

Grzegorz MZYK

INSTYTUT CYBERNETYKI TECHNICZNEJ
POLITECHNIKI WROCŁAWSKIEJ

Zastosowanie metody zmiennych instrumentalnych do identyfikacji systemów HAMMERSTEINA-WIENERA

Mgr inż. Grzegorz Mzyk



Absolwent Wydziału Elektroniki Politechniki Wrocławskiej (1998). Obecnie doktorant w Zakładzie Sterowania i Optymalizacji Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej. Jest autorem kilku artykułów na temat identyfikacji nieliniowych systemów dynamicznych w obecności zakłóceń losowych.

Streszczenie

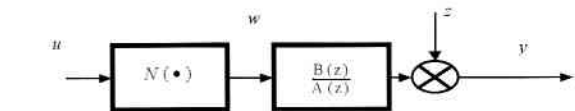
W pracy rozpatruje się zastosowanie metody najmniejszych kwadratów i metody zmiennych instrumentalnych do estymacji parametrów blokowo zorientowanych, nieliniowych systemów dynamicznych. Dla systemu o strukturze Hammersteina-Wienera zaproponowano algorytm identyfikacji metodą zmiennych instrumentalnych i porównano go z algorytmem najmniejszych kwadratów. Udowodniono zgodność zaproponowanego estymatora, nawet dla przypadku występowania skorelowanych zakłóceń. Rozwiązano problem optymalnej generacji zmiennych instrumentalnych.

Abstract

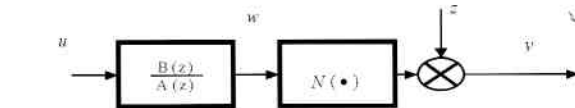
Application of least squares and instrumental variables to recovering parameters of nonlinear complex dynamic block-oriented systems is examined. For a system with the Hammerstein-Wiener structure the instrumental variable algorithm is designed and compared with the least squares algorithm for estimating system parameters. The advantages of the proposed estimator are discussed and in particular its weak consistency, even in the presence of correlated noise is shown. The problem of generating optimal values of instruments is analysed.

1. Wprowadzenie

Metoda zmiennych instrumentalnych (IV, *ang. instrumental variables*) jest jedną z najbardziej popularnych i uniwersalnych metod używanych do estymacji parametrów prostych jednoclementowych liniowych obiektów dynamicznych. Zasadniczym celem pracy jest zastosowanie jej do identyfikacji nieliniowych systemów dynamicznych o złożonej strukturze, a także porównanie z klasyczną metodą najmniejszych kwadratów (LS, *ang. least squares*). Analizy stosowalności metody dokonano na przykładzie systemu o strukturze Hammersteina-Wienera opisanego szczegółowo w pkt. 2. System Hammersteina-Wienera można traktować jako połączenie dwóch typowych struktur kaskadowych: Hammersteina (Rys. 1) i Wienera (Rys. 2). Pomimo możliwości zastosowania w wielu dziedzinach ([1], [5], [6], [13]), systemowi o strukturze Hammersteina-Wienera poświęcono w literaturze stosunkowo mało uwagi. W pkt. 3 przedstawiono algorytm oparty na metodzie najmniejszych kwadratów zaprezentowany w [1] dla przypadku występowania białych zakłóceń, a następnie pokazano jego asymptotyczne obciążenie w przypadku zakłóceń skorelowanych. Dla wyeliminowania tego obciążenia, w pkt. 4 zaproponowano nowy estyma-



Rys. 1.



Rys. 2.

tor oparty na metodzie zmiennych instrumentalnych. Postępowanie jest oparte na pomysłach zaprezentowanych w pracach [11] i [12] dotyczących identyfikacji mniej skomplikowanych systemów dynamicznych. Zaproponowaną metodę porównano następnie z algorytmem najmniejszych kwadratów. Dokonano analizy asymptotycznej błędów estymacji oraz przedyskutowano zalety zaproponowanej metody. W szczególności wykazano zbieżność estymatora w sensie prawdopodobieństwa nawet w przypadku występowania skorelowanych zakłóceń. Rozwiązano problem optymalnej generacji zmiennych instrumentalnych, a także wyliczono rząd szybkości zbieżności zaproponowanego estymatora. Jako ilustrację wykazanych własności załączono przykładowe wyniki porównawczych badań eksperymentalnych.

