

---

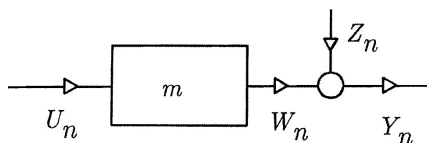
## 7. Nieparametryczne metody identyfikacji systemów

Włodzimierz Greblicki

---

W identyfikacji systemów, podobnie jak w każdym procesie wnioskowania, korzysta się z informacji apriorycznej oraz danych pomiarowych. Informacja aprioryczna dotyczy tego co jest przedmiotem wnioskowania, a więc np. odpowiedzi impulsowej dynamicznego systemu liniowego, lub charakterystyki nieliniowego systemu statycznego. Posiada się ją już przed przystąpieniem do pomiarów.

Dla przykładu, załóżmy że identyfikuje się dyskretny, stacjonarny, statyczny system nieliniowy o wejściu  $u$  i wyjściu  $y$  o nieznannej charakterystyce  $m$ , tzn. taki, że  $w = m(u)$ , rys. 7.1. Ponieważ pomiar sygnału wyjściowego wykonywany jest w obecności zakłócenia  $z$ , wynikiem pomiarów jest ciąg obserwacji  $(U_1, Y_1), (U_2, Y_2), \dots, (U_n, Y_n), \dots$ , gdzie  $Y_n = m(U_n) + Z_n$ .



Rysunek 7.1. Dyskretny system statyczny o charakterystyce  $m$ .

Informacja aprioryczna dotyczy charakterystyki  $m$  i określa co o niej wiadomo z góry, tzn. przed wykonaniem pomiarów. Niekiedy wiadomo, że charakterystyka ma postać:  $m(u) = \sum_{i=0}^k c_i \varphi_i(u)$ , gdzie  $\varphi_0, \dots, \varphi_k$  są ustalonymi i znanymi funkcjami. Prawdziwe wartości parametrów  $c_0, \dots, c_k$ , czyli współczynników rozwinięcia w skończony szereg, nie są znane i należy je wyznaczyć w procesie identyfikacji. Mówimy, że informacja aprioryczna jest wtedy parametryczna. Jest tak zawsze, gdy nasza wiedza aprioryczna polega na tym, że nie znamy jedynie prawdziwych wartości skończonej i znanej liczby parametrów.

Jeśli informacja wstępna nie jest tej natury, to mówimy, że jest ona nieparametryczna. Jest tak zawsze, gdy np. o charakterystyce  $m$  wiemy, że  $m(u) = \sum_{i=0}^{\infty} c_i \varphi_i(u)$ , gdzie  $\varphi_0, \varphi_1, \dots$  są ustalonymi i znanymi nam funkcjami, a prawdziwe wartości nieskończonego zestawu parametrów  $c_0, c_1, \dots$  nie są znane. Informacja aprioryczna jest nieparametryczna także, jeśli przykładowo  $m(u) = \sum_{i=0}^k c_i \varphi_i(u)$ , gdzie  $k$  jest nieznanne. Podobnie jest, gdy o charakterystyce  $m$  wiadomo jedynie, że jest ona np. funkcją ciągłą, albo że np.  $m(u) \leq d_0 + d_1|u|$ , gdzie  $d_0$  i  $d_1$  są pewnymi, nieznanymi liczbami. Informacja nieparametryczna jest zatem uboższa od parametrycznej. Algorytmy opracowane w celu identyfikacji systemu w takich warunkach nazywają się nieparametrycznymi.

W świetle powyższego jest oczywiste, że algorytmy parametryczne i nieparametryczne są przeznaczone do stosowania w zupełnie różnych sytuacjach i nie konkurują między sobą. Wprost przeciwnie, można powiedzieć, że uzupełniają się, ponie-

waż - co pokazaliśmy na przykładzie - stosuje się je do wnioskowania o charakterystyce systemu w całkowicie różnych okolicznościach, tzn. przy całkowicie różnych rodzajach informacji apriorycznej.

Nieparametryczne metody są skuteczne wtedy, gdy informacja aprioryczna o systemie jest niewielka, gdy algorytmy parametryczne zawodzą z tego powodu. Są one zatem ważnym narzędziem tam, gdzie nasza wiedza o systemie jest tak uboga, że nie wystarcza do stosowania klasycznych metod parametrycznych.

