

dc_1926_21

Megbízható módszerek elhelyezési és szimplex-alapú problémákra

MTA doktori értekezés tézisei

G.-Tóth Boglárka

Szeged
2022

dc_1926_21

1. fejezet

Megbízható módszerek speciális elhelyezési problémák megoldására

1.1. Maximális lefedési probléma egy hálózaton

A Maximális Lefedési Probléma (Maximal Covering Location Problem - MCLP), [7, 12, 11, 20], a vállalat-elhelyezés klasszikus problémája, amely számos területen alkalmazható, például az egészségügyben, a vészhelyzeti tervezésben, az ökológiában, a statisztikai osztályozásban, a belbiztonságban, stb. [1, 8, 10, 13, 36, 40].

Definiáljuk formálisan a klasszikus MCLP-t. Adott egy A véges halmaz, aminek az elemei a keresleti pontok, $\omega_a \geq 0$ kereslettel, minden $a \in A$ keresleti pontra. Szintén adott a lehetséges vállalatok halmaza, F , ahol p vállalatot akarunk elhelyezni, maximalizálva a lefedett keresletet. Egy a keresleti pont lefedett a p elemű $F^* \subset F$ által, ha létezik $f \in F^*$ vállalat, ami a -tól legfeljebb R távolságra van. Itt $R > 0$ egy fix érték, az úgynevezett *lefedési sugár*.

Ebben a munkában megoldottuk az MCLP következő változatát: a keresletet a hálózat élei mentén folytonos eloszlásúnak tételezzük fel, és keressük azt a p darab pontot a hálózat élein, amelyek a lefedett kereslet várható értékét maximalizálják.

Formálisan, legyen adott egy hálózat $N = (V, E)$, ahol V a csúcsok halmazát, $E \subseteq V \times V$ az irányítatlan élek halmazát jelöli. Minden $e \in E, e = (u, v), u, v \in V$ élhez adott egy l_e élhossz, ezáltal a $[0, l_e]$ intervallummal azonosítható, és így bármely $x \in [0, l_e]$ az e élen lévő pontként tekintünk, amely u -tól x távolságra és v -tól $l_e - x$ távolságra van. Ezzel a definícióval a V csúcsai közötti legrövidebb út távolsága könnyen kiterjeszthető az élek pontjainak távolságára, amit d -vel jelölünk. Továbbá minden e élnek van egy $\omega_e \geq 0$ súlya (az összkereslete), és egy f_e sűrűségfüggvénye, amely a keresletet modellezi az e él mentén. Tegyük fel, hogy adott az $R > 0$ lefedési sugár, és egy x pont az $e \in E$ élen. Az x pont a t_1, \dots, t_p vállalatok által lefedett, ha

$$\min_{1 \leq i \leq p} d(t_i, x) \leq R. \quad (1.1)$$

Az e él várható kereslete, amelyet a $\mathbf{t} = (t_1, \dots, t_p)$ vállalatok lefednek, a következőképpen adódik:

$$\omega_e \int_0^{l_e} \delta_e(x; \mathbf{t}) f_e(x) dx,$$

ahol $\delta_e(x; \mathbf{t}) = 1$, ha $x \in e$ lefedett $\mathbf{t} = (t_1, \dots, t_p)$ által, azaz ha (1.1) teljesül, egyébként 0 értéket vesz fel.

Ezzel az optimalizálási probléma a következőképpen írható fel,

$$\max_{\mathbf{t} \in E^p} C(\mathbf{t}) := \sum_{e \in E} \omega_e \int_0^{l_e} \delta_e(x; \mathbf{t}) f_e(x) dx. \quad (1.2)$$

Az (1.2) feladat elég sok nehezítő tulajdonsággal rendelkezik. Nem írható fel analitikusan, nemlineáris, tartalmaz diszkrét döntési részt, hogy melyik élekre helyezzünk vállalatot, és folytonos részt is, miszerint az adott éleken hova helyezzük a vállalatot. Sőt, az (1.2) általában nem Lipschitz folytonos, csak ha korlátosak a kereslet f_e sűrűségfüggvényei. Ez utóbbit feltételezve dolgoztuk ki a következő merőben új adathalmazra épülő módszert.

1.1.1. A globális optimalizálási megközelítés

Az (1.2) MINLP megoldására egy B&B algoritmust javasolunk. Mint minden B&B eljárásnál, az eredeti problémát részproblémákra bontjuk, és megfelelő korlátozó eljárásokkal kiiktathatjuk azokat a részproblémákat, amelyek nem tartalmazhatják a globális optimumot.

1.1.1.1. Felosztási szabály

Annak érdekében, hogy elkerüljük a p élből álló élhalmazok minden lehetséges kombinációjának felsorolását, (rész)élekből álló halmazokat hozunk létre. Ahelyett, hogy minden egyes élhez egy egészértékű változót rendelnénk, amely az adott élen található vállalatok számát jelzi, az egészértékű változókat a (rész)élek halmazaihoz rendeljük, amelyeket a továbbiakban *élhalmazok*-nak nevezünk, az élhalmazok halmazát pedig *szuperhalmazok*-nak.

Pontosabban, egy élhalmaz az E (rész)élek véges gyűjteménye; egy szuperhalmaz S bármely olyan halmaz, amely a következő formájú: $(E_1, p_1; E_2, p_2; \dots, E_k, p_k)$, ahol E_1, E_2, \dots, E_k diszjunkt élhalmazok és p_1, \dots, p_k ($\sum_k p_k = p$) szigorúan pozitív egész számok, amelyek azt jelzik, hogy pontosan p_j vállalat található az E_j élhalmazokon belül minden $j = 1, \dots, k$ esetén.

Kezdő szuperhalmazok

A B&B fa gyökere az $S_0 = (E, p)$ szuperhalmaz. Az S_0 -t E adott E_1, E_2, \dots, E_k partíciójának felhasználásával bontjuk fel: $\binom{p+k-1}{p}$ darab szuperhalmazzal helyettesíthetjük S_0 -t, amelyekre igaz, hogy $(E_{i_1}, p_{i_1}; \dots; E_{i_l}, p_{i_l})$, alakúak, ahol $\{i_1, \dots, i_l\} \subseteq \{1, \dots, k\}$ és $p_1 + \dots + p_l = p$. Megfigyelhető, hogy bár az ilyen kiindulási lista mérete exponenciálisan növekszik p -ben, a vizsgált MINLP nehézsége miatt csak olyan problémákat tudunk kezelni, amelyeknél p értéke viszonylag alacsony. Ezért a kiindulási lista mérete nem fog túl nagyra nőni.

A hálózat éleinek a felosztásához egy diszkrét MCLP-t írunk fel a következőképpen: tekintünk egy diszkrét lefedési problémát, amelyben a lehetséges helyek és a felhasználók is a hálózat élei, amelyek kereslete ω_e . Az e felhasználó és az f lehetséges vállalat (él) közötti távolságot, $d^*(e, f)$, az e és f pontjai közötti legkisebb távolságként definiálunk. Ezután egy e felhasználót akkor tekintünk lefedettnek, ha $d^*(e, f) \leq R$ valamely f élre. Ha ezt a diszkrét MCLP-t megoldottuk, és f_1^*, \dots, f_p^* egy optimális megoldás, akkor az E_1, \dots, E_p élhalmazokat úgy alakítjuk ki E összes élét felhasználva, hogy minden E_j élhalmaz azokat az e éleket tartalmazza, amelyekhez f_j^* a legközelebb van.

Egy szuperhalmaz felosztása

A B&B algoritmus konvergenciájának garantálása érdekében általában a lista elemeinek tetszőlegesen kicsivé kell válniuk. Definiáljuk egy E^* élhalmaz $\lambda(E^*)$ átmérőjét az E^* (rész)élei hosszának összegeként, és definiáljuk egy S szuperhalmaz $\lambda(S)$ átmérőjét az élhalmazainak maximális átmérőjeként,

$$\lambda(E_1, p_1; E_2, p_2; \dots; E_k, p_k) = \max_j \lambda(E_j).$$

Az $S = (E_1, p_1; E_2, p_2; \dots; E_k, p_k)$ szuperhalmazt a következőképpen osztjuk fel: először is megkeressük a legnagyobb átmérőjű E_{j^*} élhalmazát,

$$\lambda(E_1, p_1; E_2, p_2; \dots; E_k, p_k) = \lambda(E_{j^*}).$$

Ezután az E_{j^*} élek halmazát két részhalmazra osztjuk két „központi” él meghatározásával, majd az élek klaszterezésével ezen élek körül, kivéve ha csak egy élt tartalmaz, mert ez esetben kettévágjuk a középpontjánál. A folyamat, hasonlóan a kezdő szuperhalmaz felosztásához, egy segéd MCLP megoldásán alapul: egy 2 vállalatot elhelyező diszkrét fedési problémát vizsgálunk, amelyben lehetséges helyként az E_{j^*} élhalmaznak éleit tekintjük, felhasználóként pedig a hálózat $e \in E$ éleit ω_e kereslettel. Miután ezt a diszkrét MCLP-t megoldottuk, és egy optimális f^+, f^- megoldást kaptunk, felépítjük az $E_{j^*}^+$ és $E_{j^*}^-$ élhalmazokat úgy, hogy $E_{j^*}^+$ tartalmazza az f^+ -hoz közelebbi E_{j^*} -ban lévő e éleket, míg a többit $E_{j^*}^-$.

Az E_{j^*} élhalmaz $E_{j^*}^+$ -ra és $E_{j^*}^-$ -ra való felosztása esetén az S szuperhalmazt $p_{j^*} + 1$ szuperhalmazra osztjuk úgy, hogy i és $p_{j^*} - i$ vállalatot rendelünk $E_{j^*}^+$ -hoz és $E_{j^*}^-$ -hoz, minden $i = 0, 1, \dots, p_{j^*}$ esetén.

A konstrukcióból következik, hogy

1.1.1. állítás. *A szuperhalmazok felosztási szabálya kimerítő, azaz az $\{S_q\}_{q=0}^{\infty}$ szuperhalmazok végtelen egymásba ágyazott sorozata esetén $\lambda(S_q) \rightarrow 0$ -hoz, ahogy $q \rightarrow \infty$.*

1.1.1.2. Korlátozási szabályok

Az (1.2) célfüggvény C értékére vonatkozó alsó korlátokat úgy kaphatjuk meg, hogy a C értékét a kiértékelés alatt álló szuperhalmazban lévő p (rész)élek középpontjaként vett megoldáson értékeljük ki. A következőkben a felső korlátok kiszámítására irányuló különböző stratégiákat ismertetjük.

Árnyékkorlát

A legegyszerűbb módja, hogy felső korlátot kapjunk C -re adott S szuperhalmaz esetén, ha az S összes pontját, valamint azokat a pontokat, amelyek legfeljebb R távolságra vannak S valamelyik pontjától, lefedettnek tekintjük. Formálisan, az $S = (E_1, p_1; \dots, E_k, p_k)$ szuperhalmazon C -re vonatkozó árnyékkorlát, $\bar{C}_{SB}(S)$, a következőképpen számítható ki:

$$\bar{C}_{SB}(S) := \sum_{e \in E} \omega_e \int_0^{l_e} \delta_e^{SB}(x; S) f_e(x) dx, \quad (1.3)$$

ahol $\delta_e^{SB}(x; S) = 1$, ha x legfeljebb R távolságra van valamely $y \in E_j$ -től, egyébként 0.

A konstrukció alapján az árnyékkorlát monoton, azaz érvényes az

1.1.2. állítás. Ha $S = (E_1, p_1; \dots, E_k, p_k)$ és $S' = (E'_1, p_1; \dots, E'_k, p_k)$ olyan szuperhalmazok, amelyekre $E_i \supseteq E'_i$ minden $i = 1, \dots, k$ -ra, akkor

$$\bar{C}_{SB}(S) \geq \bar{C}_{SB}(S'). \quad (1.4)$$

Továbbá konvergencia is, formálisan teljesül az

1.1.3. állítás. Ha $\{(s_1^q, 1; \dots, s_p^q, 1)\}_q$ olyan szuperhalmazok sorozata, ahol minden s_j része egy e_j élnek, amely egy t_j pontba konvergál amikor $q \rightarrow \infty$, akkor $\bar{C}_{SB}((s_1^q, 1; \dots, s_p^q, 1)) \rightarrow C(t_1, \dots, t_p)$.

A korlátok tehát tetszőlegesen élesek, amikor a szuperhalmazok átmérője nullára csökken. Következésképpen fennáll az

1.1.4. következmény. Egy B&B módszerben, ha egy kimerítő felosztási szabályt használunk, ahogyan azt az 1.1.1.1 alfejezetben leírtuk, és egy konvergencia korlátozó szabályt, mint az árnyékkorlát, akkor a módszer konvergens.

MCLP korlát

A \bar{C}_{MCLP} felső korlátot a diszkrét MCLP egy változatának megoldásával kapjuk: egy diszkrét fedési problémát tekintünk, amelyben lehetséges helyként az $S = (E_1, p_1; \dots, E_k, p_k)$ szuperhalmaz (rész)élei szerepelnek, felhasználóként pedig a hálózat e élei, amelyek kereslete ω_e , illetve a $d^*(e, f)$ távolságot e és f (rész)él között az e és f pontjai közötti legkisebb távolságként számoljuk. Mint eddig, egy e felhasználót akkor tekintünk lefedettnek, ha $d^*(e, f) \leq R$ valamely $f \in E_j$ (rész)élre.

Egy e élt teljes mértékben lefedettnek tekintünk (és így az ω_e keresletet lefedettnek vesszük), amint egy $f \in E_j$ élen van R távolságra lévő pont e egy pontjához. Továbbá, mivel az egyes E_j élhalmazokon belül a vállalatok p_j száma adott, megszabjuk, hogy legfeljebb p_j különböző élet lehet kiválasztani minden E_j élhalmazból. Formálisan,

$$\begin{aligned} \max \quad & \sum_{e \in E} \omega_e z_e \\ \text{f.h.} \quad & z_e \leq \sum_{f \in \cup_j E_j} a_{ef} y_f \quad \forall e \in E \\ & \sum_{f \in E_i} y_f \leq p_i, \quad i = 1, 2, \dots, k \\ & y_f \in \{0, 1\} \quad \forall f \in \cup_j E_j \\ & z_e \in \{0, 1\} \quad \forall e \in E, \end{aligned} \quad (1.5)$$

ahol az a_{ef} paraméter értéke 1, ha létezik $f \in E_j$ valamely j -re, hogy $d^*(e, f) \leq R$, egyébként 0.

A konstrukció alapján a diszkrét fedési probléma optima helyes felső korlátot ad C -re az S szuperhalmazon. Ellentétben azzal, hogy a \bar{C}_{SB} árnyékkorlát monoton, ez a korlát kis szuperhalmazok esetén sem biztos, hogy éles, mivel ha egy él bármely pontja fedett, akkor az él összes kereslete fedettnek tekintett. Emiatt, ha csak ez a korlátozó módszer lenne adott, a B&B módszer nem lenne konvergens.

Ez a korlát azonban könnyen élesíthető, ha egy e él összes kereslete helyett egy kisebb mennyiséget, ω_e^* -t tekintjük,

$$\omega_e^* = \omega_e \int_0^{l_e} \delta_e^{SB}(x, S) f_e(x) dx, \quad (1.6)$$

ahol $\delta_e^{SB}(x, S)$ az árnyékkorlátban (1.3) meghatározottak szerint értelmezett.

A \bar{C}_{MCLP} MCLP korlátot az (1.5) optimumaként számítjuk, ahol az ω_e súlyokat az ω_e^* súlyokkal helyettesítjük (1.6) szerint.

Az így kapott MCLP korlát a konstrukció alapján monoton. Sőt, ha már minden élhalmaz egyetlen egy él része, akkor a kapott korlát pontosan megfelel az árnyékkorlátnak, és így ugyanazokkal a konvergencia tulajdonságokkal bír, mint az árnyékkorlát. Fontos tényező az is, hogy mivel felső korlátra van szükségünk, egy (durvább, de kevésbé költséges) felső korlátot kapunk, ha az IP (1.5) helyett annak LP relaxációját oldjuk meg.

Kevert korlát

A \bar{C}_{SB} és \bar{C}_{MCLP} felső korlátok általában megfelelők, ha a fedési területek nagy átfedésekkel rendelkeznek. Ha ezzel szemben a lefedett területek csaknem diszjunktak, a probléma (majdnem) független egy-egy vállalatból álló problémák sorozatára bontható, és így éles korlátok adhatók.

Pontosabban, $S = (E_1, p_1; \dots; E_k, p_k)$ esetén kombinálhatjuk a következő korlátokat: az E_j -re vonatkozó $\bar{C}_{SB}(E_j, 1)$ árnyékkorlátot bármely olyan felső korláttal, amit jelöljünk $\bar{C}_1(E_j)$ -vel, ami korlátot ad arra a problémára, amikor csak egy vállalatot helyezünk el E_j valamelyik pontján. Így az úgynevezett kevert korlát $\bar{C}_{MB}(S)$ a következőképpen definiálható.

$$\bar{C}_{MB}(S) = \sum_{j=1}^k \min \{ p_j \bar{C}_1(E_j), \bar{C}_{SB}(E_j, 1) \},$$

ahol $\bar{C}_{SB}(E_j, 1)$ az árnyékkorlát E_j -re. A probléma tehát redukálódik az 1 vállalatos probléma felső korlátjának a megkonstruálására az E_j élhalmazra. Legyen \mathcal{F}_j a hálózat kis részszegegmenteinek gyűjteménye, amelyre

$$E_j \subseteq \bigcup_{f \in \mathcal{F}_j} f,$$

ekkor $\bar{C}_1(E_j)$ felső korlátnak vehetjük az árnyékkorlátok maximumát egy vállalat f -en való elhelyezésére, minden $f \in \mathcal{F}_j$ esetén. Vagyis,

$$\bar{C}_1(E_j) = \max_{f \in \mathcal{F}_j} \bar{C}_{SB}((f, 1)),$$

így

$$\bar{C}_{MB}(S) = \sum_{j=1}^k \min \left\{ p_j \max_{f \in \mathcal{F}_j} \bar{C}_{SB}((f, 1)), \bar{C}_{SB}(E_j, 1) \right\}.$$

Megjegyezzük, hogy a konstrukció alapján a \bar{C}_{MB} kevert korlát monoton. Mivel azonban minden egyes E_j élhalmaz fedettségét külön-külön számítja ki, és adja össze, ha a fedett területek átfedésben vannak valahány élhalmazra, az átfedéseket többször számolja. Ezért a korlát önmagában nem mindenhol konvergens.

