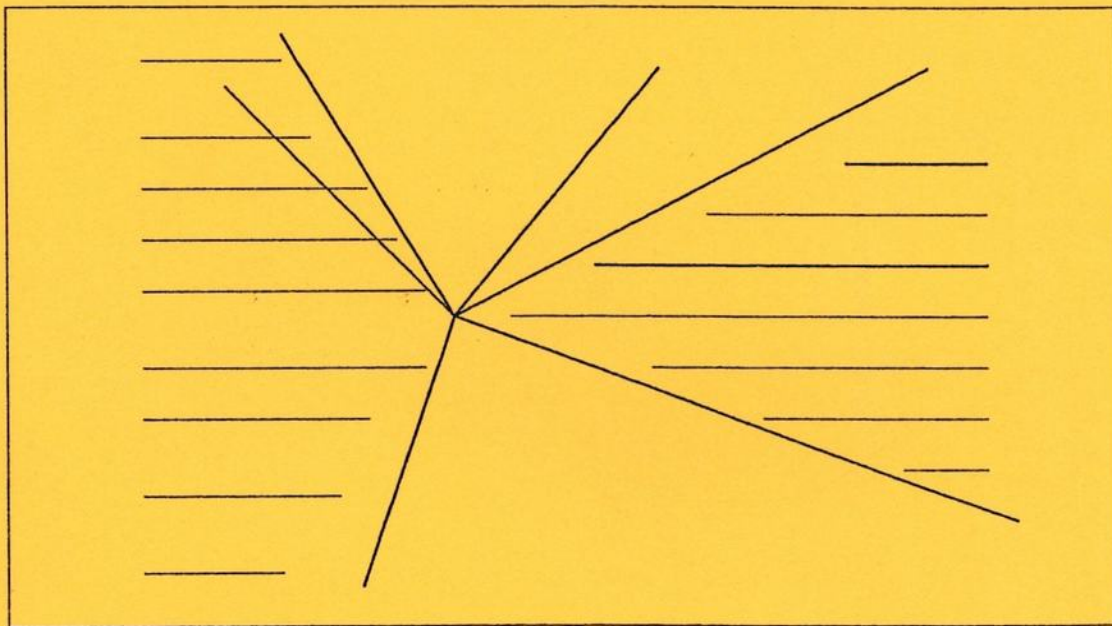


# OPERÁCIÓKUTATÁS

No.3.

**Deák István**

## **BEVEZETÉS A SZTOCHASZTIKUS PROGRAMOZÁSBA**



---

Budapest 2003

# Bevezetés a sztochasztikus programmozásba<sup>‡</sup>

Deák István

Budapesti Műszaki és Gazdaságtudományi Egyetem,

Matematika Intézet, Differenciálegyenletek Tanszék,

Operációkutatási csoport,

H-1111 Budapest, XI. Műegyetem rkp. 3.,

email: deak@math.bme.hu

2003. június 26.

---

\*Ez a jegyzet az FKFP 0231 számú pályázatának támogatásával készült, negyedéves és ötödéves matematikus, informatikus, közgazdász hallgatóknak és doktoranduszoknak.

<sup>†</sup>A jegyzetet Hans Zoltán lektorálta

# Tartalomjegyzék

<b>1. Modellépítési alapelvek</b>	<b>8</b>
1.1. A valószínűség maximalizálása . . . . .	8
1.2. A hasznossági függvény . . . . .	12
1.3. Az újságárusfiú problémája . . . . .	15
1.4. Hatékony megoldások. . . . .	15
1.5. Determinisztikusból sztochasztikus feladat . . . . .	17
1.6. Várhatóérték programozás . . . . .	22
1.7. Általános modellépítési elvek . . . . .	23
1.8. Jelölések . . . . .	25
 <b>2. Valószínűséggel korlátozott modellek</b>	 <b>27</b>
2.1. Általános elvek . . . . .	27
2.1.1. Feltételek . . . . .	27
2.1.2. Célfüggvény . . . . .	29
2.1.3. Egyedi modellek . . . . .	30
2.2. Logkonkávítási eredmények . . . . .	31
2.2.1. Konvexitási alapfogalmak . . . . .	31
2.2.2. Logkonkávítási tételek . . . . .	34
2.2.3. Többdimenziós logkonkáv eloszlások . . . . .	36
2.3. Valószínűségek kiszámítása normális eloszlás esetén . . . . .	38
2.3.1. Többdimenziós integrálok Monte Carlo kiszámítása . . . . .	39
2.3.2. Egy integráltranszformáció . . . . .	39
2.3.3. A belépési és kilépési állandók . . . . .	44
2.3.4. Ortonormalizált becslések . . . . .	45
2.4. Korlátok a valószínűségekre . . . . .	48
2.4.1. Valószínűségek korlátozása . . . . .	48
2.4.2. Szimuláció és korlátozás . . . . .	51
2.5. A STABIL modell . . . . .	53
2.5.1. Konvexitás és valószínűségek . . . . .	54

2.5.2.	Optimalizálási algoritmus . . . . .	55
2.5.3.	Iránymenti optimalizálás . . . . .	58
2.5.4.	A STABIL alkalmazása . . . . .	58
2.5.5.	Numerikus példa . . . . .	59
2.6.	Megoldó algoritmusok . . . . .	60
<b>3.</b>	<b>Kétlépcsős modellek</b>	<b>62</b>
3.1.	A kétlépcsős modelltípus alapjai . . . . .	62
3.1.1.	Lineáris modell . . . . .	62
3.1.2.	Szekvenciális döntések . . . . .	63
3.1.3.	Többlépcsős modell . . . . .	64
3.1.4.	Indukált feltételek . . . . .	65
3.1.5.	A pótló függvény tulajdonságai . . . . .	67
3.1.6.	Speciális esetek . . . . .	70
3.2.	A pótló függvény kiszámítása . . . . .	71
3.2.1.	Diszkrét eloszlás . . . . .	72
3.2.2.	Diszkrétizálás . . . . .	73
3.2.3.	Monte Carlo becslés . . . . .	74
3.3.	Korlátok a pótló függvényre . . . . .	78
3.3.1.	Jensen korlát . . . . .	78
3.3.2.	Edmundson-Madansky korlátok . . . . .	80
3.3.3.	A korlátok geometriai jelentése . . . . .	81
3.3.4.	Egy numerikus példa . . . . .	82
3.4.	Megoldó algoritmusok . . . . .	83
3.4.1.	Általános megjegyzések . . . . .	83
3.4.2.	Dantzig-Wolfe . . . . .	83
3.4.3.	Egyéb módszerek . . . . .	84
<b>4.</b>	<b>Szukcesszív regressziós approximáció</b>	<b>86</b>
4.1.	Egyenlet megoldása . . . . .	86
4.1.1.	Determinisztikus eset . . . . .	87

4.1.2.	Numerikus példa . . . . .	89
4.1.3.	Zajos függvény . . . . .	90
4.2.	Valószínűséggel korlátozott modellek . . . . .	91
4.2.1.	A STABIL megoldása regressziós közelítéssel . . . . .	91
4.2.2.	Algoritmus a STABIL feladatára . . . . .	93
4.2.3.	Egy valószínűséggel korlátozott numerikus példa . . . . .	94
4.3.	Kétlépcsős modellek . . . . .	97
4.3.1.	A kétlépcsős feladat regressziós közelítéssel . . . . .	98
4.3.2.	Algoritmus a kétlépcsős feladatra . . . . .	98
4.3.3.	Numerikus példa . . . . .	99
4.4.	Kombinált modell . . . . .	102
4.4.1.	Kétlépcsős feladat valószínűségi korláttal . . . . .	102
4.4.2.	Regresszió a kombinált modell megoldására . . . . .	104
4.4.3.	Algoritmus a kombinált modellre . . . . .	106
4.4.4.	Numerikus példa . . . . .	107
4.5.	Sztokasztikus kvadratikusan programozás . . . . .	108

<b>5. Irodalom</b>	<b>111</b>
--------------------	------------

# Bevezetés

A sztochasztikus programozás a véletlen jelenlétében való döntéshozatallal foglalkozik. Másképpen fogalmazva ez az a matematikai tudományág, amely optimális döntések vizsgálatával és meghatározásával foglalkozik olyan esetekben, amikor véletlen mennyiségeket is figyelembe kell venni a modellek felépítésében, az optimális döntés meghozásában és a döntés következményeinek kiértékelésében.

A jegyzetben a sztochasztikus programozás alapvető modelljeit és megoldási módszereit ismertetjük. Anyagunk négy nagyobb részből áll. Először az általános elvekkel foglalkozó részben különböző, a sztochasztikus programozásban használt modellépítési és döntéshozási elvekkel foglalkozunk. A 2. és a 3. Fejezetben a valószínűséggel korlátozott modellek és a kétlépcsős modellek tulajdonságait tárgyaljuk. Azért választottuk ezt a két modell típust részletesebb tárgyalásra, mert ezeknek van alaposan kidolgozott matematikai elméletük és a gyakorlatban jól használható megoldó algoritmusuk, továbbá ennek a két modell típusnak a vizsgálata és alkalmazása alkotja a sztochasztikus programozási kutatások nagy részét. A második és harmadik részben tárgyalt anyag lényegében a sztochasztikus lineáris programozás modelljeit és megoldó algoritmusait ismerteti. A negyedik részben a szukcesszív regressziós approximációk megoldási módszerét tárgyaljuk. Ez használható mindkét nagy modellcsalád esetében, azonkívül ennek használatával meg lehet oldani egy olyan kombinált feladatot is, amelyben valószínűséggel korlátozott feltétel is és a kétlépcsős feladatra jellemző pótló függvény is jelen van. Ennek az egységes megoldó algoritmusnak a kapcsán megmutatjuk, hogy olyan általános kvadratikus sztochasztikus programozási feladat is optimalizálható, amelyben mind kvadratikus feltételek, mind kvadratikus célfüggvény is szerepel.

Terjedelmi korlátok miatt nem tudunk minden, a sztochasztikus programozás témakörébe tartozó feladattal foglalkozni, a legfontosabb speciális eseteket azonban megemlítjük. Az érintett témákról és egyéb, itt nem tárgyalt lehetőségekről

az irodalomban talál az olvasó részletesebb leírást – a jegyzet végén egy rövid ismertetésben és az utána következő irodalomjegyzékben a legfontosabb könyvekre és cikkekre felhívjuk a figyelmet. Bár a jegyzet az időkorlát miatt nem lett teljes, de véleményünk szerint így is jól használható – a közeljövőben egy bővített változat megírását tervezzük.

A jegyzetet a Budapesti Műszaki és Gazdaságtudományi Egyetemen ötödéves informatikus és matematikus hallgatóinak, 1993 és 2002 között tartott előadásaim alapján állítottam össze és elsősorban a számukra írtam. Az itt összefoglalt anyag egy féléves oktatásra alkalmas, esetleg további irodalom felhasználásával bővíthető. A jegyzet olvasása az alapvető matematikai és valószínűségszámítási eredményeken kívül feltételezi a lineáris és a nemlineáris optimalizálás elméletének és gyakorlatának ismeretét is. Az utóbbi két témakör nem közismert részeire röviden utalunk, az irodalomról szóló részben az ezekre vonatkozó utalásokat is lehet találni.

Ma már tucatnyinál is több összefoglaló jellegű könyvet írtak a sztochasztikus programozásról, de ezek kevésbé használhatók negyed vagy ötödéves hallgatók oktatására, magyar nyelven pedig tudtunkkal semmilyen sztochasztikus programozási jegyzet vagy könyv nem jelent meg. A jegyzettel ezt a hiányt szeretnénk pótolni. A jegyzet használatát azoknak is ajánljuk, akik a determinisztikus optimalizálás módszerein kívül szeretnék megismerni és munkájukban felhasználni a sztochasztikus programozás alapvető modelljeit.

Egy sztochasztikus programozási modellt három szempontból értékelhetünk: matematikai, számítástechnikai és alkalmazási szempontból. Tárgyalásmódunkban mindhárom szempont, lehetőleg kiegyensúlyozottan szerepel. A közölt modellek vizsgálatában nagy súlyt fektettünk arra, hogy precíz matematikai eszközökkel mutassuk meg a megengedett megoldások létezésének, illetőleg az optimalitási kritériumoknak a teljesülését. A számítástechnikai szempont figyelembevétele azt jelenti, hogy számítástechnikailag hatékony, gyakorlatilag alkalmazható algoritmusokkal foglalko-

zunk. Tárgyalásmódunk gyakorlatias – kisméretű numerikus példákkal szemléltetjük az elveket, továbbá az algoritmusokat olyan részletességgel írjuk le, hogy a matematikai levezetések teljes megértése nélkül is fel lehessen használni a leírt módszereket. Ezek a példák azt is lehetővé teszik, hogy a sztochasztikus programozás témakörébe elemi szinten betekintést nyerjenek más szakterületek kutatói, alkalmazói. Végül az alkalmazási szempontot is figyelembe vettük, – ahol lehet, már megtörtént gyakorlati alkalmazásokat mutatunk be, illetőleg ilyenekre hívjuk fel a figyelmet. Ezek az alkalmazási példák arra is rávilágítanak, hogy mennyire szerteágazóak a sztochasztikus programozás felhasználásának lehetőségei.

Ezt a jegyzetet Prékopa Andrásnak ajánlom, akitől az első lineáris programozási előadást hallgattam 1965-ben, majd évtizedeken át szerencsém volt az általa vezetett kutatócsoportokban dolgozni.

Budapest, 2003.

Deák István



# 1. Modellépítési alapelvek

A véletlentől függő döntéseket nem tudjuk egymáshoz hasonlítani addig, amíg nem határozzuk meg, hogy milyen szempont szerint tekintünk egy döntést jobbnak a másikinál. Tehát elsősorban azt a szempontot kell meghatároznunk, amely alapján a döntéseket rangsorolhatjuk. Ezek után lehet feladatunk megengedett megoldások és az optimális megoldás létezésének (egyediségének) vizsgálata. Végül azt kell megvizsgálnunk, hogy milyen algoritmussal lehet az optimális megoldást meghatározni – itt mindig számítástechnikailag megvalósítható és lehetőleg hatékony eljárásra kell törekednünk. Ez a három szempont – döntési elv, megengedettség és kiszámíthatóság – a feladatok felépítésében és tárgyalásában döntő fontosságú lesz.

A statisztika is tekinthető a véletlen jelenlétében való döntéshozatalnak, így bizonyos értelemben a statisztika is része a sztochasztikus programozásnak. Például a statisztikai próbák ( $u$  próba, vagy  $\chi^2$  próbák) esetén is a véletlen minta alapján akarunk döntést hozni a  $H_0$  hipotézis elfogadásáról (vagy elvetéséről). Ezért az alapelvek vizsgálata során először a statisztikában használt gondolatokat tárgyaljuk meg, majd leírunk egy lineáris programozásból származó mintapéldát. A mintapélda lehetséges változatai között tárgyaljuk bevezető jelleggel azt a két modellépítési elvet, amely a sztochasztikus programozás két legfontosabb modell típusának felépítéséhez vezet: a valószínűséggel korlátozott és a kétlépcsős sztochasztikus programozási feladatot.

## 1.1. A valószínűség maximalizálása

Nagyon gyakori az az eset, amikor valamilyen számunkra kívánatos esemény valószínűségét szeretnénk minél nagyobbá (vagy legalábbis egy adott értéknél nagyobbá) tenni. Tekintsük az úgynevezett portfólió feladatot, amellyel tőzsdei befektetések alkalmával találkozhatunk.

Tegyük fel, hogy egységnyi pénzt akarunk  $n$  különböző részvénybe befektetni, az  $i$ -edik részvénybe  $x_i$  pénzt, ahol  $\sum_{i=1}^n x_i = 1, x_i \geq 0, i = 1, \dots, n$ . Valamilyen rögzített idő elteltével akarjuk a részvényeket eladni. Az  $i$ -edik részvény jövőbeli ára

egy  $\xi_i$  valószínűségi változó, melynek várható értéke  $\mathbf{E}\xi_i = p_i$ . A véletlen jövőbeli  $\xi_i$  árák kovarianciamátrixát jelölje  $R = \{r_{ij}\}_{i,j=1}^n$ ,  $r_{ij} = \mathbf{E}(\xi_i \xi_j)$ . A jövőbeli teljes összeg, amit kaphatunk egy  $\eta = \sum_{i=1}^n x_i p_i$  valószínűségi változó. Legyen célunk egy  $d > 0$  hozam elérése, és tekintsük feladatunknak az  $\{\eta > d\}$  esemény valószínűségének maximalizálását, vagyis legyen feladatunk:

$$\max_{x_1, \dots, x_n} \mathbf{P} \{ \eta > d \}.$$

Tekintettel arra, hogy ezt a valószínűséget nehéz meghatározni, ezért a Csebisev egyenlőtlenség alapján egy felső korlátot adunk a komplementer esemény valószínűségére. Jelöljük a hozam várható értékét  $m$ -el,  $m = \mathbf{E}(\eta)$ , és legyen  $\sigma^2 = \mathbf{D}^2(\eta)$ , ekkor

$$\mathbf{P} \{ \eta \leq d \} = \mathbf{P} \{ m - \eta \geq m - d \} \leq \mathbf{P} \{ |m - \eta| \geq m - d \} \leq \frac{\sigma^2}{(d - m)^2}.$$

A  $\mathbf{P} \{ \eta > d \}$  valószínűség maximalizálása helyett így megelégszünk azzal, hogy ennek a felső korlátnak a minimalizálására törekszünk, vagyis feladatunkat újra átfogalmazzuk a következő formába:

$$\begin{aligned} \min_{x_1, \dots, x_n} \quad & \frac{\sigma^2}{(d - m)^2}, \\ \text{feltéve, hogy} \quad & \sum x_i p_i = m, \\ & \sum x_i = 1, \\ & x_i \geq 0. \end{aligned}$$

A továbbiakban szereplő minimalizálási (maximalizálási) feladatokban a feltételek előtt álló „feltéve, hogy” kifejezést „f.h.” rövidítéssel jelöljük. A  $\sigma^2 = \mathbf{D}^2(\sum_{i=1}^n x_i p_i)$  meghatározható, így ez a feladat egy egyszerű kvadratikus optimalizálási feladat.

Valószínűségek maximalizálásával foglalkozunk a maximum likelihood becslések meghatározása esetén is. Tegyük fel, hogy a statisztikai sokaság valószínűségi változójának a sűrűségfüggvénye  $f(x, \vartheta)$ , ahol  $\vartheta$  az eloszlás egy ismeretlen paramétere.

Tegyük fel, hogy a mintavételi eljárás folyamán az  $x_1, x_2, \dots, x_n$  mintákat kaptuk, és tekintsük az

$$L = f(x_1, \theta) f(x_2, \vartheta) \cdots f(x_n, \vartheta)$$

függvényt, amelyet likelihood függvénynek nevezünk (ezt a  $dx_1 \cdots dx_n$  eltérésekkel szorozva a mintavétel valószínűségét kapjuk). Mivel éppen az adott  $x_1, x_2, \dots, x_n$  mintát kaptuk, ezért azt gondoljuk, hogy ez volt a legvalószínűbb az összes lehetőség közül, vagyis az  $L$  függvény maximális értéket ért el. Azt a  $\vartheta$  paraméterértéket keressük, amelyre az  $L$  függvény a maximumát éri el (vagy a  $\log L$  függvény éri el maximumát, ami ezzel ekvivalens). Tehát feladatunk azon  $\hat{\vartheta}$  érték meghatározása, amelyre

$$\hat{\vartheta} = \arg \max_{\vartheta} L(x_1, \dots, x_n, \vartheta)$$

teljesül (itt az „arg max” az az értéke az argumentumnak, amelyre az adott függvény a maximumát veszi fel).

**Példa.** Tekintsünk egy statisztikai sokaságon értelmezett normális eloszlású valószínűségi változót, amelynek az  $m$  várható értéke és az  $s$  szórásnégyzete ismeretlen, tehát a sűrűségfüggvénye

$$f(x, m, s) = \frac{1}{\sqrt{2\pi}\sqrt{s}} e^{-\frac{(x-m)^2}{2s}}.$$

A várható érték és a szórásnégyzet maximum likelihood becslése akkor azon  $\hat{m}, \hat{s}$  értékek lesznek, amelyekre a

$$\begin{aligned} \ln L(x_1, \dots, x_n, m, s) &= \ln \left[ \prod_{i=1}^n f(x_i, m, s) \right] \\ &= -\frac{n}{2} (\ln 2\pi + \ln s) - \sum_{i=1}^n \frac{(x_i - m)^2}{2s^2} \end{aligned}$$

log-likelihood függvény felveszi a maximumát. A feltétel nélküli optimalizálásra vonatkozó elsőrendű szükséges optimalitási kritériumokat felírva a log-likelihood függ-

vényre a következő egyenletrendszert kapjuk

$$\frac{\delta \ln L}{\delta m} = \sum_{i=1}^n \frac{x_i - m}{s} = 0, \quad \frac{\delta \ln L}{\delta s} = -\frac{n}{2s} + \sum_{i=1}^n \frac{(x_i - m)^2}{2s} = 0.$$

Ezt az egyenletrendszert az ismeretlen  $m, s$  változókra megoldva kapjuk a

$$\hat{m} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{s} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{m})^2$$

egyenlőségeket, amelyek szerint az empirikus közép és az empirikus szórásnégyzet a normális eloszlás esetén a maximum likelihood becslések.

Egy további lehetőség arra, hogy egy kívánatos esemény valószínűségét lehetőleg nagyra tegyük az lehet, hogy ezt a valószínűséget beépítjük egy hasznossági függvénybe (utility function, Marschak, 1951) – a hasznossági függvényeket a következő részben tárgyaljuk részletesebben. Tekintsünk egy  $\mathbf{x} = (x_1, \dots, x_n)$  befektetési stratégiát, és a következő  $u$  hasznossági függvényt (amely nem konkáv, de kvázi-konkáv):

$$u(z, \mathbf{x}) = \begin{cases} -K, & \text{ha } z < d, \\ z, & \text{ha } z \geq d, \end{cases}$$

ahol  $d$  és  $K$  adott nemnegatív konstansok,  $z$  a befektetés hozama (az eredeti befektetés a nyereséggel növelve). Tekintettel arra, hogy egy adott  $\mathbf{x}$  befektetés hozama az  $\eta$  valószínűségi változó, vagyis  $\eta$  függ  $\mathbf{x}$ -től, ennek sűrűségfüggvénye  $f(y, \mathbf{x})$  is függvénye a befektetési stratégiának. A véletlen  $u(\eta, \mathbf{x})$  hasznosság értékét az adott hasznossági függvény alapján számíthatjuk ki és ennek a várható értékét szeretnénk maximalizálni. Ezek szerint a feladatunk a következő formába írható:

$$\max_{\mathbf{x} \in X} \mathbf{E}_\eta(u(\eta, \mathbf{x})) = -K \int_{-\infty}^d f(y, \mathbf{x}) dy + \int_d^{\infty} y f(y, \mathbf{x}) dy.$$

A várható érték alsó indexe azt jelöli, hogy melyik valószínűségi változó szerint vesszük a várható értéket – ezt a jelölést főleg olyankor használjuk, ha félreértés léphet fel.

## 1.2. A hasznossági függvény

Sokszor a nyeremény értéke önmagában nem határozza meg a játékos által követett stratégiát. Hiszen kevés ember kockáztatna egy millió forintot ezer forint nyereségért, de sokan szívesen lottóznak, hiszen pár száz forint elvesztésével szemben a többmillióس nyeremény lehetősége áll (még akkor is, ha ennek csak kicsi a valószínűsége).

Általában az  $u$  hasznosságot egy, az  $x$  nyeremény függvényeként értelmezett  $u(x)$  monoton növekvő, konkáv függvénnyel definiáljuk – természetesen más paraméterektől is függhet az  $u$  függvény. A monoton növekvés mutatja azt a kívánalmat, hogy nagyobb nyereményt hasznosabbnak értékeljünk, a konkáv tulajdonság pedig kifejezi azt az elvet, hogy egy bizonyos szint feletti nyereségnövekedés már egyre kisebb növekedést jelent hasznosságban a játékos számára.

D. Bernoulli (1873) fogalmazta meg hasznossági függvények segítségével a pétervári problémát. Egy játékos Pálnak a következő játékban való részvételt ajánlja fel. A játékos feldob egy pénzérmét (a pénzérme szabályos,  $1/2$  valószínűséggel fej,  $1/2$  valószínűséggel írás lesz az eredmény). Ha az eredmény fej, akkor Pál kap 1 dukátot, egyébként a játék végetér. Ha fej volt az első dobás eredménye, akkor a játék folytatódik, a játékos megint feldobja a pénzérmét, fej esetén Pál két dukátot kap, ha a harmadik dobásra is fej jön ki, Pál négy dukátot kap, stb. Ha egyszer is írás lesz a pénzdobás eredménye, akkor végetér a játék.

A játékban a Pál által elnyerhető pénzösszeg várható értéke  $\sum_{i=1}^{\infty} (1/2)^i 2^{i-1} = \infty$  (ennek alapján fogalmaznak meg nyerőnek gondolt stratégiákat is). Mégis, ha a játék megkezdése előtt valaki hajlandó lenne megvenni Pál helyét a játékban, Pál valószínűleg egy sokkal kisebb, rögzített pénzért eladná a részvételi jogát. A kérdés az, hogy Pál milyen rögzített összegért lenne hajlandó lemondani a játékjogáról.

Tegyük fel, hogy Pálnak a játék megkezdése előtt már van  $a$  összegű pénze, és a játékban a véletlen  $x$  összeget nyerheti el. Ekkor az általa elnyert összeg hasznosságát a következő hasznossági függvénnyel definiáljuk:

$$h(x, a, b) = b \log \frac{a+x}{a},$$

a=	0	100	100 000	10 000 000
D=	2	4.3	9.2	11.7

1. táblázat. Pál induló vagyona és a hozzátartozó determinisztikus ekvivalens értéke.

ahol  $b$  egy pozitív konstans. Ez a hasznossági függvény monoton növekvő és konkáv, továbbá megfelel a hasznosságról heurisztikusan elfogadott elveinknek. A véletlen hasznosság várható értéke:

$$\begin{aligned}
\mathbf{E}h(x, a, b) &= \frac{b}{2} \log \frac{a+1}{a} + \frac{b}{4} \log \frac{a+2}{a} + \dots + \frac{b}{2^n} \log \frac{a+2^{n-1}}{a} + \dots \\
&= b \log [(a+1)^{1/2} (a+2)^{1/4} \dots (a+2^{n-1})^{1/2^n} \dots] - b \log a \sum_{n=1}^{\infty} \frac{1}{2^n} \\
&= b \log \left[ \prod_{n=1}^{\infty} \left( \frac{a+2^{n-1}}{a} \right)^{1/2^n} \right].
\end{aligned}$$

Legyen  $D$  az a rögzített összeg, amelynek ugyanakkora a hasznossága, mint a véletlen nyereség hasznosságának várható értéke (itt  $D$ -t a várható hasznosság determinisztikus ekvivalensének nevezzük). Tehát  $D$ -t a következő egyenlőség határozza meg:

$$\mathbf{E}h(x, a, b) = b \log \frac{a+D}{a}.$$

Ebből az egyenletből kaphatjuk mindkét oldalt  $e$  hatványára emelve, hogy

$$\frac{a+D}{a} = \prod_{n=1}^{\infty} \left( \frac{a+2^{n-1}}{a} \right)^{1/2^n} = \prod_{n=1}^{\infty} \left( \frac{1}{a} \right)^{1/2^n} \prod_{n=1}^{\infty} (a+2^{n-1})^{1/2^n} = \frac{1}{a} \prod_{n=1}^{\infty} (a+2^{n-1})^{1/2^n},$$

amiből a determinisztikus ekvivalens értéke kifejezhető:

$$D = \prod_{n=1}^{\infty} (a+2^{n-1})^{1/2^n} - a.$$

Néhány numerikus adatot adunk meg az 1. Táblázatban annak szemléltetésére, hogyan függ  $D$  értéke az  $a$  kezdeti vagyontól. Látható, hogy  $D$  nagyon lassan nő, tehát ha Pál nem nagyon gazdag ember (100 000 dukátnál kisebb a vagyona), akkor 20 dukátért boldogan eladná a részvételi jogát (az adott hasznossági függvény esetén).

Általában, ha a nyereség egy  $\xi$  valószínűségi változó, és a játékos az  $u(x)$  hasznossági függvény szerint értékeli az  $x$  nyereségét, akkor a hasznosság várható értéke  $\mathbf{E}[u(\xi)]$ . A determinisztikus ekvivalens értéke pedig azon  $D$ , amelyre  $u(D) = \mathbf{E}[u(\xi)]$ , és ebből az egyenletből  $D$  meghatározható:

$$D = u^{-1}(\mathbf{E}[u(\xi)]).$$

Ha adott egy  $u(x)$  hasznossági függvény, akkor ehhez definiáljuk a

$$\varrho(x) = -\frac{u''(x)}{u'(x)}$$

kockázattól való tartózkodás (risk aversion) függvényét. Pál esetében a hasznossági függvény  $u(x) = f(x, a, b) = b \log \frac{a+x}{a}$  volt, a kockázattól való tartózkodás értéke pedig  $\varrho(x) = \frac{1}{x+a}$ . Legyen  $a = 1, b = 1$ , ekkor egységnyi nyeresénytől  $1/2$  a tartózkodás,  $x = 1000$  nyereségtől pedig kb.  $0.001$  a tartózkodás mértéke. Egy másik,  $u(x) = 1 - e^{-x}$  alakú hasznossági függvény esetén a kockázattól való tartózkodás függvénye  $1$ , vagyis itt mindenféle nyereség esetén ugyanakkora kockázatot hajlandó vállalni a játékos.

Nem konkáv (de monoton növekvő) hasznossági függvényeket is használhatunk. Tegyük fel, hogy a börtönből való kikerülés feltétele egy  $K$  óvadék letétele. Ekkor a játékban való esetleges nyereséget is ezzel az óvadékkal összehasonlítva értékeli a játékos hasznosnak, vagy sem: legyen itt  $u(x) = -1$ , ha  $x < K$ , és  $u(x) = 1$ , ha  $x > K$ .

Megjegyezzük, hogy valószínűségek maximalizálására, minimalizálására, vagy valószínűségi változók várható értékére, hasznossági függvényre vonatkozó gondolatmenetet csak akkor használhatunk, ha ugyanaz a helyzet ismételt, többször is lezajlik, ugyanolyan körülmények között. Ha a valószínűség nem közelíthető jól a relatív gyakorisággal (annak határértékével), mint például ritka katasztrófák bekövetkezése esetén, akkor inkább egy olyan esemény valószínűségét szeretnénk maximalizálni, amely a számunkra kedvező.

### 1.3. Az újságárusfiú problémája

Tekintsük az újságárusfiú problémáját, amely a sztochasztikus programozás egyik klasszikus feladata. Tegyük fel, hogy egy újságárusfiú a központban újságokat vehet,  $b$  egységáron, amelyeket az utcán egy magasabb  $c$  áron adhat el ( $c > b$ ). A jelentkező igények száma  $\xi$ , amely valószínűségi változó ismert eloszlással. Az újságárusfiú alapvető problémája az, hogy hány újságot vegyen meg, ha a nyereségének a várható értékét akarja maximalizálni.

A feladat leírásához két esetet kell külön megvizsgálni. Ha kevesebb újságot vesz a központban, mint az aktuális kereslet ( $x < \xi$ ), akkor a haszna

$$d^- = c\xi - bx = -b(x - \xi) + (c - b)\xi.$$

Abban az esetben pedig, ha több újságot vesz meg előre, mint amennyit el tud adni ( $\xi < x$ ), akkor a nyeresége

$$d^+ = cx - bx = -(c - b)(\xi - x) + (c - b)\xi.$$

Tegyük fel, hogy a véletlen kereslet eloszlása diszkrét, vagyis  $p_i = P\{\xi = i\}$ ,  $i = 0, 1, \dots$  a kereslet eloszlása. Ekkor a várható nyereség értéke az  $x$  függvénye lesz:

$$\mathbf{E}(d^-) + \mathbf{E}(d^+) = (c - b)\mathbf{E}(\xi) - \sum_{i \leq x} bp_i(x - i) - \sum_{i > x} (c - b)p_i(i - x).$$

### 1.4. Hatékony megoldások.

Pénzügyi alkalmazások esetén érdekes definíciót adott Markowitz (1952). Tekintsük az előbb leírt portfólió feladatot, vagyis legyenek a feltételeink a következők:

$$x_1 + \dots + x_n = 1, \quad x_1, \dots, x_n \geq 0.$$

Tegyük fel, hogy az  $\mathbf{x} = (x_1, \dots, x_n)$  portfóliónk egy jövőbeli (véletlen) eredményt ad, amelynek várható értéke  $M$  és szórása  $\sigma$ . Az  $M$  a jövőbeli teljes nyereségünk várható értéke, a  $\sigma$  szórása pedig a bizonytalansági szintnek (rizikónak) tekinthető. Egy adott befektetés kiértékelése kapcsán nyilván mindkét mennyiségnek szerepelnie kell;  $M$ -et maximalizálni, a  $\sigma$  bizonytalanságot minimalizálni szeretnénk.



**1. Definíció.** Az  $\mathbf{x} = (x_1, \dots, x_n)$  portfóliót *hatékony*nak nevezzük, ha

- (i)  $M$  nem növelhető (nem növekvő  $\sigma$  szórás mellett),
- (ii)  $\sigma$  nem csökkenthető az  $M$  csökkentése nélkül.

A definíció bizonyos értelemben egy nyeregpontot határoz meg: az adott  $\sigma$  bizonytalansági szint mellett nem növelhető a nyereség, és az adott  $M$  nyereség mellett pedig nem csökkenthető a bizonytalanság. A megengedett, illetőleg hatékony megoldásokat egy  $\sigma, m$  koordinátarendszerben ábrázoljuk, a megengedett megoldások  $S$  halmazának bizonyos határpontjai lesznek a hatékony megoldások.

Matematikailag a hatékony megoldás keresése a következő kvadrátikus programozási feladatot jelenti. Legyen  $x_i$  az  $i$ -edik részvénybe történő befektetési hányad, mely esetén a jövőbeli nyereség egy  $\xi_i$  valószínűségi változó,  $\mathbf{E}(\xi_i) = m_i$ , a nyereségek kovarianciamátrixa pedig legyen  $\Sigma = \{\sigma_{ij}\}_{i,j=1}^n$ . Ekkor egy adott  $M$  nyereség elérésének feltételezése mellett abban az esetben lesz a befektetés minimális szórású, ha az  $\mathbf{x}$  portfólió a következő feladat optimális megoldása:

$$\begin{aligned} \min \left( \sum_{i=1}^n x_i^2 \sigma_{ii} + \sum_{i=1}^n \sum_{j=1, i \neq j}^n x_i x_j \sigma_{ij} \right) \\ \text{f.h. } \sum_{i=1}^n x_i = 1, \\ \sum_{i=1}^n x_i m_i = M, \\ x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0. \end{aligned}$$

A feladat könnyen bővíthető egyéb feltételekkel is, például megkívánhatjuk, hogy az osztalék egy minimális értéket érjen el, vagy a bizonytalanság nem lehet egy értéknél nagyobb, stb.

Általában a hatékony megoldások a megengedett megoldások halmazának egy felületét (felület részét) képezik. Így egyértelmű hatékony megoldásról nem lehet szó, bár adott nyereség esetén általában csak egy minimális szórású megoldás van. Könnyen belátható, hogy ha egy  $\mathbf{x}$  és egy  $\mathbf{y}$  pont hatékony megoldás, akkor ezeknek az összes  $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$ ,  $0 \leq \lambda \leq 1$  konvex kombinációja is vagy hatékony megoldás,

$x_2 =$	0	0.2	0.4	0.6	0.8	1.0
$M =$	8	9.2	10.4	11.6	12.8	14
$\sigma =$	3.0	3.17	3.65	4.33	5.13	6.0

2. táblázat. Két részvény kombinációjának nyeresége és szórása,  $\rho = 0.5$  esetén.

vagy csak megengedett megoldás. Tehát a megengedett megoldások  $S$  halmaza konvex halmaz, a hatékony megoldások pedig (egy adott  $const.$  állandó esetén) a  $\{\mathbf{x} \mid \sup_{\sigma=const.} M(\mathbf{x})\}$  által megadott pontok halmaza  $\sigma = const \in I$  esetén.

**Példa.** Legyen adva két részvény, amelyek közül az első egy kisebb  $m_1 = 8\%$  nyereséget, de kisebb  $\sigma_1 = 3.0$  bizonytalanságot mutat (csatornázási vállalat, vagy más szolgáltatás), míg a másik nagyobb  $m_2 = 14\%$  nyereséget és nagyobb  $\sigma_2 = 14.0$  bizonytalanságot eredményez. Ekkor egy  $\rho = 0.5$  korreláltság mellett a teljes portfólió  $M$  nyeresége és  $\sigma$  szórása a következő táblázatban látható. Az  $(x_1, x_2)$  portfólió

$$\sigma^2 = x_1^2 \sigma_1^2 + 2x_1 x_2 \sigma_1 \sigma_2 \rho + x_2^2 \sigma_2^2$$

szórásnégyzetű,  $M = 8x_1 + 14x_2$  várhatóértékű (százalék) nyereséget kínál. Ez a mintapélda az  $x_1 = 1 - x_2$  helyettesítéssel egy egyváltozós szélsőérték feladattá formálható: adott  $M$ -hez kiszámítható a  $\sigma$  (vagy fordítva: adott  $\sigma$  esetén  $M$  meghatározható). Látható, hogy adott  $\sigma_1, \sigma_2, \rho$  értékek esetén az  $M$  várhatóérték a  $\sigma$  konkáv függvénye (tehát a megengedett megoldások halmaza konvex, ennek felső burkolója adja a hatékony megoldásokat).

