk meg a számításokat. Továbbá az egyszerűség

kedvéért a levezetésben a

$$v(r) = v_{2^{-1/6},1}(r) \text{ és } E = LJ_{2^{-1/6},1}$$

jelöléseket fogjuk használni.

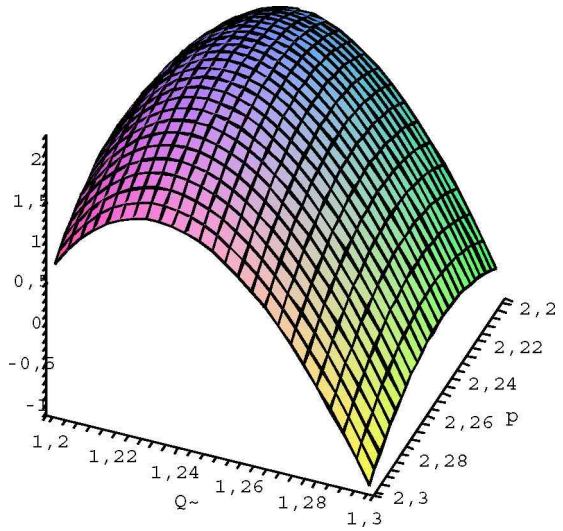
A 2. Lemma alapján $E_1^* < -1$. Az E_1^* értékre egy alsó korlát a 6. és 7. Lemmák által adható. Most választanunk kell egy alkalmas $R(Q,k)$ függvényt, amely az alsó korlátot -1 felett tartja. Ehhez definiáljuk az $R(Q,k) = pqQ^k$ $(pq \geq 1, Q > 1, k = 0, 1, 2, \ldots)$ függvényt. Könnyű látni, hogy

$$S_{LJ}(q,p,Q) := \sum_{k=0}^{\infty} \left(\frac{1}{pqQ^{12k}} - \frac{2}{pqQ^{6k}} \right) \left(\left(2pQ^{k+1}+1\right)^3 - \left(2pQ^k - 1\right)^3 \right) > -\infty \tag{4.43}$$

teljesül. Valóban, mivel $Q > 1$ igaz, ahogy k tart a végtelenbe az összeg első tagja (azaz $v(pqQ^k)$) gyorsabban tart a 0-hoz, mint ahogyan a második tag tart a végtelenbe (ez valójában a (P4) tulajdonság teljesülése). Ezért a

$$g_v(q,p,Q) := v(q) + 1 - (2p+1)^3 + S_{LJ}(q,p,Q) \tag{4.44}$$

függvény jól definiált. A 4.3. ábra mutatja ezt a függvényt, itt a $q = 0.618$ rögzített változóérték mellett ábrázoltuk azt. Jegyezzük meg, hogy a (P1) és (P2) tulajdonságok miatt a g_v függvénynek van zérushelye a $[0, s]$ intervallumban.



4.3. ábra. A $g_v(0.618, p, Q)$ függvény grafikonja.

Egy alsó korlát meghatározásához ezután a

$$\begin{aligned} \frac{\partial g_v}{\partial p}(q,p,Q) &= 0, \\ \frac{\partial g_v}{\partial Q}(q,p,Q) &= 0, \\ g_v(q,p,Q)+1 &= 0 \end{aligned}$$

háromváltozós nemlineáris egyenletrendszert kell megoldanunk. A (4.43) konvergens sorozat zárt alakját és a parciális deriváltakat egy szimbólikus-algebrai rendszerrel kaphatjuk meg. Ezek meghatározására a MAPLE 9 [38] programcsomagot használtuk. A nemlineáris rendszer megoldása

$$Q = 1.23474998, \quad p = 2.2408615800535, \quad q = 0.6184503450386, \tag{4.45}$$

lesz, amely tehát egy alsó korlátot ad az optimális skálázott Lennard-Jones feladatban az atompárok közötti minimális távolságra.

Ahogyan azt az 5. Következményben állítottuk, ez a korlát tovább javítható ha több paramétert vezetünk be. A 3 változós rendszer helyett 5 változót használva a

$$q = 0.618735677 \tag{4.46}$$

megoldást kapjuk, ami egy picit jobb alsó becslése a minimális távolságnak.

Megjegyzés. Az 5. következményt használva egyre több változó bevezetésével szignifikáns javítást nem tudunk elérni, viszont a számítások műveletigénye megnő.

Továbbfejlesztett változat

A (4.29) egyenlőtlenség a Lennard-Jones párpotenciálra a $d = 3$-ra a $[0, 0.653775s]$, míg $d = 2$-re a $[0, 0.752915s]$ intervallumokat adja. Ezért ebben az esetben $v_{t,\varepsilon}(\underline{R}) = \infty$. A (4.31) egyenlet megoldása $d = 3$-ra a

$$q = 0.654673s = 0.734846t \tag{4.47}$$

és $d = 2$-re a

$$q = 0.759006s = 0.851955t \tag{4.48}$$

alsó korlátokat adja.

Mint ahogyan azt a fejezet bevezetőjében említettük, korábbi eredmények is léteznek a Lennard-Jones problémában előforduló minimális távolságra. Ezeket az eredményeket a 4.1. táblázatban foglaltuk össze. A numerikus számításokat a Mathematica [69] programmal végeztük el (szemben az első változattal, ahol a MAPLE 9 programot használtuk), tekintettel arra, hogy a MAPLE 9 nem tudott olyan integrált kiszámolni, amelyben az integrandus tartalmaz alsó egészrész függvényt.

4.1. táblázat. Méretfüggetlen alsó korlátok a minimális atompár távolságra az optimális skálázott Lennard-Jones klaszterekben.

dimenzió	Xue [70]	Blanc [3]	13. Tétellel	15. Tétellel
2	–	0.7286	0.7284	0.7590
3	0.5	0.6108	0.6187	0.6547

Megjegyzés. A (4.41) képlet (azaz a méretfüggő alsó korlát) a skálázott Lennard-Jones klaszterre $d = 3$-ra $n < 139$ esetén, míg $d = 2$-re $n < 19$ esetén ad jobb alsó korlátot a táblázatban szereplő méretfüggetlen alsó korlátoknál.

Megjegyzés. A tézis elkészítése közben (és a [67] cikk közlésre beküldése után) jutott tudomásunkra egy friss eredmény (Schahinger *et al.* [57]), amely az itt ismertetett alsó korlátnál jobb értéket ad. Az eredmény egyelőre csak kézirat formájában létezik.

4.4.3. Lineáris alsó korlát az optimum értékére

A 13. Tétel (tehát az első változat) numerikus értékeit és a (4.42) összefüggést használva az optimális Lennard-Jones potenciál függvényre teljesül a

$$-138.6775911n \cdot \varepsilon \leq LJ^*_{\sigma,\varepsilon} \quad (n = 2, 3, \ldots)$$

lineáris alsó korlát a $d = 3$ esetben.

