

Oponentský posudek na disertační práci

Autor: Vladimír Chýna

Název: Grafické modely pro analýzu spojitých finančních dat

Problémy vzájemné závislosti v náhodném světě je problém, který neustále přitahuje pozornost.. Korelační koeficient, který je nejvíce známý a intuitivně pochopitelný, však měří jen míru lineární závislosti a navíc může dávat i klamné výsledky. Proto byla a je snaha najít jiné způsoby měření. Na nelineární závislosti byl vytvořen mocný aparát copul, na odstranění „falešných“ korelací se stále více prosazují modely založené na teorii grafů. Právě těmito se problémy se zabývá předložená práce.

Autor rozdělil práci do jedenácti kapitol.. V úvodní kapitole jsou shrnuty základní problémy vyskytující se s pojmem korelace. Autor na několika příkladech prezentuje situace, kdy použití klasického korelačního koeficientu může vést k k nesmyslným či paradoxním závěrům. Bohužel některé příklady nejsou zvoleny nejvhodněji. Příklad s měsícem a houbami není onou závislostí, kterou autor avizuje na začátku odstavce, kdy veličina A ovlivňuje veličiny B a C (a pak zjistíme falešnou korelaci mezi Ba C), ale závislostí typu A ovlivňuje B a B ovlivňuje C. Rovněž v příkladu nelineární závislosti měl autor zvolit raději funkci s lichým exponentem, aby byla zachována jednoznačnost.

Kapitola 2 je věnována vysvětlením některých pojmů z teorie grafů, které jsou dále používány. Osobně bych tuto kapitolu viděl raději jako dodatek, neboť je exkurzí do jiné oblasti matematiky a slouží spíše jako pomocný aparát. Ale to jen můj osobní názor. Totéž platí i o následující kapitole 3. Dát kapitole název „ Základní pojmy ze statistiky“ nepůsobí v disertační práci příliš dobře.

V kapitole 4 autor přechází k VAR modelům, které jsou již chápány (alespoň částečně) v souvislosti s teorií grafů, a tak se buduje aparát pro aplikace. Podobně i v kapitolách 5 a 6, kde jsou prezentovány gaussovské grafické modely a vlastnosti entropie a divergence.

V kapitole 7 autor přistupuje k analýze vstupních dat. Na příkladu odvětvových indexů českého kapitálového trhu jsou ukázány některé testy základních předpokladů (testy nezávislosti, testy normality). Jedná se ovšem o všeobecně známé procedury. V další kapitole se autor vrací opět k metodice teorie grafů, uvádí metodu *deviace*, pomocí níž lze ověřit, zda daný grafický model správně popisuje skutečnou strukturu dat. Metoda je opět aplikována na odvětvové indexy.

Devátá část pak nabízí postup, jakým způsobem vybrat grafický model. Je uvedeno několik algoritmů a jejich variant s aplikacemi a následná kapitola pak uvádí možné postupy, jak daným datům přiřadit konkrétní VAR model a odpovídající orientovaný graf.

Poslední jedenáctá kapitola se pak věnuje komplexnímu zpracování dat uvedenými postupy. Aplikace jsou provedeny jednak na výše uvedené oborové indexy českého kapitálového trhu, dále závislosti hlavních světových akciových indexů s českým indexem PX 50 a na měnové kurzy.

Hlavní přednost práce vidím v uvedených aplikacích a v odladěných algoritmech. Speciálně zjištění o provázanosti (či neprovázanosti) světových akciových indexů jsou velmi zajímavá a někdy i mírně překvapivá. Autor uvádí ekonomickou interpretaci zjištěných faktů a názorně ukazuje význam používané metodiky. Rovněž je nutno ocenit i vytvořené algoritmy, které vůbec nejsou elementární.

Práce sice neobsahuje příliš vlastních teoretických výsledků (snad jen ty věty, u kterých je uveden důkaz), ale to asi ani nebylo cílem. Objevují se občas drobné nepřesnosti či překlepy (např. na str. 32 místo Φ_1, \dots, Φ_k má být Φ_1, \dots, Φ_p), ale ty nesnižují kvalitu. Možná bylo i vhodné se stručně zmínit i o metodách nelineárních korelací (copul), přestože se jedná o dosti odlišnou metodiku. Práce by tak byla zasazena do širších souvislostí.

Otázky

- Jak je to se závislostí hlavních světových indexů Autor na základě analýzy přichází k závěru, že závislost není tak těsná, jak by se zdálo. Na druhou stranu je známo že hlavní obchodníci často s touto závislostí počítají. Platí tedy tato závislost jenom někdy? Pokud ano, je možno ji pomocí dané (případně modifikované) metodiky zjistit?
- Jaké jsou možnosti využití předkládaných algoritmů při zpracování jiných dat, než v práci použitých? Jak rychlá je konvergence algoritmů?

Práci doporučuji k obhajobě