Bizonytalan konvex optimalizálás

Szakdolgozat

Készítette: Szabó Elek Zsolt

Matematika BSc
Almalmazott matematikus szakirány

Témavezető: Dr. Csáji Balázs Csanád

Tudományos főmunkatárs, SZTAKI
Adjunktus, Matematikai Intézet, ELTE

Matematika BSc

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2022

Tartalomjegyzék

Ábrák jegyzéke

Köszönetnyilvánítás

Köszönetet szeretnék mondani Dr. Csáji Balázs Csanádnak a témaválasztás során kapott segítségéért, és a szakdolgozatom elkészítése során felmerült kérdések megválaszolásáért.

1. Bevezetés

A szakdolgozatom azzal foglalkozik, hogy hogyan tudunk optimalizálni akkor, ha bizonytalanság is jelen van az optimalizálandó feladatban. Az ezzel foglalkozó (sztochasztikus optimalizálás [4]) megoldási módszerek közül az úgynevezett szkenárió módszerrel fogok foglalkozni. Egy szkenárió az a bizonytalanság egy példánya, és a szkenárió módszer pedig ezeknek egy véges halmazát véve dolgozik (ezért lesz a gyakorlatban jól használható), és ad egy megoldást a feladatra. A későbbi fejezetekben ennek a megoldásnak a "jóságát" fogjuk megnézni, hogy milyen valószínűséggel sért meg bizonyos feltételeket. A szkenárió módszer a standart sztochasztikus programozással szemben az a hatalmas előnye, hogy eloszlás-független, így sokkal szélsesebb körben és egyszerűbben alkalmazható.

Az optimalizálás a mindennapjainkban is jelen van, hiszen amikor az igényeinket kielégítő terméket szeretnénk venni, akkor sokszor a lehető legkisebb olyan árat keressük, amihez tartozó termék még megfelelő az elvárásainknak, azaz megpróbálunk optimalizálni az ár és a tulajdonságok szerint. Ám sokszor nem ilyen egyszerű a helyzet, mert valamilyen bizonytalansági paraméter is szerepet játszhat a döntésünk során. A szkenárió módszer ilyen feladatokkal foglalkozik, és csak véges minta alapján dolgoz, amiket valószínűleg kísérletek által kapunk meg. A szakdolgozathoz tartozó forrás többnyire az [1] lesz.

1.1. Bizonytalan lineáris programozás ellipszoidon

Ebben az alfejezetben a forrásom a [3].

Egy lineáris programozási (LP) feladat a következő alakú:

$$LP: \quad \min_{x \in \mathbb{R}^n} c^{\mathrm{T}} x + d \qquad \text{ahol} \quad Ax \le b.$$

A fentiekben $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, $b \in \mathbb{R}^m$ és $A \in \mathbb{R}^{m \times n}$. A feladat adatainak a (c, d, A, b) négyest nevezzük. Ezeket kicsit kompaktabb módon is reprezálhatjuk, egy $(m+1) \times (n+1)$-es adat-mátrix segítségével

$$D = \left(\begin{array}{c|c} c^T & d \\ \hline A & b \end{array} \right).$$

Azonban a valóságban előforduló lineáris programozási feladatokban gyakran az a helyzet adódik, hogy az adatokat nem ismerjük teljesen, azaz valamilyen bizonytalanságtól függnek, vagy eleve hibaterheltnek tekintjük őket. Ezért van szükségünk a bizonytalan lineáris programozási feladatra.

1.1. Definíció. *Egy bizonytalan lineáris programozási feladat az a*

$$\left\{ \min_x \{c^{\mathrm{T}}x + d \ : Ax \leq b\} \right\}_{(c,d,A,b)\in\mathcal{U}}$$

lineáris programozási feladatok családja, ahol az adatok (az adat-mátrix) egy $\mathcal{U} \subseteq \mathbb{R}^{(m+1)\times(n+1)}$ *bizonytalansági halmazból kerülnek ki. Jelölése:* $LP_{\mathcal{U}}$.

Mindig feltételezzük, hogy a bizonytalan halmazunk affin módon paraméterezve van, azaz

$$\mathcal{U} = \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = \left(\begin{array}{c|c} c_0^{\mathrm{T}} & d_0 \\ \hline A_0 & b_0 \end{array}\right) + \sum_{\ell=1}^{L} \zeta_\ell \left(\begin{array}{c|c} c_\ell^{\mathrm{T}} & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right) : \ \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\},$$

valamilyen $\mathcal{Z}$ perturbációs halmazra és $\zeta = (\zeta_1, \zeta_2, \ldots, \zeta_L)^T \in \mathcal{Z}$ vektorra. Vezessük be a következő jelölést

$$D_i = \left(\begin{array}{c|c} c_i^{\mathrm{T}} & d_i \\ \hline A_i & b_i \end{array}\right).$$

Vegyük észre, hogy ha a $\mathcal{Z}$ perturbációs halmaz egy másik, $\hat{\mathcal{Z}}$ perturbációs halmaz képe egy M affin leképezés alatt, ahol $M : \mathcal{Z} \to \hat{\mathcal{Z}}$, $M(\xi) = \zeta = p + P\xi$, akkor a $\mathcal{Z}$ által kapott paraméterezést átírhatjuk a $\hat{\mathcal{Z}}$ által kapott paraméterezésre:

$$\begin{aligned}
\mathcal{U} &= \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = \left(\begin{array}{c|c} c_0^{\mathrm{T}} & d_0 \\ \hline A_0 & b_0 \end{array}\right) + \sum_{\ell=1}^{L} \zeta_\ell \left(\begin{array}{c|c} c_\ell^{\mathrm{T}} & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right) : \ \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\} \\
&= \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = D_0 + \sum_{\ell=1}^{L} \zeta_\ell D_\ell : \ \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\} \\
&= \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = D_0 + \sum_{\ell=1}^{L} \left(p_\ell + \sum_{k=1}^{K} P_{\ell k}\xi_k \right) D_\ell : \ \xi \in \hat{\mathcal{Z}} \right\} \\
&= \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = \left(D_0 + \sum_{\ell=1}^{L} p_\ell D_\ell \right) + \sum_{k=1}^{K} \xi_k \left(\sum_{\ell=1}^{L} P_{\ell k} D_\ell \right) : \ \xi \in \hat{\mathcal{Z}} \right\} \\
&= \left\{ \left(\begin{array}{c|c} c^{\mathrm{T}} & d \\ \hline A & b \end{array}\right) = \hat{D}_0 + \sum_{k=1}^{K} \xi_k \hat{D}_k : \ \xi \in \hat{\mathcal{Z}} \subset \mathbb{R}^K \right\}.
\end{aligned}$$

Ebből annyit használunk fel, hogy ha egy bizonytalan lineáris programozási feladat adva van egy ellipszoidon, akkor elegendő a feladatunkat csak gömbön nézni, hiszen egy ellipszoid az egységgömb $\{\xi \in \mathbb{R}^K : ||\xi||_2 \leq 1\}$ képe affin leképezés alatt.

1.2. Definíció. *Egy $x \in \mathbb{R}^n$ vektort az $LP_{\mathcal{U}}$ bizonytalan lineáris programozási feladat robosztus megoldásának (robust feasible solution) nevezzük, ha minden $(c, d, A, b) \in \mathcal{U}$ négyesre teljesül, hogy $Ax \leq b$.*

1.3. Definíció. *Egy $x \in \mathbb{R}^n$ vektornak a $\hat{c}(x)$ robosztus értéke az $LP_{\mathcal{U}}$ feladatra nézve az a legnagyobb érték, amit az $\mathcal{U}$ bizonytalansági halmazon felvehet, azaz*

$$\hat{c}(x) = \sup_{(c,d,A,b)\in\mathcal{U}} c^{\mathrm{T}}x + d.$$

Ezzel a kettő definícióval már meg tudjuk fogalmazni, hogy mit szeretnénk optimalizálni egy $LP_{\mathcal{U}}$ feladatban.

1.4. Definíció. *Egy bizonytalan $LP_{\mathcal{U}}$ feladatnak a Robosztus Megfelelőjének (RM) nevezzük az optimalizációs feladatot*

$$\min_x \left\{ \hat{c}(x) = \sup_{(c,d,A,b)\in\mathcal{U}} (c^{\mathrm{T}}x + d) : \; Ax \leq b \quad \forall (c, d, A, b) \in \mathcal{U} \right\} \tag{1}$$

ahol a robosztus értéket minimalizáljuk a robosztus megoldásokon.

Ez az úgynevezett legrosszabb esetes (Worst-case) optimalizálás, ami a következő fejezetben, a szkenárió módszernél is elő fog jönni.

Az általánosságot meghagyva, fel lehet tenni, hogy egy $LP_{\mathcal{U}}$ feladat RM-jének a célfüggvénye nem függ a bizonytalan halmaztól, úgyanis az (1) ekvivalens a

$$\min_{x,t} \left\{ t : c^{\mathrm{T}}x + d \leq t, \; Ax \leq b \quad \forall (c, d, A, b) \in \mathcal{U} \right\}$$

feladattal, de ebben már a "t", mint célfüggvény már nem függ a bizonytalansági halmaztól. Tehát az általánosság csorbítása nélkül feltehetjük, hogy egy $LP_{\mathcal{U}}$ bizonytalan lineáris programozási feladat RM-je a következő optimalizációs feladat

$$\min_x \left\{ c^{\mathrm{T}}x + d : Ax \leq b \quad \forall (A, b) \in \mathcal{U} \right\}. \tag{2}$$

(Itt már az $\mathcal{U}$ halmaz már csak az (A, b) feltételek halmaza.) Így kicsit egyszerűbben is írhatjuk az $\mathcal{U}$ halmazt

$$\mathcal{U} = \left\{ (A, b) = (A_0, b_0) + \sum_{\ell=1}^{L} \zeta_\ell (A_\ell, b_\ell) : \; \zeta \in \mathcal{Z} \right\}.$$

Továbbá úgy átalakítható (ekvivalensen) a (2), hogy csak egyetlen egy feltétel függjön a bizonytalansági halmaztól (részletesebben a [3] 16-20. oldalán), azaz a $LP_{\mathcal{U}}$ RM-je a következőre redukálódott:

$$\min_x \left\{ c^{\mathrm{T}}x + d : a^{\mathrm{T}}x \leq b \quad \forall (a,b) \in \mathcal{U} \right\}. \tag{3}$$

Felhasználva az $\mathcal{U}$ alakját, a (3) úgy is írható, hogy

$$\min_x \left\{ c^{\mathrm{T}}x + d : a^{\mathrm{T}}x \leq b \quad \forall \Big((a,b) = (a_0, b_0) + \sum_{\ell=1}^{L} \zeta_\ell (a_\ell, b_\ell) : \ \zeta \in \mathcal{Z} \Big) \right\}.$$

Most nézzük meg mit tudunk mondani, ha $\mathcal{Z}$ egy ellipszoid. Szerepelt, hogy az ellipszoid egy gömb affin képe, így feltehető, hogy $\mathcal{Z}$ egy origó középpontú R sugarú gömb. Így a fenti képlet miatt a feltételeink azok lesznek, hogy

$$\begin{aligned}
& a_0^{\mathrm{T}}x + \sum_{\ell=1}^{L} \zeta_\ell a_\ell^{\mathrm{T}} x \leq b_0 + \sum_{\ell=1}^{L} \zeta_\ell b_\ell & \forall \zeta : ||\zeta||_2 \leq R \\
\Longleftrightarrow\ & \sum_{\ell=1}^{L} \zeta_\ell (a_\ell^{\mathrm{T}} x - b_\ell) \leq b_0 - a_0^{\mathrm{T}} x & \forall \zeta : ||\zeta||_2 \leq R \\
\Longleftrightarrow\ & \max_{||\zeta||_2 \leq R} \left\{ \sum_{\ell=1}^{L} \zeta_\ell (a_\ell^{\mathrm{T}} x - b_\ell) \right\} \leq b_0 - a_0^{\mathrm{T}} x & \\
\Longleftrightarrow\ & R \left(\sqrt{\sum_{\ell=1}^{L} (a_\ell^{\mathrm{T}} x - b_\ell)^2} \right) \leq b_0 - a_0^{\mathrm{T}} x, &
\end{aligned}$$

ami egy explicit konvex feltétel, így meg tudjuk oldani konvex programozási feladatmegoldóval. Ellipszoid bizonytalansági halmaz tipikusan akkor jön elő, amikor egy többdimenziós konfidenciaintervallumot csinálunk egy becsült (pl. maximum likelihood módszer alapján) megoldás körül (részletesebben [6] 12.4.2, 508. oldal). Sajnos nem ilyen egyszerű a helyzet, ha a bizonytalan halmazunk formája nem ilyen szép, ezért lesz nagyon hasznos a szkenárió módszer.

