



Politechnika Wrocławska



Wydział Informatyki i Telekomunikacji

Katedra Automatyki, Mechatroniki i Systemów Sterowania

**Nieparametryczne metody identyfikacji
w optymalizacji procesów produkcyjnych**

Nonparametric identification in optimization
of manufacturing processes

Paweł Mielcarek

(rozprawa doktorska)

Promotor:

dr hab. inż. Grzegorz Mzyk, prof. uczelni

Słowa kluczowe:

- identyfikacja systemów,
- modelowanie nieparametryczne,
- optymalizacja procesów produkcyjnych,
- dynamiczne systemy nieliniowe,
- systemy Hammersteina i Wienera.

Wrocław, czerwiec 2023

Paweł Mielcarek

Nieparametryczne metody identyfikacji w optymalizacji procesów produkcyjnych

Nonparametric identification in optimization of manufacturing processes

– Rozprawa doktorska –

Czerwiec 2023

Promotor:

dr hab. inż. Grzegorz Mzyk, prof. uczelni



Politechnika Wroclawska

Spis treści

Streszczenie	IX
Abstract	XI

I Wstęp

1	Spis symboli	3
2	Wprowadzenie	5
2.1	Historia identyfikacji systemów	5
2.2	Teza i cele pracy	7
2.3	Potrzeba stosowania modeli w optymalizacji procesów produkcyjnych	8
2.4	Problemy ze stosowalnością metod nieparametrycznych w praktyce	8
2.4.1	Restrykcyjne założenia o sygnale wejściowym	9
2.4.2	Wielowymiarowość problemu (ang. <i>curse of dimensionality</i>)	10
2.4.3	Brak wiedzy wstępnej o strukturze modelu i poszczególnych bloków	11
2.4.4	Aspekty obliczeniowe	12

II Część teoretyczna

3	Wieloetapowa identyfikacja systemów Hammersteina-Wienera	15
3.1	Model N-L-N	15
3.2	Założenia algorytmu	16
3.3	Algorytm	17
3.3.1	Krok 1: identyfikacja odpowiedzi impulsowej	18
3.3.2	Krok 2: estymacja wyjściowej charakterystyki nieliniowej	19
3.3.3	Krok 3: estymacja sygnału wewnętrznego x_k za pomocą filtra odwrotnego	20
3.3.4	Krok 4: identyfikacja parametrów wejściowej charakterystyki nieliniowej	22
3.4	Własności algorytmu	23
3.5	Przykład numeryczny	23
4	Identyfikacja nieliniowych systemów dynamicznych o strukturze ogólnej metodą jądrową	27
4.1	Opis problemu	27
4.2	Opis i klasa rozpatrywanych systemów	28
4.3	Wielowymiarowa regresja jądrowa	28
4.4	Redukcja wymiaru problemu	29
4.4.1	Kwantyzowane wejście	30
4.4.2	Sygnał okresowy	30
4.5	Własności algorytmu	31
4.6	Przykład numeryczny	32
4.7	Propozycje potencjalnych zastosowań	34
5	Rekurencyjna identyfikacja wielopoziomowa, z dekompozycją i predykcją interakcji	37
5.1	Opis problemu	38
5.2	Algorytm off-line bez rekurencji	40
5.3	Rekurencyjny algorytm on-line	41
5.4	Uogólnienie algorytmu do systemów klasy ARMAX	44
5.5	Badania symulacyjne	45
5.5.1	Algorytm off-line	46
5.5.2	Algorytm on-line	46

6	Nieparametryczna identyfikacja systemów w zastosowaniach praktycznych	51
7	Modelowanie szkieł chalkogenidkowych w kalorymetrze różnicowym	53
7.1	Opis problemu	53
7.2	Model Hammersteina z elementem liniowym typu ARMA przy użyciu metody predykcji interakcji	54
7.3	Model nieparametryczny	55
8	Modelowanie w optymalizacji pracy obrabiarek sterowanych numerycznie	61
8.1	Układ pomiarowy oraz akwizycja danych	65
8.2	Wielowymiarowa regresja jądrowa w modelowaniu szeregów czasowych	67
8.3	Przykład detekcji niezgodności procesu technologicznego	68
8.4	Porównanie z innymi metodami uczenia maszynowego	73
8.4.1	Drzewo regresji	73
8.4.2	Sztuczna sieć neuronowa	75
8.4.3	Podsumowanie wyników	77
9	Podsumowanie	79
9.1	Rezultaty badań i ich następstwa	79
9.2	Problemy otwarte	80

IV Dodatki

A	Dowody twierdzeń i wyprowadzenia techniczne	83
A.1	Dowód Twierdzenia 3.4.1	83
A.2	Dowód Twierdzenia 3.4.2	83
A.3	Dowód Twierdzenia 4.5.1	85
A.4	Uwaga do Twierdzenia 4.2	86
A.5	Przykład do Twierdzenia 4.2	86
A.6	Linearyzacja lokalna w algorytmie IPM (Rozdział 5)	87
B	Dodatki, twierdzenia, lematy, definicje	89
B.1	Definicje typów zbieżności	89
B.2	Rozkład SVD macierzy	90
B.3	Błąd średniokwadratowy MSE	91

VIII Spis treści

B.4 Mocne prawo wielkich liczb Kołmogorowa	91
B.5 Dwu-etapowy algorytm identyfikacji z agregacją parametrów ..	91
B.6 Algorytm optymalizacji wypukłej	93

V Bibliografia

Spis rysunków	101
Spis tabel	104

Streszczenie

Rozprawa składa się z dwóch części – teoretycznej i praktycznej. Pierwsza z nich poświęcona jest uogólnieniu tradycyjnych metod nieparametrycznej identyfikacji systemów nieliniowych (rozwijanych w zespole od dziesięcioleci głównie pod kątem teoretycznym) w celu umożliwienia ich zastosowania w wybranych zagadnieniach praktycznych. W szczególności, opracowano wieloetapowy algorytm identyfikacji systemu będącego połączeniem kaskadowym systemów Hammersteina i Wienera. Tak ogólne podejście jest szczególnie ważne przy braku wiedzy wstępnej o strukturze modelu i opisie poszczególnych jego bloków. Kolejnym aspektem pracy jest propozycja specjalnych sygnałów wejściowych w eksperymentach identyfikacyjnych. Stawiane zadania prowadziły z reguły do metod pracujących w przestrzeniach o dużej wymiarowości, co przekładało się na powolną zbieżność asymptotyczną estymatorów, często wykluczającą możliwość zastosowania danego algorytmu w konkretnym problemie praktycznym. Zaproponowane rozwiązanie polega na redukcji wymiarowości, i w konsekwencji poprawie szybkości zbieżności estymatorów, poprzez użycie pobudzeń kwantowanych, lub okresowych. Podjęto także próbę zastosowania techniki dekompozycji i koordynacji w celu podziału zadania identyfikacji systemu złożonego na zadania mniejsze, stosując metodę predykcji interakcji. Naprzemienna identyfikacja bloków składowych systemu Hammersteina, była już rozpatrywana w literaturze, jednak w zaproponowanym podejściu dopuszcza się opis nieparametryczny bloku nieliniowego. Opracowano także wersje rekurencyjne zaproponowanych algorytmów.

Część praktyczna rozprawy przedstawia rezultaty zastosowania nowych podejść, po odpowiednich modyfikacjach, w dwóch zadaniach praktycznych: do

X Spis treści

modelowania procesu nagrzewania szkiełek chalcogenidkowych w kalorymetrze różnicowym (w celu zbadania procesu starzenia materiału) oraz do wykrywania sytuacji nietypowych w pracy obrabiarek sterowanych numerycznie (CNC) na podstawie pomiaru chwilowego natężenia prądu zasilającego.

Abstract

The presented doctoral dissertation was divided into two sections – theoretical and applicational. The theoretical part aimed to generalise classical identification methods of nonlinear dynamic systems to enable application in selected real-life problems. The first presented algorithm is the multi-level identification of Hammerstein-Wiener systems. This complex approach to block-oriented methods supports the lack of knowledge related to the mathematical description of the system, due to the cascade connection of two nonlinear blocks and one dynamic block. The second proposed solution is focused on using the active experiment approach to identification. Many of the practical problems faced a challenge in modelling the system, because of large dimension spaces which led to slow asymptotical convergence. The proposed solution is based on dimensionality reduction using quantized or periodical excitations. The last presented solution uses the idea of decomposition and coordination, splitting the identification of the complex system into smaller and less complex tasks, with the use of the interaction prediction method. Alternate identification of Hammerstein system blocks was the subject of many publications, recent contribution lies in the nonparametric approach to nonlinear characteristics identification, and generalisation of the method to the ARMAX systems. The recursive versions of presented algorithms were developed as well.

The applicational part is focusing on the application of such a new approach in two real-life problems: an examination of the thermal ageing process of chalcogenide glasses in a differential scanning calorimeter and monitoring the technological process in CNC machines to detect any discrepancy, based on the current load measurement.

Wstep

Spis symboli

k - numer pomiaru, dyskretna chwila czasu,
 N - liczba pomiarów,
 u_k - wejście systemu,
 $u_k^{(1)}$ - wejście systemu w pierwszym eksperymencie,
 $u_k^{(2)}$ - wejście systemu w drugim eksperymencie,
 y_k - wyjście systemu,
 v_k, w_k, x_k - niedostępne dla pomiarów sygnały wewnętrzne,
 z_k - zakłócenia,
 a, b, θ - wektory parametrów,
 Q - kryterium jakości estymatora,
 f, g - funkcje bazowe,
 ϵ - proces deterministyczny,
 γ, λ - odpowiedź impulsowa liniowego podsystemu dynamicznego,
 $\mu()$ - wejściowa charakterystyka nieliniowa,
 $\eta()$ - wyjściowa charakterystyka nieliniowa,
 $\hat{R}()$ - estymator regresji,
 $K()$ - funkcja jądra,
 h - szerokość jądra,
 MSE, δ - błąd średniokwadratowy,
 s, S, s^* - wymiar modelu,
 i, j - pomocnicze zmienne indeksujące,
 ma - średnia krocząca,
 w - okno średniej kroczącej,
 $E\{X_n\}$ - wartość oczekiwana zmiennej losowej X_n .

Wprowadzenie

2.1 Historia identyfikacji systemów

Identyfikacja i modelowanie systemów mają długą historię, ich początków możemy upatrywać w XIX wieku. W 1801 roku Friedrich Gauss, bazując na obserwacjach Josepha Piazziego, znalazł aktualne położenie planetoidy Ceres, która zniknęła w promieniach słońca. Wykorzystując jeszcze wtedy nieopublikowaną metodę najmniejszych kwadratów i prawa Keplera, Gauss wyznaczył estymatory parametrów trajektorii ruchu ciała niebieskiego i zaproponował aktualną pozycję, która okazała się poprawna. Ceres został odnaleziony, pozwalając na kontynuowanie pracy astronomom, a Gauss zyskał na popularności w środowiskach akademickich. W roku 1809 opublikował metodę na której wówczas bazował [13].

Analizując podejście Gaussa można zauważyć, że skorzystał on z dwóch rodzajów informacji: pomiarów (Piazzi) oraz wiedzy eksperckiej (Kepler). W zasadzie można uznać to za pewną ogólną definicję szkoły identyfikacji którą nazywamy parametryczną. Liczba pomiarów musi być odpowiednia do ilości parametrów, które występują w modelu, a liczba nieznanych parametrów jest znana i oczywiście skończona. Wiedza ekspercka dotyczy struktury modelu, czyli opisu matematycznego danego zjawiska. Dzięki takiemu podejściu stworzenie modelu, nawet przy niewielkiej liczbie pomiarów jest zadaniem stosunkowo łatwym. Niestety, jak się okazuje, poleganie na znanym opisie niesie ryzyko błędnej aproksymacji. Dzieje się tak dlatego, że rozwiązując rzeczywiste problemy, warunki mogą być odmienne od tych które posłużyły do konstrukcji modelu ogólnego lub stan aktualnej wiedzy o zjawisku nie jest odpowiednio szczegółowy i nazbyt uproszczony. Oczywiście zakładając, że w ogóle taka

wiedza jest dostępna, co oznacza, że jeżeli zjawisko jest nowe, przed jego identyfikacją musi nastąpić proces poznania jego natury, co znacząco ogranicza możliwość zastosowania takich algorytmów w nowych badaniach.

Dalszy rozwój matematyki na początku XX wieku, a w szczególności badania na temat systemów nieliniowych pokazały, że dotychczasowe podejście do opisu świata jest niewystarczające aby zrozumieć jego skomplikowaną naturę. W środowisku statystyków i matematyków zajmujących się analizą funkcjonalną zaczęły się pojawiać inne formy opisu matematycznego [1], które rezygnowały z przywiązania do parametrów [44] [14]. Był to początek metod nieparametrycznych, takich jak rozwinięcie w szereg Volterry [24], estymatory ortogonalne [8], estymatory jądrowe (w tym estymator Nadaraya'i-Watsona) czy zastosowanie szeregu Wienera.

W latach osiemdziesiątych XX wieku zespół polskich naukowców zainteresował się zastosowaniem tych matematycznych konstruktów do analizy systemów dynamicznych [22] [16] [17] [27]. Profesorowie Włodzimierz Greblicki, Mirosław Pawlak, Adam Krzyżak oraz ich następcy zastosowali metody nieparametryczne do identyfikacji systemów dynamicznych, skupiając się na systemach blokowych takich jak systemy Hammersteina [5], Wienera [15] [46] oraz ich kombinacjach i modyfikacjach (np. elementy autoregresyjne).

Zastosowanie metod nieparametrycznych w identyfikacji systemów nadało nowy kierunek badaniom, jednak wciąż pozostawało wiele przeszkód w ich implementacji w świecie inżynierskim. Wszelkie dowody opierają się na własnościach asymptotycznych, czyli uwzględniających bardzo duże zbiory danych wejściowych oraz na restrykcyjnych założeniach takich jak np. nieskorelowanie i odpowiedni rozkład sygnału wejściowego.

Stosunkowo świeżym podejściem, zainicjowanym na początku XXI wieku jest łączenie metod klasycznych - parametrycznych, z nowoczesnymi - nieparametrycznymi [28] [30]. Podejście to rozwiązuje wiele problemów znanych z czysto nieparametrycznych rozważań.

Zagadnienia omawiane w pracy są pewnym kolejnym uogólnieniem zaprezentowanych wcześniej metod. Adresując głównie problemy z zastosowaniem inżynierskim metod nieparametrycznych, rozprawa ta ma na celu zaproponowanie zestawu narzędzi dzięki którym możliwa staje się identyfikacja systemów o ogólnej strukturze ze zredukowanymi założeniami wstępnymi. Dostarczając tym samym formalizmów udowadniających skuteczność opisywanych algorytmów.

mów jak i przykładów ich zastosowania.

2.2 Teza i cele pracy

Niniejsza rozprawa ma na celu zaproponowanie rozwiązań problemu stosowalności teoretycznych metod identyfikacji systemów w problemach praktycznych [45] [55], z naciskiem na środowiska przemysłowe (Rozdział 8) i badawczo-rozwojowe (Rozdział 6) [34] [35]. Wiedząc, że w skomplikowanych nieliniowych systemach dynamicznych ilość potencjalnych problemów zdecydowanie przewyższa możliwości tej pracy, zdecydowano skupić się na najistotniejszych aspektach. Problemy te zebrano w podrozdziale 2.4 i tam też zaadresowano autorskie rozwiązania, które są omawiane bardziej szczegółowo w kolejnych Rozdziałach pracy, gdzie poparto je analizą teoretyczną. W celu utrzymania omawianego, uniwersalnego charakteru prezentowanych rozważań, ich wspólnym elementem jest nieparametryczna nieliniowość. Z takim założeniem wykorzystano w pracy bogatą historię badań dotyczących identyfikacji nieparametrycznej nieliniowych systemów dynamicznych, której niewielki wycinek omówiono w poprzednim punkcie (2.1).

Na uwagę zasługuje też pewien efekt uboczny niniejszych rozważań, a mianowicie wypadkowa nieparametryczności i ogólnego opisu systemu. Dzięki takiemu podejściu, znacząco skrócony zostaje czas jaki trzeba poświęcić na analizę systemu, przed modelowaniem. Ma to szczególne znaczenie w środowisku produkcyjnym, gdzie mnogość różnych procesów w ramach jednego produktu lub marszruty może być znaczna, a czas wdrożeń musi być minimalizowany, z uwagi na szybkie tempo zmian produktów oraz ciągle przyspieszanie procesów produkcyjnych.

Z uwagi na restrykcyjne założenia i wolną zbieżność algorytmów, bardzo dojrzałe i wyrafinowane metody identyfikacji nieparametrycznej, są uciążliwe lub niemożliwe do zastosowania inżynierskiego. Dlatego też podjęto próbę uogólnienia owych metod teoretycznych (Rozdziały 3, 4 i 5).

Następnie, mając już paletę rozwiązań potencjalnie nadających się do zagadnień praktycznych [2] [14], zdecydowano się je zastosować (Rozdziały 7 i 8) i porównać z istniejącymi i popularnymi metodami.

2.3 Potrzeba stosowania modeli w optymalizacji procesów produkcyjnych

Zaawansowana analityka już od wielu lat jest domeną systemów produkcyjnych. Wiele z nich z uwagi na swój charakter powstawało już na skalę przemysłową z pewnym dziedzictwem matematycznym, chociażby w postaci równań opisujących reakcje chemiczne czy też szybkości dostaw mediów do odbiorców. Dla produkcji dyskretniej lub w partiach, w takich branżach jak motoryzacja, lotnictwo czy dobra konsumenckie, rewolucją było wprowadzenie metod statystycznych do zarządzania produkcją, takich jak np. *Lean Six Sigma*. Taka analiza danych, szczególnie pod kątem jakości produkcji przyczyniła się do znaczącej poprawy wydajności oraz zmniejszenia odpadu. W czasach obecnych gdzie cyfryzacja produkcji pozwala na zbieranie ogromnej ilości danych z procesu i oprzyrządowania oraz umożliwia zastosowanie oprogramowania w celu przeprowadzenia skomplikowanej analityki, jakość stosowanych algorytmów ma szczególne znaczenie. Taka analiza danych na produkcji i zarządzając łańcuchem dostaw, nie jest dodatkiem a wymogiem, aby produkować redukując koszty, zwiększając zysk i w efekcie pozostawać konkurencyjnymi na rynku.

Obecnie możemy mówić o różnych wymiarach tego typu analiz. Np. diagnostyka problemów jakościowych, prognozowanie awarii, prognozowanie wyników produkcyjnych, proponowanie nastaw masz czy też autonomiczne podejmowanie przez system decyzji wpływając na proces tak aby zoptymalizować wydajność. Nie należy również zapinać o klasycznym zastosowaniu omawianych metod - sterowaniu urządzeniami wykonawczymi [4] [50] [54], tak aby osiągnąć jak największą precyzję procesu, bądź zminimalizować straty.

Aby jednak to wszystko zrealizować, należy zastanowić się nad stosowalnością metod teoretycznych, opisywanych w książkach i publikacjach naukowych w odniesieniu do problemów jakie występują przy pracy z systemem rzeczywistym, a nie symulowanym lub testowanym w warunkach laboratoryjnych. Właśnie o tym aspekcie niniejszych rozważań traktuje podrozdział - 2.4.

2.4 Problemy ze stosowalnością metod nieparametrycznych w praktyce

Identyfikacja parametryczna, historycznie odegrała wielką rolę w poznawaniu i opisie wszelkich zjawisk i systemów. Jej największą wadą jest założenie, że zna-

my pewien abstrakcyjny opis systemu, co w dobrze poznanych zjawiskach - lub odpowiednio uproszczonych ich modelach może nie wydawać się problemem. Jednakże w bardziej skomplikowanych przypadkach, szczególnie z uwzględnieniem wszelkich nieliniowości i dynamiki istnieje wysokie prawdopodobieństwo błędnej parametryzacji. Problemem bliźniaczym może być również potrzeba identyfikacji systemu który nie jest odpowiednio dokładnie znany, lub jest nowy w sensie znajomości jego matematycznego opisu. Dokładnie takie ograniczenia napotyka zastosowanie metod parametrycznych w praktyce.

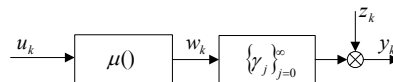
Jakie zatem problemy utrudniają zastosowanie metod literaturowych w zagadnieniach praktycznych? Kolejne podrozdziały traktują o następujących problemach których próbę rozwiązania podjęto w niniejszej pracy:

- restrykcyjne założenia o sygnale wejściowym - 2.4.1,
- wielowymiarowość problemu (ang. *curse of dimensionality*) - 2.4.2,
- brak wiedzy wstępnej o strukturze modelu i poszczególnych bloków - 2.4.3,
- aspekty obliczeniowe - 2.4.4.

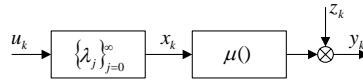
2.4.1 Restrykcyjne założenia o sygnale wejściowym

Dla popularnych w literaturze systemów o strukturze blokowej jak systemy Hammersteina (rys. 2.1), Wienera (rys. 2.2) lub Wienera-Hammersteina przyjmuje się restrykcyjne założenia o sygnale wejściowym. Jest to związane z faktem, że sygnały wewnętrzne są niedostępne dla pomiarów, a ich rozkład jest związany z rozkładem na wejściu. Dla systemów Hammersteina przy pobudzeniu białym szumem, współczynniki odpowiedzi impulsowej γ_j mogą być wyznaczone za pomocą funkcji korelacji wzajemnej wejścia-wyjścia, zaś nieliniowość $\mu(u_k)$ może być wyznaczona za pomocą metody najmniejszych kwadratów lub innego rodzaju regresji.

W systemach Wienera przy zastosowaniu na wejściu sygnału o rozkładzie normalnym, również sygnał x_k ma rozkład normalny, co prowadzi do problemu rozwiązanego wielokrotnie w literaturze i korelacja wzajemna wejścia-wyjścia jest proporcjonalna do odpowiedzi impulsowej systemu (tw. Bussganga).



Rysunek 2.1. System Hammersteina.



Rysunek 2.2. System Wienera.

W przypadku gdy procedura zostanie uruchomiona przy niespełnionych powyższych założeniach identyfikacja jest problematyczna. W systemach Hammersteina dla skorelowanego wejścia, regresja wejścia-wyjścia traci prosty związek z charakterystyką nieliniową, a w systemach Wienera niegausowskie wejście uniemożliwia oparcie się o metodę korelacyjną [40], ponieważ nie funkcjonuje wtedy twierdzenie Bussganga.

Zaproponowane rozwiązania (Rozdziały 3, 4):

1. Metoda zmiennych instrumentalnych [12] [25] zamiast metody najmniejszych kwadratów (nieparametryczna generacja zmiennych instrumentalnych).
2. Lokalna nieparametryczna linearyzacja jądrowa oraz wielowymiarowa regresja jądrowa charakterystyki nieliniowej w parametrycznej identyfikacji elementów liniowych dynamicznych systemów Wienera, Hammersteina i Wienera-Hammersteina.
3. Identyfikacja nieparametryczna za pomocą wielowymiarowej regresji jądrowej.

2.4.2 Wielowymiarowość problemu (ang. *curse of dimensionality*)

Wymiarowość problemów identyfikacji systemów o strukturze blokowej jest związana z liczbą argumentów optymalizowanej funkcji. W algorytmach z agregacją argumentów dla systemu Hammersteina [3] powstają wektory wielowymiarowe, o wymiarze równym iloczynowi liczby parametrów modeli poszczególnych bloków składowych. Dla systemów typu Wienera na wymiar problemu wpływa bezpośrednio długość odpowiedzi impulsowej podsystemu liniowego, w sposób wykładniczy (radykalny) zmniejsza ona prawdopodobieństwo selekcji lokalnej.

