Opponensi vélemény

Fábián Csaba:
On first-order methods in stochastic programming

c. MTA doktori értekezéséről

Az értekezés 11 számozott fejezetet, egy "Additional material" című függeléket és irodalomjegyzéket tartalmaz, terjedelme 132 számozott oldal, a kapcsolódó irodalomjegyzék 191 tételből áll, amelyekből 17 a szerző saját, vagy társszerzőkkel írt dolgozata ([50]–[64], [184],[190] számú munkák).

Az értekezés témaköre a sztochasztikus programozás és modellezés, tágabb körben pedig az operációkutatás körébe tartozik. Az értekezés elsősorban a számítási algoritmusokra, ezek implementálási kérdéseire és alkalmazásaira helyezi a hangsúlyt.

A Bevezetés c. fejezetben a szerző ismerteti a sztochasztikus programozás fejlődésének legfontosabb állomásait különös tekintettel a Prékopa András nevével fémjelzett magyar iskolára és azokra a fejlődési trendekre, amelyekhez az értekezés témakörei is kapcsolódnak.

Az értekezés 2. fejezetében a vágósíkos megoldási módszereket és azok javításait tárgyalja. Itt először ismerteti a Lemaréchaltól, Nemirovskii-től és Nesterovtól származó módszer alapelvét, különféle változatait, azok előnyeit és hátrányait, a vonatkozó legfontosabb elméleti konvergencia eredményeket, majd részletesen ismerteti a saját ill. társszerzőkkel elért algoritmusait és ezek konvergenciáját. Ezek a következők: Algorithm 8 (Részben inegzakt szint módszer, 10. Tétel, 11. Állítás, Algorithm 13 (A korlátos szint módszer részben inegzakt változata, 15. Tétel, 16. Állítás). A fejezet utolsó szakaszában ismerteti a szerző előbbi eredményeinek nemzetközi szakmai hatását, továbbfejlesztéseit, ill. az algoritmusokkal kapcsolatos kedvező számítógépes tesztelési eredményeket.

Az értekezés 3. fejezete kockázatkerülő feladatok vágósíkos megoldási módszereivel foglalkozik. Legyenek $R, R' \in \mathcal{L}^1(\Omega, \mathcal{M}, P)$ véletlen hozamok ($(\Omega, \mathcal{M}, P)$ valószínűségi tér). Az R hozam dominálja az R' hozamot (másodrendű sztochasztikus dominancia (SSD), jelölés: $R \succeq_{SSD} R'$), ha a következő három ekvivalens feltétel bármelyike teljesül:

(a) $\mathrm{E}(u(R)) \geq \mathrm{E}(u(R'))$ fennáll bármely olyan nemcsökkenő és konkáv u haszonossági függvényre, amelynek véges várható értéke van.

(b) $\mathrm{E}\big([t-R]_+\big) \leq \mathrm{E}\big([t-R']_+\big)$ fennáll minden $t \in \mathbb{R}$ esetén.

(c) $\mathrm{Tail}_\beta(R) \geq \mathrm{Tail}_\beta(R')$ fennáll minden $0 < \beta \leq 1$ esetén, ahol

$$\mathrm{Tail}_\beta(R) = \max_{t \in \mathbb{R}} \{\beta t - \mathrm{ES}_t(R)\}$$

és

$$\mathrm{ES}_t(R) = \mathrm{E}\left([t-R]_+\right),$$

amely a $t \in \mathbb{R}$ cél esetén várható hiányt jelöli. A feltételes kockáztatott érték (feltételes várható veszteség Conditional Value-at-Risk)

$$\mathrm{CVaR}_\beta(Q) = \min_{t \in \mathbb{R}} \left\{ t + \frac{1}{\beta}\mathrm{E}\left([Q-t]_+\right) \right\}$$

a veszteség feltételes várható értéke az esetek legrosszabb $\beta \cdot 100\%$-ára szorítkozva.

Jelölje az $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ a különböző kötvényekbe befektetett összegek arányait és legyen az R n-dimenziós véletlen vektor a különböző kötvények hozama a befektetési periódus végén. Feltesszük, hogy R komponensei $\mathcal{L}^1$-beliek. Az x portfolió hozama $R_x = x_1 R_1 + \cdots + x_n R_n$. Jelölje továbbá $X \subset \mathbb{R}^n$ a megengedett portfoliók halmazát. Az x^* porfolió SSD-hatékonynak nevezzük, ha nincs olyan $x \in X$ portfolió, amelyre $R_x \succeq_{SSD} R_{x^*}$.

