8 Identyfikacja systemów za pomocą falek – algorytmy

Zygmunt Hasiewicz

8.1 Wstęp

W tym rozdziale przedstawiamy algorytmy identyfikacji charakterystyk nieliniowych oparte na ortogonalnych układach falkowych. Prace nad zastosowaniem układów funkcji ortogonalnych do konstrukcji algorytmów identyfikacji nieliniowych charakterystyk systemów prowadzone są w Instytucie Cybernetyki Technicznej od wielu lat i w przeszłości do tego celu wykorzystane zostały klasyczne układy ortogonalne, takie jak funkcje trygonometryczne oraz ortogonalne wielomiany Legendre'a, Laguerre'a i Hermite'a. Wypracowany został ogólny schemat konstruowania i analizy tak zwanych ortogonalnych algorytmów identyfikacji, a specyfika poszczególnych układów funkcji ortogonalnych dała algorytmy identyfikacji o odmiennych postaciach i własnościach. Odpowiednie rezultaty przedstawione są między innymi w pracach [10]-[14], [26], [27] oraz w [21], [22], [24], [28], [29] powstałych we współpracujących ośrodkach.

Funkcje falkowe, a zwłaszcza układy falek ortogonalnych, stały się w ostatniej dekadzie popularnym narzędziem i znalazły liczne zastosowania w wielu rozmaitych dziedzinach nauki i techniki. O szczególnej pozycji falek ortogonalnych zadecydowały ich korzystne cechy, które w skrócie można przedstawić następująco (patrz, np., [2], [7], [23], [32]):

1. Koncepcja falek ortogonalnych jest niezwykle prosta i opiera się na dwóch, odpowiednio dobranych, funkcjach początkowych, tzw. falce-ojcu (funkcji skalującej) $\varphi(x)$ oraz falce-matce $\psi(x)$.

2. Istnieje ogólna metodologia konstruowania ortogonalnych baz falkowych, która umożliwia otrzymywanie falek ortogonalnych o zadanych z góry własnościach (np. rozmiarze nośnika, liczbie znikających momentów, itp.).

3. Dostępne są gotowe pakiety oprogramowania do szybkich obliczeń falkowych.

Ważne miejsce ze względu na swoje specyficzne własności zajmuje tu klasa falek ortogonalnych o zwartym nośniku. Typowymi reprezentantami tej klasy są układ ortogonalny Haara, rodziny falek Daubechies oraz symlety i koiflety o różnych numerach falkowych p. Uzupełnieniem tych przykładów może być szereg dalszych rodzin falkowych, jakie można skonstruować indywidualnie [6]. Podstawową wadą falek z praktycznego punktu widzenia jest to, iż z reguły nie

są one wyrażone w jawnej postaci analitycznej. Nadaje to niejawny charakter algorytmom falkowym i wymaga w praktyce bezpośredniego ich powiązania (w trybie 'on line') z odpowiednimi numerycznymi procedurami do obliczeń wartości funkcji falkowych. Dostępność gotowych pakietów oprogramowania (na przykład w środowisku Matlab lub Mathematica) jest tutaj istotnym ułatwieniem.

8.2 Modele falkowe

Rosnące zainteresowanie funkcjami falkowymi w zastosowaniu do identyfikacji systemów nieliniowych (patrz, np., [5], [8], [16]-[19], [25], [31]) ma swoje źródło w szczególnych cechach modeli falkowych, które można scharakteryzować w następujący sposób: (a) różnorodność reprezentacji i łatwość modyfikacji, (b) adaptacyjność (dostrajalność) do identyfikowanego sektora (punktu) charakterystyki nieliniowej oraz (c) możliwość swobodnego wyboru funkcji falkowych. Cechy te wynikają z ogólnych własności falek i aproksymacji falkowych [2], [7], [23], [32].

8.2.1 Klasy modeli

Każda charakterystyka całkowalna z kwadratem $F(x) \in L^2(\mathbf{R})$ może być aproksymowana za pomocą układu falek ortogonalnych $\{\varphi_{Mn}(x)\}_{n\in\mathbf{Z}} \cup \{\psi_{mn}(x)\}$, $M \le m \le K-1, n \in \mathbf{Z}$, $\mathbf{Z}-$ zbiór liczb całkowitych. Model falkowy ma postać

$$F(x;K) = \sum_{n=-\infty}^{\infty} \alpha_{Mn}^F \varphi_{Mn}(x) + \sum_{m=M}^{K-1} \sum_{n=-\infty}^{\infty} \beta_{mn}^F \psi_{mn}(x) \tag{8.1}$$

gdzie

$$\alpha_{Mn}^F = \int_{-\infty}^{+\infty} F(x)\,\varphi_{Mn}(x)\,dx\;; \quad \beta_{mn}^F = \int_{-\infty}^{+\infty} F(x)\,\psi_{mn}(x)\,dx \tag{8.2}$$

Funkcje falkowe w tym modelu

$$\varphi_{Mn}(x) = 2^{M/2}\varphi\left(2^M x - n\right)\;; \quad \psi_{mn}(x) = 2^{m/2}\psi\left(2^m x - n\right) \tag{8.3}$$

otrzymuje się w łatwy sposób przez skalowanie (czynnik m) oraz przesunięcie (czynnik n) wzajemnie odpowiadających sobie falki-ojca $\varphi(x)$ oraz falki-matki $\psi(x)$. Funkcje $\{\varphi_{Mn}(x)\}$ oraz $\{\psi_{mn}(x)\}$ dla $m = M, M+1, ..., K-1$ oraz $n \in \mathbf{Z}$ generują początkową klasę modeli $V_M = \text{span}\{\varphi_{Mn}(x), n \in \mathbf{Z}\}$ oraz przestrzenie dodatkowych (uzupełniających) detali $W_m = \text{span}\{\psi_{mn}(x), n \in \mathbf{Z}\}$ i wyznaczają w ten sposób klasę modeli $V_M \oplus W_M \oplus W_{M+1} \oplus \ldots \oplus W_{K-1} = V_K \subset L^2(\mathbf{R})$, do której należy model $F(x;K)$. Taka budowa daje możliwość stopniowej ("krok po kroku") poprawy jakości modelu falkowego przez dodawanie kolejnych detali

$$F(x;K+1) = F(x;K) + \sum_{n=-\infty}^{\infty} \beta_{Kn}^F \psi_{Kn}(x)$$

Ze względu na zależność

$$\text{span}\left\{\varphi_{Mn}(x), \psi_{mn}(x);\ M \le m \le K-1, n \in \mathbf{Z}\right\} = \text{span}\left\{\varphi_{Kn}(x),\ n \in \mathbf{Z}\right\}$$

model $F(x;K) \in V_K$ można równoważnie przedstawić w postaci

$$F(x;K) = \sum_{n=-\infty}^{\infty} \alpha_{Kn}^{F} \varphi_{Kn}(x) \tag{8.4}$$

lub w postaci całkowej (wstawiając definicję (8.2) współczynnika α_{Kn}^{F} do wzoru (8.4))

$$F(x;K) = \int_{-\infty}^{+\infty} q_K(x,v) F(v)\, dv \tag{8.5}$$

gdzie $q_K(x,v) = \sum_{n=-\infty}^{\infty} \varphi_{Kn}(x)\varphi_{Kn}(v)$ jest jądrem układu $\{\varphi_{Kn}(x), n \in \mathbf{Z}\}$.