Nawet gdy sumaryczna liczba argumentów systemu jest znacznie mniejsza od liczby pomiarów, algorytm, ze względu na swoją postać, może być źle zdefiniowany. Odwracana macierz będzie osobliwa, gdyż zagregowany wektor argu-

mentów może mimo to przekroczyć liczbę pomiarów (patrz dodatek B.5). Zaś z każdym wydłużeniem długości odpowiedzi impulsowej modelu, wymiar rozpatrywanego otoczenia rośnie co prowadzi do gwałtownego wzrostu wariancji i w efekcie powoduje radykalny spadek liczby wyselekcjonowanych, użytecznych pomiarów.

Zaproponowane rozwiązania (Rozdziały 4, 5):

1. Redukcja wymiaru za pomocą zastosowania aktywnego eksperymentu takiego jak wejście okresowe lub wejście skwantowane (systemy Wienera, NARMAX, system o strukturze ogólnej).
2. Dekompozycja zadania identyfikacji systemu złożonego (np. Hammersteina) z naprzemienną identyfikacją poszczególnych bloków i ze wzajemną koordynacją przy nieparametrycznym traktowaniu nieliniowości.

2.4.3 Brak wiedzy wstępnej o strukturze modelu i poszczególnych bloków

Z uwagi na fakt, że modele o strukturze blokowej narzucają zależność bloków między sobą oraz ich charakter istnieje ryzyko zastosowania niewłaściwej klasy/struktury do danego problemu [31]. To samo tyczy się błędu zbytniego uproszczenia modelu w stosunku do zachowania rzeczywistego systemu. Tego typu błędne założenia *a priori* mogą być przyczyną dużego błędu oraz niezastosowalności modelu na rzeczywistych danych.

Drugim aspektem jest kwestia liczby współczynników. Nie znając struktury, istnieje duże ryzyko "przeparametryzowania" (ang. *overparametrized model*) i wzrostu wariancji modelu z powodu dopasowywania się go do toru zakłóceń, zamiast do rzeczywistej charakterystyki.

Zaproponowane rozwiązania (Rozdziały 3, 4):

1. Podjęcie tematyki identyfikacji bardziej ogólnych struktur blokowo-zorientowanych takich jak systemy Hammersteina-Wienera (N-L-N) oraz Wienera-Hammersteina (L-N-L).
2. Opracowanie algorytmów semiparametrycznych, np. parametryzacja bloku liniowego przy nieparametrycznym modelu nieliniowym statycznym.
3. Zastosowanie nieparametrycznych metod wielowymiarowych. Co prowadzi do wcześniej omawianego problemu wielowymiarowości.

2.4.4 Aspekty obliczeniowe

Złożoność obliczeniowa omawianych grup algorytmów może być bardzo duża. Co samo w sobie stanowi problem co do stosowalności ich w praktyce. W literaturze możemy spotkać się z takimi problemami obliczeniowymi jak optymalizacja niewypukła czy nieodwracalna charakterystyka statycznej nieliniowości.

Zaproponowane rozwiązania (Rozdział 5):

1. Algorytmy rekurencyjne, znacząco zmniejszające złożoność obliczeniową.
2. Wyznaczanie rozwiązań minimalnych przy użyciu rozkładu SVD (Dodatek B.2).

Część teoretyczna

Wieloetapowa identyfikacja systemów Hammersteina-Wienera

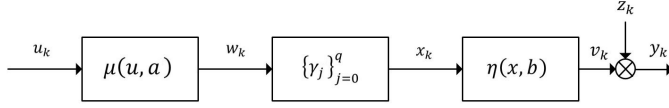
W niniejszym rozdziale zaprezentowany został algorytm identyfikacji systemu Hammersteina-Wienera, wspomagany metodami nieparametrycznymi przy częściowej wiedzy parametrycznej o poszczególnych blokach. W punkcie 3.1 pokazano strukturę systemu, założenia oraz wymaganą o systemie wiedzę *a priori*. Punkt 3.3 opisuje sposób działania algorytmu, którego celem jest w pierwszej kolejności następująca identyfikacja bloku liniowego za pomocą jądrowej metody najmniejszych kwadratów. W drugim kroku drugi (wyjściowy) blok nieliniowy jest identyfikowany przy użyciu aktywnego eksperymentu, binarnych sekwencji wejściowych które, z uwagi na swój charakter "zaślepiają" pierwszy blok nieliniowy. Następnie, cały problem, dzięki identyfikacji funkcji rozkładu prawdopodobieństwa zakłóceń, zostaje sprowadzony do systemu Hammersteina.

3.1 Model N-L-N

Rozważmy system (N-L-N) gdzie między dwoma statycznymi nieliniowymi blokami znajduje się jeden liniowy, dynamiczny, jak pokazano na rys. 3.1.

Taki system można zapisać równaniem,

$$y_k = \eta \left(\sum_{j=0}^q (\gamma_j \mu(u_{k-j})) \right) + z_k. \quad (3.1)$$



Rysunek 3.1. System Hammersteina–Wienera

Klasycznie u_k oraz y_k są pomiarami wejścia oraz wyjścia w chwili k , z_k jest zaś losowym zakłóceniem. Wektor $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_q)^T$ reprezentuje skończoną odpowiedź impulsową liniowego bloku dynamicznego, zaś funkcje $\eta(\cdot)$ oraz $\mu(\cdot)$ opisują charakterystyki nieliniowe na wejściu oraz na wyjściu.

3.2 Założenia algorytmu

1. Charakterystyki nieliniowe można opisać jako kombinacje liniowe funkcji bazowych f_i oraz g_j znanych *a priori*

$$\mu(u) = \mu(u, a) = a^T f(u), \quad (3.2)$$

$$a = (a_1, a_2, \dots, a_m)^T, a \in \mathbb{R}^m,$$

$$f(u) = (f_1(u), f_2(u), \dots, f_m(u))^T,$$

$$\eta(x) = \eta(x, b) = b^T g(x), \quad (3.3)$$

$$b = (b_1, b_2, \dots, b_n)^T, b \in \mathbb{R}^n,$$

$$g(x) = (g_1(x), g_2(x), \dots, g_n(x))^T.$$

Liczba parametrów, tj. wymiar wektorów a oraz b jest stała oraz znana. Ponadto założono, że obie statyczne charakterystyki nieliniowe spełniają warunek Lipschitza. Charakterystyki nieliniowe są podwójnie różniczkowalne i spełniają warunki $\mu'(u_0) \neq 0$ oraz $\eta'(x_0) \neq 0$. Wyjściowa charakterystyka nieliniowa jest ściśle monotoniczna, a w związku z tym odwracalna.

2. Odpowiedź impulsowa liniowego bloku dynamicznego jest opisana wektorem parametrów $\{\gamma_j\}_{j=0}^q$, gdzie q – jest długością odpowiedzi impulsowej, która jest skończona i znana,

$$\gamma = (\gamma_0, \gamma_1, \dots, \gamma_q), \gamma \in \mathbb{R}^{q+1}.$$

3. Zaproponowany algorytm bazuje na dwóch eksperymentach, gdzie wartości wejściowe $u_k^{(1)}$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie oraz takiej samej wartości oczekiwanej. Funkcja gęstości prawdopodobieństwa $\nu(\cdot)$ spełnia warunek Lipschitza oraz $\nu(0) > 0$. Pobudzenie $u_k^{(2)}$ jest zaś losowym sygnałem binarnym.
4. Sygnały u_k oraz z_k są niezależne oraz o stałych wariancjach $\sigma_u^2 < \infty$ i $\sigma_z^2 < \infty$. Ponadto wartość oczekiwana zakłóceń jest równa zero $E[z_k] = 0$, a nieliniowa charakterystyka wejściowa $\mu(u_k^{(2)})$ spełnia warunki $\mu(0) = 0$ oraz $\mu(1) \neq 0$. Ponieważ wzmocnienie w stanie ustalonym poszczególnych bloków jest niemożliwe do wyznaczenia założymy (bez utraty ogólności), że $G = \sum_{j=0}^q \gamma_j = 1$ oraz $\mu(1) = 1$.

Z założeń 1 oraz 2, wynika że system jest opisany za pomocą $n + m + q + 1$ parametrów.

3.3 Algorytm

Celem algorytmu jest identyfikacja nieznanych charakterystyk wejściowej i wyjściowej nieliniowości $\mu(\cdot)$, $\eta(\cdot)$ oraz wektora parametrów skończonej odpowiedzi impulsowej γ . Bazując tylko i wyłącznie na pomiarach wejścia oraz wyjścia systemu $\{u_k, y_k\}_{k=1}^N$, w dwóch eksperymentach: losowym $u_k^{(1)}$ oraz aktywnym $u_k^{(2)}$.

Zaproponowany algorytm bazuje zarówno na narzędziach identyfikacji parametrycznej jak i nieparametrycznej, w celu wyznaczenia parametrów każdego z bloków systemu Hammersteina-Wienera. Etapy algorytmu muszą być wykonywane w poniższej kolejności, z uwagi na zależność każdego etapu od wyniku etapu go poprzedzającego.

Opisywany algorytm złożony jest z czterech kroków:

Krok 1: Bezpośrednia identyfikacja parametrów skończonej odpowiedzi impulsowej γ liniowego bloku dynamicznego, z wykorzystaniem metody najmniejszych kwadratów wspomaganą selekcją jądrową, przy losowym sygnale wejściowym.

Krok 2: Estymacja wektora parametrów wyjściowej charakterystyki nieliniowej b z wykorzystaniem pobudzenia sygnałem binarnym oraz regresji jądrowej.

Krok 3: Wyznaczenie odwrotnej charakterystyki nieliniowej na wyjściu, z wykorzystaniem sygnału wyjściowego y_k . W efekcie wyznaczenie również pomocniczego sygnału r_k , o takiej samej wartości oczekiwanej jak niedostępny sygnał x_k .

Krok 4: Identyfikacja wektora parametrów wejściowej charakterystyki nieliniowej a , analogicznie jak dla systemu Hammersteina [28].

Zaprezentowany algorytm ma na celu minimalizację kryterium

$$Q(\hat{\gamma}, \hat{a}, \hat{b}) = E(y_k - \hat{y}_k)^2 \rightarrow \min_{\hat{\gamma}, \hat{a}, \hat{b}} \quad (3.4)$$

gdzie $y_k = y_k(\gamma, a, b)$ jest wyjściem systemu oraz $\hat{y}_k = \hat{y}_k(\hat{\gamma}, \hat{a}, \hat{b})$ jest wyjściem modelu o parametrach $\hat{\gamma}$, \hat{a} oraz \hat{b} .

3.3.1 Krok 1: identyfikacja odpowiedzi impulsowej

Rozważmy wektor regresora zawierający $q + 1$ kolejnych wejść systemu

$$\phi_k = \left(u_k^{(1)}, u_{k-1}^{(1)}, \dots, u_{k-q}^{(1)} \right)^T,$$

oraz funkcję jądra K

$$K(v) = \begin{cases} 1, & \text{jeżeli } |v| \leq 1 \\ 0, & \text{jeżeli } |v| > 0. \end{cases} \quad (3.5)$$

Zaproponowany estymator odpowiedzi impulsowej, bazujący na metodzie najmniejszych kwadratów, będzie miał postać $\hat{\gamma}$ [38]

$$\hat{\gamma} = \left(\sum_{k=1}^N \phi_k \phi_k^T K\left(\frac{\Delta_k}{h}\right) \right)^{-1} \left(\sum_{k=1}^N \phi_k y_k K\left(\frac{\Delta_k}{h}\right) \right), \quad (3.6)$$

gdzie Δ_k jest normą wektora regresji

$$\Delta_k = \|\phi_k\|_\infty = \max_{j=0,1,\dots,q} |u_{k-j}^{(1)}|. \quad (3.7)$$

Założono odwracalność macierzy $\frac{1}{N} \sum_{k=1}^N \phi_k \phi_k^T K\left(\frac{\Delta_k}{h}\right)$.

3.3.2 Krok 2: estymacja wyjściowej charakterystyki nieliniowej

W kolejnym kroku system jest pobudzony binarnym wejściem losowym $u_k^{(2)}$, gdzie prawdopodobieństwo wylosowania 0 oraz 1 jest takie same. Celem aktywnego eksperymentu jest zaślepienie wyjściowej charakterystyki nieliniowej, dzięki czemu niedostępny sygnał w_k również przyjmuje postać binarną. Możliwe kombinacje wektora regresji ϕ_k wyglądają zatem następująco

$$\begin{aligned} \varphi_1 &= (1, 0, 0, \dots, 0)^T, & \varphi_2 &= (0, 1, 0, \dots, 0)^T, \\ \varphi_3 &= (1, 1, 0, \dots, 0)^T, & \dots & \varphi_{N_0} = (1, 1, 1, \dots, 1)^T. \end{aligned}$$

Wszystkich możliwych wektorów jest N_0 , gdzie $N_0 = 2^{q+1}$. Dzięki wiedzy o parametrach odpowiedzi impulsowej z kroku 1, γ , możemy wyznaczyć siatkę deterministycznych punktów

$$x_{[i]} = \varphi_i^T \gamma, i = 1, 2, \dots, N_0$$

które reprezentują wszystkie możliwe wartości sygnału x_k . Następnie w tych wybranych punktach wyznacza się estymatory wyjściowej charakterystyki nieliniowej za pomocą estymatora jądrowego

$$\hat{\eta}(x_{[i]}) = \frac{\sum_{k=1}^N y_k \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}, \quad (3.8)$$

gdzie

$$\delta(\phi_k, \varphi_i) = \begin{cases} 1, & \text{jeżeli } \phi_k = \varphi_i \\ 0, & \text{w przeciwnym przypadku} \end{cases} \quad (3.9)$$

jest selektorem jądrowym oraz z $\frac{0}{0}$ rozumianym jako 0. Z uwagi na binarne wejście i warunek $\phi_k = \varphi_i$ parametr szerokości jądra h nie ma zastosowania. Aby uniknąć sytuacji w której mianownik estymatora jest równy 0, wprowadźmy dodatkowe oznaczenie wszystkich wyjść, dla których $\phi_k = \varphi_i$ dla danego punktu estymacji $x_{[i]}$: $y_{(1)}, y_{(2)}, \dots, y_{(L)}$ gdzie L jest losową liczbą wybranych przez estymator jądrowy wartości wyjścia. Estymator przyjmuje zatem postać

$$\hat{\eta}(x_{[i]}) = \text{Avg}(S(x_{[i]})) = \begin{cases} \frac{1}{L} \sum_{l=1}^L y(l), & \text{jeżeli } L > 0 \\ 0, & \text{w przeciwnym przypadku,} \end{cases} \quad (3.10)$$

gdzie

$$S(x_{[i]}) = \{y_k : \phi_k = \varphi_i\}. \quad (3.11)$$

Prawdopodobieństwo sytuacji w której $(\phi_k = \varphi_i)$ jest stałe i wynosi

$$P\{\delta(\phi_k, \varphi_i) = 1\} = \frac{1}{N_0} = \frac{1}{2^{q+1}}.$$

Rezultatem tego kroku jest N_0 par

$$\{(x_{[i]}, \hat{\eta}(x_{[i]}))\}_{i=1}^{N_0}. \quad (3.12)$$

Dzięki tym parom, możliwe jest znalezienie najlepszych wartości parametrów za pomocą metody najmniejszych kwadratów

$$\hat{b} = (\Psi^T \Psi)^{-1} \Psi^T \zeta, \quad (3.13)$$

gdzie

$$\begin{aligned} \Psi &= (g(x_{[1]}), g(x_{[2]}), \dots, g(x_{[N_0]})), \\ \zeta &= (\hat{\eta}(x_{[1]}), \hat{\eta}(x_{[2]}), \dots, \hat{\eta}(x_{[N_0]})). \end{aligned}$$

Odwracalność macierzy $\Psi^T \Psi$ (jak również macierzy $A_N^T A_N$ w (3.16)) zależy od pobudzenia, funkcji gęstości prawdopodobieństwa na wejściu oraz kształtu charakterystyki nieliniowej. Dodatkowo w tym kroku algorytmu, różnorodność wyjść dla każdego punktu $x_{[i]}$ zależy od obecności szumu z_k . W związku z tym funkcja gęstości prawdopodobieństwa może być estymowana bazując na różnicy względem wartości średniej $\hat{\eta}(x_{[i]})$ np. za pomocą metody jądrowej (tylko, jeżeli $L > 0$)

$$\hat{f}(z) = \frac{1}{Lh_z} \sum_{l=1}^L K\left(\frac{z^{(l)} - z}{h_z}\right). \quad (3.14)$$

3.3.3 Krok 3: estymacja sygnału wewnętrznego x_k za pomocą filtra odwrotnego

Po estymacji wektora parametrów wyjściowej charakterystyki nieliniowej b , wektor parametrów wejściowej charakterystyki nieliniowej a pozostaje nieznanym. W kroku 3 zaproponowano estymację sygnału x_k dzięki czemu pozostała

część systemu składająca się z charakterystyki nieliniowej na wejściu oraz z liniowego bloku dynamicznego może być sprowadzona do systemu Hammersteina. Synteza filtra odwrotnego nie jest częścią algorytmu i musi być wykonana niezależnie, przed uruchomieniem omawianego procesu identyfikacji. W Przykładach 1-3 przedstawiono obliczenia dla prostych charakterystyk nieliniowych. Wprowadźmy oznaczenia

$$\begin{aligned}v &= \eta(x), \\y &= v + z, \\ \xi(y) &= \eta^{-1}(y) + \varphi(y).\end{aligned}$$

Poszukujemy takiej funkcji $\varphi(y)$ zwanej dalej korektorem, aby dla ustalonego x zachodziło

$$E\xi(y) = E\xi(v + z) = x. \quad (3.15)$$

A zatem, uwzględniając, że $x = \eta^{-1}(v)$, postulujemy by

$$E\eta^{-1}(y) + E\varphi(y) = \eta^{-1}(v),$$

czyli

$$E\varphi(y) = \underbrace{\eta^{-1}(v) - E\eta^{-1}(v + z)}_{\triangleq \alpha(v)}.$$

Zadanie polega więc na tym, aby przy znanej (wyestymowanej) funkcji gęstości prawdopodobieństwa zakłóceń $f(z)$ i znanej funkcji $\alpha(v)$ znaleźć funkcję $\varphi()$, taką że

$$\int_{-\infty}^{\infty} \varphi(v + z) f(z) dz = \alpha(v).$$

Jest to tzw. zadanie rozplotu, powszechnie rozpatrywane w literaturze. Poniżej podano kilka przykładów dla postaci wielomianowych funkcji $\eta^{-1}()$ i zakłóceń o rozkładzie jednostajnym $z \sim \mathcal{U}[-1, 1]$.

Przykład 1. Dla funkcji liniowej $\eta^{-1}(v) = v$ mamy $\alpha(v) = v - E(v + z) = 0$, a zatem korekcja nie jest konieczna, tj. możemy przyjąć $\varphi(y) = 0$ i $\xi(y) = \eta^{-1}(y) = y$.

Przykład 2. Niech $\eta^{-1}(v) = v^2$. Wtedy $\alpha(v) = v^2 - E(v + z)^2 = -Ez^2 = -\frac{1}{3}$, stąd proponowane $\varphi(y) = -\frac{1}{3}$ oraz $\xi(y) = y^2 - \frac{1}{3}$.

Przykład 3. Niech $\eta^{-1}(v) = v^3$. Wtedy $\alpha(v) = v^3 - E(v + z)^3 = -3vEz^2 =$

$-v$. A ponieważ $\int_{-1}^1 -(v+z)f(z)dz = -v$ to $\varphi(y) = -y$ i stąd $\xi(y) = y^3 - y$.

W podejściu numerycznym można zastosować metodę Monte Carlo do wyznaczenia szukanej funkcji. Dzięki własności (3.15), w sytuacji, gdy funkcja $\xi(y)$ zostanie zastosowana dla wyjścia y_k systemu Hammersteina-Wienera, zachodzi

$$E\{\xi(y_k) | x_k = x\} = x.$$

3.3.4 Krok 4: identyfikacja parametrów wejściowej charakterystyki nieliniowej

W ostatnim kroku algorytmu, dzięki wygenerowanemu sygnałowi $r_k = \xi(y_k)$ identyfikacja pierwszej charakterystyki nieliniowej sprowadza się do systemu Hammersteina. Sygnał x_k można zapisać jako

$$x_k = \lambda_k^T \theta,$$

gdzie (patrz dodatek B.5)

$$\lambda_k = \left(f_1(u_k^{(1)}), \dots, f_m(u_k^{(1)}), \dots, f_1(u_{k-q}^{(1)}), \dots, f_m(u_{k-q}^{(1)}) \right)^T$$

jest wektorem regresora, oraz

$$\theta = (\gamma_0 a_1, \dots, \gamma_0 a_m, \dots, \gamma_q a_1, \dots, \gamma_q a_m)$$

jest wektorem iloczynów parametrów. Wprowadzając ogólny zapis $\Lambda_N = (\lambda_1^T, \lambda_2^T, \dots, \lambda_N^T)^T$ oraz $R_N = (r_1, r_2, \dots, r_N)^T$, estymator iloczynów $\hat{\theta}$ można zapisać z użyciem sygnału r_k jako

$$\hat{\theta} = (\Lambda_N^T \Lambda_N)^{-1} \Lambda_N^T R_N. \quad (3.16)$$

Dzięki znajomości parametrów odpowiedzi impulsowej można użyć metody SVD do estymacji parametrów \hat{a} . Wykorzystując wektor \hat{a} , estymator wejściowej charakterystyki nieliniowej można zapisać jako

$$\hat{\mu}(u) = \hat{\mu}(u, \hat{a}) = \hat{a}^T f(u), \quad (3.17)$$

gdzie $\hat{a} = (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_m)^T$.

3.4 Własności algorytmu

W Twierdzeniach 3.4.1 oraz 3.4.2 pokazano zbieżności estymatorów bloku liniowego oraz charakterystyki nieliniowej.

Twierdzenie 3.4.1 *Jeżeli system jest systemem Hammersteina–Wienera, oraz $h \sim N^{-\alpha}$, gdzie $\alpha \in (0, \frac{1}{q+3})$ to*

$$\hat{\gamma}_j \xrightarrow{N \rightarrow \infty} \gamma_j; \quad j = 0, 1, \dots, q; \quad (3.18)$$

według prawdopodobieństwa.

Dowód. Patrz dodatek A.1.

Twierdzenie 3.4.2 *Jeżeli $N \rightarrow \infty$, to*

$$E [\eta(x_{[i]}) - \hat{\eta}(x_{[i]})]^2 \rightarrow 0. \quad (3.19)$$

Dowód. Patrz dodatek A.2.

