Opinia o rozprawie doktorskiej mgr. Rafała Wieczorka pt. „ Pomiar optymalny w kwantowej teorii decyzji statystycznej”

prof. dr hab. Adam Paszkiewicz

Bayesowską teorią podejmowania deczji można określić jako połączenie teorii gier i teorii prawdopodobieństwa. Załóżmy, że całość naszych obserwacji to ustalenie punktu x w przestrzeni $\mathfrak{X}$ wyników pomiarów. Wiemy wstępnie, że punkt x podlega rozkładowi losowemu ρ_i z prawdopodobieństwem π_i dla $i \in \mathbb{N}$, oczywiście $\sum \pi_i = 1$. Należy opisać, zależną jedynie od x, ruletkę dającą wynik $j \in \mathbb{N}$ z prawdopodobieństwem $\mathcal{M}_j(x)$, oczywiście $\sum_{j\in\mathbb{N}} \mathcal{M}_j(x) = 1$. Prawdopodobieństwo, że j będzie równe i powinno być największe. Należy więc maksymalizować prawdopodobieństwo detekcji

$$\mathbb{P} = \sum_{ij} \pi_i \int_{\mathfrak{X}} \mathcal{M}_i(x)\rho_i(x)\mathrm{dx}.$$

Wynik j dany przez ruletkę interpretujemy jako decyzję, że nieznanym wstępnie rozkładem jest ρ_j. Naturalnym uogólnieniem problemu detekcji jest problem minimalizacji średniej straty. Jeżeli w opisanej sytuacji podejmujemy decyzję j, to ponosimy stratę $L(i,j)$. Należy zminimalizować wielkość

$$r = \sum_i \pi_i \int_{\mathcal{H}} \sum_j L(i,j)\mathcal{M}_j(x)\rho_i(x)\mathrm{dx}.$$

W powszechnie akceptowanej aksjomatyce mechaniki kwantowej rozkład ρ_i na $\mathfrak{X}$ zastepuje stan, czyli dodatni, normalny, unormowany funkcjonał, ρ_i na ustalonej algebrze von Neumanna $\mathfrak{M}$. Zależną od x rultkę zastepuje pomiar, rozumiany jako układ nieujemnych operatorów $\mathcal{M}_i \in \mathfrak{M}$, spełniający $\sum_i \mathcal{M}_i = 1$. Poszukujemy pomiaru, który maksymalizuje prawdopodobieństwo detekcji

$$\mathcal{P}(\pi, \mathcal{M}) = \sum_i \pi_i\rho_i(\mathcal{M}_i)$$

a ogólniej, minimalizuje średnią stratę

$$r(\pi, \mathcal{M}) = \sum_i \pi_i \sum_j L(i,j)\rho_i(\mathcal{M}_j).$$

Wyniki przedstawione w opiniowanej rozprawie obejmują znaczny obszar wspomnianej kwantowej analizy bayesowskiej. Temat wymusza ogólny charakter badań. Wymagają one wyspecjalizowania się w nieelementarnej teorii algebr operatorowych. Obszar pytań, który należy atakować jest stosunkowo wąski, podyktowany interpretacjami statystycznymi. Badania mają trochę charakter fizyki matematycznej. Autor musiał prześledzić różnorodną literaturę. Jest tam kilka zaskakujących pomysłów, lecz także przeliczenia bez pełnego aksjomatycznego umocowania. Wiele wyników w rozprawie to uogólnienia spostrzeżeń dotyczących skończenie wymiarowej przestrzeni Hilberta, innym razem autor podaje nowe dowody lub dowody prościej zapisane. Praca jest duża objętościowo (ok. 40 stron dowodów). Jej ocena wymaga prześledzenia nowych elementów.

Ostatecznie wszystkie 6 rozdziałów uznaję za cenne, a w szeczólnie twórczy i wartościowy jest rozdział 8. Przełomowy we wcześniejszej literaturze był opis pomiaru maksymalizującego prawdopodobieństwo detekcji w przypadku stanów czystych (Bielavkin, 1975 r.). W twierdzeniu 8.1 autor podaje analogiczny wynik w nieskończonym wymiarze. Podany jest ogólny wzór opisujący optymalny pomiar $\mathcal{M}$ (założenia wymagają oczywistej poprawki dotyczącej stanów ρ_i). Wzór na $\mathcal{M}$ jest nieefektywny, $\mathcal{M}$ pojawia się jako rozwiązanie pewnego równania. Można jednak rozwinąć teorię oszacowań prawdopodobieństwa detekcji. Ostateczny wynik, twierdzenie 8.3 i wniosek 8.4 są eleganckie i zaskajujące. Największym osiągnięciem w tym rozdziale jest opis asymptotyki skierowanego zbioru układów stanów $\psi^\gamma = (\psi_1^{(\gamma)}, \psi_2^{(\gamma)}, \ldots)$. Autor bada kiedy zależne od γ prawdopodobieństwo detekcji $\mathbb{P}(\pi, \psi^{(\gamma)})$ zbiega do 1, dla indeksu γ, przebiegającego zbiór skierowany. Wskazany jest jedynie warunek dostateczny

$$\lim_\gamma \sum_j \| \psi_j^{(\gamma)} - \psi_j \| = 0.$$

Ostatecznym wynikiem są twierdzenia 8.10 i wniosek 8.11. Podają one asymptotykę prawdopodobieństwa detekcji dla „dużych" ψ. Metody sprowadzają się do geometrii przestrzeni Hilberta i słabej zawartości kuli w przestrzeni Hilberta. Dowody są długie, nietrywialne, zręcznie rozbite na lematy.