Wzory (8.1), (8.4), (8.5) wyznaczają trzy różne klasy modeli falkowych w przestrzeni V_K.

8.2.2 Modele o nośniku zwartym

Dla falek ograniczonych o zwartym nośniku, to jest spełniających warunki

$$|\varphi(x)| \le M_\varphi I_{[s_1,s_2]}(x); \quad |\psi(x)| \le M_\psi I_{[t_1,t_2]}(x) \tag{8.6}$$

oraz (patrz (8.3))

$$|\varphi_{Mn}(x)| \le 2^{M/2} M_\varphi\ I_{\left[\frac{s_1+n}{2^M}, \frac{s_2+n}{2^M}\right]}(x); \quad |\psi_{mn}(x)| \le 2^{m/2} M_\psi\ I_{\left[\frac{t_1+n}{2^m}, \frac{t_2+n}{2^m}\right]}(x)$$

gdzie M_φ, M_ψ - pewne stałe dodatnie, model falkowy (8.1) osiąga zdolność do samostrojenia poprzez automatyczną adaptację granic sumowania do lokalnego punktu x oraz nośników falek zastosowanych w modelu:

$$F(x;K) = \sum_{n=n_{\min}(\varphi,x,M)}^{n_{\max}(\varphi,x,M)} \alpha_{Mn}^{F}\varphi_{Mn}(x) + \sum_{m=M}^{K-1} \sum_{n=n_{\min}(\psi,x,m)}^{n_{\max}(\psi,x,m)} \beta_{mn}^{F}\psi_{mn}(x) \tag{8.7}$$

We wzorze (8.7) zachodzi wtedy

$$\alpha_{Mn}^{F} = \int_{(s_1+n)/2^M}^{(s_2+n)/2^M} F(x)\varphi_{Mn}(x)\, dx; \quad \beta_{mn}^{F} = \int_{(t_1+n)/2^m}^{(t_2+n)/2^m} F(x)\psi_{mn}(x)\, dx$$

oraz

$$\begin{aligned} n_{\min}(\varphi,x,M) &= \lceil 2^M x - s_2 \rceil; \quad n_{\max}(\varphi,x,M) = \lfloor 2^M x - s_1 \rfloor \\ n_{\min}(\psi,x,m) &= \lceil 2^m x - t_2 \rceil; \quad n_{\max}(\psi,x,m) = \lfloor 2^m x - t_1 \rfloor \end{aligned}$$

wskazują aktywne (nieznikające) falki w punkcie x ($\lfloor . \rfloor$ oraz $\lceil . \rceil$ oznacza funkcję 'podłoga' oraz 'sufit'). Odpowiednie modele (8.4), (8.5) przybierają wówczas postać

$$F(x;K) = \sum_{n=n_{\min}(\varphi,x,K)}^{n_{\max}(\varphi,x,K)} \alpha_{Kn}^{F} \varphi_{Kn}(x) \tag{8.8}$$

oraz

$$F(x;K) = \int_{x-(s_2-s_1)/2^K}^{x+(s_2-s_1)/2^K} q_K(x,v)\, F(v)\, dv \tag{8.9}$$

gdzie $q_K(x,v) = \sum_{n=n_{\min}(\varphi,x,K)}^{n_{\max}(\varphi,x,K)} \varphi_{Kn}(x)\, \varphi_{Kn}(v)$ oraz $n_{\min}(\varphi,x,K) = \lceil 2^K x - s_2 \rceil$, $n_{\max}(\varphi,x,K) = \lfloor 2^K x - s_1 \rfloor$. Liczba składników w modelach (8.7), (8.8) oraz wyrażeniu $q_K(x,v)$ jest ograniczona i zależy od rozmiaru nośników $s = s_2 - s_1 = t_2 - t_1$ falek $\varphi(x)$ oraz $\psi(x)$. Dla każdego m oraz każdego x zachodzi oszacowanie

$$\begin{aligned} n_{\max}(\varphi,x,m) - n_{\min}(\varphi,x,m) + 1 &\leq S \\ n_{\max}(\psi,x,m) - n_{\min}(\psi,x,m) + 1 &\leq S \end{aligned} \tag{8.10}$$

gdzie $S = \lfloor s \rfloor + 1$, a nośniki modeli (8.7) oraz (8.8)-(8.9) zawarte są w przedziale $\left[x - s/2^K, x + s/2^K\right]$. Zachodzą dwie podstawowe własności:

1. Dla funkcji falkowych jak we wzorze (8.6) model $F(x;K)$ (o postaci (8.7), (8.8) lub (8.9)) zbiega ze wzrostem K punktowo do charakterystyki $F(x)$ prawie wszędzie (w sensie miary Lebesgue'a), to znaczy

$$|AE(F;x;K)| \to 0 \quad \text{gdy} \quad K \to \infty \tag{8.11}$$

dla prawie każdego punktu x, gdzie $|AE(F;x;K)| = |F(x) - F(x;K)|$ jest punktowym błędem aproksymacji dla modelu $F(x;K)$, a w szczególności zbieżność (8.11) zachodzi w punktach ciągłości charakterystyki $F(x)$ (por. Twierdzenie 2.1(ii) w [20]).

2. Dla lokalnie gładkich (w otoczeniu punktu x_0) charakterystyk $F(x) \in C^{\nu_F}(x_0)$, gdzie $\nu_F \in (0,1]$ lub $\nu_F \in \mathbf{N}$ - zbiór liczb naturalnych (tj. spełniających warunek Lipschitza lub różniczkowalnych) oraz dużych wartości skali K zachodzi oszacowanie

$$|AE(F;x_0;K)| \leq C_{AE,F\psi}\ 2^{-\lambda_{F\psi} K} \tag{8.12}$$

gdzie $C_{AE,F\psi}$ jest stałą, $\lambda_{F\psi} = \min\{\nu_F, r_\psi + 1\}$ oraz r_ψ jest liczbą znikających momentów falki-matki $\psi(x)$, tj.

$$r_\psi = \max\{\, k \geq 0 : \int_{t_1}^{t_2} x^k \psi(x)\, dx = 0\} \tag{8.13}$$

Niektóre z wymienionych charakterystyk dla podstawowych klas funkcji falkowych zebrane zostały w Tabeli 8.1 (p oznacza tam numer falki).

