Jegyzet

az Operációkutatás II című tantárgyhoz

Király Tamás előadásai alapján készítette Papp Olga

Utolsó frissítés: 2011. május 20.

Tartalomjegyzék

1. TU mátrixok: kerekítés és színezés

1.1. Emlékeztető

Egy mátrix **teljesen unimoduláris (TU)**, ha minden aldeterminánsa 0, 1, vagy -1 értékű. Néhány egyszerű tulajdonság:

- TU mátrix transzponáltja is TU,
- Ha A TU, akkor $[A, -A]$ és $[A, I]$ is TU,
- Páros gráf incidencia-mátrixa és irányított gráf incidencia-mátrixa TU.

1.1. Tétel. *Ha $A \in \mathbb{R}^{m\times n}$ TU mátrix és $b \in \mathbb{Z}^m$, akkor a $\max\{cx : Ax \le b\}$ feladatnak, ha megoldható, van egész optimális megoldása, hiszen minden erős bázismegoldás egészértékű.*

1.2. Tétel (Farkas Lemma TU mátrixokra)**.** *Ha $A \in \mathbb{R}^{m\times n}$ TU mátrix és $b \in \mathbb{Z}^m$, akkor a következő két állítás közül pontosan az egyik igaz:*

1. *Az $Ax \le b$ egyenlőtlenség-rendszernek van egész megoldása.*
2. *Az $yA = 0$, $y \ge 0$, $yb < 0$ rendszernek van $y \in \{0,1\}^m$-beli megoldása.*

Legyen $D = (V, E)$ egy irányított gráf, és A az incidencia-mátrixa. Tudjuk, hogy A TU mátrix. Legyen $b \in \mathbb{Z}^V$ tetszőleges. Az $Ax = b, x \ge 0$ rendszer másképp úgy írható, hogy $\varrho_x(v) - \delta_x(v) = b_v$ minden $v \in V$-re, és $x \ge 0$. A TU Farkas Lemma erre a feladatra a következőt adja:

1.3. Tétel. *Legyen $D = (V, E)$ egy irányított gráf, és $b \in \mathbb{Z}^V$ tetszőleges. Pontosan akkor van olyan x nemnegatív egész vektor az éleken, amire $\varrho_x(v) - \delta_x(v) = b_v$ minden $v \in V$-re, ha nincs olyan $U \subseteq V$ (esetleg üres) halmaz, amire $\varrho_D(U) = 0$ és $\sum_{u\in U} b_u > \sum_{v\notin U} b_v$.*

1.2. Kerekítés és egyenletes színezés

1.4. Tétel. *Legyen $A \in \mathbb{R}^{m\times n}$ TU mátrix és $z \in \mathbb{R}^n$. Ekkor z kerekíthető úgy (minden komponensét alsó vagy felső egészrészre változtatva), hogy Az minden komponense kevesebb mint 1-gyel változik.*

Bizonyítás. Tekintsük a következő egyenlőtlenség-rendszert:

$$\lfloor z \rfloor \le x \le \lceil z \rceil$$
$$\lfloor Az \rfloor \le Ax \le \lceil Az \rceil$$

Ennek az egyenlőtlenség-rendszernek a mátrixa TU. Tudjuk, hogy van megoldása (maga z), tehát van egész megoldása is, ami pont a tételnek megfelelő kerekítést ad. □

Ennek a tételnek például következménye, hogy tetszőleges $M \in \mathbb{R}^{k\times l}$ valós mátrix elemeit lehet úgy kerekíteni, hogy minden sorösszeg és oszlopösszeg kevesebb mint 1-gyel változzon. Legyen ugyanis $A \in \mathbb{Z}^n$ a $K_{k,l}$ teljes páros gráf incidencia-mátrixa (ahol $n = kl$), és legyen $z \in \mathbb{R}^n$ az M mátrix elemei által alkotott vektor (a mátrix elemeit természetes módon megfeleltethetjük a gráf éleinek). Mivel A TU mátrix, alkalmazhatjuk a fenti tételt, ami pont olyan kerekítést ad, ahol a sor- és oszlopösszegek kevesebb mint 1-gyel változnak.

1.5. Tétel. *Legyen $A \in \mathbb{R}^{m \times n}$ TU mátrix, $b \in \mathbb{Z}^m$, és k pozitív egész szám. Legyen $z \in \mathbb{Z}^n$ olyan egész vektor, amire $Az \le kb$. Ekkor z előáll $z = z^1 + z^2 + \dots + z^k$ alakban, ahol $z^i \in \mathbb{Z}^n$ és $Az^i \le b$ $(i = 1, \dots, k)$.*

Bizonyítás. A tételt k szerinti indukcióval bizonyítjuk. $k = 1$-re az állítás triviális. Tehát tegyük fel, hogy a tétel $(k-1)$-re igaz, és legyen $z \in \mathbb{Z}^n$ olyan egész vektor, amire $Az \le kb$. Az $\{x \in \mathbb{R}^n : \ Az - (k-1)b \le Ax \le b\}$ poliéder nemüres, hiszen z/k benne van. Mivel az egyenlőtlenség-rendszer mátrixa TU, létezik egész megoldás is; legyen ez z^k. Ekkor $Az^k \le b$, és $A(z - z^k) \le (k-1)b$, tehát a $z - z^k$ vektorra és $(k-1)$-re használva az indukciót kész vagyunk. □

A tétel segítségével megmutatjuk, hogy egy TU mátrix oszlopainak mindig van egyenletes színezése.

1.6. Tétel. *Legyen $A \in \mathbb{R}^{m \times n}$ TU mátrix, és k pozitív egész szám. Ekkor A oszlopait k csoportra lehet osztani úgy, hogy minden sorban az egyes csoportokban az elemek összege legfeljebb 1-gyel tér el egymástól.*

Bizonyítás. Legyen $d \in \mathbb{Z}^m$ az A oszlopainak összege, és legyen $b_1 = \lfloor d/k \rfloor$, $b_2 = \lceil d/k \rceil$. Ekkor az $\{x \in \mathbb{R}^n : \ kb_1 \le Ax \le kb_2, x \ge 0\}$ poliéderben benne van az $\mathbf{1}$ vektor (az azonosan 1 vektor). Az előző tétel szerint léteznek $z^1, \dots, z^k$ nemnegatív egész vektorok, amikre $z^1 + \dots + z^k = \mathbf{1}$, és $b_1 \le Az^i \le b_2$. Álljon az oszlopok i-edik csoportja azokból az oszlopokból, ahol z^i megfelelő komponense 1. Ez pont jó csoportokra osztást ad. □

Alkalmazzuk ezt a tételt egy páros gráf incidencia mátrixára!

1.7. Következmény. *Adott egy $G = (V, E)$ páros gráf. Tetszőleges k pozitív egész számra ki lehet színezni G éleit k színnel úgy, hogy minden v csúcsra minden színosztályból vagy $\lfloor d_G(v)/k \rfloor$ vagy $\lceil d_G(v)/k \rceil$ él illeszkedik.*

Ennek speciális esete két ismert tétel:

- Ha $k = \Delta$ (a maximális fokszám): Kőnig élszínezési tételét kapjuk, miszerint E felbontható Δ darab párosításra.
- Ha $k = \delta$ (a minimális fokszám): Gupta tételét kapjuk, ami szerint E felbontható δ részre úgy, hogy minden rész fedi az összes csúcsot.

2. Szimplex módszer

Tekintsük a következő primál feladatot:

$$Ax = b$$
$$x \ge 0$$
$$\max cx,$$

ahol $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m, x \in \mathbb{Q}^n, c \in \mathbb{Q}^{1 \times n}$. Legyen továbbá az A mátrix rangja: $r(A) = m$. Vegyük észre, hogy ha az A mátrix rangja $r(A) \neq m$, akkor a feladat vagy redundáns, vagy nem megoldható.

Nézzük a duál feladatot:

$$yA \geq c$$
$$\min yb,$$

ahol $y \in \mathbb{Q}^{1\times m}$.

2.1. Tétel (Gyenge dualitás tétel)**.** *Legyen x primál megengedett megoldás és y duál megengedett megoldás. Ekkor teljesül a következő:*

$$cx \leq yb$$

Bizonyítás.

$$cx \underbrace{\leq}_{x\geq 0, yA\geq c} (yA)x = y(\underbrace{Ax}_{=b}) = yb \qquad \square$$

2.2. Tétel (Erős dualitás tétel)**.** *Ha a primál feladat megoldható, és az optimuma korlátos (azaz cx nem lehet tetszőlegesen nagy), akkor*

$$\max cx = \min yb.$$

2.3. Tétel (Ekvivalens alak - komplementaritási feltétel)**.** *Ha x optimális primál megoldás, akkor $\exists$ y duál megoldás, amire $cx = yb$, azaz ha $x_j > 0$, akkor $(yA)_j = c_j$.*

2.1. Definíció (Primál feladat bázismegoldása)**.** Olyan x megoldás, amire A-nak az $x_j > 0$-hoz tartozó oszlopai lineárisan függetlenek.

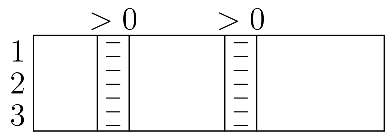
2.2. Definíció (Bázis)**.** B bázis: A-nak $m \times m$-es nemszinguláris részmátrixa.

2.1. Jelölés. $x \in \mathbb{R}^n : x = (x_B, x_N)$, ahol x_B-vel jelöljük a bázishoz tartozó koordinátákat, x_N-nel pedig a nem bázishoz tartozókat.

2.2. Jelölés. B-hez hozzárendelt $\bar{x} \in \mathbb{R}^n$: $\bar{x}_B = B^{-1}b$, $\bar{x}_N = 0$.

2.1. Megjegyzés. Ha $B^{-1}b \geq 0$, akkor $\bar{x}$ primál megoldás, azaz a B bázis megengedett.

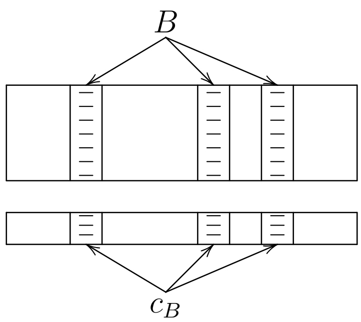
2.2. Megjegyzés.



Ha x bázismegoldás, akkor kiegészítve m db lineárisan független oszloppá bázist kapunk. Ehhez a bázishoz hozzárendelt x pedig bázismegoldás lesz, mivel tudjuk, hogy $Bx = b$-nek egy megoldása van, és mivel x megoldása, így kész is vagyunk.

2.3. Megjegyzés. Egy bázismegoldás viszont nemcsak egy bázishoz lehet hozzárendelve: a két oszlopot bárhogy kiegészítve egy harmadik, tőlük lineárisan független oszloppal több, egymástól különböző bázist kaphatunk. Két bázist különbzőnek tekintünk - még akkor is, ha ugyanaz a bázismegoldás tartozik hozzájuk.

2.3. Definíció (Bázishoz tartozó duális vektor)**.** A B bázishoz tartozó duális vektor: $\bar{y} = c_B B^{-1}$, ahol c_B a c célfüggvény B bázishoz tartozó része.



Nézzük $\bar{y}A - c$-t. Erről annyit tudunk, hogy a B-hez tartozó koordinátái nullák.

$$(\bar{y}A - c)_B = (\bar{y}B - c_B) = c_B B^{-1} B - c_B = 0$$

Ez az $\bar{y}$, $\bar{x}$ teljesíti a komplementaritási feltételeket ($\bar{y}$ nem feltétlenül megoldása a duálnak).

Ha $\bar{y}A - c \geq 0$, akkor $\bar{y}$ duál megoldás, és $\bar{x}$ és $\bar{y}$ mindig teljesíti a komplementaritási feltételeket(azaz $c\bar{x} = \bar{y}b$), így $\bar{x}$ a primál feladatnak és $\bar{y}$ a duál feladatnak optimális megoldása.

Tehát ha B primál- és duál-megengedett akkor a hozzá tartozó megoldások optimálisak.

2.1. A szimplex módszer tulajdonságai

- primál megengedett bázisokon lépked
- minden lépésben egy oszlopot cserélünk ki B-ben.
- a primál célfüggvényérték folyamatosan nő (azaz nem csökken)
- véges sok lépésben eljutunk egy duál megengedett bázishoz.

2.1. Észrevétel. *A szimplex módszer során mindig teljesül, hogy* $c\bar{x} = \bar{y}b$.

Tegyük fel, hogy B primál megengedett bázis. Tartozik hozzá egy $\bar{x}$ és egy $\bar{y}$. Írjuk fel a következő ekvivalens rendszert:

$$B^{-1}Ax = B^{-1}b$$

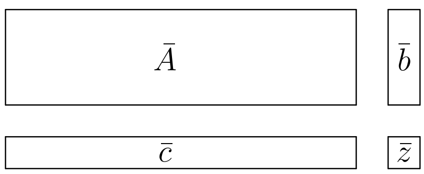
2.3. Jelölés. $\bar{A} = B^{-1}A \in \mathbb{Q}^{m \times n}$

$\bar{b} = B^{-1}b$

$\bar{c} = \bar{y}A - c$ - redukált költség.

$\bar{z} = c\bar{x} = \bar{y}b$

Nézzük a szimplex tábla ábrázolását:



2.4. Megjegyzés. $\bar{A}$-ban B egy egységmátrix, $\bar{b}$ pedig az aktuális megoldás értékeit tartalmazza.

		x_3		x_7		x_9	
x_7		0		1		0	
x_3		1		0		0	
x_9		0		0		1	

B primál megengedett bázis $\Leftrightarrow$ $\bar{b} \geq 0$

B duál megengedett bázis $\Leftrightarrow$ $\bar{c} \geq 0$

Miért nevezzük $\bar{c}$-t redukált költségnek? Írjuk át:

$$\bar{c} = \bar{y}A - c = c_B B^{-1} A - c = c_B \bar{A} - c$$

Mi történik akkor, ha $q \in N$-re $\bar{x}_q$-t növeljük δ-val, és közben $\bar{x}_B$-t változtatjuk úgy hogy $\bar{A}x = \bar{b}$ továbbra is teljesüljön?

		x_3		x_q		x_7		x_9	
x_7		0		–		1		0	
x_3		1		–		0		0	
x_9		0		–		0		1	

1. egyenletnél: x_7-et változtatjuk
2. egyenletnél: x_3-at változtatjuk
3. egyenletnél: x_9-et változtatjuk

2.4. Jelölés. Az A mátrix i. sora legyen: $a_{i.}$ vagy $A_{i.}$, j. oszlopa pedig legyen: $a_{.j}$ vagy $A_{.j}$

$$\bar{x}'_q = \bar{x}_q + \delta$$

$$\bar{x}'_B = \bar{x}_B - \delta \bar{a}_{.q}$$

$$c\bar{x}' = c\bar{x} + \delta c_q - \delta c_B \bar{a}_{.q}$$

A csökkenés mértéke: $\delta(c_B \bar{A} - c)_q = \delta \bar{c}_q$

Tehát $\bar{c}_q$ azt mondja meg, hogy mennyivel csökken a célfüggvény érték, ha $\bar{x}_q$-t növeljük. Igazából ez a csökkenés negatív, azaz a célfüggvény értéke nő (azaz nem csökken), így minden lépésben az előzőnél jobb (azaz nem rosszabb) megoldást kapunk.

2.2. A szimplex módszer egy lépése

0. ha $\bar{c} \geq 0$, akkor kész vagyunk.
1. ha nem, válasszunk egy $q \in N$-t, amire $\bar{c}_q < 0$. Ezt többféleképpen megtehetjük:
   - Bland szabály: válasszuk a legkisebb ilyen q-t. Ez a választási módszer garantálja, hogy az algoritmusunk véges lesz.
   - válasszuk a legkisebb $\bar{c}_q$ értéket. Ez nem garantálja a végességet, de a gyakorlatban gyorsabb.

   Az így választott x_q kerül majd a bázisba.
2. Ha $\bar{a}_{.q} \leq 0$, akkor

   2.1. Állítás. *Ilyenkor a célfüggvény nem korlátos.*

Bizonyítás.

$$\bar{x}'_q = \bar{x}_q + \delta$$

$$\bar{x}'_B = \bar{x}_B - \delta \bar{a}_{.q}$$

ami tetszőleges $\delta \geq 0$-ra megengedett megoldást ad. A célfüggvényérték tehát szigorúan nő ($-\delta\bar{c}_q$-vel), azaz tetszőlegesen nagy lehet. □

3. Ha $\bar{a}_{.q} \nleq 0$, akkor ki kell választani a bázisból kikerülő változót. Azt az r-et válasszuk, amire a
$$\frac{\bar{b}_r}{\bar{a}_{rq}} = \min_{i:\bar{a}_{iq}>0} \frac{\bar{b}_i}{\bar{a}_{iq}}$$

   Ekkor az r sorhoz tartozó bázisváltozó kerül ki a bázisból.

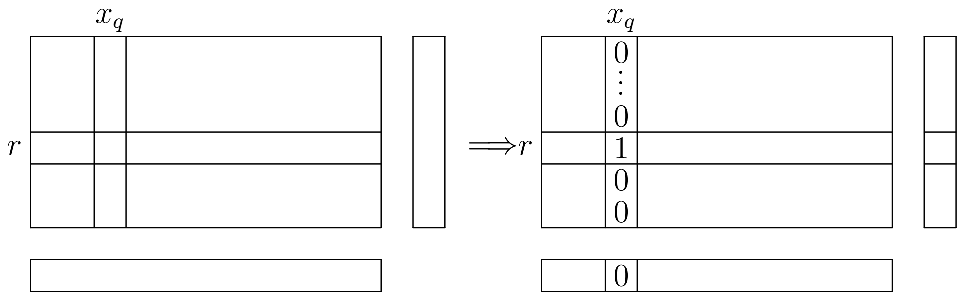
   Ha több i is minimális, akkor alkalmazzuk a Bland szabályt: az a bázisváltozó kerül ki, amelyiknek az indexe a legkisebb.

4. Új szimplex tábla kiszámítása (pivotálás): az r sor többszöröseit adjuk hozzá a többi sorhoz.(a $\bar{b}$ is hozzátartozik a sorhoz)

$$\bar{a}'_{rj} = \frac{\bar{a}_{rj}}{\bar{a}_{rq}} \qquad \bar{b}'_r = \frac{\bar{b}_r}{\bar{a}_{rq}}$$

$$i \neq r : \bar{a}'_{ij} = \bar{a}_{ij} - \bar{a}_{rj}\frac{\bar{a}_{iq}}{\bar{a}_{rq}} \qquad \bar{b}'_i = \bar{b}_i - \bar{b}_r\frac{\bar{a}_{iq}}{\bar{a}_{rq}}$$

$$\bar{c}'_j = \bar{c}_j - \bar{a}_{rj}\frac{\bar{c}_q}{\bar{a}_{rq}}$$



2.4. Tétel. *A Bland szabályt használva a szimplex módszer véges sok lépésben véget ér.*

Bizonyítás. Ha egy lépésnél változik $\bar{x}$, akkor $c\bar{x}$ szigorúan nő. Ezért csak abból lehetne probléma, hogy végtelen ciklusba kerülünk, miközben $\bar{x}$ nem változik. Tegyük fel indirekt, hogy van egy ilyen ciklus, aminek tehát az elején és a végén ugyanaz a bázis van.

Egy indexet mozgónak nevezünk, ha a hozzá tartozó változó a ciklus során ki- illetve bekerül a bázisba. A nem mozgó indexet tehát a ciklus során vagy végig a bázisban vannak, vagy végig a bázison kívül.

Legyen p a legnagyobb mozgó index, és legyen t_1 egy olyan lépés, amikor bekerül, és t_2 egy olyan lépés, ahol kikerül. Feltehetjük, hogy $t_1 < t_2$. Jelölés: a t_1 lépés előtt: $B, \bar{B}, \bar{c}, \bar{A}$; a t_2 lépés előtt: $B', \bar{B}', \bar{c}', \bar{A}'$.

Mivel p kerül be a t_1-edik lépésben, $\bar{c}_p < 0$ és $j < p$ esetén $\bar{c}_j \geq 0$.

Nézzük most a t_2-edik lépést: legyen r az x_p bázisváltozóhoz tartozó sor, és legyen q az az index, ami bekerül a bázisba. Ekkor $\bar{c}'_q < 0$, $\bar{a}'_{rq} > 0$, és $\bar{a}'_{iq} \leq 0$ az összes olyan i-re, ami mozgó bázisváltozóhoz tartozik. Az utóbbi azért igaz, mert ezekre az i-kre $\bar{b}'_i = 0$, és az ezekhez a sorokhoz tartozó változóknak p-nél kisebb az indexük.

A fent elmondottakból

$$0 < \bar{c}_q^{t_1} - \bar{c}_q^{t_2} = \bar{c}_B B^{-1} a_{.q} - \bar{c}_{B'} (B')^{-1} a_{.q} = (c_B B^{-1} B' - c_{B'}) \bar{a}'_{.q} = \bar{c}_{B'} \bar{a}'_{.q}.$$

De ha a jobb oldalon szereplő skalárszorzatot tagonként nézzük, a $\bar{c}_p \bar{a}'_{rq}$ tag szigorúan kisebb mint nulla, a többi mozgó indexhez tartozó tag legfeljebb 0, míg a nem mozgó indexekhez tartozó tagok értéke 0 (hiszen ha egy ilyen j index benne van B'-ben, akkor B-ben is benne van, tehát $\bar{c}_j = 0$).

Azt kaptuk, hogy $\bar{c}_{B'} \bar{a}'_{.q} < 0$, ellentmondás. □

3. Érzékenységvizsgálat

Legyen B optimális bázis. Mennyire változtathatjuk meg a c-nek vagy b-nek egy koordinátáját, hogy B optimális maradjon?