3.5 Przykład numeryczny

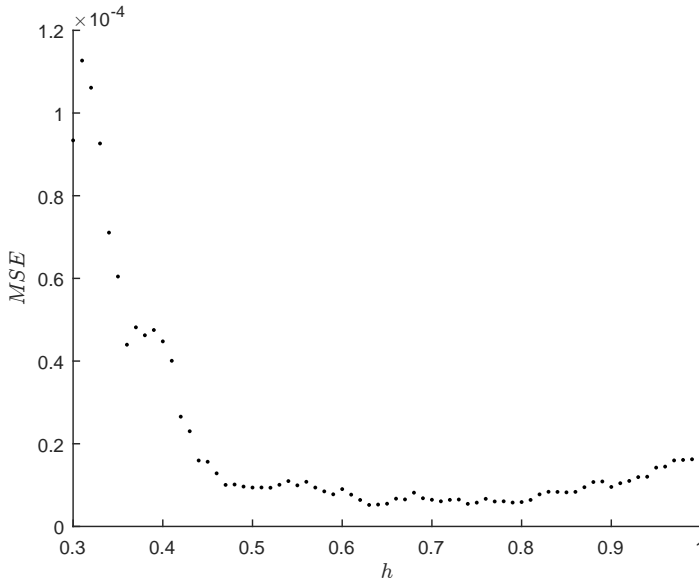
Aby zademonstrować działanie algorytmu w praktyce, zasymulowano prosty system Hammersteina–Wienera, gdzie nieliniowe charakterystyki statyczne są liniowymi kombinacjami jednomianów

$$\mu(u, a) = a_1 u + a_2 u^2, \quad a = (a_1, a_2)^T = (0.6, 0.4)^T,$$

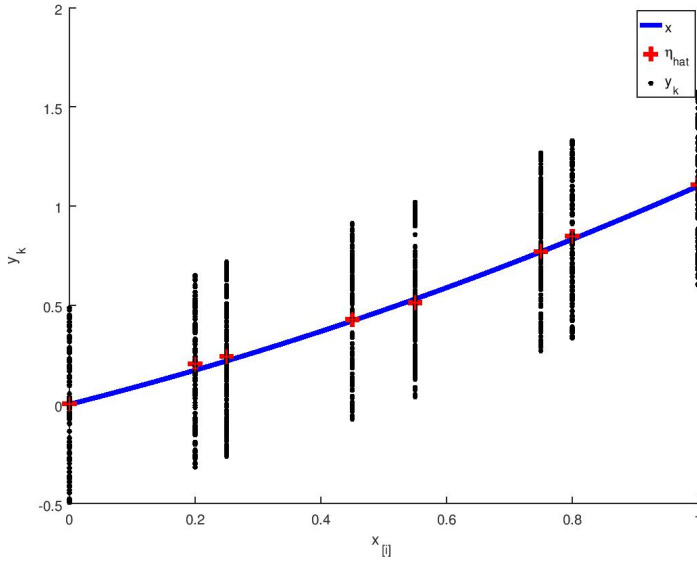
$$\eta(x, b) = b_1 x + b_2 x^2, \quad b = (b_1, b_2)^T = (0.8, 0.3)^T,$$

skończona odpowiedź impulsowa bloku dynamicznego przyjmuje zaś formę $\gamma = (0.55, 0.25, 0.2)^T$. Metoda opiera się na dwóch eksperymentach o różnych pobudzeniach na wejściu. W eksperymencie pierwszym, sygnał wejściowy ma charakter losowy o jednostajnym rozkładzie prawdopodobieństwa $u_k^{(1)} \sim \mathcal{U}(-1, 1)$. Następnie, w aktywnym eksperymencie, sygnał wejściowy $\{u_k^{(2)}\}$ przyjmuje postać losowych sekwencji binarnych. W obu przypadkach zakłócenia miały formę losowego procesu o jednostajnym rozkładzie prawdopodobieństwa $z_k \sim \mathcal{U}(-0.5, 0.5)$. W eksperymencie zastosowano $N = 10^5$ par pomiarów wejścia–wyjścia $\{(u_k, y_k)\}_{k=1}^N$. Przed uruchomieniem algorytmu, za pomocą sprawdzianu krzyżowego wybrano wartość parametru $h = 0, 63$

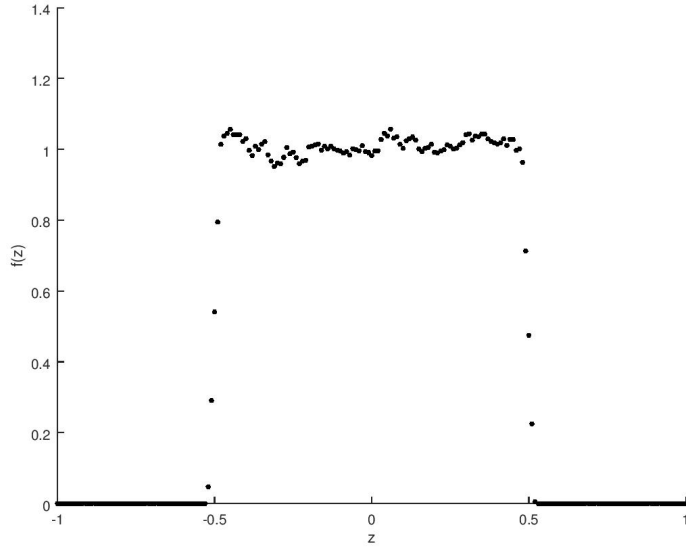
gwarantującą odpowiedni kompromis pomiędzy obciążeniem i wariancją (Rys. 3.2). W pierwszym kroku wyznaczono parametry odpowiedzi impulsowej za pomocą metody najmniejszych kwadratów wspomaganej selekcją jądrową: $\hat{\gamma} = (0.554, 0.24, 0.206)^T$. Błąd średniokwadratowy parametrów modelu i systemu wyniósł $MSE_{\gamma} = \frac{1}{q+1} \sum_{i=0}^q (\gamma_i - \hat{\gamma}_i)^2 = 1.52 \cdot 10^{-4}$. W drugim kroku przeprowadzono aktywny eksperyment. Wyjściowa charakterystyka nieliniowa została wyznaczona za pomocą estymatora jądrowego na siatce punktów $\{x_{[i]}\}_{i=1}^{N_0}$, gdzie $N_0 = 2^{q+1} = 8$. Na Rys. 3.3 przedstawiono pomiary wyjścia systemu, charakterystykę systemu oraz wyjście modelu. Wartości wyznaczonych parametrów wyniosły: $\hat{b} = (0.802, 0.298)^T$ zaś błąd średniokwadratowy $MSE_b = \frac{1}{2} \sum_{i=1}^2 (b_i - \hat{b}_i)^2 = 8 \cdot 10^{-6}$. Dodatkowo, za pomocą metody jądrowej wyznaczono estymator funkcji gęstości prawdopodobieństwa zakłóceń, w każdym punkcie pomiarowym $x_{[i]}$. Na Rys. 3.4 przedstawiono uzyskane wartości.



Rysunek 3.2. Sprawdzenie krzyżowe w celu wyznaczenia najlepszej szerokości jądra h



Rysunek 3.3. Nieparametryczny estymator wyjściowej charakterystyki nieliniowej (czerwone plusy) porównany do nieliniowej charakterystyki wyjściowej systemu $\eta(x)$ (niebieska krzywa). W tle pomiary wyjścia (czarne punkty)



Rysunek 3.4. Nieparametryczna estymacja gęstości rozkładu prawdopodobieństwa zakłóceń $f(z)$

Identyfikacja nieliniowych systemów dynamicznych o strukturze ogólnej metodą jądrową

Tradycyjne podejście do identyfikacji i modelowania systemów kojarzone jest z wiedzą na temat matematycznej reprezentacji systemu (np. równanie różniczkowe, struktura blokowa) oraz ze skończoną ilością nieznanych parametrów. Z uwagi na bogaty dorobek naukowy w tej dziedzinie, wiele algorytmów opierających się na takich założeniach, które cechują się relatywnie szybką zbieżnością estymowanych parametrów do wartości prawdziwych zostało zaprezentowanych. Niemniej, rozważając coraz bardziej skomplikowane procesy i złożone modele, przy takim założeniu można popełnić już na samym początku błąd, definiując niewłaściwie strukturę. W celu wyeliminowania tego problemu zaprezentowano poniżej opisany algorytm, skupiający się na identyfikacji systemu z bardzo ogólnymi założeniami, bez narzuconej z góry struktury, rezygnując również z blokowej reprezentacji takiej jak system Hammersteina lub system Wienera.

4.1 Opis problemu

Istotą niniejszego rozdziału jest zaproponowanie nieparametrycznego algorytmu identyfikacji, dla systemów o jak najbardziej ogólnych założeniach. W szczególności bardziej ogólnych niż systemy blokowo-zorientowane takie jak Hammersteina, Wienera i ich kombinacje. W podrozdziale 7.3 zostały przedstawione zastosowania praktyczne bazujące na opisanym algorytmie. Istotą opisywanego podejścia jest identyfikacja nieparametryczna wielowymiarowa, bazująca na aktywnym eksperymencie tj. specjalnych sekwencjach wejściowych.

4.2 Opis i klasa rozpatrywanych systemów

Rozważmy dyskretny nieliniowy system dynamiczny o ogólnej reprezentacji

$$y_k = \mathcal{F}(\{u_{k-i}\}_{i=0}^{\infty}) + z_k, \quad (4.1)$$

gdzie $\{u_{k-i}\}_{i=0}^{\infty}$ jest losową, ograniczoną sekwencją wejść ($|u_k| < u_{\max}$), z_k jest losowym zakłóceniem. Funkcjonał $\mathcal{F}(\cdot)$ cechuje się eksponencjalnym zaniżaniem pamięci oraz spełnia warunek Lipschitz'a $|\mathcal{F}(u_i) - \mathcal{F}(u_j)| \leq L|u_i - u_j|$, dla wszystkich możliwych wejść, gdzie L jest dowolną stałą rzeczywistą. Z uwagi na gasnącą pamięć systemu, zaprezentowany algorytm identyfikacji będzie się ograniczał tylko do skończonej jego odpowiedzi, gdzie s będzie "punktem odcięcia" i jednocześnie wymiarem modelu, więc w efekcie powstaje model o skończonej odpowiedzi impulsowej systemu o nieskończonej odpowiedzi impulsowej. Zatem

$$\bar{u}_{k-i} \triangleq \begin{cases} u_{k-i}, & \text{dla } i = 0, 1, \dots, s-1 \\ 0, & \text{dla } i \geq s \end{cases}. \quad (4.2)$$

Zakładając, że

$$\Delta(s) \triangleq |\mathcal{F}(\{u_{k-i}\}_{i=0}^{\infty}) - \mathcal{F}(\{\bar{u}_{k-i}\}_{i=0}^{\infty})| \leq c\lambda^s, \quad (4.3)$$

z pewną nieznaną wartością $c = \text{const}$, oraz $0 < \lambda < 1$. Celem jest identyfikacja systemu (stworzenie modelu $\hat{\mathcal{F}}(\cdot)$) wykorzystując sekwencję N pomiarów wejścia-wyjścia $\{(u_k, y_k)\}_{k=1}^N$.

Jak wspomniano we wprowadzeniu, w taki opis systemu można zaklasyfikować popularne w literaturze modele blokowe, takie jak Hammersteina, Wienera oraz Hammersteina-Wienera.

4.3 Wielowymiarowa regresja jądrowa

Podstawę algorytmu stanowi estymator jądrowy, wprowadźmy s -wymiarowy wektor wejść (tzw. regresor)

$$u_k^{(s)} = (u_k, u_{k-1}, \dots, u_{k-(s-1)})^T, \quad (4.4)$$

oraz funkcję regresji

$$R_s(x^{(s)}) = E(y_k | u_k^{(s)} = x^{(s)}), \quad (4.5)$$

z odpowiednim wektorem argumentów

$$x^{(s)} = (x_0, x_1, \dots, x_{s-1})^T. \quad (4.6)$$

Przykładowo, dla $s = 1$

$$R_1(x_0) = E(y_k | u_k = x_0), \quad (4.7)$$

oraz dla $s = 2$

$$R_2(x_0, x_1) = E(y_k | u_k = x_0 \cap u_{k-1} = x_1). \quad (4.8)$$

Estymator jądrowy [21, 23, 26, 53] dla $R_s(x^{(s)})$ przyjmuje zatem postać

$$\hat{R}_s(x^{(s)}) = \frac{\sum_{k=1}^N y_k K\left(\frac{\|u_k^{(s)} - x^{(s)}\|}{h}\right)}{\sum_{k=1}^N K\left(\frac{\|u_k^{(s)} - x^{(s)}\|}{h}\right)}, \quad (4.9)$$

gdzie $\|\cdot\|$ jest normą (np. Euklidesową), a $K(\cdot)$ pełni rolę jądra

$$K(v) = \begin{cases} 1, & \text{as } |v| \leq 1 \\ 0, & \text{as } |v| > 1 \end{cases}. \quad (4.10)$$

4.4 Redukcja wymiaru problemu

Kluczowym problemem jaki się pojawia w prezentowanej regresji jest wzrost wymiaru, co powoduje wykładnicze, ze względu na wymiar zmniejszanie się prawdopodobieństwa zaklasyfikowania nowego pomiaru do jądra, co w efekcie powoduje powolną zbieżność estymatora i dużą wariancję [41] [43] [56], co można przedstawić równaniem

$$P \left\{ K \left(\frac{\|u_k^{(s)} - x^{(s)}\|}{h} \right) = 1 \right\} = P \left\{ \|u_k^{(s)} - x^{(s)}\| \leq h \right\} \sim h^s. \quad (4.11)$$

Zaproponowane poniżej sygnały wejściowe redukują wymiar, ale należy nadmienić, że przy tej kompresji następuję pewne utrata informacji a system jest identyfikowany tylko w określonych punktach lub na danym zbiorze.

4.4.1 Kwantyzowane wejście

Założmy, że w każdym s -elementowym ciągu wartości wejściowych $u_k, u_{k+1}, \dots, u_{k+s-1}$, d wejść ma rozkład dyskretny. W związku z tym wszystkie punkty $u_k^{(s)} \in R^s$ należą do podprzestrzeni (tzw. rozmaitości) [52] o wymiarze

$$s^* = s - d, \quad (4.12)$$

Dla każdego z punktów pomiarowych, prawdopodobieństwo selekcji jądra jest zatem proporcjonalne do ch^{s^*} , gdzie s^* , ($s^* < s$), jest wewnętrznym wymiarem tej podprzestrzeni. Dla $d = s$ (wszystkie sygnały wejściowe są skwantowane) uzyskujemy $MSE(\hat{R}_s(u_k^{(s)})) \sim N^{-1}$. Przykłady możliwych przestrzeni dla $s = 2$ zostały pokazane na Rysunkach 4.1 oraz 4.2.

4.4.2 Sygnał okresowy

Założmy, że wejście ma charakter okresowy, z okresem N_0 , tzn., $u_k = u_{k+N_0}$ dla każdego $k = 1, 2, \dots, N - N_0$. Wówczas regresor $u_k^{(s)}$ ((4.4) oraz (4.5)) przyjmuje wartość dla N_0 różnych punktów w R^s , $n \in Z$,

$$x_{[1]}^{(s)} = \begin{bmatrix} u_s \\ u_{s-1} \\ \vdots \\ u_1 \end{bmatrix}, \dots, x_{[N_0]}^{(s)} = \begin{bmatrix} u_{N_0+s-1} \\ u_{N_0+s-2} \\ \vdots \\ u_{N_0} \end{bmatrix} \quad (4.13)$$

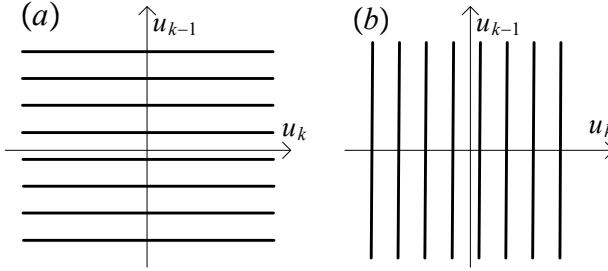
z prawdopodobieństwami

$$P \{ u_k^{(s)} = x_{[p]}^{(s)} \} = \frac{1}{N_0}, p = 1, 2, \dots, N_0 \quad (4.14)$$

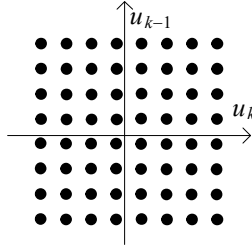
Pomiary są jednostajnie rozłożone na skończonej liczbie N_0 punktów

$$x_{[1]}^{(s)}, \dots, x_{[p]}^{(s)}, \dots, x_{[N_0]}^{(s)} \quad (4.15)$$

Zmniejszenie wartości parametru wygładzania $h(N)$ nie wpływa asymptotycznie na estymator jądrowy (ponieważ $s^* = 0$). W związku z czym, rząd szybkości zbieżności wynosi $MSE \sim N^{-1}$. Jednakże należy nadmienić, że estymator jest wyznaczony dla skończonej ilości punktów, oraz zwiększanie N_0 powoduje wzrost wariancji estymatora regresji dla poszczególnych punktów $x_{[p]}^{(s)}$.



Rysunek 4.1. Przestrzeń wejść dla $s = 2$ oraz $d = 1$, (a) k nieparzyste, (b) k parzyste.



Rysunek 4.2. Przestrzeń wejść dla $s = 2$ oraz $d = 2$.

4.5 Własności algorytmu

Uwaga 4.1 Jak pokazują badania, kształt funkcji jądra ma dużo mniejszy wpływ na zbieżność estymatora, niż parametr h [18] [23], stąd też w tym przypadku skupiono się tylko na jądrze w postaci okna (funkcja Parzena).

Twierdzenie 4.5.1 Jeżeli przyjmiemy kryterium zbieżności zgodnie z definicją MSE (Dodatek B.3.1), to

$$MSE\left(\widehat{R}_s\left(u_k^{(s)}\right)\right) = o\left(N^{-\frac{2}{s^*+2}}\right). \quad (4.16)$$

Dowód. Patrz dodatek A.3.

Przykład. Dla wejściowego sygnału w pełni skwantowanego $s = d$, a w konsekwencji $s^* = 0$, szybkość zbieżności błędu zredukowano do

$$MSE\left(\widehat{R}_s\left(u_k^{(s)}\right)\right) = o\left(N^{-1}\right). \quad (4.17)$$

4.6 Przykład numeryczny

W celu ilustracji zaproponowanej metody, wykonano symulację prostego systemu Wienera

$$x_k = 0.5x_{k-1} + u_k \quad (4.18)$$

$$v_k = \arctg(x_k) \quad (4.19)$$

$$y_k = v_k + z_k \quad (4.20)$$

popudzonego losowym wejściem o rozkładzie jednostajnym na równoodległych punktach

$$u_k \sim \{-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1\}, \quad (4.21)$$

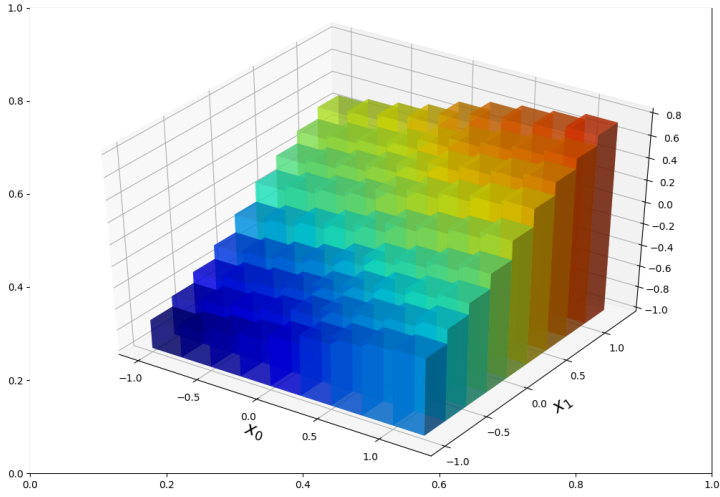
o wyjściowym zakłóceniu, również o rozkładzie jednostajnym

$$z_k \sim U[-0.1, 0.1]. \quad (4.22)$$

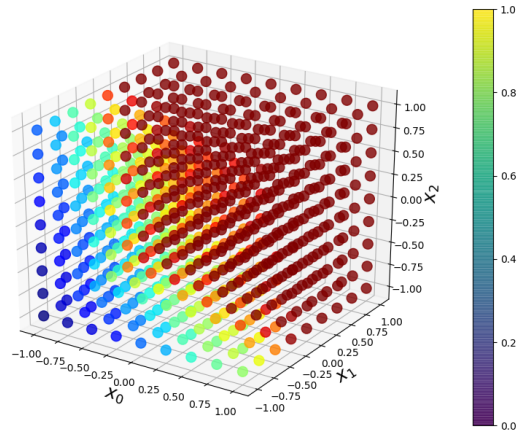
Dla $N = 10^4$ par wejść-wyjść $\{(u_k, y_k)\}_{k=1}^N$, nieparametryczny model $\widehat{R}_s(u_k)$ został wyznaczony dla $s = 1, 2, 3$. Modele zostały porównane ze względu na kryterium

$$\delta_s = \frac{1}{N - s + 1} \sum_{k=s}^N \left(y_k - \widehat{R}_s \left(u_k^{(s)} \right) \right)^2. \quad (4.23)$$

Wyniki pokazano w tabeli 4.1 oraz na rysunkach 4.3 oraz 4.4. Rysunki 4.3 oraz 4.4 ilustrują nieparametryczny oraz nieliniowy charakter modelu, pokazując również ogólny charakter związku wejścia z wyjściem. Aby uniknąć błędu związanego z nadmiernym uproszczeniem modelu, czyli zastosowaniu zbyt restrykcyjnej skali s w stosunku do ogólnego obiektu o nieskończonej odpowiedzi impulsowej, kwantyzacja wejścia (4.21) może być dobrym wyborem, kiedy $s = \text{const} < \infty$, oraz nieparametryczny estymator $\widehat{R}_s \left(u_k^{(s)} \right)$ pełni rolę wspierającą dla nieliniowego algorytmu najmniejszych kwadratów. Niemniej jednak, w czysto nieparametrycznym podejściu, tzn., dla $s(N) \rightarrow \infty$, liczba możliwych wystąpień of $u_k^{(s)}$ wzrasta eksponencjalnie. W rozważanym przykładzie, dla 9 punktów w (4.21), prawdopodobieństwo klasyfikacji jądra $x^{(s)} = u_k^{(s)}$ jest równe 9^{-s} oraz estymator staje się wrażliwy dla zakłócenia z_k .



Rysunek 4.3. 2-wymiarowy model $\widehat{R}_2(x_0, x_1)$.



Rysunek 4.4. 3-wymiarowy model $\widehat{R}_3(x_0, x_1, x_2)$.

Tabela 4.1. Porównanie błędu średniokwadratowego δ_s wyjścia modelu dla $s = 1, 2, 3$, do najlepszej możliwej aproksymacji liniowej (BLA).

	δ_1	δ_2	δ_3
\hat{R}_s	0.0656	0.0279	0.0137
BLA	0.0660	0.0304	0.0176

W tabeli 4.1 Pokazano redukcje błędu wraz ze wzrostem skali modelu. Wynik ten porównano również z najlepszą możliwą aproksymacją liniową $FIR(s)$ systemu Wienera. Warto podkreślić, że wyniki te uzyskano przy stosunkowo wysokiej ogólności systemu, tzn bazując jedynie na pomiarach wejścia i wyjścia.

4.7 Propozycje potencjalnych zastosowań

Pierwszym zastosowaniem jest identyfikacja systemu tylko w punktach istotnych z uwagi na charakter problemu, np. w punktach pracy lub dla charakterystycznych wartości.