1.1.2. Eredmények

A folytonos kereslettel adott MCLP feladatra egy újfajta korlátozás és szétválasztás módszert adtunk, ahol az elágazási eljárás meglehetősen sajátos, mivel sikeresen kihasználja azt a ténytet, hogy az elhelyezési döntések egy hálózaton születnek.

Különböző korlátozási szabályokat javasoltunk, és különböző hálózatokon teszteltük ezeknek a működését. Bizonyosodott, hogy az úgynevezett Smart stratégia, amely egy tanulási folyamat révén a B&B fa több szintjén részfeladatonként azonosítja a legígéretesebb korlátozási stratégiát, a futási idő szempontjából a legjobb eredményt szolgáltatja.

1.2. Cégbővítés a létező vállalatok változtatásával

Ebben a kutatásban egy folytonos *versenyző* vállalatelhelyezési problémát vizsgálunk, ami azt jelenti, hogy ugyanabban a régióban más vállalatok is vannak, amelyek ugyanazt a terméket vagy szolgáltatást kínálják, és verseny folyik a piaci részesedés vagy a nyereség maximalizálásáért. Ha a lánc már működik néhány vállalattal a térségben, és bővíteni kívánja jelenlétét, akkor az új vállalat elhelyezése mellett a meglévő vállalataiba is beruházhat, növelve vagy csökkentve azok minőségét, vagy bezárva néhányat közülük (hogy a rájuk szánt költségvetést a többi vállalatba fektesse). Azonban, a szakirodalomban szereplő modellek egyikében sem megengedett, hogy az elhelyező lánc meglévő vállalatai változtassák minőségüket vagy bezárják azokat. Ebben a munkában pontosan ezt a hiányt pótoljuk.

Az új modell, amelyet itt bevezetünk, a [17] modelljének kiterjesztése. Az eredeti modellben a bővülő lánc csak egyetlen lehetőséget vesz figyelembe a nyereség növelésére: egy új vállalat megnyitását, amit mindenféleképpen megnyit. Az új modellben feltételezzük, hogy a bővülő láncnak már van néhány működő vállalata a területen, egyébként, ha a lánc most tör csak be a piacra, az eredeti modelltől csak a költségvetési feltételben különbözik, illetve abban, hogy nem kötelező új vállalatot telepítenie.

A modellhez a következő jelöléseket vezetjük be:

Indexek

- i a keresleti pontok indexe, $i \in I = \{1, \dots, i_{\max}\}$,
- j a meglévő vállalatok indexe, $j \in J = \{1, \dots, j_{\max}\}$, ezek közül az első $k < j_{\max}$ a bővülő láncé, $j \in J_1 = \{1, \dots, k\}$.

Változók

- f_0 az új vállalat helyének koordinátái, $f_0 = (f_0^1, f_0^2)$,
- α_0 az új vállalat minősége,
- α_j a lánc tulajdonában lévő j -edik meglévő vállalat minősége, $j \in J_1$, jelenleg $\alpha_j = \tilde{\alpha}_j$,
- y_j bináris változó, 0, ha a j vállalat bezár ($j \in J_1$) vagy nem nyit ki ($j = 0$); 1 egyébként,
- ns a probléma változói, $ns = (f_0, \alpha_0, \alpha_1, \dots, \alpha_k, y_0, \dots, y_k)$.

Paraméterek

- p_i az i keresleti pont helye,
- w_i a p_i koncentrált éves kereslete (vagy vásárlóereje),
- f_j a meglévő j vállalat helye,
- d_{ij} távolság p_i és f_j között,
- d_i^{\min} annak a p_i körüli körnek a sugara, ahol új vállalat telepítése nem engedélyezett,
- S a vállalat helyének keresési tere,
- $\tilde{\alpha}_j$ a meglévő f_j vállalat jelenlegi minősége, $j \in J$,
- $g_i(\cdot)$ egy nem csökkenő (és nem negatív) függvény,
- \tilde{u}_{ij} a p_i vásárló számára f_j vállalat hasznossága, $\tilde{u}_{ij} = \tilde{\alpha}_j / g_i(d_{ij})$, $i \in I$, $j \in J$,
- α_{\min} az új vállalat minimális minőségi szintje,
- α_{\max} az új vállalat maximális minőségi szintje,
- α_{\max}^j maximális minőségi szint, amelyet f_j elérhet, $j \in J_1$,
- A_j az f_j vállalat fenntartásának éves költsége, $j \in J_1$,
- C_j az f_j vállalat bezárásának költsége, $j \in J_1$,
- B a lánc éves költségvetése (új vállalat megnyitására, a meglévő vállalatok minőségének változtatására vagy bezárására, valamint a nyitva tartott vállalatok működési költségeire).

Számított értékek

- $d_i(f_0)$ távolság az új vállalat és p_i között,
- $u_{i0}(f_0, \alpha_0)$ a p_i vásárló számára az új vállalat hasznossága, $u_{i0}(f_0, \alpha_0) = \alpha_0 / g_i(d_i(f_0))$,
- $u_{ij}(\alpha_j)$ a p_i vásárló számára f_j vállalat hasznossága $j \in J_1$, $u_{ij}(\alpha_j) = \alpha_j / g_i(d_{ij})$,
- $V_j(\alpha_j)$ az f_j minőségének α_j -re történő változtatásának éves költsége, $j \in J_1$,
- $M(ns)$ a lánc által megszerzett éves piaci részesedés,
- $F(M(ns))$ a lánc által elért éves várható bevétel,
- $R_j(\alpha_j)$ a j -edik vállalat éves működési költsége, $j \in J_1$,
- $T(ns)$ a lánc éves összköltsége,
- $\Pi(ns)$ a lánc által elért éves nyereség.

Mivel a létező vállalatok már megépültek, az elhelyezésük már aligha módosítható. Viszont valószínűleg a minőség $\tilde{\alpha}_j$ -ről bizonyos szintig $\alpha_{\max}^j \leq \alpha_{\max}$ -ig javítható. Az új vállalat egy f_0 helyszínen α_0 minőséggel történő megnyitásának $G(f_0, \alpha_0)$ éves költségén kívül (amennyiben megnyitjuk), most néhány

további költséget is figyelembe kell venni, nevezetesen $A_j, C_j, V_j(\alpha_j), j = 1, \dots, k$, és $R_j(\alpha_j), j = 0, \dots, k$, a fentiek szerint. Az előző jelöléssel a bővülő lánc éves költsége a következőképpen adódik,

$$T(ns) = \sum_{j=1}^k (y_j(A_j + R_j(\alpha_j) + V_j(\alpha_j)) + (1 - y_j)C_j) + y_0(G(f_0, \alpha_0) + R_0(\alpha_0)), \quad (1.7)$$

és az általa megszerzett piaci részesedés, a valószínűségi vásárlói kiválasztási szabály alkalmazásával

$$M(ns) = \sum_{i \in I} w_i \frac{y_0 u_{i0}(f_0, \alpha_0) + \sum_{j \in J_1} y_j u_{ij}(\alpha_j)}{y_0 u_{i0}(f_0, \alpha_0) + \sum_{j \in J_1} y_j u_{ij}(\alpha_j) + \sum_{j \in J \setminus J_1} \tilde{u}_{ij}}. \quad (1.8)$$

A profit-maximalizálási probléma (P) a következőképpen fogalmazható meg:

$$\max \quad \Pi(ns) = F(M(ns)) - T(ns) \quad (1.9)$$

$$\text{f.h.} \quad f_0 \in S \subset \mathbb{R}^2 \quad (1.10)$$

$$d_i(f_0) \geq d_i^{\min}, \quad i \in I \quad (1.11)$$

$$y_0 \alpha_{\min} \leq \alpha_0 \leq y_0 \alpha_{\max} \quad (1.12)$$

$$y_j \alpha_{\min} \leq \alpha_j \leq y_j \alpha_{\max}^j, \quad j \in J_1 \quad (1.13)$$

$$T(ns) \leq B \quad (1.14)$$

$$y_j \in \{0, 1\}, \quad j = 0, \dots, k. \quad (1.15)$$

Az f_j vállalat minőségének a jelenlegi $\tilde{\alpha}_j$ értékről α_j -ra történő változásából eredő költségfüggvénynek, $V_j(\alpha_j)$, nem növekedőnek kell lennie a $[\alpha_{\min}, \tilde{\alpha}_j]$ és nem csökkenőnek a $(\tilde{\alpha}_j, \alpha_{\max}^j]$ tartományban, mivel minél nagyobb a $|\alpha_j - \tilde{\alpha}_j|$ különbség, annál nagyobb a módosításokhoz szükséges beruházás. Továbbá azt várnánk, hogy $V_j(\tilde{\alpha}_j + \alpha_j) > V_j(\tilde{\alpha}_j - \alpha_j)$, mivel a minőség javítása drágább, mint a minőségromlás. Elvárjuk, hogy $V_j(\alpha_j)$ konvex legyen $(\tilde{\alpha}_j, \alpha_{\max}^j]$ -ban (és $[\alpha_{\min}, \tilde{\alpha}_j]$ -ban is), mivel minél nagyobb minőséget várunk a vállalatától, annál magasabbak lesznek ennek a költségei, növekvő mértékben. Ráadásul minden változtatás esetén $v_j > 0$ fix költséget ki kell fizetni (általában a minőség változtatása a vállalat ideiglenes bezárását igényli). Ez a fix költség az, ami megakadályozza a túlságosan kis változtatásokat; valójában egy magas v_j érték megakadályozna minden változtatást. A $V_j(\alpha_j)$ minőség változtatásának költségfüggvénye a következő lehet:

$$V_j(\alpha_j) = \begin{cases} \frac{1}{\delta_j} (G_2(2\tilde{\alpha}_j - \alpha_j) - G_2(\tilde{\alpha}_j)) + v_j & \text{ha } \alpha_j < \tilde{\alpha}_j \\ 0 & \text{ha } \alpha_j = \tilde{\alpha}_j \\ G_2(\alpha_j) - G_2(\tilde{\alpha}_j) + v_j & \text{ha } \alpha_j > \tilde{\alpha}_j. \end{cases}$$

A $\delta_j > 0$ paraméter határozza meg, hogy mennyivel kerül kevesebbe a minőség csökkentése, mint a növelése. Ami az $R_j(\alpha_j)$ üzemeltetési költségfüggvényt illeti, annak nem csökkenőnek kell lennie, bár funkcionális formája az adott vállalatípustól függően változhat. Ebben a kutatásban lineáris formát feltételezünk, $R_j(\alpha_j) = o_j \alpha_j$, adott konstans $o_j > 0$ értékkel.

A (P) probléma egy MINLP, ezért az optimalizálás szempontjából kihívást jelent, és globális optimalizálási eszközökre van szükség, hogy megbirkózzunk vele.

1.2.1. MINLP-re kiterjesztett intervallumos B&B algoritmus

A globális optimalizálásban használt B&B módszerek hasonlóak a konvex MINLP-problémákhoz használt nemlineáris B&B módszerekhez (lásd például [5] 3.1. szakaszát). Azonban a konvex MINLP relaxált feladatai konvex NLP-problémák, amelyek lokális technikákkal megoldhatók; így a konvex nemlineáris B&B-ben a szétválasztás csak az egészértékű változóknál történik. Ezzel szemben a nemkonvex MINLP relaxációja nemkonvex NLP, amelyek maguk is nehezen megoldható problémák; így a megoldásukhoz a B&B eljárás részeként az elágaztatás a folytonos változóknál is szükséges.

Ebben a munkában a fő eredmény, hogy adunk egy intervallum aritmetikán alapuló B&B módszert, ami egy általános MINLP-t és így a fenti feladatot is megoldja. A valós számokkal dolgozó valós analízissel ellentétben az intervallum analízis kompakt intervallumokkal dolgozik. Egyik fő előnye, hogy a tulajdonságok okos felhasználásával lehetővé teszi a korlátok automatikus kiszámítását, ami különösen hasznos a B&B módszerek tervezésénél. Az alapötlet az, hogy azzal a relaxált problémával dolgozunk, amelyet akkor kapunk, ha feltételezzük, hogy az egészértékű változók is folytonosak, majd a keresési

tér egyes részeinek elutasítása előtt gondoskodunk arról, hogy ne távolítsunk el olyan részeket, amelyek egészértékű optimumot tartalmazhatnak.

Az intervallum analízis Kearfott [26] által javasolt szabványos jelölését használjuk. Az intervallumokat félkövré betűkkel, azok alsó és felső határait pedig „aláhúzással”, illetve „föléhúzással” jelöljük. Egy $z = [\underline{z}, \bar{z}]$ intervallum szélességét a $\text{wid}(z) = \bar{z} - \underline{z}$, jelöli, míg egy $z = (z_1, \dots, z_n)^T$ intervallumvektor szélességét a $\text{wid}(z) = \max\{\text{wid}(z_i) : i = 1, \dots, n\}$ adja. A z intervallum középpontját a $\text{mid}(z) = (\underline{z} + \bar{z})/2$ jelöli. \mathbb{I} és \mathbb{I}^n az intervallumok, illetve n dimenziós intervallumok halmazát jelöli.

Az intervallumos B&B algoritmusok fő eszköze a befoglalófüggvény.

1.2.1. definíció. Egy $f : \mathbb{R}^n \rightarrow \mathbb{R}$ valós függvény befoglalófüggvénynek nevezzük az $f : \mathbb{I}^n \rightarrow \mathbb{I}$ intervallum értékű függvényt, ha $\{f(z) : z \in z\} \subseteq f(z)$ igaz minden f értelmezési tartományán belüli z intervallumra.

Az f befoglalófüggvény fő előnye, hogy közvetlenül kaphatunk korlátokat f -re bármely intervallumon (az értelmezési tartományon belül).

A bevezetett cégbővítési probléma további nehézsége, hogy a $V_j(\alpha_j)$ függvényeket esetszétválasztással határoztuk meg. Mégis tudunk a V_j függvényre és annak deriváltjára egy befoglalófüggvényt adni a következőképpen:

$$V_j(\alpha_j) = \begin{cases} \frac{1}{\delta_j}(G_2(2\bar{\alpha}_j - \alpha_j) - G_2(\bar{\alpha}_j)) + v_j & \text{ha } \bar{\alpha}_j < \bar{\alpha}_j \\ G_2(\alpha_j) - G_2(\bar{\alpha}_j) + v_j & \text{ha } \underline{\alpha}_j > \bar{\alpha}_j \\ [0, \max\{\frac{1}{\delta_j}(G_2(2\bar{\alpha}_j - \underline{\alpha}_j) - G_2(\bar{\alpha}_j)), \\ G_2(\bar{\alpha}_j) - G_2(\bar{\alpha}_j)\} + v_j] & \text{ha } \underline{\alpha}_j < \bar{\alpha}_j < \bar{\alpha}_j \end{cases}$$

valamint

$$V'_j(\alpha_j) = \begin{cases} -\frac{1}{\delta_j}G'_2(2\bar{\alpha}_j - \alpha_j) & \text{ha } \bar{\alpha}_j < \bar{\alpha}_j \\ G'_2(\alpha_j) & \text{ha } \underline{\alpha}_j > \bar{\alpha}_j \\ (-\infty, +\infty) & \text{ha } \underline{\alpha}_j < \bar{\alpha}_j < \bar{\alpha}_j, \end{cases}$$

ahol G_2 és G'_2 a G_2 és G'_2 befoglalófüggvényei (lásd még [25, 32]).

Ezután ismertetjük az általunk kidolgozott intervallumos B&B módszer főbb új lépéseit. Az általános jelölés érdekében z -vel jelöljük a változók n hosszú vektorát, f -el a maximalizálandó célfüggvényt, $g_j(z) \leq 0, j = 1, \dots, r$ pedig a probléma feltételeit adja. Esetünkben z az ns vektor (dimenziója $n = 2k + 4$), és az (1.10)–(1.14) feltételeket megfelelő alakúra alakítva megkaphatjuk a $g_j(z)$ függvényeket.

1.2.1.1. A felosztási szabály

Legyen z a felosztandó intervallum. Amikor egy i koordinátairányt választunk ki felosztásra, akkor azt merőlegesen vágjuk ketté a következőképpen:

- Ha a z_i változó folytonos, akkor a megfelelő részintervallumok i -edik komponense $[z_i, \text{mid}(z_i)]$, illetve $[\text{mid}(z_i), \bar{z}_i]$ lesz.
- Ha azonban z_i egész változó, akkor ezek $[z_i, \lfloor \text{mid}(z_i) \rfloor]$ és $[\lceil \text{mid}(z_i) \rceil, \bar{z}_i]$ lesznek; vegyük észre, hogy ha $\text{mid}(z_i)$ nem egész szám, a $(\lfloor \text{mid}(z_i) \rfloor, \lceil \text{mid}(z_i) \rceil)$ nyílt intervallum pontjai nem egész számok, és ezért nem lehetségesek. Így tehát biztosított, hogy az egészértékű változók határai minden intervallumra mindig egész számok.

A részintervallumok többi összetevője nem változik a felosztás végrehajtásakor. A vágandó koordinátairányok kiválasztása két lépésben történik. Először mindig az egészértékű változók közül választunk, és ha nincsenek egészértékű komponensek, amelyeket fel lehet osztani, akkor a folytonos koordináták közül választunk. A második lépésben a kiválasztás az intervallum szélessége szerint történik, azaz a legszélesebb komponenst választjuk ki a kettévágásra.

Az elhelyezési problémánkban az összes egész változó bináris, így az intervallum a bináris változó mentén történő felosztásakor az előző folyamat az részintervallumban a bináris változót 0-ra, illetve 1-re rögzíti.