Tudjuk, hogy B optimális $\Leftrightarrow$ $\bar{b} \geq 0$(primál megengedett) és $\bar{c} \geq 0$ (duál megengedett).

1. $q \in N$-re: $c'_q = c_q + \delta$
   - $\bar{b}$ nem változik, ekkor B primál megengedett marad
   - $\bar{c}' = \bar{y}A - c' = \underbrace{c_B \bar{A}}_{\text{nem változik}} - \underbrace{c'}_{\text{csak ez változik}}$

$$\bar{c}'_j = \bar{c}_j, \text{ ha } j \neq q$$

$$\bar{c}'_q = \bar{c}_q - \delta$$

$$\bar{c}' \geq 0 \Leftrightarrow \delta \leq \bar{c}_q$$

2. Az r sorhoz tartozó bázisváltozó súlyát növeljük δ-val.
   - $\bar{b}$ nem változik, ekkor B primál megengedett marad
   - $\bar{c}' = c'_B \overline{A} - c'$. Ekkor $\bar{c}'_B \equiv 0$(ez mindig igaz), $\bar{c}'_N = \bar{c}_N + (\delta \bar{a}_{r.})_N$. Azaz $j \in N : \bar{c}'_j = \bar{c}_j + \delta \bar{a}_{rj}$. Ez mikor marad nemnegatív?
     - Ha $\bar{a}_{rj} = 0$, akkor mindig.
     - Ha $\bar{a}_{rj} > 0$, akkor szükséges, hogy $\delta \geq -\frac{\bar{c}_j}{\bar{a}_{rj}}$
     - Ha $\bar{a}_{rj} < 0$, akkor szükséges, hogy $\delta \leq -\frac{\bar{c}_j}{\bar{a}_{rj}}$

   Tehát

$$\bar{c}' \geq 0$$

$$\Updownarrow$$

$$\max\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} > 0\} \leq \delta \leq \min\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} < 0\}$$

Látjuk, hogy az alsó korlát egy nempozitív szám, a felső korlát egy nemnegatív szám, de mindkettő lehet nulla is. Továbbá ha üreshalmazon maximalizálunk, akkor $-\infty$-t kapunk, illetve ha üreshalmazon minimalizálunk, akkor $+\infty$-t kapunk.

3. Legyen $b'_r = b_r + \delta$. Ekkor

   - $\bar{c}$ nem változik, ekkor B duál megengedett marad.
   - $\bar{b}' = B^{-1}b'$, tehát $\bar{b}'_i = B^{-1}_{i.}b' = \bar{b}_i + \delta B^{-1}_{ir}$. Hasonlóan az előző esethez:

$$\bar{b}' \geq 0$$

$$\Updownarrow$$

$$\max\{-\frac{\bar{b}_i}{\bar{B}^{-1}_{ir}} : \bar{b}_i > 0\} \leq \delta \leq \min\{-\frac{\bar{b}_i}{\bar{B}^{-1}_{ir}} : \bar{b}_i < 0\}$$

Tehát ha

$$A : \begin{array}{|c|c|}\hline \qquad\qquad & \mathcal{I} \\ \hline \end{array}$$

$$c : \begin{array}{|c|c|}\hline \qquad\qquad & 0 \ldots 0 \\ \hline \end{array}$$

akkor

$$\bar{A} : \begin{array}{|c|c|}\hline \qquad\qquad & B^{-1} \\ \hline \end{array}$$

$$\bar{c} : \begin{array}{|c|c|}\hline \qquad\qquad & \bar{y} \\ \hline \end{array}$$

B^{-1} és $\bar{y}$ is könnyen kiolvasható a szimplex táblából.
Most nézzük meg, hogyan változik $\bar{z}$: $\bar{z}' = \bar{y}b' = \bar{z} + \delta\bar{y}_r$.
$\bar{y}$-t árnyékárnak nevezzük, mivel $\bar{y}_r$ megmondja, hogy - élve a termelési feladat példájával - milyen egységáron érdemes az r-edik alapanyagból vásárolni (feltéve, hogy amennyit vásárolunk az adott határokon belül marad).

Mi lesz akkor, ha az elején nincs primál megengedett megoldásunk? Ehhez nézzük meg a duál szimplex módszert:

4. Duál szimplex módszer

4.1. A duál szimplex módszer tulajdonságai

- duál megengedett bázisokon lépked
- minden lépésben egy oszlopot cserélünk ki B-ben.
- ugyanazt a szimplex táblát használjuk, mint a primál szimplex módszernél
- $c\bar{x}$ folyamatosan csökken (azaz nem nő)

- véges sok lépésben eljutunk egy primál megengedett bázishoz.

A különbség annyi lesz, hogy először azt határozzuk meg, hogy melyik változó lép ki a bázisból, és csak utána azt, hogy melyik lép be.

4.2. A duál szimplex módszer egy lépése

0. Ha $\bar{b} \geq 0$, akkor kész vagyunk. Primál megengedett bázisunk van, azaz optimális bázist találtunk.

1. Ha nem, válasszunk egy r-et, amire $\bar{b}_r < 0$. Ezt többféleképpen megtehetjük:

   - Bland szabály: válasszuk azt az r-et, amihez a legkisebb indexű bázisváltozó tartozik. Ez a választási módszer garantálja, hogy az algoritmusunk véges lesz.
   - Válasszuk a legkisebb $\bar{b}_r$ értéket. Ez nem garantálja a végességet, de a gyakorlatban gyorsabb.

   Az így választott r-hez tartozó bázisváltozó lép ki a bázisból.

2. ha $\bar{a}_{r.} \geq 0$, akkor

   4.1. Állítás. *A primál feladatnak nincs megoldása.*

   Bizonyítás.
   $$\underbrace{\bar{a}_{r.}x}_{\geq 0} = \underbrace{\bar{b}_r}_{<0}$$
   egy érvényes egyenlet lenne, ami nem lehetséges. □

3. Ha $\bar{a}_{r.} \ngeq 0$, akkor ki kell választani a bázisba bekerülő változót. Azt az x_q-t válasszuk, amire $\bar{a}_{rq} < 0$ és
   $$-\frac{\bar{c}_q}{\bar{a}_{rq}} = \min\{-\frac{\bar{c}_j}{\bar{a}_{rj}} : j \in N, \bar{a}_{rj} < 0\}$$

   4.1. Megjegyzés. Ha ez a minimum nulla, akkor degeneráció lép fel a duál szimplexben, tehát $\bar{c}$ nem változik a báziscsere során.

   Ekkor x_q kerül a bázisba.

   Ha több q is minimális, akkor alkalmazzuk a Bland szabályt: válasszuk a minimális ilyen q-t.

   Miért pont így kell választani a bemenő változót? Arra van szükségünk, hogy $\bar{c} \geq 0$ maradjon.

   Báziscsere után: $\bar{c}'_j = \bar{c}_j - \bar{a}_{rj}\frac{\bar{c}_q}{\bar{a}_{rq}}$ nemnegatív marad, ha

   - $\bar{a}_{rj} \geq 0$, mivel $\bar{c}_j$-t ekkor növeljük
   - $\bar{a}_{rj} < 0$, de $-\frac{\bar{c}_j}{\bar{a}_{rj}} \geq -\frac{\bar{c}_q}{\bar{a}_{rq}}$

4. a báziscsere ugyanúgy történik, mint a primál szimplex módszernél.

□

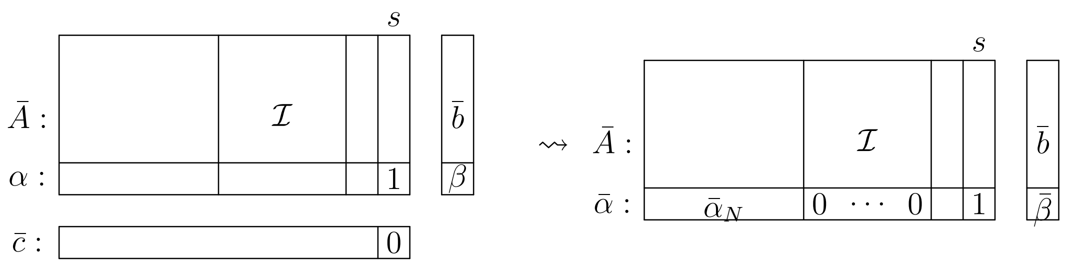
Mikor érdemes a duál szimplex módszert alkalmazni?

4.3. Alkalmazás

Tegyük fel, hogy már megoldottunk egy feladatot, és hozzá kell vennünk még a rendszerhez egy $\alpha x \leq \beta$ feltételt.

Vegyünk egy új slack változót: $s \geq 0$ úgy, hogy $\alpha x + s = \beta$. Írjuk át a feltételt ekvivalensen: $\bar{\alpha} x + s = \bar{\beta}$, ahol $\bar{\alpha}_B = 0$.

Azaz:



4.2. Megjegyzés. $\bar{c}$-on nem kell változtatni, ugyanis a slack-változóhoz 0 tartozik $\bar{c}$-ban.

A bázist kibővítjük s-sel, így az új feladatra egy duál megengedett bázist kapunk, és innentől fogva használhatjuk a duál szimplex módszert, mivel kaptunk egy kiindulási táblát a módszerhez.

4.4. Alkalmazás

A duál szimplexszel a megengedettségi feladatot azonnal meg tudjuk oldani. Nézzük a következő primál feladatot és a hozzá tartozó duált:

$$(P) \quad \begin{aligned} Ax &= b \\ x &\geq 0 \\ \max\, &0x \end{aligned} \qquad (D) \quad \begin{aligned} yA &\geq 0 \\ & \\ \min\, &yb \end{aligned}$$

A (P) feladatnak minden megengedett megoldása optimális, miközben a (D) feladatnak a $(0 \ldots 0)$ egy megengedett megoldása.

Ebben a feladatban $\bar{c} = \underbrace{c_B \bar{A}}_{=0} - \underbrace{c}_{=0} = 0$, tehát minden bázis duál-megengedett.

Induljunk ki tetszőleges bázisból, és használjuk a duál szimplex módszert. Ekkor vagy kapunk egy primál megengedett bázist, vagy kapunk egy bizonyítékot arra, hogy a feladat nem megoldható. (Ez a bizonyíték pont a Farkas lemmából következik.)

5. Kétfázisú szimplex módszer

1. fázis: primál megengedett bázis keresése

2. fázis: primál szimplex módszer alkalmazása ebből a bázisból kiindulva.

Nézzük az első fázist. Legyen a feladat: $\max\{cx : \; Ax = b, x \geq 0\}$.

Feltesszük, hogy $b \geq 0$. Különben a megfelelő sorokat (-1)-gyel be kell szorozni.

Vezessünk be minden sorhoz új mesterséges változókat: $u_i : i = 1 \dots m$.

$$\begin{aligned} Ax + Iu &= b \\ (x, u) &\geq 0 \end{aligned}$$

5.1. Megjegyzés. Vegyük észre, hogy a fenti rendszer nem ekvivalens az eredetivel.

Ahhoz, hogy az eredetivel ekvivalens rendszert kapjunk, szükség van egy olyan megoldásra, ahol $u = 0$. Legyen a célfüggvény: $\max -\sum_{i=1}^{m} u_i$.

- Ha ennek a feladatnak 0 az optimuma, akkor az eredeti feladat egy megoldását kaptuk.
- Ha ennek a feladatnak < 0 az optimuma, akkor az eredeti feladatnak nincs megoldása.

Legyen B a mesterséges változók oszlopaiból álló bázis. Ekkor B primál megengedett. A kiindulási szimplex tábla:

$$A^1 : \begin{array}{|c|c|} \hline A & \mathcal{I} \\ \hline \end{array} \quad \begin{array}{|c|} \hline b \\ \hline \end{array}$$

$$c^1 : \begin{array}{|c|c|} \hline 0 \quad \cdots \quad 0 & -1 \cdots -1 \\ \hline \end{array}$$

$$\begin{aligned} \bar{A}^1 &= A^1 \\ \bar{b}^1 &= b \\ \bar{c}^1 &= c_B^1 \bar{A}^1 - c^1 \end{aligned}$$

Tudjuk, hogy $\bar{c}_B^1 = 0$, és $\bar{c}_N^1 = -\sum_{i=1}^{m} a_{i.}$.

A kiindulási szimplex tábla tehát:

$$\begin{array}{|c|c|} \hline A & \mathcal{I} \\ \hline \end{array} \quad \begin{array}{|c|} \hline b \\ \hline \end{array}$$

$$\bar{c}^1 : \begin{array}{|c|c|} \hline -\sum a_i & 0 \cdots 0 \\ \hline \end{array}$$

Erre a szimplex táblára kell alkalmazni a primál szimplex módszert.

- Ha az optimum < 0, akkor nincs megoldás.
- Ha az optimum $= 0$, de marad mesterséges változó a bázisban (azaz a feladatunk degenerált volt), akkor tegyük a következőt a mesterséges változók kiküszöbölése érdekében:

$$\begin{array}{r} \\ (r.\text{sor })u_i \\ \bar{A}^1 : \\ \\ \end{array} \begin{array}{|c|c|c|c|} \multicolumn{4}{c}{\qquad\qquad\qquad u_i} \\ \hline & & & \\ \hline & & & \\ \hline & & & \\ & & & \\ \hline \end{array} \quad \begin{array}{|c|} \multicolumn{1}{c}{} \\ \hline \\ \hline 0 \\ \hline \\ \\ \hline \end{array}$$

Legyen a megmaradt változó a szimplex tábla r-edik sorában. Ez a sor nem lehet azonosan 0, mert A sorai lineárisan függetlenek.

Válasszunk tehát egy tetszőleges nemnulla elemet egy nem-mesterséges változó oszlopában, és ott pivotáljunk:

					u_i			
$(r.\text{sor })u_i$		$\neq 0$						0
$\bar{A}^1$:								

Mivel a jobboldalon 0 van, ezért a primál megoldás nem változik, és a célfüggvényérték sem változik.

5.2. Megjegyzés. Azért kell nemnulla elemet választanunk, hogy a pivotálást el tudjuk végezni.

- Ha az optimum = 0 és nincs mesterséges változó a bázisban (azaz az eredeti feladatra kaptunk egy megengedett bázist), akkor elhagyjuk a mesterséges változókat, és az eredeti c célfüggvényre alkalmazzuk a szimplex módszert.

5.1. Szoftveres szemléltetés

A fent bemutatott kétfázisú szimplex módszer működése megtekinthető a Paul A. Jensen által készített ORMM Excel bővítmény segítségével:
`http://www.me.utexas.edu/~jensen/ORMM/index.html`

Az összes bővítmény letölthető egy .zip fájlban innen:
`http://www.me.utexas.edu/~jensen/ORMM/frontpage/jensen.lib/pclib/jensen.lib.zip`

6. Hálózati szimplex módszer

Tekintsük a következő feladatot:

Adott egy $D = (V, E)$ irányított, gyengén összefüggő gráf, $b : V \to \mathbb{Z}$ az igények minden csúcsra, ahol $\sum_{v\in V} b_v = 0$ és $c : E \to \mathbb{R}$ az élek súlya.

Legyen $x : E \to \mathbb{R}$ a változó.

6.1. Jelölés (Befolyó áram)**.** $\rho_x(v)$ jelölje a v csúcsba belépő éleken az x-ek összegét.

6.2. Jelölés (Kifolyó áram)**.** $\delta_x(v)$ jelölje a v csúcsból kilépő éleken az x-ek összegét.

$$\rho_x(v) - \delta_x(v) = b_v, \ \forall v \in V$$
$$x \geq 0$$
$$\max \sum_{e\in E} c_e x_e$$

6.1. Megjegyzés. Ha $\sum b_v = 0$ nem lenne igaz, akkor a feladatnak nem lenne megoldása.

6.2. Megjegyzés. A hálózati folyam feladatnak mindig van optimális egész megoldása, ha a feladat egyáltalán megoldható.

Legyen v_0 egy kijelölt csúcs. Ha $\rho_x(v) - \delta_x(v) = b_v$ minden $v \in V \setminus \{v_0\}$-ra teljesül, ekkor v_0-ra is teljesül. Vezessük be tehát az eredeti egyenlőségrendszer helyetta következőt: $\rho_x(v) - \delta_x(v) = b_v \ \ \forall v \in V \setminus \{v_0\}$. Így a sorok lineárisan függetlenek lesznek, ezáltal a feladatra alkalmazhatjuk a szimplex módszert: $Ax = b, \quad x \geq 0$, ahol A sorai lineárisan függetlenek.

Az A mátrixunk a következő lesz:

$$V = \{v_0, v_1, \dots, v_m\},\ E = \{e_0, e_1, \dots, e_n\}$$
$$A \in \mathbb{Z}^{mxn}$$
$$a_{ij} = \begin{cases} -1 & \text{ha } v_i \text{ töve } e_j\text{-nek} \\ +1 & \text{ha } v_i \text{ feje } e_j\text{-nek} \\ 0 & \text{különben} \end{cases}$$

6.3. Megjegyzés. Az A mátrix minden oszlopában legfeljebb egy $+1$-es és legfeljebb egy -1-es található.

6.3. Jelölés. A továbbiakban kontextustól függően a következő ekvivalens jelöléseket fogjuk használni:

$$\begin{array}{ccc} b_{v_i} & \sim & b_i \\ c_{e_j} & \sim & c_j \\ y_{v_i} & \sim & y_i \\ x_{v_i} & \sim & x_i \end{array}$$

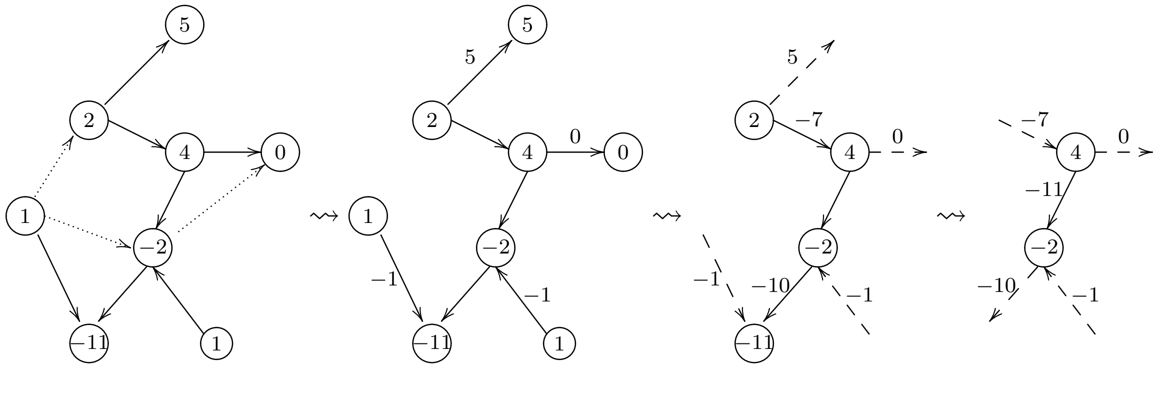
A feladatunk tehát felírható $\max\{cx, Ax = b, x \geq 0\}$ alakban.

6.1. Állítás. *B bázis $\Leftrightarrow$ $\{e_j : j \in B\}$ feszítőfa (irányítás nélkül).*

Bizonyítás. $\Leftarrow$: Belátjuk, hogy $Bx = b$ egyértelműen megoldható. Keressünk a fán olyan áramot, ami minden igényt kielégít. Minden élre adjunk olyan értéket, ahol $\rho_x(v) - \delta_x(v) = b_v$. A fa leveleire egy-egy él illeszkedik. Ezekre egyértelműen meg tudjuk adni a változó értékét.

Ha a leveleket elhagyjuk, akkor újabb fát kapunk, amely fa leveleire ugyancsak egyértelműen meghatározható a változó értéke.

Például:



$\Rightarrow$: Indirekt bebizonyítjuk, hogy ha a B-hez tartozó élek nem alkotnak feszítőfát, akkor B nem bázis. Legyen m darab él, ami nem alkot feszítőfát. Ekkor a részgráf tartalmaz kört. Legyen ez a kör C.

Legyen

$$x_e = \begin{cases} +1 & \text{, ha } e \in C \text{ előre-él} \\ -1 & \text{, ha } e \in C \text{ hátra-él} \\ 0 & \text{, ha } e \notin C \end{cases}$$

Ekkor a $Bx = 0$ feladatra $x \neq 0$, tehát B szinguláris, azaz B nem bázis. □

Legyen $\bar{x}$ a $Bx = b$ egyértelmű megoldása és legyen $\bar{y}$ a következő: $\bar{y}_0 = 0$ lesz a v_0-hoz tartozó duális változó. Továbbá ha $uv \in B$, akkor legyen $\bar{y}_v - \bar{y}_u = c_{uv}$. Ez egyértelműen meghatározza $\bar{y}$-t.

Az uv él redukált költsége: $\bar{c}_{uv} = \bar{y}_v - \bar{y}_u - c_{uv}$.

6.1. Definíció. A B bázis primál megengedett, ha $\bar{x} \geq 0$.

6.2. Definíció. A B bázis duál megengedett, ha $\bar{c} \geq 0$.

6.1. Primál hálózati szimplex módszer lépései

Tegyük fel, hogy B primál megengedett bázis. Az alábbi lépéssorozatnál nem kell az A mátrixot felhasználnunk.

0. Ha $\bar{c} \geq 0$, akkor kész vagyunk (a bázisunk primál és duál megengedett).

1. Ha nem, akkor válasszunk egy olyan e_p élt, amire $\bar{c}_p < 0$. Ha több alternatívánk van, alkalmazzuk a Bland szabályt: válasszuk a legkisebb indexű ilyet. Az így választott e_p él kerül majd a bázisba.

2. Vegyük hozzá a B feszítőfához az e_p élt. Ekkor egy egyértelmű C kört kapunk, aminek e_p éle. Legyenek a C-ben e_p-vel egyirányú élek előreélek, a többi C-beli él pedig legyen hátraél.

