

Toruń, 4 października 2024 r.

dr hab. Agnieszka Goroncy, prof. UMK
Katedra Statystyki Matematycznej i Eksploracji Danych
Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika w Toruniu

**Recenzja rozprawy doktorskiej mgr Mateusza Johna
pt. „*Estymacja i testowanie macierzy kowariancji
należących do podprzestrzeni kwadratowych*”**

Recenzowana rozprawa, zgodnie z jej tytułem, dotyczy estymacji macierzy kowariancji należących do podprzestrzeni kwadratowych w standardowym modelu wielowymiarowym, modelu podwójnie wielowymiarowym oraz w modelu krzywych wzrostu, a także testowania wspomnianych macierzy kowariancji w modelach wielowymiarowym oraz podwójnie wielowymiarowym. W całej rozprawie zakładany jest rozkład normalny rozważanych modeli.

Po zapoznaniu się z rozprawą doktorską pana mgra Mateusza Johna stwierdzam, że spełnia ona wymagania stawiane rozprawom doktorskim. Dalsza część tej recenzji zawiera omówienie zawartości rozprawy oraz uzasadnienie tej oceny.

Rozprawa doktorska pana mgra Mateusza Johna składa się z czterech rozdziałów. Wyniki zawarte w rozprawie zostały opisane w dwóch współautorskich artykułach naukowych oraz jednym rozdziale książki, przy czym Autor opublikował do tej pory łącznie cztery prace. Dwa ze zgłoszonych artykułów zostały opublikowane w bardzo dobrych (stypunktowych, zgodnie z aktualnym ministerialnym wykazem czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych) czasopismach z dziedziny statystyki matematycznej. Wszystkie trzy prace stanowiące treść rozprawy opublikowano w wydawnictwie Springer.

W rozdziale pierwszym, który stanowi wprowadzenie do rozważanych zagadnień, zaprezentowano podstawowe pojęcia, definicje i własności operatorów macierzowych, omówiono modele danych, które zostaną poddane analizie w dalszej części pracy. Przedstawione zostały również szczególne struktury kowariacyjne należące do podprzestrzeni kwadratowych w modelu wielowymiarowym (m.in. macierz o strukturze kompletnej symetrii - CS, struktura kołowej macierzy Toeplitza - CT), a także uogólnienia struktur kowariacyjnych w modelu podwójnie wymiarowym (m.in. macierz o blokowej strukturze kompletnej symetrii - BCS, macierz o blokowej strukturze kołowej macierzy Toeplitza - BCT). Autor wprowadził również tło historyczne związane z literaturą dotyczącą testowania struktury macierzy kowariancji, sięgające lat czterdziestych ubiegłego wieku. Omówiono również podstawowe testy, które rozważono w kolejnych częściach pracy, tj.

test ilorazu wiarygodności (LRT), test wynikowy Rao (RST), test Walda (WT), dla których rozkładem granicznym statystyk testowych jest rozkład chi-kwadrat z ustaloną liczbą stopni swobody. Na koniec rozdziału krótko opisano dane rzeczywiste, które posłużyły do ilustracji rozważanych w rozprawie zagadnień. W mojej ocenie rozdział ten jest zasadniczo kompletny i wyczerpuje słownik pojęć (w większości algebraicznych), własności i faktów niezbędnych do dalszych rozważań.

W rozdziale drugim Autor rozpatruje problem estymacji macierzy kowariancji o liniowej strukturze należącej do podprzestrzeni kwadratowej, w modelu wielowymiarowym z nieznaną wartością oczekiwaną, w którym przestrzeń wartości oczekiwanej komutuje z przestrzenią macierzy kowariancji. Do estymacji zastosowano metodę największej wiarygodności. Wyniki tego rozdziału stanowią rozszerzenie wyników Szatrowskiego (*Ann Stat* 8:802-810, 1980), który zakładał, że wartość oczekiwana w modelu jest znana. Uogólnienie to uwzględniające zawarte zostało w twierdzeniach 2.1 oraz 2.2. Zgodnie z tezą pierwszego z nich, estymator największej wiarygodności macierzy kowariancji o pewnej strukturze istnieje i jest przedstawiony w postaci jawnej wtedy i tylko wtedy, gdy przestrzeń macierzy kowariancji jest podprzestrzenią kwadratową. W twierdzeniu 2.2, które uważam za ważny wynik rozprawy, podana jest dokładna postać estymatora największej wiarygodności macierzy kowariancji z zadaną strukturą, którą jest rzut ortogonalny macierzy będącej estymatorem największej wiarygodności macierzy kowariancji bez struktury na przestrzeń liniową, do której należy rozpatrywana macierz kowariancji o zadanej strukturze.

