

Wydział Elektroniki
Katedra Automatyki, Mechatroniki i Systemów Sterowania
Politechnika Wrocławska

**MODELOWANIE SYSTEMÓW
NIELINIOWYCH O NIEZNANEJ
STRUKTURZE**
(rozprawa doktorska)

Szymon Łagosz

Promotor:

dr hab. inż. Przemysław Śliwiński, prof. PWr

Promotor pomocniczy:

dr hab. inż. Paweł Wachel, prof. PWr

Wrocław 2021

Podziękowania

Serdecznie dziękuję moim Promotorom – Panu Profesorowi Przemysławowi Śliwińskiemu oraz Panu Profesorowi Pawłowi Wachłowi – za możliwość rozwoju naukowego, liczne inspirujące dyskusje, a także za poświęcony mi czas i okazaną życzliwość.

Osobne podziękowania pragnę złożyć Rodzicom oraz Marcie, bez wsparcia których ta praca nigdy by nie powstała.

Spis treści

Lista symboli	v
Streszczenie	vii
Abstract	ix
1 Wprowadzenie	1
1.1 Modelowanie systemów – zadania i cele	1
1.2 Wybór reprezentacji systemu	4
1.2.1 Reprezentacja dynamicznych systemów nieliniowych	6
1.2.2 Szereg Volterra	8
1.2.3 Szereg Wienera i rozwinięcie w chaos wielomianowy	9
1.3 Modele wielomianowe i problem wielowymiarowości	12
1.4 Estymacja modeli wielomianowych	12
1.4.1 Metoda najmniejszych kwadratów	13
1.4.2 Regularyzacja	14
1.5 Metody optymalizacji stochastycznej pierwszego rzędu	18
1.5.1 Postać ogólna metod pierwszego rzędu	18
1.5.2 Metoda stochastycznego gradientu prostego	19
1.5.3 Metody stochastyczne, a metody <i>online</i> optymalizacji wypukłej	20
1.5.4 Metody lustrzane	22
1.5.5 Metody lustrzane z uśrednianiem	23
1.6 Zastosowania praktyczne modeli wielomianowych	24
1.7 Cele pracy	25
1.8 Założenia ogólne i klasyfikacja zadań	25
1.9 Teza pracy	25
2 Reprezentacja systemów oparta o szereg Volterra	27
2.1 Rozwinięcie Volterra	27
2.2 Własności aproksymacyjne	29
2.3 Studium przypadku – aproksymacja systemu LNL	33
2.3.1 Błąd aproksymacji	34
2.3.2 Podsumowanie	35
2.4 Konstrukcja modeli volterrowskich	35
Dodatki	37
2.A Dowód lematu 2.1	37

3	Reprezentacje systemów oparte o szeregi ortogonalne	41
3.1	Ogólna postać reprezentacji ortogonalnej	42
3.2	Wejście o rozkładzie gaussowskim. Szereg Wienera	42
3.2.1	Wyznaczanie jąder funkcjonałów	44
3.2.2	Własności aproksymacyjne	45
3.3	Wejście o rozkładzie nieguassowskim. Rozwinięcie w chaos wielomianowy	49
3.4	Konstrukcja modeli ortogonalnych	51
3.5	Studium przypadku – aproksymacja systemu NL	51
3.5.1	Błąd aproksymacji	52
	Dodatki	55
3.A	Wielomiany ortogonalne	55
3.A.1	Wielomiany Hermite’a	55
3.A.2	Wielomiany Legendre’a	56
3.B	Rozwinięcia Wienera – konstrukcja	56
3.B.1	Ortogonalizacja funkcjonałów Volterra	56
3.B.2	Inne własności	60
4	Modelowanie systemów z wykorzystaniem metod lustrzanych z uśrednianiem	65
4.1	Algorytm lustrzany z uśrednianiem	65
4.1.1	Dobór długości kroku	69
4.1.2	Określenie funkcji regularyzującej	71
4.2	Modelowanie systemów dynamicznych	73
4.2.1	Systemy z pamięcią skończoną	74
4.2.2	Systemy z pamięcią zanikającą	76
4.2.3	Modelowanie systemów z pamięcią zanikającą za pomocą wariantu entropijnego algorytmu DA oraz modeli opartych o reprezentacje wielomianowe	78
4.3	Modelowanie systemu LNL	81
4.3.1	Opis problemu	81
4.3.2	Błąd metody	81
4.4	Eksperyment numeryczny	83
4.4.1	Eksperyment I - zależność błędu metody od wartości stałej G	85
4.4.2	Eksperyment II - zależność błędu metody od liczby pomiarów	86
4.5	Podsumowanie	87
	Dodatki	89
4.A	Dowód twierdzenia 4.1	89
4.B	Dowód twierdzenia 4.2	91
4.C	Dowód twierdzenia 4.3	93
4.D	Lematy techniczne	94

4.E	Podstawowe fakty i definicje z optymalizacji wypukłej wykorzystane w rozdziale	97
5	Uwagi końcowe	99
5.1	Oryginalne wyniki naukowe przedstawione w pracy	99
5.2	Otwarte problemy badawcze	100
	Bibliografia	101

Lista symboli

$\ \cdot\ _p$	norma ℓ_p
$\ \cdot\ _*$	norma dualna
$\langle \cdot, \cdot \rangle$	iloczyn skalarny
$[\cdot]$	funkcja „podłoga” (część całkowita z argumentu)
\mathcal{C}	zbiór ograniczeń parametrów modelu
$C_T(\cdot)$	suma empirycznych strat modelu po T iteracjach
$\Delta_{\text{apr.}}$	błąd aproksymacji metody
$\Delta_{\text{est.}}$	błąd estymacji metody
D	rozmiar słownika
$\mathbb{E}\{\cdot\}$	operator wartości oczekiwanej
\mathcal{F}_t	σ -ciało generowane przez t pierwszych par pomiarowych
\mathbf{g}_i	gradient funkcji straty w i -tej iteracji
$\mathbf{g}_{a:b}$	suma gradientów funkcji straty o indeksach od a do b
$G_p[\cdot; \cdot]$	funkcjonał Wienera p -tego rzędu
$G_{p,M}[\cdot; \cdot]$	obcięty funkcyjonał Wienera p -tego rzędu (długości pamięci M)
η_i	długość korku algorytmu w i -tej iteracji
$H_p[\cdot; \cdot]$	funkcjonał Volterry p -tego rzędu
$H_{p,M}[\cdot; \cdot]$	obcięty funkcyjonał Volterry p -tego rzędu (długości pamięci M)
$\mathcal{H}_p(\cdot)$	wielomian Hermite’a p -tego rzędu
\mathcal{K}_p	tensor p -tego rzędu
$\hat{L}_S(\cdot)$	empiryczna funkcja straty modelu opartego o słownik S
$\hat{L}_t(\cdot)$	empiryczna strata modelu w chwili t
$L(\cdot)$	(kwadratowa) funkcja straty
$L_S(\cdot)$	(kwadratowa) funkcja straty dla modeli opartych o słownik S
$\mathcal{L}_p(\cdot)$	wielomian Legendre’a p -tego rzędu
\mathcal{M}	zbiór hipotetycznych modeli systemu
M	długość pamięci modelu
$m_{\text{sys}}[\cdot]$	charakterystyka systemu
$m^*[\cdot]$	najlepszy model w danej klasie \mathcal{M}
$m_S[\cdot; \cdot]$	model oparty o słownik S
N	liczba pomiarów wejścia-wyjścia systemu
\mathcal{O}	notacja „duże O”
o	notacja „małe o”
P	rzęd modelu

$Q(\cdot)$	funkcja kryterialna
\mathcal{S}	zbiór pomiarów wejścia-wyjścia systemu
$\tilde{S}_{P,M}^r$	rozszerzony słownik wielomianowy (przeskalowany przez stałą r)
S	słownik funkcjonałów
$S_{P,M}$	słownik wielomianowy
$\text{sgn}(\cdot)$	funkcja <i>signum</i> (znak)
$\boldsymbol{\theta}$	wektor parametrów modelu
$\boldsymbol{\theta}^*$	wektor parametrów najlepszego możliwego modelu
$\boldsymbol{\theta}_i$	rozwiązanie pośrednie algorytmu w i -tej iteracji
T	liczba iteracji algorytmu
$V[\cdot]$	szereg Volterra
$V_{P,M}[\cdot]$	obcięty szereg Volterra (rzędu P , długości pamięci M)
$\text{Var}\{\cdot\}$	wariancja
$W[\cdot]$	szereg ortogonalny (Wienera)
$W_{P,M}[\cdot]$	obcięty szereg ortogonalny (rzędu P , długości pamięci M)
\mathbf{X}_n	wektor przechowujący losowe wejścia systemu
\mathbf{x}_n	wektor przechowujący deterministyczne wejścia systemu
X_n	losowe wejście systemu w chwili n
x_n	deterministyczne wejście systemu w chwili n
Y_n	losowe wyjście systemu w chwili n
y_n	deterministyczne wyjście systemu w chwili n
Z_n	losowe zakłócenie addytywne w chwili n

Streszczenie

Rozprawa dotyczy problemu modelowania dynamicznych systemów nieliniowych w warunkach nieznajomości ich struktury wewnętrznej. Niewielka wiedza wstępna na temat badanego systemu wymaga zastosowania odpowiednio uniwersalnych metod. W związku z powyższym, w rozprawie wykorzystuje się posiadające własność uniwersalnej aproksymacji reprezentacje wielomianowe, oparte o szereg Volterra oraz szeregi ortogonalne (szeregi Wienera). Inherentnym problemem rozważanej klasy modeli jest wysoka liczba nieznanymi parametrów, rosnąca eksponencjalnie wraz ze stopniem złożoności modelu. Zaproponowane w pracy metody modelowania, wykorzystujące nowoczesne algorytmy optymalizacji wypukłej, w istotnym stopniu niwelują niekorzystny wpływ tego zjawiska na jakość uzyskanych modeli, a przy tym cechują się niską złożonością obliczeniową.

Rozdział pierwszy zawiera wprowadzenie do problemu modelowania systemów. Wyjaśniony został wpływ wiedzy wstępnej na wybór reprezentacji systemu oraz przedstawiono powód, dla którego modele wielomianowe są odpowiednim wyborem w przypadku nieznajomości struktury badanego systemu. Następnie dokonano przeglądu istniejących metod modelowania systemów oraz zaproponowano, by do celów estymacji modeli wielomianowych wykorzystać algorytmy optymalizacji stochastycznej. Na końcu rozdziału zaprezentowano tezę i cele pracy. W rozdziale drugim i trzecim omówiona została reprezentacja systemów oparta kolejno o szereg Volterra oraz szeregi ortogonalne. Szczególny nacisk położono na przedstawienie ich własności aproksymacyjnych, najpierw o charakterze ogólnym, a następnie na przykładzie wybranych systemów o strukturze blokowej z gładkimi nieliniowościami, tj. systemów Wienera-Hammersteina (w przypadku reprezentacji opartej o szereg Volterra) oraz systemów Hammersteina (w przypadku reprezentacji opartych o szeregi ortogonalne). W rozdziale czwartym zbadano możliwość wykorzystania algorytmu lustrzanego z uśrednianiem w zadaniu estymacji modeli wielomianowych, szczególną uwagę poświęcając entropijnemu wariantowi metody. Jako główne osiągnięcie wskazać można dowód zbieżności błędu estymacji metody dla systemów z pamięcią zanikającą eksponencjalnie. Działanie zaproponowanego algorytmu zostało zilustrowane eksperymentem numerycznym. Rozdział piąty zawiera podsumowanie uzyskanych w pracy rezultatów oraz przedstawienie otwartych problemów badawczych.

Abstract

The thesis concerns the problem of modelling dynamic nonlinear systems in the absence of information regarding their inner structure. Limited a priori knowledge about the examined system requires application of sufficiently universal methods. In view of the above, in the thesis representations possessing the universal approximation property have been employed, specifically those based on Volterra and orthogonal (Wiener) series. The inherent problem of the considered class of models is a huge number of unknown parameters, growing exponentially with the candidate models' complexity. The modelling methods proposed in the thesis employ modern optimization methods and to a large extent reduce the negative impact of this phenomenon on the quality of obtained models while at the same time keeping low computational cost.

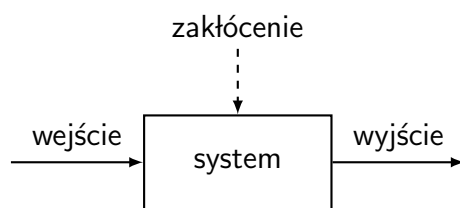
The first chapter of the thesis serves as an introduction to the systems modelling problem. It discusses the role of a priori knowledge in choosing the system's representation and explains the reason for which polynomial models are suitable candidates when the system's structure remains unknown. Then, we review existing methods for nonlinear systems modeling and propose to apply stochastic optimization algorithms for the purpose of polynomial models estimation. The last part of the chapter presents the statement and aims of the thesis. In chapters two and three we discuss systems representations based on Volterra and orthogonal series, with particular emphasis on their approximation properties, examined in details for particular block-oriented systems with smooth nonlinearities, i.e. Wiener-Hammerstein systems (in case of Volterra representation) and Hammerstein systems (in case of orthogonal representations). In chapter four we inspect the theoretical possibilities of applying the dual averaging algorithm for polynomial models estimation, paying special attention to the entropic variant of the method. In the main result of this study we demonstrate that the estimation error of the aforementioned algorithm converges for systems possessing exponentially fading memory. Numerical experiments are supplied to illustrate the effectiveness of the proposed method. Finally, chapter five summarizes the main research findings and presents open research problems.

Rozdział 1

Wprowadzenie

1.1 Modelowanie systemów – zadania i cele

W ramach przedstawionych w pracy rozważań, *systemem* będziemy nazywali bliżej nieokreślony obiekt, który przekształca pewien sygnał wejściowy (wejście systemu) w sygnał wyjściowy (wyjście systemu) i który działa w obecności niedostępnych dla pomiarów zakłóceń (rys. 1.1).



Rysunek 1.1: Schemat systemu. Linia przerywaną oznaczono grupę sygnałów niepodlegających bezpośrednim pomiarom

Jeśli dyskusję zawężymy do systemów klasy SISO (ang. *Single Input Single Output*; sygnały wejścia/wyjścia są jednowymiarowe), z czasem dyskretnym (sygnały są ciągami liczbowymi) i zakłóceniami o charakterze addytywnym, to – w ujęciu funkcjonalnym – systemy takie można opisać za pomocą ogólnego wzoru¹

$$Y_n = m_{\text{sys}}[\mathbf{X}_n] + Z_n, \quad (1.1)$$

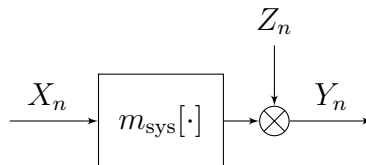
gdzie $Y_n \in \mathbb{R}$ stanowi wyjście systemu w pewnej chwili n , wektor

$$\mathbf{X}_n = [X_n, X_{n-1}, X_{n-2}, \dots] \in \mathbb{R}^\infty$$

zawiera obecne i przeszłe wartości wejść systemu, $Z_n \in \mathbb{R}$ oznacza zakłócenie, a $m_{\text{sys}} : \mathbb{R}^\infty \rightarrow \mathbb{R}$ jest funkcjonałem, zwanym dalej *charakterystyką systemu*, wyrażającym regułę, zgodnie z którą system odwzorowuje wejście w wyjście (rys. (1.2)).

¹Zgodnie z przyjętą konwencją, jeśli zmienna ma charakter losowy to oznaczana będzie wielką literą. Wytłuszczenie natomiast będzie stosowane przy oznaczaniu wektorów i macierzy (zarówno losowych, jak i deterministycznych).

Zauważmy, że przyjmując opis (1.1), wprowadziliśmy niejawnie założenie o stacjonarnym oraz przyczynowym charakterze rozważanych systemów (funkcjonał $m_{\text{sys}}[\cdot]$ nie zmienia się w czasie, a wyjście systemu Y_n nie zależy od przyszłych wartości wejść, tj. zbioru $\{X_{n+1}, X_{n+2}, \dots\}$).



Rysunek 1.2: System z addytywnym zakłóceniem na wyjściu

Ze względu na naturę charakterystyki $m_{\text{sys}}[\cdot]$, systemy klasyfikujemy jako *liniowe* lub *nieliniowe*. Tak więc, *systemami liniowymi* nazywamy systemy, których charakterystyka spełnia zasadę superpozycji, tj. dla każdego $\alpha, \beta \in \mathbb{R}$ i $\mathbf{x}, \mathbf{y} \in \mathbb{R}^\infty$ zachodzi

$$m_{\text{sys}}[\alpha\mathbf{x} + \beta\mathbf{y}] = \alpha m_{\text{sys}}[\mathbf{x}] + \beta m_{\text{sys}}[\mathbf{y}].$$

Teoria dotycząca systemów tej klasy jest dobrze poznana [10, 66, 105] i ze względu na szereg korzystnych własności, ich analiza na ogół nie sprawia trudności. Niestety jednak, dla zdecydowanej większości zjawisk występujących w przyrodzie opis liniowy jest poprawny tylko w przybliżeniu i dla ograniczonej klasy sygnałów wejściowych. W związku z powyższym, mając na względzie zastosowania praktyczne, nasza uwaga skupiona będzie na *systemach nieliniowych*, tj. takich które zasady superpozycji nie spełniają. Ich bogactwo, a zarazem złożoność w obrazowy sposób ujmują słowa Stanisława Ulama, który teorię systemów nieliniowych miał przyrównać do działu zoologii zajmującego się wszystkimi zwierzętami będącymi *nie-słoniami* [17].

Modelowanie systemów jest terminem mającym obszerne znaczenie, który w ogólności rozumieć można jako dziedzinę nauki badającą metody znajdowania *modelu matematycznego*, czyli formalnego opisu systemu, przy wykorzystaniu zbioru pomiarów sygnałów wejścia i wyjścia oraz wszelkiej dostępnej wiedzy wstępnej dotyczącej badanego systemu. W tym miejscu należy zwrócić uwagę, że choć treść tej definicji w znaczącym stopniu pokrywa się z pojęciem *identyfikacji systemów* (por. np. [105, Rozdział 1]), to jednak część badaczy (np. Wachel w [115, Rozdział 2]) dokonuje następującego rozróżnienia pomiędzy zadaniami stawianymi przed identyfikacją/modelowaniem systemów: celem identyfikacji jest znalezienie opisu możliwie dokładnie odtwarzającego strukturę, cechy fizyczne i działanie poszczególnych komponentów systemu, natomiast cel modelowania ogranicza się do odwzorowania relacji wejścia-wyjścia systemu, z uwzględnieniem pewnej, uprzednio przyjętej funkcji kryterialnej, której rolą jest ocena jakości uzyskanego opisu – w tym ujęciu, modelowanie systemów rozpatrywać można jako obszar szerszej dziedziny nauki jaką jest uczenie maszynowe. Warto jednak pamiętać, że w związku z tym, że granica pomiędzy pojęciami identyfikacji i modelowania jest umowna, bywają one używane zamiennie, a ich znaczenie często zależy od kontekstu w jakim są umieszczone.

Formalnie powiemy, że celem modelowania jest znalezienie modelu $m \in \mathcal{M}$, który minimalizuje wartość funkcji kryterialnej $Q(m)$, gdzie $\mathcal{M} \subset \{m : \mathbb{R}^\infty \rightarrow \mathbb{R}\}$ jest wcześniej przyjętym, zazwyczaj nieprzeliczalnym zbiorem hipotetycznych modeli, stanowiącym kodyfikację wiedzy apriorycznej użytkownika o badanym systemie. Środkiem do realizacji celu modelowania są *algorytmy modelowania*, które jako wejście przyjmują zbiór \mathcal{M} oraz uzyskany w ramach eksperymentu zbiór pomiarów wejścia-wyjścia systemu (zwany *zbiorem uczącym*), $\mathcal{S} = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$, natomiast na wyjściu zwracają model empiryczny, który z racji faktu, że w ogólności jest wynikiem działania na zbiorze danych losowych, sam także ma charakter stochastyczny.

Założmy teraz, że pewien algorytm A , zwraca model $\hat{m} = A(\mathcal{S}, \mathcal{M}) \in \mathcal{M}$. Ogólną postać funkcji kryterialnej takiego modelu zdefiniujemy za pomocą wzoru

$$Q(\hat{m}) = \mathbb{E} \{L(\hat{m})\} - L(m_{\text{sys}}), \quad (1.2)$$

gdzie $L(\cdot)$ oznacza tzw. funkcję straty (ang. *loss function*) i gdzie wartość oczekiwana rozpatrywana jest względem rozkładu prawdopodobieństwa elementów zbioru \mathcal{S} . W modelowaniu systemów, podobnie jak w problemach regresji, najczęstszym wyborem jest kwadratowa funkcja straty, dlatego też za $L(\cdot)$ będziemy przyjmowali

$$L(m) = \mathbb{E}\{Y_n - m[\mathbf{X}_n]\}^2, \quad (1.3)$$

gdzie wartość oczekiwana rozpatrywana jest względem rozkładu podobieństwa pary pomiarowej $\{Y_n, \mathbf{X}_n\}$; wartość funkcji straty, $L(m)$, będziemy nazywali *błędem średniokwadratowym* modelu $m[\cdot]$. Funkcja kryterialna mówi więc o *średniej* różnicy pomiędzy błędami średniokwadratowymi modeli zwracanych przez algorytm A i charakterystyki prawdziwego systemu $m_{\text{sys}}[\cdot]$ (zauważmy, że z uwagi na obecność zakłóceń, w ogólności $L(m_{\text{sys}})$ ma niezerową wartość, zob. wzór (1.1)); z tego względu wartość $Q(\hat{m})$ należy rozumieć jako cechę metody (algorytmu A), a nie konkretnej realizacji $\hat{m}[\cdot]$.

Wzór (1.2) będziemy dekomponować w sposób następujący,

$$Q(\hat{m}) = \underbrace{\mathbb{E} \{L(\hat{m})\} - L(m^*)}_{\Delta_{\text{est.}}} + \underbrace{L(m^*) - L(m_{\text{sys}})}_{\Delta_{\text{apr.}}}, \quad (1.4)$$

gdzie $m^*[\cdot]$ jest najlepszym możliwym (w odniesieniu do danej funkcji straty), ale niekoniecznie unikalnym modelem spośród wszystkich należących do zbioru \mathcal{M} , tzn. spełnia warunek

$$m^* \in \arg \min_{m \in \mathcal{M}} L(m).$$

Zauważmy, że jeśli spełnione jest standardowe założenie o zerowej wartości oczekiwanej zakłócenia, $\mathbb{E}\{Z_n\} = 0$, to oprócz funkcji straty, model $m^*[\cdot]$ minimalizuje także wartość funkcji kryterialnej, tj. $m^* \in \arg \min_{m \in \mathcal{M}} Q(m)$, a relacja pomiędzy poszczególnymi komponentami wzoru (1.4) wygląda następująco

$$L(\hat{m}) \geq L(m^*) \geq L(m_{\text{sys}}),$$

przez co $\Delta_{\text{est.}}, \Delta_{\text{apr.}} \geq 0$.

W wyrażeniu (1.4) wyróżniamy dwie główne składowe:

- *Błąd aproksymacji*,

$$\Delta_{\text{apr.}} = L(m^*) - L(m_{\text{sys}}),$$

stanowiący różnicę pomiędzy błędem średniokwadratowym najlepszego możliwego modelu $m^*[\cdot]$ i prawdziwego systemu, $m_{\text{sys}}[\cdot]$. Zasadniczo, im większa złożoność modeli znajdujących się w zbiorze \mathcal{M} , tym wartość $\Delta_{\text{apr.}}$ jest mniejsza. W szczególności, gdy charakterystyka prawdziwego systemu zawiera się w zbiorze modeli hipotetycznych, tzn. $m_{\text{sys}} \in \mathcal{M}$, błąd aproksymacji znika całkowicie, $\Delta_{\text{apr.}} = 0$.

- *Błąd estymacji*,

$$\Delta_{\text{est.}} = \mathbb{E}\{L(\hat{m})\} - L(m^*),$$

będący *średnią* różnicą pomiędzy błędem średniokwadratowym modelu wynikowego $\hat{m}[\cdot]$ i najlepszego możliwego modelu $m^*[\cdot]$. Typowo, błąd ten rośnie wraz ze wzrostem złożoności zbioru \mathcal{M} . Oprócz tego, na jego wielkość wpływ ma użyty algorytm modelowania.

W związku z powyższym, jako jedną z głównych trudności w modelowaniu systemów, oprócz doboru algorytmu modelowania, wskazać można problem strojenia poziomu złożoności rozważanej klasy modeli, tak aby zachowany został kompromis pomiędzy wartością błędu estymacji, a wartością błędu aproksymacji. W ogólności stosowanie bardziej skomplikowanych modeli powoduje, że błąd aproksymacji maleje. Jednocześnie może to prowadzić może do zjawiska zwanego *przeuczeniem* (ang. *overfitting*), którego efektem jest nadmierne dopasowanie modelu do zakłóconych danych pomiarowych i utrata zdolność uogólniania, co z kolei przejawia się wysokim błędem modelu, jeśli na wejściu pojawiają się wartości niezaobserwowane w zbiorze \mathcal{S} . Modele o prostszej strukturze sprzyjają natomiast mniejszemu błędowi estymacji, zarazem powodując, że błąd aproksymacji jest większy.

1.2 Wybór reprezentacji systemu

Jednym z pierwszych etapów modelowania systemów jest na ogół określenie matematycznej reprezentacji badanego obiektu. Rozumiemy przez to postawienie hipotezy co do jawnej postaci przekształcenia $m_{\text{sys}}[\cdot]$ ze wzoru (1.1), z dokładnością do wartości parametrów, właściwych dla danej reprezentacji, co w efekcie pozwoli na dokonanie wstępnej selekcji modeli do zbioru \mathcal{M} . Czynnikiem, który ma decydujący wpływ na przyjętą formę reprezentacji jest poziom posiadanej wiedzy wstępnej na temat rozważanego systemu (dotyczącej np. jego struktury wewnętrznej, obecności nieliniowości i sprzężeń zwrotnych, długości pamięci, a także rozkładu prawdopodobieństwa sygnałów wejściowych) oraz dostępne narzędzia formalne.

Przykładowo, jeśli wiedza aprioryczna pozwala stwierdzić, że dany system jest liniowy, to standardowym sposobem jego opisu jest reprezentacja wykorzystująca sumę spłotową [66],

$$m_{\text{sys}}[\mathbf{x}_n] = T_{\text{lin}}[\mathbf{x}_n; \boldsymbol{\lambda}] = \sum_{i=0}^{\infty} \lambda_i x_{n-i}, \quad (1.5)$$

gdzie wektor $\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \lambda_2, \dots]$ przechowuje kolejne wartości odpowiedzi impulsowej systemu. Wzór ten jest właściwy dla *prawie wszystkich* systemów liniowych, z wyłączeniem pewnych patologicznych przypadków (zob. [90]), które z powodu niewielkiego znaczenia dla zastosowań inżynierskich wyłączamy z dalszej dyskusji. W konsekwencji, jeśli dodatkowo założymy, że system jest stabilny, wiedzę wstępną o systemie zawarlibyśmy w zbiorze $\mathcal{M} = \{T_{\text{lin}}[\cdot; \boldsymbol{\lambda}] : \|\boldsymbol{\lambda}\|_1 < \infty\}$.

Przejdźmy teraz do nieliniowych systemów *statycznych*, tj. takich w których charakterystyka spełnia warunek $m_{\text{sys}}[\mathbf{x}_n] = f_{\text{sys}}(x_n)$, gdzie $f_{\text{sys}} : \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją jednoargumentową. Odpowiedź na pytanie o reprezentację systemów tej klasy przynosi teoria aproksymacji funkcji. I tak na przykład, jeśli $f_{\text{sys}}(\cdot)$ jest *funkcją analityczną*, to można przedstawić ją za pomocą szeregu potęgowego [31],

$$f_{\text{sys}}(x) \simeq \sum_{p=0}^{\infty} c_p (x - z)^p, \quad (1.6)$$

gdzie c_p oznacza p -ty współczynnik rozwinięcia, a z jest stałą zwaną środkiem szeregu. Wartości współczynników wyznacza się według wzoru Taylora, $c_p = \frac{f^{(p)}(z)}{p!}$, przy czym należy pamiętać, że reprezentacja ta obowiązuje dla punktów znajdujących się w otoczeniu środka szeregu, bowiem przedział zbieżności szeregu (1.6) (nazywanego w tym wypadku szeregiem Taylora) ograniczony jest do $(z - r, z + r)$, gdzie r to tzw. promień zbieżności szeregu, dany wzorem

$$r = \frac{1}{\limsup_{p \rightarrow \infty} \sqrt[p]{|c_p|}}.$$

Z kolei, jeśli $f_{\text{sys}}(\cdot)$ należy do zbioru funkcji całkownych z kwadratem, tj. $f_{\text{sys}} \in L^2(\mathcal{X}) = \{f : \mathcal{X} \rightarrow \mathbb{R} : \int_{\mathcal{X}} f^2(x) dx < \infty\}$, gdzie $\mathcal{X} \subseteq \mathbb{R}$ jest pewnym niekoniecznie ograniczonym przedziałem, to możemy rozwinąć ją w szereg funkcji ortogonalnych. Reprezentacja systemu przyjmuje wtedy postać

$$f_{\text{sys}}(x) \simeq \sum_{p=0}^{\infty} c_p \varphi_p(x), \quad (1.7)$$

gdzie funkcje $\{\varphi_p(\cdot)\}$, stanowiące bazę ortogonalną danej przestrzeni, spełniają warunek

$$\langle \varphi_i, \varphi_j \rangle = \begin{cases} a_i, & \text{dla } i = j \\ 0, & \text{w przeciwnym wypadku,} \end{cases}$$

dla pewnych $a_i > 0$, a kolejne współczynniki rozwinięcia (1.7) wyznacza się według wzoru

$$c_p = \frac{\int_{\mathcal{X}} f(x)\varphi_p(x)dx}{\int_{\mathcal{X}} \varphi_p^2(x)dx}.$$

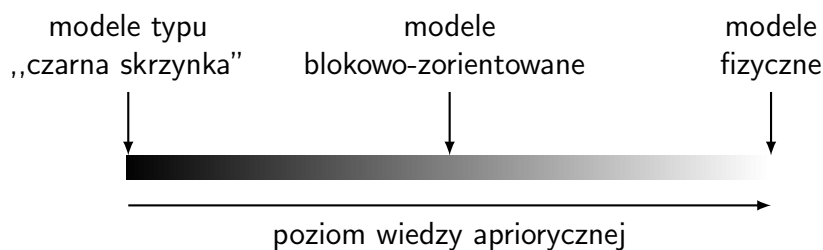
Ze względu na możliwość modelowania systemów nieciągłych, reprezentacja (1.7) doczekała się wielu zastosowań w algorytmach identyfikacji systemów [35, 37, 38]. Oprócz bazy trygonometrycznej (wykorzystywanej w szeregu Fouriera), przykładem często stosowanych baz ortogonalnych są bazy oparte o funkcje wielomianowe, takie jak wielomiany Legendre’a, Laguerre’a czy też Hermite’a, a także falki, [83, 102, 104]. Przedstawione reprezentacje systemów statycznych nie są jedynymi możliwymi, natomiast pełnią istotną rolę z punktu widzenia zagadnień poruszanych w dalszej części pracy.

1.2.1 Reprezentacja dynamicznych systemów nieliniowych

Przejdziemy teraz do problemu dużo bardziej złożonego, dotyczącego sposobów reprezentacji dynamicznych systemów nieliniowych. Jako jedno z podstawowych podejść do tego zagadnienia, wymienić można *modelowanie fizyczne*. Jest to ogólne określenie na zbiór technik, w których zakłada się dużą wiedzę aprioryczną dotyczącą procesów zachodzących w systemie. Modele fizyczne otrzymuje się w oparciu o podstawowe teorie fizyczne, korzystając na przykład z równań mechaniki lagranżowskiej i najczęściej wyraża za pomocą równań stanu. Modele tego typu cechują się wysoką interpretowalnością – ich struktura jest przejrzysta, a poszczególne parametry odnoszą się do fizycznych cech systemu, w związku z czym określa się je mianem „białych skrzynek” (ang. *white-box models*). W opisywanym podejściu ciężar modelowania przeniesiony jest na kodyfikację wiedzy wstępnej; typowo, zbiór \mathcal{M} jest jednoelementowy i w konsekwencji pozyskanie modelu nie wymaga użycia algorytmów modelowania. Jednocześnie, ze względu na wymagania dotyczące poziomu wiedzy wstępnej o systemie, zakres stosowalności tego podejścia bywa w praktyce mocno ograniczony.

Całkowicie odmienna idea stoi za modelowaniem wykorzystującym modele typu „czarna skrzynka” (ang. *black-box models*), w którym to podejściu zakłada się tylko niewielką wiedzę aprioryczną o systemie. W związku z tym, od potencjalnych modeli wymaga się własności *uniwersalnej aproksymacji*, tak aby ich matematyczna reprezentacja była odpowiednio elastyczna, a przez to właściwa dla możliwie szerokiej klasy systemów. Przykładem modeli posiadających tę cechę są np. sztuczne sieci neuronowe. W ostatnich latach zaobserwować mogliśmy gwałtowny rozwój metod opartych na tzw. uczeniu głębokim (ang. *deep learning*), czyli bazujących na wielowarstwowych sztucznych sieciach neuronowych. Choć koncepcja ta nie jest nowa, bowiem jej korzenie sięgają lat 40. ubiegłego wieku, [34], to szersze zainteresowanie zaczęła zyskiwać od momentu ukazania się przełomowej pracy Hintona *et al.* [45], w której zaproponowano efektywną metodę wyznaczania wag połączeń poszczególnych warstw sieci. Metody uczenia głębokiego szybko znalazły zastosowanie

w problemach tak skomplikowanych jak klasyfikacja obrazów, rozpoznawanie mowy, czy też analiza tekstu pisanego, wypierając przy tym rozwiązania alternatywne. Niestety, podejście to nie jest pozbawione wad. Estymacja parametrów głębokich sieci neuronowych wymaga bowiem ogromnych ilości danych oraz zasobów obliczeniowych, dlatego mimo znacznego postępu techniki komputerowej, w szczególności w dziedzinie przetwarzania równoległego, ich stosowanie może być bardzo drogie i czasochłonne. Ze strony teoretycznej, dodatkowe utrudnienie stanowi fakt, że estymacja parametrów modeli opartych o sieci neuronowe jest problemem niewypukłym, co znacznie utrudnia formalną analizę istniejących algorytmów, z kolei dobór właściwej struktury modelu (czyli architektury sieci) wciąż postrzega się bardziej jako wynik stosowania heurystyk, niż metody naukowej.

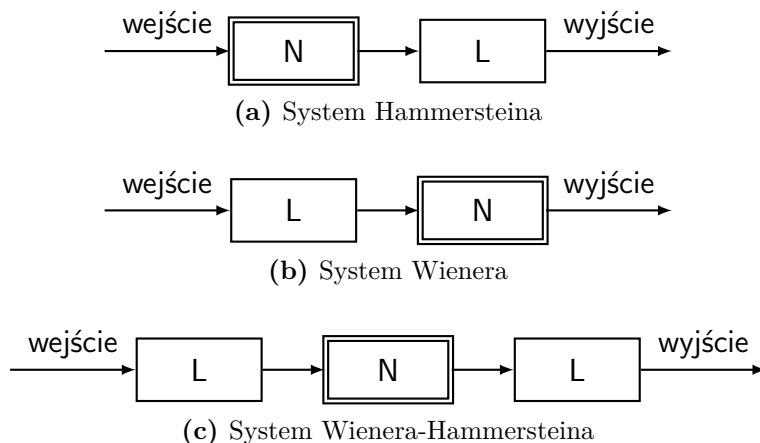


Rysunek 1.3: Uproszczona systematyka modeli otrzymywanych w ramach różnych podejść. Szczegółowy przegląd sposobów modelowania systemów nieliniowych przedstawiony został przez Ljunga i Schoukensa w pracach przekrojowych [67, 96]

Podane tu dwa sposoby postępowania znajdują się na przeciwnych krańcach spektrum możliwych podejść do problemu modelowania (rys. (1.3)). W praktyce, często wykorzystuje się rozwiązania pośrednie – z jednej strony bowiem założenie o posiadaniu pełnej wiedzy wstępnej o systemie jest idealizacją, nieznaną odzwierciedlenia w rzeczywistości, a z drugiej strony jako wyjątkową można traktować sytuację, w której pozbawieni jesteśmy jakichkolwiek przesłanek dotyczących procesów zachodzących w systemie.

W tym miejscu warto zwrócić uwagę na koncepcję systemów blokowo-zorientowanych [5, 75], która stanowi pewien kompromis pomiędzy technikami omówionymi powyżej (i stąd w literaturze nazywana jest podejściem opartym o „szarą skrzynkę”, ang. *grey-box*). Podłoże tego podejścia stanowi założenie, mówiące o tym, że modelowany system posiada ustaloną strukturę wewnętrzną i może być przedstawiony jako połączenie prostych bloków: liniowych podsystemów dynamicznych (L) i nieliniowych podsystemów statycznych (N). Dzięki takiej dekompozycji, zadanie modelowania sprowadzić można do estymacji poszczególnych bloków systemu (co pociąga za sobą konieczność estymacji sygnałów interakcyjnych, zazwyczaj niedostępnych dla pomiarów). Do najczęściej dyskutowanych modeli blokowych należą modele oparte o strukturę Hammersteina (inaczej NL, rys. 1.4a), strukturę Wienera (LN, rys. 1.4b) oraz strukturę Wienera-Hammersteina (LNL, rys. 1.4c). Ze względu na własności aproksymacyjne, na szczególne zainteresowanie zasługują struktury LN i LNL. Jak

bowiem wykazał Palm w [81], dowolny stacjonarny system ciągły, z czasem dyskretnym i o skończonej pamięci może być dowolnie dokładnie przybliżony za pomocą równoległego połączenia skończonej liczby systemów o strukturze LNL; podobny wynik, w odniesieniu do kaskad LN, odnaleźć możemy w pracy Korenberga [60]. Wyróżniona tu klasa modeli posiada własność uniwersalnej aproksymacji, dlatego można ją rozpatrywać także w kontekście modeli „czarnoskrzynkowych”.



Rysunek 1.4: Struktury systemów blokowo-zorientowanych

Niewątpliwą zaletą koncepcji systemów blokowo-zorientowanych jest fakt, że pomimo znacznych zdolności aproksymacyjnych, złożoność modeli jest dużo niższa, niż np. złożoność sztucznych sieci neuronowych, co pozwala na utrzymanie błędu estymacji na relatywnie niskim poziomie, nawet gdy liczba dostępnych pomiarów jest nieduża. Należy jednak podkreślić, że podejście oparte o reprezentacje blokowe, wymaga konstrukcji wysoce wyspecjalizowanych algorytmów, uwzględniających nie tylko strukturę systemu, ale także typ posiadanej wiedzy apriorycznej, która może mieć parametryczny, nieparametryczny lub mieszany charakter, zob. np. [42, 43, 74, 84]. Niestety, jeśli wiedza wstępna jest nieprawidłowa (tj. niezgodna z założeniami danej metody), zbieżność algorytmów nie może być dłużej zagwarantowana. W związku z powyższym, w warunkach niepewności, a także w sytuacjach gdy pozyskanie odpowiedniej wiedzy wstępnej jest niemożliwe lub zbyt kosztowne, uwaga kierowana jest w stronę jeszcze bardziej uniwersalnych reprezentacji systemu. Przejdziemy teraz do omówienia dwóch z nich, tj. reprezentacji opartej o szereg Volterry oraz szereg Wienera.

1.2.2 Szereg Volterry

W kontekście przyczynowych, stacjonarnych systemów nieliniowych z czasem ciągłym, szeregami Volterry określamy wyrażenia postaci

$$V[x(t)] = h_0 + \sum_{p=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} h_p(\tau_1, \dots, \tau_p) x(\tau_1) \cdots x(\tau_p) d\tau_1 \cdots d\tau_p, \quad (1.8)$$

nazwane tak na cześć Vito Volterra, matematyka żyjącego na przełomie XIX i XX wieku, którego znamy obecnie jako jednego z prekursorów analizy funkcjonalnej. Jedną z kluczowych idei Volterra było uogólnienie rozwinięcia Taylora na funkcjonały przekształcające funkcje rzeczywiste, do przedstawienia których używał właśnie szeregów postaci (1.8) (w nieco ogólniejszej formie, z dowolnymi granicami całkowania), [113]. Funkcjonały, które można rozwinąć w szereg Volterra określa się mianem *funkcjonałów analitycznych*, a jądra rozwinięcia, czyli funkcje $h_p(\cdot)$, wyznacza się podobnie jak współczynniki rozwinięcia Taylora [89]. Istotny wkład w rozwój teorii Volterra miał Fréchet, który sformułował twierdzenie analogiczne do twierdzenia Weierstrassa, wykazując, że rozwinięcie Volterra jest zbieżne jednostajnie do dowolnego funkcjonału ciągłego na zbiorze zwartym funkcji ciągłych [92]. Niestety, warunek dotyczący zwartości stanowił poważne ograniczenie [89], wykluczając m.in. funkcje określone na przedziale nieskończonym. Z tego względu, z punktu widzenia własności aproksymacyjnych, ważniejszy wynik uzyskany został przez Boyda i Chue [11], którzy przy dużo słabszych założeniach nałożonych na klasę sygnałów wejściowych (dotyczących ograniczoności sygnału i dodatkowo, w przypadku systemu z czasem ciągłym, jego gładkości), udowodnili, że za pomocą obciętych szeregów Volterra można dowolnie dokładnie aproksymować stacjonarne, przyczynowe systemy nieliniowe, o ile posiadają własność zwaną *pamięcią zanikającą* (ang. *fading memory*). Własność ta jest rozszerzeniem pojęcia ciągłości i stanowi formalne ujęcie idei wyrażanej już wcześniej m.in. przez Wienera, który rozważał systemy, których „wyjście jest asymptotycznie niezależne od wejścia z dalekiej przeszłości” [117].

Niestety, z reprezentacją opartą o szereg Volterra związane są pewne niedogodności. Po pierwsze, bezpośrednie wyznaczenie jąder rozwinięcia (1.8) możliwe jest tylko w przypadku systemów volterrowskich (czyli takich, które można przedstawić za pomocą funkcjonałów analitycznych), w dodatku dla ograniczonej, w sposób zależny od własności konkretnego systemu, amplitudy sygnału wejściowego. Z kolei, w przypadku nieanalitycznych systemów z pamięcią zanikającą, jesteśmy zmuszeni polegać na metodach przybliżonych. Trudność ta jest konsekwencją faktu, że w ogólności, z sygnału wyjściowego systemu nie można wydzielić składowych odpowiadających poszczególnym komponentom jąder rozwinięcia Volterra. Dodatkowo, zakres stosowalności reprezentacji Volterra jest dość ograniczony, bowiem wyniki odnoszące się do własności aproksymacyjnych dotyczą tylko systemów ciągłych, nie gwarantują zatem zbieżności do systemów o charakterystykach spoza powyższej klasy.

1.2.3 Szereg Wienera i rozwinięcie w chaos wielomianowy

Norbert Wiener nowatorsko wykorzystał szeregi Volterra w charakterze narzędzia analizy fizycznego systemu nieliniowego (zob. [94]). Jedną z jego najważniejszych koncepcji było osadzenie teorii systemów w kontekście probabilistycznym. Pierwotnie bowiem, Wiener zajmował się całkowymi przekształceniami ruchu Browna (zwanego dzisiaj również *procesem Wienera*), by następnie, opierając się o wyniki Camerona i Martina [16], objąć badaniami także systemy nieliniowe (zob.