   Ha C-ben nincsenek hátraélek, akkor tetzőleges $\delta > 0$-ra

   $$\bar{x}' = \begin{cases} \bar{x}_e + \delta, \text{ ha } e \in C \\ \bar{x}_e, \text{ különben} \end{cases}$$

   megengedett megoldás.

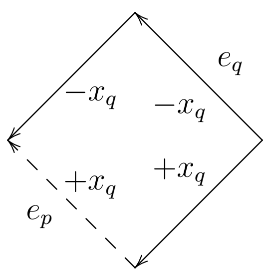
   6.2. Állítás. *Ebben az esetben a célfüggvény nem korlátos.*

   Bizonyítás. $\sum_{uv \in C} \bar{c}_{uv} = \sum_{uv \in C} (\bar{y}_v - \bar{y}_u - c_{uv}) = -\sum_{uv \in C} c_{uv}$
   $c\bar{x}' = c\bar{x} + \delta \sum_{uv \in C} c_{uv} = c\bar{x} - \delta \sum_{uv \in C} \bar{c}_{uv} = c\bar{x} - \delta \bar{c}_p \to_{\delta \to \infty} +\infty$
   Tehát a célfüggvény nem korlátos. □

3. Ha van C-ben hátraél, akkor legyen e_q az a hátraél, amire $\bar{x}_q$ minimális. Ez az él fog kikerülni a bázisból.

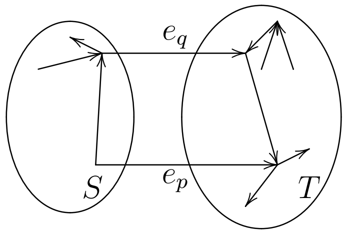
4. Vegyük hozzá a B bázishoz az e_p élt, és hagyjuk el a bázisból az e_q élt. $B' = B + \{p\} - \{q\}$

   $$\bar{x}'_j = \begin{cases} \bar{x}_j + \bar{x}_q, \text{ ha } e_j \text{ előreél } C\text{-ben} \\ \bar{x}_j - \bar{x}_q, \text{ ha } e_j \text{ hátraél } C\text{-ben} \\ \bar{x}_j, \text{ ha } e_j \notin C \end{cases}$$



Számoljuk ki az $\bar{y}'$-t:

Ha az e_p élt kihagyjuk a fából, akkor a fa két komponensre esik: S és T. Válasszuk S-t és T-t úgy, hogy e_p a T-be lépjen.



Ekkor

- Ha $v_o \in S$, akkor
$$\bar{y}'_v = \begin{cases} \bar{y}_v, \text{ ha } v \in S \\ \bar{y}_v - \bar{c}_p, \text{ ha } v \in T \end{cases}$$
- Ha $v_o \in T$, akkor
$$\bar{y}'_v = \begin{cases} \bar{y}_v + \bar{c}_p, \text{ ha } v \in S \\ \bar{y}_v, \text{ ha } v \in T \end{cases}$$

A fenti megkülönböztetés azért szükséges, hogy $y_0 = 0$ maradjon.

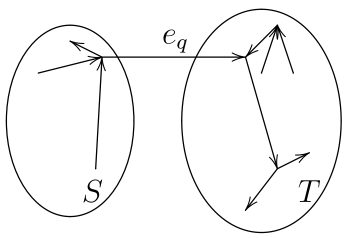
6.4. Megjegyzés. Ha b és c egészek, akkor $\bar{y}$, $\bar{x}$ és $\bar{c}$ végig egészek maradnak. Sőt, ha a költségek(súlyok) egészek, a duál végig egész lesz, és ha az igények egészek, a primál végig egész lesz.

6.2. Duál hálózati szimplex módszer

Legyen B duál megengedett bázis (azaz $\bar{c} \geq 0$).

0. Ha $\bar{x} \geq 0$, akkor kész vagyunk (a bázisunk primál és duál megengedett).
1. Ha nem, akkor legyen $q \in B$ olyan, hogy $\bar{x}_q < 0$. Ha több alternatívánk van, alkalmazzuk a Bland szabályt: válasszuk a legkisebb indexű ilyet. Az így választott változó kerül majd ki a bázisból.

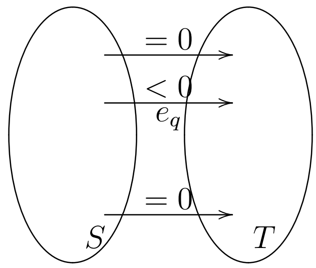
   Ha elhagyjuk az e_q élt, akkor a fa két részre esik: S és T. Válasszuk S-t és T-t úgy, hogy e_q a T-be lépjen.

2.

   6.3. Állítás. *Ha nincs T-ből S-be vezető él, akkor a primál feladatnak nincs megoldása.*

Bizonyítás. Az $\bar{x}$ aktuális primál vektor a $\rho_x(v) - \delta_x(v) = b_v$, $\forall v \in V$ feladat megoldása. Az e_q élen: $\bar{x} < 0$. A többi S-ből T-be vezető élen $\bar{x} = 0$, mivel ezek nincsenek a bázisban.



Látjuk, hogy S-ből negatív folyam megy T-be, azaz S igénye nagyobb T igényénél.

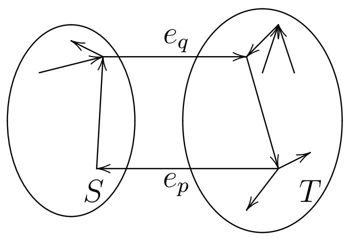
$\sum_{v\in S} b_v > \sum_{v\in T} b_v$, mivel

$\sum_{v\in T} b_v - \sum_{v\in S} b_v = \sum_{v\in T}(\rho_{\bar{x}}(v) - \delta_{\bar{x}}(v)) - \sum_{v\in S}(\rho_{\bar{x}}(v) - \delta_{\bar{x}}(v)) = 2\sum_{e:S\to T} \bar{x}_e < 0$

Nemnegatív folyammal nem lehet az igényeket kielégíteni, tehát nincs megoldás. $\square$

3. Ha van T-ből S-be él, akkor válasszuk ki azt az e_p $T \to S$ élt, amire $\bar{c}_p = \min\{\bar{c}_e$, ahol $e : T \to S$ él$\}$. Ha több ilyen van, akkor alkalmazzuk a Bland szabályt: válaszzuk a legkisebb indexűt. Ez az él fog bekerülni a bázisba.

4. Hagyjuk el a B bázisból az e_q élt, és vegyük hozzá az e_p élt. $B' = B - \{q\} + \{p\}$.



Legyen $C : E_B \cup \{e_p\}$ egyetlen köre.

Legyenek az e_p-vel egyirányú élek előreélek, a vele ellentétes irányú élek pedig hátraélek. Ekkor

$$\bar{x}'_e = \begin{cases} \bar{x}_e - \bar{x}_q, \text{ ha } e \text{ előreél } C\text{-ben} \\ \bar{x}_e + \bar{x}_q, \text{ ha } e \text{ hátraél } C\text{-ben} \\ \bar{x}_e, \text{ ha } e \notin C \end{cases}$$

Számoljuk ki az $\bar{y}'$-t:

- Ha $v_o \in S$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v, \text{ ha } v \in S \\ \bar{y}_v + \bar{c}_p, \text{ ha } v \in T \end{cases}$$

- Ha $v_o \in T$, akkor

$$\bar{y}'_v = \begin{cases} \bar{y}_v - \bar{c}_p, \text{ ha } v \in S \\ \bar{y}_v, \text{ ha } v \in T \end{cases}$$

6.3. Kezdeti primál bázis keresése

Több módszert is ismertetünk:

1. A $c \equiv 0$ súlyfüggvényre alkalmazzuk a hálózati szimplex módszert. Ilyenkor tetszőleges B bázisra $\bar{y} \equiv 0$, $\bar{c} \equiv 0$. A gyakorlatban ez lassú módszer, és csak azért nem ciklizál, mert a Bland szabályt alkalmazzuk.

2. Legyen B tetszőleges bázis. Ha ez a bázis se nem primál-, se nem duál megengedett, akkor módosítsuk c-t úgy, hogy duál megengedett legyen:

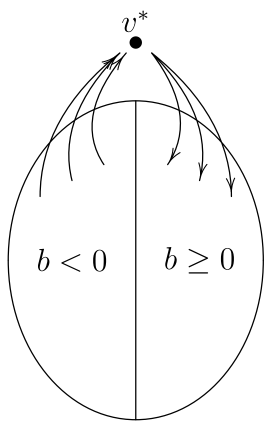
$$c'_{uv} = \begin{cases} c_{uv}, \text{ ha } \bar{c}_{uv} \geq 0 \\ \bar{y}_v - \bar{y}_u, \text{ ha } \bar{c}_{uv} < 0 \end{cases}$$

   Ezzel a c'-vel B duál megengedett lesz. Most alkalmazzuk a duál szimplex módszert, ezáltal B' optimális bázist kapunk c'-re, ami primál megengedett bázis az eredeti c célfüggvényre. Ezzel a B'-vel kezdve alkalmazzuk a primál szimplex módszert az eredeti c célfüggvényre.

3. **Kétfázisú szimplex módszer**

   Vegyünk hozzá az eredeti gráfhoz egy új csúcsot: v^* $(V' = V \cup \{v^*\})$ és a következő új éleket: $E' = E \cup \{v^*v : b_v \geq 0\} \cup \{vv^* : v_v < 0\}$.

$$c'_e = \begin{cases} -1, \text{ ha } e = v^*v \text{ vagy } e = vv^* \\ 0, \text{ ha } e \in E \end{cases}$$



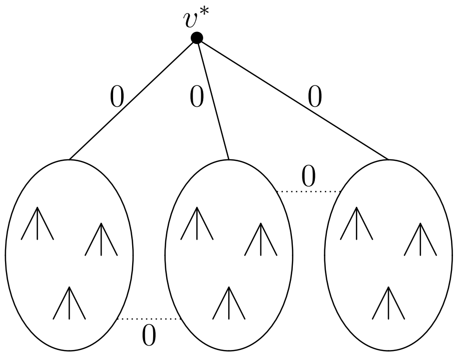
   A v^*-ra illeszkedő élek primál megengedett bázist határoznak meg, mivel:

   $\bar{x}_{v^*v} = b_v$ és $\bar{x}_{vv^*} = -b_v$, tehát $\bar{x} \geq 0$.

   Alkalmazzuk a primál hálózati szimplex módszert. Ekkor

   - Ha az optimum negatív, akkor az eredeti feladatnak nincs megoldása.
   - ha az optimum nulla, akkor hagyjuk el a v^*-os éleket, és az így keletkezett részfákat egészítsük ki feszítőfává 0-ás éleket hozzávéve. Így primál megengedett bázist kapunk

az eredeti feladatra.



7. Egészértékű lineáris programozás

7.1. Bevezetés

A következő alapfeladattal foglalkozunk:

$$\max cx$$
$$Ax \leq b$$
$$x \in \mathbb{Z}^n$$

A fenti feladat LP-relaxáltja:

$$\max cx$$
$$Ax \leq b$$
$$x \in \mathbb{R}^n$$

Speciális eset: bináris programozási feladat:

$$\max\{cx : Ax \leq b, x \in \{0,1\}^n\}$$

Például: **Bináris hátizsák feladat**

Adott n db tárgy.

A j-edik tárgy értéke legyen $c_j \geq 0$, súlya pedig $a_j > 0$.

Adott továbbá egy hátizsák, melynek teherbíró képessége $b > 0$.

Keressük max $\{\sum_{j=1}^n c_j x_j : \sum_{j=1}^n a_j x_j \leq b, x \in \{0,1\}^n\}$

A fenti feladat egy NP-teljes feladat. Az LP relaxáltja könnyen megoldható: az érték/súly arány szerint helyezzük csökkenő sorrendbe a tárgyakat. Sorra tegyük be őket a hátizsákba, amíg a hátizsák be nem telik(az utolsó betett tárgy lehet hogy csak részben fér be, de a relaxált feladatnál ez nem baj).

Az egészértékű programozási feladatnál kicsit általánosabb a **vegyes programozási feladat**, ahol nem feltétlenül az összes változónak kell egészértékűnek lenni:

$$\max\{cx + dz : \ Ax + Bz \leq b, x \in \mathbb{Z}^{n_1}, z \in \mathbb{R}^{n_2}\}.$$

Egy példa: **Szolgáltató-elhelyezési feladat**

Adott m db ügyfél, és n db lehetséges szolgáltatóhely. Minden ügyfélnek egységnyi igénye van, ezt megoszthatjuk több szolgáltatóhely között.

c_{ij}: i-edik ügyfél kiszolgálásának költsége a j-edik szolgáltatóhelyről.

f_j: j-edik szolgáltatóhely megnyitásának költsége

u_j: j-edik szolgáltatóhely kapacitása

A feladat felírása vegyes programozási feladatként:

$$
\begin{aligned}
\min \sum_{j=1}^{n}(f_j + \sum_{i=1}^{m} c_{ij}x_{ij}) & \\
0 \le x_{ij} \le y_j & \qquad i \in \{1,\dots,m\},\ j \in \{1,\dots,n\} \\
y_j \in \{0,1\} & \qquad j \in \{1,\dots,n\} \\
\sum_{j=1}^{n} x_{ij} = 1 & \qquad i \in \{1,\dots,m\} \\
\sum_{i=1}^{m} x_{ij} \le u_j y_j & \qquad j \in \{1,\dots,n\}.
\end{aligned}
$$

Ha egy egészértékű programozási feladatot meg akarunk oldani, akkor tulajdonképpen egy poliéder egész pontjainak konvex burkán akarunk egy lineáris célfüggvényt maximalizálni.

7.1. Definíció. Legyen P poliéder, $P \subseteq \mathbb{R}^n$. Ekkor $P_I = \mathrm{conv}(P \cap \mathbb{Z}^n)$ a P egész pontjainak konvex burka.

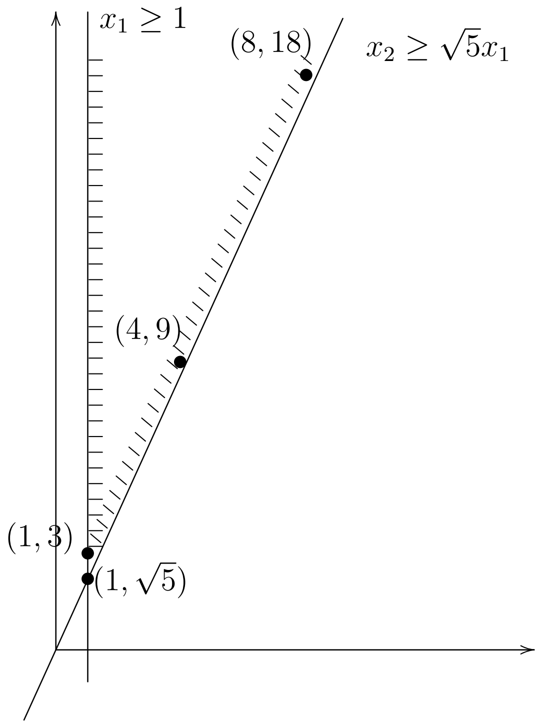
7.1. Megjegyzés. Ezzel a definicióval viszont az egész pontok konvex burka nem lesz mindig poliéder.

Például

2-dimenzióban: Legyen $P = \{(x_1, x_2) : x_1 \ge 1, x_2 \ge \sqrt{5}x_1\}$.

Ekkor az egész pontok konvex burkának végtelen sok csúcsa lesz. Például: $(1,3)$, $(4,9)$, $(8,18)$, stb.

Az irracionális meredekségű egyeneshez tetszőlegesen közel tudunk egész pontot találni. $(4,9)$ közelebb van, mint $(1,3)$, de $(8,18)$ már közelebb van, mint $(4,9)$, és így tovább, egyre közelebb kerülünk, és közben mindig új csúcsokat definiálunk, ezáltal végtelen sok csúcsot kapunk a konvex burokhoz.



7.1. Tétel (Meyer)**.** *Ha P racionális poliéder, akkor P_I poliéder.*

Bizonyítás. Ha P korlátos, akkor P_I véges sok pont konvex burka, tehát poliéder. Az tehát az érdekes eset, amikor P nem korlátos. Ekkor a Motzkin tétel szerint $P = Q + C$, ahol Q egy racionális politop (korlátos poliéder), és C egy végesen generált racionális kúp. A C kúpot véges sok egész vektor is generálja, legyenek ezek $g_1, \ldots, g_k$. Tekintsük a következő korlátos poliédert:

$$D = \left\{ \sum_{i=1}^{k} \lambda_i g_i : \ 0 \leq \lambda_i \leq 1 \ (i = 1, \ldots, k) \right\}.$$

Belátjuk, hogy $P_I = (Q+D)_I + C$. Ez elég a tétel bizonyításához, hiszen $Q+D$ korlátos, tehát $(Q+D)_I$ poliéder, és $(Q+D)_I + C$ is.

Egyik irány: $P_I \subseteq (Q+D)_I + C$. Mivel $(Q+D)_I + C$ konvex halmaz, elég belátni, hogy a P minden egész p pontja benne van. Legyen tehát $p \in P \cap \mathbb{Z}^n$.

Motzkin tétele szerint p felírható $q + c$ alakban, ahol $q \in Q, c \in C$. Ekkor léteznek olyan nemnegatív μ_i együtthatók, melyekkel:

$$p = q + c = q + \sum_{i=1}^{k} \mu_i g_i = (q + \sum_{i=1}^{k} (\mu_i - \lfloor \mu_i \rfloor) g_i) + \sum_{i=1}^{k} (\lfloor \mu_i \rfloor) g_i) =: (q+d) + c'.$$

Itt $d \in D$ és $c' \in C$ nyilvánvalóan teljesül, továbbá $q + d$ egész, mivel $q + d = p - c'$, azaz előáll két egész vektor különbségeként. Így p tényleg benne van $(Q+D)_I + C$-ben.

Másik irány: $P_I \supseteq (Q+D)_I + C$. Nyilván $Q + D \subseteq P$, így

$$(Q+D)_I + C \subseteq P_I + C = P_I + C_I.$$

Tetszőleges X és Y halmazokra igaz, hogy $\mathrm{conv}(X) + \mathrm{conv}(Y) \subseteq \mathrm{conv}(X+Y)$. Ezt használva:

$$P_I + C_I \subseteq \mathrm{conv}((P \cap \mathbb{Z}^n) + (C \cap \mathbb{Z}^n)) \subseteq (P+C)_I = P_I.$$

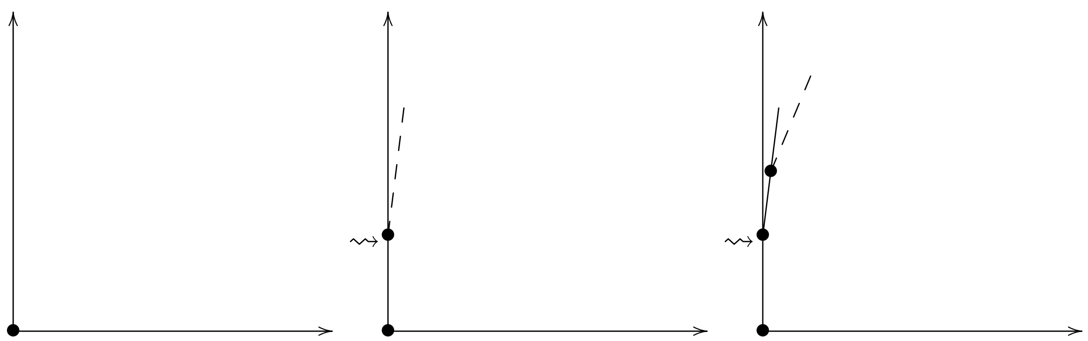
□

7.2. Megjegyzés. Lehet, hogy P_I-nek sokkal több csúcsa van, mint P-nek.

Például

Induljunk ki a következő kúpból: $\{x_1 \geq 0, x_2 \geq 0\}$. Itt az origó benne van P_I-ben. Válasszunk x_2-n tetszőleges egész pontot, majd "döntsük meg" a lapot úgy, hogy mindkét pont benne legyen, közben az újonnan bekerült részbe ne legyenek egész pontok. Keressünk egy újabb egész pontot, és folytassuk az eljárást.

Ezzel az eljárással tetszőleges N számra le tudunk gyártani olyan egy-csúcsú poliédert, amelyhez tartozó egész pontok konvex burkának N csúcsa van, és mindezt már 2-dimenzióban is megtehetjük.



A továbbiakban nézzünk néhány algoritmust az egészértékű programozási feladat megoldására:

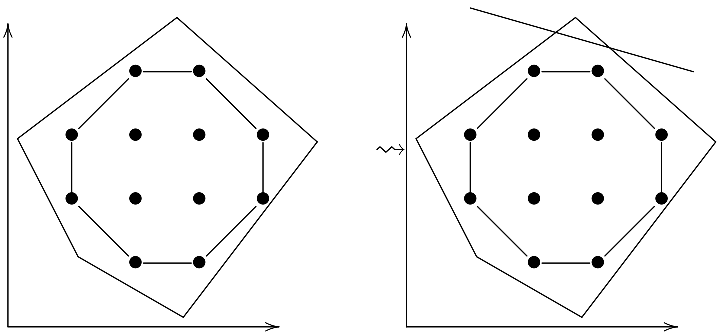
7.2. Vágósíkos eljárás

Nézzük a következő feladatot: $\max\{cx : Ax \leq b, x \in \mathbb{Z}^n\}$. Legyen $P = \{x \in \mathbb{R}^n : Ax \leq b\}$.

Algoritmus

1. Oldjuk meg a $\max\{cx, x \in P\}$ feladatot. (Például szimplex algoritmussal)
2. Ha az x^* optimális megoldás egész, akkor kész vagyunk, hiszen az optimális megoldás P-ben van.
3. Ha x^* nem egész, akkor keressünk olyan $\alpha x \leq \beta$ egyenlőtlenséget, amire $\alpha x^* > \beta$, de $\alpha x \leq \beta\ \forall x \in P \cap \mathbb{Z}^n$
4. Legyen $P := P \cap \{x : \alpha x \leq \beta\}$, és lépjünk az 1. pontra.

