

VIII Interdyscyplinarne Warsztaty
Matematyczno-Informatyczne

8-9.05.2021

Spis treści

Abstrakty	1
O statystyce, nieprecyzyjnych danych, komputerach oraz pożytkach płynących z mieszania (<i>prof. dr hab. Przemysław Grzegorzewski</i>)	1
Czy prawdopodobieństwo jest szybsze niż światło? Opowieść o miarach, przyczynowości i zakrzywionej czasoprzestrze (<i>dr Tomasz Miller</i>)	2
O pewnych zastosowaniach algebry uniwersalnej (<i>dr hab. Anna Zamojska-Dzienio</i>)	3
O przenoszeniu własności anihilatorowej na pierścień wielomianów (<i>Grzegorz Bajor</i>)	4
Kolorowania online grafów geometrycznych (<i>Joanna Chybowska-Sokół</i>)	4
Algebry Steinberga (<i>Anna Cichocka</i>)	5
Badanie przestrzennych struktur w mózgu myszy z wykorzystaniem metod statystyki bayesowskiej (<i>Agnieszka Geras</i>)	5
Odpowiedzialne uczenie głębokie na obrazach medycznych płuc w chorobie COVID-19 (<i>Weronika Hryniewska</i>)	6
Metody selekcji zmiennych uwzględniające informacje o ich kosztach - podejście oparte na teorii informacji (<i>Tomasz Klonecki</i>)	7
Matematyczna analiza modeli turbulencji (<i>Przemysław Kosewski</i>)	8
Improving Neural Machine Translation (<i>Mateusz Krubiński</i>)	8
Testowanie hipotez o warunkowej niezależności z wykorzystaniem teorii informacji (<i>Małgorzata Łazęcka, Jan Mielniczuk</i>)	9
Grupoid grafowy (<i>Paweł Matraś</i>)	9
Problem homomorfizmu w szczególnych klasach grafów (<i>Karolina Okrasa</i>)	10

Estymacja przyczynowa pod warunkiem częściowej złej specyfikacji modelu (<i>Krzysztof Rudaś</i>)	11
Co słyhać na Reddicie? Czyli analiza 180 prac (2019-2020) wykorzystujących platformę Reddit (<i>Jan Sawicki</i>)	11
Alokacja optymalna w warstwowych schematach próbkowania (<i>Wojciech Wójciak</i>)	12
Wielomiany ortogonalne (<i>Agnieszka Zięba</i>)	13
Tkaniny w geometrii symplektycznej (<i>Marcin Zubilewicz</i>)	13
O wagach, które dopuszczają jądra reprodukujące typu Szegö (<i>Tomasz Łukasz Żynda</i>)	14

Abstrakty

O statystyce, nieprecyzyjnych danych, komputerach oraz pożytkach płynących z mieszania

prof. dr hab. Przemysław Grzegorzewski

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Instytut Badań
Systemowych Polskiej Akademii Nauk

Statystyka jest sztuką podejmowania decyzji w obliczu niepewności. Dostarcza narzędzi do opisywania i wyjaśniania rzeczywistości, do prognozowania i weryfikowania hipotez. Przez długi czas jedyną powszechnie uznaną teorią opisującą niepewność był rachunek prawdopodobieństwa. W ostatnich kilkudziesięciu latach zaproponowano kilka nowych podejść do opisu i analizy niepewności, rozszerzających klasyczną teorię prawdopodobieństwa lub ortogonalnych do niej. Ich wspólną cechą jest uelastycznienie metod klasycznych, tak aby można je było łatwiej dostosować do rzeczywistego charakteru dostępnych informacji i lepiej sobie radzić z modelowaniem innych niż losowość rodzajów niepewności, takich – na przykład – jak brak precyzji.

W ten sposób statystyka została skonfrontowana z teorią zbiorów rozmytych. Po okresie początkowych antagonizmów badacze zdali sobie sprawę, że statystyka i teoria zbiorów rozmytych nie powinny być uważane za konkurencyjne, gdyż opisują inne aspekty niepewności, natomiast mogą się z powodzeniem uzupełniać. Zauważono, że poszerzenie statystyki o analizę danych rozmytych nie tylko pozwala rozwiązać niektóre problemy, ale też generuje nowe ciekawe pytania. W szczególności, rozróżnienie między tzw. zbiorami ontycznymi i epistemicznymi (por. [2]) prowadzi do odmiennych definicji wariancji, a co za tym idzie, wymaga innych narzędzi wspomagających wnioskowanie. Analitykom danych i statystykom z wydatną pomocą przyszła informatyka, która otworzyła drogę rozwojowi technik permutacyjnych oraz metod repróbkiowania, w tym bootstrapu, bez których efektywne wnioskowanie na podstawie danych rozmytych nie byłoby możliwe (por. [1, 3, 4]).

