

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Szabó Máté

**SZTOCHASZTIKUS ÉS ROBUSZTUS
OPTIMALIZÁLÁS PÉNZÜGYI
ALKALMAZÁSOKKAL**

Szakdolgozat
Matematika BSc

Témavezető:
Dr. Király Tamás
Operációkutatási tanszék



Budapest, 2025

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, Király Tamásnak, akihez bátran fordulhattam kérdéseimmel, és aki szakmai tudásával, iránymutatásával és támogatásával segítette szakdolgozatom megvalósulását.

Hálás vagyok korábbi oktatóimnak is, akik az elmúlt évek során nemcsak ismereteket adtak át, hanem gondolkodni is megtanítottak.

Köszönettel tartozom családomnak, barátaimnak is, hogy mellettem álltak, nemcsak ebben az időszakban, hanem végig az egyetemi éveim alatt.

Tartalomjegyzék

Bevezetés	4
1. Alapfogalmak	5
2. Egylépcsős sztochasztikus programozás	7
2.1. Bevezetés	7
2.2. Lehetséges minőségi mutatók	8
2.3. Valószínűségi korlátok, ahol csak a jobb oldal véletlen	9
3. Többlépcsős sztochasztikus programozás	13
3.1. Bevezetés	13
3.2. Aggregációs technikák	14
3.3. A kétlépcsős modell	16
3.4. A diszkrét véges eset	17
3.5. Benders-dekompozíció	18
3.6. Beágyazott dekompozíció	22
4. Robusztus lineáris optimalizálás	29
4.1. Alapfogalmak	29
4.2. Fontos megfigyelések	30
4.3. Robusztus feladat átírása kezelhető feladatra	31
5. Gyakorlati példák	35
5.1. Többperiódusos befektetési probléma	35
5.2. Többperiódusos adósságkezelési probléma	36
6. Összegzés és kitekintés	39

Bevezetés

A pénzügyi világ számos kihívás elé állítja az embert, mert a gazdasági bizonytalanságok és a jövőbeli kockázatok nagy hatással vannak a vállalatok és magánszemélyek stratégiáira egyaránt.

Sztochasztikus optimalizálás során az optimális döntéseket olyan modellek alapján hozzuk meg, amelyek figyelembe veszik a kockázatot és a jövőbeli bizonytalanságokat, ismert valószínűségi eloszlások alapján. Ezzel szemben a robusztus optimalizálás nem a bizonytalanság eloszlására, hanem annak egy ismert halmazára épít, olyan döntéseket keresve, amelyek a legrosszabb esetben is kielégítik az előírt feltételeket.

E szakdolgozat célja bemutatni a matematikai optimalizálásnak ezen két fontos ágazatát. A szakdolgozat elején a sztochasztikus programozás egy- és többlépcsős esetével fogok foglalkozni. Ezeknek először elméleti alapjait fogom tárgyalni, majd speciális esetek felépítéseit fogom vizsgálni. Gyakran látni fogjuk, hogy az elsőre nehezen kezelhető, véletlent tartalmazó problémákat korábban megtanult lineáris programozási vagy konvex optimalizálási problémákká lehet alakítani. Dolgozatom második részében a robusztus optimalizálásról lesz szó, ahol először bevezetem a robusztus optimalizálás alapfogalmait, majd áttérek a bizonytalansági halmazok speciális eseteire.

Szakdolgozatom utolsó fejezete a többlépcsős sztochasztikus programozás alkalmazását mutatja be gyakorlati pénzügyi feladatokon.

1. fejezet

Alapfogalmak

Az alábbiakban azokat a korábban tanult definíciókat és tételeket gyűjtöttem össze, amelyeket szakdolgozatom megértése szempontjából a legfontosabbnak tartok. Fontos azonban megjegyezni, hogy ezen kívül is léteznek olyan, meg nem említett, előismeretek, amelyek szükségesek lehetnek.

1.0.1. Definíció. (Optimalizálási feladat): Adott egy $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ függvény és egy $H \subseteq \mathbb{R}^n$ halmaz. Egy standard optimalizálási feladat célja, hogy egy olyan $x^* \in H$ -t találjunk, amely maximalizálja (minimalizálja) $f(x)$ -et. A H halmazt a megengedett megoldások halmazának, az f -et pedig célfüggvénynek nevezzük.

1.0.2. Megjegyzés. Megeshet, hogy nem létezik ilyen x^* . Ha az előbb említett H üres, akkor a probléma megoldhatatlan, ha pedig található egy olyan $x_k^* \in H$ pontsorozat, hogy $f(x_k^*) \rightarrow \infty$, ha $k \rightarrow \infty$, akkor a célfüggvény felülről nem korlátos.

1.0.3. Definíció. (Optimális megoldás): Optimális megoldásnak nevezünk egy olyan $x^* \in H$ -t, amelyre teljesül a következő: $f(x^*) \geq f(x)$, $\forall x \in H$.

1.0.4. Definíció. (Valószínűségi mező): A $(\Omega, \mathcal{A}, \mathbb{P})$ hármast valószínűségi mezőnek hívjuk, ha:

- Ω a lehetséges kimenetek halmaza
- $\mathcal{A}$ egy σ -algebra, az események rendszere (azaz Ω részhalmazaiából álló család)
- $\mathbb{P}$ valószínűségi mérték, vagyis σ -additív $\mathcal{A} \rightarrow [0, 1]$ függvény

1.0.5. Definíció. (Valószínűségi vektorváltozó): Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ egy valószínűségi mező és $d \geq 1$ egy egész szám. Egy

$$X = (X_1, X_2, \dots, X_d) : \Omega \rightarrow \mathbb{R}^d$$

függvényt valószínűségi vektorváltozónak nevezünk, ha minden komponense valószínűségi változó, vagyis olyan $X_i : \Omega \rightarrow \mathbb{R}$, $(i = 1, 2, \dots, d)$ függvény, amire teljesül, hogy minden Borel-halmaz ősképe benne van $\mathcal{A}$ -ban, vagyis mérhető vagy $\mathcal{A}$ -mérhető. Formálisan:

$$X^{-1}(H) = \{\omega \in \Omega \mid X(\omega) \in H\} \in \mathcal{A}, \quad \forall H \in \mathcal{B}(\mathbb{R}^d).$$

1.0.6. Definíció. (Primál-duál LP feladatpár általános alakjai):

Primál feladat	Duális feladat
$\max_{x_0, x_1} c_0 x_0 + c_1 x_1$	$\min_{y_0, y_1} b_0^T y_0 + b_1^T y_1$
s.t. $Px_0 + Ax_1 = b_0$	s.t. $y_0 P + y_1 Q = c_0$
$Qx_0 + Bx_1 \leq b_1$	$y_0 A + y_1 B \geq c_1$
$x_1 \geq 0$	$y_1 \geq 0$

ahol:

- $x_0 \in \mathbb{R}^r$, $x_1 \in \mathbb{R}^n$: a primál változók,
- $y_0 \in \mathbb{R}^m$, $y_1 \in \mathbb{R}^p$: a duál változók,
- $c_0 \in \mathbb{R}^r$, $c_1 \in \mathbb{R}^n$: a primál célfüggvény vektorai,
- $b_0 \in \mathbb{R}^m$, $b_1 \in \mathbb{R}^p$: a primál megszorítások jobb oldali vektorai,
- $P \in \mathbb{R}^{m \times r}$, $A \in \mathbb{R}^{m \times n}$, $Q \in \mathbb{R}^{p \times r}$, $B \in \mathbb{R}^{p \times n}$: a feltételmatrixok.

1.0.7. Tétel. (Dualitási alternatíva tétel): Az alábbi négy állítás közül pontosan az egyik teljesül:

1. A primál és duál feladat is megoldható, és

$$\max_{x_0, x_1} (c_0 x_0 + c_1 x_1) \leq \min_{y_0, y_1} (b_0^T y_0 + b_1^T y_1).$$

(Erős dualitás esetén egyenlőség áll fenn.)

2. Egyik sem oldható meg.
3. A primál feladatnak nincs megoldása, és a duális feladat célfüggvénye alulról nem korlátos.
4. A duál feladatnak nincs megoldása, és a primál feladat célfüggvénye felülről nem korlátos.

1.0.8. Definíció. (Konvex halmaz): Egy $H \in \mathbb{R}^d$ halmazt konvexnek nevezünk, ha $\forall x, y \in H$ pontpárra és $\forall \lambda \in [0, 1]$ -re a $\lambda x + (1 - \lambda)y$ pont is $\in H$.

1.0.9. Definíció. (Konvex függvény): Legyen $H \in \mathbb{R}^d$ halmaz $f : H \rightarrow \mathbb{R}$ függvény. Azt mondjuk, hogy az f függvény konvex, ha H konvex halmaz, és $\forall x, y \in H$ és $\forall \lambda \in [0, 1]$ esetén teljesül:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

1.0.10. Definíció. (Logkonkáv függvény): Legyen $f : \mathbb{R}^d \rightarrow [0, \infty)$. Azt mondjuk, hogy f logkonkáv, ha $\forall x, y \in \mathbb{R}^d$ és $\forall \lambda \in [0, 1]$ esetén teljesül:

$$f(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda}.$$

Ezzel ekvivalens, hogy az f függvény logaritmus konkáv.

1.0.11. Megjegyzés. Logkonvexitás definíciója ugyanez, fordított egyenlőtlenséggel.

2. fejezet

Egylépcsős sztochasztikus programozás

Az egylépcsős sztochasztikus programozási feladatokban a célunk egy olyan x döntési vektor megválasztása, amely megfelel a determinisztikus feltételeknek és a bizonytalanságot hordozó véletlen vektor realizációjától függő feltételeknek is. Mivel a véletlen eseményeket x megválasztása után figyeljük meg, ezért nincs lehetőség x utólagos megváltoztatására. A feladat célja tehát az, hogy úgy válasszunk megengedett x megoldást, még a bizonytalan események bekövetkezése előtt, hogy ennek a véletlen események figyelembevételével kiértékelt célfüggvényértéke a lehető legkedvezőbb legyen. A fejezet a [2], [3], [5], [8] forrásokat veszi alapul.

2.1. Bevezetés

Az egylépcsős sztochasztikus lineáris programozási modellekben, akár a feltételekben, akár a célfüggvényben, rendszerint megtalálható a következő típusú véletlen vektor: $\psi(x, \xi) = T(\xi)x - h(\xi)$, ahol $\xi : \Omega \rightarrow \mathbb{R}^r$ vektorváltozó egy $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mező felett. Vagyis $T(\xi) \in \mathbb{R}^{s \times n}$ és $h(\xi) \in \mathbb{R}^s$, így ξ -től függő véletlen mátrix és véletlen vektor. Erről az összefüggésről feltesszük, hogy lineáris, vagyis:

$$T(\xi) = T + \sum_{i=1}^r T_i \xi_i \quad h(\xi) = h + \sum_{i=1}^r h_i \xi_i$$

Itt T , T_i , h , h_i megfelelő méretű determinisztikus mátrixok és vektorok $i = 1, \dots, r$.

Továbbá feltesszük, hogy az egylépcsős modellek esetében (és később a többlépcsős modellek esetében is) a bizonytalanságot hordozó vektorváltozók eloszlása ismert. Ennek megfelelően ismertnek tekintjük azoknak a paramétereknek az együttes eloszlását is, amelyek ezekből a vektorváltozókból származnak, és amelyek a célfüggvényben vagy a feltételekben szerepelnek. Jelen esetben ξ eloszlását, valamint a $(T(\xi), h(\xi))$ páros együttes eloszlását is ismertnek feltételezzük.

Vegyük észre, hogy $T(\xi)x - h(\xi)$ egy affin leképezés rögzített T -re és h -ra, így x vektor egy adott rögzítése során egy konkrét eloszlást határozunk meg $\psi(x, \xi)$ -nek. Tehát amikor az x döntési vektort megválasztjuk, akkor gyakorlatilag arra törekszünk, hogy a $\psi(x, \xi)$ véletlen vektor eloszlása számunkra megfelelő tulajdonságokkal rendelkezzen. Ahhoz, hogy egy véletlen vektorról el tudjuk dönteni, hogy „mennyire jó” számunkra, bevezetünk egy $\mu : \mathbb{R}^n \rightarrow \mathbb{R}$ úgynevezett „minőségi mutató” függvényt. Egy adott x kiértékelését pedig úgy végezzük, hogy a μ -be behelyettesítjük a $\psi(x, \xi)$ -t, vagyis $G(x) = \mu(\psi(x, \xi))$. Az egylépcsős sztochasztikus programozási feladatok két alapvető típusa a következő alakokkal rendelkezik:

$$\begin{array}{ll} \min_x c^T x & \min_x c^T x + G(x) \\ Ax = b & Ax = b \\ G(x) \geq d & \end{array}$$

2.2. Lehetséges minőségi mutatók

A legegyszerűbb megközelítés egy véletlen vektor minőségi mutatójára az, ha egyszerűen vesszük a várható értékét, vagyis: $\mu(\psi(x, \xi)) = \mathbb{E}[(\psi(x, \xi))]$. Ekkor a $\mathbb{E}[(\psi(x, \xi))] \geq 0$ a feltételek között azt jelenti, hogy $\mathbb{E}[T(\xi)x] \geq \mathbb{E}[h(\xi)]$. Ez azonban egy meglehetősen durva rendszer, ugyanis a $(T(\xi), h(\xi))$ együttes eloszlást egyetlen mátrixszal és vektorral jellemzi, és semmilyen tekintettel nincs a szórásra.

