EGERVÁRY JENŐ ÉLETE ÉS MUNKÁSSÁGA (1891-1958) [1]

RAPCSÁK TAMÁS
MTA SZTAKI

1891. április 16-án született Debrecenben. Középfokú tanulmányait a debreceni Állami Főreáliskolában végezte 1901-1909 között, ott tett érettségi vizsgát az összes tárgyból jeles eredménnyel. Az 1909-1913 években a budapesti Tudományegyetem Bölcsészeti Karán végzett tanulmányai alapján középiskolai tanári alap- és szakvizsgát tett, az összes tárgyból kitüntetéssel. 1913 nyarán állami szünidei ösztöndíjjal angliai tanulmányúton vett részt. 1914-ben bölcsészetdoktori szigorlatot tett mennyiségtanból, elméleti fizikából, valamint kozmográfiából, summa cum laude eredménnyel. Az 1913-14 tanévben matematikai pályamunkájáért a Pasquich-alapból jutalomdíjat nyert.

Az 1914-1917 években a budapesti Földrengési Obszervatóriumban dolgozott tanársegédként. 1917-ben pedagógiai vizsgát tett és középiskolai tanári oklevelet szerzett mennyiségtan-természettan tárgycsoportra. 1918-ban nevezték ki a budapesti Állami Felsőipariskola rendes tanárává. Az 1921-1926 években a Budapesten működő kolozsvári Tudományegyetemen magántanárként tevékenykedett. 1932-től kezdődően a budapesti Középiskolai Tanárképző Intézet előadó tanára a matematikai-fizika differenciálegyenletei című tárgykörben.

1932-ben matematikai munkásságáért König Gyula jutalomban részesült. „Analízis és annak matematikai-fizikai alkalmazásai" című dolgozatát nyújtotta be magántanári pályázatként, aminek alapján 1938-ban a budapesti Tudományegyetem magántanárává, majd 1941-ben a budapesti Műegyetem nyilvános rendes tanárává nevezték ki. A Műegyetem 1955-ben bekövetkezett kettéválása után az Építőipari és Közlekedési Műszaki Egyetem Matematika Tanszékének a vezetője volt 1958. október 15-i nyugdíjazásáig. Egyidejűleg, az Eötvös Loránd Tudományegyetemen —éveken keresztül— tartott előadásokat differenciálegyenletekről az alkalmazott matematika szakos hallgatóknak.

A Magyar Tudományos Akadémia 1943-ban levelező, majd 1946-ban rendes tagjává választotta. Különösen az alkalmazott matematika széles körben való elterjesztése terén szerzett elévülhetetlen érdemeket; tudományos munkássága elismeréséül két alkalommal Kossuth-díjjal tüntették ki.

Egerváry Jenő tudományos munkásságára jellemző szerteágazó érdeklődési köre. Első eredményei Fejér Lipót kutatásaihoz kapcsolódnak; számos dolgozatot írt az analízis és függvénytan köréből, közben azonban figyelme

[1] Beérkezett: 2002. március 29. A kutatás az OTKA T029572 számú szerződés keretében folyt.

az algebrai egyenletek felé is fordult. Már első dolgozataira jellemző, hogy a determinánselmélet sok helyütt fontos és hasznos segédeszköz szerepét tölti be. Ugyanakkor jelentős érdeklődést tanúsított geometriai és elméleti fizikai kérdések iránt. 1938-tól kezdve mintegy tizenöt éven keresztül egymás után közölte eredményeit a geometria és a differenciálegyenletek tárgyköréből. A geometrián belül egy, az alkalmazásokhoz és a méréshez kapcsolódó fontos és mély kérdéssel, a görbék metrikus görbületével is foglalkozott [3], [4], [5]. Életének utolsó hat évében azután szinte kizárólag a mátrixelméleti kutatásoknak szentelte figyelmét, nagy hangsúlyt helyezve az alkalmazásokra.

Egerváry valamennyi munkáját —miként előadásait is— a világosság, szabatosság, valamint az eleganciára való törekvés jellemzi. Stílusa tömör, nem használ felesleges szavakat, de ez mégsem megy az érthetőség rovására. Egyik legfőbb eszköze, aminek segítségével mindig közel tudja hozni az olvasót, illetve hallgatót a tárgyhoz, a szemléltetés. Másik jellemző vonás, ami csaknem valamennyi művén végigvonul, az az igény, hogy eredményeinek —legyenek azok a legelvontabb területről valók is— az alkalmazási körét is megtalálja és megmutassa. Az elméletnek és a gyakorlatnak éppen ez az összefonódása késztette arra, hogy 1947-ben tevékenyen részt vegyen a Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének a megalapításában, ahol a „Mechanikai és Szilárdságtani Osztály" vezetője, majd az Intézet Matematikai Kutató Intézetté való átszervezése után a „Mátrixelmélet és Alkalmazási Osztály" vezetője volt.

Egerváry művei között különleges helyet foglal el a mátrixok kombinatorikus tulajdonságaival foglalkozó dolgozat, egyrészt tárgyánál, másrészt a legutóbbi időkig az alkalmazások révén nyert jelentőségénél fogva [2]. Erre a dolgozatra a megjelenése után több mint húsz évvel H. W. Kuhn figyelt fel az Egyesült Államokban, felismerve a dolgozatban közölt tétel alkalmazási lehetőségét az akkor új diszciplínának számító operációkutatás hozzárendelési feladatának megoldására. A hozzárendelési feladat speciális egész értékű lineáris optimalizálási feladat, aminek igen sok gyakorlati alkalmazása van. Egy példa az, ha adott bizonyos számú elvégezendő munka és ugyanannyi dolgozó, akik a munkákat különböző költségekkel tudják végrehajtani, és a dolgozók között az összes munkát úgy kell szétosztani, hogy minden dolgozó pontosan egy munkát kapjon, és a munkavégzés összköltsége minimális legyen.

A hozzárendelési feladat megoldó algoritmusát H. W. Kuhn publikálta 1955-ben, ami magyar módszer néven vált ismertté [23]. A névadás H. W. Kuhn dolgozatát [25] és Komlósi Sándor 1992-ben a Szigma folyóiratban publikált fordítását [26] alapul véve, a következőképpen történt:

1953 nyarán a Nemzeti Mérésügyi Hivatal és más amerikai kormányzati szervek megalakították egy kiváló kombinatorikusokból és algebristákból álló csoportot a Kaliforniai Egyetemen, ahol abban az időben az egyik legjobbnak tartott számítógép, a Standards Western Automatic Computer (SWAC) működött. Ismeretes, hogy az akkori gépekben a memória volt a szűk kapacitás, a SWAC egész memóriája mindössze 256 darab Williamson katódcsőből állt. Azon a nyáron Kuhn egyik kollégája, C. B. Tompkins, $10 \times$

10-es hozzárendelési feladatokat próbált a SWAC segítségével és az összes $10! = 3\,628\,800$ permutáció leszámlálásával megoldani, de egyetlen próbálkozása sem járt sikerrel. Mint ismeretes, egy 10×10-es hozzárendelési probléma olyan lineáris optimalizálási feladatra vezet, amelyben 20 egyenlőség feltétel és 100 nemnegatív változó van. 1953-ban ilyen méretű feladatot még nem tudtak megoldani.

Kuhn ebben az időszakban König gráfelméleti könyvét [22] olvasva rájött, hogy egy gráf két n szögpontú részre való felosztása és a két rész közötti párosítás problémája pontosan megegyezik egy olyan $n \times n$-es hozzárendelési feladattal, amelyben a költségmátrix elemei binárisak. De ami még ennél is fontosabb, Kuhn észrevette, hogy König megadott egy olyan algoritmust, amellyel meghatározható volt a fenti párosítási probléma optimális megoldása. Ezek után annak a kérdésnek a megoldása maradt hátra, hogyan lehet az általános hozzárendelési feladatot bináris költségmátrixú hozzárendelési feladatra visszavezetni. Kuhn felfigyelt egy lábjegyzetre König könyvében, ami Egerváry egy 1931-ben írt cikkére utalt, és ráérzett, hogy a fenti probléma megoldásának a kulcsa Egerváry cikkében található meg.

Beszerezte Egerváry cikkét, egy nagy magyar szótár és egy nyelvtankönyv kíséretében a Bryn Mawr kollégiumból, két hétig tanult magyarul és közben lefordította Egerváry cikkét angolra, amiben megtalálta König tételének az általánosítását és az általános hozzárendelési feladat visszavezetésének a módszerét bináris költségmátrixú feladatra. Felhasználva Egerváry redukciós eljárását és König maximum-párosítási algoritmusát, Kuhn 1953 őszén számos 12×12-es hozzárendelési feladat optimális megoldását számította ki kézzel. Mindegyik feladatot két órán belül megoldotta és ez meggyőzte arról, hogy a kombinált algoritmus jó. A történet ezen részéhez Kuhn az alábbi kedves megjegyzést fűzte:

> „Valószínűleg ez egyike volt azon legutolsó eseteknek, amikor papírral és tollal le lehetett győzni a világ legnagyobb és leggyorsabb elektronikus számítógépét."

Mivel a hozzárendelési probléma megoldására javasolt algoritmus két magyar matematikus, König D. és Egerváry J. eredményeire épült, ezért Kuhn az eljárást magyar módszernek nevezte el, és a szakirodalom ma is így ismeri. A Magyar Operációkutatási Társaság Kuhn professzort a hozzárendelési probléma megoldásában elért eredményeiért, az 1992. május 22-én megtartott közgyűlésén, tiszteletbeli tagjává választotta.

Egy évvel a magyar módszer megjelenése után kiderült, hogy a kidolgozott technika alkalmas a szállítási feladat megoldására is, mint ahogy ez L. R. Ford és D. R. Fulkerson 1956-ban publikált cikkében [20], valamint Egerváry [16] és [17] dolgozataiban megtalálható. Tekintettel az említett kapcsolatra, ez az eljárás is magyar módszer néven vált ismertté. Ezekről a témakörökről részletesebben lehet olvasni az [1] és [21] tankönyvekben.

Egerváry 1931-es cikkében szereplő fő eredményének jelentősége Frank A. [19] szerint:

1. A tétel a lineáris optimalizálás dualitás tételének a legelső explicit megfogalmazása speciális mátrixokra vonatkoztatva.

2. A dualitás tétel olyan esetét tárgyalja, amelyben mindig létezik egész értékű optimum is.

3. A konstruktív bizonyítás rámutat az algoritmusok —mint önálló matematikai objektumok— jelentőségére, szoros összefüggésben az algoritmusok hatékonyságával. Bár 1931-ben e fogalmak értelemszerűen fel sem vetődhettek, Egerváry módszere elméletileg és gyakorlatilag is igen hatékony. Amint azt jóval később kimutatták, valójában polinomiális, sőt erősen polinomiális futásidejű.

4. Egerváry tétele és eljárása kristálytisztán megmutatja, hogy egy szép matematikai eredmény miként nyújt megoldást számtalan, természetesen felvetődő gyakorlati kérdésre.

Egerváry és König az operációkutatáson, de talán mondhatjuk, hogy a matematikán belül sikertörténetet indított el, két addig különálló terület, a mátrixelmélet és a kombinatorika összekapcsolásával, ami nagyon szép kutatási irányt és hatékony algoritmusokat eredményezett. Az az út, amit Egerváry és König említett tételei a gráfelméleti megfogalmazástól a szállítási probléma megoldásáig befutottak, igen szépen és meggyőzően példázza az elmélet és a gyakorlat egységét, és azt a kölcsönhatást, ami mindkettő fejlődését szükségképpen előreviszi.

Egerváry egy másik kiemelkedő eredménye, ami ma is alapvető az operációkutatás és az alkalmazott matematika algoritmikus területein, a rangcsökkentő eljárás. Egerváry lineáris egyenletrendszerek megoldásával foglalkozott a [8], [10], [12], [13], [14] és [15] dolgozatokban. Érdekes és egyben jellemző, hogy a matematika, ezen sokak által teljesen lezártnak vélt területét számos új eredménnyel gazdagította. Az 1956-ban és 1960-ban megjelent [12] és [15] dolgozatokban olyan rangcsökkentő eljárást dolgozott ki, amely a diadikus felbontás általánosításának tekinthető, és lehetőséget ad a lineáris egyenletrendszerek véges iterációval történő egyszerű megoldására. Ehhez szükséges és elegendő feltételt kellett adni arra vonatkozóan, hogy egy mátrixból kivonva egy diádot, mikor csökken pontosan eggyel a mátrix rangja.

A [11] dolgozatban Egerváry a fenti rangcsökkentő eljárást alkalmazta homogén lineáris diofantoszi egyenletrendszerek megoldására, a [6] és [7] dolgozatokban projektor és Stieltjes mátrixokra. A [13] dolgozatban az inverz mátrix fogalmat tetszőleges (téglalap) alakú mátrixra terjesztette ki. A [18] dolgozatban konstruktív módszert adott kvadratikus mátrixok Jordan féle normálalakba történő átalakítására, ami rangcsökkentő eljárásának egyik legszebb alkalmazása.

Egerváry pedagógiai munkásságában kiemelkedő helyet foglal el a Budapesti Műszaki Egyetemen kifejtett tevékenysége; a mérnök- és gépészmérnök-, majd a közlekedés- és építőmérnökhallgatók számára tartott „analízis" előadásai napjainkban is példaértékűek a felsőfokú műszaki képzés terén. Nyilvánvalóan korábbi oktatói munkája, de méginkább az alkalmazott matematika iránti elkötelezettsége inspirálta olyan előadói stílusra és metodikára, amely a leghatékonyabb módon ismertethette meg a mérnökhallgatókat azzal a matematikai anyaggal, amely a további tanulmányok megalapozásához, illetve a majdani gyakorlati munkájukhoz elengedhetetlennek tűnt.

A bibliográfia közel száz publikációját tartja nyilván (lásd Egerváry Jenő tudományos munkáinak jegyzékét mellékelve), melyek a matematika és az alkalmazások legkülönbözőbb területeihez kapcsolódnak. Aktivitása életének utolsó éveiben nem hogy csökkent volna, ellenkezőleg; állandóan új eredményekkel gazdagította a nemzetközi alkalmazott matematikai irodalmat, megbecsülést szerezve a magyar matematikának világszerte. Hazánk és a matematikus társadalom nagy vesztesége, hogy Egerváry Jenő, akinek most születése száztizedik évfordulójára emlékezünk, alkotóereje teljében önkezével vetett véget életének. Ezzel kívánt pontot tenni egy, a személye ellen 1958-ban indított méltatlan, őt emberi és tudósi méltóságában megalázni hivatott, gazdasági ürügyekbe bújtatott, személyi bosszútól sem mentes hajsza végére.

Születésének századik évfordulójáról a Budapesti Műszaki Egyetem, melynek haláláig tanára volt, megemlékezést tartott és emlékkiállítást is rendezett, de erről, eléggé sajnálatos módon, sokan csak megkésve értesülhettek, s így e rendezvények meglehetősen szűk körre korlátozódtak csupán. A Magyar Operációkutatási Társaság 1991 október 8-án, a XX. Magyar Operációkutatási Konferencia keretében emlékülést rendezett a tiszteletére. A Magyar Operációkutatási Társaság 1992. április 6-án —20 műegyetemi professzor aláírásával támogatva— azzal a kéréssel fordult a Budapesti Műszaki Egyetem rektorához, hogy Egerváry Jenő emlékét megörökítendő, a mellszobrát az egyetem kertjében, a szoborparkban helyezzék el. Az Egyetemi Tanács a javaslatot támogatta, de az mind a mai napig nem valósult meg.

A XXV. Magyar Operációkutatási Konferencia (Debrecen, 2001. október 17-20) plenáris ülésen emlékezett meg Egerváry Jenő születésének 110. évfordulójáról. A plenáris ülés első előadása szólt —a jelenlegi összeállítás alapján— Egerváry Jenő életéről és munkásságáról, a második előadás Egerváry rangcsökkentő eljárásáról (Galántai Aurél) és a harmadik a magyar módszer általánosításairól (Frank András). Egerváry Jenő életéről és munkásságáról bővebb összeállítás található Rózsa Pál [27] és [28] cikkeiben.

Úgy véljük, hogy Egerváry Jenő akadémikus olyan kiemelkedő alakja a magyar és a nemzetközi tudományos életnek, hogy a róla szóló megemlékezés nem korlátozódhat egy-egy szűkebb szakmai társaságra, ezért azzal a kéréssel fordulunk a szélesebben vett szakma képviselőihez, hogy segítsenek Egerváry Jenő emlékének méltó megőrzésében és ápolásában.

Köszönetnyilvánítás

Szeretnék köszönetet mondani Székely-Doby Sándornak, Egerváry Jenő volt aspiránsának, aki értékes tanácsokkal szolgált, komoly segítséget nyújtott az anyaggyűjtésben és aki a Magyar Operációkutatási Társaság 1991. évi Egerváry emlékülésének előadójaként a pedagógust mutatta be.

Egerváry Jenő tudományos munkáinak jegyzéke

1. Az integrálegyenletek egy osztályáról, *Mathematikai és Physikai Lapok* 23 (1914) 301–355
2. A seismikus trajektóriákról s az azokkal kapcsolatos Bertrand-féle problémáról, *Mathematikai és Physikai Lapok* 26 (1917) 1–18.
3. Über seismischen Trajektorien und über das Bertrandsche Problem in der Seismologie, *Gerlands Beiträge zur Geophysik* 14 (1918) 284–299.
4. Über die charakteristischen geometrischen Eigenschaften der Legendreschen und Tschebyscheffschen Polynome, *Archiv der Mathematik und Physik* (3) 27 (1918) 17–24.
5. On a maximum-minimum problem and its connexion with the roots of equations, *Acta Litterarum ac Scientiarum* 1 (1922) 39–45.
6. Egy asymmetrikus multilineáris formára vonatkozó minimum föladat, *Mathematikai és Physikai Lapok* 29 (1922) 21–43.
7. Über gewisse Extremumprobleme der Funktionentheorie, *Mathematische Annalen* 99 (1928) 542–561.
8. Einige Extremalprobleme im Bereiche der trigonometrischen Polynome, *Mathematische Zeitschrift* 27 (1928) 641–652. (Szász Ottóval együtt)
9. A trinom egyenletről, *Matematikai és Fizikai Lapok* 37 (1930) 36–57.
10. On a generalisation of a theorem of Kakeya, *Acta Litterarum ac Scientiarum* 5 (1931) 78–82.
11. Matrixok kombinatorikus tulajdonságairól, *Matematikai és Fizikai Lapok* 38 (1931) 16–28.
12. Verschärfung eines Harnackschen Satzes und anderer Abschätzungen für nichtnegative harmonische Polynome, *Mathematische Zeitschrift* 34 (1932) 741–757.
13. Jelentés az 1934. évi König Gyula-jutalomról, *Matematikai és Fizikai Lapok* 41 (1934) 93–102.
14. Szász Pál: A differenciál- és integrálszámítás elemei (Könyvismertetés), *Matematikai és Fizikai Lapok* 42 (1935) 247–248.
15. Abbildungseigenschafften der arithmetischen Mittel der geometrischen Reihe, *Mathematische Zeitschrift* 42 (1937) 221–230.
16. A tetraéderről, *Középiskolai Matematikai és Fizikai Lapok* 14 (1937) 1–4.
17. A magasságponttal bíró tetraéderről, *Matematikai és Fizikai Lapok* 45 (1938) 18–35.
18. Über ein Minimumproblem der Elementargeometrie, *Journal für die reine und angewandte Mathematik* 178 (1938) 174–186.

19. Az elektromágneses térben elektronmozgás differenciálegyenleteiről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 57 (1938) 968–987.
20. On orthocentric simplexes, *Acta Litterarum ac Scientiarum* 9 (1910) 218–226.
21. Über ein räumliches Analogon des Sehnenvierecks, *Journal für die reine und angewandte Mathematik* 182 (1940) 122–128.
22. Az n-méretű euklidesi tér görbéiről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 59 (1940) 787–797.
23. Az n-méretű euklidesi tér görbéinek simulógömbjeiről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 59 (1940) 775–786.
24. Fondements d'une théorie générale de la courbure linéaire, *Commentarii Mathematici Helvetici* 13 (1940) 257–276 (Alexits Györggyel együtt)
25. Azonosságok alkalmazásairól, *Mennyiségtani és Természettudományi Didaktikai Lapok* (1943) 33–41.
26. Seismikus méréseknél alkalmazandó geometriai szerkesztésekről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 61 (1943) 1109- (Gerő L., Pogány B. és Vargha B.-val együtt.)
27. Forgattyús hajtómű keresztfejsebességének maximuma, *Technika* 25 (1944).
28. *Differenciálegyenletek*, Mérnöktovábbképző Intézet Kiadványai, 1945.
29. A remark on the lenght of the circle and on the exponential function, *Acta Scientiarum Mathematicarum* 11 (1946) 114–118.
30. On a new form of the differential equations of the problem of three bodies, *Hungarica Acta Mathematica* 1 (1946) 1–18.
31. On a generalization of the solution of the Lagrange problem of three bodies, *Doklady Akademii Nauk SSSR* 55 (1947) 805–807 (in Russian).
32. Sur une nouvelle solution particuliére du probléme des trois corps, *Commentarii Mathematici Helvetici* 24 (1950) 1–3.
33. On a generalisation of a theorem of Sylvester, *Hungarica Acta Mathematica* 1 (1947) 53–57.
34. *A mechanika differenciálegyenleteiről*, Mérnöktovábbképző Intézet Kiadványai, Budapest, 1948.
35. A Rayleigh-módszer alkalmazása forgó rendszerek kritikus szögsebességének megállapításánál, *Matematikai Lapok* 1 (1949) 16–26.
36. On the smallest convex cover of a simple arc of space-curve, *Publicationes Mathematicae* 1 (1949) 65–70.
37. On the mapping of the unit-circle by polynomials, *Acta Scientiarum Mathematicarum* 12 (1950) 226–230.

38. A remark on the curvature and tortuosity of space-curves, *Acta Mathematica Academiae Scientiarum Hungaricae* 1 (1950) 46–47.

39. On the Feuerbach-spheres of an orthocentric simplex, *Acta Mathematica Academiae Scientiarum Hungaricae* 1 (1950) 5–16.

40. Eine Bemerkung über definite quadratische Formen, *Publicationes Mathematicae* (1950) 193–195.

41. Az ortocentrikus koordináta-rendszerről és annak néhány alkalmazásáról, *Magyar Matematikai Kongresszus Közleményei* (1952) 387–396.

42. A matematika gyakorlati alkalmazásai, különös tekintettel a technika differenciálegyenleteire, *A Magyar Tudományos Akadémia III. Matematikai és Fizikai Osztályának Közleményei* 1 (1951) 101–115.

43. On a certain point of the kinetical gas theory, *Studia Mathematica* 12 (1951) 170- (Turán Pállal együtt).

44. A kinetikus gázelmélet bizonyos kérdéseiről, *A Magyar Tudományos Akadémia III. Osztályának Közleményei* 1 (1951) 303–314. (Turán Pállal együtt).

45. A hővezetési differenciálegyenlet megoldása az időtől lineárisan függő kerület mellett, *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 1 (1952) 11–22. (Lovass-Nagy Viktorral együtt).

46. Matrix-függvények kanonikus előállításáról és annak néhány alkalmazásáról, *A Magyar Tudományos Akadémia III. Matematikai és Fizikai Osztályának Közleményei* 3 (1953) 417–458.

47. On a property of the projector matrices and its application to the canonical representation of matrix functions, *Acta Scientiarum Mathematicarum* 15 (1953) 1–6.

48. On a lemma of Stieltjes on matrices, *Acta Scientiarum Mathematicarum* 15 (1953) 99–103.

49. Stieltjesnek egy mátrixelméleti lemmájáról, *Matematikai Lapok* 7 (1956) 271- .

50. On the hermitian normalform of a matrix and Sylvester's law of nullity, *Publicationes Mathematicae* 3 (1953) 144–149.

51. Matrixok diadikus előállításán alapuló módszer bilineáris alakok transzformációjára és lineáris egyenletrendszerek megoldására, *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 2 (1953) 11–32.

52. On the contractive linear transformations of n-dimensional vector space, *Acta Scientiarum Mathematicarum* 15 (1954) 178–182.

53. On hypermatrices whose blocks are commutable in pairs and their application in lattice-dynamics, *Acta Scientiarum Mathematicarum* 15 (1954) 211–222.

54. Páronként felcserélhető blokkokból álló hipermatrixokról és azok alkalmazásáról a rácsdinamikában, *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 3 (1954) 31–47.

55. Über die Faktorisation von Matrizen und ihre Anwendung auf die Lösung von linearen Gleichungssystemen, *Zeitschrift für angewandte Mathematik und Mechanik* 35 (1955) 111–118.

56. On the application of the matrix theory to the calculation of chain-bridges, *Acta Technica Academiac Scientiarum Hungaricae* 11 (1955) 241–256.

57. A matrix-elmélet alkalmazása lánchidak számítására, *A Magyar Tudományos Akadémia III. Matematikai és Fizikai Osztályának Közleményei* 3 (1954) 9–23.

58. A mátrix-elmélet alkalmazása lánchidak számítására, *A Magyar Tudományos Akadémia III. Matematikai és Fizikai Osztályának Közleményei* 5 (1955) 301–313.

59. A Hunyadi-Scholtz-féle matrixok alkalmazása a rácsos szerkezetek elméletében, *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 3 (1954) 289–300.

60. Auflösung eines homogenen linearen diophantischen Gleichungssystems mit Hilfe von Projektormatrizen, *Publicationes Mathematicae* 4 (1955-56) 481–483.

61. Remarques algébriques sur la solution donnée par M. Fréchet in l'équation de Kolmogoroff, II., *Publicationes Mathematicae* 5 (1957) 60–71. (Aczél Jánossal együtt).

62. Régi és új módszerek lineáris egyenletrendszerek megoldására, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 109–122.

63. Az inverz matrix általánosítása, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 315–324.

64. Begründung und Darstellung einer allgemeinen Theorie der Hängebrücke mit Hilfe der Matrizenrechnung, *Abhandlungen der Internationalen Vereinigung für Brückenbau und Hochbau (Zürich)* 16 (1956) 149–184.

65. A függőhidak általános elméletének megalapozása és felépítése matrixszámítás segítségével, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 2 (1957) 3–32.

66. A differenciálokról, *Matematikai Lapok* 8 (1957) 79–85.

67. Über einige Anwendungen von Hypermatrizen, deren Blöcke vertauschbar sind, *Zeitschrift für angewandte Mathematik und Mechanik* 37 (1957) 1-2.

68. Über eine Verallgemeinerung der Purcellschen Methode zur Auflösung linearer Gleichungssystems, *Österreichisches Ingenieur-Archiv* 11 (1957) 249–251.

69. On rank-diminishing operations and their applications to the solution of linear equations, *Zeitschrift für angewandte Mathematik und Physik* 11 (1960).

70. Bemerkungen zum Transportproblem, *MT W-Mitteilungen, Mathematisches Labor der Technischen Hochschule in Wien* 5 (1958) 278–284.

71. Kombinatorikus módszer a szállítási probléma megoldására, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 4 (1959) 15–28.

72. Notes on interpolation V. (On the stability of intrpolation), *Acta Mathematica Academiae Scientiarum Hungaricae* 9 (1958) 259–267. (Turán Pállal együtt).

73. Notes on interpolation VI. (On the stability of the intrpolation on an infinite interval), *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959) 55–62. (Turán Pállal együtt)

74. On the application of matrices and hypermatrices in engineering analysis, *Mémoires et Publications de la Société des Sciences, des Arts et des Lettres du Hainaut* (1959) 44–57.

75. Über einen konstruktive Methode zur Reduktion einer Matrix auf die Jordansche Normalform, *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959) 31–54.

76. Über eine Eigenschaft der Parabel und des Paraboloids, *Publicationes Mathematicae* 6 (1959) 269–275.

77. Über eine Methode zur numerischen Lösung der Poissonschen Differenzengleichung für beliebige Gebiete, *Acta Mathematica Academiae Scientiarum Hungaricae* 11 (1960) 341–361.

78. Discrete models and matrix methods in engineering mechanics, *Procedings of the Third Congress on Theoretical and Applied Mechanics – Bangalore*, India (1957) 259–276.

Egerváry Jenő egyéb munkái (egyetemi jegyzetek, sokszorosított tanulmányok stb.)

79. *Infinitézimális számítás elemei technikusok számára*, Budapest, Asturias, 1924 (Schrodt Istvánnal együtt).

80. *Analízis és geometria* (Műegyetemi jegyzet, 1. évf. 1. félév 1944/45), Budapest, Sztachora, 1945.

81. Analízis és geometria (Műegyetemi jegyzet, 1. évf. 2. félév 1947/48), Budapest, Diószegi sokszorosító, 1947.
82. Analízis és geometria (Műegyetemi jegyzet, 1. évf. 1. félév 1947/48), Budapest, Diószegi sokszorosító, 1948.
83. Differencálegyenletek (Természettudományi Kar jegyzetei), Budapest, 1951.
84. Differencálegyenletek (Természettudományi Kar jegyzetei), Tankönyvkiadó Jegyzetsokszorosító Üzeme, Budapest, 1952.
85. Matematika 3. (A Budapesti Műszaki Egyetem 2. éves mérnökhallgatói részére), Tankönyvkiadó Jegyzetsokszorosító Üzeme, Budapest, 1952.
86. Az Alkalmazott Matematikai Intézet munkája a matematikai fizika és annak ipari alkalmazásai terén, *A Magyar Tudományos Akadémia III. Matematikai és Fizikai Osztályának Közleményei* 3 (1953) 353–362.
87. Tanulmány a mátrixelméletnek lánchidak számításánál való alkalmazásáról, Készült a Közlekedés és Postaügyi Minisztérium 9. Út és Híd Főosztály megbízásából, Budapest, 1955. (Lovass-Nagy Viktorral együtt).

Irodalom

1. Bajalinov E. és Imreh B., *Operációkutatás*, Polygon, Szeged, 2001.
2. Egerváry J., Matrixok kombinatorius tulajdonságairól, *Matematikai és Fizikai Lapok* 38 (1931) 16–28.
3. Egerváry J., Az n-méretű euklidesi tér görbéiről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 59 (1940) 787–797.
4. Egerváry J., Az n-méretű euklidesi tér görbéinek simulógömbjeiről, *Magyar Tudományos Akadémia Matematikai és Természettudományi Értesítő* 59 (1940) 775–786.
5. Egerváry, J., Fondements d'une théorie générale de la courbure linéaire, *Commentarii Mathematici Helvetici* 13 (1940) 257–276 (Alexits Györggyel együtt)
6. Egerváry, J., On a property of the projector matrices and its application to the canonical representation of matrix functions, *Acta Scientiarum Mathematicarum* 15 (1953) 1–6.
7. Egerváry, J., On a lemma of Stieltjes on matrices, *Acta Scientiarum Mathematicarum* 15 (1953) 99–103.
8. Egerváry J., Matrixok diadikus előállításán alapuló módszer bilineáris alakok transzformációjára és lineáris egyenletrendszerek megoldására, *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 2 (1953) 11–32.
9. Egerváry, J., On combinatorial properties of matrices, translated by H. W. Kuhn, *Logistic Papers* 11, George Washington University (1955) 1–11.
10. Egerváry, J., Über die Faktorisation von Matrizen und ihre Anwendung auf die Lösung von linearen Gleichungssystemen, *Zeitschrift für angewandte Mathematik und Mechanik* 35 (1955) 111–118.

11. Egerváry, J., Auflösung eines homogenen linearen diophantischen Gleichungssystems mit Hilfe von Projektormatrizen, *Publicationes Mathematicae* 4 (1955-56) 481–483.

12. Egerváry J., Régi és új módszerek lineáris egyenletrendszerek megoldására, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 109–122.

13. Egerváry J., Az inverz matrix általánosítása, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 315–324.

14. Egerváry, J., Über eine Verallgemeinerung der Purcellschen Methode zur Auflösung linearer Gleichungssystems, *Österreichisches Ingenieur-Archiv* 11 (1957) 249–251.

15. Egerváry, J., On rank-diminishing operations and their applications to the solution of linear equations, *Zeitschrift für angewandte Mathematik und Physik* 11 (1960).

16. Egerváry, J., Bemerkungen zum Transportproblem, *MTW-Mitteilungen, Mathematisches Labor der Technischen Hochschule in Wien* 5 (1958) 278–284.

17. Egerváry J., Kombinatorikus módszer a szállítási probléma megoldására, *Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 4 (1959) 15–28.

18. Egerváry, J., Über einen konstruktive Methode zur Reduktion einer Matrix auf die Jordansche Normalform, *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959) 31–54.

19. Frank A., A magyar módszer és általánosításai, kézirat, 2001.

20. Ford, L.R. and Fulkerson, D.R., Solving the transportation problem, *Management Science* 3 (1956) 24–32.

21. Imreh B., *Kombinatorikus optimalizálás*, Novadat, Szeged, 1999.

22. König, D., *Theorie der endlichen und unendlichen Graphen: kombinatorische Topologie der Strecken-komplexe*, Akademische Verlagsgessellschaft, Leipzig, 1936.

23. Kuhn, H. W., The Hungarian method for the assigment problem, *Naval Research Logistic Quartely* 3 (1955) 253–258.

24. Kuhn, H. W., Variants of the Hungarian method for assignment problems, *Naval Research Logistic Quarterly* 3 (1956) 253–258.

25. Kuhn, H. W., On the origin of the Hungarian method, in: *History of Mathematical Programming*, Elsevier Science, 1992.

26. Kuhn, H. W., A magyar módszer eredetéről (ford. Komlósi S.), *Szigma* 23 (1992) 113–118.

27. Rózsa P., Egerváry Jenő munkásságáról, *Matematikai Lapok* 3-4 (1959) 195–225.

28. Rózsa P., Jenő Egerváry (1891–1958), a great personality of Hungarian mathematical school, *Periodica Polytechnica* 28 (1984) 287–298.

A MAGYAR MÓDSZER ÉS ÁLTALÁNOSÍTÁSAI [1]

FRANK ANDRÁS
ELTE, Operációkutatási Tanszék

A lineáris programozás dualitás tételének legkorábban előforduló speciális alakja Egerváry Jenő klasszikus tétele teljes páros gráf teljes párosításainak maximális súlyáról. Az elegáns bizonyítás elvezet egy hatékony algoritmushoz, melynek a szállítási problémára kiterjesztett alakja a nemzetközi szakirodalomban a Magyar Módszer nevet viseli. A módszer alapelvét azóta több irányban is általánosították: nempáros gráfok maximális súlyú párosításainak meghatározására, a súlyozott matroid metszet problémára, folyam és szubmoduláris áram feladatokra. Jelen munkában áttekintjük a Magyar Módszer e kiterjesztéseit.

1 Bevezetés

A Matematikai és Fizikai Lapok 1931-es 38-as számában jelent meg Egerváry Jenő: *Matrixok kombinatorius tulajdonságairól* című dolgozata. [A címben szereplő *kombinatorius* szó valószínűleg nem elírás, mert a cikk páratlan oldalainak tetején is ekként szerepel. Ugyanakkor Egerváry egy 1957-ben megjelent dolgozatának [9] irodalomjegyzékében saját cikkére hivatkozva már a *kombinatorikus* szót használja.] A dolgozat fő eredménye Kőnig Dénes páros gráfok maximális elemszámú párosításairól szóló tételének súlyozott változata, melyet eredeti, betűhív alakban idézünk:

1.1 Tétel (Egerváry). *Ha az* $||a_{ij}||$ *n-edrendű matrix elemei adott nemnegatív egészek, úgy a*

$$\lambda_i + \mu_j \geq a_{ij}, \qquad (i, j = 1, 2, \ldots n), \quad (\lambda_i, \mu_j, \text{ nem negatív egészek}) \tag{1}$$

feltételek mellett

$$\min \sum_{k=1}^{n} (\lambda_k + \mu_k) = \max(a_{1\nu_1} + a_{2\nu_2} + \ldots + a_{n\nu_n}) \,, \tag{2}$$

hol $\nu_1, \nu_2, \ldots \nu_n$ *az* $1, 2, \ldots n$ *számok összes permutációit befutják.*

Egerváry a dolgozatban megjegyzi, hogy a tétel akkor is érvényben marad, ha elejtjük az a_{ij} mátrixelemekre rótt egészértékűségi feltevést, és a λ_i, μ_j számokról nem követeljük meg az egészértékűséget.

---

[1] Beérkezett: 2002. március 29. Készült a T029772 és T037547 sz. OTKA pályázatok támogatásával. Operációkutatási Tanszék, Eötvös Loránd Tudományegyetem, Pázmány P. s. 1/c, Budapest, Hungary, H-1117 és Traffic Lab Ericsson Hungary, Laborc u.1, Budapest, H-1037. A szerző tagja az Egerváry Kutatócsoportnak (EGRES). e-mail: `frank@cs.elte.hu`.

Kőnig Dénes érdeme volt, hogy felismerte az eredetileg Frobenius által vizsgált súlyozatlan probléma gráfelméleti jellegét, Frobenius tételét (amelynek ekvivalens alakja a Hall tétel) általánosította, és a konstruktív alternáló utas módszert alkalmazta tételének bizonyítására. Ennek segítségével hatékonyan ki lehet számítani páros gráf maximális elemszámú párosítását és az éleknek pontokkal történő minimális elemszámú lefogását. Akkoriban a gráfelmélet gyerekcipőben járt, így nem csoda, hogy Egerváry a tételét még a mátrixok nyelvén fogalmazta meg. Lényegét tekintve azonban ez a tétel páros gráfokról szól, és az általánosításokhoz is könnyebben eljuthatunk, ha gráfos alakban adjuk meg. Tekintettel arra, hogy az A mátrix a_{ij} elemei költségként értelmezhetők, a továbbiakban a helyett a c jelölést használjuk. Azt is előrebocsátjuk, hogy a célfüggvényt néha költség- néha súlyfüggvénynek nevezzük, az előbbit inkább minimalizáláskor, az utóbbit maximalizáláskor használva.

1.2 Tétel. *Legyen $G = (V, E)$ teljes páros gráf, melynek pontjai két n elemű osztályba vannak sorolva. Legyen továbbá $c : E \to \mathbf{R}_+$ az éleknek egy nemnegatív számokkal történő súlyozása. Ekkor G teljes párosításainak maximális súlya egyenlő a nemnegatív súlyozott lefogások minimális súlyával, ahol nemnegatív súlyozott lefogáson egy olyan $\pi : V \to \mathbf{R}_+$ függvényt értünk, amelyre $\pi(u) + \pi(v) \geq c(uv)$ a gráf minden uv élére fennáll. Amennyiben c egészértékű, úgy az optimális π is választható egészértékűnek.*

A szakirodalomban a maximális súlyú párosítás kérdését gyakran hozzárendelési problémának nevezik. Érdemes még egy ekvivalens alakban megfogalmazni Egerváry tételét, a lineáris programozás nyelvén.