Z uwagi na sposób pobudzenia, charakterystyka systemu jest modelowana tylko w niektórych punktach, w celu rozszerzenia modelu dla dowolnego wejścia $\{u_k\}$, konieczna jest dodatkowa wiedza o systemie. W tym podrozdziale założono, że funkcjonal $\mathcal{F}(\{u_{k-i}\}_{i=0}^{\infty})$ należy do znanej, skończonej z uwagi na wymiar, klasy systemów $F(u_k^{(s)}, \theta)$

$$\mathcal{F}(\{u_{k-i}\}_{i=0}^{\infty}) \subset F(u_k^{(s)}, \theta) \quad (4.24)$$

z nieznanym wektorem parametrów θ . W prezentowanym podejściu, może być zastosowane jedno z pobudzeń zaprezentowanych w punkcie 4.4. System jest identyfikowany na skończonym zestawie N_0 punktów $x_{[1]}^{(s)}, x_{[2]}^{(s)}, \dots, x_{[p]}^{(s)}, \dots, x_{[N_0]}^{(s)}$, gdzie $x_{[p]}^{(s)} \in R^s$, oraz $p = 1, 2, \dots, N_0$. Zapiszmy rzeczywisty i nieznaną wektor parametrów jako θ^* . Załóżmy, że θ^*

$$\mathcal{F}(\{u_{k-i}\}_{i=0}^{\infty}) = F(u_k^{(s)}, \theta) \iff \theta = \theta^*. \quad (4.25)$$

Ponadto załóżmy, że kryterium

$$Q(\theta) = E \left[y_k - F \left(u_k^{(s)}, \theta \right) \right]^2 \quad (4.26)$$

jest wypukła na $\theta \in \Xi$, gdzie Ξ jest pewnym sąsiedztwem rzeczywistego wektora θ^*

$$\theta^* = \arg \min_{\theta} Q(\theta). \quad (4.27)$$

W celu takiej mieszanej identyfikacji (parametryczno-nieparametrycznej [28] [32] [39] [47]) zaproponowano poniższy 2-etapowy algorytm. Krok 1. (nieparametryczny) Wykorzystując pomiary wejścia-wyjścia $\{(u_k, y_k)\}_{k=1}^N$ dla $p = 1, 2, \dots, N_0$ wyznaczmy estymator

$$\widehat{R}_s(x_{[1]}^{(s)}), \dots, \widehat{R}_s(x_{[p]}^{(s)}), \dots, \widehat{R}_s(x_{[N_0]}^{(s)}), \quad (4.28)$$

Krok 2. (parametryczny) Minimalizacja empirycznego wariantu kryterium najmniejszych kwadratów (4.26)

$$\widehat{\theta} = \arg \min_{\theta} \frac{1}{N_0} \sum_{p=1}^{N_0} \left(\widehat{R}_s \left(x_{[p]}^{(s)} \right) - F \left(x_{[p]}^{(s)}, \theta \right) \right)^2. \quad (4.29)$$

Lemat 4.2 *Jeżeli błąd nieparametrycznego estymatora*

$$\left\| \widehat{R}_s \left(x_{[p]}^{(s)} \right) - F \left(x_{[p]}^{(s)}, \theta^* \right) \right\| = O \left(N^{-\tau} \right), \quad (4.30)$$

w rozumieniu średniokwadratowym, dla wszystkich $p = 1, 2, \dots, N_0$, wówczas w kroku 2 (parametrycznym)

$$\left\| \widehat{\theta} - \theta^* \right\| = O \left(N^{-\tau} \right). \quad (4.31)$$

Dowód. Niech $F \left(u_k^{(s)}, \theta \right)$ spełnia warunek Lipshitz'a dla wszystkich u_{k-l} zawartych w $u_k^{(s)}$ oraz wszystkich parametrów θ . Własność (4.31) można dowieść analogicznie do dowodu Twierdzenia 1 w [29] (patrz również uwaga w dodatku A.4 oraz przykład w dodatku A.5).

Rekurencyjna identyfikacja wielopoziomowa, z dekompozycją i predykcją interakcji

Celem algorytmu jest identyfikacja systemu Hammersteina (blokowy system N-L Rys. 5.1), przy uniknięciu problemu agregacji parametrów [3]. Zaproponowany algorytm bazuje na naprzemiennej identyfikacji bloków nieliniowego i dynamicznego ze wzajemną koordynacją oraz nieparametrycznie wyznaczaną nieliniowością. Dzięki takiej dekompozycji złożonego systemu blokowego, problem przekleństwa wielowymiarowości (2.4.2) nie stanowi przeszkody, ponieważ zamiast jednego zagregowanego wektora parametrów, mamy do czynienia z niezależną identyfikacją każdego z bloków, co znacząco redukuje problem występujący np. w [3], gdzie agregacja parametrów może doprowadzić do bardzo dużych wektorów co w konsekwencji może uniemożliwić zastosowanie algorytmu w praktyce. Ponadto rekurencyjny wariant algorytmu znacznie wpływa na wydajność obliczeniową metody (2.4.4).

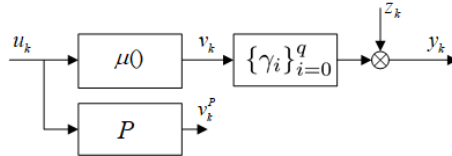
Historia naprzemiennej identyfikacji bloków w systemie Hammersteina sięga lat 60. XX wieku, gdzie zaproponowany przez Narendra'ego oraz Gallman'a algorytm [44] iteracyjnie identyfikował współczynniki wielomianu, będącego opisem systemu nieliniowego (ilość współczynników była skończona i znana) oraz parametry odpowiedzi impulsowej. W latach 60. oraz 70. nastąpił rozwój tego podejścia [5] [48], opierając się głównie o metody parametryczne, jak metoda najmniejszych kwadratów czy też zmiennych instrumentalnych, wciąż zakładając bogatą wiedzę *a priori* o systemie. Z drugiej zaś strony, równolegle w latach 80. Greblicki oraz Pawlak [20] zajmowali się zastosowaniem dokonania z pola statystyki w celu nieparametrycznej identyfikacji nieliniowości, wykorzystując jądrowy estymator funkcji regresji Nadaraya-Watson'a, a następnie rozwinięcia ortogonalne funkcji [17]. Istotną wadą metod nieparametrycznych

bazujących na estymacji za pomocą funkcji regresji jest zniekształcenie charakterystyki poprzez przeskalowanie oraz przesunięcie o pewnie nieznaną stałą.

Technika zaproponowana po raz pierwszy w (Hasiewicz, Mzyk, [58]) łączy zarówno ideę naprzemiennej estymacji bloków w systemie Hammersteina jak i nieparametrycznie wyznaczaną nieliniowość, z zastosowaniem koordynacji (patrz Dodatek A.6). Metoda ta została tutaj rozszerzona na systemy z blokiem liniowym typu ARMAX [36] oraz zastosowana w praktyce, na rzeczywistych danych. Nieparametryczne podejście do nieliniowości eliminuje ryzyko błędnego określenia struktury i ilości parametrów, a wariant rekurencyjny proponowanej metody znacząco redukuje złożoność obliczeniową. Opisywany w tym rozdziale rekurencyjny algorytm został zastosowany do modelowania termicznego procesu szkielek chalkogenidkowych, co zostało opisane w sekcji 7.2.

5.1 Opis problemu

Zaprezentowany algorytm identyfikuje system Hammersteina za pomocą predykcji niemierzalnego sygnału interakcji v_k (rys. 5.1). Predyktor pełni rolę estymatora charakterystyki nieliniowej $\mu(\cdot)$, wykorzystując regresję nieparametryczną (np. regresję jądrową lub ortogonalną). Parametrycznie zdefiniowany zostaje jedynie liniowy blok dynamiczny.



Rysunek 5.1. System Hammersteina z predykcją sygnału wewnętrznego (interakcyjnego).

Badany system jest opisany równaniem

$$y_k = \sum_{i=0}^q \gamma_i \mu(u_{k-i}) + z_k, \quad (5.1)$$

gdzie q jest długością odpowiedzi impulsowej, $\mu(\cdot)$ charakterystyką nieliniową, z_k losowymi zakłóceniami o wartości oczekiwanej równej zero, k jest numerem pomiaru, a liczbę wszystkich pomiarów będziemy oznaczać N . Wykorzystując powyższe oznaczenia oraz założenie o charakterze zakłóceń, regresję całego systemu można zapisać jako

$$\begin{aligned} R(u) &\triangleq E[y_k | u_k = u] = E \left[\sum_{i=0}^q \gamma_i \mu(u_{k-i}) + z_k | u_k = u \right] \\ &= \gamma_0 \mu(u) + E \left[\sum_{i=1}^q \gamma_i v_{k-i} | u_k = u \right] + E[z_k | u_k = u] \\ &= \mu(u) + d. \end{aligned} \quad (5.2)$$

Zatem opisywana regresja jest przesunięta względem charakterystyki $\mu(\cdot)$ o pewną nieznaną stałą d , którą można zapisać jako

$$d = E v_k \cdot \sum_{i=1}^q \gamma_i. \quad (5.3)$$

Zadaniem predyktora P (Rys. 5.1) bazującego na regresji nieparametrycznej, jest identyfikacja charakterystyki niedostępnego sygnału v_k ,

$$v_k^P(c) = P \left(u_k; \{u_k, y_k\}_{k=1}^N \right) = \hat{R}(u_k) - c, \quad (5.4)$$

gdzie c jest parametrem predyktora, którego celem jest identyfikacja stałej d .

Równoległym procesem jest identyfikacja bloku dynamicznego. Do tego celu wykorzystano metodę najmniejszych kwadratów, gdzie parami wejść i wyjść będą $\{v_k^P(c), y_k\}_{k=1}^N$, a wynikowy wektor parametrów $\hat{\gamma}_N(c)$ będzie estymatorem odpowiedzi impulsowej systemu γ . Kryterium które minimalizuje algorytm ma postać

$$Q(c, \hat{\gamma}) = \mathbf{E} \|Y_N - \bar{Y}_N(c, \hat{\gamma})\|_2^2 \rightarrow \min_{c, \hat{\gamma}}. \quad (5.5)$$

W podrozdziałach 5.2 oraz 5.3 przedstawiono dwa warianty omawianego algorytmu - klasyczny oraz rekurencyjny. W obu przypadkach do wyznaczenia estymatora regresji bloku nieliniowego można zastosować np. regresję jądrową lub regresję ortogonalną. Również wspólna jest idea identyfikacji bloku dynamicznego z użyciem metody najmniejszych kwadratów, z tą różnicą, że w

algorytmie *off-line* zastosowano jej klasyczną wersję, a w algorytmie *on-line* wersję rekurencyjną.

W podrozdziale 5.4 przedstawiono zaś uogólnienie metody rekurencyjnej dla systemów z elementem autoregresyjnym ARMAX.

5.2 Algorytm off-line bez rekurencji

Opisywany wariant algorytmu bazuje na klasycznej wersji metody najmniejszych kwadratów, stąd też identyfikacja następuje po zebraniu wszystkich N pomiarów, a następnie uruchomieniu procedury. Ostatnim etapem jest rozwiązanie dwupoziomowego zadania optymalizacji z koordynacją.

Zacznijmy od kolekcji wartości predykowanych sygnału interakcyjnego (patrz Rys. 5.1 oraz równanie (5.4)), który można zapisać jako

$$V_N^P = V_N^P(c) = \left(\bar{v}_k^P, \bar{v}_{k+1}^P, \dots, \bar{v}_{k+(N-1)}^P \right)^T, \quad (5.6)$$

gdzie \bar{v}_k^P jest wektorem aktualnego sygnału predyktora v_k^P oraz jego q -poprzednich wartości, przesuniętym o wartość parametru c ,

$$\left(\bar{v}_k^P \right)^T = \widehat{R}(\bar{u}_k) - c [1, 1, \dots, 1]_{1 \times (q+1)}, \quad (5.7)$$

gdzie

$$\widehat{R}(\bar{u}_k) = \left[\widehat{R}(u_k), \widehat{R}(u_{k-1}), \dots, \widehat{R}(u_{k-q}) \right], \quad (5.8)$$

jest tablicą aktualnej i q -poprzednich wartości nieparametrycznego estymatora regresji bloku nieliniowego $\mu(\cdot)$. Symbol k oznacza numer pomiaru wykorzystanego do obliczenia estymatora regresji. Wprowadźmy wektor zaobserwowanych wyjść systemu

$$Y_N = \left(y_k, y_{k+1}, \dots, y_{k+(N-1)} \right)^T. \quad (5.9)$$

Wektor wyjść modelu, zależny od c oraz $\hat{\gamma}$, ma postać

$$\bar{Y}_N(c, \hat{\gamma}) = V_N^P(c) \hat{\gamma}. \quad (5.10)$$

Dla systemu Hammersteina funkcja opisująca kryterium (5.5) jest wypukłą [42] ze względu na zmienną c . Wykorzystując zatem tradycyjny algorytm najmniejszych kwadratów można zapisać

$$\hat{\gamma}_N(c) = \left(V_N^{P^T}(c) V_N^P(c) \right)^{-1} V_N^{P^T}(c) Y_N. \quad (5.11)$$

Idąc dalej i traktując całą procedurę jako dwupoziomowe zadanie optymalizacji, gdzie parametr c pełni rolę zmiennej górnego poziomu a współczynniki γ optymalizują poziom dolny można sprowadzić kryterium (5.5) do postaci zależnej od jednej zmiennej

$$\bar{Q}(c) = Q(c, \hat{\gamma}_N(c)). \quad (5.12)$$

Tak przedstawiony problem może być rozwiązany z wykorzystaniem tradycyjnego programowania wypukłego, opisanym w dodatku B.6.

Własność 5.2.1 *Jeżeli $\hat{R}(u) \rightarrow R(u)$ według prawdopodobieństwa, dla $N \rightarrow \infty$, to*

$$c \rightarrow d, \quad (5.13)$$

$$V_N^P \rightarrow V_N, \quad (5.14)$$

$$\gamma_N(d) \rightarrow \gamma, \quad (5.15)$$

według prawdopodobieństwa.

Uzasadnienie. Z uwagi na wypukłość $\bar{Q}(c)$, asymptotycznie uzyskujemy (patrz Dodatek A.6)

$$\arg \min_c \bar{Q}(c) \rightarrow d \text{ według prawdopodobieństwa.} \quad (5.16)$$

Wyznaczając minimum (5.12) oraz c za pomocą algorytmu optymalizacji wypukłej (Dodatek B.6), parametr c zbiega do prawdziwej wartości d , $c \rightarrow d$ według prawdopodobieństwa. Przedstawione rozwiązanie ma jedną zasadniczą wadę – własności obliczeniowe (zajętość pamięci). Algorytm wymaga przetrzymywania wszystkich pomiarów, co powoduje obliczenia na stosunkowo długich wektorach, co przekłada się na niską wydajność w kontekście mocy obliczeniowej komputera. Fakt ten zasadniczo utrudnia wykorzystanie algorytmu w praktyce.

5.3 Rekurencyjny algorytm on-line

Rekurencyjny algorytm bazuje na naprzemiennym identyfikowaniu przesunięcia charakterystyki nieliniowej d oraz odpowiedzi impulsowej $\{\gamma\}$. Kolejna iteracja obliczeń jest uruchamiana kiedy pojawi się nowa para pomiarów $\{x_k, y_k\}$,

a parametry c oraz $\hat{\gamma}_k$ są aktualizowane, bez potrzeby przechowywania ich poprzednich wartości w pamięci. Tak zaprojektowany algorytm zapewnia wysoką wydajność obliczeniową.

Algorytm jest rozbity na osiem kroków, które są wykonywane po odczytaniu kolejnego pomiaru wejścia-wyjścia. W reprezentowanej procedurze N jest ostatnim pomiarem, a cały proces rozpoczyna się od $N = 1$, gdzie dla części zmiennych $N = 0$ oznacza wartości początkowe. Zapiszmy parametr c za pomocą ułamka

$$c_N = \frac{\mathcal{L}_N}{\mathcal{K}_N}, \quad (5.17)$$

gdzie licznik i mianownik będą wyznaczane w kolejnych krokach algorytmu.

Krok 1: Obliczenie zmiennych dodatkowych ρ_{N-1} , G_N oraz g_N ,

$$\rho_{N-1} \triangleq \hat{R}^T(\bar{u}_{k+(N-1)}) \left[y_{k+(N-1)} - \hat{R}(\bar{u}_{k+(N-1)}) L^{N-1} \right], \quad (5.18)$$

$$g_N \triangleq G_N \rho_{N-1}, \quad (5.19)$$

gdzie

$$G_N \triangleq G_{N-1} - \frac{G_{N-1} \hat{R}^T(\bar{u}_{k+(N-1)}) \hat{R}(\bar{u}_{k+(N-1)}) G_{N-1}}{1 + \hat{R}(\bar{u}_{k+(N-1)}) G_{N-1} \hat{R}^T(\bar{u}_{k+(N-1)})}, \quad (5.20)$$

$$G_0 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{N \times N}, \quad (5.21)$$

oraz

$$L_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{1 \times N}. \quad (5.22)$$

Krok 2: Aktualizacja macierzy L^N oraz \mathbf{M}^N ,

$$L^N = L^{N-1} + G_N g_N, \quad (5.23)$$

$$\mathbf{M}^N = \frac{N-1}{N} \mathbf{M}^{N-1} + \frac{1}{N} \mathbf{R}^{N-1}, \quad (5.24)$$

gdzie

$$\mathbf{M}^N = \mathbf{M}_{(q+1) \times (q+1)} = \frac{1}{2N} \left[P_N^T \mathbf{1}_{N \times (q+1)} + (P_N^T \mathbf{1}_{N \times (q+1)})^T \right], \quad (5.25)$$

$$\mathbf{M}_0 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{M+1 \times q+1}, \quad (5.26)$$

$$\mathbf{R}^{N-1} = [\widehat{R}_{ij}]_{(q+1) \times (q+1)}, \quad (5.27)$$

$$P_N \triangleq \begin{bmatrix} \widehat{R}(\bar{u}_k) \\ \widehat{R}(\bar{u}_{k+1}) \\ \vdots \\ \widehat{R}(\bar{u}_{k+(N-1)}) \end{bmatrix}, \quad (5.28)$$

$$\bar{R}_{ij} = \frac{\widehat{R}(u_{k+(N-1)-i}) + \widehat{R}(u_{k+(N-1)-j})}{2}, \quad (5.29)$$

oraz $\mathbf{1}_{(q+1) \times (q+1)}$ jest macierzą jedynek o rozmiarze $(q+1) \times (q+1)$.

Krok 3: Aktualizacja licznika $c_N - \mathcal{L}_N$,

$$\mathcal{L}_N = \left(\frac{N-1}{N} \right) \mathcal{L}_{N-1} + \mathcal{K}_{\mathcal{L}}^N, \quad (5.30)$$

gdzie

$$\mathcal{K}_{\mathcal{L}}^N = g_N^T \mathbf{M}^N g_N + 2L^{N-1T} \mathbf{M}^N g_N + \frac{1}{N} L^{N-1T} \mathbf{R}^{N-1} L^{N-1}, \quad (5.31)$$

oraz

$$L_0^{N-1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{1 \times N}. \quad (5.32)$$

Krok 4: Aktualizacja mianownika $c_N - \mathcal{M}_N$,

$$\mathcal{M}_N = \mathcal{M}_{N-1} + \mathcal{K}_{\mathcal{M}}^N, \quad (5.33)$$

gdzie

$$\mathcal{K}_{\mathcal{M}}^N = g_N^T \mathbf{1}_{(q+1) \times (q+1)} g_N + 2L^{N-1T} \mathbf{1}_{(q+1) \times (q+1)} g_N. \quad (5.34)$$

Krok 5: Obliczanie zaktualizowanej wartości parametru c_N ,

$$c_N = \frac{\mathcal{L}_N}{\mathcal{M}_N}. \quad (5.35)$$

Krok 6: Aktualizacja tabeli predyktorów $\bar{v}_{k+(N-1)}(c)$ z użyciem nowego c_N ,

$$\left(\bar{v}_{k+(N-1)}^P\right)^T = \widehat{R}^P \left(\bar{u}_{k+(N-1)}\right) - c_N [1, 1, \dots, 1]_{1 \times (q+1)}. \quad (5.36)$$

Krok 7: Zastosowanie standardowego algorytmu rekurencyjnego najmniejszych kwadratów,

$$\Gamma_N = \Gamma_{N-1} - \frac{\Gamma_{N-1} \bar{v}_{k+(N-1)}^P \left(\bar{v}_{k+(N-1)}^P\right)^T \Gamma_{N-1}}{1 + \left(\bar{v}_{k+(N-1)}^P\right)^T \Gamma_{N-1} \bar{v}_{k+(N-1)}^P}. \quad (5.37)$$

Krok 8: Aktualizacja $\widehat{\gamma}_N$,

$$\widehat{\gamma}_N = \widehat{\gamma}_{N-1} + \Gamma_N \bar{v}_{k+(N-1)}^P \left[y_{k+(N-1)} - \left(\bar{v}_{k+(N-1)}^P\right)^T \widehat{\gamma}_{N-1} \right], \quad (5.38)$$

oraz normalizacja wyniku;

$$\widehat{\gamma}_N := \frac{\widehat{\gamma}_N}{\widehat{\gamma}_N[0]}. \quad (5.39)$$

Podsumowując powyższy algorytm, poniżej przedstawiono asymptotyczne właściwości zaprezentowanego rozwiązania.

Własność 5.3.1 *Jeżeli $\widehat{R}(u) \rightarrow R(u)$, to $c_N \rightarrow d$ według prawdopodobieństwa dla $N \rightarrow \infty$.*

Uzasadnienie. Rekurencyjny estymator odpowiedzi impulsowej staje się równoważny wersji off-line, jak w równaniu (5.11).

5.4 Uogólnienie algorytmu do systemów klasy ARMAX

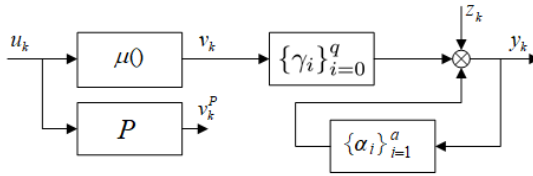
Aby uogólnić algorytm do systemów z autoregresją zmodyfikowano wektor parametrów odpowiedzi impulsowej γ oraz postać sygnału \bar{v}_N^P .

$$\widehat{\gamma}_N = (1, \widehat{\gamma}_1, \dots, \widehat{\gamma}_q, \alpha_1, \alpha_2, \dots, \alpha_a)^T. \quad (5.40)$$

Dodatkowe parametry α_i są parametrami odpowiedzi impulsowej wynikającej z dodatkowego bloku dynamicznego, co pokazano na rysunku 5.2.

Analogicznie wektor predyktora \bar{v}_N^P został rozszerzony o a poprzednich pomiarów wyjścia systemu,

$$\bar{v}_N^P = (v_{k+(N-1)}^P, v_{k+(N-1)-1}^P, \dots, v_{k+(N-1)-q}^P, y_{k+(N-1)-1}, y_{k+(N-1)-2}, \dots, y_{k+(N-1)-a})^T. \quad (5.41)$$



Rysunek 5.2. System Hammersteina z elementem liniowym typu ARMAX.

5.5 Badania symulacyjne

Dla zaprezentowania metody, wykonano symulacje poniższego systemu Hammersteina,

$$\mu(u) = \sin(\pi u) + \cos(\pi u) + \frac{\sqrt{2}}{2}, \quad (5.42)$$

$$x_k = v_k + v_{k-1} \text{ i.e., } \gamma = (1, 1)^T, \quad (5.43)$$

$$z_k \sim \mathcal{U}[-0.01, 0.01], \quad (5.44)$$

$$y_k = x_k + z_k, \quad (5.45)$$

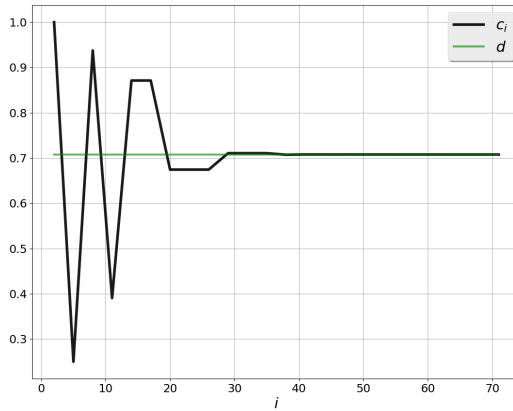
który został pobudzony sygnałem losowym o jednostajnym rozkładzie prawdopodobieństwa i niezależnych zmiennych

$$u_k \sim \mathcal{U}[-1, 1]. \quad (5.46)$$

W podrozdziałach 5.5.1 oraz 5.5.2, przedstawiono zarówno wariant off-line jak i on-line algorytmu.

5.5.1 Algorytm off-line

Z wykorzystaniem algorytmu opisanego w 5.2 estymator c osiągnął wartość $c = 0.70687$. Oczywiście prawdziwa wartość przesunięcia regresji, wynikająca ze wzoru (5.42) wynosi $d = \frac{\sqrt{2}}{2}$, zatem błąd estymatora wyniósł $c - d = -2.3 \times 10^{-4}$. Na Rysunku 5.3 pokazano zmianę estymatora podczas kolejnych kroków optymalizacji.



