Posudek oponenta bakalářské práce

Název: Kombinování odhadů
Autor: Valeriya Plotnikova

Shrnutí obsahu práce

Tématem předložené práce je odhad reálného parametru jakožto lineární kombinace (speciálně jakožto váženého průměru) předem daných odhadů tohoto parametru tak, aby výsledný odhad měl co nejmenší střední kvadratickou chybu (MSE). V úvodu práce je popsán uvažovaný problém spolu se strukturou práce. První kapitola je věnována teoretickému odvození optimálního odhadu v případě váženého průměru (tj. odvození vah, s nimiž má vážený průměr daných odhadů nejmenší MSE) pro jednorozměrný parametr a na příkladu demonstrován postup pro vícerozměrný parametr. Druhá kapitola obsahuje dva příklady, na které je metoda kombinování odhadů pomocí váženého průměru použita (odhad σ^2 na základě náhodného výběru z $\mathrm{N}(0, \sigma^2)$ a odhad λ na základě náhodného výběru z $\mathrm{Po}(\lambda)$). Ve třetí kapitole (a ve dvou dodatcích) jsou pak pro tyto dva příklady uvedeny simulace a hodnocení jednotlivých odhadů podle relativní střední kvadratické chyby a relativního vychýlení. Shrnutí poznatků a uvedených výsledků lze nakonec práce nalézt v závěru.

Celkové hodnocení práce

Téma práce. Téma práce je přiměřené a studentka jej zpracovala v souladu se zadáním.

Vlastní příspěvek. Studentka zpracovala zejména článek Lavancier a Rochet (2016) a přiblížila tak českému čtenáři problematiku kombinování odhadů.

Hlavním přínosem předložené práce je podle autorky odvození optimálního vektoru vah pro vážený průměr odhadů, který je v literatuře na toto téma pouze uveden bez detailního důkazu. Toto odvození se zdá být v pořádku, byť vyžaduje jisté doplnění (viz připomínku č. 1 níže).

Dále studentka vypracovala příklady, na kterých je ukázáno, jak lze metodu kombinování odhadů použít v praxi, což je jistě užitečné. U přehledové práce bych ovšem uvítal poněkud pestřejší škálu příkladů, ze kterých by čtenář mohl získat ucelenější pohled (viz připomínku č. 2 a otázku č. 2 níže).

Matematická úroveň. Matematická úroveň práce je vcelku obstojná. Text je psán víceméně srozumitelně, použité pojmy jsou korektně definovány a značení je ve většině případů řádně zavedeno. Některá tvrzení ovšem vyžadují jisté doplnění (viz připomínky č. 1 a 2 níže) případně drobnou opravu (viz připomínku č. 3).

Práce se zdroji. Dle mého názoru by se zdroji v práci mělo být zacházeno poněkud pečlivěji.

Hlavní referencí pro teoretickou část práce byl článek Lavancier a Rochet (2016). Konkrétně str. 3 a 4 v práci odpovídají sekci 2.1 v tomto článku a byť je článek citován v textu jako reference k pojmu „oracle“ (str. 3, ř. 10 zdola) a jako příklad literatury, která „se zabývá problémem kombinování odhadů“ (str. 4, ř. 7), domnívám se, že by měl být také uveden na začátku str. 3 jako zdroj, podle kterého se postupuje při odvození s tím, že některé kroky jsou podrobněji rozepsány. Podobně pro podkapitoly 1.3 a 1.4 v práci.

Zdrojem některých podpůrných teoretických výsledků, které jsou využity v práci ve druhé kapitole (např. definice 1 a 2, případně věta 5) je monografie Anděl (2007). Tyto výsledky jsou citovány, avšak obsahují symboly, které jsou v této knize zavedeny na předchozích stranách. Bohužel však v předložené práci již vysvětleny nejsou (např. pojem „statistika“ nebo symbol Ω v definici 2 a větě 5). Navíc odkazy na přesné stránky mnohdy nejsou správně (např. definice 1 suficientní statistky se v knize Anděl (2007) nachází na str. 124 a nikoliv na str. 136). Podobně u převzatého příkladu na str. 16 v práci zcela chybí odkaz na první část příkladu ($\boldsymbol{S}$ je suficientní statistika), která se v monografii Anděl (2007) nachází na str. 125. Odkaz na druhou část ($\boldsymbol{S}$ je úplná statistika) je již v pořádku.

V neposlední řadě se domnívám, že by práce měla obsahovat větší přehledovou část s odkazy na literaturu, která se zabývá danou problematikou a příbuznými tématy. V tomto směru předložená práce příliš opory čtenáři neposkytuje.

Formální úprava. Po formální stránce je práce v pořádku až na pár drobných detailů (např. konzistentní používání slova „váh“ místo „vah“ ve 2. pádě mn. č. a obecněji častějšího nesprávného použití „á“ namísto „a“, použití symbolu „:=“ pro definici objektu ve vzorci (1.5) a u dalších definic nikoliv a místy poněkud nepřehledné výpočty s několika kroky na jednom řádku, viz např. str. 4, ř. 7 a 8 zdola).