Az elhelyezési problémánkban a következő eljárásokat is elvégezzük:

- (1) Az $y_j = 0$ részintervallumban az α_j változót is 0-ra állítjuk (a j -edik vállalatot bezárjuk vagy nem nyitunk újat), így szélessége 0 lesz, és a továbbiakban nem választjuk ki a felosztási szabályban; analóg módon, ha $y_0 = 0$, akkor az f_0^1 és f_0^2 változókat is 0-ra állítjuk.
- (2) Amikor egy α_j változót $j = 1, \dots, k$, az *első alkalommal* választjuk ki felosztásra, ahelyett, hogy az intervallumot a fent leírtak szerint a középpontjánál feleznénk, a $\bar{\alpha}_j$ mentén három részre osztjuk, létrehozva a következő három részintervallumot: $[\alpha_j, \bar{\alpha}_j - EPS]$, $[\bar{\alpha}_j, \bar{\alpha}_j]$ és $[\bar{\alpha}_j + EPS, \bar{\alpha}_j]$, ahol EPS a számítógép által kezelhető legkisebb pozitív szám. Ez a V_j függvények miatt bizonyult hasznosnak. Természetesen ez csak a változó első kiválasztásakor történik, utána a kettévágást használjuk.
- (3) Egy intervallum mindig két koordinátáirányra merőlegesen van felosztva. Ezért mindig tetraszekciót végzünk, a szakirodalomban általánosan használt felezés helyett.
Az eljárás miatt sok különleges esetre kell ügyelnünk. Például, az y_j és α_j osztásakor, ha α_j -t először választjuk ki felosztásra, csak 4 új részintervallum keletkezik: $y_j = 1$ esetén α_j a fent leírtak szerint 3 részre van osztva, de $y_j = 0$ esetén csak $\alpha_j = 0$ eset lehetséges.

1.2.1.2. Kivágási tesztek

A legtöbb kivágási tesztet módosítani kell az egész számú változók kezeléséhez. A következőkben ezeket a teszteket ismertetjük, szükség esetén a módosításokkal együtt. Mint már említettük, a probléma korlátait általánosan $g_j(z) \leq 0, j = 1, \dots, r$ jelöli.

Fízibilitási teszt

Azt mondjuk, hogy a $g_j(z) \leq 0$ feltétel *biztosan folytonosan teljesül* egy z intervallumra, ha $\bar{g}_j(z) \leq 0$, illetve hogy *biztosan folytonosan nem teljesül*, ha $\underline{g}_j(z) > 0$. Abban az esetben, ha $0 \in g_j(z)$, nem tudjuk garantálni sem azt, hogy a feltétel biztosan folytonosan teljesül, sem azt, hogy biztosan folytonosan nem teljesül, ezért a feltételt ilyenkor *meghatározatlannak* nevezzük.

Egy z intervallumot *biztosan folytonosan fízibilisnek* nevezünk, ha az összes $g_j(z) \leq 0, j = 1, \dots, r$ feltétel biztosan folytonosan teljesül, és *biztosan folytonosan nem fízibilisnek*, ha létezik legalább egy olyan feltétel, amely biztosan folytonosan nem teljesül. Minden más esetben, amikor legalább egy feltétel meghatározatlan, de egyik feltétel sem biztosan folytonosan nem teljesül, a z intervallum *folytonosan meghatározatlan*.

Továbbá, ha az összes folytonos feltétel szigorúan teljesül, azaz $\bar{g}_j(z) < 0, j = 1, \dots, r$, akkor egy z intervallumot *szigorúan biztosan folytonosan fízibilisnek* nevezünk. Csak a biztosan folytonosan nem fízibilis intervallumokat tudjuk kizárni, így a fízibilitási teszt ezeket az intervallumokat elveti.

Megjegyezzük, hogy egy z intervallum esetében a $g_j(z) \leq 0, j = 1, \dots, r$ folytonos korlátok minden $z \in z$ pontra teljesülnek. Ezért a z minden egészértékű pontja biztosan fízibilis, például a z sarokpontjai, amelyek a felosztási szabály szerint mindig egészértékű pontok. Ez különösen hasznos az algoritmus által talált legjobb érték frissítéséhez, amelyet a levágási tesztben használunk.

Monotonitási teszt

A monotonitás segíthet kizárni azokat az intervallumokat, amelyek nem tartalmazhatnak globális optimumot a belsejükben. Ez alkalmazható mind a biztosan folytonosan fízibilis, mind a meghatározatlan intervallumokra.

Legyen $\nabla f(z) = (\nabla_1 f(z), \dots, \nabla_n f(z))^T$ az f célfüggvény gradiensének a z intervallumra vett befejezése. A monotonitási teszt akkor alkalmazható, ha létezik olyan z_i változó, amelyre a célfüggvény monoton, azaz $0 \notin \nabla_i f(z)$. Jelöljük ∂S -sel az S keresési tér határát (ahol a $g_j(z) \leq 0$ feltételeket nem vesszük figyelembe). A problémánkban $S = (S, [0, \alpha_{\max}], [0, 1], \dots, [0, \alpha_{\max}], [0, 1])$.

Folytonos, feltétel nélküli feladatok esetén a z' ($z' = (z_1, \dots, z_i, \dots, z_n)$) vagy $z' = (z_1, \dots, \bar{z}_i, \dots, z_n)$, attól függően, hogy a függvény csökkenő vagy növekvő) tartalmazza a maximumot a z intervallumban. Ha $z' \not\subset \partial S$, akkor van egy másik folytonos intervallum, aminek része z' , így a z intervallum kizárható. Ha $z' \subset \partial S$, tehát z' a határon van, akkor a z intervallum z' -re szűkíthető.

Egész változókkal rendelkező, feltételes problémák esetén azonban a z' -t tovább kell ellenőrizni. Ha egy biztosan folytonosan fízibilis z intervallumra létezik egy z_i változó amire $0 \notin \nabla_i f(z)$, akkor a következők érvényesek:

1. Ha z_i folytonos, akkor tudjuk, hogy a maximum nem lehet a belsejében, így a z intervallum vagy kizárható (ha szigorúan folytonosan fízibilis), vagy leszűkíthető a z' lapjára, ha $z' \subset \partial S$.

2. Ha z_i egész, akkor az intervallum leszűkíthető z' -re.

Vegyük észre, hogy ha a z_i változó egész, az intervallum nem távolítható el, mivel a gradiens előjelét a $(z_1, \dots, (z_i - 1, z_i), \dots, z_n)$, illetve $(z_1, \dots, (z_i, z_i + 1), \dots, z_n)$ nyitott tartományban megváltoztathatja. Mivel az osztási szabályban az ilyen, nem egészértékű pontokat tartalmazó tartományokat eltávolítjuk, ezt értelmezhetjük úgy, hogy z' a fizibilis halmaz határán van.

Tekintsünk most egy meghatározatlan intervallumot, z -t. Attól függően, hogy melyik feltételek nem biztosan folytonosan teljesülnek z -re, a monotonitási teszt kimenetele más és más lehet. Általánosságban ebben az esetben az a fontos, hogy a feltételek amelyekben z_i monoton változó megjelenik, biztosan folytonosan teljesüljenek z' -n, amely tartalmazhatja az intervallum maximumát. Az ilyen z' lapot *fizibilisnek* nevezzük z_i -re. Ha $0 \notin \nabla f(z)$ egy z_i változóra, akkor a következőket állíthatjuk:

1. Ha z_i egy folytonos változó, és a z' lap fizibilis z_i -re, akkor a z intervallum vagy kizárható (ha $z' \not\subset \partial \mathcal{S}$ -ban), vagy leszűkíthető z' -re (ha $z' \subset \partial \mathcal{S}$ -ban).
2. Ha z_i egy egész változó, és z' fizibilis z_i -re nézve, akkor az intervallum leszűkíthető z' -re.

Konkrétan a mi elhelyezési problémánk esetében, mivel az $y_j, \alpha_j, j = 0, \dots, k$ változók nem szerepelnek az új vállalat elhelyezésének fizibilis területét meghatározó feltételekben, elegendő annak ellenőrzése, hogy a megfelelő intervallumok biztosan folytonosan kielégítik-e a költségvetési feltételt. Az f_0^1, f_0^2 elhelyezési változók esetében azonban az összes feltételt ellenőrizni kell, mivel azok az összes feltételben megjelennek.

1.2.1.3. Vetített egydimenziós Newton-módszer

A [17] cikkben egy egydimenziós Newton-módszert adtunk meg, amely itt mind a biztosan folytonosan szigorúan fizibilis, mind a meghatározatlan intervallumokra, illetve azok folytonos minőségi változóira alkalmazhatunk.

Ha az intervallum folytonosan szigorúan fizibilis, akkor a módszert változtatás nélkül alkalmazhatjuk. A módszer lényege, hogy az intervallumos Newton-módszert egyszerre csak egy változót figyelembe véve futtatjuk, a többi változót az aktuális intervallumértékükön rögzítve. Vegyük észre, hogy a klasszikus intervallumos Newton-módszerrel ellentétben, amelyet általában egylépéses (vagy egyiterációs) eljárás-ként hajtunk végre, a vetített Newton-módszert addig futtatjuk, amíg az intervallum egyetlen részét sem tudjuk eltávolítani, vagy amíg az intervallumot el nem vetjük teljesen (mivel nem szerepel benne lokális optimum), vagy amíg az adott változó szélessége nem lesz kisebb, mint a megállási kritériumunkban szereplő tolerancia.

Tekintsünk most egy meghatározatlan intervallumot, z -t. Alkalmazhatjuk a vetített egydimenziós Newton-módszert a z egy z_i változójára, feltéve, hogy z_i nem szerepel egyetlen olyan feltételben sem, amelyre z nem biztosan folytonosan szigorúan fizibilis.

Azt az empirikus szabályt is alkalmaztuk, hogy a módszert csak olyan intervallumokra használjuk, amelyek szélessége kisebb, mint egy.

A bevezetett elhelyezési problémában, ha a z intervallum meghatározatlan bizonyos elhelyezési feltételekre, de szigorúan fizibilis a költségvetési feltételre, a teszt alkalmazható a minőségi változókra (a célfüggvény konkáv ezekben a változóknak [17]).

1.2.2. Eredmények

Felírtunk egy olyan új valóságközeli modellt, ami egy vállalatlánc bővítési törekvéseit szolgálja. A cég dönthet, hogy egy új vállalatot megnyit, a meglévő vállalatai minőségét felfelé vagy lefelé változtatja, esetleg néhány meglévő vállalatát bezárja (hogy az itt megspórolt pénzt a lánc tulajdonában lévő többi vállalatba vagy az új vállalatba fektesse, amennyiben az(ok) nyitva van(nak)), vagy mindezen lehetőségek egy kombinációját valósítja meg.

Az így kapott modell egy nehezen megoldható MINLP-probléma, amelynél a létező megoldók kudarcot vallanak: csak a BARON [35] és a LocSol [6] képes megoldani néhány kisebb problémát, de nagyobb esetek megoldását vagy nem találták meg, vagy túl sokáig futottak.

Kifejlesztettünk a probléma megoldására egy intervallumos B&B módszert. Ez a módszer a folytonos NLP-problémákhoz javasolt klasszikus intervallumos B&B módszerek egészértékű változók kezelésére szolgáló módosítása. Különösen a felosztási szabályt módosítottuk úgy, hogy figyelembe vegye az egészértékű változókat, valamint a monotonitási tesztet, és a vetített egydimenziós Newton-módszert. E tesztek módosításának lényege, hogy az egész változókat relaxáljuk, és amikor elérkezik a részterületek

kizárásának pillanata, az egészértékűséget újra figyelembe vesszük, hogy ne távolítsuk el a lehetséges optimumokat.

A kifejlesztett intervallumos B&B algoritmus úgynevezett Newton-változata közepes méretű problémákat nehézség nélkül, átlagosan körülbelül 30 perc CPU-idő alatt képes megoldani. A módszer garantáltan megtalálja az összes optimális megoldást.

1.3. Egy versenyző vállalat hely-ár egyensúlyai

A vállalatok optimális helyének problémája és a termék megfelelő árának meghatározása a két fő mozgatórugó az ellátási láncokban. Ez a két, egymással összefüggő logisztikai tényező versenyelőnyt biztosít, miközben jelentősen hozzájárul az ellátási lánc hatékonyságához és reagálóképességéhez. A profit maximalizálása az egyes versenytárs cégek számára egy hely-ár játéknak tekinthető, amelyet már többször is tanulmányoztak Hotelling [24] cikket követően. A legtöbb létező szakirodalom lineáris piaccal foglalkozik (lásd [14, 19, 30]), ami részben a kapcsolódó elhelyezési problémák megoldásának bonyolultságából adódik.

Az áregyensúly megléte a játék második szakaszában többek között az elfogadott árpolitikától függ. Amikor minden cég fix gyári árat határoz meg, és a vevő gondoskodik a fuvarozásról (*gyári árpolitika*), akkor ritkán létezik áregyensúly (lásd [21]). Ebben az esetben a kapcsolódó elhelyezési problémát nemlineáris elhelyezési terekben tanulmányozták az árakat paraméterként véve (lásd [16, 22, 37, 41]). Másrészt mindig létezik egy áregyensúly, amikor minden egyes vállalat minden egyes piaci területen egy adott árat számít fel, amely magában foglalja a szállítási költségeket is (*szállított árpolitika*).

Azzal a feltételezéssel élünk, hogy a piacokat n keresleti pontban aggregáljuk, ahol egy adott homogén ár-elasztikus terméket értékesítenek a konkurens cégek. A cégek gyártják és szállítják a terméket a vevőknek, akik mindig attól a cégtől vásárolnak, amelyik a legalacsonyabb árat kínálja.

Nevezetesen két olyan céget tekintünk, amelyek egy-egy üzletet helyeznek el a folytonos térben, állandó termelési határkölségek mellett. Ebben a játékban az egyes játékosok (cégek) számára rendelkezésre álló lépések száma végtelen sok, mivel a vállalatok elhelyezése a síkon (illetve annak egy részhalmazán) történik, és az árak is folytonosan választhatók. Egy cég számára a kifizetési függvény az általa elért nyereség, amelyet az egyes cégek által kiszolgált piaci szegmens határoz meg.

1.3.1. A feladat redukálása egy elhelyezési játékká

Először megmutatjuk, hogy létezik áregyensúly, majd az egyensúlyi árak felhasználásával, a hely-ár játékot elhelyezési játékká redukáljuk.

Van n vásárlónk, akiknek a kereslete w_i ($i = 1, \dots, n$), és két konkurens cégnek ($u = 1, 2$) kell kiszolgáltatnia. Mindkét u cég a sík bármely pontján elhelyezheti a vállalatait. Azonban, először arra az általános esetre utalunk, amikor a keresési tér egy tetszőleges L^u elhelyezési tér. L^u lehet egy diszkrét halmaz, egy hálózat vagy a sík egy régiója. Jelölje

$$p_{xi}^u \quad \text{a } u \text{ cég } x \text{ helyről } i \text{ piacra szállított határkölséget.}$$

Ez magában foglalja mind a szállítási, mind a termelési költségeket.

$$p_i^u(X^u) = \min\{p_{xi}^u : x \in X^u\}, \text{ a minimum ár, amelyet a } u \text{ cég a } i \text{ piacon kínálhat, ha } X^u \subset L^u$$

részhalmazát választjuk a u cég telephelyeinek halmazaként, $u = 1, 2$.

Feltételezzük továbbá, hogy minden cég csak olyan árat kínál, amely vagy megegyezik a szállított határkölségével, vagy meghaladja azt, így garantálva a nyereséget. Megmutatható, hogy az áregyensúly létrejön, mivel a két vállalat közül egyik sem próbálja meg saját versenyelőnye miatt megváltoztatni az árát. Más szóval, miután X^1 és X^2 rögzített, az egyensúlyi árak matematikailag a következőképpen határozhatók meg:

$$\bar{p}_i^1(X^1, X^2) = \begin{cases} p_i^2(X^2) & \text{ha } p_i^1(X^1) < p_i^2(X^2) \\ p_i^1(X^1) & \text{máskülönben} \end{cases}$$

$$\bar{p}_i^2(X^1, X^2) = \begin{cases} p_i^1(X^1) & \text{ha } p_i^2(X^2) < p_i^1(X^1) \\ p_i^2(X^2) & \text{máskülönben.} \end{cases}$$

Hasonlóképpen, az egyes u cégek (vagy a piac u cég által kiszolgált szegmense) által megszerzett piaci

részesezés a következő:

$$\begin{aligned} M^1(X^1, X^2) &= \{i \in \{1, \dots, n\} : p_i^1(X^1) < p_i^2(X^2)\}, \\ M^2(X^1, X^2) &= \{i \in \{1, \dots, n\} : p_i^2(X^2) < p_i^1(X^1)\}. \end{aligned}$$

A cégek profitja a fentiek alapján

$$\begin{aligned} \Pi^1(X^1, X^2) &= \sum_{i \in M^1(X^1, X^2)} (p_i^2(X^2) - p_i^1(X^1)) w_i, \\ \Pi^2(X^1, X^2) &= \sum_{i \in M^2(X^1, X^2)} (p_i^1(X^1) - p_i^2(X^2)) w_i. \end{aligned}$$

Röviden, a hely-ár játék egy olyan hely-játékra redukálódik, amelyben a döntések a helyválasztásról szólnak, és az u játékos kifizetési függvénye $\Pi^u(X^1, X^2)$, $u = 1, 2$.

Szeretnénk hangsúlyozni, hogy jelen esetben nem a Π^1 vagy Π^2 optimalizálása a célunk, hanem az egyensúlyi helyzet megtalálása. Az $(\tilde{X}^1, \tilde{X}^2)$ vállalatok akkor teljesítik az egyensúlyi helyzet tulajdonságát, ha az alábbi két feltétel egyaránt teljesül:

$$\Pi^1(\tilde{X}_1, \tilde{X}_2) \geq \Pi^1(X^1, \tilde{X}^2) \quad \forall X^1 \quad \text{és} \quad \Pi^2(\tilde{X}_1, \tilde{X}_2) \geq \Pi^2(\tilde{X}^1, X^2) \quad \forall X^2.$$

Amint arra [31] rámutatott, a következő minimalizálási probléma bármely globális optimális megoldása egyensúlyi helyzetet jelent:

$$\min \{S(X^1, X^2) : X^1 \subset L^1, X^2 \subset L^2\}, \quad (P)$$

ahol

$$S(X^1, X^2) = \sum_{i=1}^n \min \{p_i^1(X^1), p_i^2(X^2)\} w_i$$

a *társadalmi költség* néven ismert. Megjegyzendő azonban, hogy a fordítottja általában nem érvényes. Ezenkívül a társadalmi költség lokális minimumai csak *lokális* egyensúlyokhoz vezetnek. Más szóval, léteznek egyensúlyok ugyanabban a szomszédságban, amelyben lokális minimalizálók.

