

Joanna Trzęsiok

Uniwersytet Ekonomiczny w Katowicach

PRZEGLĄD METOD REGULARYZACJI W ZAGADNIENIACH REGRESJI NIEPARAMETRYCZNEJ

Streszczenie: Wiele nieparametrycznych funkcji regresji w trakcie wykonywania algorytmu jest systematycznie poprawianych, tak by końcowy model charakteryzował się jak najlepszym dopasowaniem do danych ze zbioru uczącego. W efekcie otrzymujemy modele o niskich wartościach błędów resubstytucji i wysokiej złożoności, które jednak charakteryzują się niewielką zdolnością uogólniania, rozumianą jako zdolność poprawnej predykcji na nowych obiektach. Zachodzi potrzeba przeciwdziałania temu zjawisku. Proces uproszczenia postaci modelu przy jednoczesnym kontrolowaniu jego dopasowania nazywamy regularyzacją. W artykule przedstawione i porównane zostały techniki regularyzacji wykorzystywane w nieparametrycznych metodach regresji.

Słowa kluczowe: regularyzacja, regresja nieparametryczna.

1. Wstęp

W dobie dynamicznie rozwijających się narzędzi informatycznych oraz technik numerycznych zaproponowanych zostało wiele metod nieparametrycznych, w których funkcje regresji, budowane w kolejnych krokach algorytmów, są systematycznie poprawiane, tak by końcowy model charakteryzował się jak najlepszym dopasowaniem do danych ze zbioru uczącego. W efekcie otrzymujemy modele o niewielkich wartościach błędów resubstytucji, które jednak nie gwarantują równie niskich wartości błędów predykcji. Przeciwnie, małe błędy na zbiorze uczącym zazwyczaj uzyskujemy dla modeli charakteryzujących się wysoką złożonością i tym samym dających duże błędy prognoz na zbiorze testowym. Niskie błędy resubstytucji i wysoka złożoność modelu prowadzą do nadmiernego dopasowania funkcji regresji do danych ze zbioru uczącego (*overfitting*). Zachodzi więc potrzeba przeciwdziałania temu zjawisku poprzez uproszczenie postaci modelu przy jednoczesnym kontrolowaniu jego dopasowania. Technikę tę nazywamy **regularyzacją** (*regularization*).

Celem artykułu było przedstawienie i usystematyzowanie zaproponowanych w literaturze metod regularyzacji stosowanych w budowie nieparametrycznych modeli regresji. Techniki regularyzacji są często dedykowane odpowiednim metodom, przez co niemożliwe jest ich porównanie w kontekście jakości predykcji. Dlatego też w artykule zaproponowano zestawienie i pogrupowanie tych metod jedynie pod względem zastosowanych w nich mechanizmów zapobiegania nadmiernemu dopasowaniu modeli do danych ze zbioru uczącego.

2. Regularyzacja jako kompromis między dokładnością a złożonością modelu

Metody regularyzacji, wprowadzone do większości nieparametrycznych metod regresji, oparte są najczęściej na jednym z dwóch alternatywnych podejść. Pierwszą z tych technik przedstawili Hastie, Tibshirani i Friedman [2001, s. 34]. Zaproponowali oni, aby przeciwdziałać nadmiernemu dopasowaniu funkcji regresji poprzez minimalizację funkcjonu $PRSS(f, \lambda)$ (*penalized residual sum of squares*) określającego kompromis pomiędzy dokładnością a złożonością modelu:

$$PRSS(f, \lambda) = RSS(f) + \lambda \cdot J(f). \quad (1)$$

Pierwszy składnik tego funkcjonu – $RSS(f)$, jest miarą jakości dopasowania funkcji regresji, obliczoną na zbiorze uczącym, natomiast $J(f)$ odzwierciedla stopień złożoności modelu. Parametr $\lambda \geq 0$, nazywany współczynnikiem kary (*penalty*), określa proporcje pomiędzy składowymi funkcjonu $PRSS(f, \lambda)$. Im mniejsza jest jego wartość, tym większy nacisk położony jest na dopasowanie modelu do danych ze zbioru uczącego. Natomiast dla dużych wartości λ uzyskujemy modele charakteryzujące się niską złożonością.