1.3 Tétel. *Jelölje A a G teljes páros gráf pont-él incidencia mátrixát, és legyen c nemnegatív célfüggvény. Ekkor a $\max\{ax : x \geq 0, Ax \leq 1\}$ lineáris program optimuma mindig felvétetik egész (és így $0-1$) vektoron, és van olyan egész optimális megoldás, amelyre $Ax = 1$. Továbbá, ha c egészértékű, úgy a $\min\{\sum(\pi(v) : v \in V), \pi \geq 0, \pi A \geq c\}$ lineáris program optimuma is egészen felvétetik, és a két optimum mindig megegyezik egymással.*

Egerváry eredményének úttörő jellege a következőkben foglalható össze.

1. A tétel a legelső explicit megfogalmazása, mégha speciális mátrixokra is, a lineáris programozás dualitás tételének.

2. A dualitás tételnek olyan esetét írja le, amelyben mindig létezik egészértékű optimum.

3. A konstruktív bizonyítási jelleg rámutat az algoritmusoknak, mint önálló matematikai objektumoknak jelentőségére, összefüggésben az algoritmusok hatékonyságával. Bár 1931-ben e fogalmak értelemszerűen fel sem vetődhettek, Egerváry módszere mind gyakorlatilag, mind elméletileg hatékony. Amint az kimutatható (lásd alább), polinomiális, sőt valójában erősen polinomiális futásidejű.

4. Egerváry tétele és eljárása azt demonstrálja, hogy egy elegáns matematikai eredmény miként nyújthat megoldást természetesen felvetődő gyakorlati kérdésekre.

A Magyar Módszer számos további eljárás kiindulópontja lett. Korai példaként Ford és Fulkerson minimális költségű folyam algoritmusa hozható fel. A hozzárendelési feladatban, szállítási problémában, hálózati folyamoknál fellép az a jelenség, hogy az optimum egész vektoron is felvétetik. Ennek valójában közös gyökere van, éspedig az, hogy a szóban forgó problémához tartozó lineáris program feltételi mátrixa teljesen unimoduláris. A Magyar Módszer gondolatait azonban olyan körülmények között is sikerrel fejlesztették tovább, amikor a feltételi mátrix már nem teljesen unimoduláris, mégis az optimum egészértékűsége biztosítható, ennek messze ható kombinatorikus következményeivel együtt. Mindkét eljárás kiindulópontja J. Edmondstól származik. Az első segítségével meg lehet határozni tetszőleges gráf maximális súlyú párosítását. A második kiterjesztés matroidokra vonatkozik. Ennek segítségével például ki lehet számítani egy élsúlyozott irányítatlan gráf olyan minimális súlyú feszítő fáját, amely egy megadott pontban (vagy általánosabban egy stabil ponthalmaz pontjaiban) fokszám korlátoknak tesz eleget. Vagy, a gráf élhalmazán megadott k darab költségfüggvényhez meg tudunk határozni k élidegen feszítő fát úgy, hogy minimalizáljuk a kiválasztás teljes költségét, ahol az i-dik fát az i-dik költségfüggvény szerinti költsége szerint számoljuk be.

A dolgozat célja, hogy összefoglalót adjon az ilyen irányú eredményekről. Ennek során megadok néhány olyan bizonyítást, melyek ismertek ugyan, de magyar nyelvű szakirodalomban még nem szerepeltek. A súlyozott matroid metszet tételre egy másutt még nem publikált új bizonyítást adok, amely csak a matroidok elemi tulajdonságait használja, szemben a komolyabb apparátust igénylő korábbi poliéderes vagy algoritmikus hátterű bizonyításokkal. Azt is igazolni fogom, hogy már Egerváry eredeti bizonyítása is hatékony, azaz erősen polinomiális futásidejű algoritmust szolgáltat maximális súlyú teljes párosítás megkeresésére. Ez nem ugyanaz, mint a H. W. Kuhn által javasolt eljárás (amit valójában Kuhn nyomán Magyar Módszernek neveznek), mert Kuhn a duál változók cseréjét és az alternáló utas módszert egybeötvözi, míg az Egerváry bizonyításból közvetlenül adódó algoritmusban a maximális elemszámú párosítást meghatározó Kőnig-féle algoritmus egy különálló szubrutinként szerepel.

2 Páros gráfok

2.1 Maximális elemszámú párosítások

Az egész elmélet kiindulópontja Kőnig [15] klasszikus tétele.

2.1 Tétel (Kőnig Dénes). *Egy $G = (S, T; E)$ páros gráfban a diszjunkt élek maximális $\nu = \nu(G)$ száma egyenlő az éleket lefogó pontok minimális $\tau = \tau(G)$ elemszámával.*

Bizonyítás. Egy ν elemű párosítás lefogásához kell legalább ν csúcs, így az összes élhez is kell, ezért $\nu \leq \tau$. Ebből az is következik, hogy $(*)$ *egy ν elemű L lefogás egy ν elemű párosítás minden élét pontosan egyszer fogja le.*

A fordított irányhoz lássuk be, hogy G élei lefoghatók $\nu(G)$ ponttal. Indirekt, legyen G minimális ellenpélda abban az értelemben, hogy $\nu(G) < \tau(G)$, de minden G-nél kisebb G' gráfra $\nu(G') = \tau(G')$.

Minden u csúcsot elkerül maximális párosítás, mert ha valamelyiket nem, úgy $\nu(G-u) < \nu(G)$, és miután $G-u$ élei már lefoghatók $\nu(G-u) \leq \nu(G)-1$ ponttal, ezen lefogáshoz u-t hozzávéve G éleinek egy legfeljebb $\nu(G)$ elemű lefogását kapnánk.

Legyen $e = st$ a gráf egy éle, melyre $s \in S, t \in T$. A $G - e$ éleinek létezik $\nu(G)$ elemű $L := A \cup B$ lefogása, ahol $A \subseteq S, B \subseteq T$. Az L nem fogja le e-t, mert különben G éleinek is $\nu(G)$ elemű lefogása volna. $G-s$-nek van $\nu(G)$ elemű párosítása, amelynek B-t fedő M_B része $(*)$ miatt nem fedi A egyetlen pontját sem. $G-t$-nek van $\nu(G)$ elemű párosítása, amelynek A-t fedő M_A része $(*)$ miatt nem fedi B egyetlen pontját sem. De most $M_A \cup M_B \cup \{e\}$ párosítása G-nek, melynek elemszáma $|A|+|B|+1 = \nu(G)+1$, ellentmondás. □

A most következő algoritmikus bizonyítás lényegében Kőnig eredeti bizonyítása kicsit algoritmikusabb nyelven elmondva. (Kőnig nem tekintette explicit azt a kérdést, hogy miként lehet megtalálni a szóban forgó alternáló utakat, de a bizonyításából ez közvetlenül kiolvasható.)

Algoritmikus bizonyítás. A nemtriviális $\nu \geq \tau$ irány igazolásához konstruálunk egy M párosítást és egy L lefogást, melyek elemszáma ugyanaz. Az eljárás tetszőleges M párosításból indul ki, ami kezdetben az üres halmaz is lehet. Az általános lépésben vagy találunk egy nagyobb párosítást, és ekkor a nagyobb párosításra vonatkozóan iteráljuk az eljárást, vagy pedig egy $|M|$-mel megegyező elemszámú lefogást, amikor is az algoritmus véget ér.

Irányítsuk meg M éleit T-től S felé, míg az összes többi élt fordítva. Jelölje R_S illetve R_T az S-ben illetve a T-ben az M által fedetlen pontok halmazát. Jelölje Z az R_S pontjaiból az így kapott irányított gráfban irányított úton elérhető pontok halmazát (amit szélességi kereséssel találhatunk meg).

Két eset lehetséges. Amennyiben R_T-nek esik pontja Z-be, akkor megkaptunk egy olyan R_S-t és R_T-t összekötő P utat, amely M-ben alternál. Most M és P szimmetrikus differenciája egy M-nél eggyel több élből álló M' párosítás. (Technikailag az eljárást nagyon egyszerű végrehajtani: a megtalált út éleinek irányítását egyszerűen megfordítjuk.)

A másik esetben R_T diszjunkt Z-től. Z definíciója folytán Z-ből nem lép ki irányított él. Érvényes továbbá, hogy Z-be nem lép be megirányított $uv \in M$ párosítás él, hiszen v csak u-n keresztül érhető el, így v csak akkor lehetett irányított úton elérhető R_S-ből, ha u is az volt.

Következik, hogy az $L := (T \cap Z) \cup (S - Z)$ halmaz egyrészt lefogja az összes élt, másrészt minden M-beli élnek pontosan az egyik végpontját tartalmazza, tehát $|M| = |L|$. □

A fenti bizonyítás egyúttal hatékony algoritmust is jelent a szóban forgó optimumok meghatározására. A lépésszám megbecsléséhez figyeljük meg, hogy legfeljebb $n/2$ alkalommal kell utat keresnünk. Miután egyetlen út megkeresése az élszámmal arányos időben történhet, az összlépésszám nem nagyobb, mint $O(nm)$ (ahol n a gráf pontszáma, míg m az élszáma).

A Kőnig tételhez szorosan kapcsolódik Hall tétele.

2.2 Tétel. *Egy $G = (S, T; E)$ páros gráfban akkor és csak akkor létezik S-t fedő párosítás, ha teljesül a Hall-féle feltétel:*

$$|\Gamma(X)| \geq |X| \quad \textit{minden } X \subseteq S \textit{ részhalmazra}\,, \tag{3}$$

ahol $\Gamma(X)$ jelöli azon T-beli pontok halmazát, melyeknek van szomszédja X-ben.

A tételt rögtön kicsit általánosabb alakban igazoljuk. Definiáljuk egy $X \subseteq S$ halmaz **hiányát** a $h(X) := (|X| - |\Gamma(X)|)$ értékkel és legyen $\mu = \mu(G, S)$ a maximális hiány, azaz

$$\mu := \max_{X \subseteq S} h(X)\,. \tag{4}$$

2.3 Tétel. *Egy $G = (S, T; E)$ páros gráfban egy párosítás által nem fedett S-beli pontok minimális száma egyenlő μ-vel.*

Bizonyítás. A max $\leq$ min egyenlőtlenség nyilván fennáll. A fordított irány igazolásához legyen M egy maximális (azaz ν elemű) párosítás és L egy minimális (azaz $\tau = \nu$ elemű) lefogás. Legyen $X := S - L$. Ekkor M pontosan $|S| - \nu$ elemét nem fedi S-nek. Másrészt $\Gamma(X) \subseteq L - S$ és így

$$|X| - |\Gamma(X)| \geq |S - L| - |L - S| = |S| - |L| = |S| - \tau = |S| - \nu\,.$$

□

Legyen $\mathcal{F}$ az S maximális (azaz μ) hiányú részhalmazainak rendszere, vagyis $\mathcal{F} := \{X \subseteq S : |X| - |\Gamma(X)| = \mu\}$. Az $\mathcal{F}$ tagjait röviden *max-hiányú* halmazoknak fogjuk hívni. Az előbbi bizonyítás mutatja, hogy egy-egy értelmű kapcsolat áll fenn a minimális lefogások és a max-hiányú S-beli halmazok között: Ha L minimális lefogás, akkor $S - L$ max-hiányú halmaz, míg ha $H \subseteq S$ max-hiányú halmaz, akkor $\Gamma(H) \cup (S - H)$ minimális lefogás lesz.

2.4 Lemma. *$\mathcal{F}$ zárt a metszet és unió képzésre.*

Bizonyítás. Könnyű ellenőrizni, hogy a h függvény szupermoduláris, azaz $h(X) + h(Y) \leq h(X \cap Y) + h(X \cup Y)$. Tegyük most fel, hogy X és Y két maximális hiányú halmaz (azaz $\mathcal{F}$ elemei). Ekkor $\mu + \mu = h(X) + h(Y) \leq h(X \cap Y) + h(X \cup Y) \leq \mu + \mu$, és emiatt valóban $h(X \cap Y) = \mu$, $h(X \cup Y) = \mu$. □

A lemmából következik, hogy az összes max-hiányú halmaz metszete is és uniója is max-hiányú, azaz létezik egy egyértelmű legszűkebb és egy legbővebb max-hiányú halmaz. Megjegyzendő, hogy Kőnig fenti alternáló utas algoritmusa segítségével e két halmaz könnyen megkonstruálható. Például, a legszűkebb K max-hiányú halmaz, amint azt könnyű kimutatni, éppen $Z \cap S$ lesz, ahol Z jelölte az irányított segédgráfban az R_S-ből elérhető pontok halmazát. Ez egyúttal azt is mutatja, hogy az algoritmus által produkált $(S - K \cup \Gamma(K))$ lefogás nem függ az algoritmus futásától (szemben az algoritmus által szolgáltatott párosítással). A súlyozott párosítási algoritmus bizonyításához szükségünk lesz még az alábbi hasznos megfigyelésre.

2.5 Lemma. *Legyen $K \subseteq S$ a legszűkebb max-hiányú halmaz G-ben. Ha a gráfból kitöröljük az összes olyan élt, amely $\Gamma(K)$ és $S - K$ között vezet, akkor a létrejövő G' gráfban a maximális hiány ugyanaz, mint G-ben. Továbbá G és G' max-hiányú halmazainak rendszere ugyanaz.*

Bizonyítás. Miután G egy M maximális párosításának a $\Gamma(K)$-t fedő élei mind K-ban végződnek, az M benne van G'-ben is, vagyis G' max hiánya legfeljebb akkora, mint G-é, de persze kisebb nem lehet, mert G' részgráfja G-nek. Ebből az is következik, hogy a G egy max-hiányú halmaza G'-ben is max-hiányú. Legyen most X tetszőleges max-hiányú halmaz G'-ben. Mivel K max-hiányú G'-ben is, a 2.4 lemma szerint $K \cap X$ is max-hiányú G'-ben. De akkor $K \cap X$ max-hiányú G-ben, hiszen $K \cap X$-ből induló élt nem töröltünk, és így a K minimalitása folytán $K \subseteq X$. Ekkor viszont $\Gamma(X) = \Gamma'(X)$, azaz X max-hiányú G-ben is. □

2.2 Súlyozott párosítások

Egerványnak az 1.2 tételére adott eredeti bizonyításának a váza a következő. Legyen π egy minimális súlyú nemnegatív, egészértékű súlyozott lefogás. (A π súlyán a $\sum[\pi(v) : v \in S \cup T]$ összeg értendő.) Feltehetjük, hogy π az S elemein mindenütt pozitív, mert ha nem, akkor az S-beli pontokon eggyel növelve, a T-beli pontokon eggyel csökkentve már ilyen lesz. Amennyiben azon élek G_π részgráfjában, melyekre $\pi(u) + \pi(v) = c(uv)$ létezik teljes M párosítás, úgy M maximális súlyú párosítás, melynek súlya egyenlő a $\pi(v)$ értékek összegével. Ha viszont G_π-ben nincs teljes párosítás, úgy Kőnig vagy Hall tétele alapján létezik hiányos X halmaz, azaz olyan, amelynek $|X|$-nél kevesebb szomszédja van. A π értékeit az X pontjain eggyel csökkentve, a $\Gamma_{G_\pi}(X)$ pontjain pedig eggyel növelve, egy másik nemnegatív, súlyozott lefogást kapunk, amelynek súlya kisebb, mint π-é, ellentmondásban π minimális választásával.

Ez a bizonyítás könnyen algoritmizálható, hiszen tetszőleges π súlyozott lefogásból kiindulva vagy talál egy maximális súlyú teljes párosítást, és ekkor az aktuális π is optimális, vagy pedig talál egy jobb súlyozott lefogást, amivel az eljárást iterálva véges sok lépés után az első eset következik be. Egy kézenfekvő gyorsítási lehetőség azonnal kínálkozik (amint azt Egerváry a későbbi [9] dolgozatában maga is javasolja): a π súlyozott lefogás módosításakor ne eggyel növeljünk vagy csökkentsünk, hanem a legnagyobb olyan δ értékkel, amelyre a módosított π' még súlyozott lefogás. Az így nyert eljárást nevezzük Egerváry algoritmusának. (Hangsúlyozzuk, hogy ez nem ugyanaz, mint a H. W. Kuhn által kifejlesztett primál-duál eljárás, amit ma valójában a világban Kuhn javaslata nyomán Magyar Módszernek hívnak.)

Az Egerváry algoritmusnak az az előnye is megvan, hogy nem egészértékű c súlyfüggvényre is működik, bár ekkor még az is kérdés, hogy az eljárás véges-e egyáltalán, és Egerváry valójában a nemegész c esetét nem algoritmikusan, hanem folytonossági megfontolásokkal intézte el.

Nem kevésbé fontos a másik kérdés, hogy Egerváry algoritmusa milyen hatékony, akár egész a c, akár nem. Példával megmutatható, hogy ha tetszőleges olyan X halmazt használunk a π módosítására, amely megsérti a Hall-féle feltételt, akkor az algoritmus nem polinomiális futásidejű (és ez a kellemetlenség még akkor is előfordulhat, ha X maximális hiányú halmaz.) Ráadásul irracionális költségek esetén még azt sem tudjuk, hogy az algoritmus véges sok lépés után megáll-e.

Egerváry azonban [8]-ban azt javasolja, hogy a π változtatását annak a maximális hiányú X halmaznak a segítségével végezzük, amelyet Kőnig alternáló utas algoritmusa szolgáltat, amely tehát a(z egyértelmű) legszűkebb max-hiányú halmaz. Az alábbiakban kimutatjuk, hogy Egerváry algoritmusa ilyenkor polinomiális, sőt erősen polinomiális futásidejű. Mivel nem jelent semmilyen többlet nehézséget, a bizonyítást rögtön a tétel azon csöppnyit általánosabb alakjára mondjuk el, amikor a páros gráf nem feltétlenül teljes, csupán azt írjuk elő, hogy tartalmazzon teljes párosítást.

2.6 Tétel. *Tegyük fel, hogy a $G = (S, T; E)$ páros gráfnak van teljes párosítása. Legyen továbbá $c : E \to \mathbf{R}_+$ az éleknek egy nemnegatív számokkal történő súlyozása. Ekkor G teljes párosításainak maximális súlya egyenlő a súlyozott lefogások minimális súlyával, ahol súlyozott lefogáson egy olyan $\pi : V \to \mathbf{R}$ függvényt értünk, amelyre $\pi(u) + \pi(v) \geq c(uv)$ a gráf minden uv élére fennáll. Amennyiben c egészértékű, úgy az optimális π is választható egészértékűnek. Amennyiben G teljes páros gráf, az optimális π választható nemnegatívnak.*

Bizonyítás. Tetszőleges M teljes párosításra és π súlyozott lefogásra fennáll, hogy $\sum(\pi(z) : z \in S \cup T) = \sum([\pi(u)+\pi(v)] : uv \in M) \geq \sum(c(uv) : uv \in M)$, vagyis a minimum valóban legalább a maximum. Az is kiolvasható, hogy itt egyenlőség pontosan akkor áll, ha az M párosítás minden uv éle *pontos* abban az értelemben, hogy $\pi(u) + \pi(v) = c(uv)$. Célunk tehát azt kimutatni, hogy létezik egy olyan π, amelyre nézve a pontos élek gráfjában létezik teljes párosítás.

Legyen π egy olyan súlyozott lefogás (egészértékű, ha c az), amelyre a pontos élek $G_\pi = (S, T; E_\pi)$ részgráfjában a $\mu_\pi (= |S| - \nu(G_\pi))$ érték minimális, és ezen belül, az (egyértelmű) legszűkebb max-hiányú $K \subseteq S$ halmaz a lehető legnagyobb. Készen vagyunk, ha a μ_π maximális hiány nulla, mert ez épp azt jelenti, hogy G_π-nek van teljes párosítása. Tegyük fel tehát, hogy $\mu_\pi > 0$. Mivel G-nek van teljes párosítása, biztosan van olyan e éle G-nek, amely K és $T - \Gamma_\pi(K)$ között vezet, ahol $\Gamma_\pi(K)$ jelöli a K szomszédainak halmazát a G_π-ben. Ilyen él nem pontos, így a

$$\delta := \min(\pi(u) + \pi(v) - c(uv) : uv \in E, u \in K, v \in T - \Gamma_\pi(K)) \qquad (5)$$

érték pozitív. Módosítsuk π-t úgy, hogy a K minden elemén δ-val csökkentjük, míg $\Gamma_\pi(K)$ minden elemén δ-val növeljük, azaz

$$\pi'(v) := \begin{cases} \pi(v) - \delta, & \text{ha } v \in K; \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_\pi(K); \\ \pi(v), & \text{egyébként.} \end{cases} \tag{6}$$

A δ választása miatt az így módosított π' továbbra is súlyozott lefogás, amely egészértékű, ha c az volt. A π'-ra vonatkozó pontos élek $G_{\pi'}$ gráfja abban különbözik G_π-től, hogy van legalább egy éle (ahol a δ-t definiáló minimum felvétetett) K és $T - \Gamma_\pi(K)$ között, de biztosan nincsen éle $T \cap \Gamma_\pi(K)$ és $S - K$ között (miközben G_π-nek lehetett). A 2.5 lemma szerint $\mu_{\pi'} = \mu_\pi$, és $G_{\pi'}$-ben a legszűkebb max-hiányú halmaz szigorúan bővebb, mint K, és ez ellentmond π választásának.

Beláttuk tehát, hogy van olyan π súlyozott lefogás, amelyre a pontos élek részgráfjában van teljes párosítás. Végül tegyük fel, hogy G teljes páros gráf, és igazoljuk, hogy π választható nemnegatívnak. Legyen a π legnegatívabb értéke $-\alpha$ ahol $\alpha > 0$ és legyen $\pi(v) = -\alpha$, ahol v mondjuk S-ben van. ekkor minden $u \in T$ csúcsra $\pi(u) + \pi(v) \geq c(uv) \geq 0$ miatt $\pi(u) \geq \alpha$. Így ha π-t úgy módosítjuk, hogy S elemein növeljük α-val, míg T elemein csökkentjük α-val, akkor egy másik minimális súlyozott lefogás keletkezik, amelyik már nemnegatív. □

A 3. szakaszban még az is kiderül, hogy az optimális súlyozott lefogás pontosan akkor választható nemnegatívnak, ha a maximális súlyú párosítás teljes párosításon vétetik fel (amely feltétel persze teljes páros gráf esetén fennáll.)

A bizonyítás alapján Egerváry algoritmusa a következő. Az algoritmus bemenete egy teljes párosítással rendelkező páros gráf, míg a kimenete egy maximális súlyú párosítás és egy minimális súlyú súlyozott lefogás. Szubrutinként szükségünk van a súlyozatlan esetre vonatkozó fentebb ismertetett alternáló utas algoritmusra, amely egy tetszőleges $G' = (S, T; E')$ páros gráfban megkonstruál egy maximális M' párosítást és a(z egyértelmű) legszűkebb $K' \subseteq S$ max-hiányú halmazt (melyekre tehát $|M'| = |\Gamma'(K')| + |S - K'|$). Hivatkozás kedvéért ezt Kőnig szubrutinnak nevezzük.

Az algoritmus egy általános lépésében rendelkezésre áll egy π súlyozott lefogás (amely egészértékű, ha c az, és amely kezdetben lehet például az azonosan α lefogás, ahol α a $c(e)$ költségek maximuma. Alkalmazzuk a Kőnig szubrutint a π-re nézve pontos élek G_π részgráfjára. Amennyiben G_π-ben van M teljes párosítás, úgy az algoritmus az aktuális π súlyozott lefogás és M teljes párosítás kiadásával véget ér. Ha viszont G_π-nek nincs teljes párosítása, úgy a szubrutin által szolgáltatott (G_π-re nézve) legszűkebb K_π max-hiányú halmaz segítségével (6) szerint módosítjuk π-t, és az eljárást iteráljuk.

Mi mondható az algoritmus lépésszámáról? Tekintsük egy fázisnak az algoritmus azon szakaszát, amíg a pontos élek (egyre változó) részgráfjában a maximális hiány változatlan. Nyilván legfeljebb $|S|$ fázis létezik. Egy fázis során a szubrutint legfeljebb $|S|$-szer hívjuk meg, hiszen beláttuk, hogy a π

cseréjekor a legszűkebb max-hiányú halmaz szigorúan bővül. Vagyis a súlyozatlan esetre vonatkozó alternáló utas algoritmusnak legfeljebb $|S|^2$-szeri meghívásával az algoritmus futása befejeződik. Miután Kőnig algoritmusának lépésszámára $O(|S||E|)$ korlát volt mondható, a leírt súlyozott eljárás teljes futásideje $O(|S|^3|E|)$.

Bár ez a lépésszám nem különösebben látványos (és valójában hatékonyabb eljárások is léteznek), mindenesetre azt megkaptuk, hogy az algoritmus polinomiális futásidejű, sőt erősen polinomiális is abban az értelemben, hogy a futásidő egyáltalán nem függ a szereplő c költségfüggvénytől, amennyiben feltesszük, hogy a számokkal végzett összeadást, kivonást és összehasonlítást egyetlen lépésben tudjuk elvégezni.

A jelen megközelítés előnye, hogy tisztán mutatja a súlyozatlan és a súlyozott párosítási algoritmusok viszonyát. A súlyozatlan Kőnig-féle algoritmus teljesen szeparáltan, szubrutinként kerül felhasználásra. Amint azt H. W. Kuhn megmutatta, a két eljárás összevonható, aminek talán hátránya, hogy az algoritmus összetettebbé válik, de előnye, hogy jobb lépésszám becslés adódik. Most ismertetjük a Kuhn által javasolt Magyar Módszert.

Kuhn algoritmusa súlyozott teljes párosítás meghatározására

Az általános lépésben tekintjük a pontos élek által (az $S \cup T$ ponthalmazon) alkotott G_π részgráfot. Legyen M egy már rendelkezésre álló párosítás G_π-ben. Irányítsuk meg M éleit T-től S felé, míg az összes többi G_π-beli élt fordítva. Jelölje R_S illetve R_T az S-ben illetve a T-ben az M által fedetlen pontok halmazát. Jelölje Z az R_S pontjaiból az így kapott irányított gráfban irányított úton elérhető pontok halmazát (amit például szélességi kereséssel találhatunk meg). Amennyiben R_T-nek esik pontja Z-be, akkor megkaptunk egy olyan R_S-t és R_T-t összekötő P utat, amely M-ben alternál. Most M és P szimmetrikus differenciája egy M-nél eggyel több élből álló M' párosítás. Ekkor az eljárás egy fázisa véget ér, és az új M' párosítással folytatva iteráljuk az eljárást.

Nézzük most azt az esetet, amikor R_T diszjunkt Z-től. Legyen $K_\pi := Z \cap S$ és módosítsuk π-t a (6)-ban leírtak szerint. Ekkor az R_S-ből elérhető pontok halmaza szigorúan bővül. Így egy fázis (ami alatt tehát a pontos élek gráfjában a maximális párosítás elemszáma nem nő) $|S|$ útkereső eljárás alkalmazása után véget ér. Mivel egy útkeresés $O(|E|)$ lépésben végrehajtható és legfeljebb $|S|$ fázis van, az algoritmus teljes futásideje $O(|E||S|^2)$.

A fejezet lezárásaképp megmutatjuk, hogy a maximális súlyú (nem feltétlenül teljes) párosítás meghatározásának problémája egyszerű fogással visszavezethető a maximális súlyú teljes párosítására.

2.7 Tétel. *Egy $G' = (S', T'; E')$ páros gráfban nemnegatív c súlyfüggvény esetén a párosítások maximális ν'_c súlya egyenlő a nemnegatív (!) súlyozott lefogások minimális τ'_c súlyával. Amennyiben c egészértékű, az optimális π' is választható egészértékűnek.*

Bizonyítás. A $\nu'_c \leq \tau'_c$ egyenlőtlenség nyilvánvaló, így csak a fordított iránnyal foglalkozunk. Új pontok esetleges hozzávételével elérhetjük, hogy a páros gráf két osztálya egyforma méretű legyen. Egészítsük ki a gráfot 0 súlyú élek bevételével egy G teljes páros gráffá. A súlyfüggvény ezen kiterjesztését továbbra is jelölhetjük c-vel. A 2.6 tétel (második része) szerint G-nek létezik egy M teljes párosítása és c-nek egy π nemnegatív súlyozott lefogása, melyekre $c(M) = \sum \pi(v)$. Mivel az új élek súlya 0, így az új élek kihagyásával M-ből keletkező G'-beli M' párosítás súlya változatlanul $c(M)$. Miután új pontból (ha van) csak 0 súlyú él megy ki, így π értéke az új pontokon 0, vagyis, ha π-t megszorítjuk az eredeti pontokra, akkor a keletkező π'-re $\sum \pi'(v) = \sum \pi(v)$, és így $\sum \pi'(v) = c(M')$. □

Végül megjegyezzük, hogy tetszőleges rögzített k pozitív egészre Ford és Fulkerson minimális költségű folyam algoritmusának segítségével ki lehet számolni a maximális (vagy minimális) súlyú k élű párosítást, ha ilyen párosítás egyáltalán létezik.

3 Teljesen unimoduláris mátrixok az optimalizálásban

Az alábbiakban ismertetjük azt a Hoffmantól és Kruskaltól [13] származó megközelítést, amely rávilágít arra a háttérben megbújó mélyebb okra, ami miatt Egerváry tétele fennáll. Ez pedig az, hogy egy páros gráf pont-él incidencia mátrixa teljesen unimoduláris, és a lineáris programozás dualitás tételében az optimumok egészértékű vektoron is felvétetnek, amennyiben a feltételi mátrix teljesen unimoduláris.

Az alábbiakban egy mátrixot vagy egy vektort akkor nevezünk egésznek vagy egészértékűnek, ha minden eleme (komponense) egész szám. Valamely A mátrixot akkor nevezünk *teljesen unimodulárisnak*, ha minden aldeterminánsa $(0, \pm 1)$ értékű. Speciálisan, ilyen mátrix minden eleme $0, +1$ vagy -1. Világos, hogy TU-mátrix transzponáltja is az. Sorokat vagy oszlopokat -1-gyel szorozva vagy elhagyva ismét TU-mátrixot kapunk. Továbbá, egységvektorokat sorként vagy oszlopként egy TU-mátrixhoz illesztve TU-mátrixot kapunk. Így, ha az A TU-mátrixot kiegészítjük egy I egységmátrixszal, akkor a keletkező (A, I) mátrix is TU-mátrix. Ha A TU-mátrix, úgy $(A, -A)$ is az.

Példaképp, legyen A egy $D = (V, E)$ irányított gráf incidencia mátrixa, azaz A sorai a V-nek, oszlopai E-nek felelnek meg, és az $a_{v,e}$ elem akkor $+1$ illetve -1, ha az e él belép illetve kilép v-ből (egyébként 0). Nem nehéz igazolni, hogy digráf incidencia mátrixa teljesen unimoduláris, és hogy páros gráf incidencia mátrixa teljesen unimoduláris.

Ezeket általánosítja a *hálózati mátrix*. Legyen D olyan irányított gráf, amely irányítatlan értelemben összefüggő, és legyen F egy feszítő fa. Az A mátrix sorai az F éleinek felelnek meg, míg az oszlopai az F-en kívüli éleknek. Minden uv nem-fa élre a fában egy egyértelmű (nem biztosan irányított) út vezet v-ből u-ba. Ennek egy f elemére a mátrix $a_{f,e}$ elemét definiáljuk 1-nek, ha f iránya megegyezik az útéval és -1-nek, ha azzal ellentétes. A mátrix minden más eleme 0. Számos érdekes alkalmazást tesz lehetővé az alábbi Tutte-tól való eredmény.

3.1 Tétel. *Az A hálózati mátrix teljesen unimoduláris.*

Bizonyítás. Mivel hálózati mátrix részmátrixa is az, elég azt belátni, hogy egy négyzetes hálózati mátrix determinánsa $0, 1$ vagy -1. Hálózati mátrix sorát vagy oszlopát -1-gyel szorozva hálózati mátrixot kapunk. Egy sor vagy oszlop -1-gyel való szorzása annak felel meg, hogy a megfelelő élt (akár fa-él, akár nem-fa él) átirányítjuk.

Tekintsük a fának egy v végpontját. Ha az F fa v-vel szomszédos éléhez tartozó sorban lévő nemnulla elemek α száma legfeljebb 1, akkor a determináns kifejtési szabály alapján indukcióval készen vagyunk. Tegyük fel, hogy $\alpha > 1$, vagyis v szomszédos legalább két nem-fa éllel. Átirányítás miatt feltehető, hogy ezek közül pontosan egy van v felé irányítva. Legyen ez sv és legyen vt egy másik nem-fa él. Ha az sv-nek megfelelő oszlopot, hozzáadjuk a vt-nek megfelelő oszlophoz, akkor egyrészt persze a determináns értéke nem változik, másrészt ismét hálózati mátrixot kapunk, éspedig azé a gráfét, amelyben a vt él helyett az st él szerepel.

Ilyen átalakításokkal egy olyan digráfot kaphatunk, amelyben az F feszítő fa változatlan, egyetlen nem-fa él (nevezetesen sv) szomszédos v-vel, vagyis a hozzátartozó hálózati mátrix v-nek megfelelő sorában egy nemnulla elem van. Ilyen hálózati mátrixról pedig már láttuk, hogy a determinánsa $0, \pm 1$, ugyanakkor a fenti operációk nem változatták a determináns abszolút értékét. $\square$

Most megvizsgáljuk, hogy a lineáris programozás dualitás tétele illetve a Farkas lemma miként alkalmazható olyan kombinatorikus optimalizálási feladatok esetén, mint amilyen a súlyozott párosítás problémája. A megoldás kulcsa az a megfigyelés lesz, hogy bizonyos speciális feltételi mátrixok esetén az optimum mindig egész vektoron is felvétetik.

Adott $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ mátrixhoz tekintsük a

$$Px = b_0, \quad Qx \leq b_1 \tag{7}$$

lineáris rendszert. Tegyük fel, hogy ennek megoldás halmaza az R poliéder nem üres. Az R egy elemét nevezzük *erős bázis-megoldás*nak, ha előáll valamely $M'x' = b'$ egyenletrendszer egyértelmű megoldásának nulla komponensekkel való kiegészítéseként, ahol M' az M egy $[(r(M) \times (r(M)]$-es nemszinguláris részmátrixa és b' jelöli a b azon részét, amely az M' sorainak felel meg. (Könnyű igazolni, hogy egy $\{Ax = b, x \geq 0\}$ alakú rendszer egy megoldása pontosan akkor erős bázis-megoldás, ha a pozitív komponenseihez tartozó A-oszlopok lineárisan függetlenek). Abban a speciális esetben, amikor R csúcsos, némi munkával igazolható, hogy az erős bázis-megoldások éppen az R csúcsai. A definícióból következik, hogy legfeljebb csak véges sok erős bázis-megoldás létezhet. A Caratheodory tétel egy változata szerint mindig létezik erős bázis-megoldás, sőt tetszőleges olyan c célfüggvényre, amelyre cx felülről korlátos az R poliéderen, érvényes, hogy az R bármely x' eleméhez létezik egy olyan x^* erős bázis-megoldás, amelyre $cx^* \geq cx'$, amiből adódik, hogy a $\max\{cx : x \in R\}$, ha korlátos egyáltalán, akkor egy erős bázis-megoldáson felvétetik.

3.2 Lemma. *Tetszőleges M TU-mátrixszal megadott (7) egyenlőtlenség-rendszer esetén, ha a jobboldali korlátozó b vektor egész, akkor minden erős bázis-megoldás egész.*

Bizonyítás. Egy erős bázis-megoldás előáll valamely $M'x' = b'$ egyenletrendszer egyértelmű megoldásának nulla komponensekkel való kiegészítéseként, ahol M' az M egy $[(r(M) \times (r(M)]$-es nemszinguláris részmátrixa és b' jelöli a b azon részét, amely az M' sorainak felel meg. Mármost, ha M TU-mátrix, akkor a nemszinguláris M' determinánsa $+1$ vagy -1. A Cramer szabály szerint, miután b' egész, az egyértelmű x' megoldás is az. □

3.3 Lemma. *Legyen c tetszőleges (nem feltétlenül egészértékű) vektor. Bármely*

$$M = \begin{pmatrix} P \\ Q \end{pmatrix}$$

TU-mátrixszal megadott $K := \{x : Px = 0,\ Qx \leq 0\}$ metszet-kúpnak, ha van olyan x' eleme, amelyre $cx' > 0$, akkor K-nak van ilyen $(0, \pm 1)$-értékű eleme is.

Bizonyítás. Mivel x' pozitív számszorosa is K-ban van, feltehető, hogy x' maga olyan, hogy minden komponense a $[-1, +1]$ zárt intervallumba esik. Vagyis a

$$(-1, \ldots, -1) \leq x \leq (1, \ldots, 1), \quad Px = 0, \quad Qx \leq 0 \tag{9}$$

rendszer által meghatározott korlátos poliédernek x' olyan eleme, amelyre $cx' > 0$. Ekkor az előbb említett tulajdonság szerint van olyan x^* erős bázis-megoldása (8)-nak, amelyre $cx^* \geq cx'$. A 3.2 lemma miatt x^* egészértékű, azaz minden komponense $0, \pm 1$. □

A Farkas lemma (egyik változata) azt mondja ki, hogy a (7) és az alábbi (9) lineáris rendszerek közül pontosan az egyik oldható meg. Amennyiben a szereplő M mátrix teljesen unimoduláris, a megoldásokról több mondható:

3.4 Tétel. *Tegyük fel, hogy az $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ mátrix teljesen unimoduláris. Ha a (7) primál probléma oldható meg és a korlátozó b vektor egész, akkor (7)-nek van egész megoldása is. Ha az*

$$y_1 \geq 0, \quad yM = 0, \quad yb < 0 \tag{9}$$

duális probléma oldható meg, ahol $y = (y_0, y_1)$, akkor van $(0, \pm 1)$-értékű y megoldás is (függetlenül b egészértékűségétől).

Bizonyítás. A tétel első fele következik a 3.2 lemmából, és abból a már említett eredményből, hogy ha létezik megoldás, akkor létezik erős bázis-megoldás is. A tétel második fele pedig a 3.3 lemma közvetlen folyománya. □

3.5 Tétel. *Ha a* $\max(cx : Px = b_0, Qx \leq b_1)$ *lineáris programozási problémának létezik megoldása, továbbá ha az* $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ *mátrix teljesen unimoduláris és b egész, akkor az optimum egész vektoron is felvétetik (függetlenül attól, hogy c egészértékű vagy sem).*

Bizonyítás. Miután az optimum erős bázis-megoldáson is felvétetik, a 3.2 lemmából az eredmény következik. □

Alkalmazásként először levezetjük Kőnig tételét.

A 2.1 tétel bizonyítása. Nyilván minden gráfra $\nu \leq \tau$. Az egyenlőség igazolásához azt kell kimutatnunk, hogy páros gráfban létezik olyan párosítás és olyan lefogó pontrendszer, melyek elemszáma megegyezik.