Rysunek 5.3. Zmiana estymatora c podczas kolejnych iteracji.

5.5.2 Algorytm on-line

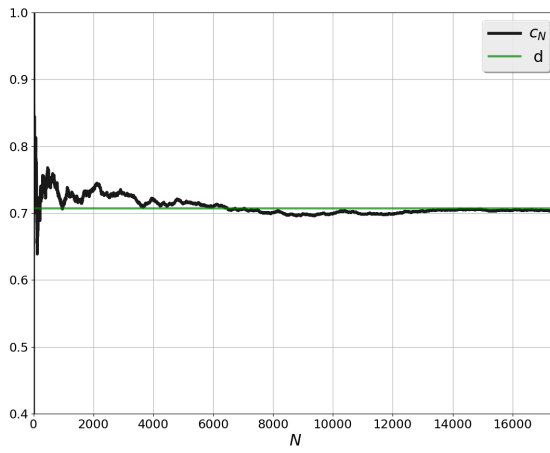
W algorytmie online (podrozdział 5.3) wartości parametrów c_N oraz $\hat{\gamma}_N$ były aktualizowane z każdym pojawieniem się nowego pomiaru wejścia-wyjścia. Na Rysunku 5.4 pokazano przebieg wyznaczania estymatora c_N na tle prawdziwej wartości przesunięcia regresji d . Wartości estymatora odpowiedzi impulsowej $\hat{\gamma}_N$ zostały pokazane na Rysunku 5.5.

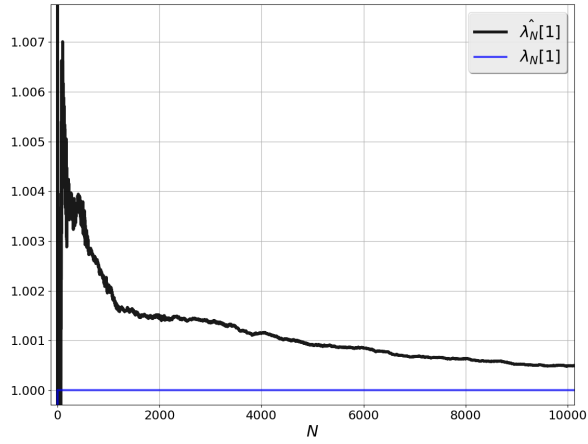
W Tabeli 5.1 przedstawiono wyniki symulacji, gdzie $\hat{\gamma}_N[0] = 1$ z uwagi na normalizację (podrozdział 5.3). Błąd został obliczony następująco $\|\hat{\gamma}_N[1] - \gamma_N[1]\|$ oraz $c_N - d$.

Na Rysunku 5.6 przedstawiono porównanie charakterystyk prawdziwej regresji, oraz estymatora. Przesunięcie tych charakterystyk względem siebie o wartość d pokazuje istotę i cel identyfikacji niniejszą metodą.

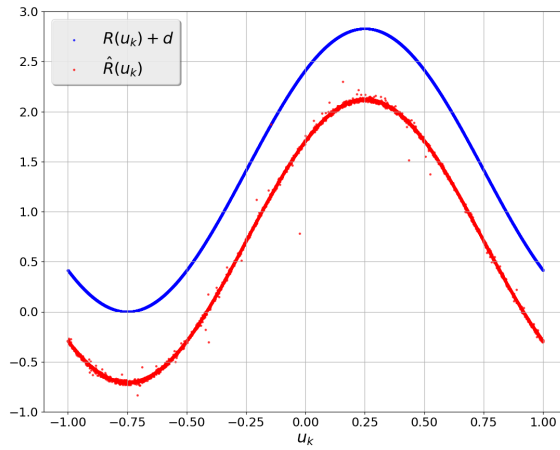
Tabela 5.1. Wyniki symulacji.

Estymator	Wartość prawdziwa	Błąd	N
c_N	$\frac{\sqrt{2}}{2}$	5×10^{-3}	10000
		5×10^{-4}	48600
$\hat{\gamma}_N[1]$	1	1×10^{-4}	6000
		7×10^{-5}	25000

**Rysunek 5.4.** Estymator c_N .



Rysunek 5.5. Zmiana estymatora $\hat{\gamma}_N[1]$.



Rysunek 5.6. Przesunięcie charakterystyki regresji.

Zastosowania praktyczne

Nieparametryczna identyfikacja systemów w zastosowaniach praktycznych

W niniejszej części przedstawione zostały propozycje rozwiązania problemów modelowania rzeczywistych systemów z zastosowaniem metod nieparametrycznych jak i wyzwania związane z zastosowaniem tych metod w problemach praktycznych. Zaproponowane algorytmy wychodzą na przeciw wyzwaniom wymienionym w punkcie 2.4. Skupiono się na dwóch realizacjach praktycznych: modelowaniu procesów termicznych (Rozdział 7) oraz modelowanie przebiegu prądu pobieranego przez maszynę produkcyjną (Rozdział 8). Oba przedstawione problemy adresują obecne i popularne w środowisku produkcyjnym zagadnienia takie jak:

- modelowanie procesów w pracy laboratoryjnej i badawczej,
- optymalizacja planowania produkcji i kontroli procesu produkcyjnego,
- przewidywanie awarii i optymalne planowanie cykli serwisowych.

Najistotniejszym czynnikiem wspólnym opisywanych metod, jest niewielka wymagana wiedza *a priori* jaką należy posiadać w odniesieniu do struktury modelowanego systemu, a szczególnie charakterystyki nieliniowej, która nawet w podejściu opartym o strukturę Hammersteina jest modelowana nieparametrycznie.

Modelowanie szkieł chalcogenidkowych w kalorymetrze różnicowym

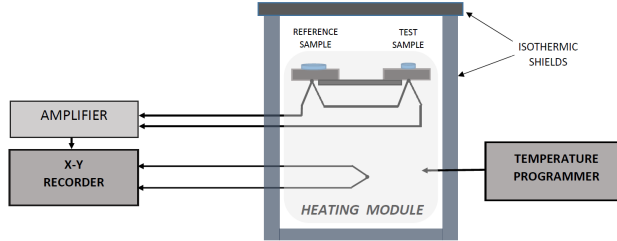
7.1 Opis problemu

Nieliniowe modele dynamiczne z powodzeniem stosuje się w badaniach zachowania nowoczesnych materiałów. Wiele z nich (np. stosowanych w elektronice) cechuje się interesującymi właściwościami, ale problemem w ich analizie jest silna nieliniowość. W takiej sytuacji, tradycyjne modele identyfikacji nie dają satysfakcjonujących rezultatów. Jako przykład, można wskazać szkiełka chalcogenidkowe. Posiadają one unikatowe własności optyczne, szczególnie w podczerwieni, co sprawia, że są one interesującym materiałem do wykorzystania w technice światłowodowej [11]. Innymi obszarami zastosowania oraz potencjalnego zastosowania mogą być: medycyna, militaria [57] [6], komunikacja optyczna [51].

Jednym z głównych problemów jakie wiążą się z używaniem szkieł chalcogenidkowych do tego typu zastosowań jest ich trwałość i zmiana parametrów w czasie eksploatacji. Problem ten jest charakterystyczny dla takich materiałów jak szkło, ceramika czy polimery. Własności fizykochemiczne oraz termodynamiczne podlegają powolnej zmianie w czasie, zależnie od temperatury materiału. Proces ten nazywa się starzeniem fizycznym [49]. Jedną z najefektywniejszych metod badania tego zjawiska jest analiza termiczna.

Najbardziej popularne metody analizy termicznej (Rys. 7.1) tego typu materiałów to DTA (eng. *Differential Thermal Analysis*) oraz DSC (eng. *Differential Scanning Calorimetry*). w metodzie DTA rejestruje się różnicę temperatur pomiędzy próbką referencyjną a próbką testowaną względem czasu lub temperatury. Metoda DSC jest w zasadzie rozwinięciem metody DTA, które polega

na pomiarze zmiany różnicy strumienia ciepłego pomiędzy próbką badaną a referencyjną. Na potrzeby badań wykonano urządzenie laboratoryjne oparte o metodę DSC [34] [35], i jako punkt wyjściowy przyjęto model Hammersteina.



Rysunek 7.1. Schemat urządzenia do pomiaru metodami DTA oraz DSC

7.2 Model Hammersteina z elementem liniowym typu ARMA przy użyciu metody predykcji interakcji

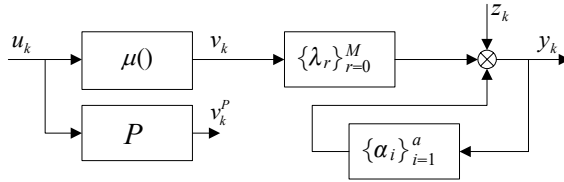
[7] [10] Bazując na obserwacjach z poprzednich prac na temat szkiełek chalcogenidkowych, proces analizy termicznej tego materiału wykazują silną zależność od poprzednich wyjść. W związku z tą zależnością zdecydowano się uogólnić prezentowany algorytm do modelu z autoregresyjnym blokiem liniowym ARMAX (Fig. 7.2) *i.e.*,

$$y_k = \sum_{r=0}^M \lambda_r v_{k-r} + \sum_{i=1}^a \alpha_i y_{k-i} + z_k. \quad (7.1)$$

Błąd aproksymacji (Tabela 7.1) zdefiniowano jako

$$Q = \frac{1}{N} \sum_{k=1}^N (y_k - y_k^P)^2. \quad (7.2)$$

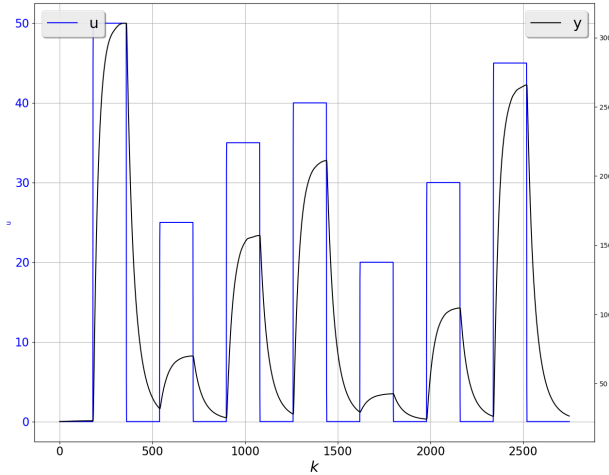
Identyfikacja i weryfikacja wyników zostały przeprowadzone na różnych zbiorach danych (Rys. 7.3). Wyniki otrzymano dla $M = 299$, $a = 1$, oraz $h = 5$. Na Rysunkach 7.4 oraz 7.5 pokazano prawdziwe wyjście systemu oraz wyjście modelu.



Rysunek 7.2. Model Hammersteina z liniowym blokiem typu ARMAX

Tabela 7.1. Błąd aproksymacji

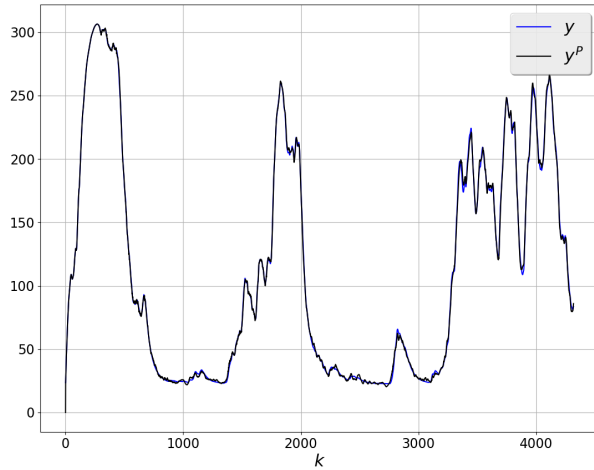
Model	Q
ARX	2.0592
ARMAX	0.6773



Rysunek 7.3. Zbiór danych weryfikacyjnych, gdzie u jest mocą grzałki a y jest temperaturą próbki

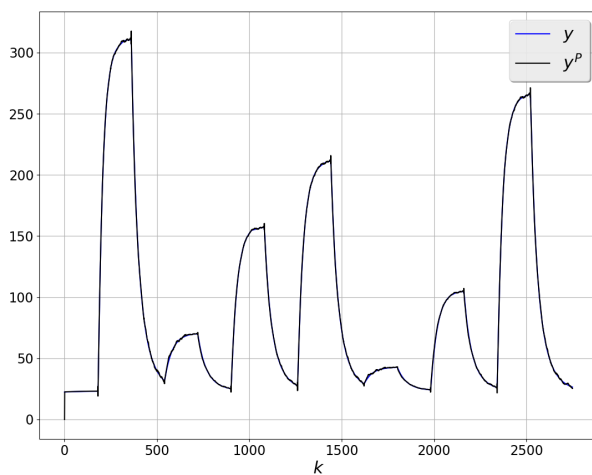
7.3 Model nieparametryczny

Oznaczmy temperaturę próbki jako wyjście systemu y_k oraz moc grzałki jako wejście u_k . Nieparametryczny model bazujący na wielowymiarowej regresji ją-

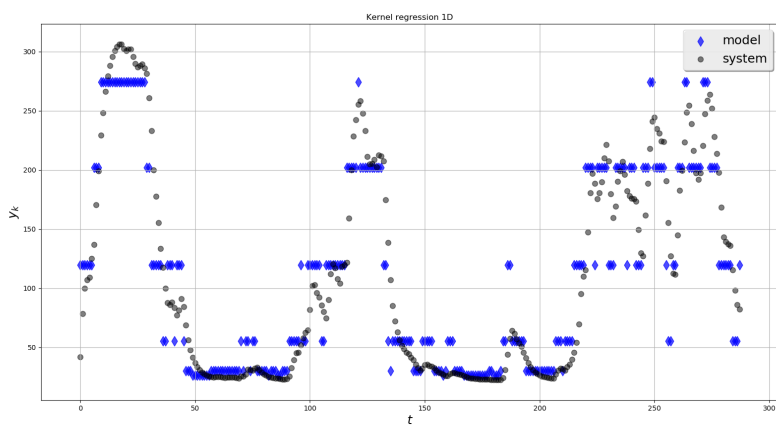


Rysunek 7.4. Identyfikacja za pomocą modelu ARMAX, prawdziwe wyjście oraz wyjście modelu

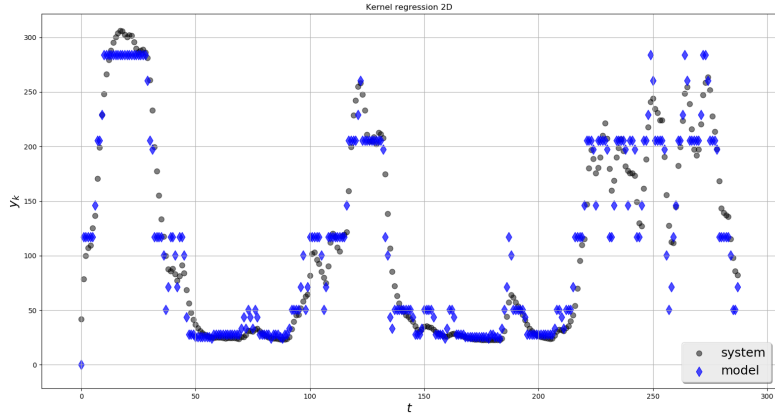
drowej $\hat{R}_s(u_k^{(s)})$ został wyznaczony dla $s = 1, 2, 3, 4$. Na Rysunkach 7.6–7.9 przedstawiono wyniki.



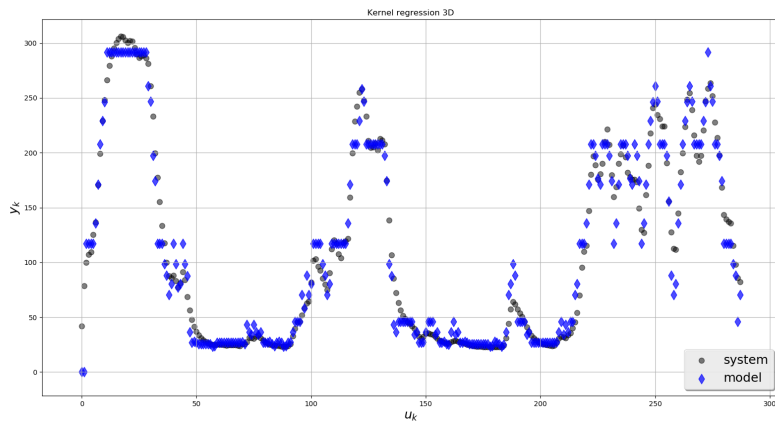
Rysunek 7.5. Weryfikacja modelu ARMAX, prawdziwe wyjście oraz wyjście modelu



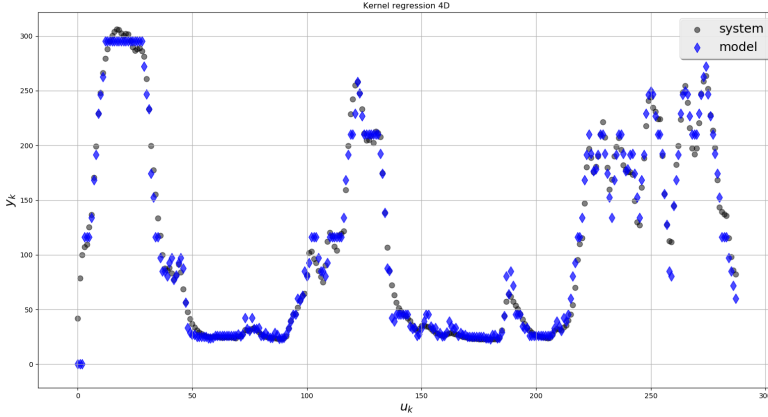
Rysunek 7.6. Eksperyment—regresja 1-wymiarowa



Rysunek 7.7. Eksperyment—regresja 2-wymiarowa



Rysunek 7.8. Eksperyment—regresja 3-wymiarowa



Rysunek 7.9. Eksperyment—regresja 4-wymiarowa

W Tabeli 7.2 pokazano błąd średniokwadratowy dla wymiarów modelu $s = 1, 2, 3, 4$. Przewagą modelu nieparametrycznego jest wyeliminowanie błędu aproksymacji dla $s(N) \rightarrow \infty$, w odróżnieniu do modeli liniowych czy Hammersteina.

Tabela 7.2. Błąd średniokwadratowy δ_s wyjścia modelu, dla $s = 1, 2, 3, 4$.

	δ_1	δ_2	δ_3	δ_4
\hat{R}_s (Prezentowana metoda)	1017	477	232	169
BLA (Liniowy FIR(s))	1710	1331	1165	1114
Hammerstein (wielomian 3-ciego rzędu + FIR(s))	1102	553	296	202

Wyniki porównano z liniowym modelem FIR(s) oraz parametrycznym modelem o strukturze Hammersteina. Warto również wspomnieć, że model Hammersteina z nieparametryczną nieliniowością zaprezentowany przez Greblickiego oraz Pawlaka w 1980s [21], w tym przypadku nie jest możliwy do zastosowania z uwagi na autokorelację wejścia. Z drugiej zaś strony, parametryczny model Hammersteina jest silnie uzależniony od arbitralnie wybranego typu funkcji bloku nieliniowego. W przedstawionym eksperymencie zastosowano wielomian trzeciego rzędu $\mu(\cdot)$ połączony z liniowym blokiem dynamicznym FIR(s). Tabela 7.1 pokazuje, że wyniki przedstawionego algorytmu, mimo zdecydowanie mniej restrykcyjnych założeń oraz bez wymaganej wiedzy *a priori*.

Modelowanie w optymalizacji pracy obrabiarek sterowanych numerycznie

Maszyny sterowane komputerowo są podstawowymi urządzeniami produkcyjnymi w niemal każdej gałęzi przemysłu. Jednym z typów tych maszyn są maszyny CNC (ang. *Computerized Numerical Control*, Rys. 8.1), czyli obrabiarki sterowane numerycznie. Można je spotkać w produkcji samochodów, maszyn i urządzeń, komponentów samolotów i przemysłu zbrojeniowego, elektroniki czy w branży medycznej. Efektywność tego typu maszyn zależy od bardzo wielu czynników, wśród których można wymienić:

- konstrukcję i sterowanie maszyny,
- używane narzędzia i ich zużycie,
- rodzaj i jakość materiału obrabianego,
- parametry procesu,
- temperaturę otoczenia,
- rodzaj i jakość chłodziwa lub jego brak,
- mocowanie obrabianej części.

Z uwagi na powyższe czynniki oraz ogólnie skomplikowaną wieloosiową konstrukcję, istnieje duży potencjał w optymalizacji pracy tych maszyn, który jest wykorzystywany zarówno przez praktyków inżynierii produkcji jak i ośrodki naukowe. Wspomnianą optymalizację można zaproponować między innymi w następujących obszarach:

- prognozowanie awarii,
- planowanie produkcji,
- planowanie przestojów serwisowych.
- dobór nastaw parametrów technologicznych (np. prędkość wrzeciona, posuw czy obroty narzędzia),

- monitorowanie i prognozowanie zużycia narzędzia.



Rysunek 8.1. Tokarko-frezarka CNC

Z uwagi na wpływ wyżej wymienionych aspektów na wartość natężenia prądu zasilania takiej maszyny postawiono hipotezę

Hipoteza 1 *Za pomocą pomiaru chwilowego natężenia prądu pobieranego przez maszynę oraz jego dynamiki, można weryfikować proces technologiczny i dysponując odpowiednim modelem, oceniać jakość procesu.*

Ze względu na wiele czynników które wpływają na zachowanie maszyny oraz w efekcie na charakterystykę poboru prądu, zdecydowano się analizować natężenie prądu w formie szeregów czasowych. Należy jednak nadmienić, że zaprezentowany tutaj algorytm nie został opisany w części teoretycznej, i mimo tego, że bazuje on na tych samych pomysłach – jest stosowany bez szczególnej analizy oraz przy braku dowodów.

