Instytut Informatyki, Automatyki i Robotyki

Seria: PREPRINTY nr 5/2008

Parametryczno-nieparametryczna identyfikacja systemów Wienera

(rozprawa doktorska)

Paweł Wachel

Promotor: prof. dr hab. inż. Zygmunt Hasiewicz

Słowa kluczowe:
- identyfikacja systemów złożonych,
- systemy nieliniowe,
- systemy Wienera,
- identyfikacja parametryczna,
- identyfikacja nieparametryczna.

Rozprawa powstała częściowo w ramach współpracy z dr hab. inż. Mirosławem Pawlakiem, profesorem Uniwersytetu Manitoby, Winnipeg, Kanada.

Wrocław, 26.03.2008

Spis treści

Spis rysunków

Wykaz ważniejszych oznaczeń

Oznaczenia podano w kolejności alfabetycznej zastosowanych symboli.

$\|\cdot\|$	– norma euklidesowa
$\lvert\cdot\rvert$	– wartość bezwzględna
$\lfloor\cdot\rfloor$	– funkcja „podłoga”
$(\cdot)^{[i]}$	– i–ta realizacja estymatora
$(\cdot)^{T}$	– transpozycja
$\hat{(\cdot)}^{AD}$	– estymator zbudowany w oparciu o metodę uśrednionej pochodnej
$\hat{(\cdot)}^{Cr}$	– estymator zbudowany w oparciu o metodę korelacyjną
$\hat{(\cdot)}^{fAD}$	– estymator zbudowany w oparciu o metodę uśrednionej pochodnej oraz wykorzystujący jądrową estymację gęstości sygnału $\{X_n\}$
$\hat{(\cdot)}^{LS}$	– estymator zbudowany w oparciu o metodę najmniejszych kwadratów
$\hat{(\cdot)}^{RC}$	– estymator zbudowany w oparciu o metodę korelacji rangowej
$\mathbf{1}[A]$	– indykator zdarzenia A
$\mathbf{1}_B(x)$	– funkcja wskaźnikowa przyjmująca wartość 1, gdy $x \in B$ i 0, gdy $x \notin B$
$\forall$	– kwantyfikator ogólny („dla każdego...”)
$\exists$	– kwantyfikator szczegółowy („istnieje...”)
α	– stała multiplikatywna w metodzie korelacyjnej
$bias\{\cdot\}$	– obciążenie estymatora
$corr\{\cdot,\cdot\}$	– korelacja
$Cov\{\cdot,\cdot\}$	– kowariancja
$D_f(\cdot)$	– gradient funkcji $f_{\underline{X}}(\cdot)$
$D_G(\cdot)$	– gradient funkcji $G(\cdot)$
$E\{\cdot\}$	– operator wartości oczekiwanej
$Err(\cdot;\cdot)$	– empiryczny błąd estymacji odp. impulsowej podsystemu dynamicznego
$f_V(\cdot)$	– funkcja gęstości prawdopodobieństwa sygnału interakcyjnego $\{V_n\}$
$f_X(\cdot)$	– funkcja gęstości prawdopodobieństwa sygnału wejściowego $\{X_n\}$
$f_{\underline{X}}(\cdot)$	– funkcja gęstości prawdopodobieństwa sygnału wektorowego $\{\underline{X}_n\}$
$\Phi(\cdot)$	– dystrybuanta standardowego rozkładu normalnego $\mathcal{N}(0,1)$
$g(\cdot)$	– nieliniowa charakterystyka podsystemu statycznego
$g(\cdot;\underline{l})$	– funkcja regresji zależna od wektorowego parametru $\underline{l} \in \Lambda$
$G(\cdot)$	– nieliniowość w reprezentacji systemu Wienera za pomocą statycznego systemu typu MISO
h	– parametr wygładzania w estymatorze charakterystyki nieliniowej
h_0	– parametr wygładzania w estymatorze pilotowym
h_1	– parametr wygładzania w estymatorze gęstości $f_X(\cdot)$

h_2 – parametr wygładzania w estymatorze pochodnej gęstości $f_X(\cdot)$
I – zbiór indeksów, $I = \{p+1, p+2, \ldots, N\}$
I_1 – zbiór indeksów, $I_1 = \{p+1, \ldots, p+N_1\}$
I_2 – zbiór indeksów, $I_2 = \{2p+1+N_1, \ldots, N\}$
κ – stała multiplikatywna w metodzie uśrednionej pochodnej
$K(\cdot)$ – funkcja jądra w estymatorze charakterystyki nieliniowej
$K_0(\cdot)$ – funkcja jądra w estymatorze pilotowym
$K_1(\cdot)$ – funkcja jądra w estymatorze gęstości prawdopodobieństwa sygnału wejściowego $\{X_n\}$
$K_2(\cdot)$ – funkcja jądra w estymatorze pochodnej gęstości prawdopodobieństwa sygnału wejściowego $\{X_n\}$
$\{\lambda_i\}_{i=0}^{p}$ – odpowiedź impulsowa podsystemu dynamicznego
$\underline{\lambda}$ – wektor utworzony z elementów odpowiedzi impulsowej podsystemu dynamicznego, $\underline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_p)^T$
Λ – zbiór możliwych odpowiedzi impulsowych podsystemu dynamicznego
$MISE(\cdot;\cdot)$ – empiryczny błąd estymacji charakterystyki nieliniowej
μ_V – wartość oczekiwana sygnału interakcyjnego $\{V_n\}$
μ_X – wartość oczekiwana sygnału wejściowego $\{X_n\}$
μ_Y – wartość oczekiwana sygnału wyjściowego $\{Y_n\}$
N – liczba dostępnych obserwacji wejścia i wyjścia systemu
N_I – liczba elementów zbioru I
N_{I_1} – liczba elementów zbioru I_1
N_{I_2} – liczba elementów zbioru I_2
$\mathcal{N}(0,1)$ – standardowy rozkład normalny
NSR – współczynnik „noise–to–signal ratio”
p – długość pamięci systemu
$\hat{Q}_N(\cdot)$ – funkcja kryterialna wykorzystywana w konstrukcji estymatora $\hat{\underline{\lambda}}_N^{LS}$
$\hat{S}_N(\cdot)$ – funkcja kryterialna wykorzystywana w konstrukcji estymatora $\hat{\underline{\lambda}}_N^{RC}$
σ_V^2 – wariancja sygnału interakcyjnego $\{V_n\}$
σ_X^2 – wariancja sygnału wejściowego $\{X_n\}$
σ_Z^2 – wariancja sygnału zakłócającego $\{Z_n\}$
T_1 – zbiór par obserwacji wejścia i wyjścia systemu, $T_1 = \{(X_n, Y_n)\}_{n \in I_1}$
T_2 – zbiór par obserwacji wejścia i wyjścia systemu, $T_2 = \{(X_n, Y_n)\}_{n \in I_2}$
T_G – zbiór par obserwacji wejścia i wyjścia systemu, $T_G = \{(\underline{X}_n, Y_n)\}_{n \in I}$
$Var\{\cdot\}$ – wariancja
$\{V_n\}$ – sygnał interakcyjny (ciąg zmiennych losowych)
$\{X_n\}$ – sygnał wejściowy (ciąg zmiennych losowych)
$\underline{X}_n$ – $(p+1)$–elementowy wektor utworzony z kolejnych obserwacji wejścia systemu, $\underline{X}_n = (X_n, X_{n-1}, \ldots, X_{n-p})^T$
$\{\underline{X}_n\}$ – wektorowy sygnał wejściowy w reprezentacji systemu Wienera za pomocą statycznego systemu typu MISO
$\{(X_n, Y_n)\}_{n=1}^{N}$ – zbiór par obserwacji wejścia i wyjścia systemu
$\{Y_n\}$ – sygnał wyjściowy (ciąg zmiennych losowych)
$\{Z_n\}$ – addytywny sygnał zakłócający (ciąg zmiennych losowych)
Ω – nośnik funkcji $f_{\underline{X}}(\cdot)$

Wstęp

W pracy przedstawiono nową klasę parametryczno–nieparametrycznych algorytmów identyfikacji systemów Wienera z czasem dyskretnym, dla których przyjąć można założenie, iż pamięć części dynamicznej systemu jest skończona. Zaproponowane metody umożliwiają pełną identyfikację rozważanych systemów, tj. pozwalają na oszacowanie opisów zarówno podsystemów liniowych, jak i nieliniowych. Przeprowadzono analizę teoretyczną własności asymptotycznych skonstruowanych algorytmów oraz wykonano symulacje komputerowe, ilustrujące ich zachowanie dla małej i umiarkowanej liczby obserwacji.

W rozdziale 1. prezentowana jest ogólna koncepcja systemów o strukturze blokowej na tle innych podejść do opisu systemów nieliniowych spotykanych w literaturze. Główną uwagę poświęcono systemom Wienera – ich konstrukcji, zastosowaniom w dziedzinie aproksymacji systemów nieliniowych oraz implementacjom praktycznym. Rozdział 2. wprowadza do zagadnienia parametryczno–nieparametrycznej identyfikacji systemów Wienera. Omówiono w nim najpierw ideę parametrycznej oraz nieparametrycznej identyfikacji systemów, ze szczególnym uwzględnieniem algorytmów opisanych w literaturze i poświęconych rozważanej klasie obiektów. Na tym tle wprowadzona została koncepcja mieszanej, parametryczno–nieparametrycznej, identyfikacji systemów Wienera oraz omówiony został wspólny schemat konstrukcji prezentowanych metod. Rozdział ten zawiera również podstawowe założenia wykorzystywane w pracy, a także tezę i cele pracy. W rozdziałach 3., 4. i 5. przedstawiono kolejno trzy różne metody identyfikacji. Szczegółowo omówiono konstrukcje odpowiednich estymatorów podsystemów liniowych i nieliniowych oraz podano warunki gwarantujące zbieżność algorytmów. Oprócz analizy zgodności estymatorów, w rozdziałach 3. i 5. przeprowadzono analizę rzędów szybkości zbieżności odpowiednich oszacowań. W uzupełnieniu analizy asymptotycznej, dla każdej z metod wykonano badania symulacyjne ilustrujące ich efektywności dla skończonej liczby obserwacji. Rozdział 3. zawiera opis korelacyjnej metody identyfikacji. W rozdziale 4. przedstawiono algorytm oparty na metodzie najmniejszych kwadratów opracowany we współpracy z Panem Profesorem Mirosławem Pawlakiem z University of Manitoba. W rozdziale tym krótko omówiono również metodę identyfikacji systemów Wienera wykorzystującą ideę estymacji korelacji rangowej. Algorytm identyfikacji, wykorzystujący koncepcję estymacji uśrednionej pochodnej, konstruowany jest z kolei w rozdziale 5. Rozdział 6. zawiera zbiorczą analizę uzyskanych w rozprawie rezultatów teoretycznych oraz badań symulacyjnych algorytmów. W rozdziale 7. przedstawiono otwarte problemy naukowe, a także określono możliwe kierunki dalszych badań.

Autor pragnie podziękować Panu Profesorowi Zygmuntowi Hasiewiczowi, promotorowi pracy oraz Panu Profesorowi Włodzimierzowi Greblickiemu, kierownikowi Zakładu Sterowania i Optymalizacji w Instytucie Informatyki, Automatyki i Robotyki Politechniki Wrocławskiej, za liczne dyskusje oraz cenne uwagi przekazane w trakcie studiów

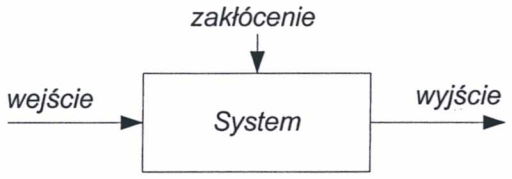
doktoranckich. Szczególne podziękowania składam również Panu Profesorowi Mirosławowi Pawlakowi z University of Manitoba za uwagi merytoryczne dotyczące materiału zawartego w rozdziale 4. oraz umożliwienie mi odbycia zagranicznego stażu naukowego w Department of Electrical and Computer Engineering, University of Manitoba. Dziękuję także wszystkim pracownikom Zakładu Sterowania i Optymalizacji, Instytutu Informatyki, Automatyki i Robotyki Politechniki Wrocławskiej. Na koniec dziękuję serdecznie moim Rodzicom oraz Kasi, bez których ta praca nie mogłaby powstać.

Rozdział 1

Wprowadzenie – systemy Wienera i ich zastosowania

1.1 Koncepcja systemów blokowo-zorientowanych

Zadanie identyfikacji systemów, jako jedno z podstawowych zagadnień automatyki, polega na wyznaczeniu matematycznego modelu zjawiska fizycznego lub procesu technologicznego na podstawie dostępnych danych pomiarowych oraz posiadanej informacji wstępnej o systemie. Terminem *system* określa się przy tym obiekt posiadający wejście i wyjście, w którym zazwyczaj występuje niedostępny dla pomiarów sygnał zakłócający, rys. 1.1.



Rysunek 1.1: System jako obiekt identyfikacji

Wykorzystywane w automatyce systemy, w zależności od stopnia skomplikowania ich struktury, określa się mianem obiektów prostych lub złożonych. Przykładowo, obiektem prostym jest dyskretny statyczny system nieliniowy opisany równaniem

$$y_n = g\left(x_n, z_n\right), \tag{1.1}$$

gdzie $\{x_n\}$ i $\{y_n\}$ są odpowiednio wejściem i wyjściem systemu, a $\{z_n\}$ jest sygnałem zakłócającym. Nieliniowa charakterystyka systemu jest reprezentowana przez funkcję $g\left(\cdot,\cdot\right)$. Innym obiektem prostym jest dyskretny liniowy system dynamiczny opisany wzorem

$$y_n = \phi_n^T \theta + z_n, \tag{1.2}$$

gdzie θ jest wektorem parametrów

$$\theta = \left(a_1, a_2, \ldots, a_q, b_1, b_2, \ldots, b_r\right)^T,$$

a ϕ_n wektorem regresorów utworzonym z pomiarów wejścia i wyjścia

$$\phi_n = \left(-y_{n-1}, \ldots - y_{n-q}, x_{n-1}, \ldots x_{n-r}\right)^T.$$

Teoria identyfikacji obiektów prostych np. typu (1.1) lub (1.2) jest dobrze ugruntowana zarówno w przypadku deterministycznych, jak i losowych sygnałów wejściowych, zob. np. Hasiewicz [48,49], Söderström i Stoica [118], Ljung [76], Mańczak i Nahorski [78], Goodwin i Payne [24], Nelles [83]. Niestety, wiele spotykanych w praktyce procesów i zjawisk nie daje się w zadowalający sposób reprezentować za ich pomocą, chociażby dlatego, że część z nich cechuje zarówno występowanie pamięci, jak i zależności nieliniowych. Do opisu takich obiektów wykorzystuje się zazwyczaj systemy złożone. Przez złożoność systemu rozumie się fakt, że system składa się z wielu elementów oraz że obiekty (podsystemy) wchodzące w skład całego systemu są ze sobą połączone i wzajemnie od siebie zależą, a ich rozdzielenie jest niemożliwe, niebezpieczne, lub zbyt kosztowne, Mzyk [80]. Z tego względu zagadnienie identyfikacji systemów złożonych jest na ogół znacznie trudniejsze od identyfikacji systemów prostych i rodzi dodatkowe problemy, które nie mają swoich odpowiedników w klasycznej teorii, Hasiewicz [48]. Utrudnienia te wynikają między innymi z faktu, że dla dynamicznych systemów złożonych nie jest znana uniwersalna reprezentacja umożliwiająca wyznaczenie matematycznej zależności między wejściem a wyjściem dowolnego systemu złożonego bez dodatkowej wiedzy apriorycznej o strukturze systemu. Pewne rezultaty można jednak otrzymać ograniczając rozważania do klasy systemów stacjonarnych (tj. takich, których parametry nie zmieniają się w czasie), przyczynowych (tzn. takich, w których wyjście w dowolnej chwili t_0 nie zależy od wejścia i wyjścia w chwilach późniejszych) oraz posiadających nieliniowości spełniające określone warunki ciągłości, Boyd i Chua [9]. Jednym ze sposobów opisu tego typu systemów jest wykorzystanie tzw. szeregów Volterry, u podstaw których leży koncepcja rozwinięcia funkcji nieliniowej w szereg Taylora oraz reprezentacja ciągłych systemów liniowych za pomocą całki splotowej postaci

$$y(t) = \int_0^t \vartheta(\tau)\, x(t-\tau)\, d\tau, \tag{1.3}$$

gdzie $\{x(t)\}$ i $\{y(t)\}$ są odpowiednio wejściem i wyjściem systemu, a $\{\vartheta(t)\}$ jest jego odpowiedzią impulsową, Greblicki [34]. W przypadku systemów złożonych z czasem ciągłym i spełniających wymienione wyżej warunki, szereg Volterry (będący uogólnieniem całki (1.3)) przyjmuje postać

$$y(t) = \vartheta_0 + \sum_{i=1}^{\infty} \int \cdots \int \vartheta_i(\tau_1, \cdots \tau_i)\, x(t-\tau_1) \cdots x(t-\tau_i)\, d\tau_1 \cdots d\tau_i, \tag{1.4}$$

gdzie $\vartheta_i(\,)$ są tzw. jądrami Volterry. Analogicznie, dla systemów z czasem dyskretnym odpowiedni szereg Volterry, jako uogólnienie splotu dyskretnego, wyraża się wzorem

$$y_n = \varrho_0 + \sum_{i=1}^{\infty} \sum_{j_1,\ldots,j_i \geq 0} \varrho_i(j_1, \ldots j_i)\, x_{n-j_1} \cdots x_{n-j_i}. \tag{1.5}$$

Niewątpliwą zaletą reprezentacji (1.4) i (1.5) jest ich uniwersalność. Rozważane szeregi pozwalają bowiem na opisanie każdego stacjonarnego i fizycznie realizowalnego (tj. przyczynowego) systemu spełniającego określone warunki ciągłości, bez względu na jego złożoność i strukturę wewnętrzną. Ponadto opis obiektu za pomocą rozwinięcia w szereg Volterry może być otrzymany zarówno dla losowych, jak i deterministycznych sygnałów wejściowych. Niestety zależności (1.4) i (1.5) posiadają również wady, z których najistotniejsze związane są ze sposobem wyznaczania współczynników $\vartheta_i(\,)$ i $\varrho_i(\,)$. Okazuje się bowiem, że w wielu przypadkach wyznaczenie jąder $\vartheta_i(\,)$ lub $\varrho_i(\,)$ jest zadaniem na tyle skomplikowanym i wymagającym znacznych mocy obliczeniowych, że praktyczne zastosowanie zależności (1.4) i (1.5) do identyfikacji systemów nie daje

zadowalających rezultatów, Billings i Fakhouri [6]. Równie istotnym ograniczeniem związanym z wykorzystaniem szeregów Volterry jest wymaganie, aby nieliniowość systemu dawała się rozwinąć w szereg Taylora.

Innym sposobem opisu szerokiej klasy dynamicznych systemów nieliniowych jest wykorzystanie reprezentacji zaproponowanej przez Wienera [136], zob. też Therrien [122]. W odróżnieniu od szeregów Volterry, koncepcja Wienera opiera się na wykorzystaniu zbioru funkcjonałów, które przy założeniu że sygnał wejściowy jest gaussowskim białym szumem, są względem siebie ortogonalne, Therrien [122], zob. też Greblicki i Pawlak [39]. Proponowana zależność wyraża się wzorem

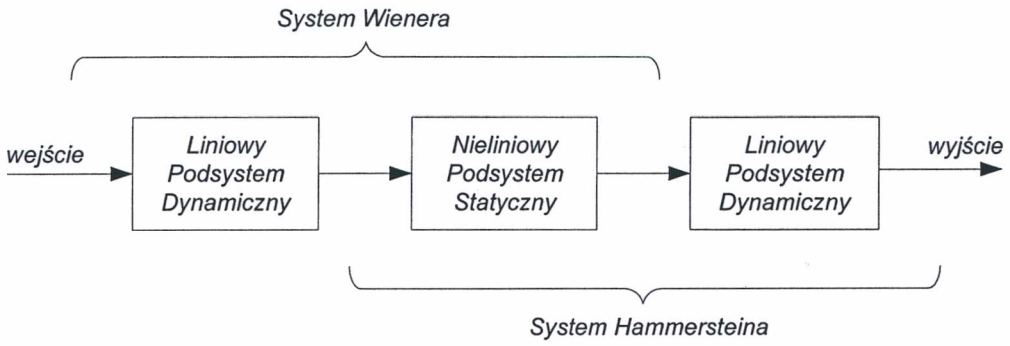
$$y(t) = \sum_{i=0}^{\infty} G_i \left[k_i(\tau) ; x(\tau) ; \tau \in (-\infty, t) \right], \tag{1.6}$$

gdzie $G_i(\,)$ są funkcjonałami o znanej postaci, ale posiadającymi nieznane jądra $k_i(\,)$. Niestety, podobnie jak zależności (1.4) i (1.5), również i to podejście nie pozwala na efektywne wykorzystanie w dziedzinie identyfikacji systemów złożonych. Zastosowanie reprezentacji (1.6) wymaga bowiem wykonania złożonych obliczeń, nawet w przypadku opisu systemów liniowych oraz nie pozwala na wykorzystanie w prosty sposób posiadanej informacji apriorycznej o systemie, Billings [5, str. 274].

Wspomniane powyżej, istotne ograniczenia reprezentacji systemów nieliniowych za pomocą funkcjonalnych zależności (1.4), (1.5) i (1.6), sprowokowały podjęcie badań nad wyznaczeniem nowych, równie uniwersalnych, ale łatwiejszych w zastosowaniach sposobów opisu systemów nieliniowych. Jednym z najczęściej wykorzystywanych obecnie podejść jest, sformułowana na początku lat osiemdziesiątych ubiegłego stulecia (ale wykorzystywana wcześniej np. przez Narendrę i Gallmana [82]), koncepcja systemów o strukturze blokowej (*ang. block–oriented systems*), Billings [5], Billings i Fakhouri [6–8]. Podejście to zakłada, że identyfikowany system można przedstawić w postaci odpowiednio ze sobą połączonych liniowych obiektów dynamicznych i nieliniowych obiektów statycznych. W odróżnieniu od szeregów funkcjonalnych umożliwiających reprezentację o charakterze czarnej skrzynki (*ang. black–box description*), Billings [5], podejście blokowo-zorientowane pozwoliło sprowadzić zadanie identyfikacji do tzw. problemu szarej skrzynki (*ang. gray–box description*), tj. zadania w którym znana jest struktura połączeń wewnętrznych systemu, Sjöberg *et al.* [117], Pearson i Pottmann [94].

Stosunkowo często rozważanym w literaturze systemem blokowo-zorientowanym jest tzw. system *sandwich* (zwany również systemem Wienera–Hammersteina) utworzony poprzez kaskadowe połączenie kolejno liniowego obiektu dynamicznego, nieliniowego obiektu statycznego oraz liniowego obiektu dynamicznego, rys. 1.2, Billings [5], Billings i Fakhouri [8], Bershad *et al.* [3], Kibangou i Favier [65], Krzyżak [71], Korenberg i Hunter [69]. Zaletą systemu sandwich jest jego elastyczność, umożliwiająca reprezentację szerokiej klasy zjawisk i procesów technologicznych o charakterze nieliniowym i posiadających pamięć. Główną wadą jest natomiast fakt występowania dwóch niedostępnych dla pomiarów sygnałów wewnętrznych, tj. wejścia i wyjścia statycznego obiektu nieliniowego. W konsekwencji, pomimo znanej struktury całego systemu, zadanie identyfikacji (na podstawie obserwacji wejścia i wyjścia systemu) jest zadaniem skomplikowanym. Okazuje się jednak, że wydzielenie z systemu sandwich prostszych obiektów, tzw. systemów Wienera i Hammersteina (zob. rys. 1.2), prowadzi do struktur równie interesujących z punktu widzenia zastosowań do opisu zjawisk nieliniowych, a przy tym łatwiejszych w identyfikacji.

Niech $\{x_n\}$, $\{y_n\}$ oznaczają odpowiednio wejście i wyjście systemu oraz $\{v_n\}$ będzie sygnałem wewnętrznym. Przez $\{z_n\}$ oznaczymy szum występujący w systemie. Przedsta-



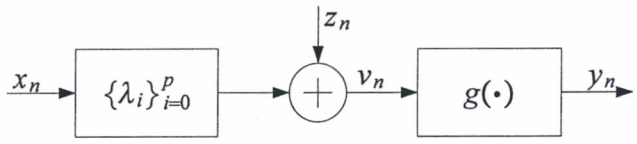
Rysunek 1.2: Struktura systemów sandwich, Wienera i Hammersteina

wiony na rysunku 1.3 system Wienera opisany jest za pomocą równań

$$v_n = \sum_{i=0}^{p} \lambda_i x_{n-i} + z_n \tag{1.7}$$

$$y_n = g(v_n), \tag{1.8}$$

gdzie $\{\lambda_i\}_{i=0}^{p}$ i $g(\cdot)$ są odpowiednio odpowiedzią impulsową podsystemu dynamicznego

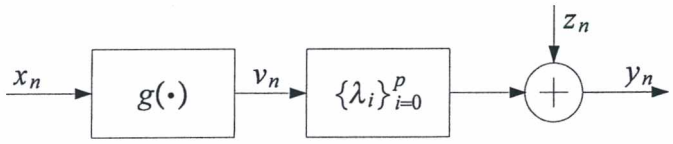


Rysunek 1.3: System Wienera z addytywnym sygnałem zakłócającym na wyjściu podsystemu dynamicznego

oraz charakterystyką nieliniową podsystemu statycznego, a $p \in \mathbf{N}$ jest długością pamięci systemu. Analogicznie system Hammersteina, zob. rys 1.4, definiowany jest za pomocą równań

$$v_n = g(x_n)$$

$$y_n = \sum_{i=0}^{p} \lambda_i v_{n-i} + z_n.$$



Rysunek 1.4: System Hammersteina

Opisane powyżej systemy Wienera i Hammersteina odgrywają kluczową rolę w teorii identyfikacji systemów nieliniowych. Ze względu na czytelną, blokowo-zorientowaną strukturę oraz liczne zastosowania, poczynając od lat siedemdziesiątych ubiegłego stulecia,

opracowano wiele różnych algorytmów ich identyfikacji oraz zaproponowano wykorzystanie ich w wielu dziedzinach naukowych i technicznych. Należy podkreślić, że istotny wkład w rozwój metod identyfikacji omawianych systemów został wniesiony przez Szkołę Identyfikacji Nieparametrycznej utworzoną przez profesora Włodzimierza Greblickiego. Jej twórca zapoczątkował rozwój nieparametrycznych metod identyfikacji systemów i wraz z profesorami Adamem Krzyżakiem, Mirosławem Pawlakiem i Leszkiem Rutkowskim opracował szereg nieparametrycznych algorytmów identyfikacji [26–29, 32, 33, 35, 37–43, 71, 91, 103]. Zaproponowane metody w istotny sposób rozwinęły teorię i wzbogaciły praktykę algorytmów, umożliwiając identyfikację systemów w sytuacjach, w których wcześniej opracowane metody parametryczne nie mogły być stosowane. Koncepcja nieparametrycznej identyfikacji systemów ze szczególnym uwzględnieniem algorytmów dedykowanych systemom Wienera omówiona jest w rozdziale 2.

Oprócz przedstawionych powyżej systemów typu sandwich, Wienera i Hammersteina, w literaturze rozważa się także inne systemy blokowo-zorientowane. Są to między innymi: systemy Urysona (będące równoległym połączeniem wielu systemów Hammersteina), Billings [5] i addytywne systemy typu NARMAX, Mzyk [80]. Jednak zasadnicza uwaga badaczy skupia się na zagadnieniach identyfikacji systemów Hammersteina i Wienera. Należy przy tym podkreślić, że pomimo podobnej konstrukcji, zadanie identyfikacji systemów Wienera jest na ogół znacznie trudniejsze od zadania identyfikacji systemów Hammersteina. Istotne utrudnienia wynikają m.in. z faktu, że w systemach Wienera charakterystyka nieliniowa przekształca *nieznany i niedostępny* dla pomiarów sygnał interakcyjny, podczas gdy w systemach Hammersteina przekształceniu nieliniowemu podlega *dostępny* dla pomiarów sygnał wejściowy.

1.2 Systemy Wienera

Jak już zaznaczono, ważną rolę wśród omówionych wyżej systemów o strukturze blokowej odgrywają systemy Wienera (rys. 1.3). Ze względu na swoją budowę, obiekty te znakomicie nadają się do modelowania zjawisk fizycznych i procesów technologicznych, posiadających nie tylko cechy układów nieliniowych, ale dodatkowo charakteryzujących się występowaniem pamięci. Przykładowe zastosowania systemów Wienera w naukach technicznych oraz medycynie, chemii i biologii omówione są w p. 1.4.
Oprócz podstawowych cech struktury charakteryzujących systemy Wienera można wyróżnić szereg własności określających dodatkowo elementy tej klasy systemów. Są to między innymi:

- długość pamięci systemu (pamięć skończona lub nieskończona),
- własności charakterystyki nieliniowej (np. ciągłość, różniczkowalność itp.),
- miejsce występowania sygnału zakłócającego (szum wewnętrzny lub na wyjściu systemu),
- stacjonarność.

Rozważane w literaturze systemy Wienera są ponadto:

- systemami z czasem ciągłym lub dyskretnym,
- systemami o strukturach typu SISO, MISO lub MIMO.

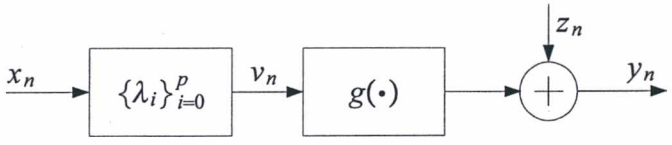
Długość pamięci systemu oraz własności charakterystyki nieliniowej w systemie mają szczególnie znaczenie przy doborze właściwego modelu w przypadku zastosowań omawianych systemów oraz w przypadku konstruowania algorytmów ich identyfikacji. Systemy ze skończoną pamięcią dyskutowane są m.in. przez Nordsjö i Zetterberga [84], Lacy'ego i Bernsteina [73], Pawlaka *et al.* [93], Wachla [127, 129], Raicha *et al.* [98] oraz Celkę *et al.* [12]. Z kolei Greblicki [26–29, 33] oraz Hughes i Westwick [56] rozważają identyfikację systemów z pamięcią nieskończoną. Typowym założeniem odnośnie charakterystyki nieliniowej w systemie jest wymaganie, aby miała ona charakter wielomianowy, Lacy i Bernstein [73], Korenberg i Hunter [67], Celka *et al.* [12], zob. też Pajunen [87]. Rozważa się także systemy z nieliniowościami modelowanymi za pomocą funkcji odcinkami liniowych, Vörös [124], Pupeikis [96], Chen [15], Wigren [137], zob. też Hagenblad [44], lub funkcji sklejanych (*ang. spline functions*), Hughes i Westwick [56]. Szczególnie szeroka, nieparametryczna klasa nieliniowości rozważana jest przez Greblickiego w pracach [26–29, 33]. W [26] zakłada się tylko, że nieliniowość jest funkcją ściśle monotoniczną i posiadającą ograniczoną pochodną, podczas gdy w [28] dopuszczalna jest każda fizycznie realizowalna charakterystyka nieliniowa. Dyskutowane w literaturze systemy Wienera z reguły są systemami stacjonarnymi. Systemy niestacjonarne (Wienera i Hammersteina) omawiane są m.in. przez Nordsjö i Zetterberga [84] oraz Celkę i Colditza [13].

W zależności od miejsca występowania sygnału zakłócającego, systemy Wienera podzielić można na trzy grupy: systemy, w których zakłócenie nie występuje, zob. np. Lacy i Bernstein [73], systemy z addytywnym sygnałem zakłócającym na wyjściu podsystemu dynamicznego (zob. rys. 1.3 i wzory (1.7)–(1.8)), Gómez i Baeyens [22, 23], Greblicki [26–29, 33] oraz systemy z szumem addytywnym na wyjściu podsystemu statycznego, rys. 1.5, opisane równaniami

$$v_n = \sum_{i=0}^{p} \lambda_i x_{n-i}, \tag{1.9}$$

$$y_n = g(v_n) + z_n, \tag{1.10}$$

zob. Pawlak *et al.* [93], Wachel [129], Hughes i Westwick [56], Bai [1].



Rysunek 1.5: System Wienera z addytywnym sygnałem zakłócającym na wyjściu systemu

Okazuje się, że obecność zakłóceń i miejsce ich występowania w systemie mają istotny wpływ na możliwości identyfikacji systemu i postać algorytmów.

Ze względu na powszechność techniki cyfrowej, większość zastosowań oraz opracowanych algorytmów identyfikacji systemów Wienera dotyczy systemów z czasem dyskretnym. Zagadnienie identyfikacji systemów z czasem ciągłym dyskutowane jest np. przez Greblickiego w pracy [29]. W kontekście liczby wejść i wyjść typową i najczęściej rozważaną w literaturze architekturą systemów Wienera jest architektura typu SISO (*ang. Single Input, Single Output*), zob. np. Billings i Fakhouri [8], Greblicki [26–29, 33]. Systemy Wienera o strukturze MIMO (*ang. Multiple Inputs, Multiple Outputs*) dyskutowane są przez Westwicka i Verhaegena [135] oraz Jeonga *et al.* [62]. Z kolei algorytmy identyfikacji systemów typu MISO (*ang. Multiple Inputs, Single Output*) omawiane są m.in. przez Westwicka i Kearney'ego [134] oraz Zhu [141]. Szczególny rodzaj systemu

Wienera utworzonego z liniowego obiektu dynamicznego typu SIMO i nieliniowego obiektu statycznego typu MISO wykorzystywany jest przez Boyda i Chuę [9] do aproksymacji szerokiej klasy systemów nieliniowych, zob. p. 1.3.1.

1.3 Własności aproksymacyjne systemów Hammersteina i Wienera

Omówiona w p. 1.1 reprezentacja systemów nieliniowych przy użyciu operatorów Volterry z reguły nie jest bezpośrednio wykorzystywana w identyfikacji systemów. Stanowi ona jednak podstawę do konstruowania i analizy systemów, o strukturach umożliwiających aproksymację lub dokładną reprezentację szerokiej klasy dynamicznych systemów nieliniowych. W [115] Shi i Sun rozważają reprezentację systemów nieliniowych za pomocą równolegle połączonych systemów o strukturze sandwich. Dowodzą oni, że systemy z czasem dyskretnym, skończoną pamięcią i posiadające reprezentację w postaci skończonego szeregu Volterry, mogą być precyzyjnie opisane za pomocą sumy skończonej liczby systemów typu sandwich. Z kolei w [68] Korenberg dyskutuje reprezentację systemów nieliniowych za pomocą odpowiednio połączonych liniowych obiektów dynamicznych z nieliniowymi obiektami statycznymi. W pracy tej wykazano, że każdy system nieliniowy o skończonej pamięci z czasem dyskretnym, dla którego istnieje reprezentacja w postaci skończonego szeregu Volterry może być dokładnie reprezentowany przez równoległe połączenie skończonej liczby systemów Wienera. Powyższy wynik prowadzi więc do ogólnego wniosku, że systemy będące równoległym połączeniem systemów Wienera typu SISO mogą być wykorzystywane do reprezentacji szerokiej klasy systemów nieliniowych. Nieco inne podejście do aproksymacji szerokiej klasy dynamicznych systemów nieliniowych proponuje Sandberg w serii prac [107–109]. W [107] autor rozważa koncepcję reprezentacji przyczynowych i stacjonarnych systemów nieliniowych z czasem dyskretnym za pomocą obiektów o strukturze Hammersteina. Omawiane podejście zakłada, że aproksymowane systemy posiadają „pamięć niemal skończoną" (*ang. approximately-finite-memory*[1]). Ten sam warunek jest również wykorzystywany przez Parka i Sandberga w pracy [89]. Na podstawie prezentowanych w literaturze rezultatów, dla szerokiej klasy systemów nieliniowych autorzy formułują następującą, nieformalną tezę: „jeżeli wyjście systemu w dowolnej chwili t_0 zależy od dalekiej przeszłości sygnału wejściowego tylko w dowolnie małym stopniu, to system może być dowolnie dokładnie aproksymowany przez proste struktury takie jak: skończone szeregi Volterry, dynamiczne wielowarstwowe sieci neuronowe lub sieci radialne". Główne rezultaty teoretyczne zawarte w pracy zawierają sformułowanie jednolitych warunków koniecznych i wystarczających jakie musi spełniać system aby istniała reprezentacja za pomocą wymienionych wyżej struktur. Własności aproksymacyjne systemów o strukturze blokowej (w szczególności systemów Hammersteina i Wienera), na tle podejść wykorzystujących rozwinięcia w szeregi Volterry i Wienera, dyskutowane są również przez Billingsa w przekrojowej pracy [5].

Innym przykładem reprezentacji złożonych systemów nieliniowych jest, oparta na własności zanikania pamięci, teoria aproksymacji systemów nieliniowych zaproponowana przez Boyda i Chuę w pracy [9].

[1] tłumaczenie autora.

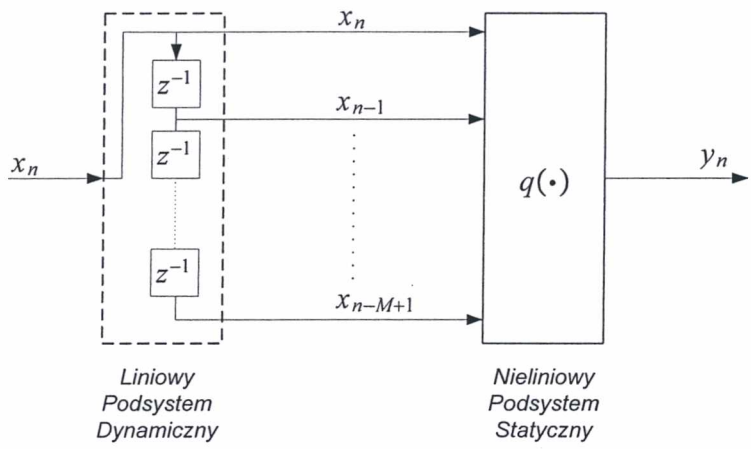
1.3.1 Aproksymacja dynamicznych systemów nieliniowych za pomocą systemów Wienera

Przedstawione przez Boyda i Chuę [9] rezultaty odnoszą się do szerokiej klasy systemów stacjonarnych z zanikającą pamięcią (*ang. fading memory*). Intuicyjnie system posiada zanikającą pamięć, jeżeli dwa sygnały wejściowe, bliskie w teraźniejszości, ale różniące się w przeszłości, powodują że odpowiadające im sygnały wyjściowe są odpowiednio bliskie w teraźniejszości (ścisłą definicję znaleźć można np. w [9]). Wymaganie, aby system posiadał zanikającą pamięć nie jest więc wymaganiem istotnie ograniczającym klasę spotykanych w praktyce systemów, a w przypadku systemów liniowych jest równoważne z wymaganiem stabilności systemu. Nieco silniejszym ograniczeniem zawężającym klasę systemów rozważanych w [9] jest natomiast wymaganie, aby nieliniowość w systemie była funkcją analityczną[2].

Rezultaty otrzymane w omawianej pracy dotyczą zarówno systemów z czasem ciągłym jak i z czasem dyskretnym. W przypadku systemów z czasem dyskretnym, w pracy wykazano, że każdy system spełniający określone powyżej warunki może być aproksymowany z dowolną dokładnością przez skończony szereg Volterry, zob. Boyd i Chua [9, Tw. 3]. Szereg ten można z kolei zapisać w wygodniejszej, blokowo-zorientowanej postaci jako tzw. operator NLMA (*ang. Nonlinear Moving–Average*) określony wzorem

$$y_n = q\left(x_n, x_{n-1}, \ldots, x_{n-M+1}\right), \tag{1.11}$$

gdzie $\{x_n\}$ i $\{y_n\}$ są odpowiednio sygnałem wejściowym i wyjściowym, $q : \mathbf{R}^M \to \mathbf{R}$ jest wielomianem, a $M < \infty$ pewną stałą.
Operator (1.11) można następnie przedstawić w postaci systemu Wienera (typu SISO)



Rysunek 1.6: System Wienera jako kaskadowe połączenie liniowego podsystemu dynamicznego typu SIMO z nieliniowym podsystemem statycznym typu MISO

utworzonego poprzez kaskadowe połączenie liniowego obiektu dynamicznego typu SIMO z nieliniowym obiektem statycznym typu MISO, zob. rys. 1.6. Pomimo stosunkowo nietypowej struktury otrzymanego systemu Wienera, powyższy rezultat prowadzi do wniosku, że systemy Wienera stanowią ważną klasę obiektów umożliwiających reprezentację szerokiej rodziny nieliniowych systemów dynamicznych.

[2] tj. funkcją, która daje się rozwinąć w szereg Taylora, Fichtenholz [20, t. II, str. 388].

1.4 Przykłady zastosowań praktycznych

Nieskomplikowana struktura systemów Wienera oraz fakt występowania w nich jednocześnie pamięci i nieliniowości, powoduje że systemy te znajdują zastosowanie w wielu, często niezwiązanych za sobą, dziedzinach nauki. Poniżej omówiono dwie przykładowe aplikacje w telekomunikacji (kompensacja zniekształceń nieliniowych w systemach modulacji sygnałów) oraz chemii (sterowanie procesem neutralizacji pH). Wymienione zostały także zastosowania m.in. w takich dziedzinach jak medycyna, optyka i przetwarzanie obrazów.

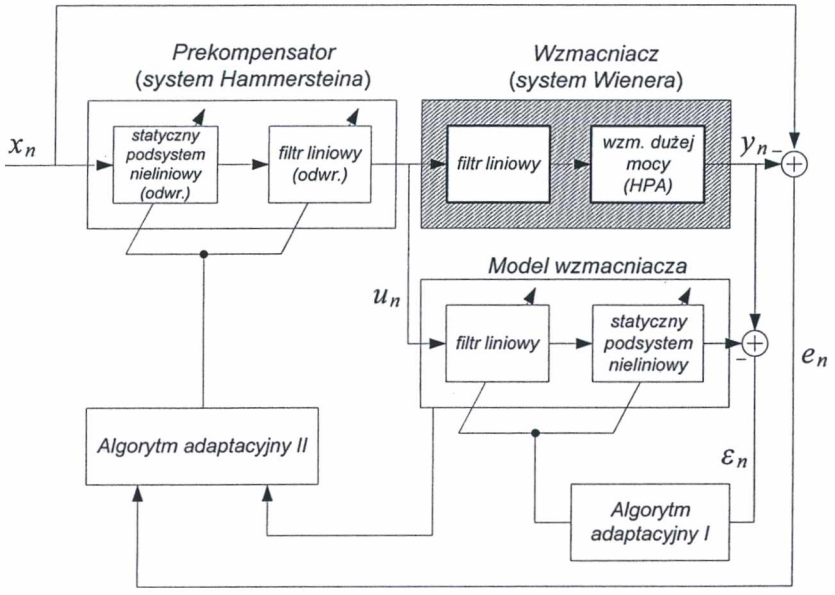
1.4.1 Kompensacja zniekształceń nieliniowych w systemach modulacji sygnałów typu OFDM

Ze względu na liczne zalety technika transmisji danych OFDM (*ang. Orthogonal Frequency Division Method*) jest obecnie powszechnie wykorzystywana m.in. w szerokopasmowych systemach cyfrowych, takich jak: systemy telewizji cyfrowej, bezprzewodowy dostęp do Internetu, ADSL itp. Transmisja OFDM, polegająca na rozdzieleniu pojedynczego strumienia danych (bitów) na wiele podnośnych, oprócz wielu istotnych zalet posiada również wady. W szczególności, istotnym problemem jest występowanie zniekształceń nieliniowych wprowadzanych przez występujące w systemie wzmacniacze.

W pracy [64] Kang *et al.* dyskutują technikę adaptacyjnej prekompensacji zniekształceń nieliniowych występujących w systemach OFDM i powodowanych przez filtr liniowy połączony kaskadowo ze wzmacniaczem dużej mocy HPA (*ang. High Power Amplifier*). W omawianym podejściu statyczny wzmacniacz HPA, poprzedzony obiektem liniowym, modelowany jest za pomocą systemu Wienera, a prekompensacja zniekształceń polega na zastosowaniu na wejściu, układu o strukturze Hammersteina z odpowiednio dobraną pamięcią i nieliniowością. Zadaniem systemu Hammersteina jest takie wstępne przetworzenie sygnału wejściowego, aby po przejściu przez nieliniowy element (system Wienera) nie występowały zniekształcenia nieliniowe. Proponowany algorytm prekompensacji przedstawiony jest schematycznie na rys. 1.7. W celu dostrojenia układu prekompensującego, tj. systemu Hammersteina, dokonywana jest identyfikacja nieznanego systemu Wienera. Zadanie to realizowane jest przez Algorytm Adaptacyjny i ustalający parametry modelu wzmacniacza na podstawie wielkości błędu ε_n (zob. rys. 1.7). Otrzymane w procesie identyfikacji parametry nieznanego systemu Wienera, wraz z informacją o błędzie e_n (zob. rys. 1.7), wykorzystywane są następnie przez Algorytm Adaptacyjny II do wyznaczenia charakterystyki nieliniowej i filtru liniowego w systemie Hammersteina.

