

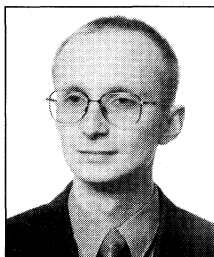
Adam Grzegorz POLAK, Janusz MROCZKA

POLITECHNIKA WROCŁAWSKA
KATEDRA METROLOGII ELEKTRONICZNEJ I FOTONICZNEJ

Ocena dokładności zregularyzowanej identyfikacji modelu płuc w czasie natężonego wydechu

Dr inż. Adam Grzegorz POLAK

– adiunkt w Katedrze Metrologii Elektronicznej i Fotonicznej na Wydziale Elektroniki Politechniki Wrocławskiej. Zajmuje się modelowaniem i identyfikacją obiektów złożonych oraz nieinwazyjnymi metodami pomiarowymi w medycynie.



Prof. dr hab. inż. Janusz MROCZKA

– kierownik Katedry Metrologii Elektronicznej i Fotonicznej na Wydziale Elektroniki Politechniki Wrocławskiej. Zajmuje się algorytmizacją problemu odwrotnego i jego aplikacją w pomiarach oraz analizą spektralną i polaryzacyjną promieniowania rozproszonego w układach dyspersyjnych.



Streszczenie

Płuca w czasie natężonego wydechu opisane mogą być nieliniowym dyskretnym modelem z dwunastoma parametrami. Identyfikacja tego modelu na podstawie zarejestrowanej krzywej natężonego wydechu jest źle uwarunkowana numerycznie. W referacie przeanalizowano zregularyzowaną estymację parametrów za pomocą regresji grzbietowej. Wyniki badań symulacyjnych pokazują, że przy odpowiednim doborze macierzy regularyzacji można spodziewać się zadowalająco dokładnej estymacji pola przekroju dróg oddechowych warunkującego ich opór oraz mniej dokładnego oszacowania podatności płuc i objętości zalegającej.

Abstract

Lungs during forced expiration can be described by a non-linear discrete model with twelve parameters. Identification of this model on the basis of a forced expiration curve is numerically ill-conditioned. Regularised estimation of parameters by means of the ridge regression has been analysed in the paper. Results of simulation studies show that, while matching the regularisation matrix properly, one can expect satisfying precision of estimation of the airway cross-sectional area which determines the airway resistance and less accurate assessment of the lung compliance and the residual volume.

Wprowadzenie

Nieinwazyjny test natężonego wydechu jest jedną z metod badania układu oddechowego. Polega on na możliwie silnym i głębokim wydechu poprzedzonym maksymalnym wdechem. Rejestrowana zależność pomiędzy natężeniem przepływu i objętością wydychanego powietrza w dużym przedziale objętości płuc nie zależy od wysiłku pacjenta i jednocześnie jest wrażliwa na zmiany właściwości mechanicznych układu oddechowego.

Zaproponowane w latach osiemdziesiątych i dziewięćdziesiątych modele matematyczne natężonego wydechu wykorzystują teorię prędkości falowej w elastycznych rurach i dychotomiczny model dróg oddechowych [2], [8], [10]. Złożoność tych modeli spowodowała, że nie były one stosowane do estymacji parametrów charakteryzujących układ oddechowy w oparciu o dane eksperymentalne.

We wcześniejszych pracach autorzy dokonali redukcji zaproponowanego przez nich kompleksowego modelu natężonego wydechu do postaci z 12 stopniami swobody oraz przeprowadzili analizę jego wrażliwości [12]. Na tej podstawie określono przybliżone właściwości estymatorów parametrów modelu uzyskiwanych metodą największej wiarygodności. Okazało się, że ten sposób identyfikacji jest źle uwarunkowany numerycznie (liczba uwarunkowania l_N odwracanej macierzy informacyjnej, czyli stosunek jej największej wartości osobliwej do wartości najmniejszej, wynosi ok. $1.3 \cdot 10^3$). Dalsze badania wykazały, że dane pomiarowe mogą być

podzielone na dwa zakresy (przed osiągnięciem szczytowego przepływu wydechowego PEF i po jego osiągnięciu), a parametry modelu zaliczone do dwóch grup. Każda z nich może być estymowana osobno na podstawie danych należących do wymienionych przedziałów. Takie podejście poprawia uwarunkowanie numeryczne identyfikacji.

Innym sposobem rozwiązywania zadań źle uwarunkowanych numerycznie jest regularyzacja procesu identyfikacji [5], [7], [9]. Jej idea polega na uwzględnieniu w minimalizowanej funkcji kryterialnej formuły zawierającej dodatkową wiedzę o identyfikowanym systemie. Zaletą regularyzacji jest pozbycie się złego uwarunkowania numerycznego estymacji (a tym samym nieprzewidywalnych błędów obliczeń numerycznych) oraz zmniejszenie wariancji estymatorów, wadą natomiast ich obciążenie.

Celem prezentowanych badań jest symulacyjna ocena dokładności estymacji parametrów modelu układu oddechowego w czasie natężonego wydechu z zastosowaniem zregularyzowanej formy minimalizowanej funkcji kryterialnej.

