ELTE TTK, Operációkutatási Tanszék

Frank András

OPERÁCIÓKUTATÁS

2011. október 29.

1. Fejezet

OPTIMALIZÁLÁS GRÁFOKON

Ebben a fejezetben gráfokra vonatkozó optimalizálási problémákat vizsgálunk, bemutatva a megoldásukra szolgáló algoritmusokat.

1.1 ALGORITMUSOK HATÉKONYSÁGÁRÓL

Egy algoritmustól elsősorban azt várjuk, hogy véges legyen, de ez még édeskevés, hiszen ha az alapstruktúra véges, akkor többnyire nem okoz nehézséget a véges sok lehetőséget mind számba venni. Ugyanakkor e 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 vagy különbözőt.

Állítás 1.1.1 *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. Általános tapasztalat, hogy exponenciális vagy nagyobb lépésszámú algoritmusok a gyakorlat számára gyakran használhatatlanok.

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.

Az egyszerűség kedvéért tekintsük a színezési problémát csak a $k = 2$ és a $k = 3$ esetben. Egy algoritmus megadásakor, vagy akár azt megelőzően, természetes kívánság, hogy legyen egy olyan eszközünk, amelynek segítségével egy rendelkezésünkre bocsátott válasz helyességét gyorsan ellenőrizni tudjuk, anélkül, hogy a válaszhoz vezető számítás részleteit át kéne vizsgálnunk. Például egy polinom valamely gyökét megtalálni nem éppen egyszerű feladat, de egy valahonnan megkapott gyök-jelölt helyessége behelyettesítéssel könnyen ellenőrizhető. Hasonlóképp, ha valaki betoppan a gráf pontjainak egy 2- vagy egy 3-színezésével, azt gyorsan tudjuk ellenőrizni, hogy a színezés jó-e, vagyis azt, hogy minden él két végpontja tényleg különböző színű-e. Ha azonban az a válasz, hogy az illető gráf pontjainak nem létezik jó 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.

TÉTEL 1.1.2 *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 szíműek, 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 $P_2[v, z]$ részút élszáma azonos paritású, tehát az általuk és az uv él által alkotott kör páratlan elemszámú. •

TÉTEL 1.1.3 *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 legfeljebb 2.*

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. •

TÉTEL 1.1.4 *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 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 bizonyításaik is nagyjából egyforma nehézségűek, 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, FIGYELEM: az NP nem a polinomialitás tagadását jelzi!) Azt mondjuk, hogy egy tulajdonság co-NP-ben van, ha a tulajdonság nem meglétére 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. (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.)

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 eszmefuttatásnak 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 ahhoz hasonlít, mint amikor egy függvény folytonosságáról beszélünk. Egyrészt él bennünk egy szemlelétes 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). **Sétán** 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. A séta **zárt**, ha $v_0 = v_k$. 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 zárt séta, amelyben a $v_0 = v_k$ egybeeséstől eltekintve a csúcsok mind különbözőek. Amennyiben a W séta egy $K = (v_i, e_{i+1}, v_{i+1}, \ldots, e_j, v_j)$ rész-sétája kör (ahol $0 \le i < j \le k$), 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.)

Állítás 1.2.1 *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.)

Gyakorlat 1.2.1 *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.

Gyakorlat 1.2.2 *Egy irányított fa akkor és csak akkor s-fenyő, ha az $s \in S$ pont befoka nulla a többi ponté pedig egy.*

Gyakorlat 1.2.3 *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.

TÉTEL 1.2.2 *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.

Feladat 1.2.4 *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.

Gyakorlat 1.2.5 *Igazoljuk, hogy a BFS algoritmus helyesen határozza meg az s-től való távolságot.*

Gyakorlat 1.2.6 *Igazoljuk, hogy irányítatlan gráfban a távolság függvény kielégíti a háromszög egyenlőtlenséget.*

Gyakorlat 1.2.7 *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 **laminárisnak** 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.

Gyakorlat 1.2.8 *Igazoljuk, hogy legalább háromtagú lamináris sorozatnak (amelyben minden elem kétszer fordul elő) van x, y, y alakú három egymást követő elemből álló része.*

A mélységi kereséssel kapott fenyőt (amely persze nem egyértelmű) DFS fenyőnek hívjuk. 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 **keresztélnek**, 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-nek számos 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.

Feladat 1.2.9 *Igazoljuk, hogy ha a gráf 2-élösszefüggő, akkor az így kapott irányítás erősen összefüggő.*

A DFS másik alkalmazása aciklikus digráfban a pontok ún. topológikus sorrendjének meghatározására szolgál (a **topológikus sorrend** olyan, amelyre vonatkozólag minden él előre felé vezet.) 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 topológikus sorrendjéből az újonnan hozzáadott s-t kihagyva megkapjuk az eredeti digráf pontjainak egy topológikus sorrendjét.)

Feladat 1.2.10 *Igazoljuk, hogy aciklikus digráf elhagyási sorrendjének megfordítása topológikus sorrendet ad.*

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.

Gyakorlat 1.2.11 *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 topológikus 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.)

Feladat 1.2.12 *Igazoljuk, hogy a második fázisban kapott fordított fenyves komponensei éppen a digráf erősen összefüggő komponensei lesznek.*

A mélységi keresésnek elméleti alkalmazásai is vannak.

Feladat 1.2.13 *Bizonyítsuk be, hogy egy legalább két pontú, összefüggő, egyszerű irányítatlan gráfban, ha minden pont foka legalább három, akkor létezik úgynevezett periférikus kör, azaz egy olyan húrmentes C kör, amelynek pontjait elhagyva összefüggő gráfot kapunk.*

(Segítség: az s gyökerű mélységi fához válasszunk egy olyan xy nem-fa élt, amelynek s-hez közelebbi x pontja s-től a lehető legtávolabb van a fában, és ezen belül y a lehető legközelebb. Ekkor az xy alapköre periférikus.)

A periférikus köröknek fontos szerepük van síkbarajzolható gráfok síkbarajzolásánál. Nevezetesen, nem nehéz belátni, hogy 3-összefüggő síkgráfban egy kör akkor és csak akkor periférikus, ha a síkbaágyazásnál tartományt határol: íly módon a tartományt tisztán kombinatorikusan lehet definiálni. W.T. Tutte bebizonyította, hogy egy 3-összefüggő síkgráfnak a következő módon lehet egy síkbarajzolását meghatározni. Keressünk egy C periférikus kört. Ennek pontjait helyezzük el a síkban úgy, hogy konvex poligont feszítsenek. Tekintsük azt az egyenletrendszert, amely azt írja le, hogy a gráf minden nem C-ben lévő csúcsa a szomszédos csúcsok súlypontjában van. Ezen egyenletrendszer megoldása, amint azt Tutte bebizonyította, a pontoknak olyan elhelyezését szolgáltatja, amelyben különböző csúcsoknak különböző pontok felelnek meg, és a gráfban szomszédos csúcsoknak megfeleltetett síkbeli pontokat szakaszokkal összekötve a gráfnak olyan síkbarajzolást kapjuk, amelyben a korlátos tartományok konvexek.

Feladat 1.2.14 *Periférikus köröket használva igazoljuk G. Dirac azon tételét, mely szerint egy gráf akkor és csak akkor soros-párhuzamos, ha nem tartalmaz részgráfként felosztott teljes négyest.* (A felosztott teljes négyes a négypontú teljes gráfból áll elő úgy, hogy az éleket [a végpontoktól eltekintve] diszjunkt utakkal helyettesítjük.)

(Definíció szerint egy **soros-párhuzamos gráf** egy pontból kiindulva az alábbi műveletek tetszés szerinti egymás utáni alkalmazásával áll elő: (1) egy új él hozzáadása, melynek egyik vége régi pont, a másik viszont új, (2) egy meglévő élt egy osztóponttal felosztunk, (3) egy meglévő éllel párhuzamos élt behúzunk.)

Feladat 1.2.15 *Tegyük fel, hogy $G = (V, E)$ irányított, összefüggő Euler-gráf, azaz minden pontnak ugyanannyi a befoka, mint a kifoka. Legyen s tetszőleges pont és F s-gyökerű fordított feszítő fenyő. Az s-ből kiindulva minden u pontnál tetszés szerint haladjunk tovább egy még nem használt uv élen, csak arra ügyelve, hogy $u \neq s$ esetén az u-ból kilépő egyetlen F-beli élt csak akkor vegyük igénybe, ha már más lehetőségünk nincsen. Az eljárás akkor ér véget, amikor egy olyan ponthoz érünk, amelynek már minden kimenő élét használtuk. Igazoljuk, hogy ez a végső pont s és az eljárás G-nek egy Euler-bejárását szolgáltatja, azaz minden élt pontosan egyszer használ.*

2011. október 29. file: linp, eler

1.3 OPTIMÁLIS UTAK, SÉTÁK, POTENCIÁLOK

Tegyük fel, hogy a $D = (V, A)$ n pontú és m élű hurokmentes irányított gráf élein adott egy $c : A \to \mathbf{R}$ 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 legolcsóbb, ha a P kezdőpontjából a végpontjába nem létezik D-ben P-nél olcsóbb út.

Gyakorlat 1.3.1 *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 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.

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

Az s rögzítettsége miatt van az, hogy 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.

1.3.1 Konzervatív költségfüggvények, megengedett potenciálok

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, hanem csupán arra az szorítkozunk, amikor nem létezik negatív összköltségű egyirányú kör. Ilyenkor a költségfüggvényről azt mondjuk, hogy **konzervatív**. Például c biztosan konzervatív, ha nem-negatív, vagy akkor is, ha tetszőleges, de a digráfban egyáltalán nincsen egyirányú kör, azaz D **aciklikus**. Nem nehéz egyéb konzervatív költségfüggvényeket konstruálni: ilyen például, ha D egy 3 élű egyirányú kör, melyen a költségek $-1, +1, +1$.

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 $\pi : V \to \mathbf{R}$ függvényt c-re nézve **megengedett potenciál**nak, vagy **c-megengedett**nek, ha

$$\pi(v) - \pi(u) \leq c(uv) \text{ fennáll minden } uv \in A \text{ élre.} \tag{1.1}$$

A π által az élhalmazon indukált $\Delta_\pi : A \to \mathbf{R}$ 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. A π c-megengedettsége tehát azt jelenti, hogy $\Delta_\pi \leq c$ vagy másképp írva, $c_\pi \geq 0$, ahol

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

Egyszerű, de annál hasznosabb a következő megfigyelés.

Lemma 1.3.1 *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: $\widetilde{\Delta}_\pi(K) = \sum_{i=1}^{k} \Delta_\pi(v_i v_{i+1}) = \sum_{i=1}^{k} [\pi(v_{i+1}) - \pi(v_i)] = 0$. •

Az állítás (b) része szerint egy pontindukált költségfüggvény semleges abban az értelemben, hogy minden st-útnak ugyanannyi a költsége. Ebből az következik, hogy egy P st-út akkor és csak akkor legolcsóbb a c-re nézve, ha legolcsóbb c_π-re nézve. Ezért ha adott c-hez tudnánk találni egy megengedett potenciált, akkor a nem-negatív költségfüggvényre vonatkozó Dijkstra algoritmus már alkalmazható lenne. A következő tétel lényegi része épp azt mondja ki, hogy konzervatív költségfüggvény esetén mindig létezik megengedett potenciál.

TÉTEL 1.3.2 (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 1.3.1 lemma nyomán $\widetilde{c}(K) = \widetilde{c}_\pi(K) \geq 0$.

A fordított irány igazolásához induljunk ki egy tetszőleges π potenciálból, amely egészértékű, ha c az. Ha $c_\pi := c - \Delta_\pi$ nem-negatív, akkor π definíció szerint megengedett. 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 ez nem sikerül, 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.

$$\pi(v) := \begin{cases} \pi(v) - \varepsilon, & \text{ha } v \in Z \\ \pi(v), & \text{ha } v \in V - Z, \end{cases} \tag{1.2}$$

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

$$\varepsilon_1 := \min\{c_\pi(e) : e \in A, e \text{ kilép } Z\text{-ből}\}. \tag{1.3}$$

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 (1.3)-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 figyeljük meg, hogy egészértékű c esetén az eljárás végig fenntartja π egészértékűségét. $\bullet$

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

A Gallai tétel segítségével a legolcsóbb st-út kiszámításának problémáját konzervatív költségfüggvény esetén könnyen visszavezethetjük arra az esetre, amikor a költségfüggvény nem-negatív, és ekkor már alkalmazhatjuk Dijkstra algoritmusát. Valóban, ha c konzervatív, akkor létezik egy megengedett π potenciál, vagyis olyan, amelyre $c_\pi \geq 0$. Miután minden P st-útra $\widetilde{c}_\pi(P) = \widetilde{c}(P) - \pi(t) + \pi(s)$, vagyis a $\widetilde{c}(P) - \widetilde{c}_\pi(P)$ érték nem függ a P úttól, a c_π nem-negatív költségfüggvényre nézve legolcsóbb st-út az eredeti c-re is legolcsóbb. E megközelítés hátránya, hogy a Dijkstra algoritmus ugyan $O(n^2)$ lépésszám igényű, de a Gallai tétel bizonyításában ismertetett algoritmus megengedett potenciál kiszámítására $O(m^2)$-es, vagyis ezen nagyobb szám jellemzi az algoritmus teljes lépésszámát. A következő szakaszban bemutatunk egy másik megközelítést, amely amely egyáltalán nem támaszkodik a Dijkstra algoritmusra és amelynek $O(mn)$-es lépésszáma jobb, mint a előbbi, bár a Dijkstra igényelte $O(n^2)$-et nem éri el.

Dijkstra algoritmusa a Véges Matematika c. előadásban az első év során már szerepelt: a fejezet végén ismétlésképp röviden áttekintjük az alapideáját.

Gyakorlat 1.3.2 *Mutassuk meg, hogy konzervatív c-re létezik nem-pozitív megengedett potenciál.*

Egy $x : A \to \mathbf{R}$ függvényt **tenzió**nak nevezünk, ha a digráf minden C köre **semleges** abban az értelemben, hogy a C egyik irányba menő élein az x-összeg egyenlő a másik irányú élek x-összegével. Speciálisan ez azt jelenti, hogy minden egyirányú K körre $\widetilde{x}(K) = 0$.

Feladat 1.3.3 *Mutassuk meg, hogy egy $x : A \to \mathbf{R}$ függvény akkor és csak tenzió, ha potenciál különbség.*

Feladat 1.3.4 *Erősen összefüggő digráfban egy $x : A \to \mathbf{R}$ függvény akkor és csak akkor tenzió, ha $\widetilde{x}(K) = 0$ minden K egyirányú körre.*

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

Feladat 1.3.5 *Egy $x : A \to \mathbf{R}$ függvény akkor és csak tenzió, ha egy adott T feszítő fához tartozó valamennyi* (tehát $m - n + 1$) *alapkör semleges.*

Feladat 1.3.6 *Adjunk az 1.3.2 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.*

Feladat 1.3.7 *Igazoljuk, hogy ha* π_1 *és* π_2 *c-megengedett potenciálok, akkor a* π*-vel jelölt maximumuk is az, ahol* $\pi(v) := \max\{\pi_1(v), \pi_2(v)\}$ $(v \in V)$.

Feladat 1.3.8 *Igazoljuk, hogy egészértékű konzervatív költségfüggvény esetén létezik egy egyértelmű legnagyobb nem-pozitív megengedett potenciál. Elhagyható-e az egészértékűségi feltevés?*

1.3.2 Legolcsóbb séták tetszőleges költségfüggvényre

Visszatérve a legolcsóbb út problémájára, egyszerűbb a helyzet, amikor út helyett legolcsóbb sétát keresünk (amelyben tehát egy ponton vagy élen többször is végighaladhatunk). Ha a digráfban léteznek negatív összköltségű egyirányú körök (röviden negatív körök), akkor az sv-séták költsége esetleg nem korlátos alulról. Ez a helyzet például, ha D maga egy kör, amelynek minden éle -1 költségű. Emiatt célszerű csupán a legfeljebb j élű sv-sétákkal foglalkozni valamely előre adott j egészre. Mivel az ilyen séták száma véges, biztosan van közöttük legolcsóbb. Két rokon, egymással ekvivalens feladatot tekintünk.

Legolcsóbb sv-séták

Egy kijelölt s pontból minden v csúcsra a legolcsóbb legfeljebb i élű sv-sétát akarjuk meghatározni $i = 0, 1, 2, \dots$ esetén. Ennek költségét jelölje $\mu_c^{(i)}(v)$. Ha nem létezik legfeljebb i élű sv-séta, akkor az $\mu_c^{(i)}(v) = \infty$ jelölést fogjuk használni.

A $\mu_c^{(i)}(v)$ költségeket könnyű kiszámítani egymás után az $i = 0, 1, 2, \dots$ értékekre, hiszen egy legfeljebb $k+1$ élű sv-séta vagy pontosan $k+1$ élből áll vagy legfeljebb k élből. Ebből adódik az alábbi rekurzió. Legyen $\mu_c^{(0)}(s) = 0$ és minden $v \in V$-re $\mu_c^{(0)}(v) \equiv \infty$. Amennyiben a $\mu_c^{(i)}(v)$ értékek már minden v csúcsra rendelkezésre állnak, úgy legyen

$$\mu_c^{(i+1)}(v) = \min\{\mu_c^{(i)}(v), \min\{\mu_c^{(i)}(u) + c(uv) : uv \in A\}\}. \tag{1.4}$$

Ugyanez a rekurzió használható maguknak a legolcsóbb legfeljebb i élű s-ből v-be vezető $S_c^{(i)}(v)$ sétáknak a megkonstruálására is. Amennyiben nem létezik legfeljebb i élű sv-séta, a $S_c^{(i)}(v) = \{*\}$ jelölést használjuk. A kiindulási $i = 0$ értékre legyen $S_c^i(s) = \{s\}$ és $S_c^i(v) = \{*\}$, ha $v \in V - s$. Ha valamilyen $i \geq 0$-re már minden v-re kiszámítottuk a $S_c^{(i)}(v)$ sétákat, akkor legyen:

$$S_c^{(i+1)}(v) := \begin{cases} \{*\}, & \text{ha } \mu_c^{(i+1)}(v) = \infty \\ S_c^{(i)}(v), & \text{ha } \mu_c^{(i+1)}(v) = \mu_c^{(i)}(v) < \infty \\ S_c^{(i)}(u) + uv, & \text{ha } \mu_c^{(i+1)}(v) = \mu_c^{(i)}(u) + c(uv) < \infty \text{ valamely } uv \in A \text{ élre.} \end{cases} \tag{1.5}$$

Magyarán, a szétválasztás annak megfelelően történik, hogy (1.4) jobb oldalán a külső minimum végtelen vagy véges: amennyiben véges, úgy az első vagy a második tagon vétetik fel, és ha a második tagon, akkor azon belül melyik uv élen. A definícióból adódóan $S_c^{(i+1)}(v)$ egy legolcsóbb legfeljebb $i+1$ élű sv-séta.

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

Legolcsóbb v-ben végződő séták

A sétákra vonatkozó másik feladatban nincs kijelölt kezdőpont. Minden v csúcsra az ott végződő legolcsóbb legfeljebb i élű sétát akarjuk meghatározni $i = 0, 1, 2, \dots$ esetén. Ennek költségét jelölje $\pi_c^{(i)}(v)$. Mivel az egyetlen v pontból álló 0 élű séta költsége 0, így $\pi_c^{(i)}(v) \leq 0$ minden i-re.

A $\pi_c^{(i)}(v)$ költségeket könnyű kiszámítani egymás után az $i = 0, 1, 2, \dots$ értékekre, hiszen egy v-ben végződő legfeljebb $k+1$ élű séta vagy pontosan $k+1$ élből áll vagy legfeljebb k élből. Ebből adódik az alábbi rekurzió. Legyen $\pi_c^{(0)}(v) \equiv 0$ és amennyiben a $\pi_c^{(i)}(v)$ értékek már minden v csúcsra rendelkezésre állnak, úgy legyen

$$\pi_c^{(i+1)}(v) = \min\{\pi_c^{(i)}(v), \min\{\pi_c^{(i)}(u) + c(uv) : uv \in A\}\}. \tag{1.6}$$

Természetesen ugyanez a rekurzió használható maguknak a legolcsóbb legfeljebb i élű v-ben végződő $W_c^{(i)}(v)$ 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:

$$W_c^{(i+1)}(v) := \begin{cases} W_c^{(i)}(v), & \text{ha } \pi_c^{(i+1)}(v) = \pi_c^{(i)}(v) \\ W_c^{(i)}(u) + uv, & \text{ha } \pi_c^{(i+1)}(v) = \pi_c^{(i)}(u) + c(uv) \text{ valamely } uv \in A \text{ élre.} \end{cases} \tag{1.7}$$

A szétválasztás annak megfelelően történik, hogy (1.6) jobb oldalán a külső minimum az első vagy a második tagon vétetik fel, és ha a második tagon, akkor azon belül melyik uv élen. A definícióból adódóan $W_c^{(i+1)}(v)$ egy legolcsóbb legfeljebb $i+1$ élű v-ben végződő séta.

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

Feladat 1.3.9 *Dolgozzunk ki eljárást a legolcsóbb v-ben végződő pontosan i élű séta meghatározására.*

A két feladat ekvivalenciája

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

Lemma 1.3.3 (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

$$\pi_c^{(i)}(v) = \mu_c^{(i+1)}(v) \text{ minden } v \in V\text{-re,} \tag{1.8}$$

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

$$\mu_c^{(i)}(v) = \pi_c^{(i+1)}(v) - M \text{ minden } v \in S\text{-re,} \tag{1.9}$$

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

1.3.3 A Bellman–Ford algoritmus

A Bellman–Ford algoritmus a fentebb ismertetett eljárás apró módosítása, amelynek révén eldönthetjük, hogy egy c költségfüggvény konzervatív-e vagy sem. Az algoritmus futása akkor ér véget, amikor vagy egy negatív kört talált, vagy amikor az i index eléri n-et, azaz amikor meghatároztuk minden v-re a legolcsóbb v-ben végződő legfeljebb n élű sétát, és ráaadásul ezen séták mindegyike út lesz.

Tegyük fel, hogy valamely adott $i \leq n-1$ értékre a $W_c^{(i)}(v)$ legolcsóbb legfeljebb i élű v-ben végződő sétákat már minden v-re meghatároztuk, és tegyük fel azt is, hogy ezen séták mindegyike út. A kezdeti $i=0$ esetben az egyetlen pontból álló $W_c^{(0)}(v) = \{v\}$ séta bizonyosan út. Tekintsük most minden $v \in V$-re az (1.7) rekurzióval definiált $W_c^{(i+1)}(v)$ sétát és vizsgáljuk meg azt az esetet, amikor valamely v-re $W_c^{(i+1)}(v)$ nem út. Miután a feltevés szerint $W_c^{(i)}(u)$ minden $u \in V$ csúcsra út, így szükségképp $W_c^{(i+1)}(v) = W_c^{(i)}(u) + uv$ valamely $uv \in A$ élre és v rajta van a $W_c^{(i)}(u)$ úton. Ezen út v-ig tartó kezdő szakaszát jelölje Q_1, míg a v-től u-ig vezető részútját Q_2.

Tekintsük a $K := Q_2 + uv$ egyirányú kört. Amennyiben $\widetilde{c}(K)$ negatív, úgy az algoritmus ezen K negatív kör kiadásával végetér. Miután $\widetilde{c}(Q_1) = \widetilde{c}(W_c^{(i+1)}(v)) - \widetilde{c}(K)$, kapjuk, hogy $\widetilde{c}(K)$ nem lehet pozitív, mert akkor a Q_1 v-ben végződő út olcsóbb lenne, mint a legolcsóbb $W_c^{(i+1)}(v)$. Szükségképpen $\widetilde{c}(K) = 0$ és emiatt Q_1 is egy v-ben végződő legolcsóbb legfeljebb $i+1$ élű séta, ami ráadásul út. Ekkor az aktuális $W_c^{(i+1)}(v)$ sétát a Q_1 útra cseréljük.

Következmény 1.3.4 *A Bellman–Ford algoritmus vagy megtalál egy negatív kört vagy minden v-re megad egy legolcsóbb v-ben végződő legfeljebb n élű sétát, ami út.* •

A leírt módosítás használható az s-ből induló legolcsóbb séták esetén is. Ekkor tehát az algoritmus vagy megtalál egy negatív kört vagy minden v-re megad egy legolcsóbb legfeljebb n élű sv-sétát, ami út. Következésképp, az utóbbi esetben kapott út egy legolcsóbb sv-út. Az így nyert algoritmus lépésszáma $O(mn)$. Erre a változatra is Bellman–Ford algoritmusként hivatkozunk.

Gallai tételének fenti bizonyításánál egy tetszőleges potenciálból kiindulva azt javítgattuk, amíg vagy egy megengedett potenciált nem kaptunk vagy pedig egy negatív kört. Az alábbiakban megmutatjuk, hogy a Bellman–Ford algoritmus által $O(mn)$ lépésben kiszámolt legolcsóbb legfeljebb n élű v-ben végződő $W_c^{(n)}(v)$ séták költsége közvetlenül megad egy megengedett potenciált. Az alábbi eredmény a Gallai tétel egyfajta finomításának tekinthető.

TÉTEL 1.3.5 *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 $\pi_c^{(n)}(v)$ $(v \in V)$ potenciál megengedett.*

Biz. (P1)→(P2) nyilvánvaló.

(P2)→(D2) A definícióból adódóan tetszőleges uv élre:

$$\pi_c^{(n)}(v) \leq \pi_c^{(n-1)}(u) + c(uv). \tag{1.10}$$

Az 1.3.4 következmény miatt ha a Bellman–Ford algoritmus nem talál negatív kört, akkor az általa talált v-ben végződő legolcsóbb legfeljebb n élű séta út, emiatt legfeljebb $n-1$ élű, és ezért

$$\pi_c^{(n)}(u) = \pi_c^{(n-1)}(u). \tag{1.11}$$

Ezekből $\pi_c^{(n)}(v) \leq \pi_c^{(n-1)}(u) + c(uv) = \pi_c^{(n)}(u) + c(uv)$, azaz $\pi_c^{(n)}$ valóban megengedett potenciál.

A (D2)→(D1) irány nyilvánvaló. A (D1)→(P1) irány pedig (amint a Gallai tételnél már kimutattuk) abból adódik, hogy egy K egyirányú körre és π megengedett potenciálra $\widetilde{c}(K) = \widetilde{c}_\pi(K) \geq 0$. •

Figyeljük meg, hogy ily módon egy $O(mn)$ lépésszámú algoritmust kaptunk a konzervativitás ellenőrzésére, ami jobb, mint az 1.3.2 tétel bizonyításában szereplő $O(m^2)$-es algoritmus.

A $\pi_c^{(n)}$ és a $\mu_c^{(n)}$ közötti kapcsolat alapján az 1.3.5 tételből rögtön kiolvasható az alábbi.

TÉTEL 1.3.6 *Tegyük fel, hogy a D digráf minden pontja elérhető s-ből. A c költségfüggvényre a következők ekvivalensek.*

(P1) *D-ben nincs negatív kör* (azaz c konzervatív).

(P3) *A Bellman–Ford algoritmus nem talál negatív kört.*

(D1) *Létezik megengedett potenciál.*

(D3) *A $\mu_c^{(n)}(v)$ ($v \in V$) potenciál megengedett.* •

Feladat 1.3.10 *Igazoljuk az* 1.3.2 *tételnek az alábbi* 1.3.7 *tételben megfogalmazott kiterjesztését, amelyben a potenciál-különbségre nem csupán felső, hanem alsó korlátot is előírunk.*

TÉTEL 1.3.7 *A $D = (V, A)$ digráf élhalmazán adott két korlátozó függvény: $c_{al} \leq c_{fel}$. Akkor és csak akkor létezik olyan $\pi : V \to \mathbf{R}$ 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 $c'(uv) := -c_{al}(vu)$.*

Feladat 1.3.11 *Fogalmazzuk meg az* 1.3.7 *tétel általánosítását, ha minden v pontban a $\pi(v)$-re alsó és felső korlát is ki van tűzve.*

Feladat 1.3.12 *Dolgozzunk ki eljárást annak eldöntésére, hogy egy digráf konzervatív súlyozására nézve létezik-e nulla súlyú kör.*

Feladat 1.3.13 *Melyek azok a digráfok, amelyek élhalmazán létezik $\{+1, -1\}$-értékű potenciál-különbség?*

Feladat 1.3.14 *Tegyük fel, hogy c nem konzervatív. Nevezzünk egy v csúcsot hibásnak, ha $\pi_n(v) < \pi_{n-1}(v)$.*
(a) *Mutassuk meg, hogy egy v-ben végződő legolcsóbb legfeljebb n élű séta indukál egy K negatív kört. Igazoljuk, hogy ha c-t minden éllel egységesen a $\widetilde{c}(K)/|K|$ értékkel megemelünk, akkor a keletkező c^+ költségfüggvényre nézve v már nem hibás, továbbá minden c-re nézve hibátlan pont c^+-ra nézve is hibátlan.*

1.3.4 Min-max tétel és optimalitási kritérium

Emlékezzünk, hogy $\pi_c^{(n)}(v)$ jelölte a legfeljebb n élű v-ben végződő legfeljebb n élű séták költségének minimumát, míg ha s előre adott gyökérpont, akkor $\mu_c^{(n)}(v)$ volt a legfeljebb n élű sv-séták költségének minimuma, $\mu_c(v)$ pedig a legolcsóbb sv-út minimális költsége.

Egy uv élt a továbbiakban **pontos**nak fogunk hívni (π-re nézve), ha (1.1)-et egyenlőséggel teljesíti. Jelölje $D_0 = (V, A_0)$ a μ_c-re nézve pontos élek részgráfját.

TÉTEL 1.3.8 *Adott egy D digráf, amelynek minden csúcsa elérhető s-ből, továbbá adott a konzervatív c költségfüggvény. Minden t csúcsra egy P st-út akkor és csak akkor legolcsóbb st-út* (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. Tegyük fel először, hogy P egy st-út D_0-ban. Legyen $c_1(uv) := c(uv) - \mu_c(v) + \mu_c(u)$ minden $uv \in A$ élre. Mivel μ_c megengedett, $c_1 \geq 0$ és így minden D-beli st-út c_1-költsége nem-negatív. Mivel P pontos élekből áll, $\widetilde{c}_1(P) = 0$, és ezért P legolcsóbb st-út c_1-re nézve. De ekkor P legolcsóbb az ekvivalens c-re nézve is.

Megfordítva, tegyük fel, hogy P legolcsóbb st-út D-ben, melynek pontjai legyenek $s = v_0, v_1, \cdots, v_k = t$. Ekkor $\mu_c(t) = \widetilde{c}(P) = \sum[c(v_{i-1}v_i) : i = 1, \ldots, k] \geq \sum[\mu_c(v_i) - \mu_c(v_{i-1}) : i = 1, \ldots, k] = \mu_c(t) - \mu_c(s) = \mu_c(t)$, 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. •

TÉTEL 1.3.9 (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, melynek pontjai $v_1 = s, v_2, \ldots, v_k = t$. Ekkor $\widetilde{c}(P) = \sum c(v_i v_{i+1}) \geq \sum[\pi(v_{i+1}) - \pi(v_i)] = \pi(t) - \pi(s)$, 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 1.3.8 tétel miatt megfelel. •

Kanonikus megengedett potenciál

Az alábbi tétel fő tartalma az, hogy mind $\pi_c^{(n)}$, mind $\mu_c^{(n)}$ egyfajta értelemben kanonikus.

