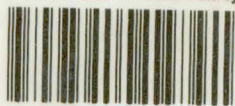


Biblioteka Główna i OINT
Politechniki Wrocławskiej



100100215747

Komitet Automatyki i Robotyki
Polskiej Akademii Nauk

Jerzy Świątek

Wybrane zagadnienia identyfikacji statycznych systemów złożonych



Wydawnictwo Naukowe
Politechniki Wrocławskiej

**Wybrane
zagadnienia
identyfikacji
statycznych
systemów złożonych**

Monografie
Tom 13

Komitet Redakcyjny serii

Tadeusz Kaczorek (przewodniczący)

Stanisław Bańka

Władysław Findeisen

Henryk Górecki

Edward Jezierski

Jerzy Klamka

Józef Korbicz

Roman Kulikowski

Antoni Niederliński

Maciej Niedźwiecki

Ryszard Tadeusiewicz

Krzysztof Tchoń

Leszek Trybus

Jan Węglarz

Jerzy Świątek

**Wybrane
zagadnienia
identyfikacji
statycznych
systemów złożonych**



Oficyna Wydawnicza Politechniki Wrocławskiej
Wrocław 2009

Recenzenci

Antoni Niederliński

Maciej Niedźwiecki

Opracowanie redakcyjne i korekta

Alicja Kordas

Projekt okładki i stron tytułowych

Wojciech J. Steifer

Wszelkie prawa zastrzeżone. Opracowanie w całości ani we fragmentach nie może być powielane ani rozpowszechniane za pomocą urządzeń elektronicznych, mechanicznych, kopiujących, nagrywających i innych bez pisemnej zgody posiadacza praw autorskich.

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2009

ISBN 978-83-7493-442-8

Oficyna Wydawnicza Politechniki Wrocławskiej

Wybrzeże Wyspiańskiego 27, 53-370 Wrocław

<http://www.oficyna.pwr.wroc.pl>

oficwyd@pwr.wroc.pl

Drukarnia Oficyny Wydawniczej Politechniki Wrocławskiej. Zam. nr /2009

*Najdroższym mi kobietom:
Joannie – mojej Mamie
Ewie – mojej Córcie*

Spis treści

Od autora	9
Wykaz ważniejszych oznaczeń	11
1. Identyfikacja systemów złożonych – pojęcia podstawowe	23
1.1. Wstęp.....	23
1.2. Rola modelu w badaniach systemowych.....	24
1.3. Zadanie identyfikacji obiektu	30
1.4. Opis systemu złożonego	38
1.5. Problemy identyfikacji systemów złożonych	47
2. Podstawowe zadania identyfikacji obiektów statycznych.....	50
2.1. Wyznaczanie parametrów charakterystyki statycznej obiektu.....	51
2.2. Estymacja parametrów charakterystyki statycznej obiektu.....	55
2.2.1. Estymacja parametrów charakterystyki statycznej obiektu na podstawie zakłóconych pomiarów wyjścia	56
2.2.2. Estymacja parametrów charakterystyki statycznej obiektu ze zmiennymi niemierzalnymi wielkościami losowymi.....	70
2.2.3. Estymacja parametrów charakterystyki statycznej obiektu ze zmiennymi niemierzalnymi wielkościami losowymi na podstawie zakłóconych pomiarów wyjścia	75
2.3. Zadanie wyboru optymalnego modelu w warunkach deterministycznych.....	81
2.3.1. Aproksymacja charakterystyki statycznej.....	81
2.3.2. Zadanie wyboru optymalnego modelu na podstawie ciągu niezakłóconych obserwacji	89
2.3.3. Zadanie wyboru optymalnego modelu z wykorzystaniem sieci neuronowej.....	94
2.4. Zadanie wyboru optymalnego modelu w warunkach losowych.....	99
2.4.1. Pełna informacja probabilistyczna	102
2.4.2. Niepełna informacja probabilistyczna.....	107
3. Identyfikacja obiektów złożonych przy ograniczonych możliwościach pomiarowych.....	115
3.1. Separowalność deterministyczna.....	118

3.2. Separowalność probabilistyczna	127
3.2.1. Estymacja parametrów charakterystyk statycznych systemu złożonego na podstawie zakłóconych pomiarów wyróżnionych wyjść	127
3.2.2. Estymacja parametrów charakterystyk statycznych systemu złożonego ze zmiennymi niemierzalnymi wielkościami losowymi	145
3.2.3. Estymacja parametrów charakterystyk statycznych systemu złożonego ze zmiennymi niemierzalnymi wielkościami losowymi na podstawie zakłóconych pomiarów wyróżnionych wyjść systemu złożonego.....	150
4. Wybór optymalnego modelu systemu złożonego.....	155
4.1. Lokalnie optymalny model systemu złożonego	157
4.2. Globalnie optymalny model systemu złożonego.....	159
4.3. Model globalnie optymalny z uwzględnieniem jakości modelu lokalnego	161
4.4. Algorytm identyfikacji dla systemu złożonego o strukturze szeregowej	163
4.4.1. Globalnie optymalny model systemu o strukturze szeregowej	153
4.4.2. Globalnie optymalny model z uwzględnieniem jakości modelu lokalnego dla systemu o strukturze szeregowej	180
4.5. Modelowanie systemów złożonych za pomocą sieci neuronowych.....	191
4.5.1. Algorytm uczenia sieci neuronowej dla systemu złożonego z lokalnym wskaźnikiem jakości identyfikacji	194
4.5.2. Algorytm uczenia sieci neuronowej dla systemu złożonego z globalnym wskaźnikiem jakości identyfikacji	196
5. Identyfikacja dwustopniowa.....	200
5.1. Dwustopniowe algorytmy estymacji parametrów obiektu	204
5.1.1. Metoda maksymalnej wiarygodności	209
5.1.2. Metoda maksymalnego prawdopodobieństwa	219
5.2. Dwustopniowe zadanie wyboru optymalnego modelu w warunkach losowych	226
5.2.1. Przypadek liniowo-kwadratowy.....	230
5.2.2. Porównanie podejścia dwustopniowego i bezpośredniego	238
5.3. Identyfikacja destylacyjnej kolumny wypełnionej z pulsacją fazy parowej	241
5.3.1. Opis danych pomiarowych.....	243
5.3.2. Algorytmy identyfikacji i wyniki obliczeń	247
5.3.3. Rozwinięcie badań z wykorzystaniem identyfikacji wielostopniowej...	254
Dodatek. Przekształcenia zmiennych losowych.....	257
Bibliografia.....	263

Od autora

Zaprezentowany w książce materiał jest wynikiem kilkuletnich prac autora w zakresie identyfikacji w Zakładzie Identyfikacji i Modelowania Instytutu Informatyki Technicznej (obecnie Instytutu Informatyki) Politechniki Wrocławskiej. Prace te są kontynuacją badań w zakresie identyfikacji systemów złożonych, prowadzonych w zespole pod kierunkiem Profesora Zdzisława Bubnickiego.

Tematem książki są wybrane problemy identyfikacji obiektów statycznych. Na wstępie przedstawiono rolę modelu w badaniach systemowych, wprowadzono pojęcia dotyczące zadania identyfikacji oraz wskazano aktualne problemy identyfikacji systemów złożonych. Następnie sformułowano podstawowe zadania identyfikacji obiektów statycznych. Omówiono zadania wyznaczania parametrów charakterystyk statycznych oraz zadania wyboru optymalnego modelu w warunkach deterministycznych i w warunkach losowych, czyli takich, gdy na obiekt działają pewne wielkości losowe, a wyniki eksperymentu są zniekształcone zakłóceniami pomiarowymi.

Zasadnicza część książki dotyczy identyfikacji statycznych systemów złożonych, czyli identyfikacji takich obiektów, w których wyróżniono elementy składowe i wskazano powiązania między nimi. Dokładniej omówione wybrane zagadnienia to:

➤ Identyfikacja przy ograniczonych możliwościach pomiarowych.

Jest to zadanie identyfikacji systemu złożonego, w którym – dla zadanych wartości wejść zewnętrznych – tylko wyróżnione wartości wyjść są dostępne do pomiaru. Rozważono problem możliwości wyznaczenia parametrów charakterystyk statycznych poszczególnych elementów systemu złożonego na podstawie dostępnych, ograniczonych pomiarów.

➤ Identyfikacja globalna.

W odróżnieniu od zadania identyfikacji lokalnej, które polega na wyznaczeniu optymalnego modelu każdego z elementów systemu złożonego z pominięciem struktury, zadanie identyfikacji globalnej polega na wyznaczeniu optymalnego modelu systemu złożonego z uwzględnieniem jego struktury.

➤ Identyfikacja dwustopniowa.

Koncepcja identyfikacji dwustopniowej wynika z metodologii prowadzenia badań, organizacji eksperymentu lub dekompozycji zadania identyfikacji. Przedstawiono dwustopniowe zadanie estymacji parametrów charakterystyki statycznej oraz dwu-

stopniowe zadanie wyboru optymalnego modelu. Zbadano warunki równoważności podejścia dwustopniowego i bezpośredniego.

Książka jest przeznaczona dla osób zainteresowanych tworzeniem modeli matematycznych obiektów złożonych. W szczególności adresowana jest do studentów i pracowników naukowych kierunków: automatyka i robotyka, ekonometria, informatyka, matematyka.

Autor pragnie podziękować recenzentom – Panu profesorowi Antoniemu Niederlińskiemu oraz Panu profesorowi Maciejowi Niedźwieckiemu – za cenne uwagi i twórczą krytykę, które ukształtowały ostateczną zawartość książki.

Dziękuję Panu dr. inż. Wojciechowi Pieniżkowi za udostępnienie wyników eksperymentu pozwalających na ilustrację zadania identyfikacji wielostopniowej oraz Kolegom z zespołu za uwagi, które wykorzystałem w redakcji książki. Oddzielne podziękowania przekazuję współpracownikom – Panu mgr. inż. Jarosławowi Drapale oraz Panu mgr. inż. Krzysztofowi Brzostowskiemu – za wsparcie podczas przygotowywania rękopisu. Pragnę wyrazić szczególną wdzięczność za owocną współpracę Pani mgr Alicji Kordas, która podjęła się trudu opracowania redakcyjnego i korekty książki.

Jerzy Świątek

Wykaz ważniejszych oznaczeń

- A – macierz powiązań pomiędzy wejściami i wyjściami elementów systemu złożonego
- B – macierz wskazująca wejścia zewnętrzne systemu złożonego
- C – macierz wskazująca wyróżnione wyjścia systemu złożonego
- D_u – podzbiór przestrzeni wejść, $D_u \subseteq U \subseteq \mathbb{R}^S$
- $F(u)$ – charakterystyka statyczna obiektu z wektorem wejść u
- $F(u, \theta)$ – charakterystyka statyczna obiektu z wektorem wejść u oraz wektorem parametrów θ
- $F(u, \theta, \omega)$ – charakterystyka statyczna obiektu z wektorem wejść u , wektorem parametrów θ oraz wektorem wielkości losowych ω
- $\tilde{F}(x)$ – charakterystyka statyczna systemu złożonego z wektorem wejść zewnętrznych x
- $\tilde{F}(x, \tilde{\theta})$ – charakterystyka statyczna systemu złożonego z wektorem wejść zewnętrznych x oraz wektorem parametrów $\tilde{\theta}$
- $F_1(u_1, \theta_1)$ – charakterystyka statyczna obiektu na pierwszym stopniu z wektorem wejść u_1 oraz wektorem parametrów θ_1
- $F_2(u_2, \theta_2)$ – charakterystyka statyczna obiektu na drugim stopniu z wektorem wejść u_2 oraz wektorem parametrów θ_2
- $F_m(u_m)$ – charakterystyka statyczna m -tego elementu systemu złożonego z wektorem wejść u_m
- $F_m(u_m, \theta_m)$ – charakterystyka statyczna m -tego elementu systemu złożonego z wektorem wejść u_m oraz wektorem parametrów θ_m
- $f(u, y)$ – funkcja gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ określona na $U \times Y$

$f_u(u)$	– funkcja gęstości rozkładu prawdopodobieństwa wejścia \underline{u} określona na \mathbf{U}
$f_w(w, \theta; u)$	– funkcja gęstości rozkładu prawdopodobieństwa zakłóconego pomiaru wyjścia \underline{w} dla obiektu z wektorem parametrów θ i ustalonym wejściu u
$f_y(y u)$	– funkcja gęstości rozkładu prawdopodobieństwa wyjścia \underline{y} pod warunkiem, że na wejściu zmienna losowa \underline{u} przyjęła wartość u
$f_z(z)$	– funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{z} określona na \mathbf{Z}
$f_\theta(\theta)$	– funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{\theta}$ określona na Θ , funkcja gęstości rozkładu prawdopodobieństwa a priori
$f'(\theta W_N; U_N)$	– funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{\theta}$ określona na Θ pod warunkiem, że na wyjściu uzyskano macierz wyników pomiarów W_N przy serii identyfikującej U_N , funkcja gęstości rozkładu prawdopodobieństwa a posteriori
$f_\omega(\omega)$	– funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej $\underline{\omega}$ określona na Ω
$g_u(u)$	– funkcja wagi jakości przybliżenia określona na zbiorze \mathbf{D}_u
J_F	– macierz Jakobiego (jakobian przekształcenia odwrotnego F_ω^{-1})
J_h	– macierz Jakobiego (jakobian przekształcenia odwrotnego h_z^{-1})
L	– wymiar przestrzeni wyjść obiektu i modelu
l	– indeks l -tego wyjścia obiektu i modelu
L_m	– wymiar przestrzeni wyjść opisu m -tego elementu systemu złożonego
l_m	– indeks l_m -tego wyjścia obiektu i modelu m -tego elementu systemu złożonego
$L(\theta, \bar{\theta})$	– funkcja strat
$L_N(W_N, \theta; U_N)$	– funkcja wiarygodności dla obiektu z wektorem parametrów θ , zadanej serii identyfikującej U_N , oraz macierzą wyników pomiarów W_N
$L_{1N_1}(W_{N_1n_2}, \theta_{1n_2}; U_{1N_1n_2})$	– funkcja wiarygodności na pierwszym stopniu dla obiektu z wektorem parametrów θ_{1n_2} , zadanej serii identyfikującej $U_{1N_1n_2}$, ustalonej wartości wejścia na drugim stopniu $u_2 = u_{2n_2}$ oraz macierzą wyników pomiarów $W_{N_1n_2}$

$L_{2N_2}(\hat{\Xi}_{1N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2})$ – funkcja wiarygodności na drugim stopniu dla obiektu z wektorem parametrów θ_2 , zadanych serii identyfikujących $\bar{U}_{1N_1N_2}$, U_{2N_2} oraz wynikami estymacji na pierwszym stopniu $\hat{\Xi}_{1N_1N_2}$

$L_{N_2N_1}(\bar{W}_{N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2})$ – funkcja wiarygodności przy podejściu bezpośrednim dla obiektu z wektorem parametrów θ_2 , zadanych seriach identyfikujących $\bar{U}_{1N_1N_2}$, U_{2N_2} oraz macierzą wyników pomiarów $\bar{W}_{N_1N_2}$

M – liczba elementów systemu złożonego

m – indeks m -tego elementu systemu złożonego

N – długość serii pomiarowej

n – indeks n -tego pomiaru

N_1 – długość serii pomiarowej na pierwszym stopniu

n_1 – indeks n_1 -tego pomiaru na pierwszym stopniu

N_2 – długość serii pomiarowej na drugim stopniu

n_2 – indeks n_2 -tego pomiaru na drugim stopniu

O_m – m -ty element obiektu złożonego

$q(y, \bar{y}) \stackrel{\text{df}}{=} q(F(u), \Phi(u, \theta))$ – ocena różnicy pomiędzy funkcją aproksymowaną a funkcją aproksymującą dla zadanego punktu $u \in D_u$

$q(y_n, \bar{y}_n) \stackrel{\text{df}}{=} q(y_n, \Phi(u_n, \theta))$ – ocena różnicy pomiędzy zmierzoną wartością wyjścia obiektu a wartością wyznaczoną z modelu dla zadanego wejścia u_n

$Q(\theta) \stackrel{\text{df}}{=} \|F(u) - \Phi(u, \theta)\|_{D_u}$ – ocena różnicy pomiędzy funkcją aproksymowaną a funkcją aproksymującą zbiorze D_u

$Q(\Phi) \stackrel{\text{df}}{=} E_{u, y} [q(y, \Phi(u))]$ – ocena różnicy pomiędzy wyjściem obiektu i wyjściem nieparametrycznego modelu w warunkach losowych przy pełnej informacji probabilistycznej

$Q_{1N_1n_2}(\theta_1)$ – wskaźnik jakości identyfikacji na pierwszym stopniu dla zadanej serii identyfikującej $U_{1N_1n_2}$ oraz ustalonego wejścia na drugim stopniu $u_2 = u_{2n_2}$ z wektorem parametrów θ_1

$Q_{2N_2}(\theta_2)$ – wskaźnik jakości identyfikacji na drugim stopniu dla obiektu z wektorem parametrów θ_2 i zadanej serii identyfikującej U_{2N_2}

$Q_{N_1 N_2}(\theta_2)$ – wskaźnik jakości identyfikacji przy podejściu bezpośrednim dla obiektu z wektorem parametrów θ_2 i zadanych seriach identyfikujących $\bar{U}_{1N_1 N_2}$ oraz U_{2N_2}

$Q(\theta_2) \stackrel{\text{df}}{=} E_{\underline{u}_1, \underline{u}_2, \underline{y}} [q_1(\underline{y}, \bar{\underline{y}} = \Phi(\underline{u}_1, \underline{u}_2, \theta_2))] -$ ocena różnicy pomiędzy wyjściem obiektu i wyjściem modelu przy podejściu bezpośrednim oraz pełnej informacji probabilistycznej

$Q_1(\theta_1, \underline{u}_2) \stackrel{\text{df}}{=} E_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \bar{\underline{y}} = \Phi_1(\underline{u}_1, \theta_1)) | \underline{u}_2] -$ ocena różnicy pomiędzy wyjściem obiektu i wyjściem modelu na pierwszym stopniu przy ustalonym wejściu na drugim stopniu \underline{u}_2 oraz pełnej informacji probabilistycznej

$Q_2(\theta_2) \stackrel{\text{df}}{=} E_{\underline{u}_2} [q_2(\theta_1^*, \bar{\underline{\theta}}_1 = \Phi_2(\underline{u}_2, \theta_2))] -$ ocena różnicy pomiędzy wyjściem obiektu i wyjściem modelu na drugim stopniu przy pełnej informacji probabilistycznej

$Q_{mN_m}(\theta_m) \stackrel{\text{df}}{=} \|Y_{mN_m} - \bar{Y}_{mN_m}(\theta_m)\|_{U_{mN_m}} -$ lokalna ocena różnicy pomiędzy macierzą zmierzonych wartości składowych wektora wyjść m -tego elementu systemu złożonego Y_{mN_m} a macierzą wartości składowych wektora wyjść wyznaczonych z modelu m -tego elementu systemu złożonego $\bar{Y}_{mN_m}(\theta)$ dla zadanej serii identyfikującej U_{mN_m} (lokalny wskaźnik jakości identyfikacji)

$Q_N(\theta) \stackrel{\text{df}}{=} \|Y_N - \bar{Y}_N(\theta)\|_{U_N} -$ ocena różnicy pomiędzy macierzą zmierzonych wartości składowych wektora wyjść obiektu Y_N a macierzą wartości składowych wektora wyjść wyznaczonych z modelu $\bar{Y}_N(\theta)$ dla zadanej serii identyfikującej U_N

$\tilde{Q}_N(\theta) \stackrel{\text{df}}{=} \|V_N - \bar{V}_N(\theta)\|_{X_N} -$ globalna ocena różnicy pomiędzy macierzą zmierzonych wartości składowych wektora wyróżnionych wyjść systemu złożonego V_N a macierzą wartości składowych wektora wyjść wyznaczonych z modelu systemu złożonego $\bar{V}_N(\theta)$ dla zadanej serii identyfikującej wejść zewnętrznych X_N (globalny wskaźnik jakości identyfikacji)

$\bar{Q}_N(\theta) \stackrel{\text{df}}{=} H(\tilde{Q}_N(\theta), Q_{1N}(\theta_1), Q_{2N}(\theta_2), \dots, Q_{MN}(\theta_M)) -$ syntetyczny wskaźnik jakości identyfikacji uwzględniający zarówno ocenę globalną, jak i oceny lokalne

$Q_u(\Phi(u)) \stackrel{\text{df}}{=} E \left[q(y, \Phi(u)) \Big|_{\underline{y}} \Big| u = u \right]$ – ocena różnicy pomiędzy wyjściem obiektu i wyjściem nieparametrycznego modelu w warunkach losowych dla ustalonego wejścia u , przy pełnej informacji probabilistycznej

$Q_{uM_n}(\Phi(u_n)) \stackrel{\text{df}}{=} \sum_{m=1}^{M_n} q(y_{nm}, \Phi(u_n))$ – empiryczne oszacowanie oceny różnicy pomiędzy wyjściem obiektu i wyjściem nieparametrycznego modelu na podstawie obserwacji wyjścia o długości M_n przy ustalonym wejściu u_n

$R(\bar{\Psi})$ – ryzyko podjęcia decyzji z wykorzystaniem algorytmu $\bar{\Psi}$ przetworzenia danych pomiarowych W_N, U_N

$r(\bar{\theta}, W_N; U_N)$ – warunkowe ryzyko podjęcia decyzji, że składowe wektora parametrów przyjmą wartość $\bar{\theta}$, pod warunkiem, iż w wyniku pomiarów wyjścia otrzymano macierz wyników pomiarów W_N , a na wejście podano serię identyfikującą U_N

$r_1(\bar{\theta}_{1n_2}, W_{N_1n_2}; U_{N_1n_2})$ – warunkowe ryzyko podjęcia decyzji, że składowe wektora parametrów przyjmą wartość $\bar{\theta}_{1n_2}$ na pierwszym stopniu, pod warunkiem, iż w wyniku pomiarów wyjścia otrzymano macierz wyników pomiarów $W_{N_1n_2}$, a na wejście podano serię identyfikującą $U_{1N_1n_2}$

$r_2(\bar{\theta}_2, \mathcal{E}_{1N_1N_2}; U_{2N_2})$ – warunkowe ryzyko podjęcia decyzji, że składowe wektora parametrów przyjmą wartość $\bar{\theta}_2$ na drugim stopniu, pod warunkiem, iż w wyniku estymacji na pierwszym stopniu otrzymano macierz oszacowań $\mathcal{E}_{N_1N_2}$, a na wejście podano serię identyfikującą U_{2N_2}

$r(\bar{\theta}_2, \bar{W}_{N_1N_2}; \bar{U}_{1N_1N_2}, U_{1N_2})$ – warunkowe ryzyko podjęcia decyzji, że składowe wektora parametrów przyjmą wartość $\bar{\theta}_2$ przy podejściu bezpośrednim, pod warunkiem, iż w wyniku pomiarów wartości składowych wektora wyjść otrzymano macierz wyników pomiarów $\bar{W}_{N_1N_2}$, a na wejścia podano serie identyfikujące $\bar{U}_{1N_1N_2}$ oraz U_{2N_2}

R – wymiar przestrzeni parametrów opisu obiektu

r – indeks r -tego parametru opisu obiektu

R – zbiór liczb rzeczywistych

R_m – wymiar przestrzeni parametrów opisu m -tego elementu systemu złożonego

r_m	– indeks r_m -tego parametru opisu m -tego elementu systemu złożonego
S	– wymiar przestrzeni wejść obiektu
s	– indeks s -tego wejścia obiektu
S_m	– wymiar przestrzeni wejść m -tego elementu systemu złożonego
s_m	– indeks s_m -tego wejścia m -tego elementu systemu złożonego
T	– transpozycja wektora

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix} \quad \text{– } S\text{-wymiarowy wektor wejść obiektu, } u \in U \subseteq \mathbb{R}^S$$

$U \subseteq \mathbb{R}^S$ – przestrzeń wejść obiektu, podzbiór S -wymiarowej przestrzeni liczb rzeczywistych

u_1 – S_1 -wymiarowy wektor wejść na pierwszym stopniu $u_1 \in U_1 \subseteq \mathbb{R}^{S_1}$

$U_1 \subseteq \mathbb{R}^{S_1}$ – przestrzeń wejść obiektu na pierwszym stopniu, podzbiór S_1 -wymiarowej przestrzeni liczb rzeczywistych

$U_{1N_1n_2} = \begin{bmatrix} u_{11n_2} & u_{12n_2} & \cdots & u_{1N_1n_2} \end{bmatrix}$ – macierz wyników pomiarów wartości składowych wektora wejść na pierwszym stopniu w n_2 -tym eksperymencie na drugim stopniu, tj. dla ustalonego wektora wejść na drugim stopniu $u_2 = u_{2n_2}$, seria identyfikująca na pierwszym stopniu

$\bar{U}_{1N_1N_2} = \begin{bmatrix} U_{1N_11} & U_{1N_12} & \cdots & U_{1N_1N_2} \end{bmatrix}$ – macierz wyników pomiarów wartości wszystkich składowych wektorów wejść na pierwszym stopniu dla zadanej serii pomiarowej na drugim stopniu

u_2 – S_2 -wymiarowy wektor wejść na drugim stopniu $u_2 \in U_2 \subseteq \mathbb{R}^{S_2}$

$U_2 \subseteq \mathbb{R}^{S_2}$ – przestrzeń wejść obiektu na drugim stopniu, podzbiór S_2 -wymiarowej przestrzeni liczb rzeczywistych

$U_{2N_2} = \begin{bmatrix} u_{21} & u_{22} & \cdots & u_{2N_2} \end{bmatrix}$ – macierz wyników pomiarów wartości składowych wektora wejść na drugim stopniu, seria identyfikująca na drugim stopniu

$$u_m = \begin{bmatrix} u_m^{(1)} \\ u_m^{(2)} \\ \vdots \\ u_m^{(S_m)} \end{bmatrix} \quad \text{– } S_m\text{-wymiarowy wektor wejść } m\text{-tego elementu systemu złożonego}$$

$$u_m \in U_m \subseteq \mathbb{R}^{S_m}$$

$U_m \subseteq \mathbb{R}^{S_m}$ – przestrzeń wejść m -tego elementu systemu złożonego, podzbiór S_m -wymiarowej przestrzeni liczb rzeczywistych

$U_N = [u_1 \ u_2 \ \dots \ u_N]$ – macierz wyników pomiarów wartości składowych wektora wejść obiektu, seria identyfikująca

$v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(L)} \end{bmatrix}, \bar{v} = \begin{bmatrix} \bar{v}^{(1)} \\ \bar{v}^{(2)} \\ \vdots \\ \bar{v}^{(L)} \end{bmatrix}$ – \tilde{L} -wymiarowy wektor wyróżnionych wyjść systemu złożonego i jego modelu, odpowiednio: $v, \bar{v} \in V \subseteq \mathbb{R}^{\tilde{L}}$

V – przestrzeń wyróżnionych wyjść systemu złożonego oraz wyjść jego modelu – podzbiór \tilde{L} -wymiarowej przestrzeni liczb rzeczywistych

Var_N – empiryczna wariancja wyznaczona na podstawie N obserwacji

$V_N = [v_1 \ v_2 \ \dots \ v_N]$ – macierz wyników pomiarów wartości składowych wektora wyróżnionych wyjść systemu złożonego

$w = h(y, z)$ – zakłócony wynik pomiaru – opis systemu pomiarowego

$W_N = [w_1 \ w_2 \ \dots \ w_N]$ – macierz zakłóconych wyników pomiarów wartości składowych wektora wyjść obiektu

$\tilde{W}_N = [\tilde{w}_1 \ \tilde{w}_2 \ \dots \ \tilde{w}_N]$ – macierz zakłóconych wyników pomiarów wartości składowych wektora wyróżnionych wyjść systemu złożonego

$W_{N_1 n_2} = [w_{1n_2} \ w_{2n_2} \ \dots \ w_{N_1 n_2}]$ – macierz zakłóconych wyników pomiarów wartości składowych wektora wyjść obiektu w n_2 -tym pomiarze na drugim stopniu, tj. dla ustalonego wejścia na drugim stopniu $u_2 = u_{2n_2}$, wyniki eksperymentu na pierwszym stopniu

$\bar{W}_{N_1 N_2} = [W_{N_1 1} \ W_{N_1 2} \ \dots \ W_{N_1 N_2}]$ – macierz zakłóconych wyników pomiarów wszystkich wartości składowych wektora wyjść na pierwszym stopniu dla zadanej serii pomiarowej na drugim stopniu

$w_{n_1 n_2} = h(y_{n_1 n_2}, z_{n_1 n_2})$ – wynik pomiaru wartości składowych wektora $y_{n_1 n_2}$ z zakłóceniem $z_{n_1 n_2}$

$x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(\tilde{S})} \end{bmatrix}$ – \tilde{S} -wymiarowy wektor wejść zewnętrznych systemu złożonego, $x \in X \subseteq \mathbb{R}^{\tilde{S}}$

$X \subseteq \mathbb{R}^{\tilde{S}}$ – przestrzeń wejść zewnętrznych systemu złożonego, podzbiór \tilde{S} -wymiarowej przestrzeni liczb rzeczywistych

$X_N = [x_1 \ x_2 \ \cdots \ x_N]$ – macierz wyników pomiarów wartości składowych wektora wejść zewnętrznych systemu złożonego, seria identyfikująca

$y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix}, \bar{y} = \begin{bmatrix} \bar{y}^{(1)} \\ \bar{y}^{(2)} \\ \vdots \\ \bar{y}^{(L)} \end{bmatrix}$ – L -wymiarowy wektor wyjść obiektu i modelu, odpowiednio,
 $y, \bar{y} \in Y \subseteq \mathbb{R}^L$

$Y \subseteq \mathbb{R}^L$ – przestrzeń wyjść obiektu i modelu – podzbiór L -wymiarowej przestrzeni liczb rzeczywistych

$y_m = \begin{bmatrix} y_m^{(1)} \\ y_m^{(2)} \\ \vdots \\ y_m^{(L_m)} \end{bmatrix}, \bar{y}_m = \begin{bmatrix} \bar{y}_m^{(1)} \\ \bar{y}_m^{(2)} \\ \vdots \\ \bar{y}_m^{(L_m)} \end{bmatrix}$ – L_m -wymiarowy wektor wejść m -tego elementu systemu złożonego i jego modelu, odpowiednio:
 $y_m, \bar{y}_m \in Y_m \subseteq \mathbb{R}^{L_m}$

$Y_m \subseteq \mathbb{R}^{L_m}$ – przestrzeń wyjść obiektu i modelu m -tego elementu systemu złożonego i jego modelu, podzbiór L_m -wymiarowej przestrzeni liczb rzeczywistych

$Y_N = [y_1 \ y_2 \ \cdots \ y_N]$ – macierz wyników pomiarów wartości składowych wektora wyjść obiektu, wyniki eksperymentu

$Y_{N_1 n_2} = [y_{1n_2} \ y_{2n_2} \ \cdots \ y_{N_1 n_2}]$ – macierz wyników pomiarów wartości składowych wektora wyjść obiektu w n_2 -tym pomiarze na drugim stopniu, tj. dla ustalonego wejścia na drugim stopniu $u_2 = u_{2n_2}$, wyniki eksperymentu na pierwszym stopniu

z_n – zakłócenia w n -tym pomiarze, wartość zmiennej losowej z ,
 $z_n \in Z \subseteq \mathbb{R}^L$

$Z \subseteq \mathbb{R}^L$ – przestrzeń zakłóceń, podzbiór L -wymiarowej przestrzeni liczb rzeczywistych

$\theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix}$ – R -wymiarowy wektor parametrów charakterystyki statycznej obiektu,
 $\theta \in \Theta \subseteq \mathbb{R}^R$

