

Bírálat

Fábián Csaba: On first order methods in stochastic programming c.

MTA doktori értekezéséről

Deák István, az MTA doktora

A bírálatomat több részre tagolva írtam meg. Az első részben áttekintem a munkát és egy vázlatos leírást adok a doktori értekezésről, a második részben egy rövid áttekintést adok a disszertációról, a harmadik részben kiemelem az általam fontosnak tartott eredményeket és megadom méltatásukat, a negyedik részben az általam kifogásolhatónak tartott részekkel foglalkozom, míg az utolsó részben összefoglalom bírálatomat és ajánlást teszek az értekezés további sorsára.

1. Az értekezés angol nyelvű, 11 fejezetre és egy függelékre van osztva és összesen 132 oldalas (ebből az irodalomjegyzék 16 oldalt foglal el).

Az első fejezet egy rövid bevezetés, amelyben a sztochasztikus programozás legismertebb kutatóit sorolja fel és méltatja a szerző. A másodiktól a hatodik fejezetig a disszertáció a kétlépcsős problémák megvitatásával és a vonatkozó megoldási eljárásokkal foglalkozik. A hetedik-től a tizedik fejezetig a valószínűséggel korlátozott modellekkel kapcsolatos eredményeit írja le a szerző. A 11. fejezetben az eddigi eredmények összefoglalását vázolja, majd egy kétoldalas függelékben az algoritmusok összehasonlítására vonatkozó, eddig nem ismertetett részleteket, és a primál-duál összefüggéseket említi meg.

A sztochasztikus programozás két nagy területre osztható: a kétlépcsős modellek és a valószínűséggel korlátozott modellek. Ezek a teljes tudományág 50 %-át illetőleg 30 %-át fedik le, a fennmaradó 20 százalék nehezen kategorizálható, többféle részkutatás között oszlik meg. Az értekezés első része (a második fejezettől a hatodikig) foglalkozik a kétlépcsős feladatokkal, míg a második, kisebbik része 7.-től a 10. fejezetig, a valószínűséggel korlátozott modellekkel.

Minden fejezetben először az alapproblémát ismerteti a szerző, aztán a legjobb, vagy általa érdekesnek tartott megoldást írja le, ezt követően a saját fejlesztéseit, módosításait vagy eljárását adja meg az alapproblémára egy „Contribution” elnevezésű alfejezetben, s végül általában az irodalomban található egyéb megoldásokkal hasonlítja össze a saját javaslatait, vagy számítástechnikai összehasonlításokról ír. Minden esetben világosan elkülöníti azt, hogy mivel járult hozzá az elmélet vagy a gyakorlati számítási eljárások területén az eddigi eredményekhez, valamint azt is pontosan leírja, ha több társszerző által elért eredményről van szó, hogy mi az egyes szerzők eredménye, mit tekint a saját eredményének. A fejezetek végén egy-egy összefoglalásban kiemeli a legfontosabb, az adott témában elért eredményeit.

A szerző külön érdemének tekintem, hogy sok különböző kutatóval dolgozott együtt, ezeknek a csoportos kutatásoknak a kezdeményezője és a kutató csoportok összefogójának bizonyult a leírt eredmények elérése alatt.

2. Általános értékelés. A szerző 15 saját cikkét sorolja fel az irodalomjegyzékben, ezek többnyire társszerzőkkel íródtak, de négyet csak a jelölt írt, további négy cikk pedig doktorandusz hallgatóival (Papp, Szőke, Veszprémi) készült. Több olyan cikk szerepel az irodalomjegyzékben, amelyek társszerzői Fábián mellett a sztochasztikus programozás neves kutatói (Prékopa, Mitra, Szántai). A cikkek főleg neves folyóiratokban jelentek meg. A szerző eredményeit számos konferencián előadta, amelyeken a jelen bíráló írója is részt vett és ezeket az eredményeket a hallgatóság elismeréssel fogadta.

Az értekezés minden részében számos idézet és hivatkozás szerepel a sztochasztikus programozás neves kutatóitól, és a szerző eredményei ezekhez illeszkednek, illetve az eddig publikált eredmények továbbfejlesztései, vagy, más szerzők Fábián itt leírt eredményeire épülő eljárásokat fejlesztettek ki.

Az értekezés angol nyelven íródott, a nyelvezete megfelel a szokásos tudományos terminológiának, nyelvtanilag helyesen van megfogalmazva. Az egész mű – némileg kibővítve a kutatók előtt közismert alapelvekkel, modellekkel, valamennyi magyarázó szöveggel ellátva és a speciális részek kihagyásával vagy lerövidítésével tankönyvként is hasznos lehetne – ennek a könyvnek a megírását a bíráló hasznosnak tartaná. Egyébként a disszertáció szép kiállítású, jól tagolt, arányos szerkezetű.

3. Mivel az értekezés igen széles témakört fed le, ezért csak az általam legfontosabbnak tartott eredményeket ismertetem.