TÉTEL 1.3.10 *Legyen c konzervatív. Ekkor $\pi_c^{(n)}$ megengedett potenciál, továbbá*

$$\pi_c^{(n)} \geq \pi \tag{1.12}$$

minden nem-pozitív megengedett π potenciálra. (Azaz $\pi_c^{(n)}$ az egyértelmű legnagyobb nem-pozitív megengedett potenciál). *Ha adott s gyökérpontból a digráf minden pontja elérhető, akkor $\mu_c^{(n)}$ megengedett potenciál, $\mu_c^{(n)} = \mu_c$ és*

$$\mu_c \geq \pi \tag{1.13}$$

fennáll minden olyan megengedett π potenciálra, amelyre $\pi(s) = 0$. (Azaz $\mu_c^{(n)}$ az egyértelmű legnagyobb megengedett potenciál, amelyre $\pi(s) = 0$).

Biz. $\pi_c^{(n)}$ megengedettségét már láttuk az 1.3.5 tételben. Ebből az 1.3.3 Lemma folytán $\mu_c^{(n)}$ megengedettsége is következik. A Bellman–Ford algoritmusból adódóan egy legolcsóbb legfeljebb n-élű sv-séta választható útnak és így valóban $\mu_c^{(n)} = \mu_c$.

(1.13) igazolásához legyen π olyan megengedett potenciál, amelyre $\pi(s) = 0$. Minden sv-út költsége legfeljebb $\pi(v) - \pi(s) = \pi(v)$, amiből a legolcsóbb sv-út költsége is legfeljebb $\pi(v)$, és így $\pi(v) \leq \mu_c(v)$.

(1.12) igazolásához egészítsük ki a digráfot egy új s ponttal valamint 0 költségű sv élekkel ($v \in V$). Egy-egy értelmű kapcsolat van a keletkezett D^+ digráf legolcsóbb sv-útjai valamint D legolcsóbb v-ben végződő útjai között. Ebből adódóan a legolcsóbb sv-út $\mu_c(v)$ költségére $\mu_c(v) = \pi_c^{(n)}(v) \quad (v \in V)$. Legyen $\pi(s) = 0$. Mivel π nem-pozitív és az új sv-élek költsége 0, ezért π megengedett potenciál a D^+-ban is. Emiatt minden sv-út költsége legfeljebb $\pi(v) - \pi(s) = \pi(v)$, amiből a legolcsóbb sv-út költsége is legfeljebb $\pi(v)$, és így $\pi(v) \leq \mu_c(v) = \pi_c^{(n)}(v)$. •

Feladat 1.3.15 *Ha P olyan st-út, amely egy megengedett potenciálra nézve pontos élekből áll, akkor P legolcsóbb st-út.*

Feladat 1.3.16 *Adott az éleken egy tetszőleges súly-függvény. Határozzunk meg egy olyan s-ből t-be vezető utat, amelyen a legnagyobb súly a lehető legkisebb.*

Feladat 1.3.17 *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.*

Feladat 1.3.18 *Igazoljuk, hogy ha egy s-ből t-be vezető P út minden éle benne van egy legolcsóbb útban, akkor P maga is legolcsóbb út.*

Feladat 1.3.19 *Igazoljuk, hogy egy éleken értelmezett költségfüggvény akkor és csak akkor pontindukált, ha minden* (irányítatlan értelemben vett) *körre az egyik irányba menő élek összköltsége egyenlő a másik irányba menő élek összköltségével.*

Feladat 1.3.20 *Igazoljuk, hogy a pontindukált költségek alteret alkotnak $\mathbf{R}^A$-ban. Határozzuk meg a tenziók ortogonális kiegészítő alterét, más szóval, jellemezzük azon $x \in \mathbf{R}^A$ vektorokat, melyeknek minden pontindukált költségfüggénnyel vett skalárszorzata nulla.*

Feladat 1.3.21 *Igazoljuk, hogy c egészértékű és pontindukált, akkor létezik olyan π egészértékű potenciál, amelyre $c = \Delta_\pi$.*

1.3.5 Speciális konzervatív költségfüggvények

Bemutatunk két speciális konzervatív költségfüggvényt, amelyek esetén a legolcsóbb utak megkeresésére hatékonyabb algoritmus adható. Az egyik esetben D aciklikus és c tetszőleges, míg a másikban D tetszőleges és c nem-negatív.

Legolcsóbb utak aciklikus digráfban

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.

TÉTEL 1.3.11 *Egy $D = (V, A)$ digráf akkor és csak akkor aciklikus, ha pontjainak létezik* **topologikus sorrend***je azaz egy olyan $v_1, v_2, \ldots, v_n$ sorrend, amelyben minden él korábbi pontból későbbibe vezet.*

Biz. Egy egyirányú kör pontjainak nem létezhet aciklikus sorrendje, hiszen minden pontjának pozitív a befoka. Emiatt egy egyirányú kört tartalmazó digráfnak se létezhet topológikus 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 $v_1, v_2, \ldots, v_n$ topológikus sorrendet. •

A bizonyításból adódó algoritmus egyetlen forráspontot $O(m)$ lépésszámban tud megtalálni, így a topológikus sorrend megkeresésének összlépésszáma $O(mn)$. Megjegyzendő azonban, hogy a mélységi keresés okos alkalmazásával a teljes topológikus sorrendet $O(m)$ lépésben meg lehet találni.

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

Tegyük most fel, hogy a D digráf aciklikus és az s pontból mindegyik másik pontba vezet út. Legyen a c súlyozás 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. Mindenesetre aciklikus digráf esetén tetszőleges c konzervatív.

Aciklikus digráfban a következő egyszerű direkt eljárással fel lehet építeni a legolcsóbb utak fenyőjét. Tegyük fel, hogy az $s = v_1, v_2, \ldots, v_n$ topológikus sorrend első $j-1$ pontja által feszített részgráfban már meghatároztuk a legolcsóbb utak egy F_{j-1} fenyőjét az $\mu_c(v_i)$ költségekkel egyetemben $(1 \leq i \leq j-1)$. 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

$$\mu_c(v_j) = \min\{\mu_c(v_i) + c(v_iv_j) : v_iv_j \in A\} \tag{1.14}$$

formula adja. Továbbá, ha v_iv_j jelöli azt az élt, amelyen a minimum felvétetik, akkor $F_j := F_{j-1} + v_iv_j$ a legolcsóbb utak fenyője az első j csúcson. Ezt a rekurziót $j = 1, \ldots, n$-re követve a végül kapott F_n feszítő fenyő egy legolcsóbb utak fenyője lesz.

Ütemezés: a PERT módszer

Alkalmazásokban előfordul, hogy aciklikus digráfban minimális helyett maximális súlyú utat kell keresni. A fenti eljárás értelemszerű módosítással ekkor is alkalmazható. Tegyük fel, hogy a topológikus sorrend első $j-1$ pontja által feszített részgráfban már meghatároztuk a legsúlyosabb utak egy F_{j-1} fenyőjét a legsúlyosabb sv_i-út $\ell'_c(v_i)$ súlyával egyetemben $(1 \leq i \leq j-1)$. A sorrendben következő v_j pontra legyen $\ell'_c(v_j) := \max\{\ell'_c(v_i) + c(v_iv_j) : v_iv_j \in A\}$ és legyen $F_j := F_{j-1} + v_iv_j$, ahol a v_iv_j egy olyan él, amelyen a maximum felvétetik.

Fogalmazzuk meg az 1.3.9 tétel ellenpárját erre az esetre.

TÉTEL 1.3.12 *Legyen $c : A \to \mathbf{R}$ a $D = (V, A)$ aciklikus irányított gráf élhalmazán egy tetszőleges súlyfüggvény, és tegyük fel, hogy az s pontból minden pontba vezet egyirányú út. Ekkor*

$$\min\{\pi(t) - \pi(s) : \pi(v) - \pi(u) \geq c(uv), uv \in A\} = \max\{\widetilde{c}(P) : P \text{ út } s\text{-ből } t\text{-be}\}. \bullet \tag{1.15}$$

Az is gyakori, hogy nem annyira az optimális útra vagyunk kiváncsiak, mint inkább az optimális π függvényre. Tekintsük a következő ütemezési feladatot. Egy projekt különféle részfeladatok elvégzéséből áll, melyeknek előre adott a végrehajtási ideje. Tudjuk továbbá, hogy technológiai előírások miatt bizonyos részfeladatok megelőznek másokat (házépítésnél 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. További elvárás az egyes részfeladatok kezdési időpontját úgy megadni, hogy minden feladat kezdésére az őt a részbenrendezésben megelőzők már elkészüljenek.

A megoldáshoz készítsünk el egy D irányított gráfot a következőképpen. Reprezentáljunk mindegyik z részfeladatot 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 pontot, amelybe minden v_x pontból vezessünk 0 súlyú élt. Feladatunk olyan $\pi(v)$ időpontok kijelölésével ekvivalens, amelyekre $\pi(s) = 0$, $\pi(t)$ minimális és minden uv élre a $\pi(v) - \pi(u)$ időpont különbség legalább akkora, mint az él súlya.

Az 1.3.12 tétel pontosan erre a kérdésre adott választ: *a projekt végrehajtásának minimális ideje egyenlő a legsúlyosabb st-út súlyával.* Egy ilyen utat szoktak néha kritikus útnak nevezni, és az előzőekben leírt minimális út problémának az itteni feladatra adaptált változatát kritikus út módszernek (angolul **PERT**: project evaluation and review technique).

Gyakorlat 1.3.22 *Adott egy út bizonyos részútjainak $\mathcal{F}$ rendszere. Válasszunk ki $\mathcal{F}$-nek diszjunkt tagjait úgy, hogy az összhosszuk maximális legyen.*

Feladat 1.3.23 *Dolgozzunk ki eljárást pontsúlyozott részbenrendezett halmaz maximális súlyú láncának megkeresésére.*

Feladat 1.3.24 *Adott két betűkből álló sorozat. Válasszunk ki algoritmikusan egy maximális sok tagból álló közös részsorozatot. A feladat még akkor is megoldható, ha az egyes betűknek különböző súlya van és a kiválasztott közös részsorozatban lévő betűk súlyösszegét akarjuk maximalizálni.*

Feladat 1.3.25 *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.*

Feladat 1.3.26 *Határozzuk meg algoritmikusan egy véges sorozat legtöbb tagból álló monoton növekvő részsorozatát!*

Feladat 1.3.27 *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?*

Nem-negatív költségek: 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 a legolcsóbb utaknak egy fenyőjét építjük fel élek egyenkénti hozzávételével. Különbség viszont, hogy a pontoknak a fenyőbe kerülési sorrendjét nem lehet előre megmondani, mert az csak menetközben derül ki, a következő lemma szerint.

Lemma 1.3.13 *Legyen T egy legolcsóbb utak s-fenyője a $V(T)$ ponthalmazon. Tegyük fel, hogy az*

$$m_T := \min\{\mu_c(u) + c(uv) : uv \text{ kilép } V(T)\text{-ből}\} \tag{1.16}$$

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 1.3.13 lemma kézenfekvő és a korábbinál hatékonyabb 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 1.3.13 lemma alapján ez az (1.16) 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 (1.16) 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 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 $O(n^2)$.

Érdekes alkalmazásként megmutatjuk, hogy a Dijkstra algoritmus hasznos lehet általános c konzervatív költségfüggvény esetén is. Tegyük ugyanis fel, hogy most egy egyszerű digráfban 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 $O(n^2 m)$. De ennél gazdaságosabban is eljárhatunk. A π_c függvényről megmutattuk, hogy $O(mn)$ lépésben kiszámítható. Mivel π_c c-megengedett potendciál, azaz $c' := c - \pi_c \geq 0$, 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 $O(n^2 m)$.

2011. október 29. file: linp, minutak

1.4 MINIMÁLIS KÖLTSÉGŰ FÁK

1.4.1 Fák

A gráfokon tekintett optimalizálási feladatok közül a legkorábbi egy $G = (V, E)$ összefüggő irányítatlan gráf minimális költségű feszítő fájának megkeresését célozza. Ez a mohó algoritmussal történik, ami valami olyasfélét akar kifejezni, hogy az algoritmus során mindig a lokálisan legjobbat választjuk. A fákra vonatkozó mohó algoritmusnak számos változata ismert: az alábbiakban ezek egységes leírását adjuk meg. Jelölje $c : E \to \mathbf{R}$ a költségfüggvényt. A következő lemma a fák egy fontos kicserélési tulajdonságát írja le.

Lemma 1.4.1 *Jelölje T_1 és T_2 két V-t feszítő fa élhalmazát. Ekkor bármely $e \in T_1$ élre van olyan $f \in T_2$ él, amelyre mind $T_1 - e + f$, mind $T_2 - f + e$ fa.*

Biz. Ha $e \in T_2$, akkor $f := e$ jó lesz. Tegyük fel, hogy $e = st \notin T_2$. $T_1 - e$-nek két komponense van, K_1 és K_2. T_2-ben van egy egyértelmű P út, amely összeköti az s és t pontokat. Legyen f a P útnak egy olyan éle, amely K_1 és K_2 között vezet. Ezen f él kielégíti a lemma kívánalmait. •

Legyen T a $G = (V, E)$ egy feszítő fája. Az $e \in F$ élhez tartozó **alapvágás**on a G azon vágását értjük, amelyet a $T - e$ két komponense határoz meg. Egy $f = uv \in E - T$ élhez tartozó **alapkör**ön azt a kört értjük, amely az f élből és az u és v pontokat a T fában összekötő egyértelmű útnak az éleiből áll. A 1.4.1 lemmából rögtön következik:

TÉTEL 1.4.2 *A G gráf valamely T feszítő fájára az alábbiak ekvivalensek:*
(a) T minimális költségű,
(b) $c(e) \leq c(f)$ fennáll minden $e \in T$ élre, ahol f az e alapvágásának egy eleme,
(c) $c(e) \geq c(f)$ fennáll minden $e \notin T$ élre, ahol f az e alapkörének egy eleme. •

MOHÓ ALGORITMUS [Boruvka, 1926], [Kruskal, 1956] *Az eljárás egy feszítő erdőt épít élek egyenkénti hozzávételével. Az V ponthalmazú élt nem tartalmazó erdővel indul, és akkor ér véget, amikor az aktuális feszítő erdő már fa. Az általános lépés abból áll, hogy az aktuális már megkonstruált erdőhöz hozzáadjuk a legolcsóbb olyan élt, amely két komponensét köti össze.*

Ismeretes a mohó algoritmusnak másik változata is.

DIJKSTRA-PRIM ALGORITMUSA [Dijkstra 1959], [Prim 1957] *Egy tetszőleges x_0 pontból indulva élek egyenkénti hozzávételével fát építünk, egészen addig, amíg feszítő fát nem kapunk. Az általános lépésben a legolcsóbb olyan éllel növelünk, amelynek pontosan az egyik végpontja tartozik a már megkonstruált fához.*

A következő algoritmus óvatosnak nevezhető; ahelyett, hogy olcsó élekből próbálna fát vagy erdőt építeni, megszabadul a drága élektől, persze ügyelve az összefüggőség megtartására.

FORDÍTOTT MOHÓ (ÓVATOS) ALGORITMUS *Az eljárás során éleket hagyunk ki a gráfból arra ügyelve, hogy a visszamaradó részgráf összefüggő legyen. Az általános lépésben kiválasztjuk a maximális költségű élt, amely nem elvágó él, és ezt elhagyjuk a gráfból.*

Ezen algoritmusok egyetlen közös általános keretbe foglalhatók.

Gyakorlat 1.4.1 *Mutassuk meg, hogy az előbbi algoritmusok mindegyike az alábbi speciális esetének tekinthető.*

ÁLTALÁNOS ALGORITMUS *Az eljárás az alábbi két művelet tetszőleges sorrendben történő egymás utáni alkalmazásából áll. Az első művelet a V csúcshalmazon egy F feszítő erdőt épít élek egyenkénti hozzávételével, míg a második bizonyos éleket kitöröl. Kezdetben $F := \emptyset$.*

1. LÉPÉS *Ha az aktuális F erdő már feszítő fa, az algoritmus befejeződik. Ha F nem összefüggő, akkor válasszunk G-nek egy tetszőleges olyan B vágását, amelyben nincs F-nek éle és legyen e a B vágás legolcsóbb eleme. Adjuk e-t F-hez.*

2. LÉPÉS *Amig van kör, válasszunk ki egy tetszőleges C kört. Legyen $e \in C$ a legdrágább éle $C - F$ -nek. Töröljük e-t G-ből.*

TÉTEL 1.4.3 *Az algoritmus által talált végső feszítő F fa minimális költségű.*

Biz. Az algoritmus futásának tetszőleges közbenső állapota egy (F, D) párral jellemezhető, ahol F az addig megkonstruált erdőt, D pedig az addig eltörölt élek halmazát jelöli. Azt igazoljuk indukcióval, hogy létezik olyan T minimális költségű feszítő fája G-nek, amelyre $F \subseteq T \subseteq E - D$. Világos, hogy bármelyik minimális költségű fa jó lesz, amikor $F = D = \emptyset$. Tegyük most fel, hogy az állítást már beláttuk valamely (F, D) párra, vagyis hogy van egy olyan T minimális költségű feszítő fa, amelyre $F \subseteq T \subseteq E - D$.

Tegyük fel először, hogy az 1. lépést alkalmaztuk, és legyen $e \in B$ az újonnan F-hez vett él. Legyen $F' := F + e$. Ha $e \in T$, akkor készen vagyunk, mert a változatlan T jó lesz az (F', D) párra nézve is. Ha $e \notin T$, akkor legyen C_e az e alapköre a T-re nézve. Az e él a szabály szerint a B vágásban van, így C_e-nek kell lennie egy másik f élének B-ben. Miután $e, f \in B$, az 1. lépés szabálya szerint $c(e) \leq c(f)$. Mivel $e, f \in C_e$ és T minimális költségű, azt kapjuk, hogy $c(f) \geq c(e)$. Ezekből $c(e) = c(f)$, és $T' := T - f + e$ is egy minimális költségű feszítő fa, amelyre $F' \subseteq T' \subseteq E - D$.

Másodszor, tegyük fel, hogy a 2. lépést alkalmazzuk, és legyen $e \in C$ a frissen eltörölt él. Ha T nem tartalmazza e-t, készen vagyunk, mert a változatlan T jó lesz az $(F, D + e)$ párra nézve is. Tegyük fel tehát, hogy T tartalmazza e-t, és legyen B_e az e-nek T-re vonatkozó alapvágása. Ekkor létezik egy $f \neq e$ él, amelyre $f \in C \cap B_e$. Mivel $e, f \in C$, a 2. lépés szabály szerint $c(e) \geq c(f)$. Mivel $e, f \in B_e$ és T minimális költségű, következik, hogy $c(e) \leq c(f)$. Ezekből $c(e) = c(f)$, és $T' := T - f + e$ is egy minimális költségű feszítő fa, amelyre $F \subseteq T' \subseteq E - (D + e)$. •

Feladat 1.4.2 *Ha minden költség különböző, akkor a min. költségű feszítő fa egyértelmű.*

Feladat 1.4.3 *Tegyük fel, hogy két költségfüggvény adott az éleken: c_1, c_2. Adjunk algoritmust olyan feszítő fa megkeresésére, amely a c_1-re nézve minimális költségű és ezen belül c_2-re nézve minimális költségű. Hogyan általánosítható az eljárás több költségfüggvényre.*

Feladat 1.4.4 *$G = (V, E)$ összefüggő gráf egy F feszitő fája akkor és csak akkor minimális költségű, ha a fa bármely e élének a költsége nem nagyobb, mint az $F - e$ két komponense között vezető bármely él költsége. Továbbá F akkor és csak akkor minimális költségű, ha bármely $uv \in E - E(F)$ él költsége legalább akkora, mint a fában az u és v között vezető egyértelmű út bármely élének a költsége.*

Feladat 1.4.5 *Legyen $r(u, v)$ szimmetrikus, nemnegatív egészértékű függvény a V alaphalmaz elempárjain. Akkor és csak akkor létezik olyan $G = (V, E)$ gráf, amelyre $\lambda(u, v; G) = r(u, v)$ minden $\{u, v\}$ pontpárra, ha $r(u_1, u_k) \geq \min\{r(u_1, u_2), r(u_2, u_3) \ldots, r(u_{k-1}, u_k)\}$ fennáll minden, a V különböző elemeiből készített $u_1, \ldots, u_k$ sorozatra.*

Útmutatás. Legyen F a maximális súlyú feszítő fa, és helyettesítsük mindegyik uv élét $r(u, v)$ párhuzamos éllel. Ez jó lesz G-nek.

Matroidok

Túl azon az örömteli megfigyelésen, hogy a mohó algoritmus helyesen működik maximális súlyú feszítő fa kiszámítására, joggal vetődik fel a kérdés, hogy vannak-e más szituációk is, amikor a Kruskal-típusú mohó algoritmus helyes eredményt szolgáltat. Igazolható például, hogy ha egy mátrix minden oszlopának adott a súlya, akkor maximális össz-súlyú lineárisan független oszlopot a mohó algoritmus segítségével kereshetünk (támaszkodva a Gauss eliminációra).

Mármost a lényeg az, hogy mind egy gráf részerdői, mind egy mátrix oszlopainak független részhalmazai matroidot alkotnak. Legyen S véges halmaz és $\mathcal{F}$ az S részhalmazainak egy rendszere, melynek tagjait függetlennek hívjuk. Az $(S, \mathcal{F})$ párt **matroid**nak nevezzük, ha *(a)* az üres halmaz független, *(b)* független halmaz részhalmaza független, és *(c)* minden $Z \subseteq S$ részhalmazra a Z-be eső, Z-ben már tovább nem bővíthető független halmazok elemszáma ugyanaz a csupán Z-től függő $r(Z)$ szám. A harmadik tulajdonság azzal ekvivalens, hogy a mohó algoritmus minden $0-1$-es súlyfüggvényre maximális súlyú független halmazt szolgáltat.

A matroidelméletben igazolják, hogy a mohó algoritmus tetszőleges súlyfüggvény esetén helyesen számolja ki a maximális súlyú független halmazt. Matroidok segítségével azonban sokkal bonyolultabb optimalizálási kérdések is megválaszolhatók. Például, mikor lehet és hogyan egy gráfban k élidegen feszítő fát találni? Amennyiben létezik k élidegen feszítő fa, hogyan lehet úgy meghatározni őket, hogy az össz-költségük minimális legyen? Vagy, hogyan lehet egy irányított gráfnak egy olyan minimális költségű részgráfját meghatározni, amelyben egy megadott gyökérpontból minden más csúcsba vezet k élidegen út?

1.5 PÁROS GRÁFOK OPTIMÁLIS PÁROSÍTÁSAI

Egy osztályba 25 gyerek jár. Egy kiránduláson készült fényképek közül 25-t 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.

Gyakorlat 1.5.1 *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.*

1.5.1 Maximális elemszámú párosítások: a javító utak módszere

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

TÉTEL 1.5.1 (Kőnig) *Egy $G = (S, T; E)$ páros gráfban a diszjunkt élek maximális $\nu = \nu(G)$ száma egyenlő az éleket lefogó pontok minimális $\tau = \tau(G)$ elemszámával.*

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).

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

$$h(X) := |X| - |\Gamma(X)| \tag{1.17}$$

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

$$\mu := \max_{X \subseteq S} h(X) \tag{1.18}$$

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.

Lemma 1.5.2 *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.

TÉTEL 1.5.3 (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:*

$$|\Gamma(X)| \geq |X| \text{ minden } X \subseteq S \text{ részhalmazra.} \bullet \tag{1.19}$$

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

$$H := Z \cap S \tag{1.20}$$

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.

Feladat 1.5.2 *Igazoljuk, hogy a Kőnig algoritmus által kiadott* (1.20)*-beli H max-hiányú halmaz független az algoritmus futásától.*

1.5.2 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 (tanusítványt) 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 $\pi : V \to \mathbf{R}$ 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.

TÉTEL 1.5.4 (Egerváry) *A $G = (S, T; E)$ teljes párosítással rendelkező páros gráfban a $c : E \to \mathbf{R}_+$ 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 megkaptunk 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

$$\delta := \min\{\pi(u) + \pi(v) - c(uv) : uv \in E, u \in H, v \in T - \Gamma_{G_\pi}(H)\} \tag{1.21}$$

érték pozitív. Módosítsuk π-t a következőképp:

$$\pi'(v) = \begin{cases} \pi(v) - \delta, & \text{ha } v \in H, \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_{G_\pi}(H), \\ \pi(v) & \text{különben.} \end{cases} \tag{1.22}$$

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 $O(|E||S|^2)$.

Gyakorlat 1.5.3 *Igazoljuk, hogy a minimális összértékű súlyozott lefogások konvex halmazt alkotnak.*

Feladat 1.5.4 *Igazoljuk, hogy ha egy M teljes párosítás minden éle eleme valamely maximális súlyú teljes párosításnak, akkor M maga is maximális súlyú teljes párosítás.*

Feladat 1.5.5 *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

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

és

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

Valójában egy olyan π-t adunk meg, amelyre

$$\text{minden } uv \in M \text{ élre } \pi(u) + \pi(v) \leq c(uv) \tag{1.25}$$

és

$$\text{minden } uv \in E - M \text{ élre } \pi(u) + \pi(v) \geq c(uv) \tag{1.26}$$

Ennek ugyanis alkalmas növelésével az (1.23)-t és (1.24)-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, 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 $\pi'(y) - \pi'(x) \leq c(st)$ 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 (1.25) és (1.26) 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.

1.5.3 Egerváry eredeti bizonyítása és algoritmusa

Egerváry az 1.5.4 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.

$$\delta := \min\{\pi(u) + \pi(v) - c(uv) : uv \in E, u \in X, v \in T - \Gamma_{G_\pi}(X)\}. \tag{1.27}$$

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

$$\pi'(v) = \begin{cases} \pi(v) - \delta, & \text{ha } v \in X, \\ \pi(v) + \delta, & \text{ha } v \in \Gamma_{G_\pi}(X), \\ \pi(v) & \text{különben.} \end{cases} \tag{1.28}$$

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

Lemma 1.5.5 *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 $\Gamma(X) \cap \Gamma(Y) \supseteq \Gamma(X \cup Y)$. $\bullet$

Lemma 1.5.6 *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 $h(X \cap Y) = \mu, h(X \cup Y) = \mu$. $\bullet$

Az 1.5.6 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.

TÉTEL 1.5.7 *Kőnig alternáló utas algoritmusa által szolgáltatott* (1.20)*-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. $\bullet$

Feladat 1.5.6 *Hogyan lehet az alternáló utas algoritmus segítségével az egyértelmű legbővebb max-hiányú halmazt megkonstruálni?*

Lemma 1.5.8 *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 1.5.6 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. $\bullet$

TÉTEL 1.5.9 *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 a (1.27) definícióval megadott minimum felvétetik. Az 1.5.8 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.

1.5.4 Maximális súlyú párosítások

Megmutatjuk, hogy a maximális súlyú (nem feltétlenül teljes) párosítás meghatározásának problémája egyszerű fogással visszavezethető a maximális súlyú teljes párosításéra.

TÉTEL 1.5.10 *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 1.5.4 tétel (második 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 $\widetilde{\pi}'(V') = \widetilde{c}(M')$. •

Végül megjegyezzük, hogy tetszőleges rögzített pozitív k egészre 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.

Feladat 1.5.7 *Tegyük fel, hogy a $G = (S, T; E)$ páros gráfban M egy j élű párosítás, amely a j élű párosítások között maximális súlyú a $c : E \to \mathbf{R}$ 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.*

1.6 ÁRAMOK ÉS FOLYAMOK HÁLÓZATOKBAN

1.6.1 Fogalmak

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

Áramok

Jelöljön $D = (V, A)$ egy irányított gráfot. Valamely $x : A \to \mathbf{R}$ 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 $c : A \to \mathbf{R}$ költségfüggvényre vonatkozólag a cx skalárszorzatot nevezzük az x áram **költség**ének.

Gyakorlat 1.6.1 (a) *Igazoljuk, hogy x 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.*

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

$$f \leq x \leq g. \tag{1.29}$$

(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 nem-negatív $x : A \to \mathbf{R}_+$ 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

$$\delta_x(s) = \delta_x(s) - \varrho_x(s) = \sum[\delta_x(v) - \varrho_x(v) : v \in S] = \delta_x(S) - \varrho_x(S). \tag{1.30}$$

Vagyis minden S $s\bar{t}$-halmazra a $\delta_x(S) - \varrho_x(S)$ értékkel definiált **tiszta kiáramlás** független 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-fonatnak nevezünk. Az x folyamot **út-folyam**nak (**kör-folyam**nak) nevezzük, ha x csak egy s-ből t-be vezető út (irányított 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.

Gyakorlat 1.6.2 *Igazoljuk, hogy minden nem-negatív folyam előállítható kör- folyamok és út-folyamok nem-negatív lineáris kombinációjaként. Speciálisan, minden fonat élidegen körök és st-utak uniója.*

1.6.2 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 termelő kapacitása (supply). Adott továbbá *ell* fogyasztóhely, melyeknek ismerjük az igényeit (demand). Tudjuk, hogy mely üzemekből mely fogyasztókhoz milyen 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 kapacitás és igény azonosan egy, 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 $g : E \to \mathbf{R}_+$ kapacitásfüggvény valamint egy $c : E \to \mathbf{R}_+$ költségfüggvény. Az első feladat annak eldöntése, hogy létezik-e olyan $x : E \to \mathbf{R}_+$ függvény, amelyre $0 \le x \le g$, $d_x(v) \le q(v)$ 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 $M := \sum_{v \in T} q(v)$ 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 élidegen (vagy pontidegen) út. 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 $S \subseteq V - t$ 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 $D' = (V, A')$ digráfban

(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 s-ből t-be (ún. 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 sokkal 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 átjutni, és valójában egy-egy ilyen átmenetek 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.

1.6.3 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.

TÉTEL 1.6.1 (Hoffman, 1960) *A $D = (V, A)$ digráfban adott $f \leq g$ kapacitásfüggvényekre vonatkozólag akkor és csak akkor létezik megengedett áram, ha*

$$\varrho_f(X) \leq \delta_g(X) \text{ minden } X \subseteq V \text{ halmazra.} \tag{1.31}$$

Továbbá, ha f és g egészértékűek és (1.31) *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 (1.31) következik.

Az elegendőség igazolásához tekintsük a következő függvényt:

$$\beta(X) := \delta_g(X) - \varrho_f(X). \tag{1.32}$$

Most (1.31) 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.

Lemma 1.6.2 $\beta(X) + \beta(Y) = \beta(X \cap Y) + \beta(X \cup Y) + d_{g-f}(X, Y)$.