Azt a problémát, ahol az L^u , $u = 1, 2$, potenciális helyek halmazai végesek, [31] megvizsgálta. Ha azonban L^u , $u = 1, 2$ folytonos, azaz a tér egy-egy régiója, a (P) probléma viszonylag nehezebben megoldhatóvá válik, mint a diszkrét eset megfelelője. Itt ezt a síkbeli versenyző elhelyezési problémát szeretnénk komplex és újszerű globális optimalizálási technikák segítségével megoldani.

Különösen arra az esetre koncentrálnak, amikor $L^u \subseteq \mathbb{R}^2$, $u = 1, 2$ és $|X^u| = 1$, $u = 1, 2$. Az egyszerűség kedvéért az x^u -t használjuk az u cég által választott vállalat helyének jelölésére, azaz $X^u = \{x^u\}$. Mivel a szállított határkölség felírható mint $p_{xi}^u = c_x^u + t_{xi}^u$, ahol c_x^u és t_{xi}^u a termelési, illetve a szállítási határkölséget jelenti, (P) átírható a következőre:

$$\min \left\{ \sum_{i=1}^n \min \{c_{x1}^1 + t_{x1i}^1, c_{x2}^2 + t_{x2i}^2\} w_i : x^1 \in L^1, x^2 \in L^2 \right\}. \quad (P')$$

Célunk a (P') probléma vizsgálata. A (P') megoldásának nehézségére rávilágítandó, tekintsük azt a speciális esetet, amikor $c_x^1 = c_x^2$ minden $x \in \mathbb{R}^2$ -ra, tehát megállapíthatjuk, hogy $p_{xi}^u = t_{xi}^u$. Tegyük fel továbbá, hogy az u cég szállítási határkölsége az u x^u helyéről az i piacra az i és x^u közötti $d_i(x^u)$ távolság lineáris függvénye, így feltételezhetjük, hogy $p_{xi}^u = d_i(x^u)$. Mivel $|X^u| = 1$, így $p_{xi}^u(x^u) = p_{x^u i}^u = d_i(x^u)$ is érvényes. E feltételezések alapján a társadalmi költség a következőképpen határozható meg:

$$S(x^1, x^2) = \sum_{i=1}^n \min \{d_i(x^1), d_i(x^2)\} w_i. \quad (1.16)$$

Ebben az esetben a (P') ezután a jól ismert, több forrásból álló, kapacitás nélküli Weber-problémára redukálódik, amelyről ismert, hogy NP-nehez (lásd Megiddo és Supowit [27]).

Mivel a (P') probléma célfüggvénye sem nem konvex, sem nem konkáv, globális optimalizálási technikákat kell alkalmazni. A kutatás egyik fő eredménye a váltakozó Weiszfeld-féle heurisztika.

1.3.2. A heurisztikus megközelítés

Először javaslunk egy váltakozó Weiszfeld-féle algoritmust a könnyebb eset megoldására, amikor a (P') probléma a több forrásból álló Weber-problémára redukálódik (az (1.16) minimalizálása). Ezután kiterjesztjük ezt a megközelítést az általános költségszerkezetre.

1.3.2.1. A váltakozó Weiszfeld-féle elhelyezési algoritmus

A váltakozó Weiszfeld-féle elhelyezési algoritmust, amit alább ismertetünk, röviden AWLA-nak hívjuk (lásd az 1.1. algoritmust). Minden egyes iterációban csak azokat a vevőket vesszük figyelembe, amelyekből a megfelelő cég pozitív nyereséget kap (lásd a 3. és 5. lépést). Miután a keresleti pontok egy vállalathoz való hozzárendelése befejeződött, a vállalat áthelyezése a megfelelő 1-medián probléma megoldásával történik az 1.1. algoritmus 4. és 6. lépésének felhasználásával. Ezeket a lépéseket Weiszfeld [39] ötlete alapján hajtuk végre, de az általunk kifejlesztett új iteratív képletet használjuk, lásd pl. az (1.18) egyenletet. Ezt a folyamatot addig ismételjük, amíg a vállalatok helyének változása egy adott tűréshatárnál kisebb nem lesz. Ez az algoritmus egyik iterációról a másikra javítja a célfüggvény értékét, amíg egy lokális minimumot nem talál, és így a keresés véget ér.

1.1. algoritmus: AWLA

- 1: Definiáljuk a két kezdőpontot x^1 és x^2 és állítsuk be a STOP=false értéket.
 - 2: **repeat**
 - 3: Számítsuk ki az $M^1(x^1, x^2) = \{i \in \{1, \dots, n\} : d_i(x^1) < d_i(x^2)\}$ halmazt.
 - 4: Alkalmazzuk a Weiszfeld-féle algoritmust az 1. cég 1-medián problémájának megoldására a $M^1(x^1, x^2)$ halmaz alapján. Tekintsük x^1 -t kiindulási pontnak, és legyen \hat{x}^1 az újonnan kapott hely.
 - 5: Számítsuk ki a $M^2(\hat{x}^1, x^2) = \{i \in \{1, \dots, n\} : d_i(x^2) < d_i(\hat{x}^1)\}$ halmazt.
 - 6: Alkalmazzuk a Weiszfeld-féle algoritmust a 2. cég 1-medián problémájának megoldására a $M^2(\hat{x}^1, x^2)$ halmaz alapján. Tekintsük x^2 -t kiindulási pontnak, és legyen \hat{x}^2 az újonnan kapott hely.
 - 7: **if** $\max(\|\hat{x}^1 - x^1\|, \|\hat{x}^2 - x^2\|) < \varepsilon$
 - 8: | STOP=true
 - 9: **else**
 - 10: | legyen $x^1 = \hat{x}^1$ és $x^2 = \hat{x}^2$
 - 11: **until** STOP
-

Az 1. lépésben a vállalatok kezdeti helyszíneinek kiválasztására többféle stratégia létezik. A legegyszerűbb módszer, hogy az x^1 és x^2 értékeket véletlenszerűen generáljuk a vizsgált területről.

1.3.2.2. Az általános költségszerkezet

Hasonló stratégia alkalmazható az általános eset kezelésére is, amikor $c_x^1 \neq c_x^2$ és/vagy $t_{xi}^1 \neq t_{xi}^2$. Az egyetlen különbség az, hogy a 4. és 6. lépésben megoldandó 1-medián problémák célfüggvénye nem a klasszikus Weber-problémáé. Ebben az alfejezetben ezt a kérdést vizsgáljuk meg.

Jelöljük a keresleti pontok helyét $a_i, i = 1, \dots, n$ -vel. Szükségünk van a c_x^u , az u cég marginális termelési költségének matematikai megfogalmazására, amikor a vállalat x helyen található. Feltételezve, hogy a működési költségek a választott helyszíntől függetlenül azonosak, akkor a c_x^u csak az x helyszíntől függ. Ez a költség általában a vállalatnak a keresleti pontokhoz viszonyított relatív elhelyezkedéséhez kapcsolódik a távolságokon keresztül. A c_x^u költség általában növekszik, ahogy x közelít bármely keresleti ponthoz. Ennek oka, hogy a nagy népességkoncentrációjú központokban a vállalat működési költségei a földterület és a telephelyek értéke miatt magasabbak. A [17] cikkben a c_x^u -ra megfelelő kifejezéseket javaslunk, beleértve a következő formulát, amelyet használni fogunk:

$$c_x^u = \sum_{i=1}^n \frac{w_i}{W} \frac{1}{d^u(x, a_i)^{\phi_{i0}^u} + \phi_{i1}^u}, \quad (1.17)$$

ahol $W = \sum_{i=1}^n w_i$ az összkereslet, és $\phi_{i0}^u, \phi_{i1}^u > 0$ olyan paraméterek, amelyek értéke adott vagy becsült. Az első a távolságot modulálja, míg a második a nullával való osztás elkerülése mellett (azaz $x = a_i$ esetén) teszi lehetővé a kiigazítást.

Az (1.17) felhasználásával az AWLA algoritmus 4. és 6. lépéseiben megoldandó megfelelő „minisum”-problémák a következő célfüggvénnyel rendelkeznek:

$$\min F^u(x) = \sum_{i \in M^u} w_i (c_x^u + t_{xi}^u) = Q^u c_x^u + \sum_{i \in M^u} w_i g^u(d^u(x, a_i)),$$

ahol M^u az $u = 1$ esetén $M^1(x^1, x^2)$, $u = 2$ esetén $M^2(\hat{x}^1, x^2)$, $Q^u = \sum_{i \in M^u} w_i$ a u cég által szállított piaci részesedést adja meg, $g^u(d^u(x, a_i))$ pedig a szállítási költséget jelenti a távolság függvényében.

Egy Weiszfeld-féle eljárást dolgozunk ki ezen új „minisum”-problémák megoldására. Az optimalitási feltételből kapjuk, hogy

$$\frac{\partial F^u}{\partial x_1} = Q^u \sum_{i=1}^n \frac{w_i}{W} \frac{-\phi_{i0}^u (d^u(x, a_i))^{\phi_{i0}^u - 1}}{((d^u(x, a_i))^{\phi_{i0}^u} + \phi_{i1}^u)^2} \frac{\partial d^u(x, a_i)}{\partial x_1} + \sum_{i \in M^u} w_i \frac{dg}{dd^u(x, a_i)} \frac{\partial d^u(x, a_i)}{\partial x_1} = 0.$$

Ha a távolságfüggvény felírható mint

$$\frac{\partial d^u(x, a_i)}{\partial x_1} = x_1 A_{i1}^u(x) - B_{i1}^u(x),$$

azt kapjuk, hogy

$$x_1 = \frac{Q^u \sum_{i=1}^n B_{i1}^u(x) \frac{w_i}{W} \frac{-\phi_{i0}^u (d^u(x, a_i))^{\phi_{i0}^u - 1}}{((d^u(x, a_i))^{\phi_{i0}^u} + \phi_{i1}^u)^2} + \sum_{i \in M^u} B_{i1}^u(x) w_i \frac{dg}{dd^u(x, a_i)}}{Q^u \sum_{i=1}^n A_{i1}^u(x) \frac{w_i}{W} \frac{-\phi_{i0}^u (d^u(x, a_i))^{\phi_{i0}^u - 1}}{((d^u(x, a_i))^{\phi_{i0}^u} + \phi_{i1}^u)^2} + \sum_{i \in M^u} A_{i1}^u(x) w_i \frac{dg}{dd^u(x, a_i)}}. \quad (1.18)$$

Az x_2 levezetése hasonló. A kapott egyenletek segítségével rekurzív módon frissítjük x_1 és x_2 értékeit, ami a Weiszfeld-féle módszerhez vezet.

1.3.1. állítás. *Az AWLA módszer minden egyes lépése nem növekvő a több forrásból álló Weber-probléma esetén.*

Kiemelnék két számítási fejlesztést a keresés felgyorsítása és irányítása érdekében.

(i) Annak elkerülése, hogy az új vállalatok a keresleti pontokon helyezkedjenek el, a vevői helyek körül nem megengedett régiókat vezetünk be. Ezeket kis tiltott körök határozzák meg, amelyek középpontja egy-egy vevőhelyen helyezkedik el, kis sugárral, ami arányos a vevői kereslettel és fordítottan arányos a vevők számának négyzetgyökével. Ezek a tiltott körök az (1.18) egyenletben a nullával való osztás problémáját is elkerülik, miközben garantálják, hogy mind a monotonitási, mind a metszési tesztek jól definiáltak maradnak az intervallumos B&B módszerben. Megjegyezzük, hogy az egyszerűség kedvéért szükség esetén egy kis konstans értéket is használhatunk minden sugárra. A klasszikus Weiszfeld algoritmus az 1-medián problémára alkalmazva nem garantáltan konvergál, ha az iterációban egy részmegoldás az egyik (nem optimális) keresleti pontba esik.

(ii) Ezen kívül, ha az egyik iterációról a másikra az új pont, mondjuk az A pont a B pontba kerül, amely történetesen egy tiltott körön belül van, akkor az A -n és B -n áthaladó egyenes és a kör két metszéspontját értékeljük ki, és azt választjuk ki új helyként, amelyik a kisebb költségű megoldást adja.

1.3.3. Eredmények

A hely-ár játék egyensúlyi pontjának meghatározásához egy a síkban történő elhelyezési feladat megoldását vizsgáljuk, ahol két egymással versengő cég egy-egy üzletet helyez el. A probléma megoldására egy általános költségfüggvény mellett egy intervallumos B&B módszert alkalmaztunk, amely azonban csak kis méretű esetekre bizonyul alkalmasnak. A két vállalat koordinátáinak frissítésére egy újonnan kifejlesztett iteratív képleten alapuló, váltakozó Weiszfeld-féle heurisztikát dolgoztunk ki, amely a nagyobb esetek megoldására is alkalmas. A számítási eredmények megmutatták, hogy a heurisztikus módszer hatékony, hiszen nagyon jó eredményeket ér el rövid idő alatt.

Speciális esetként megvizsgáltuk a 2-forrású Weber-probléma optimális megoldásainak és a (P') -nek a kapcsolatát is. Kísérleteket végeztünk meglévő adathalmazokon, ahol azt találtuk, hogy a 2-forrású Weber-probléma optimális megoldása nem feltétlenül marad optimális az új, (P') problémára.

2. fejezet

Szimplex-alapú B&B módszer

A globális optimalizálásban a korlátozás és szétválasztás (B&B) algoritmusok kimerítő keresést végeznek a fizibilis területen. Néhány speciális probléma esetén a fizibilis terület egy szimplex, például a portfólió optimalizálásban és a keverési problémákban, ahol a változók összegének egynek kell lennie. Emellett bármely konvex vagy konkáv poliéderrel megadott fizibilis halmaz esetén a szimplexekre való felbontás egyszerűen adódik [18]. Ilyen esetekben a természetes választás a szimplex, mint partíciós halmaz használata, ami egy szimplex-alapú B&B módszerre vezet.

2.1. Szimplex gömbökkel való fedése

Szimplexnek nevezzük h affin független csúcs, v_1, \dots, v_h konvex kombinációinak halmazát, azaz,

$$S = \{x \in \mathbb{R}^n : x = \sum_{j=1}^h \lambda_j v_j, \sum_{j=1}^h \lambda_j = 1; \lambda_j \in [0, 1], j = 1, \dots, h\}. \quad (2.1)$$

A csúcsokat az n dimenziós euklideszi térben határozzuk meg, ahol $n \geq h - 1$, és az euklideszi távolságot vesszük normaként. A probléma összetevői a szimplex mellett a csúcsok körüli r_i sugarú körök vagy gömbök,

$$B_i = \{x \in \mathbb{R}^n : \|x - v_i\| \leq r_i\}. \quad (2.2)$$

Most már megfogalmazható a Szimplex Fedési probléma.

Szimplex Fedési probléma (Simplex Cover – SC). Adott egy S szimplex és a csúcsaiban lévő B_1, \dots, B_h gömbök. Igazoljuk, hogy

$$\forall x \in S, \quad x \in \cup_i B_i \quad (2.3)$$

vagy másképpen, ellenőrizzük, hogy

$$\exists x \in S, \quad x \notin \cup_i B_i. \quad (2.4)$$

Az SC-problémának egy másik ekvivalens felírása is használható, amely a csúcsoktól való távolságra összpontosít. Ez közelebb áll a Laguerre-Voronoi diagramok koncepciójához (vagy más néven hatvány (power) diagramokhoz). Mi a négyzetes euklideszi távolságra összpontosítunk. Definiáljuk a következő függvényt:

$$\varphi(x) = \min_i \{\|x - v_i\|^2 - r_i^2\}, \quad (2.5)$$

ahol a négyzetes távolság additívan súlyozott [3]. Vegyük észre, hogy a φ függvény, amely szigorúan konvex függvények minimuma, nem konvex. Sőt, speciális esetben, amikor a B_i gömbök egyike sem fedti a többi $v_j, i \neq j$ csúcsot, a φ függvénynek minden egyes csúcsban $-r_i^2$ a lokális minimuma, azaz $\varphi(v_i) = -r_i^2$. Az SC probléma egyenértékű a következővel:

$$\forall x \in S \quad \varphi(x) \leq 0. \quad (2.6)$$

Az ekvivalens (2.6) probléma alapján érdekes az SCO (Szimplex Fedési Optimalizálás) vizsgálata:

$$\Phi := \max_{x \in S} \varphi(x). \quad (2.7)$$

Az SC probléma igazolt, ha $\Phi \leq 0$.

2.1.1. Távolságok, Voronoi-diagramok és szimplexek

A Voronoi-diagramok egy p_i ponthalmazra a teret $D(p_i)$ régiókra osztják, ahol a $D(p_i)$ összes pontja közelebb van p_i -hez, mint bármely p_j -hez, $j \neq i$. A Voronoi-diagramokkal kapcsolatban számos távolságfüggvény használható. Szokásos az euklideszi távolság figyelembevétele:

$$d_e(p_1, p_2) = \|p_1 - p_2\|_2. \quad (2.8)$$

A p_i Voronoi-cella az euklideszi távolság esetén a p_i -hez legközelebb eső pontok;

$$D(p_i) = \{x \in \mathbb{R}^n \mid d_e(x, p_i) \leq d_e(x, p_j), \quad j \neq i, j = 1, \dots, h\}. \quad (2.9)$$

Az SC problémában minden B_i gömbnek r_i sugara van. Ezt a sugarat tekinthetjük súlynak. A távolság súlyozásának kétféle módját szokták figyelembe venni, és az úgynevezett súlyozott pontoknál, p_i -nél lehet használni:

- a távolság súlyozása additív módon:

$$d_a(x, p_i) = d_e(x, p_i) - r_i, \quad (2.10)$$

- súlyozzuk az úgynevezett hatványtávolságot:

$$d_p(x, p_i) = d_e(x, p_i)^2 - r_i^2. \quad (2.11)$$

A d_a használata a Voronoi-fogalomkörben olyan diagramokhoz vezet, amelyek élei hiperbolikus görbék. Léteznek úgynevezett multiplikatív Voronoi-diagramok is, amelyek hiperbolikus görbékkel rendelkeznek [3].

