Operációkutatás jegyzet

Frank András és Király Tamás

2013. szeptember

Bevezetés

Az operációkutatás (Operations Research) fogalma a II. világháború során alakult ki, eredeti jelentése hadműveleti kutatás. Alapvetően nagy rendszerek (hadsereg, nagyvállalatok, hálózatok) tervezési, irányítási, üzemeltetési problémáinak matematikai módszereivel foglalkozik. Részét képezi az alkalmas matematikai modellek megtalálása illetve kifejlesztése, e modellek matematikai vizsgálata beleértve hatékony algoritmusok kidolgozását és végül az algoritmusok implementálása, megfelelő informatikai környezetbe ágyazása, és az eredeti gyakorlati feladaton történő alkalmazása.

A jelen munka az ELTE TTK matematikus alapképzésben tartott két féléves Operációkutatás c. előadás anyagát öleli fel. Az előadás elméleti és alkalmazott matematikus szakirányú hallgatóknak szól. Legfőbb célja a matematikai háttér bemutatása valamint azon gyakorlati problémák felvázolása, amelyek alapvetően motiválták a szóbanforgó matematikai modellek kialakulását és vizsgálatát.

Mindvégig azt a kettős célt tartottuk szem előtt, hogy egyrészt megadjuk a szilárd alapokat azon hallgatóknak, akik később az Operációkutatás különféle ágaival részletekbe menően akarnak majd foglalkozni, másrészt kellőképp tájékoztassuk azon hallgatókat is a matematikai alapkultúrához tartozó legfontosabb operációkutatási eredményekről, akik későbbi tanulmányaik során más irányban szakosodnak majd.

A jegyzet némileg bővebb, mint ami egy kurzus során ténylegesen bemutatásra kerül, de célunk volt, hogy az érdeklődő hallgató jobban el tudja helyezni az elhangzottakat, biztosabb alapokhoz és szélesebb kitekintéshez jusson.

Bár formálisan bevezetésre kerülnek, előfeltételként szükség van olyan gráfelméleti alapfogalmakra, mint út, kör, fa, páros gráf, aciklikus digráf, Euler gráf, párosítás, kromatikus szám, továbbá az sem bizonyul hátrányosnak, ha az olvasó találkozott már olyan alaptételekkel, mint Menger és Kőnig tételei, vagy olyan algoritmusokkal, mint Kruskal mohó algoritmusa és a Dijkstra algoritmus. Hasonlóképp támaszkodunk a lineáris algebra olyan alapfogalmaira és eredményire, mint például vektortér, determináns, lineáris egyenletrendszer megoldhatósága, Gauss elimináció. Amire valójában itt szükségünk van, azt a Lineáris egyenletrendszerek c. fejezetben összefoglaljuk. (Ez a fejezet csak segédanyag és értelemszerűen nem szerepel a kurzuson.)

Végül nélkülözhetetlen valamiféle előzetes tájékozottság az algoritmusok mibenlétéről illetve a hatékonyságuk fogalmáról, futási időről. A jegyzetben speciális hangsúlyt fektetünk a polinomiális futásidejű algoritmusok bemutatására. Céljainkhoz elegendő, ha ezt a fogalmat, akárcsak az **NP** és **NP**-teljesség fogalmát, csak szemléletesen vezetjük be és használjuk.

A jegyzet nem csupán definíciók, tételek, algoritmusok és bizonyítások akkurátusan összeállított gyűjteménye: igyekeztünk a motivációkat, keresztkapcsolatokat, a fogalmak hátterét megvilágítani. Két olyan szakasz is van (az 1.1 és a 3.1.1), amely inkább csak „mesebeszéd” abban a tekintetben, hogy nem konkrét matematikai eredmények ismertetését, mint inkább a szemlélet megalapozását szolgálják.

Nemegyszer előfordul, hogy ugyanarra a tételre több bizonyítást is adunk, nem mintha egyetlen bizonyítás nem volna kellőképp meggyőző, hanem azért, mert a bizonyítási gondolatot legalább olyan fontosnak tartjuk, mint a bizonyításra kerülő tételt magát. Néha pedig egy később bizonyításra kerülő tételt, állítást korábban feladatként kitűzünk, mert az olvasó gyakran jobban jár, ha saját maga is megpróbálkozik a bizonyítással.

Az anyag első fele az elmúlt közel húsz esztendő oktatási tapasztalatain alapul. A II. félév anyagát tartalmazza a jegyzet további három fejezete. Miután az Operációkutatás II. tárgy csak a BSc képzés indulásakor került bevezetésre, a 6-8. fejezetek mögött értelemszerűen kevesebb oktatási gyakorlat áll.

Az I. félév anyaga, amelyet az első öt fejezet dolgoz fel, tartalmilag három részre oszlik. Az Optimalizálás gráfokon című fejezet áttekinti a gráfokra, hálózatokra vonatkozó alapvető optimalizálási feladatokat, bemutatja az ezeket motiváló gyakorlati alkalmazásokat és ismerteti a legfontosabb megoldó algoritmusokat. Ezután kerülnek ismertetésre a lineáris programozás alapfogalmai és eredményei, beleértve a Farkas lemmát és a dualitás tételt. Végül a harmadik részben elmagyarázzuk, hogy a gráfra vonatkozó alaptételek (Gallai, Egerváry, Hoffman tételei) és ezek kiterjesztései miként következnek a teljesen unimoduláris mátrixok egy elegáns tulajdonságát használva a Farkas lemmaból illetve a Dualitás tételből.

Az utolsó három fejezet a II. félév anyagának nagy részét tartalmazza. Kimaradt a játékelmélet alapjainak ismertetése, mivel ez megtalálható a Játékelmélet c. online jegyzetben. A hatodik fejezet a szimplex módszer különféle változatait mutatja be, köztük a folyam feladatok megoldására szolgáló hálózati szimplex módszert. Ezután az egészértékű lineáris programozás alapvető módszereit ismertető fejezet következik, végül pedig a konvex optimalizálás elméletének alapjait (Karush-Kuhn-Tucker tétel, Lagrange duális) és a legegyszerűbb megoldási módszereket ismerheti meg az olvasó.

Ezúton mondunk köszönetet Szeszlér Dávidnak a rendkívül alapos lektori munkájáért, hasznos megjegyzéseiért és tanácsaiért, valamint Papp Olgának, aki a második féléves anyagot Király Tamás előadásai alapján lejegyzetelte; a 6-8. fejezetek az általa leírtak továbbfejlesztéseként jöttek létre. Végül felhívjuk a figyelmet az *Operációkutatás példatár* jegyzetre, amely feladatgyűjteményként szolgál a jelen anyaghoz.

Budapest, 2013. augusztus

Frank András és Király Tamás

Tartalomjegyzék

1. fejezet

Optimalizálás gráfokon

Ebben a fejezetben gráfokra vonatkozó optimalizálási problémákat vizsgálunk, bemutatva a megoldásukra szolgáló algoritmusokat. Feltételezzük, hogy az olvasó már találkozott a gráfelmélet olyan alapfogalmaival, mint csúcs, él, párhuzamos élek, kör, vágás, fa, erdő, irányított és irányítatlan gráf, stb.

1.1. Algoritmusok hatékonyságáról

Bevezetésképp az algoritmus fogalmáról, algoritmusok hatékonyságáról illetve leállási feltételekről adunk röviden szemléletes tájékoztatást. Az itt szereplő tételek és bizonyítások csupán a felvetődő fogalmak, gondolatok illusztrálására szolgálnak.

Az algoritmus szemléletes fogalma mára már a köznyelvbe is beszivárgott: valamiféle utasítás sorozatot, forgatókönyvet jelent, amely egy inputnak nevezett kiinduló helyzetből elemi lépések megadott sorozatával javasol eljutást egy megcélzott véghelyzetbe. Egy szakácskönyv algoritmusok gyűjteménye. Két szám összeszorzása éppúgy algoritmussal történik, mint mondjuk egy háromszög megszerkesztése a három súlyvonalából. Algoritmust ad meg Karinthy, amikor leírja miképp lehet eljutni a Csömöri úttól egészen a Filatori gátig.

Kérdés persze, hogy mit is értünk elemi lépésen. Ez rajtunk illetve a szituáción múlik. Amikor rántottát készítünk, a tojás feltörése, vagy a hagyma felszeletelése tekinthető elemi műveletnek. Ha egy hotel séfjeként egy száz fős társaságnak kell gondoskodni a reggelijéről, akkor maga a rántotta készítés tekinthető elemi műveletnek. Ez az egymásba ágyazási szemlélet jellemző az algoritmikus gondolkodásmódra: korábban elkészített egyszerűbb algoritmusokat gyakran használunk kész segédeszközként –szubrutinként– egy összetettebb algoritmus készítésekor.

Egy algoritmustól elsősorban azt várjuk, hogy véges legyen, de ez valójában még édeskevés, hiszen ha az alapstruktúra véges, akkor többnyire nem okoz leküzdhetetlen nehézséget a véges sok lehetőséget mind számba venni, legalábbis elvben. Ugyanakkor a lehetőségek száma tipikusan olyan nagy, hogy még viszonylag kis példákon is kilátástalan a teljes áttekintés akár a legjobb számítógépet használva. Tekintsük például azt a feladatot, amelyben egy n pontú $G = (V, E)$ gráfról el kell döntenünk, hogy a pontjait meg lehet-e színezni k színnel úgy, hogy egy színosztályon belül nem vezet él. (Ezt nevezhetjük jó színezésnek.) Ha a gráfban bármely két pont szomszédos, úgy pontosan akkor létezik jó színezés, ha $k \geq n$. Ha van két nem-szomszédos csúcs, akkor a feladatot kétfelé vághatjuk, annak megfelelően, hogy a két csúcs azonos színt kap-e majd vagy különbözőt.

1.1.1. Állítás. *Amennyiben u és v nem-szomszédos csúcsok, úgy G-nek pontosan akkor létezik jó színezése, ha a G' és G'' gráfok közül legalább az egyiknek létezik, ahol G' az u és v csúcsok összehúzásával keletkezik G-ből, míg G'' az új uv él hozzáadásával.*

Biz. Ha G-nek létezik jó színezése, úgy az u és v színe vagy megegyezik vagy különböző. Az első esetben G'-nek kapjuk egy jó színezését, a másodikban G''-nek. Megfordítva, mind G'-nek, mind G''-nek egy jó színezése természetesen kiterjeszthető G jó színezésévé. •

Az állítás közvetlenül megad egy rekurzív algoritmust, ami nyilván véges. Ugyanakkor az eljárás a gyakorlatban használhatatlan már viszonylag kis gráfok esetén is ($n \geq 30$), mert minden lépésben megduplázodik a gráfok száma. Ez az algoritmus tehát exponenciális lépésszámú a bemenő gráf méretének függvényében. (Egy gráf vagy egész szám mérete szemléletesen a számítógépes tárolásukhoz szükséges bitek száma.) Általános tapasztalat, hogy exponenciális vagy nagyobb lépésszámú algoritmusok a gyakorlat számára gyakran használhatatlanok. Ennek hátterére világít rá a következő összehasonlító táblázatocska. Feltéve, hogy a számítógépünk másodpecenként egymillió elemi utasítást képes végrehajtani, a futási idő az input n-nel jelölt méretének függvényében a következőképp alakul egy n^2-es, n^3-os, 2^n-es és egy $n!$ lépésszámú algoritmus esetén.

Lépésszám:	n^2	n^3	2^n	$n!$
$n = 10$	< 1 mp	< 1 mp	< 1 mp	4 mp
$n = 30$	< 1 mp	< 1 mp	18 p	10^{25} év
$n = 50$	< 1 mp	< 1 mp	36 év	rengeteg év
$n = 100$	< 1 mp	1 mp	10^{17} év	rengeteg év
$n = 1000$	1 mp	18 p	rengeteg év	rengeteg év
$n = 10000$	2 p	12 nap	rengeteg év	rengeteg év

Akkor tekintünk hatékonynak egy algoritmust, ha a lépésszáma a bemenő adatok méretének egy hatványával korlátozható. Az ilyen algoritmust polinomiális futásidejűnek nevezik (szemben az exponenciális vagy még nagyobb futásidejű algoritmusokkal). A továbbiakban a lépésszám és futásidő szavakat egymás szinonímáiként használjuk. Egy tulajdonságról vagy a tulajdonság meglétét firtató problémáról azt mondjuk, hogy **P**-ben van, ha létezik polinomiális futásidejű algoritmus a tulajdonság meglétének az ellenőrzésére. Közismert például, és alább látni is fogjuk, hogy egy gráf 2-színezhetősége **P**-ben van.

A főnök elv avagy a leállási szabály

Mielőtt azonban konkrét hatékony algoritmusok tervezéséről szót ejtenénk, fontos kihangsúlyozni, hogy megfogalmazható egy közbeiktatott kérdés, éspedig az, hogy meg lehet-e adni valamiféle eszközt annak ellenőrzésére, hogy az (egyelőre esetleg ismeretlen) algoritmus által szolgáltatott végeredmény vajon helyes-e, anélkül, hogy az algoritmust nekünk újra futtatnunk kéne. Ez a kérdés tökéletesen független magától az algoritmustól, vagyis attól, hogy a keresett eredményt miképp is kaptuk meg.

Kis színes illusztrációként a kérdést a következő alakban szokták megfogalmazni. Munkahelyi főnökünk kiad egy megoldandó feladatot, például egy nagyméretű $Ax = b$ lineáris egyenletrendszer egy megoldását kell meghatároznunk. Két napon át éjjel-nappal dolgozunk a problémán, és végül megtalálunk egy megoldást, amit boldogan viszünk a főnöknek. Ő természetesen kíváncsi arra, hogy a megoldásunk helyes-e. Jószerencse, hogy ehhez egyáltalán nem kell a mi számításainkat ellenőriznie, hanem egyszerű behelyettesítéssel gyorsan el tudja dönteni a helyességet. De mi van akkor, ha a konkrét rendszernek éppenséggel nincsen megoldása? Ekkor abban a kényelmetlen helyzetben találjuk magunkat, hogy hiába dolgoztunk éjjel-nappal, azt vagyunk kénytelenek főnökünknek jelenteni, hogy nem találtunk megoldást. Kérdő arckifejezésére semmi jobbal nem tudunk válaszolni, minthogy elővesszük a 25 oldalnyi számítást és elkezdjük a részleteket magyarázgatni, amire a főnök joggal válaszolja azt, hogy neki nincs ideje két napi munka minden részletét lépésről lépésre ellenőriznie. A kérdés tehát az, hogy nem létezik-e valami olyan bizonyíték (igazolvány, tanúsítvány), amit lehet ugyan, hogy csak hosszadalmas számítgatással lehetett megkapni, de ha már egyszer rendelkezésre áll, akkor a főnök arra rápillantva rögtön meggyőzve érezheti magát, hogy a konkrét egyenletrendszernek valóban nincs megoldása. Más szóval egy könnyen ellenőrizhető leállási szabályt keresünk: végül is legelőször saját magunkat kell meggyőznünk, hogy algoritmusunk helyes végeredményt adott.

Egyszerű dolgunk van, ha az egyenletrendszerben szerepel például a $2x + y = 4$ valamint a $2x+y = 5$ egyenlet, hiszen ezeket egyszerre persze, hogy nem lehet kielégíteni. Még azt is könnyen fel tudjuk ismerni, hogy az $x + y = 1$, $x + 2y = 3$ és $2x + 3y = 100$ rendszer nem oldható meg, hiszen az első két egyenlet összege $2x + 3y = 4$ ellentmond a harmadik egyenletnek. Sajnos nagyon is jól el lehet képzelni, hogy a rendszer megoldhatatlanságára nincs mindig ilyen egyszerű ok, márpedig a célunk éppen az lenne, hogy az összes ilyen okot feltárjuk.

Mindenesetre ha az A sorainak létezik olyan lineáris kombinációja, amely a sorokat a 0 vektorba viszi, de a jobboldali b vektort egy nem-nulla számba, azaz ha létezik olyan y, amelyre $yA = 0$, $yb \neq 0$, akkor egész biztos, hogy az $Ax = b$ nem oldható meg, mert egy x megoldásra azt kapnánk, hogy $0 = (yA)x = y(Ax) = yb \neq 0$. Vagyis, ha egy ilyen y-t viszünk el a főnöknek, akkor behelyettesítéssel rögtön tudja ellenőrizni, hogy $yA = 0$, $yb \neq 0$ és így biztos lehet abban, hogy az $Ax = b$-nek valóban nincs megoldása. A fő kérdés az, hogy ha a konkrét $Ax = b$ lineáris egyenletrendszernek nincs megoldása, akkor erre vajon mindig létezik-e ilyen y tanúsítvány. Éppenséggel elvileg nagyon is elképzelhető, hogy bizonyos rendszerek megoldhatatlanságáért valóban egy ilyen y felel, mások megoldhatatlanságáért viszont valami más szerkezeti baj. Az élet napfényes oldalához tartozik a lineáris algebrának az az elegáns tétele, miszerint a konkrét esetben nem lehet másféle baj, azaz:

1.1.1. Tétel. *Egy $Ax = b$ lineáris egyenletrendszer akkor és csak akkor NEM oldható meg, ha létezik olyan y, amelyre $yA = 0$ és $yb \neq 0$.*

Megjegyezzük, hogy itt a tétel csak egy gondolat illusztrációjául szolgál. Bizonyítása az 2.2.10 tételnél szerepel majd.

Térjünk most vissza a gráfszínezési problémára és az előzőek mintájára vizsgáljuk meg, hogy itt létezik-e leállási szabály, azaz egy olyan eszköz, amelynek segítségével egy rendelkezésünkre bocsátott válasz helyességét gyorsan ellenőrizni tudjuk anélkül, hogy a válaszhoz vezető számítás részleteit át kellene vizsgálnunk. Mindenesetre, ha valaki betoppan a gráf pontjainak egy k-színezésével, akkor azt gyorsan (azaz polinom időben) tudjuk ellenőrizni, hogy a szóbanforgó színezés jó-e, vagyis azt, hogy minden él két végpontja tényleg különböző színt kapott-e. Ha azonban az a válasz, hogy az illető gráf pontjainak nem létezik jó k-színezése, úgy ezt nem tudjuk másként ellenőrizni, mint a feladat újra megoldásával.

Nézzük a következő három tételt. Az első egy gráf 2-színezhetőségére, a második a 3-színezhetőségre, végül a harmadik az általános k-színezhetőségre ad szükséges és elegendő feltételt.

1.1.2. Tétel. *Egy $G = (V, E)$ gráf pontjainak akkor és csak akkor létezik jó 2-színezése, ha a gráfban nincs páratlan élszámú* (röviden páratlan) *kör.*

Biz. Ha egy kör pontjainak létezik jó piros-kék színezése, akkor az egyik piros pontjától körbe menve a pontok színei váltakozva piros-kék színűek, tehát a kör összesen páros sok pontból áll, vagyis egy páratlan kört nem lehet 2 színnel jól színezni. Emiatt egy páratlan kört tartalmazó gráfot sem lehet, tehát a feltétel szükséges.

Az elegendőség igazolásához feltehető, hogy a gráf összefüggő, mert különben a bizonyítást külön végezhetjük az összefüggő komponensekre. Tekintsük a gráfnak egy tetszőleges F feszítő fáját, és nézzük ennek egy kiválasztott s pontjától a gráf pontjainak fabeli távolságát. Színezzük meg a pontokat két színnel aszerint, hogy ez a távolság páros vagy páratlan. Ha minden él két végpontja különböző színű, akkor megkaptuk a jó 2-színezést. Ha mondjuk az uv él két végpontja egyszínű, akkor az u-ból illetve a v-ből a fában s-be vezető $P_1[u, s]$ illetve $P_2[v, s]$ útnak létezik egy egyértelmű első közös pontja. Ezt z-vel jelölve, a $P_1[u, z]$ és a $P_2[v, z]$ részutak élszáma azonos paritású, tehát az általuk és az uv él által alkotott kör páratlan elemszámú. •

1.1.3. Tétel. *Egy $G = (V, E)$ gráf pontjainak akkor és csak akkor létezik jó 3-színezése, ha az éleinek van olyan aciklikus irányítása, amelyben minden egyirányú út hossza (= élszáma) legfeljebb 2.*

[Az aciklikus digráf definícióját lásd alább az 1.2. szakasz elején. Egy G irányítatlan gráf irányításán azt értjük, hogy G minden uv élét az uv vagy a vu irányított élek valamelyikével helyettesítjük.]

Biz. Ha a gráf pontjainak $\{V_1, V_2, V_3\}$ egy jó 3-színezése, azaz minden él különböző V_i osztályok között vezet, akkor az összes élt irányítsuk az alacsonyabb indexű osztály végpontjától a magasabb indexű felé. Ily módon egy olyan aciklikus irányítást kapunk, amelyben nincsen 2-nél hosszabb egyirányú út.

Megfordítva, tekintsük a gráfnak egy aciklikus irányítását, amelyben nincsen 2-nél hosszabb út. Jelölje V_1 azon pontok halmazát, melyekből nem lép ki él, V_2 azokét, melyekből nem lép ki él $V - V_1$-be, és végül V_3 azokét, melyekből nem lép ki él $V - (V_1 \cup V_2)$-be. A konstrukció miatt semelyik él két végpontja sem lehet ugyanabban a V_i-ben. Azt kell csak kimutatnunk, hogy minden csúcs benne van a V_i-k valamelyikében.

A konstrukcióból adódik, hogy minden V_2-beli csúcsból lép ki él V_1-be és minden V_3-beli csúcsból lép ki él V_2-be, és így minden V_3-beli pontból indul ki 2-élű út. Márpedig ha indirekt egy s pont semelyik V_i-ben sincs benne, akkor lép ki belőle él valamely $v \in V_3$-ba, de akkor az sv élt a v-ből induló 2-élű úttal kiegészítve már három élű utat kapnánk, ellentmondásban a tétel feltevésével. •

1.1.4. Tétel. *Egy $G = (V, E)$ gráf pontjainak akkor és csak létezik jó k-színezése, ha létezik egy olyan $\varphi : V \to \{1, 2, \ldots, k\}$ leképezés, hogy minden $uv \in E$ élre $\varphi(u) \neq \varphi(v)$.* •

Ugyan mindhárom tétel szükséges és elegendő feltételt ad, azonban e tételek információ tartalma nagyon is különböző. Az 1.1.4 tétel semmi más, mint egy (fontoskodó) átfogalmazása az eredeti definíciónak, és így teljesen értéktelen.

Az 1.1.2 és 1.1.3 tételek már nem semmitmondóak és bizonyításuk is nagyjából egyforma nehézségű, de még ez a két tétel is jellegében alapvetően eltér egymástól. Az 1.1.2 tétel megadta azt a könnyen ellenőrizhető tanúsítványt (a páratlan kört), amely igazolja egy konkrét gráf 2-színezhetőségének lehetetlenségét. Az 1.1.3 tétel ilyesmivel nem szolgált: nem látszik, hogy mitől volna egyszerűbb (mint ahogy nem is az) a háromélű utat nem tartalmazó aciklikus irányíthatóságot ellenőrizni, mint a jó 3-színezés meglétét. Tehát az 1.1.3 tétel nem tekinthető másnak, mint a 3-színezhetőség egy ekvivalens átfogalmazásának, míg az 1.1.2 tétel a 2-színezhetőség „jó karakterizációja”. Kicsit még jobban megvilágítja a helyzetet, ha az 1.1.2 tételt „kifordítva” fogalmazzuk meg: *Egy gráf akkor és csak akkor* NEM *színezhető kettő színnel, ha tartalmaz páratlan kört.* Ez azért jó karakterizáció, mert nemcsak egy konkrét 2-színezés helyessége ellenőrizhető gyorsan, hanem egy körről is rögtön ellenőrizhető, hogy valóban a gráfban van-e és hogy tényleg páratlan sok éle van.

Azt mondjuk, hogy egy tulajdonság **NP**-ben van (nem-determinisztikusan polinomiális), ha a tulajdonság meglétére létezik polinom időben ellenőrizhető bizonyíték. (FIGYELEM, FIGYELEM, VESZÉLY: az **NP** rövidítés **NEM** a polinom időben való megoldhatóság tagadását jelzi!) Azt mondjuk, hogy egy tulajdonság co-**NP**-ben van, ha a tulajdonság hiányára létezik polinom időben ellenőrizhető bizonyíték. A k-színezhetőség **NP**-ben van (egy megadott színezésről polinom időben könnyű eldönteni, hogy jó-e). Az 1.1.2 tétel szerint a 2-színezhetőség co-**NP**-ben is van, ugyanakkor a 3-színezhetőségről ezt nem tudni (és éppenséggel az az általános vélekedés, hogy nincsen). Egy másik közismert tulajdonság, a gráfok síkbarajzolhatósága szintén **NP**-ben van (egy konkrét síkbarajzolás helyessége könnyen ellenőrizhető) és Kuratowski tétele nyomán co-**NP**-ben is van. Kuratowski tétele ugyanis azt mondja ki, hogy egy gráf pontosan akkor NEM síkbarajzolható, ha tartalmaz felosztott K_5-t (ötpontú teljes gráf) vagy $K_{3,3}$-t (3-ház 3-kút gráf). E gráfokról ugyanis az Euler formula segítségével egyszerű dolog kimutatni, hogy nem síkbarajzolhatók, és ezért ilyen részgráfok jelenléte valóban gyorsan ellenőrizhető igazolványt jelentenek a gráf síkbarajzolásának lehetetlenségére. (Járatosabb vagy különösen éles szemű olvasóink kedvéért megjegyezzük, hogy a síkbarajzolhatóságra csak az jelentene valóban gyorsan ellenőrizhető bizonyítékot, ha a gráf csúcsai nem túl nagy, egészértékű koordinátájú pontokba kerülnek. Nem-triviális tétel biztosítja ennek lehetőségét.)

Vannak olyan tulajdonságok is, amelyekről ránézésre sem az nem világos, hogy **NP**-ben vannak, sem az, hogy co-**NP**-ben. Például, egy gráfot perfektnek neveznek, ha minden feszített részgráfjának a kromatikus száma egyenlő a részgráfban lévő maximális teljes részgráf (:klikk) pontszámával. Lovász kimutatta, hogy a perfektség co-**NP**-ben van és nemrégiben az **NP**-beliségét is igazolták. (E paragrafus a veszteség érzése nélkül kihagyható, ha valaki nem hallott még perfekt gráfokról.)

Figyeljük meg, hogy a P-beli tulajdonságok (problémák) automatikusan **NP**$\cap$co-**NP**-ben vannak, hiszen egy polinomiális algoritmus teljes futása és így a szolgáltatott végeredmény helyessége is polinom időben ellenőrizhető.

Az 1.1.2 tétel fenti bizonyítása könnyen átalakítható algoritmussá, amely polinomiális futásidőben vagy megtalálja a keresett 2-színezést vagy pedig a 2-színezés lehetetlenségét igazoló páratlan kört. Nem ismeretes polinomiális algoritmus egy gráf 3-színezhetőségének eldöntésére. Ráadásul erős jelek utalnak arra, hogy ilyen algoritmus nem is létezhet. Kimutatták ugyanis, hogy a 3-színezhetőség problémája **NP**-teljes abban az értelemben, hogy ha erre létezik polinomiális algoritmus, akkor valamennyi **NP**-beli probléma megoldására is létezik. Márpedig tengernyi egyéb **NP**-teljes feladat van, amelyek egyikére sem ismert polinomiális algoritmus. Néhány **NP**-teljes tulajdonság: a gráfban van Hamilton kör, a gráf élei k ponttal lefoghatók, a gráf élei 3 színnel megszínezhetők, a gráfban létezik legalább k élű vágás.

Fontos megjegyezni, hogy a fentebb bevezetett polinomialitás fogalma a hatékonyság egy lehetséges elméleti megragadása. (Egy másik lehetőség például a legrosszabb eset lépésszámának becslése helyett az átlagos lépésszámot nézni). Tapasztalatok szerint ez legtöbbször egybeesik az algoritmus gyakorlati hatékonyságával, bár nem mindig.

Végül megjegyezzük, hogy a fenti meggondolások ebben a formában csupán a szemléletet orientáló eszmefuttatásoknak tekinthetők, hiszen valójában még azt sem vezettük be, hogy mit is értünk algoritmuson. A Turing gép (amely nem egy fizikailag létező „gép”, hanem egy matematikai definíció) segítségével mindez a Bonyolultságelmélet c. tárgy keretében kerül felépítésre. A helyzet szerkezetileg némileg ahhoz hasonló, mint amikor egy függvény folytonosságáról beszélünk. Egyrészt él bennünk egy szemléletes kép, amely szerint egy függvény akkor folytonos, ha „a ceruza felemelése nélkül” le lehet rajzolni. Ez felel meg az algoritmus fogalmáról élő szemléletes képünkek. Másrészt van a folytonosság formális definíciója, amely a szemléletes folytonosság képet akarja megragadni. Ezzel áll párhuzamban a Turing gép, amely az algoritmus intuitív fogalmát igyekszik formalizálni.

1.2. Gráfok bejárása: elérhetőség

Legyen $D = (V, A)$ irányított gráf (röviden digráf). **Sétá**n egy olyan $W = (v_0, e_1, v_1, e_2, \ldots, e_k, v_k)$ sorozatot értünk, amelyben felváltva következnek pontok és élek úgy, hogy mindegyik e_i él a v_{i-1} pontból vezet a v_i pontba. Az egyetlen v_0 tagból álló $W = (v_0)$ sorozatot is

sétának tekintjük. A séta **zárt**, ha $k > 0$ és $v_0 = v_k$. A szereplő élek száma a séta **hossza**. Azt mondjuk, hogy v_0 a séta kezdőpontja, míg v_k a séta végpontja. Azt mondjuk, hogy D-ben v_k **elérhető** v_0-ból, ha létezik v_0 kezdőpontú és v_k végpontú séta. Amennyiben a sétában nincs ismétlődés, **egyirányú** vagy **irányított út**ról, röviden, útról beszélünk. Egy s-ből t-be vezető utat st-útnak fogunk hívni. Az **egyirányú** vagy **irányított kör** olyan legalább egy élű zárt séta, amelyben a $v_0 = v_k$ egybeeséstől eltekintve a csúcsok mind különbözőek. Ezek szerint, ha e_1 egy v_0 pontban ülő hurok él, akkor (v_0, e_1, v_0) egy egyélű kör. Amennyiben a W séta egy $K = (v_i, e_{i+1}, v_{i+1}, \ldots, e_j, v_j)$ rész-sétája egy kör (ahol $0 \leq i < j \leq k$), úgy azt mondjuk, hogy a séta **tartalmazza** vagy **indukálja** a K kört. (Figyeljük meg, hogy egy séta élei által meghatározott részgráf nem minden egyirányú köre a séta indukált kör.)

Az egyirányú kör fenti definíciójának kis hátránya, hogy kitüntet egy kezdőpontot és megad egy körbejárási irányt is. Egy alternatív megközelítés, ha az egyirányú kört egy olyan digráfnak tekintjük, amely irányítatlan értelemben összefüggő, minden pontjába egy él lép be és egy él lép ki. Egy digráfot **aciklikus**nak mondunk, ha nincsen benne egyirányú kör. Egy **irányítatlan kör** olyan irányítatlan gráf, amely összefüggő és minden pontjának a befoka 2. Egy irányított gráf egy részgráfját **kör**nek nevezzük, ha irányítatlan értelemben kör. Minden legalább három pontú körnek kétféle körbejárása lehetséges. A továbbiakban lerögzítjük az egyiket és az óramutató szerinti körbejárásnak nevezzük. A kör azon éleit, melyek ebbe az irányba mutatnak **előre élek**nek, a fordított irányba mutatókat pedig **hátra élek**nek nevezzük. Megállapodunk abban, hogy egyirányú kör esetén az óramutató szerinti körbejárást aszerint rögzítjük le, hogy minden él előre él legyen.

1.2.1. Állítás. *Ha létezik s-ből t-be séta, akkor létezik út is.*

Biz. Ha a $W = (s = v_0, e_1, \ldots, e_k, v_k = t)$ st-séta maga nem út, akkor létezik W által indukált $K = (v_i, e_{i+1}, v_{i+1}, \ldots, e_j, v_j)$ egyirányú kör. A K kör éleit kihagyva W-ből a $(v_0, \ldots, v_i, e_{j+1}, \ldots, v_k)$ sétát kapjuk. Ezt a redukciós lépést mindaddig ismételhetjük, amig az aktuális st-séta indukál kört. A végső séta nem indukál kört, azaz st-út. Miután minden redukciónál csökken a séta élszáma, legfeljebb k körkihagyás után megkapjuk a keresett st-utat. •

Azt mondjuk, hogy az állítás bizonyításában kapott P st-út a W séta **egyszerűsítésével** áll elő. (Figyelem: egyszerűsítésnél a szóbanforgó K kör éleit nem a digráfból hagyjuk el, hanem csak a sétát definiáló sorozatból vágjuk ki. Kényelmesen előfordulhat ugyanis, hogy a K kör egy élét a séta később még használni fogja, tehát a gráfból nem szabad kihagyni.)

1.1. Gyakorlat. *Mutassuk meg, hogy egy séta egyszerűsítésével kapott út függhet a redukcióban használt körök választásától.*

Egy olyan irányított $F = (S, E)$ fát, amelynek minden pontja elérhető egyirányú úton s-ből s-**fenyő**nek nevezünk. Azt mondjuk, hogy F feszíti S-t. Ha a fenyő részgráfja D-nek és az egész V halmazt tartalmazza, **feszítő** s-fenyőről beszélünk. **Fenyves**nek hívunk egy olyan irányított erdőt, melynek komponensei fenyők.

1.2. Gyakorlat. *Egy irányított fa akkor és csak akkor s-fenyő, ha az $s \in S$ pont befoka nulla a többi ponté pedig egy.*

1.3. Gyakorlat. *Egy s-et tartalmazó digráf akkor és csak akkor s-fenyő, ha az s pontból kiindulva elő lehet úgy állítani irányított élek egyenkénti hozzávételével, hogy az aktuálisan hozzáadott él hegye új pont, míg a töve már meglévő.*

Kérdés, hogy miként lehet hatékonyan eldönteni, hogy egy D digráfban létezik-e st-út? Ez valójában két kérdést is jelent. Egyrészt konstruálnunk kell egy st-utat, ha ilyen egyáltalán létezik. Ha viszont nem létezik st-út, úgy ennek egy könnyen ellenőrizhető tanúsítványát kell bemutatnunk. A trükk abból áll, hogy nem csupán a t csúcs s-ből való elérhetőségét vizsgáljuk, hanem egyszerre valamennyi csúcsét.

1.2.1. Tétel. *Jelölje S a $D = (V, A)$ digráfban azon csúcsok halmazát, amelyek az s csúcsából elérhetők. Ekkor S minden valódi, s-et tartalmazó S' részhalmazából vezet ki él, de S-ből nem. Továbbá létezik S-t feszítő s-fenyő.*

Biz. Ha indirekt egy uv él kilépne S-ből, akkor v pont is elérhető volna, hiszen $u \in S$ definíció szerint az, vagyis létezik P út s-ből u-ba, amihez az uv élt hozzávéve egy sv-utat kapnánk, ellentmondásban azzal, hogy v nem elérhető. Ha az S'-ből nem lépne ki él, akkor semelyik S'-n kívüli pont nem volna elérhető s-ből, ellentmondásban S definíciójával.

Legyen F egy nem bővíthető s-fenyő. Állítjuk, hogy ennek S' csúcshalmaza éppen S. Mivel F minden pontja elérhető s-ből, így $S' \subseteq S$. Ha indirekt $S' \subset S$ állna, úgy az első rész szerint lép ki egy uv él S'-ből. De ezt F-hez véve egy nagyobb fenyőt kapnánk, ellentmondásban F maximális választásával. •

Hogyan lehet algoritmikusan megkonstruálni a szóbanforgó S halmazt és F fenyőt? Az alábbi címkézési technika segít. A digráf minden v pontjához tartozzék egy R-címke (**R**each = elér), amely azt mutatja, hogy az algoritmus futásának egy adott pillanatában v-t már elértük s-ből egy út mentén vagy sem. Amennyiben nem, akkor az R-címke tartalma NEM ELÉRT. Ha v-t már elértük, akkor R-címkéjének tartalma ELÉRT valamint azon útnak a legutolsó uv éle, amelyen elértük v-t. Az egyetlen kivétel maga az s pont, amelynek R-címkéje mindig ELÉRT.

Ezen kívül minden pontban fenntartunk egy S-címkét (**S**can = letapogat, átvizsgál), amely azt jelzi, hogy az adott pillanatban a v pontból vajon már az összes továbblépési lehetőséget átvizsgáltuk-e (azaz valamennyi $vu \in A$ élre az u csúcs már elért-e), amikoris az S-címke tartalma ÁTVIZSGÁLT, vagy pedig még van át nem vizsgált vx él. Kezdetben minden S-címke tartalma NEM ÁTVIZSGÁLT.

Az algoritmus általános lépésében kiválasztunk egy már elért, de még át nem vizsgált u pontot (ami induláskor persze csak az s pont lehet) és eldöntjük, hogy van-e olyan uv éle a digráfnak, hogy v még nem elért. Amennyiben nincs, akkor az u pontot ÁTVIZSGÁLT-nak deklaráljuk és az eljárást iteráljuk. Ha viszont találunk ilyen v pontot, akkor v-t ELÉRT-nek nyilvánítjuk, az R-címkéjébe betesszük az uv élt, és ismét az eljárást iteráljuk. Az algoritmus akkor ér véget, amikor már minden elért pont átvizsgált lesz.

Egyszerű feladat annak igazolása, hogy az algoritmus lefutása után az ELÉRT pontok S halmazából nem vezet kifelé él, továbbá, hogy az elért pontok R-címkéjébe írt élek egy s gyökerű fenyőt alkotnak, melynek ponthalmaza S.

Az eljárás az S meghatározása után folytatható egy tetszőleges S-ben nem szereplő s_2 pont gyökérnek való kijelölésével. Végül egy fenyvest kapunk, melynek gyökerei $s_1 := s, s_2, \ldots$, és amely az összes pontot tartalmazza.

Az eljárásról annyit érdemes még tudni, hogy megfelelő adatstruktúra alkalmazásával a futási idő lineáris, azaz az élek számával arányos. További megjegyzés, hogy az eljárás irányítatlan gráfokra is alkalmazható.

1.4. Feladat. *Egy páros gráf élei pirossal és kékkel vannak színezve. Fejlesszünk ki* *algoritmust annak meghatározására, hogy a gráf két megadott pontja között létezik-e alternáló piros-kék út.*

1.2.1. Szélességi keresés

Az algoritmus futtatása során szabadságunk van az aktuális már elért, de még át nem vizsgált pont kiválasztásában. Egy lehetséges stratégia azt a még nem átvizsgált pontot választani, amelyiket leghamarabb értük el. Ebben az esetben **szélességi keresés**ről beszélünk (breadth first search: BFS). A BFS például alkalmas arra, hogy segítségével a pontok s-től való távolságát egyszerűen meghatározzuk. Csupán azt a csekély módosítást kell a fenti algoritmusban végrehajtani, hogy minden v pontra fenntartunk egy $\ell(v)$ változót is, amely a már elért pontoknál megmondja az s-től való távolságot. Kezdetben ez az s-ben 0, mindenütt másutt ∞. Amikor az algoritmus során egy v pontot az uv él mentén u-ból elérünk, akkor az $\ell(v)$ értéket $\ell(u) + 1$-re állítjuk be. Valójában ez az algoritmus speciális esete Dijkstra később ismertetésre kerülő eljárásának, amely általában nem-negatív súlyozás esetén számítja ki egy v pontnak s-től való távolságát.

Gyakorlatok

1.5. *Igazoljuk, hogy a BFS algoritmus helyesen határozza meg az s-től való távolságot.*

1.6. *Igazoljuk, hogy irányítatlan gráfban a távolság függvény kielégíti a háromszög egyenlőtlenséget.*

1.7. *Legyen S és T a D digráf pontjainak két részhalmaza. Miként lehet a fenti algoritmus segítségével eldönteni, hogy létezik-e út S-ből T-be?*

1.2.2. Mélységi keresés

A címkézési eljárásban egy másik lehetséges stratégia az, amikor az algoritmus azt a még át nem vizsgált pontot választja ki, amelyiket a legkésőbb értük el. Ebben az esetben az eljárást **mélységi keresés**nek nevezzük (depth first search: DFS). A DFS-nél minden pontnak van egy **elérési időpont**ja, amikor a pont ELÉRT lesz (tehát amikor az algoritmus először találkozik az illető ponttal), és van egy **elhagyási időpont**ja, amikor a pont ÁTVIZSGÁLT lett (vagyis amikor a keresés utoljára találkozik az illető ponttal). Mind a kettő meghatározza a pontok egy sorrendjét: az **elérési** és az **elhagyási** sorrendet. A két sorrend összefésülésével kapjuk a pontok **kezelési** sorendjét. Tehát a kezelési sorrendben minden pont kétszer fordul elő, és a két előfordulás közötti ponthalmaz, amint az könnyen belátható, két különböző pontra vagy diszjunkt vagy tartalmazkodó. Az ilyen sorozatot **laminális**nak nevezzük. Következik, hogy ha s-ből minden pont elérhető, akkor a sorozat első és utolsó tagja az s gyökérpont. Egyébként egy lamináris sorozat, amelynek első és utolsó tagja s, mindig egyértelműen leír egy s gyökerű fenyőt. Ezt rekurzívan definiálva úgy kaphatjuk meg, hogy veszünk a sorozatnak egy x, y, y alakú három egymást követő eleméből álló részét [ilyen van a laminaritás miatt], a két y-t kihagyjuk, a maradékhoz megkonstruáljuk a fenyőt, és végül hozzávesszük az xy élt.

1.8. Gyakorlat. *Igazoljuk, hogy legalább háromtagú lamináris sorozatnak (amelyben minden elem kétszer fordul elő) van x, y, y alakú három egymást követő elemből álló része.*

DFS fenyők

A mélységi kereséssel kapott fenyőt (amely persze nem egyértelmű) **DFS vagy mélységi fenyő**nek hívjuk. Irányítatlan esetben DFS vagy mélységi fáról beszélünk. A DFS fenyő fontos tulajdonsága, hogy minden xy élre az y elérési időpontja megelőzi az x elhagyási időpontját. Speciálisan, összefüggő irányítatlan gráf mélységi fájához nem tartozik keresztél. (Egy s gyökerű irányítatlan fa esetén egy xy nem-fa élt akkor hívunk **keresztél**nek, ha a fában az x és y-t összekötő egyértelmű út s-hez (a fában) legközelebbi pontja különbözik x-től és y-tól.)

A DFS fának több érdekes alkalmazása van. Segítségével lehet például lineáris időben egy 2-élösszefüggő gráf erősen összefüggő irányítását megkapni: vegyünk egy s gyökerű mélységi fát, irányítsuk a fa éleit s-től kifelé, a nem-fa éleket pedig s felé. Mivel nincs keresztél, így minden élt irányítottunk.

1.9. Feladat. *Igazoljuk, hogy ha a gráf 2-élösszefüggő, akkor az így kapott irányítás erősen összefüggő.*

Topologikus sorrend

A DFS egy másik alkalmazása aciklikus digráfban a pontok ún. topologikus sorrendjének meghatározására szolgál. A digráf csúcsainak egy sorrendjéről akkor mondjuk, hogy **topologikus**, ha minden él előre mutat, azaz a töve a sorrendben megelőzi hegyét (=fejét). Egy digráfot akkor neveztünk aciklikusnak, ha nem tartalmaz egyirányú kört. Azt, hogy egy digráf nem aciklikus, egy konkrét egyirányú körének bemutatásával tanúsíthatjuk. Milyen gyorsan ellenőrizhető igazolvány adható a digráf aciklikusságának tanúsítására? Erre ad választ a következő egyszerű, de hasznos tétel.

1.2.2. Tétel. *Egy $D = (V, A)$ digráf akkor és csak akkor aciklikus, ha pontjainak létezik topologikus sorrendje, azaz egy olyan $v_1, v_2, \dots, v_n$ sorrend, amelyben minden él korábbi pontból későbbibe vezet.*

Biz. Egy egyirányú kör pontjainak nem létezhet topologikus sorrendje, hiszen a kör minden pontjának pozitív a befoka. Emiatt egy egyirányú kört tartalmazó digráfnak se létezhet topologikus sorrendje, vagyis a feltétel szükséges.

Az elegendőséghez figyeljük meg, hogy egy aciklikus digráfnak létezik forráspontja, vagyis olyan pontja, amibe nem lép be él. Ha ugyanis minden pontba lép be él, akkor egy pontból a belépő él mentén visszafelé indulva a fordított sétát mindig tudnánk folytatni és előbb-utóbb egy kört kapnánk. Válasszunk ki egy v_1 forráspontot és legyen ez a sorrend első pontja. A $D-v_1$ digráf is aciklikus, ennek is létezik egy v_2 forráspontja. Ezt az eljárást folytatva megkapjuk a keresett $v_1, v_2, \dots, v_n$ topologikus sorrendet. •

(Megjegyezzük, hogy megszámlálhatóan végtelen sok pontból álló aciklikus digráf pontjainak nem feltétlenül van topologikus sorrendje. Ezt példázza az a digráf, amelynek csúcsai a racionális számok és az u, v racionális számokra uv akkor él, ha $u < v$.)

A bizonyításból adódó algoritmus egyetlen forráspontot $O(m)$ lépésszámban tud megtalálni, így a topologikus sorrend megkeresésének összlépésszáma $O(mn)$. Mélységi keresés okos alkalmazásával a teljes topologikus sorrendet $O(m)$ lépésben meg lehet találni. Ennek érdekében feltehetjük, hogy a digráfnak van olyan s pontja, ahonnan minden más pont elérhető. Valóban, mert ha nem ez a helyzet, akkor adjunk a digráfhoz egy új s pontot, és vezessünk s-ből minden eredeti pontba élt. Így aciklikus digráfot kapunk, amely pontjainak topologikus sorrendjéből az újonnan hozzáadott s-t kihagyva megkapjuk az eredeti digráf pontjainak egy topologikus sorrendjét.

1.10. Gyakorlat. *Dolgozzunk ki algoritmust annak eldöntésére, hogy két közös csúcshalmazon lévő digráfnak létezik-e közös topologikus sorrendje.*

1.11. Feladat. *Igazoljuk, hogy aciklikus digráf elhagyási sorrendjének megfordítása topologikus sorrendet ad.*

1.12. Feladat. *Igazoljuk, hogy tetszőleges hurokmentes digráf élhalmaza felbontható két aciklikus digráf egyesítésére.*

Erősen összefüggő komponensek

Gráfelméletben igazolják, hogy tetszőleges $D = (V, A)$ irányított gráf esetén, ha két pontot ekvivalensnek tekintünk amennyiben mindkettő elérhető a másikból egyirányú úton, úgy ekvivalencia relációt kapunk. Érvényes, hogy az ekvivalencia osztályai erősen összefüggő részgráfokat feszítenek, amelyek mindegyikét egy-egy pontra összehúzva aciklikus digráfot kapunk. Az ekvivalencia osztályok által feszített digráfokat szokás a D digráf erősen összefüggő komponenseinek nevezni.

1.13. Gyakorlat. *Igazoljuk, hogy egy mélységi kereséssel kapott feszítő fenyves olyan, hogy a digráf bármely C erősen összefüggő komponensére megszorítva C-nek feszítő fenyőjét adja (amelynek gyökere a C-nek a keresés által legelőször elért pontja).*

A topologikus sorrend meghatározásánál kicsit ravaszabb módon lehet egy digráf erősen összefüggő komponenseit előállítani. Ismét feltehetjük, hogy egy s pontból minden pont elérhető. Az algoritmus két külön fázisból áll. Az első fázisban mélységi kereséssel határozzuk meg az elhagyási sorrendet. A második fázisban tetszőleges keresési eljárással (ami lehet a DFS is, de ezt már nem használjuk ki a bizonyításban) határozzunk meg egy fordított fenyvest úgy, hogy a soron következő gyökérpont mindig az első fázisban kapott elhagyási sorrend még nem szerepelt legutolsó tagja legyen. (Fordított fenyő alatt olyan irányított fát értünk, amelyben a gyökértől eltekintve minden pont kifoka egy, míg a gyökéré nulla. Fordított fenyves olyan irányított erdő, amelynek minden komponense fordított fenyő.)

1.14. Feladat. *Igazoljuk, hogy a második fázisban kapott fordított fenyves komponensei éppen a digráf erősen összefüggő komponensei lesznek.*

1.3. Optimális utak és potenciálok

1.3.1. Bevezetés

Tekintsük a következő gyakorlati jellegű kérdéseket.

1. *Legrövidebb utak.* A közkedvelt GPS (Global Positioning System) műholdak segítségével meghatározza a pozíciónkat, majd kiszámítja, hogy merre tudunk a megadott célpontba a leghamarabb vagy a legolcsóbban eljutni. Hogyan lehet egy ilyen leggyorsabb (vagy legrövidebb) utat hatékonyan kiszámolni?

2. *Házépítés.* Egy családi ház építkezése elemi munkafázisokra bomlik (alapok kiásása, betonozás, 1. szint felhúzása, 2. szint felhúzása, tető építés, belső vakolás, festés, ablakok, stb). Minden fázisnak adott a végrehajtási ideje, továbbá egy megelőzési reláció, amely azt mondja meg, hogy például a tető építése csak az alapozás befejezése után következhet. Kérdés, hogy miként ütemezzük a munkákat, ha célunk a mihamarabbi befejezés.

3. *Nyaraló kiadás.* Kiadjuk álomszép balatoni nyaralónkat a nyári hónapokra. A nyaralót egyszerre egy család használhatja és a jelentkezők megadják, hogy melyik időintervallumra szeretnék kivenni a házat. A mi feladatunk a jelentkezők közül úgy választani, hogy a nyaraló a lehető legtöbb napra ki legyen adva.

4. *Monoton növő részsorozat.* Adott számsorozatnak válasszunk ki hatékonyan maximálisan sok tagját, melyek monoton növő részsorozatot alkotnak.

5. *Optimális közös részsorozat.* Két betűsorozatnak válasszunk ki egy maximális közös részsorozatát.

E látszólag távolfekvő problémákról kimutatható, hogy matematikai gyökerük közös: legolcsóbb utat keresni egy irányított gráfban. Célunk e problémakör áttekintése.

Tegyük fel, hogy a $D = (V, A)$ n pontú és m élű hurokmentes és párhuzamos élt nem tartalmazó irányított gráf élein adott egy $c : A \to \mathbb{R}$ költség- (vagy másnéven súly-) függvény. Egy P út, séta vagy kör $\widetilde{c}(P)$-vel jelölt költségén a P éleinek költségösszegét értjük. A digráf egy P útjáról azt mondjuk, hogy c-**legolcsóbb** vagy röviden **legolcsóbb**, ha a P kezdőpontjából a végpontjába nem létezik D-ben P-nél olcsóbb út.

1.15. Gyakorlat. *Igaz-e, hogy legolcsóbb út bármely részútja is legolcsóbb út?*

Egyik célunk adott s és t csúcsokra minimális költségű, másnéven legolcsóbb s-ből t-be vezető (irányított vagy más néven egyirányú) utat, röviden st-utat keresni. Kiderül, hogy kényelmesebb azzal a többet kívánó problémával foglalkozni, amikor egy rögzített s gyökérpontból az összes többi v pontba szimultán kell legolcsóbb utat kiszámítani. Az olyan pontokat, amelyek egyáltalán nem érhetők el s-ből, kihagyhatjuk, mert ez a többi pont elérhetőségét nem befolyásolja. Emiatt amikor a legolcsóbb s-ből induló utak felől érdeklődünk mindig feltehetjük, hogy s-ből minden pont elérhető.

$$\text{A legolcsóbb } sv\text{-út költségét jelölje } \mu_c(v).$$

Az s rögzítettsége miatt e jelölésben az s nem is szerepel. A digráf egy s gyökerű F fenyőjéről azt mondjuk, hogy a **legolcsóbb utak fenyője**, ha az F minden v pontjára az F-ben lévő P egyértelmű sv-út költsége $\mu_c(v)$, azaz P a digráfnak egy legolcsóbb sv-útja. Amennyiben a c negatív is lehet, úgy a legolcsóbb út feladat már az azonosan -1 költségfüggvény esetén is **NP**-teljes, ugyanis ekkor magában foglalja a Hamilton-út problémának az előírt végpontú változatát, ami a tetszőleges végpontú Hamilton-út problémához hasonlóan **NP**-teljes. Emiatt a legolcsóbb út feladatot nem vizsgáljuk teljes általánosságban, amikor a digráf és a költségfüggvény is tetszőleges.

Először bemutatunk két speciális költségfüggvényt, amelyek esetén a legolcsóbb utak megkeresésére hatékony algoritmus adható. Az első esetben D aciklikus és c tetszőleges, míg a másikban D tetszőleges és c nem-negatív. A következő részben pedig megadjuk majd e két speciális eset közös általánosítását is (amikor D tetszőleges, de nem létezik negatív összköltségű egyirányú kör).

1.3.2. Legolcsóbb utak aciklikus digráfban

Tegyük fel, hogy a D digráf aciklikus és az s pontból mindegyik másik pontba vezet egyirányú út, másszóval s-ből minden más pont elérhető. Legyen a $c : A \to \mathbb{R}$ súlyozás (vagy költségfüggvény) tetszőleges, tehát c-nek lehetnek pozitív és negatív komponensei is. Ez azért jó, mert a c negálása révén nem csupán a minimális költségű, hanem a maximális költségű (súlyú) út problémát is meg tudjuk oldani: tetszőleges irányított gráf esetén ez a feladat **NP**-teljes volt.

Aciklikus digráfban a következő egyszerű direkt eljárással fel lehet építeni a legolcsóbb utak fenyőjét. Az 1.2.2 szakaszban láttuk, hogy miképp lehet egy topologikus sorrendet lineáris időben meghatározni. Tegyük fel, hogy az $s = v_1, v_2, \dots, v_n$ topologikus sorrend első $j-1$ pontja által feszített részgráfban már meghatároztuk a legolcsóbb utak egy F_{j-1} fenyőjét a $\mu_c(v_i)$ költségekkel egyetemben $(1 \le i \le j-1)$. Tekintsük a sorrendben következő v_j pontot. Miután v_j-be csak j-nél kisebb indexű pontból vezet él, a legolcsóbb sv_j-út költségét a

$$\mu_c(v_j) = \min\{\mu_c(v_i) + c(v_iv_j) : v_iv_j \in A\} \tag{1.1}$$

formula adja. Továbbá, ha v_iv_j jelöli azt az élt (pontosabban az egyik olyan élt), amelyen a minimum felvétetik, akkor $F_j := F_{j-1} + v_iv_j$ a legolcsóbb utak fenyője az első j csúcson. (ha több minimalizáló él van, mindegy, hogy melyikkel növeljük a fenyőt.) Ezt a rekurziót $j = 1, \dots, n$-re követve a végül kapott F_n feszítő fenyő egy legolcsóbb utak fenyője lesz. Miután a $\mu_c(v_j)$ értékek illetve a fenyőbe kerülő v_iv_j élek kiszámításához a digráf minden élét egyszer kell csak tekintetbe venni, az algoritmus lépésszáma $O(m)$, felhasználva, hogy a topologikus sorrendet is $O(m)$ lépésben lehetett megkapni.

Az (1.1) formulából adódik, hogy minden $v_iv_j \in A$ élre

$$\mu_c(v_j) \le \mu_c(v_i) + c(v_iv_j),$$

míg ha v_iv_j az F_n fenyő éle, akkor itt egyenlőség teljesül:

$$\mu_c(v_j) = \mu_c(v_i) + c(v_iv_j). \tag{1.2}$$

Az algoritmus következményeként kapjuk az alábbi tételt.

1.3.1. Tétel. *$D = (V, A)$ aciklikus digráfban, amelyben minden pont elérhető s-ből, tetszőleges költségfüggvényre létezik legolcsóbb utak fenyője.* •

A legolcsóbb út költségére vonatkozik az alábbi min-max tétel.

1.3.2. Tétel. *Legyen $c : A \to \mathbb{R}$ a $D = (V, A)$ aciklikus irányított gráf élhalmazán egy tetszőleges költségfüggvény, és tegyük fel, hogy létezik st-út. Az s-ből t-be vezető utak költségének $\mu_c(t)$ minimuma egyenlő a*

$$\max\{\pi(t) - \pi(s) : \pi : V \to \mathbb{R},\ \pi(v) - \pi(u) \le c(uv),\ uv \in A\}$$

értékkel. Ha c egészértékű, az optimális π is választható egészértékűnek.

Biz. Tetszőleges π-re és $P = \{s = v_0, v_1, \ldots, v_k = t\}$ st-útra

$$\widetilde{c}(P) = \sum_i c(v_i v_{i+1}) \geq \sum_i [\pi(v_{i+1} - \pi(v_i)] = \pi(t) - \pi(s), \tag{1.3}$$

vagyis a tételben a max $\leq$ min irány következik. (Ehhez a egyenlőtlenséghez az aciklikusságra nincs is szükség).

A fordított irányú egyenlőtlenséget elég bebizonyítani abban a speciális esetben, amikor minden csúcs elérhető s-ből. Ha esetleg nem ez a helyzet, akkor egy új s' pontot a digráfhoz veszünk egy 0 költségű $s's$ éllel valamint minden $v \in V$-re egy M költségű $s'v$-éllel, ahol M kellően nagy. Az így nyert digráfot és költségfüggvényt jelölje D' és c'. Ekkor D'-ben minden $v \in V$ pont elérhető s'-ből és $\mu(v) = \mu'(v)$, ahol $\mu'(v)$ a legolcsóbb $s'v$-út c'-költsége D'-ben. Legyen most π' egy olyan függvény V'-n, amelyre $\pi'(v) - \pi'(u) \leq c'(uv)$ minden $uv \in A'$ élre és $\pi'(t) - \pi'(s') = \mu'(t)$. Ekkor a $\pi := \pi'|V$ függvényre (ami tehát π' megszorítása V-re) $\pi(v) - \pi(u) \leq c(uv)$ minden $uv \in A$ élre. Továbbá, $\pi(s) - \pi'(s') = \pi'(s) - \pi'(s') \leq c'(s's) = 0$ miatt $\pi(s) \leq \pi'(s')$ és ezért $\mu(t) \geq \pi(t) - \pi(s) = \pi'(t) - \pi(s) \geq \pi'(t) - \pi'(s') = \mu'(t) = \mu(t)$, amiből $\mu(t) = \pi(t) - \pi(s)$. Tehát, ha D'-re igaz a tétel, akkor D-re is, és emiatt feltehetjük, hogy D-ben minden csúcs elérhető s-ből.

Tekintsük a fenti algoritmus által szolgáltatott P utat, azaz az F_n fenyőben lévő egyértelmű st-utat. Minden $v \in V$-re legyen $\pi(v) := \mu_c(v)$ a legolcsóbb sv-út költsége. Ekkor egyrészt minden uv élre egy legolcsóbb su utat az uv éllel kiegészítve (az aciklikusság miatt) egy sv-utat kapunk, és ezért $\pi(v) \leq \pi(u) + c(uv)$, azaz $\pi(v) - \pi(u) \leq c(uv)$, másrészt (1.2) miatt P minden uv élére $\pi(v) - \pi(u) = c(uv)$, vagyis (1.3)-ben egyenlőség áll, és így $\widetilde{c}(P) = \pi(t) - \pi(s)$. •

Alkalmazásokban előfordul, hogy aciklikus digráfban minimális helyett maximális súlyú st-utat kell keresni. A fenti eljárást ekkor a $-c$ költségfüggvényre kell alkalmazni, de az eredeti c nyelvén közvetlenül is megfogalmazhatjuk, a következőképpen.

Tegyük fel, hogy a topologikus sorrend első $j-1$ pontja által feszített részgráfban már meghatároztuk a legsúlyosabb utak egy F_{j-1} fenyőjét a legsúlyosabb sv_i-út $\tau_c(v_i)$-vel jelölt súlyával egyetemben ($1 \leq i \leq j-1$). A sorrendben következő v_j pontra legyen $\tau_c(v_j) := \max\{\tau_c(v_i) + c(v_i v_j) : v_i v_j \in A\}$ és legyen $F_j := F_{j-1} + v_i v_j$, ahol a $v_i v_j$ egy olyan él, amelyen a maximum felvétetik.

Fogalmazzuk meg az 1.3.2 tétel ellenpárját erre az esetre. A keveredés elkerülése érdekében π helyett τ-t használunk, és az utána következő alkalmazás érdekében felcseréljük a két oldalt.

1.3.3. Tétel. *Legyen $c : A \to \mathbb{R}$ a $D = (V, A)$ aciklikus irányított gráf élhalmazán egy tetszőleges súlyfüggvény, és tegyük fel, hogy az s pontból minden pontba vezet egyirányú út. Ekkor*

$$\min\{\tau(t) - \tau(s) : \tau : V \to \mathbb{R},\ \tau(v) - \tau(u) \geq c(uv), uv \in A\}$$
$$= \max\{\widetilde{c}(P) : P \text{ út } s\text{-ből } t\text{-be}\}. \quad \bullet$$

Egy projekt ütemezési feladat: a PERT módszer

Alkalmazásokban előfordul, hogy nem elsősorban az optimális útra vagyunk kíváncsiak, hanem inkább az optimális τ függvényre. Tekintsük a következő ütemezési feladatot.

Egy projekt különféle elemi feladatok elvégzéséből áll, melyeknek előre adott a végrehajtási ideje. (Például házépítésnél az alapok kiásása, betonozás, első szint felhúzása, második szint felhúzása, tető építés, belső vakolás, vízcső szerelés, villanyvezetékek, festés, ablakok, stb). Tudjuk továbbá, hogy technológiai előírások miatt bizonyos részfeladatok megelőznek másokat (az alapozás a fürdőszoba csempézés előtt van), azaz a részfeladatok halmazán adott egy részbenrendezés. Kérdés, hogy mi az a legrövidebb idő, amely alatt a teljes projekt elvégezhető azon kikötés mellett, hogy az egyes részfeladatok kezdési időpontját úgy megadni, hogy minden feladat kezdésére a részbenrendezésben őt megelőzők már elkészüljenek.

A megoldáshoz készítsünk el egy D irányított gráfot a következőképpen. Mindegyik z részfeladatot reprezentáljuk egy $u_z v_z$ irányított éllel, amelynek súlya legyen a z végrehajtási ideje. Amennyiben a z feladat technológiailag megelőzi az y feladatot, úgy vegyünk be D-be egy $v_z u_y$ élt 0 súllyal. Végül adjunk a digráfhoz egy s forráspontot, amelyből minden u_x pontba vezessünk 0 súlyú élt, és adjunk egy t nyelőpontot, amelybe minden v_x pontból vezessünk 0 súlyú élt. Feladatunk olyan $\tau(v)$ időpontok kijelölésével ekvivalens, amelyekre $\tau(s) = 0$, $\tau(t)$ minimális és minden uv élre a $\tau(v) - \tau(u)$ időpont-különbség legalább akkora, mint az él súlya.

Az 1.3.3 tétel pontosan erre a kérdésre adott választ: *a projekt végrehajtásának minimális össz-ideje (vagyis $\tau(t) - \tau(s)$ minimuma) egyenlő a legsúlyosabb st-út súlyával.* Egy ilyen utat szoktak néha **kritikus út**nak nevezni, míg az előzőekben leírt minimális út problémának az itteni feladatra adaptált változatát kritikus út módszernek (angolul **PERT**: project evaluation and review technique).

Az 1.3.3 tétel egy szemléletes interpretációja révén a főnökünket rögtön meg tudjuk arról győzni, hogy az általunk javasolt optimális ütemezés (amit tehát kezdési időpontokat meghatározó τ függvény ír le) valóban az elvileg legkorábbi befejezést garantálja, hiszen már a kritikus P úton lévő elemi munkák elvégzéséhez is annyi idő kell, mint a P út össz-súlya (vagyis a P úton lévő elemei munkák össz-ideje), márpedig mi a teljes projektet is ennyi idő alatt le tudjuk bonyolítani, és így az ütemezés szükségképpen optimális.

Röviden emlékeztetünk a részbenrendezett halmaz (partially ordered set, poset) fogalmára. Azt mondjuk, hogy a $(P, \preceq)$ pár egy részbenrendezett halmaz, ha $\preceq$ egy reláció a P halmaz elemein, amelyre (A) $x \preceq x$, (B) $x \preceq y$ és $x \neq y$ esetén (amit $x \prec y$ rövidít) $y \not\preceq x$, és (C) $x \preceq y$ és $y \preceq z$ esetén $x \preceq z$. A részbenrendezett halmazhoz hozzárendelhetünk egy D digráfot a P ponthalmazon, amelyben u-ból akkor vezet él v-be, ha $u \succ v$. A D digráf egyszerű, aciklikus és tranzitív abban az értelemben, hogy ha xy és yz élek, akkor xz is él. Megfordítva, minden ilyen digráf meghatároz egy részbenrendezett halmazt.

A P halmaznak egy páronként összehasonlítható elemekből álló részhalmazát **lánc**nak hívjuk, míg egy páronként összehasonlíthatatlan elemekből álló részhalmaz neve **antilánc**. Egy lánc részhalmaza is lánc, egy antilánc részhalmaza is lánc. Láncnak és antiláncnak nyilván legfeljebb csak egy közös eleme lehet. A lánc a D irányítatlan alapgráfjában egy klikknek felel meg, az antilánc pedig egy stabil halmaznak.

Dilworth tétele szerint a maximális antilánc elemszáma egyenlő a P-t fedő láncok minimális számával. A fedő láncok diszjunktaknak is választhatók.

Mirsky tétele szerint (amit néha poláris Dilworth tételnek is hívnak) a maximális lánc elemszáma egyenlő a P-t fedő antiláncok minimális számával. A fedő antiláncok diszjunktaknak is választhatók.

Feladatok

1.16. *Egy út bizonyos részútjainak adott $\mathcal{F}$ rendszeréből kell kiválasztanunk diszjunkt tagokat úgy, hogy a kiválasztottak összhossza maximális legyen. Fogalmazzuk meg a feladatot aciklikus digráf leghosszabb st-útjának problémájaként.*

1.17. *Dolgozzunk ki eljárást pontsúlyozott részbenrendezett halmaz maximális súlyú láncának megkeresésére.*

1.18. *Adott két betűkből álló sorozat maximális sok tagból álló közös részsorozatát kell kiválasztanunk. Fogalmazzuk meg a feladatot részbenrendezett halmazbeli maximális lánc meghatározásának problémájaként.*

1.19. *Igazoljuk algoritmikusan, hogy egy P részbenrendezett halmazban a leghosszabb lánc elemszáma egyenlő a P-t fedő antiláncok minimális számával! Fogalmazzuk meg és igazoljuk a megfelelő tételt maximális súlyú láncokról, ha P elemei súlyozva vannak.*

1.20. *Egy véges számsorozat legtöbb tagból álló monoton növekvő részsorozatát kell meghatároznunk. Fogalmazzuk meg a feladatot részbenrendezett halmazbeli maximális lánc meghatározásának problémájaként.*

1.21. *Egy véges számsorozat legtöbb tagból álló konvex részsorozatát kell meghatároznunk. Fogalmazzuk meg a feladatot aciklikus digráf leghosszabb st-útjának problémájaként.* (Egy számsorozat konvex, ha az egymást követő tagok különbségei monoton növő.)

1.22. *Igazoljuk, hogy egy $nm+1$ különböző tagokból álló számsorozatnak vagy van $n+1$ tagú monoton növő vagy egy $m+1$ tagú monoton csökkenő részsorozata! Létezhet-e mind a két fajta részsorozat?*

1.3.3. Legolcsóbb utak nem-negatív költségekre: Dijkstra algoritmusa

Tegyük most fel, hogy D tetszőleges, de a c költségfüggvény nem-negatív. Dijkstra algoritmusa két ötleten múlik. Az aciklikus esethez hasonlóan itt is s-ből induló legolcsóbb utaknak egy fenyőjét építjük fel élek egyenkénti hozzávételével. Lényeges különbség azonban, hogy a pontoknak a fenyőbe kerülési sorrendjét nem lehet előre megmondani (mint ahogy az aciklikus esetben a topologikus sorrenddel meg lehetett), mert az csak menetközben derül ki, a következő lemma szerint.

1.3.4. Lemma. *Legyen T egy legolcsóbb utak s-fenyője a $V(T)$ ponthalmazon. Tegyük fel, hogy az*

$$m_T := \min\{\mu_c(u) + c(uv) : uv \text{ kilép } V(T)\text{-ből}\} \tag{1.4}$$

minimum valamely $a = u_a v_a$ élen vétetik fel. Ekkor $T' := T + a$ is legolcsóbb utak s-fenyője a $V(T) + v_a$ ponthalmazon.

Biz. A T-re vonatkozó feltevés miatt csak az s-ből v_a-ba vezető T'-beli P' útról kell belátnunk, hogy D-ben legolcsóbb. A jelölések folytán $\widetilde{c}(P') = m_T$. Legyen P tetszőleges sv_a-út D-ben. Legyen ennek (s felől indulva) az első $V(T)$-ből kilépő éle $e = u_e v_e$, míg az s-től u_e-ig tartó részútja P''. A c nem-negativitása valamint m_T és $\mu_c(u_e)$ jelentése miatt $\widetilde{c}(P') = m_T \leq \mu_c(u_e) + c(e) \leq \widetilde{c}(P'') + c(e) \leq \widetilde{c}(P)$. •

Az 1.3.4 lemma kézenfekvő és hatékony megoldást kínál a legolcsóbb utak fenyőjének kiszámítására. Kiindulva az egyetlen s pontból álló fenyőből, élek egymás utáni hozzávételével, $n-1$ fázisban, felépítjük a V-t feszítő legolcsóbb utak fenyőjét. Ehhez csak az kell, hogy egy közbenső, már kiszámított T fenyőhöz meg tudjuk határozni a hozzáveendő élt. Az 1.3.4 lemma alapján ez az (1.4) minimum kiszámításával megtehető. Kérdés, hogy hány lépésben. A naív megközelítés szerint e minimum meghatározásához számba kell venni az összes $V(T)$-ből kilépő élt. Ezek számára m-nél jobb felső korlátot nem lehet biztosítani, és ezért ez a megközelítés összességében egy $O(mn)$ lépésszámú algoritmust eredményez.

Dijkstra algoritmusának második ötlete az, hogy fenntartunk és menetközben alkalmasan módosítunk bizonyos adatokat, amelyek segítségével az (1.4) minimum $O(m)$ lépés helyett már $O(n)$ lépésben kiszámítható.

Tegyük fel, hogy egy közbenső fázisban a T fenyőn, valamint a T pontjaira már kiszámolt $\mu_c(v)$ értékeken kívül rendelkezésre állnak a következő adatok. Minden $v \in V - V(T)$ fenyőn kívüli pontra a $\mu_T(v)$ címke tartalma legyen az s-ből v-be vezető, csak $V(T)$ pontjait használó sv-utak költségének minimuma, míg az $e_T(v)$ címke tartalma egy ilyen sv-út utolsó éle (amely tehát kilép T-ből és a hegye v). Ezen kívül a v_T címke tartalma az a $z \in V - V(T)$ pont, ahol a $\mu_T(v)$ értékek minimuma ($v \in V - V(T)$) felvétetik, míg m_T a minimum értéke.

Ezen címkék segítségével rögtön meg tudjuk mondani, hogy melyik élt kell T-hez venni. Nevezetesen, tekintjük a v_T-ben lévő z pontot és az $e_T(z)$-ben lévő uz élt és ezt adjuk T-hez. A lemma miatt $\mu_c(z) = \mu_T(z)$.

A keletkező T' fenyőhöz tartozó címkék módosításához figyeljük meg, hogy egy $v \in V - T'$ pontra a legolcsóbb olyan sv-út, amely csak T'-beli pontot használ vagy használja a z pontot vagy nem, attól függően, hogy a $\mu_{T'}(v) := \min\{\mu_T(v),\ \mu_c(z) + c(zv)\}$ minimum a második vagy az első tagon vétetik-e fel. Ennek eldöntése tehát konstans lépésben megtehető. Miután n pont van összesen, megállapíthatjuk, hogy a T fenyő egy újabb éllel való növelésekor minden $\mu_{T'}(v)$ címke $O(n)$ lépésben meghatározható. Hasonlóan egyszerű megfontolás nyomán a $v_{T'}$ és az $m_{T'}$ címkék is $O(n)$ lépésben meghatározhatók. Miután $n-1$ fenyő növelést hajtunk végre, a Dijkstra algoritmus teljes lépésszáma $O(n^2)$.

1.23. Feladat. *Igaz-e, hogy Dijkstra algoritmusa aciklikus digráf esetén tetszőleges költségfüggvényre a legolcsóbb utat szolgáltatja?*

1.3.4. Konzervatív költségfüggvények, megengedett potenciálok, tenziók

Amennyiben a c költségfüggvény negatív is lehet, úgy a legolcsóbb út feladatról már fentebb megjegyeztük, hogy **NP**-teljes. Emiatt a problémát nem vizsgáljuk teljes általánosságban, amikor a digráf és a költségfüggvény is tetszőleges, hanem csupán arra az esetre szorítkozunk, amikor nem létezik negatív összköltségű egyirányú kör (röviden negatív kör). Ilyenkor a költségfüggvényről azt mondjuk, hogy **konzervatív**. Például c bizonyosan konzervatív, ha nem-negatív, vagy akkor is, ha tetszőleges, de D aciklikus. Nem nehéz egyéb konzervatív költségfüggvényeket konstruálni: ilyen például, ha D egy 3 élű egyirányú kör, melyen a költségek rendre $-1, +1, +1$.

Felvetődik a kérdés, hogy miként lehet eldönteni, hogy egy költségfüggvény konzervatív-e vagy sem. Egyáltalán, milyen könnyen ellenőrizhető tanúsítványt tudunk elképzelni arra, hogy c konzervatív? Nevezzünk egy $\pi : V \to \mathbb{R}$ függvényt c-re nézve **megengedett potenciál**nak, vagy c**-megengedett**nek, ha

$$\pi(v) - \pi(u) \leq c(uv) \text{ fennáll minden } uv \in A \text{ élre.} \tag{1.5}$$

Egy uv élt akkor nevezünk **pontos**-nak (a π-re nézve), ha $\pi(v) - \pi(u) = c(uv)$. A π által az élhalmazon indukált $\Delta_\pi : A \to \mathbb{R}$ költségfüggvény definíciója a következő:

$$\Delta_\pi(uv) := \pi(v) - \pi(u) \text{ minden } uv \in A \text{ élre.}$$

Az így előálló költségfüggvényeket **pontindukált**nak nevezzük, vagy másnéven **potenciál-különbség**nek. A π c-megengedettsége tehát azt jelenti, hogy $\Delta_\pi \leq c$ vagy másképp írva, $c_\pi \geq 0$, ahol

$$c_\pi := c - \Delta_\pi,$$

azaz $c_\pi(uv) = c(uv) - \pi(v) + \pi(u)$ minden uv élre. Egy uv él akkor pontos π-re nézve, ha $c_\pi(uv) = 0$.

Egyszerű megfigyelések

Az alábbi megfigyelés azt fejezi ki, hogy a pontindukált költségfüggvények egyfajta értelemben semlegesek.

1.3.5. Lemma. *A Δ_π pontindukált költségfüggvényre nézve* **(a)** *minden st-út költsége ugyanaz az érték, éspedig $\pi(t)-\pi(s)$, továbbá* **(b)** *minden egyirányú kör költsége nulla.*

Biz. Legyen P tetszőleges st-út, melynek pontjai $v_1 = s, v_2, \ldots, v_k = t$. Ennek költsége:

$$\widetilde{\Delta}_\pi(P) = \sum_{i=1}^{k-1} \Delta_\pi(v_i v_{i+1}) = \sum_{i=1}^{k-1} [\pi(v_{i+1}) - \pi(v_i)] = \pi(v_k) - \pi(v_1) = \pi(t) - \pi(s).$$

Legyen K egyirányú kör, melynek csúcsai ciklikus sorrendben $v_{k+1} = v_1, v_2, \ldots, v_k$. Ennek költsége:

$$\widetilde{\Delta}_\pi(K) = \sum_{i=1}^{k} \Delta_\pi(v_i v_{i+1}) = \sum_{i=1}^{k} [\pi(v_{i+1}) - \pi(v_i)] = 0. \bullet$$

A következő észrevétel egyfajta lehetőséget biztosít egy st-út legolcsóbb voltának igazolására megengedett potenciálok segítségével.

1.3.6. Lemma. *Legyen π megengedett potenciál a c költségfüggvényre nézve* (azaz $c \geq \Delta_\pi$ vagy $c_\pi \geq 0$). **(a)** *Minden P st-útra $\widetilde{c}(P) \geq \pi(t) - \pi(s)$. Ha P minden éle pontos, akkor P c-legolcsóbb st-út.* **(b)** *Minden K egyirányú körre $\widetilde{c}(K) \geq 0$. Ha K minden éle pontos, akkor $\widetilde{c}(K) = 0$.*

Biz. (a) Legyenek P pontjai $v_1 = s, v_2, \ldots, v_k = t$. Ekkor

$$\widetilde{c}(P) = \sum_{i=1}^{k-1} c(v_i v_{i+1}) \geq \sum_{i=1}^{k-1} [\pi(v_{i+1}) - \pi(v_i)] = \pi(v_k) - \pi(v_1) = \pi(t) - \pi(s),$$

amiből adódik, hogy $\widetilde{c}(P) \geq \pi(t) - \pi(s)$. Vagyis minden st-út költsége legalább $\pi(t) - \pi(s)$. Mivel egy pontos élekből álló st-út költsége pontosan $\pi(t) - \pi(s)$, az ilyen út szükségképpen legolcsóbb st-út.

(b) Legyenek K pontjai ciklikus sorrendben $v_{k+1} = v_1, v_2, \ldots, v_k$. Ekkor

$$\widetilde{c}(K) = \sum_{i=1}^{k} c(v_i v_{i+1}) \geq \sum_{i=1}^{k} [\pi(v_{i+1}) - \pi(v_i)] = 0.$$

Ebből adódik, hogy minden egyirányú kör c-költsége nemnegatív és ha egy ilyen kör minden éle pontos, akkor a c-költség 0. •

1.3.7. Lemma. *Ha c konzervatív és P legolcsóbb st-út, akkor P minden P' uv-részútja legolcsóbb uv-út.*

Biz.

1.3.1. Állítás. *Ha W egy st-séta, akkor W magában foglal egy olyan Q st-utat, amelyre $\widetilde{c}(Q) \leq \widetilde{c}(W)$.*

Biz. Az s-ből indulva haladjunk a W sétán. Amikor először egy korábbi pontba visszaérünk, akkor egy K egyirányú kör keletkezik, amit a sétából kivágva olyan W' st-sétát kapunk, amelyre a konzervativitás miatt $\widetilde{c}(W') = \widetilde{c}(W) - \widetilde{c}(K) \leq \widetilde{c}(W)$. Ezt a körkivágási eljárást ismételve végül egy Q st-utat kapunk, amelyre $\widetilde{c}(Q) \leq \widetilde{c}(W)$. •

Legyen R tetszőleges uv-út. Jelölje W azt az st-sétát, amelyet P-ből kapunk azáltal, hogy a P' részutat R-rel helyettesítjük. Az 1.3.1 állítás miatt van egy Q st-út, amelyre $\widetilde{c}(Q) \leq \widetilde{c}(W)$. Ekkor a $\widetilde{c}(P) \leq \widetilde{c}(Q) \leq \widetilde{c}(W) = \widetilde{c}(P) - \widetilde{c}(P') + \widetilde{c}(R)$, amiből $\widetilde{c}(P') \leq \widetilde{c}(R)$, vagyis P' valóban legolcsóbb uv-út. • •

A konzervativitás jellemzése: Gallai tétele

Jelölje $\pi_c(v)$ a legolcsóbb v-ben végződő (röviden v-végű) út költségét (bárhol is kezdődjék) vagyis

$$\pi_c(v) := \min\{\widetilde{c}(P) : P \text{ } v\text{-végű út}\}. \tag{1.6}$$

Mivel az egyetlen $\{v\}$ pontból álló séta 0 költségű, a π_c függvény nem-pozitív. (Figyelem: a π_c függvény a csúcshalmazon van értelmezve, szemben c_π-vel, ami az élhalmazon.)

1.3.8. Tétel (Gallai). *A $D = (V, A)$ irányított gráf élein egy c költségfüggvény akkor és csak akkor konzervatív, ha létezik hozzá c-megengedett potenciál. Amennyiben c egészértékű és konzervatív, úgy létezik egészértékű megengedett potenciál is.*

Biz. Amennyiben létezik π megengedett potenciál, úgy bármely K egyirányú körre az 1.3.5 lemma nyomán $\widetilde{c}(K) = \widetilde{c}_\pi(K) \geq 0$.

A fordított irány az alábbi lemmán múlik.

1.3.9. Lemma. *Ha c konzervatív, akkor az* (1.6)*-ben definiált π_c függvény c-megengedett.*

Biz. Legyen uv a digráf egy éle és legyen P_u egy legolcsóbb u-végű út, amelyre tehát $\widetilde{c}(P_u) = \pi_c(u)$. Amennyiben v nincs rajta P-n, úgy $P_v := P + uv$ egy v-végű út és ezért $\widetilde{c}(P_v) \geq \pi_c(v)$. Ebből $\pi_c(v) \leq \widetilde{c}(P_v) = \widetilde{c}(P_u) + c(uv) = \pi_c(u) + c(uv)$, vagyis ilyenkor valóban $\pi_c(v) - \pi_c(u) \leq c(uv)$.

Tegyük most fel, hogy v rajta van a P_u-úton. A P_u út v-ig tartó kezdő szakaszát jelölje P_1, míg a v-től u-ig vezető részútját P_2. Ekkor $\widetilde{c}(P_u) = \widetilde{c}(P_1) + \widetilde{c}(P_2)$. Tekintsük a $K := P_2 + uv$ egyirányú kört. Mivel c konzervatív, $\widetilde{c}(K) \geq 0$, vagyis $c(uv) \geq -\widetilde{c}(P_2)$. Miután P_1 egy v-végű út, ezért $\pi_c(v) \leq \widetilde{c}(P_1)$. Mindezeket összevetve azt kapjuk, hogy $\pi_c(v) - \pi_c(u) \leq \widetilde{c}(P_1) - \widetilde{c}(P_u) = -c(P_2) \leq c(uv)$, vagyis π_c tényleg megengedett potenciál. •

A tétel második része következik a lemmából, hiszen ha c egészértékű, úgy π_c is az. • •

A következő megfigyelés azt mutatja, hogy a π_c függvény egyfajta értelemben kanonikus megengedett potenciál.

1.3.10. Tétel. *Legyen c konzervatív költségfüggvény. A π_c függvény az egyértelmű legnagyobb nem-pozitív megengedett potenciál* (azaz $\pi_c \geq \pi$ minden nem-pozitív megengedett π potenciálra).