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 (1.31) feltétel miatt, megengedett áram volna, hiszen az 1.6.1 gyakorlat 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 (1.31)-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 $m : V \to \mathbf{R}$ előre adott függvény. Vagy még általánosabban, adott $p : V \to \mathbf{R} \cup \{-\infty\}$, $b : V \to \mathbf{R} \cup \{\infty\}$ esetén ($p \leq b$) minden v csúcsra legyen

$$p(v) \leq \varrho_x(v) - \delta_x(v) \leq b(v). \tag{1.33}$$

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 $x : A \to \mathbf{R}$ függvényhez legyen $x' : A' \to \mathbf{R}$ a következőképpen definiálva: $x'(e) := x(e)$ ha $e \in A$, és $x'(e) := \varrho_x(v) - \delta_x(v)$ ha $e' = vs$ $(v \in V)$. Hasonlóképp terjesszük ki f-et és g-t az új élekre: $f(us) := p(u), g(us) := b(u)$. Könnyen látszik, hogy x akkor és csak akkor teljesíti (1.29)-t és (1.31)-t, ha x' megengedett áram.

Feladat 1.6.3 *Hoffman tételének általánosításaként igazoljuk, hogy akkor és csak akkor létezik az* (1.29)*-t és* (1.33)*-t kielégítő x függvény, ha $\varrho_f(X) - \delta_g(X) \leq \min\{p(X), b(V - X)\}$ fennáll minden $X \subseteq V$-re.*

Feladat 1.6.4 *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.*

1.6.4 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).

TÉTEL 1.6.3 (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 $\delta_g(S) \geq k$. Ha e feltétel teljesül, g egészértékű és k egész, úgy a folyam is választható egészértékűnek.*

A tételt néha az alábbi ekvivalens alakban említik.

TÉTEL 1.6.4 *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 (1.30) folytán $\delta_x(s) = \delta_x(S) - \varrho_x(S) \leq \delta_g(S)$, amiből $max \leq min$ 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 (1.31) fennáll, és így az 1.7.4 tétel szerint létezik megengedett áram (amely ráadásul egészértékű, ha g az). Kihagyva a hozzávett e^* élt, μ_g nagyságú folyamot kapunk. •

Feladat 1.6.5 *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.)

Gyakorlat 1.6.6 *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 (1.31)-t megsértő halmazai között.

Lemma 1.6.5 **(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* (1.31) *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 $\delta_g(X) < \varrho_f(X)$. •

Feladat 1.6.7 *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 (1.31), 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.

2011. október 29. file: lin23

1.7 MAXIMÁLIS 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 1.7.6 tételben a feltétel szükségessége illetve az 1.6.4 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.

1.7.1 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.)

1.7.2 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.

1.7.3 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.)

TÉTEL 1.7.1 *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, $\sigma_x(v) = \sigma_x(u) + 1$.

Lemma 1.7.2 *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.

Lemma 1.7.3 *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!

Feladat 1.7.1 (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.*

Feladat 1.7.2 *Adott e élre hogyan lehet eldönteni, hogy* (a) *létezik-e olyan maximális folyam, amely telíti e-t,* (b) *minden maximális folyam telíti e-t.*

Feladat 1.7.3 *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.*

Feladat 1.7.4 *Adott két kapacitásfüggvény esetén algoritmikusan döntsük el, hogy létezik-e olyan S $s\bar{t}$-halmaz, amely miondkét kapacitásfüggvényre nézve minimális vágást határoz meg.*

Feladat 1.7.5 *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.*

Feladat 1.7.6 *Készítsünk algoritmust, amely adott költségfüggvényre eldönti, hogy létezik-e k élidegen út s-ből t-be, melyek mindegyike minimális költségű.*

Feladat 1.7.7 *Digráfban keressünk két diszjunkt halmazt, melyek egyike $s\bar{t}$-halmaz, másika $t\bar{s}$-halmaz úgy, hogy befok összegük minimális.*

1.7.4 A szintező algoritmus megengedett m-áramok kiszámítására

A Ford és Fulkerson által bevezetett növelő utas módszer, illetve annak finomított változata, az Edmonds–Karp – Dinits algoritmus segítségével erősen polinomiális időben, nevezetesen $O(nm^2)$ lépésben meg tudtunk határozni egy s-ből t-be menő maximális nagyságú folyamot és egy minimális vágást.

Az alábbiakban bemutatunk egy ettől gyökeresen különböző eljárást, az úgynevezett **szintező algoritmus**t, amely minden szempontból felülmúlja a növelő utas módszert. (Az angol nyelvű szakirodalomban az ilyen típusú eljárásokat *push-relabel* algoritmusnak hívják, mi a szintező eljárás nevet használjuk). A Goldbergtől és Tarjantól származó eljárás elvileg is és gyakorlailag is hatékonyabb a növelő utas algoritmusnál. Nem használ növelő utakat, nem használ segédgráfot, sőt még folyamokat sem! Egyetlen lépése csak kicsiny, lokális változtatásból áll (szemben a növelő utas eljárásnak egy egész út mentén történő változtatásával), és helyességének illetve a lépésszámára adott korlátnak bizonyítása is egyszerű.

Legyen $D = (V, A)$ digráf élhalmazán adott az $f : A \to \mathbf{R}_+ \cup \{-\infty\}$ és $g : A \to \mathbf{R}_+ \cup \{\infty\}$ függvény, melyekre $f \leq g$. Egy $x : A \to \mathbf{R}$ függvény (vagy vektor) **megengedett**, ha $f \leq x \leq g$. Legyen $\varrho_x(Z) := \sum[x(e) : e \in A$ belép Zbe], $\delta_x(Z) := \varrho_x(V - Z)$ és $\Psi_x(Z) := \varrho_x(Z) - \delta_x(Z)$ $(Z \subseteq V)$. Könnyen belátható, hogy a Ψ_x függvény moduláris abban az értelemben, hogy

$$\Psi_x(Z) = \sum[\Psi_x(v) : v \in Z]. \tag{1.34}$$

Adott $m : V \to \mathbf{R}$ függvény esetén azt mondjuk, hogy $x : A \to \mathbf{R}$ **moduláris áram**, röviden m-**áram**, ha

$$\Psi_x(v) = m(v) \text{ minden } v \in V \text{ csúcsra.} \tag{1.35}$$

Ha $m \equiv 0$, visszajutunk a már ismert áram fogalomhoz. Egy másik speciális esetben $f \equiv 0 \leq g$ és m-et olyképp definiáljuk, hogy $m(t) = k$, $m(s) = -k$ két kijelölt s és t pontra, míg $m(v) = 0$ minden más pontra. Ekkor egy megengedett m-áram nem más, mint egy k nagyságú folyam s-ből t-be.

Gyakorlat 1.7.8 *Tegyük fel, hogy $\widetilde{m}(V) = 0$. Igazoljuk, hogy x akkor és csak akkor m-áram, ha $\varrho_x(v) - \delta_x(v) \leq m(v)$ minden v csúcsra. Ha x m-áram, akkor $\varrho_x(Z) - \delta_x(Z) = \widetilde{m}(Z)$ minden $Z \subseteq V$ halmazra.*

TÉTEL 1.7.4 (Hoffman, 1960) *Akkor és csak akkor létezik megengedett m-áram, ha $\widetilde{m}(V) = 0$ és*

$$\varrho_f(X) - \delta_g(X) \leq \widetilde{m}(X) \text{ minden } X \subseteq V\text{-re.} \tag{1.36}$$

Ha f, g és m egészértékű és (1.36) *fennáll, akkor létezik egészértékű megengedett m-áram is.*

Feladat 1.7.9 *Vezessük le a $D = (V, A)$ digráfra vonatkozó* 1.7.4 *tételt azon eggyel több pontú digráfra vonatkozó speciális alakjából, amelyben $m \equiv 0$.*

Hoffman tétele speciális esetben kiadja a következőt.

TÉTEL 1.7.5 *Akkor és csak akkor létezik k nagyságú megengedett folyam s-ből t-be, ha $\delta_g(S) \geq k$ fennáll minden S $s\bar{t}$-halmazra. Ha g egészértékű, a folyam is választható egészértékűnek.*

A tétel ekvivalens alakban is megfogalmazható.

TÉTEL 1.7.6 (Max-flow Min-cut) *Adott g kapacitás függvény és $D = (V, A)$ digráf esetén, 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, a maximális folyam is választható egészértékűnek.*

Biz. Alkalmazzuk az előző tételt a szóbanforgó minimum k értékére. •

Egyszerűség kedvéért feltesszük, hogy nincsenek párhuzamos élek. Az algoritmus fenntart egy megengedett $x : A \to \mathbf{R}$ vektort és igyekszik a (1.35) követelményt elérni. Egy $v \in V$ pontra akkor mondjuk, hogy **pozitív**, **negatív** vagy **semleges**, ha $\Psi_x(v) - m(v)$ pozitív, negatív vagy nulla. Egy él e is **csökkenthető**, ha $x(e) > f(e)$ és **növelhető**, ha $x(e) < g(e)$.

Szint tulajdonságok és megállási szabályok

A megengedett x-en kívül az algoritmus fenntart egy $\Theta : V \to \{0, 1, \ldots, n\}$ szintfüggvényt, ahol $\Theta(v)$ a v csúcs szintje. (Szokás szerint n a V csúcshalmaz elemszámát jelöli.) Adott $j \in \{0, 1, \ldots, n\}$ szintre az $L_j := \{v \in V : \Theta(v) = j\}$ halmazt **szinthalmaz**nak hívjuk. Tekintsük a következő **szint tulajdonságok**at.

(LP1) Minden negatív csúcs az L_0 szinthalmazban van.
(LP2′) $\Theta(v) \geq \Theta(u) - 1$ minden növelhető uv élre, azaz, minden növelhető él legfeljebb egy szintet lép le.
(LP2″) $\Theta(v) \leq \Theta(u) + 1$ minden csökkenthető uv élre, azaz, minden a csökkenthető él legfeljebb egy szintet lép fel.

Az algoritmus futása akkor fejeződik be, ha a következő két **megállási szabály** egyike bekövetkezik.

(A) Nincs pozitív csúcs.
(B) Létezik egy z pozitív csúcs és a szintje alatt egy üres L_j szinthalmaz (azaz, $j < \Theta(z)$).

Lemma 1.7.7 *Tegyük fel, hogy x és Θ teljesítik a szint tulajdonságokat. Ekkor* **(A)** *esetén x megengedett m-áram, míg* **(B)** *esetén a $Z := \{v \in V : \Theta(v) \geq j\}$ halmaz megsérti a* (1.36) *feltételt.*

Biz. Ha nincs pozitív csúcs, akkor $\widetilde{m}(V) = 0 = \Psi_x(V)$ miatt negatív sincs, és így x megengedett m-áram.

Tegyük fel, hogy **(B)** teljesül. Miután (LP1) miatt minden negatív csúcs L_0-ban van Z nem tartalmaz negatív csúcsot. Ugyanakkor Z tartalmazza a pozitív z-t és ezért $\Psi_x(Z) = \sum[\Psi_x(v) : v \in Z] > \widetilde{m}(Z)$. Másrészről, az L_j szinthalmaz üressége miatt minden Z-ből kilépő e él legalább két szintet lép le és ezért (LP2′) folytán $x(e) = g(e)$, vagyis $\delta_x(Z) = \delta_g(Z)$. Hasonlóképp, minden Z-be lépő e él legalább két szintet lép fel és így (LP2″) miatt $x(e) = f(e)$ vagyis $\varrho_x(Z) = \varrho_f(Z)$. A kettőből $\varrho_f(Z) - \delta_g(Z) = \varrho_x(Z) - \delta_x(Z) = \Psi_x(Z) > \widetilde{m}(Z)$ adódik, mutatva, hogy Z megsérti (1.36)-t. •

Alapműveletek egy pozitív z csúcsnál

Az algoritmus egy közbenső, általános helyzetében adott egy $x : A \to \mathbf{R}$ megengedett vektor és egy Θ szintfüggvény, melyek teljesítik a szint tulajdonságokat. Tegyük fel, hogy egyik megállási szabály sem áll fenn.

Egy z pozitív csúcsra bizonyosan $\Theta(z) < n$, mert $\Theta(z) = n$ esetén létezne z alatt üres szinthalmaz, azaz **(B)** fennállna. Két alapműveletet alkalmazhatunk z-nél: élérték csere (push) és csúcs-emelés (relabel).

Élérték csere z-nél módosítja $x(e)$-t egy z-ben kezdődő vagy végződő e élen, a következőképp.

(Növelés) Ha $e = zu$ növelhető él, amely lelép z-ből, akkor növeljük $x(e)$-t az $\alpha := \min\{g(e) - x(e), \Psi_x(z) - m(z)\}$ értékkel.
(Csökkentés) Ha $e = uz$ csökkenthető él, amely fellép z-be, akkor csökkentsük $x(e)$-t az $\alpha := \min\{x(e) - f(e), \Psi_x(z) - m(z)\}$ értékkel.

Csúcs-emelés z-nél Amennyiben élérték csere nem alkalmazható z-nél, emeljük meg z szintjét eggyel, azaz növeljük eggyel a $\Theta(z)$ értéket.

Lemma 1.7.8 *Az alapműveletek megőrzik a megengedettséget és a szint tulajdonságokat.*

Biz. Az α értelmezése miatt egy élérték csere fenntartja a megengedettséget. Mivel nem hogy létre új negatív csúcsot, az (LP1) tulajdonság is fennmarad. Miután egy élérték csere csak olyan uv élen okoz változást, amelyre $|\Theta(u) - \Theta(v)| = 1$, az (LP2) tulajdonság sem tud elromlani.

A z csúcs-emelés nem érinti (LP1)-et, hiszen $\Theta(z)$-t csak akkor növeljük, ha z pozitív csúcs. Megőrzi (LP2)-t is, mert csak akkor alkalmazhatjuk, ha már nincsen z-ből lelépő növelhető, vagy z-be fellépő csökkenthető él. •

Az eljárás és lépésszám becslés

Az algoritmus tetszőleges $x : A \to \mathbf{R}$ vektorral indul, amelyre $f \leq x \leq g$, továbbá Θ kezdetben azonosan nulla. Egy közbenső helyzetben adott egy megengedett x és egy Θ szintfüggvény, melyekre fennállnak a szint tulajdonságok. Ha nincsen pozitív csúcs, akkor az adott x megengedett m-áram és az algoritmus futása befejeződik.

Ameddig van pozitív csúcs, az algoritmus kiválaszt közülük egy legmagasabb szinten lévőt és a következőképpen kezeli. Amíg csak z pozitív marad és van belőle lelépő növelhető vagy z-be fellépő csökenthető él, alkalmazzuk az élérték csere műveletet z-nél. Ha ezek után z még mindig pozitív marad, alkalmazzuk a csúcsemelést z-nél. Ennek megfelelően, a z kezelésének a végére z vagy semlegessé vált vagy a szintjét megemeltük.

Az egyik lehetséges leállási mód az, amikor egy élérték csere nyomán **(A)** igazzá válik: a kapott x megengedett m-áram az 1.7.4 Lemma folytán. A másik lehetséges leállási mód az, amikor egy csúcs-emelés nyomán **(B)** igazzá válik, vagyis a z emelésekor a z (emelés előtti) szinthalmaza üressé válik. Ekkor a $Z = \{v : \Theta(v) \geq \Theta(z)\}$ halmaz megsérti (1.36)-t az 1.7.4 Lemma miatt.

Hoffman 1.7.4 tételének nemtriviális iránya rögtön következik, amint belátjuk, hogy a megállási szabályok egyike véges sok lépés után bizonyosan bekövetkezik. Valójában érvényes a következő élesebb becslés.

Lemma 1.7.9 *Legfeljebb n^3 alapművelet után* **(A)** *vagy* **(B)** *bekövetkezik.*

Biz. Azt mondjuk, hogy egy z-nél lévő e élen végrehajtott élérték csere **semlegesítő**, ha z-t semlegessé teszi (amely akkor fordul elő, ha $\alpha = \Psi_x(z) - m(z)$). Az algoritmus egy **fázis**án a futásnak azon szakaszát értjük, melynek során a Θ függvény változatlan. Miután egy csúcsnál legfeljebb n szintemelés lehet, a fázisok teljes száma legfeljebb n^2.

Egy z' csúcsnál végrehajtott élérték csere csak egy z' alatti csúcsot alakíthat pozitívvá, a legmagasabb szint választási szabály miatt, ha a z csúcs semlegessé válik, ugyanabban a fázisban már nem lesz újra pozitív. Emiatt egy fázison belük legfeljebb n semlegesítő élérték csere fordulhat elő, és így a semlegesítő élérték cserék teljes száma legfeljebb n^3.

Egy nem-semlegesítő élérték csere az e élen vagy felnöveli $x(e)$-et $g(e)$-re, amivel nem-növelhetővé teszi e-t, vagypedig lecsökkenti $x(e)$-t $f(e)$-re amivel nem-csökkenthetővé teszi e-t. Emiatt egy nem-semlegesítő élérték csere műveletet követően csak akkor kerülhet újra sor az e élen élérték cserére, ha a $\Theta(z) - \Theta(u)$ előjele megfordul. Ekkorra viszont a $\Theta(z) + \Theta(u)$ összeg legalább kettővel megnőtt. Emiatt a nem-semlegesítő élérték cserék száma egy e élen legfeljebb n, és így a nem-semlegesítő élérték cserék teljes száma legfeljebb $|A|n \leq n^3$. •

Összefoglalva, az algoritmus futása legfeljebb $O(n^3)$ élérték csere és csúcs-emelés után befejeződik. Megjegyzendő, hogy alkalmas adatstruktúra használatával biztosítható, hogy az alapműveletek konstans időben végrehajthatók és ezért a szintező algoritmus teljes bonyolultsága $O(n^3)$.

Feladat 1.7.10 *Dolgozzunk ki egy olyan szintező eljárást teljes párosítás megkeresésére páros gráfban, amely csak az egyik pont osztály pontjaihoz rendel szinteket.*

A leginkább sértő halmaz kiszámítása

Amnnyiben nem létezik megengedett m-áram, a fenti algoritmus megtalált egy (1.36)-t megsértő Z halmazt. Az eljárás minimális módosításával egy legjobban sértő Z halmaz is megkereshető vagyis olyan, amelyre $\varrho_f(X) - \delta_g(X) - \widetilde{m}(X)$ a lehető legnagyobb.

A módosítás abból áll, hogy a futás nem ér véget, amikor egy szintemelés nyomán egy szint kiürül. Ennek következtében lehetnek csúcsok a legfelső L_n szinten és emiatt a z választását úgy módosítjuk, hogy mindig az n-dik szint alatti legfelső pozitív csúcsot választjuk. Az algoritmus akkor ér véget, ha már nincs pozitív csúcs az n-edik szint alatt. Ha egyáltalán nincs pozitív csúcs, akkor az aktuális x megengedett m-áram és ilyenkor nincsen hiányos halmaz.

TÉTEL 1.7.10 *Ha a módosított algoritmus futásának befejezésekor van pozitív csúcs, akkor egy üres szint feletti pontok Z halmaza a legjobban sérti* (1.36)*-t, azaz*

$$\varrho_f(Z) - \delta_g(Z) - \widetilde{m}(Z) \geq \varrho_f(X) - \delta_g(X) - \widetilde{m}(X) \text{ minden } X \subseteq V\text{-re.} \tag{1.37}$$

Biz. Mivel Z minden pozitív csúcsot tartalmaz, de negatívat nem, $\Psi_x(Z) - \widetilde{m}(Z) = \sum[\Psi_x(v) - m(v) : v \in Z] \geq \sum[\Psi_x(v) - m(v) : v \in X] = \Psi_x(X) - \widetilde{m}(X)$, amiből $\varrho_f(Z) - \delta_g(Z) - \widetilde{m}(Z) = \varrho_x(Z) - \delta_x(Z) - \widetilde{m}(Z) = \Psi_x(Z) - \widetilde{m}(Z) \geq \Psi_x(X) - \widetilde{m}(X) \geq \varrho_f(X) - \delta_g(X) - \widetilde{m}(X)$. •

Maximális nagyságú st-folyam kiszámítása

Amint már említettük, Hoffman tétele az $f \equiv 0 \leq g$, $m(t) = k, m(s) = -k$, és $m(v) = 0 \quad (v \in V - \{s,t\})$ választással az 1.7.5 tételre specializálódik. Ebben az esetben egy m-áram épp egy megengedett k nagyságú folyam. Továbbá, ha az S megsérti (1.36)-t, azaz $-\delta_g(S) = \varrho_f(S) - \delta_g(S) > \widetilde{m}(S)$, akkor $0 \leq \delta_g(S) < \widetilde{m}(S)$ és ezért $s \in S \subseteq V - t$ és $\widetilde{m}(S) = k$.

Amikor a k folyamnagyság elő van írva, a fenti (módosított) algoritmus vagy megad egy k nagyságú st-folyamot vagy pedig egy olyan S $s\bar{t}$-halmazt, melyre $\delta_g(S) < k$. Ha k nincs előre megadva, hanem maximális nagyságú folyamot keresünk, akkor alkalmazzuk első menetben a fenti algoritmust a $k := \delta_g(s)$ értékre. Amennyiben létezik k nagyságú st-folyam, akkor ez bizonyosan maximális nagyságú, hiszen $S := \{s\}$ egy minimális vágást definiál. Ha nem létezik k nagyságú st-folyam, akkor a módosított algoritmus megtalál egy legjobban sértő S $s\bar{t}$-halmazt, vagyis egy olyant, amelyre $\delta_g(S)$ minimális. Legyen $k' := \delta_g(S)$. Az MFMC tétel miatt biztosan létezik k' nagyságú st-folyam és ezért az algoritmust k'-re újrafuttatva ezt kiszámíthatjuk.

Megjegyzendő, hogy Goldberg és Tarjan eredeti algoritmusa egyetlen menetben számolja ki a maximális st-folyamot (és nem kettőben), viszont meg kell engednie, hogy a csúcsok az n-dik szint fölé is kerülhessenek. Emiatt külön igazolni kell, hogy a csúcsok nem tudnak a $2n-1$-dik szint fölé menni.

Egy speciális eset

Ha valaki a fenti eljárást vagy a bizonyítást kissé bonyodalmasnak érzi, érdemes a következő speciális esetet külön átgondolnia.

Tegyük fel, hogy $f \equiv 0, g \equiv \infty$. Ekkor az x megengedettsége egyszerűen azt jelenti, hogy x nemnegatív. Az $x(e)$ mindig növelhető, és akkor csökkenthető, ha pozitív. Hoffman tétele ekképp egyszerűsödik.

TÉTEL 1.7.11 *Akkor és csak akkor létezik nemnegatív m-áram, ha $\widetilde{m}(V) = 0$ és*

$$0 \leq \widetilde{m}(X) \text{ minden olyan } X \subseteq V\text{-re, amelyből nem lép ki él.} \tag{1.38}$$

Ha m egészértékű és (1.38) *fennáll, akkor létezik egészértékű nemnegatív m-áram is.*

A szint tulajdonságok az alábbira egyszerűsödnek.

(LP1) Minden negatív csúcs az L_0 szinthalmazban van.
(LP2′) Minden él legfeljebb egy szintet lép le.
(LP2″) Minden pozitív értékű él legfeljebb egy szintet lép fel.

A lemma **(B)** részének bizonyítása is kicsit egyszerűbbé válik: Miután (LP1) miatt minden negatív csúcs L_0-ban van Z nem tartalmaz negatív csúcsot. Ugyanakkor Z tartalmazza a pozitív z-t és ezért $\Psi_x(Z) = \sum[\Psi_x(v) : v \in Z] > \widetilde{m}(Z)$. Másrészről, az L_j szinthalmaz üressége miatt minden Z-ből kilépő e él legalább két szintet lép le és ezért (LP2′) folytán ilyen él nem létezhet. Továbbá minden Z-be lépő e él legalább két szintet lép fel és így (LP2″) miatt $x(e) = 0$) vagyis $\varrho_x(Z) = 0$. A kettőből $0 = \varrho_x(Z) - \delta_x(Z) = \Psi_x(Z) > \widetilde{m}(Z)$ adódik, mutatva, hogy Z megsérti (1.36)-t. •

Az élérték csere művelet is egyszerűsödik.

(Növelés) Ha $e = zu$ él lelép z-ből, akkor növeljük $x(e)$-t az $\Psi_x(z) - m(z)$ értékkel. (Ezáltal z semlegessé válik).
(Csökkentés) Ha $e = uz$ csökkenthető él, amely fellép z-be, akkor csökkentsük $x(e)$-t az $\alpha := \min\{x(e), \Psi_x(z) - m(z)\}$ értékkel.

1.7.5 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 $c : A \to \mathbf{R}_+$ 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. A rövidség kedvéért ezt k-**fonat**nak fogjuk nevezni. Az elnevezés arra utal, hogy azon e élek halmaza, melyeken $x(e) = 1$ felbomlik k élidegen st-útra és körökre.

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 **költségét** a $cz = \sum[c(e)z(e) : e \in A]$ skaláris szorzattal definiáljuk. (Mivel c nemnegatív, 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 $\pi : V \to \mathbf{Z}_+$ olyan függvény a csúcshalmazon, amelyre $\pi(s) = 0 \le \pi(v) \le \pi(t)$ minden $v \in V$-re. Egy ilyen függvényt **potenciál**nak hívunk. Definiáljuk a $\Delta_\pi : A \to \mathbf{R}$ költségfüggvényt a következőképpen:

$$\Delta_\pi(uv) := \pi(v) - \pi(u), \tag{1.39}$$

és nevezzük Δ_π-t **pontindukált** költségfüggvénynek. 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

$$c_\pi(uv) := c(uv) - \Delta_\pi(uv) \tag{1.40}$$

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

TÉTEL 1.7.12 (Ford és Fulkerson) *A $D = (V, A)$ irányított gráf élhalmazán adott a $c : A \to \mathbf{R}_+$ költségfüggvény. A k-fonatok minimális költsége* (vagyis k élidegen st-út összköltségének a minimuma) *egyenlő a*

$$k\pi(t) + \sum[c_\pi(uv) : uv \in A, c_\pi(uv) < 0] \tag{1.41}$$

érték maximumával, ahol a maximum az összes $\pi : V \to \mathbf{R}_+$ 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:

$$c_\pi(uv) > 0 \Rightarrow z(uv) = 0, \tag{i}$$
$$c_\pi(uv) < 0 \Rightarrow z(uv) = 1. \tag{ii}$$

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

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

$$\sum c(uv)z(uv) = \sum \Delta_\pi(uv)z(uv) + \sum c_\pi(uv)z(uv) =$$
$$k\pi(t) + \sum[c_\pi(uv)z(uv) : c_\pi(uv) > 0] + \sum[c_\pi(uv)z(uv) : c_\pi(uv) < 0] \ge$$
$$k\pi(t) + 0 + \sum[c_\pi(uv) : c_\pi(uv) < 0].$$

Ebből egyrészt következik a min $\ge$ 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. 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 $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. 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 $\delta(S) = \delta_z(S) - \varrho_z(S) = \delta_z(s)$, 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:

Állítás 1.7.13 *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, 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 $z'(uv) = 1$, ha uv a P-nek előre-éle és legyen $z'(uv) = 0$, ha vu a P-nek hátra-éle.

A módosításból adódik:

Állítás 1.7.14 *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. •

Az általános eset

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

TÉTEL 1.7.15 *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbf{Z}_+$ egészértékű kapacitásfüggvény és a $c : A \to \mathbf{R}_+$ költségfüggvény. A k nagyságú egészértékű megengedett folyamok költségének minimuma egyenlő a*

$$k\pi(t) + \sum [c_\pi(uv) g(uv) : uv \in A, c_\pi(uv) < 0] \tag{1.42}$$

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:

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

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

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

Az algoritmus egy általános helyzetében, ahol adott a z folyam és a 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 $z(vu) > 0$.

2. Fejezet

LINEÁRIS ALGEBRA: LINEÁRIS EGYENLETRENDSZEREK

2.1 VEKTORTÉR, ALTÉR, LINEÁRIS FÜGGETLENSÉG

Az alábbiakban áttekintjük a V Euklideszi vektortér néhány alaptulajdonságát. Az alaptest mindig a valós ($\mathbf{R}$) vagy a racionális ($\mathbf{Q}$) számok teste. Amikor Euklideszi vektortérről beszélünk, a valós szám n-esek $\mathbf{R}^n$ terére (vagy a racionális szám n-esek $\mathbf{Q}^n$ terére) gondolunk. (Lineáris algebrában bebizonyítják, hogy bármely két n dimenziós Euklideszi vektortér egymással izomorf.) Mindig az $\mathbf{R}$-t használjuk alaptestként, de valamennyi állítás érvényes a $\mathbf{Q}$ esetén is. Remélhetőleg nem okoz majd zavart, hogy jelölésben nem teszünk különbséget a 0 szám és a vektortér nulleleme 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 $b := \lambda_1 x_1 + \ldots + \lambda_k x_k$ 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 kombinációná**k 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.

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 **lineárisan függetlenek**nek mondjuk. Az $x_1, \ldots, x_k$ vektorok **kört** 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 azt tesszük fel, hogy az $x_1, \ldots, x_k$ 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 $(a_1, a_2, \ldots, a_n)$ 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, azaz az a-nak megfelelő oszlopvektor.

Legyen X és Y két vektor-halmaz $\mathbf{R}^n$-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 $\lambda \in \mathbf{R}$ 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 $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$ vektor **skalárszorzata** az $ab := a_1b_1 + \ldots + a_nb_n$ 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 $x_1, \ldots, x_k$ vektorok mindegyikére, akkor ortogonális ezek bármely lineáris kombinációjára is.

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

Altereket más módon is előállíthatunk. Legyen $a \in \mathbf{R}^n$ nem-nulla vektor. Az a-ra ortogonális elemek $\{x \in \mathbf{R}^n : ax = 0\}$ 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-dik egységvektor, úgy az a-ra merőleges vektorok halmazát **koordináta-hipersík**nak hívjuk. Ez tehát mindazon vektorokból áll, amelyek i-dik koordinátája 0. Könnyen látszik, hogy egy homogén hipersík alteret alkot. Ebből adódik, hogy homogén hipersíkok metszete is altér. Másszóval, adott $x_1, \ldots, x_k$ vektorok mindegyikére ortogonális vektorok halmaza alteret alkot, melyet az $x_1, \ldots, x_k$ **ortogonális kiegészítő alterének,** más néven **nullteré**nek nevezünk. Ha az x_i vektorok mindegyike valamelyik $(0, 0, \ldots, 1, \ldots, 0)$ alakú egységvektor, úgy ezek ortogonális kiegészítő alterét **koordináta-altér**nek 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 vetületé**t úgy kapjuk, hogy x j-dik komponensét 0-re állítjuk. Valójában ez a **belső** vetület, megkülönböztetendő a **külső** vetülettől, amelyet úgy kapunk, hogy az x vektor j-dik 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. $x \in U, \lambda \in \mathbf{R}$ 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 **magteré**nek 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 **képteré**nek neveznek.

Mese. Egy $x = (x_1, \ldots, x_n)$ szám n-esen néha az $\mathbf{R}^n$ $(x_1, \ldots, x_n)$ koordinátájú pontját értjük, néha az origóból az x pontba mutató vektort. 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 $m \times n$-es mátrix, azaz A-nak m sora és n oszlopa van. A mátrix i-dik oszlopát a_i-vel jelöljük, a j-dik sorát pedig ${}_ja$-val. Az A **sorteré**n az $\mathbf{R}^n$-nek az A sorai által generált alterét értjük, amelynek jele $\mathcal{S}(A)$ vagy $\mathbf{R}^m A$. A sortér tehát az $\{y^t A : y \in \mathbf{R}^m\}$ halmaz. Az A mátrix **nulltere** az A sorainak ortogonális kiegészítő altere, melynek jele $\mathcal{N}(A)$. A nulltér tehát az $Ax = 0$ egyenletrendszer $\{x \in \mathbf{R}^n : Ax = 0\}$ megoldás-halmaza. Az A oszlopai által generált altér az A **oszloptere**, jelölésben $\mathcal{O}(A)$ vagy $A\mathbf{R}^n$, 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 sor vektorokat. 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-dik oszlop együtthatója a z i-dik komponense. Hasonlóképp, egy m-dimenziós y vektorra az yA tekinthető, mint az A sorainak egy lineáris kombinációja.

