



Wielomianowe modele Hammersteina

W tym rozdziale dokonamy przeglądu siedmiu metod identyfikacji układów Hammersteina, które wykorzystują modele Hammersteina z wielomianowym modelem elementu nieliniowego. Omówione metody wykorzystują modele liniowego układu dynamicznego w postaci modeli przenoszenia impulsów [21, 27, 34, 63, 120, 161] lub rozwinięcia Laguerre'a odpowiedzi impulsowej [156]. Parametry systemu Wienera są szacowane za pomocą różnych procedur, takich jak zwykle najmniejsze kwadraty [27], iteracyjne najmniejsze kwadraty [63, 120, 161], iteracyjna korelacja i najbardziej strome zejście [156], błąd predykcji [34] i regresja pseudolinearna [21].

5.1 Nieiteracyjna identyfikacja systemów Hammersteina metodą najmniejszych kwadratów

Jedna z najbardziej znanych metod identyfikacji systemów Hammersteina opiera się na przekształceniu nieliniowego w parametрах problemu identyfikacji SISO w liniowy w parametрах problem identyfikacji MISO i zastosowaniu algorytmu parametrów najmniejszych kwadratów. Metoda ta została pierwotnie zaproponowana przez Changa i Luusa [27]. Krótki opis metody znajduje się poniżej.

Najpierw wprowadzamy równoległy wielomianowy model Hammersteina, tj. model z częścią liniową zdefiniowaną przez funkcję transferu $\hat{B}(q^{-1})/\hat{A}(q)$;

$$\hat{y}(n) = \frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} \hat{v}(n), \quad (5.1)$$

$$\hat{B}(q^{-1}) = \hat{b}_0 q_1^{-1} + \dots + \hat{b}_n q_{nb}^{-nb}, \quad (5.2)$$

$$\hat{A}(q^{-1}) = 1 + \hat{a}_1 q_1^{-1} + \dots + \hat{a}_n q_{na}^{-na}, \quad (5.3)$$

i część nieliniową zdefiniowaną przez model wielomianowy:

$$\hat{v}(n) = \hat{f}^T \mathbf{u}(n) = \hat{\mu}_1 u(n) + \hat{\mu}_2 u^2(n) + \dots + \hat{\mu}_r u^r(n). \quad (5.4)$$

A. Janczak: Identification of Nonlinear Systems, LNCIS 310, pp. 143-157, 2005.
 © Springer-Verlag Berlin Heidelberg 2005

Bez utraty ogólności można przyjąć, że μ^{\wedge}_1 jest jednością, a $\mu^{\wedge}_2, \dots, \mu^{\wedge}_r$ można odpowiednio znormalizować. Następnie (5.1) można wyrazić jako

$$y^{\wedge}(n) = \frac{B(q^{-1})}{B(q^{-1})} \sum_{k=2}^r \mu^{\wedge}_k u_k(n). \quad (5.5)$$

Wprowadzając $B^{-1}(q^{-1}) = B^{\wedge}(q^{-1}), \dots, B^{-k}(q^{-1}) = \mu^{\wedge}_k B^{\wedge}(q^{-1})$ i $u_1(n) =$

$u(n), \dots, u_k(n) = u^k(n)$, (5.5) można przepisać jako

$$y^{\wedge}(n) = \sum_{k=1}^r \frac{B_k(q^{-1})}{A^{\wedge}(q^{-1})} u_k(n). \quad (5.6)$$

Model MISO Wienera (rys. 5.1) składa się z r liniowych dynamicznych podmodeli ze wspólnym mianownikiem w każdej gałęzi. Równoważnie, (5.6) można zapisać jako

$$y^{\wedge}(n) = 1 - A^{\wedge}(q^{-1}) y^{\wedge}(n) + \sum_{k=1}^r B^{-k}(q^{-1}) u_k(n). \quad (5.7)$$

Korzystając z definicji $A^{\wedge}(q^{-1})$ i $B^{-k}(q^{-1})$, otrzymujemy

$$y^{\wedge}(n) = - \sum_{m=1}^{na} a^{\wedge}_m y^{\wedge}(n-m) + \sum_{k=1}^r \sum_{m=1}^{nb} b u_{kmk}(n-m), \quad (5.8)$$

gdzie b_{km} jest m -tym parametrem wielomianu $B_k(q^{-1})$. Model równoległy (5.8) można przekształcić w model szeregowo-równoległy liniowy w parametrach

przy użyciu dostępnych opóźnionych danych wyjściowych systemu $y(n-m)$ zamiast danych wyjściowych modelu $y^{\wedge}(n-m)$:

$$y^{\wedge}(n) = - \sum_{m=1}^{na} a^{\wedge}_m y(n-m) + \sum_{k=1}^r \sum_{m=1}^{nb} b u_{kmk}(n-m). \quad (5.9)$$

Wprowadzenie wektora parametrów θ^{\wedge} i wektora regresji $x(n)$:

$$\theta^{\wedge} = a^{\wedge}_1 \dots a^{\wedge}_{na} \ b_{r1} \dots b_{rnb}^T, \quad (5.10)$$