A páros gráf incidencia mátrixát jelölje A, amelyben a soroknak a gráf pontjai, az oszlopoknak a gráf élei felelnek meg. Tekintsük a következő primál-duál lineáris program párt:

$$\max(1x : Ax \leq 1,\, x \geq 0)\,, \tag{10}$$

$$\min(1y : yA \geq 1,\, y \geq 0)\,. \tag{11}$$

A 3.5 tétel szerint mindkét programnak az optimuma egész vektoron felvétetik. Jelöljük ezeket rendre x_0-lal és y_0-lal. (10) minden egészértékű megoldása $0-1$ értékű, és rögtön látszik, hogy (11) minden optimális egészértékű megoldása is $0-1$ értékű. Legyen M azon élek halmaza, melyeken x_0 az 1 értéket veszi fel, és legyen L azon pontok halmaza, amelyeken y_0 az 1 értéket veszi fel. Az $Ax \leq 1$ feltétel azt jelenti, hogy M párosítás a gráfban, míg az $yA \geq 1$ feltétel azt jelenti, hogy L az éleket lefogó pontrendszer. A primál és duál optimum értékek egyenlősége pedig azt jelenti, hogy $|M| = |L|$, ami a célunk volt. □

Természetesen a primál programban az azonosan 1 célfüggvény helyett választhatunk tetszőleges c célfüggvényt. Ekkor a fenti módszer kiadja a 2.6 tételt, sőt annak utolsó mondatát még az alábbi erősebb alakban: *Az optimális π akkor és csak akkor választható nemnegatívnak, ha a maximális súlyú párosítás teljes párosításon is felvétetik.*

További általánosításokat kaphatunk, ha a primál feladatban a jobboldalt valamilyen (nemnegatív) b vektornak választjuk. Ennek az a kombinatorikus jelentése, hogy maximális költségű fokszám-korlátozott részgráfot keresünk. Természetesen alsó korlátokat is kitűzhetünk a fokszámokra, mint ahogy korlátozhatjuk alulról és felülről azt, hogy egy élt hány példányban vehetünk be a keresett részgráfba. Valójában nem is érdemes megfogalmazni a különböző lehetőségekre vonatkozó min-max tételeket, mert a dualitás tétel és a páros gráf incidencia mátrixának teljes unimodularitása már magában hordozza a szükséges információt.

A szakasz befejezéseként megmutatjuk, hogy a TU-mátrixokra vonatkozó Farkas lemma (3.4 tétel) miként adja ki Hoffman megengedett áram tételét. Jelöljön $D = (V, E)$ egy irányított gráfot. Legyen $f : E \to \mathbf{R} \cup \{-\infty\}$ alsó kapacitás, $g : E \to \mathbf{R} \cup \{+\infty\}$ felső kapacitás úgy, hogy $f \leq g$. Valamely $x : E \to \mathbf{R}$ vektorra és $S \subseteq V$ részhalmazra legyen $\varrho_x(S) := \sum(x(uv) : uv \in E, uv$ belép S-be) és legyen $\delta_x(S) := \varrho_x(V - S)$. Az x vektort *áramnak* (cirkulációnak) nevezzük, ha teljesül a rá *megmaradási szabály*, azaz $\varrho_x(v) = \delta_x(v)$ fennáll minden v csúcsra. Az x áramot *megengedettnek* mondjuk, ha

$$f \leq x \leq g\,. \tag{12}$$

3.6 Tétel (A. Hoffman, 1960). *Akkor és csak akkor létezik megengedett áram, ha*

$$\varrho_f(X) \leq \delta_g(X) \text{ minden } X \subseteq V \text{ halmazra}. \tag{13}$$

Továbbá, ha f és g egészértékűek, és (13) fennáll, akkor létezik egészértékű megengedett áram is.

Bizonyítás. A szükségesség igazolásához, tegyük fel, hogy x megengedett áram. Ekkor $\delta_g(X) - \varrho_f(X) \geq \delta_x(X) - \varrho_x(X) = 0$, amiből (13) következik.

Az elegendőséghez tekintsük az $\{Ax \leq 0, x \leq g, -x \leq -f\}$ rendszert. Figyeljük meg, hogy a jelen esetben egy x vektorra akkor és csak akkor teljesül az $Ax \leq 0$ egyenlőtlenség, ha $Ax = 0$. A 3.4 tételt alkalmazva kapjuk, hogy ha a fenti rendszernek nincs megoldása, akkor van olyan (y, u, v) $(0,1)$-értékű vektor, amelyre $(*)$ $yA + u - v = 0$ és $(**)$ $ug - vf < 0$. Mivel $f \leq g$, így minden élre feltehető, hogy $u(e)$ és $v(e)$ közül legalább az egyik nulla (ha ugyanis mindkettő 1, akkor mindkettőt helyettesíthetjük nullával.)

Jelölje Z azon z pontok halmazát, ahol az $y(z) = 1$. Ekkor $(*)$ miatt minden olyan e élre, amelynek mindkét vége vagy Z-ben vagy $V - Z$-ben van, $u(e) = v(e) = 0$. Továbbá minden Z-be belépő e élre $v(e) = 1$, $u(e) = 0$ és minden z-ből kilépő élre $v(e) = 0, u(e) = 1$. Miután $ug = \delta_g(Z)$ és $vf = \varrho_f(Z)$, így $(**)$ ellentmond a (13) feltételnek. □

4 Párosítások nem-páros gráfban

Nem-páros gráfokra a maximális (súlyú) párosítás meghatározásának problémája jóval nehezebb, mint a páros esetben. Ebben a fejezetben ismertetjük azt a döntően J. Edmondstól eredő elméleti hátteret, amely lehetővé tette [4] az algoritmusok kidolgozását. Maguk az algoritmusok összetettebbek annál, semhogy egy ilyen összefoglaló cikk keretében vállalkozhatnánk a bemutatásukra, ugyanakkor az alábbi fő eredményekre mára már viszonylag tömör bizonyítások állnak rendelkezésre. A közölt bizonyítások Schrijver [17] munkájából valók, némi egyszerűsítéssel.

4.1 Maximális elemszámú párosítások

A maximális elemszámú párosítás meghatározásának kérdésére Tutte tétele illetve a Berge-Tutte formula adja meg az elvi választ. A háromszög példája mutatja, hogy a páros gráfra vonatkozó esettel szemben, itt már a párosítások maximális ν elemszáma lehet szigorúan kisebb, mint a lefogó pontok minimális τ száma.

4.1 Tétel (W. T. Tutte). *Egy $G = (V, E)$ gráfban akkor és csak akkor létezik teljes párosítás, ha a csúcsok bármely X részhalmazát kihagyva a keletkező páratlan pontszámú komponensek $q(X)$ száma legfeljebb $|X|$.*

Ezt általánosítja (bárha könnyen levezethető belőle) a párosítások maximális számára vonatkozó Berge-Tutte formula.

4.2 Tétel (Berge-Tutte formula).

$$\nu(G) = \min(|V| - q(X) + |X| : X \subseteq V)/2 \,. \tag{14}$$

Vagy ekvivalens alakban, egy párosítás által fedetlenül hagyott pontok minimális száma egyenlő $q(X) - |X|/2$ érték minimumával.

Bizonyítás. Egy összefüggő gráfot *(faktor-)kritikusnak* neveznek, ha bármely pontját kihagyva a maximális párosítás elemszáma nem csökken, más szóval minden csúcsot elkerül maximális párosítás. A bizonyítás egyszerű indukcióval fog következni az alábbi lemmából.

4.3 Lemma (Gallai). *Ha $H = (U, F)$ gráf kritikus, akkor $|U|$ páratlan és $\nu(H) = (|U| - 1)/2$, azaz H-nak van olyan párosítása, amely egyetlen pontot hagy fedetlenül.*

Bizonyítás. Kritikus gráfnak természetesen nem lehet teljes párosítása. Tegyük fel indirekt, hogy H egy maximális párosítása legalább két pontot fedetlenül hagy, és válasszuk meg ezt az M párosítást és a fedetlen s és t pontokat úgy, hogy a H-beli távolságuk minimális legyen. Ez a távolság persze nem lehet egy, azaz s és t nem lehet szomszédos, mert akkor az st élt M-hez véve nagyobb párosítást kapnánk. Az s-t és t-t összekötő P legrövidebb útnak legyen x egy belső pontja. Mivel H kritikus, létezik x-t elkerülő maximális M' párosítás. M és M' két maximális elemszámú párosítás, ezért szimmetrikus differenciájuk diszjunkt alternáló körökből és páros élszámú utakból áll. Ezek közül jelölje R az x-t tartalmazó utat. Ekkor M és R szimmetrikus differenciája egy olyan M'' maximális párosítás, amely szabadon hagyja x-t, és legalább az s és t pontok egyikét, és ez ellentmond az s, t és M választásának. □

Rátérve a Berge-Tutte formula bizonyítására, látható, hogy tetszőleges M párosítás és $X \subseteq V$ halmaz esetén legalább $q(X) - |X|$ pont marad fedetlen, azaz M legfeljebb $|V| - (q(X) - |X|)$ pontot fed, így az M elemszáma legfeljebb $(|V| - q(X) + |X|)/2$. Így (14)-ben $\nu(G) \leq \min$ következik.

A fordított irány bizonyításához V elemszáma szerinti indukciót alkalmazunk. Ha $|V| = 0$, akkor (14) mindkét oldala 0. Tegyük fel tehát, hogy $|V| \geq 1$ és azt, hogy a (14) formula érvényes minden kisebb gráfra. Nyilván feltehető, hogy G összefüggő. Azt kell kimutatnunk, hogy létezik egy olyan $X_0 \subseteq V$ halmaz, amelyre

$$\nu(G) \geq (|V| - q(X_0) + |X_0|)/2 \,. \tag{15}$$

1. eset *G nem kritikus*, azaz van olyan v pontja, amelyet elhagyva a keletkező G' gráfra $\nu(G') \leq \nu(G) - 1$. Legyen $V' := V - v$. Indukciót használva kapjuk, hogy létezik olyan $X_0' \subseteq V - v$, amelyre $\nu(G') = (|V'| - q'(X_0') + |X_0'|)/2$, ahol $q'(X_0')$ a $G' - X_0'$-ben jelöli a páratlan komponensek számát. Legyen $X_0 := X_0' + v$. Nyilván $q(X_0) = q'(X_0')$. Ezeket összevetve kapjuk:

$$\nu(G) - 1 \geq \nu(G') = (|V'| - q'(X_0') + |X_0'|)/2 = (|V| - q(X_0) + |X_0| - 2)/2\,,$$

ami éppen (15).

2. eset *G kritikus.* A Gallai lemma alapján $\nu(G) = (|V| - 1)/2$. Tehát $X_0 := \emptyset$ választással

$$\nu(G) = (|V| - 1)/2 \geq (|V| - q(X_0) + |X_0|)/2\,,$$

azaz (15) fennáll. □

4.2 A súlyozott párosítások problémája

Hogyan lehet nempáros gráfban egy maximális súlyú (teljes) párosítást megtalálni, és egyáltalán mi mondható a maximális súlyú párosítás súlyáról vagy a minimális költségű teljes párosítás költségéről? Magyarán, mi Egerváry tételének nempáros gráfos megfelelője? A megoldáshoz J. Edmonds zseniális ötlete adja a kulcsot. Ennek lényege a következő. Tekintsük a teljes párosítások (incidencia vektorainak) konvex burkát. Írjuk ezt le egyenlőtlenségekkel, majd alkalmazzuk a lineáris programozás dualitás tételét. A terv nehézsége persze az egyenlőtlenségek megtalálásában van, de az adott problémára Edmondsnak ez fényesen sikerült.

Feltesszük, hogy a szereplő gráfok hurok-mentesek. Célunk tehát egyenlőtlenségekkel (azaz félterek metszeteként) megadni egy $G = (V, E)$ gráf párosításai illetve teljes párosításai incidencia vektorainak konvex burkát, melyeket rendre B_P illetve B_{TP}-vel fogunk jelölni. Továbbiakban rövidség kedvéért egyszerűen csak párosítások konvex burkáról beszélünk. Miután tetszőleges politop felírható véges sok féltér metszeteként, tudjuk, hogy létezik ilyen leírás. Ennek explicit megadása azzal az előnnyel jár, hogy a lineáris programozás dualitás tételét alkalmazva egy min-max tételt nyerhetünk a maximális súlyú párosítás illetve a minimális súlyú teljes párosítás súlyára.

Páros gráfok esetén a páros gráf incidencia mátrixának teljes unimodularitása miatt az $\{x : x \in \mathbf{R}_+^E,\ d_x(v) \leq 1 \text{ minden } v \in V \text{ csúcsra}\}$ poliéder csúcsai egészek, és így $0-1$ vektorok, vagyis párosítások incidencia vektorai, tehát ez a poliéder éppen a párosítás poliéder. Hasonlóképp, ha a $d_x(v) \leq 1$ egyenlőtlenségeket egyenlőségekre cseréljük, úgy megkapjuk a páros gráf teljes párosítás poliéderét. Nem páros gráf esetén azonban az így kapott poliéderek már nem írják le B_P-t illetve B_{TP}-t. Például egy háromszög esetén a mindenütt 1/2 vektorra $d_x(v) = 1$ teljesül, de ez nem lehet B_T-ben, mert abban bármely vektorban a komponensek összege legfeljebb 1 lehet.

Jelölje $\mathcal{O}$ a V legalább háromelemű páratlan elemszámú részhalmazainak rendszerét. Ennek bármely Z tagja olyan, hogy tetszőleges párosítás legfeljebb $(|Z| - 1)/2$ darab Z által feszített élt tartalmaz és tetszőleges teljes párosítás páratlan sok, így legalább egy élt tartalmaz a $[Z, V - Z]$ vágásból.

4.3 A teljes párosítás poliéder

Legyen $G = (V, E)$ teljes párosítással rendelkező irányítatlan gráf. Jelölje a teljes párosítások (incidencia vektorainak) konvex burkát P_G. Tekintsük a következő poliédert

$$P' := \tag{16}$$

$$\{\, x \in \mathbf{R}^E,\ x \geq 0,\ d_x(v) = 1 \text{ minden } v \in V \text{ csúcsra}, \tag{17}$$

$$d_x(Z) \geq 1 \text{ minden } Z \in \mathcal{O}\text{-ra.}\,\} \tag{18}$$

4.4 Tétel (Edmonds). $B_{TP} = P'$

Bizonyítás. Paritási megfontolásból adódik, hogy egy M teljes párosítás és egy páratlan elemszámú Z halmaz esetén $d_M(Z)$ páratlan és így legalább 1. Emiatt $B_{TP} \subseteq P'$. A fordított irányú tartalmazás igazolásához tekintsünk P'-nek egy x elemét. Erről fogjuk kimutatni, hogy előáll teljes párosítások konvex kombinációjaként.

Feltehetjük, hogy x minden élen pozitív, mert azon éleket, ahol 0, kihagyhatjuk a gráfból. Nevezzünk egy $Z \subseteq V$ halmazt *pontosnak* (x-re nézve), ha $|Z|$ páratlan és $d_x(Z) = 1$. Egy pontos halmaz komplementere is pontos. Z pontos halmaz *valódi*, ha $|Z| \geq 3, |V - Z| \geq 3$. Jelölje $d(Z, Z')$ a $Z - Z'$ és $Z' - Z$ között vezető élek számát.

4.5 Lemma. *Ha Z és Z' olyan pontos halmazok, melyek metszete páratlan elemszámú, akkor metszetük és uniójuk is pontos. Továbbá ilyenkor $d(Z, Z') = 0$.*

Bizonyítás. Ha a metszet páratlan, akkor az unió is, ezért $1 + 1 = d_x(Z) + d_x(Z') = d_x(Z \cap Z') + d_x(Z \cup Z') + 2d_x(Z, Z') \geq 1 + 1 + 0$, amiből a lemma következik. □

Nevezzünk egy M párosítást (az x-re nézve) *feszesnek*, ha M teljes párosítás és $d_M(Z) = 1$ fennáll minden Z pontos halmazra. A tétel bizonyításához arra lesz csak szükségünk, hogy létezik feszes párosítás, de kényelmesebb kicsit többet igazolni:

4.6 Lemma. *Minden $e = uv \in E$ élhez létezik e-t tartalmazó feszes párosítás.*

Bizonyítás. Először nézzük meg azt az esetet, amikor nem létezik valódi pontos halmaz. Ekkor azt kell igazolnunk, hogy létezik e-t tartalmazó teljes párosítás, hiszen ez automatikusan feszes. Ha indirekt nem létezne ilyen, akkor Tutte tétele nyomán létezik egy olyan $A \subseteq V - \{u, v\}$ halmaz, amelyre a $G - \{u, v\} - A$ gráfban a páratlan komponensek t száma nagyobb, mint A elemszáma. Legyen $A' := A \cup \{u, v\}$, jelölje a szóban forgó páratlan komponenseket $Z_1, Z_2, \ldots, Z_t$, és legyen E' az A' és a páratlan komponensek között vezető élek halmaza. Mivel $|V|$ páros, t és $|A|$ ugyanolyan paritású, így $t \geq |A| + 2 = |A'|$. Ekkor egyrészt $x(E') = \sum(d_x(Z_i) : i = 1, \ldots, t) \geq t$, másrészt $x(E') \leq \sum(d_x(w) : w \in A') - x(uv) \leq |A'| - x(uv) < |A'|$, amiből $|A'| > t$ adódik, és ez az ellentmondás bizonyítja a lemmát abban az esetben, ha nincs valódi pontos halmaz.

Tegyük most fel, hogy Z_1 valódi pontos halmaz. Legyen $Z_2 := V - Z_1$ és jelölje rendre $G_1 = (Z_1 + z_2, E_1)$ és $G_2 = (Z_2 + z_1, E_2)$ a Z_2 illetve a Z_1 halmazok összehúzásával keletkező gráfokat, ahol z_i a Z_i halmaz összehúzásával keletkező csúcsot jelöli. Jelölje x_1 illetve x_2 az x vektor megszorítását az E_1 illetve az E_2 éleire. Miután $d_x(Z_1) = 1$, a G_i minden csúcsára teljesül $d_{x_i}(v) = 1$. Továbbá G_i minden páratlan elemszámú Z halmazára $d_{x_i}(Z) \geq 1$.

Tegyük először fel, hogy az $e = uv$ él mindkét vége ugyanabban a Z_i-ben van, mondjuk Z_1-ben. Indukcióval létezik G_1-nek egy M_1 feszes párosítása, amely tartalmazza e-t. Legyen f' az M_1-nek a z_2-t fedő éle. Jelölje f a G-nek azt az élét, amelyből a Z_2 összehúzásakor f' keletkezett és jelölje f'' a G_2-nek azt az élét, amely f-ből keletkezett a Z_1 összehúzásakor. (Vagyis f'' szomszédos a z_1 csúccsal.) Indukcióval létezik G_2-ben egy f''-t tartalmazó M_2 feszes párosítása. Ekkor $M := (M_1 - f') \cup (M_2 - f'') + f$ teljes párosítása G-nek, amely tartalmazza e-t.

Abban az esetben, ha e végpontjai különböző Z_i-khez tartoznak, akkor G_1-ben létezik e'-t tartalmazó M_1 feszes párosítás és G_2-ben létezik e''-t tartalmazó M_2 feszes párosítás és ilyenkor $M := (M_1 - e') \cup (M_2 - e'') + e$ teljes párosítása G-nek, amely tartalmazza e-t.

Állítjuk, hogy mindkét esetben M feszes. Valóban, tetszőleges $Z \subseteq V$ páratlan elemszámú halmazra $|Z \cap Z_1|$ és $|Z \cap Z_2|$ egyike páratlan, mondjuk $|Z \cap Z_1|$. A 4.5 lemma alapján $|Z \cup Z_1|$ és így $|Z_2 - Z|$ is páratlan és pontos, továbbá $|Z \cap Z_1|$ pontos és $d_x(Z, Z_1) = 0$. Emiatt $d_M(Z \cap Z_1) = d_{M_1}(Z \cap Z_1) = 1$, $d_M(Z_2 - Z) = d_{M_2}(Z_2 - Z) = 1$ és $d_M(Z, Z_1) = 0$. Ebből $1 + d_M(Z) = d_M(Z_1) + d_M(Z) = d_M(Z_1 \cap Z) + d_M(Z_1 \cup Z) + 2d_M(Z_1, Z) = 1 + 1 + 0$, amiből $d_M(Z) = 1$ következik, vagyis M valóban feszes. □

A tétel bizonyításához visszatérve, azt feltettük, hogy x mindenütt pozitív. Azt is feltehetjük, hogy G összefüggő. A tétel nyilvánvalóan igaz, ha G-nek egyetlen éle van. Tegyük fel tehát, hogy nem ez a helyzet. Ekkor $(*)$ x-nek van olyan komponense, amely kisebb, mint 1.

A 4.6 lemma szerint létezik egy M feszes párosítás. Tekintsük az $x_\alpha := (x - \alpha\chi_M)/(1 - \alpha)$ vektort. Mivel M teljes párosítás minden v csúcsra $d_{x_\alpha}(v) = 1$ teljesül bármely α értékre. Mivel M feszes, kicsiny pozitív α-ra $d_{x_\alpha}(Z) \geq 1$ fennáll minden páratlan Z-re. Mivel x mindenütt pozitív, kicsiny pozitív α-ra $x_\alpha \geq 0$. Válasszuk α-t maximálisra úgy, hogy $x_\alpha \in P'$, vagyis α a legnagyobb olyan szám, amelyre $\alpha \leq x(e)$ minden $e \in E$ élre és $(d_x(Z) - \alpha d_M(Z))/(1 - \alpha) \geq 1$ (Konkrétan $\alpha := \min\{\alpha_1, \alpha_2\}$, ahol $\alpha_1 = \min(x(e) : e \in E)$, és $\alpha_2 := \min\{(d_x(Z) - 1))/(d_M(Z) - 1) : Z$ nem-pontos páratlan halmaz.$\}$) Ekkor $(*)$ miatt $0 < \alpha < 1$, $x_\alpha \in P'$ és az α maximális választása miatt vagy x_α-nak van 0 komponense vagy x_α-ra nézve szigorúan több pontos halmaz létezik. Így indukcióval feltehetjük, hogy x_α benne van B_{TP}-ben, azaz előáll teljes párosítások konvex kombinációjaként. De akkor $x = \alpha\chi_M + (1 - \alpha)x_\alpha$ miatt x is előáll teljes párosítások konvex kombinációjaként. □

Tegyük fel, hogy $c : E \to \mathbf{R}_+$ egy adott súly- (vagy költség) függvény.

A teljes párosítás poliéder leírását és a dualitás tételt felhasználva kapjuk az alábbi formulát a teljes párosítások minimális súlyára.

4.7 Tétel (Edmonds). *A teljes párosítások minimális súlya egyenlő az alábbi maximummal:*

$$\max \sum \bigl(y(Z) : Z \subseteq V,\ |Z| \text{ páratlan}\bigr) , \tag{19}$$

ahol $y(Z) \geq 0$, *ha* $|Z| > 1$ *és minden* $e = uv \in E$ *élre*

$$\sum \bigl(y(Z) : |Z \cap \{u, v\}| = 1\bigr) \leq c(e) .$$

□

4.4 A párosítás poliéder

A teljes párosítások poliéderének leírását felhasználva egyszerű elemi konstrukció segítségével megadhatjuk a párosítások poliéderét. Jelölje $i_x(Z)$ a Z által feszített éleken az x komponenseinek összegét.

4.8 Tétel. *Tetszőleges* $G = (V, E)$ *gráf párosításai incidencia vektorai által feszített politop pontosan a következő poliéder:*

$$P_0 := \{\, x \in \mathbf{R}^E : x \geq 0 , \tag{20}$$

$$d_x(v) \leq 1 \text{ minden } v \in V\text{-re}, \tag{21}$$

$$i_x(Z) \leq (|Z| - 1)/2 \text{ minden páratlan elemszámú } Z \subseteq V \text{ halmazra}\} . \tag{22}$$

Bizonyítás. Egy M párosítás $x := \chi_M$ incidencia vektorára nyilván mindkét egyenlőtlenség fennáll, így a párosítások konvex burkának elemeire is.

Megfordítva, legyen most x egy olyan vektor, amely az egyenlőtlenségeket kielégíti. Készítsük el a G gráfot két példányban, melyeket jelölje $G' = (V', E')$, $G'' = (V'', E'')$. Kössük össze a megfelelő v' és v'' pontokat egy-egy éllel és az így kapott gráfot (a $V' \cup V''$ ponthalmazon) jelöljük H-val. Készítsünk el egy y vektort a H gráf élein. G minden e élére legyen $y(e') := y(e'') := x(e)$, ezenkívül minden $v \in V$ csúcsra legyen $y(v'v'') := 1 - d_x(v)$.

Állítjuk, hogy y benne van a H teljes párosítás poliéderében. A definícióból nyilvánvaló, hogy $d_y(u) = 1$ fennáll H minden csúcsára. Legyen most $Z = A' \cup B''$ egy páratlan elemszámú halmaz. Ekkor $A - B$ és $B - A$ egyike páratlan elemszámú, mondjuk $A - B$, és így, $i_x(A - B) \leq (|A - B| - 1)/2$ miatt, $d_y(Z) \geq \sum[d_y(v') : v' \in (A' - B')] - 2i_y(A' - B') - d_y(A' - B', A' \cap B') + d_y(A'' - B'', A'' \cap B'') = |A - B| - 2i_x(A - B) \geq 1$. A 4.4 tétel miatt y a H teljes párosításainak konvex kombinációja, ami miatt ezen párosításokat a G-re megszorítva az x a G párosításainak konvex kombinációjaként áll elő. □

Figyeljük meg, hogy a tétel azzal ekvivalens, hogy a P_0 poliéder egész. Ez ugyanis azzal egyenértékű, lévén P_0 korlátos, hogy P_0 csúcsai egészek, ami a (21) miatt azt jelenti, hogy P_0 csúcsai $0-1$-vektorok. Miután a P_0-ban lévő (így (21)-et) kielégítő $0-1$-vektorok pont a párosítások incidencia vektorai, a tétel valóban azzal ekvivalens, hogy P_0 egész poliéder. A 4.8 tételben megadott leírást és a dualitás tételt összetéve, a következő eredmény adódik.

4.9 Tétel (Edmonds). *Legyen $G = (V, E)$ gráf élein adott a $c : E \to \mathbf{R}$ súlyfüggvény. A gráf maximális súlyú párosításának súlya egyenlő a*

$$\min \sum_{v\in V} \pi(v) + \sum_{Z\in\mathcal{O}} y(Z)(|Z|-1)/2) \tag{23}$$

értékével, ahol $\pi : V \to \mathbf{R}_+$, $y : \mathcal{O} \to \mathbf{R}_+$ nemnegatív és minden $uv \in E$ élre

$$\pi(u) + \pi(v) + \sum\big(y(Z) : u, v \in Z\big) \geq c(uv)\,.$$

□

Az Edmonds-féle program a párosítások esetére tehát sikerrel járt: egyenlőtlenségekkel leírtuk mind a párosítások, mind a teljes párosítások konvex burkának poliéderét, és a lineáris programozás dualitás tétele segítségével karakterizálni lehetett az optimum értékeket.

Ha mindezt algoritmikus szempontból is szeretnénk hasznosítani, rögtön nagy problémába ütközünk. A szóban forgó lineáris programokban exponenciálisan sok egyenlőtlenség szerepel (minden páratlan elemszámú halmaz definiál egyet). Tehát például egy száz pontú gráfnál egyszerűen le sem tudjuk explicit írni (elfogadható időben és helyen) ezen egyenlőtlenségeket, nemhogy mondjuk a szimplex módszer alkalmazásáról egyáltalán beszélhetnénk. J. Edmonds másik nagyszabású felismerése az volt, hogy az egyenlőtlenségek explicit felsorolására valójában nincs szükség. Megadott egy olyan direkt, konstruktív bizonyítást a 4.9 dualitás tételre, amely egyúttal polinomiális futásidejű algoritmust is eredményez. Eljárása az eredeti Magyar Módszer nagymérvű általánosításának tekinthető.

Megjegyzendő, hogy amiképp a Magyar Módszer alapgondolata kiterjeszthető a szállítási feladatra, ugyanúgy Edmonds eljárása is általánosítható minimális költségű fokszám-korlátozott részgráf keresésére tetszőleges gráf esetén.

4.5 Duális egészértékűség

Páros gráfok esetén a párosítás poliédert egy teljesen unimoduláris mátrix írta le, és emiatt egész célfüggvény esetén a duális lineáris problémának is mindig létezett egészértékű optimuma. A 4.4 tételben megadott primál poliéderről, bár a megadott leíró rendszere nem teljesen unimoduláris, mégis kimutattuk, hogy egész (merthogy a csúcsai a teljes párosítások). Ennek nyomán felvetődik a kérdés, vajon egész c esetén a duális feladatnak is mindig létezik-e egész optimuma. A válasz sajnálatos módon nemleges, amint azt a teljes négypontú gráf mutatja a $c \equiv 1$ súlyozás esetén. Ekkor a primál optimum értéke 2 (bármely teljes párosítás súlya 2), a duál optimum egyetlen megoldása (amint az könnyen ellenőrizhető) az, amikor az y mind a négy egyelemű halmazon 1/2, másutt 0. Az egészértékű duális optimum értéke 1 vagyis kisebb, mint 2.

Ennek fényében meglepő, hogy egészértékű c esetén a 4.9 tételben szereplő duális optimum mindig eléretik egész vektoron is.

4.10 Tétel (Cunningham és Marsh). *Legyen $G = (V, E)$ gráf élein adott a $c : E \to \mathbf{Z}$ súlyfüggvény. A gráf maximális súlyú párosításának súlya egyenlő a*

$$\min \sum_{v \in V} \pi(v) + \sum_{Z \in \mathcal{O}} y(Z)(|Z| - 1)/2) \tag{25}$$

értékével, ahol $\pi : V \to \mathbf{Z}_+$, $y : \mathcal{O} \to \mathbf{Z}_+$ nemnegatív, egészértékű, és minden $uv \in E$ élre

$$\pi(u) + \pi(v) + \sum(y(Z) : u, v \in Z) \geq c(uv) \,. \tag{26}$$

Létezik olyan duális optimum, amelyre az $\mathcal{O}_y := \{Z \in \mathcal{O}, y(Z) > 0\}$ halmazrendszer lamináris (vagyis bármely két tagja vagy diszjunkt, vagy az egyik tartalmazza a másikat).

Bizonyítás. Jelölje a maximális súlyú párosítás súlyát ν_c. Adott (y, π) duális megoldásra legyen

$$b(y, \pi) := \sum_{v \in V} \pi(v) + \sum_{Z \in \mathcal{O}} y(Z)(|Z| - 1)/2 \,. \tag{27}$$

A dualitás tétel triviális max $\leq$ min iránya miatt $\nu_c \leq b(y, \pi)$.

A tétel első részének igazolásához tegyük fel indirekt, hogy a min-max összefüggés nem érvényes és válasszunk egy olyan (G, c) ellenpéldát, amelyre a $|V| + |E| + \sum_{e \in E} |c(e)|$ összeg minimális. Ekkor minden e élre $c(e) \geq 1$ (különben a nempozitív élek kitörölhetők.) A most következő állítások mind erre a legkisebb ellenpéldára vonatkoznak.

4.11 Állítás. *Minden v csúcsot elkerül maximális súlyú párosítás.*

Bizonyítás. Tegyük fel indirekt, hogy a v csúcsot minden maximális súlyú párosítás fedi. Ekkor a c-értékeket a v végű éleken eggyel csökkentve a maximális súlyú párosítás súlya is csökken, éspedig eggyel, azaz a keletkező c' súlyfüggvényre nézve $\nu_{c'} = \nu_c - 1$. Mivel (G, c') már nem ellenpélda, így létezik egy egészértékű (y', π') duális optimális megoldás, amelyre $b(y', \pi') = \nu_{c'}$. Ha most π' értékét a v csúcson eggyel megnöveljük, akkor olyan (y', π'') duális megoldást kapunk c-re nézve, amelyre $\nu_c = \nu_{c'} + 1 = b(y', \pi') + 1 = b(y', \pi'') \geq \nu_c$, amiből $\nu_c = b(y', \pi'')$ adódik, ellentmondásban (G, c) ellenpélda voltával. □

4.12 Állítás. *Legyen (y, π) egy optimális duális megoldás és M egy maximális súlyú párosítás. Ekkor M fed minden olyan v csúcsot, amelyre $\pi(v) > 0$. Továbbá $y(Z) > 0$ esetén az M-nek azon élei, melyek Z-be esnek a Z-nek egy majdnem teljes (azaz $(|Z| - 1)/2$ elemű) párosítását adják.*

Bizonyítás. A 4.8 tétel szerint a primál optimum párosításon felvétetik. Az állítás nem más, mint a duál egyenlőtlenségeknek megfelelő optimalitási kritérium. □

A fenti két állítást összevetve kapjuk, hogy

4.13 Állítás. *Tetszőleges (y, π) optimális duális megoldás esetén $\pi \equiv 0$.* $\square$

4.14 Állítás. *Létezik olyan (y, π) optimális duális megoldás, amelyre az*

$$\mathcal{O}_y := \{ Z \in \mathcal{O},\ y(Z) > 0 \}$$

halmaz-rendszer lamináris.

Bizonyítás. Induljunk ki egy (y, π) duális optimumból. Az y mindenesetre választható racionálisnak, hiszen egy bázis-megoldás racionális és mindig létezik optimális bázis-megoldás. (A 4.13 állítás szerint $\pi \equiv 0$.)

Tegyük fel, hogy $\mathcal{O}_y$ nem lamináris, azaz tartalmaz két halmazt, A-t és B-t, melyekre $A \cap B, A - B, B - A$ egyike sem üres. Állítjuk, hogy $|A \cap B|$ páratlan. Ha ugyanis páros lenne, akkor egy $v \in A \cap B$ csúcsra a 4.12 állítás első része miatt létezik maximális súlyú v-t elkerülő M párosítás. Másrészt, a 4.12 állítás második része miatt M-nek A-ba illetve B-be eső élei teljes párosítását adják $A - v$-nek illetve $B - v$-nek, de akkor $|A \cap B - v|$ páros.

Jelölje az $y(A)$ és $y(B)$ értékek minimumát α, és csökkentsük mind $y(A)$-t, mind $y(B)$-t α-val. Most $|A \cap B|$ páratlan, így persze $|A \cup B|$ is az. Növeljük $y(A \cup B)$-t α-val, és amennyiben $|A \cap B| \geq 3$, úgy $y(A \cap B)$-t is növeljük α-val. Ezt az átalakítást egy kikeresztezési lépésnek nevezzük. Könnyen ellenőrizhető, hogy a létrejövő (y', π') teljesíti (26)-ot. Azt sem nehéz belátni, hogy $b(y, \pi) = b(y', \pi')$, azaz (y', π') is optimális duális megoldás.

Ismételjük ezt az eljárást egészen addig, amíg a szóban forgó $\mathcal{Z}_y$ halmazrendszer nem lamináris. Állítjuk, hogy a kikeresztezési eljárás véges sok lépés után véget ér. Ezt elég csak arra az esetre igazolni, amikor y egészértékű, mert ha nem az, akkor komponenseinek közös nevezőjével felszorozhatjuk. Egy kikeresztezési lépésnél a $\sum(y(Z)|Z| : Z \in \mathcal{O})$ összeg vagy változatlanul marad (amikor $|A \cap B| > 1$) vagy csökken (ha $|A \cap B| = 1$). Világos, hogy ezen utóbbi eset csak véges sokszor fordulhat elő. Az első esetben viszont a $\sum(y(Z)|Z|^2 : Z \in \mathcal{O})$ kísérő összeg szigorúan nő, így véges sok ilyen lépés után a $\sum(y(Z)|Z| : Z \in \mathcal{O})$ összegnek kell változnia, tehát kikeresztezési lépésből valóban csak véges sok lehet. $\square$

Rendelkezésünkre áll tehát egy $(y,\ \pi = 0)$ duális optimális megoldás, amelyre $\mathcal{O}_y$ lamináris. Mivel (G, c) ellenpélda, y nem egészértékű, azaz van olyan $A \in \mathcal{O}_y$ halmaz, amelyre $y(A)$ nem egész, azaz $\beta := y(A) - \lfloor y(A) \rfloor$ pozitív. Tegyük fel, hogy A maximális ilyen. Módosítsuk most y-t úgy, hogy $y(A)$-t csökkentjük β-val, míg $\mathcal{O}_y$ azon tagjain (ha vannak egyáltalán), melyek részhalmazai A-nak és maximális ilyenek (és melyek a laminaritás miatt diszjunktak) az y értéket növeljük β-val. Az A maximális választása és c egészértékűsége miatt így egy másik duális megoldást kapunk, de ez ellentmond y optimális voltának, hiszen ezen változtatás során a $\sum_{X \in \mathcal{O}_y} y(X)(|X|-1)/2$ összeg szigorúan csökken. Az ellentmondás mutatja, hogy nem létezhet ellenpélda.

A második részhez csak azt kell észrevennünk, hogy a fent leírt kikeresztezési eljárás az egészértékűséget nem rontja el, így egy tetszőleges egészértékű duális optimumból kiindulva végül kapott „lamináris" optimum is egészértékű. $\square\,\square$

5 Matroidok

5.1 A mohó algoritmus

A Magyar Módszer egy másik irányú általánosításának ismertetése előtt érdemes felidézni, hogy a kombinatorikus optimalizálásnak van egy másik kiinduló pontja is, éspedig Kruskal közismert algoritmusa összefüggő élsúlyozott gráf maximális vagy minimális súlyú feszítő fájának meghatározására. Ez az eljárás az úgynevezett mohó algoritmus, és szintén alapvető eredménynek számít, mégha jóval egyszerűbb is, mint a Magyar Módszer.

Nem meglepő, hogy nemcsak a Magyar Módszernek találtak kiterjesztéseit, hanem a mohó algoritmusnak is. Kiderült például, hogy a mohó algoritmus valójában minden matroidon helyes eredményt ad. A matroid egy olyan absztrakt struktúra, amely egy $(S, \mathcal{F})$ párral adható meg, ahol $\mathcal{F}$ az S alaphalmaz bizonyos, függetlennek hívott, részhalmazainak olyan rendszere, amelyre érvényesek az úgynevezett függetlenségi axiómák:

(I1) az üres halmaz független,

(I2) egy független halmaz bármely részhalmaza független,

(I3) az S valamennyi Z részhalmazára, a Z-be eső, Z-ben már nem bővíthető független halmazok elemszáma ugyanaz a csupán Z-től függő $r(Z)$ szám, (amit a Z halmaz rangjának neveznek, míg egy Z-ben maximális független halmazt Z egy bázisának).

Például egy gráf élhalmazán az erdők, mint független halmazok matroidot alkotnak, vagy egy mátrix oszlopainak halmazán a lineárisan független részhalmazok szintén. A matroidok érdekességét egyrészt az adja, hogy számos helyen felbukkannak (néha ugyancsak rangrejtve), másrészt matroidokra igen elegáns, jól használható elméletet dolgoztak ki.

A mohó algoritmus segítségével tetszőleges c súlyfüggvényre egy matroid maximális súlyú bázisa kiszámítható. Ehhez tegyük fel, hogy az S elemei súly szerint csökkenő sorrendben vannak, azaz $c(s_1) \geq c(s_2) \geq \ldots \geq c(s_n)$. Ebben a sorrendben tekintsük az elemeket, és az aktuálisan vizsgált elemet akkor válasszuk ki, ha a már eddig kiválasztottakkal együtt független halmazt alkot. Igazolni lehet, hogy így bizonyosan maximális súlyú független halmazt kapunk. Jelölje $r(c)$ a maximális súlyú bázis súlyát a c súlyfüggvényre nézve. Az $r(c)$ tehát a mohó algoritmus révén könnyen számolható, sőt valójában egy egyszerű explicit formában is felírható. Ehhez legyen $S_i := \{s_1, \ldots, s_i\}$.

