

POSUDOK VEDÚCEHO BAKALÁRSKEJ PRÁCE

Názov: Geometric approach to the estimation of scatter
Autor: Juraj Bodík

ZHRNUTIE OBSAHU PRÁCE

Práca predstavuje dva robustné odhady parametrov polohy a rozptylenia (angl. *scatter*) viacozmerných dát známe z literatúry: MCD (minimum covariance determinant) a MVE (minimum volume ellipsoid). Oba odhady sú podobné v tom, že využívajú geometriu dát, a dajú sa interpretovať tak, že dátový súbor prekladajú vhodne voleným elipsoidom ktorý popisuje tvar (časti) pozorovaní.

V kapitole 1 je problematika uvedená na príklade klasického tolerančného elipsoidu pre viacozmerné normálne rozdelenie. Je diskutovaný jeho súvis s Mahalanobisovou vzdialenosťou od centra rozdelenia. Tieto výsledky sú využité v ďalšej časti práce, kde je zavedený MCD odhad. V kapitole 3 sú skúmané základné teoretické vlastnosti MCD odhadu. Je dokázaná jeho afínna ekvivariancia a presná hodnota bodu zlomu, a čiastočne dokázaný dôležitý fakt, že MCD odhad má maximálny možný bod zlomu afínne ekvivalentných odhadov. V náznaku je diskutovaná voľba parametra MCD odhadu zaručujúca jeho konzistenciu pre normálne rozdelenie, a jeho eficiencia.

V závere práce je tento odhad v krátkosti porovnaný s MVE odhadom. Na príklade autor vysvetluje fundamentálne rozdiely medzi prístupmi.

CELKOVÉ HODNOTENIE PRÁCE

Vlastný príspevok. Jedná sa skôr o prácu kompilačného typu. Autor spracoval značné množstvo netriviálnej odbornej literatúry. Uvedené dôkazy jednoduchších tvrdení, a náznaky hlavných myšlienok dôkazov zložitejších viet, sú ale jeho vlastným príspevkom.

Matematická úroveň. Úroveň práce je uspokojivá. Volené matematické značenie a argumentácia by sa miestami dali upraviť, predvedené dôkazy sú ale podľa môjho najlepšieho vedomia vedené rigorózne a správne.

Práca so zdrojmi. Zdroje sú citované korektne.

Formálna úprava. Práca je písaná anglicky. Nejde o autorov prvý jazyk, čo sa na výklate podpisuje. Niektoré formulácie sú nešikovné a ťažkopádne, celkovo je ale text zrozumiteľný a neobsahuje zásadné chyby.

ZÁVER

Pán Bodík pristupoval k práci veľmi aktívne a so zaujatím. Nazbieraný materiál by nepochybne dokázal ďalej rozširovať — predložená práca je iba skrátenou verziou textu

ktorý pôvodne pripravil. S časťou pôvodnej práce, ktorá obsahuje návrhy viacerých originálnych metód detekcie odľahlých pozorovaní, a ktorá nie je súčasťou finálnej verzie predloženej bakalárskej práce, sa autor zúčastnil súťaže SVOČ.

Tematika študovaná v práci je obsiahla a náročná. Autor sa však s témou vysporiadal bez väčších problémov. Vo výsledku prácu považujem za dobre zvládnutý úvod do problematiky robustnej analýzy viacozmerných dát. Odporúčam ju uznáť ako bakalársku prácu na MFF UK.

Stanislav Nagy
KPMS MFF UK
20. mája 2019