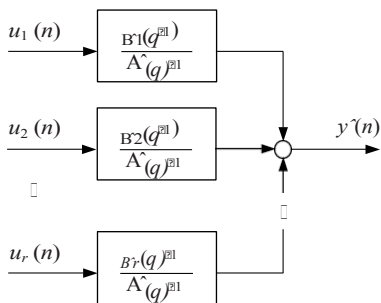
$$x(n) = -y(n-1) \dots -y(n-na) \ u_1(n-1) \dots u_1(n-nb) \dots u_r(n-1) \dots u_r(n-nb)^T, \quad (5.11)$$

mamy

$$y^{\wedge}(n) = x^T(n) \theta^{\wedge}. \quad (5.12)$$

Minimalizacja sumy N błędów kwadratowych między wyjściem systemu a wyjściem modelu względem wektora regresji θ^{\wedge} daje następujące wyniki

$$\theta^{\wedge} = (X^T X)^{-1} X^T y, \quad (5.13)$$



Rys. 5.1. Model Hammersteina przekształcony do postaci MISO

gdzie

$$X = x(1) \dots x(N)^T, \quad (5.14)$$

$$y = y(1) \dots y(N)^T. \quad (5.15)$$

W ten sposób można obliczyć parametry przekształconego modelu MISO (5.9), ale naszym głównym celem jest określenie parametrów modelu zdefiniowanego przez (5.1) i (5.4). Parametry \hat{a}_m są elementami wektora $\hat{\theta}$. Również parametry \hat{b}_m są dostępne bezpośrednio jako $\hat{b}_m = \hat{b}_{1m}$. Parametry

$\hat{\mu}_k$ można obliczyć na podstawie pozostałych elementów $\hat{\theta}$, ale istnieje pewien ilość nadmiarowości jako

$$\hat{\mu}_k = \hat{t}_{lk}^{bmk}, \quad k = 2, \dots, r, \quad m = 1, \dots, nb, \quad (5.16)$$

i można obliczyć nb różnych zestawów parametrów $\hat{\mu}$. Chang i Luus [27] zasugerowali obliczenie pierwiastka błędu średniokwadratowego (RMS) wyniku modelu dla wszystkich nb zestawów $\hat{\mu}$ i zaakceptowanie zestawu, który daje najmniejszą wartość RMS, jako bardziej wiarygodnego podejścia w porównaniu z obliczaniem średniej z nb wartości.

5.2 Iteracyjna identyfikacja systemów Hammersteina metodą najmniejszych kwadratów

W iteracyjnej metodzie najmniejszych kwadratów Narendry i Gallmana [120] parametry liniowego układu dynamicznego i nieliniowego elementu statycznego są aktualizowane oddzielnie i sekwencyjnie. Metoda ta wykorzystuje naprzemienną regulację parametrów liniowego układu dynamicznego i nieliniowego elementu statycznego w celu zminimalizowania sumy błędów podniesionych do kwadratu. Błąd $e(n)$ w czasie n wynosi

$$e(n) = y(n) - \sum_{m=1}^{na} \hat{a}_m y(n-m) + \sum_{m=1}^{nb} \hat{b}_m \hat{v}^{(n-m)}, \quad (5.17)$$

gdzie

$$\hat{v}^{(n)} = \sum_{k=1}^K \hat{\mu}^T u_k^{(n)}. \quad (5.18)$$

Wprowadzając $\hat{\theta} = \begin{bmatrix} \theta_a \\ \theta_b \end{bmatrix}$ i $x(n) = \begin{bmatrix} x_a^T(n) \\ x_b^T(n) \end{bmatrix}$, gdzie

$$\theta_a = [\hat{a}_1 \dots \hat{a}_{na}]^T, \quad (5.19)$$

$$\theta_b = [\hat{b}_1 \dots \hat{b}_{nb}]^T, \quad (5.20)$$

$$x_a(n) = [-y(n-1) \dots -y(n-na)]^T, \quad (5.21)$$

$$x_b(n) = [\hat{v}^{(n-1)} \dots \hat{v}^{(n-nb)}]^T, \quad (5.22)$$

(5.17) można zapisać jako

$$e(n) = y(n) - x^T(n) \hat{\theta} = y(n) - x_a^T(n) \theta_a - x_b^T(n) \theta_b. \quad (5.23)$$

Teraz, wprowadzając wektory $\hat{\mu} = [\hat{\mu}_1^T \dots \hat{\mu}^{T-1}]^T$ i $u(n-1) = [u(n-1) \ u^2(n-1) \dots \ u^r(n-1)]^T$, można zdefiniować macierz $U(n)$:

$$U(n) = \begin{bmatrix} u^T(n-1) \\ \vdots \\ u^T(n-nb) \end{bmatrix}. \quad (5.24)$$

Z (5.18) wynika, że wektor $x_b(n)$ można zapisać jako

$$x_b(n) = U(n) \hat{\mu}, \quad (5.25)$$

i (5.23) staje się

$$e(n) = y(n) - x_a^T(n) \theta_a - U(n) \hat{\mu}^T \theta_b. \quad (5.26)$$

W tym iteracyjnym podejściu szacowanie parametrów odbywa się zgodnie z następującą iteracyjną procedurą:

1. Na podstawie N pomiarów sygnałów wejściowych i wyjściowych systemu oraz założonego wektora początkowego $\hat{\mu}^{(1)}$, wektor parametrów $\hat{\theta}^{(1)} = [\hat{\theta}^{(1)T} \ \hat{\theta}^{(1)T}]^T$ szacowany jest model (5.23):

$$\hat{\theta}^{(1)} = (X^T X)^{-1} X^T y, \quad (5.27)$$

gdzie

$$y = [y(1) \dots y(N)]^T, \quad (5.28)$$

$$X = [x(1) \dots x(N)]^T. \quad (5.29)$$

2. $Z [\theta^{(1)T} \theta^{(1)T}]^T, \mu^{(2)}$ jest obliczane poprzez minimalizację sumy kwadratów błędów dla modelu (5.26)

$$\mu^{(2)} = U^T U U^{-1T} (y - X_a \theta^{(1)}), \quad (5.30)$$

gdzie

$$X_a = x(1) \dots x(N)^T, \quad (5.31)$$

$$U = U^T(1) \theta_b^{(1)} \dots U^T(N) \theta_b^{(1)T}. \quad (5.32)$$

3. Używając $\mu^{(2)}, \theta_a^{(2)} = [\theta_b^{(2)T} \theta^{(2)T}]^T$ oblicza się zgodnie ze schematem użyty w kroku 1 i proces jest kontynuowany.

