Ujvári Miklós

KONVEX ANALÍZIS

2009

Tartalomjegyzék

Bevezetés

A dualitás az a jelenség, mikor egy halmazon (S) értelmezve van egy önmagába történő leképezése (D) úgy, hogy a halmaz bármely elemére kétszer alkalmazva a leképezést, visszakapjuk az eredeti elemet: bizonyítható a

$$D : S \to S,\ D \circ D = \text{identitás}$$

ún. dualitástétel. Ekkor a halmazon értelmezett duálisképzésről beszélünk. Egy elem képét az elem duálisának nevezzük, az elemet magát pedig, a duálisától való megkülönböztetése érdekében, primál elemnek.

Ennek a jelenségnek a felismerése azért hasznos, mert segítségével az elemeket másik oldalukról is látjuk: mint duálisuk duálisát. Gondoljunk például az $n \times n$-es, valós elemű, invertálható mátrixok halmazán értelmezett invertálás leképezésre, ez duálisképzés. Tudjuk, hogy nyílt halmaz lineáris inverzképe is nyílt halmaz; és mivel a fentiek szerint egy invertálható mátrix az inverzének az inverze, azért az is látszik, hogy nyílt halmaz lineáris képe is nyílt halmaz lesz, ha a transzformáció mátrixa invertálható. E két tétel duális párt alkot, mivel a "nehéz" tétel "könnyű" párjából egy dualitástétel segítségével következik.

Másik például, ha a D_1, D_2 leképezésekről tudjuk, hogy duálisképzések, vagyis érvényesek a

$$D_1 \circ D_1 = \text{identitás},\ D_2 \circ D_2 = \text{identitás}$$

dualitástételek, akkor "könnyű"

$$D_1 \circ A = B \circ D_2$$

felcserélési tételekből (A, B leképezések) könnyen következik "nehéz" duális párjuk, az

$$A \circ D_2 = D_1 \circ B$$

felcserélési tétel — szorozzuk meg az első egyenlőséget balról D_1-gyel, jobbról D_2-vel!

A konvex analízis (és a konvex geometriával közös része is) igen gazdag mind duális tételpárokban, mind dualitástételekben. A leghíresebb könnyű-nehéz duális tételpár a Rockafellar-féle primál-duál programpárokra vonatkozó gyenge, illetve erős dualitási tétel. Itt a Rockafellar-féle Farkas-tétel az a dualitástétel, amelynek segítségével a könnyű (gyenge dualitási) tételből következik a nehéz (erős dualitási) tétel.

A jegyzet elsődleges célja ezeknek a mély felismeréseknek az ismertetése.

A jegyzet felépítése

Az első hét fejezetben egyelőre a kúplineáris programok általánosságáig jutunk.

1. A Gauss–Jordan-elimináció leírása; az output vizsgálata; következmények (Fredholm alternatívatétele, Hermite-féle normálalak, Moore–Penrose-féle általánosított inverz).

2. A homogén, illetve általános lineáris egyenletrendszerek megoldáshalmazainak vizsgálata: jellemzések, dimenzió, az e tulajdonságokat megtartó transzformációk. Olyan fontos ötletek tűnnek fel, mint a dualitás, a homogenizáció vagy a diagonális altér.

3. Analóg vizsgálatok lineáris egyenlőtlenség-rendszerek megoldáshalmazaival kapcsolatban. Egyik fő célunk a Motzkin-tétel bizonyítása, mely szerint a poliéderek éppen a végesen generált halmazok. Először a tétel homogén változatát, a Weyl–Minkowski-tételt igazoljuk, a Farkas-lemma és a Caratheodory-tétel azonnali következményének, a lineáris egyenlőtlenségek alaptételének segítségével. A fejezet végén az első szeparációs tételtől (első Hahn–Banach-tétel) a Farkas-lemma számtalan variánsát egységesítő kúplineáris Farkas-lemmán át eljutunk a (zártsági feltételes) kúplineáris Farkas-tételig, amelyet a hatodik fejezetben használunk majd, kúplineáris erős dualitási tételek bizonyítására.

4. Konvex halmazok relatív belsejével és lezártjával kapcsolatos alapvető kérdések (mikor üres/ relatív nyílt/ zárt/ affin?) megválaszolása után a ri, cl (és rec) operációkra vonatkozó felcserélhetőségi tételeket bizonyítunk. (Operáción itt egyszerűen $\mathcal{R}^n$-beli halmazokkal végzett műveletet értünk.) Ezek közül kiemelkedő fontosságúak a zártsági tételek, melyek elégséges feltételt adnak arra, hogy egy zárt, konvex halmaz lineáris képe is zárt legyen.

5. Szeparálhatóságok hierarchiáját definiáljuk, azután megvizsgáljuk, hogy két nemüres, diszjunkt, külön-külön relatív nyílt/ zárt/ poliédrikus, konvex halmaz "mennyire" szeparálható. Alkalmazásként zárt, konvex halmazok új jellemzéseit nyerjük bizonyos támaszféltereik metszeteként. További alkalmazásokat látunk majd a következő fejezetben.

6. A kúplineáris Farkas-lemma szigorú egyenlőtlenséges (Stiemke-tétel), illetve Slater-feltételes (Krein-tétel) változatait tárgyaljuk, különféle általánosságban. E tételek (és a 4. fejezetbeli zártsági tételek) segítségével elégséges (Slater-)feltételét adjuk a kúplineáris Farkas-lemma, a kúplineáris Farkas-tétel, és ezáltal a kúplineáris programokra vonatkozó dualitási tételek zártsági feltételeinek.

7. Példák elemzése után konvex és lezárt konvex generálási tételek (a Minkowski–Klee-tétel, illetve a Straszewicz–Klee-tétel) segítségével belátjuk, hogy a poliéderek éppen azok a zárt, konvex halmazok, amelyeknek véges sok exponált lapja van. Szó esik még a poliéderek minimális és maximális valódi lapjairól, majd felcserélhetőségi tételek segítségével konjugált lapokra vonatkozó dualitástételeket bizonyítunk. A fejezet végén kúplineáris programok regularizációjával foglalkozunk, vagyis azzal a kérdéssel, hogy hogyan szabadulhatunk meg a dualitási tételek Slater-feltételeitől.

Az utolsó hét fejezetben kerül sor a konvex, illetve sima függvényekkel leírt programok vizsgálatára.

8. Függvények epigráfjának és nívóhalmazainak tulajdonságait vizsgáljuk, például azt, hogy mikor zártak, konvexek; konvex esetben mi a lezártjuk, relatív belsejük. Végül a cl (és a rec) operátorokra vonatkozó felcserélhetőségi tételeket bizonyítunk, valamint Weierstrass tételét. (Operátoron itt $\mathcal{R}^n$-en értelmezett függvényekkel végzett műveletet értünk.)

9. Főleg a konjugálás operátorra vonatkozó felcserélhetőségi tételeket bizonyítunk, az előző fejezet eredményeinek, valamint a konvex Stiemke-tétel segítségével. Illusztrációképpen támasz- és távolságfüggvényekkel kapcsolatos észrevételek szerepelnek.

10. A fejezetet a főleg az effektív tartomány relatív belsejében meglévő, alapvető folytonossági és szubdifferenciálhatósági tulajdonságok vizsgálatával kezdjük. Ezután a függvénynek és konjugáltjának szubdifferenciálja közti kapcsolat, valamint az előző fejezet eredményeinek segítségével felcserélhetőségi tételeket bizonyítunk. Konvex függvények differenciálhatóságának, majd differenciálható függvények konvexitásának jellemzéseit adjuk.

A fejezet végén differenciálható, konvex függvények lokális (és így globális) minimumhelyeit karakterizáljuk.

11. A hatodik fejezetben kúplineáris programokra bizonyított dualitási tételek jelentős általánosítása, a Rockafellar-féle programpárra. Példaképpen Eisenberg és Dennis dualitási tételeit említjük.

12. A Lagrange-féle programpárra vonatkozó dualitási és optimalitási tételeket részben az előző fejezet eredményeiből, részben a konvex Farkas-tételből bizonyítjuk. Külön vizsgáljuk azt az esetet, mikor a primál programot differenciálható, konvex függvények írják le, ekkor a Lagrange-duál helyett tekinthetjük az egyszerűbb Wolfe-duált is. Utóbbira példaképpen kvadratikus (speciálisan lineáris) dualitási tételeket bizonyítunk. Végül az eddigi gondolatmenet egy részét megismételjük konvex függvények helyett kúpkonvex leképezésekkel leírt programokra, kitérve regularizációjukra is.

13. Konvexitás helyett simaságot megkövetelve a Lagrange-féle primál programot leíró függvényekre, dualitási tételeket már nem, de optimalitási tételeket még mindig igazolhatunk. Az itt leírt bizonyításban a feladathoz rendelt kúpoknak (belső, megengedett, csökkenési és érintő irányok kúpjai), valamint a 6. fejezet tételeinek jut fontos szerep.

14. A 11. és 12. fejezet egyes eredményeit bizonyítjuk újra, valamivel gyengébb formában, de nagyobb rálátással.

A függelékekben vizsgáljuk politóp és konvex halmaz összegének tulajdonságait, majd alkalmazásként további zártsági és szeparációs tételeket bizonyítunk.

Irodalomjegyzék és köszönetnyilvánítás

A jegyzet olvasása lineáris algebrai, analízisbeli, lineáris programozási előismereteket igényel:

1, 7, 10, 12, 16, 18, 22, 23, 25.

Az egyes fejezetekhez felhasznált irodalom:

1.	16, 22, 25
2.	10, 12, 13, 20, 24
3.	2, 18, 23
4.	2, 13, 20
5.	13, 20
6.	2, 3, 13, 20
7.	13, 20, 24, 23, 26 4, 32, 17, 31
8.	7, 13, 20
9.	13, 20
10.	3, 7, 8, 13, 15, 20
11.	20, 24
12.	3, 4, 5, 6, 9, 11 13, 18, 21, 24, 32
13.	1, 3, 7, 8, 13, 14
14.	8
A-C.	28, 29, 30.

A területtel való további ismerkedéshez ajánlott irodalom: 13, 19, 21, 27, 20 (a terület alapműve), 24 (a minimax elmélethez), 3 (a végtelen dimenzió felé).

Köszönetet mondok Kovács Margitnak, akinek Konvex analízis és nemlineáris programozás előadásai, valamint az ezekhez az előadásokhoz írt könyve nagy segítségemre voltak e jegyzet elkészítésében. Például a jegyzet 13. fejezetét szinte változtatás nélkül [13]-ból vettem át.

Köszönetet mondok Frank Andrásnak, témavezetőmnek, aki sokirányú segítséget nyújtott e jegyzet elkészítéséhez: a 3., 6. és 7. fejezetek jelentős részét az ő másodéves matematikus hallgatóknak tartott Operációkutatás előadásai alapján írtam meg; ő hívta fel a figyelmemet a [25] és [3] könyvekre; és számos konzultációval is segített. OTKA keretéből (jelenlegi száma T17580) támogatta a jegyzet elkészítését.

Nagyon sokat profitáltam Rapcsák Tamással folytatott beszélgetéseinkből. Több, a témához tartozó cikket átgondoltunk együtt. Ő hívta fel a figyelmem a [19] jegyzetre is.

Terlaky Tamás hívta fel a figyelmemet a [21] jegyzetre (ahonnan a 12. fejezet lényegét vettem át), és számos konzultációval is segített. Hozzá a Peregrinatio II. Alapítvány segítségével jutottam ki a delfti Műszaki Egyetem Operációkutatási Tanszékére.

Illés Tibor hívta fel a figyelmemet a kúplineáris programozás alapvető eredményeire és a [8] jegyzetre (ahonnan a 14. fejezetet vettem át).

Köszönettel tartozom még ifj. Böröczky Károlynak, tőle hallottam 5.7 itt leírt bizonyításának korábbi változatát és 5.15 bizonyítását.

A jegyzet megírásának idején munkáltatóim a MTA Rényi Alfréd Matematikai Kutatóintézet és az Eötvös Loránd Tudományegyetem voltak.

Jelölések

$\|S\|$	az S halmaz elemszáma
$\mathcal{N}, \mathcal{R}$	a természetes és a valós számok halmaza
$\mathcal{R}_+, \mathcal{R}_-$	a nemnegatív, illetve a nempozitív valós számok halmaza
$S^{m \times n}$	az S halmaz elemeiből álló $m \times n$-es mátrixok
S^n	$S^{n \times 1}$, az n-elemű (oszlop)vektorok halmaza
$a_{ij}, (A)_{ij}$	az $A \in S^{m \times n}$ mátrix (i,j)-edik eleme,
${}_ia, a_j$	i-edik sorvektora, illetve j-edik oszlopvektora
$x_i, (x)_i$	az x vektor i-edik eleme
A^T, A^{-1}	az A mátrix transzponáltja, illetve inverze
e_i	$(\delta_{1i}, \ldots, \delta_{ni})^T$, ahol $\delta_{ij} = 1$, ha $i = j$, 0 különben
$\mathbf{1}$	$e_1 + \ldots + e_n = (1, \ldots, 1)^T \in \mathcal{R}^n$
E	$e_1 e_1^T + \ldots + e_n e_n^T \in \mathcal{R}^{n \times n}$, $(E)_{ij} = \delta_{ij}$
$\|\|x\|\|$	az $x \in \mathcal{R}^n$ vektor euklideszi normája, $\|\|x\|\| = \sqrt{x^T x}$
$O(x, \varepsilon)$	az x középpontú, ε sugarú, nyílt gömb
$\operatorname{int} S$	az $S \subseteq \mathcal{R}^n$ halmaz belső pontjainak halmaza
$\operatorname{cl} S$	az $S \subseteq \mathcal{R}^n$ halmaz lezártja
$f\|_S$	az f függvény megszorítása az S halmazra
$\sup S$, $\inf S$	az $S \subseteq \mathcal{R}$ halmaz szuprémuma, illetve infimuma,
$\max S$, $\min S$	maximuma, minimuma (ha a sup, illetve az inf felvétetik)
$:=$	definiáló egyenlőség

Végesen generált halmazok

1. A Gauss–Jordan-elimináció

A Gauss–Jordan-elimináció fontos bizonyítási eszköz. Elméletileg lineáris egyenletrendszerek megoldására szolgál, vagyis, adott $A \in \mathcal{R}^{m \times n}$ mátrix és $b \in \mathcal{R}^m$ vektor esetén, $Ax = b$ tulajdonságú $x \in \mathcal{R}^n$ vektor(ok) megkeresésére. Már az $m = n = 1$ eset mutatja e kérdés bonyolultságát: a $0x = 0$, $0x = 1$, illetve $1x = 0$ egyenletrendszerek megoldáshalmaza rendre $\mathcal{R}$, $\emptyset$, illetve $\{0\}$. Megoldóképlet tehát nem várható. Az alábbi algoritmus eldönti, hogy az egyenletrendszernek létezik-e megoldása, és ha igen, akkor megad egy, a megoldáshalmazt "generáló" vektorrendszert.

A módszer a következő észrevételen alapszik:

1.1. Állítás: *Ha az* (A, b) *mátrixon végrehajtjuk az alábbi transzformációk egyikét, és az így kapott mátrix* (A', b')*, akkor az* $Ax = b$ *és az* $A'x = b'$ *egyenletrendszerek megoldáshalmaza megegyezik:*
T_1*: A mátrix egyik sorát egy nemnulla* λ *valós számmal szorozzuk.*
T_2*: A mátrix egyik sorához hozzáadjuk egy másik sorának* λ*-szorosát, ahol* λ *tetszőleges valós szám.* $\square$

Természetesen a fenti tulajdonsággal rendelkezik az a transzformáció is, hogy a mátrix két sorát felcseréljük. (Ezt jelöljük T_0-lal.) A T_0, T_1, T_2 transzformációkat **elemi sortranszformációk**nak nevezzük. Hasonlóan definiálhatók az elemi oszloptranszformációk is.

1.2. Állítás: *Az elemi sortranszformációk egyenértékűek a mátrix egy ún.* **elemi mátrix***szal való balról szorzásával, ahol az elemi mátrixot úgy kapjuk, hogy a megfelelő transzformációt az egységmátrixon hajtjuk végre. Az elemi mátrixok invertálhatóak.* $\square$

Hasonló állítás mondható ki az elemi oszloptranszformációk esetén is, csak itt az elemi mátrixszal jobbról kell szorozni. Ennek az állításnak a megjegyzését segítheti az angol “jobs=munkák” szó: Jobbról Oszlopot, Balról Sort transzformál az elemi mátrixszal való szorzás.

1.3. Állítás: *Az egyetlen egyenletből álló $Ax = b$ rendszernek pontosan akkor nincs megoldása, ha $A = 0$, és $b \neq 0$.*

Bizonyítás: Ha $A \neq 0$, akkor $x := A^T(AA^T)^{-1}b$ megoldás. Ha pedig $A = 0$, akkor az $Ax = b$ rendszer csak $b = 0$ esetén megoldható. □

A Gauss–Jordan-elimináció során T_1 és T_2 típusú sortranszformációkkal olyan alakra hozzuk a megoldható egyenletrendszert, amelynek megoldása(i) már könnyen kiolvasható(k). Ha az egyenletrendszernek nincs megoldása, akkor az egyik átalakított rendszernek lesz egy 1.3 szerint ellentmondó ($0x = (\neq 0)$ alakú) egyenlete.

A következőkben a **Gauss–Jordan-eliminációt** írjuk le formálisabban. Az algoritmusban szereplő egyes változók jelentése:

i	jelöli, hogy hányadik sorban áll az éppen vizsgált egyenlet;
k	jelöli az eddig talált nemnulla sorok számát;
$A^{(k)}x = b^{(k)}$	a k-szor transzformált, $Ax = b$-vel azonos megoldáshalmazú rendszer.

Gauss–Jordan-elimináció:

1. Be: m, n, A, b $[m, n \in \mathcal{N},\ A \in \mathcal{R}^{m \times n},\ b \in \mathcal{R}^m]$
2. $i := 1 : k := 0 : A^{(0)} := A : b^{(0)} := b$
3. Ciklus, amíg $i \leq m$, és az $A^{(k)}x = b^{(k)}$ rendszer i-edik egyenlete megoldható (vö. 1.3)
4. Ha $A^{(k)}$ i-edik sorának van nemnulla eleme, akkor $k := k + 1$: **Sortranszformáció**
5. $i := i + 1$
6. Ciklus vége
7. Ha $i \leq m$, akkor Ki: $Ax = b$ megoldhatatlan!
   különben Ki: $Ax = b$ megoldható!
   $r := k$
   $p := b^{(r)}_{i_1} e_{j_1} + \ldots + b^{(r)}_{i_r} e_{j_r} \in \mathcal{R}^n$,
   $q^{(j)} := e_j - a^{(r)}_{i_1 j} e_{j_1} - \ldots - a^{(r)}_{i_r j} e_{j_r} \in \mathcal{R}^n$
   $(j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\})$
8. Program vége.

Sortranszformáció:	
1.	$i_k := i$
2.	Válasszunk j_k oszlopindexet úgy, hogy az $A^{(k-1)}$ mátrix (i_k, j_k) pozícióján álló eleme nemnulla legyen.
3.	Először osszuk le az $(A^{(k-1)}, b^{(k-1)})$ mátrix i_k-adik sorát j_k-adik elemével, majd a sor megfelelő számú többszöröseit vonjuk le a többi sorból úgy, hogy a j_k-adik oszlop i_k-tól különböző indexű elemei lenullázódjanak.
4.	Az így kapott mátrix legyen $(A^{(k)},\ b^{(k)})$.
5.	Szubrutin vége.

Néhány egyszerű észrevétel az algoritmussal kapcsolatban:
– Az algoritmus során mindig $i_1 < \ldots < i_k$; és az $A^{(k)}$ mátrix i-nél kisebb, az $i_1, \ldots, i_k$ számok mindegyikétől különböző indexű sorai nullák, akárcsak a $b^{(k)}$ vektor megfelelő elemei.
– Az algoritmus során mindig az $A^{(k)}$ mátrix j_1-edik, …, j_k-adik oszlopvektorai rendre $e_{i_1} \in \mathcal{R}^m$, …, $e_{i_k} \in \mathcal{R}^m$. Az is látszik, hogy $j_k \neq j_1, \ldots, j_{k-1}$.
– Léteznek $N^{(k)}$ invertálható mátrixok úgy, hogy

$$A^{(k)} = N^{(k)}A, \text{ és } b^{(k)} = N^{(k)}b.$$

Ezeket a mátrixokat úgy kapjuk, hogy az $m \times m$-es egységmátrixon is elvégezzük ugyanazokat az elemi sortranszformációkat, amelyeket az A mátrixon végrehajtottunk, míg az $A^{(k)}$ mátrixig jutottunk (vö. 1.2). Itt felhasználtuk, hogy invertálható mátrixok szorzata is invertálható.

1.4. Tétel: *Ha az algoritmus az $Ax = b$ rendszer megoldhatatlanságát állapítja meg, akkor az valóban megoldhatatlan. Ha az algoritmus szerint $Ax = b$ megoldható, akkor az ekkor kapott p, $q^{(j)}$ ($j \neq j_1, \ldots, j_r$) vektorokkal fennáll, hogy*

$$\{x \in \mathcal{R}^n : Ax = b\} = \\ = \left\{p + \textstyle\sum_j \lambda_j q^{(j)} : \lambda_j \in \mathcal{R}\ (j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\})\right\},$$

és

$$\{z \in \mathcal{R}^n : Az = 0\} = \\ = \left\{\textstyle\sum_j \lambda_j q^{(j)} : \lambda_j \in \mathcal{R}\ (j \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\})\right\}.$$

(Semennyi vektor összegén a 0 vektort értve az $r = n$ esetben.)

Bizonyítás: Ha az algoritmus az $Ax = b$ rendszer megoldhatatlanságát állapítja meg, akkor az aktuális $A^{(k)}x = b^{(k)}$ rendszer i-edik egyenlete, és így az $A^{(k)}x = b^{(k)}$ rendszer is megoldhatatlan. Ekkor 1.1 szerint az $Ax = b$ rendszer is megoldhatatlan.

Különben a fenti megjegyzések segítségével könnyen belátható, hogy

$$A^{(r)}p = b^{(r)}, \text{ és } A^{(r)}q^{(j)} = 0 \ (j \neq j_1, \ldots, j_r).$$

Ebből 1.1 szerint

$$Ap = b, \text{ és } Aq^{(j)} = 0 \ (j \neq j_1, \ldots, j_r)$$

is következik, valamint az igazolandó halmazegyenlőségek "$\supseteq$" része.

Ha most $x \in \mathcal{R}^n$ olyan vektor, amelyre $Ax = b$ teljesül, akkor ismét a fenti megjegyzések segítségével, koordinátáról koordinátára ellenőrizve, könnyen belátható, hogy

$$x = p + \sum_{j \neq j_1, \ldots, j_r} (x - p)_j q^{(j)}.$$

Ez bizonyítja az első igazolandó halmazegyenlőség "$\subseteq$" részét is.

Végül ha $z \in \mathcal{R}^n$ olyan vektor, amelyre $Az = 0$ teljesül, akkor az $x := p + z$ vektorra $Ax = b$ teljesül, és így az előző bekezdésben igazoltak szerint

$$z = \sum_{j \neq j_1, \ldots, j_r} z_j q^{(j)}.$$

Eszerint a második igazolandó halmazegyenlőségben is fennáll a "$\subseteq$" tartalmazás. $\square$

A Gauss–Jordan-elimináció végrehajtása során a j_k oszlopindexek megválasztása többféleképpen történhet, ennek megfelelően más-más lehet az algoritmus lefolyása. Természetesen megoldható $Ax = b$ rendszer esetén az eljárás azt mindig megoldhatónak találja majd, de előfordulhatna például, hogy különböző r számokat talál különböző lehetséges lefutásai során. Megmutatjuk, hogy ez nincs így, sőt megoldható $Ax = b$ egyenletrendszer esetén talált r szám csak az A mátrixtól függ, annak "rangja". (Belátjuk például, hogy az adott rendű, négyzetes mátrixok között a lehető legrangosabbak éppen az invertálható mátrixok.)

A $v_1, \ldots, v_l \in \mathcal{R}^d$ (esetleg nem páronként különböző) vektorok **lineáris kombináció**i a $\sum_{i=1}^{l} \lambda_i v_i$ vektorok, ahol $\lambda_1, \ldots, \lambda_l \in \mathcal{R}$.

A $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorok **lineárisan függetlenek**, ha csak triviális módon lehet belőlük kikombinálni a 0 vektort, vagyis

$$\lambda_1, \ldots, \lambda_l \in \mathcal{R}, \sum_{i=1}^{l} \lambda_i v_i = 0 \text{ esetén } \lambda_1 = \ldots = \lambda_l = 0.$$

Ha a $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorrendszer nem lineárisan független, akkor **lineárisan összefüggő**nek nevezzük.

Lineárisan független például az üres vektorrendszer vagy különböző 0-1 egységvektorok tetszőleges rendszere (egységvektoron egységnyi hosszú vektort értünk). Lineárisan összefüggő bármely a nullvektort vagy két egyforma vektort tartalmazó vektorrendszer. Lineárisan független vektorrendszer bármely része is lineárisan független, míg lineárisan összefüggő vektorrendszert tartalmazó vektorrendszer is lineárisan összefüggő. Ha S és S' is lineárisan független vektorrendszer, továbbá teljesül, hogy $v \in S$ és $v' \in S'$ esetén $v^T v' = 0$, akkor $S \cup S'$ is lineárisan független vektorrendszer lesz.

További könnyen ellenőrizhető észrevételek a lineáris függetlenséggel kapcsolatban:

1.5. Állítás: *Ha a $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorok lineárisan függetlenek, de rendszerük egy v_{l+1} vektorral bővítve már lineárisan összefüggő, akkor a v_{l+1} vektor előáll a $v_1, \ldots, v_l$ vektorok lineáris kombinációjaként.* □

1.6. Állítás: *A $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorok pontosan akkor lineárisan függetlenek, ha egyikük sem áll elő a többiek lineáris kombinációjaként.* □

Hogy egy vektorrendszer lineárisan független vagy összefüggő, azt mechanikusan az alábbi módon dönthetjük el: Alkossunk a vektorokból, mint oszlopvektorokból, egy $A \in \mathcal{R}^{d \times l}$ mátrixot, és Gauss–Jordan-eliminációval vizsgáljuk meg, hogy az $Ax = 0$ homogén egyenletrendszernek van-e nullától különböző megoldása (vagyis $r < l$ teljesül-e). Ha van ilyen megoldás, akkor a vektorok lineárisan összefüggőek, különben lineárisan függetlenek. Az is látszik, hogy ha a vektorok száma, l nagyobb, mint a tér dimenziója, d, akkor a vektorrendszer biztosan lineárisan összefüggő, ekkor ugyanis $r \leq d < l$. Tehát

1.7. Állítás: *Az $\mathcal{R}^d$-beli lineárisan független vektorrendszerek legfeljebb d eleműek.* □

Az $S_g := (e_1, \ldots, e_d)$ vektorrendszer mutatja, hogy az alábbi állítás 1.7 erősítése.

1.8. Állítás: *Ha az S_f lineárisan független vektorrendszer elemei kikombinálhatók az S_g vektorrendszer elemeiből, akkor S_f elemszáma legfeljebb akkora, mint S_g elemszáma.*

Bizonyítás: Az S_g vektorrendszer elemeiből, mint oszlopvektorokból, állítsunk össze egy A mátrixot, hasonlóan S_f-ből egy B mátrixot. Ekkor létezik X mátrix úgy, hogy $AX = B$. Az X mátrix oszlopvektorai nyilván lineárisan függetlenek, így 1.7 szerint számuk, $|S_f|$ legfeljebb akkora, mint a sorok száma, $|S_g|$. □

A $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorrendszer legtöbb elemből álló, lineárisan független részrendszerét a vektorrendszer **bázis**ának nevezzük. A bázis elemszáma a vektorrendszer **rang**ja. 1.7 szerint a $v_1, \ldots, v_l \in \mathcal{R}^d$ vektorrendszer rangja legfeljebb d (és persze legfeljebb l).

1.9. Állítás: *Egy vektorrendszer valamely része pontosan akkor bázis, ha belőle a vektorrendszer minden eleme kikombinálható, méghozzá egyféleképpen.*

Bizonyítás: Az állítás 1.5, 1.6 és 1.8 egyszerű következménye. □

Mint azt 1.13-ból látni fogjuk, vektorrendszer rangja és egy bázisa is meghatározható a Gauss–Jordan-elimináció segítségével. Előkészítésképpen néhány állítást fogalmazunk meg a rangról.

1.10. Állítás: *Egy vektorrendszer rangja nem változik, ha*
a) elhagyunk belőle egy nulla vektort;
b) két vektorát felcseréljük;
c) egyik vektorát egy $\lambda \neq 0$ valós számmal megszorozzuk;
d) egyik vektorához hozzáadjuk egy másik vektorának λ-szorosát, ahol λ tetszőleges valós szám.

Bizonyítás: Csak az állítás d) részét kell igazolnunk. Itt c) miatt feltehető, hogy $\lambda = 1$, és b) miatt elég azt megmutatnunk, hogy ha S a $v_1, v_2, \ldots, v_l$ vektorrendszer lineárisan független része, akkor a $v_1 + v_2, v_2, \ldots, v_l$ vektorrendszerben is található vele azonos elemszámú S' lineárisan független részrendszer.

Ha $v_1 \notin S$, akkor $S' := S$ megfelel. Ha $v_1, v_2 \in S$, akkor könnyen belátható, hogy

$$S' := \{v_1 + v_2\} \cup (S \setminus \{v_1\})$$

megfelel. Marad az az eset, mikor $v_1 \in S$, de $v_2 \notin S$. Ekkor könnyen belátható, hogy a fenti S' rendszer, vagy ha az nem, hát

$$S' := \{v_2\} \cup (S \setminus \{v_1\})$$

megfelel. □

Adott $A \in \mathcal{R}^{m \times n}$ mátrix esetén sorvektorai rendszerének rangja a mátrix **sorrang**ja, oszlopvektorai rendszerének rangja a mátrix **oszloprang**ja, jelük $\mathrm{sr}\,(A)$, illetve $\mathrm{or}\,(A)$. Nyilván

$$\mathrm{sr}\,(A) = \mathrm{or}\,(A^T), \text{ és } \mathrm{or}\,(A) = \mathrm{sr}\,(A^T).$$

Tudjuk továbbá, hogy

$$\max\{\mathrm{sr}\,(A), \mathrm{or}\,(A)\} \leq \min\{m, n\}.$$

1.11. Állítás: *Tetszőleges (szorozható) A, B mátrixok esetén*

$$\mathrm{or}\,(AB) \leq \mathrm{or}\,(B),$$

egyenlőséggel, ha A invertálható. Hasonlóan

$$\mathrm{sr}\,(AB) \leq \mathrm{sr}\,(A),$$

egyenlőséggel, ha B invertálható.

Bizonyítás: Az AB mátrix oszlopvektorai az $Ab_1, \ldots, Ab_k$ vektorok. Ha például az $Ab_1, \ldots, Ab_r$ vektorok lineárisan függetlenek, akkor a $b_1, \ldots, b_r$ vektorok is lineárisan függetlenek, ugyanis $Bx = 0$ esetén $ABx = 0$ is fennáll. Ebből már látszik, hogy $\mathrm{or}\,(AB) \leq \mathrm{or}\,(B)$.

Ha most az A mátrix invertálható, akkor a fentiek szerint

$$\mathrm{or}\,(B) = \mathrm{or}\,(A^{-1}AB) \leq \mathrm{or}\,(AB)$$

is teljesül.

Az állítás második fele hasonlóan igazolható. □

1.12. Állítás: *Az A mátrix sor- és oszloprangja sem változik, ha*
a) a mátrix egy csupa nulla sorát vagy oszlopát elhagyjuk, illetve ha
b) a mátrixon elemi sor- vagy oszloptranszformációkat hajtunk végre.

Bizonyítás: Az állítás 1.10 és 1.11 egyszerű következménye (vö. 1.2). □

1.13. Tétel: *Az A mátrix sor- és oszloprangja is megegyezik a Gauss–Jordan-elimináció során bármely olyan b vektor mellett kiszámolt r számmal, amelyre az $Ax = b$ rendszer megoldható (például $b = 0$ ilyen vektor). Az*

$$Ae_{j_1}, \ldots, Ae_{j_r}, \text{ illetve az } A^T e_{i_1}, \ldots, A^T e_{i_r}$$

vektorrendszer az A mátrix oszlop-, illetve sorvektorai rendszerének egy bázisa.

Bizonyítás: A tétel 1.12 egyszerű következménye.

Könnyű belátni, hogy az $A^{(r)}$ mátrix, és így az A mátrix sor- és oszloprangja is r. Ha ugyanis az $A^{(r)}$ mátrixból csupa nulla sorait elhagyva csak az $i_1, \ldots, i_r$ indexű sorait hagyjuk meg, akkor az így kapott r sorból álló mátrix sor- és oszloprangja is r, hiszen j_1-edik, …, j_r-edik oszlopai különböző 0-1 egységvektorok. Ez bizonyítja a tétel első felét.

A tétel második feléhez hajtsuk végre az eliminációt csak a $j_1, \ldots, j_r$ indexű oszlopvektorokon, illetve csak az $i_1, \ldots, i_r$ indexű sorvektorokon. Ugyanolyan lefutás mellett r sor- és oszloprangú mátrixhoz, az $A^{(r)}$ mátrix megfelelő részmátrixához jutunk. Ezért 1.12 szerint az eredeti A mátrix megfelelő oszlop- és sorvektorai lineárisan függetlenek; a tétel már igazolt első fele miatt pedig bázist alkotnak. □

Eszerint az A mátrix sor- és oszloprangja megegyezik. Ezt a közös értéket a mátrix **rang**jának nevezzük, és $\mathrm{r}\,(A)$-val jelöljük.

1.14. Állítás: *Szorozható A, B mátrixok esetén*

$$\mathrm{r}\,(AB) \le \min\{\mathrm{r}\,(A), \mathrm{r}\,(B)\}.$$

Összeadható A, B mátrixok esetén

$$\mathrm{r}\,(A + B) \le \mathrm{r}\,(A) + \mathrm{r}\,(B).$$

Bizonyítás: Az állítás első fele 1.11 következménye, második fele pedig az elsőé, ugyanis

$$A + B = (A, B)(E, E)^T,$$

és

$$\mathrm{r}\,(A, B) \le \mathrm{r}\,(A) + \mathrm{r}\,(B).$$

□

1.15. Állítás: *Az invertálható mátrixok éppen azok a négyzetes mátrixok, amelyeknek a rangja a rendjével egyezik meg.*

Bizonyítás: Legyen $A \in \mathcal{R}^{n\times n}$. Tegyük fel először, hogy az A mátrix invertálható, vagyis létezik (szükségképpen egyértelmű) B mátrix úgy, hogy $AB = E = BA$. Ekkor az $Ax = 0$ egyenletrendszernek egyetlen megoldása a 0 vektor (hiszen $Ax = 0$ esetén $x = BAx = 0$), így A oszlopvektorai lineárisan függetlenek, és A rangja a rendje, n.

Ha most $\mathrm{r}\,(A) = n$, akkor az A mátrix oszlopvektorainak lineárisan független rendszerét tetszőleges e_i vektorral bővítve már lineárisan összefüggő rendszert kapunk, így 1.5 szerint $1 \le i \le n$ esetén az $Ax = e_i$ rendszer megoldható, jelölje egy megoldását x_i. Jelölje X azt a mátrixot, amelynek oszlopvektorai rendre $x_1, \ldots, x_n$. Ekkor $X \in \mathcal{R}^{n\times n}$, teljesül, hogy $AX = E$, és 1.14 szerint $\mathrm{r}\,(X) = n$. A fentiek szerint az X mátrixhoz is létezik Y mátrix úgy, hogy $XY = E$ legyen. Ekkor

$$XA = XA(XY) = X(AX)Y = XY = E,$$

tehát $XA = E = AX$, vagyis az X mátrix az A mátrix inverze. □

Most a lineáris egyenletrendszerek megoldhatóságának két ekvivalens feltételét fogalmazzuk meg.

1.16. Tétel: (Kronecker–Capelli) *Ha az $Ax = b$ rendszer megoldható, akkor*

$$\mathrm{r}\,(A, b) = \mathrm{r}\,(A),$$

különben pedig

$$\mathrm{r}\,(A, b) = \mathrm{r}\,(A) + 1.$$

Bizonyítás: A bizonyítást azzal a nyilvánvaló észrevétellel kezdjük, hogy

$$\mathrm{r}\,(A) \le \mathrm{r}\,(A, b) \le \mathrm{r}\,(A) + 1.$$

Most tegyük fel indirekt, hogy az $Ax = b$ rendszer megoldható, mégis $\mathrm{r}\,(A, b) = \mathrm{r}\,(A) + 1$. Ekkor az (A, b) mátrix egy bázisának A-beli része az A mátrix bázisa lenne (mátrix bázisa alatt az oszlopvektor-rendszerének bázisát értve), az A mátrix oszlopvektorai, és így a b vektor is előállna e vektorok lineáris kombinációjaként. Ez pedig ellentmond annak, hogy az (A, b) mátrix fenti bázisa lineárisan független.

Végül tegyük fel indirekt, hogy az $Ax = b$ rendszer nem megoldható, mégis $\mathrm{r}(A, b) = \mathrm{r}(A)$. Ekkor az A mátrix egy bázisa az (A, b) mátrix bázisa is, vagyis a b vektor kikombinálható belőle. De akkor az $Ax = b$ rendszer mégis megoldható lenne, ellentmondáshoz jutottunk. □

Ha az $Ax = b$ rendszer megoldható, akkor ennek tanúja a p vektor. Hasonlóan tömör bizonyítékát adhatjuk az $Ax = b$ rendszer megoldhatatlanságának is.

1.17. Tétel: (Fredholm alternatívatétele) *Az $Ax = b$ rendszer pontosan akkor megoldhatatlan, ha létezik $y \in \mathcal{R}^m$ vektor úgy, hogy $A^T y = 0$, és $b^T y \neq 0$. (Másképpen megfogalmazva az $Ax = b$ rendszer pontosan akkor megoldható, ha $A^T y = 0$ esetén $b^T y = 0$.)*

Bizonyítás: A tétel "akkor" iránya könnyen belátható: Ha létezne a tételben leírt tulajdonságú x és y vektor is, akkor a

$$0 = 0^T x = (A^T y)^T x = y^T A x = y^T (Ax) = y^T b \neq 0$$

ellentmondáshoz jutnánk.

A tétel másik irányát Gauss–Jordan-eliminációval bizonyíthatjuk: Ha az $Ax = b$ rendszer nem megoldható, akkor az elimináció során valamikor ellentmondó egyenletre bukkanunk, vagyis az aktuális $A^{(k)} x = b^{(k)}$ rendszer i-edik egyenlete $0^T x = (\neq 0)$ alakú lesz.

Tudjuk, hogy az $N := N^{(k)}$ mátrixszal $A^{(k)} = NA$, $b^{(k)} = Nb$. A fentiek szerint

$$e_i^T A^{(k)} = e_i^T N A = 0,\; e_i^T b^{(k)} = e_i^T N b \neq 0.$$

Legyen y az N mátrix i-edik sorvektorának, $e_i^T N$-nek a transzponáltja. Erre az y vektorra teljesül, hogy $A^T y = 0$, $b^T y \neq 0$, mint azt az előző egyenletek éppen kimondják. □

Azt mondjuk, hogy az $A \in \mathcal{R}^{m \times n}$ mátrix **Hermite-féle normálalak**ú, ha valamely r esetén kiválaszthatók páronként különböző $i_1, \ldots, i_r$ sorindexei és páronként különböző $j_1, \ldots, j_r$ oszlopindexei úgy, hogy:

a) Ha az i sorindex az $i_1, \ldots, i_r$ sorindexek mindegyikétől különbözik, akkor az A mátrix i-edik sora csupa nullából áll.

b) Bármely $k \in \{1, \ldots, r\}$ esetén az A mátrix j_k-adik oszlopának minden eleme nulla, leszámítva az i_k-adikat, ami egyes.

Ekkor automatikusan $\mathrm{r}(A) = r$. (Ez éppen úgy igazolható, mint az 1.13 bizonyításában szereplő $\mathrm{r}(A^{(r)}) = r$ állítás.)

Az $i_1, \ldots, i_r$ sorindexek és a $j_1, \ldots, j_r$ oszlopindexek választása nem egyértelmű, a következőkben mindig egy lehetséges választást jelölnek.

A P_π alakú mátrixokat, ahol

$$P_\pi := \sum_{k=1}^{m} e_{\pi(k)} e_k^T \in \mathcal{R}^{m\times m} \quad (\pi : \{1, \ldots, m\} \to \{1, \ldots, m\} \text{ permutáció}),$$

permutáció mátrixoknak nevezzük. Könnyen belátható, hogy ha π, σ permutációk, akkor $P_{\pi\circ\sigma} = P_\pi \cdot P_\sigma$, így P_π invertálható, $P_\pi^{-1} = P_\pi^T = P_{\pi^{-1}}$. Teljesül továbbá, hogy

$$(P_\pi^T A)_{ij} = a_{\pi(i)j}, \text{ sőt } (P_\pi^T A P_\sigma)_{ij} = a_{\pi(i)\sigma(j)}.$$

Ha az A mátrix Hermite-féle normálalakú, akkor $P_\pi^T A P_\sigma$ is az.

1.18. Tétel: *Tetszőleges $A \in \mathcal{R}^{m\times n}$ mátrixhoz létezik $N \in \mathcal{R}^{m\times m}$ invertálható mátrix úgy, hogy az NA mátrix Hermite-féle normálalakú. Tetszőleges $A \in \mathcal{R}^{m\times n}$ mátrixhoz léteznek $N_1 \in \mathcal{R}^{m\times m}$, $N_2 \in \mathcal{R}^{n\times n}$ invertálható mátrixok úgy, hogy az $N_1 A N_2$ mátrix és transzponáltja is Hermite-féle normálalakú.*

Bizonyítás: A tétel első feléhez oldjuk meg az $Ax = 0$ egyenletrendszert Gauss–Jordan-eliminációval, ekkor az $N := N^{(r)}$ mátrix megfelel.

A tétel második feléhez legyen $N_1 \in \mathcal{R}^{m\times m}$ olyan invertálható mátrix, amelyre az $N_1 A$ mátrix Hermite-féle normálalakú. Ekkor az $N_1 A$ mátrix $r = \mathrm{r}\,(A)$ darab kitüntetett oszlopában különböző 0-1 egységvektorok állnak, ezek egyesei által kijelölt sorokat elhagyva az $N_1 A$ mátrixból, nullmátrixhoz jutunk. A kitüntetett oszlopok megfelelő számszorosait kivonva a többi oszlopból elérhetjük, hogy a nem kitüntetett oszlopok elemei lenullázódjanak. A fenti elemi oszloptranszformációk egy-egy elemi mátrixszal való jobbról szorzásnak felelnek meg (vö. 1.2), ezen elemi mátrixok szorzata legyen N_2. Nyilvánvalóan N_1 és N_2 megfelelnek. □

Azt mondjuk, hogy az $X \in \mathcal{R}^{n\times m}$ mátrix az $A \in \mathcal{R}^{m\times n}$ mátrix **Moore–Penrose-féle általánosított inverz**e, ha teljesül rá, hogy
a) $AXA = A$;
b) $XAX = X$;
c) $(AX)^T = AX$;
d) $(XA)^T = XA$.

Ha $A \in \mathcal{R}^{n\times n}$ invertálható, akkor az $X := A^{-1}$ mátrixra a), b), c) és d) is fennáll.

Például ha az $A \in \mathcal{R}^{m \times n}$ mátrix rangja m $[n]$, akkor

$$X = A^T(AA^T)^{-1} \ [X = (A^TA)^{-1}A^T]$$

megfelel. (Ha $\mathrm{r}\,(A) = m$, akkor

$$AA^Tx = 0 \Rightarrow ||A^Tx||^2 = x^TAA^Tx = 0 \Rightarrow A^Tx = 0 \Rightarrow x = 0,$$

így $\mathrm{r}\,(AA^T) = m$, és 1.15 szerint AA^T invertálható.)

1.19. Tétel: *Minden mátrixnak egyetlen Moore–Penrose-féle általánosított inverze van.*

Bizonyítás: Legyen $A \in \mathcal{R}^{m \times n}$ tetszőleges mátrix. Először azt mutatjuk meg, hogy az A mátrixnak legfeljebb egy Moore–Penrose-féle általánosított inverze lehet.

Tegyük fel, hogy az X_1 és az X_2 mátrix is kielégíti az általánosított inverz definíciójában szereplő a), b), c) és d) mátrixegyenleteket. Megmutatjuk, hogy
1. tetszőleges $x \in \mathcal{R}^n$ esetén $(X_1 - X_2)Ax = 0$;
2. $y \in \mathcal{R}^m$, $A^Ty = 0$ esetén $(X_1 - X_2)y = 0$;
3. $X_1 = X_2$.

1. Először is a) és d) szerint, $X = X_1$ vagy $X = X_2$ esetén

$$A^T = (AXA)^T = A^TX^TA^T = (XA)^TA^T = XAA^T,$$

így $(X_1 - X_2)AA^T = 0$, amiből

$$(X_1 - X_2)AA^T(X_1 - X_2)^T = 0$$

adódik. Ha egy B mátrixra $BB^T = 0$ teljesül, akkor szükségképpen $B = 0$, ugyanis ekkor

$$\sum_{i,j} b_{ij}^2 = \sum_i (BB^T)_{ii} = 0.$$

Ezt az észrevételt a $B = (X_1 - X_2)A$ mátrixra alkalmazva látjuk, hogy $(X_1 - X_2)A = 0$. Ezért tetszőleges $x \in \mathcal{R}^n$ esetén $(X_1 - X_2)Ax = 0$.

2. Másodszor b) és c) szerint, $X = X_1$ vagy $X = X_2$ esetén

$$X = X(AX) = X(AX)^T = XX^TA^T,$$

ezért $y \in \mathcal{R}^m$, $A^Ty = 0$ esetén $(X_1 - X_2)y = 0$.

3. Legyen $N \in \mathcal{R}^{m\times m}$ olyan invertálható mátrix, amelyre az NA mátrix Hermite-féle normálalakú. Könnyen belátható, hogy az

$$N^T e_i \ (i \neq i_1, \dots, i_r),\ Ae_{j_1}, \dots, Ae_{j_r}$$

vektorok lineárisan függetlenek (lásd az 1.5 előtti észrevételt). A fentiek szerint e vektorok bármelyikét balról szorozva az $X_1 - X_2$ mátrixszal, nullvektort kapunk. Mivel a fenti m lineárisan független $\mathcal{R}^m$-beli vektor lineáris kombinációjaként tetszőleges $e_i \in \mathcal{R}^m$ vektor előáll, azért

$$(X_1 - X_2)e_i = 0 \ (i = 1, \dots, m)$$

is teljesül, és akkor persze $X_1 - X_2 = 0$, vagyis $X_1 = X_2$.

Most megmutatjuk, hogy tetszőleges $A \in \mathcal{R}^{m\times n}$ mátrixnak létezik Moore–Penrose-féle általánosított inverze.

Legyenek $N_1 \in \mathcal{R}^{m\times m}$ és $N_2 \in \mathcal{R}^{n\times n}$ invertálható mátrixok úgy, hogy az N_1AN_2 mátrix és transzponáltja is Hermite-féle normálalakú. Ekkor alkalmas $\pi : \{1, \dots, m\} \to \{1, \dots, m\}$ és $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ permutációkkal

$$P_\pi^T N_1 A N_2 P_\sigma = \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix},$$

ahol a szereplő E mátrix $r \times r$-es, ahol $r = \mathrm{r}\,(A)$. Legyen

$$U := N_1^{-1} P_\pi (E, 0)^T \in \mathcal{R}^{m\times r},\ W := (E, 0) P_\sigma^T N_2^{-1} \in \mathcal{R}^{r\times n},$$

ekkor $A = UW$, továbbá $\mathrm{r}\,(U) = \mathrm{r}\,(W) = r$, tehát az U^TU és WW^T mátrixok invertálhatók. Ezek után könnyen belátható, hogy

$$X := W^T (WW^T)^{-1} (U^TU)^{-1} U^T$$

az A mátrix egy Moore–Penrose-féle általánosított inverze. □

Az A mátrix a fentiek szerint egyértelmű Moore–Penrose-féle általánosított inverzének jele $A^\dagger$.

1.20. Következmény: *Tetszőleges $A \in \mathcal{R}^{m\times n}$ mátrix esetén*

$$(A^\dagger)^\dagger = A, \text{ és } (A^T)^\dagger = (A^\dagger)^T.$$

□

2. Alterek, affin halmazok

Ebben a fejezetben a homogén ($Ax = 0$ alakú), illetve általános ($Ax = b$ alakú) lineáris egyenletrendszerek megoldáshalmazainak szerkezetét vizsgáljuk. Ez vezet a címben említett fogalmakhoz. Három igen fontos ötlet is megjelenik, a dualitás, a diagonális altér és a homogenizáció, ezek lényegét vázoljuk először tömören.

A **dualitás** az a jelenség, mikor egy halmazon értelmezve van egy önmagába történő leképezése úgy, hogy minden elemre kétszer alkalmazva a leképezést, visszakapjuk az eredeti elemet. Ekkor a halmazon értelmezett **duálisképzés**ről beszélünk. Egy elem képét az **elem duálisá**nak nevezzük, az elemet magát pedig, a duálisától való megkülönböztetése érdekében, **primál elem**nek.

Például egy valós szám (-1)-szerese nevezhető a szám duálisának, de a halmaz és a duális leképezés lehet például (lásd a 2., 3., 6., 7. fejezeteket)
– az $\mathcal{R}^d$-beli alterek a $^\perp$-képzéssel;
– az $\mathcal{R}^d$-beli, zárt, konvex kúpok a *-képzéssel;
– az $\mathcal{R}^d$-beli végesen generált kúpok a *-képzéssel;
– az $\mathcal{R}^d$-beli, az origót tartalmazó, zárt, konvex halmazok a $^\circ$-képzéssel;
– az $\mathcal{R}^d$-beli, az origót tartalmazó poliéderek a $^\circ$-képzéssel;
– az $\mathcal{R}^d$-beli, az origót belső pontként tartalmazó, kompakt, konvex halmazok a $^\circ$-képzéssel;
– az $\mathcal{R}^d$-beli, az origót belső pontként tartalmazó politópok a $^\circ$-képzéssel;
– az ($\mathcal{R}^d$-beli zárt, konvex kúp és egy nemüres exponált lapja) párok a $^* \times {}^\triangle$-képzéssel;
– az ($\mathcal{R}^d$-beli poliéder kúp és egy nemüres exponált lapja) párok a $^* \times {}^\triangle$-képzéssel;
– az ($\mathcal{R}^d$-beli, az origót belsejében tartalmazó, kompakt, konvex halmaz és egy exponált lapja) párok a $^\circ \times {}^\diamond$-képzéssel;
– az ($\mathcal{R}^d$-beli, az origót belsejében tartalmazó politóp és egy lapja) párok a $^\circ \times {}^\diamond$-képzéssel;
– az összes lineáris program a duális program képzéssel;
– vagy akár az összes, zárt, konvex kúpokkal leírt kúplineáris program a duális program képzéssel.

Ennek a jelenségnek a felismerése azért hasznos, mert segítségével az elemeket másik oldalukról is látjuk: mint duálisuk duálisát.

A **diagonális altér** legyen

$$D := \cup \Big\{ \{x\} \times \ldots \times \{x\} \subseteq \mathcal{R}^{dk} : x \in \mathcal{R}^d \Big\}.$$

Az alábbi észrevétel segítségével bizonyos, több halmaz metszetére vonatkozó állításokat vissza lehet vezetni arra az esetre, mikor csak két halmaz szerepel az állításban (ráadásul az egyik altér). Ebben a fejezetben még nem lenne rá feltétlenül szükség, mégis használjuk, hogy ez az alapvető ötlet is minél korábban megjelenjen a jegyzetben, és ne számítson újnak, mikor már benne vagyunk a bizonyítások "sűrűjében". Az észrevétel a következő: ha $S = S_1 \times \ldots \times S_k$ alakú valamely $S_i \subseteq \mathcal{R}^d$ $(i = 1, \ldots, k)$ halmazok esetén, akkor

$$S \cap D = \cup \Big\{ \{x\} \times \ldots \times \{x\} \subseteq \mathcal{R}^{dk} : x \in \cap_{i=1}^k S_i \Big\}.$$

(Lásd még 2.13, 6. fejezet.)

A **homogenizáció** egy fogás, melyet alkalmazva egy bizonyos, halmazokra vonatkozó állítás érvényességét elegendő eggyel magasabb dimenziós, de strukturáltabb halmazokon ellenőrizni, például (lásd a 2.-7. fejezeteket)
– affin halmazok helyett altereken;
– konvex halmazok helyett konvex kúpokon;
– poliéderek helyett poliéder kúpokon.

A bizonyítások így sokszor egyszerűbbé, átláthatóbbakká válnak.

Az $L \subseteq \mathcal{R}^d$ halmazt **lineáris halmaz**nak, vagy rövidebben **altér**nek nevezzük, ha az origót, továbbá bármely két elemének lineáris kombinációit is tartalmazza, vagyis $0 \in L$, és

$$x_1, x_2 \in L,\ \lambda_1, \lambda_2 \in \mathcal{R} \text{ esetén } \lambda_1 x_1 + \lambda_2 x_2 \in L.$$

Másképpen megfogalmazva egy $L \subseteq \mathcal{R}^d$ nemüres halmaz pontosan akkor altér, ha teljesül rá, hogy $\mathcal{R}L \subseteq L$, és $L + L \subseteq L$ (és akkor $\mathcal{R}L = L$, és $L + L = L$, $1 \in \mathcal{R}$, és $0 \in L$ miatt), ahol természetesen

$$S \cdot S_1 := \{\lambda x : \lambda \in S,\ x \in S_1\}\ (S \subseteq \mathcal{R},\ S_1 \subseteq \mathcal{R}^d),$$
$$S_1 + S_2 := \{x_1 + x_2 : x_1 \in S_1,\ x_2 \in S_2\}\ (S_1, S_2 \subseteq \mathcal{R}^d).$$

Ha az S, S_1 vagy S_2 halmazok valamelyike egyelemű, akkor a fenti jelölésben a halmaz helyett egyetlen elemét is írhatjuk (például $\{x\} + S_2$ helyett $x + S_2$ szerepelhet).

Altér például a csak az origót tartalmazó halmaz, vagy az egész $\mathcal{R}^d$ tér. Könnyen belátható, hogy ha $A \in \mathcal{R}^{m\times d}$, $B \in \mathcal{R}^{d\times n}$, továbbá $L_1 \subseteq \mathcal{R}^m$, $L_2 \subseteq \mathcal{R}^n$ alterek, akkor az

$$A^{-1}(S_1) := \{x \in \mathcal{R}^d : Ax \in S_1\} \ (S_1 \subseteq \mathcal{R}^m),$$
$$BS_2 := \{By : y \in S_2\} \ (S_2 \subseteq \mathcal{R}^n)$$

jelölésekkel az $A^{-1}(L_1)$ lineáris inverzkép és a BL_2 lineáris kép is altér, speciálisan az A mátrix **nulltere** és a B mátrix **képtere**,

$$\operatorname{Ker} A := A^{-1}(0) = \{x \in \mathcal{R}^d : Ax = 0\}, \text{ illetve}$$
$$\operatorname{Im} B := B\mathcal{R}^n = \{By : y \in \mathcal{R}^n\}$$

alterek.

Felmerül az a (két) kérdés, hogy megfordítva, minden $L \subseteq \mathcal{R}^d$ altér előáll-e $\operatorname{Ker} A$, illetve $\operatorname{Im} B$ alakban, valamely $A \in \mathcal{R}^{m\times d}$, $B \in \mathcal{R}^{d\times n}$ mátrixok esetén. Ha az első kérdésre igen a válasz, akkor látható, hogy az alterek zárt halmaz (a $\{0\}$) folytonos inverz képeként zárt halmazok. Ha a második kérdésre igen a válasz, akkor az alterek egy tömör leírását nyertük véges sok elemük lineáris kombinációinak halmazaként.

Megmutatjuk, hogy a két kérdés ekvivalens, és mindkettőre igen a válasz. Először vezessünk be két fogalmat, amelyek segítségével újra fogalmazhatjuk a kérdéseket.

Alterek tetszőleges rendszerének metszete nyilván altér, így egy adott $S \subseteq \mathcal{R}^d$ halmaz esetén létezik egyértelműen az S halmazt tartalmazó legszűkebb altér, S **lineáris burka**,

$$\operatorname{lin} S := \cap\{L : S \subseteq L,\ L \text{ altér}\}.$$

2.1. Állítás: *Tetszőleges nemüres $S \subseteq \mathcal{R}^d$ halmaz esetén teljesül, hogy*

$$\operatorname{lin} S = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathcal{N},\ \lambda_i \in \mathcal{R},\ x_i \in S\ (i = 1, \dots, k) \right\}.$$

Az üres halmaz lineáris burka a triviális $\{0\}$ altér.

Bizonyítás: A jobb oldali halmaz (az S-beli elemek lineáris kombinációinak halmaza) nyilván az S halmazt tartalmazó altér, így S lineáris burkát is tartalmazza.

Másfelől k szerinti indukcióval könnyen belátható, hogy egy az S halmazt tartalmazó altér szükségképpen az S elemeiből képzett k-tagú lineáris kombinációkat is tartalmazza. □

Teljesül, hogy $0 \in S_1, S_2$ esetén

$$\operatorname{lin}(S_1 \times S_2) = (\operatorname{lin} S_1) \times (\operatorname{lin} S_2).$$

A lineárisburok-képzés tartalmazástartó:

$$S_1 \subseteq S_2 \text{ esetén } \operatorname{lin} S_1 \subseteq \operatorname{lin} S_2.$$

Adott $S \subseteq \mathcal{R}^d$ halmaz esetén S **ortogonális kiegészítő altere** az összes S-beli vektorra merőleges vektorok halmaza. (Két vektor, a és x **merőleges** vagy **ortogonális**, ha $a^T x = 0$.) Az S halmaz ortogonális kiegészítő alterének jele $S^\perp$, eszerint

$$S^\perp := \left\{ a \in \mathcal{R}^d : a^T x = 0 \ (x \in S) \right\}.$$

Bármely $S \subseteq \mathcal{R}^d$ halmaz esetén $S^\perp$ (zárt) altér. Ha $S = L$ altér, akkor az $L^\perp$ alteret szokás még az L altér **duális alter**ének nevezni (vö. 2.12). Teljesül, hogy $0 \in S_1, S_2$ esetén

$$(S_1 \times S_2)^\perp = S_1^\perp \times S_2^\perp.$$

Az ortogonális kiegészítő altér képzése tartalmazás-fordító:

$$S_1 \subseteq S_2 \text{ esetén } S_1^\perp \supseteq S_2^\perp.$$

2.2. Állítás: *Tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén $S^\perp = (\operatorname{lin} S)^\perp$.* □

Első kérdésünk úgy fogalmazható, hogy minden $L \subseteq \mathcal{R}^d$ altérhez létezik-e véges S halmaz, amelyre $L = S^\perp$, a második kérdés pedig úgy, hogy létezik-e véges S halmaz, amelyre $L = \operatorname{lin} S$. Megmutatjuk, hogy ez a két kérdés ekvivalens egymással.

2.3. Tétel: *Tetszőleges $A \in \mathcal{R}^{m \times n}$ mátrix esetén*

$$(\operatorname{Im} A)^\perp = \operatorname{Ker} A^T, \textit{ és } (\operatorname{Ker} A^T)^\perp = \operatorname{Im} A.$$

Bizonyítás: Legyen S az A oszlopvektorainak halmaza. Ekkor 2.1 szerint $\operatorname{Im} A = \operatorname{lin} S$, továbbá nyilván $\operatorname{Ker} A^T = S^\perp$. A tétel első fele ezek után 2.2 következménye.

A tétel második fele Fredholm alternatívatételéből következik. Eszerint ugyanis az $Ax = b$ rendszer pontosan akkor megoldható (vagyis $b \in \operatorname{Im} A$), ha $A^T y = 0$ esetén $b^T y = 0$ (vagyis $b \in (\operatorname{Ker} A^T)^\perp$). $\square$

Most már látjuk, hogy ha a két kérdés valamelyikére igen a válasz, akkor abból (a természetes

$$S^{\perp\perp} := (S^\perp)^\perp \ (S \subseteq \mathcal{R}^d)$$

jelöléssel) azonnal következik az alapvető

$$L^{\perp\perp} = L \ (L \subseteq \mathcal{R}^d \text{ altér})$$

azonosság. Ebből a két kérdés ekvivalenciája is látszik. Ha például már beláttuk, hogy minden $L \subseteq \mathcal{R}^d$ altérhez létezik $B \in \mathcal{R}^{d\times n}$ mátrix úgy, hogy $L = \operatorname{Im} B$, akkor tetszőleges $L \subseteq \mathcal{R}^d$ altér esetén létezik $A \in \mathcal{R}^{m\times d}$ mátrix úgy, hogy $L^\perp = \operatorname{Im} A^T$, és akkor

$$L = L^{\perp\perp} = (\operatorname{Im} A^T)^\perp = \operatorname{Ker} A.$$

Egy $S \subseteq \mathcal{R}^d$ halmazt az $L \subseteq \mathcal{R}^d$ altér **(lineáris) generáló részhalmaz**ának hívunk, ha $\operatorname{lin} S = L$. Ilyen részhalmaz például az egész L altér.

A fentiek szerint elég például azt belátni, hogy tetszőleges $L \subseteq \mathcal{R}^d$ altérnek létezik véges generáló részhalmaza. E célból érdemes az L minimális (nem szűkíthető) generáló részhalmazait vizsgálni. Kiderül, hogy ezek éppen a lineárisan független generáló részhalmazok.

2.4. Állítás: *Adott $S \subseteq \mathcal{R}^d$ halmaz esetén az alábbi állítások ekvivalensek:*
a) minden $S' \subset S$ esetén $\operatorname{lin} S' \subset \operatorname{lin} S$;
b) minden $x \in S$ esetén $x \notin \operatorname{lin}(S \setminus \{x\})$;
c) ha $k \in \mathcal{N}$, $\lambda_1, \ldots, \lambda_k \in \mathcal{R}$, $x_1, \ldots, x_k \in S$, továbbá $\sum_{i=1}^k \lambda_i x_i = 0$, akkor $\lambda_1 = \ldots = \lambda_k = 0$. $\square$

Ha az a), b) vagy c) állítások egyike (és akkor mindegyike) teljesül, akkor az S halmazt **lineárisan független**nek nevezzük, különben pedig **lineárisan összefüggő**nek. Ez a lineáris függetlenségi definíció nyilván az előző fejezetben szereplő definíció általánosítása, véges halmazokra megegyeznek. Egy lineárisan független halmaz most is legfeljebb d elemű lehet, hiszen most is teljesül, hogy lineárisan független halmaz bármely része is lineárisan független, az ellenpéldának pedig lenne $d + 1$ elemű része.

2.5. Állítás: *Adott $L \subseteq \mathcal{R}^d$ altér és $S \subseteq \mathcal{R}^d$ halmaz esetén az alábbi állítások ekvivalensek:*
a) S minimális generáló részhalmaza L-nek, vagyis $\operatorname{lin} S = L$, de $S' \subset S$ esetén $\operatorname{lin} S' \subset L$;
b) S lineárisan független generáló részhalmaza L-nek;
c) S maximális (nem bővíthető) L-beli lineárisan független halmaz, vagyis $S \subseteq L$ lineárisan független, de $S \subset \tilde{S} \subseteq L$ esetén $\tilde{S}$ már lineárisan összefüggő. □

Ha az a), b) vagy c) állítások egyike (és akkor mindegyike) teljesül, akkor az S halmazt az L altér **bázis**ának nevezzük. 1.9 segítségével könnyen belátható, hogy valamely A mátrix oszlopvektor-rendszerének (ezt jelölje S) bázisai éppen az $\operatorname{Im} A$ altér S-beli bázisai. Például:

2.6. Állítás: *Legyen $A \in \mathcal{R}^{m\times n}$ tetszőleges mátrix, és alkalmazzuk a Gauss–Jordan-eliminációt az $Ax = 0$ rendszer megoldásainak megkeresésére. Ekkor az első fejezetbeli jelölésekkel:*
a) az $\operatorname{Im} A$ altér bázisa $S_1 := \{Ae_{j_1}, \dots, Ae_{j_r}\}$;
b) az $\operatorname{Im} A^T$ altér bázisa $S_2 := \{A^T e_{i_1}, \dots, A^T e_{i_r}\}$;
c) a $\operatorname{Ker} A$ altér bázisa $S_3 := \{q^{(j)} : j \neq j_1, \dots, j_r\}$;
d) a $\operatorname{Ker} A^T$ altér bázisa $S_4 := \{N^{(r)T} e_i : i \neq i_1, \dots, i_r\}$.

Bizonyítás: A bázis fogalmának 2.5 b) jellemzését használjuk.

a) 1.13 szerint S_1 az A mátrix oszlopvektorai rendszerének bázisa. Ezért S_1 lineárisan független, továbbá 1.9-ből adódóan A oszlopvektorai, és akkor $\operatorname{Im} A$ elemei is, kikombinálhatók S_1 elemeiből.

b) hasonlóan igazolható.

c) Nyilvánvaló, hogy S_3 lineárisan független; 1.4 szerint pedig a $\operatorname{Ker} A$ altér megegyezik S_3 lineáris burkával.

d) S_4 elemei az invertálható $N^{(r)}$ mátrix sorvektorai, így lineárisan függetlenek. Az is nyilvánvaló, hogy $S_4 \subseteq \operatorname{Ker} A^T$. Azt kell még megmutatnunk, hogy $x \in \operatorname{Ker} A^T$ esetén $x \in \operatorname{lin} S_4$. 1.19 bizonyítása során láttuk, hogy $S_1 \cup S_4$ elemeiből minden $\mathcal{R}^m$-beli, speciálisan az $x \in \mathcal{R}^m$ vektor is, kikombinálható. Léteznek tehát $y \in \operatorname{lin} S_4$, $z \in \operatorname{lin} S_1$ vektorok úgy, hogy $x = y + z$. Ekkor nyilván

$$0 = z^T x = z^T y + z^T z = z^T z,$$

így $z = 0$, $x = y \in \operatorname{lin} S_4$. □

Megválaszolandó második kérdésünk így fogalmazható: Létezik-e minden altérnek bázisa? (Egy bázis véges generáló részhalmaz. Egy véges generáló részhalmaz pedig tartalmaz minimális generáló részhalmazt, vagyis bázist.) Ezt a kérdést (általánosabban) válaszolja meg a következő alapvető tétel.

2.7. Tétel: *Legyen $L \subseteq \mathcal{R}^d$ altér, S_f lineárisan független, S_g az L generáló részhalmaza. Tegyük fel, hogy $S_f \subseteq S_g$. Ekkor létezik az L-nek olyan S_b bázisa, amelyre*

$$S_f \subseteq S_b \subseteq S_g.$$

Speciálisan (legyen $S_g = L$, illetve $S_f = \emptyset$!) minden független részhalmaz bázissá bővíthető, és minden generáló részhalmaz bázissá szűkíthető. Ezért minden altérnek létezik bázisa.

Bizonyítás: S_b legyen az S_f-et tartalmazó S_g-beli lineárisan független halmazok közül egy legnagyobb elemszámú. Könnyen belátható, hogy az így választott S_b halmaz megfelel 2.5 b)-nek. □

Ezzel beláttuk, hogy második kérdésünkre igen a válasz (vagyis minden altérnek van véges generáló részhalmaza). A fentiek szerint ebből az is következik, hogy az első kérdésre is igen a válasz (vagyis minden altér valamely véges halmaz ortogonális kiegészítő altere), továbbá teljesül, hogy tetszőleges altér duális alterének duális altere az eredeti altér. Az utóbbi dualitástételt hamarosan újra bizonyítjuk.

A következő tétel, amely 1.8 azonnali következménye, az alterek egy fontos mérőszámának bevezetését teszi lehetővé.

2.8. Tétel: *Legyen $L \subseteq \mathcal{R}^d$ altér. Ha $S_f \subseteq L$ lineárisan független halmaz, továbbá S_g az L generáló részhalmaza, akkor az S_f halmaz elemszáma legfeljebb akkora, mint az S_g halmazé.* □

2.8-ból azonnal következik, hogy L minden bázisa azonos elemszámú. (2.7-ből adódóan ez az állítás nem semmitmondó!) Ezt a közös elemszámot az L altér **dimenzió**jának nevezzük, jele $\dim L$. Például $\dim\{0\} = 0$, és $\dim \mathcal{R}^d = d$, hiszen $\{0\}$ bázisa az üres halmaz, és $\mathcal{R}^d$ bázisa $\{e_1, \ldots, e_d\}$. További például 1.13 és 2.6 szerint $A \in \mathcal{R}^{m\times n}$ esetén

$$\dim(\operatorname{Im} A) = \operatorname{r}(A) = \dim(\operatorname{Im} A^T),$$
$$\dim(\operatorname{Ker} A) = n - \operatorname{r}(A), \text{ és } \dim(\operatorname{Ker} A^T) = m - \operatorname{r}(A).$$

2.7 két további, a dimenzióval kapcsolatos következményét említjük.

2.9. Állítás: *Ha $L_1, L_2 \subseteq \mathcal{R}^d$ alterek, továbbá $L_1 \subseteq L_2$, és $\dim L_1 = \dim L_2$, akkor $L_1 = L_2$.*

Bizonyítás: 2.7 szerint L_1 egy bázisa kiegészíthető L_2 egy bázisává. A dimenziók megegyeznek, így L_1 fenti bázisa rögtön L_2 bázisa is lesz. $\square$

Ha $L_1, L_2 \subseteq \mathcal{R}^d$ alterek, akkor $\mathrm{lin}\,(L_1 \cup L_2) = L_1 + L_2$, és persze ($L_1 \cap L_2$ altér) $\mathrm{lin}\,(L_1 \cap L_2) = L_1 \cap L_2$. A következő állítás szerint e négy altér dimenziói közül elég hármat ismerni.

2.10. Állítás: *Legyen $L_1, L_2 \subseteq \mathcal{R}^d$ két tetszőleges altér. Ekkor*

$$\dim(L_1 + L_2) + \dim(L_1 \cap L_2) = \dim L_1 + \dim L_2.$$

Bizonyítás: Egészítsük ki az $L_1 \cap L_2$ altér egy bázisát az L_1, illetve az L_2 alterek egy-egy bázisává. Könnyen belátható, hogy a két bázis úniója az $L_1 + L_2$ altér bázisa lesz. $\square$

Hasonlóképpen egy $L \subseteq \mathcal{R}^d$ altér dimenziójának ismeretében könnyen meghatározható $L^\perp$ dimenziója, és viszont.

2.11. Tétel: *Ha $L \subseteq \mathcal{R}^d$ altér, akkor*

$$\dim(L^\perp) = d - \dim L.$$

Bizonyítás: Legyen $B \in \mathcal{R}^{d \times n}$ olyan mátrix, amelyre $L = \mathrm{Im}\, B$. Ekkor $L^\perp = \mathrm{Ker}\, B^T$; továbbá $\dim L = \mathrm{r}\,(B)$, és $\dim L^\perp = d - \mathrm{r}\,(B)$. $\square$

Most újra bizonyíthatjuk első dualitástételünket.

2.12. Tétel: *Ha $L \subseteq \mathcal{R}^d$ altér, akkor $L^{\perp\perp} = L$.*

Bizonyítás: Nyilvánvaló, hogy $L \subseteq L^{\perp\perp}$, és mivel 2.11-ből adódóan e két altér dimenziója megegyezik, azért 2.9 szerint $L = L^{\perp\perp}$ is teljesül. $\square$

2.13. Tétel: *Ha $L_1, L_2 \subseteq \mathcal{R}^d$ alterek, akkor*
a) $(L_1 + L_2)^\perp = L_1^\perp \cap L_2^\perp$;
b) $(L_1 \cap L_2)^\perp = L_1^\perp + L_2^\perp$.

A megfelelő állítások igazak akkor is, ha kettő helyett több altér szerepel bennük.

Bizonyítás: A tétel a) része a $^{\perp}$-képzés tartalmazás-fordításának egyszerű következménye: például $L_1 \subseteq L_1 + L_2$, és így $(L_1 + L_2)^{\perp} \subseteq L_1^{\perp}$.

A tétel b) része a tétel a) részének és 2.12-nek segítségével igazolható.

$$(L_1 \cap L_2)^{\perp} = (L_1^{\perp\perp} \cap L_2^{\perp\perp})^{\perp} = (L_1^{\perp} + L_2^{\perp})^{\perp\perp} = L_1^{\perp} + L_2^{\perp}.$$

(Figyeljük meg, hogy milyen hasznos az $L^{\perp\perp} = L$ dualitástétel!)

A kettőnél több altér esete hasonlóan is elintézhető (csak jelölésben bonyolultabb), itt mégis a diagonális alteres fogást alkalmazzuk, ezt gyakorlandó.

Az állítás a) része ismét könnyű, b) részéhez legyen D a diagonális altér, $L := L_1 \times \ldots \times L_k$. Ekkor a két altérre vonatkozó b) állítás szerint

$$(D \cap L)^{\perp} = D^{\perp} + L^{\perp},$$

ahol

$$\begin{aligned} (D \cap L)^{\perp} &= \cup \left\{ \{a_1\} \times \ldots \times \{a_k\} \subseteq \mathcal{R}^{dk} : \sum a_i \in \left(\cap_{i=1}^{k} L_i\right)^{\perp} \right\}, \\ D^{\perp} &= \cup \left\{ \{a_1\} \times \ldots \times \{a_k\} \subseteq \mathcal{R}^{dk} : \sum a_i = 0 \right\}, \\ L^{\perp} &= L_1^{\perp} \times \ldots \times L_k^{\perp}. \end{aligned}$$

Szorozzuk meg ezt az egyenlőséget balról az $(E, \ldots, E) \in \mathcal{R}^{d \times (dk)}$ mátrixszal. Máris adódik a kívánt b) állítás. $\square$

A következő tétel egy fontos lineáris transzformáció, a vetítés definiálását teszi lehetővé.

2.14. Tétel: *Tetszőleges $L \subseteq \mathcal{R}^d$ altér esetén minden $\mathcal{R}^d$-beli vektor pontosan egyféleképpen áll elő, mint egy L-beli és egy $L^{\perp}$-beli vektor összege.*

Bizonyítás: Mivel nyilván $L \cap L^{\perp} = \{0\}$, azért 2.12-ből és 2.13-ból adódik, hogy $L + L^{\perp} = \mathcal{R}^d$.

Tegyük fel most, hogy $x_1, x_2 \in L$, $a_1, a_2 \in L^{\perp}$, és $x_1 + a_1 = x_2 + a_2$. Ekkor

$$x_1 - x_2 = a_2 - a_1 \in L \cap L^{\perp} = \{0\},$$

amiből az előállítás egyértelműsége is következik. $\square$

Ha egy az $\mathcal{R}^n$ téren értelmezett, az $\mathcal{R}^m$ térbe vezető leképezés minden $x \in \mathcal{R}^n$ vektorhoz az $Ax \in \mathcal{R}^m$ vektort rendeli hozzá, ahol $A \in \mathcal{R}^{m\times n}$, akkor a leképezést **lineáris leképezés**nek nevezzük, és A-val jelöljük, akárcsak a neki megfelelő mátrixot. Az $m = 1$ esetben **lineáris függvény**ről, az $m = n$ esetben **lineáris transzformáció**ról beszélünk. A $T : \mathcal{R}^d \to \mathcal{R}^d$ transzformációt tehát lineáris transzformációnak nevezzük, ha előáll

$$T(x) = Nx \ (x \in \mathcal{R}^d)$$

alakban valamely $N \in \mathcal{R}^{d\times d}$ mátrix esetén. Ha még az is teljesül, hogy az N mátrix invertálható, akkor a T transzformációt **invertálható lineáris transzformáció**nak nevezzük. Invertálható lineáris transzformáció inverze is invertálható lineáris transzformáció: éppen

$$T^{-1}(y) = N^{-1}y \ (y \in \mathcal{R}^d)$$

alakú. Fontos példa lineáris transzformációra az altérre történő vetítés.

Ha $x \in L$, és $a \in L^\perp$, akkor azt mondjuk, hogy az x vektor az $x + a$ vektor **vetület**e az L altérre. Könnyen belátható, hogy a vetületképzés (a **vetítés**) lineáris transzformáció, vagyis ekvivalens egy V mátrixszal való balról szorzással, ahol tehát

$$V(x + a) = x \ (x \in L,\, a \in L^\perp).$$

A V mátrixot ekkor az L altérre való vetítésnek megfelelő **vetítő mátrix**nak nevezzük. Tetszőleges $A \in \mathcal{R}^{m\times n}$ mátrix esetén az $\operatorname{Im} A$ altérre való vetítésnek megfelelő vetítő mátrix $V = AA^\dagger$, míg az $\operatorname{Im} A^T$ altér esetében ez a mátrix $V = A^\dagger A$. (Ennek az állításnak a megjegyzését segítheti a könnyen ellenőrizhető $\operatorname{Im} A^\dagger = \operatorname{Im} A^T$ azonosság.) Ebből már látszik, hogy a vetítő mátrixok megegyeznek transzponáltjukkal (vagyis szimmetrikusak) és négyzetükkel is (vagyis idempotensek). Megfordítva is; ha V egy idempotens, szimmetrikus mátrix, akkor $V = V^\dagger$, és így V az $\operatorname{Im} V$ altérre való vetítésnek megfelelő vetítő mátrix. A vetítő mátrixok **pozitív szemidefinit** mátrixok, vagyis $x^TVx \geq 0$ $(x \in \mathcal{R}^d)$ teljesül rájuk. (Hiszen $x^TVx = ||Vx||^2$.) Ha a V mátrix az L altérre való vetítésnek megfelelő vetítő mátrix, akkor az $L^\perp$ altérre való vetítésnek megfelelő vetítő mátrix $E - V$ lesz. A vetítő mátrixok norma-nemnövelők: $x \in \mathcal{R}^d$ esetén $||Vx|| \leq ||x||$. (Hiszen $||x||^2 - ||Vx||^2 = x^T(E - V)x$.)

A lineáris transzformációkkal kapcsolatos másik alapvető észrevétel a következő.

2.15. Tétel: *Ha $v_1, \dots, v_l \in \mathcal{R}^d$ és $v'_1, \dots, v'_l \in \mathcal{R}^d$ két lineárisan független vektorrendszer, akkor létezik $N \in \mathcal{R}^{d\times d}$ invertálható mátrix úgy, hogy*

$$Nv_1 = v'_1, \dots, Nv_l = v'_l.$$

Bizonyítás: Alkossunk a $v_1, \dots, v_l$, illetve a $v'_1, \dots, v'_l$ vektorokból, mint oszlopvektorokból, egy $A_1 \in \mathcal{R}^{d\times l}$, illetve egy $A_2 \in \mathcal{R}^{d\times l}$ mátrixot. 1.18 szerint léteznek $N_1, N_2 \in \mathcal{R}^{d\times d}$ invertálható mátrixok és $\pi_1, \pi_2 : \{1, \dots, d\} \to \{1, \dots, d\}$ permutációk úgy, hogy

$$P^T_{\pi_1} N_1 A_1 = (E, 0)^T = P^T_{\pi_2} N_2 A_2.$$

Ekkor

$$N := N_2^{-1} P_{\pi_2} P^T_{\pi_1} N_1$$

megfelel. □

Most rátérünk az általános ($Ax = b$ alakú) lineáris egyenletrendszerek megoldáshalmazai szerkezetének vizsgálatára.

Az $M \subseteq \mathcal{R}^d$ halmazt **affin halmaz**nak nevezzük, ha két elemével együtt azok 1-összegű lineáris kombinációit (a két elem által meghatározott **egyenes**t) is tartalmazza, vagyis

$$x_1, x_2 \in M,\ \lambda_1, \lambda_2 \in \mathcal{R},\ \lambda_1 + \lambda_2 = 1 \text{ esetén } \lambda_1 x_1 + \lambda_2 x_2 \in M.$$

Könnyen belátható, hogy az alterek éppen az origót tartalmazó affin halmazok. Valóban, a nemtriviális irányhoz legyen M affin halmaz, $0 \in M$. Ekkor $x \in M$, $\lambda \in \mathcal{R}$ esetén

$$\lambda x = (1 - \lambda) \cdot 0 + \lambda x \in M,$$

tehát $\mathcal{R}M \subseteq M$. Ha $x_1, x_2 \in M$, akkor

$$x_1 + x_2 = 2\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) \in 2M \subseteq M,$$

tehát $M + M \subseteq M$ is teljesül. Most már könnyen belátható, hogy

2.16. Tétel: *A nemüres, affin halmazok éppen az eltolt alterek. Minden nemüres, affin halmaz pontosan egy altér eltoltjaként áll elő, amely alteret úgy kapjuk, hogy az affin halmazból kivonjuk egy tetszőleges elemét.*

Bizonyítás: Először a tétel első felét igazoljuk. Könnyen belátható, hogy affin halmaz eltoltja is affin halmaz. Emiatt az affin halmazból kivonva egy elemét, az origót tartalmazó affin halmazt, vagyis alteret kapunk, amit ugyanezzel az elemmel eltolva visszakapjuk az eredeti affin halmazt.

A tétel második feléhez megmutatjuk, hogy M nem lehet két különböző altér eltoltja, vagyis $x_1, x_2 \in M$ esetén $M - x_1$ és $M - x_2$ ugyanaz az altér. Mivel $x_2 - x_1 \in M - x_1$, és $M - x_1$ a fentiek szerint altér, azért $x_1 - x_2 \in M - x_1$ is teljesül. Ebből pedig

$$M - x_1 = M - x_1 + x_1 - x_2 = M - x_2$$

adódik. □

Az $M \subseteq \mathcal{R}^d$ affin halmazból egy elemét kivonva nyert (egyértelmű) alteret az M affin halmazzal **párhuzamos altér**nek nevezzük, és $\operatorname{par} M$-mel jelöljük. (A "par" az angol "parallel" szó rövidítése.) Két affin halmaz **párhuzamos**, ha ugyanaz a párhuzamos alterük.

2.17. Tétel: *Az affin halmazok éppen a lineáris egyenletrendszerek megoldáshalmazai.*

Bizonyítás: Könnyen belátható, hogy az $Ax = b$ rendszer megoldáshalmaza (ha nemüres) éppen $p + \operatorname{Ker} A$, ahol p tetszőleges megoldás. Tehát a megoldáshalmaz eltolt altér, és így affin halmaz.

Megfordítva, ha M nemüres, affin halmaz, akkor tetszőleges $p \in M$ esetén az $L := M - p$ halmaz altér, és $M = p + L$. Legyen A olyan mátrix, amelyre $L = \operatorname{Ker} A$ teljesül, ekkor a fentiek szerint

$$M = \{x : Ax = Ap\},$$

tehát M egy lineáris egyenletrendszer megoldáshalmaza. □

Nyilván affin halmazok tetszőleges rendszerének metszete is affin halmaz, így tetszőleges $S \subseteq \mathcal{R}^d$ esetén létezik egyértelműen az S halmazt tartalmazó legszűkebb affin halmaz, S **affin burka**,

$$\operatorname{aff} S := \cap\{M : S \subseteq M, M \text{ affin halmaz}\}.$$

2.1 mintájára, könnyen belátható, hogy

2.18. Állítás: *Nemüres $S \subseteq \mathcal{R}^d$ esetén*

$$\operatorname{aff} S = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathcal{N}, \lambda_i \in \mathcal{R}, \sum_{i=1}^{k} \lambda_i = 1, x_i \in S \ (i = 1, \ldots, k) \right\},$$

az S-beli elemek **affin kombináció***inak halmaza. (Az üres halmaz affin, így affin burka önmaga.)* □

Bevezethető az affin generáló részhalmaz, az affin függetlenség, az affin bázis, az affin dimenzió fogalma. A megfelelő állításokat direkt bizonyításuk helyett homogenizációval visszavezetjük a lineáris párjukra vonatkozó hasonló állításokra.

Vezessük be az alábbi jelöléseket:

$$\begin{aligned} &M(L) := \left\{ x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq L \right\} && (L \subseteq \mathcal{R}^{d+1} \text{ altér}), \\ &L(M) := \operatorname{lin}\left(\{1\} \times M\right) && (M \subseteq \mathcal{R}^d \text{ affin halmaz}). \end{aligned}$$

A jelöléseknek megfelelően $M(L)$ affin halmaz, és $L(M)$ altér. Mint az a következő egyszerű állítás b) részéből látszik, tetszőleges $M \subseteq \mathcal{R}^d$ affin halmaz esetén $M = M(L(M))$. Ezek szerint az affin halmazok éppen az $M(L)$ alakú halmazok, ahol L az $\mathcal{R}^{d+1}$-beli altereken fut. Ezért affin halmazokra vonatkozó állításokat sokszor vissza lehet vezetni az alterekre vonatkozó, egyszerűbb állításokra. Ezt a fogást (illetve általánosabb változatait) hívjuk homogenizációnak. (Megjegyezzük, hogy affin halmazokra vonatkozó állítások 2.16 segítségével is visszavezethetők alterekre vonatkozó állításokra, de 2.16 nem általánosítódik konvex halmazokra, ellentétben a homogenizációval.)

2.19. Állítás: *Tetszőleges $S, S_1, S_2 \subseteq \mathcal{R}^d$ halmazok esetén*
a) $\operatorname{lin} S = \operatorname{lin} \operatorname{aff} S$*, speciálisan*

$$\operatorname{lin}\left(\{1\} \times S\right) = \operatorname{lin}\left(\{1\} \times \operatorname{aff} S\right);$$

b)

$$\left\{ x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq \operatorname{lin}\left(\{1\} \times S\right) \right\} = \operatorname{aff} S;$$

c)

$$\operatorname{lin}\left(\{1\} \times S_1\right) \subseteq \operatorname{lin}\left(\{1\} \times S_2\right) \Longleftrightarrow \operatorname{aff} S_1 \subseteq \operatorname{aff} S_2.$$

(Az állítás c) része persze érvényben marad, ha benne mindkét $\subseteq$ jel helyébe $=$-t, vagy $\subset$-t írunk.)

Bizonyítás: a) A $\text{lin}\, S \subseteq \text{lin}\,\text{aff}\, S$ tartalmazás $S \subseteq \text{aff}\, S$, míg a $\text{lin}\, S \supseteq \text{lin}\,\text{aff}\, S$ tartalmazás $\text{aff}\, S \subseteq \text{lin}\, S$ következménye. A második igazolandó halmazegyenlőség hasonlóan bizonyítható.

b) könnyű következménye annak, hogy $\text{lin}\, S$ [$\text{aff}\, S$] S elemeinek lineáris [affin] kombinációiból áll.

c) a) és b) következménye. □

2.20. Állítás: *Legyen $0 \neq a \in \mathcal{R}^d$, $0 \neq \beta \in \mathcal{R}$, az $a^T x = \beta$ egyenlet megoldáshalmazát jelölje H. Tegyük fel, hogy a H affin halmaz összemetsz egy $L \subseteq \mathcal{R}^d$ altérrel. Ekkor*

$$L = \text{lin}\,(L \cap H).$$

Bizonyítás: Természetesen $\text{lin}\,(L \cap H) \subseteq L$. A fordított tartalmazáshoz legyen $x \in L$, és tegyük fel először, hogy $a^T x \neq 0$. Ekkor $\frac{\beta}{a^T x} x \in L \cap H$, és így

$$x = \frac{a^T x}{\beta} \left(\frac{\beta}{a^T x} x \right) \in \text{lin}\,(L \cap H).$$

Tegyük fel most, hogy $a^T x = 0$, és legyen $x_0 \in L \cap H$ tetszőleges elem. Ekkor az

$$x_1 := x_0 + (1/2)x \text{ és } x_2 := -x_0 + (1/2)x$$

vektorokra teljesül, hogy $a^T x_1 \neq 0 \neq a^T x_2$, így a már elintézett eset szerint $x_1, x_2 \in \text{lin}\,(L \cap H)$. De akkor $x = x_1 + x_2$ is $\text{lin}\,(L \cap H)$-beli. □

2.21. Tétel: *Tetszőleges $L \subseteq \mathcal{R}^{d+1}$ altér és $M \subseteq \mathcal{R}^d$ affin halmaz esetén teljesül, hogy*
a) $M(L(M)) = M$;
b) $L(M(L)) = L$ valahányszor $M(L) \neq \emptyset$.

Bizonyítás: A tétel a) része 2.19 b), b) része pedig 2.20 egyszerű következménye (legyen $a := e_1 \in \mathcal{R}^{d+1}, \beta := 1$). □

Egy $S \subseteq \mathcal{R}^d$ halmazt az $M \subseteq \mathcal{R}^d$ affin halmaz **(affin) generáló részhalmaz**ának hívunk, ha $\text{aff}\, S = M$. Ilyen például az egész M. Az S halmaz pontosan akkor lesz az M affin halmaz affin generáló részhalmaza, ha az $\{1\} \times S$ halmaz lineáris generáló részhalmaza az $L(M)$ altérnek.

2.22. Állítás: *Legyen $S \subseteq \mathcal{R}^d$, ekkor az alábbi állítások ekvivalensek:*
a) $S' \subset S$ esetén $\operatorname{aff} S' \subset \operatorname{aff} S$;
b) minden $x \in S$ esetén $x \notin \operatorname{aff}(S \setminus \{x\})$;
c) S-beli elemek nulla együtthatóösszegű lineáris kombinációjaként csak úgy állhat elő a tér 0 eleme, ha a kombinációban minden együttható nulla;
d) minden $x \in S$ esetén $(S \setminus \{x\}) - x$ lineárisan független;
e) S üres, vagy létezik $x \in S$ úgy, hogy $(S \setminus \{x\}) - x$ lineárisan független;
f) $\{1\} \times S$ lineárisan független.

Bizonyítás: $a) \iff f)$, $b) \iff f)$, $c) \iff f)$ rendre 2.4 a), b), c)-ből következik 2.19 segítségével.

$c) \Rightarrow d) \Rightarrow e) \Rightarrow c)$ pedig a lineáris függetlenség 2.4 c) jellemzésének egyszerű következménye. □

Ha a 2.22-beli a), b), …, f) állítások egyike (és akkor mindegyike) teljesül S-re, akkor S-et **affin független** halmaznak nevezzük.

2.23. Állítás: *Adott $\emptyset \neq M \subseteq \mathcal{R}^d$ affin halmaz, $S \subseteq \mathcal{R}^d$ esetén az alábbi állítások ekvivalensek:*
a) S az M minimális affin generáló részhalmaza;
b) S az M affin független affin generáló részhalmaza;
c) S maximális M-beli affin független halmaz;
d) $S \neq \emptyset$, és minden $x \in S$ esetén $(S \setminus \{x\}) - x$ a $\operatorname{par} M$ bázisa;
e) létezik $x \in S$ úgy, hogy $(S \setminus \{x\}) - x$ a $\operatorname{par} M$ bázisa;
f) $\{1\} \times S$ az $L(M)$ bázisa.

Bizonyítás: $a) \iff f)$, $b) \iff f)$, $c) \iff f)$ rendre 2.5 a), b), c)-ből következik 2.19 segítségével.

$b) \Rightarrow d) \Rightarrow e) \Rightarrow b)$-ből például $b) \Rightarrow d)$-t bizonyítjuk: Ha S affin független halmaz, és $\operatorname{aff} S = M$, akkor 2.22 d) szerint az $(S \setminus \{x\}) - x$ halmaz lineárisan független minden $x \in S$ esetén. Továbbá

$$\begin{aligned} \operatorname{par} M &= M - x = (\operatorname{aff} S) - x = \\ &= \operatorname{aff}(S - x) = \operatorname{lin}((S \setminus \{x\}) - x) \ (x \in S), \end{aligned}$$

ugyanis az $\operatorname{aff}(S - x)$ affin halmaz tartalmazza az origót, és így altér. □

Ha a 2.23-beli a), b), …, f) állítások egyike (és akkor mindegyike) teljesül S-re, akkor az S halmazt az M affin halmaz **affin bázis**ának nevezzük.

Affin bázis mindig létezik: 2.23 e) szerint ha a $\operatorname{par} M$ altér bázisát és még az origót eltoljuk egy M-beli elemmel, akkor az M affin halmaz affin bázisát kapjuk. 2.23 f) segítségével belátható 2.7 analogonja is.

Az affin bázis elemszáma-1, vagyis a par M altér dimenziója az M affin halmaz **dimenzió**ja. A 0-, 1-, 2-, illetve $(n-1)$-dimenziós affin halmazok a **pont**ok, **egyenes**ek, **sík**ok, illetve **hipersík**ok. Az üres halmaz dimenzióját -1-nek szokás definiálni. Általában $S \subseteq \mathcal{R}^d$ esetén az S halmaz **dimenzió**ja legyen a

$$\operatorname{par} S := \operatorname{par}(\operatorname{aff} S) = (\operatorname{aff} S) - (\operatorname{aff} S)$$

altér dimenziója.

2.24. Állítás: *Az $(n-r)$-dimenziós $M \subseteq \mathcal{R}^n$ affin halmazok éppen az $M = \{x : Ax = b\}$ alakú nemüres halmazok, ahol A egy r-rangú mátrix.*

Bizonyítás: 2.17 bizonyításának mintájára, 2.6-ból adódik. □

Ha egy az $\mathcal{R}^n$ téren értelmezett, az $\mathcal{R}^m$ térbe vezető leképezés minden $x \in \mathcal{R}^n$ vektorhoz az $Ax - b \in \mathcal{R}^m$ vektort rendeli hozzá, ahol $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, akkor a leképezést **affin leképezés**nek nevezzük. Az $m = 1$ esetben **affin függvény**ről, az $m = n$ esetben **affin transzformáció**ról beszélünk. Tehát a $T : \mathcal{R}^d \to \mathcal{R}^d$ transzformációt affin transzformációnak nevezzük, ha előáll

$$T(x) = Nx + v \ (x \in \mathcal{R}^d)$$

alakban valamely $N \in \mathcal{R}^{d\times d}$ mátrix és $v \in \mathcal{R}^d$ vektor esetén. Ha még az is teljesül, hogy az N mátrix invertálható, akkor a T transzformációt **invertálható affin transzformáció**nak, vagy rövidebben **átkoordinátázás**nak nevezzük. Átkoordinátázás inverze is átkoordinátázás: éppen

$$T^{-1}(y) = N^{-1}y - N^{-1}v \ (y \in \mathcal{R}^d)$$

alakú.

2.25. Tétel: *Ha adott két azonos elemszámú, affin független halmaz, S és S', akkor létezik invertálható N mátrix és egy v vektor úgy, hogy*

$$Nx_i + v = x_i' \ (i = 0, 1, \ldots, k)$$

teljesül, ahol az x_i, illetve x_i' vektorok rendre S, illetve S' elemei.

Bizonyítás: Ha S és S' affin független halmazok (vagyis az $(S \setminus \{x_0\}) - x_0$ és $(S' \setminus \{x_0'\}) - x_0'$ halmazok lineárisan függetlenek), akkor 2.15 szerint létezik N invertálható mátrix úgy, hogy

$$N(x_i - x_0) = x_i' - x_0' \ (i = 1, \ldots, k)$$

teljesüljön. Ekkor ez az N mátrix és a $v := x_0' - Nx_0$ vektor megfelel. □

3. Poliéder kúpok, poliéderek

Ebben a fejezetben egyenletrendszerek helyett egyenlőtlenség-rendszerek megoldáshalmazait vizsgáljuk. Ezek a halmazok kevésbé strukturáltak, mint az affin halmazok, mégis sok eredmény átmenthető. A Fredholm-féle alternatívatételnek például a Farkas-lemma felel meg, melynek sok ekvivalens alakja közül az egyik, hogy az $Ax = b, x \geq 0$ rendszer pontosan akkor megoldható, ha az $A^T y \geq 0, b^T y < 0$ rendszer nem az. Bizonyítható például a Gauss–Jordan-eliminációnak megfelelő szimplex módszer segítségével.

Itt csak vázlatosan írunk e módszerről. Az első fejezet jelöléseit használjuk. Tekintsük az összes olyan N invertálható mátrix S halmazát, amelyre az $N(A,b)$ mátrix Hermite-féle normálalakú, és a b vektor oszlopindexe nincs benne a $\{j_1,\ldots,j_r\}$ halmazban. Az S halmaz pontosan akkor nem üres, ha az $Ax = b$ rendszer megoldható, és ekkor $r = \mathrm{r}\,(A) = \mathrm{r}\,(A,b)$. Egy ilyen N mátrixot megtalálhatunk a Gauss–Jordan-elimináció segítségével. A módszer az alábbi két észrevételen alapszik:

– Ha $Nb \geq 0$, akkor $x := p$ megoldása az $Ax = b, x \geq 0$ rendszernek;
– Ha létezik olyan i sorindex, amelyre az NA mátrix i-edik sorvektora, $e_i^T NA \geq 0$, és az Nb vektor i-edik eleme, $e_i^T Nb < 0$, akkor $y := N^T e_i$ az $A^T y \geq 0, b^T y < 0$ rendszer megoldása.

Az algoritmushoz ezek után már csak egy olyan szabály kell, amely meghatározza, hogy ha a fenti két lehetőség egyike sem állna fenn, akkor mi legyen a következő vizsgálandó $N \in S$ invertálható mátrix, ráadásul olyan, hogy ez a bolyongás S elemein véges lépésben a fenti két lehetőség egyikével befejeződik. Ilyen szabály például a Bland-szabály (lásd [23]).

Rögtön a Caratheodory-tétel is látszana (mely szerint ha az $Ax = b$, $x \geq 0$ rendszer megoldható, akkor olyan megoldása is létezik, hogy az x megoldás pozitív elemeinek megfelelő A-beli oszlopvektorok lineárisan függetlenek — a p vektor ilyen megoldás), továbbá a lineáris egyenlőtlenségek alaptétele (mely szerint ha az $Ax = b, x \geq 0$ rendszer nem megoldható, akkor létezik olyan y vektor, amelyre $A^T y \geq 0, b^T y < 0$, és y merőleges $\mathrm{r}\,(A,b) - 1$ lineárisan független A-beli oszlopvektorra — ha $Ax = b$ nem megoldható, akkor a Fredholm-féle alternatívatételben szereplő, különben a fenti y vektor megfelel, a Kronecker-Capelli-tételt is használva). Terjedelemkorlát miatt mi mégsem követjük ezt az algoritmikus utat.

A Farkas-lemmát általánosabb alakban, indukcióval bizonyítjuk. Itt megfigyelhető egy jelenség, amellyel később is találkozunk majd (lásd 4.23): általánosabbat néha könnyebb bizonyítani, mert a bizonyítás kiléphet a szűkebb keretek közül.

3.1. Tétel: (Farkas-lemma) *Tetszőleges* $A_1 \in \mathcal{R}^{m\times n_1}, A_2 \in \mathcal{R}^{m\times n_2}$ *mátrixok és* $b \in \mathcal{R}^m$ *vektor esetén az alábbi rendszerek közül pontosan az egyik oldható meg:*

$$A_1x_1 + A_2x_2 = b, x_1 \geq 0, \textit{ illetve } A_1^Ty \geq 0, A_2^Ty = 0, b^Ty < 0.$$

Bizonyítás: A tételben szereplő első rendszerre röviden (P)-ként, a második rendszerre röviden (D)-ként fogunk hivatkozni. (A "primál" és a "duál" szavak kezdőbetűi.)

Az állítást n_1 (az A_1 mátrix oszlopszáma) szerinti indukcióval igazoljuk. Azt az esetet, mikor $n_1 = 0$, elintézi a Fredholm-féle alternatívatétel. Tegyük fel most, hogy az A_1 mátrixnak van legalább egy oszlopa, és kisebb oszlopszámú A_1 mátrixok esetén igaz az állítás.

Először is hagyjuk el az A_1 mátrixból utolsó oszlopát (ennek jele legyen a), az így kapott mátrixot jelölje A_1'. Az indukciófeltétel szerint az alábbi rendszerek közül pontosan az egyik megoldható:

$$A_1'x_1' + A_2x_2 = b, x_1' \geq 0, \text{ illetve } A_1'^Ty \geq 0, A_2^Ty = 0, b^Ty < 0.$$

Az első rendszerre a rövidség kedvéért (P')-ként hivatkozunk, a másodikra (D')-ként. Vegyük észre, hogy
– ha az x_1', x_2 vektorok a (P') rendszer megoldását adják, akkor az $x_1^T := (x_1'^T, 0)$ jelöléssel x_1, x_2 a (P) rendszer megoldása lesz;
– ha az y vektor a (D') rendszer megoldása, és $a^Ty \geq 0$ is teljesül, akkor y a (D) rendszer megoldása is.

Másodszor toldjuk hozzá az a oszlopot az A_2 mátrixhoz első oszlopként, az így kapott mátrixot jelölje A_2'. Jelölje továbbá A_1' ugyanazt a mátrixot, mint az előbb. Az indukciófeltétel szerint az alábbi rendszerek közül pontosan az egyik megoldható:

$$A_1'x_1' + A_2'x_2' = b, x_1' \geq 0, \text{ illetve } A_1'^Ty \geq 0, A_2'^Ty = 0, b^Ty < 0.$$

Az első rendszerre a rövidség kedvéért (P'')-ként hivatkozunk, a másodikra (D'')-ként. Vegyük észre, hogy

– ha az y vektor a (D'') rendszer megoldása, akkor y a (D) rendszer megoldása is;
– ha az x_1', x_2' vektorok a (P'') rendszer megoldását adják, és x_2' első eleme nemnegatív, akkor az $(x_1^T, x_2^T) := (x_1'^T, x_2'^T)$ jelöléssel x_1, x_2 a (P) rendszer megoldása lesz.

Már látjuk, hogy a (P) és (D) rendszerek valamelyike megoldható, kivéve azt az esetet, mikor létezik a (P'') rendszernek olyan x_1', x_2' megoldása, hogy x_2' első eleme negatív, továbbá létezik a (D') rendszernek olyan y megoldása, amelyre $a^T y < 0$. Ez az eset viszont nem fordulhat elő, különben

$$\begin{aligned} 0 > b^T y &= (A_1' x_1' + A_2' x_2')^T y = \\ &= (\underbrace{x_1'}_{\geq 0})^T (\underbrace{A_1'^T y}_{\geq 0}) + \underbrace{(x_2')_1}_{<0} \cdot \underbrace{(a^T y)}_{<0} + \textstyle\sum_{j=2}^{n_2+1} (x_2')_j \cdot \underbrace{(((A_2')_j)^T y)}_{=0} > 0 \end{aligned}$$

lenne.

Az pedig, hogy a (P) és (D) rendszerek valamelyike nem megoldható, nyilvánvaló, különben a két megoldással

$$0 > b^T y = (A_1 x_1 + A_2 x_2)^T y = \underbrace{x_1^T}_{\geq 0} (\underbrace{A_1^T y}_{\geq 0}) + x_2^T (\underbrace{A_2^T y}_{=0}) \geq 0$$

lenne. □

A Farkas-lemmának számos ekvivalens alakja van. Például az eredeti változata (a továbbiakban ezt hívjuk Farkas-lemmának) az alábbi két rendszerről szólt

$$Ax = b, x \geq 0, \text{ illetve } A^T y \geq 0, b^T y < 0,$$

de egyszerű fogások segítségével (mint például a vektor pozitív és negatív része különbségeként való felírása, vagy eltérésváltozók bevezetése) az alábbi két rendszerről szóló általános alak is igazolható

$$Ax = b, Bx \geq c, \text{ illetve } A^T y + B^T z = 0, z \leq 0, b^T y + c^T z < 0.$$

Hasonlóképpen egyszerűen vezethető le egy olyan rendszer megoldhatóságát karakterizáló tétel, amelyben már szigorú egyenlőtlenségek is vannak (lásd [27]). Például Stiemke-tétele, mely szerint pontosan akkor létezik pozitív vektor egy mátrix nullterében, ha a mátrix transzponáltjának képterében csak az origó a nemnegatív vektor. Ezekre a kérdésekre később visszatérünk (lásd 3.32 és 6.8).

Először a Farkas-lemma néhány egyszerű következményét említjük, azután a Caratheodory-tételt is használva belátjuk a Farkas-lemma egy általánosítását, a lineáris egyenlőtlenségek alaptételét.

A $K \subseteq \mathcal{R}^d$ halmazt **konvex kúp**nak nevezzük, ha az origót, továbbá bármely két elemének nemnegatív lineáris kombinációit is tartalmazza, vagyis $0 \in K$, és

$$x_1, x_2 \in K,\ \lambda_1, \lambda_2 \geq 0 \text{ esetén } \lambda_1 x_1 + \lambda_2 x_2 \in K.$$

Másképpen megfogalmazva egy $K \subseteq \mathcal{R}^d$ halmaz pontosan akkor konvex kúp, ha teljesül rá, hogy nemüres, és $\mathcal{R}_+ K \subseteq K$ (vagyis K **kúp**), továbbá $K + K \subseteq K$ (és akkor $\mathcal{R}_+ K = K$, és $K + K = K$, $1 \in \mathcal{R}_+$, és $0 \in K$ miatt).