[117]). Wiener zaproponował, by do reprezentacji systemów nieliniowych posłużyć się szeregiem postaci

$$F[x(t)] = g_0 + \sum_{p=1}^{\infty} G_p [x(t); g_p], \quad (1.9)$$

gdzie G_p są funkcjonalami powstałymi w wyniku ortogonalizacji szeregu Volterry względem losowego wejścia będącego gaussowskim białym szumem (w skrócie GBS-em). W praktyce, oznacza to, że przy założeniu, że $x(t)$ jest GBS-em, dla dowolnych jąder g_p funkcjonały G_p są ze sobą nieskorelowane. Korzystając z faktu, że w systemach przyczynowych tzw. klasy Wienera, jądra g_p są określone na półprostej $[0, \infty)$ oraz całkwalne z kwadratem, można rozwinąć je w szereg Laguerre'a, $g_p(\tau_1, \dots, \tau_p) = \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} c_p(i_1, \dots, i_p) L_{i_1}(\tau_1) \dots L_{i_p}(\tau_p)$, gdzie L_k oznacza wielomian Laguerre'a k -tego rzędu. W konsekwencji, jak pokazuje np. Schetzen w [92], funkcjonały reprezentacji (1.9) można przedstawić za pomocą wzoru

$$G_p [x(t); g_p] = \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} c_p(i_1, \dots, i_p) \Psi_{(i_1, \dots, i_p)}(\mathbf{v}(t)), \quad (1.10)$$

gdzie $\Psi_{(i_1, \dots, i_p)}(\cdot)$ są ortogonalnymi względem GBS-u funkcjonalami, powstałymi przez wymnożenie wielomianów Hermite'a (zagadnienie to zostanie szerzej omówione w ustępie 3.2), a wektor $\mathbf{v}(t) = [v_0(t), v_1(t), \dots]$ przechowuje wartości całek splotowych wejścia systemu z kolejnymi wielomianami Laguerre'a, tj. $v_k(t) = \int_0^{\infty} x(t - \tau) L_k(\tau) d\tau$. Podejście takie ułatwia analizę własności statystycznych szeregu Wienera, a także sprawia, że jądra funkcjonałów G_p można wyrazić za pomocą przeliczalnego zbioru parametrów $\{c_p(i_1, \dots, i_p)\}$, zamiast funkcji $g_p(\tau_1, \dots, \tau_p)$.

Reprezentacja wienerowska, w kontekście zarówno systemów z czasem ciągłym, jak i dyskretnym, właściwa jest dla tzw. *klasy Wienera* systemów nieliniowych (pojęcia tego nie należy mylić z systemami o strukturze Wienera, LN), tj. takich, których odpowiedź na pobudzenie GBS-em ma skończoną wariancję. Dzięki temu, wyjście systemu należącego do klasy Wienera można rozpatrywać jako element przestrzeni Hilberta zmiennych losowych ze skończoną wariancją, co znacznie ułatwia formalną analizę ich własności. Przede wszystkim, w przestrzeni takiej można wyróżnić bazę ortogonalną (na którą, *notabene*, składają się składowe funkcjonałów G_p , czyli funkcje $\Psi_{(i_1, \dots, i_p)}(\cdot)$), dzięki czemu jądra rozwinięcia Wienera dają się wyznaczyć w sposób analityczny, za pomocą rzutowania ortogonalnego – obserwacja ta jest punktem wyjścia do konstrukcji algorytmu korelacyjnego [63]. Ponadto, rozszerzanie modelu o kolejne funkcjonały nie powoduje zmiany wartości współczynników rozwinięcia, co może być szczególnie przydatne w rekurencyjnych algorytmach modelowania. Co najważniejsze, można dowieść, że szereg Wienera zbiega do wyjścia systemu w sensie średniokwadratowym [28, 49], a co za tym idzie, charakterystyki systemów klasy Wienera, w przeciwieństwie do systemów z pamięcią zanikającą, nie muszą spełniać warunku ciągłości. Nie bez znaczenia pozostaje także fakt, że wyniki dotyczące własności aproksymacyjnych szeregów Wienera są *kompatybilne* z przyjętą przez nas kwadratową funkcją straty.

Wykorzystanie GBS-u do badania systemów nieliniowych ma głębokie uzasadnienie teoretyczne. Przede wszystkim, jak zauważają m.in. Marmarelis i Marmarelis [70], z faktu, że widmo białego szumu jest stałe w pełnym zakresie częstotliwości, wynika, że istnieje niezerowe prawdopodobieństwo, iż dowolna forma sygnału (na przedziale skończonym) będzie dowolnie dokładnie reprezentowana przez pewną realizację szumu białego. Oznacza to, że na podstawie eksperymentu, w którym badany system jest przez odpowiednio długi czas pobudzany GBS-em, jesteśmy w stanie (w przybliżeniu) scharakteryzować odpowiedź tego systemu na dowolne pobudzenie (zob. także [82]). Można więc uznać, że w systemach nieliniowych, GBS pełni rolę analogiczną do roli skoku jednostkowego w systemach liniowych. Należy jednak mieć na uwadze, że zaproponowana przez Wienera reprezentacja pozostaje konstruktem czysto teoretycznym, gdyż z faktu, iż energia GBS-u jest nieskończona, wynika iż proces ten jest fizycznie nierealizowalny. W praktyce, wymusza to użycie pobudzenia z widmem ograniczonym, które może być zrealizowane np. za pomocą szumu termicznego.

Mimo, że pierwotnie ograniczona do GBS-u, teoria Wienera z biegiem lat została rozszerzona na inne klasy sygnałów wejściowych. Schetzen przedstawił reprezentację ortogonalną oraz metodę wyznaczania jąder dla wejścia będącego kolorowym szumem gaussowskim [91], a także sygnałów będących wynikiem przekształcenia GBS-u za pomocą filtru o odwracalnej charakterystyce znanej postaci [92]. Z kolei Bose zaproponował modyfikację szeregu Wienera, w której poszczególne funkcjonały (1.10) wyrażone są za pomocą funkcji prostokątnych [8], dzięki czemu są losowo niezależne dla dowolnego sygnału wejściowego. Niestety, rozwiązanie to jest obliczeniowo nieefektywne i nie doczekało się wielu zastosowań. Jeszcze inny pomysł przedstawiony został w pracy Korenberga *et al.* [61], w której zaprezentowano algorytm *offline* estymacji modelu, którego poszczególne operatory składowe były zortogonalizowane ze względu na realizację sygnału wejściowego.

W przypadku gdy znany jest rozkład prawdopodobieństwa sygnału wejściowego i nie jest on GBS-em, ale spełnia warunki ustalone w pracy Ernsta *et al.* [28], do badania systemu możemy posłużyć się metodą *rozwinęcia w chaos wielomianowy* (ang. *polynomial chaos expansion*; w skrócie PCE), zob. [57], która stanowi uogólnienie podejścia Wienera. Formalnie, jeśli wyjście systemu, Y , jest zmienną losową mającą skończoną wariancję i będącą przekształceniem pewnego przeliczalnego zbioru zmiennych losowych X (co w szczególności odnieść można do systemów z czasem dyskretnym), to można rozwinąć ją w szereg postaci

$$Y = \sum_{\alpha} c_{\alpha} \Psi_{\alpha}(\mathbf{X}),$$

gdzie Ψ_{α} są wielowymiarowymi wielomianami, ortogonalnymi względem rozkładu $\mathcal{D}_{\mathbf{X}}$.

1.3 Modele wielomianowe i problem wielowymiarowości

W pracy będziemy wykorzystywali modele oparte o reprezentację Volterry/Wienera, zwane dalej *modelami wielomianowymi*, w których długości pamięci M , oraz rząd P , mają skończone wartości. Jak wykażemy w dalszej części pracy, modele takie przedstawić można za pomocą ogólnego wzoru

$$m_{S_{P,M}}[\cdot; \boldsymbol{\theta}] = \sum_{i=1}^D \theta_i \varphi_i[\cdot], \quad (1.11)$$

będącego liniową kombinacją funkcjonałów $\varphi_1, \dots, \varphi_D$, stanowiących unikalne elementy zbioru $S_{P,M}$, zwanego dalej *słownikiem*, którego zawartość zależy od wykorzystywanej w danym wypadku reprezentacji. W modelach wielomianowych, rozmiar słownika, D , uzależniony jest od przyjętej długości pamięci oraz rzędu modelu i wyraża się wzorem $D = \binom{P+M}{M} = \frac{(P+M)!}{P!M!}$ [53]. Jak można zaobserwować, wzrost D postępuje bardzo szybko wraz ze wzrostem wartości parametrów P i M , co jest źródłem nie tylko wysokich kosztów obliczeniowych związanych ze stosowaniem metod opartych o modele wielomianowe, ale przede wszystkim, częstokroć nierealistycznych wymagań dotyczących wielkości zbioru uczącego. Ten ostatni problem jest przejawem zjawiska nazywanego szerzej „przekleństwem wielowymiarowości” (ang. *curse of dimensionality*), wynikającego z faktu, że w wielowymiarowej przestrzeni, odległość między punktami pomiarowymi może być bardzo duża (zob. [41, rozdział 2.2]). Należy więc uznać, że największym wyzwaniem związanym ze stosowaniem modeli wielomianowych jest ograniczenie wpływu tego niekorzystnego zjawiska przy jednoczesnym zachowaniu możliwości aproksymacyjnych wykorzystywanej reprezentacji.

1.4 Estymacja modeli wielomianowych

Wprowadźmy oznaczenie $L_S(\boldsymbol{\theta}) = L(m_S[\cdot; \boldsymbol{\theta}])$. Dla klasy $\mathcal{M}_S = \{m_S[\cdot; \boldsymbol{\theta}] : \boldsymbol{\theta} \in \mathbb{R}^D\}$ modeli opartych o ustalony słownik S , zadanie estymacji sprowadza się do znalezienia wektora parametrów $\hat{\boldsymbol{\theta}}$ minimalizującego błąd estymacji

$$\Delta_{\text{est.}} = \mathbb{E}\{L(\hat{m})\} - L(m^*) = \mathbb{E}\{L_S(\hat{\boldsymbol{\theta}})\} - L_S(\boldsymbol{\theta}^*), \quad (1.12)$$

gdzie

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} L_S(\boldsymbol{\theta}) \quad (1.13)$$

jest wektorem parametrów najlepszego możliwego modelu $m^*[\cdot] = m_S[\cdot; \boldsymbol{\theta}^*]$. Z uwagi na fakt, iż w ogólnym przypadku nie mamy dostępu do funkcjonału $m_{\text{sys}}[\cdot]$ oraz rozkładu prawdopodobieństwa sygnału wejściowego, zadania tego nie możemy rozwiązać w sposób analityczny i zmuszeni jesteśmy polegać na metodach przybliżonych.

1.4.1 Metoda najmniejszych kwadratów

Typowe podejście do powyższego problemu opiera się na minimalizacji *empirycznej funkcji straty* [112], którą uzyskuje się w oparciu o zbiór obserwacji \mathcal{S} . W przypadku kwadratowej funkcji straty (1.3), której empiryczny odpowiednik przyjmuje postać

$$\hat{L}_S(\boldsymbol{\theta}) = \frac{1}{N - M + 1} \sum_{i=M}^N (m_S[\mathbf{X}_i; \boldsymbol{\theta}] - Y_i)^2, \quad (1.14)$$

opisane podejście określa się *metodą najmniejszych kwadratów* (ang. *least squares method*, w skrócie LS). Formalnie, zadanie estymacji polega więc na rozwiązaniu następującego problemu optymalizacyjnego

$$\hat{\boldsymbol{\theta}}^{\text{LS}} = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} \hat{L}_S(\boldsymbol{\theta}) = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} \|\boldsymbol{\phi}\boldsymbol{\theta} - \mathbf{Y}\|_2^2, \quad (1.15)$$

gdzie

$$\mathbf{Y} = [Y_M, Y_{M+1}, \dots, Y_N]^T$$

oraz

$$\boldsymbol{\phi} = \begin{bmatrix} \varphi_1(\mathbf{X}_M) & \dots & \varphi_D(\mathbf{X}_M) \\ \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{X}_N) & \dots & \varphi_D(\mathbf{X}_N) \end{bmatrix}$$

są odpowiednio wektorem przechowującym wyjścia systemu oraz macierzą regresorów. W tym miejscu należy zwrócić uwagę, że ze względu na pamięć modelu (i nieznaną wartość warunków początkowych), konstrukcja funkcji \hat{L}_S (a zatem i estymatora $\hat{\boldsymbol{\theta}}^{\text{LS}}$) wymaga odrzucenia $M - 1$ pierwszych par pomiarowych.

Jedną z istotniejszych własności metody jest to, że stowarzyszony problem optymalizacyjny (1.15) jest wypukły, a zarazem posiada rozwiązanie analityczne. Jeśli bowiem macierz $\boldsymbol{\phi}^T \boldsymbol{\phi}$ jest odwracalna, to estymator $\hat{\boldsymbol{\theta}}^{\text{LS}}$ przyjmuje postać

$$\hat{\boldsymbol{\theta}}^{\text{LS}} = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1} \boldsymbol{\phi}^T \mathbf{Y},$$

natomiast w przeciwnym razie można skorzystać z przekształcenia zwanego pseudoinwersją Moore'a-Penrose'a [12]. Ma to szczególne znaczenie w przypadku, gdy liczba elementów słownika przewyższa liczbę dostępnych punktów pomiarowych, $D > N - M + 1$, a więc gdy problem (1.15) jest niedookreślony, a zatem nie posiada unikalnego rozwiązania – wówczas, ze zbioru rozwiązań problemu zostaje zwrócone to, którego norma euklidesowa jest najmniejsza. Dokładna analiza własności estymatora $\hat{\boldsymbol{\theta}}^{\text{LS}}$ została przedstawiona m.in. w [85, rozdział 4] i [105, rozdział 4]. Ze względu na liczne zalety, metoda LS stała się podstawą wielu parametrycznych i parametryczno-nieparametrycznych metod identyfikacji systemów.

1.4.2 Regularyzacja

Niestety, w przypadku modeli typu *black-box* (a więc, w szczególności, modeli wielomianowych), zakres stosowalności metody LS bywa mocno ograniczony. Wskazują na to zarówno rezultaty empiryczne [62], jak i teoretyczne – np. w pracy [110] wykazano, że jeśli zakłócenie ma rozkład normalny, to dla systemów statycznych błąd estymacji metody jest nie mniejszy niż $c\frac{D}{N}$, dla pewnej stałej $c > 0$. W związku z tym, gdy wielkość zbioru danych jest stosunkowo niewielka ($N \ll D$), czy też umiarkowana ($N \approx D$), należy spodziewać się wysokiego błędu estymacji metody. Istnieje kilka rozwiązań tego problemu. Oprócz oczywistych, takich jak wykonanie dodatkowych pomiarów wejścia-wyjścia systemu (co może być kosztowne i nie zawsze jest możliwe), czy też redukcja liczby elementów słownika (co z kolei może niekorzystnie wpłynąć na błąd aproksymacji metody), dla polepszenia własności estymatora często korzysta się z techniki zwanej *regularyzacją*. Polega ona na wyznaczeniu parametrów modelu poprzez jednoczesną minimalizację empirycznej funkcji straty (1.14) oraz pewnej nieujemnej funkcji regularyzującej $\Omega(\boldsymbol{\theta})$, której rolą jest wprowadzenie *kary*, proporcjonalnej do stopnia szeroko rozumianej *złożoności* wektora parametrów $\boldsymbol{\theta}$. W konsekwencji, metody wykorzystujące regularyzację faworyzują modele o prostszej strukturze. Odpowiednio dobrana funkcja regularyzująca zmniejsza wariancję estymatora, zapobiegając tym samym zjawisku przeuczenia i stanowi element konieczny w niektórych problemach nauczania maszynowego [100], a jej postać może wynikać ze szczególnej wiedzy wstępnej o modelowanym systemie lub też preferencji o bardziej ogólnym charakterze, związanych np. z własnościami statystycznymi danego estymatora. Formalnie, w przypadku zastosowania regularyzacji, estymacja modelu sprowadza się do rozwiązania następującego problemu optymalizacyjnego

$$\hat{\boldsymbol{\theta}}^\Omega = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} \hat{L}_S(\boldsymbol{\theta}) + \lambda \Omega(\boldsymbol{\theta}), \quad (1.16)$$

gdzie $\lambda = [0, \infty)$ jest parametrem określającym wagę kary $\Omega(\boldsymbol{\theta})$. Z problemem tym silnie związane jest zagadnienie minimalizacji funkcji $\hat{L}_S(\boldsymbol{\theta})$ z ograniczeniami nałożonymi na wektor $\boldsymbol{\theta}$, tj.

$$\hat{\boldsymbol{\theta}}^C = \arg \min_{\boldsymbol{\theta} \in C} \hat{L}_S(\boldsymbol{\theta}), \quad C = \{\boldsymbol{\theta} \in \mathbb{R}^D : \Omega(\boldsymbol{\theta}) \leq c, c > 0\}. \quad (1.17)$$

Bezpośredni związek pomiędzy problemami (1.16) i (1.17) wynika z faktu, że lagranżjan problemu (1.17) jest dany wzorem $\mathcal{L}(\lambda) = \hat{L}_S(\boldsymbol{\theta}) + \lambda \Omega(\boldsymbol{\theta}) - \lambda c$ – z warunku Slatera [12, rozdział 5] wynika więc, że jeśli funkcja $\Omega(\cdot)$ jest wypukła, a rozwiązanie (1.17) jest unikalne, to dla pewnego $\lambda^* \geq 0$ zachodzi $\hat{\boldsymbol{\theta}}^C = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} \hat{L}_S(\boldsymbol{\theta}) + \lambda^* \Omega(\boldsymbol{\theta})$. Relacja pomiędzy parametrami c i λ^* jest w ogólności niemożliwa do ustalenia [111], niemniej jednak, powyższa analiza dowodzi, że estymatory (1.16), (1.17) można w pewnych warunkach uznać za tożsame. Z tego względu, mówiąc o regularyzacji, będziemy mieli na myśli zarówno problemy typu (1.16), jak i (1.17). Nadmienimy jednak, że przewagą sformułowania problemu za pomocą wzoru (1.17)

jest bezpośrednio ukazanie wpływu jakie regularyzacja wywiera na możliwości aproksymacyjne metody. Jak można bowiem zauważyć, nałożenie ograniczeń na zbiór dopuszczalnych parametrów $\boldsymbol{\theta}$ sprawia, że zbiór modeli hipotetycznych zostaje w sposób jawny ograniczony do $\mathcal{M}_{S,C} = \{m_S[\cdot; \boldsymbol{\theta}] : \boldsymbol{\theta} \in C\} \subseteq \mathcal{M}_S$ (w konsekwencji zmienia się także definicja wektora parametrów najlepszego modelu, który dany jest teraz przez $\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in C} L_S(\boldsymbol{\theta})$ (por. wzór (1.13)). Wiąże się to oczywiście z potencjalnym wzrostem błędu aproksymacji, równocześnie jednak, jak pokażemy dalej, działanie takie może w sposób znaczący zredukować błąd estymacji metody.

Regularyzacja ℓ_2

W problemach regresji liniowej, do których zalicza się estymację modeli wielomianowych, jednymi z najczęściej rozważanych są funkcje regularyzujące oparte o (pseudo)normy ℓ_p , tj. $\Omega(\boldsymbol{\theta}) = \|\boldsymbol{\theta}\|_p^p = \sum_{i=1}^D |\theta_i|^p$. Podręcznikowym przykładem jest regularyzacja ℓ_2 , bazująca na normie euklidesowej, w kontekście estymacji modeli statystycznych zaproponowana po raz pierwszy przez Hoerla i Kennarda [46]. Pewna modyfikacja tej metody, wykorzystująca funkcję regularyzującą $\Omega(\boldsymbol{\theta}) = \|\mathbf{W}\boldsymbol{\theta}\|_2^2$, została zastosowana do estymacji modeli volterrowskich w pracach [6, 7], w których diagonalna macierz wag \mathbf{W} wyrażała przypuszczenie dotyczące zanikania i „gładkości” komponentów składających się na jądra Volterra. Niestety, jak dowodzi Hsu *et al.* [48], błąd estymacji metod wykorzystujących regularyzację ℓ_2 jest tego samego rzędu co błąd estymacji metody LS, $\mathcal{O}(\frac{D}{N})$, tak więc, ze względu na powody wymienione wcześniej, w rozważanym problemie zasadność stosowania regularyzacji tego typu jest dyskusyjna.

Regularyzacja ℓ_0

W ostatnich latach dużą uwagę przyciągają metody bazujące na funkcjach regularyzujących postaci $\Omega(\boldsymbol{\theta}) = \lim_{p \rightarrow 0^+} \|\boldsymbol{\theta}\|_p^p = \#\{i : \theta_i \neq 0\}$, opartych o pseudonormę ℓ_0 , której wartość równa jest liczbie niezerowych komponentów wektora $\boldsymbol{\theta}$ [13]. Regularyzacja tego typu wymusza *rzadkość* wektora parametrów $\boldsymbol{\theta}$, co w efekcie ogranicza liczbę stopni swobody rozważanej klasy modeli [86, str. 8]. Korzenie tego podejścia sięgają dziedziny przetwarzania sygnałów, gdzie idea rzadkiej reprezentacji jest fundamentem, na którym opierają się algorytmy oszczędnego próbkowania (ang. *compressed sensing*), zaproponowane m.in. przez Candès *et al.* [18] i Donoho [22]. Ze względu na atrakcyjne własności, metody te znalazły zastosowanie także w innych gałęziach nauki [121], przyczyniając się m.in. do powstania wielu nowych algorytmów estymacji modeli wielomianowych [53, 68]. Można wykazać, że dla systemów statycznych, błąd estymacji modeli utworzonych w oparciu o estymator (1.16) z regularyzacją ℓ_0 jest rzędu $\mathcal{O}(1 + \frac{R}{N} \ln(\frac{D}{R} + 1))$ [15], gdzie $R \leq D$ stanowi górne ograniczenie normy ℓ_0 wektora $\boldsymbol{\theta}^*$. Tak więc, jeśli badany system znajduje rzadką (albo w przybliżeniu rzadką) reprezentację w słowniku S , zależność całościowego

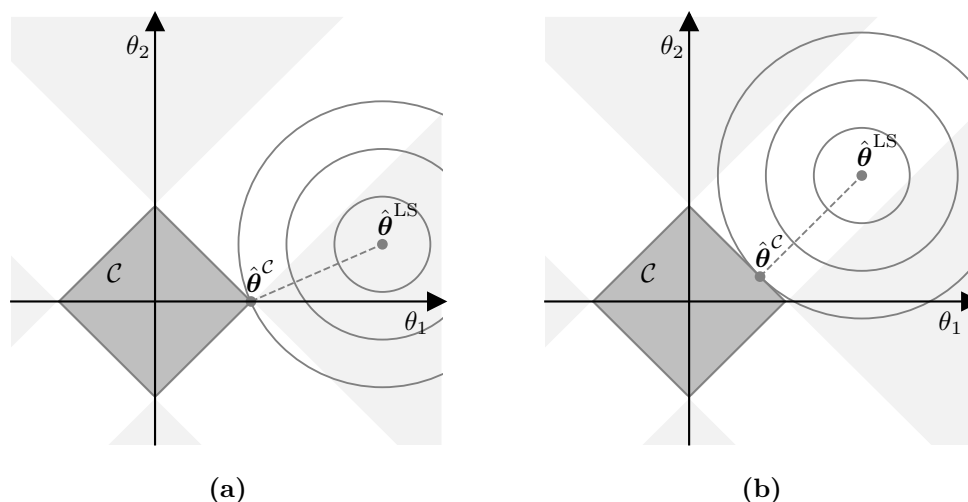
błędu metody od rozmiaru słownika staje się logarytmiczna, co drastycznie niweluje skutki „przekleństwa wielowymiarowości”. Niestety jednak, jako że pseudonorma ℓ_0 jest funkcją niewypukłą, estymatory oparte o regularyzację tego typu nie mogą zostać wyznaczone przy użyciu metod optymalizacji wypukłej, a rozwiązanie problemów (1.16), (1.17) jest NP-trudne [76]. W związku z tym, w praktyce stosowane są algorytmy zachłanne, które w każdej iteracji zwracają rozwiązanie lokalnie optymalne (zob. np. [2, rozdział 7]); w pracach [109, 120] omówiono warunki, w których metody tego typu osiągają rozwiązania globalnie optymalne.

Regularyzacja ℓ_1

Alternatywę dla metod wykorzystujących regularyzację ℓ_0 , stanowią estymatory znane pod nazwą Lasso (ang. *least absolute shrinkage and selection operator*) [108], w których funkcja regularyzująca oparta jest o normę ℓ_1 . Z uwagi na wypukłość normy ℓ_1 (jest to „najmniejsza” wypukła funkcja spośród (pseudo)norm ℓ_p), estymatory tego typu można wyznaczyć przy użyciu ogólnych metod programowania wypukłego, np. metody punktu środkowego, [79]. Istnieją również metody specjalnie do tego celu przystosowane [58, 95], nierzadko stosowane są też metody przybliżone [27, 33]. Estymatory z regularyzacją ℓ_0 i ℓ_1 łączy kilka podobieństw. Po pierwsze, można wykazać, że w pewnych warunkach osiągają to samo rozwiązanie [13]. Po drugie, podobnie jak regularyzacja ℓ_0 , także i regularyzacja ℓ_1 wymusza rzadkość modelu, choć w nieco bardziej łagodny sposób.

Rozważmy przypadek, gdy macierz regresorów ϕ jest ortonormalna. Nietrudno pokazać, że estymator (1.17) z ograniczeniem ℓ_1 przyjmuje wtedy postać, $\hat{\theta}_i^c = \text{sgn}(\hat{\theta}_i^{\text{LS}}) \max(|\hat{\theta}_i^{\text{LS}}| - \lambda^*, 0)$, $i = 1, \dots, D$ [108]. Widzimy więc, że zastosowanie kary ℓ_1 powoduje przesunięcie składowych wektora $\hat{\theta}^{\text{LS}}$ w kierunku zera, część z nich w rzeczy samej zerując. Interpretację geometryczną tej własności przedstawiono na rys. 1.5.

Chociaż rzadkość modelu postrzega się w niektórych zastosowaniach jako wartość samą w sobie (np. ze względu na niski koszt ewaluacji modeli rzadkich [68]), z punktu widzenia niniejszych rozważań bardziej interesujący jest jej wpływ na własności estymatora. Okazuje się, że tak jak estymatory oparte o normę ℓ_0 , tak i estymatory typu Lasso w sposób znaczący łagodzą problem „przekleństwa wielowymiarowości”. Nie czyniąc założeń o rzadkości reprezentacji systemu, Juditski i Nemirovski wykazali, że górne oszacowanie błędu estymacji estymatora (1.17) z ograniczeniem ℓ_1 jest rzędu $\mathcal{O}(\sqrt{\frac{\ln D}{N}})$ [51]. Oznacza to, że wpływ wielkości słownika na błąd estymatora jest porównywalny, a nawet mniejszy, niż w przypadku regularyzacji ℓ_0 . Jednocześnie jednak, dla ustalonego D szybkość zbieżności błędu metody z ograniczeniem ℓ_1 jest znacznie mniejsza i wynosi $\frac{1}{\sqrt{N}}$ ($\frac{1}{N}$ w przypadku regularyzacji ℓ_0). Zależność tę można uzasadnić nieparametryczną wiedzą wstępną o systemie [116].



Rysunek 1.5: Ilustracja działania metody LS z ograniczeniami ℓ_1 , dla ortonormalnej macierzy regresorów. Na wykresach zaznaczono poziomice empirycznej funkcji straty. Oprócz tego, ciemnoszarym kolorem zaznaczono zbiór ograniczeń (będący kulą w ℓ_1 , $\mathcal{C} = \{\boldsymbol{\theta} \in \mathbb{R}^D : \|\boldsymbol{\theta}\|_1 \leq c\}$, dla pewnego $c > 0$), a kolorem jasnoszarym zbiór rozwiązań standardowej metody LS mających tę własność, że ich odpowiednik z ograniczeniem ℓ_1 ma postać rzadką (przypadek ten przedstawiono na rys. a). W pozostałych przypadkach, zastosowanie ograniczeń nie prowadzi do rozwiązania rzadkiego (rys. b). Nie przedstawiono trywialnego przypadku, gdy $\|\hat{\boldsymbol{\theta}}^{\text{LS}}\|_1 \leq c$, tj. gdy estymatory $\hat{\boldsymbol{\theta}}^{\text{LS}}$ i $\hat{\boldsymbol{\theta}}^{\text{C}}$ są tożsame

Metoda najmniejszych kwadratów z ograniczeniami w zadaniu estymacji systemów dynamicznych

Podkreślmy raz jeszcze, że omówione do tej pory rezultaty odnoszą się do systemów statycznych – ze względu na trudności jakie niesie za sobą analiza własności estymatorów w przypadku gdy dane pomiarowe są skorelowane, w literaturze niewiele jest wyników teoretycznych dotyczących błędu estymacji nieliniowych systemów dynamicznych. Jednym z nich jest praca Wachla i Śliwińskiego [116], w której autorzy przedstawili oszacowanie błędu estymacji metody LS z ograniczeniem ℓ_1 dla systemów z pamięcią skończoną. Następnie rezultat ten został rozszerzony przez Wachla [114, 115] na klasę systemów z pamięcią zanikającą (wyniki te podsumowano w tabeli 1.1). Z kolei w [103], Śliwiński *et al.* rozważyli bardziej ogólne ograniczenia typu ℓ_p , dla $p \geq 1$, w kontekście estymacji modeli volterrowskich. Otrzymane rezultaty formalnie potwierdzają przydatność zastosowania ograniczeń typu ℓ_1 w estymacji modeli dynamicznych, a przy tym ukazują wpływ jaki wywierają skorelowane dane pomiarowe na wielkość błędu estymacji. W tym miejscu należy jednak nadmienić, że autorzy wymienionych prac skupiają się na własnościach estymatorów *teoretycznych*, nie omawiając jednak sposobów ich wyznaczania. W przypadku dużych słowników

i zbiorów danych, ze względu na złożoność obliczeniową standardowych metod optymalizacyjnych, ich zastosowanie może okazać się w praktyce problematyczne.

Klasa systemów	Rząd błędu estymacji $\Delta_{\text{est.}}$
Statyczne	$\sqrt{N^{-1} \ln D}$ [51]
Z pamięcią skończoną (długości M)	$\sqrt{(N - M)^{-1} M \ln D}$ [116]
Z pamięcią zanikającą eksponencjalnie	$\sqrt{N^{-1} \lceil \ln N \rceil \ln D} + N^{-c_1}$ [114]

Tabela 1.1: Porównanie górnych oszacowań błędów estymacji metody LS z ograniczeniem ℓ_1 dla różnych klas systemów. Pamięć zanikająca eksponencjalnie oraz stała c_1 zostały zdefiniowane w rozdziale 4 (definicja 4.1)

1.5 Metody optymalizacji stochastycznej pierwszego rzędu

Jako alternatywę wobec metod omówionych do tej pory, w niniejszej pracy do celów estymacji modeli wielomianowych proponuje się wykorzystanie stochastycznych metod optymalizacji wypukłej (ang. *stochastic convex optimization*). Inaczej niż w podejściu, w którym estymator minimalizuje empiryczną funkcję straty (regularyzowaną, bądź nie), celem metod stochastycznych jest bezpośrednia minimalizacja *prawdziwej* funkcji straty, $L_S(\boldsymbol{\theta})$. Wprowadzając dodatkowo ograniczenia na zbiór parametrów modelu, metody te zastosujemy do rozwiązania problemu

$$\min_{\boldsymbol{\theta} \in \mathcal{C}} L_S(\boldsymbol{\theta}), \quad (1.18)$$

gdzie \mathcal{C} jest pewnym wypukłym i domkniętym zbiorem ograniczeń.

Dyskusję zawężymy do algorytmów optymalizacji pierwszego rzędu, w których do poszukiwania rozwiązania problemu (1.18) wykorzystuje się zakłócony gradient funkcji straty, otrzymany na podstawie danych pomiarowych systemu. Dodatkowo, na potrzeby tego wprowadzenia, założymy że zarówno system, jak i model jest statyczny. Należy również zaznaczyć, że podrozdział ten nie ma na celu szczegółowego zapoznania z metodami optymalizacji stochastycznej, w związku z czym detale techniczne zostały pominięte. Szczegółowa analiza wybranych metod przeprowadzona została m.in. w [9, 14, 101].

1.5.1 Postać ogólna metod pierwszego rzędu

Algorytmy optymalizacji pierwszego rzędu są algorytmami iteracyjnymi, które w każdym powtórzeniu t wyznaczają tzw. *rozwiązanie pośrednie* $\boldsymbol{\theta}_{t+1}$, wykorzystując

do tego celu gradient $\mathbf{g}_t = \nabla \hat{L}_t(\boldsymbol{\theta}_t)$ funkcji

$$\hat{L}_t(\boldsymbol{\theta}) = (Y_t - m_S[X_t; \boldsymbol{\theta}])^2, \quad (1.19)$$

stanowiącej szczególną formą empirycznej funkcji straty, utworzonej na podstawie *jednej* pary pomiarowej $\{X_t, Y_t\}$. Gradient funkcji (1.19), jest D -wymiarowym wektorem, określonym przez wzór

$$\nabla \hat{L}_t(\boldsymbol{\theta}) = 2(m_S[X_t; \boldsymbol{\theta}] - Y_t) [\varphi_1(X_t), \dots, \varphi_D(X_t)]^T, \quad (1.20)$$

pozwalającym ustalić kierunek spadku prawdziwej funkcji straty. Ze względu na zakłócenia, informacja ta ma charakter przybliżony, jednak w przypadku algorytmów stochastycznych wystarczy, żeby była ona *średnio* poprawna [14, rozdział 6], tj. aby dla każdego $\boldsymbol{\theta}$, wyrażenie $\nabla \hat{L}_t(\boldsymbol{\theta})$ było nieobciążonym estymatorem prawdziwego gradientu $\nabla L_S(\boldsymbol{\theta})$, t.j. $\mathbb{E}\{\nabla \hat{L}_t(\cdot)\} = \nabla L_S(\cdot)$. W przypadku funkcji (1.20) warunek ten jest spełniony, gdy zakłócenie ma zerową wartość oczekiwaną i jest niezależne od wejścia systemu. Rozwiązanie końcowe algorytmu, $\hat{\boldsymbol{\theta}}$, stanowi zazwyczaj średnia arytmetyczna z rozwiązań pośrednich $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{N+1}$. Schematyczne działanie metod stochastycznych przedstawia algorytm 1.1.

Zauważmy, że w rozważanej klasie algorytmów każda para pomiarowa jest przetwarzana tylko raz, tak więc liczba iteracji algorytmu równa jest wielkości zbioru danych pomiarowych. Często przekłada się to na liniową względem D i N złożoność obliczeniową algorytmu, bowiem przy odpowiednim doborze metody, główną trudnością obliczeniową każdej iteracji staje się wyznaczenie gradientu \mathbf{g}_t . W konsekwencji, złożoność obliczeniowa pojedynczej iteracji wynosi $\mathcal{O}(D)$, natomiast całego algorytmu $\mathcal{O}(ND)$.

Algorytm 1.1 Schemat działania typowych algorytmów stochastycznych pierwszego rzędu (wariant *offline*)

Wejścia: Zbiór pomiarów \mathcal{S} , słownik S

Wyznacz rozwiązanie początkowe $\boldsymbol{\theta}_1$

for $t = 1, \dots, N$ **do**

 Oblicz gradient funkcji straty w punkcie $\boldsymbol{\theta}_t$, $\mathbf{g}_t \leftarrow \nabla \hat{L}_t(\boldsymbol{\theta}_t)$

 Wyznacz rozwiązanie pośrednie $\boldsymbol{\theta}_{t+1}$

end for

Wyznacz rozwiązanie końcowe (najczęściej $\hat{\boldsymbol{\theta}} \leftarrow \frac{1}{N+1} \sum_{i=1}^{N+1} \boldsymbol{\theta}_i$)

1.5.2 Metoda stochastycznego gradientu prostego

Klasyknym przedstawicielem omawianej klasy algorytmów jest metoda stochastycznego gradientu prostego (ang. *stochastic gradient descent*, w skrócie SGD),

której pierwowzór przedstawiony został w latach 60. ubiegłego wieku przez Robbinsa i Monro [87]. Rozwiązanie początkowe algorytmu SGD jest dowolnie wybranym wektorem spełniającym ograniczenia algorytmu, tj. $\boldsymbol{\theta}_1 \in \mathcal{C}$, a następujące po sobie rozwiązania pośrednie wyznacza się wzorem

$$\boldsymbol{\theta}_{t+1} = \pi_{\mathcal{C}}(\boldsymbol{\theta}_t - \eta_t \mathbf{g}_t), \quad (1.21)$$

gdzie $\eta_t > 0$ jest długością kroku algorytmu, a $\pi_{\mathcal{C}}(\boldsymbol{\theta}) = \arg \min_{\mathbf{u} \in \mathcal{C}} \|\boldsymbol{\theta} - \mathbf{u}\|_2$ jest euklidesowym rzutem wektora $\boldsymbol{\theta}$ na zbiór \mathcal{C} . Jak nietrudno sprawdzić, wzór (1.21) można zapisać w formie problemu optymalizacyjnego,

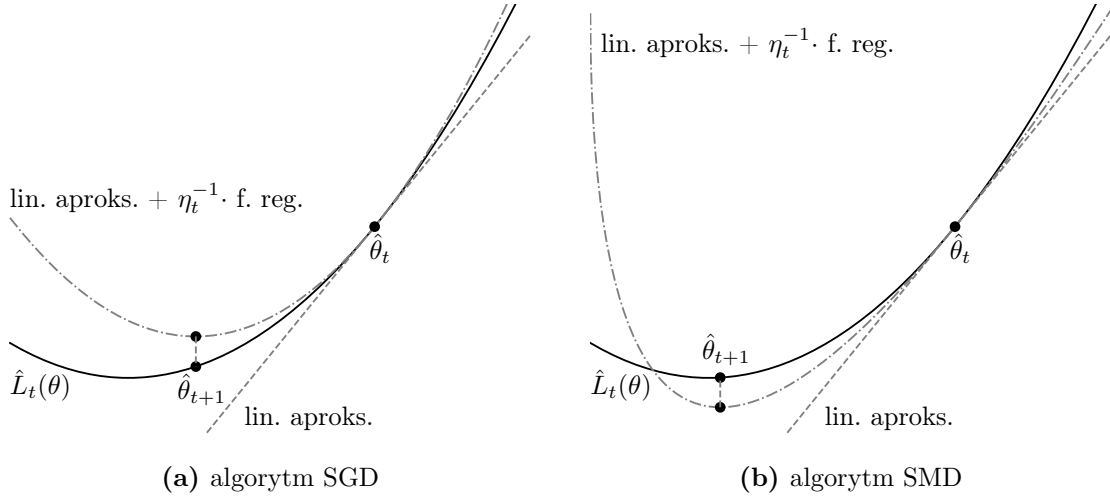
$$\begin{aligned} \boldsymbol{\theta}_{t+1} &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \overbrace{\hat{L}_t(\boldsymbol{\theta}_t) + \langle \mathbf{g}_t, \boldsymbol{\theta} - \boldsymbol{\theta}_t \rangle}^{\text{liniowa aproksymacja}} + \eta_t^{-1} \overbrace{\frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_2^2}^{\text{f. regularyzująca}} \right\} \\ &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \langle \mathbf{g}_t, \boldsymbol{\theta} \rangle + \eta_t^{-1} \frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|_2^2 \right\}, \end{aligned} \quad (1.22)$$

gdzie $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ jest standardowym iloczynem skalarnym. Tym samym, rozwiązanie pośrednie $\boldsymbol{\theta}_{t+1}$ algorytmu można interpretować jako minimum sumy liniowej aproksymacji funkcji $\hat{L}_t(\cdot)$ w punkcie $\boldsymbol{\theta}_t$ oraz przeskalowanej funkcji regularyzującej, która sprawia, że następujące po sobie rozwiązania pośrednie nie mogą znaleźć się w zbyt dużej euklidesowej odległości od siebie (rys. 1.6a).

Przejdziemy teraz do przedstawienia własności statystycznych metody. Niech stała $R = \sup_{\boldsymbol{\theta} \in \mathcal{C}} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}\|_2$ oznacza „promień” zbioru \mathcal{C} , a stała G spełnia warunek $\mathbb{E} \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_2^2 \leq G^2$ dla każdego $\boldsymbol{\theta} \in \mathcal{C}$. Wtedy, dla kroku $\eta_t = \frac{R}{G\sqrt{N}}$, błąd estymacji metody wynosi $\frac{RG}{\sqrt{N}}$ [50]. Na pierwszy rzut oka wynik ten wskazuje, że błąd metody SGD jest niezależny od wielkości słownika, co oznaczałoby, że problem sygnalizowany przy omawianiu metod bazujących na algorytmie LS został całkowicie wyeliminowany. Należy jednak zauważyć, że rozszerzanie słownika o kolejne elementy powoduje wzrost normy ℓ_2 gradientu: zakładając, że norma ℓ_∞ gradientów \mathbf{g}_t jest ograniczona i korzystając z faktu, że normy ℓ_∞ i ℓ_2 są równoważne, w szczególności zaś spełniają nierówność $\|\mathbf{g}_t\|_2 \leq \sqrt{D} \|\mathbf{g}_t\|_\infty$, dochodzimy do wniosku, że $G \sim \sqrt{D}$, a więc rząd szybkości zbieżności algorytmu wynosi $\sqrt{\frac{D}{N}}$. Zauważmy też, że manipulacja zbiorem ograniczeń \mathcal{C} ma wpływ jedynie na stałą R , nie pozwala więc na poprawienie rzędu szybkości zbieżności algorytmu. Fakt ten jest w pewien sposób sprzeczny z intuicją, którą można było nabyć we wcześniejszej części rozdziału, bowiem inaczej niż w przypadku metody LS, przyjęcie zbioru ograniczeń w postaci kuli w ℓ_1 nie niweluje wpływu wielkości słownika na błąd estymacji metody.

1.5.3 Metody stochastyczne, a metody *online* optymalizacji wypukłej

Zanim przejdziemy do omówienia kolejnych metod stochastycznych, przedstawimy ideę stojącą za algorytmami *online* optymalizacji wypukłej (ang. *online convex*



Rysunek 1.6: Ilustracja działania algorytmów SGD i SMD w przypadku problemów jednowymiarowych bez ograniczeń, $\mathcal{C} = \mathbb{R}$

optimization methods, w skrócie OCO), omówionych w m.in. w [44, 99]. W odróżnieniu do metod typu *offline*, w których zakłada się dostęp do całego zbioru danych pomiarowych przed rozpoczęciem działania algorytmu, metody OCO zakładają sekwencyjny dostęp do danych pomiarowych, a ich celem jest minimalizacja sumy empirycznych strat modeli pośrednich (tj. opartych o rozwiązania pośrednie, θ_t) względem ustalonego, niezmiennego w czasie modelu, opartego o dowolny wektor parametrów $\mathbf{u} \in \mathcal{C}$. Sumę tę przedstawia się zazwyczaj w funkcji \mathbf{u} i wyraża za pomocą następującego wzoru

$$C_T(\mathbf{u}) = \sum_{t=1}^T \left(\hat{L}_t(\theta_t) - \hat{L}_t(\mathbf{u}) \right).$$

Cechą charakterystyczną metod OCO jest to, że ich stosowanie nie wymaga wprowadzenia założeń o rozkładzie danych pomiarowych. Dane te mogą mieć więc charakter zarówno deterministyczny, jak i stochastyczny, a modelowany system może być niestacjonarny, a nawet adaptować się do zwróconych przez algorytm estymatorów θ_t . Z tego względu, metody OCO znajdują zastosowanie nie tylko w typowych problemach nauczania maszynowego jak predykcja i klasyfikacja, ale także np. w teorii gier [19].

