Dariusz BUCHCZIK
POLITECHNIKA ŚLĄSKA, INSTYTUT AUTOMATYKI,
ZAKŁAD SYSTEMÓW POMIAROWYCH

Procedura kalibracji przy wykorzystaniu metody najmniejszej mediany kwadratów w przypadku modeli wielowymiarowych

Dr inż. Dariusz BUCHCZIK

Ukończył studia na Wydziale Automatyki, Elektroniki i Informatyki Politechniki Śląskiej w 1996 r. Tytuł doktora uzyskał w 2005 r. Swoje zainteresowania naukowe koncentruje wokół odpornych procedur estymacji. Prowadzi badania związane z wykorzystaniem metody najmniejszej mediany kwadratów do estymacji charakterystyk wielowymiarowych. Interesuje się również problematyką kalibracji czujników przyspieszenia. Członek Komisji Metrologii Oddziału PAN w Katowicach.

e-mail: dariusz.buchczik@polsl.pl

Streszczenie

Odporne procedury estymacji umożliwiają zredukowanie wpływu błędów nadmiernych na wyniki pomiarów. Zastosowanie metody najmniejszej mediany kwadratów (MNMK) do wyznaczania charakterystyk aparatury pomiarowej wiąże się z koniecznością oceny otrzymanych wyników. W pracy zaproponowano metodę szacowania wariancji przewidywanej średniej wartości obserwacji dla MNMK. Przeprowadzono również badania symulacyjne w celu porównania wyników oszacowania z otrzymanymi w trakcie symulowanego eksperymentu. Uzyskane rezultaty potwierdzają poprawność proponowanych koncepcji.

Słowa kluczowe: wzorcowanie, estymatory odporne, metoda najmniejszej mediany kwadratów.

Multidimensional model calibration procedure by means of the least square median method

Abstract

Robust regression is commonly used to reduce effects caused by outliers in measurement data sets. Evaluation of calibration lines of measuring instruments determined with a use of a least median of squares (LMS) method is essential. Standard error of mean response might be used to estimate the quality of the calibration lines. An idea of estimation of the standard error of mean response in case of the LMS method is proposed in this paper. A priori procedure enables to perform the estimation before the measurements. A posteriori procedure is used after the measurements, when there are available its results. The procedures a priori and a posteriori have been verified in simulation research. The estimated and calculated on simulation values of LMS standard errors of mean response have been compared to proof correctness of the estimation method. Relatively small errors of estimation confirm that the procedure is correct.

Keywords: calibration, robust regression, least median of squares method.

1. Wstęp

Regresja liniowa należy do najważniejszych procedur statystycznych stosowanych praktycznie we wszystkich dziedzinach nauki. Wyznaczanie krzywych kalibracji jest klasycznym zadaniem, do którego rozwiązania stosuje się metodę regresji liniowej. Metoda najmniejszych kwadratów (MNK) jest najczęściej stosowana dla problemu dopasowania wartości współczynników, z góry założonego modelu, do uzyskanych w rezultacie eksperymentu pomiarowego danych.

Błędy nadmierne, wynikające z bardzo różnych przyczyn i często występujące w danych pomiarowych, mogą pozostawać niezauważone w trakcie zautomatyzowanych obliczeń komputerowych. Wpływ takich punktów na wartości współczynników regresji MNK bywa dramatyczny i prowadzi do otrzymania całkowicie zafałszowanych rezultatów. Procedury odporne należą do metod statystycznych opracowanych z myślą o redukcji zagrożenia płynącego ze strony błędów nadmiernych.

Do najbardziej obiecujących odpornych procedur estymacji należy metoda najmniejszej mediany kwadratów (MNMK). Wyznaczanie charakterystyk aparatury pomiarowej wyłącznie z wykorzystaniem MNMK, podobnie jak w przypadku MNK, wymaga oceny uzyskanych wyników, która polega na wyznaczeniu przedziału ufności dla przewidywanej średniej wartości obserwacji. Autor zaproponował koncepcję szacowania wariancji przewidywanej średniej wartości obserwacji dla MNMK, co jest niezbędne do wyznaczenia odpowiedniego przedziału ufności. Niekonstruktywny charakter kryterium metody uniemożliwia analityczne wyznaczenie oszacowania. Prowadzone badania miały charakter behawioralny.