	Daubechies/symlety	koiflety
nośnik $\varphi(x)$	$[0, 2p-1]$	$[-2p, 4p-1]$
nośnik $\psi(x)$	$[-p+1, p]$	$[-3p+1, 3p]$
$n_{\min}(\varphi, x, m)$	$\lfloor 2^m x \rfloor - 2p + 2$	$\lfloor 2^m x \rfloor - 4p + 2$
$n_{\max}(\varphi, x, m)$	$\lceil 2^m x \rceil - 1$	$\lceil 2^m x \rceil + 2p - 1$
$n_{\min}(\psi, x, m)$	$\lfloor 2^m x \rfloor - p + 1$	$\lfloor 2^m x \rfloor - 3p + 1$
$n_{\max}(\psi, x, m)$	$\lceil 2^m x \rceil + p - 2$	$\lceil 2^m x \rceil + 3p - 2$
r_ψ	$p-1$	$2p-1$

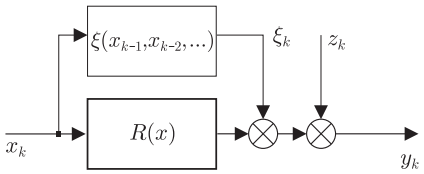
Tablica 8.1. Podstawowe charakterystyki dla typowych falek ortogonalnych o zwartym nośniku

8.3 Zadanie identyfikacji

Rozpatrujemy zagadnienie identyfikacji nieliniowości złożonego systemu dynamicznego o strukturze blokowej, zbudowanego z nieliniowych bloków statycznych oraz liniowych dynamicznych ([3], [15]), na podstawie pomiarów jego wejścia i wyjścia. Dla podkreślenia, że identyfikowana nieliniowość może się w takim wypadku różnić od charakterystyk nieliniowych statycznych części systemu, dla jej oznaczenia stosujemy symbol R. Odpowiednią dyskusję można znaleźć w pracach [1], [15], [9]. Rozważania dotyczą klasy systemów dyskretnych, dla których nieliniowość R można 'wyizolować' z reszty systemu (Rys. 8.1), tj. opis wejście-wyjście systemu jako całości przedstawić w postaci:

$$y_k = R(x_k) + \xi(x_{k-1}, x_{k-2}, \ldots) + z_k \tag{8.14}$$

W tym opisie $\{\xi_k = \xi(x_{k-1}, x_{k-2}, \ldots)\}$ reprezentuje oddziaływanie dynamicznej części systemu na wyjście i z punktu widzenia nieliniowości R jest 'szumem systemowym', $\{z_k\}$ jest zewnętrznym zakłóceniem losowym, a $\{(x_k, y_k)\}$ - stacjonarnym ciągiem losowych wejść i wyjść systemu.



Rysunek 8.1. Identyfikowany system ogólny

O systemie (8.14) przyjmuje się następujące założenia:

A1. Nieliniowość $R(x)$ spełnia ograniczenie $|R(x)| \leq C_{1R}|x| + C_{2R}$ dla pewnych stałych $C_{1R}, C_{2R} > 0$.

A2. Sygnał wejściowy $\{x_k\}_{k \in \mathbf{Z}}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie (ciąg i.i.d.) ze skończoną wariancją σ_x^2, dla których

istnieje gęstość prawdopodobieństwa $f(x)$ i gęstość ta jest ograniczona przez pewną stałą $f(x) \leq M_f < \infty$. Gęstość wejścia $f(x)$ oraz wariancja σ_x^2 nie są znane.

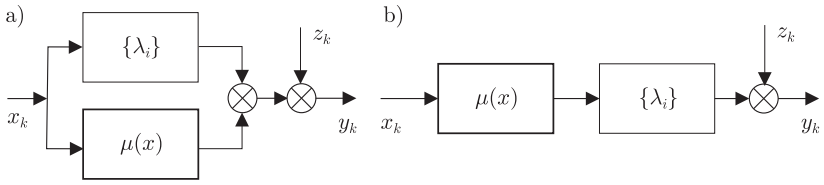
A3. Szum systemowy $\{\xi_k\}_{k\in\mathbf{Z}}$ ma postać $\xi_k = \xi_{1k} + \xi_{2k} + ... + \xi_{Jk}$, gdzie $\xi_{jk} = \sum_{i=1}^{\infty} \lambda_{ji}\zeta_j(x_{k-i})$, przy czym dla każdego $j = 1,2,...,J$ zachodzi $\sum_{i=1}^{\infty} |\lambda_{ji}| < \infty$ oraz $\zeta_j(x)$ jest funkcją nieliniową, taką że $\mathrm{E}\zeta_j(x_1) = 0$ oraz $\mathrm{E}\zeta_j^2(x_1) = \sigma_{j\zeta}^2 < \infty$. Odpowiedzi impulsowe $\{\lambda_{ji}\}$ oraz funkcje $\zeta_j(x)$ nie są znane.

A4. Losowe zakłócenie $\{z_k\}_{k\in\mathbf{Z}}$ jest procesem skorelowanym, powstałym w wyniku filtracji stacjonarnych białych szumów $\{\varepsilon_{lk}\}_{k\in\mathbf{Z}}$, $l = 1,2,...,L$, o zerowej wartości oczekiwanej $\mathrm{E}\varepsilon_{l1} = 0$ i skończonej wariancji $\mathrm{E}\varepsilon_{l1}^2 = \sigma_{l\varepsilon}^2 < \infty$, takim że $z_k = z_{1k} + z_{2k} + ... + z_{Lk}$, gdzie $z_{lk} = \sum_{i=0}^{\infty} \omega_{li}\varepsilon_{l,k-i}$. Dla każdego $l = 1,2,...,L$ zachodzi $\sum_{i=0}^{\infty} |\omega_{li}| < \infty$. Odpowiedzi impulsowe filtrów $\{\omega_{li}\}$ i wariancje zakłóceń $\sigma_{l\varepsilon}^2$ nie są znane.

A5. Procesy $\{x_k\}$ oraz $\{\varepsilon_{lk}\}$, $l = 1,2,...,L$, są wzajemnie statystycznie niezależne.

A6. Możliwy jest pomiar wyłącznie wejść i wyjść systemu jako całości $\{(x_k, y_k)\}$, a nieliniowość $R(x)$ należy wyznaczyć w zadanym przedziale $[a,b]$.

A7. Dla $x \in [a,b]$ zachodzi $0 < \delta \leq f(x)$, gdzie $\delta > 0$ jest pewną stałą.



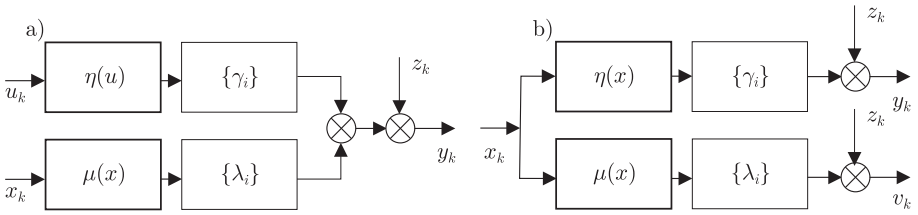
Rysunek 8.2. Przykładowe systemy **a)** prosty system równoległy, **b)** system Hammersteina

Równanie (8.14) oraz założenia A1-A7 obejmują szeroką klasę systemów o strukturach 'otwartych', która zawiera często rozpatrywaną i ważną dla zastosowań klasę systemów szeregowych Hammersteina [10], klasę systemów równoległych [4] oraz szereg klas pochodnych [25]. Przykładowe systemy przedstawione są na rysunkach 8.2-8.3. Szczegółowe omówienie tej klasy wraz z przykładami odpowiednich systemów przedstawiono w pracy [19].