Biz. Legyen π nem-pozitív megengedett potenciál. Legyen P_t egy legolcsóbb t-végű út, amelyre tehát $\widetilde{c}(P_t) = \pi_c(t)$. Jelölje P_t kezdőpontját s. Az 1.3.6 lemma és $\pi(s) \geq 0$ folytán $\pi_c(t) = \widetilde{c}(P_t) = \pi(t) - \pi(s) \geq \pi(t)$. •

Az 1.3.8 tétel alábbi kiterjeszésében a potenciál-különbségre nem csupán felső, hanem alsó korlátot is előírunk.

1.3.11. Tétel. *A $D = (V, A)$ digráf élhalmazán adott két korlátozó függvény: $c_{al} \leq c_{fel}$. Akkor és csak akkor létezik olyan $\pi : V \to \mathbb{R}$ vektor* (amely ráadásul egészértékű, ha c_{al} és c_{fel} is az), *amelyre $c_{al}(uv) \leq \pi(v) - \pi(u) \leq c_{fel}(uv)$ minden $e = uv$ élre, ha a c'-vel élsúlyozott $D' = (V, A')$ segédgráfban nincsen negatív kör, ahol uv akkor eleme A'-nek, ha vagy $uv \in A$ és ekkor $c'(uv) := c_{fel}(uv)$, vagy $vu \in A$ és ekkor $c'(uv) := -c_{al}(vu)$.* •

1.24. Feladat. *Igazoljuk az* 1.3.11 *tételt.*

Tenziók

Miként lehet egy $x : A \to \mathbb{R}$ függvényről felismerni, hogy pontindukált-e? Ha például a digráf az u és v pontokból valamint az u-ból v-be vezető e és f párhuzamos élekből áll, akkor $x(e) \neq x(f)$ esetén x nyilván nem lehet potenciál-különbség.

Egy $x : A \to \mathbb{R}$ függvényt **tenzió**nak nevezünk, ha a digráf minden C köre **semleges** abban az értelemben, hogy a C előremenő élein a $\varphi_x(C)$-vel jelölt x-összeg egyenlő a hátramenő éleken vett $\beta_x(C)$ x-összeggel. (A φ betű az angol *forward* szóra utal, míg a β a *backward*-ból jön.) Magyarul C akkor semleges, ha $\varphi_x(C) - \beta_x(C) = 0$. Speciálisan ez azt jelenti, hogy minden egyirányú K körre $\widetilde{x}(K) = 0$.

Legyen $e = uv$ egy él, míg T egy e-t nem tartalmazó feszítő fa. Az e él az u-t és v-t a T-ben összekötő egyértelmű P úttal egy kört alkot, amit az e él **T-hez tartozó alapkör**ének nevezünk.

1.3.12. Tétel. *Egy $c : A \to \mathbb{R}$ függvény akkor és csak akkor tenzió, ha potenciál-különbség. Ha c egészértékű tenzió, akkor létezik olyan egészértékű π potenciál, amelyre $c = \Delta_\pi$.*

Biz. Egy egyszerű, de hasznos megfigyeléssel kezdjük. Legyen uv és vu két szembe irányított párhuzamos él, melyekre $c(vu) = -c(uv)$. Bármely $\pi : V \to \mathbb{R}$ függvényre $c(uv) = \pi(v) - \pi(u)$ pontosan akkor teljesül, ha $c(vu) = \pi(u) - \pi(v)$. Ebből adódik, hogy $(*)$ egy költségfüggvény akkor és csak akkor potenciál-különbség, ha néhány élt megfordítva és ezeknek a költségét negálva potenciál-különbséget kapunk.

Tegyük fel először, hogy c potenciál-különbség, azaz létezik egy $\pi : V \to \mathbb{R}$ függvény, amelyre $c = \Delta_\pi$. Legyen C a digráf egy köre. Ha C egyirányú, akkor az 1.3.5 lemma nyomán tudjuk, hogy $\widetilde{c}(C) = 0$. Ha a körön vannak előremenő és hátramenő élek, akkor az utóbbiakat fordítsuk meg és költségüket negáljuk. A keletkező K egyirányú körre $0 = \widetilde{c}(K) = \varphi_c(C) - \beta_c(C)$.

A fordított irány igazolásához legyen most c tenzió, amelyről ki akarjuk mutatni, hogy potenciál-különbség. Feltehetjük, hogy D irányítatlan értelemben összefüggő. A $(*)$ tulajdonság nyomán ekvivalens feladathoz jutunk, ha néhány élt megfordítunk és költségeiket negáljuk. Ezért feltehetjük, hogy a digráfnak van egy F feszítő fenyője, melynek gyökerét jelölje r. Legyen $\pi(v)$ a fenyőben az r-ből v-be vezető egyértelmű egyirányú út költsége. Figyeljük meg, hogy ha c egészértékű, akkor π is az.

1.3.2. Állítás. $c = \Delta_\pi$.

Biz. A definícióból adódóan az F fenyő minden uv élére $c(uv) = \pi(v) - \pi(u)$. Tegyük most fel, hogy az $e = uv$ él nincs a fenyőben. Tekintsük az e él C alapkörét. Ennek létezik egy r-hez legközelebbi s pontja, és ekkor $C = uv + P_u + P_v$, ahol P_u illetve P_v az s-ből a fenyőben az u-ba illetve a v-be vezető egyirányú utakat jelöli. A π definíciójából adódik, hogy $\pi(v) - \pi(u) = \widetilde{c}(P_v) - \widetilde{c}(P_u)$. Feltehetjük, hogy uv a kör előremenő éle. Ekkor $\varphi_c(C) = c(uv) + \widetilde{c}(P_u)$ és $\beta_c(C) = \widetilde{c}(P_v)$. Emiatt $0 = \varphi_c(C) - \beta_c(C) = c(uv) + \widetilde{c}(P_u) - \widetilde{c}(P_v) = c(uv) + \pi(u) - \pi(v)$, azaz $c(uv) = \pi(v) - \pi(u)$. • •

Feladatok

1.25. *Igazoljuk, hogy a pontindukált költségek alteret alkotnak $\mathbb{R}^A$-ban. Határozzuk meg ennek ortogonális kiegészítő alterét, más szóval, jellemezzük azon $x \in \mathbb{R}^A$ vektorokat, melyeknek minden pontindukált költségfüggvénnyel vett skalárszorzata nulla.*

1.26. *Igazoljuk, hogy ha c egészértékű és pontindukált, akkor létezik olyan π egészértékű potenciál, amelyre $c = \Delta_\pi$.*

1.27. *Legyen adott egy c konzervatív költségfüggvény a D irányított gráf élhalmazán. Igazoljuk, hogy ha a K egyirányú kör minden éle benne van 0-költségű egyirányú körben, akkor $\widetilde{c}(K) = 0$.*

1.28. *Melyek azok a digráfok, amelyek élhalmazán létezik $\{+1, -1\}$-értékű potenciál-különbség?*

1.29. *Adjunk az* 1.3.8 *tételre alternatív bizonyítást, amely pontszám szerinti indukciót használ az alábbi vázlat alapján. Válasszunk ki egy tetszőleges z pontot, minden uz és zv élpárra vegyünk egy új élt u-ból z-be, amelynek költsége legyen $c(uz) + c(zu)$, majd töröljük a z pontot. A keletkező kisebb pontszámú gráfra alkalmazzunk indukciót.*

1.30. *Erősen összefüggő digráfban egy $x : A \to \mathbb{R}$ függvény akkor és csak akkor tenzió, ha $\widetilde{x}(K) = 0$ minden K egyirányú körre.*

1.31. *Egy $x : A \to \mathbb{R}$ függvény akkor és csak tenzió, ha egy adott T feszítő fához tartozó valamennyi* (tehát $m - n + 1$) *alapkör semleges.*

1.32. *Igazoljuk, hogy ha π_1 és π_2 c-megengedett potenciálok, akkor a π-vel jelölt maximumuk is az, ahol $\pi(v) := \max\{\pi_1(v), \pi_2(v)\}$ $(v \in V)$. Erre támaszkodva adjunk altenatív bizonyítást az 1.3.10 tételre.*

1.33. *Legyen π_1 és π_2 egészértékű c-megengedett potenciál. Igazoljuk, hogy $\pi_1 \sqcap \pi_2$ is c-megengedett potenciál, ahol $\pi_1 \sqcap \pi_2$ értéke a $v \in V$ csúcsban $\lfloor(\pi_1(v) + \pi_2(v)/2\rfloor$.*

1.3.5. Legolcsóbb utak: min-max tétel és optimalitási feltétel

Az alábbiakban végig feltesszük, hogy a D digráfban minden pont elérhető s-ből.

1.3.13. Tétel. *Tegyük fel, hogy a $D = (V, A)$ digráf minden csúcsa elérhető egy kijelölt s pontból. Legyen c konzervatív költségfüggvény a D élhalmazán. Ekkor μ_c megengedett potenciál. Ráadásul μ_c az egyértelmű legnagyobb olyan megengedett potenciál, amelyre $\pi(s) = 0$* (vagyis $\mu_c \geq \pi$ minden olyan megengedett π potenciálra, amelyre $\pi(s) = 0$).

Biz. Legyen uv a digráf egy éle és tekintsünk egy P legolcsóbb su-utat, amelyre tehát $\widetilde{c}(P) = \mu_c(u)$. Amennyiben v nincs rajta P-n, úgy $P + uv$ egy sv-út, és ezért a költsége legfeljebb $\mu_c(v)$. Tehát $\mu_c(v) \geq \widetilde{c}(P) + c(uv) = \mu_c(u) + c(uv)$, vagyis ilyenkor valóban $\mu_c(v) - \mu_c(u) \leq c(uv)$.

Tegyük most fel, hogy v rajta van a P-úton. A P út v-ig tartó kezdő szakaszát jelölje P_1, míg a v-től u-ig vezető részútját P_2. Ekkor tehát $\widetilde{c}(P) = \widetilde{c}(P_1) + \widetilde{c}(P_2)$. Tekintsük a $K := P_2 + uv$ egyirányú kört. Mivel c konzervatív, $\widetilde{c}(K) \geq 0$, vagyis $c(uv) \geq -\widetilde{c}(P_2)$. Miután P_1 egy sv-út, $\mu_c(v) \leq \widetilde{c}(P_1)$. Mindezeket összetéve kapjuk, hogy $\mu_c(v) - \mu_c(u) \leq \widetilde{c}(P_1) - \widetilde{c}(P) = -c(P_2) \leq c(uv)$.

A második rész igazolásához legyen π olyan megengedett potenciál, amelyre $\pi(s) = 0$ és legyen P egy legolcsóbb st-út, amelyre tehát $\widetilde{c}(P) = \mu_c(t)$. Az 1.3.6 lemma alapján $\mu_c(t) = \widetilde{c}(P) \geq \pi(t) - \pi(s) = \pi(t)$. •

Alternatív bizonyítás Megmutatjuk, hogy μ_c megengedettsége rögtön következik az 1.3.9 lemmából is. Valóban, vegyünk a digráfhoz egy új s' pontot és egy új $s's$ élt, melynek c-költsége legyen $-M$, ahol M egy nagy szám. A keletkező D' digráfban az M nagy választása miatt a legolcsóbb v-végű utak mind s'-ben kezdődnek és emiatt $\pi'_c(v) = \mu_c(v) - M$, ahol $\pi'_c(v)$ a a D' legolcsóbb v-végű útjának a költségét jelöli. Ezért $\pi'_c(v)$ megengedettségéből következik μ_c megengedettsége. •

1.3.14. Tétel (Legolcsóbb utak részgráfjának tétele)**.** *Tegyük fel, hogy a $D = (V, A)$ digráf minden csúcsa elérhető egy kijelölt s pontból. Legyen c konzervatív költségfüggvény a D élhalmazán. Jelölje $D_0 = (V, A_0)$ a μ_c-re nézve pontos élek részgráfját. Minden t csúcsra egy P st-út akkor és csak akkor legolcsóbb st-út D-ben* (azaz $\mu_c(t)$ költségű)*, ha minden éle D_0-ban van. Speciálisan, D_0-ban bármely s-fenyő a D-nek egy legolcsóbb utak fenyőjét alkotja.*

Biz. Az 1.3.13 tétel szerint μ_c megengedett potenciál. Az 1.3.6 lemma első része alapján minden μ_c-re nézve pontos élekből álló st-út legolcsóbb st-út, vagyis minden D_0-beli st-út legolcsóbb st-út D-ben.

Megfordítva, tegyük fel, hogy P legolcsóbb st-út D-ben, melynek pontjai legyenek $s = v_0, v_1, \cdots, v_k = t$. Ismét használva, hogy μ_c megengedett potenciál, kapjuk, hogy

$$\begin{aligned}\mu_c(t) = \widetilde{c}(P) = \sum[c(v_{i-1}v_i) : i = 1, \ldots, k]\\ \geq \sum[\mu_c(v_i) - \mu_c(v_{i-1}) : i = 1, \ldots, k] = \mu_c(t) - \mu_c(s) = \mu_c(t),\end{aligned}$$

amiből $c(v_{i-1}v_i) = \mu_c(v_i) - \mu_c(v_{i-1})$ minden $i = 1, \ldots, k$-re, mutatva, hogy P valamennyi éle D_0-ban van. •

A legolcsóbb utakra vonatkozó min-max tételt aciklikus digráfokra már bebizonyítottuk (1.3.2 tétel). A tétel általánosabban is érvényes.

1.3.15. Tétel (Duffin)**.** *Konzervatív c költségfüggvény esetén az s-ből t-be vezető utak költségének $\mu_c(t)$ minimuma egyenlő a $\pi(t) - \pi(s)$ érték maximumával, ahol a maximum az összes megengedett π potenciálon veendő. Amennyiben c egészértékű, úgy az optimális π is választható annak.*

Biz. Legyen π megengedett potenciál, P pedig egy tetszőleges út s-ből t-be. Az 1.3.6 lemma folytán $\widetilde{c}(P) \geq \pi(t) - \pi(s)$, vagyis a tételben a min $\geq$ max irány következik.

A fordított irányú egyenlőtlenség igazolásához kell találnunk egy P st-utat és egy π megengedett potenciált, amelyekre $\widetilde{c}(P) = \pi(t) - \pi(s)$. E célra viszont egy tetszőleges P legolcsóbb st-út és a μ_c megengedett potenciál az 1.3.14 tétel miatt megfelel. •

1.3.6. Algoritmusok

A fentiekben megteremtettük az elvi hátteret olyan algoritmusok készítéséhez, amelyek segítségével hatékonyan kereshetünk negatív kört vagy megengedett potenciált. Erre két eltérő jellegű algoritmust is leírunk.

Javító utas algoritmus

Célunk algoritmikusan újra bizonyítani Gallai tételének nem-triviális irányát (: *ha egy digráfban nincs negatív kör, akkor van megengedett potenciál.*) Ennek érdekében induljunk ki egy tetszőleges π potenciálból, amely egészértékű, ha c az. Ha $c_\pi := c - \Delta_\pi$ nem-negatív, akkor π definíció szerint megengedett és ekkor az eljárás már véget is ér. Tegyük fel tehát, hogy vannak **hibás** élek, azaz olyanok, amelyek c_π-értéke negatív.

Az alábbi eljárás egyenként megszünteti a hibás éleket anélkül, hogy újabb hibásakat kreálna, illetve ha egy él hibásságát nem sikerül megszüntetnie, akkor talál egy negatív kört. E célból tekintsünk egy hibás $st \in A$ élt, amelyre tehát $c_\pi(st) < 0$. Legyen $A_\pi := \{e \in A : c_\pi(e) \leq 0\}$ és tekintsük a D digráf $D_\pi = (V, A_\pi)$ részgráfjában a t-ből elérhető pontok Z halmazát. Amennyiben s benne van Z-ben, azaz D_π-ben létezik egy P egyirányú ts-út, úgy a $K := P + st$ egyirányú körre $\widetilde{c}(K) = \widetilde{c}_\pi(K) = \widetilde{c}_\pi(P) + c_\pi(st) < 0$, vagyis K negatív kör.

Ha s nincs Z-ben, akkor módosítsuk π-t a következőképpen.

$$\pi(v) := \begin{cases} \pi(v) - \varepsilon, & \text{ha } v \in Z \\ \pi(v), & \text{ha } v \in V - Z \end{cases} \tag{1.7}$$

ahol $\varepsilon := j \min\{|c_\pi(st)|,\ \varepsilon_1\}$ és

$$\varepsilon_1 := \min\{c_\pi(e) : e \in A, e \text{ kilép } Z\text{-ből}\}. \tag{1.8}$$

Az ε_1-t itt ∞-nek értelmezzük, ha D-nek semelyik éle sem lép ki Z-ből. Megállapíthatjuk, hogy ε a definíciója folytán pozitív és a π fenti módosításával nem keletkezik új hibás él. Ha $\varepsilon = |c_\pi(st)|$, akkor π'-re nézve st már nem hibás, vagyis a hibás élek halmaza a célnak megfelelően kisebb lett. Amennyiben $\varepsilon = \varepsilon_1$, akkor ismételjük az eljárást a változatlan st élre és a módosított π' potenciálra nézve.

Figyeljük meg egyrészt, hogy a Z által feszített A_π-élek és $A_{\pi'}$-élek ugyanazok, másrészt a Z-ből kilépő $e \in A$ él, amelyen az (1.8)-beli minimum felvétetik bekerül $A_{\pi'}$-be, hiszen $\varepsilon = \varepsilon_1 = c_\pi(e)$ miatt $c_{\pi'}(e) = 0$. Emiatt a $D_{\pi'}$-ben a t-ből elérhető pontok halmaza szigorúan bővebb, mint Z és így ilyen ismétlésre legfeljebb $n-1$-szer kerülhet sor.

Végül vegyük észre, hogy egészértékű c esetén az eljárás végig fenntartja π egészértékűségét. •

A bizonyításból adódóan egyetlen hibás él megjavítása $O(m)$ lépésben történhet, így az algoritmus teljes lépésszáma $O(m^2)$.

Legolcsóbb v-végű séták és utak: a Bellman–Ford algoritmus

Konzervatív c esetén a fenti $O(m^2)$-es javító utas algoritmus megtalál *egy* megengedett potenciált. A most következő Bellman–Ford féle eljárás kiszámítja az 1.3.9 lemmában szereplő kanonikus π_c megengedett potenciált vagy pedig talál egy negatív kört. Az algoritmus lépésszáma $O(mn)$ lesz.

Az alapötlet az, hogy az $i = 0, 1, \ldots, n$ értékek mindegyikére egymás után egy egyszerű rekurzió segítségével kiszámítjuk a v-végű ($v \in V$) legolcsóbb legfeljebb i élű sétát. Jelölje ennek költségét $\pi_c^{(i)}(v)$. Ehhez nem is kell c-ről feltenni, hogy konzervatív.

Ha viszont az, akkor az 1.3.1 állításból tudjuk, hogy a minimalizáló séta választható útnak is, és emiatt $\pi_c(v) = \pi_c^{(n)}(v)$.

Mivel az egyetlen v pontból álló él nélküli séta költsége 0, így $\pi_c^{(0)}(v) = 0$ minden v-re. Figyeljük meg, hogy egy v-ben végződő legfeljebb $i+1$ élű séta vagy pontosan $i+1$ élből áll vagy legfeljebb i élből. Ebből adódóan ha a $\pi_c^{(i)}(v)$ értékek már minden v csúcsra rendelkezésre állnak, úgy legyen

$$\pi_c^{(i+1)}(v) = \min\{\pi_c^{(i)}(v), \min\{\pi_c^{(i)}(u) + c(uv) : uv \in A\}\}. \tag{1.9}$$

Természetesen ugyanez a rekurzió használható maguknak a legolcsóbb legfeljebb i élű v-végű $W_c^{(i)}(v)$ sétáknak a megkonstruálására is. Valóban, $i = 0$-ra $W_c^i(v)$ legyen az egyetlen v pontból álló 0 élű séta. Ha pedig valamilyen $i \geq 0$-ra már minden v-re kiszámítottuk a $W_c^{(i)}(v)$ sétákat, akkor legyen:

$$W_c^{(i+1)}(v) := \begin{cases} W_c^{(i)}(v), & \text{ha } \pi_c^{(i+1)}(v) = \pi_c^{(i)}(v) \\ W_c^{(i)}(u) + uv, & \text{ha } \pi_c^{(i)}(v) > \pi_c^{(i+1)}(v) = \pi_c^{(i)}(u) + c(uv) \text{ egy } uv \in A\text{-ra.} \end{cases} \tag{1.10}$$

A szétválasztás annak megfelelően történik, hogy (1.9) jobb oldalán a külső minimum az első vagy a második tagon vétetik fel, és ha nem az elsőn, akkor a második tagon belül melyik uv élen. A definícióból adódóan $W_c^{(i+1)}(v)$ egy legolcsóbb legfeljebb $i+1$ élű v-ben végződő séta.

Az algoritmus futása kétféleképpen érhet véget.

I. eset A futás végén kapott valamennyi $W_c^{(n)}(v)$ séta út. Ekkor $W_c^{(n)}(v)$ egy legolcsóbb v-végű út és ezért $\pi_c^{(n)}(v) = \pi_c(v)$. Az 1.3.9 lemma miatt az így kiszámított π_c megengedett potenciál.

II. eset A végül kapott séták valamelyike nem út. Ekkor van olyan i index, hogy a $W_c^{(i)}(v)$ séták mindegyik $v \in V$ pontra utat alkotnak, de van olyan v csúcs, amelyre $W_c^{(i+1)}(v)$ nem út. Ekkor az (1.9) rekurzióban a minimum nem az első tagon vétetik fel és emiatt bizonyosan

$$\pi_c^{(i+1)}(v) < \pi_c^{(i)}(v).$$

Az (1.9) második tagjában a minimum egy olyan uv élen éretik el, amelyre v rajta van $P_u := W_c^{(i)}(u)$ úton. (Ha ugyanis nem volna rajta, akkor $W_c^{(i+1)}(v) = W_c^{(i)}(u) + uv$ út volna.) Jelölje P_1 a P_u út v-ig terjedő kezdő szakaszát, míg P_2 a P_v vu-részútját. Ekkor $K := P_2 + uv$ egyirányú kör.

1.3.3. Állítás. $\widetilde{c}(K) < 0$.

Biz. Miután P_1 v-ben végződik és $|P_1| < |P_u| \leq i$, ezért

$$\widetilde{c}(P_1) \geq \pi_c^{(i)}(v) > \pi_c^{(i+1)}(v) = \widetilde{c}(P_1) + \widetilde{c}(K),$$

amiből $\widetilde{c}(K) < 0$ adódik. •

Az algoritmus futása ebben az esetben a negatív K kör kiadásával végződik.

Az algoritmus minden i-re a minimumok számolásánál minden egyes élt egyszer tekint, így a teljes lépésszám $O(nm)$. Ez jobb, mint a javító utakat használó $O(m^2)$-es algoritmus. A fenti megfontolásokból adódik a Gallai tétel egyfajta finomítása.

1.3.16. Tétel. *Adott D digráfra és c költségfüggvényre a következők ekvivalensek.*
(P1) *D-ben nincs negatív egyirányú kör* (azaz c konzervatív).
(P2) *A Bellman–Ford algoritmus nem talál negatív kört.*

(D1) *Létezik megengedett potenciál.*
(D2) *A Bellman–Ford algoritmus által szolgáltatott $\pi_c^{(n)}(v)$ $(v \in V)$ potenciál megengedett és $\pi_c^{(n)} = \pi_c$.* •

1.34. Gyakorlat. *Dolgozzunk ki eljárást a legolcsóbb v-ben végződő pontosan i élű séta meghatározására.*

Legolcsóbb s-ből induló séták és utak: a Bellman–Ford algoritmus

Tegyük fel most, hogy D minden pontja elérhető s-ből és c konzervatív. Az 1.3.13 tételben láttuk, hogy μ_c megengedett potenciál (ahol $\mu_c(v)$ jelölte a legolcsóbb sv-út költségét). Fentebb leírtuk, hogy miként lehet algoritmikusan kiszámítani a π_c potenciált és az 1.3.13 tétel alternatív bizonyításában jeleztük, hogy ebből miként adható meg közvetlenül μ_c. Ennek ismeretében pedig rendelkezésre áll a μ_c-re nézve pontos élek D_0 részgráfja, amiről az 1.3.14 tételben láttuk, hogy megadja az s-ből induló összes legolcsóbb utat. Érdekes megfigyelni, hogy ebben a felépítésben nincs is szükség a legolcsóbb, legfeljebb i élű sétáknak a fenti rekurzióval történő meghatározására, mert a μ_c önmagában definiálja D_0-t.

Ez a megközelítés különösen előnyös, ha nem csupán egy rögzített s-ből induló legolcsóbb utakra vagyunk kíváncsiak, hanem valamennyi $s \in V$ pontra meg kell határoznunk a legolcsóbb utak s-gyökerű fenyőjét. Ennél mi sem egyszerűbb: n-szer kell alkalmaznunk a Bellman–Ford algoritmust. Eszerint az eljárás lépésszáma $O(n^2m)$. De ennél gazdaságosabban is eljárhatunk. A π_c függvényről megmutattuk, hogy $O(mn)$ lépésben kiszámítható. Mivel π_c c-megengedett potenciál, azaz $c' := c - \Delta_{\pi_c} \geq 0$, egy c'-re nézve legolcsóbb s-gyökerű fenyő a Dijkstra algoritmus segítségével $O(n^2)$ lépésben számítható. Mivel egy ilyen fenyő c-re nézve is legolcsóbb utak fenyője, az n darab fenyő kiszámításának lépésszáma $O(mn) + O(n^3)$, ami $m \leq n^2$ miatt $O(n^3)$ és ez jobb, mint az n darab Bellman–Fordból előbb kapott $O(n^2m)$.

Visszatérve a rögzített s pontból induló legolcsóbb utak kiszámítására egy kicsit direktebb változatot is megadunk, amikor az algoritmust közvetlenül a legolcsóbb sv-utak kiszámítására fogalmazzuk meg. Ne tegyük fel apriori, hogy c konzervatív: ha kiderül, hogy nem az, az algoritmus majd kiad egy negatív kört. Erre a változatra is Bellman–Ford algoritmusként hivatkozunk és valójában Bellman és Ford eredeti algoritmusa erre vontakozik. Ilyenkor a kijelölt s pontból minden $v \in V - s$ csúcsra a legolcsóbb legfeljebb i élű sv-sétát akarjuk meghatározni $i = 1, 2, \ldots$ esetén, amely konyervatív c esetén mind út lesz. Ennek költségét jelölje $\mu_c^{(i)}(v)$.

Feltesszük, hogy s-be nem lép él, hogy a digráfban nincs hurokél és nincs (egyirányú) párhuzamos él sem. Ennek folytán $\mu_c^{(i)}(s) = 0$ minden i-re. Azt is feltehetjük, hogy minden sv él a digráfhoz tartozik, mert különben nagy M költséggel a digráfhoz véve a legolcsóbb st-út költsége nem változik és negatív kör se keletkezik. (Választhatjuk M-et például a $\max\{c(e) : e \in A\}$ értéknek.) Ennek megfelelelően $\mu_c^{(1)}(v) = c(sv)$ $(v \in V - s)$. A $\mu_c^{(i)}(v)$ költségeket könnyű kiszámítani egymás után az $i = 2, 3, \ldots$ értékekre, hiszen egy legfeljebb $i + 1$ élű sv-séta vagy pontosan $i + 1$ élből áll vagy legfeljebb i élből. Ebből adódik az alábbi rekurzió. Amennyiben a $\mu_c^{(i)}(v)$ értékek már minden v csúcsra rendelkezésre állnak, úgy legyen

$$\mu_c^{(i+1)}(v) = \min\{\mu_c^{(i)}(v), \min\{\mu_c^{(i)}(u) + c(uv) : uv \in A\}\}. \tag{1.11}$$

Ugyanez a rekurzió használható maguknak a legolcsóbb legfeljebb i élű $S_c^{(i)}(v)$ sv-sétáknak a megkonstruálására is. A kiindulási $i = 1$ értékre álljon $S_c^i(v)$ séta az egyetlen sv élből ($v \in V - s$). Ha valamilyen $i \geq 1$-re már minden $v \in V - s$-re kiszámítottuk az $S_c^{(i)}(v)$ sétákat, akkor legyen:

$$S_c^{(i+1)}(v) := \begin{cases} S_c^{(i)}(v), & \text{ha } \mu_c^{(i+1)}(v) = \mu_c^{(i)}(v) \\ S_c^{(i)}(u) + uv, & \text{ha } \mu_c^{(i)}(v) > \mu_c^{(i+1)}(v) = \mu_c^{(i)}(u) + c(uv) \text{ egy } uv \in A\text{-ra.} \end{cases} \tag{1.12}$$

Magyarán, a szétválasztás annak megfelelően történik, hogy (1.11) jobb oldalán a külső minimum az első vagy a második tagon vétetik fel, és ha nem az elsőn, akkor a második minimumon belül melyik uv élen. A definícióból adódóan $S_c^{(i+1)}(v)$ egy legolcsóbb legfeljebb $i + 1$ élű sv-séta.

Mármost, ha egy $S_c^{(n)}(v)$ sv-séta út, akkor ez legolcsóbb sv-út. Ha pedig $S_c^{(n)}(v)$ valamelyik v-re nem út, akkor c-nem konzervatív és a v-végű sétáknál leírtakhoz hasonlóan lehet egy negatív kört megtalálni.

Az algoritmus minden i esetén a minimumok számolásánál minden egyes élt egyszer tekint, így adott n-re a teljes lépésszám $O(nm)$.

Megjegyzés Valójában az algoritmus során nincsen szükség az összes $S_c^{(i)}(v)$ séta tárolására. Elég ha csak azt az uv élt tároljuk, amelyen az (1.11) rekurzió második tagjában a minimum felvétetik, amikor a minimum nem az első tagon vétetik fel. Ha c konzervatív, akkor bármely v-pontba úgy kaphatunk meg egy legolcsóbb utat s-ből, hogy v-ből visszafelé haladva a tárolt éleken visszaérünk s-be. Ha c nem konzervatív, akkor ez a visszafelé lépkedő eljárás egy olyan v-ből indulva, amelyre $\mu_c^{(n)}(v) < \mu_c^{(n-1)}(v)$ nem s-be ér vissza, hanem egy negatív kört alkot. Ha visszont $\mu_c^{(n)}(v) = \mu_c^{(n-1)}(v)$ minden v-re, akkor a legkényelmesebb az a fentebb már említett megközelítés, amely csak a $\mu_c = \mu_c^{(n)}$-t számolja ki, mert ekkor a μ_c-pontos élek D_0 digráfja közvetlenül megadja az összes s-ből induló legolcsóbb utat.

A két feladat ekvivalenciája

A fenti két feladatról érződik, hogy nagyon közel állnak egymáshoz, ekvivalenciájukat fejezi ki az alábbi kis lemma.

1.3.17. Lemma. (a) *A $\pi_c^{(i)}$ függvény előáll, mint egy egy ponttal nagyobb digráfra vonatkozó $\mu_c^{(i+1)}$ függvény V-re való megszorítása.* (b) *$\mu_c^{(i)}$ előáll, mint egy 1 ponttal nagyobb digráfra vonatkozó $\pi_c^{(i+1)}$ függvény konstanssal való eltoltjának V-re való megszorítása.*

Biz. (a) Vegyünk a digráfhoz egy új s csúcsot és minden v csúcsra egy 0 költségű sv élt. Rögtön láthatóan

$$\pi_c^{(i)}(v) = \mu_c^{(i+1)}(v) \text{ minden } v \in V\text{-re,} \tag{1.13}$$

ahol $\mu_c^{(i+1)}$ a kibővített digráfra vonatkozik.

(b) Adjunk a digráfhoz egy új s' pontot és egy új $s's$ élt, melynek költsége legyen egy alkalmasan nagy M szám negatívja. Ekkor a kibővített digráfban minden v-re a v-ben végződő legolcsóbb legfeljebb $i+1$ élű séta s'-ben fog kezdődni, és ezért

$$\mu_c^{(i)}(v) = \pi_c^{(i+1)}(v) - M \text{ minden } v \in S\text{-re}, \tag{1.14}$$

ahol $\pi_c^{(i+1)}$ a kibővített digráfra vonatkozik. •

Feladatok

1.35. *Fogalmazzuk meg az* 1.3.11 *tétel általánosítását, ha minden v pontban a $\pi(v)$-re alsó és felső korlát is ki van tűzve.*

1.36. *Dolgozzunk ki eljárást annak eldöntésére, hogy egy digráf konzervatív súlyozására nézve létezik-e nulla súlyú kör.*

1.37. *Tegyük fel, hogy c nem konzervatív. Nevezzünk egy v csúcsot hibásnak, ha $\pi_c^{(n)}(v) < \pi_c^{(n-1)}(v)$.* (a) *Mutassuk meg, hogy egy v-ben végződő legolcsóbb legfeljebb n élű séta indukál egy K negatív kört. Igazoljuk, hogy ha c-t minden élen egységesen a $\widetilde{c}(K)/|K|$ értékkel megemeljük, akkor a keletkező c^+ költségfüggvényre nézve v már nem hibás, továbbá minden c-re nézve hibátlan pont c^+-ra nézve is hibátlan.*

1.38. *Legyen adott egy nemnegatív költségfüggvény egy irányított gráf élhalmazán. Igazoljuk, hogy ha egy s-ből t-be vezető P egyirányú út minden éle benne van egy legolcsóbb egyirányú útban, akkor P maga is legolcsóbb egyirányú út. Mutassuk meg, hogy a megfelelő állítás irányítatlan gráfra nem igaz.*

1.39. *Adott az éleken egy tetszőleges súly-függvény. Határozzunk meg egy olyan s-ből t-be vezető utat, amelyen a legnagyobb súly a lehető legkisebb.*

1.40. *Tegyük fel, hogy az éleken két konzervatív költségfüggvényünk adott: c_1 és c_2. Készítsünk algoritmust olyan st-út megkeresésére, amely a c_1-re nézve minimális költségű és ezen belül c_2-re nézve minimális költségű.*

1.41. *Egy digráfban minden v pontra meg van adva egy sv-út. Amennyiben ezek $\pi(v)$ költsége megengedett potenciál, úgy mindegyik út legolcsóbb út.*

1.4. Páros gráfok optimális párosításai

Egy osztályba 25 gyerek jár. Egy kiránduláson készült fényképek közül 25-öt hívtak elő, melyek mindegyikén a gyerekek egy csoportja látható. A képeket szeretnénk kiosztani a gyerekek között, természetesen úgy, hogy minden gyerek rajta legyen a neki juttatott képen. Mikor lehetséges ez, és hogyan tudunk hatékonyan megkeresni egy ilyen hozzárendelést? Általánosabb feladathoz jutunk, ha minden gyerek, mondjuk 0-tól 10-ig terjedő pontozással megmondja, hogy az egyes fényképek számára mennyit érnek. Ekkor a gyerekeknek és a fényképeknek egy olyan egymáshoz rendelését kell megkeresnünk, amelynek az összpontszáma a lehető legnagyobb.

Az ilyen jellegű problémák körét nevezik **hozzárendelési feladat**nak (assignment problem). Íme egy másik példa: az úszószövetség szeretné kiválasztani a válogatott négyszer százméteres vegyes-váltó négy tagját. Mind a négy úszásnemben rendelkezésre áll a szóbajövő úszók legjobb időeredménye. Válasszuk ki a négy úszót és rendeljük hozzájuk a négy különböző úszásnemet úgy, hogy az időeredmények összege minimális legyen.

Láthatjuk, hogy a hozzárendelési problémának több variációja is van. Az egyik alak egy élsúlyozott teljes páros gráfban, amelynek két pontosztálya egyforma elemszámú, maximális súlyú teljes párosítás meghatározását célozza. (**Párosítás**on [matching] olyan gráfot értünk, amelyben minden pont foka legfeljebb egy. A **teljes párosítás**ban minden pont foka pontosan egy.) Ugyanez a kérdés kicsit általánosabb, ha a szóbanforgó páros gráf nem feltétlenül teljes, hanem csak annyit teszünk fel, hogy létezik benne teljes párosítás. Ennek kapcsán vizsgálandó, hogy egy páros gráfban egyáltalán mikor létezik teljes párosítás, illetve ha nem létezik, mekkora a legnagyobb párosítás és azt miként tudjuk meghatározni. Feltehetjük a kérdést, hogy mekkora a maximális súlyú (nem feltételenül teljes) párosítás súlya, vagy a maximális súlyú k élű párosítás súlya. (Az úszóváltó összeállításánál minimális súlyú négyélű párosítást keresünk). Megjegyzendő, hogy a hozzárendelési problémát néha mátrix nyelven fogalmazzák meg. Például: adott egy nemnegatív $n \times n$-es mátrix, válasszuk ki a mátrix n elemét úgy, hogy minden sorból és minden oszlopból egy elem kerül kiválasztásra és a kiválasztott elemek összege maximális. Ez a feladat ekvivalens egy $n \times n$-es élsúlyozott teljes páros gráf maximális súlyú teljes párosításának meghatározásával.

Mindezen problémák megoldására szolgál a Magyar módszer. Az elnevezés H. Kuhn amerikai kutatótól származik, aki egy 1955-ös cikkében Kőnig Dénes és Egerváry Jenő korábbi gondolataira támaszkodva elegáns algoritmust fejlesztett ki maximális súlyú párosítás meghatározására páros gráfban. Félreértés forrása lehet, hogy a szakirodalomban néha a maximális **elemszámú** párosítás megkeresését biztosító alternáló utas eljárást is már Magyar módszernek nevezik. Hangsúlyozzuk azonban, hogy a Kuhn által Magyar módszernek nevezett eljárás maximális **súlyú** teljes párosítás megkeresésére szolgál.

1.42. Gyakorlat. *Mutassuk meg, hogy ha rendelkezésünkre áll egy olyan szubrutin, amelynek segítségével tetszőleges nemnegatív súlyozásra meg tudunk egy maximális súlyú teljes párosítást határozni, akkor egy minimális súlyú teljes párosítást is ki tudunk számítani.*

1.4.1. Maximális elemszámú párosítások: a javító utak módszere

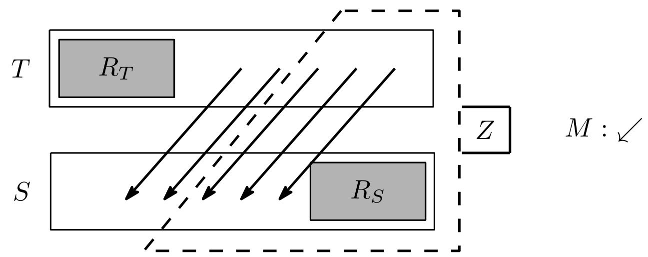
Vizsgálatainkat kezdjük a súlyozatlan eset áttekintésével. A kiindulási eredmény Kőnig Dénes tétele.

1.4.1. Tétel (Kőnig)**.** *Egy $G = (S, T; E)$ páros gráfban a maximális párosítás $\nu = \nu(G)$ elemszáma egyenlő az éleket lefogó pontok minimális $\tau = \tau(G)$ elemszámával.*

Biz. Egy ν elemű párosítás lefogásához kell legalább ν csúcs, így az összes élhez is kell ennyi, ezért $\nu \leq \tau$.

A nemtriviális $\nu \geq \tau$ irány igazolásához konstruálunk egy $M \subseteq E$ párosítást és egy $L \subseteq S \cup T$ lefogást, melyekre $|M| = |L|$. Az eljárás a G egy tetszőleges M párosításából 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 S-től T felé. 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 D_M irányított gráfban irányított úton elérhető pontok halmazát (amit például szélességi kereséssel találhatunk meg).



1.1. ábra. A Kőnig tétel bizonyításának illusztrálása

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 könnyű végrehajtani: a megtalált út éleinek irányítását egyszerűen megfordítjuk. Az így nyert átirányított gráf éppen $D_{M'}$, vagyis az a digráf, amit G-ből kapunk az M' éleinek T-ből S felé, és a többi élnek S-től T felé történő irányításával.)

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. Tehát az M minden eleme vagy teljesen Z-ben fekszik vagy teljesen kívüle.

Következik, hogy az $L := (T \cap Z) \cup (S - Z)$ halmaz egyrészt lefogja a G összes élét, másrészt minden M-beli élnek pontosan az egyik végpontját tartalmazza, tehát $|M| = |L|$, amivel a Kőnig tétel bizonyítása teljes. •

A fenti bizonyítás egyúttal hatékony eljárást is jelent a szóbanforgó optimumok meghatározására. Az algoritmust **Kőnig** (alternáló utas) **algoritmus**ának nevezzük.

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).

Érdemes megfogalmaznunk Kőnig tételét egy ekvivalens alakban. Ehhez definiáljuk egy $X \subseteq S$ halmaz **hiány**át a

$$h(X) := |X| - |\Gamma(X)| \tag{1.15}$$

értékkel, ahol $\Gamma(X) = \Gamma_G(X) = \Gamma_E(X)$ jelöli az X szomszédainak halmazát, vagyis $\Gamma(X) := \{v \in T : \text{létezik } uv \in E \text{ él, melyre } u \in X\}$. Jelölje $\mu = \mu(G, S)$ a maximális hiányt, azaz

$$\mu := \max_{X \subseteq S} h(X) \tag{1.16}$$

Mivel $h(\emptyset) = 0$, a μ értéke mindig nemnegatív. 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.

1.4.2. Lemma. *Egy-egy értelmű kapcsolat áll fenn az élek minimális elemszámú lefogásai és a max-hiányú S-beli halmazok között: ha $L \subseteq S \cup T$ minimális lefogás, akkor $S - L$ max-hiányú halmaz, míg ha $H \subseteq S$ max-hiányú halmaz, akkor $\Gamma(H) \cup (S - H)$ minimális lefogás.*

Biz. Ha $L \subseteq S \cup T$ egy éleket lefogó ponthalmaz, akkor a $H' := S - L$ halmaz $h(H')$ hiánya legalább $|S| - |L|$, hiszen $\Gamma(H') \subseteq L \cap T$ miatt $h(H') \geq |H'| - |L \cap T| = (|S| - |S \cap L|) - |L \cap T| = |S| - |L|$. Másrészt tetszőleges $H \subseteq S$ halmazra $L' := \Gamma(H) \cup (S - H)$ lefogja az éleket és $|L'| = (|S| - |H|) + |\Gamma(H)| = |S| - \mu(H)$. A kettő összevetéséből a lemma következik. •

Jelölje $\varphi = \varphi(G, S)$ azon S-beli pontok minimális számát, melyeket egy párosítás fedetlenül hagy. Nyilván $\varphi + \nu = |S|$. Miután a lemmából $\tau + \mu = |S|$ következik, érvényes Kőnig tételének alábbi, ekvivalens alakja.

1.4.3. Tétel (Kőnig – Hall)**.** *$G = (S, T; E)$ páros gráfban $\varphi = \mu$, azaz egy párosítás által fedetlenül hagyott S-beli pontok minimális száma egyenlő az S részhalmazainak maximális hiányával. Speciálisan, akkor és csak akkor létezik S-t fedő párosítás, ha nincs hiányos halmaz, azaz teljesül a Hall-féle feltétel:*

$$|\Gamma(X)| \geq |X| \text{ minden } X \subseteq S \text{ részhalmazra.} \bullet \tag{1.17}$$

A tételt néha Ore tételének is hívják, míg speciális második része a Hall tétel. Megjegyezzük, hogy Kőnig algoritmusa közvetlenül is kiad egy max-hiányú halmazt. Nevezetesen az algoritmus futásának végén kapott elérhető pontok Z halmazára egyrészt az algoritmus által kiadott (maximális) M párosítás $|R_S|$ darab pontot hagy fedetlenül S-ben, másrészt a

$$H := Z \cap S \tag{1.18}$$

halmazra $\Gamma(H) = Z \cap T$, és így H hiánya pontosan $|R_S|$, tehát H max-hiányú.

A Kőnig algoritmus által szolgáltatott M maximális párosítás természetesen függ a futás során hozott döntéseinktől, hiszen az algoritmus azt nem specifikálja, hogy ha több növelő út is rendelkezésre áll, akkor melyiket használjuk.

1.43. Feladat. *Igazoljuk, hogy a Kőnig algoritmus által kiadott* (1.18)*-beli* H *max-hiányú halmaz független az algoritmus futásától.*

1.4.2. Maximális súlyú teljes párosítások: a Magyar módszer

Tételezzük most fel, hogy a $G = (S, T; E)$ páros gráf élein adott egy c súlyfüggvény. Tegyük fel, hogy G-nek létezik teljes párosítása, és vizsgáljuk meg a maximális súlyú teljes párosítás megkeresésének problémáját. Az első ezzel kapcsolatos kérdés az, hogy milyen hatékonyan ellenőrizhető igazolványt (tanúsítványt) tudunk elképzelni egy kiválasztott teljes párosítás súlyának maximalitására, annak mintájára, ahogy egy megadott M párosítás maximális elemszámára igazolvány az éleknek egy $|M|$ elemszámú lefogása. Ennek általánosításaként nevezzünk egy csúcsokon értelmezett $\pi : V \to \mathbb{R}$ függvényt **súlyozott lefogás**nak, ha minden $uv \in E$ élre $\pi(u) + \pi(v) \geq c(uv)$. (Figyeljük meg, hogy ha c azonosan 1, akkor egy $(0, 1)$-értékű súlyozott lefogás épp az élhalmaz egy lefogásának incidencia vektora). Egy élt **pontos**nak fogunk nevezni (π-re nézve), ha itt egyenlőség áll. A π súlyozott lefogás $\widetilde{\pi}(V)$ **összértékén** a $\sum[\pi(v) : v \in V]$ összeget értjük, ahol $V = S \cup T$. A főtétel Egerváry Jenőtől származik 1931-ből.

1.4.4. Tétel (Egerváry). *A* $G = (S, T; E)$ *teljes párosítással rendelkező páros gráfban a* $c : E \to \mathbb{R}_+$ *súlyfüggvényre vonatkozó maximális súlyú teljes párosítás* ν_c *súlya egyenlő a súlyozott lefogások minimális* τ_c *összértékével. Amennyiben* c *egészértékű, úgy az optimális súlyozott lefogás is választható annak. Amennyiben* G *teljes páros gráf és* c *nemnegatív, úgy az optimális súlyozott lefogás választható nemnegatívnak is.*

Biz. Megjegyezzük, hogy a tétel implicit azt is tartalmazza, hogy a szóbanforgó minimum létezik. Mivel a maximumot a teljes párosítások véges halmazán tekintjük, így annak létezése nem kérdéses.

A harmadik rész igazolásához azt mutatjuk meg, hogy tetszőleges π súlyozott lefogás nemnegatívvá alakítható az összérték megváltoztatása nélkül. Legyen a π legkisebb értéke $-K$ (ahol $K > 0$) és legyen mondjuk az S-ben $-K$ értékű pont. Mivel G teljes páros gráf, c nemnegatív és π súlyozott lefogás, így minden $v \in T$ pontra $\pi(v) \geq K$. Az S elemein a π értékeket egységesen K-val növelve, T elemein pedig K-val csökkentve olyan nemnegatív súlyozott lefogást kapunk, melynek összértéke $|S| = |T|$ miatt szintén $\widetilde{\pi}(V)$.

A tétel min = max részének igazolásához először azt látjuk be, hogy max $\leq$ min. Tekintsük ehhez a gráf egy tetszőleges $M := \{u_1v_1, u_2v_2, \ldots, u_nv_n\}$ teljes párosítását valamint a c-nek egy π súlyozott lefogását. Ekkor $\widetilde{c}(M) := \sum c(u_iv_i) \leq \sum[\pi(u_i) + \pi(v_i) : i = 1, \ldots, n] = \widetilde{\pi}(V)$, amiből $\nu_c \leq \pi_c$ következik. Itt egyenlőség pontosan akkor áll, ha M minden élen pontos.

A fordított irányú egyenlőtlenség bizonyításához tehát kell találnunk egy alkalmas π súlyozott lefogást és egy olyan teljes párosítást, amely pontos élekből áll. Más szóval olyan π-t kell keresnünk, hogy a π-re vonatkozó pontos élekből álló részgráf tartalmazzon teljes párosítást.

Erre szolgál a H. Kuhn által bevezetett Magyar módszer. Tetszőleges π súlyozott lefogással indulunk, amely egészértékű, ha c az. (Hogyan lehet ilyent találni?) Az általános lépésben tekintjük a pontos élek által alkotott $G_\pi = (S, T; E_\pi)$ részgráfot és ezen futtatjuk a Kőnig tétel fentebb leírt bizonyításának algoritmusát. Kiindulunk tehát a G_π-nek egy tetszőleges M párosításából. Megirányítjuk az M-beli éleket T-től S felé, míg az összes többi G_π-beli élt S-től T felé. 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.

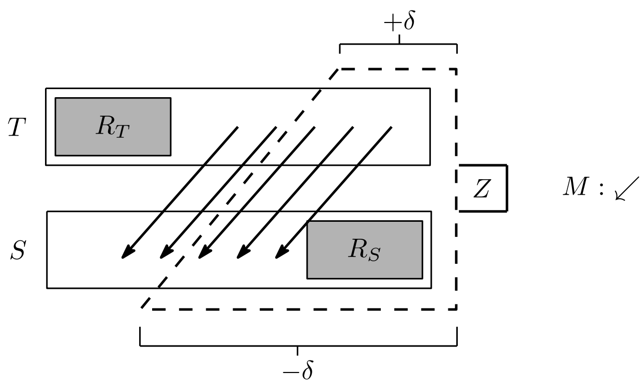
Amennyiben R_T-nek esik pontja Z-be, úgy megkapunk egy olyan R_S-t és R_T-t összekötő P utat, amely M-ben alternál. Az M és P szimmetrikus differenciája egy M-nél eggyel több élből álló M' párosítást alkot. Az M'-vel folytatva iteráljuk az eljárást.

Nézzük most a másik lehetőséget, amikor R_T diszjunkt Z-től. Z definíciója folytán Z-ből nem vezet ki irányított él és Z-be nem lép be megirányított $uv \in M$ párosítás él. Legyen $H := Z \cap S$. Mivel G-nek van teljes párosítása, biztosan van olyan e éle G-nek, amely H és $T - \Gamma_{G_\pi}(H)$ között vezet, ahol $\Gamma_{G_\pi}(H)$ jelöli a H szomszédainak halmazát a G_π-ben. Ilyen él nem lehet pontos, így a

$$\delta := \min\{\pi(u) + \pi(v) - c(uv) : uv \in E, u \in H, v \in T - \Gamma_{G_\pi}(H)\} \tag{1.19}$$

érték pozitív. Módosítsuk π-t a következőképp:

$$\pi'(v) = \begin{cases} \pi(v) - \delta, & \text{ha } v \in H, \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_{G_\pi}(H), \\ \pi(v) & \text{különben.} \end{cases} \tag{1.20}$$



1.2. ábra. A π változtatása a Magyar módszer egy lépésében

A δ választása miatt az így módosított π' továbbra is súlyozott lefogás, amely egészértékű, ha c és π az volt. A π'-re vonatkozó pontos élek $G_{\pi'}$ gráfja és G_π ugyanazon éleket feszíti Z-ben, továbbá $G_{\pi'}$-nek van legalább egy éle (ahol a δ-t definiáló minimum felvétetett) H és $T - \Gamma_{G_\pi}(H)$ között, ezért $G_{\pi'}$-ben az R_S-ből elérhető pontok halmaza szigorúan bővebb, mint G_π-ben. (Figyelem: az NEM igaz, hogy $G_{\pi'}$-ben biztosan több pontos él van, mint G_π-ben.)

Emiatt a $G_{\pi'}$-höz és M-hez rendelt irányított gráfban az R_S-ből elérhető pontok halmaza szigorúan bővebb. Így egy fázis (ami során tehát a pontos élek gráfjában a maximális párosítás elemszáma nem nő) legfeljebb $|S|$ útkereső eljárás alkalmazása után véget ér. Ezzel a min-max tétel bizonyítását befejeztük. A tétel második állításához figyeljük meg, hogy ha c egészértékű, akkor a fenti eljárás során π egészértékűsége végig megőrződik. •

Mivel egy útkeresés $O(|E|)$ lépésben végrehajtható és legfeljebb $|S|$ fázis van, az algoritmus teljes futásideje $O(|E||S|^2)$.

1.44. Gyakorlat. *Igazoljuk, hogy a minimális összértékű súlyozott lefogások konvex halmazt alkotnak.*

1.45. Feladat. *Igazoljuk, hogy ha egy M teljes párosítás minden éle eleme valamely maximális súlyú teljes párosításnak, akkor M maga is maximális súlyú teljes párosítás.*

1.46. Feladat. *Legyen π egy minimális összértékű súlyozott lefogás és $G_\pi = (S,T;E_\pi)$ a pontos élek gráfja. Igazoljuk, hogy G egy M teljes párosítása akkor és csak akkor maximális súlyú, ha $M \subseteq E_\pi$.*

Alternatív bizonyítás: javítás negatív kör mentén

Bemutatunk egy másik bizonyítási módszert is az Egerváry tétel nemtriviális max $\geq$ min irányának igazolására, azonban most nem célunk hatékony algoritmus kiolvasása. Legyen M egy maximális súlyú teljes párosítás. Célunk egy olyan π függvény megkonstruálása, amelyre

$$\text{minden } uv \in M \text{ élre } \pi(u) + \pi(v) = c(uv) \tag{1.21}$$

és

$$\text{minden } uv \in E - M \text{ élre } \pi(u) + \pi(v) \geq c(uv). \tag{1.22}$$

Valójában egy olyan π-t adunk meg, amelyre

$$\text{minden } uv \in M \text{ élre } \pi(u) + \pi(v) \leq c(uv) \tag{1.23}$$

és

$$\text{minden } uv \in E - M \text{ élre } \pi(u) + \pi(v) \geq c(uv) \tag{1.24}$$

Ennek ugyanis alkalmas növelésével az (1.21)-t és (1.22)-t kielégítő π könnyen megkapható.

Irányítsuk az M-beli éleket S felé, a többi élt T felé, majd az $E - M$-beli élek költségét negáljuk. Jelölje D' a kapott digráfot és c' a módosított súlyozást. Állítjuk, hogy c' konzervatív. Ha ugyanis létezne D'-ben egy K' negatív kör, akkor ennek G-ben egy olyan M-ben alternáló K kör felelne meg, amelyre $\widetilde{c'}(K') < 0$ miatt $\widetilde{c}(K - M) > \widetilde{c}(K \cap M)$, és ezért a K mentén az M elemeinek kicserélésével kapott $M' := M \ominus K$ teljes párosítás súlya nagyobb, mint M súlya, ellentmondásban az M választásával. (Itt $\ominus$ a szimmetrikus differenciát jelöli, azaz $M \ominus K = (M - K) \cup (K - M)$.)

Mivel c' konzervatív, Gallai 1.3.8 tétele folytán létezik egy π' megengedett potenciál. Ez azt jelenti, hogy egy M-beli xy élre ($x \in T, y \in S$) $\pi'(y) - \pi'(x) \leq c(xy)$, míg egy $E - M$-beli uv élre ($u \in S, v \in T$) $\pi'(v) - \pi'(u) \leq c'(uv) = -c(uv)$.

Negáljuk a T elemein a π' értékeit és a kapott függvényt jelölje π. Ekkor a $\pi'(y) - \pi'(x) \leq c(xy)$ egyenlőtlenségből $\pi(y) + \pi(x) \leq c(xy)$ lesz, míg $\pi'(v) - \pi'(u) \leq -c(uv)$-ből $-\pi(v) - \pi(u) \leq -c(uv)$ azaz $\pi(v) + \pi(u) \geq c(uv)$, vagyis (1.23) és (1.24) teljesül. •

A fenti bizonyításból egy algoritmus is kiolvasható. Induljunk ki egy tetszőleges M teljes párosításból. Amennyiben az ehhez rendelt D' digráfban a c' súlyozás konzervatív, úgy a bizonyítás szerint M maximális súlyú. Ha D'-ben van negatív kör, akkor ennek segítségével a bizonyításban leírtak szerint megkapunk egy M-nél nagyobb súlyú teljes párosítást, amivel iterálhatjuk az eljárást. Az így nyert algoritmus véges, hiszen mindig egy jobb teljes párosítást kapunk, sőt abban az esetben polinomiális is, amikor a c súlyfüggvény nemnegatív (ami feltehető) és kis egészekből áll. Ugyanis ha c legnagyobb értéke M, akkor a maximális súlyú teljes párosítás súlya legfeljebb $M|S|$ és ezért legfeljebb $M|S|$ párosítás javítás lehetséges. Ha viszont nincs előre adott felső korlát a c értékeire vagy ha c valós értékű, akkor a fenti algoritmusról kimutatható, hogy nem polinomiális futásidejű.

1.4.3. Egerváry eredeti bizonyítása és algoritmusa

Egerváry az 1.4.4 tételt eredetileg egészértékű c-re bizonyította. Ebből a közös nevezővel való felszorzással a tétel könnyen következik racionális súlyfüggvényekre is. Tetszőleges valós súlyfüggvényekre pedig a tételt Egerváry folytonossági megfontolásokkal vezette le.

Legyen tehát c egészértékű. Legyen π egészértékű súlyozott lefogás, melynek összértéke minimális. (Jogos minimumról beszélni, hiszen egészértékű súlyozott lefogásokról van szó, és az ilyenek összértéke korlátos alulról.) Legyen $G_\pi = (S, T; E_\pi)$ a pontos élek részgráfja. A G_π-nek bármely teljes párosítása maximális súlyú, hiszen pontos élekből áll. Belátjuk, hogy G_π-nek van teljes párosítása. Ha indirekt nem ez a helyzet, akkor a Kőnig-Hall tétel nyomán létezik egy $X \subseteq S$ hiányos halmaz, amelyre tehát $|\Gamma_{G_\pi}(X)| < |X|$. Ennek segítségével tudunk π-n javítani. Definiáljuk δ-t a következőképpen.

$$\delta := \min\{\pi(u) + \pi(v) - c(uv) : uv \in E, u \in X, v \in T - \Gamma_{G_\pi}(X)\}. \tag{1.25}$$

Miután nincsen pontos él X és $T - \Gamma_{G_\pi}(X)$ között, így δ pozitív és persze egész. Módosítsuk π-t a következőképp:

$$\pi'(v) = \begin{cases} \pi(v) - \delta, & \text{ha } v \in X, \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_{G_\pi}(X), \\ \pi(v) & \text{különben.} \end{cases} \tag{1.26}$$

Az így nyert π' továbbra is súlyozott lefogás, amelyre $\widetilde{\pi}'(V) = \widetilde{\pi}(V) - \delta|X| + \delta|\Gamma_{G_\pi}(X)| < \widetilde{\pi}(V)$, ellentmondásban π minimális választásával. •

Egerváry ezen bizonyítási módszere egyúttal algoritmust is jelent maximális súlyú teljes párosítás és minimális összértékű súlyozott lefogás kiszámítására: kiindulunk egy tetszőleges egészértékű π súlyozott lefogásból, és a pontos élek G_π gráfjában (például Kőnig algoritmusával) vagy találunk egy teljes párosítást, amely értelemszerűen maximális súlyú, vagypedig találunk egy hiányos halmazt, amelynek segítségével a bizonyításban leírt módon javítjuk a súlyozott lefogást. A módosított súlyozott lefogással folytatva iteráljuk az eljárást. Nevezzük ezt az eljárást **Egerváry algoritmus**ának.

Kimutatható, hogy az algoritmus ezen generikus alakja (amikor a π javítására használt X hiányos halmazt szabadon választjuk) egész vagy racionális c-re még akkor sem polinomiális, ha mindig max-hiányú halmazzal dolgozunk. Ráadásul valós súlyfüggvény esetén az algoritmus még csak nem is biztosan véges. Megmutatjuk azonban, hogy a max-hiányú halmazok speciális választása esetén Egerváry algoritmusa még valós c esetén is polinomiális. Ehhez szükségünk lesz majd az alábbi hasznos megfigyelésekre.

Max-hiányú halmazok

1.4.5. Lemma. *Az S halmaz részhalmazain értelmezett $\gamma(X) := |\Gamma(X)|$ függvény szubmoduláris, azaz az S bármely két X, Y részhalmazára fennáll a* **szubmodularitási egyenlőtlenség***:*

$$\gamma(X) + \gamma(Y) \geq \gamma(X \cap Y) + \gamma(X \cup Y).$$

Biz. Az egyenlőtlenség következik, amint megfigyeljük, hogy $\Gamma(X) \cup \Gamma(Y) = \Gamma(X \cup Y)$ és $\Gamma(X) \cap \Gamma(Y) \supseteq \Gamma(X \cap Y)$. •

1.4.6. Lemma. *A max-hiányú halmazok $\mathcal{F}$ rendszere zárt a metszet és unió képzésre.*

Biz. Mivel a $|\Gamma(X)|$ függvény szubmoduláris, így $h(X)+h(Y) \leq h(X\cap Y)+h(X\cup Y)$. 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 $h(X \cap Y) = \mu, h(X \cup Y) = \mu$. •

Az 1.4.6 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.

1.4.7. Tétel. *Kőnig alternáló utas algoritmusa által szolgáltatott* (1.18)*-beli H max-hiányú halmaz az egyértelmű legszűkebb max-hiányú halmaz* (és így nem függ az algoritmus futása közben tett választásoktól).

Biz. Mivel az M maximális párosítás fedi $S - R_S$ minden pontját, ezért tetszőleges max-hiányú halmaz tartalmazza R_S-t. Bármely X halmazra, amelyre $R_S \subseteq X \subset H$, a szóbanforgó irányított gráfban lép ki $X \cup \Gamma_M(X)$-ből egy uv él és így $\Gamma(X) \supseteq \Gamma_M(X) \cup \{v\}$. Így $|\Gamma(X)| > |\Gamma_M(X)| = |X| - |R_S|$, azaz $h(X) < |R_S|$, tehát X nem max-hiányú. •

1.47. Feladat. *Hogyan lehet az alternáló utas algoritmus segítségével az egyértelmű legbővebb max-hiányú halmazt megkonstruálni?*

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

Biz. Miután G egy M maximális párosításának a $\Gamma(H)$-t fedő élei mind H-ban végződnek, az M benne van G'-ben is, vagyis G' max hiánya legfeljebb akkora, mint G-é, és 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 H max-hiányú G'-ben is, az 1.4.6 lemma szerint $H \cap X$ is max-hiányú G'-ben. De akkor $H \cap X$ max-hiányú G-ben, hiszen $H \cap X$-ből induló élt nem töröltünk, és így a H minimalitása folytán $H \subseteq X$. Ekkor viszont $\Gamma(X) = \Gamma'(X)$, azaz X max-hiányú G-ben is. •

1.4.9. Tétel. *Amennyiben az Egerváry algoritmus futtatásakor a szóbanforgó π súlyozott lefogás javítására a Kőnig algoritmus által szolgáltatott egyértelmű legszűkebb max-hiányú X halmazt használjuk, úgy az algoritmus polinomiális futásidejű.*

Biz. Figyeljük meg először, hogy miként változik a pontos élek gráfja, amikor az algoritmus π-ről π'-re tér át. Mindenesetre az $X \cup \Gamma_{G_\pi}(X)$ valamint ennek komplementere által feszített pontos élek nem változnak. Az $S - X$ és $\Gamma_{G_\pi}(X)$ között vezető esetleges pontos élek megszűnnek pontosnak lenni, míg az X és $T - \Gamma_{G_\pi}(X)$ között vezető élek közül mindazok pontossá válnak, amelyeken az (1.25) definícióval megadott minimum felvétetik. Az 1.4.8 lemmából következik, hogy ezen cserénél a maximális hiány vagy csökken, vagy ha nem, úgy a legszűkebb max-hiányú halmaz szigorúan bővül.

Tekintsük egy fázisnak az algoritmus futásának azon szakaszát, amely során a pontos élek (egyre változó) részgráfjában a maximális hiány változatlan. Egyetlen fázis során a max-hiányú halmaz legfeljebb $|S|$-szer tud bővülni és nyilván legfeljebb $|S|$ fázis létezik. Vagyis a Kőnig algoritmus 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 Kuhn Magyar módszere hatékonyabb), azt mindenesetre 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 összadást, kivonást és összehasonlítást egyetlen lépésben tudjuk elvégezni.

1.4.4. Maximális súlyú párosítások

Visszavezetés teljes párosításra

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

1.4.10. 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'_c is választható egészértékűnek.*

Biz. 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. Az 1.4.4 tétel (harmadik része) szerint G-nek létezik egy M teljes párosítása és c-nek egy π nemnegatív súlyozott lefogása, melyekre $\widetilde{c}(M) = \widetilde{\pi}(V)$. Mivel az új élek súlya 0, így az új élek kihagyásával M-ből keletkező G'-beli M' párosítás súlya változatlanul $\widetilde{c}(M)$. Továbbá, mivel M minden éle pontos, ezért egy 0 súlyú uv élének végpontjaira $\pi(u) = \pi(v) = 0$. Emiatt π értéke az új pontokon 0, hiszen új pontból csak 0 súlyú él megy ki.

Ha tehát π-t megszorítjuk az eredeti $V' = S' \cup T'$ ponthalmazra, akkor a keletkező π'-re $\widetilde{\pi}'(V') = \widetilde{\pi}(V)$ és $\widetilde{\pi}'(V') = \widetilde{c}(M')$. •

Direkt eljárás

Az 1.4.10 tétel nemtriviális részének igazolására bemutatunk egy direkt algoritmust is. Célunk tehát egy $\pi : S \cup T \to \mathbb{R}_+$ nemnegatív súlyozott lefogást valamint egy π-re nézve pontos élekből álló M párosítást találni úgy, hogy M minden pozítív $\pi(v)$ értékű, röviden **pozitív pont**ot fed.

Kezdetben legyen π a T elemein azonosan 0 és minden $s \in S$ ponton $\pi(s) := \max\{c(st) : t \in T\}$. Legyen továbbá M egy pontos élekből álló párosítás, például az üres halmaz. Az algoritmus egy közbenső helyzetében rendelkezésünkre áll egy $\pi \geq 0$ súlyozott lefogás és egy pontos élekből álló M párosítás úgy, hogy M minden T-beli pozitív pontot fed. Ezt a tulajdonságot végig fenntartjuk, miközben a π és az M változtatásával egyre csökkentjük a fedetlen pozitív pontok halmazát.

Jelölje R_T a T-beli fedetlen pontok halmazát és R_S^+ az S-beli pozitív fedetlen pontok halmazát. A pontos élek G_π részgráfjában irányítsuk meg M éleit S felé, a többi élt pedig T felé. A keletkező digráfban jelölje Z az R_S^+-ból elérhető pontok halmazát. A Kőnig tétel algoritmikus bizonyításában látottakhoz hasonlóan most is M minden éle vagy teljesen Z-ben fekszik vagy teljesen kívüle.

I. eset $Z \cap R_T \neq \emptyset$, azaz D-ben létezik egy P egyirányú út R_S^+-ból R_T-be. A P mentén cserélve egy (M-nél eggyel nagyobb elemszámú) M' párosítást kapunk, amely eggyel kevesebb pozitív pontot hagy fedetlenül, mint az M (nevezetesen a P kezdőpontját M nem fedi, de M' már igen).

II. eset $Z \cap R_T = \emptyset$. Legyen

$$\delta_1 := \min\{\pi(u) + \pi(v) - c(uv) : uv \in E,\ u \in Z \cap S,\ v \in T - \Gamma_{G_\pi}(Z \cap S)\}. \quad (1.27)$$

Az üres halmazon vett minimumot $+\infty$-nek definiáljuk. (Mivel most nem tettük fel, hogy létezik teljes párosítás, előfordulhat, hogy G-nek nem létezik éle $Z \cap S$ és $T - \Gamma_{G_\pi}(Z \cap S)$ között.) Mivel a digráfban Z-ből nem lép ki él, a δ_1 érték szigorúan pozitív. Legyen

$$\delta_2 := \min\{\pi(s) : s \in Z \cap S\}. \quad (1.28)$$

Itt δ_2 véges, de előfordulhat, hogy 0. Végül legyen $\delta := \min\{\delta_1, \delta_2\}$. Módosítsuk π-t a következőképp:

$$\pi'(v) = \begin{cases} \pi(v) - \delta, & \text{ha } v \in Z \cap S, \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_{G_\pi}(Z \cap S), \\ \pi(v) & \text{különben.} \end{cases} \quad (1.29)$$

A δ választása miatt az így nyert π' is súlyozott lefogás és továbbra sincs T-ben fedetlen pozitív pont (merthogy T-ben csak fedett pontok π-jét növelhettük). Amennyiben $\delta = \delta_1 < \delta_2$, úgy $S \cap Z$ valamennyi pontja pozitív maradt. Ezért a π'-höz és a változatlan M-hez tartozó digráfban az elérhető pontok halmaza szigorúan bővebb, mint Z. Ebből következik, hogy a $\delta = \delta_1 < \delta_2$ eset egymást követően legfeljebb $|M| < |S|$-szer fordulhat elő.

Tegyük most fel, hogy $\delta = \delta_2 \leq \delta_1$, és legyen $z \in S \cap Z$ egy olyan pont, amelyre a δ_2 definíciójában a minimum eléretik, ami azt jelenti, hogy $\pi'(z) = 0$. Mivel $z \in Z$, létezett R_S^+-ból z-be P egyirányú út, amely mentén cserélve egy $|M|$-mel megegyező elemszámú M' párosítást kapunk (amely $z \in R_S^+$ esetén maga M), de M' eggyel kevesebb pozitív pontot hagy szabadon, mint M.

Összefoglalva megállapíthatjuk, hogy az algoritmus során a fedetlen pozitív pontok száma legfeljebb $|S|$-szer csökkenhet (akár mert az I. eset során növelő utat találtunk, akár mert a II. eset során a $\delta = \delta_2 \leq \delta_1$ eset következett be). Továbbá két ilyen csökkenés között a Z legfeljebb $|S|$-szer bővülhet. Miután az elérhető pontok halmazát $O(|E|)$ lépésben tudjuk kiszámolni az eljárás $O(|E||S|^2)$ lépés után megadja egy súlyozott lefogást és egy pontos élekből álló párosítást, amely minden pozitív pontot fed. •

Végül megjegyezzük, hogy tetszőleges rögzített pozitív egész k-ra Ford és Fulkerson később ismertetésre kerülő minimális költségű folyam algoritmusának segítségével ki lehet számolni a maximális (vagy minimális) súlyú k élű párosítást, ha ilyen párosítás létezik egyáltalán. Ennek a megközelítésnek az lesz majd az előnye, hogy minden $k = 1, 2, \dots, \nu(G)$ értékre megadja a legolcsóbb k élű párosítást.

Ha viszont valóban csak egyetlen rögzített k-ra kell ezt meghatároznunk, akkor közvetlenebb módon is eljárhatunk. Ehhez mindenesetre feltesszük, hogy $\nu(G) \geq k$, azaz létezik k élből álló párosítás. Speciálisan adódik, hogy mind S, mind T legalább k elemű. Feltehetjük, hogy minden él költsége szigorúan pozitív, hiszen konstans hozzáadása a k élű párosítások egymáshoz való költség viszonyát nem érinti. Adjunk S-hez $|T| - k$ új pontot és T-hez $|S| - k$ új pontot, az új pontokat kössük össze a másik osztály valamennyi régi pontjával, és az új élek költségét válasszuk azonosan nullának. Legyen a megnövelt gráf G^+.

Könnyű ellenőrizni, hogy egy eredeti k élű párosítást új élekkel kiegészíthetünk egy ugyanolyan költségű G^+-beli teljes párosítássá, és megfordítva, egy G^+-beli teljes párosításból az új éleket kihagyva G-nek nyerünk egy ugyanolyan költségű k élű párosítását. Ebből adódik, hogy egy legolcsóbb G^+-beli teljes párosítás eredeti élei egy legolcsóbb k élű párosítást adnak G-ben.

1.48. Feladat. *Tegyük fel, hogy a $G = (S, T; E)$ páros gráfban M egy j élű párosítás, amely a j élű párosítások között maximális súlyú a $c : E \to \mathbb{R}$ súlyfüggvényre nézve. Jelölje R_S illetve R_T az M által fedetlen S-beli illetve T-beli pontok halmazát. Irányítsuk M éleit S felé, a többi élt T felé. A kapott D digráf élhalmazán definiáljuk a c' költségfüggvényt úgy, hogy egy ts él ($t \in T, s \in S$) költsége legyen $c(ts)$, egy uv élé ($u \in S, v \in T$) pedig $-c(uv)$.* (a) *Igazoljuk, hogy c' konzervatív.* (b) *Igazoljuk, hogy ha P egy legolcsóbb út R_S-ből R_T-be, akkor az $M \ominus P$ $j + 1$ elemű párosítás maximális súlyú a $j + 1$ élű párosítások között.*

1.5. Áramok és folyamok hálózatokban

1.5.1. Fogalmak

Ebben a részben hálózatokra vonatkozó két rokon fogalommal foglalkozunk: áramokkal és folyamokkal.

Áramok

Jelöljön $D = (V, A)$ egy irányított gráfot. Valamely $x : A \to \mathbb{R}$ függvényre és $S \subseteq V$ részhalmazra legyen $\varrho_x(S) := \sum[x(uv) : uv \in A, uv$ belép S-be$]$ és legyen $\delta_x(S) := \varrho_x(V - S)$. Azt mondjuk, hogy x **áram** (circulation), ha teljesül rá a **megmaradási szabály** (conservation rule), azaz $\varrho_x(v) = \delta_x(v)$ fennáll minden v csúcsra. Valamely $c : A \to \mathbb{R}$ költségfüggvényre vonatkozólag a $cx := \sum[c(a)x(a) : a \in A]$ skalárszorzatot nevezzük az x áram **költség**ének.

1.5.1. Állítás. (a) *Az x függvény $D = (V, A)$ élhalmazán akkor és csak akkor áram, ha $\varrho_x(v) \leq \delta_x(v)$ fennáll minden v csúcsra.* **(b)** *Ha x áram, akkor $\varrho_x(Z) = \delta_x(Z)$ minden $Z \subseteq V$ részhalmazra is fennáll.*

Biz. (a) $\widetilde{x}(A) = \sum[\varrho_x(v) : v \in V] \leq \sum[\delta_x(v) : v \in V] = \widetilde{x}(A)$, amiből $\varrho_x(v) = \delta_x(v)$ következik minden $v \in V$-re. (b) Jelölje a Z által feszített élek halmazán az x-értékek összegét $i_x(Z)$. Ekkor $\varrho_x(Z) = \sum[\varrho_x(v) : v \in Z] - i_x(Z) = \sum[\delta_x(v) : v \in Z] - i_x(Z) = \delta_x(Z)$. •

Legyen $f : A \to \mathbb{R} \cup \{-\infty\}$ alsó kapacitás, $g : A \to \mathbb{R} \cup \{+\infty\}$ felső kapacitás úgy, hogy $f \leq g$. Azt mondjuk hogy az x áram **megengedett** (feasible), ha

$$f \leq x \leq g. \tag{1.30}$$

(Figyelem: az f-ben megengedünk $-\infty$ komponenst, ami persze csak annyit jelent, hogy az illető élen az áram értéke nincs alulról korlátozva. Analóg módon a g-nek lehetnek $+\infty$ komponensei, de az x áram komponensei mindig végesek. Az f alsó korlátban $+\infty$-t, a g felső korlátban pedig $-\infty$-t nem engedünk meg. Néha előírjuk, hogy az f vagy a g komponensei egészértékűek legyenek; ebbe beleértjük a $\pm\infty$-t is.)

A vizsgálandó fő kérdés az, hogy mikor létezik (egészértékű) megengedett áram, illetve ha létezik, miképp lehet meghatározni egy minimális költségű megengedett áramot.

Folyamok

Az árammal rokon a folyam fogalma. Ismét adott egy $D = (V, A)$ irányított gráf, továbbá D-nek egy kijelölt s forráspontja (source) és egy t nyelőpontja (sink). A továbbiakban, amikor folyamokról lesz szó, végig feltesszük, hogy s-be nem lép be él és t-ből nem lép ki él. **Folyam**on egy olyan $x : A \to \mathbb{R}_+$ nem-negatív függvényt értünk, amely minden, s-től és t-től különböző pontra teljesíti a megmaradási szabályt, azaz $\varrho_x(v) = \delta_x(v)$ fennáll minden $v \in V - \{s, t\}$ csúcsra. Amennyiben még az $x \leq g$ feltétel is teljesül, g-**megengedett** (röviden, **megengedett**) **folyam**ról beszélünk. Egy $(0, 1)$-értékű folyamot **fonat**nak nevezünk. Az x fonat azonosítható azon élek által alkotott részgráffal, melyeken az x értéke 1. Ez tehát egy olyan részgráfot alkot, amelyben az s és t kivitelével minden v csúcs befoka és kifoka megegyezik.

Egy s-et tartalmazó, de t-t nem tartalmazó X halmazt nevezzünk $s\bar{t}$-halmaznak. Tetszőleges S $s\bar{t}$-halmaz és x folyam esetén

$$\delta_x(s) = \delta_x(s) - \varrho_x(s) = \sum[\delta_x(v) - \varrho_x(v) : v \in S] = \delta_x(S) - \varrho_x(S). \tag{1.31}$$

Vagyis minden S $s\bar{t}$-halmazra a $\delta_x(S) - \varrho_x(S)$ értékkel definiált **tiszta kiáramlás** független az S választásától. Ezt a közös, $\delta_x(s)$-sel (és $\varrho_x(t) = \delta_x(V - t)$-vel) egyenlő értéket nevezzük az x folyam **nagyság**ának (flow amount). Az $x(uv)$ szám a folyam **értéke** az $uv \in A$ élen. (Figyelem: a szakirodalomban nem ritka, hogy a $\delta_x(s)$ számot az általunk használt folyam nagyság helyett az x folyam értékének [flow value] hívják. Ez amiatt nem szerencsés, mert összekeverhető a folyamnak egy e élen felvett $x(e)$ értékével.) Egy k nagyságú fonatot röviden k**-fonat**nak nevezünk. (Az elnevezés arra utal, hogy azon e élek halmaza, melyeken $x(e) = 1$ felbomlik k élidegen st-útra és körökre.) Az x folyamot **út-folyam**nak (**kör-folyam**nak) nevezzük, ha x csak egy s-ből t-be vezető egyirányú út (egyirányú kör) mentén pozitív.

A fő kérdés, hogy miként lehet meghatározni egy maximális nagyságú (egészértékű) folyamot, illetve adott költségfüggvény esetén hogyan lehet kiszámítani egy előre adott k értékre a k nagyságú folyamok közül a minimális költségűt. (Valamely $c : A \to \mathbb{R}$ költségfüggvényre nézve a $cx := \sum[c(a)x(a) : a \in A]$ skalárszorzatot nevezzük az x folyam **költség**ének.)

1.49. Gyakorlat. *Igazoljuk, hogy minden nem-negatív folyam előállítható kör- folyamok és út-folyamok nem-negatív lineáris kombinációjaként. Speciálisan, minden fonat páronként élidegen körök és st-utak uniója.*

1.5.2. Motivációk

Megemlítünk néhány természetes gyakorlati feladatot, melyek megoldása áramok vagy folyamok segítségével történhet.

A szállítási probléma

Adott k üzem mindegyike ugyanazt a terméket állítja elő és tudjuk, hogy mekkora az egyes üzemek maximális

kibocsátó képessége (supply). Adott továbbá ℓ fogyasztóhely, melyeknek ismerjük az igényeit (demand). Tudjuk, hogy mely üzemekből mely fogyasztókhoz milyen átbocsátó kapacitással lehet szállítani, és hogy mennyi az egységnyi termék szállításának a költsége. Döntsük el, hogy az adott feltételek mellett létezik-e olyan szállítási terv, amely kielégíti a fogyasztók igényeit, és ha létezik, keressük meg a legolcsóbb megoldást. Ez a **szállítási feladat** (transportation problem). Ha minden kibocsátás és igény azonosan 1, akkor a szállítási feladat a korábban megismert hozzárendelési problémára redukálódik.

Gráfnyelven a szállítási feladat a következőképpen fogalmazható meg. Adott egy $G = (S, T; E)$ páros gráf. Az S-beli pontok felelnek meg az üzemeknek, míg T elemei a felhasználóknak. Minden $v \in S$-beli ponthoz adott egy $q(v)$ szám, amely az illető üzem kibocsátóképességét jelzi. Minden $v \in T$-beli ponthoz adott egy $h(v)$ szám, amely v igényét jelzi. Adott még a gráf élein egy $g : E \to \mathbb{R}_+$ kapacitásfüggvény valamint egy $c : E \to \mathbb{R}_+$ költségfüggvény. Az első feladat annak eldöntése, hogy létezik-e olyan $x : E \to \mathbb{R}_+$ függvény, amelyre $0 \leq x \leq g$, $d_x(v) \leq q(v)$ minden $v \in S$-re, és $d_x(v) = h(v)$ minden $v \in T$-re. Amennyiben létezik ilyen x, úgy a második feladat egy olyan x meghatározásából áll, amely minimalizálja a $cx = \sum[c(e)x(e) : e \in E]$ összköltséget.

A szállítási (és speciális esetként a hozzárendelési) feladatot a következőképp lehet folyam feladatként megfogalmazni. Irányítsuk meg a G gráf éleit S-től T-felé. Adjunk a gráfhoz egy új s pontot, amelyből minden $v \in S$ pontba vezet egy $q(v)$ kapacitású él. Adjunk a gráfhoz egy új t pontot, amelybe minden $v \in T$ pontból vezet egy $h(v)$ kapacitású él. A szállítási feladatnak pontosan akkor van megoldása, ha az így keletkezett D digráfban van $M := \sum_{v \in T} h(v)$ nagyságú folyam. A költséges változat egy minimális költségű M nagyságú folyam meghatározását célozza.

Élidegen utak

Alkalmazásokban gyakran vetődik fel a kérdés, hogy mikor létezik irányított gráf valamely pontjából egy megadott másikba k páronként élidegen (vagy pontidegen) út. Nem fog zavart okozni, hogy a rövidség kedvéért (bár kissé pontatlanul) a továbbiakban mér nem tesszük ki a „páronként” határozószót.

Erre az elvi választ Menger tétele adja meg, amelynek különféle változatai vannak, annak megfelelően, hogy élidegen vagy pontidegen utakat keresünk irányított vagy irányítatlan gráfban. Az irányított élidegen verzió szerint *akkor és csak akkor létezik s-ből t-be k élidegen út, ha minden s-et tartalmazó $S \subseteq V$ $s\bar{t}$-halmaz kifoka legalább k.* Kérdés, hogy algoritmikusan miként lehet megtalálni k élidegen utat, illetve ha nincs megoldás, akkor hogyan határozható meg a Menger tétel által biztosított k-nál kisebb kifokú S halmaz. (Egy ilyen halmaz gyorsan ellenőrizhető igazolványként szolgálhat arra, hogy D-ben nem létezik k élidegen st-út.) Még összetettebb feladatot kapunk, ha az éleken adott költségfüggvényre vonatkozólag úgy akarunk k élidegen utat keresni, hogy összköltségük minimális legyen.