W dalszych rozważaniach skoncentrujemy się na dwóch rodzajach systemów, a mianowicie na systemie Hammersteina i Wienera. Struktury takie są wykorzystywane do modelowania różnych procesów, np. w przemyśle chemicznym, czy też biologii.

## 7.1. System Hammersteina

### 7.1.1. Wprowadzenie

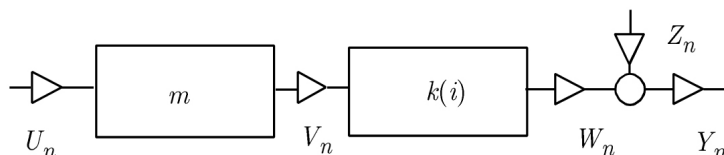
Dyskretny system Hammersteina przedstawiony na rys. 7.2 składa się z dwóch podsystemów, nieliniowego statycznego i liniowego dynamicznego. Nieliniowy ma charakterystykę  $m$ , co oznacza, że  $V_n = m(U_n)$ . O charakterystyce tej wiadomo, że

$$|m(u)| \leq d_0 + d_1|u|,$$

gdzie  $d_0$  i  $d_1$  są nieznanymi. Podsystem liniowy jest stabilny opisywany jest równaniem stanu

$$\begin{cases} X_{n+1} = AX_n \\ W_n = c^T X_n, \end{cases} \quad (7.1)$$

gdzie  $X_n$  jest wektorem stanu w chwili  $n$ . Ani jego wymiar, ani macierz  $A$ , ani wektory  $b$  i  $c$  nie są znane. Wiadomo jedynie, że jest on stabilny. Jego odpowiedź impulsowa  $k_n = c^T A^{n-1} b$ ,  $n = 1, 2, \dots$ , jest zatem zupełnie nieznaną. Informacja aprioryczna o obydwu podsystemach jest zatem nieparametryczna.



Rys. 7.2. System Hammersteina.

System pobudzany jest losowym, stacjonarnym, białym szumem. Taki sam charakter ma zakłócenie  $Z_n$  o zerowej wartości średniej. Identyfikuje się go na podstawie obserwacji

$$(U_1, Y_1), (U_2, Y_2), \dots, (U_n, Y_n)$$

wykonanych na wejściu i wyjściu całego systemu. Sygnał  $V_n$  łączący podsystemy nie jest mierzony.

Podstawą dla konstrukcji algorytmów estymacji charakterystyki  $m$  jest to, że (przy założeniu, że  $Em(U_0) = 0$ )

$$E\{Y_{n+1}|U_n = u\} = \mu(u),$$

gdzie  $\mu(u) = k_1 m(u)$ , co oznacza, że - z dokładnością do nieznannej stałej  $k_1 = c^T b$  - charakterystyka ta jest funkcją regresji. Dla identyfikacji podsystemu nieliniowego można zatem wykorzystać nieparametryczne metody estymacji funkcji regresji. Uwagę skoncentrujemy na dwóch klasach algorytmów, a mianowicie estymatorach jądrowych i ortogonalnych.

Warto w tym miejscu zauważyć, iż to że jesteśmy w stanie wykryć jedynie  $\mu$ , czyli  $m$  z dokładnością do nieznanego mnożnika, wynika z szeregowej struktury systemu i tego, że wewnętrzny sygnał systemu  $V_n$  nie jest mierzony.

Dla ułatwienia, przez  $f(u)$  będziemy oznaczać gęstość prawdopodobieństwa sygnału  $U_n$ ,  $f(u, y)$  będzie gęstością pary  $(U_n, Y_n)$ , natomiast  $f(y|u)$  gęstością warunkową  $Y_n$  względem  $U_n$ . Zamiast  $U_0$  będziemy niekiedy pisać  $U$ . Ponadto,  $g(u) = \mu(u)f(u)$ .