W prezentacji postaramy się naszkicować wzajemne relacje między statystyką i teorią zbiorów rozmytych, wskazać niektóre problemy i wyzwania, które pojawiły się na przecięciu obu teorii oraz zasygnalizować pewne problemy otwarte.

Literatura

[1] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. Garcia-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B.

Ramos-Guajardo, S. de la Rosa de Súa, B. Sinova (2014), *A distance-based statistic analysis of fuzzy number-valued data*, International Journal of Approximate Reasoning, 55, 1487–1501.

[2] I. Couso, D. Dubois (2014), *Statistical reasoning with set-valued information: Ontic vs. epistemic views*, International Journal of Approximate Reasoning, 55, 1502–1518

[3] P. Grzegorzewski (2020), *Two-sample dispersion problem for fuzzy data*, M.J. Lesot i poz. (red.), Information Processing and Management of Uncertainty in Knowledge-Based System (IPMU 2020), Springer, pp. 82–96.

Czy prawdopodobieństwo jest szybsze niż światło? Opowieść o miarach, przyczynowości i zakrzywionej czasoprzestrze

dr Tomasz Miller

Centrum Kopernika Badań Interdyscyplinarnych, Uniwersytet Jagielloński

W 1908 r. Hermann Minkowski jako pierwszy zdał sobie sprawę, że teoria opublikowana trzy lata wcześniej przez jego byłego studenta Alberta Einsteina (nazwana później szczególną teorią względności), ukazała radykalnie nowy obraz świata. Oto czas i przestrzeń – dotąd rozłączne pojęcia – okazały się względnymi, zależnymi od obserwatora aspektami czterowymiarowej „czasoprzestrzeni”. Einsteiński postulat, że nic – ani cząstki, ani oddziaływania, ani jakiegokolwiek związku przyczynowo-skutkowe – nie może rozchodzić się szybciej niż światło w próżni, w naturalny sposób wpisuje się w ten geometryczny obraz. Wyraża się go za pomocą tzw. krzywych przyczynowych oraz relacji przyczynowego poprzedzania, która wiąże te punkty czasoprzestrzeni, między którymi możliwa jest podświetlna komunikacja.

Jak jednak uczy fizyka, punktowe cząstki i zdarzenia są daleko idącą idealizacją i zasadne jest pytać, jak rygorystycznie opisywać przyczynowość dla zjawisk nielokalnych. W referacie, po popularnonaukowym wprowadzeniu w geometrię czasoprzestrzeni (także zakrzywionych!), opowiem o prowadzonych wspólnie z M. Ecksteinem badaniach nad rozszerzeniem teorii przyczynowości na miary probabilistyczne [1]. Skupię się zwłaszcza na zastosowaniu rozwiniętego formalizmu do opisu przyczynowej ewolucji miar probabilistycznych, które prowadzi do eleganckiego, bo niezależnego od obserwatora, probabilistycznego uogólnienia krzywej przyczynowej [2], a także do głębokiego związku między przyczynowością a tzw. równaniem ciągłości – jednym z najważniejszych równań w fizyce [3].

Literatura

[1] M. Eckstein, T. Miller (2017), *Causality for nonlocal phenomena*, Ann. Henri Poincaré, 18(9), 3049–3096.

[2] T. Miller (2017), *Polish spaces of causal curves*, J. Geom. Phys., 116, 295–315.

[3] T. Miller (2021), *Causal evolution of probability measures and continuity equation*, arxiv :2104.02552.