2. Sformułowanie problemu

2.1. Badany system

Rozpatrywany jest skalarny, dyskretny, nieliniowy system złożony, który można opisać następującym równaniem (por. [1] i Rys. 3):

$$y_k = \sum_{i=1}^p a_i N_2(y_{k-i}) + \sum_{j=0}^n b_j N_1(u_{k-j}) + z_k \quad (1)$$

gdzie

$$N_1(u) = \sum_{i=1}^m c_i f_i(u) \quad i \quad N_2(y) = \sum_{i=1}^q d_i g_i(y).$$

Sygnały y_k , u_k i z_k stanowią odpowiednio wyjście, wejście systemu oraz zakłócenie w chwili k . Zakładamy, że wejście u_k jest ograniczonym białym szumem, zaś losowe zakłócenie z_k ma zerową wartość oczekiwaną, jest ograniczonym ($|z_k| < z_{\max}$) procesem skorelowanym niezależnym od u_k . Założmy także, że postaci liniowo niezależnych funkcji $g_1(\dots)g_q(\dots)$ i $f_1(\dots)f_m(\dots)$ oraz rzędy p , q , n , m są znane a priori. Niech

$$\begin{aligned} \mathbf{a} &= (a_1, \dots, a_p)^T & \mathbf{b} &= (b_0, \dots, b_n)^T \\ \mathbf{c} &= (c_1, \dots, c_m)^T & \mathbf{d} &= (d_1, \dots, d_q)^T \end{aligned} \quad (2)$$

oznaczają nieznanne, prawdziwe wektory parametrów systemu. System (1) jest zatem bardziej ogólny niż spotykany często w zastosowaniach klasyczny system Hammersteina (Rys. 1). Różni się on także od szeroko omawianego w literaturze systemu Wienera-Hammersteina, gdzie dwa liniowe obiekty dynamiczne otaczają jeden nieliniowy element statyczny ([2], [3]).

2.2. Zadanie identyfikacji

Celem identyfikacji jest odkrycie prawdziwych wektorów parametrów \mathbf{a} , \mathbf{b} , \mathbf{c} i \mathbf{d} (danych wzorem (2)), na podstawie zbioru pomiarów (u_k, y_k) ($k = 1..N$) wejść i wyjść całego systemu uzyskanych w eksperymencie. Zauważmy, że opis systemu dany wzorem (1) nie jest jednoznaczny. Dla niezerowych stałych α i β systemy z wektorami parametrów \mathbf{a} , \mathbf{b} , \mathbf{c} , \mathbf{d} oraz $\beta\mathbf{a}$, $\alpha\mathbf{b}$, \mathbf{c}/α , \mathbf{d}/β są nierozróżnialne z punktu widzenia wejść i wyjść (patrz (1)). Dla uzyskania jednoznaczności przyjmijmy następujące założenia ([1]): (a) $\Theta_{ad} = \mathbf{a}\mathbf{d}^T$ i $\Theta_{bc} = \mathbf{b}\mathbf{c}^T$ nie są jednocześnie zerowe; (b) $\|\mathbf{a}\|_2 = 1$ i $\|\mathbf{b}\|_2 = 1$, gdzie $\|\cdot\|_2$ oznacza normę euklidesową wektora; (c) pierwsze niezerowe elementy wektorów \mathbf{a} i \mathbf{b} są dodatnie. Niech

$$\theta = (b_0c_1, \dots, b_0c_m, \dots, b_n c_1, \dots, b_n c_m, \\ a_1 d_1, \dots, a_1 d_q, \dots, a_p d_1, \dots, a_p d_q)^T = \\ = (\theta_1, \dots, \theta_{(n+1)m}, \theta_{(n+1)m+1}, \dots, \theta_{(n+1)m+p})^T$$

będzie wektorem zagregowanych parametrów systemu i niech ϕ_k będzie uogólnionym wektorem wejść (regresorem)

$$\phi_k = (f_1(u_k), \dots, f_m(u_k), \dots, f_1(u_{k-n}), \dots, f_m(u_{k-n}), \\ g_1(y_{k-1}), \dots, g_q(y_{k-1}), \dots, g_1(y_{k-p}), \dots, g_q(y_{k-p}))^T \quad (3)$$

Równanie (1) przyjmuje postać $y_k = \phi_k^T \theta + z_k$, co można zapisać dla $k=1..N$ w postaci macierzowej

$$\mathbf{Y}_N = \Phi_N \theta + \mathbf{Z}_N \\ \text{gdzie} \quad (4) \\ \mathbf{Y}_N = (y_1, \dots, y_N)^T, \quad \mathbf{Z}_N = (z_1, \dots, z_N)^T, \\ \Phi_N = (\phi_1, \dots, \phi_N)^T.$$

3. Metoda najmniejszych kwadratów

Dla późniejszych porównań z metodą zmiennych instrumentalnych zaczniemy od zaprezentowania dwuetapowego algorytmu opartego na metodzie najmniejszych kwadratów i dekompozycji według wartości szczególnych (patrz [1]).