Mindenesetre a következő nagyon speciális esetben még csak Menger tételére sincs szükségünk. Tegyük fel, hogy a $D' = (V, A')$ digráf maga egy fonat, azaz

(i) s-be nem lép be él,

(ii) t-ből nem lép ki él,

(iii) minden más v csúcs $\varrho(v)$ befoka egyenlő a csúcs $\delta(v)$ kifokával.

Ebben az esetben bizonyosan létezik $\delta(s)$ élidegen út st-út. Valóban, induljunk ki s-ből, majd amíg csak lehetséges, haladjunk tovább addig még nem használt él mentén. A fokszámokra tett feltételek miatt csak t-ben akadunk el. Így tehát megtaláltunk egy st-sétát, amelyből az esetleges köröket kihagyva egy P_1 st-utat nyerünk. Ezt az eljárást $\delta(s)$-szer ismételve megkapjuk a keresett $\delta(s)$ élidegen utat.

Bár ez a mohó módszer csak nagyon speciális esetben használható, az általános esetre is szolgál útmutatással. Ahelyett ugyanis, hogy D-ben a k élidegen utat közvetlenül próbálnánk megtalálni, a D egy olyan D' részgráfjának megkeresésére törekszünk, amely teljesíti a fenti három tulajdonságot és amelyben $\delta_{D'}(s) = k$. A fentiek szerint ekkor a D'-ben már könnyűszerrel megtaláljuk a k élidegen utat.

A feladat tehát egy egészértékű k nagyságú folyam megkeresése a $g \equiv 1$ kapacitásfüggvényre nézve, vagy röviden egy k-fonat megkeresése. A költséges útproblémában pedig egy minimális költségű k-fonatot keresünk. $k = 1$-re konzervatív költség esetén ez éppen a már megismert legolcsóbb út meghatározásával ekvivalens.

Az irányított kínai postás probléma

Egy áramproblémára vezető érdekes feladat a következő. Egy $D = (V, A)$ erősen összefüggő digráfot kell egy megadott pontjából kiindulva úgy bejárnunk, hogy minden élén legalább egyszer végigmenjünk és a kiindulási pontba jussunk vissza. Egyik cél a végigjárt élek számának minimalizálása, vagy általánosabban, ha az éleken adott egy végighaladási idő, akkor a teljes bejárás összidejének minimalizálása.

Például egy postásnak a postáról elindulva egy körzet minden utcáján, amelyek mindegyikéről feltesszük, hogy egyirányú, legalább egyszer végig kell haladnia majd a postára visszatérnie. (A kérdést eredetileg irányítatlan gráfra fogalmazta meg Mey-Go Guan kínai matematikus 1960-ban. Ennek megoldása, és a kínai postás elnevezés J. Edmondstól származik, és jóval mélyebb eszközöket igényel, mint az irányított változat).

Egy másik alkalmazásban egy áramkör működésének helyességét kell tesztelnünk. Ehhez meg van adva, hogy az áramkör milyen állapotokban lehet. Ezek az állapotok felelnek meg a digráf csúcsainak. Ezen kívül adott még, hogy mely állapotokból mely másokba lehet közvetlenül átjutni, és valójában egy-egy ilyen átmenetnek a helyességét tudjuk mérni. A feladat az összes lehetséges állapot-átmenet ellenőrzése minimális idő alatt. Világos, hogy a digráf bejárási probléma miért modellezi ezen tesztelési feladatot.

Annak érdekében, hogy az irányított postás problémát áram feladatként megfogalmazzuk, képzeljünk el a digráf éleinek egy adott bejárását. Jelölje $z(uv)$ azt a számot, ahányszor az uv élen áthaladtunk. Rögtön látszik, hogy z egy olyan egészértékű áram, amelynek értéke minden élen legalább 1. Megfordítva, egy olyan z egészértékű áram segítségével, amely minden élen legalább 1 megadhatunk egy bejárást, amely minden e élen pontosan $z(e)$-szer halad végig. Ugyanis ha mindegyik e élt $z(e)$ darab párhuzamos példányával helyettesítjük, akkor Euler-féle digráfot kapunk és az Euler digráfok közismerten bejárhatók úgy, hogy minden élen pontosan egyszer haladunk végig. Ezen megfigyelés alapján a D digráfban az optimális bejárási probléma egy minimális költségű megengedett egészértékű áramnak a meghatározásával egyenértékű az $f \equiv 1$ és $g \equiv +\infty$ korlátozó függvényekre vonatkozóan.

1.5.3. Megengedett áramok

A postás probléma egy másik változatában az a kérdés, hogy egy erősen összefüggő digráfban mikor létezik olyan zárt séta, amely minden élt legalább egyszer használ, de, mondjuk, legfeljebb csak kétszer. Ez azzal ekvivalens, hogy mikor létezik egészértékű megengedett áram az $f \equiv 1, g \equiv 2$ korlátozó függvények esetén. Ha például a digráf három s-ből t-be vezető diszjunkt útból valamint egy t-ből s-be vezető élből áll, akkor ezen az élen bizonyosan háromszor végig kell mennünk.

Megengedett áramok létezésére ad szükséges és elegendő feltételt az alábbi, Alan Hoffmantól származó tétel.

1.5.1. Tétel (Hoffman, 1960)**.** *A $D = (V, A)$ digráfban adott $f \leq g$ kapacitásfüggvényekre vonatkozólag akkor és csak akkor létezik megengedett áram, ha*

$$\varrho_f(X) \leq \delta_g(X) \text{ minden } X \subseteq V \text{ halmazra.} \tag{1.32}$$

Továbbá, ha f és g egészértékűek és (1.32) *fennáll, úgy létezik egészértékű megengedett áram is.*

Biz. 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 (1.32) következik.

Az elegendőség igazolásához tekintsük a következő függvényt:

$$\beta(X) := \delta_g(X) - \varrho_f(X). \tag{1.33}$$

Most (1.32) azzal ekvivalens, hogy β nem-negatív. Az $X, Y \subseteq V$ halmazokra jelölje $d_x(X, Y)$ az $x(e)$ értékek összegét mindazon e élekre, melyek $X - Y$ és $Y - X$ egy-egy pontját kötik össze (mindegy melyik irányban). A bizonyítás kulcsa a következő lemma.

1.5.2. Lemma. $\beta(X) + \beta(Y) = \beta(X \cap Y) + \beta(X \cup Y) + d_{g-f}(X, Y)$.

Biz. Könnyen ellenőrizhetjük, hogy minden lehetséges él hozzájárulása a két oldalhoz ugyanannyi. •

A Hoffman tétel bizonyításához visszatérve nevezzünk egy olyan e élt **pontos**nak, amelyre $f(e) = g(e)$. Nevezzünk **pontos**nak csúcsok egy Z részhalmazát, amelyre $\beta(Z) = 0$. Tegyük fel indirekt, hogy a D digráfra nem igaz a tétel, és válasszunk egy olyan ellenpéldát (adott D), amelyben a pontos élek és a pontos halmazok együttes száma maximális. Az nem lehet, hogy minden él pontos, mert akkor $x := f(= g)$, az (1.32) feltétel miatt, megengedett áram volna, hiszen az 1.5.1 állítás nyomán tudjuk, hogy ha $\varrho_x(v) \leq \delta_x(v)$ minden csúcsra fennáll, akkor x áram. Legyen $a = st$ olyan él, amelyre $f(a) < g(a)$.

Állítjuk, hogy a belép egy pontos T halmazba. Valóban, ha nem lépne be, akkor $f(a)$-t meg tudnánk úgy növelni, hogy a módosított f' alsó korlátra továbbra is fennállna $f' \leq g$ és $\varrho_{f'}(Z) \leq \delta_g(Z)$ minden $Z \subseteq V$-re, továbbá vagy az a él válna pontossá, vagy pedig egy olyan halmaz, amelybe az a él belép. Ez a lehetőség azonban (mivel régi pontos halmaz nem szűnik meg) ellentmondana a pontos élek és halmazok maximális együttes számára tett feltevésünknek. Tehát az a él valóban belép egy T pontos halmazba. Analóg módon látható, hogy a kilép egy S pontos halmazból.

Az a él létezése folytán tudjuk, hogy a $d_{g-f}(S, T)$ érték szigorúan pozitív. A lemmát és (1.32)-t alkalmazva kapjuk, hogy $0+0 = \beta(S)+\beta(T) > \beta(S \cap T)+\beta(S \cup T) \geq 0+0$, amely ellentmondás mutatja, hogy nem létezhet ellenpélda, és így a tétel következik. Ugyanez a gondolatmenet azt is mutatja, hogy ha f és g egészértékű, akkor van egészértékű megengedett áram is. • •

Gyakran az áramnál általánosabb fogalmat tekintenek. Például az x függvényre azt írjuk elő, hogy minden v pontra $\varrho_x(v) - \delta_x(v) = m(v)$, ahol $m : V \to \mathbb{R}$ előre adott függvény. Vagy még általánosabban, adott $p : V \to \mathbb{R} \cup \{-\infty\}$, $b : V \to \mathbb{R} \cup \{\infty\}$ esetén $(p \leq b)$ minden v csúcsra legyen

$$p(v) \leq \varrho_x(v) - \delta_x(v) \leq b(v). \tag{1.34}$$

Egy egyszerű fogással azonban ez a feladat áram problémává alakítható. Nevezetesen, vegyünk fel egy új s csúcsot, és D valamennyi pontjából vezessünk egy-egy élt s-be. Egy új vs élen legyen $f(vs) := p(v), g(vs) := b(v)$. Jelölje a kibővített élek halmazát A'. Tetszőleges $x : A \to \mathbb{R}$ függvényhez legyen $x' : A' \to \mathbb{R}$ a következőképpen definiálva: $x'(e) := x(e)$ ha $e \in A$, és $x'(e) := \varrho_x(v) - \delta_x(v)$ ha $e' = vs$ $(v \in V)$. Könnyen látszik, hogy x akkor és csak akkor teljesíti (1.30)-t és (1.32)-t, ha x' megengedett áram.

1.50. Feladat. *Hoffman tételének általánosításaként igazoljuk, hogy akkor és csak akkor létezik* (1.30)*-t és* (1.34)*-t kielégítő x függvény, ha $\varrho_f(X)-\delta_g(X) \le \min\{\widetilde{p}(X), \widetilde{b}(V-X)\}$ fennáll minden $X \subseteq V$-re.*

1.51. Feladat. *Tegyük fel, hogy rendelkezésünkre áll egy algoritmus megengedett áram megkeresésére az olyan esetekre, amikor az éleken csak alsó korlát adott. Erre támaszkodva készítsünk eljárást az általános esetre, amikor alsó és felső korlátok is adottak.*

1.5.4. Áramok és folyamok kapcsolata

Alapvető az alábbi, L.R. Fordtól és D.R. Fulkersontól származó Maximális-Folyam Minimális-Vágás tétel (Max Flow Min Cut; MFMC).

1.5.3. Tétel (Maximális-Folyam Minimális-Vágás)**.** *A $D = (V, A)$ irányított gráfban akkor és csak akkor létezik a g kapacitásra vontakozó k nagyságú megengedett folyam, ha minden S $s\bar{t}$-halmazra $\delta_g(S) \ge k$. Ha e feltétel teljesül, g egészértékű és k egész, úgy a folyam is választható egészértékűnek.*

A tételbeli S halmazból kilépő élek halmazát $\delta^+(S)$-sel jelöljük és **vágás**nak nevezzük, melynek **értéke** vagy **kapacitása** a $\delta_g(S)$ szám. A tételt néha az alábbi ekvivalens alakban említik.

1.5.4. Tétel. *A megengedett st-folyamok maximális nagysága egyenlő a $\delta_g(S)$ értékek minimumával, ahol a minimum az összes $s\bar{t}$-halmazra megy. Ha g egészértékű, úgy a maximum egészértékű folyamon is felvétetik.*

Biz. Legyen x megengedett st-folyam és S $s\bar{t}$-halmaz. Ekkor (1.31) folytán $\delta_x(s) = \delta_x(S) - \varrho_x(S) \le \delta_g(S)$, amiből $max \le min$ következik.

A fordított irány igazolásához jelölje μ_g a szóbanforgó minimum értékét. (Ez a minimum nyilvánvalóan létezik, merthogy véges sok szám minimumáról van szó. Az viszont még ezen a ponton nem világos, hogy létezik maximális nagyságú folyam.) Adjunk D-hez egy új $e^* = ts$ élt és definiáljuk $f(e^*) = \mu_g$, $g(e^*) := \{\infty\}$. Az eredeti éleken legyen f mindenhol nulla.

Könnyen látszik, hogy most (1.32) fennáll, és így az 1.5.1 tétel szerint létezik megengedett áram (amely ráadásul egészértékű, ha g az). Kihagyva a hozzávett e^* élt, μ_g nagyságú folyamot kapunk. •

1.52. Feladat. *Az MFMC tételből vezessük le Menger tételének irányított élidegen változatát.* (Amely szerint egy irányított gráfban akkor és csak akkor van az s pontból a t pontba k élidegen út, ha minden $s\bar{t}$-halmazból legalább k él lép ki.)

Hoffman tétele az MFMC tételből

Nemcsak a maximális folyam feladat vezethető vissza megengedett áramokra, hanem megfordítva, a Hoffman tétel is levezethető az MFMC tételből. Ennek érdekében bevezetjük a $\Psi_f(X) := \varrho_f(X) - \delta_f(X)$ $(X \subseteq V)$ halmazfüggvényt. (Az itt következő levezetésben feltesszük, hogy f és g is véges értékű: kis gyakorló feladat az általános eset visszavezetése a véges f, g esetre.)

1.53. Gyakorlat. *Igazoljuk, hogy* $\Psi_f(X) = \sum[\Psi_f(v) : v \in X]$.

Ha Ψ_f minden csúcson nulla, akkor f megengedett áram, és készen vagyunk. Ha Ψ_f nem azonosan nulla, akkor a $V^+ = \{v : \Psi_f(v) > 0\}$ és $V^- = \{v : \Psi_f(v) < 0\}$ halmazok nem-üresek. Készítsük el a $D' = (V', A')$ digráfot úgy, hogy $V' = V \cup \{s, t\}$ és $A' = A \cup \{sv : v \in V^+\} \cup \{vt : v \in V^-\}$. Definiáljuk a g' kapacitásfüggvényt a következőképpen. Legyen $g'(sv) := \Psi_f(v)$, ha $v \in V^+$, legyen $g'(vt) := -\Psi_f(v)$, ha $v \in V^-$, és végül legyen $g'(a) := g(a) - f(a)$, ha $a \in A$. Legyen $M = \sum[\Psi_f(v) : v \in V^+]$. A következő lemma a kapcsolatot írja le egyrészt D megengedett áramai és D' megengedett st-folyamai között, másrészt D' M-nél kisebb vágásai és D (1.32)-t megsértő halmazai között.

1.5.5. Lemma. (a) *Ha x M-nagyságú g'-megengedett st-folyam D'-ben, akkor $f + x$* (az eredeti A-ra megszorítva) *megengedett áram.*

(b) *Ha $\delta_{g'}(X + s) < M$ valamely $X \subseteq V$ halmazra, akkor X megsérti a Hoffman-féle* (1.32) *feltételt.*

Biz. (a) A megengedettség, azaz $f \leq f + x \leq g$, következik a konstrukcióból. A megmaradási szabály nyilván fennáll a $V - (V^+ \cup V^-)$ elemeire. Mivel x nagysága M, minden s-ből kilépő él telített. Így V^+-nak bármely v pontjára az eredeti digráfban érvényes, hogy $x(sv) + \varrho_x(v) = \delta_x(v)$, azaz $\varrho_f(v) - \delta_f(v) + \varrho_x(v) = \delta_x(v)$, és így $\varrho_{f+x}(v) = \delta_{f+x}(v)$, vagyis a megmaradási szabály érvényes a V^+-ban fekvő pontokra is. A bizonyítás a V^--beli pontokra analóg módon végezhető el.

(b) $\Psi_f(V^+) = M > \delta_{g'}(X + s) = \delta_{g-f}(X) + \Psi_f(V^+ - X) - \Psi_f(V^- \cap X)$, amibe $\Psi_f(V^+ - X) = \Psi_f(V^+) - \Psi_f(V^+ \cap X)$-et helyettesítve kapjuk, hogy

$$0 > \delta_{g-f}(X) - \Psi_f(V^+ \cap X) - \Psi_f(V^- \cap X) =$$

$$\delta_{g-f}(X) - \Psi_f(X) = \delta_{g-f}(X) + \delta_f(X) - \varrho_f(X),$$

vagyis $\delta_g(X) < \varrho_f(X)$. •

1.54. Feladat. *Vezessük le Hoffman tételét az MFMC tételből, ha f-nek lehetnek $\{-\infty\}$, g-nek pedig $\{+\infty\}$ komponensei.*

A fenti redukció azt is mutatja, hogy nemcsak egy megengedett áramot kiszámító algoritmus használható maximális folyam keresésére, hanem egy maximális folyam illetve minimális vágás meghatározására szolgáló algoritmus segítségével is el lehet dönteni, hogy teljesül-e (1.32), és ha igen, akkor tudunk találni megengedett áramot. E visszavezetés még arra is jó, hogy a minimális költségű folyam feladatra nemsokára bemutatásra kerülő algoritmus segítségével megoldható lesz a minimális költségű megengedett áram problémája is.

1.6. Folyam algoritmusok

Következő feladatunk a maximális folyam meghatározására szolgáló javító utas algoritmus vizsgálata lesz, majd pedig a minimális költségű k nagyságú folyamok kiszámítására vonatkozó algoritmust ismertetjük.

A Ford – Fulkerson növelő utas algoritmus segítségével az MFMC tételre új bizonyítást nyerünk. Legyen x megengedett folyam. Ekkor tetszőleges S $s\bar{t}$-halmaz esetén az x folyam nagyságára érvényes az alábbi becslés. $\delta_x(s) = \delta_x(s) - \varrho_x(s) = \sum[\delta_x(v) - \varrho_x(v) : v \in S] = \delta_x(S) - \varrho_x(S) \leq \delta_g(S)$. Ebből adódik az 1.5.3 tételben a feltétel szükségessége illetve az 1.5.4 tételben max $\leq$ min egyenlőtlenség.

Az is megállapítható, hogy egy x folyam bizonyosan maximális nagyságú, amennyiben létezik egy olyan S $s\bar{t}$-halmaz, amelyre teljesülnek az alábbi **optimalitási feltételek**.

(a) $x(a) = g(a)$ minden a élre, amely kilép S-ből, és

(b) $x(a) = 0$ minden a élre, amely belép S-be.

Jelen célunk egy ilyen x folyam és S halmaz algoritmikus megkeresése. Ezt először csak egészértékű (illetve racionális) g kapacitásfüggvény esetén tesszük meg, majd tetszőleges g-re.

1.6.1. Maximális folyamok: a növelő utak módszere

A Fordtól és Fulkersontól származó algoritmus tetszőleges x megengedett st-folyamból indul ki (például $x \equiv 0$), és azt iteratívan javítja. Készítsünk el egy $D_x = (V, A_x)$ digráfot a következőképp. Egy uv él A_x-hez tartozik, ha vagy (i) $uv \in A$ és $x(uv) < g(uv)$, és ekkor ezen élét D_x-nek **előre-él**nek hívjuk, vagy (ii) $vu \in A$ és $x(vu) > 0$, és ekkor uv neve **hátra-él**. Jelölje S az s-ből D_x-ben irányított úton elérhető pontok halmazát.

1. eset *$t \notin S$, azaz t nem érhető el s-ből.* Mivel D_x semelyik éle sem lép ki S-ből, ezért D-ben minden S-ből kilépő él telített (azaz $x(uv) = g(uv)$) és minden S-be belépő uv élen $x(uv) = 0$. Vagyis az (a) és (b) optimalitási feltételek teljesülnek és az algoritmus véget ér: az adott x folyam nagysága egyenlő $\delta_g(S)$-sel.

2. eset *$t \in S$, azaz t elérhető s-ből.* Legyen P tetszőleges s-ből t-be vezető irányított út D_x-ben.

Legyen $\Delta_1 := \min\{g(uv) - x(uv) : uv$ előre-éle P-nek$\}$ és $\Delta_2 = \min\{x(vu) : uv$ hátra-éle P-nek$\}$. Legyen $\Delta = \min\{\Delta_1, \Delta_2\}$. Ekkor Δ pozitív. Nevezzük P egy élét **kritikusnak**, ha Δ ezen az élen éretik el.

Módosítsuk x-et a következőképp. Ha uv előre-éle P-nek, úgy a D uv élén növeljük $x(uv)$-t Δ-val. Ha uv hátra-éle P-nek, úgy a D vu élén csökkentsük $x(vu)$-t Δ-val. Könnyen látható, hogy a módosított x' megengedett folyam lesz, amelynek nagysága Δ-val nagyobb, mint x-é. Következésképp, ha g egészértékű, akkor a 2. eset csak véges sokszor fordulhat elő, vagyis véges sok növelés után az 1. eset következik be, amikor is az algoritmus véget ér. Tehát egész kapacitások esetén az MFMC tétel bizonyítást nyert.

Amennyiben g racionális, a nevezők legkisebb közös többszörösével a kapacitásokat végigszorozva visszajutunk az egész kapacitású esethez. •

Megjegyzendő, hogy ha g irracionális, akkor a fenti eljárás nem biztosan ér véget véges sok lépésben (amint az példával demonstrálható). Másik hátrány, hogy még egész kapacitások esetén is az iterációk száma arányos lehet az előforduló legnagyobb kapacitás nagyságával. Így az algoritmus bonyolultsága az input méretének exponenciális függvényével arányos, azaz nem polinomiális. (Egy szám nagysága a jegyei számának exponenciális függvénye.)

1.6.2. Skálázási technika

Az alábbiakban bemutatunk egy ügyes fogást, amelynek segítségével bizonyos eljárásokat polinomiális futásidejűvé lehet tenni. Használatát a maximális folyam problémán szemléltetjük, mert ott igen egyszerű, de számos alkalommal bonyolultabb körülmények között is használható.

Tételezzük fel, hogy a kapacitások egész számok és kettes számrendszerben vannak megadva. A legnagyobb kapacitás álljon M jegyből. Összesen M darab folyam problémát fogunk megoldani, mindegyikben a megelőzően megkapott maximális folyamot használjuk kiindulási folyamként. Jelölje g_i azt a kapcitás függvényt, amely úgy áll elő, hogy minden élen az eredeti kapacitásnak (balról) az első i jegyét tekintjük, míg a többit eltöröljük. Tegyük fel, hogy a g_i kapacitásfüggvényre nézve már meghatároztuk az x_i maximális folyamot. Ekkor $2x_i$ megengedett folyam a g_{i+1}-re nézve. A $2x_i$-ből kiindulva alkalmazzuk a fent leírt növelő utas módszert a g_{i+1} kapacitásfüggvényre vonatkozólag.

Miután minden e élre $g_{i+1}(e)$ értéke vagy $2g_i(e)$ vagy $2g_i(e) + 1$, legfeljebb élszámnyi növelés után megkapjuk az x_{i+1} maximális folyamot (a g_{i+1}-re nézve). Összesen tehát legfeljebb $M|A|$ növelés segítségével megkonstruáltunk egy eredeti kapacitásokra vonatkozó maximális folyamot.

1.6.3. Legrövidebb növelő utak

A fenti eljárás hátránya, hogy csak egész (és így racionális) kapacitásokra működik. Ezen nehézség leküzdésére J. Edmonds és R. Karp [1972] és E.A. Dinits [1970] javasolták, hogy minden iterációban a legkisebb élszámú növelő utat válasszuk. Ez az egyszerű megszorítás lehetővé teszi, hogy a Ford – Fulkerson algoritmus bonyolultságát $|V|$ és $|A|$ polinomjával korlátozzuk, függetlenül a kapacitások nagyságától. (Ezt persze úgy értve, hogy a számokkal végzett alapműveleteket egyetlen lépésnek tekintjük.)

1.6.1. Tétel. *Ha a Ford – Fulkerson féle növelő utas algoritmusban mindig a legrövidebb növelő utat használjuk, úgy az eljárás tetszőleges g kapacitásfüggvény esetén legfeljebb $O(|V||A|)$ növelés után véget ér.*

Biz. Jelölje $\sigma_x(v)$ a v távolságát D_x-ben s-től. (Ha egyáltalán nincs s-ből v-be út, akkor $\sigma_x(v) := \infty$). Legyen P egy legrövidebb út D_x-ben s-ből t-be. Ekkor P mindegyik uv élére, $\sigma_x(v) = \sigma_x(u) + 1$.

1.6.2. Lemma. *Amikor P mentén végrehajtunk egy növelést, a $\sigma_x(v)$ érték semmilyen v-re sem csökken.*

Biz. Nézzük meg milyen hatással van a növelés a D_x segédgráfra. Mivel a folyamot D-nek csak olyan élein változtattuk, melyek P éleinek felelnek meg, D_x csupán P éleinél változhat. Éspedig, D'_x lehetséges új élei P élei megfordítva, ugyanakkor P kritikus élei (ahol Δ felvétetik) eltűnnek D_x-ből. A v pont s-től való távolsága csak akkor csökkenhetne, ha olyan uw éleket adnánk a segédgráfhoz, melyekre $\sigma_x(w) > \sigma_x(u)+1$, amiből a lemma következik. •

A növelések sorozatát fázisokra bontjuk. Egy fázis során $\sigma_x(t)$ ugyanaz marad. A lemma szerint legfeljebb $|V|-1$ fázis lehetséges.

1.6.3. Lemma. *Egy fázison belül legfeljebb $|A|$ növelésre kerülhet sor.*

Biz. Jelölje $\sigma_i(v)$ a v pont távolságát s-től az i fázis kezdetén az aktuális segédgráfban. Nevezzünk egy uv élt i-**szorosnak**, ha $\sigma_i(v) = \sigma_i(u)+1$. Az i-dik fázis során csupán i-szoros éleket használunk. Tudjuk, hogy egy növelés legalább egy i-szoros élt eltüntet az aktuális segédgráfból és nem hoz be új i-szoros élt. Mivel a segédgráfnak legfeljebb $|A|$ darab i-szoros éle van, a lemma következik. •

Mindezeket összetéve kapjuk, hogy legfeljebb $|V||A|$ növelésre van szükségünk, így az Edmonds – Karp és Dinits féle algoritmus össz-bonyolultsága $O(|V||A|^2)$, hiszen egyetlen növelés $O(|A|)$ lépést igényel.

Miután a fenti algoritmus futása során a rendelkezésre álló legrövidebb növelő utak közül bármelyiket választhatjuk, a végül kapott maximális folyam függ ezen választásoktól. Nem így a végső S!

Feladatok

1.55. (a) *A végül kapott minimális $\delta^+(S)$ vágás független az algoritmus futásától.* (b) *Ha X és Y minimalizálja az $\delta_g(Z)$ értéket az összes $s\bar{t}$-halmazra, akkor $X\cap Y$ és $X\cup Y$ is minimalizáló $s\bar{t}$-halmazok.* (c) *A minimalizáló halmazok metszete S.*

1.56. *Adott e élre hogyan lehet eldönteni, hogy* (a) *létezik-e olyan maximális folyam, amely telíti e-t,* (b) *minden maximális folyam telíti e-t.*

1.57. *Algoritmikusan határozzuk meg az összes $\{x,y\}$ rendezett csúcspárt, amelyre létezik olyan X halmaz, hogy $s,x\in X, y\in V-X$ és X minimális vágást határoz meg.*

1.58. *Adott két kapacitásfüggvény esetén algoritmikusan döntsük el, hogy létezik-e olyan S $s\bar{t}$-halmaz, amely mindkét kapacitásfüggvényre nézve minimális vágást határoz meg.*

1.59. *Adott c_1 és c_2 kapacitásfüggvények esetén keressünk olyan c_1-re nézve minimális $s\bar{t}$-vágást, amely a c_2-re nézve a lehető legkisebb.*

1.60. *Készítsünk algoritmust, amely adott költségfüggvényre eldönti, hogy létezik-e k élidegen út s-ből t-be, melyek mindegyike minimális költségű.*

1.61. *Digráfban keressünk két diszjunkt halmazt, melyek egyike $s\bar{t}$-halmaz, másika $t\bar{s}$-halmaz úgy, hogy befok összegük minimális.*

1.6.4. Minimális költségű folyamok

Első célunkat elértük: folyamok segítségével hatékonyan lehet egy $D=(V,A)$ digráfban k élidegen st-utat keresni. A legolcsóbb út probléma általánosításaként most vizsgáljuk meg, hogy adott $c: A\to\mathbb{R}_+$ költségfüggvény esetén hogyan lehet meghatározni k élidegen st-utat, melyek összköltsége minimális. Ehhez egy minimális költségű k nagyságú megengedett egészértékű folyamot fogunk kiszámítani a $g \equiv 1$ kapacitásfüggvényre vonatkozóan. Valójában az alábbi algoritmus általános g-re is kiterjeszthető, de az egyszerűség kedvéért, és amiatt, hogy az élidegen utakhoz amúgy is csak erre a speciális esetre van szükségünk, feltesszük, hogy $g \equiv 1$. Ilyenkor egy k nagyságú egészértékű megengedett x folyam valójában $(0,1)$-értékű, azaz k-fonat.

Korábban láttuk, hogy miként lehet egy maximális M nagyságú s-ből t-be vezető folyamot polinom időben kiszámítani. Most minden 0 és M közé eső k egészre szeretnénk találni egy olyan k-fonatot melynek költsége a k-fonatok közt minimális. Egy z fonat **költségé**t a $cz = \sum[c(e)z(e) : e \in A]$ skaláris szorzattal definiáljuk. (Mivel c nem-negatív, a legolcsóbb k-fonatban, ha vannak körök, akkor ezek 0 költségűek, így kihagyhatók.)

Megjegyezzük, hogy van egy speciális költségfüggvény osztály, amelyre a feladat szinte semmitmondó. Legyen $\pi : V \to \mathbb{Z}_+$ olyan függvény a csúcshalmazon, amelyre $\pi(s) = 0 \leq \pi(v) \leq \pi(t)$ minden $v \in V$-re. Egy ilyen függvényt **potenciál**nak hívunk. Korábban (az 1.3.4 szakaszban) már definiáltuk a $\Delta_\pi : A \to \mathbb{R}$ pontindukált költségfüggvényt, amelyre tehát

$$\Delta_\pi(uv) := \pi(v) - \pi(u). \tag{1.35}$$

Ennek lehetnek negatív értékei is, de bizonyosan konzervatív, hiszen minden kör költsége nulla. Miután egy st-út Δ_π-költsége $\pi(t) - \pi(s) = \pi(t)$, egy egyirányú körnek pedig 0, kapjuk, hogy a Δ_π pontindukált költségfüggvényre vonatkozólag minden k-fonatnak ugyanaz a költsége, éspedig $k\pi(t)$. Az $uv \in A$ élekre a

$$c_\pi(uv) := c(uv) - \Delta_\pi(uv) \tag{1.36}$$

jelölést használva azt kapjuk, hogy a minimális költségű k-fonat meghatározása szempontjából c és c_π ekvivalens.

1.6.4. Tétel (Ford és Fulkerson). *A $D = (V, A)$ irányított gráf élhalmazán adott a $c : A \to \mathbb{R}_+$ költségfüggvény. A k-fonatok minimális költsége* (vagyis k élidegen st-út összköltségének a minimuma) *egyenlő a*

$$k\pi(t) + \sum[c_\pi(uv) : uv \in A, c_\pi(uv) < 0] \tag{1.37}$$

érték maximumával, ahol a maximum az összes $\pi : V \to \mathbb{R}_+$ potenciálra megy. Egy z k-fonat akkor és csak akkor minimális költségű a k-fonatok között, ha létezik olyan π potenciál, amelyre fennállnak a következő optimalitási feltételek:

$$c_\pi(uv) < 0 \Rightarrow z(uv) = 1 \;\; \text{vagy ekvivalensen} \;\; z(uv) = 0 \Rightarrow c_\pi(uv) \geq 0, \tag{i}$$

$$c_\pi(uv) > 0 \Rightarrow z(uv) = 0 \;\; \text{vagy ekvivalensen} \;\; z(uv) = 1 \Rightarrow c_\pi(uv) \leq 0, \tag{ii}$$

Amennyiben c egészértékű, úgy az optimális π is választható egészértékűnek.

Biz. Potenciálok segítségével egy z fonat cz költségére az alábbi alsó korlátot nyerhetjük.

$$\sum c(uv)z(uv) = \sum \Delta_\pi(uv)z(uv) + \sum c_\pi(uv)z(uv) =$$

$$k\pi(t) + \sum[c_\pi(uv)z(uv) : c_\pi(uv) > 0] + \sum[c_\pi(uv)z(uv) : c_\pi(uv) < 0] \geq$$
$$k\pi(t) + 0 + \sum[c_\pi(uv) : c_\pi(uv) < 0].$$

Ebből egyrészt következik a min $\geq$ max egyenlőtlenség, másrészt az, hogy egy k-fonat bizonyosan minimális költségű a k-fonatok között, ha létezik hozzá olyan π potenciál, amelyre a fenti becslésben minden egyenlőtlenség egyenlőséggel teljesül, ami viszont pont azzal ekvivalens, hogy fennállnak a tételben megadott (i) és (ii) optimalitási feltételek. Emiatt a tétel mindkét része következik, ha kimutatjuk, hogy minden lehetséges egész k értékre létezik egy k-fonat és egy ehhez tartozó π potenciál (amely egészértékű, ha c az), melyek kielégítik az optimalitási feltételeket. Ezeket konstruálja meg Ford és Fulkerson most ismertetésre kerülő minimál-költséges folyam algoritmusa, amely a maximális nagyságú folyam kiszámítására vonatkozó Ford–Fulkerson féle növelő utas eljárás finomításának tekinthető.

Az eljárás a $z \equiv 0$ 0-fonattal és a $\pi \equiv 0$ potenciállal indul. Ezután a fonat nagyságát (vagyis az élidegen st-utak számát) növeljük egyenként, illetve menetközben néha a potenciált növeljük úgy, hogy az optimalitási feltételek végig fennállnak. Az algoritmus akkor ér véget, amikor maximális nagyságú folyamot illetve egy minimális vágást kaptunk. Az algoritmus végig megőrzi az aktuális folyam egészértékűségét és amennyiben c egészértékű, úgy az aktuális potenciálét is.

ITERATÍV LÉPÉS Az általános helyzetben adott a z fonat és a π potenciál, és ezek kielégítik az (i) és (ii) feltételeket. Megkonstruálunk egy $D' = (V, A')$ segédgráfot a következőképpen. D'-nek kétféle éle van: előre és hátra. Egy $uv \in A'$ él **előre-él**, ha $uv \in A, c_\pi(uv) = 0$ és $z(uv) = 0$. Egy $uv \in A'$ él **hátra-él**, ha $vu \in A, c_\pi(vu) = 0$ és $z(vu) = 1$. Legyen S az s-ből D'-ben egyirányú úton elérhető pontok halmaza. Emlékeztetünk, hogy D'-ben nem lép ki él S-ből. Két eset lehetséges.

1. Eset *$t \notin S$, azaz t nem elérhető s-ből.*

Legyen $\varepsilon_1 = \min\{c_\pi(uv) : uv \in A, u \in S, v \in V - S, z(uv) = 0\}$ és $\varepsilon_2 = \min\{-c_\pi(uv) : uv \in A, u \in V - S, v \in S, z(uv) = 1\}$, ahol az üres halmazon vett minimumot ∞-nek definiáljuk. Legyen $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$.

1.6.1. Állítás. $\varepsilon > 0$.

Biz. Lássuk be először, hogy $\varepsilon_1 > 0$. Ha D-nek minden S-ből kilépő uv élén $z(uv) = 1$, akkor $\varepsilon_1 = \infty$. Legyen most uv a D-nek egy olyan S-ből kilépő éle, amelyre $z(uv) = 0$. Az (i) feltétel miatt $c_\pi(uv) \geq 0$, de itt nem szerepelhet egyenlőség, mert $c_\pi(uv) = 0$ esetén uv a D'-nek egy S-ből kilépő (előre) éle volna. Tehát $c_\pi(uv) > 0$ a D minden olyan S-ből kilépő élére, amire $z(uv) = 0$, így $\varepsilon_1 > 0$.

Most megmutatjuk, hogy $\varepsilon_2 > 0$. Ha D-nek minden S-be belépő uv élén $z(uv) = 0$, akkor $\varepsilon_2 = \infty$. Legyen most uv a D-nek egy olyan S-be belépő éle, amelyre $z(uv) = 1$. Az (ii) feltétel miatt $c_\pi(uv) \leq 0$, de itt nem szerepelhet egyenlőség, mert $c_\pi(uv) = 0$ esetén a fordított vu él a D'-nek egy S-ből kilépő (hátra) éle volna. Tehát $c_\pi(uv) < 0$ a D minden olyan S-be belépő élére, amire $z(uv) = 1$, így $\varepsilon_2 > 0$. •

Amennyiben $\varepsilon = \infty$, úgy az algoritmus véget ér. Ebben az esetben az S-ből kilépő D-beli élek mind telítettek (azaz mindegyik ilyen élen a z értéke 1, míg az S-be belépő D-beli élek mindegyikén a z értéke nulla. Így tehát $\delta(S) = \delta_z(S) - \varrho_z(S) = \delta_z(s)$, vagyis az aktuális z fonat maximális nagyságú és az S-ből kilépő élek halmaza minimális vágást határoz meg.

Legyen most $\varepsilon < \infty$, és módosítsuk π-t úgy, hogy minden $v \in V - S$-re növeljük $\pi(v)$-t ε-nal. Az S és az ε definíciójából rögtön kapjuk:

1.6.2. Állítás. *A módosított potenciál és a változatlanul hagyott z fonat kielégíti az optimalitási feltételeket.* •

Készítsük el az új segédgráfot és ismételjük meg az eljárást. Figyeljük meg, hogy a segédgráfban a (régi) S által feszített élek változatlanok maradnak és legalább egy S-ből kilépő új él keletkezik az ε választása folytán. (Csupán tájékoztatásul: az összes S-be lépő él megszűnik.) Emiatt az új segédgráfban az s-ből elért pontok halmaza szigorúan bővebb lesz, mint S. Ezért az 1. eset legfeljebb $(|V| - 1)$-szeri előfordulása alatt vagy után bizonyosan vagy az $\varepsilon = \infty$ következik be, vagy pedig az alábbi 2. eset.

2. Eset *$t \in S$, vagyis t elérhető s-ből.* Legyen P a D'-ben egy s-ből t-be vezető egyirányú út. Módosítsuk z-t a következőképpen. Legyen $z'(uv) = 1$, ha uv a P-nek előre-éle és legyen $z'(uv) = 0$, ha vu a P-nek hátra-éle.

A módosításból adódik:

1.6.3. Állítás. *A módosított fonat és változatlanul hagyott potenciál kielégíti az optimalitási feltételeket.* •

Az algoritmus leírását befejeztük, és ezzel a tétel bizonyítása is teljes, hiszen az eljárás véges sok lépés után minden lehetséges k-ra megad egy k-fonatot és egy potenciált, melyek teljesítik az optimalitási feltételeket. Miután összesen $M \leq |A|$ folyamnövelésre kerül sor és két folyamnövelés között legfeljebb $|V| - 1$ potenciál változtatásra, a fenti algoritmus polinomiális futásidejű. • •

Emlékezzünk vissza Duffin tételére (1.3.15 tétel), amely azt mondta ki, hogy konzervatív c esetén a legolcsóbb st-út költsége egyenlő a $\pi(t) - \pi(s)$ érték maximumával, ahol a maximum az összes megengedett π potenciálon veendő.

Ford és Fulkerson fenti 1.6.4 tétele a $k = 1$ speciális esetben azt mondja, hogy nem-negatív c esetén a legolcsóbb st-út költsége egyenlő a

$$\pi(t) + \sum[c_\pi(uv) : uv \in A, c_\pi(uv) < 0] \tag{1.38}$$

érték maximumával, ahol a maximum az összes $\pi : V \to \mathbb{R}_+$ potenciálra megy (ahol $\pi(s) = 0$).

Miként adja vissza ez a tétel Duffin tételét? Először is figyeljük meg, hogy a Duffin tételt elegendő nem-negatív c-re igazolni, hiszen létezik π megengedett potenciál és $c - \Delta_\pi$ egy ekvivalens és nem-negatív költségfüggvény. Mármost Duffin tételéhez elegendő megmutatnunk, hogy (1.38)-ben elég megengedett π-kre szorítkoznunk, amikor tehát egyáltalán nincs olyan uv él, amelyre $c_\pi(uv) < 0$. Legyen most π optimális és tegyük fel, hogy az $e = uv$ él hibás, azaz $c_\pi(uv) < 0$. Ekkor a Ford-Fulkerson tételben megfogalmazott optimalitási kritérium miatt uv minden legolcsóbb st-útban benne van. Ez azt jelenti, hogy az s-ből induló legolcsóbb utak D' digráfjából e-t kihagyva t már nincs benne az elérhető pontok S halmazában. De ekkor a π értékét $\alpha := -c_\pi(uv)$-vel csökkentve a $V - S$ valamennyi elemén a kapott π' potenciál szintén optimális és kevesebb hibás él van.

Az általános eset

Megjegyezzük, hogy a fenti tétel és algoritmus minimális változtatással átvihető az általános esetre, amikor egy általános $g : A \to \mathbb{Z}_+$ kapacitásfüggvényre vonatkozó megengedett k nagyságú folyamok közül keressük meg a minimális költségűt. Az alábbiakban vázoljuk az eltéréseket.

1.6.5. Tétel. *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbb{Z}_+$ egészértékű kapacitásfüggvény és a $c : A \to \mathbb{R}_+$ költségfüggvény. A k nagyságú egészértékű megengedett folyamok költségének minimuma egyenlő a*

$$k\pi(t) + \sum[c_\pi(uv)g(uv) : uv \in A, c_\pi(uv) < 0] \tag{1.39}$$

érték maximumával, ahol a maximum az összes π potenciálra megy. Egy k nagyságú megengedett z folyam akkor és csak akkor minimális költségű a k nagyságú megengedett folyamok között, ha létezik olyan π potenciál, amelyre fennállnak a következő optimalitási feltételek:

$$c_\pi(uv) < 0 \Rightarrow z(uv) = g(uv), \tag{i}$$

$$c_\pi(uv) > 0 \Rightarrow z(uv) = 0. \tag{ii}$$

Amennyiben c egészértékű, úgy az optimális π is választható egészértékűnek.

Biz. Potenciálok segítségével egy z folyam cz költségére az alábbi alsó korlátot nyerhetjük.

$\sum c(uv)z(uv) = \sum[\pi(v) - \pi(u)]z(uv) + \sum c_\pi(uv)z(uv) = k\pi(t) + \sum[c_\pi(uv)z(uv) : c_\pi(uv) > 0] + \sum[c_\pi(uv)z(uv) : c_\pi(uv) < 0] \geq k\pi(t) + 0 + \sum[c_\pi(uv)g(uv) : c_\pi(uv) < 0]$.

Ebből következik, hogy egy k nagyságú z folyam bizonyosan minimális költségű a k nagyságúak között, ha létezik hozzá olyan π potenciál, amelyre a fenti becslésben minden egyenlőtlenség egyenlőséggel teljesül, ami viszont pont azzal ekvivalens, hogy fennállnak a tételben megadott (i) és (ii) optimalitási feltételek.

Emiatt a tétel mindkét része következik, ha kimutatjuk, hogy minden lehetséges egész k értékre létezik egy k nagyságú egészértékű megengedett folyam és egy ehhez tartozó π potenciál (amely egészértékű, ha c az), melyek kielégítik az optimalitási feltételeket.

Az eljárás az azonosan nulla folyammal és az azonosan nulla potenciállal indul. Ezután a folyam nagyságát növeljük egyenként, illetve menetközben néha a potenciált növeljük úgy, hogy az optimalitási feltételek végig fennállnak. Az algoritmus akkor ér véget, amikor maximális nagyságú folyamot illetve egy minimális vágást kaptunk. Az algoritmus végig megőrzi az aktuális folyam egészértékűségét és amennyiben c egészértékű, úgy az aktuális potenciálét is.

Az általános helyzetben adott egy z folyam és egy π potenciál, melyek kielégítik az (i) és (ii) feltételeket. A $D' = (V, A')$ segédgráfot a következőképpen definiáljuk. Egy $uv \in A'$ él **előre-él**, ha $uv \in A, c_\pi(uv) = 0$ és $z(uv) < g(uv)$. Egy $uv \in A'$ él **hátra-él**, ha $vu \in A, c_\pi(vu) = 0$ és $z(vu) > 0$.

1. Eset *$t \notin S$, azaz t nem elérhető s-ből.*

Legyen $\varepsilon_1 = \min\{c_\pi(uv) : uv \in A, u \in S, v \in V - S, z(uv) < g(uv)\}$ és $\varepsilon_2 = \min\{-c_\pi(uv) : uv \in A, u \in V - S, v \in S, z(uv) > 0\}$, ahol az üres halmazon vett minimumot ∞-nek definiáljuk. Legyen $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Az optimalitási feltételek és az S definíciója miatt ε pozitív.

Amennyiben $\varepsilon = \infty$, úgy az algoritmus véget ér. Ebben az esetben az S-ből kilépő eredeti élek mind telítettek, míg az S-be belépő eredeti élek mindegyikén a folyam nulla. Így tehát $\delta_g(S) = \delta_z(S) - \varrho_z(S) = \delta_z(s)$, vagyis az aktuális z folyam maximális nagyságú és az S-ből kilépő élek halmaza minimális vágást határoz meg.

Legyen most $\varepsilon < \infty$, és módosítsuk π-t úgy, hogy minden $v \in V - S$-re növeljük $\pi(v)$-t ε-nal. Az S és az ε definíciójából rögtön kapjuk:

1.6.4. Állítás. *A módosított potenciál és a változatlanul hagyott z folyam kielégíti az optimalitási feltételeket.* •

Készítsük el az új segédgráfot és ismételjük meg az eljárást. Figyeljük meg, hogy a segédgráfban a (régi) S által feszített élek változatlanok maradnak és legalább egy S-ből kilépő új él keletkezik az ε választása folytán. (Csupán tájékoztatásul: az összes S-be lépő él megszűnik.) Emiatt az új segédgráfban az s-ből elért pontok halmaza szigorúan bővebb lesz, mint S. Ezért az 1. eset legfeljebb $(|V| - 1)$-szeri előfordulása alatt vagy után bizonyosan vagy az $\varepsilon = \infty$ következik be, vagypedig az alábbi 2. eset.

2. Eset *$t \in S$, vagyis t elérhető s-ből.* Legyen P a D'-ben egy s-ből t-be vezető egyirányú út. Módosítsuk z-t a következőképpen. Legyen $z'(uv) = z(uv) + 1$, ha uv a P-nek előre-éle és legyen $z'(uv) = z(uv) - 1$, vu a P-nek hátra-éle.

A módosításból adódik:

1.6.5. Állítás. *A módosított folyam és változatlanul hagyott potenciál kielégíti az optimalitási feltételeket.* •

Ezzel az algoritmus leírását be is fejeztük. Az eljárás bizonyosan véges, hiszen összesen M folyamnövelésre kerül sor és két folyamnövelés között legfeljebb $|V|$ potenciál változtásra. Ezáltal a tétel bizonyítása teljes. • •

Mi mondható az eljárás futásidejéről? Minthogy minden $1 \leq k \leq M$ értékre ki akartuk számolni a k nagyságú minimális költségű folyamot, az algoritmus lépésszámát az M értékének és a digráf $|A|$ méretének függvényében kell becsülnünk. Mivel lényegében M darab növelésre volt szükségünk (2. eset), ezért az eljárás polinomimális M és $|A|$ függvényében.

Elképzelhető persze olyan helyzet is, amikor csak egy maximális nagyságú minimális költségű folyamot kell kiszámítanunk. Ebben az esetben az algoritmus lépésszámának a digráf méretének függvényében kellene polinomiálisnak lennie, és ez most természetesen nem áll, hiszen a lépésszám az M nagyságával arányos. De még ebben az esetben is két fontos speciális eset kezelhető polinomiális időben.

Ha a legnagyobb kapacitás nem túlságosan nagy (nevezetesen, ha $|V|$ polinomjával korlátozható), akkor a fenti eljárás nyilván polinomiális, függetlenül attól, hogy a c költség-függvény egészértékű-e vagy sem.

Tegyük most fel, hogy a kapacitások tetszőlegesek, a költség-függvény egészértékű és „kicsi". Ekkor az algoritmus következő módosítása erősen polinomiális algoritmust eredményez. Tekintsük az algoritmus egy közbenső helyzetét, amikor egy π potenciál rendelkezésre áll. A szoros élek digráfjában (egy uv él **szoros**, ha $c_\pi(uv) = 0$) számítsunk ki (a Max-flow Min-cut algoritmusssal) egy maximális nagyságú z folyamot és egy S minimális ki-kapacitású $s\bar{t}$-halmazt. Ha D-ben minden S-be belépő és S-ből kilépő él szoros, akkor S minimális vágást határoz meg és az aktuális z folyam maximális nagyságú, amely kielégíti az optimalitási feltételeket. Ebben az esetben az algoritmus véget ér.

Amennyiben D-nek létezik nem-szoros éle, amely S-be be- vagy kilép, úgy ε, ahogy azt fentebb az 1. esetnél kiszámítottuk, véges lesz. Módosítsuk π-t úgy, mint az előbb, azaz minden $v \in V - S$-re növeljük $\pi(v)$-t ε-nal. Azt állítjuk, hogy a $\pi(t)$ érték a D digráf valamely (irányítatlan értelemben vett) s-ből t-be vezető útjának a költsége. Valóban, minden folyam-növeléskor a segédgráfban van olyan st út, amely szoros élekből áll. Következik, hogy a különböző $\pi(t)$ értékek száma (vagyis az 1. eset előfordulásainak száma) felülről korlátozható az élek össz-költségével. Tehát az eljárás polinomiális, ha a c egészértékű és legnagyobb értéke $|V|$ polinomjával korlátozható. Ugyanakkor van olyan példa (sorozat), amely azt mutatja, hogy ha c-nek csak az egészértékűségét tesszük fel, akkor az algoritmus nem polinomiális.

Megjegyezzük, hogy Edmonds és Karp a skálázási technikát eredetileg a legolcsóbb folyamok polinomiális idejű meghatározására dolgozta ki. Az első erősen polinomiális algoritmus Tardos Éva nevéhez fűződik.

1.62. Feladat. *Igazoljuk, hogy tetszőleges c esetén a fenti algoritmus véges sok lépésben véget ér.*

1.63. Feladat. *Hogyan lehet egy élsúlyozott páros gráfban adott k-ra maximális súlyú k élű párosítást kiszámítani?*

2. fejezet

Lineáris egyenletrendszerek

2.1. Vektortér, altér, lineáris függetlenség

Az alábbiakban áttekintjük a V Euklideszi vektortér néhány alaptulajdonságát. Ez a fejezet nem tartozik szorosan az operációkutatáshoz (és az előadáson nem is hangzik el), mert az itt közöltek a lineáris algebra részét alkotják. Célunk az, hogy azon alapfogalmakat és alaperedményeket mutassuk be, amelyekre a lineáris programmozási fejezetek a későbbiekben támaszkodni fognak.

Az alaptest mindig a valós ($\mathbb{R}$) vagy a racionális ($\mathbb{Q}$) számok teste. Amikor Euklideszi vektortérről beszélünk, a valós szám n-esek $\mathbb{R}^n$ terére (vagy a racionális szám n-esek $\mathbb{Q}^n$ terére) gondolunk. (Lineáris algebrában bebizonyítják, hogy bármely két n-dimenziós Euklideszi vektortér egymással izomorf.) Mindig az $\mathbb{R}$-t használjuk alaptestként, de valamennyi állítás érvényes a $\mathbb{Q}$ esetén is. Remélhetőleg nem okoz majd zavart, hogy jelölésben nem teszünk különbséget a 0 szám és a vektortér nulleleme (az origó) között. A vektortér elemeit néha vektoroknak, néha pontoknak tekintjük.

Legyen x és y két azonos dimenziós vektor. Azt mondjuk, hogy $x \geq y$, ha x minden komponense nagyobb vagy egyenlő y megfelelő komponensénél. Amennyiben $x \leq y$ és $x \neq y$, úgy az $x < y$ jelölést használjuk. Amikor x minden komponense szigorúan kisebb az y megfelelő komponensénél, az $x \ll y$ jelölést használjuk.

Adott $x_1, \ldots, x_k$ vektorok és $\lambda_1, \ldots, \lambda_k$ számok esetén a $b := \lambda_1 x_1 + \cdots + \lambda_k x_k$ vektort az $x_1, \ldots, x_k$ vektorok egy **lineáris kombinációjá**nak nevezzük. Ha a λ_i számok összege 1, **affin kombináció**ról beszélünk, míg ha valamennyi λ_i nem-negatív, úgy **nem-negatív kombináció**ról van szó. Egy nem-negatív, affin kombinációt **konvex kombináció**nak mondunk. Véges sok vektor lineáris kombinációinak halmazát a vektorok **lineáris burká**nak nevezzük. Véges sok pont affin (konvex) kombinációinak halmazát a pontok **affin**, (**konvex**) **burká**nak hívjuk. Például két pont konvex burka az őket összekötő szakasz, három pont konvex burka az általuk feszített háromszög.

Amennyiben a $b \neq 0$ elem előáll az $x_1, \ldots, x_k$ elemek lineáris kombinációjaként, úgy azt mondjuk, hogy b **lineárisan függ** az $x_1, \ldots, x_k$ elemektől. Ha b-nek nincs ilyen előállítása, akkor b **lineárisan független** az $x_1, \ldots, x_k$ elemektől. A lineáris kombináció **triviális**, ha mindegyik λ_i együttható 0. Ha legalább az egyikük nem-nulla, **nem-triviális lineáris kombináció**ról beszélünk. Azt mondjuk, hogy az $x_1, \ldots, x_k$ vektorok **lineárisan összefüggnek**, ha a vektortér nulleleme előáll nem-triviális lineáris kombinációjukként. Ha nincs ilyen előállítás, úgy az $x_1, \ldots, x_k$ vektorokat **line-**

árisan függetleneknek mondjuk. Az $x_1, \dots, x_k$ vektorok **kör**t alkotnak, ha lineárisan összefüggnek, de bármely valódi részük már lineárisan független. (Az elnevezés a gráfelmélet kör-fogalmából jön és nincs köze a geometria kör-fogalmához.) Könnyű ellenőrizni, hogy egy kör bármely eleme lineárisan függ a kör többi elemétől. (Ha csak annyit tesszünk fel, hogy az $x_1, \dots, x_k$ elemek lineárisan összefüggnek, akkor nem feltétlenül igaz az, hogy mindegyik x_i elem lineárisan függ a többitől. Például, az $(1,0), (2,0), (0,1)$ kétdimenziós vektorok lineárisan összefüggnek, de a (0,1) vektor nem állítható elő az $(1,0)$ és $(2,0)$ vektorok lineáris kombinációjaként.)

Egy $(a_1, a_2, \dots, a_n)$ egymás mellé leírt szám n-est sorvektornak tekintünk, míg ha ezeket az elemeket egymás alá írjuk, oszlopvektorról beszélünk. Ha a sor-vektor, akkor a^t az a transzponáltja, vagyis az a-nak megfelelő oszlopvektor.

Legyen X és Y két vektor-halmaz $\mathbb{R}^n$-ben. Ekkor a **vektor-összeg**ükön vagy **Minkowski összeg**ükön (röviden, **összegükön**) az $X + Y := \{x + y : x \in X, y \in Y\}$ halmazt értjük. A különbségük analóg módon definiálható.

A V vektortér A **altere** egy olyan nemüres részhalmaza V-nek, amelyre fennáll, hogy
1. $x \in A \Rightarrow \lambda x \in A$ minden $\lambda \in \mathbb{R}$ számra,
2. $x, y \in A \Rightarrow x + y \in A$.

Nyilván az egyetlen nulla elemből álló halmaz altér, az ún. **triviális altér**. Maga az egész V is altér. A definícióból következik, hogy a vektortér 0 eleme minden altérben benne van. Továbbá az altér véges sok elemének bármilyen lineáris kombinációja is az altérben van. Érvényes, hogy alterek metszete is altér. Könnyen ellenőrizhető, hogy két altér összege is altér, éspedig a mindkettőt magában foglaló legszűkebb altér. Egy altér **dimenziója** az altérből kiválasztható lineárisan független elemek maximális száma. Speciálisan a triviális altér dimenziója 0.

Két n-dimenziós $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ vektor **skalárszorzata** az $ab := a_1 b_1 + \dots + a_n b_n$ szám. Természetesen $ab = ba$. Azt mondjuk, hogy a és b **ortogonális**, ha skalárszorzatuk 0. Az $A, B \subseteq V$ halmazokról azt mondjuk, hogy egymásra **ortogonálisak** (vagy **merőlegesek**), ha A mindegyik eleme ortogonális B mindegyik elemére. (Ebben az értelemben tehát a háromdimenziós tér egy vízszintes és egy függőleges síkja nem ortogonális egymásra!) Könnyen látszik, hogy ha egy y vektor ortogonális az $x_1, \dots, x_k$ vektorok mindegyikére, akkor ortogonális ezek bármely lineáris kombinációjára is.

A definícióból rögtön adódik, hogy ha $x_1, \dots, x_k$ a V vektortér véges sok eleme, akkor a lineáris burkuk (vagyis a lineáris kombinációjukként előálló elemek halmaza) alteret alkot, amit az $x_1, \dots, x_k$ elemek által **generált altér**nek is nevezünk. Ez nem más, mint az $X := \{x_1, \dots, x_k\}$ halmazt tartalmazó alterek metszete, vagyis a legszűkebb X-t magában foglaló altér.

Altereket más módon is előállíthatunk. Legyen $a \in \mathbb{R}^n$ nem-nulla vektor. Az a-ra ortogonális elemek $\{x \in \mathbb{R}^n : ax = 0\}$ halmazát, másszóval az $ax = 0$ lineáris egyenlet megoldás-halmazát **origón átmenő** vagy **homogén hipersík**nak nevezzük, melynek (egyik) **normálisa** vagy **normál vektora** a. Amennyiben a az i-edik egységvektor, úgy az a-ra merőleges vektorok halmazát **koordináta-hipersík**nak hívjuk. Ez tehát mindazon vektorokból áll, amelyek i-edik koordinátája 0. Könnyen látszik, hogy egy homogén hipersík alteret alkot. Ebből adódik, hogy homogén hipersíkok metszete is altér. Másszóval, adott $x_1, \dots, x_k$ vektorok mindegyikére ortogonális vektorok halmaza alteret alkot, melyet az $x_1, \dots, x_k$ **ortogonális kiegészítő alterének,** más néven **nullterének** nevezünk. Ha az x_i vektorok mindegyike valamelyik $(0, 0, \dots, 1, \dots, 0)$ alakú egységvektor, úgy ezek ortogonális kiegészítő alterét **koordináta-altérnek** hívjuk. Ez tehát koordináta-hipersíkok metszete, vagyis mindazon vektorok halmaza, amelyeknek k előre adott komponense 0. Egy x pontnak az x_j **koordináta** (vagy **tengely**) **menti vetületét** úgy kapjuk, hogy x j-edik komponensét 0-ra állítjuk. Valójában ez a **belső** vetület, megkülönböztetendő a **külső** vetülettől, amelyet úgy kapunk, hogy az x vektor j-edik komponensét eltöröljük. (A tengelymenti belső vetítés tehát a vektorteret egy alterére képezi le, míg a külső vetítés egy másik vektortérre).

Legyenek U és V vektorterek. Azt mondjuk, hogy a $\varphi : U \to V$ leképezés **lineáris transzformáció** (vagy **leképezés**), ha

1. $x \in U, \lambda \in \mathbb{R}$ esetén $\varphi(\lambda x) = \lambda\varphi(x)$ és
2. $x, y \in U$ esetén $\varphi(x + y) = \varphi(x) + \varphi(y)$.

Könnyen ellenőrizhető, hogy az U azon elemeinek halmaza, amelyek a V nulla elemébe képződnek le, az U-nak alterét alkotják, amely alteret a φ leképezés **magterének** neveznek. Hasonlóképpen egyszerű azt belátni, hogy V azon elemei, amelyek valamely U-beli elem képeként állnak elő (azaz a $\{\varphi(u) : u \in U\}$ halmaz elemei), a V-nek alterét alkotják, amely alteret a φ leképezés **képterének** neveznek.

Egy $x = (x_1, \dots, x_n)$ szám n-esen néha az $\mathbb{R}^n$ $(x_1, \dots, x_n)$ koordinátájú pontját értjük, néha az origóból az x pontba mutató vektort. Pontok egy $\{p_1, \dots, p_q\} \subseteq R^n$ halmazáról azt mondjuk, hogy **affin független**, ha egyik pontot sem lehet néhány másik pont affin kombinációjaként előállítani. Az affin függetlenség a pontok függetlenségét akarja megragadni, a lineáris függetlenség a vektorok függetlenségét. (Két pont éppen akkor affin független, ha különbözőek, három különböző pont affin függetlensége azzal ekvivalens, hogy nincsenek egy egyenesen.) Pontok egy k elemű halmaza pontosan akkor affin független, ha az egyikből a többibe mutató $k-1$ vektor lineárisan független. Ez azzal ekvivalens, hogy az eggyel magasabb dimenziós, az új koordinátában egy 1-essel kiegészített vektorok lineárisan függetlenek.

2.2. Mátrixok, egyenletrendszerek megoldhatósága

Legyen A egy $m \times n$-es mátrix, azaz A-nak m sora és n oszlopa van. A mátrix i-edik oszlopát a_i-vel jelöljük, a j-edik sorát pedig ${}_ja$-val. Az A **sorterén** az $\mathbb{R}^n$-nek az A sorai által generált alterét értjük, amelynek jele $\mathcal{S}(A)$ vagy $\mathbb{R}^m A$. A sortér tehát az $\{y^t A : y \in \mathbb{R}^m\}$ halmaz. Az A mátrix **nulltere** az A sorainak ortogonális kiegészítő altere, melynek jele $\mathcal{N}(A)$. A nulltér tehát az $Ax = 0$ egyenletrendszer $\{x \in \mathbb{R}^n : Ax = 0\}$ megoldás-halmaza. Az A oszlopai által generált altér az A **oszloptere**, jelölésben $\mathcal{O}(A)$ vagy $A\mathbb{R}^n$, míg az oszlopainak merőleges altere a mátrix **bal nulltere**. Ennek jele $\mathcal{BN}(A)$.

Mostantól fogva azzal a jelölési egyszerűsítéssel élünk, hogy nem különböztetjük meg az oszlop- és sorvektorokat. Ennek megfelelően, ha az Az szorzatot tekintjük, akkor a z-t oszlopvektornak képzeljük, míg az yA szorzat esetén az y-t sorvektornak. Hasonlóképp, két vektor skalárszorzata esetén sem tesszük ki a transzponálási jelet, vagyis az a és b n-dimenziós vektorok skalárszorzatát ab-vel vagy ba-val jelöljük. (Ez az egyszerűsítési megállapodás zavart okozhatna, ha az a vektort $n \times 1$-es mátrixként, a b vektort pedig $1 \times n$-es mátrixként tekintenénk, mert akkor az ab mátrix szorzat egy $n \times n$-es mátrixot jelöl. Szerencsére az a, b vektorok ilyen típusú szorzatára az alábbiakban nem lesz szükségünk, így az említett zavar sem fordulhat elő.)

Valamely n-dimenziós z vektor esetén az Az vektor tekinthető úgy, mint az A oszlopainak egy lineáris kombinációja, ahol az i-edik oszlop együtthatója $z(i)$, a z i-edik komponense. Hasonlóképp, egy m-dimenziós y vektorra az yA tekinthető, mint az A sorainak egy lineáris kombinációja.

Futólag már említettük, hogy ha egy $z \in \mathbb{R}^n$ vektor ortogonális az A soraira, vagyis ha $Az = 0$, akkor z ortogonális az A sorainak bármely yA lineáris kombinációjára is, azaz $(yA)z = 0$. Valóban, $(yA)z = y(Az) = y0 = 0$.

2.2.1. Lemma. *Ha a $z \in \mathbb{R}^n$ nem-nulla vektor ortogonális az A mindegyik sorára* (azaz benne van A nullterében, vagyis $Az = 0$), *akkor z lineárisan független az A soraitól,* (azaz z nincs benne az A sorterében).

Biz. Tegyük fel indirekt, hogy z előáll az A sorainak lineáris kombinációjaként, azaz $z = yA$ valamely $y \in \mathbb{R}^m$-re. Ekkor $0 < zz = (yA)z = y(Az) = 0$, ami ellentmondás. •

Megjegyzendő, hogy a lemmában lényeges feltétel, hogy a valós vagy a racionális test felett vagyunk (legalább is annyiban, hogy rendezett test felett). A GF(2) (vagyis a kételemű) test felett például az $(1, 1)$ vektor ortogonális saját magára.

Figyeljük meg, hogy a sortér és a nulltér az $\mathbb{R}^n$ egymásra merőleges alterei. (Ugyanis, ha egy vektor merőleges a mátrix soraira, akkor merőleges a sorokból készült lineáris kombinációkra is, vagyis a sortér minden elemére). A 2.2.1 lemma alapján a sortérnek és a nulltérnek a közös része az egy szem nulla vektorból áll. Analóg módon az oszloptér és a bal nulltér az $\mathbb{R}^m$-nek egymásra merőleges alterei, melyek metszete a triviális altér.

Az A mátrix segítségével megadható egy $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ leképezés a $\varphi_A(z) := Az$ képzési szabállyal. Könnyen látszik, hogy φ_A lineáris transzformáció, amelynek képtere az A oszloptere, míg magtere az A nulltere. Hasonlóképp bevezethetünk egy ${}_A\varphi : \mathbb{R}^m \to \mathbb{R}^n$ lineáris transzformációt a ${}_A\varphi(y) := yA$ képzési szabállyal. Ennek képtere az A sortere, míg magtere az A bal nulltere.

Az a kérdés, hogy az $Az = 0$ (homogén lineáris) egyenletrendszernek van-e nem-triviális megoldása azzal ekvivalens, hogy az A oszlopai lineárisan összefüggőek-e vagy sem. Egy másik interpretáció szerint az $Az = 0$ azt jelenti, hogy z ortogonális az A soraira, vagyis az $Az = 0$ nem-triviális megoldhatóságának kérdése azzal ekvivalens, hogy az A nulltere (azaz a φ_A leképezés magtere) nem-triviális-e.

2.2.2. Tétel. *Legyen A egy $m \times n$-es mátrix, ahol $1 \leq m < n$. Ekkor az $Az = 0$ homogén lineáris egyenletrendszernek létezik nem-triviális megoldása.* (Másszóval m-nél több m-dimenziós vektor mindig lineárisan összefüggő. Még másképp, A nulltere nem-triviális.)

Biz. m szerinti indukciót használunk. Amennyiben $m = 1$, a tétel triviális. Legyen tehát $m > 1$ és tételezzük fel, hogy a tétel érvényes minden olyan mátrixra, amelynek m-nél kevesebb sora van.

A tétel állítása triviális, ha A-nak van egy csupa 0-ból álló oszlopa, így feltesszük, hogy nem ez a helyzet. Figyeljük meg, hogy ha az A egyik sorát helyettesítjük a sor λ-szorosával valamilyen $\lambda \neq 0$ számra, akkor a mátrix nulltere, vagyis az $Az = 0$ rendszer megoldásainak halmaza változatlan marad. Hasonló kijelentés érvényes, ha a mátrix egyik sorát hozzáadjuk egy másik sorához, vagy ha két sort felcserélünk. Ezen műveletek ismételt alkalmazásával egy olyan A_1 mátrix nyerhető, amelynek nulltere ugyanaz, mint az A mátrixé, és amelynek első oszlopa az $(1, 0, 0, \ldots, 0)^t$ m-dimenziós egység-vektor. Legyen A' az a mátrix, amely az A_1-ből keletkezik az első sor és az első oszlop eltörlésével. Az A'-re az indukciós feltevés szerint érvényes a tétel, azaz létezik egy $(n-1)$-dimenziós z' nem-nulla vektor, amelyre $A'z' = 0$.

Jelölje a' azt a vektort, amely az A_1 első sorából keletkezik az első (1-es) komponens eltörlésével és legyen $\alpha = -a'z'$. Ekkor a $z := (\alpha, z')$ vektor nem-nulla vektor, amelyre $A_1 z = 0$, tehát $Az = A_1 z$ miatt $Az = 0$. •

Megjegyzendő, hogy a fenti bizonyítás könnyen algoritmussá alakítható az $Az = 0$ egy nem-triviális megoldásának megkeresésére. Ez a Gauss-elimináció speciális esete homogén lineáris egyenletrendszer egy nemtriviális megoldásának megtalálására.

2.2.3. Következmény. *Ha egy mátrix oszlopai is és sorai is lineárisan függetlenek, akkor a mátrix négyzetes.*

Biz. Valóban, ha például több oszlop volna, mint sor, akkor a 2.2.2 tétel szerint az oszlopok lineárisan összefüggnének. •

2.2.4. Következmény. *Ha az A' mátrixból lineárisan függetlenül kiválasztható sorok maximális száma kisebb, mint az oszlopok száma, akkor az $A'z = 0$ rendszernek van nem-triviális megoldása.*

Biz. Válasszunk ki maximálisan sok lineárisan független sort és jelölje a mátrixukat A. A 2.2.2 tétel szerint $Az = 0$-nak van nem-triviális megoldása. Ez az eredeti $A'z = 0$-nak is megoldása, hiszen A' minden sora lineárisan függ A soraitól. •

2.2.5. Lemma. *Tegyük fel, hogy az A mátrixban az első r oszlop lineárisan független és a többi oszlop lineárisan függ ezektől. Hasonlóképp legyen az első s sor lineárisan független és a többi sor lineárisan függjön ezektől. Ekkor az első r oszlop és az első s sor által meghatározott A_1 részmátrix oszlopai is és sorai is lineárisan függetlenek.*

Biz. Szimmetria miatt elég kimutatni, hogy az A_1 oszlopai lineárisan függetlenek. Töröljük el az $a_{r+1}, \ldots, a_n$ oszlopokat. Továbbra is érvényes, hogy mindegyik sorvektor lineárisan függ az első s sortól. Így ha egy vektor ortogonális az első s sorra, akkor ortogonális a többi sorra is, vagyis ha az A_1 oszlopai lineárisan összefüggnének, akkor az A mátrix első r oszlopa is lineárisan összefüggne, ellentmondásban a feltevéssel. •

A 2.2.3 következmény és a 2.2.5 lemma kombinációjából kapjuk a következőt.

2.2.6. Tétel. *Egy A mátrix lineárisan független oszlopainak maximális száma egyenlő a lineárisan független sorok maximális számával.* •

Ezt a közös maximális számot $r(A)$-val jelöljük és a mátrix **rangjá**nak nevezzük.

Ezen meggondolásokból még egy fontos tulajdonság kiolvasható. Tegyük fel, hogy a mátrixból egymás után választunk sorokat, csak arra ügyelve, hogy a kiválasztott sorok lineárisan függetlenek legyenek. Ezt egészen addig tesszük, amíg már több sor nem választható ki, azaz amikor már a nem kiválasztott sorok mindegyike lineárisan függ a kiválasztott soroktól. Ekkor, függetlenül a közbenső választási döntéseinktől, a kiválasztott lineárisan független sorok száma mindig ugyanaz lesz, nevezetesen az A mátrix rangja. (Másszóval a mohó algoritmus mindig maximális sok lineárisan független sort fog megtalálni.) Ennek igazolásához tegyük fel, hogy az algoritmus mondjuk az első s sort választotta ki és legyen indirekt $s < r := r(A)$. Feltehetjük, hogy az A első r oszlopa lineárisan független. De ekkor ellentmondásban vagyunk a 2.2.5 lemmával. Ebből az is következik, hogy az A sorterének a dimenziója egyenlő a mátrix rangjával (ami egyenlő a mátrix oszlopterének dimenziójával.)

2.2.7. Következmény. *Egy mátrix rangja nem változik, ha hozzáveszünk egy új oszlopot, amely lineárisan függ az oszlopoktól, vagy ha elhagyunk egy meglévő oszlopot, amely lineárisan függ a többi oszloptól. Analóg állítás érvényes sorokra.* •

2.2.8. Tétel. *Ha egy $m \times n$-es A mátrix sorai lineárisan függetlenek* (azaz $r(A) = m$)*, akkor tetszőleges n-dimenziós b vektorra az $Ax = b$ egyenletrendszernek létezik megoldása* (ami $b \neq 0$ esetben azzal ekvivalens, hogy b lineárisan függ az A oszlopaitól.) *Ha $m = n$, akkor a megoldás egyértelmű.*

Biz. Ha $b = 0$, akkor $x = 0$ megoldás, így feltesszük, hogy $b \neq 0$. A 2.2.6 tétel miatt A-nak van m lineárisan független oszlopa. Tegyük fel, hogy az első m oszlop lineárisan független. Láttuk, hogy m-nél több m dimenziós vektor lineárisan összefüggő, így az $a_1, \dots, a_m, b$ vektorok lineárisan összefüggőek. Egy ilyen nem-triviális lineáris összefüggésben a b együtthatója nem lehet 0, mert ez azt jelentené, hogy az $a_1, \dots, a_m$ vektorok lineárisan összefüggőek, ellentétben a feltevéssel. Ha viszont a b együtthatója nem nulla, akkor b kifejezhető az a_i vektorok lineáris kombinációjaként.

Az $m = n$ esetben az egyértelműség bizonyításához tegyük indirekt fel, hogy létezik két megoldás is. Ekkor ezek z különbsége nem-nulla és $Az = 0$, vagyis az A oszlopai lineárisan összefüggőek, de ekkor a sorai is, ellentétben a feltevéssel. •

Egy négyzetes mátrixot **szinguláris**nak neveznek, ha sorai (illetve ezzel ekvivalensen, oszlopai) lineárisan összefüggnek.

A 2.2.8 tétel általánosításához tegyük ismét fel, hogy az A mátrixnak m sora és n oszlopa van (lehet $m = n$), de a sorok lineáris függetlenségét nem tételezzük fel.

2.2.9. Tétel. *A következők ekvivalensek.*
(A) Az $Ax = b$ egyenletrendszernek létezik megoldása.
(B) $r(A) = r([A, b])$, (ahol $[A, b]$ az a mátrix, amely A-ból áll elő a b oszlop hozzávételével).
(C) Nem létezik olyan y, amelyre $yA = 0$, $yb \neq 0$.

Biz. (B)→(A). Tekintsünk A-nak $r(A)$ lineárisan független oszlopát. $r(A) = r([A, b])$ azt jelenti, hogy b lineárisan függ ezektől, tehát függ az A oszlopaitól, vagyis (A) fennáll.

(A)→(C). Ha létezik olyan x és y, amelyekre $Ax = b$ és $yA = 0$, akkor $0 = (yA)x = y(Ax) = yb$.

(C)→(B). Tegyük fel indirekt, hogy $r(A) < r([A, b])$. Válasszunk ki $[A, b]$-nek $r(A) + 1$ lineárisan független sorát és legyen $[A_1, b_1]$ az általuk alkotott részmátrix. A_1 sorai már lineárisan összefüggnek, hiszen A-nak nincsen $r(A) + 1$ darab lineárisan független sora. De ekkor létezik olyan $y_1 \neq 0$ vektor, amelyre $y_1 A_1 = 0$. $[A_1, b_1]$ sorainak lineáris függetlensége miatt $y_1[A_1, b_1] \neq 0$, azaz $y_1 b_1 \neq 0$. Egészítsük ki y_1-t 0 komponensekkel egy y m-dimenziós vektorrá. Erre $yA = 0, yb \neq 0$, ellentmondásban (C)-vel. •

A 2.2.9 tételben az (A)-beli problémát **primál** problémának nevezzük. Ez tehát azt kérdezi, hogy a b oszlopvektor benne van-e az A oszlopainak alterében. A (C)-beli problémát **duál** (vagy **duális**) problémának hívjuk. Ez azt kérdezi, hogy létezik-e olyan y vektor, amely A valamennyi oszlopára merőleges, ugyanakkor b-re nem, másszóval, hogy létezik-e olyan homogén hipersík (melynek normálisa y), amely tartalmazza A minden oszlopát, de b-t nem. Fontossága miatt megismételjük a 2.2.9 tételből az (A) és (C) feltételek ekvivalenciáját és direkt bizonyítást adunk rá.

2.2.10. Tétel (Fredholm féle alternatíva tétel)**.** *Az $Ax = b$ rendszernek akkor és csak akkor van megoldása, ha nem létezik olyan y, amelyre $yA = 0$, $yb \neq 0$. Ekvivalensen, egy [b] vektor vagy benne van egy altérben* [melyet az A oszlopai generálnak]*, vagy elválasztható tőle egy* [y normálisú] *homogén hipersíkkal abban az értelemben, hogy a hipersík az alteret tartalmazza, de a vektort nem.*

Biz. Egyszerre nem létezhet a szóbanforgó x és y, mert akkor $0 = (yA)x = y(Ax) = yb \neq 0$.

Annak igazolására, hogy a primál és a duál probléma egyike biztosan megoldható az A sorainak m száma szerint indukciót használunk. Az $m = 1$ eset könnyű gyakorlat. Szintén egyszerűn látszik a tétel, ha A azonosan nulla.

Tegyük most fel, hogy A-nak van nemnulla eleme, hogy $m \geq 2$, és hogy kevesebb sorú márixokra a tétel igaz. Sor- és oszlopcserékkel elérhetjük, hogy $a_{11} \neq 0$. Könnyű ellenőrizni, hogy az egyenletrendszer egy sorát nemnulla számmal szorozva a primál megoldás-halmaz nem változik. A duál megoldás-halmaz ilyenkor változhat ugyan, de a duál probléma megoldhatósága nem (tehát az eredeti duál akkor és csak akkor oldható meg, ha a módosított). Ugyanez a két kijelentés érvényes, ha az egyik egyenletet hozzáadjuk egy másikhoz. Ezek alapján feltehetjük, hogy $a_{1,1} = 1$ és az első oszlop többi eleme 0. Tekintsük az első sor törlésével keletkező $A'x' = b'$ rendszert. Indukcióból kapjuk, hogy vagy ez, vagy pedig a duális $\{y'A' = 0, y'b' \neq 0\}$ rendszer megoldható. Amennyiben létezik x', úgy ennek első komponensét szabadon változtathatjuk, hiszen A' első oszlopa nulla. Emiatt ezt a komponenst olyannak tudjuk választatni, hogy az $Ax = b$ rendszerből kihagyott első egyenlet is teljesüljön. Tehát ha a redukált primál megoldható, akkor az eredeti is. Ha viszont a redukált duálisnak létezik egy $(m-1)$ dimenziós y' megoldása, akkor az $y := (0, y')$ az eredeti duálisnak megoldása. •

Megjegyzések Az alternatíva tételt a szemléletes megfogalmazás alapján hívhatjuk szeparációs tételnek is. Később látni fogjuk, hogy jóval általánosabb szeparációs tételek is léteznek. Algebrai tételek ilyen jellegű geometriai szemléltetése sokszor hozzájárul magának a tételnek a megsejtéséhez, megkönnyíti a tétel megértését, és segíti a megjegyzést is. Ugyanakkor a három-dimenziós geometriai tartalom kézenfekvősége önmagában egyáltalán nem jelenti azt, hogy az általános tétel triviális lenne, vagy hogy egyáltalán igaz! Lásd még a 3.1.1 részbeli megjegyzéseket.

2.1. Feladat. *Igazoljuk, hogy $r(Y \cdot A) \leq r(A)$!*

2.2. Feladat. *Igaz-e, hogy az $Ax = 0$ egyenletrendszernek akkor és csak akkor létezik semelyik koordinátájában sem nulla megoldása, ha A minden oszlopa lineárisan függ a többitől?*

2.3. Egyenletrendszer megoldás-halmaza, affin alterek

Miután áttekintettük a lineáris egyenletrendszerek megoldhatóságának kérdését, vizsgáljuk meg, hogy miként lehet leírni a megoldások halmazát.

2.3.1. Tétel. *Egy A $n \times n$-es nem-szinguláris négyzetes mátrix első m sora által alkotott részmátrixot jelölje A_1, míg a maradékot A_2. Tegyük fel, hogy az A_1 minden sora ortogonális A_2 minden sorára. Ekkor A_1 sortere éppen az A_2 nulltere és A_1 nulltere éppen az A_2 sortere.*

Biz. Mivel az állítás második fele az elsőből következik az A_1 és A_2 szerepének felcserélésével, csupán az első rész bizonyítására szorítkozunk. A feltevés szerint A_1 minden sora ortogonális A_2 minden sorára, így A_1 sorainak lineáris kombinációja is ortogonális az A_2 soraira, azaz A_1 sortere része A_2 nullterének. A fordított irányú tartalmazás igazolásához legyen $z \in \mathcal{N}(A_2)$, azaz $A_2 z = 0$. Mivel A nem-szinguláris, a 2.2.8 tétel alapján létezik $y = (y_1, y_2)$, amelyre $yA = z$.

Most tehát $yA = z$ ortogonális A_2 soraira, és $y_1 A_1$ is ortogonális A_2 soraira, ezért $y_2 A_2 = yA - y_1 A_1$ is ortogonális A_2 soraira. A 2.2.1 lemma alapján az A_2 sorainak egy nemnulla lineáris kombinációja nem lehet ortogonális A_2 minden sorára, így $y_2 A_2$ szükségképpen 0, azaz $z = y_1 A_1$, vagyis z benne van A_1 sorterében. •

2.3.2. Tétel. *Legyen A_1 olyan $m \times n$-es mátrix ($m < n$), amelynek sorai lineárisan függetlenek. Ekkor létezik olyan $(n-m) \times n$ méretű A_2 mátrix, amelynek sorai ortogonálisak az A_1 soraira és amely az A_1-gyel együtt egy $n \times n$-es nem-szinguláris mátrixot alkot. Az A_1 sortere az A_2 nulltere, és A_1 nulltere az A_2 sortere.*

Biz. A 2.2.2 tétel szerint van olyan $z_1 \in \mathbb{R}^n$ nemnulla vektor, amely ortogonális az A_1 soraira (magyarul az $A_1 z = 0$ rendszernek van nem-triviális megoldása.) Természetesen ekkor z_1 lineárisan független A_1 soraitól. Egészítsük ki az A_1 mátrixot a z_1 sorvektorral. Miután z_1 lineárisan független az A_1 soraitól, a megnövelt A_1' mátrix sorai is lineárisan függetlenek. Amennyiben A_1'-nek még mindig kevesebb, mint n sora van, úgy a 2.2.2 tétel alapján ismét létezik egy olyan z_2 vektort, amely ortogonális az A_1' soraira.