Oczywistą zaletą metody iteracyjnej jest nieredundantna parametryzacja modelu. Badania symulacyjne przeprowadzone przez Gallmana [41] pokazują, że taka iteracyjna procedura daje znacznie niższą wariancję parametrów i nieco niższy błąd RMS niż nieiteracyjna metoda Changa i Luusa. Metoda iteracyjna zastosowana do różnych problemów, takich jak nieliniowość wielomianowa, półfalowy detektor liniowy lub nieliniowość nasycająca, okazała się bardzo skuteczna [120]. Pomimo doniesień o udanych doświadczeniach, metoda może nie być zbieżna w niektórych przypadkach.

5.3 Identyfikacja systemów Hammersteina w obecności skorelowanego szumu

Bezpośrednie zastosowanie metody Changa i Luusa [27] dla układów Hammersteina ze skorelowanymi addytywnymi zakłóceniami wyjściowymi prowadzi do asymptotycznie błędnych oszacowań parametrów. Aby przezwyciężyć ten problem, Haist *et al.* [63] zaproponowali iteracyjną procedurę najmniejszych kwadratów, która szacuje parametry liniowego układu dynamicznego, parametry nieliniowego elementu statycznego oraz parametry modelu szumu. Podejście to opiera się na założeniu, że system Hammersteina z zaburzeniem wyjściowym ma następujące parametry forma

$$y(n) = \frac{B(q)^{-1}}{A(q^{-1})} f u(n) + \frac{\varepsilon(n)}{D(q^{-1})}, \quad (5.33)$$

gdzie

$$D(q^{-1}) = 1 + d_1 q_1^{-1} + \dots + d_{nd} q_{nd}^{-nd}, \quad (5.34)$$

a $\varepsilon(n)$ jest dyskretnym szumem białym o zerowej średniej. Załóżmy, że element nieliniowy jest opisany wielomianem

$$f u(n) = \mu_1 u(n) + \mu_2 u^2(n) + \dots + \mu_r u^r(n). \quad (5.35)$$

Oznaczając niezakłócone wyjście systemu przez $y^-(n)$ i addytywne zakłócenie na wyjściu systemu przez $\eta(n)$, mamy

$$y(n) = \bar{y}(n) + \eta(n) \quad (5.36)$$

$$\bar{y}(n) = \sum_{k=1}^r B_k(q^{-1}) u_k(n), \quad (5.37)$$

$$\eta(n) = \frac{\varepsilon(n)}{D(q^{-1})}, \quad (5.38)$$

gdzie $B_k(q^{-1}) = \mu_k B(q^{-1})$ i $u_k(n) = u^k(n)$ lub, równoważnie,

$$\begin{aligned} y(n) = & - \sum_{m=1}^{na} a_m \bar{y}(n-m) + \sum_{k=1}^r \sum_{m=1}^{nb} b_{km} u_k(n-m) \\ & - \sum_{m=1}^{nd} d_m \eta(n-m) + \varepsilon(n). \end{aligned} \quad (5.39)$$

Wprowadzenie wektora parametrów θ i wektora regresji $x(n)$:

$$\theta = [a_1 \dots a_{na} \ b_{11} \dots b_{1nb} \dots b_{r1} \dots b_{rnb} \ d_1 \dots d_{nd}]^T, \quad (5.40)$$

$$x(n) = [-\bar{y}(n-1) \dots -\bar{y}(n-na) \ u_1(n-1) \dots u_1(n-nb) \dots \\ u_r(n-1) \dots u_r(n-nb) \ -\eta(n-1) \dots -\eta(n-nd)]^T, \quad (5.41)$$

mamy

$$y(n) = x^T(n)\theta + \varepsilon(n). \quad (5.42)$$

Teraz można dokonać aktualizacji parametrów zgodnie z następującą iteracyjną procedurą:

1. Ponieważ sekwencję $\bar{y}(n)$ i $\eta(n)$ nie są początkowo znane, oblicz parametry deterministycznej części modelu za pomocą nieiteracyjnej procedury Changa i Luusa, opisanej w sekcji 5.1. Wygenerować estymatę sekwencji $\bar{y}(n)$ z (5.37) i obliczyć estymatę sekwencji $\eta(n)$ z (5.36) używając estymaty sekwencji $\bar{y}(n)$.
2. Oblicz ulepszone oszacowania parametrów na podstawie

$$\hat{\theta} = \left(\sum_{i=1}^N x(i)x^T(i) \right)^{-1} \sum_{i=1}^N x(i)y(i), \quad (5.43)$$

gdzie oszacowane wartości $\bar{y}(n)$ i $\eta(n)$ są używane w $x(n)$ zamiast ich nieznanych prawdziwych wartości.

3. Poprawione oszacowania parametrów dają kolejne oszacowanie sekwencji $\bar{y}(n)$. Ponadto, oszacowana sekwencja $\eta(n)$ jest ponownie obliczana na podstawie (5.36).
4. Kontynuuj procedurę od kroku 2, aż zmiana znormalizowanej wartości RMS będzie mniejsza niż określone minimum.

Aby zapewnić stabilność numeryczną procedury, sugeruje się zastosowanie współczynnika krokowego ϵ . Dzięki współczynnikowi krokowemu przyjęta wartość wektora parametrów modelu $\theta^{*(j+1)}$ w iteracji $j+1$ jest obliczana zgodnie ze wzorem

$$\theta^{*(j+1)} = \epsilon \theta^{(j+1)} + (1 - \epsilon) \theta^{*(j)}, \quad (5.44)$$

gdzie $\theta^{(j+1)}$ jest oszacowaniem wektora parametrów otrzymanym z

(5.43).