## 1.5. Determinisztikusból sztochasztikus feladat

Az alábbi mintapélda lehetséges változataival a sztochasztikus programozás két legfontosabb döntési elvét szemléltetjük: a **valószínűséggel korlátozott** modellt és a **kétlépcsős sztochasztikus programozási** modellt.

Tekintsünk egy olajfinomítót, amelybe kétféle nyersanyagot (kőolajat) szállítanak,  $x_1, x_2$  mennyiségben, ezek beszerzési ára ismert  $c_1, c_2$ . A nyersanyagokból kétféle készterméket készítenek, benzint és fűtőolajat. Természetesen a kétféle kőolajból

különböző arányban tudják előállítani a két készterméket – egységnyi „nehéz” kőolajból  $a_{11}, a_{21}$ , egységnyi „könnyű” kőolajból pedig  $a_{12}, a_{22}$  benzint illetőleg fűtőolajat lehet előállítani. Az olajfinomító által készített benzin és fűtőolaj mennyiségét úgy akarják meghatározni, hogy a piaci keresletet  $(\xi_1, \xi_2)$  kielégítsék, minimális költséggel. A finomítási eljárás alatt egy 10 százalékos veszteséggel is számolnunk kell, amely adottnak tételezhető fel. Tegyük még fel, hogy a maximális terhelhetőség miatt van egy felső korlát a teljes feldolgozható mennyiségre, ami  $K_U$  egységnél nem lehet nagyobb, továbbá a minimális kihasználhatóság miatt (technológiai alapjárat) van egy  $K_L$  alsó korlát is. Ezek alapján egy lineáris programozási feladat írható fel a termelési folyamatra:

$$\begin{aligned} \min \quad & (c_1 x_1 + c_2 x_2) \\ \text{f.h.} \quad & a_{11} x_1 + a_{12} x_2 \geq \xi_1, \\ & a_{21} x_1 + a_{22} x_2 \geq \xi_2, \\ & x_1 + x_2 \leq K_U, \\ & 2x_1 + x_2 \geq K_L, \\ & x_1, x_2 \geq 0. \end{aligned}$$

Ezt a feladatot a **determinisztikus alapeladatnak** nevezzük, ha az igények adottak, vagyis ha  $\xi_1 = b_1, \xi_2 = b_2$  valamilyen  $b_1, b_2$  állandókkal. A fő kérdésünk ekkor az, hogy ha bizonyos mennyiségek a fentebbi lineáris programozási modellben valószínűségi változók, akkor hogyan kell értelmezni a feladatot, illetőleg mit tekintünk optimális megoldásnak. A valószínűségi változók egyes esetekben nem-konvex, vagy pedig üres megengedett megoldási halmazra is vezethetnek.

A lineáris programozási feladatok geometriai szemléltetése alapján a feltételek egy konvex poliédert adnak, amelyhez egy  $(c_1, c_2)$  normálvektorú egyenest illesztünk (az optimális megoldás esetében ez egy, a megengedett megoldások halmazának érintője lesz). Ha a  $\xi_1, \xi_2$  változik, akkor a feltételi egyeneseket (párhuzamosan) toljuk el, ha a  $(c_1, c_2)$  változik, akkor az illesztett egyenest forgatjuk el. A legnehezebb feladat akkor keletkezik, ha az  $A$  együtthatói változnak, mert ezzel a megengedett megoldások poliédere változik: a feltételeket eltoljuk és elforgatjuk.

Az egyszerűség kedvéért most csak a  $\xi_1, \xi_2$  igényekről tesszük fel, hogy valószínűségi változók, valamilyen adott eloszlással, a többi paramétert rögzítettnek tekintjük. Az alapmodell (változatlan formájában) nyilván értelmezhetetlen ebben az esetben, viszont többféle sztochasztikus optimalizálási modell írható fel. Mielőtt ezeket ismertetnénk, vegyük észre, hogy ha a technológia mátrix együtthatói (az  $a_{11}, a_{12}, a_{21}, a_{22}$  paraméterek) véletlenek is lehetnek, akkor még az sem biztosítható, hogy létezik megengedett megoldás. Hiszen a feltételek egy konvex poliédert írnak le, amelyek egyes esetekben üres halmazt is adhatnak. Ha a költségfüggvény  $c_1, c_2$  paraméterei véletlenek, akkor pedig a célfüggvény értéke válhat nem korlátossá.

(i) **Várható érték optimalizálás.** A legegyszerűbb megközelítés az, ha a valószínűségi változókat a várható értékükkel helyettesítjük; ekkor egy egyszerű lineáris programozási feladatot kell csak megoldanunk. Legyenek a feladatban szereplő paraméterek értékei a következők:

$$c_1 = 25, c_2 = 35, a_{11} = 0.3, a_{12} = 0.7, a_{21} = 0.6, a_{22} = 0.2, K_U = 120, K_L = 20,$$

az igényt jelentő valószínűségi változokról pedig tegyük fel, hogy egyenletes eloszlásúak a  $[20, 80]$ , illetőleg a  $[10, 110]$  intervallumban, vagyis  $\mathbf{E}(\xi_1) = 50, \mathbf{E}(\xi_2) = 60$ . Ezeket behelyettesítve a fenti feladatba egy determinisztikus lineáris programozási feladatot kapunk, amelynek optimális megoldása könnyen megadható.

(ii) **Megbízhatóság (valószínűségi korlát).** Egy másik lehetőség sztochasztikus programozási modell kialakítására egy kívánatos esemény valószínűségére előírt valószínűségi korlát (probabilistic constraint), vagyis egy megbízhatósági típusú feltétel előírása: azt kívánjuk meg, hogy az igényeket  $p$  valószínűséggel ki tudjuk

elégíteni (ahol  $p$  egy 1-hez közeli megbízhatósági szint, például  $p = 0.9$ ).

$$\begin{aligned}
 (1) \quad & \min (c_1 x_1 + c_2 x_2) \\
 \text{f.h. } & \mathbf{P} \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 \geq \xi_1 \\ a_{21}x_1 + a_{22}x_2 \geq \xi_2 \end{array} \right\} \geq p, \\
 & x_1 + x_2 \leq K_U, \\
 & 2x_1 + x_2 \geq K_L, \\
 & x_1, x_2 \geq 0.
 \end{aligned}$$

Az előírandó  $p$  megbízhatósági szint nagysága az adott alkalmazástól függ. A mindennapi életben a  $p = 0.95$  általában elégséges, de egy ország villamos energiával való ellátottságában nem túlzás  $p = 0.997$  valószínűséget megkövetelni.

(iii) **Kétlépcsős döntés.** Egy harmadik lehetőség a modellalkotásra az lehet, hogy egy pótlólagos (második lépcsős) beavatkozás lehetőségét engedjük meg. Tegyük fel, hogy az  $(x_1, x_2)$  termelési terv összeállítása és végrehajtása után megfigyelhetjük a  $\xi_1, \xi_2$  igények valódi értékét. Ezek ismeretében a tőzsdén azonnali vétellel be tudjuk szerezni a késztermékekből esetleg hiányzó mennyiséget – természetesen a szokásosnál magasabb áron (spot price).

Legyen a benzinből és/vagy fűtőolajból hiányzó mennyiség  $y_1, y_2$ , amelynek az egységára  $q_1, q_2$ , tehát a pótlólagos beszerzés költsége  $q_1 y_1 + q_2 y_2$ , amely a  $\xi_1, \xi_2$  véletlen igényektől függ, vagyis maga is valószínűségi változó. Ennek a költségnek a várható értékével növeljük meg az eredeti költségfüggvényt. Tehát feladatunk a következő lesz:

$$\begin{aligned}
 \min & c_1 x_1 + c_2 x_2 + \mathbf{E}(q_1 y_1 + q_2 y_2) \\
 \text{f.h. } & x_1 + x_2 \leq K_U, \\
 & 2x_1 + x_2 \geq K_L, \\
 & a_{11}x_1 + a_{12}x_2 + y_1 = \xi_1, \\
 & a_{21}x_1 + a_{22}x_2 + y_2 = \xi_2, \\
 & x_1, x_2, y_1, y_2 \geq 0.
 \end{aligned}$$

Ezt a kétlépcsős feladatot az általában szokásos formában úgy írjuk fel, hogy a második lépcsős tevékenységet (a később beszerzendő  $y_1, y_2$  mennyiség meghatározását) a pótló függvénybe, illetőleg annak a definíciójába tesszük. Legyen az úgynevezett **első lépcsős feladat** a következő:

$$(2) \quad \begin{aligned} \min \quad & c_1 x_1 + c_2 x_2 + \mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi})) \\ \text{f.h.} \quad & x_1 + x_2 \leq K_U, \\ & 2x_1 + x_2 \geq K_L, \\ & x_1, x_2 \geq 0, \end{aligned}$$

ahol a pótlólagos  $q(\mathbf{x}, \boldsymbol{\xi})$  költség (a második lépcsős beavatkozás költsége) a következőképpen adható meg:

$$q(\mathbf{x}, \boldsymbol{\xi}) = \min_{y_1, y_2} \left\{ q_1 y_1 + q_2 y_2 \left| \begin{array}{l} y_1 = \xi_1 - [a_{11}x_1 + a_{12}x_2], \\ y_2 = \xi_2 - [a_{21}x_1 + a_{22}x_2], \\ y_1 \geq 0, y_2 \geq 0 \end{array} \right. \right\}.$$

Az ebben a definícióban szereplő lineáris programozási feladatot a **második lépcsős feladatnak** nevezzük. A  $q(\mathbf{x}, \boldsymbol{\xi})$  függvényt a pótlólagos (recourse) vagy járulékos költségnek nevezzük, ennek várható értékét, a  $Q(\mathbf{x}) = \mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$  függvényt pedig a várható pótlás költségének nevezzük. A (??)-ben megadott feladat viszont alapjában véve egy egyszerű determinisztikus (nemlineáris) programozási feladat, amely a hagyományos determinisztikus optimalizálási eljárással (elvieken legalábbis) megoldható. A nehézség a  $Q(\mathbf{x})$  függvényben van, ennek a kiszámítása általában nem könnyű.

Ezt a modelltypust kétlépcsősnek is nevezzük, mivel az első lépcsőben az  $\mathbf{x}$  változó értékéről döntünk, majd a második lépcsőben tudunk korrigálni az  $(y_1, y_2)$  változók meghatározásával. A kétlépcsős modell bizonyos dinamizmus beépítését is lehetővé teszi, hiszen itt döntés  $\mathbf{x}$ , megfigyelés  $\boldsymbol{\xi}$ , döntés  $\mathbf{y}$  sorozatban határozzuk a  $\mathbf{x}, \mathbf{y}$  döntési változóinkról. A kétlépcsős modell nehézség nélkül általánosítható arra az esetre, amikor szekvenciálisan hozunk döntéseket: megfigyeljük a véletlen értékét és ennek alapján újabb döntést hozhatunk a véletlen hatásának korrigálására több-

szőr egymás után – ezeket a modelleket többlepcsős sztochasztikus problémáknak nevezzük.

Természetesen, hasonló büntetőfüggvényes megközelítés másmilyen járulékos költségfüggvénnyel is adható, vagyis az eredeti  $c_1x_1 + c_2x_2$  lineáris taghoz hozzáadhatjuk az  $\mathbf{E}[(\boldsymbol{\xi} - \mathbf{A}\mathbf{x})'(\boldsymbol{\xi} - \mathbf{A}\mathbf{x})]$  mennyiséget is.

## 1.6. Várhatóérték programozás

Az előző mintapélda esetén a két legfontosabb sztochasztikus programozási feladattípust, a valószínűséggel korlátozott és a kétlépcsős feladatot mutattuk be. Megemlítettük a várható érték esetén történő optimalizálást is, de erről most megmutatjuk, hogy ez a megközelítés általában – a megbízhatóság szempontjából – nagyon rossz eredményt is adhat, ezért általában nem ajánljuk az alkalmazását. Legyen feladatunk a következő:

$$\begin{aligned} \min \quad & (x + y) \\ \text{f.h. } & 2x + y \geq u_1, \\ & x + 2y \geq u_2, \\ & x, y \geq 0, \end{aligned}$$

ahol  $u_1, u_2$  valószínűségi változók, amelyek függetlenek és egyenletes eloszlásúak a  $[0, 2]$  intervallumban. Helyettesítsük most a valószínűségi változókat a várható értékükkel, ekkor feladatunk új determinisztikus alakja

$$\begin{aligned} \min \quad & (x + y) \\ \text{f.h. } & 2x + y \geq 1, \\ & x + 2y \geq 1, \\ & x, y \geq 0. \end{aligned}$$

Ha ezt a feladatot megoldjuk, akkor optimális megoldásként az  $x^* = 1/3, y^* = 1/3$  értékeket kapjuk. Az optimális megoldás megengedettségének valószínűsége (vagyis

annak a valószínűsége, hogy ezen  $x^*, y^*$  megoldás esetén a feltételek fennállnak) egyszerűen kiszámítható:

$$\mathbf{P}\{\omega | 1 > u_1(\omega), 1 > u_2(\omega)\} = \mathbf{P}\{u_1 < 1\}\mathbf{P}\{u_2 < 1\} = 1/4,$$

vagyis csak az esetek 1/4-ében fog mind a két feltétel együtt fennállni. Természetesen  $n$  (hasonló) feltétel esetén ez a valószínűség  $1/(2^n)$  lesz, tehát a várható érték szerinti optimális megoldás megengedettsége igen kicsi is lehet.

Ha nagyon biztosak akarunk lenni a feltételek teljesülésében, akkor az  $x^{**} = 1, y^{**} = 1$  megoldást alkalmazzuk. Ekkor a feltételek 1 valószínűséggel teljesülnek ugyan, de a költségünk értéke 2, míg az előző  $x^* = 1/3, y^* = 1/3$  megoldás esetén ez csak 2/3 volt. Vagyis szükségtelenül nagy költség árán érünk el nagy megbízhatóságot.

A valószínűséggel korlátozott modell esetén egy kompromisszumot keresünk a megoldás lehetőleg nagy  $p$  valószínűségű megengedettsége és a minimalizálandó célfüggvény nem túl nagyra növelése között. Ha a feladatból egy valószínűséggel korlátozott modellt akarunk készíteni, akkor ennek alakja a következő lesz:

$$\begin{array}{c} \min (x + y) \\ \text{f.h. } \mathbf{P} \left\{ \begin{array}{l} 2x + y \geq u_1 \\ x + 2y \geq u_2 \end{array} \right\} \geq p, \\ x, y \geq 0. \end{array}$$

A fentebb megadott  $u_1, u_2$  valószínűségi változók esetén a feltételben szereplő valószínűség könnyen felírható:  $(2x + y)(x + 2y)/4 = 0.5(x + y)^2 + 0.25xy$ , tehát például  $x + y \geq \sqrt{2}$  esetén már 1 valószínűséggel teljesül a megbízhatósági feltételünk – és itt a költségünk csak  $\sqrt{2}$  lesz. Ha pedig csak egy  $p = 0.95$  megbízhatóságot követelünk meg, akkor még tovább csökkenthető a költség.

## 1.7. Általános modellépítési elvek

Általában a következő eljárás szerint konstruálunk sztochasztikus programozási modellt. Kiindulunk egy determinisztikus (lineáris, vagy nemlineáris) optimalizálási feladatból. Meghatározzuk azon változók, együttthatók, mennyiségek körét, amelyek



valószínűségi változók. Valamilyen megfontolások alapján egy döntési elvet fogadunk el, hogy mikor tekintünk egy megoldást optimálisnak (vagy általában jobbnak egy másik megoldásnál) és végül a döntési elv segítségével felírjuk a sztochasztikus programozási modellt. Jegyezzük meg, hogy egy determinisztikus modellből még akkor is többféle sztochasztikus programozási modellt készíthetünk, ha lerögzítjük a valószínűségi változóként szereplő mennyiségek körét.

Az általános sztochasztikus lineáris programozás esetén a

$$\begin{aligned} \min \mathbf{c}'\mathbf{x} \\ A\mathbf{x} &\geq \mathbf{b}, \\ \mathbf{x} &\geq 0 \end{aligned}$$

determinisztikus alapfeladatból indulunk ki. Itt akármelyik paraméter lehet valószínűségi változó, tehát feltesszük, hogy van egy  $(\Omega, \mathcal{A}, P)$  valószínűségi mezőnk, és lehetséges valószínűségi változóink az  $(A(\omega), \mathbf{b}(\omega), \mathbf{c}(\omega)), \omega \in \Omega$ .

Vizsgálataink során mindig feltesszük, hogy a modellben szereplő valószínűségi változók (sűrűségfüggvényük, vagy eloszlásfüggvényük) ismertek. Tehát nem foglalkozunk azzal, hogy hogyan lehet a valóságban előforduló véletleneket modellezni, milyen illesztési eljárásokkal lehet a mintákból az egyes valószínűségi eloszlásokat meghatározni – ezek önmagukban is érdekes és általában nehéz feladatok.

A sztochasztikus programozásban beszélünk statikus és dinamikus modellekről; az első esetben egy egyszeri esetet kell megvizsgálni (a döntési változókról a véletlen realizálódása előtt kell döntenünk), míg a második esetben egy sorozat helyzetet kell modellálni és optimalizálni ( $t = 0, 1, 2, \dots$  időpontokban, vagy általánosabban egy folytonos paraméterű  $t$  segítségével felírt állapotot kell vizsgálni).

Szoktunk beszélni „itt és most” modellekről, valamint „várunk és meglátjuk” modellekről. Az első esetben most azonnal kell döntenünk, anélkül hogy a bekövetkező véletlenről (az általános eloszlásán kívül) bármit is tudnánk, míg a második esetben lehetőségünk van arra, hogy a később bekövetkező véletlen hatását is figyelembe vegyünk. Ilyen szempont szerint a valószínűséggel korlátozott modell „itt és most” típusú, tehát statikus modellnek tekinthető, míg a kétlépcsős (és az ennek analó-

giájára felépített többlépcsős) modell „várunk és meglátjuk” típusú. Ez a második típus ilyen formában tekinthető egy dinamikus tulajdonságokkal rendelkező modell-típusnak is.

A kiindulásként tekintett determinisztikus modellből elkészített sztochasztikus programozási modell esetén a következő alapkérdések merülnek fel:

- (i) milyen feltételek mellett lesz az új probléma megengedett megoldásainak halmaza (nem-üres és) konvex,
- (ii) milyen feltételek mellett lesz az új probléma célfüggvénye konvex (minimalizálási feladat esetén),
- (iii) milyen megoldó algoritmust használhatunk az optimális megoldás meghatározására, különös tekintettel arra, hogy sok esetben a valószínűségek vagy várható értékek kiszámítása pontatlan (csak valamilyen zajjal lehetséges).

## 1.8. Jelölések

A jegyzetben a vektorokat oszlopvektoroknak gondoljuk és vastag betűvel jelöljük, mint például:  $\mathbf{x}$ ,  $\mathbf{b}$ . Vektorok transzponáltját egy felső vesszővel jelöljük, mint  $\mathbf{x}'$ ,  $\mathbf{b}'$ . Ezek szerint az  $x_1, x_2, \dots, x_n$  komponensekből álló vektor  $\mathbf{x}' = (x_1, x_2, \dots, x_n)$ . Vektorok dimenzióját csak akkor adjuk meg, ha egyébként félreérthető a helyzet.

Mátrixot nagy betűkkel jelölünk, mint  $A, B$ . Ennek az  $A$  mátrixnak az  $i$ -edik sorában és  $j$ -edik oszlopában álló elem az  $a_{ij}$ , ezzel egy  $n \times n$  négyzetes mátrix  $A = \{a_{ij}\}_{i,j=1}^n$  formában is írható.

A valószínűségi változókat görög betűkkel jelöljük, például  $\xi, \eta$ . A valószínűségi vektor-változókat vastag görög betűkkel jelöljük, tehát a  $\xi_1, \xi_2, \dots, \xi_n$  komponensekből álló vektorra  $\boldsymbol{\xi}' = (\xi_1, \xi_2, \dots, \xi_n)$ . Egy valószínűségi (vektor-)változó várható értékét  $\mathbf{E}$  jelöli, a szórásnégyzetét pedig  $\mathbf{D}^2$ . A hangsúlyozás érdekében néha a várható érték alsó indexével mutatjuk meg, hogy melyik valószínűségi változó szerint képezzük a várható értéket, mint például  $\mathbf{E}_\eta(\xi\eta)$  jelenti az  $\eta$ -ra képzett várható értéket. Események valószínűségét  $\mathbf{P}$ -vel jelöljük.

Egyenlőségek és egyenlőtlenségek felírásakor az

$$A\mathbf{x} \leq \mathbf{b}$$

típusú egyenlőtlenséget soronként kell értelmezni, vagyis a baloldalon álló  $A\mathbf{x}$  vektor  $i$ -edik komponensét  $A\mathbf{x}|_i$ -vel jelölve az  $A\mathbf{x} \leq \mathbf{b}$  vektor-vektor egyenlőtlenség akkor és csak akkor teljesül, ha az  $n$  darab  $A\mathbf{x}|_i \leq b_i, i = 1, 2, \dots, n$  egyenlőtlenségek mindegyike teljesül. Az  $\mathbf{x} \geq 0$  szokásos nemnegativitási feltétel pedig minden komponens nemnegativitását jelenti, tehát az  $x_1 \geq 0, \dots, x_n \geq 0$  egyenlőtlenségek mindegyike fennáll.

A függvények jelölésére használt  $f(\mathbf{x}), g(\mathbf{x})$  kifejezések  $f : \mathbf{R}^n \rightarrow \mathbf{R}^1$  skalárfüggvényeket jelölnek (hacsak külön értelmezést nem adunk). A  $\max_{\mathbf{x}} h(\mathbf{x}, \boldsymbol{\xi})$  típusú jelölések azt mutatják, hogy a maximalizálást az  $\mathbf{x}$  változóra végezzük.

Az  $n$ -dimenziós euklideszi teret  $\mathbf{R}^n$ -el jelöljük. A  $O(n)$  jelölést pedig a szokásos értelemben használjuk: az  $f_n(x)$  függvény  $O(n)$ -es, ha  $f_n(x)/n$  korlátos minden  $n$  esetén.

## 2. Valószínűséggel korlátozott modellek

Ebben a fejezetben azokat az alaptípusokat tárgyaljuk meg, amelyekben a sztochasztikus modellben vagy valamilyen kívánatos esemény valószínűséget maximalizáljuk (vagy minimalizáljuk), vagy csak egy valószínűségi korlátot írunk elő egy esemény valószínűségére.

Az ilyen feladatok szokásos elnevezése angolul „probabilistic constrained optimization” (vagy kevésbé szerencsés módon „chance constrained programming”). Ezek a feladatok formailag egy várható érték optimalizálási feladattá is átírhatók.

### 2.1. Általános elvek

Kiindulásként tekintünk egy determinisztikus modellre. Feltesszük, hogy a feladat bizonyos paraméterei véletlenek, elkészítjük ennek egy (vagy több) sztochasztizált változatát, erre különböző feladatokat írunk fel és megvizsgáljuk, hogy különböző sztochasztikus modellek megoldásai milyen viszonyban vannak egymással.

#### 2.1.1. Feltételek

Tekintsük kiindulásként a következő determinisztikus modellt, amelyben a feladatunk adott  $\xi$  paraméterek mellett megengedett  $\mathbf{x}$  megoldás keresése:

$$\begin{aligned} g_1(\mathbf{x}, \xi) &\geq 0, \\ &\vdots \\ g_m(\mathbf{x}, \xi) &\geq 0. \end{aligned}$$

Itt és a következőkben is  $\mathbf{x}' = (x_1, \dots, x_n)$ ,  $\xi' = (\xi_1, \dots, \xi_m)$  adott dimenziós vektorok. Egy másik típusú, adott  $h$  célfüggvény minimumát kereső determinisztikus optimalizálási feladat a következő:

$$\begin{aligned}
& \min_{\mathbf{x}} h(\mathbf{x}, \boldsymbol{\xi}) \\
& \text{f.h. } g_1(\mathbf{x}, \boldsymbol{\xi}) \geq 0, \\
& \quad \vdots \\
& g_n(\mathbf{x}, \boldsymbol{\xi}) \geq 0.
\end{aligned}$$

Vegyük észre, hogy ha a második modelltípus esetén tudunk optimalizálni (rendelkezésünkre áll egy algoritmus, amelynek segítségével a fenti feladat optimumát meg tudjuk határozni), akkor az első modelltípusban megfogalmazott megengedett megoldás keresését is meg tudjuk valósítani: például ha a  $g_1(\mathbf{x}, \boldsymbol{\xi}) \geq 0$  egyenlőtlenséghez keresünk megengedett megoldást, ez ekvivalens a  $\max_{\mathbf{x}} g_1(\mathbf{x}, \boldsymbol{\xi})$  feltétel nélküli optimalizálás végrehajtásával. Természetesen itt előfordulhat, hogy egyáltalán nincsen megengedett megoldás.

Tegyük fel, hogy a  $\boldsymbol{\xi}_i$  paraméterek valószínűségi változók. Ekkor megbízhatósági (valamilyen esemény valószínűségére vonatkozó) feltételeket többféleképpen írhatunk elő ezeknek a modelleknek a sztochasztikus megfelelőire. Legyenek  $p_1, p_2, \dots$  adott valószínűségek, amelyek általában 1-hez közeliek.

**1. Egyedi valószínűségi korlátok.** Ilyen korlátokat akkor írhatunk elő, ha a  $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_m$  valószínűségi változók függetlenek.

$$\begin{aligned}
\mathbf{P} \{g_1(\mathbf{x}, \boldsymbol{\xi}_1) \geq 0\} & \geq p_1, \\
& \vdots \\
\mathbf{P} \{g_m(\mathbf{x}, \boldsymbol{\xi}_m) \geq 0\} & \geq p_m.
\end{aligned}$$

**2. Együttes valószínűségi korlát.** Ilyen típusú feltételeket akkor érdemes használni, ha az egyes valószínűségi változók között (erős) korreláció érvényesül. Legyen  $p$  egy 1-hez közeli érték, és legyen:

$$\mathbf{P} \left\{ \begin{array}{c} g_1(\mathbf{x}, \boldsymbol{\xi}) \geq 0, \\ \vdots \\ g_m(\mathbf{x}, \boldsymbol{\xi}) \geq 0 \end{array} \right\} \geq p.$$

Vegyük észre, hogy ha a feltételek  $g_i(\mathbf{x}) \geq \xi_i$  alakba írhatók és a  $\xi_i$  valószínűségi változók teljesen függetlenek, akkor egyedi korlátokat kapunk. Ha a valószínűségi változók egymástól függenek, akkor nem tudjuk dekomponálni a feladatot. (Az egyedi korlátok értékeit általában könnyebb kiszámítani.)

**3. 1 valószínűséggel megkövetelt** (biztosan bekövetkező) korlátok.

$$\begin{aligned} \mathbf{P}\{g_1(\mathbf{x}, \boldsymbol{\xi}) \geq 0\} &= 1, \\ &\vdots \\ \mathbf{P}\{g_m(\mathbf{x}, \boldsymbol{\xi}) \geq 0\} &= 1. \end{aligned}$$

### 2.1.2. Célfüggvény

Sztochasztikus feladatainkban valamilyen korlátok mellett egy  $h(\mathbf{x}, \boldsymbol{\xi})$  függvényt minimalizálunk (tehát a legegyszerűbb esetekben a célfüggvény determinisztikus  $h(\mathbf{x})$  alakú is lehet). Így feladatunk

$$\min_{\mathbf{x} \in X, \boldsymbol{\xi} \in \Xi} h(\mathbf{x}, \boldsymbol{\xi}),$$

ahol  $\mathbf{x}, \boldsymbol{\xi}$  valamilyen adott egyenlőtlenségek által leírt  $X, \Xi$  tartományokban mozoghatnak. Ha  $\boldsymbol{\xi}$  véletlen, akkor ennek a feladatnak a szokásos átírása sztochasztikus programozási feladattá:

$$\min_{\mathbf{x} \in X} \mathbf{P}\{h(\mathbf{x}, \boldsymbol{\xi}) \in [a, b]\},$$

valamilyen  $[a, b]$  intervallumra. Ilyen alakú feladattal találkozunk, ha egy valószínűségi korlátunk van egy optimalizálási feladatban, és a feladat megoldásának első fázisaként ehhez a korláthoz kell egy megengedett megoldást keresni, amit a

$$\max_{\mathbf{x} \in X} \mathbf{P}\{g(\mathbf{x}, \boldsymbol{\xi}) \geq 0\}$$

célfüggvény optimalizálásával lehet megadni. Egy ilyen feladat megoldása arra a kérdésre is választ ad, hogy mekkora valószínűséggel állhat fenn a valószínűségi korlátban szereplő feltétel eseménye.

További változat lehet az a gyakorlatban felmerülő eset, amikor valamilyen függvény várható értéket kell optimalizálni, vagyis legyen a feladatunk alakja a következő:

$$\min_{\mathbf{x} \in X} \mathbf{E}(h(\mathbf{x}, \boldsymbol{\xi})).$$

### 2.1.3. Egyedi modellek

Az itt említésre kerülő modellek főleg történetileg érdekesek. Charnes, Cooper és Symonds (1958, 1963) javasolták a következő modell típusokat arra az esetre, amikor függetlenség van az egyes valószínűségi változók között (vagy elhanyagolható a korreláció). Legyen  $T = \{t_{ij}\}$  és  $\mathbf{c}$  adott, és a kiindulásul vett determinisztikus modell alakja a következő:

$$\begin{aligned} \min \mathbf{c}'\mathbf{x} \\ \text{f.h. } T\mathbf{x} \geq \boldsymbol{\xi}. \end{aligned}$$

Tegyük fel, hogy a döntési változóink a véletlen  $\boldsymbol{\xi}$ -től lineárisan függenek, vagyis valamilyen rögzített  $D$  mátrix esetén legyen  $\mathbf{x} = D\boldsymbol{\xi}$ . Legyen feladatunk a  $D$  mátrix  $d_{ij}$  elemeinek meghatározása. Többféle lehetőségünk van arra, hogy sztochasztikus programozási modellt írjunk fel erre a feladatra, de mindegyikben előírjuk, hogy a feltétel valamilyen valószínűséggel teljesüljön.

**E-model**, a véletlen értékű célfüggvény várható értékének minimalizálása, valószínűségi feltétel mellett.

$$\begin{aligned} \min_{d_{ij}} \mathbf{E}(\mathbf{c}'D\boldsymbol{\xi}) \\ \text{f.h. } \mathbf{P}\{TD\boldsymbol{\xi} \geq \boldsymbol{\xi}\} \geq p. \end{aligned}$$

Ha  $\mathbf{c}$  és  $\boldsymbol{\xi}$  teljesen független valószínűségi változók, akkor a célfüggvény  $\mathbf{E}(\mathbf{c})D\mathbf{E}(\boldsymbol{\xi})$  alakba írható. Legyen  $\mathbf{p}' = (p_1, \dots, p_n)$  adott megbízhatósági korlát, tegyük fel, hogy a valószínűségi korlát dekomponálható az alábbi alakban, ekkor az E modell részletesen kiírva

$$\begin{aligned} \min_{d_{ij}} \mathbf{E}(\mathbf{c})D\mathbf{E}(\boldsymbol{\xi}) \\ \text{f.h. } \mathbf{P} \left\{ \sum_{k=1}^m \left( \sum_{j=1}^n t_{ij}d_{jk} \right) \xi_k \geq \xi_i \right\} \geq p_i, i = 1, \dots, m. \end{aligned}$$

**V modell**, a szórásnégyzet minimalizálása. Legyenek a feltételeink ugyanazok, mint az E modellben, de célfüggvényünkben valamilyen kívánatos állapottól való

eltérés várható értékét szeretnénk minimalizálni

$$\begin{aligned} \min \mathbf{E} [(\mathbf{c}'D\boldsymbol{\xi} - \mathbf{c}'_0\mathbf{x}_0)^2], \\ \text{f.h. } \mathbf{P}\{TD\boldsymbol{\xi} \geq \boldsymbol{\xi}\} \geq p, \end{aligned}$$

ahol  $\mathbf{c}_0, \mathbf{x}_0$  rögzített értékek.

**P modell**, a valószínűséget maximalizáljuk. Az E modellben megadott feltételek mellett a célfüggvény legyen egy kívánatos küszöbnél nagyobb és ennek az eseménynek (minimális profit elérésének) valószínűségét maximalizáljuk:

$$\begin{aligned} \max \mathbf{P}\{\mathbf{c}'D\boldsymbol{\xi} \geq \mathbf{c}'_0\mathbf{x}_0\} \\ \text{f.h. } \mathbf{P}\{TD\boldsymbol{\xi} \geq \boldsymbol{\xi}\} \geq p. \end{aligned}$$

## 2.2. Logkonkavitási eredmények

A valószínűséggel korlátozott modellek esetén alapvető fontosságú a megengedett megoldások tartományának konvexitása: bizonyos eloszlások esetén a valószínűségi feltételek által adott megengedett megoldások halmaza konvex lesz. Egydimenziós eloszlás esetén ez triviális, jónéhány többdimenziós eloszlás esetén ezt Prékopa, alábbiakban ismertetésre kerülő eredményei alapján állíthatjuk.

### 2.2.1. Konvexitási alapfogalmak

Néhány, a konvexitással kapcsolatos definíciót és állítást adunk meg először.

**2. Definíció.** (i) Egy  $\Omega$  halmazt konvexnek nevezünk, ha tetszőleges  $\mathbf{x}, \mathbf{y}$  pontokra minden  $0 \leq \lambda \leq 1$  esetén  $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in \Omega$  fennáll.

(ii) Egy  $\Omega$  konvex halmazon értelmezett  $f(\mathbf{x})$  függvényt konvexnek nevezünk, ha tetszőleges  $\mathbf{x}, \mathbf{y} \in \Omega$  pontokra minden  $0 \leq \lambda \leq 1$  esetén  $f(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y})$  fennáll.

(iii) egy  $f(\mathbf{x})$  függvényt konkávnak nevezünk, ha  $-f(\mathbf{x})$  konvex.

A definíciókból elemi levezetésekkel beláthatók az alábbi eredmények:



**3. Lemma.** (i) Egy  $\Omega$  konvex halmazon differenciálható  $f(\mathbf{x})$  függvény akkor és csak akkor konvex, ha tetszőleges  $\mathbf{x}, \mathbf{y} \in \Omega$  esetén fennáll, hogy

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

(ii) Ha az  $f(\mathbf{x})$  függvény konvex egy  $\Omega$  konvex halmazon, akkor az  $A_c = \{\mathbf{x} | f(\mathbf{x}) \leq c\}$  (alsó) nívóhalmazai konvex halmazok tetszőleges  $c$  konstans esetén (egy konkáv függvény felső nívóhalmazai is konvexek).

(iii) Ha  $f_1$  és  $f_2$  konvex függvények az  $\Omega$  konvex halmazon, akkor ezeknek a nemnegatív  $\alpha_1, \alpha_2$  súlyokkal vett  $\alpha_1 f_1 + \alpha_2 f_2$  összegei is konvexek.

A konvexitás fogalmat többféleképpen lehet általánosítani. Mivel számunkra a megengedett megoldások halmazának konvexitása fontos, ezért ilyen szempontok szerint adunk meg további definíciókat.

**4. Definíció.** Egy  $f(\mathbf{x}), \mathbf{x} \in A \subset \mathbf{R}^m$  függvényt kvázikonkávnak nevezünk egy  $A$  konvex halmazon, ha minden  $\mathbf{x}, \mathbf{y} \in A$  pontpár és tetszőleges  $0 < \lambda < 1$  esetén fennáll, hogy

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \geq \min[f(\mathbf{x}), f(\mathbf{y})].$$

Következmény. Egy  $f$  függvény akkor és csak akkor kvázikonkáv, ha tetszőleges  $(-\infty < c < \infty)$  konstans esetén az  $f$  függvény minden  $A_c = \{\mathbf{x} | f(\mathbf{x}) \geq c\}$  (felső) nívóhalmaza konvex.

**5. Definíció.** Egy  $A$  konvex halmazon értelmezett nemnegatív  $f(\mathbf{x}), \mathbf{x} \in A \subset \mathbf{R}^m$  függvényt logaritmikusan konkávnak (röviden logkonkávnak) nevezünk, ha minden  $(\mathbf{x}, \mathbf{y}) \in A$  pontpár és  $0 < \lambda < 1$  esetén fennáll, hogy

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

**6. Tétel.** Ha egy  $f(\mathbf{x}) > 0, \mathbf{x} \in A$  függvény logkonkáv, akkor a  $\log f(\mathbf{x})$  függvény konkáv.

**7. Tétel.** Egy logkonkáv függvény kvázikonkáv is.

A tétel állítása könnyen belátható. Legyen ugyanis  $f(\mathbf{x})$  logkonkáv függvény, ekkor:

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \geq [f(\mathbf{x})]^\lambda [f(\mathbf{y})]^{1-\lambda} \geq \min(f(\mathbf{x}), f(\mathbf{y})).$$

Jelöljük az  $\mathbf{R}^n$  tér nyílt intervallumai által generált Borel halmazalgebrát  $\mathcal{B}_m$ -el és legyen a  $\mathcal{B}_m$ -en értelmezett Lebesgue mérték  $\mathcal{L}_m$ .

**8. Definíció.** Egy  $\mathcal{B}_m$ -en definiált  $\mathbf{P}$  valószínűségi mértéket logkonkávnak nevezünk, ha tetszőleges  $A, B \subset \mathcal{B}_m$  konvex halmazok és  $0 < \lambda < 1$  esetén fennáll a következő egyenlőtlenség

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

A halmazok közötti összeadást itt Minkowski értelmében vesszük, vagyis az összegre  $\lambda A + (1 - \lambda) B = \{\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} | \mathbf{x} \in A, \mathbf{y} \in B\}$ .