Ezt az eljárást $(n-m)$-szer alkalmazva olyan $z_1, \ldots, z_{n-m}$ vektorokat kapunk, melyek mindegyike ortogonális az A_1 valamennyi sorára valamint egymásra is, továbbá a végül kapott $n \times n$-es mátrix sorai lineárisan függetlenek. Ezen konstrukció alapján a z_i $(i = 1, \ldots, n-m)$ sorvektorokból álló A_2 mátrix teljesíti a tétel kívánalmait. •

2.3. Feladat. *Tegyük fel, hogy a 2.3.2 tételbeli A_1 mátrix (I_m, B) alakú, ahol I_m az $m \times m$-es egységmátrixot jelöli, míg B tetszőleges $m \times (n-m)$-es mátrix. Igazoljuk, hogy az $A_2 := [B^T, -I_{n-m}]$ kielégíti a tétel kívánságait, azaz A_2 sortere az A_1 sorterének ortogonális kiegészítője.*

A két tétel összevetéséből adódik, hogy az $A_1 x = 0$ homogén lineáris egyenletrendszer megoldás-halmaza pontosan a $z_1, z_2, \ldots, z_{n-m}$ vektorok lineáris burka. Az is következik, hogy az n-dimenziós tér tetszőleges Q m-dimenziós alterének létezik egy egyértelműen meghatározott $Q^{\perp}$ $n-m$-dimenziós ortogonális kiegészítő altere. $\mathbb{R}^n$ minden eleme egyértelműen áll elő egy Q és egy Q^*-beli elem összegeként.

2.3.3. Következmény. *Minden generált altér előáll nulltérként és minden nulltér előáll generált altérként. Egy $Ax = 0$ homogén egyenletrendszer megoldás-halmaza előáll véges sok vektor lineáris burkaként, azaz $\{yB : y \in \mathbb{R}^n\}$ alakban,* (ahol n az A és a B oszlopainak száma). *Véges sok vektor lineáris kombinációinak halmaza előáll egy homogén lineáris egyenletrendszer megoldás-halmazaként.* •

A 2.2.2 tételből következően az $\mathbb{R}^n$ térben legfeljebb csak n vektor választható ki lineárisan függetlenül. Miután az n egységvektor lineárisan független, az $\mathbb{R}^n$ dimenziója n. Az is következik, hogy tetszőleges A altér véges sok elem generált altereként áll elő: válasszunk ki maximálisan sok lineárisan független elemét A-nak (ezek száma legfeljebb n), minden elem ezektől lineárisan függ, azaz A a kiválasztott elemek által generált altér.

A 2.3.1 és 2.3.2 tételekből látjuk, hogy minden altér nemcsak generált altérként áll elő, hanem nulltérként is. (Ez annak a szemléletes geometriai ténynek az általánosítása, hogy a síkban egy origón átmenő egyenest egyrészt meg lehet adni $ax+by = 0$ alakban, másrészt $\alpha(-b, a)$ „paraméteres" alakban.) Az is következik, hogy k darab lineárisan független vektor nullterének rangja $n-k$. Speciálisan, ha A $n-1$ rangú mátrix, akkor az $Az = 0$ megoldás-halmaza egy-dimenziós altér, másnéven egy **origón átmenő egyenes**, amelynek pontjai valamely a vektorra $\{\lambda a : \lambda \in \mathbb{R}\}$ alakban adhatók meg (ahol a az A soraira merőleges nem-nulla vektor).

A 2.2.9 tétel választ ad arra a kérdésre, hogy egy egyenletrendszernek mikor van megoldása. Feltéve, hogy létezik megoldás, mi mondható a megoldás-halmazról?

Affin altéren vagy **eltolt altér**en (vagy **affin halmaz**on) egy altér eltoltját értjük. Vagyis C affin altér, ha létezik olyan A altér és a vektor, amelyekre $C = \{x : x = z + a$ valamilyen $z \in A$ vektorra $\}$, vagyis $C = A + \{a\}$. Ilyenkor könnyen látható, hogy C bármely c elemére $C - \{c\} = A$, vagyis az altér, amelynek eltolásából a C keletkezik, egyértelműen meghatározott. A C affin altér **dimenzió**ján a definiáló A altér dimenzióját értjük.

2.4. Gyakorlat. *Igazoljuk, hogy affin alterek* (a) *nemüres metszete és* (b) *összege is affin altér!*

Véges sok pont affin burka affin alteret alkot, ami nem más, mint a véges sok pontot tartalmazó legszűkebb affin altér. Két (különböző) pont affin burkát a két pont **összekötő egyenes**ének nevezzük, míg a két pont konvex burka a két pontot összekötő **szakasz**. Egy egy-dimenziós affin halmazt **egyenes**nek nevezünk.

Könnyen látszik, hogy egy C halmaz akkor és csak akkor affin altér, ha van olyan c eleme, amelyre az $\{x - c : x \in C\}$ halmaz altér. A $C \neq \emptyset$ halmaz akkor és csak akkor affin altér, ha bármely két elemének affin kombinációja C-ben van, ami azzal ekvivalens, hogy bármely véges sok elemének affin kombinációja C-ben van.

2.3.4. Tétel. *Tegyük fel, hogy az $Ax = b$ egyenletrendszernek x_0 megoldása. Ekkor a megoldások $M := \{x : Ax = b\}$ halmaza az $\mathbb{R}^n$ tér affin altere, nevezetesen az A nullterének eltoltja. Másként fogalmazva, az $Ax = 0$ homogén egyenletrendszer egy tetszőleges* *megoldását x_0-hoz adva megoldást kapunk, és M minden tagja így áll elő. Megfordítva, minden affin altér előáll egy lineáris egyenletrendszer megoldás-halmazaként.*

(Figyeljük meg, hogy a 2.2.9 tételben az $Ax = b$ megoldhatóságának az oszlopok terét magában foglaló $\mathbb{R}^m$ térben volt szemléletes jelentése, a megoldások halmazát viszont a sorteret magában foglaló $\mathbb{R}^n$ térben szemléltetjük affin altérként.)

Biz. Legyen $z \in \mathcal{N}(A)$, azaz $Az = 0$. Ekkor nyilván $A(z + x_0) = b$, vagyis $\mathcal{N}(A) + \{x_0\} \subseteq M$. Legyen most $x_1 \in M$. Ekkor $z := x_1 - x_0$-ra fennáll $Az = 0$, tehát x_1 előáll mint az x_0 és az $\mathcal{N}(A)$-beli z elem összege. Ezzel a tétel első felét igazoltuk.

Tekintsünk most egy C affin alteret, amely valamely ${}_1a, \dots, {}_ma$ vektorok által generált altér x_0 vektorral történő eltolásával áll elő, vagyis az $\{yA + x_0 : y \in \mathbb{R}^m\}$ alakú vektorok halmaza, ahol A jelöli az ${}_1a, \dots, {}_ma$ sorokból álló mátrixot. A 2.3.2 tétel szerint van olyan Z mátrix, amelynek nulltere éppen az A sortere. Legyen $b := Zx_0$. Ekkor C éppen a $Zx = b$ egyenletrendszer megoldás-halmaza. •

2.3.5. Következmény. *Amennyiben az $Ax = b$ egyenletrendszernek x_0 egy megoldása, úgy a megoldások halmaza előáll véges sok vektor lineáris burkának x_0-lal történő eltolásaként, azaz $\{yB + x_0 : y \in \mathbb{R}^n\}$ alakban.* •

A következményben megfogalmazott eredményre néha úgy hivatkoznak, hogy egy lineáris egyenletrendszer megoldás-halmaza előállítható paraméteres alakban. Ennek speciális esete az a geometriában tanult eredmény, hogy egy síkot a háromdimenziós térben meg lehet adni egy lineáris egyenletrendszer megoldás-halmazaként is és paraméteres alakban is, azaz $\alpha a + \beta b + c$ alakban, ahol α, β valós paraméterek, a, b, c pedig vektorok $\mathbb{R}^3$-ban.

2.3.6. Következmény. *Amennyiben az $Ax = b$ egyenletrendszernek van megoldása, úgy az M megoldáshalmaz dimenziója $n - r(A)$, ahol n az oszlopok száma.*

Biz. Az előbbi tétel szerint M az A nullterének az eltoltja. Álljon A_1 az A-nak $r(A)$ lineárisan független sorából. Nyilván A-nak és A_1-nek ugyanaz a nulltere. A 2.3.1 és 2.3.2 tételek alapján A_1 nullterének rangja $n - r(A_1) = n - r(A)$. •

2.5. Gyakorlat. *Tegyük fel, hogy az $Ax = b$ rendszer megoldható. Legyen A' az A maximálisan sok lineárisan független sorából alkotott részmátrix és b' a b ennek megfelelő része. Ekkor az $A'x = b'$ tetszőleges x^* megoldására $Ax^* = b$-nek.*

Tegyük fel, hogy az $Ax = b$ egyenletrendszer megoldható. Azt mondjuk, hogy az $ax = \beta$ egyenlet **logikai következménye** $Ax = b$-nek, ha ennek minden megoldása kielégíti $ax = \beta$-t (másszóval, ha az $\{x : Ax = b\}$ affin altér benne van az $\{x : ax = \beta\}$ hipersíkban). Azt mondjuk, hogy $ax = \beta$ **lineáris következménye** $Ax = b$-nek, ha előáll az $Ax = b$ egyenleteinek lineáris kombinációjaként, azaz ha létezik olyan y vektor, amelyre $yA = a$ és $yb = \beta$.

2.6. Feladat. *Igazoljuk, hogy $ax = \beta$ akkor és csak akkor lineáris következmény, ha logikai.*

Befejetésül megemlítjük a Cramer szabály néven szereplő elegáns tételt, amely bizonyos esetekben determinánsok segítségével explicit formában megadja egy egyenletrendszer megoldását.

2.3.7. Tétel (Cramer szabály)**.** *Egy nem-szinguláris négyzetes A mátrix esetén az $Ax = b$ egyenletrendszer egyértelmű megoldásának i-edik komponse a $\frac{\det(A_i)}{\det(A)}$ hányadossal egyenlő, ahol A_i azt a mátrixot jelöli, amelyik az A-ból áll elő az i-edik oszlopának b-re történő cserélésével.* •

A tétel haszna nem abban van, hogy a segítségével oldjuk meg a szóbanforgó egyenletrendszert, hiszen a determinánsok kiszámítása már az $n \geq 4$ esetben is tipikusan éppúgy a Gauss eliminációval történik, mint magának az egyenletrendszernek a direkt megoldása. Az 5. fejezetben azonban a Cramer szabály döntő szerephez jut annak feltárásában, hogy a lineáris programozás ezután tárgyalandó alaperedményei miként egységesítik és általánosítják a hálózati optimalizálásban már megismert legfontosabb tételeket, mint amilyenek Egerváry, Gallai vagy Hoffman tételei.

3. fejezet

Lineáris egyenlőtlenségrendszerek megoldása

3.1. Bevezetés

Egy olajfeldolgozó üzemben kétféle nyersolaj áll rendelkezésre: az A típusból 8 millió hordó, a B típusból 5 millió. Ezekből készítenek benzint és gázolajat. Az üzemben három technológiai eljárás közül lehet választani. Az első eljárás bemeneti-kimeneti arányait az jellemzi, hogy 3 hordó A-kőolajból és 5 hordó B-kőolajból 4 hordó benzint és 3 hordó gázolajat állít elő. A második eljárás 1 hordó A-ból és 1 hordó B-ből készít 1 hordó benzint és 1 hordó gázolajat, míg a harmadik eljárásnál ezek az értékek rendre 5, 3 és 3, 4. Tudván, hogy a benzin hordójáért 4 dollárt, a gázolaj hordójáért 3 dollárt kapunk, a meglévő nyersolaj készletet miképp osszuk fel a három eljárás között, ha célunk az össz-bevétel maximalizálása. (Egyszerűség kedvéért nem vesszük most tekintetbe az eljárások esetleg eltérő üzemi költségeit).

Jelölje x_i $(i = 1, 2, 3)$ azt, hogy az egyes eljárásokat milyen mértékben használjuk. x_1 tehát azt jelenti, hogy az első eljárással $3x_1$ A-olajat és $5x_1$ B-olajat dolgozunk fel, és ennek során $4x_1$ benzint és $3x_1$ gázolajat kapunk. Az x_i értékeknek természetesen nemnegatívnak kell lenniük. Az adatok alapján az A-olajból $3x_1 + x_2 + 5x_3$ hordót használunk, és így ez az összeg legfeljebb 8 millió. A B-olajra az $5x_1 + x_2 + 3x_3 \leq 5000000$ egyenlőtlenség adódik.

Az eljárásokkal benzinből összesen $4x_1 + x_2 + 3x_3$ hordó áll elő, melynek értéke $4(4x_1 + x_2 + 3x_3)$ dollár. Gázolajból $3x_1 + x_2 + 4x_3$ hordót nyerünk, melynek értéke $3(3x_1 + x_2 + 4x_3)$. Az összbevételünk tehát $25x_1 + 7x_2 + 24x_3$ dollár. Feladatunk maximalizálni a $25x_1 + 7x_2 + 24x_2$ célfüggvényt az $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$ valamint a $3x_1+x_2+5x_3 \leq 8000000$ és $5x_1+x_2+3x_3 \leq 5000000$ feltételek mellett. (Mivel x_i ebben a modellben a hordók számát jelöli, így ki kellene kötnünk, hogy minden x_i egész. A fenti feladatban azonban a hordók száma nagy, így gyakorlati szempontból nem számít, ha elengedjük az egészértékű megkötést. Jelezzük ugyanakkor, hogy számos gyakorlati problémában szükséges lehet a változókra tett egészértékűségi megkötés. Lineáris egyenlőtlenség-rendszerek egészértékű megoldhatóságával az *egészértékű programozás* foglalkozik.)

A lineáris algebra egyik kiinduló pontja a lineáris egyenletrendszerek vizsgálata volt. A Gauss-elimináció segítségével elvi és algoritmikus választ kaptunk arra a kérdésre, hogy egy lineáris egyenletrendszernek mikor van megoldása. A **lineáris programozás** lineáris egyenlőtlenség-rendszerekkel foglalkozik. Egy egyenlőtlenség lehet szigorú vagy egyenlőséget is megengedő, de a továbbiakban egyenlőtlenségen mindig ezen utóbbit értjük, hacsak kifejezetten az ellenkezőjét nem mondjuk. A legelső kérdés az, hogy egy egyenlőtlenség-rendszernek mikor létezik megoldása, vagy másképp fogalmazva, egy egyenlőtlenség-rendszer megoldás-halmaza, melyet majd poliédernek nevezünk, mikor nemüres. Az erre vonatkozó eredmény (Farkas lemma) az egyenletrendszerekről szóló Fredholm tétel direkt általánosítása. Hasonlóképp, az egyenletredszerek megoldás-halmazára vonatkozó eredmények szépen kiterjeszthetők egyenlőtlenség-rendszerek megoldás-halmazára.

Egyenlőtlenség-rendszerekkel kapcsolatban azonban olyan új típusú kérdések is felvetődnek, amelyeknek nincs is értelmes speciális esetük egyenletrendszerekre vonatkozólag. Megkérdezhetjük, hogy valamely c vektorra a cx lineáris célfüggvény korlátos-e az R poliéderen (mondjuk) felülről. (Egy affin altéren egy lineáris célfüggvény vagy konstans vagy nem korlátos). Ha korlátos, úgy harmadik célunk meghatározni a cx függvény maximumát (vagy ha alulról korlátos, úgy minimumát) R-en. Persze most még azt (a később majd bizonyításra kerülő tényt) sem tudjuk, hogy a szóbanforgó maximum egyáltalán létezik-e: Weierstrass általános tétele szerint egy korlátos zárt halmazon folytonos függvény felveszi maximumát, így miután cx folytonos és egy poliéder bizonyosan zárt, R korlátossága esetén már most is bizonyosak lehetünk a maximum létezésében. Nemsokára ezt is és a nem-korlátos esetet is igazoljuk, Weierstrass nélkül.

3.1.1. Megjegyzések az intuícióról

A lineáris egyenletrendszerek megoldására vonatkozó elmélet megértését megkönnyítette, hogy három dimenzióban a problémához egy geometriailag szemléletes képet lehetett kapcsolni. A geometriai intuíció segítséget jelent egyenlőtlenség-rendszerek vizsgálatánál is. Egy három-változós egyenlőtlenség-rendszer megoldás-halmazát is szépen lehet ábrázolni. Egyetlen $qx \leq \beta$ ($q \neq 0$) egyenlőtlenség megoldás-halmaza az $\mathbb{R}^3$-ban egy zárt féltérként képzelhető el. Egy egyenlőtlenség-rendszer megoldáshalmaza így néhány féltér metszete. Három dimenzióban véges sok féltér metszete nem más, mint egy konvex poliéder (megengedve, hogy a poliéder nem feltétlenül korlátos). Ez a kép természetesen sugallja, hogy magasabb dimenzióban is egy egyenlőtlenség-rendszer megoldás-halmazát majd poliédernek nevezzük. Kérdés persze, hogy mennyire szerencsés ez az elnevezés abban az értelemben, hogy egy n-változós egyenlőtlenség-rendszer megoldás-halmaza valóban rendelkezik-e olyan tulajdonságokkal, melyeket 3-dimenziós szemléletünk sugall. Három dimenzióban például világos, hogy egy korlátos poliéder a csúcsainak konvex burka. Igaz-e ez magasabb dimenzióban is? A kérdés persze így eleve csalás, hiszen még azt sem tudjuk, hogy miként is kéne magasabb dimenzióban a csúcsot definiálni.

Ami esetleg kézenfekvő a 3-dimenziós szemléletünkben, az lehet, hogy n dimenzióban nem is igaz. Vagy még ha igaz is, a legkevésbé sem azért, mert három dimenzióban jól látszik. Gondoljunk arra, hogy egy n pontú gráf szerkezete mennyivel összetettebb lehet, mint egy három pontúé. Azt nyilván senki sem hiszi, hogy egy három pontú gráfra érvényes állításnak automatikusan tetszőleges gráfra is igaznak kéne lennie. A szemléletes és az igaz állítások kapcsolatának jobb megértésére tekintsük a következő (nem feltétlenül igaz) állításokat.

1. *Ha f folytonos függvény az $I = [0, 1]$ zárt intervallumon, amelyre $f(0) < 0 < f(1)$, akkor létezik olyan $x \in I$ szám, amelyre $f(x) = 0$.*

2. *Ha f folytonos függvény az $I = (0, 1)$ nyílt intervallumon, akkor f véges sok pont kivételével I minden pontjában deriválható.*

3. *Az n-dimenziós Euklideszi térben n darab páronként hegyes szöget bezáró vektorok egy halmaza mindig beforgatható a nemnegatív térszögletbe* (azaz létezik egy olyan ortonormált lineáris transzformáció, amely az n vektor mindegyikét nemnegatív vektorba képezi).

4. *Ha $\mathbb{R}^n$-ben egy P korlátos poliéder bármely két csúcsa szomszédos, akkor P-nek legfeljebb $n + 1$ csúcsa van.*

5. *Ha $\mathbb{R}^n$-ben véges sok pont P konvex burka nem tartalmazza az origót, akkor van olyan zárt féltér, amely magában foglalja P-t, de nem tartalmazza az origót.*

Ezen állítások mindegyikét többé-kevésbé szemléletesnek érezzük. Az első közülük Bolzano tétele, amit bevezető analízisben bizonyítanak. Nem ritka az a felfogás, hogy a Bolzano tétel nyilvánvaló, hiszen egy „folytonos vonal" a $(0, -1)$ és $(1, 1)$ pontok között szükségképpen metszi az $y = 0$ tengelyt, és a részletes bizonyításra csak azért van szükség, mert „a matematikában pontosnak kell lenni". Ez a nézet azonban téves. A Bolzano tétel ugyanis nem arról szól, hogy a folytonosságra bennünk élő szemléletes érzetre igaz-e valami vagy sem, hanem arról, hogy a folytonosságra bevezetett formális definíció (matematikai modell) vajon valóban teljesíti-e azokat az elvárásokat, amelyeket a szemléletes folytonosság képünk sugall. A Bolzano tétel egy ilyen elvárt tulajdonság fennállását igazolja vissza.

A fenti 2. állítás egy másik ilyen elvárt tulajdonságot fogalmaz meg, amely szintén eléggé szemléletesnek tűnik, csak hát éppen nem igaz: van olyan folytonos függvény I-n, amely egyetlen pontban sem differenciálható.

A 3. állítás nyilvánvaló a síkban, könnyen igazolható 3 dimenzióban, és némi erőfeszítéssel bebizonyítható még $\mathbb{R}^4$-ben is. Magasabb dimenzióban azonban már az állítás nem érvényes! Kis analógia: könnyen igazolható, hogy legfeljebb négy pontú gráfok kromatikus száma egyenlő a maximális teljes részgráfjuk pontszámával. Öt pontú gráfokra azonban ez már nem áll, hiszen az öt pontú kör kromatikus száma 3, de nincs benne háromszög.

A 4. állítás 2 és 3 dimenzióban kézenfekvő. A 4-dimenziós térben azonban minden $n \geq 3$-ra lehet olyan n csúcsú poliédert konstruálni, amelynek csúcsai páronként szomszédosak.

Az 5. állítás 3-dimenzióban szintén kézenfekvő, de az előbbi példák elbizonytalaníthatnak, hogy vajon magasabb dimenzióban is igaznak kell-e lennie. Mindenesetre, ha az elkövetkezőkben esetleg tényleg az derül ki, hogy érvényes, akkor ezt a tényt a legkevésbé sem szabad majd magától értetődőnek tekintenünk.

A 3-dimenziós térben szemléletes állítások $\mathbb{R}^n$-be történő átvitelének nehézségeit némileg érzékelteti az alábbi feladat.

3.1. Feladat. *Igazoljuk, hogy $\mathbb{R}^n$-ben véges sok pont konvex burka zárt.*

A következőkben olyan fogalmakat építünk ki magasabb dimenzióban, melyek három dimenzióban jól ismertek. Mi egy poliéder lapja, csúcsa, dimenziója? Azt az utat követjük, amely általában egy definíció kiterjesztésénél szokás: kiválasztjuk az ismert esetben a szóbanforgó fogalom valamely alapvető tulajdonságát, és az általánosításhoz ezt használjuk definícióként. (Például egy pozitív a szám negatív egész kitevős hatványát úgy definiáltuk, hogy érvényben maradjon a pozitív egész kitevős hatványra fennálló egyszerű szabály. Ezért lett, definíció szerint $a^0 := 1$ vagy $a^{(-n)} := 1/a^n$). Egy dologra azonban ügyelni kell. Elképzelhető, hogy az általánosítandó fogalomnak nem csak egy alapvető tulajdonsága van, így az általánosításra is több lehetőség kínálkozik. Ilyenkor meg kell vizsgálni, hogy a különböző úton kapott definíciók vajon ekvivalensek-e egymással vagy nem.

A helyzet megvilágítására álljon itt egy gráfos példa. Egy irányítatlan gráfban az a tulajdonság, hogy a gráf bármely két pontja között vezet út azzal ekvivalens, hogy a gráfnak van feszítő fája. Az ilyen tulajdonságú gráfokat nevezik összefüggőnek. Magasabb rendű összefüggőség definíciójához mindkét tulajdonságot vehetjük alapul. Egy gráfot nevezhetünk k-szor út-összefüggőnek, ha bármely két pontja között vezet k élidegen út, és beszélhetünk k-szoros fa-összefüggőségről, ha a gráfban létezik k élidegen feszítő fa. A $k = 1$ esetben a két fogalom egybeesik, de a háromszög (mint gráf) példája mutatja, hogy a két tulajdonság $k \geq 2$ esetén már nem ekvivalens.

Tegyük most fel, hogy a 3-dimenziós (konvex) poliéder csúcsának fogalmát akarjuk magasabb dimenziós poliéderekre kiterjeszteni. Egy R korlátos 3-dimenziós poliéder csúcsa az R-nek olyan pontja, amely nincs benne a poliéder két másik pontját összekötő szakaszban. Ezen tulajdonság egy lehetőség a magasabb dimenziós poliéder csúcsának definiálására. Egy másik kézenfekvő lehetőség azt mondani, hogy az R valamely z pontja akkor csúcs, ha létezik egy sík, amelynek R-rel vett metszete az egyetlen z pontból áll. Melyiket válasszuk magasabb dimenzióban a csúcs definíciójának? Netán olyan szerencsénk lesz, hogy a kétféle lehetőség ekvivalens? Az előbbi gráfos példa mindenesetre azt mutatja, hogy még ha ki is derül majd, hogy ez a helyzet, akkor sem tekinthető ezen ekvivalencia magától értetődőnek.

3.2. Kúpok, poliéderek, politópok

Az alterek (illetve az affin halmazok) éppen azok a halmazok, melyek zártak a lineáris (ill. affin) kombináció képzésre. Pontok egy halmazát akkor nevezzük **konvex**nek, ha zárt a konvex kombináció képzésre, vagyis akárhogy véve a halmaznak véges sok elemét, ezek konvex kombinációja is a halmazhoz tartozik.

Mivel $\mathbb{R}^n$ konvex, tetszőleges C halmaz benne van egy őt tartalmazó legszűkebb konvex halmazban, nevezetesen a C-t tartalmazó konvex halmazok metszetében. Könnyen kimutatható, hogy (lásd a 3.5 gyakorlatot), hogy ez a metszet nem más, mint a C konvex burka, melyet a továbbiakban $\text{konv}(C)$-vel jelölünk.

Amennyiben $a \in \mathbb{R}^n$ nem-nulla vektor, β tetszőleges szám, az $ax = \beta$ lineáris egyenlet megoldás-halmazát **hipersík**nak (hyperplane) nevezzük. Ez az $\{x \in \mathbb{R}^n : ax = 0\}$ **homogén hipersík** eltoltja. A fentiek alapján a hipersík dimenziója $n - 1$ (innen az elnevezés). Következik, hogy az affin altér hipersíkok metszetének tekinthető. Az $\{ax \leq \beta\}$ egyenlőtlenség megoldás-halmazát, vagyis az $\{x \in \mathbb{R}^n : ax \leq \beta\}$ halmazt (zárt) **féltér**nek (closed halfspace) hívjuk, melynek $\{x : ax = b\}$ a **határoló hi-**

persíkja, és amelynek normálisa a. (Ha szigorú egyenlőtlenség van, **nyílt féltér**ről beszélünk). A $\beta = 0$ esetben azt mondjuk, hogy a féltér **homogén**.

Gyakorlatok

3.2. *Ha z a $z_1, \dots, z_k$ pontok konvex kombinációja és mindegyik z_i a $v_1, \dots, v_\ell$ pontok konvex kombinációja, akkor z a $v_1, \dots, v_\ell$ pontoknak is konvex kombinációja.*

3.3. *Egy halmaz akkor és csak akkor konvex, ha bármely két elemének bármely konvex kombinációja a halmazhoz tartozik.*

3.4. *Konvex halmazok metszete is konvex.*

3.5. *A konv(C) halmaz nem más, mint a C elemeinek felhasználásával készült konvex kombinációk halmaza.*

3.2.1. Kúpok

Vektorok egy nemüres C halmazát **kúp**nak (cone) nevezzük, ha C zárt nemnegatív számmal történő szorzásra nézve, vagyis ha C bármely elemének nem-negatív számszorosa is C-hez tartozik. Ebből adódik, hogy az origó mindig a kúpban van. A kúp **triviális**, ha egyetlen pontja van (az origó). Amennyiben a kúp még az összeadásra is zárt, **konvex kúp**ról beszélünk. Ez könnyen láthatóan valóban konvex. Miután a továbbiakban csak konvex kúpokról lesz szó, kúpon automatikusan konvex kúpot fogunk érteni. Egy altér például mindig kúp. (A kúp ezen definíciója egyrészt általánosabb annál, mint amit szokásos geometriai kúp fogalmunk diktálna, hiszen megenged olyan alakzatokat is, melyeket síkok határolnak. Például a síkban a nemnegatív síknegyed kúp. Másrészt szűkebb, mert kúp eltoltja nem kúp). Két tipikus példa kúpra:

Végesen generált kúp (röviden, **generált kúp**): Véges sok $a_1, \dots, a_n \in \mathbb{R}^m$ vektor nemnegatív lineáris kombinációinak halmaza. Jelölése : $\text{kúp}(a_1, \dots, a_n) := \{z : z = \sum_i \lambda_i a_i, \lambda_i \geq 0\}$. Amennyiben A egy olyan $m \times n$-es mátrix, melynek oszlopai az a_i vektorok, úgy az a_i vektorok kúpja $\{Ax : x \geq 0\} = A\mathbb{R}^n_+$. Az A mátrix sorvektorai $\mathbb{R}^n$-ben az $\{yA : y \geq 0\} = \mathbb{R}^m_+ A$ kúpot generálják, melyet G_A-val jelölünk.

Metszetkúp (más néven **poliéder-kúp**): Véges sok homogén féltér metszete; $R := \{x : b_1x \leq 0, \dots, b_mx \leq 0\}$, ahol $b_i \in \mathbb{R}^n$. Amennyiben B egy olyan $m \times n$-es mátrix, melynek sorai a b_i vektorok, úgy $R = \{x : Bx \leq 0\}$. A B oszlopvektorai $\mathbb{R}^m$-ben az $\{y : yB \leq 0\}$ metszetkúpot definiálják.

Megjegyzendő, hogy adott $\{p_1, p_2, \dots, p_t, a_1, \dots, a_n\}$ vektorokra a $\{z : z = \sum_j \mu_j p_j + \sum_i \lambda_i a_i, \ \lambda_i \geq 0\}$ halmaz generált kúpot alkot (vagyis itt csak bizonyos együtthatókra követelünk meg nemnegativitást), éspedig a $\{p_1, -p_1, \dots, p_t, -p_t, a_1, \dots, a_n\}$ vektorok által generált kúp. Speciálisan, a $p_1, \dots, p_t$ vektorok által generált altér is generált kúp. A generált kúp tehát a generált altér általánosítása.

Hasonlóképp, $\{q_1, \dots, q_t, b_1, \dots, b_m\}$ vektorok esetén az $\{x : q_1x = 0, \dots, q_tx = 0, b_1x \leq 0, \dots, b_mx \leq 0\}$ halmaz metszetkúp, éspedig $\{x : q_1x \leq 0, -q_1x \leq 0, \dots, q_tx \leq 0, -q_tx \leq 0, b_1x \leq 0, \dots, b_mx \leq 0\}$. Speciálisan, a $q_1, \dots, q_t$ vektorok nulltere (másnéven ortogonális kiegészítő altere) metszetkúp. A metszetkúp tehát a nulltér általánosítása.

Figyeljük meg, hogy a generált kúpnak könnyen gyárthatunk egy elemét, de egyáltalán nem könnyű eldönteni egy megadott elemről, hogy a kúpban van-e. Fordított a helyzet metszet-kúp esetén: könnyű eldönteni, hogy egy megadott elem benne van-e, de nem könnyű találni egy nem-nulla elemet.

3.6. Gyakorlat. *Két metszetkúp metszete metszetkúp. Két generált kúp vektor-összege generált kúp.*

Korábban láttuk, hogy egy generált altér mindig előáll nulltérként és megfordítva. E tétel szép általánosításaként bebizonyítjuk majd, hogy egy metszetkúp mindig előáll generált kúpként és egy generált kúp metszetkúpként. Ez az ekvivalencia nem nyilvánvaló: például egy metszetkúp zártsága rögtön látszik abból, hogy a félterek zártak és zárt halmazok metszete is zárt; ugyanakkor egy generált kúp zártságának igazolása nem ilyen kézenfekvő.

Egy q nemnulla vektor esetén a $\{\lambda q : \lambda \in \mathbb{R}_+\}$ generált kúpot **végtelen irány**nak vagy röviden **irány**nak vagy másként **sugár**nak (ray) mondjuk és $\vec{q}$-val jelöljük. A $\vec{q}_1, \dots, \vec{q}_k$ irányok egy **nemnegatív kombinációj**án a $q_1, \dots, q_k$ vektorok egy nemnagatív kombinációjához tartozó irányt értjük.

A generált kúp tekinthető véges sok irány nemnegatív kombinációi halmazának. Egy z pontból induló $\vec{q}$ irányú **félegyenes**en a $z + \vec{q} := \{x : x = z + \lambda q, \lambda \in \mathbb{R}_+\}$ halmazt értjük, ahol $q \neq 0$. Tehát az irány egy origóból kiinduló félegyenes, és a félegyenes egy eltolt irány.

Adott K kúphoz hozzárendelhetjük a $K^* := \{x : xz \leq 0$ minden $z \in K$ elemre$\}$ halmazt, és ezt a K **polárisá**nak nevezzük. Könnyen látszik, hogy K^* maga is kúp, és az is, hogy a K kúp polárisának polárisa magában foglalja K-t, azaz $K \subseteq (K^*)^*$. Itt nem szükségképpen áll egyenlőség, hiszen bármely kúp polárisa könnyen ellenőrizhetően zárt, vagyis nem zárt K esetén $K \neq (K^*)^*$. Igazolható ugyanakkor, hogy a K lezártja (vagyis a K-t tartalmazó zárt halmazok metszete) éppen $(K^*)^*$. Speciálisan, zárt K-ra $K = (K^*)^*$.

3.2.2. Poliéderek és politópok

A metszetkúp fogalmánál általánosabb a következő:

Poliéder (polyhedron, tbsz: polyhedra): Véges sok féltér metszete: $R := \{x : Qx \leq b\}$, ahol Q egy $m \times n$-es mátrix, b m-dimenziós vektor. Másszóval a poliéder egy lineáris egyenlőtlenség-rendszer megoldás-halmaza. Figyeljük meg, hogy a definícióból adódóan egy poliéder mindig konvex, hiszen ha néhány vektor kielégít egy lineáris egyenlőtlenséget, akkor konvex kombinációjuk is. A 3-dimenziós térgeometriában megszokott (konvex) poliéderek megadhatók félterek metszeteként, vagyis kielégítik a fenti definíciót, ugyanakkor ez utóbbi megenged nem korlátos poliédereket is. Például egy metszetkúp vagy egy affin altér poliéder.

A formailag általánosabb, egyenlőségeket és egyenlőtlenségeket egyaránt tartalmazó $\{Px = b_0, Qx \leq b_1\}$ rendszer megoldás-halmaza is poliéder, hiszen egy $px = \beta$ egyenlet megoldás-halmaza felfogható mint a $px \leq \beta$ és a $-px \leq -\beta$ egyenlőtlenségek közös megoldás-halmaza. Nyilván az $\{x : Qx \geq b\}$ halmaz is poliéder éppúgy, mint a Q oszlopai által definiált $\{y : yQ \leq c\}$ halmaz. Ez is jelzi, hogy egy poliéder többféle módon is megadható mátrixszal. Az $\{Ax = b, x \geq 0\}$ egyenlőtlenség-rendszerről azt mondjuk, hogy **standard** alakú, vagyis ha egy olyan egyenletrendszerről van szó, amelynek változóira nemnegativitási kikötés van. Az $\{x : Ax = b, x \geq 0\}$ poliéder **standard alak**ban van adva. Egy standard alakban adott poliéder tehát egy affin altér és a nem-negatív térszöglet metszete.

Az $R := \{x : Px = b_0, Qx \leq b_1\}$ poliéder egy z elemére nézve a definiáló $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ mátrix egy sorát valamint a sor által meghatározott egyenlőtlenséget z-**aktív**nak vagy röviden csak aktívnak nevezzük, ha z egyenlőséggel teljesíti. A P sorai automatikusan aktívak. A z-re nézve aktív sorok részmátrixát az M z-**aktív részmátrixá**nak nevezzük és $M_z^=$-vel jelöljük. A z által szigorú egyenlőtlenséggel teljesülő sorok mátrixát $Q_z^<$ jelöli.

Politóp (polytope): Véges sok pont konvex burka. (Az üres halmazt is politópnak tekintjük, mint nulla darab pont konvex burka.) Azt mondjuk, hogy a politópot a szóbanforgó pontok generálják. Ezek szerint egyetlen pont is politópot alkot. Két pont által generált politóp a két pontot összekötő szakasz.

3.7. Gyakorlat. *Két poliéder metszete is poliéder. Két politóp vektor-összege is politóp.*

3.8. Gyakorlat. *Amennyiben $R := \{x : Mx \leq b\}$ nemüres, úgy R akkor és csak akkor az egész tér, ha $r(M) = 0$* (azaz M a csupa-nulla mátrix) *és $b \geq 0$.*

Természetesen vetődnek fel a következő kérdések. Mikor létezik egy egyenlőtlenségrendszernek megoldása, azaz mikor nem-üres egy poliéder? Erre válaszol majd a Farkas lemma, amely a Fredholm-féle alternatíva tételnek lineáris egyenlőtlenségekre vonatkozó kiterjesztése. Hogyan lehet „paraméteresen" megadni egy egyenlőtlenség-rendszer megoldás-halmazát, annak mintájára, ahogyan egy egyenletrendszer megoldás-halmazát meg lehetett adni így? A korlátos esetben erre válaszol majd az a bizonyításra kerülő tétel, miszerint minden korlátos poliéder politóp, és megfordítva. További kérdés, hogy két egyenlőtlenség-rendszer megoldás-halmaza mikor ugyanaz, magyarán, mikor definiálják ugyanazt a poliédert? Kezdjük egy egyszerű megfigyeléssel.

3.2.1. Lemma. *Ha az R poliéder kúp, akkor metszetkúp.*

Biz. Az R megadható $\{x : Qx \leq b\}$ alakban és feltehetjük, hogy Q-nak a lehető legkevesebb sora van. Azt igazoljuk, hogy ekkor $b = 0$. Mindenesetre $b \geq 0$, hiszen $0 \in R$ miatt $0 = Q0 \leq b$. Tegyük fel indirekt, hogy $b(i) > 0$ valamelyik i-re. A Q minimalitása miatt van olyan x' vektor, amely a $Qx \leq b$ rendszerből egyedül a ${}_iqx \leq b(i)$ egyenlőtlenséget sérti meg, azaz ${}_iqx' > b(i) > 0$. Ekkor az $\alpha := b(i)/{}_iqx'$ számra $x^* := \alpha x'$ benne van R-ben, de $2x^*$ például nincs, mert ${}_iq(2x^*) = 2\alpha_iqx' = 2b(i) > b(i)$, ellentmondásban R kúp voltával. •

Egy poliédert akkor nevezünk **korlátos**nak, ha létezik olyan K pozitív szám, amelyre $|x(i)| \leq K$ a poliéder minden x pontjának mindegyik $x(i)$ komponensére. Egy poliéder (**külső**) **dimenzió**ján, (röviden dimenzióján) az őt tartalmazó legszűkebb affin altér dimenzióját értjük. A poliéder **belső dimenziója** a benne fekvő affin alterek dimenziójának a maximuma. Például, ha a poliéder egyetlen pontból áll, akkor külső és belső dimenziója is nulla. Általában egy affin altérnek, mint poliédernek a külső és belső dimenziója megegyezik az affin altér korábban már bevezetett dimenziójával, speciálisan $\mathbb{R}^n$ egy hipersíkjának külső és belső dimenziója is $n-1$. A síkban a nemnegatív síknegyed belső dimenziója 0, külső dimenziója 2. Az n-dimenziós térben egy féltér belső dimenziója $n-1$, külső dimenziója n.

Egy q vektorról azt mondjuk, hogy a $z \in R$ elem **mozgásvektor**a, ha létezik kicsiny $\lambda > 0$ szám, amelyre mind $z + \lambda q$, mind $z - \lambda q$ R-ben van. A 0-vektor mindig ilyen, míg ha $q \neq 0$, nem-triviális mozgásvektorról beszélünk. Azt mondjuk, hogy $z \in R$ a poliéder **relatív belső pont**ja, ha létezik nem-triviális mozgásvektora. Ha nem létezik, akkor z **extrém**. Másszóval z akkor extrém, ha nincs rajta a poliéder két másik pontját összekötő szakaszon. Nem nehéz igazolni, hogy z pontosan akkor extrém, ha nem áll elő a poliéder más pontjainak konvex kombinációjaként.

Ha minden vektor mozgásvektor, akkor z **belső** pontja R-nek. A definícióból közvetlenül kiolvasható, hogy egy $z \in R$ elem mozgásvektorai alteret alkotnak, melynek neve a z **mozgástere**. Az 1-dimenziós térben egy (zárt) szakasznak végpontjaitól különböző pontjai belső pontok. Ugyanakkor a 2-dimenziós térben egy szakasz olyan poliéder, amelynek nincs belső pontja.

Egy $\{z + \lambda q : \lambda \in \mathbb{R}\}$ alakú egyenest q irányú egyenesnek nevezünk, ahol $q \neq 0$. Legyen $R = \{x : Qx \leq b\}$ nem-üres poliéder. Egy q vektorról azt mondjuk, hogy R **eltolási** vektora, ha R minden z pontjára és minden λ számra $z + \lambda q \in R$. Másszóval, az R bármely pontján átmenő q irányú egyenes R-ben van. (Néha használják a karakterisztikus vektor elnevezést, de ez nem túl szerencsés, mert ez a név már foglalt egy halmaz karakterisztikus vektorára). Rögtön látszik, hogy az eltolási vektorok alteret alkotnak, a poliéder **eltolási alterét**.

Ha egy poliéder nem tartalmaz teljes egyenest (félegyenest) akkor azt mondjuk, hogy **egyenes-mentes** (**félegyenes-mentes**). A $\vec{q}$ irányt az R **poliéder egy irányá**nak nevezzük, ha $z + \lambda q \in R$ fennáll az R minden z elemére és minden pozitív λ-ra.

3.9. Gyakorlat. *Az R poliéder irányainak nemnegatív kombinációi is az R irányai, azaz a poliéder irányai kúpot alkotnak.*

A poliéder irányainak kúpját a poliéder **irány-** (néha **recessziós**) **kúpjá**nak nevezzük. Az R poliéder egy iránya **extrém**, ha nem állítható elő tőle különböző R-beli irányok nemnegatív kombinációjaként. Egy altérnek például nincs extrém iránya.

3.10. Gyakorlat. *A poliéder egy iránya akkor és csak akkor extrém, ha nem állítható elő két tőle különböző R-beli irány nemnegatív kombinációjaként.*

Egy 3-dimenziós poliédert lapok, élek illetve csúcsok határolnak. Ezeket a fogalmakat szeretnénk magasabb dimenzióra kiterjeszteni. Egy $R \subseteq \mathbb{R}^n$ nemüres poliéder F **oldala** (face) R-nek egy

$$F := \{x \in R : cx = \delta\} \tag{3.1}$$

alakú nemüres részhalmaza, ahol $\delta := \max\{cx : x \in R\}$ valamely cx lineáris célfüggvényre, melyre a maximum létezik. A $c \neq 0$ esetben a $H = \{x : cx = \delta\}$ hipersíkot a poliéder egy **támasz-síkjá**nak nevezzük. A $c \equiv 0$ célfüggvényre a definíció azt adja, hogy R maga is oldal. **Valódi oldal**on (proper face) olyan oldalt értünk, amely nem az egész poliéder. A poliéder valódi oldala tehát az optimum helyek halmaza valamely nemnulla lineáris célfüggvényre nézve, másként szólva a poliédernek az a része, amely egy hipersíkkal érintkezik, amikor azt kívülről a poliéderhez toljuk. Amennyiben az oldal egyetlen pontból áll, úgy ezt a pontot a poliéder **csúcsá**nak nevezzük. Tehát egy $z \in R$ pont akkor csúcs, ha létezik olyan c vektor, amelyre a $cz > cx$ minden $x \in R - z$-re.

A definícióból látszik, hogy egy poliéder oldala maga is poliéder. Egy affin altér például olyan poliéder, amelynek nincs valódi oldala. Poliéder **minimális oldalá**n egy tartalmazásra nézve minimális oldalt értünk. Egy tartalmazásra nézve maximális valódi oldalt **lap**nak (facet) nevezünk.

A poliédert **csúcsos**nak (pointed) mondjuk, ha van csúcsa. Nem minden poliédernek van csúcsa, például az affin altereknek bizonyosan nincs. A 3.10 gyakorlat alapján a poliéder egy z eleme akkor extrém pont, ha nem áll elő a poliéder néhány más pontjának konvex kombinációjaként.

3.11. Gyakorlat. *Igazoljuk, hogy az R poliéder egy z pontjára a következők ekvivalensek.* (i) *z extrém,* (ii) *z nincs R-hez tartozó szakasz belsejében,* (iii) *nincs olyan $x' \neq 0$ vektor, amelyre $z + x'$ és $z - x'$ is R-ben van.*

3.3. Csúcsok és bázis-megoldások

3.3.1. Bázis-megoldások

Legyen $Q \neq 0$ egy $m \times n$-es mátrix és jelölje R a

$$Qx \leq b \tag{3.2}$$

egyenlőtlenség-rendszer megoldásainak halmazát. Ebben a részben végig feltesszük, hogy az R poliéder nem üres. Célunk megvizsgálni az R poliéder és az azt definiáló $Qx \leq b$ leíró rendszer kapcsolatát.

3.3.1. Tétel. *Valamely $q \neq 0$ vektorra a következők ekvivalensek:*
(1) $Qq = 0$.
(2) *q eltolási vektora R-nek.*
(3) *R-nek van olyan z pontja, amelyre $z + \lambda q$ minden valós λ-ra R-ben van.*

Biz. Az (1)→(2) irány nyilvánvaló, hiszen bármely $z \in R$ esetén $Q(z + \lambda q) = Qz + \lambda Qq = Qz \leq b$, azaz $z + \lambda q \in R$. A (2)→(3) irány semmitmondó. Végül, ha (3) fennáll, akkor szükségképpen $Qq = 0$, mert különben kellően nagy λ-ra a $Q(z+\lambda q) \leq b$ és $Q(z - \lambda q) \leq b$ egyenlőtlenség-rendszerek közül az egyik biztosan nem teljesülne. •

3.3.2. Következmény. *Az $R := \{x \in \mathbb{R}^n : Qx \leq b\}$ poliéder eltolási altere a Q mátrix nulltere.* •

3.3.3. Tétel. *Az $R = \{x \in \mathbb{R}^n : Qx \leq b\}$ poliéder egy z elemének mozgástere a $Q_z^=$ mátrix nulltere. Más szóval a q vektor akkor és csak akkor mozgásvektora z-nek, ha $Q_z^= q = 0$.*

Biz. Ha $Q_z^= q = 0$, akkor minden λ-ra $Q_z^=(z \pm \lambda q) = b_z^=$ és ezért kellően kicsiny pozitív λ-ra $Q(z \pm \lambda q) \leq b$, vagyis q a z mozgásvektora.

Ha $Q_z^= q \neq 0$, akkor $Q_z^=$-nek van olyan ${}_iq$ sora, amelyre ${}_iqq \neq 0$. Bármilyen pozitív λ-ra, ha ${}_iqq > 0$, akkor $Q(z + \lambda q) \not\leq b$, míg ha ${}_iqq < 0$, akkor $Q(z - \lambda q) \not\leq b$. Tehát q nem mozgásvektora z-nek. •

3.3.4. Tétel. *Tegyük fel, hogy az R poliéder nemüres és $R = \{x : Qx \leq b\} = \{x : Q'x \leq b'\}$. Ekkor Q és Q' sortere megegyezik. Tetszőleges $z \in R$ esetén $Q_z^=$ és $Q_z'^=$ sortere megegyezik.*

Biz. A 3.3.1 tétel szerint egy q vektorra akkor és csak akkor $Qq = 0$, ha $Q'q = 0$, vagyis Q és Q' nulltere megegyezik, így sorterük is.

A második részhez a 3.3.3 tételt használjuk. Ez alapján $Q_z^= q = 0$ pontosan akkor teljesül, ha a q vektor mozgásvektora z-nek az R poliéderben, ami pedig pontosan akkor, ha $Q_z'^= q = 0$. Így $Q_z^=$ és $Q_z'^=$ nulltere megegyezik, tehát a sorterük is. •

3.3.5. Tétel. *Tegyük fel, hogy az R poliéder nemüres és $R = \{x : Qx \leq b\} = \{x : Q'x \leq b'\}$. A Q valamely j oszlopa pontosan akkor lineárisan független, ha Q' megfelelő j oszlopa lineárisan független.*

Biz. Feltehetjük, hogy az első j oszlopról van szó. Azt látjuk be, hogy Q első j oszlopa akkor és csak akkor lineárisan összefüggő, ha Q' első j oszlopa az. Szimmetria miatt elég az egyik irányt belátni, így tegyük fel, hogy Q' első j oszlopa lineárisan összefügg. Ekkor létezik egy olyan $q' \neq 0$ vektor, amelynek csak az első j komponense lehet nemnulla és $Q'q' = 0$. A 3.3.4 tétel miatt Q és Q' sortere megegyezik, így valamely x-re $Qx = 0$ pontosan akkor áll fenn, ha $Q'x = 0$, amiből következik, hogy $Qq' = 0$, vagyis Q első j oszlopa is lineárisan összefüggő. •

Egy $z \in R$ elem **szint**jén a $\sigma(z) := r(Q) - r(Q_z^=)$ számot értjük. A $Qx \leq b$ lineáris rendszer egy z megoldását (azaz az $z \in R$ elemet) **bázis-megoldás**nak nevezünk, ha a z-aktív $Q_z^=$ részmátrix rangja $r(Q)$, más szóval a 0 szintű elemek a bázis-megoldások. Ha egy bázis-megoldás ráadásul olyan, hogy z nem-nulla komponenseinek megfelelő Q-oszlopok lineárisan függetlenek, akkor **erős bázis-megoldás**ról beszélünk. Speciálisan, ha Q oszlopai lineárisan függetlenek, akkor minden bázis-megoldás erős.

A 3.3.4 tételből rögtön kapjuk az alábbit.

3.3.6. Következmény. *Az R poliéder egy z elemének szintje csak a poliédertől függ és nem a poliédert meghatározó egyenlőtlenség-rendszer konkrét alakjától. Speciálisan, a bázis-megoldás fogalma is csak a poliédertől függ.* •

Érdemes kiolvasni, hogy más alakú egyenlőtlenség-rendszerek esetén mit is jelent a bázis-megoldás fogalma.

3.3.7. Tétel. (i) *Egy $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ nem-nulla mátrix esetén a*

$$Px = b_0, Qx \leq b_1 \tag{3.3}$$

lineáris rendszernek egy z megoldása akkor bázis-megoldás, ha $r(M) = r(M_z^=)$.

(ii) *Az $\{Ax = b, x \geq 0\}$ egy z megoldása akkor és csak akkor bázis-megoldás, ha a pozitív elemekhez tartozó A-beli oszlopok lineárisan függetlenek.*

(iii) *Az $\{yA \geq 0, yb = -1\}$ rendszer egy y_0 megoldása akkor és csak akkor bázis-megoldás, ha az A-ból lineárisan függetlenül kiválasztható, az y_0-ra merőleges oszlopok maximális száma $r(A, b) - 1$.*

Biz. (i) (3.3) és $\{Px \le b_0, -Px \le -b_0, Qx \le b_1\}$ megoldáshalmaza ugyanaz, továbbá $r(M) = r(M')$ és $r(M_z^=) = r(M_z'^=)$, ahol $M' := \begin{pmatrix} -P \\ M \end{pmatrix}$.

(ii) Esetleges oszlopcserével feltehetjük, hogy z-nek az utolsó j komponense pozitív. Jelölje az ezen j oszlophoz tartozó $m \times j$-es részmátrixot A'. Legyen $M := \begin{pmatrix} A \\ -I \end{pmatrix}$, ahol I az $n \times n$-es egységmátrixot jelöli. Az (i) rész szerint z akkor bázis-megoldás, ha $r(M_z^=) = r(M) = n$. Ekkor $M_z^=$ az M mátrix első $m + (n - j)$ sora (vagyis az A sorai valamint a $-I$ első $n - j$ sora). Ennek a bal alsó $(n - j) \times (n - j)$-es részmátrixa egy negatív egységmátrix, így $M_z^=$ rangja pontosan akkor n, ha az első $n - j$ oszlopának és utolsó $n - j$ sorának kitörlésével keletkező A' részmátrix rangja $n - (n - j) = j$, ami épp azt jelenti, hogy A' oszlopai lineárisan függetlenek.

(iii) Jelölje A_0 az A azon a_i oszlopaiból álló részmátrixot, melyekre y_0 merőleges, azaz $a_i y_0 = 0$. Definíció szerint y_0 akkor bázis-megoldás, ha $r(A_0, b) = r(A, b)$. A tétel állítása pedig azzal ekvivalens, hogy y_0 pontosan akkor bázis-megoldás, ha $r(A_0) = r(A, b) - 1$. Azt kell tehát csak belátnunk, hogy $r(A_0) = r(A_0, b) - 1$. De ez rögtön látszik, hiszen $y_0 A_0 = 0$ és $y_0 b = -1$ miatt a b vektor nem függ lineárisan A_0 oszlopaitól. •

Megjegyzés A szakirodalomban általában az $\{Ax = b, x \ge 0\}$ rendszerre vezetik be a bázis-megoldás fogalmát; egy z megoldást akkor *definiálva* bázis-megoldásnak, ha a pozitív komponenseihez tartozó A-oszlopok lineárisan függetlenek. Mi egy általánosabb megközelítést használtunk és ez a tulajdonság tételként adódott! Ebben az esetben ráadásul minden bázis-megoldás erős.

A 3.3.5 tételből rögtön kapjuk az alábbiakat.

3.3.8. Következmény. *Az erős bázis-megoldás fogalma csak a poliédertől függ és nem a poliédert meghatározó egyenlőtlenség-rendszertől.* •

3.3.9. Következmény. *A $\{Px = b_0, Qx \le b_1\}$ egyenlőtlenség-rendszer egy z bázis-megoldása pontosan akkor erős, ha a z nem-nulla komponenseihez tartozó M-beli oszlopok lineárisan függetlenek, ahol $M = \begin{pmatrix} P \\ Q \end{pmatrix}$.* •

3.3.10. Tétel. (A) *Minden megoldható lineáris egyenlőtlenség rendszernek létezik bázis-megoldása, nevezetesen bármely minimális szintű z elem bázis-megoldás.* (B) *Létezik erős bázis-megoldás is, nevezetesen egy maximálisan sok 0 komponenst tartalmazó x^* bázis-megoldás erős.*

Biz. (A) Belátjuk, hogy $\sigma(z) = 0$, azaz z bázis-megoldás. Ha indirekt $r(Q) > r(Q_z^=)$, úgy a Fredholm tétel szerint létezik q vektor, amelyre $Q_z^= q = 0$ és $Q_z^< q \ne 0$. A q esetleges negálásával elérhetjük, hogy a $Q_z^< q$ vektornak van szigorúan pozitív komponense. Ekkor van olyan $\lambda > 0$ érték, amelyre $z' = z + \lambda q \in R$ és ${}_i q z' = b(i)$ a $Q_z^<$ valamely ${}_i q$ sorára. $Q_z^= q = 0$ és ${}_i q q \ne 0$ miatt ${}_i q$ lineárisan független $Q_z^=$ soraitól. Így $Q_z^= z' = b_z$ miatt $r(Q_{z'}^=) > r(Q_z^=)$, ellentmondásban z választásával.

(B) Tegyük fel indirekt, hogy az x^* nemnulla komponenseinek megfelelő Q-beli oszlopok lineárisan összefüggőek. Ez azt jelenti, hogy létezik egy olyan $q \ne 0$ vektor, amelyre $Qq = 0$ (vagyis q eltolási vektor) és $x^*(i) = 0$ esetén $q(i) = 0$. Ekkor alkalmas

λ-ra $x^*_\lambda := x^* + \lambda q$-nak több nulla komponense lesz, mint x^*-nak, továbbá x^*_λ is bázis-megoldás, ellentmondva x^* választásának. •

3.3.11. Tétel. *A $Qx \leq b$ egyenlőtlenség-rendszer egy z megoldása akkor és csak akkor erős bázis-megoldás, ha létezik Q-nek egy olyan $r(Q)$ sorból és $r(Q)$ oszlopból álló nem-szinguláris Q' részmátrixa, amelyre z a $Q'x' = b'$ egyértelmű x' megoldásából áll elő 0-komponensek hozzávételével* (ahol b' a b azon részét jelöli, amely a Q' sorainak felel meg.)

Biz. Ha z a megadott módon áll elő, úgy $Q^=_z$ tartalmazza Q'-t, így rangja $r(Q)$. Továbbá a z nemnulla komponenseinek megfelelő Q-beli oszlopvektorok lineárisan függetlenek, hiszen ezek mindegyike Q' egy oszlopának kibővítése, márpedig Q' a feltevés szerint nem-szinguláris, így oszlopvektorai lineárisan függetlenek. Vagyis ilyenkor z valóban erős bázis-megoldás.

Megfordítva, legyen z erős bázis-megoldás. Ekkor $r(Q^=_z) = r(Q)$. Válasszuk ki $Q^=_z$-nek $r(Q)$ darab lineárisan független sorát, majd a z nem-nulla komponenseinek megfelelő lineárisan független oszlopokat tetszés szerint egészítsük ki a Q oszlopai közül $r(Q)$ darab lineárisan független oszloppá. Az így kapott $r(Q)$ sor és $r(Q)$ oszlop által meghatározott Q' részmátrix a 2.2.5 lemma miatt nem-szinguláris, és éppen z-t definiálja a kívánt módon. •

3.3.12. Következmény. *Tetszőleges egyenlőtlenség-rendszernek legfeljebb csak véges sok erős bázis-megoldása van.* •

Jegyezzük meg, hogy egy egyenletrendszernek minden megoldása bázis-megoldás, vagyis bázis-megoldásból kényelmesen lehet végtelen sok.

Gyakorlatok

3.12. *Igazoljuk, hogy a 3.3.7 tétel* (i) *részében z szintje $r(M) - r(M^=_z)$.*

3.13. *Igazoljuk, hogy a $Px_0 + Ax_1 = b, x_1 \geq 0$ rendszer egy megoldása akkor és csak akkor bázis-megoldás, ha az x_1 nem-nulla elemeihez tartozó P-beli oszlopokat az A-ból kiválasztott maximálisan sok lineárisan független oszloppal kiegészítve még mindig lineárisan független rendszert kapunk.*

3.14. *Igazoljuk, hogy a $\{Bx \leq b, x \geq 0\}$ rendszer egy z megoldása akkor és csak akkor bázis-megoldás, ha a B valamely B' nemszinguláris négyzetes részmátrixára z a $B'x' = b'$ egyértelmű megoldásából áll elő nullák hozzávételével.*

3.15. *Legyen $f \ll g$, (azaz f minden komponensében kisebb, mint g), ahol $f, g \in \mathbb{R}^n$. Igazoljuk, hogy az $Ax = b, f \leq x \leq g$ rendszer egy z megoldása pontosan akkor bázis-megoldás, ha az A azon a_i oszlopai lineárisan függetlenek, melyekre $f(i) < z(i) < g(i)$. Mik az erős bázis-megoldások? Mik a bázis- és az erős bázis-megoldások, ha $f \ll g$ helyett csak a gyengébb $f \leq g$ egyenlőtlenséget tesszük fel?*

Feladatok

3.16. *Az R poliéder egy z pontját tartalmazó legbővebb, R-ben fekvő affin altér az A karakterisztikus altér z-vel való eltoltja.*

3.17. *Egy* $R := \{x : Qx \leq b\} \subseteq \mathbb{R}^n$ *poliéder belső dimenziója* $n - r(Q)$. •

3.18. *Mutassunk olyan* $Qx \leq b$ *alakú egyenlőtlenség-rendszert, ahol egy erős bázis-megoldás előáll más erős bázis-megoldások konvex kombinációjaként. Bizonyítsuk be, hogy ha* Q *oszlopai lineárisan függetlenek, akkor ilyen példa nem létezik.*

3.3.2. Csúcsos poliéderek

Nézzük meg, hogy mi a kapcsolat csúcs és extrém pont között, és hogy ezek definíciója miként tükröződik a poliéder mátrixszal történő megadásában.

3.3.13. Tétel. *Az* $R = \{x : Qx \leq b\}$ *poliéder egy* z *elemére a következők ekvivalensek:*
(1) Q *oszlopai lineárisan függetlenek és* z *bázis-megoldás* (azaz $Q_z^=$ oszlopai lineárisan függetlenek, vagyis Q-nak van n lineárisan független z-aktív sora).
(2) z *csúcs.*
(3) z *extrém pont.*

Biz. (1)⇒(2) Legyen c a $Q_z^=$ sorainak az összege, azaz $c = y_1 Q$, ahol y_1 azt a $(0, 1)$-es vektort jelöli, amelyben a $Q_z^=$ sorainak megfelelő komponensek értéke 1, a többié 0. Tetszőleges $x \in R$ esetén $cx = (y_1 Q)x = y_1(Qx) \leq y_1 b = y_1(Qz) = (y_1 Q)z = cz$. Ha itt valamely $x \in R$ elemre egyenlőség szerepel, akkor $Q_z^= x = b_z^=$, ennek pedig z az egyértelmű megoldása, hiszen a feltevés szerint $Q_z^=$ oszlopai lineárisan függetlenek. (Itt a $b_z^=$ vektor a b vektor azon komponenseiből áll, melyek a Q mátrix z-aktív sorainak felelnek meg.)

(2)⇒(3) Ha z csúcs, akkor létezik egy olyan c vektor, amelyre $cz > cx$ minden $x \in R - z$ elemre. Ha z, indirekt, nem extrém, akkor létezik $x, y \in R - z$, melyekre $z = (x+y)/2$. De ekkor $cx < cz$ és $cy < cz$ és így $cz = (cx + cy)/2 < (cz + cz)/2 = cz$, ellentmondás.

(3)⇒(1) Tegyük fel, hogy z extrém. Amennyiben $Q_z^=$ oszlopai, indirekt, lineárisan összefüggőek, úgy létezik egy q nemnulla vektor, amelyre $Q_z^= q = 0$. De ekkor kicsiny pozitív ε-ra $z + \varepsilon q$ is és $z - \varepsilon q$ is benne van R-ben (merthogy kielégítik $\{Qx \leq b\}$-t), ellentmondásban a feltevéssel, hogy z extrém. •

3.3.14. Következmény. *Egy poliédernek legfeljebb véges sok csúcsa van.*

Biz. A 3.3.13 tételben az (1) tulajdonság miatt minden z csúcshoz létezik Q-nak n lineárisan független z-aktív sora, mely sorhalmaz különböző csúcsra különböző. Így R-nek legfeljebb $m!/(n!(m-n)!)$ csúcsa lehet. •

A következő eredmény jellemzi a csúcsos poliédereket.

3.3.15. Tétel. *Egy* $R = \{x : Qx \leq b\}$ *nemüres poliéderre a következők ekvivalensek:*
(1) Q *oszlopai lineárisan függetlenek.*
(2) R *egyenes-mentes.*
(3) *Az* R *eltolási altere triviális.*
(4) R *csúcsos.*

Biz. Az első három feltétel ekvivalenciája közvetlenül adódik a 3.3.1 tételből.

(4)⇒(1) Ha z csúcs, akkor a 3.3.13 tétel nyomán $Q_z^=$ oszlopai lineárisan függetlenek, így persze Q oszlopai is azok.

(1)⇒(4) A 3.3.10 tétel miatt van bázis-megoldás, és a 3.3.13 tétel miatt bármely z bázis-megoldás csúcs. •

3.3.16. Tétel. *Minden $R = \{x : Qx \leq b\}$ nemüres poliéder előáll, mint egy A altér és egy R' csúcsos poliéder összege. Nevezetesen, A az R eltolási altere* (azaz Q nulltere), *míg $R' = R \cap A^\perp$, ahol $A^\perp$ az A altér ortogonális kiegészítője* (vagyis Q sortere).

Biz. Először belátjuk, hogy $R \subseteq A + R'$, azaz bármely $z \in R$ elem előáll egy A-beli és egy R'-beli elem összegeként. Valóban, minden z elem egyértelműen előáll egy A-beli z_1 és egy $A^\perp$-beli z_2 elem összegeként. Belátjuk, hogy $z_2 \in R'$. Ha nem ez volna a helyzet, akkor $z \in A^\perp$ miatt z_2 nem volna R-ben, azaz z_2 megsértené $Qx \leq b$ valamelyik sorát. De akkor $Qz_1 = 0$ miatt $z = z_1 + z_2$ is megsértené ugyanazt a sort, ellentétben a $z \in R$ feltevéssel. Így valóban $R \subseteq A + R'$. Másrészt a definíciókból világos, hogy $A + R' \subseteq A + R \subseteq R$, amiből $A + R' = R$.

Végül belátjuk, hogy R' egyenes-mentes. Az A altér egy bázisából, mint sorvektorokból készítsük el a Q^* mátrixot. Ekkor tehát Q sorai és Q^* sorai egymásra merőlegesek, együtt kifeszítik az egész teret, azaz $\begin{pmatrix} Q \\ Q^* \end{pmatrix}$ teljes oszlop-rangú. Miután R' a $\{Q^*x = 0, Qx \leq b\}$ rendszer megoldás-halmaza, a 3.3.15 tételből adódik, hogy R' egyenes-mentes. •

3.19. Feladat. *Egy $R := \{x : Mx \leq 0\}$ metszetkúp az $A := \{x : Mx = 0\}$ eltolási altér* (ami speciális metszetkúp) *és az $R' := R \cap A^\perp$ csúcsos metszetkúp vektor-összege.* •

3.3.3. Korlátos poliéderek

Miután megtudtuk, hogy egy poliéder mikor nem tartalmaz egyenest, nézzük meg, hogy mikor nem tartalmaz félegyenest. Azt mondtuk, hogy a $\vec{q}$ irány a poliéder iránya, ha R minden z elemére a $\{z + \lambda q : \lambda \geq 0\}$ félegyenes R-ben van.

3.3.17. Tétel. *Valamely nemnulla q vektorra a következők ekvivalensek:*
(1) $Qq \leq 0$.
(2) *$\vec{q}$ a poliéder iránya.*
(3) *R-nek van olyan z pontja, amelyre a $\{z + \lambda q : \lambda \geq 0\}$ félegyenes R-ben van.*

Biz. (2)→(3) semmitmondó. A (3)→(1) és (1)→(2) irányok közvetlenül látszanak. •

3.3.18. Következmény. *Az $R := \{x \in \mathbb{R}^n : Qx \leq b\}$ poliéder iránykúpja a Q mátrix $M_Q = \{x : Qx \leq 0\}$ metszetkúpja.* •

3.20. Gyakorlat. *Igazoljuk, hogy egy $\{x : Px = b_0, Qx \leq b_1\}$ alakban adott nemüres poliéder iránykúpja $\{x : Px = 0, Qx \leq 0\}$.*

3.21. Feladat. *Egy poliédernek és iránykúpjának extrém irányai ugyanazok.* •

3.3.19. Tétel. *Egy $R = \{x : Qx \leq b\}$ nemüres poliéderre a következők ekvivalensek:*
(1) *R nem tartalmaz félegyenest.*
(2) *R-nek véges sok csúcsa van, melyek konvex burka R.*
(3) *R korlátos.*
(4) *R iránykúpja triviális.*

Biz. (1)⇒(2) Mivel R nem tartalmaz félegyenest, így egyenest még kevésbé, és ezért a 3.3.15 tétel miatt van csúcsa. A 3.3.14 következmény miatt véges sok csúcsa van. Jelölje R_K a csúcsok konvex burkát. Belátjuk, hogy $R = R_K$. Ha bizonyos vektorok kielégítenek egy egyenlőtlenség-rendszert, akkor bármely konvex kombinációjuk is kielégíti, ezért $R_K \subseteq R$.

A fordított irányú tartalmazás igazolásához indirekt tegyük fel, hogy a poliédernek van olyan z pontja, amely nem áll elő csúcsok konvex kombinációjaként. Válasszuk z-t olyannak, hogy $Q_z^=$, a z-aktív részmátrix maximális legyen. Mivel z nem csúcs, így $Q_z^=$ oszlopai lineárisan összefüggenek. Ezért létezik egy nemnulla q vektor, amelyre $Q_z^= q = 0$. Kicsiny pozitív λ-ra $z + \lambda q \in R$ és mivel R nem tartalmaz félegyenest, nagy λ értékre $z + \lambda q \notin R$. Ez azt jelenti, hogy $Q_z^<$-nek van olyan ${}_iq$ sora, amelyre ${}_iqq > 0$. Így ha λ-t nullától kezdve folyamatosan növeljük, lesz egy olyan λ_1 érték, amelyre $z_1 := z + \lambda_1 q$ benne van R-ben és aktív részmátrixa szigorúan bővebb $Q_z^=$-nél. (Nevezetesen $\lambda_1 := \min\{b_1(i) - {}_iqz)/({}_iqq)\}$, ahol a minimum a $Q_z^=$ azon ${}_iq$ soraira megy, amelyekre ${}_iqq > 0$.) Analóg módon létezik egy $z_2 := z - \lambda_2 q$ vektor R-ben ($\lambda_2 > 0$), amelynek aktív részmátrixa szigorúan bővebb $Q_z^=$-nél. A z-re tett feltevés miatt mind z_1, mind z_2 benne van R_K-ban, és ezért a $z_1 z_2$ szakasz belsejében fekvő z is, ellentmondás.

(2)⇒(3) Triviális.

(3)⇒(4) Ha indirekt létezne az iránykúpnak q nemnulla eleme, akkor bármely $z \in R$ elemre a $\{z + \lambda q : \lambda \geq 0\}$ félegyenes R-ben volna, és így R nem lenne korlátos.

(4)⇒(1) Ha indirekt valamely q nemnulla vektorra a $\{z + \lambda q : \lambda \geq 0\}$ félegyenes R-ben volna, akkor szükségképpen $Qq \leq 0$, azaz q benne volna az iránykúpban. •

A (2) tulajdonság bizonyítása mögött rejlő geometriai szemlélet a következő: nincs mit bizonyítani, ha z maga csúcs. Ha nem az, úgy tekintsük a z pont R_z oldalát, amin az R legszűkebb olyan oldalát értjük, amely tartalmazza z-t. (Ez annak felel meg, hogy a z-aktív egyenlőtlenségeket egyenlőségnek vesszük.) Keresünk egy irányt, amely mentén z-ben elmozdulva R_z-ben maradunk, és megnézzük, hogy az ilyen irányú z-n átmenő egyenes mely x_1 és x_2 pontoknál lép ki a poliéderből. Mivel az x_1 oldala és x_2 oldala is szűkebb z oldalánál, így indukcióval ők már előállnak csúcsok konvex kombinációjaként, de akkor az $[x_1, x_2]$ szakasz pontjai is előállnak, speciálisan z is.

3.4. A Fourier-Motzkin elimináció és következményei

A Gauss elimináció egyrészt hatékony algoritmust szolgáltatott lineáris egyenletrendszerek megoldására, másrészt fontos bizonyítási eszköznek bizonyult (például a Fredholm féle alternatíva tételnél.)

A Gauss-elimináció mintájára egy kézenfekvő eljárást adunk lineáris egyenlőtlenségrendszerek megoldására. A módszer eredetileg Fouriertól származik, amelyet később Motzkin elemzett, így az irodalomban Fourier-Motzkin (röviden FM) eliminációként hivatkozzák. A Gauss eliminációhoz hasonlóan az FM eljárás is véges algoritmust szolgáltat, de ez, szemben a Gauss eliminációval, szórványos kivételektől eltekintve nem hatékony a gyakorlatban. Valójában a Gauss elimináció polinomiális futásidejű algoritmus, míg az FM elimináció exponenciális. Ugyanakkor az FM eljárás is hatékony bizonyítási eszköznek bizonyul, melynek segítségével néhány alaperedmény könnyen kiadódik.

3.4.1. Oszlop elimináció

Legyen A olyan $m \cdot n$-es mátrix ($m \geq 1, n \geq 2$), amelynek első, a_1-gyel jelölt oszlopa $0, \pm 1$ értékű. Jelölje rendre I, J, K a sorok azon indexhalmazait, melyekre az $a_1(i)$ értéke $+1, -1$ illetve 0. Készítsük el az $A^{[1]}$ mátrixot a következőképpen. A K-nak megfelelő sorok változatlanul kerüljenek be $A^{[1]}$-be. Ezen kívül minden $i \in I$, $j \in J$ választásra legyen ${}_ia + {}_ja$ az $A^{[1]}$ egy sora, melyet jelöljünk ${}_{[ij]}a$-val. Ez azt jelenti, hogy ha I vagy J üres, akkor $A^{[1]}$ egyszerűen az A mátrix K-nak megfelelő részmátrixa. Általában $A^{[1]}$-nek $m-(|I|+|J|)+|I||J|$ sora van. Figyeljük meg, hogy $A^{[1]}$ első oszlopa csupa nullából áll. Hasonlóképp, tegyük fel, hogy a Q mátrix első oszlopa $0, \pm 1$ értékű, és rendeljük a

$$Qx' \leq b \tag{3.4}$$

egyenlőtlenség-rendszerhez a

$$Q^{[1]}x' \leq b^{[1]} \tag{3.5}$$

rendszert, ahol $b^{[1]}$ az $A := (Q, b)$ mátrixhoz tartozó $A^{[1]}$ mátrix utolsó oszlopa.

3.4.1. Tétel. (A) *Az*

$$Ax \leq 0 \tag{3.6}$$

egyenlőtlenség-rendszernek bármely megoldása az

$$A^{[1]}x \leq 0 \tag{3.7}$$

rendszernek is megoldása, és a (3.7) bármely megoldásának első komponensét alkalmasan megváltoztatva a (3.6) egy megoldását kapjuk.

(B) *A (3.4) bármely megoldása a (3.5) rendszernek is megoldása, és a (3.5) bármely megoldásának első komponensét alkalmasan megváltoztatva a (3.4) egy megoldását kapjuk.*

Biz. (A) Az első rész közvetlenül adódik az $A^{[1]}$ konstrukciójából, hiszen az $A^{[1]}$ minden sora az A sorainak nemnegatív lineáris kombinációja. A második részhez, legyen z megoldása (3.7)-nek. Mivel $A^{[1]}$ első oszlopa 0, feltehető, hogy $z(1) = 0$. A $z(1)$ értékét fogjuk alkalmasan megváltoztatni, hogy (3.6) egy megoldását nyerjük.

Amennyiben J üres, vagyis A első oszlopában nincsen negatív elem, úgy z első komponensét kellően kicsiny α értékre változtatva (3.6) egy megoldását kapjuk. (Nevezetesen, $\alpha := \min_{i \in I}\{-{}_iaz\}$ megteszi.) Analóg módon, ha I üres, úgy $z(1)$-t kellően nagyra változtatva kapunk (3.6)-nek egy megoldását. Tegyük most fel, hogy sem I, sem J nem üres. Állítjuk, hogy

$$\max_{j \in J}\{{}_jaz\} \leq \min_{i \in I}\{-{}_iaz\}. \tag{3.8}$$

Valóban, ha volna olyan $i \in I, j \in J$ index-pár, amelyre ${}_jaz > -{}_iaz$, úgy ${}_{[ij]}az = {}_iaz + {}_jaz > 0$ volna, ellentmondásban a feltevéssel, hogy z megoldása (3.7)-nek. Mármost, ha α tetszőleges olyan szám, amelyre

$$\max_{j \in J}\{{}_jaz\} \leq \alpha \leq \min_{i \in I}\{-{}_iaz\}, \tag{3.9}$$

úgy z első komponensét α-ra változtatva a kapott z_α-ról állítjuk, hogy megoldása (3.6)-nek. Valóban, ha $h \in K$, akkor ${}_ha$ első komponense 0, így a ${}_haz_\alpha = {}_haz \leq 0$. Ha $h \in I$, azaz ${}_ha(1) = 1$, akkor (3.8) második egyenlőtlensége folytán ${}_haz_\alpha = {}_haz + \alpha \leq 0$. Végül a $h \in J$ esetben ${}_ha(1) = -1$, és ekkor (3.8) első egyenlőtlensége folytán ${}_haz_\alpha = {}_haz - \alpha \leq 0$.

(B) Egy x' vektor pontosan akkor megoldása (3.4)-nek, ha az $x := (x', -1)$ megoldása $Ax \leq 0$-nak. Továbbá x' pontosan akkor megoldása (3.5)-nek, ha az x megoldása $A^{[1]}x \leq 0$-nak. Alkalmazhatjuk a tétel **(A)** részét. •

3.4.2. Tétel. *Metszetkúp tengelymenti* (külső vagy belső) *vetülete metszetkúp. Poliéder tengelymenti vetülete poliéder.*

Biz. A 3.4.1 tétel szerint az $\{x : Ax \leq 0\}$ metszetkúp x_1 tengelymenti belső vetülete az $\{x : A^{[1]}x \leq 0,\ x_1 = 0\}$ metszetkúp. Miután $A^{[1]}$ első oszlopa 0-vektor, a külső vetület nem más, mint az $\{x' : A'^{[1]}x' \leq 0\}$ metszetkúp, ahol $A'^{[1]}$ jelöli az $A^{[1]}$-ből az első (azonosan nulla) oszlop elhagyásával keletkező mátrixot. Az $R = \{x' : Qx' \leq b\}$ poliéder x_1 menti belső vetülete a 3.4.1 tétel alapján az $\{x' : Q^{[1]}x' \leq b^{[1]},\ x_1 = 0\}$ poliéder, míg a külső vetülete az $\{x' : Q'^{[1]}x' \leq b^{[1]}\}$ poliéder, ahol $Q'^{[1]}$ az a mátrix, amely $Q^{[1]}$ első (azonosan nulla) oszlopának elhagyásával keletkezik. •

3.22. Feladat. *Igazoljuk, hogy poliéder lineáris képe poliéder.*

3.4.2. Poliéder = politóp + generált kúp

Politóp előáll poliéderként

A 2.3.3 következmény szerint minden generált altér nulltér és minden nulltér generált altér, ami azzal ekvivalens, hogy egy $Ax = 0$ homogén egyenletrendszer megoldáshalmaza előáll véges sok vektor lineáris burkaként, és megfordítva, véges sok vektor lineáris kombinációinak halmaza előáll egy homogén lineáris egyenletrendszer megoldáshalmazaként. Ennek általánosításaként igazolni fogjuk, hogy minden metszet-kúp generált kúp és minden generált kúp metszet-kúp.

A 2.3.4 tétel szerint ha az $Ax = b$ egyenletrendszernek x_0 egy megoldása, akkor a megoldások halmaza affin (=eltolt) altér, nevezetesen az A nullterének x_0-lal való eltoltja. Ennek általánosításaként igazolni fogjuk, hogy $\mathbb{R}^n$-ben egy halmaz pontosan akkor poliéder, ha egy politóp és egy generált kúp összege. Az egyik iránnyal kezdjük.

3.4.3. Tétel. *Egy politóp és egy generált kúp összege poliéder. Speciálisan, minden politóp korlátos poliéder és minden generált kúp előáll metszetkúpként.*

Biz. Tekintsük $\mathbb{R}^m$-ben az A ($m \cdot n$-es) mátrix oszlopai által generált kúpot és az A' ($m \cdot n'$-s) mátrix oszlopai által feszített politópot. Ezek összege a $C := \{z : z = Ax + A'x', (x, x') \geq 0, \underline{1}x' = 1\}$ halmaz, ahol $\underline{1}$ a csupa egyesből álló (n' dimenziós) vektort jelöli.

Tekintsük most $\mathbb{R}^{m+n+n'}$-ben az $R := \{(z, x, x') : Ax + A'x' - Iz = 0, x \geq 0, x' \geq 0, \underline{1}x' = 1\}$ poliédert. Ha R-nek vesszük a külső vetületét az (x, x') komponenseinek megfelelő koordináták mentén, akkor (definíció szerint) azon z vektorok halmazát kapjuk, melyekhez van olyan (x, x'), hogy $Ax + A'x' - Iz = 0, x \geq 0, x' \geq 0, \underline{1}x' = 1$, azaz $z = Ax + A'x'$. Vagyis a külső vetület éppen C, és így a 3.4.2 következmény miatt C valóban poliéder. •

A Farkas lemma

A lineáris programozás egyik sarokköve a Farkas lemma. Ennek több ekvivalens megfogalmazása van: a legszemléletesebb alakkal kezdjük.

3.4.4. Tétel (Farkas lemma, geometriai alak). *Ha egy $C \subseteq \mathbb{R}^k$ generált kúp nem tartalmaz valamely $b \in \mathbb{R}^k$ elemet, akkor létezik olyan* (zárt) *homogén féltér, amely magában foglalja C-t, de nem tartalmazza b-t. Ha egy P politóp nem tartalmazza b-t, akkor létezik olyan féltér, amely magában foglalja P-t, de nem tartalmazza b-t.*

Biz. Mivel minden generált kúp metszetkúp, azaz véges sok homogén féltér metszete, így ha b nincs a kúpban, akkor nincs benne ezen félterek valamelyikében. A második rész ugyanígy következik abból, hogy minden politóp poliéder, azaz véges sok féltér metszete. •

A Farkas lemma eredeti, Farkas Gyula által kimondott alakja a következő.

3.4.5. Tétel (Farkas lemma, standard alak). *Az $\{Ax = b, x \geq 0\}$ rendszernek pontosan akkor van megoldása, ha az $\{yA \geq 0, yb < 0\}$ rendszernek nincs.*

Biz. Mindkét rendszernek nem lehet megoldása, mert akkor $0 \leq (yA)x = y(Ax) = yb < 0$ volna. A fordított irányhoz tekintsük az A oszlopai által generált K kúpot és tegyük fel, hogy az $\{Ax = b, x \geq 0\}$ rendszernek nincsen megoldása. Ez éppen azt jelenti, hogy b nincs benne K-ban A 3.4.4 tétel szerint van olyan F homogén féltér, amely K-t magában foglalja, de b-t nem. Az F normálisát y-nal jelölve ez avval ekvivalens, hogy az A mindegyik a_i oszlopára $ya_i \geq 0$ és $yb < 0$. •

Az $\{Ax = b, x \geq 0\}$ rendszert **primál** feladatnak, míg az $\{yA \geq 0, yb < 0\}$ rendszert **duál** vagy **duális** feladatnak nevezik. Megjegyzendő, hogy ha a szóbanforgó y létezik, akkor az úgy is megválasztható, hogy $yb = -1$ teljesüljön, így néha azt hívjuk duális feladatnak, amikor az $yb < 0$ helyett $yb = -1$-et követelünk. Néha a duálist az $\{yA \leq 0, yb > 0\}$ ekvivalens alakban adják meg: ennek és az eredeti duálisnak a megoldásai egymás minusz egyszeresei.