### 7.1.2. Algorytmy jądrowe

W algorytmach jądrowych wykorzystuje się to, że dla odpowiednio wybranej funkcji jądrowej  $K$ , ma miejsce poniższa zbieżność:

$$\lim_{h \rightarrow 0} \int \varphi(v) \frac{1}{h} K\left(\frac{u-v}{h}\right) dv = \varphi(u) \int K(v) dv. \quad (7.2)$$

Zachodzi ona w każdym punkcie  $u$  ciągłości funkcji  $\varphi$ , jeśli

$$\sup_v |K(v)| < \infty \quad (7.3)$$

oraz

$$\lim_{|v| \rightarrow \infty} vK(v) = 0. \quad (7.4)$$

Przykładami takich jąder są np. funkcja prostokątna

$$K(v) = \begin{cases} 1/2, & \text{dla } |v| \leq 1 \\ 0, & \text{dla } |v| > 1, \end{cases}$$

lub  $1/\pi(1+v^2)$ ,  $(1/2)e^{-|v|}$ ,  $(1/\sqrt{2\pi})e^{-v^2/2}$ .

Powyższą własność wykorzystuje się do estymacji  $f(u)$  oraz  $g(u)$ . Z (7.2) wynika bowiem, że

$$\lim_{h \rightarrow 0} \int f(v) \frac{1}{h} K\left(\frac{u-v}{h}\right) dv = f(u),$$

jesli ponadto

$$\int K(v) dv = 1. \quad (7.5)$$

Zauważając teraz, że

$$\int f(v) \frac{1}{h_n} K\left(\frac{u-v}{h_n}\right) dv = \frac{1}{h_n} EK\left(\frac{u-U}{h_n}\right), \quad (7.6)$$

możemy następująco oszacować wyrażenie po prawej stronie:

$$f_n(u) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{u-U_i}{h_n}\right),$$

gdyż

$$Ef_n(u) = \frac{1}{h_n} EK \left( \frac{u - U}{h_n} \right).$$

Dzięki (7.2) i (7.5) możemy zatem napisać, że  $\lim_{n \rightarrow \infty} Ef_n(u) = f(u)$ , w każdym punkcie  $u$  ciągłości gęstości  $f$ , jeśli tylko

$$\lim_{n \rightarrow \infty} h_n = 0. \quad (7.7)$$

Analizując wariancję stwierdzamy, że

$$\begin{aligned} \text{var}[f_n(u)] &= \frac{1}{nh_n^2} \text{var} \left[ K \left( \frac{u - U_0}{h_n} \right) \right] \leq \frac{1}{nh_n^2} EK^2 \left( \frac{u - U_0}{h_n} \right) \\ &= \frac{1}{nh_n} \frac{1}{h_n} \int K^2 \left( \frac{u - v}{h_n} \right) f(v) dv. \end{aligned}$$

Ponieważ

$$\lim_{h \rightarrow 0} \frac{1}{h} \int K^2 \left( \frac{u - v}{h} \right) f(v) dv = f(u) \int K^2(v) dv,$$

zatem  $\lim_{n \rightarrow \infty} nh_n \text{var}[f_n(u)] = f(u) \int K^2(v) dv$ , jeśli

$$\lim_{n \rightarrow \infty} nh_n = \infty. \quad (7.8)$$

W rezultacie

$$\lim_{n \rightarrow \infty} E(f_n(x) - f(x))^2 = 0,$$

co oznacza zgodność estymatora gęstości.

Przechodząc do  $g(u)$  zauważamy, że  $\mu(u) = \int y f(y|u) dy$ , skąd wynika, że  $g(u) = \int y f(u, y) dy$ . Pamiętając, że

$$\lim_{h \rightarrow 0} \int g(v) \frac{1}{h} K \left( \frac{u - v}{h} \right) dv = g(u),$$

zauważamy, iż

$$\begin{aligned} \int g(v) \frac{1}{h_n} K \left( \frac{u - v}{h_n} \right) dv &= \frac{1}{h_n} \int \int y K \left( \frac{u - v}{h_n} \right) f(u, y) dv dy \\ &= \frac{1}{h_n} E \left\{ Y K \left( \frac{u - U}{h_n} \right) \right\}, \end{aligned}$$

którą to wielkość można estymować następująco:

$$\frac{1}{nh_n} \sum_{i=1}^n Y_i K \left( \frac{u - U_i}{h_n} \right)$$

jest, w świetle powyższego, naturalnym estymatorem nieznanego  $g(u)$ .