8.4 Falkowe algorytmy identyfikacji

Falkowe algorytmy identyfikacji nieliniowości $R(x)$ wykorzystywać będą modele przedstawione w Sekcji 8.2. Opierać się one będą na spostrzeżeniu, że we

wszystkich punktach x, w których gęstość wejścia $f(x)$ jest dodatnia (por. założenie A7) identyfikowaną nieliniowość $R(x)$ można wyrazić jako iloraz $R(x) = g(x)/f(x)$, gdzie $g(x) = R(x) f(x)$. Przy losowym wejściu systemu prowadzi to do wygodnej dekompozycji zadania i zastąpienia bezpośredniej identyfikacji nieliniowości $R(x)$ prostszymi zadaniami identyfikacji funkcji $g(x)$ (tj. nieliniowości $R(x)$ zmodulowanej przez gęstość $f(x)$) oraz gęstości wejścia $f(x)$ (etap 1). Następnie wystarczy wyznaczyć ilorazowy model nieliniowości $R(x)$ (etap 2). Tego typu schemat postępowania stosowano w cytowanych na wstępie wcześniejszych pracach [10]-[14].



Rysunek 8.3. Przykłady systemów **a)** o wielu wejściach i jednym wyjściu (systemy MISO - multiple input-single output), **b)** o jednym wejściu i wielu wyjściach (systemy SIMO - single input-multiple output)

Przy zastosowaniu falek ortogonalnych o zwartym nośniku, punktem wyjścia do konstrukcji algorytmów identyfikacji są modele falkowe $g(x;K)$ oraz $f(x;K)$ funkcji $g(x)$ oraz $f(x)$ o postaci jak we wzorach (8.7), (8.8) oraz (8.9), z nieznanymi współczynnikami (por. wzór (8.2)):

$$\alpha_{mn}^{g} = \int_{-\infty}^{+\infty} R(x)\,\varphi_{mn}(x)\,f(x)dx \quad \text{dla} \quad m = M, K$$

$$\beta_{mn}^{g} = \int_{-\infty}^{+\infty} R(x)\,\psi_{mn}(x)\,f(x)dx$$

$$\alpha_{mn}^{f} = \int_{-\infty}^{+\infty} \varphi_{mn}(x)\,f(x)dx \quad \text{dla} \quad m = M, K$$

$$\beta_{mn}^{f} = \int_{-\infty}^{+\infty} \psi_{mn}(x)\,f(x)dx$$

Synteza algorytmów identyfikacji przebiega według następujących standardowych kroków:

Krok 1 Uwzględniając, że $f(x)$ jest gęstością rozkładu prawdopodobieństwa oraz biorąc pod uwagę wzór (8.14) wraz z założeniami A2-A5 zauważa się, że współczynniki falkowe α_{mn}^{g}, β_{mn}^{g}, α_{mn}^{f}, β_{mn}^{f} są odpowiednio wartościami oczekiwanymi:

$$\alpha_{mn}^{g} = E\left[R(x_1)\,\varphi_{mn}(x_1)\right] = E\left[y_1 \varphi_{mn}(x_1)\right]$$

$$\beta_{mn}^{g} = E\left[R\left(x_1\right)\psi_{mn}\left(x_1\right)\right] = E\left[y_1\psi_{mn}\left(x_1\right)\right]$$
$$\alpha_{mn}^{f} = E\left[\varphi_{mn}\left(x_1\right)\right]$$
$$\beta_{mn}^{f} = E\left[\psi_{mn}\left(x_1\right)\right]$$

Estymuje się je zatem na podstawie pomiarów $\{(x_k, y_k)\}_{k=1}^{N}$ stosując następujące średnie empiryczne (wykorzystuje się tu stacjonarność procesów $\{x_k\}$ oraz $\{y_k\}$ wynikającą z założeń A2-A5):

$$\hat{\alpha}_{mn}^{F} = N^{-1}\sum_{k=1}^{N}\vartheta_k^{F}\varphi_{mn}\left(x_k\right); \qquad \hat{\beta}_{mn}^{F} = N^{-1}\sum_{k=1}^{N}\vartheta_k^{F}\psi_{mn}\left(x_k\right) \tag{8.15}$$

gdzie $F \in \{g, f\}$ oraz $\vartheta_k^{g} = y_k$, $\vartheta_k^{f} = 1$.

Krok 2 Stosując metodę "wstaw" (plug-in), zastępuje się w modelach $g\left(x; K\right)$ oraz $f\left(x; K\right)$ nieznane współczynniki α_{mn}^{g}, β_{mn}^{g}, α_{mn}^{f}, β_{mn}^{f} ich empirycznymi oszacowaniami (8.15). Empiryczne modele falkowe $\hat{g}(x; K)$ oraz $\hat{f}(x; K)$ funkcji $g\left(x\right)$ oraz $f(x)$ przybierają jedną z trzech następujących postaci:

$$\sum_{n=n_{\min}(\varphi,x,M)}^{n_{\max}(\varphi,x,M)}\hat{\alpha}_{Mn}^{F}\varphi_{Mn}\left(x\right) + \sum_{m=M}^{K-1}\sum_{n=n_{\min}(\psi,x,m)}^{n_{\max}(\psi,x,m)}\hat{\beta}_{mn}^{F}\psi_{mn}\left(x\right) \tag{8.16}$$

$$\sum_{n=n_{\min}(\varphi,x,K)}^{n_{\max}(\varphi,x,K)}\hat{\alpha}_{Kn}^{F}\varphi_{Kn}\left(x\right) \tag{8.17}$$

$$N^{-1}\sum_{k=1}^{N}\vartheta_k^{F}q_K\left(x, x_k\right) \tag{8.18}$$

gdzie $F \in \{g, f\}$, $\vartheta_k^{g} = y_k$, $\vartheta_k^{f} = 1$ jak we wzorach (8.15). Modele (8.18) otrzymuje się przez wstawienie (8.15) do wzoru (8.17) i wykorzystanie definicji jądra $q_K\left(x, v\right)$ (por. wzór (8.9)).