Podobnie jak w przypadku metody Changa i Luusa, z (5.41) wynika, że istnieje pewna nadmiarowość w określaniu parametrów. Pomimo potwierdzonych w praktyce udanych zastosowań, nie jest dostępny dowód zbieżności tej metody.

5.4 Identyfikacja układów Hammersteina za pomocą rozwinięcia funkcji Laguerre'a

Przykładem metod identyfikacji wykorzystujących ortonormalne funkcje bazowe jest metoda zaproponowana przez Thathachara i Ramaswamy'ego [156]. W metodzie tej nieliniowa część modelu Hammersteina jest reprezentowana przez wielomian (5.4), a część liniowa - przez rozwinięcie Laguerre'a jego odpowiedzi impulsowej $h(n)$:

$$h(n) = \sum_{i=1}^n A_i l_i(n), \quad (5.45)$$

gdzie $l_i(n)$ jest i -tą funkcją Laguerre'a, a A_i , $i = 1, \dots, n$, są parametrami. Dyskretnie funkcje Laguerre'a są zdefiniowane jako

$$L_i(q^{-1}) = \frac{q^{-1} - \exp(-T)}{1 - \exp(-T)} \frac{1^{i-1}}{1 - \exp(-T)q^{-1}} \quad \text{dla } i = 1, 2, \dots, \quad (5.46)$$

gdzie T jest okresem próbkowania, a $L_i(q)$ jest dodatnią połowkową transformata transformata
 $l_i(n)$:

$$L_i(q^{-1}) = \sum_{n=0}^{\infty} l_i(n) q^{-n}. \quad (5.47)$$

Funkcje $l_i(n)$ spełniają warunek ortonormalności:

$$\sum_{n=0}^{\infty} l_i(n) l_k(n) = \begin{cases} 0 & \text{dla } i \neq k \\ 1 & \text{dla } i = k \end{cases} \quad k \leq n, \quad (5.48)$$

oraz

$$l_i(n) = 0 \quad \text{dla } n < 0 \text{ i wszystkich } i. \quad (5.49)$$

Problem identyfikacji polega na wyznaczeniu parametrów $\hat{\mu}_k$, $k = 1, \dots, r$, części nieliniowej, oraz A_i , $i = 1, \dots, n$, części liniowej. Problem

jest rozwiązywana w sposób iteracyjny, gdzie każda iteracja składa się z dwóch kroków. W pierwszym kroku parametry $\hat{\mu}_k$ są utrzymywane na stałym poziomie z poprzedniej iteracji, a A_i są obliczane poprzez rozwiązanie zestawu równań liniowych. Aby rozpocząć procedurę identyfikacji, przyjmuje się pewne arbitralne wartości początkowe $\hat{\mu}_k$, co pozwala na obliczenie wyjścia nieliniowej części modelu.

Wyrażenia dla parametrów A_i można wyprowadzić z funkcji korelacji krzyżowej wejścia-wyjścia części liniowej:

$$A_1 = \frac{C_{nl}}{B_{nl}} \quad \square$$

$$A_s = \frac{C_{nl-s+1} - \sum_{i=1}^{s-1} A_i B_{nl-s+i}}{B_{nl-s+1}} \quad \square$$

$$A_{nl} = \frac{C_1 - \sum_{i=1}^{nl-1} A_i B_{ii}}{B_{nl}} \quad \square$$
(5.50)

Współczynniki B_j , $j = 1, \dots, nl$, są obliczane jako następujące średnie:

$$B_j = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{n=0}^N u(n) f_j(n), \quad (5.51)$$

gdzie $f_j(n)$ jest wyjściem $L_j(q^{-1})$ wzbudzonym przez $\hat{f}^* u(n)$. W podobny sposób współczynniki C_j , $j = 1, \dots, nl$ są obliczane jako

$$C_j = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{n=0}^N u(n) y_{j-1}(n), \quad (5.52)$$

gdzie $y_{j-1}(n)$ jest wyjściem $L_{j-1}(q^{-1})$ wzbudzonym przez $y(n)$.

W drugim kroku parametry A_i są utrzymywane na stałym poziomie, podczas gdy $\hat{\mu}_k$ są określane.

jest regulowany za pomocą algorytmu, który dostosowuje $\hat{\mu}_k$ do ich optymalnych wartości, minimalizując średni błąd kwadratowy (MSE) między wyjściem modelu a wyjściem systemu. Minimalizacja jest przeprowadzana za pomocą metod gradientowych, tj. stromego zejścia w przypadku braku zakłóceń i aproksymacji stochastycznej w obecności addytywnych zakłóceń wyjściowych. MSE jest zdefiniowane jako

$$J = \frac{1}{N} \sum_{n=1}^N \sum_{\mu \in T}^{m \in \Sigma^1} h(m) u(n-m) - y(n)^2, \quad (5.53)$$

gdzie $u(n) = u(n) \ u^2(n) \dots \ u^r(n)^T$ i $\mu \in T = \mu^1 \ \mu^2 \dots \ \mu^r$. W

W przypadku braku zakłóceń parametry nieliniowego modelu elementu są dostosowywane zgodnie z następującą zasadą:

$$\hat{\mu}^{(new)} = \hat{\mu}^{(old)} - \hat{\mu} \eta \frac{\partial J}{\partial \mu}, \quad (5.54)$$

gdzie η jest wielkością kroku. W obecności addytywnych zakłóceń wyjściowych zamiast stałej wielkości kroku $\eta(n)$ stosuje się zależną od iteracji wielkość kroku, która spełnia następujące warunki:

$$\eta(n) \geq 0, \quad \sum_{k=1}^{\infty} \eta^2(n) < \infty, \quad \sum_{k=1}^{\infty} \eta(n) = \infty. \quad (5.55)$$