A 15. Tételből és a 6. Következményből $d = 3$-ra a

$$-68.9554\varepsilon n \leq LJ^*_{t,\varepsilon},$$

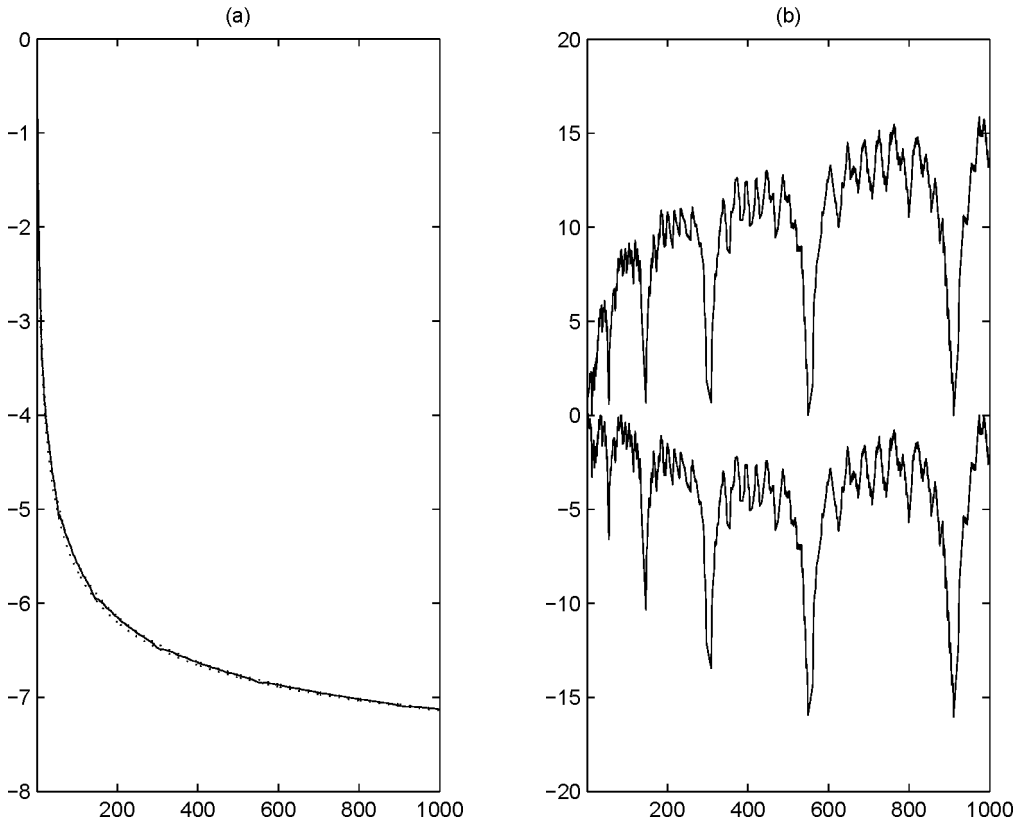
míg $d = 2$-re a

$$-9.4478\varepsilon n \leq LJ^*_{t,\varepsilon}$$

lineáris alsó korlátokat kapjuk.

4.4.4. Statisztikák empírikus adatokból

Az előző szakaszok folytatásaként mutatunk néhány statisztikát a redukált egységes Lennard-Jones feladatra. Az adatok a Cambridge Cluster Database (CCD) [6] és a Chemoinformatics Laboratory at the Department of Chemistry, University of Science and Technology of China [22] helyekről származnak. Ezeken a weboldalakon a redukált egységes Lennard-Jones feladatra megtalálhatjuk az eddigi legjobb megoldásokat (melyekről sok esetben feltehetjük, hogy globális optimumok)



4.4. ábra. (a) Az E^*/n hányados és egy $n^{1/3}$ szerinti köbös illesztés, (b) az alsó- és felső becslések eltérése.

a minimumpontok koordinátáival együtt $n \leq 1000$-re. Ebben a szakaszban $d = 3$, valamint az $E^* = LJ^*_{1,1}$ jelölést használjuk.

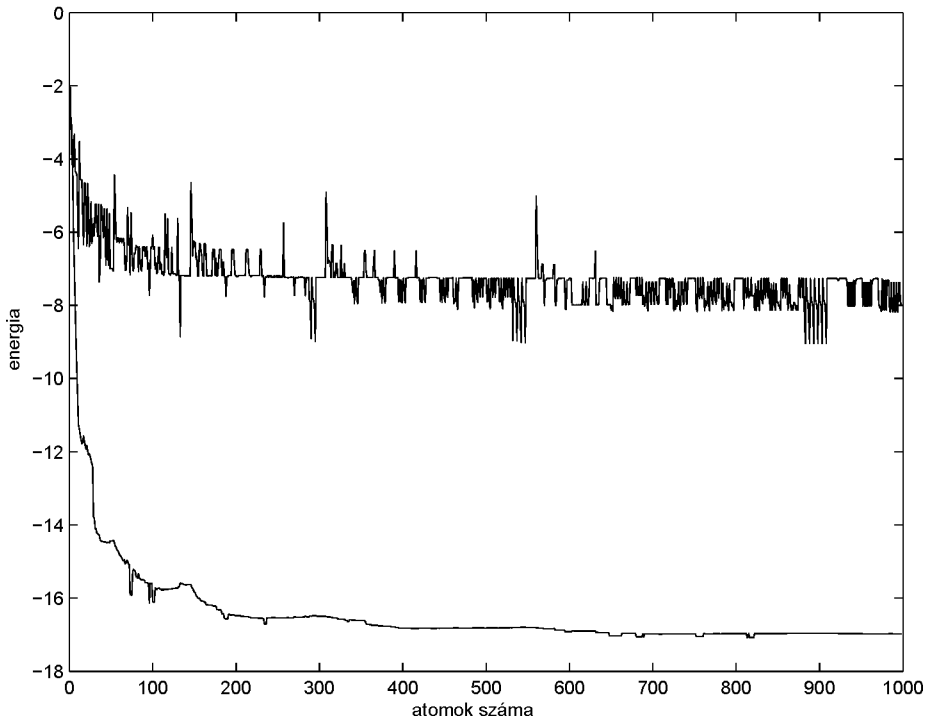
Az atomszámok függvényében a globális minimum értékeket, korlátokat az E^*_i értékekre, valamint a minimális és maximális atompár távolságokat vizsgáljuk.