**9. Tétel.** Ha a  $\mathbf{P}$  valószínűségi mérték logkonkáv és  $A$  egy konvex halmaz, akkor a  $\mathbf{P}(A + \mathbf{t})$  logkonkáv függvénye lesz a  $\mathbf{t}$  változónak.

Ugyanis legyen  $\mathbf{t}_1, \mathbf{t}_2$  adott, ekkor  $A + \lambda \mathbf{t}_1 + (1 - \lambda) \mathbf{t}_2 = \lambda(A + \mathbf{t}_1) + (1 - \lambda)(A + \mathbf{t}_2)$ . Ekkor  $\mathbf{P}$  logkonkávitásából következik, hogy

$$\mathbf{P}(A + \lambda \mathbf{t}_1 + (1 - \lambda) \mathbf{t}_2) = \mathbf{P}(\lambda(A + \mathbf{t}_1) + (1 - \lambda)(A + \mathbf{t}_2)) \geq [\mathbf{P}(A + \mathbf{t}_1)]^\lambda [\mathbf{P}(A + \mathbf{t}_2)]^{1-\lambda}.$$

Megjegyezzük, hogy ha  $f$  logkonkáv függvény egy  $A$  konvex halmazon, akkor a függvény  $f(\mathbf{x}) = 0, \mathbf{x} \notin A$  kiterjesztéssel az egész téren logkonkáv lesz.

A logkonvexitást hasonlóan lehet értelmezni és a logkonvex függvények esetén a sztochasztikus programozásban felhasználható érdekes eredményeket lehet kapni.

**10. Definíció.** Egy nemnegatív  $f(\mathbf{x}), \mathbf{x} \in A \subset \mathbf{R}^m$  függvényt logaritmikusan konvexnek (röviden logkonvexnek) nevezünk, ha  $A$  egy konvex halmaz, és minden  $(\mathbf{x}, \mathbf{y}) \in A$  pontpár és  $0 < \lambda < 1$  esetén fennáll, hogy

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq [f(\mathbf{x})]^\lambda [f(\mathbf{y})]^{1-\lambda}.$$

A Hölder egyenlőtlenség  $(\int_a^b |x(t)y(t)|dt \leq \|x\|_p \|y\|_q, \text{ ha } \frac{1}{p} + \frac{1}{q} = 1, p, q > 1)$  segítségével belátható, hogy ha a konvex  $D$  halmazon értelmezett két logkonvex függvény összegét tekintjük, az is logkonvex  $D$ -n. Ugyanis legyen  $\mathbf{x}, \mathbf{y} \in D$  és  $0 < \lambda < 1$ , akkor

$$\begin{aligned} f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) + g(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) &\leq \\ &\leq [f(\mathbf{x})]^\lambda [f(\mathbf{y})]^{1-\lambda} + [g(\mathbf{x})]^\lambda [g(\mathbf{y})]^{1-\lambda} \\ &\leq [f(\mathbf{x}) + g(\mathbf{x})]^\lambda [f(\mathbf{y}) + g(\mathbf{y})]^{1-\lambda}. \end{aligned}$$

**11. Tétel.** *Legyen  $f$  logkonvex a  $D$  konvex halmazon és Riemann integrálható az  $[A + \mathbf{t}] \in D$  mérhető halmazokon, valamint  $\mathbf{t}_1, \mathbf{t}_2$  két további vektor, amelyekre  $A + \mathbf{t}_1$  és  $A + \mathbf{t}_2$  a  $D$ -ben vannak, akkor minden  $0 < \lambda < 1$  esetén az  $f$  által generált  $P$  mértékre igaz, hogy*

$$\mathbf{P}(A + \lambda \mathbf{t}_1 + (1 - \lambda) \mathbf{t}_2) \geq [\mathbf{P}(A + \mathbf{t}_1)]^\lambda [\mathbf{P}(A + \mathbf{t}_2)]^{1-\lambda}.$$

### 2.2.2. Logkonkávítási tételek

**12. Tétel (Prékopa egyenlőtlensége).** *Legyen  $g, h$  két nemnegatív Borel mérhető logkonkáv függvény az  $\mathbf{R}^m$ -ben és definiáljuk  $\mathbf{x}, \mathbf{y} \in \mathbf{R}^m$ ,  $0 < \lambda < 1$  esetén az*

$$r(\mathbf{t}) = \sup_{\mathbf{t} = \lambda \mathbf{x} + (1-\lambda) \mathbf{y}, \mathbf{x}, \mathbf{y} \in \mathbf{R}^m} g(\mathbf{x})h(\mathbf{y})$$

*Lebesgue mérhető függvényt. Ekkor fennáll a következő egyenlőtlenség:*

$$(3) \quad \int_{\mathbf{R}^m} r(\mathbf{t}) d\mathbf{t} \leq \left[ \int_{\mathbf{R}^m} g^{1/\lambda}(\mathbf{x}) d\mathbf{x} \right]^\lambda \left[ \int_{\mathbf{R}^m} h^{1/(1-\lambda)}(\mathbf{y}) d\mathbf{y} \right]^{1-\lambda}.$$

Az egyenlőtlenség bizonyítása [?]-ben található meg. Az állítás általánosabb formában is igaz, de nekünk a kimondott alak elégséges. Megjegyezzük, hogy ha  $g$  és  $h$  logkonkáv függvények, akkor az  $r(\mathbf{t})$  függvény is logkonkáv.

**13. Tétel (A logkonkáv mértékek tétele – Prékopa).** *Legyen  $f(\mathbf{x})$ ,  $\mathbf{x} \in \mathbf{R}^m$  egy logkonkáv sűrűségfüggvény, amely generálja a  $\mathbf{P}$  valószínűségi mértéket. Ekkor  $\mathbf{P}$  egy logkonkáv mérték.*

**Bizonyítás.** Legyen  $A, B$  két konvex halmaz. Definiáljuk az  $f(\mathbf{x})$  függvény csonkítottjait és kiterjesztését a következőképpen:

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

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

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

Ekkor az  $f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x})$  függvények logkonkávok az  $\mathbf{R}^m$ -ben és fennáll a következő összefüggés:

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

mivel  $f_1^\lambda(\mathbf{x}) f_2^{1 - \lambda}(\mathbf{x}) \leq f(\mathbf{x})$  fennáll. Alkalmazzuk most Prékopa egyenlőtlenségét a  $g(\mathbf{x}) = f_1^\lambda(\mathbf{x}), h(\mathbf{y}) = f_2^{1 - \lambda}(\mathbf{y})$  függvényekre, ekkor kapjuk, hogy

$$\begin{aligned} P\{\lambda A + (1 - \lambda)B\} &= \int_{\lambda A + (1 - \lambda)B} f(\mathbf{t}) d(\mathbf{t}) = \int_{\mathbf{R}^m} f_3(\mathbf{t}) d\mathbf{t} \\ &\geq \int_{\mathbf{R}^m} \sup_{(\mathbf{x}, \mathbf{y}) | \mathbf{t} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}} f_1^\lambda(\mathbf{x}) f_2^{1 - \lambda}(\mathbf{y}) d\mathbf{t} \\ &\geq \left[ \int_A f(\mathbf{x}) d\mathbf{x} \right]^\lambda \left[ \int_B f(\mathbf{y}) d\mathbf{y} \right]^{1 - \lambda} = P(A)^\lambda P(B)^{1 - \lambda}. \end{aligned}$$

A logkonkáv mértékek tételének következményei az alábbi eredmények:

**14. Tétel.** Ha  $f(\mathbf{x}, \mathbf{y}), \mathbf{x} \in \mathbf{R}^n, \mathbf{y} \in \mathbf{R}^m$  egy  $n + m$  változóban logkonkáv függvény, akkor

$$\bar{f}(\mathbf{x}) = \int_{\mathbf{R}^m} f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$$

logkonkáv függvénye az  $\mathbf{x} \in \mathbf{R}^n$  változónak.

**15. Tétel.** Két  $\mathbf{R}^m$ -ben logkonkáv függvény konvolúciója is logkonkáv.

**Bizonyítás.** Legyen  $g$  és  $h$  két függvény, amelyek logkonkávok az  $\mathbf{R}^m$  térben. Be-  
látjuk, hogy  $g(\mathbf{x}-\mathbf{y})h(\mathbf{y})$  is logkonkáv  $\mathbf{R}^{2m}$ -ben. Felhasználjuk az előző tételt, ennek  
az integrálja az  $\mathbf{y}$  szerint szintén logkonkáv függvény az  $\mathbf{x}$  változóban.

**16. Tétel.** Tegyük fel, hogy a  $\mathbf{P}$  mértéket az  $f(\mathbf{x}), \mathbf{x} \in \mathbf{R}^m$  logkonkáv sűrűségfügg-  
vény generálja és legyen az  $A \subset \mathbf{R}^m$  egy konvex halmaz. Ekkor

- (i)  $\mathbf{P}(A + \mathbf{x})$  az  $\mathbf{x} \in \mathbf{R}^m$  változónak egy logkonkáv függvénye,
- (ii) az  $F(\mathbf{x}) = \int_{\mathbf{t} \leq \mathbf{x}} f(\mathbf{t}) d\mathbf{t}, \mathbf{x} \in \mathbf{R}^m$  eloszlásfüggvény logkonkáv,
- (iii) ha  $m = 1$ , akkor  $1 - F(x)$  is logkonkáv.

### 2.2.3. Többdimenziós logkonkáv eloszlások

A bizonyítások részbeni részletezése mellett az alábbiakban felsorolunk néhány több-  
dimenziós logkonkáv valószínűségi eloszlást (az eloszlások logkonkávitása a sűrűség-  
függvények logkonkávitásából következik).

(i) **Egyenletes eloszlás.** Legyen  $D \subset \mathbf{R}^n$  egy korlátos konvex tartomány. A  $D$   
tartományon egyenletes eloszlás sűrűségfüggvénye a

$$f(\mathbf{x}) = \begin{cases} \frac{1}{\mu(D)}, & \text{ha } \mathbf{x} \in D, \\ 0, & \text{egyébként,} \end{cases}$$

ahol  $\mu$  az  $\mathbf{R}^n$ -ben definiált Lebesgue mérték.

(ii) **Normális eloszlás.** Az  $\mathbf{R}^n$ -ben definiált nem-degenerált többdimenziós  
normális eloszlás sűrűségfüggvénye

$$\varphi(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det C}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'C^{-1}(\mathbf{x}-\boldsymbol{\mu})},$$

ahol  $\boldsymbol{\mu}$  a várható érték vektor,  $C$  pedig a kovarianciamátrix (szimmetrikus, pozitív  
definit mátrix),  $\det C$  a kovarianciamátrix determinánsa. Mivel  $C^{-1}$  is pozitív de-  
finit, továbbá minden pozitív szemidefinit kvadratikus alak egy konvex függvényt  
határoz meg, ezért az exponensben álló

$$-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'C^{-1}(\mathbf{x} - \boldsymbol{\mu})$$

függvény konkáv, tehát  $\varphi(\mathbf{x})$  valóban logkonkáv függvény az  $\mathbf{R}^n$ -ben.

(iii) Tekintsük a következő  $f$  függvényt, amely a **Dirichlet** eloszlás sűrűségfüggvénye

$$f(\mathbf{x}) = K x_1^{p_1-1} \cdots x_m^{p_m-1} (1 - x_1 - \cdots - x_m)^{p_{n+1}}, \mathbf{x} \in S,$$

ahol  $S$  az egységsimplex, vagyis  $S = \{\mathbf{x} | x_1 + \cdots + x_m < 1, x_i > 0\}$  és az  $f(\mathbf{x}) = 0$  mindenhol máshol. A  $K$  konstans értékét a következő kifejezés adja meg:

$$K = \frac{\Gamma(p_1 + \cdots + p_{n+1})}{\Gamma(p_1) + \cdots + \Gamma(p_{n+1})},$$

ahol a  $\Gamma$  függvény nem egészértékű  $p$  paraméter esetén  $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$  a nemteljes gamma függvény, egész  $p = n$  paraméterértékre pedig a faktoriális  $\Gamma(n) = (n-1)!$ . Ha a  $p_i \geq 1, i = 1, \dots, n+1$  feltétel fennáll a paraméterekre, akkor az  $f(\mathbf{x})$  logaritmus konkáv függvény lesz azon az  $S$  halmazon, ahol  $f(\mathbf{x}) > 0$ .

A Wishart, béta, lognormális eloszlások is logkonkávok, valamint egy Prékopa által definiált többdimenziós gamma eloszlás is.

A logkonvexitási eredmények alapján igaz a következő

**17. Tétel.** *Ha az  $f(\mathbf{x}), \mathbf{x} \in \mathbf{R}^m$  valószínűségi sűrűségfüggvény esetén az  $f^{-\frac{1}{m}}(\mathbf{x})$  függvény konvex az egész térben, akkor az  $f(\mathbf{x})$  által generált  $\mathbf{P}$  valószínűségi mértékre fennáll az  $A, B$  konvex halmazok és  $0 < \lambda < 1$  esetén a következő egyenlőtlenség:*

$$\mathbf{P}(\lambda A + (1 - \lambda)B) \geq \min[\mathbf{P}(A), \mathbf{P}(B)],$$

*tehát a  $\mathbf{P}$  mérték kvázikonkáv.*

(iv) **Student eloszlás.** Legyenek a  $\xi_1, \dots, \xi_m$  valószínűségi változók standard normális eloszlásúak, 0 várható értékkel és  $R$  korrelációs mátrixszal. Legyen  $\eta$  egy további valószínűségi változó, amely független az előzőktől és  $\chi$  eloszlású,  $\nu$  szabadsági fokkal. Ekkor a  $\zeta_i = \sqrt{\nu} \xi_i / \eta, i = 1, \dots, m$  valószínűségi változók együttes sűrűségfüggvénye (a többváltozós  $t$ , másnéven a Student eloszlás sűrűségfüggvénye)

$$f(\mathbf{x}) = \frac{\Gamma(\frac{1}{2}(\nu + m))}{(\pi\nu)^{\frac{1}{2}m} \Gamma(\frac{1}{2}\nu) |R|^{\frac{1}{2}}} \left( 1 + \frac{1}{\nu} \mathbf{x}' R^{-1} \mathbf{x} \right)^{-(\nu+m)/2}.$$

Az előző eredmény alapján a többdimenziós Student ( $t$ ) eloszlás kvázikonkáv. Hasonlóképp a többdimenziós Pareto eloszlás is kvázikonkáv.

Logkonkavitási (illetőleg kvázikonkavitási) definíciókat ki lehet mondani diszkrét eloszlásokra is, aminek alapján a megfelelő valószínűségi eloszlások is kvázikonkávok lesznek.

### 2.3. Valószínűségek kiszámítása normális eloszlás esetén

A sztochasztikus programozási modellekben előforduló események valószínűségének meghatározása az általános esetben igen nehéz numerikus feladat. Néhány speciális esetre az alábbi és a következő pontban adunk eljárást. Az eredményeket röviden összefoglalva: (i) normális eloszlás esetén ki tudjuk számítani néhány egyszerű konvex halmaz valószínűségét (beleértve az eloszlásfüggvény értéket) és (ii) néhány eloszlás esetén meg tudjuk határozni az eloszlásfüggvény értéket. Természetesen ezen eljárások numerikus felhasználhatósága (gyorsasága) a dimenziószámtól függ – az alábbiakban vázolt eljárások  $n = 10 - 100$  esetén adnak gyakorlatilag elfogadható idő alatt megfelelő numerikus eredményt (lásd [?]).

Megjegyezzük még, hogy a nemlineáris optimalizálási algoritmusokban gyakran van szükség a gradiens értékére. Ezt néhány esetben (valamilyen alacsonyabb dimenziós) eloszlásfüggvényértékek segítségével ki lehet számítani – például az  $n$  dimenziós normális eloszlásfüggvény gradiense kifejezhető  $(n - 1)$  dimenziós eloszlásfüggvények értékeivel, más esetekben pedig a numerikus differenciából becsült (zajos) értékeket használjuk.

Az alábbiakban ismertetendő ortonormalizált becslések módszere alkalmazható, ha adott az  $X$  halmaz, és a kiszámítandó  $\mathbf{P}\{X\}$  valószínűség az eloszlásfüggvény értéke ( $X = \{\mathbf{x} | \mathbf{x} \leq \mathbf{h}\}$ ), az  $X$  halmaz egy téglatest, poliéder, ellipszoid vagy körkúp. A NORSET számítógépes szubrutin csomagban [?] található azok a programok, amelyek segítségével három tizedes pontosságra ki lehet számítani a valószínűségeket 0.1 sec alatt  $n=20$  dimenzióig, és legfeljebb 3 sec idő alatt 100 dimenzióig.

### 2.3.1. Többdimenziós integrálok Monte Carlo kiszámítása

A többdimenziós integrálok numerikus integrálási formulái általában nem jól használhatók magasabb dimenzióban. Ennek a dimenziós robbanás-ként ismert tulajdonság az oka: ha egy  $n$ -dimenziós halmaz minden koordinátatengelyen  $k$  osztópontot veszünk fel, akkor  $k^n$  darab pontunk lesz, ami általában nem biztosít megfelelő pontosságot, és mégis igen gyorsan nő. Ezen jelenség miatt a gyakorlatban csak Monte Carlo integrálás használható  $n > 5 - 10$  esetén.

A Monte Carlo integrálási módszerek segítségével statisztikai becslés kapható egy integrál értékére. Például legyen  $\zeta$  egyenletes eloszlású valószínűségi változó az  $X \subset \mathbf{R}^n$  tartományban,  $\zeta_1, \dots, \zeta_N$  ennek független realizációi, akkor a

$$J = \int_X h(\mathbf{z}) d\mathbf{z}$$

integrál értékének egy torzítatlan becslése az

$$\Theta_0 = \frac{1}{N} [h(\zeta_1) + \dots + h(\zeta_N)]$$

átlag, mivel  $J = \mathbf{E}(h(\zeta))$ . A számított  $\Theta_0$  eredmény  $\mathbf{D}(\Theta_0) = \frac{\mathbf{D}(h(\zeta))}{\sqrt{N}}$  szórását (vagy annak háromszorosát) használjuk hibabecslésnek – természetesen ez csak egy valószínűségi értelemben vett korlát a hibára. A becslés kiszámításának számítástechnikailag nehéz (munkaigényes) része a feladatok többségében a  $h(\zeta_i)$  függvényértékek kiszámítása, vagyis a  $\Theta_0$  becslés munkaigényét az  $N$  mintaszámmal lineárisan arányosnak tételezzük fel. Ezért a  $\mathbf{D}(\Theta_0)$  hiba csak  $1/\sqrt{N}$  sebességgel csökkenthető, ami elég kedvezőtlen, így Monte Carlo számítások esetén olyan becsléseket érdemes kidolgozni, amely  $N$  növelése helyett (lényegében változatlan munka mellett) a  $\mathbf{D}(\Theta_0)$  hiba csökken (szórás-csökkentési eljárások).

### 2.3.2. Egy integráltranszformáció

Legyen a  $\mathbf{0}$  várható értékű és  $R$  korrelációs mátrixszú  $n$ -dimenziós standard normális eloszlás sűrűségfüggvénye és eloszlásfüggvénye  $\varphi$ , illetőleg  $\Phi$ , tehát



$$\begin{aligned}\varphi(\mathbf{z}) &= (2\pi)^{(-n/2)}|R|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\mathbf{z}R^{-1}\mathbf{z}\right\}, \\ \Phi(\mathbf{h}) &= \int_{-\infty}^{h_1} \cdots \int_{-\infty}^{h_n} \varphi(\mathbf{z})d\mathbf{z}.\end{aligned}$$

Feltesszük még, hogy a normális eloszlás nem-degenerált, vagyis  $R$  pozitív definit mátrix. Másmilyen, nem-standard normális eloszlás esetén a valószínűségek egy lineáris transzformációval megkaphatók. Tegyük fel, hogy feladatunk az  $n$ -dimenziós  $X$  halmaz valószínűségének meghatározása, vagyis az

$$(4) \quad I = \mathbf{P}\{X\} = \int_X \varphi(\mathbf{z})d\mathbf{z}$$

integrál kiszámítása. A fentiekben ismertetett legegyszerűbb Monte Carlo integrálási módszer nem könnyen alkalmazható az  $X$  halmaz általános volta és a  $\varphi$  függvény alakja miatt (nem korlátos  $X$  esetén nem is tudunk egyenletes eloszlású realizációkat generálni). Az  $I$  integrál egyszerűen átírható a következő alakba:

$$(5) \quad I = \int_X \varphi(\mathbf{z})d\mathbf{z} = \int_{\mathbf{R}^n} f(\mathbf{z})d\Phi(\mathbf{z}),$$

ahol  $f(\mathbf{z})$  az  $X$  halmaz karakterisztikus függvénye, azaz

$$f(\mathbf{z}) = \begin{cases} 1, & \text{ha } \mathbf{z} \in X, \\ 0, & \text{egyébként.} \end{cases}$$

Az (??) egyenlőség jobboldala alapján a következő Monte Carlo eljárást lehet megadni az  $I$  kiszámítására. Generáljunk  $\mathbf{x}_i, i = 1, \dots, N$  független mintákat a  $\varphi$  sűrűségfüggvényű  $\boldsymbol{\xi}$  valószínűségi változóból, ekkor az  $I$  integrál torzítatlan becslése lesz a következő átlag:

$$\Theta_1 = \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i).$$

Ezt a becslést a **durva becslésnek** (**crude estimator**) nevezzük (vagy más néven elfogadás-elvetés becslésnek is, mivel csak azt kell ellenőrizni, hogy a  $\mathbf{x}_i, i = 1, \dots, N$  minták közül hány lesz az  $X$  tartományon belül) és ez nem más, mint az  $X$  tartományba való beesés relatív gyakorisága.

Ismert összefüggés alapján a normális eloszlású  $\boldsymbol{\xi}$  valószínűségi változó dekomponálható a következő módon:

$$\boldsymbol{\xi} = \chi_n T \boldsymbol{\eta},$$

ahol  $\chi_n$  egy  $n$  szabadságfokú  $\chi$ -eloszlású valószínűségi változó,  $T$  egy felső háromszög mátrix, amelyre  $TT' = R$  és az  $\boldsymbol{\eta}$  vektor egyenletes eloszlású az  $S = \{\mathbf{x} | \sum_{i=1}^n x_i^2 = 1\}$  egységgömb felületén. A  $\boldsymbol{\xi}$  vektor  $\chi_n$  „hossza” és  $\boldsymbol{\eta}$  „iránya” független valószínűségi változók. Ennek a felbontásnak a segítségével integrálunkat is dekomponálhatjuk.

Legyen a  $\chi$ -eloszlású valószínűségi változónk eloszlásfüggvénye  $K(s), s \geq 0$ , az  $\boldsymbol{\eta}$  irány eloszlásfüggvénye  $V(\mathbf{y}), \mathbf{y} \in S$ , ekkor (??) átírható az

$$(6) \quad I = \int_{\mathbf{R}^n} f(\mathbf{z}) d\Phi(\mathbf{z}) = \int_{\mathbf{R}^n} f(sT\mathbf{y}) dK(s) dV(\mathbf{y}) = \int_S \left( \int_0^\infty f(sT\mathbf{y}) dK(s) \right) dV(\mathbf{y})$$

alakba. Vezessük be a  $g(\mathbf{y})$  jelölést a belső integrálra:

$$(7) \quad g(\mathbf{y}) = \int_0^\infty f(sT\mathbf{y}) dK(s).$$

Ez a  $g(\mathbf{y})$  függvény adott  $\mathbf{y}$  vektorra megadja a  $\lambda T\mathbf{y}, \lambda \geq 0$  sugár  $X$  halmazba eső részének valószínűségi tartalmát:  $g(\mathbf{y}) = P\{\lambda T\mathbf{y} \in X | \mathbf{y}, \lambda \geq 0\}$ . Először a  $g(\mathbf{y})$  függvény értékeinek meghatározásával foglalkozunk.

Tegyük fel, hogy a  $\lambda \mathbf{z} = \lambda T\mathbf{y}$  vonal elmetshi az  $X$  konvex halmazt, vagyis van egy „belépési” és egy „kilépési” pont az  $\mathbf{z}$  irány esetén. A belépési és a kilépési pontokat azon  $\lambda_L$  és  $\lambda_U$  állandók adják meg, amelyekre  $X \cap \{\mathbf{z} | \mathbf{z} = \lambda T\mathbf{y}\} = [\lambda_L T\mathbf{y}, \lambda_U T\mathbf{y}]$ , vagyis a következő összefüggésekből határozhatók meg:

$$(8) \quad \begin{aligned} \lambda_L &= \min_{\lambda} \{ \lambda | f(\lambda T \mathbf{y}) = 1 \} = \min_{\lambda} \{ \lambda | \lambda T \mathbf{y} \in X \}, \\ \lambda_U &= \max_{\lambda} \{ \lambda | f(\lambda T \mathbf{y}) = 1 \} = \max_{\lambda} \{ \lambda | \lambda T \mathbf{y} \in X \}. \end{aligned}$$

Bevezetve az  $\lambda_L^+ = \max\{0, \lambda_L\}$ ,  $\lambda_U^+ = \max\{0, \lambda_U\}$  jelöléseket, a  $\lambda_L, \lambda_U$  állandók segítségével a  $g(\mathbf{y})$  függvényérték kiszámítható

$$(9) \quad g(\mathbf{y}) = K(\lambda_U^+) - K(\lambda_L^+)$$

az egydimenziós  $K(\cdot)$   $\chi$ -eloszlásfüggvény segítségével. Ha most még a  $-\mathbf{y}$  irány esetén is fel akarjuk írni a  $g$  függvény értékét, akkor a  $\lambda_U^- = \min\{0, \lambda_U\}$ ,  $\lambda_L^- = \min\{0, \lambda_L\}$  jelölésekkel azonnal adódik

$$(10) \quad g(-\mathbf{y}) = -K(-\lambda_U^-) + K(-\lambda_L^-).$$

Összefoglalva a két félegyenesre vonatkozó eredményeket: a (??) által megadott konstansokkal egy tetszőleges  $\mathbf{y}$  vektor által megadott  $\lambda \mathbf{z} = \lambda T \mathbf{y}$  egyenesen (itt  $\lambda$  nincs nemnegatívítással korlátozva) a valószínűségi tartalom is felírható:

$$e(\mathbf{z}) = e(T \mathbf{y}) = [g(\mathbf{y}) + g(-\mathbf{y})]/2.$$

Mivel a  $\lambda_L, \lambda_U$  értékek  $\mathbf{y}$ -tól függenek, ezért a  $g(\mathbf{y}) = K(\lambda_U^+ | \mathbf{y}) - K(\lambda_L^+ | \mathbf{y})$  jelölést is használjuk a következőkben. A  $g$  függvény kiszámíthatósága azon múlik, hogy az egydimenziós  $K(\cdot)$  eloszlásfüggvény kiszámítására gyors és pontos numerikus eljárások (szubrutinok) léteznek-e. A  $\chi$  valószínűségi változó  $K(\cdot)$  eloszlásfüggvényére ilyen szubrutinok léteznek.

A fentiek szerint az

$$(11) \quad I = \int_S [K(\lambda_U^+ | \mathbf{y}) - K(\lambda_L^+ | \mathbf{y})] dV(\mathbf{y})$$

integrál kiszámítására használható a következő Monte Carlo integrálási módszer: generálunk  $\mathbf{y}_i, i = 1, \dots, N$  mintákat az  $S$  egységgömbön egyenletes eloszlásból,

és determinisztikusan kiszámítjuk minden egyes  $\mathbf{y}_i$  irány esetén a  $g$  függvényt. A becslésünk így a következő formát ölti:

$$(12) \quad \Theta_2 = \frac{1}{N} \sum_{i=1}^N g(\mathbf{y}_i) = \frac{1}{N} \sum_{i=1}^N K(\lambda_U^+ | \mathbf{y}_i) - K(\lambda_L^+ | \mathbf{y}_i).$$

Vegyük észre, hogy az így előállított algoritmus akkor hatékony, ha egy gyors numerikus eljárással meg tudjuk határozni a  $\lambda_U, \lambda_L$  belépési illetve kilépési állandókat. Természetesen elvileg tetszőleges  $X$  halmaz valószínűsége meghatározható a fenti módon (még nem-konvex halmazé is), de a gyakorlati használhatóság  $\lambda_U, \lambda_L$  gyors kiszámíthatóságán múlik. Az egyszerűség kedvéért mi csak az eloszlásfüggvény esetén írjuk le az algoritmust, de ez számítástechnikailag hatékony módon megtehető tetszőleges konvex poliéder, hiperellipszoid és körkúp esetén is.

A (??) egyenletben leírt integráltranszformációt és az eredményül kapott kettős integrált többféleképpen is felfoghatjuk.

(i) Tetszőleges  $p = \mathbf{P}\{\xi \in X\}$  valószínűség felírható, mint az  $X$  halmaz

$$(13) \quad f(\xi) = \begin{cases} 1, & \text{ha } \xi \in X, \\ 0, & \text{egyébként.} \end{cases}$$

indikátor valószínűségi változójának a várható értéke, vagyis

$$p = \mathbf{E}[f(\xi)].$$

Ez a várható érték megfelel a (??) jobboldalán álló integrálnak. Felhasználva a  $\mathbf{E}(\alpha) = \mathbf{E}[\mathbf{E}(\alpha|\beta)]$  ismételt (feltételes) várható érték összefüggést ez a várható érték átírható a

$$(14) \quad p = \mathbf{E}[f(\xi)] = \mathbf{E}[f(\chi_n T \boldsymbol{\eta})] = \mathbf{E}_{\boldsymbol{\eta}}[\mathbf{E}_{\chi}(f(\chi_n T \boldsymbol{\eta}) | \boldsymbol{\eta})]$$

alakba, ami viszont pontosan megfelel (??) kettős integráljának.

(ii) Egy másik lehetséges értelmezés adódik a numerikus integrálás szempontjainak figyelembevételével. A Monte Carlo integrálás elég jól működik, ha a feladat

dimenziója nagy, de tudjuk, hogy viszonylag lassú,  $O(N^{-1/2})$  a konvergencia sebessége. A hagyományos (determinisztikus) integrálási szabályok kis hibával dolgoznak, de ezeket nem nagyon lehet magasabb dimenzióban használni.

A kettős integrál formájába írt kifejezés a munkánkat két részre osztja: egy egydimenziós, vonal menti integrál meghatározása hagyományos numerikus integrálási technika segítségével és egy, az  $n$ -dimenziós térben elhelyezkedő  $(n - 1)$ -dimenziós felületen elvégzett Monte Carlo integrálásra. Ezt a felbontást sugaras-felületi integrálásnak, vagy iránymenti szimulációnak (directional simulation) is nevezik, s a többdimenziós  $t$  eloszlás kiszámításában, illetőleg elliptikus sűrűségfüggvényű eloszlások kiszámításában használható.

(iii) Végül a Monte Carlo integrálás szempontjából is megvizsgálhatjuk a dekompozíciót. Minden szimuláció esetén a fő kérdés az, hogyan lehet csökkenteni a becslés szórását (anélkül, hogy lényegesen megnövelnénk a szükséges munkát). Ez az eljárás éppen erre példa – a változók számának csökkentésével szóráscsökkenést érünk el.

### 2.3.3. A belépési és kilépési állandók

Egy  $\lambda \mathbf{z} = \lambda T \mathbf{y}$  egyenes és egy általános  $X$  halmaz esetén a  $\lambda_L, \lambda_U$  belépési és kilépési állandók meghatározása nem egyszerű feladat, de néhány néhány egyszerű konvex  $X$  halmaz esetén ez nem okoz nehézséget. Szemléltetésképpen most a leg-egyszerűbb esetet írjuk le a  $\lambda_L, \lambda_U$  konstansok meghatározására; legyen halmazunk az  $X = \{\mathbf{x} | \mathbf{x} \leq \mathbf{h}\}$  egyenlőtlenséggel adva, ahol  $\mathbf{h} \geq \mathbf{0}$  adott vektor. Ekkor az  $X$  halmaz valószínűsége a többdimenziós normális eloszlás eloszlásfüggvényének értéke a  $\mathbf{h}$  helyen:

$$(15) \quad I = \mathbf{P}\{X\} = \int_X \varphi(\mathbf{z}) d\mathbf{z} = \int_{\mathbf{x} \leq \mathbf{h}} \varphi(\mathbf{z}) d\mathbf{z} = \int_{-\infty}^{h_1} \cdots \int_{-\infty}^{h_n} \varphi(\mathbf{z}) d\mathbf{z} = \Phi(\mathbf{h}).$$

Tegyük fel, hogy a  $\lambda \mathbf{z} = \lambda T \mathbf{y}$  egyenes metszi az  $X$  konvex halmazt, vagyis van egy „belépési” és egy „kilépési” pont az  $\mathbf{z}$  irány esetén (ez a  $\mathbf{h} \geq \mathbf{0}$  feltevés miatt most mindig teljesül). A belépési és a kilépési pontokat azon  $\lambda_L$  és  $\lambda_U$  állandók adják meg,

amelyekre  $X \cap \{\mathbf{z} | \mathbf{z} = \lambda T\mathbf{y}\} = [\lambda_L T\mathbf{y}, \lambda_U T\mathbf{y}]$ , vagyis a következő összefüggésekből határozhatók meg:

$$(16) \quad \begin{aligned} \lambda_L &= \min_{\lambda} \{\lambda | f(\lambda T\mathbf{y}) = 1\} = \min_{\lambda} \{\lambda | \lambda T\mathbf{y} \in X\}, \\ \lambda_U &= \max_{\lambda} \{\lambda | f(\lambda T\mathbf{y}) = 1\} = \max_{\lambda} \{\lambda | \lambda T\mathbf{y} \in X\}. \end{aligned}$$

Jelöljük a  $T$  mátrix  $i$ -edik sorát  $\mathbf{t}_i$ -vel. Egy  $\lambda T\mathbf{y}$  egyenes az  $\{\mathbf{x} | x_i = h_i\}$  hipersíkot azon  $\lambda_i$  konstans esetén metszi az  $\mathbf{y}$  pontban, amelyre  $\lambda_i T\mathbf{y} |_i = \lambda_i \mathbf{t}'_i \mathbf{y} = h_i$ , amelyből

$$\lambda_i = \frac{h_i}{\mathbf{t}'_i \mathbf{y}}.$$

Az egyenesnek az  $X$  halmazból való kilépési pontja a leghamarabb elmetszett hipersíkon van, következésképpen

$$\lambda_U = \min_i \lambda_i = \min_i \frac{h_i}{\mathbf{t}'_i \mathbf{y}}.$$

Mivel most  $\lambda T\mathbf{0} = \mathbf{0} \leq \mathbf{h}$ , ezért  $\lambda_L = -\infty$  a belépési pont. Ha az origó nincsen benne az  $X$  halmazban, akkor a kilépési pont meghatározásához hasonlóan kell eljárni.

**Algoritmus (a belépési és kilépési pont meghatározására, ha  $\mathbf{0} \leq \mathbf{h}$ )**

0. [Adott  $\mathbf{y}$  vektor esetén  $\lambda_L, \lambda_U$  kiszámítása.]
1. Legyen  $\mathbf{z} = T\mathbf{y}$ .
2. Számítsuk ki a  $\lambda_i = \frac{h_i}{z_i}$  értékeket  $i = 1, \dots, n$ .
3. Legyen  $\lambda_L = -\infty$ ,  $\lambda_U = \min_i \lambda_i$ .

Látható, hogy hasonló műveleteket kell elvégezni akkor is, ha egy  $A\mathbf{z} \leq \mathbf{b}$  egyenlőtlenségekkel meghatározott poliéder valószínűségi tartalmát kell meghatározni, vagy csillag alakú, hipersíkokkal meghatározott (nemkonvex) poliéder valószínűségét keressük. Hiperellipszoid és körkúp esetén is ki lehet számítani a belépési és kilépési konstansokat.

#### 2.3.4. Ortonormalizált becslések

A  $\Theta_2$  becslést hatékonyabbá (kisebb szórásúvá) tesszük két további módosítás beépítésével. A  $\Theta_2$  szórásnégyzete kisebb, mint  $\Theta_1$  szórásnégyzete, de még mindig elég

nagy, mert az  $S$ -en egyenletes eloszlású  $\mathbf{y}_i$  vektorok „túl véletlenszerűen” szóródnak; úgy csökkentjük a szórást, hogy „egyenletesebben” vesszük fel az  $\mathbf{y}_i$  vektorokat.

Az egyik módosítás szerint egy szabályos (de véletlenszerűen elhelyezkedő) pontrendszert alkotunk. Tekintsünk egy, az  $S$  egységgömb felületén elhelyezkedő vektorokból álló véletlen ortonormalizált  $U$  rendszert, vagyis legyen  $U = \{\mathbf{u}^i, i = 1, \dots, n \mid \mathbf{u}^i \in S, \mathbf{u}^i \mathbf{u}^j = \delta_{ij}, i, j = 1, \dots, n\}$ , amely egyenletes eloszlású a véletlen ortonormalizált rendszerek halmazán. (Az  $U$ -ban lévő  $n$  darab ortonormalizált vektor az  $S$  egyenletesebb lefedését adja, mint  $n$  darab  $S$ -en egyenletes eloszlású  $\mathbf{y}_i$  vektor.)