2.2.1. Példa. Tekintsük az alábbi problémát:

$$\begin{array}{ll} \min_{x_0, x_1} x_0 + x_1 \\ x_0 \geq h_1(\xi) \\ x_1 \geq h_2(\xi), \end{array}$$

ahol $h(\xi) = (\xi, \xi)$, és ξ és $\mathbb{P}(\xi = 0) = 0,999$, $\mathbb{P}(\xi = 1000) = 0,001$. Ekkor ha $h(\xi)$ -t $\mathbb{E}[h(\xi)]$ -vel helyettesítjük a következő LP-feladatot kapjuk:

$$\begin{array}{ll} \min_{x_0, x_1} x_0 + x_1 \\ x_0 \geq 1 \\ x_1 \geq 1, \end{array}$$

amire az $(x_0, x_1) = (1, 1)$ az optimális megoldáspár, azonban vegyük észre, hogy az $(x_0, x_1) = (0, 0)$ ugyanakkora valószínűséggel lesz megengedett megoldás az eredeti problémára, mint az $(1, 1)$, de a célfüggvényértéke kisebb.

Ahelyett, hogy a vektorváltozókat a várható értékükkel helyettesítenénk, egy ennél kifinomultabb megközelítés, a valószínűségi korlát, ahol a minőségi mutatót a $\mu(\psi(x, \xi) = \mathbb{P}((\psi(x, \xi) \geq 0)) \iff \mathbb{P}(T(\xi)x \geq h(\xi))$ formában definiáljuk. Így a valószínűséggel korlátozott modellekben a megszokott feltételek ($Ax = b$) mellett,

a $\mathbb{P}(T(\xi)x \geq h(\xi)) \geq \alpha$ típusú, vagy a kifejezés zárójelén belüli feltételeket külön kiértékelő $\mathbb{P}(t_i^T(\xi)x \geq h_i(\xi)) \geq \alpha_i$ típusú feltételek jelennek meg, ahol $\alpha, \alpha_i \in [0, 1]$, $i = 1, \dots, s$. Az előbbi típusú feltételeket közös vagy együttes valószínűségi korlátoknak, míg az utóbbiakat különálló valószínűségi korlátoknak nevezzük.

2.3. Valószínűségi korlátok, ahol csak a jobb oldal véletlen

A valószínűségi feltételekkel korlátozott modellek esetén különösen fontos megvizsgálni a megengedett megoldások halmazának konvexitását. Amennyiben a célfüggvény is megfelelő tulajdonságokkal rendelkezik: például minimalizálás esetén konvex, maximalizálás esetén konkáv, akkor a feladat egy konvex optimalizálási probléma.

Az ilyen problémák előnyei közé tartozik, hogy a célfüggvénynek van globális szélsőértéke. Egyes esetekben rendelkezésünkre állnak megoldó algoritmusok, míg más esetekben csupán a konvexitás tulajdonságai alapján tudunk hasznos következtetéseket tenni, vagy becsülni a megoldást.

A következőkben azokat az eseteket fogom bemutatni, amikor a valószínűséggel korlátozott feltételek között csak olyanok vannak, melyeknek a $\mathbb{P}$ -n belüli résznek a jobb oldala véletlen, a bal oldala determinisztikus, vagyis $T(\xi) \equiv T$. Először a különálló, majd az együttes valószínűségi korlátokkal ellátott modellekkel fogok foglalkozni, melyek az alábbi módon néznek ki:

Különálló valószínűségi korlátok

$$\begin{aligned} \min_x \quad & c^T x \\ & Ax = b \\ & \mathbb{P}(t_i^T x \geq h_i(\xi)) \geq \alpha_i, \quad i = 1, \dots, s \end{aligned}$$

Együttes valószínűségi korlátok

$$\begin{aligned} \min_x \quad & c^T x \\ & Ax = b \\ & \mathbb{P}(Tx \geq h(\xi)) \geq \alpha \end{aligned}$$

2.3.1. Definíció. Legyen $H \subseteq \mathbb{R}^d$ halmaz és $f : H \rightarrow \mathbb{R}$ függvény. Azt mondjuk, hogy az f függvény kvázikonvex, ha H konvex halmaz, és $\forall x, y \in H$ és $\forall \lambda \in [0, 1]$ esetén teljesül:

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}.$$

Továbbá, ha f kvázikonvex, akkor $g = -f$ kvázikonkáv: $g(\lambda x + (1 - \lambda)y) \geq \min\{g(x), g(y)\}$.

2.3.2. Állítás. Az egydimenziós valószínűségi változók eloszlásfüggvényei kvázikonvexek és kvázikonkávak is.

Bizonyítás. Nyilvánvalóan következik abból, hogy az eloszlásfüggvények monoton növekvők. Ekkor $F(\lambda x + (1 - \lambda)y) \leq \max\{F(x), F(y)\}$ biztosan teljesül, hiszen egy $[x, y]$ intervallumon belüli tetszőleges ponton felvett függvényérték legfeljebb

annyi, mint $F(y)$. A kvázikonkáv tulajdonságot teljesen ugyanezzel a logikával tudjuk belátni. $\square$

Különálló valószínűségi korlátok

Ilyen esetben $G_i(x) = \mathbb{P}(t_i^T x \geq h_i(\xi)) = F_{h_i(\xi)}(t_i^T x)$, $i = 1, \dots, s$

2.3.3. Állítás. A $H = \{x \mid F_{h_i(\xi)}(t_i^T x) \geq \alpha\}$ halmaz konvex.

Bizonyítás. Mivel $F_{h_i(\xi)}$ kvázikonkáv függvény, ezért a következő igaz:

$$F_{h_i(\xi)}(t_i^T(\lambda x + (1 - \lambda)y)) \geq \min\{F_{h_i(\xi)}(t_i^T x), F_{h_i(\xi)}(t_i^T y)\}$$

Kell: $\forall x, y \in H$ és $\forall \lambda \in [0, 1]$ -re $\lambda x + (1 - \lambda)y \in H$, vagyis $F_{h_i(\xi)}(t_i^T(\lambda x + (1 - \lambda)y)) \geq \alpha$. Ez teljesülni fog, hiszen, ha $t_i^T x \leq t_i^T y$, akkor $F_{h_i(\xi)}(t_i^T(\lambda x + (1 - \lambda)y)) \geq \min\{F_{h_i(\xi)}(t_i^T x), F_{h_i(\xi)}(t_i^T y)\} = F_{h_i(\xi)}(t_i^T x) \geq \alpha$. $\square$

2.3.4. Megjegyzés. A $H = \{x \mid F_{h_i(\xi)}(t_i^T x) \leq \alpha\}$ halmaz is konvex, $F_{h_i(\xi)}$ kvázikonvexitása miatt, hasonló bizonyítással.

Így az egyenlőtlenségrendszer által meghatározott halmaz konvex halmazok metszete, tehát konvex. Sőt LP-re is átírható a feladat, hiszen a $F_{h_i(\xi)}(t_i^T x) \geq \alpha$ feltétel ekvivalens a $t_i^T x \geq F_{h_i(\xi)}^{-1}(\alpha)$ feltétellel, ahol a $F_{h_i(\xi)}^{-1}(\alpha) = \inf\{s \in \mathbb{R} : F_{h_i(\xi)}(s) \geq \alpha\}$ az általánosított inverz eloszlásfüggvény vagy más néven kvantilisfüggvény.

Együttes valószínűségi korlátok

A korábbiakból világos, hogy a megengedett megoldások halmazának konvexitásához $G(x) = \mathbb{P}(Tx \geq h(\xi)) = F_{h(\xi)}(Tx)$ feltételek mellett az eloszlásfüggvények kvázikonkavitása elégséges feltétel volt. Ez a tulajdonság tetszőleges $\mathbb{R} \rightarrow \mathbb{R}$ eloszlásfüggvényre teljesül, azonban az $\mathbb{R}^d \rightarrow \mathbb{R}$ eloszlásfüggvényeknek csak bizonyos halmazaira, amennyiben $2 \leq d \in \mathbb{Z}$. A következőkben ezekről az esetekről lesz szó.

Amennyiben a $h_i(\xi)$, $i = 1, \dots, s$, valószínűségi változók egymástól függetlenek, akkor az együttes eloszlásfüggvényük a marginális eloszlások eloszlásfüggvényeinek szorzataként adható meg, azaz $F_{h(\xi)} = \prod_{i=1}^s F_{h_i(\xi)}$. Sajnálatos módon azonban a kvázikonkáv függvények szorzatáról általában nem mondható el, hogy az is kvázikonkáv lenne. Például, ha $f_1(x) = -x^2$ és $f_2(x) = -1$, akkor ezek szorzata $f_1(x) \cdot f_2(x) = x^2$, ami konvex, tehát nem kvázikonkáv. Ezzel ellentétben a konkáv függvények összege konkáv, vagyis a természetes ötlet, hogy vegyük a logaritmusát az eloszlásfüggvényeknek. Így $\log F_{h(\xi)} = \sum_{i=1}^s \log F_{h_i(\xi)} \geq \log \alpha$ feltétellel helyettesítve az eredetit ekvivalens feladatot kapunk. Ebből látszik, hogy a marginális eloszlásfüggvények logkonkavitása elégséges feltétel, ahhoz, hogy konvex tartományon optimalizáljunk.

Amennyiben nem feltételezhető a valószínűségi változók függetlensége, úgy általában nem tehetünk olyan állítást a marginális eloszlások tulajdonságairól, amely

önmagában biztosítaná, hogy az optimalizálási probléma konvex marad. Ehelyett célszerű olyan többdimenziós eloszlásokat vizsgálni, amelyek együttes eloszlásfüggvénye kvázikonkáv, mivel ez közvetlenül garantálhatja a konvex tartományokat, és így a konvex optimalizálási keret megtartását.

2.3.5. Állítás. Minden logkonkáv függvény egyben kvázikonkáv is.

Bizonyítás. Legyen $f : \mathbb{R}^d \rightarrow [0, \infty)$ logkonkáv függvény, azaz

$$f(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda}, \quad \forall x, y \in \mathbb{R}^d, \lambda \in [0, 1].$$

Legyen $m = \min\{f(x), f(y)\}$. Mivel mindkét $f(x), f(y) \geq m$, ezért:

$$f(x)^\lambda \geq m^\lambda, \quad f(y)^{1-\lambda} \geq m^{1-\lambda}.$$

Ezért:

$$f(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda} \geq m^\lambda m^{1-\lambda} = m^{\lambda+(1-\lambda)} = m = \min\{f(x), f(y)\}.$$

Ez pontosan a kvázikonkáv függvény definíciója. $\square$

2.3.6. Tétel. (Prékopa-Leindler egyenlőtlenség): Legyen $0 < \lambda < 1$, és $f, g, h : \mathbb{R}^d \rightarrow [0, \infty)$ nemnegatív, valós értékű mérhető függvények, amelyekre $\forall x, y \in \mathbb{R}^d$ esetén

$$f(\lambda x + (1 - \lambda)y) \geq g(x)^\lambda h(y)^{1-\lambda}.$$

Ekkor

$$\int_{\mathbb{R}^d} f(t) dt \geq \left(\int_{\mathbb{R}^d} g(x) dx \right)^\lambda \left(\int_{\mathbb{R}^d} h(y) dy \right)^{1-\lambda}.$$

2.3.7. Tétel. Logkonkáv sűrűségfüggvényű eloszlás, eloszlásfüggvénye is logkonkáv.

Bizonyítás. Legyen $f : \mathbb{R}^d \rightarrow [0, \infty)$ logkonkáv függvény, azaz

$$f(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda f(y)^{1-\lambda}, \quad \forall x, y \in \mathbb{R}^d, \lambda \in [0, 1].$$

Legyenek A és $B \subset \mathbb{R}^d$ konvex halmazok. Defináljuk a f a következő függvényeket:

$$f_1(x) = \begin{cases} f(x), & \text{ha } x \in A, \\ 0, & \text{egyébként,} \end{cases}$$

$$f_2(x) = \begin{cases} f(x), & \text{ha } x \in B, \\ 0, & \text{egyébként,} \end{cases}$$

$$f_3(x) = \begin{cases} f(x), & \text{ha } x \in \lambda A + (1 - \lambda)B, \\ 0, & \text{egyébként.} \end{cases}$$

Ekkor f_1, f_2 és f_3 is logkonkávak. Ez azért van így, hiszen ha x, y olyan pontok, amelyekre $f_i(x) = f(x)$ és $f_i(y) = f(y)$, akkor következik abból, hogy f logkonkáv. Ha meg $f_i(x) = 0$ vagy $f_i(y) = 0$, akkor is nyilvánvalóan teljesül a logkonkávítás definíciója. A következő egyenlőtlenség fog teljesülni f_1, f_2 és f_3 függvényekre:

$$f_3(t) \geq \sup_{t=\lambda x+(1-\lambda)y} f_1(x)^\lambda f_2(y)^{1-\lambda}.$$

Ez logkonkávitásból következik, ha $x \in A$ és $y \in B$, és ha $x \notin A$ vagy $y \notin B$, akkor $f_1(x) = 0$ vagy $f_2(y) = 0$, így $f_1(x)^\lambda f_2(y)^{1-\lambda} = 0$. Alkalmazva a Prékopa–Leindler-egyenlőtlenséget a $f(t) = f_3(t)$ $g(x) = f_1(x)^\lambda$, $h(y) = f_2(y)^{1-\lambda}$ függvényekre, azt kapjuk, hogy:

$$\begin{aligned} \int_{\mathbb{R}^d} f_3(t) dt &\geq \left(\int_{\mathbb{R}^d} f_1(x) dx \right)^\lambda \left(\int_{\mathbb{R}^d} f_2(y) dy \right)^{1-\lambda}, \\ &\quad \Updownarrow \\ \int_{\lambda A + (1-\lambda)B} f(t) dt &\geq \left(\int_A f(x) dx \right)^\lambda \left(\int_B f(y) dy \right)^{1-\lambda}, \end{aligned}$$

azaz:

$$\mathbb{P}(\lambda A + (1-\lambda)B) \geq \mathbb{P}(A)^\lambda \mathbb{P}(B)^{1-\lambda}.$$

Mivel beláttuk, hogy a valószínűségi mérték logkonkáv, ezért következik, hogy az eloszlásfüggvény is logkonkáv, hiszen többdimenziós valószínűségi változók esetén az eloszlásfüggvény a következőképpen definiált:

$$F(x) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d),$$

ami a $H = (-\infty, x_1] \times \dots \times (-\infty, x_d]$ halmazhoz rendeli annak a valószínűségét, hogy $x \in H$. $\square$

Az előbb láttuk, hogy a logkonkáv sűrűségfüggvénnyel rendelkező eloszlásoknak az eloszlásfüggvényei is logkonkávak, és következésképpen kvázikonkávak is. Most pedig a d -dimenziós normális eloszlás sűrűségfüggvényéről megmutatom, hogy logkonkáv.