Futólag már említettük, hogy ha egy $z \in \mathbf{R}^n$ 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, $(yA)z = y(Az) = y0 = 0$.

Lemma 2.2.1 *Ha a $z \in \mathbf{R}^n$ 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 $y \in \mathbf{R}^m$-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 $\mathbf{R}^n$ 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 $\mathbf{R}^m$-nek egymásra merőleges alterei, melyek metszete a triviális altér.

Az A mátrix segítségével megadható egy $\varphi_A : \mathbf{R}^n \to \mathbf{R}^m$ 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 ${}_A\varphi : \mathbf{R}^m \to \mathbf{R}^n$ 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.

TÉTEL 2.2.2 *Legyen A $m \times n$-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 $(1, 0, 0, \dots, 0)^t$ 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.

Következmény 2.2.3 *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. •

Következmény 2.2.4 *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. •

Lemma 2.2.5 *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.

TÉTEL 2.2.6 *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.)

Következmény 2.2.7 *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.* •

TÉTEL 2.2.8 *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 $a_1, \ldots, a_m, b$ vektorok lineárisan összefüggőek. Egy ilyen nem-triviális lineáris összefüggésben a b együtthatója nem lehet 0, mert ez azt jelentené, hogy az $a_1, \ldots, a_m$ 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.

TÉTEL 2.2.9 *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.

TÉTEL 2.2.10 (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.

Feladat 2.2.1 *Igazoljuk, hogy $r(Y \cdot A) \leq r(A)$!*

Feladat 2.2.2 *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.

TÉTEL 2.3.1 *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. •

TÉTEL 2.3.2 *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 $z_1 \in \mathbf{R}^n$ 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. •

Feladat 2.3.1 *Tegyük fel, hogy az 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. $\mathbf{R}^n$ minden eleme egyértelműen áll elő egy Q és egy Q^*-beli elem összegeként.

Következmény 2.3.3 *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 $\{yB : y \in \mathbf{R}^n\}$ 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 $\mathbf{R}^n$ térben legfeljebb csak n vektor választható ki lineárisan függetlenül. Miután az n egységvektor lineárisan független, az $\mathbf{R}^n$ 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 $\{\lambda a : \lambda \in \mathbf{R}\}$ 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.

Gyakorlat 2.3.2 *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 $C - \{c\}$ altér. A C 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.

TÉTEL 2.3.4 *Tegyük fel, hogy az $Ax = b$ egyenletrendszernek x_0 megoldása. Ekkor a megoldások $M := \{x : Ax = b\}$ halmaza az $\mathbf{R}^n$ 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 $\mathbf{R}^m$ térben volt szemléletes jelentése, a megoldások halmazát viszont a sorteret magában foglaló $\mathbf{R}^n$ 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 ${}_1a, \ldots, {}_ma$ vektorok által generált altér x_0 vektorral történő eltolásával áll elő, vagyis az $\{yA + x_0 : y \in \mathbf{R}^m\}$ alakú vektorok halmaza, ahol A jelöli az ${}_1a, \ldots, {}_ma$ 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. •

Következmény 2.3.5 *Amennyiben az $Ax = b$ egyenletrendszernek x_0 egy megoldása, úgy a megoldások halmaza előáll véges sok vektor lineáris burkának x_0-lal történő eltolásáként, azaz $\{yB + x_0 : y \in \mathbf{R}^n\}$ 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 $\mathbf{R}^3$-ban.

Következmény 2.3.6 *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 $n - r(A_1) = n - r(A)$. •

Gyakorlat 2.3.3 *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.*

Feladat 2.3.4 *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 $yb = \beta$. Igazoljuk, hogy $ax = \beta$ akkor és csak akkor lineáris következmény, ha logikai.*

2011. október 29.file: lin10

3. Fejezet

LINEÁRIS EGYENLŐTLENSÉG-RENDSZEREK 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 bementi-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 $\mathbf{R}^3$-ban egy zárt féltérként képzelhető el. Egy egyenlőtlenség-rendszer megoldás-halmaza 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 három 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 hogyan is kéne magasabb dimenzióban a csúcsot definiálni.

Ami esetleg kézenfekvő a három dimenziós szemléletünkben, az lehet, hogy n dimenzióban nem is igaz. Vagy még ha igaz is, nem azért az, mert három dimenzióban jól látszik. Gondoljunk arra, hogy egy n pontú gráf szerkezete mennyivel összetettebb lehet, mint egy három pontúé. Azt nyilván senki nem 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 vektor mindig beforgatható a nemnegatív térszögletbe (azaz létezik egy olyan ortonormált transzformáció, amely az n vektor mindegyikét nemnegatív vektorba képezi).*

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

5. *Ha $\mathbf{R}^n$-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 vajon valóban teljesíti-e azokat az elvárásokat, amelyeket a szemléletes folytonosság képünk diktál. 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 $\mathbf{R}^4$-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. Négy dimenzióban azonban minden n-re lehet n csúcsú poliédert konstruálni úgy, hogy a csúcsai páronként szomszédosak legyenek.

Az 5. állítás három 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 három dimenzióban szemléletes állítások $\mathbf{R}^n$-be történő átvitelének nehézségeit jól érzékelteti az alábbi feladat.

Feladat 3.1.1 *Igazoljuk, hogy $\mathbf{R}^n$-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 ekvivalensek-e vajon egymással vagy sem. 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 pont között vezet k élidegen út, és beszélhetünk k-szoros fa-összefüggőségről, ha a gráfban létezik k élidegen feszítő fa. A háromszög (mint gráf) mutatja, hogy a két tulajdonság $k \geq 2$ esetén már nem ekvivalens.

Tegyük most fel, hogy a három-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?

3.2 KÚPOK, POLIÉDEREK, POLITOPOK

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.

Gyakorlat 3.2.1 *Ha z a $z_1, \dots, z_k$ pontok konvex kombinációja és mindegyik z_i a $v_1, \dots, v_l$ pontok konvex kombinációja, akkor z a $v_1, \dots, v_l$ pontoknak is konvex kombinációja.*

Gyakorlat 3.2.2 *Egy halmaz akkor és csak akkor konvex, ha bármely két elemének bármely konvex kombinációja a halmazhoz tartozik.*

Gyakorlat 3.2.3 *Konvex halmazok metszete is konvex.*

Mivel $\mathbf{R}^n$ konvex, tetszőleges C halmaz benne van egy őt tartalmazó legszűkebb konvex halmazban, nevezetesen a C-t tartalmazó konvex halmazok metszetében. Ezt a C **konvex burká**nak (convex hull) nevezzük és konv(C)-vel jelöljük.

Gyakorlat 3.2.4 *A konv(C) halmaz nem más, mint a C elemeinek felhasználásával készült konvex kombinációk halmaza.*

Amennyiben $a \in \mathbf{R}^n$ nem-nulla vektor, β tetszőleges szám, az $ax = \beta$ lineáris egyenlet megoldás-halmazát **hipersík**nak (hyperplane) nevezzük. Ez az $\{x \in \mathbf{R}^n : ax = 0\}$ 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-rendszer megoldás-halmazát, vagyis az $\{x : ax \leq \beta\}$ halmazt (zárt) **féltér**nek (closed halfspace) hívjuk, melynek $\{x : Ax = b\}$ a **határoló síkja**, és amelynek normálisa a. (Ha szigorú egyenlőtlenség van, **nyílt féltér**ről beszélünk). A $\beta = 0$ esetben azt mondjuk, hogy a féltér **homogén**.

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

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

Megjegyzendő, hogy ha $\{p_1, p_2, \dots, p_t, a_1, \dots, a_n\}$ vektorok esetén a $\{z : z = \sum_j \mu_j p_j + \sum_i \lambda_i a_i, \lambda_i \ge 0\}$ halmaz generált kúp (vagyis ha csak bizonyos együtthatókra követelünk meg nemnegativitást), éspedig a $\{p_1, -p_1, \dots, p_t, -p_t, a_1, \dots, a_n\}$ vektorok kúpja. 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 $\{x : q_1x = 0, \dots, q_tx = 0, b_1x \le 0, \dots, b_mx \le 0\}$ halmaz metszetkúp, éspedig $\{x : q_1x \le 0, -q_1x \le 0, \dots, q_tx \le 0, -q_tx \le 0, b_1x \le 0, \dots, b_mx \le 0\}$. Speciálisan, a $q_1, \dots, q_t$ vektorok nulltere (másnéven kiegészítő altere) metszetkúp. A metszetkúp tehát a nulltér általánosítása.

Gyakorlat 3.2.5 *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 $\{\lambda q : \lambda \in \mathbf{R}_+\}$ 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 kombináció**ján 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 $z + \vec{q} := \{x : x = z + \lambda q, \lambda \in \mathbf{R}_+\}$ halmazt értjük. Tehát az irány egy origóból kiinduló félegyenes, és a félegyenes egy eltolt irány.

Adott K kúphoz hozzárendelhetjük a $K^* := \{x : xz \le 0 \text{ minden } z \in K \text{ elemre}\}$ 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 $K \ne (K^*)^*$. 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 politopok

A metszetkúpénál általánosabb a következő fogalom:

Poliéder (polyhedron, tbsz: polyhedra): Véges sok féltér metszete: $R := \{x : Qx \le b\}$, 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 háromdimenziós térgeometriában megszokott (konvex) poliéderek megadhatók félterek metszeteként, vagyis kielégítik a fenti definíciót, ugyanakkor ez megenged nem korlátos poliédereket is. Például egy metszetkúp vagy egy affin altér poliéder.

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

Az $R := \{x : Px = b_0, Qx \le b_1\}$ 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.

Politop (polytope): Véges sok pont konvex burka. (Az üres halmazt is politopnak tekintjük, mint nulla darab pont konvex burka.) Azt mondjuk, hogy a politopot a szóbanforgó pontok generálják. Ezek szerint egyetlen pont is politopot alkot. Két pont által generált politop neve **szakasz**.

Gyakorlat 3.2.6 *Két poliéder metszete is poliéder. Két politop vektor-összege is politop.*

Gyakorlat 3.2.7 *Amennyiben* $R := \{x : Mx \le b\}$ *nemüres, úgy* R *akkor és csak akkor az egész tér, ha* $r(M) = 0$ (azaz M a csupa-nulla mátrix) *és* $b \ge 0$.

Természetes kérdések: Mikor létezik egy egyenlőtlenség-rendszernek megoldása, azaz mikor 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 politop, és megfordítva. További kérdés, hogy két egyenlőtlenség-rendszer megoldás-halmaza mikor ugyanaz, magyarán, mikor definiálják ugyanazt a poliédert? Kezdjük egy egyszerű megfigyeléssel.

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

Biz. Az R megadható $\{x : Qx \le b\}$ alakban és feltehetjük, hogy Q-nak a lehető legkevesebb sora van. Azt igazoljuk, hogy ekkor $b = 0$. Mindenesetre $b \ge 0$, hiszen $0 \in R$ miatt $0 = Q0 \le b$. Tegyük fel indirekt, hogy $b(i) > 0$ valamelyik i-re. A Q minimalitása miatt van olyan x' vektor, amely a $Qx \le b$ rendszerből egyedül a ${}_iqx \le b(i)$ 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 ${}_iq(2x^*) = 2\alpha_i qx' = 2b(i) > b(i)$, ellentmondásban R kúp voltával. •

Egy poliédert akkor nevezünk **korlátos**nak, ha létezik olyan K pozitív szám, amelyre $|x(i)| \le K$ 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 $\mathbf{R}^n$ 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 $q \ne 0$, nem-triviális mozgásvektorról beszélünk. Azt mondjuk, hogy z a poliéder **relatív belső pont**ja, ha létezik nem-triviális mozgásvektora. Ha nem létezik, akkor z **extrém**. Ha minden vektor mozgásvektor, akkor z **belső** pontja R-nek. A definícióból közvetlenül kiolvasható, hogy egy $z \in R$ elem mozgás vektorai alteret alkotnak, melynek neve a z **mozgástere**. Ennek dimenziója a z elem **szabadsági foka** vagy röviden **foka**. Extrém pont foka ezek szerint 0. A 2-dimenziós térben egy szakasz olyan poliéder, amelynek nincs belső pontja. Egy 3-dimenziós kocka belső pontjainak foka 3, egy lapjának belső pontja 2 fokú, míg egy élének belső pontja 1 fokú.

Gyakorlat 3.2.8 *Igazoljuk, hogy egy $z \in R$ pont akkor és csak akkor relatív belső pont, ha előáll más R-beli pontok konvex kombinációjaként.*

TÉTEL 3.2.2 *Az $R = \{x \in \mathbf{R}^n : Qx \le b\}$ 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 $Q(z \pm \lambda q) \le b$, vagyis q a z mozgásvektora.

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

Egy $\{z + \lambda q : \lambda \in \mathbf{R}\}$ alakú egyenest q irányú egyenesnek nevezünk. Legyen $R = \{x : Qx \le b\}$ 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.

Gyakorlat 3.2.9 *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.

Gyakorlat 3.2.10 *A poliéder egy iránya akkor és csak akkor extrém, ha nem állítható elő két tőle különböző R-beli irány nemnegatív kombinációjaként.*

Egy háromdimenziós poliédert lapok, élek illetve csúcsok határolnak. Ezeket a fogalmakat szeretnénk magasabb dimenzióra kiterjeszteni. Egy $R \subseteq \mathbf{R}^n$ 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 $c \equiv 0$ 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 $c \neq 0$ esetben a $H = \{x : cx = \delta\}$ hipersíkot a poliéder egy **támasz-síkjá**nak nevezzük.

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 poliéder egy z elemét **extrém pont**nak hívjuk, ha nem áll elő a poliéder néhány más pontjának konvex kombinációjaként.

Gyakorlat 3.2.11 *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.*

3.2.3 Bázis-megoldások

TÉTEL 3.2.3 *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. •

Következmény 3.2.4 *Az $R := \{x \in \mathbf{R}^n : Qx \leq b\}$ poliéder eltolási altere a Q mátrix nulltere.*

Feladat 3.2.12 *Az R poliéder egy z pontját tartalmazó legbővebb, R-ben fekvő affin altér az A eltolási altér z-vel való eltoltja.*

Feladat 3.2.13 *Egy $R := \{x : Qx \leq b\} \subseteq \mathbf{R}^n$ poliéder belső dimenziója $n - r(Q)$.* •

TÉTEL 3.2.5 *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 3.5.1 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 3.2.2 tétel szerint mind a $Q^=_z$, mind a $Q'^=_z$ nulltere a z mozgástere, így $Q^=_z$ és $Q'^=_z$ sortere is ugyanaz. •

Egy $z \in R$ elem **szint**jén a $\sigma(z) := r(Q) - r(Q^=_z)$ számot értjük. Egy $z \in R$ elemet bázis-megoldásnak nevezünk, ha a z-aktív $Q^=_z$ részmátrix rangja $r(Q)$, más szóval a 0 szintű elemek a bázis-megoldások.

Következmény 3.2.6 *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.*

TÉTEL 3.2.7 *Minden nemüres poliédernek létezik bázis-megoldása, nevezetesen bármely minimális szintű elem bázis-megoldás.*

Biz. Legyen $z \in R$ minimális szintű elem. 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 $Q^<_z q \neq 0$. 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 ${}_iqz' = b(i)$ a $Q^<_z$ valamely ${}_iq$ sorára. $Q^=_z q = 0$ és ${}_iqq \neq 0$ miatt ${}_iq$ lineárisan független $Q^=_z$ soraitól. Így $Q^=_z z' = b_z$ miatt $r(Q^=_{z'}) > r(Q^=_z)$, ellentmondásban z választásával. •

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

TÉTEL 3.2.8 (i) *Egy* $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ *nem-nulla mátrix esetén a*

$$Px = b_0, Qx \leq b_1 \tag{3.2}$$

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

(ii) Esetleges sorcserével feltehetjük, hogy z-nek az utolsó j komponense pozitív. Jelölje az ezen j oszlophoz tartozó $m \cdot j$-es részmátrixot A'. Legyen $M := \begin{pmatrix} A \\ -I \end{pmatrix}$, ahol I az $n \cdot n$-es egységmátrixot jelöli. A tétel előtti megjegyzés 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 $(n - j) \cdot (n - j)$-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. •

Gyakorlat 3.2.14 *Igazoljuk, hogy a* 3.5.6 *tétel* (i) *részében z szintje* $r(M) - r(M_z^=)$.

Megjegyzés A szakirodalomban általában az $\{Ax = b, x \geq 0\}$ standard alakú 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!

Gyakorlat 3.2.15 *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.*

Gyakorlat 3.2.16 *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.*

3.3 A FOURIER-MOTZKIN ELIMINÁCIÓ

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 egyenlőtlenség-rendszerek 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.

3.3.1 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.

TÉTEL 3.3.1 *Az*

$$Ax \leq 0 \tag{3.3}$$

egyenlőtlenség-rendszernek bármely megoldása az

$$A^{[1]}x \leq 0 \tag{3.4}$$

rendszernek is megoldása. A (3.4) *bármely megoldásának első komponensét alkalmasan megváltoztatva a* (3.3) *egy megoldását kapjuk.*

Biz. 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 (3.4)-nek. Mivel $A^{[1]}$ első oszlopa 0, feltehető, hogy $z(1) = 0$. A $z(1)$ értékét fogjuk alkalmasan megváltoztatni, hogy (3.3) egy megoldását nyerjük.

Amennyiben J üres, vagyis A első oszlopában nincsen negatív elem, úgy z első komponensét kellően kicsi α értékre változtatva (3.3) egy megoldását kapjuk. (Nevezetesen, $\alpha := \min_{i \in I}\{-{}_iaz\}$ megteszi.) Analóg módon, ha I üres, úgy $z(1)$-t kellően nagyra változtatva kapunk (3.3)-nek egy megoldását. Tegyük most fel, hogy sem I, sem J nem üres. Állítjuk, hogy

$$\max_{j \in J}\{{}_jaz\} \leq \min_{i \in I}\{-{}_iaz\}. \tag{3.5}$$

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 (3.4)-nek. Mármost, ha α tetszőleges olyan szám, amelyre

$$\max_{j \in J}\{{}_jaz\} \leq \alpha \leq \min_{i \in I}\{-{}_iaz\}, \tag{3.6}$$

úgy z első komponensét α-ra változtatva a kapott z_α-ról állítjuk, hogy megoldása (3.3)-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 (3.5) második egyenlőtlensége folytán ${}_haz_\alpha = {}_haz + \alpha \leq 0$. Végül a $h \in J$ esetben ${}_ha(1) = -1$, és ekkor (3.5) első egyenlőtlensége folytán ${}_haz_\alpha = {}_haz - \alpha \leq 0$. •

Következmény 3.3.2 *Tegyük fel, hogy a Q mátrix első oszlopa $0, \pm 1$ értékű, és rendeljük a*

$$Qx' \leq b \tag{3.7}$$

egyenlőtlenség-rendszerhez a

$$Q^{[1]}x' \leq b^{[1]} \tag{3.8}$$

rendszert, ahol $b^{[1]}$ az $A := (Q, b)$ mátrixhoz tartozó $A^{[1]}$ mátrix utolsó oszlopa. Ekkor (3.7) *bármely megoldása a* (3.8) *rendszernek is megoldása, és a* (3.8) *bármely megoldásának első komponensét alkalmasan megváltoztatva a* (3.7) *egy megoldását kapjuk.*

Biz. Egy x' vektor pontosan akkor megoldása (3.7)-nek, ha az $x := (x', -1)$ megoldása $Ax \leq 0$-nak. Továbbá x' pontosan akkor megoldása (3.8)-nek, ha az x megoldása $A^{[1]}x \leq 0$-nak. Alkalmazhatjuk a 3.3.1 tételt. •

Következmény 3.3.3 *Metszetkúp tengelymenti* (külső vagy belső) *vetülete metszetkúp. Poliéder tengelymenti vetülete poliéder.*

Biz. A 3.3.1 tétel szerint az $\{x : Ax \leq 0\}$ metszetkúp x_1 tengelymenti belső vetülete az $\{x : A^{[1]}x \leq 0\}$ 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 $R = \{x' : Qx' \leq b^{[1]}\}$ poliéder x_1 menti belső vetülete a 3.3.2 következmény alapján az $\{x' : Q^{[1]}x' \leq b^{[1]}\}$ 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. •

Feladat 3.3.1 *Igazoljuk, hogy poliéder lineáris képe poliéder.*

3.3.2 Poliéder = politop + generált kúp

A 2.3.3 következmény szerint minden generált altér nulltér és minden nulltér generált altér, ami azzal ekvivalens, hogy egy $Ax = 0$ homogén egyenletrendszer megoldás-halmaza előáll véges sok vektor lineáris burkaként, és megfordítva, véges sok vektor lineáris kombinációinak halmaza előáll egy homogén lineáris egyenletrendszer megoldás-halmazaként. 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 $\mathbf{R}^n$-ben egy halmaz pontosan akkor poliéder, ha egy politop és egy generált kúp összege. Az egyik iránnyal kezdjük.

TÉTEL 3.3.4 *Egy politop és egy generált kúp összege poliéder. Speciálisan, minden politop korlátos poliéder és minden generált kúp előáll metszetkúpként.*

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

Tekintsük most $\mathbf{R}^{m+n+n'}$-ben az $R := \{(z, x, x') : Ax + A'x' - Iz = 0, x \geq 0, x' \geq 0, ex' = 1\}$ 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 $Ax + A'x' - Iz = 0, x \geq 0, x' \geq 0, ex' = 1$, azaz $z = Ax + A'x'$. Vagyis a külső vetület éppen C, és így a 3.3.3 következmény miatt C valóban poliéder. •

A lineáris programozás egyik sarokköve a Farkas Lemma. Ennek több ekvivalens alakja van: a legszemléletesebbel kezdjük.

TÉTEL 3.3.5 (Farkas lemma, geometriai alak) *Ha egy $C \subseteq \mathbf{R}^k$ generált kúp nem tartalmaz valamely $b \in \mathbf{R}^k$ elemet, akkor létezik olyan* (zárt) *homogén féltér, amely magában foglalja C-t, de nem tartalmazza b-t. Ha egy P politop 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 politop poliéder, azaz véges sok féltér metszete. •

A 3.3.4 tétel szerint egy politop és egy generált kúp összege poliéder. Megmutatjuk, hogy érvényes a megfordítás is, azaz minden poliéder előáll, mint egy politop és egy generált kúp összege. A bizonyítás érdekessége, hogy megfordítás igazolásához magát a 3.3.4 tételt használjuk fel.

Tekintsük először a speciális esetet, amikor egy metszetkúpot akarunk előállítani generált kúpként. Az A mátrixra jelölje G_A az A sorai által generált $\{yA : y \geq 0\}$ kúpot, míg M_A az A sorai által definiált $\{x : Ax \leq 0\}$ metszetkúpot.

Lemma 3.3.6 *Amennyiben az A mátrix sorai által generált kúp megegyezik egy B mátrix sorai által definiált metszetkúppal, azaz $G_A = M_B$, úgy $M_A = G_B$.*

Biz. Először lássuk be, hogy $M_A \subseteq G_B$. Ehhez legyen $z \in M_A$, vagyis ${}_i az \leq 0$ fennáll az A minden sorára. Emiatt az A sorainak bármely $q = \sum \lambda_i \ {}_i a$ $(\lambda_i \geq 0)$ nemnegatív kombinációjára $qz \leq 0$, azaz

$$G_A \text{ minden } q \text{ elemére } qz \leq 0. \tag{3.9}$$

Ha indirekt z nincs a G_B generált kúpban, akkor a 3.3.5 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 eleme $M_B = G_A$-nak, másrészt $qz > 0$, ellentmondva (3.9)-nek. Tehát $M_A \subseteq G_B$.

A fordított $G_B \subseteq M_A$ irány igazolásához figyeljük meg, hogy $G_A = M_B$ miatt ${}_i a\, {}_j b \leq 0$ fennáll az A minden ${}_i a$ és a B minden ${}_j b$ sorára. Emiatt minden $y \in G_B$ elemre ${}_i a y \leq 0$ fennáll vagyis $y \in M_A$. Tehát $G_B \subseteq M_A$ és így $G_B = M_A$. •

Feladat 3.3.2 *Állapítsuk meg, hogy az alábbi két állítás közül melyik igaz, melyik nem.* (A) *Ha* $G_A \supseteq M_B$, *úgy* $M_A \subseteq G_B$. (B) *Ha* $G_A \subseteq M_B$, *úgy* $M_A \supseteq G_B$.

TÉTEL 3.3.7 *Minden metszetkúp előáll generált kúpként.*

Biz. Az A mátrix sorai által definiált M_A metszetkúpról fogjuk kimutatni, hogy generált kúp. A 3.3.4 tétel szerint a G_A generált kúp előáll metszetkúpként, azaz létezik egy olyan B mátrix, amelyre $G_A = M_B$. A 3.3.6 lemma nyomán $M_A = G_B$, vagyis M_A a B sorai által generált kúp. •

Egy C kúp polárisán a $C^* := \{x : xy \leq 0 \text{ minden } y \in C\text{-re}\}$ kúpot értettük.

Feladat 3.3.3 *Minden generált kúp a polárisának polárisa, azaz* $(K^*)^* = K$.

A metszetkúpok előbbi előállítására támaszkodva megadjuk a poliéderek előállítását, amely tehát a 3.3.4 tétel megfordításának tekinthető.

TÉTEL 3.3.8 *Minden nemüres poliéder előáll mint egy politop és egy generált kúp összege. Speciálisan, minden korlátos poliéder politop.*

Biz. Legyen $R = \{x : Qx \leq b\}$ nemüres poliéder. A bizonyítás ötlete az, hogy R-t beágyazzuk egy eggyel magasabb dimenziós tér $\{(x, \lambda) : x \in \mathbf{R}^n, \lambda = 1\}$ hipersíkjába, ahol R az origóval egy kúpot feszít, majd ezen kúpra alkalmazzuk a 3.3.7 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 3.3.7 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: $(x_1, 1), \ldots, (x_k, 1)$, $(x'_1, 0), \ldots, (x'_l, 0)$. Tekintsük $\mathbf{R}^n$-ben az x_i-k konvex burkaként definiált P politopot 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 $R = P + C$. •

3.3.3 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 3.3.2 következmény folytán $Qx \leq b$ megoldhatósága ekvivalens $Q_1 x_1 \leq b^{[1]}$ megoldhatóságával, és a 3.3.1 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 $t + 1 \leq j_1 < j_2 \ldots < j_s \leq n$. 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.

3.3.4 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.

Gyakorlat 3.3.4 *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).

A 1.3.5 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 $\pi : V \to \mathbf{R}$ 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.

Feladat 3.3.5 *Igazoljuk, hogy egy élsúlyozott irányított gráfban akkor és csak akkor van negatív össz-súlyú irányított kör, ha nincsen olyan $\pi := V \to \mathbf{R}$ függvény, amelyre $\pi(v) - \pi(u) \leq c(uv)$ fennáll minden uv élre. Továbbá, ha a c súlyfüggvény egészértékű, akkor π is választható annak.*

lin12 2011. október 29.

3.4 MEGOLDHATÓSÁG: A FARKAS LEMMA

Valamely feladat algoritmikus megoldásához fontos lehet egy olyan eszköz, amelynek segítségével gyorsan ellenőrizhető egy javasolt megoldás helyessége, függetlenül attól, hogy a szóbanforgó megoldáshoz milyen nehézségek árán jutottunk. Például, behelyettesítéssel könnyű ellenőrizni, hogy egy megadott x_1 vektor megoldása-e az $Ax = b$ lineáris egyenletrendszernek. Arra is van egyszerűen ellenőrizhető tanúsítvány, ha az $Ax = b$ rendszernek nincsen megoldása, hiszen a Fredholm féle alternatíva tétel szerint ilyenkor létezik egy y, amelyre $yA = 0, yb \neq 0$, és egy adott y-ról ennek fennállása szintén könnyen eldönthető.

Kérdés, hogy létezik-e hasonló jellegű, könnyen ellenőrizhető tanúsítvány annak igazolására, hogy egy egyenlőtlenség-rendszer nem oldható meg, magyarán megadható-e a Fredholm féle alternatíva tétel lináris egyenlőtlenségekre vonatkozó ellenpárja. Az igenlő választ Farkas Gyula tétele szolgáltatja, amely a lineáris programozás egyik alapkövének tekinthető, és ezért röviden csak Farkas lemmának nevezik. Az eredmény poliéderek nemürességére ad szükséges és elegendő feltételt, és 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 geometriailag szemléletes változatot a 3.3 részben már be is bizonyítottuk. Most bemutatjuk a három leggyakoribb algebrai változatot, majd közös általánosításként egy olyan formát is felírunk, amelyből mind a három alak speciális esetként kiadódik. A Farkas lemmának az eredeti, Farkas Gyula által kimondott alakja a következő.

TÉTEL 3.4.1 (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. Tekintsük az A oszlopai által generált $K = \{Ax : x \geq 0\}$ kúpot. A Farkas lemma 3.3.5 geometriai alakja szerint K pontosan akkor tartalmazza b-t (más szóval $\{Ax = b, x \geq 0\}$ pontosan akkor oldható meg), ha nincs olyan F homogén féltér, amely tartalmazza a K kúp a_i generáló elemeit (és így magát a kúpot), 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 $yb < 0$. [Egy félteret akkor neveztünk homogénnek, ha határoló hipersíkja tartalmazza az origót.] •

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 úgy is megválasztható, hogy $yb = -1$ teljesüljön, így néha azt hívjuk duális feladatnak, amikor az $yb < 0$ helyett $yb = -1$-t követelünk. Néha a duálist az $\{yA \leq 0, yb > 0\}$ ekvivalens alakban adják meg: ennek és az eredeti duálisnak a megoldásai egymás minusz egyszeresei.

A 3.5.6 és a 3.5.5 tételek alkalmazásával megkaphatjuk a Farkas lemma standard alakjának egy élesítését.

TÉTEL 3.4.2 *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 lineárisan független vektorok kúpjában is. Egy y duális bázis-megoldás geometrialilag 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.

Egy másik geometriai szemléltetés

Érdekes, hogy a 3.4.1 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 nullttereként illetve sortereként, a Farkas lemma ekvivalens megfogalmazása a következő. *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.