A másik módosítás pedig a számítási munka relatív csökkentésével jár, és a következő lépésekből áll. Tekintsük két,  $U$ -ból származó vektor normalizált összeget, vagyis legyen

$$(17) \quad \mathbf{v}^{i,j,\mathbf{s}} = \frac{1}{\sqrt{2}} (s_1 \mathbf{u}^i + s_2 \mathbf{u}^j),$$

ahol az  $i, j$  indexpár és az  $\mathbf{s}$  előjelvektor felveszi az összes lehetséges értéket a

$$J^* = \{(i, j, \mathbf{s}) \mid i = 1, \dots, n-1, j = 2, \dots, n, i < j, s_k = -1, \text{ vagy } s_k = 1, k = 1, \dots, n\}$$

halmazból. A lényegi oka annak, hogy a  $\mathbf{u}^i$  vektorok normalizált összeget használjuk az  $\mathbf{u}^i$  vektorok helyett az az, hogy egy adott  $U$  rendszerben csak  $n$  darab  $\mathbf{u}^i$  vektorunk van, de az adott  $U$  rendszerből  $2n(n-1)$  különböző  $\mathbf{v}^{i,j,\mathbf{s}}$  vektort tudunk előállítani – az egymástól különböző egyenesek száma  $n(n-1)$  lesz. Továbbá az  $n(n-1)$  egyenes előállításához ugyanúgy  $n$  darab  $T$  mátrixszal való szorzás kell csak, mint az  $n$  darab  $\mathbf{u}^i$  vektor transzformálásához, ugyanis fennáll a következő egyenlőség:

$$\mathbf{z}^{i,j,\mathbf{s}} = T \mathbf{v}^{i,j,\mathbf{s}} = \frac{1}{\sqrt{2}} (s_1 T \mathbf{u}^i + s_2 T \mathbf{u}^j).$$

Ez a második módosítás tehát csak annyi változtatással jár, hogy az  $\mathbf{u}$  generálása,  $\mathbf{v}$  vektorok előállítása és  $T$ -val való transzformálása helyett az  $\mathbf{u}$  vektorok előállítása után először ezeket  $T$ -vel transzformáljuk, és ezek után számítjuk ki – egyszerű összeadással a  $\mathbf{v}^{i,j,\mathbf{s}}$  vektorok transzformáltjait. Egy darab  $U$  rendszer esetén (amikor

$U$ -ból két vektornak vesszük minden lehetséges módon az összeget és különbségét) a teljes becslés így a következő formát ölti:

$$(18) \quad O_2 = \frac{1}{n(n-1)} \sum_{(i,j), \mathbf{s} \in J^*} e(T\mathbf{v}^{i,j,\mathbf{s}}) = \frac{1}{n(n-1)} \sum_{(i,j), \mathbf{s} \in J^*} e\left(\frac{1}{\sqrt{2}}(s_1 T\mathbf{u}^i + s_2 T\mathbf{u}^j)\right).$$

A Monte Carlo eljárás az integrál becslésére tehát abban áll, hogy  $N$  darab véletlen ortonormált  $U$  rendszeren vesszük a függvényértékek átlagát. Ezt a becslést az ortonormalizált-2-es becslésnek, vagy röviden  $O_2$  becslésnek nevezzük.

Természetesen az  $O_2$  becslésben szereplő két  $U$ -beli vektor összegeinek felhasználása helyett vehetjük  $k$  darab  $U$ -beli vektor normalizált összegét – ezáltal az egy  $U$  rendszerből előállított vektorok számát megnövelhetnénk – ezeket a becsléseket  $O_k$ -val jelöljük. Bár bizonyos esetekben ( $n \sim 10, k = 3, 4$ ) ezek az  $O_k$  becslések jobbak szoktak lenni, de az egyszerűség kedvéért a részletes ismertetéstől eltekintünk – az  $O_2$  becslések általában megfelelő numerikus hatékonyságot biztosítanak.

Az algoritmus végleges változatában a  $\lambda_L^+, \lambda_U^+$ , és a  $g$  függvény helyett az eredeti  $\lambda_L, \lambda_U$  állandókat és az  $e$  függvényt használjuk; a  $g(\mathbf{y})$  and  $g(-\mathbf{y})$  függvényeket egyszerre számítjuk ki, megfeleltetve azon vektorok számát, amelyekre a  $\lambda_L, \lambda_U$  állandókat ki kell számítanunk.

Az alábbiakban kissé rövidítve, de lényegében használható módon leírjuk az  $O_2$  becslés kiszámításának az algoritmusát.

$O_2$  algoritmus (általános eset, egy  $U$  rendszer)

1. Generáljuk  $U = \{\mathbf{u}^i\}$ -t és számítsuk ki az  $\bar{\mathbf{u}}^i = T\mathbf{u}^i, i = 1, \dots, n$  transzformáltakat.
2. Állítsuk be az  $S = 0$  kezdeti értéket.
3. Állítsuk elő az összes lehetséges  $\mathbf{z} = \mathbf{z}^{i,j,\mathbf{s}} = T\mathbf{v}^{i,j,\mathbf{s}} = \frac{1}{\sqrt{2}}(s_1 \bar{\mathbf{u}}^i + s_2 \bar{\mathbf{u}}^j), (i, j, \mathbf{s}) \in J^*$  vektort, és minden  $\mathbf{z}$  vektorra végezzük el a következő lépést.

4. Kezdjük a  $e(T\mathbf{v})$  függvény kiszámítását

számítsuk ki a  $\lambda_L = \min \{|\lambda| | \lambda \mathbf{z} \in X\}, \lambda_U = \max \{|\lambda| | \lambda \mathbf{z} \in X\}$  állandókat,

Ha  $\lambda \mathbf{z}$ -nek és  $X$ -nek üres a metszete, akkor legyen  $\lambda_L = \lambda_U = 0$ ,

ha  $\lambda_U \geq \lambda_L \geq 0$ , akkor legyen  $S = S + K(\lambda_U) - K(\lambda_L)$ ,



ha  $\lambda_U \geq 0 \geq \lambda_L$ , akkor legyen  $S = S + K(\lambda_U) + K(-\lambda_L)$ ,

ha  $0 \geq \lambda_U \geq \lambda_L$ , akkor legyen  $S = S + K(-\lambda_L) - K(-\lambda_U)$ .

befejeztük a  $e(T\mathbf{v})$  függvényérték kiszámítását.

5. A keresett valószínűség becsléseként adjuk át a  $p = S/[2n(n-1)]$  értéket.

A számítási munka (relatív) csökkentését azzal értük el, hogy az  $O_2$  becslésben  $O(n^2)$  mátrix szorzás helyett csak  $O(n)$   $T$ -vel való szorzást hajtottunk végre (még néhány összeadást és skalárszorzást). Az ortonormalizált becslések alkalmazását tekinthetjük úgy is, mint ellentétes (antithetic) változók használatát szórás csökkentés céljából. De úgy is tekinthetjük az  $O_2$  becslést, mint egy determinisztikus integrálási formula véletlenné tételét – ami által könnyen kaphatunk hibabecslést.

## 2.4. Korlátok a valószínűségekre

Az eljárás a közismert Boole–Bonferroni egyenlőtlenségeken alapul, és  $n$ -dimenziós eloszlásfüggvények értékeit lehet korlátozni (kiszámítani) a segítségével, többféle ismert eloszlás esetén. A módszer gyakorlati alkalmazhatósága azon múlik, hogy az adott  $n$ -dimenziós eloszlás egy-, két-, vagy esetleg további alacsonyabb dimenziós vetületi (perem)eloszlásfüggvényének értékeit számítástechnikailag hatékonyan meg tudjuk-e határozni.

### 2.4.1. Valószínűségek korlátozása

Kiindulunk a Boole–Bonferroni egyenlőtlenségekből:

**18. Tétel (Boole).** *Legyenek  $A_1, A_2, \dots, A_n$  tetszőleges események, ekkor*

$$\mathbf{P}\{\cup_{i=1}^n A_i\} \leq \sum_{i=1}^n \mathbf{P}\{A_i\},$$

$$\mathbf{P}\{\cap_{i=1}^n A_i\} \geq \sum_{i=1}^n \mathbf{P}\{A_i\} - (n-1),$$

**Bizonyítás.** Az első egyenlőtlenség közismert. A második levezetéséhez tekintünk először a

$$\mathbf{P}\{\overline{\cap_{i=1}^n A_i}\} = \mathbf{P}\{\cup_{i=1}^n \overline{A_i}\} \leq \sum_{i=1}^n \mathbf{P}\{\overline{A_i}\} = \sum_{i=1}^n [1 - \mathbf{P}\{A_i\}] = n - \sum_{i=1}^n \mathbf{P}\{A_i\}$$

egyenlőtlenséget. Ennek felhasználásával kaphatjuk, hogy

$$\mathbf{P}\{\cap_{i=1}^n A_i\} = 1 - \mathbf{P}\{\overline{\cap_{i=1}^n A_i}\} \geq 1 - [n - \sum_{i=1}^n \mathbf{P}\{A_i\}] = \sum_{i=1}^n \mathbf{P}\{A_i\} - (n - 1).$$

Tegyük fel, hogy az együttes események közül néhánynak a valószínűsége is adva van. (Három eseményen szemléltetjük az elgondolást.) Legyenek adva a  $\mathbf{P}\{A_1\} = 0.3, \mathbf{P}\{A_2\} = 0.4, \mathbf{P}\{A_3\} = 0.3, \mathbf{P}\{A_1 \cap A_2\} = 0.05$  valószínűségek. Vezessük be a következő jelöléseket:

$$\begin{aligned} x_1 &= \mathbf{P}\{A_1 \setminus A_2 \cup A_3\}, & x_2 &= \mathbf{P}\{A_2 \setminus A_1 \cup A_3\}, & x_3 &= \mathbf{P}\{A_3 \setminus A_1 \cup A_2\}, \\ x_4 &= \mathbf{P}\{A_1 \cap A_2 \setminus A_3\}, & x_5 &= \mathbf{P}\{A_1 \cap A_3 \setminus A_2\}, & x_6 &= \mathbf{P}\{A_2 \cap A_3 \setminus A_1\} \\ x_7 &= \mathbf{P}\{A_1 \cap A_2 \cap A_3\} \end{aligned}$$

Ekkor egy lineáris programozási feladatot írhatunk fel az együttes esemény valószínűségére:

$$p = \mathbf{P}\{A_1 \cup A_2 \cup A_3\} = x_1 + x_2 + \cdots + x_7.$$

$$\begin{array}{cccccccl} x_1 & & +x_4 & & +x_6 & +x_7 & = & 0.3 \\ & x_2 & & +x_4 & +x_5 & & +x_7 & = & 0.4 \\ & & x_3 & & +x_5 & +x_6 & +x_7 & = & 0.3 \\ & & & x_4 & & & +x_7 & = & 0.05 \end{array}$$

Ha ezt a feladatot a

$$\max(x_1 + x_2 + \cdots + x_7), \text{ illetve a } \min(x_1 + x_2 + \cdots + x_7)$$

célfüggvények esetén megoldjuk, akkor a három halmaz uniójának  $p$  valószínűségére egy felső, illetőleg egy alsó korlátot kapunk. Hasonló lineáris programozási feladatokat írhatunk fel tetszőleges  $A_1, \dots, A_n$  halmazok valószínűségére, ha valamilyen  $\mathbf{P}\{A_{i_1} \cap \dots \cap A_{i_l}\}$  valószínűségek ismertek (meghatározhatók).

Az általános esetben a következő tételt használjuk fel.

**19. Tétel (Poincare tétele.).** *Legyenek  $A_1, A_2, \dots, A_n$  tetszőleges események. Ekkor*

$$\mathbf{P}\{\cup_{i=1}^n A_i\} = S_1 - S_2 + S_3 - \dots + (-1)^{n-1} S_n,$$

ahol

$$S_k = \sum_{i_1} \sum_{i_2} \dots \sum_{i_k} \mathbf{P}\{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}\}, \forall i_1, \dots, i_k = 1, \dots, n.$$

**20. Tétel (Bonferroni egyenlőtlenségek.).** *Legyenek  $A_1, A_2, \dots, A_n$  tetszőleges események. Ekkor*

$$\begin{aligned} \mathbf{P}\{\cup_{i=1}^n A_i\} &\geq \sum_{k=1}^{2m} (-1)^{k-1} S_k, \text{ ha } m \geq 1, \\ \mathbf{P}\{\cup_{i=1}^n A_i\} &\leq \sum_{k=1}^{2m+1} (-1)^{k-1} S_k, \text{ ha } m \geq 0. \end{aligned}$$

A Bonferroni egyenlőtlenségek alapján alsó és felső korlátok adhatók a  $\mathbf{P}\{\cup_{i=1}^n A_i\}$  együttes valószínűségekre (vagy a  $\mathbf{P}\{\cap_{i=1}^n A_i\}$  metszet valószínűsége).

**21. Tétel (Takács tétele).** *Legyenek  $A_1, A_2, \dots, A_n$  tetszőleges események és legyen  $\xi$  azon események száma, amelyre az  $A_1, A_2, \dots, A_n$  események bekövetkeznek. Ekkor*

$$S_k = \mathbf{E} \left[ \binom{\xi}{k} \right].$$

**Bizonyítás.** Legyen  $\xi_i$  az indikátor valószínűségi változója az  $A_i$  eseménynek (tehát  $\xi_i = 1$ , ha  $A_i$  bekövetkezett, egyébként  $\xi_i = 0$ ). Ekkor  $\xi = \sum_{i=1}^n \xi_i$ . Felírjuk a binomiális együtthatókra vonatkozó Cauchy tételt:

$$\binom{\xi}{k} = \sum_{k_1 \geq 0} \dots \sum_{k_n \geq 0} \binom{\xi_1}{k_1} \dots \binom{\xi_n}{k_n},$$

Mivel itt  $\xi_i$  csak 0 vagy 1 lehet, ezért a  $\binom{\xi_i}{k_i} = \xi_i$ , egyébként a binomiális együttható mindig nulla. Ekkor

$$\binom{\xi}{k} = \sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} \xi_{i_1} \cdots \xi_{i_k}.$$

Mivel  $\mathbf{P}\{A_i\} = \mathbf{E}(\xi_i)$  és  $S_k = \sum_{i_1} \cdots \sum_{i_k} \mathbf{P}\{A_{i_1} \cap \dots \cap A_{i_k}\} = \sum_{i_1} \cdots \sum_{i_k} \mathbf{E}(\xi_{i_1} \cdots \xi_{i_k})$ , a tétel állítását beláttuk.

#### 2.4.2. Szimuláció és korlátozás

Tekintsük az  $\mathbf{R}^n$ -ben elhelyezkedő  $A$  halmaz valószínűségét egy  $f(\mathbf{x})$  sűrűségfüggvénnyel és  $F(\mathbf{x})$  eloszlásfüggvénnyel megadott eloszlás esetén, amely

$$\mathbf{P}\{A\} = \int_A f(\mathbf{x}) d\mathbf{x}.$$

Legyenek adva az  $A_1, \dots, A_n$  események az  $A_i = \{\mathbf{x} | -\infty < x_i \leq a_i\}$  formában adva, ahol  $a_1, \dots, a_n$  adott konstansok. Ha most  $A = A_1 \cap A_2 \cap \dots \cap A_n$  formában állítható elő, akkor a  $\mathbf{P}\{A\}$  valószínűség az eloszlásfüggvény értéke az  $\mathbf{a} = (a_1, \dots, a_n)$  pontban:  $\mathbf{P}\{A\} = F(\mathbf{a})$ . Alkalmazzuk most az előbb ismerttetett Boole–Bonferoni egyenlőtlenségeket erre a feladatra. (Az eljárás használható másmilyen  $A_i$  illetőleg  $A$  halmaz esetén – például egy téglalap valószínűségének meghatározására, de az egyszerűség kedvéért csak ezt az algoritmust írjuk le.)

A módszer gyakorlati alkalmazhatósága attól függ, hogy az egyszeres  $\mathbf{P}\{A_i\}$  valószínűségeken kívül ki tudjuk-e számítani (számítástechnikailag hatékonyan) a kétszeres  $\mathbf{P}\{A_i \cap A_j\}$ , háromszoros  $\mathbf{P}\{A_i \cap A_j \cap A_l\}$ , illetőleg általában a  $k$ -szoros  $\mathbf{P}\{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}\}$  metszet-valószínűségeket.

Általában egy  $k$  darab halmaz metszetéből álló halmaz valószínűségének meghatározása egy  $k$ -dimenziós integrált jelent, amit 2-5 dimenzióig el lehet végezni az ismert numerikus integrálási technikákkal, 5-10 dimenzió között pedig ezeknek valamilyen speciális változatával. Egy másik kérdés, ami felmerül ilyenkor, hogy hány darab ilyen valószínűséget kell kiszámítani, hiszen egy  $n$ -dimenziós  $\mathbf{P}\{A\}$  valószínűség kiszámítása esetén  $\binom{n}{k}$  darab  $k$ -szoros valószínűséget kell kiszámítani. Amikor  $n$  nagy, akkor nem számítjuk ki az összes lehetséges  $k$ -szoros valószínűséget, hanem a

cseresznyefa vagy a hipercseresznyefa módszer segítségével csak némelyeket határozzunk meg, ezzel növeljük a hatékonyságot (lásd Szántai és Bukszár eredményeit).

Legyen

$$\bar{S}_k = \sum_{1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n} \mathbf{P}\{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}\}, k = 1, 2, \dots, n$$

Ekkor

$$\begin{aligned} \bar{S}_1 &= n - S_1, \\ \bar{S}_2 &= \binom{n}{2} - (n-1)S_1 + S_2, \\ &\vdots \\ \text{általában } \bar{S}_k &= \binom{n}{k} + \sum_{i=1}^k \binom{n-i}{k-i} (-1)^i S_i, k = 1, 2, \dots, n, \end{aligned}$$

ami teljes indukcióval bizonyítható.

Egy olyan leegyszerűsített algoritmust írunk le, amelyben csak azt tesszük fel, hogy az egyszeres és a kétszeres valószínűségeket meg tudjuk határozni (valamilyen determinisztikus numerikus integrálással, amelyeknek a hibája  $10^{-10}$ -nél kisebb, ami elhanyagolható) – tehát  $\bar{S}_1$  és  $\bar{S}_2$  pontosan számítható. A  $\mathbf{P}\{A\} - [1 - \bar{S}_1 + \bar{S}_2]$  maradék érték meghatározására Monte Carlo integrálást fogunk használni, vagyis a determinisztikus integrálással kombináljuk a szimulációt. Ez a maradék érték kis hibájú (szórású) lesz, mert az értéke is kicsi.

Legyen  $A = A_1 \cap \dots \cap A_n$ , akkor fennáll a következő összefüggés:

$$\mathbf{P}\{A\} = 1 - \mathbf{P}\{\bar{A}_1 \cup \dots \cup \bar{A}_n\} = 1 - \bar{S}_1 + \bar{S}_2 - \bar{S}_3 \dots (-1)^n \bar{S}_n.$$

Három becslést fogunk megadni a  $\mathbf{P}\{A\}$  valószínűségre, amelyeket  $\nu_0, \nu_1, \nu_2$ -vel fogunk jelölni. A  $\mathbf{P}\{A\}$ -ra vonatkozó végleges becslés pedig ezeknek egy kombinációja lesz.

Generálunk az adott  $f(\mathbf{x})$  sűrűségfüggvényű eloszlást követő  $\boldsymbol{\xi}$  vektorból, ezek az  $\mathbf{x}_1, \dots, \mathbf{x}_N$  realizációk, ezek felhasználásával számítjuk a következő becsléseket:

$$\begin{aligned}
\hat{P}_0 : \nu_0 &= \mathbf{P}\{A_1 \cap \dots \cap A_n\} & \Rightarrow \hat{P}_0 &= \frac{N_0}{N} \\
\hat{P}_1 : \nu_1 &= \bar{S}_2 - \bar{S}_3 + \dots + (-1)^n \bar{S}_n, & \Rightarrow \hat{P}_1 &= 1 - \bar{S}_1 + \nu_1 \\
\hat{P}_2 : \nu_2 &= -\bar{S}_3 + \dots + (-1)^n \bar{S}_n, & \Rightarrow \hat{P}_2 &= 1 - \bar{S}_1 + \bar{S}_2 - \nu_2
\end{aligned}$$

Hasonló módon lehet eljárásokat felépíteni abban az esetben, ha olyan determinisztikus numerikus eljárásaink vannak, amelyek segítségével az  $m = 1, 2, \dots, k$ -szoros valószínűségek is kiszámíthatók. Ezekben az esetekben a  $\binom{n}{m}$  számú  $m$ -szeres valószínűségek közül csak egyeseket fogunk kiszámítani.

## 2.5. A STABIL modell

Leírjuk az egyik legegyszerűbb sztochasztikus lineáris programozási (SLP) modellt, amelyben a feltételek és a célfüggvény is lineáris, de egy együttes valószínűségi korlát van előírva bizonyos feltételek teljesülésére. Ennek a STABIL modellnek a kiindulási pontja a következő determinisztikus modell:

$$\begin{aligned}
(19) \quad & \min \mathbf{c}'\mathbf{x}, \\
& \text{f.h. } A\mathbf{x} \geq \mathbf{h}, \\
& B\mathbf{x} \leq \mathbf{b}.
\end{aligned}$$

Tegyük fel, hogy a jobboldali vektor egyes komponensei valószínűségi változók, például  $\mathbf{h}$  adott vektor helyett egy  $\boldsymbol{\xi}$  valószínűségi vektorváltozót kell tekintenünk. Ekkor a determinisztikus modell értelmét veszti és egy sztochasztikus döntési elvet kell bevezetni. Ezt meg lehet tenni együttes valószínűségi korlát bevezetésével; ezekre a feltételekre egy valószínűségi korlátot írunk elő és ezen feltétel mellett optimalizálunk:

$$\begin{aligned}
(20) \quad & \min \mathbf{c}'\mathbf{x}, \\
& \text{f.h. } G(\mathbf{x}) = \mathbf{P}\{A\mathbf{x} \geq \boldsymbol{\xi}\} \geq p, \\
& B\mathbf{x} \leq \mathbf{b}.
\end{aligned}$$

Az általánosság megszorítása nélkül feltesszük, hogy  $E(\boldsymbol{\xi}) = 0, D^2(\boldsymbol{\xi}) = 1$ . Az általános eset ebből egy  $T\boldsymbol{\xi} + \mathbf{h}$  lineáris transzformációval előállítható.

Megjegyezzük, hogy formálisan tetszőleges valószínűségi korlát átírható egy várható értékre adott korlátnak, hiszen ha az  $A$  esemény teljesülésének valószínűségére feltesszük a  $\mathbf{P}\{A\} \geq p$  korlátot, akkor ezzel ekvivalens lesz az  $\mathbf{E}(\chi_A) \geq p$  feltevés, ahol  $\chi_A$  az  $A$  esemény karakterisztikus (indikátor) változója. (Számítási nehézségeink azonban ettől nem szoktak megoldódni, de egységesen lehet kezelni a kétlépcsős típust és a valószínűségi korlátos feladatot.)

A STABIL modell általános formája a következőképpen adható meg:

$$\begin{aligned} \min f(\mathbf{x}), \\ \text{f.h. } G(\mathbf{x}) = \mathbf{P}\{g_i(\mathbf{x}) \geq \xi_i, i = 1, \dots, m\} &\geq p, \\ g_i\mathbf{x} &\geq b_i, i = m + 1, \dots, m + M. \end{aligned}$$

ahol feltesszük, hogy az  $f(\mathbf{x})$  függvény konvex, a  $g_i, i = 1, \dots, m + M$  függvények pedig kvázikonkávak, a  $\xi_1, \dots, \xi_m$  valószínűségi változók együttes eloszlása logkonkáv. Ezek a feltevések biztosítják, hogy a megengedett megoldások halmaza konvex lesz (ha nem üres).

### 2.5.1. Konvexitás és valószínűségek

A felírt sztochasztikus programozási modell esetén az első felmerülő kérdés, hogy konvex-e a megengedett  $\mathbf{x}$  megoldások halmaza. A determinisztikus lineáris feltételek nyilván konvex megengedett megoldási halmazt adnak. A nemlineáris  $G(\mathbf{x})$  függvény a  $\boldsymbol{\xi}$  valószínűségi változó eloszlásától függ. Prékopa ismertetett eredménye alapján az  $\{\mathbf{x} \mid G(\mathbf{x}) \geq p\}$  halmaz konvex lesz logaritmikusan konkáv többdimenziós eloszlások esetén.

A második felmerülő kérdés, hogyan lehet meghatározni a  $G(\mathbf{x})$  függvény értékeit. Mivel ez jelen esetben a  $\boldsymbol{\xi}$  valószínűségi változó eloszlásfüggvénye, ezért a kérdés numerikus természetű. Az előző alpontokban két numerikus eljárást ismertettünk  $n$ -dimenziós valószínűségek (eloszlásfüggvények) kiszámítására, természetesen más eljárások is használhatók. A durva Monte Carlo módszer majdnem minden esetben

alkalmazható, még akkor is, ha szóban forgó valószínűségi változó sűrűségfüggvénye nem ismert, csak a konstrukciója (mint például a Prékopa definiálta többdimenziós gamma eloszlás esetén).

### 2.5.2. Optimalizálási algoritmus

A harmadik kérdés, hogy milyen nemlineáris optimalizálási eljárással lehet az optimumot meghatározni. Erre egy egyszerű, megengedett irányok elnevezésű módszert ismertetünk, bár lényegében minden nemlineáris optimalizálási algoritmus használható lenne, amely lineáris és kvázikonkáv feltételek esetén konvergens. Ezeket azért nem ismertetjük, mert a megfelelő hatékonyság eléréséhez elég részletesen kellene a számítógépes eljárásokat leírni. (Megjegyezzük, hogy pillanatnyilag a Mayer által kidolgozott általánosított redukált gradiens módszer – GRG – tűnik a leghatékonyabb eljárásnak.)

Tekintsük még egyszer a

$$(21) \quad \begin{aligned} \min f(\mathbf{x}) &= \mathbf{c}'\mathbf{x}, \\ \text{f.h. } G(\mathbf{x}) &= \mathbf{P}\{\mathbf{A}\mathbf{x} \geq \boldsymbol{\xi}\} \geq p, \\ B\mathbf{x} &\leq \mathbf{b} \end{aligned}$$

feladatot, ahol a  $\boldsymbol{\xi}$  valószínűségi változó eloszlása legyen standard normális,  $R$  korrelációs mátrixszal. Tegyük fel az egyszerűség kedvéért, hogy a megengedett megoldások halmaza nem üres és korlátos, továbbá a  $G, f$  függvények differenciálhatóak az egész térben. Ezen feltevések mellett a feladat konvex optimalizálási probléma. Megoldó algoritmusunk a szimplex algoritmushoz hasonlóan két fázisból áll, először egy megengedett megoldást keresünk, majd a második fázisban megengedett megoldások egy olyan sorozatát állítjuk elő, amelyen a célfüggvény értéke csökken (javul).

A második fázis leírásával kezdjük az algoritmus ismertetését, amely a megengedett irányok elnevezésű iteratív megoldó algoritmus (Zoutendijk P2 algoritmus) – ez két részből áll: egy megengedett irány megkereséséből és egy irány menti optimalizálásból. Tegyük fel, hogy rendelkezésünkre áll egy  $\mathbf{x}_k$  vektor, amely a (??)



alapfeladat megengedett megoldása, akkor a következő eljárással lehet két lépésben az  $\mathbf{x}_k$ -ből az  $\mathbf{x}_{k+1}$  megengedett megoldást előállítani. Legyen  $\vartheta > 0$  egy kicsi, de rögzített szám. Az első lépés az úgynevezett iránykeresés feladata:

$$\begin{aligned}
 & \min y, \\
 (22) \quad & \text{f.h. } G(\mathbf{x}_k) + \nabla G(\mathbf{x}_k)[\mathbf{x} - \mathbf{x}_k] + \vartheta y \geq p, \\
 & B\mathbf{x} \leq \mathbf{b}, \\
 & \nabla f(\mathbf{x}_k)[\mathbf{x} - \mathbf{x}_k] \leq y,
 \end{aligned}$$

amelyben az  $(\mathbf{x}, y)$ ,  $n+1$  darab ismeretlen meghatározása a feladatunk (a feladat feltételei lényegében az eredeti feladat linearizáltjai, illetőleg egy csökkenő irány keresése). Ennek a feladatnak mindig van megengedett megoldása, hiszen a  $(\mathbf{x}_k, 0)$  mindig kielégíti a feltételeket. Az  $\mathbf{x}$  vektor egy korlátos tartományban változhat, így a célfüggvény alulról korlátos, tehát az iránykereső feladatnak van véges optimuma. A  $(??)$  feladat megoldása után kapjuk az  $(\mathbf{x}^*, y^*)$  optimális megoldást. Vegyük észre, hogy  $y^* \leq 0$ , hiszen volt olyan megengedett megoldás (nevezetesen  $(\mathbf{x}_k, 0)$ ), amelyen  $y = 0$  volt. Itt két eset lehetséges.

(i) Az  $y^* = 0$  egyenlőség fennáll, ekkor az eljárás végetér, az előző  $\mathbf{x}_k$  megengedett megoldás az eredeti feladatnak optimális megoldása is.

(ii) Az  $y^* < 0$  egyenlőtlenség igaz – ekkor újabb  $\mathbf{x}_{k+1}$  megengedett megoldást állítunk elő az  $\mathbf{x}^*$  segítségével. Minimalizáljuk az  $f(\mathbf{x})$  függvényt, az  $(??)$  feladat feltételei esetén, az  $\mathbf{x}_k + \lambda[\mathbf{x}_k^* - \mathbf{x}_k]$  félegyenes mentén, vagyis legyen feladatunk

$$\begin{aligned}
 & \min_{\lambda} (f(\mathbf{x}_k + \lambda[\mathbf{x}_k^* - \mathbf{x}_k])), \\
 (23) \quad & \text{f.h. } G(\mathbf{x}) = \mathbf{P}\{A\mathbf{x} \geq \boldsymbol{\xi}\} \geq p, \\
 & B\mathbf{x} \leq \mathbf{b}.
 \end{aligned}$$

Ez az úgynevezett irány-menti minimalizálás feladata. Legyen ennek a feladatnak egy optimális megoldása  $\lambda_k = \arg \min_{\lambda} (f(\mathbf{x}_k + \lambda[\mathbf{x}_k^* - \mathbf{x}_k]))$  és definiáljuk a

következőképpen az új megengedett megoldást:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k[\mathbf{x}_k^* - \mathbf{x}_k].$$

Kvázikonkáv feltételi függvények és célfüggvény mellett a megengedett irányok módszerének konvergenciája alapján igaz a következő:

**22. Tétel.** *Az algoritmus vagy véges sok lépésben véget ér (valamilyen  $k$  index esetén  $y^* = 0$ ), vagy pedig az előállított  $\mathbf{x}_k, \mathbf{x}_{k+1}, \dots$  megengedett megoldások sorozata konvergál a  $(??)$  probléma optimális megoldásához.*

A megoldó algoritmus teljességéhez már csak az hiányzik, hogy egy olyan eljárást adjunk, amely előállítja a  $(??)$  alapeladat egy  $\mathbf{x}_1$  megengedett megoldását (első fázis). Erre az előbb leírt, a második fázist alkotó algoritmus is megfelelő. Ugyanis tekintsük a

$$(24) \quad \begin{aligned} &\min -G(\mathbf{z}), \\ &B\mathbf{z} \leq \mathbf{b} \end{aligned}$$

feladatot. Itt nem okvetlenül szükséges egy optimális megoldás megadása, hanem elég ennek a feladatnak egy olyan  $\mathbf{z}_l$  megengedett megoldását meghatározni, amely esetén  $G(\mathbf{z}_l) > p$  fennáll.

Ezek szerint a  $(??)$  első fázis a következő lépésekből áll. Keresünk lineáris programozási eszközökkel egy olyan  $\mathbf{z}_1$  vektort, amelyre  $B\mathbf{z}_1 \leq \mathbf{b}$  teljesül. Ezek után az elsőfázisos feladat iránykereső feladatát oldjuk meg, vagyis a

$$(25) \quad \begin{aligned} &\min y \\ &-G(\mathbf{z}_k) + \nabla(-G(\mathbf{z}_k))[\mathbf{z} - \mathbf{z}_k] \leq y, \\ &B\mathbf{z} \leq \mathbf{b} \end{aligned}$$

minimalizálást. Természetesen itt vagy találunk egy megfelelő  $\mathbf{z}_l$  vektort (amelyre  $G(\mathbf{z}_l) > p$  igaz) és akkor folytathatjuk a második fázissal, vagy pedig nincs ilyen

vektor, és akkor tudjuk, hogy az eredeti (??) feladatnak sincs megengedett megoldása. Ha találtunk megfelelő  $\mathbf{z}_l$  vektort, amely az (??) alapfeladatnak megengedett megoldása, akkor áttérhetünk a második fázis algoritmusára.

### 2.5.3. Iránymenti optimalizálás

Az iránymenti minimalizálás egy lineáris  $f(\mathbf{x}) = \mathbf{c}'\mathbf{x}$  költségfüggvény esetén egyszerű: az adott,  $\mathbf{x}_k$  pontból induló félegyenesnek a megengedett megoldások tartományának határával való metszéspontját kell meghatározni. Ha a félegyenes egy lineáris feltételt metsz el először, akkor feladat egy félegyenes és egy hipersík metszésének meghatározásából áll.

Ha a félegyenes a valószínűségi feltételt metszi el először, akkor ennek a metszéspontnak (a  $\lambda_k$  állandónak) a meghatározása nem könnyű: egy félegyenes és egy „zajos” függvény metszéspontját kell előállítani; zajosnak nevezünk egy  $G(\mathbf{x})$  függvényt, ha egy adott  $\mathbf{x}$  pontban nem a pontos  $G(\mathbf{x})$  függvényérték áll a rendelkezésünkre, csak egy  $\tilde{G}(\mathbf{x}) = G(\mathbf{x}) + \varepsilon$ , additív zajjal terhelt becslés. Egy ilyen feladatra használható algoritmus az utolsó részben van leírva, az SRA gyökkereső algoritmus. Természetesen lehet másmilyen, zajos függvény esetén működő heurisztikus eljárást is alkalmazni, bár ilyenkor fontos arra figyelni, hogy az új pont biztosan megengedett legyen (egyébként az algoritmus nem működik).

Ha például az aranymetszés, vagy görbeillesztéses módszerrel keressük meg a gyököt, akkor célszerű „visszalépni” a végső megoldás megadásakor. Ezt a következőképpen valósítjuk meg. Tegyük fel, hogy a függvényérték kiszámítása egy véletlen zajjal történik, ekkor legyen ennek a szórása  $\sigma$ . Ha az  $x_k$  pontból kiindulva a metszéspont valamilyen  $y_k$  pont lenne, akkor az  $x_{k+1} = y_k - 4\sigma(y_k - x_k)$  pontot fogadjuk el metszéspontként, ami nagy valószínűséggel a megengedett megoldások tartományán belül marad.

### 2.5.4. A STABIL alkalmazása

A népgazdasági tervezés folyamán használtak lineáris programozást is: a gazdaságot egymástól lényegében független szektorokra osztották, amelyeket minden szektorban

szereplő változókkal fűztek össze. A STABIL modellt először a magyar villamosenergiaipar lineáris szektormodellje sztochasztizált változatára alkalmaztuk 1972-ben, ahol négy jobboldali változó volt véletlen. A kiindulásként vett modell egy lineáris programozási feladat volt, amelyet a Tervhivatal (teljes népgazdaságot leíró) lineáris programozási modelljének a villamosenergiaiparra vonatkozó része (szektormodellje) volt.

A változók között szerepeltek a különböző (már működő – exogén és a tervezési időszakban elkészülő – endogén) erőművekben termelt energia értékei, a különböző fűtőanyagok, az elektromos energia exportjának és importjának értékei, beruházási változók (amelyek a beruházásokra rendelkezésre álló pénzösszegre vonatkozó korlátokat kielégítették). A feltételek között szerepeltek a munkaerőre, a beruházásokra, a külkereskedelmi egyenlegre, az elektromos energia iránti igényekre vonatkozó és más pénzügyi feltételek. Négy feltételt tekintettünk sztochasztikusnak. Az első két feltétel a külkereskedelmi negatív szaldó értékére vonatkozó korlátot adta meg dollár és rubel relációban, a harmadik az input-output táblából az elektromos energiára vonatkozó feltétel volt. A negyedik feltétel a minimális elektromos energiára követelt meg korlátot. A jobboldali valószínűségi változók várható értékét és korrelációmátrixát az MVMT szakemberei adták meg.

#### **2.5.5. Numerikus példa**

Tekintsük a következő, két determinisztikus döntési változót, két determinisztikus feltételt és két feltételi sor együttes teljesülésének valószínűségére alsó korlátot tartalmazó numerikus feladatot:

$$\begin{aligned} & \min(x_1 + x_2), \\ \mathbf{P} \left\{ \begin{array}{l} 3x_1 + x_2 - 6 \geq \beta_1 \\ x_1 + 8x_2 - 8 \geq \beta_2 \end{array} \right\} & \geq p, \\ x_1 + 4x_2 & \geq 4, \\ 3x_1 + x_2 & \geq 3, \\ x_1 \geq 0, x_2 & \geq 0, \end{aligned}$$

ahol  $\beta_1, \beta_2$  normális eloszlású valószínűségi változók, amelyek várható értéke 0, szórásuk 1, korrelációs együtthatójuk pedig  $\varrho$ .

Ha a várható érték programozást alkalmazzuk a feladatra (vagyis a valószínűségi feltételt elhagyjuk,  $\beta_1 = 0, \beta_2 = 0$  értékeket helyettesítünk), akkor a keletkező lineáris programozási feladat optimális megoldása  $(\frac{40}{23}, \frac{18}{23})$  a célfüggvény értéke 2.521 lenne.