O pewnych zastosowaniach algebry uniwersalnej

dr hab. Anna Zamojska-Dzienio

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Algebrą (lub: strukturą algebraiczną, algebrą abstrakcyjną) nazywamy zbiór wraz z rodziną operacji na nim zdefiniowanych. *Algebra uniwersalna* jest działem matematyki, który zajmuje się badaniem ogólnych własności struktur algebraicznych. Szczególną rolę odgrywają *rozmaitości algebr*, czyli klasy wszystkich algebr (danego typu), które spełniają zadany zbiór równości. Wiele struktur ma swoje źródło w badaniach z innych dziedzin nauki - algebraicy nie muszą znać pochodzenia i motywacji stojącej za rozważanymi obiektami (choć bardzo wzbogaca to badania!) i mogą się skupić na czystej analizie ich własności i poszukiwaniem analogii między nimi.

Podczas referatu przedstawię różnorodne przykłady algebr, które m.in.

- pojawiają się w topologii niskich wymiarów (główną motywacją do ich badania jest znalezienie łatwo obliczalnych niezmienników węzłów),
- opisują rozwiązania równania Yanga-Baxtera (wywodzącego się z fizyki cząstek elementarnych) lub
- stanowią algebraiczną semantykę dla różnych logik.

W szczególności, pokażę, jak narzędzia algebry uniwersalnej okazały się przydatne w opisanu struktury quandle medialnych (rozmaitości generowanej przez quandle Alexandra) i stworzeniu algorytmów do ich zliczania [2], istotnie rozszerzających znane rezultaty.

Podręczniki [1] i [3] polecam jako dobry wstęp do Algebry Uniwersalnej, gdyby kogoś zainteresował temat.

Literatura

[1] C. Bergman (2011), *Universal Algebra - Fundamentals and Selected Topics*, Chapman and Hall/CRC

[2] P. Jedlička, A. Pilitowska, D. Stanovský, A. Zamojska-Dzienio (2015), *The structure of medial quandles*, J. Algebra 443, 300–334

[3] A. Romanowska (2020), *Algebra i jej zastosowania*, Oficyna Wydawnicza Politechniki Warszawskiej

O przenoszeniu własności anihilatorowej na pierścień wielomianów

Grzegorz Bajor

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Znajdowanie anihilatorów algebr, jest częstą praktyką do szukania zależności w algebrze, dlatego warto także zapytać się o pewne własności anihilatorów.

Często zadaje się pytanie, czy jeśli pewna algebra A spełnia pewną własność, to czy dla pierścienia wielomianów $A[X]$ nad wyjściową algebrą, wskazana własność jest nadal prawdziwa.

Podczas prezentacji wprowadzę definicję związane z anihilatorami algebry, a także przedstawię konstrukcję pewnego pierścienia R , dla którego własność anihilatorowa jest spełniona, ale już w $R[x]$ jej nie spełnia.

Referat oparty jest na wynikach otrzymanych razem z Michałem Ziembowskim.

Literatura

- [1] G. Bajor, M. Ziembowski (2019), *Annihilator condition does not pass to polynomials and power series*, Journal of Pure and Applied Algebra. 223, 3869–3878
-

Kolorowania online grafów geometrycznych

Joanna Chybowska-Sokół

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Kolorując graf online nie znamy całej jego struktury a priori. Wierzchołki są nam prezentowane kolejno wraz z krawędziami do wcześniej przedstawionego podgrafu i kolorujemy je natychmiast gdy są zaprezentowane, bez możliwości późniejszej zmiany koloru. Naturalnie będziemy używać więcej kolorów niż w optymalnym kolorowaniu. W niektórych modelach kolorowania wręcz ukończenie pokolorowania będzie niemożliwe. Z pomocą może nam przyjść wiedza o geometrycznej strukturze kolorowanego grafu. Pokażemy przykłady algorytmów kolorujących online wykorzystujących pewne geometryczne własności grafów.

Joanna Chybowska-Sokół: Partially supported by the National Science Center of Poland under grant no. 2016/23/N/ST1/03181

Literatura

- [1] J. Chybowska-Sokół, G. Gutowski, K. Junosza-Szaniawski, P. Mikos, A. Polak (2020), *Online Coloring of Short Intervals*, Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2020), Seria Leibniz International Proceedings in Informatics (LIPIcs) 176, 52:1—52:18
- [2] J. Fiala, A. Fishkin, F. Fomin (2004), *On distance constrained labeling of disks*, Theoret. Comput. Sci. 326, 216–292

- [3] K. Junosza-Szaniawski, P. Rzażewski, J. Sokół, K. Węsek (2018), *Online coloring and $l(2,1)$ -labeling of unit disk intersection graphs*, SIAM Journal on Discrete Mathematic 32.2, 1335–1350
-