1.4.2 Proces neutralizacji pH

Istotnym elementem wielu procesów produkcyjnych jest powstawanie ścieków o charakterze kwasowym lub zasadowym. W takich sytuacjach substancje odpadowe poddaje się neutralizacji, polegającej na wymieszaniu ścieków z odpowiednią ilością odczynnika o przeciwnym pH. Sterowanie omawianym procesem jest jednak zadaniem stosunkowo skomplikowanym, ze względu na występującą w nim inercję oraz silne zależności nieliniowe. Z tego powodu, przy konstruowaniu algorytmów sterowania wykorzystuje się modele nieliniowe omawianego zjawiska. Analiza odpowiednich reakcji chemicznych prowadzi do wniosku, że modele o strukturze Wienera szczególnie dobrze reprezentują omawiany proces zobojętniania, por. Gómez *et al.* [23], Norquay *et al.* [85], Kalafatis *et al.* [63], Pajunen [87]. Zagadnienie neutralizacji pH omówione zostanie krótko na przykładzie procesu rozważanego przez Gómeza *et al.* w pracy [23]. Do zbiornika, zob. rys. 1.8, dostarczane są trzy substancje: wodorotlenek sodu ($NaOH$), wodorowęglan sodu ($NaHCO_3$) oraz kwas azotowy(V) (HNO_3). Wielkościami wejściowymi w omawianym procesie są szybkości



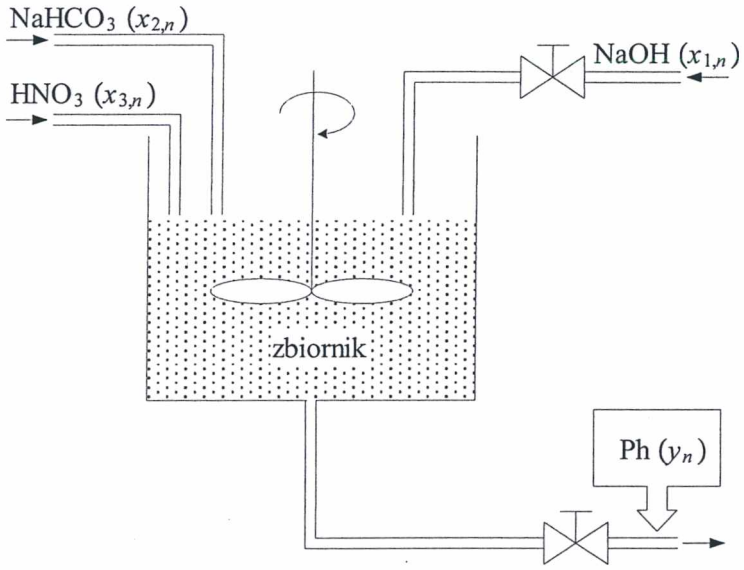
Rysunek 1.7: Schemat prekompensacji zniekształceń nieliniowych wprowadzanych przez wzmacniacz dużej mocy w systemie OFDM

przepływu (odpowiednio $x_{1,n}$, $x_{2,n}$ i $x_{3,n}$) wymienionych wyżej związków. Wyjściem y_n całego systemu jest wartość współczynnika pH substancji wypływającej ze zbiornika. Zarówno szybkość przepływu $x_{3,n}$ kwasu azotowego jak i poziom cieczy w zbiorniku są wielkościami stałymi. Celem procesu neutralizacji jest uzyskanie odpowiedniego pH substancji wypływającej ze zbiornika (ścieku) poprzez regulację przepływu $x_{1,n}$, tj. ilości zasady NaOH dostarczanej do zbiornika. Zakłada się przy tym, że szybkość przepływu $x_{2,n}$ soli $NaHCO_3$ jest nieznana.
Powyższe założenia prowadzą do wniosku, że wejściem i wyjściem rozważanego obiektu identyfikacji są odpowiednio $x_{1,n}$ i y_n, podczas gdy $x_{2,n}$ jest niedostępnym dla pomiarów szumem, a $x_{3,n}$ stałym parametrem.

1.4.3 Inne zastosowania

Oprócz omówionych wyżej aplikacji, systemy Wienera wykorzystywane są z powodzeniem również w innych dziedzinach. Dalsze przykłady ich zastosowań to m.in.:

- modelowanie procesu destylacji (identyfikacja kolumny destylacyjnej), Zhu [141], Pearson i Pottmann [94],
- sterowanie procesem polimeryzacji, Jeong *et al.* [62],
- zastosowania w neurologii, Marmarelis i Naka [79],
- modelowanie układu wzrokowego człowieka, den Brinker [16],
- modelowanie i analiza zapisu EEG u niemowląt, Celka i Colditz [13],
- zastosowania w naukach biologicznych, Hunter i Korenberg [57],
- zastosowania w optyce, Celka *et al.* [12],



Rysunek 1.8: Proces neutralizacji Ph

- identyfikacja systemów chaotycznych z czasem ciągłym i dyskretnym, Chen *et al.* [14], Xu *et al.* [140],
- modelowanie prędkości przepływu wody w rzece, Sbárbaro i Johansen [110],
- usuwanie szumów z obrazów, Sekko *et al.* [114].

1.5 Systemy Wienera – podsumowanie zastosowań praktycznych i własności teoretycznych

Przedstawiona wyżej krótka charakterystyka rozważanych w literaturze systemów Wienera dowodzi, że obiekty te cieszą się dużym zainteresowaniem badaczy. Ich własności teoretyczne oraz prosta konstrukcja powodują, że stanowią one istotny element teorii aproksymacji dynamicznych systemów nieliniowych. Jednocześnie znajdują one szereg zastosowań praktycznych w wielu często niezwiązanych ze sobą dziedzinach naukowych. Powyższe obserwacje w pełni uzasadniają więc celowość konstruowania algorytmów identyfikacji systemów Wienera.

1.6 Określenie zakresu tematycznego pracy

W pracy rozważa się zadanie identyfikacji systemów Wienera z czasem dyskretnym i skończoną pamięcią. Zakłada się, że długość pamięci systemu jest znana. Równocześnie przyjmuje się założenie, że wiedza wstępna o nieliniowej charakterystyce w systemie jest bardzo mała i ma charakter nieparametryczny. Szczególny nacisk kładzie się na opracowanie algorytmów umożliwiających identyfikację w przypadku losowych sygnałów wejściowych – o rozkładach innych niż rozkład normalny. Proponuje się parametryczno–nieparametryczne algorytmy identyfikacji, które umożliwiają identyfikację liniowego i nieliniowego podsystemu z dokładnością do nieznanej multiplikatywnej stałej. Dla skonstruowanych metod identyfikacji przeprowadza się analizę własności asymptotycznych. Wykonane badania eksperymentalne ilustrują natomiast zachowanie się algorytmów dla małej i średniej liczby obserwacji.

Rozdział 2

Parametryczno–nieparametryczna identyfikacja systemów Wienera. Teza i cele pracy

2.1 Parametryczna i nieparametryczna identyfikacja systemów

Wiedza aprioryczna o strukturze systemu oraz charakterze sygnału wejściowego i zakłócającego ma kluczowe znaczenie przy doborze właściwej metody identyfikacji, Hasiewicz [49], Greblicki [25], Söderström i Stoica [118]. Gdy jest ona na tyle duża, że zadanie identyfikacji daje się sprowadzić do wyznaczenia skończonej i znanej liczby parametrów systemu, możliwe jest zastosowanie tzw. parametrycznych metod identyfikacji. W wielu przypadkach natomiast, dostępna informacja o systemie jest znacznie bardziej ograniczona, a jej poszerzenie może być zbyt kosztowne lub niemożliwe. W sytuacjach takich, w zależności od charakteru posiadanej informacji wstępnej, konieczne jest zastosowanie tzw. algorytmów nieparametrycznych lub parametryczno–nieparametrycznych.

W rozdziale przedstawiona jest krótka charakterystyka parametrycznych, nieparametrycznych oraz nowych koncepcyjnie parametryczno–nieparametrycznych algorytmów identyfikacji systemów, ze szczególnym uwzględnieniem technik zastosowanych w pracy. Ponadto omówione są krótko, dotychczas prezentowane w literaturze, parametryczne i nieparametryczne algorytmy identyfikacji systemów Wienera.

2.1.1 Metody parametryczne

W zależności od zakresu posiadanej informacji wstępnej, metody parametryczne opracowane dla celów identyfikacji systemów bazują na metodzie najmniejszych kwadratów, metodzie największej wiarygodności lub też metodzie największego prawdopodobieństwa a posteriori, Hasiewicz [49], Söderström i Stoica [118], Ljung [76], Mańczak i Nahorski [78], Goodwin i Payne [24]. Przykładowo, dla statycznego systemu nieliniowego opisanego równaniem

$$y_n = g(x_n; \theta) + Z_n,$$

gdzie funkcja $g(\cdot; \theta)$ znana jest z dokładnością do wektora parametrów $\theta \in \mathbf{R}^k$, $k < \infty$, a $\{Z_n\}$ jest stacjonarnym białym szumem o zerowej wartości oczekiwanej $E\{Z_n\} = 0$ i skończonej wariancji $Var\{Z_n\} < \infty$, metoda najmniejszych kwadratów polega na

minimalizacji funkcji kryterialnej

$$\Gamma(\theta) = \sum_{n=1}^{N} [y_n - g(x_n; \theta)]^2,$$

gdzie N jest liczbą par obserwacji (x_n, y_n) wejścia i wyjścia systemu. Dla takiego podejścia opracowano szereg uogólnień i modyfikacji, umożliwiających stosowanie metody najmniejszych kwadratów w przypadku bardziej złożonych obiektów identyfikacji, zob. np. Mzyk [80]. Analizę własności asymptotycznych metody znaleźć można m.in. w Söderström i Stoica [118], Van der Vaart [123], Jennrich [61], Wu [139].

2.1.2 Metody nieparametryczne

W odróżnieniu od metod parametrycznych, idea nieparametrycznej identyfikacji systemów jest relatywnie nowa, a u jej podstaw leży teoria nieparametrycznego wnioskowania statystycznego i w szczególności zagadnienie nieparametrycznej estymacji funkcji regresji. **Metody nieparametryczne wymagają jedynie niewielkiej informacji wstępnej o identyfikowanym systemie i przez to umożliwiają rozwiązanie zadań identyfikacji będących poza zasięgiem jakichkolwiek podejść parametrycznych.** Koncepcja omawianych algorytmów polega na znalezieniu zależności pomiędzy szukaną charakterystyką nieliniową występującą w systemie, a odpowiednią funkcją regresji dającą się w łatwy sposób estymować na podstawie zbioru posiadanych pomiarów. W typowym zadaniu nieparametrycznej identyfikacji prostego systemu statycznego zakłada się, że wejście $\{X_n\}$ systemu jest ciągiem *i.i.d.* zmiennych losowych, a identyfikacja przeprowadzana jest na podstawie zbioru par obserwacji wejścia i wyjścia systemu, reprezentowanych przez zbiór zmiennych losowych $\{(X_n, Y_n)\}_{n=1}^{N}$. W przypadku statycznego systemu nieliniowego opisanego równaniem

$$Y_n = g(X_n) + Z_n,$$

gdzie $\{Z_n\}$ jest losowym szumem o zerowej wartości oczekiwanej, niezależnym od wejścia systemu, a charakterystyka $g(\cdot)$ jest całkowicie nieznana, prawdziwa jest następująca zależność

$$g(x) = E\{Y_n | X_n = x\}. \tag{2.1}$$

Zadanie identyfikacji nieliniowości $g(\cdot)$ można więc realizować poprzez estymację funkcji regresji (wyjścia względem wejścia) występującej w (2.1).

Gwałtowny rozwój nieparametrycznych metod identyfikacji dynamicznych systemów o złożonej strukturze zapoczątkowany został opublikowanymi przez Greblickiego i Pawlaka rezultatami badań dotyczących nieparametrycznej identyfikacji systemów Hammersteina. W [38] wykazali oni, że dla nieliniowej charakterystyki $g(\cdot)$ występującej w takim systemie prawdziwa jest następująca zależność

$$g(x) = aE\{Y_n | X_n = x\} + b,$$

gdzie a i b są stałymi, niemożliwymi niestety do wyznaczenia bez dodatkowej informacji wstępnej. Powyższy wynik stanowi podstawę konstrukcyjną nieparametrycznych algorytmów identyfikacji również innych systemów blokowo-zorientowanych (ale nie systemów Wienera), zob. np. Greblicki i Pawlak [39], Pawlak i Hasiewicz [92], Hasiewicz [50], Greblicki [31], Hasiewicz [51], Hasiewicz i Śliwiński [54]. Z kolei Greblicki, w cyklu prac [26–28], zastosował podejście nieparametryczne do identyfikacji systemów Wienera. Uzyskane tam rezultaty omówione będą szerzej w p. 2.2.2.

Nieparametryczna estymacja funkcji regresji

Koncepcja nieparametrycznej estymacji funkcji regresji, zaproponowana niezależnie przez Watsona [133] i Nadarayę [81], wywodzi się od jądrowych estymatorów funkcji gęstości prawdopodobieństwa, Parzen [90]. Dla zmiennej losowej X posiadającej gęstość prawdopodobieństwa $f(\cdot)$, jądrowy estymator funkcji gęstości, zbudowany w oparciu o N–elementowy zbiór $\{X_n\}_{i=1}^N$ niezależnych zmiennych losowych o rozkładzie[1] $f(\cdot)$, przyjmuje postać

$$\hat{f}_N(x) = \frac{1}{Nh} \sum_{n=1}^{N} K\left(\frac{x - X_n}{h}\right), \tag{2.2}$$

gdzie $K(\cdot)$ jest tzw. funkcją jądra, a $h = h(N)$ dodatnim ciągiem liczbowym (tzw. parametrem wygładzania, *ang. bandwidth parameter*), takim że $h \to 0$ i $Nh \to \infty$ przy $N \to \infty$. Idea estymacji gęstości $f(\cdot)$ w ustalonym punkcie x polega zatem na przybliżaniu wartości $f(x)$ na podstawie danych pomiarowych, które znajdują się odpowiednio blisko punktu x, zob. np. Greblicki [25], Hasiewicz [49], Kulczycki [72], Wand i Jones [132]. Przyjmując odpowiednie założenia odnośnie jądra $K(\cdot)$ można pokazać, że estymator $\hat{f}_N(\cdot)$ jest zgodnym estymatorem gęstości $f(\cdot)$ w każdym punkcie ciągłości funkcji $f(\cdot)$. Dla estymatora $\hat{f}_N(\cdot)$ uzyskano ponadto szereg wyników teoretycznych określających warunki pozwalające otrzymać inne niż punktowa rodzaje zbieżności. Przykładowo, zbieżność jednostajna dyskutowana jest przez Silvermana [116], a zbieżność estymatora w metryce L_1 omawiana jest przez Devroye'a i Györfiego w [18].

Niech X i Y będą zmiennymi losowymi o łącznej gęstości prawdopodobieństwa $f_{XY}(\cdot,\cdot)$ oraz $E|Y| < \infty$. Podstawą konstrukcji jądrowego estymatora funkcji regresji

$$\mu(x) = E\{Y|X = x\}$$

jest obserwacja, że zgodnie z definicją warunkowej wartości oczekiwanej,

$$\mu(x) = \frac{r(x)}{f(x)}, \tag{2.3}$$

gdzie $r(x) = \int_{-\infty}^{\infty} y f(x,y)\,dy$ oraz $f(x)$ jest gęstością prawdopodobieństwa zmiennej losowej X. Wykorzystując N–elementowy zbiór $\{(X_n, Y_n)\}_{n=1}^N$ zmiennych losowych o rozkładzie $f_{XY}(\cdot,\cdot)$, funkcję $\mu(\cdot)$ estymuje się przybliżając funkcje $r(\cdot)$ i $f(\cdot)$ w powyższym wyrażeniu. Szczególnie często rozważanymi estymatorami regresji $\mu(\cdot)$ zbudowanymi w oparciu o dekompozycję (2.3) są estymatory jądrowe (*ang. kernel estimates*) oraz estymatory wykorzystujące rozwinięcia funkcji w szereg ortogonalny (*ang. orthogonal series estimates*). Konstruowane dalej w pracy algorytmy identyfikacji systemów Wienera budowane są w oparciu o estymatory jądrowe.

Jądrowe estymatory regresji. Pierwsze prace poświęcone jądrowym estymatorom funkcji regresji opublikowano w latach sześćdziesiątych ubiegłego stulecia, Watson [133], Nadaraya [81], a gwałtowny wzrost zainteresowania nowo otrzymanymi metodami zaowocował szeregiem rezultatów rozwijających wstępne wyniki w wielu kierunkach, zob. np. Greblicki *et al.* [35], Devroye [17], Härdle [46], Wand i Jones [132].

Wstawiając estymator (2.2) do mianownika wyrażenia (2.3) oraz konstruując analogiczne oszacowanie dla całki w liczniku wyrażenia (2.3), otrzymuje się następujący estymator

[1] W zastosowaniach praktycznych ciągowi zmiennych losowych $\{X_n\}_{n=1}^N$ odpowiada zbiór obserwacji (pomiarów) $\{x_n\}_{n=1}^N$ będący realizacją ciągu (procesu stochastycznego) $\{X_n\}_{n=1}^N$.

funkcji regresji[2] (zob. np. Greblicki [25], Hasiewicz [49], Kulczycki [72])

$$\hat{\mu}_N(x) = \frac{\sum_{n=1}^{N} Y_n K\left(\frac{x - X_n}{h}\right)}{\sum_{n=1}^{N} K\left(\frac{x - X_n}{h}\right)}, \tag{2.4}$$

gdzie $K(\cdot)$ i h są odpowiednio funkcją jądra i parametrem wygładzania zależnym od liczby obserwacji N. Przyjmując odpowiednie założenia o funkcji jądra $K(\cdot)$ i parametrze wygładzania h można wykazać, że estymator $\hat{\mu}_N(\cdot)$ zbiega do funkcji $\mu(\cdot)$ według prawdopodobieństwa gdy $N \to \infty$, w każdym punkcie x, w którym $\mu(x)$ i $f(x)$ są ciągłe oraz $f(x) > 0$. Ze względu na nieskomplikowaną budowę i łatwość implementacji, estymator $\hat{\mu}_N(\cdot)$ stanowi istotny element konstrukcyjny wielu nieparametrycznych algorytmów identyfikacji. Podobnie jak dla jądrowego estymatora gęstości prawdopodobieństwa, dla omawianego estymatora funkcji regresji otrzymano szereg wyników teoretycznych określających warunki, przy których $\hat{\mu}_N(\cdot)$ posiada różne własności asymptotyczne. Przykładowo zagadnienie asymptotycznej normalności estymatora $\hat{\mu}_N(\cdot)$ dyskutowane jest przez Nadarayę [81] i Schustera [113], problematyka optymalnej szybkości zbieżności $\hat{\mu}_N(\cdot)$ omawiana jest przez Stona w [120], a Mack i Silverman [77] podają warunki, przy których zachodzi słaba i mocna jednostajna zgodność $\hat{\mu}_N(\cdot)$. W [35] Greblicki *et al.* rozważają z kolei punktową zgodność estymatora $\hat{\mu}_N(\cdot)$ w przypadku dowolnego rozkładu prawdopodobieństwa zmiennej wejściowej X (*ang. distribution–free consistency*).

2.1.3 Metody parametryczno–nieparametryczne

Algorytmy parametryczno–nieparametryczne umożliwiają identyfikację w sytuacji, gdy wiedza aprioryczna o systemie ma mieszany, parametryczny i nieparametryczny, charakter. Typowym przykładem tego rodzaju informacji wstępnej jest posiadanie dużej wiedzy (parametrycznej) o jednym z podsystemów i małej (nieparametrycznej) na temat drugiego podsystemu.

W kontekście identyfikacji systemów Hammersteina pierwszy tego rodzaju algorytm zaproponowany został przez Hasiewicza i Mzyka w [52]. Wykorzystując technikę nieparametrycznej estymacji funkcji regresji, w pracy skonstruowano estymator wewnętrznego sygnału interakcyjnego, który następnie wykorzystano do przeprowadzenia niezależnie parametrycznej identyfikacji nieliniowego podsystemu statycznego i liniowego dynamicznego (metodą najmniejszych kwadratów).

2.2 Aktualny stan badań na temat identyfikacji systemów Wienera

Prezentowane w literaturze metody identyfikacji systemów Wienera klasyfikować można według różnych kryteriów. Typowa i często spotykana systematyka opiera się na podziale algorytmów ze względu na ilość informacji apriorycznej niezbędnej dla zastosowania określonej metody identyfikacji (algorytmy parametryczne i nieparametryczne). Inny, właściwy dla systemów Wienera podział, związany jest z charakterem sygnału wejściowego oraz szczególnymi własnościami charakterystyki nieliniowej. Ze względu na specyficzną

[2] W przypadku gdy licznik i mianownik estymatora są równe zero, zazwyczaj przyjmuje się, że $0/0 = 0$ (por. np. Greblicki i Pawlak [39]). Dla rozważanych w pracy estymatorów z jądrem również zastosujemy tę konwencję.

konstrukcję omawianych systemów (w szczególności fakt występowania podsystemu liniowego przed podsystemem nieliniowym) zazwyczaj przyjmowanym założeniem jest wymaganie, aby sygnał wejściowy systemu miał charakter gaussowski, zob. Billings i Fakhouri [6], Bershad *et al.* [2], Celka *et al.* [11], Chen [15], Enqvist [19], Greblicki [26–28], Westwick i Verhaegen [135]. Algorytmy umożliwiające identyfikację dla niegaussowskich sygnałów wejściowych są rzadziej omawiane w literaturze, zob. np. Pawlak *et al.* [93], Taleb *et al.* [121]. Innym charakterystycznym dla systemów Wienera założeniem jest wymaganie, aby charakterystyka nieliniowa była funkcją odwracalną, zob. np. Greblicki [26,27], Gómez *et al.* [23], Gómez i Baeyens [22], Raich *et al.* [98], Solé-Casals *et al.* [119], Taleb *et al.* [121]. Z kolei Greblicki [28], Hughes i Westwick [56], Lacy i Bernstein [73] oraz Wigren [138] prezentują algorytmy nie wymagające odwracalności nieznanej charakterystyki.

Fakt występowania nieliniowego obiektu statycznego za liniowym obiektem dynamicznym może powodować istotne komplikacje w identyfikacji, w przypadku gdy na wyjściu systemu występuje addytywny sygnał zakłócający (por. p. 1.2). Z tego powodu prezentowane w literaturze algorytmy dają się klasyfikować również pod kątem „lokalizacji" niedostępnych dla pomiaru zakłóceń. Identyfikacja systemów Wienera, w których szum nie występuje rozważana jest np. przez Lacy'ego i Bernsteina [73]. Systemy z addytywnym sygnałem zakłócającym interakcje, tj. występującym na wyjściu podsystemu dynamicznego, dyskutują Gómez i Baeyens [22,23] oraz Greblicki [26–29,33]. Z kolei identyfikacja systemów z szumem na wyjściu całego systemu rozważana jest przez Baia [1], Hughesa i Westwicka [56], Pawlaka *et al.* [93] oraz Wachla [129].

W dalszej części rozdziału omówione zostaną prezentowane dotychczas w literaturze parametryczne i nieparametryczne metody identyfikacji systemów Wienera i na tym tle przedstawiona zostanie ogólna koncepcja proponowanych w pracy nowych algorytmów parametryczno–nieparametrycznych.

2.2.1 Parametryczna identyfikacja systemów Wienera

W typowym zadaniu parametrycznej identyfikacji systemów Wienera zakłada się zazwyczaj, że podsystem dynamiczny posiada skończoną pamięć (o znanej długości), a nieliniowa charakterystyka $g(\cdot)$ jest wielomianem znanego stopnia, Lacy i Bernstein [73]. Rzadziej rozważa się ogólniejsze podejścia, w których pamięć systemu nie musi być skończona. Przykładowo Norquay *et al.* [85] proponują algorytm umożliwiający parametryczną identyfikację systemów Wienera z wyjściowym szumem addytywnym (rys. 1.5) i podsystemem liniowym opisywanym równaniem różnicowym

$$V_n = \frac{B\left(q^{-1}\right)}{A\left(q^{-1}\right)} X_n,$$

gdzie $A\left(q^{-1}\right) = 1 + a_1 q^{-1} + \ldots + a_k q^{-k}$ oraz $B\left(q^{-1}\right) = b_1 q^{-1} + \ldots + b_l q^{-l}$ są wielomianami znanych stopni k i l, a q^{-1} jest operatorem przesunięcia wstecz. Nieliniowa charakterystyka $g(\cdot)$ jest wielomianem znanego stopnia m,

$$g(v) = \gamma_1 v + \gamma_2 v^2 + \cdots \gamma_m v^m.$$

Zadanie identyfikacji polega na wyznaczeniu współczynników wielomianów $A\left(q^{-1}\right)$, $B\left(q^{-1}\right)$ i $g(v)$, tj. elementów wektora $\theta_0 = [a_1, \ldots, a_k, b_1, \ldots, b_l, \gamma_1, \ldots, \gamma_2, \ldots, \gamma_m]^T$, na podstawie zbioru obserwacji wejścia i wyjścia systemu $\{(X_n, Y_n)\}_{n=1}^N$. W omawianym podejściu, na podstawie dostępnej informacji wstępnej, konstruuje się model „wejście–wyjście" systemu zależny od wektorowego parametru θ,

$$Y_n(X_n;\theta) = \sum_{i=1}^{m} \gamma_i \left(\frac{B\left(q^{-1}\right)}{A\left(q^{-1}\right)} X_n \right)^i$$

oraz funkcję kryterialną zbudowaną w oparciu o zbiór $\{(X_n, Y_n)\}_{n=1}^{N}$,

$$\Gamma_1(\theta) = \frac{1}{N}\sum_{n=1}^{N}[Y_n - Y_n(X_n;\theta)]^2 .$$

Metoda polega na wyznaczaniu takiego wektora θ który minimalizuje funkcję kryterialną $\Gamma_1(\cdot)$.

Omówiony algorytm jest jedną z wielu parametrycznych metod identyfikacji systemów Wienera zbudowanych w oparciu o metodę najmniejszych kwadratów, por. np. Pupeikis [97], Lacy i Bernstein [73], Schoukens *et al.* [111], Bruls *et al.* [10], Hughes i Westwick [56]. Inne podejścia do identyfikacji systemów Wienera proponowane są m.in. przez Xu *et al.* [140] i Chena *et al.* [14] gdzie zastosowano teorię uczenia sieci neuronowych do modelowania nieliniowości w systemie. Podejście częstotliwościowe, umożliwiające identyfikację przy deterministycznym pobudzeniu sinusoidalnym, omawiane jest przez Baia [1], a zastosowanie metod korelacyjnych proponowane jest przez Billingsa i Fakhouriego w [7, 8]. W [137] Wigren konstruuje rekurencyjny algorytm, oparty na metodzie błędu predykcji, umożliwiający identyfikację systemów z obiektem dynamicznym typu ARMA i odcinkami liniowymi nieliniowościami (*ang. piecewise linear*), zob. też Wigren [138]. Podejście rekurencyjne proponowane jest również przez Chena [15] do identyfikacji nieciągłych, odcinkami liniowych nieliniowości oraz przez Jacobsa [59], zob. też Vörös [124] i Hagenblad [44]. Na uwagę zasługują także metody typu *blind identification*, Taleb *et al.* [121].

Należy podkreślić, że w odróżnieniu od algorytmów nieparametrycznych, dla większości omawianych w literaturze metod parametrycznej identyfikacji systemów Wienera nie przeprowadzono teoretycznej analizy ich własności asymptotycznych.

2.2.2 Nieparametryczna identyfikacja systemów Wienera

Pierwszy nieparametryczny algorytm identyfikacji systemów Wienera opublikowany został w roku 1992 przez Greblickiego w [26]. Zaproponowana metoda, podobnie jak opracowane wcześniej nieparametryczne algorytmy identyfikacji systemów Hammersteina (Greblicki i Pawlak [38]), bazowała na zależności pomiędzy odpowiednią funkcją regresji, a nieliniową charakterystyką występującą w systemie. W pracy wykazano, że dla asymptotycznie stabilnego systemu Wienera (1.7)–(1.8) o dowolnej pamięci, funkcja regresji (wejścia względem wyjścia) o postaci

$$\eta(y) = E\{X_n | Y_n = y\}$$

jest związana z nieliniową charakterystyką $g(\cdot)$ następującą zależnością

$$\eta(y) = cg^{-1}(y), \tag{2.5}$$

gdzie c jest stałą zależną od podsystemu liniowego oraz od wariancji sygnału $\{X_n\}$. Otrzymany związek udowodniony został przy założeniu, że sygnał wejściowy $\{X_n\}$ oraz wewnętrzny sygnał zakłócający są stochastycznie niezależnymi ciągami typu *i.i.d.* o rozkładach normalnych z zerowymi wartościami oczekiwanymi. Ponadto, ze względu na występowanie w (2.5) funkcji odwrotnej do $g(\cdot)$ przyjęto, że nieliniowa charakterystyka jest funkcją ściśle monotoniczną i spełniającą warunek Lipschitza (tzn. $|g(v_1) - g(v_2)| \leq \alpha |v_1 - v_2|^{\beta}$, $\beta > 0, \alpha > 0$, i $v_1, v_2 \in \mathbf{R}$). Przedstawiony algorytm identyfikacji umożliwiał poprzez estymację regresji $\eta(\cdot)$ szacowanie odwrotności nieliniowej charakterystyki $g(\cdot)$. Do estymacji funkcji $\eta(\cdot)$ wykorzystano estymator jądrowy (por. (2.4)) zbudowany

w oparciu o N–elementowy zbiór obserwacji $\{(X_n, Y_n)\}_{n=1}^{N}$ pochodzących z systemu,

$$\hat{\eta}_N(y) = \frac{\sum_{n=1}^{N} X_n K\left(\frac{y - Y_n}{h}\right)}{\sum_{n=1}^{N} K\left(\frac{y - Y_n}{h}\right)}, \tag{2.6}$$

gdzie $K(\cdot)$ jest nieujemną, ograniczoną i lipschitzowską funkcją jądra, taką że $\int K(t)\,dt = 1$ i $K(t)\,t^{1+\varepsilon} \to 0$ gdy $|t| \to \infty$, dla dowolnego $\varepsilon > 0$ oraz h jest parametrem wygładzania spełniającym warunki $h \to 0, N^2 h \to \infty$ gdy $N \to \infty$. W pracy wykazano, że estymator $\hat{\eta}_N(\cdot)$ jest zgodnym estymatorem funkcji $cg^{-1}(\cdot)$ w każdym punkcie y, w którym gęstość prawdopodobieństwa wyjścia systemu jest dodatnia. Po przyjęciu dodatkowych założeń odnośnie jądra $K(\cdot)$ i gładkości funkcji $g(\cdot)$ określony został również rząd szybkości zbieżności estymatora

$$\hat{\eta}_N(y) = cg^{-1}(y) + O_P\left(N^{-\frac{1}{3}}\right),$$

gdy $h \sim N^{-1/6}$, w każdym punkcie y w którym gęstość wyjścia systemu jest dodatnia (zob. definicja symbolu $O_P(\cdot)$ na str. 87).

Omówiony wyżej estymator $\hat{\eta}_N(y)$ został następnie zastosowany do skonstruowania nieparametrycznego algorytmu identyfikacji odpowiedzi impulsowej podsystemu dynamicznego. Podejście to wykorzystywało obserwację, że dla rozważanego systemu Wienera

$$c\lambda_i = E\{X_0 \eta(Y_i)\}. \tag{2.7}$$

Na podstawie powyższego wzoru zaproponowano estymator zbudowany w oparciu o nowy ciąg pomiarów $\{(X'_n, Y'_n)\}_{n=1}^{N'}$ niezależnych od obserwacji wykorzystanych w konstrukcji $\hat{\eta}_N(\cdot)$. Ze względu na nieznajomość funkcji $\eta(\cdot)$ w (2.7), w jej miejsce wstawiono estymator $\hat{\eta}_N(\cdot)$. Jako estymator $c\lambda_i$ przyjęto zatem

$$\hat{\chi}_{i,N,N'} = \frac{1}{N'} \sum_{n=1}^{N'} X'_n \hat{\eta}_N\left(Y'_{i+n}\right).$$

Zależność (2.5) stanowiła również podstawę kolejnego nieparametrycznego algorytmu identyfikacji systemów Wienera, Greblicki [27]. W pracy tej zaproponowano klasę algorytmów ortogonalnych zbudowanych w oparciu o szeregi: trygonometryczny, Legendre'a i Hermite'a. Wykazano zgodność estymatorów oraz określono rzędy ich szybkości zbieżności. Otrzymane rezultaty wskazują, że dla dwukrotnie różniczkowalnych nieliniowości opisany powyżej algorytm jądrowy zbiega szybciej niż algorytmy ortogonalne. Niemniej jednak algorytmy ortogonalne cechują się lepszymi własnościami obliczeniowymi, zob. [27, str. 283].

Kolejny etap rozwoju teorii nieparametrycznej identyfikacji systemów Wienera polegał na opracowaniu metody identyfikacji, nie nakładającej *żadnych* ograniczeń na nieliniową charakterystykę w systemie[3], Greblicki [28]. Opracowana metoda stanowiła uogólnienie algorytmu z [26] i, podobnie jak omówione wyżej algorytmy, bazowała na zależności (2.5). Z tego powodu założenia odnośnie charakteru sygnału wejściowego i zakłócającego pozostały niezmienione. Nieznaną funkcję regresji $\eta(y)$ estymowano przy pomocy estymatora (2.6), przyjmując jednak nieco inne założenia odnośnie jądra $K(\cdot)$, zob. [28, str. 540]. Wykazano zgodność algorytmu w punktach y, w których nieliniowa charakterystyka $g(\cdot)$ jest odwzorowaniem różnowartościowym oraz w ich otoczeniu

[3] Nieliniowość musiała być jedynie funkcją mierzalną.

spełnia warunek Lipschitza. Zakładając dodatkowo, że w rozważanych punktach $g(\cdot)$ ma niezerową pochodną wykazano, iż estymator posiada szybkość zbieżności $O_P\left(N^{-1/4}\right)$, (dla $h \sim N^{-1/4}$). W omawianej pracy zaproponowano również nieparametryczny estymator odpowiedzi impulsowej podsystemu dynamicznego. W odróżnieniu od wcześniejszych wyników, algorytm nie wymaga stosowania estymatora funkcji regresji $\eta(\cdot)$ i opiera się na spostrzeżeniu, że w rozważanym systemie $c_1\lambda_i = E\{X_0 Y_i\}$, $i = 0, 1, \ldots$, gdzie c_1 jest pewną stałą. Jako estymator elementu $c_1\lambda_i$ przyjęto więc

$$\tilde{\chi}_{i,N} = \frac{1}{N}\sum_{n=1}^{N} X_n Y_{i+n}. \tag{2.8}$$

Wykazano, że dla dowolnej nieliniowości $g(\cdot)$ spełniającej warunek $E|V_1 g(V_1)| < \infty$, gdzie $\{V_n\}$ jest sygnałem interakcyjnym w systemie, powyższy estymator jest zgodnym estymatorem elementu $c_1\lambda_i$, oraz że w przypadku gdy nieliniowość spełnia warunek Lipschitza, estymator posiada parametryczną szybkość zbieżności, tj.

$$c_1\lambda_i = \tilde{\chi}_{i,N} + O_P\left(N^{-\frac{1}{2}}\right), \; i = 0, 1, 2, \ldots$$

Oprócz omówionych powyżej algorytmów umożliwiających identyfikację przy bardzo łagodnych założeniach o postaci systemu, podjęto również próby opracowania metod nieparametrycznych, pozwalających na identyfikację w przypadku sygnałów wejściowych nie będących ciągami typu *i.i.d.* (ale w dalszym ciągu o charakterze gaussowskim), Greblicki [33]. Teorię nieparametrycznej identyfikacji systemów Wienera rozwijano też w innych kierunkach. W [30] Greblicki zaproponował rekurencyjną wersję algorytmu, a Wachel [125,126] przedstawił algorytm ortogonalny, zbudowany w oparciu o rozwinięcia falkowe.

2.3 Ogólna charakterystyka proponowanych metod identyfikacji

Zarówno parametryczne jak i nieparametryczne metody identyfikacji systemów Wienera posiadają ograniczenia, wynikające ze specyficznej konstrukcji rozważanej klasy systemów. Omawiane zadanie identyfikacji jest stosunkowo trudne i zazwyczaj wymaga przyjęcia silnych założeń odnośnie samego systemu bądź też sygnałów wejściowego i zakłócającego. W praktyce, opracowane dotychczas algorytmy parametryczne wymagają dużej wiedzy o nieliniowości $g(\cdot)$ oraz występowania gaussowskiego sygnału wejściowego. Jeżeli informacja o nieliniowości jest zbyt mała, możliwe jest zastosowanie algorytmów nieparametrycznych. Niestety, w przypadku gdy sygnał wejściowy nie jest sygnałem gaussowskim, większość metod parametrycznych oraz metody nieparametryczne nie mogą być stosowane. Z tego powodu szczególnie istotne wydaje się opracowanie algorytmów umożliwiających identyfikację w sytuacjach, gdy dostępna wiedza aprioryczna nie wystarcza do zastosowania algorytmów czysto parametrycznych oraz sygnał wejściowy systemu nie jest sygnałem gaussowskim. Prezentowane w pracy podejście parametryczno–nieparametryczne ma umożliwić identyfikację w przypadku, gdy omówione powyżej parametryczne i nieparametryczne metody identyfikacji nie mogą być stosowane wprost. Koniecznym warunkiem jest jednak skończoność i znajomość długości pamięci systemu, co jak wynika z rozważań w p. 2.2.1 nie jest założeniem nietypowym, podczas gdy informacja o nieliniowości w systemie może być bardzo mała. Zakłada się zatem, że dostępna informacja aprioryczna o podsystemie dynamicznym ma charakter parametryczny, a wiedza o nieliniowości jest nieparametryczna.

Zanim przedstawiona zostanie ogólna koncepcja proponowanych metod warto podkreślić, że ze względu na kaskadową budowę systemów Wienera i nieznajomość obu podsystemów składowych, identyfikacja może być przeprowadzona tylko z dokładnością do nieznanej multiplikatywnej stałej $c_0 \neq 0$. Łatwo bowiem zauważyć, że dla dowolnego sygnału wejściowego, wyjście systemu opisanego równaniami (1.9)–(1.10) będzie identyczne z wyjściem systemu o odpowiedzi impulsowej $\{c_0\lambda_i\}_{i=0}^{p}$ i nieliniowości $g(\cdot/c_0)$. Własność powyższa jest niezależna od stosowanej metody identyfikacji i wynika z budowy systemu oraz charakteru posiadanej informacji pomiarowej, a w szczególności z faktu, że wewnętrzny sygnał interakcyjny nie jest dostępny.

2.3.1 Idea dwuetapowej parametryczno–nieparametrycznej identyfikacji systemów Wienera

Proponowane w pracy parametryczno–nieparametryczne metody identyfikacji systemów Wienera konstruowane są z wykorzystaniem różnorodnych narzędzi statystyki matematycznej, poczynając od metod korelacyjnych, poprzez metodę najmniejszych kwadratów i estymację korelacji rang, a kończąc na zastosowaniu jądrowego estymatora funkcji regresji oraz tzw. estymatora uśrednionej pochodnej (*ang. average derivative estimate*). Pomimo wielu istotnych różnic, proponowane algorytmy zbudowane są w oparciu o wspólny szkielet, którego geneza wywodzi się z opracowanej przez Hasiewicza i Mzyka parametryczno–nieparametrycznej metody identyfikacji systemów Hammersteina [52]. Należy przy tym podkreślić, że:

Uwaga 2.1 *W systemach Hammersteina (pobudzanych sygnałami typu i.i.d.) zarówno wejście jak i wyjście nieliniowego podsystemu statycznego jest sygnałem typu i.i.d.; w przypadku systemów Wienera natomiast (bez względu na fakt, czy wejście systemu jest ciągiem typu i.i.d.), wejście i wyjście nieliniowego podsystemu statycznego jest sygnałem skorelowanym, a więc nie jest ciągiem typu i.i.d.*

Uwaga 2.2 *W systemach Hammersteina przekształceniu nieliniowemu poddawany jest dostępny dla obserwacji sygnał wejściowy systemu. W systemach Wienera natomiast, argumentem (wejściem) nieliniowości jest nieznany i niedostępny dla pomiarów sygnał interakcyjny.*

Zawarte w uwagach 2.1 i 2.2 obserwacje prowadzą do wniosku, że zadanie identyfikacji systemów Wienera jest zazwyczaj znacznie bardziej skomplikowane niż zadanie identyfikacji systemów Hammersteina.

Podstawą proponowanych w pracy metod jest elementarne spostrzeżenie, że przy znajomości podsystemu dynamicznego, niedostępny dla pomiarów sygnał interakcyjny $\{V_n\}$ może być wyznaczony przy użyciu zgromadzonych obserwacji wejścia $\{X_n\}$ całego systemu. Na podstawie zbioru par obserwacji (V_n, Y_n), nieliniowa charakterystyka $g(\cdot)$ może być następnie estymowana metodami nieparametrycznymi. Dla systemu Wienera z addytywnym szumem wyjściowym niezależnym od wejścia, zachodzi bowiem związek[4]

$$g(v) = E\{Y_1 | V_1 = v\}. \tag{2.9}$$

W rzeczywistości, opracowane metody identyfikacji nie wymagają tak dużej informacji wstępnej dotyczącej części dynamicznej systemu. Wykorzystując pomiary wejścia i wyjścia całego systemu można bowiem skonstruować zgodne estymatory parametrów podsystemu

[4] Dodatkowo wartość oczekiwana zakłócenia musi być równa zero.

dynamicznego i, w konsekwencji, szacować nieznany sygnał $\{V_n\}$. Schemat proponowanych dwuetapowych algorytmów identyfikacji daje się przedstawić następująco:

Etap I

Parametryczna identyfikacja odpowiedzi impulsowej $\{\lambda_i\}_{i=0}^{p}$ podsystemu dynamicznego na podstawie dostępnych obserwacji wejścia i wyjścia całego systemu.

Etap IIa

Estymacja sygnału interakcyjnego $\{V_n\}$ na podstawie dostępnych obserwacji wejścia oraz uzyskanego w etapie I estymatora odpowiedzi impulsowej.

Etap IIb

Nieparametryczna identyfikacja charakterystyki nieliniowej $g(\cdot)$ z wykorzystaniem uzyskanego w etapie IIa estymatora sygnału $\{V_n\}$ i dostępnych obserwacji wyjścia systemu.

2.4 Teza pracy

Zastosowanie parametryczno–nieparametrycznych metod identyfikacji systemów Wienera skonstruowanych w pracy, pozwala na identyfikację liniowych i nieliniowych podsystemów składowych, przy założeniu parametrycznej wiedzy o podsystemie dynamicznym i nieparametrycznej o podsystemie statycznym. Algorytmy parametryczno–nieparametryczne umożliwiają identyfikację w sytuacji, gdy losowy sygnał wejściowy nie jest sygnałem gaussowskim, a szum zakłócający działa na wyjście systemu. Ponadto algorytmy nie wymagają odwracalności charakterystyki nieliniowej systemu i umożliwiają estymację nieliniowości wprost, tj. nie poprzez funkcję odwrotną, łagodząc w ten sposób typowe wymagania wstępne i rozszerzając zakres stosowalności metod.

2.5 Cele pracy

Celami pracy są:

- Opracowanie algorytmów identyfikacji systemów Wienera, umożliwiających identyfikację w przypadku gaussowskich i niegaussowskich sygnałów wejściowych oraz odwracalnych i nieodwracalnych charakterystyk nieliniowych.
- Przeprowadzenie analizy teoretycznej algorytmów dotyczącej ich własności asymptotycznych.
- Wykonanie badań eksperymentalnych proponowanych metod, celem ilustracji zachowania algorytmów dla małej i średniej liczby obserwacji.

2.6 Podstawowe założenia oraz klasyfikacja zadań

W pracy rozważana jest identyfikacja systemów Wienera opisanych równaniami

$$V_n = \sum_{i=0}^{p} \lambda_i X_{n-i}, \tag{2.10}$$

$$Y_n = g(V_n) + Z_n, \tag{2.11}$$

gdzie $\{X_n\}$ i $\{Y_n\}$ oznaczają odpowiednio sygnał wejściowy i wyjściowy, $\{Z_n\}$ jest addytywnym szumem na wyjściu systemu oraz $\{V_n\}$ jest sygnałem wewnętrznym reprezen-

tującym wyjście obiektu dynamicznego i zarazem wejście obiektu statycznego[5]. Skończony $(p+1)$–elementowy ciąg $\{\lambda_i\}_{i=0}^{p}$ jest odpowiedzą impulsową podsystemu dynamicznego. Funkcja $g(\cdot)$ reprezentuje nieliniową charakterystykę w systemie.

2.6.1 Założenia ogólne

Odnośnie klasy identyfikowanych systemów Wienera przyjmuje się następujące założenia ogólne, wspólne dla wszystkich rozważanych w pracy metod identyfikacji:

Założenie 2.1 *Długość p odpowiedzi impulsowej podsystemu dynamicznego jest skończona i znana. Ponadto $\lambda_0 \neq 0$.*

Założenie 2.2 *Sygnał wejściowy $\{X_n\}$ jest ciągiem i.i.d. zmiennych losowych.*

Założenie 2.3 *Addytywny szum $\{Z_n\}$ zakłócający wyjście systemu jest ciągiem i.i.d. zmiennych losowych o zerowej wartości oczekiwanej i wariancji $\sigma_Z^2 < \infty$ oraz $\{Z_n\}$ i $\{X_n\}$ są stochastycznie niezależne.*

Założenie 2.4 *Identyfikowany system znajduje się w stanie ustalonym.*

2.6.2 Klasyfikacja zadań

Z przedstawionych w p. 2.5 celów pracy wynikają następujące zadania:

- Opracowanie metody identyfikacji dla przypadku gaussowskiego sygnału wejściowego oraz dowolnej charakterystyki nieliniowej systemu.
- Opracowanie algorytmów identyfikacji dla przypadku niegaussowskich sygnałów wejściowych, dodatkowo przy braku odwracalności charakterystyki nieliniowej.
- Analiza asymptotyczna algorytmów.
- Wykonanie badań symulacyjnych proponowanych metod identyfikacji.

[5] W rozważanym zadaniu identyfikacji sygnały $\{X_n\}, \{V_n\}, \{Y_n\}$ i $\{Z_n\}$ są ciągami zmiennych losowych.