- θ – R -wymiarowy wektor parametrów charakterystyki statycznej obiektu – w modelu Bayesa wartość zmiennej losowej $\underline{\theta}$
- $\Theta \subseteq \mathbb{R}^R$ – przestrzeń parametrów charakterystyki statycznej obiektu – podzbiór R -wymiarowej przestrzeni liczb rzeczywistych
- $\tilde{\theta} = \begin{bmatrix} \tilde{\theta}^{(1)} \\ \tilde{\theta}^{(2)} \\ \vdots \\ \tilde{\theta}^{(\tilde{R})} \end{bmatrix}$ – \tilde{R} -wymiarowy wektor parametrów charakterystyki statycznej systemu złożonego $\tilde{\theta} \in \tilde{\Theta} \subseteq \mathbb{R}^{\tilde{R}}$
- $\tilde{\Theta} \subseteq \mathbb{R}^{\tilde{R}}$ – przestrzeń parametrów charakterystyki statycznej systemu złożonego – podzbiór \tilde{R} -wymiarowej przestrzeni liczb rzeczywistych
- $\tilde{\theta} = \Gamma(\theta)$ – zależność pomiędzy parametrami charakterystyki statycznej systemu złożonego, systemu złożonego a parametrami wszystkich elementów
- θ^* – optymalna wartość wektora parametrów funkcji aproksymującej przy założeniu, że $u \in D_u$ z funkcją wagi jakości przybliżenia $g_u(u)$
- Θ^* – dopuszczalny zbiór parametrów systemu złożonego zapewniający żadaną dokładność modeli lokalnych
- θ_1 – R_1 -wymiarowy wektor parametrów charakterystyki statycznej na pierwszym stopniu, $\theta_1 \in \Theta_1 \subseteq \mathbb{R}^{R_1}$
- $\bar{\theta}_1$ – R_1 -wymiarowy wektor wyjść modelu drugim stopniu, $\bar{\theta}_1 \in \Theta_1 \subseteq \mathbb{R}^{R_1}$
- $\Theta_1 \subseteq \mathbb{R}^{R_1}$ – przestrzeń parametrów charakterystyki statycznej na pierwszym stopniu oraz wyjść modelu na drugim stopniu, podzbiór R_1 -wymiarowej przestrzeni liczb rzeczywistych
- θ_1^* – optymalna wartość wektora parametrów na pierwszym stopniu przy pełnej informacji probabilistycznej
- $\hat{\theta}_{1N_1n_2}$ – oszacowanie wartości wektora parametrów θ_1 charakterystyki statycznej na pierwszym stopniu, wynik identyfikacji na pierwszym stopniu dla zadanej serii identyfikującej $U_{1N_1n_2}$ oraz ustalonego wejścia na drugim stopniu $u_2 = u_{2n_2}$
- θ_2 – R_2 -wymiarowy wektor parametrów charakterystyki statycznej na drugim stopniu, $\theta_2 \in \Theta_2 \subseteq \mathbb{R}^{R_2}$

- $\Theta_2 \subseteq \mathbb{R}^{R_2}$ – przestrzeń parametrów charakterystyki statycznej na drugim stopniu, podzbiór R_2 -wymiarowej przestrzeni liczb rzeczywistych
- θ_2^* – optymalna wartość wektora parametrów modelu na drugim stopniu przy pełnej informacji probabilistycznej
- $\hat{\theta}_{2N_2}$ – oszacowanie wartości wektora parametrów θ_2 charakterystyki statycznej na drugim stopniu, wynik identyfikacji na podstawie N_2 pomiarów na drugim stopniu
- $\tilde{\theta}_{2N_1N_2}$ – oszacowanie wartości wektora parametrów θ_2 charakterystyki statycznej, wynik identyfikacji z wykorzystaniem algorytmu bezpośredniego na podstawie N_1 pomiarów na pierwszym stopniu i N_2 pomiarów na drugim stopniu
- $\theta_m = \begin{bmatrix} \theta_m^{(1)} \\ \theta_m^{(2)} \\ \vdots \\ \theta_m^{(R_m)} \end{bmatrix}$ – R_m -wymiarowy wektor parametrów charakterystyki statycznej m -tego elementu systemu złożonego, $\theta_m \in \Theta_m \subseteq \mathbb{R}^{R_m}$
- $\Theta_m \subseteq \mathbb{R}^{R_m}$ – przestrzeń parametrów charakterystyki statycznej m -tego elementu systemu złożonego, podzbiór R_m -wymiarowej przestrzeni liczb rzeczywistych
- $\theta_{mN_m}^*$ – lokalnie optymalna wartość wektora parametrów m -tego elementu systemu złożonego wyznaczona na podstawie N_m pomiarów
- $\hat{\theta}_N$ – oszacowanie wartości wektora parametrów charakterystyki statycznej θ , wyznaczone na podstawie N pomiarów
- θ_N^* – optymalna wartość wektora parametrów modelu wyznaczona na podstawie N pomiarów
- $\tilde{\theta}_N^*$ – globalnie optymalna wartość wektora parametrów modelu systemu złożonego wyznaczona na podstawie N pomiarów
- $\tilde{\theta}_N \approx \theta_N^*$ – numeryczne przybliżenie optymalnej wartości parametrów modelu θ_N^*
- $\bar{\theta}_N^*$ – optymalna wartość wektora parametrów modelu systemu złożonego wyznaczona na podstawie N pomiarów przy syntetycznym wskaźniku jakości identyfikacji

- $\tilde{\theta}_N^{**}$ – globalnie optymalna wartość wektora parametrów modelu systemu złożonego wyznaczona na podstawie N pomiarów, przy założeniu, że modele lokalne spełniają żądaną dokładność
- $\mu_{j-1} = \begin{bmatrix} \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(l_{j-1})} \end{bmatrix}$ – wejście neuronu E_{ij}
- $\mu_j^{(i)} = \phi_{ij} \left(\sum_{s=1}^{l_{j-1}} \mu_{j-1}^{(s)} \theta_{ij}^{(s)} + \theta_{ij}^{(0)} \right)$ – wyjście neuronu E_{ij}
- $\hat{\Xi}_{1N_1N_2} = \begin{bmatrix} \hat{\theta}_{1N_11} & \hat{\theta}_{1N_12} & \cdots & \hat{\theta}_{1N_1N_2} \end{bmatrix}$ – macierz oszacowań wartości wektora parametrów θ_1 wyznaczonych na pierwszym stopniu dla zadanej serii identyfikującej U_{2N_2} na drugim stopniu
- ϕ_{ij} – funkcja aktywacji neuronu E_{ij}
- $\Phi(u, \theta)$ – model obiektu statycznego – przybliżenie charakterystyki statycznej z wektorem wejść u i wektorem parametrów θ
- $\Phi(x, \theta)$ – model obiektu statycznego – przybliżenie charakterystyki statycznej systemu złożonego z wektorem wejść zewnętrznych x i wektorem parametrów θ
- $\Phi(u)$ – nieparametryczny model – przybliżenie charakterystyki statycznej z wektorem wejść u
- Φ^* – optymalny nieparametryczny model przy pełnej informacji probabilistycznej
- $\Phi_1(u_1, \theta_1)$ – model obiektu statycznego – przybliżenie charakterystyki statycznej obiektu na pierwszym stopniu z wektorem wejść u_1 i wektorem parametrów θ_1
- $\Phi_2(u_2, \theta_2)$ – model obiektu statycznego – przybliżenie charakterystyki statycznej obiektu na drugim stopniu z wektorem wejść u_2 i wektorem parametrów θ_2
- $\Psi_N(U_N, Y_N)$ – algorytm identyfikacji, algorytm przetworzenia macierzy pomiarów wejścia i wyjścia obiektu
- $\tilde{\Psi}_N(X_N, V_N)$ – algorytm identyfikacji dla systemu złożonego, algorytm przetworzenia macierzy pomiarów wejść zewnętrznych i wyróżnionych wyjść systemu złożonego

-
- $\Psi_{1N_1}(U_{1N_1n_2}, Y_{N_1n_2})$ – algorytm identyfikacji na pierwszym stopniu
 $\Psi_{2N_2}(U_{2N_2}, \hat{\Sigma}_{1N_1N_2})$ – algorytm identyfikacji na drugim stopniu
 $\tilde{\Psi}_{2N_1N_2}(\bar{U}_{1N_1N_2}, U_{2N_1N_2}, \bar{W}_{N_1N_2})$ – algorytm identyfikacji otrzymany w wyniku złożenia algorytmów z pierwszego i drugiego stopnia
 $\tilde{\Psi}_{2N_1N_2}(\bar{U}_{1N_1N_2}, U_{2N_2}, \bar{W}_{N_1N_2})$ – bezpośredni algorytm identyfikacji
 $\Omega \subseteq \mathbf{R}^L$ – przestrzeń wielkości losowych w obiekcie, podzbiór L -wymiarowej przestrzeni liczb rzeczywistych
 ω – L -wymiarowy wektor wielkości losowych, wartość zmiennej losowej $\underline{\omega}$, $\omega \in \Omega \subseteq \mathbf{R}^L$

1. Identyfikacja systemów złożonych – pojęcia podstawowe

1.1. Wstęp

Problematyka identyfikacji obiektów, tj. ustalania modeli obiektów na podstawie badań eksperymentalnych, po pewnym okresie stagnacji przeżywa obecnie nowy intensywny etap rozwoju. Wiąże się to z jednej strony z praktyczną koniecznością rozpatrywania nowych obiektów i celów identyfikacji, a z drugiej strony – z nowymi możliwościami realizacji algorytmów identyfikacji. W pierwszym przypadku chodzi o aktualne problemy identyfikacji systemów złożonych, rozumianych jako obiekt, w którym wyróżniono podstawowe elementy składowe i wskazano powiązania między nimi. Elementy składowe takiego systemu złożonego stanowią samoistne obiekty, a ich złożenie pozwala na opis procesów złożonych. W drugim przypadku mamy na myśli nowe możliwości realizacji komputerowej, a mianowicie zastosowania technik komputerowych w rozproszonych systemach gromadzenia i przetwarzania informacji. Obszar aktualnych zagadnień metodologicznych i technicznych leży na pograniczu automatyki i informatyki, z wyraźną przewagą po stronie informatyki.

Nowe problemy identyfikacji wiążą się z aktualnymi problemami zautomatyzowanych komputerowych systemów zarządzania i sterowania. Dotyczy to zdecentralizowanego sterowania systemami złożonymi [2, 22, 42, 53–55, 71] i elastycznymi procesami produkcji [18, 19, 22, 112, 115, 130] oraz sterowania w wielopoziomowych systemach zarządzania i sterowania [3, 23, 80, 84]. Zwiększa się także rola metod i technik identyfikacji w zadaniach praktycznych, których celem jest poznanie obiektu. Dotyczy to zwłaszcza komputeryzacji prac eksperymentalnych w laboratoriach badawczych [7, 23, 70, 72], a także identyfikacji obiektów o różnej naturze, np. biologicznych [6, 45, 90, 91, 93, 97, 98, 107], ekonomicznych lub technicznych [4, 35, 103, 109, 111, 114]. Często modele obiektów są wykorzystywane do celów decyzyjnych lub diagnostycznych [11, 12, 104, 106]. Znaczący obszar aktualnych zadań identyfikacji związany jest z opracowaniem algorytmów identyfikacji dla specjalnych modeli [52, 60, 61, 113, 117, 133, 134].

Niniejsza praca dotyczy obszernej problematyki związanej z identyfikacją systemów złożonych. Wśród systemów złożonych można wyróżnić systemy wejściowo-

-wyjściowe [2, 3, 46] oraz systemy sieciowe [17–19]. Wejściowo-wyjściowe systemy złożone to takie, w których wyróżniono elementy składowe z określonymi wejściami i wyjściami, a struktura systemu zadana jest przez podanie powiązań pomiędzy wejściami i wyjściami elementów systemu. Systemy sieciowe to takie, w których wyróżniono elementy składowe, a powiązania pomiędzy elementami zadane są uwarunkowaniami czasowymi [18]. Przykładem takich systemów stosowanych do opisu systemów produkcji są kompleksy operacji, gdzie czas wykonania operacji zależy od wielkości zadania lub zasobu przydzielonego do operacji, a struktura kompleksu – od uwarunkowań czasowych. Kolejne operacje mogą się rozpocząć po zakończeniu wcześniejszych. Wynika to z procesu produkcyjnego. Uwarunkowania te określają sieciową strukturę kompleksu.

W dalszych rozważaniach ograniczymy się do systemów wejściowo-wyjściowych [122, 132, 135].

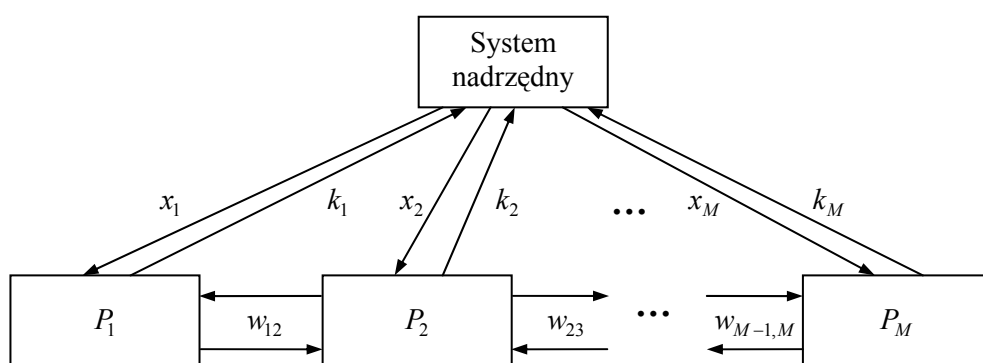
Rozdział 1. to wprowadzenie w problematykę identyfikacji obiektów. Przedstawiono w nim rolę modelu w badaniach systemowych, sformułowano zadanie identyfikacji obiektów oraz zaproponowano opis wejściowo-wyjściowych systemów złożonych. Wstępnie określono podstawowe zagadnienia identyfikacji systemów złożonych. W rozdziale 2. przedstawiono podstawowe zadania identyfikacji obiektów. Ograniczono się do identyfikacji charakterystyk statycznych. Zaznaczono w ten sposób podział na charakterystyczne, typowe zadania identyfikacji, które mogą wystąpić przy różnych założeniach dotyczących sytuacji pomiarowej oraz znajomości obiektu. W rozdziale 3. omówiono zadanie identyfikacji statycznego systemu złożonego, w którym tylko wybrane wielkości mogą być mierzone. Prowadzi to do zadania identyfikacji systemów przy ograniczonych możliwościach pomiarowych. W rozdziale 4. podano opis zadania identyfikacji lokalnej i globalnej statycznego systemu złożonego, a także algorytm identyfikacji dla struktury szeregowej oraz sformułowano zadania identyfikacji globalnej z uwzględnieniem jakości modeli lokalnych. W rozdziale 5. przedstawiono zagadnienie identyfikacji wielostopniowej. Skoncentrowano się na dwustopniowych algorytmach estymacji parametrów oraz dwustopniowym zadaniu wyboru optymalnego modelu.

1.2. Rola modelu w badaniach systemowych

Badanie obiektów o różnej naturze (np. technicznych, biologicznych, ekonomicznych) prowadzi do zebrania spostrzeżeń, danych, informacji na temat obserwowanych zjawisk. Zebrana i uporządkowana wiedza na temat obserwowanego obiektu tworzy model badanej rzeczywistości. Model ten jest stosowany do celów badawczych. Na bazie wiedzy o badanym obiekcie formułowane są nowe problemy poznawcze. Model służy do formułowania zadań projektowych. Jest podstawą do formułowania diagnozy. Wartości parametrów modelu niosą informację o aktualnym stanie badanego obiektu. Na podstawie wiedzy o obiekcie formułowane są metody i algorytmy sterowania i zarządzania.

Biorąc pod uwagę postać modelu, jego zakres oraz sposób wykonania, a także historyczny rozwój, możemy wyróżnić następujące modele:

Modele konceptualne – odpowiadają na pytania: Jak system jest zbudowany? Jaki jest cel i zakres działania danego systemu? Zwykle modele te oddają koncepcje funkcjonowania badanego procesu, obiektu czy zjawiska. Pokazują one, jak taki system jest zorganizowany. Wskazują na jego elementy składowe i powiązania między nimi. Modele te często są przedstawiane w postaci schematów blokowych lub też opisu działania systemu. Na rysunku 1.1 przedstawiono koncepcję funkcjonowania dwupoziomowego systemu zarządzania. Wyodrębniono w nim system nadrzędny oraz M podsystemów: P_1, P_2, \dots, P_M . Określono zmienne decyzyjne x_1, x_2, \dots, x_M , przekazywane z systemu nadrzędnego do podsystemów, oraz zmienne koordynacyjne k_1, k_2, \dots, k_M , pochodzące z podsystemów. Zwrócono uwagę na wzajemne powiązania pomiędzy systemami poprzez zmienne $w_{ij}, i = 1, 2, \dots, M, j = 1, 2, \dots, M$.



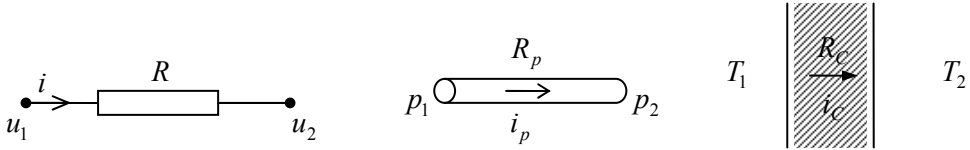
Rys. 1.1. Dwupoziomowy system zarządzania

Modele fizyczne umożliwiają badanie danego zjawiska lub obiektu w skali laboratoryjnej. Często są to modele uproszczone i pomniejszone, ale z zachowaniem fizycznej natury badanego procesu. Zmiana skali jest zwykle podyktowana kosztem budowy modelu. Uproszczenia mają na celu uwypuklenie badanego zjawiska. Przykładem zastosowania modelu fizycznego jest badanie w tunelu aerodynamicznym oporu powietrza stawianego przez dany model samochodu. Badanie nie wymaga kompletnego modelu pojazdu. Wystarczy zamodelować jego kształt.

Modele analogowe korzystają z analogii fizykochemicznych pomiędzy badanymi procesami. Na przykład: rezystancja, opór hydrauliczny oraz opór cieplny są opisane podobną zależnością. W modelu analogowym taki obiekt jest elementem proporcjonalnym i opisujemy go zależnością $y = \theta u$, gdzie: u, y są – odpowiednio – wejściem i wyjściem obiektu, natomiast θ jest współczynnikiem proporcjonalności.

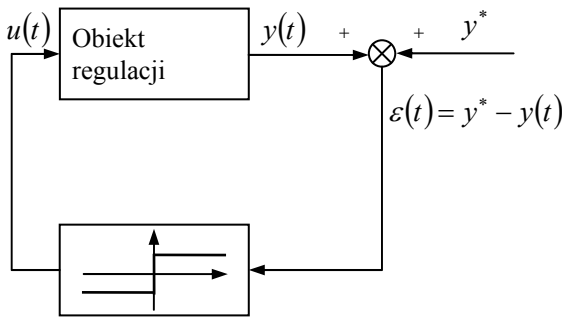
Takie ujęcie problemu pozwala na abstrahowanie od natury zjawiska. Badane zjawisko może być przedstawione za pomocą procesów o innej naturze fizycznej. Stwarza to nowe możliwości badawcze. Zmiana natury zjawiska może być podyktowana względami bezpieczeństwa lub przesłankami ekonomicznymi. Zmiana natu-

ry zjawiska może ponadto zmienić skalę czasu. Na przykład wolnozmiennie procesy ciepłe możemy zasymulować szybkimi procesami elektrycznymi. Poczytne miejsce w problemach modelowania, ze względu na ich uniwersalność i dostępność w poprzednim okresie, znalazły analogowe modele elektryczne. Opracowano zestawy typowych bloków umożliwiających modelowanie obiektów, które zyskały nazwę komputerów analogowych.



Rys. 1.2. Przykłady elementów proporcjonalnych:
a) rezystancja R , b) opór przepływu R_p , c) opór cieplny R_C

Modele matematyczne przedstawiają związki pomiędzy wielkościami w badanym procesie. Opisują własności statyczne obiektu, podawane zwykle w postaci zależności funkcyjnych, równań i nierówności. Własności dynamiczne są przedstawiane za pomocą równań i nierówności różniczkowych dla procesów ciągłych oraz równań i nierówności różnicowych dla procesów dyskretnych. Na rysunku 1.3 przedstawiono układ regulacji przekaźnikowej.



Rys. 1.3. Układ regulacji

Celem układu jest stabilizacja wyjścia $y(t)$ obiektu całkującego, z inercją o wejściu $u(t)$, na zadanym poziomie y^* . Obiekt regulacji jest opisany zależnością

$$\frac{d^2 y(t)}{dt^2} + \frac{1}{T} \frac{dy(t)}{dt} = K u(t), \quad (1.1)$$

z warunkami początkowymi

$$y(t)|_{t=0} = y_0, \quad \frac{dy(t)}{dt}|_{t=0} = y_1, \quad (1.2)$$

gdzie: parametr K jest stałą wzmocnienia, a T jest stałą czasową.

Na wyjściu regulatora przekaźnikowego otrzymujemy znak błędu regulacji, czyli

$$u(t) = \text{sign}(\varepsilon(t)), \quad (1.3)$$

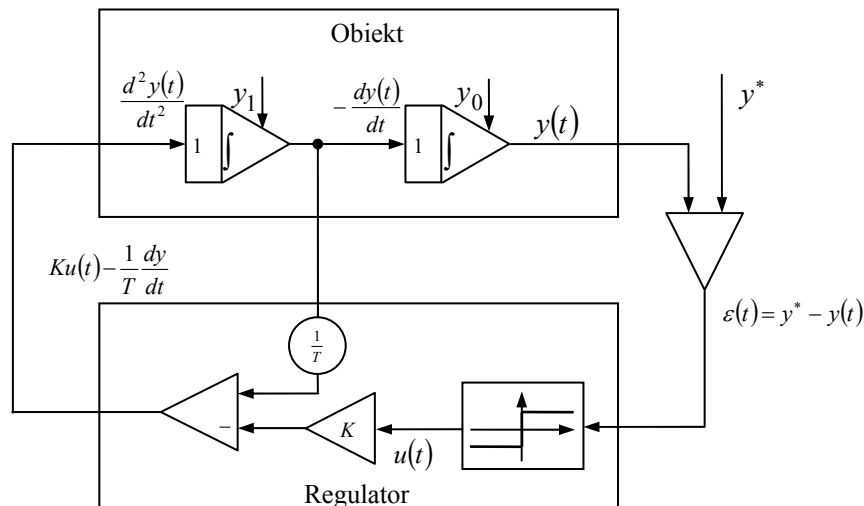
gdzie błąd regulacji

$$\varepsilon(t) = y^* - y(t). \quad (1.4)$$

Równania (1.1)–(1.4) to model matematyczny rozpatrywanego układu regulacji. Modelem matematycznym są też zestawy zdań logicznych wypowiedzianych na temat obserwowanego obiektu. Ogólniej – zebrane obserwacje dotyczące badanego obiektu tworzą bazę wiedzy, a jeśli ograniczymy się do konkretnego obszaru, zbiór ten jest wiedzą dziedzinową, przydatną do projektowania systemów ekspertowych [20, 73].

Modele komputerowe są opracowane na bazie opisów matematycznych. Biorąc pod uwagę sposób wykonania, możemy wyróżnić modele cyfrowe lub analogowe. Do symulacji ciągłych procesów dynamicznych często były stosowane tzw. komputery analogowe. Komputer analogowy stanowi zestaw specjalizowanych typowych bloków (m.in. elementów całkujących, wzmacniaczy, sumatorów, elementów nieliniowych), realizowanych za pomocą układów elektronicznych. Odpowiednie połączenie elementów umożliwi symulację badanego procesu. Takie modele badanego procesu przyjęło się nazywać modelami analogowymi.

Na rysunku 1.4 przedstawiono model analogowy obiektu opisanego równaniami (1.1)–(1.4).



Rys. 1.4. Przykład modelu analogowego

Modele cyfrowe to zwykle symulacja badanego obiektu za pomocą odpowiedniego programu komputerowego, opracowanego na podstawie modelu matematycznego. Rozwój techniki obliczeniowej, metod numerycznych oraz inżynierii oprogramowania

dostarczył wiele gotowych bibliotek i pakietów, które umożliwiają modelowanie procesów o różnej naturze, jak na przykład dyskretnych procesów produkcyjnych, procesów ekonomicznych, układów regulacji czy też procesów biologicznych. Można tutaj wymienić między innymi biblioteki dostępne na platformie programowej MATLAB, w tym pakiet SIMULINK, oraz system MULTI-EDIP z aplikacją SYMULATOR [70, 72, 87]. Wspecjalizowane pakiety pozwalają na modelowanie procesów ciągłych, zdecydowanie wypierając wcześniej wspomniane modele analogowe. Istotną przewagą modeli cyfrowych nad analogowymi polega na prostej możliwości modelowania wielowymiarowych złożonych procesów. Współczesne możliwości techniki obliczeniowej zapewniają odpowiednią dokładność oraz czas uzyskania wyników symulacji. Bardzo ważnym i aktualnym zagadnieniem jest wykorzystanie bazy wiedzy na temat obserwowanego obiektu, zwanej wiedzą dziedzinową, do modelowania [20, 73]. Wiedza dziedzinowa wraz z regułami wnioskowania są podstawą do budowy systemów ekspertowych. Zadanie eksploracji wiedzy na bazie zaobserwowanych faktów można porównać z tradycyjnie rozumianym zadaniem identyfikacji.

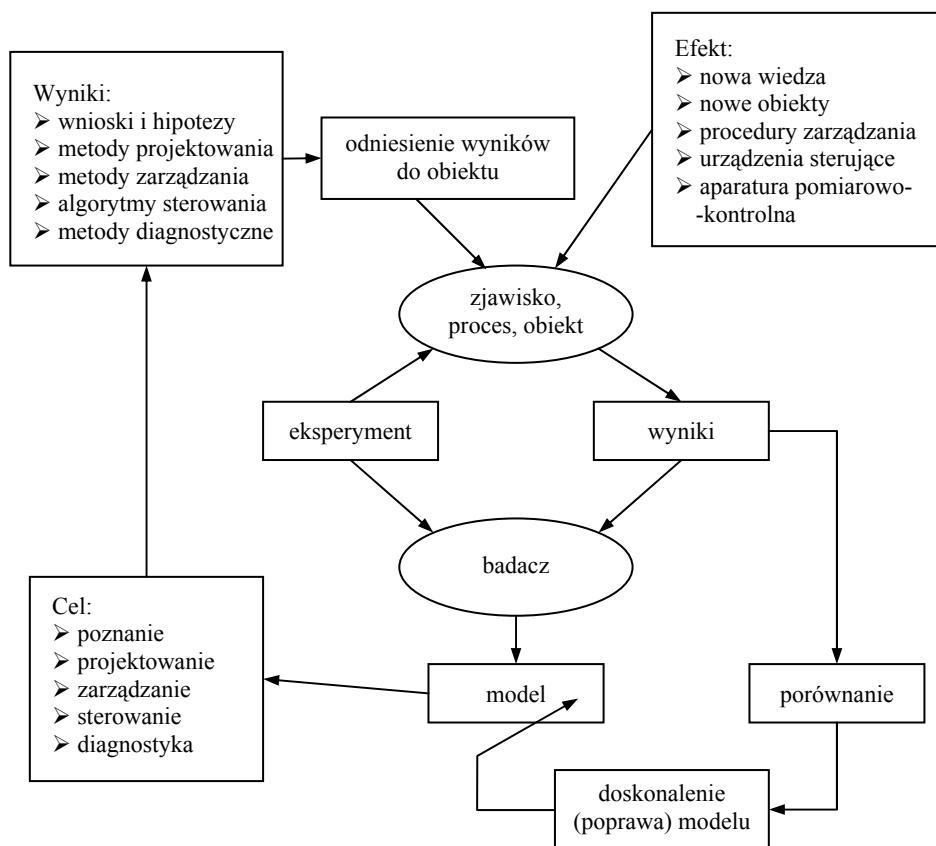
Proces tworzenia modelu oraz sposób jego wykorzystania przedstawiono na rysunku 1.5. Badacz – w celu poznania obiektu, procesu lub zjawiska – prowadzi eksperyment. W wyniku otrzymuje dane pomiarowe, informacje i spostrzeżenia na temat badanego zjawiska. Uporządkowane wyniki eksperymentu, podawane w postaci wzorów, równań, reguł i zdań logicznych, tworzą model badanego procesu. Model taki nie od razu zadowala eksperymentatora.

Porównanie wyników eksperymentu z danymi uzyskanymi z modelu często prowadzi do konieczności doskonalenia modelu. Model jest zwykle budowany w konkretnym celu. Na podstawie uzyskanego modelu formułowane są zadania poznawcze, zadania projektowe, problemy zarządzania i sterowania obiektem oraz zadania diagnozy. Po rozwiązaniu tych zadań wysnuwamy wnioski i hipotezy, uzyskujemy rekomendację dla projektów nowych obiektów, opracowujemy metodykę tworzenia algorytmów zarządzania, sterowania i diagnozy. Wnioski i hipotezy są weryfikowane przez badanie rzeczywistego obiektu. Daje to nową wiedzę o badanych zjawiskach i procesach. Algorytmy zarządzania i sterowania są wykorzystywane do projektowania systemów zarządzania oraz urządzeń sterujących. Opracowane algorytmy diagnostyczne są podstawą do budowy urządzeń pomiarowo-kontrolnych.

W celu ilustracji zadania tworzenia modelu i jego wykorzystania w procesie poznawczym i projektowym pokrótce omówimy badania kolumn destylacyjnych, prowadzone w Instytucie Inżynierii Chemicznej Urzędzeń Ciepłych Politechniki Wrocławskiej [57, 76, 77]. Do przykładu tego wrócimy w rozdziale 5.

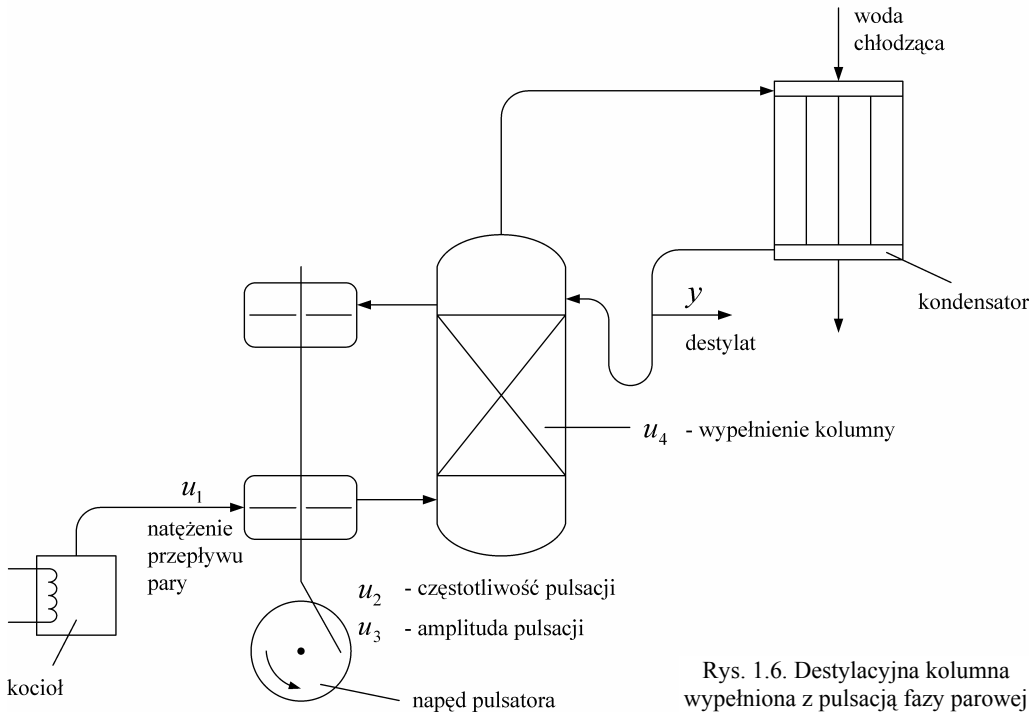
Rozdzielenie ciekłego roztworu na odpowiednie frakcje lub poszczególne czyste składniki wymaga budowy kolumn o dużej zdolności rozdzielczej. Tradycyjne rozwiązania wiązały się z projektowaniem kolumn destylacyjnych o dużej wysokości lub też kaskadą kilku niższych, co się wiąże z dużymi kosztami. Prowadzone badania nad zjawiskami towarzyszącymi przepływowi faz i wymianie masy pokazały, że intensyfikację wymiany masy w układzie ciecz–gaz osiąga się przede wszystkim przez

zwiększenie powierzchni kontaktu faz i zmniejszenie sumarycznego oporu przenikania masy, zakłócając stabilność warstw granicznych na powierzchni styku. Efekt ten można uzyskać przez zwiększenie burzliwości naturalnego przepływu faz, stawiając przeszkody na drodze naturalnego przepływu, jak również wywołanie dodatkowego, zewnętrznego zaburzenia przepływu.