3.1. Algorytm LS

Krok 1: Wyliczyć estymator LS

$$\hat{\theta}^{(LS)} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T \mathbf{Y}_N \quad (5)$$

zagregowanego wektora parametrów θ , a następnie skonstruować oszacowania $\hat{\Theta}_{ad}^{(LS)}$ i $\hat{\Theta}_{bc}^{(LS)}$ macierzy Θ_{ad} i Θ_{bc} (patrz założenie (a)).

Krok 2: Dokonać dekompozycji SVD (według wartości szczególnych) oszacowań $\hat{\Theta}_{ad}^{(LS)}$ i $\hat{\Theta}_{bc}^{(LS)}$

$$\hat{\Theta}_{bc}^{(LS)} = \sum_{i=1}^{\min(n,m)} \sigma_i \bar{\mu}_i \bar{\mu}_i^T, \quad \hat{\Theta}_{ad}^{(LS)} = \sum_{i=1}^{\min(p,q)} \delta_i \bar{\zeta}_i \bar{\zeta}_i^T,$$

a następnie wyliczyć estymatory

$$\hat{\mathbf{b}} = s_\mu \bar{\mu}_1, \quad \hat{\mathbf{c}} = s_\mu \sigma_1 \bar{\mu}_1, \quad \hat{\mathbf{a}} = s_\zeta \bar{\zeta}_1, \quad \hat{\mathbf{d}} = s_\zeta \delta_1 \bar{\zeta}_1 \quad (6)$$

gdzie s_μ i s_ζ oznaczają znaki pierwszych niezerowych elementów odpowiednio w wektorach $\bar{\mu}_1$ i $\bar{\zeta}_1$.

Dekompozycja według wartości szczególnych pozwala na rozszczepienie zagregowanej macierzy parametrów na iloczyn dwóch wektorów. W pracy [1] pokazano, że

$$(\bar{\mu}_1, \sigma_1, \bar{\mu}_1) = \arg \min_{c \in \mathbb{R}^m, b \in \mathbb{R}^n} \|\Theta_{bc}^{(LS)} - \mathbf{b}\mathbf{c}^T\|_F^2$$

3.2. Własności asymptotyczne

Udowodniono następujące twierdzenia:

T1: Dla każdego $N > 0$, gdy kolumny macierzy Φ_N są liniowo niezależne, a zakłócenia $z_k \equiv 0$, wtedy:

$$\hat{\mathbf{b}} = \mathbf{b} \quad \hat{\mathbf{c}} = \mathbf{c} \quad \hat{\mathbf{a}} = \mathbf{a} \quad \hat{\mathbf{d}} = \mathbf{d} \quad (7)$$

T2: Niech zakłócenie z_k będzie białym szumem o zerowej wartości oczekiwanej, skończonej wariancji, niezależnym od wejścia u_k . Niech ponadto ϕ_k będzie procesem wektorowym ustawnie pobudzającym. Wtedy dla $N \rightarrow \infty$:

$$\hat{\mathbf{b}} - p.l \rightarrow \mathbf{b} \quad \hat{\mathbf{c}} - p.l \rightarrow \mathbf{c} \\ \hat{\mathbf{a}} - p.l \rightarrow \mathbf{a} \quad \hat{\mathbf{d}} - p.l \rightarrow \mathbf{d} \quad (8)$$

Mnożąc lewostronnie równanie (4) przez Φ_N^T otrzymujemy:

$$\Phi_N^T \mathbf{Y}_N = \Phi_N^T \Phi_N \theta + \Phi_N^T \mathbf{Z}_N$$

Stąd błąd estymacji wektora θ można zapisać w postaci:

$$\Delta = \hat{\theta}^{(LS)} - \theta = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T \mathbf{Z}_N = \\ = \left(\frac{1}{N} \sum_{k=1}^N \phi_k \phi_k^T \right)^{-1} \left(\frac{1}{N} \sum_{k=1}^N \phi_k z_k \right)$$