2.3.8. Állítás. *Legyen X d -dimenziós normális eloszlású valószínűségi változó μ várhatóérték vektorral és Σ pozitív definit kovarianciamátrixszal. Ekkor a sűrűségfüggvénye, vagyis a*

$$f(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)$$

függvény logkonkáv.

Bizonyítás. Vegyük $f(x)$ logaritmusát:

$$\log f(x) = -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)$$

Az első két tag konstans, így nulla a második deriváltjuk, a harmadik tag pedig egy kvadratikus alak megszorozva $-\frac{1}{2}$ -el. Mivel Σ pozitív definit, ezért Σ^{-1} is pozitív definit, így a kvadratikus alak második deriváltja, azaz Hesse-mátrixa pozitív definit. Ezért az egész kifejezés $-\frac{1}{2}$ -szerese negatív definit lesz, tehát $\log f(x)$ szigorúan konkáv függvény. $\square$

2.3.9. Megjegyzés. További logkonkáv sűrűségfüggvényű eloszlások például a konvex halmazon értelmezett egyenletes eloszlás és a Dirichlet eloszlás.

3. fejezet

Többlépcsős sztochasztikus programozás

Ennek a fejezetnek a középpontjában a többlépcsős sztochasztikus lineáris programozás áll. A rövid elméleti bevezetést követően azokra az aggregációs technikákra térek ki, amelyek célja a problémák megoldásának közelítése, miközben csökkentett feladatmérettel dolgozunk.

Ezt követően bemutatom a kétlépcsős modellek felépítését, valamint a többlépcsős sztochasztikus programozás diszkrét esetét, különös tekintettel az ezekhez kapcsolódó dekompozíciós algoritmusokra, mint a Benders-dekompozíció és a beágyazott (Benders)-dekompozíció.

A fejezet szakirodalmi alapját elsősorban a következő források képezték: [1], [2], [3], [4], [6].

3.1. Bevezetés

Az olyan problémákban, ahol a véletlen szerepet játszik, a legtöbbször nem áll rendelkezésére a döntéshozónak elegendő információ, hogy optimális döntéseket hozzon egyből. Az ilyen feladatokban azonban gyakran megengedett, hogy a kezdeti döntések meghozatala után úgynevezett pótlási lépéseket tegyünk, miután a döntésekhez lényeges információból több áll rendelkezésünkre.

Számos ilyen típusú feladatra alkalmazhatók a többlépcsős sztochasztikus programozás különböző algoritmusai, megoldási módszerei, ahol a „lépcsők” száma a döntési szakaszok számára utal. A többlépcsős modellekben kezdeti döntéseket hozunk, majd felváltva megvárjuk egy véletlen esemény bekövetkezését, és ez alapján újabb döntéseket hozunk. A modell ezen struktúrája miatt megkülönböztetünk anticipatív, vagyis előrelátó, illetve adaptív változókat.

Ez formálisan leírva egy n lépcsős modellre:

Az i -edik szakaszban meghozott döntésvektor x_i függ az összes korábbi döntéstől és véletlen eseménytől, és ennek benne kell lennie egy rá vonatkozó megengedett megoldások halmazában, vagyis

$$x_i(x_1, x_2, \dots, x_{i-1}; \xi_2, \dots, \xi_i) \in H_i(x_1, x_2, \dots, x_{i-1}; \xi_2, \dots, \xi_i) \subseteq \mathbb{R}^{n_i} \quad \forall i = 1, \dots, n,$$

ahol H_i egy feltételes megengedett döntési halmaz, n_i pedig döntésvektor hossza. Továbbá minden szakaszban a döntésvektorhoz tartozik egy $c_i(\xi_1, \dots, \xi_i)^T$, és a cél, hogy maximalizáljuk a $\mathbb{E}[c_i^T x_i]$ várható értékek összegét.

Explicit leírva: Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ egy valószínűségi mező, legyenek $\xi_i : \Omega \rightarrow \mathbb{R}^{r_i}$ vektorváltozók. Defináljuk a következő vektort: $\xi = (\xi_2^T, \xi_3^T, \dots, \xi_n^T)^T : \Omega \rightarrow \mathbb{R}^R$, ahol $R = r_2 + r_3 + \dots + r_n$. Legyen továbbá ζ_i , ami a ξ vektor első i koordinátájából álló része: $\zeta_i = (\xi_2^T, \xi_3^T, \dots, \xi_i^T)^T$ az i . állapotváltozó $i = 1, \dots, n$. Ekkor az általános többlépcsős sztochasztikus lineáris program a következő formát ölti:

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & c_1^T x_1 + \sum_{i=2}^n \mathbb{E}[c_i^T(\zeta_i) x_i(\zeta_i)] \\ & A_{1,1} x_1 = b_1 \\ & A_{i,1}(\zeta_i) x_1 + \sum_{j=2}^i A_{i,j}(\zeta_i) x_j(\zeta_j) = b_i(\zeta_i), \quad (m.m.) \quad i = 2, \dots, n \\ & x_i(\zeta_i) \geq 0, \quad (m.m.) \quad i = 1, \dots, n \end{aligned} \tag{3.1}$$

Legyen $\mathcal{F}_i = \sigma(\zeta_i) \subset \mathcal{A}$, ekkor $x_i(\zeta_i) : \Omega \rightarrow \mathbb{R}^{n_i}$, $A_{i,j}(\zeta_i) : \Omega \rightarrow \mathbb{R}^{m_i \times n_j}$ $j = 1, \dots, i$, $b_i(\zeta_i) : \Omega \rightarrow \mathbb{R}^{m_i}$, $c_i(\zeta_i) : \Omega \rightarrow \mathbb{R}^{n_i}$ $\mathcal{F}_i$ -mérhetőek $i = 1, \dots, n$ -re, $\{\mathcal{F}_1, \dots, \mathcal{F}_n\}$ pedig egy filtráció, ahol $\{\emptyset, \Omega\} = \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n \subset \mathcal{A}$. Továbbá feltesszük a feladat feltételei között szereplő mátrixokról és vektorokról, hogy az egylépcsős modelleknél definiálttal megegyező módon függenek a véletlen vektoroktól.

3.1.1. Megjegyzés. Innentől az egyszerűség kedvéért (ζ_i) -ket nem fogom kiírni, valamint a problémák a következő formában lesznek:

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & \sum_{i=1}^n \mathbb{E}[c_i^T x_i] \\ & \sum_{j=1}^i A_{i,j} x_j = b_i, \quad i = 1, \dots, n \\ & x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

3.2. Aggregációs technikák

A teljes valószínűségi tér leírása a feladaton belül sokszor olyan komplex és nagy számításigényű modellekhez vezet, hogy a gyakorlatban nehéz ezekkel előrejutni. Ennek kiküszöbölésére különböző aggregációs technikákat lehet alkalmazni, amelyek

egyszerűsítik a feladatot, és bár nem feltétlenül adják vissza az optimális megoldást, gyakran jó becsléseket adnak jelentősen rövidebb futási idő mellett.

3.2.1. Definíció. (Részfiltráció): Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ egy valószínűségi mező, és legyen $\{\mathcal{F}_i\}_{i \geq 1}$ egy filtráció, azaz egy bővülő σ -algebra sorozat: $\mathcal{F}_j \subseteq \mathcal{F}_i$, ha $j \leq i$. Egy $\{\mathcal{G}_i\}_{i \geq 1}$ filtráció részfiltrációja az $\{\mathcal{F}_i\}_{i \geq 1}$ filtrációnak, ha minden i -re $\mathcal{G}_i \subseteq \mathcal{F}_i$.

Döntési aggregáció

Itt a döntést kevesebb információból hozzuk meg, vagyis az $x_i: \Omega \rightarrow \mathbb{R}^{n_i}$ változók nem csak az eredeti $\mathcal{F}_1 \subset \mathcal{F}_2 \cdots \subset \mathcal{F}_n$ filtráció megfelelő σ -algebráira, hanem ennek egy $\hat{\mathcal{F}}_1 \subset \hat{\mathcal{F}}_2 \cdots \subset \hat{\mathcal{F}}_n$ részfiltráció σ -algebráira is mérhetők. Így a döntés-aggregált probléma a következő formát ölti:

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & \mathbb{E} \left[\sum_{i=1}^n \mathbb{E}[c_i^T \mid \hat{\mathcal{F}}_i] x_i \right] \\ & \sum_{j=1}^i A_{i,j} x_j = b_i, \quad i = 1, \dots, n \\ & x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

Feltétel aggregáció

A feltétel-aggregált verzióban a kényszereket aggregáljuk az előzőhöz hasonló módon, így a probléma jelentősen kisebb nagyságú lesz és a következő alakot veszi fel:

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & \mathbb{E} \left[\sum_{i=1}^n c_i^T x_i \right] \\ & \mathbb{E} \left[\sum_{j=1}^i A_{i,j} x_j \mid \hat{\mathcal{F}}_i \right] = \mathbb{E} [b_i \mid \hat{\mathcal{F}}_i], \quad i = 1, \dots, n \\ & x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

Teljes aggregáció

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & \mathbb{E} \left[\sum_{i=1}^n \mathbb{E}[c_i^T \mid \hat{\mathcal{F}}_i] x_i \right] \\ & \mathbb{E} \left[\sum_{j=1}^i A_{i,j} x_j \mid \hat{\mathcal{F}}_i \right] = \mathbb{E} [b_i \mid \hat{\mathcal{F}}_i], \quad i = 1, \dots, n, \\ & x_i \geq 0, \quad i = 1, \dots, n. \end{aligned}$$

3.2.2. Állítás. Jelöljék a megengedett megoldások halmazát az eredeti, döntés-, feltétel- és teljesen aggregált problémában $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4$ halmazok. Ekkor a következők mondhatók el ezekről: $\mathcal{H}_2 \subseteq \mathcal{H}_1 \subseteq \mathcal{H}_3; \mathcal{H}_2 \subseteq \mathcal{H}_4 \subseteq \mathcal{H}_3$;

3.2.3. Megjegyzés. Ennek egyszerű következménye az, hogy az eredeti feladat (és a teljesen aggregált feladat) optimumára alsó becslést ad a döntés aggregált, felső becslést pedig a feltétel-aggregált feladat optima, hiszen maximalizálásnál, ha egy megengedett megoldások halmazának csak egy részhalmazán optimalizálunk, akkor az optimum legfeljebb akkora, mint az eredeti halmazon.

Bizonyítás. Elegendő $\mathcal{H}_2 \subseteq \mathcal{H}_1 \subseteq \mathcal{H}_3$ -t bizonyítani, $\mathcal{H}_2 \subseteq \mathcal{H}_4 \subseteq \mathcal{H}_3$ bizonyítása hasonló módon történik.

$\mathcal{H}_1 \supseteq \mathcal{H}_2$: A feltételek ugyanazok a döntés-aggregált problémában, mint az eredetiben, azonban a változók a rész- σ -algebrákra is mérhetők. Természetesen ekkor az eredetiekre is azok, tehát minden megengedett megoldás a döntés-aggregált feladatban megengedett megoldás lesz az eredeti feladatban is.

$\mathcal{H}_1 \subseteq \mathcal{H}_3$: A sztochasztikus programozási feladatokban x_i -k valószínűségi változók, vagyis, amikor egy feladatot megfogalmazunk, ott igazából azt mondjuk, hogy olyan valószínűségi változókat kell választani, hogy a feltételek 1-valószínűséggel teljesülnek és ezen belül a célfüggvény maximális. Ekkor a feltételes várható érték tulajdonságai miatt, $i = 1, \dots, n$ -re:

$$\sum_{j=1}^i A_{i,j}x_j = b_i \text{ (m.m.)} \Rightarrow \mathbb{E} \left[\sum_{j=1}^i A_{i,j}x_j \mid \hat{\mathcal{F}}_i \right] = \mathbb{E} [b_i \mid \hat{\mathcal{F}}_i] \text{ (m.m.)} \quad \forall \hat{\mathcal{F}}_i \subseteq \mathcal{F}_i$$

teljesül, vagyis az eredeti probléma minden megoldása benne van a feltétel-aggregált probléma megoldásai között, így az eredeti probléma megengedett megoldásainak halmaza részhalmaza a feltétel-aggregálténak. $\square$

3.3. A kétlépcsős modell

A kétlépcsős sztochasztikus lineáris programozási feladatot a többlépcsős általános formája (3.1) alapján fel lehet írni:

$$\begin{aligned} \max_{x,y} \quad & c_1^T x + \mathbb{E} [c_2^T(\xi)y(\xi)] \\ & Ax = b, \\ & B(\xi)x + C(\xi)y(\xi) = d(\xi), \\ & x \geq 0, \quad y(\xi) \geq 0. \end{aligned} \tag{3.2}$$

Egy másik megszokott felírás a következő:

$$\begin{aligned} \max_x \quad & c_1^T x + \mathbb{E}[f(x, \xi)] \\ & Ax = b \\ & x \geq 0 \end{aligned} \tag{3.3}$$

ahol $f(x, \xi) = \max_y \{c_2^T(\xi)y \mid B(\xi)x + C(\xi)y(\xi) = d(\xi), y(\xi) \geq 0\}$ az úgynevezett pótlási függvény, $f(x) = \mathbb{E}[f(x, \xi)]$ pedig a várható pótlás függvénye. Gyakorlatban az $f(x)$ függvény kiszámítása általában komoly erőfeszítéseket igényel, hiszen

ezek gyakran többdimenziós integrálok kiszámítását jelentik. Ezért fontos az a feladat, hogy felismerjük azokat a problémátípusokat, melyeknél a nagy számításigény nagymértékben leegyszerűsödik.