2. Procedura wzorcowania

Wzorcowanie, nazywane również kalibracją, jest procedurą polegającą na ustaleniu, w ściśle określonych warunkach, relacji między wartościami wielkości mierzonej wskazywanymi przez przyrząd lub układ pomiarowy a odpowiednimi wartościami wielkości realizowanymi przez wzorce jednostki miary [6]. W trakcie wzorcowania wyznacza się model matematyczny przyrządu wzorcowanego z dokładnością do wartości współczynników.

Najczęściej stosuje się model liniowy

$$y_i = x_{i1}\theta_1 + ... + x_{ip}\theta_p + e_i, \tag{1}$$

poszukiwana charakterystyka ma postać

$$\hat{y}_i = x_{i1}\hat{\theta}_1 + ... + x_{ip}\hat{\theta}_p, \tag{2}$$

natomiast błędem resztowym jest różnica

$$r_i = y_i - \hat{y}_i. \tag{3}$$

3. Błędy nadmierne

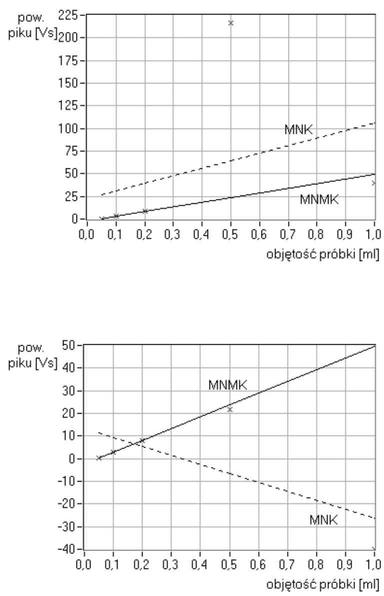
Występowanie błędów nadmiernych (grubych) jest poważnym problemem w miernictwie. Źródła powstawania tych błędów są bardzo różne i związane zarówno z pomyłkami obsługujących aparaturę jak i nieprzewidywalnymi czynnikami zewnętrznymi. Szacuje się, że w pomiarach inżynierskich do około 10% wyników jest obarczonych błędami nadmiernymi, a w niektórych dziedzinach pomiarów ich udział wzrasta nawet do około 30% [5].

Wzorcowanie wykonuje się szczególnie starannie i na ogół w warunkach laboratoryjnych, więc błędy nadmierne występują znacznie rzadziej. Jeden obarczony błędem nadmiernym wynik wzorcowania pociąga jednak za sobą wiele błędnych wyników pomiarów.

Przykład ilustrujący wpływ błędów nadmiernych na wyniki wzorcowania aparatury chromatograficznej przy wykorzystaniu procedur MNK oraz MNMK [2] przedstawiono na rys. 1.

a)

b)



Rys. 1. Przykładowe krzywe kalibracji chromatografu dla modelu jednowymiarowego dla różnych typów błędów grubych
Fig. 1. Gas chromatograph calibration lines of simple linear model for different types of outliers

3.1. Punkt załamania metody

Odporność procedury na występowanie w danych pomiarowych błędów grubych charakteryzuje się za pomocą punktu załamania [7]. Punkt załamania ε^* definiuje się jako najmniejszy iloraz liczby punktów m obarczonych błędami i liczebności zbioru danych n, taki że wartości współczynników regresji metody przyjmują wartości dowolnie różniące się od poprawnych. Wartość punktu załamania określa ile punktów pomiarowych musi być „dobrych", aby wartości współczynników regresji pozostały ograniczone.

4. Metoda najmniejszych kwadratów

Klasyczną metodą wyznaczania wartości współczynników regresji jest metoda najmniejszych kwadratów, której kryterium opiera się na minimalizacji sumy kwadratów błędów resztowych

$$\underset{\hat{\theta}}{\text{minimum}} \sum_{i=1}^{n} r_i^2 . \tag{4}$$

Szczególnie niekorzystną cechą MNK, biorąc pod uwagę jej zastosowanie do estymacji charakterystyk, jest całkowity brak odporności na błędy grube występujące w danych pomiarowych. Punkt załamania MNK wynosi 0, co oznacza, iż zaledwie jeden błąd nadmierny może spowodować, że estymowane współczynniki przyjmują dowolne wartości.