A d_p használatával kicsit egyszerűbb az élet; az elválasztó alakzatok síkok. Kapunk egy hatványdiagramot hatványsíkokkal:

$$\Pi_{ij} = \{x \in \mathbb{R}^n \mid d_p(x, p_i) = d_p(x, p_j)\}, \quad (2.12)$$

és a hatványcellák

$$D_p(p_i) = \{x \in \mathbb{R}^n \mid d_p(x, p_i) \leq d_p(x, p_j), \quad j \neq i, j = 1, \dots, h\}. \quad (2.13)$$

A B_i gömbön kívüli x pontra $d_p(x, p_i)$ pozitív, és $\sqrt{d_p(x, p_i)}$ a gömbtől való távolságot jelenti. A *hatvány Voronoi-diagram* (vagy *hatvány diagram*) cellái konvex poliéderek. A különbség az euklideszi Voronoi-diagram és a hatványdiagram között az, hogy az utóbbiban egy cella lehet üres.

2.1.2. Simplex lefedése, elemzés a θ -pont alapján

A

$$\varphi(x) = \min_i \{d_p(x, v_i)\} \quad (2.14)$$

függvény minden egyes hatványcellában darabonként konvex. Emiatt a φ maximumát egy hatványcellában annak egy extrém pontján veszi fel. Ez azt jelenti, hogy a maximumot, Φ -t keresve a hatványdiagramok határait alkotó Π_{ij} hatványsíkokra összpontosíthatunk. Jelöljük a halmaz belsejét *int*-tel. A 2.1.1 lemma azt mondja ki, hogy az SCO megoldását a hatványcellák határán kell keresni.

2.1.1. lemma. Egy $x^* \in \operatorname{argmax}_{x \in S} \varphi(x) \notin \operatorname{int} S \cap \cup_i \operatorname{int} D_p(v_i)$.

Könnyen ellenőrizhető, hogy egy gömb lefedi-e a teljes szimplexet. Ebben az esetben a $\varphi(x)$ maximumát az S egy vagy több v_i csúcsában érhetjük el. Arra az esetre, ha a maximum nem egy csúcsban van, a lemma azt mutatja, hogy a Π_{ij} hatványsíkokra koncentrálhatunk. Tulajdonképpen azt az információt adja, hogy ha a hatványsíkok metszéspontja a szimplexszel lefedett, azaz $\varphi(x) \leq 0$ az $x \in S \cap \cup_{ij} \Pi_{ij}$ pontokra, akkor az egész szimplex lefedett.

Most arra a helyzetre koncentrálunk, amikor a maximumot az S egy belső pontján érjük el; azaz $x^* = V\lambda$, ahol $V = [v_1, v_2, \dots, v_h]$ és $\lambda_i > 0$, $\sum_i \lambda_i = 1$.

2.1.2. lemma. Tekintsük az SCO problémát egy szimplexten, amelynek csúcsai v_1, v_2, \dots, v_h és a hozzájuk tartozó B_i gömbök adottak. Legyen $x^* = V\lambda$ az SCO egy belső maximális pontja. Ekkor

$$d_p(x^*, v_1) = d_p(x^*, v_i), \quad i = 2, \dots, h, \quad (2.15)$$

vagyis minden egyes csúcs hatványfüggvényének ugyanaz az értéke x^* -ban.

A (2.15) egyenletet kielégítő úgynevezett θ pont teljesíti a belső optimum szükséges feltételét. Vegyük észre, hogy a θ pont a S ponton kívül is lehet. A síkokat kiegyenlítő θ pont kiszámításának képlete a hatványfüggvények értékeinek egyenlővé tételéből adódik:

$$(\theta - v_1)^T(\theta - v_1) - r_1^2 = (\theta - v_i)^T(\theta - v_i) - r_i^2, \quad i = 2, \dots, h. \quad (2.16)$$

Algebrai átalakítások után azt kapjuk, hogy θ megtalálása h darab lineáris egyenlőség megoldásával megkapható:

$$\begin{aligned} 2(v_i - v_1)^T V \lambda &= r_1^2 - r_i^2 + v_i^T v_i - v_1^T v_1, \quad i = 2, \dots, h \\ (1, 1, \dots, 1)^T \lambda &= 1. \end{aligned} \quad (2.17)$$

Ha θ egy belső pont, megvan az SCO maximális pontja, és egyszerűen ellenőrizhetjük, hogy $d_e(\theta, v_1)^2 \leq r_1^2$ az SC megoldásához.

Miután megtaláltuk a θ -t, lehet, hogy az S -n kívül van, de minden gömb által lefedett, $\theta \in \cap_j B_j$. A következő lemma azt mutatja, hogy ez elégséges feltétel a S lefedettségének igazolására, és nincs szükség további elemzésre.

2.1.3. lemma. *Legyen adott egy $x = V\lambda$ pont úgy, hogy $x \notin S$ és $x \in \cap_j B_j$. Ekkor $\exists y \in \cap_j B_j \cap S$.*

A 2.1.1-2.1.3 lemmák kombinációjának következménye a következő tétel:

2.1.4. tétel. *Legyen θ a (2.17) egy megoldása. Ha $\varphi(\theta) \leq 0$, akkor S lefedett.*

Megjegyezzük, hogy ez a tétel még akkor is érvényes, ha $\theta \notin S$. Azonban, ha $\theta \notin S$ és $\varphi(\theta) > 0$, akkor az SCO határon lévő megoldásait részletesebben kell vizsgálnunk. Amint a következő szakaszban látható, ott is szerepet játszik a θ -pont.

2.1.3. SCO optima a határon

Ha a φ maximuma nem az S belsejében van, akkor a határán található. Szélsőséges esetben a maximum az S egyik csúcsában lehet, ha egy nagy gömb lefedi az egész S szimplexet. Alternatív megoldásként az x^* maximális pont az S egy élén, vagy akár az S egy magasabb dimenziós határoló lapján található, legyen ez C . A θ -pont ismét nagy szerepet játszik a C felület meghatározásában. Tekintsük a $T = \text{conv}\{v_1, \dots, v_h, \theta\}$ halmazt. Most láthatjuk, hogy

$$\theta = \operatorname{argmax}_{x \in T} \varphi(x), \quad (2.18)$$

mert minden $x \in T$ esetén φ a θ irányában növekszik, tehát θ a (2.18) globális maximumpontja. Továbbá, θ a hatványdiagram közös csúcsa. Mivel $S \subset T$, így $\varphi(\theta) \geq \Phi$. Ez megerősíti azt az érvelést, hogy a negatív $\varphi(\theta)$ -ból következik, hogy S fedett.

Ez segít meghatározni az S lapját, C -t is, ahol a maximumot adó x^* pont található. Az általánosság megszorítása nélkül, legyen a θ pontot definiáló $\lambda_1, \dots, \lambda_m \geq 0$, és $\lambda_{m+1}, \dots, \lambda_h < 0$. Ekkor a C lap, ahol a maximális Φ értéket elérjük, a $C = \text{conv}\{v_1, \dots, v_m\}$ által adott. Ez azt jelenti, hogy az SCO $x^* = V\mu$ (lokális vagy globális) maximumpontjában $\mu_{m+1}, \dots, \mu_h = 0$. Mivel μ -nek $h - m$ eleme nulla, az SCO ekvivalens a következővel:

$$\max_C \varphi(x) = \max\{\varphi(\hat{V}\hat{\mu}) \mid \sum_{j=1}^m \hat{\mu}_j = 1, \hat{\mu} \geq 0\}, \quad (2.19)$$

ahol $\hat{\mu}$ és \hat{V} most már μ és V első m eleméből áll csak.

A problémát tekinthetjük alacsonyabb dimenziójú fedési feladatnak. Sajnos nem egy ekvivalens SC problémával van dolgunk. A B_1, \dots, B_m gömbök mellett a $B_j, j = m + 1, \dots, h$ gömbök \hat{B}_j vágásával is foglalkoznunk kell a C síkjában.

Az x^* globális optimumot a hatványsíkok csúcsait leíró poliéderes halmaz C halmazzal metszett egyik fizibilis bázismegoldásánál érjük el.

Mivel φ darabonként konvex, a fizibilis bázismegoldásoknál lokális optimumai lesznek, de nem lehet tudni, hogy melyik felel meg a Φ globális maximumnak. Ez azt mutatja, hogy az SC egy nehezen megoldható probléma. Főleg olyan eseteknél lesz ez így, ahol a nagyszámú csúcs egy szabálytalan szimplexet ír le, és egyik gömb sem fed le más csúcsot.

2.1.4. A szabályos eset

A kérdés az, hogy az SC probléma könnyebben megoldható-e szabályos szimplexek esetén. Az ellenőrzés első lépése természetesen az, hogy az egyik sugár elég nagy-e ahhoz, hogy a teljes szimplexet lefedje. A következő lépés a θ pont kiszámítása és annak ellenőrzése, hogy fedett-e. Ha ez utóbbi nem teljesül, akkor a (2.19) problémát az alacsonyabb dimenziójú C síkban kell megvizsgálni. A szabályos eset itt egy sajátos tulajdonsággal bír. Egy $v_j, j = m + 1, \dots, h$ csúcson az összes $v_i, i = 1, \dots, m$ csúcstól való egyenlő távolsága miatt a v_j vetülete a C lapra szintén egyenlő távolságú pontot ad. Ez azt jelenti, hogy minden v_j vetített csúcs a C középpontjába esik.

2.1.5. Eredmények

Megvizsgáltuk az SC problémát, amely azt hivatott eldönteni, hogy egy adott szimplexet az n dimenziós térben lefednek-e a csúcspontjában adott gömbök. Megmutattuk, hogy ez a probléma ekvivalens egy SCO globális optimalizálási probléma megoldásával. A következőket találtuk.

- A feladattól függően az SCO-nak lehet belső optimuma, amely megegyezik egy hatványdiagram úgynevezett θ csúcspontjával. Egy eljárást adtunk meg ennek a θ pontnak a megtalálására.
- Ha a θ pontot a gömbök lefedik, akkor a szimplex lefedett, függetlenül attól, hogy a θ pont a szimplexben belül vagy kívül helyezkedik el.
- Ha az SCO-nak határoptimuma van, akkor lehetnek lokális, nem globális optimumai is.
- Ez utóbbi esetben a θ pont határozza meg a szimplex C felületét, ahol az SCO globális optimumpontjai megtalálhatók. A globális optimum a C és az ún. hatványsíkok által meghatározott poliéder egy fizibilis bázismegoldása.
- Ez utóbbi esetben az SC probléma egyenértékű azzal a feladattal, hogy a C -t több gömbbel fedjük le, mint amennyit a csúcsok középpontjában lévő gömbök adnak.
- Egy szabályos szimplex esetén, az SCO optimális pontja vagy egyedi, vagy az SCO ekvivalens azzal a kérdéssel, hogy C lefedett-e a csúcsaiban lévő gömbökkel, illetve egy gömbbel a középpontjában.

2.2. Szimplex finomítás szabályos szimplexekkel

Ezt a kutatást egy keveréktervezési probléma motiválta, ahol olyan keverékeket keresünk, amelyek maximum n nyersanyagból állnak, és a termék egy olyan $x \in \mathbb{R}^n$ vektorral reprezentálható, amely megfelel bizonyos követelményeknek. A lehetséges keverékek halmazát matematikailag az n -dimenziós Δ egységshimplex határozza meg,

$$\Delta = \{x \in \mathbb{R}^n \mid \sum_{j=1}^n x_j = 1; 0 \leq x_j \leq 1 \forall j\}, \quad (2.20)$$

más néven standard $(n - 1)$ -szimplex, ahol az x_j változók az x termékben lévő összetevők hányadát adják meg. Itt nem ugyanazt a jelölést használjuk, mint eddig, mivel a csúcsok egységvektorok. Vegyük észre azt is, hogy ebben az esetben a szimplex dimenziója és a tér dimenziója mindig megegyezik, ami fontos tényező. Így ahelyett, hogy a csúcsok számát megkülönböztetnénk a dimenzióktól, közvetlenül n -t fogunk használni mindkettőre.

A probléma megoldásának egyik megközelítése egy szabályos rács meghatározása G egyenlő távolságra lévő rácspontokkal minden egyes tengelyhez, ami egy $\alpha = 1/(G - 1)$ hálóméretet eredményez. A méret persze relatív a Δ szimplex $w(\Delta)$ méretéhez, amit annak leghosszabb éle határoz meg. Az összes rácspont kiértékelésére irányuló stratégia nem vonzó, mivel nem hatékony. Ha egy egységintervallumon végezzük, a függvénykiértékelések száma exponenciálisan nő a dimenzióval: G^n . Ez nem olyan rossz az egységshimplex, mivel tudjuk a [9] cikk alapján, hogy a rácson lévő pontok teljes száma

$$\sum_{k=1}^{n-1} \binom{G}{k} \binom{n-2}{k-1}, \quad (2.21)$$

ami sajnos azért így is gyorsan növekszik.

A keverési problémák megoldására vonzó alternatíva a B&B stratégia alkalmazása, amelynek célja nem az összes pont legenerálása, hanem a terület felosztása és azon régiók (részhalmazok) tovább bontásának az elkerülése, amelyekről tudjuk, hogy nem tartalmaznak optimális megoldást.

Megmutatjuk, hogy lehet a B&B-módszerben olyan finomítást végezni, amely a felezés helyett szabályos szimplexekeket használ, és hogyan valósítható meg egy ilyen finomítás, illetve milyen hatással van a kiértékelendő szimplexekek és a generált mintapontok (csúcsok) számára.

2.2.1. Reguláris finomítás

Általában a szimplex-alapú halmazt (2.1) szerint definiáljuk, de megfogalmazhatjuk a következőképpen is,

$$S = \{x \in \mathbb{R}^n : x = V\lambda, \lambda \in \Delta\}, \quad (2.22)$$

ahol Δ a (2.20) formulában definiált egységsszimplex, és a V mátrix v_j csúcsai affin függetlenek. Alternatív megoldásként a korábbiaknak megfelelően leírhatunk egy szimplexet a csúcsok $\text{conv}(V)$ konvex burkaként is. Az a szemlélet, hogy egy általános szimplexet az egységsszimplex alapján írunk fel mint (2.22)-ben, azt mutatja, hogy a felosztás szempontjából elegendő egyszerűen az egységsszimplexre koncentrálni, hiszen a Δ felosztása a V mátrixszal való szorzással bármelyik szimplex felosztására lefordítható.

A kérdés az, hogyan lehet a Δ -t finomítani szabályos szimplexekek segítségével. Ehhez bevezetjük az Egységes Simplex Fedést (Uniform Simplex Cover, USC), ahol a szimplexet egyforma méretű, azonos orientációjú, egymást átfedő részszimplexekek fedik le, és analizáljuk ennek tulajdonságait. Az egyforma méretű és orientációjú szimplexekek ötletének megfogalmazásához a következő fogalmakat vezetjük be.

Minden egyes szimplex egy c középponttal és egy r sugárral rendelkezik, amelyek a következőképpen határozzák meg a csúcsmátrixát:

$$V = c\mathbf{1}^T + rD, \quad (2.23)$$

ahol $\mathbf{1}$ a csupa egyes vektor, $E = (e_1, \dots, e_n)$ pedig az egységmátrix. $D = (d_1, \dots, d_n) = E - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ egy szimmetrikus mátrix, amelynek irányai a középpontból az egységsszimplex csúcsai felé mutatnak. Vegyük észre, hogy az egységsszimplex Δ középpontja $\frac{1}{n}\mathbf{1}$, míg az egységsszimplex sugara, azaz a D mátrixhoz viszonyított mérete 1.

Az átfedő tulajdonság tanulmányozásához az első kérdésünk az, hogy hogyan határozzuk meg, hogy egy szabályos szimplex tartalmaz-e egy másikat (lefedi).

2.2.1. tétel. Legyenek S_1 és S_2 szimplexekek, amelyeknek a csúcsmátrixai $V = c_1\mathbf{1}^T + r_1D$ és $W = c_2\mathbf{1}^T + r_2D$. Az S_2 szimplexet az S_1 szimplex lefedti (tartalmazza), ha

$$c_{2j} - c_{1j} + \frac{r_1 - r_2}{n} \geq 0 \quad j = 1, \dots, n. \quad (2.24)$$

A 2.2.1 tétel azt is megmutatja, hogy az $r_1 \geq r_2$ sugár szükséges feltétele annak, hogy az S_2 az S_1 részhalmaza lehessen.

A Δ egységsszimplex finomítását adjuk meg a következőképpen:

$$R = \{S_1, \dots, S_p\}, \quad (2.25)$$

ahol az egybevágó részszimplexekek használatának kérdése az, hogy mely c_1, \dots, c_p középpontokat és r sugarat használjuk úgy, hogy a Δ -t az R finomítás lefedje, azaz

$$\Delta \subset \bigcup_{i=1}^p S_i. \quad (2.26)$$

A következőkben egy olyan helyzetet vizsgálunk, ahol az m középpontot szabályos módon helyezzük el az egyes élek mentén. Ily módon egy olyan R felosztást kapunk, amely egységesen lefedti Δ -t. Ezt m USC finomításnak nevezzük.

2.2.1.1. 2USC, fedés n szimplexszel

Legyen adott egy szimplex, c középponttal és r sugárral definiálva. Az R 2USC finomítása n részszimplex, S_i , $i = 1, \dots, n$, amelyek relatív mérete $\beta \geq \frac{n-1}{n}$. Minden egyes S_i részhalmaznak a sugara $r_i = \beta r$ és a középpontja a következőképpen adható meg,

$$c_i = \beta c + (1 - \beta)v_i = \beta c + (1 - \beta)(c + rd_i) = c + (1 - \beta)rd_i. \quad (2.27)$$

Az egységsszimplexre a $\beta = \frac{n-1}{n}$ redukciót használva

$$c_i = \frac{n-1}{n^2} \mathbf{1} + \frac{1}{n} e_i. \quad (2.28)$$

2.2.2. tétel. A Δ egységsszimplex 2USC finomítása, R , lefedi Δ -t, ha $\beta \geq \frac{n-1}{n}$. Vagyis, $\forall x \in \Delta$ -hez $\exists S_i \in R$ úgy, hogy $x \in S_i$.

2.2.3. állítás. A Δ egységsszimplex 2USC finomításból származó R minden S_i részhalma tartalmazza a $c = \frac{1}{n} \mathbf{1}$ középpontot.