3.3.1. Definíció. Egy kétlépcsős sztochasztikus programozási feladatra azt mondjuk, hogy:

- rögzített pótlású, ha a pótlási változókhoz tartozó együtthatómátrix determinisztikus, vagyis $C(\xi) \equiv C$.
- teljes pótlású, ha $\forall x \in \mathbb{R}^n$ és minden ξ által felvehető vektorra a pótlási feladat megoldható.
- relatíve teljes pótlású, ha $\forall x \in H$, ahol $H = \{x \mid Ax = b\}$ és minden ξ által felvehető vektorra a pótlási feladat megoldható.

3.4. A diszkrét véges eset

Tekintsünk egy kétlépcsős sztochasztikus lineáris programozási feladatot ξ diszkrét és véges vektorváltozóval. Ilyenkor a (3.2) probléma átírható egy determinisztikus LP-változatra. Legyenek a lehetséges események $\omega_1, \omega_2, \dots, \omega_r$ és ezen események valószínűségei: $p_1, p_2, \dots, p_r$. Ekkor a kétlépcsős feladat a következő formát ölti:

$$\begin{aligned} \min_x \quad & c_0^T x + \sum_{k=1}^r p_k \min_{y_k} c_k^T y_k \\ \text{s.t.} \quad & Ax = b \\ & B_k x + C_k y_k = d_k, \quad k = 1, \dots, r \\ & x \geq 0 \\ & y_k \geq 0, \quad k = 1, \dots, r \end{aligned} \tag{3.4}$$

Ez a feladat a szimplex módszerrel megoldható, azonban vegyük észre, hogy a probléma mérete nagymértékben függ a véletlen esemény lehetséges kimeneteleinek a számától, ami gyorsan nagy számításigényű feladatokhoz vezethet. A gyorsítás érdekében érdemes megvizsgálni a feltétel mátrix alakját:

$$\begin{pmatrix} A & & & \\ B_1 & C_1 & & \\ \vdots & & \ddots & \\ B_r & & & C_r \end{pmatrix} \quad (\text{Az üres helyeken nullmátrixok vannak})$$

A mátrix alakjából látszik, hogy a különböző kimenetek csak az első döntésben (x) függenek össze. Így logikus gondolat az, ha ezeket külön feladatként oldjuk meg. A következőkben erre fogunk egy módszert látni, amit a szakirodalom Benders-dekompozíciónak vagy a sztochasztikus programozásban L-shaped method-nak is neveznek.

3.5. Benders-dekompozíció

A feladat átírása

A Benders által bevezetett dekompozíciós algoritmus a már előbb említett (3.4) féle szerkezetet használja ki. Az eljárás első lépésében az úgynevezett mesterproblémának, amelynek célja x megfelelő megválasztása, oldjuk meg egy relaxált változatát, ezzel meghatározva egy x vektort, amelyet majd a pótlási problémákba helyettesítünk. Ezen megoldás alapján megpróbáljuk megoldani a pótlási feladatokat, amelyek az adott x mellett vizsgálják a második lépésben fellépő döntéseket. Ezek a problémák nem feltétlenül megoldhatók, illetve az így kapott megoldás az eredeti probléma szempontjából még nem feltétlenül optimális.

A pótlási feladatok megoldása során azonban információt nyerünk az adott x megoldásról, és ennek segítségével vágásokat (optimalitási vagy megengedettségi vágásokat) tudunk hozzáadni a mesterproblémához. Ezután a módosított mesterproblémát ismét megoldjuk, új x megoldást nyerve, majd ezt felhasználva újra megvizsgáljuk a pótlási problémákat, és újabb vágásokat adunk a mesterproblémához.

Ezt az iteratív folyamatot mindaddig folytatjuk, amíg el nem érünk egy olyan x^* megoldást, amely az eredeti kétlépcsős probléma szempontjából optimális.

Az algoritmus során az alábbi optimalizálási feladatot vizsgáljuk:

$$\begin{aligned} \min_x \quad & c_0^T x + \sum_{k=1}^r p_k f_k(x) \\ & Ax = b \\ & x \geq 0 \end{aligned} \tag{3.5}$$

ahol $f_k(x)$ a k . eseményhez tartozó pótlási probléma, vagyis $f_k(x) = \min\{c_k^T y_k \mid C_k y_k = d_k - B_k x, y_k \geq 0\}$ $k = 1, \dots, r$.

3.5.1. Állítás. Az $\{x \mid Ax = b, Bx + Cy = d, x \geq 0, y \geq 0\} \neq \emptyset$ korlátos halmazon definiált $f(x) = \min\{c^T y \mid Cy = d - Bx, y \geq 0\}$ függvény szakaszonként lineáris, konvex és alulról korlátos függvény.

Bizonyítás. A feltevés szerint az $f(x)$ értelmezési tartománya, tehát a $H = \{x \mid Ax = b, Bx + Cy = d, x \geq 0, y \geq 0\} = \{x \mid Ax = b, x \geq 0\} \cap \{x \mid \exists y \geq 0 : Cy = d - Bx\} \neq \emptyset$ halmaz korlátos. H továbbá egy poliéder (tehát konvex), hiszen két poliéder metszete. Az $\{x \mid Ax = b, x \geq 0\}$ halmaz nyilván egy poliéder. A másik halmaz, $\{x \mid \exists y \geq 0 : Cy = d - Bx\}$, szintén poliéder, mivel ez a következő halmaz vetülete az x -változóinak terére: $\{(x, y) \mid Bx + Cy = d, y \geq 0\}$. Ez maga is egy poliéder. Poliéderek metszete pedig szintén poliéder. Egy adott $x \in H$ -ra az

$$f(x) = c_{Z_C}^T Z_C^{-1}(d - Bx), \text{ ha } Z_C^{-1}(d - Bx) \geq 0,$$

ahol Z_C egy optimális bázisa C -nek az adott x -re nézve. Legyen adott egy tetszőleges

$x \in H$, és tekintsük az ehhez tartozó optimális bázist Z_C . Amíg az x a H -n belül úgy változik, hogy az optimális bázis nem változik, az $f(x)$ függvény egy lineáris alakot vesz fel. Amint pedig az x átlépi azt a határt, ahol az optimális bázis megváltozik, a függvény egy másik lineáris szakaszra tér át. Így az $f(x)$ függvény a H -n egy véges sok lineáris szakaszból álló, azaz szakaszonként lineáris függvény.

Mivel az $f(x)$ egy korlátos halmazon értelmezett, szakaszonként lineáris, így folytonos függvény, ezért biztosan alulról korlátos is.

A konvexitás bizonyításához már láttuk, hogy konvex tartományon van értelmezve a függvény. Most pedig vegyük az $x_1, x_2 \in H$ vektorokat és a hozzájuk tartozó y_1, y_2 vektorokat. Ezekre igaz, hogy: $f(x_i) = c^T y_i$, $Cy_i = d - Bx_i$, $y_i \geq 0$, $i = 1, 2$.

$$\lambda y_1 + (1 - \lambda)y_2 \in \{Cy = d - B(\lambda x_1 + (1 - \lambda)x_2), y \geq 0\} \quad \forall \lambda \in [0, 1]$$

és

$$f(\lambda x_1 + (1 - \lambda)x_2) = \min\{c^T y \mid Cy = d - B(\lambda x_1 + (1 - \lambda)x_2), y \geq 0\} \leq$$

$$c^T(\lambda y_1 + (1 - \lambda)y_2) = \lambda f(x_1) + (1 - \lambda)f(x_2). \quad \square$$

A (3.5) feladat egy ekvivalens megfogalmazása a következő:

$$\begin{aligned} \min_x \quad & c_0^T x + \sum_{k=1}^r p_k \theta_k \\ & Ax = b \\ & \theta_k - f_k(x) \geq 0 \quad k = 1, \dots, r \\ & x \geq 0 \end{aligned}$$

3.5.2. Megjegyzés. Ez valójában így nem LP alakú, de kezelhetjük LP-ként, mert ahogy az előbb láttuk, az $f_k(x)$ szakaszonként lineáris, konvex és alulról korlátos függvény $\forall k = 1, \dots, r$ -re. Tehát létezik véges sok lineáris függvény, amelyeknek maximuma $f_k(x)$:

$$f_k(x) = \max_{m \in \{1, \dots, M\}} \rho_{k,m}(x).$$

Így a

$$\theta_k - f_k(x) \geq 0 \quad k = 1, \dots, r$$

feltételeket helyettesíthetjük a következő lineáris egyenlőtlenségekkel:

$$\theta_k - \rho_{k,m}(x) \geq 0 \quad m = 1, \dots, M, \quad k = 1, \dots, r$$

amelyek már lineáris programozási alakot eredményeznek, ezek lesznek az optimalitási vágások.

Néhány további lineáris egyenlőtlenséget is hozzá kell adni a feladathoz, annak érdekében, hogy biztosítsuk, hogy $f_k(x)$ értelmezési tartományán maradunk $\forall k = 1, \dots, r$ -re, ezek pedig a megengedettségi vágások:

$$\tau_{k,l}(x) \geq 0 \quad l = 1, \dots, L, \quad k = 1, \dots, r.$$

Algoritmus lépései

1. Találjunk θ_k -ra $k = 1, \dots, r$ -re alsó korlátot az alábbi feladat megoldásával:

$$\tilde{\theta}_k = \min\{c_k^T y_k \mid Ax = b, C_k y_k = d_k - B_k x, x \geq 0, y_k \geq 0\},$$

majd oldjuk meg a

$$\min\{c_0^T x + p^T \theta \mid Ax = b, \theta \geq \tilde{\theta}, x \geq 0\} \text{ feladatot,}$$

ahol $p^T = (p_1, \dots, p_r)$, $\theta = (\theta_1, \dots, \theta_r)^T$, $\tilde{\theta} = (\tilde{\theta}_1, \dots, \tilde{\theta}_r)^T$. Ebből kapjuk az $(\hat{x}, \hat{\theta})$ párt.

2. A kapott $\hat{x}$ alapján írjuk fel a pótlási problémákat és azok duálisait:

$$f_k(\hat{x}) = \min\{c_k^T y_k \mid C_k y_k = d_k - B_k \hat{x}, y_k \geq 0\} =$$

$$\max\{u_k^T (d_k - B_k \hat{x}) \mid u_k^T C_k \leq c_k\}, \quad k = 1, \dots, r$$

(Ez itt most egy egyszerűsítés jelölésben, mert nyilván, ha az egyik nem oldható meg, akkor ez nem igaz. Tehát vehetjük úgy, hogy a célfüggvény értelmezve van a nem megoldható problémák esetében is, és ilyenkor $+\infty$ -t vagy $-\infty$ -t vesz fel, attól függően, hogy primál vagy duál a nem megoldható feladat). A duális feladatok megoldása után két lehetőség állhat elő: minden pótlási problémában:

a) A primál feladat nem oldható meg és a duál nem korlátos.

b) A primál feladat megoldható és a primál- és duál optimumok megegyeznek.

Ha van olyan pótlási probléma, amire az a) eset következik be, akkor a 3. lépés jön, egyébként pedig a 4.

3. Ha valamelyik pótlási feladat nem megoldható, akkor az $\hat{x}$ biztosan nem megengedett megoldás. Ilyenkor a pótlási feladat duálisa nem korlátos, vagyis létezik egy $\hat{u}_k$ végtelen növelő irány: $\hat{u}_k^T (d_k - B_k \hat{x}) > 0$ és $C_k^T \hat{u}_k \leq 0$. Ezt a simplex módszer megadja. Ekkor ha a $\hat{u}_k^T (d_k - B_k \hat{x}) \leq 0$ egyenlőtlenséget hozzáadjuk a főproblémához, akkor a főproblémából azt az $\hat{x}$ -et, amiből ezt kaptuk kizártuk. Az ilyen egyenlőtlenség hozzáadását a főfeladathoz megengedettségi vágásnak nevezzük. Folytatjuk az 5. lépéssel.