Metody stochastyczne i metody OCO łączy bliskie pokrewieństwo. Mianowicie, algorytmy OCO mogą zostać w prosty sposób przekształcone w algorytmy stochastyczne, z teoretycznymi gwarancjami zbieżności. Jak pokazuje np. Shalev-Schwartz [99, rozdział 5], badając górne ograniczenie wyrażenia $\sup_{\mathbf{u} \in \mathcal{C}} C_T(\mathbf{u})$, przy założeniu, że dane pomiarowe są *i.i.d.*, można ustalić błąd estymacji metody w formie wzoru (1.12). Podobne podobieństwo zachodzi także w drugą stronę, tzn. niektóre warianty klasycznych algorytmów stochastycznych można rozpatrywać w kategoriach

algorytmów *online*. Przykładem niech będzie algorytm SGD – jeśli długość kroku wyznaczać w sposób niewymagający znajomości parametru N (np. zgodnie ze wzorem $\eta_t = \frac{R}{G\sqrt{t}}$), to metoda ta wpisuje się w ogólny schemat działania metod *online* (por. [122]), zob. algorytm 1.2. W związku z powyższym, na gruntach optymalizacji stochastycznej i optymalizacji *online* powstało wiele koncepcyjnie podobnych metod.

W ostatnich latach szczególnym zainteresowaniem cieszą się *adaptacyjne* metody OCO, tj. takie, w których długość kroku, bądź funkcja regularyzująca zmieniają się w czasie, w sposób zależny od zaobserwowanych po drodze pomiarów. Chociaż podejście to nie poprawia teoretycznej szybkości zbieżności algorytmów dla przypadku pesymistycznego, to jednak okazuje się, że w praktyce algorytmy adaptacyjne działają dużo lepiej, co widać chociażby na przykładzie klasycznego już algorytmu AdaGrad, zaproponowanego jednocześnie przez McMahana i Streetera [73] oraz Duchiego *et al.* [24, 25]. Algorytm ten rozbudził zainteresowanie badaczy metodami adaptacyjnymi i stał się inspiracją dla wielu kolejnych prac, zob. pracę przekrojową [72].

Algorytm 1.2 Schemat działania typowych algorytmów stochastycznych pierwszego rzędu (wariant *online*)

Wejścia: Słownik S

```

Wyznacz rozwiązanie początkowe  $\theta_1$ 
for  $t = 1, \dots, \infty$  do
  Dokonaj pomiaru wejścia  $X_t$  i wyjścia  $Y_t$  systemu
  Oblicz gradient funkcji straty w punkcie  $\theta_t$ ,  $\mathbf{g}_t \leftarrow \nabla \hat{L}_t(\theta_t)$ 
  Wyznacz rozwiązanie pośrednie  $\theta_{t+1}$ 
  Wyznacz aktualne rozwiązanie wyjściowe (najczęściej  $\hat{\theta} \leftarrow \frac{1}{t+1} \sum_{i=1}^{t+1} \theta_i$ )
end for

```

1.5.4 Metody lustrzane

Przedstawienie kolejnych kroków algorytmu SGD w postaci wzoru (1.22) prowokuje pytania o możliwość uogólnienia metody na inne funkcje regularyzujące, np. oparte o metryki *nieceuklidesowe*. W pracy [3] pokazano, że przyjęcie wzoru

$$\theta_{t+1} = \arg \min_{\theta \in \mathcal{C}} \left\{ \langle \mathbf{g}_t, \theta \rangle + \eta_t^{-1} B_h(\theta, \theta_t) \right\},$$

gdzie $B_h(\cdot, \cdot)$ stanowi tzw. *dywergencję Bregmana*, czyli przekształcenie postaci

$$B_h(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}) - (h(\mathbf{y}) + \langle \mathbf{x} - \mathbf{y}, \nabla h(\mathbf{y}) \rangle),$$

dla funkcji $h(\cdot)$ spełniającej pewne szczególne warunki wypukłości, prowadzi do rodziny *metod lustrzanych*² (ang. *mirror descent algorithms*, w skrócie MD), pierwotnie

²Tłumaczenie własne.

zapropionowanych przez Nemirowskiego i Yudina [77]. Zauważmy, że po dokonaniu podstawienia $h(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|_2^2$, metoda redukuje się do algorytmu SGD. Co ciekawe, w ogólnym przypadku funkcja $B_h(\mathbf{x}, \mathbf{y})$ jest niesymetryczna i nie stanowi metryki w ścisłym tego słowa znaczeniu, rys. 1.6b.

Metody lustrzane swoją nazwę zawdzięczają szczególnej interpretacji ich działania. Jeśli wektory $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t$ traktować jako elementy przestrzeni prymalnej, a gradienty $\mathbf{g}_1, \dots, \mathbf{g}_t$ jako elementy przestrzeni dualnej, to proces wyznaczania kolejnego rozwiązania pośredniego można przedstawić jako ciąg kroków, polegających na (zob. np. [14, rozdział 4]):

1. Odwzorowaniu wektora $\boldsymbol{\theta}_t$ w przestrzeń dualną, $\boldsymbol{\theta}_t \rightarrow \nabla h(\boldsymbol{\theta}_t)$.
2. Wykonaniu kroku algorytmu, $\nabla h(\psi_t) = \nabla h(\boldsymbol{\theta}_t) - \eta_t \mathbf{g}_t$.
3. Odwzorowaniu otrzymanego wektora $\nabla h(\psi_t)$ w przestrzeń prymalną, $\psi_t = (\nabla h)^{-1}(\nabla h(\boldsymbol{\theta}_t) - \eta_t \mathbf{g}_t)$.
4. Rzutowaniu ψ_t na zbiór \mathcal{C} względem pseudometryki bregmanowskiej, zgodnie ze wzorem

$$\boldsymbol{\theta}_{t+1} = \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} B_h(\boldsymbol{\theta}, \psi_t).$$

Metod SMD nie będziemy poddawać głębszej dyskusji, zwrócimy jednak uwagę na jeden szczególny, *entropijny* wariant algorytmu, który uzyskuje się przyjmując za $h(\cdot)$ funkcję ujemnej entropii, tj. $h(\boldsymbol{\theta}) = \sum_{i=1}^D \theta_i \ln \theta_i$, a za zbiór ograniczeń D -wymiarowy *simpleks probabilistyczny*, tj. $\mathcal{C} = \Delta_D := \{\boldsymbol{\theta} \in \mathbb{R}^D : \sum_{i=1}^D \theta_i = 1, \theta_i \geq 0\}$. Taki wybór prowadzi do jawnej postaci rozwiązań pośrednich $\boldsymbol{\theta}_t$ (co w ogólnym przypadku nie jest zagwarantowane, a co jest elementem koniecznym jeśli algorytm ma być efektywny obliczeniowo), a co istotniejsze, rząd szybkości zbieżności metody jest taki sam jak w przypadku metody LS z ograniczeniami ℓ_1 i wynosi $\mathcal{O}(\sqrt{\frac{\ln D}{T}})$ [50]. Niestety, wariant entropijny gwarantuje zbieżność metody tylko i wyłącznie, gdy długość kroku ma stałą wartość [80, uwaga 6.16], co wyklucza zastosowanie podejścia adaptacyjnego.

1.5.5 Metody lustrzane z uśrednianiem

Jako ostatnią, przedstawimy grupę *metod lustrzanych z uśrednianiem*³ (w literaturze angielskiej występujących najczęściej pod nazwą *dual averaging methods*, w skrócie DA), zaproponowanych przez Nesterowa [78], a następnie rozwiniętych i osadzonych w szerszym kontekście probabilistycznym przez Xiao [118]. W algoryt-

³Tłumaczenie własne.

mie DA, rozwiązania pośrednie wyznacza się wzorem

$$\begin{aligned}\boldsymbol{\theta}_{t+1} &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \overbrace{\frac{1}{t} \sum_{i=1}^t (\hat{L}_i(\boldsymbol{\theta}_i) + \langle \mathbf{g}_i, \boldsymbol{\theta} - \boldsymbol{\theta}_i \rangle)}^{\bar{L}_t(\boldsymbol{\theta})} + (t\eta_t)^{-1} h(\boldsymbol{\theta}) \right\}, \\ &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \frac{1}{t} \sum_{i=1}^t \langle \mathbf{g}_i, \boldsymbol{\theta} - \boldsymbol{\theta}_i \rangle + (t\eta_t)^{-1} h(\boldsymbol{\theta}) \right\}.\end{aligned}$$

Działanie algorytmu polega więc na poszukiwaniu minimum funkcji $\bar{L}_t(\boldsymbol{\theta})$, która stanowi *globalną* aproksymację funkcji straty, do zbudowania której wykorzystywana jest średnia ze wszystkich zaobserwowanych do danego momentu gradientów \mathbf{g}_i , $i = 1, \dots, t$; poza tym, w odróżnieniu do metod MD, kolejne rozwiązania pośrednie regularyzowane są w odniesieniu do tego samego, ustalonego punktu, a nie rozwiązania z iteracji poprzedzającej.

Działanie metod DA zostało dokładnie opisane w rozdziale 4, gdzie przedstawiono także wyniki dotyczące zastosowania algorytmu w zadaniu modelowania systemów dynamicznych. Szczególny nacisk położono na omówienie entropijnego wariantu metody, który inaczej niż w przypadku algorytmu MD, zezwala na zmianę długości kroku η_t w czasie. Fakt ten ma niebagatelne znaczenie, bowiem pozwala na adaptacyjny dobór długości kroku.

1.6 Zastosowania praktyczne modeli wielomianowych

Modele oparte o reprezentację Votlerry znajdują zastosowanie w automatyce (przede wszystkim w sterowaniu predykcyjnym [23, 39, 40, 69], ale również np. w modelowaniu procesów przemysłowych o parametrach rozłożonych [64]), przetwarzaniu sygnałów [29, 32, 65, 71, 107], a także w dziedzinach takich jak inżynieria mechaniczna i materiałowa, zob. [20] i przykłady tam zawarte.

Zastosowania modeli ortogonalnych w automatyce obejmują m.in. sterowanie optymalne [47, 54] i sterowanie predykcyjne w warunkach losowych [30, 56], zob. także pracę przekrojową [57], w której opisano więcej przykładów. Oprócz tego modele ortogonalne wykorzystywane są do rozwiązywania stochastycznych równań różniczkowych [119] oraz w analizie wrażliwości modeli fizycznych [106].

1.7 Cele pracy

Celami pracy są:

- opracowanie efektywnych obliczeniowo metod i algorytmów modelowania szerokiej klasy dynamicznych systemów nieliniowych, których stosowanie nie wymaga znajomości struktury badanego systemu, a szybkość zbieżności zależy tylko w niewielkim stopniu od poziomu złożoności rozważanej klasy modeli,
- analiza własności statystycznych algorytmów, ze szczególnym uwzględnieniem własności o charakterze nieasymptotycznym, tj. dla skończonej liczby pomiarów,
- implementacja opracowanych algorytmów oraz przeprowadzenie ich weryfikacji w ramach eksperymentów numerycznych na – o ile to możliwe – zbiorach pomiarowych fizycznie istniejących systemów.

1.8 Założenia ogólne i klasyfikacja zadań

Praca dotyczy dynamicznych systemów nieliniowych z czasem dyskretnym, które są jednocześnie stacjonarne i przyczynowe, i w których wyjście jest zakłócone przez szum addytywny, spełniający założenie:

Założenie 1.1. *Zakłócenie $\{Z_n\}$ jest ciągiem i.i.d. zmiennych losowych, niezależnych od wejścia systemu. Ponadto, wartość oczekiwana zakłócenia wynosi zero, $\mathbb{E}\{Z_n\} = 0$, a jego wariancja jest skończona, $\sigma_Z^2 < \infty$.*

Dodatkowo, o sygnale wejściowym poczynimy następujące założenie:

Założenie 1.2. *Sygnal wejściowy, $\{X_n\}$, jest ciągiem i.i.d. zmiennych losowych.*

Ze względu na sposób przeprowadzenia eksperymentu pomiarowego, w ramach którego pozyskiwany jest zbiór pomiarów wejścia-wyjścia systemu, wyróżnimy dwa rodzaje zadań modelowania:

- z *eksperymentem czynnym* – gdy pobudzenie można kontrolować i rozkład prawdopodobieństwa danych wejściowych, \mathcal{D}_X , jest znany;
- z *eksperymentem biernym* – gdy rozkład \mathcal{D}_X pozostaje nieznan.

1.9 Teza pracy

Nowoczesne gradientowe metody optymalizacji wypukłej, takie jak algorytmy lustrzane z uśrednianiem, mogą stanowić wydajne obliczeniowo narzędzie modelowania nieliniowych systemów dynamicznych. Wykorzystanie nieeuklidesowych wariantów tych metod, m.in. opartych o funkcję ujemnej entropii, pozwala dodatkowo

znacznie ograniczyć negatywny wpływ wymiaru problemu na szybkość zbieżności metody, co w konsekwencji umożliwia efektywną estymację modeli o dużych słownikach. Metody te mogą być rozpatrywane pod kątem zastosowania reprezentacji niewymagających znajomości struktury badanego systemu, opartych o rozwinięcie Volterry lub rozwinięcie Wienera.

Rozdział 2

Reprezentacja systemów oparta o szereg Volterry

W rozdziale omówiono zagadnienie reprezentacji systemów nieliniowych za pomocą szeregu Volterry. Oprócz własności aproksymacyjnych, przedstawiono także przykłady zastosowania oraz metodę konstrukcji modeli volterrowskich. W szczególności ustalono szybkość zbieżności błędu aproksymacji modeli volterrowskich dla pewnej klasy systemów LNL. Otrzymany rezultat opublikowany został w pracy [62].

2.1 Rozwinięcie Volterry

Niech $\mathcal{K}_p = \underbrace{\mathbb{R}^\infty \times \cdots \times \mathbb{R}^\infty}_p$ oznacza tensor p -tego rzędu. Reprezentacja oparta o dyskretny szereg Volterry przyjmuje postać

$$V[\mathbf{x}_n] = h_0 + \sum_{p=1}^{\infty} H_p[\mathbf{x}_n; h_p], \quad (2.1)$$

gdzie h_0 jest stałą, a przez $H_p : \mathbb{R}^\infty \times \mathcal{K}^p \rightarrow \mathbb{R}$ oznaczać będziemy *funkcjonał Volterry* p -tego rzędu, który wyraża się przez wzór

$$H_p[\mathbf{x}_n; h_p] = \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j}, \quad (2.2)$$

gdzie $h_p \in \mathcal{K}_p$ stanowi *jądro Volterry*, przy czym $h_p(i_1, \dots, i_p)$ odnosi się do elementu jądra o indeksach i_1, \dots, i_p . Dla wygody, stałą h_0 będziemy czasem interpretować jako funkcyjnał i oznaczać przez $H_0[\mathbf{x}_n; h_0]$. W celu lepszego zilustrowania wyrażenia

(2.2), przedstawiamy wzory definiujące pierwsze trzy funkcjonały Volterra:

$$\begin{aligned} H_1[\mathbf{x}_n; h_1] &= \sum_{i=0}^{\infty} h_1(i) x_{n-i}, \\ H_2[\mathbf{x}_n; h_2] &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} h_2(i, j) x_{n-i} x_{n-j}, \\ H_3[\mathbf{x}_n; h_3] &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} h_3(i, j, k) x_{n-i} x_{n-j} x_{n-k}. \end{aligned}$$

Jak można zauważyć, funkcjonal Volterra pierwszego rzędu, H_1 , tożsamy jest z reprezentacją splotową (1.5), stanowi więc liniowy komponent reprezentacji volterrowskiej (2.1). Tym samym, w przypadku gdy system nie przejawia nieliniowych zachowań, jądro h_1 pełni rolę wektora przechowującego kolejne wartości odpowiedzi impulsowej systemu. Funkcjonały drugiego i trzeciego rzędu odpowiedzialne są natomiast za odwzorowanie odpowiednio kwadratowych i sześciennych składowych nieliniowości systemu, a ich jądra zobrazować można, kolejno, jako macierz i trójwymiarową kostkę o nieskończonych wymiarach; w literaturze dotyczącej teorii systemów nazywa się je czasem odpowiedziami impulsowymi wyższych rzędów (np. [7]). W związku z tym, wzór (2.1) można interpretować dwojako: jako uogólnienie szeregu Taylora na funkcje wielu zmiennych pozwalającą na odzwierciedlenie dynamicznych własności systemu; lub też jako rozszerzenie opisu splotowego (1.5) o odpowiedzi impulsowe wyższych rzędów, celem uwzględnienia nieliniowego charakteru systemu.

W kontekście algorytmów modelowania, korzystać będziemy z *podwójnie obciętego* szeregu Volterra, którego rząd P oraz długości pamięci M są skończone, a który wyraża się poprzez

$$V_{P,M}[\mathbf{x}_n] = h_0 + \sum_{p=1}^P H_{p,M}[\mathbf{x}_n; h_p],$$

gdzie $H_{p,M}$ stanowi obcięty funkcjonal Volterra p -tego rzędu, o długości pamięci równej M , określony przez wzór

$$H_{p,M}[\mathbf{x}_n; h_p] = \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j},$$

przy czym, w tym wypadku, jądra są tensorami o skończonych wymiarach, tj. $h_p \in \mathcal{K}_p^M = \underbrace{\mathbb{R}^M \times \cdots \times \mathbb{R}^M}_p$.

Analizując przedstawione powyżej wzory, zaobserwować można pewną nadmiarowość omawianej reprezentacji, która wynika z symetrii jąder drugiego i wyższych rzędów. Zauważmy bowiem, że np. funkcjonal drugiego rzędu można przedstawić w formie tożsamej, ale wymagającej mniejszej liczby parametrów (elementów jądra). Mianowicie, dla jądra zdefiniowanego przez

$$\tilde{h}_2(i, j) = \begin{cases} h_2(i, j) + h_2(j, i), & \text{dla } i > j, \\ h_2(i, i), & \text{dla } i = j, \\ 0, & \text{dla } i < j. \end{cases} \quad (2.3)$$

funkcjonał opisany wzorem

$$H_{2,M}[\mathbf{x}_n; \tilde{h}_2] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \tilde{h}_2(i, j) x_{n-i} x_{n-j} = \sum_{i=0}^{M-1} \sum_{j=0}^i \tilde{h}_2(i, j) x_{n-i} x_{n-j} \quad (2.4)$$

spełnia tożsamość

$$H_{2,M}[\mathbf{x}_n; \tilde{h}_2] = H_{2,M}[\mathbf{x}_n; h_2],$$

przy czym do jego opisu wymaga się jedynie $\binom{M+1}{2}$, zamiast M^2 elementów jądra. W ogólności, dla funkcyjonału p -tego rzędu, modyfikacja ta pozwala zredukować ilość wymaganych parametrów z M^p do $\binom{M+p-1}{p}$. W sumie, do zdefiniowania obciętego szeregu Volterry opartego o zmodyfikowane jądra \tilde{h}_p , utworzone w sposób analogiczny do (2.3), potrzebnych jest $D = \binom{M+P}{P}$ parametrów [53]. Dla odróżnienia od standardowego opisu rozwinięcia Volterry, powyższy wariant reprezentacji będziemy nazywali *trójkątnym*.

2.2 Własności aproksymacyjne

Wprowadzimy teraz pojęcie *pamięci zanikającej*, które przy omawianiu własności aproksymacyjnych szeregów Volterry, pełni rolę analogiczną do pojęcia ciągłości w teorii funkcji rzeczywistych.

Definicja 2.1 ([11]). System o charakterystyce $m_{\text{sys}}[\cdot]$ ma pamięć zanikającą dla klasy ograniczonych sygnałów wejściowych $K \subseteq l_\infty$, jeśli dla każdego sygnału wejściowego $\mathbf{x}_n \in K$, istnieje taki ciąg malejący $w_n : \mathbb{N} \rightarrow (0, 1]$, $\lim_{n \rightarrow \infty} w_n = 0$, że dla każdego $\varepsilon > 0$, istnieje taka $\delta > 0$, że dla każdego sygnału $\mathbf{u}_n \in K$ i dla każdej chwili czasu $n \in \mathbb{N}$ zachodzi następująca implikacja

$$\sup_{k \geq 0} |x_{n-k} - u_{n-k}| w_k < \delta \implies |m_{\text{sys}}[\mathbf{x}_n] - m_{\text{sys}}[\mathbf{u}_n]| < \varepsilon.$$

Intuicyjnie, własność tą posiadają systemy w których *mała* zmiana sygnału wejściowego powoduje *małą* zmianę na wyjściu i których wyjście, wraz z upływem czasu, zależy w coraz mniejszym (zbiegającym asymptotycznie do zera) stopniu od wartości sygnału zmierzonego na wejściu systemu w pewnej ustalonej chwili n_0 . Jeśli system ma *pamięć skończoną*, tzn. spełnia zależność

$$m_{\text{sys}}[\mathbf{x}_n] = m_{\text{sys}}[\mathbf{x}_{n,M}],$$

gdzie $M < \infty$ stanowi długość pamięci systemu, a $\mathbf{x}_{n,M} = [x_n, x_{n-1}, \dots, x_{n-M+1}] \in \mathbb{R}^M$ jest obciętym wektorem wejść systemu, to warunkiem posiadania własności pamięci zanikającej jest po prostu ciągłość funkcjonału $m_{\text{sys}}[\cdot]$. Przykładem systemów posiadających nieskończoną i jednocześnie zanikającą pamięć są asymptotycznie stabilne liniowe systemy dynamiczne, posiadające reprezentację splotową (1.5).

Jako przykłady systemów nieposiadających pamięci zanikającej, można wymienić takie, których charakterystyki są funkcjami wielowartościowymi (np. mające charakter histerezy) [71, ustęp 2.1.3.3] lub takie, które mają więc niż jeden stabilny punkt równowagi, osiągalny w ramach rozpatrywanej klasy wejść [93]. Fizycznym przykładem systemu z pamięcią niezanikającą jest bezpiecznik topikowy, którego uproszczony model matematyczny wyrazić można za pomocą wzoru

$$T[\mathbf{x}_n] = \begin{cases} 1, & \text{dla } \|\mathbf{x}_n\|_\infty \geq C, \\ 0, & \text{dla } \|\mathbf{x}_n\|_\infty < C, \end{cases}$$

gdzie $C > 0$ jest pewną stałą.

Przywołamy teraz twierdzenie, które porządkuje wiedzę dotyczącą własności aproksymacyjnych szeregu Volterry.

Twierdzenie 2.1 ([11]¹). *Niech $m_{\text{sys}}[\cdot]$ będzie charakterystyką systemu stacjonarnego posiadającego pamięć zanikającą dla klasy ograniczonych sygnałów wejściowych K . Wtedy, dla dowolnego $\varepsilon > 0$, istnieje taki obcięty szereg Volterry $V_{P,\infty}$, $P < \infty$, że dla każdego sygnału wejściowego $\mathbf{x}_n \in K$, błąd modelu $V_{P,\infty}$, dany przez*

$$e_P = |m_{\text{sys}}[\mathbf{x}_n] - V_{P,\infty}[\mathbf{x}_n]|, \quad (2.5)$$

spełnia nierówność $e_P < \varepsilon$.

Dowód. Zob. [11]. □

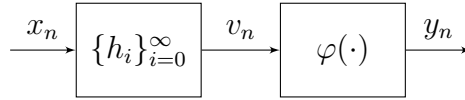
Powyższe twierdzenie można potraktować jako odpowiednik Twierdzenia Weierstrassa dla systemów z pamięcią zanikającą. Mimo iż jego treść ogranicza się do własności szeregu Volterry klasy $V_{P,\infty}$, a więc obciętych tylko z jednej strony (z pamięcią nieskończoną), to wynik ten stał się przyczynkiem do rozwoju wielu algorytmów modelowania opartych o reprezentację volterrowską.

Uwaga 2.1. *Jeżeli system posiada pamięć skończoną o długości M , w treści twierdzenia (2.1) szereg z pamięcią nieskończoną, $V_{P,\infty}$, zastąpić można szeregiem Volterry z pamięcią o długości M , tj. $V_{P,M}$.*

¹W cytowanej pracy, przez *system* autorzy rozumieją operator, który przekształca sygnał w sygnał, np. $m_{\text{sys}} : l_\infty \rightarrow l_\infty$, dlatego w celu zachowania przejrzystości prezentacji i przyjętej tu, odmiennej definicji systemu, wzór (1.1), cytowane twierdzenie zostało podane w nieco zmienionej formie.

Przykład 2.1. Rozważmy system o strukturze LN (rys. 2.1), którego podsystem dynamiczny opisany jest odpowiedzią impulsową, $\{h_i\}_{i=0}^{\infty}$, spełniającą warunek $\|h\|_1 < \infty$ (co implikuje stabilność systemu), a charakterystyka nieliniowa jest funkcją sigmoidalną, daną przez wzór $\varphi(x) = \arctan(x)$. Zależność wyjścia od wejścia systemu opisuje równanie

$$T[\mathbf{x}_n] = \arctan\left(\sum_{i=0}^{\infty} h_i x_{n-i}\right). \quad (2.6)$$



Rysunek 2.1: System Wienera

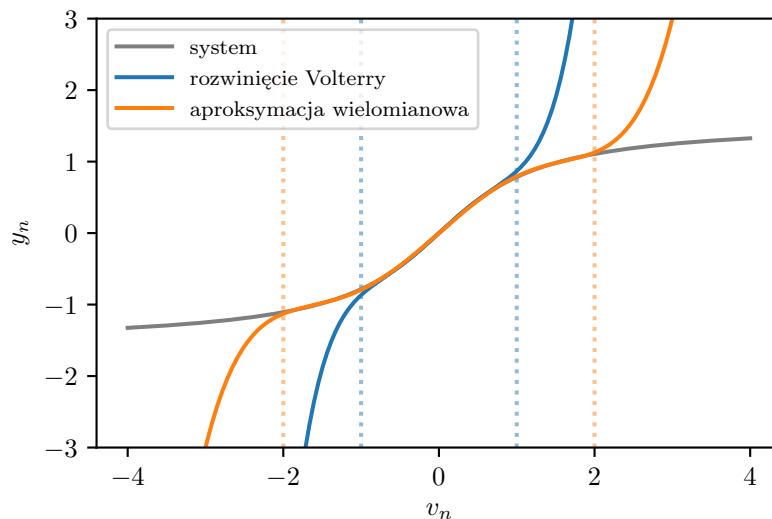
Korzystając z faktu, że $\arctan(\cdot)$ jest funkcją analityczną, rozwijamy ją w szereg Taylora wokół punktu zerowego. W ten sposób, równanie (2.6) przyjmuje postać

$$T[\mathbf{x}_n] = \sum_{p=1}^{\infty} (-1)^{p+1} \frac{v_n^{2p-1}}{2p-1} = v_n - \frac{v_n^3}{3} + \frac{v_n^5}{5} - \frac{v_n^7}{7} + \dots, \quad (2.7)$$

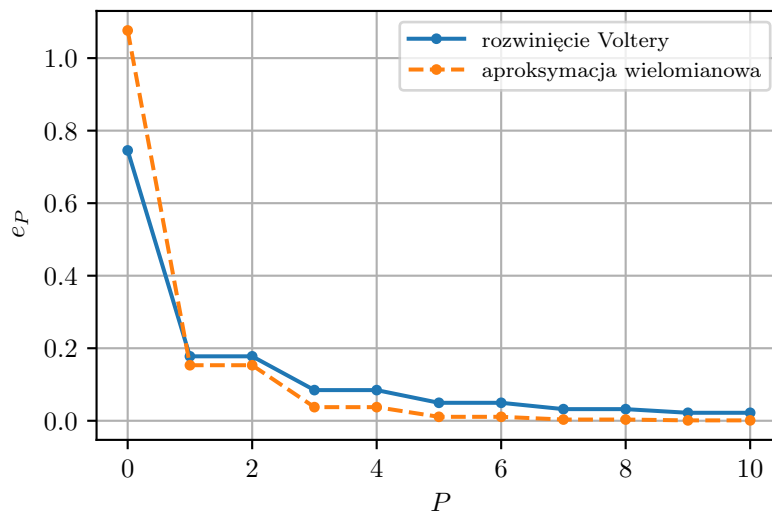
gdzie $v_n = \sum_{i=0}^{\infty} h_i x_{n-i}$ jest sygnałem pośrednim systemu. Tym samym otrzymujemy rozwinięcie Volterry systemu, którego jądra wyrażone są poprzez $g_p(i_1, \dots, i_p) = \frac{(-1)^{p+1}}{2p-1} h_{i_1} \dots h_{i_p}$ dla p nieparzystych i $g_p(i_1, \dots, i_p) = 0$ dla p parzystych. Z uwagi na ograniczony przedział zbieżności szeregu Taylora funkcji $\arctan(\cdot)$, który w tym wypadku wynosi $[-1, 1]$, reprezentacja volterrowska systemu (2.7) jest zbieżna, gdy sygnał wewnętrzny $\{v_n\}$ jest ograniczony, tzn. dla każdego $n \in \mathbb{N}$ spełnia warunek $|v_n| \leq 1$. Zauważając, że $|v_n| = |\sum_{i=0}^{\infty} h_i x_{n-i}| \leq \sum_{i=0}^{\infty} |h_i x_{n-i}| \leq \|\mathbf{x}_n\|_{\infty} \sum_{i=0}^{\infty} |h_i|$, dochodzimy do wniosku, iż wyjście szeregu (2.7) zbiega do wyjścia systemu (2.6) dla klasy sygnałów wejściowych $K = \{\{x_n\} : \|\mathbf{x}_n\|_{\infty} \leq \|h\|_1^{-1}\}$. Jak widać zatem, promień zbieżności szeregu Volterry uzależniony jest zarówno od własności charakterystyki nieliniowej, jak i od odpowiedzi impulsowej podsystemu dynamicznego.

Jednakże, zgodnie z twierdzeniem 2.1, otrzymany rezultat nie ogranicza stosowalności rozwinięcia Volterry, ponieważ dla dowolnej klasy sygnałów ograniczonych, np. $K_{\alpha} = \{\{x_n\} : \|\mathbf{x}_n\|_{\infty} \leq \alpha \|h\|_1^{-1}\}$, istnieje szereg $V_{P,\infty}$, którego wyjście zbiega (wg. normy maksimum) do wyjścia systemu (2.6). Dla ustalenia uwagi przyjmijmy, że kolejne komponenty odpowiedzi impulsowej podsystemu dynamicznego dane są przez $h_i = 2^{-i}$, $i = 0, 1, \dots$. W ten sposób $\|h\|_1 = 2$ i szereg Volterry zbiega dla klasy sygnałów $K = \{\{x_n\} : \|\mathbf{x}_n\|_{\infty} \leq \frac{1}{2}\}$. Załóżmy teraz, że naszym celem jest uzyskanie modelu volterrowskiego systemu dla szerszej klasy sygnałów wejściowych, np. $K_2 = \{\{x_n\} : \|\mathbf{x}_n\|_{\infty} \leq 1\}$. Jego jądra będą miały postać $g_p(i_1, \dots, i_p) = c_p h_{i_1} \dots h_{i_p}$, gdzie $\{c_p\}$ to kolejne współczynniki rozwinięcia funkcji $\arctan(\cdot)$ w szereg potęgowy na przedziale $[-\alpha, \alpha] = [-2, 2]$, których wartości uzyskać można korzystając np. z metody aproksymacji wielomianowej. Na wykresie

2.2 przedstawiono charakterystyki nieliniowe prawdziwego systemu, (2.6), modelu którego współczynniki $\{c_p\}$ uzyskane zostały za pomocą aproksymacji wielomianowej, a także modelu pierwotnego, (2.7). Z kolei wykres 2.3 przedstawia błąd (2.5), w zależności od rzędu modelu, P .



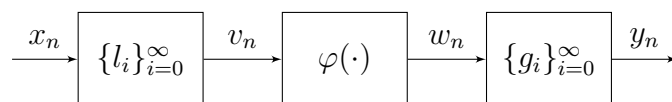
Rysunek 2.2: Charakterystyki nieliniowe modeli rzędu $P = 5$, w funkcji sygnału pośredniego. Linie przerywane oznaczają granice zbieżności poszczególnych modeli



Rysunek 2.3: Błąd modelu, w zależności od rzędu rozwinięcia

2.3 Studium przypadku – aproksymacja systemu LNL

Przedstawimy teraz wynik dotyczący szybkości zbieżności błędu aproksymacji obciążonej reprezentacji volterrowskiej dla systemów LNL.



Rysunek 2.4: System Wienera-Hammersteina

Rozważmy system Wienera-Hammersteina (LNL), którego schemat przedstawiono na rys. 2.4. Kolejne bloki systemu opisać można za pomocą wzorów

$$v_n = \sum_{i=0}^{\infty} l_i x_{n-i}, \quad w_n = \varphi(v_n), \quad y_n = \sum_{i=0}^{\infty} g_i w_{n-i},$$

gdzie przez v_n , w_n i y_n oznacza się, kolejno, wyjście pierwszego podsystemu liniowego, wyjście podsystemu nieliniowego i wyjście drugiego podsystemu liniowego, które stanowi jednocześnie wyjście całego systemu. W konsekwencji, wyjście systemu opisać można równaniem

$$y_n = m_{\text{LNL}}[\mathbf{x}_n] = \sum_{i=0}^{\infty} g_i \varphi \left(\sum_{j=0}^{\infty} l_j x_{n-(i+j)} \right), \quad (2.8)$$

gdzie $m_{\text{LNL}} : \mathbb{R}^{\infty} \rightarrow \mathbb{R}$ oznacza funkcjonal reprezentujący system LNL.

W stosunku do systemu i sygnału wejściowego poczynimy następujące założenia:

Założenie 2.1. *Oba podsystemy liniowe są asymptotycznie stabilne.*

Założenie 2.2. *Charakterystyka podsystemu nieliniowego, $\varphi(\cdot)$, jest funkcją, której wszystkie pochodne, $\varphi^{(p)}(\cdot)$, $p = 1, 2, \dots$, są ograniczone w taki sposób, że spełniają nierówność*

$$|\varphi^{(p)}(x)| \leq B^p,$$

dla pewnej stałej $B > 1$ i wszystkich $x \in \mathbb{R}$.

Założenie 2.3. *Sygnał wejściowy, $\{x_n\}$, jest ograniczony w taki sposób, że spełniona jest nierówność $\forall_n |x_n| \leq 1$.*

Założenie dotyczące stabilności podsystemów liniowych jest jednym z warunków zbieżności reprezentacji volterrowskiej badanego systemu. Dla transmitancji posiadających bieguny wielokrotne (o maksymalnej krotności równej m), kolejne składowe

odpowiedzi impulsowych bloków dynamicznych, ograniczyć można w sposób następujący [36]:

$$|l_i| \leq k_l i^{m-1} \rho_l^i \leq k'_l \rho_l^{i(1-\varepsilon)} \quad \text{oraz} \quad |g_i| \leq k_g i^{m-1} \rho_g^i \leq k'_g \rho_g^{i(1-\varepsilon)}, \quad (2.9)$$

gdzie $\rho_l = \max_i |\zeta_{l,i}| \in (0, 1)$, $\rho_g = \max_i |\zeta_{g,i}| \in (0, 1)$ (i gdzie $\zeta_{l,i}$, $\zeta_{g,i}$ są i -tymi biegunem transmitancji pierwszego i drugiego podsystemu liniowego), a $0 < k_l, k_g, k'_l, k'_g < \infty$, $0 < \varepsilon < 1$ stanowią pewne stałe.

Uwaga 2.2. Ograniczenie wprowadzone przez założenie 2.3 można złagodzić i przedstawione rozważania rozszerzyć na klasę sygnałów ograniczonych na dowolnym przedziale. Jeśli bowiem sygnał ograniczony jest przez pewną stałą $A > 1$, tj. $\forall_n |x_n| \leq A$, to wejścia systemu można znormalizować (poprzez podzielenie ich przez A), a stałą B zastąpić przez AB . Operacja ta nie zmienia promienia zbieżności rozwinięcia Taylora funkcji $\varphi(\cdot)$, gdyż jest on nieskończony (jest to konsekwencją założenia 2.2).

2.3.1 Błąd aproksymacji

Przez błąd obcięcia podwójnie obciętego szeregu Volterry (rzędu P , długości pamięci M) będziemy rozumieli maksymalną możliwą różnicę pomiędzy wyjściem systemu (2.8), a wyjściem *najlepszego* modelu $V_{P,M}$. Wielkość ta wyrażona jest wzorem

$$e_{P,M} = \inf_{V \in \{V_{P,M}\}} \sup_{\mathbf{x} \in \mathbb{R}^\infty} |m_{\text{LNL}}[\mathbf{x}] - V[\mathbf{x}]|. \quad (2.10)$$

Lemat 2.1. Dla systemu LNL zdefiniowanego przez wzór (2.8) i spełniającego założenia 2.1 i 2.2, dla klasy sygnałów wejściowych spełniających założenie 2.3, błąd obcięcia (2.10) spełnia następującą nierówność

$$e_{P,M} \leq c_1 C^{-P} + c_2 \lambda^{-M}, \quad (2.11)$$

gdzie $\lambda = (\max_i |\zeta_i|)^{-1} > 1$ (dla ζ_i będącego i -tym biegunem transmitancji pierwszego podsystemu liniowego) i gdzie $c_1 > 0, c_2 > 0, C > 1$ to pewne stałe, niezależne od P i M .

Dowód. Dowód przedstawiony został w dodatku 2.A. □

Zauważmy, że dla systemu spełniającego podane założenia zachodzi

$$\lim_{P,M \rightarrow \infty} e_{P,M} = 0,$$

a zatem lemat 2.1 pokazuje, że *podwójnie* obcięty szereg Volterry, przy założeniu ograniczonego wejścia, zbiega (w sensie sygnału wyjściowego) do prawdziwego systemu. Tak więc, lemat ten, oprócz wskazania szybkości zbieżności reprezentacji volterrowskiej, dla określonej klasy systemów rozszerza zakres stosowalności twierdzenia [11], bowiem – mimo nieskończonej pamięci systemu – dopuszcza wykorzystanie modeli z obciętą pamięcią.

Jak pokazujemy poniżej, znajomość górnego ograniczenia błędu obciążenia pozwala na oszacowanie błędu aproksymacji rozważanej klasy modeli.

Wniosek 2.1. Niech $\Delta_{apr.}^{P,M}$ oznacza błąd aproksymacji oparty o kwadratową funkcję straty dla modeli klasy $\{V_{p,m} : p \leq P, m \leq M\}$. Dla systemu spełniającego warunki podane w lemacie 2.1 zachodzi następująca nierówność

$$\Delta_{apr.}^{P,M} \leq 2c_1^2 C^{-2P} + 2c_2^2 \lambda^{-2M}, \quad (2.12)$$

gdzie poszczególne stałe odpowiadają tym z lematu 2.1.

Dowód. Niech $m^* \in \arg \min_{m \in \{V_{(P,M)}\}} L(m)$ będzie modelem minimalizującym kwadratową funkcję straty, a $m^+ \in \arg \min_{m \in \{V_{(P,M)}\}} \sup_{\mathbf{x} \in \mathbb{R}^\infty} |m_{LNL}[\mathbf{x}] - V[\mathbf{x}]|$ modelem minimalizującym błąd obciążenia. Z faktu, iż $L(m^*) \leq L(m^+)$, wynika następująca nierówność

$$\Delta_{apr.}^{P,M} = L(m^*) - L(m_{LNL}) \leq L(m^+) - L(m_{LNL}).$$

Zauważmy teraz, że prawą stronę powyższego wzoru można dalej ograniczyć przez

$$L(m^+) - L(m_{LNL}) = \mathbb{E}\{m_{LNL}[\mathbf{X}_n] - m^+[\mathbf{X}_n]\}^2 \leq e_{P,M}^2,$$

a więc dokonując podstawienia (2.11), otrzymujemy nierówność

$$\Delta_{apr.}^{P,M} \leq (c_1 C^{-P} + c_2 \lambda^{-M})^2.$$

Zastosowanie nierówności Cauchy'ego-Schwarza kończy dowód. □

2.3.2 Podsumowanie

Rozważana klasa systemów obejmuje wszystkie stabilne struktury z m.in. sinusoidalnymi, wielomianowymi i eksponencjalnymi nieliniowościami. Niemniej jednak, zastosowana technika dowodowa stawia dodatkowe wymagania dotyczące gładkości funkcji $\varphi(\cdot)$, co bez narzucenia dalszych ograniczeń na klasę sygnałów wejściowych, wyklucza np. funkcję $\arctan(\cdot)$.

Uzyskany wynik posłuży w dalszej części pracy jako punkt wyjścia do konstrukcji algorytmu modelowania systemów LNL z nieskończoną pamięcią.

2.4 Konstrukcja modeli volterrowskich

Istotną cechą funkcjonałów Voltery, która sprawia, że są one atrakcyjne z punktu widzenia modelowania systemów, jest to, że są one liniowe ze względu na parametry. Mianowicie, dla dowolnych $\alpha, \beta \in \mathbb{R}$ oraz $a_p, b_p \in \mathcal{K}_p$, p -ty funkcjonał Voltery

spełnia tożsamość

$$\begin{aligned}
H_p[\mathbf{x}_n; \alpha a_p + \beta b_p] &= \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} [\alpha a_p(i_1, \dots, i_p) + \beta b_p(i_1, \dots, i_p)] \prod_{j=1}^p x_{n-i_j} \\
&= \alpha \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} a_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \\
&\quad + \beta \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} b_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \\
&= \alpha H_p[\mathbf{x}_n; a_p] + \beta H_p[\mathbf{x}_n; b_p].
\end{aligned}$$

Dzięki temu, tak jak powiedzieliśmy w rozdziale pierwszym, modele oparte o reprezentację Volterry można przedstawić jako liniową kombinację elementów pewnego słownika S (zob. wzór (1.11)). Omówimy teraz konstrukcję słowników dla modeli volterrowskich.

Niech C_M^p oznacza zbiór zawierający unikalne p -elementowe krotki, zawierające indeksy od 0 do $M - 1$, przykładowo $C_3^2 = \{(0, 0), (0, 1), (0, 2), (1, 1), (1, 2), (2, 2)\}$. Niech teraz $\mathcal{G}(p, L)$ będzie odwzorowaniem C_M^p w zbiór przekształceń $\{\varphi_{\alpha_1}, \varphi_{\alpha_2}, \dots\}$, gdzie α_i oznacza i -ty element zbioru C_M^p . Jak łatwo sprawdzić, przyjmując

$$\varphi_{\alpha}(\mathbf{x}_n) = \prod_{\alpha \in \alpha} x_{n-\alpha},$$

p -ty obcięty operator Volterry z pamięcią długości M może być wyrażony jako liniowa kombinacja elementów $\mathcal{G}(p, M)$, w następstwie czego, słownik volterrowski rzędu P i pamięci długości M definiujemy przez

$$S_{P,M} = \{\mathbf{1}\} \cup \bigcup_{p=1}^P \mathcal{G}(p, M),$$

gdzie $\mathbf{1}(\mathbf{x}_n) = 1$, oznacza element stały, będący odpowiednikiem operatora Volterry rzędu zerowego.