Algebry Steinberga

Anna Cichocka

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Algebry Steinberga są przykładem algebr łącznych definiowanych w oparciu o obszerne grupoidy topologiczne Hausdorffa (ang. ample Hausdorff groupoid). Dla dowolnego rozważanego przemiennego pierścienia z jedyneką \mathcal{R} oraz obszernego grupoidu \mathbf{G} elementami algebry Steinberga $\mathcal{A}_{\mathcal{R}}(\mathbf{G})$ nazywamy funkcje działające z \mathbf{G} w \mathcal{R} spełniające pewne określone własności. Klasa algebr Steinberga jest uogólnieniem takich klas algebr jak na przykład grafowe algebry ścieżek Leavitta. Podczas referatu przybliżona zostanie konstrukcja algebry Steinberga oraz podstawowe fakty z nią związane.

Literatura

- [1] B. Steinberg(2010), *A groupoid approach to discrete inverse semigroup algebras*, Advances in Mathematics 223, 2, 689–727.
- [2] J. Renault(1980), *A Groupoid Approach to C^* -Algebras*, vol. 793 of Lecture Notes in Mathematics. Springer-Verlag
- [3] L. O. Clark, C. Farthing, A. Sims, M. Tomforde(2014), *A groupoid generalisation of Leavitt path algebras*, Semigroup Forum 89, 3, 501–517.
- [4] S. W. Rigby(2018), *The groupoid approach to Leavitt path algebras*, arXiv e-prints
-

Badanie przestrzennych struktur w mózgu myszy z wykorzystaniem metod statystyki bayesowskiej

Agnieszka Geras

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Sekwencjonowanie RNA i otrzymane dzięki tej technice dane na temat ekspresji genów umożliwiają badanie kompozycji typów komórek w różnych tkankach [1]. W szczególności nowa metoda przestrzennego sekwencjonowania (ang. *spatial transcriptomics*), pozwalająca na zachowanie przestrzennej informacji na temat ekspresji genów, otwiera więcej możliwości w tej dziedzinie [2].

W swoim wystąpieniu zaprezentuję nowy, probabilistyczny model poziomu ekspresji genów w wybranych punktach tkanki oraz metodę estymacji jego parametrów. Przedstawię także wyniki przestrzennej analizy typów komórek w mózgu myszy,

otrzymane przy pomocy nowo zaproponowanego modelu. Okazuje się, że możliwe jest odtworzenie struktur znanych z literatury oraz różnicowanie przestrzenne pomiędzy dwoma głównymi typami neuronów: hamującymi i pobudzającymi. Praca wspólna z Shadi Darvish Shafighi, Jensem Lagergrenem oraz Ewą Szczurek.

Literatura

- [1] Allen W. Zhang et al. (2019), *Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling*, Nature Methods 16
 - [2] Emelie Berglund et al. (2018), *Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity*, Nature Communications 9
-

Odpowiedzialne uczenie głębokie na obrazach medycznych płuc w chorobie COVID-19

Weronika Hryniewska

Wydział Matematyki i Nauk Informacyjnych Politechniki Warszawskiej

Obecnie jednym z najważniejszych problemów globalnych jest szybkie i niekontrolowane rozprzestrzenianie się wirusa COVID-19. W krótkim czasie doprowadziło ono do powstania wielu modeli uczenia głębokiego z modułami wyjaśniającymi do detekcji COVID-19.

W niniejszej prezentacji pokażę systematyczną analizę różnych aspektów proponowanych modeli. Przedstawię liczne błędy popełnione na różnych etapach pozyskiwania danych, tworzenia modelu i konstruowania wyjaśnień. Wskażę typowe błędy wynikające z braku dogłębnego zrozumienia dziedziny radiografii oraz niewystarczającej współpracy inżynierów uczenia głębokiego z radiologami.

