prof. dr hab. Jan Mielniczuk
Instytut Podstaw Informatyki PAN
i Wydział MiNI PW

25.I.2016

Recenzja rozprawy doktorskiej mgr Damiana Brzyskiego

Selecting relevant groups of explanatory variables via convex optimization methods with the false discovery rate control

Omówienie rozprawy

W standardowym modelu liniowym $y = X\beta + z$ ($X(n \times p)$- macierz eksperymentu, z-wektor niezależnych zmiennych $N(0, \sigma^2)$, β- wektor nieznanych współczynników) rozpatruje się rozbicie zbioru indeksów zmiennych $\{1, 2, \ldots, p\} = I_1 \cup \ldots \cup I_m$ i odpowiadające przedstawienie wektora wartości oczekiwanych odpowiedzi $Ey = \sum_{i=1}^m X_{I_i}\beta_{I_i}$. Problemem badawczym rozważanym w rozprawie doktorskiej jest detekcja grup indeksów odpowiadajacym niezerowym $\beta_{I_j} : ||\beta_{I_j}|| > 0$, w sytuacji, gdy liczba takich grup jest nieduża w porównaniu z liczbą wszystkich grup m. Punktem wyjścia jest metoda SLOPE, dla której

$$\hat{\beta}^S = \text{argmin}_b\{\frac{1}{2}||y - Xb||^2 + \sigma J_\lambda(b)\},$$

gdzie $J_\lambda(b) = \sum_{i=1}^p \lambda_i b_{(i)}$, $\lambda = (\lambda_1, \ldots, \lambda_p)^T$ jest wektorem uporządkowanych nierosnąco nieujemnych kar, a $|b|_{(1)} \geqslant |b|_{(2)} \geqslant \ldots \geqslant |b|_{(p)}$ są uporządkowanymi nierosnąco wartościami bezwzględnymi współrzędnych wektora b. Metodę SLOPE adaptuje się w pracy do grupowej metody SLOPE, gSLOPE, dla której

$$\hat{\beta}^{gS} = \text{argmin}_b\{\frac{1}{2}||y - Xb||^2 + \sigma J_\lambda(W[\![b]\!]_I)\},$$

gdzie $W = diag(w_1, \ldots, w_m)$ jest diagonalną macierzą wag poszczególnych grup, a $[\![b]\!]_I = (||X_{I_1}\beta_{I_1}||, \ldots, ||X_{I_1}\beta_{I_m}||)^T = (||\beta_{I_1}||, \ldots, ||\beta_{I_m}||)^T$; ostatnia równość zachodzi dla macierzy eksperymentu ortogonalnej wewnątrz każdej z grup (przypadek rozpatrywany w pracy). Nośnik wektora $\hat{\beta}^{gS}$ traktuje się jako selektor zbioru indeksów niezerowych bloków. Dla sytuacji ortogonalnej $X^TX = I$ w rozdziale 5.3 sprowadza się rozwiązanie powyższego problemu do problemu SLOPE z diagonalną macierzą eksperymentu $diag(w_1^{-1}, \ldots, w_m^{-1})$ i zastosowania praktyczne tematyki.

Główny badany w rozprawie obiekt, selektor gSLOPE, jest naturalnym przeniesieniem estymatora SLOPE na przypadek grupowy. Idea SLOPE polega na współmonotonicznym przypisaniu wektorowi statystyk pozycyjnych $|b|_{(i)}$ wektora kar λ_i, co w efekcie, w przypadku występowania wielu dużych wartości $|b|_{(i)}$, powoduje, że małe wartości bezwzględne współczynników będą karane bardziej liberalnie. Jest to idea podobna do idei procedury Benjaminiego-Hochberga, w której występowanie wielu dużych wartości statystyki z wymusza obniżenie progu akceptacji zmiennej jako istotnej dla pozostałych zmiennych. Jednocześnie jest to metodologia *przeciwna* do np. metodologii adaptacyjnej metody Lasso, w której kara jest tym mniejsza, im większa jest wartość estymatora najmniejszych kwadratów $\hat{\beta}_i^{LS}$. Konsekwencje różnic w tych podejściach dotyczące błędu selekcji, FDR i oczekiwanej frakcji poprawnych selekcji PSR nie wydają się do końca rozpoznane.

Sama rozprawa doktorska mgr Brzyskiego prezentuje bardzo dobry poziom matematyczny i wyniki teoretyczne w niej przedstawione dają w moim przekonaniu istotne rozwinięcie metody SLOPE. Problem rozwiązania dla grupowego SLOPE w przypadku macierzy ortogonalnej został pomysłowo zredukowany do problemu SLOPE z diagonalną macierzą eksperymentu, a problem wynikowy szczegółowo zanalizowany. Za główne wyniki matematyczne w pracy dotyczące gSLOPE uważam twierdzenie 5.3.1, wniosek 4.2.1 i techniczne lematy 4.2.2 i 4.2.3 oraz, dla estymatora ODS, twierdzenie 7.1.1. Podoba mi się również pomysłowa analiza siły sygnałów przedstawiona w 5.4, która pozwala na odpowiednią kalibrację trudności problemu regresyjnego. W tym kontekście nie jest dla mnie problemem pojawienie się pracy konferencyjnej Gossman i inni (2015) wymienionej na str. 23, której elementem wspólnym z rozprawą jest definicja gSLOPE dla pewnego szczególnego układu wag.

Zawartość merytoryczna rozprawy jest przedstawiona bardzo jasno, dowody są przejrzyste i formalnie poprawne, a całość napisana dobrą angielszczyzną. Autor udowadnia swoją rozprawą, że jest zdolnym badaczem, z biegłością posługującym się metodami optymalizacji, statystyki wielowymiarowej i analizy danych.

Samo przedstawienie problemu selekcji grup w problemach regresyjnych we wstępie jest w moim odczuciu zbyt wąskie, gdyż pomija się inne niż FDR istotne własności selektorów grup zmiennych, w szczególności ich zgodność i kontrolę błędu selekcji (w szczególności prace Bach (EJS, 2007), Nardi, Rinaldo (EJS, 2008)) skupiając się jedynie na kontroli

(iii) SLOPE nie jest jedyną propozycją wykorzystania procedur związanych z procedurą B-H w selekcji modelu liniowego (koniec str. 7). Naturalna metoda została zaproponowana m.in. w Bunea i inni (JSPI, 2007).

(iv)$(5.10)_2$: raczej: $||\tilde{y}_{I_i}||w_i = |1 + \mu w_i^2|c_i$.

(v) Nazywanie mocą oczekiwanej frakcji prawidłowo zidentyfikowanych grup (str. 27) jest obocznością nomenklaturową; zwykle wielkość ta określana jest jako PSR (*Positive Selection Rate*).

Podsumowując, moim przekonaniu praca zyskałaby na właczeniu analizy pewnych istotnych zagadnień, o których piszę wyżej, a przede wszystkim porównaniu, zarówno praktycznym jak i teoretycznym estymatora gSLOPE z innymi podejściami, nawet kosztem rozdziału 7, który jest oderwany od całości rozprawy. Nie ulega jednak dla mnie jednak wątpliwości, że mamy do czynienia z bardzo dobrą rozprawą doktorską.

Konkuzja

W mojej opinii rozprawa doktorska przedstawiona przez mgr Damiana Brzyskiego spełnia ustawowe wymogi stawiane rozprawom doktorskim w dziedzinie nauk matematycznych, dyscyplina matematyka i wnoszę o dopuszczenie go do dalszych etapów przewodu doktorskiego.

Jan Mielniczuk