Krok 3 Na podstawie odpowiednich modeli ($\hat{g}(x; K)$, $\hat{f}(x; K)$) (to jest pary wzorów (8.16), (8.17) lub (8.18) dla $F \in \{g, f\}$) wyznacza się empiryczny model falkowy $\hat{R}\left(x; K\right)$ nieliniowości $R(x)$ jako iloraz:

$$\hat{R}\left(x; K\right) = \hat{g}(x; K)/\hat{f}\left(x; K\right) \tag{8.19}$$

Modele (algorytmy) (8.16), (8.17) oraz (8.18) mają jednakowe własności asymptotyczne (Sekcja 8.5), ale ze względu na odmienną postać różnią się własnościami użytkowymi. Mianowicie, (8.17) nie zawierają składników falkowych $\psi_{mn}\left(x\right)$ i przez to są najprostsze pod względem obliczeniowym. Dla identyfikacji funkcji $R(x)$ w przedziale $[a,b]$ wymagają obliczenia $2[S + 2^{K}\lceil b-a\rceil]$ współczynników empirycznych $\hat{\alpha}_{Kn}^{g}$, $\hat{\alpha}_{Kn}^{f}$ (por. wzór (8.10)). Modele (8.16)

wymagają obliczenia łącznie większej liczby $2[S\,(K-M+1)+2^K\,\lceil b-a\rceil]$ oszacowań $\hat{\alpha}^g_{Mn}$, $\hat{\beta}^g_{mn}$, $\hat{\alpha}^f_{Mn}$, $\hat{\beta}^f_{mn}$, ale za to są bardziej elastyczne - umożliwiają łatwe zwiększenie skali (precyzji) przez dodanie kolejnych członów. Z kolei modele (8.18) są innym przedstawieniem modeli (8.17), łatwiejszym do analizy (por. [25], [17]).

8.5 Zbieżność i szybkość zbieżności

Empiryczne modele falkowe $\hat{g}(x;K)$, $\hat{f}\,(x;K)$ oraz $\hat{R}\,(x;K)$ zbiegają (w sensie probabilistycznym) do funkcji $g(x), f(x)$ oraz $R(x)$ ze wzrostem liczby pomiarów N, jeśli towarzyszy temu odpowiedni wzrost skali (precyzji) K. Okazuje się, że dla uzyskania punktowej zbieżności modeli falkowych nie jest konieczna żadna szczególna regularność funkcji $R(x)$ oraz $f(x)$, np. ich różniczkowalność, której generalnie się wymaga w przypadku klasycznych modeli ortogonalnych (por. [13], [14], [21], [24]). Modele falkowe zbiegają dla wszystkich ciągłych funkcji $R(x)$ oraz $f(x)$, a zakładając pewną gładkość tych funkcji można dodatkowo określić asymptotyczną szybkość zbieżności i podać rekomendacje dotyczące wyboru skali (poziomu precyzji) K.

8.5.1 Zbieżność modeli falkowych

Załóżmy, że skala K w modelach zależy od liczby pomiarów N w próbie $\{(x_k,y_k)\}$, $k=1,\ldots,N$ to znaczy jest pewną funkcją $K=K(N)$, oraz że spełnione są wszystkie warunki wymienione w Sekcji 8.3. Wówczas zachodzi następujące twierdzenie.

Twierdzenie 8.1 *Jeżeli $K=K(N)$ oraz skala (poziom precyzji) $K(N)$ rośnie wraz z N w taki sposób, że*

$$K(N)\to\infty,\ 2^{K(N)}/N\to 0 \quad gdy \quad N\to\infty \tag{8.20}$$

to dla modeli falkowych (8.16), (8.17), (8.18) zachodzi zbieżność

$$\hat{R}(x;K(N))\to R(x) \ \ wg\ prawdopodobieństwa \tag{8.21}$$

gdy $N\to\infty$ w prawie każdym punkcie $x\in[a,b]$, a w szczególności we wszystkich punktach x przedziału $[a,b]$, gdzie funkcje $R(x)$ oraz $f(x)$ są ciągłe.

Proof. Ze względu na postać modelu $\hat{R}(x;K(N))=\hat{g}(x;K(N))/\hat{f}(x;K(N))$ wystarczy wykazać zbieżność według prawdopodobieństwa modeli $\hat{g}(x;K(N))$ oraz $\hat{f}(x;K(N))$ do funkcji $g(x)=R(x)f(x)$ oraz $f(x)$ (z założenia A7, $f(x)>0$ w przedziale $[a,b]$). Na mocy nierówności Czebyszewa zbieżność taka wynika ze średniokwadratowej zbieżności $\hat{g}(x;K(N))$ oraz $\hat{f}(x;K(N))$ do $g(x)$ oraz $f(x)$. Dla wykazania zbieżności średniokwadratowej zauważamy, że dla modeli $\hat{F}(x;K)\in\{\hat{g}(x;K),\hat{f}(x;K)\}$ zachodzi $E\hat{F}(x;K)=F(x;K)$, gdzie $F(x;K)$ jest

odpowiednim elementem pary $\{g(x;K), f(x;K)\}$. Własność ta wynika generalnie z faktu, że dla $F(x) \in \{g(x), f(x)\}$ empiryczne współczynniki falkowe $\hat{\alpha}^F_{mn}$, $\hat{\beta}^F_{mn}$ we wzorze (8.15) są nieobciążonymi estymatorami współczynników α^F_{mn}, β^F_{mn}, to znaczy $E\hat{\alpha}^F_{mn} = \alpha^F_{mn}$, $E\hat{\beta}^F_{mn} = \beta^F_{mn}$. Ponadto, dla każdego $x \in [a,b]$ oraz każdej ustalonej skali K zachodzi oszacowanie

$$var\left\{\hat{F}(x;K)\right\} \leq (C_F 2^K)N^{-1} \tag{8.22}$$

przy czym stała $C_F > 0$ nie zależy od K oraz N. Wyprowadzenie tej własności dla poszczególnych modeli (8.16), (8.17) oraz (8.18) znaleźć można w pracach [25], [17], [19]. W rezultacie otrzymujemy następujące oszacowanie średniokwadratowego błędu modeli $\hat{F}(x;K) \in \{\hat{g}(x;K), \hat{f}(x;K)\}$, gdzie $AE(F;x;K) = F(x) - F(x;K)$ oznacza punktowy błąd aproksymacji:

$$E\left[\hat{F}(x;K) - F(x)\right]^2 \leq AE^2(F;x;K) + (C_F 2^K)N^{-1} \tag{8.23}$$

Z tego oszacowania oraz własności (8.11) wynika średniokwadratowa zbieżność $\hat{g}(x;K(N)) \to g(x)$ oraz $\hat{f}(x;K(N)) \to f(x)$ gdy $N \to \infty$ przy spełnieniu warunków (8.20). ■

Powyższy rezultat pokazuje, że przy bardzo ogólnych założeniach dotyczących systemu i zakłóceń, podanych w Sekcji 8.3, empiryczny model falkowy $\hat{R}(x;K(N))$ zbiega punktowo (przy $N \to \infty$) do nieliniowości $R(x)$ (prawie wszędzie) i że zbieżność ta zachodzi dla obszernej klasy charakterystyk nieliniowych i rozkładów prawdopodobieństwa wejścia. Wymaga się tu jedynie spełnienia warunków A1, A2 oraz A7. Dla zapewnienia zbieżności praktycznie niezbędny jest więc tylko odpowiedni wzrost skali $K(N)$ w modelach $\hat{g}(x;K(N))$ oraz $\hat{f}(x;K(N))$, połączony ze wzrostem liczby pomiarów N według wzoru (8.20). Warunek ten nie jest zbyt ograniczający i zachodzi dla wszystkich $K(N) = \lfloor c \log_2 N \rfloor$, z dowolnym $0 < c < 1$. Konkretny wybór stałej c decyduje o szybkości zbieżności we wzorze (8.21). Pokażemy, że przy właściwym wyborze tej stałej modele falkowe dzięki swej specyfice mogą zapewnić większą szybkość zbieżności niż klasyczne modele ortogonalne.

