prof. dr hab. Józef Banaś
Katedra Analizy Nieliniowej
Politechnika Rzeszowska

Recenzja

rozprawy doktorskiej p. mgr Agnieszki Szumera pt.
Probabilistyczne Struktury Regresji

Przedłożona przez p. mgr Agnieszkę Szumerę rozprawa doktorska jest opracowaniem dość obszernym, liczącym 104 strony. Autorka załącza w rozprawie bibliografię złożoną z 34 pozycji, które pochodzą z różnego okresu, w tym są pozycje pochodzące jeszcze z XIX wieku (tych pozycji jest 8).

W recenzowanej rozprawie doktorskiej Autorka zajmuje się uogólnieniem klasycznej koncepcji regresji rozważanej na gruncie przestrzeni pseudo-Hilberta w pracy [25] (według numeracji Bibliografii w rozprawie doktorskiej) na przestrzeń probabilistyczną. Uogólnienie zawarte w rozprawie doktorskiej przedstawione jest przy pomocy rozwiązania pewnego problemu ekstremalnego na wspomnianej przestrzeni probabilistycznej. Kluczowym pojęciem wprowadzonym w rozprawie doktorskiej jest pojęcie **probabilistycznej struktury regresji**.
W celu wprowadzenia tego pojęcia załóżmy, że $(\Omega, \mathcal{A}, P)$ jest przestrzenią probabilistyczną oraz A jest zadanym zbiorem niepustym, natomiast B oznacza zbiór liczb rzeczywistych $\mathbb{R}$ lub zbiór liczb zespolonych $\mathbb{C}$. Następnie, rozważmy dwie funkcje $x : \Omega \to A$ oraz $y : \Omega \to B$, które interpretujemy jako tzw. funkcje próbkujące, a więc funkcje wygenerowane przez dane eksperymentalne rozważanego modelu regresji.
W dalszym ciągu symbolem δ oznaczać będziemy odchylenie funkcji teoretycznej od danych empirycznych. Zatem

$$\delta : B^\Omega \times B^\Omega \to \overline{\mathbb{R}},$$

gdzie symbol B^Ω oznacza zbiór wszystkich funkcji odwzorowujących przestrzeń zdarzeń elementarnych Ω w zbiór B, natomiast $\overline{\mathbb{R}}$ oznacza rozszerzoną prostą rzeczywistą.

Wspomniane wyżej pojęcie probabilistycznej struktury regresji oznaczamy przez $\mathcal{B}$ i przedstawia ono zespół $(A, B, \delta; x, y)$, który jest rozważany nad przestrzenią probabilistyczną $(\Omega, \mathcal{A}, P)$.

Z probabilistyczną strukturą regresji $\mathcal{B}$ związuje się rodzinę funkcji $\mathcal{F}$ nazywaną **teoretycznym modelem funkcyjnym struktury regresji $\mathcal{B}$**. Rodzina $\mathcal{F}$ jest pewnym podzbiorem rodziny wszystkich funkcji odwzorowujących zbiór A w zbiór B. W pracy zakłada się, że rodzina $\mathcal{F}$ tworzy tzw. rozmaitość liniową we wspomnianej przestrzeni funkcji odwzorowujących zbiór A w zbiór B, która jest wyposażona w standardowe działania dodawania funkcji i ich mnożenia przez skalary za zbioru B.

W dalszych rozważaniach prowadzonych w recenzowanej rozprawie doktorskiej pojawia się naturalne zagadnienie polegające na wyznaczeniu optymalnych funkcji w modelu funkcyjnym $\mathcal{F}$ w tym sensie, że są one najlepiej dopasowane do danych empirycznych reprezentowanych przez funkcje x i y. Kryterium, według którego oceniamy stopień tego dopasowania, jest oczywiście funkcja δ. Ujmując rzecz nieco ściślej, rozważmy problem polegający na wyznaczeniu i zbadaniu klasy $\mathrm{Reg}(\mathcal{F}, \mathcal{B})$ złożonej ze wszystkich funkcji $f_0 \in \mathcal{F}$, które minimalizują funkcjonał $f \longrightarrow \delta(f \circ x, y)$ określony na klasie funkcji $\mathcal{F}$ o wartościach w rozszerzonej prostej rzeczywistej $\overline{\mathbb{R}}$. Funkcje takie noszą nazwę **funkcji regresji w klasie $\mathcal{F}$ względem probabilistycznej struktury regresji $\mathcal{B}$**, natomiast wyżej zdefiniowany problem ekstremalny nazwany został przez Autorkę **problemem regresji w klasie $\mathcal{F}$ dla struktury $\mathcal{B}$**.

Przedłożona rozprawa doktorska jest wzorowana na pracy [25], której autorami są prof. D. Partyka i prof. J. Zając. W nawiązaniu do rozważań tej pracy Doktorantka rozróżnia dwa podstawowe typy struktur regresji opartych na metodzie najmniejszych kwadratów. Są to tzw. asynchroniczne struktury regresji oraz synchroniczne struktury regresji. Nie będę tutaj przytaczał określeń tychże typów struktur regresji, ponieważ w swojej pracy doktorskiej p. A. Szumera skupia się na synchronicznych strukturach regresji określonych przy pomocy przestrzeni probabilistycznych. Tym niemniej, Doktorantka uzyskała wiele wyników mających ogólniejszy i alternatywny charakter w porównaniu do pracy [25]. Ponadto, rozważania prowadzone w rozprawie zawierają uproszczenia rozważań prowadzonych we wspomnianej pracy [25].