2. Megközelítések

Ebben a fejezetben forrásként használtam az [1] könyvet. A bizonytalansági tényezőnek a halmaza Δ, és egy elemét δ-val fogom jelölni. A következő példák során a minimalizálandó célfüggvény ℓ-lel lesz jelölve, ekkor egy adott δ-ra, és $\nu \in \mathbb{R}^{d-1}$ esetén az $\ell(\nu, \delta)$ értéket akarjuk minimalizálni ν szerint, amit úgy tudunk kifejezni, hogy

$$\min_{\nu \in \mathbb{R}^{d-1}} \ell(\nu, \delta). \tag{4}$$

A fenti (4) képletben viszont nincs jól definiálva , hogy mi δ szerepe az ℓ minimalizálásban, de ezt többféleképpen is lehet értelmezni, hogy mi legyen a szerepe, ami több megközelítéshez vezethet.

2.1. Legrosszabb eset

Az egyik legtermészetesebb ötletünk az lehet, hogy a legrosszabb esetünket akarjuk minimalizálni, tehát ha a "legrosszabb" (Worst-case) δ-t kapjuk a Δ halmazból, akkor azt akarjuk, hogy az $\ell(\nu, \delta)$ érték legyen a lehető legkisebb

$$\min_{\nu \in \mathbb{R}^{d-1}} [\max_{\delta \in \Delta} \ell(\nu, \delta)]. \tag{5}$$

Ezt a modellt használják a robusztus feladatok megoldásánál is, és nagyon elterjedt az irányításelméletben is.

2.2. Átlagos

Most nem a legrosszabb esetre szeretnénk minimalizálni, hanem átlagos szinten (Avarage approach). Ehhez szükségünk van egy valószínűségi mértékre a Δ halmazon, legyen ez $\mathbb{P}$. Erre a $\mathbb{P}$-re többféleképp is nézhetünk, például ez adja meg, hogy milyen eséllyel kapunk meg egy bizonyos δ-t, vagy hogy mennyire nagy szerepet akarunk adni a különféle bizonytalan kimeneteleknek. Mindegy hogyan fogjuk ezt fel, egyszerűen a $\mathbb{P}$ csak súlyozza a δ-kat, és ekkor egy átlagos érték optimalizálási feladatot kapunk:

$$\min_{\nu \in \mathbb{R}^{d-1}} \mathbb{E}_{\Delta}[\ell(\nu, \delta)] = \min_{\nu \in \mathbb{R}^{d-1}} \int_{\Delta} \ell(\nu, \delta) \ \mathrm{d}\mathbb{P}.$$

Az átlagos megközelítéses módszer elég gyakori a sztochasztikus optimalizálásban is, általában azoknál, ahol ugyanaz a döntés egymásután végrehajtódik különböző féltételek mellett.

2.3. Esély-korlátozásos

Nem csak az átlagolásra tudjuk használni a $\mathbb{P}$ valószínűségi mértékünket. Egy alternatív megközelítés az, hogy kihagyunk egy olyan területet a Δ halmazból, aminek ϵ esélye, hogy bekövetkezik, és a maradék $\Delta_\epsilon \subset \Delta$ halmazon akarunk optimalizálni (Chance-constrained approach), aminek a valószínűsége $\mathbb{P}(\Delta_\epsilon) = 1 - \epsilon$, azaz

$$\min_{\nu \in \mathbb{R}^{d-1}, \Delta_\epsilon} [\max_{\delta \in \Delta_\epsilon} \ell(\nu, \delta)]. \tag{6}$$

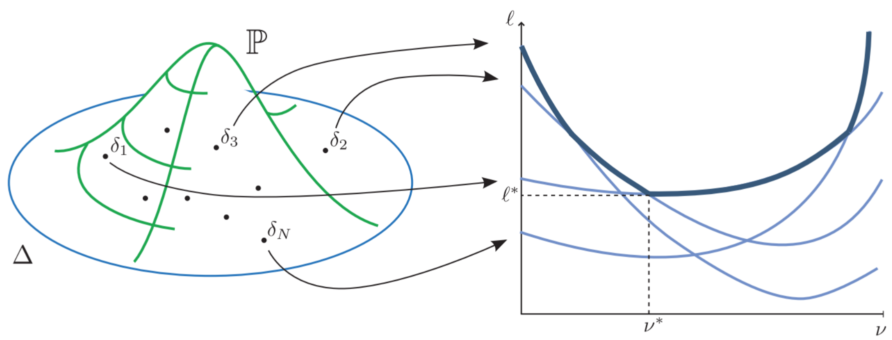
Adott ϵ értékre legyen $(\nu_\epsilon^*, \Delta_\epsilon^*)$ a (6) megoldása, és értéke ℓ_ϵ^*, tehát $\ell_\epsilon^* = \max_{\delta \in \Delta_\epsilon^*} \ell(\nu_\epsilon^*, \delta)$ egy olyan érték, hogy az ℓ függvény a Δ_ϵ halmazon csak kisebb lehet, mint ℓ_ϵ^*. A Δ_ϵ halmazt is úgy választottuk, hogy az előző részekben szereplő legrosszabb eset (5) optimumához képest a legtöbbet bírjuk az optimum értékből lefaragni, tehát az optimalizálás szempontjából a Δ_ϵ halmaz szerepe is nagyon fontos ennél a megközelítésnél.

2.4. Szkenárió módszer

Itt feltesszük, hogy adott δ-ra $\ell(\nu, \delta)$ egy konvex függvény a ν változójában. Ez azért kell, hogy majd tudjunk konvex optimalizálási feladatmegoldót használni. A szkenárió módszer leírása innentől nagyon egyszerű.

Legyen $\delta_1, \delta_2, ..., \delta_N \in \Delta$ minták Δ-ból. Ekkor a következő feladatot akarjuk megoldani:

$$\min_{\nu \in \mathbb{R}^{d-1}} [\max_{i=1,2,...,N} \ell(\nu, \delta_i)]. \tag{7}$$



1. ábra. Szkenárió módszer ábrázolása ([1] 7. oldal)

Ennek a feladatnak a neve SP_N (Scenario Program with N scenarios). Tehát kiveszünk N darab kimenetelt Δ-ból, tehát N darab szkenáriót, és az N darab szkenárió figyelembevételével alkalmazzuk a legrosszabb esetes megközelítésünket (5). Mivel konvex függvények maximuma is konvex, ezért a minimalizálandó függvény is konvex, tehát tényleg alkalmazhatunk konvex optimalizálási feladatmegoldót (7)-re. Megjegyzendő, hogy a δ-k eloszlásából nem használtuk fel semmit, tehát a szkenárió módszer eloszlásfüggetlen. Legyen a fenti SP_N feladat optimális megoldása $\nu^* \in \mathbb{R}^{d-1}$ és értéke ℓ^* $(= \max_{i=1,2,\ldots,N} \ell(\nu^*, \delta_i))$. Természetes kérdés, hogy új δ esetén mennyi az esélye annak, hogy $\ell^* < \ell(\nu^*, \delta)$, tehát hogy rosszul becsültünk, erről ki fog derülni később, hogy az ilyen δ-áknak az eloszlását az úgynevezett Béta-eloszlással bírjuk dominálni.

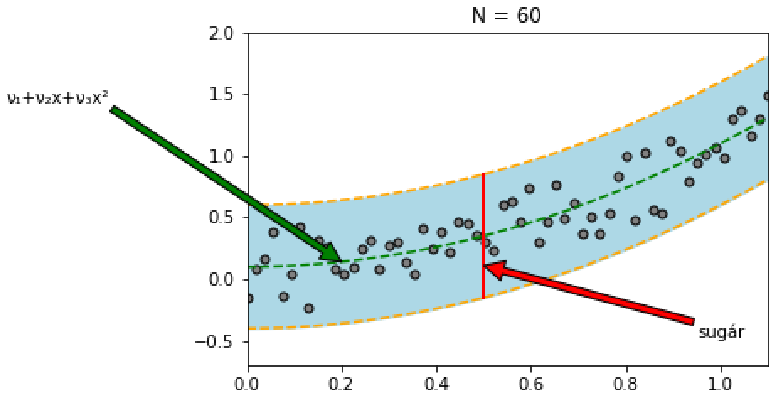
2.4.1. Példa a szkenárió módszer alkalmazására

Az előző fejezetben leírtuk, hogyan működik intuitíven a szkenárió módszer. Most egy példát mutatok rá, ami Python-ban lett megírva.

Legyen $\mathbb{R}^2$-ben N pont, ezek $(x_i, y_i) \in \mathbb{R}^2$ alakúak, és valamilyen eloszlásból kaptuk őket, amit nem ismerünk. Az a feladatunk, hogy megtaláljuk azt a (ν_1, ν_2, ν_3) hármast, amire az ezekkel definiált $y = \nu_1 + \nu_2 x + \nu_3 x^2$ egyenletű másodfokú polinomot körülvevő r sugárú környezetben az összes pont benne legyen. (Arra is kíváncsiak leszünk, hogy ha még generálunk pontokat (ugyanabból az eloszlásból), akkor mennyi fog a mi részünkbe beleesni.)

```
import matplotlib.pyplot as plt
import numpy as np
np.random.seed(20000229)
N = 60
xdata2 = np.random.normal(0,0.6,N)
xdata1 = np.linspace(0,1.1,N)
qdata = np.linspace(0,1.1,100)
plt.plot(xdata1, xdata1**2+0.1+(np.arctan(xdata2))*(1/(np.pi)),
          marker="o", markersize=5, markeredgecolor="black",
          markerfacecolor="grey",linestyle='None')
plt.plot(qdata, qdata**2+0.1, linestyle = '--', color = 'green')
plt.plot(qdata, qdata**2+0.6, linestyle = '--', color = 'orange')
plt.plot(qdata, qdata**2-0.4, linestyle = '--', color = 'orange')
plt.plot(np.linspace(0.5,0.5,20),np.linspace(-0.15,0.85,20),
          color = 'red',  linewidth=2)
plt.fill_between(qdata ,qdata**2+0.6, qdata**2-0.4, color='lightblue')
plt.annotate('\u03BD₁+\u03BD₂x+\u03BD₃x²',
ha = 'center', va = 'bottom',
xytext = (-0.4, 1.4),xy = (0.2, 0.14),arrowprops = {'facecolor' : 'green'})
plt.annotate('sugár',
ha = 'center', va = 'bottom',
xytext = (1, -0.5),xy = (0.5, 0.11),arrowprops = {'facecolor' : 'red'})
plt.xlim([0, 1.1])
plt.ylim([-0.7, 2])
plt.title("N = 60")
ax1 = plt.show()
```