Podstawy teoretyczne

Model matematyczny natężonego wydechu

Nieliniowy model matematyczny natężonego wydechu został szczegółowo przedstawiony w [10]. W jego skład wchodzi cząstkowe równanie różniczkowe pozwalające obliczyć spadek ciśnienia w obrębie każdej z dychotomicznie dzielących się generacji dróg oddechowych (model o parametrach rozłożonych) z uwzględnieniem prędkości falowej rozchodzenia się zaburzeń mechanicznych w ich ścianach, charakterystyki statyczne oskrzeli i płuc oraz algorytm obliczeniowy. W prezentowanych badaniach symulacyjnych wartości dwunastu parametrów dobrano tak, aby model odzwierciedlał lekkie zaburzenia restrykcyjno-obturacyjne funkcjonowania płuc (pojemność życiowa $VC=3.5 \text{ dm}^3$, stosunek natężonej pojemności pierwszosekundowej do pojemności życiowej $FEV_1/VC \approx 70\%$, opór dróg oddechowych $R_{aw} \approx 0.47 \text{ kPa} \cdot \text{dm}^{-3} \cdot \text{s}$).

Regularyzacja grzbietowa

W celu regularyzacji zadania identyfikacji modelu płuc wybrano regresję grzbietową (ang.: *ridge regression*, *RR*) [5], [9], polegającą

na wprowadzeniu do funkcji kryterialnej stabilizatora Ω_{RR} w postaci:

$$\Omega_{RR}(\theta) = \frac{1}{2} k (\theta - \theta^\#)^T (\theta - \theta^\#) \quad (1)$$

gdzie θ jest wektorem p parametrów, $\theta^\#$ można uważać za „typowy” punkt pracy systemu w przestrzeni parametrów, a $k > 0$ jest skalar- nym parametrem regularyzacji. Bazując na metodzie najmniej- szych kwadratów (LS), gdzie z jest wektorem n danych pomiaro- wych a $y(\theta)$ funkcją modelową, otrzymuje się następującą postać zregularyzowanej funkcji kryterialnej V_{RR} :

$$V_{RR}(\theta) = V_{LS}(\theta) + \Omega_{RR}(\theta) = \frac{1}{2} (z - y(\theta))^T (z - y(\theta)) + \frac{1}{2} k (\theta - \theta^\#)^T (\theta - \theta^\#) \quad (2)$$

Zastosowanie regresji grzbietowej powoduje, że zadanie odwracania macierzy informacyjnej może być dobrze uwarunkowane, a obliczone estymatory pozbawione są wtedy nieprzewidywalnych błędów numerycznych. Uzyskuje się to dzięki dodaniu małej liczby k do diagonalnej macierzy i „przyciągnięciu” wartości parametrów w niewielkim stopniu wpływających na funkcje modelową w kie- runku wektora $\theta^\#$. Dodatkową zaletą tego typu regularyzacji w przypadku modeli nieliniowych pod względem parametrów (NLP) jest uzyskanie kroku Levenberga-Marquardta w algoryt- mach iteracyjnych [14], który jest zalecany w tego rodzaju oblicze- niach [13].

Regularyzacja identyfikacji systemów stosowana jest głównie w przypadku współliniowości występujących w modelach liniowych (LP) typu „czarna skrzynka” (np. [4], [5], [9]), można też znaleźć przykłady z modelami NLP (np. [7], [9]). Wspomniane współlinio- wości występują w macierzy η wrażliwości modelu względem jego parametrów:

$$\eta(\theta) = \frac{\partial y(\theta)}{\partial \theta} \quad (3)$$

Przy liniowej regresji grzbietowej zaleca się standaryzację macierzy η :

$$\eta_N = \eta N^{-1}, \quad N = \text{diag}(\text{diag}(\eta^T(\theta_0) \eta(\theta_0))). \quad (4)$$

Powoduje to, że w czasie estymacji każdy element diagonalny macierzy informacyjnej zwiększany jest proporcjonalnie i jest rów- ny $(1+k) \eta_{Ni}^T \eta_{Ni} = 1+k$ (i jest numerem kolumny/wiersza). Z dru- giej strony standardowe algorytmy identyfikacji nieliniowej posłu- gują się podstawową formą macierzy wrażliwości, uwzględniając ewentualnie wektor wag odpowiadających konkretnym punktom pomiarowym. Z tego powodu w niniejszej pracy zaproponowano zastąpienie standaryzacji macierzy wrażliwości standaryzującą ma- cierz regularyzacji $K=kN^2$ dającą następującą postać stabilizatora identyfikacji

$$\Omega_{RR}(\theta) = \frac{1}{2} (\theta - \theta^\#)^T K (\theta - \theta^\#) \quad (5)$$

W tym przypadku elementy diagonalne macierzy informacyj- nej nadal zwiększane są proporcjonalnie, lecz brak standaryzacji macierzy η powoduje, że wartości tych elementów mogą się znacznie różnić, co w efekcie może wpłynąć na nieco gorsze uwarunkowanie numeryczne identyfikacji. W ten sposób otrzy- mano szczególnie przypadkiem uogólnionej regresji grzbietowej (GRR), w której K jest dowolną macierzą diagonalną $p \times p$ o ele- mentach optymalizowanych w czasie procesu estymacji parametrów.