5. Procedury odporne

Jednym z możliwych rozwiązań uprzednio opisanych problemów jest zastosowanie procedur estymacji odpornych na błędy nadmierne. Do najbardziej znanych procedur odpornych zaliczają się: metoda najmniejszych modułów błędów, procedury Hubera, Beatona-Tukeya i wiele innych. Metody odporne na występowanie w wynikach pomiarów do 50% błędów nadmiernych opierają się na wykorzystaniu kryteriów medianowych. Zalicza się do nich m.in. procedurę medianową z powtarzaniem Siegela oraz metodę najmniejszej mediany kwadratów. Szerszy przegląd procedur odpornych przedstawiono w [5, 6, 7].

5.1. Procedura najmniejszej mediany kwadratów

Metoda najmniejszej mediany kwadratów opracowana została przez Rousseeuw [8] i polega na minimalizacji mediany kwadratów błędów resztowych. Kryterium metody ma postać

$$\underset{\hat{\theta}}{\text{minimum}} \ \underset{i}{\text{med}} \left\{ r_i^2 \right\}. \tag{5}$$

Wartość punktu załamania MNMK zależy od liczby parametrów estymowanej charakterystyki p oraz liczby punktów pomiarowych n i wynosi do 50%. Prawie połowa punktów pomiarowych może być obarczona błędami nadmiernymi, a mimo tego estymowane współczynniki przyjmują prawidłowe wartości.

6. Szacowanie wariancji współczynników regresji MNMK

Procedura oszacowania wariancji współczynników regresji MNMK została zaproponowana przez autora w jego rozprawie doktorskiej [1].

W przypadku wykorzystania procedury MNK wariancję współczynników regresji [3, 4] wyznacza się z macierzy kowariancji wektora $\hat{\boldsymbol{\theta}}$

$$\operatorname{Var} \hat{\boldsymbol{\theta}} = (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x})^{-1} \sigma_e^2 , \tag{6}$$

przy czym wariancja poszczególnych współczynników regresji jest równa wartościom elementów diagonalnych macierzy kowariancji (6). Zakłada się przy tym, że losowe błędy obserwacji e_i są niezależne i mają ten sam rozkład o wartości oczekiwanej równej zero i wariancji σ_e^2.

Szacując wariancję współczynników regresji MNMK należy zwrócić uwagę na pewne konsekwencje związane z odpornym charakterem metody. Jak już wspomniano, procedura odporna cechuje się wykorzystaniem do wyznaczenia wartości współczynników regresji jedynie części danych pomiarowych. Punkty, które nie biorą bezpośredniego udziału w obliczeniach nie mają żadnego wpływu na współczynniki regresji, a więc również na ich wariancję. Dotyczy to zarówno wartości zmiennych niezależnych x_{ij}, jak i zależnych y_i.

Badania eksperymentalne przeprowadzone przez autora wykazały, że podobnie jak w przypadku MNK, wariancja współczynników regresji MNMK zależy od wariancji składnika losowego w równaniu regresji oraz wartości punktów, które bezpośrednio służą do wyznaczenia wartości współczynników regresji. Ponadto wariancja współczynników regresji MNMK od liczby danych. Powoduje to konieczność wprowadzenia dodatkowego współczynnika korygującego, o wartości zależnej od liczby punktów pomiarowych.

Pełny program badań eksperymentalnych przeprowadzonych przez autora przedstawiono w [1]. Wykorzystano model (1), przy czym dla p=3 i 4, liczba punktów pomiarowych wynosiła n=10, 20, 40 oraz 80, natomiast dla p=5 i 6, n=20, 40 oraz 80. W przypadku gdy p=5 i 6, n=10 występuje wiele równoważnych rozwiązań kryterium MNMK. Szacowanie wariancji współczynników regresji byłoby wtedy bezcelowe. Przyjęto wartości wszystkich parametrów modelu θ_j=1. Wariancje poszczególnych współczynników regresji MNMK autor proponuje szacować jako wartości elementów diagonalnych macierzy

$$\boldsymbol{V} = \kappa \, (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{w} \, \boldsymbol{x})^{-1} \sigma_e^2 , \tag{7}$$

gdzie κ jest współczynnikiem korygującym, $\boldsymbol{w}$ jest diagonalną macierzą wag o rozmiarze $n \times n$

$$\boldsymbol{w} = \begin{bmatrix} w_{11} & 0 & \cdots & 0 \\ 0 & w_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_{nn} \end{bmatrix} . \tag{8}$$