2. ábra. Példához a pontok generálása/ábrázolásának kódja



3. ábra. Illusztráció a példához

Pythonban létezik konvex optimalizálásra feladatmegoldó ($CVXPY$), és ezt fogjuk használni a $\delta_1, \delta_2, \ldots, \delta_N$ N-esre, ahol $\delta_i = (x_i, y_i) \in \mathbb{R}^2$, és $\ell(\nu, \delta_i)$ lesz az (x_i, y_i) pont távolsága az $y = \nu_1 + \nu_2 x + \nu_3 x^2$ egyenletű parabolától.Tehát a (konvex) feladat (7) alapján:

$$\min_{\nu \in \mathbb{R}^{d-1}} [\max_{i=1,2,\ldots,N} \ell(\nu, \delta_i)] = \min_{\nu_1,\nu_2,\nu_3} [\max_{i=1,2,\ldots,N} |y_i - (\nu_1 + \nu_2 x_i + \nu_3 x_i^2)|]. \quad (8)$$

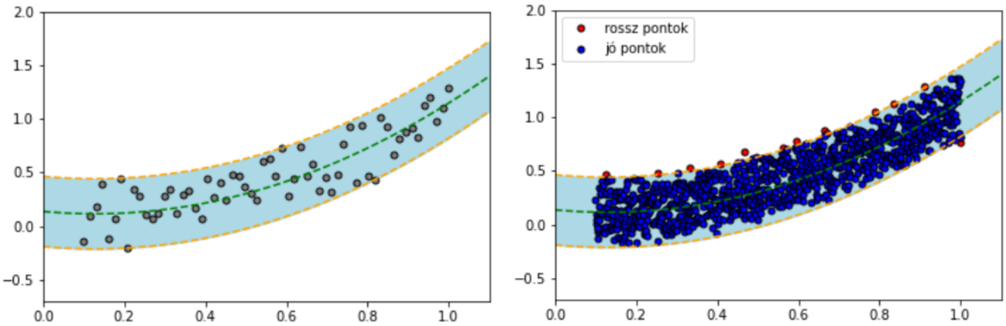
Az (8) feladat megoldása $N = 60$ esetén:

```python
import cvxpy as cp
# Változóink (keressük a 3 együtthatót)
x = cp.Variable(3)
# A feladatunk megfogalmazása
obj = cp.Minimize(cp.maximum(*(cp.abs(xdata1y-(x[0]+x[1]*xdata1+x[2]*xdata1**2)))))
# Megoldjuk
prob = cp.Problem(obj)
prob.solve()  #Visszaadja az optimális m.o-t (ha létezik)
print("státusz:", prob.status)
print("optimális érték(sugár):", prob.value)
print("optimális megoldás:", x.value)
```

```
státusz: optimal
optimális érték(sugár): 0.32692817529570123
optimális megoldás: [ 0.13702416 -0.34600724  1.35612472]
```

4. ábra. $CVXPY$ használata az $N = 60$ esetre

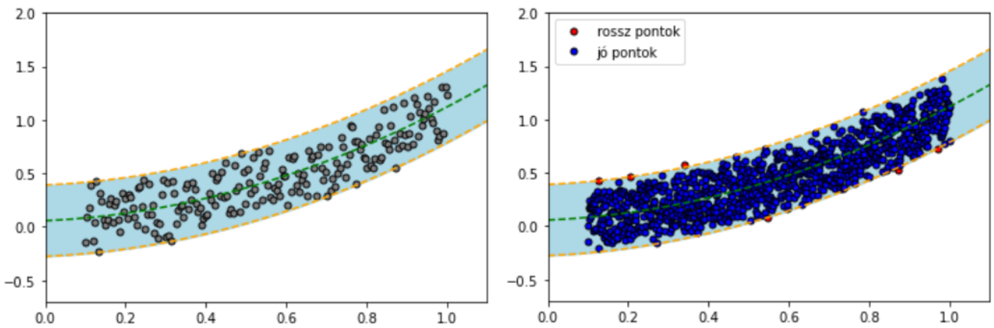
A megoldáshoz tartozó ábra:



5. ábra. $N = 60$-ra optimális megoldás (bal) és az erre tesztelt pontok (jobb)

A megkapott megoldásra 1000 pontot tesztelve (ugyanabból az eloszlásból) azt tapasztalhatjuk, hogy csak 14 pont esett ki az intervallumból, tehát a mostani jelöléseinkkel ez így fogalmazható meg:

$$|\{\delta \in \{\delta_1^{teszt1}, \delta_2^{teszt1}, \ldots, \delta_{1000}^{teszt1}\} : \ell^* < \ell(\nu^*, \delta)\}| = 14.$$



6. ábra. $N = 200$

Ha viszont az előzőeket megcsináljuk $N = 200$-ra, akkor már csak 6 darab ilyen pont esik ki a nekünk megfelelő intervallumból, azaz

$$|\{\delta \in \{\delta_1^{teszt2}, \delta_2^{teszt2}, \ldots, \delta_{1000}^{teszt2}\} : \ell^* < \ell(\nu^*, \delta)\}| = 6.$$

Tehát mintha az N növekedésével nagyobb lenne "garancia" arra, hogy a tesztpontok tényleg az intervallumunkba beleessenek. Az előbbi példával részletesebben foglalkozik a [7].

2.1. Tétel. *(Theorem 1.1, [1] 12. oldal) Minden $\epsilon \in (0,1)$ sértés paraméter (violation parameter) és $\beta \in (0,1)$ konfidencia paraméter (confidence parameter) esetén, ha szkenáriók N számára teljesül, hogy $N \geq (\ln \frac{1}{\beta} + d - 1)$, akkor legalább $1 - \beta$ valószínűséggel*

$$\mathbb{P}(\delta \in \Delta : \ell(\nu^*, \delta) > \ell^*) \leq \epsilon.$$

A következő fejezetben általánosabb formában is kimondjuk ezt a tételt (ami a (3.1) tétel lesz) és speciális feladat fajtára (teljesen támasztott feladatra) be is bizonyítjuk.

3. Matematikai felépítés

Eddig az $\ell(\nu, \delta)$ függvény minimalizálásával foglalkoztunk, ami egy δ bizonytalansági paramétertől is függött. Most egy általánosabb feladatot fogalmazunk meg. Az általánosság csorbítása nélkül feltehetjük, hogy a minimalizálandó célfüggvényünk a $c^T\theta$, ahol $\theta \in \Theta \subset \mathbb{R}^d$. Ez azért van, mert konvex, bizonytalanságtól függő célfüggvények esetét vissza lehet erre vezetni (hasonlóan, mint ahogy az első fejezetben tettük). A feltételek pedig a $\{\Theta_\delta : \delta \in \Delta\}$ halmaz lesz, ahol Θ_δ egy feltétel, ami a δ biztonytalansági paraméterrel van paraméterezve. A Δ pedig legyen a $\mathbb{P}$ valószínűségi mértékkel ellátva. Ekkor a következőképpen néz ki az SP_N feladat:

$$\begin{aligned} SP_N : \quad & \min_{\theta \in \Theta} c^{\mathrm{T}}\theta \\ & \text{ahol} \qquad \theta \in \bigcap_{i=1}^{n} \Theta_{\delta_i}. \end{aligned} \tag{9}$$

Vegyük észre, hogy amit eddig néztünk, az ennek tényleg egy speciális esete, hiszen $\theta = (\nu, x)$, $c = [0, 0, ..., 0, 1]^T$ és $\Theta_{\delta_i} = \{x : x \geq \ell(\nu, \delta_i)\}$ választással éppen az (7)-et kapjuk, mert

$$\min_{\nu \in \mathbb{R}^{d-1}, x \in \mathbb{R}} [0, 0, \ldots, 0, 1]^{\mathrm{T}}(\nu, x) \quad = \min_{\nu \in \mathbb{R}^{d-1}, x \in \mathbb{R}} x \qquad \text{és}$$

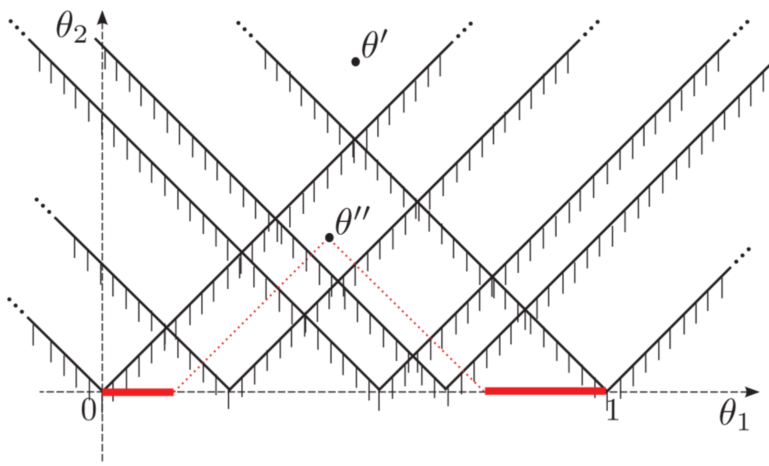
$$x \in \bigcap_{i=1}^{N} \{x : x \geq \ell(\nu, \delta_i)\} \text{ ugyanaz, mint } x \in \{x : x \geq \max_{i=1,2,\ldots,N} \ell(\nu, \delta_i)\}.$$

De x pontosan akkor lesz minimális, amikor a feltételnél (az egyenlőtlenségnél) épp egyenlőség van, tehát az egészből annyi fog megmaradni, hogy $\min_{\nu \in \mathbb{R}^{d-1}} [\max_{i=1,2,\ldots,N} \ell(\nu, \delta_i)]$, ami épp a (7). A továbbiakban jelölje θ^* a (9) megoldását.

3.1. Definíció. *(Sértő halmaz és sértés valószínűsége) Egy adott $\theta \in \Theta$ esetén a $\{\delta \in \Delta : \theta \notin \Theta_\delta\}$ halmazt a θ-hoz tartozó sértő halmaznak (violation set) nevezzük. Ennek a sértő halmaz valószínűsége a sértés valószínűsége (violation probability), azaz $V(\theta) := \mathbb{P}(\{\delta \in \Delta : \theta \notin \Theta_\delta\})$.*

Tehát a sértő halmaz az a Δ egy részhalmaza, és a sértés valószínűsége pedig ennek a halmaznak a valószínűségi mértéke.

3.1. Példa. *Tegyük fel, hogy $\theta \in \mathbb{R}^2$, a feltételek $|\theta_1 - \delta| \leq \theta_2$ alakúak és δ egyenletes eloszlású a $\Delta = [0, 1]$ invervallumon. Tekintsük a $\theta' = (\frac{1}{2}, \theta_2)$ pontot, ahol $\theta_2 \geq \frac{1}{2}$. Ekkor $V(\theta') = 0$, hiszen $|\frac{1}{2} - \delta| > \frac{1}{2}$ semmilyen $\delta \in [0, 1]$-re nem teljesül. Ha viszont a $\theta'' = (\frac{3}{7}, \frac{1}{3})$ pontot nézzük, akkor azok a $\delta \in [0, 1]$,*



7. ábra. Sértő halmaz ábrázolása ([1] 33. oldal)

amikre $|\frac{3}{7} - \delta| > \frac{1}{3}$ teljesül, azok a $\left[0, \frac{2}{21}\right) \cup \left(\frac{16}{21}, 1\right]$ halmazban vannak, tehát

$$V(\theta'') = \mathbb{P}\left(\delta \in \left\{\left[0, \frac{2}{21}\right) \cup \left(\frac{16}{21}, 1\right]\right\}\right) = \frac{2}{21} + \left(1 - \frac{16}{21}\right) = \frac{7}{21}.$$

3.1. Tétel. *(Theorem 1.3, [1] 20.oldal) Minden $\epsilon \in (0, 1)$ sértés paraméter és $\beta \in (0, 1)$ konfidencia paraméter esetén, ha szkenáriók N számára teljesül, hogy $N \geq (ln\frac{1}{\beta} + d - 1)$, akkor legalább $1 - \beta$ valószínűséggel*

$$\mathbb{P}(\delta \in \Delta : \theta^* \notin \Theta_\delta) \leq \epsilon.$$

Ezt a tételt egy későbbi tétel alapján fogjuk bebizonyítani. Viszont ezek előtt, kell két feltétel, hogy ezek teljesülhessenek.