Rozdział 3

Identyfikacja systemów Wienera z zastosowaniem metody korelacyjnej

W tym rozdziale prezentowana jest metoda identyfikacji systemów Wienera z gaussowskim sygnałem wejściowym. Omawiany algorytm bazuje na podejściu zaproponowanym przez Greblickiego [28], w którym do nieparametrycznej estymacji odpowiedzi impulsowej podsystemu dynamicznego wykorzystano korelacyjną metodę identyfikacji (zob. wzór (2.8)). Proponowany tutaj algorytm umożliwia identyfikację systemów z addytywnym szumem na wyjściu, pobudzanych gaussowskim sygnałem wejściowym o dowolnej (nieznanej) wartości średniej (zakłada się jednak, że w systemie nie występuje wewnętrzny sygnał zakłócający). W przeciwieństwie do [28], przedstawiona metoda umożliwia identyfikację nieliniowej charakterystyki *wprost*, tj. nie poprzez funkcję odwrotną.

Wprowadzone w rozdziale estymatory części liniowej, sygnału interakcyjnego i części nieliniowej systemu oznaczane będą indeksem „*Cr*" (od *ang. Correlation*).

3.1 Sformułowanie problemu

Rozważane zadanie identyfikacji dotyczy klasy systemów Wienera opisywanych równaniami (2.10)–(2.11), dla których spełnione są założenia (2.1)–(2.4). Dodatkowo zakładamy, że:

Założenie 3.1 *Sygnał wejściowy* $\{X_n\}$ *ma rozkład normalny* $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ *z nieznaną wartością oczekiwaną* μ_X *i wariancją* σ_X^2.

Założenie 3.2 *Nieliniowa charakterystyka* $g(\cdot)$ *ograniczona jest przez wielomian, tj.* $|g(v)| \leq m(v), \forall v \in \mathbf{R}$, *gdzie* $m(\cdot)$ *jest wielomianem dowolnego skończonego stopnia.*

Założenie 3.3 *Rozważana klasa systemów zawiera tylko systemy, dla których* $Cov\{X_1, Y_1\} \neq 0$.

3.2 Identyfikacja podsystemu dynamicznego metodą korelacyjną

Na mocy założenia 3.1 oraz twierdzenia B.1 w dodatku B.2 zauważmy najpierw, że sygnał $\{V_n\}$ jest ciągiem zmiennych losowych o rozkładzie normalnym z wartością oczekiwaną

i wariancją określonymi wzorami

$$\mu_V = \mu_X \sum_{i=0}^{p} \lambda_i \qquad \text{oraz} \qquad \sigma_V^2 = \sigma_X^2 \sum_{i=0}^{p} \lambda_i^2. \tag{3.1}$$

Zatem, zgodnie z twierdzeniem B.2 (zob. dodatek B.2, wzór (B.3)), dla dowolnego $i \in \{0, 1, \ldots, p\}$ prawdziwa jest następująca zależność

$$E\{X_0|V_i\} = \mu_X + \rho\frac{\sigma_X}{\sigma_V}(V_i - \mu_V), \tag{3.2}$$

gdzie ρ jest współczynnikiem korelacji zmiennych losowych X_0 i V_i, który wyznaczymy znajdując wcześniej kowariancję

$$Cov\{X_0, V_i\} = E\{(\mu_X - X_0)(\mu_V - V_i)\} = E\{X_0 V_i\} - \mu_X \mu_V.$$

Wykorzystując wzory (3.1) oraz definicję sygnału interakcyjnego $\{V_n\}$ (zob. (2.10)), otrzymujemy

$$\begin{aligned} E\{X_0 V_i\} - \mu_X \mu_V &= \lambda_i E\{X_0^2\} + \mu_X^2 \sum_{j=0, j\neq i}^{p} \lambda_j - \mu_X \mu_V \\ &= \lambda_i \sigma_X^2 + \lambda_i \mu_X^2 + \mu_X^2 \sum_{j=0, j\neq i}^{p} \lambda_j - \mu_X^2 \sum_{j=0}^{p} \lambda_j. \end{aligned}$$

Ostatecznie więc $Cov\{X_0, V_i\} = \lambda_i \sigma_X^2$, i wobec tego

$$\rho = corr\{X_0, V_i\} = \lambda_i \frac{\sigma_X}{\sigma_V}.$$

Wyrażenie (3.2) można teraz zapisać w postaci

$$E\{X_0|V_i\} = \mu_X + \lambda_i \frac{\sigma_X^2}{\sigma_V^2}(V_i - \mu_V). \tag{3.3}$$

Zauważmy dalej, że

$$\begin{aligned} E\{X_0 Y_i\} &= E\{X_0 g(V_i)\} = E\{E\{X_0 g(V_i)|V_i\}\} \\ &= E\{g(V_i) E\{X_0|V_i\}\} = E\{Y_i E\{X_0|V_i\}\}. \end{aligned} \tag{3.4}$$

Wstawiając (3.3) do (3.4) otrzymujemy kluczowe dla dalszych rozważań równanie

$$\alpha\lambda_i = E\{X_0 Y_i\} - \beta, \tag{3.5}$$

w którym stałe α oraz β dane są wzorami

$$\alpha = \frac{\sigma_X^2}{\sigma_V^2}[E\{Y_1 V_1\} - \mu_V \mu_Y] \qquad \text{oraz} \qquad \beta = \mu_X \mu_Y, \tag{3.6}$$

natomiast μ_Y jest wartością oczekiwaną wyjścia systemu, tj. $\mu_Y = E\{Y_1\}$.

Równanie (3.5) stanowiące podstawę omawianego algorytmu identyfikacji podsystemu dynamicznego jest uogólnieniem związku wprowadzonego w [28] na przypadek sygnałów wejściowych o niezerowej wartości oczekiwanej (tj. $\mu_X \neq 0$). Jak łatwo zauważyć, dla $\mu_X = 0$ stała β jest równa zero oraz $\alpha = \sigma_X^2 \backslash \sigma_V^2 E\{Y_i V_i\}$, co odpowiada sytuacji przedstawionej w [28].

Wykorzystując powyższe rezultaty możemy przeprowadzić konstrukcję estymatora odpowiedzi impulsowej podsystemu dynamicznego.

3.2.1 Konstrukcja estymatora podsystemu dynamicznego

Zgodnie z obserwacją w p. 2.3, zarówno podsystem dynamiczny jak i element statyczny systemu Wienera mogą być zidentyfikowane tylko z dokładnością do multiplikatywnej stałej $c_0 \neq 0$. Z równania (3.5) wynika (zob. założenie 2.1 na str. 22) że $\alpha = \lambda_0^{-1}[E\{X_0Y_0\} - \beta]$, co z kolei prowadzi do wniosku, że $\alpha \neq 0$ wtedy i tylko wtedy, gdy $Cov\{X_0, Y_0\} \neq 0$. Dla rozważanej klasy systemów warunek ten jest spełniony na mocy założenia 3.3, a w celu identyfikacji odpowiedzi impulsowej obiektu dynamicznego wystarczy estymować prawą stronę zależności (3.5).

Oznaczymy $\gamma_i = E\{X_0Y_i\}$. Naturalnym estymatorem wartości γ_i jest następująca średnia z próby, konstruowana w oparciu o zbiór obserwacji $\{(X_n, Y_n)\}_{n=1}^N$,

$$\hat{\gamma}_{i,N} = \frac{1}{N-i}\sum_{n=1}^{N-i} X_nY_{i+n}. \tag{3.7}$$

Drugim, koniecznym do oszacowania elementem równania (3.5), jest stała β. Ponieważ jest ona iloczynem wartości oczekiwanych wejścia i wyjścia systemu, więc odpowiedni estymator przyjmie następującą, naturalną, postać

$$\hat{\beta}_N = \hat{\mu}_{X,N} \times \hat{\mu}_{Y,N}, \tag{3.8}$$

gdzie $\hat{\mu}_{X,N}$, $\hat{\mu}_{Y,N}$ są estymatorami wartości oczekiwanych odpowiednio wejścia i wyjścia systemu, tj.

$$\hat{\mu}_{X,N} = \frac{1}{N}\sum_{n=1}^{N} X_n \qquad \text{oraz} \qquad \hat{\mu}_{Y,N} = \frac{1}{N}\sum_{n=1}^{N} Y_n.$$

W rezultacie, do oszacowania elementów przeskalowanej odpowiedzi impulsowej $\alpha\lambda_i, i = 0, 1, \ldots, p$, otrzymujemy następujący zbiór estymatorów

$$\hat{\lambda}_{i,N}^{Cr} = \hat{\gamma}_{i,N} - \hat{\beta}_N, \qquad i = 0, 1, \ldots, p, \tag{3.9}$$

których własności asymptotyczne oraz zachowanie dla małej i średniej liczby obserwacji omówione zostaną odpowiednio w p. 3.5.1 i 3.7.

3.3 Estymacja sygnału interakcyjnego

Zauważmy, że odpowiedni estymator sygnału interakcyjnego $\{V_n\}$ można łatwo uzyskać wstawiając we wzorze (2.10) otrzymane wyżej estymatory $\{\hat{\lambda}_{i,N}^{Cr}\}_{i=0}^p$ w miejsce nieznanych, prawdziwych współczynników odpowiedzi impulsowej $\{\lambda_i\}_{i=0}^p$, tj.

$$\hat{V}_{n,N}^{Cr} = \sum_{i=0}^{p} \hat{\lambda}_{i,N}^{Cr} X_{n-i}. \tag{3.10}$$

Jednak, ze względu na ograniczoną liczbę dostępnych danych pomiarowych (N–par obserwacji wejścia i wyjścia systemu, $\{(X_n, Y_n)\}_{n=1}^N$), estymator $\hat{V}_{n,N}^{Cr}$ może być wyznaczony tylko dla $n \in I$, gdzie I jest zbiorem indeksów

$$I = \{p+1, p+2, \ldots, N\}. \tag{3.11}$$

Z kolei występowanie multiplikatywnej stałej α w równaniu (3.5) oraz jej wpływ na estymatory $\{\hat{\lambda}_{i,N}^{Cr}\}_{i=0}^p$ powoduje, że $\hat{V}_{n,N}^{Cr}$ jest w rzeczywistości estymatorem przeskalowanego sygnału interakcyjnego $\{\alpha V_n\}$. Fakt ten udowodniony zostanie w p. 3.5.2.

3.4 Nieparametryczna identyfikacja nieliniowości

Konstruowanie nieparametrycznego estymatora charakterystyki nieliniowej rozpoczniemy od przytoczenia wzoru (2.9) (zob. str. 20), tj. spostrzeżenia, że funkcja $g(\cdot)$ jest równa funkcji regresji wyjścia systemu względem wartości sygnału interakcyjnego $\{V_n\}$, tj.

$$g(v) = E\{Y_1 | V_1 = v\}. \tag{3.12}$$

Związek (3.12) prowadzi więc do wniosku, że nieznaną charakterystykę $g(\cdot)$ można przybliżać poprzez estymację warunkowej wartości oczekiwanej $E\{Y_1|V_1 = v\}$. Zakładając chwilowo, że sygnał interakcyjny $\{V_n\}$ może być mierzony i stosując estymator jądrowy rozważanej funkcji regresji zbudowany w oparciu o zbiór $\{(V_n, Y_n)\}_{n=1}^N$, otrzymujemy następujące oszacowanie

$$\tilde{g}_N(v) = \frac{\sum_{n=1}^{N} Y_n K\left(\frac{v - V_n}{h}\right)}{\sum_{n=1}^{N} K\left(\frac{v - V_n}{h}\right)}, \tag{3.13}$$

gdzie $K(\cdot)$ i $h = h(N)$ są odpowiednio wybraną funkcją jądra i parametrem wygładzania zależnym od liczby obserwacji N (por. p. 2.1.2).

Niestety, ze względu na brak możliwości pomiaru sygnału interakcyjnego $\{V_n\}$, bezpośrednie zastosowanie estymatora (3.13) nie jest możliwe. Wykorzystując oszacowanie (3.10) możemy jednak tak zmodyfikować estymator $\tilde{g}_N(\cdot)$, aby był on zbudowany w oparciu o zbiór $\{(\hat{V}_{n,N}^{Cr}, Y_n)\}_{n\in I}$, gdzie I jest zbiorem indeksów zdefiniowanym w (3.11). Wystarczy w tym celu zastosować podstawienie $\hat{V}_{n,N}^{Cr}$ w miejsce V_n. W efekcie otrzymujemy następujący estymator nieznanej charakterystyki nieliniowej $g(\cdot)$

$$\hat{g}_N^{Cr}(v) = \frac{\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{Cr}}{h}\right)}{\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{Cr}}{h}\right)}. \tag{3.14}$$

Zauważmy, że w przeciwieństwie do $\tilde{g}_N(\cdot)$ wyznaczenie wartości estymatora $\hat{g}_N^{Cr}(\cdot)$ jest możliwe na podstawie dostępnych par obserwacji wejścia i wyjścia całego systemu. Jednak, jak pokazano w p. 3.5.3, ze względu na występowanie stałej α w równaniu (3.5), estymator $\hat{g}_N^{Cr}(\cdot)$ (przy pewnych ogólnych założeniach o funkcji $K(\cdot)$ i parametrze wygładzania h) jest zgodnym estymatorem *przeskalowanej* nieliniowości $g(\cdot/\alpha)$.

3.5 Własności asymptotyczne algorytmu

Analizę własności asymptotycznych omawianego algorytmu identyfikacji rozpoczniemy od określenia asymptotyki estymatorów $\{\hat{\lambda}_{i,N}^{Cr}\}_{i=0}^p$ odpowiedzi impulsowej $\{\alpha\lambda_i\}_{i=0}^p$ oraz estymatora $\hat{V}_{n,N}^{Cr}$ sygnału interakcyjnego $\{V_n\}$. Następnie zbadamy asymptotyczne zachowanie estymatora $\hat{g}_N^{Cr}(\cdot)$. Sformułowane zostaną założenia o funkcji jądra $K(\cdot)$ i parametrze wygładzenia h, przy których $\hat{g}_N^{Cr}(\cdot)$ jest zgodnym estymatorem funkcji $g(\cdot/\alpha)$. Ponadto, ograniczając klasę dopuszczalnych nieliniowości do klasy funkcji dwukrotnie różniczkowalnych, określony zostanie rząd szybkości zbieżności estymatora $\hat{g}_N^{Cr}(\cdot)$.

3.5.1 Analiza zbieżności estymatora odpowiedzi impulsowej podsystemu dynamicznego

W celu wykazania zbieżności oraz określenia rzędu szybkości zbieżności estymatorów $\hat{\lambda}_{i,N}^{Cr}$, $i = 0, \ldots, p$ zbadane zostaną własności asymptotyczne estymatorów składowych $\hat{\gamma}_{i,N}$, $i = 0, \ldots, p$ oraz $\hat{\alpha}_N$ (zob. wzory (3.7) i (3.8)).
Zgodnie z przyjętymi założeniami o wejściu $\{X_n\}$ i szumie $\{Z_n\}$, sygnały $\{X_n\}$ i $\{Y_n\}$ w stanie ustalonym są stacjonarne. Wynika stąd natychmiast, że

$$E\{\hat{\gamma}_{i,N}\} = \gamma_i. \tag{3.15}$$

Wyznaczmy teraz oszacowanie wariancji

$$Var\{\hat{\gamma}_{i,N}\} = \frac{1}{(N-i)^2} Var\left\{\sum_{n=1}^{N-i} X_n Y_{i+n}\right\}. \tag{3.16}$$

Zauważmy w tym celu, że

$$Var\{X_n Y_{i+n}\} \leq E\{X_n^2 g^2(V_{i+n})\} + \sigma_Z^2 E\{X_n^2\},$$

co na mocy założenia 3.2 daje

$$Var\{X_n Y_{i+n}\} \leq E\{X_n^2 m^2(V_{i+n})\} + \sigma_Z^2 E\{X_n^2\}. \tag{3.17}$$

Oszacowanie (3.17), ze względu na skończoność wszystkich momentów gaussowskich zmiennych losowych, prowadzi do wniosku, że $Var\{X_n Y_{i+n}\} < \infty$ dla każdego $i \in \{0, 1, \ldots, p\}$, $1 \leq n \leq N - i$.
Ze względu na skończoną długość pamięci systemu oraz na podstawie założenia 2.3 o zakłóceniu wyjściowym, prawdziwa jest następująca własność:

Własność 3.1 *Zmienne losowe $X_n Y_{i+n}$ oraz $X_k Y_{i+k}$, $(i \in \{0, 1, \ldots, p\})$ są niezależne, gdy $|k - n| > p$.*

Wobec powyższego, na mocy Lematu A.1 w dodatku A.1, otrzymujemy następujące oszacowanie wariancji estymatora $\hat{\gamma}_{i,N}$,

$$Var\{\hat{\gamma}_{i,N}\} \leq \frac{p+1}{(N-i)^2} \sum_{n=1}^{N-i} Var\{X_n Y_{i+n}\} = \frac{p+1}{(N-i)} Var\{X_0 Y_i\}.$$

Ostatecznie więc

$$Var\{\hat{\gamma}_{i,N}\} = O\left(\frac{1}{N-i}\right) = O\left(N^{-1}\right). \tag{3.18}$$

Rezultaty (3.15) i (3.18) prowadzą do wniosku, że dla każdego $i \in \{0, 1, \ldots, p\}$,

$$\hat{\gamma}_{i,N} = \gamma_i + O_P\left(N^{-1/2}\right). \tag{3.19}$$

Zbadamy obecnie asymptotyczne zachowanie estymatora $\hat{\beta}_N$. Zgodnie z przyjętymi założeniami zauważamy, że

$$\hat{\mu}_{X,N} = \mu_X + O_P\left(N^{-1/2}\right) \quad \text{oraz} \quad \hat{\mu}_{Y,N} = \mu_Y + O_P\left(N^{-1/2}\right),$$

skąd, na mocy lematu B.2 w dodatku B.1, otrzymujemy

$$\hat{\beta}_N = \mu_X \mu_Y + O_P\left(N^{-1/2}\right). \tag{3.20}$$

Uzyskane powyżej wyniki dowodzą prawdziwości następującego twierdzenia:

Twierdzenie 3.1 *Jeżeli dla systemu Wienera (2.10)–(2.11) spełnione są założenia 2.1–2.4 oraz 3.1–3.3, to dla każdego $i \in \{0,1,\ldots,p\}$ estymator $\hat{\lambda}_{i,N}^{Cr}$ jest zgodnym estymatorem przeskalowanego elementu odpowiedzi impulsowej $\alpha\lambda_i$. Ponadto*

$$\hat{\lambda}_{i,N}^{Cr} = \alpha\lambda_i + O_P\left(N^{-1/2}\right).$$

3.5.2 Analiza zbieżności estymatora sygnału interakcyjnego

Udowodnimy następujący lemat:

Lemat 3.1 *Jeżeli spełnione są założenia twierdzenia 3.1, to dla każdego $n \in I$ estymator $\hat{V}_{n,N}^{Cr}$ jest zgodnym estymatorem αV_n. Ponadto*

$$\hat{V}_{n,N}^{Cr} = \alpha V_n + O_P\left(N^{-1/2}\right). \tag{3.21}$$

Dowód. Zgodnie z definicją (2.10) sygnału $\{V_n\}$ oraz definicją (3.10) estymatora $\hat{V}_{n,N}^{Cr}$ otrzymujemy

$$\hat{V}_{n,N}^{Cr} - \alpha V_n = \sum_{i=0}^{p} X_{n-i}\left(\hat{\lambda}_{i,N}^{Cr} - \alpha\lambda_i\right). \tag{3.22}$$

Rezultat (3.21) wynika zatem bezpośrednio z (3.22) i twierdzenia 3.1. ■

W celu wykazania zbieżności estymatora charakterystyki statycznej części nieliniowej, oprócz wyniku (3.21), niezbędne jest również określenie asymptotycznej szybkości zbieżności jednostajnej estymatorów $\{\hat{V}_{n,N}^{Cr}\}_{n\in I}$, to jest zbadanie zachowania się wyrażenia

$$\sup_{n\in I}\left|\hat{V}_{n,N}^{Cr} - \alpha V_n\right|, \tag{3.23}$$

gdy $N \to \infty$[1]. W tym celu udowodnimy najpierw następujący lemat pomocniczy:

Lemat 3.2 *Niech $\{\xi_n\}_{n=1}^{N}$ będzie ciągiem zmiennych losowych typu i.i.d. o rozkładzie gaussowskim $\mathcal{N}\left(\mu_\xi, \sigma_\xi^2\right)$. Wtedy, dla dowolnie małego $\varepsilon > 0$, zachodzi*

$$\frac{1}{\sqrt{N}} \sup_{1\le n\le N} |\xi_n| = o_P\left(N^{-1/2+\varepsilon}\right). \tag{3.24}$$

Dowód lematu zamieszczony jest w dodatku A.1.

Uzyskany powyżej wynik pozwala obecnie na określenie asymptotyki wyrażenia (3.23). Lemat 3.3 podaje odpowiednie oszacowanie.

Lemat 3.3 *Jeżeli spełnione są założenia twierdzenia 3.1, to*

$$\sup_{n\in I}\left|\hat{V}_{n,N}^{Cr} - \alpha V_n\right| = o_P\left(N^{-\frac{1}{2}+\varepsilon}\right),$$

dla dowolnie małego $\varepsilon > 0$.

Dowód. Niech $\varrho_N(\omega)$ będzie ciągiem zmiennych losowych, takim że $\varrho_N(\omega) = \sup_{1\le n\le N}|X_n(\omega)|$ (lub krócej $\varrho_N = \sup_{1\le n\le N}|X_n|$). Zgodnie z (3.22) otrzymujemy

$$\sup_{n\in I}\left|\hat{V}_{n,N}^{Cr} - \alpha V_n\right| \le \varrho_N \sum_{i=0}^{p}\left|\hat{\lambda}_{i,N}^{Cr} - \alpha\lambda_i\right|.$$

Teza twierdzenia wynika więc bezpośrednio z powyższego oszacowania oraz twierdzenia 3.1 i lematu 3.2. ■

[1] chodzi o zbadanie asymptotycznego zachowania ciągu zmiennych losowych $\nu_N(\omega) = \sup_{n\in I}|\hat{V}_{n,N}^{Cr}(\omega) - \alpha V_n(\omega)|$.

3.5.3 Analiza zbieżności estymatora charakterystyki nieliniowej

Podamy obecnie warunki, jakie muszą spełniać jądro $K(\cdot)$ oraz parametr wygładzania h występujące we wzorze (3.14), aby estymator $\hat{g}_N^{Cr}(\cdot)$ był zgodnym estymatorem przeskalowanej charakterystyki nieliniowej $g(\cdot/\alpha)$. Ponadto, przyjmując dodatkowe założenia o gładkości nieliniowości, zbadamy rząd szybkości zbieżności estymatora (3.14).

Zgodność estymatora charakterystyki nieliniowej

Przyjmiemy następujące założenia odnośnie jądra $K(\cdot)$ i parametru wygładzania h:

Założenie 3.4 *Jądro $K(\cdot)$ jest nieujemną funkcją lipschitzowską (tj. $\exists L < \infty$, $\forall v_1, v_2 \in \mathbf{R}$, $|K(v_1) - K(v_2)| \leq L|v_1 - v_2|$), dla której spełnione są następujące warunki:* $\sup_{v\in\mathbf{R}} K(v) = k_0 < \infty$, $\int_{-\infty}^{\infty} K(v)\,dv = 1$ *oraz* $\lim_{|v|\to\infty} vK(v) = 0$.

Założenie 3.5 *Ciąg liczbowy $h = h(N)$ jest rzędu $N^{-1/5}$, tj. $h \sim N^{-1/5}$.*

Poniższe twierdzenie rozstrzyga o zgodności estymatora $\hat{g}_N^{Cr}(\cdot)$.

Twierdzenie 3.2 *Jeżeli spełnione są założenia 2.1–2.4 oraz 3.1–3.5, to dla systemu Wienera (2.10)–(2.11) estymator $\hat{g}_N^{Cr}(v)$ jest zgodnym estymatorem funkcji $g(v/\alpha)$ w każdym punkcie ciągłości $g(v/\alpha)$, tzn.*

$$\hat{g}_N^{Cr}(v) \to g(v/\alpha), \qquad gdy\ N \to \infty, \tag{3.25}$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą.

Dowód. Przyjmiemy następujące oznaczenia

$$\hat{r}_N^{Cr}(v) = \frac{1}{N_I h}\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{Cr}}{h}\right) \quad \text{oraz} \quad \hat{f}_N^{Cr}(v) = \frac{1}{N_I h}\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{Cr}}{h}\right).$$

gdzie N_I oznacza liczbę elementów zbioru I ($N_I = \#\{I\} = N-p$). Dodatkowo oznaczymy $r(v) = \alpha^{-1} g(v/\alpha) f_V(v/\alpha)$, gdzie $f_V(\cdot)$ jest gęstością prawdopodobieństwa sygnału interakcyjnego $\{V_n\}$. Zbieżność w (3.25) zachodzi jeżeli $\hat{r}_N^{Cr}(v)$ i $\hat{f}_N^{Cr}(v)$ zbiegają według prawdopodobieństwa odpowiednio do $r(v)$ i $\alpha^{-1} f_V(v/\alpha)$ przy $N \to \infty$ w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą. W dowodzie wykorzystamy następującą dekompozycję różnicy $\left[\hat{r}_N^{Cr}(v) - r(v)\right]$,

$$\hat{r}_N^{Cr}(v) - r(v) = \left[\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)\right] + \left[\tilde{r}_N(v) - r(v)\right], \tag{3.26}$$

gdzie

$$\tilde{r}_N(v) = \frac{1}{N_I h}\sum_{n\in I} Y_n K\left(\frac{v - \alpha V_n}{h}\right). \tag{3.27}$$

Zgodnie z (3.26) analizę wyrażenia $\left[\hat{r}_N^{Cr}(v) - r(v)\right]$ przeprowadzimy w dwóch etapach. Rozpoczniemy od zbadania zbieżności różnicy $[\tilde{r}_N(v) - r(v)]$.

Teoria estymatorów jądrowych dla ciągów zmiennych losowych typu *i.i.d.* jest dobrze ugruntowana i szeroko opisana w literaturze (zob. dyskusja w p. 2.1.2). Jednak w rozważanym przypadku ciąg $\{V_n\}_{n\in I}$ nie jest ciągiem typu *i.i.d.*, co powoduje że klasyczne rezultaty dotyczące zbieżności nie mogą być bezpośrednio zastosowane do

określenia własności granicznych wyrażenia $[\tilde{r}_N(v) - r(v)]$. Zgodnie z przyjętymi założeniami odnośnie sygnału wejściowego i szumu na wyjściu systemu, zachodzą następujące równości

$$E\{\tilde{r}_N(v)\} = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v-\alpha x}{h}\right) g(x) f_V(x)\, dx = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v-x}{h}\right) r(x)\, dx. \quad (3.28)$$

Ponieważ nieliniowa charakterystyka $g(\cdot)$ jest ograniczona przez wielomian (zob. założenie 3.2) oraz sygnał wejściowy jest gaussowski, to całka $\int_{-\infty}^{\infty} |r(v)|\, dv$ jest skończona. Stąd, zgodnie z lematem B.3 z dodatku B.3, na mocy założeń 3.4 i 3.5, ma miejsce następująca zbieżność

$$E\{\tilde{r}_N(v)\} \to r(v), \text{ gdy } N \to \infty, \quad (3.29)$$

w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą. Zauważmy dalej, że z założenia o skończonej długości pamięci i definicji sygnału $\{V_n\}$ (zob. wzór 2.10) wynika następująca własność (por. własność 3.1):

Własność 3.2 *Zmienne losowe V_n i V_k są niezależne jeżeli $|n-k| > p$.*

Wykorzystując własność 3.2, na mocy lematu A.1 z dodatku A.1, prawdziwe jest następujące oszacowanie wariancji estymatora $\tilde{r}_N(v)$,

$$Var\{\tilde{r}_N(v)\} \leq \frac{p+1}{N_I h^2} Var\left\{Y_1 K\left(\frac{v-\alpha V_1}{h}\right)\right\}. \quad (3.30)$$

Ze względu na ograniczoność jądra $K(\cdot)$ zauważamy następnie, że element

$$h^{-1} Var\{Y_1 K([v-\alpha V_1]/h)\}$$

występujący w (3.30), jest ograniczony przez (por. Greblicki [25])

$$\frac{2k_0}{h} E\left\{g^2(V_1) K\left(\frac{v-\alpha V_1}{h}\right)\right\} + \frac{2k_0\sigma_Z^2}{h} E\left\{K\left(\frac{v-\alpha V_1}{h}\right)\right\}. \quad (3.31)$$

Ponowne wykorzystanie założeń 3.1 i 3.2 prowadzi do wniosku, że $\int_{-\infty}^{\infty} \left|g^2(v/\alpha) f_V(v/\alpha)\right| dv < \infty$. Zatem, na mocy lematu B.3 z dodatku B.3, wyrażenie (3.31) dąży do

$$\frac{2k_0}{\alpha} f_V\left(\frac{v}{\alpha}\right)\left[g^2\left(\frac{v}{\alpha}\right) + \sigma_Z^2\right] < \infty,$$

gdy $N \to \infty$ w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą. Ostatecznie więc nierówność (3.30), dla h spełniającego założenie 3.5, daje

$$Var\{\tilde{r}_N(v)\} = O\left(N^{-4/5}\right) \quad (3.32)$$

w każdym punkcie ciągłości funkcji $g(v/\alpha)$.

Zgodnie z rezultatami w (3.29) i (3.32) zachodzi więc zbieżność

$$[\tilde{r}_N(v) - r(v)] \to 0, \qquad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą.

Przejdziemy teraz do drugiego etapu dowodu tj. do zbadania asymptotyki wyrażenia $[\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)]$ w dekompozycji (3.26). Korzystając z faktu, że jądro $K(\cdot)$ spełnia warunek Lipschitza (zob. zał. 3.4), otrzymujemy

$$\left|\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)\right| \leq \sup_{n\in I}\left|\alpha V_n - \hat{V}_{n,N}^{Cr}\right| \frac{L}{N_I h^2} \sum_{n\in I} |Y_n|.$$

Na mocy lematu A.1 w dodatku A zachodzi następujące oszacowanie

$$Var\left\{\frac{1}{N_I}\sum_{n\in I}|Y_n|\right\} \leq \frac{p+1}{N_I} Var\left\{|Y_n|\right\} = O\left(N^{-1}\right).$$

Wobec tego

$$\frac{1}{N_I}\sum_{n\in I}|Y_n| = E\left\{|Y_1|\right\} + O_P\left(N^{-1/2}\right). \tag{3.33}$$

Uzyskany w lemacie 3.3 rezultat dotyczący zbieżności wyrażenia $\sup_{n\in I}\left|\alpha V_n - \hat{V}_{n,N}^{Cr}\right|$ wraz z obserwacją (3.33) i założeniem 3.5, prowadzą do wniosku, że

$$\left|\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)\right| \to 0, \qquad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$. Ostatecznie więc

$$\hat{r}_N^{Cr}(v) \to r(v), \qquad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, w którym $g(v/\alpha)$ jest funkcją ciągłą. Stosując analogiczną argumentację dla estymatora $\hat{f}_N^{Cr}(v)$ otrzymujemy

$$\hat{f}_N^{Cr}(v) \to \alpha^{-1} f_V(v/\alpha), \qquad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, co kończy dowód. ■

Analiza szybkości zbieżności estymatora charakterystyki nieliniowej

Dla określenia rzędu szybkości zbieżności estymatora $\hat{g}_N^{Cr}(\cdot)$ ograniczymy klasę rozważanych systemów Wienera do klasy z nieliniowościami, dla których spełnione są założenia (por. Greblicki [26]):

Założenie 3.6 *Druga pochodna $g''(\cdot)$ nieliniowej charakterystyki $g(\cdot)$ jest funkcją ciągłą.*

Założenie 3.7 *Pochodna $g'(\cdot)$ nieliniowości $g(\cdot)$ ograniczona jest przez wielomian, tj. istnieje taki wielomian $m_0(\cdot)$ dowolnego skończonego stopnia, że $|g'(v)| \leq m_0(v)$, $\forall v \in \mathbf{R}$.*

Przyjmiemy również dodatkowe założenia o funkcji jądra $K(\cdot)$.

Założenie 3.8 *Jądro $K(\cdot)$ posiada ciągłą i ograniczoną drugą pochodną $K''(\cdot)$, tj. $|K''(v)| < k_1 < \infty$, $\forall v \in \mathbf{R}$. Ponadto $\int_{-\infty}^{\infty} vK(v)\,dv = 0$ oraz $\int_{-\infty}^{\infty} v^2 K(v)\,dv < \infty$.*

Możemy teraz sformułować następujące twierdzenie:

Twierdzenie 3.3 *Jeżeli spełnione są założenia twierdzenia 3.2 oraz dodatkowo założenia 3.6–3.8, to dla systemu Wienera (2.10)–(2.11) zachodzi*

$$\hat{g}_N^{Cr}(v) = g(v/\alpha) + O_P\left(N^{-(2/5)+\varepsilon}\right), \tag{3.34}$$

dla dowolnie małego $\varepsilon > 0$, w każdym punkcie $v \in \mathbf{R}$.

Dowód twierdzenia zamieszczony jest w dodatku A.1.

3.6 Podsumowanie wyników teoretycznych

Dla opracowanej w rozdziale metody identyfikacji przeprowadzono analizę podstawowych własności asymptotycznych. W szczególności sformułowane zostały łagodne wymagania jakie musi spełniać nieliniowa charakterystyka w systemie aby zapewnić zbieżność algorytmu (założenia 3.2 i 3.3). Ponadto, dla estymatorów odpowiedzi impulsowej podsystemu dynamicznego udowodniona została parametryczna szybkość zbieżności $O_P\left(N^{-1/2}\right)$. W przypadku odpowiednio gładkich nieliniowości natomiast (założenia 3.6 i 3.7), pokazano że estymator $\hat{g}_N^{Cr}(\cdot)$ zbiega punktowo do przeskalowanej charakterystyki nieliniowej $g(\cdot/\alpha)$ z szybkością $O_P\left(N^{-(2/5)+\varepsilon}\right)$.

3.7 Wyniki badań eksperymentalnych

Zachowanie się algorytmu dla małej i średniej liczby obserwacji badano eksperymentalnie przy użyciu środowiska obliczeniowego MATLAB. Eksperymenty wykonano dla systemów Wienera z liniowymi podsystemami dynamicznymi opisanymi odpowiedzią impulsową

$$\lambda_i = \left(\frac{1}{i+1}\right),\ i = 0, 1, \ldots, p, \tag{3.35}$$

gdzie $p = 4$. Symulacje przeprowadzono dla trzech różnych charakterystyk nieliniowych występujących w systemie (zob. rys. 3.1),

$$\begin{aligned} g_1(v) &= 2\arctan(2v), \\ g_2(v) &= 2\cos(2v), \\ g_3(v) &= \lfloor v \rfloor, \end{aligned} \tag{3.36}$$

gdzie $\lfloor\cdot\rfloor$ jest funkcją „podłoga". Sygnał wejściowy systemu wygenerowano z rozkładu $\mathcal{N}(\mu_X, \sigma_X)$ z parametrami $\mu_X = 1$ oraz $\sigma_X = 0,5$. Stąd, zgodnie z przeprowadzoną wyżej analizą teoretyczną algorytmu, sygnał interakcyjny $\{V_n\}$ posiadał rozkład normalny z wartością oczekiwaną i wariancją (por. wzory (3.1)):

$$\begin{aligned} \mu_V &= \mu_X \sum_{i=0}^{p} \lambda_i \approx 2,28, \\ \sigma_V^2 &= \sigma_X^2 \sum_{i=0}^{p} \lambda_i^2 \approx 0,36. \end{aligned}$$

Przyjmiemy następujące oznaczenia $\vartheta_1 = \mu_V - 3\sigma_V \approx 0,48$ oraz $\vartheta_2 = \mu_V + 3\sigma_V \approx 4,08$. Wtedy na mocy reguły „trzech sigm" zachodzi oszacowanie

$$P\{V_1 \in [\vartheta_1, \vartheta_2]\} \approx 1. \tag{3.37}$$

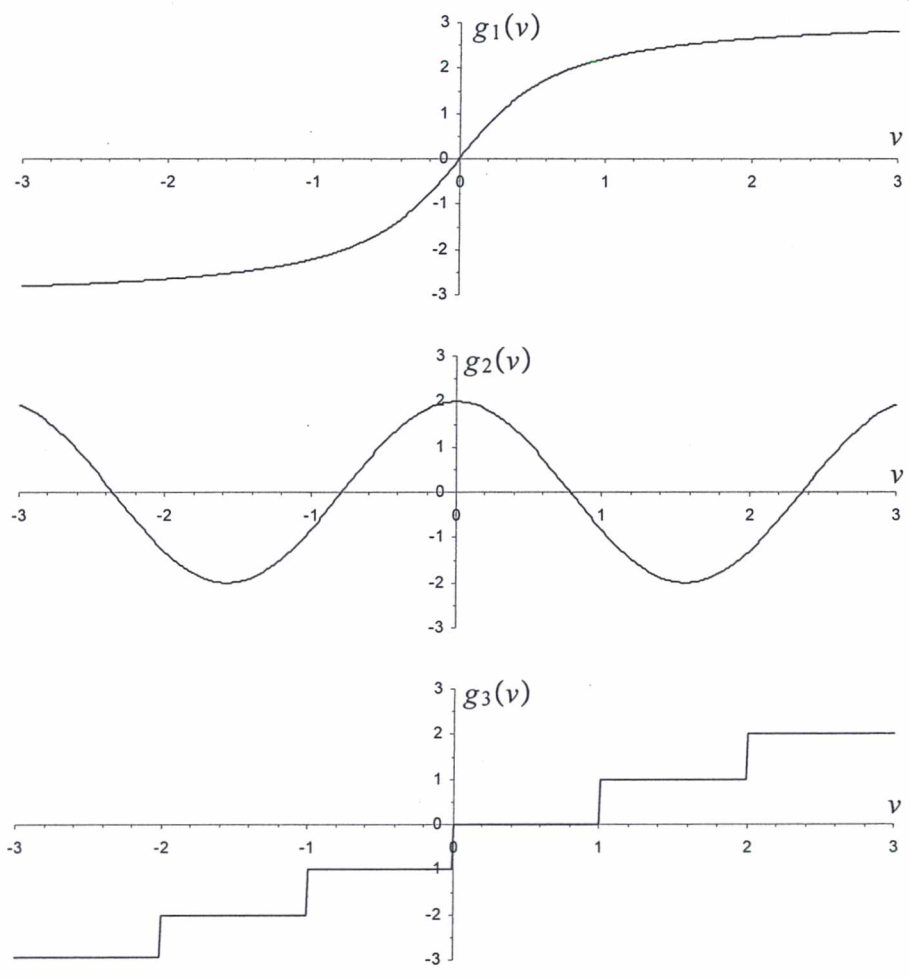
Sygnał zakłócający $\{Z_n\}$ wygenerowano z rozkładu jednostajnego na odcinku $[-\tau, \tau]$, gdzie τ jest stałą[2], dobraną tak, aby[3] NSR= $\tau/\gamma = 10\%$ (przy czym $\gamma = \max_{v\in[\vartheta_1,\vartheta_2]} g(v)$). W nieparametrycznym estymatorze $\hat{g}_N^{Cr}(\cdot)$ zastosowano gaussowską funkcję jądra, tj.

$$K(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right). \tag{3.38}$$

Dla każdego z rozważanych systemów wyznaczono numerycznie wartości multiplikatywnej stałej α (zob. wzór. (3.6)). Dokładność metody badano przy użyciu następujących błędów empirycznych:

[2] wartość stałej τ wyznaczono oddzielnie dla każdej z rozważanych nieliniowości.

[3] "Noise–to–Signal Ratio" – współczynnik określający poziom szumu względem sygnału użytecznego.



Rysunek 3.1: Wykorzystane w symulacjach nieliniowości g_1, g_2 i g_3

- Błąd (względny) estymacji odpowiedzi impulsowej podsystemu dynamicznego:

$$Err(\hat{\underline{\lambda}};\alpha)=\frac{1}{R}\sum_{i=1}^{R}\frac{\left\|\hat{\underline{\lambda}}_N^{Cr[i]}-\alpha\underline{\lambda}\right\|^2}{\|\alpha\underline{\lambda}\|^2}100\%=\frac{1}{R}\sum_{i=1}^{R}\frac{\left\|\alpha^{-1}\hat{\underline{\lambda}}_N^{Cr[i]}-\underline{\lambda}\right\|^2}{\|\underline{\lambda}\|^2}100\%. \tag{3.39}$$

- Globalny błąd (względny) estymacji charakterystyki nieliniowej:

$$MISE\left(\hat{g}_N^{Cr};\alpha\right)=\frac{\frac{1}{R}\sum_{i=1}^{R}\left[\theta\sum_{j=1}^{M}\left(\hat{g}_N^{Cr[i]}\left(\alpha v_j\right)-g\left(v_j\right)\right)^2\right]}{\int_{\vartheta_1}^{\vartheta_1}g^2\left(v\right)dv}100\%, \tag{3.40}$$

gdzie $(\cdot)^{[i]}$ oznacza i–tą realizację estymatora, $\|\cdot\|$ jest normą euklidesową oraz $R=30$ jest liczbą realizacji estymatora (powtórzeń eksperymentu). Ponadto $\hat{\underline{\lambda}}_N^{Cr}=\left(\hat{\lambda}_{0,N}^{Cr},\ldots,\hat{\lambda}_{p,N}^{Cr}\right)^T$ i $\underline{\lambda}=(\lambda_0,\ldots,\lambda_p)^T$. W definicji błędu $MISE\left(\hat{g}_N^{Cr};\alpha\right)$ przyjęto $M=1000$. Ciąg $\{v_n\}$ jest

M–elementowym zbiorem punktów równomiernie rozłożonych na odcinku $[\vartheta_1, \vartheta_2]$. Przez θ oznaczono odległość miedzy sąsiadującymi ze sobą wyrazami ciągu $\{v_n\}$.

W kontekście problematyki związanej z identyfikacją systemów nieliniowych (a w szczególności z identyfikacją systemów Wienera), nieliniowości $g_1(\cdot)$, $g_2(\cdot)$ i $g_3(\cdot)$ scharakteryzować można następująco:

- Funkcja $g_1(\cdot)$ jest typowym przykładem charakterystyki często spotykanej w literaturze (zob. np. Raich *et al.* [98], Enqvist [19], Hasiewicz *et al.* [53]). Ze względu na fakt, że jest ona funkcją ściśle monotoniczną, wielokrotnie różniczkowalną oraz nieparzystą, nieliniowość $g_1(\cdot)$ spełnia typowe wymagania występujące w identyfikacji systemów Wienera (jest np. funkcją odwracalną, por. Greblicki [26]).
- Nieliniowość $g_2(\cdot)$, podobnie jak $g_1(\cdot)$, jest funkcją wielokrotnie różniczkowalną. W odróżnieniu od $g_1(\cdot)$ charakterystyka $g_2(\cdot)$ jest jednak tylko przedziałami ściśle monotoniczna, co powoduje, że jest ona również tylko przedziałami odwracalna. Nieliniowość $g_2(\cdot)$ jest funkcją parzystą.
- W odróżnieniu od funkcji $g_1(\cdot)$ i $g_2(\cdot)$ charakterystyka $g_3(\cdot)$ jest funkcją nieciągłą (nieciągłości występują w punktach $v \in \mathbf{N}$). Ponadto, ze względu na fakt, że $g_3(\cdot)$ jest odcinkami stała, nie istnieje funkcja odwrotna do $g_3(\cdot)$.

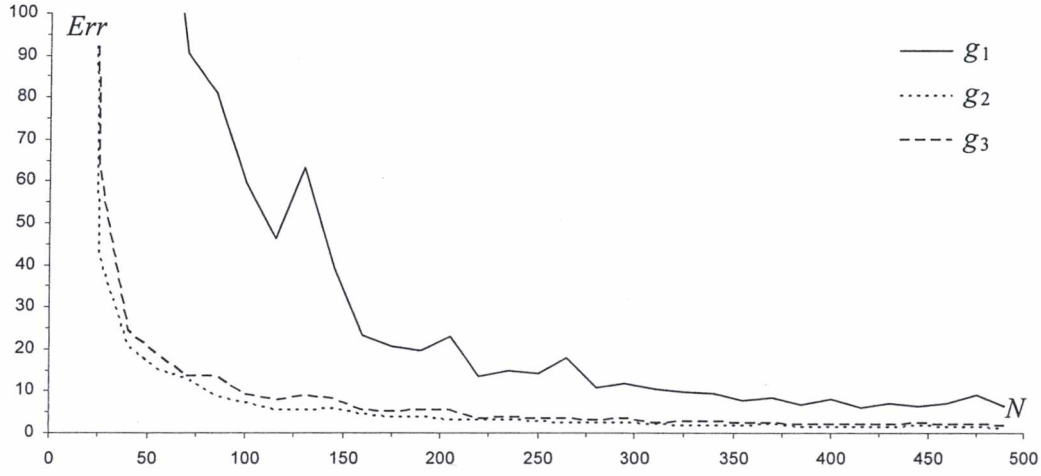
W eksperymencie badano zachowanie algorytmu identyfikacji dla nieliniowości $g_1(\cdot)$, $g_2(\cdot)$ i $g_3(\cdot)$. Rysunki 3.2 oraz 3.3 (str. 35) przedstawiają odpowiednio błędy empiryczne $Err(\hat{\lambda}_N^{Cr}; \alpha)$ i $MISE\left(\hat{g}_N^{Cr}; \alpha\right)$ dla rozważanych charakterystyk, w funkcji liczby obserwacji N zmienianej od 25 do 500. Uzyskane rezultaty pokazują, że w przypadku omawianych nieliniowości błędy te maleją wraz ze wzrostem liczby obserwacji N. W szczególności, dla charakterystyk $g_2(\cdot)$ i $g_3(\cdot)$ wartości błędów $Err(\hat{\lambda}_N^{Cr}; \alpha)$ są mniejsze od 10%, gdy liczba obserwacji N jest większa od 100 (por. rys. 3.2). Dla charakterystyki $g_1(\cdot)$ natomiast, dziesięcioprocentowy błąd $Err(\hat{\lambda}_N^{Cr}; \alpha)$ osiągany jest dopiero, gdy $N = 350$. Gorsze rezultaty uzyskane w przypadku nieliniowości $g_1(\cdot)$ wynikają z faktu, że na odcinku $[\vartheta_1, \vartheta_2]$ (zob. wzór (3.37)) funkcja $g_1(v) = \arctan(2v)$ jest bliska funkcji stałej (por. rys. 3.1), co z kolei powoduje, że kowariancja $Cov\{X_1, Y_1\}$ jest bliska zeru (por. zał. 3.3, str. 23) i tym samym bliska zeru jest również multiplikatywna stała α. Omówiony powyżej efekt nie jest natomiast widoczny na rysunku 3.3 przedstawiającym błędy estymacji nieliniowości. Zastosowana w symulacjach ciągła funkcja jądra (zob. wzór (3.38)) powoduje, że estymator $\hat{g}_N^{Cr}(\cdot)$ lepiej przybliża ciągłą nieliniowość $g_1(\cdot)$ niż nieciągłą $g_3(\cdot)$, co kompensuje większy błąd estymatora $\hat{\lambda}_N^{Cr}$ w przypadku nieliniowości $g_1(\cdot)$. W szczególności, rysunek 3.3 prowadzi do wniosku, że bez względu na charakter nieliniowości występującej w systemie, błędy empiryczne $MISE\left(\hat{g}_N^{Cr}; \alpha\right)$ osiągają wartości nie większe niż 20%, gdy liczba obserwacji N jest większa od 400.