Globális minimum értékek. A 4.4 (a) ábra a vélt globális minimum értékeket (tehát az eddig talált legjobb megoldások függvényértékeit) mutatja. Itt az $E^*(n)/n$ hányadost jelenítettük meg annak érzékeltetésére, hogy az optimális konfigurációk energiaszintje lineárisan korlátos. Az ábra tartalmaz egy köbös illesztést $n^{1/3}$-ban. Egy polinomiális alsó korlát az E^* értékre

$$-8.6263n - 59.0267n^{2/3} - 66.9958n^{1/3},$$

ezért (empirikusan) -8.6263 egy aszimptotikus alsó korlát az $E^*(n)/n$ értékre ahogyan $n \to \infty$. Láttuk, hogy a bizonyított alsó korlát -39.2205. Egy hasonló polinomiális felső korlát szintén létezik; az 4.4 (b) ábra mutatja az eltéréseket ezekhez a felső- és alsó becslésekhez képest.

Korlátok az atomokhoz tartozó energiákra. Mint azt láttuk, $E^*_i < -\varepsilon$ ($= -1$ a



4.5. ábra. Maximális és minimális E_i^*/ε értékek a klaszter méretének függvényében.

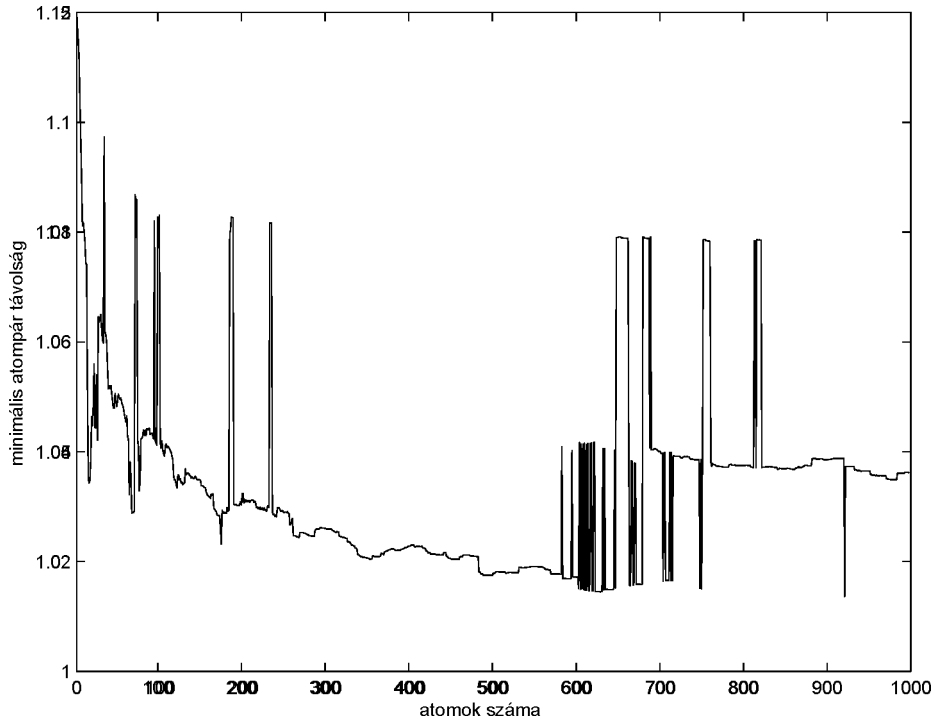
redukált és a skálázott verzióra). A rendelkezésre álló adatokból a $\min E_i^*$ és $\max E_i^*$ értékek számíthatók. A 4.5. ábrán ezek az E_i^*/ε értékekre vonatkozó minimumok és maximumokat láthatjuk az atomszámok függvényében.

Láthatjuk, hogy a $\max E_i^* < -3\varepsilon$ sejtés (vö. 4.2. alfejezet) empírikus értelmben teljesül.

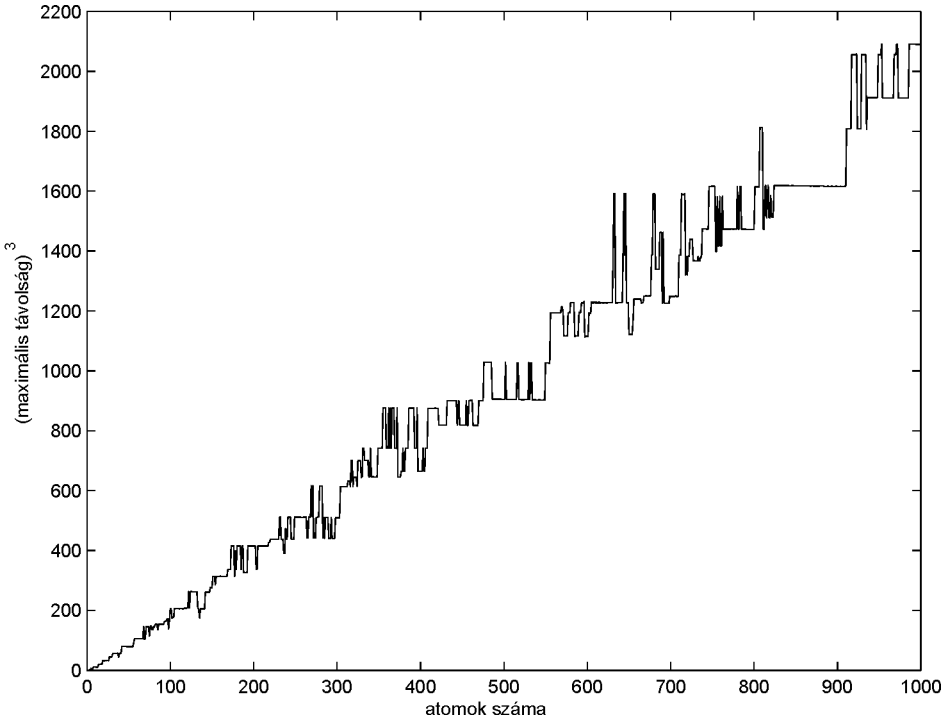
Ha $n > 30$, akkor a $\min E_i^*/\varepsilon$ hányados a -14 és -17.1 értékek között oszcillál (a pontos minimum érték $n = 823$-ra -17.0799), míg a bizonyított korlátunk -78.4410.

Korlátok a minimális távolságra. A 4.6. ábra az $r_{\min}/t$ értékeket mutatja. Ezekből az adatokból láthatjuk, hogy a minimális távolság mindig nagyobb, mint a Lennard-Jones párpotenciál zérushelye. A bizonyított eredményünk ettől az értéktől távol van; azonban az $r_{min} > t$ összefüggés teljesülésének bizonyítása reménytelennek tűnik.

Korlátok a maximális távolságra. Xue [70] sejtése szerint az optimális Lennard-Jones klaszter átmérője (azaz a maximális távolság) felülről korlátos $\mathcal{O}(n^{1/3})$ szerint. A 4.7. ábra mutatja, hogy ez a sejtés empirikusan jól megalapozott. Blanc [3] bizonyította, hogy az optimális Lennard-Jones klaszter mérete felülről korlátozható az atomok számával, tehát $\max_{i,j} r_{ij} \leq n$ teljesül.