3.1. Feltétel. *(Konvexitás) A Θ és $\Theta_\delta, \delta \in \Delta$ halmazok legyen zárt konvex halmazok.*
Megjegyzés. Ekkor a (9)*-ban szereplő* $\bigcap_{i=1}^{n} \Theta_{\delta_i}$ *halmaz is zárt konvex halmaz, hiszen zárt konvex halmazok metszete is zárt konvex halmaz, tehát a felírt* (9) *feladat konvex feladat.*

3.2. Feltétel. *(Létezés és egyértelműség) Minden $N \in \mathbb{N}$ és minden $(\delta_1, \delta_2, ..., \delta_N) \in \Delta^N$ esetén a* (9) *feladat megoldása létezik, és egyértelmű.*

A Δ^N halmazt (Δ^N az N darab Δ keresztszorzata) lássuk el a $\mathbb{P}^N$ (valószínűségi) szorzatmértékkel, ahol $\mathbb{P}$ a Δ halmazon egy valószínűségi mérték. Ekkor a következő tételt tudjuk kimondani:

3.2. Tétel. *(Theorem 3.7, [1] 39.oldal) Legyen $N \geq d$ és teljesüljenek a* (3.1) *és* (3.2) *feltételek. Ekkor a következő egyenlőtlenség érvényes:*

$$\mathbb{P}^N(V(\theta^*) > \epsilon) \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{10}$$

Ennek a tétel segítségével fogjuk belátni a (3.1) tételt.

Bizonyítás (3.1)**.** A sértő halmaz definíciója miatt a $\mathbb{P}(\delta \in \Delta : \theta^* \notin \Theta_\delta) \leq \epsilon$ egyenlőtlenséget át lehet írni $V(\theta^*) \leq \epsilon$ alakra. A (10) egyenlőtlenséggel pedig ekvivalens az, hogy

$$\mathbb{P}^N(V(\theta^*) \leq \epsilon) \geq 1 - \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^k (1-\epsilon)^{N-i}.$$

Azt fogjuk megmutatni, hogy

$$1 - \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \geq 1 - \beta$$

teljesülésére az $N \geq (ln\frac{1}{\beta} + d - 1)$ elégséges feltétel. A fenti pedig ekvivalens azzal, hogy

$$\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta$$

Ennek az egyenlőtlenségnek a bal oldalát fogjuk átalakítgatni.

$$
\begin{aligned}
\sum_{i=0}^{d-1}\binom{N}{i}\epsilon^i(1-\epsilon)^{N-i} &= 2^{d-1}\sum_{i=0}^{d-1}\binom{N}{i}\frac{\epsilon^i}{2^{d-1}}(1-\epsilon)^{N-i} \\
&\le 2^{d-1}\sum_{i=0}^{d-1}\binom{N}{i}\frac{\epsilon^i}{2^i}(1-\epsilon)^{N-i} && (d-1\ge i) \\
&= 2^{d-1}\sum_{i=0}^{d-1}\binom{N}{i}\left(\frac{\epsilon}{2}\right)^i(1-\epsilon)^{N-i} && (N\ge d \text{ feltétel}) \\
&= 2^{d-1}\left(\frac{\epsilon}{2}+(1-\epsilon)\right)^N \\
&= 2^{d-1}\left(1-\frac{\epsilon}{2}\right)^N \\
&\le 2^{d-1}\exp\Big(-\frac{\epsilon}{2}N\Big).
\end{aligned}
$$

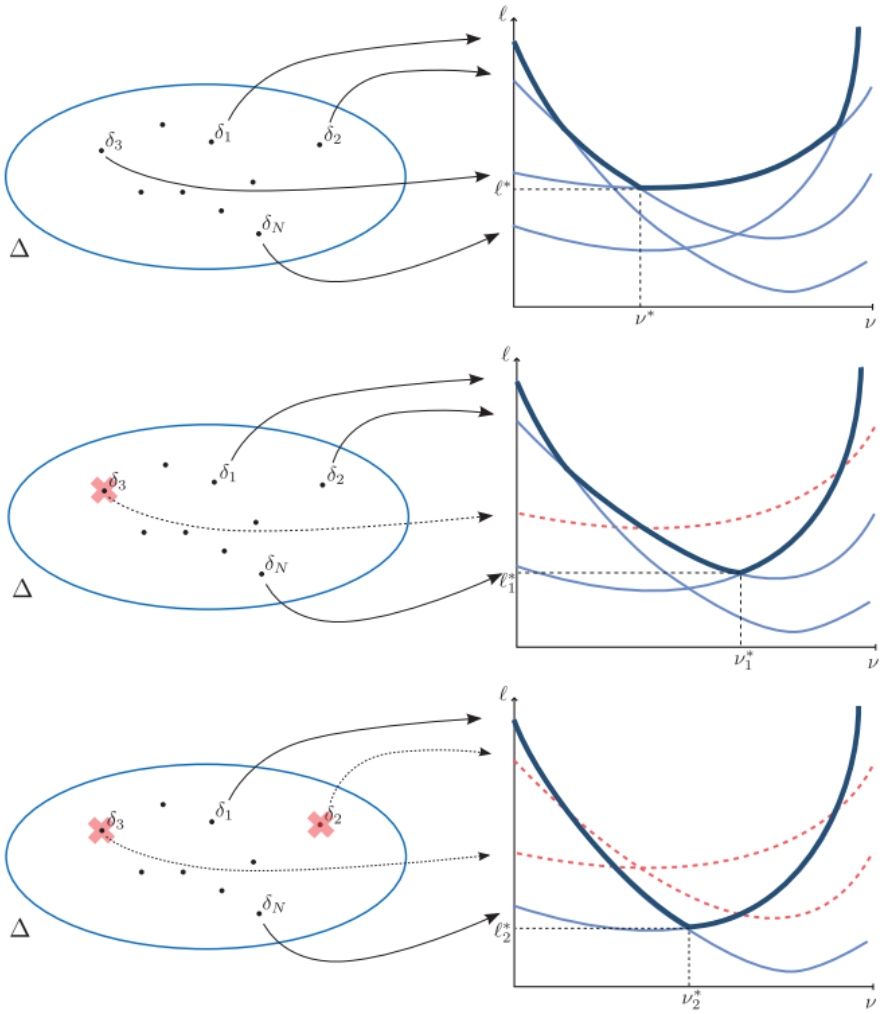
Az utolsó egyenlőtlenség azért teljesül, mert az $\exp(-x)$ függvény konvex és az $x = 0$ pontbeli érintője az $1 - x$, ami a görbe alatt marad, ezért $\exp(-x) \ge 1 - x$, amiből $\exp(-\frac{\epsilon}{2}) \ge 1 - \frac{\epsilon}{2}$. Azaz azt kaptuk, hogy

$$
\begin{gathered}
2^{d-1}\exp\big(-\tfrac{\epsilon}{2}N\big) \le \beta, \text{ebből logaritmust véve} \\
(d-1)\ln 2 - \tfrac{\epsilon}{2}N \le \ln\beta \text{ adódik, ami ekvivalens azzal, hogy} \\
N \ge \tfrac{\epsilon}{2}\big(\ln\tfrac{1}{\beta} + (d-1)\ln 2\big).
\end{gathered}
$$

Az utolsó sornál még azt felhasználva, hogy $\ln 2 > 1$, épp azt kapjuk, hogy $N \ge \frac{\epsilon}{2}\big(\ln\frac{1}{\beta} + d - 1\big)$, amit akartunk. Tehát ha ez teljesül, akkor a (10) is teljesül, azaz tényleg elégséges feltétel volt.

3.1. Szkenárió módszer feltételek elhagyásával

Tekintsük a (7) SP_N feladatot adott $(\delta_1, \delta_2, \dots, \delta_N) \in \Delta^N$ esetén. Ekkor a feladat megfelelően van definiálva, tehát implentálni is tudjuk konvex feladatmegoldóba. Legyen a kapott megoldás szokásos módon ν^* és értéke ℓ^*. Ezek után vizsgáljuk meg az aktív szkenáriókat, azaz az olyan $\delta_i, i \in \{1, 2, \dots, N\}$, amikre $\ell^* = \ell(\nu^*, \delta_i)$. Dobjunk el egy ilyen feltételt, és megint oldjuk meg a feladatot. Az új megoldást követően megint megcsinálhatjuk az előző folyamatot, így kapunk $\nu_1^*, \nu_2^*, \dots$ megoldásokat és az ehhez tartozó $\ell_1^*, \ell_2^*, \dots$ értékek pedig (reményeink szerint) javulni fognak. Jelölje ν_k^* azt a megoldást, ahol k darab szkenáriót dobtunk el és értékét jelölje ℓ_k^*. A következő oldalon lévő ábra próbálja illusztrálni a folyamatot.



8. ábra. δ_2 és δ_3 eldobása([1] 17. oldal)

3.3. Tétel. *(Theorem 1.2, [1] 18.oldal) Legalább* $1-\beta$ *valószínűséggel* ℓ^* *az* ϵ_k *megbízhatósági szintű (risk guaranteed), azaz*

$$\mathbb{P}^N\left[\mathbb{P}(\delta \in \Delta : \ell(\nu^*, \delta) > \ell_k^*) \leq \epsilon_k\right] \leq 1-\beta, \textit{ ahol}$$

$$\epsilon_k = \frac{k}{n} + \left[\frac{\sqrt{k}}{N} + \frac{\sqrt{k}+1}{N}\left((d-1)\ln(k+d+1) + \frac{d-1}{\sqrt{k}} + \ln\frac{1}{\beta}\right)\right].$$

Ennek tétel bebizonyításához szükségünk lesz egy másik tételre, és egy defininícióra. Jelölje θ_k^* azt a megoldást a (9) feladatra, ami pontosan k darab feltételt sért. Ez nem ugyanazt jelenti, hogy k darab feltételt hagytunk el és arra oldottuk meg a feladatot. Ez azért van, mert az elhagyott feltétel az elhagyása után sértetté válik, viszont ezt követően egy másik feltétel elhagyásával megint kielégítetté válhat.

3.4. Tétel. *(Theorem 3.9, [1] 45.oldal) Legyen $N \geq d$. A* (3.1) *és* (3.2) *feltételek mellett a következő teljesül*

$$\mathbb{P}^N(V(\theta_k^*) > \epsilon) \leq \binom{k+d-1}{k} \sum_{i=0}^{k+d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{11}$$

Ezzel fogjuk belátni a (3.3)-at. De előtte vegyük észre, hogy $k = 0$ eset az éppen az, amikor egy feltételt sem sértünk, tehát ez az eredeti (9) megoldása (ami θ^*), és ezért nem meglepő módon visszakapjuk a (3.2) tételt.