Nazwa „regresja grzbietowa” bierze się z kształtu zależności estymatorów od wartości parametru k . W przebiegach tych wystę- pują charakterystyczne „grzbiety”, których położenie pierwotnie wykorzystywano do wyznaczania k [5], [9]. Tak uzyskane wartości są jednak zawyżone [4], w związku z czym proponuje się inne spo- soby oszacowania parametru regularyzacji, np. $k = \hat{\sigma}^2 / \hat{\theta}_{RR}^T \hat{\theta}_{RR}$ [3] lub $k = p \hat{\sigma}^2 / \hat{\theta}_{RR}^T \hat{\theta}_{RR}$ [6], gdzie $\hat{\sigma}^2$ to estymator wariancji szumu addytywnego, a $\hat{\theta}_{RR}$ jest zregularyzowanym estymatorem wektora parametrów modelu.

Badania symulacyjne

Własności statystyczne zregularyzowanej identyfikacji modelu płuc określone zostaną na drodze symulacji numerycznych, które autorzy stosowali także we wcześniejszych pracach (np. [11], [12]). Wykorzystuje się w nich znaną strukturę modelu (przy założeniu bezbłędności tej struktury) oraz wybrany wektor wartości parame- trów θ_0 cechujący dany system. Pozwala to na symulację zachowa- nia się modelu w określonym punkcie przestrzeni parametrów i przeprowadzenie jego analizy. W szczególności, wykorzystując fakt że $\theta \approx \theta_0$, możliwe jest przewidzenie jakości estymacji parame- trów modelu bez potrzeby tworzenia algorytmów identyfikacji. Da- je to możliwość sprawdzenia np. wpływu postaci funkcji kryterial- nej, różnych schematów jej optymalizacji, ilości używanych punk- tów pomiarowych czy poziomu zakłóceń na dokładność estymacji. Należy przy tym mieć na uwadze, że wyniki przeprowadzanej ana- lizy są dokładne dla modeli LP i przybliżone (w otoczeniu punktu θ_0) dla modeli NLP, co w tym przypadku jest spowodowane zalez- nością $\eta = \eta(\theta)$.

Obciążenie, wariancja i wiarygodność estymatorów

Zastosowanie regresji grzbietowej powoduje, że estymatory pa- rametrów są obciążone. Wartość obciążenia b_{RR} można wyznaczyć linearyzując funkcję modelową w punkcie θ_0 traktowanym jako rzeczywisty wektor parametrów, i wynosi ono

$$b_{RR} = -[\eta^T(\theta_0) \eta(\theta_0) + K]^{-1} K (\theta_0 - \theta^\#) \quad (6)$$

Jak już wspomniano, zregularyzowana identyfikacja prowadzi do zmniejszenia wariancji estymatorów. W podobny jw. sposób wy- znaczyć można ich macierz kowariancji Σ_{RR} :

$$\Sigma_{RR} = \sigma^2 [\eta^T(\theta_0) \eta(\theta_0) + K]^{-1} \eta^T(\theta_0) \eta(\theta_0) [\eta^T(\theta_0) \eta(\theta_0) + K]^{-1} \quad (7)$$

gdzie σ^2 jest wariancją szumu addytywnego. Uzyskane asymptotyczne wariancje estymatorów będą mniejsze od dol- nej granicy Craméra-Rao, gdyż macierz K jest dodatnio okre- ślona [7].

Niepewność estymacji można wyznaczyć obliczając błędy śre- dniokwadratowe estymacji dla każdego z parametrów (wartość oczekiwana różnicy $\hat{\theta} - \theta_0$ w F):

$$\Delta \hat{\theta}_{RR} = \sqrt{b_{RR}^2(\hat{\theta}) + \sigma_{RR}^2(\hat{\theta})} \quad (8)$$

gdzie σ_{RR}^2 jest wariancją estymatora. Analizę porównawczą do- kładności estymacji poszczególnych parametrów łatwiej jest pro- wadzić korzystając z błędów względnych $\delta \theta_{RR} = \theta_0^{-1} \Delta \theta_{RR}$, nato- miast w celu doboru wartości parametru regularyzacji można

posłużyć się dodatkowym kryterium minimalnej wartości całkowitego względnego błędu średniokwadratowego danego wzorem:

$$\Delta_{MSE} = [\mathbf{b}_{RR}^T \Theta_0^{-1} \Theta_0^{-1} \mathbf{b}_{RR} + \text{tr}(\Theta_0^{-1} \Sigma_{RR} \Theta_0^{-1})]^{1/2} \quad (9)$$