Macierz $\boldsymbol{w}$ wskazuje punkty, które mają bezpośredni wpływ na wartości współczynników regresji. Zbiór takich punktów tworzy tzw. efektywny plan eksperymentu. Elementy w_{ii} mogą przybierać wartości

$$w_{ii}=\begin{cases}1, \text{gdy punkt pomiarowy bierze udział w tworzeniu wyniku},\\ 0, \text{gdy punkt pomiarowy nie bierze udziału w tworzeniu wyniku}.\end{cases} \tag{9}$$

Liczba elementów k_p, dla których w_{ii}=1, wynika bezpośrednio z kryterium MNMK (5) oraz sposobu obliczania mediany

$$k_p=\begin{cases}\frac{n}{2}+0{,}5 \ \text{ gdy } n \text{ jest nieparzyste},\\ \frac{n}{2}+1 \ \text{ gdy } n \text{ jest parzyste}.\end{cases} \tag{10}$$

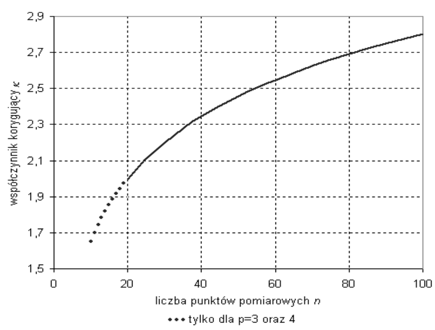
Wartości wag w_{ii} nie są znane, jeśli nieznany jest efektywny plan eksperymentu. Jego wyznaczenie jest możliwe tylko na podstawie analizy wartości błędów resztowych, czyli dopiero po wyznaczeniu wartości współczynników regresji. Efektywny plan eksperymentu tworzy k_p punktów, dla których kwadraty błędów resztowych r_i^2 są najmniejsze. Bez wyznaczenia wartości współczynników regresji znana jest jedynie liczba punktów tworzących efektywny plan eksperymentu k_p.

Na podstawie uzyskanych wyników autor stwierdził, że wartość κ jest niezależna od liczby współczynników p, natomiast rośnie wraz z liczbą punktów pomiarowych n. Wartość współczynnika korygującego κ dobrze aproksymuje funkcja

$$\kappa=0{,}5(1+\ln n), \tag{11}$$

której wykres przedstawiono na rys. 2.

Rezultaty badań autora pozwalają na ocenę dokładności metody szacowania wariancji współczynników regresji MNMK wg (7). Dla większości przypadków różnica pomiędzy wariancją oszacowaną a wyznaczoną eksperymentalnie jest rzędu kilku do kilkunastu procent wyznaczonej eksperymentalnie wariancji współczynników regresji MNMK. W szczególnych przypadkach błąd oszacowania sięga kilkudziesięciu procent.



Rys. 2. Wartość współczynnika korygującego κ w zależności od liczby punktów pomiarowych
Fig. 2. Correction factor κ plot vs. number of measurement points

6.1. Szacowanie a priori wariancji współczynników regresji

Wyznaczenie wariancji współczynników przed wykonaniem eksperymentu pomiarowego jest praktycznie niemożliwe, ze względu na nieznajomość a priori efektywnego planu eksperymentu. Można jednak oszacować graniczne wartości wariancji współczynników oraz średnią wartość wariancji współczynników. Graniczne wartości wariancji współczynników regresji szacuje się za pomocą analizy wszystkich możliwych realizacji planu eksperymentu. Należy wygenerować wszystkie możliwe macierze wag $\boldsymbol{w}$ i oszacować dla nich wartość wariancji poszczególnych współczynników regresji. Analizując oszacowane wartości wariancji dla wszystkich efektywnych planów eksperymentu można wybrać plany, dla których wariancje poszczególnych współczynników są najmniejsze i największe. Daje to możliwość oceny jakości poszczególnych efektywnych planów eksperymentu.

Perspektywa wielokrotnego powtarzania eksperymentu pomiarowego przy wykorzystaniu tego samego planu eksperymentu zachęca do innego sformułowania sposobu oszacowania wariancji współczynników regresji. Zamiast oszacowania wartości granicznych poszczególnych wariancji można wyznaczyć średnią, dla wszystkich możliwych efektywnych planów eksperymentu. Powyższy sposób szacowania nie pozwala wprawdzie na bezpośrednie wskazanie najgorszego bądź najlepszego efektywnego planu eksperymentu, umożliwia jednak oszacowanie przeciętnej wariancji współczynników, czyli ocenę całego planu eksperymentu, a nie tylko pewnych jego realizacji.