Rys. 1.5. Model w badaniach systemowych

Na tej podstawie zaprojektowano kolumnę destylacyjną wypełnioną z pulsacją fazy parowej (rys. 1.6). Zadaniem wypełnienia oraz pulsacji jest zwiększenie powierzchni kontaktu faz. Przed badaczem staje pytanie: Jaki jest wpływ rodzaju wypełnienia kolumny oraz wielkości charakteryzujących zaburzenie przepływu na ocenę sprawności kolumny w stanie ustalonym? Inaczej mówiąc, celem badań eksperymentalnych jest opracowanie modelu kolumny destylacyjnej wypełnionej z pulsacją fazy parowej, który opíše wpływ strumienia przepływu u_1 , częstotliwości u_2 oraz amplitudy pulsacji u_3 na sprawność y wypełnionej kolumny z pulsacją. Miarą sprawności kolumny y może być liczba pólek teoretycznych lub wartość objętościowego współczynnika przenikania masy.



Rys. 1.6. Destylacyjna kolumna wypełniona z pulsacją fazy parowej

W celu opracowania modelu matematycznego opisanego obiektu identyfikacji przeprowadzono szereg pomiarów dla kolumny wypełnionej konkretnym rodzajem wypełnienia. W pierwszej kolejności, dla zadanej serii wartości strumienia przepływu u_1 , mierzono skład fazy parowej i na tej podstawie określano liczbę półek teoretycznych oraz wartość objętościowego współczynnika przenikania masy. Pozwoliło to na określenie zależności sprawności kolumny z ustalonym wypełnieniem od strumienia przepływu. Następnie pomiary przeprowadzono dla zadanych serii częstotliwości u_2 oraz amplitudy pulsacji u_3 . Kolejno badania te powtórzono dla kolumn z różnym wypełnieniem. Na podstawie uzyskanych danych pomiarowych opracowano odpowiednie modele matematyczne kolumn wypełnionych z pulsacją fazy parowej.

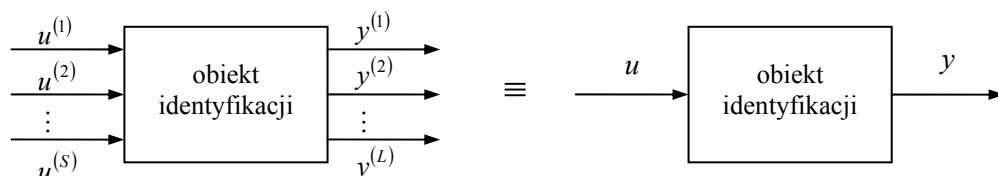
Opracowane modele matematyczne były podstawą do dalszych prac badawczych i projektowych. Wyniki badania wrażliwości wielkości charakteryzujących sprawność kolumny na zmianę strumienia przepływu, amplitudy i częstotliwości pulsacji były podstawą do opracowania rekomendacji dla projektantów wypełnionych kolumn destylacyjnych z pulsacją fazy parowej [57, 77, 138].

1.3. Zadanie identyfikacji obiektu

W celu wprowadzenia oraz ujednoczenia oznaczeń sformułujemy podstawowe zadania identyfikacji obiektu. Problematyka ta jest dokładnie przedstawiona w wielu

książkach na temat identyfikacji [8, 14, 29, 43, 44, 68, 69, 74, 83, 88]. Zawarte w tym podrozdziale rozważania oparto przede wszystkim na ujęciu przedstawionym w pracy [14] i ograniczono się do identyfikacji obiektów statycznych, stacjonarnych.

Celem identyfikacji jest stworzenie modelu matematycznego badanego obiektu na podstawie zebranych obserwacji (danych pomiarowych). Obiekt, dla którego będziemy ustalać zależność wielkości $y^{(1)}, y^{(2)}, \dots, y^{(L)}$ od $u^{(1)}, u^{(2)}, \dots, u^{(S)}$, nazwiemy obiektem identyfikacji, a wielkości $y^{(1)}, y^{(2)}, \dots, y^{(L)}$ oraz $u^{(1)}, u^{(2)}, \dots, u^{(S)}$ – odpowiednio – wielkościami wyjściowymi i wejściowymi (rys. 1.7).



Rys. 1.7. Obiekt identyfikacji z wektorem wejść u i wektorem wyjść y

W dalszych rozważaniach będziemy zakładać, że wejścia i wyjścia obiektu identyfikacji są – odpowiednio – S - oraz L -wymiarowymi wektorami, zapisywanymi w postaci:

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix}, \quad (1.5)$$

gdzie: u – wektor wejść obiektu, $u \in \mathcal{U} \subseteq \mathcal{R}^S$, \mathcal{U} – S -wymiarowa przestrzeń wejść, y – wektor wyjść obiektu, $y \in \mathcal{Y} \subseteq \mathcal{R}^L$, \mathcal{Y} – L -wymiarowa przestrzeń wyjść.

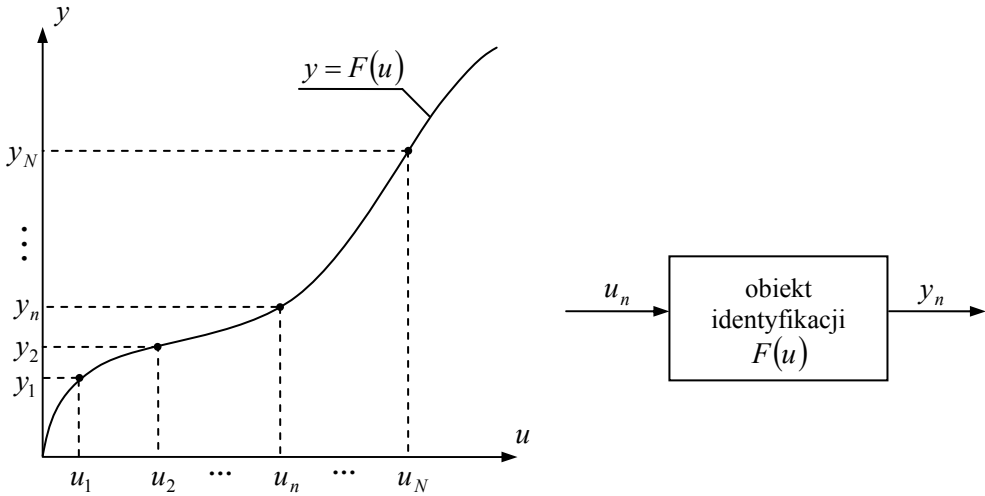
Podstawą do określenia modelu matematycznego obiektu są informacje aprioryczne oraz wyniki pomiarów. Na rysunku 1.8 przedstawiono wyniki pomiarów jednowymiarowego obiektu identyfikacji ($L = S = 1$), dla którego dokładna charakterystyka statyczna ma postać $y = F(u)$. Oznaczmy przez y_n wynik n -tego pomiaru wartości składowych wektora wyjść obiektu przy zadanej wartości składowych wektora wejść u_n , $n = 1, 2, \dots, N$, gdzie N jest liczbą wykonanych pomiarów. Wyniki pomiarów zapisujemy w postaci:

$$U_N = [u_1 \quad u_2 \quad \dots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \dots \quad y_N], \quad (1.6)$$

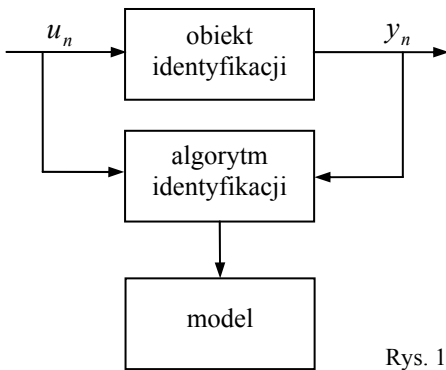
U_N, Y_N są macierzami, których kolumnami są wyniki kolejnych pomiarów – odpowiednio – wejść u i wyjść y .

Teraz naszym zadaniem jest wyznaczenie algorytmu identyfikacji, to jest algorytmu przetworzenia danych pomiarowych w celu wyznaczenia modelu obiektu. Postać algorytmu identyfikacji zależy od informacji apriorycznej o obiekcie, informacji o pomiarach

oraz przyjętej metody. Zrealizowany w postaci systemu komputerowego lub pewnego urządzenia technicznego („hardware’owego”) algorytm identyfikacji zwany jest identyfikatorem [14]. Obiekt identyfikacji wraz z identyfikatorem stanowi układ identyfikacji (rys. 1.9).



Rys. 1.8. Zestaw pomiarów charakterystyki statycznej $y = F(u)$ jednowymiarowego obiektu



Rys. 1.9. Układ identyfikacji

Na rysunku 1.9 pokazano przepływ informacji, tj. dane uzyskane w wyniku eksperymentu, przetworzone według odpowiedniego algorytmu identyfikacji, tworzą model obiektu.

Jak widać z przytoczonych rozważań (rys. 1.5), na proces tworzenia modelu matematycznego składają się różne czynności i zadania. Można je usystematyzować [14]:

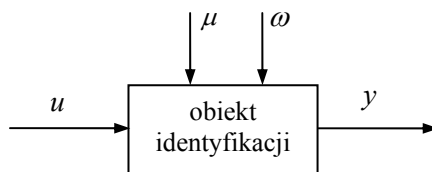
1. Określenie obiektu identyfikacji.
2. Określenie klasy modeli.
3. Organizacja eksperymentu.
4. Opracowanie algorytmu identyfikacji.
5. Realizacja algorytmu identyfikacji.

Proces tworzenia modelu matematycznego często nie jest czynnością jednorazową. Wymienione czynności są ściśle powiązane i nie mogą być wykonane niezależnie. Czasami wstępna analiza wyników eksperymentu powoduje, że konieczny jest powrót do określenia klasy modeli. Omówimy pokrótce zaproponowane czynności. Pozwoli to na przyjęcie nomenklatury i oznaczeń przydatnych w dalszych rozważaniach.

Określenie obiektu identyfikacji

Jest to bardzo ważny element zadania identyfikacji. W tym momencie należy określić, jaki jest cel identyfikacji, tzn. w jakim celu tworzymy model obiektu. Ten etap wymaga ścisłej współpracy projektanta modelu z jego użytkownikiem. Od tego zależy, jakie wielkości zostaną uwzględnione w modelu. W tym miejscu należy odpowiedzieć na pytania: Czy badamy charakterystyki statyczne czy też własności dynamiczne obiektu? Pomiedzy jakimi wielkościami szukamy zależności?

Odpowiedzi na te pytania prowadzą w konsekwencji do określenia wektora wielkości wejściowych u oraz wektora wielkości wyjściowych y (rys. 1.10). Analiza obiektu identyfikacji może ponadto wskazać na wielkości, które wpływają na wyjście obiektu, a nie są wskazanymi wielkościami wejściowymi. Mogą to być zakłócenia mierzalne μ lub niemierzalne ω .



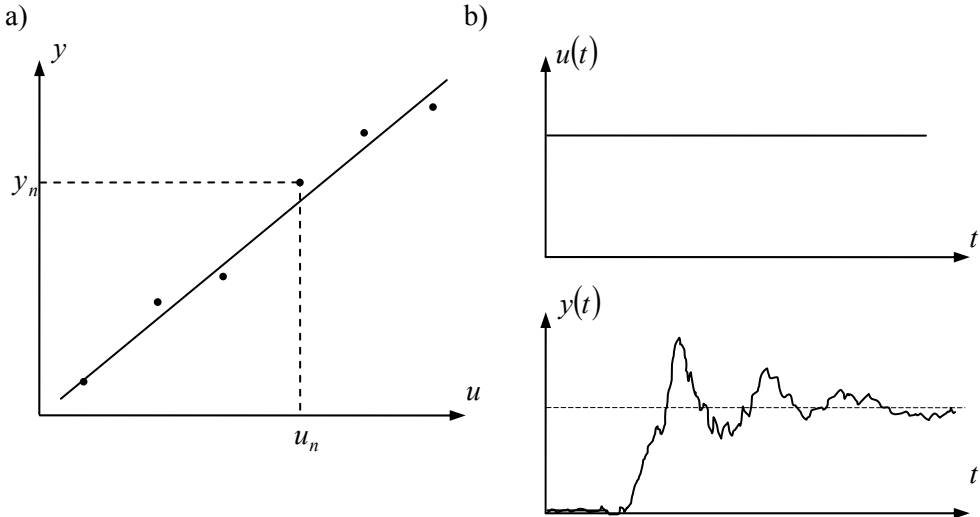
Rys. 1.10. Wielkości charakterystyczne obiektu identyfikacji

Określenie klasy modeli

Ważnym elementem zadania identyfikacji jest dobór postaci opisu do badanego obiektu. Sprowadza się to do ustalenia odpowiedniej postaci funkcji dla charakterystyk statycznych. W praktyce polega to na ustaleniu postaci opisu z dokładnością do parametrów. Podpowiedzią w doborze odpowiedniej postaci może być:

- Analiza zjawisk fizykochemicznych, która sprowadza się do opisu procesu na podstawie znanych praw fizyki, chemii, biologii, ekonomii itp. Zwykle w wyniku otrzymujemy opis obiektu znany z dokładnością do parametrów.
- Analiza wymiarowa, która na podstawie analizy jednostek występujących przy wielkościach wejściowych i wyjściowych może sugerować sposób ich przekształcenia, a tym samym postać modelu.
- Analiza danych eksperymentalnych, w wyniku której możemy uzyskać odpowiedź co do kształtu funkcji opisujących dane eksperymentalne. Na przykład dla danych przedstawionych na rysunku 1.11a proponujemy charakterystykę liniową, natomiast przedstawione na rysunku 1.11b zarejestrowane sygnały ciągłe kojarzymy z zakłóconą odpowiedzią liniowego układu drugiego rzędu z opóźnieniem, na który podano skok jednostkowy.

Często projektant modelu decyduje się na wybór arbitralny, który jest podyktowany względami praktycznymi, wynikającymi z celu modelowania. Na przykład analiza zjawisk daje złożoną nieliniową postać równań, które opisują obiekt identyfikacji, ale ze względu na przewidywane wykorzystanie modelu wygodniej jest przybliżyć obiekt modelem liniowym w okolicy punktu pracy.



Rys. 1.11. Przykładowe wyniki obserwacji: a) identyfikacja charakterystyki statycznej, b) identyfikacja charakterystyki dynamicznej

Podczas określania klasy modeli należy zwrócić uwagę na dwa istotnie różne przypadki. Zarówno analiza zjawisk fizykochemicznych, jak i analiza wymiarowa mogą dostarczyć dokładnego opisu obiektu, w którym znane są zależności funkcyjne i one dokładnie opisują badaną rzeczywistość. Jest to przypadek idealny. Problem sprowadza się tutaj do wyznaczenia parametrów w tych zależnościach. Mówimy wówczas, że opis jest znany z dokładnością do parametrów, a czasami [14] używamy określenia: „prawdziwy” opis należy do klasy modeli.

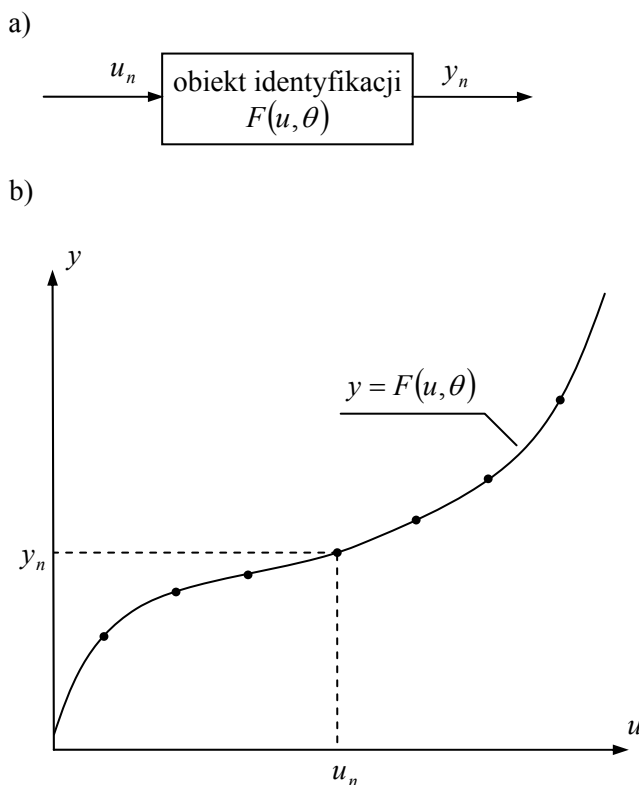
W przypadku charakterystyki statycznej dany jest opis

$$y = F(u, \theta), \quad (1.7)$$

gdzie: F jest znaną funkcją opisującą zależność wektora wyjść y od wektora wejść u , natomiast θ jest wektorem parametrów.

Na rysunku 1.12 przedstawiono jednowymiarowy obiekt identyfikacji ($L = S = 1$) o charakterystyce statycznej $y = F(u, \theta)$. W tym przypadku (rys. 1.12b) niezakłócona wartość pomiaru wyjścia y_n , przy zadanej wartości wejścia u_n , jest punktem z badanej charakterystyki. Obliczona wartość wyjścia modelu, przy znanej wartości składowych wektora parametrów θ , pokrywa się z wartością zmierzoną y_n dla tej samej wartości

wejścia u_n . Zadanie identyfikacji sprowadza się do wyznaczenia nieznanych wartości parametrów opisu. Parametry te mogą mieć sens fizyczny.



Rys. 1.12. „Prawdziwy” opis należy do klasy modeli: a) charakterystyka statyczna obiektu znana z dokładnością do parametrów, b) niezakłócone wyniki pomiarów

Przypadek drugi – znacznie częstszy – to taki, kiedy w wyniku analizy danych eksperymentalnych, lub arbitralnej decyzji, przyjmujemy model, który daje przybliżony opis badanej rzeczywistości. Rzeczywista charakterystyka statyczna obiektu $y = F(u)$ istnieje, ale nie jest znana. Dla charakterystyki statycznej proponujemy przybliżony opis, dalej zwany modelem

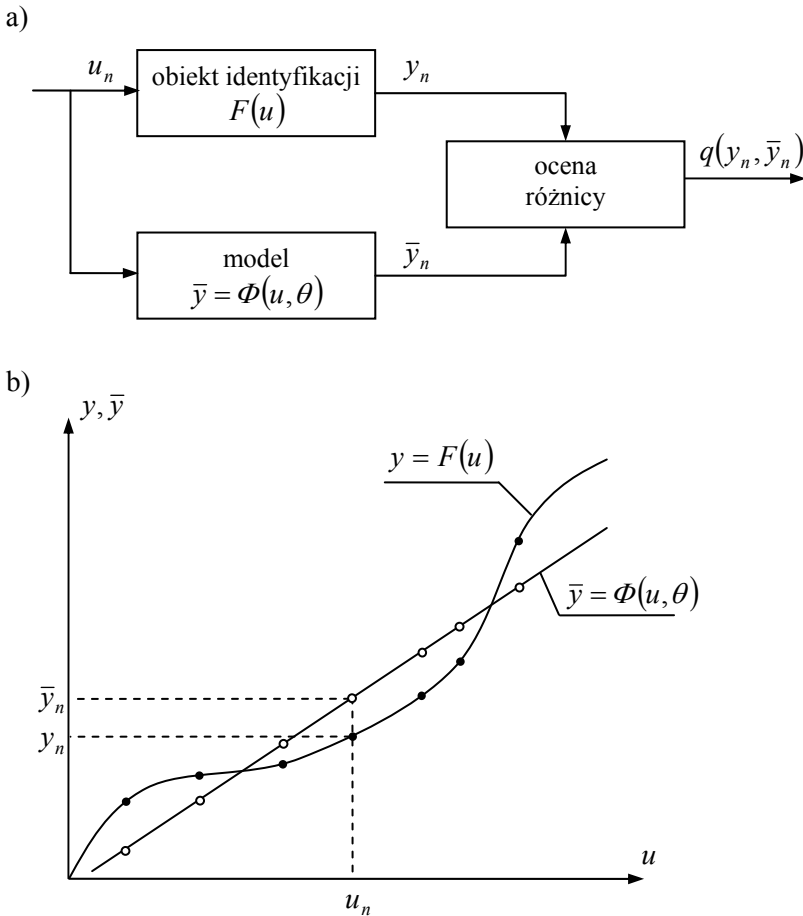
$$\bar{y} = \Phi(u, \theta), \quad (1.8)$$

gdzie: Φ jest zaproponowaną (zadaną) funkcją opisującą zależność wektora wyjść modelu \bar{y} od wektora wejść u , natomiast θ jest wektorem parametrów modelu.

W tym miejscu słowa „zaproponowana (zadana) funkcja” wymagają komentarza. Często propozycja ta jest wynikiem wielu prób doboru funkcji spośród tych nowych uniwersalnych aproksymatorów [1] oraz wynikiem badań mających na celu dobór

odpowiedniej struktury modelu. Może to być charakterystyka statyczna w postaci sieci neuronalnej [75, 82, 136].

Na rysunku 1.13 przedstawiono jednowymiarowy obiekt identyfikacji ($L = S = 1$) o charakterystyce statycznej $y = F(u)$ i proponowanym modelu $\bar{y} = \Phi(u, \theta)$. Tym razem obliczona, na podstawie modelu (1.8), wartość wyjścia \bar{y}_n może się różnić od wartości otrzymanej w wyniku niezakłóconego pomiaru y_n tej samej wartości wejścia u_n (rys. 1.13b). Zadanie identyfikacji polega na wyznaczeniu takich wartości parametrów modelu, dla których oceniona różnica pomiędzy wartościami obliczonymi z modelu a wynikami pomiarów wartości wyjścia obiektu będzie minimalna w sensie przyjętego kryterium jakości identyfikacji. Tym razem mówimy o zadaniu wyboru optymalnego modelu.



Rys. 1.13. Wybór optymalnego modelu: a) obiekt identyfikacji o nieznannej charakterystyce statycznej $y = F(u)$ oraz proponowany model $\bar{y} = \Phi(u, \theta)$, b) niezakłócone wyniki pomiarów \bullet oraz odpowiednie wartości wyjścia modelu \circ

Organizacja eksperymentu

W wyniku eksperymentu uzyskujemy następujące dane: dla zadanej sekwencji wartości składowych wektora wejść obiektu

$$U_N = [u_1 \quad u_2 \quad \dots \quad u_N], \quad (1.9)$$

mierzmy odpowiednie wartości składowych wektora wyjść i zapisujemy je w postaci macierzy

$$Y_N = [y_1 \quad y_2 \quad \dots \quad y_N], \quad (1.10)$$

gdzie N jest liczbą pomiarów.

Sekwencja (1.9) jest zwana serią identyfikującą, a macierz (1.10) – wynikiem eksperymentu [14]. W praktycznych sytuacjach wynik pomiaru jest obciążony błędem pomiarowym. Znaczący to, że w n -tym pomiarze, dla zadanych wartości składowych wektora wejść u_n , mierzmy wartości składowych wektora wyjść y_n , a na mierzonej wielkości nakładają się przypadkowe zakłócenia pomiarowe z_n . W wyniku pomiaru otrzymujemy wartość w_n różną od y_n , czyli

$$w_n = h(y_n, z_n), \quad (1.11)$$

gdzie h jest znaną funkcją opisującą sposób nakładania się zakłóceń pomiarowych na mierzoną wielkość.

W tym przypadku wyniki eksperymentu zapiszemy w postaci macierzy

$$W_N = [w_1 \quad w_2 \quad \dots \quad w_N]. \quad (1.12)$$

Przeprowadzenie eksperymentu wymaga ustalenia wielu czynników, które mają wpływ na jego wynik [14, 56]. Należą do nich:

➤ **Ustalenie charakteru eksperymentu**

wiąże się z określeniem możliwości przeprowadzenia eksperymentu czynnego lub biernego. Eksperyment czynny umożliwia wybór serii identyfikującej, co jest związane z planowaniem eksperymentu. W przypadku eksperymentu biernego pozostaje monitorowanie wartości wejść i wyjść obiektu.

➤ **Planowanie eksperymentu**

dotyczy wyłącznie eksperymentu czynnego i polega na odpowiednim wyborze serii identyfikującej. Przesłanki tego wyboru mogą być różne [56]. Zasadniczą przesłanką jest dobór serii identyfikującej, która zapewni odpowiednią jakość modelu oraz możliwość jego weryfikacji, inne mogą dotyczyć kosztów przeprowadzenia eksperymentu oraz prostoty obliczeń w algorytmie identyfikacji.

➤ **Określenie parametrów eksperymentu**

długości serii pomiarowej i zakresu zmian wielkości wejściowych.

➤ **Ustalenie techniki eksperymentu**

jest ściśle związane z poprzednimi czynnikami, a polega na technicznym zabezpieczeniu wykonania i rejestracji pomiarów, tzn. na doborze odpowiednich czujników, przyrządów i układów pomiarowych zapewniających wykonanie zaplanowanych pomiarów oraz ich przekazanie do identyfikatora.

Opracowanie algorytmu identyfikacji

Zebrane dane pomiarowe są podstawą do wyznaczenia modelu obiektu. Opracowanie algorytmu identyfikacji wiąże się z zaproponowaniem pewnego przepisu na przetworzenie danych pomiarowych w celu obliczenia parametrów modelu w przypadku modeli parametrycznych lub wyznaczenie modelu nieparametrycznego. Postać algorytmu identyfikacji zależy w dużej mierze od informacji apriorycznych o obiekcie oraz sytuacji pomiarowej. Gdy opis obiektu jest znany z dokładnością do parametrów, mówimy o zadaniu wyznaczenia parametrów w przypadku deterministycznym lub estymacji parametrów w przypadku probabilistycznym. Podobnego podziału, biorąc pod uwagę sytuację pomiarową, można dokonać dla zadania wyboru optymalnego modelu.

W rozdziale 2. dokonano przeglądu podstawowych zadań identyfikacji obiektów statycznych w warunkach deterministycznych oraz losowych.

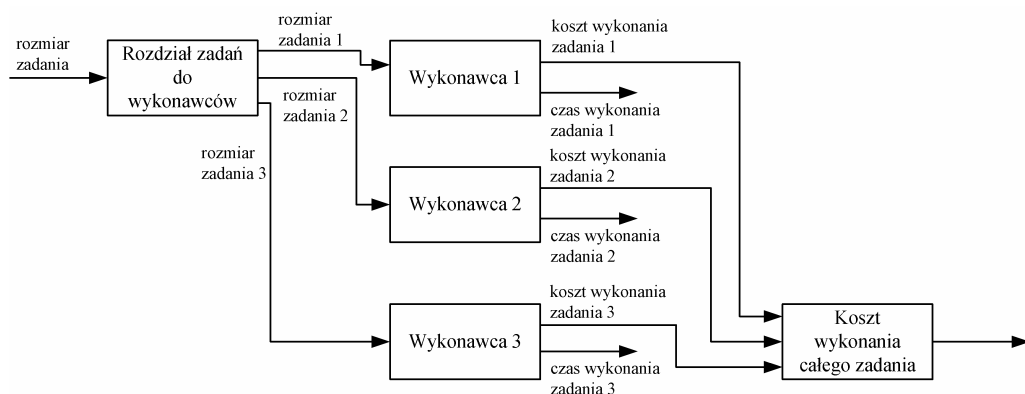
Realizacja algorytmu identyfikacji

W zależności od postaci danych pomiarowych oraz techniki ich przetwarzania algorytm identyfikacji może być zrealizowany w technice cyfrowej lub analogowej. W pierwszym przypadku jest to związane z opracowaniem odpowiedniego programu komputerowego, a w drugim – z zaprojektowaniem odpowiedniego analogowego przetwornika rejestrowanych w czasie eksperymentu sygnałów. Czasem algorytm opracowany w formie programowej lub analogowej zwany jest identyfikatorem [14]. Możemy też mówić o identyfikatorach uniwersalnych i specjalizowanych. Identyfikatory uniwersalne to na ogół oprogramowanie wspierające prace laboratoryjne i badawcze. Obecnie są projektowane systemy komputerowe, które oprócz odpowiedniego oprogramowania realizującego algorytmy identyfikacji zapewniają organizację eksperymentu, a poprzez układ wejść i wyjść analogowych i cyfrowych dla zadanych serii pomiarowych automatycznie, w czasie rzeczywistym, rejestrują wyniki pomiarów. Jako przykład możemy tutaj wymienić system MULTI-EDIP – system wspomagający prace laboratoryjne i badawcze [72], pakiet „identyfikacja systemów” na platformie MATLAB [65, 70]. Identyfikatory specjalizowane często stanowią fragmenty komputerowych systemów sterowania, np. jako elementy sterowania adaptacyjnego.

1.4. Opis systemu złożonego

Obecnie skupimy się na wejściowo-wyjściowych, statycznych systemach złożonych, czyli takich obiektach, w których wyróżniono powiązane między sobą elementarne części składowe. Wyróżnione elementy opisują fragmenty złożonego procesu i mogą stanowić niezależne obiekty. W przypadku każdego z nich wskazujemy na wielkości wejściowe i wyjściowe. Podane zależności pomiędzy poszczególnymi składowymi elementami, czyli podane zależności pomiędzy wejściami i wyjściami poszczególnych elementów, tworzą opis struktury takiego systemu złożonego.

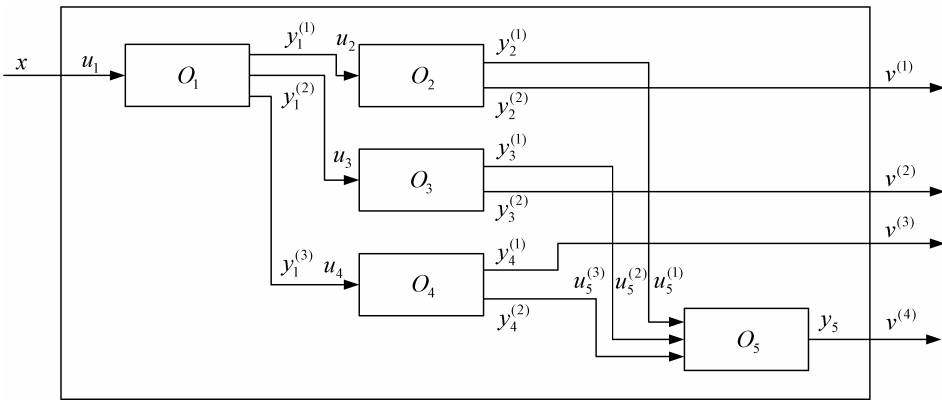
Przykładem złożonego systemu jest system produkcji. Rozważmy problem realizacji podzielonego zadania o zadanym całkowitym rozmiarze (rys. 1.14). Zadanie jest realizowane równoległe przez trzech wykonawców. Czas oraz koszt wykonania zadania jest różny dla każdego wykonawcy i zależy od rozmiaru przydzielonego zadania. W celu optymalnego zarządzania procesem wytwórczym konieczna jest znajomość charakterystyki wykonawców, a zwłaszcza znajomość zależności czasu oraz kosztu wykonania zadania w funkcji wielkości przydzielonego zadania. W powyższym procesie wytwarzania wyróżnimy elementarne czynności. Są to: rozdział zadań dla poszczególnych wykonawców, realizacja zadań przez wykonawców, ocena kosztu wykonania zadania.