Az algoritmus működése:



Olyan vágósíkokat keresünk, amelyek segítségével szűkíthetjük a keresési teret, de nem vágunk ki egész pontokat.

7.3. Megjegyzés. A fenti algoritmus ebben a formában nem véges.

Hogyan találjuk meg az algoritmus 3. pontbeli egyenlőtlenségét?

Gomory-vágás

Legyen a feladat $\max\{cx, Ax = b, x \geq 0, x \in \mathbb{Z}^n\}$.

Tudjuk, hogy $\max\{cx, Ax = b, x \geq 0\}$ megoldható a szimplex módszerrel. Kapjuk a B optimális bázist. Használjuk a korábbi jelölést: $\bar{x}$, $\bar{A}$, $\bar{b}$ ($\bar{A} = B^{-1}A$, $\bar{b} = B^{-1}b$).

Tegyük fel, hogy $\bar{x}$ nem egész: $x_q \notin \mathbb{Z}$ ($q \in B$). Tudjuk, hogy $\bar{x}_q = \bar{b}_r$ (ahol x_q az r-edik sorhoz tartozó bázisváltozó).

Az r-edik sor egyenlete:
$x_q + \sum_{j \in N} \bar{a}_{rj} x_j = \bar{b}_r$ - érvényes egyenlet minden megoldásra.

$x_q + \sum_{j \in N} \lfloor \bar{a}_{rj} \rfloor x_j \leq \bar{b}_r$ - érvényes egyenlőtlenség minden megoldásra.

$x_q + \sum_{j \in N} \lfloor \bar{a}_{rj} \rfloor x_j \leq \lfloor \bar{b}_r \rfloor$ - érvényes egyenlőtlenség minden egész megoldásra, de $\bar{x}$ megoldásra már nem igaz, mivel $\bar{x}_q = \bar{b}_r > \lfloor \bar{b}_r \rfloor$ és $\bar{x}_j = 0\ \forall j \in N$. Tehát kaptunk egy olyan egyenlőtlenséget, ami minden egész megoldásra érvényes, de a kapott megoldásra nem.

7.4. Megjegyzés. Megfelelő megvalósítás esetén a Gomory vágásokkal véges algoritmust kapunk.

7.5. Megjegyzés. A vágósíkos algoritmusok lehetnek nagyon lassúak is, ha túl "kicsiket" vágunk. Csak olyan vágásokat kellene alkalmazni, amik "jók". Például ha ez a vágás az egész pontok konvex burkának egy lapja. Az egész poliédert viszont általában nem ismerjük, így a lapjait sem tudjuk meghatározni. Vannak viszont olyan esetek, amikor ismerünk néhány lapot.

Tehát ha ismerünk néhány lapot, akkor amint az LP-re optimális megoldást találunk, akkor megnézzük, hogy van-e a lapok között olyan, amit a talált optimum nem teljesít. Ha találunk ilyet, akkor ezt a lapot hozzávesszük a feladathoz, és ismételjük az eljárást. Ezáltal biztosak lehetünk abban, hogy "jól" vágunk.

A fenti módszert az utazóügynök feladat esetében gyakran használják.

Utazóügynök feladat

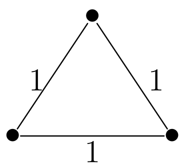
Legyen V ponthalmaz, $|V| = n$. Legyen továbbá $c \in \mathbb{R}_+^{\binom{n}{2}}$ súlyfüggvény az összes lehetséges élen. Célunk a minimális költségű Hamilton-kör keresése.

Legyen $P_I = \{$Hamilton körök karakterisztikus vektorainak a konvex burka$\}$, $P_I \subseteq \mathbb{R}^{\binom{n}{2}}$

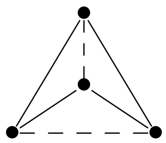
7.6. Megjegyzés. A poliéder nem függ a súlyfüggvénytől, minden n-re egy poliéderünk van.

Például

$n = 3$: egyetlen Hamilton-kör van.



$n = 4$: három Hamilton-kör van



$n \geq 3$ esetében: $d_x(v) = 2$ - egy olyan hipersíkot határoz meg, ami tartalmazza a poliédert.

Legyen tetszőleges $Z \subseteq V : 2 \leq |Z| \leq |V| - 1$. $d_x(Z) \geq 2$ - a poliéder egy lapját határozza meg.

Alkalmazzuk a vágósíkos eljárást a következőképpen: nézzük meg minél több ismert lap esetében, hogy a megoldásunk teljesíti-e a feltételeket. Amelyik lapra nem teljesíti, azt vegyük hozzá a rendszerhez.

7.7. Megjegyzés. Előfordulhat olyan eset, hogy egy nem egész megoldás az összes $d_x(Z) \geq 2$ alakú feltételt teljesíti. Ekkor Gomory vágást, vagy korlátozás és szétválasztást alkalmazhatunk (lásd később).

7.3. Dinamikus programozási algoritmusok

7.3.1. Bináris hátizsákfeladat

Legyen $k \in \{1, \dots, n\}$, $d \in \{0, \dots, b\}$. Definiáljuk a következő függvényt:

$$f_k(d) = \max\{\sum_{j=1}^{k} c_j x_j : \sum_{j=1}^{k} a_j x_j \leq d, x \in \{0,1\}^k\}.$$

Látjuk, hogy $f_n(b)$ lesz az eredeti bináris hátizsák feladat optimum értéke.

7.1. Jelölés. Legyen $f_0(d) = 0$ minden $d \in \{0, \dots, b\}$.

Számoljuk ki $f_k(d)$-t rekurzívan:

$$f_k(d) = \begin{cases} f_{k-1}(d), \text{ ha } a_k > d \\ \max\{f_{k-1}(d), f_{k-1}(d - a_k) + c_k\}, \text{ ha } a_k \leq d \end{cases}$$

Algoritmus

```
for k = 1...n
    for  d = 0...b
    do { számoljuk ki f_k(d) -t a fenti képlettel}
```

Az algoritmus lépésszáma: $\mathcal{O}(nb)$.

Az optimális megoldást a következőképpen kapjuk:

$$x_n := \begin{cases} 0, \text{ ha } f_n(b) = f_{n-1}(b) & (1) \\ 1, \text{ ha } f_n(b) = f_{n-1}(b - a_n) + c_n & (2) \end{cases}$$

$x_{n-1}, \dots, x_1$-et pedig esetszétválasztással kapjuk:

- Ha az (1) eset következik be, akkor

$$x_{n-1} := \begin{cases} 0, \text{ ha } f_{n-1}(b) = f_{n-2}(b) \\ 1, \text{ ha } f_{n-1}(b) = f_{n-2}(b - a_{n-1}) + c_{n-1} \end{cases}$$

- Ha a (2) eset következik be, akkor

$$x_{n-1} := \begin{cases} 0, \text{ ha } f_{n-1}(b - a_n) = f_{n-2}(b - a_n) \\ 1, \text{ ha } f_{n-1}(b - a_n) = f_{n-2}(b - a_n - a_{n-1}) + c_{n-1} \end{cases}$$

Így folytatjuk, amíg x_1-et el nem érjük.

7.3.2. Nemnegatív mátrixú egészértékű feladat

Most a fentihez hasonló dinamikus programozási algoritmust adunk több egyenlőtlenség esetén. Legyen a feladatunk

$$\max\{cx : \ Ax \leq b, x \geq 0, x \in \mathbb{Z}^n\},$$

ahol A, b, c nemnegatív és egész.

Adott $k \in \{1, \dots, n\}$ szám és $0 \leq d \leq b$ egész vektor esetén legyen

$$f_k(d) = \max\{\sum_{j=1}^{k} c_j x_j : \sum_{j=1}^{k} a_{ij} x_j \leq d_i \ (i = 1, \dots, m), \ x \geq 0\}.$$

Definiáljuk ezen kívül $f_0(d)$-t 0-nak minden $0 \leq d \leq b$-re. A következő rekurzióval lehet $k \geq 1$-re és $0 \leq d \leq b$-re kiszámolni a függvényértékeket:

$$f_k(d) = \begin{cases} f_{k-1}(d) & \text{ha valamilyen } i\text{-re } a_{ik} > d_i, \\ \max\{f_{k-1}(d), f_k(d - a_{.k}) + c_k\} & \text{ha } a_{.k} \leq d. \end{cases}$$

Az optimumértéket $f_n(b)$ adja meg. A lépésszám $O(n \prod_{i=1}^{m}(b_i + 1))$, ami sajnos a legtöbb feladatnál nagyon lassú algoritmust ad.

7.4. Korlátozás és szétválasztás

Tekintsük a következő alakú feladatot:

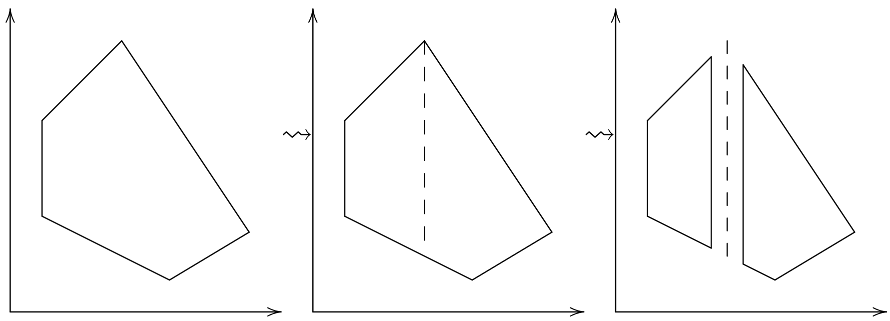
$$\max cx \tag{1}$$
$$Ax \leq b \tag{2}$$
$$x \text{ egész}, \tag{3}$$

ahol A, b, c egészek. A *korlátozás és szétválasztás* módszerének alapgondolata az, hogy a feladatot szétválasztjuk részfeladatokra úgy, hogy az eredeti feladat megengedett megoldásainak halmaza a részfeladatok megengedett megoldás-halmazainak diszjunkt uniója legyen. Ekkor az eredeti feladat optimum-értéke megegyezik a részfeladatok optimum-értékeinek maximumával. Az egyes részfeladatokat további részfeladatokra lehet szétbontani, így a vizsgált feladatok egy fát alkotnak, aminek gyökerében az eredeti feladat található.

A részfeladatok relevanciájának vizsgálatához van szükség a *korlátozásra.* A korlátozás azt jelenti, hogy minden részfeladatnál felső (és esetleg alsó) korlátot számolunk az optimum értékére. Amennyiben egy részfeladatra vonatkozó felső korlát kisebb mint egy már ismert alsó korlát, akkor azzal a részfeladattal nem kell tovább foglalkozni, törölhetjük a fából.

LP alapú korlátozás és szétválasztásról akkor beszélünk, ha a felső korlátokat az LP relaxált optimuma adja, a szétválasztás pedig lineáris egyenlőtlenségek hozzáadásával történik.

Az algoritmus működése:



A részfeladatokra osztás során kihagyunk részeket a feladat teréből, de csak olyanokat, amikben nincsenek egész pontok.

A továbbiakban a legegyszerűbb LP alapú általános algoritmust írjuk le.

Legyen $P = \{x \in \mathbb{R}^n : \; Ax \leq b\}$. Feltesszük, hogy P korlátos poliéder. A részfeladataink mindig

$$\max\{cx : \; x \in P' \cap \mathbb{Z}^n\}$$

alakúak lesznek, ahol $P' \subseteq P$ poliéder. Jelölések:

- $u(P') = \max\{cx : \; x \in P'\}$. Ez felső korlát a részfeladat optimumértékére.
- $x^*(P')$ a $\max\{cx : \; x \in P'\}$ feladat egy optimális bázismegoldása. Ez kiszámolható szimplex módszerrel.
- x^*: az eddig talált legjobb egész megoldás
- L: az eddig talált legjobb egész megoldás célfüggvényértéke, illetve $-\infty$ ha még nem találtunk egész megoldást.

- $\mathcal{F}$: A feldolgozandó poliéderek listája.

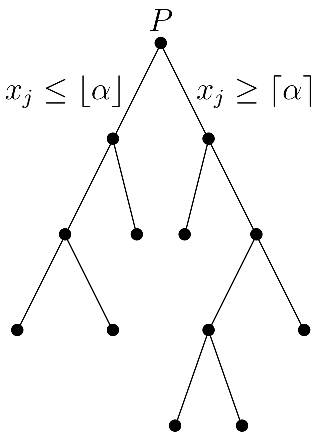
Az algoritmus inicializálása: $\mathcal{F} = \{P\}$, és $L = -\infty$. Egy általános lépés a következő:

1. Ha $\mathcal{F}$ üres: kész vagyunk. Különben az $\mathcal{F}$ listáról kiválasztunk egy P' feladatot, és kiszámoljuk az $u(P')$ értéket.
2. Ha $u(P')$ nem létezik, mert az LP relaxált nem megoldható: töröljük P'-t $\mathcal{F}$-ből.
3. Ha $u(P') \leq L$: töröljük P'-t $\mathcal{F}$-ből.
4. Ha $u(P') > L$ és $x^*(P')$ egész: $x^* := x^*(P')$, és $L := u(P')$. Töröljük P'-t $\mathcal{F}$-ből.
5. Ha $u(P') > L$ és $x^*(P')$ nem egész: válasszunk egy olyan j indexet, amire $x^*(P')$ j-edik komponense egy nem egész α szám. Legyen $P'_1 = P' \cap \{x \in \mathbb{R}^n : \ x_j \leq \lfloor \alpha \rfloor\}$, és $P'_2 = P' \cap \{x \in \mathbb{R}^n : \ x_j \geq \lceil \alpha \rceil\}$. Töröljük P'-t $\mathcal{F}$-ből, és adjuk hozzá P'_1-et és P'_2-t $\mathcal{F}$-hez.

7.1. Állítás. *Az algoritmus véges sok lépésben talál egy optimális megoldást.*

Bizonyítás. Mivel P korlátos, létezik olyan N szám, hogy tetszőleges $x \in P$-re és $j \in \{1, \dots, n\}$-re $|x_j| \leq N$.

Az algoritmus futása alapján felépíthetünk egy gyökeres bináris fát, amelynek minden csúcsa egy feldolgozott poliédernek felel meg (a gyökér P-nek) és az egyes elágazások a szétválasztásoknak.



Tekintsük a fa egy tetszőleges ágát! Egy adott x_j változó szerint ezen az ágon legfeljebb $2N$ szétválasztás lehet, mert minden szétválasztásnál egy 1 hosszú nyílt intervallum kimarad az x_j lehetséges értékei közül. Így minden ág legfeljebb $2nN$ hosszú, amiből következik, hogy az algoritmus véges sok lépésben véget ér.

Az optimális megoldás P-ben van, és ha az algoritmus egy adott szétválasztási lépésénél P'-ben ott van, akkor vagy P'_1-ben vagy P'_2-ben is ott van. Tehát a fenti fa valamelyik leveléhez tartozó poliéderben van az optimális megoldás. Mivel egy levélben nem választunk szét, ezért ott meg is találjuk. Tehát az algoritmus mindenképpen megtalálja az optimális megoldást. $\square$

Korlátozás és szétválasztás algoritmusa a bináris hátizsák feladatra

Ha bináris hátizsákfeladatra futtatjuk a fenti algoritmust, az algoritmusban szereplő poliéderek mindig a következő alakúak lesznek:

$$P(J_0, J_1) = \{0 \leq x \leq 1 : \sum_{j=1}^{n} a_j x_j \leq b, x_j = 0, \text{ ha } j \in J_0, \ x_j = 1, \text{ ha } j \in J_1\}.$$

A $\max\{cx : x \in P(J_0, J_1)\}$ feladat nagyon egyszerűen megoldható a mohó algoritmussal:

Rakjuk a nem $(J_0 \cup J_1)$-beli változókat $\frac{c_j}{a_j}$ szerint csökkenő sorrendbe, majd sorba növeljük meg őket amennyire lehet.

Az elején $d = b - \sum_{j \in J_1} a_j$ - ennyi még belefér a hátizsákba.

Ha j_1 az első index a fenti sorrendben, akkor legyen $x_{j_1} = \min\{1, \frac{d}{a_{j_1}}\}$, és legyen $d := d - a_{j_1} x_{j_1}$. Majd vegyük a következőt a sorrend szerint, és ugyanígy folytassuk.

7.8. Megjegyzés. Ha egy változó értéke nem 1, hanem $\frac{d}{a_j}$, akkor az utána következő összes többi változó értéke 0 lesz, hiszen a d kifogyott. Tehát legfeljebb egyetlen nem egész érték lesz. Ezért mindig egyértelmű, melyik változó szerint kell szétválasztani.

Az algoritmust kiegészíthetjük azzal, hogy ha egy adott lépésben az x_j változó szerint választunk szét, akkor az optimális tört megoldásban x_j-t 0-ra módosítva egy megengedett egész megoldást kapunk. Ha ez jobb, mint az eddig talált legjobb x^* megoldás, akkor lecserélhetjük x^*-ot erre.

7.5. Közelítő algoritmusok

7.5.1. Minimális lefogó csúcshalmaz

Adott $G = (V, E)$ irányítatlan gráf.

7.2. Definíció (Lefogó csúcshalmaz)**.** A lefogó csúcshalmaz minden élnek legalább egyik végpontját tartalmazza.

A feladat: keressünk minimális méretű lefogó csúcshalmazt.

7.9. Megjegyzés. A fenti feladat NP-teljes.

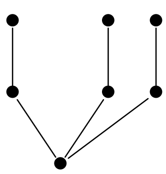
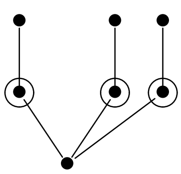
Két algoritmust ismertetünk a lefogó csúcshalmaz keresésére, amelyek nem feltétlenül szolgáltatnak optimális megoldást:

1. Algoritmus (mohó)

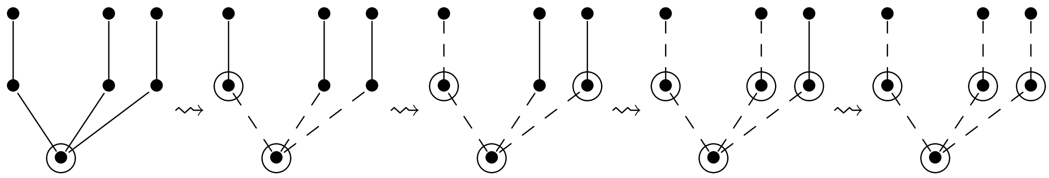
Kezdetben legyen $U = \emptyset$.

1. Válasszunk maximális fokú csúcsot a gráfban. Legyen ez v.
2. Legyen $U := U \cup \{v\}$
3. A G gráfból hagyjuk ki ezt a csúcsot, és az összes rá illeszkedő élt.
4. Ha nem marad él, akkor kész vagyunk. U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

7.10. Megjegyzés. A fenti algoritmus nem feltétlenül talál optimális megoldást.

Például:  optimális megoldása: 

Az algoritmusunk viszont a következőket találja:

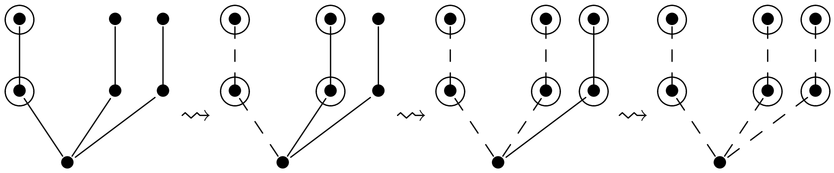


2. Algoritmus

Kezdetben legyen $U = \emptyset$.

1. Válasszunk egy tetszőleges uv élt a gráfban.
2. Legyen $U := U \cup \{u\} \cup \{v\}$
3. A G gráfból hagyjuk ki ezt a két csúcsot, és az összes rájuk illeszkedő élt.
4. Ha nem marad él, akkor kész vagyunk. U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

7.11. Megjegyzés. A fenti algoritmus nem tűnik túl jó megoldásnak, mivel feleslegesen két csúcsot is bevesz lépésenként a lefogó csúcshalmazba. Az előbbi példára rosszabban teljesít, mint a mohó algoritmus:



Hasonlítsuk össze a két algoritmust. Ehhez bevezetjük az α - közelítés fogalmát:

7.3. Definíció. Egy minimalizálási feladatra egy algoritmus α-közelítő, ha tetszőleges inputra az output értéke legfeljebb α-szorosa az optimálisnak.

7.12. Megjegyzés. Az 1-közelítő algoritmus optimális megoldást ad.

7.2. Állítás. *A 2. algoritmus 2-közelítő algoritmus.*

Bizonyítás. Legyen F a kiválasztott élek halmaza. F diszjunkt élekből áll. Ekkor $|U_{opt}| \geq |F|$, mivel az optimális lefogó csúcshalmaz lefogja F éleit.

Másrészt: $|U| = 2|F|$. Tehát $|U_{opt}| \geq \frac{1}{2}|U|$. □

7.3. Állítás. *Az 1. algoritmus nem 2-közelítő algoritmus.*

Bizonyítás. Definiálunk egy $G = (A, B; E)$ páros gráfot. Az A osztályban 60 csúcs van, mind 5 fokú. A B osztályban 12 db 5 fokú, 15 db 4 fokú, 20 db 3 fokú, 30 db 2 fokú, és 60 db 1 fokú. Minden $i \in \{1, \ldots, 5\}$-re a B-beli i fokúak szomszédságai A egy partícióját adják (hogy pontosan melyik partícióját, az mindegy).