Wprowadźmy ogólny opis dynamicznego systemu nieliniowego

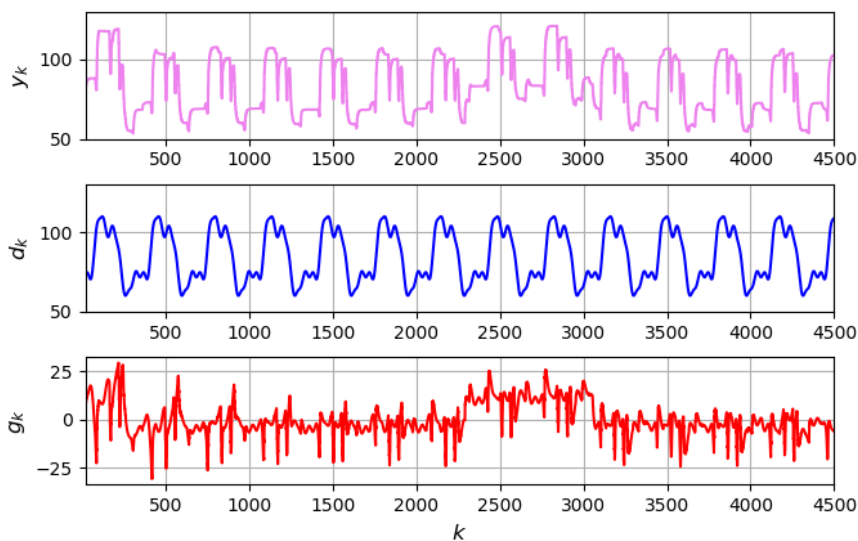
$$y_k = d_k + g_k + z_k, \quad (8.1)$$

gdzie y_k oznacza natężenie pobieranego prądu w chwili k , z_k oznaczmy jako zakłócenie. Składniki d_k oraz g_k są kolejno procesem deterministycznym oraz

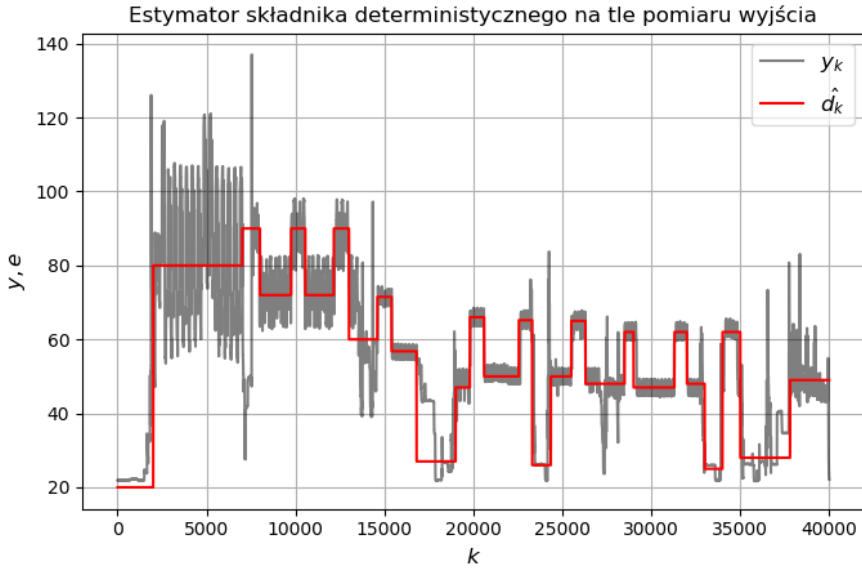
składową stochastyczną z nieliniową autoregresją, gdzie

$$g_k = \mathcal{G}(g_{k-1}, g_{k-2}, \dots, g_{k-s}). \quad (8.2)$$

Niech funkcja $\mathcal{G}()$ będzie autoregresywną częścią modelu dynamicznego, której wartość jest zależna od s poprzednich wartości tej samej funkcji. Proces deterministyczny d_k jest związany z wiedzą ekspercką o systemie i niezależnie od zastosowania może być dowolnym przybliżeniem badanego sygnału, np. prostym przesunięciem zależnym od wartości średniej sygnału (Rys. 8.3), funkcją wielomianową w dziedzinie czasu dyskretnego k bądź nawet uogólniając zaprezentowaną metodę - funkcją zależną od wejścia, np. od pomiaru odległości narzędzia skrawającego od materiału. Biorąc pod uwagę ostatni przykład i fakt trudności pomiarów tego typu wartości, można posłużyć się symulacją procesu technologicznego z programu CAM (eng. *Computer Aided Manufacturing*) z którego też może zostać wyeksportowany program NC, a zatem i wiedza o przewidywanym stanie maszyny w czasie.



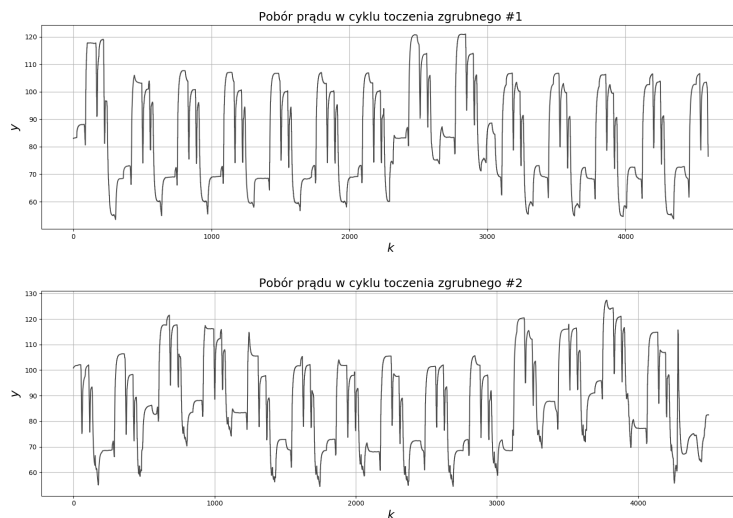
Rysunek 8.2. Rozkład procesu na składowe deterministyczną d_k oraz stochastyczną g_k



Rysunek 8.3. Składowa deterministyczna

Na rysunku 8.2 pokazano przykładowy rozkład procesu w formie szeregu czasowego na składową deterministyczną i składową stochastyczną. Wykorzystano identyfikację procesu deterministycznego w formie szeregu Furiera, proces stochastyczny g_k jest różnicą sygnału wyjściowego i d_k (8.1), który można zamodelować np. za pomocą wielowymiarowej regresji jądrowej (punkt 8.2).

Warto zauważyć, że w różnych cyklach produkcyjnych mogą występować różnice względem modelu które uznaje się za dopuszczalne, np. różna chwila złamania wióra, czyli momentu w którym w wyniku skrawania wiór zostaje odcięty. Wpływa to znacząco na siły skrawania występujące w tym czasie, w związku z czym mogą nastąpić znaczne różnice zarówno w składowej deterministycznej (chwilowe przesunięcie przebiegu). Zatem celem jest skonstruowanie takiego modelu żeby tego typu anomalie były dozwolone (Rys. 8.4), ale realne problemy pozostawały możliwe do wskazania. Wspomniane wcześniej problemy takie jak tępienie narzędzia bądź słabej jakości materiał, mogą się objawiać np. znacznie różną charakterystyką nieliniową (większe "poszarpanie" wykresu, bądź inna częstotliwość drgań sygnału) lub trendem procesu deterministycznego.



Rysunek 8.4. Pobór prądu podczas tej samej operacji technologicznej dla dwóch różnych cykli

8.1 Układ pomiarowy oraz akwizycja danych

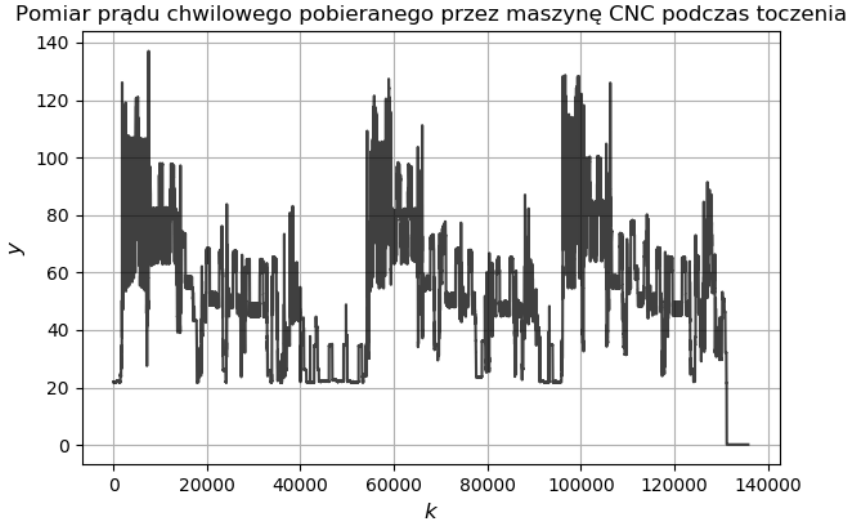
Badania eksperymentalne przeprowadzono na danych pochodzących z zakładu produkcyjnego zajmującego się produkcją części maszyn, gdzie maszyny CNC są wykorzystywane do precyzyjnej obróbki metalu i kształtowania produktu gotowego oraz półproduktów. Typ produkcji można określić jako jednostkowy, małoseryjny z cyklicznie powracającymi zleceniami. W tego typu produkcji kontrola procesu jest szczególnie trudna, ponieważ zmienność wytwarzanego produktu jest na tyle duża, że opracowanie dedykowanych metod analizy współczynników jakościowych dla każdego wariantu jest nieopłacalne.

Pomiar natężenia prądu został zrealizowany na analogowych modułach pomiarowych Beckhoff EL3433 podłączonych do sterownika PLC (ang. *Programmable Logic Controller*) siecią EtherCAT (Rys. 8.5). Pomiar natężenia prądu pobieranego przez maszyny technologiczne jest jedną z wielu funkcjonalności zbudowanego systemu. Wśród nich można wymienić badanie zaniku fazy, pomiar napięcia czy też obliczanie harmonicznych i współczynników mocy. Wszystkie dane są magazynowane w pamięci sterownika z próbkowaniem co $100ms$, a po wstępnej filtracji i obróbce część z nich jest wysyłana do serwera



Rysunek 8.5. Układ pomiarowy

zdalnego (w modelu PaaS - ang. *Platform as a Service*) za pomocą protokołu MQTT. Sterowanie maszyny bazujące na systemie Siemens SINUMERIK również wysyła dane do serwera za pomocą bramki IoT (ang. *Internet of Things*), dzięki obu źródłom danych, możliwe jest skorelowane danych pomiarowych i wydarzeń na maszynie jak np. wybór programu technologiczne lub alarm. Na Rys. 8.6 pokazano przykładowy przebieg, w którym ta sama operacja technologiczna została uruchomiona trzy razy. Każdy z cykli został uruchomiony wykorzystując te same narzędzia, program i parametry ale za każdym razem założony nowy materiał surowy.



Rysunek 8.6. Pomiar chwilowego prądu pobieranego przez maszynę podczas trzech cykli toczenia

8.2 Wielowymiarowa regresja jądrowa w modelowaniu szeregów czasowych

Zgodnie ze wzorem (8.1) niech d_k będzie procesem deterministycznym

$$d_k = \epsilon(k), \quad (8.3)$$

lub uogólniając, w przypadku dostępu do wejścia i modelu jego wpływu na ów sygnał

$$d_k = \epsilon(k, u). \quad (8.4)$$

Wprowadźmy zatem wielowymiarową regresję jądrową

$$\hat{\mathcal{G}}_s(x^{(s)}) = \frac{\sum_{k=1}^N g_k K\left(\frac{\|\phi_k^{(s)} - x^{(s)}\|}{h}\right)}{\sum_{k=1}^N K\left(\frac{\|\phi_k^{(s)} - x^{(s)}\|}{h}\right)}, \quad (8.5)$$

którą wykorzystano do estymowania nieliniowej charakterystyki autoregresyjnej będącej składową modelu systemu. Wektor autoregresji będzie zatem wyglądał następująco

$$\phi_k^{(s)} = (g_{k-1}, g_{k-2}, \dots, g_{k-s})^T. \quad (8.6)$$

8.3 Przykład detekcji niezgodności procesu technologicznego

Rozważmy przykład monitorowania operacji skrawania w kolejnych cyklach produkcyjnych. Model został nauczony na pierwszej operacji po zweryfikowaniu poprawności procesu.

W pierwszej kolejności z uwagi na okresowy charakter przebiegu, wyznaczono składową deterministyczną w postaci aproksymatora z wykorzystaniem szeregu Fouriera rzędu $S = 7$

$$\widehat{d}_k = a_0 + \sum_{k=1}^S a_n \cos(k\pi) + b_n \sin(k\pi), \quad (8.7)$$

gdzie wektory a oraz b są wektorami parametrów. Aby ułatwić identyfikację, wykorzystano wiedzę *a priori* dotyczącą częstotliwości ruchu narzędzia skrawającego. Estymator \widehat{d}_k co ważne nie pełni roli porównania modelu z aktualnym wyjściem, tylko zdejmuje charakterystykę deterministycznej z modelu, tak aby porównać jego cechy charakterystyczne. Dlatego, estymator \widehat{d}_k dla każdego z przebiegów jest dopasowywany indywidualnie, przy zachowaniu metody (np. szereg Fouriera) oraz kluczowych parametrów (rząd szeregu). W praktyce można zastosować modelowanie przedziałami w czasie rzeczywistym pracy maszyny, aktualizując składową \widehat{d}_k w oknie określonej ilości pomiarów wyjścia.

Następnie zastosowano regresję jądrową (8.5) w celu wyznaczenia składowej stochastycznej

$$\widehat{g}_k = y_k - \widehat{d}_k. \quad (8.8)$$

Ostatecznie model referencyjny sygnału wyjściowego przyjmuje postać

$$\widehat{y}_k = \widehat{d}_k + \widehat{g}_k. \quad (8.9)$$

Nauczono trzy modele, odpowiednio dla $s = 1$, $s = 2$ oraz $s = 3$. Model zweryfikowano na danych testowych pochodzących z jednego z kolejnych cykli

produkcyjnych, który został zweryfikowany jako poprawny.

Jako wskaźniki błędów zastosowano błąd typu MAE (ang. *Mean Absolute Error*)

$$MAE_s = \frac{1}{N} \sum_{k=1}^N |\hat{y}_k - y_k|, \quad (8.10)$$

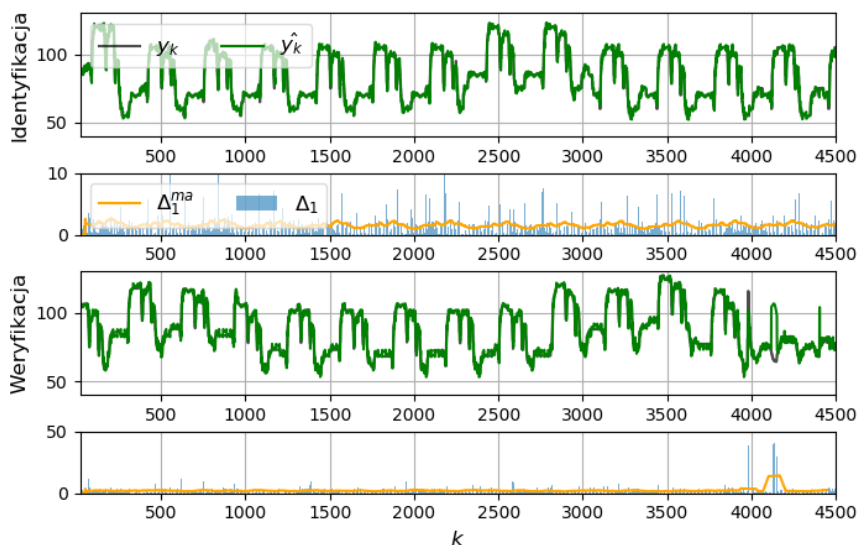
oraz błąd średniokwadratowy ze średniej kroczącej błędów chwilowych $\Delta_s(k) = |\hat{y}_k - y_k|$ postaci

$$MSE_s^{ma} = \frac{1}{N} \sum_{n=w}^n \left(\frac{1}{w} \sum_{i=N-w+1}^n \Delta_s(i) \right)^2, \quad (8.11)$$

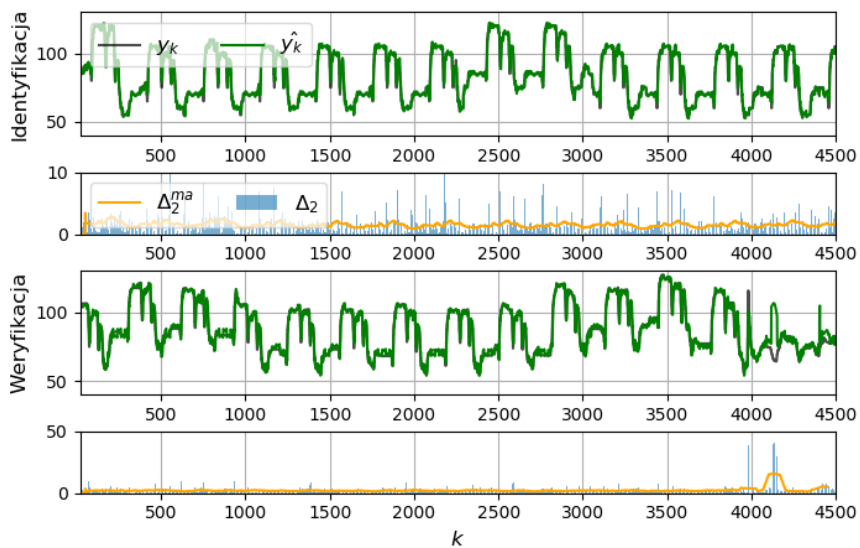
gdzie przebieg średniej kroczącej błędów chwilowych wynosi

$$\Delta_s^{ma}(n) = \frac{1}{w} \sum_{i=n-w+1}^n \Delta_s(i). \quad (8.12)$$

Na Rys. 8.7, 8.8 oraz 8.9 Przedstawiono sygnał prawdziwy, estymator oraz błędy chwilowe oraz średnią krocząca błędów chwilowych dla okna $w = 7$ oraz $N = 4600$. Każdy z eksperymentów wykonano zarówno na zbiorze uczącym jak i zbiorze testującym. Wyniki przedstawiono w Tab. 8.1 oraz 8.1.



Rysunek 8.7. Uczenie modelu $s = 1$



Rysunek 8.8. Uczenie modelu $s = 2$

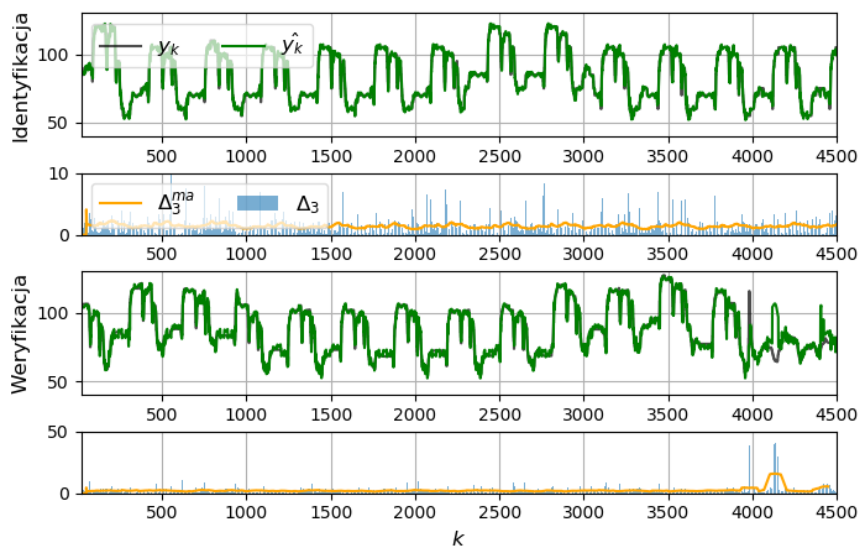
Rysunek 8.9. Uczenie modelu $s = 3$

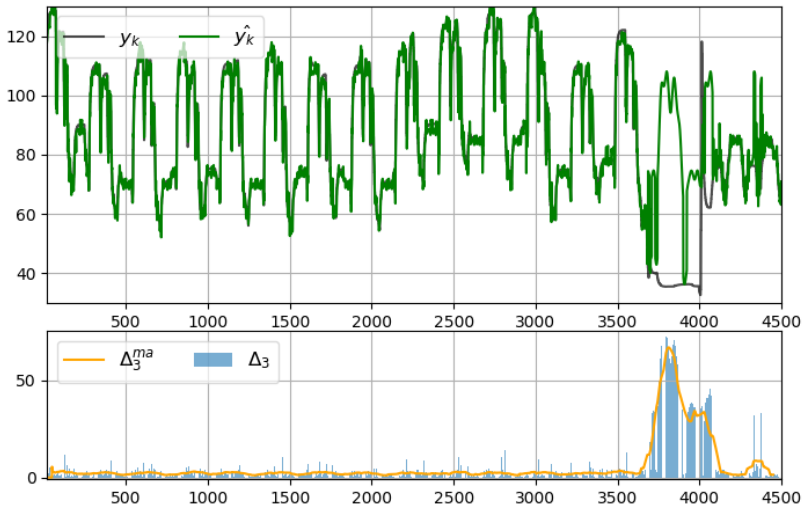
Tabela 8.1. Porównanie błędów podczas nauki modelu

	MAE_s	MSE_s^{ma}
$s = 1$	1,91	4,06
$s = 2$	1,85	3,84
$s = 3$	1,79	3,45

Tabela 8.2. Porównanie błędów podczas weryfikacji modelu

	MAE_s	MSE_s^{ma}
$s = 1$	2,40	13,66
$s = 2$	2,47	15,09
$s = 3$	2,49	15,48

Następnie zastosowano model nauczony na danych zawierających problem technologiczny, co przedstawiono na Rys. 8.10. W Tab. 8.3 pokazano wyniki eksperymentu. Jak widać, przy analizie przebiegu całego cyklu, udało się wykryć problem. Zarówno wskaźnik MSE_3^{ma} jak i MAE_3 wskazał wartość znacznie wyższą, od tej uzyskanej na próbie poprawnego przebiegu procesu. Dzięki czułości wskaźnika MSE_3^{ma} oraz funkcji przebiegu błędu $\Delta_3^{ma}(k)$ można precyzyjnie określić lokalizację błędu stosując próg wartości $\Delta_3^{ma}(k)$ przy której aktywowany jest np. alarm na maszynie, zdarzenie w systemie zarządzania produkcją bądź też następuje automatyczna korekcja parametrów technologicznych w zależności od wartości maksymalnej błędu.



Rysunek 8.10. Eksperyment wykrycia awarii

Tabela 8.3. Eksperyment wykrycia awarii

	MAE_s	MSE_s^{ma}
$s = 3$	5,68	180,48

8.4 Porównanie z innymi metodami uczenia maszynowego

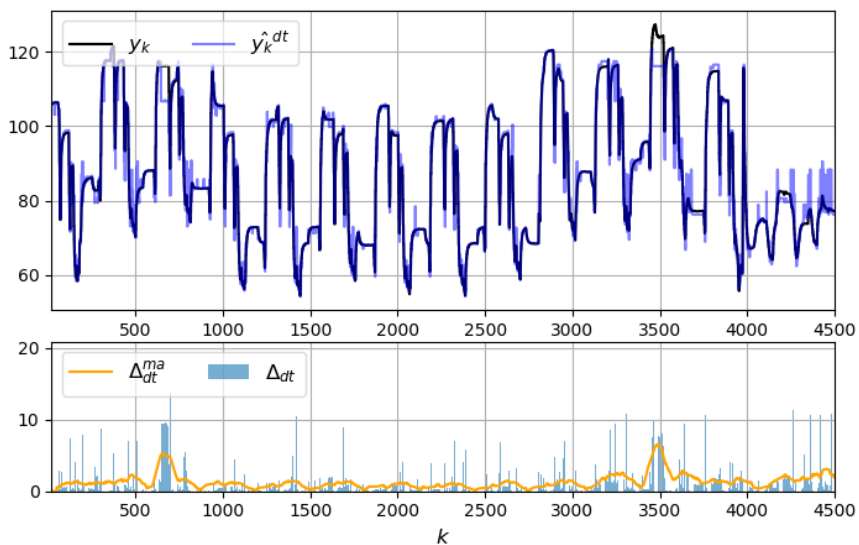
W celu zestawienia proponowanej metody z gotowymi algorytmami uczenia maszynowego używanymi z powodzeniem w problemach praktycznych wybrano dwa niezależne podejścia: drzewo decyzyjne regresji (punkt 8.4.1) oraz sztuczną sieć neuronową - perceptron wielowarstwowy (ang. *Multilayer Perceptron* - MLP, punkt 8.4.2). Celem zadania było zbudowanie takiego modelu, aby przy zastosowaniu wskaźnika (8.11) różnica weryfikacji na danych poprawnych oraz danych z awarią była jak największa.

Do obu eksperymentów wykorzystano bibliotekę języka *Python* - *scikit-learn* a dokładniej klasy *DecisionTreeRegressor* i *MLPRegressor*. Aby wyniki były porównywalne we wszystkich przykładach zastosowano $s = 3$ czyli na podstawie 3 ostatnich pomiarów model wykonywał predykcję kolejnego. Uczenie było przeprowadzane dla różnych zestawień parametrów tak, aby błąd średniokwadratowy predykcji od danych uczących był jak najmniejszy.