5.1 Tétel. *Tetszőleges c esetén a maximális súlyú bázis $r(c)$ súlyára*

$$r(c) = r(S)c(s_n) + \sum_{i=1}^{n-1} r(S_i)[c(s_i) - c(s_{i+1})] \,. \tag{28}$$

A matroidok használhatósága többek között azon múlik, hogy a matroidok osztálya számos műveletre nézve zárt. Például a gráfoknál alkalmazott élelhagyás vagy élösszehúzás fogalma átvihető matroidokra, vagy egy gráf sík-dualizálása elvezet a duális matroid fogalmához. Szintén érdekes tény, hogy a maximális c-súlyú bázisok egy M_c matroid bázisait alkotják. Ennek rangfüggvényét az alábbiakban r_c-vel fogjuk jelölni. (Egy adott Z részhalmazra $r_c(Z)$ tehát azt a maximális számot jelenti, ahány elemmel egy maximális súlyú bázis bele tud metszeni Z-be.)

5.2 Maximális elemszámú közös függetlenek

A mohó algoritmus matroidokon való helytállósága persze örömteli jelenség, az algoritmus hatósugara azonban túlságosan szűk. Például páros gráf maximális súlyú párosításának kiszámítására nem alkalmas. Másik baj, hogy már a feladat kis változtatásánál is hatástalanná válik. Fákkal kapcsolatosan például meglehetős természetességgel vetődik fel az a kérdés, hogy miként lehet egy minimális súlyú feszítő fát meghatározni, ha előírjuk, hogy egy rögzített s pontban a fának a fokszáma előre megadott szám legyen, vagy általánosabban, megadott korlátok közé essék. A következő példa mutatja, hogy ilyenkor a mohó algoritmus már nem feltétlenül ad jó eredményt. [A gráf legyen a teljes négyes az s, a, b, c pontokon, az élsúlyok legyenek $w(sa) = 1, w(sb) = 1, w(sc) = 2, w(ab) = 3, w(bc) = 10, w(ac) = 11$. A keresett fa foka az s pontnál 2 legyen. Ekkor a mohó algoritmus először kiválasztja a két 1 súlyú élt sa-t és sb-t, majd más lehetősége nem lévén választja a 10 súlyú bc élt. Az így kapott fa össz-súly 12. Ugyanakkor az sc, sa, ab élekből álló fa össz-súlya csak 5.]

Még nehezebb a következő kérdés: mikor létezik egy gráfban két diszjunkt feszítő fa, és ha létezik hogyan lehet megtalálni azt a két élidegen fát, melyek összköltsége minimális. Persze ugyanezt a kérdést kettő helyett bármely k pozitív egészre megfogalmazhatjuk. Sőt még azt az általánosítást is tekinthetjük, ahol az élhalmazon nem egy, hanem k költség-függvény adott és úgy kell kiválasztanunk k élidegen feszítő fát, hogy az i-ediknek az i-dik költség-függvényre vonatkozó költségét tekintjük, és ezek összegét akarjuk minimalizálni.

Irányított gráfokban tűzhető ki az alábbi feladat: Határozzunk meg egy minimális költségű részgráfot, amelyben minden pont elérhető egy előre adott gyökérpontból. Általánosabban olyan minimális költségű részgráfot keresünk, amelyben minden pont k élidegen úton elérhető egy előre adott gyökérpontból.

Ezek mind olyan kérdések, melyek kívül esnek a Magyar Módszer (vagy más hálózati folyamos modell) hatókörén és megoldásukhoz egy új elméletre volt szükség. A kiindulópont J. Edmonds matroid-metszet tétele, amely a Kőnig tétel nagymérvű általánosítása matroidokra.

5.2 Tétel (Edmonds metszettétele). *Az S alaphalmazon adott két matroid. A közös független halmazok maximális elemszáma egyenlő a*

$$\min_{X \subseteq S}(r_1(X) + r_2(S - X)) \tag{29}$$

értékkel.

Edmonds tételét néha az alábbi ekvivalens alakban fogalmazzák meg.

5.3 Tétel (Edmonds). *Az S alaphalmazon adott két matroid, melyek rangfüggvénye r_1 és r_2. Akkor és csak akkor létezik legalább k elemű közös független halmaz, ha a*

$$r_1(X) + r_2(S - X) \geq k \tag{30}$$

fennáll minden $X \subseteq S$ halmazra.

Bizonyítás (Woodall). A (30) feltétel szükségessége kézenfekvő, hiszen tetszőleges F közös független halmazra és az alaphalmaz $\{X_1, X_2\}$ partíciójára fennáll, hogy $r_1(X_1) \leq |X_1 \cap F|$ és $r_2(X_2) \leq |X_2 \cap F|$, amiket összeadva kapjuk, hogy $r_1(X_1) + r_2(X_2) \leq |X_1 \cap F| + |X_2 \cap F| = |F|$.

Az elegendőséghez $|S|$ szerinti indukciót használunk. $|S| = 0$-ra a tétel nyilván igaz, így tegyük fel, hogy S nemüres és hogy a tétel érvényes minden olyan esetben, amikor az alaphalmaz $|S|$-nél kevesebb elemet tartalmaz. Válasszunk ki egy tetszőleges $s \in S$ elemet és legyen $S' := S - s$.

Amennyiben $r_1(X') + r_2(S' - X') \geq k$ fennáll minden $X' \subseteq S'$ részhalmazra, indukció alapján az $M_1|S'$ és $M_2|S'$ matroidoknak létezik k elemű közös függetlenje, amely természetesen M_1 és M_2-nek is közös függetlenje. Így feltehetjük, hogy létezik egy $X' \subseteq S'$ halmaz, amelyre

$$r_1(X') + r_2(S' - X') \leq k - 1\,. \tag{31}$$

Ebből következik, hogy s nem hurok egyik matroidban sem, mert ha mondjuk M_1-ben hurok volna, akkor $X := X' + s$-re $r_1(X) + r_2(S - X) = r_1(X') + r_2(S' - X') \leq k - 1$, ellentétben (30)-cal.

Húzzuk most össze mindkét matroidban az s elemet. Állítjuk, hogy a keletkező $M_i \cdot S'$ matroidok r'_i rangfüggvényére $r'_1(Y') + r'_2(S' - Y') \geq k - 1$ teljesül minden $Y' \subseteq S'$ részhalmazra. Álljon ugyanis valamelyik Y'-re $r'_1(Y') + r'_2(S' - Y') \leq k - 2$. Ez azzal ekvivalens, hogy $r_1(Y' + s) - 1 + r_2(S' - Y' + s) - 1 \leq k - 2$, azaz

$$r_1(Y' + s) + r_2(S - Y') \leq k\,. \tag{32}$$

Figyeljük meg, hogy $X' \cap Y'$ és $X' \cup (Y' + s)$ egymás komplementerei (S-re nézve), így (30) alapján $[r_1(X' \cap Y') + r_2((S' - X') \cup (S - Y'))] \geq k$, és hasonlóképp $X' \cup (Y' + s)$ és $(S' - X') \cap (S - Y')$ egymás komplementerei és ezért $r_1(X' \cup (Y' + s)) + r_2((S' - X') \cap (S - Y'))] \geq k$. Ezt és a szubmodularitást felhasználva (31) és (32) összeadásával kapjuk, hogy $(k - 1) + k \geq r_1(X') + r_1(Y' + s) + r_2(S' - X') + r_2(S - Y') \geq r_1(X' \cap (Y' + s)) + r_1(X' \cup (Y' + s)) + r_2((S' - X') \cap (S - Y')) + r_2((S' - X') \cup (S - Y')) = [r_1(X' \cap Y') + r_2((S' - X') \cup (S - Y'))] + [r_1(X' \cup (Y' + s)) + r_2((S' - X') \cap (S - Y'))] \geq k + k$, ami ellentmondás.

Az $M_i \cdot S'$ matroidokra tehát k helyén $(k - 1)$-gyel teljesül (30), így az indukciós feltevés alapján ezen matroidoknak létezik egy $k - 1$ elemű F közös függetlenje. De akkor $F + s$ az M_1 és M_2 matroidok k elemű közös függetlenje. $\square$

5.3 A súlyozott matroid metszet probléma

Amint említettük, a mohó algoritmus segítségével adott $c : S \to \mathbf{R}$ vektor esetén meg lehet határozni egyetlen matroid maximális súlyú bázisát. A fentiekben tételt adtunk két matroid közös függetlenjének maximális elemszámára. Kérdés, mi mondható két matroid közös független halmazainak maximális súlyáról. A probléma első megoldása szintén Edmondstól származik. Mi most egy egyszerűsített utat követünk: a kapott eredmény Egerváry tételének általánosítása. Valójában itt több kérdés is kitűzhető: keressünk maximális súlyú közös függetlent vagy maximális súlyú közös bázist, esetleg minden szóba jövő i értékre kíváncsiak lehetünk a maximális súlyú i elemű közös független halmazra. E feladatok többé-kevésbé ekvivalensek és most a maximális súlyú közös bázis feladatkörével foglalkozunk.

Már említettük, hogy a maximális c-súlyú bázisok egy M_c matroid bázisait adják, és hogy $r(c)$-t az (28) formula határozza meg. Az alábbi előkészítő lemma arra ad választ, hogy miként változik az $r(c)$ függvény, ha c értékeit egy adott Z halmaz minden elemén eggyel megnöveljük vagy lecsökkentjük.

5.4 Lemma. *Legyen az M matroid rangfüggvénye r. Adott $c : S \to \mathbf{Z}$ egészértékű vektorra és $Z \subseteq S$ halmazra legyen $c^+ := c + \chi_Z$ és $c^- := c - \chi_Z$. Ekkor*

$$r(c^+) = r(c) + r_c(Z) \ , \tag{33}$$

és

$$r(c^-) = r(c) - r(S) + r_c(S - Z) \ , \tag{34}$$

ahol r_c a maximális súlyú bázisok által alkotott matroid rangfüggénye.

Bizonyítás. Az első azonosság igazolásához rendezzük az S elemeit c-szerint csökkenő sorrendbe úgy, hogy egyenlő súlyú elemek esetén a Z elemeit vesszük előbbre. Mivel c egészértékű, ez a sorrend c^+ súlyozás szerint is csökkenő (azaz korábbi elem súlya nagyobb vagy egyenlő, mint egy későbbi elemé). Így a mohó algoritmus által szolgáltatott B bázis mind a c, mind a c^+ súlyozásra nézve maximális súlyú.

Ekkor persze $r_c(Z) \geq |B \cap Z|$, de itt valójában egyenlőségnek kell állnia, mert ha létezne olyan maximális c-súlyú B' bázis, amelyre $|B' \cap Z| > |B \cap Z|$, akkor $c^+(B') = c(B') + |Z \cap B'| > c(B) + |Z \cap B| = c^+(B)$, ellentmondásban avval, hogy B maximális c^+-súlyú. Tehát $r_c(Z) = |B \cap Z|$, és így $r(c+) = c^+(B) = c(B) + |B \cap Z| = r(c)r_c(Z)$.

A második azonossághoz először figyeljük meg, hogy $r_c = r_{c-\chi_S}$, majd alkalmazzuk az első azonosságot c helyén $c - \chi_S$-re és Z helyén $(S - Z)$-re. Kapjuk, hogy $r(c^-) = r(c - \chi_S + \chi S - Z) = r(c - \chi_S) + r_{c-\chi_S}(S - Z) = r(c) - r(S) + r_c(S - Z)$. $\square$

Térjünk most rá a maximális súlyú közös bázis kérdésére. Legyen ehhez adott az S alaphalmazon két matroid, M_1 és M_2, továbbá egy $c : S \to \mathbf{Z}$ egészértékű súly-függvény. Tegyük fel, hogy az M_1 és M_2 matroidoknak van közös bázisa, melynek elemszámát jelölje k.

5.5 Tétel [10]. *Az M_1 és M_2 matroidok közös bázisainak maximális c-súlya egyenlő a* $\min(r_1(c_1)+r_2(c_2) : c_1+c_2 = c,\ c_i$ *egészértékű) értékkel, ahol $r_i(x)$ az M_i matroidban a maximális x-súlyú bázis súlyát jelöli.*

Bizonyítás. Adott B közös bázis és c_1, c_2 esetén, nyilván $c(B) = c_1(B) + c_2(B) \leq r_1(c_1) + r_2(c_2)$, amiből a max $\leq$ min irány következik. Ebből a becslésből az is kiolvasható, hogy az egyenlőség igazolásához a c-nek olyan $c_1 + c_2$ felbontását kell találnunk, amelyre létezik M_1-nek és M_2-nek olyan B közös bázisa, amely

(∗) *egyszerre az M_1-nek maximális c_1-súlyú bázisa és az M_2-nek maximális c_2-súlyú bázisa.*

Legyen c_1 és c_2 a c-nek olyan egész felbontása, amelyre $r_1(c_1) + r_2(c_2)$ minimális. (Miután c_i egészértékű és tetszőleges B közös bázisra $c(B)$ alsó korlát $r_1(c_1)+r_2(c_2)$-re, a szóban forgó minimum létezik). Jelölje M'_i az M_i matroid maximális c_i-súlyú bázisai által alkotott matroidot ($i = 1, 2$), míg a megfelelő rangfüggvényeket jelöljük r'_i-vel.

5.6 Lemma. *Az M'_1 és M'_2 matroidoknak van k elemű közös független halmaza.*

Bizonyítás. Az Edmonds féle metszet tétel szerint, ha nem létezik k elemű közös független halmaz, akkor van olyan $Z \subseteq S$ halmaz, amelyre

$$r'_1(Z) + r'_2(S - Z) < k\ . \tag{35}$$

Legyen $c_1^+ := c_1 + \chi_Z$ és $c_2^- := c_2 - \chi_Z$. Alkalmazva a (33) formulát az M_1 matroidra és c_1 súlyfüggvényre illetve a (34) formulát az M_2 matroidra és a c_2 súlyfüggvényre, azt kapjuk, hogy $r_1(c_1^+) = r_1(c_1)+r'_1(Z)$ és $r_2(c_2^-) = r_2(c_2)+r'_2(S-Z)-r_2(S) = r_2(c_2)+r'_2(S-Z)-k$. Mindezeket összevetve az adódik, hogy $r_1(c_1^+)+r_2(c_2^-) = r_1(c_1)+r_2(c_2)+r'_1(Z)+r'_2(S-Z)-k < r_1(c_1)+r_2(c_2)$, ellentmondásban c_1 és c_2 minimális választásával. □

A lemma által biztosított közös B bázis tehát teljesíti a (∗) tulajdonságot. □ □

Érdemes külön is megfogalmazni a (∗) tulajdonságot.

5.7 Következmény. *Amennyiben két matroidnak van közös bázisa, úgy tetszőleges c egészértékű súlyozáshoz létezik c-nek egy egészértékű $c_1 + c_2$ felbontása valamint a két matroidnak egy B közös bázisa úgy, hogy B maximális c_1-súlyú bázisa M_1-nek és maximális c_2-súlyú bázisa M_2-nek.* □

A maximális súlyú közös független halmaz súlyára is megállapíthatunk egy formulát. (Fekete Zsolt és Makai Márton doktorandusz hallgatók segítségét a bizonyítás kidolgozásában ezúton is köszönöm.)

5.8 Tétel. *Az S alaphalmazon adott két matroid és egy c nemnegatív egészértékű súly-függvény. A maximális súlyú közös független halmaz súlya egyenlő a* $\min\{r_1(c_1) + r_2(c_2) : c_1 + c_2 = c,\ c_1 \geq 0,\ c_2 \geq 0,\ c_i$ *egész* $\}$ *értékkel.*

Bizonyítás. Legyen F közös független és legyen $c_1 \geq 0, c_2 \geq 0$ olyan, hogy $c_1 + c_2 = c$. Ekkor $c(F) = c_1(F) + c_2(F) \leq r_1(c_1) + r_2(c_2)$, amiből a max $\leq$ min irány következik.

A fordított irányhoz azt fogjuk megmutatni, hogy létezik c-nek egy olyan c_1, c_2 egészértékű, nemnegatív felbontása valamint egy F közös független halmaz, melyekre F maximális c_i-súlyú független halmaza M_i-nek ($i = 1, 2$).

Feltehetjük, hogy minden elem c-súlya szigorúan pozitív és egyik matroidban sincs hurok elem. (Miért?) Legyen $R := \max(r_1(S), r_2(S)) + 1$ és S' egy R elemű halmaz, amely diszjunkt S-től. Legyen M_i^+ ($i = 1, 2$) az a matroid, amelyet úgy kapunk, hogy az M_i és az S'-n vett szabad matroid direkt összegét R-rel csonkoljuk. Terjesszük ki a c-t az $S \cup S'$-re úgy, hogy az S' elemein legyen c azonosan nulla. Könnyen látszik, hogy M_i^+-ban egy R elemű B halmaz akkor és csak akkor bázis, ha $B \cap S$ független M_i-ben, és ekkor persze $c(B) = c(B \cap S)$. (Speciálisan S' nulla súlyú bázis.) Ennek megfelelően M_1 és M_2 maximális súlyú közös független halmazának a súlya egyenlő az M_1^+ és M_2^+ matroidok maximális súlyú közös bázisának a súlyával.

Az 5.7 következmény szerint létezik egy olyan közös B bázis és c_1, c_2 egész felbontása a kiterjesztett c-nek, hogy B maximális c_i-súlyú bázisa M_i^+-nak.

5.9 Állítás. *c_i ($i = 1, 2$) választható olyannak, hogy S' elemein azonosan nulla.*

Bizonyítás. Kimutatjuk, hogy S' minden elemének c_1 értéke ugyanaz. Az R érték választása miatt $B \cap S'$ nemüres. Az $S - B$ sem lehet üres, vagyis B nem lehet teljesen S'-ben, mert akkor egyrészt $c(B) = 0$ volna, ugyanakkor amiatt, hogy minden S-beli elem c-súlya pozitív, és nincs hurok elem, létezik pozitív súlyú közös bázis. A B ezen elhelyezkedése miatt elég kimutatnunk, hogy egy $B \cap S'$-beli x elem és egy $S' - B$-beli y elem c_1-súlya megegyezik. Miután $B - x + y$ egy másik közös bázis, kapjuk, hogy $c_1(y) \leq c_1(x), c_2(y) \leq c_2(x)$, amiből $c_1(x) + c_2(x) = 0 = c_1(y) + c_2(y)$ miatt $c_1(y) = c_1(x)$ következik. □

Feltesszük tehát, hogy c_1 és c_2 azonosan nulla az S' elemein. Ebből következik, hogy minden $x \in S \cap B$ elemre $c_i(x) \geq 0$, hiszen egy $y \in S' - B$ elemre $B - x + y$ bázisa M_i^+-nak, és ezért $c_i(x) \geq c_i(y) = 0$.

Legyen most $x \in S - B$ egy olyan elem, amelyre, mondjuk, $c_1(x)$ negatív. Növeljük $c_1(x)$-t nullára, $c_2(x)$-t pedig csökkentsük $c(x)$-re. B továbbra is maximális súlyú bázis marad M_2^+-ban, hiszen egy bázison kívüli elemen csökkentettük a súlyt. B maximális súlyú bázis marad M_1^+-ban is, hiszen B minden elemének a c_1-súlya nemnegatív és egy negatív súlyú elem súlyát növeltük nullára.

Ilyen módosításokkal elérhetjük, hogy c_i nemnegatív. Kapjuk, hogy c_i-nek az S-re való $c_i|S$ megszorítására nézve az $F := B \cap S$ közös független halmaz maximális $c_i|S$-súlyú független az M_i matroidban. □□

J. Edmonds további érdeme, hogy mind a súlyozatlan, mind az általános súlyozott esetre kidolgozott egy polinomiális futásidejű algoritmust, amely egyúttal a tételeinek bizonyítására is használható.

Az Edmonds féle eredeti súlyozott matroid-metszet tétel a fentebb szereplőnél jóval bonyolultabb, így nem csoda, hogy Edmonds súlyozott metszet algoritmusa (és helyességének bizonyítása) pedig különösen az. A nehézség fő forrása abból ered, hogy szemben az eredeti Egerváry tétellel, az idevonatkozó duális problémát csak meglehetősen bonyolult alakban sikerült megadni. A fenti súly-szétvágós min-max tétel nemcsak egyszerű alakú, de lehetőséget teremtett egy viszonylag egyszerűen megfogalmazható és igazolható súlyozott matroid metszet algoritmus megadására [10]. Ez az algoritmus a Magyar Módszer nagyfokú és direkt általánosításának tekinthető.

6 Szubmoduláris modellek

6.1 Szubmoduláris áramok

A súlyozott matroid metszet tétel és a rá vonatkozó algoritmus számos nehéz dolog megoldását tette lehetővé (például az előbb felsorolt gráf optimalizálási feladatokét). De azért nem mindent! Nézzük például azt a feladatot, amikor egy irányított gráfot kell erősen összefüggővé tennünk minimális számú, vagy általánosabban, minimális összköltségű élének összehúzásával. Egy másik természetes feladat a következő: Határozzunk meg egy irányított gráf minimális összköltségű részgráfját, amelyben egy meghatározott gyökérponttól minden más csúcsba vezet k pontdiszjunkt út.

Ezek a feladatok már a súlyozott matroid metszet probléma segítségével sem oldhatók meg. Kifejlesztettek azonban egy még általánosabb keretet, amely magában foglalja a súlyozott matroid metszet problémát és a minimális költségű folyamok problémát is. Ez pedig a szubmoduláris áramok fogalma, amelyet Edmonds és Giles vezetett be 1976-ban [7].

Legyen $D = (V, E)$ irányított gráf és $b : V \to \mathbf{Z}$ egy keresztező szubmoduláris függvény, azaz $b(X)+b(Y) \geq b(X\cap Y)+b(X\cup Y)$ fennáll minden olyan $X, Y \subseteq V$ halmazpárra, amelyre $X \cap Y$ és $V-(X\cup Y)$ egyike sem üres. Legyen f és g a digráf élhalmazán értelmezett egészértékű alsó illetve felső korlát. Egy $x : E \to \mathbf{R}$ vektort szubmoduláris áramnak hívnak, ha $\varrho_x(Z) - \delta_x(Z) \leq b(Z)$ minden $Z \subseteq V$ halmazra fennáll. A szubmoduláris áram *megengedett*, ha $f \leq x \leq g$. Ha b azonosan nulla, visszajutunk a közönséges áram fogalmához. Edmonds és Giles többek között megmutatta, hogy a szubmoduláris áramok poliédere egész. Hoffman 3.6 megengedettségi tétele is szépen átmegy szubmoduláris áramokra [11].

6.1 Tétel. *Teljesen szubmoduláris b esetén akkor és csak akkor létezik egészértékű megengedett szubmoduláris áram, ha*

$$\varrho_f(A) - \delta_g(A) \leq b(A) \tag{36}$$

fennáll minden $A \subseteq V$-ra.

Megjegyzendő, hogy ez az eredmény nem csak Hoffman tételét foglalja magában, hanem Edmonds matroid metszet tétele is közvetlenül adódik.

Szubmoduláris áramok segítségével sikerült először levezetni az alábbi szeparációs tételt [11], amely a klasszikus konvex-konkáv elválasztási tételek diszkrét ellenpárjának tekinthető.

6.2 Tétel. *Legyen S alaphalmaz, $p^* : 2^S \to \mathbf{Z} \cup \{-\infty\}$ szupermoduláris függvény és $b^* : 2^S \to \mathbf{Z} \cup \{\infty\}$ szubmoduláris függvény, melyekre $p^*(\emptyset) = b^*(\emptyset) = 0$ és $p^* \leq b^*$. Ekkor van olyan egészértékű m moduláris függvény, amelyre $p^* \leq m \leq b^*$. Más szóval, az $\{x : x(A) \leq b^*(A)$ minden $A \subseteq S$-re, $x(S) = 0\}$ és az $\{x : x(A) \geq p^*(A)$ minden $A \subseteq S$-re, $x(S) = 0\}$ poliéderek metszete akkor és csak akkor tartalmaz egész pontot, ha $p^* \leq b^*$.*

Kiderült, hogy a Magyar Módszer alapgondolatai még ilyen általános körülmények közé is átvihetők. Evvel kapcsolatosan az egyik első dolgozat [2] egy olyan eljárást ír le, amely speciális esetként magában foglalja az eredeti Magyar Módszert, annak Ford-Fulkerson féle minimális költségű folyamokra való kiterjesztését és a fentebb említett súlyozott matroid metszet algoritmust is. Segítségével oldható meg algoritmikusan a digráf erősen összefüggővé tevésének előbb említett problémája, vagy az az érdekes kérdés, hogy egy irányítatlan gráfot mikor és miként lehet k-szor élösszefüggővé irányítani, sőt e feladatnak még az a költséges változata is kezelhető, amikor minden él lehetséges két irányának költsége különböző, és célunk a minimális költségű k-élösszefüggő irányítás megkeresése.

6.2 Még tovább

Bármennyire is általános ez az elmélet, az összes ilyen típusú eredmény azonban még ebbe sem fér bele. Például 1995-ben Jordán Tiborral bebizonyítottunk egy min-max formulát arra vonatkozólag, hogy egy irányított gráfot hány új él hozzáadásával lehet k-szor pontösszefüggővé tenni [12]. Ennek megfogalmazása érdekében legyen $D = (V, A)$ egyszerű irányított gráf és tegyük fel, hogy a $k < |V| - 1$. A V diszjunkt nemüres részhalmazaiból álló (X, Y) párra legyen $h(X, Y) := |V - (X \cup Y)|$. Azt mondjuk, hogy (X, Y) *egyirányú*, ha nem megy él X-ből Y-ba, azaz $\delta(X, Y) = 0$, ahol $\delta(X, Y)$ jelöli általában az X-ből Y-ba vezető élek számát jelöli.

A Menger tétel segítségével nem nehéz kimutatni, hogy egy $D^+ = (V, A^+)$ irányított gráf akkor és csak akkor k-összefüggő, ha $h(X, Y) \geq k$ fennáll minden (X, Y) egyirányú párra. Definiáljuk egy (X, Y) pár hiányát: $p_{hi}(X, Y) = (k - h(X, Y))^+$, ha (X, Y) egyirányú, és $:= 0$ különben. Világos, hogy $p_{hi}(X, Y) = (k - k\delta_A(X, Y) - h(X, Y))^+$ minden (X, Y) párra fennáll. Diszjunkt részhalmazokból álló párok egy $\mathcal{F}$ családját akkor nevezzük függetlennek, ha $\mathcal{F}$ bármely két $(X, Y), (X', Y')$ tagjára az $X \cap X'$ és $Y \cap Y'$ halmazok egyike legalább üres.

6.3 Tétel [12]. *A $D = (V, A)$ irányított gráf akkor és csak akkor tehető legfeljebb γ új él hozzáadásával k-összefüggővé, ha*

$$\sum (p_{hi}(X, Y) : (X, Y) \in \mathcal{F}) \leq \gamma \tag{37}$$

fennáll minden egyirányú párokból álló független $\mathcal{F}$ családra.

A [12] dolgozat valójában egy ennél jóval általánosabb, szupermoduláris függvényekre vonatkozó eredményt mutat be, amelynek számos egyéb következménye van, egyebek között Győri Ervin távolinak látszó, nehéz min-max tétele függőlegesen konvex vízszintes és függőleges szakaszok által határolt síkbeli tartomány minimális számú téglalappal történő fedéséről. Sajnos a bizonyítás nem konstruktív, és mind a mai napig nem tudjuk, hogy miként lehet algoritmikusan az összefüggőség növelés feladatát hatékonyan megoldani.

Van itt még egy újszerű jelenség. A Magyar Módszernek a minden korábban említett kiterjesztésénél lehetséges volt a súlyozott esetek kezelése is. Az összefüggőség növelésénél azonban a súlyozott változat általánosságban NP-teljes, és csak arra az esetre van megfelelő eredményünk, amikor a súlyfüggvény speciális alakú: egy pontsúlyozás indukálja.

A kérdés fennmarad: létezik-e ezen tételnek és a szubmoduláris áramok elméletének közös általánosítása. Egy másik, nagyobb lélegzetű kutatási irány annak feltárása, hogy a szubmoduláris áramok elméletét és a párosítások elméletét be lehet-e vonni valamiféle közös ernyő alá.

* * *

Összefoglalva megállapítható, hogy Egerváry tétele és az arra adott bizonyításának gondolata, bár egy szűk oldalon leírható, egy szerteágazó elméletnek lett kiindulópontja. Az egész kérdéskör azért tűnik különösképp vonzónak, mert gyakorlati kérdések, algoritmikus és tisztán elméleti vizsgálatok metszéspontjában áll, mert számos korábban reménytelennek tűnő probléma vált kezelhetővé általa, és mert még mindig nem kevés izgalmas nyitott probléma forrása.

Irodalom

1. W. H. Cunningham and A. B. Marsh A primal algorithm for optimum matching, *Mathematical Programming Study* 8 (1978) 50–72
2. W. H. Cunningham and A. Frank, A primal-dual algorithm for submodular flows, *Mathematics of Operations Research*, Vol. 10, No. 2 (1985), 251–261.
3. J. Edmonds, Paths, trees, and flowers, *Canadian Journal of Mathematics*, 17, (1965) 449–467.
4. J. Edmonds, Maximum matching and a polyhedron with 0–1 vertices, *Journal of Research of the National Bureau of Standards* (B) 69 (1965), 125–130.
5. J. Edmonds, Matroids and the greedy algorithm, *Math. Programming*, 1 (1971) 127–136
6. J. Edmonds, Matroid intersection, *Annals of Discrete Math.* 4, (1979) 39–49
7. J. Edmonds and R. Giles, A min-max relation for submodular functions on graphs, *Annals of Discrete Mathematics* 1, (1977), 185–204
8. Egerváry Jenő, Matrixok kombinatorius tulajdonságairól, *Matematikai és Fizikai Lapok*, No. 38 (1931), 16–28.
9. Egerváry Jenő, Kombinatorikus módszer a szállítási probléma megoldására, *A Matematikai Kutató Intézet Közleményei*, IV./1 (1958), 15–28.

10. A. Frank, A weighted matroid intersection algorithm, *J. Algorithms* 2 (1981) 328–336

11. A. Frank, An algorithm for submodular functions on graphs, *Annals of Discrete Mathematics* 16 (1982) 97–120.

12. A. Frank and T. Jordán, Minimal edge-coverings of pairs of sets, *J. Combinatorial Theory*, Vol. 65, No. 1 (1995, September) pp. 73–110.

13. A. J. Hoffman and J. B. Kruskal, Jr, Integral boundary points of convex polyhedra, Linear Inequalities and Related Systems, *Annals of Mathematics Study*, 38 (1956) 223–346, Princeton University Press.

14. A. Hoffman, Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, *Proceedings of the Symposia of Applied Mathematics*, 10 (1960) 113–127.

15. D. Kőnig, Graphok és matrixok, *Matematikai és Fizikai Lapok*, 38 (1931) 116-119.

16. H. W. Kuhn, The Hungarian Method for the assignment problem, *Naval Research Logistic Quarterly*, 2 (1955), 83–97.

17. A. Schrijver, Short proofs on the matching polytope, *J. Combinatorial Theory* (Ser B) 34 (1983) 104–108.

THE HUNGARIAN METHOD AND ITS EXTENSIONS

One of the earliest appearance of a special form of the linear programming duality theorem is E. Egerváry's classical result on the maximum weight of a perfect matching in a complete bipartite graph. Its elegant proof gave rise to an efficient algorithm called Hungarian Method. The basic principles of the procedure has been generalized in several directions such as the maximum weight matching problem in nonbipartite graphs, the weighted matroid intersection and the submodular flow problem. In the present work we overview these extensions of the Hungarian Method. As a new contribution, a short proof is given for the weight-splitting version of the weighted matroid intersection theorem.

EGERVÁRY RANGSZÁMCSÖKKENTŐ ALGORITMUSA ÉS ALKALMAZÁSAI [1]

GALÁNTAI AURÉL
Miskolci Egyetem, Matematikai Intézet

A rangszámcsökkentő eljárást Egerváry az ötvenes évek elejétől fejlesztette ki nemzetközileg is jelentős mátrixelméleti kutatásai keretében. A rangszámcsökkentést Egerváry többek között mátrixok rangjának megállapítására, faktorizációjára és egyenletmegoldó eljárások kifejlesztésére alkalmazta. A rangszámcsökkentő eljárás alkalmazása és vizsgálata új lendületet kapott az ABS módszerek megjelenésével [1], [3]. Dolgozatunkban az Egerváry-féle rangszámcsökkentő eljárás történetével és alkalmazásaival, valamint a rangszámcsökkentéshez kapcsolódó újabb eredményekkel foglalkozunk.

1 Bevezetés

Egerváry az ötvenes évek elején kezdett el intenzíven foglalkozni a mátrixszámítással. Akadémiai székfoglalója [7] lényegében már tartalmazza azt a szemléletet és technikát, amely egyértelműen Egerváryra és tanítványaira jellemző. Ezek közül kiemelhető a bázisfaktorizáció, a diadikus felbontás és a rangszámcsökkentés kapcsolatának felismerése és messzemenő kihasználása. Egerváry mátrixelméleti munkássága alapvetően konstruktív jellegű. Korai halála, amely a numerikus lineáris algebrában is bekövetkezett paradigmaváltás (számítógéporientált algoritmusok fejlesztésének) kezdetére esett, megakadályozta abban, hogy eredményeit ilyen irányban továbbfejlessze, ill. elterjessze. Ezért Egerváry munkásságának főként az elméleti része maradt meg a szakmai köztudatban. A nyolcvanas évek elején Abaffy, Broyden és Spedicato [1], [3] kifejlesztette az ABS-módszereket, amelyek, mint később kiderült, a rangszámcsökkentésen alapulnak. Az ABS-módszerek meglehetősen intenzív kutatása adott új lendületet a rangszámcsökkentő eljárásra vonatkozó további vizsgálatoknak.

A következőkben áttekintjük az Egerváry-féle rangszámcsökkentő eljárás történetét és alkalmazásait. Az

$$A' = A - Auv^T A / v^T Au \quad (A \in R^{m \times n},\ u \in R^n,\ v \in R^m)$$

transzformációt rangszámcsökkentésnek nevezzük. A rangszámcsökkentést, amelynek számos fontos alkalmazása van, többször is felfedezték. Különböző okok miatt a története ellentmondásos. A külföldön megjelent idevágó dolgozatok (Chu, Funderlic, Golub [6], Hubert, Meulman, Heiser [30]) hiányosak, illetve bizonyos értelemben torzítanak is.

[1] Beérkezett: 2002. március 4.

Mai ismereteink szerint a rangszámcsökkentés Wedderburn [38] 1934-ben megjelent könyvében szerepel először kvadratikus alakok Lagrange-féle redukciójával kapcsolatban. Wedderburn a következő eredményt igazolta.

1. Tétel (Wedderburn, 1934.) *Legyen $A \in R^{m\times n}$ tetszőleges $r \geq 1$ rangú mátrix. Ha $u \in R^m$, $v \in R^n$ olyan tetszőleges vektorok, amelyekre $v^T Au \neq 0$, akkor az $A' = A - Auv^T A/v^T Au$ mátrix rangja pontosan $r-1$.*

Wedderburn megjegyezte, hogy a rangszámcsökkentés ismételt alkalmazásával az r rangú A mátrix r darab 1-rangú mátrix összegére történő ún. diadikus felbontását kapjuk [38].

1. Definíció. *Egy $A \in R^{m\times n}$ $(A \neq 0)$ mátrix diadikus felbontásán az*

$$A = \sum_{i=1}^{k} b_i c_i^T \qquad (b_i \in R^m,\ c_i \in R^n)$$

előállítást értjük. A diadikus felbontás minimális, ha k a legkisebb ilyen tulajdonságú természetes szám.

A Wedderburn által javasolt algoritmus, amely megegyezik az Egerváry-féle rangszámcsökkentő eljárással, mai jelölésekkel a következő:

Rangszámcsökkentő eljárás. *Legyen $A \in R^{m\times n}$ egy tetszőleges $r \geq 1$ rangú mátrix és legyen*

$$H_1 = A \tag{1}$$

$$H_{i+1} = H_i - H_i x_i y_i^T H_i / y_i^T H_i x_i, \quad i = 1, \ldots, k\,, \tag{2}$$

ahol $x_i \in R^n$ és $y_i \in R^m$ olyan, amelyre $y_i^T H_i x_i \neq 0$ és $k \leq r$ □

Minthogy $H_{r+1} = 0$, az A mátrix diadikus felbontása:

$$A = \sum_{i=1}^{r} H_i x_i y_i^T H_i / y_i^T H_i x_i\,. \tag{3}$$

Wedderburn eredményét nem ismerve Egerváry 1953 és 1958 között egy sor dolgozatban fejlesztette ki és alkalmazta a róla elnevezett rangszámcsökkentő eljárást ([8]-[13], [15]-[18]). Egerváry a rangszámcsökkentő eljárás első speciális változatát mátrixok rangjának, illetve bázisfaktorizációjának meghatározására fejlesztette ki 1953-ban. Ismeretes, hogy Egerváry a mátrixok rangját a mátrixok minimális diadikus felbontásával definiálta [7]. Mátrixok minimális diadikus felbontása és az ún. bázisfaktorizációja között szoros kapcsolat van.

2. Definíció (Frazer, Duncan, Collar, 1938.) *Egy $A \in R^{m\times n}$ $(A \neq 0)$ mátrix bázisfaktorizációján az $A = FG^T$ szorzatfelbontást értjük, ahol $F \in R^{m\times r}$, $G \in R^{n\times r}$ és* $\operatorname{rank}(A) = \operatorname{rank}(F) = \operatorname{rank}(G)$.

A bázisfaktorizáció nem egyértelmű. Ha adott egy $A = FG^T$ bázisfaktorizáció, akkor minden más bázisfaktorizáció felírható

$$A = (FM)\left(M^{-1}G^T\right)$$

alakban, ahol $M \in R^{r \times r}$ tetszőleges.

Legyen $F = [f_1, \ldots, f_r]$ ($f_i \in R^m$) és $G = [g_1, \ldots, g_r]$ ($g_i \in R^n$). Ekkor az

$$A = FG^T = \sum_{i=1}^{r} f_i g_i^T$$

bázisfaktorizáció a mátrix egy minimális diadikus felbontását adja. Fordítva, ha ismert az

$$A = \sum_{i=1}^{r} b_i c_i^T$$

minimális diadikus felbontás, akkor ez az

$$A = BC^T = [b_1, \ldots, b_r]\,[c_1, \ldots, c_r]^T$$

bázisfaktorizációt adja.