Gdy z_k jest białym szumem o zerowej wartości oczekiwanej i skończonej wariancji, wtedy elementy macierzy \mathbf{Z}_N są niezależne od elementów w macierzy (Φ_N) , a zatem $(\mathbf{A} \mathbf{E} 0$ gdy $N \rightarrow \infty$). Jednak gdy z_k jest procesem skorelowanym, tzn. $(z_k z_l - i(0$ dla pewnego $i(0$, wtedy estymator LS (5) wektora (jest asymptotycznie obciążony z powodu zależności pomiędzy zakłóceniami z_k , a wartościami $g_i(y_{k-i})$ ($i=1..q$, $i=1..p$) występującymi w macierzy (Φ_N) . W konsekwencji, estymatory dane równaniami (6) będą także obciążone. Wniosek ten został zilustrowany w poniższym przykładzie.

3.3. Badania eksperymentalne

W eksperymencie symulowano i identyfikowano system o następujących parametrach:

$$\mathbf{b} = \mathbf{c} = \mathbf{a} = \mathbf{d} = (0.5; 0.5; 0.5; 0.5)^T \\ f(x) = \cos tx \quad g(x) = \sin tx \quad t, l = 1..4$$

Wejście u_k generowano zgodnie z rozkładem jednostajnym na przedziale $[-10; 10]$. Zakłócenie z_k generowano wprawdzie jako biały szum o rozkładzie jednostajnym na przedziale $[-0.01; 0.01]$ (ok. 1% amplitudy użytecznego sygnału), lub $[-0.05; 0.05]$ (ok. 5% amplitudy użytecznego sygnału). Eksperyment powtórzono dla z_k przepuszczonego przez filtr kolorujący $F(z) = 1 + 0.5z^{-1}$, gdzie z^{-1} oznacza operator przesunięcia o jeden krok wstecz w dziedzinie czasu. Porównywano zagregowany błąd estymacji wyliczony według następującej formuły:

$$ERR_{LS} = \left\| \begin{bmatrix} \hat{\mathbf{b}}(N), \hat{\mathbf{c}}(N), \hat{\mathbf{a}}(N), \hat{\mathbf{d}}(N) \end{bmatrix} - [\mathbf{b}, \mathbf{c}, \mathbf{a}, \mathbf{d}] \right\|_2$$

Wyniki eksperymentów pokazano na Rys. 4.

4. Metoda zmiennych instrumentalnych

4.1. Założenia

Niech zapis $\text{Plim}_{N \rightarrow \infty} X_N = X$ dla pewnych wektorów losowych X_N , X (RS oznacza, że $X_N \rightarrow X$ według prawdopodobieństwa, gdy $N \rightarrow \infty$). Założmy, że posiadamy macierz zmiennych instrumentalnych (N) , spełniającą dla przypadku skorelowanych zakłóceń z_k następujące warunki:

- (C1): $\dim \Psi_N = \dim \Phi_N$
- (C2): Ψ_k i z_k są ograniczone i wzajemnie niezależne
- (C3): macierz $\Psi_N^T \Phi_N$ jest nieosobliwa z prawdopodobieństwem 1
- (C4): $\text{Plim}(\frac{1}{N} \Psi_N^T \Phi_N)$ istnieje i jest nieosobliwa.

Lemat 1. Warunkiem koniecznym istnienia macierzy zmiennych instrumentalnych (N spełniającej (C4) jest nieosobliwość macierzy $1/N(\text{NT}(N, \text{tzn. } (k \text{ musi być ustawicznie pobudzające (patrz Dodatek A).$

4.2. Algorytm IV

Po lewostronnym pomnożeniu równania (4) przez macierz (NT otrzymujemy:

$$\Psi_N^T Y_N = \Psi_N^T \Phi_N \theta + \Psi_N^T Z_N$$

Biorąc pod uwagę powyższe warunki (C2) i (C3) postulujemy zastąpienie estymatora LS danego wzorem (5) wyliczanego w kroku 1 przez następujący estymator IV:

$$\bar{\theta}^{(IV)} = (\Psi_N^T \Phi_N)^{-1} \Psi_N^T Y_N \tag{9}$$

gdzie $\Psi_N = (\Psi_1, \dots, \Psi_N)^T$. W drugim kroku postępujemy analogicznie, dokonując rozkładu SVD uzyskanych na podstawie oszacowań i macierzy (ad i (bc.