A mohó algoritmus lefuthat úgy, hogy a teljes B halmazt választja ki lefogó csúcshalmaznak. Másrészt A is lefogó csúcshalmaz, tehát mivel $|B| > 2|A|$, ez nem 2-közelítő algoritmus. □

7.13. Megjegyzés. Az is belátható, hogy a mohó algoritmus semmilyen α-ra nem lesz α-közelítő.

7.14. Megjegyzés. Nem ismert $\alpha < 2$-re α-közelítő polinomiális futási idejű algoritmus erre a feladatra.

7.15. Megjegyzés. A 2. algoritmusnál van jobb is: például ha a végén a felesleges csúcsokat töröljük, de ettől függetlenül 2-közelítő algoritmusnál nem ismert jobb a minimális lefogó csúcshalmaz feladatra.

7.2. Tétel. *Ha $\alpha < \frac{7}{6}$-ra van α-közelítő polinomiális algoritmus a minimális lefogó csúcshalmaz feladatra, akkor $P = NP$.*

7.5.2. Minimális költségű lefogó csúcshalmaz

Legyen $G = (V, E)$ irányítatlan gráf, $c \in \mathbb{R}_+^V$ költségfüggvény. Keressük azt az U lefogó csúcshalmazt, amire $\sum_{u \in U} c_u$ minimális.

Erre a feladatra is adunk 2-közelítő algoritmust. Írjuk fel a feladatot egészértékű programozási feladatként:

$$\begin{aligned} x &\in \{0,1\}^V \\ x_u + x_v &\geq 1, \forall uv \in E \\ \min\ & cx \end{aligned}$$

A feladat LP-relaxáltja:

$$\begin{aligned} x &\in \mathbb{R}_+^V \\ x_u + x_v &\geq 1, \forall uv \in E \\ \min\ & cx \end{aligned}$$

7.16. Megjegyzés. Nincs értelme 1-nél nagyobb változót venni, mivel az lecsökkenthető 1-re.

Nézzük az LP-relaxált duálisát:

$$\begin{aligned} y &\in \mathbb{R}^E \\ y &\geq 0 \\ \sum_{u:uv \in E} y_{uv} &\leq c_v, \forall v \in V \\ \max & \sum_{e \in E} y_e \end{aligned}$$

Primál-duál algoritmus

Minden lépésben lesz egy $U \subseteq V$ és lesz egy y duális megoldás. Kezdetben legyen $U = \emptyset$, $y \equiv 0$, $E' = E$.

1. Válasszunk tetszőleges $uv \in E'$ élt.
2. Emeljük meg y_{uv}-t annyira, hogy u-nál vagy v-nél a duál egyenlőtlenség teljesüljön egyenlőséggel - előfordulhat, hogy az egyenlőség mindkettőre teljesül.
3. U-hoz vegyük hozzá u és v közül azt, amelyiknél egyenlőség van - ha mindkettőnél egyenlőség van, akkor mindkettőt vegyük hozzá. Töröljük E'-ből a lefogott éleket.
4. Ha $E' = \emptyset$, akkor kész vagyunk. U a lefogó csúcshalmaz. Különben lépjünk az 1. pontra.

7.17. Megjegyzés. Az E'-re azért van szükség, mert a duális feladatot nem akarjuk megváltoztatni.

7.18. Megjegyzés. Az y_{uv} emelése csak az u, v egyenlőtlenségeknél számít, a többit nem módosítja.

7.4. Állítás. *A fenti algoritmus 2-közelítő.*

Bizonyítás. Legyen $\pi \in \mathbb{R}_+^V$ segédváltozónk. Kezdetben legyen $\pi \equiv 0$. Amikor y_{uv}-t megemeljük δ-val, akkor $\pi(u)$-t és $\pi(v)$-t is megemeljük δ-val.

A végén $\sum_{v \in V} \pi(v) = 2 \sum_{e \in E} y_e$. Másrészt: amikor v bekerül U-ba, akkor $\pi(v) = c(v)$, és ez később sem változik.

Tehát $\sum_{u \in U} c_u \leq \sum_{v \in V} \pi(v)$, amiből

$$\sum_{u \in U} c_u \leq 2 \sum_{e \in E} y_e \underbrace{\leq}_{\text{gyenge dualitás}} \leq 2OPT_{LP} \leq 2OPT_{IP}$$

□

7.19. Megjegyzés. Többet bizonyítottunk az állításnál: az LP-relaxáltnak legfeljebb kétszerese a kapott megoldás.

8. Játékelmélet

A játékelméleten belül számos különböző megközelítés létezik a „játék” definíciójára, aszerint hogy mit feltételezünk a játékosok tudásáról és viselkedéséről. Ezen az előadáson az úgynevezett stratégiai alakú játékok elméletébe nyújtunk bepillantást. Ezekben a játékokban véges sok játékos egymástól függetlenül választ stratégiát a rendelkezésre álló stratégiák halmazából. Ismerik a többiek lehetőségeit, de nem ismerik a többiek döntését. A játékosok hasznát az összes játékos döntése együtt határozza meg.

A továbbiakban példákon keresztül mutatjuk be az alapfogalmakat.

8.1. Fogoly-dilemma

Egy bűnténynek két gyanusítottja van, de nem tudjuk eldönteni, hogy ki volt az, nekik kell vallani. Ha az egyik gyanúsított a másikra vall, akkor az előbbi 1 évre kerül börtönbe, a másik viszont 4 évre. Ha egyikük sem vall, akkor mindkettőjüket 2 évre börtönözik be, ha mindkettejük vallanak, akkor mindkettőjük 3 év börtönbüntetést kap.

Az előbbi feladat felírható hasznossági mátrix-szal a következőképpen:

1. \ 2.	vall	nem vall
vall	-3 -3	-4 -1
nem vall	-1 -4	-2 -2

Nézzük meg, mi lenne a jó stratégia a játékosok számára?

Az első játékosnak mindenképpen jobb, ha vall, bármit is választ a második játékos. Ugyanez mondható el a második játékosról is. Ezt nevezzük domináns stratégiának.

Ha mindkét játékos önző, akkor mindketten vallani fognak, ami az ő szempontjukból a legjobbnak tűnik.

De a domináns stratégia rosszabb mindkettejüknek összehasonlítva azzal a stratégiával, hogy egyikük sem vall.

8.2. Szennyezési játék

Van n ország, és mindegyik országnak el kell döntenie, hogy szennyezi vagy nem szennyezi a környezetét. A nem-szennyezés drága, mivel 3-al megnöveli az ország költségét, a szennyezés viszont mindenkit érint - 1-el növeli minden más ország költségét. Az i. ország költségét felírhatjuk tehát így:

$$i.\ \text{ország költsége} = \begin{cases} 3 + \text{szennyező országok száma} & \text{, ha nem szennyez} \\ \text{szennyező országok száma} & \text{, ha szennyez} \end{cases}$$

Van domináns stratégia?

Tegyük fel, hogy az ország, aminek a stratégiáját el akarjuk dönteni tudja, hogy mit választott a többi ország. Ekkor ha az ország nem szennyez, akkor 3-al növelné a költségét, ha viszont szennyezne, akkor csak 1-el növelné a költségét.

A domináns stratégia tehát az, hogy mindenki szennyez. Ha háromnál több ország vesz részt a játékban, akkor a domináns stratégia rosszabb annál az esetnél, ahol egyik ország sem szennyez.

8.1. Jelölés. Legyen n a játékosok száma. Továbbá legyen S_i az i. játékos stratégiáinak halmaza, az $s = (s_1, s_2, \ldots, s_n)$ vektor pedig a stratégia-vektor, ahol $s_i \in S_i$ $(i = 1, \ldots, n)$.

Legyen $S = S_1 \text{ x } S_2 \text{ x } \ldots \text{ x } S_n$ a stratégiavektorok halmaza.

Ha $s \in S$, akkor legyen s_{-i} az i. játékos kivételével a többiek stratégiája.

Legyen $u_i : S \to \mathbb{R}$ az i. játékos hasznossági függvénye.

8.1. Megjegyzés. S_i lehet végtelen halmaz is.

8.1. Definíció (Domináns stratégia)**.** $s \in S$ domináns stratégia, ha $\forall i \in \{1, 2, \ldots, n\}$ és $\forall s' \in S : u_i(s_i, s'_{-i}) \geq u_i(s')$, ahol $u_i(s_i, s'_{-i})$ jelöli azt, hogy az i. játékos az s_i stratégiát választotta, mindenki más pedig az s' szerinti stratégiát.

8.2. Definíció (Szigorúan domináns stratégia)**.** $s \in S$ szigorúan domináns stratégia, ha $\forall i \in \{1, 2, \ldots, n\}$ és $\forall s' \in S, s'_i \neq s_i : u_i(s_i, s'_{-i}) > u_i(s')$.

8.2. Megjegyzés. A fogoly-dilemma és a szennyezési játék esetében is volt szigorúan domináns stratégia.

8.3. Megjegyzés. Domináns stratégia akár több is lehet egy játéknál (ha például a hasznossági mátrixban mindenütt ugyanaz az érték van, akkor mindegy, ki hogyan választ), de szigorúan domináns csak egy lehet, ha egyáltalán van ilyen.

8.3. Vickrey árverés

Az n játékos egy adott tárgyra licitálhat. Az i-edik játékosnak v_i-t ér a tárgy. A licitálás abból áll, hogy mindenki mond egymástól függetlenül egy nemnegatív valós összeget. Formálisan, $S_i = \mathbb{R}_+$ minden i-re. A legnagyobb összeget licitáló kapja meg a tárgyat (döntetlen esetén a kisebb indexű játékos), de csak a második legnagyobb összeget kell kifizetnie.

8.1. Állítás. *Az $s_i = v_i$ $(i = 1, \ldots, n)$ domináns stratégia, azaz a Vickrey árverésen mindenkinek érdemes a valódi értéket licitálnia.*

Bizonyítás. Azt kell belátnunk, hogy minden $s' \in S$-re és minden i-re $u_i(s_i, s'_{-i}) \geq u_i(s')$. Kéte esetet különböztetünk meg.

- Ha $\max\{s'_j : j \neq i\} \geq s_i$, akkor $u_i(s_i, s'_{-i}) = 0$ és $u_i(s') \leq 0$.
- Ha $\max\{s'_j : j \neq i\} < s_i = v_i$, akkor $u_i(s_i, s'_{-i}) = v_i - \max\{s'_j : j \neq i\} > 0$ és $u_i(s') \in \{0, v_i - \max\{s'_j : j \neq i\}\}$.

Tehát mindkét esetben teljesül a dominancia feltétele. □

8.4. Közös ló játék

Van n játékosunk, akik ugyanazt a közös erőforrást szeretnék használni. Mindegyik játékos eldöntheti, hogy mennyire szeretné kihasználni az erőforrást, de az erőforrás annál kevésbé hatékony minél többet használják.

A játékosoknak végtelen sok stratégiájuk van: $S_i = [0,1]$, $i \in \{1,\dots,n\}$.

$$u_i(s) = \begin{cases} 0, \text{ ha } \sum_{j=1}^n s_j \geq 1 \\ s_i(1-\sum_{j=1}^n s_j) \text{ egyébként} \end{cases}$$

Hogyan válasszunk stratégiát az i. játékosnak, ha a többi játékos stratégiája rögzített?

Tegyük fel, hogy $\sum_{i\neq j} s_j = t \leq 1$. Ekkor $u_i(s) = s_i(1-t-s_i)$. Ennek a maximuma: $\frac{1-t}{2}$

8.4. Megjegyzés. A fenti feladatban nincs domináns stratégia, mivel az i. játékos stratégiája függ a többitől, nem tudja úgy megválasztani a stratégiáját, hogy bármit választanak a többiek, neki ez jó legyen. Ha például az i. játékos az s_i stratégiát választja, akkor tudunk olyan s' stratégiát mondani, amire $u_i(s_i, s'_{-i}) < u_i(s')$.

Beszélhetünk viszont egyensúlyi megoldásról: ilyenkor egyik játékosnak sem érdemes változtatnia a jelenlegi stratégiáján.

Ha van olyan s, ahol $s_i = \frac{1-\sum_{i\neq j} s_j}{2}$ minden i-re, akkor ez egyensúlyi helyzet, mert senki sem akarna változtatni a stratégiáján.

Felírható tehát a következő egyenlőtlenség-rendszer: $s_i + \sum_{j=1}^n s_j = 1$ $\forall i$. Összeadva őket kapjuk: $(n+1)\sum_{j=1}^n s_j = n$, tehát $\sum_{j=1}^n s_j = \frac{n}{n+1}$. Visszahelyettesítve, az egyensúlyi megoldás $s_i = \frac{1}{n+1}$, $\forall i = (1,\dots,n)$.

De tudunk ennél jobb megoldást is. Az előbbi stratégia esetében egy játékos hasznossága $u_i(s) = \frac{1}{n+1}(1-\frac{n}{n+1}) = \frac{1}{(n+1)^2}$.

Legyen $s' = (\frac{1}{2n}, \frac{1}{2n}, \dots, \frac{1}{2n})$. Ennek a hasznossága: $u_i(s') = \frac{1}{2n}(1-\frac{1}{2}) = \frac{1}{4n}$, ami nem egyensúlyi, de jobb, mint az egyensúlyi.

8.3. Definíció (Tiszta Nash-egyensúly)**.** $s \in S$ tiszta Nash-egyensúly, ha $\forall i$, $\forall s'_i \in S_i : u_i(s) \geq u_i(s'_i, s_{-i})$.

8.2. Állítás. *Minden domináns stratégia Nash-egyensúly.*

Bizonyítás. Az állítás a definicióból triviálisan bizonyítható. □

8.3. Állítás. *Ha van szigorúan domináns stratégia, akkor nincs más Nash-egyensúly.*

Bizonyítás. Legyen s szigorúan domináns stratégia. Tegyük fel, hogy s' Nash-egyensúly. Ekkor ha $s_k \neq s'_k$, akkor $u_k(s_k, s'_{-k}) > u_k(s')$, mert s szigorúan domináns, de ugyanakkor s' Nash-egyensúly, emiatt $u_k(s_k, s'_{-k}) \leq u_k(s')$. Ez ellentmondás, így csak az lehetséges, hogy $s = s'$. □

8.5. Fej vagy írás játék

Két játékos a következőképpen játszik: az első játékosnál van egy érme, azt elrejti fejjel vagy írással. A második játékosnak ki kell találnia, hogy az első játékos fejet vagy írást választott. Ha a második játékos eltalálja, akkor 1 pontot kap, és az első játékos 1 pontot veszít, ha nem találja el, akkor fordítva - az első játékos kap 1 pontot, a második pedig veszít egyet. A hasznossági mátrix:

1. \ 2.	fej	írás
fej	1 -1	-1 1
írás	-1 1	1 -1

Nincs tiszta Nash-egyensúly, mert

- Ha $s_1 = s_2$, akkor az első játékos jobban jár, ha változtat
- Ha $s_1 \neq s_2$, akkor a második játékos jár jobban, ha változtat.

8.5. Megjegyzés. Bizonyos értelemben lehet Nash-egyensúlyról beszélni - az első játékos $50-50\%$ valószínűséggel választja a fej- vagy írást, a második játékos is $50-50\%$ valószínűséggel tippel fej- vagy írásra. Ezt nevezzük kevert stratégiának

8.4. Definíció (Kevert stratégia). m_i kevert stratégiája az i. játékosnak, ha valószínűségi eloszlás az S_i halmazon.

8.5. Definíció (Kevert stratégia-vektor). $m = (m_1, \ldots, m_n)$ kevert stratégia-vektor, ha m_i kevert stratégiája az i-edik játékosnak $(i = 1, \ldots, n)$.

8.6. Definíció (Kevert hasznosság). $u_i(m) = u_i(s)$ várható értéke az m valószínűségi eloszlás szerint.

8.7. Definíció (Kevert Nash-egyensúly). m kevert Nash-egyensúly, ha $\forall i$ és tetszőleges m_i' kevert stratégiájára i-nek:$u_i(m) \geq u_i(m_i', m_{-i})$.

8.1. Tétel (Nash, 1951). *Véges sok játékos és véges stratégia-halmaz esetén mindig van kevert Nash-egyensúly.*

Nézzünk példát olyan játékra, ahol nincs kevert Nash-egyensúly:

8.6. Árazási játék

Van két eladó, akik ugyanazt a terméket árulják, de meghatározhatják a termék árát. $s_1 \in [0, 1]$, $s_2 \in [0, 1]$. Továbbá van három vevő:A, B, C, akik a következőképpen döntenek:

- A csak az első eladótól hajlandó vásárolni
- B az olcsóbb árú terméket vásárolja meg (ha a két eladónál ugyanaz az ár, akkor az elsőtől vásárol)
- C csak a második eladótól hajlandó vásárolni.

8.4. Állítás. *A fenti játéknál nincs kevert Nash-egyensúly.*

Indoklás (csak arra, hogy nincs tiszta Nash-egyensúly)

Ha például az első eladó $\frac{1}{2}$-nél nagyobb áron ad el, akkor a második eladónak elég egy olyan árat meghatározni, ami ugyancsak nagyobb $\frac{1}{2}$-nél, de kicsit kisebb az első eladó áránál. De ekkor az első is a második eladó ára alá mehet úgy, hogy $\frac{1}{2}$-nél még mindig nagyobb ára legyen, stb. Ha valamelyik eladó $\frac{1}{2}$ alá megy az árral, akkor a másik 1-re meghatározva az árat, jobban jár, mint az első.

A továbbiakban olyan játékokkal foglalkozunk, ahol minden játékosnak véges sok stratégiája van.

8.8. Definíció (Kevert stratégia)**.** Legyen $m_i : S_i \to [0,1]$ az i. játékos kevert stratégiája, amire $\sum_{x_i \in S_i} m_i(x_i) = 1$.

8.2. Jelölés. Legyen az i. játékos kevert stratégiáinak halmaza: M_i. Továbbá jelöljük $m = (m_1, m_2, \ldots, m_n)$-el a kevert stratégiavektort, ahol $m_i \in M_i$, $i = (1, \ldots, m)$.

Jelöljük M-mel a stratégiavektorok halmazát.

8.9. Definíció. $s \in S$-re, $m \in M$-re legyen $m(s) = \prod_{i=1}^{n} m_i(s_i)$.

8.5. Állítás. $\sum_{s \in S} m(s) = 1$.

Bizonyítás. Játékosok számára vett indukcióval:

$$\sum_{s \in S} m(s) = \sum_{s_n \in S_n} \sum_{s_{-n} \in S_{-n}} m(s_n, s_{-n}) = \sum_{s_n \in S_n} m_n(s_n) \sum_{s_{-n} \in S_{-n}} m_{-n}(s_{-n})$$

$$\underbrace{=}_{\text{indukció}} \sum_{s_n \in S_n} m_n(s_n) \underbrace{=}_{\text{definició}} 1 \qquad \square$$

8.10. Definíció (Várható haszon)**.** Az i. játékos várható haszna $m \in M$ kevert stratégia esetén: $u_i(m) := \sum_{s \in S} m(s) u_i(s) = \sum_{s_i \in S_i} \sum_{s_{-i} \in S_{-i}} m(s) u_i(s) = \sum_{s_i \in S_i} m(s_i) \sum_{s_{-i} \in S_{-i}} m(s_{-i}) u_i(s) = \sum_{s_i \in S_i} m(s_i) u_i(s_i, m_{-i})$.

8.11. Definíció (0 összegű játék)**.** Tetszőleges $s \in S$-re $\sum_{i=1}^{n} u_i(s) = 0$. Ilyen például a fej vagy írás játék, de nem ilyen a fogolydilemma vagy a közös ló játék.

8.12. Definíció (Legjobb válasz)**.** $m \in M$-re m_i az i. játékos legjobb válasza m_{-i}-re, ha tetszőleges $m_i' \in M_i$ esetén: $u_i(m) \geq u_i(m_i', m_{-i})$

Példa. Nézzük a fej vagy írás játékot. Hasznossági mátrixa:

1. \ 2.	fej	írás
fej	1 -1	-1 1
írás	-1 1	1 -1

Legyen a kevert stratégia a következő:

$m_1(\text{fej}) = \frac{1}{3}$ $\quad m_1(\text{írás}) = \frac{2}{3}$
$m_2(\text{fej}) = \frac{3}{4}$ $\quad m_2(\text{írás}) = \frac{1}{4}$

$(m_1, m_2) \in M$ kevert stratégia-vektor. Összesen négy stratégia-vektorunk van:

$$S = \{(\text{fej}, \text{fej}), (\text{fej}, \text{írás}), (\text{írás}, \text{fej}), (\text{írás}, \text{írás})\}.$$

Ezek valószínűségei:

$$m(\text{fej}, \text{fej}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}$$
$$m(\text{fej}, \text{írás}) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$$
$$m(\text{írás}, \text{fej}) = \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}$$
$$m(\text{írás}, \text{írás}) = \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{6}$$

Ha ezeket összeadjuk, az összegük: 1. Most nézzük az 1. játékos várható hozamát:

$u_1(m) = \frac{1}{4} \cdot (-1) + \frac{1}{12} \cdot 1 + \frac{1}{2} \cdot 1 + \frac{1}{6} \cdot (-1) = \frac{1}{6}$.

A 2. játékos hozamát is hasonlóan kiszámíthatnánk, de figyeljük meg, hogy ebben a játékban minden játékos haszna ugyanaz mint a másik játékos vesztesége, azaz 0−összegű játékról beszélünk, ezért $u_2(m) = -u_1(m) = -\frac{1}{6}$.

Figyeljük meg, hogy a 2. játékos aktuális stratégiája nem a legjobb válasz az 1. játékos stratégiájára. Helyette ha mindig írást mond, akkor $\frac{2}{3}$ -ad valószínűséggel eltalálja, hogy mit választott az 1. játékos. □

8.3. Jelölés. Ha $s_i \in S_i$, $m_{-i} \in M_{-i}$, akkor jelölje $(s_i, m_{-i}) \in M$ azt a kevert stratégiát, ahol az i. játékos mindig az s_i stratégiát választja.