Konvex kúp például minden altér (ezek éppen azok a konvex kúpok, melyek (-1)-szerese önmaga), vagy a nemnegatív vektorok $\mathcal{R}_+^d$ halmaza, a **nemnegatív ortáns**. Könnyen megmutatható, hogy ha $A \in \mathcal{R}^{m \times d}$, $B \in \mathcal{R}^{d \times n}$, továbbá $K_1 \subseteq \mathcal{R}^m$, $K_2 \subseteq \mathcal{R}^n$ konvex kúpok, akkor $A^{-1}(K_1)$ és BK_2 is konvex kúp, speciálisan az A mátrix **nemnegatívter**e és a B mátrix **nemnegatívkép-ter**e,

$$\operatorname{Ker}_+ A := A^{-1}(\mathcal{R}_+^m) = \left\{x \in \mathcal{R}^d : Ax \geq 0\right\}, \text{ illetve}$$
$$\operatorname{Im}_+ B := B\mathcal{R}_+^n = \{By : y \in \mathcal{R}_+^n\}$$

konvex kúpok. Egy K konvex kúp lineáris és affin burka megegyezik (a 0 vektor K-beli), mindkettő $K - K$. Ezért

$$\operatorname{lin} \operatorname{Im}_+ B = \operatorname{Im} B, \text{ és } \dim \operatorname{Im}_+ B = \operatorname{r}(B).$$

Felmerül az a (két) kérdés, hogy megfordítva, minden $K \subseteq \mathcal{R}^d$ konvex kúp előáll-e $\operatorname{Ker}_+ A$, illetve $\operatorname{Im}_+ B$ alakban valamely $A \in \mathcal{R}^{m \times d}$, $B \in \mathcal{R}^{d \times n}$ mátrixok esetén. A válasz ezúttal nem, ugyanis belátjuk majd, hogy

$$(\operatorname{Ker}_+ A)^{**} = \operatorname{Ker}_+ A, \text{ és } (\operatorname{Im}_+ B)^{**} = \operatorname{Im}_+ B$$

is teljesül (3.5), míg egy K konvex kúp esetén $K^{**} = K$ pontosan akkor, ha K zárt (3.28).

Ez persze még nem zárja ki, hogy esetleg minden zárt, konvex kúp előálljon $\operatorname{Ker}_+ A$, illetve $\operatorname{Im}_+ B$ alakban valamely $A \in \mathcal{R}^{m \times d}$, $B \in \mathcal{R}^{d \times n}$ mátrixok esetén. Ez sincs így: belátjuk majd, hogy a $\operatorname{Ker}_+ A$, illetve $\operatorname{Im}_+ B$ alakú konvex kúpok bármely vetülete zárt (sőt poliéder kúp (3.25)), ugyanakkor példát mutatunk olyan zárt, konvex kúpra, amelynek van nemzárt vetülete (3.31).

Ami még a konvex kúpok általánosságban is teljesül az annyi, hogy a $\mathrm{Ker}_+ A$ alakú konvex kúpok (a **poliéder kúp**ok) éppen az $\mathrm{Im}_+ B$ alakú konvex kúpok (a **végesen generált kúp**ok) (3.14, 3.15).

Megemlítjük még, hogy a 7. fejezetben látunk majd szükséges és elégséges feltételt arra, hogy egy zárt, konvex kúp poliéder kúp legyen, ez azt követeli meg, hogy véges sok extremális részhalmaza legyen (7.12). (Ez az eredmény bizonyos értelemben bezárja a fenti kérdések sorát, amelyben az jönne, hogy ha a konvex kúp minden vetülete zárt, akkor poliéder kúp-e.)

Konvex kúpok tetszőleges rendszerének metszete nyilván konvex kúp, így egy adott $S \subseteq \mathcal{R}^d$ halmaz esetén létezik egyértelműen az S halmazt tartalmazó legszűkebb konvex kúp, S **konvex kúp burka**,

$$\mathrm{cone}\, S := \cap\{K : S \subseteq K,\ K \text{ konvex kúp}\}.$$

2.1 mintájára igazolható az alábbi

3.2. Állítás: *Tetszőleges nemüres $S \subseteq \mathcal{R}^d$ halmaz esetén teljesül, hogy*

$$\mathrm{cone}\, S = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathcal{N},\ \lambda_i \in \mathcal{R}_+,\ x_i \in S\ (i = 1, \ldots, k) \right\},$$

az S-beli elemek **nemnegatív** *vagy* **konvex kúp kombináció***inak a halmaza. Az üres halmaz konvex kúp burka a triviális $\{0\}$ konvex kúp.* □

Teljesül, hogy $0 \in S_1, S_2$ esetén

$$\mathrm{cone}\,(S_1 \times S_2) = (\mathrm{cone}\, S_1) \times (\mathrm{cone}\, S_2).$$

A konvex kúp burok képzés tartalmazástartó:

$$S_1 \subseteq S_2 \text{ esetén } \mathrm{cone}\, S_1 \subseteq \mathrm{cone}\, S_2.$$

Adott $S \subseteq \mathcal{R}^d$ halmaz esetén az S halmaz **adjungált konvex kúp**ja az összes S-beli vektorral legfeljebb 90 fokos szöget bezáró vektorok halmaza. (Két nemnulla vektor, a és x (radiánban mért) **szög**ének koszinuszát az $(a^T x)/(||a|| \cdot ||x||)$ képlet adja meg.) Az S halmaz adjungált konvex kúpjának jele S^*, eszerint

$$S^* := \left\{ a \in \mathcal{R}^d : a^T x \geq 0\ (x \in S) \right\}.$$

Könnyen belátható, hogy tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén S^* zárt, konvex kúp. Ha $S = K$ zárt, konvex kúp, akkor a K^* kúpot a K kúp **duális**

kúpjának is nevezik (vö. 3.28). Ha $S = L$ altér, akkor $S^* = L^\perp$. Teljesül, hogy $0 \in S_1, S_2$ esetén

$$(S_1 \times S_2)^* = S_1^* \times S_2^*.$$

Az adjungált konvex kúp képzése tartalmazás-fordító:

$$S_1 \subseteq S_2 \text{ esetén } S_1^* \supseteq S_2^*.$$

3.3. Állítás: *Tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén*

$$S^* = (\operatorname{cone} S)^*, \textit{ és } S^* = (\operatorname{cl} S)^*.$$

□

2.3 mintájára (csak a Fredholm-féle alternatívatétel helyett a Farkas-lemmát használva) igazolható, hogy

3.4. Tétel: *Tetszőleges $A \in \mathcal{R}^{m \times n}$ mátrix esetén*

$$(\mathrm{Im}_+ A)^* = \mathrm{Ker}_+ A^T, \textit{ és } (\mathrm{Ker}_+ A^T)^* = \mathrm{Im}_+ A.$$

□

A 3.4 tételből (a természetes

$$S^{**} := (S^*)^* \ (S \subseteq \mathcal{R}^d)$$

jelölést használva) azonnal következik az alábbi dualitástétel (lásd még 3.28):

3.5. Tétel: *Ha R poliéder kúp vagy végesen generált kúp, akkor*

$$R^{**} = R.$$

Speciálisan R zárt halmaz. □

Ha $K_1, K_2 \subseteq \mathcal{R}^d$ konvex kúpok, akkor az uniójukat tartalmazó legszűkebb konvex kúp egyszerűen a halmazösszegük, $K_1 + K_2$. Igaz továbbá, hogy poliéder kúpok metszete is poliéder kúp (hiszen

$$(\mathrm{Ker}_+ A_1) \cap (\mathrm{Ker}_+ A_2) = \mathrm{Ker}_+ (A_1^T, A_2^T)^T),$$

és végesen generált kúpok összege is végesen generált (hiszen

$$(\mathrm{Im}_+ B_1) + (\mathrm{Im}_+ B_2) = \mathrm{Im}_+ (B_1, B_2)).$$

A következő tétel b) részét (lásd még 3.29) csak poliéder kúpok esetén igazoljuk, bár a Farkas-lemma általános alakjának segítségével végesen generált kúpokra is bizonyítható lenne (lásd [27]). Viszont nemsokára úgyis belátjuk, hogy a végesen generált kúpok éppen a poliéder kúpok.

3.6. Tétel: *Ha $K_1, K_2 \subseteq \mathcal{R}^d$ tetszőleges konvex kúpok, akkor*
a) $(K_1 + K_2)^* = K_1^* \cap K_2^*$.
Ha K_1, K_2 poliéder kúpok, akkor
b) $(K_1 \cap K_2)^* = K_1^* + K_2^*$.
A megfelelő állítások igazak akkor is, ha kettő helyett több kúp szerepel bennük.

Bizonyítás: A tétel a) része, akárcsak 2.13-ban, a *-képzés tartalmazásfordításának egyszerű következménye.
Ha most $K_i = \mathrm{Ker}_+ A_i$ $(i = 1, 2)$, akkor

$$\begin{aligned}(K_1 \cap K_2)^* &= (\mathrm{Ker}_+ (A_1^T, A_2^T)^T)^* = \\ &= \mathrm{Im}_+ (A_1^T, A_2^T) = \mathrm{Im}_+ A_1^T + \mathrm{Im}_+ A_2^T = K_1^* + K_2^*.\end{aligned}$$

A több, mint két konvex kúp, illetve poliéder kúp esete hasonlóan intézhető el. $\square$

Caratheodory tételéhez szükségünk lesz a bázismegoldás fogalmára.

Legyen $A \in \mathcal{R}^{m \times n}, b \in \mathcal{R}^m$. Az x vektor **bázismegoldás**a az $Ax \geq b$ rendszernek, ha megoldása, és

$$\mathrm{r}\,({}_I A) = \mathrm{r}\,(A), \text{ ahol } I := \{i \in \{1, \ldots, m\} : {}_i a x = b_i\}.$$

(Itt ${}_i a$ az A mátrix i-edik sorvektorát, ${}_I A$ pedig az A mátrix I-beli indexű sorokból álló részmátrixát jelöli.)
Általában egy

$$Ax = b, Bx \geq c, Cx \leq d$$

rendszer bázismegoldását úgy definiálhatjuk, mint az

$$Ax \geq b, -Ax \geq -b, Bx \geq c, -Cx \geq -d$$

rendszer fenti értelemben vett bázismegoldását. Például

3.7. Állítás: *Az x megoldás pontosan akkor bázismegoldása az $Ax = b$, $x \geq 0$ rendszernek, ha az $(a_j : x_j > 0)$ vektorok lineárisan függetlenek.* $\square$

3.8. Állítás: *Az y megoldás pontosan akkor bázismegoldása az $A^T y \geq 0$, $b^T y = -1$ rendszernek, ha az y merőleges $\mathrm{r}(A, b) - 1$ lineárisan független A-beli oszlopvektorra.* □

Caratheodory tétele azt mondja ki, hogy egy megoldható egyenlőtlenség-rendszernek mindig létezik bázismegoldása is. Ennek igazolásához szükségünk lesz az alábbi lemmára:

3.9. Lemma: *Legyen $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$. Ha az x vektor az $Ax \geq b$ rendszer megoldása, de nem bázismegoldása, akkor létezik $z \in \mathcal{R}^n$ vektor úgy, hogy ${}_IAz = 0$, $Az \neq 0$. Még $Az \not\leq 0$ is feltehető — z vagy $-z$ ilyen vektor.*

Legyen $x(\lambda) := x + \lambda z$ $(\lambda \in \mathcal{R})$. Ekkor az

$$\{x(\lambda) : \lambda_1 \leq \lambda \leq \lambda_2,\ \lambda \in \mathcal{R}\}$$

halmaz minden pontja az $Ax \geq b$ rendszer megoldása, ahol

$$\lambda_1 := \max_{i \notin I} \left\{ \frac{b_i - {}_iax}{{}_iaz} : {}_iaz > 0 \right\},\ \lambda_2 := \min_{i \notin I} \left\{ \frac{b_i - {}_iax}{{}_iaz} : {}_iaz < 0 \right\}.$$

Itt $-\infty < \lambda_1 < 0 < \lambda_2 \leq \infty$, továbbá

$$\{i : {}_iax(\lambda_1) = b_i\} =: I_1 \supset I.$$

Bizonyítás: A $\mathrm{r}({}_IA) = \mathrm{r}(A)$ feltétel azzal ekvivalens, hogy $\mathrm{Im}\,({}_IA)^T = \mathrm{Im}\,A^T$, vagyis (ortogonális kiegészítő altereket véve) $\mathrm{Ker}\,{}_IA = \mathrm{Ker}\,A$. Ezért ha $\mathrm{r}({}_IA) < \mathrm{r}(A)$, akkor $\mathrm{Im}\,({}_IA)^T \subset \mathrm{Im}\,A^T$, vagyis $\mathrm{Ker}\,{}_IA \supset \mathrm{Ker}\,A$, tehát létezik z vektor, amelyre ${}_IAz = 0$, $Az \neq 0$.

Teljesül, hogy $(Ax(\lambda))_i = {}_iax + \lambda({}_iaz) \geq b_i$, ha
1. ${}_iaz = 0$ (spec. minden $i \in I$ esetén); vagy
2. ${}_iaz < 0$, és $\lambda \leq \frac{b_i - {}_iax}{{}_iaz}$; vagy
3. ${}_iaz > 0$, és $\lambda \geq \frac{b_i - {}_iax}{{}_iaz}$;

tehát $Ax(\lambda) \geq b$, ha $\lambda_1 \leq \lambda \leq \lambda_2$, $\lambda \in \mathcal{R}$.

Az állítás hátralevő része nyilvánvaló abból, hogy
1. $i \notin I$ esetén $b_i - {}_iax < 0$ (ezért $\lambda_1 < 0 < \lambda_2$);
2. Az-nek van pozitív koordinátája (ezért $-\infty < \lambda_1$);
3. ${}_iaz = 0$ $(i \in I)$, ezért

$${}_iax(\lambda_1) = {}_iax + \lambda_1({}_iaz) = {}_iax = b_i\ (i \in I).$$

Ha $\lambda_1 = \frac{b_{i^*} - {}_{i^*}ax}{{}_{i^*}az}$, akkor

$$ {}_{i^*}ax(\lambda_1) = {}_{i^*}ax + \frac{b_{i^*} - {}_{i^*}ax}{{}_{i^*}az} {}_{i^*}az = b_{i^*} $$

ráadásul, ezért $x(\lambda_1)$ legalább eggyel több egyenlőtlenséget teljesít egyenlőséggel. $\square$

Az $Ax \geq b$ rendszer adott x megoldásához az alábbi módon kereshetünk meg egy olyan z vektort, amelyre ${}_IAz = 0, Az \neq 0$. Hagyjuk el az $E \in \mathcal{R}^{m \times m}$ egységmátrix I-beli indexű oszlopait, az így kapott mátrixot jelöljük E'-vel. Keressük az $E'y - Az = 0$ rendszer megoldását, amelyre $y \neq 0$. (Ha ilyen van, akkor a $q^{(j)}$ vektorok egyike ilyen.) Az így kapott z vektor megfelel.

3.10. Tétel: (Caratheodory) *Ha az $Ax \geq b$ rendszernek van megoldása, akkor bázismegoldása is van.*

Bizonyítás: Legyen x az $Ax \geq b$ rendszer megoldása. Ha az x vektor nem bázismegoldás, akkor tekintsük az előző állításbeli $x(\lambda_1)$-et. Ez a vektor is megoldás, és $I_1 \supset I$ miatt "nagyobb az esélye" bázismegoldásnak lenni. Ha mégsem lenne az, legyen $x := x(\lambda_1)$, és kezdjük a bizonyítást elölről... $\square$

A Caratheodory-tétel két következményét említjük.

3.11. Állítás: *Legyen $S \subseteq \mathcal{R}^d$, ekkor*

$$ \operatorname{cone} S = \left\{ \sum_{i=1}^{d} \lambda_i x_i : \lambda_i \geq 0,\ x_i \in S\ (i = 1, \ldots, d) \right\}. $$

Bizonyítás: Az, hogy az állításban szereplő jobb oldali halmaz $\operatorname{cone} S$ része, azonnal következik 3.2-ből. A fordított irányú tartalmazás pedig Caratheodory tételének $Ax = b, x \geq 0$ alakú rendszerekre vonatkozó változatának következménye. $\square$

Legyenek $O_1, O_2 \subseteq \mathcal{R}^3$ az $(1,1,0)$, illetve $(-1,1,0)$ pontok körüli egységsugarú zárt gömbök $\{x \in \mathcal{R}^3 : x_1 = 1\}$, illetve $\{x \in \mathcal{R}^3 : x_1 = -1\}$ síkba eső részei. Ekkor az $S := O_1 \cup O_2$ halmaz mutatja, hogy a következő állítás feltétele nem hagyható el.

3.12. Állítás: *Ha $S \subseteq \mathcal{R}^d$ kompakt halmaz, és elemeiből csak triviálisan lehet kikombinálni az origót nemnegatív együtthatókkal, akkor* $\operatorname{cone} S$ *zárt.*

Bizonyítás: Legyen $x^{(k)}$ $(k = 1, 2, \ldots)$ $\operatorname{cone} S$-beli elemek egy konvergens sorozata, limeszpontját jelölje x. Azt kell megmutatnunk, hogy $x \in \operatorname{cone} S$. 3.11 szerint minden $k = 1, 2, \ldots$ esetén léteznek $x_i^{(k)} \in S$ $(i = 1, \ldots, d)$ elemek és $\lambda_i^{(k)} \geq 0$ $(i = 1, \ldots, d)$ számok úgy, hogy

$$x^{(k)} = \sum_{i=1}^{d} \lambda_i^{(k)} x_i^{(k)} \quad (k = 1, 2, \ldots).$$

A $\lambda_i^{(k)}$ $(i = 1, \ldots, d)$ számok összegét jelölje $\lambda^{(k)}$ $(k = 1, 2, \ldots)$. Ezekről a $\lambda^{(k)}$ számokról feltehető, hogy közülük végtelen sok nem nulla, különben végtelen sok $x^{(k)} = 0$, és így $x = 0 \in \operatorname{cone} S$ lenne. Részsorozatot véve ha szükséges, feltehető, hogy az összes $\lambda^{(k)} \neq 0$. Bolzano–Weierstrass tétele szerint, a $\{\lambda_i^{(k)}/\lambda^{(k)} : k = 1, 2 \ldots\}$ és az $\{x_i^{(k)} : k = 1, 2, \ldots\}$ halmazok korlátossága miatt feltehető, hogy

$$\lambda_i^{(k)}/\lambda^{(k)} \to \lambda_i' \ (k \to \infty), \text{ és } x_i^{(k)} \to x_i \ (k \to \infty) \qquad (i = 1, \ldots, d)$$

valamely $\lambda_i' \geq 0$ számok és $x_i \in S$ elemek esetén. Mivel

$$x^{(k)} = \lambda^{(k)} \cdot \sum_{i=1}^{d} \underbrace{(\lambda_i^{(k)}/\lambda^{(k)})}_{\to \lambda_i'} \underbrace{x_i^{(k)}}_{\to x_i} \to x \quad (k \to \infty),$$

azért a $\{\lambda^{(k)} : k = 1, 2, \ldots\}$ halmaz sem lehet korlátlan. (Különben részsorozatot választva ha kell, feltehető lenne, hogy $\lambda^{(k)} \to \infty$ $(k \to \infty)$, és akkor $\sum_{i=1}^{d} \lambda_i' x_i = 0$ lenne, ellentétben a feltétellel.) Ezért feltehető, hogy $\lambda^{(k)} \to \lambda$ $(k \to \infty)$ valamely $\lambda \geq 0$ esetén. Legyen $\lambda_i := \lambda \cdot \lambda_i'$ $(i = 1, \ldots, d)$, ekkor $x = \sum_{i=1}^{d} \lambda_i x_i$ a $\operatorname{cone} S$ eleme. $\square$

Speciálisan a lineárisan független vektorok által generált kúpok zártak. Caratheodory tételének $Ax = b, x \geq 0$ alakú rendszerekre vonatkozó változata szerint tetszőleges végesen generált kúp előáll véges sok lineárisan független vektorok által generált kúp úniójaként, tehát a fentiekből adódóan a végesen generált kúpok mind zártak. (Lehetnek tehát poliédrikus kúpok.) E kétféle kúptulajdonság azonosságához még egy segédtételre lesz szükségünk, ami a Farkas-lemma és Caratheodory tételének (valamint 3.8-nak) azonnali következménye.

Az $\{x : a^T x = \beta\}$, $\{x : a^T x \geq \beta\}$, illetve $\{x : a^T x > \beta\}$ alakú halmazokat, ahol $0 \neq a \in \mathcal{R}^d$, $\beta \in \mathcal{R}$, (ha $\beta = 0$, akkor **homogén**) **hipersík**nak, **zárt**, illetve **nyílt féltér**nek nevezzük. 2.24 szerint ez a definíció nem mond ellent a második fejezetbelinek, vagyis a hipersíkok éppen az eltolt $(d-1)$-dimenziós alterek.

A Farkas-lemma alternatíváinak jelentése ezután:
a) az $Ax = b, x \geq 0$ rendszer megoldhatósága azt jelenti, hogy $b \in \mathrm{Im}_+ A$;
b) az $A^T y \geq 0, b^T y < 0$ rendszer megoldhatósága pedig azt, hogy létezik olyan homogén hipersík, hogy az általa meghatározott egyik zárt féltér tartalmazza az $\mathrm{Im}_+ A$ kúpot, de a b vektort nem.

A szemlélet azt sugallja, hogy az utóbbi esetben olyan fenti tulajdonságú hipersík is létezik, amely tartalmazza
a) az egész $\mathrm{Im}_+ A$ kúpot, ha $b \notin \mathrm{lin}\,\mathrm{Im}_+ A$;
b) a kúp egy "lapját", ha $b \in \mathrm{lin}\,\mathrm{Im}_+ A$.

Ezt igazolja a következő tétel. Vegyük még észre, hogy mivel az $\mathrm{Im}_+ A$ kúp lineáris burka az $\mathrm{Im}\, A$ altér, azért
a) $b \notin \mathrm{lin}\,\mathrm{Im}_+ A$ esetén $\mathrm{r}\,(A, b) = \mathrm{r}\,(A) + 1 = \dim \mathrm{lin}\,\mathrm{Im}_+ A + 1$;
b) $b \in \mathrm{lin}\,\mathrm{Im}_+ A$ esetén pedig $\mathrm{r}\,(A, b) = \mathrm{r}\,(A) = \dim \mathrm{lin}\,\mathrm{Im}_+ A$.

3.13. Tétel: (A lineáris egyenlőtlenségek alaptétele) *Bármely $A \in \mathcal{R}^{m \times n}$ mátrix és $b \in \mathcal{R}^m$ vektor esetén az alábbi állítások közül pontosan az egyik teljesül:*
a) az $Ax = b, x \geq 0$ rendszer megoldható, vagyis $b \in \mathrm{Im}_+ A$;
b) létezik olyan y vektor, amelyre $A^T y \geq 0, b^T y = -1$, és amely merőleges $\mathrm{r}\,(A, b) - 1$ lineárisan független A-beli oszlopvektorra. □

Most már rátérhetünk annak bizonyítására, hogy a végesen generált kúpok éppen a poliéder kúpok.

3.14. Tétel: (Weyl) *Minden végesen generált kúp poliéder kúp.*

Bizonyítás: Legyen $A \in \mathcal{R}^{d \times n}$ tetszőleges mátrix, az $\mathrm{Im}_+ A$ kúpot kell előállítanunk valamely $\mathrm{Ker}_+ B$ kúpként, ahol $B \in \mathcal{R}^{m \times d}$.

Először abban az esetben igazoljuk az állítást, mikor $\mathrm{r}\,(A) = d$. Ekkor az $Ax = b$ rendszer minden $b \in \mathcal{R}^d$ vektor esetén megoldható, vagyis ekkor $\mathrm{r}\,(A, b) - 1 = \mathrm{r}\,(A) - 1 = d - 1$. Tekintsük az összes olyan homogén zárt félteret, amely tartalmazza az $\mathrm{Im}_+ A$ kúpot, és amelynek határoló hipersíkját $d - 1$ lineárisan független A-beli oszlopvektor feszíti. A lineáris egyenlőtlenségek alaptétele szerint $\mathrm{Im}_+ A$ e véges sok féltér metszete.

Általában könnyen belátható, hogy ha $N \in \mathcal{R}^{d\times d}$ invertálható mátrix, akkor

$$\operatorname{Im}_+(NA) = \operatorname{Ker}_+ \hat{B} \iff \operatorname{Im}_+ A = \operatorname{Ker}_+(\hat{B}N).$$

Ezért elég az $\operatorname{Im}_+(NA)$ kúpot előállítani valamely $\operatorname{Ker}_+ \hat{B}$ kúpként.

Tudjuk, hogy létezik $N \in \mathcal{R}^{d\times d}$ invertálható mátrix és $\pi : \{1,\dots,n\} \to \{1,\dots,n\}$ permutáció úgy, hogy az NAP_π mátrix első r sorát elhagyva 0 mátrixot kapunk, első r sorának első r oszlopában pedig egységmátrix áll. Jelölje A' az NAP_π mátrix első r sorából álló mátrixot. A már belátottak szerint létezik B' mátrix úgy, hogy $\operatorname{Im}_+ A' = \operatorname{Ker}_+ B'$. De akkor

$$\begin{aligned}\operatorname{Im}_+(NA) &= \operatorname{Im}_+(NAP_\pi) = \\ &= (\operatorname{Im}_+ A') \times \{0\} = (\operatorname{Ker}_+ B') \times \{0\} = \operatorname{Ker}_+ \hat{B},\end{aligned}$$

ahol

$$\hat{B} := \begin{pmatrix} B' & 0 \\ 0 & E \\ 0 & -E \end{pmatrix}.$$

□

A Weyl-tétel duálisa, a Minkowski-tétel most már könnyen adódik 3.4 segítségével.

3.15. Tétel: (Minkowski) *Minden poliéder kúp végesen generált kúp.*

Bizonyítás: Legyen $A \in \mathcal{R}^{m\times d}$ tetszőleges mátrix, a $\operatorname{Ker}_+ A$ kúpot kell előállítanunk valamely $\operatorname{Im}_+ B$ kúpként, ahol $B \in \mathcal{R}^{d\times n}$. Weyl tétele szerint létezik B mátrix úgy, hogy $\operatorname{Im}_+ A^T = \operatorname{Ker}_+ B^T$. Ekkor

$$\operatorname{Ker}_+ A = (\operatorname{Im}_+ A^T)^* = (\operatorname{Ker}_+ B^T)^* = \operatorname{Im}_+ B,$$

vagyis a B mátrix megfelel. □

Következő célunk Weyl és Minkowski tételei inhomogén változatának bizonyítása.

A $C \subseteq \mathcal{R}^d$ halmazt **konvex halmaz**nak nevezzük, ha bármely két elemével együtt azok nemnegatív, 1-összegű lineáris kombinációit (vagyis a két elem által meghatározott **szakasz**t) is tartalmazza, vagyis

$$x_1, x_2 \in C,\ \lambda_1, \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1 \text{ esetén } \lambda_1 x_1 + \lambda_2 x_2 \in C.$$

(Az x_1, x_2 elemek által meghatározott szakaszt röviden $[x_1, x_2]$-vel jelöljük.) Könnyen belátható, hogy nevüknek megfelelően a konvex kúpok éppen azok a konvex halmazok, amelyek kúpok is egyben.

Konvex halmazok tetszőleges rendszerének metszete nyilván konvex halmaz, így egy adott $S \subseteq \mathcal{R}^d$ halmaz esetén létezik egyértelműen az S halmazt tartalmazó legszűkebb konvex halmaz, S **konvex burka**,

$$\operatorname{conv} S := \cap\{C : S \subseteq C,\ C \text{ konvex halmaz}\}.$$

2.1 mintájára igazolható az alábbi

3.16. Állítás: *Tetszőleges nemüres $S \subseteq \mathcal{R}^d$ halmaz esetén teljesül, hogy*

$$\operatorname{conv} S = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathcal{N},\ \lambda_i \in \mathcal{R}_+,\ x_i \in S\ (i = 1, \ldots, k),\ \sum_{i=1}^{k} \lambda_i = 1 \right\},$$

az S-beli elemek **konvex kombinációi***nak a halmaza. Az üres halmaz konvex burka önmaga.* □

Véges sok [affin független] pont konvex burkaként előálló halmazt **politóp**nak [**szimplex**nek] nevezzük.

A véges sok zárt féltér metszeteként előálló halmazokat (vagyis a lineáris egyenlőtlenség-rendszerek megoldáshalmazait) **poliéder**eknek nevezzük.

Könnyen belátható, hogy nevüknek megfelelően a poliéder kúpok éppen azok a poliéderek, amelyek kúpok is egyben.

Végesen generált halmaznak nevezzük egy politóp és egy végesen generált kúp összegeként előálló halmazokat.

A most következő Motzkin-tétel szerint a poliéderek éppen a végesen generált halmazok. A tételt homogenizációval igazoljuk a Weyl-Minkowski-tételekből.

Vezessük be az alábbi jelöléseket:

$$\begin{aligned} C(K) &:= \left\{ x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq K \right\} && (K \subseteq \mathcal{R}^{d+1} \text{ konvex kúp}),\\ K(C) &:= \operatorname{cone}\left(\{1\} \times C\right) && (C \subseteq \mathcal{R}^d \text{ konvex halmaz}). \end{aligned}$$

A jelölésnek megfelelően $C(K)$ konvex halmaz, $K(C)$ konvex kúp.

A homogenizációs azonosságokat (vö. 2.19, 2.20, 2.21) bonyolítja, hogy a konvex kúpok és halmazok már nem feltétlenül zártak. (Bár konvex kúp [konvex halmaz] lezártja is konvex kúp [konvex halmaz].)

3.17. Állítás: *Tetszőleges* $S, S_1, S_2 \subseteq \mathcal{R}^d$ *halmazok esetén*
a) $\operatorname{cone} S = \operatorname{cone} \operatorname{conv} S = \{\lambda x : \lambda \geq 0, x \in \operatorname{conv} S\}$, *speciálisan*

$$\operatorname{cone}(\{1\} \times S) = \operatorname{cone}(\{1\} \times \operatorname{conv} S) = \\ = \cup\{\{\lambda\} \times \{\lambda x\} : \lambda \geq 0, x \in \operatorname{conv} S\};$$

b)

$$\left\{x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq \operatorname{cone}(\{1\} \times S)\right\} = \operatorname{conv} S;$$

c)

$$\operatorname{cone}(\{1\} \times S_1) \subseteq \operatorname{cone}(\{1\} \times S_2) \Longleftrightarrow \operatorname{conv} S_1 \subseteq \operatorname{conv} S_2.$$

(Az állítás c) része persze érvényben marad, ha benne mindkét $\subseteq$ *jel helyébe* $=$*-t vagy* $\subset$*-t írunk.)*

Bizonyítás: a) A $\operatorname{cone} S \subseteq \operatorname{cone} \operatorname{conv} S$ tartalmazás $S \subseteq \operatorname{conv} S$, míg a $\operatorname{cone} S \supseteq \operatorname{cone} \operatorname{conv} S$ tartalmazás $\operatorname{conv} S \subseteq \operatorname{cone} S$ következménye. A második igazolandó halmazegyenlőség hasonlóan bizonyítható.

b) könnyű következménye annak, hogy $\operatorname{cone} S$ [$\operatorname{conv} S$] S elemeinek nemnegatív [konvex] kombinációiból áll.

c) a) és b) következménye. □

3.18. Állítás: *Tetszőleges* $S, S_1, S_2 \subseteq \mathcal{R}^d$ *halmazok esetén*
a) $\operatorname{cl} \operatorname{cone} S = \operatorname{cl} \operatorname{cone} \operatorname{cl} S$, *speciálisan*

$$\operatorname{cl} \operatorname{cone}(\{1\} \times S) = \operatorname{cl} \operatorname{cone}(\{1\} \times \operatorname{cl} S);$$

b)

$$\left\{x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq \operatorname{cl} \operatorname{cone}(\{1\} \times S)\right\} = \operatorname{cl} \operatorname{conv} S;$$

c)

$$\operatorname{cl} \operatorname{cone}(\{1\} \times S_1) \subseteq \operatorname{cl} \operatorname{cone}(\{1\} \times S_2) \Longleftrightarrow \operatorname{cl} \operatorname{conv} S_1 \subseteq \operatorname{cl} \operatorname{conv} S_2.$$

(Az állítás c) része persze érvényben marad, ha benne mindkét $\subseteq$ *jel helyébe* $=$*-t vagy* $\subset$*-t írunk.)*

Bizonyítás: a) A $\operatorname{cl} \operatorname{cone} S \subseteq \operatorname{cl} \operatorname{cone} \operatorname{cl} S$ tartalmazás $S \subseteq \operatorname{cl} S$, a $\operatorname{cl} \operatorname{cone} S \supseteq \operatorname{cl} \operatorname{cone} \operatorname{cl} S$ tartalmazás pedig $\operatorname{cl} S \subseteq \operatorname{cl} \operatorname{cone} S$ következménye.

b) 3.17 a) és 3.18 a) miatt feltehető, hogy $S = C$ zárt, konvex halmaz. Azt kell belátnunk, hogy $C(\mathrm{cl}\, K(C)) = C$. 3.17 b) szerint $C = C(K(C))$, így világos, hogy $C \subseteq C(\mathrm{cl}\, K(C))$. Megfordítva ha $x \in C(\mathrm{cl}\, K(C))$, akkor 3.17 a) szerint léteznek $\lambda_k \geq 0$ számok és $x_k \in C$ pontok ($k = 1, 2, \ldots$) úgy, hogy $\lambda_k \to 1$, és $\lambda_k x_k \to x$ ($k \to \infty$). Emiatt az x_k ($k = 1, 2, \ldots$) pontsorozat konvergens, limeszpontja az x pont, amely a C halmaz zártsága miatt C-beli. Ebből a $C(\mathrm{cl}\, K(C)) \subseteq C$ tartalmazás is látszik.

c) 3.17 a) és 3.18 a) miatt feltehető, hogy $S_i = C_i$ ($i = 1, 2$) zárt, konvex halmazok. Az állítás ezek után 3.18 b) és 3.17 c) következménye. □

3.19. Állítás: *Legyen* $H := \{x \in \mathcal{R}^d : a^T x = \beta\}$ *egy inhomogén hipersík (vagyis* $0 \neq a \in \mathcal{R}^d$, $0 \neq \beta \in \mathcal{R}$ *teljesüljön), továbbá* $K \subseteq \mathcal{R}^d$ *konvex kúp. Ekkor*

a) $a^T(K \setminus \{0\}) > 0$, $\beta > 0$ *esetén* $\mathrm{cone}\,(K \cap H) = K$.

Ha K *még zárt is, és* $K \cap H \neq \emptyset$, *akkor*

b) $a^T K \geq 0$, $\beta > 0$ *esetén* $\mathrm{cl}\,\mathrm{cone}\,(K \cap H) = K$.

(Itt $S \geq 0$, *illetve* $S > 0$ *alatt természetesen azt értjük, hogy az* $S \subseteq \mathcal{R}$ *halmaz minden eleme nemnegatív, illetve pozitív.)*

Bizonyítás: a) Persze $\mathrm{cone}\,(K \cap H) \subseteq \mathrm{cone}\, K = K$. Másfelől $0 \neq x \in K$ esetén $a^T x > 0$, így $\frac{\beta}{a^T x} x \in K \cap H$, és

$$x = \frac{a^T x}{\beta} \left(\frac{\beta}{a^T x} x \right) \in \mathrm{cone}\,(K \cap H).$$

b) Itt is nyilvánvaló, hogy $\mathrm{cl}\,\mathrm{cone}\,(K \cap H) \subseteq K$. Másfelől ha $x \in K$, akkor $a^T x \geq 0$. Ha még $a^T x > 0$ is teljesül, akkor mint az előbb $x \in \mathrm{cone}\,(K \cap H)$. Tegyük fel ezért, hogy $a^T x = 0$, és legyen $0 \neq x_0 \in K \cap H$. Ekkor $x_0 + kx \in K \cap H$ ($k = 1, 2, \ldots$), és

$$\mathrm{cone}\,(K \cap H) \ni [x_0, x] \cap [x_0 + kx, 0] \to x \qquad (k \to \infty),$$

amiből $x \in \mathrm{cl}\,\mathrm{cone}\,(K \cap H)$, és így a másik irányú tartalmazás is adódik. □

3.20. Tétel: *Tetszőleges* $C \subseteq \mathcal{R}^d$ *konvex halmaz esetén* $C = C(K(C))$, *és* $\mathrm{cl}\, C = C(\mathrm{cl}\, K(C))$.

Ha $K \subseteq \mathcal{R}^{d+1}$ *konvex kúp, amelyre* $e_1^T(K \setminus \{0\}) > 0$ *teljesül, akkor* $K = K(C(K))$.

Ha $K \subseteq \mathcal{R}^{d+1}$ *zárt, konvex kúp, amelyre* $C(K) \neq \emptyset$ *és* $e_1^T K \geq 0$ *teljesül, akkor* $K = \mathrm{cl}\, K(C(K))$.

Bizonyítás: A tétel 3.17 b), 3.18 b) és 3.19 következménye (3.19-et alkalmazzuk az $a := e_1 \in \mathcal{R}^{d+1}$ és $\beta := 1$ választással). $\square$

A tétel első feléből adódóan a [zárt] konvex halmazok éppen a $C(K)$ alakú halmazok, ahol K az $\mathcal{R}^{d+1}$-beli [zárt] konvex kúpokon fut. A homogenizáció tehát ebben az általánosságban is működik, míg például a konvex halmazok már nem éppen az eltolt konvex kúpok.

3.11-ből, illetve 3.12-ből is homogenizációval adódnak a megfelelő inhomogén állítások.

3.21. Állítás: *Legyen $S \subseteq \mathcal{R}^d$, ekkor*

$$\operatorname{conv} S = \left\{ \sum_{i=1}^{d+1} \lambda_i x_i : \lambda_i \geq 0,\ x_i \in S\ (i = 1, \ldots, d+1),\ \sum_{i=1}^{d+1} \lambda_i = 1 \right\}.$$

$\square$

3.22. Állítás: *Kompakt halmaz konvex burka is kompakt.* $\square$

Zárt halmaz konvex burka lehet nemzárt: legyen a halmaz egy egyenes és egy rá nem illeszkedő pont. Viszont könnyen belátható, hogy nyílt halmaz konvex burka is nyílt lesz.

Weyl és Minkowski tételeinek inhomogén változatai is homogenizációval adódnak. Később jól látjuk majd (vö. 4.20, 4.13), hogy a bizonyításban szereplő $\tilde{P}$ poliéder nem más, mint $\operatorname{cl} K(P)$. (Az extra koordináta helyének nincs jelentősége, általában a vektor elejére vagy végére kerül.)

3.23. Tétel: (Motzkin) *A poliéderek éppen a végesen generált halmazok.*

Bizonyítás: 1. Legyen $P \subseteq \mathcal{R}^d$ poliéder, és éppen $P = \{x : Ax \geq b\}$ alakú, ahol $A \in \mathcal{R}^{m \times d}$, $b \in \mathcal{R}^m$. Ekkor

$$\tilde{P} := \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} : Ax - b\xi \geq 0,\ \xi \geq 0 \right\} \subseteq \mathcal{R}^{d+1}$$

poliéder kúp, így Minkowski tétele szerint léteznek $x_1, \ldots, x_n \in \mathcal{R}^d$ vektorok és $\xi_1, \ldots, \xi_n \in \mathcal{R}$ számok úgy, hogy

$$\tilde{P} = \operatorname{Im}_+ \begin{pmatrix} x_1 \ldots x_n \\ \xi_1 \ldots \xi_n \end{pmatrix}.$$

Feltehető, hogy minden ξ_i szám 0 vagy 1, például $\xi_1 = \ldots = \xi_j = 1$, $\xi_{j+1} = \ldots = \xi_n = 0$. Legyenek

$$Q := \operatorname{conv}\{x_1, \ldots, x_j\},\ R := \operatorname{Im}_+(x_{j+1}, \ldots, x_n),$$

ekkor Q politóp, R végesen generált kúp, és $P = Q + R$ végesen generált halmaz, ugyanis

$$x \in P \Longleftrightarrow \begin{pmatrix} x \\ 1 \end{pmatrix} \in \tilde{P} \Longleftrightarrow$$
$$\Longleftrightarrow \begin{pmatrix} x \\ 1 \end{pmatrix} \in \operatorname{Im}_+ \begin{pmatrix} x_1 \ldots x_j & x_{j+1} \ldots x_n \\ 1 \ldots 1 & 0 \ldots 0 \end{pmatrix} \Longleftrightarrow x \in Q + R.$$

2. Legyen $Q \subseteq \mathcal{R}^d$ politóp, $R \subseteq \mathcal{R}^d$ végesen generált kúp, és éppen $Q = \operatorname{conv}\{x_1, \ldots, x_j\}$, illetve $R = \operatorname{Im}_+(x_{j+1}, \ldots, x_n)$ alakú valamely x_i $(i = 1, \ldots, n)$ vektorok esetén. Ekkor

$$\tilde{P} := \operatorname{Im}_+ \begin{pmatrix} x_1 \ldots x_j & x_{j+1} \ldots x_n \\ 1 \ldots 1 & 0 \ldots 0 \end{pmatrix}$$

végesen generált kúp, így Weyl tétele szerint létezik $A \in \mathcal{R}^{m \times d}$ mátrix és $b \in \mathcal{R}^m$ vektor úgy, hogy

$$\tilde{P} = \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} : Ax + (-b)\xi \geq 0 \right\}.$$

Ekkor $Q + R = \{x : Ax \geq b\}$ poliéder, ugyanis

$$x \in Q + R \Longleftrightarrow \begin{pmatrix} x \\ 1 \end{pmatrix} \in \tilde{P} \Longleftrightarrow$$
$$\Longleftrightarrow \begin{pmatrix} x \\ 1 \end{pmatrix} \in \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} : Ax + (-b)\xi \geq 0 \right\} \Longleftrightarrow Ax \geq b.$$

□

3.24. Következmény: *A politópok éppen a korlátos poliéderek.* □

3.25. Következmény: *Poliéderek [politópok, végesen generált kúpok] véges metszete, direkt szorzata, lineáris képe (így vetülete altérre, lineáris kombinációja) is poliéder [politóp, végesen generált kúp]. Poliéder [poliéder kúp] lineáris inverz képe is poliéder [poliéder kúp].* □

Lineáris képek zártságának a szeparációs és alternatívatételek témakörben találjuk meg fontos következményeit.

Legyenek $S_1, S_2 \subseteq \mathcal{R}^d$ nemüres halmazok. Azt mondjuk, hogy az S_1 és S_2 halmazok **elválasztható**k vagy **szeparálható**k, ha létezik hipersík úgy, hogy az általa meghatározott két zárt féltér egyike tartalmazza az S_1 halmazt, a másika pedig az S_2 halmazt. (Vagyis létezik $0 \neq a \in \mathcal{R}^d$ vektor úgy, hogy $\sup a^T S_1 \leq \inf a^T S_2$.) Ekkor egy fenti tulajdonságú hipersíkot az S_1 és S_2 halmazokat **elválasztó** vagy **szeparáló hipersík**nak nevezzük. (Vagyis a fenti jelöléssel szeparáló hipersík lesz minden olyan $\{x \in \mathcal{R}^d : a^T x = \beta\}$ alakú halmaz, ahol $\sup a^T S_1 \leq \beta \leq \inf a^T S_2$.)

Azt mondjuk, hogy az S_1 és S_2 halmazok **erősen szeparálható**k, ha létezik két őket szeparáló hipersík, amelyek egymás eltoltjai. (Vagyis létezik $0 \neq a \in \mathcal{R}^d$ vektor úgy, hogy $\sup a^T S_1 < \inf a^T S_2$.)

Most egy alapvető fontosságú szeparációs tételt bizonyítunk.

3.26. Tétel: (első Hahn–Banach-tétel) *Legyenek C_1, $C_2 \subseteq \mathcal{R}^d$ nemüres, diszjunkt, konvex halmazok. Tegyük fel továbbá, hogy C_1 kompakt, C_2 pedig zárt halmaz. Ekkor a C_1 és C_2 halmazok erősen szeparálhatók, sőt létezik $a \in C_2 - C_1$ vektor, amelyre* $\sup a^T C_1 < \inf a^T C_2$.

Bizonyítás: Először megmutatjuk, hogy a C_1 és a C_2 halmaz távolsága felvétetik. Válasszunk egy C_1-beli és egy C_2-beli elemet, ezek távolsága legyen τ. A C_1 halmaz τ sugarú zárt környezete (jelölje ezt $\hat{C}_1$) már belemetsz C_2-be. Mivel C_1 és $\hat{C}_1 \cap C_2$ is korlátos és zárt halmazok, azért Weierstrass tétele szerint a

$$t : C_1 \times (\hat{C}_1 \cap C_2) \to \mathcal{R}_+;\ t(x_1, x_2) := ||x_1 - x_2||\ (x_1 \in C_1, x_2 \in \hat{C}_1 \cap C_2)$$

folytonos függvény infimuma felvétetik, és ez könnyen láthatóan éppen a C_1 és a C_2 távolsága.

Tehát elég megmutatnunk, hogy ha két diszjunkt, konvex halmaz távolsága felvétetik, akkor erősen szeparálhatók.

Feltehetjük, hogy $C_1 = \{0\}$. (Ha nem így lenne, szeparáljuk erősen a $C_2 - C_1$ és a $\{0\}$ halmazokat!) Legyen $x_0 \neq 0$ a C_2-beli legrövidebb vektor, ekkor $x \in C_2$, $0 < \lambda < 1$ esetén $x_0 + \lambda(x - x_0) \in C_2$, így

$$\begin{array}{rcl} ||x_0 + \lambda(x - x_0)||^2 & \geq & ||x_0||^2, \\ ||x_0||^2 + 2\lambda x_0^T(x - x_0) + \lambda^2 ||x - x_0||^2 & \geq & ||x_0||^2, \\ 2x_0^T(x - x_0) + \lambda ||x - x_0||^2 & \geq & 0, \end{array}$$

amiből (tartsunk λ-val 0-hoz!) $x_0^T x \geq ||x_0||^2$ következik. Ezért $a := x_0$ megfelel. $\square$

Két zárt, konvex halmaz már nem erősen szeparálható általában: legyen az egyik halmaz a síkban egy hiperbolalap, a másik halmaz pedig az aszimptotája.

Viszont két poliéder mindig erősen szeparálható (ez az eredmény csak a Farkas-lemma segítségével is igazolható, abból kiindulva, hogy a poliéderek diszjunktsága miatt a poliédereket leíró lineáris egyenlőtlenség-rendszerek együttesen nem megoldhatók):

3.27. Tétel: *Legyen $P_1, P_2 \subseteq \mathcal{R}^d$ két diszjunkt poliéder, ekkor P_1 és P_2 erősen szeparálhatók.*

Bizonyítás: Mivel $P_1 \cap P_2 = \emptyset$, azért $0 \notin P_2 - P_1$. Itt $P_2 - P_1$ 3.25 szerint poliéder, így 3.26 szerint erősen szeparálható a csak az origót tartalmazó kompakt, konvex halmaztól. Létezik tehát $a \in \mathcal{R}^d$ (sőt $a \in P_2 - P_1$) vektor úgy, hogy $0 < \inf a^T(P_2 - P_1)$. De akkor $\sup a^T P_1 < \inf a^T P_2$ is teljesül, vagyis P_1 és P_2 erősen szeparálhatóak. $\square$

Ha egy $K \subseteq \mathcal{R}^d$ kúp és egy $S \subseteq \mathcal{R}^d$ nemüres halmaz szeparálhatóak, akkor szeparálhatóak homogén hipersíkkal is. (Ha ugyanis $\sup a^T K \leq \inf a^T S$, akkor $a^T K \subseteq \mathcal{R}$ felülről korlátos kúp $\mathcal{R}$-ben, vagyis $\{0\}$ vagy $\mathcal{R}_-$. Mindenképpen $\sup a^T K = 0$, és így $\{x \in \mathcal{R}^d : a^T x = 0\}$ is szeparáló hipersík.) Ennek az egyszerű észrevételnek a segítségével most belátunk egy fontos dualitástételt, 2.12, sőt 3.5 általánosítását.

3.28. Tétel: *Tetszőleges $S \subseteq \mathcal{R}^d$ nemüres halmaz esetén $S^{**} = \operatorname{cl} \operatorname{cone} S$. Speciálisan $S^{**} = S$ pontosan akkor teljesül, ha az S halmaz zárt, konvex kúp.*

Bizonyítás: 3.3 szerint feltehető, hogy az S halmaz zárt, konvex kúp. Az $S \subseteq S^{**}$ tartalmazás nyilvánvaló. Azt kell még belátnunk, hogy ha $x \notin S$, akkor $x \notin S^{**}$. Legyen tehát $x \notin S$ tetszőleges pont, ekkor 3.26 és a tétel előtti megjegyzés szerint létezik $0 \neq a \in \mathcal{R}^d$ vektor úgy, hogy $a^T x < 0 = \inf a^T S$. Ekkor $a \in S^*$, $x^T a < 0$, tehát $x \notin S^{**}$, amit igazolnunk kellett. $\square$

3.29. Tétel: *Legyenek $K_1, K_2 \subseteq \mathcal{R}^d$ konvex kúpok. Ekkor $(K_1 \cap K_2)^* \supseteq \operatorname{cl}(K_1^* + K_2^*)$, egyenlőséggel, ha K_1, K_2 még zártak is.*

A megfelelő állítások igazak akkor is, ha kettő helyett több kúp szerepel bennük.

Bizonyítás: Először is $K_1 \cap K_2 \subseteq K_1$, így $K_1^* \subseteq (K_1 \cap K_2)^*$, és hasonlóan $K_2^* \subseteq (K_1 \cap K_2)^*$. De akkor (mivel $(K_1 \cap K_2)^*$ zárt, konvex kúp)

$$K_1^* + K_2^* \subseteq (K_1 \cap K_2)^* + (K_1 \cap K_2)^* = (K_1 \cap K_2)^*,$$

és így

$$\mathrm{cl}\,(K_1^* + K_2^*) \subseteq (K_1 \cap K_2)^*$$

is teljesül, ami éppen a tétel első fele.

A tétel második fele 3.28 és 3.6 segítségével igazolható: ha ugyanis K_1, K_2 zárt, konvex kúpok, akkor

$$(K_1 \cap K_2)^* = (K_1^{**} \cap K_2^{**})^* = (K_1^* + K_2^*)^{**} = \mathrm{cl}\,(K_1^* + K_2^*).$$

A kettőnél több kúp esete hasonlóan intézhető el. (Figyeljük meg, milyen hasznos volt a $K^{**} = K$ dualitástétel!) □

A fejezet végén visszatérünk a Farkas-lemmához. Számtalan variánsát egységesítjük.

Legyen $K \subseteq \mathcal{R}^d$ konvex kúp, és $x_1, x_2 \in \mathcal{R}^d$ esetén $x_1 \geq_K x_2$ jelölje azt, hogy $x_1 - x_2 \in K$. Könnyen belátható, hogy a $\geq_K$ reláció reflexív, tranzitív, és ha $K \cap -K = \{0\}$ (amikor is a kúpot **csúcsos**nak nevezzük), akkor antiszimmetrikus is, tehát részbenrendezés.

Cseréljük le a Farkas-lemmában az "$x \geq 0$" feltételt az "$x \geq_K 0$" feltételre. Az így kapott kúplineáris Farkas-lemmához már fel kell tennünk, hogy az AK konvex kúp zárt (mivel ez nem teljesül feltétlenül, ha K már nem a nemnegatív ortáns), sőt a tétel alternatíváinak ekvivalenciája éppen ezt a zártságot jelenti:

3.30. Tétel: (kúplineáris Farkas-lemma) *Legyen $A \in \mathcal{R}^{m \times n}$, $K \subseteq \mathcal{R}^n$ tetszőleges konvex kúp. Tegyük fel, hogy az AK konvex kúp zárt. Ekkor (és csak ekkor) az alábbi állítások minden $b \in \mathcal{R}^m$ esetén ekvivalensek:*
a) létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $Ax = b, x \geq_K 0$;
b) valahányszor egy $y \in \mathcal{R}^m$ vektor esetén $A^T y \geq_{K^*} 0$ teljesül, akkor $b^T y \geq 0$ is egyben.

Bizonyítás: Az a) alternatíva nyilván azzal ekvivalens, hogy $b \in AK$. A b) alternatíva jelentése pedig az, hogy $b \in (AK)^{**}$, ugyanis

$$A^T y \in K^* \Longleftrightarrow \inf (A^T y)^T K \geq 0 \Longleftrightarrow \inf y^T (AK) \geq 0 \Longleftrightarrow y \in (AK)^*,$$

és a b) alternatíva éppen azt követeli meg, hogy minden ilyen y vektor esetén $b^T y \geq 0$ legyen.

A két alternatíva tehát pontosan akkor ekvivalens minden $b \in \mathcal{R}^m$ esetén, ha $AK = (AK)^{**}$, vagyis ha AK zárt, konvex kúp. $\square$

3.31. Megjegyzés: A zártsági feltétel nem triviális, nem következik K zártságából. Legyen ugyanis

$$K := \left\{ x \in \mathcal{R}^3 : x \text{ és } (1,0,1)^T \text{ szöge } \leq 45° \right\}.$$

Abból, hogy $a, x \in \mathcal{R}^d$ esetén az a és az x vektorok által bezárt szög koszinusza $a^T x/(||a|| \cdot ||x||)$ könnyen látható, hogy

$$K = \left\{ x \in \mathcal{R}^3 : 2x_1 x_3 \geq x_2^2,\ x_1, x_3 \geq 0 \right\}.$$

Legyen A az $\{x \in \mathcal{R}^3 : x_1 = 0\}$ altérre való vetítésnek megfelelő vetítő mátrix megfosztva első (nulla) sorától, azaz

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Megmutatjuk, hogy AK nem zárt. Nézzük hová vetülnek a kúp tengelyére merőleges síkokkal való metszetei! Legyen

$$H_\alpha := \left\{ x \in \mathcal{R}^3 : x_1 + x_3 = \alpha \right\} \ (\alpha \geq 0).$$

Ekkor $K = \cup_{\alpha \geq 0}(K \cap H_\alpha)$, amiből

$$AK = A(\cup_{\alpha \geq 0}(K \cap H_\alpha)) = \cup_{\alpha \geq 0} A(K \cap H_\alpha).$$

Nyilván

$$\begin{aligned} A(K \cap H_\alpha) &= \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : 2(\alpha - x_3)x_3 \geq x_2^2, 0 \leq x_3 \leq \alpha \right\} = \\ &= \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : \alpha^2/2 \geq x_2^2 + 2(x_3 - \alpha/2)^2 \right\}. \end{aligned}$$

Ezért

$$AK = \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : x_3 > 0 \right\} \cup \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},\ \mathrm{cl}\,(AK) = \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : x_3 \geq 0 \right\},$$

így például $b := (1,0)^T \in \mathrm{cl}\,(AK) \setminus AK$.

Kimondjuk még a kúplineáris Farkas-lemma egy (látszólag) általánosabb formáját, amelyet szokás szerint eltérésváltozókat bevezetve bizonyíthatunk 3.30-ból (vagy hasonló módon).

3.32. Tétel: (általános kúplineáris Farkas-lemma) *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá $K_1 \subseteq \mathcal{R}^m$, $K_2 \subseteq \mathcal{R}^n$ konvex kúpok. Tegyük fel, hogy az $AK_2 - K_1$ konvex kúp zárt. Ekkor (és csak ekkor) az alábbi állítások minden $b \in \mathcal{R}^m$ esetén ekvivalensek:*
a) létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $Ax \geq_{K_1} b$, $x \geq_{K_2} 0$;
b) valahányszor egy $y \in \mathcal{R}^m$ vektor esetén $A^T y \geq_{K_2^*} 0$, $y \leq_{K_1^*} 0$, akkor $b^T y \geq 0$ is egyben. □

Fontos következményként nyerjük az alábbi tételt:

3.33. Tétel: (kúplineáris Farkas-tétel) *Legyen $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$, $\delta \in \mathcal{R}$, továbbá $K_1 \subseteq \mathcal{R}^m$, $K_2 \subseteq \mathcal{R}^n$ konvex kúpok. Tegyük fel, hogy az*

$$\begin{pmatrix} A \\ -c^T \end{pmatrix} K_2 - \begin{pmatrix} K_1 \\ \mathcal{R}_+ \end{pmatrix}$$

konvex kúp zárt, továbbá hogy a

$$\mathbf{P} := \{x \in \mathcal{R}^n : Ax \geq_{K_1} b,\, x \geq_{K_2} 0\},$$
$$\mathbf{D} := \left\{y \in \mathcal{R}^m : A^T y \leq_{K_2^*} c,\, y \geq_{K_1^*} 0\right\}$$

jelölésekkel a $\mathbf{P} \cup \mathbf{D}$ halmaz nemüres. Ekkor az alábbi állítások ekvivalensek:
a) létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $Ax \geq_{K_1} b$, $x \geq_{K_2} 0$, és $c^T x \leq \delta$;
b) valahányszor egy $y \in \mathcal{R}^m$ vektor esetén $A^T y \leq_{K_2^*} c$, $y \geq_{K_1^*} 0$, akkor $b^T y \leq \delta$ is egyben.

Bizonyítás: Az a)⇒b) irány most is könnyű. Ha $x \in \mathbf{P}$, $y \in \mathbf{D}$, és $c^T x \leq \delta$, akkor

$$0 \leq y^T(Ax - b) + x^T(c - A^T y) = c^T x - b^T y \leq \delta - b^T y,$$

és így $b^T y \leq \delta$ teljesül.

A fordított irányhoz elég megmutatnunk, hogy b)-ből következik, hogy

$$A^T y - c\mu \leq_{K_2^*} 0,\ y \geq_{K_1^*} 0,\ \mu \geq_{\mathcal{R}_+} 0 \text{ esetén } b^T y - \delta\mu \leq 0;$$

ennek ugyanis 3.32 és a zártsági feltétel szerint már következménye az a) állítás. A kiemelt állítás $\mu > 0$ esetén a μ-vel való leosztás után b)-ből adódik. Ha $\mu = 0$, és $\mathbf{P} \neq \emptyset$, akkor 3.32 triviális "a)⇒b)" iránya szerint $b^T y \leq 0$. Ha $\mu = 0$, és $\mathbf{D} \neq \emptyset$, akkor rögzítsünk tetszőleges $y_0 \in \mathbf{D}$ elemet, és legyen $y \geq_{K_1^*} 0$ olyan vektor, amelyre $A^T y \leq_{K_2^*} 0$ teljesül. Legyen továbbá $y_\lambda := y_0 + \lambda y$ $(\lambda \geq 0)$, ekkor $y_\lambda \in \mathbf{D}$ $(\lambda \geq 0)$, b) szerint $b^T y_\lambda \leq \delta$ minden $\lambda \geq 0$ esetén, ami csak úgy lehetséges, ha $b^T y \leq 0$. Ezzel az igazolandó állítást a $\mu = 0$, $\mathbf{D} \neq \emptyset$ esetben is beláttuk. □

3.33 megfordításáról szólva könnyen belátható, hogy
– a) pontosan akkor nem teljesül, ha

$$\begin{pmatrix} b \\ -\delta \end{pmatrix} \notin \begin{pmatrix} A \\ -c^T \end{pmatrix} K_2 - \begin{pmatrix} K_1 \\ \mathcal{R}_+ \end{pmatrix}$$

– $\mathbf{P} \cup \mathbf{D} \neq \emptyset$ esetén b) pontosan akkor nem teljesül, ha

$$\begin{pmatrix} b \\ -\delta \end{pmatrix} \notin \mathrm{cl}\left(\begin{pmatrix} A \\ -c^T \end{pmatrix} K_2 - \begin{pmatrix} K_1 \\ \mathcal{R}_+ \end{pmatrix} \right).$$

Vegyük észre, hogy 3.25 szerint 3.32 és 3.33 zártsági feltétele is teljesülni fog, ha K_1 és K_2 is poliéder kúp. (A lineáris kép képzés, az összeadás és a direkt szorzás nem vezet ki a poliéder kúpok köréből.) Általános konvex kúpok esetében viszont nemtriviális feltételeket jelentenek. Nehezen ellenőrizhetőek, így más elégséges feltételeket keresünk (lásd a 6.26 és 6.27 zártsági tételeket).

A relatív belsőtől a regularizációig

4. A relatív belső, lezárt tulajdonságai

Könnyen belátható, hogy egyetlen pozitív sugarú gömb sem lehet egyetlen hipersík része sem, vagyis minden hipersík üres belsejű. A belsőképzés tartalmazástartása miatt ha egy halmaz része valamely hipersíknak, akkor e halmaz belseje is üres. Van azonban a belső fogalmának egy természetes módosítása, amely nemtriviális lehet akkor is, mikor a halmaz belefér egy hipersíkba, vagyis affin burka szűkebb, mint az egész tér, különben pedig visszaadja a belsőt.

Legyen $S \subseteq \mathcal{R}^d$ tetszőleges halmaz. Azt mondjuk, hogy az $x \in S$ pont az S halmaz **relatív belső pont**ja, ha létezik $\varepsilon > 0$ úgy, hogy

$$O(x,\varepsilon) \cap \operatorname{aff} S \subseteq S.$$

Az S halmaz relatív belső pontjainak halmazát az S halmaz **relatív belsejé**nek nevezzük, és $\operatorname{ri} S$-sel jelöljük. Ha az S halmaz megegyezik relatív belsejével, akkor azt mondjuk, hogy S **relatív nyílt**. Az S halmaz **relatív határ**a az $\operatorname{rb} S := (\operatorname{cl} S) \setminus (\operatorname{ri} S)$ halmaz, melynek pontjai az S halmaz **relatív határpont**jai.

Néhány megjegyzés a relatív belsővel kapcsolatban:

– Ha $\operatorname{aff} S = \mathcal{R}^d$, akkor $\operatorname{ri} S = \operatorname{int} S$.

– A relatív belső képzés (ellentétben a belsőképzéssel) már nem tartalmazástartó: abból, hogy $S_1 \subseteq S_2$, még nem feltétlenül következik, hogy $\operatorname{ri} S_1 \subseteq \operatorname{ri} S_2$. (Legyen S_2 egy zárt intervallum, S_1 pedig az egyik végpontja.) Egy elégséges feltétel, hogy $\operatorname{aff} S_1 = \operatorname{aff} S_2$ legyen.

– Könnyen belátható, hogy $x \in \operatorname{ri} S$, $\varepsilon > 0$ esetén

$$O(x,\varepsilon) \cap \operatorname{aff} S \subseteq S \Rightarrow O(x,\varepsilon) \cap \operatorname{aff} S \subseteq \operatorname{ri} S.$$

Azt mondjuk, hogy az $x \in \mathcal{R}^d$ pont **behúzható** egy $x_0 \in S$ pont mellé az $S \subseteq \mathcal{R}^d$ halmazba, ha létezik $\varepsilon > 0$ úgy, hogy $x_0 + \varepsilon(x - x_0) \in S$.

Azt mondjuk, hogy az $x \in \mathcal{R}^d$ pont **túlhúzható** az $x_0 \in S$ ponton az $S \subseteq \mathcal{R}^d$ halmazban, ha létezik $\varepsilon > 0$ úgy, hogy $x_0 - \varepsilon(x - x_0) \in S$.

A fenti legutolsó megjegyzés szerint tetszőleges $x \in \operatorname{aff} S$ pont behúzható egy $x_0 \in \operatorname{ri} S$ pont mellé az S (vagy akár a $\operatorname{ri} S$) halmazba, és hasonlóan túlhúzással.

Következő észrevételünk, hogy egy átkoordinátázás és a relatív belső képzés felcserélhető. (Az világos, hogy a belső, a lezárt, az affin burok képzés, a konvex burok képzés felcserélhető egy átkoordinátázással.)

4.1. Állítás: *Ha $S \subseteq \mathcal{R}^d$ tetszőleges halmaz, továbbá $T : \mathcal{R}^d \to \mathcal{R}^d$ átkoordinátázás, akkor $T(\operatorname{ri} S) = \operatorname{ri} T(S)$. (Speciálisan $\operatorname{ri} S \neq \emptyset$ pontosan akkor, ha $\operatorname{ri} T(S) \neq \emptyset$.)*

Bizonyítás: Először a $T(\operatorname{ri} S) \subseteq \operatorname{ri} T(S)$ tartalmazást igazoljuk. Feltehető, hogy $\operatorname{ri} S \neq \emptyset$. Legyen $x \in \operatorname{ri} S$. Ekkor létezik $\varepsilon > 0$ úgy, hogy

$$O(x, \varepsilon) \cap \operatorname{aff} S \subseteq S.$$

Ebből

$$T(O(x, \varepsilon)) \cap T(\operatorname{aff} S) = T(O(x, \varepsilon) \cap \operatorname{aff} S) \subseteq T(S)$$

következik. Nyílt halmaz képe átkoordinátázásnál nyílt (hiszen T az inverzének inverze), ezért létezik $\varepsilon' > 0$ úgy, hogy

$$O(T(x), \varepsilon') \subseteq T(O(x, \varepsilon)).$$

Továbbá $T(\operatorname{aff} S) = \operatorname{aff} T(S)$, így a fentiekből

$$O(T(x), \varepsilon') \cap \operatorname{aff} T(S) \subseteq T(S)$$

adódik, vagyis $T(x) \in \operatorname{ri} T(S)$.

Végül a már bizonyítottak szerint

$$T^{-1}(\operatorname{ri} T(S)) \subseteq \operatorname{ri} T^{-1}(T(S)) = \operatorname{ri} S,$$

amiből (alkalmazzuk T-t) adódik a $\operatorname{ri} T(S) \subseteq T(\operatorname{ri} S)$ tartalmazás is. $\square$

Most már egy nemtriviális példát is adhatunk a relatív belsőre:

4.2. Állítás: *Ha* $S = \{x_0, \dots, x_k\} \subseteq \mathcal{R}^d$ *affin független halmaz, akkor a*

$$Q := \operatorname{conv} S = \left\{ \sum_{i=0}^{k} \lambda_i x_i : \lambda_i \geq 0 \ (i = 0, \dots, k), \ \sum_{i=0}^{k} \lambda_i = 1 \right\}$$

szimplex relatív belseje $\operatorname{ri} Q = Q'$, *ahol*

$$Q' := \left\{ \sum_{i=0}^{k} \lambda_i x_i : \lambda_i > 0 \ (i = 0, \dots, k), \ \sum_{i=0}^{k} \lambda_i = 1 \right\}.$$

Bizonyítás: Vegyük észre, hogy elég az $S = S_0$ esettel foglalkozni, ahol $S_0 := \{0, e_1, \dots, e_k\}$. (A Q és Q' halmazoknak ekkor a Q_0 és Q_0' halmazok felelnek meg természetesen.) Valóban, 2.25 szerint létezik $T : \mathcal{R}^d \to \mathcal{R}^d$ átkoordinátázás úgy, hogy

$$T(0) = x_0, \, T(e_1) = x_1, \, \dots, \, T(e_k) = x_k,$$

és akkor a speciális eset ismeretében, az előző állítást is használva, már következne az általános eset:

$$\begin{aligned} \operatorname{ri} Q &= \operatorname{ri} \operatorname{conv} S = \operatorname{ri} \operatorname{conv} T(S_0) = \\ &= T(\operatorname{ri} \operatorname{conv} S_0) = T(\operatorname{ri} Q_0) = T(Q_0') = Q' \end{aligned}$$

lenne.

Elég tehát azt belátni, hogy $\operatorname{ri} Q_0 = Q_0'$, vagyis

$$\begin{aligned} \operatorname{ri} &\left\{ (\lambda_1, \dots, \lambda_k, 0, \dots, 0)^T \in \mathcal{R}^d : \lambda_i \geq 0, \textstyle\sum_{i=1}^{k} \lambda_i \leq 1 \right\} = \\ &= \left\{ (\lambda_1, \dots, \lambda_k, 0, \dots, 0)^T \in \mathcal{R}^d : \lambda_i > 0, \textstyle\sum_{i=1}^{k} \lambda_i < 1 \right\}. \end{aligned}$$

Először is könnyen belátható, hogy $\operatorname{aff} Q_0 = \mathcal{R}^k \times \{0\}$. (A nemtriviális tartalmazáshoz vegyük észre, hogy egy Q_0'-beli ponthoz bármely $\mathcal{R}^k \times \{0\}$-beli pont behúzható a Q_0 halmazba.)

Legyen most $x \in Q_0'$, ekkor a szereplő függvények folytonossága miatt létezik $\varepsilon > 0$ úgy, hogy $y \in O(x, \varepsilon)$ esetén

$$e_i^T y \geq 0 \ (i = 1, \dots, k), \text{ és } 1 - \sum_{i=1}^{k} e_i^T y \geq 0.$$

Ekkor nyilván

$$O(x, \varepsilon) \cap \operatorname{aff} Q_0 \subseteq Q_0,$$

vagyis $x \in \operatorname{ri} Q_0$. Ez mutatja a $Q_0' \subseteq \operatorname{ri} Q_0$ tartalmazást.

Megfordítva minden $Q_0 \setminus Q_0'$-beli x ponthoz tetszőleges közel találunk $(\operatorname{aff} Q_0) \setminus Q_0$-beli pontot. (Húzzunk túl egy Q_0'-beli pontot az x ponton!) Ebből a $\operatorname{ri} Q_0 \subseteq Q_0'$ tartalmazás is következik. $\square$

A következő lemma alapvető fontosságú a későbbiekhez (egyszerű következménye például, hogy egy konvex halmaz relatív belseje is konvex). Itt természetes jelöléssel legyen

$$[x, y[:= [x, y] \setminus \{y\} \ (x, y \in \mathcal{R}^d).$$

Hasonlóan definiálható $]x, y]$ és $]x, y[$ is.

4.3. Lemma: (Elérhetőségi lemma) *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz, $x_0 \in \operatorname{ri} C$, $x_1 \in \operatorname{cl} C$. Ekkor $[x_0, x_1[\subseteq \operatorname{ri} C$.*

Bizonyítás: A bizonyítás két lépésből áll.

1. Először azt igazoljuk, hogy ha $x_0 \in \operatorname{ri} C$, és $x_1 \in C$, akkor $[x_0, x_1[\subseteq \operatorname{ri} C$. Világos, hogy ha $\varepsilon > 0$ olyan szám, amellyel

$$O(x_0, \varepsilon) \cap \operatorname{aff} C \subseteq C$$

teljesül, akkor minden $0 < \lambda \leq 1$ esetén

$$O(x_1 + \lambda(x_0 - x_1), \lambda\varepsilon) \cap \operatorname{aff} C \subseteq C.$$

Ez mutatja, hogy $[x_0, x_1[\subseteq \operatorname{ri} C$.

2. Másodszor azt igazoljuk, hogy ha $x_0 \in \operatorname{ri} C$, és $x_1 \in \operatorname{cl} C$, akkor $[x_0, x_1[\subseteq \operatorname{ri} C$. Tegyük fel, hogy ε a fenti tulajdonságú pozitív szám, és legyen $0 < \lambda < 1$. Válasszunk olyan $x_1' \in C$ pontot, amelyre

$$||x_1' - x_1|| < \varepsilon \Big/ \left(\frac{1}{\lambda} - 1\right)$$

teljesül. Válasszunk továbbá egy x_0' pontot is úgy, hogy

$$x_1' + \lambda(x_0' - x_1') = x_1 + \lambda(x_0 - x_1)$$

legyen. Könnyű belátni, hogy ekkor $||x_0' - x_0|| < \varepsilon$, így $x_0' \in \operatorname{ri} C$. A már igazolt eset szerint $x_1 + \lambda(x_0 - x_1) \in \operatorname{ri} C$. Ezzel beláttuk az $[x_0, x_1[\subseteq \operatorname{ri} C$ tartalmazást is. $\square$

Ezidáig még azt sem igazoltuk, hogy egy nemüres, konvex halmaznak egyáltalán van relatív belső pontja. A most következő tételnek ez azonnali következménye.

4.4. Tétel: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha C affin halmaz, akkor*

$$\operatorname{ri} C = \operatorname{cl} C = \operatorname{aff} C,$$

különben

$$\emptyset \subset \operatorname{ri} C \subset \operatorname{cl} C \subset \operatorname{aff} C.$$

Bizonyítás: A tétel első fele nyilvánvaló, hiszen ha C affin halmaz, akkor $C = \operatorname{aff} C$. A tétel második feléhez három szigorú tartalmazást kell belátnunk.

1. Először azt igazoljuk, hogy ha C nem affin halmaz, akkor $\emptyset \subset \operatorname{ri} C$. Szűkítsük a C affin generáló részhalmazt az $\operatorname{aff} C$ affin halmaz affin bázisává, ily módon egy $S \subseteq C$ affin független halmazt kapunk, amelyre $\operatorname{aff} S = \operatorname{aff} C$ teljesül. Legyen $Q := \operatorname{conv} S$, ekkor $Q \subseteq C$ és $\operatorname{aff} Q = \operatorname{aff} C$ miatt $\operatorname{ri} Q \subseteq \operatorname{ri} C$. Mivel 4.2 szerint $\operatorname{ri} Q \neq \emptyset$, azért $\operatorname{ri} C \neq \emptyset$ is fennáll.

2. Másodszor azt igazoljuk, hogy ha C nem affin halmaz, akkor $\operatorname{ri} C \subset \operatorname{cl} C$. Legyenek $x_0 \in \operatorname{ri} C$, $x_1 \in (\operatorname{aff} C) \setminus C$ tetszőleges pontok, továbbá

$$\hat{\lambda} := \sup\{\lambda : x_0 + \lambda(x_1 - x_0) \in C\}.$$

Ekkor $0 < \hat{\lambda} \leq 1$, és

$$x_0 + \hat{\lambda}(x_1 - x_0) \in (\operatorname{cl} C) \setminus (\operatorname{ri} C).$$

3. Végül azt látjuk be, hogy ha C nem affin halmaz, akkor $\operatorname{cl} C \subset \operatorname{aff} C$. Legyen x_0, x_1 mint fent, ekkor

$$x_0 + 2(x_1 - x_0) \in (\operatorname{aff} C) \setminus (\operatorname{cl} C).$$

(Itt 2 helyett tetszőleges $\hat{\lambda}$-nál nagyobb szám megfelelne.) □

4.5. Következmény: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha $C \neq \emptyset$, akkor $\operatorname{ri} C \neq \emptyset$. Ha $C \neq \mathcal{R}^d$, akkor $\operatorname{cl} C \neq \mathcal{R}^d$. A C konvex halmaz relatív határa pontosan akkor üres, ha C affin halmaz.* □

A relatív belső pontok a későbbiek szempontjából igen fontos jellemzését nyerhetjük.

4.6. Lemma: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, konvex halmaz. Az x_0 pont éppen akkor relatív belső pontja a C halmaznak, ha a C minden pontja túlhúzható az x_0 ponton a C halmazban.*

Bizonyítás: A nemtriviális irányhoz tegyük fel, hogy a C halmaz minden pontja túlhúzható az $x_0 \in C$ ponton a C halmazban. Mivel C nemüres, konvex halmaz, azért létezik relatív belső pontja, ez is túlhúzható az x_0 ponton a C halmazban. De akkor az elérhetőségi lemma miatt $x_0 \in \operatorname{ri} C$.□

4.7. Tétel: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ekkor az* aff , *a* ri *és* cl *operációk nem különböztetik meg a* $\operatorname{ri} C$ *és* $\operatorname{cl} C$ *közötti halmazokat, vagyis például* $\operatorname{ri} C \subseteq S_1, S_2 \subseteq \operatorname{cl} C$ *esetén* $\operatorname{aff} S_1 = \operatorname{aff} S_2$.

Bizonyítás: A tétel affin burokra vonatkozó részéhez elég azt belátnunk, hogy $\operatorname{aff} \operatorname{ri} C = \operatorname{aff} C$, és $\operatorname{aff} C = \operatorname{aff} \operatorname{cl} C$. Ebből világos, hogy $\operatorname{aff} \operatorname{ri} C \subseteq \operatorname{aff} C$, és $\operatorname{aff} C \subseteq \operatorname{aff} \operatorname{cl} C$. Az $\operatorname{aff} C \subseteq \operatorname{aff} \operatorname{ri} C$ tartalmazáshoz legyen x_0 a C halmaz relatív belső pontja, ekkor tetszőleges $x \in \operatorname{aff} C$ pont behúzható az x_0 pont mellé a $\operatorname{ri} C$ halmazba, ami azt mutatja, hogy $x \in \operatorname{aff} \operatorname{ri} C$ is teljesül. Továbbá $\operatorname{aff} C$ a C halmazt tartalmazó zárt halmaz, így $\operatorname{cl} C \subseteq \operatorname{aff} C$, amiből $\operatorname{aff} \operatorname{cl} C \subseteq \operatorname{aff} C$ is következik.

Ezek után a tétel relatív belsőre vonatkozó részéhez elég azt belátnunk, hogy $\operatorname{ri} \operatorname{ri} C = \operatorname{ri} C$, és $\operatorname{ri} C = \operatorname{ri} \operatorname{cl} C$. A relatív belsővel kapcsolatos, 4.1 előtti utolsó megjegyzésből világos, hogy a $\operatorname{ri} C$ halmaz relatív nyílt. Mivel $C \subseteq \operatorname{cl} C$, és $\operatorname{aff} C = \operatorname{aff} \operatorname{cl} C$, azért $\operatorname{ri} C \subseteq \operatorname{ri} \operatorname{cl} C$. Ha most $x_0 \in \operatorname{ri} \operatorname{cl} C$, akkor tetszőleges $\operatorname{cl} C$-beli pont túlhúzható az x_0 ponton a $\operatorname{cl} C$ halmazban. Húzzunk rajta túl egy $\operatorname{ri} C$-beli pontot, az elérhetőségi lemmából adódóan $x_0 \in \operatorname{ri} C$ lesz. Ezzel a $\operatorname{ri} \operatorname{cl} C \subseteq \operatorname{ri} C$ tartalmazást is beláttuk.

A tétel lezártakra vonatkozó részéhez is elég belátnunk, hogy $\operatorname{cl} \operatorname{ri} C = \operatorname{cl} C$, és $\operatorname{cl} C = \operatorname{cl} \operatorname{cl} C$. Világos, hogy $\operatorname{cl} C$ zárt halmaz, és $\operatorname{ri} C \subseteq C$ miatt $\operatorname{cl} \operatorname{ri} C \subseteq \operatorname{cl} C$ is nyilvánvaló. Megfordítva legyen $x_1 \in \operatorname{cl} C$, $x_0 \in \operatorname{ri} C$. Mivel ekkor az elérhetőségi lemma szerint $[x_0, x_1[\subseteq \operatorname{ri} C$, továbbá az x_1 pont e halmaz torlódási pontja, azért $x_1 \in \operatorname{cl} \operatorname{ri} C$ is teljesül. Ezzel beláttuk, hogy $\operatorname{cl} C \subseteq \operatorname{cl} \operatorname{ri} C$. □

Egyszerű következményként nyerjük az alábbi lemmát.

4.8. Lemma: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ konvex halmazok. Ekkor az alábbi állítások ekvivalensek:*
a) $\operatorname{ri} C_1 = \operatorname{ri} C_2$;
b) $\operatorname{cl} C_1 = \operatorname{cl} C_2$;
c) $\operatorname{ri} C_1 \subseteq C_2 \subseteq \operatorname{cl} C_1$.

Bizonyítás: Az előző tételből világos, hogy a) és b) ekvivalensek. Másfelől c)-ből, az ott szereplő halmazokat lezárva, következik b). Végül ha a) és b) fennáll, akkor

$$\operatorname{ri} C_1 = \operatorname{ri} C_2 \subseteq C_2 \subseteq \operatorname{cl} C_2 = \operatorname{cl} C_1.$$

□

A relatív belsőről és a lezártról bizonyítandó felcserélhetőségi tételek némelyikében fontos szerep jut a recessziós kúpnak.

Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. A továbbiakban jelölje $\operatorname{rec} C$ a C **recessziós kúp**ját, vagyis

$$\operatorname{rec} C := \left\{ z \in \mathcal{R}^d : x + \lambda z \in C \ (x \in C, \lambda \geq 0) \right\}.$$

Ez a kúp mindazon irányok halmaza, amelyekben C bármely pontjából elindulva, "végig" a C halmazban maradunk. Nyilván $\operatorname{rec} C \subseteq \operatorname{par} C$.

Jelölje továbbá $\operatorname{line} C$ a C **linealitás ter**ét, vagyis

$$\operatorname{line} C := (\operatorname{rec} C) \cap (-\operatorname{rec} C) = \left\{ z \in \mathcal{R}^d : x + \lambda z \in C \ (x \in C, \lambda \in \mathcal{R}) \right\}.$$

A $\operatorname{line} C$ altér jelentőségét az adja, hogy tetszőleges $L \subseteq \operatorname{line} C$ altér esetén $C = L + (C \cap L^{\perp})$.

Például ha $K \subseteq \mathcal{R}^d$ konvex kúp, akkor (és csak akkor)

$$\operatorname{rec} K = K, \operatorname{line} K = K \cap -K.$$

A következő állítások a recessziós kúpról szóló alapvető észrevételeket foglalják össze.

4.9. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz. Ekkor*
a) $\operatorname{rec} C$ *zárt, konvex kúp;*
b) $\operatorname{rec} C = \{z \in \mathcal{R}^d : x_0 + \lambda z \in C \ (\lambda \geq 0)\}$, *ahol $x_0 \in C$ tetszőleges rögzített pont;*
c) a C halmaz pontosan akkor korlátos, ha $\operatorname{rec} C = \{0\}$.

Bizonyítás: a) Nyilvánvaló, hogy $\operatorname{rec} C$ kúp. Legyen most $z, z' \in \operatorname{rec} C$, $0 \leq \varepsilon \leq 1$, $x \in C$, $\lambda \geq 0$, ekkor

$$x + \lambda(\varepsilon z + (1-\varepsilon)z') = (x + \lambda\varepsilon z) + \lambda(1-\varepsilon)z' \in C,$$

amiből a $\operatorname{rec} C$ kúp konvexitása is látszik. (Idáig a C konvex halmaz zártságát sem használtuk.) Végül legyen z_k $(k = 1, 2, \ldots)$ $\operatorname{rec} C$-beli pontok valamely z ponthoz tartó sorozata, továbbá $x \in C$, $\lambda \geq 0$. Ekkor

$z_k \in \operatorname{rec} C$ miatt $x + \lambda z_k \in C$ $(k = 1, 2, \ldots)$, a C halmaz zártsága miatt pedig e pontsorozat limesze, az $x + \lambda z$ pont is C-beli. Eszerint $z \in \operatorname{rec} C$, tehát a $\operatorname{rec} C$ kúp zárt.

b) Legyen z az állítás b) részében szereplő jobb oldali halmaz egy eleme, $x \in C$ tetszőleges x_0-tól különböző pont. Megmutatjuk, hogy $x + z\mathcal{R}_+ \subseteq C$. Ha $z \in (x - x_0)\mathcal{R}$, akkor ez nyilvánvaló, ezért feltehető, hogy $z \notin (x - x_0)\mathcal{R}$. Ekkor

$$C \ni [x_0, x + \lambda z] \cap [x, x_0 + kz] \to x + \lambda z \ (k \to \infty),$$

így a C halmaz zártsága miatt $x + \lambda z \in C$, amiből $z \in \operatorname{rec} C$ következik. A fordított irányú tartalmazás nyilvánvaló.

c) Nyilvánvaló, hogy ha a C halmaz korlátos, akkor $\operatorname{rec} C = \{0\}$. Tegyük fel most, hogy C nem korlátos, legyen $x_k \in C$ $(k = 1, 2, \ldots)$, amelyre $||x_k|| \to \infty$ $(k \to \infty)$. Legyen továbbá x_0 rögzített, C-beli pont, ekkor feltehető, hogy

$$z_k := \frac{x_k - x_0}{||x_k - x_0||} \to z \quad (k \to \infty),$$

ahol z egységvektor. Mivel $x_0 + \lambda z_k \in C$, ha $0 \leq \lambda \leq ||x_k - x_0||$, azért fix $\lambda \geq 0$ esetén, elég nagy k-kra $x_0 + \lambda z_k \in C$. Ez a sorozat pedig $x_0 + \lambda z$-hez tart, így $x_0 + \lambda z \in C$, vagyis $0 \neq z \in \operatorname{rec} C$. □

4.7 megfelelője a rec operációval:

4.10. Állítás: *Tetszőleges $C \subseteq \mathcal{R}^d$ konvex halmaz esetén*

$$\operatorname{rec} C \subseteq \operatorname{rec} \operatorname{cl} C = \operatorname{rec} \operatorname{ri} C.$$

Bizonyítás: Először lássuk be, hogy $\operatorname{rec} C \subseteq \operatorname{rec} \operatorname{cl} C$. Legyen $z \in \operatorname{rec} C$, továbbá rögzítsünk egy $x_0 \in C$ pontot. Ekkor $x_0 + z\mathcal{R}_+ \subseteq C$. Mivel $x_0 \in \operatorname{cl} C$, és $C \subseteq \operatorname{cl} C$, azért az is látszik, hogy $z \in \operatorname{rec} \operatorname{cl} C$ (vö. 4.9 b)).

Másodszor belátjuk, hogy $\operatorname{rec} \operatorname{cl} C = \operatorname{rec} \operatorname{ri} C$. A már igazoltak szerint

$$\operatorname{rec} \operatorname{ri} C \subseteq \operatorname{rec} \operatorname{cl} \operatorname{ri} C = \operatorname{rec} \operatorname{cl} C.$$

A fordított irányú tartalmazáshoz legyen $z \in \operatorname{rec} \operatorname{cl} C$, és $x_0 \in \operatorname{ri} C$ tetszőleges elem. Ekkor $x_0 + z\mathcal{R}_+ \subseteq \operatorname{cl} C$, de akkor az elérhetőségi lemma miatt $x_0 + z\mathcal{R}_+ \subseteq \operatorname{ri} C$ is teljesül, vagyis $z \in \operatorname{rec} \operatorname{ri} C$. □

Egy $C \subseteq \mathcal{R}^d$ konvex halmaz **korlát kúp**ja (angolul barrier cone) legyen mindazon irányok halmaza, amelyekből nézve a C halmaz alulról korlátos, vagyis

$$\operatorname{bar} C := \left\{ a \in \mathcal{R}^d : \inf(a^T C) > -\infty \right\}.$$

Például ha $C = K$ konvex kúp, akkor $\operatorname{bar} K = K^*$. Ha $C \subseteq \mathcal{R}^d$ kompakt, konvex halmaz, akkor $\operatorname{bar} C = \mathcal{R}^d$. Általában is

4.11. Állítás: *Tetszőleges $C \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmaz esetén*

$$\operatorname{rec} C = (\operatorname{bar} C)^*, \text{ és } (\operatorname{rec} C)^* = \operatorname{cl} \operatorname{bar} C.$$

Továbbá a $\operatorname{line} C$ és a $\operatorname{lin} \operatorname{bar} C$ alterek egymás ortogonális kiegészítői.

Bizonyítás: Csak az első egyenlőséget kell bizonyítanunk, a második abból adjungált kúpokat véve adódik.

A $\operatorname{rec} C \subseteq (\operatorname{bar} C)^*$ tartalmazáshoz legyen $z \in \operatorname{rec} C$, és rögzítsünk egy tetszőleges $x_0 \in C$ pontot. Ekkor $x_0 + z\mathcal{R}_+ \subseteq C$, így

$$\inf a^T(x_0 + z\mathcal{R}_+) > -\infty \quad (a \in \operatorname{bar} C),$$

amiből $a^T z \geq 0$ $(a \in \operatorname{bar} C)$ adódik, vagyis $z \in (\operatorname{bar} C)^*$. Ezzel beláttuk, hogy $\operatorname{rec} C \subseteq (\operatorname{bar} C)^*$.

A fordított irányú tartalmazást indirekt módon bizonyítjuk. Tegyük fel, hogy létezik

$$z \in (\operatorname{bar} C)^* \setminus (\operatorname{rec} C)$$

pont. Rögzítsünk egy tetszőleges $x_0 \in C$ pontot, ekkor elég nagy $\lambda \geq 0$ esetén $x_0 + \lambda z \notin C$. Az első Hahn–Banach-tétel szerint létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$a^T(x_0 + \lambda z) < \inf a^T C.$$

Látszik, hogy $a \in \operatorname{bar} C$, és hogy $a^T z < 0$ (mivel $x_0 \in C$). Ez azonban ellentmond annak, hogy $z \in (\operatorname{bar} C)^*$. □

Könnyen belátható, hogy ha $S' \subseteq \mathcal{R}^d$ kompakt, $S \subseteq \mathcal{R}^d$ pedig zárt halmaz, akkor $S' + S$ zárt halmaz lesz. Legyen ugyanis $x'_k \in S'$ és $x_k \in S$ $(k = 1, 2, \ldots)$ két sorozat, amelyre $x'_k + x_k \to y$ $(k \to \infty)$. Feltehető, hogy $x'_k \to x'$ $(k \to \infty)$ valamely $x' \in S'$ esetén. De akkor szükségképpen $x_k \to y - x'$ $(k \to \infty)$, és így $y - x' \in S$, vagyis $y \in S' + S$, amit bizonyítanunk kellett. Ebből az észrevételből és 4.11-ből adódik az alábbi

4.12. Lemma: *Ha $C' \subseteq \mathcal{R}^d$ kompakt, konvex halmaz, $C \subseteq \mathcal{R}^d$ pedig zárt, konvex halmaz, akkor $C' + C$ zárt, konvex halmaz, melynek korlát, illetve recessziós kúpja*

$$\mathrm{bar}\,(C' + C) = \mathrm{bar}\, C, \textit{ és } \mathrm{rec}\,(C' + C) = \mathrm{rec}\, C.$$

□

Következésképpen poliéderek és végesen generált halmazok recessziós és korlát kúpjáról elmondható, hogy

4.13. Állítás: *Ha $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, és az $Ax \geq b$ rendszer megoldható, akkor*

$$\mathrm{rec}\,\{x : Ax \geq b\} = \{z : Az \geq 0\}, \textit{ és } \mathrm{bar}\,\{x : Ax \geq b\} = \{A^T y : y \geq 0\}.$$

Ha $Q \subseteq \mathcal{R}^d$ nemüres politóp, és $R \subseteq \mathcal{R}^d$ poliéder kúp, akkor

$$\mathrm{bar}\,(Q + R) = R^*, \textit{ és } \mathrm{rec}\,(Q + R) = R.$$

□

4.14. Állítás: *Legyen $A \in \mathcal{R}^{m\times n}$ egy mátrix, $P \subseteq \mathcal{R}^n$ poliéder. Ekkor AP poliéder, amelyre $\mathrm{rec}\,(AP) = A\mathrm{rec}\, P$ teljesül. Speciálisan, ha $c \in \mathcal{R}^n$ egy vektor, akkor $c^T P \subseteq \mathcal{R}$ poliéder, vagyis zárt intervallum. Ezért $\inf c^T P \in \mathcal{R}$ esetén létezik $x \in P$ úgy, hogy $c^T x = \inf c^T P$; $\inf c^T P = -\infty$ esetén pedig létezik $z \in \mathrm{rec}\, P$ úgy, hogy $c^T z = -1$.* □

Most már rátérhetünk a relatív belsőre és a lezártra vonatkozó felcserélhetőségi tételekre.

4.15. Tétel: *Legyenek $C_i \subseteq \mathcal{R}^{d_i}$ $(i = 1, \ldots, k)$ nemüres, konvex halmazok. Ekkor*

$$\mathrm{aff}\ \times_{i=1}^k C_i = \times_{i=1}^k \mathrm{aff}\, C_i,$$

és hasonlóan aff *helyett a* ri, cl, *illetve* rec *operációkkal.* □

4.16. Tétel: *Legyenek $C_i \subseteq \mathcal{R}^d$ $(i = 1, \ldots, k)$ konvex halmazok. Tegyük fel, hogy $\cap_{i=1}^k \mathrm{ri}\, C_i \neq \emptyset$. Ekkor*

a)

$$\mathrm{aff}\ \cap_{i=1}^k C_i = \cap_{i=1}^k \mathrm{aff}\, C_i = \mathrm{aff}\ \cap_{i=1}^k \mathrm{ri}\, C_i;$$

b)

$$\mathrm{ri}\ \cap_{i=1}^k C_i = \cap_{i=1}^k \mathrm{ri}\, C_i;$$

c)

$$\mathrm{cl} \cap_{i=1}^{k} C_i = \cap_{i=1}^{k} \mathrm{cl}\, C_i.$$

A c) állításhoz még azt sem kell feltenni, hogy csak véges sok halmaz metszetéről van szó.

Bizonyítás: A tétel a) részéhez először is vegyük észre, hogy $\cap C_i \subseteq \cap \mathrm{aff}\, C_i$, továbbá az utóbbi halmaz affin, így $\mathrm{aff} \cap C_i \subseteq \cap \mathrm{aff}\, C_i$. Tetszőleges $\cap \mathrm{aff}\, C_i$-beli pont behúzható egy $\cap \mathrm{ri}\, C_i$-beli pont mellé a $\cap \mathrm{ri}\, C_i$ halmazba. Ez mutatja, hogy $\cap \mathrm{aff}\, C_i \subseteq \mathrm{aff} \cap \mathrm{ri}\, C_i$. Végül nyilván $\mathrm{aff} \cap \mathrm{ri}\, C_i \subseteq \mathrm{aff} \cap C_i$.

A tétel b) részéhez először azt igazoljuk, hogy $\cap \mathrm{ri}\, C_i$ relatív nyílt halmaz. Mivel nyilván konvex, azért 4.6 szerint ehhez elég annyi, hogy minden pontján túlhúzható benne minden pontja, és ez nyilvánvaló. Eszerint $\mathrm{ri}\,(\cap \mathrm{ri}\, C_i) = \cap \mathrm{ri}\, C_i$. Másfelől az elérhetőségi lemmából könnyen beláthatóan

$$\cap \mathrm{ri}\, C_i \subseteq \cap C_i \subseteq \mathrm{cl} \cap \mathrm{ri}\, C_i,$$

amiből $\mathrm{ri}\,(\cap \mathrm{ri}\, C_i) = \mathrm{ri} \cap C_i$ is adódik (vö. 4.8).

A tétel c) részéhez ismét az elérhetőségi lemmát használva, könnyen belátható, hogy

$$\cap \mathrm{cl}\, C_i \subseteq \mathrm{cl} \cap \mathrm{ri}\, C_i \subseteq \mathrm{cl} \cap C_i \subseteq \cap \mathrm{cl}\, C_i.$$

□

Ha C_1 és C_2 két síkbeli, ugyanazon egyenes által meghatározott nyílt félsík megtoldva egy az egyenesen lévő közös ponttal, akkor a fenti tételbeli egyetlen egyenlőség sem áll fenn. Ez mutatja, hogy a tétel feltételére szükség van.

A recessziós kúpra vonatkozó megfelelő állítás:

4.17. Állítás: *Legyen $C_i \subseteq \mathcal{R}^d$ $(i \in I)$ zárt, konvex halmazok tetszőleges rendszere. Tegyük fel, hogy $\cap_{i \in I} C_i \neq \emptyset$. Ekkor*

$$\mathrm{rec} \cap_{i \in I} C_i = \cap_{i \in I} \mathrm{rec}\, C_i.$$

□

4.18. Tétel: *Tetszőleges $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz esetén*

$$\mathrm{cone}\, C = \{\lambda x : \lambda \geq 0, x \in C\}, \text{ és } \mathrm{ri}\,\mathrm{cone}\, C = \{\lambda x : \lambda > 0, x \in \mathrm{ri}\, C\}.$$

Bizonyítás: Az első egyenlőség nyilvánvaló: a nemnegatív kombinációból kiemelhetjük a súlyok összegét.

A második egyenlőség az alábbi módon látható be: Jelöljük az egyenlőség jobb oldalán szereplő halmazt S-sel. Először is megmutatjuk, hogy az S halmaz konvex. A $\operatorname{ri} C$ halmaz konvexitása miatt ehhez elég, hogy $\lambda_1, \lambda_2 > 0$, $x_1, x_2 \in \operatorname{ri} C$, $0 \le \varepsilon \le 1$ esetén léteznek $\lambda > 0$, $0 \le \varepsilon' \le 1$ számok úgy, hogy

$$\varepsilon\lambda_1 x_1 + (1-\varepsilon)\lambda_2 x_2 = \lambda(\varepsilon' x_1 + (1-\varepsilon')x_2).$$

Könnyen belátható, hogy

$$\lambda := \varepsilon\lambda_1 + (1-\varepsilon)\lambda_2 \text{ és } \varepsilon' := \varepsilon\lambda_1/\lambda$$

megfelelnek.

Másodszor megmutatjuk, hogy S relatív nyílt halmaz. Legyen $\lambda x \in S$ (itt $\lambda > 0$, $x \in \operatorname{ri} C$), 4.6 szerint azt kell igazolnunk, hogy tetszőleges $\lambda_1 > 0$, $x_1 \in \operatorname{ri} C$ esetén a $\lambda_1 x_1$ pont túlhúzható a λx ponton S-ben, vagyis létezik $0 < \varepsilon < 1$, $\lambda_2 > 0$ és $x_2 \in \operatorname{ri} C$ úgy, hogy

$$\lambda x = \varepsilon\lambda_1 x_1 + (1-\varepsilon)\lambda_2 x_2.$$

Tudjuk, hogy $x_1 \in \operatorname{ri} C$ túlhúzható az $x \in \operatorname{ri} C$ ponton $\operatorname{ri} C$-ben, vagyis elég kis $0 < \varepsilon' < 1$ esetén $x = \varepsilon' x_1 + (1-\varepsilon')x_2$, ahol $x_2 \in \operatorname{ri} C$. Válasszuk ε'-t ilyen kicsinek úgy, hogy ráadásul $\varepsilon' < \lambda_1/\lambda$ is teljesüljön. Ekkor

$$\varepsilon := \frac{\lambda}{\lambda_1}\varepsilon' < 1, \text{ és } \lambda_2 := \lambda\frac{1-\varepsilon'}{1-\varepsilon} > 0.$$

Könnyen belátható, hogy az így választott x_2, ε, λ_2 megfelelnek.

Ezek után a $\operatorname{ri}\operatorname{cone} C = S$ egyenlőség már következik abból, hogy nyilvánvalóan $\operatorname{cl} S \supseteq \operatorname{cone} C \supseteq S$, és így $\operatorname{ri}\operatorname{cone} C = \operatorname{ri} S = S$ (vö. 4.8). □

Mivel $\operatorname{cl}\operatorname{cone} C = \operatorname{cl}\operatorname{cone}\operatorname{cl} C$, azért a következő tételben nem jelent megszorítást, ha feltesszük, hogy C nem csak konvex, hanem zárt is.

4.19. Tétel: *Ha $C \subseteq \mathcal{R}^n$ nemüres, zárt, konvex halmaz, akkor $\operatorname{cl}\operatorname{cone} C$ zárt, konvex kúp, amelyre*

$$\operatorname{cl}\operatorname{cone} C \supseteq (\operatorname{cone} C) \cup (\operatorname{rec} C)$$

teljesül, egyenlőséggel, ha $0 \notin C$, vagy ha C poliéder.

Ha $P \subseteq \mathcal{R}^n$ nemüres poliéder, akkor $\operatorname{cl}\operatorname{cone} P$ *poliéder kúp, amelyre*

$$\operatorname{cl}\operatorname{cone} P \supseteq \operatorname{cone} P$$

teljesül, egyenlőséggel, ha $0 \in P$, *vagy ha* P *politóp.*

Bizonyítás: Legyen először C nemüres, zárt, konvex halmaz. Megmutatjuk, hogy ekkor $\operatorname{rec} C \subseteq \operatorname{cl}\operatorname{cone} C$. Legyen $0 \neq z \in \operatorname{rec} C$, továbbá $x_0 \notin z\mathcal{R}$ tetszőleges C-beli pont (a $C \subseteq z\mathcal{R}$ esetben nyilván $z \in \operatorname{cone} C$), ekkor $x_0 + kz \in C$ $(k = 1, 2, \ldots)$,

$$\operatorname{cone} C \ni [x_0, z] \cap [x_0 + kz, 0] \to z \qquad (k \to \infty),$$

így $z \in \operatorname{cl}\operatorname{cone} C$, vagyis $\operatorname{rec} C \subseteq \operatorname{cl}\operatorname{cone} C$. Ebből már látszik, hogy $\operatorname{cl}\operatorname{cone} C \supseteq (\operatorname{cone} C) \cup (\operatorname{rec} C)$, sőt az is, hogy

$$\operatorname{cl}\operatorname{cone} C = \operatorname{cl}((\operatorname{cone} C) \cup (\operatorname{rec} C)).$$

Tegyük fel most, hogy $0 \notin C$, és legyen $y \in \operatorname{cl}\operatorname{cone} C$. Ekkor léteznek $\lambda_k \geq 0$ számok és $x_k \in C$ pontok úgy, hogy $\lim \lambda_k x_k = y$. Ha $\sup \lambda_k = \infty$ teljesülne, akkor a λ_k $(k = 1, 2, \ldots)$ számsorozat egy részsorozata ∞-hez tartana, így ugyanez a részsorozata az x_k $(k = 1, 2, \ldots)$ pontsorozatnak 0-hoz tartana, amely C zártsága miatt C-beli lenne, ellentétben a feltétellel. Tehát $\sup \lambda_k < \infty$, és feltehető, hogy $\lambda_k \to \lambda \geq 0$ $(k \to \infty)$. Ha $\lambda > 0$, akkor $x_k \to y/\lambda$ $(k \to \infty)$, így $y/\lambda \in C$, és $y \in \operatorname{cone} C$. Ha $\lambda = 0$, akkor $x_0 \in C$, $\mu \geq 0$ esetén $x_0 + \mu\lambda_k(x_k - x_0) \in C$, feltéve, hogy már $\mu\lambda_k \leq 1$ is teljesül. Ezért

$$x_0 + \mu y = \lim x_0 + \mu\lambda_k(x_k - x_0) \in C,$$

vagyis ekkor $y \in \operatorname{rec} C$. Ezzel beláttuk, hogy $0 \notin C$ esetén $\operatorname{cl}\operatorname{cone} C \subseteq (\operatorname{cone} C) \cup (\operatorname{rec} C)$ is teljesül.

Lássuk be most a tétel poliéderekre vonatkozó részét. Ha P nemüres poliéder, akkor Motzkin tétele szerint léteznek $x_1, \ldots, x_k$ és $z_1, \ldots, z_l$ pontok úgy, hogy

$$P = \operatorname{conv}\{x_1, \ldots, x_k\} + \operatorname{cone}\{z_1, \ldots, z_l\}.$$

Ekkor 4.13 szerint $\operatorname{rec} P = \operatorname{cone}\{z_1, \ldots, z_l\}$, és így

$$(\operatorname{cone} P) \cup (\operatorname{rec} P) = \operatorname{cone}\{x_1, \ldots, x_k, z_1, \ldots, z_l\}.$$

Az utóbbi halmaz Weyl tétele szerint poliéder kúp, így megegyezik lezártjával, $\mathrm{cl\,cone}\, P$-vel.

Ha $0 \in P$, akkor $\mathrm{rec}\, P = 0 + \mathrm{rec}\, P \subseteq P$, így $\mathrm{rec}\, P \subseteq \mathrm{cone}\, P$. Ha P politóp, akkor $\mathrm{rec}\, P = \{0\} \subseteq \mathrm{cone}\, P$. Mindkét esetben

$$\mathrm{cl\,cone}\, P = (\mathrm{cone}\, P) \cup (\mathrm{rec}\, P) = \mathrm{cone}\, P.$$

□

Ha C egy síkbeli, az origót határában tartalmazó, zárt körlap, akkor $\mathrm{cl\,cone}\, C \supset (\mathrm{cone}\, C) \cup (\mathrm{rec}\, C)$. Ha P egy síkbeli, az origót nem tartalmazó egyenes, akkor $\mathrm{cl\,cone}\, P \supset \mathrm{cone}\, P$.

Az előző két tétel speciális eseteként adódik a

4.20. Lemma: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, konvex halmaz. Ekkor*

$$\begin{aligned} K(C) &= \cup \{\{\lambda\} \times \{\lambda x\} : \lambda \geq 0, x \in C\}, \\ \mathrm{ri}\, K(C) &= \cup \{\{\lambda\} \times \{\lambda x\} : \lambda > 0, x \in \mathrm{ri}\, C\}. \end{aligned}$$

Ha C még zárt is, akkor

$$\mathrm{cl}\, K(C) = K(C) \cup (\{0\} \times \mathrm{rec}\, C),$$

amely kúp poliéder, ha $C = P$ poliéder. □

4.1 általánosításaként megfogalmazható a

4.21. Tétel: *Legyen $A \in \mathcal{R}^{m \times n}$ mátrix, $C \subseteq \mathcal{R}^n$ konvex halmaz. Ekkor $\mathrm{ri}\,(AC) = A\,\mathrm{ri}\, C$.*

Bizonyítás: Nyilvánvaló, hogy az $A\,\mathrm{ri}\, C$ halmaz konvex, megmutatjuk, hogy relatív nyílt is. Legyen $x \in \mathrm{ri}\, C$, azt kell igazolnunk, hogy tetszőleges $x_1 \in \mathrm{ri}\, C$ esetén az Ax_1 pont túlhúzható az Ax ponton $A\,\mathrm{ri}\, C$-ben, vagyis létezik $0 < \varepsilon < 1$, $x_2 \in \mathrm{ri}\, C$ úgy, hogy $Ax = \varepsilon A x_1 + (1 - \varepsilon) A x_2$ legyen. Tudjuk, hogy az x_1 pont túlhúzható az x ponton $\mathrm{ri}\, C$-ben, vagyis létezik $0 < \varepsilon < 1$ szám és $x_2 \in \mathrm{ri}\, C$ pont úgy, hogy $x = \varepsilon x_1 + (1 - \varepsilon) x_2$ legyen. Utóbbi egyenlőségre alkalmazva A-t látjuk, hogy az így választott ε és x_2 megfelelnek. Tehát $A\,\mathrm{ri}\, C = \mathrm{ri}\,(A\,\mathrm{ri}\, C)$.

Másrészt nyilván

$$\mathrm{cl}\,(A\,\mathrm{ri}\, C) \supseteq A\,\mathrm{cl\,ri}\, C = A\,\mathrm{cl}\, C \supseteq AC \supseteq A\,\mathrm{ri}\, C,$$

amiből $\mathrm{ri}\,(AC) = \mathrm{ri}\,(A\,\mathrm{ri}\, C)$ is következik. □

4.22. Következmény: *Legyenek $C_1, \ldots, C_k \subseteq \mathcal{R}^d$ konvex halmazok, $\lambda_1, \ldots, \lambda_k \in \mathcal{R}$ tetszőleges számok. Ekkor*

$$\mathrm{ri} \sum_{i=1}^{k} \lambda_i C_i = \sum_{i=1}^{k} \lambda_i \mathrm{ri}\, C_i.$$

Bizonyítás: Legyen

$$A := (\lambda_1 E, \ldots, \lambda_k E),\; C := C_1 \times \ldots \times C_k,$$

és alkalmazzuk az előző tételt. □

Bár $A\mathrm{cl}\, C \subseteq \mathrm{cl}\,(AC)$ (továbbá $A\mathrm{rec}\, C \subseteq \mathrm{rec}\,(AC)$ miatt $A\mathrm{rec}\,(\mathrm{cl}\, C) \subseteq \mathrm{rec}\,\mathrm{cl}\,(AC)$) mindig fennáll, láttunk példát arra, hogy itt lehet szigorú tartalmazás, még akkor is, ha $C = K$ zárt, konvex kúp. Bizonyos feltétellel azonban mindig egyenlőség van e két halmaz között. Mivel $\mathrm{cl}\,(AC) = \mathrm{cl}\,(A\mathrm{cl}\, C)$, azért nem jelent megszorítást, ha feltesszük, hogy C nem csak konvex halmaz, hanem zárt is.