Korzystając z powyższego formalizmu będziemy także definiowali słowniki, na które składają się funkcjonały posiadające pamięć różnej długości, tj.

$$S_{(M_1, \dots, M_P)} = \{\mathbf{1}\} \cup \bigcup_{p=1}^P \mathcal{G}(p, M_p).$$

Dodatki

2.A Dowód lematu 2.1

Dowód. Zauważmy, że błąd obcięcia (2.10) spełnia następującą nierówność

$$\begin{aligned}
 e_{P,M} &\leq \sup_{\mathbf{x} \in \mathbb{R}^\infty} \left[\sum_{p=1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} - \right. \\
 &\quad \left. - \sum_{p=1}^P \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \right] \\
 &\leq \sum_{p=P+1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |h_p(i_1, \dots, i_p)| \|\mathbf{x}_n\|_\infty^p + \\
 &\quad + \sup_{\mathbf{x} \in \mathbb{R}^\infty} \sum_{p=1}^P \left[\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} - \right. \\
 &\quad \left. - \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \right], \tag{2.13}
 \end{aligned}$$

gdzie $h_p(\cdot)$ są jądrami reprezentacji volterrowskiej rozważanego systemu. W kolejnych krokach będziemy chcieli znaleźć zależność pomiędzy poszczególnymi składnikami wzoru (2.13), a parametrami systemu.

W tym celu zwróćmy uwagę, że funkcja $\varphi(x)$, charakteryzująca podsystem nieliniowy, jest funkcją gładką, co więcej, dzięki założeniu 2.2 jest tożsama swojemu szeregowi Taylora na dowolnym przedziale $[-r, r]$, $r > 0$, zatem w szczególności zachodzi

$$\varphi(x) = \varphi(0) + \frac{\varphi'(0)}{1!}x + \frac{\varphi''(0)}{2!}x^2 + \frac{\varphi'''(0)}{3!}x^3 + \dots = \sum_{i=0}^{\infty} c_p x^i, \tag{2.14}$$

gdzie $c_p = \frac{\varphi^{(p)}(0)}{p!}$ oznacza p -ty współczynnik rozwinięcia. Tym samym, p -te jądro reprezentacji volterrowskiej systemu jest określone przez wzór $h_p(i_1, \dots, i_p) = c_p \sum_{i=0}^{\infty} g_i \prod_{j=1}^p l_{i_j-i}$, gdzie $l_i = 0$ dla $i < 0$ (zob. np. [53, 55]).

Zauważmy, że suma modułów wszystkich elementów p -tego jądra reprezentacji

spełnia nierówność

$$\begin{aligned}
\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |h_p(i_1, \dots, i_p)| &\leq \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |c_p| \sum_{i=0}^{\infty} |g_i| \prod_{j=1}^p |l_{i_j-i}| \\
&= |c_p| \sum_{i=0}^{\infty} |g_i| \sum_{i_1=0}^{\infty} |l_{i_1-i}| \cdots \sum_{i_p=0}^{\infty} |l_{i_p-i}| \\
&= |c_p| \sum_{i=0}^{\infty} |g_i| \left[\sum_{j=0}^{\infty} |l_{j-i}| \right]^p \leq |c_p| \sum_{i=0}^{\infty} |g_i| \left[\sum_{j=0}^{\infty} |l_j| \right]^p. \quad (2.15)
\end{aligned}$$

Przypomnijmy, że na mocy założenia 2.1, odpowiedź impulsowa pierwszego bloku dynamicznego jest ograniczona przez

$$|l_i| \leq k_l \rho_l^{i(1-\varepsilon)},$$

gdzie $0 < \rho_l < 1$, $0 < k_l < \infty$ i $0 < \varepsilon < 1$ są pewnymi stałymi. W konsekwencji, sumę modułów kolejnych elementów wektora odpowiedzi impulsowej pierwszego podsystemu ograniczymy z góry przez

$$\sum_{i=0}^{\infty} |l_i| \leq \sum_{i=0}^{\infty} k_l \rho_l^{i(1-\varepsilon)} = \frac{k_l}{1 - \rho_l^{1-\varepsilon}} = K_l,$$

a sumę częściową przez

$$\sum_{i=M}^{\infty} |l_i| \leq \sum_{i=M}^{\infty} k_l \rho_l^{i(1-\varepsilon)} = \frac{k_l}{1 - \rho_l^{1-\varepsilon}} \rho_l^{M(1-\varepsilon)} = K_l \rho_l^{M(1-\varepsilon)}.$$

Postępując analogicznie z drugim podsystemem liniowym, otrzymujemy nierówności $\sum_{i=0}^{\infty} |g_i| \leq K_g$ i $\sum_{i=M}^{\infty} |g_i| \leq K_g \rho_g^{M(1-\varepsilon)}$. Z kolei, z założenia 2.2 i wzoru (2.14), wynika że współczynniki rozwinięcia nieliniowości w szereg potęgowy są ograniczone przez $|c_p| \leq \frac{B^p}{p!} \leq C_1 C_2^{-p}$, $p \geq 1$, dla pewnych $C_1, C_2 > 0$. Jak zostanie pokazane, w celu wymuszenia zbieżność górnego ograniczenia błędu obciążenia, wymaga się wprowadzenia założenia technicznego, $C_2 > 2K_l$ (zauważmy, że nierówność ta może być spełniona zawsze, gdyż jak łatwo pokazać, dla dowolnego nieujemnego C_2 , istnieje takie C_1 , że $\frac{B^p}{p!} \leq C_1 C_2^{-p}$ jest spełnione dla każdego p).

Łącząc powyższe wyprowadzenia, otrzymujemy górne ograniczenie wyrażenia (2.15)

$$\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |h_p(i_1, \dots, i_p)| \leq C_1 C_2^{-p} K_g K_l^p,$$

co pozwala oszacować pierwszy składnik wzoru (2.13), tj.

$$\begin{aligned}
 & \sum_{p=P+1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |h_p(i_1, \dots, i_p)| \|\mathbf{x}_n\|_{\infty}^p \\
 & \leq C_1 K_g \sum_{p=P+1}^{\infty} C_2^{-p} K_l^p \\
 & = C_1 K_g \frac{\left(\frac{K_l}{C_2}\right)^{P+1}}{1 - \frac{K_l}{C_2}} = \mathcal{O}\left(C^{-P}\right), \tag{2.16}
 \end{aligned}$$

gdzie $C = \frac{C_2}{K_l} > 1$. W celu uproszczenia analizy drugiego składnika wzoru (2.13), w kolejnym kroku wyprowadzimy górne ograniczenie następującego wyrażenia

$$\begin{aligned}
 & \left| \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} - \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \right| \\
 & = \left| \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} c_p \sum_{i=0}^{\infty} g_i \prod_{j=1}^p l_{i_j-i} \prod_{k=1}^p x_{n-i_k} - \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} c_p \sum_{i=0}^{\infty} g_i \prod_{j=1}^p l_{i_j-i} \prod_{k=1}^p x_{n-i_k} \right| \\
 & = \left| c_p \sum_{i=0}^{\infty} g_i \left(\sum_{j=0}^{\infty} l_{j-i} x_{n-j} \right)^p - c_p \sum_{i=0}^{\infty} g_i \left(\sum_{j=0}^{M-1} l_{j-i} x_{n-j} \right)^p \right| \\
 & = \left| c_p \sum_{i=0}^{\infty} g_i \left[\left(\sum_{j=M}^{\infty} l_{j-i} x_{n-j} + \sum_{j=0}^{M-1} l_{j-i} x_{n-j} \right)^p - \left(\sum_{j=0}^{M-1} l_{j-i} x_{n-j} \right)^p \right] \right| \\
 & = \left| c_p \sum_{i=0}^{\infty} g_i \sum_{q=1}^p \binom{p}{q} \left(\sum_{j=0}^{M-1} l_{j-i} x_{n-j} \right)^{p-q} \left(\sum_{j=M}^{\infty} l_{j-i} x_{n-j} \right)^q \right| \\
 & \leq |c_p| \|\mathbf{x}_n\|_{\infty}^p \sum_{i=0}^{\infty} |g_i| \sum_{q=1}^p \binom{p}{q} \left(\sum_{j=0}^{M-1} |l_j| \right)^{p-q} \left(\sum_{j=M}^{\infty} |l_j| \right)^q \\
 & \leq C_1 C_2^{-p} K_g K_l^p \sum_{q=1}^p \binom{p}{q} (1 - \lambda^{-M})^{p-q} (\lambda^{-M})^q \leq C_1 C_2^{-p} K_g K_l^p \sum_{q=1}^p \binom{p}{q} (\lambda^{-M})^q \\
 & = C_1 K_g \left(\frac{K_l}{C_2}\right)^p [(\lambda^{-M} + 1)^p - 1] = C_1 K_g \left(\frac{K_l}{C_2}\right)^p \lambda^{-M} [(\lambda^{-M} + 1)^{p-1} + \cdots + 1] \\
 & \leq C_1 K_g \left(\frac{K_l}{C_2}\right)^p \lambda^{-M} \sum_{i=0}^{p-1} 2^i = C_1 K_g \left(\frac{K_l}{C_2}\right)^p \lambda^{-M} (2^p - 1) \leq C_1 K_g \left(\frac{2K_l}{C_2}\right)^p \lambda^{-M},
 \end{aligned}$$

gdzie $\lambda = \rho_l^{-(1-\varepsilon)}$. Z oszacowania tego wyniku, że

$$\begin{aligned}
 & \sup_{\mathbf{x} \in \mathbb{R}^{\infty}} \sum_{p=1}^P \left| \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} - \sum_{i_1=0}^{M-1} \cdots \sum_{i_p=0}^{M-1} h_p(i_1, \dots, i_p) \prod_{j=1}^p x_{n-i_j} \right| \leq \\
 & \leq C_1 K_g \lambda^{-M} \sum_{p=1}^P \left(\frac{2K_l}{C_2}\right)^p \leq C_1 K_g \lambda^{-M} \left(\frac{2K_l}{C_2}\right) \frac{1 - \left(\frac{2K_l}{C_2}\right)^P}{1 - \left(\frac{2K_l}{C_2}\right)}
 \end{aligned}$$

$$\leq C_1 K_g \lambda^{-M} \left(\frac{2K_l}{C_2} \right) \frac{1}{1 - \left(\frac{2K_l}{C_2} \right)} = \mathcal{O}(\lambda^{-M}). \quad (2.17)$$

Ostatecznie, łącząc (2.13) z (2.16) i (2.17), otrzymujemy

$$e_{P,M} \leq c_1 C^{-P} + c_2 \lambda^{-M},$$

dla $c_1 = \frac{C_1 K_g K_l}{C_2 - K_l} > 0$ i $c_2 = \frac{2C_1 K_g K_l}{C_2 - 2K_l} > 0$, co kończy dowód. \square

Rozdział 3

Reprezentacje systemów oparte o szeregi ortogonalne

W rozdziale tym omówimy reprezentacje oparte o rozwinięcie Wienera oraz rozwinięcie w chaos wielomianowy, określane szerzej jako reprezentacje ortogonalne. Mimo, że oba te zagadnienia są ze sobą silnie związane, to ze względów historycznych potraktujemy je rozdzielnie, na końcu rozdziału proponując jednolity sposób ich opisu.

Koncepcje leżące u podstaw reprezentacji opartych o szeregi ortogonalne i szeregi Volterry są znacząco różne. W tej pierwszej bowiem, wyjście systemu traktuje się jako zmienną losową, którą następnie rozwija się w szereg nieskorelowanych zmiennych bazowych stanowiących pewne, w ogólności nieliniowe przekształcenie sygnału wejściowego. W omawianym przypadku reprezentacja ortogonalna ma więc z definicji charakter stochastyczny.

Będziemy zakładali, że wejście systemu, $\{X_n\}$, jest ciągiem *i.i.d.* zmiennych losowych określonym na pewnej przestrzeni probabilistycznej (Ω, \mathcal{F}, P) , gdzie standardowo Ω oznacza przestrzeń zdarzeń elementarnych, $\mathcal{F} = \sigma(\{X_n\})$ to σ -ciało generowane przez ciąg $\{X_n\}$, a P to miara probabilistyczna. Z kolei o (niezakłóconym) wyjściu systemu, $m_{\text{sys}}[\mathbf{X}_n]$, będziemy zakładali, że jest zmienną losową należącą do przestrzeni $L_2(\Omega, \mathcal{F}, P)$ zmiennych losowych ze skończoną wariancją, określonych na (Ω, \mathcal{F}, P) . O systemie, którego charakterystyka spełnia ten warunek będziemy mówili, że należy do *klasy Wienera* (w odniesieniu do danego rozkładu wejścia). Z przestrzenią $L_2(\Omega, \mathcal{F}, P)$ będziemy wiązali iloczyn skalarny, zdefiniowany przez

$$\langle U, V \rangle = \mathbb{E}\{UV\},$$

dla $U, V \in L_2(\Omega, \mathcal{F}, P)$.

3.1 Ogólna postać reprezentacji ortogonalnej

W ogólności, reprezentacja systemów nieliniowych oparta o szeregi ortogonalne przyjmuje postać

$$W[\mathbf{x}_n] = g_0 + \sum_{p=1}^{\infty} G_p[\mathbf{x}_n; g_p], \quad (3.1)$$

gdzie g_0 jest stałą (którą, podobnie jak w przypadku reprezentacji volterrowskiej, będziemy dla wygody oznaczali przez $G_0[\mathbf{x}_n; g_0]$), a przez $G_p : \mathbb{R}^{\infty} \times \mathcal{K}_p \rightarrow \mathbb{R}$ oznaczać będziemy *funkcjonał Wienera* p -tego rzędu. Dla dowolnych jąder, g_i i g_j , funkcyjonały te spełniają warunek ortogonalności, tj.

$$\langle G_i, G_j \rangle = \mathbb{E} \{G_i[\mathbf{X}_n; g_i] G_j[\mathbf{X}_n; g_j]\} = \begin{cases} A_i > 0, & \text{dla } i = j \\ 0, & \text{w przeciwnym wypadku,} \end{cases} \quad (3.2)$$

gdzie wartość oczekiwana rozpatrywana jest względem wektora \mathbf{X}_n . Funkcyjonały spełniające powyższy warunek powstają przez ortogonalizację funkcyjonałów Volterra, stąd przyjmują postać

$$G_p[\mathbf{x}_n; g_p] = g_{0(p)} + \sum_{r=1}^p H_r[\mathbf{x}_n; g_{r(p)}], \quad (3.3)$$

gdzie $g_{p(p)} = g_p \in \mathcal{K}_p$ stanowi jądro p -tego rzędu, a przez $g_{r(p)} \in \mathcal{K}_r$ dla $r < p$ oznacza się tzw. *jądra pochodne*, powstałe przez pewne przekształcenie jądra właściwego, g_p .

Zauważmy, że inaczej niż w przypadku reprezentacji opartej o szereg Volterra, dany system nie posiada unikalnej reprezentacji ortogonalnej, bowiem postać i jądra rozwinięcia (3.1) zależne są od rozkładu prawdopodobieństwa sygnału wejściowego.

Podobnie jak w przypadku reprezentacji volterrowskiej, także i tu będziemy rozważali podwójnie obciętą reprezentację ortogonalną

$$W_{P,M}[\mathbf{x}_n] = g_0 + \sum_{p=1}^P G_{p,M}[\mathbf{x}_n; g_p],$$

opartą o funkcyjonały

$$G_{p,M}[\mathbf{x}_n; g_p] = g_{0(p)} + \sum_{r=1}^p H_{r,M}[\mathbf{x}_n; g_{r(p)}].$$

3.2 Wejście o rozkładzie gaussowskim. Szereg Wienera

Dla zachowania przejrzystości, wyprowadzenia wzorów przedstawionych w tej sekcji zostały przeniesione do dodatku 3.B (strona 56).

Założmy teraz, że wejście systemu, $\{X_n\}$, jest ciągiem *i.i.d.* zmiennych losowych o rozkładzie normalnym, $X_n \sim \mathcal{N}(0, \sigma^2)$. Wtedy, pierwsze trzy funkcjonały Wienera wyrażone są następującymi wzorami:

$$\begin{aligned} G_1[\mathbf{x}_n; g_1] &= \sum_{i=0}^{\infty} g_1(i) x_{n-i}, \\ G_2[\mathbf{x}_n; g_2] &= -\sigma^2 \sum_{i=0}^{\infty} g_2(i, i) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j) x_{n-i} x_{n-j}, \\ G_3[\mathbf{x}_n; g_3] &= -3\sigma^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_3(i, j, j) x_{n-i} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} g_3(i, j, k) x_{n-i} x_{n-j} x_{n-k}, \end{aligned}$$

tak więc, zgodnie z formalizmem wprowadzonym w równaniu (3.3), jądra pochodne powyższych funkcjonałów (niebędące tożsamościowo równe zero) dane są przez

$$g_{0(2)} = -\sigma^2 \sum_{i=0}^{\infty} g_2(i, i) \quad \text{oraz} \quad g_{1(3)}(i) = -3\sigma^2 \sum_{j=0}^{\infty} g_3(i, j, j).$$

O ile postać funkcjonału G_1 jest tożsama z postacią funkcjonału Voltery pierwszego rzędu, o tyle kolejne funkcjonały Wienera są już istotnie różne od ich volterrowskich odpowiedników. W ogólności, dla wejścia gaussowskiego, p -ty funkcjonał Wienera przedstawia się wzorem

$$G_p[\mathbf{x}_n; g_p] = \begin{cases} g_{0(p)} + \sum_{r=1}^{\frac{p}{2}} H_r[\mathbf{x}_n; g_{2r(p)}], & \text{dla } p \text{ parzystego,} \\ \sum_{r=1}^{\frac{p+1}{2}} H_r[\mathbf{x}_n; g_{2r-1(p)}], & \text{dla } p \text{ nieparzystego,} \end{cases}$$

co w zwartej formie zapisać można jako

$$G_p[\mathbf{x}_n; g_p] = \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} H_{p-2r}[\mathbf{x}_n; g_{p-2r(p)}],$$

gdzie jądra kolejnych funkcjonałów $H_p[\cdot, \cdot]$ dane są przez (zob. stwierdzenie 3.2)

$$g_{p-2r(p)}(j_1, \dots, j_{p-2r}) = \frac{(-1)^r p! (\sigma^2)^r}{(p-2r)! r! 2^r} \sum_{i_1=0}^{\infty} \cdots \sum_{i_r=0}^{\infty} g_p(i_1, i_1, \dots, i_r, i_r, j_1, \dots, j_{p-2r}).$$

Wyrazy stojące przed pierwszym znakiem sumowania nieprzypadkowo przypominają wartości współczynników wielomianów Hermite'a (patrz dodatek 3.A.1). Niech

$$\Psi_{(i_1, \dots, i_p)}(\mathbf{x}_n) = \mathcal{H}_{k_1}(x_{n-\tau_1}) \cdots \mathcal{H}_{k_p}(x_{n-\tau_p}), \quad (3.4)$$

oznacza wielomian Hermite'a p -tego stopnia wielu zmiennych, gdzie k_j oznacza liczbę wystąpień liczby τ_j w krotce (i_1, \dots, i_p) , tak że spełniony jest warunek $\sum_{i=1}^p k_i = p$. Zgodnie z stwierdzeniem 3.5, dowolny funkcjonał Wienera, można wtedy przedstawić w formie liniowej kombinacji funkcji (3.4),

$$G_p[\mathbf{x}_n; g_p] = \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) \Psi_{(i_1, \dots, i_p)}(\mathbf{x}_n). \quad (3.5)$$

3.2.1 Wyznaczanie jąder funkcjonałów

Rozważmy iloczyn skalarny funkcjonałów G_p . Niech $\{\tau_1, \dots, \tau_t\}$ będzie zbiorem wszystkich unikalnych elementów zbioru $\{i_1, \dots, i_p, j_1, \dots, j_r\}$. Korzystając z faktu że zmienne losowe $\{X_n\}$ są *i.i.d.*, otrzymujemy

$$\begin{aligned} \langle \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n), \Psi_{(j_1, \dots, j_r)}(\mathbf{X}_n) \rangle &= \langle \mathcal{H}_{k_1}(X_{n-\tau_1}), \mathcal{H}_{l_1}(X_{n-\tau_1}) \rangle \\ &\dots \langle \mathcal{H}_{k_t}(X_{n-\tau_t}), \mathcal{H}_{l_t}(X_{n-\tau_t}) \rangle, \end{aligned} \quad (3.6)$$

gdzie $k_i = \#\{\tau_i \in \{i_1, \dots, i_p\}\}$ oznacza liczbę wystąpień elementu τ_i w zbiorze $\{i_1, \dots, i_p\}$, a $l_i = \#\{\tau_i \in \{j_1, \dots, j_r\}\}$. Rozważmy dwa przypadki:

1. Niech $p \neq r$. Wtedy co najmniej jeden element prawej strony równania (3.6) jest iloczynem skalarnym wielomianów Hermite'a różnych stopni, a więc

$$\langle \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n), \Psi_{(j_1, \dots, j_r)}(\mathbf{X}_n) \rangle = 0.$$

2. Niech $p = r$. Iloczyn (3.6) jest różny od zera wtedy i tylko wtedy, gdy dla każdego $i = 1, \dots, t$, spełniony jest warunek $k_i = l_i$.

Tak więc, z faktu, że $\langle \mathcal{H}_p^{[\sigma^2]}(X), \mathcal{H}_p^{[\sigma^2]}(X) \rangle = p!(\sigma^2)^p$, wynika, że iloczyn (3.6) wynosi

$$\langle \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n), \Psi_{(j_1, \dots, j_r)}(\mathbf{X}_n) \rangle = \begin{cases} (\sigma^2)^p \prod_{i=1}^t k_i!, & \text{dla } (i_1, \dots, i_p) \in \pi(j_1, \dots, j_r) \\ 0, & \text{w przeciwnym wypadku,} \end{cases} \quad (3.7)$$

gdzie $\pi(j_1, \dots, j_r)$ oznacza zbiór wszystkich krotek otrzymanych przez permutację elementów krotki (j_1, \dots, j_r) . Korzystając z (3.5), otrzymujemy

$$\langle W[\mathbf{X}_n], \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n) \rangle = (\sigma^2)^p \prod_{i=1}^t k_i! \sum_{(i_1, \dots, i_p) \in \pi(i_1, \dots, i_r)} g_p(i_1, \dots, i_p).$$

Jeśli więc przyjmiemy reprezentację opartą o jądra trójkątne (analogicznie jak w przypadku reprezentacji volterrowskich), wtedy element jądra p -tego rzędu o indeksie (i_1, \dots, i_p) , gdzie $i_1 \geq i_2 \geq \dots \geq i_p$ dany jest przez wzór

$$g_p(i_1, \dots, i_p) = (\sigma^2)^{-p} \left(\prod_{i=1}^t k_i! \right)^{-1} \langle W[\mathbf{X}_n], \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n) \rangle. \quad (3.8)$$

Jeżeli system o charakterystyce $m_{\text{sys}}[\cdot]$ należy do klasy Wienera to, jak pokażemy w następnej sekcji, przyjmuje reprezentację (3.1), której jądra wyznacza się analogicznie jak we wzorze (3.8), tj.

$$g_p(i_1, \dots, i_p) = (\sigma^2)^{-p} \left(\prod_{i=1}^t k_i! \right)^{-1} \langle m_{\text{sys}}[\mathbf{X}_n], \Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n) \rangle. \quad (3.9)$$

3.2.2 Własności aproksymacyjne

Opierając się o wyniki dotyczące rozwinięcia w chaos wielomianowy, omówimy teraz własności aproksymacyjne szeregu Wienera.

Twierdzenie 3.1. *Niech $m_{\text{sys}}[\cdot]$ będzie charakterystyką pewnego systemu należącego do klasy Wienera (w odniesieniu do wejścia o rozkładzie normalnym). Wtedy, obcięty szereg Wienera, $W_{P,M}[\cdot]$, którego jądra dane są wzorem (3.9), zbiega do $m_{\text{sys}}[\cdot]$ w sensie średniokwadratowym, tj. dla błędu zdefiniowanego przez*

$$e_{P,M} = \mathbb{E}\{m_{\text{sys}}[\mathbf{X}_n] - W_{P,M}[\mathbf{X}_n]\}^2,$$

zachodzi

$$\lim_{P,M \rightarrow \infty} e_{P,M} = 0.$$

Dowód. Dowód przedstawiony został m.in. w pracy [28, tw. 2.2]. □

W konsekwencji powiemy, że systemy klasy Wienera przyjmują reprezentację (3.1), tj. $m_{\text{sys}}[\mathbf{x}_n] \simeq W[\mathbf{x}_n]$. Należy zaznaczyć jednak, że zbieżność średniokwadratowa nie implikuje zbieżności punktowej, tak więc w ogólności, istnieje taki sygnał wejściowy $\mathbf{u} \in \mathbb{R}^\infty$, że $m_{\text{sys}}[\mathbf{u}] \neq W[\mathbf{u}]$.

Podobnie jak w przypadku klasy systemów z pamięcią zanikającą, klasa systemów Wienera zawiera w sobie systemy z pamięcią nieskończoną, która w pewnym, jakościowo odmiennym sensie również musi *zanikać*. Zauważmy bowiem, że wariancja szeregu Wienera wynosi

$$\text{Var}\{W[\mathbf{X}_n]\} = \mathbb{E}\{W^2[\mathbf{X}_n]\} - \mathbb{E}\{W[\mathbf{X}_n]\}^2,$$

co ze względu na ortogonalność funkcjonałów i fakt, iż $\mathbb{E}\{W[\mathbf{X}_n]\}^2 = g_0$, oznacza, że zachodzi następująca równość

$$\text{Var}\{W[\mathbf{X}_n]\} = \sum_{P=1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p^2(i_1, \dots, i_p) \mathbb{E}\{\Psi_{(i_1, \dots, i_p)}^2(\mathbf{X}_n)\}.$$

Z faktu, że wariancja reprezentacji zbiega do wariancji systemu oraz warunku (3.7) wynika więc, że jądra systemów należących do klasy Wienera są sumowalne z kwadratem, tzn.

$$\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p^2(i_1, \dots, i_p) < \infty,$$

co w praktyce oznacza, że system należący do klasy Wienera ma pamięć zanikającą, w tym sensie, że współczynniki g_p odpowiednio szybko zbiegają do zera.

Jak pokazujemy poniżej, twierdzenie 3.1 implikuje zbieżność błędu aproksymacji do zera.

Wniosek 3.1. Niech $\Delta_{apr.}^{P,M}$ oznacza oparty o funkcję kwadratową błąd aproksymacji dla klasy modeli należących do zbioru

$$\mathcal{M} = \{W_{p,m}[\cdot] : p \leq P, m \leq M\}.$$

Jeśli dla systemu opisanego równaniem (1.1), w którym zakłócenie spełnia założenie 1.1, zachodzi

$$\text{Var}\{m_{\text{sys}}[\mathbf{X}_n]\} < \infty, \quad (3.10)$$

to prawdziwe są następujące stwierdzenia:

1. Model $W_{P,M}[\cdot]$, którego jądra określone są za pomocą rzutowania ortogonalnego (jak we wzorze (3.9)), minimalizuje błąd aproksymacji $\Delta_{apr.}^{P,M}$.
2. Błąd aproksymacji $\Delta_{apr.}^{P,M}$ zbiega do zera, tj. $\lim_{P,M \rightarrow \infty} \Delta_{apr.}^{P,M} = 0$.

Dowód. Zwróćmy uwagę, że wariancja wyjścia systemów spełniających podane warunki wynosi

$$\text{Var}\{Y_n\} = \text{Var}\{m_{\text{sys}}[\mathbf{X}_n]\} + \sigma_Z^2,$$

a więc jeśli spełniony jest warunek (3.10) to system należy do klasy Wienera i przybiera reprezentację

$$m_{\text{sys}}[\mathbf{X}_n] = \sum_{p=0}^{\infty} G_p[\mathbf{X}_n; g_p]. \quad (3.11)$$

Podobnie, łatwo sprawdzić, że ze względu na fakt, iż zakłócenie ma zerową wartość oczekiwaną i jest niezależne od wejścia systemu, błąd średniokwadratowy modelu $m[\cdot]$ wynosi

$$L(m) = \mathbb{E}\{(m_{\text{sys}}[\mathbf{X}_n] + Z_n) - m[\mathbf{X}_n]\}^2 = \mathbb{E}\{m_{\text{sys}}[\mathbf{X}_n] - m[\mathbf{X}_n]\}^2 + \sigma_Z^2.$$

Dokonując postawienia (3.11) i $m[\mathbf{X}_n] = \sum_{p=0}^P G_{p,M}[\mathbf{X}_n; g_p']$, otrzymujemy

$$\begin{aligned} L(m) &= \sum_{p=0}^P \sum_{i_1=0}^M \cdots \sum_{i_p=0}^M [g_p(i_1, \dots, i_p) - g_p'(i_1, \dots, i_p)]^2 \mathbb{E}\{\Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n)\}^2 \\ &\quad + R_{P,M} + \sigma_Z^2, \end{aligned}$$

gdzie

$$\begin{aligned} R_{P,M} &= \sum_{p=0}^P \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_p=M+1}^{\infty} g_p^2(i_1, \dots, i_p) \mathbb{E}\{\Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n)\}^2 + \\ &\quad + \sum_{p=P+1}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p^2(i_1, \dots, i_p) \mathbb{E}\{\Psi_{(i_1, \dots, i_p)}(\mathbf{X}_n)\}^2 \end{aligned}$$

jest resztą, niezależną od jąder g_p' . Tym samym, w danym zbiorze \mathcal{M} , jądra modelu minimalizującego wyrażenie $L(m)$, $m^* = \sum_{p=0}^P G_{p,M}[\mathbf{x}_n; g_p^*]$, dane są przez $g_p^*(i_1, \dots, i_p) = g_p(i_1, \dots, i_p)$, czym udowodniliśmy stwierdzenie 1.

W celu udowodnienia stwierdzenia 2, zauważmy, że błąd średniokwadratowy charakterystyki m_{sys} wynosi $L(m_{\text{sys}}) = \sigma_Z^2$. W związku z powyższym, błąd aproksymacji wynosi

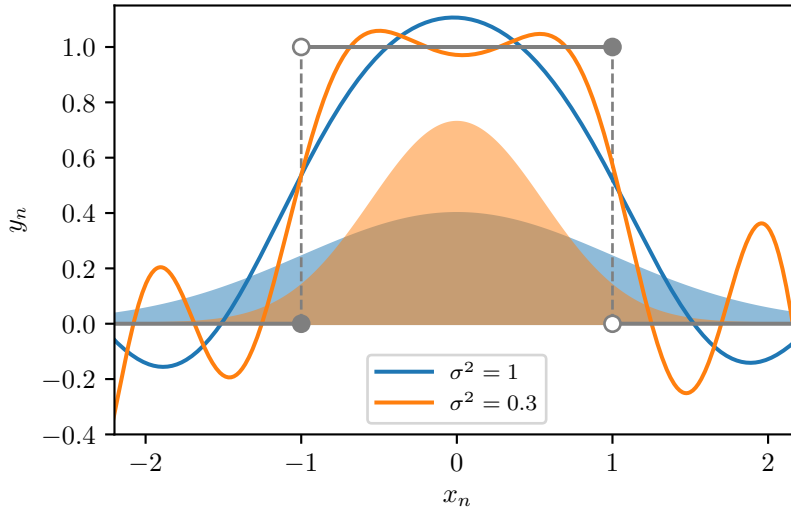
$$\Delta_{\text{apr.}}^{P,M} = \mathbb{E}\{m_{\text{sys}}[\mathbf{X}_n] - m^*[\mathbf{X}_n]\}^2.$$

Zgodnie z twierdzeniem 3.1 wyrażenie to zbiega do zera, gdy $P, M \rightarrow \infty$. \square

Przykład 3.1. Rozważmy system o strukturze LN (rys. 2.1 z przykładu 2.1), którego podsystem dynamiczny opisany jest odpowiedzią impulsową o skończonej długości, $h_i = 2^{-i}$ dla $i = 0, 1, 2$, a charakterystyka nieliniowa dana jest przez wzór

$$\varphi(x) = \begin{cases} 1, & \text{dla } x \in (-1, 1], \\ 0, & \text{w przeciwnym wypadku.} \end{cases}$$

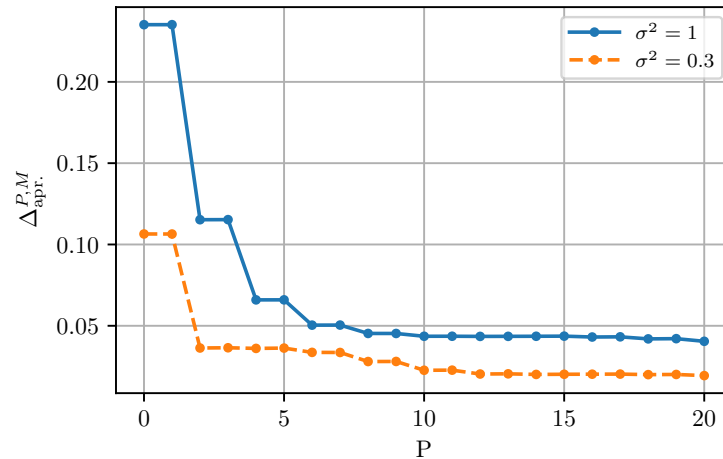
Zauważmy, że ze względu na nieciągłość funkcji $\varphi(x)$, system ten ma własność pamięci zanikającej jedynie dla sygnałów wejściowych należących do zbioru $K = \{x_n : \|x_n\|_\infty < \|h\|_1^{-1} = \frac{4}{7}\}$, w związku z czym możliwość wykorzystania reprezentacji volterrowskiej jest znacząco ograniczona. Niemniej jednak, jako że rozpatrywany system należy do klasy Wienera, istnieje możliwość jego rozwinięcia w szereg ortogonalny (3.1).



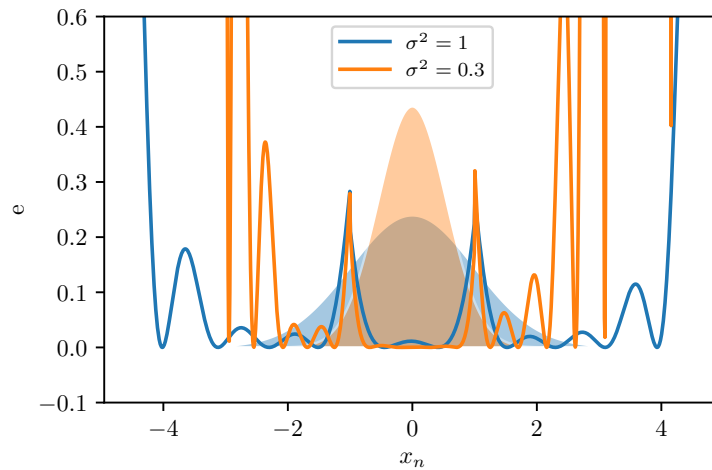
Rysunek 3.1: Przekrój charakterystyki modelu dla ustalonego $x_{n-1} = x_{n-2} = \dots = 0$. Na wykresie przedstawiono także wykresy funkcji gęstości prawdopodobieństwa sygnałów wejściowych

Na wykresie 3.1 przedstawiono przekrój charakterystyk dwóch różnych modeli wienerowskich rzędu $P = 20$, których jądra zostały wyznaczone metodą Monte Carlo, w oparciu o wzór (3.8), dla gaussowskich sygnałów wejściowych o wariancji $\sigma^2 = 1$ i $\sigma^2 = 0,3$. Pomimo, iż kwadrat różnicy pomiędzy wyjściem modelu, a systemem, $e(\mathbf{x}) = (Y_n - G_{(P,M)}[\mathbf{x}])^2$, dla przedstawionego zakresu jest średnio

duży wyższy dla modelu drugiego ($\sigma^2 = 0,3$), rys. 3.2b, jak wynika z wykresu 3.2a, błąd aproksymacji jest wyraźnie wyższy dla modelu pierwszego ($\sigma^2 = 1$). Wynika to z faktu, że błąd aproksymacji uzależniony jest od miary prawdopodobieństwa sygnału wejściowego. W związku z tym, model drugi, do estymacji którego użyto sygnału o mniejszej dyspersji, lepiej odwzorowuje system w okolicy punktu zerowego, niż model pierwszy. Przekłada się to na mniejszy błąd średniokwadratowy, gdyż wokół zera, charakterystyka rozpatrywanego systemu jest gładka. Z drugiej strony model pierwszy dokładniej aproksymuje system dla sygnałów wejściowych o większej amplitudzie.



(a)



(b)

Rysunek 3.2: Porównanie błędów modeli dla dwóch różnych wariacji sygnału wejściowego. Rysunek (a) przedstawia błąd aproksymacji modelu w zależności od rzędu rozwinięcia, P ; na rysunku (b) pokazano kwadrat różnicy między wyjściem modelu i systemem w punkcie $\mathbf{x}_n = [x_n, 0, 0, \dots]$

3.3 Wejście o rozkładzie nieguassowskim. Rozwinięcie w chaos wielomianowy

Postać szeregu Wienera wynika z założenia o rozkładzie normalnym sygnału wejściowego. Zasadnym staje się więc pytanie o możliwość konstrukcji szeregów ortogonalnych w odniesieniu do sygnałów wejściowych o innych rozkładach prawdopodobieństwa. Jak się okazuje, dla pewnej klasy rozkładów wejścia, uogólnienie takie jest możliwe. Poniższe twierdzenie przedstawia warunki zbieżności reprezentacji (3.1) do systemu.

Twierdzenie 3.2. *Niech $\{X_n\}$ będzie ciągiem i.i.d. ciągłych i ograniczonych zmiennych losowych, spełniających warunek $P(a \leq X_n \leq b) = 1$, dla pewnych $a, b < \infty$. Jeśli system opisany charakterystyką $m_{sys}[\mathbf{X}_n]$, należy do klasy Wienera (względem rozkładu zmiennej X_n), to podwójnie obcięty szereg ortogonalny, $W_{P,M}[\cdot]$, którego jądra wyznaczono za pomocą rzutowania ortogonalnego (jak we wzorze (3.9)), zbiega do $m_{sys}[\cdot]$ w sensie średniokwadratowym, gdy P i M dąży do nieskończoności.*

Dowód. Zauważmy, że założenia o sygnale wejściowym implikują nierówność

$$\mathbb{E}\{X_n^k\} = \int_b^a x^k f(x) dx \leq \int_b^a |x^k| f(x) dx \leq \max(|a|, |b|)^k,$$

co jest równoznaczne z istnieniem wszystkich momentów zmiennej losowej X_n . Wobec tego, zbieżność szeregu ortogonalnego do zmiennej Y_n wynika z [28, wniosek 3.10]. \square

W tym miejscu należy podkreślić, że warunki istnienia reprezentacji ortogonalnej, podane w powyższym twierdzeniu są *wystarczające*, ale nie *konieczne*, istnieje zatem możliwość rozważenia innych rozkładów (zob. [28]); mając jednak na względzie zastosowania praktyczne, ograniczymy się do warunków podanych powyżej.

Wniosek 3.2. *Stwierdzenia zawarte we wniosku 3.1 są prawdziwe także dla reprezentacji opartych o rozkłady niegaussowskie, spełniające warunki twierdzenia 3.2.*

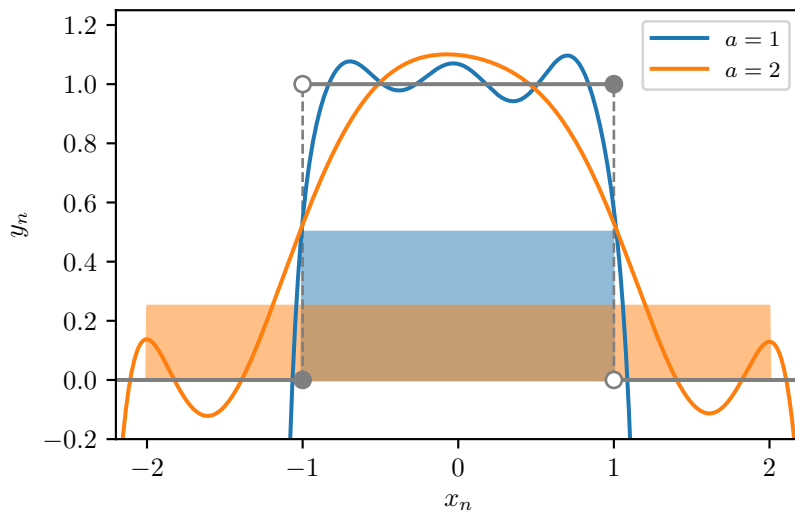
Przykład 3.2. Przykładową rodziną rozkładów spełniających warunki twierdzenia są rozkłady jednostajne, dla których reprezentacja ortogonalna opiera się o wielomiany Legendre'a (zob. dodatek (3.A.2)); dla wejścia $X_n \sim \mathcal{U}[-a, a]$, $0 < a < \infty$, funkcje składowe szeregu przyjmują wtedy postać

$$\Psi_{(i_1, \dots, i_p)}(\mathbf{x}_n) = \mathcal{L}_{k_1}^{[a]}(x_{n-\tau_1}) \dots \mathcal{L}_{k_p}^{[a]}(x_{n-\tau_p}),$$

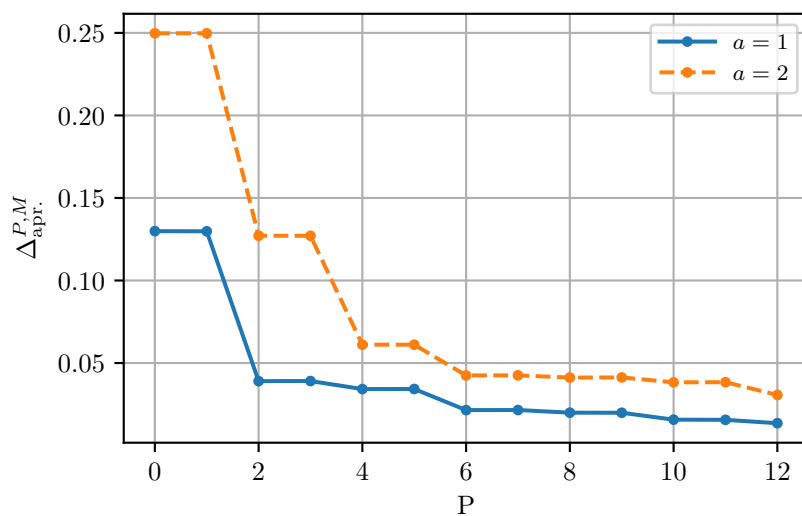
gdzie, jak poprzednio, $k_i = \#(\tau_i \in \{i_1, \dots, i_p\})$.

Wróćmy do systemu z przykładu 3.1. Na rysunku 3.3 przedstawiono charakterystykę dwóch różnych modeli ortogonalnych tego systemu, uzyskanych w oparciu o sygnał wejściowy mający rozkład jednostajny $X_n \sim \mathcal{U}[-a, a]$ i ograniczony przez

$a = 1$ oraz $a = 2$. Z kolei na rysunku 3.4 pokazano wykres przedstawiający błąd aproksymacji modelu w zależności od P .



Rysunek 3.3: Przekrój charakterystyki modelu dla ustalonego $x_{n-1} = x_{n-2} = \dots = 0$. Na wykresie przedstawiono także wykresy funkcji gęstości prawdopodobieństwa sygnałów wejściowych



Rysunek 3.4: Błąd aproksymacji modelu w zależności od rzędu rozwinięcia

3.4 Konstrukcja modeli ortogonalnych

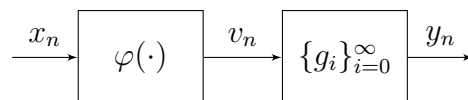
Konstrukcja słowników dla modeli ortogonalnych jest analogiczna do konstrukcji modeli volterrowskich, przedstawionej w sekcji 2.4 i różni się tylko sposobem definicji elementów słownika, $\varphi_\alpha(\cdot)$, które dla uproszczenia będziemy normalizować,

$$\varphi_\alpha(\mathbf{x}_n) = \frac{\Psi_{(\alpha_1, \dots, \alpha_p)}(\mathbf{x}_n)}{\sqrt{\langle \Psi_{(\alpha_1, \dots, \alpha_p)}, \Psi_{(\alpha_1, \dots, \alpha_p)} \rangle}},$$

gdzie funkcje $\Psi_{(\alpha_1, \dots, \alpha_p)}(\cdot)$ są ortogonalne względem rozważanego rozkładu sygnału wejściowego.