Takie podejście do zagadnienia regularyzacji funkcji regresji zostało zastosowane między innymi w przypadku krzywych sklejanых czy drzew regresyjnych.

Metoda krzywych sklejanых (zob. np. [Hastie, Tibshirani, Friedman 2001, s. 117-120]) jest przykładem regresji jednowymiarowej o charakterze lokalnym. Polega ona na podziale dziedziny zmiennej X na K rozłącznych przedziałów za pomocą uporządkowanego zbioru punktów $\{\xi_k\}_{k=1, \dots, K-1}$, nazywanych węzłami. Za pomocą klasycznej metody najmniejszych kwadratów do obserwacji należących do każdego przedziału $[\xi_{k-1}, \xi_k]$ dopasowywany jest wielomian f_k , na który nałożone zostają dodatkowe warunki zapewniające odpowiednią gładkość funkcji regresji w węzłach.

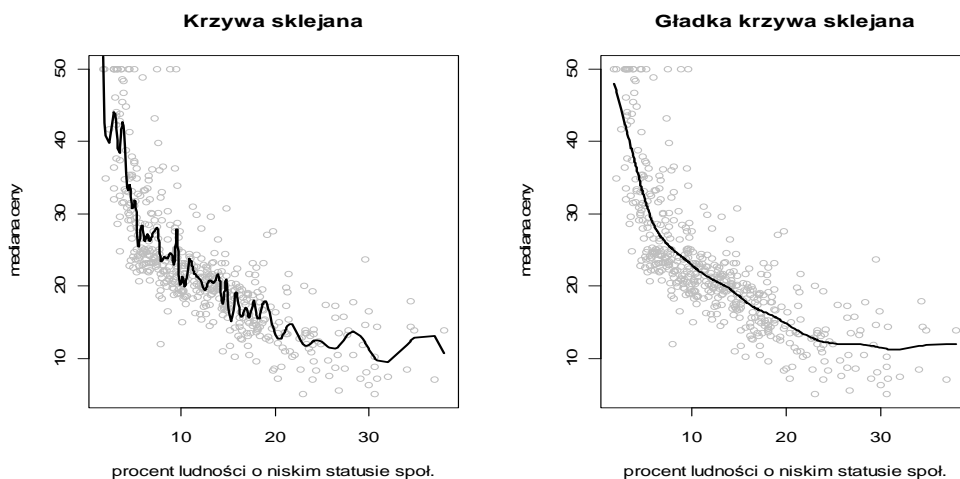
Parametrami w metodzie krzywych sklejanых są: liczba i położenie węzłów oraz stopień wielomianów f_k . Badania empiryczne pokazują, że dobrymi własnościami charakteryzują się krzywe „sklejone” z wielomianów stopnia 3, których wę-

zły są kwantylami rozkładu zmiennej objaśniającej. Największy problem stanowi jednak dobór liczby węzłów, ponieważ konsekwencją przyjęcia zbyt małych wartości tego parametru jest uzyskanie modelu o niskiej dokładności, natomiast za wysokie wartości K prowadzą do nadmiernego dopasowania funkcji regresji do danych ze zbioru uczącego. Zachodzi więc konieczność uzyskania kompromisu pomiędzy dokładnością a złożonością modelu.

Rozwiązaniem problemu nadmiernego dopasowania funkcji sklepanych do danych ze zbioru uczącego, a także dylematu związanego z doбором liczby węzłów, jest zastosowanie **gładkich krzywych sklepanych**. Są one rozwiązaniem zadania minimalizacji funkcjonału:

$$PRSS(f, \lambda) = \sum_{j=1}^N (y_j - f(x_j))^2 + \lambda \int (f''(t))^2 dt, \quad (2)$$

w którym pierwszy człon (suma kwadratów reszt) jest miarą dopasowania f do danych ze zbioru uczącego, drugi zaś określa złożoność modelu, która związana jest z gwałtownymi, oscylacyjnymi zmianami wartości estymowanej funkcji.