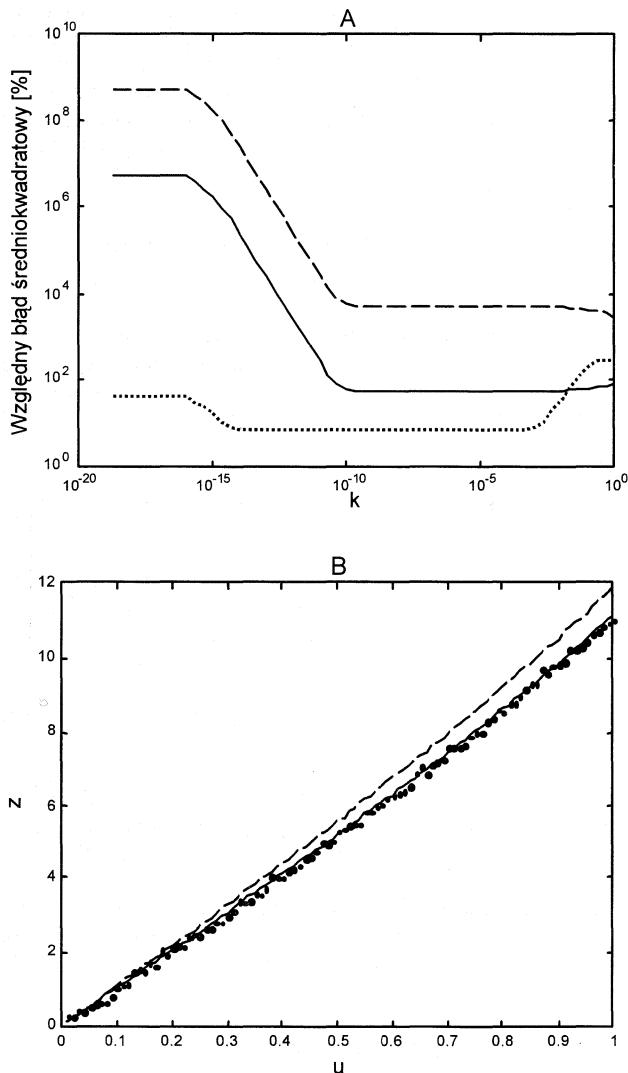
gdzie $\Theta_0 = \text{diag}(\theta_0)$.

Przykład

Korzyści płynące z zastosowania regresji grzbietowej przedstawione zostaną na przykładzie prostego modelu LP pobudzonego wejściem $\mathbf{u} = [0.01, 0.02, \dots, 1.00]^T$:

$$\begin{aligned} \mathbf{y} &= \theta_1 \mathbf{u} + \theta_2 (1 + \alpha \mathbf{u}) + \theta_3 \mathbf{u}^3, \\ \mathbf{z} &= \mathbf{y} + \mathbf{e}, \end{aligned} \quad (10)$$

gdzie $\alpha = 10^{-6}$ jest parametrem złego uwarunkowania numerycznego identyfikacji, $\theta_1 = 10$, $\theta_2 = 1.0$ i $\theta_3 = 0.1$, a \mathbf{e} to wektor realizacji addytywnego szumu losowego o odchyleniu standardowym 0.1 (mnożenie i potęgowanie wektora \mathbf{u} odnosi się do jego elementów).



Rys. 1. A) Zależność względnych błędów średniokwadratowych estymatorów parametrów od wartości współczynnika regularyzacji (θ_1 : —, θ_2 : - - -, θ_3 : ·····); B) Dopasowanie wyjścia modelu do danych (\hat{y}_{LS} : —, \hat{y}_{RR} : - - -)

W przypadku estymacji metodą LS liczba uwarunkowania numerycznego l_N wynosi ok. $7 \cdot 10^{15}$. Na rys. 1A pokazano wpływ wartości parametru k na względny błąd średniokwadratowy estymatorów zregularizowanych. Oszacowanie optymalnej wartości tego parametru na podstawie wzorów Farebrothera [3] i Hoerla *i in.* [6] daje odpowiednio wartości $k = 3 \cdot 10^{-6}$ i $k = 9 \cdot 10^{-6}$. Pozostają one w zgodzie z obszarem, w którym całkowity błąd średniokwadratowy jest minimalny (rys. 1A). Do estymacji parametrów za pomocą regresji grzbietowej i oszacowania jej dokładności wybrano $k = 5 \cdot 10^{-6}$ oraz $\theta^* = 0$, co w rezultacie zmniejszyło liczbę uwarunkowania do ok. $5 \cdot 10^5$. Wyjścia identyfikowanych modeli pokazano na rys. 1B, a estymację metodami LS i RR porównano w tab. 1.