Bizonyítás (3.3)**.** A (11) jobb oldalát fogjuk alakítgatni az $\epsilon = \epsilon_k$ szereposztással, és azt akarjuk, hogy a formula az ϵ_k-ra az elégséges feltétel legyen arra, hogy a jobboldal $\leq \beta$ (mert akkor fog az teljesülni, hogy $\mathbb{P}^N(V(\theta^*) \leq \epsilon_k) \geq 1-\beta$). Először a szummás részt írjuk át: Legyen $a = 1+\frac{1}{\sqrt{k}}$.

$$\begin{aligned}
\sum_{i=0}^{k+d-1} \binom{N}{i} \epsilon_k^i (1-\epsilon_k)^{N-i} &= a^{k+d-1} \sum_{i=0}^{k+d-1} \binom{N}{i} \frac{\epsilon_k^i}{a^{k+d-1}} (1-\epsilon_k)^{N-i} \\
&\leq a^{k+d-1} \sum_{i=0}^{k+d-1} \binom{N}{i} \frac{\epsilon_k^i}{a^i} (1-\epsilon_k)^{N-i} \\
&\leq a^{k+d-1} \sum_{i=0}^{k+d-1} \binom{N}{i} \Big(\frac{\epsilon_k}{a}\Big)^i (1-\epsilon_k)^{N-i} \\
&\leq a^{k+d-1} \sum_{i=0}^{N} \binom{N}{i} \Big(\frac{\epsilon_k}{a}\Big)^i (1-\epsilon_k)^{N-i} \\
&= a^{k+d-1} \Big(\frac{\epsilon_k}{a} + (1-\epsilon_k)\Big)^N \\
&= a^{k+d-1} \Big(1 - \Big(1-\frac{1}{a}\Big)\epsilon_k\Big)^N \\
&\leq a^{k+d-1} \exp\Big(-\Big(1-\frac{1}{a}\Big)\epsilon_k N\Big) \\
&\leq \exp(-(1-a)(k+d-1)) \exp\Big(-\Big(1-\frac{1}{a}\Big)\epsilon_k N\Big).
\end{aligned}$$

Az utóbbi 2 egyenlőtlenség pedig azért van, mert $1 - x \leq \exp(-x) \quad \forall x \in \mathbb{R}$. Most a binomiális együtthatót alakítjuk:

$$\binom{k+d-1}{k} = \frac{(k+d-1)!}{k!(d-1)!} \leq (k+d-1)^{d-1},$$

tehát ami tudjuk, hogy a (11) jobboldala kisebb, vagy egyenlő, mint

$$(k+d-1)^{d-1} \exp(-(1-a)(k+d-1)) \exp\Big(-\Big(1-\frac{1}{a}\Big)\epsilon_k N\Big).$$

Legyen az a kifejezés $\leq \beta$, ekkor

$$(k+d-1)^{d-1}\exp(-(1-a)(k+d-1)) \exp\Big(-\Big(1-\frac{1}{a}\Big)\epsilon_k N\Big) \leq \beta,$$

logaritmust véve, és az a értékét beírva kapjuk, hogy

$$(d-1)\ln(k+d-1) + \frac{1}{\sqrt{k}}(k+d-1) - \frac{1}{\sqrt{k}+1}\epsilon_k N \leq \ln\beta.$$

Ezt az egészet ϵ_k-ra rendezve adódik, hogy

$$\epsilon_k \geq \frac{\sqrt{k}+1}{N}\Big[(d-1)\ln(k+d-1) + \frac{1}{\sqrt{k}}(k+d-1) + \ln\frac{1}{\beta}\Big] =$$

$$= \frac{k}{n} + \left[\frac{\sqrt{k}}{N} + \frac{\sqrt{k}+1}{N}\left((d-1)\ln(k+d+1) + \frac{d-1}{\sqrt{k}} + \ln\frac{1}{\beta}\right)\right],$$

amit akartunk.

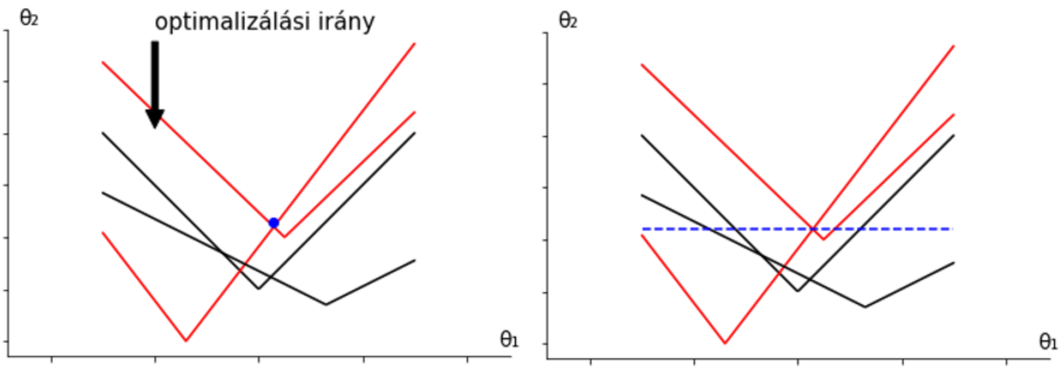
3.2. A 3.2 tétel teljesen támasztott feladatra

Tekintsük azt az SP_N feladatot, ahol $N = m, d = 2$, és $m \geq d$:

$$\begin{aligned} &\min_{\theta \in \Theta} c^{\mathrm{T}}\theta \\ &\text{ahol} \quad \theta \in \bigcap_{i=1}^{m} \Theta_{\delta_i}. \end{aligned} \tag{12}$$

ahol $(\delta_1, \delta_2, \ldots, \delta_m)$ egy független, azonos eloszlású minta a $(\Delta, \mathbb{P})$ valószínűségi mértéktérből és $\Theta \subset \mathbb{R}^d$, azaz $\Theta \subset \mathbb{R}^2$ (így $\theta = (\theta_1, \theta_2)$).

3.2. Definíció. *(Támasz feltétel) Az olyan $\theta \in \Theta_{\delta_i}$ feltételeket amiket elvéve az SP_N feladat megoldása javul, azokat támasz feltételeknek (support contstraints) nevezzük.*



9. ábra. V alakú feltételek esetén a támasz feltétel (piros) illusztrációja

3.5. Tétel. *(Theorem 5.2, [1] 57.oldal) Minden m természetes szám esetén a támasz feltételek száma legfeljebb* 2.

Bizonyítás. A Helly tételt fogjuk használni.
Helly tétel $\mathbb{R}^2$-ben [5]. Legyen adva a síkon véges sok konvex halmaz. Ha bármely három halmaz metszete nemüres, akkor az összes halmaz metszete is nemüres.
A mi esetünkben a feltételeink lesznek a konvex halmazok, azaz a $\theta \in \Theta_{\delta_i}$, $i = 1, 2, \ldots, m$ és ezeken kívül az $m+1$. feltételnek felvesszük azt a konvex halmazt, amikben azok a $\theta \in \mathbb{R}^2$ vannak, amikre $c^{\mathrm{T}}\theta < c^{\mathrm{T}}\theta^*$ (ahol θ^* az a (12) megoldása), azaz a fenti ábrán a jobb oldali képen található kék szaggatott egyenes alatti pontok. Legyen ennek a feltételnek a neve az A feltétel.
Indirekt tegyük fel, hogy a támasz feltételek száma legalább 3. Ekkor, ha nézzük a $\theta \in \Theta_{\delta_i}, \theta \in \Theta_{\delta_j}$ és A feltételeket ($i \neq j$ és $i, j \in \{1, 2, \ldots, m\}$), akkor a legalább 3 támasz feltétel közül legalább 1 nincs ezek között.
A támasz feltétel definíciója miatt, ha a kiválasztott $\theta \in \Theta_{\delta_i}, \theta \in \Theta_{\delta_j}$ feltételekre oldjuk meg a feladatot, akkor javul a megoldásunk, tehát ennek a javított megoldásnak az értéke jobb lesz, mint $c^{\mathrm{T}}\theta^*$, így ez a megoldás az A-ban is benne van, tehát bármely $\theta \in \Theta_{\delta_i}, \theta \in \Theta_{\delta_j}$ és A halmazok esetén a metszet nemüres. Nyilván bármely három $\theta \in \Theta_{\delta_i}, \theta \in \Theta_{\delta_j}, \theta \in \Theta_{\delta_k}$ metszete is nemüres ((3.2) feltétel miatt).
Ezek miatt alkalmazhatjuk a Helly tételt a $\theta \in \Theta_{\delta_i}, i \in \{1, 2, \ldots, m\}$ és A konvex halmazokra, tehát ezek metszete nemüres, ami pedig ellentmond θ^* optimalitásának, hiszen ezek a pontok, amik benne vannak a nemüres metszetben, azok eljesítik a feltételeket, és benne vannak A-ban is, azaz jobb az értékük, mint θ^*-nak.

3.3. Definíció. *(Teljesen támasztott feladat) Egy szkenárió optimalizációs feladatot amiben a költségfüggvény* $c^{\mathrm{T}}\theta$*, a tartomány* Θ*, a feltételek családja* $\{\Theta_\delta : \delta \in \Delta\}$*, a* Δ *halmazon a valószínűségi mérték* $\mathbb{P}$*, teljesen támasztott feladatnak (fully supported problem) nevezzük, ha minden* $m \geq d$ *esetén a* (12) *feladatban a támasz feltételek száma* d*-vel egyenlő* 1 *valószínűséggel.*

Megjegyzés. A támasz feltételek nyilván függnek a $(\delta_1, \delta_2, \ldots, \delta_m)$ m-estől, így az 1 valószínűség az $\mathbb{P}^m$ (valószínűségi) mérték szerint értendő.

3.6. Tétel. *([1] 58.oldal) Teljesen támasztott szkenárió optimalizációs feladatra a következő teljesül (d = 2):*

$$\mathbb{P}^N(V(\theta^*) > \epsilon) = \sum_{i=0}^{1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \tag{13}$$

Az előbbi (13)-ből következik a (10) (ami a (3.2) tétel) a $d = 2$ esetre, hiszen itt egyenlőséget követelünk meg, tehát speciálisan az egyenlőtlenség is teljesül.

Bizonyítás. Először $N = 2$ esetre látjuk be a fenti egyenlőséget, tehát a következőt szeretnénk bizonyítani ekkor:

$$\mathbb{P}^2(V(\theta^*) > \epsilon) = \sum_{i=0}^{1} \binom{2}{i} \epsilon^i (1-\epsilon)^{2-i} = 1 - \epsilon^2. \tag{14}$$

Minden $m \geq 2$ esetén a Δ^m halmazt a osszuk fel a következő eljárással. Vegyünk m elemű $(\delta_1, \delta_2, \ldots, \delta_m)$ mintát és oldjuk meg ezekkel (12) feladatot. Ezekhez 1 valószínűséggel csak kettő támasz feltétel van, legyen ezek indexe i' és j', ahol $i' < j'$. Összesen $\binom{m}{2}$ darab ilyen indexelés van, ahol $i < j$ és $i, j \in \{1, 2, \ldots, m\}$. Jelöljük $\mathcal{L}_{i,j}$-vel egy ilyet, amiben olyan $(\delta_1, \delta_2, \ldots, \delta_m) \in \Delta^m$ m-esek vannak, akikre ha megoldjuk a (12)-es feladatot, akkor a támasz feltételeiknek az indexe i, j (ahol $i < j$ és $i, j \in \{1, 2, \ldots, m\}$). Ekkor teljesül, hogy

$$\mathbb{P}^m\Big(\bigsqcup_{\substack{i<j \\ i,j\in\{1,2,\ldots,m\}}} \mathcal{L}_{i,j}\Big) = 1,$$

(mert a feladat az teljesen támasztott volt) tehát a fenti diszjunkt unió és a Δ^m halmaz csak ($\mathbb{P}^m$ szerinti) nullmértékűhalmazokban lehetnek eltérőek, így (nullmértékű halmazoktól eltekintve) partíciókra bontottuk a Δ^m-et. Jelöljük $V(\theta^*)$ eloszlásfüggvényét F_V-vel amikor $m = 2$, azaz

$$F_V(t) \;= Q_V((-\infty, t]) = \mathbb{P}^2(V(\theta^*) \leq t).$$

Először kiszámoljuk a fenti partíció egy elemének a mértékét, legyen ez az elem az $\mathcal{L}_{1,2}$.