Rys. 1. Wykresy krzywych sklepanych obrazujące zależność zm. *MEDV* od *LSTAT*. Lewy wykres – krzywa sklejana 3 rzędu z 50 węzłami. Prawy wykres – gładka krzywa sklejana

Źródło: opracowanie własne.

W celu zilustrowania omawianego zagadnienia, w programie statystycznym **R**, za pomocą funkcji `bs` oraz `smooth.spline`, zbudowane zostały dwie krzywe sklepane dla danych ze zbioru *Boston*¹, z którego wybrano: *MEDV* – medianę wartości domu (w tys. \$) jako zmienną zależną, oraz *LSTAT* – procent ludności o ni-

¹ Dane w zbiorze *Boston* zostały zebrane w roku 1978 przez Harisona i Rubinfeld. Publikacja w artykule [Harrison, Rubinfeld 1978].

skim statusie społecznym – jako zmienną objaśniającą. Dla pierwszej krzywej sklejaney (zob. lewy wykres na rys. 1) przyjęto zbyt dużą liczbę węzłów (50 węzłów), w związku z czym obserwujemy gwałtowne, oscylacyjne zmiany wartości estymowanej funkcji, co oznacza, że model jest nadmiernie dopasowany do danych. Rozwiązaniem tego problemu jest gładka krzywa sklejana (prawy wykres na rys. 1), którą otrzymujemy minimalizując funkcjonal (2).

Z analogicznym podejściem do zagadnienia regularyzacji mamy do czynienia w metodzie **drzew regresyjnych** [Breiman i in. 1984], która polega na podziale przestrzeni zmiennych objaśniających na rozłączne obszary R_p (dla $p = 1, \dots, P$) i przypisaniu wszystkim obserwacjom z danego obszaru odpowiedniej wartości γ_p :

$$\mathbf{x}_i \in R_p \Rightarrow f(\mathbf{x}_i) = \gamma_p = \frac{1}{n_p} \sum_{\mathbf{x}_i \in R_p} y_i \quad \text{dla } p = 1, \dots, P. \quad (3)$$

Model w metodzie drzew regresyjnych ma postać addytywną daną wzorem:

$$T(\mathbf{x}, \Theta) = \sum_{p=1}^P \gamma_p \cdot I(\mathbf{x} \in R_p), \quad (4)$$

z parametrami $\Theta = \{R_p, \gamma_p\}_{p=1, \dots, P}$. Jest on sumą modeli lokalnych, w postaci stałej γ_p , budowanych w rozłącznych i homogenicznych obszarach R_p .

Algorytm wyznaczania drzewa regresyjnego polega na szukaniu optymalnego podziału przestrzeni zmiennych, tj. takiego podziału, dla którego błąd predykcji modelu będzie jak najmniejszy. W pierwszym etapie tego algorytmu zazwyczaj otrzymujemy drzewo o bardzo niskim błędzie resubstytucji, który jednak nie gwarantuje równie niskich wartości błędu na zbiorze testowym. Przeciwnie, jeśli błąd resubstytucji maleje, to rośnie złożoność modelu oraz błąd predykcji. W celu wyeliminowania tego zjawiska stosuje się mechanizm regularyzacji, polegający na tzw. przycinaniu krawędzi drzewa (*pruning*) przy użyciu miary odzwierciedlającej kompromis między złożonością modelu a jego jakością (por. [Gatnar 2008, s. 43]):

$$PRSS(T, \lambda) = \sum_{j=1}^N \sum_{\mathbf{x}_j \in R_p} (y_j - \gamma_p)^2 + \lambda P. \quad (5)$$