4.3. Własności asymptotyczne

Błąd estymacji algorytmem (9) ma postać:

$$\Delta^{(IV)} = \bar{\theta}^{(IV)} - \theta = (\Psi_N^T \Phi_N)^{-1} \Psi_N^T Z_N \tag{10}$$

Twierdzenie 1. Przy spełnionych warunkach (C1)((C4), estymator (9) jest zgodny (słabo), niezależnie od budowy korelacyjnej zakłóceń, tzn.

$$P \lim (\Delta^{(IV)}) = 0$$

(patrz Dodatek B).

Twierdzenie 2. Błąd estymacji ((IV) posiada szybkość zbieżności rzędu $\frac{1}{\sqrt{N}}$ w sensie prawdopodobieństwa, dla każdego wyboru zmiennych pomocniczych spełniających warunki (C1)((C4) (patrz Dodatek C).

4.4. Optymalizacja metody IV

Przyjmijmy oznaczenie

$$\bar{\theta}_k = (f_1(u_k), \dots, f_m(u_k), \dots, f_1(u_{k-n}), \dots, f_m(u_{k-n}), \tag{11}$$

$$R_1(u_{k-1}), \dots, R_q(u_{k-1}), \dots, R_1(u_{k-p}), \dots, R_q(u_{k-p}))^T$$

gdzie

$$R_j(u) = E\{g_j(y_k) | u_k = u\} \quad j = 1 \dots q$$

są funkcjami regresji. Oznaczmy

$$Z_N^* = \frac{1}{\sqrt{N}} Z_N$$

i niech

$$Q(\Psi_N) = \max_{\|Z_N^*\| \leq 1} \|\Delta_N(\Psi_N)\|^2 \tag{12}$$

będzie wskaźnikiem jakości zmiennych instrumentalnych, gdzie jest zdefiniowane wzorem (10), zaś $\| \cdot \|$ oznacza normę euklidesową wektora.

Twierdzenie 3. Przy spełnionych warunkach (C1)((C4), (patrz (12)) osiąga swoje minimum dla

$$\Psi_N = \bar{\Phi}_N = (\bar{\theta}_1, \dots, \bar{\theta}_N)^T$$

(patrz Dodatek D).

5. Badania eksperymentalne

Dla porównania, działanie algorytmów (5) i (9) zbadano praktycznie na przykładowym systemie opisanym w pkt. 3.3.. Zmienne instrumentalne generowano następującą metodą:

$$\Psi_k = (f_1(u_k), \dots, f_m(u_k), \dots, f_1(u_{k-n}), \dots, f_m(u_{k-n}), g_1(\bar{x}_{k-1}), \dots, g_q(\bar{x}_{k-1}), \dots, g_1(\bar{x}_{k-p}), \dots, g_q(\bar{x}_{k-p}))^T$$

gdzie

$$\bar{x}_k = \sum_{i=1}^n \bar{a}_i \left\{ \sum_{j=1}^q \bar{a}_j g_j(\bar{x}_{k-1}) \right\} + \sum_{j=0}^n \bar{b}_j \left\{ \sum_{i=1}^m \bar{a}_i f_i(u_{k-1}) \right\}$$

natomiast $\bar{a}_i, \bar{b}_j, \bar{c}_i, \bar{d}_j$ oznaczają elementy wektorów $\bar{a}, \bar{b}, \bar{c}, \bar{d}$ wyliczonych przy użyciu metody LS opisaną w pkt. 3. Zagregowane błędy estymacji metodą LS (ERR_LS) i metodą IV (ERR_IV) wyliczono dla 102(N(105 pomiarów. Wyniki pokazano na Rys. 5. Jak widzimy, dla dużych N - asymptotyczne zachowanie się estymatora metodą zmiennych instrumentalnych jest lepsze niż estymatora najmniejszych kwadratów. Zbieżność jest raczej powolna, zatem fundamentalnym problemem jest wybór zmiennych instrumentalnych w optymalny sposób.