Érdekes, hogy a 3.4.5 tételnek nemcsak az oszloptérben, hanem a sortérben is szemléletes jelentés adható. Ugyanis a primál probléma megoldása azzal ekvivalens, hogy létezik olyan nemnegatív $(x, 1)$ vektor, amely ortogonális az $A' := (A, -b)$ mátrix soraira, ami viszont azzal ekvivalens, hogy az A' mátrix nullterében létezik egy olyan x' nemnegatív vektor, amelynek utolsó komponense szigorúan pozitív. A duál problémában $yb < 0$ helyett $y(-b) > 0$-t írva, a duál probléma megoldhatósága azt jelenti, hogy van olyan y vektor, amelyre $yA' \geq 0$ és yA' utolsó komponense szigorúan pozitív; magyarán azt, hogy az A' mátrix sorterében van olyan nemnegatív vektor, amelynek utolsó komponense szigorúan pozitív. Miután egy tetszőleges mátrix nulltere és sortere egymás ortogonális kiegészítő alterei, továbbá tetszőleges altér és ortogonális kiegészítő altere megadható egy mátrix nullatereként illetve sortereként, a Farkas lemma ekvivalens megfogalmazása a következő.

3.4.6. Tétel (Farkas lemma: 2. geometriai alak). *Tetszőleges altér és ortogonális kiegészítő altere közül pontosan az egyik tartalmaz olyan nemnegatív vektort, amelynek utolsó komponense pozitív.* •

Figyeljük meg, hogy két ilyen vektor skaláris szorzata biztosan pozitív, azaz nem lehetnek merőlegesek egymásra, tehát két ortogonális altér közül legfeljebb csak az egyik tartalmazhat ilyen vektort.

Korlátos poliéder előáll politópként

A 3.4.3 tétel szerint egy politóp és egy generált kúp összege poliéder. Megmutatjuk, hogy érvényes a megfordítás is, azaz minden poliéder előáll, mint egy politóp és egy generált kúp összege. A bizonyítás érdekessége, hogy megfordítás igazolásához magát a 3.4.3 tételt használjuk fel.

Tekintsük először a speciális esetet, amikor egy metszetkúpot akarunk előállítani generált kúpként. Jelölje $G_A := \{yA : y \geq 0\}$ az A mátrix sorai által generált kúpot, míg $M_A := \{x : Ax \leq 0\}$ az A sorai által definiált metszetkúpot.

3.4.7. Lemma. **(A)** *Ha $G_A \supseteq M_B$, akkor $M_A \subseteq G_B$.* **(B)** *Ha $G_A \subseteq M_B$, akkor $M_A \supseteq G_B$.* **(C)** *$G_A = M_B$ akkor és csak akkor ha $M_A = G_B$.*

Biz. **(A)** Tegyük fel, hogy $G_A \supseteq M_B$ és lássuk be, hogy $M_A \subseteq G_B$. Ehhez legyen $z \in M_A$, vagyis ${}_ia z \leq 0$ fennáll az A minden sorára. Emiatt az A sorainak bármely $q = \sum \lambda_i \ {}_ia \ \ (\lambda_i \geq 0)$ nemnegatív kombinációjára $qz \leq 0$, azaz

$$G_A \text{ minden } q \text{ elemére } qz \leq 0. \tag{3.10}$$

Másrészt, ha indirekt z nincs a G_B generált kúpban, akkor a Farkas lemma (3.4.4 tétel) szerint van olyan homogén féltér, amely tartalmazza G_B-t, de z-t nem. Vagyis létezik olyan q vektor (a féltér határoló hipersíkjának normálisa), amelyre egyrészt $Bq \leq 0$, azaz $q \in M_B \subseteq G_A$, másrészt $qz > 0$, ellentmondva (3.10)-nek. Tehát $M_A \subseteq G_B$.

(B) Tegyük most fel, hogy $G_A \subseteq M_B$. Ekkor G_A minden sorvektora M_B-ben van, azaz ${}_ia \ {}_jb \leq 0$ fennáll az A minden ${}_ia$ és a B minden ${}_jb$ sorára. Emiatt minden $y \in G_B$ elemre is érvényes ${}_iay \leq 0$ vagyis $y \in M_A$. Tehát $G_B \subseteq M_A$.

A **(C)** állítás igazolásához figyeljük meg, hogy az **(A)** és **(B)** részek összetevéséből kapjuk, hogy ha $G_A = M_B$, akkor $M_A = G_B$. Az A és a B szerepének felcseréléséből pedig adódik, hogy $G_B = M_A$, akkor $M_B = G_A$. •

3.4.8. Tétel. *Minden metszetkúp előáll generált kúpként.*

Biz. Az A mátrix sorai által definiált M_A metszetkúpról fogjuk kimutatni, hogy generált kúp. A 3.4.3 tétel szerint a G_A generált kúp előáll metszetkúpként, azaz létezik egy olyan B mátrix, amelyre $G_A = M_B$. A 3.4.7 lemma nyomán $M_A = G_B$, vagyis M_A a B sorai által generált kúp. •

A metszetkúpok előbbi előállítására támaszkodva megadjuk a poliéderek előállítását, amely tehát a 3.4.3 tétel megfordításának tekinthető.

3.4.9. Tétel. *Minden nemüres poliéder előáll mint egy politóp és egy generált kúp összege. Speciálisan, minden korlátos poliéder politóp.*

Biz. Legyen $R = \{x : Qx \leq b\}$ nemüres poliéder. A bizonyítás ötlete az, hogy R-t beágyazzuk egy eggyel magasabb dimenziós tér $\{(x, \lambda) : x \in \mathbb{R}^n, \lambda = 1\}$ hipersíkjába, ahol R az origóval egy kúpot feszít, majd ezen kúpra alkalmazzuk a 3.4.8 tételt.

Tekintsük tehát az eggyel magasabb (azaz $n+1$) dimenziós térben az $M := \{(x,\lambda) : Qx - \lambda b \leq 0, \lambda \geq 0\}$ metszetkúpot. Ez a 3.4.8 tétel szerint előáll generált kúpként. Feltehető, hogy a generáló elemek utolsó (λ-nak megfelelő) koordinátáinak mindegyike 1 vagy 0. Legyenek a generáló elemek eszerint szétválasztva: $(x_1, 1), \dots, (x_k, 1)$, $(x'_1, 0), \dots, (x'_\ell, 0)$. Tekintsük $\mathbb{R}^n$-ben az x_i-k konvex burkaként definiált P politópot valamint az x'_j vektorok által generált C kúpot.

Azt állítjuk, hogy $R = P + C$. Valóban, legyen először $z \in R$. Ekkor $(z, 1) = \sum_i \lambda_i(x_i, 1) + \sum_j \mu_j(x'_j, 0)$, ahol $\lambda_i \geq 0, \mu_j \geq 0$. Most tehát $\sum_i \lambda_i = 1$ és ezért $z_1 := \sum_i \lambda_i x_i$ benne van P-ben és $z_2 := \sum_j \mu_j x'_j$ benne van C-ben. Így z valóban előáll egy P-beli z_1 és egy C-beli z_2 elem összegeként, azaz $R \subseteq P + C$. Legyen most $z_1 \in P$, $z_2 \in C$. Ekkor $(z, 1) = (z_1, 1) + (z_2, 0)$ benne van M-ben, azaz $Qz - b \leq 0$, vagyis z benne van R-ben, azaz $P + C \subseteq R$, és így $R = P + C$. •

3.23. Feladat. *Minden generált kúp a polárisának polárisa, azaz* $(K^*)^* = K$. (Egy C kúp polárisán a $C^* := \{x : xy \leq 0 \text{ minden } y \in C\text{-re}\}$ kúpot értettük.)

A 3.4.9 tételben már láttuk, hogy minden korlátos poliéder politóp, azaz véges sok pont konvex burka, sőt a 3.3.19 tételben azt is beláttuk, hogy minden nemüres korlátos poliéder a csúcsainak konvex burka. A bizonyítás gondolatmenetét használva most megadjuk a csúcsos poliéderek előállítását.

3.4.10. Tétel. *Minden $R \neq \emptyset$ egyenes-mentes* (azaz csúcsos) *poliéder előáll, mint a csúcsai által feszített R_K politóp valamint a poliéder C karakterisztikus kúpjának a vektor-összege. Speciálisan, minden korlátos poliéder előáll csúcsainak konvex burkaként.*

Biz. Tegyük fel, hogy $R = \{x : Qx \leq b\}$ és legyen z a poliéder egy pontja. Legyen $Q_z^=$ a z aktív részmátrixa. A 3.3.13 tétel alapján R csúcsai éppen a 0 szintű pontok.

A szint szerinti indukcióval fogjuk kimutatni, hogy a poliéder minden z pontja előáll egy R_K-beli és egy C-beli pont összegeként. Ha a szint nulla, akkor tehát z csúcs, így eleme R_K-nak. Tegyük most fel, hogy $\sigma(z) > 0$ és azt, hogy minden alacsonyabb szintű pontra a szóbanforgó előállítás létezik.

Mivel $r(Q) > r(Q_z^=)$, létezik olyan q, amelyre $Q_z^= q = 0$ és $Q_z^< q \neq 0$. Feltehető, hogy $Q_z^< q$-nak van negatív komponense, különben q-t helyettesíthetjük a negatívjával.

Tegyük először fel, hogy $Q_z^< q$-nak nincsen pozitív komponense, azaz $Q_z^< q < 0$ (vagyis $q \in C$). Ekkor létezik olyan $\lambda_1 > 0$ szám, amelyre $x_1 := z - \lambda_1 q$ az R-ben van és $\sigma(x_1) < \sigma(z)$. (Nevezetesen, $\lambda_1 := \min\{(b_1(i) - {}_i a z)/(-{}_i a q)\}$, ahol a minimum az Q azon ${}_i a$ soraira megy, amelyekre ${}_i a q < 0$. Az x_1 szintje valóban kisebb, mint z szintje, hiszen egyrészt $Q_z^= x_1 = b_z^=$, másrészt $Q_z^<$ egyik ${}_i q$ sorára ${}_i q x_1 = b(i)$ és ez a sor ${}_i q q \neq 0$ és $Q_z^= q = 0$ miatt lineárisan független $Q_z^=$ soraitól.) Indukció alapján x_1 előáll $x_1 = y_1 + y'_1$ alakban, ahol $y_1 \in R_K$ és $y'_1 \in C$. De most $z = x_1 + \lambda_1 q = y_1 + y'_1 + \lambda_1 q$ és $y'_1 + \lambda_1 q \in C$, így z is előáll az R_K-beli y_1 és egy C-beli elem összegeként.

Tegyük most fel, hogy $Q_z^< q$-nak létezik pozitív komponense is, és legyen ${}_1 q$ és ${}_2 q$ a $Q_z^<$ mátrixnak két olyan sora, amelyre ${}_1 q q < 0$ és ${}_2 q q > 0$. Ekkor léteznek pozitív λ_1 és λ_2 számok, melyekre $x_1 = z - \lambda_1 q \in R$ és $x_2 = z + \lambda_2 q \in R$, és mind az x_1, mind az x_2 szintje kisebb, mint a z-é.

Indukcióval $x_i = y_i + y'_i$ $(i = 1, 2)$, ahol $y_i \in R_K$ és $y'_i \in C$. De ekkor $z = (\lambda_2 x_1 + \lambda_1 x_2)/(\lambda_1 + \lambda_2) = (\lambda_2 y_1 + \lambda_1 y_2)/(\lambda_1 + \lambda_2) + (\lambda_2 y'_1 + \lambda_1 y'_2)/(\lambda_1 + \lambda_2)$. Itt az összeg első tagja R_K-ban van, míg a második tagja C-ben. •

Megjegyzés Ezt a bizonyítást valójában a speciális korlátos esetben már korábban, a 3.3.19 tétel (2) pontjának bizonyításakor elmondtuk.

3.4.3. Az FM eljárás hatékonysága

Algoritmikus szempontból a Fourier-Motzkin eliminációt a $Qx \leq b$ egyenlőtlenség-rendszer megoldására a következőképp lehet használni. Feltehető, hogy a Q első oszlopa $0, \pm 1$ értékű, mert egy egyenlőtlenséget pozitív számmal szorozva ekvivalens rendszert kapunk. Legyen Q_1 az a mátrix, amely $Q^{[1]}$-ből keletkezik az első (csupa 0) oszlop eltörlésével. A 3.4.1 téetel folytán $Qx \leq b$ megoldhatósága ekvivalens $Q_1x_1 \leq b^{[1]}$ megoldhatóságával, és a 3.4.1 tétel bizonyítása meg is mondja, hogy egy x_1-ből hogyan lehet kiszámolni $Qx \leq b$ egy megoldását. Q_1-nek eggyel kevesebb oszlopa van mint Q-nak, így az eliminációs lépést n-szer kell használni.

Következésképp a Fourier-Motzkin algoritmus véges. Az eljárás hátránya, hogy egyetlen változó eliminálása sok új egyenlőtlenséget hoz be, vagyis menet közben az egyenlőtlenségek száma nagyon felszaporodhat. Sajnos ez nem csak elvi lehetőség, amint a következő példa mutatja. Legyen p pozitív egész és $n := 2^p + p + 2$. n darab változónk lesz: $x_1, \ldots, x_n$. $8\binom{n}{3}$ egyenlőtlenségünk van, mindegyik $\pm x_i \pm x_j \pm x_k \leq b_{ijk}$ alakú, ahol b_{ijk} adott számok. t szerinti indukcióval látható, hogy t változó kiejtése után az új egyenlőtlenségek között szerepelni fog az összes $\pm x_{j_1} \pm x_{j_2}, \ldots, \pm x_{j_s} \leq b_{j_1,\ldots,j_s}$ alakú, ahol $s = 2^t + 2$ és $t + 1 \leq j_1 < j_2 \cdots < j_s \leq n$. Így p lépés után a megmaradt változók száma $n' = n - p = 2^p + 2$ míg az egyenlőtlenségek száma legalább $2^{2^p+2} = 2^{n'}$.

Tapasztalat szerint a FM elimináció legfeljebb csak kis példákon használható, nagyobbakon tipikusan ténylegesen kezelhetetlenül sok egyenlőtlenség keletkezik.

A fenti példában olyan egyenlőtlenség-rendszer szerepelt, amelynek minden sorában három 1 abszolút értékű együttható volt. Érdekes, hogy ha csak két 1 abszolút értékű együttható van minden sorban, akkor a Fourier-Motzkin eljárás polinomiális futásidejű. Egy (sor)vektort nevezzünk **szimplá**nak, ha vagy egy nemnulla eleme van (és erről feltehető, hogy 1 abszolút értékű), vagy kettő, melyek mindegyike 1 abszolút értékű. Szimpla sorokból álló mátrixot is nevezzünk szimplának. Könnyű megfigyelni, hogy ha a kiindulási mátrix szimpla, akkor az FM eljárás során keletkező új mátrix is az lesz. Miután szimpla vektor csak kevés (legfeljebb $2m + 3m(m-1)/2$) lehet, ezért az FM eljárás szimpla mátrixokra polinomiális. Természetesen menetközben a redundáns egyenlőtlenségeket ki kell dobni: ha például az $x_1 + x_2 \leq 3$ és az $x_1 + x_2 \leq 4$ egyenlőtlenségek adódnak ki, akkor az utóbbi felesleges.

Szimpla mátrixokra, egészértékű b esetén azt is el lehet dönteni, hogy a $Qx \leq b$ rendszernek létezik-e egészértékű megoldása. Az az egyetlen különbség, hogy ha menetközben keletkezik mondjuk egy $2x_i \leq \beta$ alakú egyenlőtlenség, akkor ezt az $x_i \leq \lfloor \beta/2 \rfloor$ egyenlőtlenséggel kell helyettesíteni.

3.4.4. Alkalmazások

A 2-SAT probléma

Példaként említhetjük az ún. 2-SAT problémát (2-satisfiability = 2-kielégíthetőség). Ez gráfok nyelvén elmondva azt kívánja eldönteni, hogy egy M teljes párosítással rendelkező $G = (V, E)$ gráf pontjai közül ki lehet-e választani $|M|$ darabot, melyek az összes élt lefogják. Ennek eldöntésére, minden v csúcshoz rendeljünk egy $x(v)$ változót, és tekintsük a következő egyenlőtlenség-rendszert. $0 \leq x(v) \leq 1$ minden $v \in V$-re, $x(u) + x(v) = 1$ minden $uv \in M$ élre, és $x(u) + x(v) \geq 1$ minden egyéb uv élre. A 2-SAT problémának pontosan akkor van megoldása, ha ezen egyenlőtlenség-rendszernek van egész megoldása. Miután a feltételi mátrix szimpla, az FM elimináció polinomiális futásidejű algoritmust szolgáltat.

Érdemes az FM elimináció egy lépését közvetlenül gráfnyelven elmondani: Legyen $uv \in M$ az egyik párosítás él. A $V - \{u, v\}$ ponthalmazon definiáljuk a $G^{[1]}$ gráfot úgy, hogy xy él, ha eredetileg él, vagy ha x szomszédos u, v egyikével és y szomszédos a másikával. Könnyen ellenőrizhető közvetlenül is, hogy a 2-SAT probléma akkor és csak akkor oldható meg G-re, ha megoldható $G^{[1]}$-re.

3.24. Gyakorlat. *Tekintsük négy változóban az $x_i \geq 0$, $x_i + x_j = 1$ $(1 \leq i < j \leq 4)$ rendszert. Döntsük el az FM eliminációval, hogy van-e megoldása és van-e egész megoldása.*

Az ütemezési feladat újra

Abban a speciális esetben, amikor a mátrix minden sorában vagy egyetlen nemnulla elem szerepel és ennek az abszolút értéke 1, vagy pedig egy $+1$-es és egy -1-es, akkor az FM elimináció automatikusan fenntartja ezt az alakot. Következik, hogy ha a jobboldali b vektor egészértékű, és a $Qx \leq b$ rendszernek van megoldása, akkor van egész megoldása is (és az FM egy ilyet meg is talál).

Az 1.3.2 részben már megfogalmaztuk egy nagyobb projekt részfeladatainak ütemezési problémáját és megadtunk egy megoldást is. Érdekességként most megmutatjuk, hogy ilyenkor az FM módszer is segít. Az ott felállított matematikai modellben tehát a feladat olyan $\pi : V \to \mathbb{R}$ függvény meghatározásával ekvivalens, amelyre minden uv élnek az előre adott súlya legfeljebb $\pi(v) - \pi(v)$, minden csúcsra $f(v) \leq \pi(v) \leq g(v)$, ahol f és g előre adott korlátok, és $\pi(t) - \pi(s) \leq T$. Ha ezen egyenlőtlenségeket felírjuk, akkor mindegyikükben vagy egyetlen $+1$ vagy -1 szerepel, vagypedig egy darab $+1$ és egy darab -1. A fentiek szerint ilyen esetben az FM elimináció polinomiális algoritmust szolgáltat, rádásul, ha a súlyok egészértékűek és létezik π megoldás, akkor létezik egészértékű π is.

3.25. Feladat. *Adjunk a Fourier-Motzkin eljárás segítségével új bizonyítást Gallai 1.3.8 tételére miszerint egy élsúlyozott irányított gráfban akkor és csak akkor van negatív össz-súlyú irányított kör, ha nincsen olyan $\pi := V \to \mathbb{R}$ függvény, amelyre $\pi(v) - \pi(u) \leq c(uv)$ fennáll minden uv élre. Továbbá, ha a c súlyfüggvény egészértékű, akkor π is választható annak.*

3.5. Megoldhatóság: a Farkas lemma

A Fredholm tétel egyszerűen ellenőrizhető tanúsítványt adott arra, hogy ha az $Ax = b$ rendszernek nincsen megoldása, hiszen a ilyenkor létezik egy y, amelyre $yA = 0, yb \neq 0$, és egy adott y-ról ennek fennállása szintén könnyen eldönthető. Hasonlóképp, a Farkas lemma standard alakja arra szolgáltat tanúsítványt, hogy ha az $\{Ax = b, x \geq 0\}$ rendszernek nincsen megoldása.

Mivel egy poliédert többféle alakban is meg lehet adni, a Farkas lemmának is különböző változatai vannak. Ezek azonban egyszerű fogással következnek egymásból, ezért minden alakot Farkas lemmának nevezünk majd. A három leggyakoribb algebrai változatot adjuk meg, majd közös általánosításként egy olyan formát is felírunk, amelyből mind a három alak speciális esetként kiadódik.

A 3.3.7 és a 3.3.10 tételek alkalmazásával megkaphatjuk a Farkas lemma standard alakjának egy élesítését.

3.5.1. Tétel. *Az $\{Ax = b,\ x \geq 0\}$ rendszernek akkor és csak akkor van olyan megoldása, amelyben az x pozitív változóinak megfelelő A-beli oszlopok lineárisan függetlenek, ha nem létezik olyan y, amelyre $yA \geq 0,\ yb = -1$ és A-nak létezik $r(A, b) - 1$ lineárisan független oszlopa, amelyekre y merőleges.* (Röviden, vagy a primál, vagy a duál problémának létezik bázis-megoldása). •

A primál feltétel geometriailag azt mondja, hogy ha a b vektor benne van néhány vektor K kúpjában, akkor már benne van ezen vektorok közül vett néhány lineárisan független vektor kúpjában is. Egy y duális bázis-megoldás geometriailag a következőt jelenti. Amennyiben $r(A, b) = r(A) + 1$, úgy az y ortogonális $r(A)$ lineárisan független oszlopra, ezért $yA = 0$, vagyis az y normálisú homogén hipersík tartalmazza A oszlopait, de b-t nem. Amennyiben $r(A, b) = r(A)$, úgy az y normálisú K-t tartalmazó, b-t nem tartalmazó homogén féltér olyan, hogy határoló hipersíkja, amely a kúp egy „határoló lapját” (maximális valódi oldalát) tartalmazza. Ha a kúp teljes dimenziós, úgy a hipersík ezen határoló lap hipersíkja.

A Farkas lemma változatai

3.5.2. Tétel (Farkas lemma, (A) változat)**.** *A $Qx \leq b$ egyenlőtlenség-rendszer akkor és csak akkor oldható meg, ha nem létezik olyan $y \geq 0$, amelyre $yQ = 0, yb = -1$.*

Biz. A $Qx \leq b$ feladatot nevezzük primál problémának, az $\{yQ = 0, y \geq 0, qb = -1\}$ feladatot pedig duálnak. Mindkettő nem oldható meg, mert akkor $0 = y(Qx) \leq yb = -1$ állna. A fordított irányhoz jelölje A a $[Q, b]$ mátrixot, és legyen $c' = (0, \dots, 0, -1)$ egy $(n+1)$-dimenziós vektor. A duál megoldhatósága azt jelenti, hogy c' benne van az A sorai által generált kúpban. Ha nincs benne, akkor a Farkas lemma geometriai alakja szerint van olyan $x' = (x, \alpha)$ $(n + 1)$-dimenziós vektor, amelyre $Ax' \leq 0$ és $c'x' = 1$, ami azzal ekvivalens, hogy $\alpha = -1$ és $Qx \leq b$, vagyis a primál feladat megoldható. •

Alternatív bizonyítás A standard alakra történő visszavezetés két lépésben történik. Egyrészt az előjel kötetlen x változót a nemnegatív x' és x'' különbségeként írjuk fel. Másrészt, egy x_1 nemnegatív (m-dimenziós), úgynevezett **eltérés** (vagy **pót**) változó (slack variable) bevezetésével az egyenlőtlenség-rendszert egyenlőség-rendszerré alakítjuk. Ekkor a $Qx \leq b$ primál feladat a $Qx' - Qx'' + x_1 = b,\ (x', x'', x_1) \geq 0$ alakba megy át. Az $A := (Q, -Q, I_m)$ mátrixra alkalmazva a Farkas lemma standard alakját, azt kapjuk, hogy a megoldhatósághoz szükséges és elegendő feltétele az, hogy nem létezik y, amelyre $yb < 0$ és $yA \geq 0$, azaz $yQ \geq 0$ $y(-Q) \geq 0$, $yI_m \geq 0$, vagyis $yQ = 0,\ y \geq 0$. •

3.5.3. Tétel (Farkas lemma, (B) változat). *A* $\{Bx \leq b, x \geq 0\}$ *egyenlőtlenség-rendszer akkor és csak akkor oldható meg, ha nem létezik olyan* $y \geq 0$, *amelyre* $yB \geq 0, yb = -1$.

Biz. Mindkét rendszer nem oldható meg, mert akkor $0 \leq (yB)x = y(Ax) \leq yB = -1$. Jelölje Q azt a mátrixot, amely B-ből keletkezik azáltal, hogy aláírjuk az $n \times n$-es $-I$ egységmátrixot és jelölje b_1 azt a vektort, amely b-ből keletkezik n darab 0 komponens hozzáfűzésével. A $\{Bx \leq b, x \geq 0\}$ rendszer pontosan akkor oldható meg, ha $Qx \leq b_1$ megoldható. Az (A) változat szerint, ha $Qx \leq b_1$ nem oldható meg, akkor létezik egy olyan $(y, y') \geq 0$ vektor, amelyre $yB + y'(-I) = 0, yb = -1$. De $y' \geq 0$ miatt $yB + y'(-I) = 0$ azzal ekvivalens, hogy $yB \geq 0$. •

Egyszerű fogással a Farkas lemmát olyan alakban is megfogalmazhatjuk, amikor a Fredholm féle alternatíva tételt már explicit magában foglalja.

3.5.4. Tétel. *A*

$$\{Px = b_0, Qx \leq b_1\} \tag{3.11}$$

primál rendszer akkor és csak akkor oldható meg, ha az

$$\{y_0P + y_1Q = 0, y_1 \geq 0, yb = -1\} \tag{3.12}$$

duális nem, ahol $y = (y_0, y_1), b = (b_0, b_1)$.

Biz. Mindkét feladat nem oldható meg, mert akkor $0 = (y_0P + y_1Q)x = (y_0P)x + (y_1Q)x = y_0(Px) + y_1(Qx) \leq y_0b_0 + y_1b_1 = -1$.

Ha a primál probléma nem oldható meg, akkor a vele ekvivalens $\{Px \leq b_0, -Px \leq -b_0, Qx \leq b_1\}$ egyenlőtlenség-rendszer sem. Ekkor viszont a Farkas lemma (A) változata alapján az ehhez tartozó duál megoldható, azaz létezik $(y_0', y_0'', y_1) \geq 0$ vektor, amelyre $y_0'P + y_0''(-P) + y_1Q = 0$ és $y_0'b_0 + y_0''(-b_0) + y_1b_1 = -1$. De ekkor $y_0 := y_0' - y_0''$-re az (y_0, y_1) megoldása a (3.12) duálisnak. •

Hasznos egy olyan alakot is felírni, amely a fenti változatok mindegyikét magában foglalja.

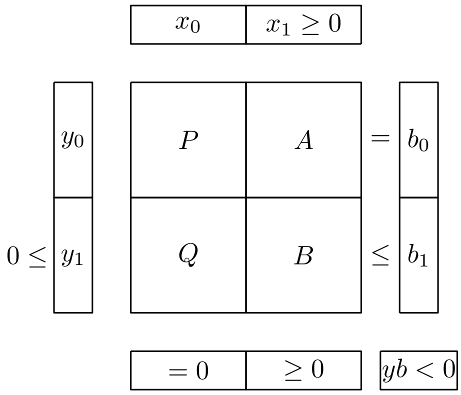
3.5.5. Tétel (Farkas lemma, általános alak). *A*

$$Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \leq b_1,\ x_1 \geq 0 \tag{3.13}$$

primál rendszernek akkor és csak akkor nincs megoldása, ha az

$$y_0P + y_1Q = 0,\ y_0A + y_1B \geq 0,\ y_1 \geq 0,\ y_0b_0 + y_1b_1 = -1 \tag{3.14}$$

duális rendszernek van.



3.1. ábra. A Farkas lemma általános alakja

Biz. Jelölje az x_0, x_1, y_0, y_1 dimenzióját rendre n_0, n_1, m_0, m_1. Az $x_1 \geq 0$ feltételt explicit beírhatjuk az egyenlőtlenségek közé $0_{n_1 n_0} x_0 + x_1(-I_{n_1 n_1}) \leq 0$ alakban. A 3.5.4 tételből közvetlenül kapjuk, hogy a primál probléma pontosan akkor oldható meg, ha az $\{y_0 P + y_1 Q = 0, y_0 A + y_1 B + y_1'(-I_{n_1 n_1}) = 0, (y_1, y_1') \geq 0, y_0 b_0 + y_1 b_1 = -1\}$ rendszer nem. Ez utóbbi megoldhatósága viszont y_1' nemnegatívitása folytán épp (3.14) duálisével ekvivalens. •

Megjegyezzük, hogy a Farkas lemmát még általánosabb alakban is fel lehetne írni. Például a primál feladatban lehetnek fordított irányú egyenlőtlenségek, vagy nempozitív változók. A megfelelő egyenlőtlenség illetve a feltételi mátrix megfelő oszlopának negálásával azonban könnyen a (3.15) alakra juthatunk, így ez a legáltalánosabb alak már nem ad igazán újat.

3.5.1. Direkt bizonyítás

Bár a Farkas lemma standard alakját már levezettük korábbi eredményekből (nevezetesen a Fourier-Motzkin eliminációra támaszkodva), érdemes egy közvetlen bizonyítást is megadni.

3.5.6. Tétel (Farkas lemma, standard alak). *Az $\{Ax = b, x \geq 0\}$ rendszernek pontosan akkor van megoldása, ha az $\{yA \geq 0, yb < 0\}$ rendszernek nincs.*

Biz. Csak a nemtriviális iránnyal foglalkozunk, és azt mutatjuk ki, hogy a primál és duál problémák egyike megoldható. A szokásnak megfelelően A-nak legyen $m \geq 1$ sora és $n \geq 1$ oszlopa. $m + n$ szerinti indukciót alkalmazunk. A lemma állítása közvetlenül látszik az $m = 1$ esetben, így feltehető, hogy $m \geq 2$. Figyeljük meg, hogy az $Ax = b$ egyenletrendszer valamely sorát nemnulla számmal szorozva a primál probléma megoldáshalmaza és a duál probléma megoldhatósága sem változik. Hasonló kijelentés érvényes, ha az egyenletrendszer egyik egyenletét hozzáadjuk egy másikhoz (vagy levonjuk belőle). Ezért, esetleges sorcserét is alkalmazva, feltehetjük, hogy az A első oszlopa az $(1, 0, \ldots, 0)$ vektor. Tegyük fel először, hogy $n = 1$. Ha $b(1) \geq 0$, akkor az n-dimenziós $x = (b(1), 0, \ldots, 0)$ primál megoldás. Ha $b(1) < 0$, akkor az m-dimenziós $y = (1, 0, \ldots, 0)$ duál megoldás. Feltehetjük tehát, hogy $n \geq 2$.

Jelölje A' az a_1 oszlop elhagyásával keletkező mátrixot. Ha az $\{A'x' = b, x' \geq 0\}$ rendszernek létezik x' megoldása, akkor ez elé egy 0 komponenst írva az eredeti primál feladat megoldását kapjuk. Ha az $\{A'x' = b, x' \geq 0\}$ rendszernek nem létezik megoldása, akkor indukció miatt létezik y', melyre $y'A' \geq 0$ és $y'b < 0$. Amennyiben $y'a_1 \geq 0$, úgy y' az eredeti duálnak is megoldása. Akkor nem vagyunk kész, ha $y'a_1 < 0$.

Tekintsük most a $\{A''x'' = b'', x'' \geq 0\}$ rendszert, ahol (A'', b'') az (A, b) első sorának törlésével keletkezik. Ha most a duálisnak van egy y'' megoldása, akkor ez elé egy 0 komponenst írva az eredeti duális megoldását kapjuk. Ha a duálisnak nincs megoldása, akkor indukció miatt létezik $x'' \geq 0$, amelyre $A''x'' = b''$. Mivel A'' első oszlopa nullvektor, $x''(1)$ bármi lehet, ezért $x''(1)$-t úgy megválaszthatjuk, hogy ${}_1ax'' = b(1)$. Belátjuk, hogy $x''(1) \geq 0$, amiből következik, hogy x'' az eredeti primál probléma megoldása. Ha ugyanis indirekt $x''(1) < 0$, akkor mind az x'', mind az $y'A$ olyan vektor, melyeknek pontosan az első komponense negatív és ezért $0 < (y'A)x'' = y'(Ax'') = y'b < 0$, amely ellentmondás bizonyítja a Farkas lemmát. •

Megjegyzés Érdemes nyomon követni a fenti bizonyítás hátterében megbúvó geometriai szemléletet. Amikor a Gauss eliminációval az a_1 oszlopot az $(1, 0, \ldots, 0)$ vektorrá alakítjuk, ennek az a jelentése, hogy az A oszlopait egy olyan bázisban írjuk fel, melynek első tagja a_1. Az $a_1 = (1, 0, \ldots, 0)$ esetben a mátrix első sorának törlése geometriailag azt jelenti, hogy az A oszlopvektorait vetítjük az a_1 normálisú hipersíkra.

A fenti bizonyítás során az a_1 törlésénél az elintézetlen eset az volt, amikor létezett egy olyan y' vektor, amelyre az $F = \{z : y'z \geq 0\}$ homogén féltér tartalmazta az $a_2, \ldots, a_n$ oszlopvektorokat, de sem a_1-et, sem és b-t nem. Ekkor F tartalmazza a $-a_1, a_2, a_3, \ldots, a_n$ vektorokat, és emiatt a b bizonyosan nem lehet benne ezen vektorok generált kúpjában. Másrészt, az a_1-re merőleges hipersíkra való vetítésnél (azaz a mátrix első sorának törlésénél) a bajos eset az volt, ha létezik egy olyan x'', amelyre $Ax'' = b$, az x'' első komponense negatív, míg a többi nem. Ez viszont éppen azt jelenti, hogy a b előáll a $-a_1, a_2, a_3, \ldots, a_n$ vektorok nemnegatív kombinációjaként, amely eshetőséget az előbb zártuk ki az F féltér segítségével.

Még egy bizonyítást leírunk, amelyben nincs szükség olyan Gauss elimináció szerű műveletre, amellyel az előbbi bizonyításban az első oszlopot egységvektorrá alakítottuk. Kényelmesebbnek bizonyul a 3.5.4 tételben megfogalmazott kicsit általánosabb alakot igazolni. (Nem ritka jelenség, hogy egy jól eltalált általánosításra az indukciós bizonyítás gördülékenyebben működik.) Szemben az ottani megközelítéssel, amely valójában speciális esetként kiadta a Fredholm féle alternatíva tételt, az itt következő bizonyítás használja azt.

3.5.7. Tétel. *A*

$$Px_0 + Ax_1 = b, \ x_1 \geq 0 \tag{3.15}$$

primál feladatnak akkor és csak akkor nincsen megoldása, ha az

$$yP = 0, \ yA \geq 0, \ yb = -1 \tag{3.16}$$

duális feladatnak van.

(Figyeljük meg, hogy (3.16) megoldhatósága ekvivalens az $\{yP = 0, yA \geq 0, yb < 0\}$ rendszer megoldhatóságával.)

Biz. A primál és a duál feladat nyilván nem oldható meg egyszerre, mert akkor $0+0 \leq 0+(yA)x_1 = (yP)x_0 + (yA)x_1 = y[(P,A)x] = yb = -1$, azaz $0 \leq -1$ következne.

Annak bizonyítására, hogy a primál és a duál feladatok egyike biztosan megoldható az A oszlopai száma szerinti indukciót alkalmazunk. Amennyiben ez az n_1-gyel jelölt szám nulla, azaz A üres, úgy a tétel következik a Fredholm féle alternatíva tételből. Tegyük tehát fel, hogy n_1 pozitív és indukció alapján azt, hogy n_1-nél kisebb oszlopszámra a tétel érvényes!

Legyen a_1 az A mátrix első oszlopa. Jelölje A' azt a mátrixot, amely A-ból keletkezik az a_1 kihagyásával. Amennyiben a

$$Px_0 + A'x_1' = b,\ x_1' \geq 0 \tag{3.17}$$

rendszernek létezik megoldása, úgy az x_1'-t egy nulla komponenssel kiegészítve (3.15) megoldásához jutunk. Ha (3.17)-nak nincs megoldása, úgy az indukciós feltevés miatt az

$$yP = 0,\ yA' \geq 0,\ yb = -1 \tag{3.18}$$

rendszernek létezik y' megoldása. Amennyiben $y'a_1 \geq 0$, úgy y' a (3.16)-nek is megoldása, és ekkor készen vagyunk.

Tegyük fel tehát, hogy $y'a_1 < 0$, azaz $y'A$ első komponense negatív, a többi nemnegatív. Jelölje P' azt a mátrixot, amelyet P-ből az a_1 oszlop hozzávételével nyerünk. Ha most az

$$yP' = 0,\ yA' \geq 0,\ yb = -1 \tag{3.19}$$

problémának van megoldása, az nyilván megoldása (3.16)-nek is, és ekkor ismét csak készen vagyunk. Ha (3.19)-nek nincs megoldása, úgy a

$$P'x_0' + A'x_1' = b,\ x_1' \geq 0 \tag{3.20}$$

rendszernek van (indukció miatt). Jelölje (x_0, x_1) azt a vektort, amely úgy keletkezik (x_0', x_1')-ből, hogy az a_1-nak megfelelő komponenst x_0'-ből x_1'-be helyezzük (ami persze azt jelenti, hogy (x_0', x_1') ugyanazt az $n_0 + n_1$ dimenziós vektort jelöli, mint (x_0, x_1)).

Állítjuk, hogy (x_0, x_1) megoldása (3.15)-nek. Ehhez csak azt kell igazolnunk, hogy $x_1(1)$ (az a_1-nek megfelelő komponens) nemnegatív. Valóban, ha ez negatív lenne, akkor x_1 is és $y'A$ is olyan, hogy első komponensük negatív, a többi pedig nem az. Emiatt $(y'A)x_1 > 0$ és így $0+0 < 0+(y'A)x_1 = (y'P)x_0 + (y'A)x_1 = y'[(P,A)x] = y'b = -1$, ami lehetetlen. •

Megmutatjuk, hogy a 3.5.7 tételből is levezethető a 3.5.5 általános alak.

Biz. (3.5.5 tételé) Jelölje az x_0, x_1, y_0, y_1 dimenzióját rendre n_0, n_1, m_0, m_1. Legyen $B' := (B, I), A' := (A, 0)$ (ahol az I egy $m_1 \times m_1$-os egység-mátrixot, a 0 pedig egy $m_0 \times m_1$-os nulla mátrixot jelöl). Most (3.13) (x_0, x_1) megoldásai és

$$Px_0 + A'x_1' = b_0,\ Qx_0 + B'x_1' = b_1,\ x_1' \geq 0 \tag{3.21}$$

(x_0, x_1') megoldásai között egy-egy értelmű kapcsolat áll fenn. Nevezetesen (x_0, x_1) az (x_0, x_1')-ből keletkezik az utolsó m_1 komponens kihagyásával, míg (x_0, x_1)-ből úgy kapjuk (x_0, x_1')-t, hogy x_1-t helyettesítjük az $x_1' := (x_1, b_1 - (Qx_0 + Bx_1))$ vektorral.

A (3.21)-hez tartozó

$$y_0P + y_1Q = 0,\ y_0A' + y_1B' \geq 0,\ yb = -1 \tag{3.22}$$

duál probléma és (3.14) ekvivalensek, hiszen csak arról van szó, hogy a (3.14)-ban explicit szereplő $y_1 \geq 0$ előjel megkötést (3.22)-ben az $y_0A' + y_1B' \geq 0$ egyenlőtlenségrendszerben rejtettük el. A (3.21) primál feladatra felírva a 3.5.7 tételt az abban szereplő (3.16) duál feladat éppen a (3.22) alakot ölti. •

További kézenfekvő megjegyzés, hogy a Farkas lemmát olyan alakban is használhatjuk, amikor a rendszer balról szorzással van adva. (Egy későbbi alkalmazás miatt néhány vektort és mátrixot vesszős betűvel jelölünk.)

3.5.8. Tétel.

$$y_0P' + y_1Q' = c_0',\ y_0A' + y_1B' \geq c_1',\ y_1 \geq 0 \tag{3.23}$$

primál rendszernek akkor és csak akkor nincs megoldása, ha a

$$P'x_0 + A'x_1' = 0,\ Q'x_0 + B'x_1' \leq 0,\ x_1' \geq 0,\ c_0'x_0 + c_1'x_1' > 0 \tag{3.24}$$

duális rendszernek van. •

Feladatok

3.26. *Az $\{yA \leq c\}$ rendszernek akkor és csak akkor van megoldása, ha nincs olyan $x \geq 0$, amelyre $Ax = 0, cx < 0$.*

3.27. *Tekintsük a (3.16) duális feladatban előforduló $\{yP = 0,\ yA \geq 0,\ yb = -1\}$ rendszert, mint primál problémát és fogalmazzuk meg erre a Farkas lemmát. Mutassuk meg, hogy a felírt duális ekvivalens a (3.15) alakkal.*

3.28. *Tegyük fel, van egy szubrutinunk, amely vagy az $\{Ax = b, x \geq 0\}$ rendszernek vagy a duális $\{yA \geq 0,\ yb = -1\}$ rendszernek kiszámít egy megoldását. Hogyan használhatjuk fel ezt a (3.15) vagy (3.16) rendszer megoldására?*

3.29. *Tegyük fel, van egy szubrutinunk a $\{Bx \leq b, x \geq 0\}$ rendszer megoldására. Hogyan használhatjuk fel ezt a $\{Qx \leq b\}$ rendszer megoldására?*

3.5.2. A szimplex algoritmus a Farkas lemmára

Ebben a szakaszban megismerkedünk a Farkas lemmával kapcsolatos fő algoritmikus eredményekkel. Azt már korábban láttuk, hogy a Fourier-Motzkin eljárás segítségével egy R poliédernek véges sok lépésben megtalálhatunk egy elemét, amennyiben R nem üres. Ez az eljárás azonban, szemben a Gauss eliminációval, nem polinomiális futásidejű és a gyakorlati tapasztalatok is kedvezőtlenek. Egy lineáris célfüggvény poliéder feletti optimalizálására is van véges algoritmus, hiszen ha cx felülről korlátos, akkor a maximum erős bázis-megoldáson is felvétetik, és ezekből csak véges sok van. Ezzel a megközelítéssel az a baj, hogy erős bázis-megoldásból igen sok lehet (exponenciálisan sok). Például, ha a poliéder, amely felett a cx célfüggvényt akarjuk maximalizálni, egy n-dimenziós egységkocka kocka, azaz $\{x : 0 \leq x \leq 1\}$, akkor ezt a feladatot mind a 2^n csúcs cx szerinti sorbarendezésével már a nem túlságosan nagy $n = 100$-as méretnél sincs semmilyen esélyünk megoldani, a legjobb számítógépet használva sem, ugyanakkor a problémát ránézésre rögtön meg lehet oldani.

A szimplex algoritmus George Dantzigtól származik. Az a lényege, hogy a poliéder bizonyos csúcsait egyre javuló sorrendben választja ki. Óriási felhalmozott tapasztalat mutatja, hogy a szimplex algoritmus a gyakorlatban hatékony, tipikusan lineáris számú csúcs átvizsgálása után megtalálja az optimumot. Annál szomorúbb, hogy konstruáltak olyan példa sorozatot, ahol a szimplex algoritmus végiglátogatja az összes, exponenciálisan sok csúcsot, mielőtt az optimálisat megtalálná. Ez azt jelenti, hogy matematikai szempontból a szimplex algoritmus nem tekinthető hatékonyabbnak, mint a durva, összes csúcsot számba vevő algoritmus. (Kimutatták azonban, hogy ha nem a legrosszabb előforduló esettel akarjuk az algoritmus hatékonyságát mérni, hanem az átlagos lépésszámot tekintjük, akkor a szimplex algoritmus polinomiális futásidejű.)

Kezdjük a Farkas lemmával, annak a 3.5.1 tételben megfogalmazott erősebb változatával, amely tehát azt mondja ki, hogy az $\{Ax = b, x \geq 0\}$ primál és az $\{yA \geq 0, yb = -1\}$ duál feladatok közül pontosan az egyiknek van bázis-megoldása:

3.5.9. Tétel. *Az $\{Ax = b,\ x \geq 0\}$ rendszernek akkor és csak akkor van olyan megoldása, amelyben az x pozitív változóinak megfelelő A-beli oszlopok lineárisan függetlenek, ha nem létezik olyan y, amelyre $yA \geq 0$, $yb < 0$ és A-nak létezik $r(A,b) - 1$ lineárisan független oszlopa, amelyekre y merőleges.* (Tömören, vagy a primál vagy a duál feladatnak létezik bázis-megoldása). •

Bizonyítás a szimplex algoritmussal. Már a Farkas lemma bizonyításánál láttuk, hogy a két lehetőség kizárja egymást. Azt látjuk be algoritmikusan, hogy legalább az egyik lehetőség fennáll. A Gauss-eliminációval először eldöntjük, hogy az $Ax = b$ rendszernek van-e egyáltalán megoldása. Ha nincs, akkor a Gauss-elimináció egy olyan y vektort szolgáltat, amelyre $yA = 0$ és $yb \neq 0$. (Lásd a Fredholm féle 2.2.10 tételt és bizonyítását.) Itt -1-gyel történő esetleges szorzás után feltehetjük, hogy $yb < 0$ azaz a második alternatívára jutottunk.

Tegyük fel tehát, hogy $Ax = b$ megoldható. Feltehetjük, hogy A sorai lineárisan függetlenek, mert ha nem, akkor az A soraiból kiválasztunk $r(A)$ lineárisan független sort (ezt valójában a Gauss-elimináció már meg is tette), és csak az ezek által alkotott részmátrixszal dolgozunk tovább. Válasszunk ki az A oszlopaiból egy B_1 bázist (ami tehát egy $m \times m$-es nemszinguláris részmátrix). Tekintsük a $B_1 x = b$ egyértelmű megoldását (amit tehát az előbbi Gauss-elimináció meghatározott), és egészítsük ki nullákkal. Így az $Ax = b$ egy x_1 megoldását kapjuk. Ha x_1 nemnegatív, akkor ez az $Ax = b, x \geq 0$ egy bázis-megoldását alkotja.

Tegyük fel most, hogy x_1-nek van negatív komponense. (Az alapalgoritmus itt egy tetszőleges negatív komponenst választ. Példával kimutatható, hogy ilyenkor az algoritmus végtelen ciklusba eshet, ezért ennek elkerülésére indokolt valamilyen megkötést tenni.) Jelölje i_1 a legkisebb indexet, amelyre $x_1(i_1) < 0$. (Ez a Bland féle legkisebb index szabály). Legyen y_1 olyan vektor, amely B minden oszlopára merőleges, kivéve, hogy $y_1 a_{i_1} = 1$. (Az $yB_1 = d$ minden m-dimenziós d-re egyértelműen megoldható.) Most $q = 1$-re

$$y_q b = y_q(Ax_q) = (y_q A)x_q = x_q(i_q) < 0. \tag{3.25}$$

Amennyiben minden a_i-re $y_1 a_i \geq 0$, úgy a duális feladat bázis-megoldását kaptuk.

Tegyük fel tehát, hogy valamely j_1 indexre $y_1 a_{j_1} < 0$ és válasszuk j_1-t a lehető legkisebbnek. (Ismét a legkisebb index szabályt alkalmazzuk). Természetesen ekkor a_{j_1} nincs a B_1 bázisban, és az is látható, hogy B_1-ben a_{i_1}-t a_{j_1}-re cserélve egy másik bázist kapunk, amit jelöljünk B_2-vel. (Valóban, az a_{j_1} vektor nem függhet lineárisan a B_1-nek a_{i_1}-től különböző oszlopaitól, hiszen az y_1 vektor ezen utóbbiak mindegyikére merőleges, míg a_{j_1}-re nem.) Iteráljuk az eljárást most a B_2 bázissal kezdve, ameddig csak lehet.

Igazolnunk kell, hogy az eljárás véges sok lépésben véget ér. Tegyük indirekt fel, hogy nem ez a helyzet. Mivel csak véges sok bázis-megoldás van, lesznek olyan oszlopvektorok, melyek időről időre ki- majd újra bekerülnek a bázisba. Legyen a_h a legnagyobb ilyen indexű. Tehát a h-nál nagyobb indexű oszlopok egy bizonyos időponttól fogva már nem változtatják helyzetüket: vagy egyszer és mindenkorra benn vannak a bázisban, vagy kívül. Legyen egy ezutáni pillanatban B_p egy olyan előforduló bázis, amelyben a_h benne van, de B_{p+1}-ben nincs. Legyen B_q $(q > p)$ egy olyan későbbi bázis, amelyben a_h nincs benne, de B_{q+1}-ben benne van. Ekkor tudjuk, hogy y_q a B_{q+1} minden oszlopára merőleges, kivéve a_h-t, amelyre $y_q a_h < 0$. A második választási szabályból az is következik minden a_h előtti a_i oszlopra (azaz $i < h$-re), hogy $y_q a_i \geq 0$.

Az első választási szabály miatt x_p minden h-nál kisebb indexű komponense nemnegatív. Így tehát az $1 \leq i < h$ indexekre $(y_q a_i) x_p(i) \geq 0$ és $(y_q a_h) x_p(h) > 0$. A B_q mátrixnak egyetlen olyan a_j oszlopa van, amelyre y_q nem merőleges, de mivel éppen ez az oszlop esik ki a B_q bázisból, az a_j szükségképpen megelőzi a_h-t. Tehát y_q merőleges a B_q-nak a_h utáni oszlopaira, de mivel a h-nál nagyobb indexű oszlopokon a B_q és a B_p bázis megegyezik (itt használva h maximális választását), y_q merőleges B_p minden h-nál nagyobb indexű oszlopára. Így felhasználva (3.25)-t, kapjuk, hogy $0 > y_q b = y_q(Ax_p) = (y_q A)x_p \geq 0$, és ez az ellentmondás a tételt és egyúttal az algoritmus végességét is bizonyítja. •

Az algoritmus általánosabb alakban is használható. Például, ha a $\{Px_0 + Ax_1 = b, x_1 \geq 0\}$ rendszer megoldhatóságát akarjuk eldönteni, akkor az előző eljárást a következőképpen kell módosítani. A kezdeti B_1 bázist úgy határozzuk meg, hogy a P maximálisan sok lineárisan független oszlopát egészítjük ki A oszlopaiból $r(P, A)$ darab lineárisan független oszloppá. Az eljárást úgy módosítjuk, hogy csak az A-beli bázis elemeket cserélhetjük, a kezdeti B_1 bázis P-ből kiválasztott elemei végig fixen maradnak. Rögtön látható, hogy az így módosított eljárás új bizonyítást ad a 3.5.7 tételre. A még általánosabb 3.5.5 alakra is kiterjeszthető az algoritmus, ha a tétel bizonyításában leírt visszavezetést alkalmazzuk. Megállapíthatjuk tehát, hogy a fenti eljárás bármely alakban adott lineáris egyenlőtlenség-rendszer megoldására alkalmas.

A név eredete

Miért hívják a fenti algoritmust szimplex algoritmusnak? Az m-dimenziós térben, ha veszünk úgy $m+1$ pontot, hogy ezek egyike sincs benne a többi konvex burkában, akkor az $m+1$ pont konvex burka definíció szerint egy **szimplex**et alkot. Egy dimenzióban ez egy szakasz, két dimenzióban háromszög, három dimenzióban tetraéder.

Tegyük fel például, hogy a síkban adottak a $p_1, \ldots, p_n$ pontok, és el akarjuk dönteni, hogy ezek konvex burkában benne van-e egy megadott b' pont. Egyszerűség kedvéért tegyük fel, hogy nincs három pont egy egyenesen. Készítsünk el egy $2 \times n$-es A' mátrixot, melynek i-edik oszlopa a p_i pont koordinátáit tartalmazza. Legyen A az a mátrix, amely A'-ből keletkezik egy csupa egyesből álló harmadik sor hozzávételével. A b'-t is egészítsük ki egy egyessel egy három dimenziós b vektorrá. Ekkor a feladat azzal ekvivalens, hogy az $\{Ax = b, x \geq 0\}$ rendszernek van-e megoldása. Egy bázis-megoldáshoz tartozó három oszlopvektor három olyan p_i pontnak felel meg, amelyek egy b'-t tartalmazó háromszöget alkotnak. Egy duális bázis-megoldás egy olyan egyenesnek felel meg (miért?), amely két p_i ponton átmegy, és az általa meghatározott egyik zárt félsík tartalmazza az összes p_i pontot, de nem tartalmazza b'-t. Mármost a szimplex algoritmus ezen a geometriai nyelven a következőképpen fut. Induljunk ki egy tetszőleges B_1-gyel jelölt háromszögből, melynek csúcsai mondjuk p_1, p_2, p_3. Állapítsuk, meg, hogy b' benne van-e B_1-ben. Ha benne van akkor készen vagyunk: b' benne van a p_i pontok konvex burkában. Ha b' nincs benne a háromszögben, akkor a háromszögnek az egyik e_1 oldalegyenese, mondjuk p_2p_3, elválasztja b'-t p_1-től. (A háromszögnek egy vagy két ilyen elválasztó oldalegyenese lehet, a b' elhelyezkedésétől függően: az általános algoritmus ezek egyikét választja, a Bland féle szabály pontosan előírja, hogy melyiket kell választani.) Amennyiben az e' egyenes által határolt, a b'-t tartalmazó nyílt félsíkban nincsen p_i pont, akkor készen vagyunk; megkaptuk az elválasztó egyenest. Ha van, mondjuk a p_4 pont, akkor p_1-t becseréljük p_4-re és a keletkező $\{p_2, p_3, p_4\}$ háromszöggel folytatva iteráljuk az eljárást.

A fenti algoritmus szemléletesen tehát azt jelenti, hogy a p_i pontok által alkotott háromszögek segítségével mintegy letapogatjuk a sík egy darabját, és eközben vagy ráakadunk a b' pontra vagy megtalálunk egy elválasztó egyenest. Magasabb dimenzióban ez azt jelenti, hogy a p_i pontjaiból készített szimplexekkel tapogatjuk le a teret, hogy megtaláljuk a b' pontot. Innen tehát az elnevezés.

Ciklizálás

A következő példa mutatja, hogy ha a futás során nem alkalmazzuk a Bland féle szabályt, akkor az algoritmus ciklizálhat. Legyen b' az origó és $n = 6$. Az origó középpontú egységkörön legyen p_1, p_3, p_5 egy egyenlő oldalú háromszög három csúcsa. $i = 1, 2, 3$-ra a p_{2i-1} pontot az origóval összekötő szakasz felező pontját az origó körül az óramutató járásával ellentétesen egy csöppnyit elforgatva kapjuk a p_{2i} pontot. Ha most B_i jelöli a p_i, p_{i+1}, p_{i+2} pontok által alkotott háromszöget (modulo 6 tekintve), akkor a szimplex algoritmus egymás után ezen háromszögeket választhatja (ezt ellenőrizzük le!), amíg vissza nem ér a kiindulási B_1-be.

Feladatok

3.30. *Hol tér el először az előbbi példában a szimplex algoritmus, ha ugyanazzal a B_1 háromszöggel kezdünk és alkalmazzuk a Bland féle legkisebb index szabályt?*

3.31. *Tegyük fel, hogy csak a bázisba bekerülő új oszlop kiválasztásánál alkalmazzuk a legkisebb index szabályt, a bázisból kikerülő oszlop meghatározásánál nem. Ciklizálhat-e ilyenkor az algoritmus vagy már ilyenkor is bizonyíthatóan mindig véges lesz?*

3.32. *Tekintsük a Farkas lemma következő alakját: Az $yA \leq c$ rendszernek pontosan akkor nincs megoldása, ha létezik olyan $x \geq 0$, amelyre $Ax = 0$ és $cx < 0$. Tegyük fel, hogy A sorai lineárisan függetlenek. Igazoljuk, hogy a következő algoritmus véges.* (A bizonyításban vagy a fenti bizonyítás lépéseit imitáljuk, vagy pedig azt mutassuk ki, hogy az alábbi algoritmus nem más, mint a fenti algoritmus adaptációja)*:*

Legyen B_1 az A egy $m \times m$-es nemszinguláris részmátrixa és legyen y_1 az $yB_1 = c_{B_1}$ egyértelmű megoldása. Amennyiben $y_1 A \leq c$, akkor készen vagyunk, megtaláltuk a kívánt y-t. Ha $y_1 A \not\leq c$, úgy legyen a_j az A mátrix legkisebb indexű oszlopa, amelyre $y_1 a_j < c(j)$. Tekintsük a $B_j x = -a_j$ egyértelmű x_1' megoldását, és jelölje x_1 azt a vektort, amely x_1'-ből keletkezik az $x_1(j)$ helyen 1-gyel, a többin pedig 0-val kiegészítve. Ekkor $Ax_1 = 0$ és $cx_1 = c_{B_1}x_1' + c(j) = (y_1 B_1)x_1' + c(j) < (y_1 B_1)x_1' + y_1 a_j = (y_1 A)x_1 = 0$. Így ha $x_1 \geq 0$, akkor x_1 teljesíti a Farkas lemma második alternatíváját. Amennyiben $x_1 \not\geq 0$, úgy legyen i a legkisebb index, amelyre $x_1(i) < 0$, és cseréljük ki a B_1-beli a_i oszlopot az a_j oszlopra. A keletkező B_2 mátrixszal folytatva iteráljuk az eljárást.

3.33. *Terjesszük ki a fenti algoritmust a Farkas lemma következő változatára. Az $\{yP = c_0, yA \leq c_1\}$ rendszernek pontosan akkor nincs megoldása, ha létezik olyan $x = (x_0, x_1)$, amelyre $(P, A)x = 0$, $x_1 \geq 0$ és $cx < 0$.*

3.5.3. Lineáris és logikai következmény

Azt mondjuk, hogy a $cx \leq \gamma$ egyenlőtlenség **logikai következménye** a $Qx \leq b$ egyenlőtlenség rendszernek, ha az utóbbinak van megoldása és minden megoldása kielégíti a $cx \leq \gamma$ egyenlőtlenséget. Ez geometriailag azt jelenti, hogy az $R := \{x : Qx \leq b\}$ poliéder teljesen a zárt $\{x : cx \leq \gamma\}$ féltérben fekszik. Azt mondjuk, hogy a $cx \leq \gamma$ egyenlőtlenség **lineáris következménye** $Qx \leq b$-nek, ha létezik olyan $y \geq 0$, amelyre $yQ = c$ és $yb \leq \gamma$.

3.5.10. Tétel. *Feltéve, hogy R nemüres, a $cx \leq \gamma$ egyenlőtlenség akkor és csak akkor lineáris következménye a $Qx \leq b$ egyenlőtlenség-rendszernek, ha logikai következménye.*

Biz. Tegyük fel először, hogy $cx \leq \gamma$ lineáris következmény, azaz létezik olyan $y \geq 0$, amelyre $yQ = c$ és $yb \leq \gamma$. Ekkor $Qx \leq b$ esetén $cx = (yQ)b = y(Qb) \leq yb \leq \gamma$, azaz $cx \leq \gamma$ valóban logikai következmény.

A fordított irányhoz tegyük fel, hogy $cx \leq \gamma$ logikai következmény. Azt kell kimutatnunk, hogy $cx \leq \gamma$ lineáris következmény, vagyis, hogy létezik olyan $y \geq 0$, amelyre $yQ = c$ és $yb \leq \gamma$. Tegyük fel, nem ez a helyzet, azaz nem létezik olyan $y \geq 0$, amelyre $yQ = c$ és $y(-b) \geq -\gamma$. Ekkor a Farkas lemma (balról szorzós alakja) miatt van olyan (x^*, α) vektor, amelyre

$$\alpha \geq 0, Qx^* - \alpha b \leq 0, cx^* - \alpha\gamma > 0. \tag{3.26}$$

Ha most $\alpha = 0$, akkor ez a

$$Qx^* \leq 0, \ cx^* > 0 \tag{3.27}$$

alakot ölti. Ebből következik, hogy az R poliéder egy z elemére bármilyen pozitív λ esetén $z + \lambda x$ benne van R-ben, ugyanakkor $c(z + \lambda x)$ bármilyen nagy lehet, ha λ nő, vagyis $cx \leq \gamma$ nem volna logikai következmény. Vagyis α-nak pozitívnak kell lennie. Feltehető, hogy $\alpha = 1$, mert α-val végigoszthatunk. Most (3.26) azzal ekvivalens, hogy

$Qx^* \leq b$, $cx^* > \gamma$ vagyis az x^* létezése cáfolja, hogy $cx \leq \gamma$ logikai következmény volna. •

A Farkas lemma különféle változatai közötti átjárásoknál megismert eszközökkel levezethetjük a 3.5.10 tétel kiterjesztését is. Tekintsük a

$$Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \leq b_1,\ x_1 \geq 0 \tag{3.28}$$

egyenlőtlenség-rendszert, melynek R megoldás-halmazáról tegyük fel, hogy nemüres. Legyen $M = \begin{pmatrix} P & A \\ Q & B \end{pmatrix}$. Legyen $c = (c_0, c_1)$ adott vektor. Az (x_0, x_1) vektort néha röviden x-szel jelöljük, és hasonlóképp a $\begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$ vektort b-vel. Azt mondjuk, hogy a

$$c_0x_0 + c_1x_1 \leq \gamma \tag{3.29}$$

egyenlőtlenség **logikai következménye** a (3.28) rendszernek, ha (3.28) minden megoldása teljesíti (3.29)-t. A (3.29) egyenlőtlenség **lineáris következménye** (3.28)-nek, ha létezik olyan $y := (y_0, y_1)$, amelyre

$$y_1 \geq 0,\ y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\quad yb \leq \gamma. \tag{3.30}$$

3.5.11. Tétel. *Feltéve, hogy R nemüres,* (3.29) *akkor és csak akkor lineáris következménye* (3.28)*-nek, ha logikai.*

Biz. Tegyük fel először, hogy $cx \leq \gamma$ lineáris következmény, azaz létezik a szóbanforgó y. Ekkor (3.28) bármely x megoldására $cx = c_0x_0 + c_1x_1 \leq [(y_0, y_1)\begin{pmatrix} P \\ Q \end{pmatrix}]x_0 + [(y_0, y_1)\begin{pmatrix} A \\ B \end{pmatrix}]x_1 = (yM)x = y(Mx) = y_0[Px_0 + Ax_1] + y_1[Qx_0 + Bx_1] \leq y_0b_0 + y_1b_1 = yb \leq \gamma$, azaz $cx \leq \gamma$ valóban logikai következmény.

A fordított irányhoz tegyük fel, hogy $cx \leq \gamma$ logikai következmény. Már beláttuk azt a speciális esetet nézzük, amikor A, B, P mindegyike üres, azaz $R = \{x : Qx \leq b\}$. Tegyük most fel, hogy A és B üres. Az R-t definiáló rendszer a

$$Px = b_0, Qx \leq b_1 \tag{3.31}$$

alakra egyszerűsödik, és ilyenkor a $cx \leq \gamma$ egyenlőtlenség definíció szerint akkor lineáris következmény, ha létezik olyan $y := (y_0, y_1)$, amelyre

$$y_1 \geq 0,\ y_0P + y_1Q = c,\ yb \leq \gamma. \tag{3.32}$$

(3.31) azzal ekvivalens, hogy $Px \leq b_0, -Px \leq -b_0, Qx \leq b_1$. Ennek a rendszernek logikai következménye a $cx \leq \gamma$ egyenlőtlenség, így a 3.5.10 tétel szerint létezik olyan $(y_0', y_0'', y_1) \geq 0$ vektor, amelyre $y_0'P + y_0''(-P) + y_1Q = c$ és $y_0'b_0 + y_0''(-b_0) + y_1b_1 \leq \gamma$, de ekkor $y_0 := y_0' - y_0''$ választással (3.32) teljesül, azaz $cx \leq \gamma$ lineáris következmény (3.31)-nek.

Az általános eset bizonyításához legyen $B^* := \begin{pmatrix} B \\ -I \end{pmatrix}$, $Q^* := \begin{pmatrix} Q \\ 0 \end{pmatrix}$, ahol I egy $n_2 \times n_2$-es egységmátrix, míg 0 egy $n_2 \times n_1$-es nulla-mátrix. Legyen $b_1^* := \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$, ahol 0 most egy n_2 dimenziós 0-vektor. A megelőző eset P és Q helyén rendre P^*-gal és Q^*-gal illetve b_1 helyén b_1^*-gal éppen az általános alakkal ekvivalens. •

3.5.4. Alkalmazások

3.5.12. Tétel. *Ha egy n-változós lineáris egyenlőtlenség-rendszernek nincsen megoldása, akkor van egy legfeljebb $n+1$ egyenlőtlenségből álló részrendszer úgy, hogy már annak sincsen megoldása.*

Biz. Amennyiben a $Qx \leq b$ rendszernek nincsen megoldása, úgy a Farkas lemma szerint létezik olyan $y \geq 0$ vektor, amelyre $yQ = 0, yb = -1$. E duális rendszer tehát megoldható, így létezik y^* bázis-megoldás is, amiből következik, hogy az y^*-nak legfeljebb $n+1$ pozitív komponense van. A Farkas lemma triviális iránya szerint az ezen komponensekhez tartozó egyenlőtlenség-rendszernek sincsen megoldása. •

3.5.13. Tétel (Caratheodory). *Ha a d-dimenziós tér egy z pontja $p \geq d+1$ darab pont konvex kombinációja, akkor ezen pontok között van legfeljebb $d+1$, amelyeknek z konvex kombinációja.*

Biz. Készítsünk el egy mátrixot, amelynek p oszlopa van és az egyes oszlopok a p pont helyvektorait tartalmazzák, majd egészítsük ki a mátrixot még egy csupa egyesből álló sorral. A keletkező $(d+1) \times p$-es mátrixot jelölje A. Az, hogy z előáll a megadott pontok konvex kombinációjaként azt jelenti, hogy az $Ax = (z, 1)$-nek létezik nem-negatív megoldása. De akkor létezik bázis-megoldása is, ami azt jelenti, hogy az előállításban legfeljebb annyi együttható nemnulla, mint ahány sora van az A mátrixnak, vagyis $d+1$. •

3.5.14. Tétel. *Ha R és R' két nemüres poliéder, melyek metszete üres, akkor van őket szigorúan elválasztó $\{x : cx = \alpha\}$ hipersík, azaz $cx < \alpha < cx'$ fennáll az R minden x és az R' minden x' elemére.*

Biz. Legyen $R = \{x : Qx \leq b\}$ és $R' := \{x : Q'x \leq b'\}$. Mivel a metszetük üres, azaz a $\{Qx \leq b, Q'x \leq b'\}$ rendszernek nincsen megoldása, a Farkas lemma szerint, létezik olyan $(y, y') \geq 0$ vektor, amelyre $yQ + y'Q' = 0$ és $yb + y'b' < 0$. Ekkor az yb és $y'b'$ számok egyike biztosan negatív, mondjuk yb. A $c := yQ$ vektor nem lehet nulla, mert akkor (a Farkas lemma triviális iránya miatt) R üres lenne. $yb + y'b' < 0$ miatt van olyan α szám, amelyre $yb < \alpha < -y'b'$. Ekkor $x \in R$-re $cx = (yQ)x = y(Qx) \leq yb < \alpha$ és $x' \in R'$-re $cx' = (yQ)x' = -(y'Q')x' = -y'(Q'x') \geq -y'b' > \alpha$. •

3.5.15. Tétel (Helly). *Az n-dimenziós térben adottak a $C_1, \ldots, C_k$ konvex halmazok, melyek metszete üres. Ekkor ezen halmazok között létezik már legfeljebb $n+1$ olyan is, amelyek metszete üres.*

Biz. Feltehető, hogy $k > n+1$. Tegyük fel indirekt, hogy bármely $n+1$ halmaz metszete nemüres. Kimutatjuk, hogy léteznek $R_i \subseteq C_i$ poliéderek úgy, hogy már ezek közül is bármely $n+1$-nek van közös pontja. E célból legyen S olyan véges halmaz, hogy a C_i-k közül bármely $n+1$-nek van közös pontja S-ben. Mindegyik C_i-re legyen R_i a C_i-be eső S-beli pontok konvex burka. Mivel C_i konvex, így $R_i \subseteq C_i$. A 3.4.3 tétel szerint R_i poliéder.

Ha most veszünk $n+1$ darab C_i halmazt, akkor az ezek metszetében lévő S-beli pontok S definíciója folytán benne vannak a megfelelő $n+1$ darab R_i metszetében is. Ebből adódik, hogy az R_i-ket definiáló egyenlőtlenség-rendszerek egyesítéséből bárhogy véve $n+1$ egyenlőtlenséget, annak van megoldása, és így a 3.5.12 tétel szerint az egész rendszernek is létezik megoldása, vagyis az összes R_i halmaznak van közös pontja, és emiatt persze az összes C_i halmaznak is van, ellentmondásban a tétel feltevésével. •

3.5.16. Tétel (Kirchberger). *Az n dimenziós térben adott k piros és ℓ zöld pont, ahol $k+\ell \geq n+2$. Amennyiben a piros pontokat nem lehet a zöld pontoktól egy hipersíkkal elválasztani, úgy a pontok között létezik legfeljebb $n+2$ olyan, hogy már ezeket sem lehet hipersíkkal elválasztani.*

Biz. Jelölje P és Z azokat a mátrixokat, amelyek oszlopai a piros illetve a zöld pontok helyvektorai. Egészítsük ki a $[P, -Z]$ mátrixot egy $(1, 1, \dots, 1, 0, \dots, 0)$ sorvektorral, amely k darab egyest tartalmaz, valamint egy $(0, \dots, 0, 1, 1, \dots, 1)$ sorvektorral, amely ℓ egyest tartalmaz. A keletkező mátrixot jelölje A. Legyen b az az $(n+2)$-dimenziós vektor, melynek utolsó két komponense 1, míg a többi 0. Az A-nak tehát $n+2$ sora van.

Állítjuk, hogy az $\{Ax = b, x \geq 0\}$ rendszernek van megoldása. Ha ugyanis nincs, akkor a Farkas lemma szerint létezik egy olyan (y, α, β) vektor, amelyre $yP + \alpha e_k \geq 0, -yZ + \beta e_\ell \geq 0$ és $\alpha + \beta < 0$, ahol e_i a csupa 1-esből álló i-dimenziós vektort jelöli. Ez viszont azt jelenti, hogy az $\{x : yx = -\alpha\}$ hipersík elválasztja a piros és a zöld pontokat, ellentétben a feltevéssel. Az $\{Ax = b, x \geq 0\}$ rendszer egy megoldása a piros illetve és a zöld pontok egy-egy konvex kombinációját adja meg, amelyek egyenlők egymással. Emiatt az ebben szereplő piros és zöld pontok nem választhatók el hipersíkkal. Miután létezik bázis-megoldás és ebben legfeljebb $n+2$ nemnulla komponens van, az ezeknek megfelelő pontok sem választhatók el hipersíkkal. •

A geometriai alkalmazások után most következzék egy érdekes eredmény a valószínűségszámítás területéről.

3.5.17. Tétel. *Ha az A $n \times n$-es nemnegatív mátrix minden oszlopában az elemek összege 1, akkor az $\{Ax = x, e_n x = 1, x \geq 0\}$ rendszernek létezik megoldása, ahol $e_n = (1, \dots, 1)$.*

Biz. Legyen $B = A - I$, ahol I jelöli a diagonális egységmátrixot. Azt kell kimutatnunk, hogy a $\{Bx = 0, e_n x = 1, x \geq 0\}$ rendszernek létezik megoldása. Ha nem létezne, úgy a Farkas lemma alapján volna olyan (y, α) vektor, amelyre $yB + \alpha e_n \geq 0$ és $\alpha < 0$, ami azzal ekvivalens, hogy létezik olyan y, amelyre $yB \gg 0$, azaz $yA \gg y$. Jelölje y legnagyobb komponensének értékét μ. A feltételek nyomán $y \ll yA \leq (\mu e_n)A = (\mu, \dots, \mu)$, ellentmondásban μ választásával. •

3.34. Feladat. *A Farkas lemma felhasználásával igazoljuk a következő eredményt.*

3.5.18. Tétel. *Legyen $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbb{R}_+$ kapacitás függvény. Akkor és csak akkor létezik az s_i pontból t_i-be α_i nagyságú folyam $(i = 1, \dots, k)$ úgy, hogy minden élre a rajta átmenő folyamértékek összege legfeljebb az él kapacitása, ha az éleken értelmezett tetszőleges c nemnegatív költségfüggvényre $\sum_{i=1}^{k} \ell_c(i)\alpha_i \leq \sum_{e \in A} c(e)g(e)$, ahol $\ell_c(i)$ jelöli az s_i-ből t_i-be vezető utak minimális c-költségét.*

4. fejezet

Lineáris optimalizálás

4.1. Iránymenti korlátosság

A Farkas lemma megadta egy lineáris egyenlőtlenség-rendszer megoldhatatlanságának, másszóval egy R poliéder ürességének az okát. A következő feladatunk annak eldöntése, hogy valamely c vektorra a cx lineáris célfüggvény korlátos-e (mondjuk) felülről egy nemüres R poliéderen.

Korábban már igazoltuk, hogy mindig létezik erős bázis-megoldás és legfeljebb véges sok van belőlük. Ki fogjuk mutatni, hogy minden olyan c vektorra, amelyre cx az R halmazon felülről korlátos, a $\sup\{cx : x \in R\}$ érték egy erős bázis-megoldáson felvétetik, azaz a maximum létezik. Túl ezen, jellemezzük majd azon c vektorokat, melyekre cx felülről korlátos R-en.

Az ideális az lenne, ha igazolni tudnánk, hogy tetszőleges c vektorra és $z \in R$ megoldásra mindig létezik olyan x^* bázis-megoldás, amelyre $cx^* \geq cz$, vagyis x^* a $\max cx$ célfüggvény szempontjából legalább olyan jó, mint z. Sajnos ez az állítás már egy dimenzióban sem igaz. Tekintsük ugyanis az $x \geq 0$ egyenlőtlenséget (ahol x most egyetlen változót jelöl). Ennek egyetlen bázis-megoldása van, az $x = 0$. Így a $c = 1$ (egydimenziós) vektor esetén az $x = 1$ ponthoz nincs nála jobb bázis-megoldás.

A bajt az okozza, hogy cx nem korlátos felülről az R-en. Emiatt érdemes megvizsgálni, hogy ez miként fordulhat elő. Könnyű megfigyelni, hogy a nem korlátosságnak egyik (és amint később majd kiderül az egyetlen) lehetséges oka, ha létezik olyan q vektor, amelyre $cq > 0, Qq \leq 0$. A poliéder egy ilyen q vektor által meghatározott irányát c-**növelő**nek nevezünk. Ekkor cx valóban nem korlátos felülről R-en, hiszen bármely pozitív λ számra $z + \lambda q$ is eleme R-nek, és $cq > 0$ miatt $c(z + \lambda q)$ bármilyen nagy lehet. A következő lemma tartalma az, hogy ha a c-növelő irányok létezését kizárjuk, akkor az előbbi példával illusztrált baj már nem fordulhat elő.

4.1.1. Lemma. *Legyen z a $Qx \leq b$ egyenlőtlenség-rendszernek egy megoldása, és c egy n-dimenziós vektor. Ha nem létezik olyan q vektor, amelyre $cq > 0, Qq \leq 0$, akkor $Qx \leq b$-nek létezik olyan x^* bázis-megoldása, amelyre $cx^* \geq cz$.*

(Megjegyzés. A lemma megfordítása nem igaz, vagyis előfordulhat, hogy mind q, mind x^* létezik. Ha például $R = \{(z_1, z_2) : -z_2 \leq 0, z_2 \leq 0\}$ a sík vízszintes tengelye, úgy minden megoldás egyúttal bázis-megoldás is, és ezért $x^* := z$ választással $cx^* = cz$ teljesül. Ugyanakkor $z := (0, 0), q = (1, 0), c = (1, 0)$ esetén $cq = 1$ és $Qq = 0$).

Biz. A $Q_z^<$ sorai száma szerinti indukció. Ha ez a szám nulla, úgy z maga bázis-megoldás, tehát jó lesz x^*-nak. Tegyük fel, hogy z nem bázis-megoldás. Ez azt jelenti, hogy $Q_z^<$-nek van olyan sora, amely lineárisan független a $Q_z^=$ soraitól, és emiatt a 2.2.9 tételből adódóan létezik olyan q, amelyre $Q_z^= q = 0, Q_z^< q \neq 0$. Tekintsük az $x_\lambda := z + \lambda q$ vektort ($\lambda \geq 0$).

1. eset $cq = 0$. Feltehető, hogy $Q_z^<$-nek van olyan ${}_iq$ sora, amelyre ${}_iqq > 0$, mert különben q-t a negatívjával helyettesíthetjük. Kicsiny λ-ra x_λ benne van R-ben, míg nagy λ-ra, ${}_iqq > 0$ miatt, nincsen. Így van olyan $\lambda' > 0$ érték, amelyre $x_{\lambda'} \in R$ és $x_{\lambda'}$ több egyenlőtlenséget teljesít egyenlőséggel, mint z. [Nevezetesen λ' a maximális olyan λ érték, amelyre $Q_z^< x_\lambda \leq b_z^<$ teljesül, vagyis $\lambda' = \min((b_z(i) - {}_iqz)/\,{}_iqq :\ {}_iq$ a $Q_z^<$ olyan sora, amelyre ${}_iqq > 0)$.] Miután $Q_{x_{\lambda'}}^<$-nek kevesebb sora van, mint $Q_z^<$-nek, az indukciós feltevést alkalmazhatjuk $x_{\lambda'}$-re. Így létezik egy olyan x^* bázis-megoldás, amelyre $cx^* \geq cx_{\lambda'} = cz + c(\lambda' q) = cz$.

2. eset $cq \neq 0$. Feltehető, hogy $cq > 0$, mert ha nem, q-t a negatívjával helyettesítjük. Amennyiben $Q_z^< q \leq 0$, úgy q léte ellentmond a lemma feltevésének. Ha viszont van olyan ${}_iq$ sora $Q_z^<$-nak, amelyre ${}_iqq > 0$, akkor ugyanúgy járunk el, mint az első esetben: indukció alapján létezik olyan x^* bázis-megoldás, amelyre $cx^* \geq cx_{\lambda'} = cz + c(\lambda' q) > cz$. (Az utolsó egyenlőtlenség érdekében kellett q-t úgy választanunk, hogy cq pozitív legyen.) •

Hasznos tudatosítani, hogy a bizonyítás algoritmikus abban az értelemben, hogy a szóbanforgó x^*-t a Gauss-elimináció segítségével polinom időben kiszámíthatjuk. A lemmát $c = 0$-ra alkalmazva visszakapjuk a 3.3.10 tételt bázis-megoldás létezéséről. A fő eredményt először speciális alakban fogalmazzuk meg.

4.1.2. Tétel (Az iránymenti korlátosság tétele, speciális alak)**.** *Tegyük fel, hogy az $R := \{x : Qx \leq b\}$ poliéder nemüres, és legyen c egy n-dimenziós vektor. A következők ekvivalensek.*

(0) *A $\{cx\}$ lineáris függvény R-en felülről korlátos.*
(1) *Minden $z \in R$ elemre létezik $Qx \leq b$-nek olyan x^* erős bázis-megoldása, amelyre $cx^* \geq cz$.*
(2) *Nem létezik olyan q vektor, amelyre $cq > 0$ és $Qq \leq 0$.*
(3) *Létezik olyan $y \geq 0$ vektor, amelyre $yQ = c$.*

Biz. Először is figyeljük meg, hogy a (2) és (3) feltételek ekvivalenciája nem más, mint a balról szorzással felírt Farkas lemma standard alakja.

Mivel a 3.3.12 következmény alapján csak véges sok erős bázis-megoldás van, (1) implikálja (0)-t. (0)-ból rögtön következik (2). Igazoljuk most a (2) $\rightarrow$ (1) irányt. A 4.1.1 lemmából tudjuk, hogy létezik olyan x^* bázis-megoldás, amelyre $cx^* \geq cz$. Válasszuk x^*-t olyannak, amelynek maximálisan sok 0 komponense van. Állítjuk, hogy x^* erős bázis-megoldás. Tegyük fel indirekt, hogy ez nem igaz, vagyis az x^* nemnulla komponenseinek megfelelő Q-beli oszlopok lineárisan összefüggőek. Ez azt jelenti, hogy létezik egy olyan $q \neq 0$ vektor, amelyre $Qq = 0$ (vagyis q eltolási vektor) és $x^*(i) = 0$ esetén $q(i) = 0$. Feltehetjük, hogy $cq \geq 0$, különben q-t a minusz egyszeresével helyettesítjük. A (2)-ből következik, hogy valójában $cq = 0$. Most alkalmas λ-ra $x^*_\lambda := x^* + \lambda q$-nak több nulla komponense lesz, mint x^*-nak, továbbá x^*_λ is bázis-megoldás, amely $cx^* = cx^*_\lambda$ miatt ellentmond x^* választásának. •

4.1.3. Következmény. *Ha az* $R = \{x : Qx \leq b\}$ *poliéder nemüres, és* $\{cx : x \in R\}$ *felülről korlátos, úgy* $\max\{cx : x \in R\}$ *létezik* (azaz létezik olyan $z \in R$, amelyre $cz = \sup\{cx : x \in R\}$).

Biz. A 3.3.12 tétel szerint véges sok erős bázis-megoldás van. A 4.1.2 tételből adódóan a maximum ezek egyikén felvétetik. •

Megjegyzés A 4.1.2 tétel három jellemzést is ad $\{cx : x \in R\}$ felülről való korlátosságára. Az első tartalma az, hogy a maximum felvétetik (véges sok erős bázis-megoldás van). A második könnyen ellenőrizhető okot mutat a nemkorlátosságra ($Qq \leq 0$ miatt $z_\lambda = z + \lambda q \in R$, így $cq > 0$ miatt cz_λ bármilyen nagy lehet.) Végül a harmadik jellemzés könnyen ellenőrizhető okot mutat a korlátosságra ($y \geq 0, yQ = c$ esetén minden $x \in R$-re $cx = (yQ)x = y(Qx) \leq yb$, magyarán az yb érték egy konkrét felső korlát.)

4.1.4. Következmény. *Ha egy egyenlőtlenség-rendszer megoldható, akkor van erős bázis-megoldása is.*

Biz. A 4.1.2 tételben $c = 0$-ra (0) fennáll, így a tétel alapján (1) is. •

Az irodalomban gyakran Caratheodory tételnek hívják a 4.1.4 következménynek az $\{Ax = b, x \geq 0\}$ alakra vonatkozó speciális esetét, ami szerint, ha van megoldás, akkor van olyan is, amelynek a nem-nulla komponenseihez tartozó A-beli oszlopok lineárisan függetlenek.

Alkalmazhatjuk a Farkas lemma balról szorzással felírt általános alakját (3.5.8 tétel) a 4.1.2 tétel kiterjesztésére arra az esetre, amikor az R poliéder a

$$Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \leq b_1,\ x_1 \geq 0 \tag{4.1}$$

egyenlőtlenség-rendszer megoldás-halmazát jelöli.