3.5 Studium przypadku – aproksymacja systemu NL

Zbadamy teraz szybkość zbieżności błędu aproksymacji obciążonej reprezentacji Wienera dla szerokiej klasy systemów Hammersteina, z nieskończoną pamięcią oraz charakterystyką nieliniową, która jest całkowalna z kwadratem.



Rysunek 3.5: System Hammersteina

Rozpatrzmy system o strukturze Hammersteina (NL, rys. (3.5)), opisany równaniem

$$y_n = m_{\text{NL}}[\mathbf{x}_n] = \sum_{i=0}^{\infty} g_i \varphi(x_{n-i}). \quad (3.12)$$

Niech

$$L_\omega^2(\mathcal{X}) = \left\{ f : \mathcal{X} \rightarrow \mathbb{R} : \int_{-\infty}^{\infty} f^2(x) \omega(x) dx < \infty \right\}$$

oznacza przestrzeń funkcji całkowalnych z kwadratem względem pewnej miary $\omega(x)$, z iloczynem skalarnym danym przez

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)\omega(x)dx$$

i normą $\|f\|_{L_\omega^2} = (\int_{-\infty}^{\infty} f^2(x)\omega(x)dx)^{\frac{1}{2}}$. Zwrócimy uwagę na podzbiór funkcji należących do przestrzeni Sobolewa $H_\omega^q(\mathcal{X}) \subset L_\omega^2(\mathcal{X})$, definiowanych następująco

$$H_\omega^q(\mathcal{X}) = \left\{ f : \mathcal{X} \rightarrow \mathbb{R} : \frac{d^p f}{dx^p} \in L_\omega^2(\mathcal{X}), 0 \leq p \leq q \right\}.$$

W stosunku do poszczególnych bloków systemu poczynimy następujące założenia:

Założenie 3.1. *Podsystem liniowy jest asymptotycznie stabilny.*

Założenie 3.2. *Charakterystyka nieliniowa $\varphi(\cdot)$ należy do przestrzeni Sobolewa, $\varphi \in H_\omega^q(\mathcal{X})$.*

Założenie 3.1 pozwala ograniczyć kolejne komponenty odpowiedzi impulsowej, analogicznie jak w równaniu (2.9),

$$|g_i| \leq k\rho^i, \quad (3.13)$$

gdzie $0 < k < \infty$ i $0 < \rho < 1$ to pewne stałe.

Dzięki założeniu 3.2, funkcja $\varphi(\cdot)$ może zostać rozwinięta w szereg wielomianów ortogonalnych. Na przykład, dla $\mathcal{X} = \mathbb{R}$, $\omega(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ funkcja $\varphi(\cdot)$ przyjmuje reprezentację opartą o wielomiany Hermite'a,

$$\varphi(x) = \sum_{i=0}^{\infty} \alpha_i \mathcal{H}_i(x),$$

ze współczynnikami określonymi przez $\alpha_i = \langle \varphi, \mathcal{H}_i \rangle$. Tym samym, równanie (3.12) przybiera postać

$$m_{\text{NL}}[\mathbf{x}_n] = \sum_{i=0}^{\infty} g_i \sum_{j=0}^{\infty} \alpha_j \mathcal{H}_j(x_{n-i}). \quad (3.14)$$

Analogicznie, dla $\mathcal{X} = [-1, 1]$, $\omega(x) = 1$, funkcja $\varphi(\cdot)$ może zostać przedstawiona za pomocą szeregu wielomianów Legendre'a,

$$m_{\text{NL}}[\mathbf{x}_n] = \sum_{i=0}^{\infty} g_i \sum_{j=0}^{\infty} \alpha_j \mathcal{L}_j(x_{n-i}).$$

W efekcie, funkcjonal $m_{\text{NL}}[\cdot]$ może zostać rozwinięty w szereg ortogonalny. Zachodzi więc tożsamość

$$m_{\text{NL}}[\mathbf{x}_n] = k_0 + \sum_{p=1}^{\infty} G_p[\mathbf{x}_n; k_p],$$

gdzie $k_0 = \alpha_0$, a dla $p \geq 1$, jądra rozwinięcia dane są przez

$$k_p(i_1, \dots, i_p) = \begin{cases} \alpha_p g_i & \text{dla } i = i_1 = \dots = i_p, \\ 0, & \text{w przeciwnym wypadku.} \end{cases} \quad (3.15)$$

3.5.1 Błąd aproksymacji

Przedstawimy teraz dwa lematy dotyczące błędów aproksymacji dla dwóch różnych klas nieliniowości oraz sygnałów wejściowych.

Lemat 3.1. Niech $\Delta_{apr.}^{P,M}$ oznacza błąd aproksymacji systemu Hammersteina opisanego równaniem (3.12) dla klasy modeli ortogonalnych opartych o wielomiany Hermite'a i sygnału wejściowego będącego ciągiem i.i.d. zmiennych losowych mających rozkład normalny, $X_n \sim \mathcal{N}(0, 1)$. Jeśli system spełnia założenia 3.1 oraz 3.2 (dla $\mathcal{X} = \mathbb{R}$, $\omega(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$) to zachodzi nierówność

$$\Delta_{apr.}^{P,M} \leq K(cP^{-q} + \|\varphi\|_{L_w^2}^2 \rho^{2M}),$$

gdzie $0 < K, c < \infty$ są pewnymi stałymi, zależnymi od własności systemu.

Dowód. Zgodnie z wnioskiem 3.2 i wzorem (3.15), najlepszy model w zbiorze $\mathcal{M} = \{W_{p,m}[\cdot] : p \leq P, m \leq M\}$ dany jest przez

$$m^*[\cdot] = \alpha_0 + \sum_{i=0}^{M-1} g_i \sum_{j=1}^P \alpha_j \mathcal{H}_j(\cdot).$$

W konsekwencji, wykorzystując równoważną definicję systemu (3.14), błąd aproksymacji można przedstawić jako

$$\Delta_{apr.}^{P,M} = \mathbb{E} \left\{ \sum_{i=0}^{M-1} g_i \sum_{j=P+1}^{\infty} \alpha_j \mathcal{H}_j(X_{n-i}) + \sum_{i=M}^{\infty} g_i \sum_{j=0}^{\infty} \alpha_j \mathcal{H}_j(X_{n-i}) \right\}^2$$

co, za sprawą ortogonalności wielomianów, upraszcza się do

$$\begin{aligned} \Delta_{apr.}^{P,M} &= \sum_{i=0}^{M-1} g_i^2 \sum_{j=P+1}^{\infty} \alpha_j^2 \mathbb{E} \{ \mathcal{H}_j^2(X_n) \} + \sum_{i=M}^{\infty} g_i^2 \sum_{j=0}^{\infty} \alpha_j^2 \mathbb{E} \{ \mathcal{H}_j^2(X_n) \} \\ &= e_P \sum_{i=0}^{M-1} g_i^2 + \|\varphi\|_{L_w^2}^2 \sum_{i=M}^{\infty} g_i^2, \end{aligned}$$

gdzie $e_P = \sum_{j=P+1}^{\infty} \alpha_j^2$ oznacza błąd aproksymacji wielomianowej. Korzystając z faktu, że dla wielomianów Hermite'a błąd ten wynosi $e_P = cP^{-q}$, dla pewnego $c > 0$ (zob. np. [4]) i dokonując podstawienia (3.13), otrzymujemy górne ograniczenie błędu aproksymacji,

$$\begin{aligned} \Delta_{apr.}^{P,M} &\leq cP^{-q} k^2 \sum_{i=0}^{M-1} \rho^{2i} + \|\varphi\|_{L_w^2}^2 k^2 \sum_{i=M}^{\infty} \rho^{2i} \\ &= cP^{-q} \frac{k^2}{1 - \rho^2} (1 - \rho^{2M}) + \|\varphi\|_{L_w^2}^2 \frac{k^2}{1 - \rho^2} \rho^{2M} \\ &\leq K(cP^{-q} + \|\varphi\|_{L_w^2}^2 \rho^{2M}), \end{aligned}$$

co kończy dowód. □

Lemat 3.2. Niech $\Delta_{apr.}^{P,M}$ oznacza błąd aproksymacji systemu Hammersteina opisanego równaniem (3.12) dla klasy modeli ortogonalnych opartych o wielomiany Legendre'a i sygnału wejściowego będącego ciągiem i.i.d. zmiennych losowych mających rozkład jednostajny, $X_n \sim \mathcal{U}[-1, 1]$. Jeśli system spełnia założenia 3.1, 3.2 (dla $\mathcal{X} = [-1, 1]$, $\omega(x) = 1$) oraz dodatkowo $\varphi^{(i)}(-1) = \varphi^{(i)}(1) = 0$, $i = 0, \dots, q-1$ to zachodzi nierówność

$$\Delta_{apr.}^{P,M} \leq K(cP^{-2q+1} + \|\varphi\|_{L_\omega^2}^2 \rho^{2M}),$$

gdzie $0 < K, c < \infty$ są pewnymi stałymi, zależnymi od własności systemu.

Dowód. Zgodnie z [38, str. 342], błąd aproksymacji funkcji $\varphi(\cdot)$ za pomocą wielomianów Legendre'a wynosi $e_P = cP^{-2q+1}$, dla pewnego $c > 0$. Pozostała część dowodu jest identyczna z dowodem lematu 3.1. \square

Wykazaliśmy zatem, że dla badanej klasy systemów modele ortogonalne mają korzystne własności aproksymacyjne. W wyrażeniu na błąd aproksymacji można wyróżnić dwie składowe. Szybkość zbieżności pierwszej z nich zależna jest od rzędu modelu i gładkości charakterystyki nieliniowej, przy czym dla rozkładu jednostajnego i wielomianów Legendre'a szybkość zbieżności jest znacznie większa, niż dla rozkładu normalnego i wielomianów Hermite'a. Natomiast szybkość zbieżności drugiej składowej zależna jest od długości pamięci modelu i z dokładnością do stałej, nie zależy od rozkładu wejścia. Zgodnie z wnioskiem 3.1 i 3.2, w obu przypadkach wraz ze wzrostem rzędu i długości pamięci modelu, błąd aproksymacji zbiega do zera, tj. $\lim_{P,M \rightarrow \infty} \Delta_{apr.}^{P,M} = 0$.

Dodatki

3.A Wielomiany ortogonalne

W dodatku przedstawiamy definicje rodziny wielomianów Hermite'a oraz Legendre'a. Wielomiany te wykorzystywane są przy omawianiu rozwinięcia ortogonalnego systemów nieliniowych.

3.A.1 Wielomiany Hermite'a

Wykorzystywane w pracy wielomiany Hermite'a, zdefiniowane są za pomocą wzorów rekurencyjnych:

$$\begin{aligned}\mathcal{H}_0^{[\sigma^2]}(x) &= 1, \\ \mathcal{H}_1^{[\sigma^2]}(x) &= x, \\ \mathcal{H}_{p+1}^{[\sigma^2]}(x) &= x\mathcal{H}_p^{[\sigma^2]}(x) - \sigma^2 \frac{d}{dx}(\mathcal{H}_p^{[\sigma^2]}(x)),\end{aligned}$$

a w postaci jawnej wyrażają się przez

$$\mathcal{H}_p^{[\sigma^2]}(x) = p! \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^r (\sigma^2)^r}{(p-2r)! r! 2^r} x^{p-2r}. \quad (3.16)$$

Wielomiany te są ortogonalne na przedziale $(-\infty, \infty)$ względem miary $\omega(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2})$, tak więc, dla zmiennych losowych o rozkładzie normalnym, $X \sim \mathcal{N}(0, \sigma^2)$, zachodzi

$$\mathbb{E}\{\mathcal{H}_p^{[\sigma^2]}(X)\mathcal{H}_r^{[\sigma^2]}(X)\} = \begin{cases} p!(\sigma^2)^p, & \text{dla } p=r, \\ 0, & \text{w przeciwnym wypadku.} \end{cases}$$

Dla uproszczenia, gdy wariancja σ^2 znana jest z kontekstu, będziemy używali krótszego zapisu $\mathcal{H}_p(\cdot) = \mathcal{H}_p^{[\sigma^2]}(\cdot)$.

3.A.2 Wielomiany Legendre'a

Wielomiany Legendre'a są zdefiniowane rekurencyjnie przez następujące wzory:

$$\begin{aligned}\mathcal{L}_0^{[a]}(x) &= 1, \\ \mathcal{L}_1^{[a]}(x) &= \frac{x}{a}, \\ (p+1)\mathcal{L}_{p+1}^{[a]}(x) &= (2p+1)\frac{x}{a}\mathcal{L}_p^{[a]}(x) - p\mathcal{L}_{p-1}^{[a]}(x).\end{aligned}$$

Wielomiany te są ortogonalne na przedziale $[-a, a]$ względem miary $\omega(x) = 1$. Tym samym, dla zmiennych losowych o rozkładzie jednostajnym, $X \sim \mathcal{U}(-a, a)$, spełniają warunek

$$\mathbb{E}\{\mathcal{L}_p^{[a]}(X)\mathcal{L}_r^{[a]}(X)\} = \begin{cases} \frac{2a}{2p+1}, & \text{dla } p=r, \\ 0, & \text{w przeciwnym wypadku.} \end{cases}$$

Dla uproszczenia, gdy wartość parametru a jest znana z kontekstu, będziemy używali krótszego zapisu $\mathcal{L}_p(\cdot) = \mathcal{L}_p^{[a]}(\cdot)$.

3.B Rozwinięcia Wienera – konstrukcja

Wyprowadzenia dla systemów z czasem ciągłym, analogiczne do zawartych w tym dodatku, zostały przedstawione w monografii [92].

3.B.1 Ortogonalizacja funkcjonałów Volterra

Dla przypomnienia, zakładamy, że wejście systemu, $\{X_n\}$, jest ciągiem *i.i.d.* zmiennych losowych o rozkładzie normalnym, tj. $X_n \sim \mathcal{N}(0, \sigma^2)$, a iloczyn skalarny zdefiniowany jest przez $\langle G_i, G_j \rangle = \mathbb{E}\{G_i[\mathbf{X}; g_i]G_j[\mathbf{X}; g_j]\}$.

Zacniemy od wyznaczenia iloczynu skalarnego dwóch dowolnych funkcjonałów Volterra,

$$\begin{aligned}\langle H_p[\mathbf{X}_n; a_p], H_r[\mathbf{X}_n; b_r] \rangle &= \\ &= \mathbb{E} \left[\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} a_p(i_1, \dots, i_p) \prod_{j=1}^p X_{n-i_j} \sum_{i_1=0}^{\infty} \cdots \sum_{i_r=0}^{\infty} b_r(i_1, \dots, i_r) \prod_{j=1}^r X_{n-i_j} \right] \\ &= \mathbb{E} \left[\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} \sum_{j_1=0}^{\infty} \cdots \sum_{j_r=0}^{\infty} a_p(i_1, \dots, i_p) b_r(j_1, \dots, j_r) \prod_{a=1}^p X_{n-i_a} \prod_{b=1}^r X_{n-j_b} \right] \\ &= \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} \sum_{j_1=0}^{\infty} \cdots \sum_{j_r=0}^{\infty} a_p(i_1, \dots, i_p) b_r(j_1, \dots, j_r) \mathbb{E} \left[\prod_{a=1}^p X_{n-i_a} \prod_{b=1}^r X_{n-j_b} \right].\end{aligned}$$

Wynik powyższego działania zawiera element $I = \mathbb{E}[\prod_{a=1}^p X_{n-i_a} \prod_{b=1}^r X_{n-j_b}]$, który dla $p+r$ będącego liczbą nieparzystą, przyjmuje wartość równą zero (jest to konsekwencją faktu, iż co najmniej jedna ze zmiennych losowych ze zbioru $\{X_n, X_{n-1}, \dots\}$ będzie podniesiona do nieparzystej potęgi). W przeciwnym wypadku, gdy $p+r$ jest liczbą parzystą, można pokazać [92, Dodatek A], że

$$I = \mathbb{E} [X_{n-i_1} \dots X_{n-i_p} X_{n-i_1} \dots X_{n-i_r}] = \sum \prod \mathbb{E} [X_i X_j],$$

gdzie $\sum \prod$ oznacza sumę iloczynów wartości oczekiwanych wszystkich dwuelementowych kombinacji zmiennych ze zbioru $\{X_{n-i_1}, \dots, X_{n-i_p}, X_{n-i_1}, \dots, X_{n-i_r}\}$, np.

$$\mathbb{E} [X_a X_b X_c X_d] = \mathbb{E} [X_a X_b] \mathbb{E} [X_c X_d] + \mathbb{E} [X_a X_c] \mathbb{E} [X_b X_d] + \mathbb{E} [X_a X_d] \mathbb{E} [X_b X_c].$$

Stwierdzenie 3.1. *Z warunku na ortogonalność (3.2), wynika, że funkcjonały Wienera są ortogonalne do każdego funkcjonału Volterry mniejszego stopnia, tj.*

$$\langle G_p, H_r \rangle = 0$$

dla $r < p$.

Dowód. Zauważmy, że z warunku $\langle G_p, G_0 \rangle = 0$, $p > 0$ wynika, iż każdy funkcjonał G_p , $p > 0$ jest ortogonalny do dowolnego funkcjonału H_0 . W konsekwencji, dla $p > 1$ zachodzi $\langle G_p, G_1 \rangle = \langle G_p, H_0 \rangle + \langle G_p, H_1 \rangle = \langle G_p, H_1 \rangle = 0$, z czego wynika, iż każdy funkcjonał G_p , $p > 1$ jest ortogonalny do dowolnego funkcjonału H_1 . Podążając tym tokiem rozumowania, można udowodnić, że dla dowolnego $r < p$ zachodzi $\langle G_p, H_r \rangle = 0$. \square

W oparciu o powyższą tożsamość i korzystając z algorytmu Grama-Schmidta, wyprowadzimy teraz wzory na funkcjonały Wienera (do trzeciego rzędu włącznie).

Funkcjonały zerowego i pierwszego rzędu

Podobnie jak w przypadku reprezentacji Volterry, funkcjonał Wienera zerowego rzędu przyjmuje wartość stałą, niezależną od argumentu (sygnału wejściowego),

$$G_0[\mathbf{x}; g_0] = g_0.$$

W celu wyznaczenia jąder funkcjonału pierwszego rzędu, należy ustalić wartość iloczyn skalarnego funkcjonału zerowego rzędu i ogólnej formy funkcjonału pierwszego rzędu (wzór (3.3)),

$$\langle G_0, G_1 \rangle = \mathbb{E} [g_0 g_{0(1)} + g_0 H_1[\mathbf{X}_n; g_1]] = g_0 g_{0(1)}.$$

Tak więc, z warunku na ortogonalność, po przyrównaniu prawej strony do zera otrzymujemy $g_{0(1)} = 0$ i ostatecznie

$$G_1[\mathbf{x}_n; g_1] = \sum_{i=0}^{\infty} g_1(i) x_{n-i}.$$

Funkcjonał drugiego rzędu

W celu ustalenia postaci jąder funkcyjonału drugiego rzędu, należy wyznaczyć iloczyny skalarne $\langle G_0, G_2 \rangle$ i $\langle G_1, G_2 \rangle$, a następnie przyrównać je do zera, by uzyskać wzory na jądra pochodne $g_{0(2)}$ i $g_{1(2)}$.

Iloczyn skalarny funkcyjonałów zerowego i drugiego rzędu dany jest przez

$$\begin{aligned} \langle G_0, G_2 \rangle &= g_0 g_{0(2)} + g_0 \sum_{i=0}^{\infty} g_{1(2)}(i) \mathbb{E}[X_{n-i}] + g_0 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j) \mathbb{E}[X_{n-i} X_{n-j}] \\ &= g_0 (g_{0(2)} + \sigma^2 \sum_{i=0}^{\infty} g_2(i, i)). \end{aligned}$$

Z warunku na ortogonalność otrzymujemy więc $g_{0(2)} = -\sigma^2 \sum_{i=0}^{\infty} g_2(i, i)$. W kolejnym kroku wyznaczamy iloczyn skalarny funkcyjonałów pierwszego i drugiego rzędu, tj.

$$\begin{aligned} \langle G_1, G_2 \rangle &= g_{0(2)} \sum_{i=0}^{\infty} g_1(i) \mathbb{E}[X_{n-i}] + \mathbb{E} \left[\sum_{i=0}^{\infty} g_1(i) X_{n-i} \sum_{i=0}^{\infty} g_{1(2)}(i) X_{n-i} \right] + \\ &+ \mathbb{E} \left[\sum_{i=0}^{\infty} g_1(i) X_{n-i} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j) X_{n-i} X_{n-j} \right] = \sigma^2 \sum_{i=0}^{\infty} g_1(i) g_{1(2)}(i), \end{aligned}$$

stąd $g_{1(2)} = 0$ i ostatecznie, po podstawieniu jąder pochodnych do wzoru (3.3), otrzymujemy

$$G_2[\mathbf{x}_n; g_2(i_1, i_2)] = -\sigma^2 \sum_{i=0}^{\infty} g_2(i, i) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j) x_{n-i} x_{n-j}.$$

Funkcjonał trzeciego rzędu

Postępując jak poprzednio, wyznaczmy kolejne iloczyny skalarne, zaczynając od

$$\begin{aligned} \langle G_0, G_3 \rangle &= g_0 g_{0(3)} + g_0 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_{2(3)}(i, j) \mathbb{E}[x_{n-i} x_{n-j}] \\ &= g_0 (g_{0(3)} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_{2(3)}(i, j) \mathbb{E}[x_{n-i} x_{n-j}]), \end{aligned}$$

skąd otrzymujemy $g_{0(3)} = -\sigma^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_{2(3)}(i, j)$. Jako następny, rozpatrzmy iloczyn

$$\begin{aligned} \langle G_1, G_3 \rangle &= \mathbb{E} \left[\sum_{i=0}^{\infty} g_1(i) x_{n-i} \sum_{i=0}^{\infty} g_{1(3)}(i) x_{n-i} \right] + \\ &+ \mathbb{E} \left[\sum_{i=0}^{\infty} g_1(i) x_{n-i} \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \sum_{i_3=0}^{\infty} g_3(i_1, i_2, i_3) x_{n-i_1} x_{n-i_2} x_{n-i_3} \right]. \end{aligned}$$

Korzystając z symetrii jąder rozwinięcia otrzymujemy więc

$$\begin{aligned}
 \langle G_1, G_2 \rangle &= \sigma^2 \sum_{i=0}^{\infty} g_1(i)g_{1(3)}(i) + \mathbb{E} \left[\sum_{i=0}^{\infty} g_1(i)g_3(i, i, i)x_{n-i}^4 \right] + \\
 &\quad + 3\mathbb{E} \left[\sum_{i=0}^{\infty} \sum_{j=0 \wedge j \neq i}^{\infty} g_1(i)g_3(i, j, j)x_{n-i}^2 x_{n-j}^2 \right] \\
 &= \sigma^2 \sum_{i=0}^{\infty} g_1(i)g_{1(3)}(i) + 3\sigma^4 \sum_{i=0}^{\infty} g_1(i)g_3(i, i, i) + 3\sigma^4 \sum_{i=0}^{\infty} \sum_{j=0 \wedge j \neq i}^{\infty} g_1(i)g_3(i, j, j) \\
 &= \sigma^2 \sum_{i=0}^{\infty} g_1(i)g_{1(3)}(i) + 3\sigma^4 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_1(i)g_3(i, j, j),
 \end{aligned}$$

tym samym, z warunku na ortogonalność wynika, że

$$\sum_{i=0}^{\infty} g_1(i) \left[g_{1(3)}(i) + 3\sigma^2 \sum_{j=0}^{\infty} g_3(i, j, j) \right] = 0,$$

co, dla dowolnych jąder g_1 i g_3 , jest spełnione wtedy i tylko wtedy, gdy $g_{1(3)}(i) = -3\sigma^2 \sum_{j=0}^{\infty} g_3(i, j, j)$. Jako ostatnim, należałoby zająć się iloczynem $\langle G_2, G_3 \rangle$. Zadanie to można jednak nieco uprościć i zamiast tego wyznaczyć iloczyn $\langle H_2, G_3 \rangle$ (patrz stwierdzenie 3.1),

$$\begin{aligned}
 \langle H_2, G_3 \rangle &= \mathbb{E} \left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} h_2(i, j)x_{n-i}x_{n-j} \right] g_{0(3)} + \\
 &\quad + \mathbb{E} \left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} h_2(i, j)x_{n-i}x_{n-j} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_{2(3)}(i, j)x_{n-i}x_{n-j} \right] \\
 &= \sigma^2 g_{0(3)} \sum_{i=0}^{\infty} g_2(i, i) + 3\sigma^4 \sum_{i=0}^{\infty} g_2(i, i)g_{2(3)}(i, i) \\
 &\quad + 2\sigma^4 \sum_{i=0}^{\infty} \sum_{j \neq i}^{\infty} g_2(i, j)g_{2(3)}(i, j) + \sigma^4 \sum_{i=0}^{\infty} \sum_{j \neq i}^{\infty} g_2(i, i)g_{2(3)}(j, j) \\
 &= \sigma^2 g_{0(3)} \sum_{i=0}^{\infty} g_2(i, i) + \\
 &\quad + 2\sigma^4 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j)g_{2(3)}(i, j) + \sigma^4 \sum_{i=0}^{\infty} g_2(i, i) \sum_{j=0}^{\infty} g_{2(3)}(j, j).
 \end{aligned}$$

Wstawiając do powyższego wyrażenia wyprowadzone we wcześniejszych krokach wyrażenia na $g_{0(3)}$ i $g_{1(3)}$ otrzymujemy

$$\langle H_2, G_3 \rangle = 2\sigma^4 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_2(i, j)g_{2(3)}(i, j).$$

Z warunku na ortogonalność wynika, że powyższe wyrażenie powinno wynosić zero dla dowolnego jądra g_2 . W szczególności, dla $g_2(i, j) = g_{2(3)}(i, j)$, powyższy iloczyn

przyjmuje postać $\langle H_2, G_3 \rangle = 2\sigma^4 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_{2(3)}^2(i, j) \geq 0$, co prowadzi do wniosku, że $g_{2(3)}(i, j) = 0$. Ostatecznie, funkcjonał trzeciego rzędu przyjmuje postać

$$G_3[\mathbf{x}_n; g_3] = -3\sigma^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} g_3(i, j, j)x_{n-i} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} g_3(i, j, k)x_{n-i}x_{n-j}x_{n-k}.$$

3.B.2 Inne własności

Stwierdzenie 3.2. *Jądra pochodne rozwinięcia Wienera dane są wzorami*

$$g_{p-2r(p)}(i_1, \dots, i_{p-2r}) = \frac{(-1)^r p! (\sigma^2)^r}{(p-2r)! r! 2^r} \sum_{j_1=0}^{\infty} \dots \sum_{j_r=0}^{\infty} g_p(j_1, j_1, \dots, j_r, j_r, i_1, \dots, i_{p-2r}), \quad (3.17)$$

oraz

$$g_{p-2r-1(p)}(i_1, \dots, i_{p-2r-1}) = 0$$

dla $r = 1, 2, \dots, \lfloor \frac{p}{2} \rfloor$.

Dowód. Ze względu na obszerność dowodu, odsyłamy do [92], gdzie przedstawiono wyprowadzenia jąder pochodnych dla funkcjonałów z czasem ciągłym. \square

Stwierdzenie 3.3. *Funkcjonały Wienera są liniowe w odniesieniu do jąder, a więc spełniają tożsamość*

$$G_p[\mathbf{x}_n; \alpha a_p + \beta b_p] = \alpha G_p[\mathbf{x}_n; a_p] + \beta G_p[\mathbf{x}_n; b_p],$$

dla dowolnych $\alpha, \beta \in \mathbb{R}$ oraz $a_p, b_p \in \mathcal{K}_p$.

Dowód. Zauważymy, że ze wzoru (3.17) wynika, iż jądra pochodne, $g_{p-2r(p)}$, są liniowym przekształceniem jądra g_p . Tym samym, zachodzi

$$\begin{aligned} G_p[\mathbf{x}_n; \alpha a_p + \beta b_p] &= \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} H_{p-2r}[\mathbf{x}_n; \alpha a_{p-2r(p)} + \beta b_{p-2r(p)}] \\ &= \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} \alpha H_{p-2r}[\mathbf{x}_n; a_{p-2r(p)}] + \beta H_{p-2r}[\mathbf{x}_n; b_{p-2r(p)}] \\ &= \alpha G_p[\mathbf{x}_n; a_p] + \beta G_p[\mathbf{x}_n; b_p], \end{aligned}$$

co kończy dowód. \square

Stwierdzenie 3.4. Niech $I_{(i_1, \dots, i_p)} \in \mathcal{K}_p$ oznacza jądro, którego element o indeksie (i_1, \dots, i_p) ma wartość równą 1, a wszystkie pozostałe są równe 0. Dowolny funkcjonal Wienera można przedstawić jako liniową kombinację iloczynu prostszych funkcjonałów Wienera, w następujący sposób

$$G_p[\mathbf{x}_n; g_p] = \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) G_{k_1} \left[\mathbf{x}_n; \underbrace{I_{(i_1, \dots, i_1)}}_{k_1 \text{ razy}} \right] \cdots G_{k_p} \left[\mathbf{x}_n; \underbrace{I_{(i_p, \dots, i_p)}}_{k_p \text{ razy}} \right],$$

gdzie k_j oznacza liczbę wystąpień liczby i_j w indeksie (i_1, \dots, i_p) , tak że $\sum_{i=1}^p k_i = p$.

Dowód. Zauważmy, że dowolne jądro g_p można przedstawić w postaci sumy

$$g_p = \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) I_{(i_1, \dots, i_p)}.$$

Korzystając z faktu, że funkcjonały Wienera są liniowe względem jąder, otrzymujemy tożsamość

$$\begin{aligned} G_p[\mathbf{x}_n; g_p] &= G_p \left[\mathbf{x}_n; \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) I_{(i_1, \dots, i_p)} \right] \\ &= \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) G_p[\mathbf{x}_n; I_{(i_1, \dots, i_p)}]. \end{aligned}$$

Wykażemy teraz, że funkcjonal $G_p[\mathbf{x}_n; I_{(i_1, \dots, i_p)}]$ można zdekomponować i przedstawić w formie iloczynów prostszych funkcjonałów, tj.

$$G_p[\mathbf{x}_n; I_{(i_1, \dots, i_p)}] = G_{k_1} \left[\mathbf{x}_n; \underbrace{I_{(i_1, \dots, i_1)}}_{k_1 \text{ razy}} \right] \cdots G_{k_p} \left[\mathbf{x}_n; \underbrace{I_{(i_p, \dots, i_p)}}_{k_p \text{ razy}} \right].$$

W tym celu rozważmy przypadek prostszy i udowodnijmy, że zachodzi tożsamość

$$G_p \left[\mathbf{x}_n; \underbrace{I_{(a, \dots, a, b, \dots, b)}}_{\substack{r \text{ razy} \\ s \text{ razy}}} \right] = G_r \left[\mathbf{x}_n; \underbrace{I_{(a, \dots, a)}}_{r \text{ razy}} \right] G_s \left[\mathbf{x}_n; \underbrace{I_{(b, \dots, b)}}_{s \text{ razy}} \right], \quad (3.18)$$

dla $p = r + s$. Zaczniemy od pokazania, że komponent wiodący, czyli wielomian stopnia p , po obu stronach równania jest identyczny. Dla lewej strony wynosi on

$$H_p \left[\mathbf{x}_n; \underbrace{I_{(a, \dots, a, b, \dots, b)}}_{\substack{r \text{ razy} \\ s \text{ razy}}} \right] = x_{n-a}^r x_{n-b}^s,$$

natomiast w przypadku prawej strony równania otrzymujemy

$$H_r \left[\mathbf{x}_n; \underbrace{I(a, \dots, a)}_{r \text{ razy}} \right] H_s \left[\mathbf{x}_n; \underbrace{I(b, \dots, b)}_{s \text{ razy}} \right] = x_{n-a}^r x_{n-b}^s.$$

Teraz, zamiast pokazywać, że jądra pochodne wyrażenia $G_r [\mathbf{x}_n; I_{(a, \dots, a)}] G_s [\mathbf{x}_n; I_{(b, \dots, b)}]$ są takie same jak jądra pochodne wyrażenia po lewej stronie równania (3.18), pokażemy, że jest ono ortogonalne do wszystkich funkcjonałów Volterry stopnia $t < r + s$, tj.

$$\begin{aligned} H_t [\mathbf{x}_n; I_{(i_1, \dots, i_t)}] &= H_u \left[\mathbf{x}_n; \underbrace{I(a, \dots, a)}_{u \text{ razy}} \right] H_w \left[\mathbf{x}_n; \underbrace{I(b, \dots, b)}_{w \text{ razy}} \right] H_{t-(u+w)} [\mathbf{x}_n; I_{(c_1, \dots, c_{t-(u+w)})}] \\ &= x_{n-a}^u x_{n-b}^w x_{n-c_1} \dots x_{n-c_{t-(u+w)}}, \end{aligned}$$

dla $c_i \neq a$ i $c_i \neq b$, $i = 1, \dots, t - (u + w)$. Z faktu, że zmienne losowe x_n są niezależne, wynika

$$\begin{aligned} \left\langle G_r [\mathbf{x}_n; I_{(a, \dots, a)}] G_s [\mathbf{x}_n; I_{(j, \dots, j)}] H_t [\mathbf{x}_n; I_{(i_1, \dots, i_t)}] \right\rangle &= \left\langle G_r [\mathbf{x}_n; I_{(a, \dots, a)}] H_u [\mathbf{x}_n; I_{(a, \dots, a)}] \right\rangle \\ &\quad \times \left\langle G_s [\mathbf{x}_n; I_{(b, \dots, b)}] H_w [\mathbf{x}_n; I_{(b, \dots, b)}] \right\rangle \\ &\quad \times \left\langle H_{t-(u+w)} [\mathbf{x}_n; I_{(c_1, \dots, c_{t-(u+w)})}] \right\rangle. \end{aligned}$$

Zauważmy, że z faktu iż $t < r + s$ wynika, że albo $u < r$, albo $s < w$ (lub oba te warunki są spełnione jednocześnie), a więc z warunku na ortogonalność, wynika że co najmniej jeden z dwóch pierwszych czynników prawej strony powyższego równania jest równy zero. \square

Stwierdzenie 3.5. *Funkcjonały Wienera można przedstawić jako kombinację liniową wielomianów Hermite'a, w następujący sposób*

$$G_p [\mathbf{x}_n; g_p] = \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} g_p(i_1, \dots, i_p) \mathcal{H}_{k_1}(x_{n-i_1}) \dots \mathcal{H}_{k_p}(x_{n-i_p}),$$

gdzie k_j oznacza liczbę wystąpień liczby i_j w indeksie (i_1, \dots, i_p) , tak że $\sum_{i=1}^p k_i = p$.

Dowód. Zauważmy, że jądra pochodne wyrażenia

$$G_p \left[\mathbf{x}_n; \underbrace{I(a, \dots, a)}_{r \text{ razy}} \right] = \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} H_{p-2r} [\mathbf{x}_n; I_{(a, \dots, a), p-2r(p)}]$$

dane są przez wzór

$$I_{(a, \dots, a), p-2r(p)}(i_1, \dots, i_{p-2r}) = \frac{(-1)^r p! (\sigma^2)^r}{(p-2r)! r! 2^r} \sum_{j_1=0}^{\infty} \dots \sum_{j_r=0}^{\infty} I_p(j_1, j_1, \dots, j_r, j_r, i_1, \dots, i_{p-2r})$$

$$= \begin{cases} \frac{(-1)^r p! (\sigma^2)^r}{(p-2r)! r! 2^r}, & \text{dla } i_1 = \dots = i_{p-2r} = a, \\ 0, & \text{w przeciwnym wypadku,} \end{cases}$$

a więc

$$G_p \left[\mathbf{x}_n; \underbrace{I(a, \dots, a)}_{r \text{ razy}} \right] = p! \sum_{r=0}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^r (\sigma^2)^r}{(p-2r)! r! 2^r} x_{n-a}^p.$$

Przyrównując powyższe wyrażenie do wzoru (3.16) otrzymujemy $G_p [\mathbf{x}_n; I_{(i, \dots, i)}] = \mathcal{H}_p(x_{n-i})$, a więc korzystając z własności przedstawionej w stwierdzeniu 3.4 kończymy dowód. \square

Rozdział 4

Modelowanie systemów z wykorzystaniem metod lustrzanych z uśrednianiem

Rozdział dotyczy zastosowania metod lustrzanych z uśrednianiem (ang. *dual averaging methods*, DA) w zadaniu modelowania dynamicznych systemów nieliniowych. W podrozdziale 4.1 przedstawiono szczegółowy opis działania algorytmu, a także przytoczono wyniki dotyczące jego własności statystycznych, w przypadku gdy badany system jest statyczny. W podrozdziale 4.2 zaprezentowane oryginalne wyniki dotyczące działania metody dla systemów z pamięcią skończoną oraz pamięcią zanikającą. Oprócz rezultatów o charakterze ogólnym, dotyczących modeli opisanych wzorem $m_S[\cdot; \boldsymbol{\theta}] = \sum_{i=1}^D \theta_i \varphi_i[\cdot]$, opartych o dowolny ustalony słownik $S = \{\varphi_1, \dots, \varphi_D\}$, przedstawiono także dokładny sposób wykorzystania entropijnego wariantu metody dla słowników wielomianowych (a więc modeli opartych o reprezentację volterrowską lub ortogonalną). W kolejnej części rozdziału, kontynuując rozważania przeprowadzone we wcześniejszej części pracy, omówiono warunki zbieżności metody dla systemów Wienera-Hammersteina. W podrozdziale 4.4 przedstawiono rezultaty eksperymentu numerycznego, potwierdzającego teoretyczne własności metody.

4.1 Algorytm lustrzany z uśrednianiem

Uproszczony opis algorytmu przedstawiony został w paragrafie 1.5.5. Zatem, bez zbędnych powtórzeń, przejdziemy teraz do szczegółowego opisu działania metody.

Algorytm definiuje się rekurencyjnie, za pomocą następujących wzorów (por. np. [118])

$$\boldsymbol{\theta}_1 = \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} h(\boldsymbol{\theta}), \quad (4.1)$$

$$\boldsymbol{\theta}_{t+1} = \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \langle \mathbf{g}_{1:t}, \boldsymbol{\theta} \rangle + \eta_t^{-1} h(\boldsymbol{\theta}) \right\}, \quad (4.2)$$

gdzie $\mathbf{g}_{1:t} = \sum_{i=1}^t \mathbf{g}_i$ stanowi sumę pierwszych t gradientów, $\{\eta_t\}$ jest nierosnącym ciągiem dodatnich długości kroku algorytmu, $\mathcal{C} \subset \mathbb{R}^D$ jest domkniętym zbiorem wypukłym, wyrażającym ograniczenia nałóżone na parametry modelu, a $h : \mathbb{R}^D \rightarrow \mathbb{R}$ jest funkcją regularyzującą¹, która jest ciągła, ograniczona i silnie wypukła na zbiorze \mathcal{C} (zob. def. 4.3, dodatek 4.E) względem pewnej ustalonej normy $\|\cdot\|$. Bez utraty ogólności założymy, że $\min_{\mathbf{u} \in \mathcal{C}} h(\mathbf{u}) = 0$ oraz że $h(\cdot)$ jest silnie wypukła ze stałą $c = 1$ (za pomocą operacji skalowania i przesunięcia, warunki te mogą być osiągnięte dla dowolnej silnie wypukłej funkcji).

Dodatkowo, celem ograniczenia wpływu szumu losowego, stochastyczna wersja algorytmu zwraca estymatę będącą uśrednieniem *rozwiązań pośrednich* $\boldsymbol{\theta}_t$, tzn. za końcowy wynik algorytmu przyjmuje się wektor

$$\bar{\boldsymbol{\theta}}_T = T^{-1} \sum_{t=1}^T \boldsymbol{\theta}_t. \quad (4.3)$$

Uwaga 4.1. *Niech*

$$f_t^*(\mathbf{w}) = \sup_{\mathbf{x}} (\langle \mathbf{w}, \mathbf{x} \rangle - f_t(\mathbf{x})) \quad (4.4)$$

oznacza funkcję będącą wynikiem transformacji Legendre'a-Fenchela funkcji

$$f_t(\mathbf{u}) = \eta_t^{-1} h(\mathbf{u}) + I_{\mathcal{C}}(\mathbf{u}), \quad (4.5)$$

gdzie

$$I_{\mathcal{C}}(\mathbf{u}) = \begin{cases} 0, & \text{dla } \mathbf{u} \in \mathcal{C}, \\ \infty, & \text{w przeciwnym wypadku} \end{cases}$$

jest funkcją charakterystyczną zbioru \mathcal{C} . Z [88, stwierdzenie 11.3] wynika, że gradient funkcji f_t^* równy jest

$$\begin{aligned} \nabla f_t^*(\mathbf{w}) &= \arg \max_{\mathbf{u}} \{ \langle \mathbf{w}, \mathbf{u} \rangle - \eta_t^{-1} h(\mathbf{u}) - I_{\mathcal{C}}(\mathbf{u}) \} = \arg \max_{\mathbf{u} \in \mathcal{C}} \{ \langle \mathbf{w}, \mathbf{u} \rangle - \eta_t^{-1} h(\mathbf{u}) \} \\ &= \arg \min_{\mathbf{u} \in \mathcal{C}} \{ \langle -\mathbf{w}, \mathbf{u} \rangle + \eta_t^{-1} h(\mathbf{u}) \}. \end{aligned}$$

Zestawienie powyższego wzoru ze wzorem (4.2) prowadzi do wniosku, że rozwiązania pośrednie algorytmu DA można wyrazić za pomocą wyrażenia

$$\boldsymbol{\theta}_{t+1} = \nabla f_t^*(-\mathbf{g}_{1:t}). \quad (4.6)$$

Tak więc, działanie algorytmu można rozumieć jako następujący po sobie ciąg od-wzorowań z przestrzeni dualnej, której elementami są wektory stanowiące gradienty funkcji straty, do przestrzeni prymalnej wektorów $\boldsymbol{\theta}_t$.

¹W literaturze anglojęzycznej, możemy spotkać się z nazwą *a distance generating function* oraz *a prox function*.

Omówienie najważniejszych własności algorytmu poprzedzimy wprowadzeniem stałej

$$R = \sqrt{\sup_{\mathbf{u} \in \mathcal{C}} h(\mathbf{u})},$$

oraz definicją σ -ciała

$$\mathcal{F}_t = \sigma(\{\mathbf{X}_i, Y_i\}_{i=1}^t),$$

generowanego przez t pierwszych par pomiarowych. Oprócz tego, przez $\|\cdot\|_*$ będziemy oznaczali *normę dualną* (względem normy $\|\cdot\|$), która jest definiowana wzorem

$$\|\mathbf{x}\|_* = \sup_{\mathbf{y} \in \mathbb{R}^D: \|\mathbf{y}\| \leq 1} \langle \mathbf{x}, \mathbf{y} \rangle.$$

Pojęcie to będzie pełniło kluczową rolę przy omawianiu rezultatów dotyczących algorytmu DA.