Rys. 1.14. Schemat ideowy realizacji podzielonego zadania przez trzech wykonawców

Zadania te stanowią elementy złożonego systemu wytwarzania (rys. 1.15). Dla każdego z nich wskażemy wielkości wejściowe i wyjściowe. Rozdział zadań dla wykonawców O_1 jest elementem, w którym wielkością wejściową u_1 jest rozmiar całkowitego zadania do wykonania, natomiast wyjściem jest wektor y_1 o składowych $y_1^{(1)}$, $y_1^{(2)}$, $y_1^{(3)}$, które są rozmiarami zadań dla poszczególnych wykonawców. Opis zasad podziału zadań – $y_1^{(1)} = F_1^{(1)}(u_1)$, $y_1^{(2)} = F_1^{(2)}(u_1)$, $y_1^{(3)} = F_1^{(3)}(u_1)$, a w zwartym zapisie $y_1 = F_1(u_1)$ – zależność pomiędzy wyjściem y_1 a wejściem u_1 jest charakterystyką statyczną tego elementu. Realizację zadań przez poszczególnych wykonawców stanowią kolejne elementy O_2 , O_3 oraz O_4 procesu wytwarzania. W elemencie O_2 wielkością wejściową u_2 jest rozmiar zadania do wykonania przez pierwszego wykonawcę, a wyjściem jest wektor y_2 , o składowych $y_2^{(1)}$, $y_2^{(2)}$, którymi są – odpowiednio – koszt $y_2^{(1)}$ oraz czas $y_2^{(2)}$ wykonania zadania przez pierwszego wykonawcę. Koszt $y_2^{(1)}$ oraz czas $y_2^{(2)}$ wykonania zadania zależy od rozmiaru zadania u_2 ; funkcje opisujące te zależności to: $y_2^{(1)} = F_2^{(1)}(u_2)$ oraz $y_2^{(2)} = F_2^{(2)}(u_2)$, a w zwartym zapisie

$y_2 = F_2(u_2)$, zależność pomiędzy wyjściem y_2 a wejściem u_2 stanowi charakterystykę statyczną tego elementu. Analogicznie są określone wejścia u_3 i u_4 oraz wyjścia y_3 i y_4 dla obiektów O_3 i O_4 , które odpowiadają drugiemu i trzeciemu wykonawcy. Kolejny element O_5 stanowi ocenę kosztów wykonania zadania. Wektorem wejść u_5 , o składowych $u_5^{(1)}, u_5^{(2)}, u_5^{(3)}$, elementu O_5 są koszty realizacji zadań przez poszczególnych wykonawców, wyjściem y_5 jest całkowity koszt wykonania zadania, czyli $y_5 = F_5(u_5) = u_5^{(1)} + u_5^{(2)} + u_5^{(3)}$, a funkcja F_5 jest charakterystyką statyczną tego elementu.



Rys. 1.15. Statyczny system złożony – realizacja podzielonego zadania przez trzech wykonawców

Zauważmy, że wyróżnione elementy stanowią odrębne obiekty wejściowo-wyjściowe, w których wyróżniono wielkości wejściowe i wyjściowe oraz wskazano na zależność statyczną pomiędzy tymi wielkościami. Wyróżnione są elementy powiązane między sobą, a konkretnie wyjścia jednych elementów są wejściami innych. Powiązania te określają strukturę systemu. W rozpatrywanym przykładzie przydział zadań do realizacji dla poszczególnych wykonawców – O_2 , O_3 oraz O_4 – jest wynikiem rozdziału zadań O_1 , tj. kolejne wyjścia $y_1^{(1)}$, $y_1^{(2)}$ oraz $y_1^{(3)}$ są wejściami do O_2 , O_3 oraz O_4 . Powiązania pomiędzy tymi elementami są dane zależnościami: $u_1 = y_1^{(1)}$, $u_2 = y_1^{(2)}$ oraz $u_3 = y_1^{(3)}$. Podstawą do oceny kosztów O_5 są koszty realizacji zadań przez wykonawców, a zatem składowymi wektora u_5 są wyjścia $y_2^{(1)}$, $y_3^{(1)}$ oraz $y_3^{(1)}$, czyli $u_5^{(1)} = y_2^{(1)}$, $u_5^{(2)} = y_3^{(1)}$ oraz $u_5^{(3)} = y_4^{(1)}$.

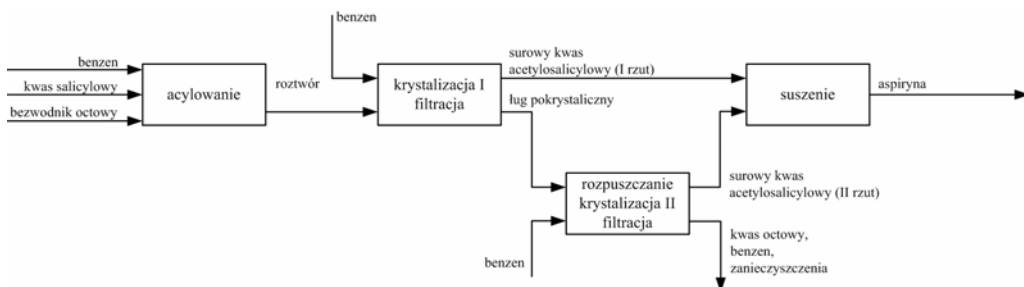
Zwróćmy uwagę na całkowity rozmiar zadania: wejście u_1 nie jest wyjściem żadnego elementu – jest ono wejściem zewnętrznym $x = u_1$.

Ze względu na ocenę rozpatrywanego procesu wytwarzania istotne są czasy wykonania poszczególnych zadań oraz koszt całkowity wykonania zadania. Wskazuje to na wyróżnione wyjścia spośród wszystkich wyjść sytemu złożonego. W naszym przykła-

dzie jest to wektor v , o składowych $v^{(1)} = y_1^{(1)}$, $v^{(2)} = y_1^{(2)}$, $v^{(3)} = y_1^{(3)}$, które są czasami wykonania poszczególnych zadań, oraz $v^{(4)} = y_5$ – całkowity koszt zadania.

Po uwzględnieniu wejść zewnętrznych x oraz wyróżnionych wyjść v system złożony może być potraktowany jako nowy obiekt o wejściach x i wyjściach v , którego charakterystykę statyczną $v = F(x)$ wyznaczamy na podstawie charakterystyk poszczególnych elementów oraz struktury systemu złożonego. ◆

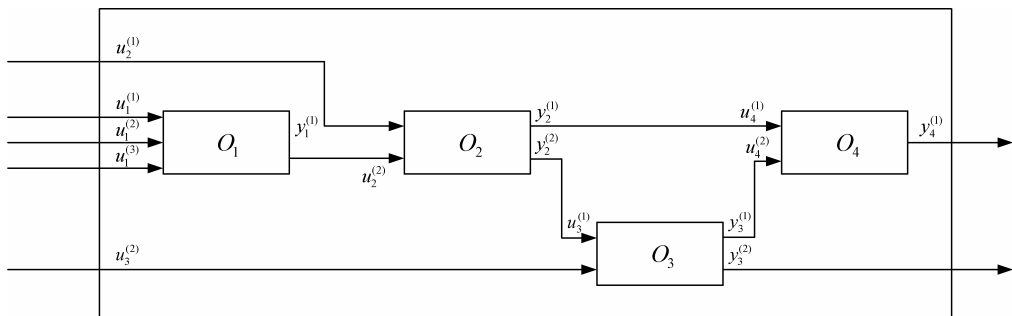
Jako kolejny przykład wejściowo-wyjściowego, statycznego systemu złożonego (rys. 1.16) rozważymy proces produkcji aspiryny [58, 62, 67, 137]. Aspirynę otrzymuje się z kwasu salicylowego przez acetylowanie bezwodnikiem octowym w roztworze benzenowym. Proces ten jest prowadzony w reaktorze przez ogrzewanie – w ciągu określonego czasu – mieszaniny bezwodnego benzenu, bezwodnika octowego oraz kwasu salicylowego. Po ukończeniu reakcji gorący roztwór przetłacza się przez filtr ciśnieniowy do krystalizatora I. Podczas chłodzenia i mieszania zachodzi krystalizacja aspiryny. Krystaliczny produkt jest filtrowany na nuczyci próżniowej i przemywany zimnym benzenem. W wyniku tego procesu otrzymujemy pierwszy rzut surowego kwasu acetylosalicylowego. Produktem ubocznym jest ług pokrystaliczny, zawierający znaczne ilości aspiryny. Ług pokrystaliczny jest rozpuszczany i kierowany do aparatu destylacyjnego, gdzie oddestylowuje się benzen, a następnie kwas octowy. Pozostałość, podobnie jak poprzednio, oczyszcza się przez krystalizację II i filtrację, w wyniku czego otrzymujemy drugi rzut surowego kwasu acetylosalicylowego. Zarówno pierwszy, jak i drugi rzut surowego kwasu acetylosalicylowego jest następnie suszony w suszarce próżniowej i w wyniku tego procesu otrzymujemy aspirynę.



Rys. 1.16. Schemat ideowy produkcji aspiryny

W omawianym procesie wyróżnimy następujące elementarne procesy jednostkowe (rys. 1.17): acylowanie – O_1 , krystalizacja I i filtracja – O_2 , rozpuszczanie, destylacja, krystalizacja II i filtracja – O_3 oraz suszenie – O_4 . Wyróżnione procesy stanowią odrębne obiekty wejściowo-wyjściowe, a ich połączenie stanowi złożony proces produkcji. Dla każdego procesu jednostkowego wskażemy jego wielkości wejściowe i wyjściowe. Acylowanie O_1 jest elementem, w którym składowymi wektora wejść u_1 są: ilość ben-

zenu $u_1^{(1)}$, ilość kwasu salicylowego $u_1^{(2)}$ oraz ilość bezwodnika octowego $u_1^{(3)}$, a wyjściem – ilość roztworu powstałego wskutek acylowania y_1 . Wektor u_2 , o składowych $u_2^{(1)}$ – ilość benzenu oraz $u_2^{(2)}$ – ilość roztworu przekazanego do pierwszej krystalizacji i filtracji, jest wektorem wejść elementu O_2 – krystalizacja I i filtracja, a wektor wyjść y_2 ma składowe: $y_2^{(1)}$ – ilość surowego kwasu acetylosalicylowego (pierwszy rzut) oraz $y_2^{(2)}$ – ilość ługu pokrystalicznego. W procesie rozpuszczanie, destylacja, krystalizacja II i filtracja O_3 wektor wejść u_3 ma składowe: $u_3^{(1)}$ – ilość ługu pokrystalicznego oraz $u_3^{(2)}$ – ilość benzenu, a wektor wyjść y_3 : $y_3^{(1)}$ – ilość surowego kwasu acetylosalicylowego (drugi rzut) oraz $y_3^{(2)}$ – ilość produktów odpadowych, tj.: kwasu octowego, benzenu i zanieczyszczeń. Składowymi wektora wejść u_4 w procesie suszenia O_4 są: ilość surowego kwasu acetylosalicylowego (pierwszy rzut) $u_4^{(1)}$ oraz ilość surowego kwasu acetylosalicylowego (drugi rzut) $u_4^{(2)}$, a ilość uzyskanej aspiryny y_4 jest wyjściem tego procesu. Zależności pomiędzy ilością produktów wejściowych a ilością produktów wyjściowych, uzyskanych w wyniku każdego z procesów elementarnych, są charakterystykami statycznymi odpowiednich elementów.



Rys. 1.17. Przykład systemu złożonego – produkcja aspiryny

Powiązania pomiędzy poszczególnymi elementarnymi procesami wynikają z faktu, że produkty uzyskane w jednych procesach są wielkościami wejściowymi kolejnych procesów. W rozpatrywanym przykładzie mamy: ilość roztworu powstałego wskutek acylowania jest przekazywana do pierwszej krystalizacji i filtracji, czyli $u_2^{(2)} = y_1^{(1)}$, ilość ługu pokrystalicznego powstałego w procesie pierwszej krystalizacji i filtracji jest przekazywana do procesu – rozpuszczanie, destylacja, krystalizacja II i filtracja, czyli $u_3^{(1)} = y_2^{(2)}$, ilość surowego kwasu acetylosalicylowego uzyskanego w pierwszym i drugim rzucie jest przekazywana do procesu suszenia, czyli $u_4^{(1)} = y_2^{(1)}$ oraz $u_4^{(2)} = y_3^{(1)}$. Zależności te określają strukturę powiązań pomiędzy elementami systemu.

Ilość kwasu salicylowego $u_1^{(2)}$, ilość bezwodnika octowego $u_1^{(3)}$ oraz ilości benzenu $u_1^{(1)}$, $u_2^{(1)}$, $u_3^{(2)}$ – potrzebne do przeprowadzenia procesów O_1 , O_2 , O_3 – są produktami podstawowymi do produkcji aspiryny i są wejściami zewnętrznymi systemu złożonego. Wektor wejść zewnętrznych x ma składowe: $x^{(1)} = u_2^{(1)}$, $x^{(2)} = u_1^{(1)}$, $x^{(3)} = u_1^{(2)}$, $x^{(4)} = u_1^{(3)}$ oraz $x^{(5)} = u_3^{(2)}$. Istotnymi, wyróżnionymi wyjściami omawianego procesu są: ilość wyprodukowanej aspiryny y_4 oraz ilość produktów odpadowych, tj.: kwasu octowego, benzenu i zanieczyszczeń $y_3^{(2)}$, zatem $v^{(1)} = y_4$ oraz $v^{(2)} = y_3^{(2)}$ są składowymi wektora wyróżnionych wyjść.

Biorąc pod uwagę wektor wejść zewnętrznych x oraz wektor wyróżnionych wyjść v , system złożony możemy potraktować jako nowy obiekt o wektorze wejść x , i wektorze wyjść v , którego charakterystykę statyczną $v = F(x)$ wyznaczamy na podstawie charakterystyk poszczególnych elementów oraz struktury systemu złożonego. ◆

Mówiąc o **wejściowo-wyjściowym statycznym systemie złożonym**, mamy na myśli obiekt, w którym wyróżniamy powiązane między sobą elementy składowe. W opisie takiego systemu należy podać charakterystyki statyczne poszczególnych elementów systemu oraz wskazać powiązania pomiędzy wejściami i wyjściami elementów.

Rozważymy wejściowo-wyjściowy system złożony, w którym wyróżniono M elementów O_1, O_2, \dots, O_M . Niech

$$y_m = F_m(u_m) \quad (1.13)$$

oznacza charakterystykę statyczną m -tego elementu o wejściach u_m i wyjściach y_m , a F_m jest znaną funkcją.

Wejście i wyjście m -tego elementu są wektorami z odpowiednich przestrzeni, tj.:

$$u_m = \begin{bmatrix} u_m^{(1)} \\ u_m^{(2)} \\ \vdots \\ u_m^{(S_m)} \end{bmatrix} \in \mathcal{U}_m \subseteq \mathcal{R}^{S_m}, \quad y_m = \begin{bmatrix} y_m^{(1)} \\ y_m^{(2)} \\ \vdots \\ y_m^{(L_m)} \end{bmatrix} \in \mathcal{Y}_m \subseteq \mathcal{R}^{L_m}, \quad (1.14)$$

gdzie: S_m oraz L_m są, odpowiednio, wymiarami przestrzeni wejść i wyjść, $m = 1, 2, \dots, M$.

Oznaczmy przez u i y , odpowiednio, wektory wszystkich wejść i wyjść obiektu, tj.

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}, \quad (1.15)$$

gdzie wektor wszystkich wejść u należy do przestrzeni

$$\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \dots \times \mathcal{U}_M \subseteq \mathcal{R}^S, \quad S = \sum_{m=1}^M S_m, \quad (1.16)$$

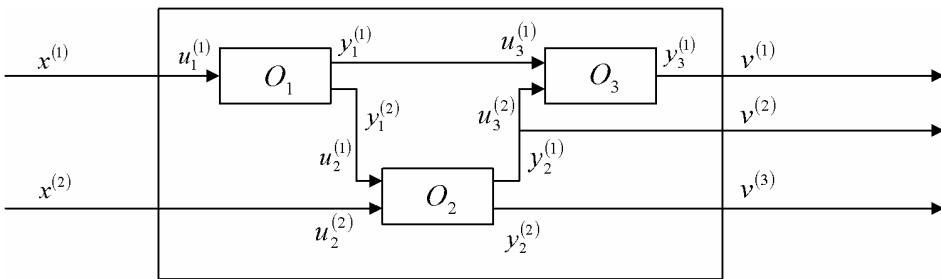
a wektor wszystkich wyjść y należy do przestrzeni

$$\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2 \times \dots \times \mathcal{Y}_M \subseteq \mathcal{R}^L, \quad L = \sum_{m=1}^M L_m. \quad (1.17)$$

Struktura systemu jest zadana przez podanie powiązań pomiędzy wyjściami i wejściami poszczególnych elementów systemu. W szczególności należy podać, które wejście dla wskazanego elementu jest wyjściem innego elementu. Zwróćmy uwagę, że w systemie pojawiają się wejścia, które nie są wyjściami innych elementów (rys. 1.18). Takie wejścia x nazwiemy wejściami zewnętrznymi systemu złożonego, gdzie

$$x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(\tilde{S})} \end{bmatrix} \quad (1.18)$$

jest \tilde{S} -wymiarowym wektorem wejść $x \in \mathcal{X} \subseteq \mathcal{U}$ z przestrzeni $\mathcal{X} \subseteq \mathcal{R}^{\tilde{S}}$.



Rys. 1.18. Przykład systemu złożonego

Strukturę systemu złożonego opisujemy zależnością

$$u = Ay + Bx, \quad (1.19)$$

w której: A jest $S \times L$ -wymiarową zero-jedynkową macierzą, definiującą powiązania pomiędzy elementami systemu, tj.:

$$A = [a_{sl}]_{\substack{s=1,2,\dots,S \\ l=1,2,\dots,L}}, \quad a_{sl} = \begin{cases} 1 & \text{gdy } u^{(s)} = y^{(l)} \\ 0 & \text{gdy } u^{(s)} \neq y^{(l)} \end{cases}, \quad (1.20)$$

natomiast zero-jedynkowa $S \times \tilde{S}$ -wymiarowa macierz B wskazuje wejścia zewnętrzne, czyli:

$$B = [b_{s\tilde{s}}]_{\substack{s=1,2,\dots,S \\ \tilde{s}=1,2,\dots,\tilde{S}}}, \quad a_{sl} = \begin{cases} 1 & \text{gdy } u^{(s)} = x^{(\tilde{s})} \\ 0 & \text{gdy } u^{(s)} \neq x^{(\tilde{s})} \end{cases} \quad (1.21)$$

Wyróżniamy ponadto pewne wyjścia, które mają znaczenie dla wykorzystania opracowanego modelu, np. wyjścia te mogą być istotne dla zadania sterowania lub zarządzania. Wyjścia te nazwiemy wyjściami wyróżnionymi przez projektanta modelu. W jednym przypadku mogą to być wyjścia, które są dostępne pomiarowo, a w innym – wyjścia systemu złożonego istotne do wykorzystania opracowanego modelu. Wyróżnienie tych dwóch przypadków jest istotne dla dalszych rozważań. Pierwszy przypadek – gdy tylko wyróżnione wyjścia systemu złożonego mogą być mierzone oraz znane są charakterystyki statyczne elementów systemu z dokładnością do wektora parametrów, tj. charakterystyki statyczne należą do klasy modeli – w dalszych rozważaniach prowadzi do sformułowania zadania identyfikacji systemu złożonego przy ograniczonych możliwościach pomiarowych. Drugi przypadek – gdy projektant modelu wyróżnił wybrane wyjścia, mimo że wszystkie mogą być mierzone i nie są znane charakterystyki statyczne elementów – w dalszych rozważaniach prowadzi do sformułowania zadania wyboru optymalnego modelu z lokalnym i globalnym wskaźnikiem jakości identyfikacji. Zadania te będą precyzyjniej sformułowane w dalszej części pracy.

Wyróżnione wyjścia oznaczymy przez

$$v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(\tilde{L})} \end{bmatrix}, \quad (1.22)$$

gdzie \tilde{L} jest liczbą wybranych wyjść spośród wszystkich L wyjść.

Wyjścia v są wskazane przez macierz zero-jedynkową C o wymiarze $\tilde{L} \times L$, tj.

$$v = Cy, \quad (1.23)$$

w której:

$$C = [c_{\tilde{l}l}]_{\substack{\tilde{l}=1,2,\dots,\tilde{L} \\ l=1,2,\dots,L}}, \quad \text{gdzie } c_{\tilde{l}l} = \begin{cases} 1 & \text{gdy } v^{(\tilde{l})} = y^{(l)} \\ 0 & \text{gdy } v^{(\tilde{l})} \neq y^{(l)} \end{cases} \quad (1.24)$$

oraz

$$v \in \mathcal{V} = \{v : \forall y \in \mathcal{Y}, v = Cy\} \subseteq \mathcal{R}^{\tilde{L}}. \quad (1.25)$$

Na rysunku 1.18 przedstawiono przykład systemu złożonego, w którym wyróżniono elementy składowe w postaci obiektów O_1 , O_2 i O_3 . Wskazano wejścia zewnętrzne x , wyróżnione wyjścia v oraz wewnętrzne powiązania.

Równania (1.19) i (1.23), dla systemu złożonego przedstawionego na rysunku 1.18, mają postać:

$$\begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \\ u_2^{(2)} \\ u_3^{(1)} \\ u_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1^{(1)} \\ y_1^{(2)} \\ y_2^{(1)} \\ y_2^{(2)} \\ y_3^{(1)} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix}, \quad (1.26)$$

$$\begin{bmatrix} v^{(1)} \\ v^{(2)} \\ v^{(3)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1^{(1)} \\ y_1^{(2)} \\ y_2^{(1)} \\ y_2^{(2)} \\ y_3^{(1)} \end{bmatrix}. \quad (1.27)$$

Zwróćmy uwagę, że na system złożony możemy spojrzeć jak na nowy obiekt identyfikacji z wektorem wejść zewnętrznych x oraz wektorem wyróżnionych wyjść v wskazanych przez (1.23). Korzystając z charakterystyk statycznych poszczególnych elementów (1.13), równania (1.19), opisującego strukturę systemu złożonego, oraz zależności (1.23), wskazującej wyróżnione wyjścia, wyznaczamy charakterystykę statyczną nowo określonego obiektu identyfikacji, czyli systemu złożonego jako całości.

Przyjmujemy oznaczenie

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} F_1(u_1) \\ F_2(u_2) \\ \vdots \\ F_M(u_M) \end{bmatrix} \stackrel{\text{df}}{=} \bar{F}(u), \quad (1.28)$$

gdzie \bar{F} jest funkcją określającą zależność pomiędzy wszystkimi wejściami i wyjściami systemu złożonego.

Po podstawieniu w miejsce u w równaniu (1.28) zależności (1.19) otrzymujemy

$$y = \bar{F}(Ay + Bx). \quad (1.29)$$

Rozwikłanie równania (1.29) względem y daje

$$y = \bar{F}_y(x; A, B), \quad (1.30)$$

gdzie \bar{F}_y oznacza rozwiązanie równania (1.29) względem y .

W dalszych rozważaniach ograniczymy się do przypadku, gdy takie rozwiązanie istnieje.

Po podstawieniu wzoru (1.30) do zależności (1.23) otrzymujemy

$$v = C\bar{F}_y(x; A, B) \stackrel{\text{df}}{=} F(x). \quad (1.31)$$

Zależność (1.31) jest charakterystyką statyczną obiektu z wektorem wejść zewnętrznych x i wektorem wyróżnionych wyjść v . Dla tak opisanego systemu złożonego są formułowane nowe zadania i problemy identyfikacji.

Łatwo zauważyć, że dla systemu złożonego o strukturze szeregowej, z wektorem wejść zewnętrznych x , będącym wektorem wejść pierwszego elementu, tj. $x = u_1$, i wektorem wyróżnionych wyjść v , którym jest wektor wyjść ostatniego M -tego elementu, tj. $v = y_M$, charakterystyka statyczna systemu złożonego (1.31) przyjmie postać

$$v = F_M(F_{M-1}(\dots F_1(x))) = F(x). \quad (1.32)$$

1.5. Problemy identyfikacji systemów złożonych

W problematyce identyfikacji systemów złożonych można wyróżnić następujące, mniej lub bardziej rozwinięte, koncepcje:

1. Identyfikacja przy ograniczonych możliwościach pomiarowych.
2. Identyfikacja globalna.
3. Identyfikacja wielostopniowa.

Identyfikacja przy ograniczonych możliwościach pomiarowych

Identyfikacja wejściowo-wyjściowego systemu złożonego przy ograniczonych możliwościach pomiarowych dotyczy obiektu, w którym wyróżniono elementy składowe. Zakłada się znajomość struktury systemu, tj. znane są powiązania pomiędzy poszczególnymi elementami. Opisy poszczególnych elementów są znane z dokładnością do parametrów. Odpowiada to problemowi wyznaczenia parametrów charakterystyki obiektu lub estymacji parametrów w przypadku losowym. Dana jest konfiguracja punktów pomiarowych, czyli zestaw tych wyjść, które są dostępne do pomiaru przy zadanych wartościach wejść zewnętrznych. Pojawia się pytanie: Czy na podstawie zadanej konfiguracji pomiarów możliwe jest jednoznaczne wyznaczenie parametrów opisów poszczególnych elementów? Odpowiedź na to pytanie prowadzi do wprowadzenia pojęcia separowalności systemu złożonego, analogicznego do pojęcia identyfikowalności dla zadania identyfikacji obiektu. Okazuje się, jak można pokazać nawet dla prostych przypadków deterministycznych, że problem ten zależy zarówno od struktury systemu złożonego, jak i od charakterystyk statycznych jego elementów [25–27, 92, 95]. Ten problem omówiono w rozdziale 3., w którym formułuje się zadanie identyfikacji z ograniczonymi możliwościami pomiarowymi. Dla różnych przypadków podano warunki separowalności. Wskazano też na możliwość uzyskania separowalności dzięki dodatkowej informacji a priori o obiekcie.

Identyfikacja globalna

Koncepcja identyfikacji globalnej systemu złożonego zakłada znajomość struktury systemu, określonej powiązaniem pomiędzy poszczególnymi elementami. W odróżnieniu od poprzedniej sytuacji, zakłada się możliwość pomiaru wszystkich wejść i wyjść poszczególnych elementów systemu, czyli dla zadanych wartości wejść zewnętrznych

wyjścia wszystkich elementów są dostępne do pomiaru. Dla każdego elementu przyjmujemy model parametryczny. W zadaniu wyboru optymalnego modelu systemu złożonego można zastosować różne koncepcje oceny jakości modelu, co prowadzi do różnych kryteriów jakości identyfikacji. Z jednej strony, niezależnie od powiązań, dla każdego elementu, traktowanego jako obiekt wyodrębniony z systemu, można sformułować osobną – lokalną ocenę jakości modelu. W ten sposób określimy lokalny wskaźnik jakości identyfikacji. Rozwiązując dla każdego elementu niezależne zadanie wyboru optymalnego modelu, otrzymamy lokalnie optymalne modele poszczególnych elementów. Korzystając ze struktury systemu złożonego, wyróżnionych wyjść oraz uzyskanych lokalnie optymalnych modeli, wyznaczamy model systemu złożonego z wejściami zewnętrznymi i wyróżnionymi wyjściami. Taki model nazwiemy lokalnie optymalnym. Z drugiej strony, biorąc pod uwagę zamierzone wykorzystanie modelu systemu złożonego, możemy być zainteresowani wyróżnionymi wyjściami systemu złożonego. Wynika stąd celowość stosowania innej niż poprzednio oceny jakości modelu. Teraz, korzystając ze struktury systemu złożonego, wyróżnionych wyjść oraz przyjętych postaci modeli, wyznaczamy model systemu złożonego z wejściami zewnętrznymi i wyróżnionymi wyjściami. Można już zatem sformułować ocenę jakości modelu jako całego systemu. Ocenę tę nazwiemy globalną. Globalny wskaźnik jakości identyfikacji ocenia różnicę pomiędzy wyróżnionymi wyjściami systemu złożonego a odpowiednimi wyjściami modelu. W wyniku rozwiązania zadania wyboru optymalnego modelu z globalnym wskaźnikiem jakości uzyskamy model, który nazwiemy modelem globalnie optymalnym. Koncepcja identyfikacji globalnej została sformułowana w pracy [14], a następnie rozwinięta w pracach [15, 17, 19, 28, 94].

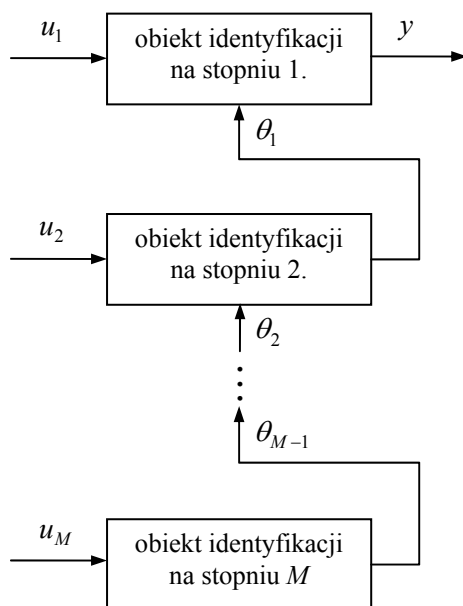
Problematykę identyfikacji lokalnej i globalnej statycznego systemu złożonego omówiono w rozdziale 4. Przedstawiono tam ogólne sformułowanie zadania identyfikacji globalnej. Dla struktury szeregowej zaproponowano algorytm identyfikacji oparty na metodzie programowania dynamicznego. Korzystając z podejścia wielokryterialnego, zaproponowano różne możliwości wyznaczenia parametrów globalnie optymalnych, z uwzględnieniem jakości modeli lokalnych. Przedstawiono też możliwość zastosowania sieci neuronowych w modelowaniu statycznych systemów złożonych.

Identyfikacja wielostopniowa

Istota koncepcji identyfikacji wielostopniowej dla dwóch stopni jest następująca: Dla pewnego obiektu (rys. 1.19) – procesu, zjawiska, urządzenia – badamy zależność wielkości wyjściowej y od wielkości wejściowej u_1 i wyznaczamy parametr w opisie tej zależności θ_1 . Jest to zadanie identyfikacji na pierwszym stopniu. Następnie zwracamy uwagę na pewną wielkość u_2 , która w czasie identyfikacji na pierwszym stopniu była stała. Zmieniamy wartość tej wielkości i powtarzamy identyfikację na pierwszym stopniu, uzyskując na ogół inną wartość parametru opisu θ_1 .