Dalej Autor rozprawy stosuje otrzymany wynik do estymacji szczególnych struktur kowariacyjnych należących do podprzestrzeni kwadratowych w standardowym modelu wielowymiarowym (przestrzeń struktury CS oraz CT), modelu podwójnie wielowymiarowym (blokowa struktura proporcjonalna do macierzy jednostkowej oraz przestrzeń struktury BCS) oraz w modelu krzywych wzrostu (przestrzeń struktury CS), przy założeniu normalności modelu. Słusznie podkreślona została tutaj waga sprawdzenia założenia dotyczącego komutatywności przestrzeni wartości oczekiwanej i przestrzeni macierzy kowariancji w estymacji macierzy kowariancji metodą największej wiarygodności. Warto w tym miejscu wspomnieć, że wyniki otrzymane przez pana mgr. Mateusza Johna i jego współautorów są zgodne z wynikami uzyskanymi wcześniej przez innych badaczy. Mimo, że omawiany rozdział dotyczący estymacji jest najmniejszy objętościowo, jest on istotny w dalszych rozważaniach. Estymatory macierzy kowariancji bez struktury są bowiem niezbędne do wyprowadzenia statystyk testowych analizowanych w kolejnej części pracy.

Dwa ostatnie rozdziały rozprawy dotyczą testowania hipotez dotyczących struktur macierzy kowariancji należących do podprzestrzeni kwadratowych, w standardowym modelu podwójnie wielowymiarowym.

W rozdziale trzecim Autor skupia się na hipotezie dotyczącej testowania struktury kowariacyjnej, należącej do podprzestrzeni kwadratowej o strukturze blokowej uwzględniającej zarówno zależności wewnątrz-, jak i międzyklasowe w modelu. Zakłada się przy tym, że zależność wewnątrzklasowa posiada pewną strukturę kowariacyjną, należącą do komutatywnej podprzestrzeni kwadratowej. Autor skupia się na konstrukcji dwóch

testów weryfikujących strukturę zależności między- i wewnątrzklasowych: teście ilorazu wiarygodności (LRT) oraz na teście wynikowym Rao (RST). Wspomniano również o fakcie pominięcia w rozważaniach testu Walda, który jest problematyczny w precyzyjnym określeniu hipotezy złożonej z uwagi na ogólność struktury kowariacyjnej. Postać statystyk testowych dla testów LRT i RST została wyprowadzona w twierdzeniach 3.2 oraz 3.5, odpowiednio. Ważną częścią tego rozdziału jest wyznaczenie rozkładu dokładnego statystyki testowej testu LRT przy założeniu prawdziwości hipotezy zerowej. Rozkład ten został przedstawiony w twierdzeniu 3.3, którego dowód wymagał niemałego wysiłku i sprawności rachunkowej.

Autor zbadał również asymptotyczną zbieżność rozkładów statystyk testowych obu testów (dokładnego dla LRT i empirycznego dla RST) przy założeniu prawdziwości hipotezy zerowej do granicznego rozkładu chi-kwadrat za pomocą symulacji numerycznych. Wyniki przeprowadzonych symulacji są zgodne z teorią, tj. wraz ze wzrostem liczebności próby można zaobserwować, że rozkłady obu statystyk dążą do granicznego rozkładu chi-kwadrat. Autor rozprawy zauważa ponadto, że zbieżność testu Rao jest szybsza niż testu ilorazu wiarygodności we wszystkich rozpatrywanych przypadkach. Ponadto, nawet dla małej próby empiryczny rozkład statystyki RST jest bardzo zbliżony do rozkładu granicznego. Nie obserwuje się jednak takiego zachowania w przypadku testu LRT. Jako konkluzję, słuszną zresztą, Autor podaje sugestię, aby w przypadku mało licznych prób dla testu LRT do podjęcia decyzji wykorzystać rozkład dokładny, z uwagi na to, że występują w tym przypadku istotne różnice między rozkładem dokładnym a rozkładem granicznym.