8.5.2 Szybkość zbieżności

Dla celów porównawczych, szybkość zbieżności modelu $\hat{R}(x;K(N))$ przedstawimy dla klas gładkości funkcji $R(x)$ oraz $f(x)$ standardowo rozpatrywanych w literaturze i ważnych dla zastosowań. Założymy mianowicie, że lokalnie (tj. w otoczeniu punktu x_0) zachodzi $R(x) \in C^{\nu_R}(x_0)$ oraz $f(x) \in C^{\nu_f}(x_0)$, gdzie $\nu_R, \nu_f \in (0,1]$ lub $\nu_R, \nu_f \in \mathbf{N}$ (zbiór liczb naturalnych), to znaczy że funkcje $R(x)$ oraz $f(x)$ są lokalnie lipschitzowskie (z indeksami ν_R oraz ν_f) lub ν_R, ν_f-krotnie różniczkowalne. Przy $R(x) \in C^{\nu_R}(x_0)$ oraz $f(x) \in C^{\nu_f}(x_0)$ funkcja $g(x) = R(x)f(x) \in C^{\nu_g}(x_0)$ gdzie $\nu_g = \min\{\nu_R, \nu_f\}$, co można łatwo sprawdzić. W dalszej części, $\vartheta_N = O(b_N)$ według prawdopodobieństwa oznacza, że dla ciągu zmiennych losowych $\{\vartheta_N\}$ oraz ciągu liczb dodatnich $\{b_N\}$ zbieżnego do zera,

wyrażenie $r_N(\vartheta_N/b_N)$ zbiega do zera według prawdopodobieństwa gdy $N \to \infty$ dla dowolnego ciągu liczbowego $\{r_N\}$ zbieżnego do zera.

Twierdzenie 8.2 *Niech $R(x) \in C^{\nu_R}(x_0)$, $f(x) \in C^{\nu_f}(x_0)$ oraz niech $(\varphi(x), \psi(x))$ będzie parą wzajemnie odpowiadających sobie funkcji falkowych taką, że falka-matka $\psi(x)$ posiada r_ψ znikających momentów (wzór (8.13)). Asymptotycznie optymalna reguła wyboru skali $K(N)$ w modelu $\hat{R}(x; K(N))$ przy modelach (8.16), (8.17) oraz (8.18) ma w punkcie x_0 postać*

$$K_{opt}(N) = \lfloor (1/(2\lambda_{g\psi} + 1)) \log_2 N \rfloor \tag{8.24}$$

gdzie $\lambda_{g\psi} = \min\{\nu_R, \nu_f, r_\psi + 1\}$. Dla takiego wyboru skali

$$\hat{R}(x_0; K_{opt}(N)) = R(x_0) + O\left(N^{-\lambda_{g\psi}/(2\lambda_{g\psi}+1)}\right) \tag{8.25}$$

według prawdopodobieństwa, i jest to asymptotycznie najlepsza gwarantowana szybkość zbieżności modelu falkowego $\hat{R}(x; K(N))$ w punkcie $x = x_0$ dla pary funkcji $(R(x), f(x))$ oraz pary falek $(\varphi(x), \psi(x))$.

Dowód. Niech $(\hat{F}(x; K(N)), F(x; K(N)), F(x))$ będzie ogólnym elementem zbioru $\{(\hat{g}(x; K(N)), g(x; K(N)), g(x)), (\hat{f}(x; K(N)), f(x; K(N)), f(x))\}$ i załóżmy, że reguła wyboru skali $K(N)$ spełnia warunki podane we wzorze (8.20). Dla założonych klas gładkości funkcji $R(x)$ oraz $f(x)$, oszacowanie asymptotycznego błędu aproksymacji $AE(F; x; K(N)) = F(x) - F(x; K(N))$ (dla dużych N, a tym samym dużych $K(N)$) funkcji $F(x) \in \{g(x), f(x)\}$, $g(x) = R(x)f(x)$, w punkcie $x = x_0$ określa wzór (8.12). Uwzględniając wzór (8.23) otrzymujemy stąd oszacowanie

$$E\left[\hat{F}(x_0; K(N)) - F(x_0)\right]^2 \leq C_{AE,F\psi}^2\, 2^{-2\lambda_{F\psi}K(N)} + (C_F 2^{K(N)})N^{-1}$$

gdzie $\lambda_{F\psi} = \min\{\nu_F, r_\psi + 1\}$ oraz ν_F jest indeksem gładkości funkcji $F(x)$. Z tego oszacowania oraz nierówności Czebyszewa wynika, że

$$\hat{F}(x_0; K(N)) = F(x_0) + O\left(\left(2^{-2\lambda_{F\psi}K(N)} + 2^{K(N)}/N\right)^{1/2}\right) \tag{8.26}$$

według prawdopodobieństwa. Ze wzoru (8.26), wykorzystując (a) fakt, że dla $\lambda_{g\psi} = \min\{\nu_g, r_\psi + 1\} = \min\{\nu_R, \nu_f, r_\psi + 1\}$ oraz $\lambda_{f\psi} = \min\{\nu_f, r_\psi + 1\}$ zachodzi nierówność $\lambda_{g\psi} \leq \lambda_{f\psi}$, oraz (b) własność, że $\hat{g}(x_0; K(N)) = g(x_0) + O(a_N)$ według prawdopodobieństwa oraz $\hat{f}(x_0; K(N)) = f(x_0) + O(b_N)$ według prawdopodobieństwa ($\{a_N\}, \{b_N\}$ - ciągi liczbowe zbieżne do zera) implikuje $\hat{g}(x_0; K(N))/\hat{f}(x_0; K(N)) = g(x_0)/f(x_0) + O(\max\{a_N, b_N\})$ według prawdopodobieństwa ([12], [19]), otrzymujemy zbieżność

$$\hat{R}(x_0; K(N)) = R(x_0) + O\left(\left(2^{-2\lambda_{g\psi}K(N)} + 2^{K(N)}/N\right)^{1/2}\right) \tag{8.27}$$

według prawdopodobieństwa. Ze względu na wzajemnie antagonistyczne zachowanie składowych błędu jako funkcji $K(N)$, największą (optymalną) gwarantowaną szybkość zbieżności sumarycznego błędu modelu otrzymuje się w warunkach zrównoważenia obu składników we wzorze (8.27), to jest dla reguły $K_{opt}(N)$ jak we wzorze (8.24) (stałej $c = 1/(2\lambda_{g\psi} + 1)$, por. Sekcja 8.5.1). Szybkość zbieżności we wzorze (8.25) wynika po wstawieniu wzoru (8.24) do wzoru (8.27). ■