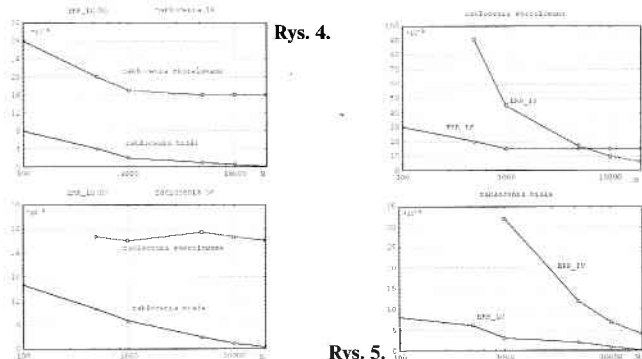
Na podstawie wniosków zawartych w pracy [8] zrodził się pomysł użycia nieparametrycznych metod identyfikacji ([4], [9]) do generowania przybliżonych wartości optymalnych zmiennych instrumentalnych. Eksperyment powtórzono dla instrumentów określonych regułą

$$\Psi_k = (f_1(u_k), \dots, f_m(u_k), \dots, f_1(u_{k-n}), \dots, f_m(u_{k-n}), \hat{R}_1(u_{k-1}), \dots, \hat{R}_q(u_{k-1}), \dots, \hat{R}_1(u_{k-p}), \dots, \hat{R}_q(u_{k-p}))^T$$

gdzie wartości $R_j(u)$ są nieparametrycznymi, jądrowymi estymatorami $R_j(u)$. Otwarty problemem pozostaje analiza, czy tak skonstruowany wektor zmiennych instrumentalnych spełnia warunki (C1)((C4). Wyniki eksperymentu przedstawiono na Rys. 6.

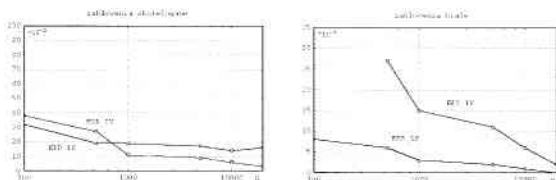


Rys. 3.



Rys. 4.

Rys. 5.



Rys. 6.

6. Podsumowanie

Złożoność obliczeniowa zaproponowanego algorytmu jest porównywalna ze złożonością algorytmu najmniejszych kwadratów. Jedynym utrudnieniem jest konieczność wygenerowania dodatkowej macierzy (zmiennych instrumentalnych). Drugi etap procedury polega na standardowym zabiegu rozkładu dwóch macierzy o ustalonych wymiarach według wartości szczególnych. Metoda zmiennych instrumentalnych może być z powodzeniem stosowana do identyfikacji nieliniowych systemów dynamicznych o złożonej strukturze (Hammersteina-Wienera) w obecności skorelowanych zakłóceń. Jest to główna zaleta metody. Należy zwrócić uwagę, że skorelowanie zakłóceń może być związane z obecnością strukturalnego sprzężenia zwrotnego w złożonym systemie. Zainteresowanych odsyłamy do pracy [7], gdzie metodę zmiennych instrumentalnych z powodzeniem zastosowano do identyfikacji złożonych liniowych systemów statycznych ze sprzężeniem zwrotnym.

7. Dodatki

7.1. Dodatek A. Szkic dowodu Lematu 1

Niech A, B e R $a \times b$ będą macierzami o skończonych elementach i niech $\det(A \ T \ A) = 0$. Wtedy istnieje niezerowy wektor e e R B taki, że $Ae = 0$, zatem $BT \ Ae = 0$ i $\det(BTA) = 0$.

Lemat 1 otrzymujemy podstawiając

$$A := \frac{1}{\sqrt{N}} \Phi_N \quad B := \frac{1}{\sqrt{N}} \Psi_N$$

7.2. Dodatek B. Szkic dowodu Twierdzenia 1

Na podstawie twierdzenia Słuckiego [10], można zapisać

$$P \lim(\Delta^{(N)}) = P \lim\left(\left(\Psi_N^T \Phi_N\right)^{-1}\right) \cdot P \lim\left(\Psi_N^T Z_N\right)$$

$$\text{Oznaczmy } G_N = \Psi_N^T Z_N, \quad G_N^i = \sum_{k=1}^N \psi_{i,k} z_k$$

Łatwo pokazać, że

$$E\left[\frac{1}{N} G_N^i\right] = 0$$

$$\text{var}\left[\frac{1}{N} G_N^i\right] \leq \frac{2\psi_{\max}^2}{N} \sum_{k=1}^N |\omega(k)|$$

gdzie $w(k) = E z_i z_i + k$ jest funkcją kowariancji zakłóceń.