A következő tétel (amely $P = \{0\}$ és $A = E$ speciális eseteinek "összege") ismét példája annak, hogy néha az általánosabb könnyebben bizonyítható.

4.23. Tétel: *Legyen $C \subseteq \mathcal{R}^n$ zárt, konvex halmaz, $P \subseteq \mathcal{R}^m$ poliéder, továbbá $A \in \mathcal{R}^{m\times n}$ tetszőleges mátrix. Ha teljesül, hogy*

$$Ax \in -\mathrm{rec}\, P,\; x \in \mathrm{rec}\, C \textit{ esetén } x \in -\mathrm{rec}\, C,$$

akkor

$$\mathrm{cl}\,(AC + P) = AC + P, \textit{ és } \mathrm{rec}\,(AC + P) = A\mathrm{rec}\, C + \mathrm{rec}\, P.$$

Bizonyítás: A bizonyítást hat lépésre bontottam, az első lépés a lényeges, a többi standard trükkök sorozata. Az első négy lépésben igazoljuk, hogy $\mathrm{cl}\,(AC + P) = AC + P$, az utolsó két lépésben, hogy $\mathrm{rec}\,(AC + P) = A\mathrm{rec}\, C + \mathrm{rec}\, P$, a tétel feltételeinek fennállása esetén. Vegyük még észre, hogy a 2. lépés alábbi bizonyításához nem elég az 1. lépést csak a $P = \{0\}$ esetben igazolni!

1. Tekintsük azt a speciális esetet, mikor $C = K$, és $P = R$, ahol K és R is kúp, K ráadásul csúcsos. A tétel feltétele ekkor azt követeli, hogy

$Ax \in -R$, $x \in K$ esetén $x = 0$ legyen. Mivel R poliéder kúp, azért létezik B mátrix úgy, hogy $R = \{z \in \mathcal{R}^m : Bz \leq 0\}$. A feltétel ekkor úgy írható le tömören, hogy $\{x \in K : BAx \geq 0\} = \{0\}$. Legyen $x_k \in K, z_k \in R$ $(k = 1, 2, \ldots)$, és tegyük fel, hogy $Ax_k + z_k \to y$ $(k \to \infty)$. Azt kell megmutatnunk, hogy ekkor $y \in AK + R$.

Ha az $\{x_k : k = 1, 2, \ldots\}$ halmaz korlátos, akkor feltehető, hogy a megfelelő sorozat konvergens, vagyis $x_k \to x$ $(k \to \infty)$ valamely $x \in K$ esetén. Ekkor $Ax_k \to Ax$ $(k \to \infty)$, így $z_k \to y - Ax$ $(k \to \infty)$. Az R zártsága miatt $y - Ax \in R$, amiből $y \in AK + R$ következik.

Megmutatjuk, hogy $\{x_k : k = 1, 2, \ldots\}$ nem lehet korlátlan. Ellenkező esetben feltehető, hogy $||x_k|| \to \infty$ $(k \to \infty)$. Legyen $x'_k := x_k/||x_k||$ $(k = 1, 2, \ldots)$, ezekről az egységvektorokról feltehetjük, hogy konvergálnak valamely $x' \in K$ egységvektorhoz. Ellentmondáshoz jutunk, ha megmutatjuk, hogy $BAx' \geq 0$, akkor ugyanis csak $x' = 0$ lehetne, nem egységvektor. Tegyük fel indirekt, hogy például $(BAx')_1 < 0$. Mivel

$$||x_k||Ax'_k + z_k \to y \ (k \to \infty),$$

azért

$$||x_k||BAx'_k + Bz_k \to By \ (k \to \infty).$$

Itt $z_k \in R$ miatt $Bz_k \leq 0$ $(k = 1, 2, \ldots)$, és elég nagy k-kra már $(BAx'_k)_1 \leq (BAx')_1/2$, így az $||x_k||(BAx'_k)_1 + (Bz_k)_1$ sorozatot egy idő után majorálja a $-\infty$-hez tartó $||x_k||(BAx')_1/2$ sorozat, az előbbi sorozat nem tarthatna a véges $(By)_1$ értékhez.

2. Azt az esetet, mikor mint az előbb $C = K$, és $P = R$, de K már nem feltétlenül csúcsos, könnyen visszavezethetjük a már elintézett esetre. Ekkor ugyanis $K = L + (K \cap L^\perp)$, ahol $L := K \cap -K$ a K kúp linealitás tere. Az

$$AK + R = A(K \cap L^\perp) + (AL + R)$$

halmaz az előző pont szerint zárt lesz, ha

$$Ax' \in -(AL + R),\ x' \in K \cap L^\perp \text{ esetén } x' = 0.$$

E feltétel fennállása pedig könnyen belátható a tétel feltételéből, ami most $Ax \in -R$, $x \in K$ esetén $x \in -K$.

3. Most még mindig legyen $P = R$, de a másik halmaz már C zárt, konvex halmaz. Ez az eset homogenizációval vezethető vissza az előző pontra. Eszerint ugyanis

$$AC + R = C\left(\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix}\right)$$

zárt, ha

$$\left.\begin{array}{r}\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}\begin{pmatrix} \lambda \\ x \end{pmatrix} \in -\begin{pmatrix} 0 \\ R \end{pmatrix}, \\ \begin{pmatrix} \lambda \\ x \end{pmatrix} \in \operatorname{cl} K(C) \end{array}\right\} \text{ esetén } \begin{pmatrix} \lambda \\ x \end{pmatrix} \in -\operatorname{cl} K(C).$$

Utóbbi feltétel pedig éppen a tétel feltétele ebben a speciális esetben (vö. 4.20).

4. Az általános esetben Motzkin tétele szerint $P = Q + R$, ahol Q politóp, R pedig poliéder kúp, a P recessziós kúpja. Minden politóp kompakt, és azt is tudjuk, hogy egy kompakt és egy zárt halmaz összege zárt lesz. Ezért az

$$AC + P = Q + AC + R$$

halmaz zárt lesz, ha $AC + R$ zárt. Ez pedig az előző pont szerint bekövetkezik, ha teljesül, hogy $Ax \in -R$, $x \in \operatorname{rec}(C)$ esetén $x \in -\operatorname{rec}(C)$, ami éppen a tétel feltétele (vö. 4.13).

5. Tekintsük ismét a $P = R$ speciális esetet. Ahogy a harmadik pont bizonyításakor láttuk, a tétel feltételének fennállása esetén

$$\operatorname{cl}\left(\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix}.$$

Másfelől könnyen ellenőrizhetően

$$\begin{aligned} &\operatorname{cl}\left(\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix}\right) = \\ &\quad = \operatorname{cl}\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & A & E \end{pmatrix} \operatorname{cl}\begin{pmatrix} K(C) \\ R \end{pmatrix}\right) = \\ &\quad = \operatorname{cl}\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & A & E \end{pmatrix} \begin{pmatrix} K(C) \\ R \end{pmatrix}\right) = \\ &\quad = \operatorname{cl}\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & A & E \end{pmatrix} K\begin{pmatrix} C \\ R \end{pmatrix}\right) = \\ &\quad = \operatorname{cl} K(AC + R). \end{aligned}$$

Ezek szerint

$$\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix} = \operatorname{cl} K(AC + R).$$

A $\{0\} \times \mathcal{R}^m$ hipersíknak az egyenlőség bal oldalán szereplő halmazzal való metszete $\{0\} \times (A \operatorname{rec} C + R)$, a jobb oldali halmazzal való metszete pedig $\{0\} \times \operatorname{rec}(AC + R)$, így $A \operatorname{rec} C + R = \operatorname{rec}(AC + R)$.

6. Az általános esetre rátérve legyen Q és R, mint a 4. lépésben. Ekkor

$$\begin{aligned} &\operatorname{rec}(AC + P) = \operatorname{rec}(Q + AC + R) = \\ &\qquad = \operatorname{rec}(AC + R) = A \operatorname{rec} C + R = A \operatorname{rec} C + \operatorname{rec} P. \end{aligned}$$

Itt felhasználtuk, hogy egy zárt, konvex halmazhoz egy kompakt, konvex halmazt hozzáadva a megfelelő korlát kúp, és így a recessziós kúp sem változik (4.12), továbbá a tétel az előző lépésben már igazolt speciális esetét. □

4.24. Következmény: *Legyen $C \subseteq \mathcal{R}^n$ zárt, konvex halmaz, $A \in \mathcal{R}^{m \times n}$ egy mátrix, és tegyük fel, hogy*

$$(\operatorname{Ker} A) \cap (\operatorname{rec} C) \subseteq -\operatorname{rec} C.$$

Ekkor

$$\operatorname{cl}(AC) = AC, \text{ és } \operatorname{rec} AC = A \operatorname{rec} C.$$

Bizonyítás: Ez éppen a tétel $P = \{0\}$ esete. □

Sőt

4.25. Következmény: *Legyenek $C_1 \subseteq \mathcal{R}^n$, $C_2 \subseteq \mathcal{R}^m$ zárt, konvex halmazok, továbbá $A \in \mathcal{R}^{m \times n}$. Ha*

$$A^{-1}(-\operatorname{rec} C_2) \cap (\operatorname{rec} C_1) \subseteq (-\operatorname{rec} C_1) \cap A^{-1}(\operatorname{rec} C_2),$$

akkor

$$AC_1 + C_2 \text{ zárt, és recessziós kúpja } A \operatorname{rec} C_1 + \operatorname{rec} C_2.$$

Bizonyítás: 4.24 szerint $AC_1 + C_2 = (A, E)(C_1 \times C_2)$ zárt halmaz lesz, ha teljesül, hogy

$$\operatorname{Ker}(A, E) \cap ((\operatorname{rec} C_1) \times (\operatorname{rec} C_2)) \subseteq -((\operatorname{rec} C_1) \times (\operatorname{rec} C_2)).$$

□

Megfogalmazható 4.22 párja is (itt $k = 0$, vagy $l = 0$ is lehet):

4.26. Következmény: *Legyenek $C_1, \ldots, C_k \subseteq \mathcal{R}^d$ zárt, konvex halmazok, $P_1, \ldots, P_l \subseteq \mathcal{R}^d$ poliéderek, továbbá $\lambda_1, \ldots, \lambda_k \in \mathcal{R}$ és $\mu_1, \ldots, \mu_l \in \mathcal{R}$ tetszőleges számok. Ha teljesül, hogy*

$$\left.\begin{array}{r} z_i \in \operatorname{rec} C_i \ (i = 1, \ldots, k), \\ z'_j \in \operatorname{rec} P_j \ (j = 1, \ldots, l), \\ \sum_{i=1}^k \lambda_i z_i + \sum_{j=1}^l \mu_j z'_j = 0 \end{array}\right\} \textit{ esetén } z_i \in -\operatorname{rec} C_i \ (i = 1, \ldots, k),$$

akkor

$$\sum_{i=1}^k \lambda_i C_i + \sum_{j=1}^l \mu_j P_j \textit{ zárt, és recessziós kúpja } \sum_{i=1}^k \lambda_i \operatorname{rec} C_i + \sum_{j=1}^l \mu_j \operatorname{rec} P_j.$$

Bizonyítás: Legyen

$$A := (\lambda_1 E, \ldots, \lambda_k E), \ C := C_1 \times \ldots \times C_k, \text{ és } P := \sum_{j=1}^l \mu_j P_j.$$

Alkalmazzuk a tételt. □

4.27. Tétel: *Legyen $B \in \mathcal{R}^{m \times n}$ mátrix, $C \subseteq \mathcal{R}^m$ konvex halmaz. Tegyük fel, hogy $B^{-1}(\operatorname{ri} C) \neq \emptyset$. Ekkor*

$$\operatorname{ri} B^{-1}(C) = B^{-1}(\operatorname{ri} C), \textit{ és } \operatorname{cl} B^{-1}(C) = B^{-1}(\operatorname{cl} C).$$

Bizonyítás: Legyen $\hat{C} := \mathcal{R}^n \times C$, $\hat{L} := \operatorname{Im}(E, B^T)^T$, továbbá $V := (E, 0)$. Ekkor

$$\hat{L} \cap \hat{C} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : y = Bx, \ Bx \in C \right\}, V(\hat{L} \cap \hat{C}) = B^{-1}(C),$$

és hasonlóan C helyett $\operatorname{ri} C$-vel, illetve $\operatorname{cl} C$-vel. A feltétel miatt $\hat{L} \cap \operatorname{ri} \hat{C} \neq \emptyset$. Most már világos (vö. 4.16, 4.21), hogy

$$\operatorname{ri} B^{-1}(C) = \operatorname{ri} V(\hat{L} \cap \hat{C}) = V \operatorname{ri}(\hat{L} \cap \hat{C}) = V(\hat{L} \cap \operatorname{ri} \hat{C}) = B^{-1}(\operatorname{ri} C),$$

és

$$\operatorname{cl} B^{-1}(C) = \operatorname{cl} V(\hat{L} \cap \hat{C}) \supseteq V \operatorname{cl}(\hat{L} \cap \hat{C}) = V(\hat{L} \cap \operatorname{cl} \hat{C}) = B^{-1}(\operatorname{cl} C).$$

Másfelől a $\operatorname{cl} B^{-1}(C) \subseteq B^{-1}(\operatorname{cl} C)$ tartalmazás nyilvánvaló B folytonossága miatt. $\square$

Legyen $C \subseteq \mathcal{R}^2$ a nemnegatív ortáns belseje, megtoldva az origóval, B pedig az x_1-tengelyre vetítés. Ez a példa mutatja, hogy a tétel feltételére szükség van.

A recessziós kúpra vonatkozó megfelelő állítás

4.28. Állítás: *Legyen $B \in \mathcal{R}^{m \times n}$ mátrix, $C \subseteq \mathcal{R}^m$ zárt, konvex halmaz. Tegyük fel, hogy a $B^{-1}(C)$ konvex halmaz nemüres. Ekkor* $\operatorname{rec} B^{-1}(C) = B^{-1}(\operatorname{rec} C)$. $\square$

Legyen $C \subseteq \mathcal{R}^n$ konvex halmaz, $A \in \mathcal{R}^{m \times n}$ mátrix. Mivel

$$A(C + \operatorname{Ker} A) = AC, \text{ és } A^{-1}(AC) = C + \operatorname{Ker} A,$$

azért 4.21, 4.22 és 4.27 segítségével könnyen belátható, hogy az AC konvex halmaz pontosan akkor relatív nyílt, ha a $C + \operatorname{Ker} A$ konvex halmaz relatív nyílt. Hasonló állítás bizonyítható relatív nyílt halmazok helyett poliéderekkel. A zárt halmazokra vonatkozó állítás Abrams tétele, amelyet itt kicsit általánosabban fogalmazunk meg.

4.29. Tétel: (Abrams) *Legyenek $S_1 \subseteq \mathcal{R}^n$, $S_2 \subseteq \mathcal{R}^m$ tetszőleges halmazok, továbbá $A \in \mathcal{R}^{m \times n}$. Ha $AS_1 + S_2$ zárt halmaz, akkor $S_1 + A^{-1}(S_2)$ is zárt halmaz. Megfordítva is, ha $S_2 \subseteq \operatorname{Im} A$.*

Bizonyítás: Először azt látjuk be, hogy ha $AS_1 + S_2$ zárt, akkor $S_1 + A^{-1}(S_2)$ is zárt. Legyen $x_k \in S_1$, $v_k \in A^{-1}(S_2)$ $(k = 1, 2, \ldots)$, és tegyük fel, hogy

$$x_k + v_k \to z \ (k \to \infty).$$

Ekkor

$$Ax_k + Av_k \to Az \ (k \to \infty),$$

és $Az \in AS_1 + S_2$ az $AS_1 + S_2$ halmaz zártsága miatt. Léteznek tehát $x \in S_1$, $y \in S_2$ pontok úgy, hogy $Az = Ax + y$. Ekkor

$$z = x + (z - x) \in S_1 + A^{-1}(S_2),$$

amit igazolni kellett.

A fordított irányhoz legyen $x_k \in S_1$, $y_k \in S_2$ $(k = 1, 2, \ldots)$, és tegyük fel, hogy

$$Ax_k + y_k \to b \ (k \to \infty).$$

Ekkor az $S_2 \subseteq \operatorname{Im} A$ feltétel miatt $b \in \operatorname{Im} A$ teljesül. Továbbá nyilván

$$A^\dagger A x_k + A^\dagger y_k \to A^\dagger b \ (k \to \infty),$$

vagyis

$$x_k + v_k \to A^\dagger b \ (k \to \infty),$$

ahol

$$v_k := A^\dagger y_k - x_k + A^\dagger A x_k \ (k = 1, 2, \ldots).$$

Mivel az $S_2 \subseteq \operatorname{Im} A$ feltételből adódóan $v_k \in A^{-1}(S_2)$ $(k = 1, 2, \ldots)$, azért $S_1 + A^{-1}(S_2)$ zártsága miatt $A^\dagger b \in S_1 + A^{-1}(S_2)$, és akkor $b = AA^\dagger b \in AS_1 + S_2$, amit igazolni kellett. □

Az A mátrix legyen $0 \in \mathcal{R}^{1\times 1}$; $S_1 \subseteq \mathcal{R}$ tetszőleges; $S_2 \subseteq \mathcal{R}$ pedig egy nem zárt halmaz úgy, hogy $0 \in S_2$. Ez a példa mutatja, hogy a tétel feltételére szükség van.

A következő tétel 4.2 általánosítása.

4.30. Tétel: *Legyenek $C_1, \ldots, C_k \subseteq \mathcal{R}^d$ nemüres, konvex halmazok. Jelölje C az uniójuk konvex burkát. Ekkor*

$$C = \cup \left\{ \sum_{i=1}^k \lambda_i C_i : \lambda_i \geq 0 \ (i = 1, \ldots, k), \ \sum_{i=1}^k \lambda_i = 1 \right\},$$

és

$$\operatorname{ri} C = \cup \left\{ \sum_{i=1}^k \lambda_i \operatorname{ri} C_i : \lambda_i > 0 \ (i = 1, \ldots, k), \ \sum_{i=1}^k \lambda_i = 1 \right\}.$$

Bizonyítás: Az első egyenlőség nyilvánvaló: $\cup C_i$-beli elemek egy konvex kombinációjáról feltehető, hogy tagjai különböző i indexeknek megfelelő C_i halmazbeli elemek konvex kombinációi. Az egyforma C_i-hez tartozó tagokat ugyanis csoportosíthatjuk, súlyaik összegét kiemelve összegükből.

Az első egyenlőségből adódik, hogy $C = C(\sum K(C_i))$, így $K(C) = \sum K(C_i)$ (vö. 3.20), és akkor 4.22 szerint $\operatorname{ri} K(C) = \sum \operatorname{ri} K(C_i)$. Az utóbbi egyenlőségre a $C(.)$ operációt alkalmazva máris adódik a kívánt második egyenlőség (vö. 4.20). □

4.31. Tétel: *Legyenek $C_1, \ldots, C_k \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmazok. Jelölje C az úniójuk konvex burkát. Tegyük fel, hogy*

$$z_i \in \operatorname{rec} C_i \ (i = 1, \ldots, k), \ \sum_{i=1}^{k} z_i = 0 \ \textit{esetén} \ z_i \in -\operatorname{rec} C_i \ (i = 1, \ldots, k).$$

Ekkor

$$\operatorname{cl} C = \cup \left\{ \sum_{i=1}^{k} \lambda_i C_i : \lambda_i \geq 0^+ \ (i = 1, \ldots, k), \ \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

(ahol a $\lambda_i \geq 0^+$ jelölés azt jelenti, hogy a $\lambda_i C_i$ halmazt $\{0\}$ helyett $\operatorname{rec} C_i$-nek tekintjük, ha $\lambda_i = 0$), és

$$\operatorname{rec} \operatorname{cl} C = \sum_{i=1}^{k} \operatorname{rec} C_i.$$

Ha az is teljesül, hogy minden C_i halmaz poliéder, akkor a tétel feltételére sincs szükség a fenti egyenlőségekhez, és $\operatorname{cl} C$ poliéder lesz.

Bizonyítás: A tétel feltétele miatt, 4.26 szerint (vö. 4.20 és az előző tétel bizonyítása)

$$\operatorname{cl} K(C) = \operatorname{cl} \sum K(C_i) = \operatorname{cl} \sum \operatorname{cl} K(C_i) = \sum \operatorname{cl} K(C_i).$$

A kiemelt egyenlőségre alkalmazva a $C(.)$ operációt, a tételben szereplő első egyenlőséget kapjuk (vö. 3.20, 4.20). Ha pedig mindkét oldalt elmetsszük a $\{0\} \times \mathcal{R}^d$ hipersíkkal, akkor a tételbeli második egyenlőséghez jutunk (felhasználva még 4.20-at, és hogy 3.18 a) szerint $\operatorname{cl} K(C) = \operatorname{cl} K(\operatorname{cl} C)$).

Ha minden C_i halmaz poliéder, akkor is

$$\operatorname{cl} K(C) = \operatorname{cl} \sum K(C_i) = \operatorname{cl} \sum \operatorname{cl} K(C_i) = \sum \operatorname{cl} K(C_i),$$

mivel az utóbbi kúp 4.20 szerint poliéder. Azt kell még belátnunk, hogy a $\operatorname{cl} C$ halmaz poliéder. A C_i poliéderekhez Motzkin tétele szerint léteznek S_i, S_i' véges halmazok úgy, hogy

$$C_i = (\operatorname{conv} S_i) + (\operatorname{cone} S_i') \ (i = 1, \ldots, k).$$

Legyen

$$P := \operatorname{conv}(\cup S_i) + \operatorname{cone}(\cup S_i'),$$

ekkor egyrészt nyilván P poliéder, és $\cup C_i \subseteq P$, tehát $\mathrm{cl}\, C \subseteq P$. Másrészt $C_i \subseteq \mathrm{cl}\, C$ miatt $\mathrm{rec}\, C_i \subseteq \mathrm{rec}\,\mathrm{cl}\, C$ (vö. 4.9 b)). 4.13 szerint $\mathrm{rec}\, C_i = \mathrm{cone}\, S_i'$. Ezért

$$\mathrm{rec}\, P = \mathrm{cone}\,(\cup S_i') = \sum \mathrm{cone}\, S_i' = \sum \mathrm{rec}\, C_i \subseteq \mathrm{rec}\,\mathrm{cl}\, C.$$

Mivel nyilvánvalóan $\mathrm{conv}\,(\cup S_i) \subseteq \mathrm{cl}\, C$, azért ebből $P \subseteq \mathrm{cl}\, C$ is látszik. □

4.32. Következmény: *Ha $C_1, \ldots, C_k \subseteq \mathcal{R}^n$ nemüres, zárt, konvex halmazok, amelyeknek ugyanaz a $K \subseteq \mathcal{R}^n$ kúp a recessziós kúpja, akkor a $C := \mathrm{conv}\,(C_1 \cup \ldots \cup C_k)$ konvex halmaz zárt, és szintén a K kúp a recessziós kúpja.*

Bizonyítás: Legyen $z_i \in K$ $(i = 1, \ldots, k)$, és tegyük fel, hogy $z_1 + \ldots + z_k = 0$. Ekkor

$$-z_1 = z_2 + \ldots + z_k \in (-K) \cap K,$$

és hasonlóan z_1 helyett a $z_2, \ldots, z_k$ vektorokkal. Ezért 4.31 alkalmazható. Ráadásul a $\lambda_i \geq 0^+$ helyett $\lambda_i \geq 0$ írható a $\mathrm{cl}\, C$ előállításában, ugyanis

$$\mathrm{rec}\, C_i + \lambda_j C_j = \lambda_j (K + C_j) = \lambda_j C_j = 0 C_i + \lambda_j C_j$$

valahányszor $\lambda_j > 0$. 4.30-at is figyelembe véve ebből már látszik, hogy $\mathrm{cl}\, C = C$. □

5. Szeparációs tételek

A harmadik fejezetben már definiáltuk halmazok szeparálhatóságát, erős szeparálhatóságát, és szerepeltek alapvető elválasztási tételek, mint például a kompakt és zárt halmaz erős elválaszthatóságáról szóló Hahn–Banach-tétel, vagy a két poliéder erős elválaszthatóságáról szóló tétel.

Most két újabb szeparálhatósági fogalmat vezetünk be.

Legyenek $S_1, S_2 \subseteq \mathcal{R}^d$ nemüres halmazok. Azt mondjuk, hogy az S_1 és S_2 halmazok **valódi módon szeparálható**ak, ha létezik őket elválasztó hipersík úgy, hogy az nem tartalmazza egyszerre mindkét halmazt. (Vagyis létezik $0 \neq a \in \mathcal{R}^d$ vektor úgy, hogy $\sup a^T S_1 \leq \inf a^T S_2$, és $\inf a^T S_1 < \sup a^T S_2$.)

Azt mondjuk, hogy az S_1 és S_2 halmazok **szigorúan szeparálható**ak, ha létezik őket elválasztó hipersík úgy, hogy az nem metsz bele az egyik vagy a másik halmazba. (Vagyis létezik $0 \neq a \in \mathcal{R}^d$ vektor úgy, hogy $a^T S_1 < a^T S_2$, ezalatt természetesen azt értve, hogy $a^T x_1 < a^T x_2$ minden $x_1 \in S_1, x_2 \in S_2$ esetén.)

Néhány megjegyzés a bevezetett szeparálhatóságokkal kapcsolatban:
– Nyilvánvaló, hogy ha az S_1 és S_2 halmazok erősen szeparálhatóak, akkor szigorúan szeparálhatóak, és ha szigorúan szeparálhatóak, akkor valódi módon is szeparálhatóak. Tehát elválaszthatóságok egy hierarchiáját definiáltuk a fentiekben.
– Az is világos, hogy például S_1 és S_2 pontosan akkor szeparálható, ha $\operatorname{conv} S_1$ és $\operatorname{conv} S_2$ szeparálható, így nem jelentett volna valódi megszorítást a fenti elválaszthatóságokat csak konvex halmazok esetén definiálni.
– A szigorú szeparálhatóság hagyományos definíciója az, hogy létezzen egy elválasztó hipersík úgy, hogy az általa meghatározott egyik nyílt féltér tartalmazza az egyik halmazt, a másik nyílt féltér pedig a másik halmazt. A fenti definíció használhatóbb. Például igaz, hogy két konvex halmaz pontosan akkor választható el valódi módon, ha relatív belsejeik szigorúan szeparálhatóak. Ehhez mindenekelőtt egy lemmára lesz szükségünk, amely azt fogalmazza meg, hogy ha egy konvex halmaz egy zárt féltér része, de nem része a félteret határoló hipersíknak, akkor a halmaz relatív belseje a zárt féltér belsejében van.

5.1. Lemma: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, konvex halmaz, $a \in \mathcal{R}^d$, $\beta \in \mathcal{R}$. Tegyük fel, hogy $a^T C \leq \beta$, és létezik $x \in C$ pont úgy, hogy $a^T x < \beta$. Ekkor $a^T \operatorname{ri} C < \beta$.*

Bizonyítás: Válasszunk egy tetszőleges $x_0 \in \operatorname{ri} C$ pontot. Ekkor az állításban szereplő $x \in C$ pont túlhúzható ezen a C halmazban, vagyis létezik $x' \in C$ pont úgy, hogy x_0 az $[x, x']$ szakasz belsejének eleme. Mivel $a^T x < \beta$, és $a^T x' \leq \beta$, azért csak $a^T x_0 < \beta$ lehet, amit igazolni kellett. $\square$

5.2. Állítás: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, konvex halmazok. A C_1 és C_2 halmazok pontosan akkor választhatók el valódi módon, ha a $\operatorname{ri} C_1$ és $\operatorname{ri} C_2$ halmazok szigorúan szeparálhatók.*

Bizonyítás: Először is ha C_1 és C_2 valódi módon szeparálhatóak, akkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$\sup a^T C_1 \leq \inf a^T C_2, \text{ és } \inf a^T C_1 < \sup a^T C_2.$$

Ekkor vagy létezik $x_1 \in C_1$ pont úgy, hogy $a^T x_1 < \sup a^T C_1$, vagy létezik $x_2 \in C_2$ pont úgy, hogy $a^T x_2 > \inf a^T C_2$ (vagy C_1 és C_2 párhuzamos hipersíkok részei, amikor is az állítás triviális). Például az előbbi esetben legyen $\beta := \sup a^T C_1$, ekkor a lemma szerint $a^T \operatorname{ri} C_1 < \beta$. Továbbá nyilván $\beta \leq a^T \operatorname{ri} C_2$, így $\operatorname{ri} C_1$ és $\operatorname{ri} C_2$ szigorúan szeparálhatóak.

Megfordítva, ha $\operatorname{ri} C_1$ és $\operatorname{ri} C_2$ szigorúan szeparálhatóak, akkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$a^T \operatorname{ri} C_1 < a^T \operatorname{ri} C_2.$$

Ekkor nyilván $\sup a^T C_1 \leq \inf a^T C_2$, sőt $\sup a^T \operatorname{cl} C_1 \leq \inf a^T \operatorname{cl} C_2$. Másrészt ha $x_1 \in \operatorname{ri} C_1, x_2 \in \operatorname{ri} C_2$, akkor

$$\inf a^T C_1 \leq a^T x_1 < a^T x_2 \leq \sup a^T C_2.$$

Látjuk, hogy C_1 és C_2 valódi módon szeparálhatóak. □

5.3. Tétel: (második Hahn–Banach-tétel) *A $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, konvex halmazok pontosan akkor választhatóak el valódi módon, ha relatív belsejeik diszjunktak. Ekkor létezik $a \in \operatorname{cl} \operatorname{cone}(C_2 - C_1)$ vektor úgy, hogy*

$$\sup a^T C_1 \leq \inf a^T C_2, \text{ és } \inf a^T C_1 < \sup a^T C_2.$$

Bizonyítás: 5.2 miatt világos, hogy ha C_1 és C_2 valódi módon szeparálhatóak, akkor $\operatorname{ri} C_1$ és $\operatorname{ri} C_2$ szigorúan szeparálhatóak, speciálisan diszjunktak.

A fordított irányra három bizonyítást is adunk, egy-egy speciális esetben igazolva azt először.

Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz, $x_0 \in \operatorname{rb} C$. A C és $\{x_0\}$ halmazokat [valódi módon] elválasztó hipersíkokat a C halmaz x_0-beli [**valódi**] **támaszhipersík**jainak nevezzük.

5.4. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz, $x_0 \in \operatorname{rb} C$. Ekkor létezik a C halmaznak x_0-beli valódi támaszhipersíkja, sőt létezik $a \in \operatorname{cl} \operatorname{cone}(C - x_0)$ vektor úgy, hogy*

$$a^T x_0 = \inf a^T C, \text{ és } a^T x_0 < \sup a^T C.$$

Bizonyítás: Legyen x_k $(k = 1, 2, \ldots)$ az x_0 ponton egy rögzített $\mathrm{ri}\, C$-beli pontot túlhúzva nyert pontok egy x_0-hoz tartó sorozata. Ekkor

$$x_k \in (\mathrm{aff}\, C) \setminus (\mathrm{cl}\, C) \ (k = 1, 2, \ldots),$$

így az első Hahn–Banach-tétel szerint léteznek $a_k \in (\mathrm{cl}\, C) - x_k$ vektorok úgy, hogy

$$a_k^T x_k < \inf a_k^T \mathrm{cl}\, C.$$

Könnyen belátható, hogy

$$a_k \in \mathrm{cone}\,(C - x_0).$$

Az a_k vektorokról még az is feltehető, hogy egységvektorok, és hogy konvergálnak valamely $a \in \mathrm{cl\,cone}\,(C - x_0)$ egységvektorhoz. Erre az a vektorra teljesül, hogy

$$a^T x_0 = \inf a^T \mathrm{cl}\, C = \inf a^T C,$$

és nem lehet, hogy az

$$\left\{ x \in \mathcal{R}^d : a^T x = a^T x_0 \right\}$$

hipersík tartalmazza a C halmazt, akkor ugyanis az a vektor merőleges lenne a $C - x_0$ halmaz, és így a $\mathrm{cl\,cone}\,(C - x_0)$ halmaz minden elemére, holott önmagára nem merőleges. $\square$

Ebből a speciális esetből már következik a második Hahn–Banach-tétel. Annak $(\mathrm{ri}\, C_1) \cap (\mathrm{ri}\, C_2) = \emptyset$ feltétele ugyanis azt jelenti, hogy

$$0 \notin (\mathrm{ri}\, C_2) - (\mathrm{ri}\, C_1), \text{ vagyis } 0 \notin \mathrm{ri}\,(C_2 - C_1).$$

Ha még az is teljesül, hogy $0 \notin \mathrm{cl}\,(C_2 - C_1)$, akkor az első Hahn–Banach-tétel szerint létezik $a \in \mathrm{cl}\,(C_2 - C_1)$ vektor úgy, hogy

$$0 < \inf a^T (C_2 - C_1), \text{ vagyis } \sup a^T C_1 < \inf a^T C_2.$$

Ha pedig csak annyi igaz, hogy $0 \in \mathrm{rb}\,(C_2 - C_1)$, akkor 5.4 szerint létezik $a \in \mathrm{cl\,cone}\,(C_2 - C_1)$ vektor úgy, hogy

$$0 = \inf a^T (C_2 - C_1), \text{ és } 0 < \sup a^T (C_2 - C_1),$$

vagyis

$$\sup a^T C_1 = \inf a^T C_2, \text{ és } \inf a^T C_1 < \sup a^T C_2.$$

Ezzel a második Hahn–Banach-tétel első bizonyítását befejeztük. □

5.5. Állítás: *Tegyük fel, hogy az $L \subseteq \mathcal{R}^d$ altér nem metsz bele a $K \subseteq \mathcal{R}^d$ konvex kúp relatív belsejébe. Ekkor létezik őket valódi módon elválasztó hipersík, sőt létezik $a \in \mathrm{cl}\,(K - L)$ vektor úgy, hogy*

$$a \in L^{\perp} \cap K^*, \textit{ és } 0 < a^T \mathrm{ri}\, K.$$

Bizonyítás: Legyen x_k $(k = 1, 2, \ldots)$ $\mathrm{ri}\, K$-beli pontok egy az origóhoz tartó sorozata. (Például egy rögzített $\mathrm{ri}\, K$-beli pont $(1/k)$-szorosai megfelelnek.) Az elérhetőségi lemma segítségével könnyen belátható, hogy ha egy konvex kúphoz hozzáadjuk a relatív belsejét, akkor eredményül a relatív belsejét kapjuk. Ezért az $x_k + \mathrm{cl}\, K$ zárt, konvex halmazok $\mathrm{ri}\,\mathrm{cl}\, K = \mathrm{ri}\, K$ részhalmazai, így diszjunktak az L altértől, speciálisan annak

$$O := \{x \in L : ||x|| \leq 1\}$$

kompakt, konvex részétől is. Az első Hahn–Banach-tétel szerint léteznek $a_k \in (x_k + \mathrm{cl}\, K) - O$ vektorok úgy, hogy

$$\sup a_k^T O < \inf a_k^T (x_k + \mathrm{cl}\, K) \; (k = 1, 2, \ldots).$$

Könnyen belátható, hogy

$$a_k \in K - L \; (k = 1, 2, \ldots).$$

Még az is feltehető, hogy az a_k vektorok egységvektorok, és sorozatuk konvergál egy $a \in \mathrm{cl}\,(K - L)$ egységvektorhoz. Megmutatjuk, hogy az a vektor megfelel az állítás kívánalmainak.

Először is ha $0 \neq x \in L$, akkor $x/||x|| \in O$, és persze $0 \in \mathrm{cl}\, K$, ezért

$$a_k^T (x/||x||) < a_k^T (x_k + 0) \; (k = 1, 2, \ldots),$$

amiből határértéket véve $a^T x \leq 0$ következik. Tehát $a \in -L^* = L^{\perp}$.

Másodszor ha $x \in K$, akkor $x \in \mathrm{cl}\, K$, és persze $0 \in O$, ezért

$$a_k^T 0 < a_k^T (x_k + x) \; (k = 1, 2, \ldots),$$

amiből határértéket véve $a^T x \geq 0$ következik. Tehát $a \in K^*$.

Most már könnyen belátható, hogy $a^T \operatorname{ri} K > 0$. Tegyük fel indirekt, hogy valamely $x_0 \in \operatorname{ri} K$ esetén $a^T x_0 = 0$. Ekkor tetszőleges $x \in \operatorname{lin} K$ pontot behúzhatunk az x_0 pont mellé a K halmazba, ugyanakkor túlhúzhatunk az x_0 ponton a K halmazban, vagyis léteznek $x', x'' \in K$ pontok úgy, hogy x_0 az $[x', x'']$ szakasz belső pontja, x pedig az x', x'' pontok által meghatározott egyenesé. Ekkor $a \in K^*$ miatt $a^T x' \geq 0$, és $a^T x'' \geq 0$; $a^T x_0 = 0$ csak úgy lehet, ha $a^T x' = 0 = a^T x''$, de akkor $a^T x = 0$ is teljesülne. Tehát $a \in (\operatorname{lin} K)^\perp$ lenne, és így

$$a \in L^\perp \cap (\operatorname{lin} K)^\perp = (L + \operatorname{lin} K)^\perp.$$

Mivel ugyanakkor nyilván $a \in \operatorname{lin}(K - L)$, és $\operatorname{lin}(K - L) = L + \operatorname{lin} K$, azért csak $a = 0$ lehetne, ami ellentmond annak, hogy a egységvektor. Ezzel beláttuk azt is, hogy $a^T \operatorname{ri} K > 0$. □

Ebből a speciális esetből is következik a második Hahn–Banach-tétel. Ugyanis, mint az előbb, annak $(\operatorname{ri} C_1) \cap (\operatorname{ri} C_2) = \emptyset$ feltétele azt jelenti, hogy $0 \notin (\operatorname{ri} C_2) - (\operatorname{ri} C_1)$, vagyis $0 \notin \operatorname{ri}(C_2 - C_1)$. Legyen

$$C := C_2 - C_1, \text{ továbbá } K := K(C), \text{ és } L := \mathcal{R} \times \{0\}.$$

Ekkor $\operatorname{ri} K(C) = K(\operatorname{ri} C) \setminus \{0\}$ miatt L nem metsz bele a K konvex kúp relatív belsejébe (különben $0 \in \operatorname{ri} C$ lenne). Ezért alkalmazható 5.5, mely szerint létezik $a \in \mathcal{R}^d$ vektor és $\alpha \in \mathcal{R}$ szám úgy, hogy

$$\begin{pmatrix} \alpha \\ a \end{pmatrix} \in \operatorname{cl}(K - L) \cap L^\perp \cap K^*, \text{ és } \begin{pmatrix} \alpha \\ a \end{pmatrix}^T \operatorname{ri} K > 0.$$

Mivel $L^\perp = \{0\} \times \mathcal{R}^d$, azért $\alpha = 0$. Könnyen belátható, hogy $K - L = \mathcal{R} \times \operatorname{cone} C$, ezért abból, hogy $(0, a^T)^T \in \operatorname{cl}(K - L)$, már következik, hogy $a \in \operatorname{cl} \operatorname{cone} C$ is teljesül. Végül abból, hogy $(0, a^T) \operatorname{ri} K > 0$, adódik, hogy $a^T C(\operatorname{ri} K) = a^T \operatorname{ri} C > 0$, vagyis $a^T \operatorname{ri} C_1 < a^T \operatorname{ri} C_2$. Eszerint az a vektor megfelel a tétel kívánalmainak (lásd még 5.2 bizonyítását). Ezzel a második Hahn–Banach-tétel második bizonyítását is befejeztük. □

Mindkét bizonyítás azon az észrevételen alapult, hogy az egyik konvex halmazról feltehető, hogy éppen a $\{0\}$ halmaz. Ezután viszont duálisak abban az értelemben, hogy míg az elsőben a pontot távolítottuk a halmaztól a halmaz külseje felé, addig a másodikban a halmazt távolítottuk a ponttól a halmaz belseje felé.

A második Hahn–Banach-tétel harmadik bizonyítása lényegesen különbözik az első kettőtől: bár nem látszik belőle, hogy $a \in \mathrm{cl\,cone}\,(C_2 - C_1)$, de nem használja a Bolzano–Weierstrass-tételt.

Bizonyítás: A bizonyítás öt lépésből áll.

1. Feltehető, hogy $C_1 = L$ altér, $C_2 = K$ pedig konvex kúp: ebből a speciális esetéből már következik a tétel, mint azt az előző bizonyítás végén láttuk.

2. Feltehető, hogy $L = \{0\}$. Jelölje ugyanis V az $L^\perp$ altérre való vetítés mátrixát. Könnyen belátható, hogy a VK konvex kúp relatív belseje nem tartalmazza az origót. Valóban, 4.21 szerint $\mathrm{ri}\,(VK) = V \mathrm{ri}\, K$. Ha létezne $x \in \mathrm{ri}\, K$ pont, amelyre $Vx = 0$, akkor $x = x - Vx \in L$ is teljesülne, holott L és $\mathrm{ri}\, K$ diszjunkt halmazok. Feltevésünk szerint létezik $a_0 \in \mathcal{R}^d$ vektor úgy, hogy $0 < a_0^T \mathrm{ri}\,(VK)$. Ismét 4.21 szerint az $a = V^T a_0$ vektorra teljesül, hogy $0 < a^T \mathrm{ri}\, K$. Továbbá $V = V^T$ miatt $a = Va_0 \in L^\perp$.

3. Feltehető, hogy $\dim K = d$. Ez egy standard fogás segítségével bizonyítható, egyszer érdemes végiggondolni. Ha ugyanis $\dim K = k < d$, akkor válasszuk az $\mathrm{aff}\, K$ affin halmaz egy K-beli $S = \{0, x_1, \ldots, x_k\}$ affin bázisát. Létezik $T : \mathcal{R}^d \to \mathcal{R}^d$ átkoordinátázás, amely a $0, x_1, \ldots, x_k \in \mathcal{R}^d$ pontokat rendre a $0, e_1, \ldots, e_k \in \mathcal{R}^d$ pontokba viszi. Legyen ez éppen $T(x) = Nx + v$ ($x \in \mathcal{R}^d$), ahol $N \in \mathcal{R}^{d\times d}$ invertálható mátrix, $v \in \mathcal{R}^d$ vektor, most $v = 0$. Legyen továbbá $W := (E, 0) \in \mathcal{R}^{k\times d}$. Jelölje K' a $WT(K) \subseteq \mathcal{R}^k$ konvex kúpot. Mivel K' tartalmazza a $WT(S)$ halmazt, a $0, e_1, \ldots, e_k \in \mathcal{R}^k$ vektorokat, azért $\dim K' = k$. Könnyen belátható, hogy $0 \notin \mathrm{ri}\, K'$. Valóban, 4.21 szerint $\mathrm{ri}\, K' = WT(\mathrm{ri}\, K)$. Ha létezne $x \in \mathrm{ri}\, K$ pont, amelyre $WT(x) = 0$, akkor x előállna az S-beli pontok affin kombinációjaként, $T(x)$ pedig a $T(S)$-beli pontok ugyanilyen együtthatókkal vett affin kombinációjaként. A $WT(x) = 0$ egyenlőségből $T(x) = 0$, és így $x = 0$ következne, holott $0 \notin \mathrm{ri}\, K$. Feltevésünk szerint létezik $a_0 \in \mathcal{R}^k$ vektor úgy, hogy $0 < a_0^T \mathrm{ri}\, K'$. Mivel 4.21 szerint $\mathrm{ri}\, K' = WN \mathrm{ri}\, K$, azért az $a := (a_0^T WN)^T \in \mathcal{R}^d$ vektorra teljesül, hogy $0 < a^T \mathrm{ri}\, K$.

4. Elég tehát azt igazolnunk, hogy ha a $K \subseteq \mathcal{R}^d$ konvex kúp belseje nem üres, de nem tartalmazza az origót, akkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $0 < a^T \mathrm{int}\, K$. (A $\dim K = d$ feltétel ugyanis azzal ekvivalens, hogy K belseje nem üres, vö. 4.7.)

Ebben a pontban megmutatjuk, hogy vagy található egy $a \in \mathcal{R}^d$ vektor, amelyre $0 < a^T \mathrm{int}\, K$, vagy létezik $x \in \mathcal{R}^d$ vektor úgy, hogy $x \notin \mathrm{line\,int}\, K$, és $(x\mathcal{R}) \cap \mathrm{int}\, K = \emptyset$.

Valóban, legyen a_0 egy K belsejében levő pont vetülete a $\mathrm{line\,int}\,K$ altér ortogonális kiegészítő alterére. Ekkor $a_0 \in \mathrm{int}\,K$, így $a_0 \neq 0$. Tekintsük a $H := a_0 + a_0^{\perp}$ hipersíkot, ennek a K konvex kúppal való metszete tartalmazza az a_0 vektort, sőt alkalmas $\varepsilon > 0$ szám esetén az $O(a_0, \varepsilon) \cap H$ halmazt is. Ha a $H \cap K$ konvex halmaz relatív határa üres, akkor 4.5 és a fenti megjegyzés miatt $H \cap K = H$. Könnyen belátható, hogy ekkor $\mathrm{int}\,K = \{x : a_0^T x > 0\}$, tehát az $a := a_0$ vektor megfelel: $0 < a^T \mathrm{int}\,K$ teljesül. Ha a $H \cap K$ konvex halmaz relatív határa nem üres, akkor legyen x_0 egy pontja. Nyilván x_0 nem eleme a $\mathrm{line\,int}\,K$ altérnek (az a_0 vektor ezen altér ortogonális kiegészítőjének eleme, így $a_0^T x_0 = 0$ lenne, holott $x_0 \in H$ miatt $a_0^T x_0 = ||a_0||^2 \neq 0$). Továbbá $x_0 \in \mathrm{rb}\,K$, így $-x_0 \notin \mathrm{int}\,K$ (különben az elérhetőségi lemma miatt $0 \in \mathrm{int}\,K$ lenne). Ezért az $x_0\mathcal{R}$ egyenes nem metsz bele K belsejébe, tehát az $x := x_0$ vektor megfelel.

5. Az előző pont eredményéből már következik, hogy ha a $K \subseteq \mathcal{R}^d$ konvex kúp belseje nem üres, de nem tartalmazza az origót, akkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $0 < a^T \mathrm{int}\,K$. Ezt a következőképpen lehet belátni.

Legyen $K_1 := K$. Az előző pontban igazoltak szerint vagy létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $0 < a^T \mathrm{int}\,K_1$ (amikor is készen vagyunk a bizonyítással), vagy létezik $x_1 \notin \mathrm{line\,int}\,K_1$ vektor úgy, hogy az $x_1\mathcal{R}$ egyenes és $\mathrm{int}\,K_1$ metszete üres.

Az utóbbi esetben legyen $K_2 := (x_1\mathcal{R}) + K_1$. Ekkor K_2 konvex kúp, belseje $\mathrm{int}\,K_2 = (x_1\mathcal{R}) + \mathrm{int}\,K_1$. Ismét az előző pont eredményét alkalmazva vagy létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $0 < a^T \mathrm{int}\,K_2$ (amikor is készen vagyunk a bizonyítással), vagy létezik $x_2 \notin \mathrm{line\,int}\,K_2$ vektor úgy, hogy $(x_2\mathcal{R}) \cap \mathrm{int}\,K_2 = \emptyset$. Utóbbi esetben legyen $K_3 := (x_2\mathcal{R}) + K_2 \ldots$ stb.

Megmutatjuk, hogy a bizonyítás legfeljebb d lépésben véget ér, ugyanis a keletkező $x_1, x_2, \ldots$ vektorok lineárisan függetlenek. Valóban, $x_1 \notin \mathrm{line\,int}\,K_1$, így $x_1 \neq 0$. Továbbá $x_2 \notin \mathrm{line\,int}\,K_2$, így $x_2 \notin x_1\mathcal{R}$. Hasonlóképpen $x_3 \notin x_1\mathcal{R} + x_2\mathcal{R} \ldots$ stb. Az $x_1, x_2, \ldots$ vektorok lineáris függetlensége ezekből az észrevételekből és 1.5-ből azonnal adódik. □

5.2 segítségével a második Hahn–Banach-tétel az alábbi ekvivalens formában is megfogalmazható:

5.6. Tétel: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, relatív nyílt, konvex halmazok. Ha C_1 és C_2 diszjunktak, akkor szigorúan szeparálhatók.* □

Hasonló tétel mondható ki, ha az egyik halmaz poliéder.

5.7. Tétel: *Legyenek $C, P \subseteq \mathcal{R}^d$ diszjunkt, nemüres, konvex halmazok, C relatív nyílt, P poliéder. Ekkor P és C szigorúan szeparálhatók.*

Bizonyítás: Tekintsük először azt a speciális esetet, mikor $C = \operatorname{ri} K$ és $P = R$, ahol K és R konvex kúpok. Megmutatjuk, hogy a tétel feltételei mellett ekkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$a^T R \leq 0 < a^T \operatorname{ri} K.$$

Ezt az állítást $\dim(R \cup K)$ szerinti indukcióval bizonyítjuk. A $d = 1$ eset nyilvánvaló (R zárt, $\operatorname{ri} K$ pedig nyílt intervallum). Általában a második Hahn–Banach-tétel szerint létezik $a_1 \in \mathcal{R}^d$ vektor úgy, hogy

$$a_1^T \operatorname{ri} R < a_1^T \operatorname{ri} K.$$

Ekkor 4.7 miatt

$$a_1^T R \leq 0 \leq a_1^T K.$$

Ha K nem része az $a_1^\perp$ hipersíknak, akkor

$$a_1^T R \leq 0 < a_1^T \operatorname{ri} K$$

(vö. 5.1), és készen vagyunk a bizonyítással. Ezért feltehetjük, hogy

$$a_1^T R \leq 0 = a_1^T K.$$

Legyen P' az R kúp és a $-a_1 + a_1^\perp$ hipersík metszete. Ekkor P' poliéder, és nem üres, mivel $a_1^T \operatorname{ri} R < 0$. Motzkin tétele szerint létezik egy Q' politóp és egy R' végesen generált kúp úgy, hogy $P' = Q' + R'$. Ekkor 4.13 és 4.17 szerint $R' = R \cap a_1^\perp$, így az $R' \subseteq R$ kúp diszjunkt K relatív belsejétől. Mivel az $R' \cup K$ halmaz dimenziója kisebb, mint az $R \cup K$ halmazé, azért az indukció feltétel szerint létezik $a_2 \in \mathcal{R}^d$ vektor úgy, hogy

$$a_2^T R' \leq 0 < a_2^T \operatorname{ri} K.$$

Válasszuk a $\lambda \geq 0$ számot úgy, hogy

$$(a_2 + \lambda a_1)^T Q' \leq 0$$

teljesüljön. (Az egyenlőtlenséget csak Q' egy olyan véges részhalmazán kell biztosítani, amelynek konvex burkaként előáll a Q' politóp. Továbbá $x \in Q'$

esetén $a_1^T x = -||a_1||^2 < 0$.) Legyen $a := a_2 + \lambda a_1$, ekkor $a^T R' \le 0$ szintén teljesül, így $a^T P' \le 0$. Ebből pedig

$$a^T R \le 0 < a^T \operatorname{ri} K$$

könnyen adódik.

A tétel teljes általánosságban a fent igazolt speciális eset következménye, ha azt az $R := \operatorname{cl} K(P) \subseteq \mathcal{R}^{d+1}$ poliéder kúpra és a $K := K(C) \subseteq \mathcal{R}^{d+1}$ konvex kúpra alkalmazzuk. (4.20-ból adódóan $R \cap \operatorname{ri} K = \emptyset$, így a fentiek valóban alkalmazhatók.) $\square$

Ezzel az alábbi szimmetrikus tételtáblázat állításainak nagy részét beláttuk:

C_1/C_2	rel. nyílt	zárt	poliéder
rel. nyílt	szigorú	valódi	szigorú
zárt	valódi	valódi	valódi
poliéder	szigorú	valódi	erős

A táblázat azt fejezi ki röviden, hogy ha $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, diszjunkt, konvex halmazok, továbbá C_1 az első oszlopban leírt tulajdonságú, C_2 pedig az első sorban leírt tulajdonságú, akkor általában legfeljebb menynyire szeparálhatóak. Azt már mindenhol beláttuk, hogy ennyire szeparálhatók, még annyi kell, hogy jobban általában nem.

Legyen C_1 egy síkbeli zárt körlap, C_2 pedig egy hozzá húzott érintő egyenes az érintési pontban kezdődő, zárt szakasza. Ekkor $\operatorname{ri} C_1$ és $\operatorname{ri} C_2$, illetve $\operatorname{ri} C_1$ és C_2 mutatja, hogy a táblázatbeli "szigorú" szavak nem cserélhetők "erős"-re. A következő állítás bizonyítja, hogy a "valódi" szavak sem cserélhetők "szigorú"-ra.

5.8. Állítás: *Legyen $C \subseteq \mathcal{R}^3$ a 3.31-beli K zárt, konvex kúp $(0,1,0)^T$ vektorral való eltoltja, továbbá legyen $L := \mathcal{R} \times \{0\} \times \{0\} \subseteq \mathcal{R}^3$. Ekkor C zárt, konvex halmaz, L altér (speciálisan relatív nyílt, zárt, poliéder), L és C diszjunktak, de nem szigorúan szeparálhatóak.*

Bizonyítás: Ha C és L összemetszenének, akkor K-nak lenne olyan pontja, amelynek második koordinátája -1, harmadik koordinátája pedig 0. Ez nem lehetséges (lásd K leírását polinom-egyenlőtlenségekkel, 3.31-ben). Most megmutatjuk, hogy C és L nem szigorúan szeparálhatók. Tegyük fel indirekt, hogy létezik $a = (a_1, a_2, a_3)^T \in \mathcal{R}^3$ vektor úgy, hogy $a^T L < a^T C$. Ekkor persze $a \in L^\perp$, és $a^T C > 0$ lenne. Mivel $a \in L^\perp = \{0\} \times \mathcal{R} \times \mathcal{R}$, azért $a_1 = 0$. Legyen $a' := (a_2, a_3)^T \in \mathcal{R}^2$, és jelölje A ugyanazt a mátrixot, mint 3.31 bizonyításában. Ekkor $a = A^T a'$, ezért $a'^T(AC) > 0$. Itt

$$0 < a'^T(AC) = a'^T(A(0,1,0)^T + AK) = a_2 + a'^T(AK).$$

Mivel $\mathcal{R} \times \{1\} \subseteq AK$, azért $a_2 = 0$. Továbbá $(0,0)^T \in AK$, ezért $a_2 > 0$, ami ellentmondás. □

Érdekességként megjegyezzük, hogy már $\mathcal{R}^2$-ben található relatív nyílt, konvex halmaz és zárt, konvex halmaz úgy, hogy nem szigorúan szeparálhatók: C_1 és $\operatorname{ri} C_2$ ilyen, ahol C_1, C_2 az állítás előtti megjegyzésben szereplő halmazok. Ugyanakkor

5.9. Állítás: *Legyen C_1, $C_2 \subseteq \mathcal{R}^2$ két nemüres, zárt, konvex halmaz, amelyek diszjunktak. Ekkor C_1 és C_2 szigorúan szeparálható.*

Bizonyítás: A második Hahn–Banach-tétel szerint valódi módon szeparálhatók valamely $H \subseteq \mathcal{R}^2$ hipersíkkal (egyenessel). Akkor nem vagyunk készen, ha $C_1 \cap H \neq \emptyset \neq C_2 \cap H$. A $C_1 \cap H$ és a $C_2 \cap H$ nemüres, zárt, diszjunkt intervallumok H-ban, távolságuk pozitív és felvétetik. Legyen ez éppen $||c_1 - c_2||$, ahol $c_i \in C_i \cap H$ $(i = 1, 2)$. A $[c_1, c_2]$ szakasz felezőpontját jelölje f. Feltehető, hogy $f = 0$. Legyen a_i a legrövidebb C_i-beli vektor $(i = 1, 2)$. Ekkor az első Hahn–Banach-tétel bizonyítása szerint $a_i^T C_i \geq ||a_i||^2$ $(i = 1, 2)$. Legyen P_i az $\{x \in \mathcal{R}^2 : a_i^T x \geq ||a_i||^2\}$ félsík elmetszve a H által meghatározott, C_i-t tartalmazó, zárt félsíkkal $(i = 1, 2)$. Könnyen belátható, hogy P_1 és P_2 diszjunkt poliéderek (csak H-ban metszhetnének össze, de akkor H-beli részük úniója lefedné H-t, a 0 eleme lenne), így erősen szeparálhatóak. Mivel $C_1 \subseteq P_1$, és $C_2 \subseteq P_2$, az állítást ebben az esetben is beláttuk. □

Persze két zárt halmaz, sőt egy zárt halmaz és egy poliéder $\mathcal{R}^2$-ben sem szeparálható erősen: ezt egy síkbeli hiperbolalap és egyik aszimptotája mutatja.

Viszont a következő állítás segítségével elegendő feltételeket adhatunk arra, mikor a táblázatban szereplő "valódi" szavakat "erős"-re cserélhetjük.

A második Hahn–Banach-tétel szerint két nemüres, konvex halmaz pontosan akkor szeparálható valódi módon, ha az origó nincs benne a különbségük relatív belsejében. Hasonló állítás mondható ki az erős szeparálhatóságról.

5.10. Állítás: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, konvex halmazok. Ekkor az alábbi állítások ekvivalensek:*
a) C_1 és C_2 erősen szeparálhatók;
b) $C_1 + O(0,\varepsilon)$ és $C_2 + O(0,\varepsilon)$ szigorúan szeparálhatók valamely $\varepsilon > 0$ esetén;
c) $0 \notin \mathrm{cl}\,(C_2 - C_1)$.

Bizonyítás: Először belátjuk, hogy az a) állításból következik a b) állítás. Ha C_1 és C_2 erősen szeparálhatók, akkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $\sup a^T C_1 < \inf a^T C_2$, még az is feltehető, hogy a egységvektor. Legyen

$$\beta := \frac{\sup a^T C_1 + \inf a^T C_2}{2}, \text{ továbbá } \varepsilon := \beta - \sup a^T C_1,$$

ekkor $\varepsilon = \inf a^T C_2 - \beta > 0$. Megmutatjuk, hogy

$$a^T(C_1 + O(0,\varepsilon)) < \beta < a^T(C_2 + O(0,\varepsilon)).$$

Például legyen $x_1 \in C_1$, és $x \in \mathcal{R}^d$ olyan vektor, amelyre $||x - x_1|| < \varepsilon$ teljesül. Azt kell belátnunk, hogy $a^T x < \beta$. Ez így van, mivel

$$\begin{aligned} a^T x &= a^T x_1 + a^T(x - x_1) \leq \\ &\leq a^T x_1 + ||a|| \cdot ||x - x_1|| < \sup a^T C_1 + 1 \cdot \varepsilon = \beta. \end{aligned}$$

Ezzel beláttuk, hogy $a^T(C_1 + O(0,\varepsilon)) < \beta$, a $\beta < a^T(C_2 + O(0,\varepsilon))$ egyenlőtlenség hasonlóan igazolható. Tehát a $C_1 + O(0,\varepsilon)$ és $C_2 + O(0,\varepsilon)$ nyílt, konvex halmazok szigorúan szeparálhatók.

Másodszor azt igazoljuk, hogy a b) állításból következik a c) állítás. Ha a $C_1 + O(0,\varepsilon)$ és $C_2 + O(0,\varepsilon)$ halmazok szigorúan szeparálhatók, akkor egy $x_1 \in C_1$ és egy $x_2 \in C_2$ pont távolsága mindig legalább $\varepsilon > 0$, ezért egyetlen $C_2 - C_1$-beli sorozat sem tarthat az origóhoz, vagyis $0 \notin \mathrm{cl}\,(C_2 - C_1)$.

Végül megmutatjuk, hogy a c) állításnak következménye az a) állítás. Az első Hahn–Banach-tétel szerint, ha $0 \notin \mathrm{cl}\,(C_2 - C_1)$, akkor az origó erősen szeparálható a $\mathrm{cl}\,(C_2 - C_1)$ halmaztól, vagyis létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $0 < \inf a^T \mathrm{cl}\,(C_2 - C_1)$. Ekkor persze $0 < \inf a^T(C_2 - C_1)$, vagyis $\sup a^T C_1 < \inf a^T C_2$, tehát a C_1 és C_2 halmazok erősen szeparálhatóak. □

5.11. Állítás: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmazok. Tegyük fel, hogy $C_1 \cap C_2 = \emptyset$, továbbá hogy*

$$(\mathrm{rec}\, C_1) \cap (\mathrm{rec}\, C_2) \subseteq (-\mathrm{rec}\, C_1) \cap (-\mathrm{rec}\, C_2).$$

Ekkor C_1 és C_2 erősen szeparálhatók.

Bizonyítás: A feltétel miatt, 4.26 szerint $C_2 - C_1$ zárt. Mivel C_1 és C_2 diszjunktak, azért $0 \notin C_2 - C_1$. Az állítás ezek után 5.10 következménye.□

5.12. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmaz, amely diszjunkt a $P \subseteq \mathcal{R}^d$ nemüres poliédertől. Tegyük fel továbbá, hogy*

$$(\operatorname{rec} C) \cap (\operatorname{rec} P) \subseteq -\operatorname{rec} C.$$

Ekkor C és P erősen szeparálhatók.

Bizonyítás: 5.11 bizonyításához hasonlóan igazolható. □

5.13. Állítás: *Legyen $C_1 \subseteq \mathcal{R}^d$ nemüres, relatív nyílt, konvex halmaz, $C_2 \subseteq \mathcal{R}^d$ pedig nemüres, zárt, konvex halmaz, amely diszjunkt a C_1 halmaz lezártjától. Tegyük fel továbbá, hogy*

$$(\operatorname{rec} C_1) \cap (\operatorname{rec} C_2) \subseteq (-\operatorname{rec} C_1) \cap (-\operatorname{rec} C_2).$$

Ekkor C_1 és C_2 erősen szeparálhatók.

Bizonyítás: 5.11-ből adódik, felhasználva még 4.10-et. □

Még a Hahn–Banach-tételek a későbbiek szempontjából fontos következményét tárgyaljuk.

5.14. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz. Ekkor*
a) C az őt tartalmazó zárt félterek metszete;
b) Ha a C halmaz nem affin, akkor itt elég azokat a zárt féltereket venni, amelyek határoló hipersíkja tartalmaz valamely $\operatorname{rb} C$-beli pontot;
c) Sőt $\operatorname{rb} C$ helyett elegendő annak egy sűrű S részhalmazát tekinteni.

Bizonyítás: a) az első Hahn–Banach-tétel következménye. A nem C-beli pontok elválaszthatók C-től a pontot nem tartalmazó hipersíkkal, és akkor az e hipersík által meghatározott, a C halmazt tartalmazó zárt féltér már nem tartalmazza a pontot. Az összes C-t tartalmazó zárt féltér metszete egyetlen nem C-beli pontot sem fog tartalmazni, viszont a C-beli pontokat mind tartalmazza majd.

b) Mivel az $\operatorname{aff} C$ affin halmazt tartalmazó hipersíkok által meghatározott zárt félterek metszete $\operatorname{aff} C$, azért csak az $(\operatorname{aff} C) \setminus C$-beli pontokat kell elválasztani relatív határpontot tartalmazó hipersíkkal C-től. Az 5.3 harmadik bizonyításában látottakhoz hasonlóan okoskodva itt is feltehető, hogy $\dim C = d$, legyen $x_0 \in \operatorname{int} C$. Ha $x \notin C$, akkor az x_0 és x pontokat összekötő szakaszon van (éppen) egy $x' \in \operatorname{rb} C$ pont, válasszuk el ezt C-től egy H hipersíkkal úgy, hogy a megfelelő H^+ zárt féltér tartalmazza C-t, 5.4 szerint ez megtehető. Ekkor $x' \in H$, és $x \notin H^+$.

c) Mint az előbb, itt is feltehetjük, hogy $\dim C = d$. Legyen x_0, x, mint b)-ben, és $x' = x + \lambda(x_0 - x) \in \operatorname{rb} C$, ahol $0 < \lambda < 1$. Ha egy x_0 körüli, $\varepsilon > 0$ sugarú, nyílt gömb $\operatorname{int} C$ része, akkor az x' körüli, $\lambda\varepsilon$ sugarú, nyílt gömb bármely pontján túlhúzható az x pont úgy, hogy $\operatorname{int} C$-beli pontba jussunk, márpedig x' ezen környezetében van S-beli pont. b) bizonyítása alkalmazható. □

5.15. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ nemaffin, zárt, konvex halmaz, és tegyük fel, hogy $\dim C = d$. Ekkor a C halmaz határának sűrű részét alkotják az úgynevezett* **sima pont***ok, vagyis azok a határpontok, amelyeket C-nek pontosan egy támaszhipersíkja tartalmazza.*

Bizonyítás: Legyen adott $p \in \operatorname{rb} C$, $\varepsilon > 0$, azt kell megmutatnunk, hogy létezik p-től legfeljebb ε távolságra lévő s sima pont. Válasszunk egy $q \in \operatorname{int} C$ pontot úgy, hogy $||q - p|| < \varepsilon/2$ legyen. Legyen $r \geq 0$ a q pont távolsága $\mathcal{R}^d \setminus C$ halmaztól, nyilván $r \leq ||p - q||$. A q pont köré írt r sugarú O zárt gömb határa tartalmaz legalább egy $s \in \operatorname{rb} C$ pontot, és ez sima, ugyanis ha s-ben lenne két támaszhipersíkja C-nek, akkor azok különböző támaszhipersíkjai lennének s-ben O-nak is ($O \subseteq C$), ami lehetetlen. Végül

$$||p - s|| \leq ||p - q|| + ||q - s|| \leq 2||p - q|| < \varepsilon.$$

□

6. Kúplineáris alternatíva- és dualitási tételek

Most már készen állunk arra, hogy könnyen ellenőrizhető elégséges feltételeket keressünk az általános kúplineáris Farkas-lemmához és a kúplineáris Farkas-tételhez. E vizsgálatok alapját képezik Stiemke és Krein tételei, valamint különféle általánosításaik.

Stiemke és Krein tételei a következő vegyes tételek azon speciális esetei, mikor $R = L$ altér. Hasonlóan vezethetők le a második Hahn–Banach-tételből, mint vegyes általánosításaik 5.7-ből.

6.1. Tétel: (vegyes Stiemke-tétel) *Legyen* $R \subseteq \mathcal{R}^d$ *poliéder kúp,* $K \subseteq \mathcal{R}^d$ *konvex kúp. Pontosan akkor teljesül* $R \cap \operatorname{ri} K \neq \emptyset$, *ha* $(-R^*) \cap (K^*) \subseteq -K^*$.

Bizonyítás: Először tegyük fel, hogy $R \cap \operatorname{ri} K \neq \emptyset$, legyen $x_0 \in R \cap \operatorname{ri} K$, és válasszunk egy tetszőleges $a \in (-R^*) \cap (K^*)$ elemet. Ekkor $a^T x_0 \leq 0$ (mivel $x_0 \in R$, $a \in -R^*$), és $a^T x_0 \geq 0$ is (mivel $x_0 \in K$, $a \in K^*$), vagyis $a^T x_0 = 0$. Tetszőleges $x \in K$ pont túlhúzható az $x_0 \in \operatorname{ri} K$ ponton a K halmazban, vagyis létezik $\varepsilon > 0$ úgy, hogy $x_0 - \varepsilon(x - x_0) \in K$. Ekkor $a \in K^*$ miatt

$$0 \leq a^T(x_0 - \varepsilon(x - x_0)) = -\varepsilon a^T x,$$

vagyis oda jutottunk, hogy tetszőleges $x \in K$ esetén $a^T x \leq 0$. Ebből $a \in -K^*$ következik, tehát $(-R^*) \cap (K^*) \subseteq -K^*$.

A megfordításhoz tegyük fel, hogy $R \cap \operatorname{ri} K = \emptyset$, ekkor 5.7 szerint az R poliéder kúp és a $\operatorname{ri} K$ relatív nyílt halmaz szigorúan szeparálhatók, azaz létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $a^T R < a^T \operatorname{ri} K$, és akkor persze

$$a^T R \leq 0 < a^T \operatorname{ri} K.$$

Erre az a vektorra teljesül, hogy $a \in (-R^*) \cap (K^*)$, de $a \notin -K^*$. Ezzel beláttuk a tétel még nem igazolt irányát is. $\square$

6.2. Tétel: (vegyes Krein-tétel) *Legyen* $R \subseteq \mathcal{R}^d$ *poliéder kúp,* $K \subseteq \mathcal{R}^d$ *konvex kúp. Ha* $R \cap \operatorname{ri} K \neq \emptyset$, *akkor* $(R \cap K)^* = R^* + K^*$.

Bizonyítás: A nemtriviális $(R \cap K)^* \subseteq R^* + K^*$ tartalmazáshoz legyen $c \in (R \cap K)^*$. Azt kell megmutatnunk, hogy $c \in R^* + K^*$.

Tekintsük először azt a speciális esetet, mikor nem csak, hogy R belemetsz a K kúp relatív belsejébe, hanem ráadásul $\{c\}^\perp \cap R \cap \operatorname{ri} K \neq \emptyset$. Legyen $x_0 \in R \cap \operatorname{ri} K$, amelyre $c^T x_0 = 0$. Ekkor tetszőleges $x \in R \cap \operatorname{lin} K$ pont behúzható az x_0 pont mellé az $R \cap K$ halmazba, vagyis létezik $\varepsilon > 0$ úgy, hogy $x_0 + \varepsilon(x - x_0) \in R \cap K$. Mivel $c \in (R \cap K)^*$, azért

$$0 \leq c^T(x_0 + \varepsilon(x - x_0)) = \varepsilon c^T x,$$

vagyis oda jutottunk, hogy tetszőleges $x \in R \cap \operatorname{lin} K$ esetén $c^T x \geq 0$. Ebből

$$c \in (R \cap \operatorname{lin} K)^* = R^* + (\operatorname{lin} K)^\perp$$

következik, és mivel $(\operatorname{lin} K)^\perp \subseteq K^*$, azért ekkor $c \in R^* + K^*$ is fennáll.

Ha pedig $(\{c\}^\perp \cap R) \cap \operatorname{ri} K = \emptyset$, akkor 5.7 szerint létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$a^T(\{c\}^\perp \cap R) \le 0 < a^T \operatorname{ri} K.$$

Ekkor

$$a \in -(\{c\}^\perp \cap R)^* = \mathcal{R}c - R^*, \text{ és } a \in K^*.$$

Választható tehát $\lambda \in \mathcal{R}$, $b \in R^*$ úgy, hogy $a = \lambda c - b$ legyen. Ekkor $x_0 \in R \cap \operatorname{ri} K$ esetén egyrészt $c^T x_0 > 0$, másrészt

$$0 < a^T x_0 = \lambda c^T x_0 - b^T x_0 \le \lambda c^T x_0,$$

amiből $\lambda > 0$ következik. De akkor

$$c = \frac{1}{\lambda} b + \frac{1}{\lambda} a \in R^* + K^*,$$

és a tételt ebben az esetben is beláttuk. $\square$

A vegyes Krein-tételnek még egy bizonyítását adjuk, most az 5.7 szeparációs tétel helyett a vegyes Stiemke-tételből vezetjük le. Hasonlóan következik Stiemke tételéből Krein tétele.

Bizonyítás: Válasszunk A mátrixot úgy, hogy $R = \operatorname{Ker}_+ A$ legyen. Ha $a \in (R \cap K)^*$, akkor az

$$Ax \ge 0,\ x \in K,\ a^T x < 0$$

rendszer megoldhatatlan. Könnyen belátható, hogy ekkor

$$\operatorname{Ker}_+ \begin{pmatrix} A & 0 \\ a^T & 1 \\ -a^T & -1 \end{pmatrix} \cap \operatorname{ri} \begin{pmatrix} K \\ \mathcal{R}_+ \end{pmatrix} = \emptyset.$$

A vegyes Stiemke-tétel szerint

$$-\operatorname{Im}_+ \begin{pmatrix} A^T & a & -a \\ 0^T & 1 & -1 \end{pmatrix} \cap \begin{pmatrix} K^* \\ \mathcal{R}_+ \end{pmatrix} \not\subseteq \begin{pmatrix} -K^* \\ -\mathcal{R}_+ \end{pmatrix},$$

vagyis létezik y vektor és ζ szám úgy, hogy

$$-A^T y - a\zeta \in K^*,\ y \ge 0,\ -\zeta \ge 0;\text{ de } A^T y + a\zeta \notin K^*,\ \text{ vagy } \zeta \not\ge 0.$$

Nem lehet, hogy $\zeta = 0$, ekkor ugyanis $-A^T y \in K^* \setminus -K^*$, $y \geq 0$ lenne, holott az $R \cap \mathrm{ri}\, K \neq \emptyset$ feltétel miatt (ismét a vegyes Stiemke-tétel szerint) $(-\mathrm{Im}_+ A^T) \cap K^* \subseteq -K^*$. Ezért $\zeta < 0$, és akkor

$$-A^T \frac{y}{-\zeta} + a \in K^*, \ \frac{y}{-\zeta} \geq 0$$

miatt $a \in (\mathrm{Im}_+ A^T) + K^*$, vagyis $a \in R^* + K^*$. Ezzel a vegyes Krein-tételbeli nemtriviális $(R \cap K)^* \subseteq R^* + K^*$ tartalmazást újra beláttuk. □

6.3. Megjegyzés: A Stiemke-tétel alternatíváit átfogalmazva kapjuk alábbi bizonyítását. Először is vegyük észre, hogy $S \subseteq \mathcal{R}^d$ konvex kúp esetén ekvivalensek: a) S altér; b) $0 \in \mathrm{ri}\, S$; c) $S \subseteq -S$; d) S^* altér; e) $S^{**} = \mathrm{cl}\, S$ altér. A bizonyítás három lépése ezek után:
1. $L \cap \mathrm{ri}\, K \neq \emptyset$ pontosan akkor, ha $L + K$ altér (b) miatt).
2. $L^\perp \cap K^* \subseteq -K^*$ pontosan akkor, ha $L^\perp \cap K^*$ altér (c) miatt).
3. $L + K$ pontosan akkor altér, ha adjungáltja, $L^\perp \cap K^*$ altér (d) miatt).

A vegyes Stiemke-tétel is bizonyítható így: pl. $R \cap \mathrm{ri}\, K \neq \emptyset$ pontosan akkor, ha $R \cap (K - K) - K$ altér. A [vegyes] Stiemke-tétel egyenértékű 5.5-tel [5.7 $P = R$, $C = \mathrm{ri}\, K$ speciális esetével], így 5.3 [5.7] egy újabb bizonyítását is nyertük.

Megjegyezzük még, hogy $R \cap \mathrm{ri}\, K \neq \emptyset$ esetén $\mathrm{cl}\,(R \cap K) = R \cap \mathrm{cl}\, K$. Ezért a vegyes Krein-tétel átfogalmazható úgy, hogy $R \cap \mathrm{ri}\, K \neq \emptyset$ esetén $R^* + K^*$ zárt.

A diagonális alteres fogást felhasználva megfogalmazhatók a (vegyes) Stiemke- és Krein-tételek többkúpos általánosításai.

6.4. Tétel: (többkúpos Stiemke-tétel) *Legyenek $K_1, \ldots, K_l \subseteq \mathcal{R}^d$ konvex kúpok. Pontosan akkor teljesül $\cap_{j=1}^l \mathrm{ri}\, K_j \neq \emptyset$, ha*

$$b_j \in K_j^* \ (j = 1, \ldots, l), \ \sum_{j=1}^l b_j = 0 \ \textit{esetén}\ b_j \in -K_j^* \ (j = 1, \ldots, l).$$

Bizonyítás: Legyen $K := K_1 \times \ldots \times K_l$, és $D \subseteq \mathcal{R}^{dl}$ a diagonális altér. Nyilván $\cap_{j=1}^l \mathrm{ri}\, K_j \neq \emptyset$ pontosan akkor, ha $D \cap \mathrm{ri}\, K \neq \emptyset$. Másrészt Stiemke tétele szerint $D \cap \mathrm{ri}\, K \neq \emptyset$ pontosan akkor, ha $D^\perp \cap K^* \subseteq -K^*$, azaz

$$\cup \left\{ \{b_1\} \times \ldots \times \{b_l\} : \textstyle\sum_{j=1}^l b_j = 0 \right\} \cap$$
$$\cap (K_1^* \times \ldots \times K_l^*) \subseteq (-K_1^*) \times \ldots \times (-K_l^*),$$

ami éppen a tétel állítása. □

6.5. Tétel: (többkúpos Krein-tétel) *Legyenek* $K_1, \ldots, K_l \subseteq \mathcal{R}^d$ *konvex kúpok. Ha* $\cap_{j=1}^l \mathrm{ri}\, K_j \neq \emptyset$*, akkor* $(\cap_{j=1}^l K_j)^* = \sum_{j=1}^l K_j^*$.

Bizonyítás: Legyen K, D, mint az előző tétel bizonyításában. A feltételből adódóan ekkor $D \cap \mathrm{ri}\, K \neq \emptyset$, így Krein tétele szerint

$$(D \cap K)^* = D^\perp + K^*.$$

Szorozzuk meg ezt az egyenlőséget balról az $(E, \ldots, E) \in \mathcal{R}^{d \times (dl)}$ mátrixszal, máris kapjuk a kívánt $(\cap K_j)^* = \sum K_j^*$ egyenlőséget, mivel

$$(D \cap K)^* = \cup \left\{ \{a_1\} \times \ldots \times \{a_l\} : \sum_{j=1}^l a_j \in (\cap K_j)^* \right\}.$$

□

6.6. Tétel: (többkúpos vegyes Stiemke-tétel) *Legyenek* $R_1, \ldots, R_k \subseteq \mathcal{R}^d$ *poliéder kúpok, továbbá* $K_1, \ldots, K_l \subseteq \mathcal{R}^d$ *konvex kúpok. Ekkor az alábbi állítások ekvivalensek:*

a)

$$(\cap_{i=1}^k R_i) \cap (\cap_{j=1}^l \mathrm{ri}\, K_j) \neq \emptyset;$$

b)

$$\left.\begin{array}{l} a_i \in R_i^* \ (i = 1, \ldots, k), \\ b_j \in K_j^* \ (j = 1, \ldots, l), \\ \sum_{i=1}^k a_i + \sum_{j=1}^l b_j = 0 \end{array}\right\} \ \textit{esetén } b_j \in -K_j^* \ (j = 1, \ldots, l).$$

Bizonyítás: Ha a) fennáll, akkor legyen x_0 az a) szerint nemüres halmaz egy eleme, továbbá a_i, b_j a b) premisszáját teljesítő vektorok. Ekkor

$$a_i^T x_0 \geq 0 \ (i = 1, \ldots, k), \ b_j^T x_0 \geq 0 \ (j = 1, \ldots, l)$$

miatt a

$$\left(\sum_{i=1}^k a_i + \sum_{j=1}^l b_j \right)^T x_0 = 0$$

egyenlőségből

$$a_i^T x_0 = 0 \ (i = 1, \ldots, k), \ b_j^T x_0 = 0 \ (j = 1, \ldots, l)$$

adódik. De akkor $x_0 \in \mathrm{ri}\, K_j$ $(j = 1, \ldots, l)$ miatt, akárcsak 6.1 bizonyításában, könnyen meggyőződhetünk róla, hogy $b_j \in -K_j^*$ $(j = 1, \ldots, l)$.

A fordított irányhoz néhány jelölést vezetünk be. Legyen

$$R := \cap R_i,\ K := \cap K_j,\ \text{továbbá } K' := \cap \mathrm{ri}\, K_j.$$

Ha b) teljesül, akkor $a_i = 0$ $(i = 1, \ldots, k)$ választással alkalmazva, a többkúpos Stiemke-tétel szerint $K' \neq \emptyset$, és akkor $K' = \mathrm{ri}\, K$ (4.16). Mivel $R^* = \sum R_i^*$ (3.6), és a többkúpos Krein-tétel szerint $K^* = \sum K_j^*$, azért látjuk, hogy b) fennállása esetén $(-R^*) \cap (K^*) \subseteq -K^*$. A vegyes Stiemke-tétel szerint $R \cap \mathrm{ri}\, K \neq \emptyset$, vagyis $R \cap K' \neq \emptyset$, a)-hoz jutottunk. □

6.7. Tétel: (többkúpos vegyes Krein-tétel) *Legyenek $R_1, \ldots, R_k \subseteq \mathcal{R}^d$ poliéder kúpok, továbbá $K_1, \ldots, K_l \subseteq \mathcal{R}^d$ konvex kúpok. Ha*

$$(\cap_{i=1}^k R_i) \cap (\cap_{j=1}^l \mathrm{ri}\, K_j) \neq \emptyset,$$

akkor

$$\Big((\cap_{i=1}^k R_i) \cap (\cap_{j=1}^l K_j)\Big)^* = \sum_{i=1}^k R_i^* + \sum_{j=1}^l K_j^*.$$

Bizonyítás: Legyen R, K, K', mint az előző tétel bizonyításában. Most is $R \cap \mathrm{ri}\, K \neq \emptyset$, amiből a vegyes Krein-tétel szerint $(R \cap K)^* = R^* + K^*$ adódik. Most már elég annyit észrevenni, hogy $R^* = \sum R_i^*$, és a többkúpos Krein-tétel szerint $K^* = \sum K_j^*$. □

A többkúpos Stiemke-tétel két egyszerű következményét említjük. Itt $x <_K y$ jelölje azt, hogy $y - x \in \mathrm{ri}\, K$, ahol $K \subseteq \mathcal{R}^d$ konvex kúp, $x, y \in \mathcal{R}^d$ vektorok.

6.8. Tétel: (Carver) *Legyen $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, továbbá $K_1 \subseteq \mathcal{R}^m$, $K_2 \subseteq \mathcal{R}^n$ konvex kúpok. Ekkor az alábbi állítások ekvivalensek:*
a) létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $Ax >_{K_1} b$, $x >_{K_2} 0$;
b) valahányszor egy $y \in \mathcal{R}^m$ vektor esetén $A^T y \geq_{K_2^*} 0$, $y \leq_{K_1^*} 0$, $b^T y \leq 0$, akkor $A^T y \leq_{K_2^*} 0$, $y \geq_{K_1^*} 0$, $b^T y \geq 0$ is egyben.

Bizonyítás: Az a) állítás azt jelenti, hogy $b \in \mathrm{ri}\,(AK_2 - K_1)$, míg a b) állítás könnyen beláthatóan azt, hogy

$$y \in (AK_2 - K_1)^*,\ b^T y \leq 0 \text{ esetén } y \in -(AK_2 - K_1)^*,\ b^T y \geq 0.$$

Ezért elég azt igazolni, hogy tetszőleges K konvex kúp esetén

$$b \in \operatorname{ri} K \iff \Big\{y \in K^* : b^T y \le 0\Big\} \subseteq \Big\{y \in -K^* : b^T y \ge 0\Big\}.$$

Mivel $b \in \operatorname{ri} K$ pontosan akkor, ha $\operatorname{ri}(\operatorname{Im}_+ b) \cap \operatorname{ri} K \neq \emptyset$, azért ez az állítás a többkúpos Stiemke-tétel egyszerű következménye. □

6.9. Tétel: (első Dubovickij–Miljutyin-tétel) *Legyenek* $K, K_1, \ldots, K_l$ $\mathcal{R}^d$*-beli, konvex kúpok úgy, hogy* $\operatorname{int} K_j \neq \emptyset$ $(j = 1, \ldots, l)$. *Ekkor az alábbi állítások ekvivalensek:*
a)

$$K \cap (\cap_{j=1}^{l} \operatorname{int} K_j) \neq \emptyset;$$

b)

$$\left.\begin{array}{r} b \in K^*,\, b_j \in K_j^* \ (j = 1, \ldots, l), \\ b + \sum_{j=1}^{l} b_j = 0 \end{array}\right\} \textit{ esetén } b = 0 = b_1 = \ldots = b_l.$$

Bizonyítás: Mivel $\operatorname{int} K_j \neq \emptyset$ esetén $\operatorname{lin} K_j = \mathcal{R}^d$, azért

$$\{0\} = (\operatorname{lin} K_j)^{\perp} = (K_j - K_j)^{\perp} = (K_j - K_j)^* = K_j^* \cap -K_j^*.$$

Igaz továbbá, hogy

$$K \cap (\cap \operatorname{int} K_j) \neq \emptyset \iff (\operatorname{ri} K) \cap (\cap \operatorname{int} K_j) \neq \emptyset,$$

mivel egy konvex halmaz pontosan akkor metsz bele egy nyílt halmazba, ha a relatív belseje belemetsz. E két észrevétel után a tétel a többkúpos Stiemke-tétel egyszerű következménye. □

Új dimenziót ad a vizsgálódásnak a poláris fogalma.

Egy $S \subseteq \mathcal{R}^d$ halmaz **polárisa** legyen az

$$S^\circ := \Big\{a \in \mathcal{R}^d : a^T x \ge -1 \ (x \in S)\Big\}$$

halmaz. Könnyen belátható, hogy tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén S° az origót tartalmazó, zárt, konvex halmaz. Ha $S \subseteq \mathcal{R}^d$ az origót tartalmazó, zárt, konvex halmaz, akkor polárisát nevezzük még az S **duális halmaz**ának is (vö. 6.13).

Néhány megjegyzés a polárisképzéssel kapcsolatban:
– Egy konvex kúp polárisa a konvex kúp adjungáltja.
– Fontos észrevétel, hogy a poláris homogenizációval is nyerhető: ha C $\mathcal{R}^d$-beli, konvex halmaz, akkor

$$C^\circ = C(K(C)^*).$$

– A polárisképzés direkt szorzattal felcserélhetősége már csak korlátozottan igaz: ha $K \subseteq \mathcal{R}^d$ kúp, és $S \subseteq \mathcal{R}^d$ tetszőleges halmaz, $0 \in S$, akkor

$$(K \times S)^\circ = K^* \times S^\circ.$$

Emiatt a diagonális alteres fogás nem működik ebben az általánosságban.
– A polárisképzés tartalmazás-fordító: ha $S_1, S_2 \subseteq \mathcal{R}^d$ tetszőleges halmazok, amelyekre $S_1 \subseteq S_2$, akkor $S_1^\circ \supseteq S_2^\circ$. Sőt

6.10. Állítás: *Legyen $S_i \subseteq \mathcal{R}^d$ $(i \in I)$ halmazoknak egy tetszőleges rendszere. Ekkor*

$$(\cup_{i\in I} S_i)^\circ = \cap_{i\in I} S_i^\circ.$$

$\square$

6.11. Állítás: *Tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén*

$$S^\circ = (\operatorname{cl} S)^\circ, \text{ és } S^\circ = (\operatorname{conv} S)^\circ = (\operatorname{conv}(S \cup \{0\}))^\circ.$$

$\square$

A következő lemma alapvető fontosságú a polárisokkal kapcsolatos vizsgálatok során, feltétele szinte minden polárisokra vonatkozó tételben megjelenik.

6.12. Lemma: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz, és tegyük fel, hogy az origó C lezártjának eleme. Ekkor*

$$\operatorname{cl} K(C^\circ) = K(C)^*, \text{ és } K(C^\circ)^* = \operatorname{cl} K(C).$$

Bizonyítás: Mivel $e_1 \in \operatorname{cl} K(C) = K(C)^{**}$, azért $e_1^T K(C)^* \geq 0$, továbbá nyilván $0 \in C(K(C)^*)$. 3.20 szerint

$$\operatorname{cl} K(C^\circ) = \operatorname{cl} K(C(K(C)^*)) = K(C)^*.$$

A második igazolandó egyenlőség az elsőből adjungált kúpokat véve adódik. $\square$

Egy $S \subseteq \mathcal{R}^d$ halmaz **bipolárisa** az

$$S^{\circ\circ} := (S^\circ)^\circ$$

halmaz. A megfelelő dualitástétel a következő

6.13. Tétel: (bipoláris tétel) *Tetszőleges $S \subseteq \mathcal{R}^d$ halmaz esetén*

$$S^{\circ\circ} = \operatorname{cl}\operatorname{conv}(S \cup \{0\}).$$

Speciálisan, ha $C \subseteq \mathcal{R}^d$ konvex halmaz, amelyre $0 \in \operatorname{cl} C$ teljesül, akkor $C^{\circ\circ} = \operatorname{cl} C$. Speciálisan minden az origót tartalmazó, zárt, konvex halmaz bipolárisa önmaga.

Bizonyítás: A tétel részállításai valójában ekvivalensek, mivel 6.11 szerint feltehető, hogy $S = C$ az origót tartalmazó, zárt, konvex halmaz. Ekkor 6.12 szerint

$$C^{\circ\circ} = C(K(C^\circ)^*) = C(\operatorname{cl} K(C)) = \operatorname{cl} C = C.$$

□

Lássuk, hogyan általánosítódik 3.6 és 3.29.

6.14. Tétel: *Tetszőleges $K \subseteq \mathcal{R}^d$ kúp és $S \subseteq \mathcal{R}^d$ halmaz esetén*

$$(K + S)^\circ = K^* \cap S^\circ.$$

Teljesül továbbá, hogy

$$(K \cap S)^\circ \supseteq \operatorname{cl}(K^* + S^\circ),$$

egyenlőséggel, ha K zárt, konvex kúp, $S = C$ pedig az origót tartalmazó zárt, konvex halmaz.

Bizonyítás: A tétel nem nyilvánvaló része az, hogy ha K zárt, konvex kúp, $S = C$ pedig az origót tartalmazó, zárt, konvex halmaz, akkor $(K \cap C)^\circ \subseteq \operatorname{cl}(K^* + C^\circ)$. E feltételekkel pedig 3.28 és 6.13 szerint

$$(K \cap C)^\circ = (K^{**} \cap C^{\circ\circ})^\circ = (K^* + C^\circ)^{\circ\circ} = \operatorname{cl}(K^* + C^\circ).$$

□

6.15. Állítás: *Poliéder polárisa is poliéder.*

Bizonyítás: Motzkin tétele szerint egy P poliéder előáll $P = Q + R$ alakban, ahol $Q = \operatorname{conv} S_1$, $R = \operatorname{cone} S_2$, ahol S_1, S_2 véges halmazok. 6.14 miatt $P^\circ = Q^\circ \cap R^*$, ahol $Q^\circ = S_1^\circ$, $R^* = S_2^*$ poliéderek. Ezért P° poliéder. □

6.16. Állítás: *Ha $R \subseteq \mathcal{R}^d$ poliéder kúp, $P \subseteq \mathcal{R}^d$ az origót tartalmazó poliéder, akkor*

$$(R \cap P)^\circ = R^* + P^\circ.$$

Bizonyítás: Most $R^* + P^\circ$ poliéder, tehát zárt, így az állítás 6.14 következménye. □

Most a (vegyes) Stiemke- és Krein-tételek általánosításait bizonyítjuk. Az általános (vegyes) Krein-tétel további feltételt ad arra, hogy $(K \cap S)^\circ = K^* + S^\circ$ legyen.

6.17. Tétel: (általános Stiemke-tétel) *Legyen $K \subseteq \mathcal{R}^d$ konvex kúp, továbbá $C \subseteq \mathcal{R}^d$ konvex halmaz. Pontosan akkor teljesül $(\operatorname{ri} K) \cap (\operatorname{ri} C) \neq \emptyset$, ha*

$$(-K^*) \cap C^* \subseteq K^* \cap -C^*.$$

Bizonyítás: Felhasználva még a 4.15, 4.20 tételeket, a többkúpos Stiemke-tétel segítségével könnyen belátható, hogy

$$\begin{aligned}
&(\operatorname{ri} K) \cap (\operatorname{ri} C) \neq \emptyset \Longleftrightarrow (\operatorname{ri}(\mathcal{R} \times K)) \cap (\operatorname{ri} K(C)) \neq \emptyset \Longleftrightarrow \\
&\quad \Longleftrightarrow -(\mathcal{R} \times K)^* \cap K(C)^* \subseteq (\mathcal{R} \times K)^* \cap -K(C)^* \Longleftrightarrow \\
&\quad \Longleftrightarrow (\{0\} \times -K^*) \cap K(C)^* \subseteq (\{0\} \times K^*) \cap -K(C)^* \Longleftrightarrow \\
&\quad \Longleftrightarrow -K^* \cap C^* \subseteq K^* \cap -C^*.
\end{aligned}$$

□

6.18. Tétel: (általános vegyes Stiemke-tétel) *Legyen $R \subseteq \mathcal{R}^d$ poliéder kúp, $C \subseteq \mathcal{R}^d$ konvex halmaz. Pontosan akkor teljesül $R \cap \operatorname{ri} C \neq \emptyset$, ha*

$$(-R^*) \cap C^* \subseteq -C^*.$$

Bizonyítás: 6.17 bizonyításának mintájára, csak most a többkúpos Stiemke-tétel helyett a vegyes Stiemke-tételt használva igazolható. □

6.19. Tétel: (általános Krein-tétel) *Legyen $K \subseteq \mathcal{R}^d$ konvex kúp, $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha $(\mathrm{ri}\, K) \cap (\mathrm{ri}\, C) \neq \emptyset$, akkor $(K \cap C)^\circ = K^* + C^\circ$, és $(K \cap C)^* = K^* + C^*$.*

Bizonyítás: 6.17 bizonyításából tudjuk, hogy $(\mathrm{ri}\, K) \cap (\mathrm{ri}\, C) \neq \emptyset$ esetén

$$(\mathrm{ri}\, (\mathcal{R} \times K)) \cap (\mathrm{ri}\, K(C)) \neq \emptyset,$$

így a többkúpos Krein-tételből

$$((\mathcal{R} \times K) \cap K(C))^* = (\mathcal{R} \times K)^* + K(C)^*$$

adódik. Ez az egyenlőség úgy fogalmazható át, hogy

$$K(K \cap C)^* = (\{0\} \times K^*) + K(C)^*.$$

Az $\{1\} \times \mathcal{R}^d$ és a $\{0\} \times \mathcal{R}^d$ hipersíkokkal metszve mindkét oldalt nyerjük a bizonyítandó egyenlőségeket. □

6.20. Tétel: (általános vegyes Krein-tétel) *Legyen $R \subseteq \mathcal{R}^d$ poliéder kúp, $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha $R \cap \mathrm{ri}\, C \neq \emptyset$, akkor $(R \cap C)^\circ = R^* + C^\circ$, és $(R \cap C)^* = R^* + C^*$.*

Bizonyítás: 6.19 bizonyításának mintájára, csak most a többkúpos Krein-tétel helyett a vegyes Krein-tételt használva igazolható. □

A Stiemke-tételt az $L := D \subseteq \mathcal{R}^{(d+1)k}$ diagonális altérre és a $K := K(C_1) \times \ldots \times K(C_k)$ konvex kúpra alkalmazva ekvivalens feltételt kapunk arra, hogy a $C_1, \ldots, C_k \subseteq \mathcal{R}^d$ konvex halmazok relatív belsejei összemetsszenek. Hasonlóképpen adódik a Krein-tételből, hogy ha $\cap_{i=1}^k \mathrm{ri}\, C_i \neq \emptyset$, akkor

$$(\cap_{i=1}^k C_i)^\circ = C(\textstyle\sum_{i=1}^k K(C_i)^*).$$

Az általános (vegyes) Stiemke- és Krein-tételekben szerepel C adjungált kúpja. Erről a kúpról szól a következő állítás.

6.21. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ekkor*
a) $\mathrm{cone}\,(C^\circ) = \mathrm{bar}\, C$;
b) $\mathrm{rec}\,(C^\circ) = (\mathrm{cone}\, C)^*$.

Bizonyítás: Az a) és a b) állítás is a definíciók egyszerű következménye. □

Fontos következményként nyerjük az alábbi állítást.

6.22. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ekkor*
a) C° pontosan akkor korlátos, ha $0 \in \operatorname{int} C$;
b) C pontosan akkor korlátos, ha $0 \in \operatorname{int}(C^\circ)$.

Bizonyítás: a) C° nemüres, zárt, konvex halmaz lévén pontosan akkor korlátos, ha recessziós kúpja a triviális $\{0\}$ kúp (4.9). Továbbá 6.21 szerint $\operatorname{rec}(C^\circ) = (\operatorname{cone} C)^*$, és $(\operatorname{cone} C)^*$ pontosan akkor a triviális kúp, ha adjungáltja, $\operatorname{cl}\operatorname{cone} C$ az egész tér. Ez persze azzal ekvivalens, hogy $\operatorname{int}\operatorname{cone} C = \mathcal{R}^d$, ami pontosan akkor áll fenn, ha $0 \in \operatorname{int} C$.

b) a)-ból adódik, azt C helyett C°-re alkalmazva. □

Most már rátérhetünk a fejezet elején ígért zártsági tételekre. Az összefoglaló zártsági tételek előtt a 4.23 és 4.25 Tételek duálisait bizonyítjuk, az alábbi lemma segítségével.

6.23. Lemma: *Legyen $A \in \mathcal{R}^{m\times n}$, $C \subseteq \mathcal{R}^m$ nemüres, zárt, konvex halmaz, és tegyük fel, hogy $A^{-1}(C) \neq \emptyset$. Ekkor*

$$A^T \operatorname{ri}\operatorname{bar} C = \operatorname{ri}\operatorname{bar} A^{-1}(C).$$

A ri *operáció elhagyható a képletből, ha $C = P$ poliéder.*

Bizonyítás: Az állítás első fele könnyen igazolható, felhasználva, hogy tetszőleges K konvex kúp esetén

$$\operatorname{cl}(A^T K) = (A^{-1}(K^*))^*,$$

valamint a 4.7, 4.11, 4.21 és 4.28 Tételeket.

Az állítás poliéderekre vonatkozó része a 4.13 és 4.28 Tételek egyszerű következménye. □

6.24. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, $C \subseteq \mathcal{R}^n$ nemüres, zárt, konvex halmaz, $P \subseteq \mathcal{R}^m$ poliéder. Tegyük fel, hogy*

$$(A^T \operatorname{bar} P) \cap \operatorname{ri}(\operatorname{bar} C) \neq \emptyset.$$

Ekkor $A^{-1}(P) + C$ zárt halmaz, és az $A^{-1}(\operatorname{rec} P) + \operatorname{rec} C$ kúp a recessziós kúpja.

Bizonyítás: A lemma szerint feltehető, hogy $A = E$. Az is feltehető, hogy $0 \in C$. (Ha nem így lenne, vonjuk ki a C halmazból egy elemét, ezáltal a tétel feltétele és konklúziója sem változik.)

Először vizsgáljuk azt a speciális esetet, mikor $P = R$ kúp. Az általános vegyes Krein-tétel szerint

$$R^* \cap \operatorname{ri} C^\circ \neq \emptyset \Rightarrow R + C \text{ zárt}, \ (R^* \cap C^\circ)^* = R + C^{\circ *},$$

felhasználva, hogy $R^{**} = R$, és $C^{\circ\circ} = C$. Mivel R^* kúp, a feltételben a $\operatorname{ri} C^\circ$ halmaz lecserélhető a $\operatorname{ri} \operatorname{cone} C^\circ$ halmazra (vö. 4.18), ami 6.21 a) miatt éppen a $\operatorname{ri} \operatorname{bar} C$ halmaz. Továbbá 6.14, 6.21 és 4.11 miatt

$$(R^* \cap C^\circ)^* = (R + C)^{\circ *} = (\operatorname{bar}(R + C))^* = \operatorname{rec}(R + C);$$

és hasonlóképpen $R + C^{\circ *} = R + \operatorname{rec} C$. Látjuk, hogy éppen a tételbeli állításhoz jutottunk a $P = R$ speciális esetben.

Általában az

$$\hat{A} := E, \ \hat{C} := C, \ \hat{P} := \operatorname{rec} P$$

választással alkalmazzuk a tétel az előbb elintézett speciális esetét. A feltételek nyilvánvalóan ekvivalensek, hiszen

$$\hat{A}^T \operatorname{bar} \hat{P} = \operatorname{bar} \operatorname{rec} P = (\operatorname{rec} P)^* = \operatorname{bar} P$$

(vö. 4.13). Tehát a feltételek fennállása esetén $(\operatorname{rec} P) + C$, és így a belőle egy politóp hozzáadásával nyert $P + C$ halmaz is zárt (lásd még Motzkin tételét, és a 4.12 előtti megjegyzést vagy 3.35-t). A recessziós kúpról szóló egyenlőség hasonlóképpen adódik 4.12 segítségével. $\square$

6.25. Következmény: *Legyen $A \in \mathcal{R}^{m \times n}$, $C_1 \subseteq \mathcal{R}^n$, $C_2 \subseteq \mathcal{R}^m$ nemüres, zárt, konvex halmazok. Tegyük fel, hogy*

$$(A^T \operatorname{ri} \operatorname{bar} C_2) \cap \operatorname{ri}(\operatorname{bar} C_1) \neq \emptyset.$$

Ekkor az $A^{-1}(C_2) + C_1$ halmaz zárt, és az $A^{-1}(\operatorname{rec} C_2) + \operatorname{rec} C_1$ kúp a recessziós kúpja.

Bizonyítás: 6.23 szerint feltehető, hogy $A = E$. Alkalmazzuk 6.24-et

$$\hat{A} := (E, E), \ \hat{C} := C_1 \times C_2, \ \hat{P} := \{0\}$$

választással. A feltételek és a konklúziók is ekvivalensek (vö. Abrams tétele; a recessziós egyenlőségek egymás $\hat{A}$-, illetve $\hat{A}^{-1}$-szeresei). $\square$

6.26. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, $C_1 \subseteq \mathcal{R}^n$, $C_2 \subseteq \mathcal{R}^m$ nemüres, zárt, konvex halmazok, $K_i := \operatorname{rec} C_i$ $(i = 1, 2)$. Ekkor*

$$\begin{array}{ccc}
a)\ (A^T \operatorname{ri} \operatorname{bar} C_2) \cap \operatorname{ri}(\operatorname{bar} C_1) \neq \emptyset & \Rightarrow & b)\ A^{-1}(C_2) + C_1 \text{ zárt} \\
\Updownarrow & & \Uparrow \\
c)\ A^{-1}(-K_2) \cap K_1 \subseteq A^{-1}(K_2) \cap -K_1 & \Rightarrow & d)\ AC_1 + C_2 \text{ zárt.}
\end{array}$$

Ráadásul b)$\Rightarrow$d) is teljesül, ha $C_2 \subseteq \operatorname{Im} A$. (Vagy ha b)-t az alábbi állításra cseréljük:

$$b')\ \operatorname{Ker}(A, E) + C_1 \times C_2 \text{ zárt.})$$

Speciálisan (és ehhez az állításhoz kétféle úton is eljuthatunk): a)$\Rightarrow$d) Ha létezik $y \in \operatorname{ri} \operatorname{bar} C_2$ úgy, hogy $A^T y \in \operatorname{ri} \operatorname{bar} C_1$, akkor $AC_1 + C_2$ zárt.

Bizonyítás: a)$\Rightarrow$b) 6.25 szerint. a)$\Longleftrightarrow$c) a többkúpos Stiemke-tételből következik (lásd még 4.11). A b) és d) állítások kapcsolatáról szóló részállítás Abrams tételének következménye, végül c)$\Rightarrow$d) a 4.25 Tételé. $\square$

6.27. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, $C \subseteq \mathcal{R}^n$ nemüres, zárt, konvex halmaz, $P \subseteq \mathcal{R}^m$ poliéder. Ekkor*

$$\begin{array}{ccc}
a)\ (A^T \operatorname{bar} P) \cap \operatorname{ri}(\operatorname{bar} C) \neq \emptyset & \Rightarrow & b)\ A^{-1}(P) + C \text{ zárt} \\
\Updownarrow & & \Uparrow \\
c)\ A^{-1}(-\operatorname{rec} P) \cap \operatorname{rec} C \subseteq -\operatorname{rec} C & \Rightarrow & d)\ AC + P \text{ zárt.}
\end{array}$$

Ráadásul b)$\Rightarrow$d) is teljesül, ha $P \subseteq \operatorname{Im} A$. (Vagy ha b)-t az alábbi állításra cseréljük:

$$b')\ \operatorname{Ker}(A, E) + C \times P \text{ zárt.})$$

Speciálisan (és ehhez az állításhoz kétféle úton is eljuthatunk): a)$\Rightarrow$d) Ha létezik $y \in \operatorname{bar} P$ úgy, hogy $A^T y \in \operatorname{ri} \operatorname{bar} C$, akkor $AC + P$ zárt.

Bizonyítás: 6.26 bizonyításához hasonlóan igazolható, csak most 6.24-et, 4.23-at és a vegyes Stiemke-tételt használva. Az a)$\Rightarrow$b') részállítás bizonyításához segítségképpen még megjegyezzük, hogy

$$\operatorname{Ker}(A, E) + C \times P = (\operatorname{Ker}(A, E) + \{0\} \times P) + C \times \{0\}.$$

$\square$

Adott, az $\mathcal{R}^d$ tér valamely részhalmazán értelmezett, valós értékkészletű f függvény, $S \subseteq f^{-1}(\mathcal{R})$ halmaz esetén **minimalizálási feladat**nak nevezzük azt a feladatot, mikor az S halmazban (a feladat **megengedett megoldás**ai között) keresünk egy olyan $\hat{x}$ vektort (a feladat **optimális megoldás**át), ahol az f függvény (a feladat **célfüggvény**e) a lehető legkisebb, vagyis $\inf\{f(x) : x \in S\}$ értéket (a feladat **optimumérték**ét) veszi fel, vagy ha ilyen $\hat{x}$ vektor nem létezik, akkor olyan $\hat{x}_k$ $(k = 1, 2, \ldots)$ sorozatát megengedett megoldásoknak (**optimalizáló sorozat**), hogy az $f(\hat{x}_k)$ sorozat a feladat optimumértékéhez tart, ahogy $k \to \infty$. Ezt a feladatot röviden így jelöljük:

$$\inf f(x),\ x \in S.$$

Ha a megengedett megoldások halmaza valamiféle, az x változóra tett $F(x)$ feltételekkel van leírva, akkor $\inf f(x)$, $x \in \{x : F(x)\}$ helyett röviden azt írjuk, hogy

$$\inf f(x),\ F(x)$$

Hasonlóan definiálható a

$$\sup f(x),\ x \in S,$$

illetve

$$\sup f(x),\ F(x)$$

maximalizálási feladat.