Literatura

- [1] W. Hryniewska, P. Bombiński, P. Szatkowski, P. Tomaszewska, A. Przelaskowski, P. Biecek (2020), *Checklist for responsible deep learning modeling of medical images based on COVID-19 detection studies*, arXiv 2012.08333
-

Metody selekcji zmiennych uwzględniające informacje o ich kosztach - podejście oparte na teorii informacji

Tomasz Klonecki

Instytut Podstaw Informatyki Polskiej Akademii Nauk

Selekcja zmiennych w problemie klasyfikacji nadzorowanej jest zadaniem kluczowym w wielu zadaniach bioinformatycznych. Większość istniejących podejść zakłada, że wszystkie zmienne mają taki sam koszt. Jednak w wielu zastosowaniach medycznych to założenie może nie być do końca właściwe, ponieważ pozyskanie jakiejś cechy może być kosztowne. W diagnostyce medycznej koszty zmiennych wiążą się zazwyczaj z kosztami wykonania badań, jednak mogą również odnosić się do aspektów poza finansowych, na przykład decyzji pomiędzy wykonaniem skomplikowanej operacji, a pobraniem krwi. W takich przypadkach celem zadania selekcji zmiennych jest uzyskanie jak największej skuteczności przewidywania zmiennej celu (np. wystąpienie choroby) w ramach założonego budżetu, określonego przez użytkownika.

Głównym celem naszych badań jest przegląd filtrowych metod selekcji zmiennych opartych o teorię informacji i zaproponowanie ich nowych wariantów uwzględniających koszty. Próbuje rozwiązać NP-trudny problem selekcji przy użyciu zachłanych algorytmów.

Nasze rozumowanie teoretyczne wspieramy eksperymentami na sztucznych danych jak i na medycznej bazie danych MIMIC-II. Eksperymenty potwierdzają, że metody przedstawione w artykule mogą usprawnić działanie algorytmów wykrywania chorób z ograniczonym budżetem.

Literatura

- [1] Paweł Teisseyre, Tomasz Klonecki (2021), *Controlling costs in feature selection: information theoretic approach*, International Conference on Computational Science (Przyjęty na konferencję)
 - [2] G. Brown, A. Pock, M. J. Zhao, M. Lujan (2018), *Conditional likelihood maximisation: A unifying framework for information theoretic feature selection*, Journal of Machine Learning Research
 - [3] V. Bolon-Canedo, I. Porto-Diaz, N. Sanchez-Marono, A. Alonso-Beranzos (2014), *A framework for cost-based feature selection*, Pattern Recognition
-

Matematyczna analiza modeli turbulencji

Przemysław Kosewski

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Najpopularniejszym modelem opisującym przepływy jest równanie Naviera-Stokesa. Numeryczne rozwiązywanie równania N-S dla „szybkich” przepływów okazują się trudnym przedsięwzięciem ze względu na konieczność odzwierciedlenia turbulencji. Powoduje to konieczność użycia bardzo drobnej siatki przez co czas obliczeń jest znacznie wydłużony. Czyni to model mało użytecznym w praktyce inżynierskiej oraz wymusza konieczność stosowania innych modeli przepływu. Jednym z takich modeli jest model turbulencji Kołmogorowa. W referacie opowiem o sposobie wyprowadzenia modelu turbulencji Kołmogorowa. Dodatkowo opowiem o aktualnym stanie wiedzy dotyczącym tego równania.

Improving Neural Machine Translation

Mateusz Krubiński

Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics

Machine Translation is a task of automatically translating text from one language into another. In recent years, solutions based on neural networks have dominated the field.

In this talk I will give a short introduction to the topic and mention a few tricks that can be used to improve the final quality of translation. I will also share my experience from participating in one of the longest-running international Machine Translation competitions.

Literatura

[1] M. Krubiński, M. Chochowski et al. (2020), *Samsung R&D Institute Poland submission to WMT20 News Translation Task*, Proceedings of the Fifth Conference on Machine Translation

Testowanie hipotez o warunkowej niezależności z wykorzystaniem teorii informacji

Małgorzata Łazęcka, Jan Mielniczuk

Wydział Matematyki i Nauk Informacyjnych PW Instytut Podstaw Informatyki PAN

Przedstawione zostanie zagadnienie dotyczące łącznego testowania warunkowej niezależności dwóch zmiennych X i Y pod warunkiem Z_i , dla p zmiennych Z_i ($i \in \{1, 2, \dots, p\}$) i bezpośrednie podejście wykorzystujące jako statystykę testową dla powyższego problemu kryterium selekcji zmiennych JMI (ang. *joint mutual information*). Kryterium JMI oparte jest na warunkowej informacji wzajemnej $I(X, Y|Z)$ - nieujemnej mierze pochodzącej z teorii informacji wyznaczającej siłę warunkowej zależności pomiędzy zmiennymi X i Y przy znanej zmiennej Z . W przypadku pojedynczego testu warunkowej niezależności jako statystyki testowej można użyć warunkowej informacji wzajemnej; kryterium JMI uśrednia wartości warunkowych informacji wzajemnych $I(X, Y|Z_i)$, zatem jest statystyką mierzącą skumulowany efekt naruszania hipotez zerowych w pojedynczych testach.

