

Posudek oponenta bakalářské práce

Jméno studenta	Michaela Zavřelová
Téma práce	Metoda nejmenších čtverců
Cíl práce	Tématem této bakalářské práce je metoda nejmenších čtverců a její využití. Metoda nejmenších čtverců je jednou ze základních metod aproximace funkcí ve statistice. Tato metoda je základní technikou používanou k vyrovnání naměřených dat, odhadu parametrů aproximačních funkcí například v regresní analýze. V práci budou popsány základní matematické vlastnosti, které jsou využity v definicích a výpočtech. Cílem této práce je zaměření na samotnou metodu nejmenších čtverců a její použití na konkrétních příkladech.
Vedoucí bakalářské práce	Mgr. Jitka Kühnová, Ph.D.

náročnost tématu na	úroveň
----------------------------	---------------

	nadprůměrná	průměrná	podprůměrná
teoretické znalosti	x		
praktické zkušenosti		x	
podkladové materiály (vstupní data) a jejich zpracování		x	

kriteria hodnocení práce	úroveň
---------------------------------	---------------

	nadprůměrná	průměrná	podprůměrná	nelze hodnotit
stupeň splnění cíle práce	x			
samostatnost při zpracování tématu				x
logická stavba práce	x			
práce s českou literaturou včetně citací	x			
práce se zahraniční literaturou včetně citací			x	
adekvátnost použitých metod		x		
hloubka provedené analýzy	x			
stupeň realizovatelnosti řešení		x		
formální úprava práce (text, grafy, tabulky)	x			
stylistická úroveň	x			
nároky BP na podkladové materiály, konzultace, průzkumy ...	vysoké	průměrné	nižší	nejsou
		x		
použití analýz, matem. statistických a jiných metod, komparací apod.	ve velké míře	přiměřené	částečné	absentuje
	x			
využitelnost námětů, návrhů a doporučení k řešení problému	ve větší míře	částečná	nižší	nevyužitelnost
	x			
obsah a relevantnost příloh v textu či příl. části BP (tabulky, grafy, propočty apod.)	vysoce funkční	funkční	méně funkční	neuspokojivé
		x		

Odpovídající hodnocení jednotlivých hledisek označte:	x
---	---

Připomínky:

- Str. 9: V definici Euklidovského skalárního součinu není uvedena váhová funkce, přitom dále v práci (např. na str. 10) se s ní pracuje.
- Str. 11: Na levé straně vztahu (1.3) by podle mého názoru měl být výraz $G_{j+1}(x)$ namísto $G_{j-1}(x)$.
- Str. 19: Na základě Odvození 2 a Odvození 3 autorka tvrdí, že vektor \mathbf{Y} má n -rozměrné normální rozdělení. Dle mého názoru ale autorka v Odvození 2 a Odvození 3 vyjádřila pouze vektor středních hodnot a kovarianční matici jakéhosi neznámého rozdělení pravděpodobnosti. Tvrzení, že se jedná o *normální* rozdělení, by mělo být zdůvodněno zcela jinak.
- Str. 20: Ve vyjádření výrazu $Q(\epsilon)$ je téměř na konci (za poslední transpozicí) závorka navíc.
- Str. 22, 3. a 4. řádek shora: Formulace „...čím blíže je hodnota R^2 rovna 1“ se mi nelíbí. Hodnota nemůže být méně nebo více rovna nějaké hodnotě. Asi bylo myšleno „...čím blíže je hodnota R^2 hodnotě 1“.
- Str. 22, 5. a 6. řádek shora: Uvedené tvrzení je dost zvláštní, neboť hodnota $S(\mathbf{Y})$ je vždy větší nebo rovna hodnotě $Q(\epsilon)$, přičemž rovnost nastává v praxi jen velmi zřídka. Dokázala by autorka zkonstruovat takový datový soubor, aby neplatilo $S(\mathbf{Y}) > Q(\epsilon)$?
- Str. 24: Vysvětlení významu hodnoty $\Pr(|t| > t)$ je trochu matoucí, pro tuto hodnotu se přitom ve statistice používá jednoznačný termín p -hodnota. Podobně je nepřesně vysvětlen pojem Adjusted R-squared, neboť se nejedná o upravenou hodnotu testového kritéria, ale o upravenou hodnotu koeficientu determinace.
- Str. 27: Tvrzení, které se dokazuje v odvození 4, mohlo být zformulováno do podoby věty a náležitě označeno. Takto se v textu ztrácí.
- Str. 37: Tvrzení, že hodnota reziduálního součtu čtverců rozhoduje o vhodnosti použitého modelu, není správné. Pro tyto účely je třeba využít koeficient determinace.
- Str. 40: Místo odkazu na vztah (3) by měl být uveden odkaz na vztah (3.1).

Celkové hodnocení:

Práce se zabývá jednou ze základních statistických metod, která je na různé úrovni matematické abstrakce popsána v mnoha učenicích. I přesto považuji práci za přínosnou, neboť je v ní kromě tradičního přístupu ukázána i zajímavá souvislost metody nejmenších čtverců s Gramovou-Schmidtovou ortogonalizací. Z hlediska teoretické náročnosti patří zvolené téma mezi náročnější (s ohledem na studovaný bakalářský studijní program). Samotné odvození metody nejmenších čtverců je provedeno velmi pečlivě, přesto jsem našel několik nesrovnalostí, které uvádím výše. Grafy vhodným způsobem doplňují jednotlivé části výkladu a jsou také zpracovány velmi pečlivě.

V praktické části autorka demonstruje metodu nejmenších čtverců na několika příkladech. Pozitivně vnímám zejména využití aktuálních ekonomických dat, která jsou pro čtenáře jistě zajímavější než například smyšlená data. Při zpracování dat byl využit volně dostupný software Octave a R, čtenář si proto může jednotlivé výpočty snadno ověřit.

Dalším kladem práce je i velmi kvalitně provedená sazba v systému TeX, oceňuji zejména vzorný zápis maticových rovnic, volbu vhodných symbolů a dodržování standardů matematické sazby.

Otázka k obhajobě:

1. Na straně 19 autorka tvrdí, že Odvození 2 a Odvození 3 implikuje fakt, že vektor \mathbf{Y} má n -rozměrné normální rozdělení. To ale dle mého názoru není pravda. Tvrzení, že se jedná o *normální* rozdělení, by mělo být zdůvodněno jinak. Prosím autorku, aby správné zdůvodnění tohoto faktu prezentovala při obhajobě práce.
2. Na straně 22 se autorka zabývá vztahem mezi celkovým a reziduálním součtem čtverců. Dokázala by autorka zkonstruovat takový datový soubor, aby neplatilo $S(\mathbf{Y}) > Q(\epsilon)$? Jaká by v takovém případě byla hodnota koeficientu determinace?

Práci doporučuji k obhajobě.

Oponent bakalářské práce:

Jméno, tituly: RNDr. Michal Čihák, Ph. D.

Podpis:

V Hradci Králové dne 11. 6. 2016