Zacniemy od przedstawienia twierdzenia, które podaje górne oszacowanie sumy pierwszych T empirycznych strat modeli pośrednich algorytmu (tj. opartych o rozwiązania pośrednie, $\boldsymbol{\theta}_t$) względem ustalonego modelu opartego o dowolny wektor parametrów $\mathbf{u} \in \mathcal{C}$. Przypominamy, że wzór ten przedstawia się następującym wzorem

$$C_T(\mathbf{u}) = \sum_{t=1}^T (\hat{L}_t(\boldsymbol{\theta}_t) - \hat{L}_t(\mathbf{u})). \quad (4.7)$$

Twierdzenie 4.1 ([78]). *Niech $\{\boldsymbol{\theta}_t\}_{t=1}^T$ będzie ciągiem rozwiązań pośrednich algorytmu DA, otrzymanym za pomocą wzorów (4.1)–(4.2). Suma empirycznych strat (wzór (4.7)) względem dowolnego wektora $\mathbf{u} \in \mathcal{C}$ spełnia nierówność*

$$C_T(\mathbf{u}) \leq \frac{1}{2} \sum_{t=1}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 + \eta_T^{-1} R^2, \quad (4.8)$$

gdzie $\eta_0 = \eta_1$.

Dowód. Zob. dodatek 4.A. □

Zaprezentowane twierdzenie dotyczy zaobserwowanego błędu (straty) algorytmu w przypadku gdy dane pomiarowe pochodzą ze zbioru uczącego, nie mówi więc o spodziewanej wartości błędu dla jeszcze niezaobserwowanych pomiarów. Mimo to, jak pokażemy dalej, twierdzenie to jest punktem wyjścia do uzyskania dalszych wyników, dotyczących błędów estymacji algorytmu (wzór (1.12)). Przedstawimy teraz jeden z nich, dotyczący systemów statycznych.

Założenie 4.1. *Długość pamięci systemu wynosi 1.*

Założenie 4.2. *Długość pamięci poszczególnych elementów słownika wynosi 1.*

Lemat 4.1. Niech $\bar{\boldsymbol{\theta}}_T$ będzie wynikiem zastosowania algorytmu DA (wzór (4.3)). Dla systemu statycznego, w którym spełnione są założenia 1.1–1.2 i 4.1, oraz słownika S spełniającego założenie 4.2, błąd estymacji spełnia nierówność

$$\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) \leq T^{-1} \mathbb{E} \left\{ \frac{1}{2} \sum_{t=1}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 + \eta_T^{-1} R^2 \right\}. \quad (4.9)$$

Dowód. Po podzieleniu obu stron nierówności (4.8) przez T , nałożeniu wartości oczekiwanej i dokonaniu podstawienia $\mathbf{u} = \boldsymbol{\theta}^*$ otrzymujemy

$$T^{-1} \sum_{t=1}^T \left(\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}_t)\} - \mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}^*)\} \right) \leq T^{-1} \mathbb{E} \left\{ \frac{1}{2} \sum_{t=1}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 + \eta_T^{-1} R^2 \right\}. \quad (4.10)$$

Zajmijmy się lewą stroną wzoru. Dla ustalonego $\boldsymbol{\theta}^*$ zachodzi $\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}^*)\} = L_S(\boldsymbol{\theta}^*)$. Zauważmy teraz, że z faktu iż system jest statyczny, wynika że wyjście systemu w chwili t , tj. Y_t , jest zmienną losową niezależną od \mathcal{F}_{t-1} (i w konsekwencji także od wektora $\boldsymbol{\theta}_t$). Podobnie, ze względu na założenie o słowniku, jego poszczególne elementy, $\varphi_i(\mathbf{X}_t)$, są także niezależne od \mathcal{F}_{t-1} . W następstwie, zachodzi równość

$$\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\} = L_S(\boldsymbol{\theta}_t)$$

i w rezultacie otrzymujemy

$$\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}_t)\} = \mathbb{E}\{\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\}\} = \mathbb{E}\{L_S(\boldsymbol{\theta}_t)\}.$$

Tym samym, nierówność (4.10) przyjmuje postać

$$T^{-1} \sum_{t=1}^T \mathbb{E}\{L_S(\boldsymbol{\theta}_t)\} - L_S(\boldsymbol{\theta}^*) \leq T^{-1} \mathbb{E} \left\{ \frac{1}{2} \sum_{t=1}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 + \eta_T^{-1} R^2 \right\}. \quad (4.11)$$

Na koniec zauważmy, że z faktu iż $L_S(\cdot)$ jest funkcją wypukłą, wynika następująca nierówność

$$\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} \leq T^{-1} \sum_{t=1}^T \mathbb{E}\{L_S(\boldsymbol{\theta}_t)\}.$$

Wstawienie powyższego oszacowania do wzoru (4.11) kończy dowód. \square

Przedstawiony rezultat pokazuje, że błąd estymacji algorytmu posiada dwa stopnie swobody i można nim manipulować poprzez dobór długości kroku oraz funkcji regularyzującej, która określa zarówno postać normy dualnej $\|\cdot\|_*$, jak i „pojemność” R^2 zbioru \mathcal{C} . Pomimo iż lemat dotyczy tylko i wyłącznie systemów statycznych, to ze względu na przejrzystość otrzymanego oszacowania, stanowi on doskonały wstęp do omówienia możliwych strategii doboru kroku i funkcji $h(\cdot)$.

4.1.1 Dobór długości kroku

Przedstawimy teraz dwa rezultaty, dotyczące kolejno deterministycznego i losowego sposobu określania sekwencji kroków $\{\eta_t\}$. Zaczniemy od wprowadzenia dwóch założeń o gradientach empirycznej funkcji straty.

Założenie 4.3. *Istnieje stała $G < \infty$, taka że dla każdej chwili czasu t i dla każdego $\theta \in \mathcal{C}$, zachodzi nierówność*

$$\mathbb{E} \left\| \nabla \hat{L}_t(\theta) \right\|_*^2 \leq G^2.$$

Założenie 4.4. *Istnieje stała $\bar{G} < \infty$, taka że dla każdej chwili czasu t i dla każdego $\theta \in \mathcal{C}$, zachodzi nierówność*

$$\left\| \nabla \hat{L}_t(\theta) \right\|_* \stackrel{p.n.}{\leq} \bar{G}.$$

Założenie 4.3 pozwoli nam ograniczyć wyrażenia $\mathbb{E} \left\| \mathbf{g}_t \right\|_*^2$. Z kolei silniejsze założenie 4.4 wymagane jest gdy długość kroku jest ustalana w sposób losowy.

Wniosek 4.1. *Jeśli spełnione jest założenie 4.3, to dla ciągu kroków, którego elementy zdefiniowane są przez*

$$\eta_t = \frac{R}{G\sqrt{t}}, \quad t \geq 1,$$

nierówność (4.9) z lematu 4.1 przyjmuje postać

$$\begin{aligned} \mathbb{E}\{L_S(\bar{\theta}_T)\} - L_S(\theta^*) &\leq T^{-1} \left(\frac{1}{2} \sum_{t=1}^T \eta_{t-1} G^2 + \eta_T^{-1} R^2 \right) \\ &\leq T^{-1} \left(RG\sqrt{T-1} + RG\sqrt{T} \right) < \frac{2RG}{\sqrt{T}}. \end{aligned} \quad (4.12)$$

Dodatkowo, jeżeli liczba kroków jest z góry znana (wariant offline algorytmu), to przyjmując stały krok $\eta_t = \frac{\sqrt{2R}}{G\sqrt{T}}$, otrzymujemy

$$\mathbb{E}\{L_S(\bar{\theta}_T)\} - L_S(\theta^*) \leq \frac{\sqrt{2}RG}{\sqrt{T}},$$

a więc polepszamy oszacowanie błędu o stałą.

Dowód. We wzorze (4.12) skorzystaliśmy z nierówności $\sum_{t=1}^{T-1} \frac{1}{\sqrt{t}} \leq 1 + \int_1^{T-1} \frac{1}{\sqrt{t}} dt \leq 2\sqrt{T-1} - 1$. \square

Wniosek 4.2. *Jeśli spełnione są założenia 4.3–4.4, to dla ciągu kroków, którego elementy zdefiniowane są przez*

$$\eta_t = \frac{R}{\sqrt{G^2 + \sum_{i=1}^t \left\| \mathbf{g}_i \right\|_*^2}}, \quad t \geq 1,$$

nierówność (4.9) z lematu 4.1 przyjmuje postać

$$\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) < \frac{2RG}{\sqrt{T}} + \frac{R\bar{G}}{T}.$$

Dowód. Zauważmy, że mimo losowej natury, z definicji poszczególnych kroków η_t wynika, że ciąg $\{\eta_t\}$ jest zawsze nierosnący, a więc nie stoi w sprzeczności z założeniami algorytmu.

Po dokonaniu podstawienia η_t w miejsce (4.9), otrzymujemy

$$\begin{aligned} Q(\bar{\boldsymbol{\theta}}_T) &\leq \frac{R}{T} \mathbb{E} \left\{ \frac{1}{2} \sum_{t=1}^T \frac{\|\mathbf{g}_t\|_*^2}{\sqrt{\bar{G}^2 + \sum_{i=1}^{t-1} \|\mathbf{g}_i\|_*^2}} + \sqrt{\bar{G}^2 + \sum_{i=1}^T \|\mathbf{g}_i\|_*^2} \right\} \\ &\stackrel{(i)}{\leq} \frac{R}{T} \mathbb{E} \left\{ \frac{1}{2} \sum_{t=1}^T \frac{\|\mathbf{g}_t\|_*^2}{\sqrt{\sum_{i=1}^t \|\mathbf{g}_i\|_*^2}} + \sqrt{\bar{G}^2 + TG^2} \right\} \\ &\stackrel{(ii)}{\leq} \frac{R}{T} \mathbb{E} \left\{ \sqrt{\sum_{t=1}^T \|\mathbf{g}_t\|_*^2} + \sqrt{\bar{G}^2 + TG^2} \right\} \\ &\leq \frac{R}{T} \left(G\sqrt{T} + \sqrt{\bar{G}^2 + TG^2} \right) < \frac{R}{T} \left(G\sqrt{T} + \sqrt{(\bar{G} + \sqrt{T}G)^2} \right) \\ &= \frac{2RG}{\sqrt{T}} + \frac{R\bar{G}}{T}, \end{aligned}$$

gdzie nierówność (i) zachodzi na podstawie założenia 4.4, a (ii) wynika z [72, lem. 4]. \square

Zwróćmy uwagę, że dla kroku η_t zaproponowanego we wniosku 4.2 i dla odpowiednio dużego T , błąd metody jest ograniczony od góry przez $\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) < \frac{2RG}{\sqrt{T}} + \epsilon$, dla dowolnie małego $\epsilon > 0$, a więc asymptotycznie ograniczenie to równe jest ograniczeniu (4.12) z wniosku 4.1. Jednakże, jeśli w danym zastosowaniu $\|\mathbf{g}_t\|_*^2 \ll G^2$ zachodzi z *dużym* prawdopodobieństwem, w praktyce możemy liczyć na lepsze rezultaty metody. Wariant oparty o krok losowy będziemy nazywali *adaptacyjnym*, gdyż w każdej iteracji długość kroku zmienia się w sposób zależny od zaobserwowanego gradientu empirycznej funkcji straty. Konsekwentnie, wariant przedstawiony we wniosku 4.1 będziemy nazywali *nieadaptacyjnym*.

Uwaga 4.2. We wniosku 4.2 poczynione zostało założenie o ograniczonej normie dualnej gradientu empirycznej funkcji straty, co z teoretycznego punktu widzenia istotnie zawęża zakres stosowalności algorytmu – wymaga ono bowiem, aby rozkład zakłócenia wyjścia systemu miał ograniczony nośnik, co stoi w sprzeczności z często spotykanym założeniem o rozkładzie gaussowskim szumu addytywnego. Założenie to możemy jednak nieco złagodzić. Jak łatwo bowiem sprawdzić, jeśli stałą \bar{G} zdefiniować jako górne ograniczenie stosunku dwóch pierwszych momentów zmiennej losowej

$\|\mathbf{g}_t\|_*$, tj. w taki sposób, żeby dla każdego $t \geq 1$ zachodziła nierówność $\frac{\mathbb{E}\|\mathbf{g}_t\|_*^2}{\mathbb{E}\|\mathbf{g}_t\|_*} \leq \bar{G}$,
nierówność

$$\mathbb{E}\left\{\frac{\|\mathbf{g}_t\|_*^2}{\sqrt{\bar{G}^2 + \sum_{i=1}^{t-1} \|\mathbf{g}_i\|_*^2}}\right\} \leq \mathbb{E}\left\{\frac{\|\mathbf{g}_t\|_*^2}{\sqrt{\sum_{i=1}^t \|\mathbf{g}_i\|_*^2}}\right\},$$

jest nadal spełniona, a przy tym wspomniane ograniczenie zostaje zniesione.

4.1.2 Określenie funkcji regularyzującej

Założmy, że funkcja $h(\cdot)$ jest silnie wypukła względem normy w ℓ_p , tj. $\|\cdot\| = \|\cdot\|_p$, $p \geq 1$. Wtedy normą dualną do $\|\cdot\|$ jest $\|\cdot\|_* = \|\cdot\|_q$, gdzie $q = (1 - p^{-1})^{-1}$ [12, str. 637].

W pierwszej kolejności rozpatrzmy podstawowy, *euklidesowy* wariant algorytmu, w którym funkcja regularyzująca oparta jest o normę w ℓ_2 , tj.

$$h(\boldsymbol{\theta}) = \frac{1}{2} \|\boldsymbol{\theta}\|_2^2.$$

Jak nietrudno sprawdzić, funkcja ta jest silnie wypukła względem normy $\|\cdot\| = \|\cdot\|_2$, zatem norma dualna dana jest przez $\|\cdot\|_* = \|\cdot\|_2$. Pośredni problem optymalizacyjny (4.2) można przekształcić do bardziej przystępnej postaci,

$$\begin{aligned} \boldsymbol{\theta}_{t+1} &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \langle \mathbf{g}_{1:t}, \boldsymbol{\theta} \rangle + \eta_t^{-1} \frac{1}{2} \|\boldsymbol{\theta}\|_2^2 \right\} \\ &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \|\eta_t \mathbf{g}_{1:t}\|_2^2 + 2 \langle \eta_t \mathbf{g}_{1:t}, \boldsymbol{\theta} \rangle + \|\boldsymbol{\theta}\|_2^2 \right\} \\ &= \arg \min_{\boldsymbol{\theta} \in \mathcal{C}} \left\{ \|\eta_t \mathbf{g}_{1:t} + \boldsymbol{\theta}\|_2^2 \right\} = \pi_{\mathcal{C}}(-\eta_t \mathbf{g}_{1:t}), \end{aligned} \quad (4.13)$$

gdzie $\pi_{\mathcal{C}}(\boldsymbol{\theta}) = \arg \min_{\mathbf{u} \in \mathcal{C}} \|\boldsymbol{\theta} - \mathbf{u}\|_2$ jest rzutem wektora $\boldsymbol{\theta}$ na zbiór \mathcal{C} . W tym wariancie, naturalnym zbiorem ograniczeń jest kula w normie ℓ_2 o środku w początku układu współrzędnych, tj. $\mathcal{C} = B_r = \{\boldsymbol{\theta} : \|\boldsymbol{\theta}\|_2 \leq r\}$ – wybór ten sprawia że, wzór (4.13) przyjmuje postać jawną, bowiem $\pi_{B_r}(\mathbf{x}) = \min\{1, r \|\mathbf{x}\|_2^{-1}\} \mathbf{x}$. Dla tak zdefiniowanej funkcji $h(\cdot)$ i zbioru \mathcal{C} , otrzymujemy $R = \sqrt{\sup_{\mathbf{u} \in \mathcal{C}} \frac{1}{2} \|\boldsymbol{\theta}\|_2^2} = \frac{\sqrt{2}}{2} r$. Zakładając, że norma ℓ_∞ gradientów \mathbf{g}_t jest ograniczona dochodzimy do wniosku, że błąd algorytmu jest rzędu $\mathcal{O}(r\sqrt{\frac{D}{T}})$ (por. ustęp 1.5.2).

Rozważymy teraz wariant *entropijny* metody, nazywany tak ze względu na funkcję regularyzującą, która oparta jest o funkcję ujemnej entropii,

$$h(\boldsymbol{\theta}) = \ln D + \sum_{i=1}^D \theta_i \ln \theta_i. \quad (4.14)$$

Obecność elementu $\ln D$, chociaż neutralna z punktu widzenia pośredniego problemu optymalizacyjnego, jest konieczna ze względu na założenia o wartości minimalnej funkcji $h(\cdot)$. Funkcja ta jest silnie wypukła względem normy $\|\cdot\| = \|\cdot\|_1$

na *sympleksie probabilistycznym* (zob. np. [78]), tj. zbiorze $\Delta_D = \{\boldsymbol{\theta} \in \mathbb{R}^D : \sum_{i=1}^D \theta_i = 1, \theta_i \geq 0\}$, który jednocześnie przyjmujemy za zbiór rozwiązań algorytmu, tj. $\mathcal{C} = \Delta_D$. Łatwo sprawdzić, że punkt początkowy algorytmu dany jest przez $\boldsymbol{\theta}_1 = \arg \min_{\mathbf{u} \in \Delta_D} \sum_{i=1}^D u_i \ln u_i = D^{-1} [1, \dots, 1]$. W celu wyznaczenia rozwiązań pośrednich, wykorzystamy warunki Karusha-Kuhna-Tuckera (w skrócie KKT, zob. np. [12, str. 244]). Zaczniemy od wyznaczenia gradientu funkcji celu (4.2). Po przemnożeniu jej przez η_t , i -ta pochodna cząstkowa dana jest przez

$$\frac{\partial \left\{ \eta_t \langle \mathbf{g}_{1:t}, \boldsymbol{\theta} \rangle + \sum_{k=1}^D \theta_k \ln \theta_k \right\}}{\partial \theta_i} = \eta_t g_{1:t,i} + \ln \theta_i + 1,$$

gdzie $g_{1:t,i}$ oznacza i -ty element wektora $\mathbf{g}_{1:t}$. Z warunków KKT otrzymujemy układ równań

$$\begin{aligned} \eta_t g_{1:t,i} + \ln \theta_{t+1,i} + 1 + \nu &= 0, \\ \sum_{i=1}^D \theta_{t+1,i} &= 1, \end{aligned}$$

gdzie ν jest mnożnikiem Lagrange'a wynikającym z ograniczenia problemu optymalizacyjnego. Rozwiązując pierwsze równanie ze względu na $\theta_{t+1,i}$, otrzymujemy

$$\theta_{t+1,i} = \exp(-\eta_t g_{1:t,i} - 1 - \nu), \quad (4.15)$$

co po podstawieniu do równania drugiego prowadzi do

$$\sum_{i=1}^D \exp(-\eta_t g_{1:t,i} - 1 - \nu) = 1$$

i w konsekwencji

$$\exp(1 + \nu) = \sum_{i=1}^D \exp(-\eta_t g_{1:t,i}). \quad (4.16)$$

Ostatecznie, łącząc równanie (4.15) z (4.16), otrzymujemy wzór wyznaczający $t + 1$ rozwiązanie algorytmu,

$$\theta_{t+1,i} = \frac{\exp(-\eta_t g_{1:t,i})}{\sum_{j=1}^D \exp(-\eta_t g_{1:t,j})}.$$

Ze względu na formę powyższego równania, omawiany wariant metody określa się czasem *eksponencjalnym*.

Nietrudno sprawdzić, że dla funkcji ujemnej entropii, stała R przyjmuje wartość

$$R = \sqrt{\sup_{\mathbf{u} \in \Delta_D} \left(\ln D + \sum_{i=1}^D u_i \ln u_i \right)} = \sqrt{\ln D}. \quad (4.17)$$

Przyjmując, jak poprzednio, założenie o ograniczonej normie ℓ_∞ gradientu, która to norma jest jednocześnie normą dualną w rozważanym wariacie, tj. $\|\cdot\|_* = \|\cdot\|_\infty$,

zauważamy, że błąd estymacji algorytmu jest rzędu $\mathcal{O}(\sqrt{\frac{\ln D}{T}})$. Oznacza to, że w porównaniu do wariantu euklidesowego, błąd wariantu entropijnego zależy w dużo mniejszym stopniu od wymiaru problemu D .

Uwaga 4.3. Funkcja regularyzująca (4.14) nie jest jedynym przykładem funkcji silnie wypukłej względem normy $\|\cdot\|_1$. Niech $D \geq 3$. Rozważmy funkcję $h(\boldsymbol{\theta}) = e^{\frac{\ln D - 1}{2}} \|\boldsymbol{\theta}\|_p^2$, $p = \frac{\ln D}{\ln D - 1}$. Zgodnie z [98, lemat 17], funkcja ta jest silnie wypukła (z modułem e) względem normy $\|\cdot\| = \|\cdot\|_{\frac{\ln D}{\ln D - 1}}$. Korzystając z nierówności $\|\mathbf{x}\|_\alpha \leq D^{\alpha - 1 - \beta^{-1}} \|\mathbf{x}\|_\beta$, która jest spełniona dla każdego $1 \leq \alpha \leq \beta$ i $\mathbf{x} \in \mathbb{R}^D$, można pokazać, że $\|\mathbf{x}\|_1 \leq D^{\frac{1}{\ln D}} \|\mathbf{x}\|_p = e \|\mathbf{x}\|_p$, co implikuje, że funkcja $h(\cdot)$ jest silnie wypukła z modułem 1 względem normy $\|\cdot\|_1$. Niestety, zgodnie z naszą najlepszą wiedzą, dla danej funkcji regularyzującej nie istnieje jawna forma rozwiązania pośredniego problemu optymalizacyjnego (4.2), a więc w każdej iteracji, algorytm wymaga korzystania z metod przybliżonych wyznaczania rozwiązania, co z kolei bezpośrednio rzutuje na koszt obliczeniowy, a także dokładność całej metody. Tak więc, jak pokazuje powyższy przykład, od postaci funkcji regularyzującej $h(\cdot)$ zależą nie tylko własności teoretyczne algorytmu, ale także praktyczna możliwość jego implementacji.

4.2 Modelowanie systemów dynamicznych

Przejdziemy teraz do omówienia algorytmu DA w zadaniu modelowania systemów dynamicznych. Rozpatrzmy dwie klasy systemów: z pamięcią skończoną oraz pamięcią zanikającą. Otrzymane rezultaty opierają się na następującym założeniu:

Założenie 4.5. Istnieje stała $G < \infty$, taka że dla każdego t i $\boldsymbol{\theta} \in \mathcal{C}$, zachodzi nierówność

$$\mathbb{E}\{\|\nabla \hat{L}_t(\boldsymbol{\theta})\|_*^2 \mid \mathcal{F}_{t-1}\} \stackrel{p.n.}{\leq} G^2.$$

Powyższe założenie pociąga za sobą założenie 4.3 (nie jest mu jednak równoważne) i dlatego, z teoretycznego punktu widzenia, jest założeniem nieco silniejszym. Jednak ze względu na pozostałe założenia, które poczynimy o systemie (w szczególności o szumie losowym), warunkowanie wartości oczekiwanej względem \mathcal{F}_{t-1} nie będzie miało istotnego wpływu na sposób szacowania stałej G .

Uwaga 4.4. Należy zwrócić uwagę, że w zadaniu modelowania systemów dynamicznych, w celu ewaluacji wyjścia modelu o długości pamięci M (i wyznaczenia gradientu \mathbf{g}_t empirycznej funkcji starty), wymagany jest wektor zawierający M ostatnich wejść systemu. Jeżeli więc, w pewnej chwili czasu T , zaobserwowaliśmy łącznie N par obserwacji wejścia-wyjścia $\{X_i, Y_i\}$, efektywna liczba pomiarów wynosi $N - (M - 1)$, gdyż pierwsze $M - 1$ obserwacje wejścia traktowane są jako warunki początkowe modelu. W związku z tym, w ogólności, liczba iteracji algorytmu jest różna od liczby pomiarów, tzn. $T \neq N$. Biorąc pod uwagę powyższy komentarz

oraz uwzględniając rozwiązanie początkowe, $\boldsymbol{\theta}_1$, dochodzimy do wniosku, iż T i N są związane następującą zależnością

$$T = N - (M - 1) + 1 = N - M + 2.$$

4.2.1 Systemy z pamięcią skończoną

Zasadniczą trudnością związaną z modelowaniem systemów dynamicznych jest fakt, że w ogólności, zarówno wyjście systemu Y_t , jak i wektor wejść modelu \mathbf{X}_t są zależne od \mathcal{F}_{t-1} , wobec czego $\mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\} \neq L_S(\boldsymbol{\theta}_t)$, a więc nie możemy bezpośrednio polegać na technice dowodowej przedstawionej w lemacie 4.1. Zauważmy jednak, że jeśli długość pamięci systemu i modelu jest nie większa niż M , to spełniona jest tożsamość $\mathbb{E}\{\hat{L}_{t+M-1}(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\} = L_S(\boldsymbol{\theta}_t)$. Obserwację tę wykorzystamy w dowodzie kolejnego twierdzenia. Najpierw jednak wprowadźmy założenia o pamięci systemu i modelu:

Założenie 4.6. Długość pamięci systemu jest znana i wynosi $M < \infty$.

Założenie 4.7. Długość pamięci poszczególnych elementów słownika jest nie większa niż długość pamięci systemu.

Zwróćmy uwagę, że wobec znajomości długości pamięci systemu (założenie 4.6), stosowanie modeli o pamięci dłuższej niż ta, jest bezzasadne, tak więc założenie 4.7 ma wymiar czysto techniczny.

Twierdzenie 4.2. Niech będą spełnione założenia 1.1–1.2 i 4.5–4.7 oraz niech $\{\eta_t\}$ stanowi ciąg niemalejących kroków, których wartość jest ustalona w sposób deterministyczny. Dla każdego $T \geq M$ błąd estymacji algorytmu DA spełnia nierówność

$$\begin{aligned} \mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) &\leq T^{-1}G(M-1) \sum_{t=1}^{T-M+1} \eta_{t+M-2} \\ &\quad + T^{-1}G \sum_{t=1}^{T-M} t(\eta_t - \eta_{t+M-1}) \\ &\quad + T^{-1}\left(\frac{G^2}{2} \sum_{t=1}^T \eta_{t-1} + \eta_T^{-1}R^2\right) \\ &\quad + 2T^{-1}(M-1)GK, \end{aligned} \tag{4.18}$$

gdzie $\eta_0 = \eta_1$ i $K = \max_{\mathbf{x} \in \mathcal{C}} \|\mathbf{x}\|$.

Dowód. Dowód przedstawiony został w dodatku 4.B. □

Wniosek 4.3. Dla ciągu kroków, którego elementy zdefiniowane są przez $\eta_t = \frac{R}{G\sqrt{t}}$, $t \geq 1$, nierówność (4.18) przyjmuje postać

$$\begin{aligned} \mathbb{E}\{L_S(\bar{\theta}_T)\} - L_S(\theta^*) &< R \frac{2G + 3(M-1)}{\sqrt{T}} + \frac{2GK(M-1)}{T} \\ &< C_1 \frac{M}{\sqrt{T}} + C_2 \frac{M}{T}, \end{aligned} \quad (4.19)$$

gdzie $C_1 = R(2G + 3)$ i $C_2 = 2GK$.

Dowód. Jak łatwo sprawdzić, dla każdego $T \geq T_0 \geq 1$ zachodzi

$$\sum_{t=T_0}^T \frac{1}{\sqrt{t}} \leq \frac{1}{\sqrt{T_0}} + \int_{T_0}^T \frac{1}{\sqrt{t}} dt = \frac{1}{\sqrt{T_0}} + 2(\sqrt{T} - \sqrt{T_0}).$$

Zgodnie z powyższym, otrzymujemy następującą nierówność

$$\begin{aligned} \sum_{t=1}^{T-M+1} \eta_{t+M-2} &= \frac{R}{G} \sum_{t=M-1}^{T-1} \frac{1}{\sqrt{t}} \\ &\leq \frac{R}{G} \left(\frac{1}{\sqrt{M-1}} + 2(\sqrt{T-1} - \sqrt{M-1}) \right) \\ &< \frac{2R\sqrt{T}}{G}, \end{aligned}$$

tym samym, pierwszy składnik prawej strony wzoru (4.18) ograniczamy z góry przez

$$T^{-1}G(M-1) \sum_{t=1}^{T-M+1} \eta_{t+M-2} \leq \frac{2R(M-1)}{\sqrt{T}}.$$

Do oszacowania drugiego składnika skorzystamy z nierówności

$$\sqrt{x+1} - \sqrt{x} < \frac{1}{2\sqrt{x}},$$

która jest spełniona dla każdego $x > 0$. Jej zastosowanie prowadzi do wniosku, że dla dowolnego $t, \tau \geq 1$ zachodzi

$$\begin{aligned} \eta_t - \eta_{t+\tau} &= \frac{R}{G} \left(\frac{1}{\sqrt{t}} - \frac{1}{\sqrt{t+\tau}} \right) < \frac{R}{G} \frac{\sqrt{t+\tau} - \sqrt{t}}{t} = \frac{R}{G} \frac{\sqrt{\tau}(\sqrt{\frac{t}{\tau} + 1} - \sqrt{\frac{t}{\tau}})}{t} \\ &< \frac{R\tau}{2Gt^{\frac{3}{2}}}, \end{aligned}$$

co prowadzi do następującej nierówności,

$$\begin{aligned} T^{-1}G \sum_{t=1}^{T-M} t(\eta_t - \eta_{t+M-1}) &< \frac{R(M-1)}{2T} \sum_{t=1}^{T-M} \frac{1}{\sqrt{t}} \\ &< \frac{R(M-1)}{2T} (2\sqrt{T-M} - 1) \\ &< \frac{R(M-1)}{\sqrt{T}}. \end{aligned}$$

Z kolei trzeci składnik został oszacowany już wcześniej, w lemacie 4.1,

$$T^{-1} \left(\frac{G^2}{2} \sum_{t=1}^T \eta_{t-1} + \eta_T^{-1} R^2 \right) < \frac{2RG}{\sqrt{T}}.$$

Ostatnim krokiem dowodu jest podstawieniu otrzymanych oszacowań do wzoru (4.18). \square

Jak możemy zaobserwować, dla *dużych* T , błąd estymacji algorytmu (4.19) jest nie większy niż $\frac{R}{\sqrt{T}}(2G + 3(M - 1))$. W porównaniu do wyniku z wniosku 4.1, oszacowanie to jest gorsze o stałą, tym większą, im pamięć systemu jest dłuższa. W szczególności, dla $M = 1$, wynik pozostaje jednak taki sam.

4.2.2 Systemy z pamięcią zanikającą

Na potrzeby tego paragrafu, przytoczymy definicję pamięci zanikającej, zaproponowaną w pracy [114]. Upřednio jednak, wprowadzimy symbol

$$\mathbf{x}_n^{[\tau]} = [x_n, x_{n-1}, \dots, x_{n-\tau}, 0, 0, \dots],$$

którym będziemy oznaczali obcięty odpowiednik wektora \mathbf{x}_n .

Definicja 4.1 ([114]). Funkcjonał $m : \mathbb{R}^\infty \rightarrow \mathbb{R}$ ma pamięć zanikającą eksponencjalnie, jeśli dla dowolnej, ograniczonej klasy sygnałów wejściowych \mathbf{x}_n , istnieją takie stałe c_0, c_1 , że nierówność

$$\left| m[\mathbf{x}_n] - m[\mathbf{x}_n^{[\tau]}] \right| \leq c_0 e^{-c_1 \tau},$$

jest spełniona dla każdego $\tau = 1, 2, \dots$

Założenie 4.8. Charakterystyka systemu $m_{sys}[\cdot]$ posiada pamięć zanikającą eksponencjalnie z dowolnymi stałymi $c_0, c_1 < \infty$.

Założenie 4.9. Elementy słownika posiadają pamięć skończoną.

Założenie 4.10. Dla danej klasy sygnałów wejściowych, charakterystyka systemu $m_{sys}[\cdot]$ jest ograniczona przez stałą $\bar{m}_{sys} < \infty$, tj.

$$P(|m_{sys}[\mathbf{X}_n]| \leq \bar{m}_{sys}) = 1.$$

Założenie 4.11. Dla danej klasy sygnałów wejściowych oraz słownika S , dla każdego $\boldsymbol{\theta} \in \mathcal{C}$ wyjście modelu $m_S[\cdot; \boldsymbol{\theta}]$ jest ograniczone przez stałą $\bar{m} < \infty$, tj.

$$P(|m_S[\mathbf{X}_n, \boldsymbol{\theta}]| \leq \bar{m}) = 1.$$

Twierdzenie 4.3. *Niech będą spełnione założenia 1.1–1.2, 4.3 i 4.8–4.11. Dla każdego $\tau \geq 0$ błąd estymacji algorytmu DA spełnia nierówność*

$$\begin{aligned} \mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) &\leq T^{-1}\tau G \sum_{t=1}^{T-\tau} \eta_{t+\tau-1} + C\epsilon_\tau \\ &\quad + T^{-1}G \sum_{t=1}^{T-\tau-1} t(\eta_t - \eta_{t+\tau}) \\ &\quad + T^{-1}\left(\frac{G^2}{2} \sum_{t=1}^T \eta_{t-1} + \eta_T^{-1}R^2\right) \\ &\quad + 2T^{-1}\tau GK, \end{aligned} \quad (4.20)$$

gdzie $\eta_0 = \eta_1$, $K = \max_{\mathbf{x} \in \mathcal{C}} \|\mathbf{x}\|$, $\epsilon_\tau = e^{-c_1\tau}$, $C < \infty$ jest pewną znaną stałą, a stała c_1 została określona w założeniu 4.8.

Dowód. Dowód przedstawiony został w dodatku 4.C. □

Zauważmy, że po dokonaniu podstawienia $\tau = M - 1$, otrzymany powyżej rezultat jest niemal identyczny z oszacowaniem otrzymanym w twierdzeniu 4.2 – jedyną różnicę stanowi komponent $C\epsilon_\tau$, który zanika do zera, gdy $\tau \rightarrow \infty$. Jednak wraz ze wzrostem τ , rośnie także wartość ostatniego składnika wzoru (4.2), tj. $2T^{-1}\tau GK$. Wynika z tego, że zbieżność błędu estymacji następuje, gdy zachowana jest relacja $\tau = o(T)$, tzn. τ musi rosnać wolniej niż T .

Wniosek 4.4. *Niech $T \geq 2$. Dla $\tau = \lceil \ln T \rceil$ i ciągu kroków, którego elementy zdefiniowane są przez $\eta_t = \frac{R}{G\sqrt{t}}$, $t \geq 1$, nierówność (4.20) przyjmuje postać*

$$\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) < C_1 \frac{\lceil \ln T \rceil}{\sqrt{T}} + C_2 \frac{\lceil \ln T \rceil}{T} + C_3 \frac{1}{T^{c_1}}, \quad (4.21)$$

gdzie $C_1 = R(2G + 3)$, $C_2 = 2GK$ i $C_3 = Ce^{-c_1}$.

Dowód. Dla danego τ , stała ϵ_τ przyjmuje postać $\epsilon_\tau = e^{-c_1 \lceil \ln T \rceil} \leq e^{-c_1(\ln T + 1)} = e^{-c_1} T^{-c_1}$. Postępując jak w dowodzie wniosku 4.3, otrzymujemy

$$\mathbb{E}\{L_S(\bar{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) \leq \frac{R}{\sqrt{T}}(2G + 3\tau) + \frac{2GK\tau}{T} + C\epsilon_\tau.$$

Dokonując podstawienia $\tau = \lceil \ln T \rceil$ uzyskujemy ostateczny rezultat. □

Jak możemy zaobserwować, szybkość zbieżności oszacowania (4.21) zależy od stałej c_1 , która z kolei określa szybkość zanikania pamięci systemu. Jeśli więc wartość stałej c_1 jest na tyle duża, że wyrażenie $\frac{1}{T^{c_1}}$ zbiega szybciej niż $\frac{\lceil \ln T \rceil}{\sqrt{T}}$, to błąd estymacji algorytmu jest rzędu $\mathcal{O}\left(\frac{\lceil \ln T \rceil}{\sqrt{T}}\right)$; w przeciwnym wypadku rząd błędu wynosi $\mathcal{O}\left(\frac{1}{T^{c_1}}\right)$. Mimo, że w porównaniu z wynikiem otrzymanym we wniosku 4.7, oszacowanie (4.21) jest ilościowo gorsze, to należy pamiętać, że obejmuje znacznie szerszą klasę systemów (w tym systemy z pamięcią skończoną), a przy tym zachodzi dla dowolnego $T \geq 2$ (por. ograniczenie nałożone na T w twierdzeniu 4.2).

4.2.3 Modelowanie systemów z pamięcią zanikającą za pomocą wariantu entropijnego algorytmu DA oraz modeli opartych o reprezentacje wielomianowe

Omówimy teraz zastosowanie algorytmu DA w zadaniu modelowania nieliniowych systemów dynamicznych za pomocą modeli opartych o słowniki wielomianowe. Ze względu na atrakcyjne własności statystyczne, skupimy się na entropijnym wariantcie metody.

Rozpatrywane zadanie modelowania dotyczy systemów, które opisać można za pomocą wzoru (1.1) i które spełniają ogólne założenia 1.1, 1.2, a także założenie o pamięci zanikającej 4.8. Dodatkowo, o sygnale wejściowym założymy, że:

Założenie 4.12. *Sygnal wejściowy jest ograniczony, w taki sposób, że*

$$P(|X_n| \leq 1) = 1.$$

Algorytm 4.1 Wariant adaptacyjny algorytm DA dla modelu wielomianowego

Wejścia: r, M, D , rozszerzony słownik S

$$\theta_{1,i} \leftarrow \frac{1}{2D}, \quad i = 1, \dots, 2D$$

$$\bar{\theta}_1 \leftarrow \theta_1$$

for $t = -M + 2, \dots, 0$ **do**

Dokonaj pomiaru wejścia X_t (warunek początkowy modelu)

end for

for $t = 1, \dots, \infty$ **do**

Dokonaj pomiaru wejścia X_t i wyjścia Y_t systemu

$$\mathbf{X}_t \leftarrow [X_t, X_{t-1}, \dots, X_{t-M+1}]$$

$$\text{Oblicz gradient funkcji straty } \mathbf{g}_t \leftarrow 2(m_S(\mathbf{X}_t; \theta_t) - Y_t) [\varphi_1(\mathbf{X}_t), \dots, \varphi_D(\mathbf{X}_t)]^T$$

Wyznacz długość kroku η_t

$$\theta_{t+1,i} \leftarrow \exp(-\eta_t g_{1:t,i}) \left(\sum_{j=1}^{2D} \exp(-\eta_t g_{1:t,j}) \right)^{-1}, \quad i = 1, \dots, 2D$$

$$\bar{\theta}_{t+1} \leftarrow \frac{1}{t+1} \theta_t + \frac{t}{t+1} \bar{\theta}_t$$

$$\text{Zwróć parametry modelu } \hat{\theta}_{t+1} \leftarrow r(\bar{\theta}_{t+1,i} - \bar{\theta}_{t+1,i+D}), \quad i = 1, \dots, D$$

end for

4.2.3.1 Rozszerzenie słownika

O ile niewielka wrażliwość błędu wariantu entropijnego na wymiar problemu D jest cechą szczególnie pożądaną, o tyle, jako znaczne ograniczenie stosowalności algorytmu, postrzegany może być fakt, iż zbiór potencjalnych rozwiązań ograniczony jest do sympleksu Δ_D . Taki stan rzeczy wyklucza bowiem ujemne parametry potencjalnego modelu, co razem z wymaganiem dotyczącym sumowalności wag do jedności sprawia, że nie mamy kontroli, innej niż manipulacja słownikiem, nad błędem aproksymacji metody.

W celu zwiększenia możliwości aproksymacyjnych algorytmu, proponuje się usprawnienie (analogiczne do zaprezentowanego w [59]), które rozszerza zbiór potencjalnych rozwiązań z Δ_D na $\Lambda_{D,r}$, gdzie

$$\Lambda_{D,r} = \left\{ \boldsymbol{\theta} \in \mathbb{R}^D : \|\boldsymbol{\theta}\|_1 \leq r \right\}$$

jest kulą ℓ_1 o promieniu r . Niech $\alpha \cdot S = \{\alpha\varphi : \varphi \in S\}$ oznacza operację, która skaluje wszystkie elementy zbioru S przez stałą $\alpha \in \mathbb{R}$. Do celów modelowania, zamiast standardowego słownika wielomianowego $S_{P,M}$, w zmodyfikowanej wersji algorytmu będziemy korzystali ze słownika rozszerzonego, określonego wzorem $\tilde{S}_{P,M}^r = r \cdot S_{P,M} \cup (-r) \cdot S_{P,M}$. Jak można zauważyć, $\tilde{S}_{P,M}^r$ składa się z przeskalowanej wersji słownika $S_{P,M}$ oraz jego odpowiednika z przeciwnym znakiem, mianowicie

$$\tilde{S}_{P,M}^r = \left\{ r\mathbf{1}, r\varphi(0), \dots, r\varphi(L-1, \dots, L-1), -r\mathbf{1}, -r\varphi(0), \dots, -r\varphi(L-1, \dots, L-1) \right\},$$

przy czym, należy pamiętać, że rozmiar zmodyfikowanego słownika jest dwa razy większy od oryginału, tj. $|\tilde{S}_{P,M}^r| = 2D$. Za pomocą prostych przekształceń można pokazać, że zbiór wszystkich kombinacji wypukłych elementów słownika $\tilde{S}_{P,M}^r$ (czyli otoczka wypukła zbioru $\tilde{S}_{P,M}^r$) jest tożsamy zbiorowi wszystkich liniowych kombinacji elementów $S_{P,M}$ o współczynnikach należących do $\Lambda_{D,r}$,

$$\begin{aligned} \left\{ \sum_{i=1}^{2D} \theta_i \varphi_i(\cdot) : \boldsymbol{\theta} \in \Delta_{2D}, \varphi_i \in \tilde{S}_{P,M}^r \right\} &= \left\{ \sum_{i=1}^D (\theta_i - \theta_{i+D}) \varphi_i(\cdot) : \boldsymbol{\theta} \in \Delta_{2D}, \varphi_i \in r \cdot S_{P,M} \right\} \\ &= \left\{ \sum_{i=1}^D r(\theta_i - \theta_{i+D}) \varphi_i(\cdot) : \boldsymbol{\theta} \in \Delta_{2D}, \varphi_i \in S_{P,M} \right\} \\ &= \left\{ \sum_{i=1}^D \theta_i \varphi_i(\cdot) : \boldsymbol{\theta} \in \Lambda_{D,r}, \varphi_i \in S_{P,M} \right\}, \end{aligned}$$

a więc, innymi słowy, zachodzi

$$\left\{ m_{\tilde{S}_{P,M}^r}(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Delta_{2D} \right\} = \left\{ m_{S_{P,M}}(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Lambda_{D,r} \right\}.$$

Oznacza to, że poprzez użycie słownika rozszerzonego, wariant entropijny algorytmu jest w stanie imitować estymację modelu związanego ze słownikiem standardowym $S_{P,M}$ na zbiorze $\mathcal{C} = \Lambda_{D,r}$. Dzięki takiemu zabiegowi, w modelu dopuszczone są ujemne wagi, a wielkość zbioru \mathcal{C} kontrolowana jest za pomocą parametru r . Zaproponowana modyfikacja wymaga dodatkowego kroku którym jest odwzorowanie rozwiązania $\bar{\boldsymbol{\theta}}_T \in \Delta_{2D}$ w $\hat{\boldsymbol{\theta}}_T \in \Lambda_{D,r}$, co czyni się w sposób następujący:

$$\forall_{i=1, \dots, D} \quad \hat{\theta}_{T,i} = r(\bar{\theta}_{T,i} - \bar{\theta}_{T,i+D}),$$

dzięki czemu, model końcowy spełnia tożsamość

$$m_{S_{P,M}}(\cdot; \hat{\boldsymbol{\theta}}_T) = m_{\tilde{S}_{P,M}^r}(\cdot; \bar{\boldsymbol{\theta}}_T).$$

4.2.3.2 Własności statystyczne algorytmu

Opierając się o twierdzenie 4.3 i wniosek 4.4, wyznaczmy teraz oszacowanie błędu estymacji algorytmu 4.1.