Lemma 1.
$$\mathbb{P}^m(\mathcal{L}_{1,2}) = \int_{[0,1]} (1-x)^{m-2}\, dQ_V(x).$$

Ennek a lemma bizonyításához kell egy definíció, és egy másik lemma.
Jelölje $\overline{\mathcal{L}_{1,2}}$ azt a halmazt, amiben olyan $(\delta_1, \delta_2, \ldots, \delta_m) \in \Delta^m$ elemek vannak, hogy ha csak a $\theta \in \Theta_{\delta_1}, \theta \in \Theta_{\delta_2}$ feltételekre oldjuk meg a (12)-es feladatot, akkor a megoldásunk a $\theta \in \Theta_{\delta_i}, i \in \{3, 4, \ldots, m\}$ feltételeket nem sérti meg.

Lemma 2. Nullmértékű halmazoktól eltekintve $\mathcal{L}_{1,2} = \overline{\mathcal{L}_{1,2}}$.

Lemma 2 bizonyítás. $\mathcal{L}_{1,2} \subseteq \overline{\mathcal{L}_{1,2}}$ (nullmértékű halmazoktól eltekintve) irány. Legyen $(\delta_1, \delta_2, \ldots, \delta_m) \in \mathcal{L}_{1,2}$ és vegyünk el egy $\theta \in \Theta_{\delta_i}, i \in \{3, 4, \ldots, m\}$ feltételt, aztán a maradék $m-1$ feltételre oldjuk meg a (12)-es feladatot. Ekkor a megoldás változatlan marad, hiszen amit eldobtunk az nem volt támasz feltétel (mert tudjuk, hogy kettő darab van, és azok a $\theta \in \Theta_{\delta_1}$ és $\theta \in \Theta_{\delta_2}$), és a támasz feltételek is ugyanazok maradtak (mert egy új feltétel felvételével vagy az új feltétel lesz az egyik támasz feltétel, vagy nem változnak a támasz feltételek, így ha nem ugyanazok maradtak volna, akkor az eldobott feltétel visszavételével nem kaphatnánk vissza azt, hogy a $\theta \in \Theta_{\delta_1}$ és $\theta \in \Theta_{\delta_2}$ a támasz feltételek az eredeti feladatban). Ezt az eldobásos folyamatot folytatva eljutunk addig, hogy csak a kettő támasz feltételre oldjuk meg a feladatot és a megoldásunk változatlan marad. Így azt kaptuk, hogy ugyanaz a megoldásunk ha a $(\delta_1, \delta_2, \ldots, \delta_m)$ -re oldjuk meg a feladatot, vagy ha csak a (δ_1, δ_2) -re, amiből következik, hogy a $\theta \in \Theta_{\delta_i}, i \in \{3, 4, \ldots, m\}$ feltételek közül semelyik sem sérül meg amikor csak a (δ_1, δ_2) -re oldjuk meg a feladatunkat, így $(\delta_1, \delta_2, \ldots, \delta_m) \in \overline{\mathcal{L}_{1,2}}$.

$\overline{\mathcal{L}_{1,2}} \subseteq \mathcal{L}_{1,2}$ (nullmértékű halmazoktól eltekintve) irány.
Legyen $(\delta_1, \delta_2, \ldots, \delta_m) \in \overline{\mathcal{L}_{1,2}}$, azaz a $\theta \in \Theta_{\delta_i}, i \in \{3, 4, \ldots, m\}$ feltételeket nem sérti meg a megoldásunk, amikor csak a (δ_1, δ_2)-re oldjuk meg a feladatot. Ezért ha a feltételek közé vesszük ezek valamelyikét, akkor nem változik meg a megoldásunk és a felvett feltétel sem válik támasz feltétellé (mert ha azzá válna, akkor elhagyásával javítanánk a megoldáson, ami ellentmondás, hiszen az változatlan maradt). Ebből következik, hogy $\theta \in \Theta_{\delta_i}, i \in \{3, 4, \ldots, m\}$ feltételek közül egyik sem lehet támasz feltétel, így a $(\delta_1, \delta_2, \ldots, \delta_m)$ nem lehet

benne olyan $\mathcal{L}_{i,j}$ halmazokban, ahol $\mathcal{L}_{i,j} \neq \mathcal{L}_{1,2}$. Tehát kaptuk, hogy

$$\overline{\mathcal{L}_{1,2}} \subseteq \Big(\bigsqcup_{\substack{i<j \\ i,j \in \{1,2,\ldots,m\}}} \mathcal{L}_{i,j} \Big)^c = \mathcal{L}_{1,2},$$

ahol A^c az A halmaz komplementerét jelöli és az egyenlőséget nullmértékű halmazoktól eltekintve értjük. Ezzel beláttuk a Lemma 2-t.

Lemma 1 bizonyításához szükségünk lesz mértékelméletből egy integrál helyettesítési formulára.

3.7. Tétel. *Legyenek (X, μ_X) és (Y, μ_Y) mértékterek, $h : X \to Y$ mérhető függvény, $\mu_X : X \to [0, +\infty]$ és jelölje $h_*\mu_X$ a μ_X mérték előretoltját h szerint. Egy g függvény Y integrálható a $h_*\mu_X$ előretolt mérték szerint akkor és csak akkor, ha $g \circ h$ integrálható X halmazon a μ_X mérték szerint. Ekkor az integrálok meg is egyeznek:*

$$\int_Y g(y)\ dh_*\mu_X(y) = \int_X (g \circ h)(x)\ d\mu_X(x).$$

Lemma 1 bizonyítás. $\mathbb{P}^m(\mathcal{L}_{1,2})$ kiszámításához már elég az is, ha $\mathbb{P}^m(\overline{\mathcal{L}_{1,2}})$ értéket számoljuk ki, hiszen beláttuk, hogy $\mathcal{L}_{1,2}$ és $\overline{\mathcal{L}_{1,2}}$ halmazok csak nullmértékű halmazokban térhetnek el, tehát $\mathbb{P}^m(\mathcal{L}_{1,2}) = \mathbb{P}^m(\overline{\mathcal{L}_{1,2}})$. Tekintsük a $(\delta_1', \delta_2', \delta_3, \ldots, \delta_m) \in \Delta^m$, ahol az első kettő elem fixálva van, és jelölje $\theta^*_{1,2}$ annak a feladatnak a megoldását, ahol csak a $\theta \in \Theta_{\delta_1'}$ és $\theta \in \Theta_{\delta_2'}$ feltételekre oldottuk meg. Annak a valószínűsége, hogy $(\delta_1', \delta_2', \delta_3, \ldots, \delta_m) \in \overline{\mathcal{L}_{1,2}}$ az éppen annyi, hogy a maradék $m-2$ darab feltétel nem sérül, azaz $(1 - V(\theta^*_{1,2}))^{m-2}$. Ezeket felhasználva:

$$\begin{aligned}
\mathbb{P}^m(\mathcal{L}_{1,2}) &= \mathbb{P}^m(\overline{\mathcal{L}_{1,2}}) \\
&= \int_{\Delta^2} (1 - V(\theta^*_{1,2}))^{m-2}\ \mathrm{d}\mathbb{P}^2(\delta_1', \delta_2') \\
&= \int_{[0,1]} (1-x)^{m-2}\ \mathrm{d}Q_V(x),
\end{aligned} \tag{15}$$

ahol az utolsó egyenlőségnél az előző (3.7) tételt használtuk fel az $X = \Delta^2$, $Y = [0,1]$, $h = V$, $\mu_X = \mathbb{P}^2$, $h_*\mu_X = Q_V$ és $g(x) = (1-x)^{m-2}$ szereposztással, így beláttuk a Lemma 1 -et is.

Most kiszámoljuk a $\mathbb{P}^m(\mathcal{L}_{1,2})$ értéket egy másik módszerrel is. Vegyük észre, hogy minden $\mathcal{L}_{i,j}$ halmaznak ugyanannyi a valószínűsége, hiszen minden

$(\delta_1, \delta_2, \ldots, \delta_m)$ esetén a sorrendet azt tetszőlegesen permutálhatjuk. Ezért, felhasználva, hogy az $\mathcal{L}_{i,j}$ halmazok a Δ^m halmazt a nullmértékű halmazoktól eltekintve partíciókra bontják, kapjuk, hogy:

$$\sum_{\substack{i<j \\ i,j\in\{1,2,\ldots,m\}}} \mathbb{P}^m(\mathcal{L}_{i,j}) = 1 \quad \Longrightarrow \binom{m}{2}\mathbb{P}^m(\mathcal{L}_{i,j}) = 1 \quad \Longrightarrow \mathbb{P}^m(\mathcal{L}_{i,j}) = \frac{1}{\binom{m}{2}}.$$

Ezt összevetve a (15)-tel:

$$\mathbb{P}^m(\mathcal{L}_{i,j}) = \frac{1}{\dbinom{m}{2}} = \int_{[0,1]} (1-x)^{m-2}\, \mathrm{d}Q_V(x). \tag{16}$$

A fenti egyenlőség minden $m \geq 2$ esetén értelmezve van. Ez egy Hausdorff-féle momentum feladat [2], amiből mi csak annyit használunk fel, hogy a momentum feladatoknál kompakt intervallumon (tehát ha a valószínűségi változó tartója kompakt) ha a megoldás létezik, akkor az egyértelmű. A $[0,1]$ halmaz kompakt, tehát ha a megoldható, akkor a megoldás egyértelmű.

3.8. Tétel. *(Theorem 7, [2] 295. oldal). Legyen $F = F(x)$ egy eloszlásfüggvény és $\mu_n = \int_{-\infty}^{\infty} |x|^n\, dF(x)$. Ha*

$$\limsup_{n\to\infty} \frac{\mu_n^{1/n}}{n} < \infty$$

akkor a $\{k_n\}_{n\geq 0}$ momentumok, ahol $k_n = \int_{-\infty}^{\infty} x^n\, dF(x)$ egyértelműen meghatározzák az $F = F(x)$ eloszlásfüggvényt.