Powtarzając wielokrotnie identyfikację na pierwszym stopniu dla różnych wartości u_2 , uzyskamy dane do zbadania zależności pomiędzy parametrem opisu na pierwszym stopniu θ_1 a wielkością wejściową u_2 . Teraz możemy zbadać zależność wartości parametru θ_1

od wartości wielkości wejściowej u_2 i wyznaczyć parametr tej zależności θ_2 . Jest to zadanie identyfikacji na drugim stopniu, gdzie θ_1 jest wyjściem, a u_2 – wejściem obiektu.



Rys. 1.19. Wielostopniowy obiekt identyfikacji

Ideę tę można łatwo uogólnić iteracyjnie na wiele stopni, zwracając uwagę na kolejne wielkości i uwzględnienie ich w opisie obiektu przez zmianę parametrów stopnia poprzedniego (rys. 1.19).

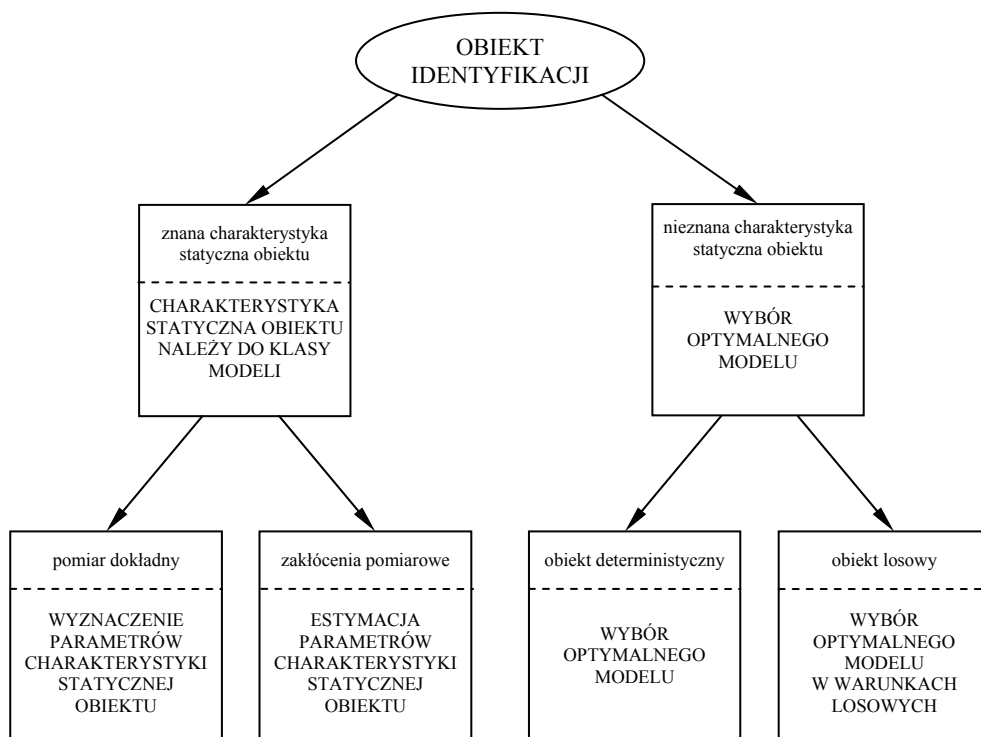
Wracając do identyfikacji dwustopniowej, zauważmy, że oba stopnie mogą dotyczyć tego samego obiektu, a zadanie identyfikacji dwustopniowej można potraktować jako swego rodzaju dekompozycję zadania identyfikacji z wyjściem y oraz wejściami u_1 i u_2 . Dekompozycja taka może być podyktowana sposobem prowadzenia eksperymentu, a także względami obliczeniowymi. Może też być naturalną konsekwencją zadania identyfikacji, zwłaszcza wtedy, gdy pierwszy stopień dotyczy badania obiektów tego samego typu, różniących się pewną wielkością u_2 , a drugi stopień – uogólnienia badań na zbiór takich obiektów.

Koncepcja identyfikacji wielostopniowej została sformułowana w pracach [16, 64], a następnie rozwinięta dla identyfikacji dwustopniowej w monografii [109].

W rozdziale 5. ograniczono się do przedstawienia zadania identyfikacji dwustopniowej, wskazując na możliwość uogólnienia na zadanie identyfikacji wielostopniowej. Omówiono tam dwustopniowe algorytmy estymacji parametrów obiektu oraz algorytmy wyboru optymalnego modelu w warunkach losowych.

2. Podstawowe zadania identyfikacji obiektów statycznych

Obecnie skupimy się na podstawowym zadaniu identyfikacji. Przytoczone tutaj rozważania ograniczają się do zadania identyfikacji obiektów statycznych. Mają one na celu omówienie zadań identyfikacji, do których będziemy się odwoływać w zadaniach identyfikacji systemów złożonych. Omówimy zadania wyznaczania algorytmów identyfikacji dla typowych przypadków.



Rys. 2.1. Typowe zadania identyfikacji

Ze względu na informacje aprioryczne o obiekcie możemy wyróżnić dwa przypadki:

1. Znamy opis obiektu z dokładnością do parametrów. Znamy zatem funkcyjną postać zależności pomiędzy wejściami i wyjściami obiektu, ale nie znamy parametrów tej zależności. Mówimy wówczas, że charakterystyka statyczna obiektu należy do klasy modeli, a zadanie identyfikacji sprowadza się do wyznaczenia nieznanymi parametrów opisu.
2. Postać funkcyjna nie jest znana. Twórca modelu przyjmuje postać tej funkcji, proponując przybliżanie zależności pomiędzy wejściami i wyjściami obiektu. W tym przypadku mówimy o zadaniu wyboru najlepszego (optymalnego) modelu. Kolejny podział dotyczy informacji o zakłóceniach pomiarowych i wielkościach losowych w obiekcie identyfikacji. Prowadzi to do wyodrębnienia kolejnych zadań:
 - Charakterystyka obiektu jest znana z dokładnością do parametrów, a na obserwowane wielkości nakładają się zakłócenia lub też w obiekcie pojawiają się losowo zmienne wielkości niemierzalne, mówimy wtedy o zadaniu estymacji parametrów.
 - Obserwacje obiektu wskazują na jego losowe zachowanie się, a charakterystyka obiektu nie jest znana, wtedy mówimy o zadaniu wyboru optymalnego modelu w warunkach losowych. W konkretnych przypadkach prowadzi to do regresji pierwszego i drugiego rodzaju.

2.1. Wyznaczanie parametrów charakterystyki statycznej obiektu

Założymy, że znamy charakterystykę statyczną obiektu, lecz nie znamy jej parametrów. Przyjęcie takiego założenia jest możliwe wówczas, gdy badany proces jest dobrze opisany i na podstawie analizy zjawisk można ustalić dokładną funkcyjną postać charakterystyki statycznej, w której pewne parametry nie są znane (rys. 2.2). Inaczej mówiąc – charakterystyka statyczna ma postać

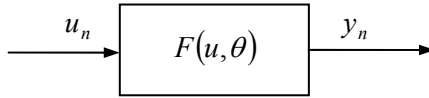
$$y = F(u, \theta), \quad (2.1)$$

gdzie: F jest znaną funkcją opisującą zależność wektora wyjść y od wektora wejść u , natomiast θ jest wektorem parametrów tej zależności.

Składowe wektora parametrów θ mają praktyczne znaczenie i odnoszą się do badanej rzeczywistości. W przypadku konkretnych obiektów parametry te mogą oznaczać wartości stałych, jak na przykład współczynnik tarcia w układzie mechanicznym lub stopa procentowa w opisie procesu ekonomicznego. W dalszych rozważaniach będziemy zakładać, że wejścia i wyjścia oraz parametry obiektu identyfikacji są – odpowiednio – S -, L - oraz R -wymiarowymi wektorami:

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix},$$

gdzie: u – wektor wejść obiektu, $u \in U \subseteq \mathbb{R}^S$, \mathcal{U} – przestrzeń wejść, y – wektor wyjść obiektu, $y \in Y \subseteq \mathbb{R}^L$, \mathcal{Y} – przestrzeń wyjść, θ – wektor parametrów charakterystyki statycznej obiektu, $\theta \in \Theta \subseteq \mathbb{R}^R$, Θ – przestrzeń parametrów.



Rys. 2.2. Obiekt identyfikacji

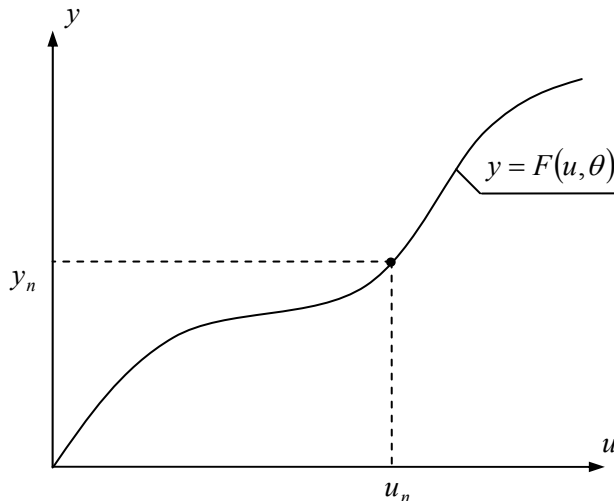
Podstawą do wyznaczenia parametrów charakterystyki statycznej są wyniki eksperymentu. W tym przypadku założymy, że pomiary nie są obarczone błędem. W n -tym pomiarze dla zadanego wektora wejść u_n mierzymy wektor wyjść y_n , $n = 1, 2, \dots, N$, gdzie N jest długością serii pomiarowej.

Wyniki pomiarów możemy zebrać w postaci macierzy:

$$U_N = [u_1 \quad u_2 \quad \dots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \dots \quad y_N], \quad (2.2)$$

gdzie: U_N, Y_N są macierzami, w których kolumnami są kolejne wyniki pomiarów wartości składowych wektora wejść u i wektora wyjść y .

Ponieważ charakterystyka statyczna obiektu należy do klasy modeli i brak jest błędu pomiarowego, uzyskane w wyniku eksperymentu pomiary są wybranymi punktami charakterystyki statycznej obiektu (rys. 2.3).

Rys. 2.3. Obserwacja obiektu o znanej z dokładnością do wektora parametrów θ charakterystyce statycznej jednowymiarowego obiektu ($L = S = 1$)

Znaczy to, że przy znanych wartościach składowych wektora parametrów θ , dla zadanych wartości składowych wektora wejść u_n , na podstawie zależności (2.1) jest

możliwe wyznaczenie wartości składowych wektora wyjść y_n . Wartość ta pokrywa się z wynikiem niezakłóconego pomiaru wartości składowych wektora wyjść, czyli jest spełniony następujący układ równań

$$y_n = F(u_n, \theta), \quad n = 1, 2, \dots, N. \quad (2.3)$$

To proste i oczywiste spostrzeżenie jest podstawą konstrukcji algorytmu identyfikacji. Zgodnie z (2.2) układ równań (2.3) możemy zapisać

$$[y_1 \quad y_2 \quad \dots \quad y_N] = [F(u_1, \theta) \quad F(u_2, \theta) \quad \dots \quad F(u_N, \theta)]. \quad (2.4)$$

Po przyjęciu oznaczeń

$$[F(u_1, \theta) \quad F(u_2, \theta) \quad \dots \quad F(u_N, \theta)] \stackrel{\text{df}}{=} \bar{F}(U_N, \theta), \quad (2.5)$$

układ równań (2.3) ma zwartą postać

$$Y_N = \bar{F}(U_N, \theta). \quad (2.6)$$

Rozwiązanie układu (2.6) względem wektora parametrów θ daje algorytm identyfikacji, tj.

$$\theta = \bar{F}_\theta^{-1}(U_N, Y_N) \stackrel{\text{df}}{=} \Psi_N(U_N, Y_N), \quad (2.7)$$

gdzie: \bar{F}_θ^{-1} oznacza funkcję odwrotną względem θ , czyli rozwiązanie równania (2.6) względem θ , natomiast Ψ_N jest algorytmem identyfikacji.

Należy zwrócić uwagę, że układ równań (2.3), zapisany w zwartej postaci (2.6), nie zawsze musi mieć rozwiązanie. Zależy to od organizacji eksperymentu, jak też od własności badanego obiektu. W tym momencie nasuwa się spostrzeżenie, że rozwiązanie układu równań (2.3) zależy od serii identyfikującej U_N . Jej długość, czyli liczba pomiarów N , musi zapewnić odpowiednią liczbę równań w układzie (2.3), która pozwala wyznaczyć wektor nieznanych parametrów. Seria powinna spełniać oczywisty warunek

$$NL \geq R, \quad (2.8)$$

który gwarantuje, że liczba równań będzie nie mniejsza niż liczba nieznanych parametrów. Serię identyfikującą U_N należy zaś dobrać tak, aby zależność (2.3) tworzyła układ równań niezależnych. Nie zawsze taka seria istnieje. Zależy to od własności obiektu. Z tym związane jest pojęcie identyfikowalności [14].

Definicja 2.1. Obiekt statyczny nazywamy identyfikowalnym, jeżeli istnieje taka seria identyfikująca $U_N = [u_1 \quad u_2 \quad \dots \quad u_N]$, która – wraz z macierzą pomiarów wyjść $Y_N = [y_1 \quad y_2 \quad \dots \quad y_N]$ – jednoznacznie określa parametry charakterystyki statycznej.

Innymi słowy – obiekt jest identyfikowalny, jeżeli istnieje taka seria identyfikująca U_N , dla której układ równań (2.3), zapisany w zwartej postaci (2.6), ma jednoznaczne rozwiązanie względem θ .

Przykład 2.1. Obiekt o jednym wyjściu ($L = 1$) jest opisany charakterystyką liniową względem parametrów, czyli opis (2.1) ma postać

$$y = F(u, \theta) = \theta^T f(u), \quad (2.9)$$

gdzie $f(u)$ jest R -wymiarowym wektorem znanych funkcji

$$f(u) \stackrel{\text{df}}{=} \begin{bmatrix} f^{(1)}(u) \\ f^{(2)}(u) \\ \vdots \\ f^{(R)}(u) \end{bmatrix}. \quad (2.10)$$

Jak łatwo zauważyć, z warunku (2.8) wynika, że konieczna jest liczba pomiarów $N = R$. W tym przypadku układ równań (2.3) ma postać

$$y_n = \theta^T f(u_n), \quad n = 1, 2, \dots, R, \quad (2.11)$$

a w zapisie zwartym (2.4)

$$Y_R \stackrel{\text{df}}{=} [y_1 \quad y_2 \quad \dots \quad y_R] = [\theta^T f(u_1) \quad \theta^T f(u_2) \quad \dots \quad \theta^T f(u_R)] \quad (2.12)$$

lub (2.6)

$$Y_R^T = \bar{f}^T(U_R) \theta, \quad (2.13)$$

gdzie: $\bar{f}(U_R)$ jest macierzą kwadratową, której kolumnami są wartości funkcji (2.10) w kolejnych punktach serii identyfikującej U_R , tj.

$$\bar{f}(U_R) \stackrel{\text{df}}{=} [f(u_1) \quad f(u_2) \quad \dots \quad f(u_R)]. \quad (2.14)$$

Warunek identyfikowalności jest tutaj następujący:

$$\det[\bar{f}^T(U_R)] \neq 0, \quad (2.15)$$

co oznacza, że wartości funkcji (2.10) w kolejnych punktach pomiarowych muszą być liniowo niezależne.

Po pomnożeniu lewostronnie równania (2.13) przez macierz odwrotną do macierzy $\bar{f}^T(U_R)$ otrzymujemy algorytm identyfikacji

$$\theta = \Psi_R(U_R, Y_R) \stackrel{\text{df}}{=} [\bar{f}^T(U_R)]^{-1} Y_R^T. \quad (2.16)$$

W szczególnym przypadku, gdy $f(u) = u$, czyli dla klasy charakterystyk liniowych, układ równań (2.13) ma postać

$$Y_R^T = U_R^T \theta, \quad (2.17)$$

a jego rozwiązanie, czyli algorytm identyfikacji

$$\theta = \Psi_R(U_R, Y_R) \stackrel{\text{df}}{=} [U_R^T]^{-1} Y_R^T. \quad (2.18)$$

W tym przypadku warunek identyfikowalności ma postać

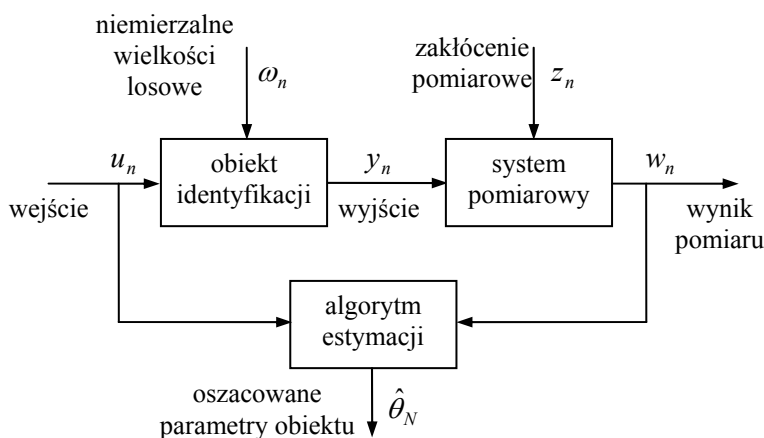
$$\det [U_R] \neq 0, \quad (2.19)$$

co oznacza, że pomiary w poszczególnych punktach pomiarowych $u_n, n=1, 2, \dots, N$, muszą być liniowo niezależne. ◆

W tym miejscu warto zwrócić uwagę na to, że wyznaczenie wektora parametrów według algorytmów (2.15) i (2.18) wymaga odwrócenia macierzy $\tilde{f}(U_R)$ oraz U_R . Nakłada to oczywisty warunek na dobór serii identyfikującej, tak aby wspomniane macierze były dobrze uwarunkowane.

2.2. Estymacja parametrów charakterystyki statycznej obiektu

W podrozdziale 2.1 przedstawiono przypadek wyznaczenia parametrów znanej charakterystyki statycznej obiektu na podstawie dokładnego pomiaru wielkości wejścia i wyjścia obiektu. Obecnie uwzględnimy fakt, że dla zadanego wejścia wynik pomiaru wyjścia obiektu jest obciążony zakłóceniami pomiarowymi (rys. 2.4). W obiekcie mogą ponadto występować niemierzalne wielkości losowe. Taka sytuacja prowadzi do nowych zadań.



Rys. 2.4. Zadanie estymacji parametrów obiektu

Ogólnie nazwiemy je zadaniami estymacji parametrów charakterystyki statycznej w warunkach losowych. Z zadaniem estymacji parametrów obiektu wiąże się,

ważne w dalszych rozważaniach, pojęcie estymowalności (odpowiadające identyfikowalności – definicja 2.1 – w przypadku deterministycznym), wprowadzone w publikacjach [25–27, 95].

Definicja 2.2. Obiekt nazywamy **estymowalnym** w sensie określonej własności, jeżeli dla każdego wektora parametrów obiektu θ istnieje taka seria identyfikująca U_N , dla której istnieje estymator $\hat{\underline{\theta}}_N = \Psi_N(U_N, \underline{W}_N)$ o określonej własności, w którym \underline{W}_N jest wielowymiarową zmienną losową, której realizacją jest macierz pomiarów wyjść W_N dla zadanej serii pomiarowej U_N .

Definicja 2.3. Obiekt nazywamy **estymowalnym** określoną metodą estymacji, jeżeli dla każdego wektora parametrów obiektu θ istnieje taka seria identyfikująca U_N , dla której wraz z macierzą pomiarów wyjść W_N istnieje oszacowanie $\hat{\theta}_N = \Psi_N(U_N, W_N)$ tą metodą.

Wyróżnimy tutaj następujące przypadki, dla których proponujemy algorytmy estymacji:

- Obiekt jest deterministyczny, lecz wynik pomiaru wartości wyjścia obiektu jest obciążony zakłóceniami pomiarowymi.
- Pomiary są dokładne, lecz w obiekcie występują niemierzalne wielkości losowe.
- W obiekcie występują niemierzalne wielkości losowe oraz wynik pomiaru wartości wyjścia obiektu jest obciążony zakłóceniami pomiarowymi.

2.2.1. Estymacja parametrów charakterystyki statycznej obiektu na podstawie zakłóconych pomiarów wyjścia

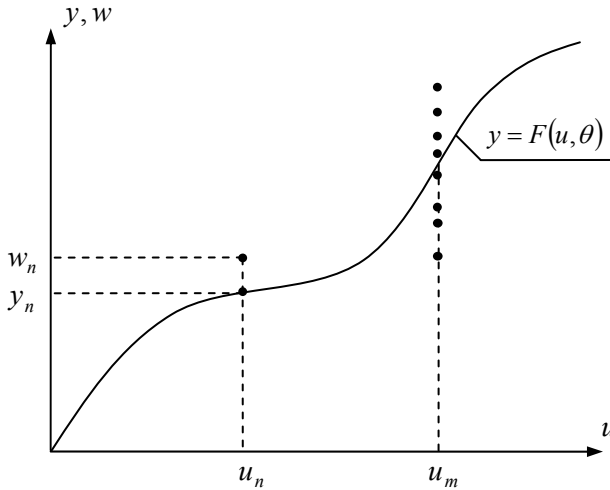
Obecnie rozważymy przypadek, gdy deterministyczny obiekt jest opisany charakterystyką statyczną, która jest znana z dokładnością do parametrów. Uwzględnimy fakt, że – dla danego wejścia – wyjście obiektu mierzymy z zakłóceniami pomiarowymi. W celu wyznaczenia nieznanymi parametrów charakterystyki obiektu wykonaliśmy eksperyment. W n -tym pomiarze, dla danego wektora wejść u_n , dokonaliśmy pomiaru wartości składowych wektora wyjść y_n . Na mierzoną wartość nałożyły się przypadkowe zakłócenia $z_n \in Z$ (Z jest przestrzenią zakłóceń) i w wyniku pomiaru otrzymujemy wartości składowych wektora w_n . W dalszych rozważaniach (dla uproszczenia zapisu) będziemy zakładać, że wymiar wektora y oraz zakłóceń z jest taki sam i równy L ($\dim y = \dim z = L$), czyli $z_n \in Z \subseteq \mathbb{R}^L$. Przypadek, gdy $\dim y > \dim z$, wymaga dodatkowych przekształceń (zob. Dodatek). Sposób nakładania się zakłóceń na mierzoną wartość może być różny – zależy on od systemu pomiarowego. Ze względu na losowy charakter błędu pomiarowego, powtarzając wielokrotnie pomiar wartości składowych wektora wyjść dla tego samego wektora wejść, otrzymamy różne wyniki pomiarów (rys. 2.5).

Opiszemy precyzyjnie aktualny układ identyfikacji (rys. 2.6). Charakterystyka statyczna obiektu jest znana z dokładnością do wektora parametrów θ i ma postać (2.1). Założymy ponadto, że znany jest opis systemu pomiarowego, a konkretnie znany jest

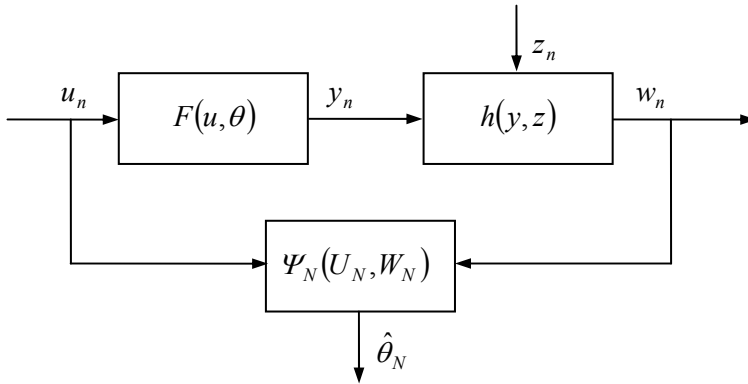
sposób nakładania się zakłóceń na mierzoną wartość składowych wektora wyjść obiektu i jest określony zależnością

$$w = h(y, z), \tag{2.20}$$

w której: h jest znaną funkcją, taką że $h: Y \times Z \rightarrow W$, $w \in W$, \mathcal{U} jest przestrzenią zakłóconego wyniku pomiaru.



Rys. 2.5. Zakłócona obserwacja jednowymiarowego ($L = S = 1$) wyjścia obiektu o znanej charakterystyce statycznej



Rys. 2.6. Estymacja parametrów charakterystyki statycznej obiektu w obecności zakłóceń pomiarowych

W przypadku zakłóceń addytywnych funkcja (2.20) ma postać

$$w = h(y, z) = y + z, \tag{2.21}$$

a zakłóceń multiplikatywnych

$$w = h(y, z) = y^T z, \quad (2.22)$$

co w jednowymiarowym przypadku ($L = 1$) daje $w = yz$.

W dalszych rozważaniach (dla uproszczenia zapisu) założymy, że funkcja h jest wzajemnie jednoznaczna względem z . Znaczy to, że istnieje funkcja odwrotna h_z^{-1} względem z , czyli

$$z = h_z^{-1}(y, w). \quad (2.23)$$

Rezygnacja z powyższego założenia wymaga dodatkowych przekształceń, a konkretnie podzielenia zbioru zakłóceń \mathcal{Z} na rozłączne podzbiory, w których funkcja h jest różnowartościowa. Uzupełniające rozważania są podane w Dodatku.

W n -tym pomiarze dla zadanej wartości wejścia u_n zarejestrowaliśmy wartość w_n , która jest wynikiem zakłóconej obserwacji wyjścia y_n . O zakłóceniach z_n w n -tym pomiarze będziemy zakładać, że są niezależnymi wartościami L -wymiarowej zmiennej losowej \underline{z} .

W dalszych rozważaniach przyjmiemy, że zmienna losowa \underline{z} przyjmuje wartości ze zbioru ciągłego \mathcal{Z} , a funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{z} określoną na \mathcal{Z} oznaczymy przez $f_z(z)$. W wyniku eksperymentu, dla zadanej serii identyfikującej o długości N , uzyskaliśmy wyniki pomiarów wartości wyjść. Dane pomiarowe zebrano w macierzach:

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad W_N = [w_1 \quad w_2 \quad \cdots \quad w_N]. \quad (2.24)$$

Teraz zadanie identyfikacji sprowadza się do oszacowania nieznanego wektora parametrów θ charakterystyki statycznej obiektu. Poszukujemy algorytmu estymacji (rys. 2.6)

$$\hat{\theta}_N = \Psi_N(U_N, W_N), \quad (2.25)$$

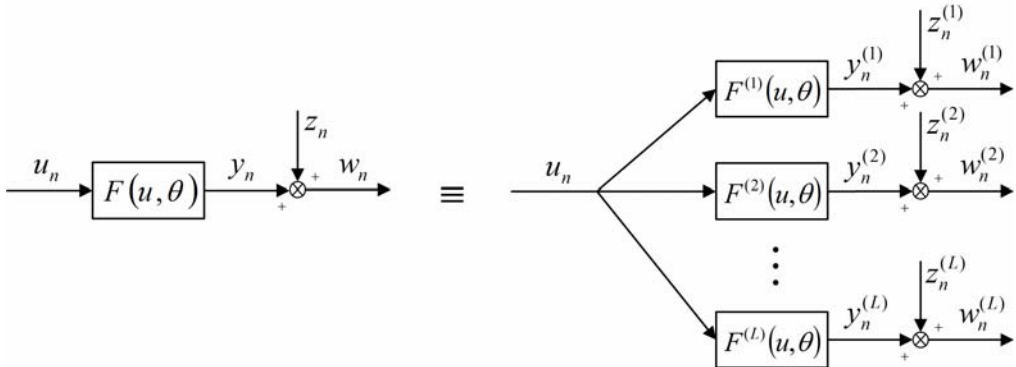
który dla zebranych danych pomiarowych U_N, W_N umożliwi wyznaczenie wartości oszacowania $\hat{\theta}_N$ wektora θ (Ψ_N oznacza algorytm estymacji dla serii pomiarowych o długości N).

W zależności od informacji, jakie posiadamy o systemie pomiarowym, oraz od warunków, jakie spełniają zakłócenia, aby uzyskać oszacowania o oczekiwanych własnościach [78] stosujemy różne metody estymacji.

Metoda najmniejszych kwadratów

Metodę najmniejszych kwadratów przedstawimy dla obiektu o jednym wyjściu ($\dim y = \dim z = L = 1$). Nie ogranicza to ogólności rozważań. Obiekt identyfikacji o L -wymiarowym wektorze wyjść można zdekomponować na L obiektów o jednym wyjściu, odpowiadającemu l -tej składowej wektora wyjściowego (rys. 2.7). W przy-

padku obiektu z l -wymiarowym wektorem wyjść wyznaczenie estymatora sprowadza się do minimalizacji empirycznego oszacowania śladu macierzy kowariancji zakłóceń pomiarowych.



Rys. 2.7. Dekompozycja obiektu identyfikacji o L -wymiarowym wektorze wyjść na L obiektów o jednowymiarowym wyjściu $y^{(l)} = F^{(l)}(u, \theta)$, $l = 1, 2, \dots, L$

Jeśli zakłócenia pomiarowe są addytywne, tj. funkcja h , która opisuje system pomiarowy, ma postać (2.21), oraz zakłócenia pomiarowe są wartościami zmiennej losowej \underline{z} , o zerowej wartości oczekiwanej i skończonej wariancji [14, 79], tj.

$$E[\underline{z}] = 0, \quad \text{Var}[\underline{z}] < \infty, \quad (2.26)$$

możemy skorzystać z metody najmniejszych kwadratów. W takim przypadku wyznaczenie oszacowania (2.25) sprowadza się do wyznaczenia takiej wartości θ_N , która minimalizuje empiryczną wariancję błędu pomiarowego. Empiryczna wariancja Var_{z_N} , wyznaczona na podstawie zbioru N obserwacji, zależy od pomiarów U_N , W_N oraz wektora parametrów obiektu θ i wyraża się wzorem

$$\text{Var}_{z_N}(U_N, W_N, \theta) \stackrel{\text{df}}{=} \frac{1}{N} \sum_{n=1}^N (w_n - y_n)^2 = \frac{1}{N} \sum_{n=1}^N (w_n - F(u_n, \theta))^2. \quad (2.27)$$

Problem wyznaczenia oszacowania **metodą najmniejszych kwadratów** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, W_N)$, dla której wyrażenie (2.27) przyjmuje wartość minimalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, W_N) : \text{Var}_{z_N}(U_N, W_N, \hat{\theta}_N) = \min_{\theta \in \Theta} \text{Var}_{z_N}(U_N, W_N, \theta). \quad (2.28)$$

Przykład 2.2. Obiekt (rys. 2.7) o jednym wyjściu ($L = 1$) jest opisany charakterystyką liniową względem parametrów jak w przykładzie 2.1 (2.9). Dla zadanej serii identyfikującej wyjście obiektu jest obserwowane z addytywnymi (2.21) zakłóceniami, które spełniają warunek (2.26).

Przyjęte założenia upoważniają do zastosowania metody najmniejszych kwadratów. Empiryczna wariancja zakłóceń (2.27) przyjmuje postać

$$\text{Var}_{z_N}(U_N, W_N, \theta) = \frac{1}{N} \sum_{n=1}^N (w_n - y_n)^2 = \frac{1}{N} \sum_{n=1}^N (w_n - \theta^T f(u_n))^2. \quad (2.29)$$

Minimalizacja (2.29) względem θ daje algorytm estymacji postaci

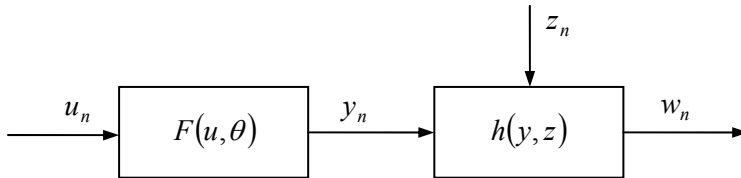
$$\hat{\theta}_N = \Psi_N(U_N, W_N) = \left[\sum_{n=1}^N f(u_n) f^T(u_n) \right]^{-1} \sum_{n=1}^N w_n f(u_n). \quad (2.30)$$

◆

Dla wybranych postaci charakterystyk (2.1) można uzyskać analityczne (np. (2.30)) rozwiązanie zadania (2.28). Dla dowolnej postaci nieliniowej funkcji F w zależności (2.1) może pojawić się problem z analitycznym rozwiązaniem zadania (2.28). W takim przypadku do rozwiązania tego zadania korzystamy z numerycznych metod optymalizacji.