3.8 Podsumowanie i wnioski

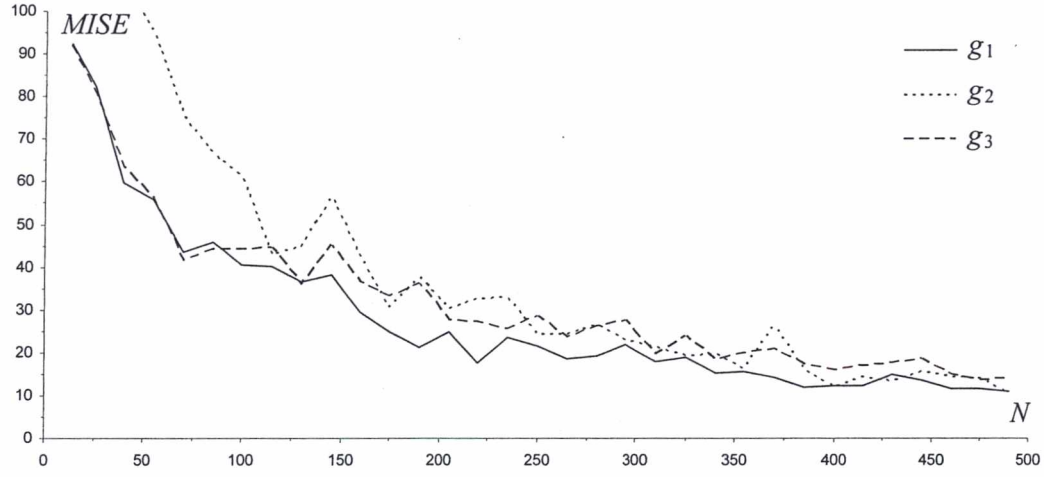
Zaproponowana w tym rozdziale metoda identyfikacji dedykowana jest systemom Wienera pobudzanym sygnałami gaussowskimi typu *i.i.d.* Analiza teoretyczna algorytmu pozwala stwierdzić, że dla stosunkowo szerokiej klasy nieliniowości estymatory części liniowej i nieliniowej są zgodne, oraz że estymator odpowiedzi impulsowej podsystemu dynamicznego posiada optymalny rząd szybkości zbieżności. Dodatkowo, dla dwukrotnie różniczkowalnych nieliniowości w systemie, estymator części nieliniowej posiada rząd szybkości zbieżności $O_P\left(N^{-(2/5)+\varepsilon}\right)$ tylko *nieznacznie* gorszy od rzędu szybkości zbieżności

estymatorów nieparametrycznych osiąganego przy identyfikacji systemów statycznych (tj. rzędu $O_P\left(N^{-(2/5)}\right)$).

Należy podkreślić, że prosta konstrukcja estymatorów $\{\hat{\lambda}_{i,N}^{Cr}\}_{i=0}^{p}$ i $\hat{g}_N^{Cr}(\cdot)$ pozwala na łatwą implementację proponowanej metody identyfikacji. Algorytm nie wymaga stosowania specjalizowanych procedur numerycznych (np. odwracania macierzy lub poszukiwania ekstremum wielowymiarowych funkcji celu) co powoduje, że jego efektywna implementacja nie jest skomplikowana.



Rysunek 3.2: Wartości błędów estymacji $Err(\hat{\lambda}_N^{Cr};\alpha)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2 i g_3



Rysunek 3.3: Wartości błędów estymacji $MISE\left(\hat{g}_N^{Cr};\alpha\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2 i g_3

Rozdział 4

Identyfikacja systemów Wienera z wykorzystaniem metody najmniejszych kwadratów

W tym rozdziale przedstawiony jest algorytm identyfikacji systemów Wienera zbudowany w oparciu o metodę najmniejszych kwadratów oraz nieparametryczną estymację funkcji regresji. W odróżnieniu od metody korelacyjnej algorytm nie wymaga, aby wejście systemu było sygnałem gaussowskim. Przyjmuje się, że sygnał $\{X_n\}$ jest ograniczony oraz ma ciągłą funkcję gęstości prawdopodobieństwa, różną od zera w otoczeniu punktu zero. Odnośnie nieliniowej charakterystyki $g(\cdot)$ zakłada się jedynie, że jest ona funkcją ciągłą. Dopuszcza się w ten sposób zarówno charakterystyki odwracalne, jak i nieodwracalne. Wyniki dotyczące metody zostały otrzymane we współpracy z prof. M. Pawlakiem i prof. Z. Hasiewiczem, i częściowo opublikowane w [93]. W rozdziale prezentowana jest także pokrewna idea metody identyfikacji skonstruowanej w oparciu o koncepcję korelacji rangowej.

Wprowadzone estymatory części liniowej, sygnału interakcyjnego i części nieliniowej systemu oznaczane będą indeksem „*LS*" w przypadku algorytmu zbudowanego w oparciu o metodę najmniejszych kwadratów (od *ang. Least Squares*) oraz indeksem „*RC*" w przypadku metody rangowej (od *ang. Rank Correlation*).

4.1 Sformułowanie problemu

Zgodnie z dyskusją w rozdziale 2, system Wienera (2.10)–(2.11) może być zidentyfikowany tylko z dokładnością do multiplikatywnej stałej. Stąd, bez utraty ogólności możemy przyjąć, że dla rozważnej klasy systemów zachodzi $\lambda_0 = 1$. Dodatkowo przyjmiemy następujące założenia o nieliniowości $g(\cdot)$ i gęstości prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego $\{X_n\}$ (zob. Pawlak *et al.* [93]):

Założenie 4.1 *Gęstość prawdopodobieństwa $f_X(\cdot)$ jest funkcją ciągłą o ograniczonym nośniku $[a_f, b_f]$ takim, że $a_f \leq 0 \leq b_f$ oraz $f_X(0) > 0$.*

Założenie 4.2 *Nieliniowa charakterystyka $g(\cdot)$ jest ciągła i różna od funkcji stałej.*

Uwaga 4.1 *Ponieważ stała a_f może być dowolnie mała, a stała b_f dowolnie duża, to występujący w założeniu 4.1 warunek ograniczonego nośnika funkcji $f_X(\cdot)$ nie jest z praktycznego punktu widzenia istotny. Przykładami gęstości spełniających założenie 4.1*

są m.in.: rozkład jednostajny, rozkład trójkątny, rozkład Epanechnikova oraz obcięty rozkład normalny.

4.2 Parametryczna identyfikacja podsystemu dynamicznego z wykorzystaniem metody najmniejszych kwadratów

4.2.1 Koncepcja metody identyfikacji

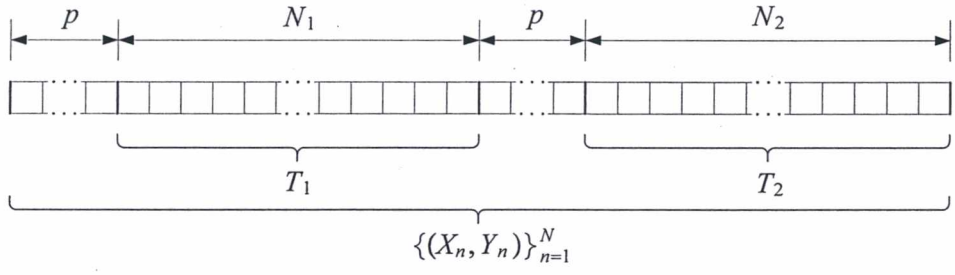
Proponowany algorytm identyfikacji wymaga wydzielenia ze zbioru obserwacji wejścia i wyjścia $\{(X_n, Y_n)\}_{n=1}^{N}$ dwóch podzbiorów T_1 i T_2 określonych następująco

$$\begin{aligned} T_1 &= \{(X_n, Y_n)\}_{n\in I_1}, \\ T_2 &= \{(X_n, Y_n)\}_{n\in I_2}, \end{aligned}$$

gdzie

$$\begin{aligned} I_1 &= \{p+1, \ldots, p+N_1\}, \\ I_2 &= \{2p+1+N_1, \ldots, N\}, \end{aligned}$$

oraz N_1 oznacza liczbę elementów zbioru I_1, (podobnie przez N_2 będziemy oznaczać liczność zbioru I_2), rys. 4.1. Ponieważ rozważana klasa systemów posiada pamięć o długości p, łatwo zauważyć, że elementy zbiorów T_1 i T_2 są wzajemnie stochastycznie niezależne[1]. Odnośnie liczności zbiorów T_1 i T_2 założymy, że $N_1 \to \infty$ i $N_2 \to \infty$, gdy $N \to \infty$, por. Pawlak *et al.* [93].



Rysunek 4.1: Wydzielenie ze zbioru obserwacji $\{(X_n, Y_n)\}_{n=1}^{N}$ stochastycznie niezależnych podzbiorów T_1 i T_2

Koncepcję metody identyfikacji omówimy w dwóch etapach. Zakładając chwilowo, że nieliniowa charakterystyka $g(\cdot)$ *jest znana*, rozpoczniemy od rozważenia zadania identyfikacji polegającego na oszacowaniu odpowiedzi impulsowej podsystemu dynamicznego na podstawie obserwacji ze zbioru T_2. W celu uproszczenia zapisu wprowadzimy następującą notację wektorową. Przez $\underline{\lambda}$ oznaczymy $(p+1)$–elementowy wektor utworzony z kolejnych wyrazów odpowiedzi impulsowej $\{\lambda_i\}_{i=0}^{p}$ podsystemu dynamicznego, tj. wektor o postaci

$$\underline{\lambda} = (1, \lambda_1, \ldots, \lambda_p)^T,$$

oraz określimy zbiór $\Lambda \subset \mathbf{R}^{p+1}$ jako podzbiór przestrzeni $\mathbf{R}^{p+1}$ utworzony z wektorów $\underline{l} = (l_0, l_1, \ldots, l_p)^T$, takich że $l_0 = 1$ oraz $l_1, \ldots, l_p \leq L_0$, gdzie $L_0 < \infty$ jest pewną stałą. Zbiór Λ jest oczywiście zbiorem zwartym. Jest on bowiem zbiorem domkniętym i ograniczonym w przestrzeni skończenie wymiarowej.

Przyjmiemy następujące założenie:

[1] tj. dowolna zmienna losowa ze zbioru T_1 nie zależy od zmiennych losowych ze zbioru T_2 i odwrotnie.

Założenie 4.3 *Wektor* $\underline{\lambda}$ *jest elementem zbioru* Λ, *tj.* $\underline{\lambda} \in \Lambda$.

Następnie, przez analogię do definicji sygnału wewnętrznego $\{V_n\}$ (zob. wzór (2.10)), dla każdego $\underline{l} \in \Lambda$ wprowadzimy sygnał $\{V_n(\underline{l})\}$ zależny od wektorowego parametru $\underline{l} \in \Lambda$ taki, że

$$V_n(\underline{l}) = \sum_{i=0}^{p} l_i X_{n-i}. \tag{4.1}$$

Zgodnie ze wzorem (4.1) oraz wzorem (2.10) definiującym sygnał $\{V_n\}$, zachodzi

$$V_n(\underline{\lambda}) = V_n, \tag{4.2}$$

i tym samym $g(V_n(\underline{\lambda})) = g(V_n)$.

Podobieństwo konstrukcji sygnałów $\{V_n\}$ i $\{V_n(\underline{l})\}$ oraz ciągłość funkcji $V_n(\underline{l})$ powoduje, że jeżeli wektor $\underline{l}$ jest bliski[2] wektorowi $\underline{\lambda}$, to różnica $V_n - V_n(\underline{l})$ jest mała (w pewnym sensie probabilistycznym[3]) i wobec tego dalej, jeżeli $g(\cdot)$ jest funkcją ciągłą, to także różnica $g(V_n) - g(V_n(\underline{l}))$ będzie odpowiednio mała. Obserwacja ta prowadzi do wniosku, że możliwe jest skonstruowanie (w oparciu o $g(V_n(\underline{l}))$ empirycznej funkcji kryterialnej $\tilde{Q}_{N_2}(\underline{l})$ określającej dla dowolnego wektora $\underline{l} \in \Lambda$ „jakość przybliżenia" wektora $\underline{\lambda}$ przez $\underline{l}$. W tym celu, na podstawie obserwacji ze zbioru T_2, można utworzyć zbiór $\{(V_n(\underline{l}), Y_n)\}_{n \in I_2}$ i funkcję kryterialną $\tilde{Q}_{N_2}(\underline{l})$ zdefiniować następująco

$$\tilde{Q}_{N_2}(\underline{l}) = \frac{1}{N_2} \sum_{m \in I_2} [Y_m - g(V_m(\underline{l}))]^2. \tag{4.3}$$

W rezultacie, wykorzystując kryterium (4.3), za estymator odpowiedzi impulsowej $\underline{\lambda}$ można przyjąć wektor $\underline{l}$, który minimalizuje funkcję $\tilde{Q}_{N_2}(\cdot)$, tj. estymator najmniejszych kwadratów o postaci

$$\underline{\tilde{\lambda}}_{N_2}^{LS} = \arg \min_{\underline{l} \in \Lambda} \tilde{Q}_{N_2}(\underline{l}).$$

W sytuacji podstawowej dla naszych rozważań, gdy nieliniowa charakterystyka $g(\cdot)$ *nie jest znana*, zastosowanie funkcji kryterialnej $\tilde{Q}_{N_2}(\underline{l})$ wprost nie jest możliwe i w konsekwencji proponowane podejście nie może być bezpośrednio zastosowane. Można jednak powyższy schemat zmodyfikować i w miejsce nieznanych wielkości $\{g(V_n(\underline{l}))\}_{n \in I_2}$ wstawić ich odpowiednie estymatory.

4.2.2 Konstrukcja pilotowego estymatora charakterystyki nieliniowej

Wprowadzony wyżej sygnał $\{V_n(\underline{l})\}$ wykorzystamy do zdefiniowania funkcji $g(\cdot; \underline{l}) : \mathbf{R} \to \mathbf{R}$, zależnej od wektorowego parametru $\underline{l}$, takiej że

$$g(v; \underline{l}) = E\{Y_n | V_n(\underline{l}) = v\}. \tag{4.4}$$

Zgodnie z obserwacją (4.2) oraz wzorem (2.9) otrzymujemy natychmiast, że $g(v; \underline{\lambda}) = g(v)$. Zauważmy następnie, że dla dowolnego $\underline{l} \in \Lambda$ możliwe jest skonstruowanie

[2] np. w sensie metryki euklidesowej.

[3] Koncepcja metody identyfikacji jest w sposób celowy omawiana z pominięciem ścisłych pojęć teorii prawdopodobieństwa, a jej precyzyjne sformułowanie oraz własności asymptotyczne zamieszczone są w dalszej części rozdziału.

estymatora funkcji $g(v;\underline{l})$ na podstawie zbioru $\{(V_n(\underline{l}), Y_n)\}_{n\in I_1}$. Wykorzystując np. nieparametryczny estymator z jądrem (por. rozdział 2, str. 15), otrzymujemy[4]

$$\tilde{g}_{N_1}(v;\underline{l}) = \frac{\sum_{n\in I_1} Y_n K_0\left(\frac{v - V_n(\underline{l})}{h_0}\right)}{\sum_{n\in I_1} K_0\left(\frac{v - V_n(\underline{l})}{h_0}\right)}, \tag{4.5}$$

gdzie $K_0(\cdot)$ i $h_0 = h_0(N_1)$ są odpowiednio funkcją jądra i parametrem wygładzania zależnym od liczby obserwacji N_1. Zauważmy, że dla dowolnego wektora $\underline{l} \in \Lambda$, ze względu na zależność (4.1), estymator $\tilde{g}_{N_1}(v;\underline{l})$ może być wyznaczony na podstawie dostępnych obserwacji wejścia i wyjścia systemu ze zbioru T_1. Warto również podkreślić, że występujące w (4.5) sygnały $\{V_n(\underline{l})\}$ i $\{Y_n\}$ nie są ciągami typu *i.i.d.*, co uniemożliwia bezpośrednie zastosowanie klasycznej teorii do określenia własności asymptotycznych estymatora $\tilde{g}_{N_1}(\cdot;\underline{l})$ (zob. też uwaga 2.1 na str. 20).

Modyfikacja funkcji kryterialnej $\tilde{Q}_{N_2}(\underline{l})$

W celu skonstruowania estymatora odpowiedzi impulsowej $\underline{\lambda}$ w przypadku nieznanej nieliniowości $g(\cdot)$, zauważmy, że jeżeli $\tilde{g}_{N_1}(\cdot;\underline{l})$ jest zgodnym estymatorem funkcji $g(\cdot;\underline{l})$ (dla każdego $\underline{l} \in \Lambda$), to im mniejsza jest odległość wektora $\underline{l}$ od $\underline{\lambda}$, tym dokładniej $\tilde{g}_{N_1}(\cdot;\underline{l})$ przybliża nieliniową charakterystykę $g(\cdot) = g(\cdot,\underline{\lambda})$. Powyższa obserwacja sugeruje więc, aby estymator $\tilde{g}_{N_1}(\cdot;\underline{l})$ wstawić w miejsce prawdziwej charakterystyki nieliniowej $g(\cdot)$ we wzorze definiującym funkcję kryterialną $\tilde{Q}_{N_2}(\cdot)$. W rezultacie otrzymujemy nowe, w pełni empiryczne, kryterium określające jakość „dopasowania" wektora $\underline{l}$ do $\underline{\lambda}$,

$$\hat{Q}_N(\underline{l}) = \frac{1}{N_2}\sum_{m\in I_2}\left[Y_m - \tilde{g}_{N_1}(V_m(\underline{l});\underline{l})\right]^2. \tag{4.6}$$

Wynikający z (4.6) estymator najmniejszych kwadratów odpowiedzi impulsowej $\underline{\lambda}$ określony jest zatem wzorem

$$\hat{\underline{\lambda}}_N^{LS} = \operatorname{argmin}_{\underline{l}\in\Lambda} \hat{Q}_N(\underline{l}).$$

Jak zaznaczono wyżej, w odróżnieniu od $\tilde{\underline{\lambda}}_{N_2}^{LS}$, do wyznaczenia estymatora $\hat{\underline{\lambda}}_N^{LS}$ nie jest wymagana znajomość nieliniowej charakterystyki $g(\cdot)$.

4.3 Estymacja sygnału interakcyjnego

Skonstruowany powyżej estymator $\hat{\underline{\lambda}}_N^{LS}$ można teraz wykorzystać do estymacji elementów sygnału interakcyjnego $\{V_n\}$. Stosując podejście wprowadzone w rozdziale 3. (por. wzór (3.10) na str. 25), otrzymujemy

$$\hat{V}_{n,N}^{LS} = \left(\hat{\underline{\lambda}}_N^{LS}\right)^T \underline{X}_n, \tag{4.7}$$

gdzie $\underline{X}_n = (X_n, X_{n-1}, \ldots, X_{n-p})^T$. Ze względu na ograniczoną liczbę obserwacji wejścia i wyjścia systemu, estymator (4.7) może być wyznaczony jedynie dla V_n takich, że $n \in I = \{p+1, p+2, \ldots, N\}$, por. wzór (3.11) w rozdziale 3.

[4] W dalszej części rozdziału estymator $\tilde{g}_{N_1}(v;\underline{l})$ będzie nazywany estymatorem pilotowym.

4.4 Nieparametryczna identyfikacja charakterystyki nieliniowej

Jako estymator charakterystyki nieliniowej przyjmiemy, podobnie jak w punkcie 4.2.2, nieparametryczny estymator z jądrem zbudowany tym razem w oparciu o zbiór $\{(\hat{V}_{n,N}^{LS}, Y_n)\}_{n\in I}$, tzn.

$$\hat{g}_N^{LS}(v) = \frac{\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{LS}}{h}\right)}{\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{LS}}{h}\right)}, \tag{4.8}$$

gdzie $K(\cdot)$ i $h = h(N)$ są odpowiednio funkcją jądra i parametrem wygładzania, por. estymatory (3.14) oraz (4.5). Własności asymptotyczne estymatora oraz wyniki eksperymentów numerycznych przeprowadzonych dla małej i średniej liczby obserwacji przedstawione są odpowiednio w p. 4.5.3 i p. 4.8.

4.5 Własności asymptotyczne algorytmu

Zbadamy obecnie asymptotyczne zachowanie omawianego algorytmu identyfikacji. W szczególności skupimy się na wykazaniu zgodności estymatora odpowiedzi impulsowej podsystemu dynamicznego i estymatora nieliniowej charakterystyki podsystemu statycznego. W tym celu przyjmiemy następujące założenia odnośnie elementów konstrukcyjnych pilotowego estymatora $\tilde{g}_{N_1}(\cdot;\underline{l})$:

Założenie 4.4 *Jądro $K_0(\cdot)$ jest funkcją ciągłą spełniającą warunek $\int_{-\infty}^{\infty} K_0(x)\,dx = 1$ oraz taką, że*

$$k_1 \mathbf{1}_{[-r,r]}(v) \leq K_0(v) \leq k_2 \mathbf{1}_{[-R,R]}(v),$$

gdzie $r \leq R$, $k_1 \leq k_2$ są dowolnymi stałymi.

Założenie 4.5 *Parametr wygładzania $h_0 = h_0(N_1)$ spełnia warunki: $h_0 \to 0$ oraz $N_1 h_0 \to \infty$ gdy $N_1 \to \infty$.*

W celu zapewnienia identyfikowalności podsystemu dynamicznego przyjmiemy również założenie typowe dla metody najmniejszych kwadratów:

Założenie 4.6 *Funkcja $Q(\underline{l}) = E\left\{[Y_1 - g(V_1(\underline{l});\underline{l})]^2\right\}$ posiada jedyne minimum globalne w punkcie $\underline{\lambda}$.*

Ponieważ w $Q(\underline{l})$ argument $\underline{l}$ wpływa równocześnie na konstrukcję funkcji $g(\cdot;\underline{l})$ (zob. wzór (4.4)) jak i na jej argument $V_1(\underline{l})$ (zob. wzór (4.1)), to automatyczne spełnienie założenia 4.6 nie jest oczywiste.

Wymagania odnośnie $\hat{g}_N^{LS}(\cdot)$ umożliwiające wykazanie zgodności estymatora charakterystyki nieliniowej zamieszczone są w p. 4.5.3.

4.5.1 Analiza zbieżności estymatora odpowiedzi impulsowej podsystemu dynamicznego

Zanim określone zostaną własności asymptotyczne estymatora $\hat{\underline{\lambda}}_N^{LS}$ zauważmy, że z założenia 4.4 wynika ciągłość[5] funkcji $\tilde{g}_{N_1}(V_n(\underline{l});\underline{l})$ względem $\underline{l}$, dla każdego $n \in I_1$.

[5] zob. przypis na stronie 15

W konsekwencji również $\hat{Q}_N(\underline{l})$ jest funkcją ciągłą. Z faktu, że Λ jest zbiorem zwartym wynika następnie, że $\hat{Q}_N(\cdot)$ osiąga wartość najmniejszą w Λ i wobec tego estymator $\hat{\underline{\lambda}}_N^{LS}$ jest dobrze określony. Należy jednak podkreślić, że dla skończonej liczby obserwacji N, empiryczna funkcja kryterialna $\hat{Q}_N(\underline{l})$ może posiadać więcej niż jedno minimum w zbiorze Λ. Założenie 4.6 gwarantuje jednak, że graniczna postać stosowanego kryterium (tj. przy $N \to \infty$) posiada jedyne minimum globalne w punkcie $\underline{\lambda}$.

Poniższe twierdzenie formułuje warunki, przy których $\hat{\underline{\lambda}}_N^{LS}$ zbiega według prawdopodobieństwa do wektora odpowiedzi impulsowej $\underline{\lambda}$, gdy liczba obserwacji N rośnie do nieskończoności.

Twierdzenie 4.1 *Jeżeli spełnione są założenia 2.1–2.4 oraz 4.2–4.6, to dla systemu Wienera (2.10)–(2.11) estymator $\hat{\underline{\lambda}}_N^{LS}$ jest zgodnym estymatorem wektora odpowiedzi impulsowej $\underline{\lambda}$ podsystemu dynamicznego, tzn.*

$$\hat{\underline{\lambda}}_N^{LS} \to \underline{\lambda}, \quad gdy\ N \to \infty,$$

według prawdopodobieństwa.

Pełny dowód, skonstruowany w oparciu o twierdzenie 5.7 w Van der Vaart [123] (przytoczone wraz z dowodem w dodatku A.2) zamieszczony jest w pracy Pawlaka *et al.* [93].

4.5.2 Analiza zbieżności estymatora sygnału interakcyjnego

Udowodnimy następujący lemat (por. analiza estymatora $\hat{V}_{n,N}^{Cr}$ w rozdziale 3.):

Lemat 4.1 *Jeżeli spełnione są założenia twierdzenia 4.1, to dla każdego $n \in I$,*

$$\left[\hat{V}_{n,N}^{LS} - V_n\right] \to 0, \quad gdy\ N \to \infty, \tag{4.9}$$

według prawdopodobieństwa.

Dowód. Zgodnie z definicją (4.1) sygnału $\{V_n\}$ oraz definicją (4.7) estymatora $\hat{V}_{n,N}^{LS}$ otrzymujemy na mocy nierówności Cauchy'ego–Schwarza (zob. dodatek B.4)

$$\left|V_n - \hat{V}_{n,N}^{LS}\right| = \left|\left[\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right]^T \underline{X}_n\right| \leq \left\|\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right\| \|\underline{X}_n\| \leq C \left\|\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right\|, \tag{4.10}$$

gdzie $C = (p+1)\max\{|a_f|, |b_f|\}$, zob. zał. 4.1. Rezultat (4.9) wynika zatem bezpośrednio z ograniczoności sygnału $\{\underline{X}_n\}$ i twierdzenia 4.1. ■

4.5.3 Analiza zbieżności estymatora charakterystyki nieliniowej

Dla wykazania zgodności estymatora charakterystyki nieliniowej niezbędne jest posiadanie informacji na temat szybkości zbieżności estymatora odpowiedzi impulsowej części liniowej. Jednak przeprowadzona powyżej analiza własności asymptotycznych estymatora $\hat{\underline{\lambda}}_N^{LS}$ dowodzi jedynie zbieżności algorytmu i nie pozwala na wnioskowanie o jego szybkości zbieżności. Należy podkreślić, że określenie rzędu szybkości zbieżności estymatora odpowiedzi impulsowej podsystemu dynamicznego jest zagadnieniem skomplikowanym i wymaga zastosowania bardziej zaawansowanych narzędzi statystyki matematycznej i teorii prawdopodobieństwa, por. np. Ichimura [58].

Z tego powodu analiza własności teoretycznych algorytmu identyfikacji nieliniowości będzie prowadzona przy dodatkowym, ogólnym, założeniu odnośnie szybkości zbieżności estymatora części liniowej. Załóżmy mianowicie, że zachodzi

$$\left\| \hat{\underline{\lambda}}_N^{LS} - \underline{\lambda} \right\| = O_P\left(a_N\right), \tag{4.11}$$

gdzie a_N jest pewnym ciągiem zbieżnym do zera.

Uwaga 4.2 *Parametryczno–nieparametryczna konstrukcja estymatora $\hat{\underline{\lambda}}_N^{LS}$ (w szczególności występowanie pilotowego estymatora nieparametrycznego w kryterium średniokwadratowym) może powodować, że $\hat{\underline{\lambda}}_N^{LS}$ zbiega do $\underline{\lambda}$ wolniej niż $O_P\left(N^{-1/2}\right)$. W szczególności ciąg a_N może więc przyjąć postać $a_N \sim N^{-\gamma}$, gdzie $\gamma \in \left(0, \frac{1}{2}\right)$.*

Odnośnie konstrukcji estymatora $\hat{g}_N^{LS}(\cdot)$ założymy, że:

Założenie 4.7 *Jądro $K(\cdot)$ jest nieujemną funkcją lipschitzowską (tj. $\exists L < \infty$, $\forall v_1, v_2 \in \mathbf{R}$, $|K(v_1) - K(v_2)| \leq L|v_1 - v_2|$), dla której spełnione są następujące warunki:* $\sup_{v \in \mathbf{R}} K(v) = k_0 < \infty$, $\int_{-\infty}^{\infty} K(v)\,dv = 1$ *oraz* $\lim_{|v| \to \infty} vK(v) = 0$.

Założenie 4.8 *Parametr wygładzania $h = h(N)$ spełnia warunki*

$$h \to 0, Nh \to \infty, \qquad gdy\ N \to \infty, \tag{4.12}$$

oraz

$$a_N/h^2 \to 0, \qquad gdy\ N \to \infty. \tag{4.13}$$

Uwaga 4.3 *Wymaganie (4.12) zawarte w założeniu 4.8 spełnione jest w szczególności gdy $h \sim N^{-\beta}$, gdzie $\beta \in (0,1)$. Z kolei, dla $a_N \sim N^{-\gamma}$ (zob. uwaga 4.2), warunek (4.13) zachodzi gdy $\beta < \gamma/2$.*

Zgodność estymatora $\hat{g}_N^{LS}(\cdot)$ przy powyższych założeniach oraz założeniach twierdzenia 4.1 wykazana jest w następującym twierdzeniu:

Twierdzenie 4.2 *Jeżeli spełnione są założenia twierdzenia 4.1 oraz założenia 4.7 i 4.8, to*

$$\hat{g}_N^{LS}(v) \to g(v), \qquad gdy\ N \to \infty, \tag{4.14}$$

według prawdopodobieństwa w każdym punkcie $v \in \mathbf{R}$, w którym gęstość $f_V(v)$ wewnętrznego sygnału interakcyjnego $\{V_n\}$ jest dodatnia.

Dowód twierdzenia zamieszczony jest w dodatku A.2.

4.6 Identyfikacja systemów Wienera z wykorzystaniem korelacji rangowej

Omówiona powyżej metoda identyfikacji skonstruowana jest w oparciu o kryterium średniokwadratowe $\hat{Q}_N(\cdot)$ utworzone na podstawie dostępnych danych pomiarowych oraz przy użyciu pilotowego estymatora charakterystyki nieliniowej. Proponowana obecnie modyfikacja algorytmu polega na zastosowaniu koncepcji korelacji rangowej, zob. np. Koronacki i Mielniczuk [70], do skonstruowania odmiennej funkcji kryterialnej $\hat{S}_N(\cdot)$, por. Geenens i Simar [21]. Podobnie jak w przypadku kryterium $\hat{Q}_N(\cdot)$, w funkcji $\hat{S}_N(\cdot)$ również

wykorzystuje się pilotowy estymator (4.5) oraz stosuje się wydzielenie ze zbioru obserwacji podzbiorów T_1 i T_2. Rozważany obecnie algorytm umożliwia identyfikację odpowiedzi impulsowej podsystemu dynamicznego poprzez minimalizację funkcji $\hat{S}_N(\cdot)$. W p. 4.6.2 i 4.8.2 wprowadzono odpowiednio koncepcję metody oraz omówiono wyniki eksperymentów numerycznych dla małej i średniej liczby obserwacji wejścia i wejścia systemu. Analiza teoretyczna własności asymptotycznych algorytmu nie została przeprowadzona.

4.6.1 Parametryczna identyfikacja podsystemu dynamicznego z wykorzystaniem korelacji rangowej

Konstruowanie algorytmu rozpoczniemy od rozważenia uproszczonego zadania identyfikacji, w którym na wyjściu systemu nie występuje szum $\{Z_n\}$. W tej sytuacji mamy

$$Y_n = g(V_n). \tag{4.15}$$

Chwilowo założymy również, że nieliniowa charakterystyka $g(\cdot)$ jest znana, a zadanie identyfikacji polega na wyznaczeniu odpowiedzi impulsowej podsystemu dynamicznego na podstawie zbioru obserwacji $\{(X_n, Y_n)\}_{n=1}^{N}$.

4.6.2 Koncepcja metody identyfikacji

Idea omawianej metody opiera się na oczywistym spostrzeżeniu, że zgodnie z (4.15),

$$\{g(V_i) \geq g(V_j)\} \text{ wtedy i tylko wtedy, gdy } \{Y_i \geq Y_j\}, \tag{4.16}$$

dla dowolnych $i, j \in I$. Wykorzystując powyższą obserwację wprowadzimy funkcję

$$\begin{aligned} s(v_1, v_2, y_1, y_2) &= \mathbf{1}[g(v_1) \geq g(v_2)]\,\mathbf{1}[y_1 \geq y_2] \\ &\quad + \mathbf{1}[g(v_1) < g(v_2)]\,\mathbf{1}[y_1 < y_2], \end{aligned}$$

gdzie $v_1, v_2, y_1, y_2 \in \mathbf{R}$ oraz $\mathbf{1}[A]$ oznacza indykator zdarzenia A. W uproszczonym zadaniu identyfikacji, tj. przy założeniu braku szumu wyjściowego, $s(V_i, V_j, Y_i, Y_j) = 1$ dla dowolnych $i, j \in I$. Jednocześnie funkcja $s(V_i(\underline{l}), V_j(\underline{l}), Y_i, Y_j)$, gdzie elementy $V_i(\underline{l})$ i $V_j(\underline{l})$ zdefiniowane są wzorem (4.1), może przyjąć wartość zero jeżeli $\underline{l} \neq \underline{\lambda}$. Własność ta pozwala na wykorzystanie funkcji $s(v_1, v_2, y_1, y_2)$ w konstrukcji kryterium określającego jakość przybliżenia wektora $\underline{\lambda}$ przez dowolny wektor $\underline{l} \in \Lambda$. Odpowiednia funkcja kryterialna zbudowana w oparciu o obserwacje ze zbioru T_2 przyjmuje postać (por. Geenens i Simar [21])

$$\tilde{S}_{N_2}(\underline{l}) = \left[\binom{N_2}{2}\right]^{-1} \sum_{i,j \in I_2; i \neq j} s(V_i(\underline{l}), V_j(\underline{l}), Y_i, Y_j), \tag{4.17}$$

gdzie czynnik wagowy $\binom{N_2}{2}$ oznacza symbol Newtona[6] normalizujący funkcję $\tilde{S}_{N_2}(\cdot)$ tak, aby jej wartości zawarte były w przedziale $[0, 1]$, por. Geenens i Simar [21], Han [45] oraz współczynnik Kendalla (zob. np. Koronacki i Mielniczuk [70, str. 474]). Jest oczywiste, że wartość maksymalną funkcja $\tilde{S}_{N_2}(\underline{l})$ osiąga gdy $\underline{l} = \underline{\lambda}$, bowiem dla każdego $i, j \in I_2$ zachodzi $s(V_i, V_j, Y_i, Y_j) = 1$. Jako estymator odpowiedzi impulsowej $\underline{\lambda}$ przyjąć więc można

$$\underline{\tilde{\lambda}}_{N_2}^{RC} = \arg\max_{\underline{l} \in \Lambda} \tilde{S}_{N_2}(\underline{l}).$$

Podobnie jak w przypadku $\tilde{Q}_{N_2}(\cdot)$, kryterium $\tilde{S}_{N_2}(\cdot)$ nie może być bezpośrednio użyte gdy nieliniowość $g(\cdot)$ nie jest znana. W konsekwencji estymator $\underline{\tilde{\lambda}}_{N_2}^{RC}$ nie może być wyznaczony.

[6] Symbol Newtona $\binom{N_2}{2}$ równy jest liczbie wszystkich kombinacji par elementów ze zbioru I_2.

Wykorzystując podejście omówione poprzednio, zastąpimy w $\tilde{S}_{N_2}(\underline{l})$ charakterystykę nieliniową $g(\cdot)$ pilotowym estymatorem (4.5). W rezultacie otrzymujemy

$$\hat{S}_N(\underline{l}) = \left[\binom{N_2}{2}\right]^{-1} \sum_{i,j\in I_2; i\neq j} \hat{s}_{N_1}(V_i(\underline{l}), V_j(\underline{l}), Y_i, Y_j),$$

gdzie

$$\begin{aligned}\hat{s}_{N_1}(v_1, v_2, y_1, y_2) &= \mathbf{1}[\tilde{g}_{N_1}(v_1; \underline{l}) \geq \tilde{g}_{N_1}(v_2; \underline{l})]\,\mathbf{1}[y_1 \geq y_2] \\ &\quad +\mathbf{1}[\tilde{g}_{N_1}(v_1; \underline{l}) < \tilde{g}_{N_1}(v_2; \underline{l})]\,\mathbf{1}[y_1 < y_2],\end{aligned}$$

i ostatecznie nowy estymator odpowiedzi impulsowej podsystemu dynamicznego przyjmuje postać

$$\hat{\underline{\lambda}}_N^{RC} = \arg\max_{\underline{l}\in\Lambda} \hat{S}_N(\underline{l}).$$

4.6.3 Estymacja sygnału interakcyjnego oraz identyfikacja charakterystyki nieliniowej

Uzyskany w p. 4.6.2 estymator $\hat{\underline{\lambda}}_N^{RC}$, zgodnie z procedurą z p. 4.3, jest następnie wykorzystany do przybliżania sygnału wewnętrznego $\{V_n\}$. Przez analogię do $\hat{V}_{n,N}^{LS}$ otrzymujemy

$$\hat{V}_{n,N}^{RC} = \left(\hat{\underline{\lambda}}_N^{RC}\right)^T \underline{X}_n,$$

gdzie $n \in I$. Następnie, postępując podobnie jak w przypadku estymatora $\hat{g}_N^{LS}(v)$, do identyfikacji nieliniowości można z kolei zastosować estymator

$$\hat{g}_N^{RC}(v) = \frac{\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{RC}}{h}\right)}{\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{RC}}{h}\right)},$$

gdzie $K(\cdot)$ i $h = h(N)$ są odpowiednio funkcją jądra i parametrem wygładzania.

4.6.4 Dyskusja algorytmu w kontekście podejścia wykorzystującego kryterium średniokwadratowe

Otrzymany powyżej estymator $\hat{\underline{\lambda}}_N^{RC}$, podobnie jak uzyskany w p. 4.2.2 estymator $\hat{\underline{\lambda}}_N^{LS}$, jest argumentem, dla którego odpowiednia funkcja celu przyjmuje wartość ekstremalną. W przypadku estymatora zbudowanego w oparciu o metodę najmniejszych kwadratów, wektor $\hat{\underline{\lambda}}_N^{LS}$ *minimalizuje* kryterium $\hat{Q}_N(\cdot)$, a w przypadku metody wykorzystującej koncepcję korelacji rangowej wektor $\hat{\underline{\lambda}}_N^{RC}$ jest punktem, w którym $\hat{S}_N(\cdot)$ przyjmuje wartość *maksymalną*. Pomimo podobnej konstrukcji obu estymatorów, w przypadku estymatora $\hat{\underline{\lambda}}_N^{RC}$ w funkcji $\hat{S}_N(\underline{l})$ następuje uśrednianie $\binom{N_2}{2} = \frac{1}{2}(N_2^2 - N_2)$ elementów losowych, natomiast w $\hat{Q}_N(\cdot)$ uśrednianych jest tylko N_2 wyrazów. Zatem, dla skończonej liczby obserwacji, oszacowanie $\hat{\underline{\lambda}}_N^{RC}$ może posiadać mniejszą wariancję niż oszacowanie $\hat{\underline{\lambda}}_N^{LS}$.

4.7 Podsumowanie wyników teoretycznych

W rozdziale przeprowadzono analizę własności asymptotycznych algorytmu zbudowanego w oparciu o metodę najmniejszych kwadratów. Wykazano zgodność estymatora części liniowej oraz zgodność estymatora nieliniowej charakterystyki $g(\cdot)$ przy założeniu, że znany jest rząd szybkości zbieżności estymatora odpowiedzi impulsowej podsystemu dynamicznego. Warunki, przy których zachodzi zgodność estymatorów umożliwiają zastosowanie metody w przypadku szerokiej klasy sygnałów wejściowych (zob. zał. 4.1) oraz szerokiej klasy charakterystyk nieliniowych występujących w systemie (zob. zał. 4.2).

4.8 Wyniki badań eksperymentalnych

Symulacje komputerowe proponowanych w rozdziale algorytmów przeprowadzono dla trzech systemów Wienera z nieliniowościami $g_1(\cdot)$, $g_2(\cdot)$ i $g_3(\cdot)$ jak w p. 3.7 (zob. wzory (3.36) na str. 32). Podsystem dynamiczny, podobnie jak w rozdziale 3, modelowany był za pomocą odpowiedzi impulsowej (3.35). Rozważane systemy pobudzano sygnałami o następujących gęstościach prawdopodobieństwa (zob. rys. 4.2):

- rozkład jednostajny

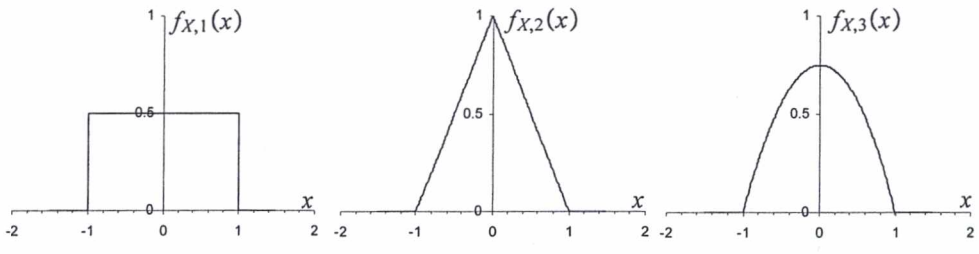
$$f_{X,1}(x) = \begin{cases} \frac{1}{2} & \text{dla} \quad x \in [-1,1] \\ 0 & \text{dla} \quad x \notin [-1,1] \end{cases}, \tag{4.18}$$

- rozkład trójkątny

$$f_{X,2}(x) = \begin{cases} 1-|x| & \text{dla} \quad x \in [-1,1] \\ 0 & \text{dla} \quad x \notin [-1,1] \end{cases}, \tag{4.19}$$

- rozkład Epanechnikova

$$f_{X,3}(x) = \begin{cases} \frac{3}{4}\left(1-x^2\right) & \text{dla} \quad x \in [-1,1] \\ 0 & \text{dla} \quad x \notin [-1,1] \end{cases}. \tag{4.20}$$



Rysunek 4.2: Wykorzystywane w eksperymentach numerycznych gęstości prawdopodobieństwa $f_{X,1}$, $f_{X,2}$ i $f_{X,3}$ sygnału wejściowego $\{X_n\}$

Do wyznaczania ekstremów funkcji kryterialnych $\hat{Q}_N(\cdot)$ i $\hat{S}_N(\cdot)$ zastosowano funkcję `fminsearch`, wchodzącą w skład pakietu obliczeniowego MATLAB i będącą implementacją bezgradientowego algorytmu optymalizacji Neldera–Meada (zob. np. Walters *et. al* [131], Lagarias [74] *et al.*). Sygnał zakłócający wygenerowano z rozkładu jednostajnego o zerowej

wartości oczekiwanej i wariancji dobranej tak, aby NSR= 10% (por. p. 3.7 w rozdziale 3). W symulacjach badano zależność błędów empirycznych $Err(\cdot;\alpha)$ i $MISE(\cdot;\alpha)$, określonych wzorami (3.39) i (3.40) na str. 33, od liczby obserwacji N wejścia i wyjścia systemu (zmienianej od 25 do 500). Zgodnie z założeniami przyjętymi w p. 4.1, multiplikatywna stała α występująca we wzorach (3.39) i (3.40) jest równa 1. Podobnie jak w rozdziale 3 w definicjach błędów $Err(\cdot;1)$ i $MISE(\cdot;1)$ przyjęto następujące wartości stałych R i M: $R = 30$, $M = 1000$. Ponieważ nośnik rozważanych gęstości prawdopodobieństwa jest równy $[-1,1]$, więc sygnał wewnętrzny $\{V_n\}$ posiadał gęstość prawdopodobieństwa o nośniku $[\vartheta_1,\vartheta_2]$, gdzie $\vartheta_1 \approx -2.28$ oraz $\vartheta_2 \approx 2.28$. W pilotowym estymatorze $\tilde{g}_{N_1}(\cdot;\underline{l})$ (zob. wzór (4.5)) zastosowano funkcję jądra określoną wzorem

$$K_0(v) = \begin{cases} \frac{15}{16}\left(1-v^2\right)^2 & \text{dla} \quad v \in [-1,1] \\ 0 & \text{dla} \quad v \notin [-1,1] \end{cases}.$$

W estymatorze $\hat{g}_N^{LS}(\cdot)$ (zob. wzór (4.8)) wykorzystano natomiast jądro gaussowskie, tj. funkcję postaci

$$K(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right).$$

Wyniki eksperymentów wykonanych dla algorytmu wykorzystującego kryterium średniokwadratowe i algorytmu zbudowanego w oparciu o ideę korelacji rangowej przedstawiono odpowiednio w p. 4.8.1 i p. 4.8.2.