4.1.5. Tétel (Az iránymenti korlátosság tétele)**.** *Tegyük fel, hogy az R poliéder nemüres, és legyen $c = (c_0, c_1)$ adott vektor. A következők ekvivalensek.*

(0) *A $\{cx\}$ lineáris függvény R-en felülről korlátos.*

(1) *Minden $z \in R$ elemre létezik* (4.1)*-nek olyan x^* erős bázis-megoldása, amelyre $cx^* \geq cz$.*

(2) *Nem létezik olyan $q = (q_0, q_1)$ vektor, amelyre $cq > 0$, és $q_1 \geq 0, Pq_0 + Aq_1 = 0, Qq_0 + Bq_1 \leq 0$.*

(3) *Létezik olyan $y = (y_0, y_1)$ vektor, amelyre*

$$y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\ y_1 \geq 0. \bullet \tag{4.2}$$

4.2. Optimalitás

Korábban megvizsgáltuk, hogy egy R poliéder mikor nemüres, majd azt, hogy egy cx lineáris célfüggvény mikor korlátos felülről R-en. Most rátérünk a lineáris programozás fő kérdésének tárgyalására: amennyiben R nemüres és $\{cx : x \in R\}$ korlátos felülről, hogyan jellemezhetjük a cx-et maximalizáló pontokat és a maximum értékét. Röviden, maximalizáljuk cx-t az R poliéder felett:

$$\max\{cx : x \in R\}. \tag{4.3}$$

A (4.3) feladatot **lineáris program**nak nevezzük. Természetesen a poliéder lehet más alakban is megadva, szorozhatunk balról, és maximalizálás helyett szerepelhet minimalizálás (lásd a (4.19) és a (4.20) alakokat). Miután R nemüres és cx felülről korlátos R-en, a 4.1.2 tétel alapján jogos (4.3)-ben maximumról beszélni.

Geometriailag egy lineáris program azt jelenti, hogy a c vektor irányában keressük az R legtávolabbi pontját, vagyis azt a pontot, amelyben egy c normálisú hipersík, ha kívülről a poliéderhez toljuk, azt megérinti. Speciális eset, amikor a c egy egységvektor (például $c = (0, 0, \dots, 0, 1)$, ekkor a lineáris programozás feladata úgy interpretálható, hogy egy poliédernek a legmagasabb pontját kell megkeresni. Ez igen egyszerűnek látszik, ráadásul az általános c esete egyszerű fogással ilyen alakra hozható. Mégsem ismert olyan hatékony (polinomiális futásidejű) eljárás, amely a Gauss-eliminációhoz hasonló egyszerű lépésekből áll. (Az olyan ismert polinomiális algoritmusok, mint az ellipszoid módszer vagy az ún. belső pontos módszerek bonyolultabb apparátust igényelnek.) Egy egyenlőtlenség-rendszer megoldására szolgáló Fourier-Motzkin eljárás ilyen egyszerű lépésekből áll, és könnyen módosítható is egy lineáris program megoldására, de nem hatékony. A szimplex algoritmus optimalizálós változatával a következő részben fogunk megismerkedni. Ez az FM eljáráshoz hasonlít abban, hogy egyszerű lépésekből áll és matematikai értelemben nem hatékony. A gyakorlatban ugyanakkor igen jól használható.

4.2.1. Optimalitási feltételek

Egy lineáris programmal kapcsolatban fontos kérdés, hogy létezik-e olyan egyszerűen ellenőrizhető eszköz, amelynek segítségével a poliéder egy megadott x^* elemének optimalitásáról gyorsan meggyőződhetünk. Amennyiben x^* nem optimális, egy olyan eszköz is kívánatos, amelynek segítségével az x^*-nál a poliédernek egy jobb eleméhez tudunk hozzájutni (x jobb: $cx > cx^*$).

Legyen x^* az $R = \{x : Qx \leq b\}$ egy adott eleme. Azt mondjuk, hogy egy $\vec{x}'$ irány x^*-nál **lehetséges elmozdulás**, ha van olyan (kicsiny) pozitív λ szám, amelyre $x^* + \lambda x' \in R$. Ha ráadásul $cx' > 0$, akkor $\vec{x}'$-t **növelő irány**nak hívjuk (c-re és x^*-re nézve). Egyszerű megfigyelni, hogy $\vec{x}'$ pontosan akkor lehetséges elmozdulás, ha $Q^=_{x^*} x' \leq 0$.

4.2.1. Tétel. *Tegyük fel, hogy az $R := \{x : Qx \leq b\}$ poliéder nemüres és $\{cx : x \in R\}$ felülről korlátos. Az R egy megadott x^* elemére a következő állítások ekvivalensek.*

(1) *$cx^* \geq cx$ minden $x \in R$-re, azaz x^* maximalizálja a cx függvényt az R-en* (röviden, x^* optimális).

(2) *Nem létezik c-növelő irány, azaz olyan x' vektor, amelyre $Q^=_{x^*} x' \leq 0$ és $cx' > 0$.*

(3) *A c vektor benne van x^* aktív sorainak kúpjában. Más szóval, van olyan y^* vektor, amely kielégíti az*

$$y^* \geq 0, y^* Q = c \tag{4.4}$$

duális feltételt, és amelyre fennáll az

$$y^*(i) > 0 \Rightarrow {}_iqx^* = b(i) \tag{4.5}$$

optimalitási kritérium. (4.4) *fennállása esetén* (4.5) *azzal ekvivalens, hogy*

$$cx^* = by^*, \tag{4.6}$$

továbbá azzal, hogy

$$y^*(b - Qx^*) = 0. \tag{4.7}$$

Biz. (1) $\Rightarrow$ (2) Ha létezik a szóbanforgó x', akkor kicsiny pozitív λ-ra az $x^* + \lambda x'$ vektor R-ben van, ami $cx' > 0$ miatt ellentmond cx^* maximalitásának.

(2) $\Rightarrow$ (3) Ha nem létezik a szóbanforgó x', akkor a Farkas lemma (balról szorzós alakja) miatt van olyan $y' \geq 0$, amelyre $y'Q^=_{x^*} = c$, így y'-t nulla komponensekkel kiegészítve egy (4.5)-t kielégítő y^*-t kapunk.

(3) $\Rightarrow$ (1) Tetszőleges $x \in R$ esetén

$$cx = (y^*Q)x = y^*(Qx) \leq y^*b, \tag{4.8}$$

vagyis az y^*b érték felső korlát $\{cx : x \in R\}$-re. Ebből adódik, hogy egy $x^* \in R$ elem bizonyosan optimális, ha (4.8)-t egyenlőséggel teljesíti. Másrészt (4.5),(4.6), (4.7) mindegyike azzal ekvivalens, hogy x^* egyenlőséggel teljesíti (4.8). •

Megjegyzés A (4.5) optimalitási kritérium szavakban kifejezve azt mondja, hogy az y^* bármely komponense csak akkor lehet pozitív, ha a neki megfelelő primál egyenlőtlenséget az x^* egyenlőséggel teljesíti.

Az előbbi bizonyítás lépéseinek a másolásával kiterjeszthetjük a tételt az általános alakra.

4.1. Gyakorlat. *Igazoljuk, hogy egy általános* $R = \{(x_0, x_1) : x_1 \geq 0, Px_0 + Ax_1 = b_0, Qx_0 + Bx_1 \leq b_1\}$ *alakban megadott poliéder* $x^* = (x^*_0, x^*_1)$ *elemére az* $x' = (x'_0, x'_1)$ *vektor* $\vec{x}'$ *iránya pontosan akkor lehetséges elmozdulás, ha* $Px'_0 + Ax'_1 = 0$ *és* $Qx'_0 + Bx'_1 \leq 0$, *és* $x^*_1(i) = 0$ *esetén* $x'_1(i) \geq 0$.

4.2.2. Tétel. *Tegyük fel, hogy a*

$$Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \leq b_1,\ x_1 \geq 0 \tag{4.9}$$

rendszerrel definiált R *poliéder nemüres és* $\{cx = c_0x_0 + c_1x_1 : x \in R\}$ *felülről korlátos. Legyen* $x^* = (x^*_0, x^*_1)$ *az* R *egy eleme, és jelölje* $(Q^=_{x^*}, B^=_{x^*})$ *a* (Q, B) *mátrix azon sorai által alkotott részmátrixot, amelyekre a hozzájuk tartozó egyenlőtlenségeket* x^* *egyenlőséggel teljesíti, míg* $b^=_{1*}$ *jelölje a* b_1 *megfelelő részét. A következő állítások ekvivalensek.*

(1) $cx^* \geq cx$ *minden* $x \in R$, *azaz* x^* *maximalizálja a* cx *függvényt az* R*-en (röviden,* x^* *optimális).*
(2) *Nem létezik* c*-növelő irány, azaz olyan* $x' = (x'_0, x'_1)$ *vektor, amelyre* $cx' > 0$,

$$Px'_0 + Ax'_1 = 0, Q^=_{x^*}x'_0 + B^=_{x^*}x'_1 \leq 0 \tag{4.10}$$

és

$$x_1^*(i) = 0 \Rightarrow x_1'(i) \geq 0. \tag{4.11}$$

(3) *Létezik olyan* $y^* = (y_0^*, y_1^*)$ *vektor, amely kielégíti az*

$$y_1^* \geq 0, y_0^* P + y_1^* Q = c_0,\ \ y_0^* A + y_1^* B \geq c_1 \tag{4.12}$$

duális feltételt, és amelyre fennáll az

$$x_1^*(j) > 0 \Rightarrow y_0^* a_j + y_1^* b_j = c_1(j) \tag{4.13}$$

valamint az

$$y_1^*(i) > 0 \Rightarrow\ \ {}_i q x_0^* +\ {}_i b x_1 = b_1(i), \tag{4.14}$$

optimalitási kritérium. (4.12) *fennállása esetén az optimalitási kritérium azzal ekvivalens, hogy*

$$cx^* = by^*, \tag{4.15}$$

és azzal, hogy

$$y^*(b - Mx^*) = 0, \tag{4.16}$$

ahol $M = \begin{pmatrix} P & A \\ Q & B \end{pmatrix}$.

Biz. (1) $\Rightarrow$ (2) Ha létezik a szóbanforgó x', akkor kicsiny pozitív λ-ra a $x^* + \lambda x'$ vektor R-ben van, ami $cx' > 0$ miatt ellentmond x^* maximalitásának.

(2) $\Rightarrow$ (3) Jelölje P' a (P, A) mátrix azon részmátrixát, amely a P-ből és azon A-beli a_i oszlopokból áll, amelyekre $x_1^*(i) > 0$, és legyen A' az A maradék része. (Tehát $(P', A') = (P, A)$). Álljon Q' a $Q_{x^*}^=$ mátrixból kiegészítve a $B_{x^*}^=$ azon oszlopaival, amelyekre az x_1^* megfelelő komponensei pozitívak, és legyen B' a $B_{x^*}^=$ maradék része. (Tehát $(Q', B') = (Q_{x^*}^=, B_{x^*}^=)$). Analóg módon definiáljuk (c_0', c_1')-t.

Ha (2) szerint nem létezik a szóbanforgó x', akkor a Farkas lemma 3.5.8 tételben megfogalmazott alakja szerint van olyan (y_0, y_1), amelyre $y_1 \geq 0$, $y_0 P + y_1 Q_{x^*}^= = c_0', y_0 A + y_1 B_{x^*}^= \geq c_1'$. Legyen $y_0^* := y_0$, és legyen y_1^* az a vektor, amelyet y_1-ből kapunk nulla komponensek hozzávételével (éspedig annyival, ahány sora $Q_{x^*}^<$-nek van). Az így kapott (y_0^*, y_1^*) vektor kielégíti a duál feltételeket és az optimalitási kritériumot.

(3) $\Rightarrow$ (1) Tetszőleges $x \in R$ esetén

$$cx = c_0 x_0 + c_1 x_1 \leq [(y_0^*, y_1^*) \begin{pmatrix} P \\ Q \end{pmatrix}] x_0 + [(y_0^*, y_1^*) \begin{pmatrix} A \\ B \end{pmatrix}] x_1 = (y^* M) x = \tag{4.17}$$

$$= y^*(Mx) = y_0^*[Px_0 + Ax_1] + y_1^*[Qx_0 + Bx_1] \leq y_0^* b_0 + y_1^* b_1 = y^* b, \tag{4.18}$$

vagyis az $y^* b$ érték felső korlát $\{cx : x \in R\}$-re. Ebből adódik, hogy az $x^* \in R$ elem bizonyosan optimális, ha (4.17) és (4.18) mindegyike egyenlőséggel teljesül. Az első azt jelenti, hogy $c_1 x_1 = (y_0^*, y_1^*) \begin{pmatrix} A \\ B \end{pmatrix} x_1^*$, ami pontosan akkor áll fenn, ha $x_1^*(j) > 0$ esetén $y_0^* a_i + y_1^* b_i = c_1(i)$, azaz (4.13) teljesül. Az x^* akkor teljesíti (4.18)-t egyenlőséggel, ha $y_1^*[Qx_0^* + Bx_1^*] = y_1^* b_1$, ami pontosan akkor áll fenn, ha $y_1^*(i) > 0$ esetén ${}_i q x_0^* +_i b x_1^* = b_1(i)$, azaz (4.14) teljesül.

Másrészt (4.15), (4.16) mindegyike azzal ekvivalens, hogy x^* mind (4.17)-t, mind (4.18)-t egyenlőséggel teljesíti. •

Megjegyzés Az optimalitási kritérium szavakkal kifejezve azt jelenti, hogy egy előjelkötött $x_1^*(j)$ primál változó vagy $y_1^*(i)$ duál változó csak akkor lehet pozitív, ha a neki megfelelő duál vagy primál egyenlőtlenség egyenlőséggel teljesül.

Megjegyezzük, még hogy a 4.2.2 tétel bizonyítására alternatív lehetőség a 4.2.1 tételből a korábban már megismert átalakításokkal jutni az általános alakra.

4.2. Gyakorlat. *Írjuk fel az optimalitási feltételeket a* $\max\{cx : Ax = b, x \geq 0\}$, $\max\{cx : Bx \leq b, x \geq 0\}$ *lineáris programokra.*

4.2.2. A dualitás tétel

A korlátossági tételben láttuk, hogy ha cx felülről korlátos az $R = \{x : Qx \leq b\}$ nemüres poliéderen, akkor tetszőleges olyan y vektorra, amelyre $y \geq 0, yQ = c$ a by érték felső korlát $\{cx : x \in R\}$ maximumára. A legjobb (ilyen típusú) felső korlátot ezen by értékek minimuma jelenti. Érdekes, hogy a legkisebb felső korlát meghatározásának feladata, vagyis a $\min\{by : y \geq 0, yQ = c\}$ problémája is egy (balról szorzással felírt) lineáris program, amit **duális program**nak hívunk, megkülönböztetendő a $\max\{cx : Qx \leq b\}$ **primál program**tól. A kérdésre, hogy az így kapott legjobb korlát vajon mindig elérhető-e, másszóval hogy a primál optimum és a duál optimum értéke mindig megegyezik-e, a dualitás tétel adja meg a választ.

4.2.3. Tétel (Dualitás tétel, speciális alak)**.** *Tegyük fel, hogy az* $R = \{x : Qx \leq b\}$ *primál poliéder nemüres. Tegyük fel továbbá, hogy* $\{cx : x \in R\}$ *felülről korlátos* (ami a 4.1.2 tétel szerint azzal ekvivalens, hogy a duális $R^* = \{y : y \geq 0, yQ = c\}$ poliéder nemüres)*. Ekkor a primál optimalizálási feladatban a maximum egyenlő a duál feladatban szereplő minimummal, azaz* $\max\{cx : Qx \leq b\} = \min\{by : y \geq 0, yQ = c\}$.

Biz. Ha $x \in R$ és $y \in R^*$, akkor $cx = (yQ)x = y(Qx) \leq yb$, és így $\max \leq \min$ következik. Az egyenlőség igazolásához egy olyan $x^* \in R$ és $y^* \in R^*$ primál és duál megoldás-párt kell találnunk, amelyekre $cx^* = by^*$. A 4.1.3 következmény szerint, ha $\{cx : x \in R\}$ felülről korlátos, akkor a maximum egy x^* erős bázis-megoldáson felvétetik. A 4.2.1 tétel szerint létezik olyan $y^* \in R^*$ vektor, amelyre $y^*(b - Qx^*) = 0$, amiből $y^*b = cx^*$ következik. •

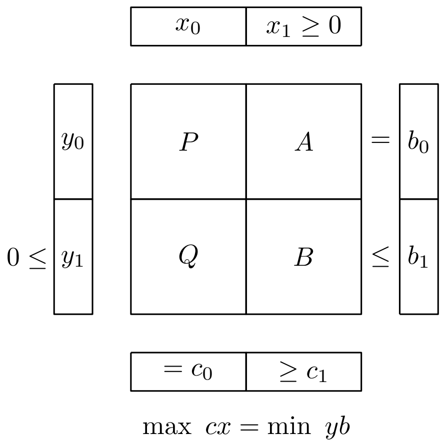
A dualitás tételt is megfogalmazhatjuk az általános alakra. A primál probléma a következő:

$$\max\{(c_0x_0 + c_1x_1) : Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \leq b_1,\ x_1 \geq 0\}. \tag{4.19}$$

A primál problémához hozzárendelt duális lineáris program a következő:

$$\min\{(b_0y_0 + b_1y_1) : y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\ y_1 \geq 0\}. \tag{4.20}$$

A (4.20)-ban szereplő poliédert R^*-gal jelöljük és **duális** poliédernek hívjuk. (Figyelem: R^* az $m := m_1 + m_2$ dimenziós térben van, míg R az $n := n_1 + n_2$ dimenziósban. Az R^* nem csak az R primál poliédertől függ, hanem a c-től is, sőt az R megadásától is!)



4.1. ábra. A dualitás tétel általános alakja

4.2.4. Tétel (Dualitás tétel, általános alak)**.** *Tegyük fel, hogy a (4.19) rendszer által definiált R primál poliéder nemüres. Tegyük fel továbbá, hogy a $cx = c_0x_0 + c_1x_1$ célfüggvényre nézve $\{cx : x \in R\}$ felülről korlátos* (vagy ekvivalensen a duális R^* poliéder nemüres)*. Ekkor a (4.19) primál optimalizálási feladatban a maximum egyenlő a (4.20) duál feladatban szereplő minimummal.*

A speciális alakhoz hasonlóan, a tétel a 4.2.2 tételből közvetlenül adódik.

A megelőző szakaszban megmutattuk, hogy a dualitás tétel miképp vezethető le a Farkas lemmából és abból a tételből, hogy a maximum (erős bázis-megoldáson) felvétetik. A lineáris és logikai következményre vonatkozó 3.5.11 tétel bizonyítása csak a Farkas lemmára támaszkodott. Most megmutatjuk, hogy a dualitás tétel következő szimmetrikus alakja könnyen levezethető a 3.5.11 tételből is.

4.2.5. Tétel (Dualitás tétel, szimmetrikus alak)**.** *Tegyük fel, hogy mind az $R := \{x : Bx \leq b, x \geq 0\}$ primál, mind az $R^* := \{y : yB \geq c, y \geq 0\}$ duál poliéder nemüres. Ekkor $cx \leq by$ fennáll minden $x \in R, y \in R^*$ esetén, és van olyan $x^* \in R, y^* \in R^*$, melyekre egyenlőség érvényes, azaz* $\max\{cx : x \in R\} = \min\{by : y \in R^*\}$.

Biz. Az x és y nem-negativitása miatt $x \in R, y \in R^*$ esetén $cx \leq (yB)x = y(Bx) \leq yb$, így mindenesetre cx felülről korlátos R-en, by pedig alulról R^*-on. Legyen $\gamma_s := \sup\{cx : x \in R\}$ és $\gamma_i := \inf\{by : y \in R^*\}$. Ekkor tetszőleges $x \in R, y \in R^*$ esetén $cx \leq \gamma_s \leq \gamma_i \leq by$. A tételhez azt kell belátnunk, hogy létezik $x^* \in R$, amelyre $cx^* = \gamma_i$ és létezik $y^* \in R^*$, amelyre $by^* = \gamma_s$. Szimmetria miatt elég y^* létezését belátnunk, x^*-é analóg módon következik.

Most tehát a $cx \leq \gamma_s$ egyenlőtlenség logikai következménye a $Bx \leq b, x \geq 0$ egyenlőtlenség-rendszernek, így a 3.5.11 tétel szerint létezik olyan $y^* \geq 0$, amelyre $y^*B \geq c$ és $y^*b \leq \gamma_s$. De itt nem szerepelhet szigorú egyenlőtlenség, mert akkor $y^*b < \gamma_s \leq \gamma_i$ ellentmondana γ_i definíciójának. Tehát valóban $y^*b = \gamma_s$. •

Megjegyzendő, hogy megfordítva, a 3.5.11 tétel is közvetlenül adódik a dualitás tételből. Nézzük ehhez a technikailag legegyszerűbb $Qx \leq b$ esetet, és tegyük fel, hogy a $cx \leq \gamma$ logikai következmény. Ez azt jelenti, hogy $\max\{cx : Qx \leq b\} \leq \gamma$, így a dualitás tétel miatt $\gamma \geq \max\{cx : Qx \leq b\} = \min\{yb : y \geq 0, yQ = c\}$. Vagyis létezik olyan $y \geq 0$, amelyre $yQ = c$ és $\gamma \geq yb$.

Feladatok

4.3. *Irjuk fel a* $\min\{\alpha : Ax - \alpha b = b, (x, \alpha) \geq 0\}$ *lineáris program duálisát, igazoljuk mind a primál, mind a duál rendszer megoldhatóságát, és a dualitás tételből vezessük le a Farkas lemmát* (3.4.5 tétel).

4.4. *Tegyük fel, hogy az* $\{x : Ax = b, x \geq 0\}$ *poliéder nemüres. Az A egy a_i oszlopát* **érdektelen***nek mondjuk, ha a poliéder minden x elemére $x(i) = 0$. Igazoljuk, hogy a_i akkor és csak akkor érdektelen, ha van olyan y vektor, amelyre $yb = 0$, $yA \geq 0$ és ya_i pozitív.*

4.5. *Tekintsük a* $\min\{cx : Ax = b, x \geq 0\}$ *lineáris programot. Legyen A' az A érdekes (azaz nem érdektelen) oszlopai által alkotott részmátrix és jelölje c' a c megfelelő részét. Igazoljuk, hogy cx akkor és csak akkor konstans R-en, ha létezik olyan y, amelyre $yA' = c'$.*

4.6. *Az* $R := \{Ax = b, x \geq 0\}$ *poliéder valamely x_0 eleme akkor és csak akkor minimalizálja cx-t R-en, ha létezik egy olyan c-vel ekvivalens nemnegatív vektor, amely merőleges x_0-ra.* A c_1 és c_2 vektor ekvivalens, ha $c_1x = c_2x$ minden $x \in R$-re.

4.7. *Tekintsük az* $R := \{x : Ax \leq b, x \geq 0\}$ *primál és* $R^* = \{y : yA \geq c, y \geq 0\}$ *duál poliédereken definiált* $\max\{cx : x \in R\}$ *és* $\min\{by : y \in R^*\}$ *primál-duál lineáris program párt, és tegyük fel, hogy R és R^* nem üres. Igazoljuk, hogy az A-nak van olyan A' nemszinguláris négyzetes részmátrixa* (mindegy milyen méretű)*, amelyre az $A'x' = b'$ egyértelmű x' megoldásából nullák hozzávételével keletkező x_1 eleme R-nek* (ahol b' azon része b-nek, amely az A' sorainak felel meg) *továbbá az $y'A' = c'$ egyértelmű y' megoldásából nullák hozzávételével keletkező y_1 eleme R^*-nak. Mutassuk meg, hogy ekkor x_1 primál optimum, y_1 duál optimum.*

4.2.3. Következmények

A játékelméletben fontos alkalmazásra lel a következő tétel. Egy vektor **tetején** értsük a legnagyobb komponensének az értékét. A vektor **alja** legyen a legkisebb komponensének az értéke.

4.2.6. Tétel (Neumann). *Tetszőleges $m \times n$-es $(m, n \geq 1)$ A mátrixra az A oszlopvektorai által feszített politopban lévő elemek tetejének a minimuma egyenlő az A sorvektorai által feszített politopban lévő elemek aljának maximumával. Formálisabban,* $\min\{(\max Ax) : x \geq 0, e_nx = 1\} = \max\{(\min yA) : y \geq 0, e_my = 1\}$, *ahol e_i az i-dimenziós csupa egyesből álló vektort jelenti.*

Biz. A primál feladat egy olyan minimális w szám keresésével ekvivalens, amelyre létezik $x \geq 0$ vektor úgy, hogy $e_nx = 1$ és $Ax \leq (w, w, \ldots, w)$ érvényes. Ez viszont éppen a

$$\min\{w : -Ax + (w, w, \ldots, w) \geq 0, x \geq 0, e_nx = 1\} \tag{4.21}$$

lineáris programmal egyenértékű.

A duális feladat egy olyan maximális z szám keresésével ekvivalens, amelyre létezik $y \geq 0$ vektor úgy, hogy $e_m y = 1$ és $yA \geq (z, z, \ldots, z)$ érvényes. Ez viszont éppen a

$$\max\{z : y(-A) + (z, z, \ldots, z) \leq 0, y \geq 0, e_m y = 1\} \tag{4.22}$$

lineáris programmal egyenértékű. Miután a (4.22) program duálisa éppen a (4.21) program, így a dualitás tételből adódik, hogy a w minimális értéke egyenlő a z maximális értékével. •

4.2.7. Tétel (Clark). *Tekintsük a* $\max\{cx : x \geq 0, Bx \leq b\}$ *és* $\min\{by : y \geq 0, yB \geq c\}$ *primál-duál program párt, és tegyük fel, hogy mindegyik megoldható. Ekkor az R primál és az R^* duál poliéderek közül az egyik nem korlátos.*

Biz. Amennyiben a $\{Bx \leq 0, x \geq 0, -1x \leq -1\}$ rendszernek létezik egy x' megoldása, akkor bármely $x \in R$ vektorra $x + \lambda x'$ minden pozitív λ-ra R-ben van, és mivel $x' \neq 0$, így R nemkorlátos. Ha a kérdéses x' nem létezik, akkor a Farkas lemma szerint van olyan $y' \geq 0$ vektor és $\alpha \geq 0$ szám, melyekre $y'B - (\alpha, \ldots, \alpha) \geq 0$, és $y'b - \alpha < 0$. Ekkor a duál poliéder bármely y elemére $y + \lambda y'$ minden pozitív λ-ra R^*-ban van, és mivel $y' \neq 0$, így R^* nem korlátos. •

Oldalak

Foglaljuk össze a poliéder oldalainak néhány tulajdonságát. Egy $R = \{x : Qx \leq b\}$ (nemüres) poliéder F oldalán az R-nek egy

$$F := \{x \in R : cx = \delta\} \tag{4.23}$$

alakú nemüres részhalmazát értettük, ahol $\delta := \max\{cx : x \in R\}$ valamely cx célfüggvényre, melyre a maximum létezik. Vagyis a poliéder oldala az optimum helyek halmaza valamely cx lineáris célfüggvényre nézve, másként szólva a poliédernek az a része, amely egy hipersíkkal érintkezik, amikor azt kívülről a poliéderhez toljuk.

4.2.8. Tétel. *Az $R = \{x : Qx \leq b\}$ poliéder egy nemüres F részhalmaza akkor és csak akkor oldala R-nek, ha létezik a Q bizonyos soraiból álló olyan Q' részmátrix, amelyre $F = \{x \in R : Q'x = b'\}$, ahol b' a Q' sorainak megfelelő részvektora b-nek.*

Biz. Tegyük fel, hogy F oldal, melyet (4.23) definiál. Tekintsük a $\min\{yb : yQ = c, y \geq 0\}$ duális lineáris programnak egy y' optimális megoldását. Legyen Q' a Q azon ${}_iq$ soraiból álló részmátrix, amelyekre a megfelelő $y'(i)$ komponens pozitív. Tetszőleges $x \in R$-re $cx = (y'Q)x = y'(Qx) \leq y'b$. A dualitás tételből következik, hogy egy $x' \in R$ vektor akkor és csak akkor primál optimum (azaz eleme F-nek), ha az y' minden pozitív komponensére a neki megfelelő primál feltétel egyenlőséggel teljesül (azaz $y'(i) > 0$-ból ${}_iqx = b(i)$ következik.) Így tehát $F = \{x \in R : Q'x = b'\}$.

Fordítva, legyen Q' a Q bizonyos sorai által alkotott mátrix, és b' a b megfelelő része, amelyekre $\{x \in R : Q'x = b'\}$ nemüres. Legyen e' a csupa egyes vektor, amelynek annyi komponense van, mint ahány sora Q'-nek. Jelölje c a Q' sorainak összegét (azaz $c = e'Q'$), míg δ a b' komponenseinek összegét ($\delta := e'b'$). Most $cx = (e'Q')x = e'(Q'x) \leq e'b' = \delta$. Ebből adódóan valamely $x \in R$ vektorra $Q'x = b'$ akkor és csak akkor teljesül, ha $cx = \delta$, amiből a tétel következik. •

4.2.4. Játékelméleti alkalmazás

Sári és Oszi a következő játékot játsszák. Egyszerre elrejtenek a kezükben egy vagy két forintot és egyúttal tippelnek arra, hogy a másik egy vagy két forintot rejtett. Amennyiben mindkettejük tippje helyes, avagy mindkettejük tippje téves, úgy a játék döntetlen. Ha viszont pontosan az egyikük tippje helyes, úgy a jól tippelő elnyeri a kettejük által elrejtett pénz *összegét* (ami tehát 2,3 vagy 4 forint). Ez a játék egy fordulója. Kérdés, hogy ha N fordulót játszanak, milyen stratégiát érdemes követni.

Mindkét játékos egy lehetséges fordulóbeli játékát egy számpárral lehet megadni, amelynek első tagja azt jelenti, hogy hány forintot rejtett, a második tagja pedig azt, hogy hány forintot tippelt. Vagyis egy fordulóban mindkét játékos előtt négy lehetséges választás van: $[1,1]$, $[1,2]$, $[2,1]$, $[2,2]$. Nevezzük ezeket **elemi** vagy **tiszta** stratégiának. A későbbiekben ezen sorrend szerint fogunk rájuk hivatkozni (tehát pl. a 3-dik elemi stratégia $[2,1]$).

Ha Sári például mindig az $[1,1]$ párt választja, akkor könnyen rosszul járhat, mert ezt ellenfele hamar kifigyelheti, és akkor a $[2,1]$ válasszal mindig nyer. Sári persze ravaszabbul is eljárhat, például mindig ugyanúgy rejt és tippel, mint ahogy Oszi tette a megelőző fordulóban, de ennek a stratégiának is az a hátulütője, hogy Oszi előbb-utóbb rájöhet az alkalmazott szabályra és akkor már könnyen nyer. Ez a veszély minden determinisztikusan meghatározott választási szabály esetén fennáll. Ezt elkerülendő Sári minden fordulóban a véletlentől teszi függővé a választását. Természetesen az a kérdés, hogy milyen valószínűséggel válasszon a négy lehetőség közül.

Tételezzük fel, hogy a lejátszott N forduló során Oszi c_i-szer játszotta meg az i-edik elemi stratégiát ($i = 1, \ldots, 4$), azaz $\sum c_i = N$. Tegyük fel, hogy Sári a következő stratégiát alkalmazta: mindig az ellenkezőjét mondja annak, mint amit tippel és ezen belül $1/2$ valószínűséggel rejt 1 vagy 2 forintot. Másként szólva a $(0, 1/2, 1/2, 0)$ valószínűségek szerint választ minden fordulóban a négy elemi stratégiából. Várható értékben mekkora nyereségre számíthat?

Oszi c_1-szer játszott $[1,1]$-t. Átlagosan ezen c_1 eset felében Sári $[1,2]$-t játszik, amikoris Sári 2 forintot veszít, a másik $c_1/2$ esetben Sári $[2,1]$-t játszik, és ekkor 3 forintot nyer. Tehát Sári várható nyeresége $3c_1/2 - 2c_1/2 = c_1/2$.

$c_2 + c_3$ esetben Oszi mást tippel, mint rejt, ezek a fordulók tehát mind döntetlenek.

Végül Oszi c_4 esetben játszik $[2,2]$-t. Ezeknek átlagosan a felében Sári $[1,2]$-t játszik és nyer 3 forintot, míg a másik felében Sári $[2,1]$-t játszik és veszít 4 forintot. Ezen c_4 esetben tehát Sári várható össz-nyeresége $3c_4/2 - 4c_4/2 = -c_4/2$.

Megállapíthatjuk tehát, hogy az N forduló során Sári várható össz-nyeresége $(c_1 - c_4)/2$ forint, ami persze veszteség, ha $c_4 > c_1$. Vagyis a fent választott $y = (0, 1/2, 1/2, 0)$ valószínűségi választás mellett Sári akkor jár a legrosszabbul, ha $c_4 = N$ és $c_1 = c_2 = c_3 = 0$. Ekkor Sári teljes vesztesége várhatólag $N/2$ forint, azaz fordulónként átlagosan $1/2$ Ft. Azaz Sári ezzel a stratégiával azt tudja biztosítani magának, hogy átlagos vesztesége Oszi bármilyen játéka esetén se haladja meg az $1/2$ forintos fordulónkénti átlagot.

Az $y = (0, 1/2, 1/2, 0)$ valószínűségek helyett természetesen választhatunk más eloszlást is. Nevezzünk egy y vektort **sztochasztikus**nak, ha nem-negatív és $1 \cdot y = 1$. Minden sztochasztikus vektor egy **kevert stratégiá**t definiál. Természetesen a fentiek mintájára tetszőleges kevert stratégiára rögzített $c := (c_1, \ldots, c_4)$ gyakoriságok esetén kiszámíthatjuk Sári várható nyereségét. Könnyű ellenőrizni, hogy ez éppen az $(yA)c$

szám lesz, ahol A az úgynevezett kifizetési mátrix (Sári szempontjából). Azaz A egy $4 \cdot 4$-s mátrix, amelynek a_{ij} eleme Sári nyereségét (másszóval Oszi veszteségét) jelzi, ha Sári az i-edik, míg Oszi a j-edik elemi stratégiát játssza.

Sári akkor fogja egy másik kevert stratégiáját jobbnak tekinteni, mint a $(0, 1/2, 1/2, 0)$ kevert stratégia, ha a fordulónkénti átlagos nyeresége nagyobb, mint az előbb adódott $-1/2$ forint. Van-e ilyen jobb stratégia és hogyan lehet a legjobbat megtalálni?

Általánosabban fogalmazva legyen adva egy A $n \times m$-es mátrix. A sorjátékos Sári és az oszlopjátékos Oszi azt játsszák, hogy minden fordulóban Sári kiválasztja A-nak egy i sorát, míg Oszi A-nak egy j oszlopát, és ennek megfelelően Oszi fizet Sárinak a_{ij} forintot (ami persze azt jelenti, hogy ténylegesen Sári fizet, amennyiben a_{ij} negatív.) Az előbbi játékhoz például a következő mátrix tartozik.

$$
\begin{array}{c c}
 & \begin{array}{cccc} [1,1] & [1,2] & [2,1] & [2,2] \end{array} \\
\begin{array}{c} [1,1] \\ [1,2] \\ [2,1] \\ [2,2] \end{array} & \begin{pmatrix} 0 & 2 & -3 & 0 \\ -2 & 0 & 0 & 3 \\ 3 & 0 & 0 & -4 \\ 0 & -3 & 4 & 0 \end{pmatrix}
\end{array}
$$

Sári egy kevert stratégiáját egy y (m-dimenziós) sztochasztikus vektor definiálja éspedig úgy, hogy Sári minden fordulóban $y(i)$ valószínűséggel választja az i-edik sort.

Tegyük fel, hogy N forduló során Oszi c_j-szer játszotta meg a j-edik oszlopot ($\sum c_j = N$). Ekkor a j-edik oszlop gyakorisága $x_j := c_j/N$. Nyilván a gyakoriságok $x := (x_1, \dots, x_n)$ vektora sztochasztikus.

Mennyire jó Sárinak egy rögzített y sztochasztikus vektor mint kevert stratégia? Adott x gyakoriság esetén Sári össz-nyeresége várható értékben $(yA)c$, vagyis fordulónkénti átlagos nyeresége $(yA)x$. Ez azon x gyakoriság mellett a legrosszabb Sárinak, amelyre $(yA)x$ legkisebb. Vagyis egy y kevert stratégia $f(y)$ jóságát az $f(y) := \min\{(yA)x : x \text{ sztochasztikus}\}$ érték méri. Sárinak tehát az az y a legjobb, amelyre $f(y)$ maximális.

Rögzített y esetén könnyű $f(y)$-t megállapítani, hiszen ez a $\min\{a_y x : x \geq 0, 1 \cdot x = 1\}$ lineáris programnak az optimuma, ahol $a_y := yA$. Mivel az optimum csúcsokban vétetik fel és az $\{x : x \geq 0, 1 \cdot x = 1\}$ poliéder csúcsai épp a $(0, \dots, 0, 1, 0, \dots, 0)$ alakú egységvektorok (n darab), ezért $f(y)$ nem más, mint az yA vektor alja (azaz legkisebb komponense).

Sári optimális kevert stratégiájának megkeresése tehát egy olyan sztochasztikus y vektor megkeresésével egyenértékű, amelyre az yA vektor alja a lehető legnagyobb. Analóg adódik, hogy Oszi optimális kevert stratégiája egy olyan sztochasztikus x vektor megkeresését igényli, amelyre az Ax vektor teteje a lehető legkisebb.

A 4.2.6 tétel alapján a két érték egyenlő, amiből kapjuk a kétszemélyes zéróösszegű játékok alaptételét.

4.2.9. Tétel (Neumann). *Tetszőleges A mátrix által meghatározott mátrixjáték esetén a sorjátékos várható nyereségének* (a legjobb kevert stratégiával elérhető) *maximuma egyenlő az oszlopjátékos várható veszteségének* (a legjobb kevert stratégiával elérhető) *minimumával.*

Visszatérve a kiindulási mátrixjátékhoz, kiszámítható (például a szimplex módszer segítségével), hogy a legjobb kevert stratégia $[0, 3/5, 2/5, 0]$. Ennek alkalmazásával Sári biztosíthatja, hogy várható értékben nem veszít. A játék szimmetrikus, ezért ugyanez a kevert stratégia Oszinak is optimális. Ha Oszi bármely más kevert statégia szerint játszik, azaz ha a megjátszott elemi stratégiáinak gyakorisága eltér a $[0, 3/5, 2/5, 0]$ gyakoriságtól, úgy Sári várható értékben nyer.

Ha ténylegesen játszani akarjuk a játékot, meg kell állapodni abban, hogy egy fordulóban ki mondja ki először a tippjét. Sári „udvariasan" felajánlja, hogy mindig Oszi mondja ki először. Tehát mindketten rejtenek, majd Oszi kimondja a tippjét, utána Sári is kimondja a tippjét. Feltéve, hogy Sári gondolkodhat Oszi tippjének ismeretében (de persze azon már nem változtathat, amennyit rejtett), ki tudja-e aknázni Sári ezt a látszólagos előnyt? Ránézésre azt hihetnénk, hogy ez nem jelent valódi előnyt, hiszen az elrejtett forintok száma minden fordulóban azelőtt kerül meghatározásra mielőtt akármelyik tipp elhangzik. Mindenesetre a fenti általános modell segítségével a kérdést precízen meg lehet válaszolni. Az eredeti mátrixot még kiegészítjük négy sorral, mivel Sárinak négy új tiszta stratégiája adódott. Nevezetesen: A: Sári 1-t rejt és ugyanazt tippeli, mint Oszi, B: Sári 1-t rejt és az ellenkezőjét tippeli, mint Oszi, C: 2-t rejt és ugyanazt tippeli, mint Oszi, D: 2-t rejt és az ellenkezőjét tippeli, mint Oszi. A szimplex módszer segítségével ki lehet számítani, hogy Sári optimális kevert stratégiáját a következő sztochasztikus vektor adja meg: $[0, 56/99, 40/99, 0, 0, 2/99, 0, 1/99]$. Ennek alkalmazásával Sári (Oszi bármilyen játéka esetén is) átlagosan $4/99$ forint nyereségre számolhat fordulónként.

5. fejezet

Lineáris programozás és hálózati optimalizálás

Mi állhat annak hátterében, hogy utakkal, folyamokkal, áramokkal, páros gráfok párosításaival kapcsolatban megannyi szép tételt tudtunk megfogalmazni és igazolni? Miként lehet ilyen tételeket megsejteni? Ebben a fejezetben megmutatjuk, hogy a szóbanforgó hálózati optimalizálási feladatok egy olyan lineáris programként írhatók fel, amelyben a feltételi mátrix teljesen unimoduláris (TU). Kiderül, hogy a tételek mindegyike úgy tekinthető, mint egy lineáris programozási tétel (Farkas lemma, korlátossági tétel, optimalitási feltétel, dualitás tétel) TU-mátrixokra felírt alakjának speciális esete. TU-mátrixokra ugyanakkor alább kimutatjuk, hogy a lineáris programozás alaperedményei erősebb, "egészértékű" alakban is fennállnak. Ennek a felismerésnek nem csak az lesz a haszna, hogy az első fejezetben már igazolt tételekre újabb bizonyítást nyerünk, hanem általa olyan hatékony eszköz birtokába jutunk, amely általánosabb ilyen irányú tételek megsejtésére és bizonyítására is alkalmas.

5.1. Teljesen unimoduláris mátrixok

Az alábbiakban egy mátrixot vagy egy vektort akkor nevezünk egésznek vagy egészértékűnek, ha minden elemük (komponensük) egész szám. Gyakran előfordul, hogy egy lineáris egyenlőtlenség-rendszernek egész megoldására vagy egy lineáris programnak egész optimális megoldására van szükségünk. Bebizonyították, hogy mindkét feladat NP-teljes, így általánosságban olyan típusú kerek megoldást nem várhatunk, mint amilyent a Farkas lemma vagy a dualitás tétel nyújt a valós (vagy racionális) esetre. Speciális feltételi mátrixok esetén azonban szavatolható egészértékű megoldás vagy optimum létezése. Ennek messzemenő következményei lesznek gráfokon megfogalmazott optimalitási feladatok megértésében.

5.1.1. Definíciók és példák

Valamely Q mátrixot akkor nevezünk **teljesen unimodulárisnak** (TU: totally unimodular), 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 a Q TU-mátrixot kiegészítjük egy I egység-mátrixszal, akkor a keletkező (Q, I) mátrix is TU-mátrix. Ha Q TU-mátrix, úgy $(Q, -Q)$ is az. (De ha mondjuk egy csupa 1 oszloppal egészítjük ki Q-t, akkor nem feltétlenül kapunk TU-mátrixot: legyen Q az $\{1, 2, 3, 4\}$ pontokon az $\{12, 13, 14\}$ élekből álló gráf 4×3-as incidencia mátrixa.)

Példaképp, legyen Q egy $D = (V, A)$ irányított gráf incidencia mátrixa, azaz Q sorai a V-nek, oszlopai A-nak felelnek meg, és az $q_{v,e}$ elem akkor $+1$ illetve -1, ha az e él belép illetve kilép v-ből (egyébként 0). Egy $G = (V, E)$ gráf (pont-él) incidencia mátrixában a soroknak a csúcsok, míg az oszlopoknak az élek felelnek meg. A mátrix egy v csúcshoz és e élhez tartozó eleme akkor 1, ha e egyik végpontja v, különben 0. Tehát az incidencia mátrix minden oszlopában két darab 1-es elem van.

5.1.1. Tétel. (a) *Digráf incidencia mátrixa teljesen unimoduláris.* (b) *Páros gráf incidencia mátrixa teljesen unimoduláris.*

Biz. (a) Vegyünk egy Q' négyzetes részmátrixot, amelyről be akarjuk látni, hogy determinánsa $0, \pm 1$. Amennyiben ennek van olyan oszlopa, amelyben legfeljebb csak egy nem-nulla elem van, akkor ezen oszlop szerint kifejtve a determinánst, indukcióval kész vagyunk. Így feltehetjük, hogy minden oszlopban pontosan két nem-nulla van (merthogy több nem lehet). Ezek közül az egyik $+1$, a másik -1, vagyis a sorokat összeadva 0-t kapunk, azaz Q' sorai lineárisan függőek, így a determináns 0.

(b) Szorozzuk meg -1-gyel a mátrix azon sorait, amelyek a páros gráf egyik osztályában lévő pontoknak felelnek meg. Ekkor egy irányított gráf incidencia mátrixát kapjuk, amiről az előbb láttuk, hogy TU. •

5.1. Feladat. *Igazoljuk, hogy ha egy páros gráf incidencia mátrixát kibővítjük egy csupa egyesekből álló sorral, akkor TU-mátrixot kapunk, míg ha az oszlopaihoz veszünk egy csupa egyes oszlopot, akkor az így keletkező mátrix nem feltételenül TU.*

5.2. Feladat. *Igazoljuk, hogy egy D digráf incidencia mátrixának oszlopai akkor és csak akkor lineárisan függetlenek, ha D irányított erdő.*

Hipergráfon egy $(V, \mathcal{F})$ párt értünk, ahol V adott alaphalmaz, $\mathcal{F}$ pedig V részhalmazainak egy rendszere, amelyben ugyanaz a részhalmaz több példányban is szerepelhet. Az $\mathcal{F}$ tagjai a hipergráf **hiperélei**. Egy H hipergráfot akkor nevezzünk **teljesen unimoduláris**nak, ha H incidencia mátrixa teljesen unimoduláris. Ez egy olyan $0-1$ értékű mátrix, amelyben a soroknak a V elemei felelnek meg, az oszlopoknak az $\mathcal{F}$ elemei, és a mátrix egy eleme pontosan akkor egy, ha az oszlopának megfelelő hiperél tartalmazza a mátrix-elem sorának megfelelő V-beli elemet. A gráfok speciális hipergráfok, ahol minden hiperél kételemű. Ezek közül már láttuk, hogy a páros gráfok teljesen unimodulárisak. Más gráfok viszont sohasem azok, hiszen egy páratlan kör incidencia mátrixának determinánsa ± 2.

Mint láttuk, minden D digráf ± 1-es incidencia mátrixa TU. Ezt általánosítja a **hálózati mátrix**. Legyen D olyan irányított gráf, amely irányítatlan értelemben összefüggő és legyen F egy feszítő fa. A H_F mátrix sorai az F éleinek felelnek meg, míg az oszlopai az F-en kívüli éleknek. Minden $e = uv$ nem-fa élre a fában egy egyértelmű (nem feltétlenül 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.

5.1.2. Lemma. *Hálózati mátrix részmátrixa is az. Hálózati mátrix sorát vagy oszlopát -1-gyel szorozva hálózati mátrixot kapunk.*

Biz. Egy oszlop eltörlése annak felel meg, hogy a megfelelő nem-fa élt a digráfból kihagyjuk. Egy sor törlése annak felel meg, hogy a megfelelő fa-élt a digráfban összehúzzuk. Egy sor vagy oszlop -1-gyel való szorzása annak felel meg, hogy a megfelelő élt (akár fa-él, akár nem-fa él) átirányítjuk. •

5.1.3. Tétel. *A H_F hálózati mátrix teljesen unimoduláris.*

Biz. A lemma alapján elég belátni, hogy egy négyzetes hálózati mátrix determinánsa $0, 1$ vagy -1. Tekintsük a fának egy v végpontját. Ha az F fa v-vel szomszédos éléhez tartozó sorban lévő nem-nulla 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 gráfot 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 nem-nulla elem van. Ilyen hálózati mátrixról pedig már láttuk, hogy a determinánsa $0, \pm 1$, ugyanakkor a fenti operációk nem változtatták a determináns abszolút értékét. •

5.1.4. Következmény. *Egy olyan hipergráf, amely egy irányított fa élhalmazán van definiálva és a hiperélek irányított utak, teljesen unimoduláris.* •

Egy hipergráfot **laminárisnak** mondunk, ha bármely két hiperéle vagy diszjunkt vagy az egyik tartalmazza a másikat. Például, ha $F = (V, E)$ egy s gyökerű fenyő és minden $e = uv$ éléhez tekintjük a v-ből a fenyőben elérhető pontok halmazát, akkor ezen halmazok lamináris rendszert alkotnak. Valójában ezen állítás megfordítását sem nehéz bebizonyítani, amely szerint minden lamináris halmazrendszer lényegében ilyen alakban áll elő.

Legyen $\mathcal{F}_1$ és $\mathcal{F}_2$ két lamináris hipergráf az S alaphalmazon. Jelölje A_i $(i = 1, 2)$ az $\mathcal{F}_i$ incidencia mátrixának transzponáltját. Ebben az oszlopok az S elemeinek felelnek meg, míg a sorok $\mathcal{F}_i$ elemeinek. Legyen $M := \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$.

5.1.5. Tétel. *M teljesen unimoduláris.*

Biz. Vegyük M-nek egy négyzetes részmátrixát. Az ebben lévő egyesek száma szerinti indukcióval ennek determinánsáról kimutatjuk, hogy 0 vagy ± 1. Mivel A_i bármely részmátrixa is egy lamináris rendszer incidencia mátrixa (miért?!), így feltehetjük, hogy a vizsgált részmátrix maga M. Ha M-ben minden elem nulla, akkor persze a determináns is nulla. Ha M-nek van olyan sora vagy oszlopa, amelyben legfeljebb egy nem-nulla elem van, akkor indukcióval (és kifejtési szabállyal) készen vagyunk.

Ha $\mathcal{F}_1$ is és $\mathcal{F}_2$ is partíció, akkor mind A_1, mind A_2 sorainak összege a csupa 1 vektor, tehát A sorai lineárisan függőek, így $\det(M) = 0$. Tegyük fel, hogy mondjuk

$\mathcal{F}_1$ nem partíció. Ekkor van egy olyan minimális Z tagja, amely része $\mathcal{F}_1$ egy másik tagjának. Ha most $\mathcal{F}_1$-nek valamennyi Z-tartalmazó tagjából kivonjuk Z-t, ami azzal ekvivalens (a laminaritás miatt), hogy a megfelelő sorokból kivonjuk Z sorát, akkor a determináns értéke nem változik. Viszont a keletkező mátrixban kevesebb egyes szerepel, így indukcióval készen vagyunk. •

5.3. Feladat. *Igazoljuk, hogy az 5.1.5 tételben szereplő M mátrix hálózati mátrix!*

5.4. Feladat. *Igazoljuk, hogy az alábbi mátrix teljesen unimoduláris, de sem ő, sem a transzponáltja nem hálózati mátrix:*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

5.1.2. Farkas lemma, dualitás tétel, optimalitási feltételek TU-mátrixokra

Az erős bázis-megoldás fogalma már eddig is hasznos volt (mert csak véges sok volt belőlük, és mert minden, a poliéderen felülről korlátos cx célfüggvény esetén $\max cx$ erős bázis-megoldáson felvétetett.) E fogalom most újabb fontos szerephez jut.

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

Biz. Legyen $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ és tekintsük a

$$Px = b_0, Qx \leq b_1 \tag{5.1}$$

rendszert. A 3.3.11 tétel szerint minden 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 nem-szinguláris részmátrixa és b' jelöli a b azon részét, amely az M' sorainak felel meg. Mármost, ha M TU-mátrix, akkor a nem-szinguláris M' determinánsa $+1$ vagy -1. A Cramer szabály szerint, miután b' egész, az egyértelmű x' megoldás is az. •

5.1.7. Lemma. *Legyen c tetszőleges (nem feltétlenül egészértékű) vektor. Bármely M TU-mátrixszal megadott K metszet-kúpnak, ha van olyan x' eleme, amelyre $cx' > 0$, akkor K-nak van ilyen $(0, \pm 1)$-értékű eleme is.*

Biz. Legyen $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ és tegyük fel, hogy a K kúp a $Px = 0, Qx \leq 0$ rendszer megoldás-halmaza. Mivel x' pozitív számszorosa is K-ban van, feltehető, hogy x' maga olyan, hogy minden komponense a $[-1, +1]$ zárt intervallumba esik. Vagyis a

$$(-1, \ldots, -1) \leq x \leq (1, \ldots, 1), \ Px = 0, Qx \leq 0 \tag{5.2}$$

rendszer által meghatározott korlátos poliédernek x' olyan eleme, amelyre $cx' > 0$. Ekkor a 4.1.2 tétel szerint van olyan x^* erős bázis-megoldása (5.2) rendszernek, amelyre $cx^* \geq cx'$. Az 5.1.6 lemma miatt x^* egészértékű, azaz minden komponense $0, \pm 1$. •

A Farkas lemma szerint az (5.1) és az alábbi (5.3) rendszerek közül pontosan az egyik oldható meg. Az alábbi tétel a Farkas lemma TU-mátrixokra vonatkozó élesítését szolgáltatja.

5.1.8. Tétel. *Tegyük fel, hogy az $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ mátrix teljesen unimoduláris. Ha az (5.1) primál probléma oldható meg és a korlátozó b vektor egész, akkor* (5.1)*-nek van egész megoldása is. Ha az*

$$y_1 \geq 0, yM = 0, yb < 0 \tag{5.3}$$

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

Biz. A tétel első fele következik az 5.1.6 lemmából, és abból a korábbi eredményből, hogy ha létezik megoldás, akkor létezik erős bázis-megoldás is. A tétel második fele pedig az 5.1.7 lemma közvetlen folyománya. •

Egy poliédert akkor nevezünk egésznek, ha minden oldala tartalmaz egész pontot. Ez nyilván azzal ekvivalens, hogy minden (tartalmazásra nézve) minimális oldal tartalmaz egész pontot, továbbá azzal (az oldal definíciója folytán), hogy minden lineáris célfüggvény optimuma egész vektoron is felvétetik. Csúcsos poliéder esetén a poliéder akkor egész, ha minden csúcsa egész. Az alábbi tételek mindegyikében az $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ mátrix teljesen unimoduláris és b egész vektor.

5.1.9. Tétel. *Ha a* $\max\{cx : Px = b_0, Qx \leq b_1\}$ *lineáris programozási problémának létezik megoldása, akkor az optimum egész vektoron is felvétetik (függetlenül attól, hogy c egészértékű vagy sem).* Ekvivalens alakban: *minden TU-mátrix és egész korlátozó vektor által megadott poliéder egész.*

Biz. Miután az optimum erős bázis-megoldáson is felvétetik, az 5.1.6 lemmából a tétel következik. •

Az alábbi tételek ugyanígy következnek a 4.1.5 és 4.2.2 tételekből az 5.1.6 és 5.1.7 lemmák segítségével.

5.1.10. Tétel. *Tegyük fel, hogy $R = \{x : Px = b_0, Qx \leq b_1\}$ nemüres. A következők ekvivalensek.*

(1) $\{cx : x \in R\}$ felülről korlátos.

(2) Nem létezik olyan $(0, \pm 1)$-értékű x' vektor, amelyre $Px' = 0, Qx' \leq 0$, és $cx' > 0$.

(3) Létezik olyan $y = (y_0, y_1)$ vektor, amelyre $y_1 \geq 0$ és $yM = c$, és amely egész, amennyiben c egész. •

5.1.11. Tétel. *Legyen x^* az $R := \{x : Px = b_0, Qx \leq b_1)\}$ poliéder egy eleme. Jelölje $Q^=_{x^*}$ a Q aktív részmátrixát. A következők ekvivalensek.*

(1) x^* maximalizálja cx-t R fölött.

(2) Nem létezik olyan $(0, \pm 1)$-értékű x' vektor, amelyre $Px' = 0$, $Q^=_{x^*} x' \leq 0$, és $cx' > 0$.

(3) Létezik olyan $y = (y_0, y_1)$ vektor, amelyre $y_1 \geq 0$, $yM = c$, $y(b - Mx^*) = 0$, és y egész, amennyiben c egész. •

5.1.3. Kerekítés és egyenletes színezés

Kerekítés

Akkor mondjuk, hogy egy z egész szám az x **szám kerekítése**, ha $|x-z|<1$. (Tehát az $1,01$-nak az 1 és a 2 is kerekítése.) Ez speciálisan azt jelenti, hogy ha x egész, akkor $x=z$. A z vektor az x **vektor kerekítése**, ha minden komponense kerekítés. Egy x nem-egész szám $\lfloor x\rfloor$ alsó egész részén a legnagyobb x-nél kisebb egész számot értjük, míg $\lceil x\rceil$ felső egész részen a legkisebb x-nél nagyob számot. Egész x-re $\lfloor x\rfloor := \lceil x\rceil := x$. Amennyiben x egy vektort jelöl, úgy $\lfloor x\rfloor$ azt a vektort jelöli, amelyet x-ből nyerünk a komponenseinek alsó egész részét véve. Az x vektor $\lceil x\rceil$ felső egész részét analóg módon definiáljuk.

5.1.12. Lemma. *Legyen A teljesen unimoduláris mátrix és x_0 egy vektor. Ekkor létezik egy olyan q egészértékű vektor, amelyre $\lfloor x_0\rfloor \le q \le \lceil x_0\rceil$ és $\lfloor Ax_0\rfloor \le Aq \le \lceil Ax_0\rceil$. Más szóval az x_0-nak van olyan q kerekítése, hogy az A minden a sorára aq kerekítése ax_0-nak.*

Biz. A feltevés szerint az $\lfloor x_0\rfloor \le z \le \lceil x_0\rceil$ és $\lfloor Ax_0\rfloor \le Az \le \lceil Ax_0\rceil$ rendszernek van megoldása, így az 5.1.8 tétel szerint van egész megoldása is. •

Érdemes megfogalmazni az alábbi következményt: *Ha $(S,\mathcal{F})$ teljesen unimoduláris hipergráf, úgy bármely $x_0 : S \to \mathbb{R}$ függvénynek létezik olyan q kerekítése, hogy minden $A\in\mathcal{F}$ hiperélre a $\sum[q(v) : v\in A]$ szám kerekítése $\sum[x_0(v) : v\in A]$-nak.*

5.1.13. Tétel. *Tetszőleges $m\times n$-es B mátrixnak van olyan kerekítése, hogy a következő mennyiségek mind egynél kevesebbel változnak: minden sorösszeg, minden oszlopösszeg, az első j sor elemeinek összege ($j=1,2,\dots,m$), az első i oszlop elemeinek összege ($i=1,2,\dots,n$).*

Biz. Legyen S a B mátrix mezőinek halmaza. B minden sorához legyen a sorban lévő mezők halmaza tagja $\mathcal{F}_1$-nek valamint minden i-re ($2\le i\le m$) az első i sor mezőinek halmaza legyen tagja $\mathcal{F}_1$-nek (összesen tehát $2m-1$ tagja van $\mathcal{F}_1$-ben). $\mathcal{F}_2$ analóg módon van definiálva az oszlopok segítségével. Ekkor $\mathcal{F}_i$ lamináris, így az 5.1.5 tétel és az 5.1.12 lemma alapján készen vagyunk. •

5.1.14. Tétel. *Egy $x_1,\dots x_n$ sorozat elemeinek létezik olyan $z_1,\dots,z_n$ kerekítése, hogy minden $1\le i\le j\le n$ indexre a $z_i+\dots+z_j$ összeg kerekítése az $x_i+\dots+x_j$ összegnek.*

Biz. A $\{v_1,\dots,v_n\}$ alaphalmazon tekintsük azt a hipergráfot, melynek élei a $\{v_i,\dots,v_j\}$ típusú halmazok minden $1\le i\le j\le n$ index párra. Amint már láttuk, ez a hipergráf teljesen unimoduláris, így az 5.1.12 lemma alkalmazható. •

Egyenletes színezések

A teljesen unimoduláris mátrixok egy másik érdekes alkalmazása hipergráfok egyenletes színezésével foglalkozik.

5.1.15. Tétel. *Legyen A TU-mátrix, b egész vektor, k pozitív egész. Legyen z olyan egész vektor, amelyre $Az\le kb$. Ekkor z előáll olyan $z_1,z_2,\dots,z_k$ egész vektorok összegeként, melyekre $Az_i\le b$.*

Biz. k szerinti indukció alapján elég egy olyan egész z_1 egész vektort találni, amelyre $Az_1 \leq b$ és $A(z - z_1) \leq (k-1)b$. Ugyanis ilyen z_1 létezése esetén $z' := z - z_1$ olyan, amelyre $Az' \leq (k-1)b$ és az indukciós feltevés alkalmazható $(k-1)$-re.

A fenti z_1 létezéséhez csak azt kell látni, hogy az $Az-(k-1)b \leq Ax \leq b$ poliédernek van egész pontja. A poliéder mindenesetre nemüres, hiszen z/k benne van. Továbbá a feltételek egy TU-mátrixszal adhatók meg, így létezik a kívánt egész pont is. •

A fenti tétel kiterjeszthető arra az esetre, amikor z nemnegativitását is megköveteljük, és az Ax-re nemcsak felső korlát van, hanem alsó is. Valóban, ha A TU-mátrix, akkor az $(A, -A, I)$ mátrix is teljesen unimoduláris. Kapjuk a következőt.

5.1.16. Következmény. *Ha $z \geq 0$ olyan egész vektor, amelyre $kb_1 \leq Az \leq kb_2$, akkor z felbomlik olyan $z_1, z_2, \ldots, z_k$ egész vektorok összegére, melyekre $z_i \geq 0$, és $b_1 \leq Az_i \leq b_2$.* •

Ezt felhasználhatjuk TU-mátrixok oszlopainak egyenletes k-színezésére. Az A oszlopainak egy partícióját („színezését") $A_1, A_2, \ldots, A_k$ részre akkor nevezzük **egyenletes**nek, ha A minden a sorára érvényes, hogy a sornak az egyes A_i részekbe eső elemeinek összege minden A_i-re lényegében ugyanaz, tehát $\lfloor e_n a/k \rfloor$ vagy $\lceil e_n a/k \rceil$.

5.1.17. Tétel. *Az A TU-mátrix oszlopainak létezik egyenletes k-színezése.*

Biz. Legyen d az A oszlopainak az összege. Legyen $b_1 := \lfloor d/k \rfloor, b_2 := \lceil d/k \rceil$. Ekkor a $z :\equiv 1$ benne van a $\{kb_1 \leq Ax \leq kb_2, x \geq 0\}$ poliéderben. Az előbbi következmény szerint z felbomlik $z_1, z_2, \ldots, z_k$ egész vektorok összegére, melyekre $z_i \geq 0$, és $b_1 \leq Az_i \leq b_2$. Világos, hogy a z_i-k $0-1$ vektorok. Legyen A_i az oszlopoknak azon halmaza, melyeknek megfelelő komponense z_i-nek 1. Ezek éppen a kívánt egyenletes színezést adják. •

Egy alkalmazás

5.1.18. Következmény. *Adott egy F irányított fa* (speciális esetben irányított út) *és F irányított részútjainak egy $\mathcal{P} := \{P_1, \ldots, P_t\}$ rendszere, ahol minden utat F-élek egy részhalmazának tekintünk. $\mathcal{P}$ tagjai megszínezhetők k színnel (minden k pozitív egészre) úgy, hogy F minden e élére az e-t tartalmazó egyszínű utak száma minden színre lényegében ugyanannyi, ahol a „lényegében ugyanannyi" azt jelenti, hogy bármely két színosztályra az eltérés legfeljebb egy lehet.* •

Ha a hálózati mátrix transzponáltjára alkalmazzuk az egyenletes színezési tételt, akkor a következőt kapjuk.

5.1.19. Következmény. *Adott egy F irányított fa és F irányított részútjainak egy $\mathcal{P} := \{P_1, \ldots, P_t\}$ rendszere, ahol minden utat F-élek egy részhalmazának tekintünk. Az F élei megszínezhetők k színnel (minden k pozitív egészre) úgy, hogy $\mathcal{P}$ minden tagjában a színek lényegében egyenletes számban fordulnak elő.* •

5.5. Feladat. *Egyszerű mohó algoritmus megadásával közvetlenül bizonyítsuk be az* 5.1.19 *következményt.*

Az 5.1.17 tétel páros gráfokra vonatkozó következményeit az 5.2.6 tételben tárgyaljuk.

5.2. A lineáris programozás alkalmazásai a hálózati optimalizálásban

Ebben a részben áttekintjük az első fejezetben megismert eredményeket a lineáris programozás szemszögéből. A csupa egyesből álló j-dimenziós vektort e_j jelöli, míg a $j \cdot j$-es identitás mátrixot I_j.

5.2.1. Páros gráfok: optimális részgráfok

Optimális párosítások

Először levezetjük Kőnig az eslő fejezetben már megismert 1.4.1 tételét:

5.2.1. Tétel (Kőnig). *A $G = (S, T; E)$ páros gráfban a független élek maximális ν száma egyenlő az éleket lefogó pontok minimális τ számával.*

Biz. A gráf pontjainak számát jelölje p az élek számát q. A páros gráf incidencia mátrixát jelölje A, amelyben a soroknak a gráf pontjai, az oszlopoknak a gráf élei felelnek meg. Ekkor tehát A egy $p \times q$ méretű $0-1$-mátrix. Tekintsük a következő primál-duál lineáris program párt:

$$\max\{e_q x : Ax \leq e_p, x \geq 0\}, \tag{5.4}$$

$$\min\{e_p y : yA \geq e_q, y \geq 0\}. \tag{5.5}$$

Az 5.1.9 tétel szerint mindkét programnak az optimuma egész vektoron felvétetik. Jelöljük ezeket rendre x_0-lal és y_0-lal. (5.4) minden egészértékű megoldása $0-1$ értékű, és rögtön látszik, hogy (5.5) 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 1-et vesz fel. Az $Ax \leq e_p$ feltétel azt jelenti, hogy M párosítás a gráfban, míg az $yA \geq e_q$ 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. •

E bizonyítás kapcsán azt mondhatjuk, hogy a Kőnig tétel nem más, mint a dualitás tétel TU-mátrixokra vonatkozó egészértékű alakja abban a speciális esetben, amikor a feltételi mátrix a páros gráf incidencia mátrixa, míg a korlátozó vektor és a célfüggvény a (megfelelő dimenziós) azonosan 1 vektor. Természetesen a primál programban az azonosan 1 célfüggvény helyett választhatunk tetszőleges c célfüggvényt. Ekkor a fenti megközelítés az 1.4.10 tételt adja meg:

5.2.2. Tétel. *Páros gráfban egy párosítás maximális költsége egyenlő* $\min\{\sum_{v \in V} \pi(v) : \pi \geq 0, \pi(u) + \pi(v) \geq c(uv)$ *minden* uv *élre*$\}$. *Ha* c *egészértékű, az optimális* π *is választható egészértékűnek.* •

Melléktermékként kapjuk:

5.2.3. Tétel. *A G páros gráf A incidencia mátrixával felírt*

$$\{x : Ax \leq e_p, x \geq 0\} \tag{5.6}$$

poliéder egész, amelynek csúcsai pontosan a gráf párosításainak incidencia vektorai. •

Egy gráf **párosítás politop**ja a párosítások incidencia vektorainak konvex burka. A 3.4.3 tétel szerint tetszőleges politop (korlátos) poliéder, azaz felírható egy lineáris egyenlőtlenség-rendszer megoldáshalmazaként. Az 5.2.3 tétel az (5.6) rendszerrel tehát konkrétan megadja a párosítás politop poliéderként történő előállítását. (Ezek miatt nem okozhat félreértést, hogy a párosítás politopot gyakran párosítás poliédernek hívják.) Megjegyzendő, hogy tetszőleges gráfra is a párosítás politop mindig része az (5.6) poliédernek, de ilyenkor lehet valódi része.

Nevezzünk egy mátrixot **bisztochasztikus**nak, ha négyzetes, nemnegatív és minden sorösszege valamint minden oszlopösszege egy. Legegyszerűbb bisztochasztikus mátrixok a permutáció mátrixok, melyeknek minden eleme 0 vagy 1 és minden oszlopában és minden sorában pontosan egy darab egyes van. Permutáció mátrixok konvex kombinációja is bisztochasztikus. A következő tétel fő mondanivalója az, hogy valójában minden bisztochasztikus mátrix előáll ilyen alakban.

5.2.4. Tétel (Birkhoff és Neumann)**.** *Egy mátrix akkor és csak akkor bisztochasztikus, ha permutáció mátrixok konvex kombinációja.*

Biz. Egy B $n \times n$-es mátrix megfelel egy G $n \times n$-es teljes páros gráf élhalmazán értelmezett x_B vektornak. Figyeljük meg, hogy a permutáció mátrixok éppen a teljes párosításoknak felelnek meg. Ha B bisztochasztikus, akkor $Ax_B = e_{n^2}, x_B \geq 0$, azaz x_B benne van a G párosítás poliéderében, vagyis előáll párosítások (incidencia vektorainak) konvex kombinációjaként. Tehát B előáll permutáció márixok konvex kombinációjaként. •

Természetesen megkaphatjuk Egerváry 1.4.4 tételét, sőt most már belefoglaljuk azt az esetet is, amikor a súlyfüggvény nem egész.

5.2.5. Tétel (Egerváry)**.** *A $G = (S, T; E)$ teljes párosítással rendelkező páros gráfban a $c \geq 0$ súlyfüggvényre vonatkozó maximális súlyú teljes párosítás ν_c súlya egyenlő a súlyozott lefogások minimális τ_c összértékével. Amennyiben G teljes páros gráf, úgy az optimális súlyozott lefogás választható nemnegatívnak is. Amennyiben c egészértékű az optimális súlyozott lefogás is választható annak.*

Biz. A fenti megközelítéshez képest csak annyit kell változtatni, hogy az $Ax \leq e_p$ egyenlőtlenség rendszer helyett az $Ax = e_p$ egyenletrendszert kell vennünk. Ekkor persze a duálisban a változókra nincs nemnegatívitás előírva. A teljes páros gráf esetén azért igaz mégis, hogy az optimális duális megoldás választható nemnegatívnak, mert ilyenkor az $\{\max cx : Ax \leq e_p, x \geq 0\}$ lineáris program optimális megoldása c nem negativitása valamint a páros gráf teljessége miatt mindig teljes párosításon is felvétetik, márpedig ezen lineáris program duálisában a változók nemnegatívak. •

Mi történik, ha adott k-ra a pontosan k élű párosítások maximális súlyára szeretnénk tételt kapni? Miután bebizonyítható, hogy egy páros gráf incidenciamátrixát egy csupa egyes sorral kiegészítve továbbra is TU-mátrixot kapunk (figyelem: csupa egyes oszloppal való kiegészítéssel nem), így a következő primál-duál lineáris program pár megadja a választ: $\max\{cx : Ax \leq e_p, e_q x = k\}$ és $\min\{\pi e_p + k\alpha : \pi A + \alpha e_q \geq c, \pi \geq 0\}$. A primál optimum tehát egészértékű, és így szükségképpen egy k elemű párosítás incidencia vektora. A duál optimum is egészértékű, feltéve, hogy c az.

Páros gráf fokszámkorlátozott részgráfjai: a szállítási probléma

További általánosításokat kaphatunk, ha a primál feladatban a jobboldalt valamilyen (nem-negatív) b vektornak választjuk. Ennek az a kombinatorikus jelentése, hogy a páros gráfban maximális súlyú fokszám-korlátozott részgráfot keresünk. Természetesen alsó korlátokat is kitűzhetünk a fokszámokra, mint ahogy korlátozhatjuk alulról és felülről azt is, hogy egy élt hány példányban vehetünk be a keresett részgráfba (megint csak amiatt, hogy az incidencia mátrixot egy csupa egyes sorral kiegészítve TU-mátrixot kapunk). Valójában nem is érdemes explicit 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. Emlékeztetünk, hogy korábban ezen feladatok körét neveztük szállítási problémának.

5.2.2. Páros gráfok: élszínezések

Közismert Kőnig élszínezési tétele, amely szerint minden Δ-reguláris páros gráf élhalmaza felbomlik Δ élidegen teljes párosításra. (Ez közvetlenül levezethető indukcióval, vagy esetleg a Hall tételre támaszkodva). Ugyanakkor a TU-mátrixokra vonatkozó 5.1.17 egyenletes színezési tételből sokkal általánosabb eredmény nyerhető. Az élszínezési tételt néha kicsit általánosabban fogalmazzák meg: *Ha egy páros gráfban a maximális fokszám Δ, akkor az éleket meg lehet Δ színnel színezni úgy, hogy minden csúcsba különböző színű élek futnak.*

5.2.6. Tétel. *Egy $G = (S, T; E)$ páros gráf éleit meg lehet k színnel úgy színezni, hogy minden v csúcsra és mindegyik j színre ($j = 1, \dots, k$) a v-be menő $d(v)$ darab él közül $\lfloor d(v)/k \rfloor$ vagy $\lceil d(v)/k \rceil$ darab színe j. Ráadásul még azt is megkövetelhetjük, hogy minden színosztály mérete közel ugyanakkora legyen, vagyis $\lfloor |E|/k \rfloor$ vagy $\lceil |E|/k \rceil$. Ha k-t a maximális Δ fokszámnak választjuk, akkor megkapjuk Kőnig élszínezési tételét, amely szerint páros gráf kromatikus indexe (élszínezési száma) a maximális fokszámmal egyenlő. Ha k-t a minimális δ fokszámnak választjuk, akkor Gupta egy tételét kapjuk, amely szerint G páros gráf élhalmaza felbontható δ részre úgy, hogy mindegyik rész fedi az összes pontot.* •

5.2.3. Megengedett potenciálok, legolcsóbb utak

Legyen $D = (V, A)$ irányított gráf, melynek $(0, 1, -1)$-es incidencia mátrixát jelölje Q. Egy $\pi : V \to \mathbb{R}$ vektort akkor neveztünk a $c : A \to \mathbb{R}$ költségfüggvényre nézve megengedett potenciálnak, ha $\pi(v) - \pi(u) \leq c(uv)$ fennáll minden $uv \in A$ élre. Figyeljük meg, hogy egy π vektor pontosan akkor megengedett potenciál, ha $\pi Q \leq c$. Egy $x : A \to \mathbb{R}$ vektor pedig pontosan akkor áram, ha $Qx = 0$. Megmutatjuk, hogy a megengedett potenciál létezésére vonatkozó 1.3.8 tétel rögtön következik a Farkas lemma TU-mátrixokra vonatkozó élesebb alakjából. Az alábbi tétel az 1.3.8 tétel más szövegezéssel.

5.2.7. Tétel. *Adott $c : A \to \mathbb{R}$ költség-függvényre akkor és csak akkor létezik olyan $\pi : V \to \mathbb{R}$ vektor, amelyre $\pi(v) - \pi(u) \leq c(uv)$ minden $e = uv \in A$ élre, ha c konzervatív, azaz ha nem létezik negatív költségű irányított kör. Amennyiben c egészértékű, úgy a potenciál is választható annak.*

Biz. A Q mátrix transzponáltja teljesen unimoduláris, így az 5.1.8 tétel miatt vagy létezik a $\pi Q \leq c$ rendszernek megoldása (amely egész, ha c az), vagy pedig a duális $\{Qx = 0, x \geq 0, cx < 0\}$ rendszernek létezik egy $(0, \pm 1)$-es megoldása. Az első eset épp egy megengedett potenciál létezését jelenti, míg a második esetben, $x \geq 0$ miatt, x egy $(0,1)$ értékű, negatív költségű áram, amely élidegen körökre bomlik, és így e körök egyike is negatív. •

A dualitás tétel TU-mátrixokra vonatkozó élesített alakjából könnyen levezethető az 1.3.15 tétel is.

5.2.8. Tétel. *Konzervatív c költségfüggvény esetén az s-ből t-be vezető utak költségének $l_c(t)$ minimuma egyenlő $\pi(t) - \pi(s)$ maximumával, ahol a maximum az összes megengedett π potenciálon veendő.*

Biz. Tegyük fel, hogy a Q mátrix első és második sora felel meg az s illetve a t pontnak. Tekintsük a $\max\{\pi(t) - \pi(s) : \pi Q \leq c\}$ lineáris programot. Ennek duálisa $\min\{cx : Qx = (-1, +1, 0, 0, \ldots, 0), x \geq 0\}$. A primál program optimális megoldása épp a tétlben szereplő maximum. Mivel Q TU-mátrix, így az 5.1.9 tétel miatt létezik egészértékű optimális π is, ha c egész. A duális programnak az 5.1.9 szerint a c egészértékűségétől függetlenül létezik egy x^* egészértékű optimuma. Figyeljük meg, hogy a $Qx = (-1, +1, 0, 0, \ldots, 0), x \geq 0$ megoldásai éppen az egy nagyságú folyamok. Mivel x^* egészértékű, így előáll, mint egy út és irányított körök (incidencia vektorainak) nemnegatív kombinációjaként. De c konzervatívitása miatt a körök költsége nemnegatív, így ezeket kihagyva feltehetjük, hogy x^* egy st út incidencia vektora. •

5.2.4. Megengedett áramok és folyamok

Korábban már megjegyeztük, hogy ha a megmaradási szabály helyett csupán a $\varrho_x(v) \leq \delta_x(v)$ egyenlőtlenséget írjuk elő minden v csúcsnál, akkor x automatikusan áram, másszóval a $Qx \leq 0$ egyenlőtlenség-rendszer megoldáshalmaza pontosan az áramok halmaza. (Ezt kellett bizonyítani az 1.5.1 gyakorlat (a) részében.)

5.2.9. Tétel. *Ha $f \leq g$ egészértékű, akkor a megengedett áramok $\{x : Qx \leq 0, f \leq x \leq g\}$ poliédere, amennyiben nemüres, egész poliéder.*

Biz. Mivel Q TU-mátrix, így ha kiegészítjük egy (negatív) egységmátrixszal, úgy továbbra is TU-mátrixot kapunk, és így az 5.1.9 tételt alkalmazhatjuk. •

Hasonló megfontolással kapjuk:

5.2.10. Tétel. *A $D = (V, A)$ digráf élhalmazán adott a $g \geq 0$ egész kapacitásfüggvény. Legyen s és t két kijelölt csúcs, melyekre $\varrho(s) = 0 = \delta(t)$. A k nagyságú megengedett folyamok $\{x \in \mathbb{R}^A : 0 \leq x \leq g, \varrho_x(v) = \delta_x(v)$ minden $v \in V - \{s,t\}$-re, $\delta_x(s) = k\}$ poliédere, amennyiben nemüres, egész poliéder.* •

Hoffman megengedett áramok létezésére vonatkozó tételét korábban már kétféleképpen is beláttuk: egyrészt adtunk rá egy direkt bizonyítást, másrészt levezettük az MFMC tételből is. Most megmutatjuk, hogy a Hoffman tétel lényegében nem más, mint a Farkas lemmának az 5.1.8 tételben TU-mátrixokra vonatkozó élesebb alakja egy digráf incidencia mátrixára felírva.

5.2.11. Tétel (Hoffman, 1960)**.** *A $D = (V, A)$ digráfban adott $f \leq g$ kapacitásfüggvényekre vonatkozólag akkor és csak akkor létezik megengedett áram, ha*

$$\varrho_f(X) \leq \delta_g(X) \text{ minden } X \subseteq V \text{ halmazra.} \tag{5.7}$$

Továbbá, ha f és g egészértékűek és (5.7) fennáll, úgy létezik egészértékű megengedett áram is.

Biz. Csak az elegendőség igazolásával foglalkozunk. Tekintsük a $Qx \leq 0, x \leq g, -x \leq -f$ rendszert. Az 5.1.8 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 az (5.7) feltételnek. •

5.2.5. Minimális költségű áramok és folyamok

Tekintsük most a költséges áram problémát, azaz adott $c : A \to \mathbb{R}$ költségfüggvény esetén keressünk minimális költségű megengedett áramot, más szóval, oldjuk meg a

$$\min\{cx : Qx = 0, f \leq x \leq g\} \tag{5.8}$$

lineáris programot. (Természetesen az $x \leq g$ egyenlőtlenség itt azt jelenti, hogy $x(e) \leq g(e)$ az olyan élekre, ahol $g(e)$ véges. Duális változó tehát csak ilyen egyenlőtlenségekhez tartozik.)

Korlátosság és optimalitás

Először vizsgáljuk meg, hogy cx mikor korlátos alulról. Készítsünk el egy $D' = (V, A')$ digráfot, és élein definiáljuk a c' költségfüggvényt a következőképpen. D'-ben uv akkor él, ha vagy $vu \in A, f(vu) = -\infty$, és ekkor $c'(uv) = -c(vu)$, vagy pedig $uv \in A, g(uv) = \infty$, és ekkor $c'(uv) = c(uv)$. Bár az 5.1.10 tételt specializálva közvetlenül is kiolvasható az alábbi eredmény, újra megadjuk az ottani bizonyítást a mostani helyzetre specializálva.

5.2.12. Tétel. *Feltéve, hogy létezik megengedett áram, a következők ekvivalensek.*

(a) *cx alulról korlátos,*

(b) *nincs negatív összköltségű irányított kör D'-ben,*

(c) *létezik egy olyan $\pi : V \to \mathbb{R}$ függvény, amelyre*

$$\pi(v) - \pi(u) \leq c(uv), \text{ ha } uv \in A \text{ és } g(uv) = \infty, \tag{5.9}$$

$$\pi(v) - \pi(u) \geq c(uv), \text{ ha } uv \in A \text{ és } f(uv) = -\infty. \tag{5.10}$$

Amennyiben c egészértékű, úgy a szóbanforgó π is választható annak.

Biz. (a)→(b) Ha létezik negatív kör D'-ben, akkor ennek egy olyan kör felel meg D-ben, melynek az előremenő élein a g végtelen, a visszamenő élein az f minusz végtelen, és az éleinek összköltsége negatív. Márpedig ha a meglévő megengedett áramot az előremenő éleken bármilyen nagy K-val egységesen megnöveljük a visszamenőkön pedig K-val csökkentjük, akkor megengedett áramot kapunk, amelynek költsége így akármilyen kicsi lehet.

(b)→(c) Ha D'-ben nincs negatív kör, akkor az 5.2.7 tétel miatt létezik egy $\pi : V \to \mathbb{R}$ függvény, amelyre $uv \in A, g(uv) = \infty$ esetén (amikor is $uv \in A'$) $\pi(v) - \pi(u) \leq c'(uv) = c(uv)$ azaz (5.9) fennáll, míg $uv \in A, f(uv) = -\infty$ esetén (amikor is $vu \in A'$) $\pi(u) - \pi(v) \leq c'(vu) = -c(uv)$ vagyis $\pi(v) - \pi(u) \geq c(uv)$, azaz (5.10) fennáll.