Tab. 1. Porównanie przykładowych wyników estymacji i oszacowanie dokładności identyfikacji metoda najmniejszych kwadratów (LS) i regresji grzbietowej (RR)

θ_0		10.00	0.10	1.00
$\hat{\theta}$	LS	$-4.41 \cdot 10^7$	$4.41 \cdot 10^7$	0.73
	RR	5.04	5.04	1.05
$\delta \hat{\theta}$ [%]	LS	$5.5 \cdot 10^5$	$5.5 \cdot 10^5$	40
	RR	$4.9 \cdot 10^1$	$4.9 \cdot 10^1$	6.5

Uzyskane wyniki obrazują główne cechy regularyzacji: poprawę uwarunkowania numerycznego identyfikacji (zmniejszenie l_N z $7 \cdot 10^{15}$ do $5 \cdot 10^5$) i zwiększenie wiarygodności estymatorów. Wprawdzie dokładność wyznaczania dwóch pierwszych parametrów nadal nie jest zadowalająca (są one związane współliniowością występującą w modelu), lecz stosunek odchylenia standardowego do jego wartości oczekiwanej zmienia się znacznie z 40 do 6.5%.

Wyniki badań symulacyjnych

Wyznaczenie macierzy wrażliwości

Macierzy wrażliwości wyjścia modelu płuc w czasie natężonego wydechu, którym jest przepływ Q , na wartość dwunastu parametrów nie można wyznaczyć analitycznie. Z tego powodu dla objętości V z zakresu pojemności życiowej elementy tej macierzy obliczone zostały na drodze symulacji numerycznych w następujący sposób:

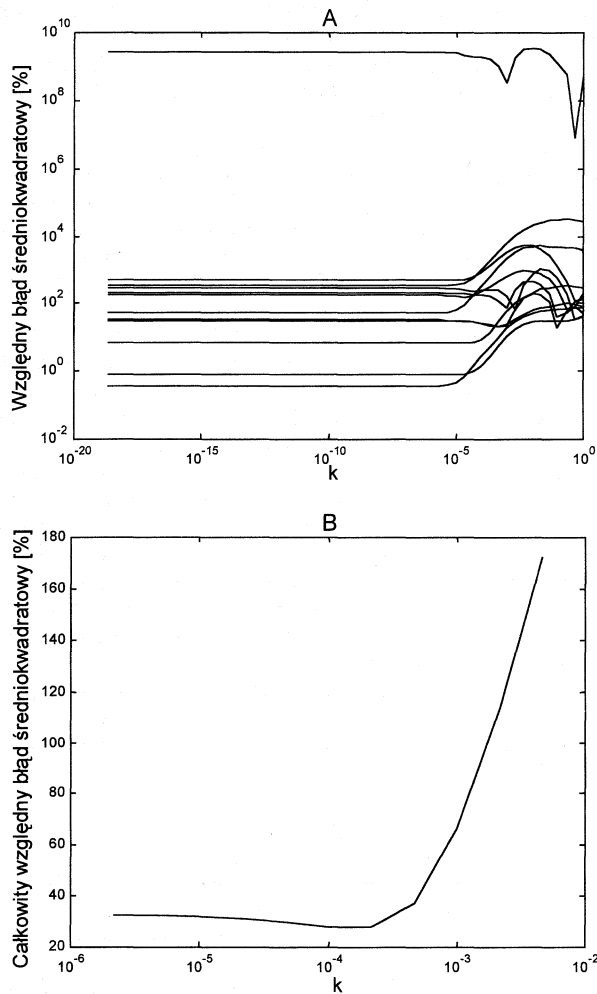
$$\eta_k(V_i) \approx \frac{Q(V_i, \theta_k + \Delta\theta_k) - Q(V_i, \theta_k - \Delta\theta_k)}{2\Delta\theta_k} \quad (11)$$

Przyrosty $\Delta\theta$ dobierano tak, aby wynikająca z nich zmiana krzywej natężonego wydechu wynosiła kilka procent [11], [12].

Ocena regresji grzbietowej z proporcjonalną macierzą K

Wyznaczona macierz wrażliwości modelu pozwala, przy założeniu, że odchylenie standardowe zakłóceń w rejestrowanym sygnale przepływu wynosi $0.05 \text{ dm}^3 \text{ s}^{-1}$ i przyjęciu $\theta^* = 0$, oszacować zależność obciążenia, wariancji i tym samym względnego błędu średniokwadratowego estymatorów od wartości parametru regularyzacji k (rys. 2A). Wartość tego parametru obliczona na podstawie wzoru Hoerla *i in.* wynosi $k = 4.4 \cdot 10^{-7}$, a całkowity względny błąd estymacji czterech pierwszych, najistotniejszych diagnostycznie parametrów (θ_1 i θ_2 określają pole przekroju poprzecznego oskrzeli, θ_3 to podatność płuc, a θ_4 jest objętością zalegającą), osiąga minimum dla $k = 1.5 \cdot 10^{-4}$ (rys. 2B). W tab. 2 porównano dokładność estymacji RR dla tych dwóch wartości k z metodą LS.