Egerváry 1953-ban definiált rangszámcsökkentő eljárása a következő:

$$H_1 = A\,,$$

$$H_{l+1} = H_l - \frac{H_l e_{j_l} e_{i_l}^T H_l}{e_{i_l}^T H_l e_{j_l}}, \quad e_{i_l}^T H_l e_{j_l} \neq 0, \quad l = 1, \ldots, r\,.$$

Ez nyilván speciális esete az általános algoritmusnak. Közvetlenül is belátható, hogy

1. $\operatorname{rank}(H_{l+1}) = \operatorname{rank}(H_l) - 1$;
2. $H_{l+1} e_{j_l} = 0$ és $e_{i_l}^T H_{l+1} = 0$;
3. $H_l e_j = 0 \Rightarrow H_{l+1} e_j = 0$ és $e_i^T H_l = 0 \Rightarrow e_i^T H_{l+1} = 0$;
4. A egy minimális diadikus felbontása és a hozzátartozó bázisfaktorizációja

$$A = \sum_{l=1}^{r} \frac{H_l e_{j_l} e_{i_l}^T H_l}{e_{i_l}^T H_l e_{j_l}} = \underbrace{[H_1 e_{j_1}, \ldots, H_r e_{j_r}]}_{B} \underbrace{\begin{bmatrix} \frac{1}{e_{i_1}^T H_1 e_{j_1}} e_{i_1}^T H_1 \\ \vdots \\ \frac{1}{e_{i_r}^T H_k e_{j_r}} e_{i_r}^T H_r \end{bmatrix}}_{C^T}$$

Érdemes megjegyezni, hogy Egerváry a fenti eljárást Jacobi bilineáris alakokra vonatkozó transzformációjának mátrixelméleti analógonjaként kapta (E. Pascal [33]). További vizsgálatainak fő eszköze az 1956-ban publikált alábbi tétel volt ([12], [17], [18]).

2. Tétel (Egerváry, 1956.) *Legyen $A \in R^{m \times n}$ tetszőleges $r \geq 1$ rangú mátrix. Ekkor*

$$\operatorname{rank}\left(A - bc^T\right) = \operatorname{rank}(A) - 1 \qquad (b \in R^m,\ c \in R^n) \tag{4}$$

akkor és csak akkor áll fenn, ha $bc^T = Auv^T A/v^T Au$*, ahol* $u \in R^m$*,* $v \in R^n$ *tetszőleges a* $v^T Au \neq 0$ *feltételt kielégítő vektorok.*

A tételnek több bizonyítása is ismert [12], [17], [19], [32]. A 2. Tételt felhasználva definiálta Egerváry az általános rangszámcsökkentési algoritmust, szintén 1956-ban. Megmutatta, hogy az algoritmus meghatározza a rangot, előállít egy

$$A = \sum_{i=1}^{r} H_i x_i y_i^T H_i / y_i^T H_i x_i \tag{5}$$

minimális diadikus felbontást és a hozzátartozó

$$A = QD^{-1}P^T \tag{6}$$

bázisfaktorizációt, ahol

$$P = \left[H_1^T y_1, \ldots, H_r^T y_r\right], \quad Q = \left[H_1 x_1, \ldots, H_r^T x_r\right]$$

és

$$D = \operatorname{diag}\left(y_1^T H_1 x_1, \ldots, y_r^T H_r x_r\right) .$$

Egerváry kimutatta, hogy eljárásával LU-típusú és egyéb szorzatfelbontások állíthatók elő. Több, a rangszámcsökkentésen alapuló eljárást javasolt homogén mátrixegyenletek [8], [11], [12], [15], [17], illetve lineáris diofantikus egyenletrendszerek általános megoldására [13]. Kimutatta továbbá, hogy eljárása kapcsolatban áll a Purcell- és a Hestenes-Stiefel-féle konjugált irány módszerekkel is [17], [17]. Észrevette a rangszámcsökkentés és a rendszámnöveléses mátrixinvertálás kapcsolatát, amelynek segítségével eljárást adott módosított diszkrét peremértékproblémák megoldására [18]. Ez az eljárás tulajdonképpen egy fordított „rendszámcsökkentéses" mátrixinvertálásnak felel meg, amelyet Brezinski és társai 1996-ban újra felfedeztek [4].

A rangszámcsökkentéssel kapcsolatban kifejlesztett technikájával Egerváry számos más eredményt is elért [9], [16], [10], [14].

Megállapíthatjuk, hogy Egerváry a rangszámcsökkentési eljárást meghatározott célokra fejlesztette ki és széles körben szisztematikusan alkalmazta. Egerváry munkásságáról részletes áttekintés kapható Rózsa Pál [34], [35], [36] cikkeiből, aki egyébként Egerváry utolsó poszthumusz cikkeit is sajtó alá rendezte.

Egerváry rangszámcsökkentéssel kapcsolatos munkássága 1986-ig nem sok nemzetközi visszhangot kapott. Householder 1964-ben megjelent klasszikus és befolyásos *The Theory of Matrices in Numerical Analysis* című könyvében ugyan idézi Egerváry több idevágó munkáját (tételesen a [9], [10], [11], [16], [17] dolgozatokat), de a 2. Tételt Wedderburn tételének nyilvánítja ([29], Exercise 34, p. 33). A rangszámcsökkentő eljárást azonban Egervárynak tulajdonítja a következő formában:

„Apply the obvious identity

$$(u_1, u_2, \ldots, u_n)(v_1, v_2, \ldots, v_n)^H = u_1 v_1^H + \ldots + u_n v_n^H$$

along with Wedderburn's theorem (Exercise 1.34) to obtain Egervary's "rank-reducing" transformations, and thereby derive the methods of triangularization and of orthogonal triangularization."

Householder könyvének megjelenése után több szerző Egerváry tételét Householdernek ([32], Thm. 2.6b, p. 200), vagy Wedderburnnak és Householdernek tulajdonította ([5]).

1984-ben Abaffy, Broyden és Spedicato kvázi-Newton technikával egy konjugált irányokon alapuló módszerosztályt fejlesztettek ki $Ax = b$ alakú lineáris egyenletrendszerek megoldására [1].

Az ABS módszerosztály:

$y_1 \in R^m$, $H_1 \in R^{m\times m}$ $(\det(H_1) \neq 0)$

for $k = 1 : m$

$p_k = H_k^T z_k \quad (z_k \in R^m,\ p_k^T A^T v_k \neq 0)$

$\alpha_k = v_k^T r_k / v_k^T A p_k \quad (r_k = Ay_k - b)$

$y_{k+1} = y_k - \alpha_k p_k$

$H_{k+1} = H_k - H_k A^T v_k w_k^T H_k / w_k^T H_k A^T v_k \quad (w_k \in R^m,\ w_k^T H_k A^T v_k \neq 0)$

end

$Ay_{m+1} = b$

Az eljárás nyilvánvalóan tartalmazza az Egerváry-féle rangszámcsökkentési eljárást. E tény felismerésére 1986-ban került sor [2], és azóta az ún. ABS módszerekkel kapcsolatos irodalomban Egerváry idevágó eredményei szerepelnek (lásd pl. [3]). Az ABS módszerek kiterjesztésre kerültek nemlineáris egyenletrendszerek megoldására (lásd [3]), valamint alkalmazásra kerültek az optimalizálásban is (lásd [39]).

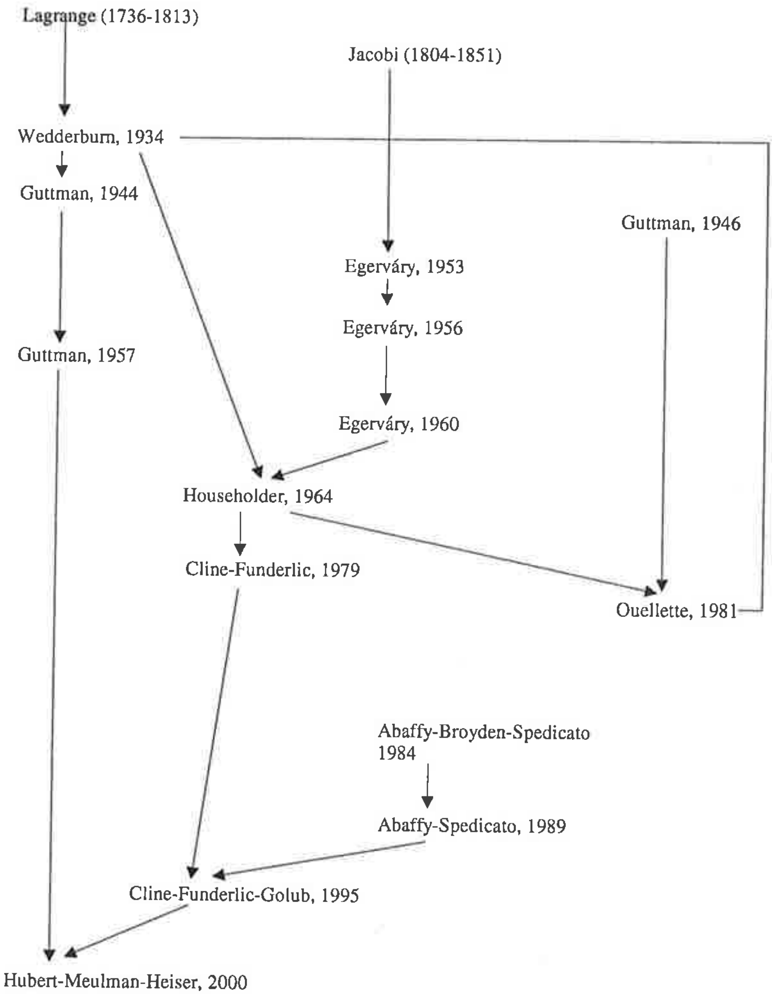
1995-ben Chu, Funderlic és Golub a rangszámcsökkentésről egy jelentős cikket publikált a SIAM Review-ban [6]. Ebben a dolgozatban Egerváry számos eredményét idézik, de ezt igen sajátos formában teszik. Például elismerik, hogy a 2. Tétel Egerváry eredménye, de meg akarnak arról győzni, hogy Householder ezt már Egerváry előtt felfedezte, noha ennek írásos nyoma nincs.

Ezt a dolgozatot követi Hubert, Meulman és Heiser 2000-ben megjelent [30] dolgozata, amelyben kimutatják, hogy a 2. Tétel blokkváltozatát Guttman 1957-ben igazolta. Tehát Ő a tétel első felfedezője. Ezt a következtetést azért vonták le, mert külföldön csak az 1960-ban megjelent Egerváry cikket ismerik.

Tény, hogy Guttman a 2. Tétel blokkváltozatát 1957-ben publikálta. Ezt megelőzően 1944-ben az 1. Tétel blokkváltozatát igazolta. Guttman a rangszámcsökkentés blokkváltozatát statisztikában alkalmazta ([30]).

Ouellette kimutatta, hogy Guttman egy 1946-os, a Schur-komplementálásnál alapvető eredményéből azonnal következik a Wedderburn és a Householder, azaz Egerváry-tétele (lásd [32]). Érdekesség, hogy Guttman ezt nem vette észre.

A most vázolt bonyolult összefüggéseket az alábbi ábrával szemléltethetjük, amely az eszmék haladási irányát mutatja.



1. ábra

Mindent összevetve megállapíthatjuk a következőket:

1. Wedderburn volt az első, aki az 1. Tételt és a rangszámcsökkentést felfedezte.

2. Egerváry a rangszámcsökkentési eljárást szisztematikusan fejlesztette ki és széles körben alkalmazta. Egerváry valójában egy egész elméletet épített ki, amelynek kulcstételét, a 2. Tételt, Ő bizonyította elsőként.

3. Egerváry után közvetlenül Guttman volt az első, aki a 2. Tételt újra felfedezte és egy más körben (statisztikában) alkalmazta.

2 Új eredmények

A következőkben bemutatok néhány eredményt, amelyet a rangszámcsökkentő eljárással kapcsolatban értem el.

Egy A^- mátrixot az A (1)-inverzének, vagy g-inverzének nevezünk, ha fennáll, hogy $AA^-A = A$. A G mátrixot az A reflexív inverzének nevezzük, ha $AGA = A$ és $GAG = G$.

Egy nemszinguláris A mátrixot erősen nemszingulárisnak nevezünk, ha létezik LU-felbontása. Egy nemszinguláris A mátrixnak akkor és csak akkor van LU-felbontása, ha minden főminor mátrixa nemszinguláris.

A rangszámcsökkentési eljárást zérusosztómentesnek nevezzük, ha

$$y_i^T H_i x_i \neq 0 \quad (i = 1, \ldots, k) \, .$$

Ez a pontos feltétele annak, hogy k egymás utáni rangszámcsökkentést végrehajthassunk. Legyen

$$X = [x_1, \ldots, x_k] \, , \quad Y = [y_1, \ldots, y_k]$$

és

$$X^{(i)} = [x_1, \ldots, x_i] \, , \quad Y^{(i)} = [y_1, \ldots, y_i] \; .$$

Ekkor a következő eredmény igaz.

3. Tétel. *A rangszámcsökkentő eljárás akkor és csak akkor zérusosztómentes, ha az* $Y^T AX = \left[y_i^T A x_j\right]_{i,j=1}^k$ *mátrix erősen nemszinguláris. Ebben az esetben a rangszámcsökkentés a következő kanonikus alakba írható:*

$$A_{i+1} = A - AX^{(i)} \left(Y^{(i)T} A X^{(i)}\right)^{-1} Y^{(i)T} A \quad (i = 1, \ldots, k) \; . \tag{7}$$

A tétel szükséges részét elsőként Abaffy, Broyden és Spedicato igazolták az ABS módszerekkel kapcsolatban [1], [3]. A kanonikus alak Egervárynál is megtalálható [17], [18], aki azt vizsgálta, hogy k egymás utáni rangredukció helyettesíthető-e egyetlen blokk rangredukcióval.

A továbbiakban tegyük fel, hogy $k = r$. Ekkor a rangszámcsökkenő eljárás az

$$A = QD^{-1}P^T \tag{8}$$

bázisfaktorizációt szolgáltatja. Felmerül a kérdés, hogy adott A, X és Y paraméterek esetén mi a kapott P, a Q és a D? Ezt néhány speciális esetben már Egerváry [17], Abaffy, Broyden, Spedicato [1], [3], valamint Chu, Funderlic és Golub [6 is megválaszolta.

Ha egy B kvadratikus mátrixnak van LDU-felbontása, akkor ezt az egyértelmű felbontást jelölje $B = L_B D_B U_B$, ahol a B index a mátrixra utal.

4. Tétel. *Legyen* $Y^T AX$ *erősen nemszinguláris. Ekkor a (6) bázisfaktorizáció tényezői*

$$P = A^T Y L_{Y^T AX}^{-T}, \quad Q = AXU_{Y^T AX}^{-1}, \quad D = D_{Y^T AX} \, , \tag{9}$$

és az A kifejezhető az

$$A = \left(AXU_{Y^TAX}^{-1}\right) D_{Y^TAX}^{-1} \left(L_{Y^TAX}^{-1} Y^T A\right) \tag{10}$$

alakban.

Az eredmény és igazolása különféle formákban a [20], [21], [23], [24] munkákban található meg. A tételnek számos fontos következménye van:

1. A rangszámcsökkentési eljárással minden bázisfaktorizáció előállítható ([6] sejtése).

2. Ha A, X és Y négyzetes nemszinguláris, akkor

$$A = \underbrace{\left(Y^{-T} L_{Y^TAX} D_{Y^TAX}\right)}_{Q} D_{Y^TAX}^{-1} \underbrace{\left(D_{Y^TAX} U_{Y^TAX} X^{-1}\right)}_{P^T} . \tag{11}$$

3. Egyértelmű jellemzés adható arra, hogy P, vagy Q mikor lesz háromszög alakú, illetve B ortogonális.

4. A Lánczos-féle tridiagonális formára történő hasonlósági transzformáció mátrixa és annak inverze egyszerre előállítható a rangszámcsökkentéssel.

3. Definíció. *Legyen $A \in R^{m\times n}$, $V \in R^{m\times r}$ és $P \in R^{n\times r}$. A (P, V) mátrixpárt A-konjugáltnak nevezzük, ha $L = V^T AP$ nemszinguláris alsó háromszögmátrix. A (P, V) pár A-bikonjugált, ha $D = V^T AP$ nemszinguláris diagonális.*

A rangszámcsökkentő eljárásnak a következő konjugálási tulajdonságai vannak.

1. Állítás. *Ha $X = B^T V$ és $Y = C^T W$, akkor a (P, V) pár B-konjugált, a (Q, W) pár pedig C-konjugált.*

2. Állítás. *Tetszőleges B-konjugált (P, V) mátrixpár esetén létezik olyan Y ortogonális mátrix, hogy a rangszámcsökkentő eljárás az $A = I$, $X = B^T V$ és Y választással éppen ezt a P mátrixot adja.*

3. Állítás. *Legyen A^- az A mátrix tetszőleges g-inverze. Ekkor a (Q, P) pár A^--bikonjugált.*

Megmutatható, hogy egy sor ismert konjugálási eljárás a rangszámcsökkentés speciális esete, vagy éppen azzal ekvivalens. A 3. Állítás miatt a rangszámcsökkentés speciális esetként tartalmazza a kétoldali Gram-Schmidt eljárást, amellyel bikonjugált párokat állíthatunk elő.

Megmutatható, hogy a rangszámcsökkentéssel kapott (Q, P) pár stabil, ha az $Y^T AX$ mátrix LDU-felbontása stabil [22].

Végül megemlítem, hogy a rangszámcsökkentő eljárás és a rendszámnöveléses mátrix invertálás között több új és érdekes összefüggést sikerült találni.

Ezek egyik következménye a kanonikus alakra adott három újabb multiplikatív reprezentáció, amelyekkel a rangszámcsökkentési eljárást különféle szempontokból jellemezhetjük. Példaként megemlítjük a következőt:

$$A_{i+1} = AX \left(\left(Y^T AX \right)^{-1} - \begin{bmatrix} \left(Y^{(i)T} AX^{(i)} \right)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \right) Y^T A \,.$$

A formulából látható, hogy az eljárás során az $Y^T AX$ mátrix inverzét közelítjük a főminor mátrixok inverzeivel.

3 Összefoglalás

Az eddig ismert tények alapján jogos Egerváry-féle rangszámcsökkentő eljárásról beszélni. Az eljárás egy széles körben használható, hatékony eszköz, amelynek rendkívül érdekes tulajdonságai vannak. Remélni lehet, hogy az eljárás végre a szakirodalomban is a megfelelő helyre kerül. Mindenesetre Egerváry munkásságának növekvő elismerését az is jelzi, hogy a Bergamói Egyetem 1998-ban Egerváry díjat adott Zhang Liwei kínai matematikusnak idevágó munkásságáért.

Irodalom

1. J. Abaffy, C. G. Broyden, E. Spedicato, A class of direct methods for linear equations, *Numerische Mathematik*, 45, 361-376, 1984.
2. Abaffy J., Galántai A.: Conjugate direction methods for linear and nonlinear systems of algebraic equations, *Colloquia Mathematica Soc. János Bolyai, 50. Numerical Methods, Miskolc (Hungary) 1986.* (ed. P. Rózsa) North-Holland, Amsterdam, 1987, pp. 481-502
3. J. Abaffy, E. Spedicato, *ABS-projection Algorithms: Mathematical Techniques for Linear and Nonlinear Algebraic Equations*, Ellis Horwood, Chichester, 1989.
4. Brezinski, C., Morandi Cecchi, M., Redivo-Zaglia, M.: The reverse bordering method, SIAM J. Matrix. Anal. Appl. 15, 1994, pp. 922-937
5. R.E. Cline, R.E. Funderlic, The rank of a difference of matrices and associated generalized inverses, *Lin. Alg. Appl.*, 24, 185-215, 1979.
6. M.T. Chu, R.E. Funderlic and G.H. Golub, A rank-one reduction formula and its applications to matrix factorizations, *SIAM Review*, 37, 512-530, 1995.
7. Egerváry J.: Mátrix-függvények kanonikus előállításáról és annak néhány alkalmazásáról, *Az MTA III. Matematikai és Fizikai Osztályának Közleményei* 3, 1953, 417-458
8. J. Egerváry, Mátrixok diadikus előállításán alapuló módszer bilineáris alakok transzformációjára és lineáris egyenletrendszerek megoldására (A method based on the dyadic representation of matrices for the transformation of bilinear forms and for solving systems of linear equations, in Hungarian), *MTA Alkalmazott Matematikai Intézetének Közleményei*, 2, 11–32, 1953.

9. J. Egerváry, On a property of the projector matrices and its application to the canonical representation of matrix functions, *Acta Sci. Math.*, 15, 1-6, 1953.

10. J. Egerváry, On a lemma of Stieltjes on matrices, *Acta Sci. Math.* 15, 99-103, 1954.

11. J. Egerváry, Über die Faktorisation von Matrizen und ihre Anwendung auf die Lösung von linearen Gleichungssystemen, *Zeitschrift für angewandte Mathematik und Mechanik* 35, 111-118, 1955.

12. J. Egerváry, Régi és új módszerek lineáris egyenletrendszerek megoldására (Old and new methods for solving linear equations, in Hungarian), *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* (Publications of the Mathematical Institute of the Hungarian Academy of Sciences), 1, 109–123, 1956.

13. J. Egerváry, Auflösung eines homogenes linearen diophantischen Gleichungssystems mit Hilfe von Projektormatrizen, *Publicationes Mathematicae*, 4, 481–483, 1956.

14. J. Egerváry, Az inverz mátrix általánosítása (On a generalized inverse for matrices, in Hungarian), *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* (Publications of the Mathematical Institute of the Hungarian Academy of Sciences), 1, 315–324, 1956.

15. J. Egerváry, Über eine Verallgemeinerung der Purcelschen Methode zur Auflösung linearer Gleichungssyteme, *Österreichisches Ingeniur-Archiv*, 11, 249–251, 1957.

16. J. Egerváry, Über einigekonstruktive Methode zur Reduktion einer Matrix auf die Jordansche Normalform, *Acta Math. Acad. Hung.*, 10, 31–54, 1959.

17. E. Egerváry, On rank-diminishing operations and their applications to the solution of linear equations, *ZAMP*, 11, 376–386, 1960.

18. J. Egerváry, Über eine Methode zur numerischen Lösung der Poissonschen Differenzengleichung für beliebige Gebiete, *Acta Mathematica Academiae Scientiarium Hungaricae*, 11, 341–361, 1960.

19. L. Elsner, P. Rózsa, On eigenvectors and adjoints of modified matrices, *Linear and Multilinear Algebra*, 10, 235–247, 1981.

20. A. Galántai, Generalized implicit LU algorithms in the class of ABS methods for linear and nonlinear systems of algebraic equations, *Quaderno DMSIA* 93/5, University of Bergamo, Bergamo, 1993.

21. A. Galántai, The global convergence of the ABS methods for a class of nonlinear problems, *Optimization Methods and Software*, 4, 283–295, 1995.

22. Galántai A.: Perturbation theory for full rank factorizations, Quaderni DMSIA, 1999/40, University of Bergamo, Bergamo, 1999.

23. Galántai A.: Rank reduction and conjugation, Mathematical Notes, Miskolc, 1 (2000), pp. 11–33.

24. Galántai A.: Rank reduction, factorization and conjugation, *Linear and Multilinear Algebra*, Vol. 49, 2001, pp. 195–207

25. L. Guttman, General theory and methods for matric factoring, *Psychometrika*, 22, 1–16, 1944.

26. L. Guttman, Enlargement methods for computing the inverse matrix, *Ann. Math. Statist.* 17, 336–343, 1946.

27. L. Guttman, A necessary and sufficient formula for matrix factoring, *Psychometrika*, 22, 79–91, 1957.
28. Hegedűs Cs., *Mátrixelmélet*, MTA Matematikai Kutató Intézete, 1974
29. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.
30. L. Hubert, J. Meulman, W. Heiser, Two purposes for matrix factorization: A historical appraisal, *SIAM Review* 42, 68–82, 2000.
31. Lovass-Nagy V., *Matrixszámítás*, Műszaki matematikai gyakorlatok, C. IV., Tankönykiadó, Budapest, 1956
32. D. V. Ouellette, Schur complements and statistics, *Lin. Alg. Appl.*, 36, 187–295, 1981.
33. E. Pascal, *Repertorium der höheren Mathematik I: Analysis*, Teubner, Leipzig, 1910.
34. Rózsa P., Egerváry Jenő munkásságáról, *Matematikai Lapok*, 10, 1959, 195–225
35. Rózsa P., Egerváry Jenő, 1891-1958, *MTA III. Matematikai és Fizikai Osztályának Közleményei*, 10, 1960, 1–7
36. P. Rózsa, *Lineáris algebra és alkalmazásai* (Linear algebra and its applications, in Hungarian), Müszaki Könyvkiadó, Budapest, 1974.
37. P. Rózsa, Jenô Egerváry(1891-1958), a great personality of the Hungarian mathematical school, *Periodica Polytechnica, Electrical Engineering*, 28, 287–298, 1984.
38. J. H. M. Wedderburn, *Lectures on Matrices*, Amer. Math. Soc. Colloquium Publications, Vol. XVII., 1934.
39. Zhang Liwei, Xia Zunquan, Feng Enmin, *Introduction to ABS Methods in Optimization*, Chinese, Dalian University of Technology Press, China, 1999.

THE RANK-REDUCTION ALGORITHM OF EGERVÁRY AND ITS APPLICATION

The history and applications of the rank-reduction algorithm of Egerváry is examined. New results concerning the rank reduction are also mentioned.

EGY HEURISZTIKUS MÓDSZER A KVADRATIKUS HOZZÁRENDELÉSI PROBLÉMA KÖZELÍTŐ MEGOLDÁSÁHOZ [1]

BORGULYA ISTVÁN
PTE Közgazdaságtudományi Kar

Tanulmányunkban egy evolúciós algoritmust ismertetünk a kvadratikus hozzárendelési probléma megoldására. Az algoritmus működése három, egymás utáni fázisra bontható: egy előkészítő, valamint két különböző helyi kereső fázisra. Az első fázisban a kezdő populáció minőségét javítja. A második fázisban egy helyi kereső eljárással felváltva javít minden megoldást (egyedet), és közben periodikusan újabb megoldásokat generál a meglévők környezetében. A harmadik fázisban a helyi kereső eljárással folytatja a megoldások javítását, és közben a legrosszabb megoldásokat periodikusan újakra cseréli. Az algoritmust a QAPLIB könyvtár jól ismert tesztfeladataival ellenőriztük. Az algoritmus a feladatok 98%-nál megtalálta a legjobb ismert megoldásokat, vagy a legjobb ismert megoldás 1%-os környezetébe került.

1 Bevezetés

A kvadratikus hozzárendelési probléma (QAP), egy klasszikus kombinatorikai optimalizációs probléma. NP-teljes feladat, és egzakt megoldása csak kisméretű feladatok eseten lehetséges, ill. gazdaságos.

A probléma a következő: adott n objektum, melyet n különböző helyre kell elhelyezni. Tudjuk, hogy f_{ij} érték áramlik az i és j objektum közt, és a k és l hely közti távolság pedig d_{kl}. Az objektumok olyan elrendezését keressük, amelynél az érték $\times$ távolság szorzatok összege minimális. Matematikailag a következőképpen formalizálhatjuk: legyen $F = \{f_{ij}\}$ és $D = \{d_{kl}\}$, két $n \times n$-es mátrix, és keresünk egy olyan π^* permutációt, amelyre minimális

$$Z(\pi) = \min_{\pi \in \Pi(n)} \sum_{i=1}^{n} \sum_{i=j}^{n} f_{ij} d_{\pi_i \pi_j}$$

ahol $\Pi(n)$ n elem permutációinak halmaza és π_i, π_j a $\pi \in \Pi(n)$ permutáció i-edik és j-edik pozíciójának értékét jelöli.

A QAP alkalmazásával számos mérnöki, ill. tudományos feladatban találkozunk. Néhány gyakorlati probléma ezek közül, pl. egyetemtervezési probléma; irodák, szobák elrendezése irodaházban, ill. korházban; VLSI lapocskák elhelyezése az alaplapon (részletesebben: [5], [8], [15]).

[1] Beérkezett: 2001. december 8.

A QAP egzakt megoldására a dinamikus programozás, különböző metszési módszerek és a korlátozás és szétválasztás módszere alkalmazhatók. $n =$ 20-nál nagyobb méretű feladatok esetén az egzakt megoldást azonban nem tudjuk meghatározni elfogadható időn belül [5]. Mivel a legtöbb gyakorlati feladat nagyméretű probléma, így számos heurisztikát alkalmaznak, melyek elfogadható időn belül jó megoldást találnak.

Sokféle heurisztika létezik a QAP megoldására. A legismertebb csoportjaik a következők: konstrukciós módszerek, javító módszerek, szimulált hűtés (SA), tabukeresés (TS), evolúciós algoritmus (EA) változatok (pl. genetikus algoritmus (GA), genetikus programozás (GP), evolúciós stratégia (ES)), hangya kolóniák (részletesebben, pl. [8], [4], [15], [10], [12], [11]). Egyes heurisztikus módszerek gyorsasága és pontossága tovább javítható, pl. különböző optimalizáló, vagy párhozamos számításokkal. Így Ahuja et al. [1] a GA teljesítményét egyes műveletek optimalizálásával, több populáció párhuzamos alkalmazásával javítja; Talbi et al. [17] párhuzamos TS algoritmust fejleszt, amely nagyméretű feladatok esetén is nagyon jó közelítést nyújt az optimumra; Bölte et al. [3] GP segítségével állít elő jobb "hő ütemező" függvényt az SA módszerhez. E mellett számos más elven működő heurisztikus módszer alkalmazható. A módszerek közt Boltzmann gép, vagy klaszterező módszerek szintén találhatók.

A heurisztikus módszerek hatékonyságát több összehasonlító tanulmány vizsgálja (pl. [13], [16], [11]). Általában megállapítható, hogy a módszerek a legjobb ismert megoldást 1%-on belüli pontossággal képesek közelíti; egyes feladatokat más-más módszerek oldanak meg sikeresebben és a módszerek egy része csak $n < 100$ méretű problémát tud eredményesen kezelni. A legeredményesebb heurisztikák a TS, SA és a "helyi javító" módszer ([1], [7]).

E heurisztikus módszerek körét egy újabb algoritmussal kívánjuk bővíteni. Célunk, hogy az algoritmus

a) kis, közepes, vagy nagyméretű problémáknál egyaránt alkalmazható legyen;

b) PC-n a szokásos feladatok esetén, elfogadható időben nyújtson jó közelítést (a legjobb ismert megoldást 1%-on belüli pontossággal közelítse); és

c) egyszerűen, néhány paraméter beállításával legyen kezelhető.

Az algoritmust a QAPLIB könyvtár [6] tesztfeladataival ellenőriztük. A feladatok 98%-át sikeresen oldotta meg: vagy megtalálta a legjobb ismert megoldást, vagy a legjobb ismert megoldást 1%-on belüli pontossággal közelítette meg. A feladatok további 2%-ban 1-1.5% pontossággal közelítette a legjobb ismert megoldást. Összehasonlításra a greedy genetikus algoritmust [1] választottuk. Megállapítható, hogy új módszerünk PC-n, általában rövidebb idő alatt, közel azonos pontosságú közelítést talált, mint a greedy genetikus algoritmus.

A továbbiakban nézzük először az új algoritmus leírását, majd a tesztfeladatok megoldását, melyek az alkalmazás lehetőségét szemléltetik.

2 Az algoritmus alapgondolata

Az algoritmus, nevezzük EF_QAP-nek (Evolutionary Framework for QAP), egy korábbi klaszter-alapú optimalizációs algoritmus [2] módosításával jött létre. Egy olyan hibrid EA, amely az evolúciós ciklust egy helyi kereső eljárással bővíti. Új algoritmusunk azonban a konvergencia gyorsítása és minél jobb minőségű eredmények elérése érdekében, a korábbi hibrid EA-tól eltérően nem egy, hanem 3 fázisból áll. A három fázis mindegyike egy-egy EA, amelyeket egymás után kell végrehajtani. Az első fázis egy "előkészítő" fázis, amely a kezdő populáció minőségét kívánja javítani. A második fázis olyan hibrid EA, amely keresés közben folyamatosan növeli a populáció méretét. E méret növeléssel gyorsítani kívánja a konvergenciát. Végül a harmadik fázis folytatja a megoldások javítását. Eltérően a második fázistól itt állandó a populáció mérete, viszont a legrosszabb megoldásokat periodikusan újakra cseréli az algoritmus.

Az algoritmus tehát három EA-ból áll. Tekintsük a permutációkat egyedeknek (vagy megoldásoknak). Minden EA-ban generációnként (iterációnként) egy szülőből másolással egy utódot hozunk létre. Mutációt (néhány párcsere a permutációban + helyi kereső eljárás) alkalmazunk, majd az utódot a leghasonlóbb egyeddel (vagy magával a szülővel) párba állítva szelektálunk az egyed és az utód közt (crowding selection). Jósági függvénynek (fitness function) a Z függvényt választhatjuk. Az EF_QAP működését meghatározó három EA feladata részletesebben a következő:

1. Az első fázisban az algoritmus először véletlenszerűen generálja a kezdő populáció egyedeit. Utána véletlenszerűen generált utódokkal próbálja javítani őket. Egy utód mindig csak a leghasonlóbb korábbi megoldást javíthatja.

2. A második fázisban a megoldások minőségét mutációval, azaz egy összetett helyi kereső eljárással javítja. Az eljárás a megoldások javításához a megoldások környezetében olyan utódokat választ, melyek a korábbi megoldásból (permutációból) 1-2 párcserével és egy szomszédságban kereső eljárással előállíthatók (A mutáció tehát az *1-2 párcsere + szomszédságban kereső eljárás* transzformációval állítja elő a szülőből az utódot). A javításra kerülő megoldás kiválasztásában prioritást élvez a legjobb, legkisebb Z függvényértékű megoldás. 0.5 valószínűséggel a legjobb, $0.5/t$ valószínűséggel (ahol t a populáció mérete) pedig tetszőleges megoldást választ az algoritmus.

   A helyi kereső eljárás tehát a véletlenszerűen választott egy-két párcsere után egy *szekvenciális, szomszédságban kereső eljárást* alkalmaz, amely végigvizsgálja az új elem szomszédos permutációit. E szekvenciális, szomszédságban kereső eljárás (sequential neighbourhood search, [8]) a permutáció pozícióin a következő sorrendben alkalmaz párcseréket:

$$(1,2),(1,3),\ldots,(1,n),(2,3),\ldots,(n-1,n),(1,2),\ldots$$

Az algoritmus a szomszédságban kereső eljárást mindaddig alkalmazza, amíg lehetséges az eredmény javítása. Ha a keresés végére ért, és keresés közben valamely párcsere javított a helyi optimum közelítés értékén, újból kezdi a szomszédságban kereső eljárást. Emellett, hogy a végrehajtási időt gyorsítsa, a függvényérték számítást Taillard [16] módszerével szintén egyszerűsíti. Ehhez az i és j objektumok cseréjénél először csak a függvényérték változását számolja ki

$$\begin{aligned}\Delta Z = {} & f_{ii}(d_{\pi_j\pi_j} - d_{\pi_i\pi_i}) + f_{ij}(d_{\pi_j\pi_i} - d_{\pi_i\pi_j}) + \\ & + f_{ji}(d_{\pi_i\pi_i} - d_{\pi_j\pi_j}) + f_{jj}(d_{\pi_i\pi_i} - d_{\pi_j\pi_j}) + \\ & + \sum_{\substack{k=1\ldots n \\ k\neq i,j}} \Big(f_{ki}(d_{\pi_k\pi_j} - d_{\pi_k\pi_i}) + f_{kj}(d_{\pi_k\pi_i} - d_{\pi_k\pi_j}) + \\ & + f_{jk}(d_{\pi_j\pi_k} - d_{\pi_i\pi_k}) + f_{ik}(d_{\pi_j\pi_k} - d_{\pi_j\pi_k}) \Big)\end{aligned}$$

Ha a $\Delta Z < 0$, akkor $-\Delta Z$-vel javul a függvény értéke. Az algoritmus a párcserét ténylegesen csak abban az esetben hajtja végre, ha javul a függvény értéke.

Általában a korábbi megoldásból az 1-2 párcsere + szomszédságban kereső eljárás transzformációval egyre jobb megoldást kapunk. Egyes feladatoknál a továbblépéshez azonban az 1-2 párcsere nem mindig elegendő, az algoritmus valamely lokális optimumnál "elakad". A megoldások javítása érdekében ezért újabb megoldásokat generál az algoritmus. Egy újabb megoldást valamely már létező megoldásból véletlenszerűen, maximum $n/8$ számú párcsere + szomszédságban kereső eljárás transzformációval állít elő. Így az új megoldások a meglévők "közelében" keletkeznek, és mint újabb változatok, növelik annak az esélyét, hogy az algoritmus közelíteni tudjon a globális optimumhoz. Az algoritmus ezért a második fázisban periodikusan, mindaddig újabb megoldásokat generál a meglévők környezetében, amíg a megoldások (egyedek) száma egy maximális értéket el nem ér.

3. A harmadik fázisban az algoritmus folytatja a második fázis helyi kereső eljárásával a megoldások javítását. Szintén generál újabb megoldásokat, de lecseréli velük a legrosszabb megoldásokat. Legrosszabbnak a legnagyobb célfüggvény értékű megoldásokat tekinti az algoritmus, és periódusonként adott százalékukat cseréli le.

3 Az algoritmus

Jelölések

Vezessük be a következő jelöléseket:

- Jelöljük az egyes EA-kat EA1, EA2 és EA3-al.

- Ugyanazon populációt alkalmazza mindhárom EA. Legyenek a $p_1, \ldots, p_t$ permutációk az egyedek (megoldások). Jelölje az it-edik generáció populációját $P(it)$, és az i-edik egyedet p_i. A jósági függvény legyen azonos a Z függvénnyel.
- Jelölje `Csere(q)` a párcsere eljárás-t. Az eljárás a q permutációban véletlenszerűen kicseréli két pozíció értékét.
- Jelölje `Nhs(q)` a szekvenciális, szomszédságban kereső eljárást, amely a q utód szomszédos permutációit mindaddig ismételten végigvizsgálja, amíg a megoldás értékén lehetséges javítani. Ha tud javítani, minden alkalommal végrehajtja a párcserét a q megoldásban.
- Jelölje `Újegyed` az új megoldást generáló eljárást. Az eljárás úgy generál egy új q megoldást, hogy véletlenszerűen választ egy p_i-t ($p_i \neq 0$, $i \in \{1, 2, \ldots, t\}$), átmásolja q-ba, alkalmazza rá maximum $n/8$-szor a `Csere(q)` eljárást, majd az `Nhs(q)` eljárást. q a $t+1$-edik egyed lesz.
- Jelölje `SortDel` a rendező eljárást, amely a prototípusokat a Z függvényértékek szerint növekvő sorba rendezi, törli a legrosszabb megoldások ddp százalékát, és helyettük új megoldásokat generál.
- Az x és z permutáció hasonlóságának mértékét $H(x, z) = 1/(1+d(x, z))$ határozza meg, ahol $d(x, z)$ a két permutáció Hamming-távolsága.
- Jelölje a véletlenszám generátort `Rnd` (egyenletes eloszlás $(0, 1)$-en).