Wniosek 4.5. *Niech będą spełnione założenia 1.1–1.2, 4.8, 4.10 oraz 4.12, i niech $\hat{\boldsymbol{\theta}}_T \in \Lambda_{D,r}$ będzie rezultatem działania algorytmu 4.1 ze słownikiem $S = \tilde{S}_{P,M}^r$ i ciągami kroków zdefiniowanymi wzorem $\eta_t = \frac{\sqrt{\ln(2D)}}{G\sqrt{t}}$, gdzie $G = 2r\sqrt{(r + \bar{m}_{\text{sys}})^2 + \sigma_Z^2}$. Dla każdego $T \geq 2$ błąd estymacji algorytmu spełnia nierówność*

$$\mathbb{E}\{L_S(\hat{\boldsymbol{\theta}}_T)\} - L_S(\boldsymbol{\theta}^*) < C_1 \frac{[\ln T]}{\sqrt{T}} \sqrt{\ln(2D)} + C_2 \frac{[\ln T]}{T} + C_3 \frac{1}{T^{c_1}}, \quad (4.22)$$

gdzie

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in \Lambda_{D,r}} L_{S_{P,M}}(\boldsymbol{\theta}),$$

$C_1 = 2G + 3$, $C_2 = 2G$, a stałe C_3 oraz c_1 zostały określone kolejno we wniosku 4.8 i w założeniu 4.8.

Dowód. Zauważmy, że za sprawą założenia 4.12 każdy element słownika $\tilde{S}_{P,M}^r$ jest jednakowo ograniczony, $|\varphi_i(\mathbf{X}_n)| < r$, co sprawia, że wyjście modelu jest również ograniczone, tj. dla każdego $\boldsymbol{\theta} \in \Delta_{2D}$ zachodzi

$$\left| m_{\tilde{S}_{P,M}^r}[\mathbf{X}_n; \boldsymbol{\theta}] \right| = \left| \sum_{i=1}^{2D} \theta_i \varphi_i(\cdot) \right| \leq r \sum_{i=1}^{2D} |\theta_i| = r \|\boldsymbol{\theta}\|_1 = r, \quad (4.23)$$

a więc założenie 4.11 jest spełnione ze stałą $\bar{m} = r$. W następstwie, nierówność (4.23) oraz założenia 4.12 i 4.10 pozwalają znaleźć górne ograniczenie następującego wyrażenia

$$\begin{aligned} & \sup_{\boldsymbol{\theta} \in \Delta_{2D}} \mathbb{E} \left\{ \left\| \nabla \hat{L}_t(\boldsymbol{\theta}) \right\|_\infty^2 \mid \mathcal{F}_{t-1} \right\} = \\ & = 4 \sup_{\boldsymbol{\theta} \in \Delta_{2D}} \mathbb{E} \left\{ \left(m_{\tilde{S}_{P,M}^r}[\mathbf{X}_n; \boldsymbol{\theta}] - Y_t \right)^2 \left\| [\varphi_1(\mathbf{X}_n), \dots, \varphi_{2D}(\mathbf{X}_n)] \right\|_\infty^2 \mid \mathcal{F}_{t-1} \right\} \\ & \leq 4r^2 \sup_{\boldsymbol{\theta} \in \Delta_{2D}} \mathbb{E} \left\{ \left(m_{\tilde{S}_{P,M}^r}[\mathbf{X}_n; \boldsymbol{\theta}] - m_{\text{sys}}[\mathbf{X}_n] \right)^2 + Z_n^2 \mid \mathcal{F}_{t-1} \right\} \\ & \leq 4r^2 \left((r + \bar{m}_{\text{sys}})^2 + \sigma_Z^2 \right). \end{aligned}$$

Tym samym stwierdzamy, że założenie 4.3 jest spełnione ze stałą

$$G = \sqrt{\sup_{\boldsymbol{\theta} \in \Delta_{2D}} \mathbb{E} \left\{ \left\| \nabla \hat{L}_t(\boldsymbol{\theta}) \right\|_\infty^2 \mid \mathcal{F}_{t-1} \right\}} = 2r\sqrt{(r + \bar{m}_{\text{sys}})^2 + \sigma_Z^2} \quad (4.24)$$

i w efekcie, algorytm spełnia wszystkie warunki twierdzenia 4.10. Zgodnie ze wzorem 4.17, w rozważanym zadaniu, stała R przyjmuje wartość $R = \sqrt{\ln(2D)}$. Z kolei stała K z twierdzenia 4.10 przyjmuje wartość $K = \max_{\mathbf{x} \in \Delta_{2D}} \|\mathbf{x}\|_1 = 1$. Ostateczny rezultat jest wynikiem zastosowania nierówności 4.21 z wniosku 4.4. \square

Podsumowując, otrzymaliśmy algorytm, którego błąd estymacji zależy tylko w niewielkim stopniu od liczby elementów słownika. W porównaniu z metodą opartą o agregację wypukłą z ograniczeniem ℓ_1 [114], algorytm 4.1 ma nieco gorszą szybkość zbieżności ($\mathcal{O}(\frac{\sqrt{\lceil \ln N \rceil}}{\sqrt{N}} \sqrt{\ln D} + \frac{1}{N^{\epsilon_1}})$ vs. $\mathcal{O}(\frac{\lceil \ln N - M + 2 \rceil}{\sqrt{N - M + 2}} \sqrt{\ln(2D)} + \frac{1}{(N - M + 2)^{\epsilon_1}})$). Jego zaletą pozostaje natomiast fakt, że dzięki odpowiednio dobranej funkcji regularyzującej, posiada on jawną formę rozwiązania i – jak wykażemy w eksperymencie numerycznym – cechuje się niską złożonością obliczeniową.

4.3 Modelowanie systemu LNL

Niniejsza część pracy stanowi uzupełnienie podrozdziału 2.3, w którym wprowadziliśmy wzór na górne ograniczenie błędu aproksymacji reprezentacji volterrowskiej dla pewnej klasy systemów LNL. Uzyskany wynik, w połączeniu z rezultatami otrzymanymi w niniejszym rozdziale, pozwala na ustalenie warunków zbieżności metody modelowania opartej o algorytm DA i słownik volterrowski dla systemów LNL z pamięcią zanikającą. Analizę błędu metody ograniczymy do wariantu entropijnego algorytmu.

4.3.1 Opis problemu

Rozpatrzmy system opisany wzorem

$$Y_n = m_{\text{LNL}}[\mathbf{X}_n] + Z_n, \quad (4.25)$$

gdzie charakterystyka $m_{\text{LNL}}[\cdot]$ została zdefiniowana we wzorze (2.8), a Z_n stanowi szum addytywny. Zakładamy, że spełnione są założenia standardowe 1.1–1.2 i wprowadzone wcześniej założenia 2.1–2.3. Wprowadzimy także dodatkowe założenie o ograniczonym wyjściu podsystemu nieliniowego, które – po uwzględnieniu stabilności podsystemów liniowych – pozwala stwierdzić nam, że rozważany system ma pamięć zanikającą eksponencjalnie.

Założenie 4.13. *Istnieje stała $\bar{\varphi} < \infty$, taka że dla każdego $x \in \mathbb{R}$, charakterystyka podsystemu nieliniowego spełnia nierówność $|\varphi(x)| \leq \bar{\varphi}$.*

4.3.2 Błąd metody

Przypomnijmy, że błąd aproksymacji przedstawiony we wniosku 2.1 dotyczy modeli klasy $\mathcal{M}_{S_{P,M}} = \{m_{S_{P,M}}[\cdot; \boldsymbol{\theta}] : \boldsymbol{\theta} \in \mathbb{R}^D\}$, natomiast wariant entropijny algorytmu DA przeznaczony jest dla modeli z ograniczeniem ℓ_1 , tj. $\mathcal{M}_{S_{P,M}^r} = \{m_{S_{P,M}^r}[\cdot; \boldsymbol{\theta}] : \boldsymbol{\theta} \in \Delta_D\} = \{m_{S_{P,M}}[\cdot; \boldsymbol{\theta}] : \boldsymbol{\theta} \in \Lambda_{D,r}\}$. Będziemy zakładali więc, że parametr r algorytmu jest dobrany w taki sposób, że najlepszy możliwy model w zbiorze $\mathcal{M}_{S_{P,M}}$

jest jednocześnie elementem zbioru $\mathcal{M}_{S_{P,M}^r}$, tj.

$$\|\boldsymbol{\theta}^*\|_1 \leq r, \quad (4.26)$$

gdzie

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^D} L_{S_{P,M}}(\boldsymbol{\theta}).$$

Jeśli warunek (4.26) jest spełniony, to błąd aproksymacji dla obu klas modeli $\mathcal{M}_{S_{P,M}}$ i $\mathcal{M}_{S_{P,M}^r}$ jest ten sam.

Niech model $\hat{m}[\cdot]$ będzie wynikiem zastosowania algorytmu 4.1. Przypomnijmy, że na wartość funkcji kryterialnej składa się błąd aproksymacji oraz błąd estymacji,

$$Q(\hat{m}) = \Delta_{\text{est.}} + \Delta_{\text{apr.}}.$$

Zauważmy, że dla podanych założeń, system spełnia wszystkie warunki twierdzenia 4.3, dzięki czemu błąd estymacji $\Delta_{\text{est.}}$ spełnia oszacowanie ze wzoru (4.22), natomiast błąd aproksymacji $\Delta_{\text{apr.}}$ został ustalony we wzorze (2.12). Po dokonaniu podstawień, otrzymujemy oszacowanie błędu całej metody

$$\begin{aligned} Q(\hat{m}) &\leq \mathcal{O}(C^{-2P} + \lambda^{-2M} + \frac{\ln T}{\sqrt{T}} \sqrt{\ln(2D)} + \frac{1}{T^{c_1}}) \\ &= \mathcal{O}(\min(C, \lambda)^{-2P} + \min(C, \lambda)^{-2M} + \frac{\ln(N-M)}{\sqrt{N-M}} \sqrt{\ln(2D)} + \frac{1}{(N-M)^{c_1}}). \end{aligned} \quad (4.27)$$

Twierdzenie 4.4. *Niech system opisany równaniem (4.25) spełnia założenia 1.1–1.2, 2.1–2.3 oraz 4.13. Jeśli spełniony jest warunek (4.26) oraz $M = P = o(N)$ to wartość funkcji kryterialnej zbiega do zera $Q(\hat{m}) \rightarrow 0$, gdy liczba pomiarów dąży do nieskończoności $N \rightarrow \infty$.*

Dowód. Z faktu, że $\min(C, \lambda) > 1$, wynika iż pierwsze dwa wyrazy wzoru (4.27) zbiegają do zera, gdy $P, M \rightarrow \infty$. Jednakże wraz ze wzrostem P i M rośnie także liczba elementów w słowniku, a więc $D \rightarrow \infty$. Z tego względu, by otrzymać optymalną (względem wzoru (4.27)) szybkość zbieżności, należy dokonać podstawienia $M = P$, co prowadzi do wzoru

$$D = \binom{P+M}{M} \Big|_{M=P} = \binom{2P}{P}. \quad (4.28)$$

Za sprawą wzoru Stirlinga otrzymujemy górne ograniczenie (4.28),

$$D = \frac{(2P)!}{(P!)^2} \leq \frac{e(2P)^{2P+\frac{1}{2}}}{2\pi P^{2P+1}} \leq 4^P. \quad (4.29)$$

Wstawiając do wzoru (4.27) wyrażenie (4.29), otrzymujemy

$$\begin{aligned} Q(\hat{m}) &= \mathcal{O}(\min(C, \lambda)^{-2P} + \frac{\ln(N-P)}{\sqrt{N-P}} \sqrt{\ln(4^{P+\frac{1}{2}})}) + \frac{1}{(N-M)^{c_1}} \\ &= \mathcal{O}(\min(C, \lambda)^{-2P} + \sqrt{P} \frac{\ln(N-P)}{\sqrt{N-P}} + \frac{1}{(N-M)^{c_1}}). \end{aligned}$$

Z powyższego oszacowania wynika, że aby wartości funkcji kryterialnej zbiegała do zera, N i P (a zatem również M) muszą dążyć do nieskończoności, jednocześnie jednak, P musi zbiegać wolniej niż N – tak więc zbieżność metody następuje, gdy zachowana jest relacja $P = o(N)$. \square

4.4 Eksperyment numeryczny

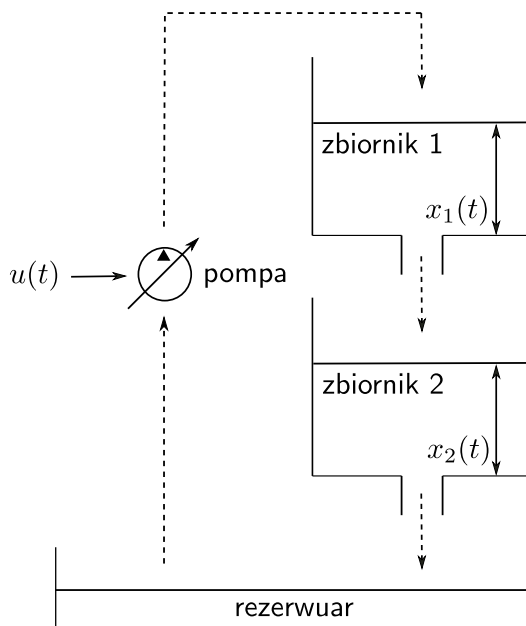
Przejdziemy teraz do przedstawienia rezultatów eksperymentu numerycznego, w którym algorytm 4.1 został użyty na zbiorze danych pomiarowych uzyskanych w ramach eksperymentu przeprowadzonego na fizycznie istniejącym systemie [97]. Głównym celem eksperymentu jest pokazanie, że mimo braku rezultatów teoretycznych odnoszących się bezpośrednio do adaptacyjnego doboru kroku w zadaniu modelowania systemów dynamicznych, w praktyce podejście to jest uzasadnione.

Badany system składał się z dwóch zamocowanych nad sobą zbiorników, umieszczonego pod nimi rezerwuaru i pompy, za pomocą której woda znajdująca się w rezerwuarze była przenoszona do górnego zbiornika. Równocześnie, przez mały otwór znajdujący się w dnie górnego zbiornika, woda spływała do zbiornika dolnego, a następnie, znowu przez mały otwór w dnie, z powrotem do rezerwuaru, tak więc układ ten działał w obiegu zamkniętym (rys. 4.1). Wejściem systemu był sygnał kontrolujący moc pompy, natomiast jego wyjściem był poziom cieczy w drugim zbiorniku.

Model fizyczny systemu przedstawia się następującym układem równań różniczkowych [97]:

$$\begin{aligned} \dot{x}_1(t) &= -k_1 \sqrt{x_1(t)} + k_4 u(t) + w_1(t), \\ \dot{x}_2(t) &= k_2 \sqrt{x_1(t)} - k_3 \sqrt{x_2(t)} + w_2(t), \\ y(t) &= x_2(t) + e(t), \end{aligned}$$

gdzie $u(t)$ oznacza sygnał wejściowy systemu, $x_1(t)$ i $x_2(t)$ to zmienne stanu (poziom cieczy w pierwszym i drugim zbiorniku), $w_1(t)$, $w_2(t)$ i $e(t)$ to zakłócenia, a k_1 , k_2 , k_3 i k_4 to stałe, zależne od własności systemu. Na podstawie powyższych równań można stwierdzić, że system nie przejawia silne nieliniowych zachowań. Model ten jednak nie uwzględnia możliwości przelewania się wody ze zbiorników, co może nastąpić (i *de facto*, jak wynika z danych pomiarowych, w czasie eksperymentu nastąpiło), gdy przepustowość pompy jest zbyt duża względem prędkości jej ubywania ze zbiornika.



Rysunek 4.1: Schemat ideowy badanego systemu. Linią przerywaną zaznaczono kierunek przepływu cieczy

W przypadku, gdy taka sytuacja nastąpi, woda zaczyna wylewać się ze zbiornika górnego i trafia częściowo do rezerwuaru, a częściowo do zbiornika dolnego, w następstwie czego, po pewnym czasie także drugi zbiornik może zostać przepięniony. Zjawisko to jest źródłem dodatkowych zakłóceń i sprawia, że w pewnych warunkach, badany system ma silnie nieliniowy charakter.

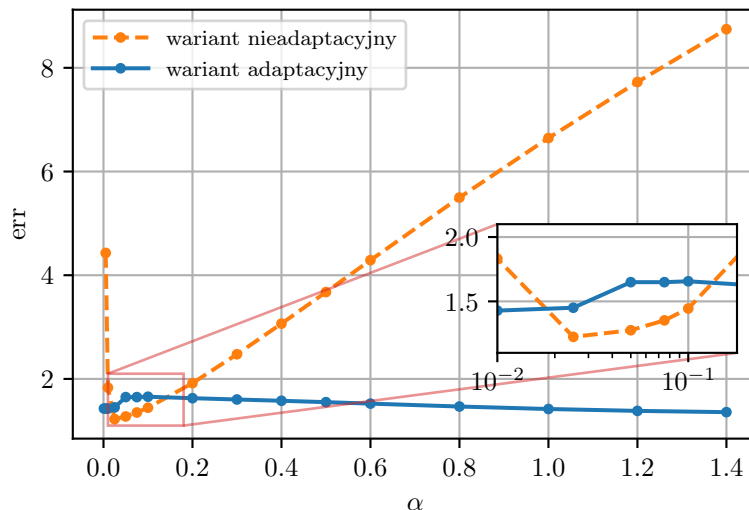
Rozpatrywany *benchmark* składa się z dwóch zbiorów pomiarowych, zawierających po 1024 par pomiarów wejścia-wyjścia; zbiór uczący $\mathcal{S}_{\text{est.}}$, przeznaczony jest do estymacji modelu, natomiast zbiór drugi $\mathcal{S}_{\text{wal.}}$, zwany zbiorem *walidującym*, jest niedostępny podczas procesu estymacji i służy do oceny jakości uzyskanego modelu za pomocą błędu empirycznego, zgodnie ze wzorem

$$\text{err}(m) = \frac{1}{1024} \sum_{i=1}^{1024} (m(\mathbf{X}_i) - Y_i)^2,$$

gdzie m to ewaluowany model, $\{X_i, Y_i\}_{i=1}^{1024}$ to pomiary należące do zbioru walidującego. W obu przypadkach, stan początkowy systemu jest ten sam, przy czym pozostaje on nieznanym użytkownikowi. W procesie walidacji modelu, przyjęliśmy, że nieznanne warunki początkowe są równe zero, tj. $X_0 = X_{-1} = \dots = 0$. Sygnał wejściowy został przeskalowany w taki sposób, aby spełnione było założenie 4.12.

Z powodu braku ilościowych informacji o sygnale wyjściowym i zakłócającym, wartość stałej G wyznaczona została w oparciu o zmodyfikowany wzór,

$$G = 2r \sqrt{(r + \hat{U})^2 + \hat{\sigma}_Z} \quad (4.30)$$



Rysunek 4.2: Eksperyment I: błąd modelu dla $N = 1024$ i stałej G przeskalowanej przez α

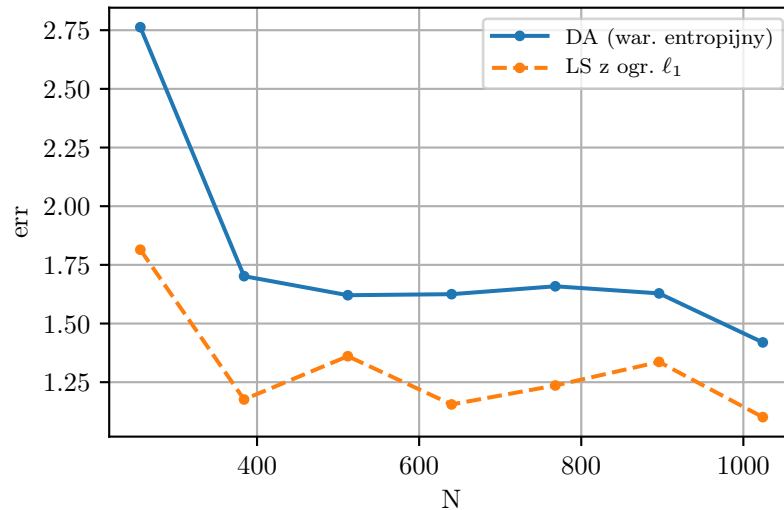
(por. (4.24)), gdzie $\hat{U} = \max_i(|Y_i| \mid Y_i \in \mathcal{S}_{\text{est.}})$, a za $\hat{\sigma}_Z$ przyjęliśmy wartość równą 10% z \hat{U} . Wartość parametru $r = 25$ i długości pamięci poszczególnych obciętych operatorów Volterra znajdujących się w użytym słowniku, $S_{(100,100,20)}$ (zawierającym $D = 6691$ elementów) została dobrana w drodze eksperymentu, podczas którego heurystycznie dobierano parametry słownika tak, aby uzyskany błąd był jak najmniejszy.

W nieadaptacyjnym wariacie algorytmu użyto ciągu kroków zdefiniowanego we wniosku 4.5, natomiast w wariacie adaptacyjnym do określenia długości kroków użyto następującego wzoru (por. wniosek 4.2)

$$\eta_t = \frac{\sqrt{\ln 2D}}{\sqrt{G^2 + \sum_{i=1}^t \|\mathbf{g}_i\|_*^2}}.$$

4.4.1 Eksperyment I - zależność błędu metody od wartości stałej G

W pierwszej części eksperymentu zbadano błąd obu wariantów metody w funkcji stałej G , która jest jednym z czynników decydujących o długości kroku algorytmu. W tym celu zmierzono jakość modeli powstałych, gdy krok metody wyznaczany był w oparciu o stałą $G_\alpha = \alpha G$, gdzie wartość G była wyznaczona za pomocą wzoru (4.30), a wartość parametru α wahała się w zakresie od 10^{-3} do 1,4. Do estymacji modeli wykorzystano wszystkie $N = 1024$ par pomiarów zbioru $\mathcal{S}_{\text{est.}}$. Wyniki symulacji, przedstawione na rys. 4.2, pokazują, że dla pierwotnej wartości G (tj. dla przypadku, gdy $\alpha = 1$), algorytm nieadaptacyjny jest wysoce nieefektywny. Niemniej jednak, istnieje pewien przedział wartości α , dla których modele uzyskane



Rysunek 4.3: Eksperyment II: błąd modelu w zależności od liczby pomiarów

za pomocą metody z sekwencją kroków znaną *a priori*, wykazują błąd mniejszy od tych uzyskanych za pomocą adaptacyjnego odpowiednika algorytmu. Należy zauważyć natomiast, że przedział ten jest bardzo wąski, co prowadzi do wniosku, iż metoda ta jest bardzo wrażliwa na dobór długości kroku. Z drugiej strony, wariant adaptacyjny wykazuje wysoką odporność na sposób wyliczania stałej G , która nie wymaga dalszego dostrajania, gdyż błąd metody jest prawie niezależny od parametru α .

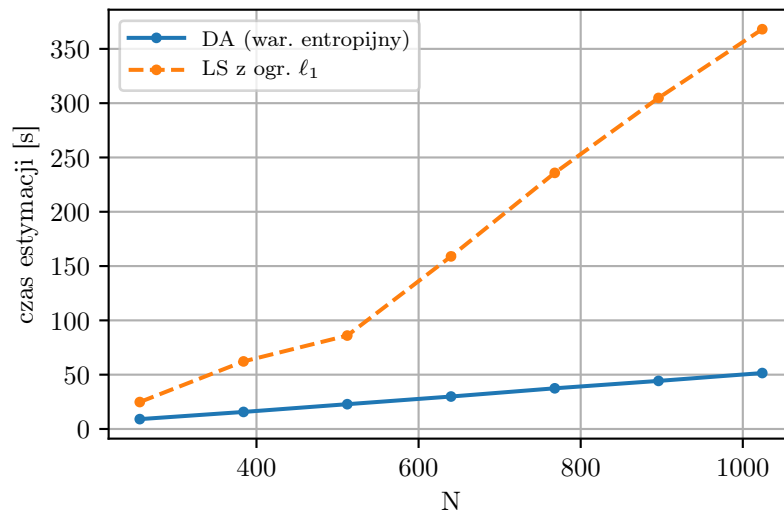
4.4.2 Eksperyment II - zależność błędu metody od liczby pomiarów

W drugiej części eksperymentu zbadana została zależność błędu adaptacyjnego wariantu metody oraz czasu estymacji modelu od liczby pomiarów N . Do celów porównawczych, analogiczny eksperyment wykonany został dla metody agregacyjnej [114, 116], w której model wyjściowy minimalizuje empiryczną funkcję straty z ograniczeniami ℓ_1 , tj.

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}: \|\boldsymbol{\theta}\|_1 \leq r} \frac{1}{N - M + 1} \sum_{i=M}^N (Y_i - m_S[\mathbf{X}_i; \boldsymbol{\theta}])^2,$$

gdzie $M = 100$ to długość pamięci modelu. W naszej implementacji, do rozwiązania powyższego problemu optymalizacyjnego posłużyliśmy się biblioteką CVXPY [1, 21].

Do estymacji modeli, wykorzystano kolejno 256, 384, 512, 640, 768, 896 i 1024 pierwszych par pomiarowych. Wynik eksperymentu przedstawiono na rys. 4.3 i 4.4, natomiast na rys. 4.5 zilustrowano sygnał wyjściowy prawdziwego systemu oraz modeli uzyskanych za pomocą obu metod, przy wykorzystaniu wszystkich $N = 1024$

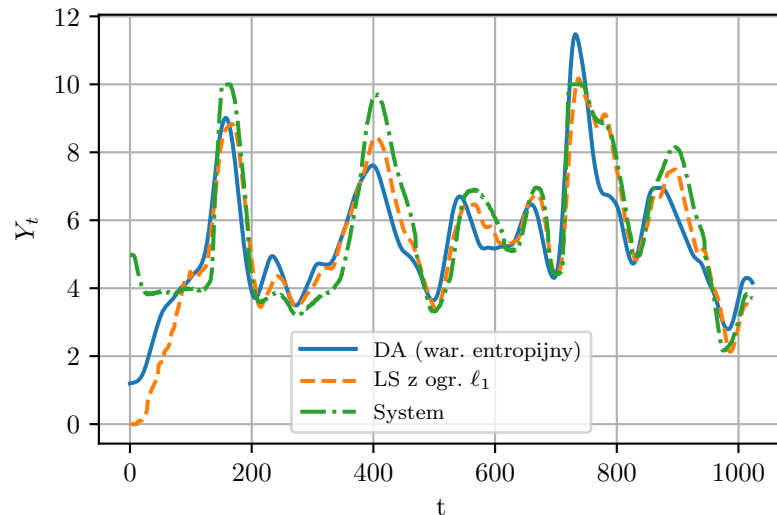


Rysunek 4.4: Eksperyment II: czas estymacji modelu (wykonywania algorytmu) w zależności od liczby pomiarów

pomiarów wejścia-wyjścia. Jak można zauważyć na wykresie 4.3, błąd metody opartej o algorytm DA jest widocznie wyższy od błędu metody opartej o agregację wypukłą. Powodu tej różnicy należy upatrywać w szybkości zbieżności algorytmów, która w przypadku algorytmu DA jest nieco wolniejsza (zob. par. 4.2.3.2). Należy przy tym pamiętać, że w przypadku algorytmu DA, ze względu na jawną formę rozwiązań pośrednich i fakt, że w każdej iteracji wykorzystywany jest jedynie gradient funkcji straty, złożoność obliczeniowa metody jest niewielka. Natomiast w przypadku korzystania z zewnętrznych *solverów*, przeznaczonych do rozwiązywania szerokiej klasy problemów optymalizacyjnych, trzeba liczyć się z dużo większym kosztem obliczeniowym, co obrazuje wykres 4.4.

4.5 Podsumowanie

Przedstawiona w rozdziale metoda modelowania oparta jest o algorytm optymalizacji wypukłej zaproponowany przez Nesterowa [78]. Jak pokazaliśmy w eksperymencie numerycznym, w wersji podstawowej, tj. z deterministycznym doбором kroku, bez dalszego strojenia stałej G błąd metody jest duży. Niedogodność ta wynika pośrednio z natury kwadratowej funkcji straty, gdyż wartość stałej G , która wyraża pesymistyczne, górne oszacowania gradientu tej funkcji w *każdej* iteracji, jest w postawionym problemie bardzo wysoka – a to z kolei pociąga za sobą niewielką długość kroku algorytmu. W związku z powyższym, mimo niepełnego uzasadnienia teoretycznego, proponuje się adaptacyjny sposób doboru kroku, który dostosowuje jego długość do obserwowanych, losowych wartości gradientu, co znacznie ogranicza, a asymptotycznie wręcz niweluje do zera wpływ wartości G na długość kroku



Rysunek 4.5: Eksperyment II: wyjście prawdziwego systemu oraz modeli uzyskanych dla $N = 1024$ (zbiór danych walidujących)

algorytmu.

W kontekście podejścia do modelowania systemów opartego o czarną skrzynkę, niewątpliwą zaletą zademonstrowanego algorytmu jest fakt, iż wpływ wielkości słownika na błąd estymacji jest pomijalnie mały, dzięki czemu niewielkim kosztem jesteśmy w stanie ograniczyć błąd aproksymacji modelu. Chociaż wyniki eksperymentu wykazały, że dla tego samego słownika błąd metody wykorzystującej entropijny wariant algorytmu DA jest nieco większy od błędu metody najmniejszych kwadratów z ograniczeniami ℓ_1 , to posiada ona kilka atrakcyjnych własności, które w pewnych zastosowaniach mogą przemawiać na jej korzyść:

- Po pierwsze, algorytm ten ma zagwarantowaną złożoność obliczeniową, która wynosi $\mathcal{O}(ND)$. Z drugiej strony, złożoność metod opartych o zewnętrzne biblioteki, takie jak CVXPY, pozostaje często nieznana użytkownikowi i zależy od ustawień programu, takich jak wartość marginesu błędu algorytmu.
- Złożoność pamięciowa algorytmu jest niezależna od liczby pomiarów N , dzięki czemu metoda może być stosowana także na dużych zbiorach danych.
- Rozwiązanie problemu optymalizacyjnego jest jawne i składa się z podstawowych operacji matematycznych, co czyni go łatwym w implementacji.
- Metoda może być rozpatrywana jako algorytm rekurencyjny, w którym koszt obliczeniowy dołączenia nowych obserwacji jest stały, co pozwala np. na zastosowanie metody w systemach czasu rzeczywistego.

Dodatki

4.A Dowód twierdzenia 4.1

Dowód. Korzystając z wypukłości funkcji $\hat{L}_t(\cdot)$ otrzymujemy nierówność

$$\hat{L}_t(\boldsymbol{\theta}_t) - \hat{L}_t(\mathbf{u}) \leq \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \mathbf{u} \rangle = \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle + \langle \mathbf{g}_t, \boldsymbol{\theta}_1 - \mathbf{u} \rangle,$$

która pozwala nam ograniczyć od góry sumę empirycznych strat (wzór (4.7)), względem dowolnego wektora $\mathbf{u} \in \mathcal{C}$, przez

$$\begin{aligned} C_T(\mathbf{u}) &\leq \sum_{t=1}^T (\langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle + \langle \mathbf{g}_t, \boldsymbol{\theta}_1 - \mathbf{u} \rangle) \\ &\leq \sum_{t=1}^T \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle + \sup_{\mathbf{w} \in \mathcal{C}} (\langle \mathbf{g}_{1:T}, \boldsymbol{\theta}_1 - \mathbf{w} \rangle) \\ &\leq \sum_{t=1}^T \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle + \sup_{\mathbf{w} \in \mathcal{C}} (\langle \mathbf{g}_{1:T}, \boldsymbol{\theta}_1 - \mathbf{w} \rangle + \eta_T^{-1} (R^2 - h(\mathbf{w}))) \\ &\leq \sum_{t=2}^T \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle + \underbrace{\sup_{\mathbf{w} \in \mathcal{C}} (\langle \mathbf{g}_{1:T}, \boldsymbol{\theta}_1 - \mathbf{w} \rangle - \eta_T^{-1} h(\mathbf{w}))}_{=V_T(-\mathbf{g}_{1:T})} + \eta_T^{-1} R^2, \end{aligned} \quad (4.31)$$

gdzie $V_t : \mathbb{R}^D \rightarrow \mathbb{R}$ jest funkcją pomocniczą, określoną przez wzór

$$V_t(\mathbf{w}) = f_t^*(\mathbf{w}) - \langle \mathbf{w}, \boldsymbol{\theta}_1 \rangle \quad (4.32)$$

i gdzie funkcja $f_t^*(\cdot)$ została zdefiniowana wcześniej, we wzorze (4.4). Podamy teraz kilka własności funkcji $V_t(\cdot)$, kluczowych z punktu widzenia dalszej części dowodu.

Po pierwsze, korzystając ze wzoru (4.6) i faktu, że gradient funkcji $V_t(\cdot)$ wynosi $\nabla V_t(\cdot) = \nabla f_t^*(\cdot) - \boldsymbol{\theta}_1$, otrzymujemy tożsamość

$$\nabla V_t(-\mathbf{g}_{1:t}) = \boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_1, \quad (4.33)$$

która jest spełniona dla dowolnego $t \geq 1$. W szczególnym przypadku, gdy $\mathbf{g}_{1:t} = 0$, wartość gradientu wynosi

$$\begin{aligned} \nabla V_t(0) &= \arg \min_{\mathbf{u} \in \mathcal{C}} h_t(\mathbf{u}) - \boldsymbol{\theta}_1 = \arg \min_{\mathbf{u} \in \mathcal{C}} h(\mathbf{u}) - \boldsymbol{\theta}_1 \\ &= 0, \end{aligned} \quad (4.34)$$

co wynika ze wzoru na rozwiązanie początkowe algorytmu, (4.1).

Przypomnijmy teraz, że zgodnie z założeniem algorytmu, funkcja $h(\cdot)$ jest silnie wypukła ze stałą 1 względem normy $\|\cdot\|$. Wobec tego, funkcja $f_t(\cdot)$ (wzór (4.5)) jest η_t^{-1} -silnie wypukła, tak więc, zastosowanie lematu 4.4 (dodatek 4.E) prowadzi do nierówności

$$f_t^*(\mathbf{x} + \boldsymbol{\delta}) - f_t^*(\mathbf{x}) \leq \langle \nabla f_t^*(\mathbf{x}), \boldsymbol{\delta} \rangle + \frac{\eta_t}{2} \|\boldsymbol{\delta}\|_*^2,$$

która jest spełniona dla dowolnych $\mathbf{x}, \boldsymbol{\delta} \in \mathbb{R}^D$. Po odjęciu od obu stron wyrażenia $\langle \boldsymbol{\delta}, \boldsymbol{\theta}_1 \rangle$ otrzymujemy

$$f_t^*(\mathbf{x} + \boldsymbol{\delta}) - \langle \mathbf{x} + \boldsymbol{\delta}, \boldsymbol{\theta}_1 \rangle - (f_t^*(\mathbf{x}) - \langle \mathbf{x}, \boldsymbol{\theta}_1 \rangle) \leq \langle \nabla f_t^*(\mathbf{x}), \boldsymbol{\delta} \rangle - \langle \boldsymbol{\delta}, \boldsymbol{\theta}_1 \rangle + \frac{\eta_t}{2} \|\boldsymbol{\delta}\|_*^2,$$

co, ze względu na przemienność działania $\langle \cdot, \cdot \rangle$, prowadzi do następującej nierówności

$$V_t(\mathbf{x} + \boldsymbol{\delta}) \leq V_t(\mathbf{x}) + \langle \nabla V_t(\mathbf{x}), \boldsymbol{\delta} \rangle + \frac{\eta_t}{2} \|\boldsymbol{\delta}\|_*^2. \quad (4.35)$$

Możemy teraz przejść do głównej części dowodu. Zauważmy, że dla $t \geq 2$ zachodzi następujące oszacowanie

$$\begin{aligned} V_t(-\mathbf{g}_{1:t}) &\stackrel{(i)}{\leq} V_{t-1}(-\mathbf{g}_{1:t}) = V_{t-1}(-\mathbf{g}_{1:t-1} - \mathbf{g}_t) \\ &\stackrel{(ii)}{\leq} V_{t-1}(-\mathbf{g}_{1:t-1}) + \langle \nabla V_{t-1}(-\mathbf{g}_{1:t-1}), -\mathbf{g}_t \rangle + \frac{1}{2} \eta_{t-1} \|\mathbf{g}_t\|_*^2 \\ &\stackrel{(iii)}{=} V_{t-1}(-\mathbf{g}_{1:t-1}) + \langle \boldsymbol{\theta}_t - \boldsymbol{\theta}_1, -\mathbf{g}_t \rangle + \frac{1}{2} \eta_{t-1} \|\mathbf{g}_t\|_*^2, \end{aligned} \quad (4.36)$$

gdzie (i) wynika z założenia o krokach $\{\eta_t\}$ oraz równania (4.32), (ii) jest rezultatem zastosowania nierówności (4.35), a (iii) wynika ze wzoru (4.33). Z kolei, dla $t = 1$ otrzymujemy następujące górne oszacowanie funkcji pomocniczej,

$$\begin{aligned} V_1(-\mathbf{g}_1) &\leq V_1(0) + \langle \nabla V_1(0), -\mathbf{g}_1 \rangle + \frac{1}{2} \eta_1 \|\mathbf{g}_1\|_*^2 \\ &= \frac{1}{2} \eta_1 \|\mathbf{g}_1\|_*^2, \end{aligned} \quad (4.37)$$

gdzie ostatnia równość zachodzi na podstawie wzoru (4.34) i faktu, że $V_1(0) = \sup_{\mathbf{x}} (-f_t(\mathbf{x})) = \inf_{\mathbf{x} \in \mathcal{C}} h(\mathbf{u}) = 0$. Ze wzoru (4.36), otrzymujemy

$$\langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle \leq V_{t-1}(-\mathbf{g}_{1:t-1}) - V_t(-\mathbf{g}_{1:t}) + \frac{1}{2} \eta_{t-1} \|\mathbf{g}_t\|_*^2,$$

co, w połączeniu ze wzorem (4.37), pozwala nam oszacować następującą sumę

$$\begin{aligned}
\sum_{t=2}^T \langle \mathbf{g}_t, \boldsymbol{\theta}_t - \boldsymbol{\theta}_1 \rangle &\leq \sum_{t=2}^T \left(V_{t-1}(-\mathbf{g}_{1:t-1}) - V_t(-\mathbf{g}_{1:t}) + \frac{1}{2} \eta_{t-1} \|\mathbf{g}_t\|_*^2 \right) \\
&= V_1(-\mathbf{g}_1) - V_T(-\mathbf{g}_{1:T}) + \frac{1}{2} \sum_{t=2}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 \\
&\leq \frac{1}{2} \eta_1 \|\mathbf{g}_1\|_*^2 - V_T(-\mathbf{g}_{1:T}) + \frac{1}{2} \sum_{t=2}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2 \\
&= -V_T(-\mathbf{g}_{1:T}) + \frac{1}{2} \sum_{t=1}^T \eta_{t-1} \|\mathbf{g}_t\|_*^2.
\end{aligned}$$

Wstawienie powyższego rezultatu do wzoru (4.31) kończy dowód. \square

4.B Dowód twierdzenia 4.2

Dowód. Zaczniemy od przedstawienia implikacji jakie niesie ze sobą założenie o ograniczonej wartości oczekiwanej kwadratu gradientu empirycznej funkcji straty. Po pierwsze, dla dowolnego $\boldsymbol{\theta} \in \mathcal{C}$ prawdziwa jest nierówność

$$\mathbb{E} \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_*^2 = \mathbb{E} \{ \mathbb{E} \{ \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_*^2 \mid \mathcal{F}_{t-1} \} \} \leq G^2,$$

w szczególności więc, zachodzi

$$\mathbb{E} \|\mathbf{g}_t\|_*^2 \leq G^2. \quad (4.38)$$

Po drugie, zgodnie z nierównością $\mathbb{E} |X| \leq \sqrt{\mathbb{E} \{X^2\}}$, z prawdopodobieństwem równym 1 zachodzi

$$\mathbb{E} \{ \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_* \mid \mathcal{F}_{t-1} \} < G. \quad (4.39)$$

Po trzecie wreszcie, zauważmy, że dla ustalonego $\boldsymbol{\theta}$, gradient chwilowej empirycznej funkcji straty stanowi nieobciążony estymator gradientu prawdziwej funkcji straty, tak więc $\nabla L_S(\boldsymbol{\theta}) = \mathbb{E} \{ \nabla \hat{L}_t(\boldsymbol{\theta}) \}$. Biorąc pod uwagę fakt, iż dla $p \geq 1$ norma $\|\cdot\|_p$ jest wypukłą, otrzymujemy zatem następującą nierówność

$$\|\nabla L_S(\boldsymbol{\theta})\|_* = \|\mathbb{E} \{ \nabla \hat{L}_t(\boldsymbol{\theta}) \}\|_* \leq \mathbb{E} \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_* = \mathbb{E} \{ \mathbb{E} \{ \|\nabla \hat{L}_t(\boldsymbol{\theta})\|_* \mid \mathcal{F}_{t-1} \} \} < G. \quad (4.40)$$

Przystąpimy teraz do właściwej części dowodu, częściowo korzystając z techniki dowodowej zaprezentowanej w pracy [26], która opiera się na następującej tożsamo-

ści,

$$\begin{aligned}
\sum_{t=1}^T \mathbb{E}\{L_S(\boldsymbol{\theta}_t) - L_S(\boldsymbol{\theta}^*)\} &= \sum_{t=1}^{T-M+1} \mathbb{E}\{L_S(\boldsymbol{\theta}_t) - \hat{L}_{t+M-1}(\boldsymbol{\theta}_t) + \hat{L}_{t+M-1}(\boldsymbol{\theta}^*) - L_S(\boldsymbol{\theta}^*)\} \\
&+ \sum_{t=1}^{T-M+1} \mathbb{E}\{\hat{L}_{t+M-1}(\boldsymbol{\theta}_t) - \hat{L}_{t+M-1}(\boldsymbol{\theta}_{t+M-1})\} \\
&+ \mathbb{E}\left\{\sum_{t=M}^T (\hat{L}_t(\boldsymbol{\theta}_t) - \hat{L}_t(\boldsymbol{\theta}^*))\right\} \\
&+ \sum_{t=T-M+2}^T \mathbb{E}\{L_S(\boldsymbol{\theta}_t) - L_S(\boldsymbol{\theta}^*)\}. \tag{4.41}
\end{aligned}$$

Fakt, że system i model mają długość pamięci nie większą niż M powoduje, że zarówno wyjście systemu Y_{t+M-1} , jak i wartości funkcjonałów stanowiących elementy słownika, $\varphi(\mathbf{X}_{t+M-1})$ są niezależne od \mathcal{F}_{t-1} , a więc także od wektora $\boldsymbol{\theta}_t$. W konsekwencji, zachodzi

$$\mathbb{E}\{\hat{L}_{t+M-1}(\boldsymbol{\theta}_t)\} = \mathbb{E}\{\mathbb{E}\{\hat{L}_{t+M-1}(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\}\} = \mathbb{E}\{L_S(\boldsymbol{\theta}_t)\}.$$

Uwzględnienie tożsamości $\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}^*)\} = L(\boldsymbol{\theta}^*)$ pozwala stwierdzić, że

$$\mathbb{E}\{L_S(\boldsymbol{\theta}_t) - \hat{L}_{t+M-1}(\boldsymbol{\theta}_t) + \hat{L}_{t+M-1}(\boldsymbol{\theta}^*) - L_S(\boldsymbol{\theta}^*)\} = 0,$$

a więc pierwsza suma po prawej stronie wzoru (4.41) jest równa zero.