Metoda maksymalnej wiarygodności

Teraz założymy, że zakłócenia z_n , w n -tym pomiarze, są niezależnymi wartościami L -wymiarowej zmiennej losowej \underline{z} , dla której jest znana funkcja gęstości rozkładu prawdopodobieństwa $f_z(z)$ określona na $\bar{\mathcal{Z}}$. Sposób nakładania się zakłóceń na mierzoną wartość na wyjściu obiektu (rys. 2.8) jest opisany znaną zależnością (2.20), w której h jest dowolną, znaną funkcją, wzajemnie jednoznaczną względem z .



Rys. 2.8. Pomiar wartości wyjścia obiektu o charakterystyce statycznej $F(u, \theta)$ z zakłóceniami pomiarowymi

Zbiór wartości w_n , $n=1, 2, \dots, N$, możemy potraktować jako zbiór niezależnych realizacji L -wymiarowej zmiennej losowej \underline{w} , która jest wynikiem przekształcenia zmiennej losowej \underline{z} zgodnie z zależnością (2.20), czyli

$$\underline{w} = h(y, \underline{z}), \quad (2.31)$$

a po uwzględnieniu (2.1)

$$\underline{w} = h(F(u, \theta), \underline{z}). \quad (2.32)$$

Funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej w możemy wyznaczyć, korzystając z przekształcenia (2.32). Zwróćmy uwagę, że wektor wejść u i wektor θ są parametrami tego przekształcenia. Dla ustalonego u funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej $w - f_w(w, \theta; u)$ z wektorem parametrów θ wyznaczamy według wzoru

$$f_w(w, \theta; u) = f_z(h_z^{-1}(F(u, \theta), w)) |J_h|, \quad (2.33)$$

w którym J_h jest macierzą Jakobiego (jakobianem) przekształcenia odwrotnego (2.23) [48, 79]

$$J_h = \frac{\partial h_z^{-1}(y, w)}{\partial w}. \quad (2.34)$$

W wyniku eksperymentu dla ustalonej serii identyfikującej U_N otrzymujemy macierz zakłóconych wyników pomiarów wyjść W_N . Jest to zbiór niezależnych realizacji zmiennej losowej w , której funkcja gęstości rozkładu prawdopodobieństwa jest dana zależnością (2.33). Dla zadanego u_n , w n -tym pomiarze, wektor parametrów θ nie jest znany w tej zależności. Możemy go wyznaczyć, korzystając z metody maksymalnej wiarygodności. Funkcja wiarygodności L_N , wyznaczona na podstawie macierzy pomiarów W_N dla zadanej serii U_N , zależy od parametrów charakterystyki statycznej obiektu θ i wyraża się wzorem

$$L_N(W_N, \theta; U_N) \stackrel{\text{df}}{=} \prod_{n=1}^N f_w(w_n, \theta; u_n) = \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \theta), w_n)) |J_h|. \quad (2.35)$$

Wyznaczenie oszacowania **metodą maksymalnej wiarygodności** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, W_N)$, dla której wyrażenie (2.35) przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, W_N) : L_N(W_N, \hat{\theta}_N; U_N) = \max_{\theta \in \Theta} L_N(W_N, \theta; U_N). \quad (2.36)$$

Przykład 2.3. Obiekt o jednym wyjściu ($L = 1$) jest opisany charakterystyką liniową względem parametrów, jak w przykładzie 2.1 (2.9). Dla zadanej serii identyfikującej wyjście obiektu jest obserwowane z addytywnymi (2.21) zakłóceniami. Zakłada się, że w kolejnych pomiarach zakłócenia są niezależnymi wartościami zmiennej losowej z o rozkładzie normalnym z wartością oczekiwaną m_z i odchyleniu standardowym σ_z , czyli

$$f_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(z - m_z)^2}{2\sigma_z^2} \right]. \quad (2.37)$$

W rozpatrywanym przykładzie funkcja odwrotna h_z^{-1} , z uwzględnieniem charakterystyki statycznej (2.1), wyrażająca się zależnością (2.9), ma postać

$$z = h_z^{-1}(F(u, \theta), w) = w - \theta^T f(u). \quad (2.38)$$

Jakobian (2.34) przekształcenia odwrotnego (2.38) jest równy

$$J_h = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw} (w - \theta^T f(u)) = 1, \quad (2.39)$$

a zatem funkcja gęstości rozkładu prawdopodobieństwa (2.33) wartości obserwowanych na wyjściu obiektu ma postać

$$f_w(w, \theta; u) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(w - \theta^T f(u) - m_z)^2}{2\sigma_z^2} \right] |1|. \quad (2.40)$$

Funkcja wiarygodności (2.35)

$$L_N(W_N, \theta; U_N) = \prod_{n=1}^N \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(w_n - \theta^T f(u_n) - m_z)^2}{2\sigma_z^2} \right], \quad (2.41)$$

a po przekształceniach

$$L_N(W_N, \theta; U_N) = \left(\frac{1}{\sigma_z \sqrt{2\pi}} \right)^N \exp \left[\sum_{n=1}^N -\frac{(w_n - \theta^T f(u_n) - m_z)^2}{2\sigma_z^2} \right]. \quad (2.42)$$

Maksymalizacja funkcji wiarygodności w tym przypadku sprowadza się do optymalizacji wykładnika wyrażenia (2.42). Przyrównanie gradientu wyrażenia (2.42) względem θ do zera daje układ równań, którego rozwiązanie

$$\hat{\theta}_N = \Psi_N(U_N, W_N) = \left[\sum_{n=1}^N f(u_n) f^T(u_n) \right]^{-1} \sum_{n=1}^N (w_n - m_z) f(u_n) \quad (2.43)$$

jest algorytmem estymacji. Zwróćmy uwagę na intuicyjnie oczywisty wynik (2.43), który od każdego pomiaru w_n nakazuje odjąć systematyczny błąd pomiaru m_z .



Przykład 2.4. Rozważmy liniowy jednowymiarowy obiekt (rys. 2.8 – o jednym wejściu i jednym wyjściu, czyli $S = L = 1$). Charakterystyka (2.1) ma postać

$$y = F(u, \theta) = \theta u. \quad (2.44)$$

W celu skupienia uwagi zakładamy, że nieznaną parametr θ ($R = 1$) przyjmuje wartości dodatnie $\Theta = \{\theta \in \mathbb{R}^1 : \theta > 0\}$. Dla zadanej serii identyfikującej wyjście obiektu jest obserwowane z multiplikatywnymi (2.22) zakłóceniami. W kolejnych pomiarach zakłócenia są niezależnymi wartościami zmiennej losowej \underline{z} o rozkładzie jednostajnym na odcinku $[0, 1]$, czyli

$$f_z(z) = \begin{cases} 1 & \text{gdy } z \in [0, 1] \\ 0 & \text{gdy } z \notin [0, 1] \end{cases} \quad (2.45)$$

Założymy ponadto, że wejście może przyjmować tylko wartości dodatnie, tj.

$$\forall n = 1, 2, \dots, N, \quad u_n > 0. \quad (2.46)$$

W rozpatrywanym przykładzie funkcja h , z uwzględnieniem charakterystyki statycznej (2.44), ma postać

$$w = h(y, z) = h(F(u, \theta), z) = \theta u z, \quad (2.47)$$

a funkcja odwrotna h_z^{-1} (2.23)

$$z = h_z^{-1}(F(u, \theta), w) = \frac{w}{\theta u}. \quad (2.48)$$

Jakobian przekształcenia odwrotnego ((2.34)) ma postać

$$J_h = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw} \left(\frac{w}{\theta u} \right) = \frac{1}{\theta u}, \quad (2.49)$$

a zatem funkcja gęstości rozkładu prawdopodobieństwa (2.33) wartości obserwowanych na wyjściu obiektu ma postać

$$f_w(w, \theta; u) = \begin{cases} \frac{1}{\theta u} & \text{gdy } \frac{w}{\theta u} \in [0, 1] \\ 0 & \text{gdy } \frac{w}{\theta u} \notin [0, 1] \end{cases}. \quad (2.50)$$

Funkcja wiarygodności (2.35)

$$L_N(W_N, \theta; U_N) = \begin{cases} \frac{1}{\theta \prod_{n=1}^N u_n} & \text{gdy } \forall n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \in [0, \theta] \\ 0 & \text{gdy } \exists n = 1, 2, \dots, N \quad \frac{w_n}{u_n} \notin [0, \theta] \end{cases}, \quad (2.51)$$

a po przekształceniach

$$L_N(W_N, \theta; U_N) = \begin{cases} \frac{1}{\theta \prod_{n=1}^N u_n} & \text{gdy } \theta \geq \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} \\ 0 & \text{gdy } \theta < \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} \end{cases}. \quad (2.52)$$

Maksymalizacja funkcji wiarygodności (2.52) względem θ daje

$$\hat{\theta}_N = \Psi_N(U_N, W_N) = \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\}. \quad (2.53)$$

Ponownie zwróćmy uwagę na wynik. Z treści problemu opisanego w przykładzie 2.4 wynika, że obserwowane wyjście jest pomnożone przez liczbę z przedziału $[0, 1]$. Inaczej mówiąc – obserwacje wyjścia są tłumione, a algorytm estymacji (2.53) proponuje wybrać taki pomiar, który jest najmniej tłumiony (pomnożony przez największą liczbę z przedziału $[0, 1]$), co łatwo zauważyć po podstawieniu równania (2.47) do wzoru (2.53), tj.

$$\hat{\theta}_N = \max_{1 \leq n \leq N} \left\{ \frac{w_n}{u_n} \right\} = \max_{1 \leq n \leq N} \left\{ \frac{\theta u_n z_n}{u_n} \right\} = \theta \max_{1 \leq n \leq N} \{z_n\}. \quad (2.54)$$



Metody Bayesa

Teraz – dodatkowo w stosunku do poprzednich rozważań – założymy, że badany obiekt został wylosowany z pewnej populacji. Może to być partia obiektów, w których parametry poszczególnych egzemplarzy różnią się między sobą. Dla danej partii obiektów wielkościami charakterystycznymi są: wartość średnia parametrów oraz wariancja. Informacja taka może być pomocna w rozwiązaniu zadania identyfikacji. Precyzyjniej rzecz ujmując, zakładamy teraz, że dla badanego obiektu (rys. 2.8) znana jest charakterystyka statyczna (2.1) z dokładnością do parametrów. System pomiarowy jest opisany zależnością (2.20), zakłócenia w n -tym pomiarze są niezależnymi wartościami zmiennej losowej z , dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_z(z)$, określona na przestrzeni zakłóceń \mathcal{Z} . Dodatkowo wektor parametrów θ jest wartością ciągłej, R -wymiarowej zmiennej losowej $\underline{\theta}$, która przyjmuje wartości ze zbioru $\Theta \subseteq \mathbf{R}^R$. Znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$ zmiennej $\underline{\theta}$, określona na Θ . Jest to dla zadania identyfikacji informacja aprioryczna, a funkcja gęstości $f_\theta(\theta)$ zwana jest funkcją gęstości rozkładu prawdopodobieństwa a priori. Do wyznaczenia oszacowania wektora pa-

parametrów charakterystyki statycznej obiektu wykorzystamy bayesowski model podejmowania decyzji.

Przez $\bar{\theta}$ oznaczymy dowolne oszacowanie wektora parametrów θ , uzyskane za pomocą algorytmu $\bar{\Psi}$, który przetwarza macierze pomiarów U_N, W_N , tj.

$$\bar{\theta} = \bar{\Psi}(U_N, W_N). \quad (2.55)$$

Wartość $\bar{\theta}$ jest oszacowaniem uzyskanym za pomocą pewnego algorytmu $\bar{\Psi}$. Do oceny jakości algorytmu wprowadzamy funkcję strat $L(\theta, \bar{\theta})$, która ocenia różnicę pomiędzy wartością nieznanego wektora parametrów θ a jego oszacowaniem $\bar{\theta}$, uzyskanym za pomocą algorytmu $\bar{\Psi}$. Za kryterium wyboru algorytmu przyjmujemy ryzyko podjęcia decyzji, zdefiniowane jako wartość oczekiwana funkcji strat, czyli

$$\begin{aligned} R(\bar{\Psi}) &\stackrel{\text{df}}{=} \mathbb{E}_{\underline{\theta}, \bar{\theta}} [L(\underline{\theta}, \bar{\theta})] \mathbb{E}_{\underline{\theta}, \underline{W}_N} [L(\underline{\theta}, \bar{\Psi}(U_N, \underline{W}_N))] \\ &= \int \int_{W_N} L(\theta, \bar{\Psi}(U_N, W_N)) f(\theta, W_N; U_N) d\theta dW_N, \end{aligned} \quad (2.56)$$

gdzie: $W_N = \overbrace{W \times W \times \dots \times W}^N$ jest przestrzenią zbioru N wyników pomiarów, a $f(\theta, W_N; U_N)$ – funkcją gęstości rozkładu prawdopodobieństwa pojawienia się wektora parametrów o wartości θ oraz macierzy wyniku pomiarów wyjścia W_N przy zadanej serii identyfikującej U_N .

Zauważmy, że ryzyko (2.56) zależy od algorytmu $\bar{\Psi}$. Wyznaczenie optymalnego algorytmu sprowadza się do wyboru takiego algorytmu Ψ_N , dla którego ryzyko (2.56) przyjmuje minimum, czyli

$$\Psi_N : R(\Psi_N) = \min_{\bar{\Psi}} R(\bar{\Psi}). \quad (2.57)$$

Funkcję gęstości rozkładu prawdopodobieństwa $f(\theta, W_N; U_N)$ można przedstawić w postaci

$$f(\theta, W_N; U_N) = f'(\theta | W_N; U_N) f''(W_N; U_N), \quad (2.58)$$

gdzie $f'(\theta | W_N; U_N)$ jest warunkową gęstością rozkładu prawdopodobieństwa zmiennej losowej $\underline{\theta}$, pod warunkiem, że seria L -wymiarowych niezależnych zmiennych losowych \underline{W} o długości N przyjęła wartość W_N , czyli realizacja wielowymiarowej zmiennej losowej \underline{W}_N jest równa macierzy W_N , przy zadanej serii identyfikującej U_N , a $f''(W_N; U_N)$ jest funkcją gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{W}_N przy zadanym U_N . Funkcja gęstości rozkładu prawdopodobieństwa $f'(\theta | W_N; U_N)$ jest zwana funkcją gęstości rozkładu prawdopodobieństwa a posteriori.

Funkcję gęstości rozkładu prawdopodobieństwa a posteriori wyznaczamy z zależności

$$f'(\theta|W_N;U_N) = \frac{f_\theta(\theta)f_{W_N}(W_N|\theta;U_N)}{f''(W_N;U_N)} = \frac{f_\theta(\theta)f_{W_N}(W_N|\theta;U_N)}{\int_{\Theta} f_\theta(\theta)f_{W_N}(W_N|\theta;U_N)d\theta}, \quad (2.59)$$

w której $f_{W_N}(W_N|\theta;U_N)$ jest warunkową gęstością rozkładu prawdopodobieństwa uzyskania macierzy pomiarów W_N , pod warunkiem, że zmienna losowa $\underline{\theta}$, opisująca wektor parametrów obiektu, przyjął wartość θ , a na wejście podano serię identyfikującą U_N .

Pomiary w_n , $n = 1, 2, \dots, N$, są niezależnymi wartościami zmiennej losowej w , pod warunkiem, że wektor parametrów przyjął wartość θ , a na wejście podano u_n . Funkcję gęstości f_{W_N} obliczymy ze wzoru

$$f_{W_N}(W_N|\theta;U_N) = \prod_{n=1}^N f_w(w_n|\theta;u_n), \quad (2.60)$$

w którym $f_w(w|\theta;u)$ jest warunkową funkcją gęstości rozkładu prawdopodobieństwa zmiennej losowej w , pod warunkiem, że zmienna losowa $\underline{\theta}$ przyjął wartość θ i składowe wektora wejść obiektu przyjął wartość u_n . Inaczej mówiąc – jest to funkcja gęstości rozkładu prawdopodobieństwa zdarzenia, że w n -tym pomiarze na wyjściu zmierzono w_n , przy wektorze parametrów charakterystyki równym θ i wektorze wejść u_n .

Funkcja ta, podobnie jak (2.33), zależy od funkcji gęstości rozkładu prawdopodobieństwa zakłóceń $f_z(z)$, opisu systemu pomiarowego (2.31) oraz charakterystyki statycznej obiektu (2.1) i jest określona zależnością

$$f_w(w|\theta;u) = f_z(h_z^{-1}(F(u,\theta),w))|J_h|, \quad (2.61)$$

w której J_h jest jakobianem przekształcenia odwrotnego h_z^{-1} (2.34).

Ostatecznie, po podstawieniu (2.60) do (2.59), z uwzględnieniem (2.61), funkcja gęstości rozkładu prawdopodobieństwa a posteriori wyraża się wzorem

$$f'(\theta|W_N;U_N) = \frac{f_\theta(\theta) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n,\theta),w_n))|J_h|}{\int_{\Theta} f_\theta(\theta) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n,\theta),w_n))|J_h|d\theta}. \quad (2.62)$$

Korzystając z zależności (2.58), ryzyko (2.56) możemy zapisać

$$R(\bar{\Psi}) = \int_{W_N} \int_{\Theta} L(\theta, \bar{\Psi}(U_N, W_N)) f'(\theta|W_N;U_N) d\theta f''(W_N;U_N) dW_N. \quad (2.63)$$

Dla zadanej serii identyfikującej U_N uzyskujemy macierz pomiarów W_N . Po wykonaniu pomiarów możemy stwierdzić, że wielowymiarowa zmienna losowa \underline{W}_N przyjęła wartość W_N . Dla ustalonych macierzy wyników pomiarów U_N, W_N funkcja $\bar{\Psi}$ przyjmuje wartość $\bar{\theta}$. Możemy zdefiniować ryzyko warunkowe podjęcia decyzji, że wektor parametrów θ przyjmie wartość $\bar{\theta}$ dla macierzy pomiarów U_N, W_N

$$r(\bar{\theta}, W_N; U_N) \stackrel{\text{df}}{=} \underset{\underline{\theta}}{\text{E}} [L(\underline{\theta}, \bar{\theta}) | W_N; U_N] = \int_{\underline{\theta}} L(\underline{\theta}, \bar{\theta} = \bar{\Psi}(U_N, W_N)) f'(\underline{\theta} | W_N; U_N) d\underline{\theta}. \quad (2.64)$$

Rozwiązanie zadania optymalizacji (2.57) jest równoważne wyznaczeniu takiej wartości oszacowania $\bar{\theta} = \hat{\theta}_N$, która minimalizuje ryzyko warunkowe (2.64). Zadanie minimalizacji ryzyka (2.56) względem funkcji $\bar{\Psi}$ jest równoważne zadaniu minimalizacji ryzyka warunkowego (2.64) względem wartości funkcji $\bar{\Psi}$ dla ustalonych macierzy wyników pomiarów U_N, W_N i sprowadza się do wyznaczenia takiej wartości oszacowania $\hat{\theta}_N = \Psi_N(U_N, W_N)$, dla której wyrażenie (2.64) przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, W_N) : r(\hat{\theta}_N, W_N; U_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, W_N; U_N). \quad (2.65)$$

W przypadku różnych postaci funkcji strat otrzymujemy różne metody estymacji. Wskażemy tutaj dwa ważne przypadki. Dla kwadratowej funkcji strat

$$L(\underline{\theta}, \bar{\theta}) = (\underline{\theta} - \bar{\theta})^T (\underline{\theta} - \bar{\theta}) \quad (2.66)$$

otrzymujemy **metodę średniej a posteriori**. Ryzyko warunkowe (2.64) ma postać

$$\begin{aligned} r(\bar{\theta}, W_N; U_N) &\stackrel{\text{df}}{=} \underset{\underline{\theta}}{\text{E}} \left[(\underline{\theta} - \bar{\theta})^T (\underline{\theta} - \bar{\theta}) | W_N; U_N \right] \\ &= \underset{\underline{\theta}}{\text{E}} \left[\underline{\theta}^T \underline{\theta} | W_N; U_N \right] - 2 \underset{\underline{\theta}}{\text{E}} \left[\underline{\theta}^T | W_N; U_N \right] + \bar{\theta}^T \bar{\theta}. \end{aligned} \quad (2.67)$$

Przyrównanie gradientu względem $\bar{\theta}$ (2.67) daje układ równań

$$\left. \text{grad} r(\bar{\theta}, W_N; U_N) \right|_{\bar{\theta} = \hat{\theta}_N} = -2 \underset{\underline{\theta}}{\text{E}} \left[\underline{\theta} | W_N; U_N \right] + 2 \hat{\theta}_N = \mathbf{0}_R, \quad (2.68)$$

gdzie $\mathbf{0}_R$ oznacza R -wymiarowy kolumnowy wektor zerowy, a jego rozwiązanie daje algorytm estymacji

$$\hat{\theta}_N = \Psi_N(U_N, W_N) = \underset{\underline{\theta}}{\text{E}} \left[\underline{\theta} | W_N; U_N \right] = \int_{\underline{\theta}} \underline{\theta} f'(\underline{\theta} | W_N; U_N) d\underline{\theta}, \quad (2.69)$$

który sprowadza się do wyznaczenia wartości oczekiwanej zmiennej losowej θ , pod warunkiem, że wielowymiarowa zmienna losowa W_N przyjęła wartość W_N przy zadanej serii identyfikującej U_N . Dla funkcji strat w postaci

$$L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta}), \quad (2.70)$$

gdzie $\delta(\cdot)$ oznacza funkcję delta Diraca, ryzyko warunkowe (2.64) przyjmuje postać

$$r(\bar{\theta}, W_N; U_N) = -\int_{\Theta} \delta(\theta - \bar{\theta}) f'(\theta | W_N; U_N) d\theta = -f'(\bar{\theta} | W_N; U_N). \quad (2.71)$$

W tym przypadku rozwiązanie zadania optymalizacji (2.65) sprowadza się do maksymalizacji funkcji gęstości rozkładu prawdopodobieństwa a posteriori względem $\bar{\theta}$, czyli

$$\hat{\theta}_N = \Psi_N(U_N, W_N) : f'(\hat{\theta}_N | W_N; U_N) = \max_{\bar{\theta} \in \Theta} f'(\bar{\theta} | W_N; U_N), \quad (2.72)$$

a metoda nosi nazwę **metody maksymalnego prawdopodobieństwa a posteriori**.

Zwróćmy uwagę, że dla ustalonych macierzy pomiarów W_N , U_N całka w mianowniku wyrażenia (2.62) jest wartością stałą i nie zależy od θ . Wyznaczenie oszacowania metodą maksymalnego prawdopodobieństwa a posteriori jest zatem równoważne maksymalizacji licznika wyrażenia (2.62) względem $\bar{\theta}$ i sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, W_N)$, dla której

$$f_{\theta}(\hat{\theta}_N) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \hat{\theta}_N), w_n)) | J_h | = \max_{\bar{\theta} \in \Theta} f_{\theta}(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) | J_h |. \quad (2.73)$$

Przykład 2.5. Rozważmy liniowy jednowymiarowy obiekt (rys. 2.8 – o jednym wejściu i jednym wyjściu, czyli $S = L = 1$). Charakterystyka statyczna (2.1) ma postać $y = F(u, \theta) = \theta u$. Wyjście obiektu jest obserwowane z addytywnymi (2.21) zakłóceniami. W kolejnych pomiarach zakłócenia są niezależnymi wartościami zmiennej losowej z , która ma rozkład normalny, o zerowej wartości oczekiwanej $\left(E[z] = 0\right)$ i odchyleniu standardowym σ_z . Parametr θ jest wartością zmiennej losowej θ o wartości oczekiwanej m_{θ} i odchyleniu standardowym σ_{θ} . Funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej z ma postać

$$f_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{z^2}{2\sigma_z^2}\right], \quad (2.74)$$

natomiast zmiennej $\underline{\theta}$

$$f_{\theta}(\theta) = \frac{1}{\sigma_{\theta}\sqrt{2\pi}} \exp\left[-\frac{(\theta - m_{\theta})^2}{2\sigma_{\theta}^2}\right]. \quad (2.75)$$

Przyjmujemy ponadto, że funkcja strat jest postaci (2.70).

Wyznaczenie oszacowania parametru θ sprowadza się do rozwiązania zadania (2.73). W rozpatrywanym przykładzie funkcja odwrotna h_z^{-1} , po uwzględnieniu charakterystyki statycznej (2.1), danej zależnością (2.44), ma postać

$$z = h_z^{-1}(F(u, \theta), w) = w - \theta u. \quad (2.76)$$

Jakobian przekształcenia odwrotnego (2.34)

$$J = \frac{\partial h_z^{-1}(F(u, \theta), w)}{\partial w} = \frac{d}{dw}(w - \theta u) = 1, \quad (2.77)$$

a zatem funkcja gęstości rozkładu prawdopodobieństwa (2.61) wartości obserwowanych na wyjściu obiektu

$$f_w(w|\theta; u) = \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{(w - \theta u)^2}{2\sigma_z^2}\right] |1|. \quad (2.78)$$

W zadaniu (2.73) wyrażenie

$$\begin{aligned} & f_{\theta}(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \theta), w_n)) |J| \\ &= \frac{1}{\sigma_{\theta}\sqrt{2\pi}} \exp\left[-\frac{(\bar{\theta} - m_{\theta})^2}{2\sigma_{\theta}^2}\right] \prod_{n=1}^N \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2}\right], \end{aligned} \quad (2.79)$$

po przekształceniach ma postać

$$\begin{aligned} & f_{\theta}(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) |J_h| \\ &= \frac{1}{\sigma_{\theta}\sqrt{2\pi}} \left(\frac{1}{\sigma_z\sqrt{2\pi}}\right)^N \exp\left[-\frac{(\bar{\theta} - m_{\theta})^2}{2\sigma_{\theta}^2} - \frac{\sum_{n=1}^N (w_n - \bar{\theta} u_n)^2}{2\sigma_z^2}\right]. \end{aligned} \quad (2.80)$$

Maksymalizacja wyrażenia (2.80) względem $\bar{\theta}$ daje

$$\hat{\theta}_N = \Psi_N(U_N, W_N) = \frac{m_\theta + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 \sum_{n=1}^N w_n u_n}{1 + \left(\frac{\sigma_\theta}{\sigma_z}\right)^2 \sum_{n=1}^N u_n^2}. \quad (2.81)$$

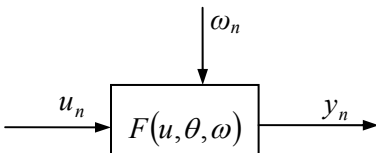
Warto zwrócić uwagę na dwa skrajne przypadki. Pierwszy, gdy N jest małą liczbą i wariancja zakłóceń jest dużo większa od wariancji parametrów ($\sigma_z^2 \gg \sigma_\theta^2$). Opisuje to sytuację, gdy pomiary są mało wartościowe i eksperyment niewiele wnosi do poznania obiektu. Dla takich danych możemy przyjąć, że $\hat{\theta}_N \approx m_\theta$. Decyzja jest oparta głównie na informacji apriorycznej. Taką decyzję można było przewidzieć przed wykonaniem eksperymentu. Odwrotny przypadek. W drugim przypadku, gdy N jest dużą liczbą ($N \rightarrow \infty$) oraz wariancja zakłóceń jest dużo mniejsza od wariancji parametrów ($\sigma_z^2 \ll \sigma_\theta^2$), w wyniku eksperymentu zebrano dużą liczbę bardzo dobrych pomiarów. Takie założenia pozwalają przyjąć oszacowanie

$$\hat{\theta}_N \approx \frac{\sum_{n=1}^N w_n u_n}{\sum_{n=1}^N u_n^2}. \quad (2.82)$$

W oszacowaniu tym istotną rolę odgrywają wyniki pomiarów, a rola informacji apriorycznej jest pomijalnie mała. Zwróćmy uwagę, że pomijając informację aprioryczną, analogiczne rozwiązanie uzyskamy metodą najmniejszych kwadratów (2.30), przyjmując $R = S = 1$ oraz $f(u) = u$. Jest to uzasadnione, gdyż informacja aprioryczna nie odgrywa roli, a pozostałe założenia są spełnione, czyli $E(\underline{z}) = 0$, $\text{Var}(\underline{z}) = \sigma_z^2 < \infty$ oraz $w = y + z$. ◆

2.2.2. Estymacja parametrów charakterystyki statycznej obiektu ze zmiennymi niemierzalnymi wielkościami losowymi

W podrozdziale 2.1 przedstawiono zadanie wyznaczania parametrów charakterystyki statycznej obiektu. Zakładano tam, że obiekt jest deterministyczny. Obecnie wzbogacimy rozważania przedstawione w podrozdziale 2.1, uwzględniając w opisie niemierzalne, zmienne wielkości losowe ω (rys. 2.9).

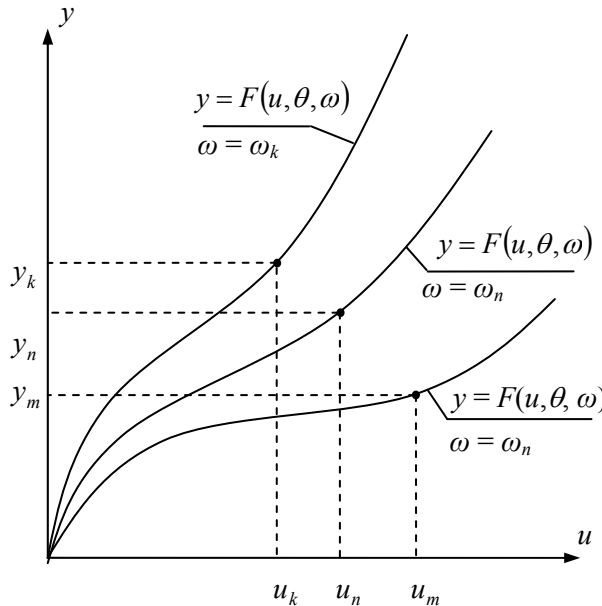


Rys. 2.9. Obiekt identyfikacji z niemierzalnymi zmiennymi wielkościami losowymi

Często się zdarza, że na wyjście badanego obiektu w znany sposób oddziałują wielkości losowe. Podając wielokrotnie tę samą wartość u_n na wejście obiektu, możemy uzyskać różne wartości wyjścia y_n , pomimo że obserwacje są bez zakłóceń. Powodem tych różnic jest losowa zmienność wielkości ω występujących w obiekcie. W konsekwencji możemy stwierdzić, że obserwujemy rodzinę charakterystyk statycznych, indeksowaną losową wielkością ω (rys. 2.10).