Przejdę teraz do krótkiego omówienia treści kolejnych pięciu rozdziałów recenzowanej rozprawy.
Rozdział 1 stanowi wprowadzenie do teorii regresji i przedstawia tę teorię w nawiązaniu do jej historycznego rozwoju. W rozdziale tym wprowadzone są opisowe definicje dotyczące pojęć używanych w teorii regresji i wyjaśniona jest rola pojęć, które te definicje wprowadzają. Szczególną uwagę Autorka zwraca na historyczny rozwój pojęć teorii regresji, podkreślając doniosłą rolę jaką w rozwoju tej teorii odgrał sir F. Galton.
Warto jeszcze wspomnieć o tym, że w Rozdziale 1 znajduje się opis asynchronicznych i synchronicznych struktur regresji przedstawiony zgodnie z prekursorską pracą [25].

Rozdział 2 ma charakter pomocniczy i zawiera fakty dotyczące wykorzystywanego w pracy całkowania funkcji zespolonych. Doktorantka napisała ten rozdział posługując się klasycznym podejściem opartym na aproksymacji funkcji zespolonych przy pomocy funkcji prostych. Dokładniej rzecz ujmując, dowolna funkcja o wartościach zespolonych przedstawiana jest jako suma części rzeczywistej tej funkcji i jej części urojonej, a następnie funkcje wyrażające część rzeczywistą i urojoną aproksymuje się odpowiednio ciągami funkcji prostych.

Podstawowym rozdziałem recenzowanej rozprawy doktorskiej jest Rozdział 3, w którym zdefiniowane jest pojęcie probabilistycznej struktury regresji nad zadaną przestrzenią probabilistyczną. Główny wynik tego rozdziału zawarty jest w Twierdzeniu 3.8, w którym podaje się przedstawienie dowolnej funkcji z teoretycznego modelu funkcyjnego struktury regresji przy pomocy wartości odpowiedniego funkcjonału z przestrzeni pseudo-Hilberta typu L^2. Doniosłość tego twierdzenia polega na tym, że przy jego pomocy charakteryzuje się klasę funkcji regresji. Warto tutaj zwrócić uwagę na fakt, że Twierdzenie 3.8 ma postać znanych z analizy funkcjonalnej twierdzeń o reprezentacji funkcjonałów liniowych w pewnych przestrzeniach Banacha.
Z Twierdzenia 3.8 wypływa szereg wniosków, które są w pracy zauważone i sformułowane. I tak, np. jednym z takich wniosków jest twierdzenie o jednoznaczności funkcji regresji (Twierdzenie 3.15) oraz twierdzenie o możliwości opisu regresji pierwszego rodzaju w terminach probabilistycznych struktur regresji (Twierdzenie 3.11). Ponadto, w Rozdziale 3 zamieszczony jest przykład uzasadniający potrzebę rozważania zespolonych probabilistycznych struktur regresji (Przykład 3.4). Przykład ten jest kontynuowany w Rozdziale 4.

Rozdział 4 omawia zagadnienie efektywnego wyznaczania funkcji regresji przy pomocy zadanej skończonej bazy rozmaitości liniowej $\mathcal{F}$ określającej model funkcyjny struktury regresji. Ponadto, w tym rozdziale znajduje się rozwiązanie problemu regresji dla jednowymiarowych, dwuwymiarowych i trójwymiarowych modeli teoretycznych $\mathcal{F}$. Znajdują się tutaj także rozważania dotyczące regresji dla wielowymiarowych modeli liniowych.

Ostatni rozdział pracy, Rozdział 5, przedstawia kilka konkretnych przykładów wyznaczania funkcji regresji przy pomocy opracowanej przez Doktorantkę metody numerycznej, która bazuje na przedstawionych w rozprawie doktorskiej rozważaniach i ich wynikach. Występujące we wspomnianych przykładach funkcje regresji są wyznaczane dla zadanych układów funkcji bazowych. Przykłady zawarte w Rozdziale 5 są bardzo dobrze dobrane i ładnie zilustrowane. Pokazują one praktyczną użyteczność wyników osiągniętych przez Doktorantkę w Jej dysertacji.

Rozprawa doktorska p. mgr A. Szumery jest starannie zredagowana i przedstawia wyniki z zakresu szeroko rozumianej probabilistyki. Wyniki te są trudne a nawet ośmielę się je nazwać wysoce nietrywialne. Na uwagę zasługuje fakt, że Doktorantka bardzo biegle posługuje się narzędziami i metodami statystyki matematycznej i rachunku prawdopodobieństwa jak również różnymi faktami z teorii miary i całki oraz algebry liniowej.

Do mankamentów pracy należy zaliczyć wprowadzany niekiedy chaos w nadmiarze różnych faktów, którymi Doktorantka się posługuje. Wydaje mi się, że przedłożoną rozprawę doktorską może czytać i rozumieć ktoś, kto jest bardzo dobrze obeznany z metodami stosowanymi w probabilistyce. Innym mankamentem pracy doktorskiej jest również to, że Doktorantka nie wskazała w rozprawie dokładnie tych rezultatów, których sama jest autorką i które zostały opublikowane w pracy [33] - jest to jedyna praca napisana przez Doktorantkę (poza pracą [27], która ma charakter dydaktyczny), cytowana w Bibliografii.

Wspomniane wyżej uwagi krytyczne nie umniejszają wartości naukowej rezultatów zawartych w recenzowanej rozprawie doktorskiej. **Z pełnym przekonaniem wnioskuję o dopuszczenie p. mgr Agnieszki Szumery do dalszych etapów przewodu doktorskiego oraz o nadanie Jej stopnia doktora nauk matematycznych.**

Rzeszów, 30 sierpnia 2022 r.