W rezultacie jądrowy algorytm estymacji  $\mu(u) = g(u)/f(u)$ , czyli  $k_1 m(u)$ , przyjmuje postać:

$$\mu_n(u) = \frac{\sum_{i=1}^n Y_i K \left( \frac{u - U_i}{h_n} \right)}{\sum_{i=1}^n K \left( \frac{u - U_i}{h_n} \right)}, \quad (7.9)$$

patrz także {13}, [14], [15], [19], [27], [30]. Jego zbieżność jest przedmiotem poniższego twierdzenia:

**Twierdzenie 7.1** Jeśli funkcja jądrowa  $K$  spełnia warunki (7.3)–(7.5), a ciąg liczbowy warunki (7.7) i (7.8), to

$$\mu_n(u) \xrightarrow{n \rightarrow \infty} \mu(u) \text{ według prawdopodobieństwa}$$

w każdym punkcie  $u$ , w którym  $m$  i  $f$  są funkcjami ciągłymi oraz  $f(u) > 0$ .

Interesujące są rekurencyjne wersje algorytmu (7.9), patrz [11], [16], [28]. W poniższym

$$\hat{\mu}_n(u) = \frac{\sum_{i=1}^n \frac{1}{h_i} Y_i K\left(\frac{u - U_i}{h_i}\right)}{\sum_{i=1}^n \frac{1}{h_i} K\left(\frac{u - U_i}{h_i}\right)} \quad (7.10)$$

licznik i mianownik można wyliczać rekurencyjnie. Oznaczając bowiem jego licznik i mianownik jako  $\hat{g}_n(u)$  i  $\hat{f}_n(u)$  zauważamy bowiem, że

$$\hat{g}_n(u) = \hat{g}_{n-1}(u) - \frac{1}{n} \left( \hat{g}_{n-1}(u) - Y_n \frac{1}{h_n} K\left(\frac{u - U_n}{h_n}\right) \right),$$

$$\hat{f}_n(u) = \hat{f}_{n-1}(u) - \frac{1}{n} \left( \hat{f}_{n-1}(u) - \frac{1}{h_n} K\left(\frac{u - U_n}{h_n}\right) \right).$$

Warunki zbieżności algorytmu podaje następujące twierdzenie:

**Twierdzenie 7.2** Jeśli funkcja jądrowa  $K$  spełnia warunki (7.3)–(7.5), a ciąg liczbowy warunek (7.7) i ponadto

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \frac{1}{h_i} = 0,$$

to

$$\hat{\mu}_n(u) \xrightarrow{n \rightarrow \infty} \mu(u) \text{ według prawdopodobieństwa}$$

w każdym punkcie  $u$ , w którym  $m$  i  $f$  są funkcjami ciągłymi oraz  $f(u) > 0$ .

Odpowiednia, całkowicie rekurencyjna wersja algorytmu ma natomiast postać jak poniżej:

$$\tilde{\mu}_n(u) = \tilde{\mu}_{n-1}(u) - \gamma_n \frac{1}{h_n} K\left(\frac{u - U_n}{h_n}\right) [\tilde{\mu}_{n-1}(u) - Y_n],$$

gdzie  $\{\gamma_n\}$  jest odpowiednio dobranym ciągiem, [10]. Jego zbieżność jest przedmiotem kolejnego twierdzenia.

**Twierdzenie 7.3** Jeśli funkcja jądrowa  $K$  spełnia warunki (7.3)–(7.5), a ciąg liczbowy warunek (7.7) i ponadto

$$\lim_{n \rightarrow \infty} \gamma_n = 0, \quad \sum_{n=1}^{\infty} \gamma_n = \infty \quad \text{oraz} \quad \lim_{n \rightarrow \infty} \frac{\gamma_n}{h_n} = 0,$$

to

$$\lim_{n \rightarrow \infty} E(\tilde{\mu}_n(u) - \mu(u))^2 = 0$$

w każdym punkcie  $u$ , w którym  $m$  i  $f$  są funkcjami ciągłymi oraz  $f(u) > 0$ .