4. Ha az összes pótlási probléma megoldható, akkor: $f_k(\hat{x}) = \hat{u}_k^T (d_k - B_k \hat{x})$, és tetszőleges x -re $f_k(x) \geq \hat{u}_k^T (d_k - B_k x)$, hiszen a pótlási probléma duálisának feltételei nem függenek x -től. A $\theta_k - f_k(x) \geq 0$ feltételek az eredeti feladatban és az előző feltételek implikálják a következő egyenlőtlenséget: $\theta_k \geq \hat{u}_k^T (d_k - B_k x)$. Ha ezek egy $(\hat{x}, \hat{\theta})$ párra $k = 1, \dots, r$ -re mind teljesülnek, akkor tudjuk, hogy az eredeti feladatnak egy optimális $x^* = \hat{x}$ megoldását találtuk meg. Ez azért van így, mert az algoritmus futása során végig igaz, hogy a megengedett megoldások halmazának részhalma az eredeti probléma megengedett megoldásainak halmaza, tehát annak az optima legalább annyi, mint az algoritmus bármely iterációjában kapott optimum. Tehát, ha az algoritmus során talált $(\hat{x}, \hat{\theta})$ pár megengedett az eredeti problémára nézve, akkor optimális is. Különben van olyan $\emptyset \neq K \subseteq \{1, \dots, r\}$,

hogy $\forall k \in K \hat{\theta}_k < f_k(\hat{x})$. Ekkor a főproblémához adjuk a következő korlátokat: $\theta_k \geq \hat{u}_k^T(d_k - B_k x)$ $k \in K$ (ezzel levágjuk a nem optimális $(\hat{x}, \hat{\theta})$ párt), majd folytatjuk az 5. lépéssel.

5. Oldjuk meg a főproblémát a hozzáadott vágásokkal együtt, hogy megkapjuk a $(\hat{x}, \hat{\theta})$ párt, majd térjünk vissza a 2. lépéshez.

3.5.3. Megjegyzés. A pótlási problémák duálisairól mindig fel van téve, hogy létezik megengedett megoldásuk, hiszen azok az esetek, amikor nincs, nem érdekesek számunkra, mert ilyenkor a primál vagy nem megoldható, vagy nem korlátos.

3.5.4. Megjegyzés. Vegyük észre, hogy ha a feladat rögzített pótlású feladat, és a pótlási lépések célfüggvényei is fixek, akkor a pótlási problémák duálisai csak a célfüggvényükben térnek el (a valószínűségek kiemelhetők még a primálnál a célfüggvényből, hiszen ezek csak számok, így az optimumhelyet nem befolyásolják). Vagyis, ha az elsőre találunk egy optimális megoldást, akkor a hozzá tartozó bázis az összes többire is duálmegengedett. Így minden pótlási problémához a bázis alapján tudunk számolni egy primál vektort: $y_k = Z^{-1}(d_k - B_k x^i)$ módon, ahol Z a duál megengedett bázis C -ben. Ha $y_k \geq 0$, akkor ez primál-megengedett is, így a simplex módszer elindítása nélkül találtunk egy optimális megoldást az x^i -hez tartozó k . pótlási problémára. Az ötlet innen már világos: mindig mikor megtalálunk egy megoldást, akkor elvégezzük az előbbi ellenőrző lépést a még megoldatlan pótlási problémákra, majd egy olyan megoldásával folytatjuk ugyanezt az eljárást, amire nem teljesült a primál-megengedettségi feltétel. A szakirodalomban ezt az ötletet bunchingnak szokták hívni.

Algoritmus helyessége

3.5.5. Állítás. *Az algoritmus véges sok lépésben megtalálja a korábban említett tulajdonságokkal rendelkező x^* -ot.*

Bizonyítás. Elegendő arra az esetre belátni, hogy csupán egyfajta realizációja van a véletlen eseménynek, a többféle realizációs eset hasonló módon történik. Nézzük egy adott $\hat{x}$ -re a lehetséges eseteket. Ha az ehhez tartozó pótlási probléma nem megoldható, vagyis a duálisa nem korlátos, akkor létezik $\hat{u}$ végtelen növelő irány ($\hat{u}^T(d - B\hat{x}) > 0$ és $C^T \hat{u} \leq 0$). Ezt a növelő irányt a simplex módszer megadja, és azt is tudjuk, hogy ez a C iránykúpjának generáló vektora. Ilyenből véges sok van, tehát amikor eljutunk addig a pontig, hogy már csak optimalitási vágást adhunk a főproblémához, tudjuk, hogy véges sok megengedettségi vágással frissítettük a főproblémát, hiszen minden lépésben legalább 1-et levágtunk a C iránykúpját generáló vektorok közül. Ha a pótlási probléma megoldható, akkor legyen az optimális duál megoldás $\hat{u}$. Mivel véges sok duál megengedett bázis van, ezért véges sok a duál

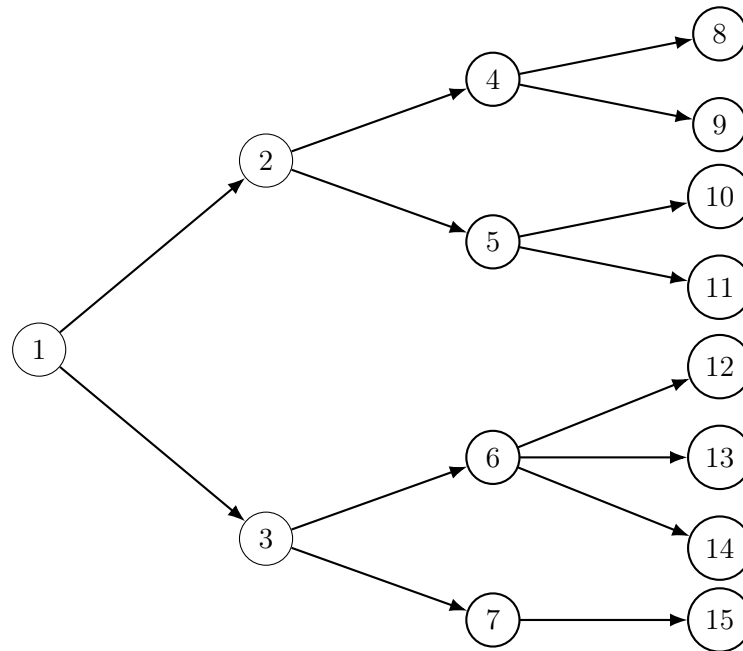
bázismegoldás is, vagyis a $\theta \geq \hat{u}^T(d_k - B_k x)$ típusú optimalitási vágásból csupán véges sokat tudunk a főproblémához adni, hiszen mindegyik bázismegoldáshoz legfeljebb 1 tartozik. Amikor eljutunk oda, hogy a főproblémából kapott $(\hat{x}, \hat{\theta})$ párra igaz a $\hat{\theta} \geq \hat{u}_k^T(d_k - B_k \hat{x})$, leáll az algoritmus és csak véges lépést tettünk. $\square$

3.6. Beágyazott dekompozíció

Feladat felírása Szenáriófával

A Benders-dekompozíció általánosítása a beágyazott dekompozíció. Ez az algoritmus a kettőnél több lépcsővel rendelkező sztochasztikus lineáris programozási feladatokra is alkalmazható a következő feltételek mellett:

- A ξ_i vektorváltozók diszkrét eloszlásúak, és véges számú értéket vehetnek fel minden $i = 1, \dots, n$ esetén, következésképp a ζ_i változók is diszkrét, véges értékkészlettel, minden $i = 1, \dots, n$ -re. Ilyenkor a véletlen vektor összes realizációját egy Szenáriófán szoktuk reprezentálni.



Ez például egy négylépcsős probléma döntési fája, ahol az első szinten hozzuk meg az x_1 döntést. Ezt követően megfigyeljük a véletlen ξ_2 kimenetelét, amely alapján a folyamat a második szint 2-es vagy 3-as csúcsába kerül. Az aktuális csúcs ismeretében hozzuk meg a következő döntést, x_2 -t. Ezután következik a ξ_3 véletlen vektor realizációja. Ha az előző szinten a 2-es csúcsba jutottunk, akkor most a 4-es vagy 5-ös csúcsok valamelyikébe lépünk tovább. Ha viszont a 3-as csúcsba kerültünk, akkor a 6-os vagy 7-es csúcs lehet a következő állapot. A megfelelő csúcs ismeretében meghozzuk a harmadik döntést, x_3 -at. Az x_4

vektor megválasztása hasonlóan történik, vagyis figyelembe vesszük ξ_4 -et és, hogy jelenleg melyik csúcán álltunk a scenáriófának.

- A másik kikötés pedig, hogy a lépcsők közötti kapcsolatokat kifejező feltételek, csak egymás utáni lépcsőkre vonatkoznak, vagyis a feladat a következő formát ölti:

$$\begin{aligned} \min_{x_1, \dots, x_n} \quad & c_1^T x_1 + \sum_{i=2}^n \mathbb{E}[c_i^T(\zeta_i) x_i(\zeta_i)] \\ & A_1 x_1 = b_1 \\ & B_i(\zeta_i) x_{i-1} + A_i(\zeta_i) x_i(\zeta_i) = b_i(\zeta_i), \quad (m.m.) \quad i = 2, \dots, n \\ & x_i(\zeta_i) \geq 0, \quad (m.m.) \quad i = 1, \dots, n \end{aligned} \quad (3.6)$$

Az algoritmus megértéséhez szükséges bevezetni néhány jelölést:

Legyen a scenáriófa csúcsainak halmaza $\mathcal{N} = \{1, \dots, r\}$, az i . szint csúcsainak halmaza $\mathcal{N}_i$, valamint legyen $|\mathcal{N}| = r$, $|\mathcal{N}_i| = r_i$, és $r_1 + r_2 + \dots + r_n = r$. Jelölje továbbá a k csúcs szülőjét $h(k)$, gyermekeit $c(k)$, $g(k)$ pedig a k gyökerű részfa csúcsainak halmazát. Mivel ξ véletlen vektor diszkrét, véges értékészletű véletlen vektor, ezért a (3.6) feladat átírható a következő formára:

$$\begin{aligned} \min \quad & c_1^T x_1 + \sum_{i=2}^n \left(\sum_{k \in \mathcal{N}_i} p_k c_k^T x_k \right) \\ & A_1 x_1 = b_1 \\ & B_k x_{h(k)} + A_k x_k = b_k, \quad k \in \mathcal{N} \\ & x_k \geq 0, \quad k \in \mathcal{N} \end{aligned} \quad (3.7)$$

Az r csúcsú scenáriófának tetszőleges $k \in \mathcal{N}$ számmal indexelt csúcsához tartozó p_k valószínűség azt fejezi ki, hogy mekkora valószínűséggel jutunk el k csúcsba, a gyökérből indulva. A scenáriófa i . szintjén lévő csúcsokra ez azt jelenti, hogy a ξ_i véletlen vektor mekkora valószínűséggel veszi fel a csúcshoz rendelt értéket. Ezen valószínűségekre igaz, hogy tetszőleges, nem utolsó szinten levő csúcs valószínűsége a gyermekei valószínűségének az összege, valamint minden szinten a valószínűségek összege 1, vagyis:

$$p_k = \sum_{j \in c(k)} p_j, \quad \forall k \in \mathcal{N} \setminus \mathcal{N}_n; \quad \sum_{k \in \mathcal{N}_i} p_k = 1$$

A (3.7) feladat alternatív felírása a következő:

$$\left. \begin{aligned} F_1 &= \min \quad c_1^T x_1 + \sum_{k \in c(1)} p_k F_k(x_1) \\ &A_1 x_1 = b_1 \\ &x_1 \geq 0 \end{aligned} \right\}$$

Tetszőleges $k \in \mathcal{N} \setminus \{1\}$ -re:

$$\left. \begin{aligned} F_k(x_{h(k)}) &= \min \quad c_k^T x_k + \sum_{j \in c(k)} \frac{p_j}{p_k} F_j(x_k) \\ A_k x_k &= b_k - B_k x_{h(k)} \\ x_k &\geq 0 \end{aligned} \right\}$$

Vagy LP formában:

$$\left. \begin{aligned} F_k(x_{h(k)}) &= \min \quad c_k^T x_k + \sum_{j \in g(k) \setminus k} \frac{p_j}{p_k} c_j^T x_j \\ A_k x_k &= b_k - B_k x_{h(k)} \\ B_j x_{h(j)} + A_j x_j &= b_j, \quad j \in g(k) \setminus k \\ x_j &\geq 0 \quad j \in g(k) \end{aligned} \right\}$$

A Benders-dekompozíció során egyetlen mesterproblémánk van, amelyhez több pótlási probléma tartozik. Azonban többlépcsős struktúrák esetében minden egyes csúcs paraméteres mesterproblémaként értelmezhető az öt megelőző szülő döntési változók függvényében.

Ezért, ha egy adott k csúcs szülőjét, $x_{h(k)}$ -t már rögzítettük, akkor a csúcshoz tartozó relaxált mesterprobléma, $\text{Rel}(k, x_{h(k)})$ a következő alakot ölti:

$$\left. \begin{aligned} \min \quad & c_k^T x_k + \theta_k \\ & A_k x_k = b_k - B_k x_{h(k)} \\ & u_{k,l}^T x_k \geq \tau_{k,l}, \quad l = 1, \dots, l_k \\ & v_{k,m}^T x_k + \theta_k \geq \rho_{k,m}, \quad m = 1, \dots, m_k \\ & x_k \geq 0 \end{aligned} \right\}$$

A harmadik és negyedik sorban található feltételek a relaxált mesterproblémához eddig hozzáadott megengedettségi és optimalitási vágások. A $\text{Rel}(k, x_{h(k)})$ duálisát a következő mátrixos alak segítségével fel tudjuk írni:

$$\begin{aligned} & \begin{array}{c} \bigcirc \\ \wedge | \\ \left[\begin{array}{cc} x_k & \theta_k \end{array} \right] \end{array} \\ \\ & \begin{array}{ccc} 0 \leq \left[\begin{array}{c} w_k \\ y_k \\ z_k \end{array} \right] & \left[\begin{array}{cc} A_k & 0 \\ U_k & 0 \\ V_k & 1 \end{array} \right] \geq & \left[\begin{array}{c} b_k - B_k x_{h(k)} \\ \tau_k \\ \rho_k \end{array} \right] \\ & \wedge | \quad \wedge | \\ & \left[\begin{array}{cc} c_k & 1 \end{array} \right] \end{array} \end{aligned}$$

Tehát $\text{Rel}(k, x_{h(k)})$ duálisa:

$$\left. \begin{aligned} \max \quad & w_k^T (b_k - B_k x_{h(k)}) + y_k^T \tau_k + z_k^T \rho_k \\ & w_k^T A_k + \sum_{l=1}^{l_k} u_{k,l} (y_k)_l + \sum_{m=1}^{m_k} v_{k,m} (z_k)_m \leq c_k \\ & \sum_{m=1}^{m_k} (z_k)_m = 1 \\ & (y_k)_l \geq 0, \quad l = 1, \dots, l_k \\ & (z_k)_m \geq 0, \quad m = 1, \dots, m_k \end{aligned} \right\}$$

Itt $u_{k,l}$ az U_k mátrix l . sorát, a $v_{k,m}$ pedig a V_k mátrix m . sorát jelenti.