2.2.1.2. Finomítás $n + 1$ szabályos szimplexszel, 2∇ USC

Vizsgálataink során olyan fedéseket is találtunk, amik kisebb átfedést mutatnak, de amelyek esetében lazítani kell az egyforma irányú és méretű tulajdonságon. Tekintsük a 2USC finomítás S_i szimplexszel azonos szerkezetű és $\frac{n-2}{n-1} \leq \rho < \frac{n-1}{n}$ sugarú részszimpleteket, \check{S}_i , $i = 1, \dots, n$ jelöléssel. Az \check{S}_i -k egyesítése egy kis „lyukat” hagy Δ -ban, amely egy $r_{n+1} = n(1 - \rho) - 1$ sugarú, „fejjel lefelé” orientált szimpex, \check{S}_{n+1} . \check{S}_{n+1} középpontja az eredeti szimplex középpontja, azaz $c_{n+1} = \frac{1}{n} \mathbf{1}$, csúcsai pedig $U_{n+1} = c_{n+1} \mathbf{1}^T - r_{n+1} D$ képlettel adhatók meg, amelyek a Δ lapjainak középpontjai, ha $\rho = \frac{n-2}{n-1}$. A $P = \{\check{S}_1, \dots, \check{S}_{n+1}\}$ finomítás $n + 1$ szimpexet tartalmaz, amik kisebb méretűek, mint S_j .

2.2.4. tétel. A Δ egységsszimplex 2∇ USC finomítása, P , lefedi Δ -t, ha $\frac{n-2}{n-1} \leq \rho < \frac{n-1}{n}$. Vagyis, $\forall x \in \Delta$ -hoz $\exists \check{S}_i \in P$ úgy, hogy $x \in \check{S}_i$.

2.2.1.3. Élenként több bázispontot használó finomítás, m USC

Az m USC koncepciója, hogy több, de kisebb, azonos irányultságú és méretű szimplexet használunk, élenként pontosan m darabot.

Az m USC finomítás, M , olyan S_i^m részszimpletekből áll, amelyek a finomítandó szimplex r sugarát $\gamma \geq \frac{n-1}{m+n-2}$ redukcióval kicsinyítik. A középpontokat az eredeti középpontból tett, csúcsok felé irányuló lépésekkel tudjuk megadni. A középpontból a C_{t_1, \dots, t_n} középpontú új részszimplext úgy írjuk le, hogy minden d_i irányban t_i egész számú relatív lépést teszünk, ahol a relatív lépések összege $m - 1$, azaz $\sum_i t_i = m - 1$.

$$C_{t_1, \dots, t_n} = c + \frac{(1 - \gamma)r}{m - 1} \sum_{i=1}^n t_i d_i. \quad (2.29)$$

A multinomiális eloszláshoz hasonlóan összesen

$$N(m, n) = \binom{m + n - 2}{m - 1} \quad (2.30)$$

ilyen t vektor van, amelyek megfelelnek a részszimpletek középpontjainak. Vegyük észre, hogy az $m - 1 = n$ speciális esetben a $t = \mathbf{1}^T$ részsimplex középpontja, azaz $C_{1, \dots, 1}$, az eredeti c középpontnak felel meg.

2.2.5. tétel. Az Δ egységsszimplex m USC finomítása, M , lefedi Δ -t, ha $\gamma \geq \frac{n-1}{m+n-2}$. Vagyis, $\forall x \in \Delta$ -hoz $\exists S_i^m \in M$, amire $x \in S_i^m$.

Bár ugyanazokat a végső szimplexekeket kaphatjuk iteratív 2USC-felosztással, a generált szimpexek száma kisebb, ha m USC-t használunk.

2.2.2. Csúcs-megosztás

A B&B szempontjából kényelmes, ha a keresési fa szimplexei közös csúcsokon osztoznak, mert:

- a csúcsokat nem kell újraértékelni, ami csökkenti a kiértékelések számát,
- nagyobb az esélye annak, hogy egy részhalma egy másik ágból származó nagyobb szimplex lefed.

A 2USC redukciós rátái bizonyos határokon belül kiigazíthatók a cél elérése érdekében. A kérdés most az, hogyan lehet ezt úgy tenni, hogy a csúcsok megosztását elősegítsük.

Belátjuk, hogy létezik olyan β redukciós ráta a 2USC-re, hogy N egymást követő finomítás után a kezdeti szimplex minden élén legalább két kiértékelt csúcs egybeessen. Ennek feltétele, hogy $1 - \beta = \beta^N$.

2.2.2.1. Redukciós ráták a csúcsok rácson való elhelyezésének elősegítésére

Tekintettel arra, hogy élenként G mintapontunk van (beleértve az egységszimplex csúcsait is), a rácsméret $\alpha = \frac{1}{G-1}$ nem választható teljesen szabadon, de nem is korlátozható a felezés $\frac{1}{2^k}$ értékére. A rácspontok száma élenként legalább $G \geq \left\lceil \frac{1}{\varepsilon} \right\rceil + 1$. Ha a pontokat élenként G pontból álló rácstra tesszük, akkor azok legfeljebb $\alpha \leq \varepsilon$ távolságra lesznek egymástól.

A következőkben a 2USC-Grid és a 2 ∇ USC-Grid módszereket mutatjuk be. Megmutatjuk, mekkora az α értéke, és hogyan lehet a szabályos fedési módszerek β, ρ redukciós rátáit úgy megadni, hogy a csúcsok a megfelelő rácstra essenek.

Legyen $\omega = \frac{w(S)}{w(\Delta)}$ a fa utolsó előtti szintjén finomítandó S szimplex relatív mérete, azaz S gyerekeinek mérete kisebb vagy egyenlő, mint ε . Tegyük fel, hogy az utolsó szimplexek mérete pontosan ε , és a lehető legnagyobb csökkentést, azaz a minimális redukciós rátát használjuk. 2USC esetén $\varepsilon = \omega \frac{n-1}{n}$ és a kívánt hálóméret $\alpha = \omega(1 - \frac{n-1}{n}) = \frac{\omega}{n} = \frac{\varepsilon}{n-1}$. 2 ∇ USC esetén $\varepsilon = \omega \frac{n-2}{n-1}$ és $\alpha = \omega(1 - \frac{n-2}{n-1}) = \frac{\omega}{n-1} = \frac{\varepsilon}{n-2}$. Vegyük észre, hogy mindkét esetben $\alpha < \varepsilon$. A gyakorlatban az utolsó szimplex mérete nem feltétlenül ε , de megmutatható, hogy a végső méret $(\lceil \frac{1}{\varepsilon} \rceil)^{-1}$, tehát 2USC esetén $\alpha = \frac{(\lceil \frac{1}{\varepsilon} \rceil)^{-1}}{n-1}$ és 2 ∇ USC esetén $\alpha = \frac{(\lceil \frac{1}{\varepsilon} \rceil)^{-1}}{n-2}$.

Legyen ϕ a minimális redukciós ráta, azaz $\frac{n-1}{n}$ a 2USC esetén és $\frac{n-2}{n-1}$ a 2 ∇ USC esetén. A finomítás csúcsait egy bizonyos, az ε értéktől függő α hálóméret esetén a rácstra helyezhetjük, ha redukciós rátaként a következő értéket használjuk fel

$$\psi(\omega) = \frac{\left\lceil \frac{\phi\omega}{\alpha} \right\rceil \cdot \alpha}{\omega}. \quad (2.31)$$

2.2.6. megjegyzés. Viszonylag könnyen belátható, hogy a 2USC-Grid finomítás minden generált csúcsa egy α méretű rácson helyezkedik el. A 2 ∇ USC-Grid esetében ez nem nyilvánvaló. Tekintsük az S szimplex $\mathcal{S}_1, \dots, \mathcal{S}_{n+1}$ 2 ∇ USC-finomítását, ahol az S csúcsai egy α -rácson vannak. Ez azt jelenti, hogy van egy olyan K érték, amire $\omega = \frac{w(S)}{w(\Delta)} = K \cdot \alpha$. A (2.31) egyenlet olyan redukciós rátát ad, amihez van egy másik egész szám $k \leq K$, amire $\psi(\omega)\omega = k \cdot \alpha$. A felfordított szimplexre $\frac{w(\mathcal{S}_{n+1})}{w(\Delta)} = \omega \cdot (n(1 - \psi(\omega)) - 1) = (n-1) \cdot K \cdot \alpha - n \cdot k \cdot \alpha$, azaz a csúcsok közötti relatív távolság szintén α többszöröse. Mivel a középpontja megegyezik az S középpontjával, a csúcsai a rácson fekszenek.

2.2.3. Eredmények

A szabályosságnak számos előnye van a gyakorlatban. Megmutattuk, hogy a szimplex finomítása hatékonyan elvégezhető szabályos részszeletekkel. Az eredmények elméleti és gyakorlati szempontból egyaránt ellentmondásosak és meglepők. Több átfedő finomítási módszert definiáltunk, bizonyítottuk helyességüket, majd megvalósítottuk őket egy szimplex-alapú B&B módszerben.

A numerikus eredmények azt mutatják, hogy a keresési tér felosztásának hagyományos ötlete néha helyettesíthető egy átfedő finomítással, amely egy rácson generálja a csúcsokat. A rácsháló ismeretében a kiértékelt szimplexek száma több nagyságrenddel csökken. A kiértékelt csúcsok számában azonban általában rosszabb, mint a leghosszabb él felezésének használata. Ha a szimplex- és a csúcsok kiértékelésének számítási költsége hasonló, akkor a 2 ∇ USC-Grid a legjobb választás.

2.3. Gradiens alapú korlátok és monotonitás szimplexeken

Ebben a szakaszban megmutatjuk, hogy az első deriváltak korlátai hogyan használhatók fel szimplexek esetén az függvényérték korlátok megadására és új monotonitási tesztek kidolgozására. A kapott korlátokat egy B&B algoritmusban implementáljuk, hogy összehasonlíthassuk a különböző már létező technikákkal.

Tekintsünk egy $f : \mathbb{R}^n \rightarrow \mathbb{R}$ függvényt, amelyet minimalizálni szeretnénk egy $D \subset \mathbb{R}^n$ halmazon, amely vagy egy intervallum, vagy egy szimplex, és amelyen f differenciálható:

$$\min_{x \in D} f(x). \quad (2.32)$$

Valójában a probléma általánosabban is feltehető, ahol a fizibilis halmaz nem teljes dimenziójú a cél-függvényhez képest. A B&B algoritmus a partíciós halmazok Λ listájával dolgozik, amelyek összességében

tartalmazzák az összes globális minimumpontot. Legyen $\mathcal{V} = \{v_1, \dots, v_h\} \subset \mathbb{R}^n$ affin független csúcsok halmaza. A j csúcs i komponensére a v_{ji} jelölést használjuk.

A vizsgálandó B&B algoritmus most a leghosszabb él felosztását használja, ahol az S partícióhalmaz (v, w) leghosszabb élét $x = \frac{v+w}{2}$ középpont segítségével felezi, ami két új szimplexhez vezet, a $\mathcal{V} \setminus \{v\} \cup \{x\}$ és $\mathcal{V} \setminus \{w\} \cup \{x\}$ csúcshalmazokkal. Az algoritmus figyelembe veszi a monotonitásból adódó dimenziócsökkenést is, amikor az S $(h-1)$ -szimplex csúcsainak \mathcal{V} halmaza $\mathcal{V} \setminus \{v\}$ -re csökken, és S helyére egy (vagy több) lapja kerül, $F := \text{conv}(\mathcal{V} \setminus \{v\})$ valamely $v \in \mathcal{V}$ -re.

Az $S = \text{conv}(\{v_1, \dots, v_h\})$ $(h-1)$ -szimplex középpontja a $c = \frac{1}{h} \sum_{j=1}^h v_j$, a relatív belsejét pedig a következő módon definiáljuk:

$$\text{rint}(S) = \{x = \sum_j \lambda_j v_j, \lambda_j > 0, j = 1, \dots, h, \sum_{j=1}^h \lambda_j = 1\}. \quad (2.33)$$

Egy S szimplex relatív határát úgy határozzuk meg, hogy eltávolítjuk belőle a relatív belsejét, azaz $\partial S = S \setminus \text{rint}(S)$. Egy S szimplexet tekintve érdekes kérdés, hogy S -nek F lapjai a D fizibilis halmaz határán vannak-e.

2.3.1. definíció. Adott egy D fizibilis halmaz, és egy S $(h-1)$ -simplex, $h \leq n$. S -t D -re nézve határ-szimplexnek nevezzük, ha létezik D -nek egy $(h-1)$ -dimenziós φ lapja, amelyre $S \subseteq \varphi$.

2.3.2. jelölés. Az S szimplex intervallumburka a $\square S = \square \text{conv}(\mathcal{V})$, azaz a legkisebb intervallum, amely az S szimplexet tartalmazza. Legyen $x = \square S$, ahol

$$x_i = [\underline{x}_i, \bar{x}_i] = [\min_{v_j \in \mathcal{V}} v_{ji}, \max_{v_j \in \mathcal{V}} v_{ji}] \quad \forall i \in \{1, \dots, n\}. \quad (2.34)$$

2.3.1. Korlátozási módszerek szimplexre

2.3.1.1. A standard intervallum-korlátok kiterjesztése a szimplexekre

Jelölje $f(x)$ egy f függvény befoglalását, $\nabla f(x)$ a gradiensek befoglalását egy x intervallumon továbbá legyen $\nabla f_i(x) = [\underline{\nabla f}_i(x), \overline{\nabla f}_i(x)]$ az intervallumgradiens i -edik komponense. Ezek a befoglalások kiszámíthatók az intervallumaritmetika és az automatikus differenciálás [33] segítségével.

Adott S szimplex esetén, $\forall x \in S \subset x = \square S$, $f(x) \in f(x)$, illetve $\forall x \in x$, $\frac{\partial f}{\partial x_i}(x) \in \nabla f_i(x)$. Ekkor $\forall x \in S \subset x = \square S$, $\frac{\partial f}{\partial x_i}(x) \in \nabla f_i(x)$.

Jelöljük a c középpontú x intervallumon vett középponti alakot $f_c(x)$ -szel. Ez tulajdonképpen az elsőrendű Taylor-féle intervallum kiterjesztés a $\nabla f(x)$ segítségével:

$$f_c(x) = f(c) + (x - c)^T \nabla f(x), \text{ ahol } c \in x.$$

Általában c a x intervallum középpontja, ilyenkor az $f_{cb}(x)$ jelölést használjuk, ahol $cb = \text{mid}(x)$. Az alsó korlát $\underline{f}_c(x)$ a következőképpen is leírható:

$$\underline{f}_c(x) = \underline{f}(c) + (x - c)^T \underline{\nabla f}(x) = \underline{f}(c) + (x - c)^T \underline{\nabla f}(x),$$

ahol az aláhúzás az intervallum aritmetika által kiszámított képlet alsó korlátját jelenti. Megjegyezzük, hogy $\forall x \in S \subset x = \square S$ pontra, $f(x) \in f_c(x)$. Így $f_c(x)$ az f alsó és felső korlátját adja S felett, még akkor is, ha $c \notin S$.

Baumann [4] a cb középpont helyett két másik alappontot javasolt a középponti formula alsó és felső korlátjának javítására. Jelöljük bb^- -szel az eredeti intervallumos Baumann-pontot, ami az optimális alsó korláthoz vezet a középponti formulában. Az alappont i . komponensét a következőképpen adjuk meg

$$bb_i^- = \begin{cases} \frac{\underline{x}_i \overline{\nabla f}_i(x) - \bar{x}_i \underline{\nabla f}_i(x)}{\text{wid}(\nabla f_i(x))} & \text{if } 0 \in \nabla f_i(x) \\ \underline{x}_i & \text{if } \underline{\nabla f}_i(x) > 0 \\ \bar{x}_i & \text{if } \overline{\nabla f}_i(x) < 0. \end{cases}$$

Bármely középponti formula (amelynek alappontja $y \in x$) az S szimplex csúcsai alapján szigorítható.

2.3.3. állítás. *Legyen*

$$\underline{f}_y(S) = \underline{f}(y) + \min_{v \in \mathcal{V}} \{(v - y)^T \nabla f(x)\}, \quad y \in x. \quad (2.35)$$

Ekkor $\underline{f}_y(S) \leq \min_{x \in S} f(x)$.

A (2.35)-ben használhatjuk az $y = cb$ vagy az $y = bb^-$ alappontokat. A bb^- Baumann-pont is általánosítható szimplex-alapú alapponttá. A cél a legjobb alappont kiválasztása a Taylor-formához, úgy, hogy az alsó korlát a lehető legnagyobb legyen. Egy szimplex esetében ahelyett, hogy az $x = \square S$ zárt intervallum határértékeit használnánk, a szimplex csúcspontjait vesszük alapul.

A (2.35) egy szimplex feletti legnagyobb alsó korlátjához a következő probléma megoldása adja az optimális alappontot:

$$\begin{aligned} \operatorname{argmax}_{y \in x} \min_{v \in \mathcal{V}} \left(\underline{f}(y) + (v - y)^T \nabla f(x) \right) = \\ \operatorname{argmax}_{y \in x} \left(\underline{f}(y) + \min_{v \in \mathcal{V}} (v - y)^T \nabla f(x) \right). \end{aligned} \quad (2.36)$$

Nyilvánvaló, hogy (2.36) egy nemlineáris probléma, mivel magában foglalja az $f(y)$ optimalizálását az y változóban. Ezért célszerű relaxálni a problémát és csak a második részt optimalizálni.

2.3.4. definíció. *Definiáljuk $bs^- = \operatorname{argmax}_{y \in x} \min_{v \in \mathcal{V}} (v - y)^T \nabla f(x)$ pontot mint Baumann-pontot a szimplex felett.*

Ez a pont egy intervallumos lineáris programmal megtalálható:

$$\begin{aligned} \max_{y \in x, z \in \mathbb{R}} \quad & z \\ \text{s.t.} \quad & z \leq (v - y)^T \nabla f(x), \quad \forall v \in \mathcal{V}. \end{aligned} \quad (2.37)$$

Legyen (z^*, y^*) a (2.37) optima. Ekkor vegyük a $bs^- = y^*$ alappontot a hozzá tartozó $\underline{f}_{bs^-}(x) = f(bs^-) + z^*$ alsó korláttal.