A minimalizálási és maximalizálási feladatokat közösen **program**nak nevezzük. Általában egy programot zárójelbe tett nyomtatott, nagy betűvel jelölünk (például (P)), ekkor a program megengedett megoldásainak halmazát a fenti betű vastagon szedett változatával jelöljük (például $\mathbf{P}$), optimális megoldásainak halmazát pedig úgy, hogy a vastagon szedett betű mellé alsó indexbe egy "o" betűt írunk (például $\mathbf{P_o}$). Az optimumérték jele például a (P) program esetében v_P, ahol a v betű az angol "value" szó rövidítése.

Azt mondjuk, hogy egy program **megoldható**, ha létezik megengedett megoldása. A megoldható programot **korlátos**nak nevezzük, ha optimumértéke véges, különben pedig **korlátlan**nak. Ha egy programnak létezik optimális megoldása, akkor azt mondjuk, hogy a program **optimumértéke felvétetik**.

Két minimalizálási feladatot **megegyező**nek tekintünk, ha ugyanaz a megengedett megoldásaik halmaza és a célfüggvényük megszorítása erre a halmazra. Két minimalizálási feladatot **ekvivalens**nek nevezünk, ha optimumértékeik megegyeznek és egyszerre vétetnek fel, vagy egyszerre nem vétetnek fel. Hasonlóan maximalizálási feladatra. Például két minimalizálási feladat ((P_1) és (P_2), f_1, illetve f_2 célfüggvényekkel) pontosan akkor ekvivalens, ha léteznek

$$\varphi : \mathbf{P_1} \to \mathbf{P_2} \text{ és } \psi : \mathbf{P_2} \to \mathbf{P_1}$$

leképezések (ún. **változótranszformációk**) úgy, hogy

$$f_2(\varphi(x)) \leq f_1(x), \text{ és } f_1(\psi(y)) \leq f_2(y)$$

valahányszor $x \in \mathbf{P_1}$, $y \in \mathbf{P_2}$.

Egy $\inf f(x)$, $x \in S$ minimalizálási feladat -1-szerese a

$$-\inf f(x),\ x \in S := \sup -f(x),\ x \in S$$

maximalizálási feladat. Hasonlóan definiálható egy maximalizálási feladat -1-szerese az inf és a sup felcserélésével.

Az

$$\inf c^T x,\ Ax \geq_{K_1} b, x \geq_{K_2} 0$$

alakú minimalizálási feladatot, ahol $A \in \mathcal{R}^{m\times n}$ mátrix, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$ vektorok, és $K_1 \subseteq \mathcal{R}^m$, $K_2 \subseteq \mathcal{R}^n$ konvex kúpok, **kúplineáris minimalizálási feladat**nak nevezzük. Ekkor azt mondjuk, hogy az

$$(A, b, c, K_1, K_2, \inf)$$

hatos a feladat egy **leírás**a. Hasonlóan definiálható a **kúplineáris maximalizálási feladat** és **leírás**ai.

A kúplineáris minimalizálási és maximalizálási feladatokat együtt **kúplineáris program**oknak nevezzük.

Egy kúplineáris program adott $(A, b, c, K_1, K_2, \inf \text{ vagy } \sup)$ leírásához tartozó **szigorúan megengedett megoldás**ai mindazon $x \in \mathcal{R}^n$ vektorok, amelyekre

$$Ax >_{K_1} b,\ x >_{K_2} 0.$$

Az adott módon leírt kúplineáris programot **szigorúan megoldható**nak nevezzük, ha létezik a leíráshoz tartozó szigorúan megengedett megoldása.

Tekintsük például az

$$\inf c^T x,\ Ax \geq_{K_1} b, x \geq_{K_2} 0$$

kúplineáris programot, melyre a rövidség kedvéért (P)-ként fogunk hivatkozni, megengedett megoldásainak halmazát $\mathbf{P}$-vel, $(A, b, c, K_1, K_2, \inf)$ leírásához tartozó szigorúan megengedett megoldásainak halmazát $\mathbf{P_s}$-sel, optimális megoldásainak halmazát $\mathbf{P_o}$-val, optimumértékét pedig v_P-vel jelöljük.

Vegyük észre, hogy ha K_1, K_2 zárt, konvex kúpok (ezt a továbbiakban feltesszük, bár nem mindig lesz rá szükség), akkor

$$\mathbf{P} = K_2 \cap A^{-1}(b + K_1)$$

zárt, konvex halmaz, és

$$\mathbf{P_s} = (\mathrm{ri}\, K_2) \cap A^{-1}(b + \mathrm{ri}\, K_1) = \mathrm{ri}\, \mathbf{P},$$

amennyiben (P) szigorúan megoldható.

Az általános kúplineáris Farkas-lemma az $AK_2 - K_1$ zártsága esetén szükséges és elégséges feltételét adja e program megoldhatóságának. 6.26-ot is felhasználva (bár általánosságáról lemondunk) kimondható, hogy

6.28. Tétel: *Tegyük fel, hogy $AK_2 - K_1$ zárt. Ekkor az alábbi állítások ekvivalensek:*
a) (P) megoldható, vagyis létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $Ax \geq_{K_1} b$, $x \geq_{K_2} 0$;
b) valahányszor egy $y \in \mathcal{R}^m$ vektor esetén $A^T y \geq_{K_2^*} 0$, $y \leq_{K_1^*} 0$, akkor $b^T y \geq 0$ is egyben.

Speciálisan, ha létezik $y_0 \in \mathcal{R}^m$ vektor, amelyre $A^T y_0 >_{K_2^*} 0$, $y_0 <_{K_1^*} 0$, akkor a) és b) ekvivalensek. (Az utóbbi feltétel azt jelenti, hogy az origó az $A^T K_1^* + K_2^*$ konvex kúp relatív belső pontja, így ez a kúp valójában altér.)$\square$

Most megvizsgáljuk, hogy egy megoldható kúplineáris program pontosan mikor korlátos.

6.29. Tétel: *Tegyük fel, hogy (P) megoldható, továbbá teljesül, hogy $A^T K_1^* + K_2^*$ zárt. Ekkor az alábbi állítások ekvivalensek:*
a) (P) korlátos;
b) nem létezik $z \in \mathcal{R}^n$ vektor, amelyre $Az \geq_{K_1} 0, z \geq_{K_2} 0, c^T z < 0$;
c) létezik $y \in \mathcal{R}^m$ vektor, amelyre $A^T y \leq_{K_2^*} c, y \geq_{K_1^*} 0$.

Speciálisan, ha létezik $z_0 \in \mathcal{R}^n$ vektor, amelyre $Az_0 >_{K_1} 0$, $z_0 >_{K_2} 0$, akkor a), b) és c) ekvivalensek. (Az utóbbi feltétel azt jelenti, hogy $0 \in$ $\mathrm{ri}\,(AK_2 - K_1)$, vagyis az $AK_2 - K_1$ konvex kúp valójában altér.)

Bizonyítás: Először megmutatjuk, hogy a)⇒b). Legyen x_0 a (P) program megengedett megoldása. Ha mégis létezne a b) állításban leírt tulajdonságokkal rendelkező z vektor, akkor $x_0 + z\mathcal{R}_+ \subseteq \mathbf{P}$ lenne, és e félegyenes mentén $-\infty$-hez tartana a célfüggvény, ami ellentmond (P) korlátosságának.

Másodszor belátjuk, hogy c)⇒a). Ha y a c) állításban leírt tulajdonságokkal rendelkező vektor, akkor tetszőleges $x \in \mathbf{P}$ esetén

$$0 \le y^T(Ax - b) + x^T(c - A^T y) = c^T x - b^T y,$$

tehát a $b^T y$ érték a (P) program optimumértékének alsó korlátja.

Végül a b) és c) állítások ekvivalenciája, és az állítás fennmaradó része 6.28 következménye. □

A tétel bizonyításában burkoltan feltűnt egy (a (P) leírásától függő) újabb kúplineáris program, amelynek optimumértéke alsó korlátja (P) optimumértékének, és amely bizonyos feltétellel megoldható, ha (P) megoldható és korlátos. Ez a

$$\sup b^T y,\ A^T y \le_{K_2^*} c, y \ge_{K_1^*} 0$$

kúplineáris program, melyet a (P) program **duál**jának nevezünk. A (P) programot ekkor a duál programtól való megkülönböztetés céljából **primál program**nak nevezzük. Ha a primál program leírása $(A, b, c, K_1, K_2, \inf)$, akkor a duál program leírása lesz

$$(-A^T, -c, b, K_2^*, K_1^*, \sup).$$

A rövidség kedvéért a duál kúplineáris programra (D)-ként fogunk hivatkozni, megengedett megoldásainak halmazát $\mathbf{D}$-vel, (fenti leírásához tartozó) szigorúan megengedett megoldásainak halmazát $\mathbf{D_s}$-sel, optimális megoldásainak halmazát $\mathbf{D_o}$-val, optimumértékét pedig v_D-vel jelöljük.

Vegyük észre, hogy

$$-(D) = \inf -b^T y,\ -A^T y \ge_{K_2^*} -c, y \ge_{K_1^*} 0$$

duálja

$$\sup -c^T x,\ -Ax \le_{K_1} -b, x \ge_{K_2} 0 = -(P).$$

(Itt kihasználtuk, hogy K_1, K_2 zárt, konvex kúpok, és hogy így $K_1^{**} = K_1$, $K_2^{**} = K_2$.) Ennek a fontos észrevételnek két következményét említjük. Egyrészt a (D) duáljának tekinthető a (P) program, így kúplineáris maximalizálási feladat duálisát is definiáltuk, ráadásul úgy, hogy bármely kúplineáris program biduálisa (duálisának duálisa) önmaga. (Egészen precízen a leírások halmazán értelmezett duálisképzésről van szó.) Másrészt (D) megoldhatósága, illetve megoldhatósága esetén korlátossága is karakterizálható bizonyos zártsági feltételek mellett. Az erről szóló tételek megfelelőik, 6.28 és 6.29 azonnali következményei.

6.30. Tétel: *Tegyük fel, hogy $A^T K_1^* + K_2^*$ zárt. Ekkor az alábbi állítások ekvivalensek:*
a) (D) megoldható, vagyis létezik $y \in \mathcal{R}^m$ vektor úgy, hogy $A^T y \leq_{K_2^*} c$, $y \geq_{K_1^*} 0$;
b) valahányszor egy $x \in \mathcal{R}^n$ vektor esetén $Ax \geq_{K_1} 0$, $x \geq_{K_2} 0$, akkor $c^T x \geq 0$ is egyben.

Speciálisan, ha létezik $x_0 \in \mathcal{R}^n$ vektor, amelyre $Ax_0 >_{K_1} 0$, $x_0 >_{K_2} 0$, akkor a) és b) ekvivalensek. (Az utóbbi feltétel azt jelenti, hogy az origó az $AK_2 - K_1$ konvex kúp relatív belső pontja, így ez a kúp valójában altér.) □

6.31. Tétel: *Tegyük fel, hogy (D) megoldható, továbbá teljesül, hogy $AK_2 - K_1$ zárt. Ekkor az alábbi állítások ekvivalensek:*
a) (D) korlátos;
b) nem létezik $w \in \mathcal{R}^m$ vektor, amelyre $A^T w \leq_{K_2^*} 0, w \geq_{K_1^*} 0, b^T w > 0$;
c) létezik $x \in \mathcal{R}^n$ vektor, amelyre $Ax \geq_{K_1} b, x \geq_{K_2} 0$.

Speciálisan, ha létezik $w_0 \in \mathcal{R}^m$ vektor, amelyre $A^T w_0 <_{K_2^*} 0$, $w_0 >_{K_1^*} 0$, akkor a), b) és c) ekvivalensek. (Az utóbbi feltétel azt jelenti, hogy az origó az $A^T K_1^* + K_2^*$ konvex kúp relatív belsejében van, így ez a kúp valójában altér.) □

Ezzel a primál és duál program közötti kapcsolatokról szóló első tételeket lényegében már be is láttuk.

6.32. Tétel: (gyenge dualitási tétel) *A (P) program optimumértéke legalább akkora, mint a (D) program optimumértéke.*

Bizonyítás: Ez persze így van, ha (P) vagy (D) nem megoldható (ekkor ugyanis definíció szerint $v_P = \infty$, illetve $v_D = -\infty$). Különben (mint azt 6.29 bizonyításakor láttuk) tetszőleges $x \in \mathbf{P}$, $y \in \mathbf{D}$ esetén $c^T x \geq b^T y$. Ebből az észrevételből az állítás megoldható programok esetében is következik. □

6.29 és 6.31 következményeként adódik az alábbi

6.33. Tétel:
a) Tegyük fel, hogy $A^T K_1^* + K_2^*$ *zárt. Ha* (P) *megoldható és korlátos, akkor* (D) *megoldható.*
b) Tegyük fel, hogy $AK_2 - K_1$ *zárt. Ha* (D) *megoldható és korlátos, akkor* (P) *megoldható.*
c) Tegyük fel, hogy létezik $z_0 \in \mathcal{R}^n$ *vektor amelyre* $Az_0 >_{K_1} 0$, $z_0 >_{K_2} 0$. *Ha* (P) *megoldható és korlátos, akkor* (D) *megoldható, és* (P) *szigorúan megoldható.*
d) Tegyük fel, hogy létezik $w_0 \in \mathcal{R}^m$ *vektor amelyre* $A^T w_0 <_{K_2^*} 0$, $w_0 >_{K_1^*} 0$. *Ha* (D) *megoldható és korlátos, akkor* (P) *megoldható, és* (D) *szigorúan megoldható.*

Bizonyítás: Például c)-hez elég annyit észrevenni, hogy ha $x \in \mathbf{P}$, akkor $x + z_0 \in \mathbf{P_s}$, mivel $K + \operatorname{ri} K = \operatorname{ri} K$ tetszőleges K konvex kúp esetén. □

6.34. Tétel: (erős dualitási tétel)
a) Tegyük fel, hogy az

$$\begin{pmatrix} A \\ -c^T \end{pmatrix} K_2 - \begin{pmatrix} K_1 \\ \mathcal{R}_+ \end{pmatrix}$$

konvex kúp zárt. (E feltétel teljesül például, ha K_1 *és* K_2 *is poliéder kúp.) Ha* (P) *vagy* (D) *megoldható, akkor* $v_P = v_D$, *és* (P) *optimumértéke felvétetik, ha véges.*
b) Tegyük fel, hogy az

$$\begin{pmatrix} A^T \\ -b^T \end{pmatrix} K_1^* + \begin{pmatrix} K_2^* \\ \mathcal{R}_+ \end{pmatrix}$$

konvex kúp zárt. (E feltétel teljesül például, ha K_1 *és* K_2 *is poliéder kúp.) Ha* (P) *vagy* (D) *megoldható, akkor* $v_D = v_P$, *és* (D) *optimumértéke felvétetik, ha véges.*
c) Ha (D) *szigorúan megoldható, akkor* $v_P = v_D$. *Ha* (D) *még korlátos is, akkor* $\mathbf{P_o} \neq \emptyset$.
d) Ha (P) *szigorúan megoldható, akkor* $v_D = v_P$. *Ha* (P) *még korlátos is, akkor* $\mathbf{D_o} \neq \emptyset$.

Bizonyítás: a) 3.33 azonnali következménye. Ha $v_P > v_D$ lenne, akkor egy $v_P > \delta > v_D$ számot választva 3.33 b) teljesülne, így a zártsági és megoldhatósági feltétel miatt létezne $x \in \mathbf{P}$ úgy, hogy $c^T x \leq \delta < v_P$, ami lehetetlen. Ezért $v_P = v_D$. Ha ez a közös érték véges, akkor jelölje δ, ismét teljesül 3.33 b), így 3.33 a) szerint $\mathbf{P_o} \neq \emptyset$.

b) a) következménye, hiszen a duál program (-1-szeresének) duálja a primál program (-1-szerese).

c) 6.26 szerint az a) pont zártsági feltétele teljesül, ha

$$-\begin{pmatrix} A \\ -c^T \end{pmatrix}^T \operatorname{ri}(K_1 \times \mathcal{R}_+)^* \cap \operatorname{ri} K_2^* \neq \emptyset,$$

vagyis létezik y vektor és μ szám úgy, hogy

$$-A^T y + c\mu >_{K_2^*} 0,\ y >_{K_1^*} 0,\ \mu > 0.$$

Legyen $y \in \mathbf{D_s}$, $\mu := 1$, ekkor y, μ megfelel.

c) Carver tétele segítségével is igazolható (vö. 6.2 második bizonyításával). Nyilván nem létezik $y \in \mathbf{D_s}$, amelyre $b^T y > v_D$, vagyis a

$$\begin{pmatrix} -A^T \\ b^T \end{pmatrix} y - \begin{pmatrix} -c \\ v_D \end{pmatrix} >_{K_2^* \times \mathcal{R}_+} 0,\ y >_{K_1^*} 0$$

rendszer nem megoldható. Carver tétele szerint létezik x vektor és λ szám, amelyre

$$-Ax + b\lambda \geq_{K_1} 0,\ x \leq_{K_2} 0,\ \lambda \leq 0,\ -c^T x + v_D \lambda \leq 0,$$

de nem mindegyik egyenlőtlenségjel fordítható meg. Nem lehet $\lambda = 0$, különben Carver tétele szerint ellentmondásba keverednénk azzal, hogy (D) szigorúan megoldható. Ezért $\lambda < 0$, de akkor $x/\lambda \in \mathbf{P}$, és $c^T(x/\lambda) \leq v_D$, vagyis $x/\lambda \in \mathbf{P_o}$.

d) c) következménye, ugyanis a duál program (-1-szeresének) duálja a primál program (-1-szerese). □

Vegyük észre, hogy az erős dualitási tétel már inhomogén tétel, bizonyításához az előtte használt homogén segédtételek (a Stiemke-tétel és a Farkas-lemma) inhomogén változatait (a Carver-tételt és a Farkas-tételt) használtuk.

A gyenge és erős dualitási tételek egyszerű következménye az alábbi

6.35. Tétel: (kiegészítő eltérések tétele) *Tegyük fel, hogy az erős dualitási tétel a) vagy b) részének zártsági feltétele teljesül (ez így van, ha (P) vagy (D) szigorúan megoldható). Ekkor tetszőleges $x \in \mathbf{P}$, $y \in \mathbf{D}$ megengedett megoldások esetén az alábbi állítások ekvivalensek:*
a) $x \in \mathbf{P_o}, y \in \mathbf{D_o}$;
b) $c^T x = b^T y$;
c) $x^T(c - A^T y) = 0 = y^T(Ax - b)$. □

A következő tétel szintén a (P) program szigorú megoldhatóságát karakterizáló Carver-tétel segítségével igazolható.

6.36. Tétel: *Tegyük fel, hogy* $\mathbf{D} \neq \emptyset$, *és legyen*

$$\mathbf{D}_\alpha := \left\{ y \in \mathbf{D} : b^T y \geq \alpha \right\} \quad (\alpha \in \mathcal{R}).$$

(P) pontosan akkor szigorúan megoldható, ha tetszőleges $\alpha \in \mathcal{R}$ esetén $\operatorname{rec} \mathbf{D}_\alpha = \operatorname{line} \mathbf{D}_\alpha$, *vagyis a $\mathbf{D}_\alpha$ halmazok egyeneseiket leszámítva korlátosak.*

Bizonyítás: Könnyen belátható, hogy egy nemüres $\mathbf{D}_\alpha$ halmaz recessziós kúpja

$$\operatorname{rec} \mathbf{D}_\alpha = \left\{ y : b^T y \geq 0, -A^T y \in K_2^*, y \in K_1^* \right\},$$

lineralitás tere pedig

$$\operatorname{line} \mathbf{D}_\alpha = \left\{ y : b^T y = 0, -A^T y \in \operatorname{line} K_2^*, y \in \operatorname{line} K_1^* \right\}.$$

Carver tétele szerint (P) pontosan akkor szigorúan megoldható, ha $\mathbf{D}_\alpha$ recessziós kúpja a $\mathbf{D}_\alpha$ linealitás terének része (és akkor egyenlőség is fennáll), ami éppen a kívánt állítás. □

6.37. Következmény: *Pontosan akkor teljesül, hogy $\mathbf{D} \neq \emptyset \neq \mathbf{P_s}$, ha $\mathbf{D_o} \neq \emptyset$, és* $\operatorname{rec} \mathbf{D_o} = \operatorname{line} \mathbf{D_o}$.

Bizonyítás: 6.34 és 6.36 következménye. □

Természetesen 6.36 és 6.37 megfelelő változtatásokkal fennáll akkor is, ha (P) és (D) szerepét felcseréljük.

6.37-ből adódik, hogy ha az egyik program megengedett megoldásainak halmaza nemüres, korlátos, akkor a másik program szigorúan megoldható. Viszont kevés primál-duál programpár van, amely emiatt szigorúan megoldható:

6.38. Állítás: *Ha* **P**, **D** *nemüres, korlátos halmazok, akkor* K_1, K_2 *altér.*

Bizonyítás: A feltétel szerint **P** nemüres, kompakt, konvex halmaz, így recessziós kúpja csak a 0 vektorból áll, vagyis

$$Az \in K_1, z \in K_2 \text{ esetén } z = 0.$$

Carver tétele szerint ekkor létezik w_0 vektor úgy, hogy

$$-A^T w_0 \in \operatorname{ri} K_2^*, w_0 \in \operatorname{ri} K_1^*.$$

Mivel a **D** konvex halmaz kompaktsága miatt

$$\operatorname{rec} \mathbf{D} = \left\{ w \in K_1^* : -A^T w \in K_2^* \right\} = \{0\},$$

azért w_0 csak a 0 vektor lehet, de akkor $0 \in \operatorname{ri} K_2^*, 0 \in \operatorname{ri} K_1^*$ miatt K_1^*, K_2^*, és így K_1, K_2 is altér. □

7. Extremális és exponált részhalmazok

Könnyen belátható, hogy lineáris függvényt optimalizálni egy halmaz és annak konvex burka felett ekvivalens programokat jelent. Ha tehát $C \subseteq \mathcal{R}^d$ konvex halmaz, és $C = \operatorname{conv} S$ valamely $S \subseteq \mathcal{R}^d$ halmaz esetén, akkor az $\inf c^T x$, $x \in C$ és $\inf c^T x$, $x \in S$ minimalizálási feladatok ekvivalensek ($c \in \mathcal{R}^d$), akárcsak a megfelelő maximalizálási feladatok. Ezért keressük egy konvex halmaz lehetőleg minél szűkebb konvex generáló részhalmazát (vagyis egy olyan halmazt, amelynek konvex burka az eredeti konvex halmaz), amelyen esetleg könnyebb optimalizálni. Így jutunk az extremális pont fogalmához.

Adott $C \subseteq \mathcal{R}^d$ konvex halmaz esetén az $x \in C$ pont a C konvex halmaz **extremális pont**ja, ha nincs benne C-beli, nem elfajuló szakasz belsejében, vagyis

$$x_1, x_2 \in C,\ \lambda_1, \lambda_2 > 0,\ \lambda_1 + \lambda_2 = 1,\ x = \lambda_1 x_1 + \lambda_2 x_2 \text{ esetén } x_1 = x_2 = x.$$

(Vagy még másképpen $C \setminus \{x\}$ konvex halmaz.) A C konvex halmaz extremális pontjainak halmazát $\operatorname{ext} C$-vel jelöljük.

A pontok elemszáma szerinti indukcióval könnyen belátható, hogy ha az x pont a C konvex halmaz extremális pontja, akkor a definiáló tulajdonság kettő helyett akárhány pontot véve is fennáll. Ebből már következik, hogy ha az S halmaz a C konvex halmaz konvex generáló részhalmaza, akkor $\operatorname{ext} C \subseteq S$. Minkowski tétele szerint kompakt, konvex C halmaz esetén az extremális pontok halmaza elég is a C halmaz konvex generálásához, tehát ekkor $\operatorname{ext} C$ a C halmaz legszűkebb konvex generáló részhalmaza.

Ha a C zárt, konvex halmaz tartalmaz egyenest (vagyis linealitás tere nemtriviális), akkor a C halmaznak nem létezhet extremális pontja. A megfelelő konvex generálási tételhez az extremális pontoknál általánosabb fogalomra van szükség, ez az extremális részhalmaz fogalma.

Adott $C \subseteq \mathcal{R}^d$ konvex halmaz esetén ennek egy $F \subseteq C$ konvex részhalmaza a C halmaz **extremális részhalmaz**a (vagy röviden **lap**ja), ha minden belsejével belemetsző C-beli szakaszt tartalmaz, vagyis

$$x_1, x_2 \in C,\ \lambda_1, \lambda_2 > 0,\ \lambda_1 + \lambda_2 = 1,\ \lambda_1 x_1 + \lambda_2 x_2 \in F \text{ esetén } x_1, x_2 \in F.$$

(Vagy még másképpen F, $C \setminus F$ konvex halmazok.) Az extremális pontok éppen az egyelemű extremális részhalmazok elemei. Annak jelölésére, hogy az F halmaz a C konvex halmaz extremális részhalmaza, az $F \triangleleft C$ jelölést alkalmazzuk. A C konvex halmaz extremális részhalmazainak halmazát jelölje $\mathcal{F}(C)$. Nyilván $\emptyset, C \in \mathcal{F}(C)$, az e két laptól különböző lapokat a C konvex halmaz **valódi** lapjainak nevezzük.

A Minkowski-tételnek megfelelő konvex generálási tétel Klee első tétele.

Ha most konvex generáló részhalmaz helyett lezárt konvex generáló részhalmazokat tekintünk, akkor az exponált pontok fogalmához jutunk.

Adott $C \subseteq \mathcal{R}^d$ konvex halmaz esetén az $x \in C$ pont a C halmaz **exponált pont**ja, ha a C halmazból annak egy támaszhipersíkja metszi ki, vagyis létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T x = \beta < a^T (C \setminus \{x\}).$$

A C konvex halmaz exponált pontjainak halmazát $\exp C$-vel jelöljük.

Straszewicz első tétele szerint minden kompakt, konvex halmaz az exponált pontjainak lezárt konvex burka. Könnyen belátható, hogy tetszőleges C konvex halmaz esetén $\exp C \subseteq \operatorname{ext} C$ (sőt Straszewicz második tétele szerint $\exp C$ az $\operatorname{ext} C$ halmaz sűrű része), így egyenest tartalmazó, zárt, konvex halmaz esetén az extremális pontok halmazával együtt az exponált pontok halmaza is üres. A megfelelő lezárt konvex generálási tételhez itt is egy általánosabb fogalomra van szükség.

Adott $C \subseteq \mathcal{R}^d$ konvex halmaz esetén ennek egy $F \subseteq C$ konvex részhalmaza a C halmaz **exponált részhalmaza** (vagy röviden **exponált lap**ja), ha a C halmazból annak egy támaszhipersíkja metszi ki, vagyis létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T C \geq \beta, \text{ és } F = \left\{x \in C : a^T x = \beta\right\}.$$

Ekkor azt mondjuk, hogy az a vektor exponálja az F részhalmazt. Annak jelölésére, hogy az F halmaz a C konvex halmaz exponált részhalmaza, az $F \triangleleft^e C$ jelölést alkalmazzuk. A C konvex halmaz exponált részhalmazainak halmazát jelölje $\mathcal{EF}(C)$. Nyilván tetszőleges C konvex halmaz esetén

$$\mathcal{EF}(C) \subseteq \mathcal{F}(C),$$

de lehet szigorú a tartalmazás, mint azt a

$$C := \left\{x \in \mathcal{R}^2 : x_1 \leq 0,\, x_2 \geq 0\right\} \cup \left\{x \in \mathcal{R}^2 : x_1 \geq 0,\, x_2 \geq x_1^2\right\}$$

halmaz mutatja.

Az első Straszewicz-tételnek megfelelő lezárt konvex generálási tétel Klee második tétele.

A fejezetben először, példák elemzése után, a Minkowski–Klee-tétel és a Straszewicz–Klee-tétel segítségével belátjuk, hogy a poliéderek éppen azok a zárt, konvex halmazok, amelyeknek véges sok exponált részhalmaza van.

Először három tételt bizonyítunk, rendre konvex kúpok, konvex halmazok, illetve poliéderek extremális és exponált részhalmazairól.

7.1. Tétel: *Legyen $K \subseteq \mathcal{R}^d$ konvex kúp. Ekkor*
a) a K konvex kúp nemüres lapjai is konvex kúpok.
b) Egy $F \subseteq K$ konvex kúp pontosan akkor lesz a K konvex kúp lapja, ha az F-beli elemek által majorált K-beli elemek is F-beliek, vagyis $x \in F$, $0 \leq_K y \leq_K x$ esetén $y \in F$.
c) A K konvex kúp legszűkebb nemüres, extremális részhalmaza $K \cap -K$.
d) A K konvex kúp legszűkebb nemüres, exponált részhalmaza $K \cap -\operatorname{cl} K$.
e) Ha a K konvex kúp nem tartalmaz egyenest, akkor az origó exponált pontja. Megfordítva is, ha a K kúp még zárt is.

Bizonyítás: a) Legyen F a K kúp lapja. Tetszőleges $x \in F$ esetén

$$x = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot (2x) \in F.$$

Itt $0 \in K$, $2x \in K$, így $F \lhd K$ miatt $0, 2x \in F$. Ez az F halmaz konvexitása miatt elég ahhoz, hogy F kúp legyen.

b) Először is legyen $F \lhd K$. Ha $x \in F$, $0 \leq_K y \leq_K x$, akkor a) miatt

$$F \ni \frac{1}{2} \cdot x = \frac{1}{2} \cdot y + \frac{1}{2} \cdot (x - y).$$

Mivel $y, x - y \in K$, azért ebből F extremalitása miatt $y \in F$ következik.

Megfordítva tegyük fel, hogy az $F \subseteq K$ konvex kúp rendelkezik az állításban leírt tulajdonsággal. Megmutatjuk, hogy ekkor $F \lhd K$. Legyen $x_1, x_2 \in K$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén $\varepsilon x_1 + (1 - \varepsilon)x_2$ F-ben van. Ekkor F kúpsága miatt

$$x := x_1 + \frac{1 - \varepsilon}{\varepsilon} \cdot x_2 \in F.$$

Mivel $x_1, x_2 \in K$, így $0 \leq_K x_1 \leq_K x$, amiből a feltétel szerint $x_1 \in F$ következik. Hasonlóan $x_2 \in F$. Ezzel beláttuk, hogy F valóban a K konvex kúp lapja.

c) Először is belátjuk, hogy $K \cap -K \lhd K$. Legyen $x_1, x_2 \in K$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén az $x := \varepsilon x_1 + (1 - \varepsilon)x_2$ pont a $K \cap -K$ altér eleme. Ekkor $-x \in K$, továbbá $(1 - \varepsilon)x_2 \in K$, így összegük, $-\varepsilon x_1$ is K-beli. De akkor $x_1 \in -K$, és hasonlóan $x_2 \in -K$ is teljesül.

Azt kell még belátnunk, hogy ha F a K konvex kúp nemüres lapja, akkor $K \cap -K \subseteq F$. Legyen $x \in K \cap -K$. Ekkor $x, -x \in K$, továbbá ezen elemek átlaga, $0 \in F$. Mivel $F \lhd K$, azért ebből $x, -x \in F$ következik. Ezzel beláttuk, hogy $K \cap -K \subseteq F$.

d) Először is belátjuk, hogy $K \cap -\operatorname{cl} K \lhd^e K$. Legyen $a \in \operatorname{ri} K^*$, ekkor, mint azt a Carver-tétel bizonyításakor láttuk,

$$\left\{x \in \operatorname{cl} K : a^T x \leq 0\right\} \subseteq \left\{x \in -\operatorname{cl} K : a^T x \geq 0\right\}.$$

Ebből pedig már könnyen következik, hogy a exponálja a $K \cap -\operatorname{cl} K$ részhalmazt, vagyis

$$a^T K \geq 0, \text{ és } K \cap -\operatorname{cl} K = \left\{x \in K : a^T x = 0\right\}.$$

Azt kell még megmutatnunk, hogy ha F a K konvex kúp nemüres, exponált lapja, akkor $K \cap -\operatorname{cl} K \subseteq F$. Tegyük fel, hogy az F lapot az a vektor exponálja, vagyis

$$a \in K^*, \text{ és } F = \left\{x \in K : a^T x = 0\right\}.$$

Azt kell belátnunk, hogy $x_0 \in K \cap -\mathrm{cl}\, K$ esetén $a^T x_0 = 0$. Mivel $x_0 \in K$, azért $a^T x_0 \geq 0$. Másrészt $-x_0$ előáll, mint valamely x_k $(k = 1, 2, \ldots)$ K-beli pontsorozat limesze, ebből pedig

$$a^T(-x_0) = \lim_{k \to \infty} a^T x_k \geq 0,$$

vagyis $a^T x_0 \leq 0$ is adódik.

e) Azt kell megmutatnunk, hogy ha K nem tartalmaz egyenest, akkor $K \cap -\mathrm{cl}\, K = \{0\}$ (vö. d)). Tegyük fel indirekt, hogy a $K \cap -\mathrm{cl}\, K$ halmaznak volna nemnulla eleme. Ekkor persze lenne nemnulla elem a bővebb $\mathrm{line}\,\mathrm{cl}\, K = (\mathrm{cl}\, K) \cap -(\mathrm{cl}\, K)$ halmazban is. Mivel $\mathrm{rec}\,\mathrm{cl}\, K = \mathrm{rec}\,\mathrm{ri}\, K$, azért oda jutnánk, hogy $\mathrm{ri}\, K$, és így K is tartalmazna egyenest, ellentmondásban a feltétellel.

Ha K zárt, és az origó exponált pontja, akkor világos, hogy nem mehet át K-beli egyenes az origón, így K zártsága miatt K egyetlen más pontján sem. $\square$

7.2. Tétel: *Legyenek $C, C' \subseteq \mathcal{R}^d$ konvex halmazok, $C' \subseteq C$. Ekkor*
a) $F \lhd C$, $F \subseteq C'$ esetén $F \lhd C'$ (és ez az állítás fennáll $\lhd$ helyett $\lhd^e$-vel is);
b) $F \lhd C'$ és $C' \lhd C$ esetén $F \lhd C$ (ez az állítás nem marad igaz, ha benne a $\lhd$ jelet $\lhd^e$-re cseréljük);
c) $F \lhd C$, $(\mathrm{ri}\, C') \cap F \neq \emptyset$ esetén $C' \subseteq F$, speciálisan $F \neq C$ esetén $F \subseteq \mathrm{rb}\, C$ (és akkor $\dim F < \dim C$);
d) $F \lhd C$ esetén $F = C \cap \mathrm{aff}\, F$;
e) $F \lhd C$ esetén $F = C \cap \mathrm{cl}\, F$;
f) ha C zárt is, akkor lapjai zártak, $F \lhd C$ esetén $\mathrm{rec}\, F \lhd \mathrm{rec}\, C$, továbbá $\mathrm{line}\, F = \mathrm{line}\, C$;
g) $x \in C$, $\dim(\mathrm{cone}\,(C - x))^* = d$ esetén $x \in \exp C$.

Bizonyítás: Az a) és b) állítások nyilvánvalóak a definíciókból.

c) Legyen $x_0 \in (\mathrm{ri}\, C') \cap F$. Ekkor minden $x \in C'$ pont túlhúzható az x_0 ponton a C' halmazban, vagyis létezik $y \in C'$ pont úgy, hogy x_0 az x és y közti szakasz belsejének az eleme. Mivel $F \lhd C$, azért ebből $x \in F$ adódik.

Ha most $F \neq C$, akkor $F \cap \mathrm{ri}\, C = \emptyset$ (különben $C \subseteq F$ lenne a már igazoltak szerint), így $F \subseteq \mathrm{rb}\, C$. Nem lehet $\dim F = \dim C$, abból ugyanis $\mathrm{aff}\, F = \mathrm{aff}\, C$, és így $F \subseteq C$ miatt $\mathrm{ri}\, F \subseteq \mathrm{ri}\, C$ következne, holott F diszjunkt a C relatív belsejétől.

d) és e) közül például d)-t látjuk be, e) hasonlóan igazolható. Mivel

$$F \subseteq C \cap \mathrm{aff}\, F \subseteq \mathrm{aff}\, F,$$

azért $\operatorname{aff} F = \operatorname{aff}(C \cap \operatorname{aff} F)$, amiből $\operatorname{ri} F \subseteq \operatorname{ri}(C \cap \operatorname{aff} F)$ adódik. A $C' := C \cap \operatorname{aff} F$ választással alkalmazzuk c)-t.

f) d)-ből és e)-ből is következik, hogy zárt, konvex halmaz lapjai is zártak. Az állítás hátralévő részéből lássuk be például azt, hogy ha $F \lhd C$ nemüres, akkor $\operatorname{rec} F \lhd \operatorname{rec} C$.

Legyen $z_1, z_2 \in \operatorname{rec} C$, $z \in \operatorname{rec} F$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén $\varepsilon z_1 + (1-\varepsilon) z_2 = z$. Ekkor tetszőleges $x \in F$, $\lambda \geq 0$ esetén

$$x + \lambda z_1,\ x + \lambda z_2 \in C,\ x + \lambda z \in F,$$
$$\varepsilon(x + \lambda z_1) + (1-\varepsilon)(x + \lambda z_2) = x + \lambda z,$$

amiből F extremalitása miatt $x + \lambda z_1, x + \lambda z_2 \in F$, és így $z_1, z_2 \in \operatorname{rec} F$ következik. Itt C zártsága ahhoz kellett, hogy $\operatorname{rec} F \subseteq \operatorname{rec} C$ legyen.

g) Legyen $K := \operatorname{cone}(C - x)$. A feltétel szerint $K^* - K^* = \mathcal{R}^d$, vagyis adjungált kúpokat véve $(\operatorname{cl} K) \cap -(\operatorname{cl} K) = \{0\}$. Ez azt jelenti, hogy $\operatorname{line} \operatorname{cl} K$ a triviális kúp, vagyis $\operatorname{cl} K$ nem tartalmaz egyenest. De akkor K sem tartalmazhat egyenest, és 7.1 e) szerint $0 \in \exp K$. Ekkor $0 \in \exp(C - x)$, vagyis $x \in \exp C$ is teljesül. $\square$

7.3. Tétel: *Legyen $P \subseteq \mathcal{R}^d$ poliéder, továbbá F a P poliéder nemüres lapja. Legyen $A \in \mathcal{R}^{m \times d}$, $b \in \mathcal{R}^m$, amelyre $P = \{x \in \mathcal{R}^d : Ax \geq b\}$, és jelölje I az A mátrix azon i sorindexeinek halmazát, amelyekre ${}_i a x = b_i$ $(x \in F)$. Jelölje $A_F^=$ az A mátrix I-beli indexű soraiból álló részmátrixát, $A_F^>$ pedig azt a mátrixot, amelyet úgy kapunk az A mátrixból, hogy elhagyjuk annak I-beli indexű sorait. Definiáljuk hasonló módon a $b_F^=, b_F^>$ vektorokat. Ekkor*

a) $\operatorname{aff} F = \{x \in \mathcal{R}^d : A_F^= x = b_F^=\}$*;*
b) $F = \{x \in \mathcal{R}^d : A_F^= x = b_F^=, A_F^> x \geq b_F^>\}$*;*
c) $\operatorname{ri} F = \{x \in \mathcal{R}^d : A_F^= x = b_F^=, A_F^> x > b_F^>\}$*;*
d) F a P poliéder exponált lapja.

Bizonyítás: Először is vegyük észre, hogy létezik $x_0 \in F$ pont úgy, hogy $A_F^= x_0 = b_F^=, A_F^> x_0 > b_F^>$ legyen. Ha ugyanis $i \notin I$, akkor létezik $x_i \in F$ pont, amelyre az i-edik egyenlőtlenség szigorúan teljesül, és akkor ezeknek az x_i pontoknak az átlaga megfelelő x_0 pont.

a) Jelölje S az állításban az egyenlőség jobb oldalán szereplő halmazt. Ekkor egyrészt S affin halmaz, és tartalmazza az F halmazt. Ebből látszik az $\operatorname{aff} F \subseteq S$ tartalmazás. Másrészt ha x_0 a bizonyítás elején leírt tulajdonságú pont, akkor tetszőleges S-beli pont behúzható az x_0 mellé a P

halmazba, és túlhúzható az x_0 ponton a P halmazban. Az így kapott két P-beli pont által meghatározott szakasz a belsejében tartalmazza az F-beli x_0 pontot, tehát végpontjai is F-beliek, így a végpontok által meghatározott egyenes minden pontja aff F-beli. Ez mutatja az $S \subseteq \operatorname{aff} F$ tartalmazást is.

b) az a) állítás következménye, mivel $F = P \cap \operatorname{aff} F$ 7.2 d) szerint.

c) Az F lap b) szerint előáll, mint aff F és az $\{x : {}_i a x \geq b_i\}$ $(i \notin I)$ zárt félterek metszete. E halmazok relatív belsejei összemetszenek (x_0 mutatja), így F relatív belseje a relatív belsők metszete.

d) Legyen a^T az $A_F^=$ mátrix sorainak összege, β pedig a $b_F^=$ vektor elemeinek összege. Könnyen belátható, hogy ekkor $\{x : a^T x = \beta\}$ a P poliéder támaszhipersíkja, amely P-ből éppen az F lapot metszi ki. $\square$

Máris látszik, hogy poliéderek esetén az extremális részhalmazok éppen az exponált részhalmazok. Továbbá minden nemüres extremális részhalmaz előáll úgy, hogy a poliédert leíró egyenlőtlenség-rendszer néhány egyenlőtlenségét egyenlőségnek kötjük meg, speciálisan egy poliéder zárt, konvex halmaz, amelynek véges sok extremális (azaz exponált) részhalmaza van. E tétel megfordításának igazolása a következő célunk, ehhez az extremális/exponált részhalmazok konvex/lezárt konvex generáló tulajdonságait vizsgáljuk.

7.4. Tétel: (Minkowski) *Minden kompakt, konvex halmaz az extremális pontjainak konvex burka.*

Bizonyítás: Legyen $C \subseteq \mathcal{R}^d$ kompakt, konvex halmaz, továbbá $C' := \operatorname{conv} \operatorname{ext} C$. Nyilvánvaló, hogy $C' \subseteq C$. A fordított irányú tartalmazást $\dim C$ szerinti teljes indukcióval bizonyítjuk. Ha $\dim C \leq 1$, akkor C zárt intervallum, amelynek extremális pontjai a végpontjai, így ekkor $C = C'$.

Az általános esetre rátérve legyen $x \in C$, $0 \neq z \in \operatorname{par} C$ tetszőleges. A C halmaz korlátossága miatt létezik $\lambda \geq 0$ úgy, hogy $x_1 := x + \lambda z$ a C halmaz relatív határpontja. Legyen H a C halmaz x_1 pontbeli, a C halmazt nem tartalmazó támaszhipersíkja, ekkor $F := H \cap C$ exponált részhalmaza, így lapja is C-nek. Mivel $C \nsubseteq H$, azért $F \neq C$, tehát F kompakt, konvex halmaz, amelynek dimenziója kisebb, mint a C halmazé. Az indukciófeltétel miatt $F = \operatorname{conv} \operatorname{ext} F$. Másfelől 7.3 szerint $\operatorname{ext} F \subseteq \operatorname{ext} C$. Mindebből $x_1 \in C'$ adódik (a C halmaz zártsága miatt $x_1 \in F$). A fenti gondolatmenetet z helyett $-z$-re ismételve jutunk egy $x_2 := x - \mu z \in \operatorname{rb} C$ ponthoz, amelyre szintén teljesül, hogy $x_2 \in C'$, így $x \in [x_1, x_2]$ is C'-beli. $\square$

Minkowski kompakt, konvex halmazokra vonatkozó tételének egyszerű következménye Minkowski poliéder kúpokra vonatkozó tétele. Ha ugyanis az $R \subseteq \mathcal{R}^d$ poliéder kúpot elmetsszük például a

$$\left\{x \in \mathcal{R}^d : -1 \leq x_i \leq 1 \ (i = 1, \ldots, d)\right\}$$

kockapoliéderrel, akkor az így kapott poliéder extremális pontjai konvex burka, és könnyen látható, hogy ezen extremális pontok konvex kúp burka viszont az R kúp.

Vegyük még észre, hogy 7.4 bizonyításában kevéssé használtuk ki a C halmaz korlátosságát. Olyan $z = z(x)$ irányra volt szükségünk, hogy x-ből z és $-z$ irányba indulva is kijuthassunk a C relatív határáig. Nevezzük a C zárt, konvex halmazt **primitív**nek, ha ilyen irány nincs valamely $x \in (\mathrm{ri}\,)C$ pont esetén. Ekkor a fenti bizonyítás megismétlésével az alábbi tételhez jutunk: Minden C zárt, konvex halmaz előáll primitív lapjainak konvex burkaként. Egy affin halmaz, vagy annak a fele (tehát az affin halmaz elmetszve egy őt nem tartalmazó zárt féltérrel) nyilván primitív. Megmutatjuk, hogy éppen ezek a halmazok primitívek.

Az S halmaz az L **altér fele**, ha L és egy L-et nem tartalmazó homogén zárt féltér metszeteként áll elő, vagyis $S = \{x \in L : a^T x \geq 0\}$ alakú, ahol $0 \neq a \in L$.

Az S halmaz az M **affin halmaz fele**, ha M és egy M-et nem tartalmazó zárt féltér metszeteként áll elő, vagyis $S = \{x \in M : a^T x \geq \beta\}$ alakú, ahol $0 \neq a \in \mathrm{par}\, M$, $\beta \in \mathcal{R}$.

7.5. Lemma: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmaz. Ekkor az alábbi állítások ekvivalensek:*
a) $\mathrm{par}\, C = \mathrm{rec}\, C \cup -\mathrm{rec}\, C$ *(vagyis C primitív);*
b) $\mathrm{rec}\, C$ *a* $\mathrm{par}\, C$ *altér, vagy annak a fele;*
c) C *az* $\mathrm{aff}\, C$ *affin halmaz, vagy annak a fele;*
d) az $\mathrm{rb}\, C$ *halmaz konvex;*
e) $C \supset \mathrm{conv}\,\mathrm{rb}\, C$.

Bizonyítás: Először azt látjuk be, hogy a)⇒b). Tegyük fel, hogy a $\mathrm{rec}\, C$ kúp nem az egész $L := \mathrm{par}\, C$ altér, legyen $x_0 \in L \setminus \mathrm{rec}\, C$. Mivel $\mathrm{rec}\, C$ zárt, konvex kúp, azért az első Hahn–Banach-tétel szerint létezik $a \in L$ vektor úgy, hogy

$$a^T x_0 < 0 \leq a^T \mathrm{rec}\, C.$$

Tehát a $H^+ := \{x : a^T x \geq 0\}$ zárt féltér tartalmazza a $\operatorname{rec} C$ kúpot. Megmutatjuk, hogy $\operatorname{rec} C = H^+ \cap L$. Teljesül, hogy

$$z \in L,\ a^T z > 0 \text{ esetén } z \in \operatorname{rec} C,$$

különben sem z, sem $-z$ nem lenne $\operatorname{rec} C$-ben, ellentétben a feltétellel. Tehát $L \cap \operatorname{int} H^+ \subseteq \operatorname{rec} C$, amiből a $\operatorname{rec} C$ kúp zártsága miatt $L \cap H^+ \subseteq \operatorname{rec} C$ következik. A másik irányú tartalmazás nyilvánvaló.

Másodszor azt igazoljuk, hogy b)$\Rightarrow$c). Legyen $x_0 \in C$, ekkor $\operatorname{aff} C = x_0 + \operatorname{par} C$, így ha $\operatorname{rec} C = \operatorname{par} C$, akkor

$$\operatorname{aff} C = x_0 + \operatorname{rec} C \subseteq C \subseteq \operatorname{aff} C$$

miatt $C = \operatorname{aff} C$ affin halmaz. Tegyük fel most, hogy $\operatorname{rec} C = \{z \in \operatorname{par} C : a^T z \geq 0\}$ valamely $0 \neq a \in \operatorname{par} C$ esetén. Mivel ekkor $-a \notin \operatorname{rec} C$, azért tetszőleges C-beli pontból $-a$ irányba kiindulva találunk utolsó C-beli pontot, legyen egy ilyen x_0. Könnyen belátható, hogy

$$C = \left\{x \in \operatorname{aff} C : a^T x \geq a^T x_0\right\}$$

az $\operatorname{aff} C$ fele.

A tétel c)$\Rightarrow$d), d)$\Rightarrow$e) része nyilvánvaló.

Végül bebizonyítjuk, hogy e)$\Rightarrow$a). Legyen $x_0 \in C \setminus \operatorname{conv} \operatorname{rb} C$. Ekkor bármely $z \in \operatorname{par} C$ esetén z vagy $-z$ $\operatorname{rec} C$-beli, különben x_0-ból elindulva mindkét irányban kiérnénk a C relatív határáig, és ekkor x_0 előállna $\operatorname{rb} C$-beli pontok konvex kombinációjaként, ellentétben a feltétellel. □

A fentiek azonnali következménye az alábbi

7.6. Tétel: (első Klee-tétel) *Minden zárt, konvex halmaz előáll azon extremális részhalmazai úniójának konvex burkaként, amelyek affin halmazok, vagy affin halmazok felei.* □

Speciális esetként adódik Minkowski tétele, vagy hogy egyenesmentes zárt, konvex halmaz előáll extremális pontjainak és extremális sugarainak konvex kombinációjaként. (**Extremális sugár** alatt az $\{x + \lambda z : \lambda \geq 0\}$ alakú extremális részhalmazokat értjük, úniójukat jelölje $\operatorname{rext} C$. Egy extremális sugár végpontja természetesen extremális pont.)

Nyilván ha $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz, akkor

$$C = \operatorname{line} C + C \cap (\operatorname{line} C)^\perp,$$

speciálisan

$$\operatorname{rec} C = \operatorname{line} C + (\operatorname{rec} C) \cap (\operatorname{line} C)^\perp = \operatorname{line} C + \operatorname{rec} C',$$

ahol $C' := C \cap (\operatorname{line} C)^\perp$ a C egyenesmentes része.

Mivel az első Klee-tétel szerint

$$C' = \operatorname{conv}(\operatorname{ext} C' \cup \operatorname{rext} C') \subseteq \operatorname{conv}(\operatorname{ext} C') + \operatorname{rec} C' \subseteq C',$$

azért a

$$C = \operatorname{line} C + \operatorname{conv}(\operatorname{ext} C') + \operatorname{rec} C' = \operatorname{conv}(\operatorname{ext} C') + \operatorname{rec} C$$

felbontás is látszik.

Ha $P \subseteq \mathcal{R}^d$ poliéder, akkor előáll $P = Q + R$ alakban, ahol Q politóp, R poliéder kúp, szükségképpen a P recessziós kúpja. Most azt is látjuk, hogy $Q := \operatorname{conv}(\operatorname{ext} P')$ megfelel, ahol a $P' := P \cap (\operatorname{line} P)^{\perp}$ poliéder a P poliéder egyenesmentes része. Sőt, ez a Q politóp a legszűkebb a $(\operatorname{line} P)^{\perp}$-beli politópok közül, amelyre a P poliéder előáll, mint a politóp és a $\operatorname{rec} P$ kúp összege:

7.7. Állítás: *Ha $Q \subseteq \mathcal{R}^d$ politóp, $R \subseteq \mathcal{R}^d$ végesen generált kúp, továbbá $L = R \cap -R$ az R kúp linealitás tere, akkor*

$$\operatorname{ext}((Q + R) \cap L^{\perp}) \subseteq \operatorname{ext}((Q + L) \cap L^{\perp}).$$

Itt $(Q+L)\cap L^{\perp}$ a Q politóp vetülete az $L^{\perp}$ altérre, tehát politóp. Speciálisan Minkowski tétele szerint extremális pontjainak konvex burka, így a fenti tartalmazásból

$$\operatorname{conv}\operatorname{ext}((Q + R) \cap L^{\perp}) \subseteq (Q + L) \cap L^{\perp}$$

adódik. Ebből az is látszik, hogy ha $P \subseteq \mathcal{R}^d$ poliéder, akkor

$$Q_0 := \operatorname{conv}\operatorname{ext}(P \cap (\operatorname{line} P)^{\perp})$$

a legszűkebb $(\operatorname{line} P)^{\perp}$-beli Q politóp, amelyre $P = Q + \operatorname{rec} P$.

Bizonyítás: Először $L = \{0\}$ esetén bizonyítjuk az állítást. Azt kell belátnunk, hogy $\operatorname{ext}(Q + R) \subseteq \operatorname{ext} Q$, ami $Q \subseteq Q + R$ miatt azzal ekvivalens, hogy $\operatorname{ext}(Q + R) \subseteq Q$. Márpedig ha $q + r \in \operatorname{ext}(Q + R)$ (ahol $q \in Q$ és $r \in R$), akkor a $q + r = \frac{1}{2}q + \frac{1}{2}(q + 2r)$ egyenlőségből $q + r$ extremalitása miatt $q + r = q$, és így $r = 0$ adódik.

Az általános eset erre vezethető vissza, ugyanis

$$(Q + R) \cap L^{\perp} = ((Q + L) \cap L^{\perp}) + (R \cap L^{\perp}).$$

□

Most már rátérhetünk a Minkowski–Klee-tételnek megfelelő lezárt konvex generálási tételekre.

7.8. Tétel: (első Straszewicz-tétel) *Minden kompakt, konvex halmaz az exponált pontjainak lezárt konvex burka.*

Bizonyítás: Legyen $C \subseteq \mathcal{R}^d$ kompakt, konvex halmaz, $C' := \operatorname{cl}\operatorname{conv}\exp C$. Ekkor nyilván C' kompakt, konvex halmaz; nemüres, hiszen egy tetszőleges rögzített ponttól legtávolibb C-beli pont eleme; továbbá $C' \subseteq C$. A fordított tartalmazáshoz tegyük fel indirekt, hogy létezik $x_0 \in C \setminus C'$ pont. Az első Hahn–Banach-tétel szerint ekkor létezik H hipersík úgy, hogy az általa meghatározott H^+ zárt féltér tartalmazza a C' halmazt, de az x_0 pontot nem. Feltehető, hogy H homogén hipersík, sőt az is, hogy az x_0 pont vetülete a H altérre éppen az origó. A (C'-t tartalmazó) $C \cap H^+$ halmaz kompakt, ezért belefoglalható egy $O(0, \varrho)$ gömbbe. Ekkor az

$$S := (H \cap O(0, \varrho)) + [0, -(\varrho/||x_0||)x_0]$$

henger tartalmazza a $C \cap H^+$ halmazt, és így C'-t is. Könnyen belátható, hogy elég nagy $\mu > 0$ szám esetén

$$S \subseteq O(-\mu x_0, (\mu + 1)||x_0||)$$

Az utóbbi gömb középpontját jelölje y, és legyen x_1 az y ponttól legtávolabbi C-beli pont. Ekkor egyrészt $x_1 \in \exp C$, másrészt

$$||y - x_1|| \geq ||y - x_0|| = (\mu + 1)||x_0||$$

miatt $x_1 \notin H^+$, ami ellentmond annak, hogy a H^+ zárt féltér a C halmaz minden exponált pontját tartalmazza. $\square$

Straszewicz tételének még egy bizonyítását adjuk, amely az előző bizonyítás duálisa.

Bizonyítás: A C halmaznak van exponált pontja (a valamely ponttól legtávolabbi C-beli pont ilyen), feltehető, hogy ez éppen az origó. A C halmaz kompaktsága, illetve $0 \in C$ miatt $0 \in \operatorname{int} C^\circ$, és $C^{\circ\circ} = C$. Jelölje S a C° halmaz határa sima pontjainak halmazát, H_s^+ pedig azt az egyértelműen meghatározott, a C° halmazt tartalmazó zárt félteret, amelynek határa támaszhipersíkja a C° halmaznak, és tartalmazza az $s \in S$ sima pontot. Mivel $0 \in \operatorname{int} C^\circ$, azért létezik egyértelműen $c_s \in \mathcal{R}^d$ vektor úgy, hogy

$$H_s^+ = \{c_s\}^\circ = \left\{x \in \mathcal{R}^d : c_s^T x \geq -1\right\} \ (s \in S).$$

5.14, 5.15 szerint $C^\circ = \cap_{s\in S} H_s^+$. Könnyen belátható, hogy minden c_s pont a $C^{\circ\circ} = C$ halmaz exponált pontja lesz (az s vektor exponálja), így

$$\begin{aligned} C &= C^{\circ\circ} = (\cap_{s\in S} H_s^+)^\circ = (\cap_{s\in S}(\{c_s\}^\circ))^\circ = (\cup_{s\in S}\{c_s\})^{\circ\circ} = \\ &= \operatorname{cl}\operatorname{conv}(\{0\}\cup\{c_s : s\in S\}) \subseteq \operatorname{cl}\operatorname{conv}\exp C = C', \end{aligned}$$

ami a tétel nemtriviális része. □

Minkowski tétele és az első Straszewicz-tétel segítségével könnyen belátható, hogy tetszőleges zárt, konvex halmaz esetén a halmaz exponált pontjai a halmaz extremális pontjainak sűrű részét alkotják (persze ez a tétel semmitmondó, ha az alaphalmaz tartalmaz egyenest). Előbb az alábbi lemmát igazoljuk.

7.9. Lemma: *Legyenek $C_1, C_2 \subseteq \mathcal{R}^d$ konvex halmazok, $x \in C_1 \cap C_2$. Tegyük fel, hogy valamely $\varepsilon > 0$ esetén $C_1 \cap O(x,\varepsilon) = C_2 \cap O(x,\varepsilon)$. Ekkor*
a) $x \in (\operatorname{ext} C_1) \cap (\operatorname{ext} C_2)$, vagy $x \notin (\operatorname{ext} C_1) \cup (\operatorname{ext} C_2)$;
b) $x \in (\exp C_1) \cap (\exp C_2)$, vagy $x \notin (\exp C_1) \cup (\exp C_2)$.

Bizonyítás: a) Szimmetriaokok miatt elég azt igazolni, hogy ha $x \in \operatorname{ext} C_1$, akkor $x \in \operatorname{ext} C_2$ is teljesül. Legyen $x \in \operatorname{ext} C_1$, és tegyük fel, hogy valamely $y_1, y_2 \in C_2$, $0 < \lambda < 1$ esetén $x = \lambda y_1 + (1-\lambda) y_2$ teljesül. Húzzuk be az y_1, y_2 pontokat az x pont mellé az $O(x,\varepsilon)$ halmazba, az így kapott pontokat jelölje y_1', y_2'. Ekkor x az y_1', y_2' pontok által meghatározott szakasz belsejében van, továbbá $y_1', y_2' \in C_2 \cap O(x,\varepsilon)$ miatt $y_1', y_2' \in C_1 \cap O(x,\varepsilon)$ is teljesül. Mivel $x \in \operatorname{ext} C_1$, azért ebből $y_1' = y_2' = x$, és így $y_1 = y_2 = x$ adódik.

b) Itt is elég csak annyit igazolni, hogy ha $x \in \exp C_1$, akkor $x \in \exp C_2$. Legyen $a \in \mathcal{R}^d$ olyan vektor, amelyre $a^T(C_1 \setminus \{x\}) > a^T x$. Megmutatjuk, hogy ekkor $a^T(C_2 \setminus \{x\}) > a^T x$ is teljesül. Legyen $y \in C_2 \setminus \{x\}$. Húzzuk be az y pontot az x pont mellé az $O(x,\varepsilon)$ gömbbe, az így kapott pontot jelölje y'. Ekkor valamely $\lambda > 0$ esetén $y' = x + \lambda(y - x)$, és most is könnyen belátható, hogy $y' \in C_1$ teljesül. Ezért $a^T y' > a^T x$, amiből $a^T y > a^T x$ is következik. □

7.10. Tétel: (második Straszewicz-tétel) *Legyen $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz. Ekkor* $\operatorname{ext} C \subseteq \operatorname{cl}\exp C$.

Bizonyítás: Ha C még kompakt is, akkor az első Straszewicz-tétel szerint $C = \operatorname{cl}\operatorname{conv}\exp C$, továbbá $\operatorname{cl}\exp C$, és így $\operatorname{conv}\operatorname{cl}\exp C$ is kompakt (vö. 3.22). Ezért ekkor

$$C = \operatorname{cl}\operatorname{conv}\exp C \subseteq \operatorname{conv}\operatorname{cl}\exp C \subseteq C$$

miatt $\mathrm{cl}\exp C$ a C konvex generáló részhalmaza, és így tartalmazza az $\mathrm{ext}\, C$ halmazt.

Általában legyen $x \in \mathrm{ext}\, C$, az x pontot kell előállítanunk $\exp C$-beli pontsorozat limeszeként. Legyen $C' := C \cap \mathrm{cl}\, O(x, 1)$, ekkor C' kompakt, konvex halmaz, továbbá a lemma szerint (alkalmazzuk $\varepsilon = 1$ választással) $x \in \mathrm{ext}\, C'$. A tétel már igazolt esete szerint x előáll, mint $x = \lim_{k\to\infty} x_k$, ahol $x_k \in \exp C'$ $(k = 1, 2, \ldots)$. Elég nagy k indexek esetén a sorozat elemei kevesebb, mint egységnyi távolságra vannak az x ponttól, és akkor a lemma szerint $\exp C$-beliek, mivel a C és a C' halmazt elmetszve az x_k pont körüli $1 - ||x_k - x||$ sugarú nyílt gömbbel, ugyanazt a halmazt kapjuk. $\square$

A $C \subseteq \mathcal{R}^d$ konvex halmaz **exponált sugar**ai az $\{x + \lambda z : \lambda \geq 0\}$ alakú exponált részhalmazai, ahol $x, z \in \mathcal{R}^d$. Az exponált sugarak unióját jelölje $\mathrm{rexp}\, C$.

7.11. Tétel: (második Klee-tétel) *Minden zárt, konvex halmaz előáll azon exponált részhalmazai uniójának lezárt konvex burkaként, amelyek affin halmazok, vagy affin halmazok felei.*

Bizonyítás: A tételt először abban az esetben bizonyítjuk, mikor a szóban forgó $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz nem tartalmaz egyenest.

Az egyszerűség kedvéért feltesszük, hogy $\dim C = d$, és $d \geq 2$ (ha $\dim C < 2$, akkor triviálisan igaz a tétel). Jelölje S az $(\exp C) \cup (\mathrm{rexp}\, C)$ halmazt, ekkor nyilván $\mathrm{cl}\,\mathrm{conv}\, S \subseteq C$ zárt, konvex halmaz, amely tartalmazza a C halmaz összes extremális pontját, és így nem üres. Tegyük fel indirekt, hogy $\mathrm{cl}\,\mathrm{conv}\, S$ nem az egész C, ekkor az első Hahn–Banach-tétel szerint létezik H hipersík, amely belemetsz az $\mathrm{int}\, C$ halmazba, de elkerüli a $\mathrm{cl}\,\mathrm{conv}\, S$ halmazt. Az első Klee-tétel egyszerű következménye, hogy ekkor a $C \cap H$ halmaznak van legalább egy extremális pontja, és akkor a második Straszewicz-tételből adódóan létezik $x \in \exp(C \cap H)$ exponált pontja is. Az exponált pontság definíciója szerint létezik H-ban egy $d-2$-dimenziós M affin halmaz, amely az x pontot metszi ki a $C \cap H$ halmazból. Speciálisan M nem metsz bele a nemüres $\mathrm{int}\, C$ halmazba (hiszen $x \notin \mathrm{int}\, C$), így a második Hahn–Banach-tétel szerint létezik őt tartalmazó H' hipersík, ami nem metsz bele az $\mathrm{int}\, C$ halmazba. Ez a H' a C halmaz támaszhipersíkja, így $C' := C \cap H'$ a C halmaz exponált részhalmaza. A "$\lhd$" reláció tranzitivitása miatt $\mathrm{ext}\, C' \subseteq \mathrm{ext}\, C$, így $\mathrm{ext}\, C \subseteq \mathrm{cl}\,\mathrm{conv}\, S$ miatt $(\mathrm{ext}\, C') \cap H = \emptyset$. Mivel $H \cap H' = M$ ($H \cap H' \supset M$ esetén $H = H'$ lenne), azért $H \cap C' = \{x\}$. Az $x \in H$ pont nem lehet $\exp C'$-beli (hiszen $\exp C' \subseteq \mathrm{ext}\, C'$, és $(\mathrm{ext}\, C') \cap H = \emptyset$), így $H \cap \mathrm{ri}\, C' = \{x\}$ (különben

5.6 szerint létezne a vektor, amelyre $a^T H < a^T \operatorname{ri} C'$; ez az a vektor exponálná az x pontot), és akkor (affin burkot véve) $H \cap \operatorname{aff} C' = \{x\}$, amiből $\dim C' \leq 1$ adódik (vö. 2.10). A C halmazra tett egyenesmentességi feltétel szerint C' nem lehet egyenes. Ugyanakkor nem lehet szakasz sem, annak végpontjai ugyanis $\operatorname{ext} C'$-beliek, és így $\operatorname{cl} \operatorname{conv} S$-beliek lennének, ebből pedig $x \in \operatorname{cl} \operatorname{conv} S$ következne. Marad az a lehetőség, hogy C' egy exponált sugara a C halmaznak, amiből ismét csak $x \in S$ következne, ellentmondásban azzal, hogy $x \notin \operatorname{cl} \operatorname{conv} S$.

Általában, ha a C halmazról nem tesszük fel, hogy egyenesmentes, akkor jelölje linealitás terét L. A tétel már igazolt része szerint a C halmaz egyenesmentes része, $C \cap L^\perp$ előáll, mint exponált pontjai és exponált sugarai lezárt konvex burka. Ha most egyenként az exponált pontokhoz és az exponált sugarakhoz is hozzáadjuk az L alteret, akkor könnyen láthatóan a C halmaz olyan exponált részhalmazait kapjuk, amelyek affin halmazok, illetve affin halmazok felei. Az is belátható, hogy ekkor $C = L + (C \cap L^\perp)$ előáll ezen exponált részhalmazok uniójának lezárt konvex burkaként. $\square$

Ha $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz, amelynek csak véges sok exponált részhalmaza van, akkor a második Klee-tétel szerint C véges sok poliéder uniójának lezárt konvex burkaként áll elő, tehát poliéder. Ezzel beláttuk az alábbi tétel még nem igazolt irányát is.

7.12. Tétel: *A poliéderek éppen azok a zárt, konvex halmazok, amelyeknek véges sok exponált részhalmaza van.* $\square$

Megjegyezzük, hogy a tétel nehezebbik irányára közvetlenebb, duális bizonyítás is adható 5.14 és 5.15 segítségével. Legyen ugyanis $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz, amelynek véges sok exponált részhalmaza van. Megmutatjuk, hogy C poliéder. Feltehető, hogy $\dim C = d$, ekkor C határa sima pontjaihoz tartozó, egyértelmű támaszhipersíkok által meghatározott zárt félterek metszete. Ezekhez a zárt félterekhez injektív módon hozzárendelhető a határuk által a C halmazból kimetszett exponált lap. Mivel C-nek csak véges sok exponált lapja van, azért a fentiek szerint véges sok zárt féltér metszeteként is előáll, vagyis poliéder.

A fejezet most következő részében a poliéderek minimális és maximális oldalaival foglalkozunk. Ezen belül megmutatjuk, hogy a bázismegoldásnak, ennek az algebrai módon definiált fogalomnak mi a geometriai tartalma. Az is kiderül, hogy a bázismegoldások halmaza csak a megoldáshalmaztól függ, annak leírásától nem.

7.13. Állítás: *Legyen $P \subseteq \mathcal{R}^n$ poliéder, továbbá $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, amelyre $P = \{x : Ax \geq b\}$. Jelölje L a P poliéder linealitás terét, továbbá legyen $P' := P \cap L^\perp$ a P egyenesmentes része. Ekkor egy x_0 vektor pontosan akkor bázismegoldása az $Ax \geq b$ rendszernek, ha $x_0 \in L + \operatorname{ext} P'$.*

Bizonyítás: Válasszunk $x \in L$, $x' \in P'$ pontokat úgy, hogy $x_0 = x + x'$ legyen. (Mivel $L = \operatorname{Ker} A$, azért speciálisan $Ax_0 = Ax'$.) Azt kell megmutatnunk, hogy x_0 pontosan akkor bázismegoldás, ha $x' \in \operatorname{ext} P'$. Legyen I az A mátrix azon i sorindexeinek a halmaza, amelyekre ${}_i a x_0 = b_i$ teljesül.

Tegyük fel először, hogy x_0 bázismegoldás, és hogy $x' = \varepsilon x_1 + (1-\varepsilon)x_2$, ahol $0 < \varepsilon < 1$, és $x_1, x_2 \in P'$. Ekkor

$$b_I = {}_I A x_0 = {}_I A x' = \varepsilon {}_I A x_1 + (1-\varepsilon) {}_I A x_2 \geq b_I$$

miatt $x_1 - x_2 \in \operatorname{Ker} {}_I A = \operatorname{Ker} A$, vagyis $x_1 - x_2 \in L \cap L^\perp = \{0\}$. Ezért $x' \in \operatorname{ext} P'$.

A másik irányhoz azt kell megmutatnunk, hogy ha $x' \in \operatorname{ext} P'$, akkor ${}_I A z = 0$ esetén $Az = 0$. Ha ${}_I A z = 0$, akkor elég kis $\varepsilon > 0$ esetén $x_1, x_2 := x_0 \pm \varepsilon z \in P$. Ekkor

$$x' = \frac{1}{2}(x' + \varepsilon z') + \frac{1}{2}(x' - \varepsilon z'),$$

ahol z' a z vektor vetülete az $L^\perp$ altérre. Az x' pont extremalitása miatt $z' = 0$, vagyis $z \in L = \operatorname{Ker} A$. □

7.14. Következmény: *Legyen $P \subseteq \mathcal{R}^n$ poliéder, és tegyük fel, hogy előáll*

$$P = \{x \in \mathcal{R}^n : Ax = b, x \geq 0\}$$

alakban, ahol $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$. Ekkor egy x_0 vektor pontosan akkor bázismegoldása az $Ax = b, x \geq 0$ rendszernek, ha extremális pontja P-nek.

Bizonyítás: Az előző állítás következménye, ugyanis x_0 pontosan akkor bázismegoldása az $Ax = b, x \geq 0$ rendszernek, ha bázismegoldása az $Ax \geq b$, $-Ax \geq -b$, $Ex \geq 0$ rendszernek, és a megoldások halmaza most egyenesmentes. □

7.15. Állítás: *Legyen P, L, P', mint 7.13-ban. Ekkor P minimális (szűkebb valódi lapot nem tartalmazó, nemüres) lapjai éppen az $x + L$ alakú halmazok, ahol $x \in \operatorname{ext} P'$. Speciálisan az $Ax \geq b$ rendszer bázismegoldásai éppen a P poliéder minimális lapjainak elemei.*

Bizonyítás: Először megmutatjuk, hogy az $L^\perp$ altérrel való metszés tartalmazástartó bijekció a P és a P' poliéderek lapjai között. Ehhez elevenítsük fel az alábbi két észrevételt:
– Ha F lapja P-nek, akkor $\operatorname{line} F = \operatorname{line} P$.
– P' lapjai éppen az $F' = F \cap L^\perp$ alakú halmazok, ahol F a P lapja.

Mindkét észrevétel 7.3 egyszerű következménye (az elsőt általánosabban is beláttuk 7.2-ben, a másodikat általánosabban is belátjuk 7.23-ban). Ezek után már könnyen igazolható, hogy a fenti leképezés P lapjaihoz P' lapjait rendeli, és szuperjektív. Injektivitása abból következik, hogy ha $F_1 \cap L^\perp = F_2 \cap L^\perp$, akkor

$$F_1 = L + (F_1 \cap L^\perp) = L + (F_2 \cap L^\perp) = F_2.$$

Az, hogy a leképezés tartalmazástartó, nyilvánvaló.

Másodszor P' egyenesmentes poliéder, így lapjai is azok, speciálisan van extremális pontjuk (az első Klee-tételnek, vagy Caratheodory tételének és 7.13-nak a következménye), melyek P' extremális pontjai is egyben. Ezért P' minimális lapjai az extremális pontjai, vagyis P minimális lapjai ezek inverz képei. □

7.16. Következmény: *Legyen $P \subseteq \mathcal{R}^n$ nemüres poliéder, az $Ax \geq b$ rendszer megoldáshalmaza, ahol $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$. Ekkor a P poliéder minimális lapjainak dimenziója $n - \mathrm{r}\,(A)$.* □

Legyen $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$. Az $Ax \geq b$ lineáris egyenlőtlenségrendszert **irredundáns**nak hívjuk, ha egyetlen egyenlőtlensége sem logikai következménye a többinek, vagyis minden i sorindex esetén létezik $x_i \in \mathcal{R}^n$ vektor, amelyre a rendszer minden egyenlőtlensége teljesül, kivéve az i-ediket. Az $Ax \geq b$ rendszert **redundáns**nak hívjuk, ha nem irredundáns. Redundáns rendszer például az $x \geq 0$, $x \leq 1$, $x \leq 1$ rendszer. A redundáns egyenlőtlenségeket (a többi egyenlőtlenség logikai következményeit) egyesével elhagyva az eredeti rendszerrel azonos megoldáshalmazú, irredundáns rendszerhez juthatunk.

A következő állításban 7.3 jelöléseit használjuk.

7.17. Állítás: *Tegyük fel, hogy az $Ax \geq b$ rendszer irredundáns. Ekkor a $P = \{x : Ax \geq b\}$ poliéder maximális valódi lapjai éppen az*

$$\{x \in P : ax = \beta\}$$

alakú lapok, ahol az a vektor az $A_P^{\gtrless}$ mátrix sorvektora, a β szám pedig a $b_P^{\gtrless}$ vektor megfelelő eleme.

Bizonyítás: Legyen $\emptyset \subset F \subset P$ maximális valódi lap, ekkor 7.3 szerint $F = \{x \in P : A_F^= x = b_F^=\}$. Legyen $x_0 \in P \setminus F$, ekkor $ax_0 > \beta$ valamely $A_P^>$ és $A_F^=$ mátrixbeli a sorvektorral, és megfelelő $b_P^>$ és $b_F^=$ vektorbeli β elemmel. Legyen $\hat{F} := \{x \in P : ax = \beta\}$, ekkor $\hat{F} \subset P$ (az x_0 pont mutatja), és $\hat{F}$ a P poliéder az F lapot tartalmazó (exponált) lapja, amiből F maximalitása miatt $F = \hat{F}$ következik.

A fordított irányhoz elég megmutatnunk, hogy ha F az állításban leírt alakú lap, akkor $F \neq \emptyset$, és hogy $A_F^=$ nem más, mint az $A_P^=$ mátrix, megtoldva még az a sorvektorral. Ehhez azt kell igazolnunk, hogy létezik $x_0 \in F$ vektor úgy, hogy $A'x_0 > b'$, ahol A' azt a mátrixot jelöli, amelyet akkor kapunk, ha az $A_P^>$ mátrixból elhagyjuk az a sorvektorát, b' pedig a $b_P^>$ vektor megfelelő része. Legyen $x_1 \in \operatorname{ri} P$, ekkor 7.3 szerint $A_P^> x_1 > b_P^>$. Mivel az $ax \geq \beta$ egyenlőtlenség nem redundáns, azért létezik x_2 vektor úgy, hogy

$$A_P^= x_2 \geq b_P^=,\ A'x_2 \geq b',\ ax_2 < \beta.$$

Ekkor az $]x_1, x_2[$ szakaszon találunk megfelelő x_0 pontot. □

7.18. Következmény: *A* $P \subseteq \mathcal{R}^n$ *poliéder minden maximális valódi lapjának dimenziója* $\dim(P) - 1$.

Bizonyítás: Legyen $Ax \geq b$ a P poliéder irredundáns leírása. Ha F a P poliéder maximális valódi lapja, akkor az előző állítás jelöléseivel $F = \{x \in P : ax = \beta\}$. 7.3 szerint, figyelembe véve még a 7.17 bizonyításában látottakat

$$\operatorname{aff} F = \{x : A_P^= x = b_P^=,\ ax = \beta\},\ \text{és}\ \operatorname{aff} P = \{x : A_P^= x = b_P^=\}.$$

Az állítás ezek után 2.24 következménye, mivel az a vektor nem állhat elő az $A_P^=$ mátrix sorvektorainak lineáris kombinációjaként ($F \neq P$). □

Az extremális pont definíciója "csak" egy kvantorban különbözik a relatív belső pont definíciójától: míg ez utóbbi azt követeli, hogy minden tőle különböző pont túlhúzható legyen az alaphalmazban a ponton, addig az extremális pont definíciójában éppen ellenkezőleg: egyetlen tőle különböző pontot se lehessen túlhúzni az alaphalmazban a ponton. Ennek a különbözőségnek köszönhetően az extremális részhalmazokra bizonyítható felcserélhetőségi tételek is más jellegűek lesznek, mint a relatív belső esetében. E tételek bizonyítására térünk rá a lap- és exponáltlap-burkokkal kapcsolatos néhány alapvető észrevétel után.

7.19. Állítás: *Legyen F_i $(i \in I)$ a $C \subseteq \mathcal{R}^d$ konvex halmaz lapjainak [exponált lapjainak] egy tetszőleges rendszere. Ekkor $\cap_{i \in I} F_i$ is a C halmaz lapja [exponált lapja] lesz.*

Bizonyítás: A lapokra vonatkozó állítás triviális. Legyenek most az F_i halmazok C exponált lapjai, méghozzá

$$F_i = \left\{ x \in C : a_i^T x = \beta_i \right\},$$

ahol $a_i \in \mathcal{R}^d$, $\beta_i \in \mathcal{R}$, $a_i^T C \geq \beta_i$ $(i \in I)$. Feltehető, hogy $\cap_{i \in I} F_i \neq \emptyset$ (mivel $\emptyset \lhd^e C$), ekkor $\cap_{i \in I} F_i = \cap_{j \in J} F_j$, valahányszor $J \subseteq I$ olyan indexhalmaz, hogy az $\{a_j : j \in J\}$ halmaz a $\operatorname{lin}\{a_i : i \in I\}$ altér bázisa. Feltehető tehát, hogy I véges indexhalmaz. Jelölje az a_i $(i \in I)$ vektorok átlagát a, és hasonlóan a β_i $(i \in I)$ számok átlagát β. Könnyen belátható, hogy ekkor $a^T C \geq \beta$, és hogy

$$\cap_{i \in I} F_i = \left\{ x \in C : a^T x = \beta \right\}.$$

$\square$

Ezért tetszőleges $C \subseteq \mathcal{R}^d$ konvex halmaz (univerzum), $S \subseteq C$ esetén definiálható az S halmaz C-beli **lap-**[**exponáltlap-**]**burka**:

$$\Phi_C(S) := \cap\{F : S \subseteq F \in \mathcal{F}(C)\}$$
$$[\Phi_C^e(S) := \cap\{F : S \subseteq F \in \mathcal{EF}(C)\}].$$

7.20. Tétel: *Legyenek $C, C' \subseteq \mathcal{R}^d$ konvex halmazok, $C' \subseteq C$, továbbá $F \lhd C$. Ekkor*
a) $\operatorname{ri} C' \subseteq \operatorname{ri} \Phi_C(C')$ *(speciálisan $x \in C$ esetén $x \in \operatorname{ri} \Phi_C(x)$);*
b) $(\operatorname{ri} C') \cap (\operatorname{ri} F) \neq \emptyset$ *esetén $F = \Phi_C(C')$ (speciálisan ha $x \in \operatorname{ri} F$, akkor $F = \Phi_C(x)$);*
c) lapjainak relatív belsejei a C konvex halmaz egy partícióját alkotják, vagyis diszjunktak, és úniójuk a C halmaz.

Bizonyítás: a) Tegyük fel indirekt, hogy létezik $x \in (\operatorname{ri} C') \cap (\operatorname{rb} \Phi(C'))$ pont. Ekkor 5.4 szerint létezik a $\Phi(C')$ halmaznak egy őt nem tartalmazó H támaszhipersíkja úgy, hogy $x \in H$. Legyen $F := H \cap \Phi(C')$, ez a halmaz exponált részhalmaza, így lapja is $\Phi(C')$-nek, amiből a $\lhd$ reláció tranzitivitása miatt $F \lhd C$ következik. Továbbá $\Phi(C') \not\subseteq H$ miatt $F \subset \Phi(C')$, és $x \in (\operatorname{ri} C') \cap F$, így 7.2 c) miatt $C' \subseteq F$. Összefoglalva $C' \subseteq F \lhd C$, $F \subset \Phi(C')$, ami ellentmondás.

b) a) miatt feltehető, hogy $C' \lhd C$. Az állítás ezek után 7.2 c) egyszerű következménye.

c) a) és b) következménye. $\square$

Speciálisan C pontosan akkor véges lapú, ha véges sok relatív nyílt, konvex halmaz úniója. Ezért ha C véges lapú, akkor AC, $B^{-1}(C)$ is véges lapú; és AC pontosan akkor véges lapú, ha $C + \operatorname{Ker} A$ véges lapú (A, B mátrixok).

Most már rátérhetünk az említett felcserélhetőségi tételekre.

7.21. Tétel: *Legyenek $C_i \subseteq \mathcal{R}^{d_i}$ $(i = 1, \dots, k)$ konvex halmazok. Ekkor*
a) $\mathcal{F}(\times_{i=1}^k C_i) = \{\times_{i=1}^k F_i : F_i \in \mathcal{F}(C_i)\ (i = 1, \dots, k)\}$;
b) $\mathcal{EF}(\times_{i=1}^k C_i) = \{\times_{i=1}^k F_i : F_i \in \mathcal{EF}(C_i)\ (i = 1, \dots, k)\}$.

Bizonyítás: A bizonyításban a $k = 2$ esetre szorítkozunk, az általános k esete csak jelölésben bonyolultabb.

a) A "$\supseteq$" tartalmazás a definíciók könnyű következménye.

A fordított irányú tartalmazáshoz legyen $\emptyset \neq F \lhd C_1 \times C_2$, $x \in \operatorname{ri} F$, továbbá $x_1 \in C_1$, $x_2 \in C_2$, amelyre $x \in \{x_1\} \times \{x_2\}$. Definiáljuk az F_i halmazokat a következőképpen: $F_i := \Phi_{C_i}(x_i)$ $(i = 1, 2)$, ekkor $x_i \in \operatorname{ri} F_i$ (7.20 a)), így

$$x \in (\operatorname{ri} F_1) \times (\operatorname{ri} F_2) = \operatorname{ri} (F_1 \times F_2).$$

Itt $F_1 \times F_2 \lhd C_1 \times C_2$ a már igazolt "$\supseteq$" tartalmazás miatt, és a relatív belsejében tartalmaz egy $\operatorname{ri} F$-beli pontot, következésképpen $F = F_1 \times F_2$ (vö. 7.20 b)).

b) A "$\supseteq$" tartalmazást igazoljuk először. Tegyük fel, hogy $F_i \lhd^e C_i$ $(i = 1, 2)$, ekkor léteznek $a_i \in \mathcal{R}^{d_i}$ vektorok és $\beta_i \in \mathcal{R}$ számok úgy, hogy

$$a_i^T C_i \geq \beta_i, \text{ és } F_i = \left\{x \in C_i : a_i^T x = \beta_i\right\} \ (i = 1, 2).$$

Legyen $a \in \{a_1\} \times \{a_2\}$, $\beta := \beta_1 + \beta_2$. Nyilvánvaló, hogy

$$a^T (C_1 \times C_2) \geq \beta, \text{ és } F_1 \times F_2 = \left\{x \in C_1 \times C_2 : a^T x = \beta\right\},$$

tehát $F_1 \times F_2$ exponált lapja a $C_1 \times C_2$ halmaznak.

A "$\subseteq$" tartalmazáshoz legyen $\emptyset \neq F \lhd^e C_1 \times C_2$, ekkor az állítás a) része szerint léteznek $\emptyset \neq F_i \lhd C_i$ $(i = 1, 2)$ halmazok úgy, hogy $F = F_1 \times F_2$. Azt kell belátnunk, hogy $F_i \lhd^e C_i$ $(i = 1, 2)$. Ez az előző érvelés megfordításával igazolható. Tudjuk, hogy létezik $a \in \mathcal{R}^{d_1} \times \mathcal{R}^{d_2}$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T (C_1 \times C_2) \geq \beta, \text{ és } F = \left\{x \in C_1 \times C_2 : a^T x = \beta\right\}.$$

Legyen $a_i \in \mathcal{R}^{d_i}$ $(i = 1, 2)$, amelyre $a \in \{a_1\} \times \{a_2\}$, valamint $\beta_i := \inf a_i^T C_i$ $(i = 1, 2)$. Nyilván mindegyik β_i véges, különben nem lehetne $a^T(C_1 \times C_2) \geq \beta$. Megmutatjuk, hogy $\beta = \beta_1 + \beta_2$. Egyrészt tetszőleges $f \in F$, $f \in \{f_1\} \times \{f_2\}$ esetén

$$\beta = a^T f = a_1^T f_1 + a_2^T f_2 \geq \beta_1 + \beta_2,$$

másrészt ha $c_i^{(j)}$ olyan C_i-beli pont, amelyre

$$a_i^T c_i^{(j)} < \beta_i + 1/j \ (i = 1, 2; j = 1, 2, \ldots),$$

és $c^{(j)} \in \{c_1^{(j)}\} \times \{c_2^{(j)}\}$ $(j = 1, 2 \ldots)$, akkor

$$\beta \leq a^T c^{(j)} = a_1^T c_1^{(j)} + a_2^T c_2^{(j)} \leq \beta_1 + \beta_2 + \frac{2}{j} \ (j = 1, 2, \ldots),$$

amiből $\beta \leq \beta_1 + \beta_2$ is adódik. Megmutatjuk, hogy

$$a_i^T C_i \geq \beta_i, \text{ és } F_i = \left\{x \in C_i : a_i^T x = \beta_i\right\} \ (i = 1, 2).$$

Az $a_i^T C_i \geq \beta_i$ egyenlőtlenségek a β_i számok definíciójából következnek. Az egyenlőséghez tegyük fel először, hogy valamely $i \in \{1, 2\}$, és $f_i \in F_i$ esetén $a_i^T f_i > \beta_i$ lenne. Legyen $f_j \in F_j$ $(j \neq i)$ tetszőleges elem, továbbá $f \in \{f_1\} \times \{f_2\}$. Ekkor $f \in F$, a

$$\beta = a^T f = a_1^T f_1 + a_2^T f_2 > \beta_1 + \beta_2 = \beta$$

ellentmondáshoz jutottunk. Másfelől ha valamely $i \in \{1, 2\}$ esetén $x \in C_i$, $a_i^T x = \beta_i$, akkor egy tetszőleges $f \in F$ elem i-edik blokkját x-re cserélve a kapott f' pont is F-beli lesz, hiszen $a^T f = a^T f'$ a már bizonyított $F_i \subseteq \{x \in C_i : a_i^T x = \beta_i\}$ tartalmazás miatt. Ezért az f' i-edik blokkja, $x \in F_i$, ami a fordított tartalmazást igazolja. $\square$

7.22. Tétel: *Tetszőleges $C \subseteq \mathcal{R}^d$ konvex halmaz esetén*
a) $\mathcal{F}(K(C)) = \{K(F) : F \in \mathcal{F}(C)\} \cup \{\emptyset\}$;
b) $\mathcal{EF}(K(C)) = \{K(F) : F \in \mathcal{EF}(C)\} \cup \{\emptyset\}$.

Bizonyítás: a) Először a "$\supseteq$" tartalmazást igazoljuk. Mivel $K(\emptyset) = \{0\} \triangleleft K(C)$ (sőt $\{0\} \triangleleft^e K(C)$), azért azt kell megmutatnunk, hogy ha $\emptyset \neq F \triangleleft C$, akkor $K(F) \triangleleft K(C)$. Legyen $\hat{c}_1, \hat{c}_2 \in K(C)$, $\hat{f} \in K(F)$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén $\varepsilon\hat{c}_1 + (1 - \varepsilon)\hat{c}_2 = \hat{f}$. 4.20 triviális első fele szerint léteznek $\lambda_1, \lambda_2, \lambda \geq 0$ számok és $c_1, c_2 \in C$, $f \in F$ vektorok úgy, hogy $\hat{c}_i \in \lambda_i(\{1\} \times \{c_i\})$ $(i = 1, 2)$, és $\hat{f} \in \lambda(\{1\} \times \{f\})$. Ekkor

$$\varepsilon\lambda_1 + (1-\varepsilon)\lambda_2 = \lambda, \text{ és } \varepsilon\lambda_1 c_1 + (1-\varepsilon)\lambda_2 c_2 = \lambda f.$$

Ha $\lambda = 0$, akkor

$$\lambda_1 = \lambda_2 = 0, \text{ és } \hat{c}_1 = \hat{c}_2 = 0 \in K(F).$$

Ha $\lambda > 0$, akkor legalább az egyik a λ_1, λ_2 számok közül pozitív. Ha például $\lambda_1 > 0 = \lambda_2$, akkor $c_1 = f$, amiből $\hat{c}_1 \in K(F)$, továbbá $\hat{c}_2 = 0 \in K(F)$ adódik. Ha λ_1 és λ_2 is pozitív, akkor

$$0 < \varepsilon' := \varepsilon\lambda_1/\lambda < 1, \text{ és } \varepsilon' c_1 + (1-\varepsilon')c_2 = f.$$

Az F extremalitása miatt ebből $c_1, c_2 \in F$, és így $\hat{c}_1, \hat{c}_2 \in K(F)$ következik.

A fordított irányú tartalmazáshoz legyen $\emptyset \neq \hat{F} \triangleleft K(C)$, azt kell megmutatnunk, hogy ekkor $\hat{F} = K(F)$ valamely $F \triangleleft C$ esetén. Legyen $F := C(\hat{F})$, ekkor 3.20 szerint $\hat{F} = K(F)$. Megmutatjuk, hogy $F \triangleleft C$. Legyen $c_1, c_2 \in C$, $f \in F$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén $\varepsilon c_1 + (1-\varepsilon)c_2 = f$. Legyen $\hat{c}_i \in \{1\} \times \{c_i\}$ $(i = 1, 2)$, és $\hat{f} \in \{1\} \times \{f\}$. Ekkor

$$\hat{c}_1, \hat{c}_2 \in K(C), \ \hat{f} \in K(F), \text{ és } \varepsilon\hat{c}_1 + (1-\varepsilon)\hat{c}_2 = \hat{f}.$$

Ebből $\hat{F}$ extremalitása miatt $\hat{c}_1, \hat{c}_2 \in \hat{F}$, vagyis $c_1, c_2 \in F$ adódik.

b) A "$\subseteq$" tartalmazáshoz legyen $\emptyset \neq \hat{F} \triangleleft^e K(C)$, azt kell megmutatnunk, hogy ekkor létezik $F \triangleleft^e C$ úgy, hogy $\hat{F} = K(F)$. Mivel $\emptyset \neq \hat{F} \triangleleft^e K(C)$, azért létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy $\hat{a} \in \{-\beta\} \times \{a\}$ esetén

$$\hat{a}^T K(C) \geq 0, \text{ és } \hat{F} = \left\{\hat{x} \in K(C) : \hat{a}^T \hat{x} = 0\right\}.$$

Legyen $F := C(\hat{F})$. Mint az előbb, most is belátható, hogy $\hat{F} = K(F)$. Megmutatjuk, hogy

$$a^T C \geq \beta, \text{ és } F = \left\{x \in C : a^T x = \beta\right\}.$$

Ebből $a^T C \geq \beta$ nyilvánvaló abból, hogy $\hat{a}^T(\{1\} \times C) \geq 0$. Másrészt $x \in F$ pontosan akkor, ha $\{1\} \times \{x\} \subseteq \hat{F}$ pontosan akkor, ha $x \in C$, $a^T x = \beta$.

A "$\supseteq$" tartalmazáshoz legyen $\emptyset \neq F \triangleleft^e C$, ekkor létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T C \geq \beta, \text{ és } F = \left\{ x \in C : a^T x = \beta \right\}.$$

Ekkor $\hat{a} \in \{-\beta\} \times \{a\}$ esetén $\hat{a}^T K(C) \geq 0$, így

$$\hat{F} := \left\{ \hat{x} \in K(C) : \hat{a}^T \hat{x} = 0 \right\} \triangleleft^e K(C).$$

Megmutatjuk, hogy $\hat{F} = K(F)$. 3.20 miatt elég belátnunk, hogy $F = C(\hat{F})$, és ez nyilvánvaló a definíciókból. □

7.23. Tétel: *Legyenek $C_i \subseteq \mathcal{R}^d$ $(i = 1, \ldots, k)$ konvex halmazok. Ekkor*
a) $\mathcal{F}(\cap_{i=1}^k C_i) = \{\cap_{i=1}^k F_i : F_i \in \mathcal{F}(C_i)\ (i = 1, \ldots, k)\}$;
b) $\mathcal{EF}(\cap_{i=1}^k C_i) \supseteq \{\cap_{i=1}^k F_i : F_i \in \mathcal{EF}(C_i)\ (i = 1, \ldots, k)\}$, egyenlőséggel, ha $\cap_{i=1}^k \mathrm{ri}\, C_i \neq \emptyset$.

Bizonyítás: A bizonyításban a $k = 2$ esetre szorítkozunk, az általános k esete csak jelölésben bonyolultabb.

a) A "$\supseteq$" tartalmazás a definíciók könnyű következménye. A "$\subseteq$" tartalmazáshoz legyen $F \triangleleft C_1 \cap C_2$, $x \in \mathrm{ri}\, F$. Legyen továbbá

$$F_i := \Phi_{C_i}(x) \ (i = 1, 2).$$

Ekkor $F, F_1 \cap F_2 \triangleleft C_1 \cap C_2$ (a már igazolt "$\supseteq$" tartalmazás szerint),

$$x \in \mathrm{ri}\, F, \text{ és } x \in (\mathrm{ri}\, F_1) \cap (\mathrm{ri}\, F_2) = \mathrm{ri}\,(F_1 \cap F_2),$$

tehát

$$F = \Phi_{C_1 \cap C_2}(x) = F_1 \cap F_2.$$

(Itt felhasználtuk 7.20 a)-t, a konvex halmazok metszetének relatív belsejéről szóló 4.16-ot és 7.20 b)-t is.)

b) Tekintsük először a $C_1 = K_1$, $C_2 = K_2$ esetet, ahol $K_1, K_2 \subseteq \mathcal{R}^d$ konvex kúp. A "$\supseteq$" tartalmazáshoz legyen $\emptyset \neq F_i \triangleleft^e K_i$ $(i = 1, 2)$, ekkor létezik $a_i \in K_i^*$ vektor úgy, hogy

$$F_i = \left\{ x \in K_i : a_i^T x = 0 \right\} \ (i = 1, 2).$$

Nyilvánvalóan $a_1 + a_2 \in (K_1 \cap K_2)^*$, és

$$F_1 \cap F_2 = \left\{ x \in K_1 \cap K_2 : (a_1 + a_2)^T x = 0 \right\},$$

így $F_1 \cap F_2 \triangleleft^e K_1 \cap K_2$. A "$\subseteq$" tartalmazáshoz legyen $F' \triangleleft^e K_1 \cap K_2$, ekkor létezik $a \in (K_1 \cap K_2)^*$ vektor úgy, hogy

$$F' = \left\{ x \in K_1 \cap K_2 : a^T x = 0 \right\}.$$

A többkúpos Krein-tétel szerint $a = a_1 + a_2$ valamely $a_1 \in K_1^*$ és $a_2 \in K_2^*$ esetén. Ekkor

$$F_i := \left\{ x \in K_i : a_i^T x = 0 \right\} \triangleleft^e K_i \ (i = 1, 2),$$

és könnyen belátható, hogy $F' = F_1 \cap F_2$.

Az általános eset homogenizációval vezethető vissza a már elintézett esetre. Legyen

$$K_i := K(C_i) \ (i = 1, 2),$$

ekkor $(\operatorname{ri} C_1) \cap (\operatorname{ri} C_2) \neq \emptyset$ esetén $(\operatorname{ri} K_1) \cap (\operatorname{ri} K_2) \neq \emptyset$, vö. 4.20. A már elintézett eset szerint $K_1 \cap K_2$ exponált lapjai éppen az egy K_1-beli és egy K_2-beli exponált lap metszeteként előálló halmazok. 7.22 szerint ebből, figyelembe véve még, hogy a $K(.)$ operáció felcserélhető a metszetképzéssel

$$K(\mathcal{EF}(C_1 \cap C_2)) = \{K(F_1 \cap F_2) : F_i \in \mathcal{EF}(C_i) \ (i = 1, 2)\}$$

adódik, és innen (alkalmazva a $C(.)$ operációt) a kívánt egyenlőség az általános esetben is. $\square$

7.24. Megjegyzés: Definiáljuk a $C_1, C_2 \subseteq \mathcal{R}^2$ halmazokat a következőképpen:

$$C_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 \leq 0,\ x_2 \geq 0 \right\} \cup \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 \geq 0,\ x_2 \geq x_1^2 \right\}$$
$$C_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 \leq 0,\ x_2 \leq 0 \right\} \cup \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 \geq 0,\ x_2 \leq -x_1^2 \right\}.$$

Ekkor a C_1, C_2 halmazok mutatják, hogy az előző tétel b) részében a C_i halmazok relatív belsejeinek diszjunktsága esetén előfordulhat szigorú tartalmazás.

A következő tétel a $K(C)$-nél "$\operatorname{rec} C$-vel nehezebben" kezelhető $\operatorname{cl} K(C)$ lapjairól szól.

7.25. Tétel: *Legyen $C \subseteq \mathcal{R}^d$ nemüres, zárt, konvex halmaz. Ekkor*
a) $\mathcal{F}(\mathrm{cl}\, K(C)) = \{\mathrm{cl}\, K(F) : \emptyset \neq F \in \mathcal{F}(C)\} \cup \{\{0\} \times G : G \in \mathcal{F}(\mathrm{rec}\, C)\}$*;*
b) $\mathcal{EF}(\mathrm{cl}\, K(C)) \subseteq \{\mathrm{cl}\, K(F) : \emptyset \neq F \in \mathcal{EF}(C)\} \cup \{\{0\} \times G : G \in \mathcal{EF}(\mathrm{rec}\, C)\}$*;*
c) $\mathcal{EF}(\mathrm{cl}\, K(C)) \supseteq \{\mathrm{cl}\, K(F) : \emptyset \neq F \in \mathcal{EF}(C)\}$*;*
d) $\mathcal{EF}(\mathrm{cl}\, K(C)) \supseteq \{\{0\} \times G : G \in \mathcal{EF}(\mathrm{rec}\, C)\}$*, ha* $\mathrm{bar}\, C$ *zárt (speciálisan ha C poliéder, kúp vagy kompakt halmaz).*

Bizonyítás: a) A "$\subseteq$" tartalmazáshoz legyen $\hat{F} \lhd \mathrm{cl}\, K(C)$. Először is ha $\hat{F} \subseteq \{0\} \times \mathcal{R}^d$, akkor 7.2 a) és 4.20 szerint

$$\hat{F} \lhd (\{0\} \times \mathcal{R}^d) \cap \mathrm{cl}\, K(C) = \{0\} \times \mathrm{rec}\, C.$$

7.21 szerint ekkor létezik $G \lhd \mathrm{rec}\, C$, amelyre $\hat{F} = \{0\} \times G$. Másodszor ha $\hat{F}$-nak létezik pozitív első koordinátájú pontja, akkor $F := C(\hat{F})$ nemüres, zárt, konvex halmaz, és 3.20 használatával a szokásos módon belátható, hogy $\hat{F} = \mathrm{cl}\, K(F)$, $F \lhd C$.

A "$\supseteq$" tartalmazáshoz legyen először $G \lhd \mathrm{rec}\, C$. Ekkor $\{0\} \times G$ lapja $\{0\} \times \mathrm{rec}\, C$-nek, ami (exponált) lapja $\mathrm{cl}\, K(C)$-nek, így a $\lhd$ reláció tranzitivitása miatt $\{0\} \times G \lhd \mathrm{cl}\, K(C)$. Másodszor legyen $\emptyset \neq F \lhd C$. Megmutatjuk, hogy $\mathrm{cl}\, K(F) \lhd \mathrm{cl}\, K(C)$. Legyen $\hat{c}_1, \hat{c}_2 \in \mathrm{cl}\, K(C)$, $\hat{f} \in \mathrm{cl}\, K(F)$, és tegyük fel, hogy valamely $0 < \varepsilon < 1$ esetén $\varepsilon \hat{c}_1 + (1 - \varepsilon)\hat{c}_2 = \hat{f}$. Azt kell belátnunk, hogy ekkor $\hat{c}_1, \hat{c}_2 \in \mathrm{cl}\, K(F)$. Azt az esetet, mikor $\hat{c}_1, \hat{c}_2 \in K(C)$, $\hat{f} \in K(F)$, már elintéztük 7.22 bizonyításában, ekkor még $\hat{c}_1, \hat{c}_2 \in K(F)$ is fennáll. Feltehető tehát, hogy a $\hat{c}_1, \hat{c}_2, \hat{f}$ vektorok közül legalább az egyiknek az első koordinátája 0 (vö. 4.20). Ha például $\hat{f} \in \{0\} \times \{f\}$, ahol $f \in \mathrm{rec}\, F$, akkor $e_1^T \hat{c}_1 = e_1^T \hat{c}_2 = 0$, vagyis $\hat{c}_i \in \{0\} \times \{c_i\}$ $(i = 1, 2)$ valamely $c_1, c_2 \in \mathrm{rec}\, C$ esetén. Ekkor az $\varepsilon c_1 + (1 - \varepsilon) c_2 = f$ egyenlőségből 7.2 f) segítségével $c_1, c_2 \in \mathrm{rec}\, F$ adódik, tehát $\hat{c}_1, \hat{c}_2 \in \mathrm{cl}\, K(F)$. Ezek szerint feltehető, hogy $\hat{f} \in \lambda(\{1\} \times \{f\})$, $\hat{c}_1 \in \mu(\{1\} \times \{c\})$, $\hat{c}_2 \in \{0\} \times \{z\}$ valamely $\lambda, \mu > 0$, $f \in F$, $c \in C$, $z \in \mathrm{rec}\, C$ esetén. Ekkor

$$\varepsilon\mu = \lambda, \text{ és } \varepsilon\mu c + (1 - \varepsilon)z = \lambda f,$$

amiből

$$c + ((1 - \varepsilon)/\lambda)z = f$$

adódik. Mivel $z \in \mathrm{rec}\, C$, azért $c + 2((1 - \varepsilon)/\lambda)z \in C$. Az

$$f = c + ((1 - \varepsilon)/\lambda)z = (1/2)c + (1/2)(c + 2((1 - \varepsilon)/\lambda)z)$$

egyenlőségből F extremalitása miatt $c \in F$, $c + 2((1-\varepsilon)/\lambda)z \in F$ adódik, és itt kettő helyett akármilyen egynél nagyobb számot írhattunk volna, így $z \in \operatorname{rec} F$ is teljesül. Most is $\hat{c}_1, \hat{c}_2 \in \operatorname{cl} K(F)$.

b) Legyen $\emptyset \neq \hat{F} \lhd^e \operatorname{cl} K(C)$, ekkor a) szerint két lehetőség van: 1) $\hat{F} = \operatorname{cl} K(F)$ valamely $\emptyset \neq F \lhd C$ esetén, vagy 2) $\hat{F} = \{0\} \times G$ valamely $G \lhd \operatorname{rec} C$ esetén. Az alábbiakban külön megvizsgáljuk mindkét esetet.

1) Azt kell megmutatnunk, hogy $F \lhd^e C$. Mivel $\emptyset \neq \hat{F} \lhd^e \operatorname{cl} K(C)$, azért létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy $\hat{a} \in \{-\beta\} \times \{a\}$ esetén

$$\hat{a}^T \operatorname{cl} K(C) \geq 0, \text{ és } \hat{F} = \left\{\hat{x} \in \operatorname{cl} K(C) : \hat{a}^T \hat{x} = 0\right\}.$$

Speciálisan $\hat{a}^T(\{1\} \times C) \geq 0$, amiből $a^T C \geq \beta$ következik. Másrészt

$$\begin{aligned} F = C(\hat{F}) &= \\ &= \left\{x \in \mathcal{R}^d : \{1\} \times \{x\} \subseteq \operatorname{cl} K(C),\, \hat{a}^T(\{1\} \times \{x\}) = 0\right\} = \\ &= \left\{x \in C : a^T x = \beta\right\}, \end{aligned}$$

így valóban $F \lhd^e C$.