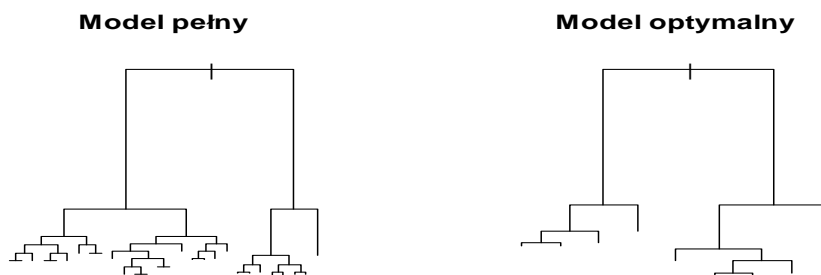
Pierwszy składnik funkcjonau (5) jest miarą dokładności drzewa, P to liczba liści i zarazem ocena złożoności modelu, natomiast parametr $\lambda \geq 0$ określa proporcje pomiędzy tymi składowymi. Duże wartości λ prowadzą do uzyskania podziału przestrzeni na niewiele segmentów, niskie zaś dają drzewo o dużej liczbie liści.

Proces przycinania rozpoczyna się od zbudowania modelu pełnego (maksymalnego) T_0 , który w kolejnych krokach algorytmu jest systematycznie przycina-

ny w ten sposób, by nie nastąpiło znaczne pogorszenie jego dokładności. Powstaje sekwencja drzew:

$$T_0, T_1, T_2, \dots, T_M, \quad (6)$$

z której wybierane jest drzewo optymalne pod względem przyjętej miary jakości, którą może być np. błąd średniokwadratowy obliczony za pomocą sprawdzania krzyżowego, czy reguła jednego odchylenia standardowego, zaproponowana przez Breimana w 1984 (por. [Gatnar 2008, s. 44]).



Rys. 2. Wykresy drzew regresyjnych zbudowanych na zbiorze *Boston*. Lewy wykres – model pełny z 28 liśćmi. Prawy wykres – model optymalny, po zastosowaniu procedury przycinania, o 9 liściach

Źródło: opracowanie własne.

Dla zilustrowania mechanizmu regularyzacji wyznaczone zostały dwa modele drzew regresyjnych na zbiorze danych *Boston*. Podobnie jak w poprzednim przykładzie zmienną zależną była *MEDV* (mediana wartości domu), natomiast jako zmienne objaśniające przyjęto wszystkie pozostałe (13) zmienne. W programie statystycznym **R**, za pomocą funkcji `rpart`, zbudowane zostało pełne drzewo regresyjne (lewy wykres na rys. 2). Drzewo to miało 28 liści, a jego błąd średniokwadratowy, obliczony na zbiorze testowym, wynosił $MSE(T_0) = 21,59$. Model ten został poddany procedurze przycinania (`prune`), w wyniku której uzyskano drzewo o 9 liściach oraz błędzie $MSE(T_{opt}) = 22,44$ (zob. prawy wykres na rys. 2). Nastąpił zatem niewielki wzrost błędu predykcji przy jednoczesnym znacznym uproszczeniu modelu.

3. Regularyzacja jako etap budowy modelu regresji

Sposobem zapobiegania nadmiernemu dopasowaniu funkcji do danych ze zbioru uczącego są również mechanizmy regularyzacji będące integralną częścią algorytmu budowy modelu. Z podejściem takim mamy do czynienia między innymi w metodach PPR oraz POLYMARS [Hastie i in. 2001, s. 347-350, 283-287], gdzie w

pierwszym etapie zbudowany zostaje model jak najlepiej dopasowany do danych ze zbioru uczącego. W kolejnym kroku zostaje on uproszczony, tak by w jak najmniejszym stopniu pogorszyć jakość predykcji. Jest to postępowanie podobne do wyznaczania modelu optymalnego (poprzez przycinanie) w metodzie drzew regresyjnych, przy czym w tamtym przypadku badacz miał możliwość zbudowania również modelu pełnego i porównania obu drzew, co w metodach PPR czy POLYMARS niestety nie jest możliwe.