4.6. ábra. Minimális atompár távolság a klaszter méretének függvényében.



4.7. ábra. A maximális távolság harmadik hatványa (redukált egységben, $t = 1$) a klaszter méretének függvényében.

4.5. Morse klaszterek

Az előző alfejezetben tanulmányozott Lennard-Jones klaszter egyik hiányosságának szokták felróni, hogy az optimális szerkezetben n függvényében nem mutat változatosságot. Emiatt (bizonyos, nem teljes) optimalizáló eljárások viszonylag hamar megtalálják az optimumot (láttuk, hogy $n = 1000$-ig léteznek jó megoldások). A másik népszerű modell a *Morse klaszter*, ahol a párpotenciál függvényt a

$$v_\rho(r) = e^{\rho(1-r)}\left(e^{\rho(1-r)} - 2\right) \tag{4.49}$$

függvénnyel definiáljuk, ahol $\rho > 0$ egy paraméter. A (4.49) és (4.2) képleteket használva a *Morse potenciált* a

$$M_\rho(x) = \sum_{1\le i<j\le n} v_\rho(\|x_i - x_j\|) \tag{4.50}$$

függvénnyel definiáljuk.

A v_ρ függvényben a ρ paraméter lehetővé teszi többféle anyag szerkezetének modellezését. A $\rho = 6$ értékre a Morse- és a skálázott Lennard-Jones potenciál hasonlóságot mutat: mindkét függvény ugyanazt a görbét írja le az $r = 1$ minimumpont környékén. A C_{60} molekulák közötti interakciót párpotenciállal szimulálva a $\rho = 13,6$ értéket kapjuk, míg például alkáli fémek szerkezetének modellezéshez a $\rho = 3,1$ választás bizonyul megfelelőnek.

A v_ρ függvény zérushelye és minimumpontja

$$t = 1 - \frac{\ln 2}{\rho} \quad \text{és} \quad s = 1$$

pont. Jegyezzük meg, hogy $\rho < \ln 2$ esetén a (4.49) függvénynek nincs zérushelye. Mindazonáltal globális optimalizálási környezetben általában a $\rho > 6$ esetek az érdekesek: ekkor az M_ρ globális optimumának megkeresése nehezebb feladat, mint a Lennard-Jones függvényre (lásd Doye *el al.* [16]). Másrészről viszont a ρ csökkenésével a méretfüggetlen minimális atompár távolság alsó korlátjának meghatározása egyre nehezebbé válik.

4.5.1. Méretfüggő alsó korlát a minimális atompár távolságra

A 3. Lemma használatával kapjuk, hogy $(\exp(\rho(1-r))-1)^2-1 \le (n-2-e_d)|v_\rho(s)|$. Ebből az egyenlőtlenségből adódik, hogy

$$q(n) = \max\left\{1 - \rho^{-1}\ln\left(\sqrt{|v_\rho(s)|(n-2-e_d)} + 1\right), 0\right\}, \tag{4.51}$$

amely egy alsó korlát a minimális atompár távolságra az optimális Morse klaszterben, amennyiben $n > 2 + e_d$. Ez a formula az

$$n \leq \left\lfloor (2 + e_d) + \frac{e^\rho(e^\rho - 2)}{|v_\rho(s)|} \right\rfloor$$

esetekben ad pozitív alsó korlátot.

4.5.2. Méretfüggetlen alsó korlát a minimális atompár távolságra

Első változat

Ki kell hangsúlyoznunk, hogy a Morse potenciál nem teljesíti a (P3) feltételt. A magyarázat az, hogy a v_ρ függvény az $r = 0$ esetben is definiált (tehát amikor két atom a tér ugyanazon pontjában van). Más szóval a (4.13) képletben szereplő G függvénynek két gyöke is van, azaz kicsi q értékekre negatívvá válik. Ezért a 4.3.1. szakasz általános módszere itt közvetlenül nem alkalmazható. Ebben az esetben a minimális atompár távolságra vonatkozó előzetes információ segíthet. Locatelli & Schoen [37] az optimális Morse klaszterek ilyen tulajdonságát vizsgálta, és bebizonyította, hogy ha $6 \leq \rho \leq 15$, akkor a minimális atompár távolság határozottan pozitív. Az ismertetett módszerük nagyban különbözik a Xue [70] által a Lennard-Jones klaszterekre adott módszertől, illetve a jelen értekezésben ismertett általános módszertől. Azonban ha felhasználjuk azt, hogy $r^* > 0$, ha $6 \geq \rho \geq 15$, akkor ez kiválthatja a (P3) tulajdonságot.

A fejezet hátralevő részében egy adott $\rho > 0$-ra az $M := M_\rho$ jelölést használjuk. A 2. Lemmából tudjuk, hogy $M_i^* < -1$ minden $i = 1, \ldots, n$-re és $\rho > 0$-ra. Mint a Lennard-Jones potenciálra, definiáljuk a $R(Q, k) := pqQ^k$ $(pq > 1, Q > 1, k = 0, 1, \ldots)$ függvényt. Az

$$S_M(q, p, Q) := \sum_{k=0}^{\infty} \left(\left(e^{\rho(1-pqQ^k)} - 1 \right)^2 - 1 \right) \left(\left(2pQ^{k+1} + 1\right)^3 - \left(2pQ^k - 1\right)^3 \right) \tag{4.52}$$

végtelen sorozat konvergens – az első tag (azaz $v_\rho(pqQ^k)$) gyorsabban tart nullához, mint a második végtelenbe – (és ez valójában megint a (P4) tulajdonság miatt van), ezért a

$$g_v(q, p, Q) := v_\rho(q) + 1 - (2p + 1)^3 + S_M(q, p, Q) \tag{4.53}$$

függvény jól definiált.

A 4.2. táblázatban a [37] cikkben közölt eredményeket hasonlítjuk össze az általános módszer használatával számolt eredményekkel. Hangsúlyozzuk viszont, hogy itt kihasználtuk, hogy a megfelelő ρ értékekre q nagyobb, mint a táblázat második

4.2. táblázat. Alsó korlátok a Morse klaszterekben található minimális atompár távolságokra különböző ρ paraméterek esetén.

ρ	q értéke [37] alapján	q értéke a 13. Tétel használatával
6	0.114	0.4985948046
7	0.376	0.6113121449
8	0.468	0.6796501438
9	0.528	0.7268978345
10	0.574	0.7618207355
11	0.613	0.7887781722
12	0.644	0.8102494106
13	0.672	0.8277671751
14	0.695	0.8423362542
15	0.715	0.8546451536

oszlopában szereplő érték. Látható, hogy az általános módszer így sokkal jobb eredményeket produkált.