2) Azt kell megmutatnunk, hogy $G \lhd^e \operatorname{rec} C$. Mivel $\emptyset \neq \hat{F} \lhd^e \operatorname{cl} K(C)$, azért létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy $\hat{a} \in \{-\beta\} \times \{a\}$ esetén

$$\hat{a}^T \operatorname{cl} K(C) \geq 0, \text{ és } \hat{F} = \left\{\hat{x} \in \operatorname{cl} K(C) : \hat{a}^T \hat{x} = 0\right\}.$$

Speciálisan $\hat{a}^T(\{0\} \times \operatorname{rec} C) \geq 0$, amiből $a^T \operatorname{rec} C \geq 0$ következik. Másrészt

$$\begin{aligned} G &= \left\{x \in \mathcal{R}^d : \{0\} \times \{x\} \subseteq \hat{F}\right\} = \\ &= \left\{x \in \mathcal{R}^d : \{0\} \times \{x\} \subseteq \operatorname{cl} K(C),\, \hat{a}^T(\{0\} \times \{x\}) = 0\right\} = \\ &= \left\{x \in \operatorname{rec} C : a^T x = 0\right\}, \end{aligned}$$

így valóban $G \lhd^e \operatorname{rec} C$.

c) Legyen $\emptyset \neq F \lhd^e C$, továbbá $\hat{F} := \operatorname{cl} K(F)$. Azt kell megmutatnunk, hogy ekkor $\hat{F} \lhd^e \operatorname{cl} K(C)$. Mivel $F \lhd^e C$, azért létezik $a \in \mathcal{R}^d$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T C \geq \beta, \text{ és } F = \left\{x \in C : a^T x = \beta\right\}.$$

Ekkor $\hat{a} \in \{-\beta\} \times \{a\}$ esetén $\hat{a}^T \operatorname{cl} K(C) \geq 0$, ezért

$$\hat{F}' := \left\{\hat{x} \in \operatorname{cl} K(C) : \hat{a}^T \hat{x} = 0\right\} \lhd^e \operatorname{cl} K(C).$$

Az pedig, hogy $\hat{F} = \hat{F}'$, mostanra már rutin alkalmazása 3.20-nak.

d) Legyen $\emptyset \neq G \lhd^e \operatorname{rec} C$, továbbá $\hat{F} := \{0\} \times G$. Azt kell megmutatnunk, hogy ekkor $\hat{F} \lhd^e \operatorname{cl} K(C)$. Mivel $\emptyset \neq G \lhd^e \operatorname{rec} C$, azért létezik $a \in \mathcal{R}^d$ vektor úgy, hogy

$$a^T \operatorname{rec} C \geq 0, \text{ és } G = \left\{ x \in \operatorname{rec} C : a^T x = 0 \right\}.$$

Ekkor 4.11 és a feltétel szerint

$$a \in (\operatorname{rec} C)^* = \operatorname{cl} \operatorname{bar} C = \operatorname{bar} C,$$

létezik tehát $\beta \in \mathcal{R}$ szám úgy, hogy $a^T C \geq \beta$. Ekkor $\hat{a} \in \{-\beta\} \times \{a\}$ esetén $\hat{a}^T \operatorname{cl} K(C) \geq 0$ (4.20), így

$$\hat{F}' := \left\{ \hat{x} \in \operatorname{cl} K(C) : \hat{a}^T \hat{x} = 0 \right\} \lhd^e \operatorname{cl} K(C).$$

Ezért

$$\hat{F} = \hat{F}' \cap \left\{ \hat{x} \in \operatorname{cl} K(C) : e_1^T \hat{x} = 0 \right\}$$

exponált lapok metszeteként maga is exponált lapja $\operatorname{cl} K(C)$-nek (7.19). □

Most már megfogalmazható 7.23 vegyes változata.

7.26. Tétel: *Legyenek $C_i \subseteq \mathcal{R}^d$ $(i = 1, \ldots, k)$ konvex halmazok, $P_j \subseteq \mathcal{R}^d$ $(j = 1, \ldots, l)$ poliéderek. Tegyük fel, hogy $(\cap_j P_j) \cap (\cap_i \operatorname{ri} C_i) \neq \emptyset$. Ekkor*

$$\mathcal{EF}((\cap_i C_i) \cap (\cap_j P_j)) = \left\{ (\cap_i F_i) \cap (\cap_j G_j) \;\middle|\; \begin{array}{l} F_i \in \mathcal{EF}(C_i) \ (i = 1, \ldots, k), \\ G_j \in \mathcal{EF}(P_j) \ (j = 1, \ldots, l) \end{array} \right\}.$$

Bizonyítás: A $k = 0$ eset 7.23 bizonyításához hasonlóan intézhető el: Az egyszerűség kedvéért feltesszük, hogy $l = 2$. Ha P_1, P_2 poliéder kúpok, akkor itt a többkúpos Krein-tétel helyett 3.6-ot használjuk, általában az $R_i := \operatorname{cl} K(P_i)$ $(i = 1, 2)$ poliéder kúpokra (4.20) alkalmazzuk a már elintézett esetet.

Ha $k \neq 0$, akkor feltehető, hogy $k = l = 1$, hiszen 7.23 és a már elintézett eset miatt elég a $C = \cap C_i$ és $P = \cap P_j$ halmazokra belátni a tételt. Az az eset, mikor C és P is kúpok, 7.23 bizonyításának megfelelő részéhez hasonlóan intézhető el, csak most a többkúpos Krein-tétel helyett a vegyes Krein-tételt használva. Az általános eset ismét homogenizációval vezethető vissza az előbbi speciális esetre, a $K := K(C)$ és $R := \operatorname{cl} K(P)$ kúpokra alkalmazva azt. A bizonyítás hasonlóan fejezhető be, mint 7.23 esetében, felhasználva 7.25-öt is. □

A felcserélhetőségi tételek alkalmazásaként homogenizációval bevezethetjük a konjugált lap fogalmát, és bizonyíthatjuk a megfelelő dualitástételeket.

Ha $K \subseteq \mathcal{R}^d$ zárt, konvex kúp, és F ennek lapja, akkor az F lap **(konvex kúp) konjugált**ja az

$$F^{\triangle} := \left\{x \in K^* : f^T x = 0 \ (f \in F)\right\}$$

zárt, konvex kúp. Például

$$\emptyset^{\triangle} = K^*, \text{ és } K^{\triangle} = K^* \cap -K^*.$$

Megmutatjuk, hogy az $F^{\triangle}$ konvex kúp a K^* konvex kúp nemüres, exponált lapja. Ha $F = \emptyset$, akkor ez nyilvánvaló. Különben legyen $f_0 \in \operatorname{ri} F$, ekkor persze $f_0 \in K = K^{**}$, és könnyen belátható, hogy

$$F^{\triangle} = \left\{x \in K^* : f_0^T x = 0\right\},$$

vagyis az f_0 vektor exponálja az $F^{\triangle}$ lapot. Definiálható

$$F^{\triangle\triangle} := (F^{\triangle})^{\triangle},$$

amely a $K^{**} = K$ konvex kúp nemüres, exponált lapja lesz. Például

$$\emptyset^{\triangle\triangle} = K \cap -K, \text{ és } K^{\triangle\triangle} = K.$$

7.27. Tétel: *Legyen $K \subseteq \mathcal{R}^d$ zárt, konvex kúp. Pontosan akkor teljesül $F^{\triangle\triangle} = F$, ha az F halmaz a K konvex kúp nemüres, exponált lapja.*

Bizonyítás: A "csak akkor" irányt az imént láttuk be. Az "akkor" irányhoz először vegyük észre, hogy

$$F^{\triangle\triangle} = (F^{\perp} \cap K^*)^{\perp} \cap K \supseteq (\operatorname{lin} F) \cap K = F$$

mindig teljesül. Másrészt ha F nemüres, exponált lapja K-nak, akkor létezik $a \in K^*$ vektor úgy, hogy $F = \{x \in K : a^T x = 0\}$. Azt kell megmutatnunk, hogy ekkor

$$((\{a\}^{\perp} \cap K)^{\perp} \cap K^*)^{\perp} \cap K \subseteq \{a\}^{\perp} \cap K.$$

Ez abból adódik, hogy most is

$$(\{a\}^{\perp} \cap K)^{\perp} \cap K^* \supseteq (\mathrm{lin}\,\{a\}) \cap K^* \ni a.$$

$\square$

Most általánosítjuk (homogenizációval) ezt a fogalmat tetszőleges C $\mathcal{R}^d$-beli, az origót tartalmazó, zárt, konvex univerzum esetére. Ha F a C halmaz nemüres (exponált) lapja, akkor láttuk, hogy $\mathrm{cl}\,K(F)$ a $\mathrm{cl}\,K(C)$ (exponált) lapja, így értelmes a $\mathrm{cl}\,K(C)$ univerzumra vonatkozó (konvex kúp) konjugáltjáról beszélni. Legyen a nemüres $F \triangleleft C$ lap **(konvex) konjugált**ja az

$$F^{\diamondsuit} := C((\mathrm{cl}\,K(F))^{\triangle}) = \left\{x \in C^{\circ} : f^T x = -1 \ (f \in F)\right\}$$

C°-beli, zárt, konvex halmaz, $\emptyset^{\diamondsuit} := C^{\circ}$. Például $C^{\diamondsuit} = \emptyset$, mivel 6.12 és 4.20 szerint

$$\begin{aligned}(\mathrm{cl}\,K(C))^{\triangle} = K(C)^* \cap -K(C)^* = \\ = \mathrm{cl}\,K(C^{\circ}) \cap -\mathrm{cl}\,K(C^{\circ}) = \{0\} \times \mathrm{rec}\,(C^{\circ}).\end{aligned}$$

Megmutatjuk, hogy $F^{\diamondsuit}$ exponált lapja C°-nek. (Ez a konvex kúp konjugáltra vonatkozó hasonló állítás és 7.25 b) következménye is.) Ha $F = \emptyset$, akkor ez nyilvánvaló. Nemüres $F \triangleleft C$ esetén legyen $f_0 \in \mathrm{ri}\,F$, ekkor $f_0 \in C$ miatt $f_0^T C^{\circ} \geq -1$, másrészt könnyen belátható, hogy

$$F^{\diamondsuit} = \left\{x \in C^{\circ} : f_0^T x = -1\right\},$$

vagyis az f_0 vektor exponálja az $F^{\diamondsuit}$ halmazt. (Itt a "$\subseteq$" tartalmazás nyilvánvaló. Másrészt ha x a jobb oldali halmaz eleme, és valamely $f \in F$ pont esetén $f^T x > -1$ lenne, akkor az f-et túlhúzva f_0-on F-ben, olyan $f' \in F$ pontot találnánk, amelyre $f'^T x < -1$ teljesülne, ami ellentmond annak, hogy $x \in C^{\circ}$.)

Ezért definiálható

$$F^{\diamondsuit\diamondsuit} := (F^{\diamondsuit})^{\diamondsuit},$$

amely $C^{\circ\circ} = C$ exponált lapja lesz. A bikonjugált tételhez felhasználjuk az alábbi lemmát:

7.28. Lemma: *Tegyük fel, hogy még* $0 \in \mathrm{int}\,C$ *is teljesül. Ekkor a* C *zárt, konvex halmaz tetszőleges* F *valódi, exponált lapjára*
a) $\mathrm{cl}\,K(F^{\diamondsuit}) = (\mathrm{cl}\,K(F))^{\triangle}$;
b) $(\mathrm{cl}\,K(F^{\diamondsuit}))^{\triangle} = \mathrm{cl}\,K(F)$.

Bizonyítás: Először azt igazoljuk, hogy a)⇒ b). Ha $\emptyset \neq F \lhd^e C$, akkor $\emptyset \neq \mathrm{cl}\, K(F) \lhd^e \mathrm{cl}\, K(C)$, ezért 7.27 szerint $(\mathrm{cl}\, K(F))^{\triangle\triangle} = \mathrm{cl}\, K(F)$, így b) az a)-beli egyenlőséget konjugálva adódik. Elég tehát a)-t belátnunk.

A tétel a) része pedig 3.20 rutin alkalmazásával adódik. Egyrészt

$$\mathrm{cl}\, K(F^{\diamondsuit}) \subseteq \mathrm{cl}\, K(C^{\circ}), \text{ és } (\mathrm{cl}\, K(F))^{\triangle} \subseteq (\mathrm{cl}\, K(C))^{*} = \mathrm{cl}\, K(C^{\circ})$$

(6.12), így

$$e_1^T \mathrm{cl}\, K(F^{\diamondsuit}) \geq 0, \text{ és } e_1^T (\mathrm{cl}\, K(F))^{\triangle} \geq 0.$$

Mindkét kúp zárt, így egyenlőségükhöz elég, hogy ugyanaz a nemüres halmaz a $C(.)$ operációnál vett képük (vö. 3.20). Könnyen belátható, hogy

$$C(\mathrm{cl}\, K(F^{\diamondsuit})) = C((\mathrm{cl}\, K(F))^{\triangle}) = F^{\diamondsuit},$$

így azt kell igazolnunk, hogy ha F a C valódi, exponált lapja, és $0 \in \mathrm{int}\, C$, akkor $F^{\diamondsuit} \neq \emptyset$. Létezik $0 \neq a \in \mathcal{R}^n$ vektor és $\beta \in \mathcal{R}$ szám úgy, hogy

$$a^T C \geq \beta, \text{ és } F = \left\{ x \in C : a^T x = \beta \right\}.$$

Mivel $0 \in \mathrm{int}\, C$, azért létezik (kis) $\lambda > 0$ úgy, hogy $-\lambda a \in C$, ekkor $-\lambda ||a||^2 \geq \beta$, vagyis β negatív. Feltehető, hogy $\beta = -1$, ekkor

$$F = \left\{ x \in C : a^T x = -1 \right\},$$

és

$$F^{\diamondsuit} = \left\{ x \in C^{\circ} : f^T x = -1 \ (f \in F) \right\} \ni a.$$

□

7.29. Tétel: (bikonjugált tétel) *Tegyük fel, hogy $C \subseteq \mathcal{R}^d$ zárt, konvex halmaz, amely a belsejében tartalmazza az origót. Ekkor $F^{\diamondsuit\diamondsuit} = F$ pontosan akkor, ha F exponált lapja C-nek.*

Bizonyítás: Ebből a "csak akkor" irányt már beláttuk. Az "akkor" irányhoz vegyük észre, hogy $\emptyset^{\diamondsuit\diamondsuit} = \emptyset$, és $C^{\diamondsuit\diamondsuit} = C$, így feltehető, hogy F valódi, exponált lapja C-nek. Ekkor 7.28 szerint

$$F^{\diamondsuit\diamondsuit} = C((\mathrm{cl}\, K(F^{\diamondsuit}))^{\triangle}) = C(\mathrm{cl}\, K(F)) = F.$$

□

A bikonjugált tétel és 6.22 következménye az alábbi

7.30. Állítás: *Legyen $C \subseteq \mathcal{R}^d$ az origót belső pontként tartalmazó kompakt, konvex halmaz. Ekkor $\diamondsuit$ tartalmazás-fordító bijekció $\mathcal{EF}(C)$ és $\mathcal{EF}(C^\circ)$ között.* $\square$

Ha $Q \subseteq \mathcal{R}^d$ az origót belső pontként tartalmazó politóp, akkor a Q° politópot a Q politóp **duális politóp**jának nevezzük. A fenti dualitás miatt politópokra vonatkozó bizonyos tételek párba állíthatók, "duálisak", akárcsak a Weyl- és a Minkowski-tétel poliéder kúpok esetében. Például egy ilyen duális tételpár:
– Minden valódi lap az extremális pontjainak konvex burka (7.4 következménye).
– Minden valódi lap az őt tartalmazó maximális valódi lapok metszeteként áll elő (7.17 következménye).

A harmadik fejezetben felmerült az a kérdés, hogy ha egy konvex kúp minden vetülete zárt, akkor poliéder kúp-e. Waksman és Epelman [31] cikkükben bizonyítják, hogy ha K zárt, konvex kúp, amely nem poliéder kúp, akkor létezik $x \in K$ pontja úgy, hogy a $\operatorname{cone}(K - x)$ kúp nem zárt. Ugyanakkor könnyen belátható, hogy

$$\operatorname{cone}(K - x) = K + \operatorname{lin} \Phi_K(x).$$

Ezért a K nempoliéder kúpot a $(\operatorname{lin} \Phi_K(x))^\perp$ altérre vetítve, Abrams tétele miatt, a vetület nem lesz zárt halmaz. Így a fenti kérdésre igenlő választ adhatunk. Tetszőleges konvex halmazra már nem igaz a tétel: a gömb minden vetülete zárt, mégsem poliéder. Ebben az általánosságban a 7.12-ben leírt kritérium alkalmazható.

A fejezet végén kúplineáris programok regularizációjával foglalkozunk.

Az $(A, b, c, K_1, K_2, \inf)$ leírású (P) kúplineáris program **regularizált**ja az $(A, b, c, K_1', K_2', \inf)$ leírású (P') kúplineáris program, ahol a hatodik fejezetbeli jelöléseket használva

$$K_1' := \Phi_{K_1}(A\mathbf{P} - b),\ K_2' := \Phi_{K_2}(\mathbf{P}).$$

A megfelelő duál kúplineáris programokat jelölje (D), illetve (D').

Néhány észrevétel a regularizált programmal kapcsolatban (vö. 7.20):
– $\mathbf{P} = \mathbf{P}'$, és $\mathbf{D} \subseteq \mathbf{D}'$;
– $\mathbf{P_s} \subseteq \operatorname{ri} \mathbf{P} \subseteq \mathbf{P}'_\mathbf{s}$, speciálisan ha (P) megoldható, akkor (P') szigorúan megoldható;
– ha (P) szigorúan megoldható, akkor $K_1 = K_1'$, és $K_2 = K_2'$, vagyis a (P) és (P'), illetve a (D) és (D') programok leírásaikkal együtt megegyeznek.

A fenti észrevételek és az erős dualitási tétel azonnali következménye az alábbi

7.31. Tétel: *Ha a fenti jelölésekkel a* (P) *program megoldható és korlátos, akkor a* (P) *és* (D') *programok optimumértékei megegyeznek, és a* (D') *program optimumértéke felvétetik.* □

Az állítás nem szimmetrikus: abból, hogy a (D') program megoldható és korlátos, még nem következik, hogy $v_P = v_{D'}$, és $\mathbf{P_o} \neq \emptyset$. Legyen ugyanis (P) és (D) az alábbi kúplineáris programpár:

$$\begin{array}{ll} (P): & \inf e_1^T x,\ e_2^T x = \sqrt{2},\ x \in K, \\ (D): & \sup \sqrt{2} y,\ e_2 y \leq_{K^*} e_1,\ y \in \mathcal{R}, \end{array}$$

ahol $K \subseteq \mathcal{R}^3$ ugyanaz a konvex kúp, mint 3.31-ben. Ekkor $\operatorname{int} K \neq \emptyset$, $\sqrt{2}(1,1,1)^T \in \mathbf{P_s}$, ezért $(P) = (P')$, és $(D) = (D')$. Ugyanakkor $\mathbf{P_o} = \emptyset$, és $0 \in \mathbf{D}$. Az erős dualitási tétel szerint $v_P = v_D \in \mathcal{R}$, mégis $\mathbf{P_o} = \emptyset$.

Operátorok a konvex analízisben

8. Zárt/konvex függvények

A következőkben szereplő függvények többnyire vagy egy adott $S \subseteq \mathcal{R}^n$ halmazon értelmezett, valós értékű; vagy az egész $\mathcal{R}^n$ téren értelmezett, azt a bővített számegyenesbe képező függvények lesznek. Ezért ebben a fejezetben az alapvető definíciókat (alulról félig folytonosság, konvexitás) e két típus közös általánosítására; egy $S \subseteq \mathcal{R}^n$ halmazon értelmezett, azt a bővített számegyenesbe képező függvényekre mondjuk ki.

Itt a bővített számegyenes az

$$\overline{\mathcal{R}} := \mathcal{R} \cup \{-\infty, +\infty\}$$

halmaz. Az összeadást úgy értelmezzük, hogy érvényben maradjon a sorozatok limeszének összege a sorozatok összegének limesze tétel, vagyis például $\alpha + (+\infty) = +\infty$ ($\alpha \in \mathcal{R}$), míg $\infty - \infty$, $-\infty + \infty$ nem értelmes. Hasonlóan értelmezzük a bővített számegyenesen a valós skalárral való szorzást is, azonban $0 \cdot (\pm\infty) = 0$. A rendezésről: természetesen $-\infty < \mathcal{R} < +\infty$.

Adott $S \subseteq \mathcal{R}^n$ halmaz esetén az $f : S \to \overline{\mathcal{R}}$ függvényt egyértelműen meghatározza **epigráf**ja [**hipográf**ja],

$$\operatorname{epi} f := \cup \{\{x\} \times \{\mu\} : x \in S,\, \mu \in \mathcal{R},\, f(x) \leq \mu\},$$
$$[\operatorname{hipo} f := \cup \{\{x\} \times \{\mu\} : x \in S,\, \mu \in \mathcal{R},\, f(x) \geq \mu\}],$$

vagy **alsó [felső] nívóhalmaz**ai, az

$$\{x \in S : f(x) \leq \alpha\} \ (\alpha \in \mathcal{R})$$
$$[\{x \in S : f(x) \geq \alpha\} \ (\alpha \in \mathcal{R})]$$

halmazok. Sokszor ennek az észrevételnek segítségével vezethetünk vissza függvényekről szóló állításokat halmazokról szóló állításokra. Ezért fontos annak vizsgálata, hogy az epigráf, illetve a nívóhalmazok mikor rendelkeznek olyan, az előző fejezetekből már jól ismert halmaztulajdonságokkal, mint például a zártság (8.8) vagy a konvexitás (8.12, 8.13).

Először a zártság kérdésével foglalkozunk. Az erről szóló 8.8 tétel előtt elevenítsünk fel néhány fogalmat a valós analízisből.

Legyen $\alpha_k \in \overline{\mathcal{R}}$ $(k = 1, 2, \ldots)$ egy számsorozat. Az α_k sorozat **alsó határértéke (limesz inferiorja)**

$$\liminf_{k\to\infty} \alpha_k := \sup\{\inf\{\alpha_k : k \geq l\} : l \in \mathcal{N}\},$$

és hasonlóan az α_k sorozat **felső határértéke (limesz szuperiorja)**

$$\limsup_{k\to\infty} \alpha_k := \inf\{\sup\{\alpha_k : k \geq l\} : l \in \mathcal{N}\}.$$

Könnyen belátható, hogy $\alpha = \liminf_{k\to\infty} \alpha_k$ $[\alpha = \limsup_{k\to\infty} \alpha_k]$ pontosan akkor, ha $\beta, \gamma \in \mathcal{R}$, $\beta < \alpha < \gamma$ esetén a sorozat véges sok eleme kisebb [nagyobb], mint β $[\gamma]$, és végtelen sok eleme kisebb [nagyobb], mint γ $[\beta]$.

Ha $\beta, \gamma \in \mathcal{R}$, $\beta < \alpha < \gamma$ esetén az α_k sorozat véges sok eleme kisebb, mint β, és véges sok eleme nagyobb, mint γ, akkor azt mondjuk, hogy az $\alpha \in \overline{\mathcal{R}}$ szám az α_k sorozat **határérték**e, jele $\alpha = \lim_{k\to\infty} \alpha_k$.

Azt mondjuk, hogy az α_k számsorozat **konvergens**, ha létezik véges $\alpha = \lim_{k\to\infty} \alpha_k$ határértéke.

8.1. Állítás: *Legyen*

$$\beta_l := \inf\{\alpha_k : k \geq l\}, \ \gamma_l := \sup\{\alpha_k : k \geq l\} \ (l \in \mathcal{N}).$$

Ekkor

$$\lim_{l\to\infty} \beta_l = \liminf_{k\to\infty} \alpha_k, \ \lim_{l\to\infty} \gamma_l = \limsup_{k\to\infty} \alpha_k.$$

Bizonyítás: Például β_l $(l = 1, 2, \ldots)$ monoton növő sorozat, így limesze elemeinek szuprémuma, $\liminf \alpha_k$. □

8.2. Állítás: *Mindig teljesül, hogy* $\liminf \alpha_k \leq \limsup \alpha_k$. *Itt pontosan akkor van egyenlőség, ha létezik* $\lim_{k\to\infty} \alpha_k$, *és akkor*

$$\lim_{k\to\infty} \alpha_k = \liminf_{k\to\infty} \alpha_k = \limsup_{k\to\infty} \alpha_k.$$

Bizonyítás: Mivel $\beta_l \le \gamma_l$, azért e sorozatok limeszeire is fennáll az egyenlőtlenség.

Ha $\liminf \alpha_k = \limsup \alpha_k$, akkor ezt a közös értéket α-val jelölve, tetszőleges $\beta, \gamma \in \mathcal{R}$, $\beta < \alpha < \gamma$ esetén, elég nagy l-ekre

$$\beta < \beta_l \le \gamma_l < \gamma,$$

így $\beta_l \le \alpha_l \le \gamma_l$ miatt a közös érték $\alpha = \lim \alpha_k$.

A fordított irányhoz vegyük észre, hogy ha létezik $\lim \alpha_k$, akkor $\beta, \gamma \in \mathcal{R}$, $\beta < \lim \alpha_k < \gamma$ esetén létezik $l_0 \in \mathcal{N}$ index úgy, hogy $\beta < \alpha_k < \gamma$ valahányszor $k \ge l_0$. Ekkor $l \ge l_0$ esetén

$$\beta \le \beta_l \le \gamma_l \le \gamma.$$

A $\liminf, \limsup$ definíciója szerint ebből

$$\beta \le \liminf \alpha_k \le \limsup \alpha_k \le \gamma$$

következik, és így $\liminf \alpha_k = \limsup \alpha_k$. $\square$

8.3. Állítás: *A* $\liminf \alpha_k$ [$\limsup \alpha_k$] *érték a legkisebb [legnagyobb] olyan* $\alpha \in \overline{\mathcal{R}}$ *szám, amelyhez létezik az* α_k $(k = 1, 2, \ldots)$ *sorozatnak* α_{k_i} $(i = 1, 2, \ldots)$ *részsorozata úgy, hogy* $\alpha = \lim_{i\to\infty} \alpha_{k_i}$.

Bizonyítás: Legyen α_{k_i} az α_k sorozat létező limeszű részsorozata. Ekkor $k_i \ge i$ miatt

$$\inf\{\alpha_{k_i} : i \ge l\} \ge \inf\{\alpha_k : k \ge l\} \ (l = 1, 2, \ldots),$$

amiből limeszeket véve

$$\lim_{i\to\infty} \alpha_{k_i} \ge \liminf_{k\to\infty} \alpha_k$$

következik.

Azt kell még megmutatnunk, hogy az α_k sorozatnak van a $\liminf \alpha_k$ értékhez konvergáló részsorozata. Válasszuk meg k_1-et úgy, hogy

$$k_1 \ge 1, \ \beta_1 \le \alpha_{k_1} \le \beta_1 + 1$$

legyen, majd k_2-t úgy, hogy

$$k_2 \geq k_1 + 1, \; \beta_{k_1+1} \leq \alpha_{k_2} \leq \beta_{k_1+1} + \frac{1}{2}$$

legyen, és így tovább, általában k_i-t úgy, hogy

$$k_i \geq k_{i-1} + 1, \; \beta_{k_{i-1}+1} \leq \alpha_{k_i} \leq \beta_{k_{i-1}+1} + \frac{1}{i}$$

legyen. A két szélső sorozattal együtt az α_{k_i} részsorozat is a $\liminf \alpha_k$ értékhez fog tartani.

A $\limsup \alpha_k$-ra vonatkozó állítás hasonlóan bizonyítható. □

Analóg eredmények bizonyíthatók az alsó [felső] burkolókról. Az analógia okát [7] írja le.

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, $f : S \to \overline{\mathcal{R}}$, továbbá $S' \subseteq S$, $x \in \operatorname{cl} S'$. Az f függvény **alsó [felső] burkoló**ja az x pontban az S' halmazra megszorítva

$$\begin{array}{l} \underline{f}_{S'}(x) := \sup\{\inf f(O(x,\delta) \cap S') : \delta > 0\} \\ [\overline{f}_{S'}(x) := \inf\{\sup f(O(x,\delta) \cap S') : \delta > 0\}]. \end{array}$$

Könnyen belátható, hogy $\alpha = \underline{f}_{S'}(x)$ $[\alpha = \overline{f}_{S'}(x)]$ pontosan akkor, ha $\beta, \gamma \in \mathcal{R}$, $\beta < \alpha < \gamma$ esetén

$$\begin{array}{l} \inf\{\delta > 0 : \inf f(O(x,\delta) \cap S') < \gamma\} = 0, \\ \inf\{\delta > 0 : \inf f(O(x,\delta) \cap S') < \beta\} > 0 \\ [\inf\{\delta > 0 : \sup f(O(x,\delta) \cap S') > \beta\} = 0, \\ \inf\{\delta > 0 : \sup f(O(x,\delta) \cap S') > \gamma\} > 0]. \end{array}$$

Ha $\beta, \gamma \in \mathcal{R}$, $\beta < \alpha < \gamma$ esetén létezik $\delta > 0$ úgy, hogy

$$\beta < \inf f(O(x,\delta) \cap S') \leq \sup f(O(x,\delta) \cap S') < \gamma,$$

akkor az $\alpha \in \overline{\mathcal{R}}$ érték az f függvény S' halmazra megszorított **burkoló**jának x pontban felvett értéke, melynek jele $f_{S'}(x)$.

Azt mondjuk, hogy az f függvény **folytonos** az $x \in S'$ pontban az S' halmazra megszorítva, ha létezik és véges az $f_{S'}(x)$ érték. Ekkor $x \in S'$ miatt $f_{S'}(x) = f(x)$.

8.4. Állítás: *A fenti definíciókkal teljesül, hogy*

$$\underline{f}_{S'}(x) = \lim_{\delta \to 0+} \inf f(O(x,\delta) \cap S') \; [\overline{f}_{S'}(x) = \lim_{\delta \to 0+} \sup f(O(x,\delta) \cap S')].$$

Bizonyítás: Azt kell megmutatnunk, hogy $\beta, \gamma \in \mathcal{R}$, $\beta < \underline{f}_{S'}(x) < \gamma$ esetén létezik $\hat{\delta} > 0$ úgy, hogy $\beta < \inf f(O(x,\delta) \cap S') < \gamma$ valahányszor $0 < \delta < \hat{\delta}$. Ha $\beta < \underline{f}_{S'}(x)$, akkor létezik $\hat{\delta} > 0$, amelyre $\inf f(O(x,\hat{\delta}) \cap S') > \beta$, ami persze a $\hat{\delta}$-nél kisebb pozitív δ-kra is érvényben marad. Ha $\underline{f}_{S'}(x) < \gamma$, akkor persze $\inf f(O(x,\delta) \cap S') < \gamma$ minden $\delta > 0$ esetén. Mindezekből látszik, hogy $\hat{\delta}$ megfelel. $\square$

8.5. Állítás: *Mindig teljesül, hogy $\underline{f}_{S'}(x) \leq \overline{f}_{S'}(x)$. Itt pontosan akkor van egyenlőség, ha létezik $f_{S'}(x)$, és akkor*

$$f_{S'}(x) = \underline{f}_{S'}(x) = \overline{f}_{S'}(x).$$

Bizonyítás: Megfelelő változtatásokkal 8.2 bizonyítása alkalmazható. $\square$

8.6. Állítás: *Az $\underline{f}_{S'}(x)$ $[\overline{f}_{S'}(x)]$ érték a legkisebb [legnagyobb] olyan $\alpha \in \overline{\mathcal{R}}$ szám, amelyhez található*

$$x_k \to x,\ x_k \in S',\ f(x_k) \to \alpha\ (k \to \infty)$$

tulajdonságú pontsorozat.

Bizonyítás: Legyen x_k $(k = 1, 2, \ldots)$ az állításban leírt tulajdonságú sorozat. Ekkor minden $\gamma \in \mathcal{R}, \gamma > \alpha$ és minden $\delta > 0$ esetén létezik $l \in \mathcal{N}$ úgy, hogy $f(x_l) < \gamma$, és $x_l \in O(x,\delta) \cap S'$. Ebből $\inf f(O(x,\delta) \cap S') < \gamma$, és így $\underline{f}_{S'}(x) \leq \gamma$ adódik. Mivel ez az egyenlőtlenség teljesül minden $\gamma \in \mathcal{R}$, $\gamma > \alpha$ esetén, azért $\underline{f}_{S'}(x) \leq \alpha$.

Azt kell még megmutatnunk, hogy $\alpha := \underline{f}_{S'}(x)$ esetén is létezik az állításban leírt tulajdonságú x_k $(k = 1, 2, \ldots)$ sorozat. 8.4 szerint tetszőleges $\beta, \gamma \in \mathcal{R}$, $\beta < \underline{f}_{S'}(x) < \gamma$ esetén létezik $\hat{\delta} > 0$ úgy, hogy

$$\beta < \inf f(O(x,\delta) \cap S') < \gamma\ (0 < \delta < \hat{\delta}).$$

Válasszuk a β_k, γ_k értékeket úgy, hogy limeszük $\underline{f}_{S'}(x)$ legyen, és a δ_k számokat úgy, hogy a megfelelő $]0, \hat{\delta}_k[$ intervallum elemei legyenek, és a δ_k számsorozat nullához tartson. Ekkor a fentiek szerint választhatók x_k $(k = 1, 2, \ldots)$ pontok úgy, hogy

$$x_k \in O(x,\delta_k) \cap S',\ \beta_k < f(x_k) < \gamma_k\ (k = 1, 2, \ldots)$$

teljesüljön, és akkor az x_k pontsorozat megfelel. $\square$

A következő tételhez szükségünk lesz még két fogalomra, ezek az f függvény **epigráf**ja [**hipográf**ja] az S' halmazra megszorítva:

$$\operatorname{epi}_{S'} f := \{(x,\mu) \in \mathcal{R}^{n+1} : x \in S',\, \mu \geq f(x)\}$$
$$[\operatorname{hipo}_{S'} f := \{(x,\mu) \in \mathcal{R}^{n+1} : x \in S',\, \mu \leq f(x)\}].$$

Az (x,μ) kicsit pontatlanul az $(x^T,\mu)^T$ vektort jelöli (a későbbiekben is). Itt sem tüntetjük fel az S' halmazt, ha $S' = S$ az egész értelmezési tartomány. Nyilván $\operatorname{hipo}_{S'} f = -\operatorname{epi}_{-S'}(-f)$, ezért például a következő két tételben elég az epigráfra vonatkozó részt igazolni.

8.7. Tétel: *Teljesül, hogy*

$$\operatorname{epi}(\underline{f}_{S'}) = \operatorname{cl}(\operatorname{epi}_{S'} f) \qquad [\operatorname{hipo}(\overline{f}_{S'}) = \operatorname{cl}(\operatorname{hipo}_{S'} f)].$$

Bizonyítás: Legyen $(x,\mu) \in \operatorname{cl}(\operatorname{epi}_{S'} f)$, ekkor létezik $x_k \in S'$ $(k = 1,2,\ldots)$ pontsorozat és $\mu_k \in \mathcal{R}$ $(k = 1,2,\ldots)$ számsorozat úgy, hogy

$$f(x_k) \leq \mu_k,\ x_k \to x,\ \mu_k \to \mu\ (k \to \infty).$$

Eszerint $x \in \operatorname{cl} S'$, és

$$\underline{f}_{S'}(x) \leq \liminf f(x_k) \leq \lim \mu_k = \mu,$$

vagyis $(x,\mu) \in \operatorname{epi}(\underline{f}_{S'})$ is teljesül.

A fordított irányú tartalmazáshoz legyen $(x,\mu) \in \operatorname{epi}(\underline{f}_{S'})$, ekkor létezik x_k $(k = 1,2,\ldots)$ S'-beli pontsorozat úgy, hogy az x_k pontsorozat limesze az x pont, míg az $f(x_k)$ számsorozat limesze az $\underline{f}_{S'}(x) \leq \mu$ szám. Ha $\lim f(x_k) < \mu$, akkor feltehető, hogy $f(x_k) < \mu$ $(k = 1,2,\ldots)$, ezért $(x_k,\mu) \in \operatorname{epi}_{S'} f$ $(k = 1,2,\ldots)$, így limesze, $(x,\mu) \in \operatorname{cl}(\operatorname{epi}_{S'} f)$. Ha $\lim f(x_k) = \mu$, akkor feltehető, hogy $f(x_k) \in \mathcal{R}$ $(k = 1,2,\ldots)$, ezért $(x_k, f(x_k)) \in \operatorname{epi}_{S'} f$ $(k = 1,2,\ldots)$, így limesze, $(x,\mu) \in \operatorname{cl}(\operatorname{epi}_{S'} f)$. $\square$

Ahogy a sorozat konvergenciájának a folytonosság felel meg, úgy felel meg a $\liminf, \limsup$ létezésének a félig folytonosság.

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, $f : S \to \overline{\mathcal{R}}$, továbbá $S' \subseteq S$, $x \in S'$. Az f függvény **alulról [felülről] félig folytonos** (röviden a.f.f., illetve f.f.f.) az x pontban az S' halmazra megszorítva, ha

$$f(x) = \underline{f}_{S'}(x)\ [f(x) = \overline{f}_{S'}(x)].$$

Mivel $x \in S'$ miatt $f(x) \geq \underline{f}_{S'}(x)$ $[f(x) \leq \overline{f}_{S'}(x)]$, azért csak az

$$f(x) \leq \underline{f}_{S'}(x) \; [f(x) \geq \overline{f}_{S'}(x)]$$

egyenlőtlenséget kell ellenőrizni. (Amely egyenlőtlenség teljesül, ha $f(x) = -\infty$ $[f(x) = \infty]$, vagy ha $x \in S'$ izolált pont.)

Az $\underline{f}_{S'}(x)$ $[\overline{f}_{S'}(x)]$ definíciójából látszik, hogy ez még azzal ekvivalens, hogy $\alpha \in \mathcal{R}$, $f(x) > \alpha$ $[f(x) < \alpha]$ esetén létezik $\delta > 0$ úgy, hogy

$$f(y) > \alpha \; [f(y) < \alpha] \; (y \in O(x, \delta) \cap S').$$

8.6-ból következik, hogy a definíció ekvivalens azzal, hogy minden, az x ponthoz konvergáló, S'-beli x_k $(k = 1, 2, \ldots)$ pontsorozatra

$$f(x) \leq \liminf_{k \to \infty} f(x_k) \; [f(x) \geq \limsup_{k \to \infty} f(x_k)].$$

Nyilván az f függvény pontosan akkor alulról félig folytonos az x pontban az S' halmazra megszorítva, ha a $-f$ függvény felülről félig folytonos az x pontban az S' halmazra megszorítva. (Mivel $\overline{f}_{S'}(x) = -(\underline{-f})_{S'}(x)$.)

Látjuk, hogy az f függvény pontosan akkor folytonos az $x \in S'$ pontban az S' halmazra megszorítva, ha alulról és felülről is félig folytonos az x pontban az S' halmazra megszorítva, továbbá $f(x)$ véges. Sorozatokkal megfogalmazva ez még azzal is ekvivalens, hogy $f(x)$ véges, és minden x ponthoz konvergáló S'-beli x_k pontsorozatra teljesül, hogy $\lim f(x_k) = f(x)$.

Ha valamelyik folytonossági tulajdonság az S' halmaz minden pontjában teljesül az S' halmazra megszorítva, akkor azt mondjuk, hogy a folytonossági tulajdonság teljesül az S' halmazon. Általában az S' halmazt nem említjük, ha $S' = S$ az egész értelmezési tartomány.

8.8. Tétel: *Tegyük fel, hogy S' zárt halmaz. Ekkor a következő állítások ekvivalensek:*
a) az f függvény a.f.f. [f.f.f.] az S' halmazon;
b) az $\operatorname{epi}_{S'} f$ $[\operatorname{hipo}_{S'} f]$ halmaz zárt;
c) az $\{x \in S' : f(x) \leq \alpha\}$ $[\{x \in S' : f(x) \geq \alpha\}]$ nívóhalmazok zártak minden $\alpha \in \mathcal{R}$ esetén.

Bizonyítás: Először azt igazoljuk, hogy a)$\Rightarrow$b). Legyen $(x, \mu) \in \operatorname{cl} \operatorname{epi}_{S'} f$, ekkor létezik $x_k \in S'$ $(k = 1, 2, \ldots)$ pontsorozat és $\mu_k \in \mathcal{R}$ $(k = 1, 2, \ldots)$

számsorozat úgy, hogy $f(x_k) \leq \mu_k$, és $x_k \to x$ $(k \to \infty)$, $\mu_k \to \mu$ $(k \to \infty)$. Az alulról félig folytonosság miatt

$$\mu = \lim \mu_k \geq \liminf f(x_k) \geq f(x),$$

továbbá S' zártsága miatt $x \in S'$. Ezért $(x, \mu) \in \operatorname{epi}_{S'} f$, amit igazolni kellett.

A b)⇒c) irány abból következik, hogy

$$\{x \in S' : f(x) \leq \alpha\} \times \{\alpha\} = (\operatorname{epi}_{S'} f) \cap (\mathcal{R}^n \times \{\alpha\}),$$

és itt a jobb oldali halmaz a feltétel szerint zárt.

Végül belátjuk, hogy c)⇒a). Tegyük fel indirekt, hogy f nem alulról félig folytonos valamely $x \in S'$ pontban az S' halmazra megszorítva, vagyis $\underline{f}_{S'}(x) < f(x)$. Ekkor létezik olyan, az x ponthoz konvergáló, S'-beli x_k $(k = 1, 2, \ldots)$ pontsorozat, amelyre

$$\lim f(x_k) = \underline{f}_{S'}(x) < f(x).$$

Legyen $\alpha \in \mathcal{R}$ az $]\underline{f}_{S'}(x), f(x)[$ intervallum eleme, ekkor elég nagy k index esetén $f(x_k) \leq \alpha$, így ezek az x_k pontok a megfelelő nívóhalmaz elemei. A nívóhalmaz zártsága miatt az $f(x) \leq \alpha$ egyenlőtlenségnek is teljesülni kellene, holott $f(x) > \alpha$. □

Az előző két tételből az is látszik, hogy

$$\underline{f}_{S'} : \operatorname{cl} S' \to \overline{\mathcal{R}} \ [\overline{f}_{S'} : \operatorname{cl} S' \to \overline{\mathcal{R}}]$$

alulról [felülről] félig folytonos függvény. Mivel nyilvánvalóan az f függvényt az S' halmazon minoráló [majoráló], a $\operatorname{cl} S'$ halmazon értelmezett, alulról [felülről] félig folytonos függvények közül a legnagyobb [legkisebb], azért az S' halmazra megszorított f függvény **alulról [felülről] félig folytonos burká**nak is nevezik. E fogalom kissé módosított formájára is szükségünk lesz a későbbiekben (vö. 9.2).

Az $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény **alsó [felső] lezárt**ja az

$$\underline{f} : \mathcal{R}^n \to \overline{\mathcal{R}} \ [\overline{f} : \mathcal{R}^n \to \overline{\mathcal{R}}]$$

függvény, ha f nem veszi fel a $-\infty$ $[\infty]$ értéket, különben pedig az azonosan

$$-\infty \ [\infty]$$

függvény. Az alsó [felső] lezárt jele $\underline{\operatorname{cl}} f$ $[\overline{\operatorname{cl}} f]$. Az f függvényt **alulról [felülről] zárt**nak nevezzük, ha

$$f = \underline{\operatorname{cl}} f \ [f = \overline{\operatorname{cl}} f].$$

Most rátérünk a konvexitás kérdésére.

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, $f : S \to \overline{\mathcal{R}}$, továbbá $S' \subseteq S$. Ha az $\operatorname{epi}_{S'} f$ halmaz konvex, akkor azt mondjuk, hogy az f függvény **konvex** az S' halmazon. A $g : S \to \overline{\mathcal{R}}$ függvény **konkáv** az S' halmazon, ha $-g$ konvex az S' halmazon. A konvex és konkáv függvények közötti egyszerű kapcsolat miatt a továbbiakban csak a konvex függvényeket vizsgáljuk.

Az $f : S \to \overline{\mathcal{R}}$ függvény **(alsó) effektív tartomány**a a

$$\operatorname{dom} f := \{x \in S : f(x) < \infty\}$$

halmaz, **szigorú epigráf**ja az S' halmazra megszorítva a

$$\operatorname{sepi}_{S'} f := \left\{(x, \mu) \in \mathcal{R}^{n+1} : x \in S',\ f(x) < \mu\right\}$$

halmaz. Néhány egyszerű észrevétel a most bevezetett fogalmakkal kapcsolatban:

– Nyilván

$$\operatorname{sepi}_{S'} f \subseteq \operatorname{epi}_{S'} f \subseteq \operatorname{cl} \operatorname{sepi}_{S'} f,$$

tehát például az epigráf és a szigorú epigráf lezártja megegyezik. (Ha konvexek, akkor a relatív belsejük is.)

– Az $\operatorname{epi}_{S'} f$ és $\operatorname{sepi}_{S'} f$ halmazok vetülete az $\mathcal{R}^n \times \{0\}$ altérre egyaránt $(S' \cap \operatorname{dom} f) \times \{0\}$, speciálisan ha az $\operatorname{epi}_{S'} f$ és $\operatorname{sepi}_{S'} f$ halmazok egyike konvex, akkor az $S' \cap \operatorname{dom} f$ halmaz is konvex.

– Nyilván

$$\operatorname{epi}_{S'} f = \operatorname{epi}_{S' \cap \operatorname{dom} f} f, \text{ és } \operatorname{sepi}_{S'} f = \operatorname{sepi}_{S' \cap \operatorname{dom} f} f,$$

speciálisan az f függvény pontosan akkor konvex az S' halmazon, ha az $S' \cap \operatorname{dom} f$ halmazon konvex.

8.9. Állítás: *Az* $\operatorname{epi}_{S'} f$ *halmaz pontosan akkor konvex, ha a* $\operatorname{sepi}_{S'} f$ *halmaz konvex.*

Bizonyítás: Először belátjuk, hogy ha $\operatorname{epi}_{S'} f$ konvex, akkor $\operatorname{sepi}_{S'} f$ is konvex. Legyen $(x, \mu), (y, \nu) \in \operatorname{sepi}_{S'} f$, továbbá $0 \leq \lambda \leq 1$. Ekkor $x, y \in S' \cap \operatorname{dom} f$, így az $S' \cap \operatorname{dom} f$ halmaz konvexitása miatt $\lambda x + (1-\lambda) y \in S' \cap \operatorname{dom} f$. Legyen $f(x) < \mu' < \mu$, $f(y) < \nu' < \nu$, ekkor $\operatorname{epi}_{S'} f$ konvexitása miatt

$$f(\lambda x + (1-\lambda) y) \leq \lambda \mu' + (1-\lambda) \nu' < \lambda \mu + (1-\lambda) \nu,$$

vagyis $\lambda(x, \mu) + (1-\lambda)(y, \nu) \in \operatorname{sepi}_{S'} f$, amit bizonyítani kellett.

A fordított irányhoz legyen $(x, \mu'), (y, \nu') \in \operatorname{epi}_{S'} f$, továbbá $0 \leq \lambda \leq 1$. Ekkor $x, y \in S' \cap \operatorname{dom} f$, így az $S' \cap \operatorname{dom} f$ halmaz konvexitása miatt $\lambda x + (1-\lambda)y \in S' \cap \operatorname{dom} f$. Legyen $\mu' < \mu \in \mathcal{R}, \nu' < \nu \in \mathcal{R}$, ekkor $\operatorname{sepi}_{S'} f$ konvexitása miatt

$$f(\lambda x + (1-\lambda)y) < \lambda\mu + (1-\lambda)\nu.$$

Tartsunk μ-vel μ'-höz, ν-vel ν'-höz. Ebből

$$f(\lambda x + (1-\lambda)y) \leq \lambda\mu' + (1-\lambda)\nu'$$

adódik, vagyis $\lambda(x, \mu') + (1-\lambda)(y, \nu') \in \operatorname{epi}_{S'} f$, amit bizonyítani kellett.□

8.10. Következmény: *Az $f : S \to \overline{\mathcal{R}}$ függvény pontosan akkor konvex a konvex $C \subseteq S$ halmazon, ha $x, y \in C$, $\mu, \nu \in \mathcal{R}$, $f(x) < \mu$, $f(y) < \nu$, $0 \leq \lambda \leq 1$ esetén*

$$f(\lambda x + (1-\lambda)y) < \lambda\mu + (1-\lambda)\nu.$$

Bizonyítás: Mert az $\operatorname{epi}_C f$ halmaz pontosan akkor konvex, ha a $\operatorname{sepi}_C f$ halmaz konvex. □

8.11. Következmény: *Az $f : S \to]-\infty, \infty]$ függvény pontosan akkor konvex a konvex $C \subseteq S$ halmazon, ha $x, y \in C$, $0 \leq \lambda \leq 1$ esetén*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Bizonyítás: Mert az $\operatorname{epi}_C f$ halmaz éppen akkor konvex, ha az $\operatorname{epi}_{C \cap \operatorname{dom} f} f$ halmaz konvex. □

8.12. Következmény: (Jensen-egyenlőtlenség) *Az $f : S \to]-\infty, \infty]$ függvény pontosan akkor konvex a konvex $C \subseteq S$ halmazon, ha $x_1, \ldots, x_k \in C$, $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$, $\lambda_1 + \ldots + \lambda_k = 1$ esetén*

$$f(\lambda_1 x_1 + \ldots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \ldots + \lambda_k f(x_k).$$

Bizonyítás: Az előző következményből adódik, k szerinti indukcióval. □

A konvexitás további jellemzéseit keresve jutunk a 8.13, 8.14 és 8.17 állításokhoz.

Az alábbi állítás szerint konvex függvény nívóhalmazai is konvexek. Megfordítva nem: egy $f_0 : \mathcal{R} \to \mathcal{R}$ monoton függvény nem feltétlenül konvex, bár nívóhalmazai konvexek.

8.13. Állítás: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \overline{\mathcal{R}}$, továbbá $C \subseteq S$ konvex halmaz. Ha az f függvény konvex a C halmazon, akkor az $\{x \in C : f(x) \leq \alpha\}$ és $\{x \in C : f(x) < \alpha\}$ nívóhalmazok konvexek minden $\alpha \in \mathcal{R}$ esetén.*

Bizonyítás: Mivel

$$\{x \in C : f(x) \leq \alpha\} \times \{\alpha\} = (\operatorname{epi}_C f) \cap (\mathcal{R}^n \times \{\alpha\}),$$

azért $\{x \in C : f(x) \leq \alpha\}$ konvex halmaz. Végül ha $x_1, x_2 \in \{x \in C : f(x) < \alpha\}$, továbbá $0 \leq \lambda \leq 1$, akkor 8.10 miatt

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda\alpha + (1-\lambda)\alpha = \alpha,$$

amiből az $\{x \in C : f(x) < \alpha\}$ nívóhalmazok konvexitása is látszik. □

Az alábbi állítás szerint (amely 8.10 egyszerű következménye) egy függvény konvexitása ekvivalens bizonyos egyváltozós függvények konvexitásával.

8.14. Következmény: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \overline{\mathcal{R}}$, továbbá $C \subseteq S$ konvex halmaz. Jelölje $S_{x,y}$ mindazon $\lambda \in \mathcal{R}$ számok halmazát, amelyre $x + \lambda y \in S$ $(x, y \in \mathcal{R}^n)$, és definiáljuk a $C_{x,y}$ halmazokat hasonlóképpen. Jelölje továbbá $f_{x,y}$ azt az $S_{x,y}$ halmazon értelmezett függvényt, amely a $\lambda \in S_{x,y}$ helyen az $f(x+\lambda y) \in \overline{\mathcal{R}}$ értéket veszi fel. Az f függvény pontosan akkor konvex a C halmazon, ha minden $x, y \in \mathcal{R}^n$ esetén az $f_{x,y}$ függvény konvex a $C_{x,y}$ halmazon.* □

Ezért fontos, hogy egyváltozós függvények konvexitásának további jellemzéseit adjuk.

8.15. Állítás: *Legyen $C \subseteq \mathcal{R}$ konvex halmaz (vagyis intervallum), továbbá $f : C \to \mathcal{R}$. Jelölje g az alábbi függvényt:*

$$g(\lambda', \lambda) := \frac{f(\lambda) - f(\lambda')}{\lambda - \lambda'} \quad (\lambda, \lambda' \in C,\ \lambda' < \lambda).$$

Az f függvény pontosan akkor konvex a C halmazon, ha a g függvény értelmezési tartományából való $(\lambda_1', \lambda_1) \leq (\lambda_2', \lambda_2)$ számpárokra

$$g(\lambda_1', \lambda_1) \leq g(\lambda_2', \lambda_2)$$

teljesül.

Bizonyítás: Az állítás "csak akkor" részét igazoljuk, a gondolatmenet megfordításával bizonyítható az "akkor" irány is.

8.11-et használjuk. Először tekintsük a $\lambda_1' = \lambda_2'$ esetet. Ekkor $\lambda_1 = \lambda_1' + \varepsilon(\lambda_2 - \lambda_1')$, ahol

$$0 < \varepsilon := \frac{\lambda_1 - \lambda_1'}{\lambda_2 - \lambda_1'} \leq 1,$$

ezért $f(\lambda_1) \leq \varepsilon f(\lambda_2) + (1-\varepsilon) f(\lambda_1')$, amiből $g(\lambda_1', \lambda_1) \leq g(\lambda_2', \lambda_2)$ adódik.

A $\lambda_1 = \lambda_2$ eset ebből f-nek az ordinátatengelyre való tükrözésével adódik, az általános eset pedig a tárgyalt esetekből, hiszen ezek szerint

$$g(\lambda_1', \lambda_1) \leq g(\lambda_1', \lambda_2) \leq g(\lambda_2', \lambda_2).$$

□

Minden $C \subseteq \mathcal{R}^n$ konvex halmaz meghatároz egy $\operatorname{ind}_C : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvényt, a C halmaz **indikátorfüggvény**ét, amely a C halmaz pontjain 0 értéket vesz fel, másutt pedig ∞ értéket. Kevésbé nyilvánvalóan minden $C \subseteq \mathcal{R}^{n+1}$ konvex halmaz meghatároz egy $\operatorname{epi}^{-1}(C) : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvényt az alábbi állításban leírt módon:

8.16. Állítás: *Legyen $C \subseteq \mathcal{R}^{n+1}$ tetszőleges konvex halmaz, továbbá*

$$(\operatorname{epi}^{-1}(C))(x) := \inf\{\mu : (x, \mu) \in C\} \ (x \in \mathcal{R}^n).$$

Ekkor $\operatorname{epi}^{-1}(C) : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvény. (Emlékeztetünk rá, hogy $\inf \emptyset = \infty$.) Továbbá $(0,1) \in \operatorname{rec} C$ esetén

$$\operatorname{epi}(\operatorname{epi}^{-1}(C)) = \Big\{(x, \mu) \in \mathcal{R}^{n+1} : (x, \hat{\mu}) \in C \ (\mu < \hat{\mu} < \infty)\Big\}$$

a C-t tartalmazó, $\operatorname{cl} C$-beli halmaz.

Bizonyítás: 8.10 egyszerű következménye. □

Ha $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ tetszőleges függvény, akkor a 8.16-ban leírt módon a $\operatorname{conv}(\operatorname{epi} f)$ konvex halmazból meghatározott függvényt f **konvex burká**nak nevezzük, és $\operatorname{conv} f$-fel jelöljük. Általában adott f_i $(i \in I)$ függvényrendszer esetén $\operatorname{conv}\{f_i : i \in I\}$ jelöli a 8.16 segítségével az $\operatorname{conv}(\cup_{i \in I} \operatorname{epi} f_i)$

konvex halmazból meghatározott függvényt. (Vagyis a $\operatorname{conv}(\inf_{i\in I} f_i(.))$ függvényt.)

8.17. Állítás: *Az $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény pontosan akkor konvex, ha* $\operatorname{conv} f = f$.

Bizonyítás: Az állítás a

$$\operatorname{sepi}\operatorname{conv} f = \operatorname{conv}\operatorname{sepi} f$$

azonosság következménye. (Epigráfokra hasonló azonosság nem áll fenn — gondoljunk a $||.||^{-1}$ függvényre! —, csak

$$\operatorname{conv}\operatorname{epi} f \subseteq \operatorname{epi}\operatorname{conv} f \subseteq \operatorname{cl}\operatorname{conv}\operatorname{epi} f$$

bizonyítható.) □

Következő célunk konvex függvény epigráfja, nívóhalmaza relatív belsejének, lezártjának meghatározása. Mivel tetszőleges $f : S \to \overline{\mathcal{R}}$ konvex függvényt epigráfja megváltoztatása nélkül kiterjeszthetünk az egész $\mathcal{R}^n$ téren értelmezett, konvex függvénnyé, értékeit az S halmazon kívül ∞-nek definiálva, azért a következőkben, az epigráffal kapcsolatos vizsgálódások során, az általánosság csorbítása nélkül csak az egész $\mathcal{R}^n$ téren értelmezett függvényeket vizsgálunk.

8.7 leírja egy tetszőleges $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény epigráfjának lezártját, mint az alulról félig folytonos burok epigráfját. A következő tétel az epigráf relatív belsejéről szól konvex $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény esetén.

8.18. Tétel: *Bármely $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvény esetén*

$$\operatorname{ri}\operatorname{epi} f = \Big\{(x,\mu) \in \mathcal{R}^{n+1} : x \in \operatorname{ri}\operatorname{dom} f,\ \mu > f(x)\Big\}.$$

Bizonyítás: A $\subseteq$ tartalmazás nyilvánvaló: ha (x,μ) a $\operatorname{ri}\operatorname{epi} f$ eleme, de például $x \notin \operatorname{ri}\operatorname{dom} f$, akkor a jobb oldali halmaz egy elemét túlhúzva rajta kilépnénk az $\operatorname{epi} f$ halmazból. Az $x \in \operatorname{ri}\operatorname{dom} f$, $f(x) = \mu$ eset hasonlóan intézhető el, ekkor az $(x, \mu+1)$ pontot túlhúzva az (x,μ) ponton lépünk ki az $\operatorname{epi} f$ halmazból.

A $\supseteq$ tartalmazáshoz legyen $\hat{x} \in \operatorname{ri}\operatorname{dom} f$, és $\infty > \hat{\mu} > f(\hat{x})$. Válasszuk a $\operatorname{dom} f$ halmazból az $\operatorname{aff}\operatorname{dom} f$ halmaz egy affin bázisát, és toljuk el úgy, hogy az $\hat{x}$ pont az affin bázis konvex burkának relatív belsejébe kerüljön (például pontjainak átlaga legyen, vö. 4.2). Ha szükséges, az $\hat{x}$ ponthoz közelebb húzva a pontokat, az aff dom f olyan dom f-beli $\{a_1, \ldots, a_r\}$ affin bázisához jutunk, hogy az általa feszített szimplexet Q-val jelölve $\hat{x} \in \operatorname{ri} Q$. Legyen

$$\infty > \alpha > \max\{f(a_i) : i = 1, \ldots, r\}.$$

Tetszőleges $x \in Q$ esetén x előáll, mint

$$x = \lambda_1 a_1 + \ldots + \lambda_r a_r, \text{ ahol } \lambda_i \geq 0,\ \lambda_1 + \ldots + \lambda_r = 1,$$

és így sepi f konvexitása miatt

$$f(x) < \lambda_1 \alpha + \ldots + \lambda_r \alpha = (\lambda_1 + \ldots + \lambda_r)\alpha = \alpha.$$

Ezért a $(\operatorname{ri} Q) \times]\alpha, \infty[$ relatív nyílt halmaz epi f része, és nyilván megegyezik az affin burkuk. Mivel $(\hat{x}, \alpha+1)$ az előbbi halmaz eleme, azért ri epi f eleme is. Ez a pont túlhúzható az epi f halmazban az $(\hat{x}, \hat{\mu})$ ponton, így az utóbbi pont is ri epi f eleme, amit igazolni kellett. □

8.19. Következmény: *Legyen α valós szám, f konvex függvény úgy, hogy valamely x esetén $f(x) < \alpha$. Ekkor már valamely $x \in \operatorname{ri} \operatorname{dom} f$ esetén $f(x) < \alpha$.*

Bizonyítás: Ha az $\{(x, \mu) : x \in \mathcal{R}^n, \mu < \alpha\}$ nyílt féltér belemetsz a konvex epi f halmazba, akkor a ri epi f halmazba is belemetsz. □

8.18-nak (és a második Hahn–Banach-tételnek) egy további egyszerű következménye, hogy ha $f : C \to \mathcal{R}$ a $C \subseteq \mathcal{R}^n$ konvex halmazon értelmezett függvény, továbbá f és $-f$ is konvex a C halmazon, akkor létezik g affin függvény úgy, hogy az f függvény a g affin függvény megszorítása a C halmazra. Ugyanis könnyen belátható, hogy a diszjunkt relatív belsejű, konvex

$$\operatorname{epi}_C f \text{ és } \begin{pmatrix} E & 0 \\ 0 & -1 \end{pmatrix} \operatorname{epi}_C(-f)$$

halmazokat szeparáló hipersík egy megfelelő affin függvény gráfja lesz.

A következő, a nívóhalmazok relatív belsejét és lezártját leíró 8.20 tétel például a konvex függvények folytonosságának vizsgálatánál, 10.1-ben használható fel.

8.20. Tétel: *Legyen f konvex függvény, továbbá $\alpha \in \mathcal{R}$ olyan szám, amelyre $\alpha > \inf f$. Ekkor az $\{x : f(x) \leq \alpha\}$ és $\{x : f(x) < \alpha\}$ konvex nívóhalmazoknak ugyanaz a lezártjuk és a relatív belsejük, nevezetesen*

$$\{x : \underline{f}(x) \leq \alpha\}, \text{ illetve } \{x \in \mathrm{ri}\,(\mathrm{dom}\, f) : f(x) < \alpha\}.$$

Továbbá ugyanaz a dimenziójuk, mint a $\mathrm{dom}\, f$ *halmaznak.*

Bizonyítás: Legyen M az $\{(x, \alpha) : x \in \mathcal{R}^n\}$ vízszintes hipersík. 8.18 és 8.19 szerint M belemetsz a $\mathrm{ri}\,\mathrm{epi}\, f$ halmazba. 4.16 szerint az

$$M \cap \mathrm{epi}\, f = \{x : f(x) \leq \alpha\} \times \{\alpha\}$$

halmaz lezártja és relatív belseje $M \cap \mathrm{cl}\,(\mathrm{epi}\, f)$ és $M \cap \mathrm{ri}\,(\mathrm{epi}\, f)$. Természetesen $\mathrm{cl}\,(\mathrm{epi}\, f) = \mathrm{epi}\,(\underline{f})$. Ezért

$$\mathrm{cl}\,\{x : f(x) \leq \alpha\} = \{x : \underline{f}(x) \leq \alpha\},$$
$$\mathrm{ri}\,\{x : f(x) \leq \alpha\} = \{x \in \mathrm{ri}\,(\mathrm{dom}\, f) : f(x) < \alpha\}.$$

Az utóbbi egyenlőség szerint

$$\mathrm{ri}\,\{x : f(x) \leq \alpha\} \subseteq \{x : f(x) < \alpha\} \subseteq \{x : f(x) \leq \alpha\},$$

és így az $\{x : f(x) < \alpha\}$ konvex halmaznak ugyanaz a relatív belseje és lezártja, mint az $\{x : f(x) \leq \alpha\}$ konvex halmaznak. Az is látszik, hogy a szereplő nívóhalmazok dimenziója megegyezik. A dimenziókra vonatkozó állításhoz ezek után elég észrevenni, hogy ha $x_0 \in \mathrm{ri}\,(\mathrm{dom}\, f)$, és $f(x_0) < \alpha$, akkor tetszőleges $x_1 \in \mathrm{dom}\, f$ esetén létezik olyan x_1' pont $]x_0, x_1[$-ből (és így $\mathrm{ri}\,\mathrm{dom}\, f$-ből), amelyre $f(x_1') < \alpha$ teljesül (alkalmazzuk 8.19-et az f függvény $[x_0, x_1]$ szakaszra való megszorítására). Ekkor az x_1 pont az x_0 és x_1' pontok által meghatározott egyenes eleme lesz, és máris látszik, hogy a tételben szereplő nívóhalmazok relatív belsejének (és így lezártjának is), valamint a $\mathrm{dom}\, f$ halmaznak ugyanaz az affin burka. □

8.21. Következmény: *Ha f a.f.f., konvex függvény, amelynek effektív tartománya relatív nyílt, akkor* $\inf f < \alpha < \infty$ *esetén*

$$\mathrm{ri}\,\{x : f(x) \leq \alpha\} = \{x : f(x) < \alpha\},$$
$$\mathrm{cl}\,\{x : f(x) < \alpha\} = \{x : f(x) \leq \alpha\}.$$

Bizonyítás: Itt $\underline{f} = f$, és $\mathrm{ri}\,(\mathrm{dom}\, f) = \mathrm{dom}\, f$. □

Következő célunk a negyedik fejezetben konvex halmazok esetében bizonyított felcserélhetőségi tételek megfelelőinek igazolása a konvex függvények általánosságában. Ehhez a cl, majd a rec operátort vizsgáljuk.

Az $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvényt **valódi** konvex függvénynek nevezzük, ha epigráfja nem üres, és nem tartalmaz függőleges egyenest, vagyis az f függvény nem az azonosan ∞ függvény, és sehol sem veszi fel a $-\infty$ értéket. A $g : \mathcal{R}^n \to \overline{\mathcal{R}}$ konkáv függvény **valódi** konkáv függvény, ha $-g$ valódi konvex függvény. Nyilván egy valódi konvex [konkáv] függvény pontosan akkor alulról [felülről] zárt, ha alulról [felülről] félig folytonos, míg egy nem valódi konvex [konkáv] függvény pontosan akkor alulról [felülről] zárt, ha az azonosan ∞ vagy az azonosan $-\infty$ függvény. Konvex függvények esetében mindig az alsó lezártról, konkáv függvények esetében mindig a felső lezártról lesz szó, így az alá-, illetve föléhúzást elhagyjuk a $\underline{\mathrm{cl}}$, illetve $\overline{\mathrm{cl}}$ jelölésből (vö. a 8.19 utáni megjegyzéssel).

8.22. Állítás: *Ha f nemvalódi, konvex függvény, akkor $f(x) = -\infty$ minden $x \in \mathrm{ri}\,\mathrm{dom}\, f$ esetén. Így egy nemvalódi, konvex függvény szükségképpen végtelen értékeket vesz fel, leszámítva esetleg effektív tartománya egyes relatív határpontjait.*

Bizonyítás: Ha $\mathrm{dom}\, f$ nemüres, akkor f nemvalódi volta miatt van olyan x_0 eleme is, ahol $f(x_0) = -\infty$. Legyen x tetszőleges $\mathrm{ri}\,\mathrm{dom}\, f$-beli pont, húzzuk ezen túl az x_0 pontot a $\mathrm{dom}\, f$ halmazban, az így kapott pontot jelölje x_1. Ekkor x az x_0 és x_1 pontok pozitív konvex kombinációja, vagyis létezik $1 > \lambda > 0$ szám úgy, hogy $x = \lambda x_0 + (1-\lambda)x_1$. 8.10 szerint tetszőleges $\alpha > f(x_0)$ és $\beta > f(x_1)$ valós számok esetén

$$f(x) = f(\lambda x_0 + (1-\lambda)x_1) < \lambda\alpha + (1-\lambda)\beta.$$

Mivel $f(x_0) = -\infty$, és $f(x_1) < \infty$, azért szükségképpen $f(x) = -\infty$. □

8.23. Következmény: *Ha f nemvalódi, konvex függvény, amelynek effektív tartománya relatív nyílt, akkor nem vehet fel véges értékeket.* □

8.24. Állítás: *Legyen f valódi konvex függvény. Ekkor* $\mathrm{cl}\, f$ *zárt, valódi konvex függvény. Továbbá* $\mathrm{cl}\, f$ *mindenütt megegyezik az f függvénnyel, leszámítva esetleg* $\mathrm{dom}\, f$ *néhány relatív határpontját.*

Bizonyítás: Mivel $\mathrm{epi}\,(\mathrm{cl}\, f) = \mathrm{cl}\,(\mathrm{epi}\, f)$, és $\mathrm{epi}\, f$ konvex, azért $\mathrm{epi}\,(\mathrm{cl}\, f)$ zárt, konvex halmaz, és $\mathrm{cl}\, f$ alulról félig folytonos függvény. A $\mathrm{cl}\, f$ függvény valódisága, és így zártsága is az állítás hátralévő részének következménye 8.22 szerint, mivel f véges a $\mathrm{dom}\, f$ halmazon. Adott $x \in \mathrm{ri}\,(\mathrm{dom}\, f)$ esetén tekintsük az $M := \{x\} \times \mathcal{R}$ affin halmazt. 8.18 szerint M belemetsz az $\mathrm{epi}\, f$ halmaz relatív belsejébe, így

$$M \cap \mathrm{cl}\,(\mathrm{epi}\, f) = \mathrm{cl}\,(M \cap \mathrm{epi}\, f) = M \cap \mathrm{epi}\, f.$$

Eszerint $(\mathrm{cl}\, f)(x) = f(x)$. Tegyük fel most, hogy $x \notin \mathrm{cl}\,(\mathrm{dom}\, f)$. Mivel $\mathrm{cl}\, f = \underline{f}$, azért

$$\mathrm{cl}\,(\mathrm{dom}\, f) \supseteq \mathrm{dom}\,(\mathrm{cl}\, f) \supseteq \mathrm{dom}\, f,$$

így ekkor $(\mathrm{cl}\, f)(x)$ és $f(x)$ is ∞-nel egyezik meg. □

8.25. Következmény: *Ha f egy valódi konvex függvény, amelynek effektív tartománya affin halmaz, akkor f zárt.*

Bizonyítás: Most $\mathrm{rb}\,\mathrm{dom}\, f = \emptyset$, így $\mathrm{cl}\, f = f$. □

8.26. Tétel: *Legyen f valódi konvex függvény, továbbá $x \in \mathrm{ri}\,(\mathrm{dom}\, f)$. Ekkor*

$$(\mathrm{cl}\, f)(y) = \lim_{\lambda \to 1-} f((1-\lambda)x + \lambda y)$$

minden y esetén.

Bizonyítás: Legyen λ_k $(k = 1, 2, \ldots)$ $]0,1[$-beli számoknak egy 1-hez tartó sorozata, továbbá $x_k := (1-\lambda_k)x + \lambda_k y$. Azt kell megmutatnunk, hogy $(\mathrm{cl}\, f)(y) = \lim f(x_k)$. Mivel $\mathrm{cl}\, f = \underline{f}$ alulról félig folytonos, azért

$$(\mathrm{cl}\, f)(y) \le \liminf f(x_k).$$

Másfelől legyen β olyan szám, amelyre $\beta \ge (\mathrm{cl}\, f)(y)$, továbbá $\alpha > f(x)$. Ekkor

$$(y, \beta) \in \mathrm{epi}\,(\mathrm{cl}\, f) = \mathrm{cl}\,(\mathrm{epi}\, f), \text{ és } (x, \alpha) \in \mathrm{ri}\,(\mathrm{epi}\, f),$$

így

$$(1-\lambda_k)(x, \alpha) + \lambda_k(y, \beta) \in \mathrm{ri}\,(\mathrm{epi}\, f) \ (k = 1, 2, \ldots),$$

amiből

$$f(x_k) < (1-\lambda_k)\alpha + \lambda_k\beta \ (k = 1, 2, \ldots),$$

és így

$$\limsup f(x_k) \le \limsup((1-\lambda_k)\alpha + \lambda_k\beta) = \beta$$

adódik. Eszerint

$$\limsup f(x_k) \le (\mathrm{cl}\, f)(y)$$

is fennáll. □

8.27. Következmény: *Ha f zárt, valódi konvex függvény, akkor*

$$f(y) = \lim_{\lambda \to 1-} f((1-\lambda)x + \lambda y)$$

minden $x \in \mathrm{dom}\, f$ és minden y esetén.

Bizonyítás: Legyen

$$\varphi(\lambda) := f((1-\lambda)x + \lambda y) \ (\lambda \in \mathcal{R}).$$

Ekkor φ valódi konvex függvény a valós számok halmazán, amelyre

$$\varphi(0) = f(x) < \infty, \text{ és } \varphi(1) = f(y).$$

Továbbá φ alulról félig folytonos 8.8 szerint, mivel a $\{\lambda : \varphi(\lambda) \le \alpha\}$ nívóhalmazok a $\{z : f(z) \le \alpha\}$ zárt halmazok inverz képei a

$$\lambda \mapsto (1-\lambda)x + \lambda y = z$$

folytonos leképezésnél. A φ függvény effektív tartománya egy intervallum. Ha ennek az intervallumnak a relatív belseje belemetsz a $]0,1[$ intervallumba, akkor 8.26 szerint

$$\lim_{\lambda \to 1-} \varphi(\lambda) = (\mathrm{cl}\, \varphi)(1) = \varphi(1),$$

különben pedig $\mathrm{dom}\, \varphi \subseteq]-\infty, 0]$, és így

$$\lim_{\lambda \to 1-} \varphi(\lambda) = \infty = \varphi(1).$$

□

Egy $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény **kúpfüggvény**, ha epigráfja kúp, **pozitív homogén** függvény, ha minden $x \in \mathcal{R}^n$ esetén

$$f(\lambda x) = \lambda f(x) \ (0 < \lambda < \infty).$$

Egy kúpfüggvény az origóban nyilván csak 0 vagy $-\infty$ értéket vehet fel. Könnyen belátható az is, hogy a kúpfüggvények éppen azok a pozitív homogén függvények, amelyek az origóban nempozitív értéket vesznek fel.

Továbbá egy nem azonosan ∞, alulról félig folytonos, pozitív homogén függvény kúpfüggvény is. Speciálisan zárt, valódi konvex függvényekre a két tulajdonság megegyezik, és ekkor $f(0) = 0$.

8.28. Állítás: *Ha f valódi konvex kúpfüggvény, akkor $f(0) = 0$, és*

$$f(\lambda_1 x_1 + \ldots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \ldots + \lambda_k f(x_k)$$
$$(\lambda_1 > 0, \ldots, \lambda_k > 0; x_1, \ldots, x_k \in \mathcal{R}^n).$$

Bizonyítás: Feltehető, hogy $x_i \in \operatorname{dom} f$ $(i = 1, \ldots, k)$. Ekkor

$$(x_1, f(x_1)), \ldots, (x_k, f(x_k)) \in \operatorname{epi} f,$$

és így

$$\lambda_1(x_1, f(x_1)) + \ldots + \lambda_k(x_k, f(x_k)) \in \operatorname{epi} f,$$

mivel az $\operatorname{epi} f$ halmaz konvex kúp. Éppen a kívánt egyenlőtlenséghez jutottunk. □

8.29. Következmény: *Ha f valódi konvex kúpfüggvény, akkor $f(-x) \geq -f(x)$ minden x esetén.*

Bizonyítás: 8.28 szerint

$$f(x) + f(-x) \geq f(x - x) = f(0) = 0.$$

□

Legyen $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvény, amely nem azonosan ∞. Ekkor az f függvény epigráfja egy nemüres, konvex halmaz, beszélhetünk a recessziós kúpjáról. Definíció szerint $(y, \nu) \in \operatorname{rec}(\operatorname{epi} f)$ pontosan akkor, ha

$$(x, \mu) + \lambda(y, \nu) = (x + \lambda y, \mu + \lambda\nu) \in \operatorname{epi} f$$

minden $(x, \mu) \in \operatorname{epi} f$, és $\lambda \geq 0$ esetén. Ez azt jelenti, hogy az

$$f(x + \lambda y) \leq f(x) + \lambda\nu$$

egyenlőtlenség fennáll minden x és $\lambda > 0$ esetén. Az y vektort rögzítve a fenti egyenlőséget kielégítő ν számok halmaza alulról zárt, felülről korlátlan intervallum, így $\operatorname{rec}(\operatorname{epi} f)$ valamely konvex kúpfüggvény epigráfja. Ezt a függvényt nevezzük az f függvény **recessziós függvény**ének, jele $\operatorname{rec} f$.

8.30. Tétel: *Legyen f zárt, valódi konvex függvény. Ekkor recessziós függvénye zárt, valódi konvex kúpfüggvény, és tetszőleges $x \in \operatorname{dom} f$ esetén az alábbi képlet határozza meg*

$$(\operatorname{rec} f)(y) = \sup_{\lambda>0} \frac{f(x+\lambda y) - f(x)}{\lambda} = \lim_{\lambda\to\infty} \frac{f(x+\lambda y) - f(x)}{\lambda}.$$

Bizonyítás: Csak a kiemelt egyenlőségeket kell igazolnunk, ezek közül az első a definíció előtti fejtegetés következménye (lásd még 4.9), a második pedig 8.14-é és 8.15-é, melyek szerint a

$$\lambda \mapsto \frac{f(x+\lambda y) - f(x)}{\lambda} \ (\lambda > 0)$$

függvény monoton növekszik. □

Például legyen $A^T A \in \mathcal{R}^{n\times n}$ egy (szimmetrikus, pozitív szemidefinit) mátrix, továbbá $b \in \mathcal{R}^n$, $\gamma \in \mathcal{R}$. Legyen továbbá

$$f(x) := x^T A^T A x + b^T x + \gamma \ (x \in \mathcal{R}^n),$$

ekkor 8.11 segítségével könnyen belátható, hogy f valódi konvex függvény. Nyilván folytonos is, speciálisan zárt. 8.30 szerint, $0 \in \operatorname{dom} f$ miatt

$$(\operatorname{rec} f)(y) = \lim_{\lambda\to 0+} (\lambda^{-1} y^T A^T A y + b^T y + \lambda\gamma) = \begin{cases} b^T y, & \text{ha } Ay = 0, \\ \infty, & \text{egyébként.} \end{cases}$$

Speciálisan ha $A^T A$ nemszinguláris is (azaz ha $\mathrm{r}\,(A) = n$), akkor $\operatorname{rec} f$ a $\{0\}$ halmaz indikátorfüggvénye.

Most már rátérhetünk a felcserélhetőségi tételekre.

Ha f valódi konvex függvény, akkor tetszőleges $\lambda > 0$ esetén a λf függvény is valódi, konvex függvény (8.11), amely zárt is, ha f zárt (8.8 c)). Hasonló állítás mondható ki konvex függvények összegéről is.

8.31. Tétel: *Legyenek $f_1, \ldots, f_k$ valódi konvex függvények, jelölje f az összegüket. Ha minden f_i zárt, és f nem azonosan ∞, akkor f zárt, valódi konvex függvény, és*

$$\operatorname{rec} f = (\operatorname{rec} f_1) + \ldots + (\operatorname{rec} f_k).$$

Ha az f_i függvények nem mind zártak, de a $\mathrm{ri}\,(\mathrm{dom}\, f_i)$ *halmazok összemetszenek, akkor*

$$\mathrm{cl}\, f = (\mathrm{cl}\, f_1) + \ldots + (\mathrm{cl}\, f_k).$$

Bizonyítás: Legyen

$$x \in \mathrm{ri}\,(\mathrm{dom}\, f) = \mathrm{ri}\,(\cap_{i=1}^{k} \mathrm{dom}\, f_i).$$

8.26 szerint minden y esetén

$$(\mathrm{cl}\, f)(y) = \lim_{\lambda \to 1-} f((1-\lambda)x + \lambda y) = \lim_{\lambda \to 1-} \sum_{i=1}^{k} f_i((1-\lambda)x + \lambda y).$$

Ha minden f_i zárt, akkor az utóbbi limesz éppen $f(y)$ 8.27 szerint, így ekkor $\mathrm{cl}\, f = f$. Ha pedig nem minden f_i zárt, de effektív tartományaik relatív belsejei összemetszenek, akkor

$$\mathrm{ri}\,(\cap_{i=1}^{k} \mathrm{dom}\, f_i) = \cap_{i=1}^{k} \mathrm{ri}\,(\mathrm{dom}\, f_i)$$

miatt x mindegyik f_i függvény effektív tartományának relatív belső pontja, és 8.26 szerint az utóbbi limesz $\sum_i \mathrm{cl}\, f_i(y)$, így ekkor $\mathrm{cl}\, f = \sum_i \mathrm{cl}\, f_i$. A recessziós függvényekről szóló egyenlőség a 8.30-beli limeszképlet következménye. □

8.32. Tétel: *Legyen I tetszőleges indexhalmaz, továbbá f_i $(i \in I)$ valódi konvex függvények, és jelölje f az f_i függvények (pontonkénti) szuprémumát. Ha f véges valahol, és minden f_i zárt, akkor f zárt, valódi konvex függvény, és*

$$\mathrm{rec}\, f = \sup\{\mathrm{rec}\, f_i : i \in I\}.$$

Ha nem minden f_i zárt, de létezik a $\cap_{i \in I} \mathrm{ri}\,(\mathrm{dom}\, f_i)$ *halmazban olyan $\hat{x}$ elem, amelyre $f(\hat{x})$ véges, akkor*

$$\mathrm{cl}\, f = \sup\{\mathrm{cl}\, f_i : i \in I\}.$$

Bizonyítás: Mivel $\mathrm{epi}\, f$ az $\mathrm{epi}\, f_i$ halmazok metszete, azért konvex, és zárt is, ha minden f_i zárt. A recessziós függvényekről szóló képlet 4.17 következménye. A lezártakról szóló képlet 4.16 c) és 8.18 következménye: a $\mathrm{ri}\,(\mathrm{epi}\, f_i)$ halmazok metszetében benne lesz az $(\hat{x}, f(\hat{x}) + 1)$ pont. □

A $C \subseteq \mathcal{R}^n$ konvex halmaz **támaszfüggvény**e a

$$\mathrm{supp}_C(x) := \sup\left\{x^T y : y \in C\right\}$$

függvény. 8.32 szerint egy nemüres $C \subseteq \mathcal{R}^n$ konvex halmaz támaszfüggvénye zárt, valódi konvex kúpfüggvény.

8.33. Tétel: *Legyenek $f_1, \dots, f_m$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények, és jelölje f a konvex burkukat. Ekkor*

$$f(x) = \inf\left\{\sum_{i=1}^m \lambda_i f_i(x_i) \;\middle|\; \begin{array}{l} \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1, \\ x_i \in \mathrm{dom}\, f_i,\ \sum_{i=1}^m \lambda_i x_i = x \end{array}\right\} \quad (x \in \mathcal{R}^n).$$

Ha még az is teljesül, hogy az f_i függvények mind zártak, és mindnek ugyanaz az f_0 függvény a recessziós függvénye, akkor f is zárt, valódi konvex függvény, amelynek szintén f_0 a recessziós függvénye. Továbbá a fenti képletben az infimum felvétetik minden $x \in \mathrm{dom}\, f = \mathrm{conv}\,(\cup \mathrm{dom}\, f_i)$ esetén.

Bizonyítás: Definíció szerint $f(x)$ azoknak a μ értékeknek az infimuma, amelyekre $(x, \mu) \in C$, ahol C a $C_i := \mathrm{epi}\, f_i$ konvex halmazok uniójának konvex burka. 4.30 szerint $(x, \mu) \in C$ pontosan akkor, ha (x, μ) kifejezhető az alábbi konvex kombinációként

$$(x, \mu) = \sum_{i=1}^m \lambda_i (x_i, \mu_i) = (\textstyle\sum \lambda_i x_i, \sum \lambda_i \mu_i),$$

ahol $(x_i, \mu_i) \in C_i$. Tehát $f(x)$ a $\sum \lambda_i \mu_i$ értékek infimuma, miközben $x = \sum \lambda_i x_i$ és $\mu_i \geq f_i(x_i)$. Ebből már látszik a tételbeli képlet érvényessége.

Ha most a C_i halmazok zártak is, továbbá recessziós kúpjuk ugyanaz a $K := \mathrm{epi}\, f_0$ zárt, konvex kúp, akkor 4.32 szerint C is zárt, konvex halmaz, amelynek recessziós kúpja szintén a K kúp. Mivel $(0, 1) \in K$, és $(0, -1) \notin K$, azért az is látszik, hogy a C halmaz egy zárt, valódi konvex függvény epigráfja. Ez a függvény nem lehet más, mint f, így $\mathrm{rec}\, f = f_0$. Mivel $x \in \mathrm{dom}\, f$ esetén $\hat{x} := (x, f(x)) \in C$, azért $f(x)$ kifejezhető, mint $\sum \lambda_i \mu_i$ konvex kombináció, ahol $\sum \lambda_i x_i = x$ és $\mu_i \geq f_i(x_i)$ (lásd a bizonyítás első részét), és itt pozitív λ_i esetén μ_i nem lehet nagyobb, mint $f_i(x_i)$. Ezért az infimum felvétetik. $\square$

8.34. Tétel: *Legyen* $A \in \mathcal{R}^{m \times n}$*, és* $g : \mathcal{R}^m \to \overline{\mathcal{R}}$ *valódi konvex függvény úgy, hogy* $g \circ A$ *nem azonosan* ∞*. Ha* g *zárt, akkor* $g \circ A$ *is zárt, valódi konvex függvény, és* $\operatorname{rec}(g \circ A) = (\operatorname{rec} g) \circ A$*. Ha* g *nem zárt, de* $Ax \in \operatorname{ri}(\operatorname{dom} g)$ *valamely* x *esetén, akkor* $\operatorname{cl}(g \circ A) = (\operatorname{cl} g) \circ A$*.*

Bizonyítás: A $g \circ A$ függvény epigráfja az $\operatorname{epi} g$ halmaz $B : (x, \mu) \mapsto (Ax, \mu)$ (folytonos) lineáris transzformációnál vett inverz képe, így konvex, ha g konvex, és zárt is, ha g zárt. A recessziós függvényekről szóló képlet 4.28, míg a lezártakról szóló képlet 4.27 és 8.18 következménye. □

Legyen $h : \mathcal{R}^n \to \overline{\mathcal{R}}$, $A \in \mathcal{R}^{m \times n}$. Jelölje Ah az alábbi képlettel leírt függvényt:

$$(Ah)(y) := \inf\{h(x) : Ax = y\} \ (y \in \mathcal{R}^m),$$

Könnyen belátható, hogy

$$\operatorname{sepi} Ah = B \operatorname{sepi} h, \text{ ahol } B := \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix},$$

ezért h konvexitása esetén Ah is konvex függvény lesz. Epigráfokra általában csak

$$B \operatorname{epi} h \subseteq \operatorname{epi} Ah \subseteq \operatorname{cl}(B \operatorname{epi} h)$$

áll fenn. Az Ah függvény zártságára ad elégséges feltételt az alábbi

8.35. Tétel: *Legyen* $h : \mathcal{R}^n \to \overline{\mathcal{R}}$ *zárt, valódi konvex függvény, továbbá* $A \in \mathcal{R}^{m \times n}$*. Tegyük fel, hogy* $Az \neq 0$ *minden olyan* z *esetén, amelyre* $(\operatorname{rec} h)(z) \leq 0$*, és* $(\operatorname{rec} h)(-z) > 0$*. Ekkor az* Ah *függvény zárt, valódi konvex függvény, és* $\operatorname{rec}(Ah) = A(\operatorname{rec} h)$*. Továbbá az* $(Ah)(y)$ *definíciójában szereplő infimum felvétetik minden* $y \in \operatorname{dom} Ah = A \operatorname{dom} h$ *esetén.*

Bizonyítás: Tekintsük h epigráfját (ez nemüres, zárt, konvex halmaz), és a $B : (x, \mu) \mapsto (Ax, \mu)$ lineáris transzformációt. Az $\operatorname{epi} h$ halmaz recessziós kúpja $\operatorname{epi}(\operatorname{rec} h)$, linealitás tere pedig olyan (z, μ) vektorokból áll, amelyekre $(\operatorname{rec} h)(z) \leq \mu$, és $(\operatorname{rec} h)(-z) \leq -\mu$. Ezért $\operatorname{epi} h$ és B kielégítik 4.24 feltételeit. Oda jutunk, hogy $B(\operatorname{epi} h)$ zárt, konvex halmaz, amelynek recessziós kúpja $B(\operatorname{epi}(\operatorname{rec} h))$. Továbbá

$$B(\operatorname{epi} h) = \operatorname{epi}(Ah), \text{ és } B(\operatorname{epi}(\operatorname{rec} h)) = \operatorname{epi}(A(\operatorname{rec} h)).$$

Látszik, hogy Ah alulról félig folytonos, konvex függvény, és a recessziós függvényekről szóló képlet is. Hogy belássuk az Ah függvény valódiságát, azt kell megmutatnunk, hogy az epi (Ah) halmaz nem tartalmaz függőleges egyenest. A függőleges egyenes jelenlétéből következne, hogy epi (Ah) recessziós kúpja, $B(\mathrm{epi}\,(\mathrm{rec}\,h))$ tartalmazna $(0,\mu)$ alakú vektort, ahol $\mu < 0$. Ekkor az epi $(\mathrm{rec}\,h)$ halmaz tartalmazna (z,μ) alakú vektort, amelyre $Az = 0$, és $\mu < 0$. Erre a z vektorra teljesülne, hogy $(\mathrm{rec}\,h)(z) < 0$, és 8.29 szerint

$$(\mathrm{rec}\,h)(-z) \geq -(\mathrm{rec}\,h)(z) > 0,$$

ellentmondásban a tétel feltételével. Az van hátra, hogy megmutassuk, hogy minden olyan y esetén, amelyre $(Ah)(y) \neq \infty$ (vagyis az $Ax = y$, $x \in \mathrm{dom}\,h$ rendszer megoldható), az $(Ah)(y)$ definíciójában szereplő infimum felvétetik valamely x esetén. Ez pedig világos abból, hogy

$$(y,(Ah)(y)) \in \mathrm{epi}\,Ah = B\mathrm{epi}\,h,$$

és így $(y,(Ah)(y)) = (Ax, h(x))$ valamely x esetén. $\square$

Legyenek $f_1,\ldots,f_k$ konvex függvények az $\mathcal{R}^n$ téren. Az $f_1,\ldots,f_k$ függvények **konvolúció**ja az

$$f(x) := \inf\{f_1(x_1) + \ldots + f_k(x_k) : x_1 + \ldots + x_k = x\} \ (x \in \mathcal{R}^n)$$

képlettel definiált függvény, jele $f = f_1\square\ldots\square f_k$. A könnyen ellenőrizhető

$$\mathrm{sepi}\,(f_1\square f_2) = \mathrm{sepi}\,f_1 + \mathrm{sepi}\,f_2$$

azonosságból adódik, hogy a $\square$ művelet kommutatív, asszociatív és konvexitás-megőrző.

Például ha f az euklideszi norma, és g egy $C \subseteq \mathcal{R}^n$ konvex halmaz indikátorfüggvénye, akkor az

$$(f\square g)(x) = \inf_{y\in C} ||x-y|| \ (x \in \mathcal{R}^n)$$

függvény a C halmaz által megadott **távolságfüggvény**, melynek jele dist_C.

8.36. Következmény: *Legyenek $f_1,\ldots,f_k$ zárt, valódi konvex függvények az $\mathcal{R}^n$ téren, jelölje f a konvolúciójukat. Tegyük fel, hogy $z_1+\ldots+z_k \neq 0$ minden olyan $z_1,\ldots,z_k$ vektor-k-as esetén, amelyre*

$$\begin{aligned}(\mathrm{rec}\,f_1)(z_1) + \ldots + (\mathrm{rec}\,f_k)(z_k) &\leq 0,\\ (\mathrm{rec}\,f_1)(-z_1) + \ldots + (\mathrm{rec}\,f_k)(-z_k) &> 0.\end{aligned}$$

Ekkor f zárt, valódi konvex függvény, és az $f(x)$ definíciójában szereplő infimum felvétetik minden $x \in \sum \operatorname{dom} f_i$ esetén (másutt pedig ∞). Továbbá

$$\operatorname{rec} f = (\operatorname{rec} f_1)\Box \dots \Box(\operatorname{rec} f_k).$$

Bizonyítás: Legyen A az $\mathcal{R}^{kn}$ térről a $\mathcal{R}^n$ térbe vezető "összeadás" lineáris leképezés, vagyis

$$A : (x_1, \dots, x_k) \mapsto x_1 + \dots + x_k \ (x_i \in \mathcal{R}^n),$$

és legyen h az az $\mathcal{R}^{kn}$ téren értelmezett, zárt, valódi konvex függvény, amelyet a

$$h(x_1, \dots, x_k) = f_1(x_1) + \dots + f_k(x_k) \ (x_i \in \mathcal{R}^n)$$

egyenlőség definiál. Alkalmazzuk 8.35-öt erre az A leképezésre és a h függvényre. $\Box$

A poliédrikus függvényekre vonatkozó felcserélhetőségi tételeket egy tételbe gyűjtöttük össze. Az $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ konvex függvény **poliédrikus** függvény, ha epigráfja poliéder.

8.37. Tétel: *Legyenek $f_1, \dots, f_k$ poliédrikus, konvex függvények, ekkor*
a) $f_1 + \dots + f_k$ is poliédrikus függvény;
b) $\max_{1 \le i \le k} f_i$ is poliédrikus függvény.
c) ha még $\operatorname{rec} f_i = f_0$ $(i = 1, \dots, k)$ is teljesül, akkor $\operatorname{conv}\{f_i : 1 \le i \le k\}$ is poliédrikus függvény.

Legyen most $A \in \mathcal{R}^{m \times n}$, továbbá f az $\mathcal{R}^n$ téren, g pedig az $\mathcal{R}^m$ téren értelmezett, poliédrikus, konvex függvény. Ekkor
d) a $g \circ A$ konvex függvény is poliédrikus;
e) az Af konvex függvény is poliédrikus, és a definíciójában szereplő infimum felvétetik, ha véges.

Legyen f_1, f_2 az $\mathcal{R}^n$ téren értelmezett, poliédrikus, konvex függvény. Ekkor
f) az $f_1 \Box f_2$ függvény is poliédrikus konvex függvény, amely ha valódi, akkor a definíciójában szereplő infimum felvétetik minden $x \in \operatorname{dom} f_1 + \operatorname{dom} f_2$ esetén (másutt pedig ∞).

Bizonyítás: a) Elég a $k = 2$ esettel elbánni. Teljesül, hogy $f_i = h_i + \operatorname{ind}_{P_i}$ $(i = 1, 2)$, ahol P_1, P_2 poliéderek, és

$$h_1(x) = \max\left\{a_i^T x - \alpha_i : i = 1, \dots, l_1\right\},$$
$$h_2(x) = \max\left\{b_j^T x - \beta_j : j = 1, \dots, l_2\right\}.$$

Legyen $P := P_1 \cap P_2$, $c_{ij} := a_i + b_j$, és $\gamma_{ij} := \alpha_i + \beta_j$. Ekkor P poliéder, és $f_1 + f_2 = h + \mathrm{ind}_P$, ahol

$$h(x) := \max\left\{c_{ij}^T x - \gamma_{ij} : i = 1, \ldots, l_1;\, j = 1, \ldots, l_2\right\}.$$

Ezért $f_1 + f_2$ poliédrikus függvény.

b) 8.32 bizonyításából látszik, mivel poliéderek metszete is poliéder.

Hasonlóan bizonyítható c) 8.33 és 4.31 segítségével, d) 8.34 és 3.25, e) 8.35 és 3.25, f) 8.36 és 3.25 segítségével. □

A felcserélési tételek azért különlegesen fontosak, mert a bennük szereplő transzformációk megőrzik a függvények zártságát és konvexitását, míg magában a zártságot esetleg nem. Ez a tény Weierstrass alábbi tétele miatt jelentős.

8.38. Tétel: (Weierstrass) *Legyen $S \subseteq \mathcal{R}^n$, $S' \subseteq S$ nemüres halmaz, továbbá $f : S \to]-\infty, \infty]$ az S' halmazon alulról félig folytonos és nem azonosan ∞ függvény. Tekintsük a*

$$(P) : \qquad \inf f(x),\ x \in S'$$

programot. Tegyük fel, hogy az alábbi a), b) és c) feltételek legalább egyike teljesül
a) az S' halmaz kompakt;
b) az S' halmaz zárt, és az $\{x \in S' : f(x) \le f(y)\}$ nívóhalmaz korlátos valamely nem optimális $y \in S'$ pont esetén;
c) az S' halmaz zárt, és $\lim_{k\to\infty} f(x_k) = \infty$ minden olyan S'-beli x_k $(k = 1, 2, \ldots)$ pontsorozatra, amelyre $\lim_{k\to\infty} ||x_k|| = \infty$.

Ekkor a (P) program optimumértéke felvétetik, az optimális megoldások halmaza kompakt, és minden optimalizáló sorozat az optimális megoldások halmazához tart.

Bizonyítás: Tegyük fel először, hogy teljesül az a) feltétel, és legyen x_k $(k=1,2,\ldots)$ optimalizáló sorozat, tehát $x_k \in S'$ $(k=1,2,\ldots)$, és $f(x_k) \to v_P$ $(k \to \infty)$, ahol v_P szokás szerint a (P) program optimumértékét jelöli. Mivel az S' halmaz kompakt, azért az x_k $(k = 1, 2, \ldots)$ sorozatnak létezik konvergens részsorozata, legyen ez éppen x_{k_i} $(i = 1, 2, \ldots)$, limeszpontja pedig $x \in S'$. Ekkor az f függvény alulról félig folytonossága miatt

$$v_P \le f(x) \le \liminf_{i\to\infty} f(x_{k_i}) = \lim_{i\to\infty} f(x_{k_i}) = v_P,$$

amiből már látszik, hogy a (P) program optimumértéke felvétetik (és az is, hogy az optimalizáló sorozatok konvergens részsorozatai optimális megoldáshoz tartanak). Mivel az optimális megoldások halmaza, $\{x \in S' : f(x) \leq v_P\}$ a kompakt S' halmaz f alulról félig folytonossága miatt zárt részhalmaza, azért az optimális megoldások halmaza kompakt. Legyen most x_k $(k = 1, 2, \ldots)$ optimalizáló sorozat, azt kell belátnunk, hogy

$$\lim_{k\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_k) = 0,$$

ahol szokás szerint $\mathbf{P_o}$ a (P) program optimális megoldásainak halmazát jelöli. Először is

$$\mathrm{dist}_{\mathbf{P_o}}(x_k) = \inf_{x\in\mathbf{P_o}} ||x - x_k|| \geq 0,$$

ezért

$$\liminf_{k\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_k) \geq 0.$$

Másrészt legyen x_{k_i} $(i = 1, 2, \ldots)$ az x_k $(k = 1, 2, \ldots)$ sorozat olyan részsorozata, amelyre

$$\limsup_{k\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_k) = \lim_{i\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_{k_i}),$$

az S' halmaz kompaktsága miatt az is feltehető, hogy az x_{k_i} $(i = 1, 2, \ldots)$ sorozat konvergál valamely $x \in S'$ (a fentiek szerint $x \in \mathbf{P_o}$) ponthoz. A távolságfüggvény folytonossága miatt

$$\lim_{i\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_{k_i}) = \mathrm{dist}_{\mathbf{P_o}}(x) = 0,$$

vagyis összefoglalva

$$0 \leq \liminf_{k\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_k) \leq \limsup_{k\to\infty} \mathrm{dist}_{\mathbf{P_o}}(x_k) = 0,$$

az x_k sorozat valóban a $\mathbf{P_o}$ halmazhoz tart.

Tegyük fel most, hogy a b) feltétel teljesül. Mivel nyilván

$$\mathbf{P_o} \subseteq \{x \in S' : f(x) \leq f(y)\},$$

és ez az utóbbi halmaz a feltételek és 8.8 miatt zárt és korlátos, vagyis kompakt, azért feltehető, hogy megegyezik az S' halmazzal. A tétel már igazolt részéből következik, hogy $\mathbf{P_o}$ ekkor is nemüres, kompakt halmaz, amelyhez minden optimalizáló sorozat konvergál.

Végül tegyük fel, hogy a c) feltétel teljesül. Legyen y olyan nem optimális S'-beli pont, amelyre $f(y) < \infty$. Ekkor a c) feltétel miatt az $\{x \in S' : f(x) \leq f(y)\}$ nívóhalmaz nem lehet korlátlan (különben valamely belőle vett sorozat elemeinek normája ∞-hez tartana és akkor az f függvény e sorozat elemein felvett értékei is ∞-hez tartanának, ellentmondásban azzal, hogy mind legfeljebb $f(y)$). Ezért y mutatja, hogy a b) feltétel is fennáll, azt az esetet pedig már elintéztük. $\square$

Természetesen kimondható a Weierstrass-tétel megfelelője felülről félig folytonos függvények maximalizálásáról is.

9. A Fenchel-konjugált

Ebben a fejezetben egy olyan operátort definiálunk és vizsgálunk, amely azt a szerepet játssza majd a Rockafellar-féle dualitáselméletben (11. fejezet), mint a * operáció a kúplineáris programok dualitáselméletében (6. fejezet). A definíció felé vezető első természetes lépésként általánosítjuk 5.14-et, mely szerint egy zárt, konvex halmaz az őt tartalmazó zárt félterek metszete.

Ezt a tételt egy zárt, konvex függvény epigráfjára alkalmazva látjuk, hogy az epigráf előáll, mint az alábbi típusú zárt félterek metszete:
1. $\{(x, \mu) : a^T x \leq \beta\}$ $(a \neq 0)$;
2. $\{(x, \mu) : \mu \geq a^T x - \beta\}$;
3. $\{(x, \mu); \mu \leq a^T x - \beta\}$.

Az ilyen típusú zárt féltereket rendre **függőleges**, **felső**, illetve **alsó** zárt féltereknek nevezzük. Alsó zárt féltér nyilván nem tartalmazhat egy olyan halmazt, mint az epigráf, amelynek $(0, 1)$ recessziós iránya. Ezért az epigráf előáll, mint az őt tartalmazó függőleges és felső zárt félterek metszete. Az alapvető észrevétel az, hogy itt a függőleges zárt félterekre nincs is szükség. Mivel az f függvény epigráfját tartalmazó felső zárt félterek nyilván az f függvényt minoráló h affin függvények epigráfjai, azért a fenti észrevétel még úgy is fogalmazható, hogy

9.1. Tétel: *Egy f zárt, konvex függvény az olyan h affin függvények pontonkénti szuprémuma, amelyekre $h \leq f$.*

Bizonyítás: Feltehető, hogy f valódi konvex függvény, mivel különben nyilvánvaló a tétel (hiszen ekkor f vagy azonosan ∞, vagy azonosan $-\infty$ a zártság definíciója szerint). Ahogy azt a tétel kimondása előtt már kifejtettük, $\operatorname{epi} f$ felső és függőleges zárt félterek metszete. Az epigráf nem lehet csupa függőleges zárt féltér metszete, ez ugyanis ellentmondana f valódiságának. Azt kell megmutatnunk, hogy az epigráfot tartalmazó függőleges és felső zárt félterek metszete megegyezik az epigráfot tartalmazó felső zárt félterek metszetével. Tegyük fel, hogy

$$H^+ = \left\{(x,\mu) : 0 \geq a_1^T x - \beta_1 =: h_1(x)\right\}$$

egy az epigráfot tartalmazó függőleges zárt féltér, és $(x_0, \mu_0) \notin H^+$. Elegendő annyit belátni, hogy létezik h affin függvény, amelyre $h \leq f$, és $\mu_0 < h(x_0)$. Tudjuk, hogy létezik legalább egy h_2 affin függvény úgy, hogy $\operatorname{epi} h_2 \supseteq \operatorname{epi} f$, vagyis $h_2 \leq f$. Tetszőleges $x \in \operatorname{dom} f$ esetén $h_1(x) \leq 0$, és $h_2(x) \leq f(x)$, amiből

$$\lambda h_1(x) + h_2(x) \leq f(x) \ (\lambda \geq 0)$$

adódik. Ugyanez az egyenlőtlenség triviálisan fennáll $x \notin \operatorname{dom} f$ esetén is, mivel akkor $f(x) = \infty$. Ha most a λ számot olyan nagynak választjuk, hogy $\lambda h_1(x_0) + h_2(x_0) > \mu_0$ teljesüljön ($h_1(x_0) > 0$, ezért ez megtehető), akkor a $h := \lambda h_1 + h_2$ affin függvény megfelel. $\square$

Vegyük észre, hogy 9.1 valóban általánosítja 5.14-et: Ha C konvex halmaz, f pedig ennek indikátorfüggvénye, akkor egy $h = a^T. - \beta$ affin függvény pontosan akkor minorálja az f függvényt, ha $h(x) \leq 0$ minden $x \in C$ esetén, vagyis $C \subseteq \{x : a^T x \leq \beta\}$.

Legyen most f tetszőleges az $\mathcal{R}^n$ téren értelmezett, zárt, konvex függvény. 9.1 szerint van egy duális módja is f leírásának: az S halmaz megadása, ahol S mindazon (a, β) párokból áll, hogy a $h := a^T. - \beta$ affin függvény minorálja az f függvényt. Pontosan akkor lesz $h \leq f$, ha

$$\beta \geq \sup\left\{a^T x - f(x) : x \in \mathcal{R}^n\right\}.$$

Ezért S valójában egy az $\mathcal{R}^n$ téren értelmezett f^c függvény epigráfja, ahol

$$f^c(a) := \sup\left\{a^T x - f(x) : x \in \mathcal{R}^n\right\} \ (a \in \mathcal{R}^n).$$

Az f^c függvényt az f függvény **konjugált**jának hívjuk.

Mivel f^c valójában a $g := x^T. - \mu$ $((x,\mu) \in \operatorname{epi} f)$ alakú affin függvények pontonkénti szuprémuma, azért f^c zárt, konvex függvény. Mivel f is előáll, mint a $h := a^T. - \beta$ $((a,\beta) \in S = \operatorname{epi} f^c)$ affin függvények pontonkénti szuprémuma, azért

$$f(x) = \sup\left\{x^T a - f^c(a) : a \in \mathcal{R}^n\right\} \quad (x \in \mathcal{R}^n),$$

vagyis f^c konjugáltja, f^{cc} újra az f függvény.

A ∞ és $-\infty$ konstans függvények nyilván egymás konjugáltjai, így az összes többi (f, f^c) párban szereplő függvények valódiak.

Egy tetszőleges $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény konjugáltját is a fenti képlettel definiálhatjuk. Fennáll a **Fenchel-egyenlőtlenség**, mely szerint

$$a^T x - f(x) \le f^c(a).$$

Nyilván $f_1 \le f_2$ esetén $f_1^c \ge f_2^c$. Továbbá mivel

$$\operatorname{epi} f^c = \left\{(a,\beta) : \operatorname{epi} f \subseteq \{(x,\mu) : a^T x - \beta \le \mu\}\right\},$$

azért láthatóan (vö. 8.7, 8.16)

$$f^c = (\operatorname{cl} f)^c, \text{ és } f^c = (\operatorname{conv} f)^c.$$

A konjugálással kapcsolatos főbb észrevételeket az alábbi tételben foglaljuk össze.