Paraméterek

7 paraméter befolyásolja az algoritmus futását: $tmax$, t, itt, kn, ddp, m és $timeend$. Szerepük a következő:

- $tmax$ - a populáció maximális mérete.
- t - a populáció maximális mérete az első fázisban.
- itt - az 1. fázis paramétere. Ha az it iteráció szám nagyobb, mint itt, megkezdődik a 2. fázis.
- kn - az ellenőrzések időpontját meghatározó paraméter. A legjobb megoldások megkeresése, vagy a legrosszabb megoldások törlése, újakra cserélése csak minden kn-dik iterációban történik.
- m - a 2. fázis paramétere. A 2. fázisban az algoritmus minden m-dik iterációban növeli a populáció méretét.
- ddp - a 3. fázis paramétere. A legrosszabb megoldások ddp százalékát törli a 3. fázisban.
- $timeend$ - a megállási feltétel paramétere. Az eljárás befejeződik, ha a futási idő (CPU idő másodpercben) nagyobb, mint $timeend$.

Az algoritmus lépései

```
Procedure EF_QAP(tmax, t, itt, kn, ddp, m, timeend, opt, optp)
** EA1 **
it := 0, glob := 1.                                        /* Kezdőértékek beállítása
Legyen p_i ∈ Π(n) (i = 1,...,t), P(it) ← {p_1,...,p_t}     /*Kezdő populáció
Z(p_1),...,Z(p_t) kiszámítása
Repeat
  i := [t * Rnd] + 1, q := p_i                             /* Utód generálás
  For j := 1 to [n/2] do Csere(q) od                       /*Mutáció
  Z(q) kiszámítása
  Legyen H(q, p_z) = max_j H(q, p_j); j, z ∈ {1, 2, ..., t}.   /*Hasonlóság vizsgálat
  If Z(q) < Z(p_z) then p_z := q fi                        /*Szelekció
  it := it + 1, P(it) ← P(it − 1)
until itt < it                                             /* 1. fázis ismétlési feltétele
** EA2 **
Repeat
  Repeat
    If 0.5 < Rnd then i := glob else i := [t * Rnd] + 1 fi   /* Utód generálás
    q := p_i
    For j := 1 to [Rnd * 2] + 1 do Csere(q) od Nhs(q)        /*Mutáció
    Z(q) kiszámítása
    If Z(q) < Z(p_i) then p_i := q fi                        /*Szelekció
    it := it + 1, P(it) ← P(it − 1)
    If (it mod m) = 0 then Újegyed fi              /* Az egyedek számának növelése
  until (it mod kn) ≠ 0
  Legyen Z(p_i) = min_j Z(p_j), i, j ∈ {1, 2, ..., t}, glob := i /*Legjobb közelítés kiválasztása
  opt = Z(glob), optp = p_glob
  If "futási idő" > timeend then exit fi                   /* Megállási feltétel
until t < tmax                                             /* 2. fázis ismétlési feltétele
** EA3 **
Repeat
  Repeat
    i := [t * Rnd] + 1, q := p_i                           /* Utód generálás
    For j := 1 to [Rnd * 2] + 1 do Csere(q) od, Nhs(q)     /* Mutáció
    Z(q) kiszámítása
    If Z(q) < Z(p_i) then p_i := q fi                      /* Szelekció
    it := it + 1, P(it) ← P(it − 1)
  until (it mod kn) ≠ 0                                    /* Ismétlési feltétel
  SortDel                                  /* Legrosszabb egyedek újracserélése
  Legyen Z(p_i) = min_j Z(p_j), i, j ∈ {1, 2, ..., t}, glob := i /* Legjobb közelítés kiválasztása
  opt = Z(glob), optp = p_glob
until "futási idő" > timeend                               /* Megállási feltétel
exit
end
```

4 Tesztproblémák megoldása

4.1 Paraméterek megválasztása

QAPLIB könyvtár[2] feladatainak megoldásához először néhány nehezebb tesztproblémát választottunk ki, amelyeknél a konvergencia sebességét, a megoldás

[2] http://fmatbhp1.tu.graz.ac.at/~karisch/qaplib

minőségét a paraméter értékek függvényében vizsgáltuk. A megfelelő paraméter értékek megválasztásához minden paraméter szerepét külön vizsgáltuk, tanulmányoztuk külön-külön a három fázis szerepét, és végül a legjobb paraméter-kombináció kiválasztását kíséreltük meg.

E körültekintő folyamat végeredményét a következőképpen foglalhatjuk össze:

- Az első fázis szerepe fontos, javítja a megoldások minőségét. Általában kisebb szórással kapjuk a végeredményeket e fázis alkalmazása esetén. Általában 400 véletlenszerűen generált megoldás elegendő a kezdő populáció minőségének javításához ($itt = 400$).

- A második fázis gyorsítja a konvergenciát. $t = 10$ kezdőpopuláció méret alkalmazásával a tesztproblémák többségénél a leggyorsabb, és legjobb közelítéseket lehet elérni. (A $t = 2$ érték az esetek 30%-ban nagyon jó eredmény nyújt, a többiben a $t = 10$ megfelelőbb.)

- Az m és kn paraméterek, amelyek a populáció méret növelés és az ellenőrzések időpontját határozzák meg, függnek a dimenziótól. Értékeik azonosak lehetnek és kis, közepes és nagy dimenziók esetén (10-30, 31-89, és 90-) rendre 50, 25 és 5 értékek a megfelelőek.

- A harmadik fázisban a populáció mérete szintén dimenziófüggő. Kis, középes és nagy dimenziónál rendre 30, 30 (vagy 60) és 90 elemű populációt ($tmax$) célszerű alkalmazni. A ddp paraméter értéke feladattól függetlenül 0.3-nak választható.

- A futásidők feladatfüggők. A különböző nehézségű feladatok megoldásához ugyanazon dimenzió esetén is más-más futásidők szükségesek a megkívánt pontosság eléréséhez. Így minden feladatnál külön kell a futásidőt ($timeend$) becsülni.

E dimenziótól függő, háromféle paraméterkombinációt alkalmaztuk tehát a QAPLIB feladatainak megoldásánál:

kis dimenzióknál:
$t = 10, \quad tmax = 30, \quad m = 50, \quad kn = 50, \quad ddp = 0.3, \quad itt = 400;$
közepes dimenzióknál:
$t = 10, \quad tmax = 30, \quad m = 25, \quad kn = 25, \quad ddp = 0.3, \quad itt = 400;$
nagy dimenzióknál:
$t = 10, \quad tmax = 90, \quad m = 5, \quad kn = 5, \quad ddp = 0.3 \quad itt = 400.$

Néhány nehéz feladat esetében (3%-a a feladatoknak), amelyeknél csak 1-1.5%-os pontossággal tudtuk a legjobb ismert megoldást közelíteni, a $t = tmax = 60$ paraméterértékek megfelelő eredményeket adtak.

4.2 Teszteredmények

Algoritmusunk tehát a QAPLIB könyvtár [6] tesztproblémáival ellenőriztük. EF_QAP a problémák 98%-át sikeresen oldotta meg; a legjobb ismert megoldást 1%-on belüli pontossággal közelítette meg, vagy megtalálta az ismert legjobb megoldást.

Bemutatásra három csoportot válogattunk a problémák közül. Arra törekedtünk, hogy eredményeink összehasonlíthatók legyenek más módszer(ek) eredményeivel, és különböző dimenziójú, típusú feladatok egyaránt szerepeljenek közöttük. Az első csoport egy ilyen válogatás (Bölte et al. [3] 14 válogatott tesztproblémája), a második csoport a QAPLIB további 100-256 dimenziós problémái és a harmadik csoport, néhány időigényesebb probléma, melyet sikeresen oldott meg az EF_QAP.

Összehasonlításhoz Ahuja et al. [1] greedy GA módszerét (jelölés: gGA) választottuk, mellyel a QAPLIB maximum 100 dimenziós problémáinak túlnyomó részét megoldották (100 dimenzió feletti problémák megoldására nem alkalmas a gGA). A gGA-t C-ben programozták és egy HP 6000 számítógépen futtatták. Mi az EF_QAP-t Visual Basic-ben programoztuk és egy HP Brio (400 Mhz) számítógépen futtattuk. Bár a mi algoritmusunk lassabban futott, az időeredmények mégis lehetővé tették a módszerek összehasonlítását.

A 14 válogatott tesztprobléma eredményét az *1. táblázat*, a további, magasabb dimenziós problémák eredményét a *2. táblázat*, az "időigényesebb" problémák eredményeit pedig a *3. táblázat* tartalmazza. Mindhárom táblánál a bemutatott eredmények több futtatás átlagos eredményei: a 70 dimenzió alatti problémákat 10-szer, a magasabb dimenziósakat pedig 5-ször futtattuk le. Az egyes problémák futási idejét a *timeend* paraméterrel adtuk meg. (Egy probléma minden egyes futásánál eredménynek a végső pontosságot és a pontosság eléréséhez szükséges futási időt tekintettük). A táblákban megadtuk a probléma nevét és dimenzióját, az ismert legjobb megoldást, az ismert legjobb megoldástól való átlagos eltérést százalékban (hiba), az átlagos futási időt (CPUsec), a *timeend* paraméter értékét és az ismert legjobb megoldás 1%-os környezetébe eső megoldások százalékát (Opt1%). Az EF_QAP és a gGA összehasonlításához a gGA adatait is megadtuk (amennyiben ismertek).

Név	Dim	Ismert legjobb	Hiba (%)		CPUsec			$\text{Opt}_{1\%}$
			gGA	EF_QAP	gGA	EF_QAP	*timeend*	EF_QAP
Nug20	20	2670	0.00	0.00	48.9	10.3	12	100
Nug30	30	6124	0.07	0.11	177.1	73.6	105	100
Tho30	30	149936	0.00	0.27	197.8	80.2	150	100
Tho40	40	240516	0.32	0.28	479.0	129.7	150	100
Sko42	42	15812	0.25	0.20	503.1	302.7	500	100
Sko49	49	23386	0.21	0.29	626.1	338.0	500	100
Wil50	50	48816	0.07	0.10	1057.6	430.2	650	100
Sko56	56	34458	0.02	0.15	1488.0	1077.0	1350	100
Sko64	64	48498	0.22	0.36	1894.1	1654.0	1800	100
Sko72	72	66256	0.29	0.23	2539.0	1383.0	1800	100
Sko81	81	90998	0.20	0.33	5482.1	4789.0	5000	100
Sko90	90	115534	0.27	0.34	6348.9	6336.0	6500	100

1. táblázat. Válogatott tesztproblémák

Név	Dim	Ismert legjobb	Hiba (%)		CPUsec			$\text{Opt}_{1\%}$
			gGA	EF_QAP	gGA	EF_QAP	*timeend*	EF_QAP
Sko100a	100	152002	0.21	0.31	16608	2348	2500	100
Sko100b	100	153890	0.14	0.14	14735	1829	2000	100
Sko100c	100	147862	0.20	0.18	20314	3818	4200	100
Sko100d	100	149570	0.17	0.34	20302	7751	9000	100
Sko100e	100	149150	0.24	0.30	20127	7942	9000	100
Sko100f	100	149036	0.29	0.46	20479	6112	7500	100
Wil100	100	273038	0.20	0.26	20544	4525	8000	100
Tai100a	100	21125314		0.98		16328	20000	100
Tai100b	100	1185996137		0.88		5449	6000	100
Esc128	128	64		0.00		359	1200	100
Tho150	150	8133484		0.43		23564	26000	100
Tai150b	150	498896643		0.80		11809	13000	100
Tai256c	256	44759294		0.14		35458	36000	100

2. táblázat. Magasabb dimenziós tesztproblémák

Név	Dim	Ismert legjobb	Hiba (%)		CPUsec			$\text{Opt}_{1\%}$
			gGA	EF_QAP	gGA	EF_QAP	*timeend*	EF_QAP
Chr20b	20	2298	5.13	0.50	96	279	450	80
Ste36a	36	9526	0.27	0.06	710	690	1000	100
Lipa50a	50	62093	0.95	0.68	15486	678	1200	100
Lipa50b	50	1210244	0.00	0.00	1509	282	500	100
Lipa60a	60	107218	0.77	0.79	3057	1518	2000	100
Lipa60b	60	2520135	0.00	0.00	3047	1289	2000	100
Lipa70a	70	169755	0.71	0.76	6148	1757	2000	100
Lipa70b	70	4603200	0.00	0.00	6122	3026	5000	100
Lipa80a	80	253195	0.61	0.59	9518	3600	5000	100
Lipa80b	80	7763962	0.00	0.00	94989	4463	7000	100
Lipa90a	90	360630	0.58	0.56	12358	4123	6500	100
Lipa90b	90	12490441	0.00	0.00	12319	8147	14000	100
Tai50a	50	4941410		1.07		1490	3000	50
Tai60a	60	7208572		0.85		6948	7200	100
Tai80a	80	13557864		1.00		8862	14000	60

3. táblázat. Időigényesebb tesztproblémák eredményei

Az 1. táblázat eredményeit vizsgálva megállapítható, hogy az EF_QAP gyorsabb a gGA-nál. Általában az EF_QAP 2-3-szor hamarabb találja meg a hasonló pontosságú megoldást (0.1%-os az eltérés a két módszer eredményeiben). A megoldás menetét vizsgálva megállapítható, hogy a kívánt pontosságot már a második fázisban, néhány újabb megoldás generálásával, gyorsan eléri az algoritmus.

A 2. táblázat eredményei hasonlóak az előzőhöz: az EF_QAP ismét gyorsabban találta meg a hasonló pontosságú eredményt (A QAPLIB magasabb dimenziós problémáinak csak a felénél ismertek a gGA eredményei). A megoldás menete ugyancsak hasonló az előző feladatokéhoz; itt is a kívánt pontosságot már a második fázisban elérte az algoritmus, nem volt szükség a harmadik fázis helyi keresésére.

A 3. táblázat azon feladatokat, feladatcsoportokat gyűjti össze, amelyeket az EF_QAP algoritmus átlagosnál hosszabb idő alatt oldott meg. Az időadatokat összehasonlítva ennek ellenére most is az EF_QAP bizonyult gyorsabbnak, csak most nagyobb a szórás az egyes módszerek időadatai közt. E feladatoknál minden esetben szükség volt a harmadik fázis helyi keresésére is. Ilyen problémák pl. az Ste36a, Lipa50a, Lipa50b, ..., Lipa90a, Lipa90b voltak.

Az időigényesebb problémák egy kis részénél viszont a második fázis "gyors" helyi keresése bizonyult eredménytelennek (azaz, a szokásos módszerekhez képest csak lassabban tudta a globális optimumot közelíteni). Feltételezve azt, hogy a helyi optimumok értékei alig különböznek egymástól, sikerült e feladatokat is gyorsabban megoldani. E feladatoknál az első fázisban mindjárt a maximális populáció méretet alkalmaztuk ($t = tmax$). Evvel automatikusan kihagytuk a második fázist, és a harmadik fázis helyi keresése már kellő gyorsasággal adta a kívánt eredményt. Ilyen problémák, pl. tai50a, tai60a, tai80a voltak. E tábla problémái közt találjuk ugyancsak azt a három feladatot (chr20b, tai50a és tai80a), amelyeket nem tudott az algoritmus minden esetben kellő pontossággal megoldani.

A QAPLIB további feladatait szintén sikerrel oldotta meg az EF_QAP. Minden feladatot átlagosan 0.00 (vagy nulla) hibával oldott meg és általában 2-3-szor hamarabb találta meg a gGA-hoz képest a hasonló pontosságú megoldást.

Összességében megállapítható, hogy az EF_QAP a QAPLIB tesztproblémáit meg tudja oldani, és három probléma kivételével, a megoldások mindig az ismert legjobb megoldás 1%-os környezetébe esnek. A futási idők összehasonlítása igen nehéz különböző gépek és programozási nyelvek esetén, de kísérletet tettünk a gGA-nál publikált futási idők [1] összehasonlítására. Ezek alapján az EF_QAP általában 2-3-szor gyorsabban találja meg a hasonló pontosságú eredményt, mint a gGA.

5 Összefoglalás

Az EF_QAP egy háromfázisú evolúciós algoritmus a QAP megoldására. Olyan heurisztika, amellyel kis, közepes és nagyméretű feladat egyaránt megoldható. Az algoritmus a QAPLIB könyvtár jól ismert teszt problémáit sikeresen oldotta meg, a feladatok 98%-ban az ismert legjobb megoldás 1%-os környezetébe került, vagy megtalálta az ismert legjobb megoldást. Összehasonlítva egy másik heurisztikus módszerrel, Ahuja et al. [1] gGA módszerével, megállapíthatjuk, hogy az EF_QAP, a 100 dimenziósnál nem nagyobb feladok esetén, 2-3-szor rövidebb idő alatt nyújtja a hasonló pontosságú eredményeket.

Az algoritmus működését 7 paraméter befolyásolja, de többségük a különböző feladatoknál azonos, rögzített értékkel használható. Az algoritmus működését vizsgálva megállapítható, hogy a tesztproblémák esetén az ismert legjobb megoldást 1-2%-os pontossággal viszonylag hamar meg tudja közelíteni. Utána fokozatosan, egyre lassabban javítja a közelítéseket. A konvergencia gyorsítását az algoritmus első és második fázisa segíti. A gyorsaság szempontjából különösen a 2. fázis fontos: az újabb közelítések generálásával lényegesen meg tudja növelni a keresés gyorsaságát és hatékonyságát.

Köszönetnyilvánítás

A kutatás az OTKA T 030861 támogatásával készült.

Irodalom

1. Ahuja K. R., Orlin J. M., Tiwari A.: A greedy genetic algorithm for the quadratic assignment problem. *Computer & Operations Research* 27 (2000) 917–934.
2. Borgulya I.: Constrained optimization using a clustering algorithm. *Central European Journal of Operations Research.* 8 (1) (2000) 13–34.
3. Bölte A., Thonemann U. W.: Optimizing simulated annealing schedules with genetic programming. *European Journal of Operational Research.* 92 (1996) 402–416.
4. Burkard R. E., Rendl F.: A thermodynamically motivated simulation procedure for combinatorial optimization problems. European Journal of Operations Research, 17(2) (1984) 169–174.
5. Burkard R. E.: Locations with spatial interactions: the quadratic assignment problem. In: Mirchandani P. B., Francis R. L.(eds), *Discrete Location Theory.* Wiley, Berlin, 1991
6. Burkard R. E., Karisch S. E., Rendl F.: QAPLIB - A quadratic assignment library. *Journal of Global Optimization.* 10 (1997) 391–403.
7. Burkard R. E., Çela E., Pardalos P. M., Pitsoulis L. S.: The Quadratic Assignment Problem. In: Pardalos P. M., Du D. Z. (eds): *Handbook of Combinatorial Optimization* Vol. 3 Kluwer Academic Publishers, Dordrecht 1998, pp. 241–339.
8. Çela E.: *The Quadratic Assignment Problem: Theory and Algorithms.* Kluwer Academic Publishers, Dordrecht, 1998.
9. Connolly D. T.: An improved annealing scheme for the QAP. *European Journal of Operational Research.* 46 (1990) 93–100.
10. Fleurent C., Ferland J. A.: *Genetic Hybrids for the Quadratic Assignment problem*, DIMACS Series in Mathematics and Theorical Computer Science, 16 (1994) 173–187.
11. Gambardella L. M., Taillard E. D., Dorigo M.: Ant Colonies for the QAP. *Journal of Operation Research Society* 50 (1999) 167–176.
12. Li Y., Pardalos P. M., Resende M. G. C.: A greedy randomized adaptive search procedure for the quadratic assignment problem. In Pardalos P. M., Wolkowicz H. (eds): *Quadratic assignment and related problems*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 16 (1994) 237–261
13. Maniezzo V., Dorigo M., Colorni A.: Algodesk: An experimental comparison of eight evolutionary heuristics applied to the Quadratic Assignment Problem. *European Journal of Operational Research.* 81 (1995) 188–204.
14. Nissen V.: Quadratic assignment, In: Bäck T., Fogel D. B., Michalewicz Z.: *Handbook of Evolutionary Computation.* Oxford University Press, 1997. pp. G9.10:1–9.
15. Pardalos P. M., Rendl F., Wolkowicz H.: The quadratic assignment problem: a survey of recent developments. In Pardalos P. M., Wolkowicz H.(eds): *Quadratic assignment and related problems*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 16 (1994) 1–42.

16. Taillard D.: Comparison of iterative searches for the quadratic assingnment problem. *Location Science*, 3 (1995) 87–105.

17. Talbi E. G., Hafidi Z., Geib J-M.: *Parallel adaptive tabu search for large optimization problems.* Research report LIFL URA 369 CNRS, University of Lille, March 1997.

A HEURISTIC ALGORITHM FOR THE QUADRATIC ASSIGNMENT PROBLEM

In this paper we describe a new evolutionary algorithm for the Quadratic Assignment Problem. This method can be divided into three stages: a preparatory and two local search stages. The first stage improves the quality of the initial population. The second stage improves the quality of the solutions with a local search procedure while periodically generating new solutions. In the third stage the algorithm continues the application of the local search procedure and replaces the weakest solutions with new ones. We tested our algorithm on all the benchmark problems of QAPLIB. The algorithm managed to find solutions which are either best known or within 1% of the best known solutions for 98% of all tasks.

KÉSZLETELOSZTÁSI JÁTÉK ÉS OPTIMÁLIS ÖSSZMENNYISÉG MESTERSÉGES PIACOKON [1]

UJHELYI GERGELY
CEU Economics Department és Rajk László Szakkollégium, Budapest

A dolgozat különböző piacszerkezetek esetén tárgyalja a mesterséges piacok kialakításának két kulcsproblémáját: a piaci jószág összes mennyiségének meghatározását és e jószágmennyiség szétosztását a szereplők között. Az elosztás modellezéséhez használt "készletelosztási játék" segítségével a hatékonyságtól különböző elosztási elvek tulajdonságai is vizsgálhatók Megmutatom, hogy ezek a tulajdonságok a piaci szerkezettel változnak. Az összes mennyiség meghatározását egy társadalmi tervező feladataként tárgyalom. Belátom, hogy az optimális összmennyiség módosul a piaci szerkezet és az elosztási szabály változásával. A következtetések alkalmazási lehetőségeit egy gyakorlati példán, a szennyezési jogok nemzetközi kereskedelmén keresztül mutatom be.

1 Bevezetés

A gazdasági szabályozás egyes területein egyre nagyobb szerephez jut a mesterséges piacok kialakítása. Ilyen a környezetszabályozás, ahol a tulajdonjogok (szennyezési jogok, vízhasználati jogok) kereskedelmére épülő különböző rendszerek terjedőben vannak. Egy másik terület a mezőgazdaság, ahol egyes termékek (pl. tej, dohány) termelési kvótáinak kereskedelmére van sokhelyütt lehetőség. A szerződéselvű politikai filozófia megközelítésében az emberi társadalom egésze is egy mesterségesen (az emberek szerződésén keresztül) létrehozott piacot alkot.

E piacok kialakításának három kulcslépése a következő: 1. a piaci jószág összes mennyiségének meghatározása; 2. e jószágmennyiség szétosztása a szereplők között; 3. a piac működése (a piaci szerkezet). A különböző alkalmazásokban az összmennyiséggel kapcsolatos legfontosabb kérdések, hogy mekkora annak optimális nagysága, valamint hogy ezt milyen körülmények befolyásolják. Mekkora legyen a szétosztandó szennyezési jogok összes mennyisége? A gazdasági környezet mely jellemzőit vegyük figyelembe a teljes tejkvóta megállapításakor? Előfordulhat-e, hogy a teljes jószágmennyiség

[1]Beérkezett: 2002. február 19. A tanulmány a BKÁE-n készített szakdolgozatom átdolgozott változata (Ujhelyi, 2001). Szeretném megköszönni Forgó Ferenc észrevételeit és támogatását, köszönöm továbbá Kaderják Péter, Reiff Ádám, Simonovits András, Szabó Andrea, Szakadát László, Váradi Balázs és Virág Gábor megjegyzéseit a korábbi változatokhoz. A kutatás során sokat segített a Rajk László Szakkollégium inspiráló légköre is.
Levelezési cím: 1118 Budapest, Háromszék u. 17/B. E-mail: ujhelyi@rajk.bke.hu

optimális marad, ha változnak a piac működési feltételei – például tranzakciós költségek jelennek meg? A jószág kezdeti szétosztásával kapcsolatban elsősorban a különböző elosztási elvek tulajdonságai fontosak. Milyen kapcsolatban áll a szennyezési jogok vagy a termelési kvóták hatékony elosztása az „igazságosság" különböző kritériumaival? Hogyan változnak ezeknek az elveknek a jellemzői, ha módosul a piac működése? Ezek olyan, a gyakorlati alkalmazásokban lényeges kérdések, amelyek általában kimaradnak a közgazdasági elemzésekből.

A mesterséges piacokkal foglalkozó közgazdasági kutatások jellemzően a 3. lépcsőre összpontosítanak, és a különböző piacszerkezetek tulajdonságait vizsgálják – az olyan alapművekre, mint Hahn (1984) vagy Stavins (1995) részletesen fogok hivatkozni a későbbiekben. Ezek az összmennyiséget egzogén adottságként kezelik, az elosztási elvek tárgyalásánál pedig nem lépnek túl a hatékonyságon. Ebben a tanulmányban ezért a piacszerkezetek hagyományos modelljeire építve az összmennyiség meghatározását és az elosztást, illetve ezek kölcsönhatásait vizsgálom. Az összmennyiség meghatározását egy társadalmi tervező optimumfeladataként írom fel, az elosztás modellezésére pedig egy kooperatív alkujátékot, „készletelosztási játékot" használok. Ez utóbbi segítségével a hatékonyság mellett lehetőség nyílik különféle „méltányos" elosztási elvek vizsgálatára is.

E tanulmány szelleméhez közel áll Amacher-Malik (1996) és (1998), akik kooperatív alkujátékot használnak vállalatok és egy környezetszabályozó hatóság közti kapcsolat leírására. Náluk az alku a vállalatok szennyezésére és a szabályozás módjára vonatkozik. Egy másik előzménynek a klímaváltozási egyezmények elméleti irodalma tekinthető, ahol a kooperatív játékelmélet az elemzés hagyományos eszközei közé tartozik (lásd pl. Carraro - Siniscalco, 1993; Barrett, 1994; Chander - Tulkens, 1997). Ebben a megközelítésben az országok koalíciókat alkotva, vagy kooperatív alkufolyamatokon keresztül határozzák meg a szennyéscsökkentéseket. Ugyanakkor a szennyezési jogok kereskedelme, vagyis a piac hiányzik ebből az irodalomból. A modellekben az országok nem indulókészletekről, hanem magáról a szennyezőanyag-fogyasztásról alkudoznak. Ez az irodalom tesz utalást az összes szennyezés endogén meghatározására is, azonban itt sem használja a piac modelljeit.

A dolgozat felépítése a következő. A 2. szakaszban megadom a vizsgált három piac-modellt. A 3. szakasz a jogok elosztását elemzi: bevezeti a jogok elosztását leíró "készletelosztási játékot" és a vizsgált megoldási elveket. Ezek után leírom a különböző megoldások alapján nyert eredményeket a három piacon. A 4. szakasz tárgyalja az összmennyiség meghatározását versenyző, majd tökéletlen piacok esetén. Az 5. szakaszban az elmondottak egy fontos alkalmazási lehetőségét mutatom be a szennyezési jogok nemzetközi kereskedelmével kapcsolatban. Végül a 6. szakasz az összefoglalást és az elemzés további lehetőségeit tartalmazza. A 3. és 4. szakaszok végén egy számpélda szemlélteti a leírtakat. A kevésbé érdekes bizonyítások függelékben szerepelnek.

2 A piac

Ebben a szakaszban bevezetem a (mesterséges) piac három egyszerű modelljét, amelyekre a későbbiek során építeni fogok. Ezek a versenyzői, a tranzakciós költségeket tartalmazó, és az ármeghatározó játékossal működő modellek.

A versenyzői piac

A piac résztvevője n db $i = 1, \ldots, n$ indexszel jelölt játékos. A piacon adott a q jószág $\overline{q}$ összes mennyisége és $\mathbf{q}^0 = (q_1^0, \ldots, q_n^0)$ elosztása, $\sum q_i^0 = \overline{q}$. Termelés nincs. Jelölje $u_i(q_i)$ folytonos, kétszer deriválható, szigorúan konkáv függvény a q_i nagyságú (pozitív) fogyasztásból származó hasznosságot az i-edik játékos számára. Legyen q_i^u az i-edik játékos telítődési pontja, méghozzá $u_i'(q_i^u) = 0$, $u_i'(q_i) > 0$ $q_i \in [0, q_i^u)$, $u_i'(q_i) \leq 0$ $q_i \in (q_i^u, \infty)$, és legyen $u_i(0) = 0$. A piacon a játékosok értékesíthetik az el nem fogyasztott készletet, illetve további jószágegységeket vásárolhatnak a p piaci áron. Felteszem, hogy a játékosok hasznossága a pénzben lineáris, vagyis az i-dik játékos feladata

$$\max_{q_i} u_i(q_i) - p(q_i - q_i^0) \,.$$

A versenyző piacok egyensúlyát jellemző szokásos elsőrendű feltétel a következő:

$$u_i'(q_i) = p \,, \tag{2.1}$$

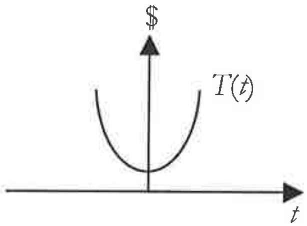
amely szerint optimumban a határhasznok kiegyenlítődnek. Látszik továbbá, hogy a játékosok indulókészletüktől függetlenül hozzák meg fogyasztási döntéseiket.

Tranzakciós költségek

A piac működésével járó tranzakciós költségek nagyok lehetnek a mesterségesen létrehozott piacok esetén. A piacra való belépés, az információhoz való hozzájutás és a tranzakciók lebonyolítása nehézkes lehet (Stavins, 1995). Különösen a működés kezdetén, a játékosok számára magas költségekkel járhat a szabályok megismerése, elsajátítása (OECD, 1998).

A tranzakciós költségeket tartalmazó modell Stavins (1995)-ön alapul. Tegyük fel, hogy a piaci tranzakcióban résztvevő játékosok a tranzakció volumenétől függően költségeket szenvednek el. Az i-edik játékos által vásárolt (vagy eladott) összes mennyiség legyen $t_i := q_i - q_i^0$. $T(t_i)$ jelöli a nemnegatív, folytonos és differenciálható tranzakciós költségfüggvényt, amelyre $T(t_i) \equiv T(-t_i)$, vagyis a költség szempontjából mindegy, hogy a játékos vevőként vagy eladóként jelenik meg a piacon. Ez utóbbi feltételből könnyen látszik, hogy $T'(0) = 0$, vagyis a tranzakciós határköltség 0 tranzakció esetén zérus.[2]

[2] $T(t) \equiv T(-t)$ miatt $dT(t)/dt = -dT(-t)/dt$ teljesül. Ha $T'(0)$ létezik, akkor $dT(0)/dt = -dT(0)/dt$, ahonnan $T'(0) = 0$.



2.1. ábra. Monoton tranzakciós költségfüggvény

Kétféle tranzakciós költséget vizsgálok: *állandó tranzakciós költség:* $T'(t) \equiv 0$ $\forall t$-re; *monoton tranzakciós költség:* $T'(t)t > 0$, ha $t \neq 0$. Egy monoton tranzakciós költségfüggvényt szemléltet a *2.1. ábra.*

Az i-edik játékos feladata most

$$\max_{q_i} u_i(q_i) - p(q_i - q_i^0) - T(q_i - q_i^0) ,$$

az elsőrendű feltétel pedig

$$u_i'(q_i) - T'(q_i - q_i^0) = p . \tag{2.2}$$

(A másodrendű feltétel teljesüléséhez tegyük fel, hogy $u_i'' - T'' < 0$.) (2.2) szerint tranzakciós költségek esetén a piac nem a piaci jószágra vonatkozó határhasznokat egyenlíti ki, hanem a határhasznok és a tranzakciós határköltségek összegét. (2.2)-ből láthatóan monoton tranzakciós költségek esetén a játékosok optimális fogyasztása indulókészletüktől is függ, míg állandó tranzakciós költségek esetén továbbra is független tőle.

Ármeghatározó játékos

A tranzakciós költségek mellett az irodalomban sokat tárgyalt másik piaci tökéletlenség a piaci erőfölény esete. Bár a gyakorlatban megfigyelt mesterséges piacok esetében eddig sehol nem tapasztaltak jelentős erőfölényt, laboratóriumi szimulációk esetén erre volt példa (Muller - Mestelman, 1998). A nemzetközi szennyezési piaccal kapcsolatban is sokan hangot adtak azon aggodalmuknak, hogy a piacot néhány nagy eladó/vevő dominálhatja (Nordhaus - Boyer, 1998). Az itt megadott modell Hahn (1984)-et követi, aki az erőfölényt az ármeghatározás képességeként definiálta.

Tegyük föl, hogy az 1-gyel jelölt játékos ármeghatározóként viselkedik a piacon abban az értelemben, hogy képes megválasztani a számára legkedvezőbb árat, feltéve, hogy nem sérül a $\overline{q}$ aggregált korlát. Az $i = 2, \ldots, n$ játékosok feladata továbbra is

$$\max_{q_i} u_i(q_i) - p(q_i - q_i^0) ,$$

és a hasznosságfüggvény szigorú konkavitása miatt lokálisan létezik az (2.1) elsőrendű feltételek által definiált $q_i^* = q_i(p)$ negatív meredekségű,[3] folytonos

[3] Alkalmazva az implicit függvény tételt a $p - u_i'(q_i) = 0$ egyenletre $q_i'(p) = 1/u_i''(q_i) < 0$ adódik.

és differenciálható függvény (a játékosok keresleti függvénye). Az 1-gyel indexelt játékos ekkor a többi játékos optimális reakcióját adottnak véve határozza meg a költségét minimalizáló árat és fogyasztást, feltéve, hogy a teljes fogyasztás nem lépi túl az aggregált korlátot:

$$\begin{aligned} \max_{p,q_1} \quad & u_1(q_1) - p(q_1 - q_1^0) \\ \text{k. f.} \quad & q_1 + \sum_{i=2}^{n} q_i(p) = \overline{q}\,. \end{aligned}$$

Behelyettesítve a korlátot a célfüggvénybe és megoldva, a következőt kapjuk (vö. Hahn (1984) 756. o.):

$$[u_1'(q_1^*) - p] \cdot \sum_{i=2}^{n} q_i'(p) + q_1^* - q_1^0 = 0 \tag{2.3}$$

$$q_1^* = \overline{q} - \sum_{i=2}^{n} q_i(p)\,. \tag{2.4}$$

A feltételekből jól látszik, hogy a piaci árra, következésképpen az egyensúlyi fogyasztásokra döntő befolyást gyakorol az ármeghatározó játékos q_1^0 indulókészlete. Ennek fontos következményei lesznek a későbbiekben.

3 A kezdeti elosztás

Ebben a szakaszban egy kooperatív alkujátékot javaslok az elosztás modellezésére. Először felírom a „készletelosztási játékot", majd ismertetem a vizsgálandó megoldásokat és a hozzájuk tartozó elosztási elveket. Ezután mutatom be az elemzés eredményeit a versenyző, a tranzakciós költségeket tartalmazó, majd az ármeghatározó játékossal zajló készletelosztási játékra.

A készletelosztási játék

A készletelosztási játékban a játékosok egy alkufolyamat során rögzített nagyságú összkészletet osztanak szét egymás között. Ez a szétosztás egy piaci játék indulókészleteit alkotja, amelyre az előző részben tárgyalt forgatókönyvek valamelyike mellett kerül sor. Itt nyílik lehetőség a szétosztott készletek adás-vételére. A játékosok tökéletesen informáltak: a készletek szétosztásánál figyelembe veszik a következő időszakban várható piaci helyzetet. Mint a 2. szakaszban láttuk, a piacon az i-edik játékos hasznát maximalizálja, adottnak véve a $\mathbf{q}^0 = (q_1^0, \ldots, q_n^0)$ készletvektort ($\sum q_i^0 = \overline{q}$).

Jelölje $v_i(\mathbf{q}^0)$ a célfüggvény optimális értékét a $\mathbf{q}^0$ elosztás függvényében. $v_i(\mathbf{q}^0)$ lesz az i-edik játékos kifizetőfüggvénye a készletelosztási játékban, ahol a játékosok a $\mathbf{q}^0$ készletvektorról alkudoznak.[4] Tegyük fel, hogy amennyiben

[4]Elosztáson a továbbiakban mindig megvalósítható elosztást fogunk érteni, vagyis teljesül, hogy $\sum q_i^0 = \overline{q}$ és $q_i^0 \geq 0$.

nem jön létre egyezség, az i-edik játékos kifizetése $-D_i$ $(D_i \geq 0)$.[5] Hogy az alkunak értelme legyen, feltesszük, hogy van olyan $\mathbf{q}^0$ elosztás, amelyben $-D_i < v_i(\mathbf{q}^0)$ minden i-re, vagyis mindig létezik olyan elosztás, amely mellett minden játékosnak érdekében áll részt venni az alkuban. A későbbiekben fontos szerepe lesz annak a nettó nyereségnek, amelyre a játékosok a megegyezés következtében szert tesznek. Jelöljük ezt R-rel:

$$R_i(\mathbf{q}^0) = v_i(\mathbf{q}^0) + D_i\,.$$

Mindez az első periódusra egy $(V, -\mathbf{D})$ alkujátékot definiál, ahol

$$V = \mathrm{con}\left(\{\, \mathbf{v}(\mathbf{q}^0) = (v_1(\mathbf{q}^0), \ldots, v_n(\mathbf{q}^0)) \mid \sum q_i^0 = \overline{q},\ q_i^0 \geq 0 \,\}\right)$$

a lehetséges kifizetésvektorok konvex burka, $-\mathbf{D} = (-D_1, \ldots, -D_n)$ pedig a fenyegetési pont. Jelölje $\mathcal{B}$ az összes alkujáték halmazát. A $\varphi : \mathcal{B} \to \Re^n$ függvény az alkujáték egy (lehetséges) megoldása, amennyiben $\varphi(V, -\mathbf{D}) \in V$ teljesül. A megoldás tehát egy függvény, amely adott játékhoz egyértelműen hozzárendeli a lehetséges kifizetésvektorok valamelyikét.

Definíció. *Ha $\mathbf{q}^0$-ra $\mathbf{v}(\mathbf{q}^0) = \varphi(V, -\mathbf{D})$, akkor azt mondjuk, hogy a $\mathbf{q}^0$ elosztás a $\varphi(V, -\mathbf{D})$ megoldást valósítja meg.*

A készletelosztási játékban tehát a játékosok adott nagyságú összkészletet osztanak szét egymás között, az így kapott indulókészletek pedig a következő időszakban a piacon keresztül valósítják meg az alkujáték megoldásait. 1950-es alapművében Nash bizonyos értelemben éppen fordított logikával teremtett kapcsolatot kooperatív és nem-kooperatív játékok között. Nála a kooperatív alkujátékot megelőzően a játékosok egy nem-kooperatív játék keretében rögzítették a D_i fenyegetési pontokat. Ha az alku nem valósult meg, a játékosok csak a nem-kooperatív játék megoldásaként kapott fenyegetési pontokat használhatták fel (fogyaszthatták el). Az itt leírt modell éppen fordítva működik. Itt a nem kooperatív piaci játékot előzi meg egy kooperatív alkufolyamat, amelynek során kialakulnak a piac indulókészletei. Amennyiben valamelyik játékos ezek után nem vesz részt a piacon, csak indulókészletét használhatja föl. Itt tehát a nem kooperatív játék „fenyegetési pontjai" alakulnak ki egy kooperatív játékon keresztül.[6]

Elosztási elvek

Háromfajta elosztási szabályt vizsgálok: az *utilitarista*, a *Nash-*, és az *egyenlő nyereségeket biztosító (Equal Gains - EG)* elosztást.