Rozdział 3. omawia istnienie pomiaru minimalizującego średnią stratę. Podany jest udoskonalony dowód znanego twierdzenia 3.1. Mówi ono, że pomiar optymalny istniej, jeżeli tylko funkcja straty $L(i,j)$ spełnia założenia uniemożliwiające straty nieskończenie ujemne. Elegancja dowodu jest uzyskana przez badanie ciagów stanów poprzez użycie przeliczalnej sumy identycznych kopii algebry von Neumanna $\mathfrak{M}$. Pięknym dopełnieniem teorii dostrzeżonym przez Doktoranta jest twierdzenie 3.5.

Mówi ono, że choć istnieją być może liczne optymalne pomiary $(\widetilde{\mathcal{M}}_j, j \in \mathbb{N})$, to bardzo użyteczny funkcjonał Lagrange'a

$$\varphi = \sum_j \widetilde{\mathcal{M}}_j \sum_i \pi_i L(i,j)\rho_j$$

jest tylko jeden. Elegancka jest końcówka dowodu. Jeżeli funkcjonał $\varphi_i - \psi$ jest dodatni i znika na dotatnim operatorze $\mathcal{M}_j$ to funkcjonał $\mathcal{M}_j(\varphi_i - \psi)$ znika na każdym opertorze (wystarczyły elementarne przekształcenie i nierówność Swcharza).

Rozdziały 4. i 6. warto omówić łącznie. Minimalizacja średniej straty sprowadza się naturalnie do minimalizacji sumy $\sum_{j\in\mathbb{N}} \varphi_j^c(\mathcal{M}_j)$ dla pewnych nieujemnych funkcjonałów φ_j^c. Dalszą strategię wskazują twierdzenia 4.2 oraz 4.4. Nieujemność funkcjonałów φ_j^c pozwala sprawnie udowodnić, że optymalny pomiar $(\mathcal{M}_j, j \in \mathbb{N})$ składa się ze wzajemnie ortogonalnych projekcji. Jest on jednoznacznie wyznaczony w ramach odpowiednio obciętej algebry. Autor ogranicza się do sytuacji, gdy nośniki funkcjonałów są wzajemnie ortogonalne (warunek 6.1), ale dowód wymaga ciekawych pomysłów również przy tym założeniu. W końcówce dowodu twierdzenia 6.3 identyczność projekcji $\mathcal{M}_{j0} = \mathcal{N}_{j0}$ w dwóch optymalnych pomiarach $(\mathcal{M}_j, j \in \mathbb{N})$, $(\mathcal{N}_j, j \in \mathbb{N})$ uzyskuje się z algebraicznych własności operatora $\frac{1}{2}(\mathcal{M}_{j0} + \mathcal{N}_{j0})$. Ładny jest lemat 5.5 wiążący wierność sumy funkcjonałów z wiernością rodziny składników.

W rozdziale 5., za cenne, samodzielne osiągnięcie Doktoranta uważam zbadanie pomiarów optymalnych dla układów 2-stanowych $(\mathcal{M}_j\ ;\ j = 1, 2)$. Twierdzenie 5.1 stanie się trwałym elementem teorii. Punkt 5.2 jest już tylko interpretacją statystyki klasycznej jako szczególnego przypadku statystyki kwantowej.

Rozdział 7. o nierównościach dla średniej straty i prawdopodobieństwa detekcji jest obszerny i wymagał dużego nakładu pracy. Uzyskane przez autora uogólnienia w punkcie 7.2 to jednak dość formalne zastosowania zaczerpniętych z literatury własności relatywnej entropii Arakiego i własności operatorowej monotoniczności. Nienaturalna jest kolejność twierdzeń 7.9 i 7.10 (twierdzenie 7.10 mówi, że zapis użyty w twierdzeniu 7.9 ma sens). Punkt 7.1 podaje szacowanie dolne średniej straty, wzorowane na pomyśle Qiu z 2008 r. Przy braku komentarza i przykładów rola szacowania jest niejasna. Postępem byłoby dopiero poradzenie sobie z nieskończoną ilością stanów. Bardziej wartościowy wydaje się punkt 7.3. Uogólnione jest tam szacowanie Holewy-Curlandera. Jest to górne szacowanie prawdopodobieństwa detekcji poprzez normę pewnego wektora zbudowanego ze stanów i ich początkowych prawdopodobieństw. Doktorant w pełni opisał, kiedy zachodzi równość (twierdzenie 7.17). Rozumowanie jest długie. Ciekawy jest dowód lematu 7.15 z użyciem przemnożenia tensorowego algebry $\mathfrak{M}$ przez$B(K)$ dla zbadania ciągu funkcjonałów, a dokładniej gestości funkcjonałów.