5.5 Metoda błędu przewidywania

W metodzie błędu predykcji przyjmuje się, że system Hammersteina jest określony przez

$$y(n) = \frac{B(q^{-1})}{A(q^{-1})} \sum_{k=1}^{\infty} u^k(n) + \varepsilon(n), \quad (5.56)$$

gdzie $\varepsilon(n)$ jest białym szumem o zerowej średniej używanym do modelowania zakłóceń. W przypadku szumu innego niż biały należy również określić parametry filtra szumów, ale ze względu na złożoność problemu nie jest on tutaj rozważany. Predyktor z jednokrokovym wyprzedzeniem dla wyjścia systemu ma postać [34]:

$$\hat{y}(n/n-1) = \frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} \sum_{k=1}^{\infty} \hat{\mu}^k u^k(n). \quad (5.57)$$

Wektor parametrów $\hat{\theta} = \hat{a}_1 \dots \hat{a}_{na} \hat{b}_1 \dots \hat{b}_{nb} \hat{\mu}_1^r \dots \hat{\mu}^T$ powinien zostać oszacowany.

na podstawie zestawu N pomiarów wejścia-wyjścia, aby zminimalizować sumę kwadratów błędów predykcji

$$J = \frac{1}{N} \sum_{i=1}^N e^2(n, \hat{\theta}), \quad (5.58)$$

gdzie

$$e(n, \hat{\theta}) = y(n) - \hat{y}(n/n-1). \quad (5.59)$$

Metoda błędu predykcji wymaga obliczenia gradientu (5.58)

w.r.t. jego parametrów. W przypadku modelu Hammersteina pochodne cząstkowe (5.58) są nieliniowe w parametrach. W związku z tym, ponieważ nie można znaleźć bezpośredniego rozwiązania problemu optymalizacji, należy zastosować metody iteracyjne. Metoda błędu predykcji jest techniką gradientową drugiego rzędu, w której

wektor parametrów $\hat{\theta}$ jest dostosowywany wzdłuż ujemnego gradientu (5.58):

$$\theta^{(j)} = \theta^{(j-1)} - \eta H_j^{-1} (\theta^{(j-1)})_G (\theta^{(j-1)}), \tag{5.60}$$

gdzie η_j to rozmiar kroku, $H(-)$ to Hessjan z (5.58) lub jego przybliżenie, a $G(-)$ to gradient z (5.58). Gradient funkcji (5.58) można obliczyć jako

$$G(\hat{\theta}^{(j)}) = \frac{dJ}{d\hat{\theta}} = 2 \sum_{i=1}^N e(n, \hat{\theta}) \frac{\partial e(n, \hat{\theta})}{\partial \hat{\theta}} = -2^N \sum_{i=1}^N e(n, \hat{\theta}) \psi(n), \quad (5.61)$$

gdzie

$$\psi(n) = \frac{\partial \hat{y}(n/n-1)}{\partial \hat{a}_1} \dots \frac{\partial \hat{y}(n/n-1)}{\partial \hat{a}_{na}} \frac{\partial \hat{y}(n/n-1)}{\partial \hat{b}_1} \dots \frac{\partial \hat{y}(n/n-1)}{\partial \hat{b}_{nb}} \quad (5.62)$$

$$\frac{\partial \hat{y}(n/n-1)}{\partial \hat{\mu}_1} \dots \frac{\partial \hat{y}(n/n-1)}{\partial \hat{\mu}_r} = - \frac{1}{A^{(q-1)}_m} \sum_{m=1}^r B^{(q-1)}_m \hat{\mu}^m u^m(n-k), \quad (5.63)$$

$$\frac{\partial \hat{y}(n/n-1)}{\partial \hat{b}_k} = A^{(q-1)}_m \sum_{m=1}^r \mu^m u^m(n-k), \quad (5.64)$$

$$\frac{\partial \hat{y}(n/n-1)}{\partial \hat{\mu}_m} = B^{(q-1)}_m \mu^m u^m(n-k). \quad (5.65)$$

Główną wadą zastosowania funkcji Hessjana w (5.60) jest to, że wymaga ona pochodnych drugiego rzędu. Aby uniknąć obliczania pochodnych drugiego rzędu, można użyć przybliżonego Hessjana. W metodzie Levenberga-Marquardta przybliżony Hessjan jest obliczany jako

$$H(\hat{\theta}^{(j)}) = \frac{1}{N} \sum_{i=1}^N \frac{\partial e(n, \hat{\theta})}{\partial \hat{\theta}} \frac{\partial e(n, \hat{\theta})}{\partial \hat{\theta}}^T + \mu I = \frac{1}{N} \sum_{i=1}^N \psi(n) \psi^T(n) + \mu I, \quad (5.66)$$

gdzie μ jest nieujemnym małym skalarą, a I jest macierzą tożsamości o odpowiednim wymiarze. Iteracyjny algorytm predykcji składa się z następujących kroków:

1. Rozpocznij iteracje z początkowym oszacowaniem parametrów $\hat{\theta}^{(0)}$ i ustaw $\hat{\mu}_1 = 1$.
2. Wybierz małą wartość μ (typowy wybór to 0,0001).
3. Oblicz $\hat{y}(n \text{ i } 1)$ i $J(\hat{\theta}^{(j-1)})$.
4. Oblicz gradient i Hessjana za pomocą (5.61) - (5.66).
5. Zaktualizuj oszacowania parametrów za pomocą (5.60) i oblicz $J(\hat{\theta}^{(j)})$.
6. Jeśli $J(\hat{\theta}^{(j)}) > J(\hat{\theta}^{(j-1)})$, zmniejsz μ o współczynnik (powiedzmy 10) i przejdź do kroku 4. $J(\hat{\theta}^{(j)}) < J(\hat{\theta}^{(j-1)})$, zaktualizuj rozwiązanie i zwiększ μ o współczynnik (powiedzmy 10) i przejdź do kroku 2.