9.2. Tétel: *Legyen f konvex függvény. Ekkor konjugáltja, f^c zárt, konvex függvény, amely pontosan akkor valódi, ha f valódi. Továbbá $(\operatorname{cl} f)^c = f^c$, és $f^{cc} = \operatorname{cl} f$.* □

A konjugált függvény definíciójában az $x \in \mathcal{R}^n$ feltétel nyilván helyettesíthető az $x \in \operatorname{dom} f$ feltétellel. Sőt

9.3. Állítás: *Tetszőleges f, az $\mathcal{R}^n$ téren értelmezett, konvex függvény esetén*

$$f^c(a) = \sup\left\{a^T x - f(x) : x \in \operatorname{ri}\operatorname{dom} f\right\} \quad (a \in \mathcal{R}^n).$$

Bizonyítás: Az állításban szereplő szuprémum a $g^c(a)$ érték, ahol g az a függvény, amely a $\operatorname{ri}\operatorname{dom} f$ halmazon megegyezik az f függvénnyel, másutt pedig ∞. 8.18 szerint $\operatorname{ri}\operatorname{epi} f = \operatorname{ri}\operatorname{epi} g$, így $\operatorname{cl}\operatorname{epi} f = \operatorname{cl}\operatorname{epi} g$, és akkor $f^c = g^c$, amit bizonyítani kellett. □

Például egy $C \subseteq \mathcal{R}^n$ konvex halmaz esetén

$$(\mathrm{ind}_C)^c(a) = \sup\left\{a^Tx - \mathrm{ind}_C(x)\right\} = \sup\left\{a^Tx : x \in C\right\} = \mathrm{supp}_C(a)$$

ezért 9.2 szerint

$$(\mathrm{supp}_C)^c = (\mathrm{ind}_C)^{cc} = \mathrm{cl}\,(\mathrm{ind}_C) = \mathrm{ind}_{\mathrm{cl}\,C}.$$

Látható, hogy $\mathrm{supp}_C = \mathrm{supp}_{\mathrm{cl}\,C}$, hogy $\mathrm{supp}_{C_1} \leq \mathrm{supp}_{C_2}$ pontosan akkor, ha $\mathrm{cl}\,C_1 \subseteq \mathrm{cl}\,C_2$, továbbá hogy

9.4. Tétel: *Egy zárt, konvex halmaz indikátor- és támaszfüggvénye egymás konjugáltjai.* □

9.5. Következmény: *A zárt, valódi konvex kúpfüggvények éppen a nemüres (zárt) konvex halmazok támaszfüggvényei. Ha f valódi konvex kúpfüggvény, akkor* $\mathrm{cl}\,f$ *az alábbi C zárt, konvex halmaz támaszfüggvénye*

$$C := \left\{a \in \mathcal{R}^n : a^Tx \leq f(x)\ (x \in \mathcal{R}^n)\right\}.$$

Bizonyítás: Az állítás első feléhez elég megmutatnunk, hogy ha f valódi konvex kúpfüggvény, akkor f^c indikátorfüggvény. Valóban, ekkor egyrészt $f(0)=0$ miatt $f^c \geq 0$, másrészt ha valamely $a \in \mathcal{R}^n$ vektor esetén $f^c(a) > 0$, akkor létezik $x \in \mathcal{R}^n$ vektor úgy, hogy $a^Tx > f(x)$. De akkor

$$f^c(a) \geq a^T(\lambda x) - f(\lambda x) = \lambda(a^Tx - f(x))\ (\lambda > 0)$$

miatt (tartsunk λ-val ∞-hez) $f^c(a) = \infty$ adódik.

A feltételekből adódóan $\mathrm{cl}\,f$ zárt, valódi konvex kúpfüggvény, így az állítás már igazolt része szerint valamely $C \subseteq \mathcal{R}^n$ zárt, konvex halmaz támaszfüggvénye. Ekkor

$$f^c = (\mathrm{cl}\,f)^c = \mathrm{ind}_C, \text{ és } C = \{a : f^c(a) \leq 0\},$$

ami f^c definíciójából adódóan éppen a tétel állítása. □

További például

9.6. Tétel: *Egy poliédrikus konvex függvény konjugáltja is poliédrikus konvex függvény.*

Bizonyítás: Ha f poliédrikus függvény, vagyis epi f poliéder, akkor f "végesen generált" is, vagyis léteznek $x_1, \ldots, x_m$ vektorok és $\mu_1, \ldots, \mu_m$ számok úgy, hogy

$$\begin{aligned} \operatorname{epi} f &= \operatorname{conv}\{(x_1, \mu_1), \ldots, (x_k, \mu_k)\} + \\ &\quad + \operatorname{cone}\{(x_{k+1}, \mu_{k+1}), \ldots, (x_m, \mu_m)\} \end{aligned}$$

valamely $1 \le k \le m$ esetén. Ekkor

$$\begin{aligned} \operatorname{epi} f^c &= \Big\{(a, \beta) : \operatorname{epi} f \subseteq \{(x, \mu) : a^T x - \beta \le \mu\}\Big\} = \\ &= \left\{ (a, \beta) \left| \begin{array}{c} \sum_{i=1}^m \lambda_i (a^T x_i - \mu_i) \le \beta \\ (\lambda_i \ge 0, \sum_{i=1}^k \lambda_i = 1) \end{array} \right. \right\} = \\ &= \left\{ (a, \beta) \left| \begin{array}{l} x_i^T a - \beta \le \mu_i \ (i = 1, \ldots, k), \\ x_i^T a \le \mu_i \ (i = k+1, \ldots, m) \end{array} \right. \right\} \end{aligned}$$

poliéder. □

9.7. Következmény: *A poliédrikus, valódi konvex kúpfüggvények éppen a nemüres poliéderek támaszfüggvényei.*

Bizonyítás: 9.4, 9.5 és 9.6 következménye, mivel egy konvex halmaz pontosan akkor poliéder, ha indikátorfüggvénye poliédrikus.

Például ha $P \subseteq \mathcal{R}^n$ nemüres poliéder, akkor 9.5 szerint supp_P zárt, valódi konvex kúpfüggvény. Ezért 9.4-et is felhasználva látszik, hogy

$$\operatorname{supp}_P = \operatorname{supp}_P^{cc} = \operatorname{ind}_P^c.$$

Itt ind_P poliédrikus, valódi (speciálisan zárt) konvex függvény, így 9.6 szerint konjugáltja, supp_P is poliédrikus.

A fordított irány hasonlóan igazolható. □

A fejezet hátralévő részében a konjugálás operátorral való felcserélhetőségi tételeket bizonyítunk, egy egyszerű észrevétellel kezdve.

9.8. Állítás: *Legyen h az $\mathcal{R}^n$ téren értelmezett konvex függvény, továbbá legyen*

$$f(x) := h(A(x - c)) + b^T x + \alpha \ (x \in \mathcal{R}^n),$$

ahol $A \in \mathcal{R}^{n \times n}$ invertálható mátrix, $b, c \in \mathcal{R}^n$ vektorok, $\alpha \in \mathcal{R}$ szám. Ekkor

$$f^c(a) = h^c((A^T)^{-1}(a - b)) + c^T a + \delta,$$

ahol $\delta := -\alpha - c^T b$.

Bizonyítás: Az $y = A(x-c)$ helyettesítéssel f^c az alábbi módon számolható ki

$$\begin{aligned}
f^c(a) &= \sup_x \left\{a^T x - h(A(x-c)) - b^T x - \alpha\right\} = \\
&= \sup_y \left\{(A^{-1}y + c)^T a - h(y) - (A^{-1}y + c)^T b - \alpha\right\} = \\
&= \sup_y \left\{(A^{-1}y)^T (a-b) - h(y)\right\} + c^T(a-b) - \alpha = \\
&= h^c((A^T)^{-1}(a-b)) + c^T a + \delta,
\end{aligned}$$

amit bizonyítani kellett. □

9.8-ból már részben következik, illetve hasonlóan egyszerű számolással igazolható, hogy

9.9. Tétel: *Tetszőleges f valódi konvex függvény és $\lambda > 0$ szám esetén*

$$f(\lambda.)^c(a) = f^c(a/\lambda) \text{ és } (\lambda f)^c(a) = \lambda f^c(a/\lambda) \ (a \in \mathcal{R}^n).$$

□

9.10. Következmény: *Tetszőleges C nemüres, konvex halmaz és $\lambda > 0$ szám esetén* $\mathrm{supp}_{\lambda C} = \lambda \mathrm{supp}_C$.

Bizonyítás: 9.4 és 9.9 szerint

$$\mathrm{supp}_{\lambda C} = (\mathrm{ind}_{\lambda C})^c = (\mathrm{ind}_C(./\lambda))^c = \mathrm{supp}_C(\lambda.) = \lambda \mathrm{supp}_C,$$

felhasználva még a támaszfüggvény pozitív homogenitását. □

9.11. Következmény: *Tetszőleges C nemüres, konvex halmaz és $\lambda > 0$ szám esetén $(\lambda C)^\circ = (1/\lambda)C^\circ$.*

Bizonyítás: 9.10 egyszerű következménye, mivel

$$C^\circ = -\{a : \mathrm{supp}_C(a) \leq 1\}.$$

□

A bonyolultabb felcserélhetőségi tételekhez szükségünk lesz a konvex Stiemke-tételre, illetve speciális eseteire. Mindenekelőtt a recessziós (kúp)-függvényt írjuk fel támaszfüggvényként (vö. 9.5). Ez a lemma később is hasznosnak bizonyul (vö. 11.9).

9.12. Lemma: *Legyen f valódi konvex függvény. Ekkor* $\operatorname{dom} f$ *támaszfüggvénye az f^c függvény recessziós függvénye,* $\operatorname{rec} f^c$*. Ha f még zárt is, akkor* $\operatorname{dom} f^c$ *támaszfüggvénye az f függvény recessziós függvénye,* $\operatorname{rec} f$*.*

Bizonyítás: A definíciókból, egyszerű számolás árán az

$$\operatorname{epi}(\operatorname{supp}_{\operatorname{dom} f}) = \left\{(a,\beta) : a^T \operatorname{dom} f \leq \beta\right\}$$

halmazról könnyen belátható, hogy éppen az $\operatorname{epi}(f^c)$ halmaz recessziós kúpja. A tétel második fele az így igazolt első feléből, azt az f^c konjugált függvényre alkalmazva adódik. □

9.13. Tétel: (konvex Stiemke-tétel) *Legyen $L \subseteq \mathcal{R}^n$ altér, továbbá f valódi konvex függvény. Ekkor az L altér pontosan akkor metsz bele a* $\operatorname{ri} \operatorname{dom} f$ *halmazba, ha nem létezik $a \in L^\perp$ vektor, amelyre* $(\operatorname{rec} f^c)(a) \leq 0$, *és* $(\operatorname{rec} f^c)(-a) > 0$.

Bizonyítás: Mivel L relatív nyílt, azért a második Hahn–Banach-tétel szerint $L \cap \operatorname{ri} \operatorname{dom} f = \emptyset$ pontosan akkor, ha L és $\operatorname{dom} f$ valódi módon szeparálhatók, vagyis létezik $a \in \mathcal{R}^n$ vektor, amelyre

$$\begin{aligned}
\inf\left\{a^T x : x \in L\right\} &\geq \sup\left\{a^T x : x \in \operatorname{dom} f\right\}, \\
\sup\left\{a^T x : x \in L\right\} &> \inf\left\{a^T x : x \in \operatorname{dom} f\right\}.
\end{aligned}$$

Itt a jobb oldalon szereplő szuprémum és infimum éppen $(\operatorname{rec} f^c)(a)$, illetve $-(\operatorname{rec} f^c)(-a)$, mivel $\operatorname{rec} f^c$ a $\operatorname{dom} f$ halmaz támaszfüggvénye a lemma szerint. A bal oldalon szereplő infimum 0, ha $a \in L^\perp$, és $-\infty$, ha $a \notin L^\perp$. Most már látszik, hogy a valódi szeparálhatóság épp azt jelenti, hogy létezik $a \in L^\perp$ vektor, amelyre $0 \geq (\operatorname{rec} f^c)(a)$, és $0 > -(\operatorname{rec} f^c)(-a)$. □

9.14. Következmény: *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá g az $\mathcal{R}^m$ téren értelmezett, valódi konvex függvény. Pontosan akkor nem létezik $y \in \mathcal{R}^m$ vektor, amelyre*

$$A^T y = 0,\ (\operatorname{rec} g^c)(y) \leq 0,\ (\operatorname{rec} g^c)(-y) > 0,$$

ha $Ax \in \operatorname{ri} \operatorname{dom} g$ *legalább egy $x \in \mathcal{R}^n$ vektor esetén.*

Bizonyítás: Alkalmazzuk 9.13-at az $L := \operatorname{Im} A$ altérre és a g függvényre. □

9.15. Következmény: *Legyenek* $f_1, \dots, f_m$ *az* $\mathcal{R}^n$ *téren értelmezett, valódi konvex függvények. Pontosan akkor nem léteznek* $a_1, \dots, a_m$ *vektorok, amelyekre*

$$\begin{gathered} a_1 + \dots + a_m = 0, \\ (\operatorname{rec} f_1^c)(a_1) + \dots + (\operatorname{rec} f_m^c)(a_m) \le 0, \\ (\operatorname{rec} f_1^c)(-a_1) + \dots + (\operatorname{rec} f_m^c)(-a_m) > 0, \end{gathered}$$

ha teljesül, hogy

$$\operatorname{ri}(\operatorname{dom} f_1) \cap \dots \cap \operatorname{ri}(\operatorname{dom} f_m) \neq \emptyset.$$

Bizonyítás: Alkalmazzuk 9.13-at a diagonális altérre és az

$$f(x_1, \dots, x_m) := f_1(x_1) + \dots + f_m(x_m) \ (x_1, \dots, x_m) \in \mathcal{R}^{nm}$$

függvényre. (Ekkor f konjugáltja

$$f^c(a_1, \dots, a_m) = f_1^c(a_1) + \dots + f_m^c(a_m),$$

és konjugáltjának recessziós függvénye

$$(\operatorname{rec} f^c)(a_1, \dots, a_m) = (\operatorname{rec} f_1^c)(a_1) + \dots + (\operatorname{rec} f_m^c)(a_m).)$$

□

9.16. Tétel: *Legyen* $A \in \mathcal{R}^{m \times n}$. *Tetszőleges* f, *az* $\mathcal{R}^n$ *téren értelmezett, konvex függvény esetén teljesül, hogy*

$$(Af)^c = f^c \circ A^T.$$

Tetszőleges g, *az* $\mathcal{R}^m$ *téren értelmezett konvex függvény esetén*

$$((\operatorname{cl} g) \circ A)^c = \operatorname{cl}(A^T g^c).$$

Ha létezik $x \in \mathcal{R}^n$ *vektor, amelyre* $Ax \in \operatorname{ri} \operatorname{dom} g$, *akkor a második képletből elhagyható a lezárás operátor, ekkor*

$$(g \circ A)^c(a) = \inf \left\{ g^c(y) : A^T y = a \right\} > -\infty \ (a \in \mathcal{R}^n),$$

ahol az infimum minden $a \in A^T \operatorname{dom} g^c$ *esetén felvététetik (másutt pedig* ∞*).*

Bizonyítás: Az első képletet egyszerű számolással bizonyíthatjuk:

$$(Af)^c(b) = \sup_y \left\{ b^T y - \inf_{Ax=y} f(x) \right\} = \sup_y \sup_{Ax=y} \left\{ b^T y - f(x) \right\} =$$
$$= \sup_x \left\{ b^T Ax - f(x) \right\} = \sup_x \left\{ (A^T b)^T x - f(x) \right\} = f^c(A^T b).$$

Alkalmazzuk ezt a képletet az A^T mátrixra és a g^c függvényre, oda jutunk, hogy

$$(A^T g^c)^c = g^{cc} \circ A^{TT} = (\mathrm{cl}\, g) \circ A,$$

és következésképpen

$$((\mathrm{cl}\, g) \circ A)^c = (A^T g^c)^{cc} = \mathrm{cl}\, (A^T g^c).$$

A tétel hátralévő része nyilvánvaló, ha g felvesz valahol $-\infty$ értéket, akkor ugyanis 8.22 szerint $g(y) = -\infty$ minden $y \in \mathrm{ri}\,\mathrm{dom}\, f$ esetén, így $g(Ax) = -\infty$, tehát g^c és $(g \circ A)^c$ is azonosan ∞. Feltehető tehát, hogy $g(y) > -\infty$ minden y esetén, és valamely x esetén $Ax \in \mathrm{ri}\,\mathrm{dom}\, g$. 8.34 szerint ekkor $(\mathrm{cl}\, g) \circ A = \mathrm{cl}\, (g \circ A)$, amiből $((\mathrm{cl}\, g) \circ A)^c = (g \circ A)^c$ adódik. Másfelől 9.14 szerint g^c és A^T teljesítik 8.35 feltételeit, ekkor tehát $\mathrm{cl}\, (A^T g^c) = A^T g^c$ valódi konvex függvény, és a definíciójában szereplő infimum felvétetik, ha véges. □

9.17. Következmény: *Legyen $A \in \mathcal{R}^{m\times n}$. Bármely $C_1 \subseteq \mathcal{R}^n$ konvex halmaz esetén teljesül, hogy*

$$\mathrm{supp}_{AC_1}(y) = \mathrm{supp}_{C_1}(A^T y) \ \ (y \in \mathcal{R}^m).$$

Tetszőleges $C_2 \subseteq \mathcal{R}^m$ konvex halmaz esetén

$$\mathrm{supp}_{A^{-1}(\mathrm{cl}\, C_2)} = \mathrm{cl}\, (A^T \mathrm{supp}_{C_2}).$$

Ha létezik x, amelyre $Ax \in \mathrm{ri}\, C_2$, akkor a lezárás operáció, illetve operátor elhagyható a második képletből, ekkor

$$\mathrm{supp}_{A^{-1}(C_2)}(a) = \inf \left\{ \mathrm{supp}_{C_2}(y) : A^T y = a \right\} > -\infty \ \ (a \in \mathcal{R}^n),$$

ahol az infimum minden $a \in -A^T \mathrm{bar}\, C_2$ esetén felvétetik (másutt pedig ∞).

Bizonyítás: Alkalmazzuk 9.16-ot az $f := \mathrm{ind}_{C_1}$, $g := \mathrm{ind}_{C_2}$ függvényekre. □

9.18. Következmény: *Legyen $A \in \mathcal{R}^{m\times n}$. Bármely $C_1 \subseteq \mathcal{R}^n$ konvex halmaz esetén teljesül, hogy*

$$(AC_1)^\circ = A^{T-1}(C_1^\circ).$$

Tetszőleges, az origót lezártjában tartalmazó $C_2 \subseteq \mathcal{R}^m$ konvex halmaz esetén

$$(A^{-1}(\mathrm{cl}\, C_2))^\circ = \mathrm{cl}\,(A^T(C_2^\circ)).$$

Bizonyítás: 9.17 és 8.20 következménye. A $0 \in \mathrm{cl}\, C_2$ feltétel ahhoz kell, hogy $A^T\mathrm{supp}_{C_2}$ valódi konvex függvény legyen. □

8.37 segítségével, 9.16 bizonyításának mintájára, könnyen belátható, hogy abban az esetben, mikor g poliédrikus függvény, 9.16 azon feltétele, hogy létezzen x, amelyre $Ax \in \mathrm{ri}\,\mathrm{dom}\, g$, gyöngíthető a következő feltétellé: létezzen x, amelyre $Ax \in \mathrm{dom}\, g$. Ugyanis ha egy g poliédrikus, konvex függvény valahol $-\infty$ értéket vesz fel, akkor az effektív tartományán mindenütt $-\infty$ értéket vesz fel (hiszen $(0,-1) \in \mathrm{rec}\,\mathrm{epi}\, g$ lesz), különben pedig g zárt függvény. Továbbá $Ax \in \mathrm{dom}\, g$ esetén

$$(A^T g^c)(a) \geq a^T x - g(Ax) > -\infty \ (a \in \mathcal{R}^n)$$

is fennáll, ezért a bizonyításban $A^T g^c$ poliédrikus, valódi (és így zárt) konvex függvény.

9.19. Tétel: *Legyenek $f_1, \ldots, f_m$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények. Ekkor*

$$(f_1 \Box \ldots \Box f_m)^c = f_1^c + \ldots + f_m^c,$$
$$(\mathrm{cl}\, f_1 + \ldots + \mathrm{cl}\, f_m)^c = \mathrm{cl}\,(f_1^c \Box \ldots \Box f_m^c).$$

Ha a $\mathrm{ri}\,\mathrm{dom}\, f_i$ $(i = 1, \ldots, m)$ halmazok összemetszenek, akkor a lezárás operátor elhagyható a második képletből, ekkor

$$(f_1 + \ldots + f_m)^c(a) = \inf \left\{ \sum_{i=1}^m f_i^c(a_i) : \sum_{i=1}^m a_i = a \right\} > -\infty \ (a \in \mathcal{R}^n),$$

ahol az infimum minden $a \in \sum \mathrm{dom}\, f_i^c$ esetén felvétetik (másutt pedig ∞).

Bizonyítás: Definíció szerint

$$\begin{aligned}
&(f_1\Box\ldots\Box f_m)^c(a) = \\
&\qquad = \sup\left\{a^Tx - \inf_{x_1+\ldots+x_m=x}\{f_1(x_1)+\ldots+f_m(x_m)\}\right\} = \\
&\qquad = \sup_x \sup_{x_1+\ldots+x_m=x}\{a^Tx - f_1(x_1)-\ldots-f_m(x_m)\} = \\
&\qquad = \sup_{x_1,\ldots,x_m}\{a^Tx_1+\ldots+a^Tx_m - f_1(x_1)-\ldots-f_m(x_m)\} = \\
&\qquad = f_1^c(a)+\ldots+f_m^c(a).
\end{aligned}$$

Ebből az is következik, hogy

$$(f_1^c\Box\ldots\Box f_m^c)^c = f_1^{cc}+\ldots+f_m^{cc} = \mathrm{cl}\, f_1+\ldots+\mathrm{cl}\, f_m,$$

és így

$$(\mathrm{cl}\, f_1+\ldots+\mathrm{cl}\, f_m)^c = (f_1^c\Box\ldots\Box f_m^c)^{cc} = \mathrm{cl}\,(f_1^c\Box\ldots\Box f_m^c).$$

Ha a $\mathrm{ri}\,\mathrm{dom}\, f_i$ $(i=1,\ldots,m)$ halmazok összemetszenek, akkor 8.31 szerint a $\mathrm{cl}\, f_1+\ldots+\mathrm{cl}\, f_m$ és $\mathrm{cl}\,(f_1+\ldots+f_m)$ függvények megegyeznek. Az utóbbi függvény konjugáltja $(f_1+\ldots+f_m)^c$. Másfelől 9.15 szerint $f_1^c,\ldots,f_m^c$ teljesíti 8.36 feltételeit, így ekkor $f_1^c\Box\ldots\Box f_m^c$ zárt, és a definíciójában szereplő infimum felvétetik a $\sum \mathrm{dom}\, f_i^c$ halmazon (másutt pedig ∞). $\Box$

9.20. Következmény: *Legyenek $C_1,\ldots,C_m \subseteq \mathcal{R}^n$ nemüres, konvex halmazok. Ekkor*

$$\begin{aligned}
\mathrm{supp}_{C_1+\ldots+C_m} &= \mathrm{supp}_{C_1}+\ldots+\mathrm{supp}_{C_m}, \\
\mathrm{supp}_{\mathrm{cl}\, C_1\cap\ldots\cap \mathrm{cl}\, C_m} &= \mathrm{cl}\,(\mathrm{supp}_{C_1}\Box\ldots\Box \mathrm{supp}_{C_m}).
\end{aligned}$$

Ha a $\mathrm{ri}\, C_i$ $(i=1,\ldots,m)$ halmazok összemetszenek, akkor a lezárás operáció, illetve operátor elhagyható a második képletből, ekkor

$$\mathrm{supp}_{C_1\cap\ldots\cap C_m}(a) = \inf\left\{\sum_{i=1}^m \mathrm{supp}_{C_i}(a_i) : \sum_{i=1}^m a_i = a\right\} > -\infty \;\; (a\in\mathcal{R}^n),$$

ahol az infimum minden $a \in -\sum \mathrm{bar}\, C_i$ esetén felvétetik (másutt pedig ∞).

Bizonyítás: Alkalmazzuk 9.19-et az $f_i := \mathrm{ind}_{C_i}$ függvényekre. $\Box$

9.21. Következmény: *Legyenek* $K_1, \ldots, K_m \subseteq \mathcal{R}^n$ *konvex kúpok. Ekkor*

$$(K_1 + \ldots + K_m)^* = K_1^* \cap \ldots \cap K_m^*,$$
$$(\mathrm{cl}\, K_1 \cap \ldots \cap \mathrm{cl}\, K_m)^* = \mathrm{cl}\,(K_1^* + \ldots + K_m^*).$$

Ha a $\mathrm{ri}\, K_i$ $(i = 1, \ldots, m)$ *halmazok összemetszenek, akkor a lezárás operáció elhagyható a második képletből.*

Bizonyítás: Alkalmazzuk 9.19-et az $f_i := \mathrm{ind}_{K_i}$ függvényekre. Könnyen utánaszámolhatunk, hogy ekkor $f_i^c = \mathrm{ind}_{-K_i^*}$. □

9.19 egy alkalmazásaként meghatározzuk a dist_C távolságfüggvény konjugáltját. Már említettük, hogy tetszőleges $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz esetén $\mathrm{dist}_C = ||.|| \Box \mathrm{ind}_C$. Határozzuk meg a $||.||$ függvény konjugáltját (pozitív homogén, így konvexitása a háromszög-egyenlőtlenség következménye). Ha $||a|| \leq 1$, akkor

$$||.||^c(a) = \sup a^T x - ||x|| \leq$$
$$\leq \sup ||a|| \cdot ||x|| - ||x|| \leq \sup 1 \cdot ||x|| - ||x|| = 0,$$

vagyis ekkor $||.||^c(a) = 0$. Ha pedig $||a|| > 1$, akkor

$$||.||^c(a) = \sup a^T x - ||x|| \geq$$
$$\geq \sup_{\lambda \geq 0} a^T(\lambda a) - ||\lambda a|| = ||a||(||a|| - 1) \sup_{\lambda \geq 0} \lambda = \infty,$$

vagyis ekkor $||.||^c(a) = \infty$. Tehát $||.||^c = \mathrm{ind}_O$, ahol O az $\mathcal{R}^n$ térbeli zárt egységgömb. Mindezekből 9.19 szerint

$$(\mathrm{dist}_C)^c(a) = ||.||^c(a) + (\mathrm{ind}_C)^c(a) = \begin{cases} \mathrm{supp}_C(a), & \text{ha } ||a|| \leq 1, \\ \infty, & \text{ha } ||a|| > 1 \end{cases}$$

adódik.

Most megvizsgáljuk azt az esetet, mikor 9.19-ben minden f_i függvény poliédrikus. Megmutatjuk, hogy ekkor a relatív belsők összemetszése helyett a

$$\mathrm{dom}\, f_1 \cap \ldots \cap \mathrm{dom}\, f_m \neq \emptyset$$

feltétel is elegendő a tétel második felének konklúziójához. Ekkor ugyanis $\mathrm{cl}\, f_i = f_i$, és $f_1 + \ldots + f_m$ valódi konvex, poliédrikus függvény (8.37). Szükségképpen az $f_1 + \ldots + f_m$ függvény konjugáltja is valódi konvex függvény, és mivel 9.19 szerint $f_1^c \Box \ldots \Box f_m^c$ lezártja, azért maga a konvolúció is valódi konvex függvény. Minden f_i^c poliédrikus, konvex függvény

(9.6), így $f_1^c \Box \ldots \Box f_m^c$ is poliédrikus, konvex függvény (8.37), az eddigiekből adódóan zárt is. 9.19 szerint ekkor

$$(f_1 + \ldots + f_m)^c = f_1^c \Box \ldots \Box f_m^c,$$

és 8.37 szerint az utóbbi konvolúció definíciójában szereplő infimum felvétetik effektív tartományán, a $\sum \operatorname{dom} f_i^c$ halmazon.

Most megfogalmazzuk 9.19 vegyes változatát.

9.22. Tétel: *Legyenek $f_1, \ldots, f_m$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények úgy, hogy $f_1, \ldots, f_k$ poliédrikus. Tegyük fel, hogy*

$$\operatorname{dom} f_1 \cap \ldots \cap \operatorname{dom} f_k \cap \operatorname{ri}(\operatorname{dom} f_{k+1}) \cap \ldots \cap \operatorname{ri}(\operatorname{dom} f_m) \neq \emptyset.$$

Ekkor

$$\begin{aligned}(f_1 + \ldots + f_m)^c(a) &= (f_1^c \Box \ldots \Box f_m^c)(a) = \\ &= \inf\{f_1^c(a_1) + \ldots + f_m^c(a_m) : a_1 + \ldots + a_m = a\} > -\infty \ (a \in \mathcal{R}^n),\end{aligned}$$

ahol az infimum minden $a \in \sum \operatorname{dom} f_i^c$ esetén felvétetik (másutt pedig ∞).

Bizonyítás: A $k = 0$, és $k = m$ esetekben már beláttuk a tételt, feltehető tehát, hogy $1 \leq k < m$. Legyen

$$g_1 := f_1 + \ldots + f_k, \text{ és } g_2 := f_{k+1} + \ldots + f_m.$$

A tétel segítségével kiszámolhatjuk a g_1 és g_2 függvények konjugáltjait, ezek

$$\begin{aligned}g_1^c(b_1) &= \inf\{f_1^c(a_1) + \ldots + f_k^c(a_k) : a_1 + \ldots + a_k = b_1\},\\ g_2^c(b_2) &= \inf\{f_{k+1}^c(a_{k+1}) + \ldots + f_m^c(a_m) : a_{k+1} + \ldots + a_m = b_2\},\end{aligned}$$

ahol az infimumok minden b_1 és b_2 esetén felvétetnek (esetleg ∞-ként). Így elegendő annyit belátni, hogy

$$(g_1 + g_2)^c(a) = \inf\{g_1^c(b_1) + g_2^c(b_2) : b_1 + b_2 = a\},$$

ahol az infimum minden a esetén felvétetik. Itt g_1 és g_2 valódi konvex függvények, g_1 ráadásul poliédrikus (8.37). Mivel

$$\begin{aligned}\operatorname{dom} g_1 &= \operatorname{dom} f_1 \cap \ldots \cap \operatorname{dom} f_k\\ \operatorname{dom} g_2 &= \operatorname{dom} f_{k+1} \cap \ldots \cap \operatorname{dom} f_m,\end{aligned}$$

azért teljesül, hogy

$$\begin{aligned}\operatorname{ri}(\operatorname{dom} g_2) &= \operatorname{ri}(\operatorname{dom} f_{k+1} \cap \ldots \cap \operatorname{dom} f_m) = \\ &= \operatorname{ri}(\operatorname{dom} f_{k+1}) \cap \ldots \cap \operatorname{ri}(\operatorname{dom} f_m),\end{aligned}$$

és így

$$\operatorname{dom} g_1 \cap \operatorname{ri}(\operatorname{dom} g_2) \neq \emptyset.$$

Ebből már következik (vizsgáljuk relatív belső pontjait), hogy az $M := \operatorname{aff}(\operatorname{dom} g_2)$ jelöléssel

$$\operatorname{ri}(M \cap \operatorname{dom} g_1) \cap \operatorname{ri}(\operatorname{dom} g_2) \neq \emptyset.$$

Legyen $h := \operatorname{ind}_M + g_1$, ekkor h valódi konvex függvény, amelynek effektív tartománya $M \cap \operatorname{dom} g_1$, így

$$\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{ri}(\operatorname{dom} g_2) \neq \emptyset,$$

és a tételben szereplő képlet alkalmazható $(h+g_2)^c$ kiszámítására. Továbbá $h + g_2 = g_1 + g_2$, így

$$(g_1 + g_2)^c(a) = (h^c \square g_2^c)(a) = \inf\{h^c(c) + g_2^c(b) : c + b = a\},$$

ahol az infimum minden a esetén felvétetik. Másfelől mivel ind_M és g_1 poliédrikus függvények, azért a tételben szereplő képletet alkalmazhatjuk h^c kiszámítására is, eszerint

$$h^c(c) = \inf\{\operatorname{supp}_M(u) + g_1^c(b_1) : u + b_1 = c\},$$

ahol az infimum minden c esetén felvétetik. Ezért

$$(g_1 + g_2)^c(a) = \inf\{\operatorname{supp}_M(u) + g_1^c(b_1) + g_2^c(b) : u + b_1 + b = a\},$$

ahol az infimum minden a esetén felvétetik. Az ind_M és g_2 függvények effektív tartományainak relatív belsejei nyilvánvalóan összemetszenek, így a tételben szereplő képletet még egyszer alkalmazva látjuk, hogy

$$\inf\{\operatorname{supp}_M(u) + g_2^c(b) : u + b = b_2\} = (\operatorname{ind}_M + g_2)^c(b_2) = g_2^c(b_2),$$

és itt az infimum mindig felvétetik. Összefoglalva oda jutottunk, hogy

$$(g_1 + g_2)^c(a) = \inf\{g_1^c(b_1) + g_2^c(b_2) : b_1 + b_2 = a\},$$

ahol az infimum minden a esetén felvétetik — és éppen ezt kellett bizonyítani. $\Box$

9.23. Következmény: *Legyenek $f_1, \ldots, f_m$ az $\mathcal{R}^n$ téren értelmezett, zárt, valódi konvex függvények úgy, hogy $f_1, \ldots, f_k$ poliédrikus. Tegyük fel, hogy*

$$\operatorname{dom} f_1^c \cap \ldots \cap \operatorname{dom} f_k^c \cap \operatorname{ri}(\operatorname{dom} f_{k+1}^c) \cap \ldots \cap \operatorname{ri}(\operatorname{dom} f_m^c) \neq \emptyset.$$

Ekkor $f_1 \Box \ldots \Box f_m$ zárt, valódi konvex függvény, és a definíciójában szereplő infimum mindig felvétetik, ha véges.

Bizonyítás: Alkalmazzuk 9.22-t az $f_1^c, \ldots, f_m^c$ függvényekre. $\Box$

9.22 felhasználható 7.26 bizonyítására. Legyenek ugyanis $C_1, \ldots, C_m$ $\mathcal{R}^n$-beli, nemüres, konvex halmazok úgy, hogy $C_1, \ldots, C_k$ poliéder. Tegyük fel, hogy

$$C_1 \cap \ldots \cap C_k \cap (\operatorname{ri} C_{k+1}) \cap \ldots \cap (\operatorname{ri} C_m) \neq \emptyset.$$

Ekkor 9.22 szerint

$$\begin{aligned}(\operatorname{ind}_{\cap C_i})^c(a) = (\textstyle\sum \operatorname{ind}_{C_i})^c(a) = \\ \inf \{\textstyle\sum_{i=1}^m (\operatorname{ind}_{C_i})^c(a_i) : \sum_{i=1}^m a_i = a\},\end{aligned}$$

ahol az infimum minden a esetén felvétetik (ha véges). Legyen F a $\cap C_i$ halmaz nemüres, exponált lapja, és tegyük fel, hogy a exponálja, vagyis

$$a^T(\cap C_i \setminus F) < \operatorname{supp}_{\cap C_i}(a) = a^T F.$$

A fentiek szerint léteznek $a_1, \ldots, a_m$ vektorok úgy, hogy $a = a_1 + \ldots + a_m$, és

$$\operatorname{supp}_{\cap C_i}(a) = \operatorname{supp}_{C_1}(a_1) + \ldots + \operatorname{supp}_{C_m}(a_m).$$

Legyen F_i a C_i halmaz az a_i vektor által exponált része, vagyis

$$F_i := \left\{x \in C_i : a_i^T x = \operatorname{supp}_{C_i}(a_i)\right\} \quad (i = 1, \ldots, m).$$

Könnyen belátható, hogy ekkor $F = \cap F_i$, amivel 7.26 nemtriviális irányát igazoltuk.

9.24. Tétel: *Legyen f_i $(i \in I)$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények egy tetszőleges rendszere. Ekkor*

$$\begin{aligned}(\operatorname{conv}\{f_i : i \in I\})^c = \sup\{f_i^c : i \in I\}, \\ (\sup\{\operatorname{cl} f_i : i \in I\})^c = \operatorname{cl}(\operatorname{conv}\{f_i^c : i \in I\}).\end{aligned}$$

Ha az I indexhalmaz véges, és a $\mathrm{cl}\,(\mathrm{dom}\, f_i)$ *halmazok mind megegyeznek egy C halmazzal (ez a helyzet például, ha minden f_i véges az egész $\mathcal{R}^n$ téren), akkor a lezárás operátor elhagyható a második képletből. Továbbá ebben az esetben*

$$(\sup\{f_i : i \in I\})^c(a) =$$
$$= \inf \left\{ \textstyle\sum_{i\in I} \lambda_i f_i^c(a_i) \,\middle|\, \begin{array}{l} \lambda_i \geq 0,\ \sum_{i\in I} \lambda_i = 1, \\ a_i \in \mathrm{dom}\, f_i^c,\ \sum_{i\in I} \lambda_i a_i = a \end{array} \right\} > -\infty \ (a \in \mathcal{R}^n),$$

ahol az infimum minden $a \in \mathrm{conv}\,(\cup \mathrm{dom}\, f_i^c)$ *esetén felvétetik (másutt pedig ∞).*

Bizonyítás: Legyen $f := \mathrm{conv}\,\{f_i : i \in I\}$. Az $\mathrm{epi}\, f^c$ epigráf minden (a, β) eleme egy $h := a^T. - \beta$ affin függvénynek felel meg, amelyre $h \leq f$. Ezek a h affin függvények éppen azok, amelyekre $h \leq f_i$ teljesül minden $i \in I$ esetén. Ezért $\beta \geq f^c(a)$ pontosan akkor, ha $\beta \geq f_i^c(a)$ minden $i \in I$ esetén, amivel az első képletet beláttuk. Ezt a képletet az f_i^c függvényekre alkalmazva kapjuk, hogy

$$(\mathrm{conv}\,\{f_i^c : i \in I\})^c = \sup\{f_i^{cc} : i \in I\} = \sup\{\mathrm{cl}\, f_i : i \in I\},$$

és következésképpen

$$(\sup\{\mathrm{cl}\, f_i : i \in I\})^c = (\mathrm{conv}\,\{f_i^c : i \in I\})^{cc} = \mathrm{cl}\,(\mathrm{conv}\,\{f_i^c : i \in I\}).$$

Ha a $\mathrm{ri}\,(\mathrm{dom}\, f_i)$ halmazok metszetében van olyan pont, ahol az f_i függvények szuprémuma véges, akkor 8.32 szerint

$$(\sup\{\mathrm{cl}\, f_i : i \in I\})^c = (\mathrm{cl}\,(\sup\{f_i : i \in I\}))^c = (\sup\{f_i : i \in I\})^c.$$

Speciálisan ez az egyenlőség fennáll, ha az I indexhalmaz véges, és a $\mathrm{dom}\, f_i$ halmazok lezártja ugyanaz a C halmaz minden $i \in I$ esetén. Az utóbbi esetben a $\mathrm{dom}\, f_i$ halmazok támaszfüggvényei, amelyek éppen a $\mathrm{rec}\, f_i^c$ recessziós függvények 9.12 szerint, mind a supp_C támaszfüggvénnyel egyeznek meg, ezért 8.33 szerint $\mathrm{conv}\,\{f_i^c : i \in I\}$ zárt, és éppen a tételbeli infimumos képlet adja meg. □

9.25. Következmény: *Legyenek $C_i \subseteq \mathcal{R}^n$ $(i \in I)$ konvex halmazok. Jelölje C az uniójuk konvex burkát, C' pedig a lezártjaik metszetét. Ekkor*

$$\mathrm{supp}_C = \sup\{\mathrm{supp}_{C_i} : i \in I\},$$
$$\mathrm{supp}_{C'} = \mathrm{cl}\,(\mathrm{conv}\,\{\mathrm{supp}_{C_i} : i \in I\}).$$

Bizonyítás: Legyen $f_i := \mathrm{ind}_{C_i}$, és alkalmazzuk 9.24-et. $\square$

9.26. Következmény: *Legyenek $C_i \subseteq \mathcal{R}^n$ $(i \in I)$ konvex halmazok. Ekkor*

$$(\mathrm{conv}\,\{C_i : i \in I\})^\circ = \cap\{C_i^\circ : i \in I\}.$$

Továbbá ha $0 \in \mathrm{cl}\,C_i$ $(i \in I)$, akkor

$$(\cap\{\mathrm{cl}\,C_i : i \in I\})^\circ = \mathrm{cl}\,(\mathrm{conv}\,\{C_i^\circ : i \in I\}).$$

$\square$

A fejezetben szereplő tételek természetesen megfelelő változtatásokkal kimondhatók lennének konvex helyett konkáv függvényekkel, ekkor az f^c (konvex) konjugált helyett a g_c ún. **konkáv konjugált** szerepelne, ahol $g : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény esetén

$$g_c(b) := \inf\left\{b^T y - g(y) : y \in \mathcal{R}^n\right\} \ (b \in \mathcal{R}^n).$$

A konvex és konkáv konjugált közti egyszerű kapcsolatot az alábbi állítás írja le:

9.27. Állítás: *Tetszőleges $g : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény esetén*

$$-g_c = (-g)^c(-.), \ \textit{így}\ \mathrm{dom}\, g_c = -\mathrm{dom}\,(-g)^c,$$

konkáv függvény effektív tartománya alatt természetesen **felső effektív tartomány***át, a (-1)-szeresének alsó effektív tartományát értve.* $\square$

Teljesül például, ha $g : \mathcal{R}^n \to [-\infty, \infty[$ nem azonosan $-\infty$ konkáv függvény (vagyis valódi konkáv függvény), akkor g_{cc} a g felső lezártja. Egy másik észrevétel a kétféle konjugálás kapcsolatáról:

9.28. Állítás: *Legyen $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konvex függvény, $g : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konkáv függvény. Ekkor*

$$(-f^c)_c = -f^{cc}(-.), \ \textit{és}\ (-g_c)^c = -g_{cc}(-.).$$

$\square$

10. Folytonosság, (szub)differenciálhatóság

Ebben a fejezetben konvex függvények folytonosságát és differenciálhatóságát vizsgáljuk. Akárcsak a folytonosság esetében a 8. fejezetben, itt is megjelenik egy "félig differenciálhatósági" tulajdonság, a szubdifferenciálhatóság. (Csak itt a differenciálhatósághoz szükséges másik "felet" nem a szuperdifferenciálhatóság, hanem a szubdifferenciál egyeleműsége jelenti.)

Először két alapvető folytonossági tételt tárgyalunk.

10.1. Tétel: *Egy az $\mathcal{R}^n$ téren értelmezett f valódi konvex függvény effektív tartománya minden C relatív nyílt, konvex részhalmazára megszorítva folytonos. Speciálisan f folytonos a $\operatorname{ri}\operatorname{dom} f$ halmazra megszorítva.*

Bizonyítás: Legyen g az a függvény, amely megegyezik az f függvénnyel a C halmazon, és másutt ∞ értéket vesz fel, ekkor g effektív tartománya éppen C. Az f függvényt a g függvénnyel helyettesítve, ha szükséges, feltehető, hogy $C = \operatorname{dom} f$. A bizonyítás ezek után két lépésből áll.

1. Először megmutatjuk, hogy ha $x_0 \in C$, és $f(x_0) < \gamma$ valamely $\gamma \in \mathcal{R}$ szám esetén, akkor létezik $C' \subseteq C$ relatív nyílt, konvex halmaz, amelyre $x_0 \in C'$, $\operatorname{aff} C' = \operatorname{aff} C$, és $f|_{C'} < \gamma$ teljesül. Valóban, 8.20 szerint

$$C' := \operatorname{ri}\{x : f(x) < \gamma\}$$

megfelel.

2. Másodszor megmutatjuk, hogy ha $x_0 \in C$, és $f(x_0) > \beta$ valamely $\beta \in \mathcal{R}$ szám esetén, akkor létezik $C'' \subseteq C$ relatív nyílt, konvex halmaz, amelyre $x_0 \in C''$, $\operatorname{aff} C'' = \operatorname{aff} C$, és $f|_{C''} \geq \beta$ teljesül. Valóban, ha $\inf f \geq \beta$, akkor $C'' := C$ megfelel. Ha pedig $\inf f < \beta$, akkor 8.20 szerint $x_0 \notin \operatorname{cl}\{x : f(x) < \beta\}$, mivel az utóbbi halmaz éppen $\{x : (\operatorname{cl} f)(x) \leq \beta\}$, és $x_0 \in \operatorname{ri}\operatorname{dom} f$ miatt $(\operatorname{cl} f)(x_0) = f(x_0)$. Ezért az első Hahn–Banach-tétel szerint létezik $a \in \mathcal{R}^n$ vektor úgy, hogy

$$a^T x_0 < \inf\left\{a^T x : f(x) < \beta\right\}.$$

A fenti infimumot δ-val jelölve könnyen belátható, hogy

$$C'' := \left\{x \in C : a^T x < \delta\right\}$$

megfelel. □

10.2. Következmény: *Az egész $\mathcal{R}^n$ téren véges f konvex függvény szükségképpen folytonos.* □

Legyen f olyan függvény, amely az $S \subseteq \mathcal{R}^n$ halmazon valós értékeket vesz fel. Azt mondjuk, hogy f **Lipschitz-függvény** az S halmazon az $\alpha \geq 0$ **Lipschitz-állandó**val, ha

$$|f(y) - f(x)| \leq \alpha ||y - x|| \ (x, y \in S).$$

10.3. Tétel: *Legyen f valódi konvex függvény, és legyen S a* $\mathrm{ri}\,\mathrm{dom}\, f$ *halmaz valamely kompakt részhalmaza. Ekkor f Lipschitz-függvény az S halmazon.*

Bizonyítás: Feltehető, hogy $\mathrm{dom}\, f$ n-dimenziós, tehát S valójában a $\mathrm{dom}\, f$ halmaz belsejének része. Legyen O az $\mathcal{R}^n$ térbeli zárt egységgömb. Ekkor minden $\varepsilon > 0$ esetén $S + \varepsilon O$ kompakt halmaz (lévén az $S \times O$ kompakt halmaz az $(x, u) \mapsto x + \varepsilon u$ folytonos leképezésnél vett képe). Az

$$(S + \varepsilon O) \cap (\mathcal{R}^n \setminus \mathrm{int}\,(\mathrm{dom}\, f)) \ (\varepsilon > 0)$$

kompakt halmazok az ε csökkentével szűkülnek, továbbá metszetük üres, így Cantor metszettétele szerint egyikük üres. Ezért bizonyos $\varepsilon > 0$ esetén

$$S + \varepsilon O \subseteq \mathrm{int}\,(\mathrm{dom}\, f).$$

10.1 szerint f folytonos az $S+\varepsilon O$ halmazon. Mivel $S+\varepsilon O$ kompakt halmaz, azért f korlátos az $S + \varepsilon O$ halmazon. Legyen α_1 és α_2 alsó, illetve felső korlát. Legyen x és y két különböző S-beli pont, és legyen

$$z := y + (\varepsilon/||y - x||)(y - x).$$

Ekkor $z \in S + \varepsilon O$, és

$$y = (1 - \lambda)x + \lambda z, \text{ ahol } \lambda = ||y - x||/(\varepsilon + ||y - x||).$$

Az f konvexitásából adódóan

$$f(y) \leq (1 - \lambda)f(x) + \lambda f(z) = f(x) + \lambda(f(z) - f(x)),$$

következésképpen

$$f(y) - f(x) \leq \lambda(\alpha_2 - \alpha_1) \leq \alpha||y - x||, \text{ ahol } \alpha = (\alpha_2 - \alpha_1)/\varepsilon.$$

Ez az egyenlőtlenség fennáll tetszőleges $x, y \in S$ esetén, így f Lipschitz-függvény az S halmazon. □

Legyen $S \subseteq \mathcal{R}^n$, továbbá $f : S \to \overline{\mathcal{R}}$. Azt mondjuk, hogy az f függvény **differenciálható** az $x \in S$ pontban, ha f valós értékeket vesz fel az x egy környezetén, továbbá létezik $a \in \mathcal{R}^n$ vektor, amelyre

$$\lim_{z \to x} \frac{f(z) - f(x) - a^T(z - x)}{||z - x||} = 0.$$

Ha ilyen a, ún. **gradiens** vektor létezik, akkor szükségképpen egyértelmű, jele $\nabla f(x)$. Azt mondjuk, hogy f differenciálható egy S' halmazon, ha az f függvény differenciálható az S' halmaz minden pontjában. Ekkor a definícióból adódóan S' az S halmaz azon része belsejének része, ahol f véges értékeket vesz fel.

Ha az f függvény differenciálható az $x \in S$ pontban, akkor tetszőleges $y \in \mathcal{R}^n$ esetén

$$0 = \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x) - (\nabla f(x))^T(\lambda y)}{\lambda ||y||},$$

tehát léteznek a

$$\lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

iránymenti derivál'tak, és értékük éppen $(\nabla f(x))^T y$.

Konvex függvények esetén fontosabb szerep jut az

$$f'(x; y) := \lim_{\lambda \to 0+} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

(esetleg végtelen) **féloldali iránymenti derivált**aknak és a szubgradienseknek. Konvex függvények féloldali iránymenti deriváltjaival kapcsolatos néhány alapvető észrevételt fogalmaz meg az alábbi

10.4. Tétel: *Legyen f konvex függvény, és legyen x egy olyan pont, ahol f véges értéket vesz fel. Ekkor tetszőleges $y \in \mathcal{R}^n$ esetén létezik $f'(x; y)$, és*

$$f'(x; y) = \inf_{\lambda > 0} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

Továbbá $f'(x; y)$ konvex kúpfüggvény az y változóban, és $f'(x; 0) = 0$ mellett teljesül, hogy

$$-f'(x; -y) \leq f'(x; y) \ (y \in \mathcal{R}^n),$$

egyenlőséggel, ha f differenciálható az x pontban.

Bizonyítás: 8.14 és 8.15 szerint a féloldali iránymenti derivált definíciójában szereplő hányados konvex f függvény esetén a λ nemcsökkenő függvénye, így $f'(x;y)$ létezik, és éppen a tételben szereplő infimum. Világos, hogy $f'(x;.)$ pozitív homogén, továbbá $f'(x;0)=0$, így $f'(x;.)$ kúpfüggvény is. Konvexitása 8.10 egyszerű következménye. Most belátjuk a tételbeli utolsó egyenlőtlenséget. Adott $\mu_1 > f'(x;-y)$ és $\mu_2 > f'(x;y)$ számok esetén

$$(1/2)\mu_1 + (1/2)\mu_2 \geq f'(x;(1/2)(-y)+(1/2)y) = 0$$

az $f'(x;.)$ függvény konvexitása miatt. Ezért $-f'(x;-y) \leq f'(x;y)$ minden y esetén. Ha f differenciálható az x pontban, akkor

$$\begin{aligned} -f'(x;-y) &= \lim_{\lambda\to 0-} \tfrac{f(x+\lambda y)-f(x)}{\lambda} = \\ &= \lim_{\lambda\to 0} \tfrac{f(x+\lambda y)-f(x)}{\lambda} = \lim_{\lambda\to 0+} \tfrac{f(x+\lambda y)-f(x)}{\lambda} = f'(x;y). \end{aligned}$$

□

Az a vektor az $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ függvény **szubgradiens**e az $x \in \mathcal{R}^n$ pontban, ha teljesül az ún. **szubgradiens egyenlőtlenség**,

$$f(z) \geq f(x) + a^T(z-x) \ (z \in \mathcal{R}^n).$$

Az f függvény x pontbeli szubgradienseinek halmaza az f függvény x pontbeli **szubdifferenciál**ja, melynek jele $\partial f(x)$. Nyilván $\partial f(x)$ zárt, konvex halmaz, mivel definíció szerint $a \in \partial f(x)$ pontosan akkor, ha az a vektor néhány zárt féltér eleme egyszerre. Általában $\partial f(x)$ lehet üres is; ha nem az, akkor azt mondjuk, hogy az f függvény **szubdifferenciálható** az x pontban.

Például ha $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, akkor $a \in \partial \mathrm{ind}_C(x)$ pontosan akkor, ha

$$\mathrm{ind}_C(z) \geq \mathrm{ind}_C(x) + a^T(z-x) \ (z \in \mathcal{R}^n).$$

Ez a feltétel azt jelenti, hogy $x \in C$, és $0 \geq a^T(z-x)$ minden $z \in C$ esetén. Tehát ha $x \in C$, akkor $\partial \mathrm{ind}_C(x)$ az

$$\left\{a : a^T(C-x) \leq 0\right\}$$

(zárt, konvex) kúp, a C halmaz **normális kúp**ja az x pontban; $x \notin C$ esetén pedig $\partial \mathrm{ind}_C(x) = \emptyset$.

10.5. Állítás: *Legyen f poliédrikus, konvex függvény, és legyen x olyan pont, ahol f véges értéket vesz fel. Ekkor f szubdifferenciálható az x pontban, és $\partial f(x)$ poliéder.*

Bizonyítás: Az $\operatorname{epi} f$ halmaz poliéder, így létezik A mátrix és b vektor úgy, hogy $\operatorname{epi} f = \{z : Az \geq b\}$. Az is feltehető, hogy A minden sorvektorának utolsó eleme 0 vagy 1. Legyen $\hat{x} := (x, f(x))$, és jelölje I az A mátrix azon sorindexeinek halmazát, amelyeknek megfelelő egyenlőtlenségeket az $\hat{x}$ pont egyenlőséggel teljesíti. Az I halmaz nem lehet üres, akkor ugyanis $\hat{x}$ az $\operatorname{epi} f$ halmaz relatív belső pontja lenne (7.3), pedig nem az (8.18). Ha az A mátrix I-beli sorindexű sorvektorainak utolsó elemei mind nullák lennének, akkor elég kis $\varepsilon > 0$ esetén $\hat{x} - (0, \varepsilon) \in \operatorname{epi} f$ lenne, ellentmondásban azzal, hogy $\hat{x}$ utolsó eleme $f(x)$. Ezért létezik $i \in I$ sorindex úgy, hogy az A mátrix i-edik sorának utolsó eleme 1. Legyen ez a sor éppen $(a^T, 1)$. Ekkor könnyen belátható, hogy $-a \in \partial f(x)$, tehát $\partial f(x)$ nemüres. A $\partial f(x)$ halmaz poliédrikussága a nyilvánvaló

$$\partial f(x) = -C((\operatorname{cone}(\operatorname{epi} f - \hat{x}))^*)$$

egyenlőség egyszerű következménye (vö. 4.19). □

A szubdifferenciál és a féloldali iránymenti deriváltak közti kapcsolatról szól a következő három tétel.

10.6. Tétel: *Legyen f konvex függvény, és legyen x egy olyan pont, ahol az f függvény véges értéket vesz fel. Ekkor az a vektor pontosan akkor szubgradiense az f függvénynek az x pontban, ha*

$$f'(x; y) \geq a^T y \ (y \in \mathcal{R}^n).$$

Teljesül továbbá, hogy az $f'(x; .)$ konvex kúpfüggvény lezártja a $\partial f(x)$ zárt, konvex halmaz támaszfüggvénye.

Bizonyítás: A $z := x + \lambda y$ értékadással a szubgradiens egyenlőtlenség az alábbi feltétellé fordítható le:

$$(f(x + \lambda y) - f(x))/\lambda \geq a^T y$$

minden y és $\lambda > 0$ esetén. Mivel az itt szereplő hányados λ nemcsökkenő függvénye, és az $f'(x; y)$ értékhez tart, ahogy $\lambda \to 0+$, azért az utóbbi egyenlőtlenség a tételbelivel ekvivalens. A tétel hátralévő része 9.5 következménye, ha az $f'(x; .)$ konvex kúpfüggvény valódi is (különben pedig a tétel már igazolt részéé). □

10.7. Tétel: *Legyen f konvex függvény, és legyen x olyan pont, ahol az f függvény véges értéket vesz fel. Ha f szubdifferenciálható az x pontban, akkor f valódi konvex függvény. Ha f nem szubdifferenciálható az x pontban, akkor létezik y, amelyre*

$$f'(x;y) = -f'(x;-y) = -\infty;$$

sőt az utóbbi egyenlőségek minden $y \in (\mathrm{ri}\,\mathrm{dom}\, f) - x$ esetén fennállnak.

Bizonyítás: Az f függvény x pontbeli szubdifferenciálhatósága esetén f majorál egy affin függvényt, tehát f valódi. A $\partial f(x)$ halmaz pontosan akkor üres, ha támaszfüggvénye a konstans $-\infty$ függvény. Ez a támaszfüggvény éppen $\mathrm{cl}\, f'(x;.)$ az előző tétel szerint. Egy konvex függvény lezártja pontosan akkor azonosan $-\infty$, ha a függvény maga valahol $-\infty$ értéket vesz fel. Így ha az f függvény nem szubdifferenciálható az x pontban, akkor létezik y, amelyre $f'(x;y) = -\infty$ (és akkor $-f'(x;-y) = -\infty$ is teljesül, mivel 10.4 szerint $-f'(x;-y) \leq f'(x;y)$). Ekkor az $f'(x;.)$ függvény $-\infty$ értéket vesz fel effektív tartománya (jelölje C) minden relatív belső pontjában (8.22). De a C halmaz, 10.4-ből könnyen láthatóan, éppen a $\lambda C'$ ($\lambda > 0$) halmazok úniója, ahol $C' := (\mathrm{dom}\, f) - x$, és mivel $0 \in C'$, azért ebből

$$C' \subseteq C \subseteq \mathrm{aff}\, C'$$

is következik. Ezért $\mathrm{aff}\, C' = \mathrm{aff}\, C$, így $C' \subseteq C$ miatt $\mathrm{ri}\, C' \subseteq \mathrm{ri}\, C$. Látjuk, hogy az $f'(x;.)$ függvény $-\infty$ értéket vesz fel minden $\mathrm{ri}\, C' = (\mathrm{ri}\,\mathrm{dom}\, f) - x$ halmazbeli ponton. □

10.8. Tétel: *Legyen f valódi konvex függvény. Ha $x \notin \mathrm{dom}\, f$, akkor $\partial f(x)$ üres. Ha $x \in \mathrm{ri}\,(\mathrm{dom}\, f)$, akkor $\partial f(x)$ nemüres, $f'(x;.)$ zárt, valódi konvex kúpfüggvény, és*

$$f'(x;y) = \sup\left\{a^T y : a \in \partial f(x)\right\} = \mathrm{supp}_{\partial f(x)}(y).$$

Végül $\partial f(x)$ pontosan akkor nemüres, korlátos halmaz, ha $x \in \mathrm{int}\,(\mathrm{dom}\, f)$, amikor is $f'(x;y)$ véges minden y esetén, és a fenti szuprémum felvétetik.

Bizonyítás: A z vektort a $\mathrm{dom}\, f$ halmazból választva, a szubgradiens egyenlőtlenségből látszik, hogy $f(x) = \infty$ esetén $\partial f(x) = \emptyset$. Ha $x \in \mathrm{ri}\,(\mathrm{dom}\, f)$, akkor az $f'(x;.)$ függvény effektív tartománya affin halmaz, $\mathrm{par}\,\mathrm{dom}\, f$. Mivel $f'(x;0) = 0$, azért az $f'(x;.)$ függvény nem lehet azonosan $-\infty$ a par dom f halmazon. Ezért $f'(x;.)$ valódi (8.23) és zárt (8.25). De akkor 10.6 szerint $f'(x;.)$ maga a $\partial f(x)$ halmaz támaszfüggvénye, ahonnan már látszik a tételbeli szuprémumos képlet, és hogy $\partial f(x)$ nemüres. A tétel hátralévő része a nyilvánvaló

$$\mathrm{rec}\,(\partial f(x)) = -(\mathrm{cone}\,(\mathrm{dom}\, f - x))^*$$

egyenlőség egyszerű következménye (vö. 4.9). □

Ezzel elégséges feltételt nyertünk arra, hogy a szubgradiensek halmaza nemüres legyen, vagyis az üres halmaz "szigorú alsó korlátja" legyen a $\partial f(x)$ halmaznak. Most "felső korlátot" keresünk ugyanehhez a halmazhoz.

10.9. Tétel: *Legyen f valódi konvex függvény, továbbá x olyan pont, ahol f szubdifferenciálható, de f nem az x pontban veszi fel az infimumát (vagyis $0 \notin \partial f(x)$). Ekkor a $C := \{z : f(z) \leq f(x)\}$ halmaz normális kúpja az x pontban a $\partial f(x)$ halmaz által generált konvex kúp lezártja.*

Bizonyítás: 8.20 szerint a $\{z : f(z) < f(x)\}$ halmaznak ugyanaz a lezártja, mint a C halmaznak, mivel feltettük, hogy $f(x) > \inf f$. Ezért az a vektor akkor és csak akkor eleme a C halmaz x pontbeli normális kúpjának, ha $a^T(z - x) \leq 0$ valahányszor $f(z) < f(x)$. A $\lambda(z - x)$ ($\lambda > 0$, $f(z) < f(x)$) alakú y vektorok éppen azok, amelyekre $f'(x; y) < 0$ (10.4). A C halmaz x pontbeli K_0 normális kúpja tehát éppen $-K^*$, ahol

$$K := \{y : f'(x; y) < 0\} \neq \emptyset.$$

Újra 8.20 és 10.6 szerint

$$\begin{aligned} \mathrm{cl}\, K &= \{y : \mathrm{cl}_y f'(x; y) \leq 0\} = \{y : \mathrm{supp}_{\partial f(x)}(y) \leq 0\} = \\ &= \{y : a^T y \leq 0 \ (a \in \partial f(x))\} = -K_1^*, \end{aligned}$$

ahol K_1 a $\partial f(x)$ konvex halmaz által generált konvex kúp (amely éppen a $\partial f(x)$ halmaz elemeinek nemnegatív konstansszorosaiból áll). Ezek szerint

$$K_0 = -K^* = -(\mathrm{cl}\, K)^* = K_1^{**} = \mathrm{cl}\, K_1,$$

amit bizonyítanunk kellett. □

10.10. Következmény: *Legyen f valódi konvex függvény, és legyen $x \in \mathrm{int}\,(\mathrm{dom}\, f)$ olyan pont, ahol az f függvény nem veszi fel a minimumát. Ekkor az a vektor pontosan akkor eleme a $C := \{z : f(z) \leq f(x)\}$ halmaz x pontbeli normális kúpjának, ha létezik $\lambda \geq 0$ szám úgy, hogy $a \in \lambda \partial f(x)$.*

Bizonyítás: A feltevés miatt, 10.8 szerint $\partial f(x)$ nemüres, kompakt, konvex halmaz, amelynek nem eleme az origó. Ebben az esetben a $\partial f(x)$ halmaz által generált konvex kúp lezártja egyszerűen a $\lambda\partial f(x)$ $(\lambda \geq 0)$ halmazok úniója (4.19). $\square$

A szubgradiensekre vonatkozó felcserélhetőségi tételek következnek, a bizonyításukhoz alapvető alábbi segédtétel és következményeinek kimondása után.

10.11. Tétel: *Legyen f valódi konvex függvény, továbbá x tetszőleges pont. Ekkor az alábbi négy, valamely a vektorra vonatkozó feltétel ekvivalens egymással:*
a) $a \in \partial f(x)$;
b) az $a^T . - f(.)$ függvény az x pontban veszi fel a maximumát;
c) $f(x) + f^c(a) \leq a^T x$;
d) $f(x) + f^c(a) = a^T x$.

Ha $(\mathrm{cl}\, f)(x) = f(x)$, akkor még az alábbi három feltétel is ekvivalens a fentiekkel:
a^c) $x \in \partial f^c(a)$;
b^c) az $x^T . - f^c(.)$ függvény az a pontban veszi fel a maximumát;
a^{cc}) $a \in \partial(\mathrm{cl}\, f)(x)$.

Bizonyítás: Az a) feltételt definiáló szubgradiens egyenlőtlenség átírható, mint

$$a^T x - f(x) \geq a^T z - f(z) \ (z \in \mathcal{R}^n).$$

Ez éppen a b) feltétel. Mivel a b) feltételben szereplő szuprémum definíció szerint éppen $f^c(a)$, azért b) ugyanaz, mint c), vagy d). Másfelől a^c), b^c) és a^{cc}) ekvivalensek az

$$f^{cc}(x) + f^c(a) = a^T x$$

feltétellel, amely viszont éppen d), ha $f(x) = (\mathrm{cl}\, f)(x) = f^{cc}(x)$. $\square$

10.12. Következmény: *Ha f zárt, valódi konvex függvény, akkor $x \in \partial f^c(a)$ pontosan akkor, ha $a \in \partial f(x)$.* $\square$

10.13. Következmény: *Ha f valódi konvex függvény, és x olyan pont, ahol f szubdifferenciálható, akkor*

$$(\mathrm{cl}\, f)(x) = f(x), \text{ és } \partial(\mathrm{cl}\, f)(x) = \partial f(x).$$

Bizonyítás: Általában

$$f(x) \geq (\mathrm{cl}\, f)(x) = f^{cc}(x) \geq a^T x - f^c(a).$$

Ha f szubdifferenciálható az x pontban, akkor létezik legalább egy a vektor úgy, hogy teljesül a 10.11-beli d) feltétel. Ekkor $f(x) = (\mathrm{cl}\, f)(x)$, következésképpen $\partial(\mathrm{cl}\, f)(x) = \partial f(x)$ az a) és a^{cc}) feltételek ekvivalenciája miatt. □

10.14. Következmény: *Legyen C nemüres, zárt, konvex halmaz. Ekkor bármely a vektor esetén $\partial\mathrm{supp}_C(a)$ azokból az $x \in C$ vektorokból áll (ha ilyen van egyáltalán), amelyekben az $a^T.$ lineáris függvény felveszi a maximumát.*

Bizonyítás: Legyen $f := \mathrm{ind}_C$, ekkor $f^c = \mathrm{supp}_C$. A 10.11-beli a^c) és b) feltételek ekvivalenciája bizonyítja az állítást. □

10.15. Következmény: *Legyen K zárt, konvex kúp. Ekkor $a \in \partial\mathrm{ind}_K(x)$ pontosan akkor, ha $x \in \partial\mathrm{ind}_{-K^*}(a)$. Ezek a feltételek még azzal is ekvivalensek, hogy*

$$x \in K,\ a \in -K^*,\ a^T x = 0.$$

Bizonyítás: Legyen $f := \mathrm{ind}_K$, ekkor $f^c = \mathrm{ind}_{-K^*}$. A 10.11-beli a), a^c) és d) feltételek ekvivalenciája bizonyítja az állítást. □

A szubgradiens definíciójának azonnali következménye, hogy

$$\partial(\lambda f)(x) = \lambda \partial f(x)\ (x \in \mathcal{R}^n, \lambda > 0).$$

Egy meglepőbb felcserélhetőségi tényt fogalmaz meg az alábbi

10.16. Tétel: *Legyenek $f_1, \ldots, f_m$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények, és legyen $f := f_1 + \ldots + f_m$. Ekkor*

$$\partial f(x) \supseteq \partial f_1(x) + \ldots + \partial f_m(x),$$

egyenlőséggel, ha a $\mathrm{ri}\,(\mathrm{dom}\, f_i)$ $(i = 1, \ldots, m)$ halmazok összemetszenek. Az egyenlőség eme elégséges feltétele kicsit gyengíthető, ha a szereplő függvények némelyike, mondjuk $f_1, \ldots, f_k$ poliédrikus. Ekkor az egyenlőséghez elegendő, hogy a $\mathrm{dom}\, f_i$ $(i = 1, \ldots, k)$ és a $\mathrm{ri}\,(\mathrm{dom}\, f_i)$ $(i = k+1, \ldots, m)$ halmazok összemetszenek.

Bizonyítás: Ha $a = a_1 + \ldots + a_m$, ahol $a_i \in \partial f_i(x)$, akkor minden z esetén

$$\begin{aligned} f(z) &= f_1(z) + \ldots + f_m(z) \geq \\ &\geq f_1(x) + a_1^T(z-x) + \ldots + f_m(x) + a_m^T(z-x) = \\ &= f(x) + (a_1 + \ldots + a_m)^T(z-x) = f(x) + a^T(z-x), \end{aligned}$$

és így $a \in \partial f(x)$. Ezzel beláttuk az általános tartalmazást. Tegyük fel most, hogy a $\mathrm{ri}\,(\mathrm{dom}\, f_i)$ halmazok összemetszenek, ekkor 9.19 szerint

$$f^c(a) = \inf\{f_1^c(a_1) + \ldots + f_m^c(a_m) : a_1 + \ldots + a_m = a\},$$

ahol minden a esetén az infimum felvétetik (esetleg ∞-ként). Így 10.11 szerint $a \in \partial f(x)$ pontosan akkor, ha

$$\begin{aligned} a^T x &= f_1(x) + \ldots + f_m(x) + \\ &\quad + \inf\{f_1^c(a_1) + \ldots + f_m^c(a_m) : a_1 + \ldots + a_m = a\}, \end{aligned}$$

ahol minden a esetén az infimum felvétetik valamely $a_1, \ldots, a_m$ vektorok mellett. Ezért $\partial f(x)$ pontosan az olyan $a_1 + \ldots + a_m$ alakú vektorokból áll, amelyekre

$$a_1^T x + \ldots + a_m^T x = f_1(x) + \ldots + f_m(x) + f_1^c(a_1) + \ldots + f_m^c(a_m).$$

Mivel $a_i^T x \leq f_i(x) + f_i^c(a_i)$ mindig teljesül, egyenlőséggel pontosan akkor, ha $a_i \in \partial f_i(x)$, azért $\partial f(x)$ ugyanaz, mint a $\partial f_1(x) + \ldots + \partial f_m(x)$ halmaz. Az az eset, mikor néhány f_i függvény poliédrikus, hasonlóan intézhető el, csak 9.19 helyett most 9.22-t használva. □

Megjegyezzük, hogy 10.16 (és 10.19) a következő fejezet eredményeiből is levezethető.

10.17. Következmény: *Legyenek $C_1, \ldots, C_m \subseteq \mathcal{R}^n$ konvex halmazok, amelyeknek relatív belsejei összemetszenek. Ekkor a $C_1 \cap \ldots \cap C_m$ halmaz normális kúpja bármely adott x pontban megegyezik a $K_1 + \ldots + K_m$ kúppal, ahol K_i a C_i halmaz normális kúpja az x pontban. Ha a szereplő konvex halmazok némelyike, mondjuk $C_1, \ldots, C_k$ poliéderek, akkor a konklúzió fennáll akkor is, ha csak a $C_1, \ldots, C_k, \mathrm{ri}\, C_{k+1}, \ldots, \mathrm{ri}\, C_m$ halmazok metszenek össze.*

Bizonyítás: Alkalmazzuk 10.16-ot az $f_i := \mathrm{ind}_{C_i}$ indikátorfüggvényekre. □

10.18. Tétel: *Legyenek $f_1, \ldots, f_m$ az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvények, és legyen $f := \max\{f_1, \ldots, f_m\}$. Ekkor*

$$\partial f(x) \supseteq \operatorname{conv}\left(\cup\{\partial f_i(x) : 1 \le i \le m,\ f_i(x) = f(x)\}\right),$$

egyenlőséggel, ha a $\operatorname{dom} f_i$ halmazok lezártja ugyanaz a C halmaz.

Bizonyítás: Az általános tartalmazáshoz 4.30 szerint annyit kell megmutatnunk, hogy tetszőleges λ_i $(1 \le i \le m,\ f_i(x) = f(x))$ nemnegatív, 1-összegű számok, továbbá $a_i \in \partial f_i(x)$ vektorok esetén

$$f(z) \ge f(x) + (\sum \lambda_i a_i)^T (z - x) \ (z \in \mathcal{R}^n).$$

Ez az egyenlőtlenség pedig a nyilvánvaló

$$\begin{aligned} f(z) \ge f_i(z) \ge f_i(x) + a_i^T(z - x) = f(x) + a_i^T(z - x) \\ (z \in \mathcal{R}^n;\ 1 \le i \le m,\ f_i(x) = f(x)) \end{aligned}$$

egyenlőtlenségek konvex kombinációja a λ_i együtthatókkal.

Ha most minden i esetén $\operatorname{cl}(\operatorname{dom} f_i) = C$, akkor az $f^c(a)$ érték a 9.24-beli képlettel számolható, ahol az infimum minden a vektor esetén felvétetik (esetleg ∞-ként). Mivel 10.11 szerint $a \in \partial f(x)$ pontosan akkor, ha $a^T x = f(x) + f^c(a)$, azért $a \in \partial f(x)$ esetén léteznek λ_i $(i = 1, \ldots, m)$ nemnegatív, 1-összegű számok és $a_i \in \mathcal{R}^n$ vektorok úgy, hogy

$$a = \lambda_1 a_1 + \ldots + \lambda_m a_m, \text{ és } f^c(a) = \lambda_1 f_1^c(a_1) + \ldots + \lambda_m f_m^c(a_m).$$

De akkor a Fenchel-egyenlőtlenség és a fentiek szerint

$$\begin{aligned} &\sum \lambda_i (f_i(x) + f_i^c(a_i)) \ge \sum \lambda_i a_i^T x = \\ &\quad = a^T x = f(x) + \sum \lambda_i f_i^c(a_i) \ge \sum \lambda_i (f_i(x) + f_i^c(a_i)). \end{aligned}$$

Látszik, hogy végig egyenlőségnek kell lenni, amiből már következik, hogy $\lambda_i = 0$, ha $f(x) > f_i(x)$, továbbá $a_i^T x = f_i(x) + f_i^c(a_i)$, vagyis $a_i \in \partial f_i(x)$, ha $\lambda_i > 0$. □

10.19. Tétel: *Legyen $f := h \circ A$, ahol h az $\mathcal{R}^m$ téren értelmezett, valódi konvex függvény, $A \in \mathcal{R}^{m \times n}$. Ekkor*

$$\partial f(x) \supseteq A^T \partial h(Ax) \ (x \in \mathcal{R}^n),$$

egyenlőséggel, ha az $\operatorname{Im} A$ altér belemetsz a $\operatorname{ri}(\operatorname{dom} h)$ halmazba, vagy ha h poliédrikus függvény, és az $\operatorname{Im} A$ altér belemetsz a $\operatorname{dom} h$ halmazba.

Bizonyítás: Ha $a \in A^T \partial h(Ax)$, akkor $a = A^T b$ valamely $b \in \partial h(Ax)$ esetén. Ekkor minden $z \in \mathcal{R}^n$ esetén

$$f(z) = h(Az) \geq h(Ax) + b^T(Az - Ax) = f(x) + a^T(z - x),$$

ami az általános tartalmazást bizonyítja. Másfelől tegyük fel, hogy az $\operatorname{Im} A$ altér belemetsz a $\operatorname{ri}(\operatorname{dom} f)$ halmazba. Ekkor f valódi, és

$$f^c(a) = \inf\left\{h^c(b) : A^T b = a\right\}$$

9.16 szerint, ahol az infimum minden olyan a esetén felvétetik valamely b mellett, amelyre $f^c(a) \neq \infty$. Adott $a \in \partial f(x)$ esetén

$$f(x) + f^c(a) = a^T x$$

10.11 szerint, így létezik b vektor, amelyre $A^T b = a$, és

$$f(x) + h^c(b) = x^T A^T b.$$

Ez a feltétel úgy fogalmazható át, hogy

$$h(Ax) + h^c(b) = (Ax)^T b,$$

vagyis más szóval $b \in \partial h(Ax)$ ismét 10.11 szerint. Ezért $a \in A^T \partial h(Ax)$. Az az eset, mikor h poliédrikus, és $Ax \in \operatorname{dom} h$ valamely x esetén, hasonlóan intézhető el, vö. a 9.18 utáni megjegyzéssel. □

Rátérünk a differenciálható, konvex függvények és különféle jellemzéseik vizsgálatára. Megjegyezzük, hogy ha egy konvex függvény differenciálható valamely x pontban, akkor szükségképpen valódi, hiszen x valamely nyílt környezetén véges értékeket vesz fel, és ez a környezet biztosan belemetsz a $\operatorname{ri}(\operatorname{dom} f)$ halmazba (sőt az $\operatorname{int}\operatorname{dom} f$ halmaz része).

A következő tétel a differenciálhatóság a szubdifferenciálhatóságon kívüli másik "feléről" szól, ahogy a fejezet elején megígértük.

10.20. Tétel: *Legyen f konvex függvény. Ha az f függvény differenciálható egy x pontban, akkor $\nabla f(x)$ az egyetlen szubgradiense az x pontban, speciálisan*

$$f(z) \geq f(x) + (\nabla f(x))^T (z - x) \ (z \in \mathcal{R}^n).$$

Megfordítva, ha az f függvénynek éppen egy szubgradiense létezik egy x pontban, akkor f differenciálható az x pontban.

Bizonyítás: Tegyük fel először, hogy f differenciálható az x pontban. Ekkor $f'(x;.)$ nem más, mint a $(\nabla f(x))^T.$ lineáris függvény. 10.6 szerint az x pontbeli szubgradiensek éppen azok az a vektorok, amelyekre

$$(\nabla f(x))^T y \geq a^T y \ (y \in \mathcal{R}^n),$$

ez a feltétel pedig pontosan akkor teljesül, ha $a = \nabla f(x)$. Ezért $\nabla f(x)$ az egyetlen szubgradiens az x pontban. Másfelől tegyük fel, hogy az f függvénynek egyértelmű x pontbeli szubgradiense az a vektor. Ekkor a

$$g : y \mapsto f(x+y) - f(x) - a^T y \ (y \in \mathcal{R}^n)$$

konvex függvénynek a 0 vektor az egyértelmű szubgradiense a 0 pontban. Azt kell megmutatnunk, hogy ebből

$$\lim_{y \to 0} \frac{g(y)}{||y||} = 0$$

következik. A $g'(0;.)$ függvény lezártja éppen a $\partial g(0)$ halmaz támaszfüggvénye, ami most az azonosan 0 függvény (10.6). Ezért $g'(0;.)$ maga is az azonosan 0 függvény, mivel $g'(0;.)$ nem különbözhet a lezártjától máshol, mint effektív tartománya relatív határpontjaiban. Oda jutottunk, hogy

$$0 = g'(0;u) = \lim_{\lambda \to 0+} (g(\lambda u) - g(0))/\lambda \ (u \in \mathcal{R}^n).$$

Itt $g(0) = 0$, és a differenciahányados a λ változó nemcsökkenő függvénye. A

$$h_\lambda : u \mapsto g(\lambda u)/\lambda \ (u \in \mathcal{R}^n), \ (\lambda > 0)$$

függvények konvexek (epigráfjuk az epi g konstansszorosa), és pontonként 0-hoz csökkennek, ahogy $\lambda > 0$ a nullához csökken. Legyen O az $\mathcal{R}^n$ térbeli zárt egységgömb, és legyen $\{a_1, \ldots, a_m\}$ $\mathcal{R}^n$-beli pontoknak egy véges halmaza úgy, hogy konvex burkuk tartalmazza az O gömböt. Ekkor minden $u \in O$ pont kifejezhető, mint az a_i pontok konvex kombinációja valamely λ_i együtthatókkal, és akkor

$$0 \leq h_\lambda(u) \leq \sum_{i=1}^{m} \lambda_i h_\lambda(a_i) \leq \max\{h_\lambda(a_i) : i = 1, \ldots, m\}.$$

Mivel a $h_\lambda(a_i)$ értékek 0-hoz tartanak, ahogy $\lambda \to 0+$, azért látjuk, hogy adott $\varepsilon > 0$ esetén létezik $\delta > 0$ úgy, hogy

$$g(\lambda u)/\lambda \leq \varepsilon \ (0 < \lambda \leq \delta, \, u \in O).$$

Mivel minden y vektor, amelyre $0 < ||y|| \leq \delta$, kifejezhető, mint λu, $\lambda = ||y||$ és $u \in O$ mellett, azért $g(y)/||y|| \leq \varepsilon$ valahányszor $0 < ||y|| \leq \delta$. Ez már azt bizonyítja, hogy $g(y)/||y||$ limesze a 0-ban valóban 0, amit bizonyítanunk kellett. $\square$

Bizonyítás nélkül kimondjuk a konvex függvények majdnem mindenütt differenciálhatóságáról szóló alábbi tételt (lásd [20], Theorem 25.5).

10.21. Tétel: *Legyen f az $\mathcal{R}^n$ téren értelmezett, valódi konvex függvény, és legyen S azoknak a pontoknak a halmaza, ahol f differenciálható. Ekkor S az $\mathrm{int}\,(\mathrm{dom}\, f)$ halmaz sűrű részhalmaza, és kivonva az $\mathrm{int}\,(\mathrm{dom}\, f)$ halmazból, nullmértékű halmazt kapunk. Továbbá a $\nabla f : x \mapsto \nabla f(x)$ gradiens-leképezés folytonos az S halmazon.*

Az iránymenti deriváltak segítségével nyerhetjük a differenciálható, konvex függvények alábbi fontos jellemzéseit.

10.22. Tétel: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \mathcal{R}$, $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, és tegyük fel, hogy f differenciálható a C halmazon (speciálisan $C \subseteq \mathrm{int}\, S$). Ekkor az f függvény pontosan akkor konvex a C halmazon, ha teljesül, hogy*

$$f(z) - f(x) \geq (\nabla f(x))^T (z - x) \ (x, z \in C).$$

Bizonyítás: Tegyük fel először, hogy az f függvény differenciálható és konvex a C halmazon. Ekkor tetszőleges $x, z \in C$ vektorok, $0 < \lambda < 1$ szám esetén

$$f(x + \lambda(z - x)) \leq \lambda f(z) + (1 - \lambda) f(x),$$

amiből

$$\frac{f(x + \lambda(z - x)) - f(x)}{\lambda} \leq f(z) - f(x)$$

adódik. Tartsunk λ-val 0-hoz, ekkor f differenciálhatósága miatt a kívánt

$$f(z) - f(x) \geq (\nabla f(x))^T (z - x) \ (x, z \in C)$$

egyenlőtlenséghez jutunk.

A megfordításhoz tegyük fel, hogy teljesül a tételbeli egyenlőtlenség. Legyen $x, y \in C$, $0 \leq \lambda \leq 1$, és $z := \lambda x + (1 - \lambda) y$. Ekkor a feltétel szerint

$$f(x) - f(z) \geq (\nabla f(z))^T (x - z), \text{ és } f(y) - f(z) \geq (\nabla f(z))^T (y - z).$$

Az első egyenlőtlenséget λ-val, a második egyenlőtlenséget $(1-\lambda)$-val szorozva, majd a kapott egyenlőtlenségeket összeadva oda jutunk, hogy

$$\lambda f(x) + (1-\lambda)f(y) - f(z) \geq (\nabla f(z))^T(\lambda x + (1-\lambda)y - z) = 0,$$

vagyis f konvexitása látszik. $\square$

10.23. Tétel: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \mathcal{R}$, $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, és tegyük fel, hogy f differenciálható a C halmazon (speciálisan $C \subseteq \operatorname{int} S$). Ekkor az f függvény pontosan akkor konvex a C halmazon, ha teljesül, hogy*

$$(\nabla f(z) - \nabla f(x))^T(z-x) \geq 0 \ (x, z \in C).$$

Bizonyítás: Először is, ha az f függvény konvex és differenciálható a C konvex halmazon, akkor 10.22 szerint tetszőleges $x, z \in C$ esetén

$$\begin{array}{rcl} f(z) - f(x) & \geq & (\nabla f(x))^T(z-x), \\ f(x) - f(z) & \geq & -(\nabla f(z))^T(z-x), \end{array}$$

és ezeket az egyenlőtlenségeket összeadva a kívánt

$$(\nabla f(z) - \nabla f(x))^T(z-x) \geq 0 \ (x, z \in C)$$

egyenlőtlenséghez jutunk.

Másodszor tegyük fel, hogy teljesül az utóbbi egyenlőtlenség, és legyenek $x, z \in C$ tetszőleges pontok. Ekkor a Lagrange-féle középértéktétel szerint létezik $y \in]x, z[$ pont úgy, hogy

$$f(z) - f(x) = (\nabla f(y))^T(z-x),$$

legyen ez a pont éppen $y = x + \kappa(z-x)$, ahol $0 < \kappa < 1$. A fenti egyenlőtlenséget az $y, x \in C$ pontpárra alkalmazva kapjuk, hogy

$$(\nabla f(x + \kappa(z-x)) - \nabla f(x))^T(\kappa(z-x)) \geq 0,$$

ahonnan a pozitív κ számmal osztva a

$$(\nabla f(y))^T(z-x) \geq (\nabla f(x))^T(z-x)$$

egyenlőtlenség adódik. A fentiek szerint

$$f(z) - f(x) \geq (\nabla f(x))^T (z - x) \ (x, z \in C),$$

ami 10.22 szerint éppen azt jelenti, hogy az f függvény konvex a C halmazon. $\square$

Speciálisan ha $S \subseteq \mathcal{R}$ nyílt halmaz, $f : S \to \mathcal{R}$ a $C \subseteq S$ nem elfajuló intervallumon differenciálható és konvex függvény, akkor az f' deriváltfüggvény monoton növekedő a C halmazon. Ha még az is teljesül, hogy C belső pontjaiban létezik az f' deriváltja, továbbá f' folytonos C-n, akkor, mint azt az egyváltozós függvények analíziséből tudjuk, f' monoton növekedése a C halmazon azzal ekvivalens, hogy C minden belső x pontjában $f''(x) \geq 0$. (A folytonossági feltétel nem szükséges, ha C nyílt intervallum.)

8.14 szerint egy az $\mathcal{R}^n$ téren értelmezett függvény konvexitásának szükséges és elégséges feltétele bizonyos egyváltozós függvények konvexitása. Ezért a fenti észrevétel segítségével a konvexitás egy újabb jellemzését nyerhetjük.

Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \mathcal{R}$, továbbá $x_0 \in S$. Azt mondjuk, hogy az f függvény **kétszer differenciálható** az x_0 pontban, ha f differenciálható az x_0 pont egy környezetén (speciálisan $x_0 \in \operatorname{int} S$) és létezik $N \in \mathcal{R}^{n \times n}$ mátrix úgy, hogy

$$\lim_{z \to 0} \frac{\nabla f(x_0 + z) - \nabla f(x_0) - Nz}{||z||} = 0.$$

Ilyen N mátrixból legfeljebb egy lehet, az f függvény x_0 pontbeli **Hesse-mátrix**ának nevezzük, jele $\nabla^2 f(x_0)$. (Ha az f függvény kétszer differenciálható az x_0 pontban, akkor a $\nabla^2 f(x_0)$ Hesse-mátrix szimmetrikus (Young tétele), és

$$\lim_{z \to 0} \frac{f(x_0 + z) - f(x_0) - (\nabla f(x_0))^T z - \frac{1}{2} z^T \nabla^2 f(x_0) z}{||z||^2} = 0$$

(infinitezimális Taylor-formula). Ha f kétszer differenciálható az x_0 pontban, akkor $\nabla f(x_0)$ helyébe más vektort, vagy $\nabla^2 f(x_0)$ helyébe más szimmetrikus mátrixot írva a formula nem marad érvényben (lásd [7], [15]).)

Például legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $f : S \to \mathcal{R}$ a $C \subseteq S$ konvex halmazon (vagyis minden pontjában) kétszer differenciálható függvény. Adott $x, z \in \mathcal{R}^n$ vektorok esetén jelölje $S_{x,z}$ a $\{\lambda \in \mathcal{R} : x + \lambda z \in S\}$ halmazt, és definiáljuk hasonlóan a $C_{x,z}$ intervallumot. Jelölje továbbá $g_{x,z}$ azt az $S_{x,z}$ halmazon értelmezett függvényt, amely $\lambda \in S_{x,z}$ esetén $f(x + \lambda z)$ értéket vesz fel. A definíciókból könnyen belátható, hogy $g_{x,z}$ a $C_{x,z}$ intervallumon kétszer differenciálható függvény lesz, amelynek gradiense és Hesse-mátrixa az alábbi módon számolható egy $\lambda \in C_{x,z}$ pontban:

$$\nabla g_{x,z}(\lambda) = (\nabla f(x + \lambda z))^T z, \text{ és } \nabla^2 g_{x,z}(\lambda) = z^T \nabla^2 f(x + \lambda z) z.$$

10.24. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $f : S \to \mathcal{R}$ a $C \subseteq S$ konvex halmazon kétszer differenciálható, konvex függvény. Tegyük fel még, hogy a $\nabla^2 f(.) : C \to \mathcal{R}^{n^2}$ leképezés folytonos a C halmazon (erre a feltételre nincs szükség, ha a C halmaz relatív nyílt). Ekkor az f függvény pontosan akkor konvex a C halmazon, ha teljesül, hogy*

$$z^T \nabla^2 f(x) z \geq 0 \ (x \in C, \ z \in \operatorname{par} C)$$

Speciálisan ha $\operatorname{int} C \neq \emptyset$, akkor a feltétel a $\nabla^2 f(x)$ $(x \in C)$ Hesse–mátrixok pozitív szemidefinitását követeli.

Bizonyítás: A fenti jelölésekkel, 8.14 szerint f pontosan akkor konvex a C halmazon, ha a $g_{x,z}$ függvények konvexek a $C_{x,z}$ intervallumon, vagyis $g'_{x,z}$ monoton növekedő a $C_{x,z}$ intervallumon, vagyis $g''_{x,z}$ nemnegatív a $C_{x,z}$ intervallum belső pontjaiban. A folytonossági feltétel miatt itt a "belső" szó el is hagyható. Összefoglalva az f függvény pontosan akkor konvex a C halmazon, ha

$$z^T \nabla^2 f(x + \lambda z) z \geq 0 \ (\lambda \in C_{x,z})$$

valahányszor $C_{x,z} \subseteq \mathcal{R}$ belseje nemüres. Az utóbbi feltétel pedig könnyen láthatóan éppen a tételbelivel ekvivalens. $\square$

Legyen $S \subseteq \mathcal{R}^n$, továbbá $C \subseteq S$ konvex halmaz. Azt mondjuk, hogy egy $f : S \to \mathcal{R}$ függvény **szigorúan konvex** a C halmazon, ha

$$f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y)$$

valahányszor $x, y \in C$, $x \neq y$ és $0 < \lambda < 1$.

10.22, 10.23 és 10.24 bizonyítása megfelelő változtatásokkal elmondható konvex helyett szigorúan konvex függvényekre is. Így jutunk az alábbi tételekhez:

10.25. Tétel: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \mathcal{R}$, $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, és tegyük fel, hogy f differenciálható a C halmazon (speciálisan $C \subseteq \operatorname{int} S$). Ekkor az f függvény pontosan akkor szigorúan konvex a C halmazon, ha teljesül, hogy*

$$f(z) - f(x) > (\nabla f(x))^T (z - x) \ (x, z \in C, \, x \neq z).$$

$\square$

10.26. Tétel: *Legyen $S \subseteq \mathcal{R}^n$, $f : S \to \mathcal{R}$, $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, és tegyük fel, hogy f differenciálható a C halmazon (speciálisan $C \subseteq \operatorname{int} S$). Ekkor az f függvény pontosan akkor szigorúan konvex a C halmazon, ha teljesül, hogy*

$$(\nabla f(z) - \nabla f(x))^T (z - x) > 0 \ (x, z \in C, \, x \neq z).$$

$\square$

10.27. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $f : S \to \mathcal{R}$ a $C \subseteq S$ konvex halmazon kétszer differenciálható, konvex függvény. Tegyük fel még, hogy a $\nabla^2 f(.) : C \to \mathcal{R}^{n^2}$ leképezés folytonos a C halmazon (erre a feltételre nincs szükség, ha a C halmaz relatív nyílt). Ha teljesül, hogy*

$$z^T \nabla^2 f(x) z > 0 \ (x \in C, \, 0 \neq z \in \operatorname{par} C)$$

(speciálisan ha a $\nabla^2 f(x)$ $(x \in C)$ Hesse–mátrixok pozitív definitek, vagyis a fenti egyenlőtlenség minden $z \neq 0$ esetén teljesül), akkor az f függvény szigorúan konvex a C halmazon. $\square$

Az erősen konvex függvény definíciója és a 10.22, 10.23 és 10.24 tételeknek megfelelő jellemzései megtalálhatók [13]-ban.

10.24 [10.27] és a következő észrevétel segítségével (amelyet a benne szereplő monotonitási feltétel miatt tárgyalunk a differenciálhatósági témakörben) már számos alapvető példa adható [szigorúan] konvex függvényekre. (Az egyik legegyszerűbb ezek közül az $f := .^T N.$ [szigorúan] konvex kvadratikus függvény, ahol N pozitív szemidefinit [pozitív definit] mátrix.)

10.28. Állítás: *Legyen $C \subseteq \mathcal{R}^n$ konvex halmaz, és $f : C \to \mathcal{R}$ a C halmazon konvex függvény, amelynek értékkészlete egy $S \subseteq \mathcal{R}$ intervallum része. Ha $g : S \to \mathcal{R}$ az S intervallumon monoton növő, konvex függvény, akkor $g \circ f$ konvex függvény a C halmazon.* $\square$

A fejezet végén differenciálható, konvex függvények lokális optimumait jellemezzük.

Tekintsük az
$$\inf f(x),\ x \in C$$
problémát, ahol $C \subseteq \mathcal{R}^n$ konvex halmaz, $f : C \to \mathcal{R}$ a C halmazon konvex függvény. Azt mondjuk, hogy az $x_0 \in C$ pont a fenti program **lokális optimuma**, ha létezik $\varepsilon > 0$ szám úgy, hogy

$$f(x) \geq f(x_0)\ (x \in C \cap O(x_0, \varepsilon))$$

teljesül. A fenti probléma lokális optimumai globális optimumok is (vagyis optimális megoldások). Ha ugyanis $x_0 \in C$ lokális optimum (a fenti egyenlőtlenségek fennállnak), akkor tetszőleges $x \in C$ esetén létezik elég kis $\lambda > 0$ szám úgy, hogy

$$y := x_0 + \lambda(x - x_0) \in C \cap O(x_0, \varepsilon),$$

és akkor $f(x) < f(x_0)$ esetén az

$$f(x_0) \leq f(y) \leq \lambda f(x) + (1 - \lambda) f(x_0) < f(x_0)$$

ellentmondáshoz jutnánk. Megjegyezzük még, hogy az optimális megoldások halmaza nívóhalmaz, így konvex, továbbá (nyilvánvalóan) legfeljebb egy eleme lehet, ha f szigorúan konvex függvény.

Láttuk, hogy ha $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konvex függvény, akkor $x_0 \in \mathcal{R}^n$ pontosan akkor optimális megoldása az $\inf f(x)$, $x \in \mathcal{R}^n$ programnak, ha $0 \in \partial f(x_0)$. (Ezért ha f még zárt is, akkor az optimális megoldások halmaza $\partial f^c(0)$, vö. 10.12.) Vizsgáljuk most az

$$\inf f(x),\ x \in C$$

programot, ahol $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, amelynek relatív belseje belemetsz f effektív tartományának relatív belsejébe. A fentiek szerint x_0 ennek pontosan akkor optimális megoldása, ha $0 \in \partial(f + \mathrm{ind}_C)(x_0)$, vagyis (vö. 10.16)
$$0 \in \partial f(x_0) + \partial \mathrm{ind}_C(x_0),$$
azaz $x_0 \in C$, és $y^T(x - x_0) \geq 0$ $(x \in C)$ valamely $y \in \partial f(x_0)$ esetén. Mindez, együtt 10.20-szal sejteti ($x_0 \in \mathrm{int}\, C$ esetén bizonyítja) a következő igen hasznos eredményt:

10.29. Tétel: *Legyen* $S \subseteq \mathcal{R}^n$*,* $C \subseteq S$ *konvex halmaz,* $f : S \to \mathcal{R}$ *a* C *halmazon konvex függvény, továbbá* $x_0 \in C$*. Tegyük fel még, hogy az* f *függvény differenciálható az* x_0 *pontban. Ekkor* x_0 *pontosan akkor optimális megoldása az* $\inf f(x)$, $x \in C$ *programnak, ha teljesül, hogy*

$$(\nabla f(x_0))^T (x - x_0) \geq 0 \ (x \in C).$$

Speciálisan $x_0 \in \operatorname{int} C$ *esetén a fenti feltétel* $\nabla f(x_0)$ *eltűnését követeli meg.*

Bizonyítás: Először is, ha $x_0 \in C$ optimális megoldás, akkor tetszőleges $x \in C$ esetén

$$0 \leq \frac{f(x_0 + \lambda(x - x_0)) - f(x_0)}{\lambda} \ (0 < \lambda < 1),$$

ahonnan (tartsunk λ-val 0-hoz, jobbról!) az f függvény x_0 pontbeli differenciálhatósága miatt

$$0 \leq f'(x_0; x - x_0) = (\nabla f(x_0))^T (x - x_0)$$

adódik. (Eddig még a függvény konvexitását sem használtuk.)

A megfordításhoz tegyük fel, hogy fennáll a tételbeli egyenlőtlenség tetszőleges $x \in C$ esetén. Ekkor az f függvény konvexitásából adódóan (vö. 10.4)

$$\begin{aligned} 0 \leq (\nabla f(x_0))^T (x - x_0) \leq \tfrac{f(x_0 + \lambda(x - x_0)) - f(x_0)}{\lambda} \leq \\ \leq f(x) - f(x_0) \ (0 < \lambda < 1, \, x \in C), \end{aligned}$$

tehát x_0 optimális megoldás. □

Dualitási és optimalitási tételek

11. Rockafellar dualitási tétele

A Rockafellar primál-duál programpárra vonatkozó dualitáselmélet a kúplineáris dualitáselmélet jelentős általánosítása, lényegében konvex kúpindikátorfüggvényekről tetszőleges konvex függvényekre. A fejezet végén két példát említünk, és a következő fejezetbeli Lagrange-programpár is egy speciális eset.

Legyen $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$, továbbá $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konvex függvény, $g : \mathcal{R}^m \to \overline{\mathcal{R}}$ valódi konkáv függvény. (A szimmetria kedvéért két konvex függvény helyett egy konvex és egy konkáv függvénnyel dolgozunk. Emiatt a szokásosnál jobban kell ügyelni a definíciókra. Emlékeztetünk arra, hogy egy konkáv függvény zártsága felülről zártságát jelenti, effektív tartománya pedig a (-1)-szeresének effektív tartománya.) Tekintsük a

$$\begin{array}{ll} (P): & \inf f(x) - g(Ax-b) + c^T x,\ x \in \mathcal{R}^n, \\ (D): & \sup g_c(y) - f^c(A^T y - c) + b^T y,\ y \in \mathcal{R}^m \end{array}$$

Rockafellar primál-duál programpárt. Szokásos módon (P) (véges célfüggvényértékű) megengedett megoldásait jelölje $\mathbf{P}$, vagyis

$$\mathbf{P} := \{x \in \mathcal{R}^n : Ax - b \in \operatorname{dom} g,\ x \in \operatorname{dom} f\};$$

szigorúan megengedett megoldásainak halmazát jelölje $\mathbf{P_s}$, vagyis

$$\mathbf{P_s} := \{x \in \mathcal{R}^n : Ax - b \in \operatorname{ri}\operatorname{dom} g,\ x \in \operatorname{ri}\operatorname{dom} f\};$$

optimumértékét jelölje v_P, optimális megoldásainak halmazát pedig $\mathbf{P_o}$. Hasonlóan definiálható $\mathbf{D}$, $\mathbf{D_s}$, v_D és $\mathbf{D_o}$ is a duál program esetében.

Vegyük észre, hogy a (P) és a (D) program valóban primál-duál programpárt alkot abban az értelemben, hogy f, g zártsága esetén a $-(D)$ program duálja a $-(P)$ program (vö. 9.28). Ha f a K_2 kúp indikátorfüggvénye, $-g$ pedig a K_1 kúpé (amikor is $f^c = \mathrm{ind}_{-K_2^*}$, és $g_c = -\mathrm{ind}_{K_1^*}$), akkor a (P), (D) programpár éppen a hatodik fejezetben tárgyalt kúplineáris programpár.

11.1. Tétel: (gyenge dualitás) *A (P) program optimumértéke legalább akkora, mint a (D) program optimumértéke.*

Bizonyítás: A $v_P \geq v_D$ egyenlőtlenség triviálisan teljesül, ha (P) vagy (D) nem megoldható. Legyen most x a (P) program megengedett megoldása, y pedig a (D) programé. Ekkor f^c és g_c definíciójából adódik, hogy

$$\begin{aligned} f^c(A^T y - c) &\geq y^T A x - f(x) - c^T x, \\ g_c(y) &\leq y^T A x - g(Ax - b) - y^T b, \end{aligned}$$

amiből már következik, hogy az x primál megengedett megoldáshoz tartozó primál célfüggvényérték legalább akkora, mint az y duál megengedett megoldáshoz tartozó duál célfüggvényérték. □

Az is látszik, hogy ha x primál, y pedig duál megengedett megoldás, akkor a hozzájuk tartozó célfüggvényértékek pontosan akkor egyeznek meg, ha a fenti bizonyításbeli egyenlőtlenségek egyenlőséggel teljesülnek (amikor is x és y a megfelelő programok optimális megoldásai).

Az erős dualitási tételt most is egy megfelelően általános Farkas-tételből vezetjük le. Megvizsgálva a már elintézett speciális esetre vonatkozó 3.33 zártsági feltételét, természetesen adódik az alábbi

Primál zártsági feltétel: Tegyük fel, hogy az

$$\begin{pmatrix} A & 0 \\ c^T & 1 \end{pmatrix} \mathrm{epi}\, f - \mathrm{hipo}\, g$$

halmaz zárt.

Megjegyezzük (ez könnyen bizonyítható), hogy a (b, δ) vektor pontosan akkor eleme a primál zártsági feltételben szereplő különbséghalmaznak, ha létezik olyan $x \in \mathbf{P}$ vektor, amelyhez tartozó primál célfüggvényérték legfeljebb δ (vö. a 3.33 utáni megjegyzéssel). A 3.33 utáni megjegyzés analogonja itt is érvényes, mint az a következő bizonyításból kiolvasható.

11.2. Tétel: (Rockafellar-féle Farkas-tétel) *Tegyük fel, hogy a primál zártsági feltétel teljesül, továbbá* $\mathbf{P} \cup \mathbf{D} \neq \emptyset$*. Ekkor tetszőleges* $\delta \in \mathcal{R}$ *szám esetén az alábbi állítások ekvivalensek:*
a) létezik $x \in \mathcal{R}^n$ *vektor, amelyre* $f(x) - g(Ax - b) + c^T x \leq \delta$*;*
b) minden $y \in \mathcal{R}^m$ *vektor esetén* $g_c(y) - f^c(A^T y - c) + b^T y \leq \delta$*.*

Bizonyítás: A tétel "a)⇒b)" iránya a gyenge dualitási tételhez hasonlóan igazolható.

A fordított irányhoz tegyük fel indirekt, hogy b) teljesül, de a) nem. Ellentmondáshoz kell jutnunk.

A tétel előtti megjegyzés szerint, mivel a) nem teljesül, azért a (b, δ) vektor nem eleme a primál zártsági feltételben szereplő különbséghalmaznak.

Az első Hahn–Banach-tétel szerint létezik $a \in \mathcal{R}^n$ vektor és $\alpha \in \mathcal{R}$ szám úgy, hogy az $\hat{a} := (a^T, \alpha)^T$ jelöléssel

$$\hat{a}^T \begin{pmatrix} b \\ \delta \end{pmatrix} < \inf \hat{a}^T \left(\begin{pmatrix} A & 0 \\ c^T & 1 \end{pmatrix} \operatorname{epi} f - \operatorname{hipo} g \right),$$

vagyis

$$a^T b + \alpha\delta < \inf((a^T A + \alpha c^T, \alpha)\operatorname{epi} f) - \sup((a^T, \alpha)\operatorname{hipo} g).$$

Könnyen belátható (mivel $(0,1) \in \operatorname{rec} \operatorname{epi} f$), hogy itt $\alpha \geq 0$.

Ha $\alpha = 0$, akkor az

$$a^T b < \inf(a^T A \operatorname{dom} f) - \sup(a^T \operatorname{dom} g)$$

egyenlőtlenséghez jutunk. Vizsgáljuk először azt az esetet, mikor $\mathbf{P} \neq \emptyset$, legyen $x_0 \in \mathbf{P}$. Ekkor $x_0 \in \operatorname{dom} f$, és $Ax_0 - b \in \operatorname{dom} g$ miatt az $a^T b < a^T A x_0 - a^T(Ax_0 - b)$ ellentmondáshoz jutunk.

Tegyük fel most, hogy $\mathbf{D} \neq \emptyset$, legyen $y_0 \in \mathbf{D}$. Jelölje ε a fenti kiemelt egyenlőtlenségben szereplő két szám pozitív különbségét. Legyen $\lambda \geq 0$ olyan nagy szám, hogy $\delta < g_c(y_0) - f^c(A^T y_0 - c) + b^T y_0 + \lambda\varepsilon$ teljesül. Az utóbbi (δ-nál nagyobb) érték felülről becsülhető az $y_0 - \lambda a$ vektorhoz tartozó duál célfüggvényértékkel, ami ellentmond b)-nek.