W referacie pokazany zostanie asymptotyczny rozkład statystyki \widehat{JMI} przy hipotezie zerowej oraz zostanie przytoczony wynik dotyczący asymptotycznego zachowania \widehat{JMI} dla binarnej zmiennej Y przy hipotezie alternatywnej. Test oparty na \widehat{JMI} zostanie porównany z ogólnymi metodami testowania w przypadku wielu hipotez (procedura Bonferroniego i Simesa).

Literatura

- [1] M. Kubkowski, M. Łazęcka, J. Mielniczuk (2020), *Distributions of a General Reduced-Order Dependence Measure and Conditional Independence Testing*, Lecture Notes in Computer Science 12143, 692–706

Grupoid grafowy

Paweł Matraś

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Referat zaczniemy od określenia topologii na ścieżkach grafu. Pozwoli nam to zdefiniować na algebraicznej strukturze grupoidu topologię. Mając tak określony grupoid możemy rozpatrywać algebrę Steinberga, której bazą są funkcje charakterystyczne pewnych podzbiorów. Pokażemy izomorfizm tej algebry z algebrą Leavitta, której definicja jest bezpośrednio związana z ścieżkami grafu.

Takie podejście pozwala uzyskać dalsze zależności pewnych własności algebraicznych z własnościami topologicznymi. Daje także nowe metody badania algebr związanych z grafami, które w ostatnich latach cieszą się dużym zainteresowaniem w różnych dziedzinach matematyki.

Przedstawione wyniki pochodzą z pracy „The groupoid approach to Leavitt path algebras”, której autorem jest Simon W. Rigby.

Literatura

- [1] Simon W. Rigby (2018), *The groupoid approach to Leavitt path algebras*, arXiv, 17–31
-

Problem homomorfizmu w szczególnych klasach grafów

Karolina Okrasa

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Dla grafu G i liczby naturalnej k , funkcja $c : V(G) \rightarrow \{1, \dots, k\}$ jest *poprawnym k -kolorowaniem* G , jeśli dla każdej krawędzi uv grafu G zachodzi $c(u) \neq c(v)$. Dla dwóch grafów G i H , funkcja $f : V(G) \rightarrow V(H)$ jest *homomorfizmem* G w H , jeśli dla każdej krawędzi uv grafu G , $f(u)f(v)$ jest krawędzią grafu H . W szczególności, jeśli H jest grafem pełnym o k wierzchołkach, każdy homomorfizm z G z H odpowiada k -kolorowaniu grafu G , homomorfizmy grafów stanowią więc uogólnienie problemu kolorowania.

Dla ustalonego grafu H , problem znajdowania homomorfizmu z dowolnego grafu G w H jest wielomianowy, jeśli H jest dwudzielny, i NP-zupełny w przeciwnym przypadku [1]. Co więcej, gdy problem jest NP-zupełny, nie potrafimy znajdować homomorfizmów z G w H istotnie szybciej niż metodą *brute-force*, jeśli nie założymy nic o grafie G .

Podczas referatu opowiem o tym, jak pewne dodatkowe założenia na klasę grafów, z których pochodzi G , wpływają na obliczeniową złożoność problemu homomorfizmu, pozwalając rozwiązać go w czasie podwykładniczym lub nawet wielomianowym [2,3]. Wyniki przedstawione w referacie zostały otrzymane wspólnie z Pawłem Rzażewskim.