Opiszmy precyzyjnie aktualny obiekt identyfikacji (rys. 2.9). Charakterystyka statyczna obiektu o wektorze wejść u i wektorze wyjść y jest znana z dokładnością do wektora parametrów θ . Na obiekt działają pewne niemierzalne wielkości losowe ω z przestrzeni $\Omega \subseteq \mathbb{R}^L$. Charakterystyka statyczna (2.1) jest uzupełniona o wektor wielkości losowych ω , czyli dana jest zależnością

$$y = F(u, \theta, \omega). \quad (2.83)$$



Rys. 2.10. Obserwacja charakterystyki statycznej obiektu jednowymiarowego ($L = S = 1$) z losową wielkością ω

W dalszych rozważaniach (dla uproszczenia zapisu) będziemy zakładać, że wymiar wektora y oraz wektora wielkości losowych ω jest taki sam i równy L ($\dim y = \dim \omega = L$). Funkcja F jest wzajemnie jednoznaczna względem ω , czyli istnieje funkcja odwrotna F_ω^{-1}

$$\omega = F_\omega^{-1}(u, \theta, y). \quad (2.84)$$

Powyższe założenia można osłabić. Gdy $\dim y > \dim \omega$, wystarczy przekształcenie F uzupełnić o dodatkowe przekształcenie prowadzące do jednoznaczności, a dla funkcji F , która nie jest wzajemnie jednoznaczna, zbiór określoności wielkości losowej Ω należy podzielić na rozłączne podzbiory, w których funkcja F jest różnowartościowa. Rozważania te można znaleźć w Dodatku.

W celu wyznaczenia nieznanego wektora parametrów charakterystyki statycznej (2.83), w n -tym pomiarze, dla zadanej wartości składowych wektora wejść u_n , zmierzono wartości składowych wektora wyjść y_n . W wyniku eksperymentu, dla zadanej serii identyfikującej o długości N , uzyskano wyniki pomiarów, które zebrano w macierzach

$$U_N = [u_1 \quad u_2 \quad \cdots \quad u_N], \quad Y_N = [y_1 \quad y_2 \quad \cdots \quad y_N]. \quad (2.85)$$

O losowo zmiennym wektorze wielkości w obiekcie zakładamy, że w n -tym pomiarze jest on równy ω_n i jest wartością L -wymiarowej zmiennej losowej $\underline{\omega}$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\omega(\omega)$. Zadanie identyfikacji sprowadza się do wyznaczenia oszacowania wektora nieznanymi parametrami θ charakterystyki (2.83). Poszukujemy algorytmu estymacji

$$\hat{\theta}_N = \Psi_N(U_N, Y_N), \quad (2.86)$$

który dla zebranych danych pomiarowych w macierzach U_N, Y_N umożliwi wyznaczenie wartości oszacowania $\hat{\theta}_N$ wektora θ .

W zależności od informacji, jaką posiadamy o zmiennej losowej $\underline{\omega}$, charakterystyce (2.83) oraz jej parametrach θ , w celu uzyskania oszacowania o oczekiwanych własnościach stosujemy różne metody estymacji.

Metoda najmniejszych kwadratów

Podobnie jak w punkcie 2.2.1, metodę najmniejszych kwadratów przedstawimy dla obiektu o jednym wyjściu ($\dim y = \dim z = L = 1$).

Metoda najmniejszych kwadratów w tym przypadku ma bardzo ograniczone zastosowanie. Może być wykorzystana dla szczególnej postaci charakterystyki statycznej (2.83)

$$y = F(u, \theta, \omega) = \tilde{F}(u, \theta) + \omega, \quad (2.87)$$

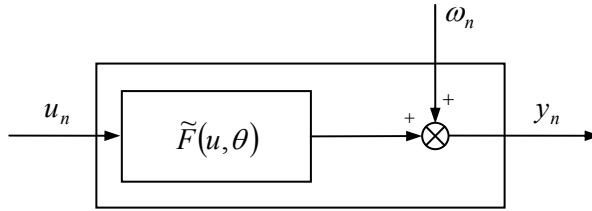
w której wyróżniono (rys. 2.11) część deterministyczną, opisaną funkcją \tilde{F} , do której dodaje się losową wielkość ω występującą w obiekcie.

Losowa wielkość jest wartością zmiennej losowej $\underline{\omega}$, która spełnia warunki

$$E[\underline{\omega}] = 0, \quad \text{Var}[\underline{\omega}] < \infty. \quad (2.88)$$

Jest to przypadek analogiczny do przedstawionego w punkcie 2.2.1, z tym że wówczas była mowa o addytywnych zakłóceniach pomiarowych, a teraz mówimy o wielkości

losowej występującej w obiekcie, która powoduje, że obiekt na te same wymuszenia odpowiada w sposób losowy.



Rys. 2.11. Obiekt identyfikacji z addytywnymi niemierzalnymi zmiennymi wielkościami losowymi

Wyznaczenie oszacowania (2.86) sprowadza się do wyznaczenia takiej wartości $\hat{\theta}_N$, która minimalizuje empiryczną wariancję losowego parametru względem θ . Empiryczna wariancja Var_{ω_N} , wyznaczona na podstawie N obserwacji, zależy od macierzy wyników pomiarów U_N, Y_N oraz wektora parametrów charakterystyki obiektu θ i wyraża się wzorem

$$\text{Var}_{\omega_N}(U_N, Y_N, \theta) \stackrel{\text{df}}{=} \frac{1}{N} \sum_{n=1}^N (y_n - \tilde{F}(u_n, \theta))^2. \quad (2.89)$$

Wyznaczenie oszacowania **metodą najmniejszych kwadratów** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, Y_N)$, dla której wyrażenie (2.89) przyjmuje wartość minimalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : \text{Var}_{\omega_N}(U_N, Y_N, \hat{\theta}_N) = \min_{\theta \in \Theta} \text{Var}_{\omega_N}(U_N, Y_N, \theta). \quad (2.90)$$

Metoda maksymalnej wiarygodności

Teraz założymy, że wartości składowych wektora wielkości losowych ω w obiekcie opisanym charakterystyką statyczną (2.83) w n -tym pomiarze są niezależnymi realizacjami ω_n L -wymiarowej zmiennej losowej $\underline{\omega}$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_{\omega}(\omega)$. Obserwowany zbiór wartości składowych wektora $y_n, n=1, 2, \dots, N$, możemy potraktować jako zbiór niezależnych realizacji L -wymiarowej zmiennej losowej \underline{y} , która jest wynikiem przekształcenia zmiennej losowej $\underline{\omega}$ zgodnie z zależnością (2.83), czyli

$$\underline{y} = F(u, \theta, \underline{\omega}). \quad (2.91)$$

Funkcję gęstości rozkładu prawdopodobieństwa obserwowanej zmiennej losowej \underline{y} możemy wyznaczyć, korzystając z przekształcenia (2.91). Zwróćmy uwagę, że wektor wejść u i wektor θ są parametrami tego przekształcenia. Dla ustalonego wektora wejść u funkcję gęstości rozkładu prawdopodobieństwa $f_y(y, \theta; u)$ zmiennej losowej \underline{y} wyznaczamy według wzoru

$$f_y(y, \theta; u) = f_\omega(F_\omega^{-1}(u, \theta, y)) |J_F|, \quad (2.92)$$

w którym J_F jest macierzą Jakobiego (jakobianem) przekształcenia odwrotnego (2.84) [48, 79]

$$J_F = \frac{\partial F_\omega^{-1}(u, \theta, y)}{\partial y}. \quad (2.93)$$

Oszacowanie nieznanego wektora parametrów θ charakterystyki statycznej (2.83) otrzymamy, maksymalizując funkcję wiarygodności

$$L_N(Y_N, \theta; U_N) \stackrel{\text{df}}{=} \prod_{n=1}^N f_y(y_n, \theta; u_n) = \prod_{n=1}^N f_\omega(F_\omega^{-1}(u_n, \theta, y_n)) |J_F| \quad (2.94)$$

względem θ .

W tym przypadku zadanie wyznaczenia oszacowania **metodą maksymalnej wiarygodności** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, Y_N)$, dla której wyrażenie (2.94) przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : L_N(Y_N, \hat{\theta}_N; U_N) = \max_{\theta \in \Theta} L_N(Y_N, \theta; U_N). \quad (2.95)$$

Metody Bayesa

Załóżmy, dodatkowo w stosunku do rozważań z poprzedniego punktu, że badany obiekt jest wylosowany z pewnej populacji. Oznacza to, że wektor parametrów θ jest wartością ciągłej zmiennej losowej $\underline{\theta}$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$. Teraz obydwa wektory – parametrów θ oraz wielkości losowych ω – występujące w opisie (2.83) są wartościami R -, L -wymiarowych zmiennych losowych – odpowiednio – $\underline{\theta}$ i $\underline{\omega}$. Istotna różnica pomiędzy tymi parametrami polega na tym, że z chwilą wylosowania obiektu do badań wektor parametrów θ jest stały, natomiast wektor wielkości losowych ω zmienia się. Dalej zakładamy, że znane są funkcje gęstości rozkładów prawdopodobieństwa zmiennych losowych $\underline{\theta}$ oraz $\underline{\omega}$. Podobnie jak w punkcie 2.2.1, dla ustalonych macierzy wyników pomiarów U_N, Y_N warunkowe ryzyko podjęcia decyzji $\bar{\theta} = \bar{\Psi}(U_N, Y_N)$ ma postać

$$r(\bar{\theta}, Y_N; U_N) \stackrel{\text{df}}{=} E_{\underline{\theta}} [L(\underline{\theta}, \bar{\theta}) | Y_N; U_N] = \int_{\Theta} L(\theta, \bar{\theta} = \bar{\Psi}(U_N, Y_N)) f'(\theta | Y_N; U_N) d\theta. \quad (2.96)$$

W tym przypadku warunkową funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej y , obserwowanej na wyjściu, pod warunkiem, że wektor parametrów przyjął wartość θ , a wektor wejść przyjął wartość u , wyznaczamy na podstawie $f_{\omega}(\omega)$ oraz charakterystyki (2.83), zgodnie ze wzorem

$$f_y(y | \theta; u) = f_{\omega}(F_{\omega}^{-1}(u, \theta, y)) |J_F|, \quad (2.97)$$

w którym J_F jest jacobianem określonym zależnością (2.93).

Odpowiednio – funkcja gęstości rozkładu prawdopodobieństwa a posteriori $f'(\theta | Y_N; U_N)$ dana jest zależnością

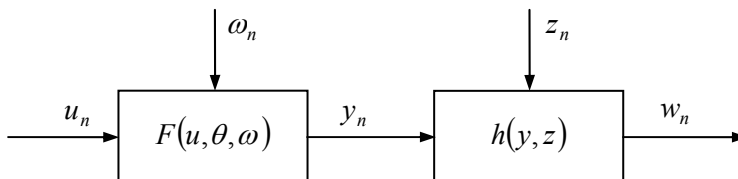
$$f'(\theta | Y_N; U_N) = \frac{f_{\theta}(\theta) \prod_{n=1}^N f_{\omega}(F_{\omega}^{-1}(u_n, \theta, y_n)) |J_F|}{\int_{\Theta} f_{\theta}(\theta) \prod_{n=1}^N f_{\omega}(F_{\omega}^{-1}(u_n, \theta, y_n)) |J_F| d\theta}, \quad (2.98)$$

a wyznaczenie oszacowania sprowadza się do wyznaczenia takiej wartości $\hat{\theta}_N = \Psi_N(U_N, Y_N)$, dla której wyrażenie (2.96) przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : r(\hat{\theta}_N, Y_N; U_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, Y_N; U_N). \quad (2.99)$$

2.2.3. Estymacja parametrów charakterystyki statycznej obiektu ze zmiennymi niemierzalnymi wielkościami losowymi na podstawie zakłóconych pomiarów wyjścia

Teraz przeanalizujemy zadanie estymacji parametrów charakterystyki statycznej obiektu z losowo zmiennymi niemierzalnymi wielkościami oraz zakłóconym pomiarem wartości wyjścia (rys. 2.12).



Rys. 2.12. Zakłócony pomiar wartości wyjścia obiektu z losowo zmiennymi niemierzalnymi wielkościami

Jest to przypadek, który łączy zadania przedstawione w punktach 2.2.1 oraz 2.2.2. Charakterystyka statyczna obiektu o wektorze wejść u i wektorze wyjść y jest znana z dokładnością do wektora parametrów θ . Na obiekt działają pewne niemierzalne wielkości losowe ω . Obiekt jest opisany zależnością (2.83). Znany jest opis systemu pomiarowego. Sposób nakładania się zakłóceń na mierzoną wartość wyjścia jest dany zależnością (2.20). Podobnie jak w punktach 2.2.1 i 2.2.2, w celu uproszczenia zapisu, przyjęto, że $\dim y = \dim \omega = \dim z = L$ oraz funkcja F (2.83) jest wzajemnie jednoznaczna względem z . Rezygnacja z powyższych założeń wymaga dodatkowych przekształceń, przedstawionych w Dodatku. W n -tym pomiarze dla zadanej wartości składowych wektora wejść u_n zarejestrowano wartość w_n , która jest wynikiem zakłóconej obserwacji wartości składowych wektora wyjść y_n . Przyjęto, że zakłócenia z_n w n -tym pomiarze są niezależnymi wartościami zmiennej losowej z , dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa. W dalszych rozważaniach założono, że zmienna losowa z przyjmuje wartości ze zbioru ciągłego \mathcal{Z} , a funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej z , określonej na \mathcal{Z} , oznaczono przez $f_z(z)$. Losowo zmienny wektor wielkości ω jest wartością zmiennej losowej ω , dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\omega(\omega)$, określona na zbiorze Ω . Zmienne losowe ω oraz z są niezależne. W wyniku eksperymentu, dla zadanej serii identyfikującej o długości N , uzyskano wyniki pomiarów wartości składowych wektora wyjść, które zebrano w macierzach U_N, W_N (2.24). Zadanie identyfikacji sprowadza się do wyznaczenia oszacowania wektora nieznanych parametrów θ charakterystyki statycznej obiektu.

Poszukujemy algorytmu estymacji

$$\hat{\theta}_N = \Psi_N(U_N, W_N), \quad (2.100)$$

który dla zebranych danych pomiarowych w macierzach U_N, W_N umożliwi wyznaczenie wartości oszacowania $\hat{\theta}_N$ wektora θ , gdzie Ψ_N .

Podobnie jak poprzednio, w zależności od informacji, jakie posiadamy o obiekcie i systemie pomiarowym, oraz od warunków, jakie spełniają zakłócenia, możemy stosować różne metody estymacji.

Metoda najmniejszych kwadratów

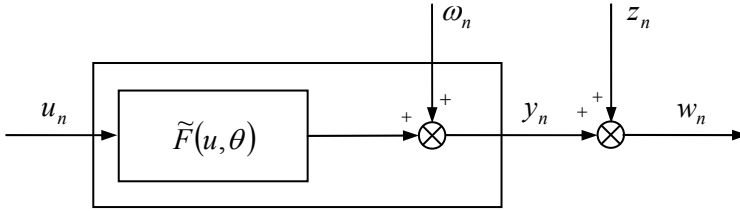
Podobnie jak w punkcie 2.2.1, metodę najmniejszych kwadratów przedstawimy dla obiektu o jednym wyjściu ($\dim y = \dim z = L = 1$).

Metoda najmniejszych kwadratów może być zastosowana jedynie do obiektu, którego charakterystyka statyczna ma postać (2.87) oraz zakłócenia są addytywne (rys. 2.13).

Zakłócenia z i losowo zmienna wielkość są wartościami zmiennych losowych – odpowiednio – z i ω , które spełniają warunki (2.26) oraz (2.88). Wyniki pomiarów wartości składowych wektora wyjść obiektu są wartościami zmiennej losowej

$$w = \tilde{F}(u, \theta) + \omega + z, \quad (2.101)$$

w której można wyróżnić część deterministyczną oraz przypadkową.



Rys. 2.13. Obiekt identyfikacji z addytywną, niemierzalną, zmienną wielkością losową oraz addytywnymi zakłóceniami

Część losowa jest sumą wartości losowej występującej w obiekcie i losowych zakłóceń. Korzystając z warunków (2.26) oraz (2.88), zauważmy, że suma zmiennych losowych $\underline{\omega}$ i \underline{z} spełnia warunek

$$\mathbb{E} [\underline{\omega} + \underline{z}] = 0, \quad \text{Var} [\underline{\omega} + \underline{z}] < \infty. \quad (2.102)$$

Jest to przypadek analogiczny do przedstawionego w punkcie 2.2.1, z tym że teraz losowym składnikiem jest suma zakłóceń i niemierzalnej, zmiennej wielkości losowej. Wyznaczenie oszacowania (2.100) sprowadza się do wyznaczenia takiej wartości $\hat{\theta}_N$, która minimalizuje empiryczną wariancję sumy błędów pomiarowego i losowego parametru. Empiryczna wariancja $\text{Var}_{(\omega+z)N}$, wyznaczona na podstawie zbioru N obserwacji, zależy od macierzy pomiarów U_N, W_N oraz wektora parametrów charakterystyki statycznej obiektu θ i wyraża się wzorem

$$\text{Var}_{(\omega+z)N}(U_N, W_N, \theta) \stackrel{\text{df}}{=} \frac{1}{N} \sum_{n=1}^N (w_n - \tilde{F}(u_n, \theta))^2. \quad (2.103)$$

Wyznaczenie oszacowania **metodą najmniejszych kwadratów** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, W_N)$, dla której wyrażenie (2.103) przyjmuje wartość minimalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : \text{Var}_{(\omega+z)N}(U_N, W_N, \hat{\theta}_N) = \min_{\theta \in \Theta} \text{Var}_{(\omega+z)N}(U_N, W_N, \theta). \quad (2.104)$$

Metoda maksymalnej wiarygodności

Teraz założmy, że wartości wektora wielkości losowych ω w obiekcie opisanym charakterystyką (2.83) w n -tym pomiarze są niezależnymi wartościami ω_n L -wymiarowej zmiennej losowej $\underline{\omega}$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\omega(\omega)$ określona na zbiorze Ω . Dodatkowo na mierzoną wartość

składowych wektora wyjść nakładają się zakłócenia opisane zależnością (2.20) i w n -tym pomiarze zakłócenia z_n są niezależnymi wartościami ciągłej zmiennej losowej \underline{z} o znanej funkcji gęstości rozkładu prawdopodobieństwa $f_z(z)$ określonej na zbiorze $\underline{\mathcal{Z}}$.

Zbiór wartości y_n , $n=1, 2, \dots, N$, możemy potraktować jako zbiór niezależnych wartości zmiennej losowej L -wymiarowej \underline{y} , która jest wynikiem przekształcenia zmiennej losowej $\underline{\omega}$ zgodnie z zależnością (2.83), czyli

$$\underline{y} = F(\underline{u}, \theta, \underline{\omega}), \quad (2.105)$$

a zbiór wartości w_n , $n=1, 2, \dots, N$, jest zbiorem realizacji L -wymiarowej zmiennej losowej \underline{w} , która jest wynikiem przekształcenia zmiennych losowych \underline{y} oraz \underline{z} zgodnie z zależnością (2.20), czyli

$$\underline{w} = h(\underline{y}, \underline{z}). \quad (2.106)$$

Funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{w} należy wyznaczyć, korzystając z przekształcenia (2.106), z uwzględnieniem (2.105). Zwróćmy uwagę, że wektor wejść \underline{u} i wektor θ są parametrami tego przekształcenia. Dla ustalonego wejścia \underline{u} funkcję gęstości rozkładu prawdopodobieństwa $f_y(y, \theta; \underline{u})$ zmiennej losowej \underline{y} wyznaczamy według wzoru (2.92). Przekształcenie h (2.106) nie jest wzajemnie jednoznaczne ze względu na parę zmiennych losowych \underline{y} oraz \underline{z} . Możemy je uczynić wzajemnie jednoznaczными, dopisując do (2.106) tożsamość $\underline{y} = \underline{y}$, i tak otrzymamy

$$\begin{aligned} \underline{w} &= h(\underline{y}, \underline{z}), \\ \underline{y} &= \underline{y}. \end{aligned} \quad (2.107)$$

Przekształcenie odwrotne przyjmuje postać

$$\begin{aligned} \underline{z} &= h_z^{-1}(\underline{y}, \underline{w}), \\ \underline{y} &= \underline{y}, \end{aligned} \quad (1.108)$$

a jacobian tego przekształcenia

$$J = \begin{bmatrix} \frac{\partial h_z^{-1}(\underline{y}, \underline{w})}{\partial \underline{w}} & \frac{\partial h_z^{-1}(\underline{y}, \underline{w})}{\partial \underline{y}} \\ \frac{\partial \underline{y}}{\partial \underline{w}} & \frac{\partial \underline{y}}{\partial \underline{y}} \end{bmatrix} = \begin{bmatrix} \frac{\partial h_z^{-1}(\underline{y}, \underline{w})}{\partial \underline{w}} & \frac{\partial h_z^{-1}(\underline{y}, \underline{w})}{\partial \underline{y}} \\ O_{L \times L} & I_{L \times L} \end{bmatrix}, \quad (2.109)$$

gdzie: $O_{L \times L}$ jest L -wymiarową macierzą zerową, $I_{L \times L}$ jest L -wymiarową macierzą jednostkową.

Wyznacznik jacobianu (2.109)

$$|J| = \left| \begin{bmatrix} \frac{\partial h_z^{-1}(y, w)}{\partial w} & \frac{\partial h_z^{-1}(y, w)}{\partial y} \\ O_{L \times L} & I_{L \times L} \end{bmatrix} \right| = \left| \frac{\partial h_z^{-1}(y, w)}{\partial w} \right| = |J_h|. \quad (2.110)$$

Ostatecznie funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej w ma postać

$$f_w(w, \theta; u) = \int_Y f_z(h_z^{-1}(y, w)) f_y(y, \theta; u) |J| dy, \quad (2.111)$$

a po uwzględnieniu (2.92) oraz (2.110) otrzymujemy

$$f_w(w, \theta; u) = \int_Y f_z(h_z^{-1}(y, w)) f_\omega(F_\omega^{-1}(u, \theta, y)) |J_F| |J_h| dy. \quad (2.112)$$

Oszacowanie nieznanego wektora parametrów θ charakterystyki statycznej (2.83) otrzymamy w wyniku maksymalizacji funkcji wiarygodności

$$\begin{aligned} L_N(W_N, \theta; U_N) &\stackrel{\text{df}}{=} \prod_{n=1}^N f_w(w_n, \theta; u_n) \\ &= \prod_{n=1}^N \int_Y f_z(h_z^{-1}(y, w_n)) f_\omega(F_\omega^{-1}(u_n, \theta, y)) |J_F| |J_h| dy \end{aligned} \quad (2.113)$$

względem θ .

W tym przypadku zadanie wyznaczenia oszacowania **metodą maksymalnej wiarygodności** sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(U_N, Y_N)$, dla której wyrażenie (2.113) przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : L_N(W_N, \hat{\theta}_N; U_N) = \max_{\theta \in \Theta} L_N(W_N, \theta; U_N). \quad (2.114)$$

Metody Bayesa

Założenia dotyczące obiektu są identyczne z podanymi w punkcie 2.2.2, dotyczącymi metod Bayesa. Znaczy to, że wektor parametrów θ charakterystyki statycznej (2.83) jest wartością ciągłej zmiennej losowej $\underline{\theta}$, dla której istnieje funkcja gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$. Teraz wektory parametrów θ oraz losowo zmiennych wielkości ω (2.83) są wartościami R -, L -wymiarowych zmiennych losowych – odpowiednio – $\underline{\theta}$ i $\underline{\omega}$. Istotna różnica pomiędzy tymi wektorami polega na tym, że z chwilą wylosowania obiektu do badań wektor parametrów θ jest stały, na-

tomiast wektor ω zmienia się w sposób losowy. Dalej zakładamy, że znane są funkcje gęstości rozkładów prawdopodobieństwa zmiennych losowych $\underline{\theta}$ oraz $\underline{\omega}$. W odróżnieniu od rozważań z punktu 2.2.2, na mierzoną wartość wyjścia nakładają się zakłócenia opisane zależnością (2.20) i w kolejnych pomiarach zakłócenia z_n są niezależnymi wartościami ciągłej zmiennej losowej \underline{z} o znanej funkcji gęstości rozkładu prawdopodobieństwa $f_z(z)$. Zbiór wartości wyjść obiektu y_n , $n=1, 2, \dots, N$, należy traktować jako zbiór niezależnych wartości warunkowej zmiennej losowej \underline{y} , pod warunkiem, że zmienna losowa $\underline{\theta}$ przyjęła wartość θ , a wartość składowych wektora na wejściu wynosi u_n . Funkcję gęstości rozkładu prawdopodobieństwa $f_y(y|\theta; u)$ wyznaczamy według wzoru (2.97). Zaobserwowany zbiór wartości w_n , $n=1, 2, \dots, N$, jest zbiorem wartości warunkowej zmiennej losowej \underline{w} , pod warunkiem, że zmienna losowa $\underline{\theta}$ przyjęła wartość θ , a wartość składowych wektora wejść wynosi u_n . Jest ona wynikiem przekształcenia warunkowej zmiennej losowej \underline{y} pod warunkiem, że zmienna losowa $\underline{\theta}$ przyjęła wartość θ , oraz zmiennej \underline{z} zgodnie z zależnością (2.20). Funkcję gęstości rozkładu prawdopodobieństwa $f_w(w|\theta; u)$ wyznaczamy tak jak $f_w(w, \theta; u)$, z tym że w zależności (2.111) w miejsce funkcji gęstości $f_y(y, \theta; u)$ wstawiamy funkcję gęstości warunkowej $f_y(y|\theta; u)$, czyli

$$f_w(w|\theta; u) = \int_{\underline{Y}} f_z(h_z^{-1}(y, w)) f_y(y|\theta; u) |J| dy, \quad (2.115)$$

a po uwzględnieniu równania (2.97) mamy

$$f_w(w|\theta; u) = \int_{\underline{Y}} f_z(h_z^{-1}(y, w)) f_\omega(F_\omega^{-1}(u, \theta, y)) |J_F| |J_h| dy. \quad (2.116)$$

Ostatecznie funkcja gęstości rozkładu prawdopodobieństwa a posteriori $f'(\theta|Y_N; U_N)$ jest dana zależnością

$$f'(\theta|W_N; U_N) = \frac{\int_{\underline{Y}} f_\theta(\theta) \prod_{n=1}^N \int f_z(h_z^{-1}(y, w_n)) f_\omega(F_\omega^{-1}(u_n, \theta, y)) |J_F| |J_h| dy}{\int_{\Theta} \int_{\underline{Y}} f_\theta(\theta) \prod_{n=1}^N \int f_z(h_z^{-1}(y, w_n)) f_\omega(F_\omega^{-1}(u_n, \theta, y)) |J_F| |J_h| dy d\theta}. \quad (2.117)$$

Wyznaczenie oszacowania sprowadza się do wyznaczenia takiej wartości $\hat{\theta}_N = \mathcal{P}_N(U_N, Y_N)$, dla której ryzyko warunkowe

$$r(\bar{\theta}, Y_N; U_N) \stackrel{\text{def}}{=} E_{\underline{\theta}} [L(\underline{\theta}, \bar{\theta}) | Y_N; U_N] = \int_{\Theta} L(\theta, \bar{\theta} = \bar{\Psi}(U_N, Y_N)) f'(\theta | Y_N; U_N) d\theta \quad (2.118)$$

przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(U_N, Y_N) : r(\hat{\theta}_N, Y_N; U_N) = \min_{\theta \in \Theta} r(\bar{\theta}, Y_N; U_N). \quad (2.119)$$

2.3. Zadanie wyboru optymalnego modelu w warunkach deterministycznych

Obecnie skupimy się na przypadku, gdy proponowany opis (model) różni się od badanej rzeczywistości. Różne mogą być tego przyczyny. Jedną z nich to fakt, że analiza zjawisk fizykochemicznych nie prowadzi do określenia charakterystyki statycznej obiektu identyfikacji. Postać charakterystyki statycznej nie jest znana. Twórca modelu, posilkując się wynikami wstępnych danych pomiarowych, narzuca pewien opis, arbitralnie wybierając klasę modeli. W innym przypadku analiza zjawisk może prowadzić do nieliniowego opisu, a taki może być nieużyteczny. Do dalszych zastosowań należy opis uprościć, wówczas, pomimo znajomości charakterystyki statycznej obiektu, poszukujemy uproszczonego, przybliżonego modelu. W przypadku znajomości charakterystyki statycznej obiektu wyznaczenie uproszczonego modelu prowadzi do zadania aproksymacji. Gdy charakterystyka statyczna obiektu nie jest znana, ale w zamian dysponujemy wynikami pomiarów wejścia i wyjścia obiektu, zadanie sprowadza się do najlepszego przybliżenia wyników pomiarów modelem z zadanej klasy modeli. Mówimy wówczas o zadaniu wyboru optymalnego modelu [14].

W tym miejscu należy skomentować sformułowanie „twórca modelu narzuca pewien opis”. Za tym sformułowaniem kryje się metodyka wyboru klasy modeli spośród szerokiej gamy uniwersalnych funkcji aproksymujących [1, 8, 29], popartych doбором struktury modelu.

Zanim sformulujemy zadanie wyboru optymalnego modelu przytoczymy zadanie aproksymacji. Jest ono bardzo podobne do zadania wyboru optymalnego modelu, z tą różnicą, że w zadaniu aproksymacji charakterystyka statyczna obiektu jest znana, a w zadaniu identyfikacji dysponujemy tylko pomiarami wartości wejścia i wyjścia obiektu, czyli wybranymi punktami charakterystyki statycznej.

2.3.1. Aproksymacja charakterystyki statycznej

Teraz rozważymy przypadek, kiedy charakterystyka statyczna obiektu jest znana i ma postać

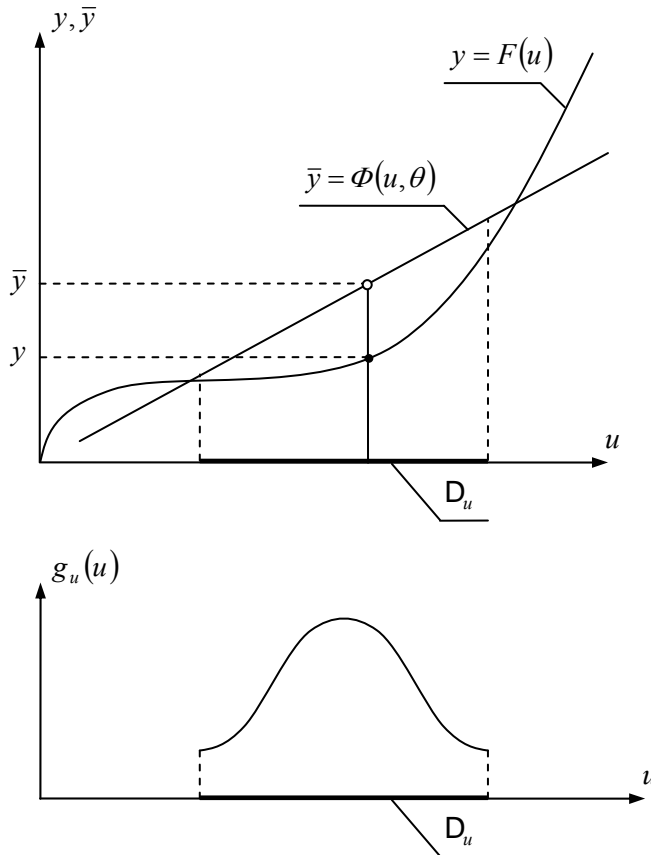
$$y = F(u), \quad (2.120)$$

gdzie: u – wektor wejść obiektu, $u \in U \subseteq \mathbb{R}^S$, U – przestrzeń wejść, y – wektor wyjść obiektu, $y \in Y \subseteq \mathbb{R}^L$, Y – przestrzeń wyjść.

Charakterystykę (2.120) proponuje się aproksymować charakterystyką

$$\bar{y} = \Phi(u, \theta), \quad (2.121)$$

gdzie: Φ jest przyjętą przez twórcę modelu funkcją opisującą zależność wektora wyjść modelu \bar{y} od wektora wejść u , natomiast θ jest wektorem parametrów modelu (rys. 2.14).