Ha $\alpha > 0$, akkor feltehető, hogy $\alpha = 1$. Ekkor a

$$\delta < g_c(-a) - f^c(A^T(-a) - c) + b^T(-a)$$

egyenlőtlenséghez jutunk, ami ismét ellentmond b)-nek. □

Most a primál zártsági feltételhez keresünk elégséges feltételeket. Az erről szóló állítás bizonyítása előtt kimondunk egy egyszerűen igazolható lemmát:

11.3. Lemma: *Legyen $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konvex függvény. Ekkor*

$$\mathrm{bar}\,(\mathrm{epi}\, f) = \begin{pmatrix} \mathrm{bar}\,(\mathrm{dom}\, f) \\ 0 \end{pmatrix} \cup \mathrm{cone} \begin{pmatrix} -\mathrm{dom}\, f^c \\ 1 \end{pmatrix}.$$

Továbbá

$$\mathrm{ri}\,\mathrm{bar}\,(\mathrm{epi}\, f) = \{\lambda(a,1) : a \in -\mathrm{ri}\,(\mathrm{dom}\, f^c),\ \lambda > 0\}\,.$$

Bizonyítás: Az első egyenlőség egyszerű számolással igazolható. A második egyenlőséghez vegyük észre, hogy a $\mathrm{bar}\,(\mathrm{dom}\, f) \times \{0\}$ kúp a $\mathrm{bar}\,(\mathrm{epi}\, f)$ halmaz valódi exponált részhalmaza, így része az utóbbi halmaz relatív határának, elhagyva belőle, a maradék konvex halmaznak ugyanaz a relatív belseje, mint a $\mathrm{bar}\,(\mathrm{epi}\, f)$ halmaznak. A második egyenlőség ezek után 4.20 következménye. □

11.4. Állítás: *A primál zártsági feltétel teljesül, ha az alábbi állítások legalább egyike igaz*
a) f, g zárt függvények, és $\mathbf{D_s} \neq \emptyset$;
b) f zárt, g poliédrikus függvény, továbbá létezik $y_0 \in \mathcal{R}^m$ vektor, amelyre $y_0 \in \mathrm{dom}\, g_c$, $A^T y_0 - c \in \mathrm{ri}\,\mathrm{dom}\, f^c$;
c) f poliédrikus, g zárt függvény, továbbá létezik $y_0 \in \mathcal{R}^m$ vektor, amelyre $y_0 \in \mathrm{ri}\,\mathrm{dom}\, g_c$, $A^T y_0 - c \in \mathrm{dom}\, f^c$;
d) f, g poliédrikus függvények, továbbá $\mathbf{D} \neq \emptyset$.

Bizonyítás: Az a) állításhoz legyen

$$C_1 := \mathrm{epi}\, f,\ C_2 := \mathrm{hipo}\, g, \ \ \text{továbbá } \hat{A} := \begin{pmatrix} A & 0 \\ c^T & 1 \end{pmatrix}.$$

Ekkor a zártsági feltételek miatt C_1, C_2 zárt, konvex halmazok, és $\hat{A}C_1 - C_2$ a primál zártsági feltételben szereplő különbséghalmaz. Az $\hat{A}C_1 - C_2$ konvex halmaz 6.26 szerint zárt lesz, ha teljesül, hogy

$$(-\hat{A}^T \mathrm{ri}\,\mathrm{bar}\, C_2) \cap (\mathrm{ri}\,\mathrm{bar}\, C_1) \neq \emptyset.$$

A lemma segítségével könnyen belátható, hogy

$$\operatorname{ri}\operatorname{bar}\operatorname{epi} f = \operatorname{ri}\operatorname{cone}\begin{pmatrix} -\operatorname{dom} f^c \\ 1 \end{pmatrix},$$
$$\operatorname{ri}\operatorname{bar}\operatorname{hipo} g = \operatorname{ri}\operatorname{cone}\begin{pmatrix} \operatorname{dom} g_c \\ -1 \end{pmatrix}.$$

Legyen $y_0 \in \mathbf{D_s}$, $\hat{y}_0 := (y_0^T, -1)^T$. Könnyen ellenőrizhető, hogy e jelölésekkel

$$-\hat{A}^T \hat{y}_0 \in \operatorname{ri}\operatorname{bar}\operatorname{epi} f, \text{ és } \hat{y}_0 \in \operatorname{ri}\operatorname{bar}\operatorname{hipo} g.$$

Ekkor tehát a fentiek szerint teljesül a primál zártsági feltétel.

b), c) az a)-hoz hasonlóan intézhető el (csak most 6.27-et használva), d) nyilvánvaló abból, hogy poliéderek lineáris képe, összege is poliéder, és így zárt. □

Mivel a $-(D)$ program duálja a $-(P)$ program, azért 11.2 és 11.4 dualizálhatók.

Duál zártsági feltétel: Tegyük fel, hogy az

$$\begin{pmatrix} A^T & 0 \\ b^T & 1 \end{pmatrix} \operatorname{hipo} g_c - \operatorname{epi} f^c$$

halmaz zárt.

11.5. Tétel: *Tegyük fel, hogy az f, g függvények zártak, teljesül a duál zártsági feltétel, továbbá $\mathbf{P} \cup \mathbf{D} \neq \emptyset$. Ekkor tetszőleges $\delta \in \mathcal{R}$ szám esetén az alábbi állítások ekvivalensek:*
a) létezik $y \in \mathcal{R}^m$ vektor, amelyre $g_c(y) - f^c(A^T y - c) + b^T y \geq \delta$;
b) minden $x \in \mathcal{R}^n$ vektor esetén $f(x) - g(Ax - b) + c^T x \geq \delta$. □

11.6. Állítás: *A duál zártsági feltétel teljesül, ha az alábbi állítások legalább egyike igaz*
a) f, g zárt függvények, és $\mathbf{P_s} \neq \emptyset$;
b) f zárt, g poliédrikus függvény, továbbá létezik $x_0 \in \mathcal{R}^n$ vektor, amelyre $x_0 \in \operatorname{ri}\operatorname{dom} f$, $Ax_0 - b \in \operatorname{dom} g$;
c) f poliédrikus, g zárt függvény, továbbá létezik $x_0 \in \mathcal{R}^n$ vektor, amelyre $x_0 \in \operatorname{dom} f$, $Ax_0 - b \in \operatorname{ri}\operatorname{dom} g$;
d) f, g poliédrikus függvények, továbbá $\mathbf{P} \neq \emptyset$. □

Most már kúplineáris speciális esetéhez (6.34) hasonlóan bizonyítható az alábbi

11.7. Tétel: (Rockafellar erős dualitási tétele)
a) Tegyük fel, hogy teljesül a primál zártsági feltétel. Ekkor $\mathbf{P} \cup \mathbf{D} \neq \emptyset$ *esetén* $v_P = v_D$*, és a primál optimumérték felvétetik, ha véges.*
b) Tegyük fel, hogy f, g *zártak, továbbá teljesül a duál zártsági feltétel. Ekkor* $\mathbf{P} \cup \mathbf{D} \neq \emptyset$ *esetén* $v_P = v_D$*, és a duál optimumérték felvétetik, ha véges.*
c) Ha a 11.4-beli a), b), c), d) feltételek legalább egyike fennáll, akkor $v_P = v_D$*, és a primál optimumérték felvétetik, ha véges.*
d) Ha a 11.6-beli a), b), c), d) feltételek legalább egyike fennáll, akkor $v_P = v_D$*, és a duál optimumérték felvétetik, ha véges.* □

11.8. Tétel: (kiegészítő eltérések tétele) *Tegyük fel, hogy 11.7 a), b), c) vagy d) pontjának feltétele teljesül. Ekkor* $x \in \mathbf{P}$*,* $y \in \mathbf{D}$ *esetén az alábbi állítások ekvivalensek*
a) az x *és* y *vektorok a megfelelő programok optimális megoldásai;*
b) az x *és* y *vektorokhoz tartozó primál, illetve duál célfüggvényértékek megegyeznek;*
c) teljesül, hogy

$$\begin{aligned} f^c(A^T y - c) &= y^T A x - f(x) - c^T x, \\ g_c(y) &= y^T A x - g(Ax - b) - y^T b. \end{aligned}$$

Bizonyítás: Az erős dualitási tétel szerint a primál és duál optimumértékek megegyeznek. Az x, y megengedett megoldások ezért pontosan akkor optimálisak, ha a hozzájuk tartozó célfüggvényértékek megegyeznek. Ez mutatja a) és b) ekvivalenciáját, b) és c) ekvivalenciája a gyenge dualitási tétel utáni megjegyzésből adódik. □

A következő két tétel 6.29 és 6.31 általánosítása. Bizonyításukhoz szükségünk lesz az alábbi lemmára:

11.9. Lemma: *Legyen* f *az* $\mathcal{R}^n$ *téren értelmezett, valódi konvex függvény. Ekkor*

$$\operatorname{bar}\operatorname{dom} f = -\operatorname{dom}\operatorname{rec} f^c.$$

Bizonyítás: 9.12 nyilvánvaló következménye, hogy

$$\operatorname{bar}\operatorname{dom} f = -\operatorname{dom}\operatorname{supp}_{\operatorname{dom} f} = -\operatorname{dom}\operatorname{rec} f^c.$$

□

11.10. Tétel: *Tegyük fel, hogy* $f, -g$ *zárt, valódi konvex függvények,* (P) *megoldható, továbbá az*

$$A^T \operatorname{dom} g_c - \operatorname{dom} f^c$$

konvex halmaz zárt. Ekkor az alábbi állítások ekvivalensek:
a) a (P) *program korlátos, vagyis* $v_P > -\infty$*;*
b) nem létezik $z \in \mathcal{R}^n$ *vektor, amelyre* $(\operatorname{rec} f)(z) + (\operatorname{rec}(-g))(Az) + c^T z < 0$*;*
c) a duál program megoldható.

Speciálisan ha a $\operatorname{dom} f^c$, $\operatorname{dom} g_c$ *konvex halmazok zártak, továbbá*

$$(A \operatorname{ri}\operatorname{dom}\operatorname{rec} f) \cap \operatorname{ri}\operatorname{dom}\operatorname{rec}(-g) \neq \emptyset,$$

akkor a fenti a), b) és c) állítások ekvivalensek.

Bizonyítás: Az állítás "c)⇒a)" része a gyenge dualitási tétel következménye.

Az "a)⇒b)" irány szokásos módon indirekt igazolható. Ha létezne b)-ben leírt tulajdonságú z vektor, akkor egy tetszőleges primál megengedett megoldásból induló, z irányú félegyenes mentén $-\infty$-hez tartana a primál célfüggvény.

Most megmutatjuk, hogy a tétel zártsági feltétele mellett b)⇒c). Tegyük fel indirekt, hogy bár b) fennáll, mégsincs megengedett megoldása a (D) programnak, vagyis

$$(A^T \operatorname{dom}(g_c) - c) \cap \operatorname{dom}(f^c) = \emptyset.$$

5.10 szerint ekkor létezik $a \in \mathcal{R}^n$ vektor úgy, hogy

$$\inf a^T (A^T \operatorname{dom}(g_c) - c) > \sup a^T \operatorname{dom}(f^c).$$

9.12 és 9.27 segítségével könnyen belátható, hogy erre az a vektorra

$$-(\operatorname{rec}(-g))(Aa) - c^T a > (\operatorname{rec} f)(a),$$

ami ellentmond b)-nek.

A tétel utolsó állítása 11.4 mintájára igazolható, 11.9 segítségével. □

11.10 duális változata, 6.31 megfelelője az alábbi

11.11. Tétel: *Tegyük fel, hogy* $f, -g$ *zárt, valódi konvex függvények,* (D) *megoldható, továbbá az*

$$A\mathrm{dom}\, f - \mathrm{dom}\, g$$

konvex halmaz zárt. Ekkor az alábbi állítások ekvivalensek:
a) a (D) *program korlátos, vagyis* $v_D < \infty$*;*
b) nem létezik $w \in \mathcal{R}^m$ *vektor, amelyre* $(\mathrm{rec}\,(-g_c))(w) + (\mathrm{rec}\,(f^c))(A^T w) + b^T w > 0$*;*
c) a primál program megoldható.

Speciálisan ha a $\mathrm{dom}\, f$, $\mathrm{dom}\, g$ *konvex halmazok zártak, továbbá*

$$(A^T \mathrm{ri}\,\mathrm{dom}\,\mathrm{rec}\,(-g_c)) \cap \mathrm{ri}\,\mathrm{dom}\,\mathrm{rec}\,(f^c) \neq \emptyset,$$

akkor a fenti a), b) és c) állítások ekvivalensek. □

6.33 általánosításához a fenti két tételen kívül szükségünk lesz még az alábbi lemmára is:

11.12. Lemma: *Legyen* f *az* $\mathcal{R}^n$ *téren értelmezett, zárt, valódi konvex függvény. Tegyük fel, hogy a* $\mathrm{dom}\, f^c$ *konvex halmaz zárt és egyenesmentes (vagyis linealitástere a triviális* $\{0\}$ *altér). Ekkor*

$$x \in \mathrm{dom}\, f,\ z \in \mathrm{ri}\,\mathrm{dom}\,\mathrm{rec}\, f \ \textit{esetén}\ x + z \in \mathrm{ri}\,\mathrm{dom}\, f.$$

Bizonyítás: Először is 11.9, 4.7 és 4.11 szerint

$$\begin{aligned} z \in \mathrm{ri}\,\mathrm{dom}\,\mathrm{rec}\, f &= -\mathrm{ri}\,\mathrm{bar}\,\mathrm{dom}\, f^c = \\ &= -\mathrm{ri}\,\mathrm{cl}\,\mathrm{bar}\,\mathrm{dom}\, f^c = \\ &= -\mathrm{ri}\,(\mathrm{rec}\,\mathrm{dom}\, f^c)^*, \end{aligned}$$

így Carver tétele szerint

$$\left\{ y \in \mathrm{rec}\,\mathrm{dom}\, f^c : z^T y \geq 0 \right\} \subseteq \left\{ y \in -\mathrm{rec}\,\mathrm{dom}\, f^c : z^T y \leq 0 \right\}.$$

Most tegyük fel indirekt, hogy $x + z \notin \mathrm{ri}\,\mathrm{dom}\, f$, ekkor a második Hahn–Banach-tétel szerint létezik $a \in \mathcal{R}^n$ nemnulla vektor úgy, hogy

$$a^T z + a^T x \leq \inf a^T \mathrm{dom}\, f.$$

Erre az a vektorra nyilván teljesül, hogy $a^T z \leq 0$ (hiszen $x \in \operatorname{dom} f$), továbbá

$$a \in \operatorname{bar} \operatorname{dom} f = -\operatorname{dom} \operatorname{rec} f^c \subseteq -\operatorname{rec} \operatorname{dom} f^c$$

(itt újra 11.9-et használtuk, és azt a tényt, hogy az effektív tartomány az epigráf lineáris képe). A fentiek szerint $\pm a \in \operatorname{rec} \operatorname{dom} f^c$, ami ellentmond a $\operatorname{dom} f^c$ halmaz egyenesmentességének. (A bizonyításból az is látszik, hogy még az egyenesmentességi feltételt elhagyva is igaz lesz a gyengébb

$$\operatorname{ri} \operatorname{dom} \operatorname{rec} f \subseteq \operatorname{rec} \operatorname{ri} \operatorname{dom} f$$

konklúzió.) □

11.13. Tétel: *Tegyük fel, hogy az f, g függvények zártak.*
a) Ha az

$$A^T \operatorname{dom} g_c - \operatorname{dom} f^c$$

konvex halmaz zárt, akkor (P) megoldhatósága és korlátossága esetén (D) megoldható lesz.
b) Ha az

$$A \operatorname{dom} f - \operatorname{dom} g$$

konvex halmaz zárt, akkor (D) megoldhatósága és korlátossága esetén (P) megoldható lesz.
c) Ha a $\operatorname{dom} f^c$, $\operatorname{dom} g_c$ konvex halmazok zártak, továbbá

$$(A \operatorname{ri} \operatorname{dom} \operatorname{rec} f) \cap \operatorname{ri} \operatorname{dom} \operatorname{rec} (-g) \neq \emptyset,$$

akkor (P) megoldhatósága és korlátossága esetén (D) megoldható lesz. Ha még az is teljesül, hogy a $\operatorname{dom} f^c$, $\operatorname{dom} g_c$ halmazok egyenesmentesek, akkor (P) szigorúan megoldható.
d) Ha a $\operatorname{dom} f$, $\operatorname{dom} g$ konvex halmazok zártak, továbbá

$$(A^T \operatorname{ri} \operatorname{dom} \operatorname{rec} (-g_c)) \cap \operatorname{ri} \operatorname{dom} \operatorname{rec} (f^c) \neq \emptyset,$$

akkor (D) megoldhatósága és korlátossága esetén (P) megoldható lesz. Ha még az is teljesül, hogy a $\operatorname{dom} f$, $\operatorname{dom} g$ halmazok egyenesmentesek, akkor (D) szigorúan megoldható. □

Megjegyezzük, hogy 11.12-ben, és így 11.13-ban is elhagyható az effektív tartományok egyenesmentességének feltétele, ha $f, -g$ konvex kúpfüggvények, vagyis $f = \operatorname{rec} f$ és $-g = \operatorname{rec}(-g)$ teljesül. Általában azonban a fenti feltétel nem hagyható el, mint azt az alábbi példa mutatja: Legyen $C \subseteq \mathcal{R}^m$ kompakt, konvex halmaz, $f := \operatorname{ind}_{\{0\}}$, $g := -\operatorname{ind}_C$, továbbá A, c tetszőleges, $-b \in C \setminus \operatorname{ri} C$. Ekkor $z = 0$ mutatja, hogy 11.13 c) feltételei teljesülnek (az egyenesmentességi feltételt leszámítva), a 0 vektor primál megengedett megoldás, a duál megengedett megoldások halmaza, $\operatorname{dom} g_c$ nemüres, (P) mégsem szigorúan megoldható.

A Rockafellar-féle erős dualitási tétel két további következményét említjük a fejezet végén. Az első Eisenberg dualitási tétele pozitív homogén függvények esetén, a második Dennis dualitási tétele zárt, konvex függvény optimalizálásáról poliéderen.

Legyen először h az $\mathcal{R}^n$ téren értelmezett, zárt, valódi konvex kúpfüggvény, k az $\mathcal{R}^m$ téren értelmezett zárt, valódi konkáv kúpfüggvény. Legyen továbbá $A \in \mathcal{R}^{m\times n}$, és $f := h$, $g := k_c$. Ekkor $-g$ a $\operatorname{dom}(k_c)$ halmaz indikátorfüggvénye, f^c pedig a $\operatorname{dom}(h^c)$ halmazé (vö. 9.5 bizonyításával). A k függvény zártsága miatt $g_c = k$. Ezért a Rockafellar-féle (P) és (D) program ekkor $b = 0$ és $c = 0$ esetén a következő alakot ölti:

$$(EP)\begin{cases} \inf h(x), \\ Ax \in \operatorname{dom} g = \operatorname{dom} k_c, \end{cases} \quad \text{azaz} \quad \begin{cases} \inf h(x), \\ y^T A x \geq k(y)\ (y \in \operatorname{dom} k) \end{cases}$$
$$(ED)\begin{cases} \sup k(y), \\ A^T y \in \operatorname{dom} f^c = \operatorname{dom} h^c, \end{cases} \quad \text{azaz} \quad \begin{cases} \sup k(y), \\ y^T A x \leq h(x)\ (x \in \operatorname{dom} h). \end{cases}$$

11.14. Tétel: (Eisenberg dualitási tétele) *Tegyük fel, hogy h az $\mathcal{R}^n$ téren értelmezett, zárt, pozitív homogén, valódi konvex függvény, k az $\mathcal{R}^m$ téren értelmezett, zárt, pozitív homogén, valódi konkáv függvény, és $A \in \mathcal{R}^{m\times n}$. Ekkor a következő állítások teljesülnek:*
a) $h(x) \geq k(y)$, ha x és y az (EP), illetve (ED) programok megengedett megoldásai;
b) ha létezik $x \in \operatorname{ri}\operatorname{dom} h$ vektor, amelyre $Ax \in \operatorname{ri}\operatorname{dom} k_c$, akkor az (EP) és (ED) programok optimumértékei megegyeznek, és az (ED) program optimumértéke felvétetik, ha véges;
c) ha létezik $y \in \operatorname{ri}\operatorname{dom} k$ vektor, amelyre $A^T y \in \operatorname{ri}\operatorname{dom} h^c$, akkor az (EP) és (ED) programok optimumértékei megegyeznek, és az (EP) program optimumértéke felvétetik, ha véges. □

Most Dennis eredményét vezetjük le a Rockafellar-féle erős dualitási tételből.

Legyen $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ zárt, valódi konvex függvény, $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$, $c = 0$, továbbá $-g$ az $\mathcal{R}^m_+$ halmaz indikátorfüggvénye. Ekkor $g_c = g$, és a Rockafellar-féle (P) és (D) programok az alábbi alakot öltik:

$$\begin{array}{ll} (DP) & \inf f(x),\ Ax \geq b,\ x \in \operatorname{dom} f, \\ (DD) & \sup b^T y - f^c(A^T y),\ y \geq 0,\ A^T y \in \operatorname{dom}(f^c). \end{array}$$

Az utóbbi program ekivalens módon még úgy is felírható, mint

$$(DD) \quad \sup b^T y - f^c(z),\ z = A^T y,\ y \geq 0,\ z \in \operatorname{dom}(f^c).$$

11.15. Tétel: (Dennis dualitási tétele) *Legyen* $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ *zárt, valódi konvex függvény,* $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$. *Vezessük be a* $P := \{x : Ax \geq b\}$ *és* $R := \operatorname{Im}_+ A^T$ *jelöléseket. Ekkor*
a) A (DP) *program optimumértéke legalább akkora, mint a* (DD) *program optimumértéke.*
b) Ha $P \cap \operatorname{ri} \operatorname{dom} f \neq \emptyset$, *akkor a* (DP) *és* (DD) *programok optimumértéke megegyezik, és a* (DD) *program optimumértéke felvétetik, ha véges.*
c) Ha $R \cap \operatorname{ri} \operatorname{dom}(f^c) \neq \emptyset$, *akkor a* (DP) *és* (DD) *programok optimumértéke megegyezik, és a* (DP) *program optimumértéke felvétetik, ha véges.*
d) A b) és c) esetben is (általában ha az optimumértékek megegyeznek) az x *primál megengedett és* (y, z) *duál megengedett megoldások pontosan akkor optimálisak, ha teljesítik az alábbi egyenlőségeket*

$$y^T Ax = b^T y = z^T x \text{ és } f^c(z) = z^T x - f(x).$$

□

12. Lagrange-, Wolfe-duál

Tekintsük most az

$$\begin{array}{ll} (LP): & \inf f(x),\ g_i(x) \leq 0\ (i = 1, \ldots, m),\ x \in C, \\ (LD): & \sup \inf\{f(x) + \sum_{i=1}^m y_i g_i(x) : x \in C\},\ y \in \mathcal{R}^m_+ \end{array}$$

Lagrange primál-duál programpárt, ahol $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, $f : C \to \mathcal{R}$ konvex függvény, $g_i : C \to \mathcal{R}$ $(i = 1, \ldots, m)$ konvex függvények. Jelölje **LP**, illetve **LD** a primál, illetve a duál program megengedett megoldásainak halmazát, vagyis legyen

$$\begin{aligned}
&\mathbf{LP} := \{x \in C : g_i(x) \leq 0 \ (i = 1, \ldots, m)\},\\
&\mathbf{LD} := \{y \in \mathcal{R}^m_+ : \inf\{f(x) + \textstyle\sum_{i=1}^m y_i g_i(x) : x \in C\} > -\infty\}.
\end{aligned}$$

Jelölje v_{LP}, illetve v_{LD} a megfelelő programok optimumértékeit, $\mathbf{LP_o}$, illetve $\mathbf{LD_o}$ pedig optimális megoldásaik halmazát. Jelölje továbbá g a $(g_1(.), \ldots, g_m(.))^T$ leképezést.

Először megmutatjuk, hogy itt a Rockafellar-féle dualitás egy speciális esetéről van szó.

Legyen ugyanis

$$\begin{aligned}
\hat{C}_1 &:= \{(x, b) \in \mathcal{R}^n \times \mathcal{R}^m : x \in C,\ g(x) + b \leq 0\},\\
\hat{C}_2 &:= \{(x, b) \in \mathcal{R}^n \times \mathcal{R}^m : b \geq 0\}
\end{aligned}$$

($\hat{C}_2$ definíciójában $b = 0$ is írható), továbbá

$$\hat{f}(x, b) := \begin{cases} f(x), & \text{ha } (x, b) \in \hat{C}_1,\\ \infty & \text{egyébként,} \end{cases}$$

$$\hat{g}(x, b) := \begin{cases} 0, & \text{ha } (x, b) \in \hat{C}_2,\\ -\infty & \text{egyébként.} \end{cases}$$

Ha $\hat{A}$ az $(n+m) \times (n+m)$-es egységmátrix, $\hat{b} = 0$, $\hat{c} = 0$, akkor az $\hat{A}, \hat{b}, \hat{c}, \hat{f}, \hat{g}$ inputhoz tartozó Rockafellar primál-duál programpár

$$\begin{aligned}
(\hat{P}) : &\quad \inf \hat{f}(x, b) - \hat{g}(x, b)\\
(\hat{D}) : &\quad \sup \hat{g}_c(a, y) - \hat{f}^c(a, y).
\end{aligned}$$

Megmutatjuk, hogy az (LP) és $(\hat{P})$, illetve (LD) és $(\hat{D})$ programok ekvivalensek. Az (x, b) vektor pontosan akkor lesz a $(\hat{P})$ program megengedett megoldása, ha $x \in C$, $g(x) + b \leq 0$, $b \geq 0$. Ekkor az x vektor az (LP) program azonos célfüggvényértékű megengedett megoldása. Megfordítva ha az x vektor az (LP) program megengedett megoldása, akkor az $(x, 0)$ vektor a $(\hat{P})$ program azonos célfüggvényértékű megengedett megoldása. Ezért az (LP) és $(\hat{P})$ programok ekvivalensek. Az (LD) és $(\hat{D})$ programok ekvivalenciájának bizonyítása érdekében számoljuk ki a $\hat{g}_c$ és $\hat{f}^c$ függvényeket:

$$\begin{aligned}
\hat{f}^c(a, y) &= \sup\left\{a^T x + y^T b - \hat{f}(x, b) : x \in \mathcal{R}^n,\ b \in \mathcal{R}^m\right\} =\\
&= \sup\left\{a^T x + y^T b - f(x) : x \in C,\ g(x) + b \leq 0\right\}\\
\hat{g}_c(a, y) &= \inf\left\{a^T x + y^T b - \hat{g}(x, b) : x \in \mathcal{R}^n,\ b \in \mathcal{R}^m\right\} =\\
&= \inf\left\{a^T x + y^T b : b \geq 0\right\} = \begin{cases} 0, & \text{ha } a = 0,\ y \geq 0,\\ -\infty & \text{egyébként.} \end{cases}
\end{aligned}$$

Ezért

$$
\begin{aligned}
&\hat{g}_c(a,y) - \hat{f}^c(a,y) = \\
&= \begin{cases} -\sup\Big\{y^T b - f(x) : x \in C,\ g(x) + b \le 0\Big\}, & \text{ha } a = 0,\ y \ge 0, \\ -\infty & \text{egyébként} \end{cases} = \\
&= \begin{cases} \inf\Big\{y^T g(x) + f(x) : x \in C\Big\}, & \text{ha } a = 0,\ y \ge 0, \\ -\infty & \text{egyébként,} \end{cases}
\end{aligned}
$$

amiből az (LD) és $(\hat{D})$ programok ekvivalenciája is látszik.

Most már könnyen belátható Rockafellar erős dualitási tételének segítségével az alábbi

12.1. Tétel: *Tegyük fel, hogy az $f, g_1, \ldots, g_m : C \to \mathcal{R}$ konvex függvények alulról félig folytonosak, továbbá hogy valamely $\lambda_0 \ge 0$ szám és $\hat{\lambda} \in \mathcal{R}_+^m$ vektor esetén a $\lambda_0 f + \hat{\lambda}^T g$ függvény (alsó) nívóhalmazai kompaktak. Ekkor* $\mathbf{LP} \cup \mathbf{LD} \neq \emptyset$ *esetén $v_{LP} = v_{LD}$, és az (LP) program optimumértéke felvétetik, ha véges.*

Bizonyítás: Elég azt megmutatnunk, hogy a $(\hat{P})$, $(\hat{D})$ programokra teljesül a primál zártsági feltétel. Kiszámolva az $\operatorname{epi} \hat{f} - \operatorname{hipo} \hat{g}$ halmazt láthatjuk, hogy ez a halmaz pontosan akkor zárt, ha az

$$
S := \left\{ \begin{pmatrix} g(x) \\ f(x) \end{pmatrix} : x \in C \right\} + \mathcal{R}_+^{m+1}
$$

halmaz zárt.

Megmutatjuk, hogy az S halmaz zárt. Legyen x_k $(k = 1, 2, \ldots)$ C-beli elemek egy sorozata, $y_k \in \mathcal{R}_+^m$ nemnegatív vektorok, $\rho_k \ge 0$ nemnegatív számok $(k = 1, 2, \ldots)$, és tegyük fel, hogy

$$
g(x_k) + y_k \to b \ (k \to \infty), \text{ és } f(x_k) + \rho_k \to \gamma \ (k \to \infty).
$$

Ekkor természetesen

$$
\lambda_0 f(x_k) + \sum_{i=1}^m \hat{\lambda}_i g_i(x_k) + \lambda_0 \rho_k + \hat{\lambda}^T y_k \to \lambda_0 \gamma + \hat{\lambda}^T b \ (k \to \infty),
$$

így elég nagy k indexek esetén x_k eleme a feltétel szerint kompakt

$$
\Big\{x \in C : \lambda_0 f(x) + \hat{\lambda}^T g(x) \le \lambda_0 \gamma + \hat{\lambda}^T b + 1\Big\}
$$

halmaznak. A Bolzano–Weierstrass-tétel szerint feltehető, hogy az x_k sorozat konvergál valamely $x \in C$ ponthoz. A szereplő függvények alulról félig folytonossága miatt az

$$\{f(x_k) : k = 1, 2, \ldots\}, \{g_i(x_k) : k = 1, 2, \ldots\} \ (i = 1, \ldots, m)$$

halmazok korlátosak (felülről persze korlátosak; ha alulról nem lenne korlátos az egyikük, például az első halmaz, akkor az maga után vonná, hogy $f(x) = -\infty$, ellentmondásban azzal, hogy $x \in C$). Feltehető tehát, hogy a megfelelő sorozatok konvergensek, limeszüket jelölje $\gamma - \rho$, illetve $b_i - \hat{y}_i$ $(i = 1, \ldots, m)$. Ekkor $\rho_k \to \rho$ $(k \to \infty)$, és $y_k \to \hat{y}$ $(k \to \infty)$, tehát $\rho \geq 0$, és $\hat{y} \in \mathcal{R}^m_+$. A szereplő függvények alulról félig folytonossága miatt $\gamma - \rho \geq f(x)$, és $b - \hat{y} \geq g(x)$. Mindebből az is látszik, hogy $(b, \gamma) \in S$, amit bizonyítanunk kellett. □

Bizonyítás nélkül megjegyezzük: C zártsága, $\lambda_0 = 1$, $\hat{\lambda} > 0$ esetén létezik $y > 0$ vektor úgy, hogy $0 \in \operatorname{ri} \operatorname{dom} (f + y^T g)^c$ (lásd [20], 13.3.4, 8.7), azaz a (primál zártsági feltételnél erősebb) $\operatorname{dom} \hat{g}_c \cap \operatorname{ri} \operatorname{dom} \hat{f}^c \neq \emptyset$ feltétel teljesül.

12.1 duális megfelelőjét direktebb módon igazoljuk (lásd még 12.18). 12.5-höz mindent elölről kezdve 11.5 megfelelőjét, a konvex Farkas-tételt bizonyítjuk.

Azt mondjuk, hogy az (LP) programra teljesül az **erős Slater-feltétel**, ha létezik $x_0 \in C$ pont, amely az (LP) program összes feltételét szigorú egyenlőtlenséggel teljesíti, vagyis amelyre $g(x_0) < 0$. Ekkor az x_0 pontot **erős Slater-pont**nak nevezzük. Ha x_0 erős Slater-pont, akkor 8.19 szerint feltehető, hogy $x_0 \in \operatorname{ri} C$ (ez mutatja azt is, hogy az erős Slater-feltétel valóban erősebb, mint az alábbi gyenge Slater-feltétel).

Azt mondjuk, hogy az (LP) programra teljesül a **(gyenge) Slater-feltétel**, ha létezik $x \in \operatorname{ri} C$ megengedett megoldása, amelyre $g_i(x) < 0$, ha g_i nemaffin függvény; és $g_i(x) \leq 0$, ha g_i affin függvény. Az ilyen x pontot **Slater-pont**nak nevezzük.

A $g_i(x) \leq 0$ feltételt **reguláris**nak nevezzük, ha az (LP) programnak létezik olyan megengedett megoldása, amely ezt a feltételt szigorúan teljesíti. Indexhalmazukat jelölje $I(\text{reg})$. A nem reguláris feltételeket **szinguláris** feltételeknek nevezzük. Indexhalmazukat jelölje $I(\text{sing})$. Ha teljesül a Slater-feltétel, nyilván minden szinguláris feltételbeli g_i függvény affin.

12.2. Állítás: *Ha teljesül a Slater-feltétel, akkor létezik olyan x Slater-pont is, amelyre*

$$g_i(x) < 0 \ (i \in I(reg)),\ g_i(x) = 0 \ (i \in I(sing)).$$

Az ilyen Slater-pontot **ideális Slater-pont***nak nevezzük.*

Bizonyítás: Minden $i \in I(\text{reg})$ esetén válasszunk $x_i \in \mathbf{LP}$ pontot, amelyre $g_i(x_i) < 0$. Ezeknek a pontoknak és egy Slater-pontnak a pozitív, konvex kombinációja megfelel. □

12.3. Tétel: (konvex Farkas-tétel) *Tegyük fel, hogy az (LP) programra teljesül a Slater-feltétel. Az*

$$\begin{aligned} & f(x) < 0, \\ & g_i(x) \leq 0 \ (i = 1, \ldots, m), \\ & x \in C \end{aligned}$$

rendszernek pontosan akkor nincs megoldása, ha létezik nemnegatív $y = (y_1, \ldots, y_m)^T$ vektor úgy, hogy

$$f(x) + \sum_{i=1}^{m} y_i g_i(x) \geq 0 \ (x \in C).$$

Utóbbi esetben a szinguláris feltételekhez tartozó y_i $(i \in I(sing))$ szorzókról az is feltehető, hogy pozitívak.

Bizonyítás: Nyilvánvaló, hogy a két rendszernek nem lehet egyszerre megoldása, ez a regularitási feltétel nélkül is fennáll. Megmutatjuk, hogy ha az első rendszernek nincs megoldása, akkor a másodiknak van.

Tekintsük az

$$S := \cup_{x \in C} \left\{ b \in \mathcal{R}^{m+1} \;\middle|\; \begin{array}{l} f(x) < b_0,\ g_i(x) \leq b_i \ (i \in I(\text{reg})), \\ g_i(x) = b_i \ (i \in I(\text{sing})) \end{array} \right\}$$

halmazt. Ez nyilván az

$$\left\{ \begin{pmatrix} x \\ b \end{pmatrix} \in \begin{pmatrix} C \\ \mathcal{R}^{m+1} \end{pmatrix} \;\middle|\; \begin{array}{l} f(x) - b_0 < 0,\ g_i(x) - b_i \leq 0 \ (i \in I(\text{reg})), \\ g_i(x) - b_i = 0 \ (i \in I(\text{sing})) \end{array} \right\}$$

8.13 szerint konvex halmaz vetülete, így konvex (a Slater-feltétel miatt a szinguláris feltételekben szereplő függvények affinok!), és az első rendszer megoldhatatlansága miatt nem tartalmazza az origót. A második Hahn–Banach-tétel szerint létezik $y = (y_0, y_1, \ldots, y_m)$ vektor úgy, hogy

$$\sum_{i=0}^{m} y_i b_i \geq 0 \ (b \in S),$$

és valamely $\overline{b} \in S$ esetén (valójában valamennyi $\overline{b} \in \operatorname{ri} S$ esetén)

$$\sum_{i=0}^{m} y_i \overline{b}_i > 0.$$

A bizonyítás hátralévő része négy részre osztható.

1. Először megmutatjuk, hogy $y_0 \geq 0$, és $y_i \geq 0$ $(i \in I(\text{reg}))$. Vegyünk egy tetszőleges b elemet S-ből. Ekkor $b(\lambda) := (b_0 + \lambda, b_1, \ldots, b_m)$ is S-beli minden $\lambda \geq 0$ esetén. Ha $y_0 < 0$ lenne, $y^T b(\lambda) \geq 0$ nem teljesülne elég nagy λ esetén. Hasonlóan bizonyítható, hogy $y_i \geq 0$ $(i \in I(\text{reg}))$.

2. Másodszor megmutatjuk, hogy

$$y_0 f(x) + \sum_{i=1}^{m} y_i g_i(x) \geq 0 \ (x \in C).$$

Ez abból az észrevételből következik, hogy minden $x \in C$, $\lambda > 0$ esetén $(f(x) + \lambda, g_1(x), \ldots, g_m(x)) \in S$, és így

$$y_0(f(x) + \lambda) + \sum_{i=1}^{m} y_i g_i(x) \geq 0 \ (x \in C).$$

Tartsunk λ-val nullához.

3. Harmadszor megmutatjuk, hogy $y_0 > 0$. Indirekt bizonyítunk. Már tudjuk, hogy $y_0 \geq 0$. Tegyük fel, hogy $y_0 = 0$. Legyen x^* ideális Slater-pont. A bizonyítás második részéből adódik, hogy $\sum_{i \in I(\text{reg})} y_i g_i(x^*) \geq 0$. Másfelől (lásd az első részt) $y_i \geq 0, g_i(x^*) < 0$ $(i \in I(\text{reg}))$, így $y_i = 0$ $(i \in I(\text{reg}))$. Az $\overline{b}$ ponthoz létezik $\overline{x} \in C$ úgy, hogy $\overline{b}_i = g_i(\overline{x})$ $(i \in I(\text{sing}))$. A $\sum_{i \in I(\text{sing})} y_i g_i(x)$ összeg az x változó affin függvénye, a C-beli $\overline{x}$-ban pozitív, a $\operatorname{ri} C$-beli x^*-ban nulla, így az $\overline{x} \in C$ pontot az $x^* \in \operatorname{ri} C$ ponton túlhúzva a C halmazban, olyan $x' \in C$ ponthoz jutunk, amelyben az összeg negatív, ami ellentmond a bizonyítás második részének.

Ezzel beláttuk, hogy $y_0 > 0$. Feltehető, hogy $y_0 = 1$. (Ha nem így lenne, leoszthatjuk vele az y vektort.)

4. Végül $I(\text{sing})$ elemszáma szerinti teljes indukcióval megmutatjuk, hogy az y_i $(i \in I(\text{sing}))$ szorzók pozitívvá tehetők. Ha $I(\text{sing}) = \emptyset$, akkor készen vagyunk. Az általános lépéshez tegyük fel, hogy $|I(\text{sing})| = k + 1 \geq 1$, és $|I(\text{sing})| \leq k$ esetén már bizonyítottuk a 4.-beli állítást. Tekintsük minden $y_s \leq 0$, $s \in I(\text{sing})$ esetén az alábbi rendszert:

$$\begin{aligned}
&g_s(x) < 0,\\
&g_i(x) \leq 0 \ (i \neq s),\\
&x \in C.
\end{aligned}$$

Az s-edik feltétel szingularitása miatt ennek a rendszernek nincs megoldása, szinguláris feltételeinek száma legfeljebb k, továbbá teljesül rá a Slater-feltétel, így az indukciófeltétel szerint léteznek $\hat{y}_i^{(s)} \geq 0$ $(i \neq s)$ szorzók úgy, hogy

$$g_s(x) + \sum_{i \neq s} \hat{y}_i^{(s)} g_i(x) \geq 0 \ (x \in C).$$

Legyen $\hat{y}_s^{(s)} := 1$, ekkor az $\hat{y}^{(s)}$ vektorok alkalmas pozitív számú többszörösét (például $(1 - y_s)$-szeresét) hozzáadva az y vektorhoz, megfelelő szorzókhoz jutunk. □

12.4. Megjegyzés: Az az eset, mikor egyetlen $f(x) < 0$ feltétel helyett több $f_j(x) < 0$ feltétel van a rendszerben, könnyen visszavezethető a fenti tételre. Vizsgáljuk ugyanis a

$$\begin{aligned}
&\tau < 0,\\
&f_j(x) - \tau \leq 0,\\
&g_i(x) \leq 0,\\
&x \in C,\ \tau \in \mathcal{R}
\end{aligned}$$

rendszert. Ez pontosan akkor megoldható, ha az

$$\begin{aligned}
&f_j(x) < 0,\\
&g_i(x) \leq 0,\\
&x \in C
\end{aligned}$$

rendszer megoldható, és ha az utóbbi rendszerre teljesül a Slater-feltétel, akkor az előbbire is teljesül. Ha az utóbbi rendszer nem megoldható, akkor az előbbire alkalmazva a konvex Farkas-tételt, az f_j függvények szorzóinak összege 1 lesz.

Most már beláthatjuk 12.1 duális megfelelőjét.

12.5. Tétel: *Tegyük fel, hogy az (LP) programra teljesül a Slater-feltétel. Ekkor $v_{LP} = v_{LD}$, és a duál optimumérték felvétetik, ha véges.*

Bizonyítás: Ha $v_{LP} = -\infty$, akkor a gyenge dualitási tétel miatt $v_{LD} = -\infty$ is teljesül, tehát ekkor $v_{LP} = v_{LD}$.

Mivel a Slater-pont egyben primál megengedett megoldás is, azért már csak azt az esetet kell megvizsgálnunk, mikor $v_{LP} \in \mathcal{R}$. Ekkor persze nem létezik primál megengedett megoldás, amelyre $f(x) - v_{LP} < 0$ lenne, így a konvex Farkas-tétel szerint létezik $y \in \mathcal{R}^m_+$ vektor, amelyre

$$f(x) + \sum_{i=1}^{m} y_i g_i(x) \geq v_{LP} \ (x \in C).$$

Mivel ugyanakkor ennek az egyenlőtlenségnek a bal oldalán infimumot véve x-ben egy legfeljebb v_{LD} nagyságú szám áll, azért az is látszik, hogy $v_{LP} = v_{LD}$, és y duál optimális megoldás. □

A konvex Farkas-tétel (vagy 12.5) segítségével nevezetes optimalitási tételeket bizonyíthatunk. Az

$$L(x, y) := f(x) + \sum_{i=1}^{m} y_i g_i(x) \ (x \in C, y \in \mathcal{R}^m_+)$$

függvényt az (LP) feladathoz tartozó **Lagrange-függvény**nek nevezzük.

Azt mondjuk, hogy az $(\hat{x}, \hat{y}) \in C \times \mathcal{R}^m_+$ vektorpár az L Lagrange-függvény **nyeregpont**ja, ha teljesül rá az alábbi **nyeregpont egyenlőtlenség**

$$L(\hat{x}, y) \leq L(\hat{x}, \hat{y}) \leq L(x, \hat{y}) \ (x \in C, y \in \mathcal{R}^m_+).$$

Könnyen belátható, hogy a nyeregpont egyenlőtlenség ekvivalens az alábbival

$$L(\hat{x}, y) \leq L(x, \hat{y}) \ (x \in C, y \in \mathcal{R}^m_+).$$

12.6. Állítás: *Az $(\hat{x}, \hat{y}) \in C \times \mathcal{R}^m_+$ vektorpár pontosan akkor lesz az L Lagrange-függvény nyeregpontja, ha teljesül rá, hogy*

$$\sup_{y \geq 0} L(\hat{x}, y) = \min_{x \in C} \sup_{y \geq 0} L(x, y) = \max_{y \geq 0} \inf_{x \in C} L(x, y) = \inf_{x \in C} L(x, \hat{y}).$$

Ekkor a fenti közös érték $L(\hat{x}, \hat{y})$.

Bizonyítás: Könnyen belátható, hogy az állításban szereplő egyenlőségekből következik a nyeregpont egyenlőtlenség.

Megfordítva bármely $(\tilde{x}, \tilde{y}) \in C \times \mathcal{R}_+^m$ vektorpár esetén

$$\inf_{x \in C} L(x, \tilde{y}) \leq L(\tilde{x}, \tilde{y}) \leq \sup_{y \geq 0} L(\tilde{x}, y),$$

ezért a bal oldalon szuprémumot véve, a jobb oldalon pedig infimumot, oda jutunk, hogy

$$\sup_{y \geq 0} \inf_{x \in C} L(x, y) \leq \inf_{x \in C} \sup_{y \geq 0} L(x, y).$$

Továbbá a nyeregpont egyenlőtlenséget használva kapjuk, hogy

$$\inf_{x \in C} \sup_{y \geq 0} L(x, y) \leq \sup_{y \geq 0} L(\hat{x}, y) \leq L(\hat{x}, \hat{y}) \leq \inf_{x \in C} L(x, \hat{y}) \leq \sup_{y \geq 0} \inf_{x \in C} L(x, y).$$

A fentiekből már következik az állítás még nem bizonyított iránya is. $\square$

12.7. Tétel: (Karush–Kuhn–Tucker) *Tegyük fel, hogy az (LP) programra teljesül a Slater-feltétel. Ekkor az $\hat{x} \in C$ vektor pontosan akkor lesz az (LP) program optimális megoldása, ha létezik $\hat{y} \in \mathcal{R}_+^m$ vektor úgy, hogy az $(\hat{x}, \hat{y})$ vektorpár az L Lagrange-függvény nyeregpontja.*

Bizonyítás: A tétel könnyű része annak bizonyítása, hogy ha $(\hat{x}, \hat{y})$ a Lagrange-függvény nyeregpontja, akkor $\hat{x}$ az (LP) program optimális megoldása. A bizonyítás ezen részéhez még a regularitási feltételre sincs szükség. A nyeregpont egyenlőtlenségből kapjuk, hogy

$$\begin{aligned} f(\hat{x}) + \textstyle\sum_{i=1}^m y_i g_i(\hat{x}) &\leq f(\hat{x}) + \textstyle\sum_{i=1}^m \hat{y}_i g_i(\hat{x}) \leq \\ &\leq f(x) + \textstyle\sum_{i=1}^m \hat{y}_i g_i(x) \ (x \in C, y \in \mathcal{R}_+^m). \end{aligned}$$

Az első egyenlőtlenségből könnyen belátható, hogy $g_i(\hat{x}) \leq 0$ $(i = 1, \ldots, m)$ (ha valamelyik i index esetén $g_i(\hat{x}) > 0$ lenne, akkor y_i szorzóját eléggé megnövelve ellentmondáshoz jutnánk), tehát $\hat{x} \in \mathbf{LP}$. A fenti egyenlőtlenséglánc két szélét tekintve $y = 0$ helyettesítéssel az

$$f(\hat{x}) \leq f(x) + \sum_{i=1}^m \hat{y}_i g_i(x) \leq f(x) \ (x \in \mathbf{LP})$$

egyenlőtlenséghez jutunk, vagyis $\hat{x} \in \mathbf{LP_o}$ is teljesül.

A fordított irányhoz már szükség lesz a regularitási feltételre. Legyen $\hat{x} \in \mathbf{LP_o}$, ekkor persze az

$$\begin{aligned}
&f(x) - f(\hat{x}) < 0, \\
&g_i(x) \leq 0 \ (i = 1, \ldots, m), \\
&x \in C
\end{aligned}$$

egyenlőtlenség-rendszer nem megoldható. Ezért a konvex Farkas-tétel szerint létezik $\hat{y} \in \mathcal{R}_+^m$ vektor, amelyre

$$f(x) - f(\hat{x}) + \sum_{i=1}^{m} \hat{y}_i g_i(x) \geq 0 \ (x \in C).$$

Ebből az egyenlőtlenségből $x = \hat{x}$ helyettesítéssel látható, hogy

$$\sum_{i=1}^{m} \hat{y}_i g_i(\hat{x}) = 0,$$

és ezután a nyeregpont egyenlőtlenség,

$$f(\hat{x}) + \sum_{i=1}^{m} y_i g_i(\hat{x}) \leq f(\hat{x}) \leq f(x) + \sum_{i=1}^{m} \hat{y}_i g_i(x) \ (x \in C, y \in \mathcal{R}_+^m)$$

is könnyen adódik. □

A Karush–Kuhn–Tucker-tétel alábbi két egyszerű következménye elvezet minket a tétel leggyakrabban alkalmazott speciális esetéhez.

12.8. Következmény: *A 12.7 Tétel feltételeinek fennállása esetén az $\hat{x} \in C$ vektor pontosan akkor lesz az (LP) program optimális megoldása, ha létezik $\hat{y} \in \mathcal{R}_+^m$ vektor úgy, hogy*

a) $f(\hat{x}) + \sum_{i=1}^{m} \hat{y}_i g_i(\hat{x}) = \min_{x \in C} f(x) + \sum_{i=1}^{m} \hat{y}_i g_i(x)$, *és*
b) $\sum_{i=1}^{m} \hat{y}_i g_i(\hat{x}) = \max_{y \geq 0} \sum_{i=1}^{m} y_i g_i(\hat{x})$.

Bizonyítás: A tétel átfogalmazása. □

12.9. Következmény: *A 12.7 Tétel feltételeinek fennállása esetén az $\hat{x} \in \mathbf{LP}$ vektor pontosan akkor lesz az (LP) program optimális megoldása, ha létezik $\hat{y} \in \mathcal{R}_+^m$ vektor úgy, hogy*

a) $f(\hat{x}) = \min_{x \in C} f(x) + \sum_{i=1}^{m} \hat{y}_i g_i(x)$, *és*
b) $\sum_{i=1}^{m} \hat{y}_i g_i(\hat{x}) = 0$.

Bizonyítás: A 12.8 átfogalmazása (vagy 12.5 egyszerű következménye). □

12.10. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nemüres, nyílt halmaz, $C \subseteq S$, továbbá $f, g_i : S \to \mathcal{R}$ $(i = 1, \ldots, m)$ a C halmazon differenciálható függvények. Ekkor a 12.7 Tétel feltételeinek fennállása esetén az $\hat{x} \in$* **LP** *vektor pontosan akkor lesz az (LP) program optimális megoldása, ha létezik $\hat{y} \in \mathcal{R}^m_+$ vektor úgy, hogy*

a)	$(\nabla f(\hat{x}) + \sum_{i=1}^m \hat{y}_i \nabla g_i(\hat{x}))^T (x - \hat{x}) \geq 0$ $(x \in C)$, *és*
b)	$\sum_{i=1}^m \hat{y}_i g_i(\hat{x}) = 0$.

Ekkor az $(\hat{x}, \hat{y})$ vektort a (LP) feladat **Karush–Kuhn–Tucker-pont***jának nevezzük.*

Bizonyítás: 12.9 és 10.29 következménye, mivel az $L(., \hat{y})$ függvény konvex a C halmazon. □

Mi a kapcsolat az $\mathbf{LP_o} \times \mathbf{LD_o}$ halmaz és a nyeregpontok (illetve a differenciálható esetben a Karush–Kuhn–Tucker-pontok) között?

12.11. Állítás: *Ha a Lagrange primál-duál programpárra teljesül, hogy $v_{LP} = v_{LD}$, akkor $\mathbf{LP_o} \times \mathbf{LD_o}$ éppen a Lagrange-függvény nyeregpontjainak halmaza. A differenciálható speciális esetben $\mathbf{LP_o} \times \mathbf{LD_o}$ éppen a Karush–Kuhn–Tucker-pontok halmaza.*

Bizonyítás: Az optimumértékek egyenlősége miatt, a kiegészítő eltérések tétele szerint $\hat{x} \in \mathbf{LP_o}$, $\hat{y} \in \mathbf{LD_o}$ pontosan akkor, ha $\hat{x} \in$ **LP**, $\hat{y} \geq 0$, és a megfelelő célfüggvényértékek megegyeznek, vagyis

$$\inf_{x \in C} f(x) + \sum_{i=1}^m \hat{y}_i g_i(x) = f(\hat{x}).$$

Könnyen belátható, hogy ekkor (és csak ekkor) $\hat{x} \in$ **LP**, $\hat{y} \geq 0$, továbbá a 12.9-beli a) és b) feltételek is teljesülnek; ami viszont azzal ekvivalens, hogy $\hat{x} \in C$, $\hat{y} \geq 0$, továbbá a 12.8-beli a) és b) feltételek teljesülnek; vagyis $(\hat{x}, \hat{y})$ nyeregpont.

A tétel Karush–Kuhn–Tucker-pontokra vonatkozó része hasonlóan igazolható. □

Tegyük fel most, hogy $C \subseteq \mathcal{R}^n$ nemüres, nyílt, konvex halmaz, továbbá $f, g_i : C \to \mathcal{R}$ $(i = 1, \dots, m)$ differenciálható, konvex függvények. Ekkor az (LP) program **Wolfe-duálja** a

$$(WD): \qquad \sup f(x) + y^T g(x), \nabla f(x) + y^T \nabla g(x) = 0,\ x \in C,\ y \in \mathcal{R}^m_+$$

program. A Lagrange-duál nem mindig ekvivalens a Wolfe-duállal. (Tekintsük például az

$$(LP_0): \qquad \inf e^x,\ x \le 0,\ x \in \mathcal{R}$$

programot. Ennek Lagrange-, illetve Wolfe-duálja

$$\begin{aligned} (LD_0): &\qquad \sup \inf\{e^x + yx : x \in \mathcal{R}\},\ y \in \mathcal{R}_+, \\ (WD_0): &\qquad \sup e^x + yx,\ e^x + y = 0,\ x \in \mathcal{R},\ y \in \mathcal{R}_+. \end{aligned}$$

A Lagrange-duál egyetlen megengedett megoldása az $y = 0$, míg a Wolfe-duálnak nincs megengedett megoldása.) Általában csak $v_{LP} \ge v_{LD} \ge v_{WD}$.

12.12. Állítás: *Ha az (LP) és (LD) programok optimumértékei megegyeznek és felvétetnek, akkor a Lagrange-duál és a Wolfe-duál ekvivalensek egymással.*

Bizonyítás: Nyilvánvaló, hogy ha $(x, y) \in \mathbf{WD}$, akkor $y \in \mathbf{LD}$, és az y-hoz tartozó, (LD) programbeli célfüggvényérték legalább akkora, mint az (x, y)-hoz tartozó, (WD) programbeli célfüggvényérték. A nemtriviális irányhoz legyen $\hat{y} \in \mathbf{LD_o}$, és válasszunk tetszőleges $\hat{x} \in \mathbf{LP_o}$ elemet. Ekkor 12.11 szerint $(\hat{x}, \hat{y})$ Karush–Kuhn–Tucker-pont, speciálisan megengedett megoldása a (WD) programnak, és célfüggvényértéke legalább akkora, mint az $\hat{y} \in \mathbf{LD}$ vektorhoz tartozó célfüggvényérték. □

12.13. Következmény: *Ha az (LP) programra teljesül a Slater-feltétel, továbbá optimumértéke felvétetik, akkor $v_{LP} = v_{WD}$, és a (WD) program optimumértéke is felvétetik.* □

Példaképpen megvizsgáljuk a konvex kvadratikus programokra vonatkozó dualitási tételeket.

Legyen $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$, továbbá $N \in \mathcal{R}^{n \times n}$ szimmetrikus, pozitív szemidefinit mátrix. A

$$(KP): \qquad \inf \frac{1}{2} x^T N x + c^T x,\ Ax \le b$$

konvex kvadratikus program Wolfe-duálja a

$$(KD): \qquad \sup -\frac{1}{2}z^TNz - b^Ty,\ Nz + A^Ty + c = 0,\ y \geq 0$$

program. Itt a szokásos módon jelölje **KP**, **KD** a megfelelő programok megengedett megoldásainak halmazát, v_{KP}, v_{KD} optimumértékeiket, $\mathbf{KP_o}$, $\mathbf{KD_o}$ optimális megoldásaik halmazát.

12.14. Tétel: (kvadratikus gyenge dualitási tétel) *Ha x a (KP) program megengedett megoldása, (y, z) pedig a (KD) programé, akkor a megfelelő célfüggvényértékek között fennáll az*

$$\frac{1}{2}x^TNx + c^Tx \geq -\frac{1}{2}z^TNz - b^Ty$$

egyenlőtlenség, egyenlőséggel éppen akkor, ha

$$x - z \in \operatorname{Ker} N,\ \textit{és}\ y^T(Ax - b) = 0.$$

Bizonyítás: Legyen $x \in \mathbf{KP}$, $(y, z) \in \mathbf{KD}$. Ekkor

$$\begin{aligned} 0 \leq \tfrac{1}{2}(x-z)^TN(x-z) + y^T(b - Ax) = \\ = \tfrac{1}{2}x^TNx + c^Tx + \tfrac{1}{2}z^TNz + b^Ty, \end{aligned}$$

amiből már következik a tétel. (A szimmetrikus, pozitív szemidefinit mátrixok éppen a Gram-mátrixok, vagyis a B^TB alakban előálló mátrixok, ezért (vagy 10.29 miatt) $\operatorname{Ker} N = \{x : x^TNx = 0\}$. Fordítva is, vö. [27], Cholesky-faktorizáció.) □

Lényegében a Farkas-lemmából, a szokásos módon következik, hogy

12.15. Tétel: *Megoldható (KP) program esetén az alábbi állítások ekvivalensek:*
a) a (KP) program korlátos;
b) nem létezik w, amelyre $Aw \leq 0, Nw = 0, c^Tw < 0$;
c) $\mathbf{KD} \neq \emptyset$. □

12.16. Tétel: (kvadratikus erős dualitási tétel) *Ha $v_{KP} \in \mathcal{R}$, akkor $v_{KP} = v_{KD}$, és $\mathbf{KD_o} \neq \emptyset$. Hasonlóan ha $v_{KD} \in \mathcal{R}$, akkor $v_{KP} = v_{KD}$, és $\mathbf{KP_o} \neq \emptyset$.*

Bizonyítás: A konvex Farkas-tételt használjuk. Ha $v_{KP} \in \mathcal{R}$, akkor nem létezik x vektor, amelyre

$$\frac{1}{2}x^T Nx + c^T x - v_{KP} < 0,\ Ax \leq b,$$

így létezik $y \geq 0$ vektor úgy, hogy

$$\frac{1}{2}x^T Nx + c^T x - v_{KP} + y^T(Ax - b) \geq 0\ (x \in \mathcal{R}^n).$$

Az előző tételt a $0x \leq 0$ feltétellel alkalmazva látjuk, hogy az itt szereplő konvex kvadratikus függvény csak úgy lehet alulról korlátos, ha létezik z vektor, amelyre $Nz + A^T y + c = 0$. Ekkor $-\frac{1}{2}z^T Nz - b^T y \geq v_{KP}$, így $(y, z) \in \mathbf{KD_o}$, és $v_{KP} = v_{KD}$.

Hasonlóan igazolható a tétel második fele is. □

Látjuk, hogy a konvex kvadratikus programokra ugyanolyan dualitási tételek állnak fenn, mint **lineáris program**okra (vagyis a $\{0\}$, $\pm\mathcal{R}_+$, $\mathcal{R}$ kúpok direkt szorzataival leírt kúplineáris programok esetén).

A lineáris programokra vonatkozó erős dualitási tétel erősítése az alábbi Goldman–Tucker-tétel, amelyet szintén a konvex Farkas-tétel segítségével bizonyíthatunk. Alapokhoz közelibb bizonyítása megtalálható [18]-ban.

12.17. Tétel: (Goldman-Tucker) *Tekintsük például a*

$$\begin{array}{rl} (LIP): & \inf c^T x,\ Ax = b,\ x \geq 0, \\ (LID): & \sup b^T y,\ A^T y \leq c,\ y \in \mathcal{R}^m \end{array}$$

lineáris primál-duál programpárt, ahol $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$. Ha a duál programnak létezik optimális megoldása, akkor léteznek olyan $\hat{x}$ primál, illetve $\hat{y}$ duál optimális megoldások, amelyekre $\hat{x} + c - A^T\hat{y} > 0$.

Bizonyítás: Nyilván nem létezik y vektor, amelyre

$$-b^T y + v_{LID} < 0,\ -b^T y + v_{LID} \leq 0,\ A^T y - c \leq 0,$$

ezért létezik $x \geq 0$, $\xi \geq 0$, amelyre

$$-b^T y + v_{LID} + x^T(A^T y - c) + \xi(-b^T y + v_{LID}) \geq 0$$

minden y vektor esetén. Ebből $Ax = (\xi + 1)b$ és $v_{LID}(\xi + 1) \geq c^T x$ adódik. Ekkor tehát $\hat{x} := \frac{x}{\xi+1}$ primál optimális megoldás. Láttuk, hogy feltehető, hogy $\hat{x}_i > 0$, ha a feltétel szinguláris. Ezért ha $\hat{x}_i = 0$, akkor létezik $y_i \in \mathbf{LID_o}$ vektor, amelyre $(A^T y_i - c)_i < 0$. Ezeknek az y_i vektoroknak pozitív konvex kombinációja legyen $\hat{y}$. Könnyen belátható, hogy $\hat{x}$ és $\hat{y}$ megfelel. □

A fejezet végén a konvex függvényekkel leírt (LP) programra vonatkozó 12.5 dualitási tételt általánosítjuk a kúpkonvex leképezésekkel leírt (GP) programra. Foglalkozunk a (GP) program regularizációjával is: megmutatjuk, hogy egy a (GP) programmal ekvivalens programra teljesül az erős Slater-feltétel.

Legyen $C \subseteq \mathcal{R}^n$ konvex halmaz, továbbá legyen $P \subseteq \mathcal{R}^n$ poliéder. Legyen $K \subseteq \mathcal{R}^m$ konvex kúp, és legyen $R \subseteq \mathcal{R}^l$ poliéder kúp. Legyen $g : C \to \mathcal{R}^m$ K-konvex leképezés, és legyen $h : P \to \mathcal{R}^l$ R-poliédrikus leképezés. (A $g : C \to \mathcal{R}^m$ leképezés **K-konvex**, ha K-epigráfja, az

$$\{(x, \mu) : x \in C,\ g(x) \leq_K \mu\}$$

halmaz konvex. Más szavakkal a $g : C \to \mathcal{R}^m$ leképezés K-konvex, ha tetszőleges $x_1, x_2 \in C$, $0 \leq \varepsilon \leq 1$ esetén teljesül, hogy

$$g(\varepsilon x_1 + (1-\varepsilon)x_2) \leq_K \varepsilon g(x_1) + (1-\varepsilon)g(x_2).$$

A h leképezés **R-poliédrikus**, ha R-epigráfja poliéder. Például minden affin leképezés R-poliédrikus.)

Tekintsük az alábbi programpárt:

$$\begin{array}{lcl} (GP) & : & \inf f(x),\ g(x) \leq_K 0,\ h(x) \leq_R 0,\ x \in C \cap P \\ (GD) & : & \sup \inf\{f(x) + y^T g(x) + z^T h(x) : x \in C \cap P\},\ y \in K^*,\ z \in R^*. \end{array}$$

Hasonlóan a fejezet elején az (LP) programra belátottakhoz, a (GP) program is megfogalmazható primál Rockafellar-programként: (GP) ekvivalens az alábbi $(\hat{P})$ programmal:

$$(\hat{P}) \qquad \inf \hat{f}(\hat{x}) - \hat{g}(\hat{x}),\ \hat{x} = (x, b_1, b_2, b_3, b_4),$$

ahol

$$\begin{array}{lcl} \hat{f}(\hat{x}) & := & \begin{cases} f(x), & \text{ha } \hat{x} \in \hat{C}_1, \\ \infty & \text{különben}, \end{cases} \\ \hat{g}(\hat{x}) & := & \begin{cases} 0, & \text{ha } \hat{x} \in \hat{C}_2, \\ -\infty & \text{különben}, \end{cases} \end{array}$$

ahol

$$\begin{array}{lcl} \hat{C}_1 & := & \{\hat{x} : x \in C,\ g(x) + b_1 \leq_K 0,\ b_2 = b_4,\ b_3 \in K\}, \\ \hat{C}_2 & := & \{\hat{x} : x \in P,\ h(x) + b_2 \leq_R 0,\ b_1 = b_3,\ b_4 \in R\}. \end{array}$$

(Itt $\hat{f}$ konvex függvény, $\hat{g}$ konkáv poliédrikus függvény, $\hat{C}_1$ konvex halmaz, $\hat{C}_2$ poliéder. A bizonyítás $b_3 = 0$, $b_4 = 0$ mellett is működik.)

Mi a kapcsolat a (GD) program és a $(\hat{P})$ program duálja, a

$$(\hat{D}) \qquad \sup \hat{g}_c(\hat{y}) - \hat{f}^c(\hat{y}),\ \hat{y} = (a, y_1, y_2, y_3, y_4)$$

program között? Számítsuk ki a $(\hat{D})$ program célfüggvényét. Definíció szerint

$$\hat{g}_c(\hat{y}) = \inf\{\hat{y}^T\hat{x} - 0 : \hat{x} \in \hat{C}_2\}.$$

Könnyen belátható, hogy

$$\hat{g}_c(\hat{y}) = \begin{cases} \inf\{a^T x + y_2^T b_2 : x \in P,\ h(x) + b_2 \leq_R 0\}, \\ \qquad \text{ha } y_1 = -y_3,\ y_4 \in R^*, \\ -\infty \qquad \text{különben.} \end{cases}$$

Hasonlóan

$$\hat{f}^c(\hat{y}) = \begin{cases} \sup\{a^T x - f(x) + y_1^T b_1 : x \in C,\ g(x) + b_1 \leq_K 0\}, \\ \qquad \text{ha } y_2 = -y_4,\ y_3 \in -K^*, \\ \infty \qquad \text{különben.} \end{cases}$$

Ezért

$$\begin{aligned} &\hat{g}_c(\hat{y}) - \hat{f}^c(\hat{y}) = \\ &= \begin{cases} \inf\{a^T x + y_2^T b_2 : x \in P,\ h(x) + b_2 \leq_R 0\} + \\ \qquad + \inf\{-a^T x + f(x) + y_3^T b_1 : x \in C,\ g(x) + b_1 \leq_K 0\}, \\ \qquad \text{ha } -y_3 = y_1 \in K^*,\ -y_2 = y_4 \in R^*, \\ -\infty \qquad \text{különben} \end{cases} \\ &= \begin{cases} \inf\{a^T x + y_4^T h(x) : x \in P\} + \\ \qquad + \inf\{-a^T x + f(x) + y_1^T g(x) : x \in C\}, \\ \qquad \text{ha } -y_3 = y_1 \in K^*,\ -y_2 = y_4 \in R^*, \\ -\infty \qquad \text{különben.} \end{cases} \end{aligned}$$

Innen látszik, hogy a (GD) program a $(\hat{D})$ program relaxációja: ha $\hat{y}$ a $(\hat{D})$ program megengedett megoldása, akkor $y := y_1$, $z := y_4$ a (GD) programnak megengedett megoldása legalább akkora célfüggvényértékkel.

Könnyen belátható a (GP), (GD) programpár optimumértékei közötti $v_{GP} \geq v_{GD}$ reláció (a gyenge dualitás). Ezt figyelembe véve a fentiekből már következnek az alábbi észrevételek:

– a fenti programok optimumértékeire fennáll: $v_{\hat{P}} = v_{GP} \geq v_{GD} \geq v_{\hat{D}}$;

– ha $v_{\hat{P}} = v_{\hat{D}}$, akkor $v_{GP} = v_{GD}$;

– ha $v_{\hat{P}} = v_{\hat{D}}$ véges, és a $(\hat{D})$ program optimumértéke felvétetik, akkor a (GD) program optimumértéke is felvétetik.

Ahhoz, hogy a 11.7 b) erős dualitási tételt alkalmazhassuk a $(\hat{P})$, $(\hat{D})$ programpárra, biztosítanunk kellene az $\hat{f}$, $\hat{g}$ függvények zártságát, ami $\hat{f}$ esetében nem garantálható. Következő észrevételünk az, hogy ha a (P) programnak létezik 11.6 a), b), c) vagy d) pontjában leírt tulajdonságokkal rendelkező megengedett megoldása, akkor a nempoliédrikus f vagy g függvény(ek) zártságának feltétele nélkül is fennáll, hogy $v_P = v_D$, és a duál optimumérték felvétetik, ha véges (tehát 11.7 d) kicsit erősíthető).

Lássuk be a fenti állítást például abban az esetben, mikor (P) szigorúan megoldható, a többi eset is hasonlóan intézhető el 8.26 és 8.27 segítségével. Jelölje $(\bar{P})$ azt a Rockafellar primál programot, amelyet úgy kapunk a (P) programból, hogy az f, g függvényeket lezártjaikkal helyettesítjük. Nyilván $(D) = (\bar{D})$, és $\mathbf{P_s} = \bar{\mathbf{P}}_\mathbf{s}$, így ha (P) szigorúan megoldható, akkor $v_{\bar{P}} = v_D$, és (D) optimumértéke felvétetik, ha véges. Általában persze $v_{\bar{P}} \leq v_P$, de ha (P) szigorúan megoldható, akkor a két optimumérték megegyezik. Ez az állítás 8.26 egyszerű következménye: legyen $x_0 \in \mathbf{P_s}$, $x_1 \in \bar{\mathbf{P}}$, ekkor

$$x_0 + \lambda(x_1 - x_0) \in \mathbf{P_s} \ (0 \leq \lambda < 1),$$

és a fenti szigorúan megengedett megoldásokhoz tartozó (P) programbeli célfüggvényérték az $x_1 \in \bar{\mathbf{P}}$ ponthoz tartozó $(\bar{P})$ programbeli célfüggvényértékhez tart, ahogy $\lambda \to 1$. Mindebből már látszik, hogy (P) szigorúan megoldhatósága esetén $v_P = v_D$, és a duál program optimumértéke felvétetik, ha véges.

Most a (GP) programra vonatkozó elégséges feltételét adjuk annak, hogy a $(\hat{P})$ programnak legyen olyan $\hat{x}_0$ megengedett megoldása, amelyre $\hat{x}_0 \in (\operatorname{ri}\operatorname{dom}\hat{f}) \cap (\operatorname{dom}\hat{g})$, vagyis $\hat{x}_0 \in (\operatorname{ri}\hat{C}_1) \cap \hat{C}_2$ (vö. 11.6 b)).

Azt mondjuk, hogy a (GP) programra teljesül a **gyenge Slater-feltétel**, ha létezik $x_0 \in \mathcal{R}^n$ pont (ún. **gyenge Slater-pont**) úgy, hogy

$$x_0 \in P \cap \operatorname{ri} C,\ g(x_0) <_K 0,\ h(x_0) \leq_R 0.$$

Megmutatjuk, hogy e feltétel fennállása esetén

$$\hat{x}_0 := (x_0, -g(x_0)/2, -h(x_0), -g(x_0)/2, -h(x_0)) \in (\operatorname{ri}\hat{C}_1) \cap \hat{C}_2.$$

Elég azt igazolnunk, hogy

$$\operatorname{ri}\hat{C}_1 = \{\hat{x} : x \in \operatorname{ri} C,\ g(x) + b_1 <_K 0,\ b_2 = b_4,\ b_3 \in \operatorname{ri} K\}.$$

Elegendő annyit belátnunk, hogy

$$\operatorname{ri}\,\{(x,b) : x \in C,\ g(x) + b \leq_K 0\} = \{(x,b) : x \in \operatorname{ri} C,\ g(x) + b <_K 0\}.$$

Jelölje $\hat{C}_1'$ az $\{(x,b) : x \in C,\ g(x) + b \leq_K 0\}$ halmazt. Egyrészt $\hat{C}_1'$ vetülete az $\mathcal{R}^n \times \{0\}$ altérre $C \times \{0\}$, ezért 4.21 szerint $\operatorname{ri} \hat{C}_1'$ vetülete az $\mathcal{R}^n \times \{0\}$ altérre $\operatorname{ri} C \times \{0\}$. Másrészt $x \in \operatorname{ri} C$ esetén jelölje M_x az $\{(x,b) : b \in \mathcal{R}^m\}$ affin halmazt. Ekkor 4.16 szerint

$$\emptyset \neq M_x \cap \operatorname{ri} \hat{C}_1' = \operatorname{ri}(M_x \cap \hat{C}_1') = \{(x,b) : g(x) + b <_K 0\}.$$

Mindebből már adódik a $\hat{C}_1$ halmaz relatív belsejére vonatkozó képlet helyessége, és így az is, hogy a gyenge Slater-feltétel fennállása esetén $(\operatorname{ri} \hat{C}_1) \cap \hat{C}_2 \neq \emptyset$.

Összefoglalva, az alábbi tételhez jutottunk:

12.18. Tétel: *Tegyük fel, hogy létezik gyenge Slater-pont. Ekkor $v_{GP} = v_{GD}$, és a (GD) program optimumértéke felvétetik, ha véges.* □

A következőkben elfeledkezünk a poliédrikus feltételekről, és csak az alábbi programpárt vizsgáljuk:

$$\begin{array}{ll} (GP): & \inf f(x),\ g(x) \leq_K 0,\ x \in C \\ (GD): & \sup \inf \left\{ f(x) + y^T g(x) : x \in C \right\},\ y \in K^*. \end{array}$$

Azt mondjuk, hogy a (GP) program teljesíti az **erős Slater-feltétel**t, ha létezik $x_0 \in C$ pont, amelyre $g(x_0) \in -\operatorname{int} K$ (ún. **erős Slater-pont**).

12.19. Tétel: *A (GP) programra vonatkozó erős Slater-feltételből következik a gyenge Slater-feltétel.*

Bizonyítás: Legyen $x_0 \in C$, amelyre $g(x_0) \in -\operatorname{int} K$, továbbá legyen $x_1 \in \operatorname{ri} C$. Az Elérhetőségi lemma szerint $0 < \varepsilon < 1$ esetén $x_0 + \varepsilon(x_1 - x_0) \in \operatorname{ri} C$. Mivel $g(x_0) \in -\operatorname{int} K$, azért ε megválasztható úgy, hogy

$$0 < \varepsilon < 1,\ g(x_0) + \varepsilon(g(x_1) - g(x_0)) \in -\operatorname{int} K$$

teljesüljön. Mivel a g leképezés K-konvex, azért

$$g(x_0 + \varepsilon(x_1 - x_0)) \leq_K g(x_0) + \varepsilon(g(x_1) - g(x_0)).$$

Itt a jobb oldalon álló pont $-\operatorname{int} K$-beli, így a bal oldalon álló pont a $(-\operatorname{int} K) - K = -\operatorname{int} K$ halmaz eleme. Ezért az $x_0 + \varepsilon(x_1 - x_0)$ pont teljesíti a gyenge Slater-feltétel követelményeit. □

12.18 és 12.19 azonnali következménye az alábbi

12.20. Tétel: *Tegyük fel, hogy létezik erős Slater-pont. Ekkor $v_{GP} = v_{GD}$, és a (GD) program optimumértéke felvétetik, ha véges.*

12.20 az alábbi kúpkonvex Farkas-tétel segítségével is belátható 12.5 mintájára (lásd még [9]: Karush–Kuhn–Tucker-tétel; [11]: duál zártsági feltétel).

12.21. Tétel: *Legyen $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, $K_1 \subseteq \mathcal{R}^k$, $K_2 \subseteq \mathcal{R}^m$ konvex kúpok, $f : C \to \mathcal{R}^k$ K_1-konvex leképezés, $g : C \to \mathcal{R}^m$ K_2-konvex leképezés. Tegyük fel, hogy* $\operatorname{int} K_1 \neq \emptyset$, *továbbá létezik erős Slater-pont. Ekkor az*

$$f(x) <_{K_1} 0,\ g(x) \leq_{K_2} 0,\ x \in C$$

rendszernek pontosan akkor nem létezik megoldása, ha létezik $z \in K_1^* \setminus \{0\}$, $y \in K_2^*$ vektor úgy, hogy

$$z^T f(x) + y^T g(x) \geq 0 \ (x \in C).$$

Bizonyítás: Egyrészt ha egyszerre léteznének a tételben leírt tulajdonságú x és (y, z) vektorok, akkor abból

$$z^T f(x) = 0 = y^T g(x)$$

következne, és akkor a szokásos módon (az $f(x)$ vektorhoz húzva a $-K_1$ kúpba, és rajta túlhúzva a $-K_1$ kúpban egy tetszőleges $\mathcal{R}^k$-beli vektort) $z = 0$ adódna, ami ellentmondás.

A nemtriviális irányhoz tegyük fel, hogy az első rendszer nem megoldható, és legyen

$$S := \cup_{x \in C} \left\{ \{a\} \times \{b\} \subseteq \mathcal{R}^k \times \mathcal{R}^m : a >_{K_1} f(x),\ b \geq_{K_2} g(x) \right\}.$$

Könnyen belátható, hogy az S halmaz konvex (konvex halmaz vetülete, lásd még 12.3 bizonyítását), nemüres, továbbá az első rendszer megoldhatatlansága miatt az origó nem eleme. A második Hahn–Banach-tétel szerint léteznek $z \in \mathcal{R}^k$, $y \in \mathcal{R}^m$ vektorok úgy, hogy nem mind a kettő nullvektor, és

$$z^T a + y^T b \geq 0 \ (\{a\} \times \{b\} \subseteq S).$$

A bizonyítás ezután három lépésre bontható.

1. Először megmutatjuk, hogy $z \in K_1^*$, és $y \in K_2^*$. E célból legyen $\{a\} \times \{b\} \subseteq S$, ekkor tetszőleges $w_1 \in K_1$, $w_2 \in K_2$ vektorok, $\lambda, \mu \geq 0$ számok esetén $\{a + \lambda w_1\} \times \{b + \mu w_2\} \subseteq S$, amiből

$$z^T(a + \lambda w_1) + y^T(b + \mu w_2) \geq 0$$

adódik. Ez az egyenlőtlenség csak úgy állhat fenn tetszőleges $\lambda, \mu \geq 0$ számok esetén, ha $z^T w_1 \geq 0$, és $y^T w_2 \geq 0$, vagyis $z \in K_1^*$, $y \in K_2^*$ adódott.

2. Másodszor megmutatjuk, hogy

$$z^T f(x) + y^T g(x) \geq 0 \ (x \in C).$$

Legyen $x \in C$ tetszőleges elem, és válasszunk $a >_{K_1} 0$ elemet, ekkor

$$\{f(x) + a\} \times \{g(x)\} \subseteq S,$$

amiből

$$z^T(f(x) + a) + y^T g(x) \geq 0$$

adódik. Tartsunk a-val 0-hoz, máris megkapjuk a kívánt egyenlőtlenséget.

3. Végül belátjuk, hogy $z \neq 0$. Tegyük fel indirekt, hogy $z = 0$, ekkor az éppen bizonyított egyenlőtlenségből, abba az $x_0 \in C$ pontot helyettesítve adódik, hogy $y^T g(x_0) \geq 0$. Mivel $y \in K_2^*$, $g(x_0) \in -\operatorname{int} K_2$, azért $y^T g(x_0) = 0$, és akkor (ismét a behúzás-túlhúzás fogással igazolható) az y vektor merőleges az egész $\mathcal{R}^m$ térre, vagyis $y = 0$, ami $z = 0$ mellett ellentmond annak, hogy y és z legalább egyike nem nullvektor. □

Megszabadulhatunk a regularitási feltételtől, ha az alábbi módosított (regularizált) duál feladatot tekintjük

$$(GD') : \qquad \sup \inf \left\{ f(x) + y^T g(x) : x \in C' \right\}, \ y \in (K')^*,$$

ahol

$$K' := \Phi_K(-g(\mathbf{GP})), \text{ és } C' := C \cap g^{-1}(K' - K).$$

(Ha a gyenge Slater-feltétel teljesül, továbbá x_0 gyenge Slater-pont, akkor $-g(x_0) \in -g(\mathbf{GP}) \subseteq K'$, továbbá $-g(x_0) \in \operatorname{ri} K$. Tehát K egy lapja belemetsz K relatív belsejébe, amiből 7.2 c) miatt $K' = K$ adódik.)

12.22. Lemma: *Tegyük fel, hogy* $\mathbf{GP} \neq \emptyset$*, ekkor az alábbi állítások teljesülnek*
a) $g(\mathbf{GP}) + K'$ *nemüres, konvex halmaz;*
b) $g(\mathbf{GP}) \cap -\mathrm{ri}\, K' \neq \emptyset$*;*
c) $C' = g^{-1}(K' - K') \cap C$*;*
d) a g *leképezés* K'*-konvex a* C' *konvex halmazon.*

Bizonyítás: Az a) állításhoz elég annyit belátnunk, hogy a g leképezés K'-konvex a (konvex) $\mathbf{GP}$ halmazon. Legyen $x_1, x_2 \in \mathbf{GP}$, és legyen $0 \leq \lambda \leq 1$. Ekkor

$$g(\lambda x_1 + (1-\lambda)x_2) \in g(\mathbf{GP}),\ \ \lambda g(x_1) + (1-\lambda)g(x_2) \in -K',$$

mivel $-K'$ konvex és tartalmazza a $g(\mathbf{GP})$ halmazt. Így

$$\lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \in K' - K'.$$

Továbbá, mivel a g leképezés K-konvex,

$$\lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \in K.$$

Mivel 7.2 d) szerint $(K' - K') \cap K = K'$, azért

$$\lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \in K',$$

amit igazolni kellett.

b) Tegyük fel indirekt, hogy b) nem teljesül, ekkor a $g(\mathbf{GP}) + K'$ és $-\mathrm{ri}\, K'$ konvex halmazok sem metszenek össze, mivel $K' + \mathrm{ri}\, K' = \mathrm{ri}\, K'$. A második Hahn–Banach-tétel szerint létezik $a \in \mathcal{R}^m$ vektor úgy, hogy

$$a^T(g(\mathbf{GP}) + K') \leq 0 \leq -a^T K',\ \ a^T \mathrm{ri}\,(g(\mathbf{GP}) + K') < -a^T \mathrm{ri}\, K'.$$

Mivel $0 \in K'$, azért $a^T g(\mathbf{GP}) \leq 0$, továbbá (mivel $g(\mathbf{GP})$ a $-K'$ kúp része) $a^T g(\mathbf{GP}) \geq 0$ is teljesül. 5.1 segítségével könnyen belátható, hogy $a^T \mathrm{ri}\, K' < 0$. De akkor $\{x : a^T x = 0\} \cap K'$ a K' kúp a $-g(\mathbf{GP})$ halmazt tartalmazó valódi lapja, ami lehetetlen.

A c) állítás a következőképpen igazolható. Legyen x a $C \cap g^{-1}(K' - K)$ halmaz eleme, és válasszunk $\hat{x}$ elemet a $C \cap g^{-1}(-\mathrm{ri}\, K')$ halmazból (a már igazolt b) rész szerint ez megtehető). Ekkor léteznek $x_1' \in K'$, $x_1 \in K$ pontok úgy, hogy $g(x) = x_1' - x_1$ legyen. Tetszőleges $0 < \lambda < 1$ esetén jelölje x_λ a $\lambda x + (1-\lambda)\hat{x} \in C$ pontot, ekkor a g leképezés K-konvexitásából adódóan

$$g(x_\lambda) \in \lambda g(x) + (1-\lambda)g(\hat{x}) - K \subseteq \lambda x_1' + (1-\lambda)g(\hat{x}) - (K + \lambda x_1).$$

Elég kis $\lambda > 0$ számok esetén $\lambda x_1' + (1-\lambda)g(\hat{x})$ a $-\operatorname{ri} K'$, és így $-K$ eleme. Ezért x_λ primál megengedett megoldás, és így $g(x_\lambda)$ a $-K'$ halmaz eleme. Ismét g K-konvexitása miatt kis $\lambda > 0$ számok esetén

$$g(x) \in \lambda^{-1} g(x_\lambda) - \lambda^{-1}(1-\lambda)g(\hat{x}) + K \subseteq K - K'.$$

Oda jutottunk, hogy léteznek $x_2 \in K$ és $x_2' \in K'$ vektorok úgy, hogy $g(x) = x_1' - x_1 = x_2 - x_2'$. Ekkor $x_1 + x_2 = x_1' + x_2' \in K'$, és mivel K' a K kúp lapja, azért $x_1 \in K'$ (felhasználva 7.1 b)-t). Így $g(x) = x_1' - x_1$ az $K' - K'$ halmaz eleme, amit bizonyítani akartunk. A másik irányú tartalmazás nyilvánvaló.

d) Legyenek az x_1, x_2 vektorok a C' halmaz elemei. Legyen továbbá $0 \leq \lambda \leq 1$. Ekkor az $x_\lambda := \lambda x_1 + (1-\lambda)x_2$ vektor is a C' halmaz eleme (g K-konvexitása miatt), így a c) pont szerint $g(x_1), g(x_2), g(x_\lambda) \in K' - K'$. Ebből adódóan

$$\lambda g(x_1) + (1-\lambda)g(x_2) - g(x_\lambda) \in K' - K'$$

is teljesül. Mivel g K-konvex a C halmazon, azért

$$\lambda g(x_1) + (1-\lambda)g(x_2) - g(x_\lambda) \in (K' - K') \cap K = K'$$

(vö. 7.2 d)), amint azt igazolni kellett. □

12.23. Tétel: *Ha a (GP) program optimumértéke véges, akkor a (GP) és (GD') programok optimumértékei megegyeznek, és a (GD') program optimumértéke felvétetik.*

Bizonyítás: A (GP) programot ekvivalens módon úgy is felírhatjuk, mint

$$\inf f(x),\ g(x) \in -K',\ x \in C',$$

ugyanis a

$$\mathbf{GP} \subseteq g^{-1}(-K') \cap C' \subseteq g^{-1}(-K) \cap C = \mathbf{GP}$$

tartalmazáslánc miatt megengedett megoldásaik halmaza megegyezik. Ha most a g függvényt mint a C' halmazon értelmezett leképezést tekintjük, akkor a lemma szerint K'-konvex, és a C' halmazt a $K' - K'$ halmazba képezi. A $g(x) \in -K'$ feltételt alkalmasan átkoordinátázva, és elhagyva belőle a $0 \in 0$ feltételeket oda jutunk, hogy a lemma szerint létező $\hat{x}$ vektor, amelyre $\hat{x} \in \mathbf{GP}$, és $g(\hat{x}) \in -\mathrm{ri}\, K'$, valójában erős Slater-pontja az átkoordinátázott feladatnak. Erre alkalmazva 12.20-at, az átkoordinátázott program duáljának optimális megoldását "visszakoordinátázva", a kívánt tulajdonságú duál optimális megoldást nyerünk. □

13. Kuhn–Tucker-típusú tételek

Ebben a fejezetben az

$$(LP): \qquad \begin{cases} \inf f(x), \\ g_i(x) \le 0 \ (i = 1, \ldots, m), \\ h_j(x) = 0 \ (j = 1, \ldots, l), \\ x \in C \end{cases}$$

problémát vizsgáljuk, ahol $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz; $S \subseteq \mathcal{R}^n$ egy a C konvex halmazt tartalmazó, nyílt halmaz; $f, g_i, h_j : S \to \mathcal{R}$ $(i = 1, \ldots, m,\ j = 1, \ldots, l)$ differenciálható függvények, melyeknek gradiensleképezései folytonosak egy $x_0 \in C$ pontban. Szükséges feltételét keressük annak, hogy az x_0 pont az (LP) program lokális optimuma legyen.

E célból vizsgáljuk először a belső irányok, a csökkenési irányok és az érintő irányok halmazát.

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, továbbá $x_0 \in \mathcal{R}^n$. Ekkor az S-beli,

$$\begin{aligned} &\mathrm{conv}\,(]x_0, x_0 + \lambda z[\cup ((\mathrm{aff}\, S) \cap O(x_0 + \lambda z, \lambda\varepsilon))) = \\ &\qquad = \{x_0 + \lambda' z' : ||z - z'|| < \varepsilon,\ 0 < \lambda' \le \lambda,\ z' \in \mathrm{par}\, S\} \end{aligned}$$

alakú (aff S-sel megegyező affin burkú, relatív nyílt) halmazokat, ahol $\varepsilon > 0$, $\lambda > 0$, $z \in (\mathrm{par}\, S) \setminus \{0\}$, az S halmaz (x_0 csúcsú, z irányú) **csepp**jének nevezzük. A z irányt ekkor az S halmaz x_0 pontbeli **belső irány**ának nevezzük, (üres, vagy par S-sel megegyező affin burkú, relatív nyílt) halmazukat jelölje $ID(x_0, S)$. Nyilván $0 < \lambda' \le \lambda$ esetén $x_0 + \lambda' z \in \mathrm{ri}\, S$, speciálisan

$$ID(x_0, S) \subseteq (\mathrm{cone}\,((\mathrm{ri}\, S) - x_0)) \setminus \{0\}.$$

Ha $x_0 \in \operatorname{ri} S$, akkor

$$ID(x_0, S) = (\operatorname{par} S) \setminus \{0\},$$

az előbbi tartalmazás ekkor nem szigorú. Egy másik speciális esetben is garantálható az egyenlőség.

Azt mondjuk, hogy a $z \in \mathcal{R}^n$ vektor az S halmaz **megengedett irány**a az $x_0 \in \mathcal{R}^n$ pontban, ha létezik $\lambda > 0$ szám úgy, hogy $0 < \lambda' \leq \lambda$ esetén $x_0 + \lambda' z \in S$ legyen. Az S halmaz x_0 pontbeli megengedett irányainak kúpját jelölje $FD(x_0, S)$. Nyilván $ID(x_0, S) \subseteq FD(x_0, S)$.

13.1. Állítás: *Legyen $C \subseteq \mathcal{R}^n$ konvex halmaz, továbbá $x_0 \in C \setminus \operatorname{ri} C$. Ekkor*

$$FD(x_0, C) = \operatorname{cone}(C - x_0),$$
$$ID(x_0, C) = (\operatorname{cone}((\operatorname{ri} C) - x_0)) \setminus \{0\} = \operatorname{ri} FD(x_0, C),$$

speciálisan

$$ID(x_0, C)^* = FD(x_0, C)^* = (C - x_0)^*.$$

□

13.2. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $g : S \to \mathcal{R}$ differenciálható függvény, melynek gradiensleképezése folytonos az $x_0 \in S$ pontban. Legyen*

$$S' := \{x \in S : g(x) \leq 0\},$$

továbbá tegyük fel, hogy $x_0 \in S'$.

Ha $g(x_0) < 0$, akkor $x_0 \in \operatorname{int} S'$, így ekkor

$$ID(x_0, S') = \mathcal{R}^n \setminus \{0\}, \text{ és } ID(x_0, S')^* = \{0\}.$$

Tegyük fel most, hogy $g(x_0) = 0$, és $\nabla g(x_0) \neq 0$. Ekkor

$$ID(x_0, S') = \operatorname{int}\{-\nabla g(x_0)\}^*, \text{ és } ID(x_0, S')^* = -\nabla g(x_0)\mathcal{R}_+.$$

Bizonyítás: A tétel nemtriviális részéhez tegyük fel, hogy $x_0 \in S'$ olyan pont, amelyre $g(x_0) = 0$, $\nabla g(x_0) \neq 0$. Ekkor 10.29-ből adódóan létezik $x \in S$ pont, amelyre $g(x) < 0$. Ez az x pont $\operatorname{int} S'$-beli; látjuk, hogy $\operatorname{aff} S' = \mathcal{R}^n$. Ezért az $ID(x_0, S')$ halmaz affin burka is az egész tér (ha nem üres), a belső irányok halmaza nyílt lesz.

Válasszunk először tetszőleges $z \in ID(x_0, S')$ irányt. Ekkor speciálisan z az S' halmaz megengedett iránya is, így létezik $\lambda > 0$, hogy $0 < \lambda' \leq \lambda$ esetén $x_0 + \lambda' z \in S'$, vagyis $g(x_0 + \lambda' z) \leq 0$. A Lagrange-féle középértéktétel szerint létezik $0 < \kappa < 1$ úgy, hogy

$$0 \geq g(x_0 + \lambda z) = g(x_0 + \lambda z) - g(x_0) = (\nabla g(x_0 + \kappa \lambda z))^T (\lambda z)$$

legyen. Ezt az egyenlőségláncot a $\lambda > 0$ számmal leosztva, majd λ-val nullához tartva kapjuk, hogy $z \in \{-\nabla g(x_0)\}^*$. Ebből

$$ID(x_0, S') \subseteq \{-\nabla g(x_0)\}^*,$$

és így ($ID(x_0, S')$ nyílt)

$$ID(x_0, S') \subseteq \operatorname{int} \{-\nabla g(x_0)\}^*$$

következik.

A másik irányú tartalmazáshoz legyen $z \in \operatorname{int} \{-\nabla g(x_0)\}^*$ tetszőleges elem. Ekkor $-\alpha := z^T \nabla g(x_0) < 0$. Rögzítsünk egy $\hat{\delta} > 0$ számot úgy, hogy $O(x_0, \hat{\delta}) \subseteq S$ legyen. Válasszunk $0 < \delta_1, \delta_2 < \hat{\delta}$ számokat úgy, hogy

$$\begin{array}{ll} ||\nabla g(x)|| < ||\nabla g(x_0)|| + 1 & (x \in O(x_0, \delta_1)), \\ ||\nabla g(x) - \nabla g(x_0)|| < \frac{\alpha}{3||z||} & (x \in O(x_0, \delta_2)) \end{array}$$

teljesüljön (a folytonossági feltétel miatt ez megtehető). Legyen most

$$\delta := \min\{\delta_1, \delta_2\},\ \varepsilon := \frac{\alpha}{3(||\nabla g(x_0)|| + 1)},\ \lambda := \frac{\delta}{\varepsilon + ||z||}.$$

Ekkor $||z' - z|| < \varepsilon$, $0 < \lambda' \leq \lambda$ esetén $||\lambda' z'|| < \delta$, hiszen

$$||z'|| = ||z' - z + z|| \leq ||z' - z|| + ||z|| < \varepsilon + ||z||.$$

Ezért a

$$\operatorname{conv} (]x_0, x_0 + \lambda z[\cup O(x_0 + \lambda z, \lambda \varepsilon))$$

csepp az S halmaz része. A Lagrange-féle középértéktétel szerint $||z' - z|| < \varepsilon$, $0 < \lambda' \leq \lambda$ esetén létezik $0 < \kappa < 1$ szám úgy, hogy

$$g(x_0 + \lambda' z') - g(x_0) = (\nabla g(x_0 + \kappa \lambda' z'))^T (\lambda' z')$$

legyen. Ezért ekkor a $g(x_0+\lambda' z')$ érték az alábbi módon becsülhető felülről:

$$\begin{aligned} g(x_0+\lambda' z') &= g(x_0+\lambda' z') - g(x_0) = (\nabla g(x_0+\kappa\lambda' z'))^T(\lambda' z') = \\ &= \lambda'[(\nabla g(x_0))^T z + (\nabla g(x_0+\kappa\lambda' z') - \nabla g(x_0))^T z + \\ &\qquad + (\nabla g(x_0+\kappa\lambda' z'))^T(z'-z)] \leq \\ &\leq \lambda'[-\alpha + \tfrac{\alpha}{3||z||}||z|| + (||\nabla g(x_0)||+1)\varepsilon] \leq \\ &\leq -\tfrac{\lambda'\alpha}{3} < 0. \end{aligned}$$

Tehát egy x_0 csúcsú, z irányú csepp az S' halmaz része. Ebből látszik az $\operatorname{int}\{-\nabla g(x_0)\}^* \subseteq ID(x_0, S')$ tartalmazás is. □

Speciális belső irányok a csökkenési irányok.

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, $x_0 \in S$, továbbá $f : S \to \mathcal{R}$. Azt mondjuk, hogy a z irány az f függvény **csökkenési iránya** az x_0 pontban, ha létezik az

$$S(f) := \{x \in S : f(x) < f(x_0)\}$$

halmaznak x_0 csúcsú, z irányú cseppje. A csökkenési irányok halmazát jelölje $DD(x_0, S, f)$. Más szóval

$$DD(x_0, S, f) := ID(x_0, S(f)).$$

13.3. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz; $f : S \to \mathcal{R}$ differenciálható függvény, melynek gradiensleképezése folytonos egy $x_0 \in S$ pontban. Ha az $x_0 \in S$ pontban $\nabla f(x_0) \neq 0$, akkor*

$$DD(x_0, S, f) = \operatorname{int}\{-\nabla f(x_0)\}^*, \text{ és } DD(x_0, S, f)^* = -\nabla f(x_0)\mathcal{R}_+.$$

Bizonyítás: Legyen

$$\hat{S}(f) := \{x \in S : f(x) - f(x_0) \leq 0\},$$

ekkor nyilván $S(f) \subseteq \hat{S}(f)$. Az x_0 pontból $-\nabla f(x_0)$ irányba indulva találunk $S(f)$ halmazbeli pontot (különben

$$0 \leq f'(x_0; -\nabla f(x_0)) = -||\nabla f(x_0)||^2 < 0$$

lenne). Az f függvény folytonossága miatt az $S(f)$ halmaz belseje nemüres, így az $\hat{S}(f)$ halmaz belseje sem üres. De akkor felhasználva még az előző tételt is

$$DD(x_0, S, f) = ID(x_0, S(f)) \subseteq ID(x_0, \hat{S}(f)) = \operatorname{int}\{-\nabla f(x_0)\}^*.$$

A fordított irányú tartalmazás abból az észrevételből következik, hogy 13.2 bizonyítása során nem csak azt láttuk be, hogy $\text{int}\{-\nabla f(x_0)\}^* \subseteq ID(x_0, \hat{S}(f))$, hanem azt is, hogy $\text{int}\{-\nabla f(x_0)\}^* \subseteq ID(x_0, S(f))$. $\square$

Legyen $S \subseteq \mathcal{R}^n$ nemüres halmaz, továbbá $x_0 \in S$. Azt mondjuk, hogy a $z \in \mathcal{R}^n$ irány az S halmaz **érintő iránya** az x_0 pontban, ha léteznek $x_k \in S$ $(k = 1, 2, \ldots)$ pontok és $\lambda_k > 0$ $(k = 1, 2, \ldots)$ számok úgy, hogy $\lambda_k \to 0$ $(k \to \infty)$, és

$$\frac{x_k - x_0}{\lambda_k} \to z \ (k \to \infty).$$

(Speciálisan $x_k \to x_0$ $(k \to \infty)$.) Az érintő irányok halmazát (nyilván kúpját) ekkor jelölje $TD(x_0, S)$.

13.4. Állítás: *Az érintő irányok kúpja zárt kúp.*

Bizonyítás: Nyilvánvaló, hogy tetszőleges $S \subseteq \mathcal{R}^n$ nemüres halmaz, $x_0 \in S$ esetén $TD(x_0, S)$ kúp. Zártságának igazolásához legyen $z_l \in TD(x_0, S)$ $(l = 1, 2, \ldots)$, és tegyük fel, hogy $z_l \to z$ $(l \to \infty)$. Az érintő irányok kúpjának definíciójából adódóan minden $l = 1, 2, \ldots$ esetén léteznek $\lambda_{kl} > 0$ $(k = 1, 2, \ldots)$ számok és $x_{kl} \in S$ $(k = 1, 2, \ldots)$ elemek úgy, hogy

$$\lambda_{kl} \to 0 \ (k \to \infty), \text{ és } \frac{x_{kl} - x_0}{\lambda_{kl}} \to z_l \ (k \to \infty).$$

Tetszőleges $l = 1, 2, \ldots$ esetén választható $k(l)$ index úgy, hogy $k(1) < k(2) < \ldots$ legyen, továbbá

$$0 < \lambda_{k(l)l} < \frac{1}{l}, \text{ és } \left\| \frac{x_{k(l)l} - x_0}{\lambda_{k(l)l}} - z_l \right\| < \frac{1}{l}$$

teljesüljön. Ekkor

$$\lambda_{k(l)l} \to 0 \ (l \to \infty), \text{ és } \frac{x_{k(l)l} - x_0}{\lambda_{k(l)l}} \to z \ (l \to \infty),$$

vagyis $z \in TD(x_0, S)$. $\square$

Konvex halmazok esetén az érintő irányok kúpja egy egyszerű formulával írható le.

13.5. Állítás: *Legyen $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz, továbbá $x_0 \in C$. Ekkor $TD(x_0, C) = \mathrm{cl}\,((C - x_0)\mathcal{R}_+)$.*

Bizonyítás: Tetszőleges $x \in C$ esetén

$$x_k := x_0 + \frac{1}{k}(x - x_0) \in C \ (k = 1, 2, \ldots),$$

ami $\lambda_k := 1/k$ $(k = 1, 2, \ldots)$ választással mutatja, hogy $x - x_0 \in TD(x_0, C)$, tehát $C - x_0 \subseteq TD(x_0, C)$. Az előző állítás szerint ebből $\mathrm{cl}\,((C - x_0)\mathcal{R}_+) \subseteq TD(x_0, C)$ is következik.

A fordított irányú tartalmazáshoz legyen $z \in TD(x_0, C)$. Ekkor léteznek $\lambda_k > 0$ $(k = 1, 2, \ldots)$ számok és $x_k \in C$ $(k = 1, 2, \ldots)$ pontok úgy, hogy $\lambda_k \to 0$ $(k \to \infty)$, és

$$\frac{x_k - x_0}{\lambda_k} \to z \ (k \to \infty).$$

Mivel $x_k \in C$, azért $(1/\lambda_k)(x_k - x_0) \in (C - x_0)\mathcal{R}_+$ $(k = 1, 2, \ldots)$, és akkor e sorozat limeszpontja, $z \in \mathrm{cl}\,((C - x_0)\mathcal{R}_+)$. □

Az érintő irányok kúpjáról szóló alapvető tétel előtt elevenítsük fel az implicitfüggvény-tételt és két egyszerű következményét.

13.6. Tétel: (implicitfüggvény-tétel, [7]) *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $h_j : S \to \mathcal{R}$ $(j = 1, \ldots, l < n)$ differenciálható függvények, melyeknek gradiensleképezései folytonosak egy $s_0 \in S$ pontban. Legyen továbbá $s_0 = (x_0, y_0)$, ahol $x_0 \in \mathcal{R}^l$, $y_0 \in \mathcal{R}^{n-l}$, és tegyük fel, hogy $h(s_0) = 0$, és a $h'(s_0) \in \mathcal{R}^{l \times n}$ mátrix első l oszlopából álló részmátrixa (ennek jele természetesen legyen $h'_x(s_0)$) nemszinguláris.*

Ekkor létezik $\varepsilon_1, \varepsilon_2 > 0$ úgy, hogy $O(x_0, \varepsilon_1) \times O(y_0, \varepsilon_2) \subseteq S$, továbbá létezik egyértelműen φ leképezés, amelyre
a) a φ leképezés az $O(y_0, \varepsilon_2)$ gömbön van értelmezve, és értékeit az $O(x_0, \varepsilon_1)$ gömbből veszi fel, és
b) $y \in O(y_0, \varepsilon_2)$ esetén $h(\varphi(y), y) = 0$.

Erre a φ leképezésre teljesül, hogy $\varphi(y_0) = x_0$, továbbá, hogy φ differenciálható egy $O(y_0, \varepsilon_3)$ nyílt gömbön, és

$$\varphi'(y_0) = -(h'_x(s_0))^{-1} \cdot h'_y(s_0).$$

13.7. Tétel: *Legyen* $S \subseteq \mathcal{R}^n$ *nyílt halmaz,* $h_j : S \to \mathcal{R}$ $(j = 1, \ldots, n)$ *differenciálható függvények, melyeknek gradiensleképezései folytonosak egy* $x_0 \in S$ *pontban. Tegyük fel, hogy* $h(x_0) = 0$, *és* $h'(x_0) \in \mathcal{R}^{n \times n}$ *nemszinguláris.*

Ekkor léteznek $\varepsilon_1, \varepsilon_2 > 0$ *számok úgy, hogy* $O(x_0, \varepsilon_1) \subseteq S$, *továbbá létezik egyértelműen* φ *leképezés, amelyre*
a) a φ *leképezés az* $O(0, \varepsilon_2)$ *gömbön van értelmezve, és értékeit az* $O(x_0, \varepsilon_1)$ *gömbből veszi fel, és*
b) $y \in O(0, \varepsilon_2)$ *esetén* $h(\varphi(y)) = y$.

Erre a φ *leképezésre teljesül, hogy* $\varphi(0) = x_0$, *továbbá, hogy* φ *differenciálható egy* $O(0, \varepsilon_3)$ *nyílt gömbön, és*

$$\varphi'(0) = (h'(x_0))^{-1}.$$

Bizonyítás: Alkalmazzuk az implicitfüggvény-tételt a

$$\tilde{h}(x, y) := h(x) - y \ (x \in S, \ y \in \mathcal{R}^n)$$

függvényre, az $\tilde{S} := S \times \mathcal{R}^n$ nyílt halmazra és az $\tilde{s} := (x_0, 0)$ pontra. $\square$

13.8. Lemma: *Legyen* $S \subseteq \mathcal{R}^n$ *nyílt halmaz,* $h_j : S \to \mathcal{R}$ $(j = 1, \ldots, n)$ *differenciálható függvények, melyeknek gradiensleképezései folytonosak egy* $x_0 \in S$ *pontban. Tegyük fel, hogy* $h(x_0) = 0$, *és* $h'(x_0) \in \mathcal{R}^{n \times n}$ *nemszinguláris.*

Ekkor x_0 *az*

$$S' := \{x \in S : h(x) = 0\}$$

halmaz izolált pontja.

Bizonyítás: Legyen $\varepsilon_1, \varepsilon_2, \varphi$ az előző tételben leírt tulajdonságú számok, illetve leképezés. Ha x_0 nem izolált pontja az S halmaznak, akkor létezik az x_0 ponttól különböző $\tilde{x}_0 \in O(x_0, \varepsilon_1)$ pont úgy, hogy $h(\tilde{x}_0) = 0$. Változtassuk meg φ értékét a 0 pontban, legyen $\tilde{x}_0$. Ekkor egy a φ leképezéstől különböző $\tilde{\varphi}$ leképezést kapunk, amelyre teljesülnek az előző tételben leírt a) és b) tulajdonságok. Ez pedig ellentmond φ egyértelműségének. $\square$

13.9. Tétel: *Legyen $S \subseteq \mathcal{R}^n$ nyílt halmaz, $h_j : S \to \mathcal{R}$ $(j = 1, \ldots, l)$ differenciálható függvények, melyeknek gradiensleképezései folytonosak egy $x_0 \in S$ pontban. Legyen*

$$S' := \{x \in S : h(x) = 0\},$$

és tegyük fel, hogy az $x_0 \in S'$ pontban a $h'(x_0) \in \mathcal{R}^{l \times n}$ mátrix rangja l (speciálisan csak $l \leq n$ lehet). Ekkor

$$TD(x_0, S') = \operatorname{Ker} h'(x_0), \text{ és } TD(x_0, S')^* = \operatorname{Im}(h'(x_0)^T).$$

Bizonyítás: Tekintsük először azt az esetet, mikor $l = n$. Ekkor 13.8 szerint x_0 izolált pontja az S' halmaznak, tehát

$$TD(x_0, S') = \{0\}.$$

Mivel $h'(x_0)$ nemszingularitása miatt

$$\operatorname{Ker} h'(x_0) = \{0\},$$

azért a kívánt egyenlőség ebben az esetben fennáll.