Mając wyprowadzone zasady obliczania gradientu, można łatwo

zaimplementować rekurencyjną wersję algorytmu, patrz równania (2.97) - (2.99).

5.6 Identyfikacja systemów MISO za pomocą metody regresji pseudolinearnej

Rozważmy model MISO Hammersteina z *wieloma* wejściami i wyjściem zakłóconym addytywnie przez skorelowany szum pomiarowy [22]:

$$y(n) = \sum_{j=1}^m \frac{B_j(q^{-1})}{A_j(q^{-1})} f_j u_j(n) + \frac{C(q^{-1})}{D(q^{-1})} \varepsilon(n), \quad (5.67)$$

oraz wielomianowe modele nieliniowości statycznych

$$f_j u_j(n) = \mu_{j1} u_{j1}(n) + \dots + \mu_{jF_j} u_{jF_j}(n), \quad (5.68)$$

gdzie

$$A_j(q^{-1}) = 1 + a_{j1} q^{-1} + \dots + a_{jna} q^{-na_j}, \quad (5.69)$$

$$B_j(q^{-1}) = 1 + b_{j1} q^{-1} + \dots + b_{jnb} q^{-nb_j}, \quad (5.70)$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \dots + c_{nc} q^{-nc}, \quad (5.71)$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \dots + d_{nd} q^{-nd}. \quad (5.72)$$

Założmy, że $\varepsilon(n)$ jest białym szumem o zerowej średniej, model (5.67) jest asymptotycznie stabilny, a rzędy wielomianów na_j , nb_j , r_j , nc i nd są znane. Model (5.67) można przekształcić do następującej równoważnej postaci:

$$A^-(q^{-1}) y(n) = \sum_{j=1}^m \frac{B_j^-(q^{-1}) f_j u_j(n)}{D(q^{-1})} + C(q^{-1}) \varepsilon(n), \quad (5.73)$$

gdzie

$$A^-(q^{-1}) = \prod_{k=1}^m A_k(q^{-1}) = 1 + a_1^- q^{-1} + \dots + a_{na}^- q^{-na}, \quad (5.74)$$

$$B_j^-(q^{-1}) = B_j(q^{-1})^{-1} \prod_{k=1, k \neq j}^m A_k(q^{-1}) = 1 + b_{j1}^- q^{-1} + \dots + b_{jnb}^- q^{-nb_j}, \quad (5.75)$$

$$C^-(q^{-1}) = C(q^{-1}) A^-(q^{-1}) = 1 + c_1^- q^{-1} + \dots + c_{nac}^- q^{-nac}, \quad (5.76)$$

z $na = na_1 + \dots + na_{nu}$, $nab_j = na_1 + \dots + na_{j-1} + nb_j + na_{j+1} + \dots + na_{nu}$, $nac = na + nc$. Wykonanie kolejnej transformacji daje model w postaci pseudolinearnej w parametrach:

$$A(q^{-1}) y(n) = \sum_{j=1}^m B_j(q^{-1}) f_j u_j(n) + C^-(q^{-1}) \varepsilon(n), \quad (5.77)$$

gdzie

$$A(q^{-1}) = A^-(q^{-1})D(q^{-1}) = 1 + \alpha q_1^{-1} + \dots + \alpha q_{n\alpha}^{-n\alpha}, \quad (5.78)$$

$$B_j(q^{-1}) = B^-_j(q^{-1})D(q^{-1}) = \theta q_{j1}^{-1} + \dots + \theta_{jn\beta} q^{-n\beta} j, \quad (5.79)$$

z $n\alpha = n\alpha + nd$ i $n\beta_j = nab_j + nd$. Wprowadzenie wektora parametrów θ ,

$$\begin{aligned} \theta = & \alpha_1 \dots \alpha_{n\alpha} \beta_{11} \mu_{11} \dots \beta_{11} \mu_{1r_1} \dots \beta_{1n\beta_1} \mu_{11} \dots \beta_{1n\beta_1} \mu_{1r_1} \dots \\ & \beta_{nu1} \mu_{nu1} \dots \beta_{nu1} \mu_{nur_{nu}} \dots \beta_{nun\beta_{nu}} \mu_{nu1} \dots \beta_{nun\beta_{nu}} \mu_{nur_{nu}} \\ & \bar{c}_1 \dots \bar{c}_{n\alpha}^T, \end{aligned} \quad (5.80)$$

and the vector $x_0(n)$,

$$\begin{aligned} x_0(n) = & y(n-1) \dots y(n-n\alpha) u_1(n) \dots u^r_1(n) \dots u_1(n-\beta_1) \\ & \dots u^r_1(n-\beta_1) \dots u_{nu}(n) \dots \mu^{r_{nu}}_{nu}(n) \dots u_{nu}(n-\beta_1) \\ & \mu^{r_{nu}}_{nu}(n-\beta_1) \epsilon(n-1) \dots \epsilon(n-n\alpha)^T, \end{aligned} \quad (5.81)$$