(c)→(a) Tetszőleges x áram költsége bármely $\Delta_\pi(uv) := \pi(v) - \pi(u)$ pontindukált költségfüggvény esetén nulla. A $c_\pi(uv) := c(uv) - \pi(v) + \pi(u)$ eltolt költségfüggvényre (5.9) azzal ekvivalens, hogy $c_\pi(uv) > 0$ esetén $g(uv) < \infty$, míg (5.10) azzal, hogy $c_\pi(uv) < 0$ esetén $f(uv) > -\infty$. Ezek alapján egy x megengedett áramra és (c)-t kielégítő π-re $cx = \sum_{uv\in A} c_\pi(uv)x(uv) = \sum_{uv\in A}[c_\pi(uv)x(uv) : c_\pi(uv) > 0] + \sum_{uv\in A}[c_\pi(uv)x(uv) : c_\pi(uv) < 0] = \sum_{uv\in A}[c_\pi(uv)g(uv) : c_\pi(uv) > 0] + \sum_{uv\in A}[c_\pi(uv)f(uv) : c_\pi(uv) < 0]$, ami a cx-re véges alsó korlát. (Most tehát részletesen kiírogatva azt a már korábban látott egyszerű tényt igazoltuk újfent, hogy ha mind a primál, mind a dual poliéder nemüres, akkor cx alulról korlátos a primál poliéderen.) •

Tegyük most fel, hogy x megengedett áram. Készítsünk el egy $D_x = (V, A_x)$ digráfot és az élhalmazán egy c_x költségfüggvényt a következőképpen. Az uv él akkor tartozzék A_x-hez, ha vagy $uv \in A, x(uv) < g(uv)$, és ekkor legyen $c_x(uv) := c(uv)$, vagy pedig $vu \in A, x(vu) > f(vu)$, és ekkor legyen $c_x(uv) := -c(vu)$. Az 5.1.11 tételt specializálva kapjuk a következőt.

5.2.13. Tétel. *Adott x megengedett áram esetén a következők ekvivalensek.*

(a) *x optimális megoldása az (5.8) minimális költségű megengedett áram feladatnak,*
(b) *D_x-ben nem létezik negatív összköltségű irányított kör,*
(c) *létezik egy olyan $\pi : V \to \mathbb{R}$ függvény, amelyre*

$$\pi(v) - \pi(u) \leq c(uv), \text{ ha } uv \in A \text{ és } x(uv) < g(uv),$$

$$\pi(v) - \pi(u) \geq c(uv), \text{ ha } uv \in A \text{ és } x(uv) > f(uv).$$

Amennyiben c egészértékű, úgy a szóbanforgó π is választható annak. •

5.6. Feladat. *Az 5.2.12 tétel fenti direkt bizonyításának mintájára adjuk meg az 5.2.13 tétel közvetlen bizonyítását is.*

5.7. Feladat. *Fogalmazzuk meg és bizonyítsuk be az 5.2.12 és az 5.2.13 tételek megengedett potenciálokra vonatkozó ellenpárját.*

Az áramokra megfogalmazott optimalitási feltételt könnyen átvihetjük folyamokra.

5.2.14. Tétel. *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbb{R}_+$ kapacitásfüggvény és a $c : A \to \mathbb{R}$ költségfüggvény. Egy k nagyságú megengedett z folyam akkor és csak akkor minimális költségű a k nagyságú megengedett folyamok között, ha létezik olyan π potenciál, amelyre fennállnak a következő optimalitási feltételek:*

$$\pi(v) - \pi(u) < c(uv) \Rightarrow z(uv) = 0, \qquad (i)$$

$$\pi(v) - \pi(u) > c(uv) \Rightarrow z(uv) = g(uv). \qquad (ii)$$

Biz. Adjunk a digráfhoz egy ts élt és definiáljuk a költségét 0-nak. Legyen $g(ts) := f(ts) := k$. Minden régi élen legyen $f(e) := 0$. Az így kibővített $D' = (V, A')$ digráfban a megengedett áramok éppen a D-beli k nagyságú folyamoknak felelnek meg, így az 5.2.13 tételt D'-re alkalmazva az (i) és (ii) feltételeket kapjuk. •

A minimális költségű folyamokra vonatkozó algoritmus segítségével már igazoltuk az alábbi tételt, legalábbis abban az esetben, amikor g egészértékű és c nemnegatív (1.6.5 tétel). Megmutatjuk, hogy a háttérben most is az 5.1.9 tételben megfogalmazott TU-mátrixokra vonatkozó egészértékű dualitás tétel áll.

5.2.15. Tétel. *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbb{R}_+$ kapacitásfüggvény és a $c : A \to \mathbb{R}$ költségfüggvény. A k nagyságú megengedett folyamok költségének minimuma egyenlő a*

$$k\pi(t) + \sum[c_\pi(uv)g(uv) : uv \in A, c_\pi(uv) < 0] \qquad (5.11)$$

érték maximumával, ahol a maximum az összes $\pi : V \to \mathbb{R}$ függvényre megy, amelyre $\pi(s) = 0$. Amennyiben g egészértékű, az optimális folyam választható egésznek. Amennyiben c egészértékű, az optimális π választható egészértékűnek.

Biz. Tegyük fel, hogy a digráf Q incidencia-mátrixának első és második sora felel meg az s illetve a t pontnak. Tekintsük a $\min\{cx : x \geq 0, Qx = (-k, +k, 0, 0, \ldots, 0), x \leq g\}$ primál programot. Az $x \leq g$ feltételt az ekvivalens $(-I_m)x \geq -g$ alakba téve felírhatjuk a duális problémát: $\max\{k(\pi(t) - \pi(s)) - gz : \pi Q - zI_m \leq c, z \geq 0\}$, ahol $m = |A|$. A primál poliéder elemei a k nagyságú folyamok. Az 5.1.9 tétel szerint egész g esetén a primál poliéder egész, függetlenül c egészértékűségétől. Hasonlóképp a duális poliéder is egész, amennyiben c egész. Figyeljük meg, hogy tetszőleges π meghatároz egy hozzá tartozó legjobb z-t: $z(uv) := \pi(v) - \pi(u) - c(uv)$, ha $\pi(v) - \pi(u) > c(uv)$, és $z(uv) = 0$, ha $c(uv) \leq \pi(v) - \pi(u)$. Így tehát adott π-hez tartozó $k(\pi(t) - \pi(s)) - gz$ célfüggvény értéke nem más, mint az (5.11) képletben megadott érték, hiszen a π eltolásával feltehetjük, hogy $\pi(s) = 0$. •

5.2.6. Hálózati mátrixokkal adott lineáris programok

Fontos megjegyezni, hogy a hálózati mátrixokkal megadott lineáris programok megoldhatók áram problémaként. Legyen $D = (V, A)$ irányított gráf, F feszítő fa és legyen $N := A - F$ a nem-fa élek halmaza. Legyen adott $f = (f_F, f_N)$ és $g = (g_F, g_N)$ korlát, melyekre $f \leq g$. Legyen továbbá $c = (c_F, c_N)$ egy olyan vektor, amelyre $c_F = 0$. Jelölje az F-hez tartozó $(0, \pm 1)$-es hálózati mátrixot B, míg a D digráf $(0, \pm 1)$-es pont-él incidencia mátrixát Q_D. Legyen továbbá $x = (x_F, x_N)$. Tekintsük a $\max\{c_N x_N : f_F \leq Bx_N \leq g_F, f_N \leq x_N \leq g_N\}$ lineáris programot. Belátjuk, hogy ez ekvivalens a $\max\{cx : Q_D x = 0, f \leq x \leq g\}$ maximális költségű áram feladattal.

Amennyiben $x = (x_F, x_N)$ áram (azaz $Q_D x = 0$), úgy könnyen látszik, hogy $x_F = Bx_N$, és persze $cx = c_N x_N$. Emiatt $f \leq x \leq g$ ekvivalens a $f_F \leq Bx_N \leq$

$g_F, f_N \leq x_N \leq g_N$ feltételekkel. Fordítva, tegyük fel, hogy x_N kielégíti ezen utóbbi egyenlőtlenségeket. Minden $e \in N$ nem-fa élhez legyen $\underline{\chi}_e$ az $(1, a_e)$ vektor, ahol a_e az A mátrix e-hez tartozó oszlopa. (Másszóval, $\underline{\chi}_e$ az e élhez tartozó C_e alapkör $0, \pm 1$-es incidencia vektora.) Ekkor persze $\underline{\chi}_e$ áram, és így az $x := \sum [x_N(e)\underline{\chi}_e : e \in N]$ is áram, méghozzá olyan, hogy $x(e) = x_N(e)$, ha $e \in N$. Látható, hogy $f_F \leq Bx_N \leq g_F$ azzal ekvivalens, hogy $f_F(e) \leq x(e) \leq g_F(e)$ minden $e \in F$ élre fennáll. •

Következik például, hogy páros gráfok éleinek vagy az irányított fák irányított részútjainak egyenletes színezéseire vonatkozó tételeket egy maximális folyamot kiszámító algoritmussal tudjuk algoritmikusan kezelni. Hasonlóképp a kerekítési eredményeket. A minimális költségű megengedett potenciál meghatározásának problémáját pedig úgy lehet algoritmikusan megoldani, hogy felírjuk a hozzátartozó duális feladatot. Ez minimális költségű megengedett áram problémának tekinthető, majd ennek megoldásaként előállítjuk az optimális áramot és ennek optimális duális megoldását, ami éppen az eredeti potenciál probléma megoldása.

6. fejezet

A szimplex módszer változatai

A 3.5.2 fejezetben szerepelt a szimplex algoritmus a Farkas lemmára, ami a gyakorlatban általában hatékonyan eldönti egy egyenlőtlenség-rendszerről, hogy megoldható-e (bár valójában nem polinomiális futási idejű). Ebben a részben kicsit más szemszögből, optimalizálási feladatok megoldási módszereként tárgyaljuk a szimplex módszert. Míg a 3.5.2 fejezetben szereplő módszer a duális feladat bázismegoldásain lépkedett, itt először egy olyan változatot tekintünk, ami primál bázismegoldásokból talál egyre jobbakat.

6.1. Primál szimplex módszer

Tekintsük a következő primál feladatot:

$$\begin{aligned} Ax &= b \\ x &\geq 0 \\ \max\, &cx, \end{aligned}$$

ahol $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m, c \in \mathbb{Q}^{1\times n}$, és a változók vektora $x \in \mathbb{Q}^n$. Ha az A mátrix rangja $r(A) < m$, akkor vagy már az $Ax = b$ egyenletrendszer sem oldható meg, vagy valamelyik egyenlet redundáns. Tehát feltehetjük, hogy $r(A) = m$. A duális feladat:

$$\begin{aligned} yA &\geq c \\ \min\, &yb, \end{aligned}$$

ahol $y \in \mathbb{Q}^{1\times m}$, azaz minden egyenlethez egy duál változó tartozik. Idézzük fel az ilyen alakú feladatokra vonatkozó dualitás tételeket.

6.1.1. Tétel (Gyenge dualitás tétel)**.** *Legyen x primál megengedett megoldás és y duál megengedett megoldás. Ekkor teljesül*

$$cx \leq yb.$$

Bizonyítás. $cx \underbrace{\leq}_{x\geq 0, yA\geq c} (yA)x = y(\underbrace{Ax}_{=b}) = yb.$ •

6.1.2. Tétel (Erős dualitás tétel)**.** *Ha a primál feladat megoldható, és az optimuma korlátos (azaz cx nem lehet tetszőlegesen nagy), akkor*

$$\max cx = \min yb.$$

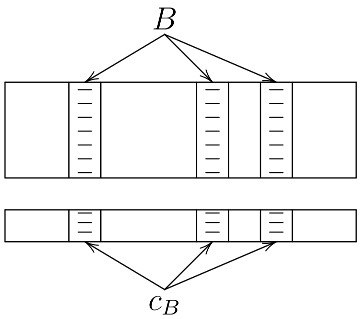
6.1.3. Tétel (Ekvivalens alak – komplementaritási feltétel)**.** *Ha x^* optimális primál megoldás, akkor $\exists\ y^*$ duál megoldás, amire $cx^* = y^*b$, azaz ha $x_j^* > 0$, akkor $(y^*A)_j = c_j$.*

Definíció (Bázis, bázismegoldás)**.** A primál feladat **bázismegoldása** egy olyan x megoldás, amire A-nak az $x_j > 0$-khoz tartozó oszlopai lineárisan függetlenek (lásd 3.3.7 Tétel). Az A-nak egy $m \times m$-es nemszinguláris részmátrixát **bázis**nak nevezzük. Formálisan ebbe beleértjük, hogy a részmátrix oszlopainak egy sorrendje is adott.

Rögzített B bázis esetén egy $x \in \mathbb{R}^n$ vektort $x = (x_B, x_N)$ alakban írhatunk, ahol x_B-vel jelöljük a bázishoz tartozó koordinátákat, x_N-nel pedig a többit, azaz a nembázis koordinátákat. A B-hez tartozó primál vektor: $\bar{x} \in \mathbb{R}^n$: $\bar{x}_B = B^{-1}b$, $\bar{x}_N = 0$. Ha $B^{-1}b \geq 0$, akkor $\bar{x}$ primál megoldás, azaz a B bázis **primál megengedett**.

Megjegyzés. Ha x bázismegoldás, akkor az $x_j > 0$-khoz tartozó oszlopokat kiegészítve m db lineárisan független oszloppá bázist kapunk. Ehhez a B bázishoz pedig pont x lesz a hozzárendelt bázismegoldás, mivel tudjuk, hogy $Bx = b$-nek egyetlen megoldása van. Egy bázismegoldás viszont nem csak egy bázishoz lehet hozzárendelve: ha m-nél kevesebb helyen pozitív, akkor ezeket bárhogy kiegészíthetjük m lineárisan független oszloppá, így több, egymástól különböző bázist kaphatunk. Két bázist különbözőnek tekintünk akkor is, ha ugyanaz a bázismegoldás tartozik hozzájuk.

Definíció (Bázishoz tartozó duális vektor)**.** A B bázishoz tartozó duális vektor: $\bar{y} = c_B B^{-1}$, ahol c_B a c célfüggvény B bázishoz tartozó része.



Az $\bar{y}$ vektor nem feltétlenül megoldása a duál feladatnak.

Észrevétel. *Az $\bar{x}$ és $\bar{y}$ vektorok teljesítik a komplementaritási feltételeket.*

Nézzük $\bar{y}A - c$-t. Erről annyit tudunk, hogy a B-hez tartozó koordinátái nullák.

$$(\bar{y}A - c)_B = (\bar{y}B - c_B) = c_B B^{-1} B - c_B = 0$$

Ha $\bar{y}A - c \geq 0$, akkor $\bar{y}$ duál megoldás. Mivel a komplementaritási feltételek teljesülnek, $\bar{x}$ a primál feladatnak és $\bar{y}$ a duál feladatnak optimális megoldása.

Definíció. A B bázishoz tartozó **redukált költség**: $\bar{c} = \bar{y}A - c$. A bázis **duál megengedett** ha $\bar{c} \geq 0$, és **optimális** ha primál és duál megengedett.

Először a szimplex módszernek azt az egyszerűbb változatát tárgyaljuk, ahol kiindulásként rendelkezésre áll egy primál megengedett bázis, és a cél egy optimális bázis megtalálása.

6.1.1. A szimplex módszer tulajdonságai

- primál megengedett bázisokon lépked
- minden lépésben egy oszlopot cserélünk ki B-ben.
- a primál célfüggvényérték folyamatosan nő (azaz nem csökken)
- véges sok lépésben eljutunk egy optimális bázishoz.

Tegyük fel, hogy B primál megengedett bázis. Tartozik hozzá egy $\bar{x}$ bázismegoldás és egy $\bar{y}$ duál vektor. Figyeljük meg, hogy a $B^{-1}Ax = B^{-1}b$ egyenletrendszer ekvivalens az eredeti, $Ax = b$ egyenletrendszerrel. Vezessük be a következő jelöléseket:

$$\begin{aligned} \bar{A} &= B^{-1}A \\ \bar{b} &= B^{-1}b \\ \bar{c} &= \bar{y}A - c \qquad && \text{– a bázishoz tartozó } \textbf{redukált költség}. \\ \bar{z} &= c\bar{x} = \bar{y}b \end{aligned}$$

A B bázis pontosan akkor primál megengedett ha $\bar{b} \geq 0$, és pontosan akkor duál megengedett ha $\bar{c} \geq 0$. Az itt bevezetett mátrixokat és vektorokat szokás egyetlen táblázatban ábrázolni, amit a B bázishoz tartozó szimplex táblának nevezünk:

$$\begin{array}{ccc} \boxed{\qquad\qquad \bar{A} \qquad\qquad} & & \boxed{\bar{b}} \\ \\ \boxed{\qquad\qquad \bar{c} \qquad\qquad} & & \boxed{\bar{z}} \end{array}$$

Megjegyzés. $\bar{A}$-ban B helyén egységmátrix van, $\bar{b}$ pedig $\bar{x}_B$ értékeit tartalmazza. Tehát ha pl. $\bar{x}_B = (\bar{x}_7, \bar{x}_3, \bar{x}_9)$, akkor $\bar{b}_1 = \bar{x}_7$, $\bar{b}_2 = \bar{x}_3$, $\bar{b}_3 = \bar{x}_9$, és az $\bar{A}$ mátrixban így néz ki a megfelelő rész:

$$\begin{array}{c|c|c|c|c|c|c|c|} & & x_3 & & x_7 & & x_9 & \\ \hline x_7 & & 0 & & 1 & & 0 & \\ x_3 & & 1 & & 0 & & 0 & \\ x_9 & & 0 & & 0 & & 1 & \\ \hline \end{array}$$

A $\bar{c}$ vektort nem véletlenül nevezzük redukált költségnek. Írjuk át a következő alakra:

$$\bar{c} = \bar{y}A - c = c_B B^{-1}A - c = c_B\bar{A} - c$$

Jelölje N a bázisban nem szereplő indexek halmazát. Mi történik akkor, ha egy adott $p \in N$-re $\bar{x}_p$-t növeljük δ-val, és közben $\bar{x}_B$-t úgy változtatjuk, hogy $\bar{A}x = \bar{b}$ továbbra is teljesüljön?

$$\begin{array}{c|c|c|c|c|c|c|c|c|c|} & & x_3 & & x_p & & x_7 & & x_9 & \\ \hline x_7 & & 0 & & - & & 1 & & 0 & \\ x_3 & & 1 & & - & & 0 & & 0 & \\ x_9 & & 0 & & - & & 0 & & 1 & \\ \hline \end{array}$$

1. egyenletnél: $\bar{x}_7$-et változtatjuk

2. egyenletnél: $\bar{x}_3$-at változtatjuk

3. egyenletnél: $\bar{x}_9$-et változtatjuk

Jelölés. Az A mátrix i-edik sorát $a_{i.}$ vagy $A_{i.}$ jelöli, j-edik oszlopát pedig $a_{.j}$ vagy $A_{.j}$.

$$\begin{aligned}\bar{x}'_p =& \bar{x}_p + \delta\\ \bar{x}'_B =& \bar{x}_B - \delta\bar{a}_{.p}\\ c\bar{x}' =& c\bar{x} + \delta c_p - \delta c_B\bar{a}_{.p}.\end{aligned}$$

Tehát a célfüggvény-érték csökkenésének mértéke: $\delta(c_B\bar{A} - c)_p = \delta\bar{c}_p$. Azaz az x_p változó redukált költsége azt adja meg, hogy lokálisan mi a költsége a változó egységnyi növelésének. A szimplex módszer során olyan p-t választunk, amire $\bar{c}_p < 0$, így ez a csökkenés negatív, azaz a célfüggvény értéke nő (pontosabban nem csökken, mert majd látjuk, hogy $\delta = 0$ is előfordulhat), így minden lépésben az előzőnél jobb (azaz nem rosszabb) megoldást kapunk.

6.1.2. A szimplex módszer egy lépése

Feltesszük, hogy kiindulásként adott egy B primál megengedett bázis.

0. ha $\bar{c} \geq 0$, akkor kész vagyunk, hiszen a bázis optimális.

1. ha nem, válasszunk egy $p \in N$-t, amire $\bar{c}_p < 0$. Ezt többféleképpen megtehetjük:

   - Bland szabály: válasszuk a legkisebb ilyen p-t. Ez a választási módszer garantálja, hogy az algoritmusunk véges lesz (bizonyítás később).
   - válasszuk a legkisebb $\bar{c}_p$ értéket. Ez nem garantálja a végességet, de a gyakorlatban sokszor gyorsabb.

   Az így választott x_p kerül majd a bázisba.

2. Ha $\bar{a}_{.p} \leq 0$, akkor

   6.1.1. Állítás. *Ilyenkor a célfüggvény nem korlátos.*

   Bizonyítás.
   $$\bar{x}'_p = \bar{x}_p + \delta$$
   $$\bar{x}'_B = \bar{x}_B - \delta\bar{a}_{.p}$$
   ami tetszőleges $\delta \geq 0$-ra megengedett megoldást ad, mivel $\bar{a}_{.p} \leq 0$. A célfüggvényérték tehát szigorúan nő ($-\delta\bar{c}_p$-vel), azaz tetszőlegesen nagy lehet. •

3. Ha $\bar{a}_{\cdot p} \nleq 0$, akkor ki kell választani a bázisból kikerülő változót. Azt az r-et válasszuk, amire a
$$\frac{\bar{b}_r}{\bar{a}_{rp}} = \min_{i:\bar{a}_{ip}>0} \frac{\bar{b}_i}{\bar{a}_{ip}}$$

   Ekkor a szimplex tábla r. sorához tartozó bázisváltozó kerül ki a bázisból.

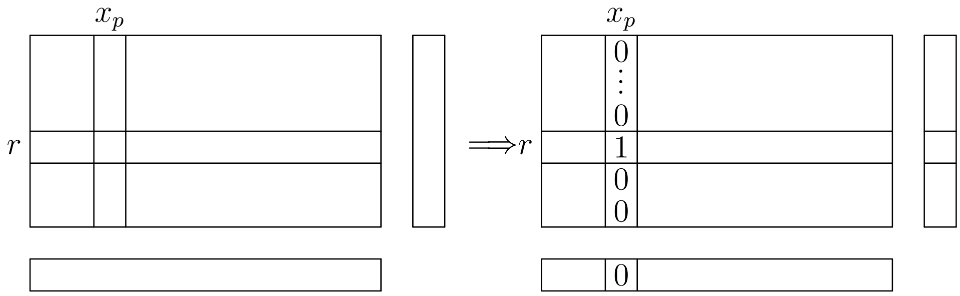
   Ha több i is minimális, akkor alkalmazzuk a Bland szabályt: az a bázisváltozó kerül ki, amelyiknek az indexe a legkisebb.

4. Új szimplex tábla kiszámítása (pivotálás): a szimplex tábla r. sorának többszöröseit adjuk hozzá a többi sorhoz. ($\bar{b}_r$ is hozzátartozik a sorhoz, és a redukált költség sorát is módosítjuk.)

$$\bar{a}'_{rj} = \frac{\bar{a}_{rj}}{\bar{a}_{rp}} \qquad \bar{b}'_r = \frac{\bar{b}_r}{\bar{a}_{rp}}$$

$$i \neq r: \bar{a}'_{ij} = \bar{a}_{ij} - \bar{a}_{rj}\frac{\bar{a}_{ip}}{\bar{a}_{rp}} \qquad \bar{b}'_i = \bar{b}_i - \bar{b}_r\frac{\bar{a}_{ip}}{\bar{a}_{rp}}$$

$$\bar{c}'_j = \bar{c}_j - \bar{a}_{rj}\frac{\bar{c}_p}{\bar{a}_{rp}}$$



6.1.4. Tétel. *A Bland szabályt használva a szimplex módszer véges sok lépésben véget ér.*

Bizonyítás. Ha egy lépésnél változik $\bar{x}$, akkor $c\bar{x}$ szigorúan nő. Ezért csak abból lehetne probléma, hogy végtelen ciklusba kerülünk, miközben $\bar{x}$ nem változik. Tegyük fel indirekt, hogy van egy ilyen ciklus, aminek tehát az elején és a végén ugyanaz a bázis van.

Egy indexet mozgónak nevezünk, ha a hozzá tartozó változó a ciklus során ki- illetve bekerül a bázisba. A nem mozgó indexek tehát a ciklus során vagy végig a bázisban vannak, vagy végig a bázison kívül.

Legyen p a legnagyobb mozgó index, és legyen t_1 egy olyan lépés, amikor bekerül, és t_2 egy olyan lépés, ahol kikerül. Feltehetjük, hogy $t_1 < t_2$. Jelölés: a t_1 lépés előtt: $B, \bar{B}, \bar{c}, \bar{A}$; a t_2 lépés előtt: $B', \bar{B}', \bar{c}', \bar{A}'$.

Mivel p kerül be a t_1-edik lépésben, $\bar{c}_p < 0$ és $j < p$ esetén $\bar{c}_j \geq 0$.

Nézzük most a t_2-edik lépést: legyen r az x_p bázisváltozóhoz tartozó sor, és legyen q az az index, ami bekerül a bázisba. Ekkor $\bar{c}'_q < 0$, $\bar{a}'_{rq} > 0$, és $\bar{a}'_{iq} \leq 0$ az összes olyan

i-re, ami mozgó bázisváltozóhoz tartozik. Az utóbbi azért igaz, mert ezekre az i-kre $\bar{b}'_i = 0$, és az ezekhez a sorokhoz tartozó változóknak p-nél kisebb az indexük.

A fent elmondottakból

$$0 < \bar{c}_q^{t_1} - \bar{c}_q^{t_2} = \bar{c}_B B^{-1} a_{.q} - \bar{c}_{B'} (B')^{-1} a_{.q} = (c_B B^{-1} B' - c_{B'}) \bar{a}'_{.q} = \bar{c}_{B'} \bar{a}'_{.q}.$$

De ha a jobb oldalon szereplő skalárszorzatot tagonként nézzük, a $\bar{c}_p \bar{a}'_{rq}$ tag szigorúan kisebb mint nulla, a többi mozgó indexhez tartozó tag legfeljebb 0, míg a nem mozgó indexekhez tartozó tagok értéke 0 (hiszen ha egy ilyen j index benne van B'-ben, akkor B-ben is benne van, tehát $\bar{c}_j = 0$).

Azt kaptuk, hogy $\bar{c}_{B'} \bar{a}'_{.q} < 0$, ellentmondás. •

6.1.3. Érzékenységvizsgálat

Legyen B optimális bázis. A gyakorlatban előforduló feladatoknál sokszor hasznos tudni, hogy a megoldásunk mennyire érzékeny a bemeneti adatok változásaira. Ebben a részben azt vizsgáljuk, hogy mennyire változtathatjuk meg a c-nek vagy b-nek egy adott koordinátáját, hogy B optimális maradjon.

Tudjuk, hogy B pontosan akkor optimális, ha $\bar{b} \geq 0$ (primál megengedett) és $\bar{c} \geq 0$ (duál megengedett).

Nem-bázis változó súlyának változtatása: $p \in N$-re: $c'_p = c_p + \delta$ valamilyen valós δ-ra.

- $\bar{b}$ nem változik, ezért B primál megengedett marad
- $\bar{c}' = \bar{y}A - c' = \underbrace{c_B \bar{A}}_{\text{nem változik}} - \underbrace{c'}_{\text{csak ez változik}}$

Tehát $\bar{c}'_j = \bar{c}_j$ ha $j \neq p$ és $\bar{c}'_p = \bar{c}_p - \delta$, vagyis B akkor és csak akkor marad optimális bázis, ha $\delta \leq \bar{c}_p$.

Bázisváltozó súlyának változtatása: Az r sorhoz tartozó bázisváltozó súlyát növeljük δ-val.

- $\bar{b}$ nem változik, ezért B primál megengedett marad
- $\bar{c}' = c'_B \overline{A} - c'$. Ekkor $\bar{c}'_B \equiv 0$ (ez mindig igaz), $\bar{c}'_N = \bar{c}_N + (\delta \bar{a}_{r.})_N$. Azaz $j \in N$ esetén $\bar{c}'_j = \bar{c}_j + \delta \bar{a}_{rj}$. Ez mikor marad nemnegatív?
  - Ha $\bar{a}_{rj} = 0$, akkor mindig.
  - Ha $\bar{a}_{rj} > 0$, akkor szükséges, hogy $\delta \geq -\frac{\bar{c}_j}{\bar{a}_{rj}}$
  - Ha $\bar{a}_{rj} < 0$, akkor szükséges, hogy $\delta \leq -\frac{\bar{c}_j}{\bar{a}_{rj}}$

Tehát

$$\bar{c}' \geq 0 \Leftrightarrow \max\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} > 0\} \leq \delta \leq \min\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} < 0\}.$$

Látjuk, hogy az alsó korlát egy nempozitív szám, a felső korlát egy nemnegatív szám, de mindkettő lehet nulla is. Továbbá ha üreshalmazon maximalizálunk, akkor $-\infty$ az alsó korlát, és ha üreshalmazon minimalizálunk, akkor $+\infty$ a felső korlát.

Jobboldal változtatása: Legyen $b'_r = b_r + \delta$. Ekkor

- $\bar{c}$ nem változik, ezért B duál megengedett marad.
- $\bar{b}' = B^{-1}b'$, tehát $\bar{b}'_i = B^{-1}_{i.}b' = \bar{b}_i + \delta B^{-1}_{ir}$.

Hasonlóan az előző esethez:

$$\bar{b}' \geq 0 \Leftrightarrow \max\{-\frac{\bar{b}_i}{B^{-1}_{ir}} : \bar{b}_i > 0\} \leq \delta \leq \min\{-\frac{\bar{b}_i}{B^{-1}_{ir}} : \bar{b}_i < 0\}.$$

Ha a feladatunkat egyenlőtlenség-rendszerből kaptuk kiegészítő változók hozzávételével, akkor B^{-1} és $\bar{y}$ is könnyen kiolvasható a szimplex táblából:

$$A: \begin{array}{|c|c|}\hline \quad\quad & \mathcal{I} \\ \hline \end{array} \qquad \bar{A}: \begin{array}{|c|c|}\hline \quad\quad & B^{-1} \\ \hline \end{array}$$

$$c: \begin{array}{|c|c|}\hline \quad\quad & 0\dots 0 \\ \hline \end{array} \qquad \bar{c}: \begin{array}{|c|c|}\hline \quad\quad & \bar{y} \\ \hline \end{array}$$

Nézzük meg, hogyan változik a célfüggvényérték a fenti változtatás során, ha B optimális bázis marad:

$$\bar{z}' = \bar{y}b' = \bar{z} + \delta\bar{y}_r.$$

Definíció. A $\bar{y}$ vektort **árnyékár** vektornak is nevezzük, mivel $\bar{y}_r$ meghatározza, hogy – élve a termelési feladat példájával – milyen egységáron érdemes az r-edik alapanyagból vásárolni (feltéve, hogy a vásárolt mennyiség a fenti határokon belül marad).

6.1.4. Módosított szimplex módszer

A szimplex módszer számítógépes implementációjakor nem érdemes a teljes szimplex táblát nyilvántartani. Vegyük észre, hogy a bázisba belépő x_p változó kiválasztásához csak a $\bar{c}$ vektorra van szükség. Ha ez megvan, a kilépő változót a $\bar{b}$ vektor és az $\bar{a}_{.p}$ oszlop segítségével határozzuk meg. Ez összesen $2m + n$ adat az $(m + 1) \times (n + 1)$-es szimplex táblából! Kérdés, hogy ezeket ki tudjuk-e számolni anélkül, hogy az egész táblát nyilvántartanánk.

A válasz az, hogy igen, feltéve hogy ismerjük az aktuális bázis inverzét. Valóban, az ismert képletek alapján

$$\begin{aligned}\bar{c} =& c_B B^{-1}A - c \\ \bar{b} =& B^{-1}b \\ \bar{a}_{.p} =& B^{-1}a_{.p}\end{aligned}$$

A **módosított szimplex módszer** lényege, hogy a szimplex tábla fenntartása helyett mindig csak az aktuális bázis inverzét számoljuk ki, és ennek segítségével számoljuk a fenti mennyiségeket. Az előző részben láttuk, hogy ha a feladatunkat egyenlőtlenség-rendszerből kaptuk kiegészítő változók hozzávételével, akkor B^{-1} nem más, mint a szimplex táblának a kiegészítő változókhoz tartozó része.

$$A: \begin{array}{|c|c|}\hline \quad\quad & \mathcal{I} \\ \hline \end{array} \qquad \bar{A}: \begin{array}{|c|c|}\hline \quad\quad & B^{-1} \\ \hline \end{array}$$

Tehát elég a táblának ezt az $m \times m$-es részét fenntartani; cserébe viszont a $\bar{c}, \bar{b}, \bar{a}_{.p}$ vektorok kiszámolásához mátrix-szorzás kell. Ha n jóval nagyobb mint m, akkor ez jelentős időmegtakarítást eredményez.

Általános esetben a módosított szimplex módszer az úgynevezett LU-felbontás segítségével valósítható meg hatékonyan.

6.2. Duál szimplex módszer

Ha kezdetben nem ismerünk primál megengedett bázist, de duál megengedettet igen, akkor használhatjuk a duál szimplex módszert. Mint később látni fogjuk, ez a helyzet például akkor, amikor egy megoldott feladatnál kiderül, hogy újabb feltételeket kell hozzávenni.

6.2.1. A duál szimplex módszer tulajdonságai

- duál megengedett bázisokon lépked
- minden lépésben egy oszlopot cserélünk ki B-ben.
- ugyanazt a szimplex táblát használjuk, mint a primál szimplex módszernél
- $c\bar{x}$ folyamatosan csökken (azaz nem nő)
- véges sok lépésben eljutunk egy primál megengedett bázishoz.

A fő különbség a primál szimplex módszerhez képest, hogy először a bázisból kilépő változót határozzuk meg, és csak utána a belépőt.

6.2.2. A duál szimplex módszer egy lépése

0. Ha $\bar{b} \geq 0$, akkor kész vagyunk. Primál megengedett bázisunk van, azaz optimális bázist találtunk.

1. Ha nem, válasszunk egy r-et, amire $\bar{b}_r < 0$. Ezt többféleképpen megtehetjük:

   - Bland szabály: válasszuk azt az r-et, amihez a legkisebb indexű bázisváltozó tartozik. Ez a választási módszer garantálja, hogy az algoritmusunk véges lesz.
   - Válasszuk a legkisebb $\bar{b}_r$ értéket. Ez nem garantálja a végességet, de a gyakorlatban gyorsabb.

   Az így választott r-hez tartozó bázisváltozó lép ki a bázisból.

2. Ha $\bar{a}_{r.} \geq 0$, akkor

   6.2.1. Állítás. *A primál feladatnak nincs megoldása.*

Bizonyítás.

$$\underbrace{\bar{a}_{r.}x}_{\geq 0} = \underbrace{\bar{b}_r}_{<0}$$

egy érvényes egyenlet lenne, ami nem lehetséges. •

3. Ha $\bar{a}_{r.} \not\geq 0$, akkor ki kell választani a bázisba bekerülő változót. Azt az x_p-t válasszuk, amire $\bar{a}_{rp} < 0$ és

$$-\frac{\bar{c}_p}{\bar{a}_{rp}} = \min\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} < 0\}$$

Megjegyzés. Ha ez a minimum nulla, akkor degeneráció lép fel, tehát $\bar{c}$ nem változik a báziscsere során.

Ha több p is minimális, akkor alkalmazzuk a Bland szabályt: válasszuk a minimális ilyen p-t. Az x_p változó kerül a bázisba.

Miért pont így kell választani a bemenő változót? Arra van szükségünk, hogy $\bar{c} \geq 0$ maradjon. Báziscsere után: $\bar{c}'_j = \bar{c}_j - \bar{a}_{rj}\frac{\bar{c}_p}{\bar{a}_{rp}}$ nemnegatív marad, ha

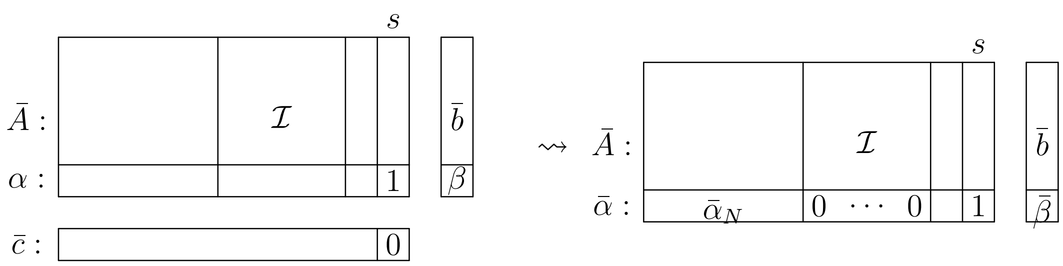
- $\bar{a}_{rj} \geq 0$, mivel $\bar{c}_j$-t ekkor növeljük
- $\bar{a}_{rj} < 0$, de $-\frac{\bar{c}_j}{\bar{a}_{rj}} \geq -\frac{\bar{c}_p}{\bar{a}_{rp}}$

4. A báziscsere ugyanúgy történik, mint a primál szimplex módszernél.

6.2.3. Alkalmazás: új feltétel hozzávétele

Tegyük fel, hogy már megoldottunk egy feladatot, és kiderül, hogy hozzá kell vennünk még a rendszerhez egy $\alpha x \leq \beta$ feltételt.

Vegyünk egy új kiegészítő változót: $s \geq 0$ úgy, hogy $\alpha x + s = \beta$. Írjuk át a feltételt ekvivalensen: $\bar{\alpha}x + s = \bar{\beta}$, ahol $\bar{\alpha}_B = 0$. Azaz:



Megjegyzés. $\bar{c}$-on nem kell változtatni, ugyanis a kiegészítő változóhoz 0 tartozik $\bar{c}$-ban.

A bázist kibővítjük s-sel, így az új feladatra egy duál megengedett bázist kapunk, és innentől fogva használhatjuk a duál szimplex módszert, mivel kaptunk egy kiindulási táblát.

6.2.4. Alkalmazás: primál megengedett bázis keresése

A duál szimplex módszert használhatjuk egy primál megengedett bázis megkeresésére is. Nézzük a következő primál feladatot és a hozzá tartozó duált:

$$(P)\quad \begin{aligned} Ax &= b \\ x &\geq 0 \\ \max\ &0x \end{aligned} \qquad\qquad (D)\quad \begin{aligned} yA &\geq 0 \\ \\ \min\ &yb \end{aligned}$$

A (P) feladatnak minden megengedett megoldása optimális, miközben a (D) feladatnak a $(0\dots 0)$ egy megengedett megoldása.

Ebben a feladatban $\bar{c} = \underbrace{c_B\bar{A}}_{=0} - \underbrace{c}_{=0} = 0$, tehát minden bázis duál-megengedett.

Induljunk ki tetszőleges bázisból, és használjuk a duál szimplex módszert. Ekkor vagy kapunk egy primál megengedett bázist, vagy kapunk egy bizonyítékot arra, hogy a feladat nem megoldható. (Ez a bizonyíték pont a Farkas lemmából következik.)

6.2.5. A duál szimplex módszer egy másfajta interpretációja

A duál szimplex módszer úgy is értelmezhető, hogy a 3.5.2. fejezetben leírt megengedettségi szimplex módszert használjuk szubrutinként az optimalizálási feladat megoldására.

Jelölje $R := \{x : Ax = b, x \geq 0)$ a primál, $R^* := \{y : yA \geq c\}$ pedig a duál poliédert. Feltesszük, hogy A sorai lineárisan függetlenek, ami azt jelenti, hogy R^* csúcsos.

A megengedettségre vonatkozó szimplex algoritmussal először megkeresünk R^*-nak egy $\bar{y}$ csúcsát (azaz $yA \geq c$ egy bázis-megoldását). Amennyiben R^* üres, úgy az eljárás egy olyan $x' \geq 0$ vektort szolgáltat, amelyre $Ax' = 0$, $cx' > 0$, és ilyenkor vagy a primál poliéder is üres, vagy ha van is egy x pontja, akkor $x + \lambda x'$ minden pozitív λ-ra R-ben van, így a célfüggvényérték nem korlátos alulról. Ekkor tehát az algoritmus futása befejeződik.

Tegyük fel tehát, hogy rendelkezésünkre áll $\bar{y}$. Jelölje $A^=$ az A-nak azon $a_{.j}$ oszlopai által alkotott részmátrixát, amelyekre $\bar{y}a_{.j} = c_j$ (vegyük észre, hogy $A^=$ tartalmaz bázist), míg a maradék oszlopok részmátrixa legyen $A^<$. A 3.5.2. fejezetben leírt eljárással döntsük el, hogy az $\{A^=x' = b,\ x' \geq 0\}$ rendszernek létezik-e megoldása. Amennyiben létezik, úgy x'-t nulla komponensekkel kiegészítve R-nek egy olyan $\bar{x}$ elemét kapjuk, amely teljesíti az optimalitási feltételeket (azaz, ha valamely j-re $\bar{x}_j$ szigorúan pozitív, akkor $\bar{y}a_j = c_j$). Ekkor $\bar{x}$ primál optimum, $\bar{y}$ duál optimum és az eljárás véget ér.

Ha a szóbanforgó x' nem létezik, akkor a 3.5.2. fejezet eljárása megtalálja $A^=$-nek egy $m-1$ lineárisan független oszlopból álló A' részmátrixát valamint egy olyan y' vektort, amelyekre $y'A^= \geq 0, y'A' = 0$ és $y'b < 0$.

Amennyiben $y'A^< \geq 0$, úgy az adódik, hogy $y'A \geq 0$, $y'b < 0$ és így (a Farkas lemma triviális irányát alkalmazva) a primál feladat nem megoldható, vagy ekvivalensen a duál feladat nem korlátos. Ilyenkor az algoritmus véget ér.

Tegyük most fel, hogy $y'A^< \not\geq 0$. Válasszuk λ-t a legnagyobb olyan számnak, amelyre $(\bar{y} + \lambda y')A \geq c$ teljesül, azaz $(\bar{y} + \lambda y')a_{.j} \geq c_j$ fennáll az $A^<$ mindegyik $a_{.j}$ oszlopára. Vagyis λ a legnagyobb szám, amelyre $\lambda y'a_{.j} \geq c_j - \bar{y}a_{.j}$ teljesül az $A^<$ valamennyi olyan oszlopára, amelyre $y'a_{.j} < 0$.

Legyen $\bar{y}' = \bar{y} + \lambda y'$. A λ választásából adódóan $\bar{y}'$ eleme R^*-nak.

6.2.1. Lemma. *$\bar{y}'$ csúcsa R^*-nak.*

Biz. Azt kell látni, hogy A-nak van m lineárisan független oszlopa, melyekre $\bar{y}'a_{.j} = c_j$. Mindenesetre $y'A' = 0$ miatt A'-nek az $m-1$ oszlopa ilyen. Legyen $a_{.q}$ egy olyan oszlop, ahol a λ definíciójában szereplő minimum felvétetik. Ekkor nyilván $\bar{y}'a_{.q} = c_j$, így csak azt kell látnunk, hogy $a_{.q}$ lineárisan független az A' oszlopaitól. De ez valóban így van, hiszen $y'A' = 0$ és $y'a_{.q} \neq 0$. •

Az $\bar{y}'$ tehát valóban csúcsa R^*-nak, és ráadásul $\bar{y}$-nál jobb csúcsa, hiszen $y'b < 0$ miatt $\bar{y}'b < \bar{y}b$. Miután R^*-nak véges sok csúcsa van, az eljárás véges sok iteráció után befejeződik.

Az algoritmusban a Farkas lemmára vonatkozó algoritmust szubrutinként használtuk, aminek belsejében persze alkalmazzuk a Bland féle legkisebb index szabályt. A fenti algoritmusban azonban, amikor az $\bar{y}$-ról áttértünk $\bar{y}'$-re, a szóban forgó $a_{.q}$ oszlop meghatározásánál nem volt szükség a Bland-szabályra.

6.3. Kétfázisú szimplex módszer

A gyakorlatban, ha nem áll rendelkezésre kezdeti primál megengedett bázis, az úgynevezett kétfázisú szimplex módszert szoktuk használni. Ennek első fázisában primál megengedett bázist keresünk, míg a második fázisban ebből a bázisból kiindulva alkalmazzuk a primál szimplex módszert.

Nézzük az **első fázist**. Legyen a feladat: $\max\{cx : \ Ax = b, x \geq 0\}$. Feltehető, hogy $b \geq 0$, mert egyenleteket szorozhatunk (-1)-gyel. Vezessünk be minden sorhoz új mesterséges változókat: u_i $(i = 1 \dots m)$.

$$\begin{aligned} Ax + Iu &= b \\ (x, u) &\geq 0 \end{aligned}$$

A fenti rendszer nem ekvivalens az eredetivel. Ahhoz, hogy az eredeti feladat megoldását kapjuk, olyan megoldást kell keresni, ahol $u = 0$. Ennek megtalálásához az első fázisban legyen a célfüggvény $\max - \sum_{i=1}^{m} u_i$.

- Ha ennek a feladatnak 0 az optimuma, akkor az eredeti feladat egy megoldását kaptuk.
- Ha ennek a feladatnak < 0 az optimuma, akkor az eredeti feladatnak nincs megoldása.

Legyen B a mesterséges változók oszlopaiból álló bázis. Ekkor B primál megengedett, tehát használható az első fázisban kiindulási bázisként.

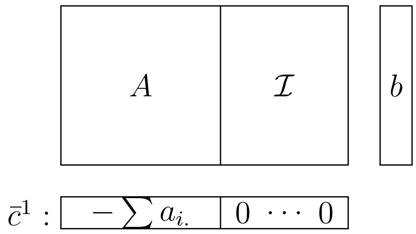
$$A^1 : \boxed{\quad A \quad \Big| \quad \mathcal{I} \quad} \quad \boxed{b}$$

$$c^1 : \boxed{0 \ \cdots \ 0 \,\big|\, -1 \cdots -1}$$

$$\begin{aligned}
\bar{A}^1 &= A^1 \\
\bar{b}^1 &= b \\
\bar{c}^1 &= c_B^1 \bar{A}^1 - c^1
\end{aligned}$$

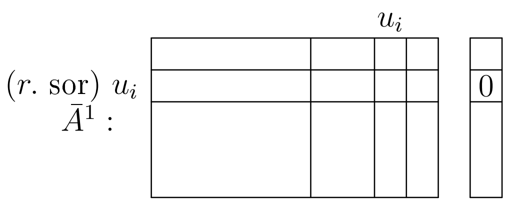
Tudjuk, hogy $\bar{c}_B^1 = 0$, és $\bar{c}_N^1 = -\sum_{i=1}^m a_{i.}$.

A kiindulási szimplex tábla tehát:

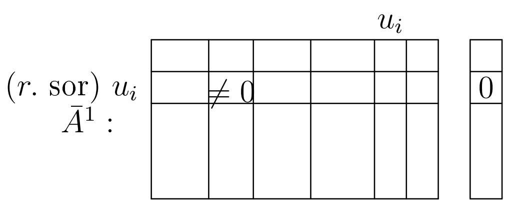


Erre a szimplex táblára kell alkalmazni a primál szimplex módszert.

- Ha az optimum < 0, akkor nincs megoldás.
- Ha az optimum $= 0$, de marad mesterséges változó a bázisban (azaz a feladatunk degenerált volt), akkor tegyük a következőt a mesterséges változók kiküszöbölése érdekében:



Legyen a megmaradt változó a szimplex tábla r-edik sorában. Ez a sor nem lehet azonosan 0, mert A sorai lineárisan függetlenek. Válasszunk tehát egy tetszőleges nemnulla elemet egy nem-mesterséges változó oszlopában, és ott pivotáljunk:



Mivel a jobboldalon 0 van, ezért a primál megoldás nem változik, és a célfüggvényérték sem változik.

- Ha az optimum $= 0$ és nincs mesterséges változó a bázisban (azaz az eredeti feladatra kaptunk egy megengedett bázist), akkor áttérhetünk a **második fázisra:** elhagyjuk a mesterséges változókat, és az eredeti c célfüggvényre alkalmazzuk a szimplex módszert.

6.4. Hálózati szimplex módszer

Adott egy $D = (V, E)$ irányított, gyengén összefüggő gráf, $c : E \to \mathbb{R}$ élsúlyokkal. Ezen túl adott egy $b : V \to \mathbb{Z}$ igényfüggvény, amelyik minden csúcsra előírja, hogy mennyi legyen a bemenő és a kimenő folyam különbsége. Feltesszük, hogy $\sum_{v \in V} b_v = 0$.

Legyen $x : E \to \mathbb{R}$ a változók vektora.

Jelölés (Bemenő folyam). $\varrho_x(v)$ jelöli a v csúcsba belépő éleken az x-ek összegét.

Jelölés (Kimenő folyam). $\delta_x(v)$ jelöli a v csúcsból kilépő éleken az x-ek összegét.

A következő alakú feladatot szeretnénk megoldani:

$$\begin{aligned} \varrho_x(v) - \delta_x(v) = b_v \qquad & \forall v \in V \\ x \geq 0 \\ \max \sum_{e \in E} c_e x_e \end{aligned}$$

Megjegyzés. Ha $\sum b_v = 0$ nem lenne igaz, akkor a feladatnak nem lenne megoldása.

Legyen v_0 egy kijelölt csúcs. Ha $\varrho_x(v) - \delta_x(v) = b_v$ minden $v \in V \setminus \{v_0\}$-ra teljesül, akkor v_0-ra is teljesül. Vezessük be tehát az eredeti egyenlőségrendszer helyett a következőt: $\varrho_x(v) - \delta_x(v) = b_v \ \forall v \in V \setminus \{v_0\}$. Így a sorok lineárisan függetlenek lesznek, ezáltal a feladatra alkalmazhatjuk a szimplex módszert: $Ax = b, \ x \geq 0$, ahol A sorai lineárisan függetlenek. Az A mátrixunk a következő lesz:

$$V = \{v_0, v_1, \dots, v_n\}, \ E = \{e_1, \dots, e_m\}$$

$$A \in \mathbb{Z}^{nxm}$$

$$a_{ij} = \begin{cases} -1 & \text{ha } v_i \text{ töve } e_j\text{-nek} \\ +1 & \text{ha } v_i \text{ feje } e_j\text{-nek} \\ \ \ 0 & \text{különben} \end{cases}$$

Megjegyzés. Az A mátrix minden oszlopában legfeljebb egy $+1$-es és legfeljebb egy -1-es található, ezért hálózati mátrix.

Jelölés. A továbbiakban kontextustól függően a következő ekvivalens jelöléseket fogjuk használni:

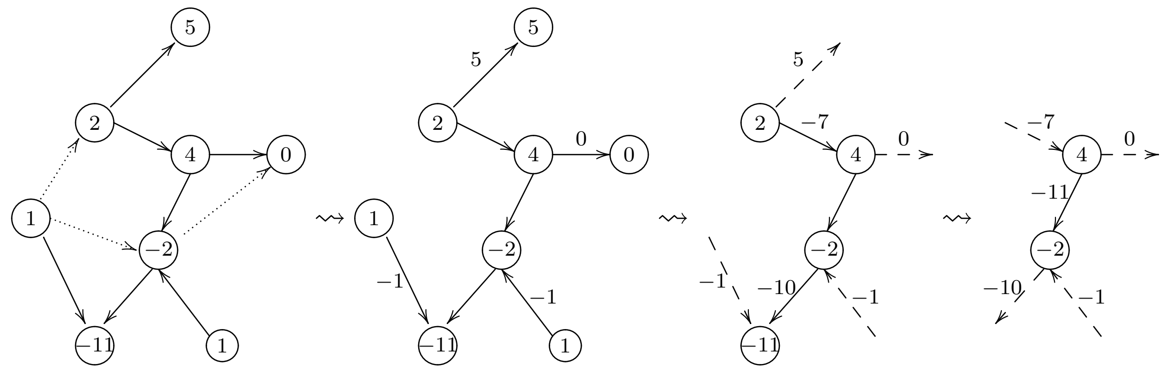
$$\begin{aligned} b_{v_i} &\sim b_i \\ c_{e_j} &\sim c_j \\ y_{v_i} &\sim y_i \\ x_{v_i} &\sim x_i \end{aligned}$$

A feladatunk tehát felírható $\max\{cx, Ax = b, x \geq 0\}$ alakban.

6.4.1. Állítás. *B bázis $\Leftrightarrow$ $\{e_j : j \in B\}$ feszítőfa (irányítás nélkül).*

Bizonyítás. $\Leftarrow$: Belátjuk, hogy $Bx = b$ egyértelműen megoldható. Keressünk a fán olyan folyamot, ami minden igényt kielégít, azaz minden élre adjunk olyan értéket, ahol $\varrho_x(v) - \delta_x(v) = b_v$. A fa leveleire egy-egy él illeszkedik. Ezekre egyértelműen meg tudjuk adni a változó értékét.

Ha a leveleket elhagyjuk, akkor újabb fát kapunk, amely fa leveleire illeszkedő élekre ugyancsak egyértelműen meghatározható a változó értéke, és így tovább.



$\Rightarrow$: Indirekt bebizonyítjuk, hogy ha a B-hez tartozó élek nem alkotnak feszítőfát, akkor B nem bázis. Adott tehát n darab él, ami nem alkot feszítőfát. Ekkor a részgráf tartalmaz kört. Legyen ez a kör C, és legyen

$$x_e = \begin{cases} +1 & \text{, ha } e \in C \text{ előre-él} \\ -1 & \text{, ha } e \in C \text{ hátra-él} \\ 0 & \text{, ha } e \notin C \end{cases}$$

Ekkor a $Bx = 0$ és $x \neq 0$, tehát B szinguláris, azaz B nem bázis. •

Az alábbiakban ismertetett hálózati szimplex módszer tulajdonképpen egyszerűen a szimplex módszer alkalmazása a feladatunkra, de mátrixok helyett gráfelméleti fogalmakkal elmondva. Amint látni fogjuk, ennek előnye, hogy az algoritmus során nem kell szorzást és osztást végezni, csak összeadást és kivonást, ezért nem merülhetnek fel numerikus pontatlanságok.

Legyen $\bar{x}$ a $Bx = b$ egyértelmű megoldása és legyen $\bar{y}$ a következő: $\bar{y}_0 = 0$ lesz a v_0-hoz tartozó duális változó. Ha $uv \in B$, akkor legyen $\bar{y}_v - \bar{y}_u = c_{uv}$; ez egyértelműen meghatározza $\bar{y}$-t (v_0-ból kiindulva kiszámolható). A továbbiakban az egyszerűség kedvéért B-vel jelöljük az $\{e_j : j \in B\}$ feszítőfát is.

Az uv él redukált költsége: $\bar{c}_{uv} = \bar{y}_v - \bar{y}_u - c_{uv}$. A korábbiaknak megfelelően a B bázis primál megengedett, ha $\bar{x} \geq 0$, és duál megengedett, ha $\bar{c} \geq 0$.

6.4.1. Primál hálózati szimplex módszer lépései

Tegyük fel, hogy B primál megengedett bázis. Az alábbi lépéssorozatnál nem kell az A értékeivel műveleteket végezni, csak az $\bar{x}, \bar{y}, \bar{c}$ vektorokkal.

0. Ha $\bar{c} \geq 0$, akkor kész vagyunk (a bázisunk primál és duál megengedett).

1. Ha nem, akkor válasszunk egy olyan e_p élt, amire $\bar{c}_p < 0$. Az így választott e_p él kerül majd a bázisba.

2. Vegyük hozzá a B feszítőfához az e_p élt. Ekkor egy egyértelmű C kört kapunk, aminek e_p éle. Nevezzük a C-ben e_p-vel egyirányú éleket előreéleknek, a többi C-beli élt pedig hátraélnek.

   Ha C-ben nincsenek hátraélek, akkor tetszőleges $\delta > 0$-ra

$$\bar{x}' = \begin{cases} \bar{x}_e + \delta, \text{ ha } e \in C \\ \bar{x}_e, \text{ különben} \end{cases}$$

   megengedett megoldás.

   6.4.2. Állítás. *Ebben az esetben a célfüggvény nem korlátos.*

   Bizonyítás. $\sum_{uv \in C} \bar{c}_{uv} = \sum_{uv \in C}(\bar{y}_v - \bar{y}_u - c_{uv}) = -\sum_{uv \in C} c_{uv}$, tehát

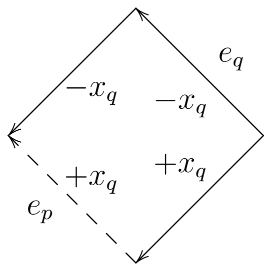
$$c\bar{x}' = c\bar{x} + \delta \sum_{uv \in C} c_{uv} = c\bar{x} - \delta \sum_{uv \in C} \bar{c}_{uv} = c\bar{x} - \delta \bar{c}_p \xrightarrow[\delta \to \infty]{} +\infty$$

   •

3. Ha van C-ben hátraél, akkor legyen e_q az a hátraél, amire $\bar{x}_q$ minimális. Ez az él fog kikerülni a bázisból.

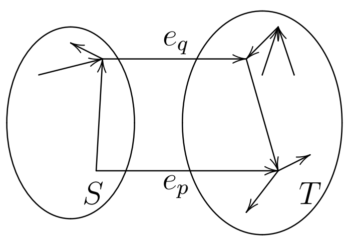
4. Vegyük hozzá a B bázishoz az e_p élt, és hagyjuk el a bázisból az e_q élt: $B' = B + \{p\} - \{q\}$.

$$\bar{x}'_j = \begin{cases} \bar{x}_j + \bar{x}_q, \text{ ha } e_j \text{ előreél } C\text{-ben} \\ \bar{x}_j - \bar{x}_q, \text{ ha } e_j \text{ hátraél } C\text{-ben} \\ \bar{x}_j, \text{ ha } e_j \notin C \end{cases}$$



   Számoljuk ki az $\bar{y}'$-t:

   Ha az e_p élt kihagyjuk a fából, akkor a fa két komponensre esik: S és T. Válasszuk S-t és T-t úgy, hogy e_p a T-be lépjen.



   Ekkor

   - Ha $v_o \in S$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v, \text{ ha } v \in S \\ \bar{y}_v - \bar{c}_p, \text{ ha } v \in T \end{cases}$$

- Ha $v_o \in T$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v + \bar{c}_p, \text{ ha } v \in S \\ \bar{y}_v, \text{ ha } v \in T \end{cases}$$

A fenti megkülönböztetés azért szükséges, hogy $y_0 = 0$ maradjon.

Megjegyzés. Ha b és c egészek, akkor $\bar{y}$, $\bar{x}$ és $\bar{c}$ végig egészek maradnak. Sőt, ha a költségek(súlyok) egészek, a duál végig egész lesz, és ha az igények egészek, a primál végig egész lesz. Tehát egész igények esetén a hálózati feladatnak mindig van egész optimális megoldása, ha egyáltalán megoldható. A ciklizálás elkerülése érdekében használhatjuk a Bland szabályt.

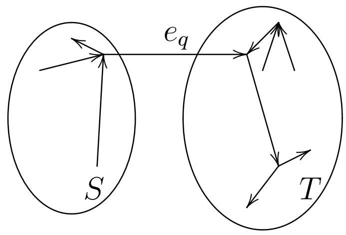
6.4.2. Duál hálózati szimplex módszer

Természetesen a hálózati szimplex módszernek is van duál változata. Tegyük fel, hogy kezdetben van egy B duál megengedett bázis (azaz $\bar{c} \geq 0$). Az általános lépés a következő:

0. Ha $\bar{x} \geq 0$, akkor kész vagyunk (a bázisunk primál és duál megengedett).

1. Ha nem, akkor legyen $q \in B$ olyan, hogy $\bar{x}_q < 0$. Ha több alternatívánk van, alkalmazzuk a Bland szabályt: válasszuk a legkisebb indexű ilyet. Az így választott változó kerül majd ki a bázisból.

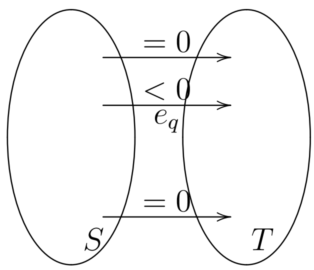
   Ha elhagyjuk az e_q élt, akkor a fa két részre esik: S és T. Válasszuk S-t és T-t úgy, hogy e_q a T-be lépjen.



2.

6.4.3. Állítás. *Ha nincs T-ből S-be vezető él, akkor a primál feladatnak nincs megoldása.*

Bizonyítás. Az $\bar{x}$ aktuális primál vektor a $\varrho_x(v) - \delta_x(v) = b_v$, $\forall v \in V$ feladat megoldása. Az e_q élen: $\bar{x} < 0$. A többi S-ből T-be vezető élen $\bar{x} = 0$, mivel ezek nincsenek a bázisban.



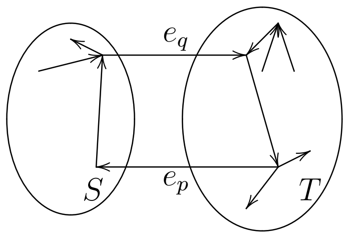
Emiatt S igénye nagyobb T igényénél, hiszen

$$\sum_{v\in T} b_v - \sum_{v\in S} b_v = \sum_{v\in T}(\varrho_{\bar{x}}(v) - \delta_{\bar{x}}(v)) - \sum_{v\in S}(\varrho_{\bar{x}}(v) - \delta_{\bar{x}}(v)) = 2 \sum_{e:S\to T} \bar{x}_e < 0.$$

Nemnegatív folyammal $T \to S$ élek híján nem lehet az igényeket kielégíteni, tehát nincs megoldás. •

3. Ha van T-ből S-be él, akkor válasszuk ki azt az e_p $T \to S$ élt, amire $\bar{c}_p = \min\{\bar{c}_e$, ahol $e : T \to S$ él$\}$. Ez az él fog bekerülni a bázisba.

4. Hagyjuk el a B bázisból az e_q élt, és vegyük hozzá az e_p élt: $B' = B - \{q\} + \{p\}$.



Az $\bar{x}'$ kiszámolásához legyen C a $B \cup \{e_p\}$ egyetlen köre. Nevezzük az e_p-vel egyirányú éleket előreéleknek, a vele ellentétes irányú éleket pedig hátraéleknek. Ekkor

$$\bar{x}'_e = \begin{cases} \bar{x}_e - \bar{x}_q, \text{ ha } e \text{ előreél } C\text{-ben} \\ \bar{x}_e + \bar{x}_q, \text{ ha } e \text{ hátraél } C\text{-ben} \\ \bar{x}_e, \text{ ha } e \notin C \end{cases}$$

Számoljuk ki az $\bar{y}'$-t:

- Ha $v_o \in S$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v, \text{ ha } v \in S \\ \bar{y}_v + \bar{c}_p, \text{ ha } v \in T \end{cases}$$

- Ha $v_o \in T$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v - \bar{c}_p, \text{ ha } v \in S \\ \bar{y}_v, \text{ ha } v \in T \end{cases}$$

6.4.3. Kezdeti primál bázis keresése

Több módszert is ismertetünk:

1. A $c \equiv 0$ súlyfüggvényre alkalmazzuk a duál hálózati szimplex módszert. Ilyenkor tetszőleges B bázisra $\bar{y} \equiv 0$, $\bar{c} \equiv 0$. A gyakorlatban ez lassú módszer, és csak azért nem ciklizál, mert a Bland szabályt alkalmazzuk.

2. Legyen B tetszőleges bázis. Ha ez a bázis se nem primál-, se nem duál megengedett, akkor módosítsuk c-t úgy, hogy duál megengedett legyen:

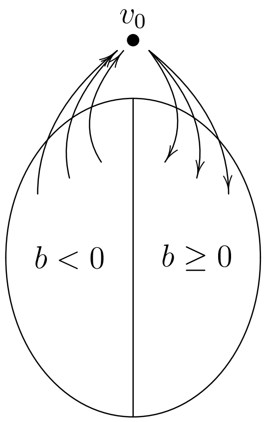
$$c'_{uv} = \begin{cases} c_{uv}, \text{ ha } \bar{c}_{uv} \geq 0 \\ \bar{y}_v - \bar{y}_u, \text{ ha } \bar{c}_{uv} < 0 \end{cases}$$

Ezzel a c'-vel B duál megengedett lesz. Most alkalmazzuk a duál hálózati szimplex módszert, ezáltal B' optimális bázist kapunk c'-re, ami primál megengedett bázis az eredeti c célfüggvényre. Ezzel a B'-vel kezdve alkalmazzuk a primál szimplex módszert az eredeti c célfüggvényre.

3. **Kétfázisú hálózati szimplex módszer**

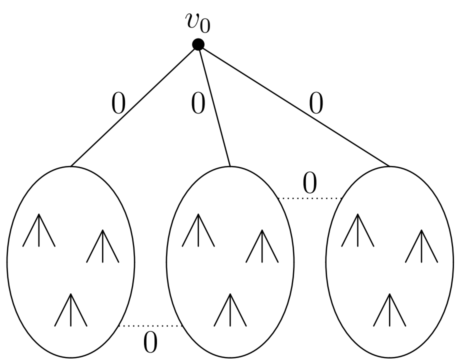
   Vegyük hozzá az eredeti gráfhoz a következő új éleket: $E' = E \cup \{v_0 v : b_v \geq 0\} \cup \{v v_0 : v_v < 0\}$. Legyen

$$c'_e = \begin{cases} -1, \text{ ha } e \in E' \setminus E \\ 0, \text{ ha } e \in E. \end{cases}$$



   Az új élek primál megengedett bázist határoznak meg, mivel $\bar{x}_{v_0 v} = b_v$ és $\bar{x}_{v v_0} = -b_v$, tehát $\bar{x} \geq 0$. Alkalmazzuk a primál hálózati szimplex módszert. Ekkor

   - Ha az optimum negatív, akkor az eredeti feladatnak nincs megoldása.
   - ha az optimum nulla, akkor hagyjuk el az $E' \setminus E$-beli éleket, és az így keletkezett részfákat egészítsük ki feszítőfává 0-ás éleket hozzávéve. Ez lehetséges, mert az eredeti gráf gyengén összefüggő. Így primál megengedett bázist kapunk az eredeti feladatra.



   Második fázis: alkalmazzuk ezzel a kiinduló feszítőfával a hálózati szimplex módszert az eredeti célfüggvényre.

Megjegyzés. A hálózati szimplex módszer kis módosítással használható arra az általánosabb feladatra is, ahol minden élre adott egy alsó és egy felső korlát az él értékére.

Ez már magában foglalja a minimális költségű maximális folyam (1.6.4 fejezet) és a minimális költségű áram feladatot. Fordítva is igaz, hogy a 1.6.4 fejezetben leírt algoritmus is használható a jelen fejezetben vizsgált feladat megoldására. A tapasztalatok azonban azt mutatják, hogy a gyakorlatban felmerülő súlyozott folyam feladatoknál sokszor a hálózati szimplex módszer különféle változatai közül kerül ki a leghatékonyabb algoritmus.

6.4.4. Erősen megengedett bázisok

A hálózati szimplex módszer esetében van a Bland szabálynál természetesebb és hatékonyabb pivotálási szabály, ami garantálja az algoritmus végességét. Ehhez azonban be kell vezetni az erősen megengedett bázis fogalmát.

Definíció. Egy primál megengedett B bázis **erősen megengedett**, ha a feszítő fa összes olyan uv élére, amire $\bar{x}_{uv} = 0$, v közelebb van a fában v_0-hoz mint u.

6.4.4. Állítás. *Ha a hálózati feladatnak van megoldása, de nincs erősen megengedett bázisa, akkor szétbontható két részfeladatra.*

Bizonyítás. Adott bázisnál nevezzünk egy uv élt **tiltottnak**, ha $\bar{x}_{uv} = 0$, és u közelebb van a fában v_0-hoz mint v. Vegyük azt a B primál megengedett bázist, ahol irányítatlan értelemben a legtöbb csúcs elérhető v_0-ból nem tiltott élen, és legyen U az elérhető pontok halmaza. Ha U-ba belépne D-nek egy e éle, akkor e-t hozzávéve a bázishoz, és egy U-ból kilépő, $B + e$ körén lévő tiltott élt kihagyva olyan bázist kapnánk, ahol U-nál nagyobb halmaz érhető el v_0-ból nem tiltott élen. Mivel ez nem lehet, U-ba nem lép be él D-ben, és az összes kilépő élre $\bar{x}_e = 0$. Ez viszont azt jelenti, hogy teszőleges x megengedett megoldásban $x_e = 0$ az összes U-ból kilépő élen, tehát a feladat szétbontható a $D[U]$ és a $D[V \setminus U]$ digráfokon értelmezett feladatra. •

Tegyük fel tehát, hogy kiindulásként adott egy erősen megengedett bázis. A primál hálózati szimplex módszer lépését a következőképpen változtatjuk meg. Tegyük fel, hogy $e_p = uv$ lép be a bázisba, és C a keletkező kör. Legyen v_C a kör v_0-hoz legközelebbi pontja. Legyen $e_q = u'v'$ az a hátra-él, amin $\bar{x}_q$ minimális, és ezek közül az, ami v_C-től előre-irányba elindulva a legutolsó a körön.

6.4.5. Állítás. *A* $B' = B + \{p\} - \{q\}$ *bázis erősen megengedett.*

Bizonyítás. Két esetet különböztetünk meg.

1. eset: $\bar{x}_q = 0$. Ekkor, mivel B erősen megengedett, e_q a $v_C u$ szakaszon van, és ezen a szakaszon az utolsó olyan hátra-él, amire $\bar{x}_e = 0$. Ezért a báziscsere után sem keletkezik tiltott él, hiszen az csak az $u'u$ szakaszon lévő 0-ás hátra-élekből keletkezhetne.

2. eset: $\bar{x}_q > 0$. Ekkor a körön lévő 0-ás élek a báziscsere után már nem 0-ásak, tehát csak olyan e hátra-él válhat tiltottá, amire $\bar{x}_e = \bar{x}_q$. Az ilyen hátra-élek a $v_C v'$ szakaszon vannak. Ha e_q a $v_C u$ szakaszon van, akkor ezek az élek a báziscsere után is v_0 fele mutatnak. Ha pedig e_q a $v v_C$ szakaszon van, akkor a báziscsere során pontosan akkor fordulnak meg, ha előtte nem v_0 fele mutattak, tehát a báziscsere után v_0 fele mutatnak. •

Most belátjuk, hogy ezzel a választással nem lehet ciklizálás.

6.4.6. Állítás. *Ha a fenti báziscsere során $\bar{x}_q = 0$, akkor $\sum_{v \in V} \bar{y}_v$ szigorúan csökken.*

Bizonyítás. Ebben az esetben e_q a $v_C u$ szakaszon van, tehát $\bar{y}_v$ az $u'u$ szakasz pontjaiban (és a belőlük induló részfákon) változik, mégpedig $\bar{c}_p$-vel nő, azaz szigorúan csökken. •

Mivel $\sum_{v \in V} \bar{y}_v$ szigorúan csökken ha a célfüggvényérték nem nő, az algoritmus során nem térhetünk vissza ugyanahhoz a bázishoz, tehát nem lehet ciklizálás.

7. fejezet

Egészértékű lineáris programozás

7.1. Bevezetés

A következő alapfeladattal foglalkozunk:

$$\begin{aligned} &\max cx \\ &Ax \leq b \\ &x \in \mathbb{Z}^n \end{aligned}$$

A fenti feladat LP-relaxáltja:

$$\begin{aligned} &\max cx \\ &Ax \leq b \\ &x \in \mathbb{R}^n \end{aligned}$$

Speciális eset: bináris programozási feladat:

$$\max\{cx : Ax \leq b, x \in \{0,1\}^n\}$$

Példa: Bináris hátizsák feladat

Adott n db tárgy. A j-edik tárgy értéke legyen $c_j \geq 0$, súlya pedig $a_j > 0$. Adott továbbá egy hátizsák, melynek teherbíró képessége $b > 0$. A lehető legnagyobb összértékű tárgyat akarunk a hátizsákba pakolni úgy, hogy az még elbírja őket. Azaz:

$$\max\{\sum_{j=1}^{n} c_j x_j : \sum_{j=1}^{n} a_j x_j \leq b,\ x \in \{0,1\}^n\}$$

A feladat NP-nehéz, tehát nem várható rá polinom idejű algoritmus. Az LP relaxáltja könnyen megoldható: az érték/súly arány szerint helyezzük csökkenő sorrendbe a tárgyakat. Sorra tegyük be őket a hátizsákba, amíg a hátizsák be nem telik (az utolsó betett tárgy lehet hogy csak részben fér be, de a relaxált feladatnál ez nem baj).

7.1.1. Állítás. *Ez a mohó algoritmus optimálisan megoldja az LP relaxáltat.*

Bizonyítás. Az algoritmus teljesen megtölti a hátizsákot. Mivel érték/súly arány szerinti csökkenő sorrendben vettük a tárgyakat, nyilvánvaló, hogy minden olyan tárgynak, ami (egészen vagy részben) kimaradt, legfeljebb annyi az érték/súly aránya, mint bárminek, ami (egészen vagy részben) bekerült. Ezért semmilyen bent lévő rész kicserélésével nem járhatunk jobban, tehát a megoldás optimális. •

Az egészértékű programozási feladatnál kicsit általánosabb a **vegyes programozási feladat**, ahol nem feltétlenül az összes változónak kell egészértékűnek lenni:

$$\max\{cx + dz : \; Ax + Bz \leq b, x \in \mathbb{Z}^{n_1}, z \in \mathbb{R}^{n_2}\}.$$

Példa: Szolgáltató-elhelyezési feladat

Adott m db ügyfél, és n db lehetséges szolgáltatóhely. Minden ügyfélnek egységnyi igénye van, ezt megoszthatjuk több szolgáltatóhely között.

c_{ij}: i-edik ügyfél kiszolgálásának költsége a j-edik szolgáltatóhelyről.

f_j: j-edik szolgáltatóhely megnyitásának költsége

u_j: j-edik szolgáltatóhely kapacitása

Ez a feladat is NP-nehéz. Vegyes programozási feladatként úgy tudjuk felírni, hogy minden szolgáltatóhelyhez bevezetünk egy bináris változót (y_j), ami azt „dönti el”, hogy megnyitjuk-e a szolgáltatóhelyet:

$$\begin{aligned}
\min \sum_{j=1}^{n}(f_j y_j + \sum_{i=1}^{m} c_{ij}x_{ij}) & & \\
0 \leq x_{ij} \leq y_j & \qquad & i \in \{1,\dots,m\},\ j \in \{1,\dots,n\} \\
y_j \in \{0,1\} & & j \in \{1,\dots,n\} \\
\sum_{j=1}^{n} x_{ij} = 1 & & i \in \{1,\dots,m\} \\
\sum_{i=1}^{m} x_{ij} \leq u_j y_j & & j \in \{1,\dots,n\}.
\end{aligned}$$

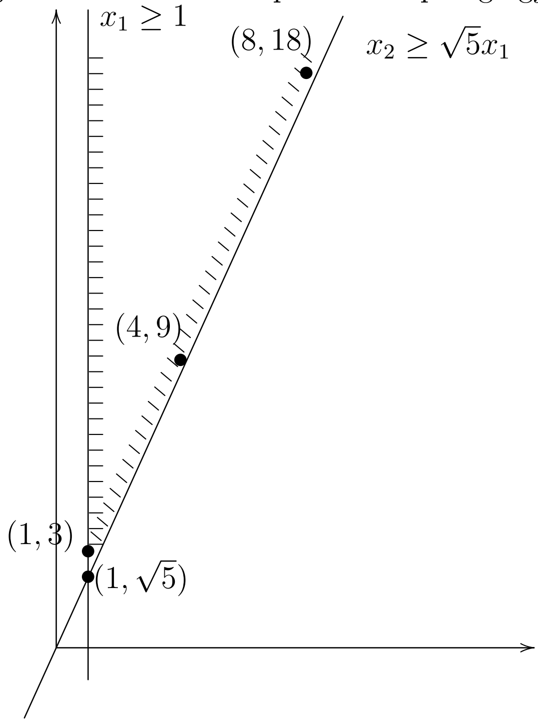
Észrevétel. *Ha rögzítjük, hogy melyik szolgáltatóhelyeket nyitjuk meg, akkor az optimális ügyfél-hozzárendelés feladatában a feltételi mátrix TU, tehát egész kapacitások esetén a szolgáltató-elhelyezési feladatnak van olyan optimális megoldása, ahol minden ügyfelet 1 szolgáltatóhoz rendelünk.*

Ha egy egészértékű programozási feladatot meg akarunk oldani, akkor tulajdonképpen egy poliéder egész pontjainak konvex burkán akarunk egy lineáris célfüggvényt maximalizálni.

Definíció. Legyen P poliéder, $P \subseteq \mathbb{R}^n$. Ekkor $P_I = \mathrm{konv}(P \cap \mathbb{Z}^n)$ a P egész pontjainak konvex burka.

Megjegyzés. Ha P leírásában irracionális együtthatók is szerepelhetnek, akkor az egész pontok konvex burka nem lesz mindig poliéder. Például 2-dimenzióban: Legyen $P = \{(x_1, x_2) : x_1 \geq 1, x_2 \geq \sqrt{5}x_1\}$. Ekkor az egész pontok konvex burkának végtelen sok csúcsa lesz. Például: $(1,3)$, $(4,9)$, $(8,18)$, stb.

Az irracionális meredekségű egyeneshez tetszőlegesen közel tudunk egész pontot találni. $(4, 9)$ közelebb van, mint $(1, 3)$, de $(8, 18)$ már közelebb van, mint $(4, 9)$, és így tovább, egyre közelebb kerülünk, és közben mindig új csúcsokat definiálunk, ezáltal végtelen sok csúcsot kapunk. Márpedig egy poliédernek csak véges sok csúcsa lehet.



7.1.1. Tétel (Meyer)**.** *Ha P racionális poliéder (azaz megadható racionális együtthatós lineáris egyenlőtlenségrendszerrel), akkor P_I poliéder.*

Bizonyítás. Ha P korlátos, akkor P_I véges sok pont konvex burka, tehát poliéder. Az tehát az érdekes eset, amikor P nem korlátos. Ekkor a 3.4.9 tétel szerint $P = Q + C$, ahol Q egy racionális politop (korlátos poliéder), és C egy végesen generált racionális kúp. A C kúpot véges sok egész vektor is generálja, legyenek ezek $g_1, \dots, g_k$. Tekintsük a következő ponthalmazt, ami politóp, hiszen k darab szakasz vektor-összege (lásd 3.7 gyakorlat):

$$D = \left\{ \sum_{i=1}^{k} \lambda_i g_i : \ 0 \leq \lambda_i \leq 1 \ (i = 1, \dots, k) \right\}.$$

Belátjuk, hogy $P_I = (Q + D)_I + C$. Ez elég a tétel bizonyításához, hiszen $Q + D$ korlátos, tehát $(Q + D)_I$ poliéder, és így $(Q + D)_I + C$ is.

Egyik irány: $P_I \subseteq (Q + D)_I + C$. Mivel $(Q + D)_I + C$ konvex halmaz, elég belátni, hogy P minden egész p pontja benne van. Legyen tehát $p \in P \cap \mathbb{Z}^n$.

A 3.4.9 tétel szerint p felírható $q + c$ alakban, ahol $q \in Q, c \in C$. Ekkor léteznek olyan nemnegatív μ_i együtthatók, melyekkel:

$$p = q + c = q + \sum_{i=1}^{k} \mu_i g_i = (q + \sum_{i=1}^{k} (\mu_i - \lfloor \mu_i \rfloor) g_i) + \sum_{i=1}^{k} \lfloor \mu_i \rfloor g_i =: (q + d) + c'.$$

Itt $d \in D$ és $c' \in C$ nyilvánvalóan teljesül, továbbá $q + d$ egész, mivel $q + d = p - c'$, azaz előáll két egész vektor különbségeként. Így p tényleg benne van $(Q + D)_I + C$-ben.

Másik irány: $P_I \supseteq (Q+D)_I + C$. Nyilván $Q+D \subseteq P$, így

$$(Q+D)_I + C \subseteq P_I + C = P_I + C_I.$$

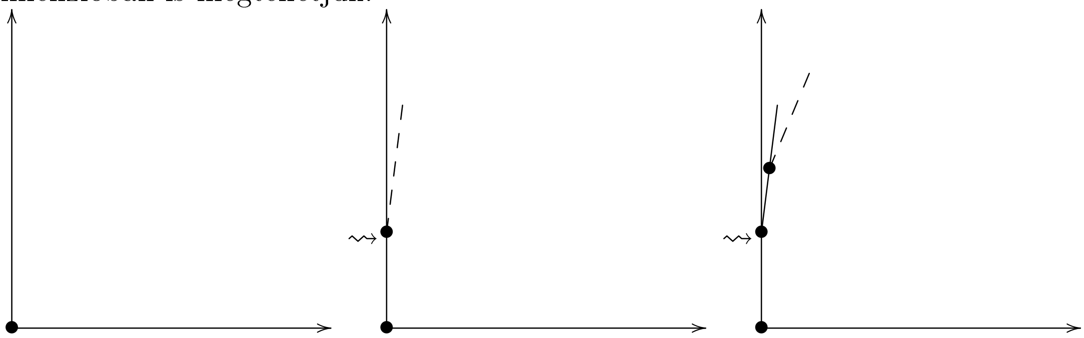
Tetszőleges X és Y halmazokra igaz, hogy $\mathrm{konv}(X) + \mathrm{konv}(Y) \subseteq \mathrm{konv}(X+Y)$. Ezt használva:

$$P_I + C_I \subseteq \mathrm{konv}((P \cap \mathbb{Z}^n) + (C \cap \mathbb{Z}^n)) \subseteq (P+C)_I = P_I.$$

•

Megjegyzés. Ugyan P_I is poliéder, de lehet, hogy sokkal bonyolultabb, mint P, sokkal több csúcsa van. Például induljunk ki a következő kúpból: $\{x_1 \geq 0, x_2 \geq 0\}$. Itt az origó benne van P_I-ben. Válasszunk az $x_2 = 0$ egyenesen tetszőleges egész v_1 pontot, majd „forgassuk el" az $x_1 \geq 0$ félsíkot a v_1 körül úgy, hogy az origó benne legyen, de az újonnan bekerült részben ne legyenek egész pontok, és a meredekség racionális maradjon. Keressünk egy újabb egész pontot az új határoló egyenesen, és folytassuk az eljárást.

Ezzel az eljárással tetszőleges N pozitív egész számra le tudunk gyártani olyan egycsúcsú poliédert, ahol az egész pontok konvex burkának N csúcsa van, és mindezt már 2 dimenzióban is megtehetjük.



A továbbiakban olyan algoritmusokat ismertetünk, amikkel egészértékű programozási feladatokat lehet többé-kevésbé hatékonyan megoldani.

7.2. Vágósíkos eljárás

Célunk a $\max\{cx : Ax \leq b, x \in \mathbb{Z}^n\}$ optimalizálási feladat megoldása. Legyen $P = \{x \in \mathbb{R}^n : Ax \leq b\}$.