6.2. Szacowanie a posteriori wariancji współczynników regresji

Oszacowanie wariancji współczynników regresji wyznaczonych na podstawie danych uzyskanych z przeprowadzonego wcześniej eksperymentu pomiarowego jest podstawą do oceny otrzymanych wyników. W przypadku, gdy wartości wariancji współczynników regresji są zbyt duże można przyjąć, że zastosowany model niedostatecznie opisuje zjawisko, bądź źle dobrany jest plan eksperymentu lub zbyt duże losowe błędy obserwacji dyskwalifikują uzyskane rezultaty.

W przeciwieństwie do szacowania a priori, gdy brak jest informacji na temat wykorzystanego efektywnego planu eksperymentu, istnieje możliwość wyznaczenia efektywnych punktów pomiarowych za pomocą analizy błędów resztowych. Różnica w sposobie szacowania, w porównaniu do metody a priori, polega na zastosowaniu konkretnego efektywnego planu eksperymentu w miejsce szacowania wartości granicznych i średniej planu.

7. Szacowanie wariancji przewidywanej średniej wartości obserwacji

Wariancję przewidywanej średniej wartości zmiennej zależnej (predykcji obserwacji) $\hat{y}_k$ dla wybranej wartości zmiennej niezależnej $\boldsymbol{x}_k$ w przypadku wykorzystania procedury MNK wyznacza się z zależności [3, 4]

$$\mathrm{Var}\ \hat{y}_\mathrm{k}=\boldsymbol{x}_k^\mathrm{T}(\boldsymbol{x}^\mathrm{T}\boldsymbol{x})^{-1}\boldsymbol{x}_k\sigma_e^2\,. \tag{12}$$

Oszacowanie wariancji przewidywanej średniej wartości obserwacji w przypadku procedury MNMK jest problemem bardzo złożonym, przede wszystkim ze względu konsekwencje związane z odpornym charakterem metody. Wykorzystanie do wyznaczenia wartości współczynników regresji jedynie części danych pomiarowych musi zostać uwzględnione w procedurze szacowania wariancji $\hat{y}_k$. Powodzenie koncepcji szacowania wariancji współczynników regresji MNMK skłoniło autora do wykorzystania pewnych jej elementów przy szacowaniu wariancji $\hat{y}_k$. Przy szacowaniu wariancji współczynników regresji MNMK wykorzystuje się elementy diagonalne macierzy $\kappa(\boldsymbol{x}^\mathrm{T}\mathbf{w}\boldsymbol{x})^{-1}$, natomiast do wyznaczania wariancji współczynników regresji MNK – macierzy $(\boldsymbol{x}^\mathrm{T}\boldsymbol{x})^{-1}$. Naturalną konsekwencją tego faktu jest koncepcja, aby potraktować całą macierz $\kappa(\boldsymbol{x}^\mathrm{T}\mathbf{w}\boldsymbol{x})^{-1}$ MNMK (nie tylko jej elementy diagonalne) jako odpowiednik macierzy $(\boldsymbol{x}^\mathrm{T}\boldsymbol{x})^{-1}$ MNK przy szacowaniu wariancji $\hat{y}_k$.

Wariancję przewidywanej średniej wartości zmiennej zależnej (obserwacji) $\hat{y}_k$ dla wybranej wartości zmiennej niezależnej $\boldsymbol{x}_k$ w przypadku MNMK autor proponuje szacować za pomocą zależności

$$\mathrm{Var}\ \hat{y}_\mathrm{k}=\boldsymbol{x}_k^\mathrm{T}\ \kappa\,(\boldsymbol{x}^\mathrm{T}\boldsymbol{w}\,\boldsymbol{x})^{-1}\boldsymbol{x}_k\sigma_e^2\,. \tag{13}$$

Rezultaty przeprowadzonych przez autora badań symulacyjnych, przy wykorzystaniu takiej samej metodologii jak przy badaniu wariancji współczynników regresji MNMK [1], pozwalają na ocenę dokładności proponowanej procedury szacowania wariancji $\hat{y}_k$ MNMK wg (13). Różnica pomiędzy wariancją oszacowaną a wyznaczoną eksperymentalnie dla większości przypadków jest rzędu kilkunastu do kilkudziesięciu procent wyznaczonej eksperymentalnie wariancji obserwacji $\hat{y}_k$ MNMK. Analogiczna różnica pomiędzy wariancją obliczoną a wyznaczoną eksperymentalnie dla MNK jest mniejsza i nie przekracza kilkunastu procent. Stosunkowo niewielkie błędy oszacowania wariancji $\hat{y}_k$ MNMK wskazują, że proponowana przez autora procedura szacowania wg jest poprawna i daje zadowalające rezultaty.