A módszer csak $\rho \geq 6$ esetén működik. Ez egyrészt azért van, mert a megfelelő nemlineáris egyenletrendszernek nincs nemnegatív megoldása; másrészt a [37] cikkben is csak a fenti táblázatban szereplő ρ értékekre számolták ki az alsó korlátokat a szerzők (az ottani módszer további finomítása nem triviális).

Továbbfejlesztett változat

A továbbfejlesztett változattal kapott eredményeket a 4.3. táblázat tartalmazza. Jegyezzük meg, hogy $\rho = 6.3532$ esetén a Morse és a Lennard-Jones párpotenciáloknak ugyanaz a zérushelye. Az utolsó sor ($\rho = 4.967$) mutatja azt a legkisebb értéket, amelyre a 15. Tétel még alkalmazható. A táblázatban a minimális atompár távolságra kapott alsó korlátokon felül feltüntettük, hogy a különböző ρ paraméterre mi lesz a v_ρ párpotenciál zérushelye (t érték), valamint a $\underline{R}$ és $\overline{R}$ értékeket is. Fontos kiemelnünk, hogy itt nincs szükség előzetes információra az $r_{\min}$ értékére vonatkozóan.

Megjegyzés. A 4.4.2. szakaszban a Lennard-Jones klaszterek eredményeinek közlésekor említett Schachinger *et al.* [57] kézirat közöl eredményeket a Morse klaszterekre is. Habár az eredmények valamivel jobbak az itt ismertetetteknél az ottani módszer hátránya, hogy közvetlenül nem használható a Morse klaszterre (illetve általában olyan v párpotenciállal definiált modellre, amelyben a $v(0)$ értéke véges), csak abban az esetben, ha arra alkalmas módszerrel (mint a Locatelli & Schoen [37] vagy az itt ismertetett továbbfejlesztett változat) már ki tudjuk mutatni, hogy a minimális atompár távolság nagyobb, mint 0.

4.3. táblázat. Javított alsó korlátok az optimális Morse klaszterek minimális atompár távolságaira.

ρ	t	$\underline{R}$	$\overline{R}$	q értéke a 15. Tételből	q értéke a 13. Tételből
15	0.95379	0.00001	0.86424	0.865683	0.854645
14	0.95049	0.00197	0.85320	0.854691	0.842336
13	0.94668	0.00039	0.84018	0.841725	0.827767
12	0.94224	0.00077	0.82460	0.826193	0.810249
11	0.93699	0.00152	0.80559	0.807236	0.788778
10	0.93068	0.00302	0.78187	0.783551	0.761821
9	0.92298	0.00608	0.75135	0.753054	0.726898
8	0.91336	0.01250	0.71045	0.712129	0.679650
7	0.90097	0.02663	0.65212	0.653727	0.611312
6.353	0.89090	0.04058	0.59809	0.599581	(0.546518)
6	0.88448	0.06167	0.55928	0.560668	0.498595
5	0.86137	0.20982	0.33235	0.333473	–
4.967	0.86045	0.23439	0.30471	0.306227	–

4.5.3. Lineáris alsó korlát az optimum értékére

Ruelle [56] bizonyította, hogy ha $\rho > \ln 16 \approx 2.7726$, akkor a v_ρ párpotenciál Fourier transzformáltja pozitív típusú, ezért Bochner tételéből [4] és a 4. Állításból adódik, hogy v_ρ stabilis. A lineáris korlát,

$$-\frac{v_\rho(0)}{2}n \leq M_\rho^* \qquad (\rho > \ln 16) \tag{4.54}$$

meglehetősen gyenge, $\rho = 4.967$ értékre (ami a legkisebb olyan érték, amire a (P3") tulajdonság még teljesül) valamint $\rho = 15$ értékre a (4.54) formula rendre a $-1.0166 \cdot 10^4 n$ és $-5.3432 \cdot 10^{12} n$ értékeket adja. Ruelle gondolatmenetéből nem lehet korlátot kinyerni a minimális atompár távolságra.

A 4.4. táblázat tartalmazza az ismertetett módszereink használatával kapott lineáris alsó korlátokat különböző ρ paraméterekre.

4.6. Konklúzió és további feladatok

Ebben a fejezetben általános eljárásokat adtunk párpotenciál függvénnyel definiált atomklaszter feladatok optimális szerkezetének vizsgálatára. Az atompárok közötti (méretfüggő és méretfüggetlen) minimális távolságra és az optimumra adott lineáris alsó korlát hasznos információ lehet a (főleg nagyméretű) molekulák optimális szerkezetének meghatározásában.

4.4. táblázat. Alsó korlátok a Morse klaszterek optimumaira.

ρ	14. Tétel használatával	6. Következmény használatával
15	$-30.370n$	$-21.6176n$
14	$-32.240n$	$-22.5917n$
13	$-34.581n$	$-23.8037n$
12	$-37.594n$	$-25.3520n$
11	$-41.617n$	$-27.3977n$
10	$-47.255n$	$-30.2230n$
9	$-55.712n$	$-34.3707n$
8	$-69.762n$	$-41.0345n$
7	$-97.522n$	$-53.4416n$
6	$-177.619n$	$-84.4438n$
5	–	$-365.2798n$
4.967	–	$-461.7701n$

Érdekes kérdés még a maximális távolság (átmérő) meghatározása már egy másik feladat, az arra adható korlát az atomszámok függvénye lesz. Ide vonatkozó használható eredmény eddig nem ismert (kivéve Blanc [3] eredményét, ami azt mondja, hogy az n atomos Lennard-Jones klaszterben a maximális átmérő kisebb, mint n). Láthattuk, hogy a Lennard-Jones klaszterek esetén empirikus eredményekből a maximális távolság $\mathcal{O}(n^{1/3})$ nagyságrendű. A pontos becslés megadása azonban egy további kihívást jelent a kutatásoknak.

A végső cél egy olyan módszer kidolgozása, amely a jelenleg ismert legjobb megoldások globalitásának leellenőrzését végezné el matematikai szigorúsággal. Ehhez azonban tapasztalataink szerint (Vinkó & Neumaier [66]) nem elég egyszerűen csak a 2. fejezetben ismertetett intervallumos B&B módszert alkalmazni a (4.3) feladatra. Eredményre vezethet viszont például az eddig ismert legjobb megoldások struktúrális szerkezetének vizsgálata.

Összefoglalás

Az értekezés tárgya megbízható globális optimalizálási módszerek továbbfejlesztése, teljes globális optimalizálási feladatokat megoldó programok összehasonlításának elvégzésére alkalmas módszer kidolgozása, valamint atomklaszterek optimális szerkezetének vizsgálata. Az 1. Fejezetben ismertettük a tárgyalt feladatok általános definícióit, valamint a globális optimalizálási módszerek egy lehetséges osztályozását.