Warto zwrócić uwagę, że wszystkie algorytmy zbiegają się w każdym punkcie  $u$ , w którym zarówno nieznaną charakterystykę  $m$  jak i gęstość  $f$  jest ciągła i ponadto  $f(u) > 0$ . Pierwsze dwa zbiegają się według prawdopodobieństwa, ostatni według średniej drugiego rzędu. Ponadto  $\mu_n(u)$  i  $\hat{\mu}_n(u)$  mają postać ułamka, nie ma jej natomiast  $\tilde{\mu}_n(u)$ .

Odnosnie do szybkości zbieżności, jeśli  $m$  i  $f$  są dwa razy różniczkowalne oraz  $h_n \sim n^{-1/4}$ , to

$$|\mu_n(u) - \mu(u)| = O(n^{-4/5}) \quad \text{w/g prawdopodobieństwa.}$$

Taką samą szybkość ma także algorytm  $\hat{\mu}_n(u)$ . Za pełną rekurencję płaci się jednak pewną cenę, bowiem

$$|\tilde{\mu}_n(u) - \mu(u)| = O(n^{-4/5+\varepsilon}) \quad \text{w/g prawdopodobieństwa,}$$

gdzie  $\varepsilon > 0$ , co oznacza pewne pogorszenie szybkości zbieżności.

Podane rezultaty dotyczące szybkości zbieżności wymagają aby nie tylko charakterystyka  $m$  lecz także gęstość  $f$  sygnału wejściowego były funkcjami odpowiednio gładkimi, aby były dwukrotnie różniczkowalne. Wady tej nie mają algorytmy, w których wykorzystuje się tzw. statystyki porządkowe, patrz [4], [20]. Szybkość ich zbieżności nie zależy od gładkości gęstości  $f$ . Innymi słowy, brak jej różniczkowości jej nie pogarsza.

Algorytmy jądrowe, po odpowiedniej modyfikacji, można także z powodzeniem stosować do wykrywania nieliniowości także w systemach o czasie ciągłym, [7], [9].

### 7.1.3. Algorytmy ortogonalne

Zauważmy, że funkcję gęstości  $f$  można rozwinąć w ortogonalny szereg Hermite'a  $h_k, k = 0, 1, \dots$ , co zapiszemy jako

$$f(u) \sim \sum_{k=0}^{\infty} \alpha_k h_k(u),$$

gdzie  $\alpha_k = \int h_k(u) f(u) du = E h_k(U)$ . Funkcja  $h_k$  określona jest wzorem  $h_k(u) = 2^{-1/2} (k!)^{-1/2} \pi^{-1/4} H_k(u) e^{-u^2/2}$ , gdzie  $H_k$  jest  $k$ -tym wielomianem Hermite'a. Spełnia on wzór Rodriguesa:  $H_k(u) = e^{u^2} (-1)^k (d^k / du^k) e^{-u^2}$ , czyli  $H_0(u) = 1$ ,  $H_1(u) = 2u$ ,  $H_2(u) = 4u^2 - 2$ ,  $H_3(u) = 8u^3 - 12u$  itd. Ponieważ współczynnik rozwinięcia  $\alpha_k$  można szacować następująco  $\check{\alpha}_k = n^{-1} \sum_{i=1}^n h_k(U_i)$ , to

$$\check{f}_n(u) = \sum_{k=0}^{N(n)} \check{\alpha}_k h_k(u)$$

można przyjąć jako estymator gęstości  $f(u)$ . W estymatorze tym  $\{N(n)\}$  jest ciągiem liczb całkowitych takim, że

$$\lim_{n \rightarrow \infty} N(n) = \infty. \quad (7.11)$$

Podobnie,

$$g(u) \sim \sum_{k=0}^{\infty} \beta_k h_k(u),$$

gdzie  $\beta_k = \int g(u) h_k(u) du$  skąd wynika, że  $\beta_k = \int \int y h_k(u) f(u, y) du dy = E Y_n h_k(U_n)$ . Zatem  $\hat{\beta}_k = n^{-1} \sum_{i=1}^n Y_i h_k(U_i)$  jest naturalnym oszacowaniem współczynnika  $\beta_k$  a

$$\check{g}_n(u) = \sum_{k=0}^{N(n)} \hat{\beta}_k h_k(u)$$

estymatorem  $g(u)$ . W rezultacie, ortogonalny estymator nieliniowości przyjmuje postać:

$$\check{\mu}_n(u) = \frac{\sum_{k=0}^{N(n)} \check{\beta}_k h_k(u)}{\sum_{k=0}^{N(n)} \check{\alpha}_k h_k(u)}.$$