4.8.1 Identyfikacja z wykorzystaniem metody najmniejszych kwadratów

Rysunki 4.3 i 4.4 (zob. str. 48) przedstawiają wyniki eksperymentu, tj. zależności błędów $Err\left(\hat{\underline{\lambda}}_N^{LS};1\right)$ i $MISE\left(\hat{g}_N^{LS};1\right)$ od liczby obserwacji N, dla gęstości sygnału wejściowego określonej funkcją $f_{X,1}(\cdot)$. Na rysunkach 4.5, 4.6 (zob. str. 48) oraz 4.7 i 4.8 przedstawiono z kolei rezultaty symulacji komputerowych odpowiednio dla gęstości wejścia $f_{X,2}(\cdot)$ oraz $f_{X,3}(\cdot)$. Otrzymane wyniki pokazują, że w przypadku systemów z nieliniowościami $g_1(\cdot)$ i $g_2(\cdot)$ (tj. nieliniowościami spełniającymi założenie 4.2), pobudzanych sygnałami o gęstościach $f_{X,1}(\cdot)$, $f_{X,2}(\cdot)$ i $f_{X,3}(\cdot)$, błędy empiryczne maleją wraz ze wzrostem liczby obserwacji N. Podobne wartości błędów uzyskano również dla nieciągłej nieliniowości $g_3(\cdot)$ nie spełniającej wymagania 4.2 co, nie przesądzając o zbieżności, sugeruje że przyjęte w rozdziale założenia dotyczące charakterystyki $g(\cdot)$ są jedynie warunkami wystarczającymi i przypuszczalnie mogą być osłabione. Rysunki 4.3, 4.5 oraz 4.7 prowadzą ponadto do wniosku, że charakter gęstości wejścia $f_X(\cdot)$ nie ma istotnego wpływu na zachowanie się estymatorów części liniowej systemu. Przykładowo, dyskutowane błędy $Err\left(\hat{\underline{\lambda}}_N^{LS};1\right)$ przyjmują wartości mniejsze niż 10%, gdy liczba obserwacji N jest większa od 250, niezależnie od zastosowanej gęstości wejścia $f_X(\cdot)$. Podobnie, w przypadku błędów empirycznych $MISE\left(\hat{g}_N^{LS};1\right)$ uzyskane wartości (zob. rys 4.4, 4.6, 4.8) są mniejsze niż 10% gdy liczba obserwacji jest większa od 200.

4.8.2 Identyfikacja z wykorzystaniem korelacji rangowej

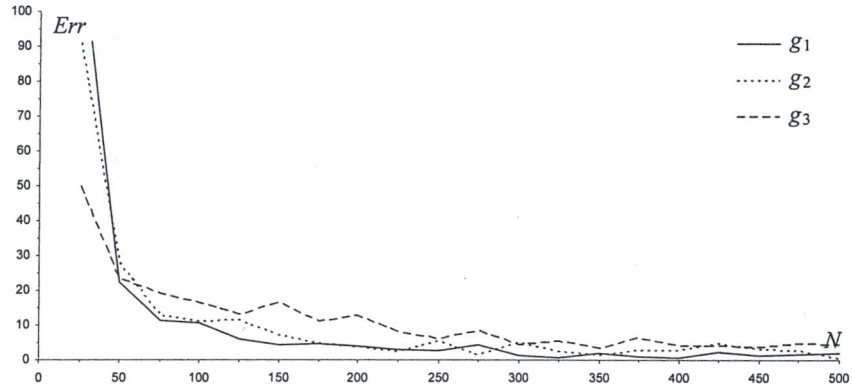
Rezultaty eksperymetów wykonanych dla algorytmu zbudowanego w oparciu o koncepcję korelacji rangowej zamieszczone są na rysunkach 4.9–4.14 (zob. str. 50–51). Rysunki 4.9, 4.10 przedstawiają błędy $Err\left(\hat{\underline{\lambda}}_N^{RC};1\right)$ i $MISE\left(\hat{g}_N^{RC};1\right)$ dla sygnału wejściowego o gęstości $f_{X,1}(\cdot)$. Podobnie rysunki 4.11, 4.12 oraz 4.13, 4.14 obrazują wyniki badań odpowiednio dla gęstości wejścia $f_{X,2}(\cdot)$ oraz $f_{X,3}(\cdot)$. Otrzymane wyniki pokazują, że

badane błędy maleją wraz ze wzrostem liczby obserwacji N. W przypadku małej liczby obserwacji (tj. mniejszej od 200) wartości błędów $Err\left(\hat{\underline{\lambda}}_N^{RC};1\right)$ i $MISE\left(\hat{g}_N^{RC};1\right)$ są jednak większe od błędów uzyskanych w badaniach algorytmu wykorzystującego kryterium średniokwadratowe. W konsekwencji, w przypadku małej liczby obserwacji systemu, estymatory $\hat{\underline{\lambda}}_N^{LS}$ i $\hat{g}_N^{LS}(\cdot)$ posiadają lepsze własności empiryczne od estymatorów $\hat{\underline{\lambda}}_N^{RC}$ i $\hat{g}_N^{RC}(\cdot)$ niezależnie od charakteru nieliniowości występującej w systemie oraz typu gęstości sygnału wejściowego $\{X_n\}$. W odróżnieniu od estymatorów $\hat{\underline{\lambda}}_N^{LS}$ i $\hat{g}_N^{LS}(\cdot)$ omawiane obecnie estymatory $\hat{\underline{\lambda}}_N^{RC}$ i $\hat{g}_N^{RC}(\cdot)$ są ponadto bardziej wrażliwe na charakter nieliniowości występującej w systemie. Otrzymane wyniki (zob. rys. 4.9, 4.11, 4.13) pokazują, że najmniejsze wartości błędów estymacji części liniowej systemu uzyskano dla gładkiej i monotonicznej nieliniowości $g_1(\cdot)$. Nieco większe wartości otrzymano dla gładkiej, ale niemonotonicznej nieliniowości $g_2(\cdot)$. W przypadku odcinkami stałej i nieciągłej nieliniowości $g_3(\cdot)$ błędy $Err\left(\hat{\underline{\lambda}}_N^{RC};1\right)$ przyjmowały natomiast wartości największe.

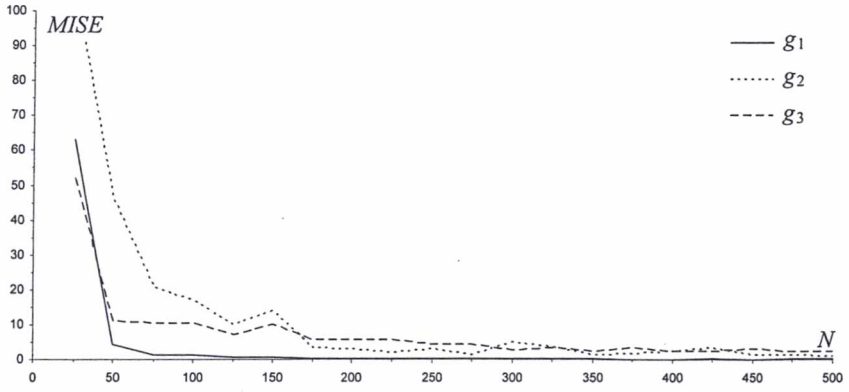
4.9 Podsumowanie i wnioski

W rozdziale zaprezentowano dwie metody identyfikacji systemów Wienera. Algorytm zbudowany w oparciu o metodę najmniejszych kwadratów dedykowany jest systemom pobudzanym ograniczonym sygnałem losowym typu *i.i.d.* o ciągłej funkcji gęstości prawdopodobieństwa. Metoda pozwala na identyfikację obiektów, w których nieliniowość $g(\cdot)$ spełnia warunek ciągłości. W celu estymacji odpowiedzi impulsowej liniowego podsystemu dynamicznego niezbędne jest wyznaczenie minimum nieliniowej funkcji kryterialnej $\hat{Q}_N(\cdot)$. Z tego powodu praktyczne zastosowanie algorytmu wymaga użycia dodatkowych technik optymalizacji nieliniowej. Podobna sytuacja ma miejsce w przypadku algorytmu zbudowanego w oparciu o koncepcję korelacji rangowej, w którym estymator odpowiedzi impulsowej podsystemu dynamicznego definiowany jest jako maksimum funkcji kryterialnej $\hat{S}_N(\cdot)$.

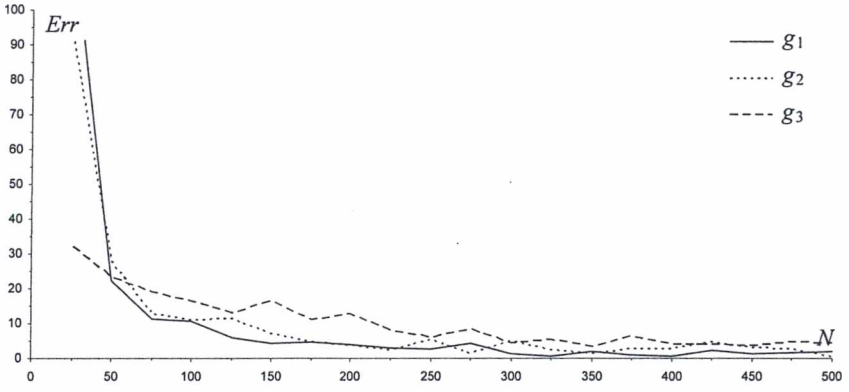
W rozdziale przeprowadzono analizę teoretyczną własności asymptotycznych estymatorów $\hat{\underline{\lambda}}_N^{LS}$ i $\hat{g}_N^{LS}(\cdot)$ oraz omówiono wyniki eksperymentów numerycznych przeprowadzonych dla algorytmu wykorzystującego kryterium średniokwadratowe oraz algorytmu zbudowanego w oparciu o koncepcję korelacji rangowej.



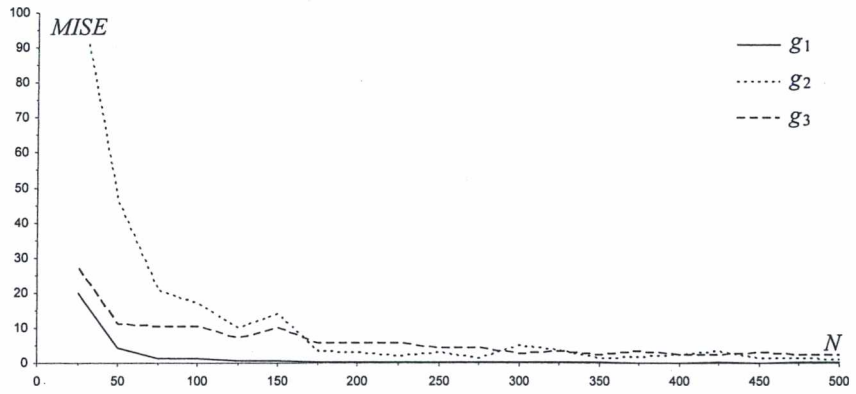
Rysunek 4.3: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{LS}; 1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,1}$



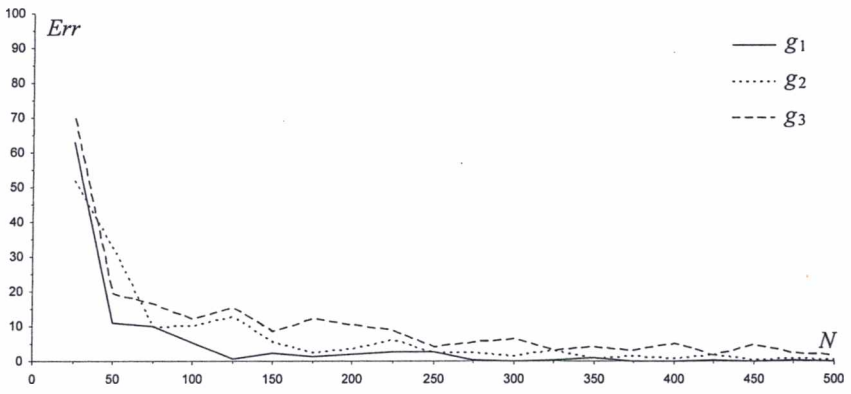
Rysunek 4.4: Błędy estymacji $MISE\left(\hat{g}_N^{LS}; 1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,1}$



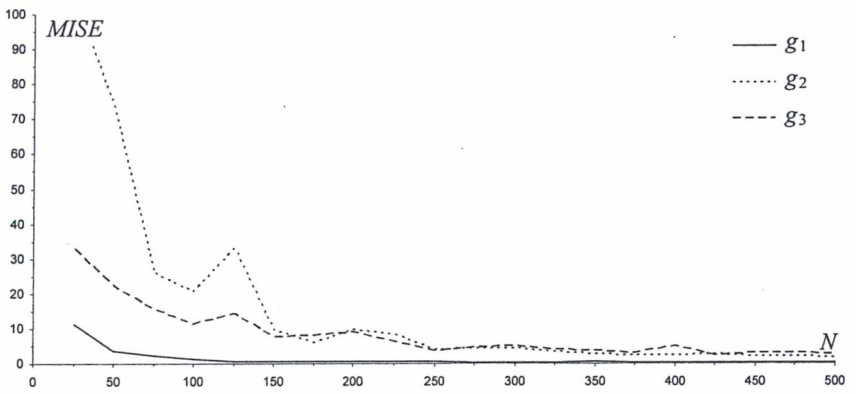
Rysunek 4.5: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{LS}; 1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,2}$



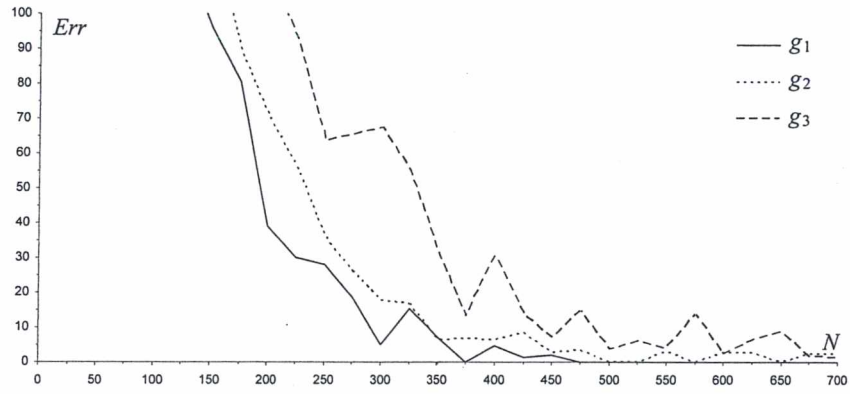
Rysunek 4.6: Błędy estymacji $MISE\left(\hat{g}_N^{LS};1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,2}$



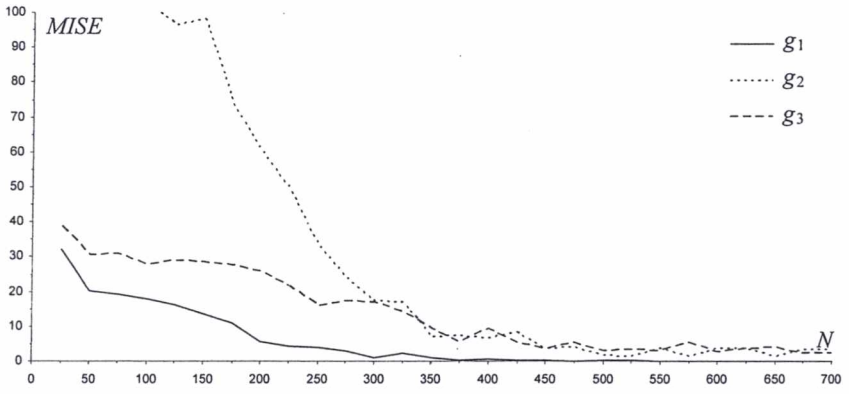
Rysunek 4.7: Błędy estymacji $Err(\hat{\lambda}_N^{LS};1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,3}$



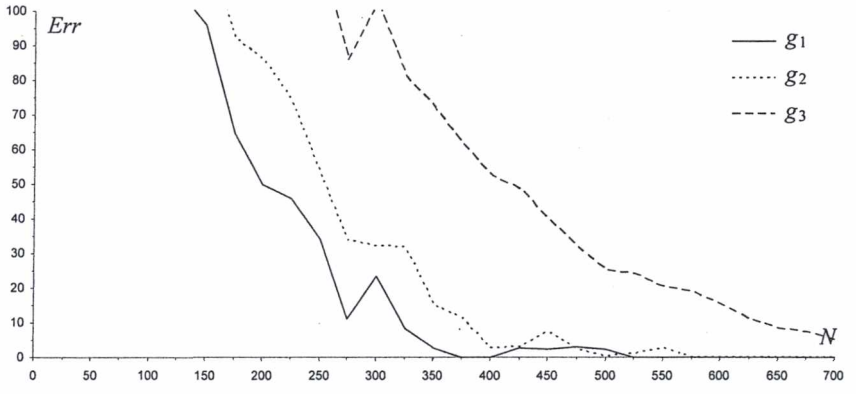
Rysunek 4.8: Błędy estymacji $MISE\left(\hat{g}_N^{LS};1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,3}$



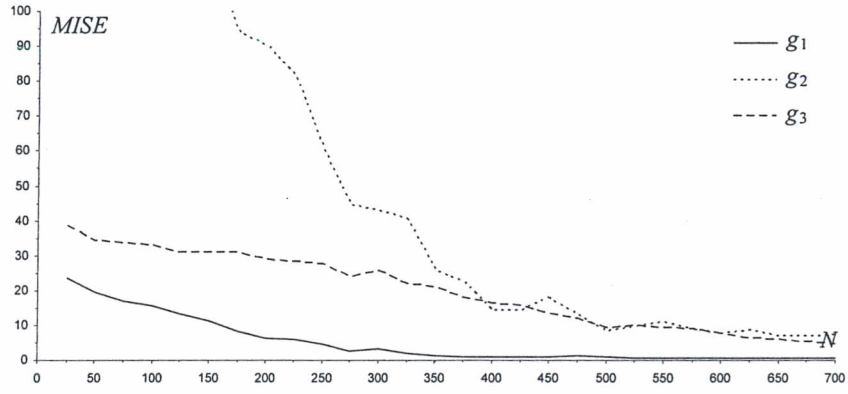
Rysunek 4.9: Błędy estymacji $Err(\hat{\lambda}_N^{RC}; 1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,1}$



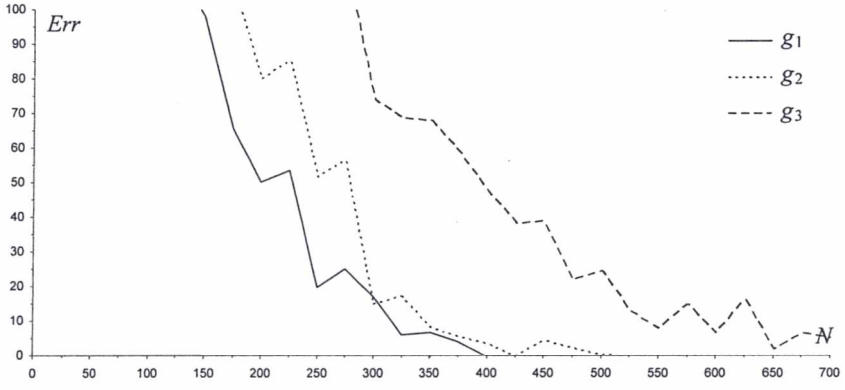
Rysunek 4.10: Błędy estymacji $MISE\left(\hat{g}_N^{RC}; 1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,1}$



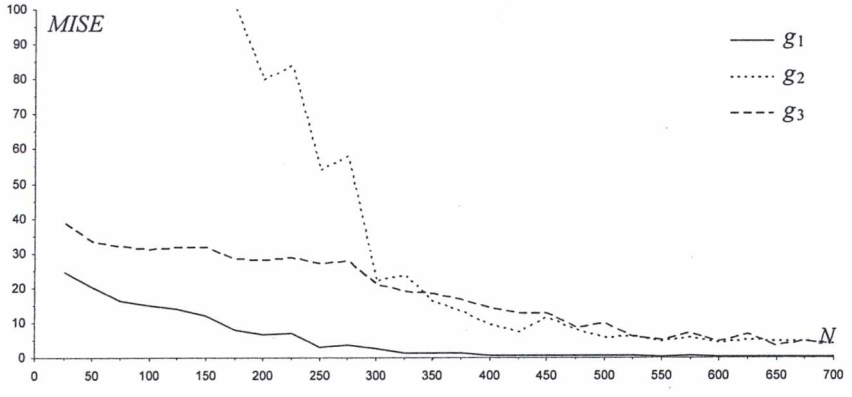
Rysunek 4.11: Błędy estymacji $Err(\hat{\lambda}_N^{RC}; 1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,2}$



Rysunek 4.12: Błędy estymacji $MISE\left(\hat{g}_N^{RC};1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,2}$



Rysunek 4.13: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{RC};1)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,3}$



Rysunek 4.14: Błędy estymacji $MISE\left(\hat{g}_N^{RC};1\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_2, g_3 i gęstości sygnału wejściowego $f_{X,3}$

Rozdział 5

Identyfikacja systemów Wienera metodą uśrednionej pochodnej

W tym rozdziale przedstawiono algorytm identyfikacji systemów Wienera ze skończoną pamięcią i niegaussowskim sygnałem wejściowym, skonstruowany w oparciu o koncepcję estymacji uśrednionej pochodnej (*ang. Average Derivative estimate*, zob. np. Powell *et al.* [95], Härdle i Stoker [47]). W odróżnieniu od metody zaproponowanej w rozdziale 4, algorytm pozwala na identyfikację podsystemu dynamicznego bez konieczności stosowania dodatkowych procedur optymalizacji nieliniowej, a więc jest prostszy pod względem obliczeniowym. Nie jest również dokonywana dekompozycja zbioru pomiarów.

Wprowadzone w rozdziale estymatory części liniowej, sygnału interakcyjnego i części nieliniowej systemu oznaczane będą indeksem „*AD*" (od *ang. Average Derivative*), gdy gęstość sygnału wejściowego $\{X_n\}$ jest *a priori* znana oraz indeksem „*fAD*", gdy gęstość $f_X(\cdot)$ jest nieznana.

5.1 Sformułowanie problemu

Ze względu na charakter zastosowanego podejścia kluczowymi założeniami odnośnie konstrukcji systemu i sygnału wejściowego są wymagania dotyczące gładkości nieliniowości $g(\cdot)$ i gęstości prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego $\{X_n\}$. Przyjmuje się w szczególności, że:

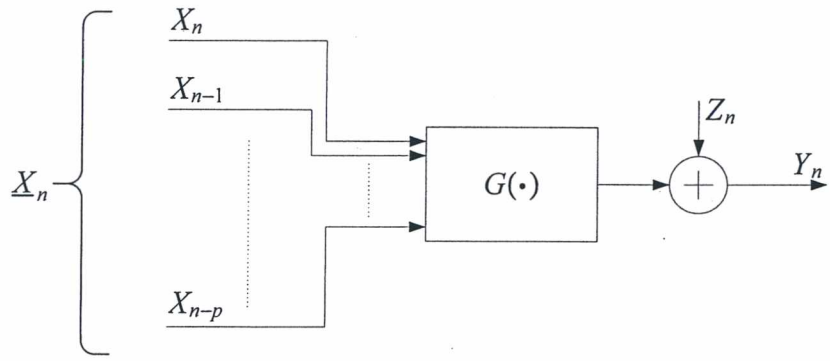
Założenie 5.1 *Funkcja gęstości prawdopodobieństwa* $f_X(\cdot)$ *sygnału wejściowego* $\{X_n\}$ *posiada ograniczony nośnik,* $\operatorname{supp} f_X = [a_f, b_f]$, $|a_f|, |b_f| < \infty$ *oraz ciągłą pochodną* $f'_X(x)$ *dla* $x \in [a_f, b_f]$. *Ponadto* $f_X(a_f) = f_X(b_f) = 0$.

Założenie 5.2 *Nieliniowa charakterystyka* $g(\cdot)$ *posiada ciągłą pochodną* $g'(\cdot)$.

5.2 Reprezentacja systemu Wienera za pomocą statycznego nieliniowego systemu typu MISO

Niech $\underline{X}_n = (X_n, X_{n-1}, \ldots, X_{n-p})^T$ będzie $(p+1)$–elementowym wektorem utworzonym z kolejnych obserwacji sygnału wejściowego $\{X_n\}$. Wykorzystując powyższą notację, system Wienera (2.10)–(2.11) można zapisać w równoważnej postaci za pomocą równania

$$Y_n = G(\underline{X}_n) + Z_n, \tag{5.1}$$



Rysunek 5.1: Reprezentacja systemu Wienera ze skończoną pamięcią przy pomocy statycznego systemu nieliniowego o strukturze MISO

gdzie funkcja $G : \mathbf{R}^{p+1} \to \mathbf{R}$ dana jest wzorem

$$G(\underline{x}) = g\left(\underline{\lambda}^T \underline{x}\right), \qquad \underline{x} \in \mathbf{R}^{p+1}, \tag{5.2}$$

a wektor $\underline{\lambda} \in \mathbf{R}^{p+1}$ utworzony jest z elementów odpowiedzi impulsowej podsystemu dynamicznego, tj. $\underline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_p)^T$. Reprezentacja (5.1) pozwala przedstawić rozważany system Wienera jako statyczny system nieliniowy o strukturze MISO z wektorowym wejściem $\{\underline{X}_n\}$ i skalarnym wyjściem $\{Y_n\}$, (zob. rys. 5.1).

Wykorzystując zbiór obserwacji wejścia i wyjścia $\{(X_n, Y_n)\}_{n=1}^N$ systemu (2.10)–(2.11) skonstruujemy analogiczny zbiór pomiarów wejścia i wyjścia, właściwy dla reprezentacji (5.1). Zbiór ten przyjmuje postać

$$T_G = \{(\underline{X}_n, Y_n)\}_{n \in I}, \tag{5.3}$$

gdzie $I = \{p+1, p+2, \ldots, N\}$ jest $(N-p)$–elementowym zbiorem indeksów (standardowo oznaczymy również $N_I = \#\{I\} = N - p$).

W celu wprowadzenia koncepcji omawianej metody identyfikacji konieczne jest określenie podstawowych własności statystycznych sygnału $\{\underline{X}_n\}$. Zgodnie z przyjętymi założeniami odnośnie wejścia $\{X_n\}$ (zob. założenie 2.2, str. 22), sygnał $\{\underline{X}_n\}$ jest ciągiem wektorów losowych o tym samym rozkładzie, ale nie jest ciągiem typu *i.i.d.* Z kolei elementy składowe każdego wektora $\underline{X}_n$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie. W konsekwencji prawdziwa jest poniższa własność:

Własność 5.1 *Sygnał $\{\underline{X}_n\}$ jest ciągiem wektorów losowych o funkcji gęstości prawdopodobieństwa $f_{\underline{X}} : \mathbf{R}^{p+1} \to \mathbf{R}$ danej wzorem*

$$f_{\underline{X}}(\underline{x}) = \prod_{j=1}^{p+1} f_X\left(\underline{x}_{(j)}\right), \qquad \underline{x} \in \mathbf{R}^{p+1}, \tag{5.4}$$

gdzie $\underline{x}_{(j)}$ oznacza j–ty element wektora[1] $\underline{x}$.

Z założenia 5.1 oraz wzoru (5.4) wynika, że nośnik $\Omega = \operatorname{supp} f_{\underline{X}}$ funkcji $f_{\underline{X}}(\cdot)$ jest kostką w przestrzeni $\mathbf{R}^{p+1}$ o bokach $[a_f, b_f]$. Ponadto $f_{\underline{X}}(\cdot)$ jest funkcją różniczkowalną w zbiorze Ω.

[1] W rozdziale zapis $\underline{w}_{(j)}$ oznaczać będzie ogólnie j–ty element wektora $\underline{w}$.

5.3 Idea metody uśrednionej pochodnej

Oznaczmy przez $d_{G,l}(\underline{x})$ i $d_{f,l}(\underline{x})$ następujące pochodne cząstkowe

$$d_{G,l}(\underline{x}) = \frac{\partial G(\underline{x})}{\partial \underline{x}_{(l)}} \quad \text{oraz} \quad d_{f,l}(\underline{x}) = \frac{\partial f_{\underline{X}}(\underline{x})}{\partial \underline{x}_{(l)}}, \tag{5.5}$$

gdzie $l \in \{1, 2, \ldots, p+1\}$ oraz $\underline{x} \in \mathbf{R}^{p+1}$. Niech dalej $D_G(\underline{x})$ i $D_f(\underline{x})$ będą odpowiednio gradientami funkcji $G(\cdot)$ i $f_{\underline{X}}(\cdot)$ w punkcie $\underline{x}$, tj.

$$\begin{aligned} D_G(\underline{x}) &= (d_{G,1}(\underline{x}), d_{G,2}(\underline{x}), \ldots, d_{G,p+1}(\underline{x}))^T, \\ D_f(\underline{x}) &= (d_{f,1}(\underline{x}), d_{f,2}(\underline{x}), \ldots, d_{f,p+1}(\underline{x}))^T. \end{aligned}$$

Na mocy założenia 5.2, gradient $D_G(\underline{x})$ jest dobrze określony w każdym punkcie $\underline{x} \in \mathbf{R}^{p+1}$. Podobnie, zgodnie z założeniem 5.1, gradient $D_f(\underline{x})$ jest dobrze określony w każdym punkcie $\underline{x} \in \Omega$.

Podstawą omawianej metody identyfikacji jest obserwacja, że na mocy reprezentacji (5.2) prawdziwa jest zależność

$$D_G(\underline{x}) = g'(\underline{\lambda}^T \underline{x})\,\underline{\lambda}, \tag{5.6}$$

gdzie $g'(x)$ oznacza pochodną funkcji $g(\cdot)$ w punkcie $x \in \mathbf{R}$. Wzór (5.6) prowadzi do kluczowego dla dalszych rozważań spostrzeżenia:

Wniosek 5.1 *Gradient $D_G(\underline{x})$ w dowolnym punkcie $\underline{x} \in \mathbf{R}^{p+1}$ jest proporcjonalny do wektora odpowiedzi impulsowej $\underline{\lambda}$.*

Powyższy wniosek sugeruje więc, aby jako oszacowanie przeskalowanej odpowiedzi impulsowej podsystemu dynamicznego przyjąć estymator gradientu $D_G(\underline{x})$ w dowolnym, ustalonym punkcie $\underline{x}$. Zauważmy jednak, że dowolność wyboru argumentu $\underline{x}$ prowadzi w szczególności do następującej równości

$$D_G(\underline{X}_n) = g'(\underline{\lambda}^T \underline{X}_n)\,\underline{\lambda}, \quad n \in I,$$

i wobec tego prawdziwa jest również zależność

$$E\{D_G(\underline{X}_n)\} = \kappa_0 \underline{\lambda}, \tag{5.7}$$

gdzie $\kappa_0 = E\{g'(\underline{\lambda}^T \underline{X}_n)\}$. Z równania (5.7) wynika, że wartość oczekiwana $E\{D_G(\underline{X}_n)\}$ jest proporcjonalna do wektora $\underline{\lambda}$ ze stałym współczynnikiem proporcjonalności równym κ_0. Zatem identyfikacja podsystemu dynamicznego może być zrealizowana poprzez estymację wyrażenia $E\{D_G(\underline{X}_n)\}$. Niestety, ze względu na nieznajomość nieliniowości $g(\cdot)$, gradient D_G również nie jest znany i w konsekwencji bezpośrednie zastosowanie równania (5.7) nie jest możliwe. Pewnym rozwiązaniem, umożliwiającym ominięcie wspomnianego problemu, jest utworzenie estymatora $\hat{D}_{G,N}$ gradientu D_G i wstawienie go do (5.7) w miejsce D_G, a następnie estymacja wartości oczekiwanej $E\left\{\hat{D}_{G,N}(\underline{X}_n)\right\}$. Podejście to posiada jednak istotną wadę w postaci konieczności szacowania gradientu nieznanej, wieloargumentowej funkcji $G(\cdot)$.

W celu ominięcia powyższych trudności wprowadzimy do równania (5.6) czynnik wagowy $f_{\underline{X}}(\underline{x})$, tj.

$$f_{\underline{X}}(\underline{x})\, D_G(\underline{x}) = f_{\underline{X}}(\underline{x})\, g'(\underline{\lambda}^T \underline{x})\, \underline{\lambda}. \tag{5.8}$$

Równość (5.8) prowadzi teraz do wniosku, że

$$E\left\{f_{\underline{X}}\left(\underline{X}_n\right) D_G\left(\underline{X}_n\right)\right\} = \kappa_1 \underline{\lambda}, \tag{5.9}$$

gdzie $\kappa_1 = E\left\{f_{\underline{X}}\left(\underline{X}_n\right) g'\left(\underline{\lambda}^T \underline{X}_n\right)\right\}$.

Dalsze rozważania poprzedzone zostaną następującym lematem (por. lemat 2.1 w Powell *et al.* [95]).

Lemat 5.1 *Dla nieliniowej charakterystyki $g(\cdot)$ oraz funkcji gęstości prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego $\{X_n\}$ spełniających odpowiednio założenia 5.1 i 5.2, zachodzi następujący związek*

$$E\left\{f_{\underline{X}}\left(\underline{X}_n\right) D_G\left(\underline{X}_n\right)\right\} = -2E\left\{D_f\left(\underline{X}_n\right) Y_n\right\}. \tag{5.10}$$

Dowód lematu zamieszczony jest w dodatku A.3.

Równość (5.10) pozwala przepisać wzór (5.9) w znacznie dogodniejszej dla dalszych rozważań formie, wykorzystującej obok sygnału wejściowego $\{\underline{X}_n\}$, także sygnał wyjściowy $\{Y_n\}$, tj.

$$E\left\{Y_n D_f\left(\underline{X}_n\right)\right\} = \kappa \underline{\lambda}, \tag{5.11}$$

gdzie

$$\kappa = -\kappa_1/2.$$

W reprezentacji skalarnej zależność (5.11) przyjmuje postać $p+1$ równań

$$E\left\{Y_n d_{f,i+1}\left(\underline{X}_n\right)\right\} = \kappa\lambda_i, \qquad i = 0, 1, \ldots, p. \tag{5.12}$$

Zauważmy teraz, że zgodnie z (5.4) oraz (5.5),

$$d_{f,i}(\underline{x}) = f'_X\left(\underline{x}_{(i)}\right) \prod_{k=1, k\neq i}^{p+1} f_X\left(\underline{x}_{(k)}\right), \qquad i \in \{1, 2, \ldots, p+1\}. \tag{5.13}$$

Zatem, gdy gęstość prawdopodobieństwa $f_X(\cdot)$ wejścia $\{X_n\}$ jest znana, możliwe jest bezpośrednie wyznaczenie pochodnych cząstkowych $d_{f,i}(\cdot), i = 1, 2, \ldots, p+1$. W tym przypadku zależność (5.12) pozwala więc na prostą estymację elementu $\kappa\lambda_i$ poprzez estymację odpowiedniej wartości oczekiwanej występującej we wzorze (5.12). Zagadnienie to omawiane jest poniżej.

5.4 Identyfikacja systemów Wienera w przypadku znanej gęstości prawdopodobieństwa sygnału wejściowego

5.4.1 Konstrukcja estymatora podsystemu dynamicznego

Zakładamy znajomość *a priori* funkcji gęstości prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego $\{X_n\}$. Proponowany estymator konstruowany jest w oparciu o zależność (5.12). Naturalnym oszacowaniem wartości oczekiwanej po lewej stronie równania (5.12) jest średnia z próby utworzona na podstawie elementów zbioru T_G, zob. wzór (5.3). W rezultacie estymator przeskalowanej odpowiedzi impulsowej $\{\kappa\lambda_i\}_{i=0}^{p}$ przyjmuje postać

$$\hat{\lambda}_{i,N}^{AD} = \frac{1}{N_I} \sum_{n\in I} Y_n d_{f,i+1}\left(\underline{X}_n\right), \qquad i = 0, 1, \ldots, p, \tag{5.14}$$

gdzie funkcja $d_{f,i+1}(\cdot)$ określona jest wzorem (5.13).

Własności asymptotyczne estymatorów $\hat{\lambda}_{i,N}^{AD}, i = 0, 1, \ldots, p$ oraz wyniki badań empirycznych dla małej i umiarkowanej liczby obserwacji omówiono odpowiednio w p. 5.4.4 i p. 5.7.

5.4.2 Estymacja sygnału interakcyjnego

Estymator przeskalowanego sygnału interakcyjnego $\{\kappa V_n\}$ konstruowany jest podobnie jak w przypadku metod omawianych w rozdziałach 3. i 4. Obecnie jest on zatem zbudowany w oparciu o estymatory $\hat{\lambda}_{i,N}^{AD}$, $i = 0, 1, \ldots, p$ i wyraża się wzorem

$$\hat{V}_{n,N}^{AD} = \sum_{i=0}^{p} \hat{\lambda}_{i,N}^{AD} X_{n-i}. \tag{5.15}$$

$$V_n = \underline{\lambda}^T \underline{X}_n.$$

Ze względu na ograniczoną liczbę dostępnych obserwacji wejścia i wyjścia systemu, estymator (5.15) może być wyznaczony dla $n \in I$.

5.4.3 Nieparametryczna identyfikacja charakterystyki nieliniowej

W przypadku znanej gęstości prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego, estymator przeskalowanej charakterystyki nieliniowej $g(\cdot/\kappa)$ konstruowany jest w oparciu o ciąg oszacowań $\{\hat{V}_{n,N}^{AD}\}_{n\in I}$ sygnału $\{\kappa V_n\}$ i określony jest wzorem

$$\hat{g}_N^{AD}(v) = \frac{\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{AD}}{h}\right)}{\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{AD}}{h}\right)}. \tag{5.16}$$

Podobnie jak w przypadku estymatorów $\hat{g}_N^{Cr}(\cdot)$ i $\hat{g}_N^{LS}(\cdot)$ (zob. rozdziały 3. i 4.), $K(\cdot)$ i $h = h(N)$ są odpowiednio funkcją jądra i parametrem wygładzania zależnym od liczby obserwacji N.

5.4.4 Własności asymptotyczne algorytmu

Analiza zbieżności estymatora odpowiedzi impulsowej podsystemu dynamicznego

Poniższe twierdzenie podaje warunki, przy których estymator $\hat{\lambda}_{i,N}^{AD}$ dany wzorem (5.14) zbiega do $\kappa\lambda_i, i = 0, 1, \ldots, p$ według prawdopodobieństwa z szybkością zbieżności rzędu $O_P\left(N^{-1/2}\right)$, tj. z szybkością typową dla metod parametrycznych.

Twierdzenie 5.1 *Jeżeli dla systemu Wienera (2.10)-(2.11) spełnione są założenia 2.1-2.4 oraz 5.1, 5.2, to dla każdego $i \in \{0, 1, 2, \ldots, p\}$ estymator $\hat{\lambda}_{i,N}^{AD}$ zbiega według prawdopodobieństwa do $\kappa\lambda_i$, gdy $N \to \infty$. Ponadto*

$$\hat{\lambda}_{i,N}^{AD} = \kappa\lambda_i + O_P\left(N^{-1/2}\right).$$

Dowód. Niech $\eta_{i,n} = Y_n d_{f,i+1}(\underline{X}_n)$ dla każdego $n \in I$ oraz $i \in \{0, 1, \ldots, p\}$. Wtedy estymator $\hat{\lambda}_{i,N}^{AD}$ można zapisać w postaci

$$\hat{\lambda}_{i,N}^{AD} = \frac{1}{N_I} \sum_{n\in I} \eta_{i,n}.$$

Ponieważ $\{X_n\}$ i $\{Z_n\}$ są stacjonarnymi ciągami zmiennych losowych, więc $\{\eta_{i,n}\}_{n\in I}$ również jest stacjonarnym ciągiem zmiennych losowych. Zatem wartość oczekiwana estymatora (5.14) jest równa

$$E\left\{\hat{\lambda}_{i,N}^{AD}\right\} = E\left\{\eta_{i,n}\right\} = E\left\{Y_n d_{f,i+1}\left(\underline{X}_n\right)\right\},$$

co na mocy równości (5.12) prowadzi do wniosku, że

$$E\left\{\hat{\lambda}_{i,N}^{AD}\right\} = \kappa\lambda_i,$$

tzn. $\hat{\lambda}_{i,N}^{AD}$ jest nieobciążonym estymatorem elementu $\kappa\lambda_i$. Analizę wariancji estymatora $\hat{\lambda}_{i,N}^{AD}$ rozpoczniemy od wykazania że $Var\left\{\eta_{i,n}\right\} < \infty$. Ze względu na stochastyczną niezależność sygnału $\{X_n\}$ i szumu $\{Z_n\}$ oraz wzór (5.1), prawdziwa jest następująca nierówność

$$Var\left\{\eta_{i,n}\right\} \leq 2Var\left\{G\left(\underline{X}_n\right) d_{f,i+1}\left(\underline{X}_n\right)\right\} + 2\sigma_Z^2 Var\left\{d_{f,i+1}\left(\underline{X}_n\right)\right\}. \tag{5.17}$$

Ponieważ $G(\cdot)$ jest funkcją ciągłą (na mocy założenia 5.2), a $d_{f,i+1}(\cdot)$ jest funkcją ograniczoną (na mocy założenia 5.1), więc wariancje po prawej stronie nierówności (5.17) są skończone i tym samym $Var\left\{\eta_{i,n}\right\} < \infty$. Zauważmy następnie, że dla dowolnych indeksów $k, l \in I$ zmienne losowe $\eta_{i,k}$ i $\eta_{i,l}$ są stochastycznie niezależne, jeżeli $|k-l| > p$. Zatem na mocy lematu A.1 w dodatku A.1 otrzymujemy

$$Var\left\{\hat{\lambda}_{i,N}^{AD}\right\} \leq \frac{p+1}{N_I} Var\left\{\eta_{i,n}\right\}$$

i wobec tego

$$Var\left\{\hat{\lambda}_{i,N}^{AD}\right\} = O\left(N^{-1}\right),$$

co kończy dowód twierdzenia. ■

Analiza zbieżności estymatora sygnału interakcyjnego

Poniższy lemat podaje warunki, przy których estymator $\hat{V}_{n,N}^{AD}$ zbiega według prawdopodobieństwa do κV_n.

Lemat 5.2 *Jeżeli spełnione są założenia twierdzenia 5.1, to dla każdego $n \in I$ estymator $\hat{V}_{n,N}^{AD}$ jest zgodnym estymatorem elementu κV_n oraz*

$$\hat{V}_{n,N}^{AD} = \kappa V_n + O_P\left(N^{-1/2}\right). \tag{5.18}$$

Ponadto

$$\sup_{n\in I}\left|\hat{V}_{n,N}^{AD} - \kappa V_n\right| = O_P\left(N^{-1/2}\right). \tag{5.19}$$

Dowód. Rząd szybkości zbieżności w (5.18) wynika z definicji sygnału $\{V_n\}$, postaci estymatora $\hat{V}_{n,N}^{AD}$ oraz z szybkości zbieżności $\hat{\lambda}_{i,N}^{AD}$ do $\kappa\lambda_i$, (por. dowód lematu 3.1 na str. 28). Odnośnie (5.19) zauważmy, że

$$\sup_{n\in I}\left|\hat{V}_{n,N}^{AD} - \kappa V_n\right| \leq \sup_{1\leq n\leq N}|X_n| \sum_{i=0}^{p}\left|\hat{\lambda}_{i,N}^{AD} - \kappa\lambda_i\right|.$$

Zgodnie z założeniem 5.1, ciąg $\left\{\sup_{1\leq n\leq N}|X_n|\right\}$ jest ograniczony z prawdopodobieństwem jeden, bowiem $P\left\{\sup_{1\leq n\leq N}|X_n| \geq c_0\right\} = 0$, gdzie $c_0 = \max\{|a_f|, |b_f|\}$. Zbieżność (5.19) wynika zatem z twierdzenia 5.1. ■

Analiza zbieżności estymatora charakterystyki nieliniowej

Postępując podobnie jak w rozdziale 3. przyjmiemy następujące założenia o funkcji jądra $K(\cdot)$ i parametrze wygładzania h w estymatorze nieliniowości $\hat{g}_N^{AD}(\cdot)$:

Założenie 5.3 *Jądro $K(\cdot)$ jest nieujemną funkcją lipschitzowską (tj. $\exists L < \infty$, $\forall v_1, v_2 \in \mathbf{R}$, $|K(v_1) - K(v_2)| \leq L|v_1 - v_2|$), dla której spełnione są następujące warunki:* $\sup_{v\in\mathbf{R}} K(v) = k_0 < \infty$, $\int_{-\infty}^{\infty} K(v)\,dv = 1$ *oraz* $\lim_{|v|\to\infty} vK(v) = 0$.

Założenie 5.4 *Zbieżny do zera ciąg liczbowy $h = h(N)$ posiada rząd szybkości zbieżności $N^{-1/5}$, tj. $h \sim N^{-1/5}$.*

Warunki przy których estymator $\hat{g}_N^{AD}(\cdot)$ zbiega do przeskalowanej nieliniowości $g(\cdot/\kappa)$ podaje poniższe twierdzenie.