A sztochasztikus példára vonatkozó numerikus számítási eredményeket az 3. Táblázatban közöljük. Az első sor az eredetileg megadott optimális megoldást, a második Szántai (egy általa kifejlesztett sztochasztikus programozási programcsomaggal számított) eredménye, a harmadik pedig az SRA módszer segítségével kapott megoldás jellemzőit tartalmazza (a legpontosabb eredmény a második sorban van). (A harmadik sorban megadott eredmények esetén a  $\beta_1, \beta_2$  valószínűségi változók együttes eloszlásfüggvényét egy duplapontosságú, sorfejtésen alapuló függvénnyel számítottuk ki.) Az első sorban adott eredményeket egy CDC 3300 számítógépen mintegy 20 perc alatt lehetett kiszámítani 1971-ben. Ma egy ilyen eredmény 5 másodpercnél rövidebb idő alatt kiszámítható.

## 2.6. Megoldó algoritmusok

A sztochasztikus programozás valószínűségi korlátos modelljei a szokásos (a megengedett megoldások tartományára és a célfüggvényre vonatkozó) konvexitási feltételek esetén konvex nemlineáris programozási feladatok. Ezért lényegében akármilyen

$k =$	módszer	$\mathbf{x}_{STO}$	$f(\mathbf{x}_{STO})$	$\mathbf{P}\{\mathbf{x}_{STO}\}$
1000	Deák	(2.0650, 0.8880)	2.9530	0.8000
	Szántai	(1.9977, 0.9015)	2.8992	0.8000
	$SRA_{det}$	(1.9686, 0.9455)	2.9140	0.8003

3. táblázat. Valószínűséggel korlátozott példa:  $p = 0.8, \varrho = 0.9$ .

konvergens nemlineáris programozási megoldó algoritmus használható a sztochasztikus modellek optimális megoldásának meghatározására. A gyakorlatban a legjobb megoldó algoritmusnak a vetített gradienses eljárás bizonyult [?], bár logaritmikus büntetőfüggvénnyel [?], primál-duál módszerrel [?], [?] is megoldottak valószínűséggel korlátozott feladatokat. Ez utóbbi cikkek érdekessége, hogy a valószínűséggel korlátozott modell duálját is megadja.

Külön kiemeljük a SUMT módszer alkalmazását, amikor is a  $\log\left(\frac{1}{F(x)-p}\right)$  büntetőfüggvényt használjuk. Ha ugyanis az eloszlás logkonkáv, akkor az  $F(x)$  eloszlásfüggvény is az, és az  $F(x) - p \geq 0$  valószínűségi korlát függvény is logkonkáv, aminek a logaritmusa a megengedett megoldások halmazán belül egy konkáv függvény lesz.

A pillanatnyilag ismert számítógépes lehetőségek ismeretében a számítástechnikailag megoldható feladatok méretére vonatkozó korlátokat röviden úgy foglalhatjuk össze, hogy olyan valószínűségi korláttal rendelkező feladatok oldhatók meg nehézség nélkül, amelyekben a valószínűségi vektorváltozó 20-30 dimenziós. A determinisztikus feltételekre és változókra vonatkozóan nincs mennyiségi korlát – pár száz feltétel és változó kezelhetőnek tűnik.

### 3. Kétlépcsős modellek

Kétlépcsős modellnek nevezzük azokat a sztochasztikus programozási modelleket, amelyekben kétszer hozhatunk döntést. Ezekben az esetekben lehetőségünk van arra, hogy az első döntésünk meghozatala ( $\mathbf{x}$  meghatározása) után megfigyeljük a véletlen  $\xi$  értékét, és ennek alapján újabb, pótlólagos  $\mathbf{y}$  döntéssel korrigáljuk a véletlen által kialakult helyzetet. Ilyen feladatot fogalmaztunk meg az 1.5 fejezetben leírt mintapélda kapcsán is. (A feladatot pótlási – recourse – feladatnak is nevezzük.)

Ennek a gondolatmenetnek az analógiájára lehet többlépcsős modelleket is felírni, amelyekre akkor van szükségünk, ha a döntéshozatal szekvenciálisan történik: döntünk  $\mathbf{x}$  felől, megfigyeljük  $\xi_1$  értékét, döntünk  $\mathbf{y}_1$ -ről, az első pótlólagos változó-csoportba tartozó változókról, megfigyeljük  $\xi_2$  értékét, döntünk  $\mathbf{y}_2$ -ről és így tovább.

#### 3.1. A kétlépcsős modelltípus alapjai

Először megfogalmazzuk a feladat alaptípusát, aztán a többlépcsős modellt, majd az alapmodell néhány fontos tulajdonságát állapítjuk meg, végül néhány speciális esetet ismertetünk ebben a részben.

##### 3.1.1. Lineáris modell

Kétféleképpen lehet bevezetni a kétlépcsős modell tárgyalását: egy kiindulásként vett determinisztikus lineáris programozási feladatot átfogalmazhatunk véletlen változók jelenléte esetén egy sztochasztikus döntési elv segítségével, vagy pedig a két részben hozott döntés egymástól való elválasztásának segítségével írjuk fel a modellt. Tekintsük a következő lineáris programozási feladatot:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{q}'\mathbf{y} \\ \text{f.h.} \quad & A\mathbf{x} = \mathbf{b}, \\ & T\mathbf{x} + W\mathbf{y} = \mathbf{h}, \\ & \mathbf{x}, \mathbf{y} \geq 0. \end{aligned} \tag{26}$$

Legyenek itt az  $A, T, W$  mátrixok rendre  $m_1 \times n_1, m_2 \times n_1, m_2 \times n_2$  dimenziósok,  $\mathbf{x} \in \mathbf{R}^{n_1}, \mathbf{y} \in \mathbf{R}^{n_2}, \mathbf{b} \in \mathbf{R}^{m_1}, \mathbf{h} \in \mathbf{R}^{m_2}$  dimenziós vektorok, ekkor a feladat jól meghatározott; ezt a továbbiakban mindig feltesszük. Mivel a feltételi egyenlőségek átfogalmazhatók egyenlőtlenségekké, illetőleg az egyenlőtlenségek kiegészítő változók segítségével a fenti alakra hozhatók a szokásos módon, ezért a továbbiakban mi csak az egyenlőségek formájában megfogalmazott feladattal foglalkozunk.

Tegyük fel, hogy a jobboldali  $\mathbf{h}$  vektor véletlen komponensekből áll, akkor a  $\mathbf{h}$  helyett egy  $\boldsymbol{\xi}$  valószínűségi vektorváltozót helyettesítve egy sztochasztikus döntési elv segítségével át kell fogalmaznunk az értelmét vesztett feladatot. Legyen célunk a véletlentől függő  $\mathbf{q}'\mathbf{y}(\mathbf{x}, \boldsymbol{\xi})$  célfüggvényérték várható értékének és valamilyen  $\mathbf{c}'\mathbf{x}$  költség összegének minimalizálása, ekkor a feladat a

$$(27) \quad \begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{E}[\mathbf{q}'\mathbf{y}] \\ \text{f.h.} \quad & A\mathbf{x} = \mathbf{b}, \\ & T\mathbf{x} + W\mathbf{y} = \boldsymbol{\xi}, \\ & \mathbf{x}, \mathbf{y} \geq \mathbf{0} \end{aligned}$$

formát ölti. Ezt a modellt nevezzük a kétlépcsős sztochasztikus programozási modell alapfeladatának.

Kihangsúlyozzuk, hogy bár formailag az  $\mathbf{y}$  változó nem függ az  $\mathbf{x}$  változótól, de  $T\mathbf{x} + W\mathbf{y} = \boldsymbol{\xi}$  feltétel összekapcsolja ezeket. Tehát bár a fenti alapfeladatban a minimalizálás az  $\mathbf{x}$  változóra történik, de a célfüggvény nem dekomponálható az összeadás két tagja szerint!

### 3.1.2. Szekvenciális döntések

A döntések egymásutánosságát kifejező formát a következőképpen vezethetjük le. A (??) feladat második egyenlőségéből  $W\mathbf{y} = \boldsymbol{\xi} - T\mathbf{x}$  kifejezhető, ahol  $\mathbf{y}$  függ az  $\mathbf{x}$ -től és a véletlen  $\boldsymbol{\xi}$ -től egyaránt. A  $q(\mathbf{x}, \boldsymbol{\xi}) = \mathbf{q}'\mathbf{y} = \mathbf{q}'\mathbf{y}(\mathbf{x}, \boldsymbol{\xi})$  függvényt pótló függvénynek, a  $Q(\mathbf{x}) = \mathbf{E}[\mathbf{q}'\mathbf{y}]$  függvényt a várható pótlás függvényének (expected recourse function) nevezzük, ahol:

$$(28) \quad Q(\mathbf{x}) = \mathbf{E}[q(\mathbf{x}, \boldsymbol{\xi})] = \mathbf{E} \left\{ \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \mid T\mathbf{x} + W\mathbf{y} = \boldsymbol{\xi}, \mathbf{y} \geq \mathbf{0} \right\}.$$



Ennek a  $Q(\mathbf{x})$  várható pótlás függvénynek a segítségével az alapmodellünk a következő formába írható:

$$(29) \quad \begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + Q(\mathbf{x}) \\ \text{f.h.} \quad & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq 0. \end{aligned}$$

Ezt a feladatot tekintjük a kétlépcsős modell első-lépcsős feladatának, míg a  $Q(\mathbf{x})$  függvény definíciójában szereplő, adott  $\mathbf{x}$  és  $\boldsymbol{\xi}$  paraméterek esetén adott

$$(30) \quad \begin{aligned} \min_{\mathbf{y}} \quad & \mathbf{q}'\mathbf{y} \\ T\mathbf{x} + W\mathbf{y} \quad & = \boldsymbol{\xi}, \\ \mathbf{y} \quad & \geq 0 \end{aligned}$$

lineáris programozási feladatot a második lépcsős feladatnak (vagy pótló feladatnak) nevezzük; a második lépcső optimális célfüggvény értéket jelöljük a  $q(\mathbf{x}, \boldsymbol{\xi})$  függvénnyel. A teljes kétlépcsős feladatot a sztochasztikus programozás pótló (recourse) feladatának is nevezik. A két lépcsőből álló feladatnak az optimalizálhatóságához szükséges, hogy a (??)-ban szereplő  $Q(\mathbf{x})$  függvény jól meghatározott (és konvex) legyen, illetőleg a (??) feladatnak legyen megengedett megoldása.

### 3.1.3. Többlépcsős modell

Tegyük fel, hogy többször javíthatjuk az előző döntéseink és a bekövetkezett véletlen hatását, tehát döntéseink és a véletlenek megfigyelése a következő sorrendben vannak: az első  $\mathbf{x}$  döntés után megfigyeljük a  $\boldsymbol{\xi}_1$  véletlen realizációját, ekkor egy  $\mathbf{y}_1$  pótlólagos döntést hozunk, a  $\boldsymbol{\xi}_2$  véletlent megfigyeljük,  $\mathbf{y}_2$ -vel korrigálunk, stb. Ha a teljes költség várható értéket akarjuk minimalizálni, akkor az  $r$ -lépcsős modellünk

a következő formát ölti:

$$\begin{aligned}
 (31) \quad & \min \mathbf{c}'\mathbf{x} + \mathbf{E}[\mathbf{q}'_1\mathbf{y}_1 + \mathbf{q}'_2\mathbf{y}_2 + \cdots + \mathbf{q}'_r\mathbf{y}_r] \\
 & \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
 & T_1\mathbf{x} + W_1\mathbf{y}_1 = \boldsymbol{\xi}_1, \\
 & T_2\mathbf{x} + W_2\mathbf{y}_2 = \boldsymbol{\xi}_2, \\
 & \vdots \\
 & T_r\mathbf{x} + W_r\mathbf{y}_r = \boldsymbol{\xi}_r, \\
 & \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_r \geq \mathbf{0}.
 \end{aligned}$$

Ennek a modellalkotási elvnek a mintájára lehet másmilyen többlépcsős feladatokat felírni – például ha az  $r$ -edik lépcsőben az  $\mathbf{y}_r$  döntési változónk nemcsak  $\mathbf{x}$ -től és a véletlenektől függ, hanem az előző  $\mathbf{y}_1, \dots, \mathbf{y}_{r-1}$  döntésektől is.

### 3.1.4. Indukált feltételek

Tekintsük a kétlépcsős feladat első lépcsőjének a következő alakját:

$$\begin{aligned}
 (32) \quad & \min \mathbf{c}'\mathbf{x} + Q(\mathbf{x}) \\
 & \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
 & \mathbf{x} \geq \mathbf{0},
 \end{aligned}$$

ahol a  $Q(\mathbf{x})$  várható pótlás függvényt (??) határozza meg. Ez a probléma általában nem tekinthető jól meghatározottnak, mert ehhez teljesülnie kell annak is, hogy a  $Q(\mathbf{x})$  függvény minden  $\mathbf{x}$  esetén jól értelmezett, vagyis a pótló függvény definíciójában szereplő (??) második lépcsős lineáris programozási feladatnak minden  $\mathbf{x}$  és  $\boldsymbol{\xi}$  esetén létezik megengedett megoldása és véges optimuma. Definíáljuk a  $K$  halmazt a következőképpen:

$$(33) \quad K = \{\mathbf{x} \mid \exists \mathbf{y}, \text{ amelyre } W\mathbf{y} = \boldsymbol{\xi} - T\mathbf{x}, \forall \boldsymbol{\xi}\}.$$

Ezt a  $K$  halmazt, illetőleg az ezt előállító feltételeket nevezzük indukált feltételeknek. Tehát az első lépcsős  $\mathbf{x}$  döntési változónak még egy  $\mathbf{x} \in K$  implicite

megfogalmazott feltételt is ki kell elégítenie. Ezzel az eredeti elsőlépcsős feladat (változatlan  $Q(\mathbf{x})$  függvény mellett) a következő alakú lesz:

$$\begin{aligned}
 (34) \quad & \min \mathbf{c}'\mathbf{x} + Q(\mathbf{x}) \\
 & \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
 & \mathbf{x} \geq 0, \\
 & \mathbf{x} \in K.
 \end{aligned}$$

**Példa.** A következő kisméretű modellen szemléltetjük az indukált feltételeket. Tekintsünk két területet, amelyeket két erőmű lát el árammal. Az  $i$ -edik területen lévő erőmű  $x_i$  villamos energiát fejleszthet, és  $\xi_i$  igényt kell kielégítenie,  $i = 1, 2$ . Van egy távvezeték is a két terület között, amelyen legfeljebb  $x_3$  mennyiségű energiát lehet átvinni, a tényleges átvitelt pedig az első területről a másodikra átvitt  $y$  energiával jelöljük (az  $y = y^+ - y^-$  jelölés bevezetésével  $y$  negatív is lehet, ha a második erőmű által generált áram egy részét az első terület ellátására visszük át, tehát  $y^+ \geq 0, y^- \geq 0$ ). Feladatunk (az első lépcső) a két erőmű  $x_1, x_2$  kapacitásának és az  $x_3$  átviteli kapacitás meghatározásából áll, adott költségek mellett. A véletlen igény jelentkezése után kell döntenünk az  $y^+, y^-$  átvitelről, úgy, hogy az igényeket minimális költséggel elégítsük ki, ami a következőképpen fogalmazható meg:

$$\begin{aligned}
 (35) \quad & \min (q_1 y^+ + q_2 y^-) \\
 & x_1 - y^+ + y^- \geq \xi_1, \\
 & x_2 + y^+ - y^- \geq \xi_2, \\
 & x_3 \geq y^+ \geq 0, x_3 \geq y^- \geq 0.
 \end{aligned}$$

Ez az adott  $x_1, x_2, x_3$  illetőleg  $\xi_1, \xi_2$  esetén a második lépcsős feladat. Természetesen az első és a második területen, illetőleg a teljes területen fellépő igények kielégítéséhez teljesülnie kell a következő egyenlőtlenségeknek:

$$\begin{aligned}
 (36) \quad & \xi_1 - x_1 \leq x_3 \\
 & \xi_2 - x_2 \leq x_3, \\
 & \xi_1 - x_1 + \xi_2 - x_2 \leq 0.
 \end{aligned}$$

Jelöljük a  $\xi_1, \xi_2, \xi_1 + \xi_2$  igények szuprémumát rendre  $h_1, h_2, h_3$ -al, akkor a  $K$  halmaz a következőképpen adható meg:

$$K = \left\{ (x_1, x_2, x_3) \left| \begin{array}{l} x_1 + x_3 \geq h_1 \\ x_2 + x_3 \geq h_2 \\ x_1 + x_2 \geq h_3 \end{array} \right. \right\}.$$

Vegyük észre, hogy ezek a feltételek a gyakorlatban teljesíthetők, de elméletileg kizárják a nem korlátos valószínűségi változók használatát. Véletlen  $\xi_1, \xi_2$  igények esetén a költségek várható értéket szeretnénk minimalizálni. Feltesszük, hogy  $X_1, X_2, X_3$  korlátaink vannak a kapacitásokra, ekkor elsőlépcsős feladatunk

$$(37) \quad \begin{aligned} \min & (c_1 x_1 + c_2 x_2 + c_3 x_3 + E(q^+ y^+ + q^- y^-)) \\ & X_i \geq x_i \geq 0, i = 1, 2, 3, \\ & (x_1, x_2, x_3) \in K, \end{aligned}$$

ahol a célfüggvényben szereplő  $(q^+ y^+ + q^- y^-)$  függvényt (??) határozza meg. Megfelelő kiegészítő változók segítségével az általunk egyenlőtlenséges formában felírt feltételek átírhatók egyenlőséges formába, vagyis az alapfeladat formájára.

### 3.1.5. A pótló függvény tulajdonságai

A  $q(\mathbf{x}, \boldsymbol{\xi}) = \{\min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \mid T\mathbf{x} + W\mathbf{y} = \boldsymbol{\xi}, \mathbf{y} \geq \mathbf{0}\}$  függvény tulajdonságait vizsgáljuk meg. Legyen a  $\boldsymbol{\xi}$  valószínűségi változó lehetséges értékeinek  $\Xi$  halmaza adott.

**23. Tétel.** *Feltesszük, hogy a  $\boldsymbol{\xi} \in \Xi, \mathbf{x} \in K$  feltételek fennállnak, és az első lépcső  $A\mathbf{x} = \mathbf{b}$  feladatának létezik  $\mathbf{x}$  megengedett megoldása. Ekkor a*

$$(38) \quad \begin{aligned} & \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \\ & T\mathbf{x} + W\mathbf{y} = \boldsymbol{\xi}, \\ & \mathbf{y} \geq \mathbf{0} \end{aligned}$$

*második lépcsős feladatnak akkor és csak akkor van véges optimuma, ha  $\exists \mathbf{z}$ , amelyre*

$$(39) \quad W^T \mathbf{z} \leq \mathbf{q}.$$

**Bizonyítás.** A második lépcső (??) feladatának a duálisa a következőképpen írható

$$\begin{aligned} \max_{\mathbf{z}} (\boldsymbol{\xi} - T\mathbf{x})^T \mathbf{z} \\ W^T \mathbf{z} \leq \mathbf{q}, \end{aligned}$$

A primál feladatnak a feltevés miatt van megengedett megoldása. A lineáris programozás dualitás tétele miatt akkor van a primálnak véges optimuma, ha a duál problémának van megengedett megoldása, vagyis ha létezik olyan  $\mathbf{z}$  vektor, melyre  $W^T \mathbf{z} \leq \mathbf{q}$ .

**24. Tétel.** Tegyük fel, hogy létezik az  $W^T \mathbf{z} \leq \mathbf{q}$  egyenlőtlenséget kielégítő  $\mathbf{z}$  vektor, továbbá létezik az  $\mathbf{E}(\boldsymbol{\xi})$  várható érték. Ekkor  $\forall \mathbf{x} \in K$  esetén létezik  $\mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$  is.

**Bizonyítás.** Az előző tétel miatt a primál és a duál problémának is van megengedett megoldása és véges optimuma; ez a véges optimumérték  $q(\mathbf{x}, \boldsymbol{\xi})$ . Felírjuk ezt a primál optimum-értéket a duális célfüggvénnyel. Tegyük fel, hogy a  $Z = \{\mathbf{z} \mid W^T \mathbf{z} \leq \mathbf{q}\}$  duál feladat megengedett megoldásai halmazának létezik legalább egy extrémális pontja. Jelöljük  $\mathbf{z}^1, \mathbf{z}^2, \dots, \mathbf{z}^r$ -el a  $Z$  halmaz extrémális pontjait, ekkor

$$(40) \quad q(\mathbf{x}, \boldsymbol{\xi}) = \max_{i=1, \dots, r} (\boldsymbol{\xi} - T\mathbf{x})' \mathbf{z}^i.$$

Ha  $\mathbf{E}(\boldsymbol{\xi})$  létezik, akkor a jobboldali lineáris függvénynek minden  $i$ -re létezik várható értéke, továbbá létezik a maximum várható értéke is, amely pedig  $\mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$ .

Tegyük fel, hogy nincs egyetlen extrémális pontja sem  $Z$ -nek. Ekkor a  $Z$  halmaz csak egy kúp lehet, és minden  $\mathbf{z} \in Z$  esetén fenn kell állnia a  $(\boldsymbol{\xi} - T\mathbf{x})' \mathbf{z} \leq 0$  egyenlőtlenségnek – ugyanis ellenkező esetben létezne  $\mathbf{z}_0 \in Z$ , amelyre az  $\lambda \mathbf{z}_0 \in Z, \lambda \geq 0$  esetén, továbbá  $(\boldsymbol{\xi} - T\mathbf{x})' \mathbf{z}_0 > 0$  is fennáll, tehát ezen a  $\lambda \mathbf{z}_0$  sugáron a duál feladat célfüggvénye nem lenne korlátos. Az egyenlőtlenség miatt  $\mathbf{z} = \mathbf{0} \in Z$  természetesen igaz, tehát a duál célfüggvény optimális értéke azonosan 0 minden lehetséges  $\boldsymbol{\xi}$  értékre, tehát a  $q$  várható értéke triviálisan létezik.

Ennek a tételnek egy következménye, hogy az adott feltételek mellett a  $Q(\mathbf{x})$  függvény konvex a  $K$  halmazon az  $\mathbf{x}$  változóban. A nem-triviális esetet vizsgáljuk csak, amikor a  $Z$  halmaznak vannak extrémális pontjai. A függvény konvexitásához

$$q(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \boldsymbol{\xi}) \leq \lambda q(\mathbf{x}_1, \boldsymbol{\xi}) + (1 - \lambda) q(\mathbf{x}_2, \boldsymbol{\xi})$$

egyenlőtlenségnek kell fennállnia, minden  $\mathbf{x}_1, \mathbf{x}_2 \in K$  pontpárra. A előző tételben felírt (??) maximum alak kihasználása, illetőleg a linearitás miatt:

$$\begin{aligned} q(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \boldsymbol{\xi}) &= \max_{i=1, \dots, r} (\boldsymbol{\xi} - T[\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2])' \mathbf{z}^i \\ &= \max_{i=1, \dots, r} (\boldsymbol{\xi} - [\lambda T \mathbf{x}_1 + (1 - \lambda) T \mathbf{x}_2])' \mathbf{z}^i \\ &\leq \lambda \left[ \max_{i=1, \dots, r} (\boldsymbol{\xi} - T \mathbf{x}_1)' \mathbf{z}^i \right] + (1 - \lambda) \left[ \max_{i=1, \dots, r} (\boldsymbol{\xi} - T \mathbf{x}_2)' \mathbf{z}^i \right] \\ &= \lambda q(\mathbf{x}_1, \boldsymbol{\xi}) + (1 - \lambda) q(\mathbf{x}_2, \boldsymbol{\xi}), \end{aligned}$$

hiszen lineáris függvények maximuma konvex marad. Hasonlóan belátható, hogy a  $q(\mathbf{x}, \boldsymbol{\xi})$  függvény adott  $\mathbf{x}$  esetén konvex a  $\boldsymbol{\xi}$  változóban is.

**25. Tétel.** *Ha a  $K$  halmaz nem üres, akkor  $K$  egy konvex poliéder.*

**Bizonyítás.** Tegyük fel, hogy  $\boldsymbol{\xi} \in \Xi$  fennáll és  $\mathbf{x}$  megengedett megoldása az  $A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$  feltételeknek. Tekintsük azon  $(\mathbf{x}, \boldsymbol{\xi})$  vektorokat, amelyek esetén létezik  $\mathbf{y}$  megengedett megoldása a második lépcsős feladatnak, vagyis amelyekre  $W\mathbf{y} = \boldsymbol{\xi} - T\mathbf{x}$  fennáll. Ezek a  $(\mathbf{x}, \boldsymbol{\xi})$  vektorok az előző tétel bizonyításában leírtak alapján valamilyen  $D$  és  $H$  mátrixok segítségével felírt homogén lineáris egyenlőtlenségrendszernek tesznek eleget:

$$D\boldsymbol{\xi} \leq H\mathbf{x}.$$

(Ezek a  $D, H$  mátrixok explicite is előállíthatók, lásd a kombinált feladat leírását a 4.4.1-ben.) Tegyük fel, hogy az egyenlőtlenség-rendszerben  $r$  egyenlőtlenség áll, akkor minden komponensnek, vagyis a

$$\sup_{\boldsymbol{\xi} \in \Xi} [D\boldsymbol{\xi}]_i = h_i, i = 1, \dots, r$$

értéknek végesnek kell lennie (egyébként a  $K$  halmaz üres lenne). Ebből következik, hogy a  $K$  felírható a  $K = \{\mathbf{x} \mid \mathbf{h} \leq H\mathbf{x}\}$  alakban, ami a tétel állítása.

Az itt kimondott tételek alapján mondhatjuk, hogy a pótló függvény konvex (szakaszonként lineáris) függvény egy konvex poliéder felett. A szakaszonként lineáris kifejezés azt jelenti, hogy a  $K$  konvex poliéder felbontható olyan  $K_i$  poliéderekre, amelyek egyesítése  $K$  és egy  $K_i$  poliéderen a  $q(\mathbf{x}, \boldsymbol{\xi})$  függvény lineáris mind  $\mathbf{x}$ -ben, mind  $\boldsymbol{\xi}$ -ben. Ha a valószínűségi változónk egy dimenziós, akkor adott  $\mathbf{x}$  esetén a  $q(\mathbf{x}, \boldsymbol{\xi})$  pótló függvény valóban szakaszonként lineáris, konvex függvény.

### 3.1.6. Speciális esetek

Az előző részben részletesen foglalkoztunk az indukált feltételekkel. A továbbiakban élünk azzal a feltevessel, hogy a második lépcsős feladatnak minden  $\mathbf{x} \in \mathbf{R}^n$  esetén van megengedett megoldása, vagyis a  $K$  halmaz az egész tér – az ilyen feladatokra azt mondjuk, hogy **teljes pótlású feladat**. Előfordulhat az is, hogy csak azon  $\mathbf{x}$  értékekre van megengedett megoldása a második lépcsőnek, amely  $\mathbf{x}$  értékek megengedett megoldásai az első lépcsőnek – ezeket a feladatokat **relatív teljes pótlású** feladatoknak nevezzük.

A kétlépcsős feladatnak egy, az általunk megadottnál általánosabb alakját kapjuk, ha megengedjük hogy a  $W$  pótlás-mátrix értékei valószínűségi változók lehetnek. Ennek a feladattípusnak a vizsgálata messzire vezetne, ezért ezt nem tárgyaljuk részletesen (itt előfordulhat, hogy a megengedett  $\mathbf{x}$  megoldások halmaza nem lesz konvex). A továbbiakban mindig feltesszük, hogy az  $W$  mátrix elemei adottak; ezt a feladattípust a **rögzített pótlás** feladatának nevezzük.

Vezessük be a  $posW = \{\mathbf{t} \mid \mathbf{t} = W\mathbf{y}, \mathbf{y} \geq \mathbf{0}\}$  jelölést, ekkor a teljes pótlás esetén  $posW = \mathbf{R}^n$ , a relatív teljes pótlás esetén pedig  $\boldsymbol{\xi} - T\mathbf{x} \in posW$ , ha  $\mathbf{x} \in \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$  fennáll. Összefoglalva: a továbbiakban csak teljes (vagy legalább relatív teljes) és rögzített pótlású feladatokkal fogunk foglalkozni.

Lehet a  $\mathbf{q}$  költségvektor is véletlentől függő – ha a  $\boldsymbol{\xi}$  nem véletlen, akkor a primál-duál átírással az eredeti feladatban megfogalmazott második lépcsős feladatot kapjuk. Teljes általánosságban tekintve a kétlépcsős feladat második lépcsős részében minden paraméter lehet véletlen, vagyis valószínűségi változóink  $(T(\omega), \mathbf{q}(\omega), \boldsymbol{\xi}(\omega))$  alakba írhatók, ahol  $\omega \in \mathcal{A}$  valamilyen  $\sigma$ -algebra eleme.

A kétlépcsős feladatnak azt a gyakorlatban fontos speciális esetet, amikor az  $\mathbf{y}$  pótlás egyszerűen az eltérés értéke, az **egyszerű pótlásnak** nevezzük. Ekkor feltesszük, hogy a célfüggvény együtthatói nemnegatívak,  $\mathbf{q}^+ \geq 0, \mathbf{q}^- \geq 0$  fennáll, a pozitív irányú eltérések kompenzálására szolgálnak az  $\mathbf{y}^+$ , a hiányok pótlására a  $\mathbf{y}^-$  változók. A feladat a  $W = (I, -I)$  helyettesítés után a

$$(41) \quad \begin{aligned} & \min_{\mathbf{y}^+, \mathbf{y}^-} \mathbf{q}^+ \mathbf{y}^+ + \mathbf{q}^- \mathbf{y}^- \\ & \text{f.h. } T\mathbf{x} + \mathbf{y}^+ - \mathbf{y}^- = \boldsymbol{\xi}, \\ & \mathbf{y}^+, \mathbf{y}^- \geq 0 \end{aligned}$$

alakot ölti. Ez a feladat általában egyszerűen megoldható, mivel a második lépcsős, nemnegatív  $\mathbf{y}^+, \mathbf{y}^-$  változók a  $\boldsymbol{\xi} - T\mathbf{x}$  különbségből számíthatók, hiszen  $\mathbf{y}^+ = [\boldsymbol{\xi} - T\mathbf{x}]^+, \mathbf{y}^- = [T\mathbf{x} - \boldsymbol{\xi}]^+$ . Így az optimális célfüggvényértékek függvénye

$$q(\mathbf{x}, \boldsymbol{\xi}) = \mathbf{q}^{+'} \mathbf{y}^+ + \mathbf{q}^{-'} \mathbf{y}^-$$

Vegyük észre, hogy a feladat teljes pótlású (mindig létezik a második lépcső  $\mathbf{y}^+, \mathbf{y}^-$  megoldása), tehát az első lépcsőbe beépíthető a második lépcső. Az egyszerű pótlás feladata ezek szerint a következő alakot ölti:

$$\begin{aligned} \min (\mathbf{c}'\mathbf{x} + Q(\mathbf{x})) &= \min \mathbf{c}'\mathbf{x} + \mathbf{q}^{+'} \mathbf{E} [\boldsymbol{\xi} - T\mathbf{x}]^+ + \mathbf{q}^{-'} \mathbf{E} [T\mathbf{x} - \boldsymbol{\xi}]^+ \\ \text{f.h. } A\mathbf{x} &= \mathbf{b}, \\ \mathbf{x} &\geq 0. \end{aligned}$$

Egyszerű eloszlások esetén a várható érték kiszámítható, vagyis az egyszerű pótlás esete megoldható explicite.

### 3.2. A pótló függvény kiszámítása

Tekintsük a pótlási feladat első lépcsős feladatát: a

$$\begin{aligned} \min \mathbf{c}'\mathbf{x} + Q(\mathbf{x}) \\ \text{f.h. } A\mathbf{x} &= \mathbf{b}, \\ \mathbf{x} &\geq 0, \end{aligned}$$



minimalizálási feladat a  $Q(\mathbf{x})$  függvénytől eltekintve igen egyszerű, tehát számítástechnikai szempontból csak ezzel a függvénnyel kell foglalkoznunk.

Vegyük észre, hogy a  $Q(\mathbf{x}) = \mathbf{E}[q(\mathbf{x}, \boldsymbol{\xi})]$  függvény kiszámítása általában egy  $n$ -dimenziós integrál meghatározása, vagyis ha a  $\boldsymbol{\xi}$  valószínűségi változó eloszlásfüggvénye  $F(\cdot)$ , lehetséges értékeinek tartománya  $\Xi$ , akkor

$$Q(\mathbf{x}) = \int_{\Xi} q(\mathbf{x}, \boldsymbol{\xi}) dF(\boldsymbol{\xi}).$$

A várható pótlás függvényének kiszámítása, mint említettük, nehéz numerikus feladat. A gyakorlatban két általános módszert követhetünk: diszkrétizálás és Monte Carlo becslések. Ha a  $\boldsymbol{\xi}$  valószínűségi vektorváltozó csak véges, vagy megszámlálható sok értéket vesz fel, akkor a pótló függvény várható értéke pontosan kiszámítható. Ezért a folytonos eloszlású valószínűségi változók esetén diszkrétizáljuk a folytonos eloszlást és a kapott diszkrét eloszlásra határozzuk meg a várható pótlás függvény értékét. A második használható számítási eljárásban statisztikai becslést adunk a várható értékre, ami mind diszkrét, mind folytonos eloszlású jobboldal esetén alkalmazható.

### 3.2.1. Diszkrét eloszlás

Tegyük fel, hogy a  $\boldsymbol{\xi}$  valószínűségi vektorváltozó a  $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^n$  értékeket veszi fel rendre  $p_1, p_2, \dots, p_n$  valószínűséggel. Legyen a

$$(42) \quad \begin{aligned} & \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \\ W\mathbf{y} &= \boldsymbol{\xi}^i - T\mathbf{x}, \\ \mathbf{y} &\geq 0 \end{aligned}$$

feladat egy optimális megoldása az  $\mathbf{y}^i$  vektor, amelyen a célfüggvényérték  $\mathbf{q}'\mathbf{y}^i$ . Mivel ezt az értéket  $p_i$  valószínűséggel veszi fel a második lépcsős célfüggvény, ezért az első lépcsős célfüggvényében szereplő várható érték  $\sum_{i=1}^n p_i \mathbf{q}'\mathbf{y}^i$  alakba írható. Tekintettel arra, hogy az (??)-ban felírt  $i$ -edik feladat független az  $j = 1, 2, \dots, n, j \neq i$  feladatoktól, ezért a teljes első és második lépcső együtt felírható, mint

$$\begin{aligned}
(43) \quad & \min \mathbf{c}'\mathbf{x} + \mathbf{q}'[p_1\mathbf{y}^1 + p_2\mathbf{y}^2 + \cdots + p_n\mathbf{y}^n] \\
& \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
& T\mathbf{x} + W\mathbf{y}^1 = \boldsymbol{\xi}^1, \\
& T\mathbf{x} + W\mathbf{y}^2 = \boldsymbol{\xi}^2, \\
& \vdots \\
& T\mathbf{x} + W\mathbf{y}^n = \boldsymbol{\xi}^n, \\
& \mathbf{x}, \mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^n \geq 0.
\end{aligned}$$

Tehát ezt az egy lineáris programozási feladatot kell csak megoldani a kétlépcsős feladat optimumának meghatározására. Hasonlóan egyszerű feladatot kapunk (az első és a második lépcső együtt írható fel), ha nemcsak a  $\boldsymbol{\xi}$  valószínűségi változó, hanem  $T, W, \mathbf{q}$  is véletlen, de diszkrét eloszlásúak. Tegyük fel, hogy a  $(\boldsymbol{\xi}, T, W, \mathbf{q})$  valószínűségi változó  $p_i$  valószínűséggel a  $(\boldsymbol{\xi}^i, T^i, W^i, \mathbf{q}^i)$  értéket veszi fel, ahol  $i = 1, 2, \dots, N$ . A kétlépcsős feladat ekkor a következő:

$$\begin{aligned}
(44) \quad & \min \mathbf{c}'\mathbf{x} + \sum_{i=1}^N p_i \mathbf{q}^{i'} \mathbf{y}^i \\
& \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
& T^1\mathbf{x} + W^1\mathbf{y}^1 = \boldsymbol{\xi}^1, \\
& T^2\mathbf{x} + W^2\mathbf{y}^2 = \boldsymbol{\xi}^2, \\
& \vdots \\
& T^N\mathbf{x} + W^N\mathbf{y}^N = \boldsymbol{\xi}^N, \\
& \mathbf{x}, \mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^N \geq 0.
\end{aligned}$$

### 3.2.2. Diszkrétizálás

Ha a  $\boldsymbol{\xi}$  valószínűségi változó folytonos eloszlású, akkor diszkrétizáljuk az eloszlást, és a közelítő diszkrét eloszlásra írjuk fel a fenti nagyméretű lineáris feladatot (nem korlátos valószínűségi változók esetén még csonkolásra is szükség van). Bár ez elvileg mindig megoldható, de a gyakorlatban 10-nel több komponensű valószínűségi

vektor esetén a diszkrét közelítésben szereplő pontok száma oly nagy lesz, hogy már számítástechnikai nehézséget okoz a keletkező lineáris feladat mérete. Öt osztópontot felvéve minden egyes koordinátatengelyen, egy tízdimenziós folytonos eloszlású valószínűségi vektorváltozó esetén összesen  $n = 5^{10}$  számú  $\xi^i$  diszkrét pontot kapunk, vagyis a lineáris feladat  $m_2 \times 5^{10}$  sorból fog állni, ahol  $m_2$  a  $T\mathbf{x} + W\mathbf{y} = \xi$  feltételeinek száma.