Megjegyzések. Az előző tételben, ha a $\mu_n = \int_{-\infty}^{\infty} |x|^n\, dF(x)$ helyett a $\mu_n = \int_0^1 |x|^n\, dF(x)$ nézünk (tehát a valószínűségi változónk tartója a $[0,1]$), akkor a tétel feltétele teljesül, hiszen

$$\mu_n = \int_0^1 |x|^n\, dF(x) \leq \int_0^1 |1|^n\, dF(x) = 1, \text{ ezért}$$

$$\limsup_{n\to\infty} \frac{\mu_n^{1/n}}{n} \leq \limsup_{n\to\infty} \frac{1^{1/n}}{n} = 0 < \infty.$$

De ez igaz minden $K \subset \mathbb{R}$ kompakt halmazra, hiszen ehhez a K-hoz létezik $M \in \mathbb{N}$, hogy $K \subset [-M, M]$, ezért

$$\mu_n = \int_K |x|^n\, dF(x) \leq \int_{-M}^{M} |x|^n\, dF(x) \leq \int_{-M}^{M} |M|^n\, dF(x) = M^n, \text{ emiatt}$$

$$\limsup_{n\to\infty} \frac{\mu_n^{1/n}}{n} \leq \limsup_{n\to\infty} \frac{(M^n)^{1/n}}{n} = \limsup_{n\to\infty} \frac{M}{n} = 0 < \infty.$$

Azaz minden kompakt halmazon a momentum feladat megoldása egyértelmű. Ez nekünk majdnem jó, hiszen a V valószínűségi változó tartója a $[0,1]$ halmaz, viszont a (16)-as egyenlet jobb oldalán nem a momentmumok vannak, hanem azoknak egy lineáris kombinánciója minden m-re. A jó hír az, hogy kifejezhető belőlük az összes momentum, így alkalmazható a fenti tétel.

$$\frac{1}{\binom{m}{2}} = \int_{[0,1]} (1-x)^{m-2} \,\mathrm{d}Q_V(x)$$

$$\begin{aligned}
m = 2 \implies \frac{1}{\binom{2}{2}} &= \int_0^1 1 \,\mathrm{d}F_V(x) = k_0 \implies k_0 = \frac{1}{\binom{2}{2}} \\
m = 3 \implies \frac{1}{\binom{3}{2}} &= \int_0^1 1 - x \,\mathrm{d}F_V(x) = k_0 - k_1 \implies k_1 = k_0 - \frac{1}{\binom{3}{2}} \\
m = 4 \implies \frac{1}{\binom{4}{2}} &= \int_0^1 (1-x)^2 \,\mathrm{d}F_V(x) = k_0 - 2k_1 + k_2 \\
&\implies k_2 = \frac{1}{\binom{4}{2}} - k_0 + 2k_1 \\
&\vdots \\
m = n+2 \implies \frac{1}{\binom{n+2}{2}} &= \int_0^1 (1-x)^n \,\mathrm{d}F_V(x) = \sum_{i=0}^{n} \binom{n}{i} (-1)^i k_i \\
&\implies (-1)^n k_n = \frac{1}{\binom{n+2}{2}} - \sum_{i=0}^{n-1} \binom{n}{i} (-1)^i k_i
\end{aligned}$$

Minden k_n esetén csak a az olyan k_i momentumokat szükséges kiszámítani, ahol $i \in \{0, 1, 2, \ldots, n-1\}$. Mivel a $k_0 = 1$, ezért minden n-re k_n értéket ki tudjuk számolni, és így már látjuk, hogy tényleg a momentum feladattal ekvivalens a (16).

A $Q_V((-\infty, t]) = F_V(t) = t^2$ megoldása a (16)-as egyenletnek, hiszen felhasználva, hogy $(F_V(t))' = 2t$ adódik, hogy:

$$\int_{[0,1]} (1-x)^{m-2} \,\mathrm{d}Q_V(x) = \int_{[0,1]} 2x(1-x)^{m-2} \,\mathrm{d}\lambda(x).$$

Mivel kompakt intervallumon vagyunk és az integrálandó függvényünk korlátos rajta, ezért ha a Riemann integrál létezik, akkor az egyenlő a fenti Lebesgue integrállal.

$$\int_0^1 2x(1-x)^{m-2}\,\mathrm{d}x = \left[\frac{-2x(1-x)^{m-1}}{m-1}\right]_0^1 - \int_0^1 \frac{-2(1-x)^{m-1}}{m-1}\,\mathrm{d}x$$

$$= \frac{2}{m-1}\int_0^1 (1-x)^{m-1}\,\mathrm{d}x \qquad (\text{ahol } \ m \geq 2)$$

$$= \frac{2}{m(m-1)} = \frac{1}{\binom{m}{2}}.$$

Ezért a $Q_V((-\infty,t]) = F_V(t) = t^2$ valóban megoldása a (16) egyenletnek (és ez egyértelmű is a (3.8) tétel alatti megjegyzés miatt). Így az F_V definiciójába a megkapott értéket beírva kapjuk, hogy

$$F_V(t) = \mathbb{P}^2(V(\theta^*) \leq t) = t^2, \quad t := \epsilon \quad \text{helyettesítve}$$
$$F_V(\epsilon) = \mathbb{P}^2(V(\theta^*) \leq \epsilon) = \epsilon^2, \ \text{ ami ekvivalens azzal, hogy}$$
$$\mathbb{P}^2(V(\theta^*) > \epsilon) = 1-\epsilon^2, \ \text{ ami épp a (14).}$$

Most már mindenünk készen áll ahhoz, hogy általános N-re lássuk be a (3.6) tételt. Tekintsünk egy teljesen támasztott (12)-es feladatot, $m = N$-re, ahol $N \geq 2$ tetszőleges. Mivel a feladat teljesen támasztott, így megint partíciókra tudjuk bontani a Δ^N halmazt az $\mathcal{L}_{i,j}$ halmazok szerint. Most csak azt a részét nézzük a Δ^N halmaznak, ami a $\{(\delta_1,\delta_2,\ldots,\delta_N) \in \Delta^N : \ \ V(\theta^*) > \epsilon\}$ elemekből áll. A partícionálás miatt a következő teljesül (felhasználva az előző számolásokat és hogy minden $\mathcal{L}_{i,j}$ valószínűsége megegyezik):

$$\mathbb{P}^N(V(\theta^*) > \epsilon) = \mathbb{P}^N\Bigg(\bigsqcup_{\substack{i<j\\ i,j\in\{1,2,\ldots,N\}}} \Big(\mathcal{L}_{i,j} \cap (V(\theta^*) > \epsilon)\Big)\Bigg)$$
$$= \binom{N}{2}\int_{\Delta^2} (1-V(\theta^*_{1,2}))^{N-2}\mathbb{1}_{\{V(\theta^*_{1,2})>\epsilon\}}\,\mathrm{d}\mathbb{P}^2(\delta_1',\delta_2')$$
$$= \binom{N}{2}\int_{[0,1]} (1-x)^{N-2}\mathbb{1}_{\{x>\epsilon\}}\,\mathrm{d}Q_V(x)$$
$$= \binom{N}{2}\int_{[0,1]} 2x(1-x)^{N-2}\mathbb{1}_{\{x>\epsilon\}}\,\mathrm{d}\lambda(x) =$$

$$
\begin{aligned}
&= \binom{N}{2} \int_{(\epsilon,1]} 2x(1-x)^{N-2}\,\mathrm{d}\lambda(x) \\
&= \binom{N}{2} \int_{\epsilon}^{1} 2x(1-x)^{N-2}\,\mathrm{d}x \\
&= \binom{N}{2} \left(\left[\frac{-2x(1-x)^{N-1}}{N-1} \right]_{\epsilon}^{1} - \int_{\epsilon}^{1} \frac{-2(1-x)^{N-1}}{N-1}\,\mathrm{d}x \right) \\
&= \frac{2}{N-1}\binom{N}{2} \left(\epsilon(1-\epsilon)^{N-1} + \frac{(1-\epsilon)^N}{N} \right) \\
&= (1-\epsilon)^N + N\epsilon(1-\epsilon)^{N-1} \\
&= \sum_{i=0}^{1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \quad \text{amit be akartunk látni.}
\end{aligned}
$$

(A fenti levezetésben $\mathbb{1}_A$ az a A halmaz indikátorfüggvényét jelöli.) Ezzel a (3.6) tétel bizonyítása véget ért, és speciálisan a (3.2) tételt is beláttuk a $d = 2$ esetre, teljesen támasztott feladat esetén.

Általános $d \geq 2$-re hasonlóan járunk el. Jelölje $\mathcal{L}_{i_1,i_2,\ldots,i_d}$ azt a halmazt, amiben az olyan $(\delta_1, \delta_2, \ldots, \delta_N) \in \Delta^N$ elemek vannak, akiknek a támasz feltételeiknek az indexe $i_1, i_2, \ldots, i_d$, ahol $i_1 < i_2 < \ldots < i_d$ és $i_j \in \{1, 2, 3 \ldots, N\}$ $\forall j \in \{1, 2, \ldots, d\}$. Ekkor megint az $\mathcal{L}_{i_1,i_2,\ldots,i_d}$ halmazok partíciókra bontják Δ^N halmazt (nullmértékű halmazoktól eltekintve). A $d = 2$ esetben elhangzottak miatt megint felírhatjuk a következőt:

$$
\mathbb{P}^N(\mathcal{L}_{1,2,\ldots,d}) = \frac{1}{\binom{N}{d}} = \int_{\Delta^d} (1 - V(\theta^*_{1,2,\ldots,d}))^{N-d} \mathrm{d}\mathbb{P}^d(\delta'_1, \delta'_2, \ldots, \delta'_d). \tag{17}
$$

Ez a (17)-es egyenletet Q_V mérték szerinti Lebesgue integrálra átírva megint egy momentum feladatot kapunk kompakt (jelen esetben a $[0, 1]$) intervallumra. Jelölje $F_V(t) = \mathbb{P}^d(V(\theta^*) \leq t)$ valószínűséget és vegyük észre, hogy $Q_V((-\infty, t]) = F_V(t) = t^d$ megoldja a (17)-es egyenletet:

$$
\begin{aligned}
\mathbb{P}^N(\mathcal{L}_{1,2,\ldots,d}) &= \int_{\Delta^d} (1 - V(\theta^*_{1,2,\ldots,d}))^{N-d}\,\mathrm{d}\mathbb{P}^d(\delta'_1, \delta'_2, \ldots, \delta'_d) \\
&= \int_{[0,1]} (1-x)^{N-d}\,\mathrm{d}Q_V(x) \\
&= \int_{[0,1]} dx^{d-1}(1-x)^{N-d}\,\mathrm{d}\lambda(x) \\
&= \int_0^1 dx^{d-1}(1-x)^{N-d}\,\mathrm{d}x =
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\frac{-dx^{d-1}(1-x)^{N-d+1}}{N-d+1}\right]_0^1 - \int_0^1 \frac{-d(d-1)x^{d-2}(1-x)^{N-d+1}}{N-d+1}\,\mathrm{d}x \\
&= \frac{d(d-1)}{N-d+1}\int_0^1 x^{d-2}(1-x)^{N-d+1}\,\mathrm{d}x \\
&= \frac{d(d-1)(d-2)}{(N-d+1)(N-d+2)}\int_0^1 x^{d-3}(1-x)^{N-d+2}\,\mathrm{d}x \\
&\quad\vdots \\
&= \frac{d!}{(N-d+1)(N-d+2)\cdot\ldots\cdot(N-1)}\int_0^1 (1-x)^{N-1}\,\mathrm{d}x \\
&= \frac{d!}{(N-d+1)(N-d+2)\cdot\ldots\cdot(N-1)N} = \frac{1}{\binom{N}{d}}.
\end{aligned}
$$