A 2. Fejezetben az intervallum aritmetikán alapuló globális optimalizálás alapvető foglamainak ismertetése után egy új intervallumos befoglaló függvényt vezettünk be, a kite befoglalást. Először az egydimenziós esettel foglalkoztunk. A módszer két korábbról ismert befoglalófüggvény szimultán használatán alapszik. Az 5. Tételben megmutattuk, hogy a kite egy differenciálható valós függvény értékkészletének alsó korlátjára mindig legalább olyan jó eredményt ad, mint a másik két befoglalófüggvény. Megvizsgáltuk a kite középpontjának optimális választásának lehetőségét, amelynek létezését és tulajdonságait a 6. Tétel mondja ki. Az intervallumos befoglaló függvények két fontos tulajdonsága, az izotonitás és a négyzetes konvergencia itt is teljesül, ezt a 7. és 8. Tétel bizonyítja. Az optimalizálás szempontjából hatékony tulajdonság továbbá a metszés, ami lehetővé teszi, hogy eliminájuk a keresési tartomány olyan részeit, amelyek garantáltan nem tartalmaznak globális minimumot. A 9. Tételben megmutattuk, hogy a kite előállításához szükséges információ hogyan alkalmazható a metszési tulajdonság kihasználásához. Végül standard tesztfüggvényeken elvégzett numerikus vizsgálatokkal kimutattuk, hogy az intervallumos korlátozás és szétválasztás típusú algoritmusba történő implementálással a feladatok kisebb számítási költséggel oldhatók meg.

A 2. Fejezet második felében a kite magasabb dimenzióba történő kiterjesztését tárgyaltuk. Egy lehetséges módszer a komponensenkénti kiterjesztés, ami az egydimenziós esetre támaszkodik. A 10. és 11. Tétel a kiterjesztés konstrukcióját és az optimális középpont választását tárgyalja. Ugyanúgy, mint az egydimenziós esetben, itt is bevezettünk egy metszési eljárást, aminek hatását a 12. Tétel ismerteti. Mivel a komponensenkénti kite, mint befoglaló függvény rendkívül költséges, ezért javasoltuk, hogy a metszési tulajdonsága miatt azt mint gyorsító technikát alkalmazzuk az intervallumos globális optimalizáló algoritmusban. Többdimenziós tesztfeladatokon végzett numerikus vizsgálatokkal kimutattuk, hogy a hagyományos eljáráshoz képest (főleg a nehezebben megoldható feladatokra) érdemes az ismertetett eljárást használni.

Az értekezés 3. Fejezete egy új módszertant ismertet (teljes) globális optimalizáló programok tesztelésére, azok megbízhatóságának vizsgálatára és egymással való összehasonlítására. Bemutattuk a teszteléshez felhasznált feladatok előkészítését, valamint a futtatásokhoz az időzítés megválasztásának kérdéskörét. Fontos szempont volt, hogy a különböző megoldó programok által megkívánt input formátumok előállításához elkészített konverterek helyes működését biztosítsuk. A futtatási eredmények alapján elkészítettük a legjobb megoldások listáját; ez összesen több, mint 1000 darab globális optimalizálási és feltétel kielégítési tesztfeladat globális optimumának megadását jelentette. A listát egy internetes oldalon elérhetővé tettük. Ismertettük továbbá, hogy a módszertan alapján elkészített számítógépes program segítségével a futtatási eredményekből automatikusan milyen táblázatokat, ábrákat készíthetünk. Ezek a kimutatások lehetővé teszik a tesztelt programok általános viselkedésének elemzését. A fejezet végén rövid összefoglalást adtunk néhány kurrens globális optimalizáló program teszteléséről és összehasonlításukról.

Az utolsó, 4. Fejezetben párpotenciállal definiált atomklaszter feladatok megoldásai optimális szerkezetének néhány meghatározó tulajdonságát vizsgáltuk. Az ismertetett módszerek általánosak abban az értelemben, hogy a párpotenciáltól csak bizonyos tulajdonságok meglétének teljesülését követeljük meg és azok megléte esetén lényegében algoritmikusan határozzuk meg az atomklaszter tulajdonságait. A 4.2. szakaszban az optimális szerkezetben előforduló minimális atompár távolságra vonatkozó méretfüggetlen alsó korlátok előállítására alkalmas formulákat mutattunk. Ezek a korlátok kisebb méretű klaszterek esetében a tapasztalati értékhez közeli számokat adnak. A 4.3. szakaszban ugyancsak a minimális távolságot vizsgáltuk, de már az atomok számától függetlenül. Itt két módszert mutattunk be, amelyek alkalmasak arra is, hogy az optimális szerkezet energiaszintjének lineáris alsó korlátját is meghatározzák. A 13. Tételben megmutattuk, hogy a minimális atompár távolságra az első esetben egy nemlineáris egyenletrendszer megoldásával kaphatunk alsó korlátot. A továbbfejlesztett változatban egy integrál formulát tartalmazó nemlineáris egyenlet megoldása ad alsó korlátot a méretfüggetlen minimális atompár távolságra. A fejezet végén két, a szakirodalomban legtöbbet vizsgált párpotenciálra (Lennard-Jones és Morse klaszterek) adtunk a bevezetett formulák használatával az addig ismert legjobb korlátokra jobb eredményeket.

Summary

The thesis deals with the development of rigorous global optimization techniques, a proposition of a methodology is given for benchmarking complete global optimization solvers and for the investigation of atomic cluster structures. In Chapter 1 the definitions were introduced and a classification of global optimization methods was listed.

In Chapter 2, after the definitions of global optimization based on interval arithmetic were given, a new interval inclusion function called kite was introduced. First, the one dimensional case was studied. The construction is based on a simultaneous use of two earlier inclusion functions. By Theorem 5 it was shown that the kite method gives an at least as good lower bound for the inclusion function as the better of the two earlier ones. We investigated the optimal choice of the center of the kite. The existence and the properties of this optimal center is provided in detail in Theorem 6. Two important properties of the interval inclusion functions are the isotonicity and the quadratic convergence. These properties hold for the kite method which are proved in Theorems 7 and 8. The pruning effect makes it possible to eliminate those parts of the search space which are guaranteed not to include global minimizer points. Theorem 9 shows how the available information can be used to make the pruning. Numerical studies were made on a large set of standard one dimensional test functions to show that the implementation of the kite method in a branch-and-bound type interval global optimization algorithm enables to solve the problems with reduced computational effort.