Másképp vegyük észre, hogy a keletkezett lineáris programozási feladat igen egyszerű struktúrájú, amit többszörösen ki lehet használni (dekompozíciós módszerek a szimplex megoldására, speciális bázisok, speciális újraszámolási algoritmusok). Ezekkel külön nem foglalkozunk, de az irodalomban részletesen tárgyalják ezeket a nagyméretű lineáris programozási feladatokat, például [?] tartalmaz ilyen leírásokat.

A folytonos eloszlás diszkrétizálása, a (későbbiekben leírandó) adaptív korlátozás és az  $n$ -dimenziós integrálások esetén megfigyelhető numerikus nehézségek mind a dimenziós robbanásnak nevezett jelenség miatt következnek be.

### 3.2.3. Monte Carlo becslés

Legyen  $\xi$  eloszlásfüggvénye  $F(\cdot)$  és tekintsük a várható pótlás függvényének integrálalakját:

$$Q(\mathbf{x}) = \mathbf{E}(q(\mathbf{x}, \xi)) = \int_{\Xi} q(\mathbf{x}, \xi) dF(\xi).$$

Könnyen adhatunk statisztikai becslést a várható pótlás értékére Monte Carlo módszerrel, hiszen adott  $\mathbf{x}$  paraméter és a  $\xi$  valószínűségi vektorváltozó egy  $\xi_i$  realizációja esetén a  $q(\mathbf{x}, \xi_i)$  függvényérték a

$$(45) \quad \begin{aligned} & \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \\ & W\mathbf{y} = \xi_i - T\mathbf{x}, \\ & \mathbf{y} \geq 0 \end{aligned}$$

lineáris programozási feladat optimum értékének meghatározását kívánja csak. A Monte Carlo integrálás durva módszere szerint egy torzítatlan becslést kaphatunk a  $Q(\mathbf{x}) = \mathbf{E}q(\mathbf{x}, \xi_i)$  értékére a

$$\Theta_1 = \sum_{i=1}^N q(\mathbf{x}, \boldsymbol{\xi}_i)$$

átlaggal, ahol  $\boldsymbol{\xi}_i, i = 1, \dots, N$  a  $\boldsymbol{\xi}$  valószínűségi változó  $N$  darab független realizációja. Ez az eljárás minden olyan  $\boldsymbol{\xi}$  valószínűségi vektorváltozó esetén használható, amikor a rendelkezésünkre áll egy véletlenszám generátor, amely  $\boldsymbol{\xi}$  realizációit előállítja.

Csak egy esetre részletezzük az eljárást – normális eloszlású valószínűségi változók esetén, ez az eloszlás a gyakorlatban fontos szerepet játszik. Az eljárás logikája több szempontból is azonos azzal a gondolatmenettel, amelyet a többdimenziós normális eloszlású valószínűségek kiszámításánál követtünk.

Az algoritmus kialakításánál felhasználjuk a  $q(\mathbf{x}, \boldsymbol{\xi})$  függvény konvexitását, a normális eloszlású vektorok dekomponálhatóságát és a dekomponálás miatt az integrál dekompozícióját, majd végül az ortonormalizált pontok rendszerét.

Tekintsük a  $\Phi(\cdot)$  eloszlásfüggvényű, normális eloszlású  $\boldsymbol{\xi}$  vektor dekompozícióját

$$\boldsymbol{\xi} = \chi_n R \boldsymbol{\eta},$$

ahol  $\chi_n$  egy  $n$ -szabadságfokú,  $\chi$ -eloszlású,  $F_n(k)$  eloszlásfüggvényű valószínűségi változó,  $\boldsymbol{\eta}$  egyenletes eloszlású az  $S_n = \{\mathbf{x} \mid \sum_{i=1}^n x_i^2 = 1\}$  egységgyömbön, eloszlásfüggvénye  $V_n(\mathbf{y})$ . Ekkor

$$Q(\mathbf{x}) = \int_{\mathbf{R}^n} q(\mathbf{x}, \boldsymbol{\xi}) d\Phi(\boldsymbol{\xi}) = \int_{S_n} \int_0^\infty q(\mathbf{x}, k R \mathbf{y}) dF_n(k) dV_n(\mathbf{y}).$$

Legyen  $u$  egyenletes eloszlású a  $[0, 1)$  intervallumban, akkor  $F_n^{-1}(u)$  eloszlásfüggvénye  $F_n(\cdot)$  (ez a standard inverziós technika véletlen számok generálására). Ezzel a felbontásunk

$$(46) \quad Q(\mathbf{x}) = \int_{S_n} \int_0^1 q(\mathbf{x}, F_n^{-1}(u) R \mathbf{y}) du dV_n(\mathbf{y})$$

alakú lesz, amiből viszont következik, hogy ha  $u_j, j = 1, \dots, N_1$  a  $[0, 1]$ -ben egyenletes eloszlású valószínűségi változó realizációi,  $\mathbf{y}_i, i = 1, \dots, N_2$  pedig az  $S_n$  egységgyömbön egyenletes eloszlású valószínűségi vektorváltozó realizációi, akkor a

$$\Theta_2 = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} q(\mathbf{x}, F_n^{-1}(u_j) R \mathbf{y}_i),$$

becslés használható a várható érték közelítésére. Ezek után szimmetrizáljuk az integrál becslését, tehát a becslésünk alakja adott  $\boldsymbol{\xi}$  esetén:

$$\Theta_3 = \frac{1}{N} \sum_{i=1}^N g_3(\mathbf{x}, \boldsymbol{\xi}_i), \quad g_3(\mathbf{x}, \boldsymbol{\xi}) = \frac{1}{2} [q(\mathbf{x}, \boldsymbol{\xi}) + q(\mathbf{x}, -\boldsymbol{\xi})].$$

Ez a módosítás csökkenti a becslés szórását, mert a  $q$  függvény lineáris (vagy majdnem lineáris) egy kis környezetben, tehát esetenként 0 szórású (pontos) becslést is kaphatunk.

A  $q(\mathbf{x}, k R \mathbf{y})$  függvény konvex a  $k$  változóban adott  $\mathbf{x}, \mathbf{y}$  esetén. Mivel  $F_n^{-1}(\cdot)$  monoton növekvő függvény, ezért a  $q(\mathbf{x}, F_n^{-1}(u) R \mathbf{y})$ , mint  $u$  függvénye tartalmazni fog egy monoton növekvő (vagy monoton csökkenő) vagy pedig mind egy monoton növekvő és egy monoton csökkenő részt is (vegyük észre azonban, hogy a  $q$  függvény konvexitását a transzformáció általában nem őrzi meg). Ezen monotonitás miatt érdemes úgynevezett antitetikus (ellentétes) változókat használni, vagyis olyan pontokban venni a függvényértékeket, amelyekben a függvényértékek negatívan korreláltak. Ennek alapján – rögzített  $\mathbf{y} \in S_n$  irányt feltételezve – kapjuk a következő becslést:

$$\Theta_4 = \frac{1}{N} \sum_{j=1}^N g_4(\mathbf{x}, u_j, \mathbf{y}),$$

$$g_4(\mathbf{x}, u, \mathbf{y}) = 0.5 [g_3(\mathbf{x}, F_n^{-1}(0.5u) R \mathbf{y}) + g_3(\mathbf{x}, F_n^{-1}(1 - 0.5u) R \mathbf{y})].$$

Ez a becslés a statisztikában használatos rétegezés eljárásának (stratification) a felezéses változata, amely monoton függvények esetén szóráscsökkenéssel jár.

Az utolsó ajánlott módosítás az algoritmuson a felhasznált irányok ortogonalizálása. Egymástól független  $\boldsymbol{\eta}_i$  irányok helyett egy darab véletlen elhelyezkedésű ortogonális vektor-rendszert használunk a (??) külső integráljának kiszámítására.

Legyen  $\mathbf{U} = \{\mathbf{u}_i\}_{i=1}^n$  egy ortonormált vektorrendszer, amely egyenletes eloszlású az  $S_n$  egységgömbön,  $\mathbf{u}_i \mathbf{u}_j = \delta_{ij}$ , ahol  $\delta_{ii} = 1, \delta_{ij} = 0, i \neq j, i, j = 1, 2, \dots, n$ . Az ortonormalizált becslés a következő:

$$\Theta_5 = \frac{1}{N} \sum_{i=1}^N g_5(\mathbf{x}, k, \mathbf{U}^{(i)}), \quad g_5(\mathbf{x}, k, \mathbf{U}) = \frac{1}{n} \sum_{i=1}^n q(\mathbf{x}, k R \mathbf{u}_i),$$

ahol  $\mathbf{U}^{(i)}$  az  $\mathbf{U} = \{\mathbf{u}_i\}_{i=1}^n$  vektorrendszer véletlen realizációit jelöli. Lényegében itt az történik, hogy az ortonormalizált pontok használata, valamint az integrál dekompozíciója a becslésnek a korrelációból következő szórását semlegesíti.

Egy további módszer arra, hogy „egyenletes” pontokat állítsunk elő az  $S_n$  felületen az lehet, hogy meghatározzuk az összes olyan vektort, amely az  $\{\mathbf{u}_i\}$  halmazrendszerből kiválasztott tetszőleges két vektor normalizált összegét adja. Ilyenkor az  $n$  darab  $\mathbf{u}_i, i = 1, \dots, n$  vektor helyett az  $(\mathbf{u}_i + \mathbf{u}_j)/\sqrt{2}, i = 1, \dots, n-1, j = i+1, \dots, n$  alakú, összesen  $n(n-1)/2$  vektort használjuk. Ez lesz az  $\Theta_6$  becslés, amely megfelel a valószínűségek kiszámításánál alkalmazott  $O_2$  becslésnek. Itt is megtehetjük azt, hogy az  $R$  korrelációs mátrixszal történő szorzás elvégzésének sorrendjét felcseréljük az összeadással, vagyis az  $R(\mathbf{u}_i + \mathbf{u}_j)/\sqrt{2}$  mátrixszorzás  $n(n-1)/2$ -szeri elvégzése helyett csak az  $n$  darab  $R\mathbf{u}_i$  szorzást hajtjuk végre, és utána adjuk össze a transzformált vektorokat.

Mindezeket összefoglalva, a számítógépes tapasztalatok alapján a szimmetrizálás, a felezés és az ortogonalizálás ötletnek a keveréke adja meg a végső becslést. Ezt a gyakorlatban használható  $\Theta_7$  becslést egy  $u \in [0, 1]$ -ben egyenletes eloszlású és egy  $\mathbf{U}$  ortonormált rendszer esetére a következőképpen adjuk meg:

$$\begin{aligned} \Theta_7 &= \frac{1}{N} \sum_{j=1}^N g_7(\mathbf{x}, u_j, \mathbf{U}^{(j)}), \\ (47) \quad g_7(\mathbf{x}, u, \mathbf{U}) &= \frac{1}{n} \frac{1}{4} \sum_{i=1}^n q(\mathbf{x}, F_n^{-1}(0.5u) R \mathbf{u}_i) + q(\mathbf{x}, -F_n^{-1}(0.5u) R \mathbf{u}_i) + \\ &\quad q(\mathbf{x}, F_n^{-1}(1.0 - 0.5u) R \mathbf{u}_i) + q(\mathbf{x}, -F_n^{-1}(1.0 - 0.5u) R \mathbf{u}_i). \end{aligned}$$

Természetesen a gyakorlatban a becslést az első sorban megadott módon határozzuk meg: az  $u$  és az  $\mathbf{U}$  valószínűségi változóknak  $N$  darab független realizációját generáljuk, ezekre számítjuk ki a  $g_7(\cdot)$  függvényértéket és ennek az átlagát vesszük.

### 3.3. Korlátok a pótló függvényre

A konvex függvényekre ismertek egyenlőtlenségek, ezek felhasználásával a pótló függvény értékére korlátokat adhatunk. Az eredeti konvex függvények helyett pedig a korlátokat használva olyan közelítő feladatokat kaphatunk, amelyek az eredeti feladat optimális értékének korlátait szolgáltatják.

#### 3.3.1. Jensen korlát

Egy konvex függvényekre vonatkozó egyenlőtlenséget fogalmazzunk meg, amely segítségével a  $q(\mathbf{x}, \boldsymbol{\xi})$  pótló függvény várható értékére alsó korlátot adhatunk.

**26. Tétel (Jensen egyenlőtlenség).** *Legyen  $\psi(\cdot)$  konvex függvény és  $\xi$  tetszőleges valószínűségi változó, melynek létezik a várható értéke. Ekkor*

$$\psi(\mathbf{E}(\xi)) \leq \mathbf{E}(\psi(\xi)).$$

**Bizonyítás.** A tétel teljes indukcióval bizonyítható, de itt egy szemléletesebb gondolatmenetet alkalmazunk. Két részben látjuk be az állítást, a diszkrét és a folytonos eloszlás esetét külön tárgyaljuk. Tegyük fel, hogy  $\xi$  diszkrét eloszlású valószínűségi változó és az  $\mathbf{P}\{\xi = x_i\} = p_i, i = 1, \dots, n$  egyenlőségekkel van meghatározva. Legyen  $x_1$  az eloszlás első pontja,  $y_1$  tetszőleges. Ekkor a konvexitás definíciója miatt  $x_1$  és  $y_1$  esetén igaz, hogy a  $p_1, 1 - p_1$  súlyokkal

$$\psi(p_1 x_1 + (1 - p_1) y_1) \leq p_1 \psi(x_1) + (1 - p_1) \psi(y_1).$$

Állítsuk most elő  $y_1$ -t a rögzített  $x_2$  és a tetszőleges  $y_2$  keverékeként,  $y_1 = \frac{p_2}{1-p_1} x_2 + (1 - \frac{p_2}{1-p_1}) y_2$  formában ( $x_2$  az eloszlás második pontja). Helyettesítsük be ezt az  $y_1$ -et az előző egyenlőtlenségbe:

$$\begin{aligned} & \psi \left( p_1 x_1 + (1 - p_1) \left[ \frac{p_2}{1 - p_1} x_2 + \left( 1 - \frac{p_2}{1 - p_1} \right) y_2 \right] \right) \leq \\ & \leq p_1 \psi(x_1) + (1 - p_1) \psi \left( \frac{p_2}{1 - p_1} x_2 + \left( 1 - \frac{p_2}{1 - p_1} \right) y_2 \right). \end{aligned}$$

A jobboldal második tagjára alkalmazzuk még egyszer a konvexitás definícióját, ekkor átrendezés után kapjuk a

$$\psi(p_1x_1 + p_2x_2 + (1 - p_1 - p_2)y_2) \leq p_1\psi(x_1) + p_2\psi(x_2) + (1 - p_1 - p_2)\psi(y_2)$$

egyenlőtlenséget. Megismételjük ezt az eljárást minden  $x_i$  pontra, végül (teljes indukcióval) kapjuk a

$$\psi\left(\sum_i p_i x_i\right) \leq \sum_i p_i \psi(x_i)$$

egyenlőtlenséget. Itt a baloldalon  $\psi(\sum_i p_i x_i) = \psi(\mathbf{E}(\xi))$  áll, a jobboldalon pedig  $\sum_i p_i \psi(x_i) = \mathbf{E}(\psi(\xi))$ , tehát a tétel állítása igaz a diszkrét esetre.

Folytonos  $f(x)$  sűrűségfüggvényű  $\xi$  valószínűségi változó esetén tekintsük azt az  $l(x) = mx + b$  egyenest, amely a  $\psi$  görbét az  $x_0 = \mathbf{E}(\xi)$  pontban érinti. Ennek az egyenesnek a meredeksége  $m = \psi'(x_0)$  és ez az egyenes átmegy az  $x_0, \psi(x_0)$  ponton, vagyis  $\psi(x_0) = mx_0 + b$ , amiből  $b$  meghatározható. Tehát az érintő egyenes  $y = l(x)$  egyenlete  $y = \psi'(x_0)x + [\psi(x_0) - \psi'(x_0)x_0]$ .

Felhasználjuk most a differenciálható konvex függvényeknek a jellemzését: tetszőleges  $x$  érték esetén a  $\psi(x)$  függvényérték az érintő egyenese felett van, tehát  $\psi'(x_0)x + [\psi(x_0) - \psi'(x_0)x_0] \leq \psi(x)$ , minden lehetséges  $x$  értékre. Ezek szerint helyettesíthetjük  $x$ -et  $\xi$ -vel;  $\xi$  minden értékére fennáll, hogy

$$\psi'(x_0)\xi + [\psi(x_0) - \psi'(x_0)x_0] \leq \psi(\xi).$$

Várható értéket véve mindkét oldalon (az egyenlőtlenség mindkét oldalát  $f(x)$ -el szorozva és végigintegrálva) a tétel állítását kapjuk, hiszen a baloldalon lévő kifejezés:

$$\mathbf{E}(\psi'(x_0)\xi + [\psi(x_0) - \psi'(x_0)x_0]) = \psi'(x_0)\mathbf{E}(\xi) + [\psi(x_0) - \psi'(x_0)x_0] = \psi(x_0) = \psi(\mathbf{E}(\xi))$$

Bizonyítás nélkül megjegyezzük, hogy a tétel a feltételes várhatóértékre is igaz konvex  $\psi$  függvényre:  $\psi(\mathbf{E}(\xi | \eta)) \leq \mathbf{E}(\psi(\xi) | \eta)$ . Ez azt jelenti, hogy ha a  $\xi$  valószínűségi változó  $I$  támaszát két részre osztjuk, a két részben definiáljuk a két feltételes eloszlást, akkor mindegyikre külön-külön felírható az egyenlőtlenség. (Természetesen a gyakorlati alkalmazáshoz szükség van a feltételes eloszlások meghatározására



és a várható értékének kiszámíthatóságára is.) Szintén bizonyítás nélkül közöljük az alábbi állítást: tetszőleges  $\mathbf{R}^m$ -ben megadott  $\boldsymbol{\xi}$  valószínűségi vektorváltozóra is fennáll a  $\psi(\mathbf{E}(\boldsymbol{\xi})) \leq \mathbf{E}(\psi(\boldsymbol{\xi}))$  egyenlőtlenség, ha a  $\psi(\cdot)$  függvény konvex.

### 3.3.2. Edmundson-Madansky korlátok

Egy felső korlátot konstruálunk most egy konvex függvény várható értékéhez.

**27. Tétel (Edmundson-Madansky egyenlőtlenség).** *Legyen  $\psi(\cdot)$  konvex függvény és  $\xi$  tetszőleges valószínűségi változó, melynek létezik a várható értéke. Ekkor*

$$\mathbf{E}(\psi(\xi)) \leq \frac{b - \mathbf{E}(\xi)}{b - a} \psi(a) + \frac{\mathbf{E}(\xi) - a}{b - a} \psi(b).$$

**Bizonyítás.** Tekintsük az  $a, b$  pontokat a konvex függvény értelmezési tartományában és legyen  $\xi \in [a, b]$ . Ekkor az  $(a, \psi(a))$  és  $(b, \psi(b))$  pontokon át húzott  $U(x)$  egyenes egyenlete

$$\begin{aligned} U(x) &= cx + d, \\ \text{ahol } c &= \frac{\psi(b) - \psi(a)}{b - a}, \quad d = \frac{b}{b - a} \psi(a) - \frac{a}{b - a} \psi(b). \end{aligned}$$

Az így kapott  $U(x)$  függvényre konvex  $\psi(x)$  függvény esetén nyilvánvalóan fennáll a következő egyenlőtlenség:

$$\psi(x) \leq U(x), x \in [a, b].$$

Tekintsük most a  $\xi \in [a, b]$  valószínűségi változót, és számítsuk ki az  $U(\xi)$  várható értéket:

$$\begin{aligned} \mathbf{E}(U(\xi)) &= \frac{\psi(b) - \psi(a)}{b - a} \mathbf{E}(\xi) + \frac{b}{b - a} \psi(a) - \frac{a}{b - a} \psi(b) \\ &= \psi(a) \frac{b - \mathbf{E}(\xi)}{b - a} + \psi(b) \frac{\mathbf{E}(\xi) - a}{b - a}. \end{aligned}$$

Mivel  $\psi(x) \leq U(x)$  fennáll, ezért  $\mathbf{E}(\psi(x)) \leq \mathbf{E}(U(x)) = U(\mathbf{E}(\xi))$ , hiszen  $U(\cdot)$  lineáris függvény, következésképpen  $U(x)$  felső korlát az  $x = \mathbf{E}(\xi)$  helyen is:

$$\mathbf{E}(\psi(x)) \leq U(\mathbf{E}(\xi)) = \frac{b - \mathbf{E}(\xi)}{b - a} \psi(a) + \frac{\mathbf{E}(\xi) - a}{b - a} \psi(b).$$

Az itt szereplő  $p = \frac{\mathbf{E}(\xi) - a}{b - a}$  tekinthető valószínűségnek (nyilván  $\mathbf{E}(\xi) \in [a, b]$ ), tehát szemléletesen a  $\xi$  eloszlását egy  $\eta$  valószínűségi változó kétpontos (az  $a$  és a  $b$  pontokra koncentrált) eloszlásával helyettesítettük, amelyre

$$\mathbf{P}\{\eta = a\} = 1 - p, \quad \mathbf{P}\{\eta = b\} = p.$$

Az Edmundson-Madansky korlát fennáll egy poliéderen értelmezett  $\xi$  valószínűségi vektorváltozóra – a felső korlát explicite meghatározható.

A Jensen és az Edmundson-Madansky korlát tovább javítható, ha újabb osztópontot veszünk fel az  $[a, b]$  intervallumban.

Tekintsük azt az esetet, amikor az  $[a, b]$  intervallumban egy további pontot felvesszünk, legyen ez  $c = \mathbf{E}(\xi)$ . Ekkor a felosztás miatt feltételes eloszlásaink lesznek az  $[a, c]$ ,  $[c, b]$  intervallumokban, amelyekre egyenként írjuk fel a korlátokat. Belátható, hogy ha  $U(x) - L(x)$  volt az eredeti két korlát között a különbség, akkor a finomabb felosztás esetén ez az eltérés csökken (nem növekszik). A tényleges felosztás a kiindulásként vett  $[\mathbf{a}, \mathbf{b}]$  téglatestet  $n$ -dimenziós téglatestek uniójaként adja meg. A felosztás elvégzése  $n$ -dimenziós esetekben (melyik téglatestet bontjuk tovább, melyik koordináta mentén kell újabb osztópontot felvenni, hogyan határozhatók meg az új osztópontok, mely téglatestet nem bontunk tovább) nem egyszerű feladat – ennek részletezése [?]-ben található.

### 3.3.3. A korlátok geometriai jelentése

A Jensen egyenlőtlenség elég szemléletes egy dimenzióban. Az  $n$ -dimenziós esetben az egyenlőtlenség azt mutatja, hogy a várható érték helyen a  $\psi$  függvény érintő-síkja teljes egészében a függvény alatt van. Ha egy konvex poliéderen értelmezett valószínűségi változóra adaptív felosztást alkalmazunk (a poliédernek egy rész-poliéderekből álló felbontását használva), akkor a Jensen egyenlőtlenségek a  $\psi$  függvénynek egy szakaszonként (poliéderenként) lineáris alsó korlátját adják meg.

Az Edmundson-Madansky felső korlát az egy-dimenziós esetben a konvex  $\psi$  függvény által a határokon felvett függvényértékeken át húzott egyenest adjak. Több-dimenziós esetben a konvex poliéder csúcspontjait összekötő egyenesek, valamint a

határpontokat összekötő egyenesek adják meg a felső korlátot; ennek alkotói lineárisak, de összességükben mint egy megcsavart (lineáris) felület jelennek meg. Az adaptív korlátozás esetén rész-poliéderenként ilyen csavart lineáris függvények adják a felső korlátot.

### 3.3.4. Egy numerikus példa

Egy véletlenszerűen generált kétlépcsős sztochasztikus programozási feladat a következő. Az első lépcső feladata

$$\begin{aligned} \min \quad & (9.0x_1 + 8.1x_2 + \mathbf{E}(3.6y_1 + 7.4y_2 + 6.9y_3)) \\ \text{f.h.} \quad & 2.5x_1 + 1.6x_2 \geq 1.8, \\ & 9.4x_1 + 9.0x_2 \geq 8.0, \\ & x_1, x_2, y_1, y_2, y_3 \geq 0. \end{aligned}$$

A második lépcső feladata pedig

$$\begin{aligned} \min \quad & (3.6y_1 + 7.4y_2 + 6.9y_3) \\ & 6.0x_1 - 9.2x_2 - 0.9y_1 - 0.7y_2 + 1.7y_3 = \xi_1, \\ & -6.3x_1 - 1.2x_2 + 3.9y_1 + 9.0y_2 - 13.0y_3 = \xi_2, \\ & y_1, y_2, y_3 \geq 0. \end{aligned}$$

A második lépcsőben szereplő két valószínűségi változóról feltesszük, hogy együttes eloszlásuk normális:  $(\xi_1, \xi_2) \in N((5.8, -8.7), \Sigma)$ , ahol  $\Sigma_{1,1} = \mathbf{D}^2(\xi_1) = 1$ ,  $\Sigma_{2,2} = \mathbf{D}^2(\xi_2) = 1$ ,  $\Sigma_{1,2} = \text{Corr}(\xi_1, \xi_2) = 0.9$ . Vegyük észre, hogy a feladat teljes pótlású: tetszőleges  $x_1, x_2$  és  $\xi_1, \xi_2$  esetén találhatunk olyan  $y_1, y_2, y_3$  értékeket, amelyekre a második lépcső egyenlőségei fennállnak.

Az adaptív korlátozás algoritmusát használva ennek a feladatnak a megoldására két eredményt kapható, az egyik esetén egy 319, a másik esetén egy 522 pontból álló diszkrét eloszlással közelítettük meg a (csonkolt) normális eloszlást. A két közelítésre kapott megoldások és célfüggvényértékek a fenti táblázatban találhatóak. A feladat nehéz megoldhatóságát mutatta, hogy  $\mathbf{D}^2[q(\mathbf{x}_k, \boldsymbol{\xi})] \sim 15^2$  volt az optimális megoldás környékén.

	$\mathbf{x}$	$\mathbf{c}'\mathbf{x} + \mathbf{E}(\mathbf{q}'\mathbf{y})$
$discr_{319}$	(0.9655, 0.0)	26.9805
$discr_{522}$	(0.9621, 0.0)	26.9847

4. táblázat. Kétlépcsős példa, adaptív korlátozás

### 3.4. Megoldó algoritmusok

A kétlépcsős sztochasztikus programozás első lépcsőjének megoldása önmagában véve egy konvex célfüggvény optimalizálását jelenti, lineáris feltételek mellett. A nehézség a  $Q(\mathbf{x})$  várható pótlás kiszámításában (és a pótló függvény meghatározásában) rejlik. Ezért fontos a pótló függvény meghatározása; a kétlépcsős feladat megoldó algoritmusai is a pótló függvény kiszámítási lehetőségeit követik.

#### 3.4.1. Általános megjegyzések

A kétlépcsős sztochasztikus programozási feladat egy optimális vagy közel optimális megoldásának előállítása többféle módon elvégezhető. Az eljárások túlnyomó többségében a valószínűségi eloszlás diszkrétizálása jelenti a kulcsot az algoritmusokhoz – ez még az adaptív korlátozás módszerére is igaz, hiszen impliciten az is egy diszkrétizálási eljárás. A diszkrétizálás után pedig egy nagyméretű lineáris programozási modellt kapunk, amelyet valamilyen speciális lineáris programozási eljárással oldunk meg.

A teljesség igénye nélkül felsorolunk néhány eljárást: Dantzig-Wolfe (vagy Benders) dekompozíció, az L-alakú módszer, korlátozás és adaptív dekompozíció, sztochasztikus dekompozíció, regularizált dekompozíció, sztochasztikus kvázigradienses eljárás. A következő fejezetben ismertetjük a szukcesszív regressziós eljárást, amely szintén alkalmas a kétlépcsős feladat megoldására.

#### 3.4.2. Dantzig-Wolfe

Az általánosan ajánlott eljárás a kétlépcsős sztochasztikus feladat megoldására a következő. Feltesszük, hogy a valószínűségi változó diszkrét eloszlású (vagy pedig egy

közelítő diszkrét eloszlású valószínűségi változóval helyettesítjük a folytonos valószínűségi változót). Ennek esetében a pótló függvény értéke explicite kiszámítható.

Tegyük fel, hogy a  $\xi$  valószínűségi vektorváltozó a  $\xi^1, \xi^2, \dots, \xi^n$  értékeket veszi fel rendre  $p_1, p_2, \dots, p_n$  valószínűséggel. Ekkor, mint azt a diszkrétizálásról szóló részben láttuk, a kétlépcsős probléma a következő (egyetlen) lineáris probléma megoldására redukálódik:

$$\begin{aligned}
 (48) \quad & \min \mathbf{c}'\mathbf{x} + \mathbf{q}'[p_1\mathbf{y}^1 + p_2\mathbf{y}^2 + \dots + p_n\mathbf{y}^n] \\
 & \text{f.h. } A\mathbf{x} = \mathbf{b}, \\
 & T\mathbf{x} + W\mathbf{y}^1 = \xi^1, \\
 & T\mathbf{x} + W\mathbf{y}^2 = \xi^2, \\
 & \vdots \\
 & T\mathbf{x} + W\mathbf{y}^n = \xi^n, \\
 & \mathbf{x}, \mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^n \geq 0.
 \end{aligned}$$

A feladat nagyméretű, de speciális szerkezetű és ritka mátrixú. Ezért vagy erre a feladatra alkalmazzuk a Benders dekompozíciót, vagy tekintjük a feladat duálját, és arra alkalmazzuk a Dantzig-Wolfe dekompozíciós eljárást. A feladat megoldásában szereplő együttható mátrix igen speciális szerkezetű, ezért a bázisban szereplő vektorok is speciálisak; ezt is ki lehet használni azzal, hogy egy bázis-dekompozíciós eljárást használunk a nagyméretű bázis tárolására és felújítására.

### 3.4.3. Egyéb módszerek

Sok, fentebb is említett módszer leírható a nemlineáris optimalizálás metszősík módszerének terminológiájával. Ezek közös tulajdonsága, hogy lényegében az adott  $Q(\mathbf{x})$  függvény esetén újabb feltételeket (metszősíkokat) generálnak, amelyeket hozzáadnak az eredeti, elsőlépcsős feladathoz. Ilyen eljárás a sztochasztikus dekompozíció (Higle és Sen [?]), amelyben véletlenszerűen választanak ki új feltételeket az optimum közelében. Ehhez hasonlít a regularizált dekompozíciós eljárás (Ruszczynski

[?]), amelyben szintén újabb és újabb metszősíkokat állítunk elő, de ennek az az előnye, hogy a metszősíkok (hozzáadott feltételek) száma egy konstans alatt tartható, tehát a feladat mérete nem nőhet minden határon túl.

Az előző részekben leírt Jensen illetőleg Edmundson-Madansky egyenlőtlenségek segítségével a  $Q(\mathbf{x})$  várható pótlás függvényre lineáris alsó és felső korlátokat állíthatunk elő. Ezeket a korlátokat az értelmezési tartományként szolgáló poliéder egyre finomodó, adaptív felosztással meghatározható felbontása esetén egyre szorosabbá tehetjük. Ha most az első lépcső feladatát tekintjük valamilyen  $L_k(\mathbf{x}) \leq Q(\mathbf{x}) \leq U_k(\mathbf{x})$  függvényekkel, akkor a

$$\left. \begin{array}{ll} \min & \mathbf{c}'\mathbf{x} + L_k(\mathbf{x}) \\ \text{f.h. } & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq 0, \end{array} \right\} \text{ és a } \left\{ \begin{array}{ll} \min & \mathbf{c}'\mathbf{x} + U_k(\mathbf{x}) \\ \text{f.h. } & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq 0, \end{array} \right.$$

feladatok optimális megoldásai közrefogják majd az eredeti célfüggvény értékeit. A megengedett megoldások halmazának egy adaptívan végzett dekompozíciója fokozatosan finomodó alsó és felső korlátokból álló probléma-sorozatot állít elő, amelyek egyre kisebb intervallumot adnak a kétlépcsős feladat optimumára, illetőleg közelítő megoldást adnak az optimális megoldásra.

Végül alkalmazhatjuk a kvázigradienses megközelítést is: a célfüggvényben szereplő  $Q(\mathbf{x})$  gradiensének (szubgradiensének) meghatározását egy statisztikai becsléssel helyettesítik és egy nemlineáris optimalizálási eljárást alkalmaznak a feladat megoldására.

Bizonyos esetekben megfigyelhető, hogy a modellben szereplő valószínűségi változók csak igen kisszámú, más valószínűségi változóktól függenek, azok lineáris kombinációjaként állíthatók elő. Ilyenkor a feladat dimenziószáma lényegesen lecsökkenthető, mert a feladat átírható a kisszámú valószínűségi változók függvényére.

Egy másik speciális esetre is utalunk, az egyszerű pótlás feladatára. A 3.1.6-ban már felírtuk azt a feladatot, amelyet szintén elég könnyen meg lehet oldani, mind folytonos, mind diszkrét eloszlás esetén.

## 4. Szukcesszív regressziós approximáció

Ebben a fejezetben egy olyan numerikus algoritmust ismertetünk, amely segítségével nemlineáris optimalizálási feladatokat tudunk megoldani zajos feltételek, vagy célfüggvény esetén. A módszert szukcesszív regressziós approximációnak nevezzük és SRA-val jelöljük. Az eljárás lényege az, hogy (kétszer differenciálható) regressziós függvények egy sorozatával közelítjük a zajos függvényeket és ezek a regressziós közelítések egyre pontosabbá tehetők az optimális megoldások környezetében. Vegyük észre, hogy míg az eddig ismertetett eljárások túlnyomó része egy folytonos függvényt (a pótló függvényt, vagy a valószínűségi változó eloszlását) diszkrétizálja, itt a nem-differenciálható függvényt kétszer differenciálható függvénnyel helyettesítjük.

Az SRA algoritmus konvergens egy egydimenziós egyenlet megoldására, mind determinisztikus (pontos), mind zajos függvényértékek esetén. Az így kapott elgondolásokat általánosítva az  $n$ -dimenziós esetre heurisztikus algoritmusokat adunk a valószínűséggel korlátos és a kétlépcsős feladatra, valamint egy kombinált modell esetén is. Ezeknek az eljárásoknak a konvergenciáját a számítógépes tapasztalatok és példák alapján mutatjuk be. A sztochasztikus lineáris programozási feladatok optimalizálásán túlmenően ez az eljárás alkalmas olyan sztochasztikus programozási feladatok megoldására is, amelyekben kvadratikus feltételek, illetőleg kvadratikus tagot is tartalmazó célfüggvény is van.

### 4.1. Egyenlet megoldása

Az algoritmust egy egydimenziós egyenlet megoldására fogalmazzuk meg először és ezen a példán szemléltetjük lényeges pontjait. Legyen feladatunk az

$$f(x) = 0, \quad x \in \mathbf{R}^1$$

egyenlet megoldása. Ezt egy olyan iteratív eljárással végezzük el, amelyben egy pontokból és függvényértékekből álló halmazból meghatározható regressziós egyenlessel közelítjük az  $f(x)$  függvényt, és a közelítés gyökével kiegészítjük az eredetileg használt ponthalmazt, ebből újabb regressziós egyenest határozunk meg, stb.

#### 4.1.1. Determinisztikus eset

Tegyük fel, hogy tetszőleges  $x$  esetén az  $f(x)$  függvényérték pontosan kiszámítható. A következő feltevést tesszük a függvényre:

**A(f-1)** az  $f(x), x \in \mathbf{R}^1$  egy folytonos függvény és tetszőleges két  $x_i, x_j, x_i \neq x_j$  pontra fennállnak a  $0 < \delta_L \leq \frac{f(x_j) - f(x_i)}{x_j - x_i} \leq \delta_U < \infty$  egyenlőtlenségek, valamilyen  $\delta_L > 0, \delta_U < \infty$  állandókkal.

A feltevésből következik, hogy az  $f(x) = 0$  egyenletnek egy  $\Theta$  gyöke létezik. Az egyszerűség kedvéért feltesszük azt is, hogy

**A(x-2)** az összes  $x_i$  pont egy korlátos  $[L, U]$  intervallumban van.

Tegyük fel, hogy adva van pontok és függvényértékek egy  $S_k = \{x_i, f_i\}_{i=0}^{k-1}$  halmaza, ahol  $f_i = f(x_i)$ . Ezen  $S_k$  halmaz segítségével az  $L^2$  normában legjobban közelítő lineáris  $g_k(x) = \alpha_k x + \beta_k$  függvény együtthatóit a következő feladat megoldásaként kaphatjuk:

$$\min_{\alpha_k, \beta_k} \sum_{i=0}^{k-1} [f_i - (\alpha_k x_i + \beta_k)]^2.$$

Az optimalitás elsőrendű szükséges feltételeit felírva, az ismeretlen  $\alpha_k, \beta_k$  értékeket a következő egyenletrendszerből kaphatjuk

$$\begin{aligned} \sum_{i=0}^{k-1} x_i [f_i - (\alpha_k x_i + \beta_k)] &= 0, \\ \sum_{i=0}^{k-1} [f_i - (\alpha_k x_i + \beta_k)] &= 0, \end{aligned}$$

Vezessük be a következő jelöléseket:

$$\begin{aligned} m_0 &= \frac{1}{k} \sum_{i=0}^{k-1} f_i, & m_1 &= \frac{1}{k} \sum_{i=0}^{k-1} x_i f_i, \\ M_0 &= \frac{1}{k} \sum_{i=0}^{k-1} 1 = 1, & M_1 &= \frac{1}{k} \sum_{i=0}^{k-1} x_i, & M_2 &= \frac{1}{k} \sum_{i=0}^{k-1} x_i^2, \end{aligned}$$

akkor

$$(49) \quad \alpha_k = \frac{-m_0 M_1 + m_1}{M_2 - M_1^2}, \quad \beta_k = \frac{m_0 M_2 - m_1 M_1}{M_2 - M_1^2}.$$



A  $g_k(x) = \alpha_k x + \beta_k = 0$  közelítő egyenlet gyöke lesz az  $x_k = -\frac{\beta_k}{\alpha_k} = -\frac{m_0 M_2 - m_1 M_1}{-m_0 M_1 + m_1}$  érték; ezzel a ponttal és a hozzátartozó  $f_k = f(x_k)$  függvényértékkel bővítjük az  $S_k$  halmazt, az új halmaz felhasználásával megint meghatározunk egy regressziós egyenest, és ismételjük az eljárást.

