

POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Název: Vybrané aplikace rekurentní metody nejmenších čtverců
Autor: Timotej Mičuda

SHRnutí OBSAHU PRÁCE

Práce se zabývá rekurentními rovnicemi pro výpočet odhadu metodou nejmenších čtverců. Kromě standardního postupu je uvažován i odhad s fixní počáteční podmínkou a odhad založený na diskontních váhách. Práce obsahuje také ilustraci metody na simulovaných datech a aplikaci na CAPM model pro tři různé akcie.

CELKOVÉ HODNOCENÍ PRÁCE

Téma práce mi přijde na bakalářskou práci vhodné.

Vlastní příspěvek autora spočívá zejména v praktické části, a to v simulační studii a aplikaci metody na reálná data. V rámci teoretické části (kapitola 2) se autor poměrně těsně drží předlohového textu, ale podařilo se mu sjednotit značení a na několika místech výrazy více rozepsat.

Matematická úroveň by mohla být podstatně lepší. Do vlastních formulací matematického textu se autor až na výjimku nepouští. Touto výjimkou je Věta 1, která ale nemá uvedeny žádné potřebné předpoklady, a rozhodně by pro její platnost nestačily předpoklady uvedené na str. 5–6. Předpoklad tvrzení 3 z této věty je zmatený. V části 2.4.2 se tvrdí něco o asymptotickém chování odhadu metodou nejmenších čtverců, argumentuje se použitím zákona velkých čísel. Opět ale chybí jakékoliv předpoklady, které by zajistily jeho platnost. Autor se nad touto pasáží evidentně sám příliš nezamyslel, protože jinak by mu bylo jasné, že je situace poněkud komplikovanější. Na různých místech v textu se píší věci po matematické stránce zcela nesmyslné (viz např. bod 7 níže).

Práce se zdroji. Použité zdroje jsou řádně citovány. V kapitole 2 autor sleduje předlohovou knihu poměrně těsně, text ale formuluje vlastními slovy a doslovně přeložené pasáže jsem nenašla.

Formální úprava poskytuje určitý prostor pro zlepšení, a to jak z pohledu různých drobných nedostatků (viz bod 7 níže), tak i z jazykové stránky. Jako čtenáři mi dost vadily zejména chybějící čárky v souvětích. Důkaz věty 1 je psán formou volného textu, kdy oznamovací věta nezačíná velkým písmenem a nekončí tečkou. Usazení obrázku 3.2 a jeho popisku je dost podivné, stejně tak počet prezentovaných desetinných míst v tabulce 3.3.

PŘIPOMÍNKY A OTÁZKY

1. V kapitole 1 není jasné, co vlastně autor předpokládá o základním regresním modelu. Na str. 4 mluví o nezávislých kopiích, následně na str. 5–6 požaduje něco slabšího. Poslední uvedená odrážka musí platit i pro náhodné regresory, ne jen pro fixní, jinak by řada výrazů nedávala smysl. A pokud autor předpokládá náhodné regresory, měly by první dvě odrážky být formulovány ve formě podmíněných momentů. Dále tyto uvedené předpoklady nezajišťují platnost Věty 1.

Otázka: Jak by měly znít předpoklady Věty 1, aby skutečně platily zde uvedené vztahy 1.–3.?

2. Na str. 6 se tvrdí, že vztah (1.1), tj. $E[Y - g(\mathbf{X})]^2$, lze přepsat do tvaru $\frac{1}{N} \sum_{t=1}^N [Y(t) - g(\mathbf{X}(t))]^2$ a že toto plyne z definice střední hodnoty.

Otázka: Prosim o detailnější vysvětlení, co tím chtěl autor říci. Co by musely být $Y(t)$ a $\mathbf{X}(t)$, aby se daná střední hodnota skutečně rovnala uváděnému výrazu?