2006-ban több szerző is dolgozott ki modelleket és hatékony eljárásokat az ES és CVaR mértékek kezelésére optimalizálási feladatok esetén.

Legyen $\widehat{R}$ egy véletlen referencia hozam. Dentcheva és Ruszczynski SSD korlátozott portfólió optimalizálási modellje a következő:

$$\begin{gathered} \max f(x) \\ x \in X,\ R_x \succeq_{SSD} \widehat{R} \end{gathered}$$

ahol $f = E(R_x)$ konkáv függvény. Roman és társainak ([147]) (c) kritériumon

alapuló optimalizálási modellje

$$\begin{gathered} \max \vartheta \\ \vartheta \in \mathbb{R},\ x \in X \\ \mathrm{Tail}_{i/S}(R_x) \geq \mathrm{Tail}_{i/S}\left(\widehat{R}\right) + \vartheta \quad (i = 1, \ldots, S). \end{gathered}$$

Véges diszkrét eloszlások esetén Fábián az **ES** és **CVaR** mértékek közti konvex konjugálási viszonyt egyszerűsítette egy lineáris programozási dualitási viszonyra ([52]) és később egy duális megoldási algoritmust is javasolt a feladat megoldására ([63]). Az utóbbi ([63]) dolgozatára alapozva a Brunel Egyetem CARISMA (Centre for the Analysis of Risk and Optimisation Modelling Applications) kutatócsoportjával közösen hatékony vágósíkos eljárásokat dolgozott ki SSD korlátos feladatok megoldására, amelyet számítógépre is implementáltak ([57]). Az eljárásokat beépítették az OptiRisk Systems által fejlesztett pénzügyi elemző eszközökbe is. Fábián Csaba és társszerzői (Mitra, Roman, Zverovich) kidolgozták a Roman et al. [147] modell egy javított skálázott változatát

$$\begin{gathered} \max \vartheta \\ \vartheta \in \mathbb{R},\ x \in X \\ R_x \succeq_{SSD} \widehat{R} + \vartheta, \end{gathered}$$

amelynek a megoldására egy vágósikos módszert is javasoltak. A számítógépes tesztelés a módszer, ill. a modell hatékonyságát igazolta. Egy későbbi munkájuk ([59]), amelyben a tesztelés input adatokként az FTSE adatokon alapult megerősítette a skálázott model alkalmazhatóságát.

Fábián Csaba és szerzőtársai fejezetbeli eredményeit számos neves külföldi szerző továbbfejlesztette, ill. alkalmazta.

Az értekezés 4. fejezete dekompozíciós módszereket vizsgál kétlépcsős sztochasztikus programozási feladatokra, amelyekben az első lépcső döntése deteminisztikus, míg a második már véletlenszerű. Legyen az első lépcső döntése $x \in X$ (X nemüres konvex korlátos poliéder) és tegyük fel, hogy S lehetséges véletlen kimenet van, ahol az s-edik kimenet valószínűsége p_s. Az első lépcső alakja

$$c^T x + \sum_{s=1}^{S} p_s q_s(x) \to \min \quad (x \in X),$$

a második lépcső alakja

$$\mathcal{R}_s(x): \qquad \begin{aligned} & q_s^T y \to \min \\ & T_s x + W_s y = h_s \\ & y \geq 0, \end{aligned}$$

ahol q_s, h_s adott vektorok, T_s, W_s adott mátrixok. A fejezetben feltételezi, hogy a második lépcső mindig megengedett bármely $x \in X$ esetén. A két probléma egy közös LP feladatban foglalható össze, amelyben a részproblémákat az első lépcső döntési változói kötik össze. Az ismert számítási eljárások összefoglalása után részletesen ismerteti a saját eredményeit. Ezek egy közelítő szint módszer adaptálása, ismert eljárások tesztelése és szoftverfejlesztés (a Brunel Egyetem CARISMA csapatával), valamint egy dekompozíción alapuló keretrendszer (a paderborni egyetem DS&OR laboratóriumával), amely egyesíti az aggregált és diszaggregált modellek előnyeit. A számítógépes tesztelések a javasolt módszerek hatékonyságát mutatják (lásd pl. a 4.3 ábrán mutatott teljesítményprofilt). A fejezet végén az alkalmazásokról is beszámol.