Zatem gdy $k \omega(k) \rightarrow 0$ otrzymujemy

$$\lim_{N \rightarrow \infty} \text{var}\left[\frac{1}{N} G_N^i\right] = 0.$$

Z nierówności Czebyszewa:

$$P \lim\left(\frac{1}{N} G_N^i\right) = P \lim\left(\frac{1}{N} \Psi_N^T Z_N\right) = 0$$

$$\text{oraz } P \lim(\Delta^{(N)}) = 0.$$

7.3. Dodatek C. Szkic dowodu Twierdzenia 2

Oznaczmy

$$A_N = \frac{1}{N} \Psi_N^T \Phi_N = \frac{1}{N} \sum_{k=1}^N \psi_k \phi_k^T$$

$$B_N = \frac{1}{\sqrt{N}} \Psi_N^T Z_N = \frac{1}{\sqrt{N}} \sum_{k=1}^N \psi_k z_k$$

$$C_N = \Delta^{(N)} \quad D_N = A_N^{-1} B_N.$$

Wtedy $C_N = \frac{1}{\sqrt{N}} D_N$. Można pokazać, że

$$|A_N^{i,j}| < \infty, \quad E B_N^i = \frac{1}{\sqrt{N}} \sum_{k=1}^N E \psi_{i,k} E z_k = 0$$

$$\text{var } B_N^i \leq 2\psi_{\max}^2 \sum_{k=1}^N |\omega(k)| < \infty$$

$$E D_N^{i,j} < \infty \quad \text{var}(D_N^{i,j}) < \infty$$

Zatem dla każdego deterministycznego ciągu liczbowego X_n takiego, że $\lim_{N \rightarrow \infty} X_n = 0$ zachodzi

$$X_n D_N^{i,j} = \frac{C_N^{i,j}}{1/\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0$$

według prawdopodobieństwa.

7.3. Dodatek D. Szkic dowodu Twierdzenia 3

Niech $\Gamma_N = \left(\frac{1}{N} \Psi_N^T \Phi_N\right)^{-1} \frac{1}{\sqrt{N}} \Psi_N^T$ wtedy

$$\Delta_N = z_{\max} \Gamma_N Z_N^*, \text{ więc}$$

$$\|\Delta_N(\Psi_N)\|^2 = \Delta_N^T(\Psi_N) \Delta_N(\Psi_N) = Z_N^{*T} \Gamma_N^T \Gamma_N Z_N$$

oraz

$$\begin{aligned} Q(\Psi_N) &= \max_{\|z_N^*\| \leq 1} (\Delta_N^T(\Psi_N) \Delta_N(\Psi_N)) = \\ &= \|\Gamma_N\|^2 = \lambda_{\max}(\Gamma_N^T \Gamma_N) \end{aligned}$$

gdzie $\lambda_{\max}(\cdot)$ oznacza największą wartość własną macierzy. Z własności

$$\lambda_{\max}(\Gamma_N^T \Gamma_N) = \lambda_{\max}(\Gamma_N \Gamma_N^T)$$

otrzymujemy

$$\begin{aligned} \max_{\|z_N^*\| \leq 1} (\Delta_N^T(\Psi_N) \Delta_N(\Psi_N)) &= \\ &= \max_{\|\xi\| \leq 1} \left\langle \xi, \left(\frac{1}{N} \Psi_N^T \Phi_N\right)^{-1} \left(\frac{1}{N} \Psi_N^T \Psi_N\right) \left(\frac{1}{N} \Phi_N^T \Psi_N\right)^{-1} \xi \right\rangle \end{aligned}$$

$$\text{Można pokazać, że } \frac{1}{N} \Psi_N^T (\Phi_N - \bar{\Phi}_N) \rightarrow 0$$

z prawdopodobieństwem 1, gdy $N \rightarrow \infty$. Stąd dla $N \rightarrow \infty$

$$\begin{aligned} \max_{\|z_N^*\| \leq 1} (\Delta_N^T(\Psi_N) \Delta_N(\Psi_N)) &= \\ &= \max_{\|\xi\| \leq 1} \left\langle \xi, \left(\frac{1}{N} \Psi_N^T \bar{\Phi}_N\right)^{-1} \left(\frac{1}{N} \Psi_N^T \Psi_N\right) \left(\frac{1}{N} \bar{\Phi}_N^T \Psi_N\right)^{-1} \xi \right\rangle \end{aligned}$$

z prawdopodobieństwem 1.