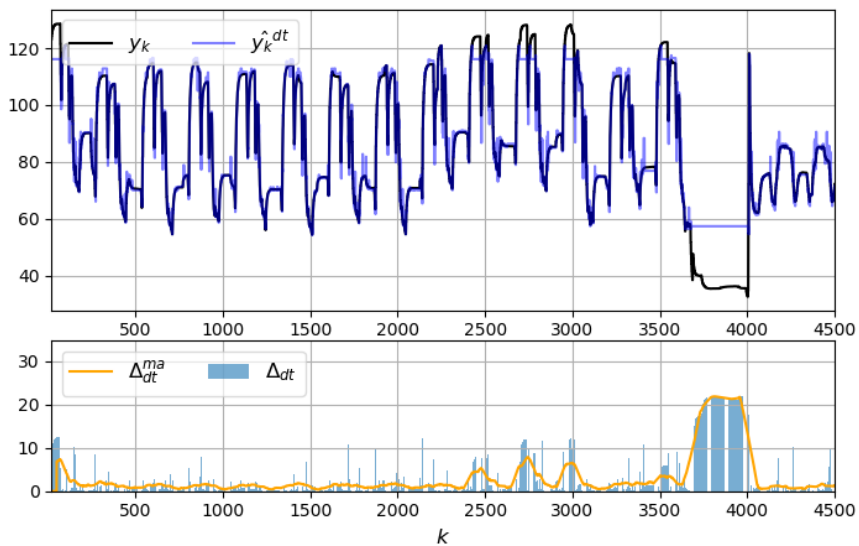
8.4.1 Drzewo regresji

Prezentowany algorytm uczenia maszynowego jest wariantem algorytmu oryginalnie zwanego drzewem decyzyjnym służącym do dzielenia danych na kategorię, w którym zamiast kategorii następuje uśrednianie wyniku w zależności do której części zbioru parametrów, badany parametr przynależy.

Na Rys. 8.11 oraz 8.12 pokazano eksperymenty weryfikacji modelu w procesie zgodnym oraz w procesie niezgodnym. Wyniki podsumowano w Tabeli 8.4. Analogicznie do wzoru (8.12) pokazano przebiegi chwilowe błędów Δ_{dt} oraz ich średnią krocząca Δ_{dt}^{ma} .



Rysunek 8.11. Weryfikacja drzewa regresji na danych z procesu zgodnego



Rysunek 8.12. Weryfikacja drzewa regresji na danych z procesu niezgodnego

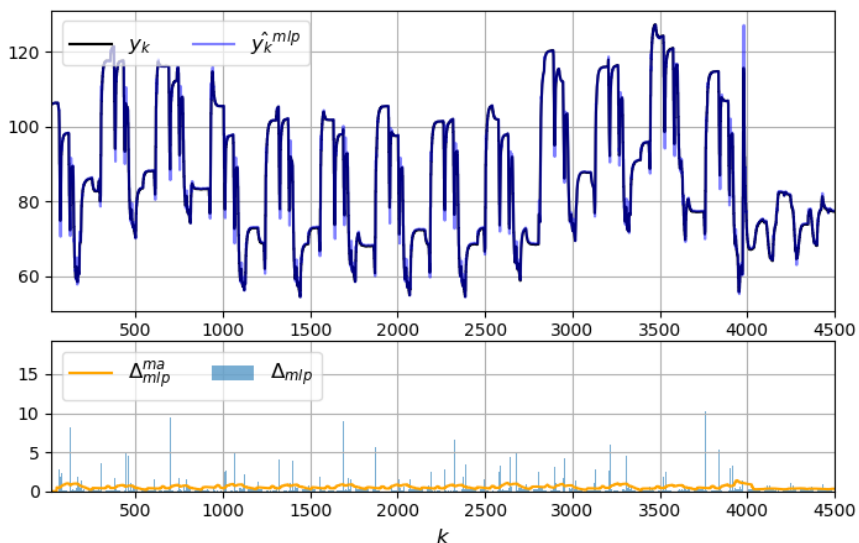
Tabela 8.4. Weryfikacja drzewa regresji

	MAE_3	MSE_3^{ma}
Proces zgodny (OK)	1,34	2,78
Proces niezgodny (NOK)	3,12	33,46
Stosunek $\frac{NOK}{OK}$	2,33	12,04

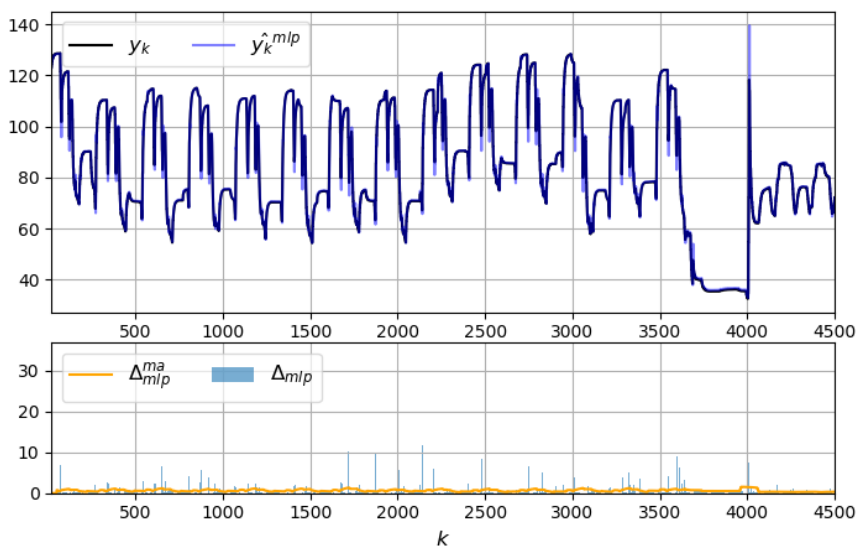
Jak widać na Rys. 8.12 drzewo regresji stosunkowo sprawnie wykrywało chwilowe odchylenia procesu produkcyjnego. Analiza wskaźnika jakości MSE_3^{ma} dla całego przebiegu wykazała stosunek procesu niezgodnego do zgodnego na poziomie **12,04**, co można uznać za wynik satysfakcjonujący dla tego konkretnego scenariusza, ponieważ model wykrył bez problemu niezgodność procesu.

8.4.2 Sztuczna sieć neuronowa

W celu porównania ze sztuczną siecią neuronową, wykorzystano perceptron wielowarstwowy o liczbie warstw ukrytych $n = 40$, co oznacza, że mamy do czynienia ze złożeniem ze sobą n funkcji aktywacji (wyjście jednej jest argumentem drugiej itd.). Jest to powszechnie używany algorytm zaliczany do sztucznych sieci neuronowych, gdzie jest warstwa wejściowa (w tym przypadku trzy wejścia - będące ostatnimi trzema pomiarami wyjść szeregu czasowego), warstwa wyjściowa (predykcja jednej, kolejnej wartości) oraz n warstw ukrytych. Jako funkcję aktywacji wybrano jednostronnie obciętą funkcję liniową $f(x) = \max(0, x)$. Na Rys. 8.13, 8.14 oraz w Tab. 8.5 pokazano weryfikację modelu w procesie zgodnym oraz w procesie niezgodnym. Analogicznie do wzoru (8.12) pokazano przebiegi chwilowe błędów Δ_{mlp} oraz ich średnią krocząca Δ_{mlp}^{ma} .



Rysunek 8.13. Weryfikacja algorytmu MLP na danych z procesu zgodnego



Rysunek 8.14. Weryfikacja algorytmu MLP na danych z procesu niezgodnego

Tabela 8.5. Weryfikacja sztucznej sieci neuronowej

	MAE_3	MSE_3^{ma}
Proces zgodny (OK)	0,53	0,33
Proces niezgodny (NOK)	0,65	0,49
Stosunek $\frac{NOK}{OK}$	1,23	1,48

Mimo, że dla algorytmu *MLP* błąd średniokwadratowy dopasowania modelu do danych uczących wyszedł najmniejszy ze wszystkich trzech algorytmów, to wykrycie procesu niezgodnego przy zastosowanych parametrach nie byłoby możliwe.

8.4.3 Podsumowanie wyników

Przed analizą ilościową prezentowanego porównania należy zaznaczyć, że przewagą niniejszej metody nad drzewem decyzyjnym oraz sztuczną siecią neuronową jest istnienie wzoru analitycznego modelu. W związku z tym, możliwa jest interpretacja procesowa (lub nawet fizyczna) nauczonego modelu, co daje dużo więcej możliwości niż tylko wykrywanie określonych sytuacji. Przykładami takich interpretacji mogą być częstotliwość, moc, przesunięcie fazowe czy wariancja zakłóceń.

Jak pokazano w Tabeli 8.6 skuteczność wykrywania procesu niezgodnego dla drzewa regresji i zaproponowanego algorytmu wyszła bardzo porównywalna. Warto jednak zaznaczyć, że eksperyment dotyczył całej operacji technologicznej, fragment niezgodny występujący pomiędzy próbką $k = 3500$ a $k = 4100$ w zaproponowanym algorytmie jest znacząco uwydatniony przebiegiem błędu chwilowego.

Tabela 8.6. Porównanie algorytmów

	Stosunek $\frac{NOK}{OK}$	MAE_3	MSE_3^{ma}
Zaproponowany algorytm		2,28	11,66
Drzewo regresji		2,33	12,04
MLP		1,23	1,48

Aplikacja pokazanego przykładu, w kontekście inżynierskim może być wykonana na wiele różnych sposobów np. monitorowanie "okna" w którym obliczany jest wybrany wskaźnik jakościowy i porównywany z zakresem OK/NOK, wykrywanie przekroczenia limitu wartości błędu chwilowego, czy też weryfikacja całej operacji po jej zakończeniu za pomocą wybranego wskaźnika.

Podsumowanie

9.1 Rezultaty badań i ich następstwa

Podsumowując, w rozprawie zebrano oraz zaproponowano metody identyfikacji, które mimo swojej formalnej natury (Część II), mogą być zastosowane do rzeczywistych aplikacji (Część III). Głównie dlatego, że skupiają się na usunięciu pewnych ograniczeń narzucanych przez metody klasyczne (podrozdział 2.4). Dlatego też, jest to swoisty kompromis pomiędzy nauką a inżynierią. Oryginalne rozwiązania można wypunktować jako:

- zaproponowano aktywny eksperyment opierający się na sekwencjach wejść (binarnych, lub na siatce w przestrzeni wielowymiarowej), które umożliwiają zaślepienie wejściowej nieliniowości w system Hammersteina-Wienera (Rozdział 3), lub pozwalają na redukcję wymiaru w regresji wielowymiarowej (podrozdział 4.4);
- opracowano metodę identyfikacji systemów o ogólnej strukturze, gdzie ze stosunkowo słabymi założeniami system może być identyfikowany bez, lub z ograniczoną wiedzą *a priori*, dzięki czemu znacząco jest zredukowane ryzyko wyboru niewłaściwej struktury modelu (Rozdział 4);
- wykazano wpływ redukcji wymiaru na poprawę szybkości zmniejszania się błędów estymatora (Twierdzenie 4.5.1);
- uogólniono algorytm predykcji interakcji do systemu autoregresyjnego AR-MAX (podrozdział 5.4);
- zaprezentowano opisywane idee w symulacjach oraz w zastosowaniach praktycznych (Część III), ze szczególnym uwzględnieniem systemu o charakterze szeregu czasowego, gdzie opracowano metodę dekompozycji systemu na składową deterministyczną i losową (Rozdział 8).

9.2 Problemy otwarte

Z uwagi na poruszone w niniejszej pracy kwestie, można wskazać następujące problemy otwarte oraz możliwe kierunki dalszych badań:

- rozwijanie możliwości jakie dają aktywne eksperymenty (kwantyzacja, okresowość itp.) w taki sposób aby osiągnąć kompromis pomiędzy wymiarem problemu (czyli jakością identyfikacji charakteru dynamicznego systemu) a zagęszczeniem punktów estymacji, w celu uniknięcia ryzyka zbytniego przefiltrowania sygnału, co może grozić utratą ważnych informacji o jego charakterze;
- kontynuacja pracy nad dekompozycją systemów o charakterze autoregresyjnym, aby uzyskać odpowiedni kompromis pomiędzy składowymi deterministycznymi i losowymi, tak aby model został nauczony odróżniać sytuację poprawne od błędnych;
- opracowanie metod interpretacji modeli bazujących na podejściu z Rozdziału 8, tak aby mimo braku wzoru, mogły przynieść pewne ogólne informacje o konkretnych systemach, wykorzystywanych w zastosowaniach praktycznych (np. zaprezentowany pomiar prądu i cechy elektrotechniczne maszyny, analiza drgań, analiza EEG i inne zastosowania medyczne, charakterystyki sygnałów telekomunikacyjnych itp.);
- zastosowywanie metod identyfikacji parametryczno-nieparametrycznej w celu identyfikacji złożonych systemów takich jak np. system Hammersteina-Wienera.

Dodatki

A

Dowody twierdzeń i wyprowadzenia techniczne

A.1 Dowód Twierdzenia 3.4.1

Dla założeń 1-4 z punktu 3.2, $c = \mu'(u_0)\eta'(x_0) \neq 0$ oraz $\alpha \in (0, \frac{1}{d})$, $d = q + 3$, weźmy rozwinięcie Taylora i zastosujmy je dla wejściowej i wyjściowej charakterystyki nieliniowej wokół punktów u_0 oraz x_0

$$\mu(u_k) = \mu(u_0) + c_1(u_k - u_0) + \rho(u_k), \quad (\text{A.1})$$

$$\eta(x_k) = \eta(x_0) + c_2(x_k - x_0) + \xi(x_k). \quad (\text{A.2})$$

Parametry c_1 oraz c_2 są stałymi różnymi od zera, o wartości pierwszych pochodnych funkcji $\mu(\cdot)$ oraz $\eta(\cdot)$ w punktach u_0 oraz x_0 . Ze względu na to, że $|\rho(u_k)| = o(h)$ oraz $|\xi(x_k)| = o(h)$, Twierdzenie można udowodnić analogicznie do Twierdzenia 3 w [41].

A.2 Dowód Twierdzenia 3.4.2

Niech założenia 1-4 z 3.1 pozostaną w mocy. Dla każdego punktu $x_{[i]}$ tworzącego siatkę

$$\begin{aligned}
E[\hat{\eta}(x_{[i]})] &= E\left[\frac{\sum_{k=1}^N y_k \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)} \Big| L > 0\right] P \\
&= E\left[\frac{\sum_{k=1}^N (v_k + z_k) \cdot \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)} \Big| L > 0\right] P \\
&= E\left[\frac{\sum_{k=1}^N \eta(x_{[i]}) \cdot \delta(\phi_k, \varphi_i) + z_k \cdot \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)} \Big| L > 0\right] P \quad (\text{A.3}) \\
&= E\left[\eta(x_{[i]}) + \frac{\sum_{k=1}^N z_k \cdot \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)} \Big| L > 0\right] P \\
&= \eta(x_{[i]}) + E\left[\frac{\sum_{k=1}^N z_k \cdot \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)} \Big| L > 0\right] P.
\end{aligned}$$

Niech $y_{(1)}, y_{(2)}, \dots, y_{(L)}$ będzie pomiarem wyjścia z (3.11). Liczba pomiarów L jest losowa, ale jej wartość oczekiwania zbiega do nieskończoności wraz ze wzrostem pomiarów

$$EL = P(\phi_k = \varphi_i) \cdot N = \frac{1}{N_0} \cdot N = \frac{N}{2^{q+1}}, \quad (\text{A.4})$$

$$P = P(L > 0) \rightarrow 1, \text{ dla } N \rightarrow \infty. \quad (\text{A.5})$$

Ostatecznie, bazując na tożsamości Walda, obciążenie estymatora asymptotycznie zbiega do zera

$$\begin{aligned}
E\left[\frac{\sum_{k=1}^N z_k \cdot \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] &= \frac{E\left[\sum_{k(l)=1}^L z_{k(l)}\right]}{EL} \\
&= \frac{EL \cdot E z_1}{EL} = 0.
\end{aligned} \quad (\text{A.6})$$

Wariancja pomiarów wyjścia zależy tylko od wariancji zakłócenia i wynosi

$$\begin{aligned}
\text{var}[y_{(l)}] &= \text{var}[\eta(x_{[i]}) + z_{(l)}] \\
&= \text{var}[\eta(x_{[i]})] + \text{var}[z_{(l)}] = \sigma_z^2.
\end{aligned} \quad (\text{A.7})$$

Zakładając, że losowa liczba pomiarów jest większa od zera, wariancja warunkowa estymatora przyjmie postać

$$\begin{aligned}
 \text{var} [\hat{\eta}(x_{[i]})] &= \sum_{k=1}^N P(L = k) \cdot \text{var} \left[\frac{1}{k} \sum_{l=1}^k y_{(l)} \right] \\
 &= \sum_{k=1}^N P(L = k) \cdot \frac{\text{var} [y_{(1)}]}{k} \\
 &= \sum_{k=1}^N P(L = k) \cdot \frac{\sigma_z^2}{k} = \frac{N_0 \sigma_z^2}{N} = \frac{c}{N} \sim N^{-1}.
 \end{aligned} \tag{A.8}$$

Ostatecznie, z (A.6) oraz (A.8), zbieżność asymptotyczna estymatora zostaje udowodniona [37][41]. Biorąc pod uwagę powyższe, Twierdzenie pozostaje w mocy, w każdym estymowanym punkcie $x_{[i]} = \varphi_i^T \gamma, i = 1, 2, \dots, N_0$ takim, że $x_{[i]} \in \text{cont}(\eta(), g())$, gdzie $\text{cont}(\eta(), g())$ jest zbiorem wszystkich punktów na $\eta()$ oraz $g()$.

A.3 Dowód Twierdzenia 4.5.1

Zdefiniujmy obciążenie i wariancję jako

$$\text{bias} \hat{R}_s(u_k^{(s)}) \triangleq E \hat{R}_s(u_k^{(s)}) - R_s(u_k^{(s)}), \tag{A.9}$$

$$\text{var} \hat{R}_s(u_k^{(s)}) \triangleq E \left\{ \hat{R}_s(u_k^{(s)}) - E \hat{R}_s(u_k^{(s)}) \right\}^2. \tag{A.10}$$

Z uwagi na spełnienie warunku Lipschitza oraz faktu, iż $\left\| \mathbf{u}_k^{(s)} - \mathbf{x}_{[p]}^{(s)} \right\| \leq h$ można zapisać

$$\text{bias} \hat{R}_s(\mathbf{x}^{(s)}) = o(h(N)), \tag{A.11}$$

$$\text{bias}^2 \hat{R}_s(\mathbf{x}^{(s)}) = o(h^2(N)), \tag{A.12}$$

$$\text{var} \hat{R}_s(\mathbf{x}^{(s)}) = o\left(\frac{1}{Nh^s(N)}\right). \tag{A.13}$$

Dla jądra w postaci funkcji okna oraz funkcjonau \mathcal{F} spełniającego warunek Lipschitza prawdopodobieństwo selekcji s -wymiarowej przestrzeni, może zostać wyznaczone za pomocą ch^s , gdzie c jest pewną stałą. W celu osiągnięcia zbieżności $\hat{R}_s(\mathbf{x}^{(s)}) \rightarrow R_s(\mathbf{x}^{(s)})$, dla $N \rightarrow \infty$, muszą zostać spełnione poniższe warunki

$$h \rightarrow 0 \text{ oraz } Nh^s \rightarrow \infty, \text{ dla } N \rightarrow \infty, \tag{A.14}$$

wówczas

$$h(N) \sim N^{-\alpha}, \text{ dla } \alpha \in \left(0, \frac{1}{s}\right). \quad (\text{A.15})$$

Aby uzyskać jak najlepszy kompromis pomiędzy kwadratem obciążenia oraz wariancją uzyskujemy

$$h^2(N) = \frac{1}{Nh^s(N)}, \quad (\text{A.16})$$

$$h^{s+2}(N) = \frac{1}{N}, \quad (\text{A.17})$$

$$h_{opt}(N) = o\left(N^{-\frac{1}{s+2}}\right), \quad (\text{A.18})$$

$$MSE = o\left(N^{-\frac{2}{s+2}}\right). \quad (\text{A.19})$$

A.4 Uwaga do Twierdzenia 4.2

Spełnienie (4.25) oraz nieliniowa optymalizacja w (4.29) są ściśle zależne od charakteru problemu. W aktywnym eksperymencie, gdzie wejście może być dowolnie wygenerowane, odpowiedni wybór punktów $\left\{x_{[p]}^{(s)}\right\}_{p=1}^{N_0}$ może znacząco uprościć operacje wykonywane w kroku 2 (przykład poniżej).

A.5 Przykład do Twierdzenia 4.2

Dla systemu

$$y_k = e^{\theta_1 u_k} + \theta_2 u_k u_{k-1} + z_k, \quad (\text{A.20})$$

w kroku 1 możemy przyjąć dwuwymiarową ($s = 2$) funkcję regresji $\widehat{R}_2(x^{(2)})$ w $N_0 = 2$ punktach $x_{[1]}^{(2)} = (1, 0)^T$, oraz $x_{[2]}^{(2)} = (1, 1)^T$, tzn.

$$\widehat{R}_2\left(x_{[1]}^{(2)}\right), \quad \widehat{R}_2\left(x_{[2]}^{(2)}\right). \quad (\text{A.21})$$

Ponieważ rzeczywiste wartości funkcji regresji wynoszą odpowiednio $R_2\left(x_{[1]}^{(2)}\right) = e^{\theta_1}$ oraz $R_2\left(x_{[2]}^{(2)}\right) = e^{\theta_1} + \theta_2$, w kroku 2 uzyskujemy trywialny estymator parametrów

$$\widehat{\theta}_1 = \log \widehat{R}_2\left(x_{[1]}^{(2)}\right), \quad (\text{A.22})$$

$$\widehat{\theta}_2 = \widehat{R}_2\left(x_{[2]}^{(2)}\right) - \widehat{R}_2\left(x_{[1]}^{(2)}\right). \quad (\text{A.23})$$

A.6 Linearyzacja lokalna w algorytmie IPM (Rozdział 5)

Dla kryterium (5.5)

$$\begin{aligned} W_N^P(c) &\triangleq V_N^P(c) \left[V_N^{P^T}(c) V_N^P(c) \right]^{-1} V_N^{P^T}(c) = \\ &= (P_N - c \mathbf{1}_{N \times (M+1)}) \left[V_N^{P^T}(c) V_N^P(c) \right]^{-1} (P_N - c \mathbf{1}_{N \times (M+1)})^T, \end{aligned}$$

zatem $V_N^{P^T}(c) V_N^P(c) = NT_N$, gdzie

$$T_N = \frac{1}{N} P_N^T P_N - c \frac{1}{N} P_N^T \mathbf{1}_{N \times (M+1)} + \frac{1}{N} (P_N^T \mathbf{1}_{N \times (M+1)})^T + c^2 \mathbf{1}_{(M+1) \times (M+1)}.$$

Zapiszmy $T_N^{-1} = (A + B)^{-1}$, gdzie

$$\begin{aligned} A &= \frac{1}{N} P_N^T P_N, \\ B &= c^2 \mathbf{1}_{(M+1) \times (M+1)} - c \left[\frac{1}{N} P_N^T \mathbf{1}_{N \times (M+1)} + \frac{1}{N} (P_N^T \mathbf{1}_{N \times (M+1)})^T \right], \end{aligned}$$

stosując linearyzację $T_N^{-1} \simeq (I - A^{-1}B) A^{-1}$, uzyskujemy

$$\frac{1}{N} P_N T_N^{-1} P_N^T = \frac{1}{N} P_N A^{-1} P_N^T - \frac{1}{N} P_N (A^{-1}B) A^{-1} P_N^T,$$

zatem (dla $B = B(c)$) kryterium można sprowadzić do

$$\begin{aligned} Q_N(c) &= Y_N^T Y_N - \frac{1}{N} Y_N^T (P_N A^{-1} P_N^T) Y_N \\ &\quad + \frac{1}{N} Y_N^T (P_N A^{-1}) B(c) (P_N A^{-1})^T Y_N. \end{aligned}$$

Stąd,

$$\frac{dQ_N(c)}{dc} = \frac{1}{N} Y_N^T (P_N A^{-1}) \frac{dB(c)}{dc} (P_N A^{-1})^T Y_N,$$

gdzie

$$\frac{dB(c)}{dc} = 2c \mathbf{1}_{(M+1) \times (M+1)} - \left(\frac{1}{N} P_N^T \mathbf{1}_{N \times (M+1)} + \frac{1}{N} (P_N^T \mathbf{1}_{N \times (M+1)})^T \right),$$

oraz $\frac{dQ_N(c)}{dc} = 0$ z uwagi na wypukłość $Q_N(c)$.

