

Diplomamunka témajavaslatok

Bolla Marianna
Sztochasztika Tanszék

- Többváltozós és diszkrét varianciaanalízis:** A klasszikus varianciaanalízis feladata egy vagy több szempont alapján képzett csoportok összehasonlítása egy egydimenziós normális eloszlású mérés alapján (ez része a Többváltozós statisztika tárgy anyagának). A módszer általánosítható többdimenziós normális eloszlásból származó megfigyelésekre (többváltozós varianciaanalízis, angol mozaikszóval MANOVA) ill. diszkrét megfigyelésekre. Előbbihez többdimenziós Steiner-tétel, Wilks Λ -eloszlás, utóbbihoz rangstatisztikák kellenek. A feldolgozandó könyvfejezet ill. cikk az utóbbi 20-30 évből valók.
- Véletlen mátrixok a statisztikában:** Az empirikus kovarianciamátrix eloszlása (Wishart) a Többváltozós statisztika tárgy anyaga. A többváltozós statisztikai módszerekben előkerülő egyéb mátrixfelbontások mátrixainak eloszlása hasonlóan levezethető. Nagyméretű adatrendszerekre alkalmazott algoritmusok gyakran csak a legnagyobb sajátértékek/szinguláris értékek becslését, nagyságrendjét igénylik. Ehhez nagy eltérés típusú eredmények kellenek. Az utóbbi 10-20 év ilyen témájú cikkei használhatók.
- Inhomogén kvadratikus alakok optimumfeladatai és statisztikai alkalmazása:** Legyenek A_1, \dots, A_k szimmetrikus, valós, $n \times n$ -es mátrixok ($k \leq n$). A feladat $\sum_{i=1}^k x_i^T A_i x_i$ maximalizálása az $x_1, \dots, x_k \in R^n$ ortonormált rendszereken. A feladat pontos megoldásához az $|A - I_n \otimes S| = 0$ általánosított karakterisztikus egyenlet írható fel, ahol az $nk \times nk$ -as A mátrix az A_i mátrixok Kronecker-összege, a $k \times k$ -as S szimmetrikus mátrixot pedig keressük (Bolla és társszerzők cikk, 1998). A többváltozós polinom gyökeinek megkeresése megoldatlan, egyelőre numerikus közelítő algoritmust használunk. Akár az egzakt megoldás megtalálása, akár az algoritmus továbbfejlesztése érdekes feladat, akárcsak a maximum- és a duális minimum-feladat alkalmazása statisztikai problémákban, pl. rétegzett megfigyelések ún. kompromisszum-faktorainak meghatározására.
- EM-algoritmus:** Az algoritmust maximum likelihood becslések meghatározására dolgozták ki hiányos megfigyelések alapján, ahol a hiányos adatok nem feltétlenül hiányos adatrendszer, hanem becsülendő modellparamétereket is jelenthetnek. A Baum–Welch algoritmus általánosításaként 1985-ben publikált algoritmus (mely egy E: feltételes várható érték képzés és egy M: maximalizáló lépésből áll) konvergencia-bizonyítása, továbbá speciális alkalmazási lehetőségeinek áttekintése a feladat.