A Farkas lemma változatai

TÉTEL 3.4.3 (Farkas lemma, (A) változat) *A $Qx \le b$ egyenlőtlenség-rendszer akkor és csak akkor oldható meg, ha nem létezik olyan $y \ge 0$, amelyre $yQ = 0, yb = -1$.*

Biz. A $Qx \le b$ feladatot nevezzük primál problémának, az $\{yQ = 0, y \ge 0, qb = -1\}$ feladatot pedig duálnak. Mindkettő nem oldható meg, mert akkor $0 = y(Qx) \le yb = -1$ 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 $Ax' \le 0$ és $c'x' = 1$, ami azzal ekvivalens, hogy $\alpha = -1$ és $Qx \le b$, 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 $Qx \le b$ primál feladat a $Qx' - Qx'' + x_1 = b$, $(x', x'', x_1) \ge 0$ 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 $yA \ge 0$, azaz $yQ \ge 0$ $y(-Q) \ge 0$, $yI_m \ge 0$, vagyis $yQ = 0$, $y \ge 0$. •

TÉTEL 3.4.4 (Farkas lemma, (B) változat) *A $\{Bx \le b, x \ge 0\}$ egyenlőtlenség-rendszer akkor és csak akkor oldható meg, ha nem létezik olyan $y \ge 0$, amelyre $yB \ge 0, yb = -1$.*

Biz. Mindkét rendszer nem oldható meg, mert akkor $0 \le (yB)x = y(Ax) \le yB = -1$. Jelölje Q azt a mátrixot, amely B-ből keletkezik azáltal, hogy aláírjuk az $n \cdot n$-es $-I$ egységmátrixot és jelölje b_1 azt a vektort, amely b-ből keletkezik n darab 0 komponens hozzáfűzésével. A $\{Bx \le b, x \ge 0\}$ rendszer pontosan akkor oldható meg, ha $Qx \le b_1$ megoldható. Az (A) változat szerint, ha $Qx \le b_1$ nem oldható meg, akkor létezik egy olyan $(y, y') \ge 0$ vektor, amelyre $yB + y'(-I) = 0, yb = -1$. De $y' \ge 0$ miatt $yB + y'(-I) = 0$ azzal ekvivalens, hogy $yB \ge 0$. •

Egyszerű fogással a Farkas lemmát olyan alakban is megfogalmazhatjuk, amikor a Fredholm féle alternatíva tételt már explicit magában foglalja.

TÉTEL 3.4.5 *A*

$$\{Px = b_0, Qx \le b_1\} \tag{3.10}$$

primál rendszer akkor és csak akkor oldható meg, ha az

$$\{y_0P + y_1Q = 0, y_1 \ge 0, yb = -1\} \tag{3.11}$$

duális nem, ahol $y = (y_0, y_1), b = (b_0, b_1)$.

Biz. Mindkét feladat nem oldható meg, mert akkor $0 = (y_0P + y_1Q)x = (y_0P)x + (y_1Q)x = y_0(Px) + y_1(Qx) \le y_0b_0 + y_1b_1 = -1$.

Ha a primál probléma nem oldható meg, akkor a vele ekvivalens $\{Px \le b_0, -Px \le -b_0, Qx \le b_1\}$ egyenlőtlenség-rendszer sem. Ekkor viszont a Farkas lemma (A) változata alapján az ehhez tartozó duál megoldható, azaz létezik $(y_0', y_0'', y_1) \ge 0$ 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 (3.11) duálisnak. •

Hasznos egy olyan alakot is felírni, amely a fenti változatok mindegyikét magában foglalja.

TÉTEL 3.4.6 (Farkas lemma, általános alak) *A*

$$Px_0 + Ax_1 = b_0,\ Qx_0 + Bx_1 \le b_1,\ x_1 \ge 0 \tag{3.12}$$

primál rendszernek akkor és csak akkor nincs megoldása, ha az

$$y_0P + y_1Q = 0,\ y_0A + y_1B \ge 0,\ y_1 \ge 0,\ yb := y_0b_0 + y_1b_1 = -1 \tag{3.13}$$

duális rendszernek van.

Biz. Jelölje az x_0, x_1, y_0, y_1 dimenzióját rendre n_0, n_1, m_0, m_1. Az $x_1 \ge 0$ feltételt explicit beírhatjuk az egyenlőtlenségek közé $0_{n_1n_0}x_0 + x_1(-I_{n_1n_1}) \le 0$ alakban. A 3.4.5 tételből közvetlenül kapjuk, hogy a primál probléma pontosan akkor oldható meg, ha az $\{y_0P + y_1Q = 0, y_0A + y_1B + y_1'(-I_{n_1n_1}) = 0, (y_1, y_1') \ge 0, y_0b_0 + y_1b_1 = -1\}$ rendszer nem. Ez utóbbi megoldhatósága viszont y_1' nemnegatívitása folytán épp (3.13) duálisával 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 megfelelő oszlopának negálásával azonban könnyen a (3.14) alakra juthatunk, így ez a legáltalánosabb alak már nem ad igazán újat.

3.4.1 Direkt bizonyítás

Bár a Farkas lemma standard alakját az előbb levezettük korábbi eredményekből (nevezetesen a Fourier-Motzkin eliminációra támaszkodva), érdemes egy közvetlen bizonyítást is megadni. Kényelmesebbnek bizonyul egy kicsit általánosabb alakot igazolni: persze még ez is speciális esete a fentebbi általános alaknak. (Nem ritka jelenség, hogy egy jól eltalált általánosításra az indukciós bizonyítás gördülékenyebben működik.) Szemben a fenti megközelítéssel, amely valójában speciális esetként kiadta a Fredholm féle alternatíva tételt, az itt következő bizonyítás használja azt.

TÉTEL 3.4.7 *A*

$$Px_0 + Ax_1 = b,\ x_1 \geq 0 \tag{3.14}$$

primál feladatnak akkor és csak akkor nincsen megoldása, ha az

$$yP = 0,\ yA \geq 0,\ yb = -1 \tag{3.15}$$

duális feladatnak van.

(Figyeljük meg, hogy (3.15) 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 oszlop-számra 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

$$Px_0 + A'x_1' = b,\ x_1' \geq 0 \tag{3.16}$$

rendszernek létezik megoldása, úgy az x_1'-t egy nulla komponenssel kiegészítve (3.14) megoldásához jutunk. Ha (3.16)-nak nincs megoldása, úgy az indukciós feltevés miatt az

$$yP = 0,\ yA' \geq 0,\ yb = -1 \tag{3.17}$$

rendszernek létezik y' megoldása. Amennyiben $y'a_1 \geq 0$, úgy y' a (3.15)-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

$$yP' = 0,\ yA' \geq 0,\ yb = -1 \tag{3.18}$$

problémának van megoldása, az nyilván megoldása (3.15)-nek is, és ekkor ismét csak készen vagyunk. Ha (3.18)-nek nincs megoldása, úgy a

$$P'x_0' + A'x_1' = b,\ x_1' \geq 0 \tag{3.19}$$

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 (3.14)-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 3.4.7 tételből is levezethető a 3.4.6 általános alak.

Biz. (3.4.6 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 (3.12) (x_0, x_1) megoldásai és

$$Px_0 + A'x_1' = b_0,\ Qx_0 + B'x_1' = b_1,\ x_1' \geq 0 \tag{3.20}$$

(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 (3.20)-hez tartozó

$$y_0P + y_1Q = 0,\ y_0A' + y_1B' \geq 0,\ yb = -1 \tag{3.21}$$

duál probléma és (3.13) ekvivalensek, hiszen csak arról van szó, hogy a (3.13)-ban explicit szereplő $y_1 \geq 0$ előjel megkötést (3.21)-ben az $y_0 A' + y_1 B' \geq 0$ egyenlőtlenség-rendszerben rejtettük el. A (3.20) primál feladatra felírva a 3.4.7 tételt az abban szereplő (3.15) duál feladat éppen a (3.21) 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.)

TÉTEL 3.4.8

$$y_0 P' + y_1 Q' = c'_0, \; y_0 A' + y_1 B' \geq c'_1, \; y_1 \geq 0 \tag{3.22}$$

primál rendszernek akkor és csak akkor nincs megoldása, ha a

$$P' x_0 + A' x'_1 = 0, \; Q' x_0 + B' x'_1 \leq 0, \; x'_1 \geq 0, \; c'_0 x_0 + c'_1 x'_1 > 0 \tag{3.23}$$

duális rendszernek van. •

Gyakorlat 3.4.1 *Az* $\{yA \leq c\}$ *rendszernek akkor és csak akkor van megoldása, ha nincs olyan* $x \geq 0$*, amelyre* $Ax = 0, cx < 0$.

Feladat 3.4.2 *Tekintsük a* (3.15)*-ben 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* (3.14) *alakkal.*

Feladat 3.4.3 *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* (3.14) *vagy* (3.15) *rendszer megoldására?*

Feladat 3.4.4 *Tegyük fel, van egy szubrutinunk a* $\{Bx \leq b, x \geq 0\}$ *rendszer megoldására. Hogyan használhatjuk fel ezt a* $\{Qx \leq b\}$ *rendszer megoldására?*

3.4.2 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 $yb \leq \gamma$.

TÉTEL 3.4.9 *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

$$\alpha \geq 0, Qx^* - \alpha b \leq 0, cx^* - \alpha\gamma > 0. \tag{3.24}$$

Ha most $\alpha = 0$, akkor ez a

$$Qx^* \leq 0, \; cx^* > 0 \tag{3.25}$$

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 (3.24) 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 3.4.9 tétel kiterjesztését is. Tekintsük a

$$Px_0 + Ax_1 = b_0, \; Qx_0 + Bx_1 \leq b_1, \; x_1 \geq 0 \tag{3.26}$$

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

$$c_0 x_0 + c_1 x_1 \leq \gamma \tag{3.27}$$

egyenlőtlenség **logikai következménye** a (3.26) rendszernek, ha (3.26) minden megoldása teljesíti (3.27)-t. A (3.27) egyenlőtlenség **lineáris következménye** (3.26)-nek, ha létezik olyan $y := (y_0, y_1)$, amelyre

$$y_1 \geq 0,\ y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\quad yb \leq \gamma. \tag{3.28}$$

TÉTEL 3.4.10 *Feltéve, hogy R nemüres,* (3.27) *akkor és csak akkor lineáris következménye* (3.26)*-nek, ha logikai.*

Biz. Tegyük fel először, hogy $cx \leq \gamma$ lineáris következmény, azaz létezik a szóbanforgó y. Ekkor (3.26) 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

$$Px = b_0, Qx \leq b_1 \tag{3.29}$$

alakra egyszerűsödik, és ilyenkor a $cx \leq \gamma$ egyenlőtlenség definíció szerint akkor lineáris következmény, ha létezik olyan $y := (y_0, y_1)$, amelyre

$$y_1 \geq 0,\ y_0P + y_1Q = c,\ yb \leq \gamma. \tag{3.30}$$

(3.29) 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 az első rész szerint létezik olyan $(y'_0, y''_0, y_1) \geq 0$ vektor, amelyre $y'_0P + y''_0(-P) + y_1Q = c$ és $y'_0b_0 + y''_0(-b_0) + y_1b_1 \leq \gamma$, de ekkor $y_0 := y'_0 - y''_0$ választással (3.30) teljesül, azaz $cx \leq \gamma$ lineáris következmény (3.29)-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. •

lin13 2011. október 29.

3.5 POLIÉDEREK

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

$$Qx \leq b \tag{3.31}$$

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.

3.5.1 Bázis-megoldások

TÉTEL 3.5.1 *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. •

Következmény 3.5.2 *Az $R := \{x \in \mathbf{R}^n : Qx \leq b\}$ poliéder eltolási altere a Q mátrix nulltere.*

Feladat 3.5.1 *Az R poliéder egy z pontját tartalmazó legbővebb, R-ben fekvő affin altér az A karakterisztikus altér z-vel való eltoltja.*

Feladat 3.5.2 *Egy $R := \{x : Qx \leq b\} \subseteq \mathbf{R}^n$ poliéder belső dimenziója $n - r(Q)$.* •

TÉTEL 3.5.3 *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 3.5.1 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 indirekt tegyük fel, hogy mondjuk $Q^=_z$ sortere nem tartalmazza $Q'^=_z$ valamely q' sorát. Ekkor a Fredholm féle alternatíva tételből következik, hogy létezik olyan q vektor, amelyre $Q^=_z q = 0$, $qq' > 0$. De ekkor kicsiny pozitív λ-ra $z' := z + \lambda q$ olyan, hogy $Qz' \leq b$, de $qq' > 0$ miatt $Q'z' \not\leq b'$, ellentmondásban a feltevéssel. •

Egy $z \in R$ elem szintjén a $\sigma(z) := r(Q) - r(Q^=_z)$ számot értettük. Egy $z \in R$ elemet bázis-megoldásnak neveztünk, ha a z-aktív $Q^=_z$ részmátrix rangja $r(Q)$, más szóval a 0 szintű elemek a bázis-megoldások.

Következmény 3.5.4 *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.*

TÉTEL 3.5.5 *Minden nemüres poliédernek létezik bázis-megoldása, nevezetesen bármely minimális szintű elem bázis-megoldás.*

Biz. Legyen $z \in R$ minimális szintű elem. 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 $Q^<_z q \neq 0$. 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 ${}_iqz' = b(i)$ a $Q^<_z$ valamely ${}_iq$ sorára. $Q^=_z q = 0$ és ${}_iqq \neq 0$ miatt ${}_iq$ lineárisan független $Q^=_z$ soraitól. Így $Q^=_z z' = b_z$ miatt $r(Q^=_{z'}) > r(Q^=_z)$, ellentmondásban z választásával. •

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

TÉTEL 3.5.6 (i) *Egy $M = \begin{pmatrix} P \\ Q \end{pmatrix}$ nem-nulla mátrix esetén a*

$$Px = b_0, Qx \leq b_1 \tag{3.32}$$

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

(ii) Esetleges sorcserével feltehetjük, hogy z-nek az utolsó j komponense pozitív. Jelölje az ezen j oszlophoz tartozó $m \cdot j$-es részmátrixot A'. Legyen $M := \begin{pmatrix} A \\ -I \end{pmatrix}$, ahol I az $n \cdot n$-es egységmátrixot jelöli. A tétel előtti megjegyzés 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 $(n-j) \cdot (n-j)$-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. •

Gyakorlat 3.5.3 *Igazoljuk, hogy a* 3.5.6 *tétel* (i) *részében z szintje* $r(M) - r(M_z^=)$.

Megjegyzés A szakirodalomban általában az $\{Ax = b, x \geq 0\}$ 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!

A 3.5.6 és a 3.5.5 tételek alkalmazásával megkaphatjuk a Farkas lemma standard alakjának egy élesítését.

TÉTEL 3.5.7 *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 lineárisan független vektorok kúpjában is. Egy y duális bázis-megoldás geometrialilag 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.

Gyakorlat 3.5.4 *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.*

Gyakorlat 3.5.5 *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.*

3.5.2 Csúcsos poliéderek

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

TÉTEL 3.5.8 *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. A (0) és (1) feltételek ekvivalenciája a definíciókból közvetlenül kiolvasható.

(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.

(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 $x, y \in P - z$, 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. •

Következmény 3.5.9 *Egy poliédernek legfeljebb véges sok csúcsa van.*

Biz. A 3.5.8 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.

TÉTEL 3.5.10 *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 3.5.1 tételből.

(4)⇒(1) Ha z csúcs, akkor a 3.5.8 tétel nyomán $Q_z^=$ oszlopai lineárisan függetlenek, így persze Q oszlopai is azok.

(1)⇒(4) A 3.5.5 tétel miatt van bázis-megoldás, és a 3.5.8 tétel miatt bármely z bázis-megoldás csúcs. •

TÉTEL 3.5.11 *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 3.5.10 tételből adódik, hogy R' egyenes-mentes. •

Feladat 3.5.6 *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.*

3.5.3 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.

TÉTEL 3.5.12 *Valamely nemnulla q vektorra a következők ekvivalensek:*
(1) $\vec{q}$ a poliéder iránya.
(2) R-nek van olyan z pontja, amelyre a $\{z + \lambda q : \lambda \geq 0\}$ félegyenes R-ben van.
(3) $Qq \leq 0$.

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

Következmény 3.5.13 *Az $R := \{x \in \mathbf{R}^n : Qx \leq b\}$ poliéder iránykúpja a Q mátrix $M_Q = \{x : Qx \leq 0\}$ metszetkúpja.*

Gyakorlat 3.5.7 *Igazoljuk, hogy egy $\{x : Px = b_0, Qx \leq b_1\}$ alakban adott nemüres poliéder iránykúpja $\{x : Px = 0, Qx \leq 0\}$.*

Feladat 3.5.8 *Egy poliédernek és iránykúpjának extrém irányai ugyanazok.* •

TÉTEL 3.5.14 *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 3.5.10 tétel miatt van csúcsa. A 3.5.9 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_1z_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) tulajdonsága 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.

A 3.3.8 tételben már láttuk, hogy minden korlátos poliéder politop, azaz véges sok pont konvex burka, sőt a 3.5.14 tételben azt is beláttuk, hogy minden nemüres korlátos poliéder a csúcsainak konvex burka. A bizonyítás gondolatmenetét használva most megadjuk a csúcsos poliéderek előállítását.

TÉTEL 3.5.15 *Minden $R \neq \emptyset$ egyenes-mentes* (azaz csúcsos) *poliéder előáll, mint a csúcsai által feszített R_K politop valamint a poliéder C karakterisztikus kúpjának a vektor-összege.*

Biz. Tegyük fel, hogy $R = \{x : Qx \leq b\}$ és legyen z a poliéder egy pontja. Legyen $Q_z^=$ a z aktív részmátrixa. A 3.5.8 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, $\lambda_1 := \min((b_1(i) - {}_iaz)/(-{}_iaq)$, ahol a minimum az Q azon ${}_ia$ soraira megy, amelyekre ${}_iaq < 0$. Az x_1 szintje valóban kisebb, mint z szintje, hiszen egyrészt $Q_z^= x_1 = b_z^=$, másrészt $Q_z^<$ egyik ${}_iq$ sorára ${}_iqx_1 = b(i)$ és ez a sor ${}_iqq \neq 0$ 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 ${}_1q$ és ${}_2q$ a $Q_z^<$ mátrixnak két olyan sora, amelyre ${}_1qq < 0$ és ${}_2qq > 0$. 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 Valójában a fenti bizonyítást a speciális korlátos esetben már korábban, a 3.5.14 tétel (2) pontjának bizonyításakor elmondtuk.

3.5.4 Alkalmazások

TÉTEL 3.5.16 *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. •

TÉTEL 3.5.17 (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 $d+1$. •

TÉTEL 3.5.18 *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 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 $cx' = (yQ)x' = -(y'Q')x' = -y'(Q'x') \geq -y'b' > \alpha$. •

TÉTEL 3.5.19 (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 3.3.4 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 3.5.16 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. •

TÉTEL 3.5.20 (Kirchberger) *Az n dimenziós térben adott k piros és l zöld pont, ahol $k + l \geq n+2$. 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 $(1, 1, \ldots, 1, 0, \ldots, 0)$ sorvektorral, amely k darab egyest tartalmaz, valamint egy $(0, \ldots, 0, 1, 1, \ldots, 1)$ 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 $yP + \alpha e_k \geq 0, -yZ + \beta e_l \geq 0$ és $\alpha + \beta < 0$, ahol e_i a csupa 1-esből álló i-dimenziós vektort jelöli. Ez viszont azt jelenti, hogy $\{x : yx = -\alpha\}$ 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 fontos eredmény a valószínűségszámítás területéről.

TÉTEL 3.5.21 *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 $e_n = (1, \ldots, 1)$.*

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 $y \ll yA \leq (\mu e_n)A = (\mu, \ldots, \mu)$, ellentmondásban μ választásával. •

Feladat 3.5.9 *A Farkas lemma felhasználásával igazoljuk a következő eredményt.*

TÉTEL 3.5.22 *Legyen $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbf{R}_+$ kapacitás függvény. Akkor és csak akkor létezik az s_i pontból t_i-be α_i nagyságú folyam $(i = 1, \ldots, k)$ úgy, hogy minden élre a rajta átmenő folyamértékek összege legfeljebb az él kapacitása, ha az éleken értelmezett tetszőleges c nemnegatív költségfüggvényre $\sum_{i=1}^{k} l_c(i)\alpha_i \leq \sum_{e \in A} c(e)g(e)$, ahol $l_c(i)$ jelöli az s_i-ből t_i-be vezető utak minimális c-költségét.*

lin14 2011. október 29.

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.

4.1.1 Erős bázis-megoldások

A $Qx \leq b$ egyenlőtlenség-rendszer egy z bázis-megoldását **erős bázis-megoldás**nak nevezzük, ha a z nem-nulla komponenseihez tartozó Q-beli oszlopok lineárisan függetlenek. Speciálisan, ha Q oszlopai lineárisan függetlenek, akkor minden bázis-megoldás erős. Megmutatjuk, hogy az erős bázis-megoldás fogalma is csak a poliédertől függ. Többet látunk be.

TÉTEL 4.1.1 *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 3.5.3 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. •

Következmény 4.1.2 *Az erős bázis-megoldás fogalma csak a poliédertől függ és nem a poliédert meghatározó egyenlőtlenség-rendszertől.* •

Következmény 4.1.3 *A $\{Px = b_0, Qx \leq b_1\}$ egyenlőtlenség-rendszer egy z bázis-megoldása pontosan akkor erős, ha a z nem-nulla komponenseihez tartozó M-beli oszlopok lineárisan függetlenek, ahol $M = \begin{pmatrix} P \\ Q \end{pmatrix}$.* •

Feladat 4.1.1 *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.*

Gyakorlat 4.1.2 *Legyen $f \ll g$, (azaz f minden komponensében kisebb, mint g), ahol $f, g \in \mathbf{R}^n$. 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 $f \leq g$-t tesszük fel?*

Ki fogjuk mutatni, hogy mindig létezik erős bázis-megoldás, legfeljebb véges sok van belőlük, továbbá 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-n. Először lássuk be, hogy legfeljebb csak véges sok erős bázis-megoldás létezhet.

TÉTEL 4.1.4 *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 az Q oszlopai közül $r(Q)$ darab lineárisan független oszloppá. Az így kapott $r(Q)$ sor és $r(Q)$ oszlop által meghatározott Q' részmátrix az 2.2.5 lemma miatt nem-szinguláris, és éppen z-t definiálja a kívánt módon. •

Következmény 4.1.5 *Tetszőleges egyenlőtlenség-rendszernek legfeljebb csak véges sok erős bázis-megoldása van.* •

4.1.2 Az iránymenti korlátosság feltétele

Következő célunk megmutatni, hogy ha egy egyenlőtlenség-rendszer megoldható, akkor van erős bázis-megoldása is. Valójában ennél többet is belátunk. 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.

Lemma 4.1.6 *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 az 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 λ' 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 3.5.5 tételt bázis-megoldás létezéséről. A fő eredményt először speciális alakban fogalmazzuk meg.

TÉTEL 4.1.7 (Az iránymenti korlátossági 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-n 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 4.1.5 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 4.1.6 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. •

Következmény 4.1.8 *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 4.1.5 tétel szerint véges sok erős bázis-megoldás van. A 4.1.7 tételből adódóan a maximum ezek egyikén felvétetik. •

Megjegyzés A 4.1.7 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 $x \in R$-ra $cx = (yQ)x = y(Qx) \leq yb$, magyarán az yb érték egy konkrét felső korlát.)

Következmény 4.1.9 *Ha egy egyenlőtlenség-rendszer megoldható, akkor van erős bázis-megoldása is.*

Biz. A 4.1.7 tételben $c = 0$-ra (0) fennáll, így a tétel alapján (1) is. •

Az irodalomban gyakran Caratheodory tételnek hívják a 4.1.9 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 (3.4.8 tétel) a 4.1.7 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.

TÉTEL 4.1.10 (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-n 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*

$$y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\ y_1 \geq 0. \bullet \tag{4.2}$$

lin15 2011. október 29.

4.2 OPTIMALITÁS: A DUALITÁS TÉTEL

Korábban megvizsgáltuk, hogy egy R poliéder mikor nemüres, majd azt, hogy egy cx lineáris célfüggvény mikor korlátos felülről R-n. Most rátérünk a lineáris programozás fő kérdésének tárgyalására: amennyiben R nemüres és $\{cx : x \in R\}$ korlátos felülről, hogyan jellemezhetjük a cx-t 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-n, a 4.1.7 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 algoritmussal 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 $R = \{x : Qx \le b\}$ 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 $Q^=_{x^*} x' \le 0$.

TÉTEL 4.2.1 *Tegyük fel, hogy az $R := \{x : Qx \le b\}$ poliéder nemüres és $\{cx : x \in R\}$ felülről korlátos. Az R egy megadott x^* elemére a következő állítások ekvivalensek.*

(1) *$cx^* \ge cx$ minden $x \in R$, azaz x^* maximalizálja a cx függvényt az R-n (röviden, x^* optimális).*

(2) *Nem létezik c-növelő irány, azaz olyan x' vektor, amelyre $Q^=_{x^*} x' \le 0$ és $cx' > 0$.*

(3) *A c vektor benne van x^* aktív sorainak kúpjában. Más szóval, van olyan y^* vektor, amely kielégíti az*

$$y^* \ge 0, y^*Q = c \tag{4.4}$$

duális feltételt, és amelyre fennáll az

$$y^*(i) > 0 \Rightarrow \ {}_iqx^* = b(i) \tag{4.5}$$

optimalitási kritérium (ami szavakban: az y^* bármely komponense csak akkor lehet pozitív, ha a neki megfelelő primál egyenlőtlenséget x^* egyenlőséggel teljesít). (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 $y' \ge 0$, 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

$$cx = (y^*Q)x = y^*(Qx) \le y^*b, \tag{4.8}$$

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). •

Az előbbi bizonyítás lépéseinek a másolásával kiterjeszthetjük a tételt az általános alakra.

Gyakorlat 4.2.1 *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* $x_1'(i) \geq 0$.

TÉTEL 4.2.2 *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*-n (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

$$y_1^*(i) > 0 \Rightarrow\ {}_iqx_0^* +\ {}_ibx_1 = b_1(i), \tag{4.14}$$

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 $M = \begin{pmatrix} P & A \\ Q & B \end{pmatrix}$.

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.

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 3.4.8 tételben megfogalmazott alakja szerint van olyan (y_0, y_1), amelyre $y_1 \geq 0$, $y_0P + y_1Q_{x^*}^= = c_0', y_0A + y_1B_{x^*}^= \geq c_1'$. 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

$$cx = c_0x_0 + c_1x_1 \leq [(y_0^*, y_1^*)\begin{pmatrix} P \\ Q \end{pmatrix}]x_0 + [(y_0^*, y_1^*)\begin{pmatrix} A \\ B \end{pmatrix}]x_1 = (y^*M)x = \tag{4.17}$$

$$= y^*(Mx) = y_0^*[Px_0 + Ax_1] + y_1^*[Qx_0 + Bx_1] \leq y_0^*b_0 + y_1^*b_1 = y^*b, \tag{4.18}$$

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 $c_1x_1 = (y_0^*, y_1^*)\begin{pmatrix} A \\ B \end{pmatrix}x_1^*$, 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 ${}_iqx_0^* +_i bx_1^* = b_1(i)$, azaz (4.14) teljesül.

Másrészt (4.15), (4.16) mindegyike azzal ekvivalens, hogy x^* mind (4.17)-t, mind (4.18)-t egyenlőséggel teljesíti. •

Megjegyezzük, hogy a 4.2.2 tétel bizonyítására alternatív lehetőség a 4.2.1 tételből a korábban már megismert átalakításokkal jutni az általános alakra.

Gyakorlat 4.2.2 *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 az ilyen 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.

TÉTEL 4.2.3 (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* 4.1.7 *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 $cx^* = by^*$. Az 4.1.8 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:

$$\min\{(c_0y_0 + c_1y_1) : y_0P + y_1Q = c_0,\ y_0A + y_1B \geq c_1,\ y_1 \geq 0\}. \tag{4.20}$$

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!)

TÉTEL 4.2.4 (Dualitás tétel) *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.

Feladat 4.2.3 *Irjuk fel a dualitás tételt* $\max\{cx : Bx \leq b, x \geq 0\}$ *illetve a* $\max\{cx : Ax = b, x \geq 0\}$ *primál lineáris programokra.*

Feladat 4.2.4 *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* (3.4.1 tétel).

Feladat 4.2.5 *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.*

Azt mondjuk, hogy egy $c'x$ célfügvény vagy a c' vektor valamely R poliéderre nézve **neutrális** (semleges), ha R minden x elemére $c'x$ értéke ugyanaz. Két vektort (célfüggvényt) akkor mondunk **ekvivalens**nek, ha különbségük neutrális.

Feladat 4.2.6 *Tekintsük a* $\min\{cx : Ax = b, x \geq 0\}$ *lineáris programot. Legyen* A' *az* A *érdekes oszlopai által alkotott részmátrix és jelölje* c' *a* c *megfelelő részét. Igazoljuk, hogy* c *akkor és csak akkor neutrális, ha létezik olyan* y*, amelyre* $yA' = c'$.

Feladat 4.2.7 *Az $R := \{Ax = b, x \geq 0\}$ poliéder valamely x_0 eleme akkor és csak akkor minimalizálja cx-t R-n, ha létezik egy olyan c-vel ekvivalens nemnegatív vektor, amely merőleges x_0-ra.*

Feladat 4.2.8 *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 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 x_1 primál optimum, y_1 duál optimum.*

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 3.4.10 tétel bizonyítása csak a Farkas lemmára támaszkodott. Most megmutatjuk, hogy a dualitás tétel könnyen levezethető a 3.4.10 tételből.

TÉTEL 4.2.5 (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-n, by pedig alulról R^*-n. 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 3.4.10 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 3.4.10 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 $\gamma \geq yb$.

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.

TÉTEL 4.2.6 (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, $\min\{(\max Ax) : x \geq 0, e_n x = 1\} = \max\{(\min yA) : y \geq 0, e_m y = 1\}$, 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 $e_n x = 1$ és $Ax \leq (w, w, \dots, w)$ érvényes. Ez viszont éppen a

$$\min\{w : -Ax + (w, w, \dots, w) \geq 0, x \geq 0, e_n x = 1\} \tag{4.21}$$

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 $yA \geq (z, z, \dots, z)$ érvényes. Ez viszont éppen a

$$\max\{z : y(-A) + (z, z, \dots, z) \leq 0, y \geq 0, e_m y = 1\} \tag{4.22}$$

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. •

TÉTEL 4.2.7 (Clark) *Tekintsük a $\max\{cx : x \geq 0, Bx \leq b\}$ és $\min\{by : y \geq 0, yB \leq c\}$ 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 $y'B - (\alpha, \dots, \alpha) \geq 0$, 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. •

TÉTEL 4.2.8 (Tucker) *Tegyük fel, hogy a Q négyzetes márix ferdén szimmetrikus. Ekkor a $Qx \ll 0$ rendszernek mindig van megoldása* (ahol $a \ll b$ azt jelöli, hogy az a vektor minden komponensében szigorúan kisebb, mint a b).

Biz. Először kimutatjuk, hogy a Q bármely ${}_iq$ sorára a $Qx \leq 0$ rendszernek van olyan megoldása, amelyre ${}_iqx \leq -1$. Ha ugyanis nem volna, akkor a Farkas lemma nyomán létezik olyan $y \geq 0$, vektor amelyre $yQ \geq 0$ és $y(i)(-1) < 0$, azaz $y(i) > 0 \ldots \alpha < 0 \ldots$

4.2.4 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.