Literatura

- [1] P. Hell, J. Nešetřil, (1990), *On the complexity of H -coloring*, J. Comb. Theory, Ser. B 48, 92–110, [https://doi.org/10.1016/0095-8956\(90\)90132-J](https://doi.org/10.1016/0095-8956(90)90132-J)
- [2] K. Okrasa, P. Rzażewski, (2021), *Fine-grained complexity of the graph homomorphism problem for bounded-treewidth graphs*, SIAM Journal on Computing 50(2), 487–508, <https://doi.org/10.1137/20M1320146>
- [3] K. Okrasa, P. Rzażewski, (2021), *Complexity of the List Homomorphism Problem in Hereditary Graph Classes*, STACS 2021 Proc., 54:1–54.17, <https://doi.org/10.4230/LIPIcs.STACS.2021.54.1-54.17>
-

Estymacja przyczynowa pod warunkiem częściowej złej specyfikacji modelu

Krzysztof Rudaś

Wydział Matematyki i Nauk Informatycznych PW, Instytut Podstaw Informatyki PAN

Modelowanie przyczynowości zajmuje się przewidywaniem efektu podejmowanej przez nas akcji (nowej kampanii marketingowej, metody leczenia pacjentów) dla pojedynczej obserwacji. W tym celu dzielimy populację na grupę eksperymentalną (poddaną działaniu) i kontrolną (nie poddaną działaniu). Następnie konstruujemy estymator który liczy różnicę efektów między grupami dla pojedynczej obserwacji. W [1] pokazaliśmy, przy założeniu liniowości w grupie eksperymentalnej i kontrolnej, asymptotyczne własności dwóch popularnych estymatorów, a także nowego podejścia łączącego zalety dwóch poprzednich.

W moim referacie przedstawię asymptotyczne wyniki uzyskane dla wyżej wymienionych estymatorów przy założeniu, że odpowiedź w grupie kontrolnej i część odpowiedzi w grupie eksperymentalnej, niezwiązana z efektem akcji, są nieliniowe.

Literatura

[1] K. Rudaś, S. Jaroszewicz (2018), *Linear regression for uplift modeling*, Data Mining and Knowledge Discovery 32, 1275–1305

Co słyhać na Reddicie? Czyli analiza 180 prac (2019-2020) wykorzystujących platformę Reddit

Jan Sawicki

Wydział Matematyki i Nauk Informatycznych, Politechnika Warszawska

Reddit to tematyczne forum internetowe z aspektami medium społecznościowego. Warsztat przedstawi ogólną i szczególną analizę blisko 180 publikacji z lat 2020 i 2021, które wykorzystywały tę platformę jako główne lub poboczne źródło danych. Przedstawione zostaną główne nurty tematyczne pojawiające się w wybranych pracach, takich jak generacja dialogów, analiza mowy nienawiści, czy diagnozowanie depresji na podstawie komunikacji w internecie. Podane zostaną najczęściej wykorzystane metody oraz praktyczne technologie, które przejawiają się najczęściej w literaturze. Dzięki eksperymentowi kończącemu prelekcje, słuchacze będą mieli możliwość przekonać się na własne oczy, że Reddit posiada “subreddit“ na dowolny temat.

Literatura

[1] Medvedev, Alexey N., Renaud Lambiotte, and Jean-Charles Delvenne (2017), *The anatomy of Reddit: An overview of academic research.*, In Dynamics on and of Complex Networks, pp. 183-204. Springer

- [2] Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn (2020), *The pushshift reddit dataset.*, International AAAI Conference on Web and Social Media, vol. 14, pp. 830-839
- [3] Zannettou, Savvas, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn (2019), *Who let the trolls out? towards understanding state-sponsored trolls.*, In Proceedings of the 10th acm conference on web science, pp. 353-362
-

Alokacja optymalna w warstwowych schematach próbkowania

Wojciech Wójciak

Wydział Matematyki i Nauk Informacyjnych Politechnika Warszawska

Klasycznym zagadnieniem w badaniach statystycznych ze schematem losowania warstwowego jest alokacja próby, tj. podział całkowitej zadanej liczebności próby pomiędzy warstwy. Jedną z możliwych metod alokacji jest *alokacja optymalna*, definiowana jako alokacja minimalizująca wariancję estymatora wartości globalnej badanej cechy populacji. Problem *alokacji optymalnej* może być rozszerzony o dodatkowe warunki nakładane na liczebność próby na poziomie warstwy, np. ograniczenie górne równe wielkości warstwy. W dotychczasowej literaturze przedstawiono propozycję rozwiązań tego problemu z pewnymi dodatkowymi warunkami ograniczającymi. Jedne z tych rozwiązań zostały zaproponowane bez formalnego dowodu, inne zaś są niepoprawne, co wykażę w referacie.