Chociaż zasada konstrukcji tego estymatora jest odmienna od jądrowego, ma on jednak pewne cechy podobne. Oznaczając bowiem przez  $k_N(u, y)$  tzw. jądro szeregu Hermite'a

$$k_N(u, y) = \sum_{k=0}^N h_k(u) h_k(y),$$

zauważamy, że można go przedstawić w postaci zbliżonej do estymatora jądrowego, a mianowicie:

$$\check{\mu}_n(u) = \frac{\sum_{k=0}^{N(n)} Y_i k_{N(n)}(u, U_i)}{\sum_{k=0}^{N(n)} k_{N(n)}(u, U_i)}.$$

O jego zbieżności stanowi poniższe twierdzenie, [18]:

**Twierdzenie 7.4** Jeśli ciąg liczbowy  $\{N(n)\}$  spełnia (7.11) i ponadto

$$\lim_{n \rightarrow \infty} \frac{N^{1/2}(n)}{n} = \infty,$$

to

$$\check{\mu}_n(u) \xrightarrow[n \rightarrow \infty]{} m(u) \text{ według prawdopodobieństwa}$$

w każdym punkcie  $u$ , w którym  $m$  i  $f$  są różniczkowalne oraz  $f(u) > 0$ .

Jeśli  $m$  i  $f$  mają dwie pochodne oraz  $N(n) \sim n^{1/2}$ , to

$$|\check{\mu}_n(u) - \mu(u)| = O(n^{-3/8}) \text{ w/g prawdopodobieństwa.}$$

Wykorzystanie innych szeregów ortogonalnych np. trygonometrycznego, Legendre'a czy Laguerre'a prowadzi do kolejnych algorytmów, [1], [17], [21], [26], [29], [31], [32]. Algorytmy, w których zastosowano falki zbadano w pracach [23], [24], [25], [33].

## 7.2. System Wienera

### 7.2.1. Wprowadzenie

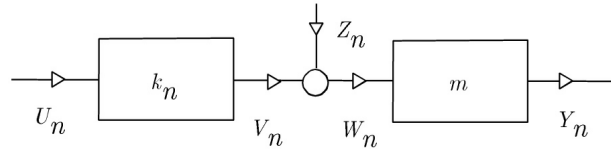
W systemie Wienera, rys. 7.3, liniowy podsystem dynamiczny o opisie (7.1) poprzedza statyczną część nieliniową o charakterystyce  $m$ . Jeśli sygnał wejściowy i zakłócenia są białymi szumami o rozkładzie Gaussa, to

$$E\{U_n | Y_{n-1} = y\} = \nu(y), \quad (7.12)$$

gdzie  $\nu(y) = \alpha_0 m^{-1}(y)$ , przy założeniu, że  $m$  jest funkcją odwracalną, [2]. W równości powyższej

$$\alpha_1 = k_1 \frac{\sigma_U^2}{\sigma_U^2 \sum_{i=0}^{\infty} k_i^2 + \sigma_Z^2},$$

gdzie  $\sigma_U^2$  i  $\sigma_Z^2$  są wariancjami sygnałów  $U_n$  i  $Z_n$ . Zatem estymując funkcję regresji (7.12) można wykryć, z dokładnością do stałego mnożnika  $\alpha_1$ , odwrotność nieliniowej charakterystyki.



Rys. 7.3. System Wienera.

### 7.2.2. Algorytmy jądrowe

Jądrowy algorytm przyjmuje zatem postać

$$\nu_n(y) = \frac{\sum_{i=1}^n U_i K\left(\frac{y - Y_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{y - Y_i}{h_n}\right)}, \quad (7.13)$$

patrz [2], [5]. Oznaczmy jeszcze przez  $f$  gęstość prawdopodobieństwa sygnału wyjściowego  $Y_n$ .