W ostatniej części omawianego rozdziału porównano moce testów LRT i RST w zależności od rozbieżności między hipotezą zerową a alternatywną oraz pod kątem wielkości próby, wykorzystując do tego entropijną funkcję straty. Ponieważ moc testu odpowiada prawdopodobieństwu odrzucenia fałszywej hipotezy zerowej, moc empiryczna obliczona została jako procent odrzuceń hipotez zerowych na tle wszystkich przeprowadzonych testów. Autor obserwuje rosnący trend mocy testów w zależności od wielkości rozbieżności, przy czym odchylenia między mocami są większe w przypadku testów RST niż LRT. Ponadto, moc wykazuje tendencję rosnącą wraz ze wzrostem liczebności próby, co również nie jest zaskakujące. Konkludując, oba testy są porównywalne, ich moce nie są istotnie różne i nie da się tym przypadku wyznaczyć testu „zwycięskiego”.

Poza symulacjami, przeprowadzona została również analiza przykładu rzeczywistego, dotyczącego płatków pewnego kwiatu. Przeprowadzony został test dotyczący struktury kołowej macierzy Toeplitza zależności wewnątrzklasowej między płatkami na pojedynczym kwiecie, przy czym do opisu zależności międzyklasowej zastosowano strukturę CS, równoważną w tym przypadku strukturze CT. Przeprowadzono testy trzech hipotez zerowych (różniące się wymiarami odpowiednich przestrzeni rozważanych w hipotezach zerowych wraz z macierzami bazowymi) z wykorzystaniem statystyk LRT i RST, przy użyciu rozkładu dokładnego statystyki LRT, rozkładu empirycznego RST oraz granicznego rozkładu chi-kwadrat obu statystyk. Decyzje nie we wszystkich przypadkach były takie same. Zgodnie z konkluzjami z wcześniej przeprowadzonych symulacji, rozkład dokładny statystyki LRT istotnie się różnił od rozkładu granicznego chi-kwadrat

w przypadku dwóch z testowanych hipotez. W przypadku trzeciej hipotezy, nawet dla tak małej próby jak rozważana, rozkłady statystyk LRT i RST nie różniły się istotnie i były bliskie rozkładowi granicznemu, prawdopodobnie z uwagi na specyfikę rozważanej struktury.

Czwarta, ostatnia część pracy dotyczy testowania niezależności międzyklasowej w modelu normalnym z macierzą kowariancji o pewnej szczególnej strukturze, mianowicie blokowej strukturze kompletnej symetrii (BCS). Skupiono się w niej na dwóch testach: teście RST oraz teście Walda (WT), a celem przeprowadzonych analiz było porównanie wyników z testami wcześniej w literaturze rozważanymi: LRT, testem Fonseca i in.(2008), oznaczanym dalej jako FT, oraz testem Roya (Mardia i in., 1979), oznaczanym dalej jako RLRT. W omawianej rozprawie dokonano również wyprowadzenia postaci statystyki WT (której niewyprowadzona postać, niezgodna z definicją testu Walda była znana od 2018 r.). Autor rozprawy podaje również, jak poprzednio, dokładny rozkład statystyki LRT, która posłuży później do porównania testów w badaniach symulacyjnych. Do badania rozbieżności między hipotezami posłużono się ponownie entropijną funkcją straty. Dodatkowo przeprowadzono badanie odporności rozważanych statystyk testowych na zaburzenie założenia o normalności modelu.

Rozkłady dokładne statystyk testowych testów FT i RLRT są znane, dlatego Autor rozprawy skupił się najpierw na wyznaczeniu dokładnego rozkładu statystyki testu LRT (jednoznacznie identyfikowanego za pomocą funkcji charakterystycznej), z uwagi na wcześniej zaobserwowaną prawidłowość dotyczącą rozbieżności rozkładu dokładnego i granicznego w przypadku małej liczby prób. Z tego powodu uznaję twierdzenie 4.1 za kolejny ważny wynik omawianej rozprawy doktorskiej. Kolejnym ważnym wynikiem tego rozdziału jest twierdzenie 4.2, opisujące postać statystyki testu RST, co nie było banalnym zadaniem. Na koniec przeprowadzono badania symulacyjne badające zbieżność rozkładów dokładnych statystyki LRT oraz rozkładów empirycznych statystyk RST i WT do granicznego rozkładu chi-kwadrat. Otrzymane wyniki potwierdzają, że wraz ze wzrostem liczebności próby rozkłady uwzględnionych statystyk testowych zbiegają do rozkładu granicznego, przy czym ponownie najszybsza zbieżność jest obserwowana w przypadku testu RST.