Metoda POLYMARS została zaproponowana przez Kooperberga, Bose'a i Stone'a [1997]. Model w tej metodzie ma postać addytywną:

$$f(\mathbf{X}) = \alpha_0 + \sum_{k=1}^K \alpha_k h_k(\mathbf{X}), \quad (7)$$

w której funkcje składowe h_k mają postać iloczynów tensorowych:

$$h_k(\mathbf{X}) = u_k \prod_{l=1}^{L_k} (X_{v(k,l)} - \xi_{v(k,l)})_+, \quad (8)$$

gdzie $u_k \in \{-1, 1\}$, $L_k \in \{1, 2\}$ oznacza liczbę czynników (funkcji sklepanych pierwszego rzędu) tworzących k -ty składnik modelu, a $v(k, l) \in \{1, \dots, m\}$ wskazuje numer zmiennej tworzącej l -ty czynnik k -tego składnika modelu.

Algorytm budowy modelu POLYMARS składa się z dwóch głównych etapów: dołączania zmiennych do modelu oraz ich eliminacji. Pierwszy z tych etapów polega na systematycznym wprowadzaniu do modelu takich funkcji składowych, które w największym stopniu redukują błąd średniokwadratowy oraz zachowują postać iloczynu tensorowego (8). W wyniku zastosowania tej procedury otrzymujemy model (maksymalny) o niskim błędzie resubstytucji i wysokiej złożoności, który jest nadmiernie dopasowany do danych ze zbioru uczącego, więc wymaga zastosowania mechanizmu regularyzacji.

Proces regularyzacji polega na stworzeniu sekwencji modeli. W każdym kroku z modelu maksymalnego zostaje wyeliminowana jedna funkcja składowa, ta której usunięcie powoduje najmniejszy wzrost sumy kwadratów reszt. Model taki zapamiętujemy jako \hat{f}_K , gdzie K jest liczbą funkcji składowych. Z sekwencji modeli wybrany zostaje ten, który jest najlepszy w sensie przyjętego kryterium, którym w tym przypadku jest uogólnione kryterium sprawdzania krzyżowego:

$$\text{GCV}(K) = \frac{\sum_{i=1}^n (y_i - \hat{f}_K(\mathbf{x}_i))^2}{n \left(1 - \frac{Z(K)}{n}\right)^2}. \quad (9)$$

Funkcja $Z(K)$ jest miarą złożoności modelu regresyjnego, który zawiera K funkcji składowych. Proponuje się przyjęcie jej wartości na poziomie:

$$Z(K) = (1 + \gamma) \cdot K, \quad (10)$$

gdzie $\gamma \in \{2, 3, 4\}$ jest stałą oznaczającą liczbę parametrów modelu, związaną z funkcjami składowymi (por. [Hastie i in. 2001, s. 286]).

Kryterium $GCV(K)$ jest miarą wyrażającą kompromis pomiędzy dopasowaniem modelu a jego złożonością. Licznik formuły zapisanej wzorem (9) to miara dopasowania, natomiast mianownik wyraża stopień złożoności modelu. Funkcja \hat{f}_K , dla której wartość uogólnionego kryterium sprawdzania krzyżowego jest największa, jest tym samym rozwiązaniem problemu regularyzacji. Procedura eliminacji zmiennych pozwala na wyznaczenie modelu optymalnego, zazwyczaj znacznie uproszczonego, kosztem niewielkiej straty dokładności.

W celu zilustrowania omówionej techniki regularyzacji zbudowany został, za pomocą funkcji `polymars` w programie **R**, model regresyjny na zbiorze *Boston*. Zmienną zależną jest w tym przypadku również *MEDV*, natomiast jako zmienne objaśniające przyjęto: *LSTAT* (procent ludności o niskim statusie społecznym), *RM* (średnią liczbę pokoi), *DIS* (odległość od miejsc zatrudnienia), *NOX* (koncentrację tlenu azotu), *CRIM* (wskaźnik przestępstw) i *PTRATIO* (liczbę uczniów przypadających na jednego nauczyciela). Etapy budowy modelu przedstawiono w tab. 1.