7.1. Szacowanie a priori wariancji przewidywanej wartości obserwacji

Wariancję MNK można obliczyć wprost z (12). Oszacowanie wariancji $\hat{y}_k$ MNMK jest praktycznie niemożliwe, ze względu na nieznajomość a priori efektywnego planu eksperymentu. Analogicznie do koncepcji szacowania a priori wariancji współczynników regresji można oszacować graniczne wartości wariancji $\hat{y}_k$ oraz średnią (dla wszystkich możliwych efektywnych planów eksperymentu) wartość wariancji $\hat{y}_k$ MNMK dla wybranej wartości zmiennej niezależnej $\boldsymbol{x}_k$.

7.2. Szacowanie a posteriori wariancji przewidywanej wartości obserwacji

Dla szacowania a posteriori wariancji $\hat{y}_k$ MNMK bezpośrednio wykorzystuje się zależność (13). W przeciwieństwie do szacowania a priori, gdy brak jest informacji na temat wykorzystanego efektywnego planu eksperymentu, analizy błędów resztowych umożliwia wyznaczenie efektywnych punktów pomiarowych planu. Różnica w sposobie szacowania, w porównaniu do metody a priori, polega na zastosowaniu konkretnego efektywnego planu eksperymentu w miejsce szacowania wartości średniej wariancji $\hat{y}_k$ planu, podobnie jak w przypadku szacowania wariancji współczynników regresji MNMK.

Tab. 1. Przykład oszacowania a posteriori wariancji przewidywanej średniej wartości obserwacji

Tab. 1. Exemplary results for a posteriori estimation of the standard error of mean response

Wybrana wartość $\boldsymbol{x}$	Var $\hat{y}_k$ MNK z oszac.	Var $\hat{y}_k$ MNK z eksper.	Var $\hat{y}_k$ MNMK z oszac.	Var $\hat{y}_k$ MNMK z eksper.
Zestaw danych I				
$\boldsymbol{x}_k^T = [1{,}8\ 1{,}8\ 1{,}8\ 1]$	0,107	0,0990	0,375	0,299
$\boldsymbol{x}_k^T = [2{,}5\ 2{,}5\ 2{,}5\ 1]$	0,425	0,405	0,913	1,04
$\boldsymbol{x}_k^T = [1{,}1\ 1{,}8\ 2{,}5\ 1]$	9,52	9,09	19,2	16,4
$\boldsymbol{x}_k^T = [1{,}1\ 2{,}5\ 1{,}1\ 1]$	9,13	8,58	21,5	16,7
Zestaw danych II				
$\boldsymbol{x}_k^T = [1{,}8\ 1{,}8\ 1{,}8\ 1]$	0,153	0,150	0,505	0,307
$\boldsymbol{x}_k^T = [2{,}5\ 2{,}5\ 2{,}5\ 1]$	0,490	0,486	0,985	0,672
$\boldsymbol{x}_k^T = [1{,}1\ 1{,}8\ 2{,}5\ 1]$	18,9	19,2	32,5	24,4
$\boldsymbol{x}_k^T = [1{,}1\ 2{,}5\ 1{,}1\ 1]$	42,6	44,4	88,1	44,4
Zestaw danych III				
$\boldsymbol{x}_k^T = [1{,}8\ 1{,}8\ 1{,}8\ 1]$	0,134	0,121	0,435	0,265
$\boldsymbol{x}_k^T = [2{,}5\ 2{,}5\ 2{,}5\ 1]$	1,19	1,17	2,61	2,44
$\boldsymbol{x}_k^T = [1{,}1\ 1{,}8\ 2{,}5\ 1]$	0,855	0,913	1,70	1,53
$\boldsymbol{x}_k^T = [1{,}1\ 2{,}5\ 1{,}1\ 1]$	1,19	1,10	3,14	2,14