**Twierdzenie 7.5** Jeśli funkcja jądrowa  $K$  spełnia warunki (7.3)–(7.5) a ciąg liczbowy warunek (7.7) i ponadto

$$\lim_{n \rightarrow \infty} n h_n^2 = \infty,$$



to

$$\nu_n(y) \xrightarrow{n \rightarrow \infty} \nu(y) \text{ według prawdopodobieństwa}$$

w każdym punkcie  $y$ , w którym  $m^{-1}$  i  $f$  są funkcjami ciągłymi oraz  $f(y) > 0$ .

Interesujące są rekurencyjne wersje algorytmu (7.9), patrz [8], [11]. W poniższym

$$\hat{\nu}_n(u) = \frac{\sum_{i=1}^n \frac{1}{h_i} U_i K\left(\frac{y - Y_i}{h_i}\right)}{\sum_{i=1}^n \frac{1}{h_i} K\left(\frac{y - Y_i}{h_i}\right)}$$

licznik i mianownik można wyliczać rekurencyjnie. Oznaczając bowiem jego licznik i mianownik jako  $\hat{g}_n(u)$  i  $\hat{f}_n(u)$  zauważamy bowiem, że

$$\begin{aligned} \hat{g}_n(y) &= \hat{g}_{n-1}(y) - \frac{1}{n} \left( \hat{g}_{n-1}(y) - U_n \frac{1}{h_n} K\left(\frac{y - Y_n}{h_n}\right) \right), \\ \hat{f}_n(y) &= \hat{f}_{n-1}(y) - \frac{1}{n} \left( \hat{f}_{n-1}(y) - \frac{1}{h_n} K\left(\frac{y - Y_n}{h_n}\right) \right). \end{aligned}$$

Warunki zbieżności algorytmu podane następujące twierdzenie:

**Twierdzenie 7.6** Jeśli funkcja jądrowa  $K$  spełnia warunki (7.3)–(7.5) a ciąg liczbowy warunek (7.7) i ponadto

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \frac{1}{h_i^2} = 0,$$

to

$$\hat{\nu}_n(y) \xrightarrow{n \rightarrow \infty} \nu(y) \text{ według prawdopodobieństwa}$$

w każdym punkcie  $y$ , w którym  $m^{-1}$  i  $f$  są funkcjami ciągłymi oraz  $f(y) > 0$ .

Przechodząc do szybkości zbieżności, należy stwierdzić, że jeśli  $m^{-1}$  i  $f$  są dwa razy różniczkowalne oraz  $h_n \sim n^{-1/6}$ , to

$$|\nu_n(u) - \nu(u)| = O(n^{-1/3}) \text{ w/g prawdopodobieństwa.}$$

Taką samą szybkość ma także algorytm  $\hat{\nu}_n(u)$ .

Algorytm (7.13) zbiega się także, jeśli sygnał wejściowy nie jest białym szumem, [12].

Algorytmy jądrowe, po odpowiedniej modyfikacji, można także z powodzeniem stosować do wykrywania nieliniowości w systemach o czasie ciągłym, [6], [9].

Do wykrywania nieliniowości można także zastosować algorytmy ortogonalne, [3].

### 7.3. Zakończenie

Metody nieparametryczne mają zastosowanie tam gdzie informacja aprioryczna o systemie jest niewielka, gdzie podejście parametryczne zawodzi. Stanowi to o istotnej ich wartości aplikacyjnej.