W drugiej kolumnie tab. 1 występują wartości 0 lub 1, które wskazują, z jakim etapem mamy do czynienia (1 – następuje dodanie nowej zmiennej do modelu, 0 – usunięcie wyrażenia z modelu). Kolumna *RSS* zawiera sumy kwadratów reszt liczone w kolejnych krokach procedury, natomiast kolumna *GCV* przedstawia kolejne wartości uogólnionego kryterium sprawdzania krzyżowego. W naszym przypadku najmniejsza wartość *GCV* to 11,98 dla modelu zawierającego 20 funkcji składowych. Model ten jest więc rozwiązaniem zadania regularyzacji. Składa się z mniejszej ilości składowych niż model maksymalny (z etapu 25 w tab. 1), a charakteryzuje się niewiele większą wartością *RSS*.

Analogiczną technikę regularyzacji wykorzystano w **metodzie PPR** [Friedman, Stuetzle 1981]. Metoda ta oparta jest na rzutowaniu danych z przestrzeni wielowymiarowej w przestrzeń o dużo niższym wymiarze, w której łatwiej jest badaczowi zaobserwować własności badanego zbioru. Funkcja regresji w metodzie PPR ma postać addytywną:

$$Y = f(\mathbf{X}) = \alpha_0 + \sum_{k=1}^K \beta_k g_k(\mathbf{a}_k^T \cdot \mathbf{X}), \quad (11)$$

gdzie $\mathbf{a}_k \in \mathbf{R}^n$ (dla $k = 1, \dots, K$) są unormowanymi wektorami, nazywanymi kierunkami rzutowania, a g_k to funkcje jednej zmiennej o parametrach β_k .

Tabela 1. Etapy budowy modelu POLYMARS na zbiorze *Boston*

Etap	0/1	l. skład.	RSS	GCV	Etap	0/1	l. skład.	RSS	GCV
1	1	1	42 716,30	85,77	26	0	24	4 118,36	12,40
2	1	2	19 472,38	39,73	27	0	23	4 143,62	12,23
3	1	3	14 026,45	29,08	28	0	22	4 175,90	12,09
4	1	4	12 364,69	26,06	29	0	21	4 219,72	11,99
5	1	5	11 146,50	23,88	30	0	20	4 297,83	11,98
6	1	6	8 980,56	19,56	31	0	19	4 419,02	12,09
7	1	7	8 578,76	19,00	32	0	18	4 567,30	12,27
8	1	8	8 148,36	18,35	33	0	17	4 743,39	12,51
9	1	9	7 812,78	17,90	34	0	16	4 915,39	12,73
10	1	10	6 495,92	15,14	35	0	15	5 120,61	13,03
11	1	11	6 163,56	14,61	36	0	14	5 278,56	13,19
12	1	12	5 814,67	14,03	37	0	13	5 526,93	13,57
13	1	13	5 721,55	14,05	38	0	12	5 920,79	14,28
14	1	14	5 394,27	13,48	39	0	11	6 620,06	15,69
15	1	15	5 201,65	13,23	40	0	10	7 014,39	16,34
16	1	16	5 123,51	13,27	41	0	9	7 240,77	16,59
17	1	17	5 040,04	13,29	42	0	8	7 956,30	17,92
18	1	18	4 957,22	13,32	43	0	7	8 432,11	18,67
19	1	19	4 880,58	13,36	44	0	6	9 609,75	20,93
20	1	20	4 791,33	13,36	45	0	5	10 355,47	22,18
21	1	21	4 656,88	13,23	46	0	4	10 447,67	22,02
22	1	22	4 585,70	13,28	47	0	3	15 439,31	32,01
23	1	23	4 531,71	13,38	48	0	2	19 472,38	39,73
24	1	24	4 175,35	12,57	49	0	1	42 716,30	85,77
25	1	25	4 094,35	12,57					

Źródło: opracowanie własne.