8.13. Definíció (Legjobb tiszta válasz)**.** Legyen $m_{-i} \in M_{-i}$ kevert stratégia az i-n kívül. Ekkor $s_i \in S_i$ az i. játékos legjobb tiszta válasza m_{-i}-re, ha tetszőleges $m'_i \in M_i$ kevert stratégia esetén: $u_i(s_i, m_{-i}) \geq u_i(m'_i, m_{-i})$.

8.6. Állítás. *Legyen $m_{-i} \in M_{-i}$. Ekkor $m_i \in M_i$ a legjobb válasz m_{-i}-re $\Leftrightarrow$ minden $s_i \in S_i$-re, amire $m_i(s_i) > 0$: s_i legjobb tiszta válasz m_{-i}-re.*

Bizonyítás. $\Leftarrow$: Legyen U az i. játékos legjobb válaszának haszna. Ekkor $u_i(m) = \sum_{s_i \in S_i} m_i(s_i) u(s_i, m_{-i}) = U$, mivel ha s_i a legjobb tiszta válasz, akkor $u(s_i, m_{-i}) = U$.

$\Rightarrow$: $u_i(m) = \sum_{s_i \in S_i} m_i(s_i) u(s_i, m_{-i}) \leq U$ és csak akkor lesz egyenlő U-val, ha $m_i(s_i) > 0$ esetén $u(s_i, m_{-i}) = U$. □

8.14. Definíció (Kevert Nash-egyensúly)**.** $m \in M$ kevert Nash-egyensúly, ha tetszőleges i-re m_i legjobb válasz m_{-i}-re.

8.2. Következmény. *$m \in M$ Nash-egyensúly $\Leftrightarrow$ tetszőleges i-re és tetszőleges $s_i \in S_i$ tiszta stratégiára, ha $m_i(s_i) > 0$, akkor s_i a legjobb tiszta válasz m_{-i}-re.*

8.7. Kő-papír-olló játék

Nézzük a hasznossági-mátrixot:

1. \ 2.	kő	papír	olló
kő	0 0	1 -1	-1 1
papír	-1 1	0 0	1 -1
olló	1 -1	-1 1	0 0

8.15. Definíció (Szimmetrikus játék)**.** Egy játék szimmetrikus, ha $S_1 = S_2 = \cdots = S_n$, és tetszőleges π permutációra és $s_1 \in S_1, \ldots, s_n \in S_n$-re $u_{\pi(i)}(s_1, \ldots, s_n) = u_i(s_{\pi(1)}, \ldots, s_{\pi(n)})$. Például a kő-papír-olló szimmetrikus, de a fej-vagy-írás játék nem szimmetrikus.

Keressünk Nash-egyensúlyt:

Tegyük fel, hogy $m = (m_1, m_2)$ kevert Nash-egyensúly. Mivel a játék szimmetrikus, ezért ugyanazok a hasznosságok, ha felcseréljük a játékosokat.

Tegyük fel továbbá, hogy m_1 nem azonosan $\frac{1}{3}$, és $m_1(\text{kő})$ a legkisebb valószínűség. Ekkor a 2. játékos számára a papír nem legjobb tiszta válasz és a következmény miatt: $m_2(\text{papír}) = 0$.

Emiatt az 1. játékos számára az olló nem legjobb tiszta válasz, és a következmény miatt: $m_1(\text{olló}) = 0$. Emiatt a 2. játékos számára a kő nem legjobb tiszta válasz, és a következmény miatt: $m_2(\text{kő}) = 0$. Emiatt az 1. játékos számára a papír nem legjobb tiszta válasz, és a következmény miatt: $m_1(\text{papír}) = 0$.

Azt kaptuk tehát, hogy az 1. játékosra $m_1(\text{kő})$ a legkisebb, de $m_1(\text{papír}) = 0$ és $m_1(\text{olló}) = 0$, ami ellentmondás.

Tehát a Nash-egyensúly: $m_1 \equiv \frac{1}{3}$, $m_2 \equiv \frac{1}{3}$. Ez tényleg Nash-egyensúly, mivel ha az 1. játékos $\frac{1}{3}$ valószínűséggel választ, akkor az 2. játékosnak minden stratégia legjobb válasz.

8.8. Nash-egyensúly 0-összegű kétszemélyes játékokra

8.16. Definíció (Legjobb nyilvános stratégia)**.** $m_i \in M_i$ kevert stratégia legjobb nyilvános stratégia az i. játékos számára, ha

$$\min_{m_{-i} \in M_{-i}} u_i(m_i, m_{-i}) = \max_{m'_i \in M_i} \min_{m_{-i} \in M_{-i}} u_i(m'_i, m_{-i}).$$

8.4. Jelölés. Legyen $|S_1| = k$, $|S_2| = l$. Ekkor $A \in \mathbb{R}^{k \times l}$ az 1. játékos hasznossági mátrixa.

Példa. Nézzük a következő 0-összegű kétszemélyes játékot:

Legyen A az első játékos és B a második játékos hasznossági mátrixa (mindkettőnél a sorok tartoznak az első játékoshoz). Ekkor tudjuk, hogy $B = -A$, mivel 0-összegű a játék. Legyen az A mátrix a következő:

1. \ 2.	X	Y	Z
E	0	1	2
F	8	3	0

A fenti játék nem igazán jó a második játékosnak, mert az soha sem nyer, legjobb esetben is csak nem kell fizetnie semmit.

Legyen az első játékos kevert stratégiája:

$$m^*(E) = \frac{3}{4}$$
$$m^*(F) = \frac{1}{4}$$

Nézzük meg a második játékosnak mennyi a várható nyereménye tiszta stratégiák esetén:

$$u(m^*, S_X) = -(\frac{3}{4} \cdot 0 + \frac{1}{4} \cdot 8) = -2$$
$$u(m^*, S_Y) = -(\frac{3}{4} \cdot 1 + \frac{1}{4} \cdot 3) = -1.5$$
$$u(m^*, S_Z) = -(\frac{3}{4} \cdot 2 + \frac{1}{4} \cdot 0) = -1.5$$

A második játékos legjobb tiszta válaszai: Y és Z. Legjobb kevert válasza pedig a legjobb tiszta válaszok konvex kombinációja.

Általánosan: Tegyük fel, hogy az 1. játékosra: $m_1 = (p_1, p_2, \ldots, p_k)$.

A pA (l hosszú) vektor j. eleme nem más, mint: $u_1(m_1, j) = -u_2(m_1, j)$.

Ekkor a 2. játékos legjobb válasza $p-re$: a pA legkisebb elemeinek indexe. Tehát az 1. játékos legjobb nyilvános stratégiája: olyan p keresése, amire pA legkisebb eleme a lehető legnagyobb.

8.6. Megjegyzés. A fenti feladat egy LP feladat:

$$(P)\qquad \begin{array}{l} \max w \\ p \geq 0 \\ \sum_{i=1}^{k} p_i = 1 \\ pA \geq w\mathbf{1}^T \end{array}$$

Az LP feladat optimális megoldása megadja a legjobb nyilvános stratégiát.

Nézzük a fenti LP feladat duálisát:

$$(D)\qquad \begin{array}{l} \min z \\ Aq \leq \mathbf{1}z \\ q \geq 0 \\ \mathbf{1}^T q = 1 \end{array}$$

8.3. Tétel. *Legyen (p^*, w^*) a (P) feladat, (q^*, z^*) a (D) feladat optimuma. Ekkor*

(i) $w^* = z^*$

(ii) A $(p^*, q^*) \in M$ kevert stratégia Nash-egyensúly.

Bizonyítás. $w^* = z^*$ következik a dualitás-tételből. A (ii) rész bizonyításához tekintsük a következő állítást:

8.7. Állítás. *q^* legjobb válasz p^*-ra.*

Bizonyítás. A komplementaritási feltételek miatt ahol $q^* > 0$, ott $p^*A - w^* = 0$, azaz w^* a p^*A vektor legkisebb eleme. Tehát q^* a legjobb tiszta válaszok konvex kombinációja. (1) □

Hasonlót állíthatunk p^*-ra: ahol $p^* > 0$, ott $Aq^* = z$, azaz z^* az Aq^* vektor legnagyobb eleme. Tehát p^* is legjobb válasz q^*-ra. (2)

(1) és (2)-ből következik tehát, hogy (p^*, q^*) Nash-egyensúly. □

8.7. Megjegyzés. A duális feladat adja a második játékos optimális nyilvános stratégiáit.

Most nézzük meg mit tudunk mondani szimmetrikus 0-összegű játékokról:

8.17. Definíció (Szimmetrikus játék)**.** Legyen A az első játékos és B a második játékos hasznossági mátrixa (mindkettőben a sorok felelnek meg az első játékos stratégiáinak). A játék szimmetrikus, ha $B = A^T$.

8.18. Definíció (Szimmetrikus stratégia)**.** Szimmetrikus játéknál egy kevert stratégia-vektort szimmetrikusnak nevezünk, ha minden játékoshoz ugyanaz a kevert stratégia tartozik.

8.19. Definíció (Szimmetrikus Nash-egyensúly)**.** Egy kevert stratégia szimmetrikus Nash-egyensúly, ha a belőle kapott szimmetrikus stratégia-vektor Nash-egyensúly.

Szimmetrikus 0-összegű játék esetén a fent szereplő (P) és (D) feladat ugyanaz, így optimális megoldásaik is ugyanazok. Ebből következik, hogy szimmetrikus 0-összegű játéknál mindig van szimmetrikus Nash-egyensúly. A következő részben megmutatjuk, hogy ez nem 0-összegű játékokra is igaz.

8.9. Lemke-Howson algoritmus kétszemélyes szimmetrikus játékra

Először megmutatjuk, hogy ha szimmetrikus játékban tudunk szimmetrikus Nash-egyensúlyt keresni, akkor tetszőleges játékban tudunk Nash-egyensúlyt keresni.

8.4. Tétel. *Tekintsünk egy olyan játékot, ahol az első játékos hasznossági mátrixa A, a második játékos hasznossági mátrixa B, és A és B összes eleme pozitív (ez feltehető, mert A és B minden elemét ugyanannyival megnövelve ekvivalens játékot kapunk). Legyen C a következő mátrix:*

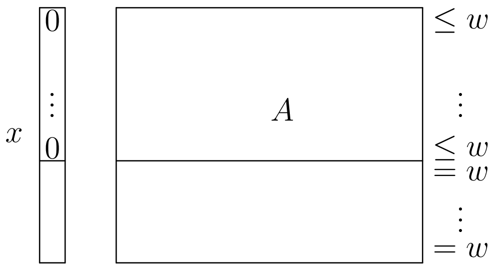
$$C = \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix} \qquad \begin{array}{c|cc} & x & y \\ \hline x & 0 & A \\ y & B^T & 0 \end{array}$$

Ha a C mátrix-szal adott játéknak (x, y) szimmetrikus Nash-egyensúlya, akkor $m_1 = \frac{x}{\sum x_i}$, $m_2 = \frac{y}{\sum y_i}$ Nash-egyensúly az eredeti játékban.

Bizonyítás. Mivel A és B minden eleme pozitív, sem x, sem y nem lehet azonosan 0, hiszen akkor (x, y) nem lenne legjobb válasz önmagára. Tehát $m_1 = \frac{x}{\sum x_i}$ és $m_2 = \frac{y}{\sum y_i}$ értelmes. Mikor $x_i > 0$, akkor s_i legjobb tiszta válasz (x, y)-ra a szimmetrikus játékban, ami pont azt jelenti, hogy s_i legjobb tiszta válasz m_2-re az eredeti játékban. Fordítva is, ha $y_j > 0$, akkor s_j legjobb tiszta válasz (x, y)-ra a szimmetrikus játékban, és így s_j legjobb tiszta válasz m_1-re az eredeti játékban. Tehát a 8.3. Tétel alapján (m_1, m_2) Nash-egyensúly. □

A következőkben ismertetjük a Lemke-Howson algoritmust, ami szimmetrikus kétszemélyes játékban talál egy szimmetrikus Nash-egyensúlyt ($m_1 = m_2$). Ez egyben bizonyítja Nash tételét a kétjátékosos esetre.

Tekintsük az első játékos A hasznossági mátrixát. Jelölje w az Ax maximumát. Ekkor lesznek olyan koordináták, amikre $Ax \leq w$, és lesznek olyanok, amikre az Ax eléri a maximumát.



Vizsgáljuk a következő poliédert: $P = \{x : Ax \leq 1, x \geq 0\}$

Feltehető, hogy

(i) $a_{ij} \geq 0$, mert ha lenne egy negatív elem, akkor minden elemhez ugyanazt az értéket hozzáadhatjuk, így a várható értéket konstanssal toljuk csak el, a Nash-egyensúly nem változik.

(ii) A-nak nincs csupa 0 oszlopa - ha lenne ilyen, azt töröljük.

8.8. Állítás. *P korlátos.*

Bizonyítás. Minden x_j korlátos mert ≥ 0 és $\leq \frac{1}{a_{ij}}$, ahol $a_{ij} > 0$. □

Ha P korlátos, akkor politóp.

8.5. Jelölés. Jelöljük a P politop csúcsait V-vel. Jelöljük továbbá minden $z \in V$ csúcsra:

$$N_z = \{j : z_j = 0\},$$
$$K_z = \{i : (Az)_i = 1\}.$$

Tudjuk, hogy $(|N_z| + |K_z|) \geq n$, mert egy csúcsnál legalább n egyenlőtlenség egyenlőséggel teljesül.

8.5. Lemma. *Ha $|N_z \cup K_z| = n$ és $z \neq 0$, akkor $x = \frac{z}{\sum z_j}$ egy szimmetrikus Nash-egyensúly.*

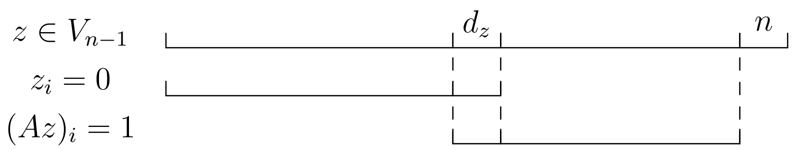
Bizonyítás. Ha $x_i > 0$, akkor $(Ax)_i = \max_{1 \leq j \leq n}(Ax)_j$, tehát x legjobb tiszta válaszok konvex kombinációja. □

8.9. Állítás. *Tegyük fel, hogy P nem degenerált, azaz minden csúcsában pontosan n db oldal találkozik ($|N_z| + |K_z| = n$). Ekkor z-nek pontosan n szomszédja van (két csúcs szomszédos ha egy élen vannak), és mindegyik más-más $n-1$ feltételt teljesít egyenlőséggel az z egyenlőséggel teljesülő feltételei közül.*

A továbbiakban az egyszerűség kedvéért feltesszük, hogy P nem degenerált. (Bizonyítható, hogy egy degenerált poliéderből nem degeneráltat kaphatunk az A mátrix kis perturbálásával, és elég kis perturbálás esetén az új játék Nash-egyensúlyának az eredeti játékban is Nash-egyensúly felel meg.)

8.6. Jelölés. Jelölje V_{n-1} az olyan $z \in V$-ket, amelyekre $N_z \cup K_z = \{1, 2, \ldots, n-1\}$, V_n pedig azokat, amelyekre $|N_z \cup K_z| = n$.

Jelölje továbbá d_z a $z \in V_{n-1}$ dupla koordinátáját.



Nevezzük a $z \in V_{n-1}$ csúcs d_z-szomszédainak a z-nek azt a két szomszédját, amit a d_z-hez tartozó egy-egy feltétel kidobásával kapunk.

8.10. Állítás. *A z csúcs d_z-szomszédai $(V_{n-1} \cup V_n)$-ben vannak.*

8.11. Állítás. *Ha $z \in V_{n-1}$ és z' egy d_z-szomszédja, amire $z' \in V_{n-1}$, és az új egyenlő feltétel F, akkor z'-ből indulva és F-et törölve z-t kapjuk vissza (mert ekkor $d_{z'}$ lesz az F dupla koordinátája). Tehát z a z'-nek $d_{z'}$-szomszédja.*

Vegyük észre, hogy $0 \in V_n$, de ez sajnos nem ad Nash-egyensúlyt. Az origónak viszont van egy $V_{n-1} \cup V_n$-beli z^1 szomszédja, amit a $z_n = 0$ feltétel kidobásával kapunk. Ha ez V_n-beli, akkor találtunk egy szimmetrikus Nash-egyensúlyt a 8.5. Lemma szerint. Ha nem, akkor lépjünk tovább a z^1 másik d_{z^1}-szomszédjára, amit a másik d_{z^1}-hez tartozó feltétel kidobásával kapunk. Ezt ismételjük, egészen addig amíg egy V_n-beli csúcsban nem érünk.

Figyeljük meg, hogy a lépések során nem érhetünk körbe: ha egy V_{n-1}-beli z csúcsba érnénk vissza, akkor annak a 8.11. Állítás alapján három különböző d_z-szomszédja lenne, ami lehetetlen. A 0 csúcsba sem érhetünk vissza, hiszen 0-nak egyetlen szomszédja van V_{n-1}-ben.

Az algoritmus tehát a következő:

0 Legyen z^1 az origónak a $z_n = 0$ feltétel kidobásával kapott szomszédja, és legyen $i = 1$.

1 Ha $z^i \in V_n$, akkor kész vagyunk.

2 Ha nem, dobjuk ki azt a d_{z^i}-hez tartozó feltételt, amelyik nem az utolsó lépésnél jött be, és legyen az így kapott szomszéd z^{i+1}.

3 Legyen $i := i + 1$, és lépjünk az 1. lépésre.

8.6. Következmény (Nash)**.** *Véges stratégiahalmazú 2 személyes játékban mindig van Nash-egyensúly.*

8.8. Megjegyzés. A Lemke-Howson algoritmus véges, de lehet exponenciális futási idejű az A mátrix méretében, hiszen a poliédernek lehet exponenciálisan sok csúcsa.

9. Konvex optimalizálás

9.1. Konvex halamzok

9.1.1. Alaptulajdonságok

9.1. Definíció (Konvex halmaz). $C \subseteq \mathbb{R}^n$ konvex, ha tetszőleges $x, y \in C$ és $0 \leq \lambda \leq 1$-re $\lambda x + (1-\lambda)y \in C$.

9.2. Definíció (Konvex kombináció). $\sum_{i=1}^{k} \lambda_i x_i$ az $x_1, \ldots, x_k \in \mathbb{R}^n$ vektorok konvex kombinációja, ahol $0 \leq \lambda_i \leq 1$, $\forall i = 1, \ldots, k$ és $\sum_{i=1}^{k} \lambda_i = 1$.

9.1. Állítás (Konvex halmaz ekvivalens definiciója). *C konvex $\Leftrightarrow$ tetszőleges véges sok pontjának konvex kombinációját tartalmazza.*

Bizonyítás. Indukció k-ra: 2-re igaz.

Tegyük fel, hogy $k-1$-re igaz. Legyen $\lambda_i > 0$, $i = 1, \ldots, k$. Ha valamelyik $\lambda_i = 0$, akkor azt kihagyva $k-1$-re már igaz.

Legyen $\lambda := \sum_{i=1}^{k-1} \lambda_i$, $x := \sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i$. Az x konvex kombinációja 1-től (k-1)-ig C-beli pontoknak, így indukció szerint $x \in C$. Ekkor $\sum_{i=1}^{k} \lambda_i x_i = \lambda x + \underbrace{\lambda_k}_{=1-\lambda} x_k = \lambda \underbrace{x}_{\in C} + (1-\lambda) \underbrace{x_k}_{\in C} \in C$.

□

9.2. Állítás. *Konvex halmazok metszete konvex.*

9.1. Megjegyzés. Az állítás végtelen sok konvex halmazra is igaz.

Bizonyítás. Legyen $C = \cap C_i$, $x, y \in C$. Ekkor $x, y \in C_i$ minden i-re. Tehát $0 \leq \lambda \leq 1$-re $\lambda x + (1-\lambda)y \in C_i, \forall i \Rightarrow \in C$. □

9.1. Jelölés. aff(C): C-t tartalmazó legszűkebb affin altér

9.3. Definíció (Konvex halmaz dimenziója). $\dim(C) := \dim(\mathrm{aff}(C))$.

9.3. Állítás. *Konvex C-re* $\mathrm{aff}(C) = \{\mu x + (1-\mu)y : x, y \in C, \mu \in \mathbb{R}\}$.

Bizonyítás. Legyen $x_1, \ldots, x_k \in C$, $z = \sum_{i=1}^{k} \mu_i x_i$, ahol $\sum \mu_i = 1$. Be kell látnunk, hogy z előáll $\mu x + (1-\mu)y$ alakban. Ha $\mu_i \geq 0$, $i = 1, \ldots, k$, akkor $z \in C$ és kész vagyunk.

Legyen $\mu = \sum_{i:\mu_i > 0} \mu_i$, $x = \sum_{i:\mu_i > 0} \frac{\mu_i}{\mu} x_i$. x konvex kombináció, tehát $\in C$.

Legyen $y = \sum_{i:\mu_i < 0} \frac{\mu_i}{1-\mu} x_i \in C$. Ekkor $z = \mu x + (1-\mu)y$. □

9.4. Definíció (Extremális halmaz). $E \subseteq C$ extremális halmaza a C konvex halmaznak, ha $x, y \in C$ és $0 < \lambda < 1$-re $\lambda x + (1-\lambda)y \in E \Leftrightarrow x \in E$ és $y \in E$.

9.4. Állítás. *Ha* $\max\{wx : x \in C\}$ *megoldható, akkor az optimális megoldások halmaza extremális halmaz lesz.*

Bizonyítás. $\Leftarrow$: Legyen S az optimális megoldások halmaza, legyen továbbá $z \in S$, α pedig az optimális megoldás értéke. $\alpha = wx$, $x \in S$. S konvex, mert $wx = \alpha$, $wy = \alpha$, tehát $w(\lambda x + (1-\lambda)y) = \alpha$, azaz $\lambda(wx) + (1-\lambda)(wy) = \alpha$.