TÉTEL 4.2.9 *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. •

Az R poliéder maga is oldal (pl. $c = 0$ célfüggvényre az R minden pontja maximalizálja cx-t, vagy másként, amikor semmilyen egyenlőtlenséget nem kötünk meg egyenlőségként.) A poliédernek egy önmagától különböző oldalát **valódi oldal**nak nevezzük. Egy tartalmazásra nézve maximális valódi oldalt **lap**nak hívunk. Fontos szerepet játszanak a (tartalmazásra nézve) **minimális oldal**ak, vagyis az olyan oldalak, melyek valódi részhalmazként már nem tartalmaznak más oldalt. Az R poliéder egy $R = \{x : Px = b_0, Qx \leq b_1\}$ leírásában szereplő $qx \leq \beta$ egyenlőtlenségről azt mondtuk, hogy **lényeges**, ha kihagyása megváltoztatja (bővíti) a poliédert. Az egyenlőtlenség **igazi**, ha egyenlőséggel történő cseréje megváltoztatja (szűkíti) a poliédert.

TÉTEL 4.2.10 *Egy R nemüres poliédernek akkor és csak akkor nincs valódi oldala, ha R affin altér.*

Biz. Ha $R = \{x : Qx = b\}$ affin altér, úgy az 5.3.1 tétel szerint nincs valódi oldala.

Megfordítva, tegyük fel, hogy az R poliédernek nincs valódi oldala. Tekintsük a poliédernek egy olyan $R = \{x : Px = b_0, Qx \leq b_1\}$ megadását, amelyben minden egyenlőtlenség igazi és lényeges. (Ilyen persze van, hiszen R egy tetszőleges leírásából kiindulva egymás után kihagyhatjuk az aktuálisan lényegtelen egyenlőtlenségeket, majd az implicit egyenlőségeket explicitté alakíthatjuk). Azt látjuk be, hogy Q üres, és így R valóban affin altér. Tegyük fel indirekt, hogy létezik egy $qx \leq \beta$ igazi és lényeges egyenlőtlenség. Jelölje Q' a q sor kihagyásával Q-ból keletkező részmátrixot és b_1' a megfelelő jobboldalt. Ekkor egyrészt van olyan $x' \in R$ pontja, amelyre $Px' = b_0$, $Q'x' \leq b_1'$ és $qx' > \beta$, másrészt R-nek van olyan x'' pontja, melyre $qx'' < \beta$. Így az $x'x''$ szakasznak van olyan z pontja, amelyre $qz = \beta$ és $z \in R$, vagyis $\{x : Px = b_0, Q'x \leq b_1', qx = \beta\}$ valódi nemüres oldala R-nek, ellentmondásban a feltevéssel, hogy ilyen oldal nem létezik. •

Következmény 4.2.11 *Egy poliéder minimális oldala affin altér.*

Biz. Az 5.3.1 tétel miatt az R poliéder egy oldalának oldala R-nek is oldala, így a minimális oldal olyan poliéder, amelynek már nincs valódi oldala. Alkalmazzuk az 5.3.2 tételt. •

lin16 2011. október 29.

4.2.5 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-dik 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 $4\cdot4$-s mátrix, amelynek a_{ij} eleme Sári nyereségét (másszóval Oszi veszteségét) jelzi, ha Sári az i-dik, míg Oszi a j-dik 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.

$$\begin{array}{c} \\ [1,1] \\ [1,2] \\ [2,1] \\ [2,2] \end{array} \begin{array}{c} \begin{array}{cccc} [1,1] & [1,2] & [2,1] & [2,2] \end{array} \\ \begin{pmatrix} 0 & 2 & -3 & 0 \\ -2 & 0 & 0 & 3 \\ 3 & 0 & 0 & -4 \\ 0 & -3 & 4 & 0 \end{pmatrix} \end{array}$$

Sári egy kevert stratégiáját egy y (m-dimenziós) sztochasztikus vektor definiálja éspedig úgy, hogy Sári minden fordulóban $y(i)$ valószínűséggel választja az i-dik sort.

Tegyük fel, hogy N forduló során Oszi c_j-szer játszotta meg a j-dik oszlopot ($\sum c_j = N$). Ekkor a j-dik oszlop gyakorisága $x_j := c_j/N$. Nyilván a gyakoriságok $x := (x_1,\ldots,x_n)$ 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 $(0, \ldots, 0, 1, 0, \ldots, 0)$ 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.

TÉTEL 4.2.12 (Neumann János) *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 stratágiáját 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.

2011. október 29.file: ljatek

4.3 POLIÉDEREK ELŐÁLLÍTÁSA

4.3.1 Oldalak

Egy poliéder oldalát a 3.2.2 részben definiáltuk, mint a poliédernek az a része, amely egy hipersíkkal érintkezik, amikor azt kívülről a poliéderhez toljuk. E definíció előnye, hogy szemléletes, ugyanakkor nem könnyű vele dolgozni. Például nem látszanak olyan elvárt tulajdonságok, minthogy két oldal nemüres metszete oldal, vagy hogy egy oldal oldala az eredeti poliédernek is oldala. Ezen segít a következő tétel, amely megmondja, hogy a poliéder egy konkrét egyenlőtlenség-rendszerrel történő megadásán miképp tükröződik az oldal fogalma.

TÉTEL 4.3.1 *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 (3.1) 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 vektorra $x \in R, Q'x = b'$ akkor és csak akkor teljesül, ha $x \in R, cx = \delta$, amiből a tétel következik. •

Tekintsük az $R := \{x : Px = b_0, Qx \leq b_1\}$ nemüres poliédert. Az alábbiakban megvizsgáljuk az R néhány tulajdonságát a

$$Px = b_0, Qx \leq b_1 \tag{4.24}$$

leíró rendszer függvényében.

A $Qx \leq b$ rendszer egy $qx \leq \beta$ egyenlőtlenségéről azt mondjuk, hogy **implicit egyenlőség**, ha R minden eleme egyenlőséggel teljesíti (azaz, ha a fordított $qx \geq \beta$ egyenlőtlenség logikai következménye az R-t meghatározó rendszernek. Geometrialilag ez azt jelenti, hogy az R poliéder teljesen az $\{x : \ qx = \beta\}$ hipersíkban fekszik.) Az implicit egyenlőségek által alkotott egyenlőség-rendszert $Q^{=}x = b^{=}$-vel jelöljük. Egy nem implicit egyenlőtlenséget **valódinak** mondunk, és az általuk alkotott rendszert $Q^{<}x \leq b_1^{<}$-vel jelöljük. (Az explicit $Px = b_1$ egyenlőség-rendszer tagjait automatikusan implicit egyenlőségeknek tekintjük.)

A $Qx \leq b_1$ egyenlőtlenség-rendszer egyik $qx \leq \beta$ tagját **feleslegesnek** nevezzük (a (4.24)-re nézve), ha az elhagyásával keletkező rendszer ugyanazt a poliédert definiálja. (Másszóval, $qx \leq \beta$ logikai következménye annak a rendszernek, amelyet (4.24)-ból kapunk $qx \leq \beta$ kihagyásával.) Egy nem felesleges egyenlőtlenség neve **lényeges**. Hasonlóképpen beszélhetünk arról, hogy egy $Px = b_0$-beli egyenlőség **lényeges** vagy **felesleges** annak megfelelően, hogy kihagyása R-nél bővebb poliédert eredményez-e vagy R-t változatlanul hagyja.

Egy poliéder leírásából egymás után kihagyva az (aktuálisan) felesleges egyenlőtlenségeket és egyenlőségeket olyan leírást kapunk, amelyben már minden egyenlőség és egyenlőtlenség lényeges. A kapott rendszer persze függhet az elhagyás sorrendjétől, hiszen egy eredetileg felesleges egyenlőtlenség egy másik elhagyásakor lényegessé válhat. Például, ha a poliéder a három dimenziós tér egy e egyenese, amely három (különböző) e-t tartalmazó sík metszeteként van adva, akkor e ezek közül bármelyik kettő metszeteként is megadható. Azt is feltehetjük, hogy minden szereplő $qx \leq \beta$ egyenlőtlenség valódi, mert különben helyettesíthetjük az $qx = \beta$ explicit egyenlőséggel. Nevezzük az R poliéder (4.24) alakú leírását **minimálisnak**, ha minden egyenlőség és egyenlőtlenség lényeges és minden egyenlőtlenség valódi. Látjuk tehát, hogy létezik minimális leírás.

Lemma 4.3.2 *Ha $qx \leq \beta$ valódi és lényeges egyenlőtlensége* (4.24)*-nek, akkor R-nek van olyan z pontja, amelyre $qz = \beta$.*

Biz. Miután $qx \leq \beta$ valódi, létezik olyan $x' \in R$, amelyre $qx' < \beta$. Miután $qx \leq \beta$ lényeges, létezik olyan x'', amely (4.24) rendszerből a $qx \leq \beta$ egyenlőtlenséget megsérti, de a többit mind kielégíti. Az $x'x''$ szakasz egyik vége R-ben van, másik nincs, így a szakasznak van olyan z pontja, amely R-ben van és $qz = \beta$. •

TÉTEL 4.3.3 *Tegyük fel, hogy az R poliéder egy minimális* (4.24) *alakú rendszerrel van adva. A következő állítások ekvivalensek.*
(i) *R affin altér.*
(ii) *R-nek nincs valódi oldala.*
(iii) *Minden lényeges egyenlőtlenség implicit egyenlőség (azaz Q üres).*

Biz. Tegyük fel, hogy R affin altér. Ekkor van $\{x : Ax = d\}$ alakú leírása. A 4.3.1 tétel szerint az oldal fogalma nem függ a poliéder leíró rendszerétől. Mivel itt csak egyenlőségek szerepelnek, R-nek nem lehet valódi oldala.

Tegyük most fel, hogy (ii) teljesül. Ha indirekt (iii) nem állna, akkor a (4.24)-ben létezne egy $qx \leq \beta$ valódi, lényeges egyenlőtlenség. Ehhez R-nek van olyan x' pontja, amelyre $qx' < \beta$, és a 4.3.2 lemma szerint olyan z pontja is, amelyre $qz = \beta$. De ekkor $\{x \in R : qx = \beta$ valódi (nemüres) oldala lenne R-nek.

A (iii)$\rightarrow$(i) irány triviális. •

Miután minimális oldalnak nincsen valódi oldala, az előző tételből kapjuk:

Következmény 4.3.4 *Minden minimális oldal affin altér.* •

Hasznos a következő lemma.

Lemma 4.3.5 *Az R poliédernek van olyan eleme, amely minden valódi egyenlőtlenséget szigorúan teljesít.*

***Biz.** Minden valódi egyenlőtlenséghez van R-nek olyan eleme, amely azt szigorúan teljesíti. Ezen elemek számtani közepe (vagy bármely pozitív együtthatós konvex kombinációja) R-ben van és valamennyi valódi egyenlőtlenséget szigorúan teljesíti. •

TÉTEL 4.3.6 *Tegyük fel, hogy* (4.24) *minimális leírása R-nek. Az R akkor és csak akkor van benne egy $H := \{x : ax = \beta\}$ hipersíkban (másszóval $ax = \beta$ akkor és csak akkor logikai következménye* (4.24)*-nek), ha létezik olyan y_0, amelyre $y_0P = a$ és $y_0b_0 = \beta$ (azaz, $ax = \beta$ lineáris következménye $Px = b_0$-nak.)*

***Biz.** Ha létezik y_0, akkor persze R minden x elemére $ax = \beta$, azaz R benne van a H hipersíkban. Tegyük most fel, hogy R benne van H-ban, azaz mind az $ax \leq \beta$, mind az $ax \geq \beta$ egyenlőlenség logikai következménye a (4.24) rendszernek. Korábban láttuk (3.4.10 tétel), hogy ha $ax \leq \beta$ logikai következmény, akkor létezik olyan $y = (y_0, y_1)$, amelyre $y_1 \geq 0$, $y_0P + y_1Q = a$, és $yb \leq \beta$. Itt nem állhat szigorú egyenlőtlenség, hiszen akkor $ax < \beta$ is következne (4.24)-ből, holott $ax \geq \beta$ következik. Tehát $yb = \beta$.

Belátjuk, hogy $y_1 = 0$. Valóban, a 4.3.5 lemma szerint R-nek van olyan x_0 eleme, amelyre $Qx_0 \ll b_1$. Mivel $x_0 \in R \subseteq H$, így $\beta = qx_0 = (y_0P + y_1Q)x_0 = y_0(Px_0) + y_1(Px_0) \leq y_0b_0 + y_1b_1 = yb = \beta$. Szükségképpen az egyenlőtlenség egyenlőséggel teljesül, ami $Qx_0 \ll b_1$ miatt azt jelenti, hogy $y_1 = 0$. •

Feladat 4.3.1 *Igazoljuk, hogy ha egy nemüres poliéder minimális leírásában lévő $px = \beta$ egyenlőséget helyettesítjük a $px \leq \beta, -px \leq -\beta$ egyenlőtlenségekkel, akkor ezek lényegesek.*

Jelölje $Q^=_F$ a Q azon sorai által alkotott részmátrixot, amelyeknek megfelelő egyenlőtlenséget az F minden eleme egyenlőséggel teljesíti. Ezeket hívjuk **implicit F-egyenlőségeknek**.

Lemma 4.3.7 *Tegyük fel, hogy R egy minimális* (4.24) *rendszerrel van megadva. Legyen F az R-nek egy oldala. Legyen Q' a Q-nak egy $m' \times n$-es részmátrixa. A Q' által meghatározott $F' := \{x \in R : Q'x = b'_1\}$ oldal akkor és csak akkor ugyanaz mint F, ha Q' része $Q^=_F$-nek és $r(Q') = r(Q^=_F)$.*

Biz. Tegyük fel először, hogy $F = F'$. Miután F-nek a 4.3.5 lemma alapján van olyan eleme, amely minden valódi F-egyenlőtlenséget szigorúan teljesít, ezért a $Q'x \leq b'_1$ minden sora implicit F-egyenlőség, azaz Q' része $Q^=_F$-nek. Ha indirekt $r(Q') < r(Q^=_F)$, akkor van olyan x', amelyre $Q'x' = 0$, és a $Q^=_F$ valamely ${}_iq$ sorára ${}_iqx' < 0$. Ekkor kicsiny pozitív ε-ra és az F' valamely z elemére a $z + \varepsilon x'$ pont F'-ben van, de szigorúan teljesíti a ${}_iqx \leq b(i)$ egyenlőtlenséget, ellentmondásban ennek implicit F-egyenlőség voltával.

A fordított irányhoz tegyük fel, hogy Q' részmátrixa $Q^=_F$-nek és $r(Q') = r(Q^=_F)$. Mivel F minden z elemére $Q^=_F z = b^=_F$, így $Q'z = b'$ és ezért $F \subseteq F'$. Az $F' \subseteq F$ tartalmazás igazolásához legyen $x' \in F'$ és legyen $z \in F$. $r(Q') = r(Q^=_F)$ miatt az $Q^=_F$ tetszőleges ${}_iq$ sorához létezik y, amelyre ${}_iq = yQ'$. Ekkor ${}_iqx' = (yQ')x' = y(Q'x') = yb' = y(Q'z) = (yQ')z =_i qz = b(i)$, tehát $Q^=_F x' = b^=_F$, azaz $x' \in F$, és így $F' \subseteq F$. •

Most már abban a helyzetben vagyunk, hogy pontosan megmondjuk a bázis-megoldások geometriai jelentését.

TÉTEL 4.3.8 *A $Qx \leq b$ egyenlőtlenség-rendszer egy megoldása akkor és csak akkor bázis-megoldás, ha eleme az $R = \{x : Qx \leq b\}$ poliéder egy minimális oldalának.*

Biz. Tekintsük az $R = \{x : Qx \leq b\}$ poliéder egy minimális F oldalát. Amint láttuk, F affin altér, ezért $F = \{x : Q^=_F x = b^=_F\}$. Az F minimalitásából következik, hogy F minden x elemére $Q^<_F x \ll b^<_F$. Tehát F tetszőleges z elemére éppen a $Q^=_F x \leq b^=_F$ rendszer az, amelynek tagjait a z egyenlőséggel teljesíti.

Állítjuk, hogy $r(Q) = r(Q^=_F)$. Valóban, ha a ${}_jq$ sor lineárisan független volna $Q^=_F$ soraitól, akkor volna olyan x', amelyre $Q^=_F x' = 0$, ${}_jqx' = 1$. De ekkor egyrészt $f + \lambda x'$ tetszőleges λ-ra teljesíti a $Q^=_F x = b^=_F$ egyenlőséget, azaz benne van F-ben, és így R-ben is, másrészt viszont alkalmas λ-ra $f + \lambda x'$ nem teljesíti az ${}_jqx \leq b(j)$ egyenlőtlenséget, azaz még sincs R-ben. Ez az ellentmondás mutatja, hogy $r(Q) = r(Q_F^=)$, tehát az F minden eleme valóban bázis-megoldás.

Legyen most z bázis-megoldás és tekintsük a $Qx \leq b$ rendszernek a z által egyenlőséggel teljesített $Q_z^= x \leq b_z^=$ részét. Tekintsük az $F_z := \{x : Qx \leq b, Q_z^= x = b_z^=\}$ oldalát R-nek. Belátjuk, hogy ez minimális oldal. Ha indirekt létezne valódi oldala, akkor az $Qx \leq b$ rendszerben volna olyan $qx \leq \beta$ egyenlőtlenség, amelyre $qz < \beta$, de F_z valamely x' elemére $qx' = \beta$. Miután z bázis-megoldás, $r(Q) = r(Q_z^=)$, és így létezik egy y', amelyre $q = y'Q_z^=$. De akkor $0 = \beta - \beta > qx' - qz = q(x' - z) = (y'Q_z^=)(x - z) = y'(Q_z^= x) - y'(Q_z^= z) = y'b' - y'b' = 0$, amely ellentmondás mutatja, hogy F_z-nek nem lehet valódi oldala. •

4.3.2 Dimenzió, lapok

Egy altér **dimenziója** a belőle kiválasztható lineárisan független elemek maximális száma. A Z affin altér, definíció szerint, egy altér eltoltja. Az első fejezetben szerepelt, hogy az affin altér előállítható $Z = \{x : Qx = b\}$ alakban és az altér, amelynek eltolásából Z keletkezett egyértelmű, éspedig $\{x : Qx = 0\}$ (ami tehát Z karakterisztikus altere). A Z affin altér **dimenziója** az őt meghatározó altér dimenziója. Például a pont nulla dimenziós, az egyenes egy dimenziós, stb. Affin alterek metszete is affin altér, feltéve, hogy nem üres. Az is rögtön látszik, hogy minden affin altér hipersíkok metszete (egy **hipersík** valamely $qx = \beta$ egyenlet megoldás halmaza.)

Az $\{x : Qx = 0\}$ altér tekinthető a Q sorai által generált altér ortogonális kiegészítő alterének, így ennek dimenziója $n - r(Q)$. Ebből kapjuk, hogy az $\{x : Qx = b\}$ affin altér dimenziója $n - r(Q)$. Egy tetszőleges $\mathbf{R}^n$-beli R nemüres halmaz **dimenzióján** az R-t tartalmazó legszűkebb affin altér dimenzióját értjük. (A legszűkebb affin altér nem más, mint az R-t tartalmazó affin elterek metszete.)

TÉTEL 4.3.9 *Tegyük fel, hogy a nemüres R poliédert definiáló* (4.24) *rendszer minimális. Ekkor a legszűkebb R-t tartalmazó affin altér $Z := \{x : Px = b_0\}$*

***Biz.** Nyilvánvaló, hogy $R \subseteq Z_R$. Tegyük fel indirekt, hogy létezik egy Z_R-nél szűkebb affin altér, amely tartalmazza R-t. Ekkor van olyan $H := \{x : qx = \beta\}$ hipersík, amelyre $R \subseteq H$, $Z_R \not\subseteq H$. A 4.3.6 tétel szerint $R \subseteq H$ azt jelenti, hogy létezik y_0, amelyre $q = y_0 P$ és $\beta = y_0 b_0$, amiből következik, hogy bármely olyan x'-re, amelyre $Q^= x' = b^=$ teljesül, fennáll $qx' = \beta$, azaz $Z_R \subseteq H$, ellentmondás. •

Következmény 4.3.10 *Az $R = \{x : Qx \leq b\}$ nemüres poliédert tartalmazó legszűkebb affin altér $Z_R := \{x : Q^= x = b^=\}$, ahol $Q^=$ jelöli az implicit egyenlőségeknek megfelelő részmátrixot. A poliéder dimenziója $n - r(Q^=)$.* •

A tételt alkalmazhatjuk a poliéder egy F oldalára.

Következmény 4.3.11 *Az $R = \{x : Qx \leq b\}$ poliéder egy F oldalát tartalmazó legszűkebb affin altér $\{x : Q_F^= x = b_F^=\}$. Az F oldal dimenziója $n - r(Q_F^=)$. Minimális oldal dimenziója $n - r(Q)$.* •

Egy c vektort illetve az általa meghatározott cx célfüggvényt akkor mondjuk **semlegesnek** (vagy neutrálisnak) az $R =: \{x : Qx \leq b\}$ poliéderre nézve, ha R minden x elemére cx értéke ugyanannyi. A semleges vektorok nyilván alteret alkotnak, melyet jelöljünk S_R-rel.

TÉTEL 4.3.12 *A semleges vektorok S_R altere éppen a Q-t tartalmazó legszűkebb Z_R affin altér A_R karakterisztikus alterének ortogonális kiegészítője. Egy c vektor akkor és csak akkor semleges, ha benne van $Q^=$ sorterében.*

***Biz.** A 4.3.10 tétel szerint $A_R = \{x : Q^= x = 0\}$. Mivel $Q^=$ sortere $\{x : Q^= x = 0\}$ altér ortogonális kiegészítője, a tétel két állítása egymással ekvivalens.

Ha $c = yQ^=$ valamilyen y-ra, azaz c benne van A_R-ben, akkor nyilván R minden x elemére $cx = (yQ^=)x = y(Q^= x) = yb^=$. Mivel a jobboldal független x-től, c valóban semleges. Fordítva, ha c nincs benne $Q^=$ sorterében, akkor létezik egy olyan x', amelyre $Q^= x' = 0$ és $cx' \neq 0$. Az R egy x_0 elemére és kicsiny $\lambda > 0$ számra $x_0 + \lambda x'$ benne van R-ben. Mivel $cx' \neq 0$, kapjuk, hogy $cx_0 \neq c(x_0 + \lambda x')$, azaz c nem semleges. •

Egy poliéder két csúcsát nevezzük **szomszédosnak**, ha az összekötő szakaszuk a poliéder oldala. (Három dimenzióban ez annak felel meg, hogy az összekötő szakaszuk a poliédernek éle.)

TÉTEL 4.3.13 *Az $R = \{x : Qx \leq b\}$ csúcsos poliéder u és v csúcsaira a következők ekvivalensek.*
(i) u és v szomszédosak.
(ii) A Q azon soraiból alkotott $Q_{uv}^=$ részmátrix, melyeknek megfelelő egyenlőtlenségeket mind u, mind v egyenlőséggel teljesíti, $n - 1$ rangú.
(iii) Léteznek Q-nak olyan Q_u és Q_v $n \times n$-es nemszinguláris részmátrixai, amelyekre u a $Q_u x = b_u$ rendszer, míg v a $Q_v x = b_v$ rendszer egyértelmű megoldásai, és amelyeknek $n - 1$ soruk közös.

*__Biz.__ $(i) \to (ii)$ Legyenek u és v szomszédosak, vagyis az összekötő F szakaszuk oldal. Ekkor F előáll $F = \{x : Qx \le b, Q_F^= x = b_F\}$ alakban. Mivel szakasz dimenziója 1, a 4.3.11 következményből kapjuk, hogy $r(Q_F^=) = n-1$. Mivel u, v benne vannak F-ben, így $Q_{uv}^=$ tartalmazza $Q_F^=$-t, tehát $r(Q_{uv}^=) \ge n-1$. Itt egyenlőségnek kell állnia, mert ha $r(Q_{uv}^=) = n$ volna, akkor a $Q_{uv}^= x = b_{uv}$ = rendszernek nem lehetne két különböző megoldása (mármint u és v.)
$(ii) \to (iii)$ Válasszunk ki a $Q_{uv}^=$ mátrixnak $n-1$ lineárisan független sorát, majd ezt egészítsük ki a $Q_u^=$ egy sorával egy nemszinguláris Q_u mátrixszá, illetve a $Q_v^=$ egy sorával egy nemszinguláris Q_v mátrixszá. A $Q_u x = b_u$ rendszernek ekkor létezik egyértelmű megoldása és u megoldása, és ugyanígy $Q_v x = b_v$-nek v a megoldása.
$(iii) \to (i)$ Az $n-1$ közös sor által alkotott mátrixot jelöljük Q'-vel. Ekkor az $F := \{x : Qx \le b, Q'x = b'\}$ poliéderben u is, v is benne van, tehát F oldala R-nek, melynek dimenziója 1, vagyis F az u-t és v-t összekötő szakasz. •

Fentebb láttuk, hogy az implicit egyenlőségek éppen az R-t tartalmazó legszűkebb affin alteret határozzák meg. Mi a valódi egyenlőtlenségek geometriai tartalma?

TÉTEL 4.3.14 *Tegyük fel, hogy a $\{Px = b_0,\ Qx \le b_1\}$ rendszer minimális, és hogy az R megoldás halmaz nemüres. Ekkor az R-t tartalmazó legszűkebb affin altér $\{x : Px = b_0\}$. Továbbá egy-egy értelmű kapcsolat áll fenn az R lapjai és a $Qx \le b_0$ egyenlőtlenségei között, azaz $\{x \in R :\ {}_iqx = b_1(i)\}$ lapot alkot, és minden lap előáll ilyen alakban.*

Biz. Először tegyük fel, hogy F lapja R-nek. Ekkor $F = \{x \in R : Q_F^= x = b_F^=\}$. Mivel $F \subset R$, létezik egy $x_0 \in R - F$ pont. Ez az $Q_F^= x \le b_F^=$ rendszer egyik $qx \le \beta$ egyenlőtlenségét szigorúan teljesíti: $qx_0 < \beta$. Tehát $qx \le \beta$ a $Qx \le b$ rendszernek valódi egyenlőtlensége, így $F' := \{x \in R : qx = \beta\}$ valódi oldala R-nek, amely tartalmazza F-t, és miután F tartalmazásra nézve maximális oldal, $F = F'$ következik. Tehát R lapja előáll a kívánt alakban.

Megfordítva, tegyük fel, hogy $qx \le \beta$ a $Qx \le b$ rendszernek valódi egyenlőtlensége. Jelölje $Q'x \le b'$ azt a részét $Q^< x \le b^<$-nak, amelyet az $qx \le \beta$ kihagyásával kapunk. A feltevés szerint $qx \le \beta$ lényeges, így létezik $x_2 \in R$, amelyre $Q'x_2 \le b'$ és $qx_2 > \beta$. Másrészt láttuk már, hogy létezik olyan $x_1 \in R$, amelyre $Q^< x_1 \ll b^<$. Az x_1 és x_2 alkalmas konvex kombinációját véve kapjuk, hogy létezik olyan $x_0 \in R$, amelyre $Q'x_0 < b'$ és $qx_0 = \beta$. Ebből következik, hogy $F := \{x \in R : qx = \beta\}$ valódi oldal. Mivel pedig egyetlen olyan eredetileg valódi egyenlőtlenség van, amely implicit F-egyenlőség, F maximális valódi oldal. •

Következmény 4.3.15 *Minden valódi oldal lapok metszete.* •

Következmény 4.3.16 *Tegyük fel, hogy mind a $\{Px = b_0,\ Qx \le b_1\}$ rendszer, mind a $\{P'x = b_0',\ Q'x \le b_1'\}$ rendszer minimális. Az ezek által definiált nemüres R és R' poliéderek akkor és csak akkor egyenlők, ha P és P' sortere ugyanaz, továbbá a $Qx \le b_1$ és $Q'x \le b_1'$ rendszer egyenlőtlenségei között egy-egy értelmű kapcsolat van, amelyben az egymásnak megfelelő $qx \le \beta$ és $q'x \le \beta'$ egyenlőtlenségekre fennáll, hogy a $(q, \beta) - (q', \beta')$ vektor benne van a (P, b_0) mátrix sorterében (ami ugyanaz, mint a (P', b_0') mátrix sortere).* •

Feladat 4.3.2 *Egy poliéder minden lapjának dimenziója eggyel kisebb, mint a poliéder dimenziója.*

lin17 2011. október 29.

4.4 A SZIMPLEX ALGORITMUS

4.4.1 Megengedettség

Ebben a szakaszban megismerkedünk a Farkas lemma illetve a dualitás tétellel 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 csúcsait szisztematikusan, javuló sorrendben tekinti át. 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 3.5.7 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:

TÉTEL 4.4.1 *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 $yb \neq 0$. 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

$$y_q b = y_q(Ax_q) = (y_q A)x_q = x_q(i_q) < 0. \tag{4.25}$$

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_r a legnagyobb ilyen indexű. Tehát az r-nél 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_r benne van, de B_{p+1}-ben nincs. Legyen B_q $(q > p)$ egy olyan későbbi bázis, amelyben a_r nincs benne, de B_{q+1}-ben benne van. Ekkor tudjuk, hogy y_q a B_{q+1} minden oszlopára merőleges, kivéve a_r-t, amelyre $y_q a_r < 0$. A második választási szabályból az is következik minden a_r előtti a_i oszlopra (azaz $i < r$-re), hogy $y_q a_i \geq 0$.

Az első választási szabály miatt x_p minden r-nél kisebb indexű komponense nem-negatív. Így tehát az $1 \leq i < r$ indexekre $(y_q a_i)x_p(i) \geq 0$ és $(y_q a_r)x_p(r) > 0$. 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_r-t. Tehát y_q merőleges a B_q-nak a_r utáni oszlopaira, de mivel az r-nél nagyobb indexű oszlopokon a B_q és a B_p bázis megegyezik (itt használva r maximális választását), y_q merőleges B_p minden r-nél nagyobb indexű oszlopára. Így felhasználva (4.25)-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 egészítjuk 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 3.4.7 tételre. A még általánosabb 3.4.6 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 burokában. 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.

Feladat 4.4.1 *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?*

Feladat 4.4.2 *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?*

Feladat 4.4.3 *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 $y_1A \leq c$, akkor készen vagyunk, megtaláltuk a kívánt y-t. Ha $y_1A \not\leq c$, úgy legyen a_j az A mátrix legkisebb indexű oszlopa, amelyre $y_1a_j < c(j)$. Tekintsük a $B_jx = -a_j$ 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 $cx_1 = c_{B_1}x'_1 + c(j) = (y_1B_1)x'_1 + c(j) < (y_1B_1)x'_1 + y_1a_j = (y_1A)x_1 = 0$. 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.

Feladat 4.4.4 *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 $cx < 0$.*

4.4.2 Optimalizálás

Térjünk rá arra, hogy a fenti eljárás hogyan használható a $\min\{cx : Ax = b, x \geq 0\}$ primál lineáris program és $\max\{by : yA \leq c\}$ duális lineáris program megoldására. Jelölje $R := \{x : Ax = b, x \geq 0)$ a primál, $R^* := \{y : yA \leq c\}$ 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 y_0 csúcsát (azaz $yA \leq c$ 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_0 pontja, akkor $x_0 + \lambda x'$ minden pozitív λ-ra R-ben van, így cx nem korlátos alulról. Ekkor tehát az algoritmus futása befejeződik.