[5]Ebben a témakörben a fenyegetési pontban a kifizetés tipikusan nem-pozitív lesz – gondoljunk például a szabályozás elmaradása esetén a játékosokat érő környezeti kárra.

[6]Mindez hasonlóságot mutat a szerződéselvű politikai filozófia szemléletmódjával. Rawls (1971) híres elméletében például a játékosok a tudatlanság fátyla mögött (kooperatívan) állapodnak meg a kialakuló társadalom (a piac) alapelveiről, és bizonyos alapvető javak elosztásáról.

Definíció (utilitarista elosztás). *Egy* $\mathbf{q}^0$ *elosztást utilitaristának fogunk nevezni, ha maximalizálja a játékosok kifizetéseinek összegét, vagyis*

$$\mathbf{q}^0 = \arg\max_{\mathbf{q}^0} \sum_{i=1}^{n} v_i(\mathbf{q}^0) \, .$$

A társadalmi hasznot maximalizáló utilitarista elosztás alapvető jelentőségű a közgazdaságtanban, vizsgálata viszonyítási pontként fog szolgálni más típusú elosztások elemzéséhez.

Definíció (Nash elosztás). *Egy* $\mathbf{q}^0$ *elosztást Nash-elosztásnak fogunk nevezni, ha az általa megvalósított* $\varphi(V, -\mathbf{D})$ *megoldás a* $(V, -\mathbf{D})$ *alkujáték Nash-megoldása.*[7]

Mint ismeretes, a Nash-megoldást a következő axiómák határozzák meg egyértelműen (lásd pl. Forgó et al., 1999):

1. (Racionalitás) $\varphi(V, -\mathbf{D}) \geq -\mathbf{D}$.
2. (Pareto-hatékonyság) Ha $\mathbf{v} \in V$ és $\mathbf{v} \geq \varphi(V, -\mathbf{D})$ akkor $\mathbf{v} = \varphi(V, -\mathbf{D})$.
3. (Kedvezőtlen alternatíváktól való függetlenség) Ha $V^1 \subset V$, $\varphi(V, -\mathbf{D}) \in V^1$, akkor $\varphi(V^1, -\mathbf{D}) = \varphi(V, -\mathbf{D})$.
4. (Kovariancia a hasznosságreprezentációk lineáris transzformálására nézve) Ha $V' = \{(\alpha_1 v_1 + \beta_1, \ldots, \alpha_n v_n + \beta_n) \mid (v_1, \ldots, v_n) \in V\}$ és $\mathbf{D}' = (\alpha_1 D_1 + \beta_1, \ldots, \alpha_n D_n + \beta_n)$, akkor $\varphi(V', -\mathbf{D}') = (\alpha_1 \varphi_1 + \beta_1, \ldots, \alpha_n \varphi_n + \beta_n)$.
5. (Szimmetria) Ha valamely (i, j) indexpárra igaz, hogy $\mathbf{v} \in V$ akkor és csak akkor, ha $\mathbf{v}' \in V$ (ahol $v_k = v'_k$ $k \neq i, j$, $v_i = v'_j$, $v_j = v'_i$), továbbá $D_i = D_j$, akkor teljesül, hogy $\varphi_i = \varphi_j$.

Ezek mindegyike fontos elv lehet a mesterséges piacok kialakítása során. Az 1-2. axiómák alapvető fontosságúak. Az adott elosztási szabályt a „Racionalitás" teszi egyénileg, míg a „Pareto-hatékonyság" társadalmilag elfogadhatóvá. A 3-4. követelmények egyfajta stabilitást testesítenek meg. A 3. axióma azt a követelményt fogalmazza meg, hogy amennyiben a lehetséges kifizetések halmaza irreleváns alternatívákkal bővül, ez ne változtasson az elosztási szabályon. Ehhez hasonlóan a 4. axióma azt írja elő, hogy a lehetséges kifizetések és a fenyegetési pontok változásakor (pl. más pénznemre történő átszámításkor, vagy a hasznok lineáris változásával) az elosztás „természetes módon" változzon. Mesterséges piacok létrehozásánál, ahol a modell számos paramétere igen bizonytalan lehet, sokat nyerhetünk azzal, ha új, esetleg irreleváns információk esetén nem kell, vagy nem kell radikálisan

[7] Az irodalomban a Nash-megoldást is az „utilitarista" jelzővel szokták ellátni, miután kiszámítása egy (speciális) jóléti függvény maximalizálását teszi szükségessé. Én az „utilitarista elosztást" az imént definiált szűkebb értelemben használom, vagyis a Nash-elosztás itt nem feltétlenül utilitarista.

változtatni a szabályokon. Végül az 5. axióma egyfajta igazságosságot fejez ki. Egyenlő bánásmódot ír elő olyan azonos helyzetben lévő játékosok esetén, akiknek a lehetőségei is hasonlóak.

Definíció (EG elosztás). *Egy* $\mathbf{q}^0$ *elosztást egyenlő nyereségeket biztosító (EG) elosztásnak fogunk nevezni, ha az általa megvalósított megoldásra* $\forall j, k \in \{1, \ldots, n\}$ *esetén teljesül, hogy*

$$\varphi_j(V, -\mathbf{D}) + D_j = \varphi_k(V, -\mathbf{D}) + D_k$$

vagy másképpen

$$R_j(\mathbf{q}^0) = R_k(\mathbf{q}^0)\,.$$

Az EG elosztás azt írja elő, hogy minden játékosnak egyformán kell részesednie a piac felállításával keletkező összes haszonból. Ez az elv egyfajta méltányosságot testesít meg.

A versenyző készletelosztási játék

Az imént megadott három elosztást és ezek kapcsolatát először versenyző készletelosztási játék esetén jellemzem. Mint az 1. szakaszban láttuk, tökéletes verseny esetén a szereplők célfüggvénye a piacon

$$\max_{q_i} \; u_i(q_i) - p(q_i - q_i^0)\,. \tag{3.1}$$

Az optimális q_i^* minden i-re eleget tesz az elsőrendű feltételnek:

$$u_i'(q_i^*) = p\,. \tag{3.2}$$

A célfüggvény optimális értéke ekkor csak a saját indulókészlettől függ, mégpedig lineárisan:

$$v_i(\mathbf{q}^0) \equiv v_i(q_i^0) \equiv u_i(q_i^*) - p(q_i^* - q_i^0)\,.$$

Az első periódus készletelosztási játékában a játékosok a fent leírt módon a $\mathbf{q}^0 = (q_1^0, \ldots, q_n^0)$ készletvektorról alkudoznak. Amennyiben nem jutnak egyezségre, D_i nagyságú kárt szenvednek el ($D_i \geq 0$). A játékosoknak a megegyezésből származó nettó nyeresége

$$R_i(\mathbf{q}^0) \equiv R_i(q_i^0) \equiv u_i(q_i^*) - p(q_i^* - q_i^0) + D_i \equiv v_i(q_i^0) + D_i\,. \tag{3.3}$$

Mindez az első periódusra egy $(V, -\mathbf{D})$ alkujátékot definiál, ahol

$$V = \{\, \mathbf{v}(\mathbf{q}^0) = (v_1(q_1^0), \ldots, v_n(q_n^0)) \mid \sum q_i^0 = \overline{q},\; q_i^0 \geq 0 \,\}$$

a lehetséges kifizetésvektorok kompakt, konvex halmaza,[8]

$$-\mathbf{D} = (-D_1, \ldots, -D_n)$$

[8] V kompakt és konvex, hiszen q_i^0 $\forall i$-re a $[0, \overline{q}]$ szakaszról vesz fel értékeket, és $v_i(q_i^0)$ lineáris q_i^0-ban.

pedig a fenyegetési pont. V elemeinek definíciójából világos, hogy versenyző esetben az elosztások és az általuk megvalósított megoldások kölcsönösen egyértelmű viszonyban állnak egymással. Ha ugyanis $\varphi(V, -\mathbf{D}) = [\varphi_i(V, -\mathbf{D})]$ egy megoldás, akkor $\varphi(V, -\mathbf{D}) \in V$ miatt $\varphi_i(V, -\mathbf{D})$ fölírható a következő alakban:

$$\varphi_i(V, -\mathbf{D}) = u_i(q_i^*) - p(q_i^* - q_i^0) \quad i = 1, \dots, n\,.$$

Innen q_i^0 és $\varphi_i(V, -\mathbf{D})$ kölcsönösen egyértelműen meghatározhatók. Az alábbi lemma az EG és a Nash elosztások létezésével kapcsolatos.

3.1 Lemma. *A versenyző készletelosztási játékban legfeljebb egy EG elosztás létezik, továbbá mindig létezik pontosan egy Nash elosztás.*

Bizonyítás. A lemma első fele abból adódik, hogy a nyereségfüggvények lineárisak az indulókészletekben. A Nash elosztás egyértelmű létezése a Nash megoldás egyértelmű létezéséből és az elosztások és a megoldások egy-egyértelmű megfeleltetéséből következik. □

Mint az a következő állításból kiderül, utilitarista elosztásból végtelen sok van. Az állítás a közgazdaságtan híres Első jóléti tételének adaptációja.

3.2 Állítás. *A versenyző készletelosztási játékban minden elosztás utilitarista.*

Bizonyítás. Minden elosztás esetén $\sum v_i(q_i^0) = \sum u_i(q_i^*)$ teljesül. □

A 3.2. állítás szerint a piac tetszőleges indulókészletekből kiindulva hatékony (sőt, társadalmi összhasznot maximalizáló) eredményre vezet. Ez ugyanakkor azt is jelenti, hogy az utilitarista szempont nem segít a lehetséges elosztások közötti választásban. Annál inkább segítségünkre lehet a következő állítás.

3.3 Állítás. *Egy versenyző készletelosztási játékban a $\mathbf{q}^N$ Nash elosztásban minden $q_i^N > 0$ esetén teljesül, hogy*

$$|R(q_j^N) - R(q_i^N)| \leq |R(q_j^0) - R(q_i^N)| \quad \forall q_j^0\text{-ra}\,. \tag{3.4}$$

Továbbá, ha létezik EG elosztás, akkor az pontosan a Nash elosztás, vagyis ilyenkor

$$R(q_j^N) - R(q_i^N) = 0 \quad \forall (i,j)\text{-re}\,.$$

Bizonyítás. A Nash-megoldást a következő optimumfeladat adja:

$$\begin{aligned} \max_{q_1^0, \dots, q_n^0} \quad & \prod_{i=1}^{n} R_i(q_i^0) \\ \text{k. f.} \quad & \sum_{i=1}^{n} q_i^0 = \overline{q} \\ & q_i^0 \geq 0 \quad i = 1, \dots, n \end{aligned}$$

λ-val jelölve a korlát Lagrange-szorzóját, az elsőrendű feltételek a következők:

$$p\prod_{j\neq i} R_j(q_j^0) - \lambda \leq 0 \quad q_i^0 \geq 0 \text{ komplementaritással}, \quad i=1,\dots,n \tag{3.5}$$

$$\sum_{i=1}^{n} q_i^0 = \overline{q}\,.$$

(3.5)-öt átrendezve kapjuk, hogy

$$R(\mathbf{q}^N) := \frac{\prod_{j=1}^{n} R_j}{\lambda} \leq R_i(q_i^0) \quad q_i^0 \geq 0 \text{ komplementaritással}, \quad i=1,\dots,n\,.$$

Vagyis egyrészt

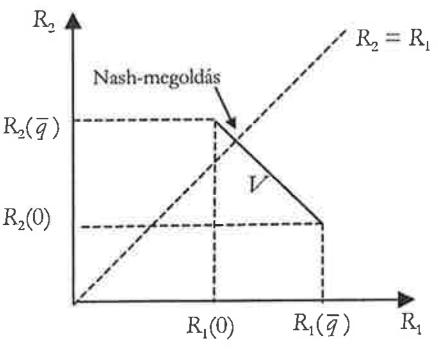
$$R_i(q_i^N) = R(\mathbf{q}^N)$$

minden i: $q_i^N > 0$ mellett, másrészt

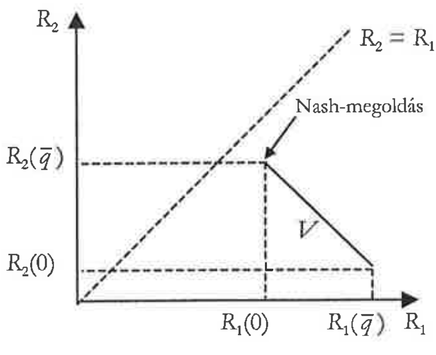
$$R_j(q_j^N) > R(\mathbf{q}^N)$$

esetén $q_j^N = 0$. Eszerint az $R(\mathbf{q}^N)$ értéknél nagyobb nyereségeket a Nash elosztás minimálisra csökkenti, vagyis (3.4) teljesül. Továbbá ha létezik EG elosztás, ez (3.5)-ből adódóan egyben a —3.1. lemma szerint egyértelmű— Nash elosztás is. □

A 3.3 állítás szerint a Nash-megoldás EG elosztásra vezet, amennyiben létezik ilyen. Ellenkező esetben biztosítja, hogy a pozitív indulókészletben részesülők nyereségéhez viszonyítva a játékosok nyereségének eltérése minimális legyen. A Nash elosztás e szerint egyfajta irigység-mentesség *(no-envy)* tulajdonsággal rendelkezik, ami számos alkalmazásban igen vonzóvá teheti a Nash megoldást a mesterséges piacok készletelosztási problémáiban. A Nash-megoldást $n = 2$-re a *3.1. ábra* szemlélteti.



(a) létezik EG elosztás



(b) nem létezik EG elosztás

3.1. ábra. Nash-megoldás a készletelosztási játékban

Készletelosztási játék tranzakciós költségek esetén

Áttérünk a készletelosztási játék vizsgálatára abban az esetben, ha a piacon tranzakciós költségek jelentkeznek. Emlékeztetőül, a piaci egyensúly feltételei tranzakciós költségek mellett a következők voltak:

$$u_i'(q_i) - T'(q_i - q_i^0) = p \quad i = 1, \dots, n \tag{2.2}$$

és

$$\sum_{i=1}^{n} q_i^0 = \overline{q} \, .$$

A másodrendű feltételek teljesülése esetén lokálisan léteznek az alábbi folytonos és differenciálható függvények:

$$q_i^* = q_i(p, q_i^0) \quad i = 1, \dots, n,$$

és

$$p = p(\mathbf{q}^0) \, .$$

Mint a 2. szakaszban láttuk, a szereplők optimális fogyasztása az ár mellett indulókészletüknek is függvénye lesz, ebből adódóan az ár is a $\mathbf{q}^0$ szétosztástól függően fog alakulni. Tranzakciós költségek esetén a célfüggvény optimális értékét ez alapján

$$v_i(\mathbf{q}^0) \equiv u_i(q_i(p(\mathbf{q}^0), q_i^0)) - p(\mathbf{q}^0) \cdot [q_i(p(\mathbf{q}^0), q_i^0) - q_i^0] - T(q_i(p(\mathbf{q}^0), q_i^0) - q_i^0)$$

adja, az R_i nyereségfüggvények definíciója pedig

$$R_i(\mathbf{q}^0) \equiv u_i(q_i(p(\mathbf{q}^0), q_i^0)) - p(\mathbf{q}^0) \cdot [q_i(p(\mathbf{q}^0), q_i^0) - q_i^0] - T(q_i(p(\mathbf{q}^0), q_i^0) - q_i^0) + D_i \, .$$

A $(V, -\mathbf{D})$ alkujátékban

$$V = \operatorname{con}\left(\{\, \mathbf{v}(\mathbf{q}^0) = (v_1(\mathbf{q}^0), \dots, v_n(\mathbf{q}^0)) \mid \sum q_i^0 = \overline{q},\ q_i^0 \geq 0 \,\}\right)$$

a lehetséges kifizetések halmazának konvex burka.

Az alábbi állítás Stavins (1995) eredményének megfelelője a készletelosztási játékban.

3.4 Állítás. *Monoton tranzakciós költségek esetén a készletelosztási játékban egy $\mathbf{q}^0$ elosztás pontosan akkor utilitarista, ha $q_i^0 = q_i^*$ teljesül minden i-re.*

Bizonyítás. Pontosan ez biztosítja, hogy (2.2) az $u_i'(q_i) = p$ optimumfeltételre egyszerűsödjön. □

A 3.4. állítás meglehetősen intuitív. Monoton tranzakciós határköltségek következtében egyéni és társadalmi szempontból egyaránt költségessé válik a piacon való részvétel. Minden egyes lezajló tranzakció erőforrások elpazarlásával jár. A társadalmi optimumot egy olyan elosztás biztosítja, amely mellett a játékosoknak nem kell részt venniük a piacon, hanem pusztán indulókészletük elfogyasztásával egyensúlyi helyzetbe juthatnak.

Miközben versenyző esetben minden lehetséges elosztás utilitarista volt, addig a 3.4. állításból adódóan monoton tranzakciós költségek esetén pontosan egy utilitarista elosztásunk lesz.[9] Az EG és a Nash elosztások egyértelműsége ezzel szemben nem biztosított. Ez abból adódik, hogy nincs többé egy-egyértelmű megfeleltetés az elosztások és az általuk megvalósított megoldás között.

A továbbiakhoz szükség lesz a Nash illetve az utilitarista elosztásokat jellemző következő két lemmára.

3.5 Lemma. *Tranzakciós költségek esetén a készletelosztási játék Nash-megoldását meghatározó feltételek a következők:*

$$-\frac{dp}{dq_i^0}\sum_{j=1}^{n}\frac{q_j-q_j^0}{R_j}+\frac{p+T'(t_i)}{R_i}\leq\frac{\lambda}{R_1\cdot\ldots\cdot R_n} \qquad (3.6)$$

$$q_i^0\geq 0 \text{ komplementaritással, } i=1,\ldots,n$$

$$\sum_{i=1}^{n}q_i^0=\overline{q}\,, \qquad (3.7)$$

ahol λ a (3.7) korláthoz tartozó Lagrange-szorzó.

Bizonyítás. Lásd A függelék. □

3.6 Lemma. *Monoton tranzakciós költségek esetén $T'(t_j)=T'(t_k)\ \forall j,k\in N$ pontosan akkor teljesül, ha az elosztás utilitarista.*

Bizonyítás. Elégségesség. A 3.4. állításból és a tranzakciós költségfüggvény tulajdonságaiból utilitarista elosztás esetén $T'(t_i)=T'(0)=0$ minden i-re. *Szükségesség.* Tegyük fel, hogy $T'(t_j)=T'(t_k)\ \forall j,k\in N$, miközben az elosztás nem utilitarista. Ekkor szükségképpen létezik $t_u>0$ és $t_v<0$ is, de a monoton tranzakciós költségek definíciójából ekkor $T'(t_u)>0$ és $T'(t_v)<0$, vagyis a tranzakciós határköltségek nem egyezhetnek meg. □

Az iménti lemmák segítségével megvizsgálhatjuk, hogyan alakul a vizsgált elosztások kapcsolata tranzakciós költségek esetén.

3.7 Állítás. *Monoton tranzakciós költségek esetén az alábbi két eset közül pontosan az egyik teljesül:*

(a) *Létezik egy és csakis egy utilitarista és EG Nash-elosztás.*

(b) *Egy adott elosztás vagy utilitarista, vagy EG, vagy Nash-elosztás, vagy ezek egyike sem.*

Bizonyítás. Megmutatom, hogy ha egy elosztás az állításban említett három típus közül kettőbe tartozik, akkor a harmadikba is tartozik.

[9] Az utilitarista elosztás egyértelmű létezését a versenyzői egyensúly egyértelmű léte biztosítja.

(i) *Az utilitarista Nash-elosztás EG elosztás.* Ekkor, mivel az utilitarista elosztásban minden indulókészlet szigorúan pozitív, (3.6) minden i-re egyenlőségre teljesül. Az egyenletek a 3.4. állítás és a 3.6. lemma értelmében a következő alakra egyszerűsödnek:

$$\frac{(p+T')R_1\cdot\ldots\cdot R_n}{\lambda}=R_i \quad \forall i\text{-re},$$

vagyis EG elosztást kapunk.[10]

(ii) *Az utilitarista EG elosztás Nash-elosztás.* $q_j^*=q_j^0$ és $R_j=R_k$ $\forall(j,k)$-ra kielégíti (3.6)-ot.

(iii) *Az EG Nash-elosztás utilitarista.* (3.6)-ot ekkor kielégíti $q_j^*=q_j^0$ (amely mellett a 3.5. lemma szerint $T'(t_j)=T'(t_k)$ teljesül minden (j,k) párra). □

Megjegyzés. *Monoton tranzakciós költségek esetén pontosan akkor létezik egyenlő nyereségeket biztosító Nash-elosztás, ha a* $\mathbf{q}^0$ *utilitarista elosztásnál* $\forall j,k$ *indexpárra*

$$-c_j(q_j(\mathbf{q}^0))+D_j=-c_k(q_k(\mathbf{q}^0))+D_k\,.$$

Hátra van még az állandó tranzakciós költségeket tartalmazó eset jellemzése, amit a következő egyszerű állítással tehetünk meg.

3.8 Állítás. *A 3.2 és 3.3 állítások állandó tranzakciós költségek esetén is fennállnak.*

Bizonyítás. Nyilvánvaló. □

A fentiek értelmében a Nash elosztás versenyző esetben tapasztalt kedvező tulajdonságai tranzakciós költségek mellett csak speciális esetekben érvényesülnek. Egyrészt állandó tranzakciós költségek esetén mindig teljesül, hogy a Nash-megoldás axiómái EG elosztást eredményeznek. Másrészt monoton költségek mellett, ha a speciális paraméter-együttes következtében az utilitarista elosztás éppen egyenlő nyereségeket ad, akkor ez egyben Nash elosztás is. Általános esetben azonban nem várhatjuk, hogy a Nash elosztás rendelkezzen az EG és az utilitarista elosztások vonzó tulajdonságaival is. Ilyenkor választani kell a nyereségek egyenlősége, a Nash-megoldás stabilitási és igazságossági axiómái, és az utilitarista hatékonyság között. Valamennyi szempont együttes kielégítésére nincs lehetőség.

Készletelosztási játék ármeghatározó játékossal

Végül az ármeghatározó játékost tartalmazó piacmodell esetén jellemzem a készletelosztási játékban vizsgált elosztásokat. Tegyük föl, miként a 2. szakasz harmadik modelljében, hogy az 1-gyel indexelt játékosnak hatalmában

[10] Mivel a Nash és az utilitarista elosztások egzisztenciája biztosított, ilyenkor EG elosztás is biztosan létezik.

áll ármeghatározóként viselkedni a piacon. Mint láttuk, elsőrendű feltétele ilyenkor

$$(u_1'(q_1^*) - p) \cdot \sum_{i=2}^{n} q_i'(p) + \left(\overline{q} - \sum_{i=2}^{n} q_i(p) - q_1^0\right) = 0 , \tag{3.8}$$

ahol $q_i(p)$ a többi játékos (2.1) elsőrendű feltételéből adódó keresleti függvény. A másodrendű feltételek teljesülése esetén lokálisan létezik a $p = p(q_1^0)$ folytonos és differenciálható, monoton növekvő függvény.[11] Az első időszakban a játékosok most is egy alku során osztják el maguk között a $\overline{q}$ nagyságú összkészletet, és most is D_i nagyságú kárral jár, ha nem sikerül megegyezniük. Ármeghatározó játékos esetén a következő módon alakul a játékosok nettó nyeresége az alkuból:

$$R_i(\mathbf{q}^0) \equiv R_i(q_1^0, q_i^0) \equiv u_i(q_i(p(q_1^0))) - p(q_1^0)[q_i(p(q_1^0)) - q_i^0] + D_i \, . \tag{3.9}$$

A Nash-megoldást az alábbi lemmával jellemezhetjük.

3.9 Lemma. *Ármeghatározó játékos esetén a készletelosztási játék Nash-megoldását a következő feltételek határozzák meg.*

$$R(\mathbf{q}^N) := \frac{p \prod_{j=1}^{n} R_j}{\lambda} \leq R_i(q_i^0) \quad q_i^0 \geq 0 \text{ komplementaritással}, i = 2, \ldots, n \tag{3.10}$$

$$\frac{p}{R_1} - \frac{dp}{dq_1^0} \sum_{j=2}^{n} \frac{q_j - q_j^0}{R_j} \leq \frac{p}{R(\mathbf{q}^N)} \quad q_1^0 \geq 0 \text{ komplementaritással} \tag{3.11}$$

$$\sum_{i=1}^{n} q_i^0 = \overline{q} , \tag{3.12}$$

ahol λ a (3.12) korláthoz tartozó Lagrange-szorzó.

Bizonyítás. Lásd B függelék. □

A következő állítás Hahn (1984) eredményének felel meg.

3.10 Állítás. *Ármeghatározó játékos mellett egy elosztás pontosan akkor utilitarista, ha teljesül, hogy $q_1^0 = q_1^*$.*

Bizonyítás. Pontosan ez biztosítja, hogy (2.3) $u_1' = p$ alakúra egyszerűsödjön. □

Miként az imént tranzakciós költségek esetén, úgy most is igen intuitív eredményt kaptunk. Az ármeghatározó játékos minden egyes piaci tranzakciója társadalmi veszteséget okoz, hiszen hasznát maximalizálva túl alacsony vagy túl magas árat fog kiszabni az egyes jószágegységekért. A társadalmi optimumot az biztosíthatja, ha az ármeghatározó játékos egyensúlyban nem vesz részt a piacon, hanem pusztán indulókészlete elfogyasztásával jut optimális helyzetbe. Belátható, hogy minél távolabb esik a kritikus játékos

[11] (3.8)-ból $\partial p / \partial q_1^0 = \left[(u_1' - p) \sum q_i'' - \sum (q_i')^2 u_1'' - 2 \sum q_i'\right]^{-1} > 0$, ha teljesül a másodrendű feltétel.

indulókészlete az egyensúlyi fogyasztástól, annál messzebb kerülünk a társadalmi optimumtól, vagyis annál nagyobb lesz a társadalmi haszonveszteség.

3.11 Állítás. *Tegyük föl, hogy az utilitarista elosztásban minden indulókészlet pozitív. Ekkor ármeghatározó játékos esetén az alábbi két eset közül pontosan az egyik teljesül:*

(a) *Létezik egy és csakis egy utilitarista és EG Nash-elosztás.*

(b) *Egy adott elosztás vagy utilitarista, vagy EG, vagy Nash elosztás, vagy ezek egyike sem.*

Bizonyítás. Ugyanúgy járunk el, mint a 3.7. állítás bizonyításánál.

(i) *Az utilitarista Nash-elosztás EG elosztás.* Ekkor, mivel az utilitarista elosztásban minden indulókészlet pozitív, a (3.10) és (3.11) kifejezések egyenlőségre teljesülnek. Az $R_i = R(\mathbf{q}^0)$ $i = 2, \ldots, n$ egyenleteket behelyettesítve (3.11)-be, és kihasználva a 3.10 állítást, kapjuk, hogy

$$\frac{p}{R_1} = \frac{p}{R(\mathbf{q}^N)} ,$$

vagyis az elosztás valóban egyenlő nyereségeket biztosít.

(ii) *Az utilitarista EG elosztás Nash-elosztás.* $q_1^* = q_1^0$ és $R_j = R_k$ $\forall(j, k)$-ra kielégíti (3.10) és (3.11)-et.

(iii) *Az EG Nash-elosztás utilitarista.* A (3.10) egyenleteket behelyettesítve (3.11)-be

$$\frac{dp}{dq_1^0} \frac{q_1 - q_1^0}{R(\mathbf{q}^N)} = 0$$

adódik. Mivel $p'(q_1^0) > 0$ (lásd 11. lábjegyzet), a Nash-megoldásban $q_j^* = q_j^0$ szükségképpen. □

Megjegyzés. *Ármeghatározó játékos esetén pontosan akkor létezik EG Nash-elosztás, ha az utilitarista elosztás* $q_1^0 = q_1^*$ *komponenséhez létezik* $(q_2^0, \ldots, q_n^0) \mid \sum q_i^0 = \overline{q}$, *amelyre*

$$u_1(q_1^*) + D_1 = u_i(q_i(p(q_1^*))) - p(q_1^*)[q_i(p(q_1^*)) - q_i^0] + D_i \quad i = 2, \ldots, n\,.$$

A 3.11 állítás szerint tehát akárcsak monoton tranzakciós költségek esetén, ugyanúgy ármeghatározó játékos mellett is két eset lehetséges. Speciális helyzetben egyértelműen létezik egy elosztás, amely mindhárom általunk vizsgált elosztás vonzerejét egyesíti. Általában azonban ha kiválasztunk egy elosztási szabályt a három közül, ezzel egyúttal le kell mondanunk a másik két szabály előnyeiről.

Ebben a szakaszban bevezettem a „készletelosztási játékot", amely lehetővé tette különböző elosztási elvek vizsgálatát a mesterséges piacokon. A versenyző, a tranzakciós költségeket tartalmazó, majd az ármeghatározó játékossal zajló játékokat elemezve azt találtam, hogy a vizsgált elosztások tulajdonságai különböznek a piac eltérő forgatókönyveiben. Láttuk például, hogy a Nash (illetve az EG) elosztás igen vonzó jellemzőkkel bír versenyző piacon, ám ezek csak speciális helyzetben teljesülnek tökéletlen piacok esetén. Mindez a piac kialakításának két szakasza —az elosztás és a piac működése— közti összefüggésekre világít rá. A következő részben megvizsgálom, hogyan függ e két szakasztól egy még korábbi fázis: a szétosztandó összes mennyiség meghatározása. Mielőtt azonban továbbmennénk, tekintsünk egy példát az eddigiek szemléltetésére.

Egy példa

Az elmondottak szemléltetésére tekintsünk egy kétszereplős szennyezési piacot, ahol a q jószág a szennyezés (illetve az erre vonatkozó jog), a hasznosság pedig a szennyezéssel járó elhárítási költségek ellentettje: $u_i(q_i) = -c_i(q_i)$. Legyenek ezek a költségfüggvények kvadratikusak:

$$c_i(q_i) = (q_i - a_i)^2 \quad i = 1, 2\ ,$$

ahol a_i pozitív paraméter.[12] Ekkor a számunkra releváns fogyasztások (szennyezések) a $(0, a_i)$ intervallumba esnek. Az i-dik játékos feladata versenyzői piacon

$$\max_{q_i} \; -(q_i - a_i)^2 - p(q_i - q_i^0)\ ,$$

a játékosok szennyezésére és a $q_1 + q_2 = \overline{q}$-t biztosító piactisztító p árra

$$q_i^* = (a_i - a_j + \overline{q})/2 \quad i, j = 1, 2;\ i \neq j$$
$$p = a_1 + a_2 - \overline{q}$$

adódik. Ezek biztosítják, hogy a $\overline{q}$ összmennyiség megvalósításának összes haszna maximális (összköltsége minimális) legyen, méghozzá

$$c_1(q_1^*) + c_2(q_2^*) = 1/2(a_1 + a_2 - \overline{q})^2$$

az indulókészletektől függetlenül.

A játékosok nettó nyereségfüggvénye:

$$R_i(q_i^0) = -1/4(A - \overline{q})^2 - (A - \overline{q})(a_i - a_j + \overline{q} - 2q_i^0)/2 + D_i \quad i = 1, 2$$

(ahol $A = a_1 + a_2$). A készletelosztási játék Nash-megoldását (pozitív indulókészleteket feltételezve) a következő feladat adja:

$$\max_{q_1^0, q_2^0} R_1(q_1^0) R_2(q_2^0)$$

$$\text{k. f.}\ \ q_1^0 + q_2^0 = \overline{q}\ ,$$

a Nash elosztás innen

$$q_i^0 = \frac{a_i - a_j + \overline{q}}{2} + \frac{D_j - D_i}{2(A - \overline{q})} \quad i, j = 1, 2 \quad i \neq j\ .$$

[12] $a_i = q_i^u$, értelmezése pedig: a szabályozás hiányában a játékosok számára egyénileg optimális szennyezés.

Egyszerű behelyettesítéssel ellenőrizhető, hogy a Nash elosztás valóban egyenlő nettó nyereségeket biztosít a játékosoknak.

Vezessünk most be egyszerű (monoton) tranzakciós költségfüggvényeket, a következő módon:

$$T_i(t_i) = 1/2(q_i - q_i^0)^2 \quad i = 1, 2 \, .$$

Az i-edik játékos feladata ekkor

$$\max_{q_i} \; -(q_i - a_i)^2 - p(q_i - q_i^0) - 1/2(q_i - q_i^0)^2 \, ,$$

a játékosok szennyezésére és a piactisztító p árra pedig

$$q_i^* = (a_i - a_j + q_i^0 + \overline{q})/3 \quad i = 1, 2; \; i \neq j$$
$$p = a_1 + a_2 - \overline{q}$$

adódik. Látható, hogy míg ebben a speciális esetben a jogok egyensúlyi ára független az elosztástól,[13] addig az optimális szennyezéseket befolyásolják az indulókészletek. A szennyezéscsökkentés összes költsége az elosztástól függően (eltekintve maguktól a tranzakciós költségektől)

$$c_1(q_1^*) + c_2(q_2^*) = 1/9(-a_2 - 2a_1 + q_1^0 + \overline{q})^2 + 1/9(-a_1 - 2a_2 + q_2^0 + \overline{q})^2 \, .$$

Ebben a modellben

$$p(\mathbf{q}^0) = p = A - \overline{q}$$

és

$$q_i(p(\mathbf{q}^0), q_i^0) = q_i(q_i^0) = (2a_i + q_i^0 - A + \overline{q})/3 \quad i = 1, 2 \, ,$$

ahonnan a játékosok nettó nyereségfüggvénye

$$R_i(\mathbf{q}^0) = R_i(q_i^0) = -\frac{(-2a_i - a_j + q_i^0 + \overline{q})^2}{9} - \frac{(A - \overline{q})(a_i - a_j - 2q_i^0 + \overline{q})}{3} - \frac{(a_i - a_j - 2q_i^0 + \overline{q})^2}{18} \, .$$

Az EG elosztást a következő egyenletek határozzák meg:

$$R_1(q_1^0) = R_2(q_2^0) \qquad \text{és} \qquad q_1^0 + q_2^0 = \overline{q} \, .$$

Innen

$$q_i^0 = \frac{a_i - a_j + \overline{q}}{2} + \frac{D_j - D_i}{2(A - \overline{q})} \quad i, j = 1, 2; i \neq j \, .$$

Ebben a példában az EG elosztás egyértelmű, és megegyezik a tökéletes piac EG elosztásával. Ugyanakkor az EG elosztás mellett adódó nettó nyereség különbözik a versenyző esetben kapott eredménytől:

$$R = \frac{D_1 + D_2}{2} - \frac{(A - \overline{q})^2}{4} - \frac{(D_1 - D_2)^2}{12(A - \overline{q})^2} \, .$$

[13] Ez minden olyan esetben teljesül, amikor a játékosok elhárítási és tranzakciós határköltség függvényeinek együttes meredeksége megegyezik.

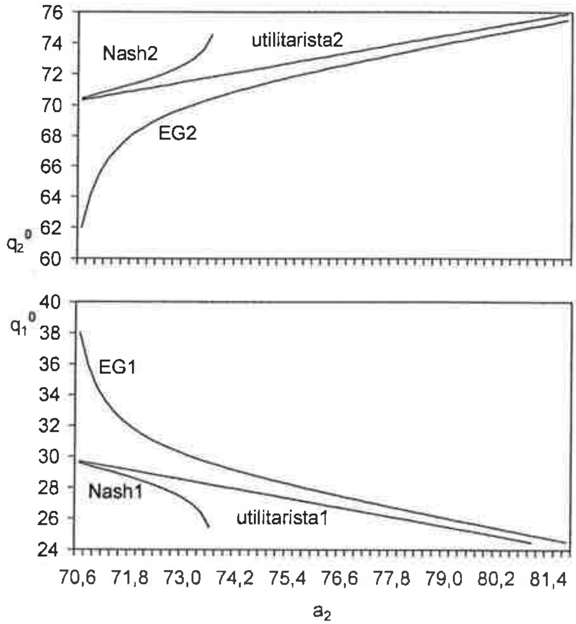
Mint látható, a nettó nyereségek kisebbek lesznek, mint versenyző esetben – ahogyan ezt el is vártuk. Az utilitarista elosztásban

$$q_i^0 = a_i - (A - \overline{q})/2 \quad i = 1, 2 ,$$

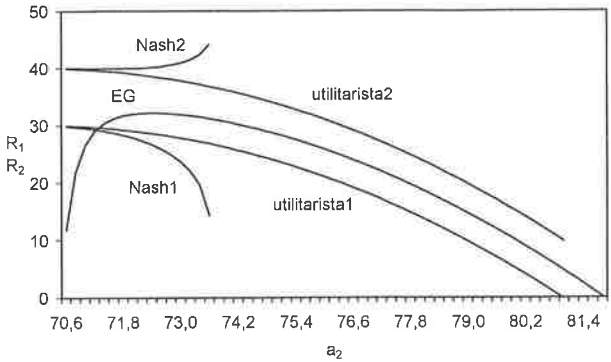
amely megegyezik a játékosok versenyző piacon kapott optimális szennyezéseivel. A Nash-megoldás feltételei egy q_1^0-ban harmadfokú egyenletet eredményeznek – mint láttuk, tranzakciós költségek esetén a Nash-elosztás általában nem egyértelmű.

Tekintsünk egy számpéldát. Legyen $\overline{q} = 100$, $D_1 = 30$, $D_2 = 40$ és $a_1 = 30$. Ekkor ahhoz, hogy létezzen a modell feltételei szerint értelmes EG elosztás (nem-negatív változók, $R_i > 0$, $c_i'(q_i) < 0$), $a_2 \in (70.49, 81.83)$ kell. A három különböző elosztás mellett e paraméterek esetén adódó indulókészleteket és nettó nyereségeket mutatják a *3.2 – 3.3. ábrák.*

Mint az ábrákból látható, bizonyos paraméter-értékek esetén nem létezik Nash-elosztás (komplex indulókészletek adódnak), és a_2 növelésével egy idő után az utilitarista elosztás is értelmetlenné válik (az első játékos nyeresége negatív lesz). Látszik továbbá, hogy ebben a példában a 3.7. állítás (b) része teljesül: vizsgált elosztásaink közül semelyik kettő nem esik egybe. Ha ilyenkor az EG elosztás mellett döntünk, ezzel lemondunk az utilitarista hatékonyságról és a Nash-féle axiómákról is. A nyereségek közötti különbség nem elhanyagolható az egyes elosztások esetén, vagyis a játékosok szempontjából sem lényegtelen kérdés, hogy milyen elveket követ majd az elosztás.