Badanie odporności rozważanych testów pod kątem zaburzeń normalności rozkładu obserwacji implikuje, że żadna z rozważanych statystyk nie jest odporna na to zaburzenie. Jest to niestety smutna konkluzja, ponieważ w praktyce często obserwuje się skośność danych, mocno zaburzającą ich normalność. W symulacjach posłużono się danymi wygenerowanymi z wielowymiarowego rozkładu t -studenta, gamma oraz jednostajnego.

Analiza mocy testów w zależności od wielkości rozbieżności pomiędzy hipotezami, przeprowadzona na podstawie badań symulacyjnych przy użyciu entropijnej funkcji straty pozwala Autorowi wyciągnąć następujące wnioski. Wraz ze wzrostem rozbieżności, wartości mocy rosną dla testów LRT (rozkład dokładny statystyki testowej), RST i RLRT (rozkłady empiryczne statystyki testowej), z wyjątkiem testów FT oraz WT. Dla dwóch równorozbieżnych struktur BCS moc testu FT istotnie się różni, zaś moc WT jest często niższa niż nominalny poziom istotności. Dlatego w dalszych badaniach słusznie Autor uwzględnia takie testy, których zachowanie jest „prawidłowe”, tj. LRT,

RST i RLRT, przy czym uwzględniono w tym przypadku tylko rozkłady empiryczne statystyk testowych przy założeniu prawdziwości hipotezy zerowej. Wyniki symulacji wskazują na podobne zachowanie mocy testów w zależności od wielkości próby, co ponownie nie pozwala na wyłonienie testu „najlepszego”.

Ostatnia część tego rozdziału dotyczy analizy tych samych danych rzeczywistych dot. płatków kwiatów. Tym razem badano niezależność długości płatków pomiędzy dowolnymi dwoma kwiatami, wykorzystując dokładny rozkład statystyki LRT (dość mała próba), empiryczne rozkłady statystyk RST i WT oraz rozkład graniczny chi-kwadrat, a także dokładne rozkłady FT i RLRT. Dla analizowanych danych żaden z rozważonych testów nie odrzucił hipotezy zerowej mówiącej o niezależności długości płatków dla różnych kwiatów. Warto podkreślić, że mimo, iż LRT nie pozwolił na odrzucenie hipotezy zerowej, p -wartości obliczone na podstawie rozkładu dokładnego i granicznego są dość odległe. Nie ma w tym nic zaskakującego, z uwagi na zaobserwowane już wcześniej wolniejsze tempo zbieżności rozkładu statystyki LRT do rozkładu granicznego w przypadku mało licznych prób. Z praktycznego punktu widzenia, z uwagi na konieczność wykorzystania specjalistycznego oprogramowania (rozkład LRT, RLRT) oraz obciążoność testu WT, Autor rozprawy słusznie sugeruje wykorzystanie wartości granicznych rozkładu chi-kwadrat i testu RST.

Tematyka badań przedstawionych w rozprawie nie jest mi szczególnie bliska i nieco odbiega od moich zainteresowań naukowych, ale po zapoznaniu się z zawartością pracy doktorskiej i skrupulatnej analizie wyników w niej zawartych jestem pod wrażeniem umiejętności prowadzenia badań naukowych przez pana mgr. Mateusza Johna. Analizując dowody twierdzeń zawartych w rozprawie, które są przeprowadzone poprawnie, mogę stanowczo stwierdzić, że jej Autor jest biegły w operacjach rachunkowych wymaganych do uzasadnienia tezy, posiada dobre intuicje matematyczne, co świadczy również o dojrzałości naukowej w posługiwaniu się językiem matematyki.

Uwagi

Moje uwagi, pytania i wątpliwości, szczególnie te, które nie zostały rozwiane przez Autora w treści rozprawy, przedstawione zostały poniżej.

- Dlaczego modele, w których nie zakłada się normalności rozkładu obserwacji nie były rozważane? Symulacje przedstawione w rozdziale 4. sugerują, że odstępstwa od normalności może mocno zaburzyć wnioskowanie. W szczególności, interesująca jest sytuacja, gdy dane pochodzą z rozkładów dyskretnych.
- Ważną częścią rozdziału 3. jest wyznaczenie rozkładu dokładnego statystyki testowej testu LRT przy założeniu prawdziwości hipotezy zerowej. Dlaczego nie przeprowadzono analogicznej analizy dla rozkładu dokładnego statystyki testowej testu wynikowego Rao (Autor bazuje na rozkładzie empirycznym), wiedząc, że w badaniach symulacyjnych zaobserwowano, że nawet dla małej próby empiryczny rozkład statystyki RST jest bardzo zbliżony do granicznego rozkładu chi-kwadrat?