$\Rightarrow$: $z \in S$, $z = \lambda x + (1-\lambda)y$, ahol $0 < \lambda < 1$, és $x, y \in C$. Be kell látnunk, hogy x és y S-ben is benne vannak.

$\alpha = wz = \lambda \underbrace{wx}_{\leq \alpha} + (1-\lambda) \underbrace{wy}_{\leq \alpha} \leq \lambda\alpha + (1-\lambda)\alpha = \alpha$. Egyenlőség csak akkor lehet, ha $wx = \alpha$ és $wy = \alpha$, azaz $x, y \in S$. □

Példa. Henger extremális halmazai:

0. dimenziós extremális halmaz elemei (extremális pontok): a két körlap határpontjai;
1. dimenziós extremális halmazok: a henger alkotói;
2. dimenziós extremális halmazok: a két körlap;
3. dimenziós extremális halmazok: az egész henger.

9.5. Állítás. *Ha E extremális halmaz, akkor $E = \operatorname{aff}(E) \cap C$.*

Bizonyítás. Legyen $z \in \operatorname{aff}(E) \cap C$. Ekkor $z = \mu x + (1-\mu)y$, ahol $x, y \in E$.

Ha $0 \leq \mu \leq 1$, akkor az E konvexitása miatt $z \in E$ és kész vagyunk.

Ha $\mu > 1$, akkor $x = \frac{1}{\mu} \underbrace{z}_{\in C} + \frac{\mu - 1}{\mu} \underbrace{y}_{\in E}$. Mivel y $\in E$ és $E \subseteq C$, ezért $y \in C$, továbbá mivel $x \in E$, és E extremális halmaz, ezért $z, y \in E$. □

9.2. Megjegyzés. Ha E_1, E_2 extremális C-ben és $E_1 \subsetneqq E_2$, akkor E_1 extremális halmaza E_2-nek.

9.6. Állítás. *Ha E_1, E_2 extremális halmazai C-nek és $E_1 \subsetneqq E_2$, akkor $\dim(E_1) < \dim(E_2)$.*

Bizonyítás. Mivel E_1 extremális halmaza E_2-nek, ezért $E_1 = \operatorname{aff}(E_1) \cap E_2$, tehát $\operatorname{aff}(E_1) \subsetneqq \operatorname{aff}(E_2)$, emiatt $\dim(E_1) < \dim(E_2)$. □

9.7. Állítás. *Ha E_1 extremális halmaza C-nek, és E_2 extremális halmaza E_1-nek, akkor E_2 extremális halmaza C-nek.*

Bizonyítás. Legyen $x, y \in C$, $0 < \lambda < 1$, $z = \lambda x + (1-\lambda)y \in E_2$. Ekkor $z \in E_1$, mert $E_2 \subseteq E_1$. E_1 extremális C-ben, emiatt $x, y \in E_1$. Továbbá $z \in E_2$, E_2 extremális E_1-ben, emiatt $x, y \in E_2$. □

9.1. Tétel (Krein-Millman, bizonyítás nélkül)**.** *Ha C kompakt, korlátos és zárt konvex halmaz, akkor C az extrém pontjainak konvex burka.*

9.5. Definíció (Konvex kúp)**.** C konvex kúp, ha konvex és tetszőleges $x \in C$ és $\lambda \geq 0$ esetében $\lambda x \in C$.

9.6. Definíció (Csúcsos kúp)**.** Konvex kúp csúcsos, ha 0 extremális pontja.

9.8. Állítás. *C konvex kúp csúcsos $\Leftrightarrow$ nem tartalmaz egydimenziós lineáris alteret.*

Bizonyítás. $\Rightarrow$: Ha tartalmaz $\{\mu x : \mu \in \mathbb{R}\}$ egyenest, akkor $x \in C$, $-x \in C$, tehát 0 nem extremális pontja.

$\Leftarrow$: Ha 0 nem extremális, akkor $\exists x, y \in C : 0 < \lambda < 1$, amire $\lambda x + (1-\lambda)y = 0$, tehát az x-et és y-t összekötő egyenes egydimenziós lineáris altér C-ben. □

9.7. Definíció (Megengedett irányok kúpja)**.** Legyen C konvex halmaz, $z \in C$. Ekkor

$$K(C,z) = \{x \in \mathbb{R}^n : \exists \varepsilon > 0, z + \delta x \in C, \forall 0 \leq \delta \leq \varepsilon\}$$

9.9. Állítás. *$K(C,z)$ konvex kúp.*

Bizonyítás. Legyen $x, y \in K(C,z)$. Az x-hez tartozó ε-t jelöljük ε_x-el, az y-hoz tartozót pedig ε_y-al. Legyen $\varepsilon = \min\{\varepsilon_x, \varepsilon_y\}$. Legyen továbbá $\delta < \varepsilon$. Ekkor

$$z + \delta(\lambda x + (1-\lambda)y) = \lambda(\underbrace{z + \delta x}_{\in C}) + (1-\lambda)(\underbrace{z + \delta y}_{\in C}) \Rightarrow z + \delta(\lambda x + (1-\lambda)y) \in C$$

□

9.10. Állítás. *$K(C,z)$ csúcsos $\Leftrightarrow$ z extremális pontja C-nek.*

Bizonyítás. Ha z nem extremális pont, akkor $(x-z) \in K(C,z)$, $(y-z) \in K(C,z)$, tehát tartalmaz 0-n átmenő egyenest. Ha viszont tartalmaz 0-n átmenő egyenest, akkor $K(C,z)$ tartalmazza a $\{\mu s : \mu \in \mathbb{R}\}$ egyenest, emiatt $\exists \delta_1 : z + \delta_1 s \in C$, $\exists \delta_2 : z - \delta_2 s \in C$. □

9.1.2. Konvex halmazok szeparációja

9.2. Tétel. *Ha C zárt konvex halmaz, és $z \notin C$, akkor van egy egyértelmű z-hez legközelebbi pont C-ben.*

Bizonyítás. Legyen $\mu = \inf\{\|x-z\| : x \in C\}$. Ez véges, és létezik $x^k \in C$ sorozat, amire $\|x^k - z\| \to \mu$. Ennek van konvergens részsorozata, ami egy $x \in C$ ponthoz tart, tehát $\|x - z\| = \mu$.

Ha $y \in C$-re $\|y - z\| = \mu$,akkor $\mu^2 \leq \|\frac{x+y}{2} - z\|^2 = \mu^2 - \frac{1}{4}\|y - x\|^2$, tehát $y = x$. □

9.8. Definíció (Vetítés konvex zárt halmazra)**.** A $z \notin C$ ponthoz legközelebbi pontot C-ben a z pont C-re való vetítésének nevezzük, jelölése $\pi_C(z)$.

9.11. Állítás. *Tetszőleges $x \in C$-re $(x - \pi_C(z))^T(z - \pi_C(z)) \leq 0$.*

9.3. Tétel (Konvex szeparációs tétel)**.** *Legyen C zárt konvex halmaz, $z \notin C$. Ekkor létezik $s \in \mathbb{R}^n$ és $\epsilon > 0$: $s^T x \leq s^T z - \epsilon$ tetszőleges $x \in C$-re.*

Bizonyítás. Legyen $s = z - \pi_C(z)$. Az előző állítás szerint tetszőleges $x \in C$-re $s^T x \leq s^T \pi_C(z)$. Azaz: $s^T x \leq s^T \pi_C(z) = s^T z + s^T(\pi_C(z) - z) = s^T z - \|s\|^2$. Tehát $\epsilon = \|s\|^2$ jó választás. □

9.4. Tétel. *Legyen C konvex halmaz, $z \notin C$. Ekkor létezik $s \in \mathbb{R}^n$ nemnulla vektor: $s^T x \leq s^T z$ tetszőleges $x \in C$-re.*

Bizonyítás. Létezik $z^k \to z$ sorozat, amire $z^k \notin \mathrm{cl}(C)$ semmilyen k-ra. Az előző tétel alapján vannak s^k nemnulla vektorok, amikre $(s^k)^T x \leq (s^k)^T z^k$ tetszőleges $x \in C$-re. Feltehetjük, hogy $\|s^k\| = 1$ minden k ra. Legyen s az s^k sorozat tetszőleges torlódási pontja; errre $s^T x \leq s^T z$ tetszőleges $x \in C$-re. □

9.5. Tétel. *Legyenek C_1 és C_2 diszjunkt konvex halmazok. Létezik $s \in \mathbb{R}^n$ nemnulla vektor: $s^T x \leq s^T y$ tetszőleges $x \in C_1$-re és $y \in C_2$-re.*

Bizonyítás. Legyen $C = C_1 - C_2$; ez konvex halmaz, és $0 \notin C$, tehát az előző tétel értelmében létezik $s \in \mathbb{R}^n$ nemnulla vektor: $s^T x \leq 0$ tetszőleges $x \in C$-re. Így $s^T x \leq s^T y$ tetszőleges $x \in C_1$-re és $y \in C_2$-re. □

9.2. Konvex függvények

9.9. Definíció (Konvex függvény)**.** Legyen $C \in \mathbb{R}^n$ konvex halmaz. Az $f : C \to \mathbb{R}$ függvény konvex, ha tetszőleges $x, y \in C$ és $0 \le \lambda \le 1$-re $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$.

9.3. Megjegyzés. Mivel C konvex, ezért $\lambda x + (1-\lambda)y \in C$.

9.6. Tétel (Jensen egyenlőtlenség)**.** *Legyen C konvex halmaz, $f : C \to \mathbb{R}$ konvex függvény, továbbá $x_1, \ldots, x_k \in C$, $\lambda_i \ge 0$, $i = 1, \ldots, k$, $\sum_{i=1}^k \lambda_i = 1$. Ekkor $f(\sum_{i=1}^k \lambda_i x_i) \le \sum_{i=1}^k \lambda_i f(x_i)$.*

Bizonyítás. Indukció k-ra; $k = 2$-re jó, mert f konvex.

Tegyük fel, hogy $k > 2$. Legyen $\lambda = \sum_{i=1}^{k-1} \lambda_i$, $x = \sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i$. Mivel x C-beli pontok konvex kombinációja, ezért $x \in C$. Ekkor $f(\sum_{i=1}^k \lambda_i x_i) = f(\lambda x + \lambda_k x_k) = f(\lambda x + (1-\lambda)x_k) \underbrace{\le}_{f \text{ konvex}} \lambda f(x) + (1-\lambda)f(x_k) = \lambda f(\sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i) + \lambda_k f(x_k) \underbrace{\le}_{\text{indukció}} \sum_{i=1}^{k-1} \lambda_i f(x_i) + \lambda_k f(x_k) = \sum_{i=1}^k \lambda_i f(x_i)$. □

9.10. Definíció (Szinthalmaz)**.** $D_\alpha(f) = \{x \in C : f(x) \le \alpha\}$ az f függvény α-hoz tartozó szinthalmaza.

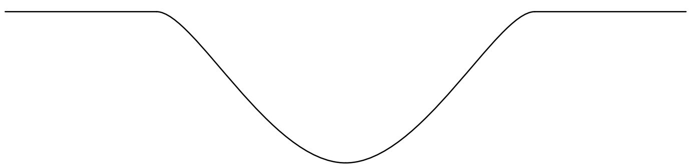
9.7. Lemma. *Ha f konvex, akkor D_α konvex tetszőleges α-ra.*

Bizonyítás. Legyen $x, y \in D_\alpha$, $0 \le \alpha \le 1$.

Ekkor $f(\lambda x + (1-\lambda)y) \underbrace{\le}_{f \text{ konvex}} \lambda f(\underbrace{x}_{\in D_\alpha}) + (1-\alpha)f(\underbrace{y}_{\in D_\alpha}) \le \lambda\alpha + (1-\lambda)\alpha = \alpha$. □

9.4. Megjegyzés. A lemma fordítottja nem igaz: Ha egy függvény minden szinthalmaza konvex $\not\Rightarrow$ a függvény konvex.

Példa



9.11. Definíció (Kvázi-konvex függvény)**.** Ha egy függvény minden szinthalmaza konvex, akkor a függvényt kvázi-konvexnek nevezzük.

9.8. Lemma. *Legyen C konvex halmaz, $f : C \to \mathbb{R}$. f konvex $\Leftrightarrow$ tetszőleges $x, y \in C$-re, $x \ne y$-ra a $\phi(\lambda) := f(x + \lambda(y-x))$ függvény konvex a $[0,1]$-en.*

Bizonyítás. $\Leftarrow$: $x, y \in C$, $0 \le \lambda \le 1$. Ekkor $f(\lambda x + (1-\lambda)y) = \phi(1-\lambda) = \phi(0 \cdot \lambda + 1 \cdot (1-\lambda)) \underbrace{\le}_{\phi \text{ konvex}} \lambda \underbrace{\phi(0)}_{=f(x)} + (1-\lambda) \underbrace{\phi(1)}_{=f(y)} = \lambda f(x) + (1-\lambda)f(y)$.

$\Rightarrow$: $0 \le \nu \le 1$, $\lambda_1, \lambda_2 \in [0,1]$. Ekkor

$$\begin{aligned}
\phi(\nu\lambda_1 + (1-\nu)\lambda_2) &= f(x + (\nu\lambda_1 + (1-\nu)\lambda_2)(y-x)) \\
&= f(\nu(x + \lambda_1(y-x)) + (1-\nu)(x + \lambda_2(y-x))) \\
&\underbrace{\le}_{f \text{ konvex}} \nu f(x + \lambda_1(y-x)) + (1-\nu)f(x + \lambda_2(y-x)) \\
&= \nu\phi(\lambda_1) + (1-\nu)\phi(\lambda_2).
\end{aligned}$$

□

Tudjuk, hogy ha ϕ folytonosan differenciálható, akkor ϕ konvex $\Leftrightarrow$ $\phi'(x)$ monoton növő.

9.9. Tétel. *Legyen C konvex, nyílt halmaz, $f : C \to \mathbb{R}$ folytonosan differenciálható függvény. Ekkor f konvex $\Leftrightarrow$ tetszőleges $x, y \in C$-re, $x \neq y$: $\nabla f(x)^T(y-x) \leq f(y)-f(x) \leq \nabla f(y)^T(y-x)$.*

Bizonyítás. $\Rightarrow$: $f(x + (1-\lambda)(y-x)) = f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$. Ekkor $\frac{f(x+(1-\lambda)(y-x))-f(x)}{1-\lambda} \leq f(y) - f(x)$. Ha $\lambda \to 1$, akkor a baloldal tart az iránymenti deriválthoz: $\nabla f(x)^T(y-x)$. A másik egyenlőtlenséget szimmetrikusan bizonyítjuk x-et és y-t felcserélve.

$\Leftarrow$: Tetszőleges $x, y \in C$-re $\nabla f(x)^T(y-x) \leq \nabla f(y)^T(y-x)$. Ekkor tetszőleges $0 \leq \lambda_1 < \lambda_2 \leq 1$ számokra: $\nabla f(\lambda_2 x + (1-\lambda_2)y)^T(x-y) \leq \nabla f(\lambda_1 x + (1-\lambda_1)y)^T(x-y) \Rightarrow \phi\,'(\lambda)$ monoton növekedő $\Rightarrow \phi$ konvex $\Rightarrow f$ konvex. □

9.12. Definíció (Hesse mátrix)**.** Legyen C konvex és nyílt halmaz, f kétszer folytonosan differenciálható függvény a C halmazon. f Hesse mátrixa: $(\nabla^2 f(x))_{ij} = \frac{\delta^2 f(x)}{\delta x_i \delta x_j}$.

9.10. Tétel. *f konvex $\Leftrightarrow$ $\nabla^2 f(x)$ pozitív szemidefinit minden $x \in C$-re.*

Bizonyítás. $\Rightarrow$: $x \in C$, $y \in C$, $s = y - x$. Ekkor $s^T \nabla^2 f(x) s = \phi\,''(0)$. $\phi\,''(0) \geq 0$, mert $\phi\,'$ monoton növő. C nyílt, így tartalmaz x körüli gömböt, ezért $s^T \nabla^2 f(x) s \geq 0\ \forall s \in \mathbb{R}^n \Leftrightarrow \nabla^2 f(x)$ pozitiv szemidefinit.

$\Leftarrow$: $0 \leq s \nabla^2 f(x + \lambda s)s = \phi\,''(\lambda)$. $\phi\,'' \geq 0$, $\forall \lambda \in [0,1] \Rightarrow \phi\,'$ monoton növő $\Rightarrow$ f konvex. □

9.3. Feltétel nélküli optimalizálás

9.13. Definíció (Lokális minimum)**.** $f : C \to \mathbb{R}$, $\bar{x} \in C$ lokális minimum, ha $\exists\ \varepsilon > 0$ amire $f(x) \geq f(\bar{x})\ \forall x \in C : ||x - \bar{x}|| \leq \varepsilon$.

9.14. Definíció (Globális minimum)**.** $\bar{x} \in C$ globális minimum, ha $f(x) \geq f(\bar{x})\ \forall x \in C$.

9.11. Tétel. *Ha f konvex és $\bar{x}$ lokális minimum, akkor $\bar{x}$ globális minimum.*

Bizonyítás. $f(x) < f(\bar{x}) \Rightarrow f(\lambda x + (1-\lambda)\bar{x}) \leq \lambda f(x) + (1-\lambda) f(\bar{x}) < f(\bar{x})$ tetszőleges $0 < \lambda \leq 1$ esetén. Tehát $\bar{x}$ tetszőleges környezetében van kisebb értékű pont. □

9.12. Tétel. *Legyen C nyílt, konvex halmaz, f konvex, folytonosan differenciálható függvény. Ekkor $\bar{x}$ globális minimum $\Leftrightarrow \nabla f(\bar{x}) = 0$.*

Bizonyítás. $\Leftarrow$: legyen $x \in C$ tetszőleges. Ekkor

$$0 = \nabla f(\bar{x})^T(x - \bar{x}) \underbrace{\leq}_{\text{9.9. tétel}} f(x) - f(\bar{x}) \Rightarrow f(x) \geq f(\bar{x}).$$

$\Rightarrow$: tetszőleges $s \in \mathbb{R}^n$-re: $f(\bar{x} + \lambda s) \geq f(\bar{x})$ ha λ elég kicsi ($\bar{x} + \lambda s \in C$). Ekkor $\frac{f(\bar{x}+\lambda s)-f(\bar{x})}{\lambda} \geq 0$ tart az iránymenti deriválthoz: $\nabla f(\bar{x})^T s$-hez ha $\lambda \to 0$. Tehát $\nabla f(\bar{x})^T s \geq 0\ \forall s \in \mathbb{R}^n \Rightarrow \nabla f(\bar{x}) = 0$. □

9.15. Definíció (Konvex halmaz relatív belseje)**.** $\text{rint}(C) = \{x \in C :$ tetszőleges $x \neq y, y \in C$-re $\exists\ x' \in C$ és $0 < \lambda < 1$: $x = \lambda x' + (1-\lambda)y\}$.

9.16. Definíció (Relatív nyílt halmaz)**.** C relatív nyílt, ha $C = \text{rint}(C)$.

9.17. Definíció. lin(C): aff(C) lineáris eltoltja.

9.13. Tétel. *Legyen C relatív nyílt, konvex halmaz, f konvex, folytonosan differenciálható függvény. Ekkor $\bar{x}$ globális minimum $\Leftrightarrow \nabla f(\bar{x})^T s = 0\ \forall s \in \text{lin}(C)$.*

Bizonyítás. $\Leftarrow$: legyen $x \in C$ tetszőleges. Ekkor

$$0 = \nabla f(\bar{x})^T \underbrace{(x - \bar{x})}_{\in \text{lin}(C)} \underbrace{\leq}_{\text{9.9. tétel}} f(x) - f(\bar{x}) \Rightarrow f(x) \geq f(\bar{x}).$$

$\Rightarrow$: tetszőleges $s \in \text{lin}(C)$-re: ha λ elég kicsi, akkor $\bar{x} + \lambda s \in C$, tehát $f(\bar{x} + \lambda s) \geq f(\bar{x})$. Ekkor $\frac{f(\bar{x}+\lambda s)-f(\bar{x})}{\lambda} \geq 0$ tart az iránymenti deriválthoz, $\nabla f(\bar{x})s$-hez ha $\lambda \to 0$. Tehát $\nabla f(\bar{x})s \geq 0$ $\forall s \in \text{lin}(C)$. $\square$

9.4. Feltételes optimalizálás

A feladat a következő:

$$(CO)\quad \begin{array}{l} \min f(x) \\ x \in C \\ g_i(x) \leq 0\ \forall i = 1, \ldots, m \end{array}$$

Tegyük fel továbbá, hogy

i C konvex, $C \subseteq \mathbb{R}^n$;

ii f, g_i konvexek, folytonosan differenciálhatóak.

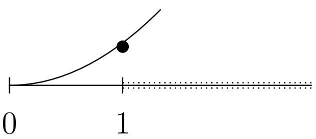
9.18. Definíció (Megengedett megoldások halmaza)**.** $F := \{x \in C : g_i(x) \leq 0, i = 1, \ldots, m\}$.

9.5. Megjegyzés. A megengedett megoldások halmaza konvex, mivel konvex függvény szinthalmazai konvexek, és konvex függvények metszete is konvex.

9.6. Megjegyzés. Általában F nem relatív nyílt. (ha az lenne, akkor az optimalitás szükséges és elégséges feltétele az lenne hogy a gradiens merőleges a lineáris altérre).