Feltehető tehát, hogy $l < n$, és az általánosság megszorítása nélkül az is, hogy a $h'(x_0)$ mátrixnak éppen az első l oszlopa által alkotott részmátrixa nemszinguláris.

Lássuk be először a $TD(x_0, S') \subseteq \operatorname{Ker} h'(x_0)$ tartalmazást. E célból legyen $z \in TD(x_0, S')$ tetszőleges nemnulla vektor. Ekkor léteznek pozitív λ_k $(k = 1, 2, \ldots)$ számok és $x_k \in S'$ $(k = 1, 2, \ldots)$ elemek úgy, hogy $\lambda_k \to 0$ $(k \to \infty)$, és

$$\frac{x_k - x_0}{\lambda_k} \to z \ (k \to \infty).$$

A $h'(x_0)$ deriváltleképezés definíciójából adódóan ugyanakkor

$$\frac{h(x_k) - h(x_0) - h'(x_0)(x_k - x_0)}{||x_k - x_0||} \to 0 \ (k \to \infty).$$

Mivel $h(x_k) = 0 = h(x_0)$, azért a számlálót és a nevezőt is leosztva a $\lambda_k > 0$ számmal, majd k-val a végtelenhez tartva a $-h'(x_0)z = 0$ egyenlőséghez jutunk, vagyis $z \in \operatorname{Ker} h'(x_0)$ teljesül.

A másik irányú tartalmazáshoz legyen $z \in \operatorname{Ker} h'(x_0)$ nemnulla elem, $z = (a, b)$, ahol $a \in \mathcal{R}^l$, $b \in \mathcal{R}^{n-l}$. Hasonlóan tagolva legyen $x_0 = (u_0, v_0)$, ahol $u_0 \in \mathcal{R}^l$, $v_0 \in \mathcal{R}^{n-l}$. Ekkor természetes jelölésekkel

$$h'_u(x_0)a + h'_v(x_0)b = 0,$$

és $h'_u(x_0) \in \mathcal{R}^{l \times l}$ nemszinguláris mátrix. Az implicitfüggvény-tétel szerint létezik $\varepsilon > 0$ szám és az $O(v_0, \varepsilon)$ gömbön differenciálható $\varphi : O(v_0, \varepsilon) \to \mathcal{R}^l$ leképezés úgy, hogy

$$u_0 = \varphi(v_0), \text{ és } h(\varphi(v), v) = 0, \ (\varphi(v), v) \in S \ (v \in O(v_0, \varepsilon)),$$

továbbá

$$h'_u(x_0)\varphi'(v_0) + h'_v(x_0) = 0.$$

Ezt az egyenletet jobbról megszorozva a b vektorral, oda jutunk, hogy

$$h'_u(x_0)(\varphi'(v_0)b - a) = 0,$$

amiből $h'_u(x_0)$ nemszingularitása miatt $\varphi'(v_0)b - a = 0$ következik. Speciálisan $b \neq 0$, különben $z = (a, b) = 0$ lenne. Legyen

$$u_k := \varphi(v_0 + \frac{1}{k}b), \ v_k := v_0 + \frac{1}{k}b, \ x_k := (u_k, v_k) \ (1/k < \varepsilon/||b||).$$

Ekkor $v_k \in O(v_0, \varepsilon)$ miatt $h(\varphi(v_k), v_k) = 0$, vagyis $h(x_k) = 0$, tehát $x_k \in S'$. Ugyanakkor φ differenciálhatósága miatt

$$\frac{\varphi(v_0 + (1/k)b) - \varphi(v_0) - \varphi'(v_0)((1/k)b)}{(1/k)||b||} \to 0 \ (k \to \infty),$$

vagyis

$$\frac{u_k - u_0}{1/k} \to a \ (k \to \infty).$$

Mivel ugyanakkor triviálisan

$$\frac{v_k - v_0}{1/k} \to b \ (k \to \infty),$$

azért

$$\frac{x_k - x_0}{1/k} \to z \ (k \to \infty),$$

és láthatóan $z \in TD(x_0, S')$. $\square$

13.10. Következmény: *Legyen* $S \subseteq \mathcal{R}^n$ *nyílt halmaz,* $a_j \in \mathcal{R}^n$, $\beta_j \in \mathcal{R}$ $(j = 1, \ldots, l)$. *Tegyük fel, hogy az*

$$S' := \left\{x \in S : a_j^T x = \beta_j \ (j = 1, \ldots, l)\right\}$$

halmaz nemüres, legyen $x_0 \in S'$. *Az* a_j *vektorokból, mint oszlopvektorokból alkotott mátrixot jelölje* A. *Ekkor*

$$TD(x_0, S') = \operatorname{Ker}(A^T), \text{ és } TD(x_0, S')^* = \operatorname{Im} A.$$

Bizonyítás: Azt az esetet, mikor az a_j vektorok lineárisan függetlenek, elintézi 13.9. Továbbá S' nemüres volta miatt az a_j vektorok közül csak $\operatorname{Im} A$ egy bázisát meghagyva sem S', sem $\operatorname{Ker}(A^T)$ nem változik, így a lineáris függetlenség feltehető. □

13.11. Tétel: (második Dubovickij–Miljutyin-tétel) *Tegyük fel, hogy* x_0 *az* (LP) *program lokális optimuma, továbbá teljesülnek az alábbi a), b), c) és d) feltételek*
a) ha legalább egy nemaffin h_j *függvény szerepel a programban, akkor a* C *halmaz belseje nem üres;*
b) a $K_{-1} := TD(x_0, \{x \in S : h_j(x) = 0 \ (j = 1, \ldots, l)\})$ *halmaz konvex kúp;*
c) a $K_0 := DD(x_0, S, f)$ *halmaz konvex kúp nemüres belseje;*
d) a $K_i := ID(x_0, \{x \in S : g_i(x) \leq 0\})$ $(i = 1, \ldots, m)$ *halmazok mindegyike konvex kúp nemüres belseje.*

Ekkor léteznek $c_i \in K_i^*$ $(i = -1, \ldots, m)$ *nem mind nulla vektorok úgy, hogy*

$$\begin{array}{ll} (-\sum_{i=-1}^m c_i)^T (x - x_0) \geq 0 \ (x \in C), & \text{ha } x_0 \notin \operatorname{ri} C \\ \sum_{i=-1}^m c_i \in (\operatorname{par} C)^\perp, & \text{ha } x_0 \in \operatorname{ri} C. \end{array}$$

Bizonyítás: Tegyük fel először, hogy $x_0 \notin \operatorname{ri} C$.

Tekintsük azt az esetet, mikor a programban szereplő h_j függvények legalább egyike nemaffin, ekkor az a) feltétel szerint $\operatorname{int} C \neq \emptyset$, és akkor $ID(x_0, C)$ a konvex $FD(x_0, C)$ kúp nemüres belseje.

Elég megmutatnunk, hogy

$$K_{-1} \cap ID(x_0, C) \cap \bigcap_{i=0}^m K_i = \emptyset,$$

akkor ugyanis az első Dubovickij–Miljutyin-tétel (6.9) szerint léteznek

$$c_i \in K_i^* \ (i=-1,\ldots,m),\ \hat{c} \in ID(x_0,C)^*$$

nem mind nulla vektorok úgy, hogy

$$\hat{c} + \sum_{i=-1}^{m} c_i = 0,$$

és akkor $c_i \in K_i^*$ $(i=-1,\ldots,m)$ megfelel (vö. 13.1).

Tegyük fel indirekt, hogy létezik

$$z \in K_{-1} \cap ID(x_0,C) \cap \bigcap_{i=0}^{m} K_i$$

vektor. Ekkor létezik a C halmaznak olyan x_0 csúcsú, z irányú cseppje, amely az

$$\{x \in C : f(x) < f(x_0),\ g_i(x) \le 0\ (i=1,\ldots,m)\}$$

halmaz része. Legyen ez a csepp éppen a

$$\operatorname{conv}\left(]x_0, x_0+\lambda z[\cup O(x_0+\lambda z, \lambda\varepsilon)\right)$$

halmaz, ahol $\varepsilon, \lambda > 0$. Léteznek továbbá $\lambda_k > 0$ $(k=1,2,\ldots)$ számok és $x_k \in S$ $(k=1,2,\ldots)$ elemek úgy, hogy

$$\lambda_k \to 0\ (k\to\infty),\ h_j(x_k) = 0\ (j=1,\ldots,l; k=1,2,\ldots)$$
$$\tfrac{x_k - x_0}{\lambda_k} \to z\ (k \to \infty).$$

Nyilván

$$x_k = x_0 + \lambda_k z + (x_k - x_0 - \lambda_k z),$$

továbbá elég nagy k esetén

$$\frac{x_k - x_0 - \lambda_k z}{\lambda_k} \in O(0,\varepsilon), \text{ és } \lambda_k < \lambda,$$

így elég nagy k esetén

$$x_k \in x_0 + \lambda_k z + \lambda_k O(0,\varepsilon) = x_0 + \lambda_k O(z,\varepsilon),$$

vagyis x_k a csepp eleme. Tehát az x_k vektorok elég nagy k esetén az (LP) program $f(x_0)$-nál kisebb célfüggvényértékű megengedett megoldásai, ugyanakkor $x_k \to x_0$ $(k \to \infty)$, ami ellentmond annak, hogy x_0 lokális optimum.

Vizsgáljuk most azt az esetet, mikor az (LP) program minden egyenlőség-feltétele lineáris, vagyis $h_j(x) = a_j^T x - \beta_j$ valamely $a_j \in \mathcal{R}^n$, $\beta_j \in \mathcal{R}$ esetén $(j = 1, \ldots, l)$. Az a_j vektorokból, mint oszlopvektorokból, alkotott mátrixot jelölje A, ekkor 13.10 szerint $K_{-1} = \mathrm{Ker}\,(A^T)$ altér, és továbbra is $ID(x_0, C) = \mathrm{ri}\, FD(x_0, C)$.

Megmutatjuk, hogy

$$K_{-1} \cap ID(x_0, C) \cap \bigcap_{i=0}^{m} K_i = \emptyset.$$

Tegyük fel indirekt, hogy létezik

$$z \in K_{-1} \cap ID(x_0, C) \cap \bigcap_{i=0}^{m} K_i.$$

Most is létezik a C halmaznak x_0 csúcsú, z irányú cseppje, amely az

$$\{x \in C : f(x) < f(x_0),\ g_i(x) \leq 0\ (i = 1, \ldots, m)\}$$

halmaz része. Mivel $z \in \mathrm{Ker}\,(A^T)$, azért $0 < \lambda' \leq \lambda$ esetén $x_0 + \lambda' z$ az (LP) program $f(x_0)$-nál kisebb célfüggvényértékű megengedett megoldása, ami ellentmond annak, hogy x_0 lokális optimum. Ezért

$$(\mathrm{ri}\, K_{-1}) \cap (\mathrm{ri}\,\mathrm{cone}\,(C - x_0)) \cap \bigcap_{i=0}^{m} K_i = \emptyset,$$

és akkor a többkúpos Stiemke-tétel szerint léteznek

$$c_i \in K_i^*\ (i = -1, \ldots, m),\ \hat{c} \in (C - x_0)^*$$

elemek úgy, hogy

$$\hat{c} + \sum_{i=-1}^{m} c_i = 0,$$

de vagy valamelyik $c_i \notin -K_i^*$, vagy $\hat{c} \notin -(C - x_0)^*$. Most sem lehet minden $c_i = 0$ $(i = -1, \ldots, m)$, különben $\hat{c}$ is nullvektor lenne, ezért c_i $(i = -1, \ldots, m)$ megfelel.

Végül az $x_0 \in \operatorname{ri} C$ eset az $x_0 \notin \operatorname{ri} C$ esethez hasonlóan intézhető el, mivel itt $ID(x_0, C) = (\operatorname{par} C) \setminus \{0\}$, és így

$$K_{-1} \cap (\operatorname{par} C) \cap \bigcap_{i=0}^{m} K_i = \emptyset.$$

□

13.12. Tétel: *Tegyük fel, hogy x_0 az (LP) program lokális optimuma, továbbá teljesül a 13.11-beli a) feltétel.*

(Fritz John) Ekkor léteznek $\lambda_0, \lambda_1, \ldots, \lambda_m \geq 0$ és $\mu_1, \ldots, \mu_l \in \mathcal{R}$ nem mind nulla szorzók úgy, hogy a

$$\hat{c} := \lambda_0 \nabla f(x_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{l} \mu_j \nabla h_j(x_0)$$

jelöléssel

$$\begin{array}{ll} \hat{c}^T(x - x_0) \geq 0 \ (x \in C), & \textit{ha } x_0 \notin \operatorname{ri} C \\ \hat{c} \in (\operatorname{par} C)^{\perp}, & \textit{ha } x_0 \in \operatorname{ri} C, \end{array}$$

továbbá

$$\lambda_i g_i(x_0) = 0 \ (i = 1, \ldots, m).$$

(Karush–Kuhn–Tucker) Ha $x_0 \in \operatorname{int} C$, akkor még $\lambda_0 = 1$ is feltehető, ha az alábbi feltételek egyike fennáll:
a) (Arrow–Hurwicz–Uzawa-feltétel) a $\nabla h_j(x_0)$ $(j = 1, \ldots, l)$ vektorok lineárisan függetlenek, továbbá létezik $z \in \mathcal{R}^n$ vektor úgy, hogy

$$\begin{array}{l} (\nabla g_i(x_0))^T z < 0 \ (1 \leq i \leq m,\ g_i(x_0) = 0), \textit{ és} \\ (\nabla h_j(x_0))^T z = 0 \ (1 \leq j \leq l); \end{array}$$

vagy az erősebb
b) (lineáris függetlenségi feltétel) a $\nabla g_i(x_0)$ $(1 \leq i \leq m,\ g_i(x_0) = 0)$, $\nabla h_j(x_0)$ $(j = 1, \ldots, l)$ vektorok lineárisan függetlenek;
vagy
c)

$$\left. \begin{array}{l} (\nabla g_i(x_0))^T y \leq 0 \ (1 \leq i \leq m,\ g_i(x_0) = 0), \\ (\nabla h_j(x_0))^T y = 0 \ (1 \leq j \leq l) \end{array} \right\} \Rightarrow y \in FD(x_0, \mathbf{LP}).$$

(Itt **LP** *az (LP) program megengedett megoldásainak halmazát jelöli.)*

Bizonyítás: Vegyük észre, hogy az alábbi esetekben könnyen találunk $\lambda_i \geq 0$ $(i = 0, \ldots, m)$, $\mu_j \in \mathcal{R}$ $(j = 1, \ldots, l)$ nem mind nulla szorzókat úgy, hogy

$$\hat{c} = 0,\ \lambda_i g_i(x_0) = 0\ (i = 1, \ldots, m)$$

legyen:
– $\nabla f(x_0) = 0$ (ekkor legyen $\lambda_0 := 1$, a többi szorzó nulla);
– $\nabla g_i(x_0) = 0$ valamely $1 \leq i \leq m$, $g_i(x_0) = 0$ esetén (ekkor legyen $\lambda_i := 1$, a többi szorzó nulla);
– a $\nabla h_j(x_0)$ $(j = 1, \ldots, l)$ vektorok lineárisan összefüggnek (ekkor a 0 vektor előáll lineáris kombinációjukként nem mind nulla $\mu_j \in \mathcal{R}$ $(j = 1, \ldots, l)$ súlyokkal, a többi szorzó legyen nulla).

Feltehető tehát, hogy a fenti esetek egyike sem áll fenn, vagyis
– $\nabla f(x_0) \neq 0$;
– $1 \leq i \leq m$, $g_i(x_0) = 0$ esetén $\nabla g_i(x_0) \neq 0$;
– a $\nabla h_j(x_0)$ $(j = 1, \ldots, l)$ vektorok lineárisan függetlenek.

Ekkor 13.9, 13.3 és 13.2 szerint, az előző tételbeli jelölésekkel
– K_{-1} konvex kúp, és $K_{-1}^* = \operatorname{Im}(h'(x_0))^T$;
– K_0 konvex kúp nemüres belseje, és $K_0^* = -\nabla f(x_0)\mathcal{R}_+$;
– $K_i = \mathcal{R}^n \setminus \{0\}$, ha $1 \leq i \leq m, g_i(x_0) < 0$, továbbá K_i konvex kúp nemüres belseje, és $K_i^* = -\nabla g_i(x_0)\mathcal{R}_+$, ha $1 \leq i \leq m, g_i(x_0) = 0$.

Tekintsük először az $x_0 \notin \operatorname{ri} C$ esetet. Az előző tétel bizonyítása során láttuk, hogy ekkor

$$K_{-1} \cap ID(x_0, C) \cap \bigcap_{i=0}^{m} K_i = \emptyset,$$

amiből

$$K_{-1} \cap ID(x_0, C) \cap K_0 \cap \bigcap \{K_i : 1 \leq i \leq m, g_i(x_0) = 0\} = \emptyset$$

adódik. Mint az előző tétel bizonyításában, az első Dubovickij–Miljutyin-, illetve a többkúpos Stiemke-tételből adódik, hogy léteznek $c_{-1} \in K_{-1}^*$, $c_0 \in K_0^*$, $c_i \in K_i^*$ $(1 \leq i \leq m,\ g_i(x_0) = 0)$ nem mind nulla vektorok úgy, hogy $c_i := 0$ $(1 \leq i \leq m,\ g_i(x_0) < 0)$ mellett

$$\left(-\sum_{i=-1}^{m} c_i\right)^T (x - x_0) \geq 0\ (x \in C).$$

A fentiek szerint választhatók $\lambda_0 \geq 0$, $\lambda_i \geq 0$ $(1 \leq i \leq m,\ g_i(x_0) = 0)$ és $\mu_1, \ldots, \mu_l \in \mathcal{R}$ nem mind nulla szorzók úgy, hogy

$$c_{-1} = -\sum_{j=1}^{l} \mu_j \nabla h_j(x_0),\ c_0 = -\lambda_0 \nabla f(x_0),$$
$$c_i = -\lambda_i \nabla g_i(x_0)\ (1 \leq i \leq m, g_i(x_0) = 0).$$

Legyenek még $\lambda_i := 0$ $(1 \leq i \leq m,\ g_i(x_0) < 0)$, ekkor ezek a szorzók megfelelnek.

Az $x_0 \in \operatorname{ri} C$ eset hasonlóan intézhető el, mivel ekkor $ID(x_0, C) = (\operatorname{par} C) \setminus \{0\}$, és így

$$K_{-1} \cap (\operatorname{par} C) \cap K_0 \cap \bigcap \{K_i : 1 \leq i \leq m, g_i(x_0) = 0\} = \emptyset.$$

Most megmutatjuk, hogy ha $x_0 \in \operatorname{int} C$, továbbá teljesül az Arrow–Hurwicz–Uzawa-feltétel, akkor feltehető, hogy $\lambda_0 \neq 0$ (és akkor persze az is, hogy $\lambda_0 = 1$).

Tegyük fel indirekt, hogy $\lambda_0 = 0$. Mivel $x_0 \in \operatorname{int} C$, azért

$$\sum \{\lambda_i \nabla g_i(x_0) : 1 \leq i \leq m,\ g_i(x_0) = 0\} + \sum_{j=1}^{l} \mu_j \nabla h_j(x_0) = 0.$$

Legyen $z \in \mathcal{R}^n$ olyan vektor, amelynek létezését az Arrow–Hurwicz–Uzawa-feltétel garantálja, és szorozzuk meg az előbbi egyenlőséget skalárisan a z vektorral. Ekkor persze egyrészt nullát kapunk, másrészt a

$$\sum \left\{ \lambda_i z^T \nabla g_i(x_0) : 1 \leq i \leq m,\ g_i(x_0) = 0 \right\}$$

számot, ami kisebb lenne, mint nulla, ha valamelyik benne szereplő λ_i szorzó pozitív lenne. Ezért minden itt szereplő $\lambda_i = 0$. Ebből

$$\sum_{j=1}^{l} \mu_j \nabla h_j(x_0) = 0$$

adódik, így a $\nabla h_j(x_0)$ $(j = 1, \ldots, l)$ vektorok lineáris függetlensége miatt $\mu_j = 0$ $(j = 1, \ldots, l)$, ami ellentmond annak, hogy a λ_i, μ_j szorzók nem mind nullák.

Most megmutatjuk, hogy a lineáris függetlenségi feltételből következik az Arrow–Hurwicz–Uzawa-feltétel.

Ehhez elég azt az állítást belátnunk, hogy ha $a_1, \ldots, a_m, b_1, \ldots, b_l \in \mathcal{R}^n$ lineárisan független vektorok, akkor létezik $z \in \mathcal{R}^n$ vektor úgy, hogy

$$a_1^T z < 0,\ \ldots,\ a_m^T z < 0,\ b_1^T z = 0,\ \ldots, b_l^T z = 0.$$

Legyen

$$L_i := \mathrm{lin}\,(\{a_1, \dots, a_m, b_1, \dots, b_l\} \setminus \{a_i\}),$$

és jelölje V_i az $L_i^\perp$ altérre való vetítésnek megfelelő vetítő mátrixot. Láttuk, hogy ez szimmetrikus, és a négyzete önmaga. Ezért $z_i := V_i a_i$ ($i = 1, \dots, m$) mellett, $a_i \notin L_i$ miatt

$$a_i^T z_i = a_i^T V_i a_i = a_i^T V_i^T V_i a_i = ||V_i a_i||^2 = ||z_i||^2 > 0,$$

és természetesen $a_j^T z_i = 0$ ($i \neq j$). Ebből már látszik, hogy a

$$z := -\sum_{i=1}^m z_i$$

vektor megfelel az Arrow–Hurwicz–Uzawa-feltételben.

Végül belátjuk, hogy a c) feltétel fennállása esetén is feltehető $\lambda_0 = 1$. Először is vegyük észre, hogy ha x_0 lokális optimuma az (LP) feladatnak, és $y \in FD(x_0, \mathbf{LP})$, akkor $\nabla f(x_0)^T y \geq 0$. (Valóban, elég kis $\lambda > 0$ számokra $x_0 + \lambda y \in \mathbf{LP}$ teljesül, így $f(x_0 + \lambda y) \geq f(x_0)$, és akkor

$$0 \leq \lim_{\lambda \to 0+} \frac{f(x_0 + \lambda y) - f(x_0)}{\lambda} = \nabla f(x_0)^T y.)$$

Összevetve ezt a c) feltétellel, látjuk, hogy

$$A_1^T y \geq 0,\ A_2^T y = 0 \text{ esetén } b^T y \geq 0,$$

ahol

$$\begin{aligned} A_1 &:= (-\nabla g_i(x_0) : 1 \leq i \leq m,\ g_i(x_0) = 0), \\ A_2 &:= (-\nabla h_j(x_0) : 1 \leq j \leq l),\ b := \nabla f(x_0). \end{aligned}$$

A Farkas-lemma (3.1) szerint léteznek x_1, x_2 vektorok úgy, hogy

$$A_1 x_1 + A_2 x_2 = b,\ x_1 \geq 0,$$

amivel a megfelelő λ_i, μ_j szorzók létezését igazoltuk. □

A c) feltétel teljesül például, ha az ott szereplő g_i, h_j függvények affinok. Bizonyítás nélkül megjegyezzük, hogy c)-nél gyengébb elégséges feltételt kapunk, ott $FD(x_0, \mathbf{LP})$ helyett a bővebb $TD(x_0, \mathbf{LP})$ halmazt írva. Az így nyert feltétel gyengébb a)-nál, b)-nél és a Kuhn–Tucker-feltételnél is (amelyben a megengedett és az érintő irányok halmaza közé eső ún. elérhető irányok halmaza szerepel), lásd [14], [1].

A Karush–Kuhn–Tucker-tétel 13.12-vel analóg bizonyítása (differenciál helyett szubdifferenciállal) megtalálható [13]-ban.

A következő állítás mutatja, hogy az erős Slater-feltételen kívül az Arrow–Hurwicz–Uzawa-feltétel és a lineáris függetlenségi feltétel is elégséges 12.10-hez és 12.13-hoz. Hasonló állítás mondható ki erős helyett gyenge Slater-pontra.

13.13. Állítás: *Tegyük fel, hogy az (LP) programban nincsenek egyenlőség-feltételek, továbbá az egyenlőtlenség-feltételeket leíró függvények konvexek is, és $\mathbf{LP} \cap \operatorname{int} C \neq \emptyset$. Ekkor az alábbi állítások ekvivalensek*
a) az $\hat{x} \in \mathbf{LP} \cap \operatorname{int} C$ pontban teljesül az Arrow–Hurwicz–Uzawa-feltétel;
b) létezik erős Slater-pont, vagyis $\tilde{x} \in C$ pont úgy, hogy $g(\tilde{x}) < 0$;
c) minden $\hat{x} \in C$ pontban teljesül az Arrow–Hurwicz–Uzawa-feltétel.

Bizonyítás: Megmutatjuk, hogy a)-ból következik b). Legyen $\hat{x} \in \operatorname{int} C$ az Arrow–Hurwicz–Uzawa-feltételt teljesítő pont, ekkor létezik $z \in \mathcal{R}^n$ vektor úgy, hogy

$$(\nabla g_i(\hat{x}))^T z < 0 \ (1 \leq i \leq m,\, g_i(\hat{x}) = 0).$$

Elég kis $\varepsilon > 0$ esetén minden $\tilde{x} \in O(\hat{x}, \varepsilon)$ pontra teljesül, hogy

$$\begin{gathered}\tilde{x} \in C,\, g_i(\tilde{x}) < 0 \ (1 \leq i \leq m,\, g_i(\hat{x}) < 0), \\ (\nabla g_i(\tilde{x}))^T z < 0 \ (1 \leq i \leq m,\, g_i(\hat{x}) = 0).\end{gathered}$$

Legyen $\lambda > 0$ olyan kicsi szám, hogy $\tilde{x} := \hat{x} + \lambda z \in O(\hat{x}, \varepsilon)$. Megmutatjuk, hogy $\tilde{x}$ erős Slater-pont. Világos, hogy $\tilde{x} \in C$, és hogy $1 \leq i \leq m$, $g_i(\hat{x}) < 0$ esetén $g_i(\tilde{x}) < 0$. Továbbá a g_i függvények konvexitásából adódóan

$$0 = g_i(\hat{x}) \geq g_i(\tilde{x}) + (\nabla g_i(\tilde{x}))^T(-\lambda z) > g_i(\tilde{x}) \ (1 \leq i \leq m,\, g_i(\hat{x}) = 0),$$

vagyis $\tilde{x}$ valóban erős Slater-pont.

Azt kell még igazolnunk, hogy b)-ből következik c) ("c)⇒a)" nyilvánvaló). Legyen $\tilde{x}$ erős Slater-pont, $\hat{x} \in C$ tetszőleges. Ekkor ismét a g_i függvények konvexitásából adódóan

$$\begin{aligned}0 > g_i(\tilde{x}) &\geq g_i(\hat{x}) + (\nabla g_i(\hat{x}))^T(\tilde{x} - \hat{x}) = \\ &= (\nabla g_i(\hat{x}))^T(\tilde{x} - \hat{x}) \ (1 \leq i \leq m,\, g_i(\hat{x}) = 0),\end{aligned}$$

vagyis $z := \tilde{x} - \hat{x}$ mutatja, hogy az $\hat{x}$ pontban teljesül az Arrow–Hurwicz–Uzawa-feltétel. □

14. Absztrakt dualitás: a perturbált duál

Az értékfüggvény és a perturbált duál fogalmát először az elméletileg könnyebben kezelhető lineáris programok esetében vezetjük be. Tekintsük az alábbi standard alakú lineáris programot:

$$(LIP) \qquad \inf c^T x,\ Ax = b,\ x \geq 0,$$

ahol $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$. Tegyük fel, hogy a (LIP) program v_{LIP} optimumértéke véges (és így felvétetik). Olyan (LID) programot szeretnénk találni, amelynek ugyanaz az optimumértéke, mint a (LIP) programnak, vagyis $v_{LIP} = v_{LID}$ teljesül. A (LID) program konstrukciójához vezető első lépésként legyen $p \in \mathcal{R}^m$, és tekintsük az alábbi lineáris programot, a (LIP) program perturbáltját:

$$(LIP_p) \qquad \inf c^T x,\ Ax = b + p,\ x \geq 0.$$

Jelölje a (LIP_p) program optimumértékét $v(p)$, ekkor a v függvény a (LIP) programhoz tartozó **értékfüggvény**.

14.1. Állítás: *Ha a (LIP) program optimumértéke felvétetik, akkor a hozzá tartozó $v : \mathcal{R}^m \to \overline{\mathcal{R}}$ értékfüggvény valódi, poliédrikus (és így zárt) konvex függvény. Továbbá a $\partial v(0)$ szubgradiens halmaz nemüres.*

Bizonyítás: A v értékfüggvény epigráfja nyilván éppen a

$$\Big\{(p, \alpha) :\ \text{létezik } x \in \mathcal{R}^n_+, \text{ amelyre } Ax = b + p,\ \beta \geq c^T x\ (\beta > \alpha)\Big\}$$

halmaz. Legyen $V := (0, E) \in \mathcal{R}^{(m+1)\times(n+m+1)}$, továbbá

$$P := \Big\{(x, p, \beta) : x \geq 0,\ Ax = b + p,\ c^T x \leq \beta\Big\},$$

ekkor nyilván

$$\operatorname{epi} v = \{(p, \alpha) : (p, \beta) \in VP\ (\beta > \alpha)\}.$$

Mivel a P poliéder lineáris képe, azért VP is poliéder, speciálisan zárt, ezért $\operatorname{epi} v = VP$. Ebből már látszik, hogy v poliédrikus (és így alulról félig folytonos) konvex függvény. A v konvex függvény valódi is, különben nem lehetne véges a $v(0) = v_{LIP}$ érték (a $(0, -1)$ vektor $\operatorname{rec} \operatorname{epi} v$ eleme lenne, így a $(0, v(0))$ epigráfbeli pontból ebbe az irányba indulva az epigráfban maradnánk). Végül 10.5 szerint $\partial v(0)$ nemüres. $\square$

A (LID) program következő választása természetesnek tűnik:

$$(LID) \qquad \sup f(0),\ f \text{ affin függvény és } f \le v.$$

Ezek szerint (LID) optimumértéke

$$\begin{aligned} v_{LID} = \sup\Big\{y^T 0 - \alpha : y \in \mathcal{R}^m,\ \alpha \in \mathcal{R},\ y^T p - \alpha \le v(p)\ (p \in \mathcal{R}^m)\Big\} = \\ = \sup\{-\alpha : y \in \mathcal{R}^m,\ \alpha \in \mathcal{R},\ v^c(y) \le \alpha\} = \\ = \sup\{-v^c(y) : y \in \mathcal{R}^m\} = v^{cc}(0). \end{aligned}$$

Látszik, hogy $v_{LIP} = v_{LID}$ pontosan akkor, ha $v(0) = v^{cc}(0)$. Mivel 14.1 szerint a v értékfüggvény zárt, azért $v(0) \in \mathcal{R}$ esetén $v(0) = v^{cc}(0)$, a primál és a duál optimumértékek megegyeznek. Továbbá az is könnyen ellenőrizhető, hogy y pontosan akkor optimális megoldása a (LID) programnak, ha $y \in \partial v(0)$. Mivel 14.1 szerint $\partial v(0) \ne \emptyset$ is teljesül, azért ezzel beláttuk a lineáris programokra vonatkozó erős dualitási tételt.

Az van hátra, hogy a (LID) programot a (LIP) programot leíró A, b, c függvényében írjuk fel. Ehhez számoljuk ki az értékfüggvény konjugáltját:

$$\begin{aligned} v^c(y) &= \sup\Big\{y^T p - v(p) : p \in \mathcal{R}^m\Big\} = \\ &= \sup_p \sup\Big\{y^T p - c^T x : x \ge 0,\ Ax = b + p\Big\} = \\ &= \sup\Big\{y^T p - c^T x : x \ge 0,\ Ax = b + p,\ p \in \mathcal{R}^m\Big\} = \\ &= \sup\Big\{y^T (Ax - b) - c^T x : x \ge 0\Big\} = \\ &= \sup\Big\{(A^T y - c)^T x : x \ge 0\Big\} - b^T y = \\ &= \begin{cases} -b^T y, & \text{ha } A^T y - c \le 0, \\ \infty & \text{különben.} \end{cases} \end{aligned}$$

Eszerint a duál program

$$(LID) \qquad \sup -v^c(y),\ y \in \mathcal{R}^m = \sup b^T y,\ A^T y \le c,\ y \in \mathcal{R}^m,$$

a standard alakú lineáris program szokásos duálja.

Általában tekintsük az

$$(AP) \qquad \inf f(x),\ x \in \mathcal{R}^n$$

absztrakt primál programot, ahol $f : \mathcal{R}^n \to \overline{\mathcal{R}}$ tetszőleges függvény. Tegyük fel, hogy adott még egy $\hat{f} : \mathcal{R}^n \times \mathcal{R}^m \to \overline{\mathcal{R}}$ függvény úgy, hogy

$\hat{f}(x, 0) = f(x)$ $(x \in \mathcal{R}^n)$ teljesül. Az ennek a "beágyazásnak" megfelelő **értékfüggvény** legyen v, ahol

$$v(p) := \inf \left\{ \hat{f}(x, p) : x \in \mathcal{R}^n \right\} \ (p \in \mathcal{R}^m).$$

A definíció nyilvánvaló következménye, hogy $v_{AP} = v(0)$. Az absztrakt primál programnak és a beágyazásnak megfelelő **absztrakt duál program**

$$(AD) \qquad \sup -v^c(y), \, y \in \mathcal{R}^m.$$

A gyenge dualitási tétel a fenti definíciók nyilvánvaló következménye.

14.2. Állítás: (absztrakt gyenge dualitás) *Minden* (AP) *absztrakt program esetén* $v^{cc}(0) = v_{AD} \leq v_{AP}$.

Bizonyítás: Nyilván minden $y \in \mathcal{R}^m$ esetén

$$v^c(y) = \sup \left\{ y^T p - v(p) : p \in \mathcal{R}^m \right\} \geq -v(0) = -v_{AP},$$

így a bikonjugált függvény definíciójából

$$v^{cc}(0) = \sup \{ -v^c(y) : y \in \mathcal{R}^m \} \leq v(0)$$

következik. □

Az (AP) absztrakt primál programot, adott beágyazása mellett, **normális**nak hívunk, ha $v_{AP} = v_{AD}$. A $v_{AP} - v_{AD}$ érték a **dualitási rés**.

14.3. Tétel: *Tegyük fel, hogy* v_{AP} *véges, és a* v *értékfüggvény konvex. Ekkor az* (AP) *program pontosan akkor normális, ha a* v *függvény alulról félig folytonos a* 0 *pontban.*

Bizonyítás: Először is ha (AP) normális program, akkor

$$v(0) = v^{cc}(0) = (\mathrm{cl}\, v)(0) \leq \underline{v}(0) \leq v(0),$$

és így $v(0) = \underline{v}(0)$, vagyis v alulról félig folytonos a 0 pontban.

Másrészt ha v konvex függvény, $v(0)$ véges, és v alulról félig folytonos a 0 pontban, akkor v valódi konvex függvény (különben $v(p) = -\infty$ lenne minden $p \in \mathrm{ri}\,\mathrm{dom}\, v$ esetén, és a 0 pontot $\mathrm{ri}\,\mathrm{dom}\, v$-beli pontsorozattal közelítve, az alulról félig folytonosság miatt, $v(0) = -\infty$ adódna). Ezért $v^{cc} = \mathrm{cl}\, v = \underline{v}$, és így $v^{cc}(0) = v(0)$. □

Most a v értékfüggvény konvexitásának egy elégséges feltételét adjuk.

14.4. Állítás: *Ha az $\hat{f} : \mathcal{R}^n \times \mathcal{R}^m \to \overline{\mathcal{R}}$ függvény konvex, akkor a $v : \mathcal{R}^m \to \overline{\mathcal{R}}$ értékfüggvény is konvex.*

Bizonyítás: A v függvény definíciójának nyilvánvaló következménye, hogy

$$\begin{aligned}\operatorname{epi} v &= \{(p,\alpha) : \alpha \geq v(p),\ \alpha \in \mathcal{R}\} = \\ &= \Big\{(p,\alpha) : \beta \geq \hat{f}(x_\beta, p) \text{ valamely } x_\beta \text{ esetén } (\beta > \alpha)\Big\} = \\ &\Big\{(p,\alpha) : (p,\beta) \in V(\operatorname{epi} \hat{f})\ (\beta > \alpha)\Big\},\end{aligned}$$

ahol $V := (0, E) \in \mathcal{R}^{(m+1)\times(n+m+1)}$. Mivel az $\operatorname{epi} \hat{f}$ halmaz konvex, azért $V(\operatorname{epi} \hat{f})$ is konvex halmaz, és így

$$v = \operatorname{epi}^{-1}(V(\operatorname{epi} \hat{f}))$$

konvex függvény (vö. 8.16). $\square$

Ha $v : \mathcal{R}^m \to \overline{\mathcal{R}}$ konvex függvény, amely a 0 pontban szubdifferenciálható, akkor $y \in \partial v(0)$ mellett teljesül, hogy $v(p) \geq v(0) + y^T p$ $(p \in \mathcal{R}^m)$. Ezért ekkor $\underline{v}(0) \geq v(0)$, amiből nyilvánvalóan $\underline{v}(0) = v(0)$ adódik, vagyis v alulról félig folytonos a 0 pontban, és 14.3 szerint az (AP) program normális. Valójában erősebb eredmény is igazolható, ha $\partial v(0)$ nemüres.

14.5. Tétel: *Ha v_{AP} véges, akkor pontosan akkor teljesül, hogy az $\hat{y}$ vektor optimális megoldása a duál programnak és $v_{AP} = v_{AD}$, ha $\hat{y} \in \partial v(0)$.*

Bizonyítás: Legyen $\hat{y} \in \partial v(0)$, ekkor a tétel előtti megjegyzés szerint $v_{AP} = v_{AD}$. Ezért ekkor

$$v(p) \geq v(0) + \hat{y}^T p = v^{cc}(0) + \hat{y}^T p = v_{AD} + \hat{y}^T p \ (p \in \mathcal{R}^m),$$

és így

$$-v^c(\hat{y}) = \inf\Big\{v(p) - \hat{y}^T p : p \in \mathcal{R}^m\Big\} \geq v_{AD},$$

ami az $\hat{y}$ vektor duál optimalitását is mutatja.

Megfordítva ha

$$-v^c(\hat{y}) = v_{AD} = v_{AP} = v(0),$$

akkor a Fenchel-egyenlőtlenség szerint

$$v(p) \geq -v^c(\hat{y}) + \hat{y}^T p = v(0) + \hat{y}^T p \ (p \in \mathcal{R}^m),$$

és így $\hat{y} \in \partial v(0)$. $\square$

14.4-ben kapcsolatot találtunk az $\operatorname{epi}\hat{f}$ és $\operatorname{epi}v$ epigráfok között. E kapcsolatot írjuk le most pontosabban.

14.6. Állítás: *Legyen* $V := (0, E) \in \mathcal{R}^{(m+1)\times(n+m+1)}$, *ekkor*

$$V \operatorname{epi}\hat{f} \subseteq \operatorname{epi}v \subseteq \operatorname{cl} V \operatorname{epi}\hat{f}.$$

Továbbá ha $\hat{f} : \mathcal{R}^n \times \mathcal{R}^m \to \overline{\mathcal{R}}$ *konvex függvény, akkor*

$$\operatorname{ri}\operatorname{epi}v = \operatorname{ri} V \operatorname{epi}\hat{f} = V \operatorname{ri}\operatorname{epi}\hat{f}.$$

Bizonyítás: Ha (p, α) a $V \operatorname{epi}\hat{f}$ halmaz eleme, akkor létezik $\hat{x} \in \mathcal{R}^n$ vektor úgy, hogy $(\hat{x}, p, \alpha)$ az $\operatorname{epi}\hat{f}$ halmaz eleme, vagyis $\alpha \geq \hat{f}(\hat{x}, p)$. Ezért

$$\alpha \geq v(p) = \inf \left\{ \hat{f}(x, p) : x \in \mathcal{R}^n \right\},$$

és így $(p, \alpha) \in \operatorname{epi}v$. Továbbá ha $(p, \alpha) \in \operatorname{epi}v$, akkor, mint azt 14.4 bizonyítása során láttuk, (p, β) a $V \operatorname{epi}\hat{f}$ halmaz eleme minden $\beta > \alpha$ esetén, és így $(p, \alpha) \in \operatorname{cl} V \operatorname{epi}\hat{f}$ is teljesül. Végül 14.4 szerint $\hat{f}$ konvexitása esetén v is konvex, így $\operatorname{epi}v$ nemüres, konvex halmaz. A már bizonyítottakból

$$\operatorname{ri}(V \operatorname{epi}\hat{f}) \subseteq \operatorname{epi}v \subseteq \operatorname{cl}(V \operatorname{epi}\hat{f})$$

adódik, és így 4.8 szerint $\operatorname{ri}(\operatorname{epi}v) = \operatorname{ri}(V \operatorname{epi}\hat{f})$. Az utolsó igazolandó egyenlőség 4.21 következménye. □

3.25 szerint ha $\operatorname{epi}\hat{f}$ poliéder, akkor $V \operatorname{epi}\hat{f}$ is poliéder, speciálisan zárt, és így $\operatorname{epi}v = V \operatorname{epi}\hat{f}$ poliéder. Ebből már következik, hogy ha v_{AP} véges, akkor $v_{AP} = v_{AD}$, és a duál optimális megoldások nemüres halmaza éppen $\partial v(0)$ (vö. 10.5, 14.5). Ha azonban $\hat{f}$ nem poliédrikus függvény, akkor a szubdifferenciálhatóság más elégséges feltételét kell keresnünk.

Tegyük fel, hogy az (AP) programhoz tartozó $\hat{f} : \mathcal{R}^n \times \mathcal{R}^m \to \overline{\mathcal{R}}$ függvény valódi konvex függvény. Azt mondjuk, hogy az (AP) program teljesíti a **Slater-feltétel**t, ha létezik $x \in \mathcal{R}^n$ vektor úgy, hogy

$$(x, 0) \in \operatorname{ri}\operatorname{dom}\hat{f}.$$

14.7. Állítás: *Az* (AP) *program pontosan akkor teljesíti a Slater-feltételt, ha* $0 \in \operatorname{ri}\operatorname{dom}v$.

Bizonyítás: 8.18 szerint

$$\operatorname{ri}\operatorname{epi} v = \{(p,\alpha) : \alpha \in \mathcal{R},\ \alpha > v(p),\ p \in \operatorname{ri}\operatorname{dom} v\}$$

és

$$\operatorname{ri}\operatorname{epi} \hat{f} = \left\{(x,p,\alpha) : \alpha \in \mathcal{R},\ \alpha > \hat{f}(x,p),\ (x,p) \in \operatorname{ri}\operatorname{dom} \hat{f}\right\}.$$

Továbbá 14.6 szerint $\operatorname{ri}\operatorname{epi} v = V \operatorname{ri}\operatorname{epi} \hat{f}$, ahol szokás szerint $V := (0, E) \in \mathcal{R}^{(m+1)\times(n+m+1)}$. Ebből pedig

$$(x,0) \in \operatorname{ri}(\operatorname{dom} \hat{f}) \Longleftrightarrow (x,0,\alpha) \in \operatorname{ri}(\operatorname{epi} \hat{f})\ (\alpha > f(x)) \Longleftrightarrow \\ \Longleftrightarrow (0,\alpha) \in \operatorname{ri}(\operatorname{epi} v) \Longleftrightarrow 0 \in \operatorname{ri}(\operatorname{dom} v)$$

következik, amit igazolni kellett. □

Most már könnyen ellenőrizhető a következő eredmény.

14.8. Tétel: (absztrakt erős dualitás) *Ha v_{AP} véges, és az (AP) program teljesíti a Slater-feltételt, akkor $v_{AP} = v_{AD}$, és a duál program optimumértéke felvétetik. Továbbá a duál program optimális megoldásainak halmaza éppen a $\partial v(0)$ halmaz.*

Bizonyítás: Mivel az (AP) program teljesíti a Slater-feltételt, azért 14.7 szerint $0 \in \operatorname{ri}(\operatorname{dom} v)$, így $v(0) = (\operatorname{cl} v)(0) = v^{cc}(0)$, és a v valódi konvex függvény szubdifferenciálható a 0 pontban (10.8). □

A fejezet hátralévő részében két példát elemzünk, és valamivel gyengébb formában igazoljuk 12.5-öt és 11.7 d)-t.

Tekintsük először a Lagrange-féle primál feladatot, az

$$\inf f_0(x),\ f_i(x) \le 0\ (1 \le i \le k);\ f_i(x) = 0\ (k+1 \le i \le m);\ x \in C$$

programot, ahol $f_i : C \to \mathcal{R}$ $(0 \le i \le k)$ konvex függvények, $f_i : \mathcal{R}^n \to \mathcal{R}$ $(k+1 \le i \le m)$ affin függvények, és $C \subseteq \mathcal{R}^n$ nemüres, konvex halmaz. A beágyazó függvény, $\hat{f}$ legyen

$$\hat{f}(x,p) := \begin{cases} f_0(x), & \text{ha } (x,p) \text{ megengedett,} \\ \infty & \text{különben,} \end{cases}$$

ahol az (x,p) megengedettsége alatt azt értjük, hogy

$$f_i(x) \le p_i\ (1 \le i \le k),\ f_i(x) = p_i\ (k+1 \le i \le m),\ x \in C.$$

Az értékfüggvény, v tehát

$$v(p) = \inf\{f_0(x) : (x,p) \text{ megengedett}\} \ (p \in \mathcal{R}^m).$$

Ekkor

$$\operatorname{dom} \hat{f} = \{(x,p) : (x,p) \text{ megengedett}\} = C' \cap M,$$

ahol

$$C' := \{(x,p) : f_i(x) \le p_i \ (1 \le i \le k),\ x \in C\},$$
$$M := \{(x,p) : f_i(x) = p_i \ (k+1 \le i \le m),\ x \in \mathcal{R}^n\}.$$

Itt $M \subseteq \mathcal{R}^{n+m}$ affin halmaz, így $\operatorname{ri} M = M$. Továbbá $C' \subseteq \mathcal{R}^{n+m}$ konvex halmaz, és nem nehéz belátni, hogy

$$\operatorname{ri} C' = \{(x,p) : f_i(x) < p_i \ (1 \le i \le k),\ x \in \operatorname{ri} C\}.$$

(A $\subseteq$ tartalmazás 8.18 megfelelő részének mintájára igazolható. A fordított tartalmazáshoz azt kell belátnunk, hogy C' minden eleme túlhúzható a jobb oldali halmaz bármely elemén a C' halmazban. Ez pedig 10.1 egyszerű következménye.)

Mivel $M \cap \operatorname{ri} C'$ nemüres ($x \in \operatorname{ri} C$ esetén

$$p := (f_1(x) + 1, \ldots, f_k(x) + 1, f_{k+1}(x), \ldots, f_m(x))^T$$

mellett (x,p) a metszet eleme), azért $\operatorname{ri}(\operatorname{dom} \hat{f}) = M \cap \operatorname{ri} C'$. Ezért ha létezik $\hat{x} \in \mathcal{R}^n$ vektor, amelyre

$$f_i(\hat{x}) < 0 \ (1 \le i \le k),\ f_i(\hat{x}) = 0 \ (k+1 \le i \le m),\ \hat{x} \in \operatorname{ri} C$$

teljesül, akkor $(\hat{x}, 0) \in \operatorname{ri}(\operatorname{dom} \hat{f})$, vagyis a primál program teljesíti a Slater-feltételt.

14.9. Állítás: *A fenti program értékfüggvényének konjugáltját az alábbi képlet határozza meg*

$$v^c(y) = \begin{cases} \sup\left\{\sum_{i=1}^m y_i f_i(x) - f_0(x) : x \in C\right\}, & \textit{ha } y_i \le 0 \ (1 \le i \le k), \\ \infty & \textit{különben.} \end{cases}$$

Bizonyítás: Legyen

$$C'' := \left\{(x,p) \in C \times \mathcal{R}^k : f_i(x) \le p_i \ (1 \le i \le k)\right\},$$

továbbá $y \in \mathcal{R}^m$ tetszőleges vektor. Ekkor a konjugált függvény definíciójából kapjuk, hogy

$$\begin{aligned}
v^c(y) &= \sup\left\{y^T p - v(p) : p \in \mathcal{R}^m\right\} = \\
&= \sup\nolimits_{p\in\mathcal{R}^m} \sup\left\{y^T p - f_0(x) : (x,p) \text{ megengedett}\right\} = \\
&= \sup\left\{y^T p - f_0(x) : (x,p) \text{ megengedett}\right\} = \\
&= \sup\left\{\textstyle\sum_{i=1}^k y_i p_i + \sum_{i=k+1}^m y_i f_i(x) - f_0(x) : (x,p) \in C''\right\} = \\
&= \sup\left\{\textstyle\sum_{i=1}^m y_i f_i(x) - f_0(x) + \sum_{i=1}^k y_i(p_i - f_i(x)) : (x,p) \in C''\right\}.
\end{aligned}$$

Ha létezik $1 \le i \le k$ index úgy, hogy $y_i > 0$, akkor, mivel

$$(x, (f_1(x), \dots, f_i(x) + \lambda, \dots, f_k(x))^T) \in C''$$

minden $\lambda \ge 0$ esetén, az utolsó szumma tetszőleges nagy lehet, és így ebben az esetben $v^c(y) = \infty$. Ha pedig $y_i \le 0$ $(1 \le i \le k)$, akkor az utolsó szumma legfeljebb 0 minden $(x,p) \in C''$ esetén, tehát ebben az esetben

$$v^c(y) = \sup\left\{\sum_{i=1}^m y_i f_i(x) - f_0(x) : x \in C\right\}.$$

□

Ezek szerint a fenti primál program absztrakt duálisa

$$\begin{aligned}
&\sup\left\{-v^c(y) : y \in \mathcal{R}^m\right\} = \\
&= \sup\left\{-\sup\left\{\textstyle\sum_{i=1}^m y_i f_i(x) - f_0(x) : x \in C\right\} : y \in \mathcal{R}_-^k \times \mathcal{R}^{m-k}\right\} = \\
&= \sup\left\{\inf\left\{f_0(x) + \textstyle\sum_{i=1}^m y_i f_i(x) : x \in C\right\} : y \in \mathcal{R}_+^k \times \mathcal{R}^{m-k}\right\},
\end{aligned}$$

a szokásos Lagrange-duál. Most már kimondható 12.5 alábbi speciális esete:

14.10. Tétel: *Ha az $f_i : C \to \mathcal{R}$ $(0 \le i \le k)$ függvények konvexek, az $f_i : \mathcal{R}^n \to \mathcal{R}$ $(k+1 \le i \le m)$ függvények affinok, és $C \subseteq \mathcal{R}^n$ konvex halmaz, továbbá létezik $\hat{x} \in \operatorname{ri} C$ vektor úgy, hogy*

$$f_i(\hat{x}) < 0 \ (1 \le i \le k), \text{ és } f_i(\hat{x}) = 0 \ (k+1 \le i \le m)$$

teljesül rá, akkor a fenti Lagrange-féle primál és duál programok optimumértékei megegyeznek, és a duál program optimumértéke felvétetik, ha véges.

Bizonyítás: A beágyazó függvény epigráfja a Lagrange primál program esetében

$$\operatorname{epi} \hat{f} = \left\{(x, p, \alpha) : f_0(x) - \alpha \leq 0,\ (x, p) \in \operatorname{dom} \hat{f}\right\}$$

könnyen láthatóan konvex halmaz, mivel konvex függvények nívóhalmazainak metszete. Ezért 14.4 szerint a v értékfüggvény is konvex. Továbbá tudjuk, hogy a tételben leírt tulajdonságú $\hat{x}$ pont létezése ekvivalens azzal, hogy $(\hat{x}, 0) \in \operatorname{ri}(\operatorname{dom} \hat{f})$, vagyis azzal, hogy a primál program teljesíti a Slater-feltételt. A $v_{AP} \in \mathcal{R}$ esetben ezek után a tétel 14.8 következménye, a $v_{AP} = -\infty$ esetben pedig a gyenge dualitási tétel miatt egyezik meg az (AP) és (AD) programok optimumértéke. □

Vizsgáljuk most az

$$\inf f_0(x) + f_1(x),\ x \in \mathcal{R}^n$$

Fenchel primál programot, ahol $f_0, f_1 : \mathcal{R}^n \to \overline{\mathcal{R}}$ legyenek valódi konvex függvények. A beágyazó $\hat{f}$ függvény most legyen az alábbi

$$\hat{f}(x, p) := f_0(x - p) + f_1(x)\ (x \in \mathcal{R}^n,\ p \in \mathcal{R}^n).$$

Ekkor $\hat{f}$ nyilván szintén valódi konvex függvény. Továbbá

14.11. Állítás: *Pontosan akkor teljesül, hogy* $(x, 0) \in \operatorname{ri}(\operatorname{dom} \hat{f})$ *valamely* $x \in \mathcal{R}^n$ *esetén, ha a* $\operatorname{ri}(\operatorname{dom} f_0)$ *és a* $\operatorname{ri}(\operatorname{dom} f_1)$ *halmazok összemetszenek.*

Bizonyítás: Vegyük észre, hogy

$$\begin{aligned} \operatorname{dom} \hat{f} &= \left\{(x, p) : \hat{f}(x, p) < \infty\right\} = \\ &= \{(x, p) : x - p \in \operatorname{dom} f_0, \text{ és } x \in \operatorname{dom} f_1\}. \end{aligned}$$

Ebből már könnyen adódik, hogy a $V := (0, E) \in \mathcal{R}^{n \times 2n}$ jelöléssel

$$V \operatorname{dom} \hat{f} = \operatorname{dom} f_1 - \operatorname{dom} f_0,$$

és így

$$\begin{aligned} V \operatorname{ri}(\operatorname{dom} \hat{f}) &= \operatorname{ri}(V \operatorname{dom} \hat{f}) = \\ &= \operatorname{ri}(\operatorname{dom} f_1 - \operatorname{dom} f_0) = \\ &= \operatorname{ri}(\operatorname{dom} f_1) - \operatorname{ri}(\operatorname{dom} f_0). \end{aligned}$$

Látjuk, hogy $(\operatorname{ri} \operatorname{dom} f_1) \cap (\operatorname{ri} \operatorname{dom} f_0) \neq \emptyset$ pontosan akkor, ha a 0 pont $V \operatorname{ri}(\operatorname{dom} \hat{f})$-beli, pontosan akkor, ha az $(x, 0) \in \operatorname{ri}(\operatorname{dom} \hat{f})$ valamely $x \in \mathcal{R}^n$ esetén. □

Most kiszámoljuk a Fenchel primál program absztrakt duálját. Ehhez vegyük észre, hogy

$$\begin{aligned}
v^c(y) &= \sup\left\{y^T p - v(p) : p \in \mathcal{R}^n\right\} = \\
&= \sup_{p\in\mathcal{R}^n}\left\{\sup\left\{y^T p - f_0(x-p) - f_1(x) : x \in \mathcal{R}^n\right\}\right\} = \\
&= \sup\left\{y^T p - f_0(x-p) - f_1(x) : (x,p) \in \mathcal{R}^{2n}\right\} = \\
&= \sup\left\{y^T (p-x) - f_0(x-p) + y^T x - f_1(x) : (x,p) \in \mathcal{R}^{2n}\right\} = \\
&= \sup\left\{-y^T z - f_0(z) + y^T x - f_1(x) : (x,z) \in \mathcal{R}^{2n}\right\} = \\
&= \sup\left\{-y^T z - f_0(z) : z \in \mathcal{R}^n\right\} + \sup\left\{y^T x - f_1(x) : x \in \mathcal{R}^n\right\} = \\
&= f_0^c(-y) + f_1^c(y).
\end{aligned}$$

Ebből már látszik, hogy a Fenchel primál program absztrakt duálja

$$\sup -v^c(y),\ y \in \mathcal{R}^n = \sup -f_0^c(-y) - f_1^c(y),\ y \in \mathcal{R}^n,$$

a szokásos Rockafellar duál program ebben a speciális esetben. A megfelelő erős dualitási tétel 14.10 mintájára igazolható (vö. 11.7 d)).

14.12. Tétel: *Legyenek $f_0, f_1 : \mathcal{R}^n \to \overline{\mathcal{R}}$ valódi konvex függvények, és tegyük fel, hogy $(\operatorname{ri}\operatorname{dom} f_0) \cap (\operatorname{ri}\operatorname{dom} f_1)$ nemüres. Ekkor a fenti Fenchel primál és duál programok optimumértékei megegyeznek, és a duál program optimumértéke felvétetik, ha véges.* □

Függelékek

A. Eltolás politóppal: zártság

A 6.27 összefoglaló zártsági tételig vezető utunk során érdekes megfigyelni, hogy mely tételek bizonyítása használja a Bolzano–Weierstrass-tételt. A második Hahn–Banach-tételt például az első Hahn–Banach-tétel segítségével igazoljuk, amelynek bizonyítása Weierstrass tételére, és így végső soron a Bolzano–Weierstrass-tételre alapszik ([7]). Ismeretes azonban a második Hahn–Banach-tételnek egy a Bolzano–Weierstrass-tételt nem használó bizonyítása is (a harmadikként leírt bizonyítás ilyen, lásd az 5. fejezetet). A második Hahn–Banach-tételből pedig már azonnal következik az első Hahn–Banach-tétel azon speciális esete, mikor C_1 egy pont (szeparáljuk a C_1 körüli, a C_2 halmaztól diszjunkt, nyílt gömböt és C_2 relatív belsejét!), tehát 3.28 is. Ezen észrevételek segítségével igazolható az első Hahn–Banach-tételnek az a speciális esete, mikor C_1 politóp (A.1), továbbá hogy egy politóp és egy zárt, konvex halmaz összege zárt, konvex halmaz (A.2) — mindez a Bolzano–Weierstrass-tétel felhasználása nélkül! Ezek után az is könnyen belátható, hogy a 6.27-ben a)⇒d)-re adott két bizonyítás közül a)⇒b')⟺d) "algebraibb", mint az a)⟺c)⇒d), hiszen az előbbi végigvihető a Bolzano–Weierstrass-tétel nélkül is, míg az utóbbi, úgy tűnik, nem.

A most következő két állítást már beláttuk (általánosabban is, lásd 3.26 és 4.12), itt leírt bizonyításuk azonban nem használja a Bolzano–Weierstrass-tételt.

A.1. Lemma: *Ha a $Q \subseteq \mathcal{R}^d$ politóp diszjunkt a $C \subseteq \mathcal{R}^d$ zárt, konvex halmaztól, akkor Q és C erősen szeparálhatók.*

Bizonyítás: Feltehető, hogy $C = K$ kúp. (Ha nem így lenne, akkor szeparáljuk erősen a $\{1\} \times Q$ politópot és a $\operatorname{cl} K(C)$ zárt, konvex kúpot!)

Mivel Q politóp, azért léteznek $x_1, \ldots, x_k \in \mathcal{R}^d$ pontok úgy, hogy Q ezek konvex burka. Jelölje X azt a mátrixot, amelynek oszlopvektorai rendre $x_1, \ldots, x_k$. Legyen

$$S := \left\{ a \in \mathcal{R}^d : a^T x_1 < 0, \ldots, a^T x_k < 0 \right\}.$$

Azt kell megmutatnunk, hogy S és K^* összemetsz, akkor ugyanis $a \in S \cap K^*$ esetén

$$\max a^T Q = \max\{a^T x_1, \ldots, a^T x_k\} < 0 = \min a^T K,$$

vagyis Q és K erősen szeparálhatók.

Először is S nem üres, különben nem létezne olyan $a \in \mathcal{R}^d$ vektor sem, amelyre

$$X^T a \leq -\mathbf{1}$$

teljesül, és akkor 3.32 szerint létezne $y \in \mathcal{R}^k$ vektor úgy, hogy

$$Xy = 0,\ y \geq 0,\ \mathbf{1}^T y > 0,$$

amikor is $X(y/(\mathbf{1}^T y)) = 0 \in Q \cap K$ lenne, holott Q és K diszjunkt halmazok.

Következő észrevételünk, hogy

$$\operatorname{cl} S = -\operatorname{Ker}_+ X^T.$$

A nemtriviális irányhoz legyen $a \in -\operatorname{Ker}_+ X^T$. Ha most $a_0 \in S$, akkor $[a_0, a] \setminus \{a\} \subseteq S$, így $a \in \operatorname{cl} S$ is teljesül.

A második Hahn–Banach-tételből adódik, hogy diszjunkt konvex halmazok szeparálhatóak. Ezért ha S és K^* diszjunkt lenne, akkor létezne $x \in \mathcal{R}^d \setminus \{0\}$ vektor úgy, hogy

$$x^T S \leq 0 \leq x^T K^*$$

teljesül. Ekkor

$$x \in K^{**} = K,\ x \in -S^* = -(\operatorname{cl} S)^* = (\operatorname{Ker}_+ X^T)^* = \operatorname{Im}_+ X$$

lenne; $x = Xz$, $z \geq 0$, $z \neq 0$ mellett $x/(\mathbf{1}^T z) \in K \cap Q$ teljesülne, ami K és Q diszjunktsága miatt ellentmondás. □

A.2. Tétel: *Ha $Q \subseteq \mathcal{R}^d$ politóp, $C \subseteq \mathcal{R}^d$ pedig zárt, konvex halmaz, akkor $Q + C$ is zárt, konvex halmaz.*

Bizonyítás: Mivel Q politóp, azért léteznek $x_1, \dots, x_k \in \mathcal{R}^d$ pontok úgy, hogy Q ezek konvex burka. Egy pont és egy őt nem tartalmazó zárt, konvex halmaz erősen szeparálható, ezért minden zárt, konvex halmaz az őt tartalmazó zárt félterek metszete. Tegyük fel, hogy C-nek ez az előállítása

$$C = \cap_{i \in I} \left\{ x \in \mathcal{R}^d : a_i^T x \leq \beta_i \right\}.$$

Megmutatjuk, hogy ekkor

$$Q + C = \cap_{i \in I} \left\{ x \in \mathcal{R}^d : a_i^T x \leq \beta_i + \max_{j=1,\dots,k} \{a_i^T x_j\} \right\}.$$

A jobb oldali (zárt, konvex) halmazt jelölje S. A nemtriviális $S \subseteq Q + C$ tartalmazáshoz legyen $x \notin Q + C$, ekkor az $x - Q$ politóp és a C zárt, konvex halmaz diszjunktak, és így erősen szeparálhatók. Létezik tehát $i \in I$ index úgy, hogy

$$\beta_i = \sup a_i^T C < \min a_i^T (x - Q) = \min_{j=1,\dots,k} a_i^T (x - x_j),$$

amikor is

$$\beta_i + \max_{j=1,\dots,k} a_i^T x_j < a_i^T x$$

mutatja, hogy $x \notin S$. Ezért $S \subseteq Q + C$, és így $Q + C = S$ zárt, konvex halmaz. $\square$

Most A.2 megfordítását, A.4-et igazoljuk, ismét csak a Bolzano–Weierstrass-tétel felhasználása nélkül. Szükségünk lesz az alábbi geometriailag szemléletes, könnyen vizualizálható lemmára.

A.3. Lemma: *Legyen $Q \subseteq \mathcal{R}^d$ politóp, a $q_1, \dots, q_t$ pontok konvex burka. Legyenek továbbá $p_1, \dots, p_t$ olyan pontok, amelyekre $p_i \in Q - q_i$ $(1 \leq i \leq t)$. Ekkor léteznek nemnegatív, nem mind nulla $\lambda_1, \dots, \lambda_t$ szorzók úgy, hogy $\sum_{i=1}^t \lambda_i p_i = 0$.*

Bizonyítás: A p_i pontok

$$p_i = \sum_{j=1}^t \varepsilon_{ij} q_j - q_i \ (1 \leq i \leq t)$$

alakúak, ahol $\varepsilon_{ij} \geq 0$ $(1 \leq i, j \leq t)$, és $\sum_{j=1}^t \varepsilon_{ij} = 1$ $(1 \leq i \leq t)$. Jelölje N az ε_{ij} szorzókból alkotott mátrixot. Ekkor $N \in \mathcal{R}^{t \times t}$ nemnegatív mátrix, amelynek minden sorösszege 1, így nem létezhet $x \in \mathcal{R}^t$ vektor, amelyre $Nx > x$. (Ugyanis különben az $x =: (\xi_1, \ldots, \xi_t)^T$ jelöléssel

$$x < Nx \leq (\max_{1 \leq i \leq t} \xi_i) N\mathbf{1} = (\max_{1 \leq i \leq t} \xi_i)\mathbf{1}$$

lenne, ami ellentmondás.) Ezért nem létezik olyan $x \in \mathcal{R}^t$ vektor sem, amelyre

$$(N - E)x \geq \mathbf{1},$$

amikor is 3.32 szerint létezik $y \in \mathcal{R}^t$ vektor úgy, hogy

$$N^T y = y,\ y \geq 0,\ \mathbf{1}^T y > 0.$$

Az y vektor i-edik elemét λ_i-vel jelölve a λ_i $(1 \leq i \leq t)$ szorzók megfelelnek a kívánalmaknak: teljesülnek a

$$\begin{aligned}
\sum_{i=1}^{t} \lambda_i p_i &= \sum_{i=1}^{t}\sum_{j=1}^{t} \lambda_i \varepsilon_{ij} q_j - \sum_{i=1}^{t} \lambda_i q_i \\
&= \sum_{j=1}^{t} \left(\sum_{i=1}^{t} \lambda_i \varepsilon_{ij} - \lambda_j \right) q_j \\
&= \sum_{j=1}^{t} 0 \cdot q_j = 0
\end{aligned}$$

egyenlőségek; a lemmát bebizonyítottuk. □

A.4. Tétel: *Legyen $Q \subseteq \mathcal{R}^d$ politóp, a $q_1, \ldots, q_t$ pontok konvex burka, és legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha a $Q + C$ halmaz zárt, akkor C is zárt.*

Bizonyítás: Legyen $c_k \in C$ $(k = 1, 2, \ldots)$, és tegyük fel, hogy $c_k \to c_\infty$ $(k \to \infty)$. Megmutatjuk, hogy $c_\infty \in C$. Minden $1 \leq i \leq t$ esetén a $q_i + c_k$ $(k = 1, 2, \ldots)$ sorozat a $q_i + c_\infty$ ponthoz konvergál. Mivel a $Q + C$ halmaz zárt, azért $q_i + c_\infty \in Q + C$. Léteznek tehát $q_i' \in Q$ és $c_i' \in C$ pontok úgy, hogy $q_i + c_\infty = q_i' + c_i'$ $(1 \leq i \leq t)$. A.3 szerint léteznek nemnegatív, nem mind nulla $\lambda_1, \ldots, \lambda_t$ szorzók, amelyekkel $\sum_{i=1}^{t} \lambda_i (q_i' - q_i) = 0$. De akkor

$$\sum_{i=1}^{t} \lambda_i c_\infty = \sum_{i=1}^{t} \lambda_i c_i' + \sum_{i=1}^{t} \lambda_i (q_i' - q_i) = \sum_{i=1}^{t} \lambda_i c_i'$$

is fennáll; látjuk, hogy a c_∞ pont a C-beli c_i' pontok konvex kombinációjaként maga is C-beli. Ezzel beláttuk a C halmaz zártságát. □

Az alábbi lemmában leírt egyszerű észrevétel segítségével az A.4 tétel jelentősen általánosítható.

A.5. Lemma: *Legyen $S \subseteq \mathcal{R}^d$ tetszőleges halmaz. Ekkor*

$$S + d \cdot \operatorname{conv} S = (d+1) \cdot \operatorname{conv} S.$$

Bizonyítás: 3.21 szerint a conv S-beli elemek legfeljebb $d+1$ S-beli elem konvex kombinációjaként is előállnak. Egy ilyen előállításban legalább az egyik súly legalább $1/(d+1)$. A megfelelő S-beli elemet kivonva a $(d+1)\operatorname{conv} S$-beli elemből, $d\operatorname{conv} S$-beli elemet kapunk. Ezzel beláttuk a nemtriviális tartalmazást. □

A.6. Tétel: *Legyen $S \subseteq \mathcal{R}^d$ véges sok politóp úniója, és legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha az $S + C$ halmaz zárt, akkor C is zárt.*

Bizonyítás: Adjuk a zárt $S + C$ halmazhoz a $d\operatorname{conv} S$ politópot. Ekkor a politóp kompaktsága és az A.5 lemma miatt a zárt $C+(d+1)\operatorname{conv} S$ halmazt kapjuk. Mivel az utóbbi összegben C konvex halmaz, $(d+1)\operatorname{conv} S$ politóp, azért a már elintézett speciális eset, A.4 alkalmazható. A C konvex halmaz zártsága adódik; a bizonyítást befejeztük. □

Megjegyezzük, hogy C konvexitásának feltétele nem hagyható el a tételben:

$$C_0 := \{0\} \cup \{x \in \mathcal{R} : x > 1\} \text{ nem zárt halmaz,}$$

bár

$$\begin{aligned} S_0 &:= \{0,1\} \subseteq \mathcal{R} \text{ véges halmaz,} \\ S_0 + C_0 &= \{0\} \cup \{x \in \mathcal{R} : x \geq 1\} \text{ zárt halmaz.} \end{aligned}$$

B. Eltolás politóppal: egyéb tulajdonságok

Az előző fejezet tételei kis módosítással érvényben maradnak akkor is, ha bennük a "zártság" tulajdonságot formálisan a "relatív nyílt", "véges exponált lapú", illetve "poliéder" tulajdonságokra cseréljük.

Először relatív nyílt halmazokra bizonyítjuk A.2, A.4 és A.6 megfelelőjét.

Egy nyílt halmazhoz tetszőleges halmazt hozzáadva nyílt halmazt kapunk. Ezért nyilvánvaló, hogy

B.1. Tétel: *Ha* $Q \subseteq \mathcal{R}^d$ *politóp,* $C \subseteq \mathcal{R}^d$ *relatív nyílt, konvex halmaz, továbbá* $\operatorname{par} Q \subseteq \operatorname{par} C$ *(vagyis* Q *egy eltoltja* $\operatorname{aff} C$*-beli), akkor a* $Q + C$ *halmaz relatív nyílt.* □

Az (A.4-nek megfelelő) megfordításhoz szükségünk lesz az alábbi lemmára.

B.2. Lemma: *Legyen* $S \subseteq \mathcal{R}^d$ *kompakt halmaz, és legyen* $C \subseteq \mathcal{R}^d$ *konvex halmaz. Ha az* $S + C$ *halmaz relatív nyílt, konvex, akkor*

$$\operatorname{par} S \subseteq \operatorname{par} C = \operatorname{par}(S + C).$$

Bizonyítás: Lássuk be először, hogy $\operatorname{par} S \subseteq \operatorname{par} C$. Tegyük fel indirekt, hogy létezik $\ell \in (\operatorname{par} S) \setminus (\operatorname{par} C)$. Ekkor létezik $a \in \mathcal{R}^d$ vektor úgy, hogy $a^T \operatorname{par} C = 0 < a^T \ell$, és létezik egy $\beta \in \mathcal{R}$ konstans is úgy, hogy a C konvex halmaz a $H := \{x : a^T x = \beta\}$ hipersík része. Legyen $s^* \in S$, amelyre $a^T s^* = \max\{a^T s : s \in S\}$, az S halmaz kompaktsága miatt ilyen s^* választható. Mivel $\ell \in \operatorname{par}(S + C)$ is teljesül, azért az $S + C$ halmaz relatív nyíltsága miatt tetszőleges $c \in C$, és elég kis $\varepsilon > 0$ esetén $s^* + c + \varepsilon\ell \in S + C$. Márpedig

$$a^T(s^* + c + \varepsilon\ell) > a^T s^* + \beta = \max a^T(S + C),$$

ami ellentmondás. Ezzel beláttuk, hogy $\operatorname{par} S \subseteq \operatorname{par} C$.

Ezek után a $\operatorname{par} C = \operatorname{par}(S + C)$ egyenlőséget a nyilvánvaló

$$\begin{aligned} \operatorname{par}(S + C) &= \operatorname{aff}(S + C) - s_0 - c_0 \subseteq \operatorname{aff} S + \operatorname{aff} C - s_0 - c_0 = \\ &= \operatorname{par} S + \operatorname{par} C = \operatorname{par} C \subseteq \operatorname{par}(S + C) \end{aligned}$$

tartalmazáslánc bizonyítja. (Itt $s_0 \in S$, $c_0 \in C$ tetszőleges.) □

B.3. Tétel: *Legyen* $Q \subseteq \mathcal{R}^d$ *politóp, a* $q_1, \ldots, q_t$ *pontok konvex burka. Legyen továbbá* $C \subseteq \mathcal{R}^d$ *konvex halmaz. Ha a* $Q + C$ *halmaz relatív nyílt, akkor a* C *halmaz is relatív nyílt, és* $\operatorname{par} Q \subseteq \operatorname{par} C$*.*

Bizonyítás: A $Q + C$ halmaz relatív nyíltsága azt jelenti, hogy

$$Q + C = (\operatorname{ri} Q) + (\operatorname{ri} C).$$

Legyen $c \in C$, ekkor minden $1 \leq i \leq t$ esetén léteznek $q_i' \in \mathrm{ri}\, Q$ és $c_i' \in \mathrm{ri}\, C$ pontok úgy, hogy $q_i + c = q_i' + c_i'$. A.3 szerint választhatók nemnegatív, nem mind nulla $\lambda_1, \ldots, \lambda_t$ szorzók, amelyekkel $\sum_{i=1}^t \lambda_i(q_i' - q_i) = 0$. Mint A.4 bizonyításában, most is

$$\sum_{i=1}^t \lambda_i c = \sum_{i=1}^t \lambda_i c_i' + \sum_{i=1}^t \lambda_i (q_i' - q_i) = \sum_{i=1}^t \lambda_i c_i',$$

a c pont a $c_i' \in \mathrm{ri}\, C$ pontok konvex kombinációja, így $c \in \mathrm{ri}\, C$. Ezzel beláttuk, hogy $C \subseteq \mathrm{ri}\, C$, azaz a C halmaz relatív nyíltságát, a $\mathrm{par}\, Q \subseteq \mathrm{par}\, C$ tartalmazás pedig B.2 azonnali következménye. □

Most már A.6 mintájára könnyen belátható B.3 alábbi általánosítása.

B.4. Tétel: *Legyen $S \subseteq \mathcal{R}^d$ véges sok politóp úniója, és legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha az $S + C$ halmaz relatív nyílt, konvex [nyílt], akkor a C halmaz is relatív nyílt [nyílt], és* $\mathrm{par}\, S \subseteq \mathrm{par}\, C$. □

A "véges exponált lapú" tulajdonság esetében A.2 analogonját az alábbi tétel írja le.

B.5. Tétel: *Legyen $Q \subseteq \mathcal{R}^d$ politóp, a $q_1, \ldots, q_t$ pontok konvex burka. Legyen továbbá $C \subseteq \mathcal{R}^d$ konvex halmaz, amelynek csak véges sok exponált lapja van. Ekkor a $Q + C$ halmaznak is csak véges sok exponált lapja van.*

Bizonyítás: Mivel

$$Q + C = \mathrm{conv} \cup_{i=1}^t (q_i + C),$$

azért tetszőleges $a \in \mathcal{R}^d$ vektor esetén

$$\begin{aligned} \inf a^T(Q + C) &= \inf \cup_{i=1}^t a^T(q_i + C) = \\ &= \min_{1 \leq i \leq t} \inf a^T(q_i + C) = \min a^T\{q_1, \ldots, q_t\} + \inf a^T C. \end{aligned}$$

Innen már látszik, hogy ha $\hat{F}$ a $Q + C$ halmaz egy $a \in \mathcal{R}^d$ vektor által exponált lapja, akkor az a vektor a C [Q] halmaz egy F [G] lapját exponálja, amelyre $\hat{F} = F + G$. Utóbbi típusú összegből csak véges sok féle lehet (Q és C exponált lapjai is véges sokan vannak), ezért az $\hat{F}$ exponált lap is csak véges sok féle lehet. □

Megjegyezzük, hogy politóp + véges exponált lapú, konvex halmaz lapjainak száma lehet végtelen. Legyen például a politóp a csak az origót tartalmazó halmaz, a véges exponált lapú, konvex halmaz pedig $\mathcal{R}^3$-ben egy nyílt féltér únió a határának egy zárt körlap része. Zárt példa nincs, mert zárt, véges lapú = poliéder = zárt, véges exponált lapú 7.12 szerint.

B.5 megfordítása, A.4 analogonja a

B.6. Tétel: *Legyen $Q \subseteq \mathcal{R}^d$ politóp, és legyen $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha a $Q + C$ konvex halmaznak csak véges sok exponált lapja van, akkor a C konvex halmaznak is csak véges sok exponált lapja van.*

Bizonyítás: Jelölje $\mathcal{S}$ azoknak az (F, G) pároknak a halmazát, amelyekben az $F \in \mathcal{EF}(C)$, $G \in \mathcal{EF}(Q)$ lapokat ugyanaz az $a \in \mathcal{R}^d$ vektor exponálja. Könnyen belátható, hogy ha $(F, G) \in \mathcal{S}$, akkor $F+G \in \mathcal{EF}(Q+C)$, továbbá

$$F = C \cap (F + G - Q),\ G = Q \cap (F + G - C).$$

Ezért az $(F, G) \mapsto F + G$ leképezés injektív az $\mathcal{S}$ halmazon, és $\mathcal{S}$-et a véges $\mathcal{EF}(Q+C)$ halmazba képezi. Következésképpen az $\mathcal{S}$ halmaz elemszáma is véges. A C halmaz minden F exponált lapjához létezik a Q halmaz egy G exponált lapja úgy, hogy $(F, G) \in \mathcal{S}$. Innen látszik, hogy $\mathcal{EF}(C)$ elemszáma sem lehet végtelen, amit igazolnunk kellett. $\square$

Itt is megfogalmazható B.6 általánosítása, amely azonban összetettebb, mint az eddigi esetekben (lásd A.6, B.4). Szükségünk lesz az alábbi lemmára.

B.7. Lemma: *Legyenek $C_1, \dots, C_l \subseteq \mathcal{R}^d$ véges exponált lapú, konvex halmazok. Ha az úniójuk konvex, akkor véges számú exponált lapja van.*

Bizonyítás: Legyen F az $\cup C_j$ konvex halmaz exponált lapja, és jelölje F_j az $F \cap C_j$ halmazt $(1 \leq j \leq l)$. Könnyen belátható, hogy az $F \mapsto (F_1, \dots, F_l)$ leképezés injekció az $\mathcal{EF}(\cup C_j)$ halmazról a véges $\mathcal{EF}(C_1) \times \dots \times \mathcal{EF}(C_l)$ halmazba. Szükségképpen az $\mathcal{EF}(\cup C_j)$ halmaz véges. $\square$

B.8. Tétel: *Legyenek $Q_1, \dots, Q_k \subseteq \mathcal{R}^d$ politópok, és jelölje S az úniójukat. Legyen továbbá $C \subseteq \mathcal{R}^d$ konvex halmaz. Ha az $S+C$ halmaz a $C_1, \dots, C_l \subseteq \mathcal{R}^d$ véges exponált lapú, konvex halmazok úniója, akkor C-nek csak véges sok exponált lapja van.*

Bizonyítás: Adjuk az $S + C = \cup C_j$ halmazhoz a $d\operatorname{conv} S$ politópot, ekkor A.5 szerint a konvex

$$C + (d + 1) \cdot \operatorname{conv} S = \cup(C_j + d \cdot \operatorname{conv} S)$$

halmazt kapjuk. B.5 szerint a $C_j+d\operatorname{conv} S$ halmazoknak véges sok exponált lapja van. Ezért B.7 miatt konvex úniójuk, a $C + (d + 1)\operatorname{conv} S$ halmaz maga is véges exponált lapú. A már igazolt speciális eset, B.6 alkalmazható; adódik a C halmaz véges exponált lapúsága is. $\square$

7.12 szerint egy konvex halmaz pontosan akkor poliéder, ha zárt és véges sok exponált lapja van. Ezért (figyelembe véve az A.2, A.4 és A.6 tételeket) látható, hogy a B.5, B.6, B.7 és B.8 tételek érvényben maradnak akkor is, ha bennük a "véges exponált lapú" tulajdonság helyét a "poliéder" tulajdonság veszi át. Itt egyszerűsége miatt csak a B.7-nek megfelelő állítást mondjuk ki.

B.9. Állítás: *Ha véges sok poliéder úniója konvex, akkor poliéder.* □

B.9 annak a nyilvánvaló ténynek a nemtriviális párja, hogy véges sok poliéder metszete is poliéder.

C. Zártsági feltételek, erős szeparáció

Ebben a fejezetben Abrams tételének változatait, valamint erős szeparációs tételeket bizonyítunk.

Abrams tétele (4.29) szerint az $AS_1 + S_2$ halmaz zártsága esetén $S_1 + A^{-1}(S_2)$ zárt, továbbá megfordítva $S_2 \subseteq \operatorname{Im} A$ esetén az $S_1 + A^{-1}(S_2)$ halmaz zártságából következik $AS_1 + S_2$ zártsága. (Itt $A \in \mathcal{R}^{m\times n}$ mátrix, $S_1 \subseteq \mathcal{R}^n$, $S_2 \subseteq \mathcal{R}^m$ tetszőleges halmazok.)

Nincs szükség az $S_2 \subseteq \operatorname{Im} A$ feltételre, ha $S_2 = L$ altér:

C.1. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá legyen $S \subseteq \mathcal{R}^n$ tetszőleges halmaz, és legyen $L \subseteq \mathcal{R}^m$ altér. Ekkor az $AS + L$ halmaz pontosan akkor zárt, ha az $S + A^{-1}(L)$ halmaz zárt.*

Bizonyítás: Válasszunk egy B mátrixot, amelyre $L = B^{-1}(0)$. Abrams tétele szerint az $AS + L$ halmaz pontosan akkor zárt, ha a BAS halmaz zárt. Másfelől $A^{-1}(L) = (BA)^{-1}(0)$, így ismét csak Abrams tétele miatt az $S + A^{-1}(L)$ halmaz is pontosan akkor zárt, ha a BAS halmaz zárt. □

A C.1 tétel érvényét veszti, ha S_2-ről csak annyit teszünk fel, hogy politóp, mint azt az alábbi példa mutatja: Legyen $S_2 \subseteq \mathcal{R}^2$ egyenlő oldalú háromszög, $S_1 \subseteq \mathcal{R}^2$ az alapja, megfosztva felezőpontjától, az origótól, $A \in \mathcal{R}^{2\times 2}$ pedig a vetítés erre az alapra. Ekkor S_1 nem konvex halmaz, S_2 politóp, és könnyen ellenőrizhető, hogy bár $S_1 + A^{-1}(S_2)$ zárt, az $AS_1 + S_2$ halmaz nem az. (A homogenizált változat olyan ellenpélda, amelyben S_2 poliéder kúp.)

Az előző példában S_1 nem konvex halmaz: konvex $S_1 = C$ halmaz esetén ugyanis a C.1 tétel az $S_2 = L$ altér helyett tetszőleges $S_2 = P$ poliéderrel is fennáll.

C.2. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá legyen $C \subseteq \mathcal{R}^n$ konvex halmaz, és legyen $P \subseteq \mathcal{R}^m$ poliéder. E feltételek mellett az $AC + P$ halmaz pontosan akkor zárt, ha a $C + A^{-1}(P)$ halmaz zárt.*

Bizonyítás: Elég azt megmutatnunk, hogy a $C+A^{-1}(P)$ halmaz zártságából következik az $AC + P$ halmaz zártsága. (A másik irányt már igazoltuk Abrams tételének bizonyításakor.)

Először tegyük fel, hogy $P = R$ poliéder kúp. Mivel

$$A^{-1}(R) = A^{-1}(R \cap A\mathcal{R}^n),$$

azért Abrams tétele szerint a $C + A^{-1}(R)$ halmaz zártságából következik az $AC + (R \cap A\mathcal{R}^n)$ halmaz zártsága. Nyilvánvalóan

$$(-R) \cap \operatorname{rec}(AC + R \cap A\mathcal{R}^n) \subseteq -(R \cap A\mathcal{R}^n) \subseteq -\operatorname{rec}(AC + R \cap A\mathcal{R}^n).$$

Ezért 4.23 szerint az $AC + (R \cap A\mathcal{R}^n)$ és R halmazok összege, az $AC + R$ halmaz zárt. Ezzel a tétel $P = R$ speciális esetét beláttuk.

Az általános esetben tegyük fel, hogy a $C + A^{-1}(P)$ halmaz zárt. Az $A^{-1}(P)$ halmaz poliéder (3.25), melynek recessziós kúpja $A^{-1}(\operatorname{rec} P)$ (4.28). Motzkin tétele és 4.13 szerint az $A^{-1}(P)$ poliéder valamely politóp és az $A^{-1}(\operatorname{rec} P)$ poliéder kúp összege. Ezért az A.4 tételből adódóan a $C + A^{-1}(P)$ halmaz zártságából következik a $C + A^{-1}(\operatorname{rec} P)$ halmaz zártsága. A tétel első felében belátottak szerint a $C + A^{-1}(\operatorname{rec} P)$ halmaz csak úgy lehet zárt, ha az $AC + (\operatorname{rec} P)$ halmaz zárt. Ismét Motzkin tétele és 4.13 miatt a P poliéder egy politóp és a $\operatorname{rec} P$ poliéder kúp összege. Tehát A.2 (vagy 4.12) szerint az $AC+(\operatorname{rec} P)$ halmaz zártságából következik az $AC+P$ halmaz zártsága. Ily módon a tételt az általános esetben is igazoltuk. □

Ahogyan azt 4.23 bizonyításakor láttuk, az $AC+P$ halmaz zártságának igazolása után további (az 5. és 6.) lépésekre volt szükség a recessziós kúpokról szóló részállítás igazolásához. A C.2 tétel esetében is lehetséges hasonló kiterjesztés, C.4.

A C.4 tétel bizonyítása, akárcsak a fent említett bizonyítás, két lépésben történik. A $P = R$ poliéder kúp speciális esetet tárgyalja az alábbi C.3 lemma. (Emlékeztetünk rá, hogy 4.20 szerint bármely $C \subseteq \mathcal{R}^d$ konvex halmaz esetén a $\operatorname{cl} K(C)$ konvex kúp és az $\{1\} \times \mathcal{R}^d$ hipersík metszete $\{1\} \times (\operatorname{cl} C)$, míg a $\operatorname{cl} K(C)$ konvex kúp és a $\{0\} \times \mathcal{R}^d$ hipersík metszete $\{0\} \times \operatorname{rec}(\operatorname{cl} C)$.)

C.3. Lemma: *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá legyen $C \subseteq \mathcal{R}^n$ zárt, konvex halmaz, és legyen $R \subseteq \mathcal{R}^m$ poliéder kúp. Jelölje $\hat{K}_1$ és $\hat{K}_2$ a következő $\mathcal{R}^{n+1}$-beli, illetve $\mathcal{R}^{m+1}$-beli konvex kúpokat:*

$$\begin{aligned}\hat{K}_1 &:= (\operatorname{cl} K(C)) + \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ R \end{pmatrix};\\ \hat{K}_2 &:= \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \operatorname{cl} K(C) + \begin{pmatrix} 0 \\ R \end{pmatrix}.\end{aligned}$$

Ezekkel a jelölésekkel a következő állítások ekvivalensek:

a) $\hat{K}_2$ *zárt;*

b) $\hat{K}_2 = \operatorname{cl} K(AC + R)$*;*

c) $AC + R$ *zárt, és* $\operatorname{rec}(AC + R) = (A\operatorname{rec} C) + R$*;*

d) $\hat{K}_1$ *zárt;*

e) $\hat{K}_1 = \operatorname{cl} K(C + A^{-1}(R))$*;*

f) $C + A^{-1}(R)$ *zárt, és* $\operatorname{rec}(C + A^{-1}(R)) = (\operatorname{rec} C) + A^{-1}(R)$.

Bizonyítás: a)⇔b): A 4.23 tétel bizonyításának 5. lépésében beláttuk, hogy a $\hat{K}_2$ lezártja $\operatorname{cl} K(AC + R)$.

b)⇔c): A $\hat{K}_2$ és $\operatorname{cl} K(AC+R)$ konvex kúpok pontosan akkor egyenlőek, ha metszeteik az $\{1\} \times \mathcal{R}^m$ és $\{0\} \times \mathcal{R}^m$ hipersíkokkal megegyeznek.

d)⇔e) és e)⇔f) hasonlóan bizonyítható.

Végül a)⇔d) a C.2 tétel következménye. □

Most már a szokásos módon (4.23 bizonyítása 4. és 6. lépésének mintájára) C.2, C.3, A.2 és megfordítása, A.4 segítségével, könnyen belátható a

C.4. Tétel: *Legyen $A \in \mathcal{R}^{m\times n}$, továbbá legyen $C \subseteq \mathcal{R}^n$ zárt, konvex halmaz, és legyen $P \subseteq \mathcal{R}^m$ poliéder. Ekkor a következő állítások ekvivalensek:*

a) $AC + P$ *zárt, és* $\operatorname{rec}(AC + P) = (A\operatorname{rec} C) + \operatorname{rec} P$*;*

b) $C + A^{-1}(P)$ *zárt, és* $\operatorname{rec}(C + A^{-1}(P)) = (\operatorname{rec} C) + A^{-1}(\operatorname{rec} P)$. □

Megjegyezzük, hogy C.2 [C.4] érvényét veszti a zárt, konvex halmazok körében: léteznek C_1, C_2 zárt, konvex halmazok [kúpok], amelyekre $C_1 + A^{-1}(C_2)$ halmaz zárt, ám az $AC_1 + C_2$ halmaz nem zárt valamely A lineáris leképezéssel.

Legyen PSD_2 a 3.31-beli K zárt, konvex kúp képe az

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 & x_2/\sqrt{2} \\ x_2/\sqrt{2} & x_3 \end{pmatrix}, \ \mathcal{R}^3 \to \mathcal{R}^{2\times 2}$$

lineáris leképezésnél, vagyis legyen

$$PSD_2 := \left\{ \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} : \alpha, \gamma \geq 0, \ \alpha\gamma - \beta^2 \geq 0 \right\}.$$

Ekkor PSD_2 (a 2×2-es, szimmetrikus, pozitív szemidefinit mátrixok halmaza) zárt, konvex kúp. (Bár PSD_2 2×2-es mátrixokból áll, tekinthetünk rá úgy is, mintha elemei 4×1-es vektorok lennének, és csak a tömörség kedvéért használjuk a mátrixos formát.)

Legyen most

$$A : \lambda \mapsto \lambda \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \ (\lambda \in \mathcal{R}), \ C_1 := \mathcal{R}, \ C_2 := PSD_2.$$

Ekkor egyrészt

$$\begin{pmatrix} 1 + k + 1/k & 1/2 + k \\ 1/2 + k & k \end{pmatrix} - k \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \ (k \to \infty)$$

mutatja, hogy

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \in \mathrm{cl}\,(AC_1 + C_2).$$

Másfelől

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} = \begin{pmatrix} 1 - \lambda & 1/2 - \lambda \\ 1/2 - \lambda & -\lambda \end{pmatrix}, \ \lambda \in \mathcal{R}$$

esetén $\alpha\gamma - \beta^2 = -1/4 \ngeq 0$, ezért

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \notin AC_1 + C_2.$$

Látjuk, hogy az $AC_1 + C_2$ halmaz nem zárt, bár a $C_1 + A^{-1}(C_2)$ halmaz zárt.

4.23, 6.24 és C.4 összeolvasásával azonnal adódik

C.5. Tétel: *Legyen* $A \in \mathcal{R}^{m\times n}$*, továbbá legyen* $C \subseteq \mathcal{R}^n$ *zárt, konvex halmaz, és legyen* $P \subseteq \mathcal{R}^m$ *poliéder. Ekkor az*

a) $(A^T \operatorname{bar} P) \cap \operatorname{ri}(\operatorname{bar} C) \neq \emptyset$*,*

b) $C + A^{-1}(P)$ *zárt, és* $\operatorname{rec}(C + A^{-1}(P)) = (\operatorname{rec} C) + A^{-1}(\operatorname{rec} P)$*,*

c) $A^{-1}(-\operatorname{rec} P) \cap \operatorname{rec} C \subseteq -\operatorname{rec} C$*,*

d) $AC + P$ *zárt, és* $\operatorname{rec}(AC + P) = (A\operatorname{rec} C) + \operatorname{rec} P$

állítások a következő logikai relációban vannak egymással: a)⇒b), c)⇒d), a)⇔c) és b)⇔d). □

Az 5. fejezetben láttuk, hogy zártsági tételek alkalmazásaként hogyan nyerhetők erős szeparációs tételek (5.10-13). Hasonlóképpen az alábbi erős szeparációs tétel 6.27 és C.2 azonnali következménye.

C.6. Tétel: *Legyen* $A \in \mathcal{R}^{m\times n}$*, továbbá legyen* $C \subseteq \mathcal{R}^n$ *konvex halmaz, és legyen* $P \subseteq \mathcal{R}^m$ *poliéder. Ekkor az*

a) $0 \notin AC + P$ *(vagyis* AC *és* $-P$ *diszjunktak),*

b) $0 \notin C + A^{-1}(P)$ *(vagyis* $-C$ *és* $A^{-1}(P)$ *diszjunktak),*

c) $0 \notin \operatorname{cl}(AC + P)$ *(vagyis* AC *és* $-P$ *erősen szeparálhatók),*

d) $0 \notin \operatorname{cl}(C + A^{-1}(P))$ *(vagyis* $-C$ *és* $A^{-1}(P)$ *erősen szeparálhatók)*

állítások a következő logikai relációban vannak egymással: a)⇔b); a)⇔c), ha az $AC + P$ *halmaz zárt; b)⇔d), ha a* $C + A^{-1}(P)$ *halmaz zárt.*

Speciálisan mind a négy állítás ekvivalens, ha a 6.27 zártsági tételbeli a), b), c) vagy d) állítás fennáll. □

A C.6-beli c) és d) állítások általában is ekvivalensek:

C.7. Tétel: *Legyen* $A \in \mathcal{R}^{m\times n}$*, és legyenek* $C_1 \subseteq \mathcal{R}^n$*,* $C_2 \subseteq \mathcal{R}^m$ *konvex halmazok. Ekkor*

a) $0 \notin \operatorname{cl}(AC_1 + C_2)$ *esetén* $0 \notin \operatorname{cl}(C_1 + A^{-1}(C_2))$*. (Más szavakkal ha* AC_1 *és* $-C_2$ *erősen szeparálhatóak, akkor* $-C_1$ *és* $A^{-1}(C_2)$ *is erősen szeparálhatóak.)*

b) Az a) pontban megfogalmazott állítás megfordítható, ha a C_2 *halmaz része az* A *mátrix képterének.*

c) Az a) pontban megfogalmazott állítás megfordítható, ha a C_2 *halmaz poliéder.*

Bizonyítás: a) A bizonyítás indirekt: megmutatjuk, hogy $0 \in \mathrm{cl}\,(C_1 + A^{-1}(C_2))$ esetén $0 \in \mathrm{cl}\,(AC_1 + C_2)$. Legyen $x_k \in C_1$, $v_k \in A^{-1}(C_2)$ $(k = 1, 2, \ldots)$ olyan pontsorozat, amelyre $x_k + v_k \to 0$ $(k \to \infty)$. Ekkor az A leképezés folytonossága miatt $A(x_k + v_k) \to 0$ $(k \to \infty)$ szintén fennáll. Mivel az Av_k vektorok definíció szerint a C_2 halmaz elemei, látjuk, hogy $0 \in \mathrm{cl}\,(AC_1 + C_2)$.

b) Tegyük fel most, hogy a C_2 halmaz része az A mátrix képterének. Megmutatjuk, hogy ha $0 \notin \mathrm{cl}\,(C_1 + A^{-1}(C_2))$, akkor $0 \notin \mathrm{cl}\,(AC_1 + C_2)$. Az origó és a $C_1 + A^{-1}(C_2)$ konvex halmaz erősen szeparálhatóak, létezik tehát egy $a \in \mathcal{R}^d$ vektor úgy, hogy

$$0 < \inf \left\{ a^T x : x \in C_1 + A^{-1}(C_2) \right\}.$$

Az $A^{-1}(C_2)$ halmaz recessziós kúpja tartalmazza az A mátrix nullterét, így az előző egyenlőtlenségből az is adódik, hogy az a vektor eleme az A^T mátrix képterének: létezik $z \in \mathcal{R}^m$ vektor úgy, hogy $a = A^T z$.

Tegyük fel indirekt, hogy $0 \in \mathrm{cl}\,(AC_1 + C_2)$. Ekkor léteznek $x_k \in C_1$, $y_k \in C_2$ $(k = 1, 2, \ldots)$ pontsorozatok úgy, hogy

$$Ax_k + y_k \to 0 \ (k \to \infty).$$

Mivel a C_2 halmaz A képterének része feltevésünk szerint, azért valamely $v_k \in \mathcal{R}^n$ vektorokkal (valójában $v_k \in A^{-1}(C_2)$) $y_k = Av_k$ $(k = 1, 2, \ldots)$ fennáll. De akkor

$$a^T(x_k + v_k) = z^T(Ax_k + y_k) \to 0 \ (k \to \infty),$$

ami $x_k + v_k \in C_1 + A^{-1}(C_2)$ miatt ellentmondás.

c) Tegyük fel, hogy C_2 poliéder. Megmutatjuk, hogy ekkor $-C_1$ és $A^{-1}(C_2)$ erős szeparálhatósága esetén az AC_1 és $-C_2$ halmazok is erősen szeparálhatók. Vegyük észre, hogy

$$A^{-1}(C_2) = A^{-1}(C_2 \cap A\mathcal{R}^n).$$

Itt a $C_2 \cap A\mathcal{R}^n$ halmaz már az A mátrix képterének része, így a b) részállítás miatt $-C_1$ és $A^{-1}(C_2)$ erős szeparálhatóságából az AC_1 és a $-C_2 \cap A\mathcal{R}^n$ halmazok erős szeparálhatósága adódik. Az AC_1 halmaz tehát belefoglalható egy zárt féltérbe úgy, hogy annak az A mátrix képterével való metszete (poliéder!) diszjunkt a $-C_2$ poliédertől. E diszjunkt poliéderek erősen szeparálhatók, amiből már a kívánt állítás, AC_1 és $-C_2$ erős szeparálhatósága, is adódik. $\square$

Végül megjegyezzük, hogy a C.7 tételbeli a) állítás általában nem megfordítható, még akkor sem, ha a C_1, C_2 halmazokról feltesszük, hogy zártak, konvexek: léteznek C_1, C_2 zárt, konvex halmazok, amelyekkel

$$0 \in \mathrm{cl}\,(AC_1 + C_2),\; 0 \notin \mathrm{cl}\,(C_1 + A^{-1}(C_2))$$

valamely A lineáris leképezés esetén.

Valóban, legyen

$$A : (\lambda, \mu) \mapsto \lambda \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \mu \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \;\; (\lambda, \mu \in \mathcal{R});$$
$$C_1 := \mathcal{R} \times \{0\} \subseteq \mathcal{R}^2;\; C_2 := PSD_2 - \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix}.$$

Ekkor, mint azt a C.4 után leírt példában már igazoltuk,

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \in \mathrm{cl}\,(AC_1 + PSD_2).$$

Ezért $0 \in \mathrm{cl}\,(AC_1 + C_2)$.

Másfelől könnyen belátható, hogy

$$A^{-1}(C_2) = \{(\lambda, \mu) : \lambda \geq -1/2,\; \mu = -1/2\}\,,$$

így valóban $0 \notin \mathrm{cl}\,(C_1 + A^{-1}(C_2))$, a C_1, C_2 halmazok megfelelnek a kívánalmaknak.

Irodalomjegyzék

1. M. S. Bazaraa és C. M. Shetty: *Nonlinear programming, theory and algorithms.* John Wiley & Sons, New York, 1979.

2. A. Berman: *Cones, matrices and mathematical programming.* Springer-Verlag, Berlin, 1973.

3. J. M. Borwein és A. S. Lewis: *Convex analysis and nonlinear optimization, theory and examples.* Springer-Verlag, New York, 2000.

4. J. M. Borwein és H. Wolkowicz, Characterization of optimality for the abstract convex program with finite dimensional range, *J. Austral. Math. Soc.* (Series A) 30 (1981), 390-411.

5. J. M. Borwein és H. Wolkowicz, Facial reduction for a cone-convex programming problem, *J. Austral. Math. Soc.* (Series A) 30 (1981), 369-380.

6. R. I. Boţ, S. M. Grad és G. Wanka, A new constraint qualification and conjugate duality for composed convex optimization problems, *J. Opt. Th. Appl.* 135 (2007), 241-255.

7. Császár Ákos: *Valós analízis I-II.* Tankönyvkiadó, Budapest, 1989.

8. J. B. G. Frenk: *Convexity and optimization* (egyetemi jegyzet). Erasmus Universiteit Rotterdam, Rotterdam, 1993.

9. J. B. G. Frenk és G. Kassay, On classes of generalized convex functions, Gordan–Farkas type theorems, and Lagrange duality, *J. Opt. Th. Appl.* Vol. 102, No. 2 (1999), 315-343.

10. Fried Ervin: *Klasszikus és lineáris algebra.* Tankönyvkiadó, Budapest, 1991.

11. V. Jeyakumar, Constraint qualifications characterizing Lagrangian duality in convex optimization, *J. Opt. Th. Appl.* Vol. 136 (2008), 31-41.

12. P. R. Halmos: *Véges dimenziós vektorterek.* Műszaki Könyvkiadó, Budapest, 1984.

13. Kovács Margit: *A nemlineáris programozás elmélete.* TYPOTEX Kft., Budapest, 1997.

14. Kósa András: *Optimumszámítási modellek.* Műszaki Könyvkiadó, Budapest, 1979.

15. Kristóf János: *Az analízis elemei I-IV* (egyetemi jegyzet). ELTE Kiadó, Budapest, 1994-1998.

16. A. G. Kuros: *Felsőbb algebra.* Tankönyvkiadó, Budapest, 1968.

17. G. Pataki, On the closedness of the linear image of a closed convex cone, *Math. Oper. Res.* Vol. 32, No. 2 (2007), 395-412.

18. Prékopa András: *Lineáris programozás I.* Bolyai János Matematikai Társulat, Budapest, 1968.

19. Rapcsák Tamás: *Nemlineáris optimalizálás.* Operációkutatás 8, Aula Kiadó, Budapest, 2006.

20. R. T. Rockafellar: *Convex analysis.* Princeton University Press, Princeton, New Jersey, 1970.

21. C. Roos és T. Terlaky: *Nonlinear optimization (WI387)* (egyetemi jegyzet). TU Delft, Delft, 1998. (Magyarul: E. de Klerk, C. Roos és T. Terlaky: *Nemlineáris optimalizálás.* Operációkutatás 5, Aula Kiadó, Budapest, 2004.)

22. Rózsa Pál: *Lineáris algebra és alkalmazásai.* Tankönyvkiadó, Budapest, 1991.

23. A. Schrijver: *Theory of linear and integer programming.* John Wiley & Sons, New York, 1986.

24. J. Stoer és C. Witzgall: *Convexity and optimization in finite dimensions I.* Springer-Verlag, Berlin, 1970.

25. G. Strang: *Linear algebra and its applications.* Academic Press, New York, 1980.

26. Szabó László: *Konvex geometria* (egyetemi jegyzet). ELTE Kiadó, Budapest, 1996.

27. Ujvári Miklós: *A szemidefinit programozás alkalmazásai a kombinatorikus optimalizálásban* (egyetemi jegyzet). ELTE Kiadó, Budapest, 2001.

28. M. Ujvári, On a closedness theorem, *Pure Mathematics and Applications* Vol. 15, No. 4 (2006), 469-486.

29. M. Ujvári, On Abrams' theorem, *Pure Mathematics and Applications* Vol. 18, No. 1-2 (2008), 177-187.

30. M. Ujvári, On closedness conditions, strong separation, and convex duality, benyújtva a *Pure Mathematics and Applications* folyóirathoz (2008).

    URL: http://www.uni-corvinus.hu/puma

31. Z. Waksman és M. Epelman, On point classification in convex sets, *Math. Scand.* 38 (1976), 83-96.

32. H. Wolkowicz, Some applications of optimization in matrix theory, *Linear Algebra and its Applications* 40 (1981), 101-118.

Tárgymutató

Mellékelem a 2002/03 tanév II. félévében tartott előadásom tematikáját, ami segíthet a jegyzet feldolgozásában.

1. $AC + P$ zártságának elégséges feltétele; Abrams-tétel.
2. Hahn–Banach-tételek (első, második, vegyes); sima pontok.
3. Stiemke és Krein tétele, többkúposításuk; Carver tétele, Dubovickij–Miljutyin-tétel.
4. Kúplineáris Farkas-lemma; összefoglaló zártsági tétel; kúplineáris programok korlátossága.
5. Kúplineáris Farkas-tétel; kúplineáris gyenge/erős dualitás; szigorú megoldhatóság; regularizáció.
6. Minkowski–Klee-tétel; Straszewicz–Klee-tétel; poliéderek új jellemzése.
7. Epigráf/nívóhalmazok lezártja/relatív belseje; konvex függvények folytonossága.
8. Fenchel-konjugált; kúpfüggvények támaszfüggvényként ($\operatorname{rec} f$, $\operatorname{cl} f'(x; .)$).
9. Differenciálható függvények konvexitása (első- és másodrendű jellemzések); lokális optimum.
10. Rockafellar-dualitás (gyenge/erős dualitási tétel, korlátosság, szigorú megoldhatóság).
11. Lagrange-dualitás; Karush–Kuhn–Tucker-tétel.
12. Konvex Farkas-tétel; Wolfe-dualitás; kvadratikus dualitás; Goldman–Tucker-tétel.