Az 5. fejezet kétlépcsős sztochasztikus optimalizálási problémák megengedettségi kérdéseit vizsgálja. Itt nem teszi fel, hogy hogy a második lépcső mindig megengedett bármely $x \in X$ esetén. Az első lépcső feladata ennek megfelelően

$$c^T x + \sum_{s=1}^{S} p_s q_s\left(x\right) \to \min \quad \left(x \in X \cap K\right),$$

ahol $X \cap K$ nem üres, $K = \cap_{s=1}^{S} K_s$ és K_s $(s = 1, \ldots, S)$ azon x vektorok halmazát jelöli, amelyre az $\mathcal{R}_s\left(x\right)$ feladat megengedett. A szerző egy regularizációs jellegű eljárást javasol a második lépcső megengedettségi problémáinak kezelésére és vizsgálja a regularizációs paraméterek értékét, ill.becslését is. A javasolt eljárást implementálta és eredményesen tesztelte is.

Az értekezés 6. fejezete kétlépcsős kockázatkerülő sztochasztikus programozási feladatokat vizsgálja két modell esetére.

Adott $x \in X$ esetén tekintsük a $q_s\left(x\right)$ függvény értékeit a véletlen $Q\left(x\right)$ függvény realizálásainak. Az első modellben az első lépcső alakja:

$$\begin{aligned} c^T x + E\left(Q\left(x\right)\right) &\to \min \\ x &\in X, \\ \beta \mathrm{CVar}_\beta\left(Q\left(x\right)\right) &\le \rho, \end{aligned}$$

ahola β és ρ paraméterek adottak. Legyen Q és $\widehat{Q}$ két integrálható véletlen változó (költség). $Q \preceq_{IC} \widehat{Q}$ akkor és csak akkor, ha $-Q \succeq_{SSD} -\widehat{Q}$. A második modell első lépcsőjének alakja

$$\begin{aligned} c^T x + E\left(Q\left(x\right)\right) &\to \min \\ x &\in X, \\ Q\left(x\right) &\preceq_{IC} \widehat{Q}, \end{aligned}$$

ahol $\widehat{Q}$ adott integrálható véletlen változó (veszteség, vagy benchmark költség). Mindkét feladatra adaptálja a 2. fejezetében javasolt algoritmusát (Algorithm 25, Algorithm 26). A paderborni egyetem DS&OR laboratóriumának munkatársaival közösen implementálta és tesztelte a javasolt algoritmust. A részletezett számítási eredmények világosan mutatják a javasolt eljárásának hatékonyságát és azt hogy módszere versenyképes Lemaréchal, Nemirovskii, Nesterov módszerével (lásd pl. a 6.1 ábra teljesítményprofilját).

A 7. fejezetben

$$\max_{x \in X} P\left(g\left(x\right) \geq Z\right),$$

illetve

$$\begin{aligned} h\left(x\right) &\to \min \\ x &\in X, \\ P\left(g\left(x\right) \geq Z\right) &\geq p \end{aligned}$$

alakú valószínűség maximalizálási feladatokat vizsgál, ahol Z egy ismert eloszlású n-dimenziós véletlen vektor, $x \in \mathbb{R}^m$, $X \subset \mathbb{R}^m$, $g : \mathbb{R}^m \to \mathbb{R}^n$, $h : \mathbb{R}^m \to \mathbb{R}$, $p > 0$ adott valószínűség és a véletlen paraméterek eloszlása log-konkáv.

Prékopa András belső közelítési sémájához kapcsolódva olyan közelítő eljárást dolgozott ki, amely a valószínűségi függvény epigráfjának belső approximációját használja. Eljárása abban különbözik Dentcheva, Prékopa, Ruszczynski kúpgeneráló eljárásától, hogy nem a valószínűségi függvény szinthalmazát közelíti, hanem az epigráfját. Az új próbapontok meghatározása korlátozás nélküli konvex minimalizálással (gradiens módszerrel) történik, szemben a klasszikus sémával, ahol a valószínűségi függvény szinthalmaza felett kell optimalizálni. Az új algoritmust több heurisztikus részmegoldással implementálták és az ezzel végzett kezdeti tesztelés azt mutatja, hogy az eljárás nem érzékeny a gradiensek számítási hibáira, valamint azt, hogy a valószínűség előírt pontosságú maximalizálásához szükséges próbapontok száma soha nem növekedett jelentősen.