Żaden z przebadanych sposobów identyfikacji modelu płuc w natężonym wydechu nie dał zadowalających wyników. Wprawdzie regresja grzbietowa z $k = 1.5 \cdot 10^{-4}$ pozwalała na lepsze oszacowanie θ_4 (objętość zalegająca), dzieje się to jednak kosztem dokład-



Rys. 2. Właściwości regresji grzbietowej z macierzą regularyzacji $\mathbf{K}=\mathbf{kN}^2$: A) zależność względnych błędów średniokwadratowych estymatorów 12 parametrów modelu płuc od wartości parametru regularyzacji; B) zależność całkowitego względnego błędu estymacji czterech pierwszych parametrów od wartości parametru regularyzacji

ności wyznaczenia trzech pozostałych, istotnych diagnostycznie parametrów. Co gorsza, obydwie wartości parametru k nie polepszają uwarunkowania numerycznego estymacji, które dla LS i dwóch przypadków RR wynosi odpowiednio $1.28 \cdot 10^{31}$, $1.27 \cdot 10^{31}$ i $4.07 \cdot 10^{30}$.

Tab. 2. Zestawienie oszacowanej dokładności estymacji parametrów modelu płuc metodami LS i RR z macierzą regularyzacji $\mathbf{K}=\mathbf{kN}^2$

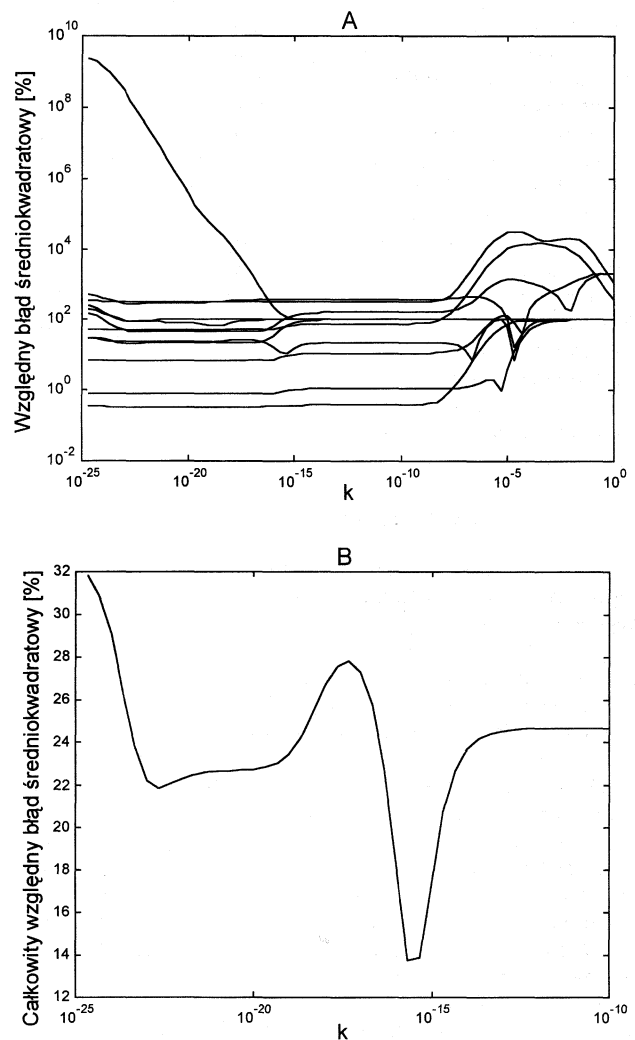
	$\delta\hat{\theta}_1$ [%]	$\delta\hat{\theta}_2$ [%]	$\delta\hat{\theta}_3$ [%]	$\delta\hat{\theta}_4$ [%]	$\delta\hat{\theta}_5$ [%]	$\delta\hat{\theta}_6$ [%]	$\delta\hat{\theta}_7$ [%]	$\delta\hat{\theta}_8$ [%]	$\delta\hat{\theta}_9$ [%]	$\delta\hat{\theta}_{10}$ [%]	$\delta\hat{\theta}_{11}$ [%]	$\delta\hat{\theta}_{12}$ [%]
LS	0.81	0.35	7.1	32	31	52	180	210	300	350	540	2.8×10^9
RR $k=4.4 \cdot 10^{-7}$	0.81	0.35	7.1	32	31	52	180	210	300	350	540	2.7×10^9
RR $k=1.5 \cdot 10^{-4}$	1.9	3.8	10	25	24	480	270	180	260	1.4×10^3	1.1×10^3	1.9×10^3

Ocena regresji grzbietowej z nieproporcjonalną macierzą \mathbf{K}

Wyniki powyższych badań pokazują, że zastosowanie regresji grzbietowej z macierzą \mathbf{K} zwiększającą proporcjonalnie wszystkie elementy diagonalne odwracanej macierzy nie daje zadowalają-

cych rezultatów, a w szczególności nie poprawia uwarunkowania numerycznego identyfikacji. Główną przyczyną są znaczne różnice we wrażliwości wyjścia modelu na wartości parametrów. To właśnie mała wrażliwość związana z niektórymi parametrami (np. θ_{12}) sprawia, że identyfikacja analizowanego modelu jest źle uwarunkowana [12]. Poprawy uwarunkowania numerycznego można się zatem spodziewać regularyzując przede wszystkim te elementy diagonalne macierzy informacyjnej, które związane są z małą wrażliwością modelu. Można to osiągnąć dodając np. do całej diagonalnej tą samą wartość, czyli stosując $\mathbf{K}=\mathbf{kI}$.