- Dlaczego w rozprawie analizowano tylko jeden rzeczywisty zbiór danych? Czy takie dane są trudno dostępne, albo nie są odpowiednio ustrukturyzowane, czy też problem tkwi w czymś innym?
- Czy analiza rozbieżności pomiędzy rozważanymi hipotezami w porównaniu testów pod kątem mocy mogłaby zostać przeprowadzona przy użyciu innej miary niż entropijna funkcja straty (rozbieżność Kullbacka-Leiblera), a jeżeli tak, to czy to rozważano?
- Rozdział poświęcony estymacji uwzględnia również model krzywych wzrostu. Dlaczego nie rozważano problemu testowania struktury macierzy kowariancji w modelu krzywych wzrostu?
- Najważniejsza moja refleksja dotyczy istoty i zasadności postawionego problemu. Autor bada macierz kowariancji, która odpowiada mierze współzmienności wielu rozważanych cech. Można ją wykorzystać do dalszych rozważań pod warunkiem, gdy zmienne są wyrażone w tych samych jednostkach. Nie jest to niestety sytuacja w praktyce często obserwowana. Jedną z możliwości obejścia tego problemu byłaby standaryzacja cech (wówczas macierz kowariancji odpowiada macierzy korelacji), bądź rozważenie odpowiadającej macierzy korelacji. W pracy brakuje mi uzasadnienia i komentarza odnoszącego się do tej kwestii.

Podsumowanie

Oceniana przeze mnie rozprawa doktorska została napisana starannie i przejrzyście, jej Autor posługuje się naukowo dojrzałym językiem i poprawnie przeprowadził rozumowania matematyczne opisane w rozprawie. Numeracja wzorów oraz ich wykorzystanie są poprawne, bibliografia została dobrana starannie, a wszystkie jej pozycje zostały właściwie wykorzystane w treści pracy. W rozprawie znajduje się co prawda kilka drobnych usterek, ale nie mają one wpływu na jej wartość merytoryczną, m.in. literówki (m.in. „operaotr”), pozostawianie pojedynczych liter, najczęściej spójników na końcu wierszy, pozostawienie zakończenia dowodu na str. 68, itp. Jednak zarówno moje powyższe uwagi i komentarze, jak i drobne usterki nie mają negatywnego wpływu na wartość merytoryczną rozprawy, którą oceniam wysoko.

Uważam, że cel, jaki postawiła pani promotor rozprawy przed panem mgr. Mateuszem Johnem nie był łatwy do osiągnięcia, przede wszystkim z uwagi na to, że podjęty temat nie jest nowy i ma długą historię. Aby osiągnąć pożądany wynik, należało rozważyć przypadki dotąd nieuwzględnione w literaturze. Wiele pytań, które nasuwały mi się w trakcie czytania pracy, w zasadzie od razu były w jej dalszej części rozwiewane bądź komentowane przez Autora rozprawy. W mojej opinii praca, jaką wykonał pan mgr Mateusz John stanowi porządny, matematycznie poprawny i ważny wkład w teorię wnioskowania statystycznego na temat macierzy kowariancji należących do podprzestrzeni kwadratowych w normalnych modelach wielowymiarowych.

Konkluzja

Podsumowując, w mojej ocenie przedstawiona mi do recenzji rozprawa doktorska pana mgr. Mateusza Johna pt. „*Estymacja i testowanie macierzy kowariancji należących do podprzestrzeni kwadratowych*” spełnia wymagania stawiane pracom doktorskim w dyscyplinie matematyka. Recenzowana rozprawa stanowi oryginalne rozwiązanie problemu naukowego, jakim jest problem estymacji i testowania macierzy kowariancji należących do podprzestrzeni kwadratowych w modelach wielowymiarowych, dowodzi ogólnej wiedzy teoretycznej Autora pracy w zakresie matematyki oraz umiejętności samodzielnego prowadzenia badań naukowych. Wnoszę zatem o dopuszczenie pana mgr. Mateusza Johna do dalszych etapów postępowania o nadanie stopnia doktora w dyscyplinie matematyka.

Gerona Aguiar