Głównym celem i osiągnięciem prezentowanej pracy jest systematyczne podejście do rozwiązania opisanego powyżej problemu *alokacji optymalnej* z dodatkowymi więzami postaci $0 < l_h \leq n_h \leq u_h \leq N_h$. (N_h - liczebność warstwy h , l_h/u_h - ograniczenie dolne/górne na liczebność próby n_h w warstwie h). W pierwszej części referatu, sformułuję przedmiotowy problem w terminach optymalizacji matematycznej. Następnie, z wykorzystaniem warunków Karusha-Kuhna-Tuckera, wyprowadzę docelowe warunki konieczne i dostateczne optymalności. W drugiej części przedstawię algorytmy numeryczne, wywiedzione z przedstawionych warunków optymalności, wraz z przykładową implementacją w R.

Referat jest wynikiem wspólnej pracy z Prof. dr hab. Jackiem Wesołowskim (PW, GUS) i dr Robertem Wieczorkowskim (GUS).

Wielomiany ortogonalne

Agnieszka Zięba

Wydział Matematyki i Nauk Informatycznych Politechniki Warszawskiej

Wielomiany i pojęcie ortogonalności są dobrze znanymi matematycznymi terminami, ale czy ich połączenie może dać jakieś nietrywialne rezultaty?

W mojej prezentacji wprowadzę podstawowe pojęcia i udowodnię własności wspomnianych obiektów. Zaprezentuję także efektywną metodę sprawdzania czy wielomiany są ortogonalne.

Literatura

[1] T. S. Chihara (2016), *An introduction to orthogonal polynomials*, Cambridge University Press

Tkaniny w geometrii symplektycznej

Marcin Zubilewicz

Wydział Matematyki i Nauk Informatycznych, Politechnika Warszawska

Temat mojego wystąpienia złożony jest z dwóch odrębnych gałęzi geometrii. Jedną z nich to teoria tkanin – klasyczna już dziedzina matematyki zapoczątkowana przez W. Blaschkego, G. Thomsena i pozostałych przedstawicieli szkoły hamburskiej w latach 20-tych XX wieku [1,2], badająca geometrię przecięć między kilkoma różnymi rodzinami krzywych lub wyżej-wymiarowych podrozmaitości. (Skojarzenie z rzeczywistymi tkaninami jest tu całkiem na miejscu!) Z drugiej strony mamy geometrię symplektyczną, stanowiącą współczesne podwaliny mechaniki hamiltonowskiej [3], w której iloczyn skalarny, pozwalający mierzyć długości i kąty, zostaje zastąpiony pewną niezdegenerowaną i *antysymetryczną* formą dwuliniową silnie związaną z fizycznym pojęciem działania. Czy z połączenia tych dwóch dziedzin może wyniknąć coś istotnie interesującego?

Aby przybliżyć nas do właściwej odpowiedzi na to pytanie, opowiem krótko i w sposób niezbyt formalny o produktach powyższej fuzji: *tkaninach lagranżowskich* (znanych również jako *struktury bilagranżowskie*) oraz o *tkaninach symplektycznych*, Skupię się głównie na strukturze lokalnej wymienionych obiektów, starając się nakreślić tym samym ich głęboki związek z szeroko pojętą mechaniką (nie tylko klasyczną).

Literatura

[1] W. Blaschke and G. Bol (1938), *Geometrie der Gewebe: Topologische Fragen der Differentialgeometrie*, Die Grundlehren der Mathematischen Wissenschaften 49, Springer

- [2] J.V. Pereira and L. Pirio (2010), *An Invitation to Web Geometry*, IMPA Monographs 2, Springer
- [3] V.I. Arnold (1981), *Metody matematyczne mechaniki klasycznej*, PWN
-

O wagach, które dopuszczają jądra reprodukujące typu Szegö

Tomasz Łukasz Żynda
Wojskowa Akademia Techniczna

Słuchacze dowiedzą się, czym jest przestrzeń Hilberta z jądrem reprodukującym. Poznają definicję przestrzeni i jądra Szegö, a także usłyszą przykładowe warunki wystarczające na wagę całkowania w iloczynie skalarnym, by istniało jądro reprodukujące ważonej przestrzeni Szegö. Dowiedzą się też, że istnieją ważne przestrzenie Szegö bez jądra reprodukującego i jak je konstruować.

Literatura

- [1] T. Ł. Żynda, On weights which admit reproducing kernel of Szegö type, *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)* 55 (2020),
-