Oszacowane dokładności estymacji dla różnych wartości parametru regularyzacji pokazano na rys. 3A, a jego wpływ na całkowity względny błąd estymacji czterech pierwszych parametrów na rys. 3B.



Rys. 3. Właściwości regresji grzbietowej z macierzą regularyzacji $\mathbf{K}=\mathbf{kI}$: A) zależność względnych błędów średniokwadratowych estymatorów 12 parametrów modelu płuc od wartości parametru regularyzacji; B) zależność całkowitego względnego błędu estymacji czterech pierwszych parametrów od wartości parametru regularyzacji

Dalsze badania symulacyjne przeprowadzono dla trzech wartości parametru k : obliczonej wg wzoru Hoerla i in. ($k=4.4 \cdot 10^{-7}$), wyznaczonej dla minimum całkowitego błędu estymacji czterech pierwszych parametrów (rys. 3B, $k=3.0 \cdot 10^{-16}$) oraz dobranej tak, by liczba uwarunkowania numerycznego była nieco mniejsza od 10^{15} , którą to wartość można przyjąć za granicę złego uwarunkowania

Tab. 3. Zestawienie oszacowanej dokładności estymacji parametrów modelu płuc metodami LS i RR z macierzą regularyzacji $K=kI$

	$\delta\theta_1$ [%]	$\delta\theta_2$ [%]	$\delta\theta_3$ [%]	$\delta\theta_4$ [%]	$\delta\theta_5$ [%]	$\delta\theta_6$ [%]	$\delta\theta_7$ [%]	$\delta\theta_8$ [%]	$\delta\theta_9$ [%]	$\delta\theta_{10}$ [%]	$\delta\theta_{11}$ [%]	$\delta\theta_{12}$ [%]
LS	0.81	0.35	7.1	32	31	52	180	210	300	350	540	2.8×10^9
RR $k=3.0 \cdot 10^{-16}$	0.89	0.35	8.4	10	56	61	96	100	100	320	370	130
RR $k=1.5 \cdot 10^{-8}$	1.1	0.72	12	21	100	130	180	100	100	410	390	100
RR $k=4.4 \cdot 10^{-7}$	1.5	12	38	23	100	1.6×10^3	450	100	100	5.3×10^3	450	100

($k=1.5 \cdot 10^{-8}$). Wyniki oszacowania dokładności estymacji, wraz z dokładnością metody LS, umieszczono w tab. 3.

W analizowanym przypadku (nieproporcjonalna macierz K) za każdym razem uzyskiwano poprawę numerycznego uwarunkowania identyfikacji. Kryterium minimalnej wartości całkowitego błędu średniokwadratowego daje najbardziej wiarygodne estymatory wybranych czterech parametrów, lecz jednocześnie $I_N=4.7 \cdot 10^{22}$ wskazuje, że otrzymywane wartości mogą być obciążone nieprzewidywalnymi błędami numerycznymi. Z drugiej strony oszacowanie Hoerla zwiększa obciążenie estymatorów i pogarsza jakość identyfikacji czterech parametrów, dając zadanie dobrze uwarunkowane numerycznie ($I_N=3.2 \cdot 10^{13}$). Przypadkiem pośrednim jest regresja grzbietowa z parametrem k dającym graniczne uwarunkowanie ($I_N=9.4 \cdot 10^{14}$). Uzyskuje się wtedy dużą precyzję estymacji θ_1 i θ_2 z gorszą jakością oszacowania θ_3 i θ_4 . Pamiętać jednak należy, że w tym przypadku błędy obliczeń numerycznych nie powinny mieć znaczącego wpływu na wartości estymatorów.

Podsumowanie i wnioski

W referacie przedstawiono zagadnienia związane z identyfikacją nieliniowego modelu matematycznego płuc w czasie natężonego wydechu, skupiające się na jakości wyznaczania czterech jego parametrów o największym znaczeniu diagnostycznym. Wcześniejsze prace pokazały, że estymacja parametrów modelu nieliniową wersją metody najmniejszych kwadratów (lub analogiczną metodą największej wiarygodności) jest źle uwarunkowana numerycznie.

Jako metodę poprawy uwarunkowania numerycznego identyfikacji wybrano regularyzację grzbietową. Przedstawiono jej wersję ze standaryzacją macierzy wrażliwości, oraz jako alternatywę w optymalizacji nieliniowej, wprowadzono standaryzującą macierz regularyzacji. Linearyzując model w wybranym punkcie przestrzeni parametrów otrzymano wzory na przybliżoną wartość obciążenia, wariacji i błędu średniokwadratowego zregulowanych estymatorów modelu nieliniowego.