2.3.5. jelölés. Az x intervallum minden w csúcspontjához tartozik egy gradiens korlát, amit jelöljön $\nabla^w f(x) \in \mathbb{R}^n$, és aminek az i . komponense $\nabla^w f_i(x) = \underline{\nabla} f_i(x)$ ha $w_i = \underline{x}_i$ és $\nabla^w f_i(x) = \overline{\nabla} f_i(x)$ ha $w_i = \overline{x}_i$.

A (2.37) lineáris programként való felírása 2^n korlátozást igényel minden egyes v csúcsra a \mathcal{V} -ban:

$$\begin{aligned} \max_{y \in x, z \in \mathbb{R}} \quad & z \\ \text{s.t.} \quad & z \leq (v - y)^T \nabla^w f(x), \quad \forall v \in \mathcal{V}, \quad x \quad \forall w \text{ csúcsára.} \end{aligned} \quad (2.38)$$

A (2.38) korlátozások egyszerűsíthetők úgy, hogy összesen 2^n lineáris egyenlőtlenséget adjuk meg.

$$\begin{aligned} \max_{y \in x, z \in \mathbb{R}} \quad & z \\ \text{s.t.} \quad & z + y^T \nabla^w f(x) \leq \min_{v \in \mathcal{V}} v^T \nabla^w f(x), \quad x \text{ minden } w \text{ csúcsára.} \end{aligned} \quad (2.39)$$

Vegyük észre, hogy (2.37), (2.38) és (2.39) ugyanannak a problémának egyenértékű leírásai, így ugyanazt az optimumot adják meg.

2.3.1.2. Korlátok Affin Aritmetikával

Ez a szakasz az Affin Aritmetika (lásd [2, 15, 28, 29, 34, 38]) használatát írja le a f függvény lineáris alsókorlátjának előállítására $x = \square S$ felett. Az $x = \square S$ intervallumvektor átalakítható affin formájú vektorra, amelyet \hat{x} jelöl, a következőképpen,

$$\hat{x} = \operatorname{mid}(x) + \operatorname{rad}(x)\varepsilon$$

ahol $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$, és $\varepsilon_i \in [-1, 1]$ minden $i \in \{1, \dots, n\}$. Az affin alak \hat{x} visszaalakítható intervallummá, ha ε_i -t a $[-1, 1]$ intervallummal helyettesítjük. Továbbá, minden $x \in x$ -re pontosan egy megfelelő értéke van az ε -nak az affin leírásban,

$$x = T(x, \varepsilon) = \operatorname{mid}(x) + \operatorname{rad}(x)\varepsilon,$$

ahol $\varepsilon_i = \frac{x_i - \text{mid}(x_i)}{\text{rad}(x_i)}, i = 1, \dots, n$.

Ha az x_i változó minden előfordulását az \hat{x}_i affin formával helyettesítjük az f kifejezésben, és a számításokat az affin aritmetika segítségével végezzük el, akkor a következő affin formát kapjuk:

$$\hat{f}(T(\mathbf{x}, \varepsilon)) = r_0 + \sum_{i=1}^n r_i \varepsilon_i + \sum_{k=n+1}^N r_k \varepsilon_k, \quad (2.40)$$

ahol ε_j a $[-1, 1]$ tartományban van minden $j \in \{1, \dots, N\}$ esetén. Megjegyezzük, hogy minden $k \in \{n+1, \dots, N\}$ esetén $r_k \varepsilon_k$ hibatagok az f nem affin műveleteiből származnak.

Az affin aritmetika segítségével (2.40) ad egy alsókorlátot f -re x felett,

$$f(x) = f(T(\mathbf{x}, \varepsilon)) \geq \underline{\hat{f}}(T(\mathbf{x}, \varepsilon)) = r_0 + \sum_{i=1}^n r_i \varepsilon_i - \sum_{k=n+1}^N |r_k|, \quad (2.41)$$

mert minden hibatagot a legrosszabb értékükkel vesszünk figyelembe.

Lineáris program az alsó korlát megadására

Annak érdekében, hogy az f alsó korlátját az S szimplexén (és nem csak az $x \in S$ -n) számítsuk ki, az x pontot úgy korlátozzuk, hogy az S -en belül legyen (2.1) hozzáadásával. Ebben az esetben x -et a $T(\mathbf{x}, \varepsilon)$ affin alakjával írjuk le, és így a következő lineáris programot kapjuk:

$$\begin{aligned} \min_{\substack{\varepsilon \in [-1, 1]^n \\ \lambda \in [0, 1]^{n+1}}} & \sum_{i=1}^n r_i \varepsilon_i \\ \text{s.t.} & T(\mathbf{x}, \varepsilon) = \sum_{j=1}^h \lambda_j v_j \\ & \sum_{j=1}^h \lambda_j = 1. \end{aligned} \quad (2.42)$$

Ha az (2.42) egzakt megoldását $(\varepsilon^*, \lambda^*)$ -gal jelöljük, akkor azt kapjuk, hogy

$$f(x) = f(T(\mathbf{x}, \varepsilon)) \geq \sum_{i=1}^n r_i \varepsilon_i^* + r_0 - \sum_{k=n+1}^N |r_k|, \quad \forall x \in S \quad (2.43)$$

és így ez az f alsó korlátja S felett.

Megjegyezzük, hogy a fenti lineáris program megoldásához csak ki kell értékelnünk f -et az S minden egyes csúcán, majd a (2.41) lineáris alsókorlát függvény minimális értékét kell vennünk. Ha $\text{rad}(x_i) \neq 0$, akkor $\varepsilon_i(v_j)$ a v_j csúcs i komponensének a ε változóba történő transzformációja (különben $r_i = 0$), azaz

$$\varepsilon_i(v_j) = \begin{cases} \frac{v_{ji} - \text{mid}(x_i)}{\text{rad}(x_i)}, & \text{ha } \text{rad}(x_i) > 0 \\ 0, & \text{ha } \text{rad}(x_i) = 0. \end{cases} \quad \forall i \in \{1, \dots, n\}, \text{ és } \forall v_j \in \mathcal{V}.$$

Ekkor az (2.43) alsó korlátja a következő lesz:

$$\min_{j \in \{1, \dots, h\}} \sum_{i=1}^n r_i \varepsilon_i(v_j) + r_0 - \sum_{k=n+1}^N |r_k|. \quad (2.44)$$

Tehát a (2.42) lineáris program megoldása helyett a fenti (2.44) egyenlettel így közvetlenül megkapjuk az f alsó korlátját S felett. A megoldás azonos a (2.42) lineáris program megoldásával.

2.3.2. Monotonitási teszt

A célunk egy hatékony monotonitási teszt kidolgozása, amely kizárja az S minden olyan részét, amely nem tartalmazhat stacionárius pontot.

2.3.6. állítás. Legyen $S \subset \text{int}(D)$ egy szimplex a D keresési tér belsejében. Ha $\exists i \in \{1, \dots, n\}$ és $0 \notin \nabla f_i(\square S)$, akkor S nem tartalmazza a (2.32) globális minimumpontját.

Ennek gyakorlati következménye, hogy S belseje nem tartalmazhatja az optimumot, így S helyettesíthető a lapjaival.

2.3.7. következmény. Legyen $S \subseteq D$ egy szimplex mint partíciós halmaz egy B&B algoritmusban a $\nabla f(\square S)$ gradiensbefoglalással. Ha $\exists i \in \{1, \dots, n\}$ hogy $0 \notin \nabla f_i(\square S)$ -ban, és S -nek nincsenek határ-lapjai, akkor S eldobható.

Akkor is eltávolíthatjuk S -t, ha S csak néhány csúcsa érinti D határát a 2.3.7 következmény szerint. Ellenkező esetben az S határ-lapjait, azaz amelyek a keresési tér határára esnek, új, kisebb dimenziós szimplex partícióként tárolunk S helyett.

Megmutatható, hogy egy F lap relatív belseje nem tartalmazhat minimumpontot. Ehhez a tulajdonsághoz a d iránymenti derivált korlátaira összpontosíthatunk, amelyeket $\underline{d^T \nabla f(\square S)}$ és $\overline{d^T \nabla f(\square S)}$ jelöl. Például, az iránymenti derivált felső korlátja

$$\overline{d^T \nabla f(\square S)} = \sum_{i=1}^n \max\{d_i \underline{\nabla f}_i(\square S), d_i \overline{\nabla f}_i(\square S)\}. \quad (2.45)$$

Vegyük észre, hogy (2.45) szerint a $\overline{d^T \nabla f(\square S)} \leq 0$ szükséges feltétele, hogy f monoton legyen $\square S$ -en, azaz $0 \notin \nabla f(\square S)$.

A B&B eljárásban, ha a keresési tér egy teljes dimenziós szimplex, akkor a partíciós halmazok teljes dimenziós n -szimplexek, egészen a lapjaira való redukálásáig. A $0 \notin \nabla f(\square S)$ tulajdonságot használva, ha $\text{rint}(S)$ eltávolítható, akkor redukálhatjuk a határ-lapjaira (azokra a lapjaira, amik a D határán vannak). Azonban az alacsonyabb dimenziójú $(h-1)$ -szimplexek esetében, amikor $h \leq n$, a $0 \notin \nabla f(\square S)$ feltétel szükséges feltétele a monotonitásnak, de nem elégséges. Ezért az iránymenti derivált korlátokra kell összpontosítanunk. A jelölések megkönnyítése érdekében a továbbiakban az S szimplex intervallumburkához tartozó gradiens befoglalását G -vel jelöljük, $G = \nabla f(\square S)$.

Most tekintsünk egy d irányvektort az S két pontjának különbségeként. A megfelelő $d^T \nabla f(x)$ iránymenti derivált is szerepelni fog a belső szorzatban, ami

$$\langle d, G \rangle = \left[\underline{d^T G}, \overline{d^T G} \right] = \left[\sum_{i=1}^n \min\{d_i \underline{G}_i, d_i \overline{G}_i\}, \sum_{i=1}^n \max\{d_i \underline{G}_i, d_i \overline{G}_i\} \right]. \quad (2.46)$$

A legáltalánosabb eredmény egy $(h-1)$ -szimplexre a következő.

2.3.8. állítás. Legyen $S \subseteq D$ egy $(h-1)$ -szimplex, $h \leq n+1$ és gradiens befoglalása G . Ha $\exists x, y \in S$ úgy, hogy a $d = x - y$ iránynak megfelelő (2.46) iránymenti derivált korlátjára $0 \notin [\underline{d^T G}, \overline{d^T G}]$ teljesül, akkor $\text{rint}(S)$ nem tartalmazza a (2.32) globális minimumpontját.

Az algoritmus kidolgozása a d irány megválasztásától és a számítási módjától függ. Az 2.3.8 állításnak is van némi gyakorlati haszna egy B&B algoritmusban.

2.3.9. következmény. Legyen $S \subseteq D$ egy $(h-1)$ -simplex mint partíciós halmaz egy B&B algoritmusban a megfelelő G gradiensbefoglalással. Ha a 2.3.8 állítás feltételei érvényesek, és S -nek nincsenek határ-lapjai, akkor S eldobható.

Az érv az, hogy S relatív határa tartalmazhat egy globális minimumpontot, de ugyanez a pont egy másik partícióhalmaz relatív határában is benne van. Mivel a minimum nem az $\text{rint}(S)$ -ben van, el kell döntenünk, hogy S melyik lapjára koncentráljunk.

2.3.10. állítás. Legyen adott egy $(h-1)$ -szimplex, $S = \text{conv}(\mathcal{V})$, $1 \leq h \leq n+1$ és annak egy F lapja, amelyet a v csúcsnak a \mathcal{V} -ből való eltávolításával kapunk. Tekintsük a $d = v - y$ irányt, ahol $y \in S$. Ha $\underline{d^T G} > 0$, azaz a d iránymenti derivált pozitív, akkor az F lapon az S összes minimum pontja megtalálható, vagyis $\text{argmin}_{x \in S} f(x) \subseteq F$.

2.3.11. megjegyzés. Vegyük észre, hogy ha több olyan lap van, amely teljesíti a 2.3.10 állítás feltételeit, akkor az állítás a lapok metszetére is érvényes, amely az S egy még kisebb dimenziós lapja. Tehát ha a 2.3.10 állítás a F és H lapokra érvényes, akkor a $F \cap H$ lapra is érvényes.

2.3.12. állítás. Legyen adott S $(h - 1)$ -szimplex, amely megfelel a 2.3.10 állítás feltételeinek. Ha F egy nem-határ lap, akkor a (2.32) globális minimumpontjai nem lehetnek $\text{rint}(F)$ -ben, azaz globális minimumpontok csak F relatív határán találhatók.

2.3.13. megjegyzés. Először is jegyezzük meg, hogy az F határát más szomszédos partícióhalmazok tartalmazzák, így azt is eltávolíthatjuk. Következésképpen az S szimplexet is figyelmen kívül hagyhatjuk a B&B eljárásban.

Másodszor, vegyük észre, hogy több olyan lap is lehet, amely teljesíti a 2.3.10 állítás feltételét, azonban ahhoz, hogy az S szimplexet kihagyhassuk a további vizsgálatból, elegendő egy olyan nem-határos lap, amely teljesíti a feltételt.

2.3.14. állítás. Legyen adott egy szimplex, $S = \text{conv}(\mathcal{V})$, $1 \leq h \leq n + 1$, és tekintsük a F lapját, amely a v csúcson a \mathcal{V} -ből való eltávolításával kapunk. Ha egy $y \in S$ pontra és $d = v - y$ irányra $\bar{d}^T \bar{G} < 0$, akkor $\text{rint}(F) \subset S$ nem tartalmazhatja az (2.32) minimum pontját.

Ez egy szép teszthez vezet arra az esetre, ha a szimplexben van monoton irány, de nem tudjuk megmutatni, hogy bármelyik lapja tartalmazza az összes minimumot a 2.3.10 állítás segítségével. Ilyen esetben a szimplexet az összes határ-lapjával kell helyettesítenünk, kivéve, ha a 2.3.14 állítás érvényes. Így csak azokat a határ-lapokat kell figyelembe vennünk, ahol nem találunk olyan y pontot, amelyre a 2.3.14 állítás érvényes.

2.3.3. A határ-lapok címkézése

Ahhoz, hogy megtudjuk, hogy egy adott lap határon van-e vagy sem, egy címkézési rendszert használunk, ami követi, hogy az adott lap vagy szimplex a D melyik minimális dimenziójú lapjába tartozik. Itt csak azt az esetet vesszük figyelembe, amikor a fizibilis halmaz egy szimplex, mivel [23] tárgyalja az intervallumos fizibilis halmazok esetét, ami könnyebben kezelhető.

Egy $D = \text{conv}(\mathcal{W})$ szimplex fizibilis halmaz esetén ez úgy történik, hogy D minden φ felületéhez hozzárendelünk egy $\mathcal{B}(\varphi)$ címkét, kezdve a csúcsokkal, amiket $\mathcal{B}(v_j) = 0..010..0$ címkével látunk el (ami egy $n + 1$ hosszú bitsztring), ahol v_j csúcsra a j -edik bit az egyetlen 1-es, $j = 1, \dots, n + 1$. Minden φ lapra, amely a $\mathcal{V} \subseteq \mathcal{W}$ csúcsok konvex kombinációja, a megfelelő $\mathcal{B}(\varphi)$ bitsztringet úgy kapjuk, hogy minden egyes $v \in \mathcal{V}$ csúcsára a címke 1 lesz abban a pozícióban, mint ahol a $\mathcal{B}(v)$ az.

Az eredeti \mathcal{W} csúcshalmaz kettéosztása után a kettéosztási pontokat az X halmazban tároljuk. Az összes generált pontot $x \in X$ -ben felcímkézzük, amelyek a partíciós halmazok csúcsaiként szolgálnak. A x pont címkéje megegyezik a φ legkisebb dimenziós felületének címkéjével, ami tartalmazza x -et. A kettéosztás során egy új csúcs $x = \frac{v+w}{2}$ címkéje $\mathcal{B}(x) = \mathbf{BitOr}(\mathcal{B}(v), \mathcal{B}(w))$.

Adott egy $S = \text{conv}(\mathcal{V})$ szimplex, a kérdés az, hogy mi a (legkisebb dimenziójú) φ felület címkéje, amelybe beletartozik. Ezt a $\mathcal{B}(\varphi) = \mathbf{BitOr}(\mathcal{B}(\mathcal{V}))$ címke határozza meg, amelyet az összes csúcscímke bitenkénti vagy műveletével kaphatunk meg. A $\mathcal{B}(\varphi)$ bitsorozat egyeseinek számát $|\mathcal{B}(\varphi)|$ jelöli, ami a φ csúcsainak számát adja meg. A 2.3.1 definíció szerint az S $(h - 1)$ -szimplex akkor határ-szimplex, ha létezik egy $(h - 1)$ -simplex φ lapja az fizibilis halmaznak ($h \leq n$) ami S -t tartalmazza.

2.3.15. állítás. Adott $S = \text{conv}(\mathcal{V})$, $h \leq n$ $(h - 1)$ -szimplexre, ha $|\mathbf{BitOr}(\mathcal{B}(\mathcal{V}))| = h$, akkor az S határ-szimplex.

2.3.4. Eredmények

A szimplex-alapú B&B módszerekben az alsó korlát meghatározása nagy jelentőséggel bír. Több módszert adtunk a gradiensbefoglalást használva. A középponti formulához több alappontot adtunk meg, illetve bemutattuk az affin aritmetika és a lineáris relaxáció használatát az intervallumburok és a szimplex felett is. Ezen túlmenően beláttunk néhány elméleti eredményt a monotonitással kapcsolatban, amelyeket új kivágási tesztek konstruálására alkalmaztunk.

A jól ismert teszt példákon elért eredmények azt mutatják, hogy az affin aritmetika ígéretes módszer a szimplexek feletti alsó korlátok meghatározására. Sok problémánál a legkevesebb szimplex kiértékelést igényli. Számítási ideje azonban nagyobb, mint számos más módszeré. Általában a lineáris programozási megoldót használó módszerek ugyanezt a hátrányt szenvedik el, és több időt igényelnek. Megállapítottuk, hogy a monotonitásvizsgálatok elengedhetetlenek a számítási idő csökkentéséhez.

A legkevesebb számítási időt igénylő módszer a szimplex középponti formuláján alapul, és a szimplex legmagasabb függvényértékkel rendelkező csúcsát használja bázispontként.