Twierdzenie 5.2 *Jeżeli spełnione są założenia twierdzenia 5.1 oraz założenia 5.3, 5.4 i dodatkowo $\kappa \neq 0$, to dla systemu Wienera (2.10)-(2.11) zachodzi zbieżność*

$$\hat{g}_N^{AD}(v) \to g(v/\kappa) \quad gdy\ N \to \infty \tag{5.20}$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, w którym $f_V(v/\kappa) > 0$.

Dowód. Dowód przebiega analogicznie do dowodu twierdzenia 3.2. Przyjmiemy standardowe oznaczenia

$$\hat{r}_N^{AD}(v) = \frac{1}{N_I h}\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{AD}}{h}\right), \qquad \hat{f}_N^{AD}(v) = \frac{1}{N_I h}\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{AD}}{h}\right),$$

oraz $r(v) = \kappa^{-1} g(v/\kappa) f_V(v/\kappa)$. Zbieżność w (5.20) zachodzi jeżeli $\hat{r}_N^{AD}(v)$ i $\hat{f}_N^{AD}(v)$ zbiegają według prawdopodobieństwa odpowiednio do $r(v)$ i $\kappa^{-1} f_V(v/\kappa)$ przy $N \to \infty$, w każdym punkcie $v \in \mathbf{R}$, w którym $f_V(v/\kappa) > 0$. Wykorzystując dekompozycję

$$\hat{r}_N^{AD}(v) - r(v) = \left[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right] + \left[\tilde{r}_N(v) - r(v)\right], \tag{5.21}$$

gdzie

$$\tilde{r}_N(v) = \frac{1}{N_I h}\sum_{n\in I} Y_n K\left(\frac{v - \kappa V_n}{h}\right),$$

zbadamy asymptotykę wyrażeń $\left[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right]$ i $\left[\tilde{r}_N(v) - r(v)\right]$. Rozpoczniemy od określenia zachowania asymptotycznego różnicy $\left[\tilde{r}_N(v) - r(v)\right]$. Zauważmy, że

$$E\left\{\tilde{r}_N(v)\right\} = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v - \kappa x}{h}\right) g(x) f_V(x)\,dx = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v - x}{h}\right) r(x)\,dx. \tag{5.22}$$

Ponieważ nieliniowa charakterystyka $g(\cdot)$ jest funkcją ciągłą (założenie 5.2) oraz gęstość sygnału wejściowego ma ograniczony nośnik (założenie 5.1), to całka $\int_{-\infty}^{\infty} |r(v)|\,dv$ jest skończona. Stąd, zgodnie z lematem B.3 z dodatku B.3, na mocy założeń 5.3 i 5.4 ma miejsce następująca zbieżność

$$E\left\{\tilde{r}_N(v)\right\} \to r(v), \text{ gdy } N \to \infty, \tag{5.23}$$

w każdym punkcie $v \in \mathbf{R}$.
Wykorzystując własność 3.2 z rozdziału 3, na mocy lematu A.1 z dodatku A.1 zauważamy następnie, że prawdziwe jest następujące oszacowanie wariancji estymatora $\tilde{r}_N(v)$,

$$Var\left\{\tilde{r}_N(v)\right\} \leq \frac{p+1}{N_I h^2} Var\left\{Y_1 K\left(\frac{v - \kappa V_1}{h}\right)\right\}. \tag{5.24}$$

Ze względu na ograniczoność jądra $K(\cdot)$ widzimy dalej, że element $h^{-1}Var\{Y_1K([v - \kappa V_1]/h)\}$ występujący w (5.24) jest ograniczony przez

$$\frac{2k_0}{h}E\left\{g^2(V_1)K\left(\frac{v-\kappa V_1}{h}\right)\right\}+\frac{2k_0\sigma_Z^2}{h}E\left\{K\left(\frac{v-\kappa V_1}{h}\right)\right\}. \tag{5.25}$$

Ponowne wykorzystanie założeń 5.1 i 5.2 prowadzi do wniosku, że $\int_{-\infty}^{\infty} g^2(v/\kappa) f_V(v/\kappa)\,dv < \infty$. Zatem, na mocy lematu B.3 z dodatku B.3, wyrażenie (5.25) dąży do

$$\frac{2k_0}{\kappa}f_V\left(\frac{v}{\kappa}\right)\left[g^2\left(\frac{v}{\kappa}\right)+\sigma_Z^2\right]<\infty,$$

gdy $N \to \infty$ w każdym punkcie $v \in \mathbf{R}$. Ostatecznie więc nierówność (5.24), dla h spełniającego założenie 5.4, daje

$$Var\{\tilde{r}_N(v)\} = O\left(N^{-4/5}\right), \tag{5.26}$$

w każdym punkcie $v \in \mathbf{R}$.
Zgodnie z rezultatami w (5.23) i (5.26) zachodzi więc zbieżność

$$[\tilde{r}_N(v) - r(v)] \to 0, \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$.
Przejdziemy teraz do określenia asymptotyki wyrażenia $[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)]$. Korzystając z faktu, że jądro $K(\cdot)$ spełnia warunek Lipschitza (zob. zał. 5.3), otrzymujemy

$$\left|\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right| \le \sup_{n\in I}\left|\kappa V_n - \hat{V}_{n,N}^{AD}\right| \frac{L}{N_I h^2}\sum_{n\in I}|Y_n|.$$

Uzyskany w lemacie 5.2 rezultat dotyczący szybkości zbieżności wyrażenia (5.19) wraz ze zbieżnością[2] $(1/N_I)\sum_{n\in I}|Y_n| = E\{|Y_1|\} + O_P\left(N^{-1/2}\right)$ i założeniem 5.4 prowadzi do wniosku, że

$$\left|\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right| \to 0, \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa w każdym punkcie $v \in \mathbf{R}$. Ostatecznie więc

$$\hat{r}_N^{AD}(v) \to r(v), \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa w każdym punkcie $v \in \mathbf{R}$.
Stosując analogiczną argumentację dla estymatora $\hat{f}_N^{AD}(v)$ otrzymujemy zbieżność

$$\hat{f}_N^{AD}(v) \to \kappa^{-1} f_V(v/\kappa), \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, co kończy dowód twierdzenia. ■

Uwaga 5.1 *Należy zwrócić uwagę, że spełnienie wymagania $\kappa \neq 0$ zawartego w twierdzeniu 5.2 zależy zarówno od postaci gęstości wejścia systemu jak i typu charakterystyki nieliniowej $g(\cdot)$. Stąd jego ogólna dyskusja jest zadaniem trudnym. Niemniej jednak warto zauważyć, że niezależnie od typu gęstości $f_X(\cdot)$ spełniającej założenie 5.1, warunek ten zachodzi dla klasy ściśle monotonicznych charakterystyk nieliniowych.*

[2] por. dowód tw. 3.2

Analiza rzędu szybkości zbieżności estymatora charakterystyki nieliniowej

Określimy teraz rząd szybkości zbieżności estymatora $\hat{g}_N^{AD}(\cdot)$. W tym celu ograniczymy klasę rozważanych systemów do systemów z dwukrotnie różniczkowalnymi nieliniowościami (por. założenia u Greblickiego [26]). Założymy również, że $f_X(\cdot)$ jest funkcją dwukrotnie różniczkowalną.

Założenie 5.5 *Drugie pochodne $g''(\cdot)$ i $f_X''(\cdot)$ nieliniowej charakterystyki $g(\cdot)$ i gęstości $f_X(\cdot)$ są funkcjami ciągłymi.*

Założenie 5.6 *Pochodna $g'(\cdot)$ nieliniowości $g(\cdot)$ jest ograniczona przez wielomian, tj. istnieje taki wielomian $m_0(\cdot)$ dowolnego skończonego stopnia, że $|g'(v)| \leq m_0(v)$, $\forall v \in \mathbf{R}$.*

Przyjmiemy również dodatkowe założenia o funkcji jądra $K(\cdot)$.

Założenie 5.7 *Jądro $K(\cdot)$ posiada ciągłą i ograniczoną drugą pochodną $K''(\cdot)$, tj. $|K''(v)| < k_1 < \infty$, $\forall v \in \mathbf{R}$. Ponadto $\int_{-\infty}^{\infty} vK(v)\,dv = 0$ oraz $\int_{-\infty}^{\infty} v^2 K(v)\,dv < \infty$.*

Można teraz sformułować następujące twierdzenie.

Twierdzenie 5.3 *Jeżeli spełnione są założenia twierdzenia 5.2 oraz dodatkowo założenia 5.5–5.7, to dla systemu Wienera (2.10)–(2.11) zachodzi zbieżność*

$$\hat{g}_N^{AD}(v) = g(v/\kappa) + O_P\left(N^{-(2/5)}\right), \tag{5.27}$$

w każdym punkcie $v \in \mathbf{R}$, w którym $f_V(v/\kappa) > 0$.

Dowód twierdzenia zamieszczony jest w dodatku A.3.

5.5 Zastosowanie metody uśrednionej pochodnej w przypadku nieznanej gęstości prawdopodobieństwa sygnału wejściowego

Omówiony powyżej algorytm umożliwia identyfikację systemów Wienera przy założeniu, że gęstość prawdopodobieństwa sygnału wejściowego jest *a priori znana*. Obecnie rozważana modyfikacja metody pozwala z kolei na identyfikację, w przypadku gdy funkcja $f_X(\cdot)$ nie jest dana *a priori*, lecz wiadomo jedynie, że $f_X(\cdot)$ oraz jej pochodna $f'(\cdot)$ są funkcjami jednostajnie ciągłymi (zob. np. Rudin [99]).

Szczególnie wygodnym kryterium przy rozstrzyganiu czy dana funkcja spełnia warunek jednostajnej ciągłości jest twierdzenie, w myśl którego każda funkcja ciągła w zbiorze zwartym jest w tym zbiorze jednostajnie ciągła, Rudin [99, str. 78]. Wynika stąd bezpośrednio, że na mocy założenia 5.1 $f_X(\cdot)$ oraz $f_X'(\cdot)$ są jednostajnie ciągłe.

5.5.1 Konstrukcja estymatora podsystemu dynamicznego

Proponowana wersja algorytmu polega na zastąpieniu funkcji $f_X(\cdot)$ występującej we wzorze definiującym estymatory $\hat{\lambda}_{i,N}^{AD}$, $i = 0, 1, \ldots, p$, (zob. wzory (5.13) i (5.14)) przez estymator $\hat{f}_{X,N}(\cdot)$ gęstości prawdopodobieństwa sygnału $\{X_n\}$ zbudowany w oparciu o dostępne obserwacje wejścia systemu $\{X_n\}_{n=1}^N$. Podobnie, w miejsce pochodnej $f_X'(\cdot)$

wstawiony zostanie estymator $\hat{\phi}_{X,N}(\cdot)$ również utworzony na podstawie obserwacji wejścia systemu. W rezultacie, zmodyfikowany estymator przeskalowanej odpowiedzi impulsowej $\{\kappa\lambda_i\}_{i=0}^{p}$ przyjmie postać

$$\hat{\lambda}_{i,N}^{fAD} = \frac{1}{N_I}\sum_{n\in I} Y_n \hat{d}_{f,i+1,N}(\underline{X}_n), \qquad i = 0,1,\ldots,p, \tag{5.28}$$

gdzie $\hat{d}_{f,i,N}(\cdot)$, $i = 0,1,\ldots,p$, są estymatorami pochodnych cząstkowych $d_{f,i}$, $i = 0,1,\ldots,p$, i wyrażają się wzorem, por. (5.13),

$$\hat{d}_{f,i,N}(\underline{x}) = \hat{\phi}_{X,N}\left(\underline{x}_{(i)}\right)\prod_{j=1,j\neq i}^{p+1}\hat{f}_{X,N}\left(\underline{x}_{(j)}\right), \qquad i = 0,1,\ldots,p.$$

W rozważanym podejściu $\hat{f}_{X,N}(\cdot)$ jest estymatorem jądrowym (por. wzór (2.2) w rozdziale 2) postaci

$$\hat{f}_{X,N}(x) = \frac{1}{h_1 N}\sum_{n=1}^{N} K_1\left(\frac{x - X_n}{h_1}\right), \tag{5.29}$$

gdzie $h_1 = h_1(N)$ i $K_1(\cdot)$ są odpowiednio parametrem wygładzania i funkcją jądra. W celu oszacowania pochodnej $f'_X(\cdot)$ wykorzystamy zaproponowany przez Bhattacharya'ę [4] estymator $\hat{\phi}_{X,N}(\cdot)$ określony następująco

$$\hat{\phi}_{X,N}(x) = \frac{1}{h_2^2 N}\sum_{n=1}^{N} K_2'\left(\frac{x - X_n}{h_2}\right). \tag{5.30}$$

W przypadku gdy $\hat{f}_{X,N}(\cdot)$ i $\hat{\phi}_{X,N}(\cdot)$ posiadają te same jądra oraz ciągi h_1, h_2 są równe, łatwo zauważyć, że zachodzi zależność

$$\hat{\phi}_{X,N}(x) = \frac{d\hat{f}_{X,N}(x)}{dx}.$$

W rezultacie, w celu oszacowania pochodnej gęstości wejścia, wykorzystano estymator będący pochodną klasycznego, jądrowego estymatora gęstości prawdopodobieństwa.

Uwaga 5.2 *Do estymacji funkcji gęstości prawdopodobieństwa $f_X(\cdot)$ oraz jej pochodnej $f'_X(\cdot)$ można zastosować również inne podejścia. Przykładowo, w pracy [36] Greblicki i Pawlak wprowadzają estymatory funkcji gęstości i jej pochodnych konstruowane przy użyciu szeregów Hermite'a.*

5.5.2 Estymacja sygnału interakcyjnego i charakterystyki nieliniowej

Przez analogię do (5.15) wprowadzimy następujący estymator przeskalowanego sygnału interakcyjnego $\{\kappa V_n\}$,

$$\hat{V}_{n,N}^{fAD} = \sum_{i=0}^{p}\hat{\lambda}_{i,N}^{fAD} X_{n-i}, \quad n \in I.$$

Podobnie, postępując tak jak w p. 5.4.3, estymator przeskalowanej charakterystyki nieliniowej $g(\cdot/\kappa)$ przyjmie postać

$$\hat{g}_N^{fAD}(v) = \frac{\sum_{n\in I} Y_n K\left(\frac{v - \hat{V}_{n,N}^{fAD}}{h}\right)}{\sum_{n\in I} K\left(\frac{v - \hat{V}_{n,N}^{fAD}}{h}\right)},$$

gdzie $K(\cdot)$ i $h = h(N)$ są odpowiednio funkcją jądra i parametrem wygładzania.

5.5.3 Własności asymptotyczne algorytmu

Prezentowana obecnie analiza teoretyczna algorytmu dotyczy zbieżności metody. W celu wykazania zgodności estymatora $\hat{g}_N^{fAD}(\cdot)$ konieczna jest jednak znajomość szybkości zbieżności estymatora $\hat{\lambda}_{i,N}^{fAD}$. Ze względu na trudności analityczne związane z określeniem szybkości zbieżności estymatora $\hat{\lambda}_{i,N}^{fAD}$, będziemy dalej zakładać, że szybkość zbieżności $\hat{\lambda}_{i,N}^{fAD}$ jest znana. Analogiczne podejście zastosowano w rozdziale 4.

Analiza zgodności estymatora odpowiedzi impulsowej podsystemu dynamicznego

Określimy warunki, przy których $\hat{\lambda}_{i,N}^{fAD}$, $i = 0, \ldots, p$ są zgodnymi estymatorami elementów przeskalowanej odpowiedzi impulsowej $\kappa\lambda_i, i = 0, \ldots, p$. W tym celu zauważmy najpierw, że wobec zbieżności badanego wcześniej estymatora $\hat{\lambda}_{i,N}^{AD}$ (zob. twierdzenie 5.1), nierówność

$$\left|\hat{\lambda}_{i,N}^{fAD} - \kappa\lambda_i\right| \leq \left|\hat{\lambda}_{i,N}^{fAD} - \hat{\lambda}_{i,N}^{AD}\right| + \left|\hat{\lambda}_{i,N}^{AD} - \kappa\lambda_i\right|,$$

sprowadza rozważany obecnie problem do wykazania, że $\left|\hat{\lambda}_{i,N}^{fAD} - \hat{\lambda}_{i,N}^{AD}\right| \to 0$, gdy $N \to \infty$, według prawdopodobieństwa.
Zauważmy, że

$$\left|\hat{\lambda}_{i,N}^{fAD} - \hat{\lambda}_{i,N}^{AD}\right| \leq \frac{1}{N_I}\sum_{n\in I}|Y_n|\left|\hat{d}_{f,i+1,N}(\underline{X}_n) - d_{f,i+1}(\underline{X}_n)\right|,$$

i dlatego

$$\left|\hat{\lambda}_{i,N}^{fAD} - \hat{\lambda}_{i,N}^{AD}\right| \leq \sup_{\underline{x}\in\Omega}\left|\hat{d}_{f,i+1,N}(\underline{x}) - d_{f,i+1}(\underline{x})\right| \frac{1}{N_I}\sum_{n\in I}|Y_n|$$

W celu wykazania zgodności estymatora wystarczy zatem dowieść, że

$$\sup_{\underline{x}\in\Omega}\left|\hat{d}_{f,i+1,N}(\underline{x}) - d_{f,i+1}(\underline{x})\right| \to 0, \quad \text{gdy } N \to \infty, \tag{5.31}$$

według prawdopodobieństwa. Na mocy lematu A.2 w dodatku A.3 zbieżność w (5.31) zachodzi, jeżeli estymatory $\hat{f}_{X,N}(x)$ i $\hat{\phi}_{X,N}(x)$ zbiegają jednostajnie według prawdopodobieństwa odpowiednio do $f_X(x)$ i $f'_X(x)$, gdy $N \to \infty$, tzn.

$$\sup_{x\in\mathbf{R}}\left|\hat{f}_{X,N}(x) - f_X(x)\right| \to 0, \quad \text{gdy } N \to \infty, \tag{5.32}$$

oraz

$$\sup_{x\in\mathbf{R}}\left|\hat{\phi}_{X,N}(x) - f'_X(x)\right| \to 0, \quad \text{gdy } N \to \infty, \tag{5.33}$$

według prawdopodobieństwa.

Zagadnienie zbieżności jednostajnej jądrowych estymatorów gęstości prawdopodobieństwa dyskutowane jest między innymi przez Nadaraya'ę [81], Schustera [112], Van Ryzina [106] i Silvermana [116]. Zgodnie z wynikami otrzymanymi np. przez Silvermana, zbieżność w (5.32) i (5.33) zachodzi jeżeli funkcje $f_X(\cdot)$ i $f'_X(\cdot)$ są jednostajnie ciągłe oraz jądra $K_1(\cdot)$, $K_2(\cdot)$ i parametry wygładzania h_1, h_2 spełniają odpowiednio założenia twierdzeń A i C w [116].

Uzyskane powyżej rezultaty dowodzą prawdziwości następującego twierdzenia:

Twierdzenie 5.4 *Jeżeli dla systemu Wienera (2.10)-(2.11) spełnione są założenia twierdzenia 5.1 oraz zachodzą zbieżności (5.32) i (5.33), to dla każdego $i \in \{0, \ldots, p\}$ estymator $\hat{\lambda}_{i,N}^{fAD}$ jest zgodnym estymatorem elementu $\kappa\lambda_i$, tzn.*

$$\hat{\lambda}_{i,N}^{fAD} \to \kappa\lambda_i, \quad gdy\ N \to \infty,$$

według prawdopodobieństwa.

Analiza zgodności estymatora odpowiedzi impulsowej podsystemu dynamicznego

Postępując tak jak w przypadku estymatora $\hat{g}_N^{LS}(\cdot)$ (zob. rozdział 4 i wzór (4.11)), założymy że zachodzi

$$\left\| \hat{\underline{\lambda}}_N^{fAD} - \kappa\underline{\lambda} \right\| = O_P(a_N),$$

gdzie $\hat{\underline{\lambda}}_N^{fAD} = \left(\hat{\lambda}_{0,N}^{fAD}, \hat{\lambda}_{1,N}^{fAD} \ldots, \hat{\lambda}_{p,N}^{fAD}\right)^T$ oraz a_N jest pewnym ciągiem zbieżnym do zera.

Odnośnie konstrukcji estymatora $\hat{g}_N^{fAD}(\cdot)$ założymy, że:

Założenie 5.8 *Jądro $K(\cdot)$ jest nieujemną funkcją lipschitzowską (tj. $\exists L < \infty$, $\forall v_1, v_2 \in \mathbf{R}$, $|K(v_1) - K(v_2)| \leq L|v_1 - v_2|$), dla której spełnione są następujące warunki:* $\sup_{v\in\mathbf{R}} K(v) = k_0 < \infty$, $\int_{-\infty}^{\infty} K(v)\,dv = 1$ *oraz* $\lim_{|v|\to\infty} vK(v) = 0$.

Założenie 5.9 *Parametr wygładzania $h = h(N)$ spełnia warunki*

$$h \to 0, Nh \to \infty, \qquad gdy\ N \to \infty,$$

oraz

$$a_N/h^2 \to 0, \qquad gdy\ N \to \infty.$$

Zgodność estymatora $\hat{g}_N^{fAD}(\cdot)$ przy powyższych założeniach oraz założeniach twierdzenia 5.4, wykazana jest w następującym twierdzeniu:

Twierdzenie 5.5 *Jeżeli spełnione są założenia twierdzenia 5.4, założenia 5.8, 5.9, oraz $\kappa \neq 0$, to*

$$\hat{g}_N^{fAD}(v) \to g(v/\kappa), \qquad gdy\ N \to \infty,$$

według prawdopodobieństwa w każdym punkcie $v \in \mathbf{R}$, w którym $f_V(v/\kappa) > 0$.

Dowód twierdzenia przebiega analogicznie do dowodu twierdzenia 4.2 w rozdziale 4.

5.6 Podsumowanie wyników teoretycznych

W rozdziale omówiono algorytm identyfikacji systemów Wienera zbudowany w oparciu o koncepcję estymacji uśrednionej pochodnej. Dla skonstruowanej w dwóch wersjach metody (tj. dla przypadku ze znaną oraz z nieznaną gęstością prawdopodobieństwa sygnału wejściowego $\{X_n\}$) przeprowadzono analizę teoretyczną odnoszącą się do asymptotycznego zachowania się estymatorów. W szczególności wykazano zgodność estymatorów $\{\hat{\lambda}_{i,N}^{AD}\}_{i=0}^p$, $\hat{g}_N^{AD}(\cdot)$ oraz określono ich rzędy szybkości zbieżności. W przypadku estymatorów $\{\hat{\lambda}_{i,N}^{fAD}\}_{i=0}^p$, $\hat{g}_N^{fAD}(\cdot)$, tj. w sytuacji gdy gęstość prawdopodobieństwa $f_X(\cdot)$ sygnału wejściowego $\{X_n\}$ jest nieznana, wykazano zgodność estymatorów odpowiedzi impulsowej części liniowej systemu oraz (przy założeniu, że znany jest rząd szybkości zbieżności oszacowań $\{\hat{\lambda}_{i,N}^{fAD}\}_{i=0}^p$) zgodność estymatora nieliniowej charakterystyki $g(\cdot)$.

5.7 Wyniki badań eksperymentalnych

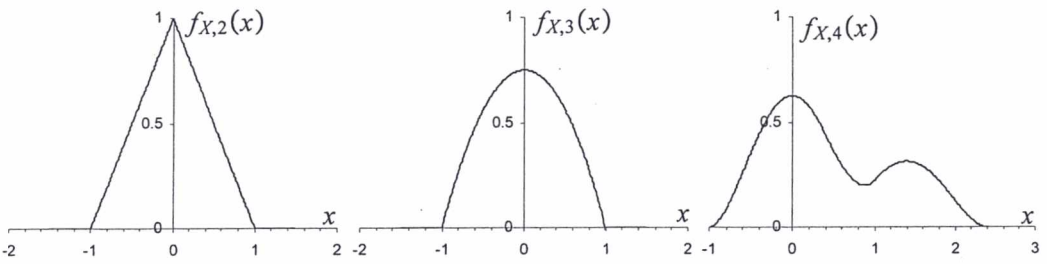
Dla opracowanego algorytmu identyfikacji wykonano badania symulacyjne ilustrujące zachowanie estymatorów $\hat{\lambda}_N^{AD}$, $\hat{g}_N^{AD}(\cdot)$ oraz $\hat{\lambda}_N^{fAD}$, $\hat{g}_N^{fAD}(\cdot)$ dla małej i umiarkowanej liczby obserwacji wejścia i wyjścia systemu.

Badania przeprowadzono dla systemów pobudzanych sygnałami wejściowymi o gęstościach $f_{X,2}(\cdot)$, $f_{X,3}(\cdot)$ (zob. wzory (4.19)–(4.20) na str. 45) oraz $f_{X,4}(\cdot)$, zob. rys. 5.2. Funkcja $f_{X,4}(\cdot)$ określona jest wzorem

$$f_{X,4}(x) = \frac{2}{3}\varphi(x) + \frac{1}{3}\varphi(x - 1.4),$$

gdzie $\varphi(x)$ jest gęstością prawdopodobieństwa $\varphi(x) = (15/16)\left(x^4 - 2x^2 + 1\right)\mathbf{1}_{[-1,1]}(x)$. Podobnie jak w rozdziałach 3. i 4., podsystem dynamiczny modelowany był za pomocą odpowiedzi impulsowej (3.35), str. 32. Ze względu na zerowanie się stałej κ w przypadku nieliniowości $g_2(\cdot)$ i $g_3(\cdot)$ (por. uwaga 5.1 na str. 59), w badaniach wykorzystano następujące charakterystyki nieliniowe (zob. rys. 5.3)

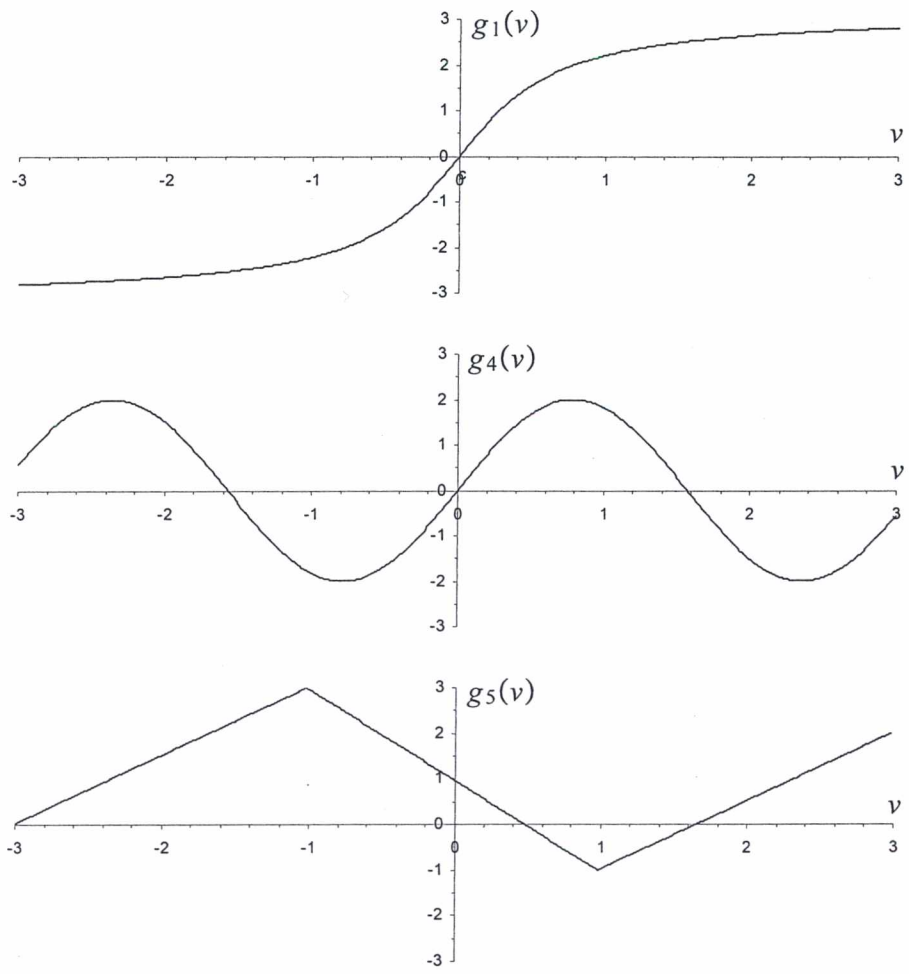
$$\begin{aligned} g_1(v) &= 2\arctan(2v), \\ g_4(v) &= 2\sin(2v), \\ g_5(v) &= \begin{cases} 1.5v + 4.5 & \text{dla} \quad v < -1 \\ -2v + 1 & \text{dla} \quad -1 \le v \le 1 \\ 1.5v - 2.5 & \text{dla} \quad v > 1 \end{cases}. \end{aligned}$$



Rysunek 5.2: Wykorzystywane w eksperymentach numerycznych gęstości prawdopodobieństwa $f_{X,2}$, $f_{X,3}$ i $f_{X,4}$

W symulacjach badano zależność błędów empirycznych $Err(\cdot;\kappa)$ i $MISE(\cdot;\kappa)$, określonych wzorami (3.39) i (3.40) na str. 33, od liczby obserwacji N wejścia i wyjścia systemu (zmienianej od 25 do 500). W definicji błędów $Err(\cdot;\kappa)$ i $MISE(\cdot;\kappa)$ przyjęto następujące wartości stałych R i M: $R = 30$, $M = 1000$. Wartości stałej κ wyznaczono numerycznie dla każdej z zastosowanych nieliniowości (oraz dla gęstości $f_{X,2}(\cdot)$, $f_{X,3}(\cdot)$ i $f_{X,4}(\cdot)$). W przypadku gęstości wejścia $f_{X,2}(\cdot)$ i $f_{X,3}(\cdot)$ (o nośniku $[-1,1]$), sygnał wewnętrzny $\{V_n\}$ posiadał gęstość prawdopodobieństwa o nośniku $[\vartheta_1, \vartheta_2]$, gdzie $\vartheta_1 \approx -2.28$ i $\vartheta_1 \approx 2.28$. Dla gęstości $f_{X,4}(\cdot)$ (z nośnikiem $[-1, 2.4]$) stałe ϑ_1 i ϑ_2 przyjmowały odpowiednio wartości -2.28 i 5.48. Sygnał zakłócający $\{Z_n\}$ wygenerowano z rozkładu jednostajnego o zerowej średniej i wariancji dobranej tak, aby NSR= 10% (por. rozdział 3). W estymatorach $\hat{g}_N^{AD}(\cdot)$ i $\hat{g}_N^{fAD}(\cdot)$ zastosowano gaussowską funkcję jądra

$$K(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right).$$



Rysunek 5.3: Wykorzystywane w symulacjach nieliniowości g_1, g_4 i g_5

5.7.1 Rezultaty eksperymentów w przypadku znanej gęstości prawdopodobieństwa sygnału wejściowego

Rysunki 5.4–5.9 (zob. str. 67) przedstawiają wyniki eksperymentów numerycznych, tj. zależności błędów empirycznych $Err(\hat{\lambda}_N^{AD};\kappa)$ i $MISE\left(\hat{g}_N^{AD};\kappa\right)$ od liczby obserwacji wejścia i wyjścia systemu. Uzyskane rezultaty prowadzą do wniosku, że wraz ze wzrostem liczby obserwacji N wartości badanych błędów maleją. Gdy sygnał wejściowy posiada gęstość $f_{X,2}(\cdot)$ lub $f_{X,3}(\cdot)$ błędy empiryczne przyjmują zbliżone wartości, niezależnie od nieliniowości występującej w systemie (zob. rys. 5.4–5.7). Należy jednak podkreślić, że dla charakterystyki $g_5(\cdot)$ (nie spełniającej założenia 5.1) uzyskano nieznacznie gorsze rezultaty eksperymentu (zob. rys. 5.4 i 5.6). Przykładowo, błąd $Err(\hat{\lambda}_N^{AD};\kappa)$ o wartości 5% dla nieliniowości $g_1(\cdot)$ i $g_4(\cdot)$ otrzymano przy $N = 150$. Dla nieliniowości $g_5(\cdot)$, natomiast, tę samą wartość błędu uzyskano przy $N = 300$ (dla $N = 150$ błąd przyjmował wartość 12%, por. rys. 5.4). W przypadku gęstości $f_{X,4}(\cdot)$ uzyskane błędy są istotnie większe (zob. rys. 5.8 i 5.9) niż w przypadku gęstości $f_{X,2}(\cdot)$ i $f_{X,3}(\cdot)$, co prowadzi do wniosku, że badany algorytm jest wrażliwy na charakter gęstości sygnału wejściowego. Rezultaty otrzymane dla gęstości $f_{X,4}(\cdot)$ wskazują ponadto, że typ nieliniowości występującej w systemie może również istotnie wpływać na efektywność algorytmu. Dla ciągłej i monotonicznej nielin-

iowości $g_1(\cdot)$, prezentowane na wykresach 5.8 i 5.9 błędy $Err(\hat{\lambda}_N^{AD};\kappa)$ i $MISE\left(\hat{g}_N^{AD};\kappa\right)$ przyjmowały wartości najmniejsze. Dla nieliniowości $g_5(\cdot)$ z kolei (tj. nieliniowości nie spełniającej założenia 5.1), otrzymano większe wartości błędów.

5.7.2 Rezultaty eksperymentów w przypadku nieznanej gęstości prawdopodobieństwa sygnału wejściowego

Badany algorytm identyfikacji w przypadku nieznanej gęstości wejścia $f_X(\cdot)$ wymaga zastosowania estymatorów $\hat{\phi}_{X,N}(\cdot)$ i $\hat{f}_{X,N}(\cdot)$ (zob. wzory (5.29) i (5.30)). W eksperymentach numerycznych przyjęto, że funkcje $K_1(\cdot)$ i $K_2(\cdot)$ występujące w oszacowaniach $\hat{\phi}_{X,N}(\cdot)$ i $\hat{f}_{X,N}(\cdot)$ są jądrami gaussowskimi, tj.

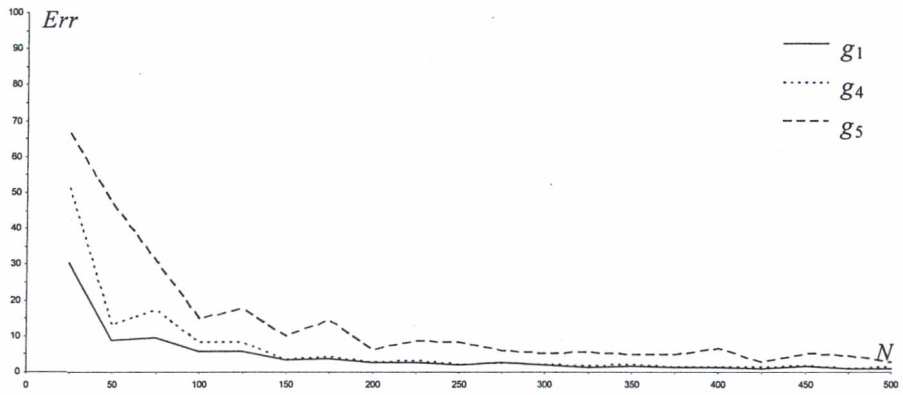
$$K_1(v) = K_2(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right).$$

Wykresy 5.10–5.15 (zob. str. 69) obrazują zmiany wielkości błędów empirycznych $Err(\hat{\lambda}_N^{fAD};\kappa)$ i $MISE\left(\hat{g}_N^{fAD};\kappa\right)$ otrzymanych dla systemów z nieliniowościami $g_1(\cdot)$, $g_4(\cdot)$, $g_5(\cdot)$ i pobudzanych sygnałami o gęstościach prawdopodobieństwa $f_{X,2}(\cdot)$, $f_{X,3}(\cdot)$ i $f_{X,4}(\cdot)$. Na podstawie otrzymanych rezultatów można stwierdzić, że badane błędy maleją wraz ze wzrostem liczby obserwacji N zarówno dla charakterystyk $g_1(\cdot)$ i $g_4(\cdot)$ spełniających przyjęte w rozdziale założenia, jak i dla charakterystyki $g_5(\cdot)$, która nie jest różniczkowalna w punktach -1 i 1.

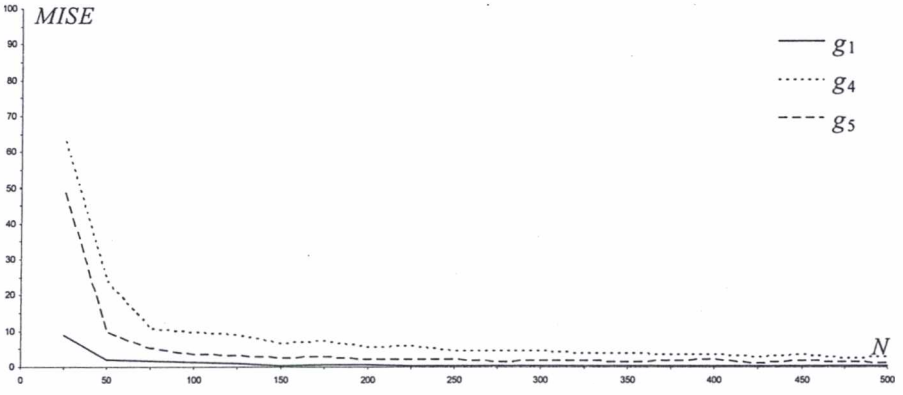
Wartości uzyskanych błędów są na ogół większe od odpowiadających im błędów otrzymanych w przypadku algorytmu wykorzystującego aprioryczną znajomość gęstości prawdopodobieństwa sygnału wejściowego. Przykładowo, dla gęstości $f_{X,2}(\cdot)$ i nieliniowości $g_1(\cdot)$ błąd estymatora $\hat{\lambda}_N^{fAD}$ nie przekraczał 30%, gdy liczba obserwacji N była większa od 220 (zob. rys. 5.10). Analogiczny eksperyment wykonany dla estymatora $\hat{\lambda}_N^{AD}$ (tj. przy założeniu znajomości gęstości wejścia) pozwolił uzyskać 30% błąd, już przy $N = 30$. Dla $N = 220$ błąd ten wynosił natomiast tylko około 3%.

5.8 Podsumowanie i wnioski

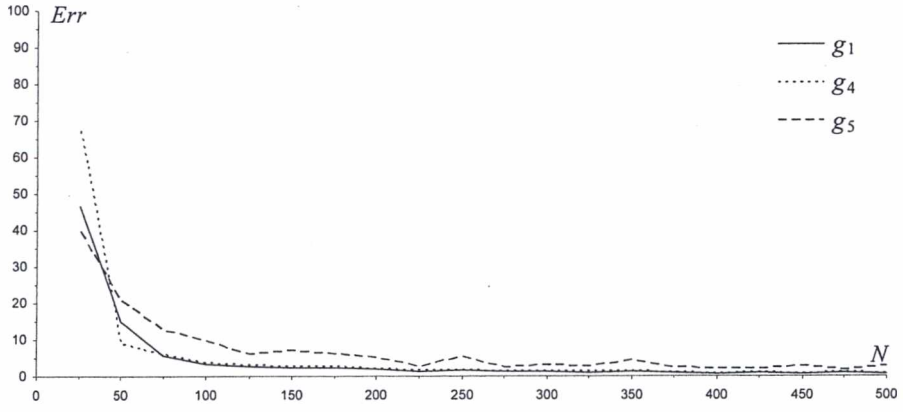
Przedstawiony w rozdziale algorytm identyfikacji skonstruowany został w oparciu o koncepcję estymacji uśrednionej pochodnej, zob. Powell *et al.* [95]. W wersji podstawowej metoda umożliwia szacowanie obu podsystemów przy założeniu, że gęstość prawdopodobieństwa sygnału wejściowego $\{X_n\}$ jest znana. W rozdziale przedstawiono również modyfikację algorytmu pozwalającą na identyfikację w przypadku, gdy gęstość wejścia nie jest *a priori* znana. Przeprowadzono analizę teoretyczną własności asymptotycznych obu algorytmów. W szczególności wykazano zgodność oraz określono rzędy szybkości zbieżności estymatorów $\{\hat{\lambda}_{i,N}^{AD}\}_{i=0}^{p}$ i $\hat{g}_N^{AD}(\cdot)$ (tj. estymatorów umożliwiających identyfikację, gdy gęstość wejścia jest znana) oraz wykazano zgodność estymatorów $\{\hat{\lambda}_{i,N}^{fAD}\}_{i=0}^{p}$ i $\hat{g}_N^{fAD}(\cdot)$ (tj. estymatorów umożliwiających identyfikację w przypadku nieznanej gęstości wejścia). Wykonane badania eksperymentalne proponowanych podejść ilustrują zachowanie się obu metod w przypadku małej i umiarkowanej liczby obserwacji wejścia i wyjścia systemu.