Feltesszük, hogy a kezdetben rendelkezésünkre áll a tetszőleges  $k$  pontból és a hozzátartozó függvényértékekből álló  $S_k$  halmaz; a gyökmeghatározás részletes algoritmus a következő:

SRA algoritmus (determinisztikus egyenletmegoldás)

0. Előkészítés. Legyen  $k$  az  $S_k = \{x_i, f_i\}_{i=0}^{k-1}$  halmazban adott pontok száma.
1. Számítsuk ki a  $g_k(x) = \alpha_k x + \beta_k$  együtthatóit  $S_k$  segítségével.
2. Legyen a közelítő gyök  $x_k = -\beta_k/\alpha_k$ .
3. Ha  $x_k$  „elég jó”, akkor STOP, egyébként adjuk hozzá  $x_k$ -t és  $f_k$ -t az eddigi halmazhoz:  $S_{k+1} = S_k \cup \{x_k, f_k\}$ , legyen  $k = k+1$  és menjünk vissza az 1. lépésre.

Megjegyezzük, hogy az  $\alpha_k > 0$  és  $M_2 - M_1^2 > 0$  egyenlőtlenségek mindig teljesülnek. Az  $x_k$  közelítő gyök meghatározására két további formulát is megadunk – ezeket egyszerű átalakításokkal kaphatjuk az  $x_k = -\frac{\beta_k}{\alpha_k}$  kifejezésből. Az egyik szerint

$$(50) \quad x_k = M_1 - \frac{m_0}{\alpha_k}.$$

Ez az összefüggés azt mutatja, hogy  $M_1$ , a pontok átlaga (várható értéke) a függvényértékek átlaga előjelével ellentétesen a gyök felé mozdul el ( $m_0 < 0$  esetén például megnöveli  $M_1$ -et, vagyis  $x_k > M_1$  lesz) – de túl is léphet rajta.

Tegyük fel most, hogy az SRA algoritmus segítségével  $S_k$ -ből kiszámítjuk  $x_k$ -t és miután az új ponttal és függvényértékkel bővítettük az  $S_k$  halmazt, az így kapott  $S_{k+1}$  halmazból meghatározzuk az  $x_{k+1}$  pontot. Ekkor a két egymásutáni pontra fennáll a következő összefüggés:

$$x_{k+1} = x_k - f_k \frac{\sum_{i=0}^{k-1} (x_k - x_i)^2}{\sum_{i=0}^{k-1} \sum_{j=i+1}^k (x_j - x_i)(f_j - f_i)}.$$

Ez a rekurzió arra világít rá, hogy az SRA hasonlít a sztochasztikus approximációból ismert Robbins-Monro eljárásra, amely  $x_{k+1} = x_k - c_k f_k$  alakú, de itt az  $f_k$  függvényérték együtthatója nem előre megadott állandó, hanem az eddigi pontoktól és függvényértékektől függő mennyiség.

Belátható az így előállított  $\{x_n\}_{n=k}^{\infty}$  pontsorozat konvergenciája (sőt a korlátossága is igazolható, vagyis az **A(x-2)** feltevés felesleges):

**28. Tétel.** *Tegyük fel, hogy az  $f(x)$  függvényértékek determinisztikusak és fennáll az A(f-1) feltétel. Ekkor az SRA által előállított  $\{x_n\}_{n=k}^{\infty}$  pontsorozat tetszőleges  $S_k$  kezdeti halmaz esetén konvergál a  $\Theta$  gyökhöz.*

Megjegyezzük, hogy ez az eljárás nem monoton, vagyis nem biztosítható, hogy az egymásutáni pontok egyre közelebb kerülnek a gyökhöz. A megszokott Newton-féle gyökkereséstől (és más determinisztikus gyökkereső algoritmusoktól) abban különbözik az SRA algoritmus, hogy az  $x_k$  előállítása közben, a  $k$ -adik iterációban nem csak az  $x_{k-1}$  pontban felvett függvényértéket használja, hanem az összes eddigi pontot és függvényértéket.

#### 4.1.2. Numerikus példa

Egy kis numerikus példát közlünk annak szemléltetésére, hogy az SRA által generált pontsorozat korlátos lesz. Tegyük fel, hogy a  $[-1, +1]$  intervallumban megadott  $f(x)$  függvény a következő értékeket veszi fel:  $f(-1) = -1, f(0) = 0, f(\varepsilon_1) = 2\varepsilon_1, \varepsilon_1 > 0$ , ahol  $\varepsilon_1 > 0$ . A kezdeti pont-függvény halmaz legyen  $S_2 = \{(-1, -1), (\varepsilon_1, 2\varepsilon_1)\}$ . Itt fontos megfigyelni, hogy  $x_0 = -1, x_1 = \varepsilon_1$  egy un. befogó pontpár, vagyis ezek között van a  $\Theta = 0$  gyök. Megmutatjuk, hogy  $x_3$  a  $[-1, +1]$  intervallumban marad.

Ugyanis az SRA az  $S_2$ -ből előállítja az  $x_2 = -\varepsilon_2, \varepsilon_2 > 0$  soronkövetkező pontot; feltesszük, hogy a függvényérték ebben a pontban nagyon közel van -1-hez, vagyis  $f(x_2) = f(-\varepsilon_2) = -1 + \varepsilon_3, \varepsilon_3 > 0$ , ahol  $\varepsilon_3$  kicsi. Ezek után az  $S_3 = S_2 \cup (-\varepsilon_2, -1 + \varepsilon_3)$  halmazból az SRA kiszámítja az  $x_3$  pontot. Most megpróbáljuk a lehető legrosszabb

eredményt előállítani (vagyis olyan értékeket adni az  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  állandóknak, amelyekre  $x_3$  a lehető legtávolabb kerül a  $\Theta$  gyöktől), akkor azt tapasztaljuk, hogy  $x_3 \in [-1, +1]$  mindig fennáll. Ugyanis konvergáljon  $\varepsilon_1, \varepsilon_3$  a 0 értékhez, az  $x_3 \leq 1$  egyenlőtlenség mindig fennáll  $\varepsilon_3 \geq 0$  esetén, – attól függetlenül, hogy milyen „lapos” az  $f(x)$  függvény a -1 és a 0 között.

#### 4.1.3. Zajos függvény

Tegyük most fel, hogy a pontos függvényértékek helyett most csak valamilyen torzítatlan becslést tudunk meghatározni a függvény értékére, ahol az zaj additív, vagyis az  $f_i$  függvényérték helyett csak az  $\tilde{f}_i = f(x_i) + \varepsilon_i$  értékek állnak rendelkezésünkre. Itt és a továbbiakban az eddigi jelölések feletti  $\sim$  jelölést használjuk arra, hogy a megfelelő valószínűségi változót megkülönböztessük a determinisztikus mennyiségtől. A zajra vonatkozóan a következő feltevést tesszük:

**A( $\varepsilon$ -3)** Legyenek az  $\varepsilon_i, i = 0, 1, \dots$  valószínűségi változók azonos eloszlásúak és egymástól teljesen függetlenek, amelyekre  $\mathbf{E}(\varepsilon_i) = 0, \mathbf{D}^2(\varepsilon_i) = \sigma^2, i = 0, 1, \dots, \mathbf{E}(\varepsilon_i \varepsilon_j) = 0, i, j = 0, 1, \dots, i \neq j$  fennáll.

Egy  $\tilde{S}_k = \{x_i, \tilde{f}_i\}_{i=0}^{k-1}$  kezdeti halmazból kiindulva (a determinisztikus esetre leírt egyenletmegoldó eljárást minimálisan módosítva) szintén konvergens eljárást kapunk: az  $\tilde{S}_k$ -ből az SRA eljárással számított  $\{\tilde{x}_n\}_{n=k}^{\infty}$  valószínűségi változók sorozata sztochasztikusan konvergál a  $\Theta$  gyökhöz.

Az eljárás konvergenciájának bizonyítása lényegében a következő tételben van megadva, amely szerint a sztochasztikus függvényértékek esetén alkalmazott SRA algoritmus asszimptotikusan a determinisztikus algoritmussá válik.

**29. Tétel.** *Legyen adva az  $S_k = \{x_i, f_i\}_{i=0}^{k-1}$  és az  $\tilde{S}_k = \{x_i, \tilde{f}_i\}_{i=0}^{k-1}$  halmaz. Az  $S_k$ -ből az SRA segítségével az  $x_k$  pontot, az  $\tilde{S}_k$ -ből az SRA algoritmus az  $\tilde{x}_k = -\tilde{\beta}_k/\tilde{\alpha}_k$  valószínűségi változót állítja elő. Ekkor igaz az, hogy*

$$\mathbf{D}^2(\tilde{x}_k \mid \mathbf{x}_{k-1}) \leq \mathbf{E}((x_k - \tilde{x}_k)^2) \leq \frac{4\sigma^2}{\delta_L^2} \left[ \frac{1}{k} + \frac{(U - L)^2}{k(M_2 - M_1^2)} \right],$$

és ez a korlát 0-hoz tart  $k \rightarrow \infty$  esetén.

Részletes tárgyalás helyett csak utalunk arra, hogy az SRA algoritmus két nagy algoritmuscsaláddal is közeli kapcsolatban áll. Az SRA tekinthető egy természetes, önjavító sztochasztikus approximációs eljárásnak – minél távolabb van  $\tilde{x}_k$  a gyöktől, annál jobban csökken a  $\mathbf{D}^2(\tilde{x}_k \mid \mathbf{x}_{k-1})$  feltételes szórásnégyzetre adott korlát. Az SRA hasonlít az úgynevezett válasz-felületi (RSM – response surface methodology) algoritmusok családjához is, annak egy aszimptotikus, visszacsatolásos formája. Az SRA algoritmus a tágabb értelemben vett adaptív algoritmusok közé tartozik.

## 4.2. Valószínűséggel korlátozott modellek

Az egyre finomodó regressziós közelítések eljárását alkalmazhatjuk a valószínűséggel korlátozott modellekre is. Az SRA algoritmust az egyik legegyszerűbb valószínűséggel korlátozott modell esetén, a STABIL modellre írjuk le; bonyolultabb feladatok megoldására hasonlóan adható meg a megfelelő algoritmus.

### 4.2.1. A STABIL megoldása regressziós közelítéssel

A STABIL modell a következő formában adható meg:

$$\begin{aligned}
 & \min \mathbf{c}'\mathbf{x} \\
 (51) \quad & \text{f.h. } G(\mathbf{x}) = P(\mathbf{t}'_i \mathbf{x} \geq \xi_i, i = 1, \dots, M) \geq p, \\
 & \mathbf{Ax} \geq \mathbf{b}, \\
 & \mathbf{x} \geq \mathbf{0},
 \end{aligned}$$

ahol feltesszük, hogy a  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_M)$  valószínűségi vektorváltozó többdimenziós normális eloszlású (logkonkáv). Az SRA algoritmus valószínűségi korlátos feladatra való alkalmazása azon az elgondoláson alapszik, hogy a  $G(\mathbf{x}) \geq p$  feltételt egy adott  $S_k$  pont-függvény halmazból számított kvadratikus regressziós függvénnyel helyettesítjük, megoldjuk az ezzel felírt közelítő problémát, ennek optimális megoldását és az ehhez tartozó zajos függvényértéket hozzáadjuk az  $S_k$  pont-függvény halmazhoz, amelyből a regressziós függvényt meghatároztuk, aztán egy új regressziós közelítést

határozzuk meg a kibővített  $S_{k+1}$  halmaz segítségével, és az egész iteratív lépést megismétljük.

Legyen  $S_k = \{\mathbf{x}_i, p_i\}_{i=0}^{k-1}$  adott, ahol  $\mathbf{x}_i, i = 0, 1, \dots, k-1$  adott pontok és  $p_i \sim G(\mathbf{x}_i)$  a megfelelő zajos függvényértékek. Az  $\varepsilon_i$  zajra feltesszük, hogy  $p_i = G(\mathbf{x}_i) + \varepsilon_i$  és  $\mathbf{E}(p_i) = G(\mathbf{x}_i), \mathbf{D}^2(p_i) = \sigma^2(\mathbf{x}_i)$  fennáll valamilyen  $\sigma^2(\mathbf{x}_i) = \mathbf{D}^2(\varepsilon_i) = \sigma^2 > 0$  állandóval – a szórás  $\mathbf{x}_i$ -től való függőségét elhanyagoljuk, hiszen az optimum környékén ez általában közel állandó lesz. Feltesszük még, hogy az  $\varepsilon_i$  zajok teljesen függetlenek:  $\mathbf{E}(p_i p_j) = 0, \forall i \neq j$ . Vegyük észre, hogy ez a feltevérendszer fennáll abban az esetben, amikor a  $G(\mathbf{x})$  függvényértéket egy olyan Monte Carlo módszerrel határozzuk meg, amely a függvényértékek torzítatlan becslését állítja elő konstans szórással – Monte Carlo számítások esetén a becslés szórása számítható, és a mintaszám növelésével vagy csökkentésével állandóvá tehető.

Adott  $S_k = \{\mathbf{x}_i, p_i\}_{i=0}^{k-1}$  pontok és függvényértékek halmaza esetén keressük azt a  $q_k(\mathbf{x}) = \mathbf{x}' D_k \mathbf{x} + \mathbf{b}_k' \mathbf{x} + c_k$  alakú regressziós függvényt, amelyre a függvények eltéréseinek  $L_2$  normája minimális, vagyis a függvény  $D_k, \mathbf{b}_k, c_k$  paraméterei a következő minimalizálási probléma optimális megoldását adják:

$$(52) \quad \min_{D_k, \mathbf{b}_k, c_k} \sum_{i=0}^{k-1} [p_i - q_k(\mathbf{x}_i)]^2 .$$

Az elsőrendű szükséges optimalitási feltételekből a közelítő kvadratikus forma  $D_k, \mathbf{b}_k, c_k$  paraméterei meghatározhatók. A  $G(\mathbf{x}) \geq p$  feltétel helyett a  $q_k(\mathbf{x}) \geq p$  közelítő feltételt használva, a STABIL közelítő feladatának a formája a következő lesz:

$$(53) \quad \begin{aligned} \min \mathbf{c}' \mathbf{x} \\ \text{f.h. } q_k(\mathbf{x}) &\geq p, \\ \mathbf{A} \mathbf{x} &\geq \mathbf{b}, \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned}$$

Ezen közelítő feladat megengedett megoldásainak halmaza akkor lesz konvex, ha a  $q_k(\mathbf{x})$  függvény konkáv – ehhez a közelítő kvadratikus alak  $-D_k$  paraméterének

pozitív definitisége szükséges. Általában nem biztosítható a pozitív definitiség (kivéve azt az esetet, amikor az (??) feladatot szemidefinit programozási módszerrel oldjuk meg). Mégis, ha a (??) feladatot a  $\mathbf{L} \leq \mathbf{x} \leq \mathbf{U}$  pótlólagos feltétellel kiegészítjük, (valamilyen elég kis  $\mathbf{L}$  illetőleg elég nagy  $\mathbf{U}$  állandókból álló vektorokkal), akkor néhány (gyakorlatilag elhanyagolható) kivételtől eltekintve a pozitív definitiség (konvexitás) teljesül. A számítási munkák csökkentése céljából a  $D_k$  mátrixról feltesszük, hogy szimmetrikus, ez a követelmény minden további nélkül beépíthető a (??) feladat megoldó algoritmusába.

#### 4.2.2. Algoritmus a STABIL feladatára

Az egyenletmegoldási algoritmushoz hasonló módon adható meg az SRA leírása a STABIL modellre:

SRA algoritmus (valószínűséggel korlátozott feladat)

0. Előkészítés. Tegyük fel, hogy  $S_k = \{\mathbf{x}_i, p_i\}_{i=0}^{k-1}$  adott és legyen a pontok száma  $k$ .
1. Számítsuk ki a  $q_k(\mathbf{x})$  függvény  $D_k, \mathbf{b}_k, c_k$  együtthatóit az  $S_k$  halmaz segítségével.
2. Legyen feladatunk

$$\begin{aligned}
 & \min \mathbf{c}'\mathbf{x} \\
 (54) \quad & \text{f.h. } q_k(\mathbf{x}) = \mathbf{x}'D_k\mathbf{x} + \mathbf{b}_k'\mathbf{x} + c_k \geq p, \\
 & \mathbf{Ax} \geq \mathbf{b}, \\
 & \mathbf{L} \leq \mathbf{x} \leq \mathbf{U}, \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned}$$

és jelöljük ennek a feladatnak egy optimális megoldását  $\mathbf{x}_k$ -val.

3. Ha  $\mathbf{x}_k$  „elég jó”, akkor STOP, egyébként határozzuk meg a  $p_k \sim G(\mathbf{x}_k)$  zajos függvényértéket és bővítsük az  $S_k$  halmazt: legyen  $S_{k+1} = S_k \cup \{\mathbf{x}_k, p_k\}$ , növeljük meg az iteráció számlálóját  $k := k+1$  és menjünk vissza az 1.

lépésre.

Néhány, számítástechnikai szempontból fontos megjegyzést teszünk. Az „elég jó” kifejezés itt és a továbbiakban is többféleképpen értelmezhető. Az egyik lehetőség az, hogy a  $\mathbf{x}_k$  pontban kiszámítjuk a  $G(\mathbf{x})$  függvény érintősíkját a zajos gradiens segítségével (illetőleg az aktív feltételek érintősíkját), és ha  $\mathbf{c}$  merőleges (egy adott hibahatáron belül) erre a síkra, akkor megállhatunk. Másmilyen, statisztikai mintavételi eljárásokkal is „elég jó”-nak minősíthető egy adott  $\mathbf{x}_k$  pont, ezek leírásától eltekintünk.

Az  $\mathbf{x}_i, p_i$  értékek különböző átlagai –  $m_0, m_1, M_1, M_2$ -nek megfelelő és hasonló mennyiségek – közvetlenül is kiszámíthatók, de az (??) minimalizálási feladat megoldásában szereplő egyenletrendszer együtthatómátrixa (illetőleg inverze) a Morrison-Sherman formula alapján is újraszámítható, viszonylag csekély munkával.

Végül megjegyezzük még, hogy a felírt közelítő (??) feladat (illetőleg a (??) feladat) tetszőleges olyan numerikus optimalizálási eljárással megoldható, amely alkalmas arra, hogy kvadratikus célfüggvény és/vagy feltételek esetén optimális megoldást adjon (ilyen például a közismert MINOS nemlineáris optimalizálási programcsomag).

#### 4.2.3. Egy valószínűséggel korlátozott numerikus példa

Legyenek a  $\beta_1, \beta_2$  normális eloszlású valószínűségi változó, amelyek várható értéke 0, szórása 1, korrelációs együtthatójuk  $\rho$ . Ha a valószínűségi változók nem függetlenek ( $\rho \neq 0$ ), akkor egy együttes valószínűségi korlát mellett optimalizálunk. Tekintsük a következő numerikus példát:

$$\begin{aligned}
& \min(x_1 + x_2), \\
& \text{f.h. } P \left\{ \begin{array}{l} 3x_1 + x_2 - 6 \geq \beta_1 \\ x_1 + 8x_2 - 8 \geq \beta_2 \end{array} \right\} \geq p, \\
& x_1 + 4x_2 \geq 4, \\
& 3x_1 + x_2 \geq 3, \\
& x_1 \geq 0, \quad x_2 \geq 0.
\end{aligned}$$

A feladat optimális megoldása az  $(1.98, 0.96)$  pont környékén van, ebben a környezetben a valószínűségi korlát aktív (ha nem lenne aktív, akkor egy egyszerű lineáris programozási feladatunk lenne csak).

A feladatra vonatkozó számítógépes eredményeket közöljük a mellékelt táblázatokban. Az 5. Táblázatban az első sor a Szántai által kifejlesztett sztochasztikus programozási programcsomag felhasználásával kapott optimális megoldás. A második sorban vannak azok az eredmények, amelyeket az SRA algoritmussal úgy kaptunk, hogy a függvényértékek meghatározását determinisztikusan végeztük (egy kétdimenziós normális eloszlásfüggvény értékeit kiszámító duplapontosságú szubrutint használtunk).

Az első két sor által adott  $f(\mathbf{x}_{STO})$  optimális célfüggvény értékek közötti  $\Delta = 2.9140 - 2.8992 = 0.015$  különbség azt mutatja, hogy az SRA algoritmusnak, mint nemlineáris optimalizálási eljárásnak mekkora a hibája  $k = 1000$  lépés esetén. A táblázat második része a  $k = 20, 50, 200$  és  $1000$  pont felhasználásával az SRA algoritmus által kapott eredményeket tartalmazza zajos függvényértékek esetén (a zaj nagyságára most  $\sigma_1 = D(\varepsilon_i) = 0.01$  volt igaz). Az  $\mathbf{x}_{STO}$  oszlopban a végeredmény, vagy a  $k$ -adik iteráció eredménye van,  $f(\mathbf{x}_{STO})$  pedig a célfüggvény értéke ezen a helyen.

Egy érdekes sejtésre mutatnak rá az adatok. Tegyük fel, hogy az eljárás folyamán szükséges zajos függvényértékeket, a  $p_i \sim G(\mathbf{x}_i)$  értékeket egy Monte Carlo eljárással határozzuk meg, a becslésnek  $\sigma_1 = \mathbf{D}(p_i), i = 1, 2, \dots$  szórása van. Ekkor a  $k$  iterációs lépés után kapott  $\mathbf{x}_k$  pontban a  $G(\mathbf{x}_k)$  pontos függvény érték és a  $p_k$



$k =$	módszer	$\mathbf{x}_{STO}$	$f(\mathbf{x}_{STO})$	$\sigma_P$	$\mathbf{P}\{\mathbf{x}_{STO}\}$
1000	Szántai	(1.9977, 0.9015)	2.8992	–	0.8000
	$SRA_{det}$	(1.9686, 0.9455)	2.9140	–	0.8003
20	SRA	(1.9831, 0.9678)	2.9509	0.0030	0.8305
50	SRA	(1.9663, 0.9732)	2.9392	0.0014	0.8019
200	SRA	(1.9543, 0.9809)	2.9353	0.0010	0.7923
1000	SRA	(1.9555, 0.9795)	2.9351	0.0004	0.8025

5. táblázat. Valószínűségi korlátos feladat: Szántai és az SRA eredményei,  $p = 0.8$ ,  $\varrho = 0.9$  esetén,  $\mathbf{D}^2(G(\mathbf{x}_i)) = \sigma_1^2 = (0.01)^2$  szórásnégyzetű függvényértékek esetén.

becsült érték közötti eltérés (vagyis a tényleges hiba) körülbelül  $3\sigma_1/\sqrt{k}$  nagyságú volt az esetek többségében. Az utolsó sorban például az aktuális optimum helyén már három tizedesjegyre pontos ( $3\sigma_P = 0.0012$ ), pedig az egyes függvényértékek kiszámításának hibája 0.03 körüli.

Más szóval a numerikus eredmények alapján azt tapasztaltuk, hogy  $k$  lépés után majdnem mindig fennáll az  $|G(\mathbf{x}_k) - p_k| \leq 3\sigma_P$  egyenlőtlenség a  $\sigma_P = \sigma_1/\sqrt{k}$  jelöléssel. A táblázatban a pontos  $G(\mathbf{x}_k)$  függvényértéket  $\mathbf{P}\{\mathbf{x}_{STO}\}$  kifejezéssel jelöltük – ezeket az értékeket egy determinisztikus eljárással határoztuk meg. Ezt a sejtést az optimális szóráscsökkenés sejtésének nevezzük – ugyanis ez a legjobb, amit statisztikai mintavételezéssel elérhetünk (ha minden pont ugyanott lenne, az  $\mathbf{x}_k$  pontban, pedig nincs ott). Mint majd látni fogjuk, ez a sejtés fennáll az SRA más alkalmazásainál is, nagyobb iterációszámnál jobban látható ennek a sejtésnek a teljesülése.

A 6. Táblázatban további számítási eredményeket adunk meg : itt a feladat  $p$  és  $\varrho$  paramétereit változtattuk meg. Az első három sorban kapott eredmények szépen demonstrálják azt a valószínűségszámítási tételt, hogy növekvő korreláció esetén egy adott pontban csökken a valószínűség.

$p =$	$\varrho =$	$\mathbf{x}_{STO}$	$f(\mathbf{x}_{STO})$	$\mathbf{P}\{\mathbf{x}_{STO}\}$
0.80	-0.9	(2.0087, 0.9708)	2.9794	0.8003
0.80	0.0	(2.0104, 0.9629)	2.9733	0.8080
0.80	0.9	(2.0820, 0.8581)	2.9402	0.7940
0.95	0.9	(2.3405, 0.9232)	3.2638	0.9510

6. táblázat. Valószínűségi korlátos feladat különböző  $p$ ,  $\varrho$  értékekre,  $k = 50$ ,  $\sigma_1 = 0.01$ ,  $\sigma_P = \sigma_1/\sqrt{k} = 0.0015$ .

### 4.3. Kétlépcsős modellek

Tekintsük a szokásos módon felírt kétlépcsős sztochasztikus programozási feladat első lépcsős részét:

$$\begin{aligned}
 (55) \quad & \min \mathbf{c}'\mathbf{x} + Q(\mathbf{x}) \\
 & \text{f.h. } A\mathbf{x} \leq \mathbf{b}, \\
 & \mathbf{x} \geq \mathbf{0}.
 \end{aligned}$$

Tekintettel arra, hogy a  $Q(\mathbf{x})$  függvény konvex, továbbá a függvény egyes értékeinek kiszámítása Monte Carlo becslések használata esetén zajos, ezért ez a függvény is közelíthető egy konvex kvadratikus regressziós függvénnyel. Bár a  $Q(\mathbf{x})$  várható pótlás költségfüggvénye szakaszonként lineáris, a közelítő regressziós függvény pedig kétszer folytonosan deriválható, a közelítés hibája a gyakorlati számítási eredmények alapján elhanyagolható. Minél nagyobb lesz a kétlépcsős feladata mérete, annál jobb a közelítés (minél több feltétel szerepel a  $q(\mathbf{x}, \boldsymbol{\xi})$  függvény meghatározásában, az optimum közelében, annál „simább” lesz a függvény). Az itt követett eljárás éppen ellentéte a szokásos eljárásnak, amikor folytonos eloszlásokat diszkrét eloszlásokkal helyettesítenek (lásd az előző fejezetet).

#### 4.3.1. A kétlépcsős feladat regressziós közelítéssel

A várható pótlás függvényét a  $q(\mathbf{x}, \boldsymbol{\xi})$  pótló függvényből számíthatjuk ki egy Monte Carlo eljárással. Amint azt a kétlépcsős problémák bevezetésénél láttuk, a

$$(56) \quad Q(\mathbf{x}) = \mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi})) = \mathbf{E}(\min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \mid W\mathbf{y} = \boldsymbol{\xi} - T\mathbf{x}, \mathbf{y} \geq 0).$$

függvénynek egy torzítatlan becslését könnyen megkaphatjuk, a megfelelő második lépcsős feladatok megoldásával. Tehát a továbbiakban feltesszük, hogy adott  $\mathbf{x}_i$  esetén kiszámíthatók a  $q_i$  zajos függvényértékek, amelyekre  $\mathbf{E}q_i = Q(\mathbf{x}_i)$ , vagyis a zajos és a determinisztikus függvényértékek közti szokásos jelöléssel  $q_i \sim Q(\mathbf{x}_i)$ . Itt a zajos függvényértékek szórásnégyzete  $\sigma^2(\mathbf{x}_i) = \mathbf{D}^2(q(\mathbf{x}_i, \boldsymbol{\xi}))$  függ az  $\mathbf{x}$  paramétertől, de – az SRA algoritmus valószínűségi feltételes modellekre való alkalmazásánál mondottakhoz hasonlóan – feltesszük, hogy  $\sigma^2(\mathbf{x}) = \sigma^2 > 0$  fennáll valamilyen pozitív  $\sigma$  állandóval.

A  $q(\mathbf{x}, \boldsymbol{\xi})$  függvény konvexitása alapján a közelítésre használandó regressziós függvényről feltesszük, hogy  $q_k(\mathbf{x}) = \mathbf{x}'D_k\mathbf{x} + \mathbf{b}'_k\mathbf{x} + c_k$  alakú, ahol  $D_k$  pozitív definit, szimmetrikus mátrix. A feltételezés jogossága ugyanazon érveken alapul, mint a valószínűséggel korlátos modellek esetében. A szimmetrikusság kívánalma könnyen beépíthető az algoritmusba, a pozitív definitisége nem.

#### 4.3.2. Algoritmus a kétlépcsős feladatra

Az SRA alkalmazása formálisan a következő algoritmusban fogalmazható meg. Tegyük fel, hogy rendelkezésünkre áll egy  $S_k = \{\mathbf{x}_i, q_i\}_{i=0}^{k-1}$  pontokból és zajos függvényértékekből álló halmaz, ahol  $q_i \sim Q(\mathbf{x}_i)$ .

SRA algoritmus (kétlépcsős feladat)

0. Előkészítés. Legyen az iterációs számláló értéke  $k$ , az  $S_k$ -ban adott pontok száma.
1. Számítsuk ki a  $q_k(\cdot)$  regressziós függvény  $D_k, \mathbf{b}_k, c_k$  együtthatóit az  $S_k$ -ből.

2. Oldjuk meg a

$$(57) \quad \begin{aligned} \min_{\mathbf{x}} (\mathbf{c}'\mathbf{x} + q_k(\mathbf{x})), \\ A\mathbf{x} \leq \mathbf{b}, \\ \mathbf{x} \geq 0 \end{aligned}$$

közelítő feladatot és jelöljük ennek egy optimális megoldását  $\mathbf{x}_k$ -val.

3. Ha  $\mathbf{x}_k$  „elég jó”, akkor STOP, egyébként számítsuk ki a

$q_k = \frac{1}{N_k} \sum_{i=1}^{N_k} q(\mathbf{x}_k, \boldsymbol{\xi}_{k_i})$  zajos függvényértéket valamilyen  $N_k$  mintaszámmal (ahol a  $\boldsymbol{\xi}_{k_i}$  értékek a  $\boldsymbol{\xi}$  vektor független realizációi). Bővítsük a ponthalmazt ezzel a ponttal, vagyis legyen  $S_{k+1} = S_k \cup \{\mathbf{x}_k, q_k\}$ , növeljük meg az iterációs számlálót  $k := k+1$  és menjünk vissza az 1. lépésre.

A teljes algoritmus numerikus megvalósítása folyamán egy korlátos tartományt is meg kell adni,  $\mathbf{L} \leq \mathbf{x}_i \leq \mathbf{U}$  alakban, ahol az  $\mathbf{x}_k$  pontok mozoghatnak; ezt feltételként hozzátesszük a (??) feladat feltételeihez. A megállási szabály, a megfelelő mátrix inverzének újraszámítási módszere tekintetében, és a legkisebb négyzetekből adódó feladat megoldása folyamán ugyanazokat az elveket követhetjük, mint amelyeket a valószínűségi korlátos feladat megoldása folyamán már megvitattunk.

#### 4.3.3. Numerikus példa

A teljesség kedvéért mégegyszer megadjuk azt a numerikus példát, amely már szerepelt a kétlépcsős modell tárgyalása esetében. Az első lépcső feladata

$$\begin{aligned} \min \quad & (9.0x_1 + 8.1x_2 + \mathbf{E}(3.6y_1 + 7.4y_2 + 6.9y_3)) \\ \text{f.h.} \quad & 2.5x_1 + 1.6x_2 \geq 1.8 \\ & 9.4x_1 + 9.0x_2 \geq 8.0 \\ & x_1, x_2, y_1, y_2, y_3 \geq 0 \end{aligned}$$

Legyen  $(\xi_1, \xi_2)$  normális eloszlású valószínűségi vektorváltozó, amelyre  $\mathbf{E}(\xi_1) = 5.8$ ,  $\mathbf{E}(\xi_2) = -8.7$ ,  $\mathbf{D}^2(\xi_1) = 1$ ,  $\mathbf{D}^2(\xi_2) = 1$ , a két valószínűségi változó közötti korreláció  $\text{Corr}(\xi_1, \xi_2) =$

	$\mathbf{x}$ ill. $\mathbf{x}_k$	$\mathbf{c}'\mathbf{x} + \mathbf{E}(\mathbf{q}'\mathbf{y})$	"pontos"	$\bar{\mathbf{x}}$
$discr_{319}$	(0.9655, 0.0)	26.9805		
$discr_{522}$	(0.9621, 0.0)	26.9847		
$k = 1$	(1.1479, 0.0)	34.79( $\pm 1.41$ )	32.53( $\pm 0.14$ )	(1.056,-0.11)
$k = 10$	(1.0867, 0.0)	29.81( $\pm 1.41$ )	29.67( $\pm 0.14$ )	(1.095,-0.05)
$k = 100$	(1.0439, 0.0)	29.40( $\pm 1.41$ )	28.26( $\pm 0.14$ )	(1.063,-0.01)
$k = 1000$	(1.0188, 0.0)	29.95( $\pm 1.41$ )	27.68( $\pm 0.14$ )	(1.029, 0.00)
$k = 4000$	(1.0042, 0.0)	29.44( $\pm 1.41$ )	27.43( $\pm 0.14$ )	(1.015, 0.00)
$k = 10000$	(0.9842, 0.0)	26.93( $\pm 1.41$ )	27.19( $\pm 0.14$ )	(1.003, 0.00)
$k = 40000$	(0.9615, 0.0)	27.59( $\pm 1.41$ )	27.09( $\pm 0.14$ )	(0.975, 0.00)

7. táblázat. Kétlépcsős feladat: Az SRA algoritmus egy futásának eredményei a  $k$ -adik iterációban: itt a mintaszám  $N_k = 100$ ,  $\mathbf{D}^2(q_i) = \sigma_1^2 = (1.41)^2$ .

0.9. A második lépcső feladata:

$$\begin{aligned}
& \min_{y_1, y_2, y_3} \quad (3.6y_1 \quad +7.4y_2 \quad +6.9y_3) \\
\text{f.h.} \quad & 6.0x_1 \quad -9.2x_2 \quad -0.9y_1 \quad -0.7y_2 \quad +1.7y_3 = \xi_1 \\
& -6.3x_1 \quad -1.2x_2 \quad +3.9y_1 \quad +9.0y_2 \quad -13.0y_3 = \xi_2 \\
& y_1, y_2, y_3 \geq 0
\end{aligned}$$

A feladat numerikus megoldására vonatkozó eredményeket a fenti 7. táblázatban adjuk meg. Az első két sorban a Mayer által adott, adaptív korlátozás segítségével kapott eredmények találhatók. A  $\mathbf{x}$  oszlopban adtuk meg a  $k$ -adik iterációban kapott  $\mathbf{x}_k$  közelítő megoldást, a  $\mathbf{c}'\mathbf{x} + \mathbf{E}(\mathbf{q}'\mathbf{y})$  oszlop mutatja az adott pontban a  $q_k$  zajos függvényértékeket – ezek szórása az  $N_k = 100$  mintaszám esetén  $\sigma_1 = 1.41$  volt, amit  $(\pm 1.41)$  formában írtunk a táblázatba és ezeket a  $q_k$  értékeket használtuk az  $S_k$  halmazban.

A  $k$ -adik iterációban kapott  $\mathbf{x}_k$  pont esetén, megnövelt  $N = 100\,000$  mintaszámmal kiszámítottuk még egyszer a  $Q(\mathbf{x})$  függvény értéket, az így kapott függvényérté-

kek a „pontos” oszlopban vannak. A legutolsó iterációban kapott (0.9615, 0.0) pontban egy még ennél is nagyobb mintával újraszámított függvényérték  $26.92(\pm 0.02)$  volt – itt a zárójeles érték a kapott eredmény szórása. Az adaptív korlátozás módszerével kétféle diszkretizálás esetén kapott optimális függvényértékek 26.9805 és 26.9847 voltak, az általunk kapott függvényérték pontosabbnak (és jobbnak) tekinthető. Az utolsó,  $\bar{\mathbf{x}}$  fejlécű oszlopban található a súlypontok, vagyis az  $\frac{1}{k} \sum_{i=0}^{k-1} \mathbf{x}_i$  átlagok, amelyek lassabb, de határozott konvergenciát mutatnak.