Analiza wyników badań symulacyjnych przeprowadzonych przez autora wykazała, że różnica pomiędzy wariancją oszacowaną a wyznaczoną eksperymentalnie zazwyczaj waha się od kilkunastu do kilkudziesięciu procent wyznaczonej eksperymentalnie wariancji obserwacji $\hat{y}_k$ MNMK, przy czym w większości przypadków wartość oszacowana jest większa od wyznaczonej eksperymentalnie. Analogiczna różnica pomiędzy wariancją obliczoną a wyznaczoną eksperymentalnie dla MNK jest mniejsza i nie przekracza kilkunastu procent. Przykład oszacowania a posteriori wariancji przewidywanej średniej wartości obserwacji dla 3 zestawów danych wg [9] przedstawiono w tab. 1. Pochodzące z procedury kalibracji dane obejmują trzy plany eksperymentu charakteryzujące się różną jakością, przy czym liczba parametrów modelu p=4, a liczba punktów pomiarowych n=10. Ich dokładną analizę przedstawiają również Piotrowski i Kostyrko [6].

Proponowana przez procedura szacowania a posteriori wariancji $\hat{y}_k$ MNMK jest zdaniem autora poprawna i daje zadowalające rezultaty.

8. Podsumowanie i wnioski

Autor zaproponował procedurę szacowania wariancji przewidywanej średniej wartości obserwacji MNMK. Opisana procedura jest rozwinięciem opracowanej przez niego wcześniej koncepcji szacowania wariancji współczynników regresji MNMK.

Koncepcja szacowania a priori wariancji przewidywanej średniej wartości obserwacji MNMK polega na przeprowadzeniu oszacowania jeszcze przed wykonaniem eksperymentu pomiarowego. Dokładność takiego oszacowania jest niewielka, gdyż możliwe jest jedynie wyznaczenie średniej wartości oszacowania z bardzo dużej liczby możliwych realizacji efektywnego planu eksperymentu MNMK.

W przypadku oszacowania a posteriori wykorzystuje się dane uzyskane z przeprowadzonego wcześniej eksperymentu pomiarowego. Znajomość efektywnego planu eksperymentu umożliwia w takim przypadku uzyskanie znacznie dokładniejszego, w porównaniu z metodą a priori, oszacowania.

Autor przeprowadził badania symulacyjne w celu porównania wyników oszacowania przewidywanej średniej wartości obserwacji MNMK z wynikami otrzymanymi z symulowanego eksperymentu. Podobne obliczenia i badania symulacyjne autor wykonał również dla MNK w celu otrzymania wyników odniesienia. Uzyskane wyniki potwierdzają poprawność proponowanych koncepcji. Różnice pomiędzy wartościami szacowanymi a uzyskanymi z symulowanego eksperymentu są rzędu kilkunastu do kilkudziesięciu procent, co należy uznać za zadowalające. Konieczne jest jednak rozszerzenie zakresu badań symulacyjnych o modele z większą liczbą parametrów, co autor zamierza wykonać w trakcie swoich dalszych prac.

Dalsze prace autora będą zmierzały do opracowania koncepcji szacowania przedziałów ufności dla przewidywanej średniej wartości obserwacji MNMK.

9. Literatura

[1] Buchczik D.: Estymacja wielowymiarowych krzywych kalibracji metodą najmniejszej mediany kwadratów. Rozprawa doktorska, Gliwice, 2005.

[2] Buchczik D.: Wykorzystanie metody najmniejszej mediany kwadratów w kalibracji aparatury chromatograficznej, Joint IMECO TC-1 and XXXIV MKM, vol. II., 231-238, Wrocław, 2002.

[3] Dobosz M.: Wspomagana komputerowo statystyczna analiza wyników badań, EXIT, Warszawa, 2001.

[4] Draper N., Smith H.: Analiza regresji stosowana, PWN, Warszawa, 1973.

[5] Piotrowski J.: Procedury pomiarowe i estymacja sygnałów, Wydawnictwo Politechniki Śląskiej, Gliwice, 1994.

[6] Piotrowski J., Kostyrko K.: Wzorcowanie aparatury pomiarowej, PWN, Warszawa, 2000.

[7] Rousseeuw P., Leroy A.: Robust regression and outlier detection, Wiley, New York, 1987.

[8] Rousseeuw P.: Least median of squares regression, Journal of the American Statistical Association, vol. 79., 871-880, 1984.

[9] Sergent M., Mathieu D., Phan-Tan-Luu R., Drava G.: Correct and incorrect use of multilinear regression, Chemometrics and Inteligent Laboratory Systems, vol. 27., 153-162, 1995.

Artykuł recenzowany