3.6.1. Megjegyzés. A $\text{Rel}(k, x_{h(k)})$ jelölésében, ha $k = 1$, akkor az $x_{h(1)} = 0$ teljesül, valamint $\forall k \in \mathcal{N}_i$ -re m_k és l_k mindvégig 0 $\text{Rel}(k, x_{h(k)})$ -ban.

Megengedettségi vágások

Ha egy tetszőleges k csúcshoz tartozó, $\text{Rel}(k, \hat{x}_{h(k)})$ feladatból kapott $\hat{x}_k$ -ra $\exists j \in c(k)$, amelyre $\text{Rel}(j, \hat{x}_k)$ nem oldható meg, akkor ennek a duálisában van egy végtelen növelő irány. Ilyenkor egy ezt kizáró megengedettségi vágást adunk a $\text{Rel}(k, \hat{x}_{h(k)})$ feladathoz. A végtelen növelő irány az az $(\dot{w}_j, \dot{y}_j)$, $\dot{y}_j \geq 0$, amelyre:

$$\dot{w}_j^T A_j + \sum_{l=1}^{l_j} u_{j,l} (\dot{y}_j)_l \leq 0, \quad \text{és} \quad \dot{w}_j^T (b_j - B_j \hat{x}_k) + \dot{y}_j^T \tau_j > 0$$

A $\dot{z}_j$ vektort azért lehet kihagyni, mert $\sum_{m=1}^{m_j} (\dot{z}_j)_m = 1$ és $(\dot{z}_j)_m \geq 0$, $m = 1, \dots, m_j$, vagyis $\dot{z}_j$ minden koordinátája 0. Az $\hat{x}_k$ megoldást, amiből a növelő irányt kaptuk, ki tudjuk zárni a következő feltétel segítségével:

$$\dot{w}_j^T (b_j - B_j x_k) + \dot{y}_j^T \tau_j \leq 0$$

Ezt átírva kapjuk a:

$$u_{k,l_k+1}^T x_k \geq \tau_{k,l_k+1} \quad (3.8)$$

feltételt, ahol

$$u_{k,l_k+1}^T = \dot{w}_j^T B_j,$$

valamint

$$\tau_{k,l_k+1} = \dot{w}_j^T b_j + \dot{y}_j^T \tau_j.$$

Az $l_k + 1$ azt jelenti, hogy eddig l_k darab megengedettségi vágása volt k -nak, de ezekhez most hozzáveszünk egyet, ez lesz az $l_k + 1$.

A (3.8) feltételt adjuk hozzá a $\text{Rel}(k, \hat{x}_{h(k)})$ -hoz.

Optimalitási vágások

Ha egy tetszőleges k csúcs relaxált feladatára kapott $\hat{x}_k$ mellett $\text{Rel}(j, \hat{x}_k)$ megoldható $\forall j \in c(k)$ esetén, akkor megvizsgáljuk, hogy adható-e optimalitási vágás a $\text{Rel}(k, \hat{x}_{h(k)})$ feladathoz. Minden $j \in c(k)$ esetén oldjuk meg a $\text{Rel}(j, \hat{x}_k)$ problémát, és legyen $(\hat{x}_j, \hat{\theta}_j)$ az ebből kapott optimális megoldás, valamint $(\hat{w}_j, \hat{y}_j, \hat{z}_j)$ a hozzá tartozó duális megoldás. Tetszőleges j -re ezekről elmondható, hogy:

$$\hat{F}_j(\hat{x}_k) = c_j^T \hat{x}_j + \hat{\theta}_j = \hat{w}_j^T (b_j - B_j \hat{x}_k) + \hat{y}_j^T \tau_j + \hat{z}_j^T \rho_j$$

Továbbá tudjuk, hogy:

$$\hat{F}_j(x_k) \geq \hat{w}_j^T (b_j - B_j x_k) + \hat{y}_j^T \tau_j + \hat{z}_j^T \rho_j$$

igaz tetszőleges x_k -ra, hiszen $\text{Rel}(j, x_k)$ duálisának feltételei nem függenek x_k -tól, így $(\hat{w}_j, \hat{y}_j, \hat{z}_j)$ megengedett megoldás, de x_k -hoz nem feltétlenül ez az optimális a $\text{Rel}(j, x_k)$ duálisára. Ezek alapján az alábbi optimalitási vágás fogalmazható meg a $\text{Rel}(k, x_{h(k)})$ problémához:

$$\theta_k \geq \sum_{j \in c(k)} \frac{p_j}{p_k} [\hat{w}_j^T (b_j - B_j x_k) + \hat{y}_j^T \tau_j + \hat{z}_j^T \rho_j]. \quad (3.9)$$

Ezután ellenőrizni kell, hogy az $(\hat{x}_k, \hat{\theta}_k)$ megoldásvektor teljesíti-e a fenti optimalitási vágást. Amennyiben teljesül az egyenlőtlenség, úgy a vágás hozzáadása szükségtelen, mivel az nem szűkíti tovább a megengedett megoldások halmazát. Ha azonban az egyenlőtlenség nem teljesül, akkor a vágást hozzá kell adni a $\text{Rel}(k, \hat{x}_{h(k)})$ feladathoz, hogy kizárjuk az aktuális nem optimális megoldást.

Az optimalitási vágás ekvivalens felírása:

$$v_{k, m_k+1}^T x_k + \theta_k \geq \rho_{k, m_k+1},$$

ahol:

$$v_{k, m_k+1} = \sum_{j \in c(k)} \frac{p_j}{p_k} \hat{w}_j^T B_j,$$

és

$$\rho_{k, m_k+1} = \sum_{j \in c(k)} \frac{p_j}{p_k} [\hat{w}_j^T b_j + \hat{y}_j^T \tau_j + \hat{z}_j^T \rho_j]$$

Algoritmus lépései

1. $\forall k \in \mathcal{N}$ -re eleinte: $\text{Rel}(k, \hat{x}_{h(k)})$ -ben az $l_k = m_k = \omega_k = 0$, ahol ω_k egy újonnan bevezett bináris logikai változó, mely 1 értéket vesz fel, ha a $\text{Rel}(k, \hat{x}_{h(k)})$ -nak van olyan megoldása jelenleg, amit tudunk használni, ha a k csúccsal találkozunk az algoritmus során. Amennyiben ez nem igaz, ω_k 0-t vesz fel. Az algoritmus megkezdéséhez még állítsuk be az $i = 1$ -et, a $\hat{x}_{h(1)} = 0$ -t és adjuk hozzá $\theta_k = 0$ feltételt a $\text{Rel}(k, x_{h(k)})$ problémához minden k -ra. Folytassuk az algoritmust a 2. szakaszával.

2. Ha $i < n$, akkor a 3. lépés következik, egyébként a 4.

3. Definiáljuk a következő ciklust:

Menjünk végig a szcenáriófa i . szintjének, vagyis $\mathcal{N}_i$ -nek a csúcsain a következő lépéssorozattal:

- Ha $\omega_k = 0$, akkor oldjuk meg a $\text{Rel}(k, \hat{x}_{h(k)})$ feladatot és folytassuk az iteráció következő lépésével, egyébként pedig lépünk a következő iterációba.
- Ha $\text{Rel}(k, \hat{x}_{h(k)})$ nem oldható meg, akkor írjuk fel a megengedettségi vágást, majd lépünk az 5. lépésébe az algoritmusnak. Egyébként folytatjuk.
- Tároljuk az $(\hat{x}_k, \hat{\theta}_k)$ megoldást, továbbá, ha $i = n$, akkor a $(\hat{w}_k, \hat{y}_k, \hat{z}_k)$ duális megoldást is.
- Beállítjuk az $\omega_k = 1$ -et és az $\omega_j = 0$ -t minden $j \in g(k) \setminus i$ -re.

Ha cikluson végigjutottunk, az azt jelenti, hogy egyik vizsgált csúcsnál sem mentünk át az algoritmus 5. lépésébe. Ilyenkor ha $i = n$, akkor a 2. számú lépéssel folytatjuk, egyébként pedig $i = i + 1$ -et beállítjuk és megismételjük a 3. lépést.

4. Ez a lépés az optimalitási vágások hozzáadására szolgál. Legyen κ az a logikai változó, ami igaz, ha még nem adtunk optimalitási vágást semelyik feladathoz sem, az $(i - 1)$. szinten. A lépés során a következő ciklust futtatjuk:

For $k \in \mathcal{N}_{i-1}$:

- Ellenőrizzük, hogy (3.9) teljesül-e $(\hat{x}_k, \hat{\theta}_k)$ -ra. Ha igen, akkor átlépünk a ciklus következő iterációjába. Egyébként pedig folytatjuk a következő lépéssel.
- κ -t hamisra állítjuk.
- Ha $m_k = 0$, akkor a $\theta_k = 0$ feltételt eltávolítjuk $\text{Rel}(k, \hat{x}_{h(k)})$ -ből.
- A gyerekekből generált optimalitási vágást hozzáadjuk a $\text{Rel}(k, x_{h(k)})$ feladathoz.
- Megoldjuk a $\text{Rel}(k, x_{h(k)})$ -t és eltároljuk a hozzá tartozó $(\hat{w}_k, \hat{y}_k, \hat{z}_k)$ duális megoldást.

Ez a ciklus mindig végig fog menni, hiszen ha egy optimalitási vágást hozzáadásunk egy problémához, akkor annak a megoldása továbbra is megengedett marad.

Miután végimentünk a cikluson ellenőrizzük a következőt:

- Ha $i = 1$ és a κ igaz, akkor az azt jelenti, hogy nem „adtunk” optimalitási vágást az első csúcsból visszafelé, hiszen az a (3.9) feltételt teljesítette az $(\hat{x}_1, \hat{\theta}_1)$, tehát megtaláltuk az optimális megoldást $\text{Rel}(1, x_{h(1)})$ -re. ami a legelső lépcső.
- Különben, ha $i > 1$, akkor legyen $i = i - 1$ és ismételjük meg a 4. lépést. Egyébként pedig a 2. lépés következik.

5. Ehhez a lépéshez egy k csúccsal, és egy ehhez tartozó megengedettségi vágással érkezünk. Két esetet tudunk megfogalmazni ekkor:

- Ha $k = 1$, akkor a többlépcsős probléma nem megoldható
- Egyébként pedig vegyük a $h(k)$ -hoz tartozó $\text{Rel}(h(k), \hat{x}_{h(h(k))})$ -t. Ehhez adjuk

hozzá a kiszámolt megengedettségi vágást és állítsuk ω_j -t 0-ra $\forall j \in g(h(k))$. Ezután pedig oldjuk meg $\text{Rel}(h(k), \hat{x}_{h(k)})$ -t. Ha ez nem megoldható, akkor erre is számoljunk egy megengedettségi vágást és térjünk vissza az algoritmus 5. lépéséhez. Ha pedig megoldható, akkor állítsuk be $\omega_{h(k)}$ -t 1-re és folytassuk a 2. lépéssel.

4. fejezet

Robusztus lineáris optimalizálás

A robusztus optimalizálásnál a bizonytalanság a sztochasztikus programozással ellentétben nem valószínűségi eloszlások formájában jelenik meg.

A feladat paramétereinek együttes eloszlását ugyanis nem ismerjük pontosan, csupán egy bizonytalansági halmazt feltételezünk, amelyből ezek a paraméterek származnak. Az optimalizálás célja ebben az esetben egy olyan megoldás megtalálása, amely minden lehetséges paraméterérték mellett teljesíti a feltételeket, és ezek közül a célfüggvény szerint legkedvezőbb. A fejezethez a [7], [9] forrásokat használtam.

4.0.1. Megjegyzés. Ez a megközelítés nagymértékben hasonlít az olyan sztochasztikus programozási modellekre, ahol a valószínűségi feltételeknek 1 valószínűséggel való teljesülését követeljük meg. A robusztus optimalizálás azonban ennél még konzervatívabb, mivel a megoldás érvényességét minden lehetséges adatrealizációra garantálni kell, nem csupán 1 valószínűséggel.

4.1. Alapfogalmak

4.1.1. Definíció. Tekintsünk egy lineáris optimalizálási feladatot a következő formában:

$$\min_x \{c^T x : Ax \leq b\}$$

Ennek a bizonytalan megfelelője az ugyanilyen alakban felírható feladatok halmaza, ahol a feladat paramétereit egy úgynevezett bizonytalansági halmazból származnak, vagyis $(A, b, c) \in \mathcal{H}$.