Szybkość zbieżności we wzorze (8.25) zależy poprzez indeks $\lambda_{g\psi}$ od lokalnej regularności mniej gładkiej funkcji w parze $(R(x), f(x))$ oraz liczby znikających momentów falki-matki $\psi(x)$. Wzrost gładkości $R(x)$ oraz $f(x)$ wraz z odpowiednim wyborem pary falek $(\varphi(x), \psi(x))$, dopasowanym do regularności $R(x)$ i $f(x)$, zwiększają szybkość zbieżności. Optymalne dopasowanie falek w modelu $\hat{R}(x; K_{opt}(N))$ do gładkości funkcji $R(x)$ oraz $f(x)$ otrzymuje się, gdy $r_\psi = \lceil \min\{\nu_R, \nu_f\} \rceil - 1$. Dla danych indeksów gładkości ν_R oraz ν_f uzyskujemy wówczas najszybszą zbieżność rzędu $O\left(N^{-\min\{\nu_R,\nu_f\}/(2\min\{\nu_R,\nu_f\}+1)}\right)$. Ponieważ generalnie, dla dowolnego indeksu $\lambda_{g\psi}$, zachodzi $\lambda_{g\psi} \,/\, (2\lambda_{g\psi} + 1) < 1/2$, stąd gwarantowana szybkość zbieżności $O\left(N^{-\lambda_{g\psi}\,/\,(2\lambda_{g\psi}+1)}\right)$ we wzorze (8.25) jest zawsze mniejsza od $O(N^{-1/2})$, to jest parametrycznej szybkości zbieżności według prawdopodobieństwa. Parametryczną szybkość zbieżności $O(N^{-1/2})$ można osiągnąć jedynie w specyficznej sytuacji, gdy jest możliwe dokładne (to znaczy pozbawione błędu aproksymacji) przedstawienie funkcji $g(x)$ oraz $f(x)$ w określonej klasie modeli falkowych V_K. Wtedy problem identyfikacji sprowadza się do parametrycznego zadania estymacji skończonej liczby współczynników falkowych funkcji $g(x)$ oraz $f(x)$ metodą momentów. Przy identyfikacji nieliniowości $R(x)$ w przedziale $[a, b]$ oraz r_ψ znikających momentach falki-matki jest tak dla przykładu przy wielomianowej charakterystyce

$$R(x) = c_0 + c_1 x + ... + c_r x^r$$

stopnia $r \leq r_\psi$ oraz lokalnie stałej ($f(x) = \text{const}$) gęstości prawdopodobieństwa wejścia systemu na przedziale $[a - \varepsilon, b + \varepsilon] \supset [a, b]$, gdzie $\varepsilon \geq s/2^K$, s jest długością nośnika falki $\psi(x)$, a K - przyjętą w modelu skalą. Odnosi się to do modelu falkowego o dowolnej z trzech rozpatrywanych postaci.

8.5.3 Zalety modeli

Ze wzoru (8.25) wynikają następujące zalety ogólnego modelu $\hat{R}(x; K_{opt}(N))$:

1. Jeżeli $R(x)$ oraz $f(x)$ spełniają warunek Lipschitza na przedziale $[a, b]$, wówczas model ten gwarantuje szybkość zbieżności $O(N^{-1/3})$ według prawdopodobieństwa dla dowolnej pary falek $(\varphi(x), \psi(x))$. Jest to najlepsza możliwa nieparametryczna szybkość zbieżności dla lipschitzowskich funkcji $R(x)$ ([30]).

2. Jeżeli $R(x)$ oraz $f(x)$ są ν_R-krotnie oraz ν_f-krotnie różniczkowalne na $[a, b]$ i dodatkowo gęstość wejścia jest gładsza $\nu_f \geq \nu_R$ (co w praktyce zwykle ma miejsce) oraz $r_\psi + 1 \geq \nu_R$, to wówczas model osiąga szybkość zbieżności

$O(N^{-\nu_R/(2\nu_R+1)})$ według prawdopodobieństwa. Jest to najlepsza możliwa nieparametryczna szybkość zbieżności dla różniczkowalnych nieliniowości $R(x)$ ([30]).

Powyższa optymalność oraz swoboda w doborze funkcji falkowych (można zastosować falki o dowolnie wybranej liczbie znikających momentów r_ψ) powodują, iż model $\hat{R}(x; K_{opt}(N))$ może osiągnąć przewagę nad tradycyjnymi modelami ortogonalnymi. Na przykład, dla $R(x)$ oraz $f(x)$ spełniających warunek Lipschitza osiągana przez $\hat{R}(x; K_{opt}(N))$ szybkość zbieżności $O(N^{-1/3})$ jest lepsza od $O(N^{-1/4})$ gwarantowanej przez model trygonometryczny oraz model Hermite'a dla gładszych, różniczkowalnych funkcji $R(x), f(x) \in C^1[a,b]$ (por. [10], [12]). Dla różniczkowalnych $R(x), f(x) \in C^d[a,b]$, $d \in \mathbf{N}$, oraz pary falek $(\varphi(x), \psi(x))$ z $r_\psi \geq d-1$ model falkowy $\hat{R}(x; K_{opt}(N))$ osiąga zbieżność rzędu $O(N^{-d/(2d+1)})$, która ponownie jest szybsza od $O(N^{-(2d-1)/4d})$ gwarantowanej przez model trygonometryczny oraz model Hermite'a (patrz prace cytowane wyżej). W szczególności, dla funkcji $R(x)$ oraz $f(x)$ wielomianowych w otoczeniu punktu x_0 model $\hat{R}(x; K_{opt}(N))$ może osiągnąć parametryczną szybkość zbieżności w tym punkcie.

8.5.4 Zagadnienie wyboru skali

Asymptotycznie optymalna reguła wyboru skali $K_{opt}(N)$ we wzorze (8.24) zależy od indeksów ν_R oraz ν_f lokalnej gładkości funkcji $R(x)$ oraz $f(x)$ i dlatego (i) wymaga zaawansowanej wiedzy wstępnej dotyczącej regularności tych funkcji oraz (ii) może ulegać zmianom, wraz z lokalną gładkością funkcji $R(x)$ oraz $f(x)$, dając niestabilne wartości skali w modelach falkowych. Z tych powodów jej znaczenie praktyczne jest ograniczone. Dlatego bardziej pożądana jest reguła suboptymalna, która nie zależy od lokalnej gładkości funkcji $R(x)$ oraz $f(x)$, gwarantując przy tym satysfakcjonującą szybkość zbieżności modelu $\hat{R}(x; K(N))$ w większości typowych sytuacji. Postulaty te spełnia reguła

$$\bar{K}(N) = \lfloor (1/3) \log_2 N \rfloor \tag{8.28}$$

która zapewnia asymptotyczną szybkość zbieżności

$$\hat{R}(x_0; \bar{K}(N)) = R(x_0) + O\left(N^{-\min\{\nu_R, \nu_f, 1\}/3}\right) \tag{8.29}$$

według prawdopodobieństwa dla każdej pary falek $(\varphi(x), \psi(x))$. Dla klasy funkcji $R(x)$ oraz $f(x)$ co najmniej lipschitzowskich w otoczeniu punktu x_0, to jest takich że $\nu_R, \nu_f \geq 1$, gwarantuje ona zbieżność modelu falkowego $\hat{R}(x_0; \bar{K}(N))$ z szybkością rzędu $O(N^{-1/3})$. Spełnienia warunku $\min\{\nu_R, \nu_f\} \geq 1$ (przynajmniej lokalnie) można oczekiwać w wielu sytuacjach praktycznych, a szybkość zbieżności $O(N^{-1/3})$ w porównaniu z najlepszą osiągalną (parametryczną) zbieżnością $O(N^{-1/2})$ według prawdopodobieństwa można uznać za zadowalającą, biorąc pod uwagę znikomy (nieparametryczny) zakres wstępnej wiedzy o funkcjach $R(x)$ oraz $f(x)$.