3. Rekurentní rovnice pro výpočet odhadu metodou nejmenších čtverců OLS jsou v praxi jistě velmi užitečné. Autor se však v práci tváří, že se jedná o jiný typ odhadu. To je pravda pouze při použití fixních počátečních podmínek. Bohužel toto není nijak odlišeno (značení, název). Řada formulací v textu působí dojmem, že i autor v tomto směru poněkud tápe. Např. v části 2.4.2 se tváří, že zkoumá asymptotické vlastnosti svého odhadu, rozepsán je ale jen klasický odhad metodou nejmenších čtverců OLS. Značení $\hat{\beta}(N)$ ve (2.19) tedy neodpovídá (2.17). Škoda, že se autor nezamyslel nad konzistencí svého odhadu z (2.17), jelikož příslušný krok od OLS je již dost jednoduchý. Přitom v části 3.3 se na tuto pasáž odkazuje v kontextu vlastností odhadu RLS.

Počáteční podmínky jsou zmíněny v části 2.4.1, mě ale stále není jasné, proč bych v praxi dělala cokoli jiného než (2.16), což vede ale na stejný odhad jako je OLS. V práci se argumentuje tím, že chceme odhady od času $t = 1$, ne až od času $t = t_0$, nicméně samotné simulace ukazují, že takto získané odhady jsou pro $t < t_0$ stejně zcela nepoužitelné a naprosto zcestné (a toto navíc platí i pro několik dalších $t > t_0$, viz např. výsledky v části 3.4).

4. V lemmatu 2 se chybně předpokládá, že existují inverze matic A, B, C, D . To by znemožňovalo aplikaci uvedenou za větou.
5. Ve všech použitých příkladech v simulační studii se využívá regresor, který je rostoucí v indexu t , nejčastěji jde o lineární závislost v t . Toto nastavení mi společně se sekvenčním přístupem přijde poněkud nešťastné a odpovídající výsledky jsou značně specifické pro tento případ.

Dále se mi zdá, že v simulacích mohl být více využit potenciál metody založené na diskontování. Příslušné odhady jsou počítány jen pro statický model, což vůbec neposkytuje prostor pro ilustraci toho, co se o nich tvrdí v teoretické části práce (schopnost reagovat dynamicky na změny). Přitom např. data z části 3.2. by se dala použít spolu s prokládáním lineární přímkou a výsledek by mohl být velmi zajímavý.

6. Na str. 29 se tvrdí, že obrázky 3.11 a 3.12 zobrazují rozptyly odhadů.

Otázka: Co je skutečně vyobrazeno na těchto obrázcích? A jak z nich lze vidět rozptyly?

7. Formální nedostatky a drobné chyby: \mathbf{x} na str. 3 dole nemá být tlustým fontem; na str. 5 je \mathbf{X} v jedné větě vektor, v následující matici; n v (1.8) má být N ; $X^T X$ v tvrzení 3 na str. 8 na má být tlustým fontem; ve větě před (2.24) má být k místo t . Značení $r(\mathbf{X})$ není vysvětleno. Na str. 16 se jako dimenze regresoru zavádí t_0 , ale toto již bylo označeno v kapitole 1 jako d . Značení a, b na str. 22 není zavedeno. Na str. 8 v důkazu tvrzení 1. chybí v jednom výrazu podmínění \mathbf{X} a ε má být tlustým fontem; na str. 11 by namísto $\beta(t)$ mělo všude být $\hat{\beta}(t)$, protože $\beta(t)$ nedává smysl; ve (2.11 a) by mělo být $\hat{\beta}$, v (2.15 c) je chyba ve jmenovateli. Velikost závorek by měla být přizpůsobena výrazu uvnitř a na konci řádků by se neměly vyskytovat jednopísmenné předložky.

Příklady nesmyslných tvrzení z textu:

- 1. odstavec na str. 1) je naprosto zmatený a plný tvrzení, která neplatí.
- Na str. 15 autor tvrdí, že vektorový prostor je matice R .

ZÁVĚR

Práci doporučuji uznat jako bakalářskou.

RNDr. Šárka Hudecová, Ph.D.
KPMS MFF UK
V Praze dne 28. srpna 2023