4.1.2. Megjegyzés. A fejezet során feltesszük, hogy a bizonytalansági halmazban lévő adatok felírása külön-külön $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ és $c \in \mathbb{R}^n$ paraméterekre a következő módon történik:

$$c(\xi) = c_0 + \sum_{s=1}^S \xi_s c_s$$

$$A(\xi) = A_0 + \sum_{s=1}^S \xi_s A_s$$

$$b(\xi) = b_0 + \sum_{s=1}^S \xi_s b_s,$$

ahol az $A_i \in \mathbb{R}^{m \times n}$, $b_i \in \mathbb{R}^m$ és $c_i \in \mathbb{R}^n$ mind ismertek és determinisztikusak $i = 0, \dots, S$ -re, a $\xi \in \mathcal{S} \subset \mathbb{R}^S$ meg egy úgynevezett perturbációs vektor. A bizonytalanságot a perturbációs vektor nem pontos ismerete idézi elő, ezért $\forall \xi \in \mathcal{S}$ -re megköveteljük a feltételek teljesülését.

4.1.3. Definíció. Egy $x \in \mathbb{R}^n$ vektor robusztus megengedett megoldásnak nevezünk, ha $\forall (A, b, c) \in \mathcal{H}$ hármásra az $Ax \leq b$ feltételek teljesülnek.

4.1.4. Definíció. Adott $x \in \mathbb{R}^n$ robusztus megengedett megoldáshoz tartozó $\hat{c}(x) = \max_{(A, b, c) \in \mathcal{H}} c^T x$ értéket robusztus célfüggvényértéknek nevezzük.

4.1.5. Definíció. Egy bizonytalan lineáris optimalizálási feladat robusztus megfelelőjének vagy robusztus ekvivalensének nevezzük a következő feladatot:

$$\min_{x \in \mathbb{R}^n} \{\hat{c}(x) : Ax \leq b\} = \min_{x \in \mathbb{R}^n} \left\{ \max_{(A, b, c) \in \mathcal{H}} cx : Ax \leq b, \quad \forall (A, b, c) \in \mathcal{H} \right\}$$

Vagyis azt az $(x^* \in \mathbb{R}^n)$ vektort keressük, amelyre igaz, hogy a bizonytalansági halmazból minden $x \in \mathbb{R}^n$ -hez a számunkra legrosszabb (A, b, c) hármast választva, minden más $x \in \mathbb{R}^n$ esetén a „legrosszabb” esetben felvett célfüggvény érték nem kisebb, mint az x^* esetén.

4.2. Fontos megfigyelések

Robusztus optimalizálási feladatok során a fő nehézséget az jelenti, hogy a feladat legtöbb esetben egy félig végtelen optimalizálási probléma, vagyis véges sok változója, de végtelen sok feltétele van a bizonytalansági halmaz megléte miatt.

Ezért a fő célunk a feladat kezelhetősége érdekében, hogy a bizonytalansági halmazt megfelelően tudjuk modellezni véges sok feltétellel.

Az első hasznos megfigyelés az, hogy a célfüggvényt tekinthetjük determinisztikusnak, hiszen a

$$\min_x \{c^T x : Ax \leq b, \quad (A, b, c) \in \mathcal{H}\}$$

feladat átírható a következő formára:

$$\begin{aligned} \min_{x, t} \quad & t \\ & c^T x - t \leq 0 \\ & Ax \leq b \\ & \forall (A, b, c) \in \mathcal{H}. \end{aligned}$$

Sőt a jobb oldalt is tekinthetjük annak, hiszen $Ax \leq b$ -t helyettesíteni tudjuk a $z = 1, \quad A_i x - b_i z \leq 0, \quad i = 1, \dots, m$ feltételekkel.

Legyen $\hat{\mathcal{H}} = \mathcal{H}_1 \times \mathcal{H}_2 \times \dots \times \mathcal{H}_m$, ahol $\mathcal{H}_i$ a $\mathcal{H}$ halmaz vetülete az i . feltételben szereplő adatok terére, azaz $\mathcal{H}_i = \{[A_i, b_i] : [A, b] \in \mathcal{H}\}$. A második megfigyeléshez pedig tekintsük a következő példát:

4.2.1. Példa. *Tegyük fel, hogy az alábbi két bizonytalan korlátból álló rendszert vizsgáljuk:*

$$x_1 \geq b_1,$$

$$x_2 \geq b_2,$$

ahol a bizonytalansági halmaz a következő:

$$\mathcal{H} = \{(b_1, b_2) \mid b_1 \geq 0, b_2 \geq 0, b_1 + b_2 \leq 1\}.$$

A robusztus megfelelő feltételrendszerében a korlátoknak minden lehetséges $(b_1, b_2) \in \mathcal{H}$ esetén teljesülniük kell:

$$x_1 \geq b_1, \quad x_2 \geq b_2 \quad \forall (b_1, b_2) \in \mathcal{H}.$$

Ez egy végtelen sok korlátból álló rendszer, azonban a rendszer ekvivalens az alábbi két korláttal:

$$x_1 \geq \max_{(b_1, b_2) \in \mathcal{H}} b_1 = 1, \quad x_2 \geq \max_{(b_1, b_2) \in \mathcal{H}} b_2 = 1.$$

A bizonytalansági halmaz vetületei az egyes feltételek terére:

$$\mathcal{H}_1 = \{b_1 \mid 0 \leq b_1 \leq 1\}, \quad \mathcal{H}_2 = \{b_2 \mid 0 \leq b_2 \leq 1\}.$$

Ha a teljes bizonytalansági halmazt kiterjesztjük a $\hat{\mathcal{H}} = \mathcal{H}_1 \times \mathcal{H}_2$ halmazra, azaz eltekintünk az $b_1 + b_2 \leq 1$ feltételtől, akkor a robusztus ekvivalens továbbra is:

$$x_1 \geq 1, \quad x_2 \geq 1.$$

4.2.2. Következmény. *Mivel a robusztus optimalizálási feladatok megoldhatóságát a „legrosszabb” esetben is megköveteljük, ezért a robusztus ekvivalens bizonytalansági halmazát kiterjeszthetjük a feltételek vetületeinek direkt szorzatára. Nem számít, hogy a különböző korlátok bizonytalan paramétereik között van-e összefüggés. Tehát a bizonytalanság soronként szeparálható.*

4.3. Robusztus feladat átírása kezelhető feladatra

Ebben a szakaszban három esetre koncentrálok, ahol a robusztus feltételek átalakíthatók hatékonyan kezelhető formába. Ezek az intervallum vagy doboz, az ellipszoidális, és a poliéderes bizonytalanság.

4.3.1. Megjegyzés. Az átalakítások során mindig egyetlen feltételre koncentrálok,

és azt mutatom meg, hogy miként vezethető vissza egy kezelhető, determinisztikus alakra. Ennek oka, hogy robusztus optimalizálásban a feltételek szeparálhatók.

Az egyszerűbb jelölés érdekében a továbbiakban az a_s és b_s ($s = 0, \dots, S$) jelölések a robusztus feltételek perturbált alakjában szereplő determinisztikus adatokat jelölik:

- az $a_s \in \mathbb{R}^n$ vektor a bal oldali mátrix adott sorának perturbációs felírásában szereplő determinisztikus vektor,
- míg a $b_s \in \mathbb{R}$ skalár ugyanazon sor jobb oldali tagjának determinisztikus komponense.

Intervallum bizonytalanság

Tegyük fel, hogy a perturbációs halmaz egy S dimenziós hiperkocka, azaz:

$$\mathcal{S} := \{\xi \in \mathbb{R}^S : \|\xi\|_\infty \leq 1\}.$$

Ekkor tetszőleges robusztus feltétel:

$$a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \leq b_0 + \sum_{s=1}^S \xi_s b_s \quad \forall \xi \in \mathcal{S}$$

Átírható:

$$\sum_{s=1}^S \xi_s (a_s^T x - b_s) \leq b_0 - a_0^T x \quad \forall |\xi_s| \leq 1$$

Ez:

$$\max_{|\xi_s| \leq 1} \sum_{s=1}^S \xi_s (a_s^T x - b_s) \leq b_0 - a_0^T x$$

Vegyük észre, hogy:

$$\max_{|\xi_s| \leq 1} \sum_{s=1}^S \xi_s (a_s^T x - b_s) = \sum_{s=1}^S |a_s^T x - b_s|$$

Így a robusztus feltétellel ekvivalens a következő:

$$a_0^T x + \sum_{s=1}^S |a_s^T x - b_s| \leq b_0$$

Az abszolút értéket pedig a t_s segédváltozók bevezetésével a következő módon tudjuk kezelni:

$$-a_s^T x - t_s \leq -b_s$$

$$a_s^T x - t_s \leq b_s$$

$$a_0^T x + \sum_{s=1}^S t_s \leq b_0$$

$$t_s \geq 0$$

egy lineáris programozási feladatot kaptunk.

Ellipszoidális bizonytalanság

Legyen a perturbációs halmaz egy gömb:

$$\mathcal{S} := \{\xi \in \mathbb{R}^S : \|\xi\|_2 \leq \Omega\}$$

Ekkor tetszőleges robusztus feltétel az alábbi formát ölti:

$$a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \leq b_0 + \sum_{s=1}^S \xi_s b_s \quad \forall \xi \in \mathcal{S}$$

Ez átírható a következő alakra:

$$\sum_{s=1}^S \xi_s (a_s^T x - b_s) \leq b_0 - a_0^T x \quad \forall \|\xi\|_2 \leq \Omega$$

Ami ezzel ekvivalens:

$$\max_{\|\xi\|_2 \leq \Omega} \sum_{s=1}^S \xi_s (a_s^T x - b_s) \leq b_0 - a_0^T x$$

Vezessük be az $u := (a_1^T x - b_1, a_2^T x - b_2, \dots, a_S^T x - b_S)^T \in \mathbb{R}^S$ vektort. Ekkor

$$\max_{\|\xi\|_2 \leq \Omega} \sum_{s=1}^S \xi_s (a_s^T x - b_s) = \max_{\|\xi\|_2 \leq \Omega} \xi^T u$$

A Cauchy–Schwarz egyenlőtlenség alapján tudjuk, hogy:

$$|\xi^T u| \leq \|\xi\|_2 \cdot \|u\|_2.$$

Ha egyenlőség van, akkor $|\xi^T u|$ nyilván maximális. A feltételben nem szerepel abszolútérték, tehát az egyenlőség akkor teljesül, ha ξ párhuzamos és egyirányú u -val, vagyis:

$$\xi = \Omega \cdot \frac{u}{\|u\|_2} \quad \text{esetén.}$$

Ekkor a maximális érték:

$$\max_{\|\xi\|_2 \leq \Omega} \xi^T u = \Omega \cdot \frac{u^T u}{\|u\|_2} = \Omega \cdot \|u\|_2.$$

Tehát a feladatban

$$\max_{\|\xi\|_2 \leq \Omega} \sum_{s=1}^S \xi_s (a_s^T x - b_s) = \Omega \cdot \sqrt{\sum_{s=1}^S (a_s^T x - b_s)^2}.$$

Ezért a robusztus feltétellel ekvivalens a következő egyenlőtlenség:

$$a_0^T x + \Omega \cdot \sqrt{\sum_{s=1}^S (a_s^T x - b_s)^2} \leq b_0,$$

ami egy nem lineáris, de konvex egyenlőtlenség, tehát az ilyenek által definiált megengedett halmaz konvex. Ennek köszönhetően a megoldáshalmaz konvex, így hatékonyan alkalmazhatók rá a konvex optimalizálási módszerek és algoritmusok.

Poliéderez bizonytalanság

Tegyük fel, hogy a perturbációs halmaz egy poliéder, azaz:

$$\mathcal{S} := \{\xi \in \mathbb{R}^S : D\xi \leq d\}.$$

A korábbi megfigyelés miatt azt is feltehetjük, hogy a feltételek jobb oldala is determinisztikus. Ekkor tetszőleges robusztus feltétel:

$$a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \leq b \quad \forall \xi \in \mathcal{S}$$

Átírható:

$$\max_{\xi: D\xi \leq d} a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \leq b \quad (4.1)$$

Vegyük észre, hogy ez a feltétel azt jelenti, hogy a

$$\begin{aligned} \max \quad & a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \\ & D\xi \leq d \end{aligned}$$

LP feladat optimuma legfeljebb b . Ahhoz, hogy a maximumot eltüntessük a kifejezésből, az erős dualitás tételt fogjuk alkalmazni. Írjuk fel a feladat duálisát:

$$\begin{aligned} \min \quad & a_0^T x + \lambda^T d \\ & \lambda^T D = u \\ & \lambda \geq 0, \end{aligned}$$

ahol $u := (a_1^T x, a_2^T x, \dots, a_S^T x)$. A robusztus feltétel helyettesíthető a:

$$\begin{aligned} a_0^T x + \lambda^T d &\leq b \\ \lambda^T D &= u \\ \lambda &\geq 0 \end{aligned} \quad (4.2)$$

feltételekkel. Ez az átalakítás ekvivalens, hiszen a $D\xi \leq d$ poliéderről feltehető, hogy nem üres, vagyis ilyenkor a duál:

- Vagy üres: Ilyenkor a primál nem korlátos, vagyis a (4.1) feltétel sem teljesül, meg az ezt helyettesítő (4.2) sem, hiszen a duál üres.
- Vagy nemüres: Ilyenkor az erős dualitás fennáll, vagyis $\max a_0^T x + \sum_{s=1}^S \xi_s a_s^T x = \min a_0^T x + \lambda^T d$. Mivel $a_0^T x + \sum_{s=1}^S \xi_s a_s^T x \leq a_0^T x + \lambda^T d$, tetszőleges ξ, λ primál és duál megoldásokra, ezért a (4.1) és (4.2) feltételek ugyanakkor teljesülnek.