Další připomínky

1. V odvození optimálního odhadu v případě váženého průměru odhadů v podkapitole 1.2 na str. 5 je použita metoda Lagrangeových multiplikátorů. Vyřešením rovnic (1.8) a (1.9) nalezneme stacionární body, tedy body podezřelé z extrémů. V dalším kroku je pak nutné ukázat, že nalezený stacionární bod je skutečně bodem extrému. Tento krok v práci ale chybí. Podobné ověření chybí i v jednorozměrném případě v příkladě na str. 13.
2. V práci se na několika místech vyskytují tvrzení, která nejsou podpořena odkazem na literaturu, argumenty nebo příklady. Například na str. 3, ř. 19 se píše

   > „V literatuře se nejčastěji můžeme setkat s následujícím vyjádřením odhadu $\hat{\theta}_\lambda$ parametru θ pomocí lineární kombinace ...“

   což by si jistě zasloužilo odkaz na literaturu. Nebo na str. 12, ř. 14 se tvrdí:

   > „Často se může stát, že není možné vyjádřit matici Σ pomocí nějakého násobku odhadovaného parametru. V této situaci je důležité odhadnout Σ pomocí $\hat{\Sigma}$, obsahující konzistentní počáteční odhad.“

   Přitom ani v jednom z hlavních příkladů (podkapitoly 2.2 a 2.3) nebylo odhad $\hat{\Sigma}$ matice Σ nutné uvažovat a v práci byl nalezen přímo optimální odhad $\boldsymbol{\lambda}^*$. Čtenář by zde jistě ocenil příklad, kde by bylo nutné odhad matice Σ uvažovat.
3. Na str. 15, ř. 10 zdola se zdá, že chybí střední hodnoty. Konkrétně, je-li matice Σ daná vzorcem (2.5), potom vektor $\Sigma^{-1}\mathbf{1} = (n/\lambda, 0)^T$ a $\lambda = \mathsf{E}\,\bar{X}_n$.

Otázky

1. Existence a jednoznačnost optimálního vektoru vah $\boldsymbol{\lambda}^*$ je úzce spjatá s volbou množiny Λ, přes kterou se minimalizuje. Zde je v úvodní teoretické části práce poněkud zmatek. Na str. 3, ř. 22 se píše, že

   > „ (...) Λ je nějaká neprázdná podmnožina prostoru $\mathbb{R}^k$. Vektor $\boldsymbol{\lambda}$ v tomto případě získáme minimalizací střední čtvercové (resp. kvadratické) chyby (MSE, mean squared error)
   > $$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda}\in\Lambda} \mathsf{E}\,[(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)^T].\text{“}$$

Musí v tomto případě optimální vektor vah $\boldsymbol{\lambda}^*$ existovat (a případně být jednoznačně určen)? Na str. 4, ř. 12 se zase píše, že

> „Je-li zapotřebí přidat další omezení pro váhový vektor, například nezápornost vah, platí uvedená metoda pro všechny uzavřené neprázdné podmnožiny množiny
>
> $$\Lambda_{max} := \{\boldsymbol{\lambda} \in \mathbb{R}^k : \boldsymbol{\lambda}^T \mathbf{1} = 1\}.\text{“}$$

Tento případ zřejmě zahrnuje i, pro práci zásadní, případ $\Lambda = \Lambda_{max}$. Existuje v tomto případě optimální vektor vah a je jednoznačně určen? A posledně na str. 4, ř. 15 zdola se píše, že

> „Optimální vektor vah, (...), má potom tvar
>
> $$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda}\in\Lambda} \mathsf{E}\left[(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)^T\right] = \arg\min_{\boldsymbol{\lambda}\in\Lambda} \boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda},$$
>
> kde $\Lambda \subseteq \Lambda_{max}$ je nějaká neprázdná podmnožina.“

Jak se to má s existencí a jednoznačností optimálního vektoru vah v tomto případě?

2. V obou hlavních příkladech (podkapitoly 2.2 a 2.3) jsou počáteční odhady vytvořeny na základě náhodného výběru. Na str. 3, ř. 9 se však píše

> „Předpokládejme, že máme k dispozici pozorování $\boldsymbol{X} \sim P_\theta$, například náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ pro $n \geq 2$ z parametrického modelu P_θ. Mějme pak posloupnost $\boldsymbol{T} = (T_1, \ldots, T_k)^T$ počátečních odhadů parametru θ, kde $\theta \in \mathbb{R}$.“

Tedy se předpokládá, že odhady $T_1, T_2, \ldots, T_k$ byly vytvořeny na základě pozorování $\boldsymbol{X}$. Může být toto pozorování závislé? Lze uvést nějaký příklad takového použití?

Závěr

Domnívám se, že předložená práce splňuje požadavky kladené na bakalářskou práci a považuji ji za průměrnou.

V Praze dne 7. června 2019

Petr Čoupek
KPMS MFF UK
coupek@karlin.mff.cuni.cz