Példa

$$\begin{array}{l} n = 1: \\ f(x) = x^2 \\ g_1(x) = 1 - x \\ C = \mathbb{R} \end{array}$$



9.19. Definíció (Megengedett irányok kúpja)**.** Legyen $K(x)$ a megengedett irányok kúpja.

$$K(x) := \{s \in \mathbb{R}^n : \exists \varepsilon > 0 : \text{ tetszőleges } 0 \leq \lambda \leq \varepsilon\text{-ra } x + \lambda s \in F\}.$$

$K(x)$ tulajdonságai:

i $K(x)$ konvex kúp;

ii $K(x)$ csúcsos $\Leftrightarrow x$ extremális pontja F-nek;

iii $K(x) \subseteq \text{lin}(F)$;

iv $K(x) = \text{lin}(F) \Leftrightarrow x \in \text{rint}(F)$

9.14. Tétel. *$\bar{x}$ optimális megoldása a CO feladatnak*

$$\Updownarrow$$

$$\nabla f(\bar{x})^T s \geq 0 \ \forall s \in K(\bar{x})$$

Bizonyítás. $\Uparrow$: Legyen $\bar{x} \neq x \in F$, $s = x - \bar{x}$. Ekkor $s \in K(\bar{x})$.

Tudjuk, hogy $f(x) - f(\bar{x}) \underbrace{\geq}_{f\ konvex} \nabla f(\bar{x})^T s \underbrace{\geq 0}_{s \in K(\bar{x})} \Rightarrow f(x) \geq f(\bar{x})$.

$\Downarrow$: Legyen $s \in K(\bar{x})$. Ekkor definició szerint $\exists \varepsilon > 0$: $x_\lambda := \bar{x} + \lambda s \in F$, $\forall 0 < \lambda \leq \varepsilon$.

$f(x_\lambda) \geq f(\bar{x}) \Rightarrow \frac{f(x_\lambda) - f(\bar{x})}{\lambda} \geq 0$.

Ha $\lambda \to 0$, akkor $0 \leq \frac{f(\bar{x}+\lambda s) - f(\bar{x})}{\lambda} \to \nabla f(\bar{x})^T s$. □

Szeretnénk jellemezni az optimumot a ∇f, ∇g_i $i = 1, \ldots, m$-mel. De ez általában nem jellemzi az optimumokat.

Példa

$$\begin{array}{rl} & C = \mathbb{R}^2 \\ & f(x_1, x_2) = x_1 \\ & g_1(x_1, x_2) = x_2 \\ (a) & g_2(x_1, x_2) = -x_2 \\ (b) & g_2(x_1, x_2) = x_1^2 - x_2 \end{array}$$

(a) (b)

A fenti példában (a)-ban és (b)-ben a gradiensek 0-ban ugyanazok, míg (a)-nál $-\infty$ az infimum, (b)-nél pedig $F = \{0\}$, tehát az optimum 0.

9.20. Definíció (Slater pont)**.** A CO feladatnak x Slater pontja, ha

i $x \in \text{rint}(C)$;

ii $g_i(x) \leq 0$ $i = 1, \ldots, m$;

iii $g_i(x) < 0$ ha g_i nemlineáris.

9.21. Definíció (Slater-reguláris feladat)**.** A CO feladat Slater reguláris, ha van Slater pont.

Példa

Az előző példánál (a) esetben tetszőleges $(x_1, 0)$ pont Slater pont, tehát a feladat Slater reguláris, míg a (b) esetben nincs Slater pont, mert 0 az egyetlen megengedett pont, ami nem teljesíti a (iii) tulajdonságot: $g_2(0) = 0$, de g_2 nemlineáris.

9.15. Tétel (Karush-Kuhn-Tucker). *Legyen C relatív nyílt és CO Slater reguláris. Ekkor $\bar{x} \in F$ pontosan akkor optimális megoldása CO-nak, ha*

$$\begin{aligned}
&\exists \mu_i \geq 0 \quad (i = 1, \ldots, m): \\
&\qquad \nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) \text{ merőleges } \mathrm{lin}(C)\text{-re}, \\
&\qquad \text{ha } g_i(\bar{x}) < 0, \text{ akkor } \mu_i = 0.
\end{aligned}$$

Bizonyítás. $\Uparrow$: Jelölés: $I = \{i \in \{1, \ldots, m\} : g_i(\bar{x}) = 0\}$.

Legyen $s \in K(\bar{x}) \Rightarrow s \in \mathrm{lin}(C)$. Ekkor tudjuk, hogy

$$0 = s^T(\nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x})) = s^T \nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}).$$

De $\sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}) \leq 0$ mert $\sum_{i=1}^{m} \mu_i s^T \nabla g_i(\bar{x}) = \sum_{i \in I} \mu_i s^T \nabla g_i(\bar{x})$ és $i \in I$: $s^T \nabla g_i(\bar{x}) \leq 0$ mert s megengedett irány és $g_i(\bar{x}) = 0$.

Tehát $0 \leq s^T \nabla f(\bar{x}) \Rightarrow \bar{x}$ optimális megoldás.

$\Downarrow$: Indirekt tegyük fel, hogy $\nexists \mu_i \geq 0$, $i \in I$ amire $\nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x})$ merőleges $\mathrm{lin}(C)$-re. A Farkas lemma szerint ekkor $\exists v \in \mathrm{lin}(C)$:

$$\begin{aligned}
&v^T \nabla g_i(\bar{x}) \leq 0 \ (i = 1, \ldots, m) \text{ és} \\
&v^T \nabla f(\bar{x}) < 0
\end{aligned}$$

Legyen x^s Slater pont. Ekkor $g_i(x^s) \leq 0 \ \forall i$, így

- ha $i \in I$: $\underbrace{g_i(x^s)}_{\leq 0} + \underbrace{g_i(\bar{x})}_{=0} \leq 0 \underbrace{\Rightarrow}_{g_i \text{ konvex}} \nabla g_i(\bar{x})(x^s - \bar{x}) \leq 0.$
- ha $i \in I$ és g_i nemlineáris: $g_i(x^s) < 0 \Rightarrow \nabla g_i(\bar{x})(x^s - \bar{x}) < 0$.

Legyen $s = v + \delta(x^s - \bar{x})$, $\delta > 0$. Ekkor tudjuk:

- $i \in I$: $\nabla g_i(\bar{x}) s^T \leq 0$;
- $i \in I$ és g_i nemlineáris: $\nabla g_i(\bar{x}) s^T < 0$;
- elég kicsi δ-ra $\nabla f(\bar{x}) s^T < 0$ mert $\nabla f(\bar{x}) v^T < 0$
- Ha δ olyan kicsi mint az előző feltételben, akkor $s = v + \delta(x^s - \bar{x}) \in \mathrm{lin}(C)$.

Ennek segítségével találhatunk egy $\bar{x}$-nál kisebb értékű megengedett pontot:

- C relatív nyílt $\Rightarrow$ elég kicsi $\varepsilon > 0$-ra $\bar{x} + \varepsilon s \in C$;
- elég kicsi ε-re: $f(\bar{x} + \varepsilon s) < f(\bar{x})$
  - $i \in I$, g_i nemlineáris: $g_i(\bar{x} + \varepsilon s) < g_i(\bar{x}) = 0$
  - $i \in I$, g_i lineáris: $g_i(\bar{x} + \varepsilon s) \leq g_i(\bar{x}) = 0$
- $i \notin I$: elég kis ε-re: $g_i(\bar{x} + \varepsilon s) < 0$ mert $g_i(\bar{x}) < 0$ és g_i folytonos függvény.

Tehát $\bar{x} + \varepsilon s \in F$ és $f(\bar{x} + \varepsilon s) < f(\bar{x})$, azaz $\bar{x}$ nem optimális. $\square$

9.22. Definíció (Lagrange függvény)**.** $x \in C$-re és $\mu \in \mathbb{R}^m_+$-re

$$L(x, \mu) := f(x) + \sum_{i=1}^{m} \mu_i g_i(x)$$

A Lagrange függvény tulajdonságai:

i Ha $x \in F$ akkor $L(x, \mu) \leq f(x)$;

ii Ha $x \in C \setminus F$ akkor $\exists \mu$: $L(x, \mu) > f(x)$;

Adott μ-re: $L(\mu) := \inf\{L(x, \mu) : x \in C\}$.

9.7. Megjegyzés. Ez egy konvex optimalizálási feladat.

9.8. Megjegyzés. Az első tulajdonságból következik, hogy $L(\mu) \leq \inf_{x \in F} f(x)$.

9.23. Definíció (Langrange duális feladat)**.**

$$\sup_{\mu \in \mathbb{R}^m_+} L(\mu)$$

9.9. Megjegyzés. Ez a feladat is egy konvex optimalizálási feladat.

9.16. Tétel (Gyenge dualitás)**.** *Ha $L(\bar{\mu}) = f(\bar{x})$, $\bar{\mu} \in \mathbb{R}^m_+$, $\bar{x} \in F$, akkor $\bar{x}$ optimális megoldása a (CO)-nak, $\bar{\mu}$ optimális megoldása a Lagrange duális feladatnak.*

Bizonyítás. $L(\mu) \leq f(\bar{x}) = L(\bar{\mu}) \leq f(x)$ tetszőleges μ-re és x-re. $\square$

9.17. Tétel (Erős dualitás)**.** *Ha (CO) Slater reguláris és $\bar{x}$ optimális megoldás, akkor $\exists \bar{\mu} \in \mathbb{R}^m_+$, amire $L(\bar{\mu}) = f(\bar{x})$.*

Bizonyítás. Ha $\bar{x}$ optimális, akkor a Karush-Kuhn-Tucker tétel szerint $\exists \bar{\mu} \in \mathbb{R}^m_+$, amire

(1) $\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i \nabla g_i(\bar{x})$ merőleges $\mathrm{lin}(C)$-re;

(2) ha $g_i(\bar{x}) < 0$, akkor $\bar{\mu}_i = 0$.

9.12. Állítás. *$\bar{x}$ optimális megoldása a $\min_{x \in C} L(x, \bar{\mu})$ feladatnak.*

Bizonyítás. $\min_{x \in C} L(x, \bar{\mu})$: konvex függvény minimalizálása relatív nyílt halmazon. Tudjuk, hogy $\bar{x}$ optimális $\Leftrightarrow$ $\nabla_{x=\bar{x}} L(x, \bar{\mu})$ merőleges $\mathrm{lin}(C)$-re.

$$\nabla_{x=\bar{x}} L(x, \bar{\mu}) = \nabla_{x=\bar{x}}(f(x) + \sum_{i=1}^{m} \bar{\mu}_i g_i(x)) = \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i \nabla g_i(\bar{x})$$

ami merőleges $\mathrm{lin}(C)$-re (1) szerint.

$\triangle$

9.13. Állítás. *$L(\bar{x}, \bar{\mu}) = f(\bar{x})$.*

Bizonyítás. $L(\bar{x}, \bar{\mu}) = f(\bar{x}) + \sum_{i=1}^{m} \bar{\mu}_i g_i(\bar{x})$. De $\sum_{i=1}^{m} \bar{\mu}_i g_i(\bar{x}) = 0$, mert

- (2) szerint: ha $g_i(\bar{x}) < 0$, akkor $\bar{\mu}_i = 0$;
- $\bar{x} \in F \Rightarrow g_i(\bar{x}) \leq 0 \ \forall i \in \{1, \dots, m\}$.

$\triangle$

Tehát $L(\bar{\mu}) = L(\bar{x}, \bar{\mu}) = f(\bar{x})$. $\square$

9.24. Definíció (Megengedett csökkenési irány)**.** Az s vektor megengedett csökkenési irány x^0-ban, ha $\exists \varepsilon > 0$, hogy tetszőleges $0 < \lambda \leq \varepsilon$-ra $f(x^0 + \lambda s) < f(x^0)$ és $x^0 + \lambda s \in F$.

Megengedett csökkenési irányok keresése x^0-ban (LP feladat)

$$\begin{aligned}
&\max u \\
&\nabla f(x^0)^T s + u \leq 0 \\
&\nabla g_i(x^0)^T s + u \leq 0, \text{ ha } g_i(x^0) = 0 \text{ és } g_i \text{nem lineáris} \\
&\nabla g_i(x^0)^T s \leq 0, \text{ ha } g_i(x^0) = 0 \text{ és } g_i \text{lineáris} \\
&u \geq 0 \\
&-1 \leq s_j \leq 1 \ j = 1, \dots, n \\
&s \in \mathrm{lin}(C)
\end{aligned}$$

9.14. Állítás. *Ha $\exists (s, u)$ megoldás, ahol $u > 0$, akkor s megengedett csökkenési irány.*

Bizonyítás. $\nabla f(x^0)^T s < 0$, tehát s csökkenési irány. Másrészt $\nabla g_i(x^0)^T s \leq 0$, ha $g_i(x^0) = 0$ és $\nabla g_i(x^0)^T s < 0$, ha g_i nemlineáris, tehát s megengedett irány. $\square$

9.15. Állítás. *Ha a feladat Slater reguláris, akkor ha x^0 nem optimális, akkor $\exists (s, u)$ megoldás, amire $u > 0$.*

Bizonyítás. Ha x^0 nem optimális, akkor $\exists v \in \mathrm{lin}(C)$:

$$\begin{aligned}
&v^T \nabla g_i(\bar{x}) \leq 0 \ i = 1, \dots, m \text{ és} \\
&v^T \nabla f(\bar{x}) < 0
\end{aligned}$$

Legyen x^s Slater pont. Ekkor $g_i(x^s) \leq 0 \ \forall i$, így

- ha $g_i(\bar{x}) = 0$: $\underbrace{g_i(x^s)}_{\leq 0} + \underbrace{g_i(\bar{x})}_{=0} \leq 0 \underbrace{\Rightarrow}_{g_i \text{ konvex}} \nabla g_i(\bar{x})(x^s - \bar{x}) \leq 0$.
- ha $g_i(\bar{x}) = 0$ és g_i nemlineáris: $g_i(x^s) < 0 \Rightarrow \nabla g_i(\bar{x})(x^s - \bar{x}) < 0$.

Legyen $s = v + \delta(x^s - \bar{x})$, $\delta > 0$. Ekkor tudjuk:

- $g_i(\bar{x}) = 0$: $\nabla g_i(\bar{x}) s^T \leq 0$;
- $g_i(\bar{x}) = 0$ és g_i nemlineáris: $\nabla g_i(\bar{x}) s^T < 0$;
- elég kicsi δ-ra $\nabla f(\bar{x}) s^T < 0$ mert $\nabla f(\bar{x}) v^T < 0$
- Ha δ olyan kicsi mint az előző feltételben, akkor $s = v + \delta(x^s - \bar{x}) \in \mathrm{lin}(C)$.

Szorozzuk meg s-t egy skalárral úgy, hogy $-1 \leq s_j \leq 1$. Ehhez az s-hez választhatunk megfelelő $u > 0$-t. $\square$

9.5. Gradiens módszer

Legyen $f : \mathbb{R}^n \to \mathbb{R}$ konvex, folytonosan differenciálható függvény. Tegyük fel, hogy $x^0 \in \mathbb{R}^n$ pontban vagyunk.

9.16. Állítás. *x^0-ban az iránymenti derivált a $-\nabla f(x^0)$ irányban a legkisebb (azonos normájú vektorok közül).*

Bizonyítás. s irányban az iránymenti derivált: $\nabla f(x^0)^T s$.

$$-\nabla f(x^0)^T \nabla f(x^0) = \min_{||s||=||\nabla f(x^0)||} \nabla f(x^0)^T s$$

Tehát $-\nabla f(x^0)$ a legmeredekebb csökkenési irány. □

Algoritmus

$x^0 \in \mathbb{R}^n$, $k = 0$.

1. Kiszámoljuk $s = -\nabla f(x^k)$-t. Ha $s = 0$, akkor kész vagyunk.
2. Legyen $\lambda^k = \operatorname{argmin}_{\lambda \in \mathbb{R}_+} f(x^k + \lambda s)$.
3. Legyen $x^{k+1} = x^k + \lambda^k s$.
4. $k := k + 1$, és tovább az 1. lépésre.

9.17. Állítás. *$f(x^{k+1}) < f(x^k)$*

Bizonyítás. s irányban az iránymenti derivált: $-\nabla f(x^k)^T \nabla f(x^k) < 0$. □

9.18. Tétel. *Ha $D = \{x \in \mathbb{R}^n : f(x) \leq f(x^0)\}$ kompakt, akkor az $x^0, x^1, x^2, \ldots$ tetszőleges $\bar{x}$ torlódási pontjára $\bar{x}$ optimális megoldása a feladatnak.*

Bizonyítás. Vegyük $x^{j_1}, x^{j_2}, \cdots \to \bar{x}$ konvergens részsorozatot. Mivel f folytonos, ezért $\lim_{i\to\infty} f(x^{j_i}) = f(\bar{x})$ és $\lim_{i\to\infty} \nabla f(x^{j_i}) = \nabla f(\bar{x})$. Legyen $s = -\nabla f(\bar{x})$. Ekkor $\nabla f(\bar{x})^T s \leq 0$ és $= 0 \Leftrightarrow s = 0$.

Másrészt: tetszőleges $\lambda \geq 0$-ra és i-re $f(x^{j_{i+1}}) \leq f(x^{j_i+1}) \leq f(x^{j_i} - \lambda \nabla f(x^{j_i}))$. Az $i \to \infty$ határátmenet mutatja, hogy $f(\bar{x}) \leq f(\bar{x} + \lambda s)$ tetszőleges $\lambda \geq 0$-ra. Tehát $\frac{f(\bar{x}+\lambda s) - f(\bar{x})}{\lambda} \geq 0$. Ha $\lambda \to 0$, akkor azt kapjuk, hogy $\nabla f(\bar{x})^T s \geq 0 \Rightarrow s = 0 \Rightarrow \nabla f(\bar{x}) = 0 \Rightarrow \bar{x}$ optimális. □

9.6. Aranymetszés módszer

A fenti módszerben a $\lambda^k = \operatorname{argmin}_{\lambda \in \mathbb{R}_+} f(x^k + \lambda s)$ kiszámolásához meg kell oldanunk egy egydimenziós konvex függvény minimalizálási feladatot. Persze ezt is csak közelítőleg tudjuk megoldani. Egy lehetséges megoldási módszer az Aranymetszés módszer, ami nem csak differenciálható függvényekre alkalmazható.

Legyen f konvex függvény $\mathbb{R}^n$-en. Az algoritmus egyre csökkenő méretű intervallumokat ad, amik garantáltan tartalmaznak optimális megoldást. Minden lépésben adott lesz $\alpha_i < \beta_i < \gamma_i < \delta_i$, amikre $f(\alpha_i) \geq f(\beta_i)$, $f(\gamma_i) \leq f(\delta_i)$, és

$$\frac{\gamma_i - \alpha_i}{\delta_i - \alpha_i} = \frac{\delta_i - \beta_i}{\delta_i - \alpha_i} = q := \frac{\sqrt{5} - 1}{2}.$$

Kiindulásként $i = 0$-ra találunk ilyeneket. A konvexitásból következik, hogy van optimális megoldás az $[\alpha_i, \delta_i]$ intervallumban.

Az aranymetszés szabálya miatt az is teljesül, hogy

$$\frac{\beta_i - \alpha_i}{\gamma_i - \alpha_i} = \frac{\delta_i - \gamma_i}{\delta_i - \beta_i} = q.$$

Az általános lépésben összehasonlítjuk az $f(\beta_i)$ és $f(\gamma_i)$ értékeket. Két eset van:

- Ha $f(\beta_i) < f(\gamma_i)$: legyen

$$\begin{aligned}
\alpha_{i+1} &= \alpha_i \\
\beta_{i+1} &= q\alpha_i + (1-q)\gamma_i \\
\gamma_{i+1} &= \beta_i \\
\delta_{i+1} &= \gamma_i
\end{aligned}$$

- Ha $f(\beta_i) \geq f(\gamma_i)$: legyen

$$\begin{aligned}
\alpha_{i+1} &= \beta_i \\
\beta_{i+1} &= \gamma_i \\
\gamma_{i+1} &= q\delta_i + (1-q)\beta_i \\
\delta_{i+1} &= \delta_i
\end{aligned}$$

Könnyen ellenőrizhető, hogy az új pontok is teljesítik a feltételeket, az $[\alpha_k, \delta_k]$ intervallum hossza q-szorosára csökken, és csak egy új függvényértéket kell kiszámolni.

9.7. Newton módszer

Tegyük fel, hogy $f : \mathbb{R}^n \to \mathbb{R}$ kétszer folytonosan differenciálható konvex függvény.

Legyen f másodrendű közelítése:

$$q_{x^0}(x) = f(x^0) + \nabla f(x^0)^T(x - x^0) + \frac{1}{2}(x - x^0)\nabla^2 f(x^0)(x - x^0).$$

Ekkor $\nabla^2 q_{x^0}(x) = \nabla^2 f(x^0)$ pozitiv szemidefinit, mert f konvex $\Rightarrow q_{x^0}(x)$ konvex.

Minimalizáljuk $q_{x^0}(x)$-et.

$$\bar{x} \text{ optimális} \Leftrightarrow \underbrace{\nabla q_{x^0}(\bar{x})}_{=\nabla f(x^0)+\nabla^2 f(x^0)(\bar{x}-x^0)} = 0$$

Tegyük fel, hogy $\nabla^2 f(x^0)$ invertálható:

$$\bar{x} \text{ optimális} \Leftrightarrow \bar{x} = x^0 - (\nabla^2 f(x^0))^{-1}\nabla f(x^0).$$

Algoritmus. Tegyük fel, hogy f szigorúan konvex (tehát $\nabla^2 f(x)$ pozitiv definit). $x^0 \in \mathbb{R}^n$, $k = 0$.

1. Kiszámoljuk $\nabla f(x^k)$-t és $\nabla^2 f(x^k)$-t. Ha $\nabla f(x^k) = 0$, akkor kész vagyunk.

2. $x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$.

3. $k := k + 1$ és tovább az 1. lépésre.

9.10. Megjegyzés. 1-dimenzióban $x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$.

9.11. Megjegyzés. Ha $\nabla^2 f(x^k)$ nem invertálható, akkor vegyük helyette $\nabla^2 f(x^k) + \alpha I$-t valamilyen megfelelően választott α-ra.