5. fejezet

Gyakorlati példák

Ennek a fejezetnek a többlépcsős sztochasztikus programozás pénzügyi feladatokban való alkalmazása a témája. A fejezet alapjául az [1]-es forrás szolgál.

5.1. Többperiódusos befektetési probléma

A többperiódusos befektetési probléma a pénzügyi döntéshozatal egyik fontos feladata, amely azt vizsgálja, miként osszuk el rendelkezésre álló erőforrásainkat több időszakon keresztül úgy, hogy maximalizáljuk a várható hozamot, miközben a periódusokhoz kapcsolódó pénzügyi kötelezettségeket is teljesítjük.

A modellt T időszakra bontjuk és a feladat lényege, hogy minden időszakban a korábban megszerzett vagyon újrabefektetéséről döntsünk úgy, hogy a folyamat végére maximalizáljuk a várható profitunkat és az aktuális időszakra előírt fizetési kötelezettségeket is teljesítsük. Egy ilyen probléma megfogalmazható a következő többlépcsős sztochasztikus programozási feladattal:

$$\begin{aligned} \max \quad & \sum_i \mathbb{E}[x_{i,T}] \\ & \sum_i (1 + R_{i,t})x_{i,t-1} - \sum_i x_{i,t} = L_t, \quad t = 1, \dots, T \\ & x_{i,t} \geq 0. \end{aligned}$$

Itt $x_{i,t-1}$ az i -edik részvénybe fektetett pénzösszeget jelenti a t -edik periódus előtt, vagyis

$$\sum_i x_{i,0}$$

a kezdőtőkének. $R_{i,t}$ adja meg az i . részvény hozamát a t . periódusban. Az L_t pedig a t . periódus végén teljesítendő kötelezettséget, ami akár negatív is lehet, ha a vállalat az adott időszakot többlettel fejezi be. Ezen két típusú változó, a hozamok és a fizetési kötelezettségek, általában valószínűségi változók, vagyis előre nem ismertek,

egy adott valószínűségi eloszlás szerint vesznek fel értékeket.

A feladatban a:

$$\sum_i (1 + R_{i,t})x_{i,t-1} - \sum_i x_{i,t} = L_t, \quad t = 1, \dots, T$$

feltételek biztosítják, hogy az előző periódus végén megszerzett vagyomból kifizessük az adott periódus kötelezettségét, majd újra befektessük a megmaradt vagyont. A nemnegativitási feltételek pedig, hogy egy részvénybe csak nemnegatív összeget fektethetünk.

A felírást az alábbi módokon tudjuk bővíteni:

- Amennyiben szeretnénk olyan lehetőséget is, hogy a pénzünk egy részét nem fektetjük be, akkor bevezethetünk egy 0 értékű hozamot, azaz $R_{0,t} = 0$, $t = 1, \dots, T$.
- Ha a kölcsönök felvételét is megengedjük, akkor a

$$\sum_i (1 + R_{i,t})x_{i,t-1} - \sum_i x_{i,t} = L_t, \quad t = 1, \dots, T$$

feltételeket módosítanunk kell a következőképpen:

$$\sum_i (1 + R_{i,t})x_{i,t-1} - \sum_i x_{i,t} + \sum_{j=t}^T k_{t,j} = L_t + \sum_{h=1}^t (1 + Q_{h,t})k_{h,t}, \quad t = 1, \dots, T,$$

ahol $k_{a,b}$ az a időpontban felvett, b időpontban visszafizetendő hitelt jelöli, $Q_{a,b}$ pedig az ehhez tartozó a kamat.

5.2. Többperiódusos adósságkezelési probléma

Amikor egy vállalatnak pénzre van szüksége, gyakori megoldásként kötvényeket bocsát ki vagy hitelt vesz fel. Mivel azonban a piaci kamatlábak folyamatosan változnak, kulcsfontosságú optimalizálási kérdés, hogy mikor és mekkora összegeket érdemes visszafizetni ahhoz, hogy a cég pénzügyileg a lehető legkedvezőbb helyzetbe kerüljön.

A modellt T időszakra bontva vizsgáljuk. Hitelből K különböző típus áll rendelkezésre, amelyeket a $k = 1, \dots, K$ indexek jelölnek. A cél annak meghatározása, hogy a T periódus alatt minden egyes időszakban melyik hiteltípust és mekkora összegben célszerű felvenni, illetve mikor érdemes visszafizetni azokat, figyelembe véve a hozzájuk kapcsolódó kamatfizetési kötelezettségeket, kibocsátási költségeket, valamint visszaváltási prémiumokat vagy diszkontokat. A modell célja, hogy a szükségletek fedezése mellett minimalizálja a teljes adósságvisszafizetés költségét a T . periódus végén. Ehhez előbb be kell vezetni a szükséges döntési változókat és paramétereket, amelyek segítségével fel lehet írni a feladatot, mint többlépcsős sztochasztikus programozási problémát.

Döntési változók

- B_t^k : A k típusú hitelből ennyit veszünk a t . periódus elején.
- $O_{s,t}^k$: Az s időszakban felvett k típusú hitel még törlesztendő része t . periódus elején.
- $R_{s,t}^k$: Az s időszakban kibocsátott k típusú hitel visszafizetett része t . szakasz elején.
- S_t : A többletpénz értéke t . periódus elején (hitelek felvétele, törlesztések és kifizetések után).

Paraméterek

- $r_{s,t}^k$: A k típusú, s -ben felvett hitel, t -ben törlesztett részének kamata.
- f_t^k : A k típusú hitel felvételének költsége a t . periódus elején.
- $g_{s,t}^k$: Prémium vagy diszkont, tekinthetünk rá úgy, mint a k típusú, s -ben felvett hitel t -ben visszafizetett részéhez tartozó visszafizetési költsége.
- i_t : Többletpénz banki kamata a t . szakaszban.
- L_t : Fizetési kötelezettség a t . szakasz elején.

Ezen paraméterek mind függhetnek a véletlen esemény realizációjától, ezért ezek valószínűségi változók.

5.2.1. Megjegyzés. Itt a $g_{t,T+1}^k$, $r_{t,T+1}^k$ és L_{T+1} paramétereket is megengedjük, ezek az utolsó periódus elején felvett hitelekhez tartozó visszafizetési költség, és kamat, valamint a fizetési kötelezettség.

Célfüggvény

$$\min \mathbb{E} \left[L_{T+1} - (1 + i_T)S_T + \sum_{k=1}^K \sum_{t=1}^T (1 + g_{t,T+1}^k + r_{t,T+1}^k)(O_{t,T}^k - R_{t,T}^k) \right]$$

Vagyis a célfüggvény annak a várható értékét minimalizálja, hogy a $T + 1$. periódus kezdetén mekkora pénzügyi teher marad fenn, ha az összes adósságot és kötelezettséget azonnal vissza kell fizetni. A célfüggvény figyelembe veszi a fennmaradó adósságokhoz kapcsolódó kamatokat és prémiumokat, valamint a rendelkezésre álló készpénzt is.

Feltételek

Készpénz

$$S_t = \sum_{k=1}^K \left\{ (1 - f_t^k)B_t^k - \sum_{s=0}^{t-1} [r_{s,t}^k O_{s,t}^k + (1 + g_{s,t}^k)R_{s,t}^k] \right\} + (1 + i_{t-1})S_{t-1} - L_t,$$

$$t = 1, \dots, T$$

A t -edik periódus elején a fennmaradó készpénz a következő forrásokból származik: az előző periódusból megmaradt többletpénz kamatozott értékéből $((1 + i_{t-1})S_{t-1})$, az újonnan felvett hitelekkel felvételi költségekkel együtt számolva $((1 - f_t^k)B_t^k)$, ebből kivonásra kerül a fennálló adósságokra történő kamatfizetés $(r_{s,t}^k O_{s,t}^k)$ és a visszafizetett adósságok prémiummal vagy diszkonttal $((1 + g_{s,t}^k)R_{s,t}^k)$, valamint ebből még levonjuk az adott periódusra előírt pénzügyi kötelezettségeket is (L_t) .

Adósságok

$t = 1, \dots, T, \quad k = 1, \dots, K, \quad s = 0, \dots, (t-1)$ -re:

$$O_{s,t}^k = O_{s,t-1}^k - R_{s,t-1}^k, \quad \text{ha } s < t-1$$

$$O_{s,t}^k = B_s^k \quad \text{ha } s = t-1$$

$$O_{t,t}^k = B_t^k$$

$$R_{t,t}^k = 0$$

$$O_{s,t}^k \geq R_{s,t}^k$$

Nemnegativitási feltételek

$t = 1, \dots, T, \quad k = 1, \dots, K, \quad s = 0, \dots, (t-1)$ -re:

$$B_t^k \geq 0, \quad O_{s,t}^k \geq 0, \quad R_{s,t}^k \geq 0, \quad S_t \geq 0$$

5.2.2. Megjegyzés. Láttuk, hogy az előbb bemutatott két pénzügyi probléma megfogalmazható többletcsős sztochasztikus programozási feladatként. Így a gyakorlatban az olyan bizonytalan tényezők alakulása, mint a kamatlábak, diszkontok vagy prémiumok, hozamok és fizetési kötelezettségek, amelyek a döntéshozatalt közvetlenül befolyásolják, scenáriófa segítségével modellezhetők, feltéve, hogy együttes eloszlásuk diszkrét, vagy diszkrétrel közelíthető. Ilyen esetben a 3. fejezetben tárgyalt beágyazott dekompozíciós módszer hatékonyan alkalmazható.

6. fejezet

Összegzés és kitekintés

A dolgozatban bemutatam a sztochasztikus és robusztus optimalizálás alapelveit, valamint a sztochasztikus programozás néhány pénzügyi alkalmazását. Az egylépcsős sztochasztikus programozásnál különös figyelmet fordítottam a valószínűségi korlátozásokkal rendelkező feladatok megoldáshalmazainak konvexitására, ennek kapcsán vizsgáltam a modellek sűrűségfüggvényeinek logkonkavitását. A többlépcsős sztochasztikus programozásról szóló fejezetben először az egyszerűsítést szolgáló aggregációs technikákat mutattam be, majd a diszkrét eset került előtérbe, amelynek kapcsán leírtam a Benders- és beágyazott dekompozíciós módszereket. A robusztus optimalizálásról szóló fejezetben a feladatok kezelhető alakra írása állt a középpontban, és néhány speciális bizonytalansági halmaz esetén bemutatam az átírási menetét. Az utolsó fejezetben két, többlépcsős sztochasztikus programozást alkalmazó pénzügyi problémát vizsgáltam: a többperiódusú portfólióoptimalizálást és a többperiódusú adósságkezelési feladatot.

A dolgozat több irányban is tovább bővíthető. Egyrészt az egylépcsős sztochasztikus modellek esetén érdemes lenne továbbvizsgálni a valószínűségi modellek összetettebb eseteit és azok megoldási módszereit. Emellett releváns lehet a Value-at-Risk és Conditional Value-at-Risk alapú modellek vizsgálata is, amelyek gyakorlati szempontból kiemelten fontosak a pénzügyi alkalmazásokban. A többlépcsős sztochasztikus programozás terén bővítési lehetőséget jelent a diszkrét megközelítések általánosítása a folytonos eloszlású vektorváltozókra. Robusztus optimalizálás esetében pedig további lehetőséget jelenthet az összetettebb bizonytalansági halmazok vizsgálata.

Irodalomjegyzék

- [1] Gerard Cornuéjols, Reha Tütüncü, *Optimization Methods in Finance*, Carnegie Mellon University, Pittsburgh, 2005.
- [2] Peter Kall, Janos Mayer, and others, *Stochastic Linear Programming*, Springer, 1976.
- [3] John R. Birge and Francois Louveaux, *Introduction to Stochastic Programming*, Springer Science & Business Media, 2011.
- [4] János Mayer, *Stochastic Linear Programming Algorithms: A Comparison Based on a Model Management System*, Routledge, 2022.
- [5] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński, *Lectures on Stochastic Programming*, 2009.
- [6] Peter Kall and Stein W. Wallace, *Stochastic Programming*.
- [7] Aharon Ben-Tal, Arkadi Nemirovski, and Laurent El Ghaoui, *Robust Optimization*, Princeton University Press, 2009.
- [8] István Deák, *Bevezetés a Sztochasztikus Programozásba*, Budapesti Corvinus Egyetem Operációkutatás Tanszék, 2003.
- [9] Vincent Leclère, *Robust Optimization (lecture slides)*, Available at: https://leclere.github.io/files/talks/robust_optimization.pdf

Alulírott Szabó Máté nyilatkozom, hogy szakdolgozatom elkészítése során az alább felsorolt feladatok elvégzésére a megadott MI alapú eszközöket alkalmaztam:

Feladat	Felhasznált eszköz	Felhasználás helye	Megjegyzés
Irodalomkeresés	GPT-4o	2., 3., 4., 5. fejezet	-
Overleaf szintaktika segítség	GPT-4o	Teljes dolgozat	Ábrák, táblázatok, egyenletek megfelelő formázása, irodalomjegyzék és hivatkozások kialakítása
Angol szakszavak magyar megfelelőjének keresése, egyéb fordítások	GPT-4o, Google Translate	2., 3., 4., 5. fejezet	-
Nyelvhelyesség és stilisztika ellenőrzése	Writefull, GPT-4o	Teljes dolgozat	-

A felsoroltakon túl más MI alapú eszközt nem használtam.