The second part of Chapter 2 deals with the higher dimensional kite. A componentwise extension was introduced which is based on the one dimensional case. Theorems 10 and 11 give the construction and the optimal choice of the center of the componentwise kite, respectively. Similar to the one dimensional case, the pruning effect plays important role here. The formulas for the pruning were given in Theorem 12. Since the computation of the componentwise kite is quite expensive, a proposition was made to use it as an accelerating tool in the global optimization context. Numerical comparisons with the traditional method and some recently proposed methods were made on a large set of standard test functions. These test results show that the usage of the componentwise kite is recommended, especially for the hard to solve problems.

In Chapter 3 a method for benchmarking complete global optimization solvers was developed. This method enables us to test and compare different global optimization solvers and to study their reliability. We discussed the preparation of the test problems and how the timing method was chosen. An important part of the testing process was to assure the correctness of the converters (to produce the different input formats for the solvers). A ranking of more than 1000 global optimization and constraint satisfaction test problems was made based on the benchmarking. This list was made available online. A testing environment was implemented based on the proposed methodology. The tables and figures can be obtained by using the introduced environment. These reports help one to study the general behaviour of the tested solvers. Finally, a short review on the benchmarking of some current state-of-the-art solvers was given.

The last chapter deals with the structural attributes of optimal atom cluster problems defined by pair potential functions. The proposed methods are general in the sense that only some properties are assumed for the pair potential function. Section 4.2 introduces formulas for size dependent bounds on the minimal inter-particle distances. These bounds are useful for relatively small clusters. In Section 4.3 size independent bounds were given. Two methods were proposed and these are able to derive linear lower bounds for the optimal energy level, too. For the first method, Theorem 13 shows that the solution of a nonlinear system of equations leads to a lower bound for the minimal interatomic distance. In the improved method a solution of a nonlinear equation gives a much better bound. Finally, these methods were applied for two well studied atomic cluster problems (Lennard-Jones and Morse clusters) and explicit results were reported.

Irodalomjegyzék

A hivatkozások végén található szám az adott előfordulás oldalszámát jelzi.

[1] R. S. Barr, B. L. Golden, J. P. Kelly, M. G. C. Resende, and W. R. Stewart. Designing and reporting on computational experiments with heuristic methods. *Journal of Heuristics*, 1:9–32, 1995. [46]

[2] E. Baumann. Optimal centered forms. *BIT*, 28:80–87, 1988. [12, 13]

[3] X. Blanc. Lower bounds for the interatomic distance in Lennard-Jones clusters. *Computational Optimization and Application*, 29:5–12, 2004. [63, 67, 84, 86, 92]

[4] S. Bochner. *Lectures on Fourier Integrals.* Princeton University Press, 1959. [78, 91]

[5] BogoMips Mini-Howto. `http://www.clifton.nl/bogomips.html` [47]

[6] Cambridge Cluster Database. `http://brian.ch.cam.ac.uk/CCD.html` [84]

[7] L. G. Casado, I. García, J. A. Martínez, and Ya. D. Sergeyev. New interval analysis support functions using gradient information in a global minimization algorithm. *Journal of Global Optimization*, 25:345–362, 2003. [1, 26, 31, 32, 43]

[8] COCONUT, COntinuous CONstraints Updating the Technology. `http://www.mat.univie.ac.at/~neum/glopt/coconut.html` [45]

[9] The COCONUT Benchmark. `http://www.mat.univie.ac.at/~neum/glopt/coconut/Benchmark/` [53]

[10] H. Crowder, R. S. Dembo, and J. M. Mulvey. On reporting computational experiments with mathematical software. *ACM Transactions on Mathematical Software*, 5:193–203, 1979. [46]

[11] A. E. Csallner, T. Csendes, and M. Cs. Markót. Multisection in interval branch-and-bound methods for global optimization I. Theoretical results. *Journal of Global Optimization*, 16:371–392, 2000. [11]

[12] T. Csendes. Automatikus differenciálás. *Polygon*, 6:33–41, 1996. [11]

[13] T. Csendes and D. Ratz. Subdivision direction selection in interval methods for global optimization. *SIAM Journal on Numerical Analysis*, 34:922–938, 1997. [11, 39]

[14] L. C. W. Dixon and G. P. Szegő. The global optimization problem: An introduction. In *Towards Global Optimization 2*, pages 1–15. North-Holland, Amsterdam, 1978. [47]

[15] E. D. Dolan and J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002. [46, 59]

[16] J. P. K. Doye, R. H. Leary, M. Locatelli, and F. Schoen. The global optimization of Morse clusters by potential energy transformations. *INFORMS Journal on Computing*, 2004. [88]

[17] R. Fourer, D. M. Gay, and B. W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming.* Duxbury Press, Brooks/Cole Publishing Company, 1993. [49]

[18] Frontline Systems. `http://www.solver.com` [55, 58]

[19] GAMS. `http://www.gams.com` [50]

[20] GAMS Solver descriptions, GAMS/OQNLP. `http://www.gams.com/solvers.htm#OQNLP` [58]

[21] Global Library. `http://www.gamsworld.org/global/globallib.htm` [46]

[22] Global minimal energies and coordinates of the LJ clusters. `http://chinfo.ustc.edu.cn/chmm/pubmats/LJ/` [84]

[23] GlobSol entry page. `http://www.mscs.mu.edu/~globsol/` [55, 58]

[24] N.I.M. Gould, D. Orban, and Ph.L. Toint. CUTEr, a constrained and unconstrained testing environment, revisited. `http://cuter.rl.ac.uk/cuter-www/problems.html`. [46]

[25] R. Hammer, M. Hocks, U. Kulisch, and D. Ratz. *Numerical toolbox for verified computing. I*, volume 21. Springer-Verlag, Berlin, 1993. [11, 26]

[26] R. Hammer, M. Hocks, U. Kulisch, and D. Ratz. *C++ Toolbox for Verified Computing I.* Springer-Verlag, Berlin, 1995. [26, 27, 40]

[27] E. Hansen. *Global Optimization Using Interval Analysis.* Marcel Decker, New York, 1992. [11]

[28] LINDO Systems Inc. LINGO 9. `http://www.lindo.com/lingom.html` [58]

[29] E. Janka. A comparison of stochastic methods for global optimization. `http://www.mat.univie.ac.at/~vpk/math/gopt_eng.html` [46]

[30] R. B. Kearfott. *Rigorous Global Search: Continuous Problems.* Kluwer, Boston, 1996. [6, 7, 11]

[31] R. Krawczyk and K. Nickel. Die zentrische Form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie. *Computing*, 28:117–137, 1982. [22]

[32] J.-L. Lagouanelle and G. Soubry. Optimal multisections in interval branch-and-bound methods of global optimization. *Journal of Global Optimization*, 30:23–38, 2004. [21]

[33] Y. Lebbah. ICOS (Interval COnstraint Solver). `http://www-sop.inria.fr/coprin/ylebbah/icos/` [58]