A metszősíkos eljárások újabb változatait ismertető második (valamint az itt leírt, egészen általános eljárásnak speciális feladatokra való alkalmazásait tartalmazó harmadik, negyedik és hatodik) fejezetet tartom az értekezés egyik legérdekesebb részének. A metszősíkos eljárások általában cikk-cakk módra haladnak az optimum felé, ezért regularizált változatokat használnak a konvergencia gyorsítására. Ezen regularizált eljárásokban azzal is számolni kell az optimalizálás folyamán, hogy orákulum függvényeket használhatunk az iterációk folyamán kapott megoldásokra, valamint a függvényértékek csak becslések (csak a várható értékük a függvényérték). Eredetileg Lemarechal, Nemirovski és Nesterov dolgozták ki a szintvonalas eljárást, amelyben felső és alsó korlátokat adnak a szinthalmazokra – ez az eljárás jobban viselkedett, mint az addigi algoritmusok. Fábián egy adaptív verziót dolgozott ki a pontatlanságok kezelésére: minél kisebb a felső és alsó korlátok közötti különbség, annál pontosabban becsüli meg a függvényértékeket. Az eljárást részben aszimptotikusan pontos eljárásnak nevezte. Ennek az eljárásnak nem túl szigorú feltételek mellett a konvergenciáját is bebizonyította. A második fejezetben leírt általános eljárásokat alkalmazta a szerző a harmadik fejezetben kockázat kerülő (CVaR) problémák megoldására, valamint a negyedik fejezetben kétlépcsős modellekre, valamint a hatodik fejezetben kockázati feltételekkel kibővített kétlépcsős modellekre. (Az ötödik fejezetben második lépcsős megengedettségi problémákkal foglalkozik a szerző.)

A másik, általam kiemelt rész a hetedik fejezetben tárgyalt eljárás a valószínűséggel korlátozott modelltípus megoldási algoritmusának egy verziója. Tekintjük az eredeti modellben a valószínűségi feltételt és annak a transzformált (negatív logaritmusát) változatát, amit aztán minimalizálnánk. Ennek a Lagrange duálisát oldjuk meg, egy lineáris programozási feladatként. A végső megoldó algoritmusban heurisztikus ötletek is vannak, amelyek a

számítási eljárást gyorsítják. Az eljárások alapján véve egy módosított Newton módszeről indulnak ki, a szerző ezek konvergenciáját is bebizonyítja, viszonylag enyhe feltételek mellett (korlátos Hesse mátrixok). Ezek az átalakítások végülis egy poliédeses approximációval helyettesítik a nehezen kiszámítható valószínűségi függvényt, amit ötletes elgondolásnak tartok. Ezt az alapalgoritmust aztán többféle feladatra (valószínűség maximalizálására, és „nehéz” függvénnyel megadott korlátozó feltétel kezelésére) alkalmazza a nyolcadik, kilencedik és tizedik fejezetekben.

A bíráló a fentebb felsorolt és leírt ötleteket (különösen a második és a hetedik fejezet általános algoritmikus módosításait és elméleti eredményeit) eredetinek ítéli meg és a jelölt saját fejlesztéseinek tartja.

4. Az értekezés hiányosságának tartom, hogy a számítástechnikai eredmények nincsenek megfelelően leírva. Annak ellenére, hogy a hivatkozások az eredeti cikkekre, ahol az eredmények és számítási feladatok megjelentek, pontosak, de a disszertáció olvasója szívesen látott volna valamennyi számítási eredményt is a disszertációban magában – egy-egy fejezet végén egy négy-öt soros összefoglalást például. A feladatok méretei, az eloszlások típusa és dimenziója, a vektorok és mátrixok dimenziói fontos mutatói a megoldhatóság nehézségeinek. A tesztfeladatok méretei nem nyilvánvalóak minden esetben, többször csak a szöveg többszöri átnézése után találhatók meg erre vonatkozó utalások. Nyilvánvalóan ez a kritika nem érinti a hatalmas tesztelési munkát bemutató negyedik és hatodik fejezetet, bár itt is jól esett volna látni néhány számot, például a legkisebb méretű feladatok négy-öt paraméterét, a legnagyobb megoldott feladatok méretei, stb. Hasonlóképpen örültem volna annak, ha legalább egy-két mondatban kitért volna a szerző arra, hogy milyen számítógépet használtak a feladatok megoldására, mennyi volt a futási idő és ehhez hasonló adatokra (9. és 10. Fejezet például).

Egy másik, kisebb problémának látom, hogy a saját eredmények elhatárolása, illetőleg az eredmények mások eredményeihez való kapcsolódásainak bemutatása többször is előfordul, így általában az egyes fejezetek elején is és a végén is, valamint még egyszer a 11. fejezetben, ahol az eredmények összefoglalásainak összefoglalását (sic) találjuk meg. Ezek a kritikai megjegyzések a dolgozat érdemeit csak kevésbé csorbítják.

A teljes disszertáció átolvasása után sem sikerült felfedeznem az értekezés szövegében sajtóhibát, elütést, értelemzavaró megfogalmazást, ami ritkaságnak számít a mai elektronikus publikálási versenyben.

5. Összességében a dolgozat jól megírt és fontos hozzájárulásokat tartalmaz a sztochasztikus programozás elméletéhez és gyakorlatához. Ezért a doktori értekezés vitára bocsátását támogatom, valamint a fokozat odaítélését is támogatom.

Budapest, 2020. augusztus 31.

Deák István, az MTA doktora