Tegyük fel tehát, hogy rendelkezésünkre áll y_0. Jelölje $A_0^=$ az A-nak azon a_i oszlopai által alkotott részmátrixát, amelyekre $y_0a_i = c(i)$, míg a maradék oszlopok részmátrixa legyen $A_0^<$. A fenti eljárással döntsük el, hogy az $\{A_0^= x' = b,\ x' \geq 0\}$ rendszernek létezik-e megoldása. Amennyiben létezik, úgy x'-t nulla komponensekkel kiegészítve R-nek egy olyan x_0 elemét kapjuk, amely teljesíti az optimalitási feltételeket (azaz, ha valamely i-re $x_0(i)$ szigorúan pozitív, akkor $y_0a_i = c(i)$). Ekkor x_0 primál optimum, y_0 duál optimum és az eljárás véget ér.

Ha a szóbanforgó x' nem létezik, akkor a fenti eljárás megtalálja $A_0^=$-nak egy $m-1$ lineárisan független oszlopból álló A'_0 részmátrixát valamint egy olyan y' vektort, amelyekre $y'A_0^= \leq 0, y'A'_0 = 0$ és $y'b > 0$.

(**Figyelem!** Az $\{y'A_0^= \leq 0, y'b \geq 1\}$ rendszer egy bázis-megoldása azt jelenti, hogy y' az $A_0^=$ mátrixnak $m-1$ lineárisan független oszlopára merőleges. Az $yA \leq c$ rendszer egy bázis-megoldása azt jelenti, hogy A-nak van m lineárisan független a_i oszlopa, amelyre $ya_i = c(i)$.)

Amennyiben $y'A_0^< \leq 0$, úgy az adódik, hogy $y'A \leq 0$, $y'b > 0$ 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 $y'A_0^< \not\leq 0$. Válasszuk λ-t a legnagyobb olyan számnak, amelyre $(y_0 + \lambda y')A \leq c$ teljesül, azaz $(y_0 + \lambda y')a_i \leq c(i)$ fennáll az $A_0^<$ mindegyik a_i oszlopára. Vagyis λ a legnagyobb szám, amelyre $\lambda y'a_i \leq c(i) - y_0a_i$ teljesül az $A_0^<$ valamennyi olyan a_i oszlopára, amelyre $a_iy' > 0$. Egyszóval legyen $\lambda_0 := \min\{c(i) - y_0a_i)/y'a_i\}$, ahol a minimum az $A_0^<$ olyan a_i oszlopaira megy, melyekre $y'a_i > 0$.

Legyen $y_1 := y_0 + \lambda_0y'$. A λ_0 választásából adódóan y_1 eleme R^*-nak.

Lemma 4.4.2 *y_1 csúcsa R^*-nak.*

Biz. Azt kell látni, hogy A-nak van m lineárisan független oszlopa, melyekre $y_1a_i = c(i)$. Mindenesetre $y'A'_0 = 0$ miatt A'_0-nak az $m-1$ oszlopa ilyen. Legyen a_j egy olyan oszlop, ahol a λ_0 definíciójában szereplő minimum felvétetik. Ekkor nyilván $y_1a_j = c(j)$, így csak azt kell látnunk, hogy a_j lineárisan független az A'_0 oszlopaitól. De ez valóban így van, hiszen $y'A'_0 = 0$ és $y'a_j \neq 0$. •

Az y_1 tehát valóban csúcsa R^*-nak, és ráadásul y_0-nál jobb csúcsa, hiszen $y'b > 0$ miatt $y_1b > y_0b$. Miután R^*-nak véges sok csúcsa van, az eljárás véges sok iteráció után befejeződik.

A algoritmusban a Farkas lemmára vonatkozó algoritmust szubrutinként használtuk, aminek belsejében persze alkalmazzuk a Bland féle legkisebb index szabályt. A fenti algoritmusban azonban, amikor az y_0-ról áttértünk y_1-re, a szóban forgó a_j oszlop meghatározásánál nem volt szükség a Bland-szabályra.

Az egész eljárás összevonható és elmondható egységes keretben:

Kiindulunk egy y_0 duális csúcsból. Ez meghatározza A-nak valamely $m\times m$-es nemszinguláris A_0 részmátrixát, amelyre $y_0A_0 = c_0$ (ahol c_0 jelöli a c-nek az A_0 oszlopaihoz tartozó részét.) Az $A_0x = b$ egyértelmű megoldását nullákkal kiegészítve kapjuk az $Ax = b$ egy x_0 megoldását.

Ha $x_0 \geq 0$, úgy x_0 primál megoldás és y_0 duál megoldás teljesítik az optimalitási feltételeket, és az algoritmus véget ér. Amennyiben $x_0 \not\geq 0$, úgy legyen i a legkisebb index, amelyre $x_0(i) < 0$. Legyen y' olyan, hogy $y'A_0$ vektornak egy kivételével minden komponense 0 és a kivételes komponens $y'a_i = -1$. Most $y'b = y'(Ax_0) = (y'A)x_0 > 0$.

Amennyiben $y'A \leq 0$, akkor $y_0 + \lambda y'$ minden pozitív λ-ra R^*-ban van és az $(y_0 + \lambda y')b$ célfüggvény-érték a λ-val tart a végtelenbe, azaz a duál probléma nem korlátos (vagy ekvivalens módon a primál poliéder üres). Ha $y'A \not\leq 0$, akkor legyen $\lambda_0 := \min[(c(i) - y_0a_i)/y'a_i]$, ahol a minimum olyan i indexekre megy, ahol $y'a_i > 0$. (Most előfordulhat, hogy $\lambda_0 = 0$). Legyen j a legkisebb index, amelyre a minimum felvétetik. Jelölje A_1 azt a mátrixot, amely A_0-ból keletkezik az a_i kitevésével és az a_j bevételével, és legyen $y_1 := y_0 + \lambda_0 y'$. Iteráljuk az eljárást.

A leírt algoritmus az R^* duális poliéder csúcsain vándorolva haladt az optimális csúcs felé, és csak amikor ezt megtalálta, akkor állította elő a primál feladatnak egy optimumát. Ezért jogos a fenti eljárást duál-szimplex algoritmusnak hívni. Természetesen elkészíthetjük a primál-szimplex változatot is. Ebben az $R := \{x : Ax = b, x \geq 0\}$ primál poliéder egy x_1 csúcsából (azaz egy bázis-megoldásból) indulunk ki. Jelölje A_1^+ az A mátrix azon oszlopaiból álló részmátrixot, amelyekhez tartozó x_1-komponense pozitívak, továbbá c_1^+ a c megfelelő részét. A 4.4.3 és 4.4.4 feladatokban leírt algoritmus segítségével döntsük el, hogy létezik-e olyan y, amelyre $yA \leq c$ és $yA_1^+ = c_1^+$, vagy pedig létezik-e olyan $x' = \ldots$

Az algoritmust kényelmesen mechanikussá lehet tenni. tekintsük a kiinduális B_1 bázist. Legyen $A_1 = (B_1)^{-1}A$ és $x_1 = (B_1)^{-1}b$. Az A_1-ben B_1 helyén természetesen egy egység mátrix szerepel. Az algoritmus tehát először azt kérdezi, hogy x_1 nem-negatív-e és ha igen akkor befejeződik. Ha x_1-nek van negatív komponense, akkor kiválasztja a legkisebb ilyen i indexet. Majd tekinti azt az y_1 vektort, amelyre y_1B_1 Ha ennek inverzével megszorozzuk A-t, akkor a B_1 helyén egy egység mátrix keletkezik ...

2011. október 29.lin18

5. Fejezet

LINEÁRIS POGRAMOZÁS 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 unimoduláris**nak (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 E-nek 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.

TÉTEL 5.1.1 (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. •

Feladat 5.1.1 *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.*

Feladat 5.1.2 *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 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.

Lemma 5.1.2 *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. •

TÉTEL 5.1.3 *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. •

Következmény 5.1.4 *Egy olyan hipergráf, amely egy irányított fa élhalmazán van definiálva és a hiperélek irányított utak, teljesen unimoduláris.* •

Egy hipergráfot **laminháris**nak 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 $M := \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$.

TÉTEL 5.1.5 *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. •

Feladat 5.1.3 *Igazoljuk, hogy az* 5.1.5 *tételben szereplő M mátrix hálózati mátrix!*

Feladat 5.1.4 *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.

Lemma 5.1.6 *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. Az Iránymenti korlátosság című szakaszban megfigyeltük, hogy 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. •

Lemma 5.1.7 *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 4.1.7 tétel szerint van olyan x^* erős bázis-megoldása (5.2) rendszernek, amelyre $cx^* \geq cx'$. A 5.1.6 lemma miatt x^* egészértékű, azaz minden komponense $0, \pm 1$. •

A Farkas lemma szerint a (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.

TÉTEL 5.1.8 *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 a 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.

TÉTEL 5.1.9 *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, a 5.1.6 lemmából a tétel következik. •

Az alábbi tételek ugyanígy következnek a 4.1.10 és 4.2.2 tételekből az 5.1.6 és 5.1.7 lemmák segítségével.

TÉTEL 5.1.10 *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.* •

TÉTEL 5.1.11 *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.

Lemma 5.1.12 *Legyen A teljesen unimoduláris mátrix és* x_0 *egy vektor. Ekkor létezik egy olyan q egészértékű vektor, amelyre* $\lfloor x_0 \rfloor \leq q \leq \lceil x_0 \rceil$ *és* $\lfloor Ax_0 \rfloor \leq Aq \leq \lceil Ax_0 \rceil$*. 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 $\lfloor x_0 \rfloor \leq z \leq \lceil x_0 \rceil$ és $\lfloor Ax_0 \rfloor \leq Az \leq \lceil Ax_0 \rceil$ 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* $x_0 : S \to \mathbf{R}$ *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.*

TÉTEL 5.1.13 *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* ($j = 1, 2, \ldots, m$)*, az első i oszlop elemeinek összege* ($i = 1, 2, \ldots, n$).

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 ($2 \leq i \leq m$) 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. •

TÉTEL 5.1.14 *Egy* $x_1, \ldots x_n$ *sorozat elemeinek létezik olyan* $z_1, \ldots, z_n$ *kerekítése, hogy minden* $1 \leq i \leq j \leq n$ *indexre a* $z_i + \ldots + z_j$ *összeg kerekítése az* $x_i + \ldots + x_j$ *összegnek.*

Biz. A $\{v_1, \ldots, v_n\}$ alaphalmazon tekintsük azt a hipergráfot, melynek élei a $\{v_i, \ldots, v_j\}$ típusú halmazok minden $1 \leq i \leq j \leq n$ index párra. Amint már láttuk, ez a hipergráf teljesen unimoduláris, így az 5.1.3 lemma alkalmazható. •

Egyenletes színezések

A teljesen unimoduláris mátrixok egy másik érdekes alkalmazása hipergráfok egyenletes színezésével foglalkozik.

TÉTEL 5.1.15 *Legyen A TU-mátrix, b egész vektor, k pozitív egész. Legyen z olyan egész vektor, amelyre $Az \leq kb$. Ekkor z előáll olyan $z_1, z_2, \ldots, z_k$ egész vektorok összegeként, melyekre $Az_i \leq b$.*

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.

Következmény 5.1.16 *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 $\lceil e_n a/k \rceil$.

TÉTEL 5.1.17 *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

Következmény 5.1.18 *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.

Következmény 5.1.19 *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.* •

Feladat 5.1.5 *Egyszerű mohó algoritmus megadásával közvetlenül bizonyítsuk be az* 5.1.19 *következményt.*

Az 5.1.17 tétel páros gráfokra vonatkozó következményeit a 5.2.6 tételben tárgyaljuk.

TU-mátrixok jellemzése

Bemutatjuk a TU-mátrixoknak egy hasznos jellemzését. A 5.1.17 tételben láttuk már, hogy TU-mátrixok oszlopai egyenletesen k-színezhetők. Kérdés, hogy ez a tulajdonság mennyire a csak TU-mátrixok sajátja. Az alábbi tétel szerint, már az egyenletes 2-színezhetőségből is következik a TU-ság. Az oszlopok egyenletes 2-színezhetőségének azt az ekvivalens definícióját használjuk, amely egy olyan z $\{\pm 1\}$-es vektor létezését követeli, amelyre Qz $\{0, \pm 1\}$-es vektor.

TÉTEL 5.1.20 (Ghouila-Houri) (ejtsd: Gujla-úri) *Egy Q mátrix akkor és csak akkor teljesen unimoduláris, ha oszlapainak bármely részhalmaza egyenletesen 2-színezhető.*

Biz. TU-mátrix egyenletes k-színezhetőségét már láttuk korábban, így csak a fordított iránnyal foglalkozunk. Megjegyezzük, hogy ha kihagyjuk Q néhány sorát, akkor Q oszlopainak egy egyenletes 2-színezése automatikusan egyenletes 2-színezése a maradéknak. Így Q mindenesetre $\{0, \pm 1\}$-es mátrix.

Tegyük fel, hogy Q minimális méretű ellenpélda. Ekkor tehát Q nem TU, de minden valódi részmátrixa az. Vagyis Q maga egy négyzetes mátrix, amelyre $K := \det Q \not\in \{0, \pm 1\}$. Így Q nem egyelemű és Q minden valódi aldeterminánsa $\{0, \pm 1\}$ értékű. Tekintsük a Q mátrix Q^{-1} inverzét. A Cramer szabály szerint

$$Q^{-1} \text{ minden nem-nulla eleme } \pm 1/K \text{ alakú.} \tag{5.4}$$

Legyen q_1^* a Q^{-1} első oszlopa és jelölje R azon j indexek halmazát, melyekre $q_1^*(j) \neq 0$. Jelölje ${}_iq$ a Q mátrix i-dik sorát és $e_1 = (1, 0, \ldots, 0)$ az első egységvektort. Mivel minden $i \geq 2$ indexre ${}_iqq_1^* = 0$, ezért (5.4) miatt a ${}_iq$ sornak páros sok olyan q_{ij} nem-nulla eleme van, amelyre $j \in R$.

A feltevés szerint a Q mátrix R-hez tartozó oszlopai egyenletesen 2-színezhetők, vagyis létezik egy olyan z $\{0, \pm 1\}$-es vektor, amelyre Qz $\{0, \pm 1\}$ értékű és $z(j)$ pontosan akkor nem nulla, ha $j \in R$. Az előbbi paritási megfigyelés miatt minden $i \geq$ 2-re ${}_iqz = 0$. De ekkor ${}_1qz \neq 0$, mert különben $Qz = 0$ volna, ellentmondásban $|\det Q| \geq$ 2-vel. z esetleges negálásával feltehetjük, hogy ${}_1qz = 1$, vagyis $Qz = e_1$, és emiatt $z = q_1^*$, ellentmondásban a (5.4) tulajdonsággal. •

2011. október 29. file: lin26

5.2 HÁLÓZATI OPTIMALIZÁLÁS ÉS LINEÁRIS PROGRAMOZÁS

Ebben a részben áttekintjük át 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 1.5.1 tételét:

TÉTEL 5.2.1 (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:

$$\max\{e_q x : Ax \leq e_p, x \geq 0\}, \tag{5.5}$$

$$\min\{e_p y : yA \geq e_q, y \geq 0\}. \tag{5.6}$$

Az 5.1.9 tétel szerint mindkét programnak az optimuma egész vektoron felvétetik. Jelöljük ezeket rendre x_0-lal és y_0-lal. (5.5) minden egészértékű megoldása $0-1$ értékű, és rögtön látszik, hogy (5.6) minden optimális egészértékű megoldása is $0-1$ értékű. Legyen M azon élek halmaza, melyeken x_0 az 1 értéket vesz fel, és legyen L azon pontok halmaza, amelyeken y_0 egyet 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 1.5.10 tételt adja meg:

TÉTEL 5.2.2 *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:

TÉTEL 5.2.3 *A G páros gráf A incidencia mátrixával felírt*

$$\{x : Ax \leq e_p, x \geq 0\} \tag{5.7}$$

poliéder egész, amelynek csúcsai pontosan a gráf párosításainak incidencia vektorai. •

Egy gráf **párosítás politop**ja a párosítások incidencia vektorainak konvex burka. A 3.3.4 tétel szerint tetszőleges politop (korlátos) poliéder, azaz felírható egy lineáris egyenlőtlenség-rendszer megoldáshalmazaként. A 5.2.3 tétel a (5.7) 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 a (5.7) 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.

TÉTEL 5.2.4 (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 1.5.4 tételét, sőt most már belefoglaljuk azt az esetet is, amikor a súlyfüggvény nem egész.

TÉTEL 5.2.5 (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 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 fokszámkorlátozott 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 kiegészítva 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 szállítás 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.*

TÉTEL 5.2.6 *Egy $G = (S,T;E)$ páros gráf éleit meg lehet k színnel úgy színezni, hogy minden v csúcsra és mindegyik j színre ($j = 1,\ldots,k$) 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.* •

A lineáris programozási megközelítés eredményességét egy kevésbé közismert tételen is bemutatjuk.

TÉTEL 5.2.7 (Folkman és Fulkerson) *Egy $G = (S,T;E)$ páros gráfban akkor és csak akkor létezik l darab élidegen k élű párosítás, ha*

$$i_G(Z) \geq l(k + |Z| - |U|) \tag{5.8}$$

fennáll $U := S \cup T$ minden Z részhalmazára, ahol $i_G(Z)$ jelöli a Z által feszített élek számát.

Biz. Mivel egy M párosítás legfeljebb $|U| - |Z|$ olyan élt tartalmaz, amelynek legalább egyik végpontja nincs Z-ben, így legalább $|M| - (|U| - |Z|)$ darab $|Z|$ által feszített élt tartalmaz. Így, ha létezik l darab k élű párosítás, akkor Z legalább $l(k + |Z| - |U|)$ élt feszít, vagyis (5.8) szükséges.

Az elegendőséghez jelölje A a páros gráf pont-él incidencia márixát, p a csúcsok számát, q az élek számát. Az xe_q és ye_q skalárszorzatot szemléletesebben $x(E)$-vel illetve $y(E)$-vel jelöljük, míg a πe_p-t $\pi(U)$-val. Tekintsük a

$$\max\,\{x(E):\ x \geq 0,\ Ax \leq le_q,\ I_q x \leq e_q\} \tag{5.9}$$

primál és a

$$\min\,\{l\pi(U) + y(E):\ (\pi, y) \geq 0,\ \pi A + yI_q \geq e_q\} \tag{5.10}$$

duális lineáris programot, ahol $\pi : U \to \mathbf{R}_+$ az A sorainak megfelelő duális változók vektora, míg $y : E \to \mathbf{R}_+$ az I_m sorainak megfelelőké. Az A mátrix TU-sága miatt mind a primál, mind a duál optimum egész vektoron felvétetik, sőt $0-1$ vektoron is, hiszen a primál feltételek között explicit szerepel a $0 \leq x \leq 1$ kikötés, míg a duálisban a jobboldalon azonosan 1 áll, így egy optimális (π, y) vektor minden komponense legfeljebb 1.

Amennyiben a (dualitás tétel miatt létező) közös optimum-érték legalább kl, úgy a $0-1$ értékű optimális primál vektor 1 értékű komponensei egy olyan legalább kl élű $G' = (U, E')$ részgráfot határoznak meg, amelyben minden pont foka legfeljebb l. Élek esetleges törlésével elérhetjük, hogy G' pontosan kl darab élből álljon. Kőnig élszínezési tétele miatt E' felbomlik l darab párosításra, amelyek szükségképpen pontosan k eleműek.

Tételezzük most fel, hogy a közös optimum értéke kisebb, mint kl. Ekkor létezik egy $0-1$ értékű (π, y) duális optimális megoldás, amelyre $l\pi(U) + y(E) < kl$. Jelölje Z a gráf azon v pontjainak halmazát, melyekre $\pi(v) = 0$. A duális feltételek miatt minden Z által feszített e élre $y(e) = 1$, és ezért $i_G(Z) \leq y(E)$. Miután $\pi(U) = |U| - |Z|$, így $l(|U| - |Z|) + i_G(Z) \leq l\pi(U) + y(E) < kl$, ellentmondásban a (5.8) feltétellel. •

5.2.3 Megengedett potenciálok, legolcsóbb utak

Legyen $D = (V, A)$ irányított gráf, melynek $(0, 1, -1)$-es incidencia mátrixát jelölje Q. Egy $\pi : V \to \mathbf{R}$ vektort akkor neveztünk a $c : A \to \mathbf{R}$ 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 $x : A \to \mathbf{R}$ vektor pedig pontosan akkor áram, ha $Qx = 0$. Megmutatjuk, hogy a megengedett potenciál létezésére vonatkozó **??** tétel rögtön következik a Farkas lemma TU-mátrixokra vonatkozó élesebb alakjából. Az alábbi tétel a **??** tétel más szövegezéssel.

TÉTEL 5.2.8 *Adott $c : A \to \mathbf{R}$ költség-függvényre akkor és csak akkor létezik olyan $\pi : V \to \mathbf{R}$ 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 a 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 1.3.9 tétel is.

TÉTEL 5.2.9 *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 a 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, mással a $Qx \leq 0$ egyenlőtlenség-rendszer megoldáshalmaza pontosan az áramok halmaza. (Ezt kellett bizonyítani a 1.6.1 gyakorlat (a) részében.)

TÉTEL 5.2.10 *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 a 5.1.9 tételt alkalmazhatjuk. •

Hasonló megfontolással kapjuk:

TÉTEL 5.2.11 *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 $\{x \in \mathbf{R}^A : 0 \leq x \leq g, \varrho_x(v) = \delta_x(v)$ 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.

TÉTEL 5.2.12 (Hoffman, 1960) *A $D = (V, A)$ digráfban adott $f \leq g$ kapacitásfüggvényekre vonatkozólag akkor és csak akkor létezik megengedett áram, ha*

$$\varrho_f(X) \leq \delta_g(X) \text{ minden } X \subseteq V \text{ halmazra.} \tag{5.11}$$

Továbbá, ha f és g egészértékűek és (5.11) *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. A 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 a (5.11) feltételnek. •

5.2.5 Minimális költségű áramok és folyamok

Tekintsük most a költséges áram problémát, azaz adott $c : A \to \mathbf{R}$ költségfüggvény esetén keressünk minimális költségű megengedett áramot, más szóval, oldjuk meg a

$$\min\{cx : Qx = 0, f \leq x \leq g\} \tag{5.12}$$

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.

TÉTEL 5.2.13 *Feltéve, hogy létezik megengedett áram, a következők ekvivalensek.*
(a) *cx alulról korlátos,*
(b) *nincs negatív összköltségű irányított kör D'-ben,*
(c) *létezik egy olyan $\pi : V \to \mathbf{R}$ függvény, amelyre*

$$\pi(v) - \pi(u) \leq c(uv), \text{ ha } uv \in A \text{ és } g(uv) = \infty, \tag{5.13}$$

$$\pi(v) - \pi(u) \geq c(uv), \text{ ha } uv \in A \text{ és } f(uv) = -\infty. \tag{5.14}$$

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 5.2.8 tétel miatt létezik egy $\pi : V \to \mathbf{R}$ függvény, amelyre $uv \in A, g(uv) = \infty$ esetén (amikor is $uv \in A'$) $\pi(v) - \pi(u) \leq c'(uv) = c(uv)$ azaz (5.13) 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 (5.14) fennáll.

(c)→(b) 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 (5.13) azzal ekvivalens, hogy $c_\pi(uv) > 0$ esetén $g(uv) < \infty$, míg (5.14) 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 $c_x(uv) := -c(vu)$. A 5.1.11 tételt specializálva kapjuk a következőt.

TÉTEL 5.2.14 *Adott x megengedett áram esetén a következők ekvivalensek.*
(a) *x optimális megoldása a* (5.12) *minimális költségű megengedett áram feladatnak,*
(b) *D_x-ben nem létezik negatív összköltségű irányított kör,*
(c) *létezik egy olyan $\pi : V \to \mathbf{R}$ 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. •

Feladat 5.2.1 *A* 5.2.13 *tétel fenti direkt bizonyításának mintájára adjuk meg az* 5.2.14 *tétel közvetlen bizonyítását is.*

Feladat 5.2.2 *Fogalmazzuk meg és bizonyítsuk be a* 5.2.13 *és a* 5.2.14 *tételek megengedett potenciálokra vonatkozó ellenpárját.*

Az áramokra megfogalmazott optimalitási feltételt könnyen átvihetjük folyamokra.

TÉTEL 5.2.15 *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbf{R}_+$ kapacitásfüggvény és a $c : A \to \mathbf{R}$ 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:*

$$\pi(v) - \pi(u) < c(uv) \Rightarrow z(uv) = 0, \tag{i}$$

$$\pi(v) - \pi(u) > c(uv) \Rightarrow z(uv) = g(uv). \tag{ii}$$

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 a 5.2.14 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 (1.7.15 tétel). Megmutatjuk, hogy a háttérben most is az 5.1.9 tételben megfogalmazott TU-mátrixokra vonatkozó élesített dualitás tétel áll.

TÉTEL 5.2.16 *A $D = (V, A)$ irányított gráf élhalmazán adott a $g : A \to \mathbf{R}_+$ kapacitásfüggvény és a $c : A \to \mathbf{R}$ költségfüggvény. A k nagyságú megengedett folyamok költségének minimuma egyenlő a*

$$k\pi(t) + \sum [c_\pi(uv)g(uv) : uv \in A, c_\pi(uv) < 0] \tag{5.15}$$

érték maximumával, ahol a maximum az összes $\pi : V \to \mathbf{R}$ 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 $\min\{cx : x \geq 0, Qx = (-k, +k, 0, 0, \dots, 0), x \leq g\}$ 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. A 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 a (5.15) 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 $f\le g$. Legyen továbbá $c=(c_F,c_N)$ egy olyan vektor, amelyre $c_F=0$. Jelölje az F-hez tartozó $(0,\pm1)$-es hálózati mátrixot B, míg a D digráf $(0,\pm1)$-es pont-él incidencia mátrixát Q_D. Legyen továbbá $x=(x_F,x_N)$. Tekintsük a $\max\{c_Nx_N : f_F\le Bx_N\le g_F, f_N\le x_N\le g_N\}$ lineáris programot. Belátjuk, hogy ez ekvivalens a $\max\{cx : Q_Dx=0, f\le x\le g\}$ maximális költségű áram feladattal.

Amennyiben $x=(x_F,x_N)$ áram (azaz $Q_Dx=0$), úgy könnyen látszik, hogy $x_F=Bx_N$, és persze $cx=c_Nx_N$. Emiatt $f\le x\le g$ ekvivalens a $f_F\le Bx_N\le g_F, f_N\le x_N\le g_N$ feltételekkel. Fordítva, tegyük fel, hogy x_N kielégíti ezen utóbbi egyenlőtlenségeket. Minden $e\in N$ nem-fa élhez legyen χ_e az $(1,a_e)$ vektor, ahol a_e az A mátrix e-hez tartozó oszlopa. (Másszóval, χ_e az e élhez tartozó C_e alapkör $0,\pm1$-es incidencia vektora.) Ekkor persze χ_e áram, és így az $x := \sum[x_N(e)\chi_e : e\in N]$ is áram, méghozzá olyan, hogy $x(e)=x_N(e)$, ha $e\in N$. Látható, hogy $f_F\le Bx_N\le g_F$ azzal ekvivalens, hogy $f_F(e)\le x(e)\le g_F(e)$ 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 megoldást, ami éppen az eredeti potenciál probléma megoldása.

2011. október 29. file: lin27

5.3 EGÉSZ POLIÉDEREK

5.3.1 Oldalak

Foglaljuk össze a poliéder oldalainak néhány tulajdonságát. Egy $R = \{x : Qx \leq b\}$ (nemüres) poliéder F oldalán az R-nek egy olyan F részhalmazát értjük, amelyhez létezik egy c vektor és egy δ szám úgy, hogy minden $x \in F$ esetén $cx = \delta$ és minden $x \in R - F$-re pedig $cx < \delta$. 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. E definíció előnye, hogy nem függ a poliéder konkrét megadásától, hátránya viszont, hogy olyan szemléletesnek tűnő tulajdonságok nem látszanak belőle, mint hogy az oldal is poliéder vagy hogy R oldalának oldala az R-nek is oldala. A következő jellemzés ezt a nehézséget küszöböli ki.

TÉTEL 5.3.1 *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. •

Az R poliéder maga is oldal (pl. $c = 0$ célfüggvényre az R minden pontja maximalizálja cx-t, vagy másként, amikor semmilyen egyenlőtlenséget nem kötünk meg egyenlőségként.) A poliédernek egy önmagától különböző oldalát **valódi oldal**nak nevezzük. Egy tartalmazásra nézve maximális valódi oldalt **lap**nak hívunk. Fontos szerepet játszanak a (tartalmazásra nézve) **minimális oldal**ak, vagyis az olyan oldalak, melyek valódi részhalmazként már nem tartalmaznak más oldalt. Az R poliéder egy $R = \{x : Px = b_0, Qx \leq b_1\}$ leírásában szereplő $qx \leq \beta$ egyenlőtlenségről azt mondtuk, hogy **lényeges**, ha kihagyása megváltoztatja (bővíti) a poliédert. Az egyenlőtlenség **igazi**, ha egyenlőséggel történő cseréje megváltoztatja (szűkíti) a poliédert.

TÉTEL 5.3.2 *Egy R nemüres poliédernek akkor és csak akkor nincs valódi oldala, ha R affin altér.*

Biz. Ha $R = \{x : Qx = b\}$ affin altér, úgy az 5.3.1 tétel szerint nincs valódi oldala.

Megfordítva, tegyük fel, hogy az R poliédernek nincs valódi oldala. Tekintsük a poliédernek egy olyan $R = \{x : Px = b_0, Qx \leq b_1\}$ megadását, amelyben minden egyenlőtlenség igazi és lényeges. (Ilyen persze van, hiszen R egy tetszőleges leírásából kiindulva egymás után kihagyhatjuk az aktuálisan lényegtelen egyenlőtlenségeket, majd az implicit egyenlőségeket explicitté alakíthatjuk). Azt látjuk be, hogy Q üres, és így R valóban affin altér. Tegyük fel indirekt, hogy létezik egy $qx \leq \beta$ igazi és lényeges egyenlőtlenség. Jelölje Q' a q sor kihagyásával Q-ból keletkező részmátrixot és b'_1 a megfelelő jobboldalt. Ekkor egyrészt van olyan $x' \in R$ pontja, amelyre $Px' = b_0$, $Q'x' \leq b'_1$ és $qx' > \beta$, másrészt R-nek van olyan x'' pontja, melyre $qx'' < \beta$. Így az $x'x''$ szakasznak van olyan z pontja, amelyre $qz = \beta$ és $z \in R$, vagyis $\{x : Px = b_0, Q'x \leq b'_1, qx = \beta\}$ valódi nemüres oldala R-nek, ellentmondásban a feltevéssel, hogy ilyen oldal nem létezik. •

Következmény 5.3.3 *Egy poliéder minimális oldala affin altér.*

Biz. Az 5.3.1 tétel miatt az R poliéder egy oldalának oldala R-nek is oldala, így a minimális oldal olyan poliéder, amelynek már nincs valódi oldala. Alkalmazzuk az 5.3.2 tételt. •

5.3.2 Egész megoldások

Egy poliédert akkor nevezünk egésznek, ha minden oldala tartalmaz egész pontot. Ez avval ekvivalens, hogy minden minimális oldala tartalmaz egész pontot, ami abban a speciális esetben, amikor a poliéder csúcsos (vagy ekvivalensen egyenes-mentes) azzal egyenértékű, hogy minden csúcs egész. Először azt vizsgáljuk meg, hogy egy affin altér mikor tartalmaz egész pontot. Az alábbi eredmény érdekes analógiát mutat a Farkas lemmával, amely arra adott jellemzést, hogy egy affin altérnek mikor nincs nemnegatív eleme.