Algorytm budowy modelu (11) opiera się na minimalizacji błędu empirycznego. Najmniej ugruntowanym elementem tej metody jest wybór liczby funkcji składowych – K . Wartość tego parametru zazwyczaj podaje użytkownik. Proste podejście, wykorzystywane w algorytmie SMART (*smooth multiply additive regression*) [Friedman 1984], polega na tym, że badacz podaje dwie wartości tego parametru: K_{pocz} – początkową (maksymalną) liczbę funkcji składowych oraz K_{konc} – liczbę funkcji użytych w końcowym modelu.

W pierwszym kroku stworzony zostaje model (maksymalny) złożony z K_{pocz} funkcji g_k , który w oparciu o przyjęte kryterium jest w kolejnym etapie stopniowo przycinany, aż do uzyskania funkcji f , która stanowi sumę K_{konc} składowych g_k .

Wykorzystanie algorytmu SMART oznacza stworzenie sekwencji modeli, z których każdy kolejny jest pewną modyfikacją poprzedniego. Końcowy model otrzymujemy po zastosowaniu techniki regularyzacji, jaką jest tutaj przycinanie, czyli redukcja kolejnych funkcji składowych z modelu.

4. Podsumowanie

Przedstawione w artykule nieparametryczne metody regresji podzielone zostały na dwie grupy w kontekście wykorzystywanych w nich technik regularyzacji. Pierwsza z nich to metody, w których model optymalny budowany jest poprzez minimalizację funkcjonału określającego kompromis pomiędzy złożonością a dokładnością modelu. Do drugiej grupy zaliczono metody, w których mechanizm regularyzacji jest integralną częścią algorytmu budowy modelu.

Techniki regularyzacji dedykowane są odpowiednim metodom, dlatego nie jest możliwe porównanie, pod względem jakości predykcji, modeli zbudowanych tą samą metodą regresji z wykorzystaniem różnych mechanizmów regularyzacji. Jednak w przypadku metod, takich jak: krzywe sklepane, drzewa regresyjne czy POLYMARS, możliwe jest porównywanie modeli budowanych bez jakichkolwiek mechanizmów regularyzacji, a także z wykorzystaniem takich technik. Przedstawione przykłady pokazują, że zastosowanie technik regularyzacji prowadzi do uproszczenia modelu, przy jednoczesnym niewielkim wzroście błędu predykcji w stosunku do modelu uzyskanego bez wprowadzenia mechanizmów regularyzacji. Mechanizmy te zapobiegają więc nadmiernemu dopasowaniu funkcji regresji do danych ze zbioru uczącego, a zatem w wyniku takiego postępowania uzyskujemy modele o większej zdolności uogólniania, rozumianej jako zdolność poprawnej predykcji na nowych obiektach.

Literatura

- Breiman L., Friedman J., Stone C., Olshen R., *Classification and Regression Trees*, CRC Press, London 1984.
- Gatnar E., *Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji*, PWN, Warszawa 2008.
- Harrison D., Rubinfeld D.L., *Hedonic prices and the demand for clean air*, „Journal of Environmental Economics and Management” 1978, no. 8.
- Hastie T., Tibshirani R., Friedman J., *The Elements of Statistical Learning*, Springer-Verlag, New York, 2001.
- Friedman J., *SMART user's Guide*, Technical Report 1, Department of Statistics, Stanford University, 1984.
- Friedman J., Stuetzle W., *Projection pursuit regression*, „Journal of the American Statistical Association” 1981, no. 76, s. 817-823.
- Kooperberg C., Bose S., Stone C., *Polychotomous regression*, „Journal of the American Statistical Association” 1997, no. 92, s. 117-127.

VARIOUS REGULARIZATION ISSUES IN NONPARAMETRIC REGRESSION

Summary: It is well known in statistics that fitting the training data too well can increase prediction risk on future predictions. In other words too large flexibility of the regression function would cause a learner to overfit the data, i.e. the learner would be able to model the noise in the data as well as the generating process and it leads to poor generalization. The process of finding the trade-off between minimizing the training error and controlling capacity is called regularization. The paper presents the issue and gives examples of regularization technique in case of nonparametric methods of regression.