3.2. ábra. Indulókészletek tranzakciós költségek esetén
("Nash1" = q_1^0 Nash-elosztás mellett, "Nash2" = q_2^0 Nash-elosztás mellett stb.)



3.3. ábra. Nettó nyereségek tranzakciós költségek esetén
("Nash1" = R_1 Nash-elosztás mellett, "Nash2" = R_2 Nash-elosztás mellett stb.)

4 Az összmennyiség meghatározása

Az eddigiekben külső adottságként kezeltem a $\overline{q}$ összes mennyiséget. Mesterséges piacok esetén azonban az egyik fontos kérdés éppen ennek a nagysága. Ebben a szakaszban ezért megvizsgálom, hogyan határozza meg a $\overline{q}$ összmennyiséget egy társadalmi összhasznot maximalizáló társadalmi tervező a három különböző piacmodell mellett. Ehhez felteszem, hogy egy $\overline{q}$ összmennyiséget biztosító piac felállításának költsége $d_i(\overline{q})$ az i-dik játékos számára, ahol d_i kétszer folytonosan differenciálható, monoton növekvő, szigorúan konvex függvény. A környezetszabályozás témakörében d_i tipikusan a környezetszennyezés következtében a játékosokat érő kár, egy mezőgazdasági kvótarendszernél pl. a nagyobb kínálat okozta árcsökkenésből származó profitkiesés.

A társadalmilag kívánatos összmennyiség versenyzői piacon

Mint azt a 2. szakaszban láttuk, versenyző piacon a játékosok egyénileg optimális fogyasztását a p ár függvényében $q_i(p)$ alakban kaptuk. Ha most kihasználjuk a

$$\sum q_i(p) = \overline{q}$$

egyenletet, úgy a megfelelő másodrendű feltételek teljesülése esetén lokálisan létezik a piactisztító árat a $\overline{q}$ összmennyiség függvényében meghatározó $p(\overline{q})$ folytonos és differenciálható függvény. Ezek alapján az új modellben a játékosok hasznosságát az összmennyiség függvényében így írhatjuk:

$$\begin{aligned} U_i(\overline{q}) &= u_i(q_i^*) - d_i(\overline{q}) \\ \text{ahol } \; q_i^* &= q_i(p(\overline{q})) \,. \end{aligned} \tag{4.1}$$

A társadalmi tervező feladata ezután a következő:

$$\max_{\overline{q}} \sum_{i=1}^{n} \Big[u_i(q_i(p(\overline{q}))) - d_i(\overline{q}) \Big] \,.$$

A minimum szükséges és elégséges feltétele[14]

$$\sum_{i=1}^{n}\left[u_i'(q_i)\frac{dq_i(p)}{d\overline{q}} - d_i'(\overline{q})\right] = 0\,. \tag{4.2}$$

Kihasználva, hogy a szennyezési piacon egyensúlyban $u_i'(q_i) = p$ teljesül $\forall i$-re, továbbá $\sum q_i(p) = \overline{q}$, kapjuk, hogy

$$\sum_j d_j'(\overline{q}) = p = u_i'(q_i^*) \quad i = 1, \ldots, n\,. \tag{4.3}$$

A (4.3) kifejezés bal oldalán az i-edik játékos egységnyi fogyasztás-növekedésének társadalmi határköltsége, míg a jobb oldalon ennek határhaszna áll. Az optimális $\overline{q}$ összmennyiség kiegyenlíti a társadalmi határköltségeket és határhasznokat. A piactisztító p ár megegyezik mindkettővel. A társadalmi optimum meghatározásában sehol nem játszik szerepet a jogok elosztása. Mint láttuk, versenyzői piacon az elosztás nem befolyásolja a piac kimenetét, ezért ilyenkor az optimális összmennyiség tetszőleges elosztási szabály esetén ugyanaz lesz. Ez nem más, mint a híres Coase-tétel:

4.1 Állítás (Coase, 1963). *(a) Az összes szennyezés optimális meghatározását követően a szennyezési jogok tökéletes versenyzői piaca implementálja a társadalmi optimumot.*

(b) Az összmennyiség optimuma független az elosztási szabálytól.

A 4.1. állítás értelmében, ha versenyző piacra számítunk, egymástól függetlenül kezelhető az összmennyiség meghatározása és annak szétosztása. A következő szakaszokban megvizsgálom, hogy vajon ez az eredmény tökéletlen piacok esetén is áll-e.

A társadalmilag kívánatos összmennyiség tranzakciós költségek esetén

Ebben a szakaszban megvizsgálom a társadalmilag kívánatos összmennyiséget, amennyiben a játékosok tranzakciós költségekkel szembesülnek a piacon. Megmutatom, hogy az optimális összmennyiség általában nem független az ezt követően megvalósuló elosztási szabálytól. Mindenekelőtt tegyük fel, hogy létezik egy folytonos és differenciálható általános elosztás-függvény, $\mathbf{q}^0(\overline{q}) = [q_1^0(\overline{q}), \ldots, q_n^0(\overline{q})]$ alakban, amely adott összmennyiséghez megadja az egyes játékosok részesedését, egy rögzített elosztási szabály szerint. Idézzük fel ezek után a tranzakciós költségek mellett zajló készletelosztási játékot. Mint láttuk, az optimális fogyasztások az ár mellett az indulókészletektől is függtek, $q_i^* = q_i(p, q_i^0)$, míg az ár nem csupán az összes jószág, hanem az elosztás függvénye volt: $p = p(\mathbf{q}^0, \overline{q})$. Ezek alapján a (4.1) hasznosság most

[14]Feltesszük, hogy $0 < \overline{q} < \sum q_i^u$, ahol $u_i'(q_i^u) = 0$ $\forall i$-re.

a következő alakot ölti:

$$
\begin{aligned}
U_i(\overline{q}) = u_i(q_i^*) - T(q_i^* - q_i^0) - d_i(\overline{q}) \\
\text{ahol}\ \ q_i^0 = q_i^0(\overline{q}) \\
p = p(\mathbf{q}^0, \overline{q}) \\
q_i^* = q_i(p, q_i^0)\ .
\end{aligned}
$$

A társadalmi haszonmaximumot ennek megfelelően a következő feladat megoldása adja:

$$
\begin{aligned}
\max_{\overline{q}} \quad & \sum_{i=1}^{n} \Big[u_i(q_i^*) - T(q_i^* - q_i^0) - d_i(\overline{q}) \Big] \\
\text{k. f.} \quad & \mathbf{q}^0 = \mathbf{q}^0(\overline{q}) \\
& p = p(\mathbf{q}^0, \overline{q}) \\
& q_i^* = q_i(p, q_i^0) \quad i = 1, \ldots, n\ .
\end{aligned}
$$

Az elsőrendű feltétel $0 < \overline{q} < \sum q_i^u$ esetén:

$$
\sum_{i=1}^{n} \left[\frac{du_i}{dq_i} \frac{dq_i}{d\overline{q}} - \frac{\partial T}{\partial q_i} \frac{dq_i}{d\overline{q}} - \frac{\partial T}{\partial q_i^0} \frac{dq_i^0}{d\overline{q}} - d_i'(\overline{q}) \right] = 0\ . \tag{4.4}
$$

Kihasználva, hogy a játékosok egyéni optimumában (2.2) szerint

$$
\frac{du_i}{dq_i} - \frac{\partial T}{\partial q_i} = p\ ,
$$

és átrendezve, (4.4) az alábbi egyenletre egyszerűsödik:

$$
\sum_{i=1}^{n} d_i'(\overline{q}) + \sum_{i=1}^{n} \frac{\partial T}{\partial q_i^0} \frac{dq_i^0}{d\overline{q}} = p(\mathbf{q}^0, \overline{q})\ . \tag{4.5}
$$

A $\overline{q}$ összmennyiség növelésekor tranzakciós költségek esetén is találkozunk a versenyző esetben tapasztalt hatásokkal: a $\sum d_i$ tagon keresztül jelentkező költséggel és az áron keresztül jelentkező haszonnal. Megjelenik azonban egy új hatás is: az összmennyiség hat a tranzakciós költségekre, méghozzá az elosztási szabálytól függően. Ha az adott szabály mellett a megnövekedett $\overline{q}$ újraosztásával növekszik a kezdőkészletek és a q_i^* egyensúlyi fogyasztások közti távolság, akkor a tranzakciós költségek emelkedni fognak. Ha ez a távolság csökken, vagyis az összmennyiség növelése csökkenti a piaci tranzakciók volumenét, ezáltal a tranzakciós költségek is csökkenni fognak. Minthogy ezek a hatások az elosztási szabálytól függnek, az optimális összmennyiség meghatározása tranzakciós költségek esetén általában nem lehet független az elosztási szabálytól.

Piactervezési szempontból érdeklődésre tarthatnak számot azok a speciális esetek, amikor az optimális $\overline{q}$ nem változik az elosztás függvényében, vagy esetleg a piac működési feltételeitől is független. Ha nem ismerjük pontosan a kialakuló piac jellegzetességeit, sokat nyerhetünk azzal, ha egy téves várakozás (pl. a tranzakciós költségek jellegére vonatkozóan) nem rontja el a folyamat legelején meghatározott összmennyiség optimális voltát. Az alábbi állítás két ilyen esetre mutat rá.

4.2 Állítás. *Tranzakciós költségek mellett az optimális összmennyiség pontosan akkor egyezik meg a versenyzői optimummal, ha az alábbi feltételek közül legalább az egyik teljesül:*

(a) az elosztás utilitarista;

(b) állandóak a tranzakciós költségek.

Bizonyítás. Kihasználjuk a (2.2) egyensúlyi feltételeket, amikből

$$p - u_i'(q_i^*) = -T'(q_i^* - q_i^0) = \partial T(q_i^* - q_i^0)/\partial q_i^0 \,.$$

Ezt behelyettesítve (4.5)-be a következőt kapjuk:

$$\sum_{i=1}^{n} d_i'(\overline{q}) - \sum_{i=1}^{n} \frac{du_i(q_i^*)}{dq_i^*} \frac{dq_i^0}{d\overline{q}} = 0 \,.$$

Összehasonlítva (4.3)-mal, a két optimális összmennyiség egyenlőségének feltétele, hogy

$$\sum_{i=1}^{n} \frac{du_i(q_i^*)}{dq_i^*} \frac{dq_i^0}{d\overline{q}} = u_j'(q_j^*)$$

teljesüljön $\forall j \in N$-re. (2.2) szerint ez pontosan akkor áll fenn, ha

$$T'(q_u^* - q_u^0) = T'(q_v^* - q_v^0) \quad \forall u, v \in N \,,$$

aminek a 3.6. lemma értelmében szükséges és elégséges feltétele, hogy (a) és (b) közül legalább az egyik teljesüljön. □

Ha versenyző piacot várva határoztuk meg a társadalmilag kívánatos összmennyiséget, az csak akkor marad optimális tranzakciós költségek fölbukkanása esetén, ha ezek a költségek állandóak és/vagy az elosztás utilitarista volt. Ellenkező esetben $\overline{q}$ már nem biztosítja a jólét-maximumot.

A társadalmilag kívánatos összmennyiség ármeghatározó játékos esetén

Most megvizsgálom, hogyan határozódik meg a társadalmilag optimális összmennyiség, amennyiben ismert, hogy az egyik játékos ármeghatározóként fog viselkedni a piacon. Megmutatom, hogy a versenyző esettel ellentétben általában itt sem lesz közömbös a jogok elosztása.

Az 1. szakasz harmadik modelljében az 1-gyel indexelt játékosnak hatalmában áll megválasztani a számára optimális piactisztító árat, amely így a $p = p(q_1^0, \overline{q})$ alakban írható. Az ár az összmennyiségen kívül az első játékos indulókészletétől is függ. Az $i = 2, \ldots, n$ játékosok egyensúlyi szennyezését ezek után a $q_i^* = q_i(p(q_1^0, \overline{q}))$ módon kaptuk, míg az első játékosra a

$$q_i + \sum_{i=2}^{n} q_i(p) = \overline{q} \tag{4.6}$$

feltételből $q_1^* = q_1(p(q_1^0, \overline{q}), \overline{q})$ adódott. A társadalmi feladat ezek után a következő alakba írható:

$$\begin{aligned} \max_{\overline{q}} \quad & \sum_{i=1}^{n} \Big[u_i(q_i) - d_i(\overline{q})\Big] \\ \text{k. f.} \quad & \mathbf{q}^0 = \mathbf{q}^0(\overline{q}) \\ & p = p(q_1^0, \overline{q}) \\ & q_i^* = q_i(p) \quad i = 2, \ldots, n \\ & q_1^* = q_1(p, \overline{q}) \,, \end{aligned}$$

ahol $\mathbf{q}^0(\overline{q})$ ismét egy tetszőleges egyértelmű elosztást jelöl. Az elsőrendű feltétel

$$\frac{du_1(q_1^*)}{dq_1} \left[\frac{\partial q_1}{\partial \overline{q}} + \frac{\partial q_1}{\partial p} \frac{dp}{d\overline{q}} \right] + \sum_{i=2}^{n} \frac{du_i(q_i^*)}{dq_i} \frac{dq_i}{dp} \frac{dp}{d\overline{q}} = \sum_{i=1}^{n} d_i'(\overline{q}) \,. \tag{4.7}$$

Kihasználva a (2.1) egyensúlyi feltételeket, valamint hogy (4.6)-ből $\partial q_1 / \partial \overline{q} = 1$ adódik, (4.7) a következő alakban írható:

$$\frac{du_1(q_1^*)}{dq_1} - (u_1' - p) \sum_{i=2}^{n} \frac{dq_i}{dp} \frac{dp}{d\overline{q}} = \sum_{i=1}^{n} d_i'(\overline{q}) \,.$$

Az ármeghatározó játékos (2.3) optimumfeltételének segítségével tovább egyszerűsíthetünk:

$$\frac{du_1(q_1^*)}{dq_1} + (q_1 - q_1^0) \frac{dp}{d\overline{q}} = \sum_{i=1}^{n} d_i'(\overline{q}) \,. \tag{4.8}$$

Ármeghatározó játékos esetén a $\overline{q}$ összmennyiség egységnyi növelésének társadalmi költsége itt is a $\sum d_i$ tagon keresztül jelentkezik, és a haszon itt is a játékosok magasabb hasznosságaiban ölt testet. Ez az esetleges haszon egyfelől az első játékos egyensúlyi fogyasztására gyakorolt közvetlen, másrészt a valamennyi játékos fogyasztására gyakorolt közvetett, az áron keresztül jelentkező hatásként adódik. A társadalmi optimum feltétele az így értelmezett határköltség és határhaszon (4.8) által előírt egyenlősége. Egyáltalán nem biztos azonban, hogy az összmennyiség növekedése esetén az egyensúlyi fogyasztások valóban növekedni fognak, növelve ezáltal a hasznosságokat. Az egyensúlyi fogyasztások természetesen az ár függvényében alakulnak, amelyet viszont az első játékos indulókészlete határoz meg. Minden olyan elosztás esetén, amikor a növekvő összmennyiség az első játékos indulókészletének növekedését vonja maga után, növelve ezáltal az egyensúlyi árat, a többi játékos egyensúlyi fogyasztása csökkeni fog, ami csökkenti haszosságukat. Ebből adódóan az elosztási szabály ármeghatározó játékos esetén is döntő befolyással lesz az optimális összmennyiségre. Akárcsak az imént tranzakciós költségek esetén, most is érdekes megvizsgálni, előfordulhat-e, hogy a tökéletes verseny esetén adódó összmennyiség tökéletlen piacon is optimális marad. (4.8)-ból adódik a következő állítás.

4.3 Állítás. *Ármeghatározó játékos esetén az optimális összmennyiség pontosan akkor egyezik meg a versenyzői optimummal, ha az alábbi feltételek közül legalább az egyik teljesül.*

(a) az elosztás utilitarista;

(b) teljesül, hogy

$$-\sum_{i=2}^{n}\frac{u_1''(q_1^*)}{u_i''(q_i^*)}=\sum_{j=2}^{n}\frac{dq_j^0}{d\overline{q}}\,. \tag{4.9}$$

Bizonyítás. Az optimális összmennyiség (4.3) szerint pontosan akkor a versenyzői optimum, ha (4.8)-ban

$$(q_1-q_1^0)\frac{dp}{d\overline{q}}=0\,.$$

A szorzat első tagja pontosan (a) esetén zérus. Belátható ezen kívül, (lásd C függelék), hogy $dp/d\overline{q}=0 \Leftrightarrow$ (4.9) fönnáll. □

Hasonlítsunk össze például egy versenyző és egy ármeghatározó esetet, amelyekben a kritikus játékos elhárítási hasznossága lineáris. A 4.3 (b) állítás szerint ekkor annak feltétele, hogy a versenyző esetben kapott optimális összmennyiség valamely (nem utilitarista) elosztás esetén optimális maradjon az ármeghatározó esetben is, az, hogy $\sum_{i=2}^{n} dq_i^0/d\overline{q}=0$ teljesüljön. Ez azt jelenti, hogy az adott elosztási szabálynak az optimális $\overline{q}$-nál az összmennyiség teljes növekményét az 1. játékosra kell osztania ($dq_1^0/d\overline{q}=1$). A többi játékos között csak a már korábban is náluk lévő készletek oszthatók újra.

Ebben a részben az összmennyiség endogén meghatározásával bővítettem a korábbi modellt. A tökéletes és a tökéletlen piacokat elemezve azt találtam, hogy az optimális összmennyiség meghatározása általában függ az elosztási szabálytól és a piac jellemzőitől. Míg tökéletes verseny esetén a kívánatos összmennyiség minden elosztás estén ugyanakkora, addig ez tökéletlen piacokon általában nem teljesül. Elmozdulva továbbá a versenyző piac feltevésétől, az optimális összmennyiség csak utilitarista elosztás, vagy speciális függvényformák mellett marad változatlan. Mindezek szemléltetésére szolgál az alábbi példa.

Egy példa

Korábbi példánkat kibővítve válasszunk kvadratikus $d_i(\overline{q})$ kárfüggvényeket a következő módon:

$$d_i(\overline{q})=b_i\overline{q}^2/2 \quad i=1,2\,,$$

ahol b_i pozitív paraméter. Versenyző esetben ekkor az optimális $\overline{q}$ összmennyiséget a következő feladat megoldásaként kapjuk

$$\min_{\overline{q}}\sum_{i=1}^{2}\Big[c_i(q_i(p(\overline{q})))+d_i(\overline{q})\Big]=\min_{\overline{q}}\sum_{i=1}^{2}\left[\left(\frac{A-\overline{q}}{2}\right)^2+\frac{b_i}{2}\overline{q}^2\right]\,,$$

ahonnan

$$\overline{q} = \frac{A}{B+1}$$

$(B = b_1 + b_2)$ az elosztástól függetlenül.

Tökéletlen piacon az elmondottak alapján általában más optimális összmennyiségek adódnak. Tekintsük például a tranzakciós költségek esetét, és válasszuk a paramétereket a 3. szakasz példájához hasonlóan, valamint legyen $B = b_1 + b_2 = 0{,}1$. A *4.1. táblázat* szemlélteti az EG, az utilitarista és a Nash elosztás esetén adódó társadalmilag optimális összes szennyezést. A versenyző eset optimális összmennyiségei a 4.2. állításnak megfelelően pontosan a középső oszlop értékei.

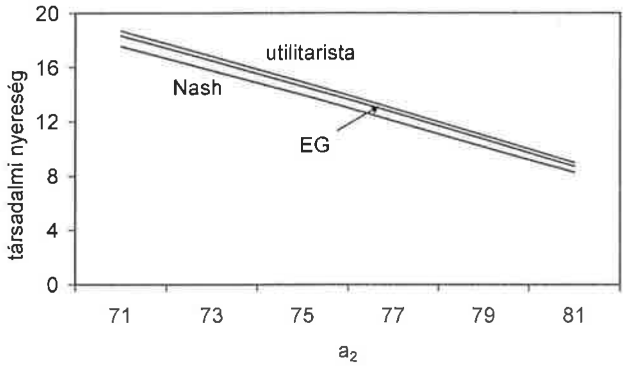
a_2	EG	utilitarista	Nash
71	91,8183	91,8182	91,7419
73	93,6365	93,6364	93,5662
75	95,4546	95,4545	95,3910
77	97,2728	97,2727	97,2160
79	99,0909	99,0909	99,0352
81	100,9090	100,9091	100,8625

4.1. táblázat. Optimális összmennyiség tranzakciós költségek esetén

Mivel az utilitarista és a másik két vizsgált elosztás közti fő különbség abból fakad, hogy az utóbbiak figyelembe veszik a D_i károk különbségeit, és példánkban ez a különbség viszonylag kicsi $(40 - 30 = 10)$, az optimális $\overline{q}$ értékek nagyon közel esnek egymáshoz. A *4.1. ábra* mutatja az egyes elosztások mellett a szennyezési piac révén elérhető maximális társadalmi nyereséget, vagyis a

$$D_1 + D_2 - \min_{\overline{q}} \sum_{i=1}^{2} \left[c_i(q_i^*) + T(q_i^* - q_i^0) + d_i(\overline{q}) \right]$$

értékeket. Az utilitarista elosztás mellett kapott nyereségek egyúttal a versenyző esetben, tetszőleges elosztás mellett elérhető társadalmi nyereségeket adják.



4.1. ábra. Maximális társadalmi nyereségek tökéletes verseny, illetve tranzakciós költségek esetén (A versenyző eset nyereségeit az utilitarista elosztás értékei adják)

5 Alkalmazás: szennyezési jogok nemzetközi kereskedelme

A leírtak segítségével módunk nyílik értékelni a nemzetközi környezetszabályozás intézményeinek néhány jellemzőjét. Az 1992-ben a New York-i ENSZ találkozón elfogadott Klímaváltozási Keretegyezmény jegyében zajló nemzetközi környezetszabályozás a következő módon jellemezhető. Az egyezmény célja „az üvegházhatású gázok koncentrációjának stabilizálása az atmoszférában egy olyan szinten, amely még nem képez veszélyes emberi beavatkozást a klímarendszerbe" (UNFCCC, 1992). Ennek érdekében a tárgyalópartnerek elfogadtak egy, a globális szinten elérni kívánt szennyezéscsökkentésre vonatkozó ajánlást, vagyis meghatározták az összes szennyezést ($\overline{q}$). E „szennyezési jogok" szétosztására ($\mathbf{q}^0$) 1997-ben a kiotói tárgyaláson (kooperatív környezetben) került sor.

Az elosztással kapcsolatban legtöbbet hangoztatott elv mindig is a méltányosság volt: a szabályozásnak e szerint „az erőfeszítések és kötelezettségvállalások kiegyensúlyozott és méltányos megosztásán kell alapulnia" (UNEP, 1995, p. 71, idézi Bierman, 1999), és a terhek egyenlő elosztását (equitable burden-sharing) kell szem előtt tartania (Tóth, 1999). Az ún. Kiotói Jegyzőkönyv rendelkezik egy nemzetközi szennyezési piac kialakításáról is, amelynek a részleteit azonban későbbi egyeztetések tárgyává teszi – ezek az egyeztetések ma is tartanak.

Hogyan értékelhetjük a nemzetközi környezetszabályozásnak ezt a folyamatát? Először is, amennyiben versenyző szennyezési piac kialakulása várható, számos pozitívumot tudunk kiemelni. Bárhogyan is osztották el Kiotóban az országok egymás között a szennyezési jogokat, a tárgyalásokat megelőzően meghatározott összmennyiség ilyenkor minden elosztás mellett optimális marad. Továbbá, ha a jogok elosztásának méltányossága az itt definiált egyenlő nyereségeket jelentette, akkor ez egyúttal kielégíti Nash vonzó axiómáit is. Sajnos azonban ezek a megállapítások általában érvényüket vesztik, ha a szennyezési piacon nem tökéletes verseny érvényesül majd. Mint láttuk, ilyenkor az összmennyiség optimális volta az elosztási szabályon múlik, például egy versenyző piacot feltételezve megállapított, majd az EG szabály alapján szétosztott összmennyiség nem lesz optimális tökéletlen piacokon. Ráadásul az egyenlő nyereségeket biztosító elosztás sem a Nash-axiómáknak, sem az utilitarista kritériumnak nem tesz eleget.

Mivel a Kiotói Jegyzőkönyv csupán általános irányelveket tartalmaz a szennyezési piacra vonatkozóan, feltehetőleg mind a globálisan kívánatos összes szennyezéscsökkentés meghatározása, mind pedig az elhárítási kötelezettségek szétosztása anélkül történt, hogy ismertek lettek volna a kialakuló piac működési feltételei, becslések születtek volna a várható tranzakciós költségekről stb. Nehéz tehát elképzelni, hogy a kiotói egyezményben rögzített elvek és mennyiségek valóban a kívánt eredményre vezessenek. Eredményeink szerint ez egyedül utilitarista elosztás esetén lenne biztosítva, azonban világos, hogy a jelenlegi nemzetközi környezetpolitikában az utilitarista elosztásban testet öltő költséghatékonyság háttérbe szorul az igazságosság különböző megközelítéseivel és a politikai megvalósíthatóság szempontjaival szemben (Nordhaus - Boyer, 1998).

6 Összefoglalás

A dolgozatban a versenyző és a tökéletlen cserepiacok hagyományos modelljeit kibővítettem az indulókészletek kooperatív alkufolyamat során történő elosztásával, és az összmennyiség endogén meghatározásával. A különböző piacmodellek esetén jellemeztem három fontos készletelosztási szabályt, az utilitarista, a Nash-, és az egyenlő nyereségeket biztosító elosztást. Megvizsgáltam továbbá az optimális összmennyiség meghatározását ezeken a piacokon. Fő eredményeim a következők voltak.

1. *Az elosztás elvei és a piac jellemzői nem kezelhetők egymástól függetlenül.* Az adott elvek szerint történő elosztás más eredményre vezet attól függően, hogy milyen tulajdonságokkal rendelkezik majd az elosztás nyomán kialakuló piac. Megfordítva, ha adott elosztási elveket érvényesítetve jött létre a jogok elosztása valamilyen piaci forgatókönyv mellett, ezek az elvek már nem érvényesek többé, ha megváltoznak a piac működési feltételei. Láttuk, hogy míg tökéletes piacon minden elosztás utilitarista (3.1 állítás), addig tranzakciós költségek vagy ármeghatározó játékos esetén a költséghatékonyság csak speciális elosztások mellett teljesül (3.4 és 3.10 állítás). Megmutattam, hogy versenyzői piacon a Nash elosztás igen vonzó tulajdonságokkal bír, nevezetesen egyenlő (vagy „legegyenlőbb") nyereségeket biztosít a játékosoknak (3.3 állítás). Tökéletlen piacokon ez azonban csak speciális költségfüggvények és fenyegetési pontok esetén teljesül (3.7 és 3.11 állítás).

2. *A társadalmilag kívánatos összmennyiség nem határozható meg az elosztási szabály és a piac jellemzőitől függetlenül.* A társadalmi összhaszon maximalizálása általában más összmennyiséget kíván attól függően, hogy melyik piaci forgatókönyv valósul meg, és hogy milyen szabály szerint fogják szétosztani az adott összkészletet. Ebből adódóan egy adott piacra, adott elosztási szabály figyelembevételével meghatározott összmennyiség nem lesz optimális, amennyiben módosulnak a piac jellemzői, vagy a játékosok más elosztási elvekre térnek át. Láttuk, hogy tökéletes piacon, vagy állandó tranzakciós költségek esetén az optimális összmennyiség minden elosztás esetén ugyanakkora (4.1 és 4.2 (b) állítás). Tökéletlen piacon azonban az optimális összmennyiség csak utilitarista elosztás, vagy speciális függvényformák és paraméterek mellett egyezik meg a versenyző piacnál kapott összmennyiséggel (4.2 és 4.3 állítás). Az eredmények segítségével értékelhettük a nemzetközi szennyezési piaccal kapcsolatban felvetett elveket, és a Kiotói Jegyzőkönyv néhány rendelkezését.

Az elmondottak kiterjesztésének egy kézenfekvő módja más típusú elosztások figyelembe vétele a készletelosztási játékban. A kooperatív alku-elmélet sok egyéb megoldást tárgyal, és ezek axiomatizálása is előrehaladott. A fontos kérdés itt az, vajon melyik megoldás-koncepció elveit részesítsük előnyben, illetve melyik írja le jobban az adott helyzetben (pl. a klímaváltozási tárgyalásokon) követett szempontokat. Egy elméleti és gyakorlati szempontból egyaránt izgalmas kiterjesztési lehetőséget jelent, ha bevezetjük a termelést is a piac modelljeibe. Végül a piac ábrázolásánál figyelembe lehet venni számos további forgatókönyvet, amelyek döntően befolyásolhatják az elosztások és az optimális összmennyiség tulajdonságait. Két, az irodalomban megtalálható jelölt lehet a rendelkezésre álló készletnél magasabb fogyasztás lehetősége (a szabályok áthághatósága) és az ehhez kapcsolódó büntetés figyelembe vétele (Malik, 1990), valamint a vizsgált jószág piacának kölcsönhatása más termékpiacokkal, és a tökéletlen termékpiacok következményei (Sartzetakis, 1997).

Függelékek

A függelék. *A 3.5 lemma bizonyítása.*

A játék Nash-megoldását az alábbi optimumfeladat adja.

$$\begin{aligned} \max_{q_1^0,\ldots,q_n^0} \quad & \prod_{i=1}^{n} R_i(\mathbf{q}^0) \\ \text{k.f.} \quad & \sum_{i=1}^{n} q_i^0 = \overline{q} \\ & q_i^0 \geq 0 \quad i = 1, \ldots, n \, . \end{aligned}$$

Az elsőrendű feltételek (λ-val jelölve a pozitív Lagrange-szorzót)

$$\sum_{i=1}^{n} \prod_{k \neq j} R_k \frac{dR_j}{dq_i^0} - \lambda \leq 0 \quad q_i^0 \geq 0 \text{ komplementaritással}, \quad i = 1, \ldots, n \, , \tag{A.1}$$

és

$$\sum_{i=1}^{n} q_i^0 = \overline{q} \, . \tag{A.2}$$

Az (A.1) alatti elsőrendű feltételeket leosztva $R_1 \cdot \ldots \cdot R_n$-nel és átrendezve ezt írhatjuk:

$$\sum_{j=1}^{n} \frac{1}{R_j} \frac{dR_j}{dq_u^0} \leq \frac{\lambda}{R_1 \cdot \ldots \cdot R_n} \, . \tag{A.3}$$

(A.3) egyszerűsítéséhez kihasználjuk, hogy a burkológörbe-tétel szerint

$$\frac{dR_i}{dq_i^0} = \frac{\partial R_i}{\partial p} \frac{dp}{dq_i^0} + \frac{\partial R_i}{\partial q_i^0} = -\frac{dp}{dq_i^0}(q_i - q_i^0) + p + T'(q_i - q_i^0) \quad \forall i\text{-re} \tag{A.4}$$

és

$$\frac{dR_j}{dq_i^0} = \frac{\partial R_j}{\partial p} \frac{dp}{dq_i^0} = -\frac{dp}{dq_i^0}(q_j - q_j^0) \quad i \neq j \, . \tag{A.5}$$

(A.4)-et és (A.5)-öt behelyettesítve (A.3)-ba és átrendezve adódik (3.6). □

B függelék. *A 3.9 lemma bizonyítása.*

A játék Nash-megoldását az alábbi feladat optimuma adja:

$$\begin{aligned} \max_{q_1^0,\dots,q_n^0} \quad & \prod_{i=1}^{n} R_i(q_1^0, q_i^0) \\ \text{k.f.} \quad & \sum_{i=1}^{n} q_i^0 = \overline{q} \\ & q_i^0 \geq 0 \quad i = 1, \dots, n\,. \end{aligned}$$

Az elsőrendű feltételek $i = 2, \dots, n$-re a következők:

$$\prod_{j \neq i} R_j(q_1^0, q_i^0)p - \lambda \leq 0 \quad q_i^0 \geq 0 \text{ komplementaritással}\,. \tag{B.1}$$

Innen, $R_1 \cdot \ldots \cdot R_n$-nel leosztva és átrendezve adódik (3.10).

Az elsőrendű feltétel $i = 1$-re

$$\sum_{j=1}^{n} \frac{dR_j}{dq_1^0} \prod_{k \neq j} R_k - \lambda \leq 0 \quad q_1^0 \geq 0 \text{ komplementaritással}\,. \tag{B.2}$$

(B.2)-t leosztva $R_1 \cdot \ldots \cdot R_n$-nel kapjuk, hogy

$$\sum_{j=1}^{n} \frac{dR_j}{dq_1^0} \frac{1}{R_j} \leq \frac{\lambda}{R_1 \cdot \ldots \cdot R_n} = \frac{p}{R(\mathbf{q}^0)}\,. \tag{B.3}$$

(B.3) egyszerűsítéséhez kihasználjuk, hogy a burkológörbe tétel értelmében

$$\frac{dR_1}{dq_1^0} = \frac{\partial R_1}{\partial q_1} \frac{dq_1}{dq_1^0} + \frac{\partial R_1}{\partial p} \frac{dp}{dq_1^0} + \frac{\partial R_1}{\partial q_1^0} = \frac{\partial R_1}{\partial q_1^0} = p\,, \tag{B.4}$$

és $i = 2, \dots, n$-re

$$\frac{dR_i}{dq_1^0} = \left(\frac{\partial R_i}{\partial q_i} \frac{dq_i}{dp} + \frac{\partial R_i}{\partial p} \right) \frac{dp}{dq_1^0} = \frac{\partial R_i}{\partial p} \frac{dp}{dq_1^0} = -(q_i - q_i^0) \frac{dp}{dq_1^0}\,. \tag{B.5}$$

(B.4)-et és (B.5)-öt behelyettesítve (B.3)-ba, és átrendezve adódik (3.11). □

C függelék. $dp/d\overline{q} = 0 \Leftrightarrow$ (4.9) *teljesül.*

Először is vegyük észre, hogy

$$\frac{dp}{d\overline{q}} = \frac{\partial p}{\partial \overline{q}} + \frac{\partial p}{\partial q_1^0} \frac{dq_1^0}{d\overline{q}}\,. \tag{C.1}$$

Jelöljük Z-vel az ármeghatározó játékos (2.3) határhasznának p szerinti deriváltját (ez a másodrendű feltétel értelmében negatív). Ekkor (2.3)-ból adódnak a következők:

$$\frac{\partial p}{\partial \overline{q}} = -\frac{1 + u_1''(q_1^*) \sum_{i=2}^{n} q_i'(p)}{Z}$$

és

$$\frac{\partial p}{\partial q_1^0} = \frac{1}{Z} \, .$$

Ezt kihasználva, (C.1) szerint

$$\frac{dp}{d\bar{q}} = 0 \Leftrightarrow \frac{1 + u_1''(q_1^*) \sum_{i=2}^n q_i'(p)}{Z} = \frac{1}{Z} \frac{dq_1^0}{d\bar{q}} \, . \qquad \text{(C.2)}$$

Használjuk most fel, hogy $q_i'(p) = 1/u_i''(q_i)$ (lásd a 3. lábjegyzetet) és hogy $dq_1^0/d\bar{q} = 1 - \sum_{i=2}^n dq_i^0/d\bar{q}$. (C.2)-ben Z-vel egyszerűsítve, és behelyettesítve adódik (4.9). □

Irodalom

1. Amacher, G. S. - Malik, A. S. (1996): Bargaining in Environmental Regulation and the Ideal Regulator, *Journal of Environmental Economics And Management*, (30) 233–253.
2. Amacher, G. S. – Malik, A. S. (1998): Instrument Choice When Regulators and Firms Bargain, *Journal of Environmental Economics And Management*, (35) 225–241.
3. Barrett, S. (1994): Self-enforcing international environmental agreements, *Oxford Economic Papers*, (46) 878–894.
4. Biermann, F. (1999): Justice in the Greenhouse: Perspectives from International Law, in: Tóth (1999), 160–173.
5. Carraro, C. - Siniscalco, D. (1993): Strategies for the international protection of the environment, *Journal of Public Economics*, (52) 309–328.
6. Chander, P. - Tulkens, H. (1997): The Core of an Economy with Multilateral Environmental Externalities, International *Journal of Game Theory*, (26) 379–401.
7. Coase, R. (1963): The problem of social cost, *Journal of Law and Economics*, (3) 1–44.
8. Forgó F. - Szép J. - Szidarovszky F. (1999): *Introduction to the Theory of Games*: Concepts, Methods, Applications, Kluwer Academic Publishers, Dordrecht.
9. Hahn, R. W. (1984): Market Power and Transferable Property Rights, *The Quarterly Journal of Economics*, November 753–765.
10. Malik, A. (1990): Markets for pollution control when firms are noncompliant, *Journal of Environmental Economics and Management*, (18) 97-106.
11. Muller, R. A. - Mestelman, S. (1998): What Have We Learned from Emissions Trading Experiments?, *Managerial and Decision Economics*, (19) 225–238.
12. Nash, J. F. (1950): The Bargaining Problem, *Econometrica*, (18) 155–162.
13. Nordhaus, W. D. - Boyer, J. (1998): Requiem for Kyoto: An Economic Analysis of the Kyoto Protocol, *The Energy Journal*, Kyoto Special Issue.
14. OECD (1998): *Lessons from Existing Trading Systems for International GHG Emission Trading*, Paris.
15. Rawls, J. (1971): *A Theory of Justice*, Harvard University Press, Cambridge MA.

16. Sartzetakis, E. S. (1997): Tradeable Emission Permits Regulations in the Presence of Imperfectly Competitive Product Markets: Welfare Implications, *Environmental and Resource Economics*, (9) 65–81.
17. Stavins, R. N. (1995): Transaction Costs and Tradeable Permits, *Journal of Environmental Economics and Management*, (29) 133–148.
18. Tóth F. L. (1999) (szerk): *Fair Weather? Equity Concerns in Climate Change*, Earthscan, London.
19. Ujhelyi G. (2001): Készletelosztási játék és optimális összmennyiség a szennyezési jogok piacán (szakdologzat), Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Matematikai Közgazdaságtan és Ökonometria tanszék.
20. UNEP (1995), Report of the Seventh Meeting of the Parties to the Montreal Protocol on Substances that Deplete the Ozone Layer (Vienna, 5-7 December 1995), 27 December
21. UNFCCC (1992): United Nations Framework Convention on Climate Change, New York, http://www.unfccc.org/resource/conv/

ENDOWMENT GAME AND OPTIMAL TOTAL QUANTITY IN ARTIFICIAL MARKETS

The paper deals with twó key issues related to the creation of artificial markets: determining the total quantity of the market good, and the initial allocation of these goods among the participants. Various market structures are considered, and the "Endowment Game" allows for allocation principles different from efficiency to be examined. I show that the characteristics of these principles vary with the market structure. Total quantity is chosen by a social planner. I show that the optimal total quantity varies with market structure and the allocation principle. The results have implications for the international emission trading system.