

POSUDEK VEDOUČÍHO DIPLOMOVÉ PRÁCE

Název: Halfspace depth for location and scatter: robustness and minimax optimality

Autor: Filip Bočinec

Práca pojednáva o polopriestorovej (Tukeyho) hĺbke viacrozmerných rozdelení v dvoch situáciách: (i) klasickej hĺbke pre parameter polohy, a (ii) rozšírení polopriestorovej hĺbky na prípad odhadovania variančnej štruktúry rozdelenia v \mathbb{R}^d . Práca vychádza z nedávneho prelomového článku Chen, Gao, Ren (2018) v ktorom bolo dokázané, že mediány (tj. najhlbší bod, resp. matica) indukované polopriestorovými hĺbkami sú minimax optimálnymi odhadmi pre viacrozmerné normálne rozdelenie v štandardných kontaminačných modeloch. Ide o zďaleka najvýznamnejší výsledok v teórii štatistickej hĺbky v poslednom období. Takmer 50 rokov po zavedení polopriestorovej hĺbky do štatistiky sa tým plne legitimizuje použitie hĺbky v robustnej analýze dát. Dôkaz tohto výsledku je pochopiteľne veľmi komplikovaný, a vyžaduje zvládnutie celej rady moderných štatistických techník. Hlavným cieľom predloženej práce bolo overenie, a spracovanie tohto dôkazu.

Kapitoly 1 a 2 predstavujú polopriestorové hĺbky pre parameter polohy a variancie, a diskutujú ich teoretické vlastnosti. Dôraz je kladený najmä na robustnosť indukovaných mediánov. V tejto časti práce je zaujímavá veta 23, kde autor nachádza a opravuje pomerne významnú chybu v hlavnom výsledku článku Chen a Tyler (2002) o robustnosti polopriestorového mediánu. Všetky tieto výsledky sú prípravou na hlavnú časť práce v kapitole 3, kde je detailne dokázaná minimax optimalita polopriestorových mediánov. V úvode kapitoly sú prehľadne zhrnuté základy teórie optimality odhadu založenej na rýchlosti konvergencie. Jedná sa o veľmi pokročilú teóriu, ktorá je ešte aj v súčasnosti objektom aktívneho výskumu. Hlavné dôkazy minimax optimality nájdeme v sekcii 3.4 (pre parameter polohy) a v sekcii 3.5 (pre odhad variancie/tvaru). Tieto dôkazy sú značne komplikované — využívajú nielen nástroje pravdepodobnosti a štatistiky, ale napríklad aj analytické, geometrické a kombinatorické argumenty. Rovnako sa tu rutinne využíva pokročilá teória empirických procesov.

Pri spracovávaní hlavných dôkazov p. Bočinec narazil na radu (drobných i väčších) nezrovnalostí v pôvodnom zdroji. Všetky tieto problémy sa mu samostatne podarilo vyriešiť, a to aj za cenu úpravy častí pôvodného dôkazu. Dôležitá je veta 57, ktorá je originálnym výsledkom s vlastným dôkazom p. Bočina. Toto tvrdenie je v článku Chen et al. (2018) neúplné, a nesprávne odkázané do literatúry. Ďalšou významnou úpravou pôvodného dôkazu je jemný argument o použití lipschitzovskej spojitosti na kompaktnej množine v závere dôkazu hlavnej vety 58. Tento krok je v pôvodnom dôkaze neúplný. Originálna je aj časť dôkazu vety 55, kde p. Bočinec zjednodušuje pôvodný argument pomocou „Le Camovej dvoj-bodovej metódy“, a ukazuje, že výsledok je možné získať aj elementárne.

Matematická úroveň práce je vynikajúca. Všetky dôležité tvrdenia sú kompletne dokázané, dôkazy sú detailné, jasné, a prehľadné. Práca je pomerne obsiahla, veľmi dobre sa však číta. V texte som nenašiel preklepy ani významnejšie formálne chyby. V budúcnosti bude zaujímavé napríklad rozšírenie tohto prístupu na ďalšie triedy rozdelení.

Jedná sa o výnimočnú teoretickú prácu na vysokej úrovni. Rozhodne ju odporúčam uznať ako diplomovú prácu na MFF UK.



Stanislav Nagy
KPMS MFF UK
24. července 2024