mamy

$$y(n) = x_0^T(n)\theta + \epsilon(n). \quad (5.82)$$

Wprowadzenie błędu predykcji $e(n)$

$$e(n) = y(n) - x^T(n)\hat{\theta}^{(n-1)}, \quad (5.83)$$

gdzie $\hat{\theta}$ jest wektorem regulowanych parametrów

$$\begin{aligned} \hat{\theta} = & \hat{\alpha}_1 \dots \hat{\alpha}_{n\alpha} \hat{\beta}_{11} \hat{\mu}_{11} \dots \hat{\beta}_{11} \hat{\mu}_{1r_1} \dots \hat{\beta}_{1n\beta_1} \hat{\mu}_{11} \dots \hat{\beta}_{1n\beta_1} \hat{\mu}_{1r_1} \dots \\ & \hat{\beta}_{nu1} \hat{\mu}_{nu1} \dots \hat{\beta}_{nu1} \hat{\mu}_{nur_{nu}} \dots \hat{\beta}_{nun\beta_{nu}} \hat{\mu}_{nu1} \dots \hat{\beta}_{nun\beta_{nu}} \hat{\mu}_{nur_{nu}} \\ & \hat{c}_1 \dots \hat{c}_{n\alpha}^T, \end{aligned} \quad (5.84)$$

and $x(n)$ is defined as

$$\begin{aligned} x(n) = & y(n-1) \dots y(n-n\alpha) u_1(n) \dots u^r_1(n) \dots u_1(n-\beta_1) \dots \\ & u^r_1(n-\beta_1) \dots u_{nu}(n) \dots \mu^{r_{nu}}_{nu}(n) \dots u_{nu}(n-\beta_1) \\ & \dots \mu^{r_{nu}}_{nu}(n-\beta_1) e(n-1) \dots e(n-n\alpha)^T. \end{aligned} \quad (5.85)$$

Oszacowanie $\hat{\theta}(n)$ dla θ_0 , które minimalizuje sumę kwadratów błędów między systemem a wyjściami modelu, można uzyskać za pomocą rekurencyjnego algorytmu regresji pseudolosowej w następujący sposób:

$$\hat{\theta}(n) = \hat{\theta}^{(n-1)} + K(n)e(n), \quad (5.86)$$

$$K(n) = P(n)x(n) = \frac{P(n-1)x(n)}{1 + x^T(n)P(n-1)x(n)}, \quad (5.87)$$

$$P(n) = P(n-1) - K(n)x^T(n)P(n-1). \quad (5.88)$$

5.7 Identyfikacja systemów z dwusegmentowymi nieliniowościami

W niektórych przypadkach tylko wielomiany wyższego rzędu mogą odpowiednio aproksymować charakterystyki nieliniowe. Wzrost rzędu wielomianu r powoduje wielokrotny wzrost ogólnej liczby parametrów w modelu liniowym w parametrach (5.12). Co więcej, pojedynczy model wielomianowy może być niedokładny w całym zakresie sygnału wejściowego układu. Alternatywne podejście do rozwiązań z pojedynczym modelem wielomianowym, omówione wcześniej, zostało zaproponowane przez Vörös [161]. W metodzie tej stosuje się dwusegmentowy opis charakterystyki nieliniowej, złożony z oddzielnych map wielomianowych dla wejść dodatnich i ujemnych. Główną motywacją dla takiego podejścia jest lepsze dopasowanie bez zwiększania rzędów wielomianów dla niektórych typów nieliniowości.

Przyjmijmy, że nieliniowa charakterystyka $f(-)$ wynosi

$$\hat{f}(n) = f(u(n)) = \begin{cases} f_1(u(n)), & \text{jeśli } u(n) \geq 0 \\ f_2(u(n)), & \text{jeśli } u(n) < 0 \end{cases} \quad (5.89)$$

gdzie

$$f_1(u(n)) = \sum_{k=1}^r \hat{\mu}_{1k} u_{1k}^k(n), \quad (5.90)$$

$$f_2(u(n)) = \sum_{k=1}^r \hat{\mu}_{2k} u_{2k}^k(n). \quad (5.91)$$

Wprowadzenie sekwencji przełączania $\{g(n)\}$,

$$g(n) = g(u(n)) = \begin{cases} 0, & \text{jeśli } u(n) \geq 0 \\ 1, & \text{jeśli } u(n) < 0 \end{cases}, \quad (5.92)$$

(5.89) można zapisać jako

$$\hat{v}(n) = f_1(u(n)) + f_2(u(n)) - f_1(u(n))g(n). \quad (5.93)$$

Następnie biorąc pod uwagę (5.90) i (5.91),

$$\hat{v}(n) = \sum_{k=1}^r \hat{\mu}_{1k} u_{1k}^k(n) + \sum_{k=1}^r \hat{p}_k u_k^k(n)g(n), \quad (5.94)$$

gdzie

$$\hat{p}_k = \hat{\mu}_{2k} - \hat{\mu}_{1k}. \quad (5.95)$$

Podstawienie (5.94) do (5.1) daje model, który jest nieliniowy w parametrach. Zakładając jednak, że $\hat{b}_1 = 1$, model (5.1) można wyrazić jako

$$\hat{y}(n) = \hat{v}(n-1) + \hat{B}(q^{-1})^{-1} \hat{v}(n) + 1 - \hat{A}(q^{-1})^{-1} \hat{y}(n). \quad (5.96)$$

Następnie podstawiając (5.94) tylko dla $v^{\wedge}(n-1)$, otrzymujemy następującą pseudoliniową-in-
można uzyskać model parametrów typu równoległego:

$$y^{\wedge}(n) = \sum_{k=1}^r \mu^{\wedge} u_{1k}^k(n-1) + \sum_{k=1}^r p u_k^k(n-1)g(n-1) + B^{\wedge}(q^{-1}) - 1 \cdot v^{\wedge}(n) + 1 - A^{\wedge}(q^{-1}) y^{\wedge}(n). \quad (5.97)$$