Przejdę do uwag krytycznych i nasuwających się pytań. Wstępne rozdziały w tego typu pracy powinny być bardziej dopracowane. Autor pominął wyjaśnienia bardziej skomplikowanych pojęć (jak płóskończoność algebry i stanu, silna niezależność, relatywna entropia stanów). Mam wiele sugestii, zauważyłem pomyłki, pechowe jest sformułowanie i początek dowodu twierdzenia 8.1. Mimo to redakcję tak obszernej pracy należy uznać za staranną.

Pytania i sugestie:

1. W definicji 2.3 i paru innych miejscach warto wyjaśnić konwencję działania operatora na stan.
2. Jak rozumieć półskończoność śladu (początek strony 3)? W którym miejscu dowodu jest ona użyta?
3. Kiedy operator jest mierzalny (koniec strony 3)?
4. Czym różnią się warunki (i) twierdzenie 3.1 oraz (ii) twierdzenie 3.3?
5. W warunkach (i), (ii), (iii) tw. 3.4 należy lepiej opisać kwantyfikację $\widetilde{\mathbb{M}}$ oraz φ. Podobna uwaga dotyczy twierdzeń 4.2 i 4.4.
6. Warto wyjaśnić równoważność słabej i σ-słabej ciągłości (strona 11, wiersz 7–8).
7. Jak uzasadnić równość (strona 22, wiersz 4)?
8. W skomplikowanych rachunkach warto pisać $(\tau \otimes \mathrm{tr})$ zamaiast $\tau \otimes \mathrm{tr}$ (strona 33, wiersz 1 od dołu i strona 34).
9. Jak wyjaśnić równość i zawieranie operatorów (strona 34, wiersz 7 i 8)?
10. Należy uzupełnić założenia twierdzenia 8.1. Teza nie zachodzi dla $\psi_1 = \cdots = \psi_n$, $\pi_1 = \cdots = \pi_n - \frac{1}{n}$.
11. Dlaczego z lematu 6.5 wynika odwaracalność $\widehat{\varphi}$ (strona 40, wiersz 2 od dołu)?
12. Jak rozumieć złożenie $\Lambda^{\frac{1}{2}}\ T$ (strona 41, wiersz 3 od dołu)?
13. Jak rozumieć zbieżność $\{\psi_i^{(\gamma)}\}$ (strona 43, wiersz 17 od dołu)?
14. Jak uzasadnić równość (strona 44, wiersz 13)? Jak wybrana jest sieć $\{\gamma'\}$ (strona 44, wiersz 13 od dołu)?

Drobne pomyłki:

strona, wiersz:	jest:	sugeruję:
1, 17	ryzyka ryzyka	ryzyka
3, 11 od dołu	miara	miarą
4, 2 od dołu	Dana jest	W dalszym ciągu pracy dana jest
8, 5	$j = 1, 2, \ldots$	$j = 1, 2, \ldots,$
12, 2	$\widetilde{\mathfrak{N}}$ w $\widetilde{\mathfrak{M}}$	$\widetilde{\mathfrak{N}}$ w $\mathfrak{M}$
16, 5	$\sum_{j=1}^{\infty} \varphi_j(\mathcal{M}_j)$	$\sum_{j=1}^{\infty} \varphi_j^c(\mathcal{M}_j)$
29, 19	$\tau(A \log B) \geqslant$	$\tau(A \log A) \geqslant$
30, 17		ten wzór zanumerować (7.9), zmienić numerację następnych wzorów
36, 3 od dołu	$(\mathcal{M}_j)]^{\frac{1}{2}}$	$(\mathcal{M}_i)]^{\frac{1}{2}}$
36, 2 od dołu	$\mathcal{M}_j^{\frac{1}{2}} =$	$\mathcal{M}_i^{\frac{1}{2}} =$
38, 1		zacząć akapitem
40, 8 od dołu	$(\widehat{\varphi} - \pi_i \rho_i) =$	$(\varphi - \pi_i \widehat{\rho_i})\widehat{\xi_i} =$
40, 6 od dołu	$= \langle \psi_i \,\vert\, \xi_i \rangle \psi_i$	$= \pi_i \langle \psi_i \,\vert\, \xi_i \rangle \psi_i$
41, 12	$= 0,$	$= 0$ dla $i \neq j,$
45, 7 od dołu	$= \sum_i$	$= \lim_{\gamma} \sum_i$

Konkluzje:

Opiniowana rozprawa dowodzi dużych zdolności Doktoranta. Jest wynikiem paru lat pracy i wyspecjalizowania się w ważnej tematyce statystyki oraz uczciwego opanowania metod analizy funkcjonalnej. Uzyskany postep w kwantowej analizie bayesowskiej zinteresuje wielu specjalistów i jest już częściowo opublikowany. Jest to bardzo dobra rozprawa doktorska, wnoszę o dopuszczenie mgr. Rafała Wieczorka do dalszych etapów przewodu doktorskiego.

Łódź, dnia 3.08.2017

A. Paszkiewicz