

Oponentský posudek diplomové práce

Název: Náhodné kótované mozaiky s aplikacemi ve výzkumu polykrystalických materiálů

Autor: Bc. Iva Karafiátová

Předložená práce se zabývá vlastnostmi modelů náhodných kótovaných mozaik a představuje možnosti neparametrického testování nezávislosti kótování. Tyto testy jsou pak použity při analýze reálných dat získaných z polykrystalického materiálu, kde kótami jsou orientace jednotlivých zrn. Autorka definuje potřebné testové statistiky, jejich odhady a odvozuje podílovou nestrannost těchto odhadů.

Jazyková i formální úroveň práce je vysoká. Zdroje jsou řádně citovány na patřičných místech. Práce je psaná podrobně a pečlivě, včetně nutného úvodu do fyzikální podstaty daného aplikačního problému a způsobů popisu vlastností zrn. V některých místech by však bylo vhodné doplnit více podrobností a komentářů, konkrétně v případech, kdy autorka zavádí nové, nestandardní objekty (např. volba vah w_v na s. 39, omezení na sousední zrna v definici 52).

Pro lepší přehlednost by také bylo vhodné více využívat odkazy napříč jednotlivými kapitolami práce, například při zavádění modelů závislého kótování odkázat na pasáž o náhodných polích na mříži. Ta jinak působí izolovaně a není jasně uvedeno, proč je do práce zahrnuta.

Podobně kapitola o náhodných polích na spojitě oblasti – proč je v práci zahrnuta? Zjevně proto, aby v části 5.1 mohl být odhadován variogram z reálných dat. Jenže proč je zde vůbec variogram odhadován, když to nesouvisí s testy nezávislosti kótování mozaiky? Stejně tak v části 5.2, proč autorka odhaduje korelační funkci kót z reálných dat? Jak to souvisí se zbytkem práce?

Dále uvádím některé podrobnější komentáře a připomínky:

- V kapitole 3 schází odkazy, odkud jsou definice přebírány.
- V definici 39 je uvedeno, že referenční bod je invariantní vzhledem k posunu, což zřejmě není pravda.
- V úvodu části 3.5 o náhodných mozaikách autorka za mozaiky bere jen ty, které mají buňky tvořené mnohoúhelníky, resp. mnohostěny. Tím ale opomíjí další zajímavé mozaiky, například Johnsonův-Mehlův model.
- V kapitole 4 by bylo žádoucí uvést definici pojmu zaměnitelnost, případně alespoň uvést odkaz. Jde zde o klíčový pojem.
- Část 4.1 o simulačních testech je mírně matoucí v tom, že mluví o nezávislých simulacích za platnosti nulové hypotézy, není zde jasné, zda autorka mluví obecně nebo o kótovaných mozaikách. Dá se ale uhadnout, že mluví obecně, protože u mozaik bude dělat neparametrické testy a tedy se nedá mluvit o simulacích. Dále by bylo vhodné explicitně uvést, že pro jednorozměrnou testovou statistiku autorka popisuje oboustranný test.
- V posledním odstavci sekce 4.2, opravdu autorka počítala průměrné hodnoty na subintervalech (jak je uvedeno v části 4.1.2) nebo jednoduše vyhodnotila hodnoty vážené empirické distribuční funkce v diskrétních bodech?
- Na s. 45 autorka uvádí, že *hladina testu je více méně dodržena*. Bylo by vhodné toto tvrzení formalizovat, například s využitím kvantilů vhodného binomického rozdělení, případně s přihlédnutím k problému mnohonásobného porovnávání.
- Dole na s. 47 autorka uvádí k jako celé číslo, přestože v definici bylo $k \in [0, 1]$.
- Na s. 50 autorka komentuje rozdílnost výsledků pro obě mozaiky. Uvádí však pouze porovnání, v čem se mozaiky liší, ne vysvětlení, co rozdíly způsobuje.

- K obrázku 5.7(b) – bylo by vhodné komentovat, zda pro jiné počty skupin by výsledky dopadly stejně/jinak.
- V závěru práce autorka komentuje, že *nulová hypotéza byla zamítnána méně často, než bylo třeba*. Bylo by vhodné závěr formulovat přesněji, zde není jasné, zda autorka mluví o konzervativnosti testu nebo o malé síle testu.

Dotazy a připomínky, které mohou zaznít v průběhu obhajoby:

- Čím je motivováno omezení se na sousední zrna v definici 52? Zjevně souvisí s definicí 53. Měly by tyto definice smysl i bez omezení na sousední zrna?
- Jsem přesvědčen, že v tvrzení 11 na s. 32 schází jeden předpoklad na vlastnosti váhové funkce, bez kterého důkaz v této formě nefunguje, konkrétně pátá rovnost v první části důkazu, a podobně v druhé části důkazu. Který to je?
- Model podmíněně nezávislého kótování je velmi specifický. Dává takový model smysl pro praxi? Model (zcela) nezávislého kótování jde chápat jako benchmark, hypotézu, kterou v úvodu analýzy dat zamítneme, což zdůvodní použití složitějších a zajímavějších modelů. Lze podobně chápat modely podmíněně nezávislého kótování? Co když bude nevhodně zvolený počet skupin / budou nevhodně zvolené / nebudou se řídit objemem buněk ale jinou vlastností?
- Čím je motivováno zavedení vah w_v ve vzorci (4.4)? Používají se někde standardně?
- Na konci důkazu lemmatu 15 autorka uvádí, čemu je rovna určitá podmíněná střední hodnota. Tento krok však není úplně zřejmý. Je možné rozebrat toto místo podrobněji?
- V sekci 4.3.1 autorka zavádí model rovnoměrného bílého šumu, viz definici 57. Zřejmě by hodnoty měly být nekorelované či nezávislé? Dávalo by smysl uvažovat jiné než rovnoměrné rozdělení?
- V definici 61 autorka zavádí vážený autoregresní model WAR. Za jakých podmínek je model dobře definovaný? Jsou známy nějaké teoretické podmínky postačující k existenci řešení?
- V posledním odstavci na s. 45 autorka uvádí, že *čím vyšších hodnot nabývá váhová funkce, tím nižší je hodnota parametru ρ* . To se mi jeví jako moc silné tvrzení, protože váhy w_a se mohou změnit škálováním, přičemž na závislostech se nic nezmění. Ovlivnilo by tedy výsledky simulační studie přeskálování na větší/menší mozaiku?
- Dole na s. 48 autorka uvádí, že jsme schopni rozeznat, kdy počet skupin není v nulové hypotéze zvolen správně. Jak? Jak to udělat pro jedna reálná data? Můžeme něco usoudit z rozdílných p-hodnot testů s různými váhami?
- Na s. 56 autorka zmiňuje konfiguraci minimalizující střední disorientaci mezi zrny. Jak se takové uspořádání najde a proč je zajímavé?

Shrnutí:

Předložená práce zpracovává zajímavé a obtížné téma, se kterým se autorka se ctí vyrovnala. Dle mého názoru práce splňuje požadavky kladené na diplomovou práci na MFF UK, proto ji doporučuji přijmout jako diplomovou práci k obhajobě.

V Praze, dne 29. 6. 2020

RNDr. Jiří Dvořák, Ph.D.