A 8. fejezetben az előző fejezet eljárására (oszlop generálási séma) alapozva szimulációs eljárást ad meg

$$\begin{aligned} \phi(Tx) &\to \min \\ Ax &\leq b \end{aligned}$$

alakú feladatok megoldására, ahol $x \in \mathbb{R}^n$, $T \in \mathbb{R}^{n \times m}$, $A \in \mathbb{R}^{r \times m}$ és $b \in \mathbb{R}^r$. A ϕ függvény konvex, kétszer folytonosan differenciálható és Hesse mátrixának sajátértékei egy fix $[\alpha, \beta]$ $(\alpha > 0)$ intervallumba esnek minden $x \in \mathbb{R}^n$ esetén. Itt javasolja az előző fejezet oszlop generálási sémájának egy randomizált változatát, ill. kidolgozza a gradiens módszer egy sztochasztikus változatát, amelynek hibájára a 35. Tételt igazolja.

A 9. fejezetben Fábián Csaba társszerzőkkel közösen szimulációs eljárást mutat be

$$\begin{aligned} c^T x &\to \min \\ Ax &\leq b, \\ \phi(Tx) &\leq \alpha \end{aligned}$$

alakú feladatok esetén nehezen számítható, de jól kondicionált ϕ függvény kezelésére. Az előző feladat helyett a

$$\begin{aligned} \phi(Tx) &\to \min \\ Ax &\leq b, \\ c^T x &\leq d \end{aligned}$$

feladatok sorozatát oldja meg csökkenő tűréssel. Jelölje $\chi(d)$ az utóbbi feladat optimumát. Alkalmas feltevések mellett $\chi(d)$ monoton csökkenő függvény és a $\chi(d) = \alpha$ feladat megoldása adja az eredeti feladat megoldását. A feladatsorozat elemeinek megoldására az előző fejezetben ismertetett eljárást használja kombinálva a $\chi(d) = \alpha$ feladatra alkalmazott Newton-módszerrel. A 49. tételben az eljárás konvergenciáját igazolja.

Az értekezés utolsó 10. fejezetében a 7. fejezetben bemutatott eljárásra alapozva szimuációs eljárást ad meg olyan valószínűség-maximalizálási feladatokra, amelyekben a véletlen paraméterek nemdegenerált normális eloszlásúak. Az eljárást társszerzőkkel számítógépre is implementálta és tesztelte. A kapott teszteredmények az eljárás gyors konvergenciáját mutatják.

Az értekezés utolsó 4 fejezetében bemutatott egymáshoz szorosan kapcsolódó eredmények lényegében az utóbbi 2-3 évben kerültek kifejlesztésre és publikálásra. A bemutatott elméleti megfontolások és számítógépes tesztelések alapján várható hogy jelentős hatást fognak a sztochasztikus programozás területén kifejteni.

Összegzés

Fábián Csaba a sztochasztikus programozás területén, elsősorban a hatékony számítógépes algoritmusok fejlesztésében és alkalmazásaiban ért el nemzetközileg is számontartott jelentős eredményeket. A szerző saját tesztelései és a független számítógépes tesztelések és gyakorlati alkalmazások is alátámasztják Fábián Csaba módszereinek versenyképességét. Algoritmusainak egy részét - elsősorban a pénzügyi szektorban - szoftvercsomagokba is beépítették. Ezen a területen a jelölt kiemelkedően színvonalas alkalmazási tevékenységet végzett. Az értekezésben bemutatott eredmények nemtriviálisak, azok kifejlesztéséhez elsősorban a sztochasztikus programozás algoritmusainak, ezek számítógépes változatainak, valamint a professzionális programozásuknak mély elméleti és gyakorlati ismeretére volt szükség.

A doktori munka tudományos eredményei alapján a nyilvános vita kitűzését és az MTA doktori cím odaítélését javaslom.

Budapest, 2020. október 6.

Galántai Aurél
az MTA doktora