Pełne wyprowadzenie znajduje się w [58].

B

Dodatki, twierdzenia, lematy, definicje

B.1 Definicje typów zbieżności

Definicja B.1.1 [9] Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \rightarrow \infty$ zbieżny z prawdopodobieństwem 1 (mocno) do \varkappa^* jeśli zachodzi

$$P(\lim_{k \rightarrow \infty} \varkappa_k = \varkappa^*) = 1$$

Definicja B.1.2 [9] Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \rightarrow \infty$ zbieżny według prawdopodobieństwa (słabo) do $\varkappa^\#$ jeśli dla każdego $\varepsilon > 0$ zachodzi

$$\lim_{k \rightarrow \infty} P(|\varkappa_k - \varkappa^\#| > \varepsilon) = 0$$

Wartość $\varkappa^\#$ nazywamy granicą stochastyczną ciągu $\{\varkappa_k\}$ i zapisujemy

$$P \lim_{k \rightarrow \infty} \varkappa_k = \varkappa^\# \tag{B.1}$$

Zapis $P \lim_{N \rightarrow \infty} \mathbf{X}_N = \mathbf{X}$ dla sekwencji wektorów losowych $\{\mathbf{X}_N\}$, oznacza, że $\mathbf{X}_N \rightarrow \mathbf{X}$ według prawdopodobieństwa, gdy $N \rightarrow \infty$.

Lemat B.1 Ze zbieżności z prawdopodobieństwem 1 wynika zbieżność według prawdopodobieństwa.

Definicja B.1.3 [9] Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \rightarrow \infty$ zbieżny według średniej z kwadratem do \varkappa^* jeśli zachodzi

$$\lim_{k \rightarrow \infty} \mathbf{E}(\varkappa_k - \varkappa^*)^2 = 0$$

Definicja B.1.4 [19] Ciąg zmiennych losowych $\{\varkappa_k\}$ ma **szybkość zbieżności rzędu** $O(e_k)$ **według prawdopodobieństwa** przy $k \rightarrow \infty$ (tj. asymptotycznie), gdzie $\{e_k\}$ jest ciągiem liczb dodatnich zbieżnym do zera, tzn.

$$\varkappa_k = O(e_k) \text{ według prawdopodobieństwa}$$

wtedy i tylko wtedy, gdy $\left\{\frac{\varkappa_k}{e_k} \chi_k\right\}$ jest zbieżny według prawdopodobieństwa do zera dla każdego ciągu liczbowego $\{\chi_k\}$, takiego że $\lim_{k \rightarrow \infty} \chi_k = 0$.

Definicja B.1.5 [19] Ciąg zmiennych losowych $\{\varkappa_k\}$ ma **szybkość zbieżności rzędu** $O(e_k)$ **według średniej z kwadratem** przy $k \rightarrow \infty$ jeżeli stała $0 \leq c < \infty$, taka, że

$$\mathbf{E} \varkappa_k^2 \leq c e_k$$

Lemat B.2 [19] Jeżeli $\varkappa_k = O(e_k)$ według średniej z kwadratem, to $\varkappa_k = O(\sqrt{e_k})$ według prawdopodobieństwa.

B.2 Rozkład SVD macierzy

Twierdzenie B.2.1 [33] Dla każdej macierzy $\mathbf{A} \in \mathbf{R}^{m,n}$ istnieją takie macierze ortogonalne $\mathbf{U} \in \mathbf{R}^{m,m}$, $\mathbf{V} \in \mathbf{R}^{n,n}$, że

$$\mathbf{U}^T \mathbf{A} \mathbf{V} = \mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_l) \quad (\text{B.2})$$

gdzie $l = \min(m, n)$, oraz przy $r = \text{rank}(\mathbf{A})$

$$\begin{aligned} \sigma_1 &\geq \sigma_2 \geq \dots \geq \sigma_r > 0 \\ \sigma_{r+1} &= \dots = \sigma_l = 0 \end{aligned}$$

Liczby $\sigma_1, \dots, \sigma_l$ są określone przez \mathbf{A} jednoznacznie i nazywają się wartościami szczególnymi macierzy \mathbf{A} . Rozwiązując równanie (B.2) względem \mathbf{A} , otrzymujemy równość

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T = \sum_{i=1}^r \mathbf{u}_i \sigma_i \mathbf{v}_i^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (\text{B.3})$$

gdzie \mathbf{u}_i i \mathbf{v}_i oznaczają i -te kolumny odpowiednio macierzy \mathbf{U} i \mathbf{V} . Rozkład SVD macierzy \mathbf{A} dany wzorem (B.3) jest jednoznaczny [33].

B.3 Błąd średniokwadratowy MSE

Definicja B.3.1

$$\widehat{MSE}(\varkappa) = E \left\{ (\widehat{\varkappa} - \varkappa)^2 \right\} = \text{bias}^2 \widehat{\varkappa} + \text{var} \widehat{\varkappa}, \quad (\text{B.4})$$

gdzie bias jest obciążeniem estymatora, var jest wariancją estymatora, a wartość prawdziwą (estymowaną) zapiszemy jako MSE .

B.4 Mocne prawo wielkich liczb Kołmogorowa

Twierdzenie B.4.1 *Jeżeli $\{\mathbf{X}_N\}$ będzie ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie, to*

$$\frac{1}{N} \sum_{k=1}^N \mathbf{X}_k \rightarrow \mathbf{EX} \quad (\text{B.5})$$

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

B.5 Dwu-etapowy algorytm identyfikacji z agregacją parametrów

Algorytm zaprezentowany w [3] składa się z dwóch kroków – pierwszym jest użycie rekurencyjnej metody najmniejszych kwadratów a drugim zastosowanie rozkładu SVD (patrz dodatek B.2). W niniejszym dodatku pokazano wariant algorytmu dla systemu Hammersteina, ale warto podkreślić, że można go również zastosować w identyfikacji Hammersteina-Wienera [3]. Zapiszmy charakterystykę nieliniową jako

$$\mu(u) = \sum_{i=1}^m c_i f_i(u). \quad (\text{B.6})$$

Wprowadźmy wektory parametrów systemu Hammersteina

$$\Gamma = (\gamma_0, \dots, \gamma_n)^T, \quad (\text{B.7})$$

$$c = (c_1, \dots, c_m)^T, \quad (\text{B.8})$$

gdzie n oraz m są długościami wektorów parametrów. Zagregowany wektor parametrów zapiszmy jako

$$\theta = (\gamma_0 c_1, \dots, \gamma_0 c_m, \dots, \gamma_n c_1, \dots, \gamma_n c_m). \quad (\text{B.9})$$

Estymator $\hat{\theta}$ jest wyznaczany klasycznie, metodą najmniejszych kwadratów. Macierz $\hat{\Theta}$, powstałą wskutek rozcięcia kolumn poddaje się dekompozycji *SVD* (patrz dodatek B.2), w celu dezagregacji

$$\hat{\Theta}_{\Gamma,c}^{(LS)} = \sum_{i=1}^{\min(n,m)} \sigma_i \hat{\mu}_i \hat{\mathbf{v}}_i^T. \quad (\text{B.10})$$

Co ostatecznie prowadzi do

$$\hat{\Gamma}_N^{(LS)} = \text{sgn}(\hat{\mu}_1[\kappa_{\mu_1}]) \hat{\mu}_1, \quad (\text{B.11})$$

$$\hat{c}_N^{(LS)} = \text{sgn}(\hat{\mu}_1[\kappa_{\mu_1}]) \sigma_1 \hat{\mathbf{v}}_1. \quad (\text{B.12})$$

Jak widać, w przypadku opisywanej metody dla $n = 20$ oraz $m = 20$, wymiar po agregacji wyniesie 400, co jest znaczącą wadą algorytmu. Opisywany w Rozdziale 5 algorytm radzi sobie z tym problemem za pomocą dekompozycją zadania z predykcji interakcji.

B.6 Algorytm optymalizacji wypukłej

Iteracyjny algorytm minimalizujący funkcję wypukłą jednej zmiennej $f(x)$,
 $x^* = \arg \min_x f(x)$ [36].

Inicjalizacja

Ustaw wartości początkowe $[\kappa_0, \xi_0]$, oraz ε_0

tak żeby $\kappa_0 < x^* < \xi_0$ oraz $0 < \varepsilon_0 < \frac{\kappa_0 + \xi_0}{2}$.

$n := 0$

Powtarzaj aż $|\xi_n - \kappa_n|$ będzie odpowiednio małe

$n := n + 1$

Jeżeli $f\left(\frac{\kappa_n + \xi_n}{2} - \varepsilon_n\right) \geq f\left(\frac{\kappa_n + \xi_n}{2} + \varepsilon_n\right)$ **to**

$$\kappa_{n+1} := \frac{\kappa_n + \xi_n}{2} - \varepsilon_n$$

$$\xi_{n+1} := \xi_n$$

Jeżeli $f\left(\frac{\kappa_n + \xi_n}{2} - \varepsilon_n\right) < f\left(\frac{\kappa_n + \xi_n}{2} + \varepsilon_n\right)$ **to**

$$\kappa_{n+1} := \kappa_n$$

$$\xi_{n+1} := \frac{\kappa_n + \xi_n}{2} + \varepsilon_n$$

$$\varepsilon_{n+1} := \varepsilon_n / 2$$

Bibliografia

- [1] H. Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 716–723.
- [2] D. Aryani, L. Wang, and T. Patikirikorala. “On identification of Hammerstein and Wiener model with application to virtualised software system”. In: *International Journal of Systems Science* 48.6 (2017), pp. 1146–1161.
- [3] E.W. Bai. “An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems”. In: *Automatica* 34.3 (1998), pp. 333–338.
- [4] M. Bodson et al. “High-performance nonlinear feedback control of a permanent magnet stepper motor”. In: *IEEE Transactions on Control Systems Technology* 1.1 (1993), pp. 5–14.
- [5] F. Chang and R. Luus. “A noniterative method for identification using Hammerstein model”. In: *IEEE Transactions on Automatic Control* 16.5 (1971), pp. 464–468.
- [6] Joël Charrier et al. “Evanescent wave optical micro-sensor based on chalcogenide glass”. In: *Sensors and Actuators B: Chemical* 173 (2012), pp. 468–476.
- [7] S. Chen and S.A. Billings. “Representations of non-linear systems: the NARMAX model”. In: *International Journal of Control* 49.3 (1989), pp. 1013–1032.
- [8] S. Chen, S.A. Billings, and W. Luo. “Orthogonal least squares methods and their application to non-linear system identification”. In: *International Journal of Control* 50.5 (1989), pp. 1873–1896.
- [9] Y.S. Chow and H. Teicher. *Probability Theory: Independence, Interchangeability, Martingales*. Springer Verlag, 2003.
- [10] F. Ding and T. Chen. “Identification of Hammerstein nonlinear ARMAX systems”. In: *Automatica* 41.9 (2005), pp. 1479–1489.
- [11] Robert Fairman and Boris Ushkov. *Semiconducting chalcogenide glass III: applications of chalcogenide glasses*. Elsevier, 2004.
- [12] B. Finigan and I. Rowe. “Strongly consistent parameter estimation by the introduction of strong instrumental variables”. In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 825–830.
- [13] C. F. Gauss and C. H. Davis. *Theory of the motion of the heavenly bodies moving about the sun in conic sections a translation of Gauss’s*. <https://www.biodiversitylibrary.org/bibliography/19023>. Boston, Little, Brown and company, 1857, p. 414. URL: <https://www.biodiversitylibrary.org/item/58729>.
- [14] G.B. Giannakis and E. Serpedin. “A bibliography on nonlinear system identification”. In: *Signal Processing* 81.3 (2001), pp. 533–580.

- [15] F. Giri et al. “Frequency identification of nonparametric Wiener systems containing backlash nonlinearities”. In: *Automatica* 49.1 (2013), pp. 124–137.
- [16] W. Greblicki. “Nonparametric identification of Wiener systems”. In: *IEEE Transactions on Information Theory* 38.5 (1992), pp. 1487–1493.
- [17] W. Greblicki. “Nonparametric orthogonal series identification of Hammerstein systems”. In: *International Journal of System Science* 20 (1989), pp. 2355–2367.
- [18] W. Greblicki, A. Krzyżak, and M. Pawlak. “Distribution-free pointwise consistency of kernel regression estimate”. In: *The Annals of Statistics* 12.4 (1984), pp. 1570–1575.
- [19] W. Greblicki and M. Pawlak. “Fourier and Hermite series estimates of regression functions”. In: *Annals of the Institute of Statistical Mathematics* 37.1 (1985), pp. 443–454.
- [20] W. Greblicki and M. Pawlak. “Identification of discrete Hammerstein systems using kernel regression estimates”. In: *IEEE Transactions on Automatic Control* 31 (1986), pp. 74–77.
- [21] W. Greblicki and M. Pawlak. *Nonparametric System Identification*. Cambridge University Press, 2008.
- [22] W. Greblicki, D. Rutkowska, and L. Rutkowski. “An orthogonal series estimate of time-varying regression”. In: *Annals of the Institute of Statistical Mathematics* 35.1 (1983), pp. 215–228.
- [23] L. Györfi et al. *A Distribution-Free Theory of Nonparametric Regression*. New York: Springer Verlag, 2002.
- [24] R. Haber. “Structural identification of quadratic block-oriented models based on estimated Volterra kernels”. In: *International Journal of Systems Science* 20.8 (1989), pp. 1355–1380.
- [25] L.P. Hansen and K.J. Singleton. “Generalized instrumental variables estimation of nonlinear rational expectations models”. In: *Econometrica: Journal of the Econometric Society* (1982), pp. 1269–1286.
- [26] W. Härdle. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press, 1990.
- [27] Z. Hasiewicz. “Identification of a linear system observed through zero-memory non-linearity”. In: *International Journal of Systems Science* 18.9 (1987), pp. 1595–1607.
- [28] Z. Hasiewicz and G. Mzyk. “Combined parametric-nonparametric identification of Hammerstein systems”. In: *IEEE Transactions on Automatic Control* 48.8 (2004), pp. 1370–1376.

- [29] Z. Hasiewicz and G. Mzyk. “Hammerstein system identification by non-parametric instrumental variables”. In: *International Journal of Control* 82.3 (2009), pp. 440–455.
- [30] Z. Hasiewicz, M. Pawlak, and P. Śliwiński. “Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support”. In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 52.2 (2005), pp. 427–442.
- [31] Z. Hasiewicz, P. Śliwiński, and G. Mzyk. “Nonlinear system identification under various prior knowledge”. In: *Proceedings of the 17th World Congress. The International Federation of Automatic Control*. 2008, pp. 7849–7858.
- [32] Z. Hasiewicz et al. “Mixed Parametric-Nonparametric Identification of Hammerstein and Wiener Systems-a Survey”. In: *Proceedings of the 16th IFAC Symposium on System Identification, Brussels, Belgium*. Vol. 16. 2012, pp. 464–469.
- [33] A. Kiełbasiński and H. Schwetlick. *Numeryczna Algebra Liniowa: Wprowadzenie do Obliczeń Zautomatyzowanych, (in Polish)*. Wydawnictwa Naukowo-Techniczne, Warszawa, 1992.
- [34] B. Kozdraś and G. Mzyk. “Identification of the heating process in Differential Scanning Calorimetry with the use of Hammerstein model”. In: *submitted to Control Engineering Practice* (2016).
- [35] B. Kozdraś, G. Mzyk, and P. Mielcarek. “Identification of the heating process in Differential Scanning Calorimetry with the use of Hammerstein model”. In: IEEE. 2020.
- [36] Paweł Mielcarek and Grzegorz Mzyk. “Recursive algorithm for interaction prediction in Hammerstein system identification with experimental studies”. In: (2023).
- [37] G. Mzyk. “A censored sample mean approach to nonparametric identification of nonlinearities in Wiener systems”. In: *IEEE Transactions on Circuits and Systems – II: Express Briefs* 54.10 (2007), pp. 897–901.
- [38] G. Mzyk. “Application of instrumental variable method to the identification of Hammerstein-Wiener systems”. In: *Proceedings of the 6th International Conference MMAR*. 2000, pp. 951–956.
- [39] G. Mzyk. *Combined Parametric-Nonparametric Identification of Block-Oriented Systems*. Lecture Notes in Control and Information Sciences 454. Springer, 2014.
- [40] G. Mzyk. “Wiener-Hammerstein system identification with non-Gaussian input”. In: *IFAC International Workshop on Adaptation and Learning in Control and Signal Processing, Antalya, Turkey*. 2010.

- [41] G. Mzyk and P. Wachel. “Kernel-based identification of Wiener-Hammerstein system”. In: *Automatica* 83 (2017), pp. 275–281.
- [42] Grzegorz Mzyk. “Efficient identification of Hammerstein systems by two-level optimization with decomposition”. In: (2018), pp. 89–92.
- [43] E.A. Nadaraya. “On estimating regression”. In: *Teoriya Veroyatnostei i ee Primeneniya* 9.1 (1964), pp. 157–159.
- [44] K.S. Narendra and P. G. Gallman. “An iterative method for the identification of nonlinear systems using the Hammerstein model”. In: *IEEE Transactions on Automatic Control* 11 (1966), pp. 546–550.
- [45] T. Patikirikorala et al. “Hammerstein–Wiener nonlinear model based predictive control for relative QoS performance and resource management of software systems”. In: *Control Engineering Practice* 20.1 (2012), pp. 49–61.
- [46] A. Radouane et al. “System identification of a class of Wiener systems with hysteretic nonlinearities”. In: *International Journal of Adaptive Control and Signal Processing* 31.3 (2017), pp. 332–359.
- [47] D. Ruppert, M.P. Wand, and R.J. Carroll. *Semiparametric Regression*. Vol. 12. Cambridge University Press, 2003.
- [48] P. Stoica and T. Söderström. “Instrumental-variable methods for identification of Hammerstein systems”. In: *International Journal of Control* 35.3 (1982), pp. 459–476.
- [49] Leendert Cornelis Elisa Struik. “Physical aging in amorphous polymers and other materials”. In: (1977).
- [50] L. Sun, A. Sano, and X. Liu. “Direct closed-loop identification approach to magnetic levitation system”. In: *IFAC-PapersOnLine* 51.15 (2018), pp. 610–615.
- [51] Vahid G Ta’eed et al. “Ultrafast all-optical chalcogenide glass photonic circuits”. In: *Optics Express* 15.15 (2007), pp. 9205–9221.
- [52] P. Wachel, P. Śliwiński, and Z. Hasiewicz. “Nonparametric identification of MISO Hammerstein system from structured data”. In: *Journal of Systems Science and Systems Engineering* 24.1 (2015), pp. 68–80.
- [53] M.P. Wand and H.C. Jones. *Kernel Smoothing*. London: Chapman and Hall, 1995.
- [54] J. Wang and Q. Zhang. “Identification of an extended Hammerstein system with input hysteresis nonlinearity for control valve stiction characterization”. In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 268–273.
- [55] D. T. Westwick and R. E. Kearney. *Identification of Nonlinear Physiological Systems*. Vol. 7. John Wiley & Sons, 2003.
- [56] W.B. Wu and J. Mielniczuk. “Kernel density estimation for linear processes”. In: *The Annals of Statistics* 30.5 (2002), pp. 1441–1459.

- [57] Matthias Wuttig and Noboru Yamada. “Phase-change materials for rewriteable data storage”. In: *Nature materials* 6.11 (2007), pp. 824–832.
- [58] Z. Hasiewicz and G. Mzyk. “Two-Level Recursive Identification of Hammerstein System by Interaction Prediction Method”. In: *International Conference on Reliability and Statistics in Transportation and Communication* (2022), pp. 119–129.

Spis rysunków

2.1	System Hammersteina.	9
2.2	System Wienera.	10
3.1	System Hammersteina–Wienera	16
3.2	Sprawdzian krzyżowy w celu wyznaczenia najlepszej szerokości jądra h	24
3.3	Nieparametryczny estymator wyjściowej charakterystyki nieliniowej (czerwone plusy) porównany do nieliniowej charakterystyki wyjściowej systemu $\eta(x)$ (niebieska krzywa). W tle pomiary wyjścia (czarne punkty)	25
3.4	Nieparametryczna estymacja gęstości rozkładu prawdopodobieństwa zakłóceń $f(z)$	26
4.1	Przestrzeń wejść dla $s = 2$ oraz $d = 1$, (a) k nieparzyste, (b) k parzyste.	31
4.2	Przestrzeń wejść dla $s = 2$ oraz $d = 2$	31
4.3	2-wymiarowy model $\hat{R}_2(x_0, x_1)$	33
4.4	3-wymiarowy model $\hat{R}_3(x_0, x_1, x_2)$	33
5.1	System Hammersteina z predykcją sygnału wewnętrznego (interakcyjnego).	38
5.2	System Hammersteina z elementem liniowym typu ARMAX....	45
5.3	Zmiana estymatora c podczas kolejnych iteracji.	46
5.4	Estymator c_N	47
5.5	Zmiana estymatora $\hat{\gamma}_N[1]$	48
5.6	Przesunięcie charakterystyki regresji.	48

7.1	Schemat urządzenia do pomiaru metodami DTA oraz DSC	54
7.2	Model Hammersteina z liniowym blokiem typu ARMAX	55
7.3	Zbiór danych weryfikacyjnych, gdzie u jest mocą grzałki a y jest temperaturą próbki	55
7.4	Identyfikacja za pomocą modelu ARMAX, prawdziwe wyjście oraz wyjście modelu	56
7.5	Weryfikacja modelu ARMAX, prawdziwe wyjście oraz wyjście modelu	57
7.6	Eksperyment—regresja 1-wymiarowa	57
7.7	Eksperyment—regresja 2-wymiarowa	58
7.8	Eksperyment—regresja 3-wymiarowa	58
7.9	Eksperyment—regresja 4-wymiarowa	59
8.1	Tokarko-frezarka CNC	62
8.2	Rozkład procesu na składowe deterministyczną d_k oraz stochastyczną g_k	63
8.3	Składowa deterministyczna	64
8.4	Pobór prądu podczas tej samej operacji technologicznej dla dwóch różnych cykli	65
8.5	Układ pomiarowy	66
8.6	Pomiar chwilowego prądu pobieranego przez maszynę podczas trzech cykli toczenia	67
8.7	Uczenie modelu $s = 1$	70
8.8	Uczenie modelu $s = 2$	70
8.9	Uczenie modelu $s = 3$	71
8.10	Eksperyment wykrycia awarii	72
8.11	Weryfikacja drzewa regresji na danych z procesu zgodnego	74
8.12	Weryfikacja drzewa regresji na danych z procesu niezgodnego . .	74
8.13	Weryfikacja algorytmu MLP na danych z procesu zgodnego	76
8.14	Weryfikacja algorytmu MLP na danych z procesu niezgodnego .	76

Spis tabel

4.1	Porównanie błędu średniokwadratowego δ_s wyjścia modelu dla $s = 1, 2, 3$, do najlepszej możliwej aproksymacji liniowej (BLA).	34
5.1	Wyniki symulacji.	47
7.1	Błąd aproksymacji.	55
7.2	Błąd średniokwadratowy δ_s wyjścia modelu, dla $s = 1, 2, 3, 4$. ..	59
8.1	Porównanie błędów podczas nauki modelu	71
8.2	Porównanie błędów podczas weryfikacji modelu	71
8.3	Eksperyment wykrycia awarii	72
8.4	Weryfikacja drzewa regresji	75
8.5	Weryfikacja sztucznej sieci neuronowej	77
8.6	Porównanie algorytmów	77