3. fejezet

A vásárlói választási szabály modellezése versenyképes vállalat-elhelyezési problémákban

Annak a piaci részesedésnek a becslése, amelyet egy vállalat meg tud szerezni egy olyan versenykörnyezetben, ahol létezik más, ugyanazt a terméket kínáló vállalat, nagyon fontos feladat, mivel egy vállalat túlélése az általa elérhető bevételektől függ, és ezek a bevételek nagymértékben a piaci részesedésen múlnak.

Ehhez azt is meg kell határozni, hogy mi a *vonzási (vagy hasznossági) függvénye* a vásárlónak egy adott vállalat felé. Általában a vonzási függvény függ az ügyfél és a vállalat közötti távolságtól, valamint a vállalat egyéb jellemzőitől, amelyek meghatározzák a *minőségét*. A vevők választási magatartását is figyelembe kell venni. A fogyasztók egynél több vállalatot is felkereshetnek a keresletük kielégítése érdekében. A fogyasztói választási magatartással foglalkozó szakirodalom azokat a kulcsfontosságú változókat vizsgálja, amelyeket egy vásárló figyelembe vesz, amikor egyik vagy másik vállalatot választja, illetve azt, hogy ezek a változók hogyan hatnak egymásra.

A szakirodalomban általánosan használt két fogyasztói választási szabály a következő:

Determinisztikus (vagy bináris) szabály: feltételezi, hogy egy vevő teljes keresletét *csak egy* központ elégíti ki, amelyik a leginkább vonzza, figyelmen kívül hagyva minden más, kevésbé vonzó vállalatot, sőt még azokat is, amelyek vonzereje kicsit tér csak el a maximálisétól.

Valószínűségi szabály: feltételezi, hogy az ügyfél valószínűségi alapon osztja fel a keresletét az *összes* vállalat között, a piacon az egyes vállalatok iránti vonzalmával arányosan.

Az 1.2 alfejezetben bevezetett jelölések alapján, az egyszerű valószínűségi szabály alkalmazása esetén a lánc által megszerzett piaci részesedés a meglévő és az x helyre α minőséggel telepített új vállalatra

$$M_P(x, \alpha) = \sum_{i \in I} w_i \frac{u_{i0}(x, \alpha) + \sum_{j \in J_1} u_{ij}}{u_{i0}(x, \alpha) + \sum_{j \in J} u_{ij}}.$$

A valószínűségi szabályban a p_i vevők vonzódását egy lánc felé az összes vállalat meghatározza amely a lánchoz tartozik. Amint látható, azt feltételezzük, hogy a hasznosság *additív*: például az első lánc számára a hasznosság a következő: $U_i^1(x, \alpha) = u_{i0}(x, \alpha) + \sum_{j \in J_1} u_{ij}$.

A megoldandó probléma így

$$\begin{cases} \max & \Pi_P(x, \alpha) = F(M(x, \alpha)) - G(x, \alpha) - G^{ex} \\ \text{s.t.} & d_i(x) \geq d_i^{\min} \quad \forall i \\ & \alpha \in [\alpha_{\min}, \alpha_{\max}] \\ & x \in S \subset \mathbb{R}^2, \end{cases} \quad (3.1)$$

ahol az F, G függvények az 1.2 alfejezetben bevezetettek, G^{ex} pedig a meglévő vállalatokhoz tartozó fix költség.

Ehhez a modellhez képest, két újfajta vásárlói választási szabályt vezetünk be.

3.1. A részlegesen valószínűségi választási szabály

Jelöljük most \bar{u} -val a vonzerő minimális szintjét, amellyel egy vállalatnak rendelkeznie kell ahhoz, hogy a vásárló a vásárlóereje egy részét ott költse el. Az egyszerűség kedvéért feltételezzük, hogy ez a minimális szint az összes keresleti pontra nézve azonos, de minden egyes i keresleti pontra más \bar{u}_i értéket is megadhatunk. Legyen

$$\tilde{u}_{ij} = \begin{cases} u_{ij} & \text{ha } u_{ij} \geq \bar{u} \\ 0 & \text{máskülönben,} \end{cases}$$

illetve

$$\tilde{u}_{i0}(x, \alpha) = \begin{cases} u_{i0}(x, \alpha) & \text{ha } u_{i0}(x, \alpha) \geq \bar{u} \\ 0 & \text{máskülönben.} \end{cases}$$

Ekkor a módosított választási szabály szerint a lánc által megszerzett piaci részesedés a következő:

$$M_{PP}(x, \alpha) = \sum_{i \in I_{PP}} w_i \frac{\tilde{u}_{i0}(x, \alpha) + \sum_{j \in J_1} \tilde{u}_{ij}}{\tilde{u}_{i0}(x, \alpha) + \sum_{j \in J} \tilde{u}_{ij}}, \quad (3.2)$$

ahol

$$I_{PP} = \{i \in I : \max\{u_{i0}(x, \alpha), \max\{u_{ij} : j \in J\}\} \geq \bar{u}\}.$$

Vegyük észre, hogy ha egy adott i keresleti pont esetében $i \notin I_{PP}$, azaz

$$\max\{u_{i0}(x, \alpha), \max\{u_{ij} : j \in J\}\} < \bar{u},$$

akkor az i pont keresletét egyetlen vállalat sem szolgálja ki. Ezért ebben a modellben a kereslet egy része kielégítetlen maradhat.

A megfelelő folytonos versenyképes vállalat-elhelyezési és minőség-tervezési probléma megegyezik a (3.1) problémával, kivéve, hogy $M(x, \alpha)$ helyébe $M_{PP}(x, \alpha)$ lép, ahogyan azt a (3.2) adja.

Ismert, hogy ezek a problémák nehezen megoldható globális optimalizálási feladatok, sok olyan lokális maximummal, amelyek nem globális optimumok [17]. A megfelelő probléma a módosított választási szabállyal még nehezebb, mivel ezen kívül sok szakadást is tartalmaz.

3.2. Multi-determinisztikus kiválasztási szabály

Ebben a szakaszban egy, a fogyasztói magatartás egy másik aspektusán alapuló fogyasztói választási szabályt vezetünk le. Vegyük például egy olyan vásárló esetét, akinek a heti bevásárlást kell elvégeznie. Van öt szupermarket a lakóhelye körül, amelyek közül kettő az A lánc, a másik három pedig egy másik, B lánc, B lánc, B lánchoz tartozik. Valószínűleg nem fogja az összes heti bevásárlását egyetlen szupermarketben intézni, mivel egyes termékek nem biztos, hogy elérhetők, vagy azok ára alacsonyabb a másik lánc szupermarketjeiben. Nem fog azonban az összes szupermarketbe elmenni, mivel ugyanazokat a termékeket, akár azonos áron is, az azonos láncba tartozó szupermarketekben fogja megtalálni. Ezért úgy dönt, hogy az A lánc egyik szupermarketjébe és a B lánc egyik szupermarketjébe megy. Az egyes láncok közül azt a szupermarketet fogja választani, amelyik a leginkább vonzza. A heti bevásárlását pedig nem 50%-ban fogja elvégezni ebben a két szupermarketben: valószínűleg több pénzt fog költeni abban a szupermarketben, amelyikhez nagyobb vonzalmat érez. Az általunk bevezetett multi-determinisztikus szabályt amit a következőkben bemutatunk, ezt a viselkedést próbálja modellezni.

Az előzőkhöz hasonlóan a síkban egyetlen vállalat elhelyezésének problémáját vizsgáljuk, statikus verseny mellett és rugalmatlan kereslettel, ahol a hasznosságfüggvény függ a vállalatok elhelyezkedésétől és minőségétől is. Ez utóbbi két tényező a probléma változó. A cél ismét a következő módon kapott profit maximalizálása a lánc által megszerzett piaci részesedésből származó bevétel, csökkentve a lánc működési költségeivel. A korábbiakhoz hasonlóan több vállalkozás van jelen a piacon, de most a vásárlók a keresletüket a vállalkozások között úgy osztják meg, hogy a következő vállalkozásokat választják: minden cégtől csak egy vállalatot vesznek igénybe, azt, amelyiknek a legnagyobb a haszna, és a kereslet megoszlik ezen vállalatok között a vonzerejük arányában.

Jelölések

- c a versenyző láncok indexe, $c \in C = \{1, \dots, c_{\max}\}$ ahol $c = 1$ a bővítő lánc, a $c \in C$ láncokhoz tartozó vállalatok $j \in J_c = \{j_{\min}^c, \dots, j_{\max}^c\}$, ahol $j_{\min}^1 = 1$, és j_{\max}^1 a bővítő lánc vállalatainak a száma, $j_{\max}^c = j_{\min}^{c+1} - 1$ és $j_{\max}^{c_{\max}} = j_{\max}$. Természetesen $J = \cup_{c \in C} J_c$
- u_i^c c lánc vállalatainak maximális hasznossága p_i számára, $u_i^c = \max\{u_{ij} : j \in J_c\}$.

A lánc ($c = 1$ lánc) által megszerzett piaci részesedés így a következő:

$$M_{MD}(x, \alpha) = \sum_{i \in I} w_i \frac{\max\{u_{i0}(x, \alpha), u_i^1\}}{\max\{u_{i0}(x, \alpha), u_i^1\} + \sum_{c=2}^{c_{\max}} u_i^c}. \quad (3.3)$$

Amint az a képletből is látható, itt azt feltételezzük, hogy a p_i -ben lévő vásárlók vonzódását egy lánc felé csak az a vállalat határozza meg, amelyikhez a legjobban vonzódnak. A lánc többi vállalata nem játszik szerepet. A determinisztikus szabálytól eltérően azonban most az összes a lánc megkapja a p_i keresletének egy részét.

Az új vállalat által megszerzett piaci részesedés

$$m_{MD_0}(x, \alpha) = \sum_{i \in I} w_i \frac{\tilde{u}_{i0}(x, \alpha)}{\max\{u_{i0}(x, \alpha), u_i^1\}} + \sum_{c=2}^{c_{\max}} u_i^c,$$

ahol

$$\tilde{u}_{i0}(x, \alpha) = \begin{cases} u_{i0}(x, \alpha) & \text{ha } u_{i0}(x, \alpha) \geq u_i^1 \\ 0 & \text{máskülönben.} \end{cases}$$

A megfelelő folytonos, versenyző vállalat-elhelyezési és minőség-tervezési probléma az (3.1), ahol $M(x, \alpha)$ a (3.3) által adott. Ez a probléma ismét egy erősen nemlineáris optimalizálási feladat, amelynek megoldásához globális optimalizálási technikákra van szükség. Megjegyezzük, hogy ha az egyes versengő láncok vállalatainak száma egy, és az elhelyező lánc új belépő, azaz nincs már létező vállalata, a modell a [17]-ben bemutatott standard valószínűségi modellre redukálódik.

3.3. Eredmények

Elemeztük a választási szabály hatását a versenyképes vállalatok optimális elhelyezésével kapcsolatban mind a részlegesen véletlenszerű mind a multi-determinisztikus szabályt vizsgálva. Megfogalmaztuk a profit-maximalizálás megfelelő elhelyezési problémáját, és egy szigorú intervallumos B&B módszert alkalmaztunk a probléma megoldására.

Megmutattuk a számítási eredmények segítségével, hogy az új vállalat optimális elhelyezkedése és minősége, valamint az új vállalat által elért nyereség jelentősen eltérhet az alkalmazott vásárlói választási szabály függvényében. Ezért körültekintően kell eldönteni, hogy a valós alkalmazásokban mi legyen az alkalmazandó választási szabály.

A versenyképes vállalat-elhelyezés nem csak azért nehéz terület, mert meglehetősen összetett matematikai modelleket foglal magában, hanem azért is, mert a vásárlói magatartás nem írható át könnyen formulákká. A modellek csak közelítést adják a valóságnak.

Irodalomjegyzék

- [1] B. Adenso-Díaz and F. Rodríguez. A simple search heuristic for the MCLP: Application to the location of ambulance bases in a rural region. *Omega*, 25(2):181–187, 1997.
- [2] A. Andrade, J. Comba, and J. Stolfi. Affine arithmetic. *International Conf. on Interval and Computer-Algebraic Methods in Science and Engineering (INTERVAL/94)*, 1994.
- [3] F. Aurenhammer. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405, 1991.
- [4] E. Baumann. Optimal centered forms. *BIT Numerical Mathematics*, 28(1):80–87, 1988.
- [5] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan. Mixed-integer nonlinear optimization. *Acta Numerica*, 22:1–131, 2013.
- [6] T. Benoist, B. Estellon, F. Gardi, R. Megel, and K. Nouioua. LocalSolver 1.x: a black-box local-search solver for 0-1 programming. *4OR*, 9(3):299, 2011.
- [7] O. Berman, Z. Drezner, and D. Krass. Generalized coverage: New developments in covering location models. *Computers & Operations Research*, 37(10):1675–1687, 2010.
- [8] A. Bonn, A. S. L. Rodrigues, and K. J. Gaston. Threatened and endemic species: are they good indicators of patterns of biodiversity on a national scale?: Threat, endemism and biodiversity. *Ecology Letters*, 5(6):733–741, 2002.
- [9] L. G. Casado, E. M. T. Hendrix, and I. García. Infeasibility spheres for finding robust solutions of blending problems with quadratic constraints. *Journal of Global Optimization*, 39(4):577–593, 2007.
- [10] C.-H. Chung. Recent applications of the maximal covering location planning (M.C.L.P.) model. *Journal of the Operational Research Society*, 37(8):735–746, 1986.
- [11] R. Church and C. ReVelle. The maximal covering location problem. *Papers of the Regional Science Association*, 32(1):101–118, 1974.
- [12] R. L. Church and M. E. Meadows. Location modeling utilizing maximum service distance criteria. *Geographical Analysis*, 11(4):358–373, 1979.
- [13] M. S. Daskin and L. K. Dean. Location of health care facilities. In M. Brandeau, F. Sainfort, and W. Pierskalla, editors, *Operations Research and Health Care*, volume 70 of *International Series in Operations Research & Management Science*, pages 43–76. Springer, Boston, MA, 2005.
- [14] C. D’Aspremont, J. Gabszewicz, and J. Thisse. On Hotelling’s stability in competition. *Econometrica*, 47:1145–1150, 1979.
- [15] L. de Figueiredo and J. Stolfi. Affine arithmetic: Concepts and applications. *Numerical Algorithms*, 37(1-4):147–158, 2004.
- [16] H. Eiselt. Hotelling’s duopoly on a tree. *Annals of Operations Research*, 40:195–207, 1992.
- [17] J. Fernández, B. Pelegrín, F. Plastria, and B. Tóth. Solving a Huff-like competitive location and design model for profit maximization in the plane. *European Journal of Operational Research*, 179(3):1274–1287, 2007.

- [18] J. Fernández, B. Tóth, L. Cánovas, and B. Pelegrín. A practical algorithm for decomposing polygonal domains into convex polygons by diagonals. *Top*, 16(2):367–387, 2008.
- [19] J. Gabszewicz and J. Thisse. Location. In R. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, pages 281–304. Elsevier Science Publishers, 1992.
- [20] R. D. Galvão and C. ReVelle. A Lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, 88(1):114–123, 1996.
- [21] M. García, P. Fernández, and B. Pelegrín. On price competition in location-price models with spatially separated markets. *Top*, 12(2):351–374, 2004.
- [22] M. García and B. Pelegrín. All Stackelberg location in the Hotelling’s duopoly model on a tree with parametric prices. *Annals of Operations Research*, 122:177–192, 2003.
- [23] E. Hendrix, B. Tóth, F. Messine, and L. Casado. On derivative based bounding for simplicial branch and bound. *RAIRO - Operations Research*, 55(3):2023–2034, 2021.
- [24] H. Hotelling. Stability in competition. *Economic Journal*, 39:41–57, 1929.
- [25] R. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer, Dordrecht, 1996.
- [26] R. Kearfott, M. Nakao, A. Neumaier, S. M. Rump, S. P. Shary, and P. van Hentenryck. Standardized notation in interval analysis. *TOM*, 15(1):7–13, 2010.
- [27] M. Megiddo and K. Supowit. On the complexity of some common geometric location problems. *SIAM Journal on Computing*, 13:182–196, 1984.
- [28] F. Messine. Extensions of affine arithmetic: application to unconstrained global optimization. *Journal of Universal Computer Science*, 8(11):992–1015, 2002.
- [29] F. Messine and A. Touhami. A general reliable quadratic form: An extension of affine arithmetic. *Reliable Computing*, 12(3):171–192, 2006.
- [30] M. Osborne and C. Pitchik. Equilibrium in Hotelling’s model of spatial competition. *Econometrica*, 55:911–922, 1987.
- [31] B. Pelegrín, P. Dorta-González, and P. Fernández. Finding location equilibria for competing firms under delivered pricing. *Journal of the Operational Research Society*, 62:729–741, 2010.
- [32] B. Pelegrín, J. Fernández, and B. Tóth. The 1-center problem in the plane with independent random weights. *Computers & Operations Research*, 35(3):737–749, 2008.
- [33] L. Rall. *Automatic Differentiation, Techniques and Applications*, volume 120 of *Lecture Notes in Computer Science*. Springer, Berlin, 1981.
- [34] S. M. Rump and M. Kashiwagi. Implementation and improvements of affine arithmetic. *Nonlinear Theory and Its Applications*, 6(3):341–359, 2015.
- [35] N. V. Sahinidis. *BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Programs*, User’s Manual, 2017. <http://www.minlp.com/downloads/docs/baron%20manual.pdf> [Accessed 17-03-2022].
- [36] D. Schilling, V. Jayaraman, and R. Barkhi. A review of covering problem in facility location. *Location Science*, 1(1):25–55, 1993.
- [37] D. Serra and C. ReVelle. Competitive location and pricing on networks. *Geographical Analysis*, 31(2):109–129, 1999.
- [38] J. Stolfi and L. de Figueiredo. *Self-Validated Numerical Methods and Applications*. Monograph for 21st Brazilian Mathematics Colloquium. IMPA/CNPq, 1997.
- [39] E. Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematical Journal*, 43:355–386, 1937.
- [40] P. D. Wright, M. J. Liberatore, and R. L. Nydick. A survey of operations research models and applications in homeland security. *INFORMS Journal on Applied Analytics*, 36(6):514–529, 2006.
- [41] S. Zhang. On a profit maximizing location mode. *Annals of Operations Research*, 103:251–260, 2001.