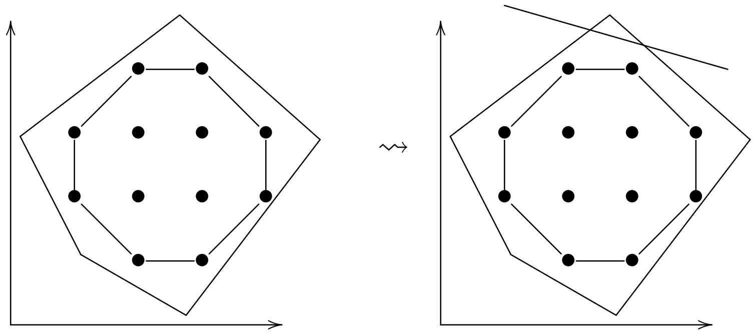
Általános vágósíkos algoritmus

1. Oldjuk meg a $\max\{cx : x \in P\}$ feladatot (például szimplex módszerrel)

2. Ha az x^* optimális megoldás egész, akkor kész vagyunk.

3. Ha x^* nem egész, akkor keressünk olyan $\alpha x \leq \beta$ egyenlőtlenséget, amire $\alpha x^* > \beta$, de $\alpha x \leq \beta$ $\forall x \in P \cap \mathbb{Z}^n$

4. Legyen $P := P \cap \{x : \alpha x \leq \beta\}$, és lépjünk az 1. pontra.



Megjegyzés. A fenti algoritmus ebben a formában nem feltétlenül véges, még két dimenzióban sem. A másik probléma, hogy hogyan találunk az algoritmus 3. pontjában egy megfelelő $\alpha x \leq \beta$ egyenlőtlenséget. Erre ad megoldást (kicsit más alakú feladatnál) Gomory módszere.

7.2.1. Gomory-vágás

Legyen a feladat $\max\{cx : Ax = b, x \geq 0, x \in \mathbb{Z}^n\}$. Tudjuk, hogy $\max\{cx : Ax = b, x \geq 0\}$ megoldható szimplex módszerrel. Legyen B az így kapott optimális bázis. Használjuk a korábbi jelölést: $\bar{x}$, $\bar{A}$, $\bar{b}$.

Tegyük fel, hogy $\bar{x}$ nem egész: $\bar{x}_q \notin \mathbb{Z}$ valamilyen $q \in B$ indexre. Tudjuk, hogy ha x_q az r-edik sorhoz tartozó bázisváltozó, akkor $\bar{x}_q = \bar{b}_r$. Az r-edik sor egyenlete az $\bar{A}x = \bar{b}$ egyenletrendszerben:

$$x_q + \sum_{j \in N} \bar{a}_{rj} x_j = \bar{b}_r \quad \text{– érvényes egyenlet minden megoldásra.}$$

Az egyenletből kaphatunk egy érvényes egyenlőtlenséget, kihasználva, hogy minden megoldás nemnegatív:

$$x_q + \sum_{j \in N} \lfloor \bar{a}_{rj} \rfloor x_j \leq \bar{b}_r \quad \text{– érvényes egyenlőtlenség minden megoldásra.}$$

Egészértékű megoldás esetén a baloldal egész, tehát a jobboldalt is kerekíthetjük:

$$x_q + \sum_{j \in N} \lfloor \bar{a}_{rj} \rfloor x_j \leq \lfloor \bar{b}_r \rfloor \quad \text{– érvényes egyenlőtlenség minden egész megoldásra.}$$

Észrevétel. *Az $\bar{x}$ megoldásra nem teljesül az egyenlőtlenség, mert $\bar{x}_q = \bar{b}_r > \lfloor \bar{b}_r \rfloor$ és $\bar{x}_j = 0 \ \forall j \in N$. Tehát kaptunk egy olyan egyenlőtlenséget, ami minden egész megoldásra érvényes, de az $\bar{x}$ megoldásra nem, így ezt használhatjuk a vágósíkos algoritmus 3. pontjában.*

Megjegyzés. Megfelelő megvalósítás esetén a Gomory vágásokkal véges algoritmust kapunk. Azonban az algoritmus még így is lehet nagyon lassú.

Egy lehetséges gyorsítás, ha olyan vágást próbálunk találni, ami a „lehető legjobb” olyan értelemben, hogy az egész pontok konvex burkának egy lapját határozza meg. A probléma az, hogy ezt a konvex burkot nem ismerjük, így a lapjait sem tudjuk meghatározni. Ez azonban nem zárja ki, hogy néhány lapját a feladat struktúrájából adódóan ismerjük, és azokkal próbálkozzunk.

Tehát ha ismerjük a konvex burok lapjainak egy részét, akkor amint az LP-re optimális megoldást találunk, megnézzük, hogy van-e a lapok között olyan, amit a talált optimum nem teljesít (ez persze nehéznek tűnik, ha exponenciálisan sok lapot ismerünk, de néha még akkor is megoldható!) Ha találunk ilyet, akkor ezt a lapot hozzávesszük a feladathoz, és ismételjük az eljárást.

Bináris hátizsák-feladat

Nézzünk egy bináris hátizsák-feladatot, ahol $b \geq a_1 \geq a_2 \geq \cdots \geq a_n$. Vegyünk egy olyan $\{j_1, \ldots, j_k\}$ indexhalmazt ($j_1 \leq j_2 \leq \cdots \leq j_k$), amire a következő négy feltétel teljesül:

1. $\sum_{i=1}^{k} a_{j_i} > b$,
2. $\sum_{i=1}^{k-1} a_{j_i} \leq b$,
3. Ha $j_1 \neq 1$, akkor $a_1 + \sum_{i=3}^{k} a_{j_i} \leq b$,
4. Ha $j > j_1$ és $j \notin \{j_1, \ldots, j_k\}$, akkor $a_j + \sum_{i=2}^{k} a_{j_i} \leq b$.

7.1. Feladat. *Mutassuk meg, hogy ha nem fér be az összes tárgy a hátizsákba, akkor mindig létezik ilyen indexhalmaz, és a* $\sum_{j<j_1} x_j + \sum_{i=1}^{k} x_{j_i} \leq k-1$ *egyenlőtlenség az egész megoldások konvex burkának lapját határozza meg. (Tipp: próbáljunk* n *darab affin független megoldást mutatni, ami ezt az egyenlőtlenséget egyenlőséggel teljesíti!)*

Utazó ügynök feladat

Adott n város, és bármely kettő között a távolság. Az utazó ügynök szeretné a legrövidebb olyan körsétát megtalálni, ami minden várost pontosan egyszer érint. Megengedett megoldásai tehát az n pontú teljes gráf Hamilton körei, és ezek közül szeretné a legrövidebbet kiválasztani. Jelölje $c(uv)$ az u és v városok távolságát, és legyen V az összes város halmaza.

Első próbálkozásként tekintsük a következő bináris feladatot:

$$\begin{aligned} \min cx& \\ d_x(v) = 2 &\quad \forall v \in V \\ x \in \{0,1\}^{\binom{n}{2}}& \end{aligned}$$

A Hamilton-körök ugyan megoldások, de sajnos több diszjunkt kör uniója is lehet megoldás, ha együtt az összes várost tartalmazzák. Ezért még hozzá kell venni az összefüggőséget biztosító egyenlőtlenségeket:

$$d_x(U) \geq 2 \quad \forall U \subsetneq V,\ U \neq \emptyset.$$

Ennek a megoldásai már kizárólag a Hamilton-körök. A feladat LP relaxáltja:

$$\min cx \tag{7.1}$$
$$d_x(v) = 2 \quad \forall v \in V \tag{7.2}$$
$$0 \leq x \leq 1 \tag{7.3}$$
$$d_x(U) \geq 2 \quad \forall U \subsetneq V,\ U \neq \emptyset. \tag{7.4}$$

Legyen P a (7.2), (7.3), (7.4) egyenlőtlenségeket kielégítő vektorok halmaza. Ekkor P_I a Hamilton körök karakterisztikus vektorainak a konvex burka.

Megjegyzés. A poliéder nem függ a súlyfüggvénytől, minden n-re egy poliéderünk van. A (7.2) egyenletek egy olyan hipersíkot határoznak meg, ami tartalmazza a P poliédert.

Ha a vágósíkos eljárást akarjuk alkalmazni, már az elején egy problémába ütközünk: a (7.4) típusú egyenlőtlenségekből exponenciálisan sok van, ezért a szimplex módszerben exponenciális méretű mátrixokkal kellene dolgozni, ami túl lassú. Ezért már az LP relaxáció megoldásához is egy vágósíkos eljárást használunk, a következőképpen:

Keressük meg az optimális olyan vektort, ami a (7.2), (7.3) feltételeket teljesíti. Legyen ez x^*.

7.2. Feladat. *Mutassuk meg, hogy $n-1$ maximális folyam kereséssel eldönthető, hogy x^* teljesíti-e az összes (7.4) típusú egyenlőtleséget, és ha nem, található egy olyan, amit nem teljesít.*

A feladat alapján el tudjuk dönteni, hogy x^* optimális megoldása-e az LP relaxációnak, és ha nem, találunk egy egyenlőtlenséget, amit nem teljesít. Ezt hozzávéve az eddigiekhez újra megoldjuk az LP-t, és ezt addig folytatjuk, amíg a (7.4) feltételek nem teljesülnek.

Megjegyzés. Ha a végén kapott megoldás nem egész, akkor alkalmazhatunk Gomory vágást, vagy korlátozás és szétválasztást (lásd később).

7.3. Dinamikus programozási algoritmusok

Dinamikus programozást olyan feldatok megoldásánál használhatunk, ahol a megoldás előállítható egyszerűbb részfeladatok megoldása segítségével. A dinamikus programozás lényege, hogy egy bizonyos részfeladatot csak egyszer oldunk meg, és a megoldást tároljuk, így ezt a megoldást a nagyobb részfeladatok megoldásához már további számolás nélkül használhatjuk.

7.3.1. Bináris hátizsákfeladat

Legyen $k \in \{1, \dots, n\}$, $d \in \{0, \dots, b\}$. Definiáljuk a következő függvényt:

$$f_k(d) = \max\{\sum_{j=1}^{k} c_j x_j : \sum_{j=1}^{k} a_j x_j \leq d,\ x \in \{0,1\}^k\}.$$

Látjuk, hogy $f_n(b)$ lesz az eredeti bináris hátizsákfeladat optimumértéke.

Legyen $f_0(d) = 0$ minden $d \in \{0, \dots, b\}$ értékre. Írjunk fel rekurzív képletet $f_k(d)$ értékére, ha $k \geq 1$:

$$f_k(d) = \begin{cases} f_{k-1}(d), \text{ ha } a_k > d \\ \max\{f_{k-1}(d), f_{k-1}(d - a_k) + c_k\}, \text{ ha } a_k \leq d \end{cases}$$

Algoritmus

```
for  k = 1...n
     for  d = 0...b
     do { számoljuk ki f_k(d) -t a fenti képlettel}
```

Ez a dinamikus programozási algoritmus $\mathcal{O}(nb)$ lépésben kiszámolja az optimumértéket. Megjegyzendő, hogy ez nem polinomiális futási idő, hiszen az input mérete $\mathcal{O}(n \log b)$.

Az optimális megoldást a következőképpen kapjuk:

$$x_n := \begin{cases} 0, \text{ ha } f_n(b) = f_{n-1}(b) & (1) \\ 1, \text{ ha } f_n(b) = f_{n-1}(b - a_n) + c_n & (2) \end{cases}$$

$x_{n-1}, \dots, x_1$-et pedig esetszétválasztással kapjuk:

- Ha az (1) eset következik be, akkor

$$x_{n-1} := \begin{cases} 0, \text{ ha } f_{n-1}(b) = f_{n-2}(b) \\ 1, \text{ ha } f_{n-1}(b) = f_{n-2}(b - a_{n-1}) + c_{n-1} \end{cases}$$

- Ha a (2) eset következik be, akkor

$$x_{n-1} := \begin{cases} 0, \text{ ha } f_{n-1}(b - a_n) = f_{n-2}(b - a_n) \\ 1, \text{ ha } f_{n-1}(b - a_n) = f_{n-2}(b - a_n - a_{n-1}) + c_{n-1} \end{cases}$$

Így folytatjuk, amíg x_1-et el nem érjük.

7.3.2. Nemnegatív mátrixú egészértékű feladat

Most a fentihez hasonló dinamikus programozási algoritmust adunk több egyenlőtlenség esetén. Legyen a feladatunk

$$\max\{cx : \ Ax \leq b, x \geq 0, x \in \mathbb{Z}^n\},$$

ahol A, b, c nemnegatív és egész.

Adott $k \in \{1, \dots, n\}$ szám és $0 \leq d \leq b$ egész vektor esetén legyen

$$f_k(d) = \max\{\sum_{j=1}^{k} c_j x_j : \sum_{j=1}^{k} a_{ij} x_j \leq d_i \ (i = 1, \dots, m), \ x \geq 0\}.$$

Definiáljuk ezen kívül $f_0(d)$-t 0-nak minden $0 \le d \le b$-re. A következő rekurzióval lehet $k \ge 1$-re és $0 \le d \le b$-re kiszámolni a függvényértékeket (fontos eltérés a bináris hátizsákfeladathoz képest, hogy a maximumban $f_k(d - a_{.k})$ szerepel):

$$f_k(d) = \begin{cases} f_{k-1}(d) & \text{ha valamilyen } i\text{-re } a_{ik} > d_i, \\ \max\{f_{k-1}(d), f_k(d - a_{.k}) + c_k\} & \text{ha } a_{.k} \le d. \end{cases}$$

Az $f_k(d)$ értékeket k szerint növekvő sorrendben, azon belül d szerint lexikografikusan növekvő sorrendben számoljuk ki. Az optimumértéket $f_n(b)$ adja meg. A lépésszám $O(n \prod_{i=1}^{m}(b_i + 1))$, ami sajnos a legtöbb feladatnál nagyon lassú algoritmust ad, de ha b kicsi, akkor használható.

7.4. Korlátozás és szétválasztás

Tekintsük a következő alakú feladatot:

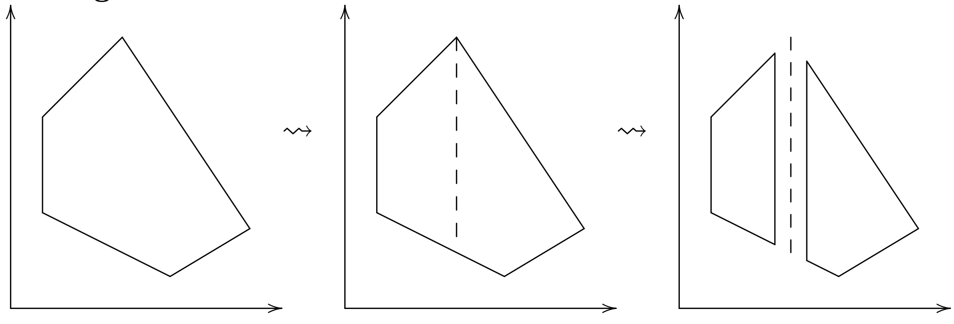
$$\max cx \tag{7.5}$$
$$Ax \le b \tag{7.6}$$
$$x \text{ egész}, \tag{7.7}$$

ahol A, b, c egészek. A *korlátozás és szétválasztás* módszerének alapgondolata az, hogy a feladatot szétválasztjuk részfeladatokra úgy, hogy az eredeti feladat megengedett megoldásainak halmaza a részfeladatok megengedett megoldás-halmazainak diszjunkt uniója legyen. Ekkor az eredeti feladat optimum-értéke megegyezik a részfeladatok optimum-értékeinek maximumával. Az egyes részfeladatokat további részfeladatokra lehet szétbontani, így a vizsgált feladatok egy fát alkotnak, aminek gyökerében az eredeti feladat található.

A részfeladatok relevanciájának vizsgálatához van szükség a *korlátozásra.* A korlátozás azt jelenti, hogy minden részfeladatnál felső (és esetleg alsó) korlátot számolunk az optimum értékére. Amennyiben egy részfeladatra vonatkozó felső korlát kisebb mint egy már ismert alsó korlát, akkor azzal a részfeladattal nem kell tovább foglalkozni, törölhetjük a fából.

LP alapú korlátozás és szétválasztásról akkor beszélünk, ha a felső korlátokat az LP relaxált optimuma adja, a szétválasztás pedig lineáris egyenlőtlenségek hozzáadásával történik.

Az algoritmus működése:



A részfeladatokra osztás során kihagyunk részeket a feladat teréből, de csak olyanokat, amikben nincsenek egész pontok. A továbbiakban a legegyszerűbb LP alapú általános algoritmust írjuk le. Legyen $P = \{x \in \mathbb{R}^n : \ Ax \leq b\}$. A részfeladataink mindig $\max\{cx : \ x \in P' \cap \mathbb{Z}^n\}$ alakúak lesznek, ahol $P' \subseteq P$ poliéder. Jelölések:

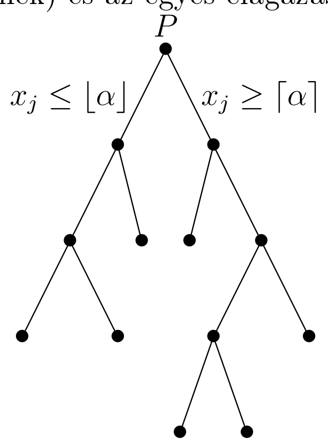
- $u(P') = \max\{cx : \ x \in P'\}$. Ez felső korlát a részfeladat optimumértékére.
- $x^*(P')$ a $\max\{cx : \ x \in P'\}$ feladat egy optimális bázismegoldása. Ez kiszámolható szimplex módszerrel.
- x^*: az eddig talált legjobb egész megoldás
- L: az eddig talált legjobb egész megoldás célfüggvényértéke, illetve $-\infty$ ha még nem találtunk egész megoldást.
- $\mathcal{F}$: A feldolgozandó poliéderek listája.

Az algoritmus inicializálása: $\mathcal{F} = \{P\}$, és $L = -\infty$. Egy általános lépés a következő:

1. Ha $\mathcal{F}$ üres: kész vagyunk. Különben az $\mathcal{F}$ listáról kiválasztunk egy P' feladatot, és kiszámoljuk az $u(P')$ értéket.
2. Ha $u(P')$ nem létezik, mert az LP relaxált nem megoldható: töröljük P'-t $\mathcal{F}$-ből.
3. Ha $u(P') \leq L$: töröljük P'-t $\mathcal{F}$-ből.
4. Ha $u(P') > L$ és $x^*(P')$ egész: $x^* := x^*(P')$, és $L := u(P')$. Töröljük P'-t $\mathcal{F}$-ből.
5. Ha $u(P') > L$ és $x^*(P')$ nem egész: válasszunk egy olyan j indexet, amire $x^*(P')$ j-edik komponense egy nem egész α szám. Legyen $P'_1 = P' \cap \{x \in \mathbb{R}^n : \ x_j \leq \lfloor \alpha \rfloor\}$, és $P'_2 = P' \cap \{x \in \mathbb{R}^n : \ x_j \geq \lceil \alpha \rceil\}$. Töröljük P'-t $\mathcal{F}$-ből, és adjuk hozzá P'_1-et és P'_2-t $\mathcal{F}$-hez.

7.4.1. Állítás. *Ha P korlátos, akkor az algoritmus véges sok lépésben talál egy optimális megoldást.*

Bizonyítás. Mivel P korlátos, létezik olyan N szám, hogy tetszőleges $x \in P$-re és $j \in \{1, \dots, n\}$-re $|x_j| \leq N$. Az algoritmus futása alapján felépíthetünk egy gyökeres bináris fát, amelynek minden csúcsa egy feldolgozott poliédernek felel meg (a gyökér P-nek) és az egyes elágazások a szétválasztásoknak.



Tekintsünk a fában egy tetszőleges utat a gyökérből egy levélbe! Egy adott x_j változó szerint ezen az úton legfeljebb $2N$ szétválasztás lehet, mert minden szétválasztásnál egy 1 hosszú nyílt intervallum kimarad az x_j lehetséges értékei közül. Így minden ilyen út legfeljebb $2nN$ hosszú, amiből következik, hogy az algoritmus véges sok lépésben véget ér.

Az optimális megoldás P-ben van, és ha az algoritmus egy adott szétválasztási lépésénél P'-ben ott van, akkor vagy P_1'-ben vagy P_2'-ben is ott van. Tehát a fenti fa valamelyik leveléhez tartozó poliéderben van az optimális megoldás. Mivel egy levélben nem választunk szét, ezért ott meg is találjuk. Tehát az algoritmus mindenképpen megtalálja az optimális megoldást. •

Korlátozás és szétválasztás algoritmusa a bináris hátizsák feladatra

Ha bináris hátizsákfeladatra futtatjuk a fenti algoritmust, az algoritmusban szereplő poliéderek mindig a következő alakúak lesznek:

$$P(J_0, J_1) = \{0 \le x \le 1 : \sum_{j=1}^{n} a_j x_j \le b,\ x_j = 0 \text{ ha } j \in J_0,\ x_j = 1 \text{ ha } j \in J_1\}.$$

A $\max\{cx : x \in P(J_0, J_1)\}$ feladat nagyon egyszerűen megoldható a mohó algoritmussal:

- Rakjuk a nem $(J_0 \cup J_1)$-beli változókat $\frac{c_j}{a_j}$ szerint csökkenő sorrendbe.
- Az elején $d = b - \sum_{j\in J_1} a_j$ - ennyi még belefér a hátizsákba.
- Ha j_1 az első index a fenti sorrendben, akkor legyen $x_{j_1} = \min\{1, \frac{d}{a_{j_1}}\}$, és legyen $d := d - a_{j_1} x_{j_1}$. Majd vegyük a következőt a sorrend szerint, és ugyanígy folytassuk.

Megjegyzés. Ha egy változó értéke nem 1, hanem $\frac{d}{a_j}$, akkor az utána következő összes többi változó értéke 0 lesz, hiszen a hátizsák megtelt. Tehát legfeljebb egyetlen nem egész érték lesz. Ezért mindig egyértelmű, melyik változó szerint kell szétválasztani.

Az algoritmust érdemes kiegészíteni azzal, hogy ha egy adott lépésben az x_j változó szerint választunk szét, akkor az optimális tört megoldásban x_j-t 0-ra módosítva egy megengedett egész megoldást kapunk. Ha ez jobb, mint az eddig talált legjobb x^* megoldás, akkor lecserélhetjük x^*-ot erre.

7.5. Közelítő algoritmusok

7.5.1. Minimális lefogó csúcshalmaz

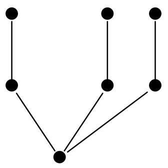
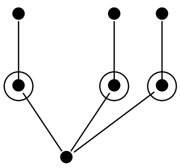
Adott egy $G = (V, E)$ irányítatlan gráf. Csúcsok egy $U \subseteq V$ halmaza **lefogó csúcshalmaz**, ha minden élnek legalább egyik végpontját tartalmazza. A minimális méretű lefogó csúcshalmaz megkeresése NP-nehéz feladat. Most két algoritmust ismertetünk lefogó csúcshalmaz keresésére, amik nagyon gyorsak, de nem feltétlenül szolgáltatnak optimális megoldást.

1. Algoritmus (mohó)

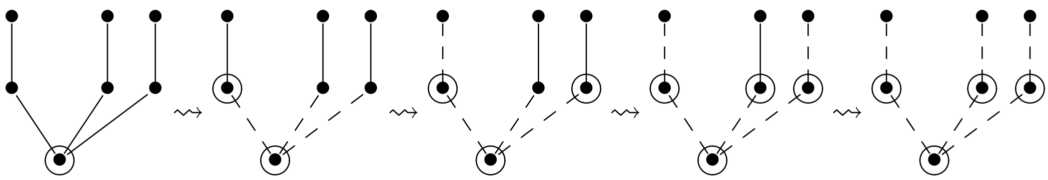
Kezdetben legyen $U = \emptyset$.

1. Válasszunk maximális fokú csúcsot a gráfban. Legyen ez v.
2. Legyen $U := U \cup \{v\}$, és a G gráfból hagyjuk ki a v csúcsot és az összes rá illeszkedő élt.
3. Ha nem marad él, akkor kész vagyunk, U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

Megjegyzés. A fenti algoritmus nem feltétlenül talál optimális megoldást. Például:


optimális megoldása:


Az algoritmusunk egy lehetésges futása viszont a következő:

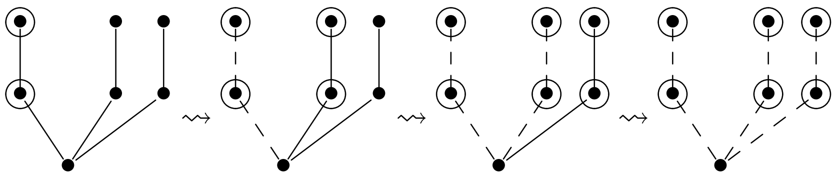


2. Algoritmus (párosító)

Kezdetben legyen $U = \emptyset$.

1. Válasszunk egy tetszőleges uv élt a gráfban.
2. Legyen $U := U \cup \{u\} \cup \{v\}$
3. A G gráfból hagyjuk ki ezt a két csúcsot, és az összes rájuk illeszkedő élt.
4. Ha nem marad él, akkor kész vagyunk. U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

Megjegyzés. A párosító algoritmus nem tűnik túl jó megoldásnak, mivel feleslegesen két csúcsot is bevesz lépésenként a lefogó csúcshalmazba. Az előbbi példára rosszabban teljesít, mint a mohó algoritmus:



Elsőre nem is könnyű olyan példát találni, ahol a párosító algoritmus jobban teljesít, mint a mohó. Azonban megmutatjuk, hogy bizonyos szempontból a párosító számít a jobb algoritmusnak. Ehhez bevezetjük az α-közelítés fogalmát:

Definíció. Egy minimalizálási feladatra egy algoritmus α-**közelítő**, ha tetszőleges inputra az output értéke legfeljebb α-szorosa az optimálisnak.

7.5.1. Állítás. *A párosító algoritmus 2-közelítő.*

Bizonyítás. Legyen F az algoritmus során kiválasztott élek halmaza. F független élekből áll, tehát egy párosítás. Ekkor $|U_{opt}| \geq |F|$, mivel az optimális lefogó csúcshalmaz lefogja F éleit. Másrészt: $|U| = 2|F|$. Tehát $|U_{opt}| \geq \frac{1}{2}|U|$. •

7.5.2. Állítás. *A mohó algoritmus nem 2-közelítő.*

Bizonyítás. Definiálunk egy $G = (A, B; E)$ páros gráfot. Az A osztályban 60 csúcs van, mind 5 fokú. A B osztályban 12 db 5 fokú, 15 db 4 fokú, 20 db 3 fokú, 30 db 2 fokú, és 60 db 1 fokú. Minden $i \in \{1, \ldots, 5\}$-re a B-beli i fokúak szomszédságai A egy partícióját adják (hogy pontosan melyik partícióját, az mindegy).

A mohó algoritmus lefuthat úgy, hogy a teljes B halmazt választja ki lefogó csúcshalmaznak. Másrészt A is lefogó csúcshalmaz, tehát mivel $|B| > 2|A|$, ez nem 2-közelítő algoritmus. (Hasonló konstrukcióval az is belátható, hogy a mohó algoritmus semmilyen α-ra nem lesz α-közelítő.) •

Megjegyzés. Érdekes módon nem ismert $\alpha < 2$-re α-közelítő polinomiális futási idejű algoritmus erre a feladatra, tehát ilyen értelemben a párosító algoritmus a „legjobb ismert". Persze az algoritmus nyilvánvalóan javítható, például ha a végén a felesleges csúcsokat töröljük, de ettől nem kapunk garantáltan 2-nél jobb közelítést.

7.5.1. Tétel (Dinur és Safra, 2005)**.** *Ha $P \neq NP$, akkor semmilyen $\alpha < 1.3607$-re nincs α-közelítő polinomiális algoritmus a minimális lefogó csúcshalmaz feladatra. (Bizonyítás nélkül)*

7.5.2. Minimális költségű lefogó csúcshalmaz

Legyen $G = (V, E)$ irányítatlan gráf, és $c \in \mathbb{R}_+^V$ költségfüggvény. A minimális költségű lefogó csúcshalmaz feladatban olyan U lefogó csúcshalmazt keresünk, amire $\sum_{u \in U} c_u$ minimális.

A következőkben erre a feladatra adunk egy 2-közelítő algoritmust, ami a párosító algoritmus általánosításnak tekinthető. Írjuk fel a feladatot egészértékű programozási feladatként:

$$\begin{aligned} &\min cx \\ x_u + x_v &\geq 1 \qquad \forall uv \in E \\ x &\in \{0,1\}^V \end{aligned}$$

A feladat LP-relaxáltja:

$$\begin{aligned} &\min cx \\ x_u + x_v &\geq 1 \qquad \forall uv \in E \\ x &\in \mathbb{R}_+^V \end{aligned}$$

Megjegyzés. Azért nem vettük hozzá az $x \leq 1$ feltételt, mert felesleges: az 1-nél nagyobb értékek lecsökkenthetők 1-re a feltételek megsértése nélkül, és a költség csak csökkenhet.

Nézzük az LP-relaxált duálisát:

$$\max \sum_{e\in E} y_e$$
$$\sum_{u:uv\in E} y_{uv} \le c_v \qquad \forall v \in V$$
$$y \in \mathbb{R}^E_+$$

Primál-duál algoritmus

Minden lépésben lesz egy $U \subseteq V$ csúcshalmazunk, egy y duális megoldás. Kezdetben legyen $U = \emptyset$, $y \equiv 0$, $E' = E$.

1. Válasszunk tetszőleges $uv \in E'$ élt.
2. Emeljük meg y_{uv}-t annyira, hogy u-nál vagy v-nél a duál egyenlőtlenség teljesüljön egyenlőséggel – előfordulhat, hogy az egyenlőség mindkettőre teljesül.
3. U-hoz vegyük hozzá u és v közül azt, amelyiknél egyenlőség van – ha mindkettőnél egyenlőség van, akkor mindkettőt. Töröljük E'-ből a lefogott éleket.
4. Ha $E' = \emptyset$, akkor kész vagyunk, U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

Megjegyzés. Az y_{uv} emelése csak az u-hoz és v-hez tartozó duál egyenlőtlenségek teljesülését befolyásolja, a többit nem. Ezért y végig duál megengedett megoldás marad. Könnyen látható, hogy ha minden csúcs költsége 1, akkor pont a párosító algoritmust kapjuk.

7.5.3. Állítás. *A fenti primál-duál algoritmus 2-közelítő.*

Bizonyítás. Vezessünk be $\pi \in \mathbb{R}^V_+$ segédváltozókat. Kezdetben legyen $\pi \equiv 0$. Az algoritmus során, amikor y_{uv}-t megemeljük δ-val, növeljük meg $\pi(u)$-t és $\pi(v)$-t is δ-val.

Az algoritmus befejezésekor $\sum_{v\in V} \pi(v) = 2\sum_{e\in E} y_e$. Másrészt, amikor v bekerül U-ba, akkor $\pi(v) = c(v)$, és ez később sem változik. Tehát $\sum_{u\in U} c_u \le \sum_{v\in V} \pi(v)$, amiből

$$\sum_{u\in U} c_u \le 2\sum_{e\in E} y_e \underbrace{\le}_{\text{gyenge dualitás}} 2OPT_{LP} \le 2OPT_{IP}$$

•

Megjegyzés. Valójában többet bizonyítottunk, mint amit a tétel állít: a kapott megoldás költsége legfeljebb kétszerese az LP-relaxált optimum-értékének.

8. fejezet

Konvex optimalizálás

8.1. Konvex halmazok

8.1.1. Alaptulajdonságok

A **konvex halmaz**, **konvex burok** és **konvex kúp** fogalmát már definiáltuk a 3.2 fejezetben. Egy $X \subseteq \mathbb{R}^n$ halmaz konvex burkát $\text{konv}(X)$ jelöli.

8.1.1. Állítás. *Konvex halmazok metszete konvex, akár végtelen soké is.*

Bizonyítás. Legyen $C = \cap C_i$, $x, y \in C$. Ekkor $x, y \in C_i$ minden i-re. Tehát $0 \leq \lambda \leq 1$-re $\lambda x + (1-\lambda)y \in C_i\ \forall i \Rightarrow \in C$. •

Az affin burok definíciója is szerepelt már a 2.1 fejezetben, itt most megismételjük és új jelöléseket is bevezetünk.

Definíció. Az $x_1, \ldots, x_k \in \mathbb{R}^n$ vektorok egy **affin kombinációja** egy olyan vektor, ami előáll $\sum_{i=1}^k \lambda_i x_i$ alakban, ahol $\sum_{i=1}^k \lambda_i = 1$. Egy $X \subseteq \mathbb{R}^n$ halmaz **affin burka** az összes olyan vektorból áll, ami X véges sok elemének affin kombinációja. Ez egy affin altér, jelölése: $\text{aff}(X)$. Az affin burok lineáris eltoltját $\text{lin}(X)$ jelöli. Az X halmaz **dimenziója**: $\dim(X) := \dim(\text{lin}(X))$ (az altér dimenzióját már definiáltuk a 2.1 fejezetben).

8.1.2. Állítás. *Konvex C-re* $\text{aff}(C) = \{\mu x + (1-\mu)y : x, y \in C,\ \mu \in \mathbb{R}\}$.

Bizonyítás. Legyen $x_1, \ldots, x_k \in C$, $z = \sum_{i=1}^k \mu_i x_i$, ahol $\sum \mu_i = 1$. Be kell látnunk, hogy z előáll $\mu x + (1-\mu)y$ alakban, ahol $x, y \in C$. Ha $\mu_i \geq 0$ $(i = 1, \ldots, k)$, akkor $z \in C$ és kész vagyunk.

Legyen $\mu = \sum_{i:\mu_i > 0} \mu_i$, $x = \sum_{i:\mu_i > 0} \frac{\mu_i}{\mu} x_i$. Ekkor x konvex kombináció, tehát $x \in C$. Legyen $y = \sum_{i:\mu_i < 0} \frac{\mu_i}{1-\mu} x_i \in C$. Ekkor $z = \mu x + (1-\mu)y$. •

Minden poliéder konvex halmaz, hiszen konvex halmazok (félterek) metszete. A következőkben azt vizsgáljuk, hogy poliéderek oldalainak mi a megfelelője általános konvex halmazok esetében.

Definíció. Egy $E \subseteq C$ halmaz **extremális halmaza** a C konvex halmaznak, ha E konvex, és teljesül, hogy ha $x, y \in C$-re és $0 < \lambda < 1$-re $\lambda x + (1-\lambda)y \in E$, akkor $x \in E$ és $y \in E$. Azaz: ha egy C-beli szakasz egy belső pontja benne van E-ben, akkor az egész szakasz benne van.

8.1.3. Állítás. *Ha* $\max\{wx : x \in C\}$ *megoldható, akkor az optimális megoldások halmaza extremális halmaz.*

Bizonyítás. Legyen S az optimális megoldások halmaza, α pedig az optimumérték. Először belátjuk, hogy S konvex: ha $x, y \in S$, akkor $wx = wy = \alpha$, tehát $w(\lambda x + (1 - \lambda)y) = \alpha$.

Legyen most $z \in S$, $z = \lambda x + (1-\lambda)y$, ahol $0 < \lambda < 1$, és $x, y \in C$. Be kell látnunk, hogy $x, y \in S$.

$$\alpha = wz = \lambda \underbrace{wx}_{\leq\alpha} + (1-\lambda)\underbrace{wy}_{\leq\alpha} \leq \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

Egyenlőség csak akkor lehet, ha $wx = \alpha$ és $wy = \alpha$, azaz $x, y \in S$. •

Példa. Henger extremális halmazai:

0 dimenziós extremális halmaz elemei (extremális pontok): a két körlap határpontjai;

1 dimenziós extremális halmazok: a henger alkotói;

2 dimenziós extremális halmazok: a két körlap;

3 dimenziós extremális halmazok: az egész henger.

Tehát még kompakt konvex halmazokra sem teljesül az a tulajdonság, ami poliéderek oldalaira igaz, hogy i dimenziós extrém halmaz mindig része egy $i + 1$ dimenziósnak. Azonban az oldalak néhány fontos tulajdonsága általánosítható extremális halmazokra.

8.1.4. Állítás. *Ha* E *extremális halmaza* C*-nek, akkor* $E = \operatorname{aff}(E) \cap C$.

Bizonyítás. Legyen $z \in \operatorname{aff}(E) \cap C$. Ekkor $z = \mu x + (1-\mu)y$, ahol $x, y \in E$. Ha $0 \leq \mu \leq 1$, akkor az E konvexitása miatt $z \in E$ és kész vagyunk. Ha $\mu > 1$, akkor $x = \frac{1}{\mu}\underbrace{z}_{\in C} + \frac{\mu-1}{\mu}\underbrace{y}_{\in E}$. Mivel $y \in E$ és $E \subseteq C$, ezért $y \in C$, továbbá mivel $x \in E$, és E extremális halmaz, ezért $z \in E$. •

Megjegyzés. Ha E_1, E_2 extremális C-ben és $E_1 \subsetneqq E_2$, akkor E_1 extremális halmaza E_2-nek.

8.1.5. Állítás. *Ha* E_1, E_2 *extremális halmazai* C*-nek és* $E_1 \subsetneqq E_2$*, akkor* $\dim(E_1) < \dim(E_2)$.

Bizonyítás. Mivel E_1 extremális halmaza E_2-nek, ezért $E_1 = \operatorname{aff}(E_1) \cap E_2$, tehát $\operatorname{aff}(E_1) \subsetneqq \operatorname{aff}(E_2)$, emiatt $\dim(E_1) < \dim(E_2)$. •

8.1.6. Állítás. *Ha* E_1 *extremális halmaza* C*-nek, és* E_2 *extremális halmaza* E_1*-nek, akkor* E_2 *extremális halmaza* C*-nek.*

Bizonyítás. Legyen $x, y \in C$, $0 < \lambda < 1$, $z = \lambda x + (1-\lambda)y \in E_2$. Ekkor $z \in E_1$, mert $E_2 \subseteq E_1$. E_1 extremális C-ben, emiatt $x, y \in E_1$. Továbbá $z \in E_2$, E_2 extremális E_1-ben, emiatt $x, y \in E_2$. •

8.1.1. Tétel (Minkowski, bizonyítás nélkül). *Ha C korlátos és zárt konvex halmaz, akkor C az extrém pontjainak konvex burka.*

Definíció. A **konvex kúp** fogalmát már definiáltuk a 3.2 fejezetben: olyan részhalmaza $\mathbb{R}^n$-nek, ami zárt a nemnegatív számmal való szorzásra és az összeadásra. Egy konvex kúp **csúcsos**, ha a 0 extremális pontja.

8.1.7. Állítás. *C konvex kúp csúcsos $\Leftrightarrow$ nem tartalmaz egydimenziós lineáris alteret.*

Bizonyítás. $\Rightarrow$: Ha tartalmaz $\{\mu x : \mu \in \mathbb{R}\}$ egyenest, akkor $x \in C$, $-x \in C$, tehát 0 nem extremális pontja.

$\Leftarrow$: Ha 0 nem extremális, akkor létezik $x, y \in C$ és $0 < \lambda < 1$, amire $\lambda x + (1-\lambda)y = 0$, tehát az x-et és y-t összekötő egyenes egydimenziós lineáris altér C-ben. •

Definíció (Megengedett irányok kúpja). Legyen C konvex halmaz, $z \in C$. Ekkor a megengedett irányok kúpja z-ben

$$K(C,z) = \{x \in \mathbb{R}^n : \exists \varepsilon > 0, z + \delta x \in C\ \forall 0 \le \delta \le \varepsilon\}.$$

8.1.8. Állítás. *$K(C,z)$ konvex kúp.*

Bizonyítás. Legyen $x, y \in K(C,z)$. Az x-hez tartozó ε-t jelöljük ε_x-el, az y-hoz tartozót pedig ε_y-al. Legyen $\varepsilon = \min\{\varepsilon_x, \varepsilon_y\}$. Legyen továbbá $\delta \le \varepsilon$. Ekkor

$$z + \delta(\lambda x + (1-\lambda)y) = \lambda(\underbrace{z + \delta x}_{\in C}) + (1-\lambda)(\underbrace{z + \delta y}_{\in C}) \Rightarrow z + \delta(\lambda x + (1-\lambda)y) \in C.$$

•

8.1.9. Állítás. *$K(C,z)$ csúcsos $\Leftrightarrow$ z extremális pontja C-nek.*

Bizonyítás. Ha z nem extremális pont, akkor belső pontja egy $[x,y]$ szakasznak, és $(x-z) \in K(C,z)$, $(y-z) \in K(C,z)$, tehát tartalmaz 0-n átmenő egyenest. Ha viszont $K(C,z)$ tartalmaz 0-n átmenő, $\{\mu s : \mu \in \mathbb{R}\}$ egyenest, akkor $\exists \delta_1 : z + \delta_1 s \in C$, $\exists \delta_2 : z - \delta_2 s \in C$, tehát z nem extremális. •

8.1.2. Konvex halmazok szeparációja

8.1.2. Tétel. *Ha C zárt konvex halmaz, és $z \notin C$, akkor van egy egyértelmű z-hez legközelebbi pont C-ben.*

Bizonyítás. Legyen $\mu = \inf\{\|x - z\| : x \in C\}$. Ez véges, és létezik $x^k \in C$ sorozat, amire $\|x^k - z\| \to \mu$. Ennek van konvergens részsorozata, ami egy $x \in C$ ponthoz tart, tehát $\|x - z\| = \mu$.

Ha $y \in C$-re $\|y - z\| = \mu$,akkor $\mu^2 \le \|\frac{x+y}{2} - z\|^2 = \mu^2 - \frac{1}{4}\|y - x\|^2$, tehát $y = x$. •

Definíció (Vetítés konvex zárt halmazra). A $z \notin C$ ponthoz legközelebbi pontot C-ben a z pont C-re való vetítésének nevezzük, jelölése $\pi_C(z)$.

8.1.3. Tétel (Konvex szeparációs tétel)**.** *Legyen C zárt konvex halmaz, és $z \notin C$. Ekkor létezik $s \in \mathbb{R}^n$ és $\epsilon > 0$: $s^T x \leq s^T z - \epsilon$ tetszőleges $x \in C$-re.*

Bizonyítás. Legyen $s = z - \pi_C(z)$, és legyen $\epsilon = \|s\|^2$. Figyeljük meg, hogy ha $x \in C$, akkor $(x - \pi_C(z))^T(z - \pi_C(z)) \leq 0$, hiszen különben az $[x, \pi_C(z)]$ szakaszon lenne olyan pont, ami közelebb van z-hez, mint $\pi_C(z)$. Így

$$s^T x \leq s^T \pi_C(z) = s^T z + s^T(\pi_C(z) - z) = s^T z - \epsilon.$$

•

8.1. Feladat. *Bizonyítsuk be a konvex szeparációs tétel segítségével a Farkas Lemmát!*

8.1.4. Tétel. *Legyen C konvex halmaz, és $z \notin C$. Ekkor létezik $s \in \mathbb{R}^n$ nemnulla vektor, amire $s^T x \leq s^T z$ tetszőleges $x \in C$-re.*

Bizonyítás. Jelölje $\mathrm{cl}(C)$ a C halmaz lezártját. Mivel $z \notin C$, létezik $z^k \to z$ sorozat, amire $z^k \notin \mathrm{cl}(C)$ semmilyen k-ra. Az előző tétel alapján vannak s^k nemnulla vektorok, amikre $(s^k)^T x \leq (s^k)^T z^k$ tetszőleges $x \in C$-re. Feltehetjük, hogy $\|s^k\| = 1$ minden k ra. Legyen s az s^k sorozat tetszőleges torlódási pontja; errre $s^T x \leq s^T z$ tetszőleges $x \in C$-re. •

8.1.5. Tétel. *Legyenek C_1 és C_2 diszjunkt konvex halmazok. Létezik $s \in \mathbb{R}^n$ nemnulla vektor, amire $s^T x \leq s^T y$ tetszőleges $x \in C_1$-re és $y \in C_2$-re.*

Bizonyítás. Legyen $C = C_1 - C_2$; ez konvex halmaz, és $0 \notin C$, tehát az előző tétel értelmében létezik $s \in \mathbb{R}^n$ nemnulla vektor: $s^T x \leq 0$ tetszőleges $x \in C$-re. Így $s^T x \leq s^T y$ tetszőleges $x \in C_1$-re és $y \in C_2$-re. •

8.2. Konvex függvények

Definíció (Konvex függvény)**.** Legyen $C \subseteq \mathbb{R}^n$ konvex halmaz. Az $f : C \to \mathbb{R}$ függvény konvex, ha tetszőleges $x, y \in C$ és $0 \leq \lambda \leq 1$ esetén $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$.

Szemléletesen ezt úgy is mondhatjuk, hogy teszőleges $x, y \in C$ esetén az $[x, y]$ szakaszon a függvény grafikonja a két végpontot összekötő szakasz alatt megy.

8.2.1. Tétel (Jensen egyenlőtlenség)**.** *Legyen C konvex halmaz, $f : C \to \mathbb{R}$ konvex függvény, továbbá $x_1, \ldots, x_k \in C$, $\lambda_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^k \lambda_i = 1$. Ekkor $f(\sum_{i=1}^k \lambda_i x_i) \leq \sum_{i=1}^k \lambda_i f(x_i)$.*

Bizonyítás. Indukció k-ra; $k = 2$-re igaz az állítás a konvexitás definíciója szerint. Tegyük fel, hogy $k > 2$. Legyen $\lambda = \sum_{i=1}^{k-1} \lambda_i$, $x = \sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i$. Mivel x C-beli pontok konvex kombinációja, ezért $x \in C$. Ekkor

$$f(\sum_{i=1}^k \lambda_i x_i) = f(\lambda x + \lambda_k x_k) = f(\lambda x + (1-\lambda)x_k) \underbrace{\leq}_{f \text{ konvex}} \lambda f(x) + (1-\lambda) f(x_k)$$

$$= \lambda f(\sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i) + \lambda_k f(x_k) \underbrace{\leq}_{\text{indukció}} \sum_{i=1}^{k-1} \lambda_i f(x_i) + \lambda_k f(x_k) = \sum_{i=1}^{k} \lambda_i f(x_i).$$

•

Definíció (Szinthalmaz). $D_\alpha(f) = \{x \in C : f(x) \le \alpha\}$ az f függvény α-hoz tartozó szinthalmaza.

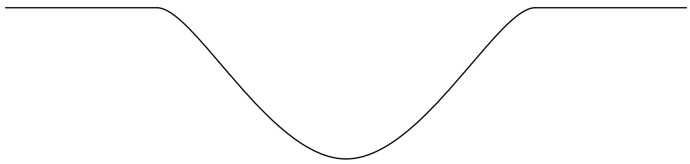
8.2.2. Lemma. *Ha f konvex, akkor D_α konvex tetszőleges α-ra.*

Bizonyítás. Legyen $x, y \in D_\alpha$, $0 \le \alpha \le 1$. Ekkor

$$f(\lambda x + (1-\lambda)y) \underbrace{\le}_{f\ konvex} \lambda f(\underbrace{x}_{\in D_\alpha}) + (1-\alpha) f(\underbrace{y}_{\in D_\alpha}) \le \lambda\alpha + (1-\lambda)\alpha = \alpha$$

. •

Megjegyzés. A lemma fordítottja nem igaz: egy függvény minden szinthalmaza konvex $\not\Rightarrow$ a függvény konvex. Az olyan függvényt, aminek minden szinthalmaza konvex, kvázi-konvexnek nevezzük. Példa:



8.2.3. Lemma. *Legyen C konvex halmaz, és $f : C \to \mathbb{R}$. Az f függvény konvex $\Leftrightarrow$ tetszőleges $x, y \in C$-re a $\phi(\lambda) := f(x + \lambda(y - x))$ függvény konvex a $[0, 1]$ intervallumon.*

Bizonyítás. $\Leftarrow$: Legyen $x, y \in C$, $0 \le \lambda \le 1$. Ekkor $f(\lambda x + (1-\lambda)y) = \phi(1-\lambda) = \phi(0 \cdot \lambda + 1 \cdot (1-\lambda)) \underbrace{\le}_{\phi\ konvex} \lambda \underbrace{\phi(0)}_{=f(x)} + (1-\lambda) \underbrace{\phi(1)}_{=f(y)} = \lambda f(x) + (1-\lambda) f(y)$.

$\Rightarrow$: Legyen $0 \le \nu \le 1$, $\lambda_1, \lambda_2 \in [0, 1]$. Ekkor

$$\begin{aligned}
\phi(\nu\lambda_1 + (1-\nu)\lambda_2) &= f(x + (\nu\lambda_1 + (1-\nu)\lambda_2)(y - x)) \\
&= f(\nu(x + \lambda_1(y - x)) + (1-\nu)(x + \lambda_2(y - x))) \\
&\underbrace{\le}_{f \text{ konvex}} \nu f(x + \lambda_1(y - x)) + (1-\nu) f(x + \lambda_2(y - x)) \\
&= \nu\phi(\lambda_1) + (1-\nu)\phi(\lambda_2).
\end{aligned}$$

•

Analízis tanulmányainkból tudjuk, hogy egy nyílt I intervallumon értelmezett, folytonosan differenciálható $\phi : I \to \mathbb{R}$ függvény pontosan akkor konvex, ha $\phi'(x)$ monoton növő. Ennek az állításnak egy többdimenziós változatát bizonyítjuk alább.

Definíció. Ha f folytonosan differenciálható a C konvex nyílt halmazon, akkor f **gradiense** x-ben: $\nabla f(x) = (\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x))$.

A gradiens segítségével kiszámolhatjuk az iránymenti deriváltat tetszőleges s irányban:

$$\lim_{\lambda \to +0} \frac{f(x + \lambda s) - f(x)}{\lambda} = \nabla f(x)^T s.$$

8.2.4. Tétel. *Legyen C konvex, nyílt halmaz, $f : C \to \mathbb{R}$ folytonosan differenciálható függvény. Ekkor f konvex $\Leftrightarrow$ tetszőleges $x, y \in C$-re $\nabla f(x)^T(y - x) \le f(y) - f(x) \le \nabla f(y)^T(y - x)$.*

Bizonyítás. $\Rightarrow$: $f(x+(1-\lambda)(y-x)) = f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$. Ekkor

$$\frac{f(x+(1-\lambda)(y-x))-f(x)}{1-\lambda} \leq f(y)-f(x).$$

Ha $\lambda \to 1$, akkor a baloldal tart az iránymenti deriválthoz az x pontban, azaz $\nabla f(x)^T(y-x)$-hez. A másik egyenlőtlenséget ugyanígy bizonyíthatjuk x-et és y-t felcserélve.

$\Leftarrow$: Tetszőleges $x, y \in C$-re $\nabla f(x)^T(y-x) \leq \nabla f(y)^T(y-x)$. Ekkor tetszőleges $0 \leq \lambda_1 < \lambda_2 \leq 1$ számokra: $\nabla f(\lambda_2 x+(1-\lambda_2)y)^T(x-y) \leq \nabla f(\lambda_1 x+(1-\lambda_1)y)^T(x-y)$ Ezek szerint $\phi'(\lambda)$ monoton növő, tehát ϕ konvex, és így f konvex. $\bullet$

Definíció (Hesse mátrix)**.** Legyen C konvex és nyílt halmaz, f kétszer folytonosan differenciálható függvény a C halmazon. f Hesse mátrixa: $(\nabla^2 f(x))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$.

8.2.5. Tétel. *Ha f kétszer folytonosan differenciálható, akkor konvex $\Leftrightarrow$ $\nabla^2 f(x)$ pozitív szemidefinit minden $x \in C$-re.*

Bizonyítás. $\Rightarrow$: $x \in C$, $y \in C$, $s = y-x$. Ekkor $s^T\nabla^2 f(x)s = \phi''(0)$. A konvexitásból következik, hogy $\phi''(0) \geq 0$, hiszen ϕ' monoton növő. A C halmaz nyílt, így tartalmaz x körüli gömböt, ezért $s^T\nabla^2 f(x)s \geq 0$ $\forall s \in \mathbb{R}^n$, tehát $\nabla^2 f(x)$ pozitiv szemidefinit.

$\Leftarrow$: $0 \leq s\nabla^2 f(x+\lambda s)s = \phi''(\lambda)$. Így $\phi'' \geq 0$ $\forall \lambda \in [0,1] \Rightarrow \phi'$ monoton növő $\Rightarrow$ f konvex. $\bullet$

8.3. Feltétel nélküli optimalizálás

Definíció. Az $f : C \to \mathbb{R}$ függvénynek $\bar{x} \in C$ **lokális minimuma**, ha $\exists\ \varepsilon > 0$ amire $f(x) \geq f(\bar{x})$ $\forall x \in B(\bar{x}, \varepsilon) \cap C$, ahol $B(\bar{x}, \varepsilon)$ az $\bar{x}$ körüli ε sugarú gömb.

$\bar{x} \in C$ **globális minimum**, ha $f(x) \geq f(\bar{x})$ $\forall x \in C$.

8.3.1. Tétel. *Ha f konvex és $\bar{x}$ lokális minimuma f-nek, akkor globális minimuma is.*

Bizonyítás. $f(x) < f(\bar{x}) \Rightarrow f(\lambda x+(1-\lambda)\bar{x}) \leq \lambda f(x)+(1-\lambda)f(\bar{x}) < f(\bar{x})$ tetszőleges $0 < \lambda \leq 1$ esetén. Tehát $\bar{x}$ tetszőleges környezetében van kisebb értékű pont. $\bullet$

8.3.2. Tétel. *Legyen C nyílt, konvex halmaz, és f konvex, folytonosan differenciálható függvény C-n. Ekkor $\bar{x}$ globális minimum $\Leftrightarrow$ $\nabla f(\bar{x}) = 0$.*

Bizonyítás. $\Leftarrow$: legyen $x \in C$ tetszőleges. Ekkor

$$0 = \nabla f(\bar{x})^T(x-\bar{x}) \underbrace{\leq}_{\text{8.2.4. tétel}} f(x)-f(\bar{x}) \Rightarrow f(x) \geq f(\bar{x}).$$

$\Rightarrow$: tetszőleges $s \in \mathbb{R}^n$-re: $f(\bar{x}+\lambda s) \geq f(\bar{x})$ ha λ elég kicsi ($\bar{x}+\lambda s \in C$). Ekkor $\frac{f(\bar{x}+\lambda s)-f(\bar{x})}{\lambda} \geq 0$ tart az iránymenti deriválthoz: $\nabla f(\bar{x})^T s$-hez ha $\lambda \to 0$. Tehát $\nabla f(\bar{x})^T s \geq 0$ $\forall s \in \mathbb{R}^n \Rightarrow \nabla f(\bar{x}) = 0$. $\bullet$

Definíció. A C konvex halmaz **relatív belseje**

$$\text{rint}(C) = \{x \in C : \text{ tetszőleges } y \in C\text{-re } \exists x' \in C \text{ és } 0 < \lambda < 1\text{: } x = \lambda x' + (1-\lambda)y\}.$$

A C halmaz **relatív nyílt**, ha $C = \text{rint}(C)$.

8.3.3. Tétel. *Legyen C relatív nyílt, konvex halmaz, és f konvex, folytonosan differenciálható függvény C-n. Ekkor $\bar{x}$ globális minimum $\Leftrightarrow \nabla f(\bar{x})^T s = 0\ \forall s \in \text{lin}(C)$.*

Bizonyítás. $\Leftarrow$: legyen $x \in C$ tetszőleges. Ekkor

$$0 = \nabla f(\bar{x})^T \underbrace{(x - \bar{x})}_{\in \text{lin}(C)} \underbrace{\leq}_{\text{8.2.4. tétel}} f(x) - f(\bar{x}) \Rightarrow f(x) \geq f(\bar{x}).$$

$\Rightarrow$: tetszőleges $s \in \text{lin}(C)$-re: ha λ elég kicsi, akkor $\bar{x} + \lambda s \in C$, tehát $f(\bar{x} + \lambda s) \geq f(\bar{x})$. Ekkor $\frac{f(\bar{x}+\lambda s) - f(\bar{x})}{\lambda} \geq 0$ tart az iránymenti deriválthoz, $\nabla f(\bar{x})s$-hez ha $\lambda \to 0$. Tehát $\nabla f(\bar{x})s \geq 0\ \forall s \in \text{lin}(C)$. •

8.4. Feltételes optimalizálás

A konvex optimalizálási (KO) feladat a következő:

$$\begin{aligned} \min f(x)& \\ x \in C& \\ g_i(x) \leq 0& \qquad (i = 1, \ldots, m), \end{aligned}$$

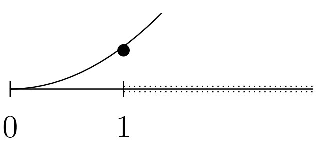
ahol $C \subseteq \mathbb{R}^n$ konvex, és f, g_i $(i = 1, \ldots, m)$ konvexek C-n és folytonosan differenciálhatóak C egy nyílt környezetében.

Definíció. A megengedett megoldások halmaza $C_g = \{x \in C : g_i(x) \leq 0\ (i = 1, \ldots, m)\}$.

A megengedett megoldások halmaza konvex, mivel konvex függvény szinthalmazai konvexek, és konvex halmazok metszete is konvex.

Megjegyzés. Általában C_g nem relatív nyílt (ha az, akkor az optimalitás szükséges és elégséges feltétele, hogy a gradiens merőleges legyen $\text{lin}(C_g)$-re). **Például** $n = m = 1$,

$$\begin{aligned} &f(x) = x^2 \\ &g_1(x) = 1 - x \\ &C = \mathbb{R} \end{aligned}$$



Definíció. Az x pontban a **megengedett irányok kúpja**

$$K(x) := \{s \in \mathbb{R}^n : \exists \varepsilon > 0 : \text{ tetszőleges } 0 \leq \lambda \leq \varepsilon\text{-ra } x + \lambda s \in C_g\}.$$

$K(x)$ tulajdonságai:

i) $K(x)$ konvex kúp;

ii) $K(x)$ csúcsos $\Leftrightarrow$ x extremális pontja C_g-nek;

iii) $K(x) \subseteq \mathrm{lin}(C_g)$;

iv) $K(x) = \mathrm{lin}(C_g) \Leftrightarrow x \in \mathrm{rint}(C_g)$

8.4.1. Tétel. *$\bar{x}$ optimális megoldása a (KO) feladatnak* $\Leftrightarrow \nabla f(\bar{x})^T s \geq 0 \ \forall s \in K(\bar{x})$

Bizonyítás. $\Leftarrow$: Legyen $\bar{x} \neq x \in C_g$, $s = x - \bar{x}$. Ekkor $s \in K(\bar{x})$. Tudjuk, hogy

$$f(x) - f(\bar{x}) \underbrace{\geq}_{f\ konvex} \nabla f(\bar{x})^T s \underbrace{\geq 0}_{s \in K(\bar{x})} \Rightarrow f(x) \geq f(\bar{x}).$$

$\Rightarrow$: Legyen $s \in K(\bar{x})$. Ekkor definició szerint $\exists \varepsilon > 0$: $x_\lambda := \bar{x} + \lambda s \in C_g$ $\forall 0 < \lambda \leq \varepsilon$. $f(x_\lambda) \geq f(\bar{x}) \Rightarrow \frac{f(x_\lambda) - f(\bar{x})}{\lambda} \geq 0$. Ha $\lambda \to 0$, akkor $0 \leq \frac{f(\bar{x} + \lambda s) - f(\bar{x})}{\lambda} \to \nabla f(\bar{x})^T s$. $\bullet$

8.4.1. A Karush-Kuhn-Tucker tétel

Feltételes optimalizálás esetén is hasznos lenne a globális optimumot a gradiensekkel jellemezni, azaz a $\nabla f(x)$, $\nabla g_i(x)$ $(i = 1, \ldots, m)$ vektorok ismeretében eldönteni, hogy x globális optimum-e. Sajnos van arra példa, hogy két feladatban adott x pontra a $\nabla f(x)$, $\nabla g_i(x)$ $(i = 1, \ldots, m)$ értékek megegyeznek, de az egyikben x globális optimum, a másikban pedig nem.

$$\begin{array}{ll} & C = \mathbb{R}^2 \\ & f(x_1, x_2) = x_1 \\ & g_1(x_1, x_2) = x_2 \\ (a) & g_2(x_1, x_2) = -x_2 \\ (b) & g_2(x_1, x_2) = x_1^2 - x_2 \end{array}$$

(a) (b)

A fenti példában az (a) feladatnál és a (b) feladatnál a gradiensek 0-ban ugyanazok, de míg (a)-nál $-\infty$ az infimum, addig (b)-nél $C_g = \{0\}$, tehát az optimum 0.

Definíció. A (KO) feladatnak x **Slater pontja**, ha

i) $x \in \mathrm{rint}(C)$,

ii) $g_i(x) \leq 0$ $(i = 1, \ldots, m)$,

iii) $g_i(x) < 0$ ha g_i nemlineáris.

A (KO) feladat **Slater reguláris**, ha van Slater pont.

Az előző példánál (a) esetben tetszőleges $(x_1, 0)$ pont Slater pont, tehát a feladat Slater reguláris, míg a (b) esetben nincs Slater pont, mert 0 az egyetlen megengedett pont, ami nem teljesíti a *iii)* tulajdonságot: $g_2(0) = 0$, de g_2 nemlineáris.

8.4.2. Tétel (Karush-Kuhn-Tucker). *Legyen C relatív nyílt és (KO) Slater reguláris. Ekkor $\bar{x} \in C_g$ optimális megoldása a (KO) feladatnak $\Leftrightarrow \exists \mu_i \geq 0 \quad (i = 1, \ldots, m)$: $\mu_i = 0$ ha $g_i(\bar{x}) < 0$, és*

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) \text{ } \textit{merőleges} \text{ } \mathrm{lin}(C)\textit{-re.}$$

Bizonyítás. $\Leftarrow$: Legyen $I = \{i \in \{1, \ldots, m\} : g_i(\bar{x}) = 0\}$. Vegyünk egy tetszőleges $s \in K(\bar{x})$ vektort, ami persze $\mathrm{lin}(C)$-ben van. Ekkor tudjuk, hogy

$$0 = s^T(\nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x})) = s^T \nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}).$$

De $\sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}) \leq 0$ mert $\sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}) = \sum_{i \in I} \mu_i s^T \nabla g_i(\bar{x})$, és $i \in I$ esetén $s^T \nabla g_i(\bar{x}) \leq 0$, mert s megengedett irány és $g_i(\bar{x}) = 0$. Tehát $0 \leq s^T \nabla f(\bar{x}) \Rightarrow \bar{x}$ optimális megoldás.

$\Rightarrow$: Indirekt tegyük fel, hogy $\nexists \mu_i \geq 0$, $i \in I$ amire $\nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x})$ merőleges $\mathrm{lin}(C)$-re. A Farkas lemma szerint ekkor $\exists v \in \mathrm{lin}(C)$:

$$\begin{aligned} v^T \nabla g_i(\bar{x}) &\leq 0 \qquad (i \in I), \\ v^T \nabla f(\bar{x}) &< 0. \end{aligned}$$

Legyen x^s Slater pont. Ekkor $g_i(x^s) \leq 0$ $\forall i$, így

- ha $i \in I$: $\underbrace{g_i(x^s)}_{\leq 0} - \underbrace{g_i(\bar{x})}_{=0} \leq 0 \underbrace{\Rightarrow}_{g_i \text{ } konvex} \nabla g_i(\bar{x})(x^s - \bar{x}) \leq 0$,
- ha $i \in I$ és g_i nemlineáris: $g_i(x^s) < 0 \Rightarrow \nabla g_i(\bar{x})(x^s - \bar{x}) < 0$.

Legyen $s = v + \delta(x^s - \bar{x})$, $\delta > 0$. Ekkor tudjuk:

- $s \in \mathrm{lin}(C)$,
- $i \in I$ esetén $\nabla g_i(\bar{x}) s^T \leq 0$,
- ha $i \in I$ és g_i nemlineáris: $\nabla g_i(\bar{x}) s^T < 0$,
- elég kicsi δ-ra $\nabla f(\bar{x}) s^T < 0$, mert $\nabla f(\bar{x}) v^T < 0$.

Válasszuk δ-t ilyen kicsinek. Az így kapott s segítségével találhatunk egy $\bar{x}$-nál kisebb értékű megengedett pontot. Elég kis $\varepsilon > 0$-ra a következők mind teljesülnek:

- $\bar{x} + \varepsilon s \in C$ (mert C relatív nyílt),
- $f(\bar{x} + \varepsilon s) < f(\bar{x})$,

- $i \in I$, g_i nemlineáris: $g_i(\bar{x} + \varepsilon s) < g_i(\bar{x}) = 0$,
- $i \in I$, g_i lineáris: $g_i(\bar{x} + \varepsilon s) \leq g_i(\bar{x}) = 0$,
- ha $i \notin I$: $g_i(\bar{x} + \varepsilon s) < 0$ mert $g_i(\bar{x}) < 0$ és g_i folytonos függvény.

Tehát $\bar{x} + \varepsilon s \in C_g$ és $f(\bar{x} + \varepsilon s) < f(\bar{x})$, azaz $\bar{x}$ nem optimális. •

8.4.2. Lagrange duális

A következőkben a lineáris programozás dualitás-tételének konvex általánosítását fogjuk bebizonyítani a Karush-Kuhn-Tucker tétel segítségével. Ehhez először definiáljuk a Lagrange függvényt, amit úgy is tekinthetünk, hogy a $g_i(x) \leq 0$ feltételek megsértését valamilyen szorzóval (lineárisan) büntetjük. A linearitás miatt viszont jutalmazni kell, ha a feltételek bőven teljesülnek.

Definíció (Lagrange függvény)**.** $x \in C$-re és $\mu \in \mathbb{R}_+^m$-re

$$L(x, \mu) := f(x) + \sum_{i=1}^{m} \mu_i g_i(x)$$

A Lagrange függvény tulajdonságai:

i) Ha $x \in C_g$ akkor $L(x, \mu) \leq f(x)$;

ii) Ha $x \in C \setminus C_g$ akkor $\exists \mu$: $L(x, \mu) > f(x)$;

Adott μ-re legyen $L(\mu) := \inf\{L(x, \mu) : x \in C\}$. Az első tulajdonságból következik, hogy $L(\mu) \leq \inf_{x \in C_g} f(x)$. Az $L(\mu)$ érték kiszámolása egy feltétel nélküli konvex optimalizálási feladat.

Definíció. A $\sup_{\mu \in \mathbb{R}_+^m} L(\mu)$ feladatot **Lagrange duális feladatnak** nevezzük.

8.2. Feladat. *Mutassuk meg, hogy a Lagrange duális feladat is egy konvex optimalizálási feladat.*

8.4.3. Tétel (Gyenge dualitás)**.** *Ha $L(\bar{\mu}) = f(\bar{x})$, $\bar{\mu} \in \mathbb{R}_+^m$, $\bar{x} \in C_g$, akkor $\bar{x}$ optimális megoldása a (KO) feladatnak, és $\bar{\mu}$ optimális megoldása a Lagrange duális feladatnak.*

Bizonyítás. $L(\mu) \leq f(\bar{x}) = L(\bar{\mu}) \leq f(x)$ tetszőleges $\mu \in \mathbb{R}_+^m$-ra és $x \in C_g$-re. •

8.4.4. Tétel (Erős dualitás)**.** *Ha C relatív nyílt, a (KO) feladat Slater reguláris és $\bar{x}$ optimális megoldás, akkor $\exists \bar{\mu} \in \mathbb{R}_+^m$, amire $L(\bar{\mu}) = f(\bar{x})$.*

Bizonyítás. Ha $\bar{x}$ optimális, akkor a Karush-Kuhn-Tucker tétel szerint $\exists \bar{\mu} \in \mathbb{R}_+^m$, amire

(1) $\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i \nabla g_i(\bar{x})$ merőleges $\mathrm{lin}(C)$-re;

(2) ha $g_i(\bar{x}) < 0$, akkor $\bar{\mu}_i = 0$.

Először belátjuk, hogy $\bar{x}$ optimális megoldása a $\min_{x\in C} L(x,\bar{\mu})$ feladatnak, azaz $L(\bar{\mu}) = L(\bar{x},\bar{\mu})$.

A $\min_{x\in C} L(x,\bar{\mu})$ feladat konvex függvény minimalizálása relatív nyílt halmazon. Tudjuk, hogy $\bar{x}$ optimális $\Leftrightarrow \nabla_{x=\bar{x}} L(x,\bar{\mu})$ merőleges $\mathrm{lin}(C)$-re.

$$\nabla_{x=\bar{x}} L(x,\bar{\mu}) = \nabla_{x=\bar{x}}(f(x) + \sum_{i=1}^{m} \bar{\mu}_i g_i(x)) = \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i \nabla g_i(\bar{x})$$

ami merőleges $\mathrm{lin}(C)$-re (1) szerint.

Ezután belátjuk, hogy $L(\bar{x},\bar{\mu}) = f(\bar{x})$. Definíció szerint $L(\bar{x},\bar{\mu}) = f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i g_i(\bar{x})$. De $\sum_{i=1}^{m} \bar{\mu}_i g_i(\bar{x}) = 0$, mert

- (2) szerint: ha $g_i(\bar{x}) < 0$, akkor $\bar{\mu}_i = 0$;
- $\bar{x} \in F \Rightarrow g_i(\bar{x}) \le 0 \ \forall i \in \{1,\dots,m\}$.

Tehát beláttuk, hogy $L(\bar{\mu}) = L(\bar{x},\bar{\mu}) = f(\bar{x})$. •

Megmutatjuk, hogy lineáris programozási feladat esetén ez a tétel tényleg az erős LP dualitásnak (4.2.3 tétel) felel meg. Tekintsük az $R := \{x : Qx \le b\}$ nemüres poliédert, ahol $Q \in \mathbb{R}^{m\times n}$, és legyen $c \in \mathbb{R}^n$ a maximalizálandó célfüggvény. A mostani jelöléseinkkel $f(x) = -c^T x$ és $g_i(x) = q_{i.}x - b_i$, tehát

$$L(\mu) = \inf_{x\in\mathbb{R}^n} -c^T x + \mu^T(Qx - b) = \begin{cases} -\infty \text{ ha } \mu^T Q \ne c^T, \\ -\mu^T b \text{ ha } \mu^T Q = c^T. \end{cases}$$

A feladat Slater-reguláris, mert R nemüres és minden g_i lineáris. Ha $\bar{x}$ optimális megoldás, akkor a fenti 8.4.4 tétel szerint létezik $\bar{\mu} \in \mathbb{R}^m_+$, amire $L(\bar{\mu}) = f(\bar{x})$. Ez pont azt jelenti, hogy $\bar{\mu}^T Q = c^T$, tehát $\bar{\mu}$ megoldása a duális feladatnak, és $\bar{\mu}^T b = c^T x$.

8.5. Megoldási módszerek

Konvex programozási feladatokat általában nem lehet a szimplex módszerhez hasonló módon megoldani, mert az optimum nem feltétlenül éretik el extremális pontban, és ha igen, akkor is lehet végtelen sok extremális pont. Ennek ellenére a konvexitás sokat segít a megoldhatóságban, és fontos speciális eseteket (például a szemidefinit feladatokat) meg lehet oldani polinom időben. Az alábbiakban néhány egyszerű általános módszert ismertetünk, amik konvergálnak az optimumhoz.

8.5.1. Megengedett csökkenési irány keresése

Definíció. Az s vektor **megengedett csökkenési irány** x^0-ban, ha $\exists \varepsilon > 0$, hogy tetszőleges $0 < \lambda \le \varepsilon$-ra $f(x^0 + \lambda s) < f(x^0)$ és $x^0 + \lambda s \in F$.

Megengedett csökkenési irány x^0-ban a következő LP feladat megoldásával kereshető:

$$\begin{aligned} &\max u && \\ \nabla f(x^0)^T s + u &\le 0 && \\ \nabla g_i(x^0)^T s + u &\le 0 && \text{ha } g_i(x^0)=0 \text{ és } g_i \text{ nem lineáris,} \\ \nabla g_i(x^0)^T s &\le 0 && \text{ha } g_i(x^0)=0 \text{ és } g_i \text{ lineáris,} \\ u &\ge 0 && \\ -1 \le s_j &\le 1 && (j=1,\dots,n), \\ s &\in \mathrm{lin}(C) && \end{aligned}$$

8.5.1. Állítás. *Ha $\exists (s,u)$ megoldás, ahol $u>0$, akkor s megengedett csökkenési irány.*

Bizonyítás. $\nabla f(x^0)^T s < 0$, tehát s csökkenési irány. Másrészt $\nabla g_i(x^0)^T s \le 0$ ha $g_i(x^0)=0$, és $\nabla g_i(x^0)^T s < 0$ ha g_i nemlineáris, tehát s megengedett irány. •

8.5.2. Állítás. *Ha a feladat Slater reguláris, akkor ha x^0 nem optimális, akkor $\exists(s,u)$ megoldás, amire $u>0$.*

Bizonyítás. Ha x^0 nem optimális, akkor $\exists v \in \mathrm{lin}(C)$:

$$\begin{aligned} v^T \nabla g_i(\bar{x}) &\le 0 \qquad (i=1,\dots,m), \\ v^T \nabla f(\bar{x}) &< 0. \end{aligned}$$

Legyen x^s Slater pont. Ekkor $g_i(x^s) \le 0\ \forall i$, így

- ha $g_i(\bar{x})=0$: $\underbrace{g_i(x^s)}_{\le 0} - \underbrace{g_i(\bar{x})}_{=0} \le 0 \underbrace{\Rightarrow}_{g_i \text{ konvex}} \nabla g_i(\bar{x})(x^s-\bar{x}) \le 0.$

- ha $g_i(\bar{x})=0$ és g_i nemlineáris: $g_i(x^s)<0 \Rightarrow \nabla g_i(\bar{x})(x^s-\bar{x})<0$.

Legyen $s = v + \delta(x^s-\bar{x})$, $\delta>0$. Ekkor tudjuk:

- $s \in \mathrm{lin}(C)$;

- ha $g_i(\bar{x})=0$, akkor $\nabla g_i(\bar{x}) s^T \le 0$;

- ha $g_i(\bar{x})=0$ és g_i nemlineáris, akkor $\nabla g_i(\bar{x}) s^T < 0$;

- elég kicsi δ-ra $\nabla f(\bar{x}) s^T < 0$, mert $\nabla f(\bar{x}) v^T < 0$.

Válasszuk δ-t elég kicsinek az utolsó feltételnek megfelelően. Szorozzuk meg s-t egy skalárral úgy, hogy $-1 \le s_j \le 1\ \forall j$. Ehhez az s-hez választhatunk megfelelő $u>0$-t. •

8.5.2. Gradiens módszer

Legyen $f : \mathbb{R}^n \to \mathbb{R}$ konvex, folytonosan differenciálható függvény. Ennek szeretnénk a minimumát megtalálni. Tegyük fel, hogy $x^0 \in \mathbb{R}^n$ pontban vagyunk.

8.5.3. Állítás. *x^0-ban az iránymenti derivált a $-\nabla f(x^0)$ irányban a legkisebb (azonos normájú vektorok közül).*

Bizonyítás. s irányban az iránymenti derivált: $\nabla f(x^0)^T s$.

$$-\nabla f(x^0)^T \nabla f(x^0) = \min_{||s||=||\nabla f(x^0)||} \nabla f(x^0)^T s$$

Tehát $-\nabla f(x^0)$ a legmeredekebb csökkenési irány. •

Algoritmus
$x^0 \in \mathbb{R}^n$, $k = 0$.

1. Kiszámoljuk $s = -\nabla f(x^k)$-t. Ha $s = 0$, akkor kész vagyunk.
2. Legyen $\lambda^k = \text{argmin}_{\lambda \in \mathbb{R}_+} f(x^k + \lambda s)$.
3. Legyen $x^{k+1} = x^k + \lambda^k s$.
4. $k := k + 1$, és tovább az 1. lépésre.

8.5.4. Állítás. $f(x^{k+1}) < f(x^k)$

Bizonyítás. s irányban az iránymenti derivált: $-\nabla f(x^k)^T \nabla f(x^k) < 0$. •

8.5.1. Tétel. *Ha $D = \{x \in \mathbb{R}^n : f(x) \leq f(x^0)\}$ kompakt, akkor az $x^0, x^1, x^2, \dots$ tetszőleges $\bar{x}$ torlódási pontja optimális megoldása a feladatnak.*

Bizonyítás. Vegyünk egy $x^{j_1}, x^{j_2}, \dots \to \bar{x}$ konvergens részsorozatot. Mivel f folytonos, ezért $\lim_{i\to\infty} f(x^{j_i}) = f(\bar{x})$ és $\lim_{i\to\infty} \nabla f(x^{j_i}) = \nabla f(\bar{x})$. Legyen $s = -\nabla f(\bar{x})$. Ekkor $\nabla f(\bar{x})^T s \leq 0$ és $= 0 \Leftrightarrow s = 0$.

Másrészt: tetszőleges $\lambda \geq 0$-ra és i-re $f(x^{j_{i+1}}) \leq f(x^{j_i+1}) \leq f(x^{j_i} - \lambda \nabla f(x^{j_i}))$. Az $i \to \infty$ határátmenet mutatja, hogy $f(\bar{x}) \leq f(\bar{x} + \lambda s)$ tetszőleges $\lambda \geq 0$-ra. Tehát $\frac{f(\bar{x}+\lambda s) - f(\bar{x})}{\lambda} \geq 0$. Ha $\lambda \to 0$, akkor azt kapjuk, hogy $\nabla f(\bar{x})^T s \geq 0 \Rightarrow s = 0 \Rightarrow \nabla f(\bar{x}) = 0 \Rightarrow \bar{x}$ optimális. •

8.5.3. Aranymetszés módszer

A fenti módszerben a $\lambda^k = \text{argmin}_{\lambda \in \mathbb{R}_+} f(x^k + \lambda s)$ kiszámolásához meg kell oldanunk egy egydimenziós konvex függvény minimalizálási feladatot. Persze ezt is csak közelítőleg tudjuk megoldani. Egy lehetséges megoldási módszer az Aranymetszés módszer, ami nem csak differenciálható függvényekre alkalmazható.

Legyen f konvex függvény $\mathbb{R}$-en. Az algoritmus egyre csökkenő méretű intervallumokat ad, amik garantáltan tartalmaznak optimális megoldást. Minden lépésben adott lesz $\alpha_i < \beta_i < \gamma_i < \delta_i$, amikre $f(\alpha_i) \geq f(\beta_i)$, $f(\gamma_i) \leq f(\delta_i)$, és

$$\frac{\gamma_i - \alpha_i}{\delta_i - \alpha_i} = \frac{\delta_i - \beta_i}{\delta_i - \alpha_i} = q := \frac{\sqrt{5} - 1}{2}.$$

Kiindulásként $i = 0$-ra találunk ilyeneket. A konvexitásból következik, hogy van optimális megoldás az $[\alpha_i, \delta_i]$ intervallumban.

Az aranymetszés szabálya miatt az is teljesül, hogy

$$\frac{\beta_i - \alpha_i}{\gamma_i - \alpha_i} = \frac{\delta_i - \gamma_i}{\delta_i - \beta_i} = q.$$

Az általános lépésben összehasonlítjuk az $f(\beta_i)$ és $f(\gamma_i)$ értékeket. Két eset van:

- Ha $f(\beta_i) < f(\gamma_i)$: legyen

$$\begin{aligned} \alpha_{i+1} &= \alpha_i \\ \beta_{i+1} &= q\alpha_i + (1-q)\gamma_i \\ \gamma_{i+1} &= \beta_i \\ \delta_{i+1} &= \gamma_i \end{aligned}$$

- Ha $f(\beta_i) \geq f(\gamma_i)$: legyen

$$\begin{aligned} \alpha_{i+1} &= \beta_i \\ \beta_{i+1} &= \gamma_i \\ \gamma_{i+1} &= q\delta_i + (1-q)\beta_i \\ \delta_{i+1} &= \delta_i \end{aligned}$$

Könnyen ellenőrizhető, hogy az új pontok is teljesítik a feltételeket, az $[\alpha_k, \delta_k]$ intervallum hossza q-szorosára csökken, és csak egy új függvényértéket kell kiszámolni.

8.5.4. Newton módszer

Tegyük fel, hogy $f : \mathbb{R}^n \to \mathbb{R}$ kétszer folytonosan differenciálható konvex függvény.

Legyen f másodrendű közelítése:

$$q_{x^0}(x) = f(x^0) + \nabla f(x^0)^T(x - x^0) + \frac{1}{2}(x - x^0)^T \nabla^2 f(x^0)(x - x^0).$$

Ekkor $\nabla^2 q_{x^0}(x) = \nabla^2 f(x^0)$ pozitiv szemidefinit, mert f konvex, tehát $q_{x^0}(x)$ konvex. Minimalizáljuk $q_{x^0}(x)$-et:

$$\bar{x} \text{ optimális} \Leftrightarrow \underbrace{\nabla q_{x^0}(\bar{x})}_{=\nabla f(x^0) + \nabla^2 f(x^0)(\bar{x} - x^0)} = 0$$

Tegyük fel, hogy $\nabla^2 f(x^0)$ invertálható:

$$\bar{x} \text{ optimális} \Leftrightarrow \bar{x} = x^0 - (\nabla^2 f(x^0))^{-1} \nabla f(x^0).$$

Algoritmus. Tegyük fel, hogy f szigorúan konvex (tehát $\nabla^2 f(x)$ pozitiv definit). $x^0 \in \mathbb{R}^n$, $k = 0$.

1. Kiszámoljuk $\nabla f(x^k)$-t és $\nabla^2 f(x^k)$-t. Ha $\nabla f(x^k) = 0$, akkor kész vagyunk.
2. $x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$.
3. $k := k + 1$ és tovább az 1. lépésre.

Megjegyzés. Az egydimenziós esetben $x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$.

Megjegyzés. Ha $\nabla^2 f(x^k)$ nem invertálható, akkor vegyük helyette $\nabla^2 f(x^k) + \alpha I$-t valamilyen megfelelően választott α-ra.

Ajánlott irodalom

- *Operációkutatás példatár*, ELTE, szerk.: Bérczi Kristóf, Frank András, Kaszanitzky Viktória, Király Csaba, Király Tamás, Kovács Erika Renáta, Pap Gyula, Pap Júlia, `www.tankonyvtar.hu` (2013)

- Gáspár László, Temesi József, *Lineáris programozási gyakorlatok*, Nemzeti Tankönyvkiadó (2002)

- Gáspár László, Temesi József, *Matematikai programozási gyakorlatok*, Nemzeti Tankönyvkiadó (1999)

- Alexander Schrijver, *Theory of Linear and Integer Programming*, Wiley Series in Discrete Mathematics & Optimization (1998)

- George B. Dantzig, Mukund N. Thapa, *Linear Programming 1: Introduction*, Springer Series in Operations Research and Financial Engineering (1997)

- Dimitris Bertsimas, John N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific Series in Optimization and Neural Computation, 6 (1997)

- Jiří Matoušek, Bernd Gärtner, *Understanding and Using Linear Programming*, Springer (2006)