Przyjęta metodologia oparta o modelowanie i symulacje komputerowe pozwoliła ocenić przydatność regresji grzbietowej w poprawie jakości estymacji parametrów. Analiza identyfikacji z macierzą regularyzacji proporcjonalną do wrażliwości modelu pokazała, że w przypadku tym nie da się osiągnąć dobrego uwarunkowania numerycznego estymacji. W dodatku jakość estymatorów (stosunek ich błędów średniokwadratowych do wartości rzeczywistych) ulegała pogorszeniu. Było to spowodowane dużymi różnicami we wrażliwości modelu natężonego wydechu na jego parametry. Z tego powodu zaproponowana została nieproporcjonalna macierz regularyzacji, stabilizująca przede wszystkim estymatory parametrów związanych z małą wrażliwością. To podejście dało zadowalające rezultaty. Pokazano, że dla wartości parametru regularyzacji $k=1.5 \cdot 10^{-8}$ uzyskać można dobre oszacowanie parametrów θ_1 i θ_2 (określają one pole przekroju poprzecznego drzewa oskrzelowego): $\delta\theta_1=1.1\%$, $\delta\theta_2=0.72\%$ oraz mniej dokładne wartości podatności płuc ($\delta\theta_3=12\%$) i objętości

zalegającej ($\delta\theta_4=21\%$). Jednocześnie uwarunkowanie numeryczne zmniejszyło się z $1.3 \cdot 10^{31}$ do $9.4 \cdot 10^{14}$.

Porównując regresję grzbietową o $k=1.5 \cdot 10^{-8}$ z metodą najmniejszych kwadratów (tab. 3) wydawać by się mogło, że zwiększenie dokładności estymacji θ_4 odbywa się kosztem gorszego oszacowania pozostałych trzech istotnych diagnostycznie parametrów. Tak jednak nie jest, gdyż przy złym uwarunkowaniu numerycznym LS rzeczywiste odchylenia estymatorów od ich wartości oczekiwanych mogą być wielokrotnie większe niż obliczone odchylenia standardowe, co jest spowodowane nakładającymi się błędami w iteracyjnych obliczeniach numerycznych. Problem ten nie występuje w przypadku regularyzacji identyfikacji.

Głównym efektem przeprowadzonych badań jest wniosek, że zastosowanie zregulowanej identyfikacji powinno pozwolić na zadawalająco dokładną estymację parametrów określających pole przekroju dróg oddechowych (a tym samym opór przepływu) oraz mniej dokładne wyznaczenie podatności płuc i pojemności zalegającej. Należy w tym miejscu zwrócić uwagę na możliwość „automatycznego” doboru wartości parametru regularyzacji w czasie identyfikacji modelu nieliniowego. Można to osiągnąć stosując odpowiednie kryterium w czasie obliczeń iteracyjnych [13]. Wydaje się, iż kryterium tym może być minimalizacja całkowitego błędu średniokwadratowego estymacji wybranych czterech parametrów modelu, lub jej postać ważona uwzględniająca ich znaczenie diagnostyczne.

Otrzymane wyniki zachęcają i pozwalają w przyszłości podjąć próbę połączenia analizowanej wcześniej selekcji parametrów modelu natężonego wydechu z ich zregulowaną estymacją.

Literatura

- [1] D.M. ALLEN: The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* 1974, 16.
- [2] D. ELAD, R.D. KAMM, A.H. SHAPIRO: Mathematical simulation of forced expiration. *J. Appl. Physiol.* 1988, 65.
- [3] R.W. FAREBROTHER: The minimum mean square error linear estimator and ridge regression. *Technometrics* 1975, 17.
- [4] R.R. HOCKING: The analysis and selection of variables in linear regression. *Biometrics* 1976, 32.
- [5] A.E. HOERL, R.W. KENNARD: Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* 1970, 12.
- [6] A.E. HOERL, R.W. KENNARD, F.K. BALDWIN: Ridge regression: some simulations. *Comm. Statist.* 1975, 4.
- [7] A. JOHANSEN: On Tikhonov regularization, bias and variance in nonlinear system identification. *Automatica* 1997, 33.
- [8] R.K. LAMBERT, T.A. WILSON, R.E. HYATT, J.R. RODARTE: A computational model for expiratory flow. *J. Appl. Physiol.: Respirat. Environ. Exercise Physiol.* 1982, 52.
- [9] D.W. MARQUARDT: Generalized inverses, ridge regression, biased linear and non-linear estimation. *Technometrics* 1970, 12.
- [10] A.G. POLAK: A forward model for maximum expiration. *Comp. Biol. Med.* 1998, 28.
- [11] A.G. POLAK, J. MROCZKA: A metrological model for maximum expiration. *Measurement* 1998, 23.
- [12] A.G. POLAK, J. MROCZKA: Analiza wrażliwości modelu natężonego wydechu. *Materiały I Sympozjum „Modelowanie i Pomiary w Medycynie”*, Krynica, kwiecień 1999.
- [13] W.H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY, W.T. VETTERLING: *Numerical Recipes. The Art of Scientific Computing*. Cambridge University Press, New York 1989.
- [14] J. SJÖBERG, Q. ZHANG, L. LJUNG, A. BENVENISTE, B. DELYON, P.-Y. GLORENNEC, H. HJALMARSSON, A. JUDITSKY: Nonlinear black-box modeling in system identification: a unified overview. *Automatica* 1995, 31.