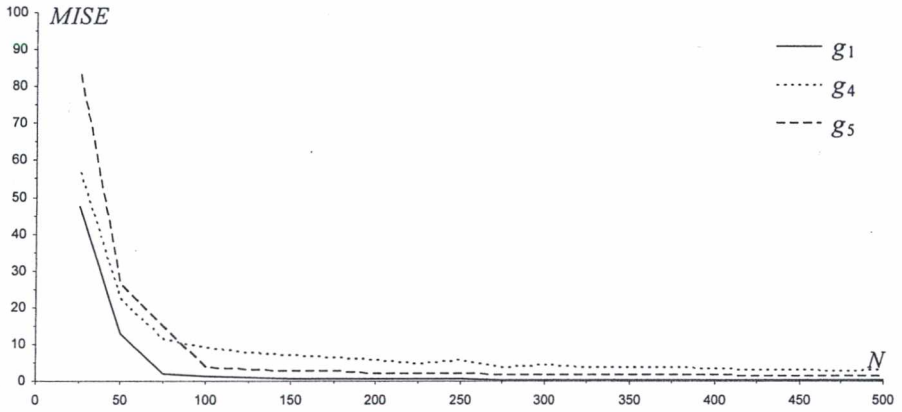
Rysunek 5.4: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{AD}; \kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,2}$



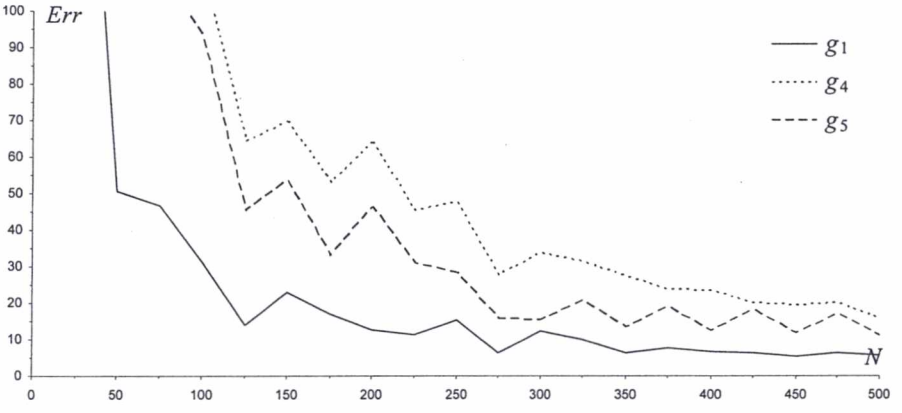
Rysunek 5.5: Błędy estymacji $MISE\left(\hat{g}_N^{AD}; \kappa\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,2}$



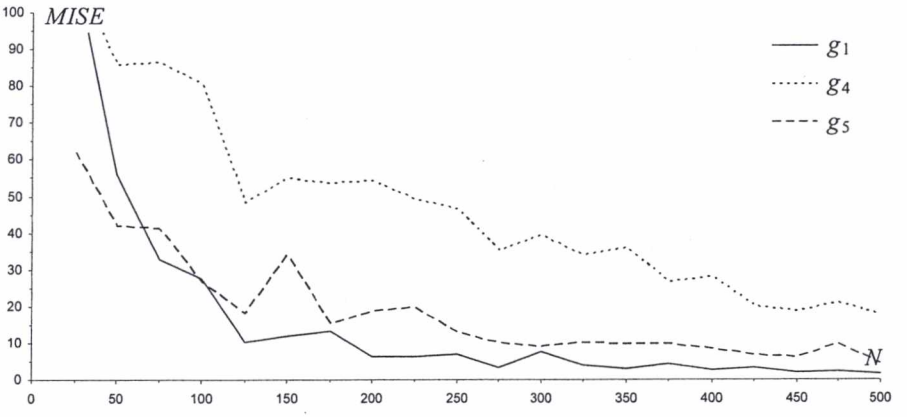
Rysunek 5.6: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{AD}; \kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,3}$



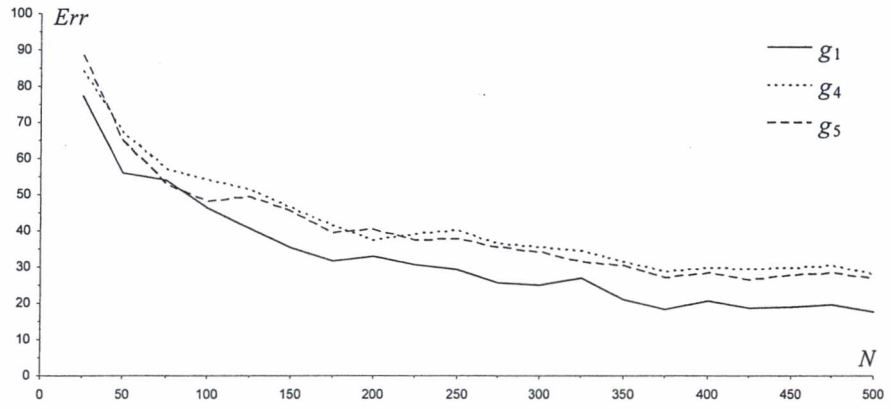
Rysunek 5.7: Błędy estymacji $MISE\left(\hat{g}_N^{AD};\kappa\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,3}$



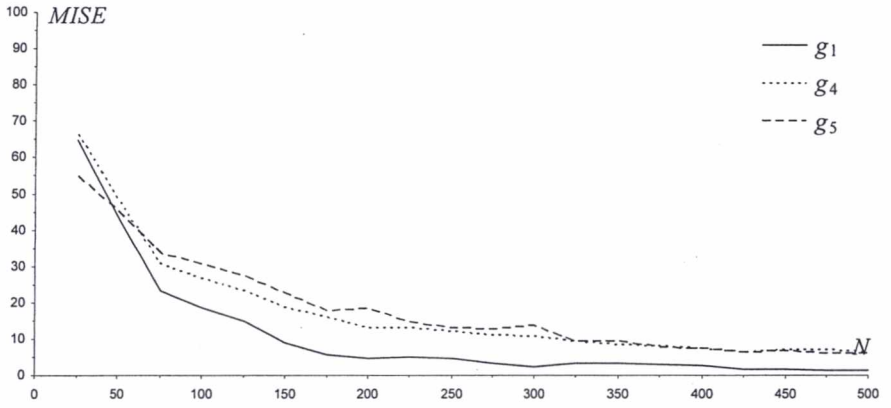
Rysunek 5.8: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{AD};\kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,4}$



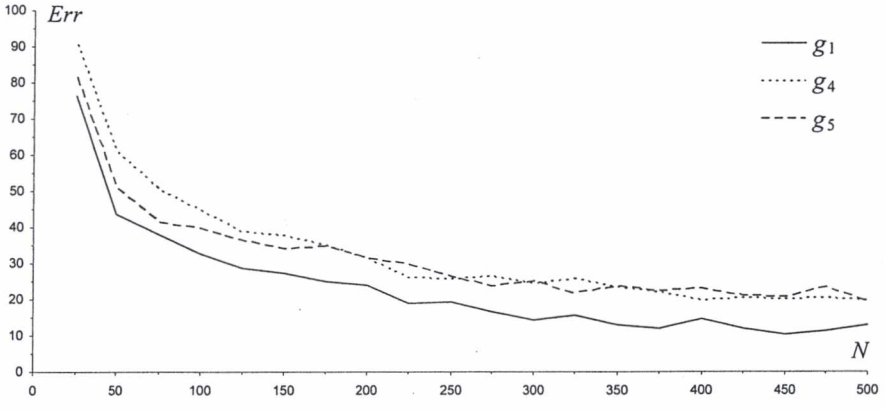
Rysunek 5.9: Błędy estymacji $MISE\left(\hat{g}_N^{AD};\kappa\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,4}$



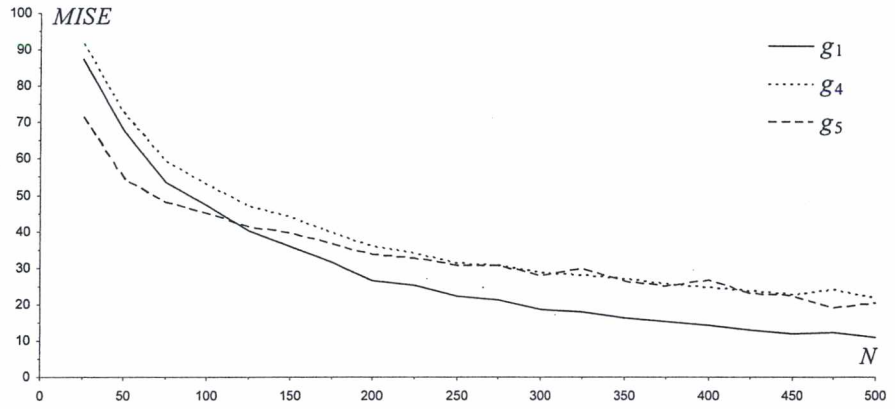
Rysunek 5.10: Błędy estymacji $Err(\hat{\lambda}_N^{fAD}; \kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,2}$



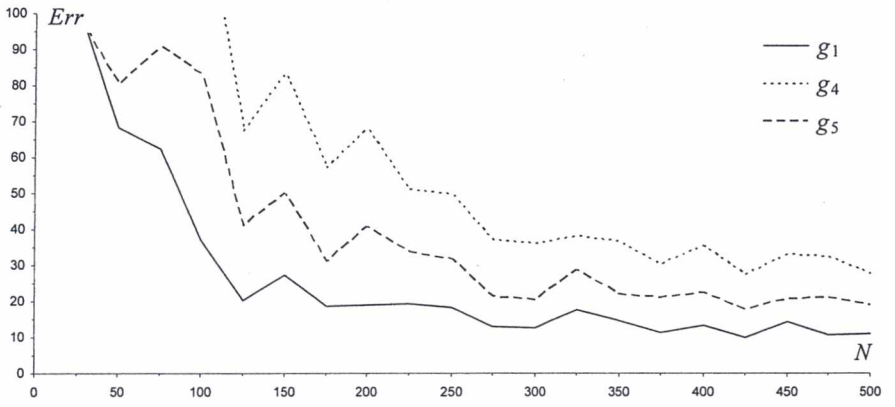
Rysunek 5.11: Błędy estymacji $MISE(\hat{g}_N^{fAD}; \kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,2}$



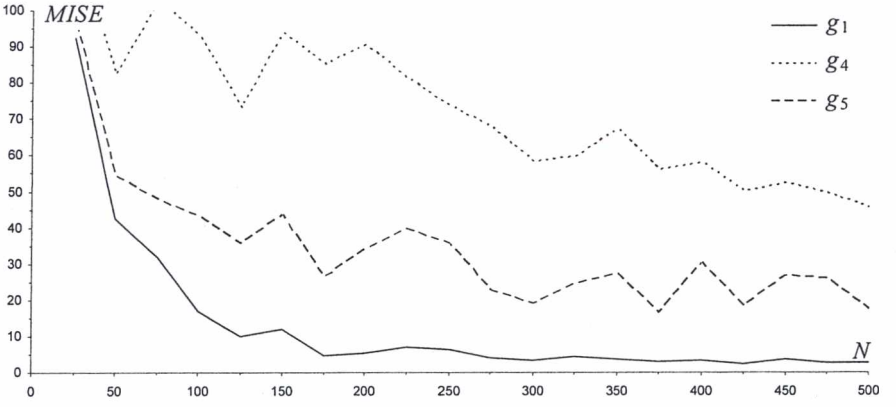
Rysunek 5.12: Błędy estymacji $Err(\hat{\lambda}_N^{fAD}; \kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,3}$



Rysunek 5.13: Błędy estymacji $MISE\left(\hat{g}_N^{fAD};\kappa\right)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,3}$



Rysunek 5.14: Błędy estymacji $Err(\hat{\underline{\lambda}}_N^{fAD};\kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,4}$



Rysunek 5.15: Błędy estymacji $MISE(\hat{g}_N^{fAD};\kappa)$ w funkcji liczby obserwacji wejścia i wyjścia systemu dla nieliniowości g_1, g_4, g_5 i gęstości sygnału wejściowego $f_{X,4}$

Rozdział 6

Porównanie opracowanych metod identyfikacji systemów Wienera

W tym rozdziale przedstawiona jest zbiorcza analiza porównawcza uzyskanych wyników oraz krótka charakterystyka cech wspólnych proponowanych metod identyfikacji, na tle innych podejść omawianych w literaturze. Dyskutowane w pracy algorytmy porównywane są również ze sobą pod kątem zakresu stosowalności, własności asymptotycznych oraz własności dla małej i umiarkowanej liczby obserwacji wejścia i wyjścia systemu.

6.1 Opracowane metody identyfikacji, a metody parametryczne i nieparametryczne

Skonstruowane w pracy algorytmy i ich analiza rozszerzają dotychczasową teorię oraz dają możliwość identyfikacji systemów Wienera w przypadku, gdy bezpośrednie zastosowanie opracowanych wcześniej metod parametrycznych i nieparametrycznych jest niemożliwe. W odróżnieniu od większości wcześniejszych podejść, omawiane w pracy algorytmy (z pominięciem algorytmu korelacyjnego) pozwalają bowiem na identyfikację systemów Wienera, gdy 1) wejście systemu nie jest sygnałem gaussowskim, 2) charakterystyka nieliniowa systemu nie jest odwracalna. W przypadku gaussowskiego wejścia $\{X_n\}$, zaproponowana w rozdziale 3. metoda korelacyjna funkcjonuje dla szerokiej klasy charakterystyk nieliniowych w systemie. Nie zakłada się mianowicie, w odróżnieniu od większości metod parametrycznych, że nieliniowość jest typu wielomianowego (por. 2.2.1, str. 16), a jedynie że jest funkcją ograniczoną przez wielomian dowolnego (skończonego) stopnia.

Ze względu na jednolity szkielet konstrukcyjny (zob. p. 2.3.1, str. 20), proponowane podejścia posiadają następujące cechy wspólne:

- Zakładana jest parametryczna informacja wstępna o podsystemie dynamicznym oraz nieparametryczna o podsystemie statycznym (w odróżnieniu od metod parametrycznych i nieparametrycznych, w których zakłada się czysto parametryczną i czysto nieparametryczną informację wstępną o systemie).
- Algorytmy umożliwiają nieparametryczną identyfikację charakterystyki nieliniowej „wprost", tj. nie poprzez funkcję odwrotną (w odróżnieniu od dotychczasowych podejść nieparametrycznych, w których szacowana jest funkcja odwrotna do nieliniowości).

- Identyfikacja nieliniowości przeprowadzana jest przy użyciu estymatora odpowiedzi impulsowej podsystemu dynamicznego (w odróżnieniu od algorytmów nieparametrycznych, w których oba podsystemy identyfikowane są niezależnie).
- Dopuszcza się występowanie addytywnego i niedostępnego dla pomiarów szumu na wyjściu całego systemu (por. dyskusja w p.2.2).

6.2 Zakres stosowalności

Proponowane w pracy podejścia różnią się między sobą pod względem zastosowanych do ich konstrukcji metod matematycznych. Algorytm omawiany w rozdziale 3. zbudowany jest w oparciu o metodę korelacyjną. Z kolei algorytmy przedstawione w rozdziale 4. bazują na metodzie najmniejszych kwadratów oraz idei estymacji korelacji rangowej, a metoda prezentowana w rozdziale 5. wykorzystuje koncepcję estymacji uśrednionej pochodnej. Zróżnicowana baza konstrukcyjna estymatorów powoduje, że dyskutowane algorytmy mogą być stosowane przy różnych założeniach odnośnie nieliniowości występującej w systemie oraz dla różnych sygnałów wejściowych. Ponadto algorytmy cechują się różnymi własnościami dla małej i umiarkowanej liczby obserwacji wejścia i wyjścia systemu. Określając zakres stosowalności, otrzymane w pracy algorytmy można więc charakteryzować według następujących kryteriów:

Wymagania dotyczące sygnału wejściowego. Szczególna konstrukcja systemów Wienera (zob. uwaga 2.1 i 2.2 na str. 20) powoduje, że charakter sygnału wejściowego $\{X_n\}$ ma kluczowe znaczenie podczas konstruowania algorytmów identyfikacji. Z tego względu parametryczne i nieparametryczne metody omawiane w literaturze z reguły wymagają, aby wejście systemu było sygnałem gaussowskim. Proponowane w pracy algorytmy pozwalają natomiast na identyfikację zarówno w przypadku gaussowskich (rozdział 3) jak i niegaussowskich (rozdziały 4 i 5) sygnałów wejściowych. Wspólnym wymaganiem dla proponowanych metod jest założenie, że sygnał wejściowy jest ciągiem typu *i.i.d.* Porównanie opracowanych algorytmów pod kątem założeń dotyczących sygnału wejściowego przedstawia tabela 6.1.

Algorytm identyfikacji	**Wymagania dotyczące sygnału wejściowego**
Algorytm korelacyjny	Gaussowski sygnał wejściowy o dowolnej wartości średniej i wariancji
Algorytm wykorzystujący metodę najmniejszych kwadratów	Ciągła funkcja gęstości o ograniczonym (chociaż dowolnie dużym) nośniku
Algorytm wykorzystujący korelację rangową	Nie przeprowadzono analizy teoretycznej algorytmu
Algorytm uśrednionej pochodnej ze znaną gęstością wejścia	Różniczkowalna gęstość o zwartym (chociaż dowolnie dużym) nośniku
Algorytm uśrednionej pochodnej z nieznaną gęstością wejścia	Różniczkowalna gęstość o zwartym (chociaż dowolnie dużym) nośniku

Tabela 6.1: Porównanie proponowanych algorytmów identyfikacji pod względem głównych wymagań dotyczących sygnału wejściowego

Wymagania dotyczące sygnału zakłócającego. Proponowane w pracy metody umożliwiają identyfikację systemów Wienera w przypadku, gdy sygnał wyjściowy (tzn. sygnał będący wyjściem nieliniowego podsystemu statycznego) zakłócany jest przez addytywny i niedostępny dla pomiarów szum $\{Z_n\}$ typu *i.i.d.* o zerowej średniej i skończonej wariancji. Wymaganie to jest wspólne dla wszystkich omawianych w pracy algorytmów.

Wymagania dotyczące podsystemu liniowego. Charakter proponowanych metod wymaga posiadania parametrycznej wiedzy o liniowym podsystemie dynamicznym (tj. znajomości długości pamięci, czyli stałej $p < \infty$). Podobnie jak w przypadku założeń odnośnie szumu wyjściowego, wymaganie to jest wspólne dla wszystkich prezentowanych w pracy algorytmów.

Wymagania dotyczące podsystemu nieliniowego. W odróżnieniu od założeń dotyczących podsystemu liniowego, dyskutowane metody różnią się między sobą pod względem warunków jakie musi spełniać nieliniowa charakterystyka występująca w systemie. Tabela 6.2 zawiera wykaz podstawowych wymagań odnośnie nieliniowości, gwarantujących zbieżność poszczególnych algorytmów.

Algorytm identyfikacji	Wymagania dotyczące nieliniowości
Algorytm korelacyjny	Nieliniowa charakterystyka ograniczona jest przez wielomian dowolnego skończonego stopnia
Algorytm wykorzystujący metodę najmniejszych kwadratów	Nieliniowa charakterystyka jest ciągła i różna od funkcji stałej
Algorytm wykorzystujący korelację rangową	Nie przeprowadzono analizy teoretycznej algorytmu
Algorytm uśrednionej pochodnej ze znaną gęstością wejścia	Nieliniowa charakterystyka posiada ciągłą pochodną
Algorytm uśrednionej pochodnej z nieznaną gęstością wejścia	Nieliniowa charakterystyka posiada ciągłą pochodną

Tabela 6.2: Porównanie proponowanych algorytmów identyfikacji pod względem głównych wymagań dotyczących nieliniowej charakterystyki w systemie

6.3 Własności asymptotyczne

Dla proponowanych w pracy algorytmów identyfikacji systemów Wienera przeprowadzona została analiza teoretyczna[1] określająca asymptotyczne własności estymatorów liniowego podsystemu dynamicznego oraz nieliniowego podsystemu statycznego. W przypadku estymatorów odpowiedzi impulsowej części liniowej systemu badano zbieżność oraz szybkość zbieżności oszacowań według prawdopodobieństwa, a w przypadku estymatorów nieliniowej charakterystyki $g(\cdot)$, badano punktową zbieżność i szybkość zbieżności (również

[1] W pracy nie przeprowadzono analizy teoretycznej własności asymptotycznych algorytmu zbudowanego w oparciu o koncepcję korelacji rangowej.

według prawdopodobieństwa). Uzyskane wyniki prowadzą do wniosku, że proponowane metody identyfikacji są zbieżne.

6.4 Numeryczna realizacja algorytmów

Ważnym, z punktu widzenia zastosowań, kryterium porównawczym omawianych metod, jest łatwość ich implementacji. Zarówno algorytm korelacyjny, jak i metoda wykorzystująca koncepcję estymacji uśrednionej pochodnej (ze znaną oraz nieznaną gęstością wejścia), nie wymagają stosowania skomplikowanych procedur numerycznych (jak np. odwracanie macierzy czy wyznaczanie minimum wielowymiarowej funkcji celu). Z tego względu wymienione wyżej algorytmy mogą być łatwo zaimplementowane w wielu środowiskach programistycznych. Z kolei algorytmy zbudowane w oparciu o kryterium średniokwadratowe i koncepcję korelacji rangowej są algorytmami, w których konieczne jest wyznaczenie ekstremów wielowymiarowych funkcji celu co powoduje, że do ich efektywnej implementacji konieczne jest zastosowanie specjalizowanych procedur optymalizacyjnych.

Przedstawiona tu analiza pokazuje, że cele i zadania postawione w punktach 2.5 i 2.6.2 zostały w pracy zrealizowane.

Rozdział 7

Uwagi końcowe

W rozdziale podsumowano oryginalne wyniki naukowe uzyskane w pracy, przedstawiono otwarte problemy badawcze oraz zaproponowano kierunki dalszych badań.

7.1 Oryginalne wyniki naukowe przedstawione w pracy

W pracy otrzymano następujące oryginalne wyniki naukowe:

- Zaproponowano ogólny schemat dwuetapowej, parametryczno–nieparametrycznej identyfikacji systemów Wienera wykorzystujący koncepcję estymacji nieznanego sygnału interakcyjnego $\{V_n\}$ występującego w systemie.
- Zaproponowano klasę parametryczno–nieparametrycznych algorytmów umożliwiających identyfikację systemów Wienera przy różnorodnych założeniach odnośnie sygnału wejściowego oraz dla szerokiej klasy charakterystyk nieliniowych.
  W szczególności opracowano:
  - algorytm korelacyjny (rozdział 3),
  - algorytm wykorzystujący metodę najmniejszych kwadratów (rozdział 4),
  - algorytm zbudowany w oparciu o koncepcję korelacji rangowej (rozdział 4),
  - algorytm wykorzystujący ideę estymacji uśrednionej pochodnej (rozdział 5).
- Przeprowadzono analizę teoretyczną własności asymptotycznych algorytmu korelacyjnego, algorytmu zbudowanego w oparciu o metodę najmniejszych kwadratów oraz algorytmu wykorzystującego ideę estymacji uśrednionej pochodnej.
  W szczególności:
  - dla metody korelacyjnej określono warunki, przy których estymatory podsystemów liniowego i nieliniowego są zgodne,
  - określono rzędy szybkości zbieżności estymatorów części liniowej i nieliniowej w metodzie korelacyjnej,
  - wykazano zgodność estymatorów podsystemów liniowego i nieliniowego opracowanych w oparciu o metodę najmniejszych kwadratów,
  - przy założeniu znajomości gęstości sygnału wejściowego dla metody wykorzystującej estymację uśrednionej pochodnej wykazano zgodność oraz określono rzędy szybkości zbieżności estymatorów podsystemów liniowego i nieliniowego.

- w przypadku nieznanej gęstości wejścia dla metody wykorzystującej estymację uśrednionej pochodnej wykazano zgodność estymatorów części liniowej i nieliniowej systemu.

Ponadto:

- Opracowano implementację proponowanych algorytmów w środowisku obliczeniowym MATLAB.
- Przeprowadzono badania eksperymentalne algorytmów ilustrujące zachowanie się metod dla małej i umiarkowanej liczby obserwacji wejścia i wyjścia systemu.

7.2 Otwarte problemy badawcze

Prezentowane w rozprawie algorytmy, wraz z rezultatami teoretycznymi dotyczącymi ich własności asymptotycznych i wynikami eksperymentów numerycznych, prowadzą do następujących, nowych problemów badawczych:

- Określenie dalszych własności teoretycznych proponowanych algorytmów. W szczególności:
  - opracowanie wersji algorytmu wykorzystującego kryterium średniokwadratowe umożliwiającej identyfikację systemów Wienera bez konieczności dzielenia zbioru dostępnych obserwacji wejścia i wyjścia systemu,
  - określenie rzędu szybkości zbieżności algorytmu wykorzystującego kryterium średniokwadratowe,
  - określenie własności asymptotycznych algorytmu wykorzystującego ideę korelacji rangowej,
  - analiza własności asymptotycznych algorytmów w przypadku sygnałów wejściowych innych niż sygnały typu *i.i.d.*,
  - analiza własności asymptotycznych algorytmów w przypadku szumu $\{Z_n\}$ nie będącego ciągiem typu *i.i.d.*
- Modyfikacja zaproponowanych algorytmów umożliwiająca identyfikację systemów Wienera w przypadku skończonej pamięci o *nieznanej długości.*
- Analiza zachowania algorytmów dla systemów Wienera o *nieskończonej pamięci.*

Dodatek A

Uzupełnienia i dowody twierdzeń

A.1 Uzupełnienia do rozdziału 3

Lemat A.1 *(por. Pawlak et al. [93]) Niech* $\{\xi_n\}_{n=1}^N$ *będzie ciągiem zmiennych losowych, takim że* $Var\{\xi_n\} < \infty$, $n = 1, 2, \ldots, N$ *oraz* ξ_n *i* ξ_m *są niezależne jeżeli* $|n - m| > p$, $p \in \mathbf{N}, p < \infty$. *Wtedy*

$$Var\left\{\sum_{n=1}^{N} \xi_n\right\} \leq (p+1) \sum_{n=1}^{N} Var\{\xi_n\}. \tag{A.1}$$

Dowód. Rozważmy przypadek gdy $E\{\xi_n\} = 0, n = 1, 2, \ldots, N$. Zapiszemy sumę wyrazów ciągu $\{\xi_n\}_{n=1}^N$ w następujący sposób

$$\sum_{n=1}^{N} \xi_n = \sum_{m=1}^{p+1} A_m,$$

gdzie $A_m = (\xi_m + \xi_{m+(p+1)} + \xi_{m+2(p+1)} + \ldots + \xi_{m+s(p+1)})$ i $s = \left\lfloor \frac{N-m}{p+1} \right\rfloor$. Stosując nierówność Cauchy'ego (zob. np. Leja [75], Jakubowski i Sztencel [60]), otrzymujemy

$$E\left\{\left(\sum_{m=1}^{p+1} A_m\right)^2\right\} \leq (p+1) \sum_{m=1}^{p+1} E\{A_m^2\},$$

co przy założeniu że $E\{\xi_n\} = 0, n = 1, 2, \ldots, N$, daje

$$Var\left\{\sum_{n=1}^{N} \xi_n\right\} \leq (p+1) \sum_{m=1}^{p+1} Var\{A_m\}.$$

Zauważając następnie, że zmienne losowe wchodzące w skład elementu A_m, $m \in \{1, 2, \ldots, p+1\}$ są niezależne, otrzymujemy nierówność (A.1). W przypadku ogólnym, tj. gdy nie zachodzi warunek $E\{\xi_n\} = 0, n = 1, 2, \ldots, N$, nierówność (A.1) wynika z faktu, że $Var\{\xi_m - a\} = Var\{\xi_m\}$, $\forall a \in \mathbf{R}$. ■

Dowód Lematu 3.2. Rozpoczniemy od wykazania tezy twierdzenia w przypadku gdy $\mu_\xi = 0$ oraz $\sigma_\xi^2 = 1$. Równość (3.24) oznacza[1], że dla dowolnie małego $\varepsilon > 0$ oraz dowolnie małego $\delta > 0$,

$$\lim_{N \to \infty} P\left\{N^{-\varepsilon} \sup_{1 \leq n \leq N} |\xi_n| > \delta\right\} = 0.$$

[1] definicja symbolu $o_P(\cdot)$ zamieszczona jest w dodatku B na str. 87.

Zauważmy, że

$$P\left\{N^{-\varepsilon}\sup_{1\leq n\leq N}|\xi_n|>\delta\right\} = 1-P\left\{N^{-\varepsilon}\sup_{1\leq n\leq N}|\xi_n|\leq\delta\right\}$$
$$= 1-P\left\{\left(N^{-\varepsilon}|\xi_1|\leq\delta\right)\cap\ldots\cap\left(N^{-\varepsilon}|\xi_N|\leq\delta\right)\right\}.$$

Ze względu na niezależność zmiennych losowych $\{\xi_n\}_{n=1}^N$ mamy dalej

$$P\left\{N^{-\varepsilon}\sup_{1\leq n\leq N}|\xi_n|>\delta\right\} = 1-\left[P\left\{N^{-\varepsilon}|\xi_1|\leq\delta\right\}\right]^N = 1-\left[P\left\{|\xi_1|\leq\delta N^{\varepsilon}\right\}\right]^N.$$

Oznaczając przez $\Phi(\cdot)$ dystrybuantę rozkładu normalnego $\mathcal{N}(0,1)$ otrzymujemy następnie, że

$$P\left\{N^{-\varepsilon}\sup_{1\leq n\leq N}|\xi_n|>\delta\right\} = 1-\left[2\Phi\left(\delta N^{\varepsilon}\right)-1\right]^N.$$

W celu udowodnienia tezy (3.24) wystarczy więc pokazać, że dla dowolnie małego $\varepsilon>0$ i dowolnie małego $\delta>0$

$$\lim_{N\to\infty}\left[2\Phi\left(\delta N^{\varepsilon}\right)-1\right]^N = 1.$$

Zauważmy, że ciąg $\left[2\Phi\left(\delta N^{\varepsilon}\right)-1\right]^N$ jest ciągiem ograniczonym z góry przez 1. Wykorzystamy następujące oszacowanie dystrybuanty rozkładu normalnego, zob. Jakubowski i Sztencel [60, str. 119],

$$\Phi(x) > 1-\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\frac{1}{x}, \qquad x>0.$$

Zatem

$$2\Phi\left(\delta N^{\varepsilon}\right)-1 > 1-\frac{1}{\delta N^{\varepsilon}}\exp\left(-\frac{\delta^2N^{2\varepsilon}}{2}\right),$$

i tym samym

$$\left[2\Phi\left(\delta N^{\varepsilon}\right)-1\right]^N > \left[1-\frac{1}{\delta N^{\varepsilon}}\exp\left(-\frac{\delta^2N^{2\varepsilon}}{2}\right)\right]^N. \tag{A.2}$$

Zauważmy następnie, że ciąg $\frac{1}{\delta N^{\varepsilon}}\exp\left(-\frac{\delta^2N^{2\varepsilon}}{2}\right)$ dąży do zera asymptotycznie szybciej niż ciąg $1/N^2$. Stąd, jeżeli zachodzi zbieżność

$$\lim_{N\to\infty}\left[1-\frac{1}{N^2}\right]^N = 1, \tag{A.3}$$

to również prawa strona nierówności (A.2) dąży do 1, gdy $N\to\infty$. Aby pokazać zbieżność w (A.3) wystarczy z kolei zauważyć, że

$$\lim_{N\to\infty}\left[1-\frac{1}{N^2}\right]^N = \left(\lim_{N\to\infty}\left[1-\frac{1}{N}\right]^N\right)\left(\lim_{N\to\infty}\left[1+\frac{1}{N}\right]^N\right) = e^{-1}e^1 = 1.$$

Na koniec, dla dowolnego μ_ξ i σ_ξ^2, otrzymujemy $N^{-1/2}\sup_{1\leq n\leq N}|\xi_n| = \sigma_\xi N^{-1/2}\times$ $\times\sup_{1\leq n\leq N}\left|\frac{\xi_n-\mu_\xi}{\sigma_\xi}\right|+\mu_\xi N^{-1/2} = o_P\left(N^{-(1/2)+\varepsilon}\right)+O\left(N^{-1/2}\right) = o_P\left(N^{-(1/2)+\varepsilon}\right)$, co kończy dowód. ■

Dowód Twierdzenia 3.3. W dowodzie wykorzystamy dekompozycję (3.26), określając najpierw rząd szybkości zbieżności estymatora $\tilde{r}_N(v)$ do $r(v)$, a następnie rząd szybkości

zbieżności do zera różnicy $[\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)]$. Całkując przez podstawienie, zapiszemy wartość oczekiwaną estymatora $\tilde{r}_N(v)$ w postaci

$$E\{\tilde{r}_N(v)\} = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v-x}{h}\right) r(x)\,dx = \int_{-\infty}^{\infty} K(x)\, r(v-hx)\,dx.$$

Stosując twierdzenie Taylora (zob. np. Leja [75], Wand i Jones [132]), na mocy założenia 3.6, zauważamy że

$$r(v-hx) = r(v) - hxr'(v) + \frac{1}{2}h^2x^2r''(v) + o\left(h^2\right),$$

co z kolei prowadzi do wniosku, że (zob. założenia 3.4 i 3.8 o funkcji jądra)

$$E\{\tilde{r}_N(v)\} = r(v) + \frac{1}{2}h^2r''(v)\int_{-\infty}^{\infty} x^2K(x)\,dx + o\left(h^2\right).$$

Stąd, zgodnie z założeniem 3.5,

$$bias\{\tilde{r}_N(v)\} = E\{\tilde{r}_N(v)\} - r(v) = O\left(N^{-2/5}\right).$$

W rezultacie, wykorzystując oszacowanie wariancji estymatora $\tilde{r}_N(v)$ uzyskane w dowodzie twierdzenia 3.2 (patrz wzór (3.32)), otrzymujemy rząd szybkości zbieżności

$$[\tilde{r}_N(v) - r(v)] = O_P\left(N^{-2/5}\right),$$

w każdym punkcie $v \in \mathbf{R}$. Udowodnimy z kolei, że

$$[\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)] = O_P\left(N^{(-2/5)+\varepsilon}\right). \tag{A.4}$$

W tym celu wykorzystamy fakt dwukrotnej różniczkowalności jądra $K(\cdot)$. Rozwijając funkcję $K(\cdot)$ w szereg Taylora

$$K(u_1) = K(u_2) + K'(u_2)(u_1-u_2) + \frac{1}{2}K''(u_0)(u_1-u_2)^2,$$

gdzie u_0 jest punktem pośrednim leżącym pomiędzy u_1 i u_2, możemy różnicę $[\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)]$ zapisać w postaci

$$\frac{1}{N_I h}\sum_{n\in I} Y_n\left[\frac{1}{h}K'\left(\frac{v-\alpha V_n}{h}\right)\left(\alpha V_n - \hat{V}_{n,N}^{Cr}\right) + \frac{1}{2h^2}K''\left(\frac{v-\xi_{n,N}}{h}\right)\left(\alpha V_n - \hat{V}_{n,N}^{Cr}\right)^2\right], \tag{A.5}$$

gdzie $\xi_{n,N}$ jest zmienną losową przyjmującą wartości[2] pomiędzy αV_n i $\hat{V}_{n,N}^{Cr}$. Reprezentacja (A.5) prowadzi po zastosowaniu nierówności trójkąta do następującego oszacowania

$$\left|\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)\right| \le \left|\tilde{r}_N'(v)\right| \sup_{n\in I}|d_{n,N}| + \sup_{n\in I}\left\{d_{n,N}^2\right\}\left|\frac{1}{N_I h^3}\sum_{n\in I} Y_n K''\left(\frac{v-\xi_{n,N}}{h}\right)\right|,$$

gdzie $d_{n,N} = \alpha V_n - \hat{V}_{n,N}^{Cr}$ oraz $\tilde{r}_N'(v)$ jest pochodną estymatora $\tilde{r}_N(v)$, tj. $\tilde{r}_N(v) = d\tilde{r}_N(v)/dv$. Na podstawie założenia 3.8 o ograniczoności drugiej pochodnej jądra $K(\cdot)$, prawdziwe jest oszacowanie

$$\left|\hat{r}_N^{Cr}(v) - \tilde{r}_N(v)\right| \le \left|\tilde{r}_N'(v)\right| \sup_{n\in I}|d_{n,N}| + \sup_{n\in I}\left\{d_{n,N}^2\right\}\frac{k_1}{N_I h^3}\sum_{n\in I}|Y_n|. \tag{A.6}$$

Odnośnie drugiego składnika po prawej stronie nierówności (A.6) zauważamy, że:

[2] tj. $\xi_{n,N} = \xi_{n,N}(\omega)$ jest taką zmienną losową, że dla każdego zdarzenia ω, $\min\{\alpha V_n(\omega), \hat{V}_{n,N}^{Cr}(\omega)\} \le \xi_{n,N}(\omega) \le \max\{\alpha V_n(\omega), \hat{V}_{n,N}^{Cr}(\omega)\}$.

- Na mocy lematu 3.3, dla dowolnie małego $\varepsilon > 0$,

$$\sup_{n\in I}\left\{d_{n,N}^2\right\} = \left(\sup_{n\in I}|d_{n,N}|\right)^2 = o_P\left(N^{-1+\varepsilon}\right).$$

- Zgodnie z założeniem 3.5,

$$h^{-3} \sim N^{3/5}.$$

- Stacjonarność sygnału $\{Y_n\}$ oraz skończona długość pamięci systemu prowadzą do wniosku, że

$$N_I^{-1}\sum_{n\in I}|Y_n| = E|Y_1| + O_P\left(N^{-1/2}\right).$$

Wobec tego

$$\sup_{n\in I}\left\{d_{n,N}^2\right\}\frac{k_1}{N_I h^3}\sum_{n\in I}|Y_n| = O_P\left(N^{-(2/5)+\varepsilon}\right).$$

W celu określenia rzędu szybkości zbieżności w (A.4) wystarczy więc udowodnić, że estymator $\tilde{r}'_N(v)$ jest ograniczony według prawdopodobieństwa[3] w każdym punkcie $v \in \mathbf{R}$. W tym celu wykażemy zbieżność $\tilde{r}'_N(v)$ według prawdopodobieństwa do $r'(v)$.
Zauważmy, że zgodnie z przyjętymi założeniami

$$\int_{-\infty}^{\infty}\frac{1}{h}K'\left(\frac{v-x}{h}\right)r(x)\,dx = \int_{-\infty}^{\infty}K\left(\frac{v-x}{h}\right)r'(x)\,dx, \tag{A.7}$$

oraz że całka $\int_{-\infty}^{\infty}|r'(v)|\,dv$ jest skończona. Stąd na mocy lematu B.3 w dodatku B.2,

$$E\left\{\tilde{r}'_N(v)\right\} = \frac{1}{h^2}\int K'\left(\frac{v-x}{h}\right)r(x)\,dx \to r'(v), \qquad \text{gdy } N\to\infty,$$

w każdym punkcie $v \in \mathbf{R}$. W celu określenia własności asymptotycznych wariancji estymatora $\tilde{r}'_N(v)$, zastosujemy lemat A.1 z dodatku A.1. Otrzymujemy

$$Var\left\{\tilde{r}'_N(v)\right\} \le \frac{p+1}{N_I h^4}Var\left\{Y_1 K'\left(\frac{v-V_1}{h}\right)\right\}.$$

Wystarczy zatem pokazać, że funkcja

$$\phi_N(v) \overset{def.}{=} \frac{1}{h^2}E\left\{Y_1^2 K'^2\left(\frac{v-V_1}{h}\right)\right\},$$

zbiega do stałej, gdy $N\to\infty$, dla każdego $v \in \mathbf{R}$. Zauważmy, że

$$\phi_N(v) \le \frac{k}{h^2}E\left\{g^2(V_1)K'\left(\frac{v-V_1}{h}\right)\right\} + \frac{k\sigma_Z^2}{h^2}E\left\{K'\left(\frac{v-V_1}{h}\right)\right\}.$$

Wykorzystując argumentację jak w (A.7) oraz lemat B.3 otrzymujemy $\phi_N(v)\to c(v) < \infty$, gdy $N\to\infty$. Uzyskane rezultaty prowadzą do wniosku że

$$\hat{r}_N^{Cr}(v) = r(v) + O_P\left(N^{-(2/5)+\varepsilon}\right), \tag{A.8}$$

dla dowolnie małego $\varepsilon > 0$ w każdym punkcie $v \in \mathbf{R}$. Zauważmy dalej, że analogiczna argumentacja dla estymatora $\hat{f}_N^{Cr}(v)$ prowadzi do wniosku, że

$$\hat{f}_N^{Cr}(v) = \alpha^{-1}f(v/\alpha) + O_P\left(N^{-(2/5)+\varepsilon}\right), \tag{A.9}$$

dla dowolnie małego $\varepsilon > 0$ w każdym punkcie $v \in \mathbf{R}$. Zbieżność w (3.34) wynika zatem bezpośrednio z (A.8) i (A.9) oraz lematu B.1 w dodatku B.1, co kończy dowód. ■

[3] Ciąg zmiennych losowych $\{X_n\}$ jest ograniczony według prawdopodobieństwa, jeżeli dla każdego $\varepsilon > 0$ istnieje stała M taka, że $\sup_n P\{|X_n| > M\} < \varepsilon$, Van der Vaart [123, str. 8].

A.2 Uzupełnienia do rozdziału 4

Twierdzenie A.1 *(zob. Van der Vaart [123, tw. 5.7, str. 45])* *Niech $M_N(\theta)$ będzie ciągiem funkcji losowych argumentu $\theta \in \Theta$, gdzie Θ jest przestrzenią metryczną z metryką $d(\cdot,\cdot)$. Niech ponadto $M(\theta)$ będzie funkcją deterministyczną, taką że*

$$\sup_{\theta\in\Theta} |M_N(\theta) - M(\theta)| \to 0, \quad \text{gdy } N \to \infty, \tag{A.10}$$

według prawdopodobieństwa oraz dla każdego $\varepsilon > 0$,

$$\inf_{\theta: d(\theta,\theta_0)\geq\varepsilon} M(\theta) > M(\theta_0), \tag{A.11}$$

gdzie θ_0 jest elementem przestrzeni Θ. Wtedy dowolny ciąg estymatorów $\left\{\hat{\theta}_N\right\}$, dla którego istnieje zbieżny do zera (według prawdopodobieństwa) ciąg zmiennych losowych $\{\xi_N\}$, taki że

$$M_N\left(\hat{\theta}_N\right) \leq M_N(\theta_0) + \xi_N, \tag{A.12}$$

zbiega według prawdopodobieństwa do θ_0, gdy $N \to \infty$.

Dowód. (por. Van der Vaart [123, tw. 5.7, str. 46]) Należy wykazać, że $d\left(\hat{\theta}_N, \theta_0\right) \to 0$ według prawdopodobieństwa gdy $N \to \infty$. Z warunku A.11 wynika, że dla każdego $\varepsilon > 0$ istnieje takie $\eta > 0$, że dla każdego $\theta : d(\theta, \theta_0) \geq \varepsilon$,

$$M(\theta) > M(\theta_0) + \eta.$$

Wobec tego zdarzenie $\{d(\hat{\theta}_N, \theta_0) \geq \varepsilon\}$ implikuje zdarzenie $A = \{M(\hat{\theta}_N) - M(\theta_0) > \eta\}$ i w celu udowodnienia tezy twierdzenia wystarczy pokazać że $P\{A\} \to 0$, gdy $N \to \infty$. Mamy

$$M\left(\hat{\theta}_N\right) - M(\theta_0) = \left[M\left(\hat{\theta}_N\right) - M_N\left(\hat{\theta}_N\right)\right] + \left[M_N\left(\hat{\theta}_N\right) - M(\theta_0)\right], \tag{A.13}$$

Z warunku (A.10) wynika, że $\left[M\left(\hat{\theta}_N\right) - M_N\left(\hat{\theta}_N\right)\right] \to 0$ według prawdopodobieństwa, gdy $N \to \infty$. Odnośnie drugiego składnika po prawej stronie równości (A.13) zauważamy, że zgodnie z założeniem A.12,

$$\begin{aligned} M_N\left(\hat{\theta}_N\right) - M(\theta_0) &= \left[M_N\left(\hat{\theta}_N\right) - M_N(\theta_0)\right] + [M_N(\theta_0) - M(\theta_0)] \\ &\leq \xi_N + [M_N(\theta_0) - M(\theta_0)]. \end{aligned}$$

Wykorzystując ponownie warunek (A.10) otrzymujemy zbieżność $[M_N(\theta_0) - M(\theta_0)] \to 0$ według prawdopodobieństwa, gdy $N \to \infty$ i wobec tego $P\{A\} \to 0$ gdy $N \to \infty$, co kończy dowód twierdzenia.