8.6 Zakończenie

Falkowe algorytmy identyfikacji wykazują korzystne własności przy rekonstrukcji charakterystyk nieliniowych i zachowują je dla szerokiej klasy systemów, pobudzeń oraz zakłóceń, scharakteryzowanej w Sekcji 8.3. Własności te wynikają głównie z dużej elastyczności modeli falkowych. Elastyczność tę uzyskuje się dzięki swobodzie w wyborze funkcji falkowych oraz adaptacyjności falek, to jest ich zdolności do wpasowywania się w kształt identyfikowanych charakterystyk przy wzroście skali. Cechy te powodują, że modele falkowe mogą z dowolną precyzją odtwarzać szczegóły charakterystyk, być zbieżne punktowo i osiągać optymalną nieparametryczną szybkość zbieżności, a w szczególnych przypadkach (np. dla charakterystyk wielomianowych) także zbieżność typu parametrycznego. Dla zastosowania modeli falkowych wystarczy wyznaczyć empiryczne współczynniki modeli z danych pomiarowych, jednak ze względu na brak jawnej, analitycznej postaci falek, obliczenia wartości funkcji falkowych trzeba przeprowadzić numerycznie. Metody obliczeniowe, wraz z wynikami badań eksperymentalnych, przedstawione są w następnym rozdziale.

8.7 Bibliografia

[1] J. S. Bendat. *Nonlinear System Analysis and Identification from Random Data.* Wiley, New York, 1990.

[2] J. T. Białasiewicz. *Falki i aproksymacje.* WNT, Warszawa, 2000.

[3] S. A. Billings, S. Y. Fakhouri. Identification of systems containing linear dynamic and static non-linear elements. *Automatica*, 18(1):15-26, 1982.

[4] H. W. Chen. Modeling and identification of parallel non-linear systems: structural classification and parameter estimation methods. *Proceedings of the IEEE*, 83(1):39-66, 1995.

[5] D. Coca, S. A. Billings. Non-linear system identification using wavelet multiresolution models. *International Journal of Control*, 74:1718-1736, 2001.

[6] A. Cohen. *Numerical Analysis of Wavelet Methods.* Elsevier, Amsterdam, 2003.

[7] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Philadelphia, 1992.

[8] R. Ghanem, F. Romeo. A wavelet-based approach for model and parameter identification of non-linear systems. *International Journal of Nonlinear Mechanics*, 36:835-859, 2001.

[9] G. B. Giannakis, E. Serpedin. A bibliography on nonlinear system identification. *Signal Processing*, 81:533-580, 2001.

[10] W. Greblicki. Non-parametric orthogonal series identification of Hammerstein systems. *International Journal of Systems Science*, 20(12):2355-2367, 1989.

[11] W. Greblicki. Nieparametryczna identyfikacja systemów. *Archiwum Automatyki i Robotyki*, 36(2):277-290, 1991.

[12] W. Greblicki, M. Pawlak. Fourier and Hermite series estimates of regression functions. *Annals of the Institute of Statistical Mathematics*, 37(3, Part A):443-454, 1985.

[13] W. Greblicki, M. Pawlak. Nonparametric identification of a cascade nonlinear time series system. *Signal Processing*, 22(1):61-75, 1991.

[14] W. Greblicki, M. Pawlak. Nonparametric recovering nonlinearities in block-oriented systems with the help of Laguerre polynomials. *Control-Theory and Advanced Technology*, 10(4, Part 1):771-791, 1994.

[15] R. Haber, L. Keviczky. *Nonlinear System Identification: Input-Output Modeling Approach*. Kluwer, Dordrecht-Boston-London, 1999.

[16] Z. Hasiewicz. Hammerstein system identification by the Haar multiresolution approximation. *International Journal of Adaptive Control and Signal Processing*, 13(8):691-717, 1999.

[17] Z. Hasiewicz. Non-parametric estimation of non-linearity in a cascade time-series system by multiscale approximation. *Signal Processing*, 81(4):791-807, 2001.

[18] Z. Hasiewicz, P. Śliwiński. Identification of non-linear characteristics of a class of block-oriented non-linear systems via Daubechies wavelet-based models. *International Journal of Systems Science*, 33(14):1121-1144, 2002.

[19] 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-Fundamental Theory and Applications*, 52, 2005 (w druku).

[20] S. Kelly, M. Kon, L. Raphael. Pointwise convergence of wavelet expansions. *Bulletin of the American Mathematical Society*, 30(1):87-94, 1994.

[21] A. Krzyżak. Identification of discrete Hammerstein systems by the Fourier series regression estimate. *International Journal of Systems Science*, 20(9):1729-1744, 1989.

[22] A. Krzyżak, J. Z. Sąsiadek, B. Kégl. Non-parametric identification of dynamic non-linear systems by a Hermite series approach. *International Journal of Systems Science*, 32(10):1261-1285, 2001.

[23] S. G. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.

[24] M. Pawlak. On the series expansion approach to the identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, AC-36:763-767, 1991.

[25] 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.

[26] E. Rafajłowicz. Nonparametric orthogonal series estimators of regression: A class attaining the optimal convegence rate in L_2. *Statistics and Probability Letters*, 5(3):219-224, 1987.

[27] E. Rafajłowicz. Nonparametric least squares estimation of a regression function. *Statistics*, 19(3):349-358, 1988.

[28] L. Rutkowski. Application of multiple Fourier series to identification of multivariate non-stationary systems. *International Journal of Systems Science*, 20:1993-2002, 1989.

[29] L. Rutkowski. Multiple Fourier series procedures for extraction of nonlinear regressions from noisy data. *IEEE Transactions on Signal Processing*, 41:3062-3065, 1993.

[30] C. J. Stone. Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, 8(6):1348-1360, 1980.

[31] N. Sureshbabu, J. A. Farrell. Wavelet-based system identification for nonlinear control. *IEEE Transactions on Automatic Control*, 44(2):412-417, 1999.

[32] P. Wojtaszczyk. *Teoria falek*. PWN, Warszawa, 2000.