W celu oszacowania drugiej sumy, korzystamy z lematu 4.3 (którego założenie o ograniczonym gradiencie jest spełnione za sprawą nierówności (4.39)) i w ten sposób otrzymujemy nierówność

$$\begin{aligned}
\sum_{t=1}^{T-M+1} \mathbb{E}\{\hat{L}_{t+M-1}(\boldsymbol{\theta}_t) - \hat{L}_{t+M-1}(\boldsymbol{\theta}_{t+M-1})\} &\leq G(M-1) \sum_{t=1}^{T-M+1} \eta_{t+M-2} + \\
&+ G \sum_{t=2}^{T-M+1} (t-1)(\eta_{t-1} - \eta_{t+M-2}) \\
&= G(M-1) \sum_{t=1}^{T-M+1} \eta_{t+M-2} + \\
&+ G \sum_{t=1}^{T-M} t(\eta_t - \eta_{t+M-1}).
\end{aligned}$$

Trzeci składnik szacujemy za pomocą twierdzenia 4.1,

$$\mathbb{E}\left\{\sum_{t=M}^T (\hat{L}_t(\boldsymbol{\theta}_t) - \hat{L}_t(\boldsymbol{\theta}^*))\right\} \leq \mathbb{E}\{C_T(\boldsymbol{\theta}^*)\} \leq \frac{G^2}{2} \sum_{t=1}^T \eta_{t-1} + \eta_T^{-1} R^2,$$

wykorzystując przy tym nierówność (4.38). Wreszcie, aby oszacować czwarty składnik, zauważmy, że na podstawie nierówności (4.40) i lematu 4.5, funkcja $L(\cdot)$ jest

G -lipschitzowska względem normy $\|\cdot\|$ (patrz def. 4.4, dodatek 4.E), przez co poszczególne elementy sumy możemy ograniczyć w następujący sposób

$$\sum_{t=T-M+2}^T (L(\boldsymbol{\theta}_t) - L(\boldsymbol{\theta}^*)) \leq G \sum_{t=T-M+2}^T \|\boldsymbol{\theta}_t - \boldsymbol{\theta}^*\| \leq 2(M-1)GK,$$

gdzie $K = \max_{\boldsymbol{\theta} \in C} \|\boldsymbol{\theta}\|$. Ostatnim krokiem dowodu jest podstawienie uzyskanych oszacowań do wzoru (4.41). \square

4.C Dowód twierdzenia 4.3

Dowód. Chcemy wykazać, że dla dużych τ wartość wyrażenia $\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t)\}$ jest bliska $\mathbb{E}\{L_S(\boldsymbol{\theta}_t)\}$. Zaczniemy od przedstawienia następującej tożsamości

$$\begin{aligned} \mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\} &= \mathbb{E}\{Y_{t+\tau}^2 - 2Y_{t+\tau}m[\mathbf{X}_{n+\tau}; \boldsymbol{\theta}_t] + m^2[\mathbf{X}_{n+\tau}] \mid \mathcal{F}_{t-1}\} \\ &= \mathbb{E}\{Y_{t+\tau}^2 - 2m_{\text{sys}}[\mathbf{X}_{n+\tau}]m[\mathbf{X}_{n+\tau}; \boldsymbol{\theta}_t] + m^2[\mathbf{X}_{n+\tau}] \mid \mathcal{F}_{t-1}\}, \end{aligned} \quad (4.42)$$

która wynika z założenia o szumie losowym systemu. Przypomnijmy, że $\mathbf{X}_n = [X_n, X_{n-1}, \dots]$, $\mathbf{X}_n^{[\tau]} = [X_n, X_{n-1}, \dots, X_{n-\tau}, 0, 0, \dots]$ i wprowadźmy wektor losowy

$$\tilde{\mathbf{X}}_{n,\tau} = [X_n, X_{n-1}, \dots, X_{n-\tau}, \tilde{X}_{n-\tau-1}, \tilde{X}_{n-\tau-2}, \dots],$$

gdzie $\tilde{X}_{n-\tau-1}, \tilde{X}_{n-\tau-2}, \dots$ są ciągiem *i.i.d.* zmiennych losowych niezależnych od \mathcal{F}_{t-1} i o takim samym rozkładzie, co wejście systemu X_n . Zgodnie z założeniem o pamięci zanikającej, zachodzą następujące nierówności

$$\left| m_{\text{sys}}[\mathbf{X}_n] - m_{\text{sys}}[\mathbf{X}_n^{[\tau]}] \right| \leq c_0 e^{-c_1 \tau}$$

oraz

$$\left| m_{\text{sys}}[\tilde{\mathbf{X}}_{n,\tau}] - m_{\text{sys}}[\mathbf{X}_n^{[\tau]}] \right| \leq c_0 e^{-c_1 \tau},$$

wobec czego, z nierówności trójkąta wynika następująca zależność

$$\left| m_{\text{sys}}[\mathbf{X}_n] - m_{\text{sys}}[\tilde{\mathbf{X}}_{n,\tau}] \right| \leq 2c_0 e^{-c_1 \tau}.$$

Zauważmy teraz, że dla każdego a zachodzi nierówność

$$a(m_{\text{sys}}[\mathbf{X}_n] - m_{\text{sys}}[\tilde{\mathbf{X}}_{n,\tau}]) \leq |a| \left| m_{\text{sys}}[\mathbf{X}_n] - m_{\text{sys}}[\tilde{\mathbf{X}}_{n,\tau}] \right| \leq |a| 2c_0 e^{-c_1 \tau}$$

i w konsekwencji

$$am_{\text{sys}}[\mathbf{X}_n] \leq am_{\text{sys}}[\tilde{\mathbf{X}}_{n,\tau}] + |a| 2c_0 e^{-c_1 \tau}. \quad (4.43)$$

Na podstawie powyższego wzoru i założenia o ograniczonym wyjściu modelu, otrzymujemy nierówność

$$\begin{aligned} \mathbb{E}\{-2m_{\text{sys}}[\mathbf{X}_{n+\tau}]m[\mathbf{X}_{n+\tau}; \boldsymbol{\theta}_t] \mid \mathcal{F}_{t-1}\} &\leq \mathbb{E}\{-2m_{\text{sys}}[\tilde{\mathbf{X}}_{n+\tau,\tau}]m[\mathbf{X}_{n+\tau}; \boldsymbol{\theta}_t] \mid \mathcal{F}_{t-1}\} \\ &\quad + 4\bar{m}c_0 e^{-c_1 \tau}. \end{aligned} \quad (4.44)$$

Ze względu na skończoną pamięć i ograniczone wyjście modelu, istnieje pewna stała $c'_0 < \infty$, taka że

$$\left| m[\mathbf{x}_n] - m[\mathbf{x}_n^{[\tau]}] \right| \leq c'_0 e^{-c_1 \tau},$$

a więc podobnie jak we wzorze (4.43), dla charakterystyki modelu zachodzi nierówność

$$am[\mathbf{X}_n] \leq am[\tilde{\mathbf{X}}_{n,\tau}] + |a| 2c'_0 e^{-c_1 \tau},$$

która w połączeniu z założeniem o ograniczonym wyjściu systemu, pozwala nam doprowadzić nierówność (4.44) do postaci

$$\begin{aligned} \mathbb{E}\{-2m_{\text{sys}}[\mathbf{X}_{n+\tau}]m[\mathbf{X}_{n+\tau}; \boldsymbol{\theta}_t] \mid \mathcal{F}_{t-1}\} &\leq \mathbb{E}\{-2m_{\text{sys}}[\tilde{\mathbf{X}}_{n+\tau,\tau}]m[\tilde{\mathbf{X}}_{n+\tau,\tau}; \boldsymbol{\theta}_t] \mid \mathcal{F}_{t-1}\} \\ &\quad + Ce^{-c_1 \tau}, \end{aligned}$$

gdzie $C = 4(\bar{m}_{\text{sys}}c'_0 + \bar{m}c_0)$. Wstawiając powyższe oszacowanie do wzoru (4.42), otrzymujemy

$$\begin{aligned} \mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) \mid \mathcal{F}_{t-1}\} &\leq \mathbb{E}\{Y_{t+\tau}^2 - 2m_{\text{sys}}[\tilde{\mathbf{X}}_{n+\tau,\tau}]m[\tilde{\mathbf{X}}_{n+\tau,\tau}; \boldsymbol{\theta}_t] + m^2[\mathbf{X}_{n+\tau}] \mid \mathcal{F}_{t-1}\} \\ &\quad + Ce^{-c_1 \tau}. \end{aligned}$$

Zauważmy, że wektor $\tilde{\mathbf{X}}_{n,\tau}$ jest niezależny od \mathcal{F}_{t-1} (a więc także od $\boldsymbol{\theta}_t$) oraz że ma ten sam rozkład co wektor $\mathbf{X}_{n+\tau}$. W konsekwencji, nałożenie operatora wartości oczekiwanej na obie strony nierówności prowadzi do następującego rezultatu

$$\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t)\} \leq \mathbb{E}\{L_S(\boldsymbol{\theta}_t)\} + Ce^{-c_1 \tau}.$$

Pozostała część dowodu jest identyczna z dowodem twierdzenia 4.2. \square

4.D Lematy techniczne

Zaprezentowane poniżej lematy zostały wykorzystane w dowodzie twierdzenia 4.2.

Lemat 4.2. *Dla każdego $t \geq 2, \tau \geq 1$, rozwiązania pośrednie algorytmu DA spełniają nierówność*

$$\|\boldsymbol{\theta}_{t+\tau} - \boldsymbol{\theta}_t\| \leq \eta_{t+\tau-1} \|\mathbf{g}_{t:t+\tau-1}\|_* + (\eta_{t-1} - \eta_{t+\tau-1}) \|\mathbf{g}_{1:t-1}\|_*,$$

natomiast dla $t = 1, \tau \geq 1$ zachodzi nierówność

$$\|\boldsymbol{\theta}_{1+\tau} - \boldsymbol{\theta}_1\| = \|\boldsymbol{\theta}_{1+\tau} - \boldsymbol{\theta}_1\| \leq \eta_\tau \|\mathbf{g}_{1:\tau}\|_*.$$

Dowód. Gradient funkcji celu problemu optymalizacyjnego (4.2) dany jest przez

$$\nabla_{\boldsymbol{\theta}} \left\{ \langle \mathbf{g}_{1:t}, \boldsymbol{\theta} \rangle + \eta_t^{-1} h(\boldsymbol{\theta}) \right\} = \mathbf{g}_{1:t} + \eta_t^{-1} \nabla h(\boldsymbol{\theta}).$$

Warunek optymalności pierwszego rzędu implikuje, że dla dowolnego $\mathbf{x} \in \mathcal{C}$ zachodzi

$$\langle \eta_t \mathbf{g}_{1:t} + \nabla h(\boldsymbol{\theta}_{t+1}), \boldsymbol{\theta}_{t+1} - \mathbf{x} \rangle \leq 0, \quad (4.45)$$

w szczególności więc, spełniona jest nierówność

$$\langle \eta_{t+\tau-1} \mathbf{g}_{1:t+\tau-1} + \nabla h(\boldsymbol{\theta}_{t+\tau}), \boldsymbol{\theta}_{t+\tau} - \boldsymbol{\theta}_t \rangle \leq 0$$

oraz

$$\langle \eta_{t-1} \mathbf{g}_{1:t-1} + \nabla h(\boldsymbol{\theta}_t), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \leq 0.$$

Sumując obie nierówności stronami i przenosząc pierwszą z nich na drugą stronę otrzymujemy

$$\langle \eta_{t-1} \mathbf{g}_{1:t-1} + \nabla h(\boldsymbol{\theta}_t), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \leq \langle \eta_{t+\tau-1} \mathbf{g}_{1:t+\tau-1} + \nabla h(\boldsymbol{\theta}_{t+\tau}), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle.$$

W kolejnym kroku odejmujemy od obu stron nierówności wyrażenie

$$\langle \eta_{t-1} \mathbf{g}_{1:t-1} + \nabla h(\boldsymbol{\theta}_{t+\tau}), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle,$$

by otrzymać

$$\begin{aligned} \langle \nabla h(\boldsymbol{\theta}_t) - \nabla h(\boldsymbol{\theta}_{t+\tau}), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle &\leq \langle \eta_{t+\tau-1} \mathbf{g}_{1:t+\tau-1} - \eta_{t-1} \mathbf{g}_{1:t-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \\ &= \eta_{t+\tau-1} \left\langle \mathbf{g}_{1:t+\tau-1} - \frac{\eta_{t-1}}{\eta_{t+\tau-1}} \mathbf{g}_{1:t-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \right\rangle \\ &= \eta_{t+\tau-1} \left\langle \mathbf{g}_{t:t+\tau-1} - \frac{\eta_{t-1} - \eta_{t+\tau-1}}{\eta_{t+\tau-1}} \mathbf{g}_{1:t-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \right\rangle \\ &= \eta_{t+\tau-1} \langle \mathbf{g}_{t:t+\tau-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \\ &\quad + (\eta_{t-1} - \eta_{t+\tau-1}) \langle -\mathbf{g}_{1:t-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle. \end{aligned} \quad (4.46)$$

Następnie, korzystając ze stwierdzenia 4.1 oraz faktu, że funkcja $h(\cdot)$ jest silnie wypukła, otrzymujemy nierówność

$$\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|^2 \leq \langle \nabla h(\boldsymbol{\theta}_t) - \nabla h(\boldsymbol{\theta}_{t+\tau}), \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle. \quad (4.47)$$

Z nierówności Höldera otrzymujemy górne ograniczenia iloczynów skalarnych wyrażenia (4.46),

$$\langle \mathbf{g}_{t:t+\tau-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \leq \|\mathbf{g}_{t:t+\tau-1}\|_* \|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|$$

oraz

$$\langle -\mathbf{g}_{1:t-1}, \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau} \rangle \leq \|\mathbf{g}_{1:t-1}\|_* \|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|.$$

Po wstawieniu uzyskanych oszacowań do wzoru (4.46) i wykorzystaniu nierówności (4.47), otrzymujemy

$$\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|^2 \leq \eta_{t+\tau-1} \|\mathbf{g}_{t:t+\tau-1}\|_* \|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\| + (\eta_{t-1} - \eta_{t+\tau-1}) \|\mathbf{g}_{1:t-1}\|_* \|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|.$$

Podzielenie obu stron przez $\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|$ kończy dowód pierwszej nierówności

W celu udowodnienia drugiej nierówności, zauważmy że z warunku optymalności pierwszego rzędu otrzymujemy nierówność

$$\langle \nabla h(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1 - \boldsymbol{x} \rangle \leq 0,$$

która jest spełniona dla każdego $\boldsymbol{x} \in \mathcal{C}$. W szczególności więc, otrzymujemy

$$\langle \nabla h(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1 - \boldsymbol{\theta}_{\tau+1} \rangle \leq 0.$$

Oprócz tego, korzystając z wzoru (4.45), otrzymujemy

$$\langle \eta_\tau \mathbf{g}_{1:\tau} + \nabla h(\boldsymbol{\theta}_{\tau+1}), \boldsymbol{\theta}_{\tau+1} - \boldsymbol{\theta}_1 \rangle \leq 0.$$

Postępując jak poprzednio doprowadzamy do nierówności

$$\langle \nabla h(\boldsymbol{\theta}_1) - \nabla h(\boldsymbol{\theta}_{\tau+1}), \boldsymbol{\theta}_1 - \boldsymbol{\theta}_{\tau+1} \rangle \leq \langle -\eta_\tau \mathbf{g}_{1:\tau}, \boldsymbol{\theta}_{\tau+1} - \boldsymbol{\theta}_1 \rangle,$$

Zastosowanie nierówności Höldera oraz stwierdzenia 4.1 kończy dowód. \square

Lemat 4.3. Niech $\{\eta_t\}$ będzie ciągiem niemalejących kroków, których długość jest ustalona w sposób deterministyczny. Załóżmy że dla każdego $\boldsymbol{\theta} \in \mathcal{C}$, gdzie \mathcal{C} to zbiór ograniczeń algorytmu DA, prawie na pewno zachodzi nierówność $\mathbb{E}\{\|\nabla \hat{L}_t(\boldsymbol{\theta})\|_* \mid \mathcal{F}_{t-1}\} \leq G$. Wtedy, dla każdego $t \geq 2, \tau \geq 1$ zachodzi

$$\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} < \tau \eta_{t+\tau-1} G + (t-1)(\eta_{t-1} - \eta_{t+\tau-1})G,$$

natomiast dla $t = 1, \tau \geq 1$ spełniona jest nierówność

$$\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} = \mathbb{E}\{\hat{L}_{1+\tau}(\boldsymbol{\theta}_1) - \hat{L}_{1+\tau}(\boldsymbol{\theta}_{1+\tau})\} < \tau \eta_\tau G.$$

Dowód. Niech $\hat{l}_t(\boldsymbol{\theta}) = \mathbb{E}\{\hat{L}_t(\boldsymbol{\theta}) \mid \mathcal{F}_{t-1}\}$. Zaczniemy od wykazania, że funkcja ta jest lipschitzowska. Dla argumentów mierzalnych względem \mathcal{F}_{t-1} możemy zastosować regułę Leibniza, by gradient funkcji $\hat{l}(\cdot)$ przedstawić jako $\nabla \hat{l}_t(\boldsymbol{\theta}) = \mathbb{E}\{\nabla \hat{L}_t(\boldsymbol{\theta}) \mid \mathcal{F}_{t-1}\}$. Z faktu, że dla $p \geq 1$ norma $\|\cdot\|_p$ jest wypukła, wynika że

$$\|\nabla \hat{l}_t(\boldsymbol{\theta})\|_* \leq \mathbb{E}\{\|\nabla \hat{L}_t(\boldsymbol{\theta})\|_* \mid \mathcal{F}_{t-1}\},$$

a więc, zgodnie z założeniem lematu, zachodzi

$$\|\nabla \hat{l}_t(\boldsymbol{\theta})\|_* \leq G.$$

Tym samym, zgodnie z lematem 4.5, funkcja $\hat{l}(\cdot)$ jest G -lipschitzowska względem normy $\|\cdot\|$. Powyższy wniosek prowadzi do następującej nierówności

$$\begin{aligned} \mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} &= \mathbb{E}\{\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau}) \mid \mathcal{F}_{t+\tau-1}\}\} \\ &= \mathbb{E}\{\hat{l}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{l}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} \\ &\leq G \mathbb{E}\{\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t+\tau}\|\}. \end{aligned}$$

Po podstawieniu wyniku z lematu 4.2, dla $t \geq 2$ otrzymujemy

$$\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} \leq \eta_{t+\tau-1} \mathbb{E} \|\mathbf{g}_{t:t+\tau-1}\|_* + (\eta_{t-1} - \eta_{t+\tau-1}) \mathbb{E} \|\mathbf{g}_{1:t-1}\|_*,$$

z kolei dla $t = 1$

$$\mathbb{E}\{\hat{L}_{t+\tau}(\boldsymbol{\theta}_t) - \hat{L}_{t+\tau}(\boldsymbol{\theta}_{t+\tau})\} \leq \eta_\tau \mathbb{E} \|\mathbf{g}_{1:\tau}\|_*.$$

Zastosowanie nierówności $\mathbb{E} \|\mathbf{g}_{a:b}\|_* \leq \mathbb{E} \sum_{i=a}^b \|\mathbf{g}_i\|_* \leq (b - a + 1)G$, która jest spełniona dla dowolnych $a \leq b$ kończy dowód dla obu przypadków. \square

4.E Podstawowe fakty i definicje z optymalizacji wypukłej wykorzystane w rozdziale

Definicja 4.2 (Dziedzina efektywna). Niech $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$. Efektywną dziedziną funkcji $f(\cdot)$ nazywamy zbiór

$$\text{dom}(f) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < \infty\}.$$

Definicja 4.3 (Funkcja silnie wypukła). Funkcja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ jest silnie wypukła ze stałą α względem normy $\|\cdot\|$ na zbiorze wypukłym $\mathcal{C} \subseteq \text{dom}(f)$ jeśli jest różniczkowalna na \mathcal{C} i dla każdego $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ spełniona jest nierówność

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\alpha}{2} \|\mathbf{y} - \mathbf{x}\|^2.$$

Stwierdzenie 4.1. Niech funkcja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ będzie silnie wypukła na zbiorze \mathcal{C} . Wtedy, dla każdego $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ spełniona jest nierówność

$$\langle \nabla f(\mathbf{y}) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \alpha \|\mathbf{y} - \mathbf{x}\|^2.$$

Dowód. Z faktu, że $f(\cdot)$ jest silnie wypukła wynika nierówność

$$f(\mathbf{x}) - f(\mathbf{y}) + \langle \nabla f(\mathbf{y}), \mathbf{y} - \mathbf{x} \rangle \geq \frac{\alpha}{2} \|\mathbf{y} - \mathbf{x}\|^2$$

oraz

$$f(\mathbf{y}) - f(\mathbf{x}) + \langle -\nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \frac{\alpha}{2} \|\mathbf{y} - \mathbf{x}\|^2.$$

Dodanie powyższych nierówności stronami kończy dowód. \square

Lemat 4.4. Niech $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ będzie funkcją silnie wypukłą ze stałą α oraz niech $\text{dom}(f)$ będzie zbiorem niepustym. Wtedy dla każdego $\mathbf{x}, \boldsymbol{\delta} \in \mathbb{R}^n$ zachodzi

$$f^*(\mathbf{x} + \boldsymbol{\delta}) - f^*(\mathbf{x}) \leq \langle \nabla f^*(\mathbf{x}), \boldsymbol{\delta} \rangle + \frac{\alpha^{-1}}{2} \|\boldsymbol{\delta}\|_*^2.$$

Dowód. Zob. np. [98, lem. 15]. Dowód bardziej ogólnego twierdzenia znaleźć można w [52, tw. 3]. \square

Stwierdzenie 4.2 (Warunek optymalności pierwszego rzędu). *Niech $f(\cdot)$ będzie funkcją wypukłą, a \mathcal{C} zbiorem wypukłym, na którym funkcja ta jest różniczkowalna. Wtedy*

$$\mathbf{x}^* \in \arg \min_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x}),$$

wtedy i tylko wtedy, gdy dla każdego $\mathbf{y} \in \mathcal{C}$ zachodzi nierówność

$$\langle \nabla f(\mathbf{x}^*), \mathbf{x}^* - \mathbf{y} \rangle \leq 0.$$

Dowód. Zob. np. [14]. \square

Definicja 4.4 (Funkcja lipschitzowska). Funkcję $f : \mathbb{R}^n \rightarrow \mathbb{R}$ nazywamy funkcją lipschitzowską ze stałą G , na zbiorze \mathcal{C} , względem normy $\|\cdot\|$, jeśli dla każdego $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ zachodzi

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq G \|\mathbf{x} - \mathbf{y}\|.$$

Lemat 4.5. *Niech $f : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją wypukłą. Funkcja ta jest funkcją lipschitzowską ze stałą G , na zbiorze \mathcal{C} , względem normy $\|\cdot\|$ wtedy i tylko wtedy, gdy dla każdego $\mathbf{x} \in \mathcal{C}$, norma dualna gradientu funkcji jest ograniczona przez G , tj. $\|\nabla f(\mathbf{x})\|_* \leq G$.*

Dowód. Zob. [99, lem. 2.6]. \square

Rozdział 5

Uwagi końcowe

Na zakończenie podsumujemy oryginalne wyniki zaprezentowane w pracy oraz przedstawimy otwarte problemy badawcze.

5.1 Oryginalne wyniki naukowe przedstawione w pracy

Wyniki otrzymane w pracy można podsumować następująco:

- Przeprowadzono analizę własności aproksymacyjnych reprezentacji wielomianowych dla wybranych systemów o strukturze blokowej z gładkimi nieliniowościami:
 - systemu Wienera-Hammersteina – w przypadku reprezentacji opartej o szereg Volterra,
 - systemu Hammersteina – w przypadku reprezentacji opartych o szeregi ortogonalne.
- Dokonano syntezy znanych z literatury wyników dotyczących rozwinięcia zmiennych losowych w chaos wielomianowy i formalnie potwierdzono zasadność stosowania reprezentacji ortogonalnych w zadaniu modelowania systemów nieliniowych należących do klasy Wienera (w odniesieniu do rozkładów normalnych oraz szerokiej klasy rozkładów o ograniczonym nośniku).
- Zaproponowano metodę estymacji modeli wielomianowych bazującą na algorytmie lustrzanym z uśrednianiem i przeprowadzono jej analizę teoretyczną. W szczególności:
 - wykazano zbieżność błędu estymacji dla klasy systemów z pamięcią skończoną oraz systemów z pamięcią zanikającą eksponencjalnie,
 - przeprowadzono analizę własności nieasympotycznych wariantu algorytmu opartego o funkcję ujemnej entropii,
 - wykazano teoretyczną zasadność stosowania adaptacyjnego doboru kroku w przypadku danych o charakterze losowym oraz systemów statycznych,

- przedstawiono warunki, w których całosciowy błąd metody (wartość funkcji kryterialnej $Q(\cdot)$) zbiega do zera w przypadku estymacji systemów o strukturze LNL za pomocą entropijnego wariantu metody oraz modeli volterrowskich.

Ponadto:

- Opracowano implementację entropijnego wariantu zaproponowanej metody w języku Python.
- Przeprowadzono symulację numeryczną, ilustrującą działanie metody na zbiorze danych pomiarowych uzyskanych w ramach eksperymentu wykonanego na fizycznie istniejącym systemie.

5.2 Otwarte problemy badawcze

W toku badań przeprowadzonych nad metodą modelowania z rozdziału 4 wyłoniono następujące problemy otwarte:

- Eksperyment numeryczny wykazał, że adaptacyjny dobór kroku może być elementem koniecznym satysfakcjonującego działania metody. Jednakże potrzebne jest ustalenie warunków, w których zastosowanie adaptacyjnego wariantu metody znajduje uzasadnienie *formalne* (problem dotyczy systemów dynamicznych).
- Zbadanie możliwości wykorzystania metody dla sygnałów wejściowych o nieograniczonej amplitudzie. Obecne wyniki teoretyczne dotyczące zastosowania metody w celu estymacji modeli wielomianowych wykluczają wejścia o rozkładzie normalnym, co z kolei eliminuje z rozważań reprezentację ortogonalną opartą o wielomiany Hermite'a.
- Zbadanie możliwości złagodzenia założeń, w szczególności dotyczących sygnału wejściowego i zakłócenia.

Bibliografia

- [1] A. Agrawal, R. Verschueren, S. Diamond, S. Boyd. A rewriting system for convex optimization problems. *Journal of Control and Decision*, 5(1):42–60, 2018.
- [2] F. Bach, R. Jenatton, J. Mairal, G. Obozinski. Optimization with Sparsity-Inducing Penalties. *Foundations and Trends in Machine Learning*, 4(1):1–106, 2011.
- [3] A. Beck, M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.
- [4] G. Ben-Yu. Error estimation of Hermite spectral method for nonlinear partial differential equations. *Mathematics of Computation*, 68:1067–1078, 1999.
- [5] S. A. Billings. Identification of nonlinear systems – a survey. *IEE Proceedings*, 127(6):272–285, 1980.
- [6] G. Birpoutsoukis, P. Z. Csurcsia, J. Schoukens. Efficient multidimensional regularization for Volterra series estimation. *Mechanical Systems and Signal Processing*, 104:896–914, 2018.
- [7] G. Birpoutsoukis, A. Marconato, J. Lataire, J. Schoukens. Regularized non-parametric Volterra kernel estimation. *Automatica*, 82:324–327, 2017.
- [8] A. G. Bose. *A theory of nonlinear systems*. Praca doktorska, Massachusetts Institute of Technology, 1956.
- [9] L. Bottou, F. E. Curtis, J. Nocedal. Optimization methods for large-scale machine learning, 2018.
- [10] G. E. P. Box, G. M. Jenkins, G. C. Reinsel. *Time Series Analysis*. Wiley Series in Probability and Statistics. Wiley, 2008.
- [11] S. Boyd, L. O. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, 32(11):1150–1161, 1985.
- [12] S. Boyd, L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- [13] A. M. Bruckstein, D. L. Donoho, M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, 2009.

-
- [14] S. Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3-4):231–357, 2015.
- [15] F. Bunea, A. B. Tsybakov, M. H. Wegkamp. Aggregation for Gaussian regression. *The Annals of Statistics*, 35(4):1674–1697, 2007.
- [16] R. H. Cameron, W. T. Martin. The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *The Annals of Mathematics*, 48(2):385, 1947.
- [17] D. Campbell, D. Farmer, J. Crutchfield, E. Jen. Experimental mathematics: The role of computation in nonlinear science. *Communications of the ACM*, 28(4):374–384, 1985.
- [18] E. J. Candès, J. Romberg, T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, feb 2006.
- [19] N. Cesa-Bianchi, G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- [20] C. Cheng, Z. Peng, W. Zhang, G. Meng. Volterra-series-based nonlinear system modeling and its engineering applications: A state-of-the-art review. *Mechanical Systems and Signal Processing*, 87:340–364, 2017.
- [21] S. Diamond, S. Boyd. CVXPY: A Python-Embedded Modeling Language for Convex Optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.
- [22] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [23] F. J. Doyle, R. K. Pearson, B. A. Ogunnaike. *Identification and Control Using Volterra Models*. Springer London, 2002.
- [24] J. Duchi, E. Hazan, Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*, 2010.
- [25] J. Duchi, E. Hazan, Y. Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.
- [26] J. C. Duchi, A. Agarwal, M. Johansson, M. I. Jordan. Ergodic mirror descent. *SIAM Journal on Optimization*, 22(4):1549–1578, 2012.
- [27] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani. Least angle regression. *The Annals of Statistics*, 32(2):407–499, 2004.
- [28] O. G. Ernst, A. Mugler, H.-J. Starkloff, E. Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM Mathematical Modelling and Numerical Analysis*, 46(2):317–339, 2012.

- [29] C. Eun, E. J. Powers. A new Volterra predistorter based on the indirect learning architecture. *IEEE Transactions on Signal Processing*, 45(1):223–227, 1997.
- [30] L. Fagiano, M. Khammash. Nonlinear stochastic model predictive control via regularized polynomial chaos expansions. *Proceedings of the IEEE Conference on Decision and Control*, strony 142–147, 2012.
- [31] G. M. Fichtenholz. *Rachunek różniczkowy i całkowy, Tom 3*. Państwowe Wydawnictwo Naukowe, Warszawa, 1975.
- [32] W. A. Frank. An efficient approximation to the quadratic Volterra filter and its application in real-time loudspeaker linearization. *Signal Processing*, 45(1):97–113, 1995.
- [33] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani. Pathwise coordinate optimization. *The Annals of Applied Statistics*, 1(2):302–332, 2007.
- [34] I. Goodfellow, Y. Bengio, A. Courville. *Deep Learning*. MIT Press, 2016.
- [35] W. Greblicki. Nonparametric system identification by orthogonal series. *Problems of Control and Information Theory*, 8(1):67–73, 1979.
- [36] W. Greblicki. *Podstawy teoretyczne automatyki*. Oficyna Wydawnicza Politechniki Wrocławskiej, 2001.
- [37] W. Greblicki, M. Pawlak. Nonparametric identification of Hammerstein systems. *IEEE Transactions on Information Theory*, 35(2):409–418, 1989.
- [38] W. Greblicki, M. Pawlak. *Nonparametric system identification*. Cambridge University Press Cambridge, 2008.
- [39] J. K. Gruber, C. Bordons, A. Oliva. Nonlinear MPC for the airflow in a PEM fuel cell using a Volterra series model. *Control Engineering Practice*, 20(2):205–217, 2012.
- [40] J. K. Gruber, J. L. Guzmán, F. Rodríguez, C. Bordons, M. Berenguel, J. A. Sánchez. Nonlinear MPC based on a Volterra series model for greenhouse temperature control using natural ventilation. *Control Engineering Practice*, 19(4):354–366, 2011.
- [41] L. Györfi, M. Kohler, A. Krzyżak, H. Walk. *A Distribution-Free Theory of Nonparametric Regression*. Springer Science & Business Media, 2006.
- [42] Z. Hasiewicz, G. Mzyk. Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 49(8):1370–1375, 2004.
- [43] Z. Hasiewicz, M. Pawlak, P. Śliwiński. Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(2):427–442, feb 2005.

-
- [44] E. Hazan. *Introduction to Online Convex Optimization*.
- [45] G. E. Hinton, S. Osindero, Y. W. Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554, 2006.
- [46] A. E. Hoerl, R. W. Kennard. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55–67, 1970.
- [47] F. S. Hover. Gradient dynamic optimization with Legendre chaos. *Automatica*, 44(1):135–140, 2008.
- [48] D. Hsu, S. M. Kakade, T. Zhang. Random design analysis of ridge regression. *Foundations of Computational Mathematics*, 14:569–600, 2014.
- [49] S. Janson. *Gaussian Hilbert Spaces*. Cambridge University Press, 1997.
- [50] A. Juditsky, G. Lan, A. Nemirovski, A. Shapiro. Robust Stochastic Approximation Approach to Stochastic Programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.
- [51] A. Juditsky, A. Nemirovski. Functional aggregation for nonparametric regression. *The Annals of Statistics*, 28(3):681–712, 2000.
- [52] S. M. Kakade, S. Shalev-Shwartz, A. Tewari. Regularization Techniques for Learning with Matrices. *Journal of Machine Learning Research*, 13:1865–1890, 2012.
- [53] V. Kekatos, G. B. Giannakis. Sparse Volterra and polynomial regression models: Recoverability and estimation. *IEEE Transactions on Signal Processing*, 59(12):5907–5920, 2011.
- [54] K. K. Kevin Kim, R. D. Braatz. Probabilistic analysis and control of uncertain dynamic systems: Generalized polynomial chaos expansion approaches. *Proceedings of the American Control Conference*, strony 44–49, 2012.
- [55] A. Y. Kibangou, G. Favier. Wiener-Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE Signal Processing Letters*, 13(6):381–384, 2006.
- [56] K. K. K. Kim, R. D. Braatz. Generalised polynomial chaos expansion approaches to approximate stochastic model predictive control. *International Journal of Control*, 86(8):1324–1337, aug 2013.
- [57] K. K. K. Kim, D. E. Shen, Z. K. Nagy, R. D. Braatz. Wiener’s polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties. *IEEE Control Systems*, 33(5):58–67, 2013.
- [58] S. J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky. An interior-point method for large-scale ℓ_1 -regularized least squares. *IEEE Journal on Selected Topics in Signal Processing*, 1(4):606–617, 2007.
- [59] J. Kivinen, M. K. Warmuth. Exponentiated Gradient versus Gradient Descent for Linear Predictors. *Information and Computation*, 132(1):1–63, 1997.

- [60] M. J. Korenberg. Parallel cascade identification and kernel estimation for nonlinear systems. *Annals of Biomedical Engineering*, 19(4):429–455, 1991.
- [61] M. J. Korenberg, S. B. Bruder, P. J. McLlroy. Exact orthogonal kernel estimation from finite data records: Extending Wiener’s identification of nonlinear systems. *Annals of Biomedical Engineering*, 16(2):201–214, 1988.
- [62] S. Łagosz, P. Śliwiński, P. Wachel. Identification of Wiener-Hammerstein systems by ℓ_1 -constrained Volterra series. *European Journal of Control*, 58:53–59, 2021.
- [63] Y. W. Lee, M. Schetzen. Measurement of the Wiener kernels of a non-linear system by cross-correlation. *International Journal of Control*, 2:3:237–254, 1965.
- [64] H. X. Li, C. Qi, Y. Yu. A spatio-temporal Volterra modeling approach for a class of distributed industrial processes. *Journal of Process Control*, 19(7):1126–1142, 2009.
- [65] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, Y. Cai. Intra-channel nonlinearity compensation by inverse Volterra series transfer function. *Journal of Lightwave Technology*, 30(3):310–316, 2012.
- [66] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, 1999.
- [67] L. Ljung. Perspectives on system identification. *Annual Reviews in Control*, 34(1):1–12, 2010.
- [68] N. Lüthen, S. Marelli, B. Sudret. Sparse Polynomial Chaos Expansions: Literature Survey and Benchmark. Raport instytutowy, 2020.
- [69] B. R. Maner, F. J. Doyle. Polymerization reactor control using autoregressive-plus Volterra-based MPC. *AIChE Journal*, 43(7):1763–1784, 1997.
- [70] P. Z. Marmarelis, V. Z. Marmarelis. *Analysis of physiological systems: the white-noise approach*. 1978.
- [71] V. J. Mathews, G. L. Sicuranza. *Polynomial Signal Processing*. Wiley, New York, 2000.
- [72] H. B. McMahan. A Survey of Algorithms and Analysis for Adaptive Online Learning. *Journal of Machine Learning Research*, 18:1–50, 2017.
- [73] H. B. McMahan, M. Streeter. Adaptive Bound Optimization for Online Convex Optimization. *Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*, 2010.
- [74] G. Mzyk. *Combined Parametric-Nonparametric Identification of Block-Oriented Systems*, wolumen 454 serii *Lecture Notes in Control and Information Sciences*. Springer, 2014.
- [75] K. S. Narendra, P. G. Gallman. An Iterative Method for the Identification of Nonlinear Systems Using a Hammerstein Model. *IEEE Transactions on Automatic Control*, 11(3):546–550, 1966.

- [76] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.
- [77] A. S. Nemirovsky, D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley, 1983.
- [78] Y. Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120:221–259, 2009.
- [79] Y. Nesterov, A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.
- [80] F. Orabona. A Modern Introduction to Online Learning. [Online; dostępny 23 czerwca 2021] <http://arxiv.org/abs/1912.13213v3>, 2020.
- [81] G. Palm. On representation and approximation of nonlinear systems, Part II: Discrete systems. *Biological Cybernetics*, 34(1):49–52, 1979.
- [82] G. Palm, T. Poggio. The Volterra Representation and the Wiener Expansion: Validity and Pitfalls. *SIAM Journal on Applied Mathematics*, 33(2):195–216, 1977.
- [83] M. Pawlak, Z. Hasiewicz. Nonlinear system identification by the Haar multiresolution analysis. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45(9):945–961, 1998.
- [84] M. Pawlak, Z. Hasiewicz, P. Wachel. On nonparametric identification of Wiener systems. *IEEE Transactions on Signal Processing*, 55(2):482–492, 2007.
- [85] C. Rao. *Linear Statistical Inference and its Applications*. Wiley Series in Probability and Statistics. Wiley, 1973.
- [86] P. Rigollet, J.-C. Hütter. High Dimensional Statistics. [Online; dostępny 23 czerwca 2021] <http://www-math.mit.edu/~rigollet/PDFs/RigNotes17.pdf>, 2019.
- [87] H. Robbins, S. Monro. A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.
- [88] R. T. Rockafellar, R. J. B. Wets. *Variational Analysis*. Springer, New York, 1998.
- [89] W. J. Rugh. *Nonlinear System Theory*. 1981.
- [90] I. W. Sandberg. A Note on the Convolution Scandal. 8(7):210–211, 2001.
- [91] M. Schetzen. A theory of non-linear system identification. *International Journal of Control*, 20(4):577–592, 1974.
- [92] M. Schetzen. *The Volterra and Wiener theories of nonlinear systems*. Wiley, 1980.
- [93] M. Schetzen. Nonlinear system modeling based on the Wiener theory. *Proceedings of the IEEE*, 69(12):1557–1573, 1981.

- [94] M. Schetzen. *Nonlinear System Modelling and Analysis from the Volterra and Wiener Perspective*, strony 13–24. Springer London, London, 2010.
- [95] M. Schmidt, G. Fung, R. Rosales. Fast Optimization Methods for L1 Regularization: A Comparative Study and Two New Approaches. *Machine Learning: ECML 2007*, strony 286–297, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [96] J. Schoukens, L. Ljung. Nonlinear System Identification: A User-Oriented Road Map. *IEEE Control Systems*, 39(6):28–99, 2019.
- [97] M. Schoukens, P. Mattsson, T. Wigren, J. P. Noël. Cascaded tanks benchmark combining soft and hard nonlinearities. *20th World Congress of the International Federation of Automatic Control*, strony 448–453, Toulouse, France, 2017.
- [98] S. Shalev-Shwartz. *Online Learning: Theory, Algorithms, and Applications*. Praca doktorska, Hebrew University of Jerusalem, 2007.
- [99] S. Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.
- [100] S. Shalev-Shwartz, O. Shamir, N. Srebro, K. Sridharan. Learnability, Stability and Uniform Convergence. *Journal of Machine Learning Research*, 11:2635–2670, 2010.
- [101] A. Shapiro, D. Dentcheva, A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Society for Industrial and Applied Mathematics, 2014.
- [102] P. Śliwiński. *Nonlinear System Identification by Haar Wavelets*. 2009.
- [103] P. Śliwiński, A. Marconato, P. Wachel, G. Birpoutsoukis. Non-linear system modelling based on constrained Volterra series estimates. *IET Control Theory & Applications*, 11(15):2623–2629, 2017.
- [104] P. Śliwiński, P. Wachel, S. Łagosz. Nonlinearity recovery by standard and aggregative orthogonal series algorithms. *Applied Stochastic Models in Business and Industry*, 34(5), 2018.
- [105] T. Söderström, P. Stoica. *System Identification*. Prentice-Hall International, 1989.
- [106] B. Sudret. Global sensitivity analysis using polynomial chaos expansions, 2008.
- [107] L. Tan, J. Jiang. Adaptive Volterra filters for active control of nonlinear noise processes. *IEEE Transactions on Signal Processing*, 49(8):1667–1676, 2001.
- [108] R. Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.
- [109] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, 2004.

-
- [110] A. B. Tsybakov. Optimal rates of aggregation. *Learning Theory and Kernel Machines. Lecture Notes in Computer Science*, wolumen 2777, strony 303–313, Berlin, Heidelberg, 2003. Springer.
- [111] E. Van Den Berg, M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2):890–912, 2008.
- [112] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 2000.
- [113] Volterra Vito. *Theory of Functionals and of Integral and Integro-Differential Equations*. Dover Publications, New York, 1959.
- [114] P. Wachel. Convex aggregative modelling of infinite memory nonlinear systems. *International Journal of Control*, 89(8):1613–1621, 2016.
- [115] P. Wachel. *Identyfikacja i agregacyjne modelowanie nieliniowych systemów dynamicznych*. Akademicka Oficyna Wydawnicza EXIT, 2017.
- [116] P. Wachel, P. Śliwiński. Aggregative Modeling of Nonlinear Systems. *IEEE Signal Processing Letters*, 22(9):1482–1486, 2015.
- [117] N. Wiener. *Nonlinear problems in random theory*. 1966.
- [118] L. Xiao. Dual Averaging Methods for Regularized Stochastic Learning and Online Optimization. *Journal of Machine Learning Research*, 11:2543–2596, 2010.
- [119] D. Xiu, G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *Society for Industrial and Applied Mathematics*, 24(2):619–644, 2002.
- [120] T. Zhang. On the Consistency of Feature Selection using Greedy Least Squares Regression. *Journal of Machine Learning Research*, 10:555–568, 2009.
- [121] Z. Zhang, Y. Xu, J. Yang, X. Li, D. Zhang. A Survey of Sparse Representation: Algorithms and Applications. *IEEE Access*, 3:490–530, 2015.
- [122] M. Zinkevich. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. *Proceedings of the 20th international conference on machine learning (icml-03)*, strony 928–936, 2003.