Lemat 2. ([12]) Niech M_1 i M_2 będą macierzami o tych samych wymiarach. Jeżeli istnieją $(M_1^T M_1)^{-1}$, $(M_1^T M_1)^{-1}$ i $(M_1^T M_1)^{-1}$ wtedy macierz $(M_1^T M_1)^{-1}$, $(M_1^T M_1)^{-1}$ i $(M_1^T M_1)^{-1}$ jest nieujemnie określona, tzn. dla każdego wektora η zachodzi $M_1^T M_1 \eta \eta^T \geq 0$.

Na podstawie Lematu 2 dla $M_1 = \frac{1}{\sqrt{N}} \bar{\Phi}_N$,

$$M_2 = \frac{1}{\sqrt{N}} \Psi_N \quad \text{i} \quad \eta = Z_N^{*T}$$

otrzymujemy (z prawdopodobieństwem 1):

$$Z_N^{*T} \Gamma_N \Gamma_N^T Z_N^* \geq Z_N^{*T} (\bar{\Phi}_N^T \bar{\Phi}_N)^{-1} Z_N^*$$

$$Q(\Psi_N) = \max_{\|Z_N^*\| \leq 1} (Z_N^{*T} \Gamma_N \Gamma_N^T Z_N^*) \geq$$

$$\geq \max_{\|Z_N^*\| \leq 1} (Z_N^{*T} (\bar{\Phi}_N^T \bar{\Phi}_N)^{-1} Z_N^*)$$

$$\text{czyli } Q(\Psi_N) = \max_{\|Z_N^*\| \leq 1} \|\Delta_N(\Psi_N)\|^2$$

osiąga swoje dolne ograniczenie dla wyboru $F_n = F_n$.

Literatura

[1] E.W. Bai: An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica* 1998, Vol. 34, No. 3, pp. 333-338.

- [2] S.A. Billings: Identification of nonlinear systems - a survey. *IEEE Proceedings* 1980, Vol. 127, No. 6, pp. 272-285.
- [3] S.A. Billings, S.Y. Fakhouri: Identification of systems containing linear dynamic and static nonlinear elements. *Automatica* 1982, Vol. 18, No. 1, pp. 15-26.
- [4] W. Greblicki, M. Pawlak: Cascade Non-Linear System Identification by a Non-Parametric Method. *International Journal of System Science* 1994, Vol. 25, No. 1, pp. 129-153.
- [5] R. Haber, P. Zeirfuss: Identification of an electrically heated heat exchanger by several nonlinear models using different structures and parameter estimation methods. Report 1988, Institute of Machine and Process Automation, Technical University of Vienna, Austria.
- [6] P.Z. Marmarelis, U.I. Naha: Identification of biological systems. *IEEE Trans. Biomed. Eng.* 1974, BME-21, pp. 88-101.
- [7] G. Mzyk: Zastosowanie metody zmiennych pomocniczych do identyfikacji systemów o złożonej strukturze. XIII Krajowa Konferencja Automatyki, Opole, 1999.
- [8] G. Mzyk: Hammerstein system identification by a semi-parametric method. *International Conference on Signals and Electronic Systems*, Ustroń, 2000.
- [9] M. Pawlak, Z. Hasiewicz, Nonlinear System Identification by the Haar Multiresolution Analysis. *IEEE Transactions on Circuits and Systems* 1998, Vol. 45, No. 9, pp. 945-961.
- [10] T. Söderström, P. Stoica: *System Identification*, Englewood Cliffs, NJ: Prentice Hall 1989.
- [11] P. Stoica, T. Söderström: Instrumental variable methods for identification of Hammerstein systems. *International Journal of Control* 1982, Vol. 35, No. 3, pp. 459-476.
- [12] K.Y. Wong, E. Polak: Identification of linear discrete time systems using the instrumental variable method, *IEEE Transactions on Automatic Control* 1967, Vol. AC-12, No. 6, pp. 707-717.
- [13] Y.K. Zhang, E.W. Bai: Simulation of spring discharge from a limestone aquifer in Iowa. *Hydrogeol. J.* 1996, No. 4, pp. 41-54.