[34] E. Lee and C. Mavroidis. Solving the geometric design problem for spatial 3R robot manipulators using polynomial homotopy continuation. *Journal of Mechanical Design*, 124:652–661, 2002. [vi]

[35] Linpack Benchmark Java Version. `http://www.netlib.org/benchmark/linpackjava/` [47]

[36] M. Locatelli and F. Schoen. Fast global optimization of difficult Lennard-Jones clusters. *Computational Optimization and Applications*, 21:55–70, 2002. [63]

[37] M. Locatelli and F. Schoen. Minimal interatomic distance in Morse-clusters. *Journal of Global Optimization*, 22:175–190, 2002. [63, 89, 90]

[38] The Maplesoft Product Site. `http://www.maplesoft.com` [83]

[39] C. Maranas and C. Floudas. A global optimization approach for Lennard-Jones microclusters. *Journal of Chemical Physics*, 97:7667–7678, 1992. [63, 65, 81]

[40] M. Cs. Markót, T. Csendes, and A. E. Csallner. Multisection in interval branch-and-bound methods for global optimization. II. Numerical tests. *Journal of Global Optimization*, 16:219–228, 2000. [11]

[41] F. Messine and J.-L. Lagouanelle. Enclosure methods for multivariate differentiable functions and application to global optimization. *Journal of Universal Computer Sciences*, 4:589–603, 1998. [14, 33]

[42] H. Mittelmann. Benchmarks. `http://plato.la.asu.edu/topics/benchm.html` [46]

[43] M. Mongeau, H. Karsenty, V. Rouzé, and J.-B. Hiriart-Urruty. Comparison of public-domain software for black box global optimization. *Optimization Methods and Software*, 13:203–226, 2000. [46]

[44] R. E. Moore. *Interval Analysis.* Prentice-Hall, Englewood Cliffs, 1966. [8, 10]

[45] A. Neumaier. *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge, 1990. [7, 8, 13]

[46] A. Neumaier. Molecular modeling of proteins and mathematical prediction of protein structure. *SIAM Review*, 39:407–460, 1997. [61]

[47] A. Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica*, 13:271–369, 2004. [vi, 3]

[48] A. Neumaier, O. Shcherbina, W. Huyer, and T. Vinkó. A comparison of complete global optimization solvers. *Mathematical Programming*, 103:335–356, 2005. [45, 46]

[49] J. A. Martínez, L. G. Casado, I. García, Ya. D. Sergeyev, and B. Tóth. On an efficient use of gradient information for accelerating interval global optimization algorithms. *Numerical Algorithms*, 37:61–69, 2004. [43]

[50] J. A. Martínez, L. G. Casado, I. García, and B. Tóth. AMIGO: advanced multidimensional interval analysis global optimization algorithm. In C. A. Floudas and P. M. Pardalos, editors, *Frontiers in Global Optimization*, pages 313–326, Kluwer, Boston, 2004. [43]

[51] J. D. Pintér. *Global Optimization in Action.* Kluwer, Dordrecht, 1996. [58]

[52] H. Ratschek and J. Rokne. *Computer Methods for the Range of Functions.* Horwood, Chichester, England, 1984. [9]

[53] D. Ratz. *Automatische Ergebnisverifikation bei globalen Optimierungsproblemen.* PhD thesis, Universität Karlsruhe, 1992. [38]

[54] D. Ratz. *Automatic Slope Computation and its Application in Nonsmooth Global Optimization.* Shaker-Verlag, Aachen, 1998. [33]

[55] D. Ratz. A nonsmooth global optimization technique using slopes – the one dimensional case. *Journal of Global Optimization*, 14:365–393, 1999. [8, 22]

[56] D. Ruelle. *Statistical mechanics: Rigorous results.* W. A. Benjamin, Inc., New York-Amsterdam, 1969. [63, 78, 91]

[57] W. Schachinger, B. Addis, I. M. Bomze, and F. Schoen. New results for molecular formation under pairwise potential minimization. *Manuscript, submitted for publication*, 2005. [84, 90]

[58] H. Schichl and A. Neumaier. Interval analysis on directed acyclic graphs for global optimization. *Journal of Global Optimization*, 33:541–562, 2005. [49]

[59] O. Shcherbina, A. Neumaier, D. Sam-Haroud, X.-H. Vu, and T.-V. Nguyen. Benchmarking global optimization and constraint satisfaction codes. In Ch. Bliek et al., editor, *Global Optimization and Constraint Satisfaction*, pages 211–222. Springer, Berlin, 2003. [46, 47]

[60] S. Skelboe. Computation of rational interval functions. *BIT*, 14:87–95, 1974. [10]

[61] D. G. Sotiropoulos and T. N. Grapsa. Optimal centers in branch-and-prune algorithms for global optimization. *Applied Mathematics and Computation*, 169:247–277, 2005. [31, 32]

[62] M. Tawarmalani and N.V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical Programming*, 99:563–591, 2004. [58]

[63] B. Tóth and T. Csendes. Empirical investigation of the convergence speed of inclusion functions in a global optimization context. *Reliable Computing*, 11:253–273, 2005. [9]

[64] T. Vinkó. Minimal inter-particle distance in atom clusters. *Acta Cybernetica*, 17:105–119, 2005. [61, 64, 67, 68, 69, 71, 72]

[65] T. Vinkó, J.-L. Lagouanelle, and T. Csendes. A new inclusion function for optimization: Kite – the one dimensional case. *Journal of Global Optimization*, 30:435–456, 2004. [14, 15, 17, 21, 22, 26]

[66] T. Vinkó and A. Neumaier. Lower bounds for the optimization problems related to atom clusters. In *SCAN2004 Book of Abstracts*, page 117, 2004. [92]

[67] T. Vinkó and A. Neumaier. New bounds for atomic clusters. *Közlésre benyújtva*, 2005. [61, 64, 65, 66, 67, 73, 74, 75, 76, 77, 84]

[68] T. Vinkó and D. Ratz. A multidimensional branch-and-prune method for interval global optimization. *Numerical Algorithms*, 37:391–399, 2004. [33, 34, 35, 37]

[69] Wolfram Research Inc. `http://www.wolfram.com` [83]

[70] G. L. Xue. Minimum inter-particle distance at global minimizers of Lennard-Jones clusters. *Journal of Global Optimization*, 11:83–90, 1997. [63, 67, 84, 86, 89]

[71] G. L. Xue. An $\mathcal{O}(n)$ time hierarchical tree algorithm for computing force field in n-body simulations. *Theoretical Computer Science*, 197:157–169, 1998. [63]

[72] G. L. Xue, R. S. Maier, and J. B. Rosen. Minimizing the Lennard-Jones potential function on a massively parallel computer. In *ICS '92: Proceedings of the 6th International Conference on Supercomputing*, pages 409–416. ACM Press, 1992. [63]