

Válasz Abaffy József opponensi bírálatára

Köszönöm Abaffy József részletes és alapos értékelését és támogatását. Kérdéseire az alábbiakban válaszolok.

1. kérdés: Több algoritmus leírásban használja az ε kilépési feltételt. Pl. Algorithm 8, Algorithm 13, Algorithm 20. Kérem beszéljen ennek pontos értékéről és a kilépési feltételekről általában, amikor különböző módszereket hasonlít össze.

Válasz: Előzetesen megjegyzem, hogy az algoritmusok leírásakor áttekinthetőségre törekedtem, azonban az implementációk bonyolultabbak, a felmerülő numerikus és technikai nehézségek miatt.

Az implementációkban relatív toleranciával dolgozunk, mivel a különböző feladatok optimális célfüggvény-értékei eltérő nagyságrendűek. Az eljárás akkor áll le, ha $(F^* - F_*)/(|F_*| + 10^{-10}) \leq \varepsilon$ teljesül, ahol F^* és F_* az optimális célfüggvény-értéknek felső illetve alsó közelítése. A tesztfeladatokon általában $\varepsilon = 10^{-5}$ értékkel dolgoztunk.

Egyszerű vágósíkos módszer és level típusú módszerek esetén a szükséges alsó és felső közelítések minden lépésben rendelkezésre állnak.

A bundle típusú módszerek a célfüggvény vágósíkos modelljéhez még egy kvadratikus kifejezést hozzáadnak. Így a modell-feladat optimuma nem feltétlenül szolgáltatja az eredeti optimum alsó korlátját. – A bundle típusú célfüggvény mellett külön minimalizálhatjuk a modell-függvényt is, ez azonban növeli a számításigényt, ezért nem akartuk minden iterációban végrehajtani. Azt tapasztaltuk, hogy a közbelső iterációkban egy nagyságrenddel szűkebb toleranciát érdemes alkalmazni, pld. 10^{-5} helyett 10^{-6} .

2. kérdés: A Dennis Moré (1983) könyvben p.123 található konvergencia tétel szerint kétszer folytonosan differenciálható függvények minimalizálásakor a line search lépésben bizonyos (Armijo) feltételek teljesülése esetén egy bizonyos iterációs szám után elegendő egyetlen lépést számolni ahhoz, hogy szuperlineáris konvergencia teljesüljön.

Kérdésem, ehhez a konvergencia tételhez hasonló numerikus eredményt talált-e az Algoritmusainak futtatása során?

Válasz: Kísérleteink során szuperlineáris konvergenciát nem tapasztaltunk, azonban a Bíráló által említetthez hasonló jelenséget igen. Köszönöm az érdekes párhuzamot. Az alábbiakban bővebben válaszolok a kérdésre, kifejtem a párhuzamot, és egy lehetőséget vázolok arra, hogy a Bíráló által hivatkozott kvázi-Newton módszert miként lehetne eljárásainkban alkalmazni.

Eddig implementált megoldó eljárásaink csak első rendű információt használnak. Ugyanis minél egyszerűbb eljárás kidolgozására törekedtünk (miközben a gyakorlati használhatóság követelményét is szem előtt tartottuk.) Valószínűségi függvény esetén már a gradiens komponenseinek közelítő számítása során is nehézségek lépnek fel, ezért eddig meg sem kíséreltünk olyan eljárást implementálni, amely a Hesse mátrix approximációjával dolgozik.

A Bíráló által említett jelenséggel párhuzamba állítható az a futtatási tapasztalatunk, hogy a valószínűség maximalizáláshoz alkalmazott belső közelítő eljárás akkor is konvergál, ha minden iterációban csak egyetlen vonalmenti keresési lépést hajtunk végre. Ezt a tapasztalt jelenséget időközben elméleti megfontolással is sikerült alátámasztanom. A konvergencia bizonyításához elegendő, ha a célfüggvénynek az optimális megoldásban vett Hesse mátrixa jól kondicionált. – A bizonyítás azon alapszik, hogy a célfüggvény gráfjára alkalmazott belső közelítés tulajdonképpen oszlopgeneráló eljárás. Ez pedig duális nézőpontból egyszerű vágósíkos eljárás, ahol azonban a vágások nem pontosak. – A konvergencia bizonyítását leíró cikk a doktori pályázatom benyújtása óta jelent meg, egyelőre elektronikus formában:

Gaining traction: on the convergence of an inner approximation scheme for probability maximization. Central European Journal of Operations Research, 2020. DOI: 10.1007/s10100-020-00697-3.

A Bíráló által hivatkozott kvázi-Newton módszer új lehetőségeket nyithat, azonban a valószínűség-maximalizálási feladatban szereplő korlátozó feltételek nehézséget jelentenek. A célfüggvény gráfját belülről közelítő modell megtartásával ez a nehézség elkerülhető. A modellt iteratívan javító keretben ugyanis a belső közelítés újabb próbapontjait egy-egy korlátozás nélküli konvex minimalizálási feladat (közelítő) megoldásával keressük meg. A javítás abban állna, hogy az újabb próbapontokat egy-egy kvázi-Newton lépéssel keresnénk meg, nem pedig a legmeredekebb csökkenés irányában végrehajtott vonalmenti lépéssel, mint most. A valószínűségi függvény gráfját belülről közelítő modell pedig itt is lehetővé teszi egyszerű véletlenített eljárás kidolgozását és konvergenciájának bizonyítását. (A modelfüggvénynek köszönhetően egy-egy, különösen zajos adatokkal számolt próbapont nem téríti el az eljárást.) Megítélésem szerint ezen az úton a mostani eljárásunknál hatékonyabbat lehet kifejleszteni.

Budapest, 2020. november 5.

Fábián Csaba