Zastąpienie $y^{\wedge}(n)$ przez $y(n)$ przekształca model (5.97) do postaci szeregowo-równoległej:

$$y^{\wedge}(n) = \sum_{k=1}^r \mu^{\wedge} u_{1k}^k(n-1) + \sum_{k=1}^r p u_k^k(n-1)g(n-1) + B^{\wedge}(q^{-1}) - 1 \cdot v^{\wedge}(n) + 1 - A^{\wedge}(q^{-1}) y(n) \quad (5.98)$$

$$\text{lu} \quad b \quad y^{\wedge}(n) = x^T(n) \theta^{\wedge}, \quad (5.99)$$

gdzie wektor parametrów jest zdefiniowany jako

$$\theta^{\wedge} = \mu^{\wedge}_{11} \dots \mu^{\wedge}_{1r} p_1 \dots p_2 \hat{b} \dots \hat{b}_{nb} a^{\wedge}_{11} \dots a^{\wedge}_{na}^T, \quad (5.100)$$

and the regression vector is

$$x(n) = u(n-1) \dots u^r(n-1) u(n-1)g(n-1) \dots u^r(n-1)g(n-1) v^{\wedge}(n-2) \dots v^{\wedge}(n-nb) - y(n-1) \dots - y(n-na)^T. \quad (5.101)$$

Model (5.98) ma również postać pseudoliniową, ponieważ $v^{\wedge}(n)$ jest zmienną niemierzalną, która zależy od parametrów nieliniowej funkcji $f^{\wedge}(\cdot)$. Dlatego nie można zastosować algorytmu nieiteracyjnego do oszacowania θ^{\wedge} . Wówczas proponujemy algorytm iteracyjny, który wykorzystuje poprzednie oszacowania parametrów

μ^{\wedge}_{1k} , oraz p_k w celu oszacowania $v^{\wedge}(n)$. Oznaczmy przez $v^{\wedge(j)}(n)$, $\mu^{\wedge(j)}$ i $p^{(j)}$ oszacowania

$v^{\wedge}(n)$, μ^{\wedge}_{1k} i p_k uzyskanych w kroku j :

$$v^{\wedge(j)}(n) = \sum_{k=1}^r \mu^{\wedge}_{1k} u^{(j)k}(n) + \sum_{k=1}^r p u_k^{(j)k}(n)g(n). \quad (5.102)$$

Następnie błąd, który ma zostać zminimalizowany, można wyrazić jako

$$e(n) = y(n) - x^{(j)T}(n) \theta^{\wedge}(j+1), \quad (5.103)$$

gdzie $x^{(j)}(n)$ jest wektorem regresji z oszacowaniami $v^{\wedge}(n)$ obliczonymi zgodnie z (5.102), a $\theta^{\wedge}(j+1)$ jest $(j+1)$ trzecim oszacowaniem θ^{\wedge} .

Iteracyjną procedurę identyfikacji można podzielić na następujące etapy:

1. Używając wektora regresji $x^{(j)}(n)$, zminimalizuj odpowiednie kryterium oparte na (5.103), aby oszacować $\theta^{\wedge}(j+1)$.

2. Korzystając z (5.102), oblicz $v^{(i+1)}(n)$.
3. Powtarzaj kroki 1 i 2, aż oszacowania parametrów zbiegną się do stałych wartości.

Testowanie powyższej procedury identyfikacji z różnie ukształtowanymi dwusegmentowymi wielomianami i nieliniowościami wykładniczymi ujawniło jej dobre właściwości zbieżności, chociaż formalny dowód zbieżności nie jest dostępny. Chociaż procedura iteracyjna, która wykorzystuje cały zestaw danych wejściowych i wyjściowych, została pierwotnie zaproponowana przez Vořosa, wyprowadzenie sekwencyjnej wersji metody jest proste.

Podobne podejście można również zastosować do identyfikacji nieciągłych układów Hammersteina, tj. układów z elementem nieliniowym opisanym funkcją nieciągłą, lub układów Wienera z tego typu nieliniowością [160].

5.8 Podsumowanie

Przedstawiliśmy siedem różnych metod identyfikacji układów Hammersteina, które wykorzystują modele Hammersteina z wielomianowym modelem elementu nieliniowego. Iteracyjna metoda najmniejszych kwadratów Narendry i Galimana oraz nieiteracyjna metoda najmniejszych kwadratów Changa i Luusa są najstarsze i należą do najbardziej znanych metod. Obie mają jednak pewne wady. Podczas gdy zbieżność pierwszej z nich nie jest gwarantowana, wadami drugiej są redundancja parametrów i ogromna liczba parametrów w przypadku modeli wysokiego rzędu. Metoda iteracyjna Haist *et al.* [63] również wykorzystuje nieunikalną parametryzację, ale w przeciwieństwie do jednoetapowej metody Changa i Luusa, jej zaletą są spójne oszacowania parametrów w przypadku skorelowanego szumu wyjściowego. W metodzie Thathachara i Ramaswamy'ego rozszerzenie Laguerre'a jest używane do reprezentowania liniowej części modelu. Parametry modelu są dostosowywane iteracyjnie za pomocą metody opartej na gradiencie. Parametry modelu funkcji przenoszenia impulsu liniowego układu dynamicznego i modelu wielomianowego elementu nieliniowego są szacowane iteracyjnie metodą Levenberga-Marquardta w podejściu błędu predykcji omówionym przez Eskinata i in [34]. Alternatywą dla metody błędu predykcji jest metoda regresji pseudolinearnej zaproponowana przez Boutayeb i in. Wreszcie, iteracyjna metoda Vořosa pozwala na identyfikację układów Hammersteina przy użyciu dwusegmentowego modelu wielomianowego elementu nieliniowego. Metoda ta wykorzystuje podejście regresji liniowej, ale formalny dowód zbieżności nie jest dostępny.