TÉTEL 5.3.4 *Legyen A egész mátrix és b egész vektor. Az $Ax = b$ rendszernek akkor és csak akkor van egész megoldása, ha minden y vektorra, amelyre yA egész, yb is egész. Ha van olyan y, amelyre yA egész, de yb nem, akkor van ilyen nemnegatív y is.*

Biz. Ha létezik egész x megoldás, és valamely y-ra yA egész vektor, akkor $(yA)x = y(Ax) = yb$, azaz yb is egész.

Az ellenkező irányú következtetés bizonyításához feltehetjük, hogy az $Ax = b$-nek létezik egyáltalán megoldása, mert ha nem létezne, akkor van olyan y, hogy $yA = 0$ és $yb \neq 0$. De akkor y úgy is választható, hogy yb nem egész.

Az is feltehető, hogy az A sorai lineárisan függetlenek, mert ha valamelyik sor lineárisan függ a többiektől, akkor ezt kihagyhatjuk (b megfelelő komponensével együtt), és az új rendszernek, ha van megoldása, akkor az az eredetinek is megoldása, ha pedig nincs, akkor indukcióval van olyan y', amelyre $y'A'$ egész, de $y'b'$ nem. Ha most a kihagyott komponenst 0-nak vesszük, akkor egy olyan y-t kapunk y'-ből, amelyre yA egész, de yb nem az.

Könnyen ellenőrizhető, hogy az eredetivel ekvivalens feladatot kapunk (mind az x, mind az y létezése szempontjából), ha az A egy oszlopát egy másik oszlopához adjuk vagy abból levonjuk. Ezen művelet ismételt alkalmazásával (és esetleges oszlop cserékkel) az A mátrix $[B, 0]$ alakra hozható, ahol B egész háromszög mátrix (azaz a főátló elemei felett minden elem 0) és a főátlóban nem nulla elemek állnak.

Mivel $B^{-1}(B, 0) = (I, 0)$ egész mátrix, a feltevésből következik, hogy $B^{-1}b$ is egész vektor. Így az $x^* := \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ egész vektorra $(B, 0)x^* = b$, azaz $Ax^* = b$.

A tétel utolsó részéhez tegyük fel, hogy valamely y'-re $y'A$ egész, de $y'b$ nem. Legyen y'' olyan egész vektor, amelyre $y^* := y' + y''$ nemnegatív. Mivel $y''A$ és $y''b$ is egész, így y^*A egész, y^*b nem az. •

Alapvető fontosságú az egész poliédereknek a következő jellemzése.

TÉTEL 5.3.5 *Az $R := \{x : Qx \leq b\}$ nemüres poliéder* (Q, b egész) *akkor és csak akkor egész, ha minden olyan c egész vektorra, amelyre $\{cx : x \in R\}$ felülről korlátos, a* $\max\{cx : x \in R\}$ *szám egész.*

Biz. A feltétel szükségessége nyilvánvaló, így csak az elegendőség bizonyításával foglalkozunk. Tekintsük az R-nek cx-t maximalizáló oldalát és legyen R' ennek egy minimális oldala. Az 5.3.3 következmény miatt R' affin altér, azaz megadható $\{x : Q'x = b'\}$ alakban, ahol Q' az Q bizonyos soraiból álló részmátrix.

A tételhez azt kell bebizonyítani, hogy R' tartalmaz egész pontot. Ha indirekt nem tartalmaz, akkor az 5.3.4 tétel alapján létezik olyan $y' \geq 0$ vektor, amelyre $c' := y'Q'$ egész, de $y'b'$ nem. Ekkor tetszőleges $x \in R'$-re $c'x = y'(Q'x) = y'b'$, és tetszőleges $x \in R$-re $c'x = y'(Q'x) \leq y'b'$, amiből következik, hogy a $\max\{c'x : x \in R\}$ az R' elemein felvétetik és így a maximum értéke $y'b'$. De ez ellentmond a tétel feltevésének, hiszen $y'b'$ nem egész. •

Alkalmazásként megadjuk egy $D = (V, A)$ digráf gyökeresen összefüggő részgráfjainak a leírását. Legyen r_0 adott gyökérpont, egy digráfot akkor mondunk gyökeresen összefüggőnek az r_0-ra nézve, ha minden $X \subseteq V - r_0$ nem-üres részhalmazba lép be él. Közismert, hogy ez azzal ekvivalens, hogy r_0-ból minden pontba vezet irányított út és még azzal, hogy a digráf tartalmaz r_0 gyökerű feszítő fenyőt. Következik, hogy egy élelhagyásra minimális gyökeresen összefüggő digráf maga egy fenyő.

A D digráfhoz rendeljük hozzá a Q mátrixot, amelynek a sorai a D éleinek felelnek meg, az oszlopai pedig a $Z \subset V - r_0$ nem-üres részhalmazoknak. Egy e élnek és Z halmaznak megfelelő elem 1, ha e belép Z-be, különben 0. Ekkor az

$$R = \{x \in \mathbf{R}^A : 0 \leq x \leq 1, \varrho_x(Z) \geq 1 \text{ minden } R \subseteq V - r_0\text{-re}\} \tag{5.16}$$

poliéder egy egész eleme $0 - 1$ értékű és így a D egy gyökeresen összefüggő részgráfjának felel meg.

TÉTEL 5.3.6 *A* (5.16) *poliéder egész.*

Biz. Legyen $c : A \to \mathbf{Z}$ egészértékű, amelyre a primál optimum létezik (ami most azzal ekvivalens, hogy $c(e) < 0$ esetén $g(e)$ véges).

Ekkor a primál probléma $\min\{cx : 0 \leq x \leq 1, Qx \geq 1\}$, míg a duális:

$$\max\{\sum[y(Z) : Z \subseteq V - r_0] - z(A) : yQ - z \leq c, y \geq 0, z \geq 0\}, \tag{5.17}$$

ahol az $x(e) \leq 1$ egyenlőtlenségnek megfelelő duális változót $z(e)$ jelöli.

Adott y egyértelműen meghatározza a z optimális választását: $z(e) = (yq_e - c(e))^+$, ahol q_e a Q e-nek megfelelő oszlopa. Így beszélhetünk arról, hogy egy y a (5.17) optimális megoldása.

Belátjuk, hogy (5.17) optimuma egész vektoron is felvétetik. Legyen y_0 egy optimális (racionális) megoldás. Ameddig csak létezik két valódi metsző halmaz X és Y melyek y_0-értéke pozitív, módosítsuk y_0 a következőképpen.

Az $\alpha := \min\{y_0(X), y_0(Y)\}$ értékkel csökkentsük $y_0(X)$-t és $y_0(Y)$-t és egyúttal növeljük α-val mind $X \cap Y$, mind $X \cup Y$ y_0-értékét. Könnyű ellenőrizni, hogy (a ϱ függvény szubmodularitása miatt) ismét duális megoldást kapunk, amely optimális. A duális optimum ezen megváltoztatását nevezzük egy kikeresztezési lépésnek.

Könnyen igazolható, hogy kikeresztezési lépésből csak véges sok lehet.

Feltehetjük tehát, hogy azon halmazok $\mathcal{H}$ rendszere, melyeken az y_0 értéke pozitív, lamináris. Ebből következik, hogy a Q mátrixnak a $\mathcal{H}$ sorai által meghatározott részmátrixa teljesen unimoduláris és emiatt y_0 egész vektor (hiszen a duálist korlátozó c vektor egész). Következik, hogy a duális optimum értéke egész és így a dualitás tétel folytán ezzel megegyező primál optimum is egész szám.

A 5.3.5 tételből következik, hogy R valóban egész poliéder. •

A bizonyításból kiadódott az alábbi következmény.

TÉTEL 5.3.7 *A* (5.17) *duális poliéder egész.* •

Mivel mind a primál, mind a duál poliéder egész, a lineáris programozás dualitás tétele megadja a legolcsóbb fenyő költségét. Ennek kimondásához nevezzünk egy $y : 2^{V-s} \to \mathbf{R}_+$ halmazfüggvényt c-**megengedettnek**, ha

$$c(e) \geq \sum(y(X) : e \text{ belép } X\text{-be}) \quad \text{minden } e \in E \text{ élre.} \tag{5.18}$$

TÉTEL 5.3.8 *[Fulkerson, 1974] Az s-fenyők minimális költsége egyenlő* $\max\{\sum[y(X) : X \subseteq V - s] : y$ c*-megengedett*$\}$. *Továbbá, ha c egészértékű, akkor az optimális y választható egészértékűnek.*

Az alábbiakban Fulkerson tételére megadunk egy alternatív, algoritmikus bizonyítást.

Biz. Legyen F feszítő r_0-fenyő és y egy c-megengedett vektor. Ekkor

$$c(F) = \sum(c(a) : e \in F) \geq \sum(\sum(y(X) : e \text{ belép } X- \text{ be}) : e \in F) \geq \sum(y(X) : X \subseteq V - s) \tag{5.19}$$

amiből max $\leq$ min következik. (5.19)-ben akkor van végig egyenlőség, ha a következő optimalitási feltételek teljesülnek.

$$c(e) = \sum(y(X) : e \text{ belép } X\text{-be}) \text{ minden } e \in F \text{ élre,} \tag{5.20}$$

$$y(X) > 0 \text{ esetén } \varrho_F(X) = 1. \tag{5.21}$$

Az alábbi algoritmus egy olyan F feszítő r_0-fenyőt és megengedett y-t konstruál, amelyre (5.20) és (5.21) teljesül. Két fázisból áll. Az elsőben y-t konstruáljuk meg, míg a másodikban F-t. Mindkét rész egyfajta értelemben mohó lesz. Az első fázis minden lépésében módoítjuk a költségfüggvényt, és az aktuális költségfüggvényt c'-vel fogjuk jelölni. Egy e élt a c' aktuális költségfüggvényre nézve **0-élnek** nevezzük, ha $c'(e) = 0$.

1. fázis *Amíg van* $V - r_0$*-nak olyan nemüres részhalmaza, amelybe nem lép be 0-él, ismételjük a következő lépést. Válasszunk egy olyan minimális nemüres* $X \subseteq V - r_0$ *részhalmazt, amelybe nem lép 0-él. Legyen* $y(X) := \min(c'(e) : e$ *belép* X*-be) és csökkentsük* $c'(e)$*-t a* $y(X)$ *értékkel az összes* X*-be lépő* e *élen.*

A módosított c' továbbra is nemnegatív, és mindazokon az X-be belépő éleken 0-vá vált, amelyeken az előbbi minimum eléretik. Az 1. fázis tehát akkor fejeződik be, amikor már minden $X \subseteq V - r_0$ részhalmazba lép be 0-él, vagyis amikor már létezik 0-élekből álló feszítő r_0-fenyő.

2. fázis *Az* r_0 *pontból kiindulva, 0-élek egymás utáni hozzávételével felépítünk egy* F r_0*-fenyőt. Ha az építési eljárás egy lépésében több, mint egy olyan 0-él van, amely a már megkonstruált részfenyő ponthalmazából kilép, akkor azt az élt adjuk a részfenyőhöz, amely az 1 fázis során a leghamarabb vált 0-éllé.*

Az előállítás szabályai miatt világos, hogy a megkonstruált y vektor c-megengedett, továbbá, hogy (5.20) fennáll.

Lemma 5.3.9 *y és F kielégítik (5.21)-t.*

Biz. Legyen X olyan halmaz, amelyre $y(X) > 0$ és tegyük fel indirekt, hogy $\varrho_F(X) > 1$. Ekkor a 2. fázisnak van egy olyan pillanata, amikor az aktuális F' részfenyő olyan, hogy $\varrho_{F'}(X) = 1$ és az F'-höz éppen hozzáadásra kerülő e él belép X-be.

Tekintsük most az 1. fázisnak azt a pillanatát, amikor $y(X)$ pozitív lett. Ekkor az X-be még nem lépett 0-él, ugyanakkor az X-nek már minden valódi nemüres részhalmazába lépett. Speciálisan, az $X - V(F')$ halmazba is lépett egy f 0-él, és mivel ez X-be nem léphetett, így f töve $X \cap V(F')$-ben van. Az f tehát már $y(X)$ pozitívvá válásának pillanatában 0-él volt, amikor még e nem volt az. Miután az f éllel is lehetne növelni az F' részfenyőt, ellentmondásra jutottunk a 2. fázis választási szabályával, amely szerint a legkorábban 0-éllé vált éllel kell növelni az aktuális részfenyőt. •

A lemmából következik, hogy az algoritmus által megtalált F feszítő r_0-fenyő és y megengedett duális megoldás kielégítik az optimalitási feltételeket, amiből az algoritmus helyessége, valamint Fulkerson tétel is következik. • •

2011. október 29.file: egesz

5.4 Dinamikus programozás

A dinamikus programozás egy lehetséges algoritmikus megközelítési mód kombinatorikus optimalizálási feladatok megoldására. Az általános ideát néhány feladaton mutatjuk be.

1. *Adott egy $G = (V, E)$ irányított gráf, élein egy c konzervatív súlyozás, határozzunk meg egy s-ből t-be vezető minimális súlyú utat.*

Ezt a problémát már korábban tárgyaltuk. A megoldás egyik kulcs észrevétele az a megfigyelés, hogy ha P egy s-ből t-be vezető legrövidebb út és v ennek belső pontja, akkor P-nek az s-től v-ig vezető kezdő szakasza egy s-ből v-be vezető legrövidebb út. (Ez az állítás csak konzervatív súlyozás esetén áll fenn és ekkor is csak irányított gráfra.)

A másik fontos ötlet az volt, hogy nem csupán a t-be vezető legrövidebb utat határoztuk meg, hanem az összes többi pontba vezetőt, sőt még ennél is többet: Meghatároztuk az összes v pontra és az összes $0 \leq k \leq n$ egészre az s-ből v-be vezető legfeljebb k élt használó utak közül a legkisebb költségűt.

Tegyük fel, hogy már ismertek a legfeljebb k élt használó legrövidebb utak $w_k(v)$ költségei. Tekintsünk egy v be menő legfeljebb $k + 1$ élű utat, amelynek költsége minimális. Amennyiben ez legfeljebb k élből áll, akkor $w_{k+1}(v) = w_k(v)$. Ha viszont pontosan $k + 1$ élből áll, akkor az utolsó előtti v' pontjáig tartó részút k élű és így költsége éppen $w_k(v')$. Ennek alapján $w_{k+1}(v) = \min\{w_k(v),\ \min\{w_k(u) + c(uv) : uv \in E\}$: minden $v \in V$-re$\}$.

Ennek alapján az algoritmus n fázisból áll: a $(k + 1)$-edikben a fenti rekurzió alapján határozhatjuk meg a $w_{k+1}(v)$ értékeket minden v pontra.

Általában is, dinamikus programozási algoritmus kidolgozására három jelenség egyidejű fennállásakor van lehetőségünk. Egyrészt a keresett optimális struktúra olyan legyen, hogy valamely része egy alkalmas részprobléma optimális megoldását adja, másrészt a szóbajövő részproblémák teljes száma ne legyen túl nagy (azaz az input méretének polinomjával legyen korlátozható), és végül a részproblémák optimális megoldásaiból az eredeti probléma optimális megoldása hatékonyan kiszámítható legyen. Nézzünk további példákat ezen gondolat megvalósítására.

2. *Adott m mátrix, $A_1, A_2, \ldots, A_m$, és ki akarjuk számítani a szorzatukat. Határozzuk meg a szorzásoknak azt a sorrendjét (vagyis a zárójelezést), amelyre az elemi szorzások össz-száma a lehető legkisebb lesz!*

Természetesen az összeszorozhatóság érdekében feltesszük, hogy a mátrixok méretei stimmelnek, azaz álljon az A_i mátrix p_{i-1} sorból és p_i oszlopból. Egy ab-s és egy bc-s mátrix szorzata ac-s. Ennek egy elemének kiszámításához b szorzásra van szükség, így a szorzások teljes száma abc. Ezt tekintjük két mátrix összeszorzási költségének.

Az A_i mátrixok összeszorzási sorrendjét egy zárójelezés határozza meg. Tegyük fel például, hogy $m = 4$. Három lehetőség (összesen 5 van) és a hozzájuk tartozó szorzási költségek:

$(A_1A_2)(A_3A_4)$; $p_0p_1p_2 + p_2p_3p_4 + p_0p_2p_4$,
$(((A_1A_2)A_3)A_4)$; $p_0p_1p_2 + p_0p_2p_3 + p_0p_3p_4$,
$((A_1(A_2A_3))A_4)$; $p_1p_2p_3 + p_0p_1p_3 + p_0p_3p_4$.

Jelölje $A_{i,\ldots,k}$ az $A_i, \ldots, A_k$ mátrixok szorzatát ($1 \leq i \leq k \leq m$) és $c_{i,\ldots,k}$ a kiszámítás minimális költségét. Tehát mi a $c_{1,\ldots,m}$-re vagyunk kiváncsiak.

Tegyük fel, hogy egy optimális zárójelezés olyan, hogy az A_j-t követően bezáródik egy maximális zárójel és elkezdődik egy másik maximális (egy összetartozó zárójel-pár "maximális", ha nincs nagyobban benne). Ekkor nyilvánvalóan a zárójelezést akár az $A_1, \ldots, A_j$, akár az $A_{j+1}, \ldots, A_m$ részsorozatra megszorítva optimális zárójelezést kapunk, vagyis teljesül a fenti két elv közül az első.

Az összes lehetséges $A_i, \ldots, A_k$ részsorozatra ($1 \leq i \leq k \leq m$) meg fogjuk határozni az optimális zárójelezés $c_{i,\ldots,k}$ költségét. Szerencsére ilyen részsorozatból csak kevés van ($m(m + 1)/2$), így a fentebb megfogalmazott elvek közül a második is teljesül.

Nyilvánvalóan az egy-tagú részsorozatokra $c_{i,\ldots,i} = 0$. Tegyük fel, hogy már meghatároztuk az összes legfeljebb t tagú részsorozatra az optimális zárójelezés költségét és most a $t + 1$ tagú részsorozatokra akarjuk ugyanezt kiszámítani. Ennek érdekében vizsgáljuk az $A_i, \ldots, A_k$ $t + 1$ tagú részsorozatot, ahol $k = i + t$.

Tekintsük ennek egy optimális zárójelezését, és tegyük fel, hogy ebben a két utoljára összeszorzott tényező az $A_i \ldots A_j$ szorzat és az $A_{j+1} \ldots A_k$ szorzat, ahol $i \leq j \leq k$. Ezen utolsó szorzásnak a költsége $p_{i-1}p_jp_k$, amiből a következő rekurzió adódik.

$$c_{i,\ldots,k} = \min\{c_{i,\ldots,j} + c_{j+1,\ldots,k} + p_{i-1}p_jp_k : j = i, \ldots, k - 1\}.$$

Vagyis a legfeljebb t tagú részsorozatok összeszorzásának minimális költségéből egyszerűen ki tudjuk számítani a $t + 1$ tagúakét is.

A most következő feladat tulajdonképpen könnyen visszavezethető az elsőre, de érdemes külön is megfogalmazni, mert érdekes alkalmazásai vannak.

3. *Egy P n elemű részbenrendezett halmaz elemein adott egy c súlyozás, keressünk maximális súlyú láncot (:teljesen rendezett részhalmazt.)!*

Figyeljük meg, hogy ha C maximális súlyú lánc és p egy eleme, akkor C-ből a p-nél nagyobb elemeket elhagyva olyan láncot kapunk, amely maximális súlyú azon láncok közül, melyek maximális eleme p.

P minden p elemére meghatározzuk a legnagyobb súlyú olyan láncot, melynek maximális eleme p. A maximumot jelölje $m(p)$. Az így adódó n lánc közül a legnagyobb súlyú lesz a keresett.

Jelölje P_1 a P legkisebb elemeinek halmazát, illetve általában P_{i+1} a $P - P_1 - \ldots - P_i$ legkisebb elemeinek halmazát. Nyilván P_1 tagjaira $m(p) = c(p)$. Tegyük fel, hogy már meghatároztuk $m(p)$-t a $P_1 \cup \ldots \cup P_i$ elemeire és most $p \in P_{i+1}$. Ekkor nyilván $m(v) = c(v) + \min m(u)$, ahol a minimum azon $u \in P_i$ elemekre megy, melyek a részbenrendezésben kisebbek v-nél.

Két speciális esetet említünk.

3.1 *Különböző számokból álló $a_1, \ldots, a_m$ sorozatból válasszunk ki egy maximális monoton növő részsorozatot.*

Definiáljunk egy P részbenrendezést a $p_1, \ldots, p_m$ elemeken úgy, hogy p_i kisebb, mint p_j, ha $i < j$ és $a_i < a_j$. (Ez nyilván részbenrendezés, hiszen két lánc szorzata.) Alkalmazható az előző eljárás az azonosan 1 súlyozás esetére.

3.2 *Adott egy $G = (A, B; E)$ páros gráf, ahol $A = \{a_1, \ldots, a_k\}$ és $B = \{b_1, \ldots, b_l\}$, keressünk maximális párosítást, amely nem tartalmaz "keresztező élt", azaz ha $a_i b_j$ és $a_{i'} b_{j'}$ a párosításhoz tartozik és $i < i'$, akkor $j < j'$.*

Definiáljunk az éleken egy részbenrendezést: Az $a_i b_j$ él megelőzi az $a_{i'} b_{j'}$ élt, ha $i < i'$ és $j < j'$. A feladat éppen azt kívánja, hogy ebben a részbenrendezésben keressünk maximális elemszámú láncot.

További érdekes speciális eset a következő. Azt mondjuk, hogy a nem feltétlenül különböző betűkből készített $X = \{x_1, \ldots, x_k\}$ sorozatnak a $Z = \{z_1, \ldots, z_j\}$ sorozat részsorozata, ha léteznek olyan $i_1 < \ldots < i_j$ indexek, melyekre $x_{i_1} = z_1, \ldots, x_{i_j} = z_j$.

3.2.1 *Határozzuk meg az $X = \{x_1, \ldots, x_k\}$ és $Y = \{y_1, \ldots, u_l\}$ sorozatoknak egy maximális közös részsorozatát.*

Készítsük el a $G = (A, B; E)$ páros gráf, ahol $A = \{a_1, \ldots, a_k\}$, $B = \{b_1, \ldots, b_l\}$ és az a_i illetve b_j pontokat akkor kössük össze éllel, ha $x_i = y_j$. A feladat éppen egy maximális nem-keresztező párosítás megtalálásával ekvivalens, így alkalmazható 3.2.

A következő példa ismét irányított gráfok legrövidebb útjaival foglalkozik.

4. *$G = (V, E)$ irányított gráf élein adott egy c nemnegatív súlyozás, határozzuk meg minden s, t pontpárra az s-ből t-be vezető legrövidebb út hosszát.*

Elvileg itt nincs nehézség, hiszen $n(n-1)$ pontpár van és ezek mindegyikére a Dijkstra algoritmus segítségével például meg tudjuk a legrövidebb utat határozni. A következő Floyd és Warshalltól származó algoritmus egyszerre számolja ki a kívánt minimumokat. Jelölje C a költség mátrixot: c_{ij} az v_i pontból a v_j-be menő él költsége. Az eljárás n fázisból áll. A k-dik fázisban meghatározzuk minden v_i, v_j pontpárra a v_i-ből v_j-be menő legkisebb költségű olyan utat, amely belső pontot csak a $V_k := \{v_1, \ldots, v_k\}$ halmazból használ. Jelölje $w_k(i, j)$ ezt a minimumot illetve W_k az ezen értékekből alkotott mátrixot. Ekkor $W_1 = C$.

Jelölje P a v_i-ből v_j-be vezető minimális utat, amelynek minden belső pontja V_k-hoz tartozik. Amennyiben P nem tartalmazza v_k-t, azaz P minden belső pontja V_{k-1}-hez is már hozzá tartozik, akkor $w_k(v_i, v_j) = w_{k-1}(v_i, v_j)$.

Amennyiben v_k rajta van P-n, úgy a P-nek mind a v_i-től v_k-ig, mind a v_k-tól v_j-ig terjedő szakasza olyan út, amelynek minden belső pontja már V_{k-1}-ben van. Ezek hossza tehát $w_{k-1}(v_i, v_k)$ illetve $w_{k-1}(v_k, v_j)$. Ebből adódik, hogy érvényes a következő rekurzió:

$$w_k(v_i, v_j) = \min\{w_{k-1}(v_i, v_j),\ w_{k-1}(v_i, v_k) + w_{k-1}(v_k, v_j)\}.$$

Vagyis a W_k mátrixot $O(n^2)$ lépésben ki tudjuk számítani W_{k-1}-ből.

Végezetül bemutatjuk a dinamikus programozásnak egy eddigieknél bonyolultabb helyzetben történő alkalmazását.

5. *Legyen $G = (V, E)$ irányítatlan gráf és T a pontjainak egy részhalmaza, mely pontokat terminálóknak nevezünk. Adott még az éleken egy nem-negatív c költség-függvény. Keressünk minimális össz-súlyú összefüggő részgráfot, amely T minden pontját tartalmazza!*

Világos, hogy a minimum egy részfán vétetik fel, mely részfát **Steiner** fának nevezik. A minimális súlyú Steiner fa feladata közös általánosítása a legrövidebb út problémának (amikor T kételemű) és a minimális költségű feszítő fa problémájának (amikor $T = V$). Bár ez a két szélső eset polinom időben megoldható, R. Karp bebizonyította, hogy a Steiner fa probléma NP-teljes, még azonosan 1 súlyok esetén is.

Amint azt Dreyfus és Wagner kimutatta, dinamikus programozás segítségével megadható olyan algoritmus, amely n-ben (a gráf pontszámában) polinomiális és csak $|T|$-ben exponenciális. Vagyis ha T kicsi, pontosabban rögzített, akkor ez az algoritmus polinomiális futásidejű. Például, ha T három elemű, akkor minden v pontból meghatározunk mindhárom terminálba egy-egy legrövidebb utat. Amelyik v-re a legkisebb ezen utak össz-hosszúsága, a három út uniója, amint az egyszerűen látható, az optimális Steiner fát adja.

Az általános eset tárgyalásához először is feltesszük, hogy a költség-függvény kielégíti a háromszög egyenlőtlenséget. Ha nem ez a helyzet, akkor minden u, v pontpárra helyettesítjük a költségüket a köztük vezető legkisebb kölségű út költségével. Jelölje $s(T' + v)$ az olyan Steiner fa minimális költségét, melynek terminál halmaza $T' \cup \{v\}$ ahol T' a T-nek valódi részhalmaza míg v tetszőleges pont (tehát lehet T-n belül is és kívül is). Jelölje $s_b(T' + v)$ az olyan Steiner fa minimális költségét, amelynek terminál halmaza ismét $T' \cup \{v\}$, de megköveteljük, hogy v a Steiner-fának belső, azaz legalább 2 fokú pontja legyen. (Ha ilyen fa nem létezik, a költséget végtelennek definiáljuk.) T minden T' részhalmazára és G minden v pontjára ki fogjuk számítani mind az $s(T' + v)$, mind az $s_b(T' + v)$ értéket. Ezek száma csak T méretétől függ exponenciálisan, $|V|$-ben polinomiális.

Legyen K a terminálóknak egy részhalmaza és tekintsünk egy ezekre vonatkozó F minimális Steiner fát, melynek valamely v pontja a fában legalább 2 fokú. Osszuk tetszőlegesen a v-ből kiinduló fa-éleket két nem-üres osztályba. Ez meghatározza a fának is egy olyan két részre osztását, amely részeknek csak v a közös pontjuk. Jelölje F' és F'' a két részfát továbbá K' és K'' a hozzájuk tartozó terminál halmazt. Világos, hogy F' (illetve F'') minimális Steiner fa a $K' + v$ ($K'' + v$) terminál halmazra nézve. Ebből adódik, hogy érvényes a következő rekurzió:

$$s_b(K + v) = \min\{s(K' + v) + s(K - K' + v) : \emptyset \subset K' \subset K\}. \tag{4.1}$$

Tekintsünk most egy $K + v$ terminálhalmazhoz tartozó F minimális Steiner fát. Három eset lehetséges.

(i) v az F belső pontja. Ekkor nyilván $s(K + v) = s_b(K + v)$. Legyen

$$\alpha_1 := s_b(K + v).$$

(ii) v a fának első fokú csúcsa, és az egyetlen fabeli szomszédja, u nincs K-ban. Ekkor a háromszög egyenlőtlenség miatt feltehető, hogy u legalább harmadfokú a fában, azaz v elhagyása után legalább 2 fokú. Ezért ilyenkor $s(K + v) = s_b(K + u) + c(uv)$. Legyen

$$\alpha_2 := \min\{s_b(K + u) + c(uv) : u \notin K\}.$$

(iii) v a fának első fokú csúcsa, és az egyetlen fabeli szomszédja, u a K-ban van. Ekkor $s(K+v) = s(K)+c(uv)$. Legyen

$$\alpha_3 := \min\{s(K) + c(uv) : u \notin K\}.$$

Ezen megfigyelésekből adódik a következő rekurzió.

$$s(K + v) = \min\{\alpha_1,\ \alpha_2,\ \alpha_3\}. \tag{4.2}$$

Az algoritmus k-adik fázisában ($k = 1, 2, \ldots, |T| - 1$) feltesszük, hogy minden v csúcsra és T-nek minden legfeljebb k elemű K részhalmazára már kiszámítottuk $s(K + v)$-t és $s_b(K + v)$-t. Ekkor (4.1) segítségével először kiszámítjuk $s_b(K + v)$-t a T-nek minden $k + 1$ elemű részhalmazára, majd ezek ismeretében (4.2) segítségével $s(K + v)$-t is. •

A következő érdekes speciális esetben a fenti algoritmus nem csak n-ben lesz polinomiális, hanem $|T|$-ben is. Tegyük fel, hogy a gráf síkba-rajzolható és a terminálok mindegyike a végtelen tartomány határán van. Vegyünk egy tetszőleges F Steiner fát. Ha ennek bármely élét kihagyjuk, akkor a fa két komponensre esik. Mármost a döntő megfigyelés az, hogy a két komponens a terminálpontokat egymást követően tartalmazza, azaz nem létezhet négy terminál pont $t_1, \ldots, t_4$, melyek ebben a sorrendben következnek a külső ablakon és t_1, t_3 az egyik fa-komponenshez tartozik, míg t_2, t_4 a másikhoz. Valóban, ha indirekt ez volna a helyzet, akkor a két komponensben vegyük az egyértelmű t_1-t t_3-mal illetve a t_2-t t_4-gyel összekötő utat. A síkbeliség miatt ezen utaknak szükségképpen van közös pontja ellentmondásban a szóbanforgó két fa-komponens diszjunktságával.

Ezen "folytonossági" tulajdonságból következik, hogy az egész algoritmus során a T-nek csak olyan K részhalmazaira kell kiszámítani (4.1) és (4.2)-t, melyek a külső ablakon folytonosak (azaz K és $T - K$ nem "keresztezi" egymást.) Ilyen K kevesebb mint $|T|^2$ van, tehát (4.1) és (4.2) kiszámítása is polinom időben történhet. •

file: kombo ldinam

Tartalom