Jegyezzük meg, hogy a valószínűségi korlátos modellnél leírt, a hibára vonatkozó optimális szórásnövekedés sejtése a numerikus adatok szerint itt is igaznak látszik: a  $k$ -adik iterációban kapott  $\mathbf{x}_k$  pontban a  $Q(\mathbf{x}_k)$  pontos függvényérték minimumtól való eltérése kisebb, mint  $3\sigma_Q = 3\sigma_1/\sqrt{k}$  a legtöbb esetben. Például a  $26.92(\pm 0.02)$  értéket elfogadva a  $Q(\mathbf{x})$  pontos minimum értéknek, látható, hogy a  $k = 10000$  lépés után

$$\begin{aligned} |27.19 - 26.92| &= 0.27 \leq 3\sigma_1/\sqrt{k} \pm 3 \times 0.14 \\ &= 3 \times 1.41 \times \frac{1}{100} \pm 3 \times 0.14 = 0.04 \pm 3 \times 0.14 \leq 0.46. \end{aligned}$$

vagyis  $0.27 \leq 0.46$  fennáll. Itt a  $\pm 3 \times 0.14$  mennyiséget azért kellett hozzáadni a hibához, mert ennyi volt a 27.19 függvényérték kiszámításának a hibája (0.95 megbízhatósági szinten) a „pontos” függvényérték oszlopában. (Ha ki tudnánk számítani az  $\mathbf{x}_k$  pontban a célfüggvény  $f(\mathbf{x}_k) = \mathbf{c}'\mathbf{x} + \mathbf{E}(\mathbf{q}'\mathbf{y})_{\mathbf{x}=\mathbf{x}_k}$  értékét pontosan, akkor a sejtés szerint  $|f(\mathbf{x}_k) - 26.92| \leq 0.04$  lenne.)

A feladat numerikus megoldásának analizálása során a következőt vehetjük észre: a  $Q(\mathbf{x})$  függvényt közelítő konvex kvadratikus alak, amelynek a (nemnegatívítási feltételek nélküli) minimuma a (0.6, −0.3) pont körül van, a szintvonalat jellemző kvadratikus alak pedig egy ellipszoid, amelynek féltengelyei mintegy 30, illetőleg 120 fokos szöget zárnak be az  $x_1$  pozitív féltengelyével és az ellipszoid első féltengelye mintegy háromszor nagyobb a másodiknál. Így evidens, hogy a feltétel melletti, első lépcsős feladat minimuma az  $x_1 > 0$  féltengelyen, a (0.9, 1.0) intervallumon van.

#### 4.4. Kombinált modell

A kétlépcsős modelltípus alkalmazásának egyik nehézsége abban áll, hogy nem minden esetben számítható ki a második lépcső feladata (mert adott  $\mathbf{x}, \boldsymbol{\xi}$  esetén nincsen megengedett megoldás). Ezt a gyakorlatban nehezen ellenőrizhető teljes pótlás feltételezésével szokták kikerülni (vagy a második lépcsős feladat átalakításával – újabb segédváltozók bevezetésével történő kibővítésével).

Prékopa megfogalmazott egy olyan modellt ([?], p. 417–418), amely matematikailag, a feladatba beépítve kezeli ezt a jelenséget. Ezt azzal éri el, hogy egy valószínűségi korlátot ír elő annak az eseménynek a valószínűségére, hogy a második lépcső feladatának van megengedett megoldása. (Ha nincsen megengedett megoldása a második lépcső feladatának, akkor az eltérést további költséggel bünteti, amelyet a pótló függvényhez ad hozzá.) Az így adott feladatban mind a kétlépcsős modellre jellemző pótló függvény, mind pedig egy valószínűségi korlát is megjelenik, ezért kombinált modellnek nevezzük.

##### 4.4.1. Kétlépcsős feladat valószínűségi korláttal

Tekintsük először az eredeti második lépcsős feladatot:

$$\begin{aligned} q(\mathbf{x}, \boldsymbol{\xi}) &= \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \\ \text{f.h. } T\mathbf{x} + W\mathbf{y} &= \boldsymbol{\xi}, \\ \mathbf{y} &\geq 0. \end{aligned} \tag{58}$$

Ez a feladat akkor és csak akkor oldható meg egy  $(\mathbf{x}, \boldsymbol{\xi})$  paraméterpár esetén, ha a következő feladatnak van megengedett megoldása

$$\begin{aligned} \min_{\mathbf{y}} \quad & \mathbf{0}'\mathbf{y}, \\ \text{f.h. } W\mathbf{y} &= \boldsymbol{\xi} - T\mathbf{x}, \\ \mathbf{y} &\geq 0. \end{aligned}$$

Ennek a duális feladata a következő formát ölti:

$$\begin{aligned} \max_{\mathbf{u}} \quad & (\boldsymbol{\xi} - T\mathbf{x})'\mathbf{u}, \\ \text{f.h. } W'\mathbf{u} &\leq 0. \end{aligned} \tag{59}$$

Ha a primál és a duál feladatnak is van megengedett megoldása, akkor a duál feladatnak korlátos a célfüggvénye a lineáris programozás dualitási tétele miatt. Jelöljük  $U = \{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(N)}\}$ -val az  $\{\mathbf{u} \mid W'\mathbf{u} \leq 0\}$  homogén lineáris egyenlőtlenség által adott kúp extrémális sugarainak halmazát. A  $(\boldsymbol{\xi} - T\mathbf{x})'\mathbf{u}^{(i)}$  függvényérték nem lehet pozitív egyetlen  $i = 1, \dots, N$  index esetén sem – ugyanis ebben az esetben a duál célfüggvény optimális értéke végtelen lenne. Tehát az extrémális sugarakkal adott  $(\boldsymbol{\xi} - T\mathbf{x})'U \leq \mathbf{0}$  egyenlőtlenségrendszer minden komponensre teljesül, következésképpen ennek az egyenlőtlenségrendszernek a teljesülése ekvivalens a primál probléma (vagyis a második lépcsős feladat) megoldhatóságával. Tehát a második lépcsős feladatnak akkor és csak akkor van megengedett megoldása, ha az

$$(60) \quad U'\boldsymbol{\xi} \leq U'T\mathbf{x}$$

egyenlőtlenségek teljesülnek. A kombinált modellben ezeknek az egyenlőtlenségeknek a teljesülését követeljük meg egy előírt  $p > 0$  valószínűséggel. Ha a  $\boldsymbol{\xi}$  valószínűségi vektorváltozó logkonkáv eloszlású, akkor a  $\mathbf{P}\{U'\boldsymbol{\xi} \leq U'T\mathbf{x}\}$  valószínűségi korlát az  $\mathbf{x}$  logkonkáv függvénye, tehát a megengedett megoldások halmaza konvex lesz.

Ezek szerint összefoglalhatjuk a kombinált modellt; az első lépcsős feladat

$$(61) \quad \begin{aligned} \min \mathbf{c}'\mathbf{x} &+ \mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi})), \\ \text{f.h. } A\mathbf{x} &\leq \mathbf{b}, \\ \mathbf{P}\{U'\boldsymbol{\xi} \geq U'T\mathbf{x}\} &\geq p, \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned}$$

ahol  $q(\mathbf{x}, \boldsymbol{\xi})$  függvényt a következő – módosított – második lépcsős feladat adja meg:

$$(62) \quad \begin{aligned} q(\mathbf{x}, \boldsymbol{\xi}) &= \min_{\mathbf{y}, \mathbf{z}^+, \mathbf{z}^-} (\mathbf{q}'\mathbf{y} + \mathbf{d}^{+'}\mathbf{z}^+ + \mathbf{d}^{-'}\mathbf{z}^-), \\ \text{f.h. } W\mathbf{y} + \mathbf{z}^+ - \mathbf{z}^- &= \boldsymbol{\xi} - T\mathbf{x}, \\ \mathbf{y}, \mathbf{z}^+, \mathbf{z}^- &\geq \mathbf{0}. \end{aligned}$$



Ennek a feladatnak a megadásában a  $W\mathbf{y}$  és a  $\boldsymbol{\xi} - T\mathbf{x}$  közti  $\mathbf{z}$  eltérés kényelmes kezelése céljából bevezettük a  $\mathbf{z} = \mathbf{z}^+ - \mathbf{z}^-$  jelölést; ez a felbontás lehetővé teszi, hogy különböző mértékben büntessük meg a pozitív, illetőleg negatív irányú eltérést. A modell leírásához hozzátartozik az a feltevés is, hogy ha egy adott  $(\mathbf{x}, \boldsymbol{\xi})$  esetén az (??) eredeti második lépcsős feladatnak létezik megengedett megoldása, akkor azt kell megoldani, nem pedig ezt a módosított második lépcsős feladatot. Ebben a modellben a  $\mathbf{d}^+, \mathbf{d}^- \geq 0$  költségeknek pozitívaknak kell lenniük mindig és elég nagynak ahhoz, hogy az eltérések minimalizálása „fontosabb” legyen, mint a  $\mathbf{q}'\mathbf{y}$  tag minimalizálása.

A kombinált modell megoldása ezek szerint az (??) feladat optimalizálása, miközben a célfüggvényben szereplő  $\mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$  tag a (??) feladatból számítható. Erre a feladatra az SRA algoritmuson kívül nem ismeretes használható megoldó algoritmus.

#### 4.4.2. Regresszió a kombinált modell megoldására

A kombinált modellben két nehezen kiszámítható függvény van: az  $\mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$  várható pótlás függvénye és a  $\mathbf{P}\{U'\boldsymbol{\xi} \leq U'T\mathbf{x}\}$  valószínűségi korlát, mindkettőnek csak zajos függvényértékeit tudjuk kiszámítani. Az SRA algoritmus alkalmazásának szellemében ezt a két nehezen (zajosan) kiszámítható függvényt helyettesítjük két regressziós függvénnyel, az  $r_k(\mathbf{x})$  illetőleg az  $f_k(\mathbf{x})$  függvénnyel. Az első függvény kiszámításának (becslésének) módszerét a kétlépcsős feladatnál írtuk le, a  $\mathbf{P}\{U'\boldsymbol{\xi} \leq U'T\mathbf{x}\}$  valószínűséget helyettesítő függvény meghatározása pedig a valószínűségi korlátok meghatározásánál van megadva – de ez a két eljárás lényegében ugyanaz.

Az adott  $\mathbf{x}_k$  pontban a pótló függvény egy zajos  $q_k$  függvényértéket a következőképpen határozhatjuk meg. Legyenek a  $\boldsymbol{\xi}_{k_i}, i = 1, \dots, N^{(q)}$  értékek a  $\boldsymbol{\xi}$  valószínűségi realizációi, és határozzuk meg a következő feladatok optimális  $q_{k_i}$  célfüggvényértékét:

$$\begin{aligned}
q_{k_i} &= \min \left( \mathbf{q}' \mathbf{y} + \mathbf{d}^{+'} \mathbf{z}^+ + \mathbf{d}^{-'} \mathbf{z}^- \right) \\
(63) \quad W \mathbf{y} + \mathbf{z}^+ - \mathbf{z}^- &\geq \boldsymbol{\xi}_{k_i} - T \mathbf{x}_{k_i}, \\
\mathbf{y}, \mathbf{z}^+, \mathbf{z}^- &\geq 0.
\end{aligned}$$

Legyen  $q_k = \frac{1}{N^{(q)}} \sum_{i=1, \dots, N^{(q)}} q_{k_i}$ , ami torzítatlan becslése a  $Q(\mathbf{x}_k)$  függvényértéknek. Adott  $\{\mathbf{x}_i\}$  pontsorozatra ezzel meghatározhatjuk az  $S_k = \{\mathbf{x}_i, q_i\}_{i=0}^{k-1}$  pontokból és függvényértékekből álló halmazt. Ennek segítségével a  $Q(\mathbf{x})$  függvényt közelítő  $r_k(\mathbf{x}) = \mathbf{x}' D_k \mathbf{x} + \mathbf{b}'_k \mathbf{x} + c_k$  alakú regressziós függvény meghatározható. Itt feltesszük, hogy  $D_k$  pozitív definit szimmetrikus mátrix.

A valószínűségi korlátnak egy adott  $\mathbf{x}_k$  esetén felvett függvényértéke a durva Monte Carlo módszerrel becsülhető. Legyen egy  $N^{(p)}$  elemű mintánk a  $\boldsymbol{\xi}$  valószínűségi változóból a  $\boldsymbol{\xi}_{k_i}, i = 1, \dots, N^{(p)}$ . A  $\chi_{k_i}$  valószínűségi változót definiáljuk a következőképpen:

$$(64) \quad \chi_{k_i} = \begin{cases} 1, & \text{ha } \{U' \boldsymbol{\xi}_{k_i} \leq U' T \mathbf{x}_k\}, \\ 0, & \text{egyébként.} \end{cases}$$

Az adott mintához meghatározzuk a  $\chi_{k_i}, i = 1, \dots, N^{(p)}$  értékeket. A zajos  $p_k \sim \mathbf{P} \{U' \boldsymbol{\xi} \geq U' T \mathbf{x}_k\}$  függvényértéket pedig a relatív gyakorisággal kaphatjuk meg, vagyis legyen

$$p_k = \frac{1}{N^{(p)}} \sum_{i=1}^{N^{(p)}} \chi_{k_i}.$$

Az így kapott zajos függvényértékekkel pedig előállítottuk a  $P_k = \{\mathbf{x}_i, p_i\}_{i=0}^{k-1}$  pontokból és zajos függvényértékekből álló halmazt, amelynek segítségével meghatározható az az  $f_k(\mathbf{x})$  regressziós függvény, amely a  $\mathbf{P} \{U' \boldsymbol{\xi} \geq U' T \mathbf{x}\}$  függvényt közelíti. Természetesen a regressziós függvényt  $f_k(\mathbf{x}) = \mathbf{x}' F_k \mathbf{x} + \mathbf{e}'_k \mathbf{x} + h_k$  alakban keressük, ahol a  $-F_k$  mátrixról feltesszük, hogy szimmetrikus és pozitív definit.

Mindkét zajos függvényérték szórása függ attól az  $\mathbf{x}_k$  ponttól, amelyben meghatározzuk, de ezeknek a szórását itt is konstansnak tesszük fel (természetesen  $D^2(q_i) \neq D^2(p_i)$  általában).

#### 4.4.3. Algoritmus a kombinált modellre

Tegyük fel, hogy a rendelkezésünkre áll egy  $\{\mathbf{x}_k\}_{i=0}^{k-1}$  pontsorozat esetén két pontfüggvény halmaz, az

$$S_k = \{\mathbf{x}_i, q_i\}_{i=0}^{k-1}, \quad P_k = \{\mathbf{x}_i, p_i\}_{i=0}^{k-1}$$

halmazok, amelyek a fenti módon határozhatók meg. Az SRA algoritmus formális leírása a következő:

SRA algoritmus (kombinált modell)

0. Előkészítés. Legyen  $k$  az eredetileg adott pontok száma.
1. a.) Számítsuk ki a  $r_k(\mathbf{x})$  regressziós függvény  $D_k, \mathbf{b}_k, c_k$  együtthatóit az  $S_k$  halmaz segítségével.
1. b.) Számítsuk ki a  $f_k(\mathbf{x})$  regressziós függvény  $F_k, \mathbf{e}_k, h_k$  együtthatóit a  $P_k$  halmaz segítségével.
2. Oldjuk meg az (??) feladatot közelítő következő optimalizálási problémát:

$$(65) \quad \begin{aligned} \min_{\mathbf{x}} \quad & (\mathbf{c}'\mathbf{x} + r_k(\mathbf{x})) \\ & A\mathbf{x} = \mathbf{b}, \\ & f_k(\mathbf{x}) \geq p, \\ & \mathbf{x} \geq 0 \end{aligned}$$

és legyen ennek egy optimális megoldása az  $\mathbf{x}_k$  pont.

3. Ha  $\mathbf{x}_k$  „elég jó”, akkor STOP, egyébként valamilyen módszerrel határozzuk meg a  $q_k \sim Q(\mathbf{x}_k)$  és  $f_k \sim \mathbf{P}\{U'\boldsymbol{\xi} \geq U'T\mathbf{x}_k\}$  zajos függvényértékeket. Bővítsük a pontokból és függvényértékekből álló halmazainkat: legyen  $S_{k+1} = S_k \cup \{\mathbf{x}_k, q_k\}$ ,  $P_{k+1} = P_k \cup \{\mathbf{x}_k, f_k\}$ , növeljük meg az iterációs számlálót  $k := k + 1$  és menjünk vissza az 1. lépésre.

	$\mathbf{x}$	$\mathbf{c}'\mathbf{x} + \mathbf{E}(\mathbf{q}'\mathbf{y})$	$\mathbf{P}\{\cdot\}$	$N$	No.iter
$\mathbf{d}^1$	(0.768, 0.087)	50.75( $\pm 2.3$ )	0.921( $\pm 0.030$ )	100	100
		49.18( $\pm 0.26$ )	0.909( $\pm 0.003$ )		
$\mathbf{d}^2$	(0.768, 0.086)	50.76( $\pm 2.3$ )	0.918( $\pm 0.030$ )	100	100
		49.14( $\pm 0.25$ )	0.906( $\pm 0.003$ )		

8. táblázat. Kombinált feladat: SRA eredmények különböző paraméterértékekre; a megengedettségre  $p = 0.9$  valószínűséget írtunk elő.

#### 4.4.4. Numerikus példa

A kombinált feladat szemléltetésére használt numerikus példát az előző szakaszban tárgyalt kétlépcsős példából alakítottuk ki. Az ott felírt feladat teljes pótlású volt, az abban szereplő  $W$  mátrix első oszlopát kitöröltük, hogy egyes esetekben a második lépcsős problémának ne legyen megengedett megoldása. Az ottani feladat egyéb konstansait nem változtattuk meg. Tehát legyen most a mátrixunk

$$W = \begin{pmatrix} -0.7 & 1.7 \\ 9.0 & -13.0 \end{pmatrix}.$$

A  $W'\mathbf{u} \leq 0$  poliedrikus kúp extrémális sugarai ekkor a következők lesznek:  $\mathbf{u}^{(1)'} = (-9.0, -0.7)$ ,  $\mathbf{u}^{(2)'} = (-13.0, -1.7)$ . A kombinált modell valószínűségi korlátjában szereplő két mátrix a következő:

$$U' = \begin{pmatrix} -9.0 & -0.7 \\ -13.0 & -1.7 \end{pmatrix}, \quad U'T = \begin{pmatrix} -49.53 & 83.64 \\ -67.29 & 121.64 \end{pmatrix}.$$

Az második lépcső eltéréseit kiegyenlítő kiegészítő  $\mathbf{z}^+$  és  $\mathbf{z}^-$  vektorokhoz tartozó költségtényezők legyenek  $\mathbf{d}^1 = (2.0, 10.0)$ ,  $\mathbf{d}^2 = (10.0, 2.0)$ , amelyek pozitív és negatív irányú eltérésekre ugyanazok. Itt szándékosan két lényegesen különböző költséggel dolgozunk, hogy a feladat numerikus érzékenységét vizsgálni lehessen.

A számítási eredményeket két táblázatban adtuk meg. A 8. táblázatban adott eredmények azt mutatják, hogy a két különböző  $\mathbf{d}^1$  és  $\mathbf{d}^2$  költségfaktor nem befo-

$k =$	$\mathbf{x}$	$\mathbf{c}'\mathbf{x} + E(\mathbf{q}'\mathbf{y})$	$r_k(\mathbf{x})$	$\mathbf{P}\{\cdot\}$	$f_k(\mathbf{x})$
1	(0.7741, 0.0804)	47.01	45.44	0.940	0.89
10	(0.7677, 0.0871)	51.07	50.13	0.930	0.90
100	(0.7677, 0.0871)	45.94	49.54	0.860	0.90
1000	(0.7708, 0.0838)	46.87	48.55	0.860	0.90
4000	(0.7722, 0.0823)	45.88	48.29	0.910	0.90
9999	(0.7718, 0.0828)	51.54	48.03	0.910	0.90

9. táblázat. Kombinált feladat: az SRA algoritmus  $k$ -adik iterációinak eredményei, költségvektor  $\mathbf{d} = (9.0, 10.0)$ , a mintaszám  $N^{(p)} = 100$ ,  $\mathbf{D}^2(q_i) = 2.3^2$ ,  $\mathbf{D}^2(p_1) = 0.03^2$ . „Pontos” értékek az utolsó pontban: költségfüggvény értéke  $48.18(\pm 0.08)$ , a megbízhatóság értéke  $0.9005(\pm 0.0009)$ .

lyásolta az optimális megoldás értékét. Itt mindkét költségfüggvény esetén két-két sort adtunk meg; az első sor az aktuális futás eredménye, a második sor a „pontos” függvényértékeket mutatja, vagyis az utolsó pontban a megnövelt mintaszámmal végrehajtott becslés eredményét adtuk meg.

A másodikként adott 9. táblázatban azt szemléltetjük, hogy egy számítógépes futás alatt hogyan csökken a célfüggvény értéke és hogyan lesz egyre pontosabb a korlát. Az  $r_k(\mathbf{x})$  és  $f_k(\mathbf{x})$  oszlopok mutatják, hogy az aktuális regressziós függvény értéke mennyi az adott pontban. A „pontos” értékek hibája megfelel a sejtés alapján számított hibának: például az utolsó iterációban a valószínűségi korlát értéke  $0.9005(\pm 0.0009)$ , amelynek a  $p = 0.9$ -tol való eltérés kisebb mint  $3\sigma_P = 3\sigma_1/\sqrt{10000} \sim 3 \cdot 0.03/100 = 0.001$  – vagyis az optimális szórásnövekedés sejtése itt is igaznak látszik.

#### 4.5. Sztochasztikus kvadratikus programozás

A gyakorlatban használt és számítástechnikailag kezelhető sztochasztikus programozási feladatok túlnyomó része lineáris feltételeket és lineáris célfüggvényt tartalmaz,

ezért ezeket összefoglalóan sztochasztikus lineáris programozási modelleknek nevezzük. Vannak olyan sztochasztikus programozási feladatok, amelyekben valamilyen egyszerűsítő feltétel mellett van adva kvadratikus célfüggvény, például ha szeparálható részekből áll. A fenti kombinált modell megoldhatósága alapján a gyakorlatban optimalizálható a következő, általános kvadratikus sztochasztikus programozási modell, amelyben nem teszünk fel dekomponálhatóságot:

$$\begin{aligned}
 & \min (F(\mathbf{x}) + Q(\mathbf{x})) \\
 (66) \quad & \text{f.h. } G(\mathbf{x}) = \mathbf{P}\{A\mathbf{x} \geq \boldsymbol{\eta}\} \geq p, \\
 & H(\mathbf{x}) \leq 0, \\
 & \mathbf{x} \geq 0.
 \end{aligned}$$

A feladatban adott  $F(\mathbf{x}), H(\mathbf{x})$  függvényekről feltesszük, hogy konvex kvadratikus függvények, a  $Q(\mathbf{x}) = \mathbf{E}(q(\mathbf{x}, \boldsymbol{\xi}))$  várható pótlás függvényét pedig a szokásos módon a következő feladatból számíthatjuk ki:

$$\begin{aligned}
 q(\mathbf{x}, \boldsymbol{\xi}) &= \min_{\mathbf{y}} \mathbf{q}'\mathbf{y} \\
 \text{f.h. } T\mathbf{x} + W\mathbf{y} &= \boldsymbol{\xi}, \\
 \mathbf{y} &\geq 0.
 \end{aligned}$$

Ennek a (??) sztochasztikus kvadratikus programozási feladatnak az elvi megoldhatóságához a megengedett megoldások halmazának konvexitása szükséges. Ehhez az  $F(\cdot)$  és  $H(\cdot)$  függvények Hesse mátrixainak pozitív definitésén kívül a következő feltevéseket kell tenni:

- (i) a  $\boldsymbol{\xi}, \boldsymbol{\eta}$  valószínűségi vektorváltozók logkonkáv eloszlásúak,
- (ii) a (??) feladat legalább relatíve teljes pótlású.

Természetesen az előzőekben leírt kombinált feladat kapcsán megfogalmazott elgondolások szerint a második feltételezés elhagyható, csak akkor a második lépcsős feladatot módosítani kell és egyúttal az első lépcsős feladatába be kell építeni egy valószínűségi korlátot a második lépcsős feladatának megoldhatóságára.

A (??) kvadratikus modellnek a gyakorlati megoldhatóságát az SRA algoritmus biztosítja, hiszen a közelítő kvadratikus feladatok a fenti feltételek mellett konvex optimalizálási feladatok, amelyek például a MINOS szubrutin-csomaggal megoldhatók.

## 5. Irodalom

A sztochasztikus programozás irodalma nagyon terjedelmes és sokrétű. A hivatkozások közül mi csak a legfontosabbakat említjük meg, további hivatkozások a megadott művekben, illetőleg bibliográfiákban találhatók.

Alapvető műnek tartjuk Prékopa András nagyon sok részletre kiterjedő, mondhatnánk enciklopédikus könyvét [?], amelyben több mint ötszáz irodalmi hivatkozás is található. Ezt kutatók és hallgatók egyaránt haszonnal forgathatják. A valószínűségi korlátozott modellek leírása ilyen mélységben nem található meg más könyvekben. Az általunk érintett témákon kívül sok más téma és jónéhány részletesen kidolgozott alkalmazás is le van írva könyvében. Jelölés rendszere is találó és konzekvens – bevezetésünk megírásakor nagyrészt az általa használt jelöléseket alkalmaztuk.

A bevezető fejezetben leírt modellek nagy része megtalálható a [?], [?] könyvekben. A felsőbbéves hallgatóknak érdemes a Kall és Wallace által írt, kissé didaktikusabb, de kevesebb témát (nagyrészt a kétlépcsős modellt) tárgyaló könyvét [?] forgatni – az adaptív korlátozás elméletének kidolgozása jó. További összefoglaló művek közül a kétlépcsős modellt és annak speciális eseteit részletesen tárgyaló Birge [?] könyvet ajánljuk, valamint a megoldó algoritmusokkal, különösen az adaptív korlátozással foglalkozó Mayer által írt [?] könyvet ajánljuk.

A valószínűséggel korlátozott feladatok leírásánál Prékopa [?] könyvét követtük, ott több, részben egyszerűbb, részben bonyolultabb modell és megoldó algoritmus is megtalálható. A logkonkavitással kapcsolatos eredmények Prékopa [?] könyvében és [?] cikkében, valamint az ezekben megadott hivatkozásokban találhatók. A valószínűségek korlátozására vonatkozóan a [?], [?] cikkeket ajánljuk. A normális valószínűségek kiszámítására vonatkozó részek részletesebb kifejtése megtalálható [?], [?], [?], az általános eloszlásokra vonatkozó algoritmusok a [?] cikkben. Az eddigi, normális valószínűségek kiszámítására vonatkozó eredményeket összefoglaló műként



a Gassmann, Deák, Szántai által írt [?] cikket ajánljuk, valamint az ott található hivatkozásokra utalunk. Monte Carlo módszerek általános leírása Hammersley és Handscomb [?] alapvető fontosságú könyvében és a [?] könyvben található. A STABIL modell eredetileg a [?] cikkben jelent meg – ez volt az első együttes valószínűségi korlátot alkalmazó ipari feladat megoldását bemutató cikk.

A kétlépcsős modellek vonatkozásában Kall [?] és Mayer [?] könyveit és egyéb, a témában írt cikkeit, Wets [?], [?], [?], Strazicky [?] műveit tartjuk alapvetőnek. A kétlépcsős modellel kapcsolatos eredmények leírása [?]-ban és [?]-ben található. A célfüggvény korlátozásának vonatkozásában [?]-t tartjuk mérvadónak, további eredményeket Frauendorfer ért el.

A szukcesszív regressziós approximációk elméletének és számítástechnikai eredményeinek leírása Deák különböző cikkeiben található [?], [?], [?] – ezekben a cikkekben az SRA alkalmazásának és numerikus megfontolásainak fontos részletei megtalálhatók.

A lineáris programozás elmélete és módszerei sok helyen megtalálhatók, ezek közül Prékopa jegyzetét tudjuk ajánlani [?], vagy a jelen, „Bevezetés a sztochasztikus programozásba” c. jegyzettel azonos sorozatban megjelent, Komáromi által írt [?] jegyzetet javasoljuk. Zajos függvények melletti optimalizálási módszereket írnak le Ermoliev (a kvázigradiens alkalmazása), Robbins és Monro, Kushner, Ruszczyński [?], Hagle és Sen [?]. A nemlineáris (determinisztikus) optimalizálás tárgyalását például a Luenberger által írott, nagyon jól olvasható [?], [?] könyvekben találhatjuk meg.

A legújabb kutatási irányokról és témákról jó áttekintést adnak a Stochastic Programming, háromévente megrendezésre kerülő konferenciák, illetőleg az itt elhangzott előadásokat tartalmazó gyűjtemények. Az utolsó két ilyen konferencia-kötet a Vancouverben és Berlinben tartott konferenciák összefoglalása (lásd a [?] és [?]

adatait).

A web-en is sok érdekes anyag található a sztochasztikus programozás témaköréből. Ezek közül elsősorban Maarten van der Vlerk (<http://mally.eco.rug.nl/spbib.html>) honlapját ajánljuk megnézésre, amelyen egyrészt egy több mint háromezer cikket tartalmazó bibliográfiát lehet találni sztochasztikus programozási cikkekből, továbbá jónéhány értékes linket. Innen lehet elérni a Mathematical Programming Society, Committee on Stochastic Programming (COSP) társaságát, s egyéb kapcsolati lehetőségeket. Egyedi keresések közül a „stochastic programming”, „probabilistic constrained optimization”, „chance constraint”, „two-stage problems” címszavakat, valamint egyes fentebb említett kutatók (Wets, Birge, Rockafellar, Mayer, Shapiro, Morton) nevét érdemes keresni. Rockafellar az interneten közzétett egy mintegy száz oldalas anyagot, amelyben a sztochasztikus programozás konvex analízisbeli eszközökkel való leírását tartalmazza.

## Hivatkozások

- [1] Birge, J., Louveaux, F.: Introduction to stochastic programming, Springer, 1997, pp. 421.
- [2] Box, G.E.P., Draper, N.R.: Empirical model-building and response surface, J. Wiley and Sons, 1987.
- [3] Bukszár, J., Prékopa, A.: Probability bounds with cherry trees, Mathematics of Operations Research, 26 (2001) 174-192.
- [4] Bukszár, J., Szántai, T.: probability bounds given by hyper-cherry trees, Optimization Methods and Software, 17 (2002) 409-422.
- [5] Deák, I.: Computer evaluation of a stochastic programming model, MTA Számítástechnikai Központ Közleményei, 9 (1972) 33-49 (in Hungarian).
- [6] Deák, I.: Three digit accurate multiple normal probabilities, Numerische Math. 35 (1980) 369-380.
- [7] Deák, I.: Véletlenszámgenerátorok és alkalmazásuk, Akadémiai Kiadó, 1986.
- [8] Deák, I.: Random Number Generators and Simulation, in: Mathematical Methods of Operations Research, ed. A.Prékopa, Akadémiai Kiadó, Budapest, 1990, pp. 342.
- [9] Deák, I.: Linear regression estimators for multinormal distributions in optimization of stochastic programming models, European J. Operational Research, 111 (1998) 555-568.
- [10] Deák, I.: Subroutines for computing normal probabilities of sets - computer experiences, Ann. of Oper. Res., V 100 (2000), 103-122.
- [11] Deák, I.: Computer experiences with successive regression approximations for solving equations, Optimization Theory (F. Gianessi, P. Pardalos, T. Rapcsák, eds.) Kluwer Academic Publishers, Dordrecht-Boston-London (2001) 65-80.

- [12] Deák, I.: Successive regression approximations for solving equations, *Pure Mathematics and Applications*, 12 (2001), 25-50.
- [13] Deák, I.: Computing two-stage stochastic programming problems by successive regression approximations, *Stochastic Optimization Techniques – Numerical Methods and Technical Applications* (ed. K. Marti) Springer LNEMS V. 513 (2002) 91-102.
- [14] Deák, I.: Probabilities of simple  $n$ -dimensional sets for the normal distribution, *IIE Transactions on Operations Engineering*, V. 35 (2003) 285-293.
- [15] Deák, I.: Solving stochastic programming problems by successive regression approximations, *Proceedings of a Conference on Stochastic Programming*, 2002, Laxenburg, Springer, megjelenés alatt.
- [16] Deák, I.: Two-stage stochastic problems with correlated normal variables: computational experiences, *Annals of Operations Research*, 2002, megjelenés alatt.
- [17] Deák, I.: Convergence of successive regression approximations for solving equations with noisy functions, *kézirat* (2003).
- [18] Ermoliev, Ju.M.: Methods of stochastic programming, in: *Optimization and Operations Research*, Nauka, 1976, p. 239 (in Russian).
- [19] Ermoliev, Ju.M.: Stochastic quasigradient methods and their applications to systems optimization, *Stochastics* 9 (1983) 1-36.
- [20] Gassmann, H., Deák, I., Szántai, T.: Computing multivariate normal probabilities: a new look, *J. Computational and Graphical Statistics*, 11 (2002) 920-949.
- [21] Hammersley and Handscomb: *Monte Carlo methods*, Methuen, London, 1964.
- [22] Higle, J.L., Sen, S.: *Stochastic decomposition: a statistical method for large scale stochastic linear programming*, Kluwer Academic Publishers, 1996.
- [23] Kall, P., Wallace, S.W.: *Stochastic Programming*, Wiley, 1994, pp. 307.

- [24] Kall, P., Mayer, J.: SLP-IOR: An interactive model management system for stochastic linear programs, *Math. Programming* 75 (1996) 221-240.
- [25] Kiefer, J., Wolfowitz, J.: Stochastic estimation of a regression function, *Ann. Math. Stat.* 23 (1952) 462-466.
- [26] Komáromi, É.: A dual method of probabilistic constrained problem, *Mathematical Programming Study* 28 (1986) 94-112.
- [27] Komáromi, É.: On properties of the probabilistic constrained linear programming problem and its dual, *J. Optimization Theory and Applications* 55 (1987) 337-390.
- [28] Komáromi É.: Lineáris programozás, BKÁE, Operációkutatási Tanszék, Operációkutatás No.2. 2002.
- [29] Kushner, H.J., Clark, D.S.: Stochastic approximation methods for constrained and unconstrained systems, in: *Applied Mathematical Sciences* 26, Springer, 1978, pp. 261.
- [30] Luenberger, D.G.: *Optimization by vector space methods*, Reading, MA., Addison-Wesley, 1969. pp. 326.
- [31] Luenberger, D.G.: *Linear and nonlinear programming*, Reading, MA., Addison-Wesley, 1984. pp.
- [32] Marti, K.: Semi-stochastic approximation by the response surface methodology (RSM), *Optimization* 25 (1992) 209-230.
- [33] Mak, W.K., Morton, D.P., Wood, R.K. (1999): Monte Carlo bounding techniques for determining solution quality in stochastic programs, *Operations Research Letters* 24, 47-56.
- [34] Mayer, J.: Computational techniques for probabilistic constrained optimization problems, in: *Lecture Notes on Economics and Mathematical Systems*, Springer, V379 (1992) 141-164

- [35] Mayer, J.: Stochastic linear programming algorithms, Gordon and Breach, 1998, pp. 153.
- [36] Prékopa, A.: Lineáris programozás I., Bolyai J. társulat kiadványa, 1968.
- [37] Prékopa, A.: Logarithmic concave measures with applications to stochastic programming, *Acta Scientiarum Mathematicarum* (Szeged) 32 (1971) 301-316.
- [38] Prékopa, A.: On logarithmic concave measures and functions, *Acta Scientiarum Mathematicarum* (Szeged) 34 (1973) 335-343.
- [39] Prékopa, A.: Contributions to the theory of stochastic programming, *Mathematical Programming* 4 (1973) 202-221.
- [40] Prékopa, A.: Numerical solution of probabilistic constrained programming problems, In: Numerical techniques for stochastic optimization, ed. Y. Ermoliev, R.Wets, Springer series in computational mathematics, Springer Verlag, (1988) 123-139.
- [41] Prékopa, A.: Stochastic Programming, in: Mathematics and its Applications 324, Kluwer, 1995.
- [42] Prékopa, A., Ganczer, S. Deák, I., Patyi, K.: The STABIL stochastic programming model and its experimental application to the electrical energy sector of the Hungarian economy, in: Proc. of the International Symp. on Stochastic Programming, ed. M.Dempster (1980) Academic Press 369-385.
- [43] Prékopa, A., Szántai, T.: A new multivariate gamma distribution and its fitting to empirical data, *Water Resources Research* 14 (1978) 19-24.
- [44] Rapcsák, T.: A SUMT módszer alkalmazása logaritmikusan konkáv feltételi függvényeket tartalmazó nem-lineáris programozási feladat megoldására, MTA Számítástechnikai és Automatizálási Kutató Intézet közleményei, 19 (1978) 17-27.

- [45] Robbins, H., Monro, S.: A stochastic approximation method, *Ann. Math. Stat.* 22 (1951), 400-407.
- [46] Ruszczyński, A.: Feasible direction methods for stochastic programming problems, *Math. Programming* 19 (1980) 220-229.
- [47] Shapiro, A., Homem-de-Mello, T. (1998): A simulation based approach to two-stage stochastic programming with recourse, *Math. Programming* 81, 301-325.
- [48] Strazicky, B.: On an algorithm for solution of the two-stage stochastic programming problem, *Methods of Operation Research* 19 (1974) 142-156.
- [49] Szántai, T.: Improved bounds and simulation procedures on the value of the multivariate normal probability distribution function, *Annals of Operations Research* 100 (2000) 85-101.
- [50] Szántai, T.: Evaluation of a special multivariate gamma distribution function, *Math. Programming Study*, 27 (1986) 1-16.
- [51] Van Slyke, R., Wets, R.J.-B.: L-shaped linear programs: algorithmic techniques and numerical results, *SIAM J. Appl. Mat.* 17 (1969) 638-663.
- [52] Wets, R.J.-B.: Stochastic programs with fixed recourse– the equivalent deterministic program, *SIAM Review* 26 (1974) 309-339.
- [53] Wets, R.J.-B.: Large scale linear programming techniques in stochastic programming, in: *Numerical techniques for stochastic optimization 1984*, Springer series in computational mathematics, (eds. Y. Ermoliev, R. Wets.), Springer Verlag, 1988, 65-94.
- [54] Wets, R.J.-B., Ziemba, W.T.(eds.): *Stochastic programming: State of the Art*, *Ann. of Operations Research*, Baltzer, 85 (1999) pp. 285.