Tehát tényleg jó az $F_V(t) = t^d$ választás. Mindenünk kész, hogy általános d-re kiszámoljuk pontosan a $\mathbb{P}^N(V(\theta^*) > \epsilon)$ valószínűséget. A számolás lényegében ugyanaz, mint a $d = 2$ esetben:

$$
\begin{aligned}
\mathbb{P}^N(V(\theta^*) > \epsilon) &= \mathbb{P}^N\Bigg(\bigsqcup_{\substack{i_1<i_2<\ldots<i_d \\ i_1,i_2,\ldots,i_d\in\{1,2,\ldots,N\}}} \Big(\mathcal{L}_{i_1,I_2,\ldots,i_d} \cap (V(\theta^*) > \epsilon)\Big)\Bigg) \\
&= \binom{N}{d}\int_{\Delta^d} (1 - V(\theta^*_{1,2,\ldots,d}))^{N-d}\mathbb{1}_{\{V(\theta^*_{1,2,\ldots,d})>\epsilon\}}\,\mathrm{d}\mathbb{P}^d(\delta'_1,\delta'_2,\ldots,\delta'_d) \\
&= \binom{N}{d}\int_{[0,1]} (1-x)^{N-d}\mathbb{1}_{\{x>\epsilon\}}\,\mathrm{d}Q_V(x) \\
&= \binom{N}{d}\int_{[0,1]} dx^{d-1}(1-x)^{N-d}\mathbb{1}_{\{x>\epsilon\}}\,\mathrm{d}\lambda(x) \\
&= \binom{N}{d}\int_{(\epsilon,1]} dx^{d-1}(1-x)^{N-d}\,\mathrm{d}\lambda(x) \\
&= d\binom{N}{d}\int_\epsilon^1 x^{d-1}(1-x)^{N-d}\,\mathrm{d}x \\
&= d\binom{N}{d}\left(\left[\frac{-x^{d-1}(1-x)^{N-d+1}}{N-d+1}\right]_\epsilon^1 - \int_\epsilon^1 \frac{-(d-1)x^{d-2}(1-x)^{N-d+1}}{N-d+1}\,\mathrm{d}x\right) \\
&= \binom{N}{d-1}\epsilon^{d-1}(1-\epsilon)^{N-d+1} + (d-1)\binom{N}{d-1}\int_\epsilon^1 x^{d-2}(1-x)^{N-d+1}\,\mathrm{d}x \\
&\quad\vdots
\end{aligned}
$$

$$
\begin{aligned}
&\vdots \\
&= \left(\sum_{i=1}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \right) + \binom{N}{1} \int_\epsilon^1 (1-x)^{N-1} \, \mathrm{d}x \\
&= \left(\sum_{i=1}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \right) + (1-\epsilon)^N = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}.
\end{aligned}
$$

Ezzel beláttuk azt, hogy minden $d \geq 2$ egész szám és $N \geq d$ egész szám esetén a teljesen támasztott feladatoknál teljesül, hogy:

$$
\mathbb{P}^N(V(\theta^*) > \epsilon) = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \tag{18}
$$

aminek következménye a (3.2) tétel teljesen támasztott feladatok esetén.

3.3. A 3.4 tétel teljesen támasztott feladatra

Csak $d = 2$-re bizonyítjuk a (3.4) tételt, viszont általános d-re ugyanez a gondolatmenet. Legyen $I = \{i_1, i_2, \dots, i_k\}$ olyan indexhalmaz, aminek k darab eleme van az $\{1, 2, \dots, N\}$ halmazból, és legyen θ_I^* a következő szkenárió feladat megoldása:

$$\begin{aligned} &\min_{\theta\in\Theta} c^{\mathrm{T}}\theta \\ &\text{ahol} \quad \theta \in \bigcap_{i\in(\{1,2,\dots,N\}\setminus I)} \Theta_{\delta_i} \end{aligned} \tag{19}$$

Teljesen támasztott feladatot feltételezünk, mert ekkor tudjuk a korábbi eredményt felhasználni. Legyen $\Delta_I^N := \{(\delta_1, \delta_2, \dots, \delta_N) \in \Delta^N : \forall j \in \{1, 2, \dots, k\} \quad \theta_I^* \notin \Theta_{\delta_{i_j}}\}$, azaz a Δ_I^N halmaz olyan $(\delta_1, \delta_2, \dots, \delta_N) \in \Delta^N$ elemekből áll, amikből ha kivesszük az I halmaz elemeit és a maradékra oldjuk feladatot (tehát a (19)-ös feladatot), akkor a θ_I^* megoldás sérti az összes kivett feltételt. A θ_k^* egy olyan megoldás volt, ami k darab feltételt sért 1 valószínűséggel, így az előző definíciók miatt létezik egy I, hogy $\theta_k^* = \theta_I^*$, ahol $(\delta_1, \delta_2, \dots, \delta_N) \in \Delta_I^N$. Ebből következik, hogy:

$$\{(\delta_1, \delta_2, \dots, \delta_N) \in \Delta^N : V(\theta_k^*) > \epsilon\} \subseteq \bigcup_{I\in\mathcal{A}} \{(\delta_1, \delta_2, \dots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\},$$

ahol $\mathcal{A}$ halmaz tartalmazza a $\{1, 2, \dots, N\}$ halmaz összes k elemű részhalmazát. (A fenti relációt nullmértékű halmazoktól eltekintve értjük.) Legyen $I \in \mathcal{A}$ egy adott indexhalmaz, ahol $I = \{i_1, i_2, \dots, i_k\}$ és számoljuk ki a fenti szummában az ehhez az I-hez tartozó tagnak a valószínűségét. Jelölje F_V a $V(\theta_I^*)$ eloszlásfüggvényét, majd felhasználva, hogy a δ_i-k függetlenek:

$$\mathbb{P}^N\big((\delta_1, \delta_2, \dots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\big) = \int_{(\epsilon,1]} \mathbb{P}^N\big(\Delta_I^N |\, V(\theta_I^*) = x\big)\, \mathrm{d}F_V(x)$$

$$\begin{aligned} &= \int_{(\epsilon,1]} \mathbb{P}^N\big(\theta_I^* \notin \Theta_{\delta_{i_1}}, \theta_I^* \notin \Theta_{\delta_{i_2}}, \dots, \theta_I^* \notin \Theta_{\delta_{i_k}} |\, V(\theta_I^*) = x\big)\, \mathrm{d}F_V(x) \\ &= \int_{(\epsilon,1]} \left(\prod_{j=1}^{k} \mathbb{P}^N(\theta_I^* \notin \Theta_{\delta_{i_j}} |\, V(\theta_I^*) = x)\right) \mathrm{d}F_V(x) \\ &= \int_{(\epsilon,1]} \left(\prod_{j=1}^{k} x\right) \mathrm{d}F_V(x) = \int_{(\epsilon,1]} x^k\, \mathrm{d}F_V(x). \end{aligned}$$

Az utolsó előtti egyenlőségénél azt felhasználva, hogy definíció szerint

$V(\theta_I^*) = x \iff \mathbb{P}(\{\delta \in \Delta : \theta_I^* \notin \Theta_\delta\}) = x$, amiből $\mathbb{P}^N(\theta_I^* \notin \Theta_{\delta_{i_j}} \mid V(\theta_I^*) = x) = x$. Mivel F_V az egy olyan megoldás eloszlásfüggvénye, ami $N-k$ darab feltétellel jött létre, ezért $F_V(x)$ értéke a (18) miatt a $d = 2$ esetre:

$$F_V(x) = \mathbb{P}^N(V(\theta_I^*) \le x) = 1 - \sum_{i=0}^{1} \binom{N-k}{i} x^i (1-x)^{N-k-i}$$

$$\implies \quad (F_V(x))' = 2\binom{N-k}{2} x(1-x)^{N-k-2}.$$

Így már megvan minden, már csak számolni kell:

$$\begin{aligned}
&\mathbb{P}^N\big((\delta_1, \delta_2, \ldots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\big) = \int_{(\epsilon,1]} x^k \, \mathrm{d}F_V(x) \\
&= \int_{(\epsilon,1]} 2x^k \binom{N-k}{2} x(1-x)^{N-k-2} \, \mathrm{d}\lambda(x) \\
&= 2\binom{N-k}{2} \int_\epsilon^1 x^{k+1}(1-x)^{N-k-2} \, \mathrm{d}x \\
&= 2\binom{N-k}{2} \left(\left[\frac{-x^{k+1}(1-x)^{N-k-1}}{N-k-1} \right]_\epsilon^1 + \int_\epsilon^1 \frac{(k+1)x^k(1-x)^{N-k-1}}{N-k-1} \, \mathrm{d}x \right) \\
&= \frac{2\binom{N-k}{2}}{N-k-1} \epsilon^{k+1}(1-\epsilon)^{N-k-1} + \frac{2\binom{N-k}{2}(k+1)}{N-k-1} \int_\epsilon^1 x^k(1-x)^{N-k-1} \, \mathrm{d}x \\
&\qquad\qquad\vdots \\
&= \frac{2\binom{N-k}{2}}{(k+2)\binom{N}{k+2}} \sum_{i=0}^{k+1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}.
\end{aligned}$$

Tudjuk, hogy

$$\{(\delta_1, \delta_2, \ldots, \delta_N) \in \Delta^N : V(\theta_k^*) > \epsilon\} \subseteq \bigcup_{I \in \mathcal{A}} \{(\delta_1, \delta_2, \ldots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\},$$

és mivel $\binom{N}{k}$ darab k elemű részhalmaza van $\{1, 2, \ldots, N\}$ halmaznak, ezért

$$\begin{aligned}
\mathbb{P}^N\big((\delta_1, \delta_2, \ldots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\big) &\le \sum_{I \in \mathcal{A}} \mathbb{P}^N\big((\delta_1, \delta_2, \ldots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\big) \\
&= \binom{N}{k} \mathbb{P}^N\big((\delta_1, \delta_2, \ldots, \delta_N) \in \Delta_I^N : V(\theta_I^*) > \epsilon\big) \\
&= \binom{N}{k} \frac{2\binom{N-k}{2}}{(k+2)\binom{N}{k+2}} \sum_{i=0}^{k+1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} =
\end{aligned}$$

$$= \binom{k+1}{k} \sum_{i=0}^{k+1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}.$$

Azaz megkaptuk, hogy

$$\mathbb{P}^N \big(V(\theta_I^*) > \epsilon \big) \leq \binom{k+1}{k} \sum_{i=0}^{k+1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i},$$

ami épp a (3.4) a $d = 2$ esetre.

Összefoglalás

A szakdolgozat során megnéztük, hogy egyszerű bizonytalansági halmaz esetén (ellipszpiod) a bizonytalan lineáris programozási feladatunk egy explicit konvex programozási feladattá válik. Ismertetve lett a szkenárió módszer, ami nagyon hasznosnak bizonyul akkor, amikor a bizonytalansági halmazról nem tudunk egyszerű alakot feltételezni, és láthattuk az alkalmazását egy illesztési feladatra. A szkenárió módszer megoldásáról bebizonyítottunk megbízhatósági szintű jellegű állításokat, hogy milyen valószínűséggel sért meg esetleg feltételeket. Definiáltuk a teljesen támasztott feladatokat, és azokra is beláttuk a kimondott tételeket.

Hivatkozások

[1] Marco C. Campi, Simone Garatti. Introduction to the Scenario Approach. Society for Industrial and Applied Mathematics, 2018.

[2] Albert N. Shiryaev. Probability. Springer, New York, 1996.

[3] Aharon Ben-Tal,Laurent El GhaouiArkadi Nemirovski. Robust Optimization. Princeton University Press, 2009.

[4] András Prékopa. Stochastic Programming. Springer Netherlands, 1995.

[5] Stephen Boyd, Lieven Vandenberghe. Convex optimization. Cambridge University Press, 2004.

[6] Erich Lehmann, Joseph P. Romano. Testing Statistical Hypotheses (Springer Texts in Statistics). Springer, 2005.

[7] Simone Garatti, Marco C. Campi, Algo Carè. On a Class of Interval Predictor Models with Universal Reliability. Automatica, Vol. 110, December 2019.