Dowód Twierdzenia 4.2. (por. dowód tw. 3.2) Przyjmiemy następujące oznaczenia

$$\hat{r}_N^{LS}(v) = \frac{1}{N_I h} \sum_{i\in I} Y_i K\left(\frac{v - \hat{V}_{i,N}^{LS}}{h}\right) \quad \text{oraz} \quad \hat{f}_N^{LS}(v) = \frac{1}{N_I h} \sum_{i\in I} K\left(\frac{v - \hat{V}_{i,N}^{LS}}{h}\right).$$

Dodatkowo oznaczymy $r(v) = g(v) f_V(v)$, gdzie $f_V(\cdot)$ jest gęstością prawdopodobieństwa sygnału interakcyjnego $\{V_n\}$. Zbieżność w (4.14) zachodzi, jeżeli $\hat{r}_N^{LS}(v)$ i $\hat{f}_N^{LS}(v)$ zbiegają

według prawdopodobieństwa odpowiednio do $r(v)$ i $f_V(v)$ przy $N \to \infty$ w każdym punkcie $v \in \mathbf{R}$, w którym $f_V(v) > 0$. W dowodzie wykorzystamy dekompozycję

$$\hat{r}_N^{LS}(v) - r(v) = \left[\hat{r}_N^{LS}(v) - \tilde{r}_N^{LS}(v)\right] + \left[\tilde{r}_N^{LS}(v) - r(v)\right], \tag{A.14}$$

gdzie

$$\tilde{r}_N^{LS}(v) = \frac{1}{N_I h} \sum_{i \in I} Y_i K\left(\frac{v - V_i}{h}\right).$$

Zgodnie z przyjętymi założeniami odnośnie sygnału wejściowego i szumu na wyjściu otrzymujemy

$$E\left\{\tilde{r}_N^{LS}(v)\right\} = \frac{1}{h} \int_{-\infty}^{\infty} K\left(\frac{v - x}{h}\right) r(x)\, dx.$$

Ponieważ nieliniowa charakterystyka $g(\cdot)$ jest funkcją ciągłą (zob. założenie 4.2) oraz sygnał wejściowy jest ograniczony (założenie 4.1), to $\int_{-\infty}^{\infty} |r(v)|\, dv < \infty$. Stąd, zgodnie z lematem B.3 z dodatku B.3, na mocy założeń 4.7 i 4.8, ma miejsce następująca zbieżność

$$E\left\{\tilde{r}_N^{LS}(v)\right\} \to r(v), \quad \text{gdy } N \to \infty, \tag{A.15}$$

w każdym punkcie $v \in \mathbf{R}$. Lemat A.1 z dodatku A.1 oraz własność 3.2 (str. 30) prowadzą następnie do wniosku, że

$$Var\left\{\tilde{r}_N^{LS}(v)\right\} \leq \frac{p+1}{N_I h^2} Var\left\{Y_1 K\left(\frac{v - V_1}{h}\right)\right\}. \tag{A.16}$$

Ze względu na ograniczoność jądra $K(\cdot)$, element $h^{-1} Var\left\{Y_1 K\left([v - V_1]/h\right)\right\}$ występujący w (A.16) jest ograniczony przez (por. Greblicki [25, dowód tw. 1])

$$\frac{2k_0}{h} E\left\{g^2(V_1) K\left(\frac{v - V_1}{h}\right)\right\} + \frac{2k_0 \sigma_Z^2}{h} E\left\{K\left(\frac{v - V_1}{h}\right)\right\}. \tag{A.17}$$

Ponowne wykorzystanie założeń 4.2 i 4.1 prowadzi do wniosku, że $\int_{-\infty}^{\infty} g^2(v) f_V(v)\, dv < \infty$. Zatem, na mocy lematu B.3 z dodatku B.3, wyrażenie (A.17) dąży do $k_0 f_V(v)\left[g^2(v) + \right.$ $\left. + \sigma_Z^2\right] < \infty$, gdy $N \to \infty$, w każdym punkcie $v \in \mathbf{R}$. Ostatecznie więc nierówność (A.16), dla h spełniającego założenie 4.8, daje

$$Var\left\{\tilde{r}_N^{LS}(v)\right\} \to 0, \quad \text{gdy } N \to \infty, \tag{A.18}$$

w każdym punkcie $v \in \mathbf{R}$.
Zgodnie z rezultatami w (A.15) i (A.18) zachodzi więc zbieżność

$$\left[\tilde{r}_N^{LS}(v) - r(v)\right] \to 0, \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$.
Przejdziemy teraz do do drugiego etapu dowodu, tj. do zbadania asymptotyki wyrażenia $\left[\hat{r}_N^{LS}(v) - \tilde{r}_N^{LS}(v)\right]$ w dekompozycji (A.14). Korzystając z faktu, że jądro $K(\cdot)$ spełnia warunek Lipschitza (zob. zał. 4.7) otrzymujemy

$$\left|\hat{r}_N^{LS}(v) - \tilde{r}_N^{LS}(v)\right| \leq \sup_{n \in I} \left|V_n - \hat{V}_{n,N}^{LS}\right| \frac{L}{N_I h^2} \sum_{i \in I} |Y_i|.$$

Wzór (4.10) oraz nierówność Cauchy'ego–Schwarza (zob. dodatek B.4) prowadzą do wniosku, że

$$\begin{aligned} \sup_{n \in I} \left|V_n - \hat{V}_{n,N}^{LS}\right| &= \sup_{n \in I} \left|\left[\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right]^T \underline{X}_n\right| \\ &\leq \sup_{n \in I} \left\{\left\|\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right\| \|\underline{X}_n\|\right\} \\ &= C \left\|\underline{\lambda} - \hat{\underline{\lambda}}_N^{LS}\right\|, \end{aligned}$$

gdzie $C = (p+1)\max\{|a_f|, |b_f|\}$. Zatem, na mocy (4.11), $\sup_{n\in I}\left|V_n - \hat{V}_{n,N}^{LS}\right| = O_P(a_N)$. Równocześnie (por. wzór (3.33))

$$\frac{1}{N_I}\sum_{i\in I}|Y_i| = E\{|Y_1|\} + O_P\left(N^{-1/2}\right).$$

Otrzymujemy więc zbieżność (zob. założenie 4.8)

$$\left|\hat{r}_N^{LS}(v) - \tilde{r}_N^{LS}(v)\right| \to 0, \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$. W rezultacie

$$\hat{r}_N^{LS}(v) \to r(v), \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa w każdym punkcie $v \in \mathbf{R}$.
Stosując analogiczną argumentację dla estymatora $\hat{f}_N(v)$ otrzymujemy zbieżność

$$\hat{f}_N^{LS}(v) \to f_V(v), \quad \text{gdy } N \to \infty,$$

według prawdopodobieństwa, w każdym punkcie $v \in \mathbf{R}$, co kończy dowód twierdzenia. ■

A.3 Uzupełnienia do rozdziału 5

Dowód Lematu 5.1. (zob. też dowód lematu 2.1 w [95]) Należy wykazać, że

$$\int_{\mathbf{R}^{p+1}} f_{\underline{X}}^2(\underline{x})\, D_G(\underline{x})\, d\underline{x} = -2\int_{\mathbf{R}^{p+1}} D_f(\underline{x})\, G(\underline{x})\, f_{\underline{X}}(\underline{x})\, d\underline{x}. \tag{A.19}$$

Zgodnie z założeniem 5.1 funkcje podcałkowe w (A.19) mają nośnik Ω. Ponadto $G(\underline{x})$ i $f_{\underline{X}}(\underline{x})$ są różniczkowalne w zbiorze Ω. Stąd, na mocy lematu B.4 w dodatku B.3, otrzymujemy tezę twierdzenia. ■

Dowód Twierdzenia 5.3. W dowodzie wykorzystamy dekompozycję (5.21), określając najpierw rząd szybkości zbieżności estymatora $\tilde{r}_N(v)$, a następnie rząd szybkości zbieżności do zera różnicy $\left[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right]$.
Całkując przez podstawienie zapiszemy wartość oczekiwaną estymatora $\tilde{r}_N(v)$ w postaci

$$E\{\tilde{r}_N(v)\} = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{v-x}{h}\right) r(x)\, dx = \int_{-\infty}^{\infty} K(x)\, r(v-hx)\, dx.$$

Zgodnie z rozwinięciem Taylora, na mocy założenia 5.5,

$$r(v-hx) = r(v) - hxr'(v) + \frac{1}{2}h^2x^2r''(v) + o\left(h^2\right),$$

co z kolei prowadzi do wniosku, że (zob. założenia 5.3 i 5.7 o funkcji jądra)

$$E\{\tilde{r}_N(v)\} = r(v) + \frac{1}{2}h^2r''(v)\int_{-\infty}^{\infty} x^2K(x)\, dx + o\left(h^2\right).$$

Stąd, ponieważ $h \sim N^{-1/5}$,

$$bias\{\tilde{r}_N(v)\} = E\{\tilde{r}_N(v)\} - r(v) = O\left(N^{-2/5}\right).$$

W rezultacie, wykorzystując oszacowanie wariancji estymatora $\tilde{r}_N(v)$ uzyskane w dowodzie twierdzenia 5.2, (patrz wzór (5.26)) otrzymujemy rząd szybkości zbieżności

$$\left[\tilde{r}_N(v) - r(v)\right] = O_P\left(N^{-2/5}\right),$$

w każdym punkcie $v \in \mathbf{R}$. Udowodnimy następnie, że

$$\left[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right] = O_P\left(N^{-2/5}\right). \tag{A.20}$$

W tym celu wykorzystamy fakt dwukrotnej różniczkowalności jądra $K(\cdot)$. Rozwijając funkcję $K(\cdot)$ w szereg Taylora

$$K(u_1) = K(u_2) + K'(u_2)(u_1 - u_2) + \frac{1}{2}K''(u_0)(u_1 - u_2)^2,$$

gdzie u_0 jest punktem pośrednim leżącym pomiędzy u_1 i u_2, różnicę $\left[\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right]$ zapiszemy w postaci

$$\frac{1}{N_I h}\sum_{n\in I} Y_n\left[\frac{1}{h}K'\left(\frac{v-\kappa V_n}{h}\right)\left(\kappa V_n - \hat{V}_{n,N}^{AD}\right) + \frac{1}{2h^2}K''\left(\frac{v-\xi_{n,N}}{h}\right)\left(\kappa V_n - \hat{V}_{n,N}^{AD}\right)^2\right], \tag{A.21}$$

gdzie $\xi_{n,N}$ jest zmienną losową przyjmującą wartości[4] pomiędzy κV_n i $\hat{V}_{n,N}^{AD}$. Reprezentacja (A.21) prowadzi do następującego oszacowania

$$\left|\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right| \le \left|\tilde{r}'_N(v)\right| \sup_{n\in I}|d_{n,N}| + \sup_{n\in I}\left\{d_{n,N}^2\right\}\left|\frac{1}{N_I h^3}\sum_{n\in I} Y_n K''\left(\frac{v-\xi_{n,N}}{h}\right)\right|,$$

gdzie $d_{n,N} = \kappa V_n - \hat{V}_{n,N}^{AD}$ oraz $\tilde{r}'_N(v)$ jest pochodną estymatora $\tilde{r}_N(v)$, tj. $\tilde{r}'_N(v) = d\tilde{r}_N(v)/dv$. Po uwzględnieniu założenia 5.7 otrzymujemy dalej

$$\left|\hat{r}_N^{AD}(v) - \tilde{r}_N(v)\right| \le \left|\tilde{r}'_N(v)\right| \sup_{n\in I}|d_{n,N}| + \sup_{n\in I}\left\{d_{n,N}^2\right\}\frac{k_1}{N_I h^3}\sum_{n\in I}|Y_n|. \tag{A.22}$$

Na mocy lematu 5.2 $\sup_{n\in I}\left\{d_{n,N}^2\right\} = \left(\sup_{n\in I}|d_{n,N}|\right)^2 = O_P\left(N^{-1}\right)$. Zachodzi również zbieżność (por. wzór (3.33) w dowodzie tw. 3.2, str. 29)

$$N_I^{-1}\sum_{n\in I}|Y_n| = E|Y_1| + O_P\left(N^{-1/2}\right).$$

Stąd, na mocy założenia 5.4,

$$\sup_{n\in I}\left\{d_{n,N}^2\right\}\frac{k_1}{N_I h^3}\sum_{n\in I}|Y_n| = O_P\left(N^{-2/5}\right).$$

W celu wykazania rzędu szybkości zbieżności jak we wzorze (A.20) wystarczy więc udowodnić, że estymator $\tilde{r}'_N(v)$ jest ograniczony według prawdopodobieństwa[5] w każdym punkcie $v \in \mathbf{R}$. Wykażemy więc zbieżność $\tilde{r}'_N(v)$ według prawdopodobieństwa do $r'(v)$. Zgodnie z przyjętymi założeniami

$$\int_{-\infty}^{\infty}\frac{1}{h}K'\left(\frac{v-x}{h}\right)r(x)\,dx = \int_{-\infty}^{\infty}K\left(\frac{v-x}{h}\right)r'(x)\,dx, \tag{A.23}$$

oraz całka $\int_{-\infty}^{\infty}|r'(v)|\,dv$ jest skończona. Stąd na mocy lematu B.3 w dodatku B.3,

$$E\left\{\tilde{r}'_N(v)\right\} = \frac{1}{h^2}\int K'\left(\frac{v-x}{h}\right)r(x)\,dx \to r'(v), \qquad \text{gdy } N \to \infty,$$

[4] zob. przypis na stronie 79.

[5] zob. przypis na stronie 80.

w każdym punkcie $v \in \mathbf{R}$. W celu określenia własności asymptotycznych wariancji estymatora $\tilde{r}'_N(v)$, zastosujemy lemat A.1 z dodatku A.1. Otrzymujemy

$$Var\left\{\tilde{r}'_N(v)\right\} \leq \frac{p+1}{N_I h^4} Var\left\{Y_1 K'\left(\frac{v-\kappa V_1}{h}\right)\right\}.$$

Wystarczy zatem pokazać, że funkcja

$$\phi_N(v) \stackrel{def.}{=} \frac{1}{h^2} E\left\{Y_1^2 K'^2\left(\frac{v-V_1}{h}\right)\right\},$$

zbiega do stałej gdy $N \to \infty$, dla każdego $v \in \mathbf{R}$. Zauważmy, że (zob. założenie 5.3)

$$\phi_N(v) \leq \frac{L}{h^2} E\left\{g^2(V_1) K'\left(\frac{v-V_1}{h}\right)\right\} + \frac{L\sigma_Z^2}{h^2} E\left\{K'\left(\frac{v-V_1}{h}\right)\right\}.$$

Wykorzystując argumentację jak w (A.23) oraz lemat B.3, otrzymujemy $\phi_N(v) \to c(v) < \infty$, gdy $N \to \infty$. Uzyskane rezultaty prowadzą do wniosku, że

$$\hat{r}_N^{AD}(v) = r(v) + O_P\left(N^{-2/5}\right), \tag{A.24}$$

w każdym punkcie $v \in \mathbf{R}$. Analogiczna argumentacja dla estymatora $\hat{f}_N^{AD}(v)$ prowadzi do wniosku, że

$$\hat{f}_N^{AD}(v) = \kappa^{-1} f(v/\kappa) + O_P\left(N^{-2/5}\right), \tag{A.25}$$

w każdym punkcie $v \in \mathbf{R}$. Zbieżność w (A.20) wynika zatem ze wzorów (A.24) i (A.25) oraz lematu B.1 w dodatku B.1, co kończy dowód twierdzenia. ■

Lemat A.2 *Jeżeli spełnione jest założenie 5.1 oraz zachodzą zbieżności (5.32) i (5.33), to*

$$\sup_{\underline{x} \in \mathbf{\Omega}} \left|\hat{d}_{f,i+1,N}(\underline{x}) - d_{f,i+1}(\underline{x})\right| \to 0, \text{ gdy } N \to \infty,$$

według prawdopodobieństwa.

Dowód. Stosując elementarną nierówność[6] $|ab - cd| \leq |b|\,|a-c| + |c|\,|b-d|$ otrzymujemy następujące oszacowanie

$$\begin{aligned}\left|\hat{d}_{f,i+1,N}(\underline{x}) - d_{f,i+1}(\underline{x})\right| \leq\ & \left|\beta_{N,i}(\underline{x})\right| \left|\hat{\phi}_{X,N}\left(\underline{x}_{(i)}\right) - f'_X\left(\underline{x}_{(i)}\right)\right| + \\ & + \left|f'_X\left(\underline{x}_{(i)}\right)\right| \left|\beta_{N,i}(\underline{x}) - \delta_i(\underline{x})\right|,\end{aligned}$$

gdzie

$$\delta_i(\underline{x}) = \prod_{k=1,k\neq i}^{p+1} f_X\left(\underline{x}_{(k)}\right) \quad \text{oraz} \quad \beta_{N,i}(\underline{x}) = \prod_{j=1,j\neq i}^{p+1} \hat{f}_{X,N}\left(\underline{x}_{(j)}\right).$$

Zatem

$$\begin{aligned}\sup_{\underline{x} \in \mathbf{\Omega}} \left|\hat{d}_{f,i+1,N}(\underline{x}) - d_{f,i+1}(\underline{x})\right| \leq\ & \sup_{\underline{x} \in \mathbf{R}^p} \left|\beta_{N,i}(\underline{x})\right| \sup_{x \in \mathbf{R}} \left|\hat{\phi}_{X,N}(x) - f'_X(x)\right| + \\ & + c \sup_{\underline{x} \in \mathbf{R}^p} \left|\beta_{N,i}(\underline{x}) - \delta_i(\underline{x})\right|,\end{aligned}$$

gdzie $c = \sup_{x \in \mathbf{R}} |f'_X(x)|$. Zauważmy, że wobec założenia 5.1, $c < \infty$. z kolei zbieżność wyrażenia $\sup_{\underline{x} \in \mathbf{R}^p} \left|\beta_{N,i}(\underline{x})\right|$ do $\sup_{\underline{x} \in \mathbf{R}^p} |\delta_i(\underline{x})| < \infty$ wynika ze zbieżności jednostajnej

[6] Nierówność wynika z dekompozycji $ab - cd = b(a-c) + c(b-d)$ oraz nierówności trójkąta.

(tzn. w normie supremum) wyrażenia $|\beta_{N,i}(\underline{x}) - \delta_i(\underline{x})|$ do zera oraz ciągłości normy (w szczególności normy supremum). Zgodnie z nierównością (B.4) z dodatku B.4 otrzymujemy dla każdego $\underline{x} \in \mathbf{R}^p$, że

$$|\beta_{N,i}(\underline{x}) - \delta_i(\underline{x})| \leq \gamma_N^{p-1}(\underline{x}) \sum_{j=1}^{p} \left| \hat{f}_{X,N}\left(\underline{x}_{(j)}\right) - f_X\left(\underline{x}_{(j)}\right) \right|, \tag{A.26}$$

gdzie $\gamma_N(\underline{x}) = \max\{|\hat{f}_{X,N}(\underline{x}_{(1)})|, \ldots, |\hat{f}_{X,N}(\underline{x}_{(p)})|, |f_X(\underline{x}_{(1)})|, \ldots, |f_X(\underline{x}_{(p)})|\}$. Oszacowanie (A.26) prowadzi z kolei do wniosku, że

$$\sup_{\underline{x} \in \mathbf{R}^p} |\beta_{N,i}(\underline{x}) - \delta_i(\underline{x})| \leq p \sup_{\underline{x} \in \mathbf{R}^p} \left| \gamma_N^{p-1}(\underline{x}) \right| \sup_{x \in \mathbf{R}} \left| \hat{f}_{X,N}(x) - f_X(x) \right|$$

co, po ponownym zastosowaniu nierówności (B.4) z dodatku B.4, kończy dowód lematu. ■

Dodatek B

Wykorzystywane w pracy fakty z teorii prawdopodobieństwa i analizy matematycznej

B.1 Szybkość zbieżności ciągów zmiennych losowych

Definicja B.1 *Dla ciągu zmiennych losowych $\{\xi_n\}$ oraz ciągu liczb dodatnich $\{b_n\}$ zbieżnego do zera, symbol $\xi_n = O_P(b_n)$ oznacza, że ciąg zmiennych losowych $\{r_n(\xi_n/b_n)\}$ zbiega do zera według prawdopodobieństwa gdy $N \to \infty$, przy dowolnie wolno zbieżnym do zera ciągu liczbowym $\{r_n\}$.*

Definicja B.2 *Dla ciągu zmiennych losowych $\{\xi_n\}$ oraz ciągu liczb dodatnich $\{b_n\}$ zbieżnego do zera, symbol $\xi_n = o_P(b_n)$ oznacza, że ciąg zmiennych losowych $\{\xi_n/b_n\}$ zbiega do zera według prawdopodobieństwa gdy $N \to \infty$.*

Lemat B.1 *Niech $\{\xi_n\}$ oraz $\{\zeta_n\}$ będą ciągami zmiennych losowych takimi że $\xi_n = a + O_P(a_n)$ oraz $\zeta_n = b + O_P(b_n)$, gdzie $b \neq 0$. Wówczas*

$$\frac{\xi_n}{\zeta_n} = \frac{a}{b} + O_P(\max\{a_n, b_n\})$$

Dowód. zob. Hasiewicz i Śliwiński [55, str. 259]. ■

Lemat B.2 *Niech $\{\xi_n\}$ oraz $\{\zeta_n\}$ będą ciągami zmiennych losowych, takimi że $\xi_n = a + O_P(a_n)$ oraz $\zeta_n = b + O_P(b_n)$. Wówczas*

$$\xi_n\zeta_n = ab + O_P(\max\{a_n, b_n\})$$

Dowód. Należy pokazać, że dla dowolnie wolno zbieżnego do zera ciągu r_n oraz dowolnego $\varepsilon > 0$ zachodzi

$$\lim_{n\to\infty} P\left\{|r_n| \frac{|\xi_n\zeta_n - ab|}{\max\{a_n, b_n\}} > \varepsilon\right\} = 0.$$

Na podstawie dekompozycji

$$\xi_n\zeta_n - ab = \xi_n(\zeta_n - b) + b(\xi_n - a)$$

otrzymujemy następujące oszacowanie

$$|\xi_n\zeta_n - ab| \leq |\xi_n||\zeta_n - b| + |b||\xi_n - a|. \tag{B.1}$$

Niech $M \in \mathbf{R}_+$ będzie dowolną stałą, taką że $M > |a|$ i $M > |b|$. Załóżmy, że dla dowolnego $\varepsilon > 0$ i ciągu r_n dowolnie wolno zbieżnego do zera, zaszło zdarzenie określone za pomocą nierówności

$$|\xi_n - a| < \frac{1}{2M|r_n|} a_n \varepsilon \quad \text{i} \quad |\zeta_n - b| < \frac{1}{2M|r_n|} b_n \varepsilon \quad \text{i} \quad |\xi_n| < M.$$

Wtedy, zgodnie z nierównością (B.1), otrzymujemy

$$|\xi_n \zeta_n - ab| \leq M \frac{1}{2M|r_n|} b_n \varepsilon + M \frac{1}{2M|r_n|} a_n \varepsilon = \frac{1}{2|r_n|} \varepsilon (a_n + b_n),$$

co prowadzi do wniosku, że

$$|\xi_n \zeta_n - ab| \leq \varepsilon \frac{\max\{a_n, b_n\}}{|r_n|}.$$

Zachodzi więc następujące oszacowanie

$$P\left\{|\xi_n \zeta_n - ab| > \varepsilon \frac{\max\{a_n, b_n\}}{|r_n|}\right\} \leq \tag{B.2}$$
$$\leq P\left\{|\xi_n - a| \geq \frac{a_n}{2M|r_n|}\varepsilon\right\} + P\left\{|\zeta_n - b| \geq \frac{b_n}{2M|r_n|}\varepsilon\right\} + P\{|\xi_n| \geq M\}.$$

Na mocy założeń twierdzenia dwa pierwsze wyrazy po prawej stronie nierówności (B.2) dążą do zera gdy $n \to \infty$. Odnośnie wyrażenia $P\{|\xi_n| \geq M\}$ zachodzi natomiast

$$P\{|\xi_n| \geq M\} \leq P\{|\xi_n - a| + |a| \geq M\} = P\{|\xi_n - a| \geq M - |a|\}.$$

Ostatecznie zatem, ponieważ $\xi_n = a + O_P(a_n)$, otrzymujemy

$$\lim_{n \to \infty} P\{|\xi_n| \geq M\} = 0$$

co kończy dowód. ■

B.2 Wybrane własności zmiennych losowych gaussowskich

Twierdzenie B.1 *Niech $\xi_i, i = 1, 2, \ldots, n$ będą niezależnymi gaussowskimi zmiennymi losowymi o średnich odpowiednio μ_i oraz wariancjach σ_i^2, tj. $\xi_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Wtedy zmienna losowa $\xi = \sum_{i=1}^n \xi_i$ ma rozkład normalny $\mathcal{N}(\mu, \sigma^2)$ z parametrami*

$$\mu = \sum_{i=1}^n \mu_i \qquad oraz \qquad \sigma^2 = \sum_{i=1}^n \sigma_i^2$$

Dowód. Zob. np. Ostasiewicz *et al.* [86] str. 169. ■

Twierdzenie B.2 *Niech X i Y będą zmiennymi losowymi o łącznym rozkładzie normalnym. Niech $E\{X\} = \mu_X$, $E\{Y\} = \mu_Y$ oraz $Var\{X\} = \sigma_X^2$ i $Var\{Y\} = \sigma_Y^2$ wtedy*

$$E\{X|Y = y\} = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(y - \mu_X) \tag{B.3}$$

gdzie $\rho = Corr\{X, Y\}$.

Dowód. Zob. np. Papoulis [88] str. 201. ■

B.3 Wybrane zagadnienia teorii całki i analizy matematycznej

Lemat B.3 *(zob. Parzen [90])* *Niech $\varphi(\cdot)$ oznacza dowolną funkcję, dla której całka $\int_{-\infty}^{\infty} |\varphi(x)|\,dx$ jest skończona oraz $K(\cdot)$ będzie dodatnią funkcją borelowską, spełniającą warunki:*

1. *$K(\cdot)$ jest funkcją ograniczoną, tj. $\sup_{v\in\mathbf{R}} K(v) < \infty$,*
2. *$\int_{-\infty}^{\infty} K(v)\,dv = 1$,*
3. *$\lim_{|v|\to\infty} vK(v) = 0$.*

Wtedy dla dodatniego ciągu liczbowego h, zbieżnego do zera,

$$\lim_{h\to 0}\frac{1}{h}\int_{-\infty}^{\infty}\varphi(x)\,K\left(\frac{v-x}{h}\right)dx = \varphi(v),$$

w każdym punkcie $v \in \mathbf{R}$, w którym $\varphi(v)$ jest ciągła.

Dowód. Zob. np. Parzen [90]. ■

Lemat B.4 *(zob. Kołodziej [66, str. 340])* *Jeżeli u i v są funkcjami klasy C^1 zmiennych $\xi_1, \xi_2, \ldots, \xi_m$ określonymi na ograniczonym zbiorze otwartym Ω w przestrzeni euklidesowej $\mathbf{R}^m$ i nośnik przynajmniej jednej z tych funkcji zawarty jest w Ω, to*

$$\int_{\Omega} u\frac{\partial v}{\partial \xi_k}dx = -\int_{\Omega} v\frac{\partial u}{\partial \xi_k}dx \quad k = 1, 2, \ldots, m.$$

Dowód. Zob. Kołodziej [66, str. 340]. ■

B.4 Wykorzystywane nierówności

Lemat B.5 *(nierówność Cauchy'ego–Schwarza)* *Jeśli $\underline{v} \in H$ i $\underline{w} \in H$, gdzie H jest przestrzenią z iloczynem skalarnym, to*

$$|(\underline{v}, \underline{w})| \le \|\underline{v}\|\,\|\underline{w}\|,$$

gdzie $(\underline{v}, \underline{w})$ jest iloczynem skalarnym elementów $\underline{v}$ i $\underline{w}$.

Dowód. Zob. np. Rudin [100, str. 327] ■

Lemat B.6 *Dla dowolnych p–elementowych ciągów liczbowych $\{a_i\}$ i $\{b_i\}$ prawdziwa jest nierówność*

$$\left|\prod_{i=1}^{p} a_i - \prod_{i=1}^{p} b_i\right| \le \gamma^{p-1}\sum_{i=1}^{p}|a_i - b_i|, \tag{B.4}$$

gdzie $\gamma = \max\{|a_1|, \ldots, |a_p|, |b_1|, \ldots, |b_p|\}$.

Dowód. Nierówność (B.4) wynika z nierówności trójkąta oraz równości

$$\prod_{i=1}^{p} a_i - \prod_{i=1}^{p} b_i = \sum_{i=1}^{p} B_i A_i (a_i - b_i), \tag{B.5}$$

gdzie

$$B_i = \begin{cases} b_1 b_2 \dots b_{i-1} & \text{dla} \quad i > 1 \\ 1 & \text{dla} \quad i = 1 \end{cases} \qquad \text{oraz} \qquad A_i = \begin{cases} a_{i+1} a_{i+2} \dots a_p & \text{dla} \quad i < p \\ 1 & \text{dla} \quad i = p \end{cases}.$$

Wzór (B.5) można z kolei udowodnić stosując zasadę indukcji matematycznej i dekompozycję $a_1 a_2 - b_1 b_2 = a_2 (a_1 - b_1) + b_1 (a_2 - b_2)$. ■

Literatura

[1] E. W. Bai. Frequency domain identification of Hammerstein models. *IEEE Transactions on Automatic Control*, 48:530–542, 2003.

[2] N. J. Bershad, P. Celka, S. McLaughlin. Analysis of stochastic gradient tracking of time-varying polynomial Wiener systems. *IEEE Transactions on Signal Processing*, 48(6):1676–1686, 2000.

[3] N. J. Bershad, P. Celka, S. McLaughlin. Analysis of stochastic gradient identification of Wiener–Hammerstein systems for nonlinearities with Hermite polynomial expansions. *IEEE Transactions on Signal Processing*, 49:1060–1072, 2001.

[4] P. K. Bhattacharya. Estimation of a probability density function and its derivatives. *Sankhya Ser. A*, 29:373–382, 1967.

[5] S. A. Billings. Identification of nonlinear systems – a survey. *IEE Proceedings*, 127:272–285, 1980.

[6] S. A. Billings, S. Y. Fakhouri. Identification of nonlinear systems using the Wiener model. *Electronics Letters*, 13:502–504, 1977.

[7] S. A. Billings, S. Y. Fakhouri. The theory of separable processes with applications to the identification of nonlinear systems. *IEE Proceedings*, 125:1051–1058, 1978.

[8] S. A. Billings, S. Y. Fakhouri. Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15–26, 1982.

[9] S. Boyd, L. O. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, CAS-32(11):1150–1161, 1985.

[10] J. Bruls, C. T. Chou, B. R. J Haverkamp, M. Verhaegen. Linear and non–linear system identification using separable least–squares. *European Journal of Control*, 5:116–128, 1999.

[11] P. Celka, N. J. Bershad, J. M. Vesin. Fluctuation analysis of stochastic gradient identification of polynomial Wiener systems. *IEEE Transactions on Signal Processing*, 48(6):1820–1825, 2000.

[12] P. Celka, N. J. Bershad, J. M. Vesin. Stochastic gradient identification of polynomial Wiener systems: analysis and application. *IEEE Transactions on Signal Processing*, 49:301–313, 2001.

[13] P. Celka, P. Colditz. Nonlinear nonstationary Wiener model of infant EEG seizures. *IEEE Transactions on Biomedical Engineering*, 49(3):556–564, 2002.

[14] G. Chen, Y. Chen, H. Ogmen. Identifying chaotic systems via a Wiener-type cascade model. *IEEE Control Systems Magazine*, 17(5):29–36, 1997.

[15] H. Chen. Recursive identification for Wiener model with discontinuous piece-wise linear function. *IEEE Transactions on Automatic Control*, 51(3):390–400, 2006.

[16] A. C. den Brinker. A comparision of results from parameter estimations of impulse responses of the transient visual system. *Biological Cybernetics*, 61:139–151, 1989.

[17] L. Devroye. On the almost everywhere convergence of nonparametric regression function estimates. *The Annals of Statistics*, 9(6):1310–1319, 1981.

[18] L. Devroye, L. Györfi. *Nonparametric Density Estimation: the L1 view.* Wiley, New York, 1985.

[19] M. Enqvist. Linear models of nonlinear FIR systems with gaussian inputs. *Report no.: LiTH-ISY-R-2462, Linköpings Universitet, Sweden*, 2002.

[20] G. M. Fichtenholz. *Rachunek różniczkowy i całkowy, Tom I i II.* PWN, Warszawa, 1978.

[21] G. Geenens, L. Simar. Index coefficients estimation in Single–Index Models: the Generalized Maximum Rank Correlation Estimator. *Technical report*, 0560, Institut de Statistique Université Catholique de Louvain, Belgium, 2005.

[22] J. C. Gómez, E. Baeyens. Subspace identification of multivariable Hammerstein and Wiener models. *IFAC, 15th Triennial World Congress, Barcelona, Spain*, 2002.

[23] J. C. Gómez, A. Jutan, E. Baeyens. Wiener model identification and predictive control of a pH neutralisation process. *Control theory and applications, IEE Proc.*, 151(3):329–338, 2004.

[24] G. C. Goodwin, R. L. Payne. *Dynamic system identification. Experiment design and data analysis.* Academic Press, New York, USA, 1977.

[25] W. Greblicki. Nieparametryczna identyfikacja systemów. *Archiwum Automatyki i Robotyki*, XXXVI(2):277–290, 1991.

[26] W. Greblicki. Nonparametric identification of Wiener systems. *IEEE Transactions on Information Theory*, 38:1487–1493, 1992.

[27] W. Greblicki. Nonparametric identification of Wiener systems by orthogonal series. *IEEE Transactions on Automatic Control*, 39(10):2077–2086, 1994.

[28] W. Greblicki. Nonparametric approach to Wiener system identification. *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, 44(6):538–545, 1997.

[29] W. Greblicki. Continuous–time Wiener system identification. *IEEE Transactions on Automatic Control*, 43(10):1488–1493, 1998.

[30] W. Greblicki. Recursive identification of Wiener systems. *International Journal of Applied Mathematics and Computer Science*, 11(4):977–991, 2001.

[31] W. Greblicki. Recursive identification of continuous–time Hammerstein systems. *International Journal of Systems Science*, 33(12):969–977, 2002.

[32] W. Greblicki. Stochastic approximation in nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 47:100–111, 2002.

[33] W. Greblicki. Nonlinearity recovering in Wiener system driven with correlated signal. *IEEE Transactions on Automatic Control*, 49:1805–1810, 2004.

[34] W. Greblicki. *Podstawy automatyki.* Wydawnictwo Politechniki Wrocławskiej, Wrocław, 2006.

[35] W. Greblicki, A. Krzyżak, M. Pawlak. Distribution–free pointwise consistency of kernel regression estimate. *The Annals of Statistics*, 12(4):1570–1575, 1984.

[36] W. Greblicki, M. Pawlak. Hermite series estimates of a probability density and its derivatives. *Journal of Multivariate Analysis*, 15:174–182, 1984.

[37] W. Greblicki, M. Pawlak. Fourier and Hermite series estimates of regressin functions. *Annals of the Institute of Statistical Mathematics*, 37A:443–454, 1985.

[38] W. Greblicki, M. Pawlak. Identification of discrete Hammerstein systems using kernel regression estimates. *IEEE Transactions on Automatic Control*, 31:74–77, 1986.

[39] W. Greblicki, M. Pawlak. Nonparametric identification of Hammerstein systems. *IEEE Transactions on Information Theory*, 35:409–418, 1989.

[40] W. Greblicki, M. Pawlak. Nonparametric identification of a cascade nonlinear time series system. *Signal Processing*, 22:61–75, 1991.

[41] W. Greblicki, M. Pawlak. Nonparametric identification of particular nonlinear time series system. *IEEE Transactions on Signal Processing*, 40:985–989, 1992.

[42] W. Greblicki, M. Pawlak. Cascade non–linear system identification by a non–parametric method. *International Journal of Systems Science*, 25:129–153, 1994.

[43] W. Greblicki, M. Pawlak. Nonparametric recovering nonlinearities in block–oriented systems with the help of Laguerre polynomials. *Control–Theory and Advanced Technology*, 10:771–791, 1994.

[44] A. Hagenblad. *Aspects of the identification of Wiener systems, Thesis no.793.* Link opings Universitet, SE-581 83 Link oping, Sweden, 1999.

[45] A. K. Han. Non–parametric analysis of a generalized regression model. *Journal of Econometrics*, 35:303–316, 1987.

[46] W. Härdle. *Applied nonparametric regression.* Cambridge University Press, 1994.

[47] W. Härdle, M. Stoker. Investigating smooth multiple regression by the method of average derivative. *Journal of the American Statistical Association*, 84(408):986–995, 1989.

[48] Z. Hasiewicz. *Identyfikacja sterowanych systemów o złożonej strukturze, Prace naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej.* Wydawnictwo Politechniki Wrocławskiej, Wrocław, 1993.

[49] Z. Hasiewicz. Informacja wstępna o systemie a metody identyfikacji. *Instytut Cybernetyki Technicznej Politechniki Wrocławskiej, Preprinty 11/94*, 1994.

[50] Z. Hasiewicz. Hammerstein system identification by the Haar multiresolution approximation. *International Journal of Adaptive Control and Signal Processing*, 13:691–717, 1999.

[51] Z. Hasiewicz. Non–parametric estimation of non–liearity in a cascade time–series system by multiscale approximations. *Signal Processing*, 81:791–807, 2001.

[52] Z. Hasiewicz, G. Mzyk. Combined parametric–nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 49(8):1370–1375, 2004.

[53] Z. Hasiewicz, M. Pawlak, P. Śliwiński. Non–parametric identification of non–linearities in block–oriented complex systems by orthogonal wavelets with compact support. *IEEE Transactions on Circuits and Systems I: Fundametal Theory and Applications*, 52(2):427–442, 2005.

[54] Z. Hasiewicz, P. Śliwiński. Identification of non–linear characteristics of a class of block–oriented non–linear systems via Daubechies wavelet–based models. *International Journal of Systems Science*, 33(14):1121–1144, 2002.

[55] Z. Hasiewicz, P. Śliwiński. *Falki ortogonalne o zwartym nośniku.* Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2005.

[56] M. C. Hughes, D. T. Westwick. Identification of IIR Wiener systems with spline nonlinearities that have variable knots. *IEEE Transactions on Circuits and Systems I: Fundametal Theory and Applications*, 52(2):427–442, 2005.

[57] I. W. Hunter, M. J. Korenberg. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biological Cybernetics*, 55:135–144, 1986.

[58] H. Ichimura. Semiparametric least squares (SLS) and weighted SLS estimation of single–index models. *Journal of Econometrics*, 58:71–120, 1993.

[59] O. L. R. Jacobs. Recursive estimation for non–linear Wiener systems by on–line implementation of Bayes' rule. *Transactions of the Institute of Measurement and Control*, 7:245–250, 1985.

[60] J. Jakubowski, R. Sztencel. *Wstęp do teorii prawdopodobieństwa.* SCRIPT, Warszawa, 2004.

[61] R. I. Jennrich. Asymptotic properties of non–linear least squares estimators. *IEEE Transactions on Automatic Control*, 50(10):1617–1622, 2005.

[62] B. Jeong, K. Yoo, H. Rhee. Nonlinear model predictive control using a Wiener model of a continuous methyl methacrylate polymerization reactor. *Industrial and Engineering Chemistry Research*, 40:5968–5977, 2001.

[63] A. Kalafatis, N. Arifin, L. Wang, W. R. Cluett. A new approach to the identification of pH processes based on the Wiener model. *Chemical Engineering Science*, 50(23):3693–3701, 1995.

[64] H. Kang, Y. Cho, D. Youn. On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter. *IEEE Transactions on Communications*, 47(4):522–526, 1999.

[65] A. Y. Kibangou, G. Favier. Wiener–Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE Signal Processing Letters*, 13(6):381–384, 2006.

[66] W. Kołodziej. *Wybrane rozdziały analizy matematycznej.* PWN, Warszawa, 1970.

[67] M. Korenberg., I. Hunter. Two methods for identifying Wiener cascades having non-invertible static nonlinearities. *Annals of Biomedical Engineering*, 27(6):793–8004, 1999.

[68] M. J. Korenberg. Parallel cascade identification and kernel estimation for nonlinear systems. *Annals of Biomedical Engineering*, 19:429–455, 1991.

[69] M. J. Korenberg, I. W. Hunter. The identification of nonlinear biological systems: LNL cascade models. *Biological Cybernetics*, 55:125–134, 1986.

[70] J. Koronacki, J. Mielniczuk. *Statystyka dla studentów kierunków technicznych i przyrodniczych.* WNT, Warszawa, 2006.

[71] A. Krzyżak. Identification of non–linear systems by recursive kernel regression estimates. *International Journal of Systems Science*, 24(3):577–598, 1993.

[72] P. Kulczycki. *Estymatory jądrowe w analizie systemowej.* WNT, Warszawa, 2005.

[73] S. L. Lacy, D. S. Bernstein. Identification of FIR Wiener systems with unknown, non-invertible, polynomial non-linearities. *International Journal of Control*, 76(15):1500–1507, 2003.

[74] J. C. Lagarials, J. A. Reeds, M. H. Wright, P. E. Wright. Convergence properties of the Nelder–Mead simplex method in low dimesions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

[75] F. Leja. *Rachunek różniczkowy i całkowy.* Państwowe Wydawnictwo Naukowe, 1965.

[76] L. Ljung. *System Identification Theory for the User.* Prentice–Hall, Englewood Cliffs, NJ, 1987.

[77] Y. P. Mack, B. W. Silverman. Weak and strong uniform consistency of kernel regression estimates. *Probability Theory and Related Fields*, 61(3):405–415, 1982.

[78] K. Mańczak, M. Nahorski. *Komputerowa identyfikacja obiektów dynamicznych.* PWN, Warszawa, 1981.

[79] V. Z. Marmarelis, K. I. Naka. White–noise analysis of a neuron chain: an application of the Wiener theory. *Science*, 175:1276–1278, 1978.

[80] G. Mzyk. *Parametryczna identyfikacja systemów o złożonej strukturze, (rozprawa doktorska).* Instytut Cybernetyki Technicznej Politechniki Wrocławskiej, Wrocław, 2002.

[81] É. A. Nadaraya. On non–parametric estimates of density functions and regression curves. *Theory of Probability and its Applications*, 10:186–190, 1965.

[82] K. S. Narendra, P. G. Gallman. An iterative method for the identification on non–linear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, 11:546–550, 1966.

[83] O. Nelles. *Nonlinear System Identification.* Springer–Verlag, New York, 2001.

[84] A. E. Nordsjö, L. H. Zetterberg. Identification of certain time-varying nonlinear Wiener and Hammerstein systems. *IEEE Transactions on Signal Processing*, 49(3):577–592, 2001.

[85] S. J. Norquay, A. Palazoglu, J. A. Romagnoli. Model predictive control based on Wiener models. *Chemical Engineering Science*, 53(1):75–84, 1998.

[86] S. Ostasiewicz, Z. Rusnak, U. Siedlecka. *Statystyka, elementy teorii i zadania.* Wyd. Akademii Ekonomicznej we Wrocławiu, Wrocław, 1999.

[87] G. A. Pajunen. Recursive identification of Wiener type nonlinear systems. *Proceedings of the 1985 American Control Conference*, 3:1365–1370, 1985.

[88] A. Papoulis. *Prawdopodobieństwo, zmienne losowe i procesy stochastyczne.* WNT, Warszawa, 1972.

[89] J. Park, I. W. Sandberg. Criteria for the approximation of nonlinear systems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(8):673–676, 1992.

[90] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33:1065–1076, 1962.

[91] M. Pawlak. On the series expansion approach to the identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 36:763–767, 1991.

[92] M. Pawlak, Z. Hasiewicz. Nonlinear system identification by the Haar multiresolution analysis. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45:945–961, 1998.

[93] M. Pawlak, Z. Hasiewicz, P. Wachel. On non–parametric identification of Wiener systems. *IEEE Transactions on Signal Processing*, 55(2):482–492, 2007.

[94] R. K. Pearson, M. Pottmann. Gray–box identification of block–oriented nonlinear models. *Journal of Process Control*, 10:301–315, 2000.

[95] J. L. Powell, J. H. Stock, T. M. Stoker. Semiparametric estimation of index coefficients. *Econometrica*, 57(6):1403–1430, 1989.

[96] R. Pupeikis. Identification of piecewise affine Wiener systems using data partition. *Department of Electrical Engineering, Link oping University, Sweden*, Report no.: LiTH-ISY-R-2523, 2003.

[97] R. Pupeikis. On the identification of wiener systems having saturation–like functions with positive slopes. *Informatica*, 16(1):131–144, 2005.

[98] R. Raich, G. Tong Zhou, M. Viberg. Subspace based approaches for Wiener system identification. *IEEE Transactions on Automatic Control*, 50(10):1629–1634, 2005.

[99] W. Rudin. *Podstawy analizy matematycznej.* WNT, Warszawa, 2000.

[100] W. Rudin. *Analiza funkcjonalna.* WNT, Warszawa, 2002.

[101] L. Rutkowski. Sequential estimates of probability densities by orthogonal series and their applications in pattern classification. *IEEE Transactions on System, Man, and Cybernetics*, SMC-10(12):918–920, 1980.

[102] L. Rutkowski. On-line identification of time-varying systems by nonparametric techniques. *IEEE Transactions on Automatic Control*, AC-27:228–230, 1982.

[103] L. Rutkowski. On system identification by nonparametric function fitting. *IEEE Transactions on Automatic Control*, AC-27:225–227, 1982.

[104] L. Rutkowski. On nonparametric identification with prediction of time-varying systems. *IEEE Transactions on Automatic Control*, AC-29:58–60, 1984.

[105] L. Rutkowski. Generalized regression neural networks in time–varying environment. *IEEE Transactions on Neural Networks*, 15(3):576–596, 2004.

[106] J. Van Ryzin. On strong consistency of density estimates. *The Annals of Mathematical Statistics*, 40(5):1765–1772, 1969.

[107] I. W. Sandberg. Approximation theorems for discrete–time systems. *IEEE Transactions on Circuits and Systems*, 38(5):564–566, 1991.

[108] I. W. Sandberg. Structure theorems for nonlinear systems. *Multidimensional Systems and Signal Processing*, 2(3):267–286, 1991.

[109] I. W. Sandberg. Approximations for nonlinear functionals. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(1):65–67, 1992.

[110] D. Sbárbaro, T. A. Johansen. Multiple local Laguerre models for modelling nonlinear dynamic systems of the Wiener class. *IEE Proceedings – Control Theory and Applications*, 144(5):375–380, 1997.

[111] J. Schoukens, J. G. Nemeth, P. Crama, Y. Rolain, R. Pintelon. Fast approximate identification of nonlinear systems. *Automatica*, 39:1267–1274, 2003.

[112] E. F. Schuster. Estimation of a probability density function and its derivatives. *The Annals of Mathematical Statistics*, 40(4):1187–1195, 1969.

[113] E. F. Schuster. Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *The Annals of Mathematical Statistics*, 43(1):84–88, 1972.

[114] E. Sekko, G. Thomas, A. Boukrouche. A deconvolution technique using optimal Wiener filtering and regularization. *Signal Processing*, 72:23–32, 1999.

[115] J. Shi, H.H. Sun. Nonlinear system identification via parallel cascaded structure. *Proceedings of the Twelfth Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 12(4):1897–1898, 1990.

[116] B. W. Silverman. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, 6(1):177–184, 1978.

[117] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P. Glorennec, H. Hjalmarsson, A. Juditsky. Nonlinear black–box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

[118] T. Söderström, P. Stoica. *System Identification.* Upper Saddle River NJ:Prentice Hall, 1989.

[119] J. Solé-Casals, C. Jutten, D. T. Pham. Fast approximation of nonlinearities for improving inversion algorithms of PNL mixtures and Wiener systems. *Signal Processing*, 85:1780–1786, 2005.

[120] C. J. Stone. Optimal rates of convergence for nonparametric estimators. *The Annals of Statistics*, 8(6):1348–1360, 1980.

[121] A. Taleb, J. Sole, C. Jutten. Quasi–nonparametric blind inversion of Wiener systems. *IEEE Transactions on Signal Processing*, 49(5):917–924, 2001.

[122] C. W. Therrien. The Lee–Wiener legacy, a history of the statistical theory of communication. *IEEE Signal Processing Magazine*, 19(6):33–34, 2002.

[123] A.W. van der Vaart. *Asymptotic Statistics.* Cambridge University Press, Cambridge, 1998.

[124] J. Vörös. Parameter identification of Wiener systems with discontinuous nonlinearities. *Systems and Control Letters*, 44(5):363–372, 2001.

[125] P. Wachel. Falkowe algorytmy nieparametrycznej estymacji nieliniowości w systemie Wienera. *Praca magisterska, specjalność: Komputerowe Sieci Sterowania, kierunek: Automatyka i Robotyka, Wydział Elektroniki, Politechnika Wrocławska*, 2004.

[126] P. Wachel. Non–parametric recovery of non–linearity in Wiener systems by wavelet–based kernel algorithm. *10th IEEE International Conference on Methods and Models in Automation and Robotics*, strony 491–496, 2004.

[127] P. Wachel. A mixed parametric–non-parametric method of Wiener system identification. *11th IEEE International Conference on Methods and Models in Automation and Robotics*, 2005.

[128] P. Wachel. Parametric–non-parametric identification of Wiener systems with non-inverse nonlinearities. *Raport Instytutu Informatyki, Automatyki i Robotyki Politechniki Wrocławskiej*, Preprints No.1/2007, 2007.

[129] P. Wachel. Two–stage identification of Wiener systems with finite memory and noninverse nonlinearities. *13th IEEE International Conference on Methods and Models in Automation and Robotics*, 2007.

[130] P. Wachel. Noninverse identification of Wiener systems. *Referat przyjęty do wygłoszenia na Krajowej Konferencji Automatyki, Szczyrk*, 2008.

[131] F. H. Walters, L. R. Parker, S. L. Morgan, S. N. Deming. *Sequential Simplex Optimization.* CRC Press, Boca Raton, FL, 1991.

[132] M. P. Wand, M. C. Jones. *Kernel Smoothing.* Chapman and Hall, 1995.

[133] G. S. Watson. Smooth regression analysis. *Sankhya, Ser. A*, 26:356–359, 1964.

[134] D. T. Westwick, R. E. Kearney. A new algorithm for the identification of multiple input Wiener systems. *Biological Cybernetics*, 68:75–85, 1992.

[135] D. T. Westwick, M. Verhaegen. Identifying MIMO Wiener systems using subspace model identification methods. *Signal Processing*, 52(2):235–258, 1996.

[136] N. Wiener. *Nonlinear problems in random theory.* Wiley, New York, 1958.

[137] T. Wigren. Recursive prediction error identification using the nonlinear Wiener model. *Automatica*, 29(4):1011–1025, 1993.

[138] T. Wigren. Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model. *IEEE Transactions on Automatic Control*, 39(11):2191–2206, 1994.

[139] C. Wu. Asymptotic theory of nonlinear least squares estimation. *The Annals of Statistics*, 9:501–513, 1981.

[140] M. Xu, G. Chen, Y. Tian. Identifying chaotic systems using Wiener and Hammerstein cascade models. *Mathematical and Computer Modelling*, 33:483–493, 2001.

[141] Y. Zhu. Distillation column identification for control using Wiener model. *Proceedings of the 1999 American Control Conference*, 5:3462–3466, 1999.

Mgr inż. Paweł Wachel

Instytut Informatyki, Automatyki i Robotyki
Politechniki Wrocławskiej
ul. Janiszewskiego 11/17
50-372 Wrocław

Niniejszy raport otrzymują:

1. OINT	1 egz.
2. Biblioteka Główna PWr	1 egz.
3. Z-ca Dyrektora Instytutu	1 egz.
4. Promotor	1 egz.
5. Recenzenci	2 egz.
6. Autor	2 egz.
	Razem: 8 egz.

Raport wpłynął do Redakcji I-6
26.03.2008 r.