**Literatura**

- [1] W. Greblicki. Nonparametric orthogonal series identification of Hammerstein systems. *International Journal of Systems Science*, 20(12):2355–2367, 1989.
- [2] W. Greblicki. Nonparametric identification of Wiener systems. *IEEE Transactions on Information Theory*, 38(5):1487–1493, 1992.
- [3] W. Greblicki. Nonparametric identification of Wiener systems by orthogonal series. *IEEE Transactions on Automatic Control*, 39(10):2077–2086, 1994.
- [4] W. Greblicki. Nonlinearity estimation in Hammerstein systems based on ordered observations. *IEEE Transactions on Signal Processing*, 44(5):1224–1233, 1996.
- [5] W. Greblicki. Nonparametric approach to Wiener system identification. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 44(6):538–545, 1997.
- [6] W. Greblicki. Continuous-time Wiener system identification. *IEEE Transactions on Automatic Control*, 43(10):1488–1493, 1998.
- [7] W. Greblicki. Continuous-time Hammerstein system identification. *IEEE Transactions on Automatic Control*, 45(6):1232–1236, 2000.
- [8] W. Greblicki. Recursive identification of Wiener systems. *International Journal of Applied Mathematics and Computer Science*, 11(4):977–991, 2001.
- [9] W. Greblicki. Recursive identification of continuous-time Hammerstein systems. *International Journal of Systems Science*, 33(12):969–977, 2002.
- [10] W. Greblicki. Stochastic approximation in nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 47(11):1800–1810, 2002.
- [11] W. Greblicki. Hammerstein system identification with stochastic approximation. *International Journal of Modelling and Simulation*, 24(3):131–138, 2004.
- [12] W. Greblicki. Nonlinearity recovering in Wiener system driven with correlated signal. *IEEE Transactions on Automatic Control*, 2004, w druku.
- [13] W. Greblicki and M. Pawlak. Identification of discrete Hammerstein systems using kernel regression estimates. *IEEE Transactions on Automatic Control*, AC-31(1):74–77, 1986.
- [14] W. Greblicki and M. Pawlak. Hammerstein system identification by nonparametric regression estimation. *International Journal of Control*, 45(1):343–354, 1987.
- [15] W. Greblicki and M. Pawlak. Nonparametric identification of Hammerstein systems. *IEEE Transactions on Information Theory*, IT-35(2):409–418, 1989.
- [16] W. Greblicki and M. Pawlak. Recursive identification of Hammerstein systems. *Journal of the Franklin Institute*, 326(4):461–481, 1989.

- [17] W. Greblicki and M. Pawlak. Nonparametric identification of a cascade non-linear time series system. *Signal Processing*, 22:61–75, 1991.
- [18] W. Greblicki and M. Pawlak. Nonparametric identification of a particular non-linear time series system. *IEEE Transactions on Signal Processing*, 40(4):985–989, 1992.
- [19] W. Greblicki and M. Pawlak. Cascade non-linear system identification by a non-parametric method. *International Journal of Systems Science*, 25(1):129–153, 1994.
- [20] W. Greblicki and M. Pawlak. Dynamic system identification with order statistics. *IEEE Transactions on Information Theory*, 40(5):1474–1489, 1994.
- [21] W. Greblicki and M. Pawlak. Nonparametric recovering nonlinearities in block-oriented systems with the help of Laguerre polynomials. *Control-Theory and Advanced Technology*, 10(4, Part A):771–791, 1994.
- [22] Z. Hasiewicz. Hammerstein system identification by the Haar multiresolution approximation. *International Journal of Adaptive Control and Signal Processing*, 13:691–717, 1999.
- [23] Z. Hasiewicz. Modular neural networks for non-linearity recovering by the Haar approximation. *Neural Networks*, 13:1107–1133, 2000.
- [24] Z. Hasiewicz. Non-parametric estimation of non-linearity in a cascade time-series system by multiscale approximation. *Signal Processing*, 81:791–807, 2001.
- [25] Z. Hasiewicz and 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*, 33:1121–1141, 2002.
- [26] A. Krzyżak. Identification of discrete Hammerstein systems by the Fourier series regression estimate. *International Journal of Systems Science*, 20:1729–1744, 1989.
- [27] A. Krzyżak. On estimation of a class of nonlinear systems by the kernel regression estimate. *IEEE Transactions on Information Theory*, 36(1):141–152, February 1990.
- [28] A. Krzyżak. Global convergence of the recursive kernel regression estimates with applications in classification and nonlinear system estimation. *IEEE Transactions on Information Theory*, 38(4):1323–1338, July 1992.
- [29] A. Krzyżak. On nonparametric estimation of nonlinear dynamic systems by the Fourier series estimate. *Signal Processing*, 52:299–321, 1996.
- [30] A. Krzyżak and M. A. Partyka. On identification of block orientated systems by non-parametric techniques. *International Journal of Systems Science*, 24(6):1049–1066, 1993.
- [31] A. Krzyżak, J. Z. Sasiadek, and 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.

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