Rys. 2.14. Charakterystyka statyczna $F(u)$, funkcja aproksymująca $\Phi(u, a)$, funkcja wagi jakości przybliżenia $g_u(u)$ dla jednowymiarowego obiektu ($L = S = 1$)

Wektory u , \bar{y} i θ są elementami – odpowiednio – przestrzeni wejść \mathcal{U} , wyjść \mathcal{Y} oraz parametrów Θ , czyli: $u \in \mathcal{U} \subseteq \mathbb{R}^S$, $\bar{y} \in \mathcal{Y} \subseteq \mathbb{R}^L$ oraz $\theta \in \Theta \subseteq \mathbb{R}^R$. Często funkcja $F(u)$ jest nazywana funkcją aproksymowaną, a funkcja $\Phi(u, \theta)$ – funkcją aproksymującą. Dalej zakładamy, że aproksymacji dokonujemy w pewnym podzbio-

rze D_u przestrzeni wejść, tj. $D_u \subseteq U \subseteq \mathbb{R}^S$. Na zbiorze D_u jest określona funkcja wagi jakości przybliżenia $g_u(u)$. Funkcja ta nadaje wagę jakości przybliżenia w poszczególnych punktach zbioru D_u (rys. 2.14). W celu porównania funkcji aproksymowanej z funkcją aproksymującą, dla każdego punktu zbioru D_u określamy miarę różnicy pomiędzy wartościami funkcji w tym punkcie, czyli

$$\forall u \in D_u \quad q(y, \bar{y}) = q(F(u), \Phi(u, \theta)), \quad (2.122)$$

gdzie q jest funkcją, której wartość jest oceną różnicy pomiędzy wartością funkcji aproksymowanej $F(u)$ a wartością funkcji aproksymującej $\Phi(u, \theta)$ w każdym punkcie zbioru D_u .

Funkcja q może być dowolną funkcją spełniającą własności funkcji miary. Przykładowe funkcje q mają postać:

$$q(y, \bar{y}) = \sum_{l=1}^L (y^{(l)} - \bar{y}^{(l)})^2 \quad (2.123)$$

lub

$$q(y, \bar{y}) = \sum_{l=1}^L |y^{(l)} - \bar{y}^{(l)}|, \quad (2.124)$$

a dla obiektów jednowyjściowych ocena (2.123), (2.124) ma postać, odpowiednio:

$$q(y, \bar{y}) = (y - \bar{y})^2 = (F(u) - \Phi(u, \theta))^2, \quad (2.125)$$

$$q(y, \bar{y}) = |y - \bar{y}| = |F(u) - \Phi(u, \theta)|. \quad (2.126)$$

Jako miarę jakości przybliżenia funkcji $F(u)$ funkcją $\Phi(u, \theta)$ w zbiorze D_u wprowadzamy miarę w przestrzeni funkcyjnej, określonej na zbiorze D_u

$$Q(\theta) = \|F(u) - \Phi(u, \theta)\|_{D_u}. \quad (2.127)$$

Przykładowe funkcje $Q(\theta)$ mają postać:

$$Q(\theta) = \int_{D_u} q(F(u), \Phi(u, \theta)) g_u(u) du \quad (2.128)$$

lub

$$Q(\theta) = \max_{u \in D_u} \{q(F(u), \Phi(u, \theta)) g_u(u)\}. \quad (2.129)$$

Funkcja Q jest miarą jakości przybliżenia, zależy od parametrów θ i umożliwia wybór najlepszego przybliżenia. Funkcja $Q(\theta)$ stanowi kryterium wyboru optymalnej

funkcji aproksymującej, a konkretnie jej parametrów. Problem wyznaczenia optymalnej funkcji aproksymującej, z danej klasy funkcji, sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość wektora parametrów θ^* , dla której miara jakości przybliżenia $Q(\theta)$ przyjmuje wartość minimalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\theta^* : Q(\theta^*) = \min_{\theta \in \Theta} Q(\theta), \quad (2.130)$$

gdzie: θ^* jest optymalną wartością wektora parametrów funkcji aproksymującej, a funkcja (2.121) z optymalnymi parametrami θ^* , tj.

$$\bar{y} = \Phi(u, \theta^*), \quad (2.131)$$

jest optymalną funkcją aproksymującą z danej klasy funkcji.

Dla obiektów jednowymiarowych ($L = S = 1$), po przyjęciu konkretnej postaci miary różnicy pomiędzy funkcją aproksymowaną i aproksymującą (2.122), wartości kryterium $Q(\theta)$ można nadać interpretację geometryczną. I tak, dla $q(y, \bar{y})$ postaci (2.126) i kryterium (2.128), tj.

$$Q(\theta) = \int_{D_u} |F(u) - \Phi(u, \theta)| g_u(u) du, \quad (2.132)$$

w przypadku stałej funkcji wagi jakości przybliżenia $g_u(u)$, miarą różnicy pomiędzy funkcją aproksymowaną a funkcją aproksymującą jest pole powierzchni pomiędzy $F(u)$ i $\Phi(u, \theta)$ nad zbiorem D_u . Dla dowolnej funkcji $g_u(u)$ pole to jest modyfikowane zgodnie z zależnością (2.132). W tym przypadku wybór optymalnych parametrów funkcji aproksymującej sprowadza się do minimalizacji zmodyfikowanego przez $g_u(u)$ pola powierzchni pomiędzy $F(u)$ i $\Phi(u, \theta)$ nad zbiorem D_u . Dla $Q(\theta)$ postaci (2.129) i $q(y, \bar{y})$ (2.126), tj.

$$Q(\theta) = \max_{u \in D_u} \{|F(u) - \Phi(u, \theta)| g_u(u)\}, \quad (2.133)$$

oceną jakości przybliżenia jest największa różnica pomiędzy $F(u)$ i $\Phi(u, \theta)$ w zbiorze D_u , skorygowana wartością funkcji wagi $g_u(u)$, a wybór optymalnej charakterystyki sprowadza się do minimalizacji – skorygowanej funkcją wagi $g_u(u)$ – największej różnicy pomiędzy funkcją aproksymowaną i aproksymującą.

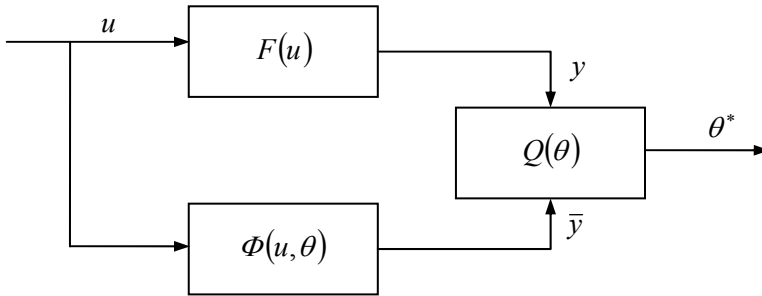
Ze względu na liczne zastosowania jednym z ważniejszych przypadków dla obiektów o jednowymiarowym wyjściu ($L = 1$) jest funkcja aproksymująca postaci

$$\bar{y} = \Phi(u, \theta) = \theta^T \varphi(u), \quad (2.134)$$

gdzie: θ^T jest transpozycją R -wymiarowego wektora parametrów θ , natomiast

$$\varphi(u) \stackrel{\text{df}}{=} \begin{bmatrix} \varphi_1(u) \\ \varphi_2(u) \\ \vdots \\ \varphi_R(u) \end{bmatrix} \quad (2.135)$$

jest R -wymiarowym wektorem funkcji, stanowiącym jądro przybliżenia.



Rys. 2.15. Aproksymacja charakterystyki statycznej

Odpowiedni dobór funkcji $\varphi_1(u), \varphi_2(u), \dots, \varphi_R(u)$ prowadzi do różnych funkcji aproksymujących. Po przyjęciu

$$\varphi(u) = \begin{bmatrix} 1 \\ u \\ \vdots \\ u^{R-1} \end{bmatrix} \quad (2.136)$$

otrzymujemy aproksymację za pomocą wielomianów.

Przyjmujemy kryterium jakości przybliżenia (2.128) postaci

$$Q(\theta) = \int_{D_u} (F(u) - \Phi(u, \theta))^2 g_u(u) du. \quad (2.137)$$

Po podstawieniu do (2.137) funkcji aproksymującej postaci (2.134) otrzymujemy

$$Q(\theta) = \int_{D_u} (F(u) - \theta^T \varphi(u))^2 g_u(u) du. \quad (2.138)$$

Jest to tak zwany przypadek liniowo-kwadratowy, w którym funkcja aproksymująca postaci (2.134) jest liniowa względem wektora parametrów, a w kryterium (2.138) ocena jakości przybliżenia jest kwadratem różnicy pomiędzy funkcją aproksymującą i aproksymowaną.

Optymalny wektor parametrów θ^* funkcji aproksymującej spełnia układ równań

$$\operatorname{grad}_{\theta} Q(\theta) = -2 \left[\int_{D_u} (F(u) - \theta^{*T} \varphi(u)) \varphi(u) g_u(u) du \right] = \mathbf{0}_R, \quad (2.139)$$

gdzie $\mathbf{0}_R$ jest R -wymiarowym wektorem zerowym.

Po rozwiązaniu powyższego układu względem θ^* otrzymujemy algorytm

$$\theta^* = \left[\int_{D_u} \varphi(u) \varphi^T(u) g_u(u) du \right]^{-1} \int_{D_u} F(u) \varphi(u) g_u(u) du. \quad (2.140)$$

Zwróćmy uwagę na postać macierzy $\int_{D_u} \varphi(u) \varphi^T(u) g_u(u) du$ w równaniu (2.140) po podstawieniu (2.135), czyli

$$\begin{aligned} & \int_{D_u} \varphi(u) \varphi^T(u) g_u(u) du \\ &= \begin{bmatrix} \int_{D_u} \varphi_1(u) \varphi_1(u) g_u(u) du & \int_{D_u} \varphi_1(u) \varphi_2(u) g_u(u) du & \cdots & \int_{D_u} \varphi_1(u) \varphi_R(u) g_u(u) du \\ \int_{D_u} \varphi_2(u) \varphi_1(u) g_u(u) du & \int_{D_u} \varphi_2(u) \varphi_2(u) g_u(u) du & \cdots & \int_{D_u} \varphi_2(u) \varphi_R(u) g_u(u) du \\ \vdots & \vdots & \ddots & \vdots \\ \int_{D_u} \varphi_R(u) \varphi_1(u) g_u(u) du & \int_{D_u} \varphi_R(u) \varphi_2(u) g_u(u) du & \cdots & \int_{D_u} \varphi_R(u) \varphi_R(u) g_u(u) du \end{bmatrix}. \end{aligned} \quad (2.141)$$

W algorytmie (2.140) należy wyznaczyć odwrotność macierzy (2.141). Operacja ta ulega znacznemu uproszczeniu po odpowiednim doborze funkcji jądra przybliżenia $\varphi(u)$.

Przyjmujemy, że zbiór funkcji $\{\varphi_1(u), \varphi_2(u), \dots, \varphi_R(u)\}$ tworzy układ ortonormalny, czyli spełnia warunek

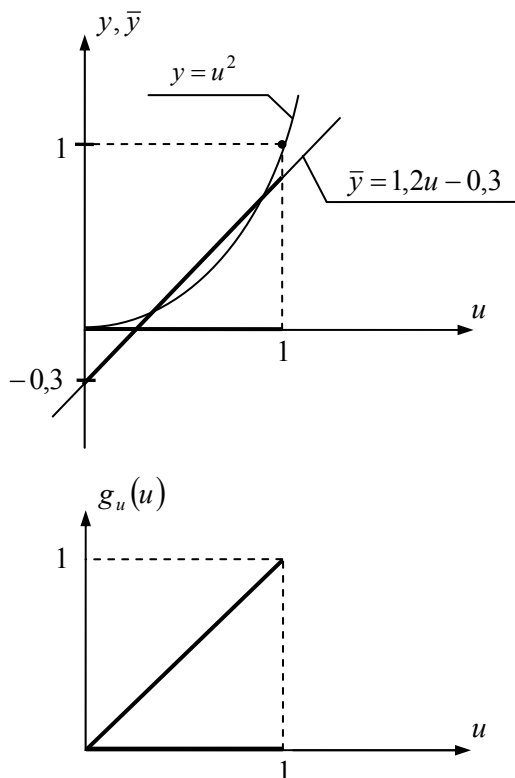
$$\int_{D_u} \varphi_i(u) \varphi_j(u) g_u(u) du = \begin{cases} 1 & \text{dla } i = j \\ 0 & \text{dla } i \neq j \end{cases}, \quad i, j = 1, 2, \dots, R, \quad (2.142)$$

wówczas $\int_{D_u} \varphi(u) \varphi^T(u) g_u(u) du$ (2.141) jest macierzą jednostkową, a algorytm aproksymacji (2.140) przyjmuje postać

$$\theta^* = \int_{D_u} F(u) \varphi(u) g_u(u) du. \quad (2.143)$$

Przykład 2.6. Charakterystykę statyczną obiektu $y = F(u) = u^2$ należy przybliżyć funkcją $\bar{y} = \Phi(u, \theta) = \theta^T \varphi(u) = \theta^{(1)} + \theta^{(2)}u$, w której: $\varphi(u) = \begin{bmatrix} 1 \\ u \end{bmatrix}$, $\theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \end{bmatrix}$, w przedziale $D_u = \{u \in \mathbb{R} : 0 \leq u \leq 1\}$ z funkcją wagi w postaci

$$g_u(u) = \begin{cases} u & \text{dla } u \in [0, 1] \\ 0 & \text{dla } u \notin [0, 1] \end{cases}$$



Rys. 2.16. Przykład aproksymacji charakterystyki statycznej

W rozpatrywanym przypadku kryterium jakości przybliżenia (2.138) ma postać

$$Q(\theta) = \int_{D_u} (F(u) - \theta^T \varphi(u))^2 g_u(u) du = \int_0^1 \left(u^2 - [\theta^{(1)} \ \theta^{(2)}] \begin{bmatrix} 1 \\ u \end{bmatrix} \right)^2 u du, \quad (2.144)$$

równanie (2.139)

$$\operatorname{grad}_{\theta} Q(\theta) = -2 \int_0^1 \left(u^2 - [\theta^{(1)*} \ \theta^{(2)*}] \begin{bmatrix} 1 \\ u \end{bmatrix} \right) \begin{bmatrix} 1 \\ u \end{bmatrix} u \, du = \mathbf{0}_R, \quad (2.145)$$

a jego rozwiązanie (2.140)

$$\theta^* = \left[\int_0^1 \begin{bmatrix} 1 \\ u \end{bmatrix} \begin{bmatrix} 1 & u \end{bmatrix} u \, du \right]^{-1} \int_0^1 u^2 \begin{bmatrix} 1 \\ u \end{bmatrix} u \, du = \begin{bmatrix} \int_0^1 u \, du & \int_0^1 u^2 \, du \\ \int_0^1 u^2 \, du & \int_0^1 u^3 \, du \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 u^3 \, du \\ \int_0^1 u^4 \, du \end{bmatrix}. \quad (2.146)$$

Po prostych przekształceniach algebraicznych otrzymujemy

$$\theta^* = \begin{bmatrix} \theta^{(1)*} \\ \theta^{(2)*} \end{bmatrix} = \begin{bmatrix} \frac{\int_0^1 u^3 \, du \int_0^1 u^3 \, du - \int_0^1 u^4 \, du \int_0^1 u^2 \, du}{\left(\int_0^1 u \, du \int_0^1 u^3 \, du - \left(\int_0^1 u^2 \, du \right)^2 \right)} \\ \frac{\int_0^1 u^4 \, du \int_0^1 u \, du - \int_0^1 u^3 \, du \int_0^1 u^2 \, du}{\int_0^1 u \, du \int_0^1 u^3 \, du - \left(\int_0^1 u^2 \, du \right)^2} \end{bmatrix}, \quad (2.147)$$

a po obliczeniu odpowiednich całek mamy

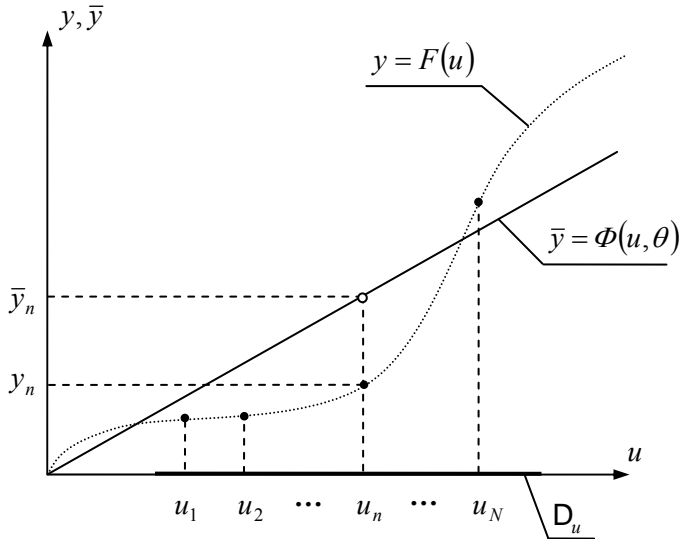
$$\theta^* = \begin{bmatrix} \theta^{(1)*} \\ \theta^{(2)*} \end{bmatrix} = \begin{bmatrix} \frac{\frac{1}{4}u^4 \Big|_0^1 \frac{1}{4}u^4 \Big|_0^1 - \frac{1}{5}u^5 \Big|_0^1 \frac{1}{3}u^3 \Big|_0^1}{\frac{1}{2}u^2 \Big|_0^1 \frac{1}{4}u^4 \Big|_0^1 - \left(\frac{1}{3}u^3 \Big|_0^1 \right)^2} \\ \frac{\frac{1}{5}u^5 \Big|_0^1 \frac{1}{2}u^2 \Big|_0^1 - \frac{1}{4}u^4 \Big|_0^1 \frac{1}{3}u^3 \Big|_0^1}{\frac{1}{2}u^2 \Big|_0^1 \frac{1}{4}u^4 \Big|_0^1 - \left(\frac{1}{3}u^3 \Big|_0^1 \right)^2} \end{bmatrix} = \begin{bmatrix} \frac{\frac{1}{4} \frac{1}{4} - \frac{1}{5} \frac{1}{3}}{\frac{1}{2} \frac{1}{4} - \left(\frac{1}{3} \right)^2} \\ \frac{\frac{1}{5} \frac{1}{2} - \frac{1}{4} \frac{1}{3}}{\frac{1}{2} \frac{1}{4} - \left(\frac{1}{3} \right)^2} \end{bmatrix} = \begin{bmatrix} -0,3 \\ 1,2 \end{bmatrix}. \quad (2.148)$$

Optymalna funkcja aproksymująca z klasy $\bar{y} = \theta^{(1)}u + \theta^{(2)}$ ma postać:
 $\bar{y} = 1,2u - 0,3$ (rys. 2.16).



2.3.2. Zadanie wyboru optymalnego modelu na podstawie ciągu niezakłóconych obserwacji

Teraz rozważymy przypadek, kiedy dokładna charakterystyka statyczna obiektu $F(u)$ istnieje i ma postać (2.120), ale nie jest znana. W zamian dysponujemy wynikami eksperymentu, w którym dla zadanej serii wejść ze zbioru D_u zmierzono, bez zakłóceń, wartość składowych wektora wyjść badanego obiektu (rys. 2.17).



Rys. 2.17. Badanie nieznanej charakterystyki statycznej jednowymiarowego obiektu identyfikacji ($L = S = 1$)

W wyniku eksperymentu dla zadanej serii identyfikującej, o zadanej długości N , dokonano pomiaru wartości składowych wektora wyjść. Wyniki przedstawiono w macierzach:

$$U_N = [u_1 \ u_2 \ \dots \ u_N], \quad Y_N = [y_1 \ y_2 \ \dots \ y_N], \quad (2.149)$$

których kolumnami są – odpowiednio – kolejne pomiary wejścia i wyjścia obiektu.

Twórca modelu proponuje przybliżyć wyniki pomiarów funkcją $\Phi(u, \theta)$ (2.121). Sytuacja jest analogiczna do opisanej w punkcie 2.3.1, z tą różnicą, że zamiast pełnej znajomości charakterystyki statycznej $F(u)$ w zbiorze D_u dysponujemy serią pomiarów w wybranych punktach tego zbioru, czyli serią: $u_n, y_n = F(u_n)$, $n = 1, 2, \dots, N$.

W celu porównania wyników pomiarów charakterystyki statycznej badanego obiektu z proponowanym modelem, dla każdego punktu pomiarowego określamy miarę różnicy pomiędzy wartością wyjścia obiektu i wartością wyjścia wyznaczoną z modelu, czyli

$$\forall n = 1, 2, \dots, N \quad q(y_n, \bar{y}_n) = q(y_n, \Phi(u_n, \theta)). \quad (2.150)$$

gdzie q (analogicznie do (2.122)) jest funkcją, której wartość ocenia różnicę pomiędzy zmierzonymi wartościami składowych wektora wyjść obiektu a wartościami składowych wektora wyjść modelu \bar{y}_n , wyznaczonymi dla wektora wejść u_n , czyli $\bar{y}_n = \Phi(u_n, \theta)$.

Jako miarę oceny jakości przybliżenia wyników pomiarów wartości składowych wektora wyjść obiektu zebranych w macierzy Y_N dla zadanej serii identyfikującej U_N z wartościami składowych wektora wyjść modelu zawartych w macierzy $\bar{Y}_N(\theta)$, wyznaczonych dla zadanej serii identyfikującej U_N , wprowadzamy

$$Q_N(\theta) = \|Y_N - \bar{Y}_N(\theta)\|_{U_N}, \quad (2.151)$$

gdzie

$$\bar{Y}_N(\theta) \stackrel{\text{df}}{=} [\Phi(u_1, \theta) \quad \Phi(u_2, \theta) \quad \dots \quad \Phi(u_N, \theta)]. \quad (2.152)$$

Jest to kryterium jakości przybliżenia, często zwanym kryterium jakości identyfikacji. Przykładowe funkcje $Q_N(\theta)$ (analogiczne do (2.128), (2.129)) mają postać, odpowiednio:

$$Q_N(\theta) = \sum_{n=1}^N q(y_n, \bar{y}_n) = \sum_{n=1}^N q(y_n, \Phi(u_n, \theta)) \quad (2.153)$$

lub

$$Q_N(\theta) = \max_{1 \leq n \leq N} \{q(y_n, \bar{y}_n)\} = \max_{1 \leq n \leq N} \{q(y_n, \Phi(u_n, \theta))\}. \quad (2.154)$$

Zwróćmy uwagę, że w kryteriach (2.153) i (2.154) brak jest funkcji wagi jakości przybliżenia $g_u(u)$, która występuje w (2.128) i (2.129). Funkcja ta może być uwzględniona poprzez odpowiednią organizację eksperymentu, a konkretnie – seria identyfikująca powinna być generowana zgodnie z funkcją wagi $g_u(u)$.

Bardzo często stosowaną oceną różnicy pomiędzy wektorem wyjść badanego obiektu a wektorem wyjść modelu jest wartość funkcji $q(y_n, \bar{y}_n)$ postaci (2.123). Przy tej postaci funkcji $q(y_n, \bar{y}_n)$ kryterium (2.153) zwane jest kwadratowym wskaźnikiem jakości identyfikacji. Wyznaczenie optymalnych modeli z wykorzystaniem wskaźnika jakości identyfikacji (2.153) z oceną różnicy pomiędzy wyjściem obiektu i modelu postaci (2.123) zwane jest metodą najmniejszych kwadratów. W tym miejscu zwróćmy uwagę na różnicę pomiędzy zadaniem estymacji parametrów obiektu z wykorzystaniem metody najmniejszych kwadratów (omówioną w punkcie 2.2.1) a zadaniem wyboru optymalnego modelu z kryterium kwadratowym. W pierwszym przypadku mówimy o znajomości charakterystyki statycznej z dokładnością do wektora parametrów. Różnica pomiędzy mierzoną wartością składowych wektora wyjść obiektu a faktyczną wartością składowych wektora wyjść wynika z zakłóceń pomiarowych (rys. 2.5). W drugim przypadku charakterystyka statyczna obiektu nie jest znana. Różnica pomiędzy bezbłędnie mierzoną wartością składowych wektora wyjść obiektu a wartością wyznaczoną z modelu wynika

z nieznanymi charakterystyki statycznej (rys. 2.17). Pomimo że w obu przypadkach otrzymujemy podobne algorytmy identyfikacji, różna jest ich interpretacja. W pierwszym przypadku otrzymujemy oszacowania wektorów parametrów charakterystyk statycznych, które mają fizyczną interpretację. W drugim przypadku, tj. zastosowania kryterium kwadratowego do wyboru optymalnego modelu, uzyskujemy wektor optymalnych parametrów, który nie ma fizycznych interpretacji. Jedyne porównanie, jakiego możemy oczekiwać, to odniesienie wyniku identyfikacji do zadania aproksymacji.

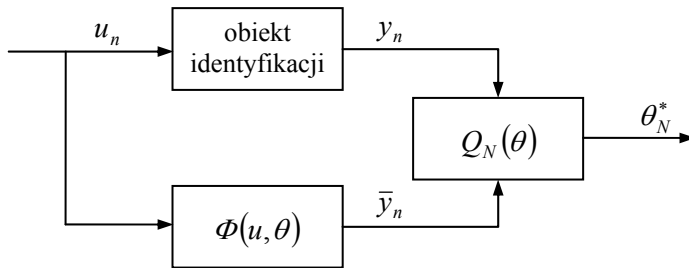
W przypadku obiektów jednowymiarowych ($L = S = 1$), po przyjęciu konkretnej postaci miary różnicy pomiędzy wyjściem obiektu i modelu (2.150), wartości kryterium $Q_N(\theta)$ można nadać interpretację geometryczną. I tak, dla funkcji $q(y_n, \bar{y}_n)$ postaci (2.126) i kryterium (2.153), tj.

$$Q_N(\theta) = \sum_{n=1}^N |y_n - \bar{y}_n| = \sum_{n=1}^N |y_n - \Phi(u_n, \theta)|, \quad (2.155)$$

miarą różnicy pomiędzy zaobserwowanym zbiorem wartości wyjść obiektu Y_N a zbiorem wartości wyjść modelu \bar{Y}_N , dla zadanej serii identyfikującej U_N , jest suma różnic pomiędzy wyjściem obiektu i modelu, czyli suma odcinków o długości $|y_n - \bar{y}_n|$, a dla kryterium (2.154), tj.

$$Q_N(\theta) = \max_{1 \leq n \leq N} \{|y_n - \bar{y}_n|\} = \max_{1 \leq n \leq N} \{|y_n - \Phi(u_n, \theta)|\}, \quad (2.156)$$

jest to największa zaobserwowana różnica pomiędzy wartością wyjścia obiektu a obliczoną wartością wyjścia modelu, czyli najdłuższym odcinkiem $|y_n - \bar{y}_n|$. Funkcja $Q_N(\theta)$ stanowi kryterium wyboru optymalnego modelu, a konkretnie kryterium wyboru jego parametrów.



Rys. 2.18. Zadanie wyboru optymalnego modelu

Problem wyznaczenia optymalnego modelu sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość wektora parametrów θ_N^* , dla której miara jakości przybliżenia $Q_N(\theta)$ przyjmuje wartość minimalną po θ ze zbioru dopuszczalnego Θ , tj.

$$\theta_N^* : Q_N(\theta_N^*) = \min_{\theta \in \Theta} Q_N(\theta), \quad (2.157)$$

gdzie θ_N^* jest optymalną wartością wektora parametrów modelu, a funkcja (2.121) z parametrami θ_N^* , tj.

$$\bar{y} = \Phi(u, \theta_N^*), \quad (2.158)$$

jest nazywana modelem optymalnym z danej klasy modeli.

Należy podkreślić, że model (2.158) jest optymalny dla:

- uzyskanych wyników pomiarów (2.149),
- przyjętej klasy modeli (2.121),
- przyjętego kryterium (2.150) i (2.151).

Znaczy to, że zmiana dowolnego z wyżej wymienionych elementów daje inny, nieporównywalny model. Pozostają pytania: Jaka jest relacja pomiędzy optymalnym modelem (2.158), uzyskanym na podstawie ciągu obserwacji, a optymalną funkcją aproksymującą (2.131), wyznaczoną przy pełnej znajomości charakterystyki statycznej obiektu? Czy po zwiększeniu liczby punktów pomiarowych uzyskamy lepsze przybliżenie θ^* ? W świetle powyższych uwag zwiększenie liczby pomiarów da inny, a nie lepszy model. O porównaniu modeli (2.158) i (2.131) można mówić jedynie wówczas, gdy eksperyment zostanie odpowiednio zaplanowany. Dobierając odpowiednią serię identyfikującą, generowaną zgodnie z funkcją wagi jakości przybliżenia $g_u(u)$, można badać relację pomiędzy θ_N^* (2.157) i θ^* (2.130). Można pokazać, że jeśli seria identyfikująca U_N jest generowana zgodnie z funkcją wagi jakości przybliżenia $g_u(u)$, to przy zwiększającej się liczbie pomiarów N optymalna wartość wektora parametrów θ_N^* , wyznaczona przy kryterium (2.153), jest zbieżna do wartości wektora parametrów θ^* (2.130) optymalnej funkcji aproksymującej (2.131) przy kryterium (2.128).

Przykład 2.7. Rozważmy obiekt o jednym wyjściu ($L=1$), w przypadku którego dla zadanej serii identyfikującej U_N zebrano wyniki pomiarów Y_N . Do opracowania wyników zaproponowano model liniowy względem parametrów postaci (2.134). Przyjęto kryterium kwadratowe, tj. w kryterium (2.153) zastosowano funkcję q postaci (2.125). Jest to tak zwany przypadek liniowo-kwadratowy. W tym przykładzie kryterium jakości modelu przyjmuje postać

$$Q_N(\theta) = \sum_{n=1}^N (y_n - \bar{y}_n)^2 = \sum_{n=1}^N (y_n - \theta^T \varphi(u_n))^2. \quad (2.159)$$

Przyrównanie gradientu kryterium (2.159) względem θ do R -wymiarowego wektora zerowego ($\mathbf{0}_R$) daje układ równań

$$\text{grad}_{\theta} Q_N(\theta) = -2 \sum_{n=1}^N (y_n - \theta_N^{*T} \varphi(u_n)) \varphi(u_n) = \mathbf{0}_R, \quad (2.160)$$

z którego wyznaczamy wektor parametrów optymalnego modelu

$$\theta_N^* = \left[\sum_{n=1}^N \varphi(u_n) \varphi^T(u_n) \right]^{-1} \sum_{n=1}^N y_n \varphi(u_n). \quad (2.161)$$

Dla szczególnego przypadku (rys. 2.19) przy jednowymiarowym wejściu ($S=1$) oraz po przyjęciu

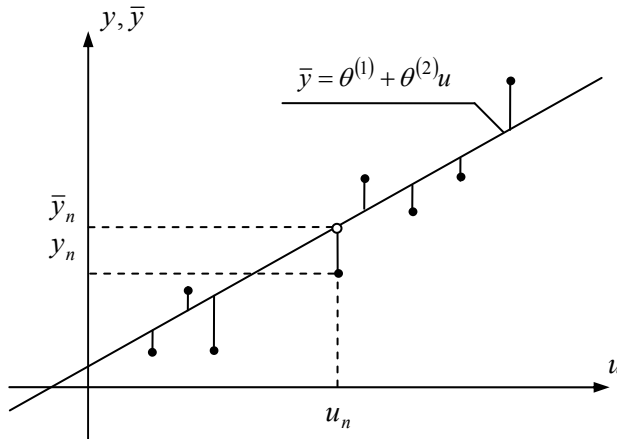
$$\varphi(u) = \begin{bmatrix} 1 \\ u \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \end{bmatrix}, \quad (2.162)$$

model (2.134) ma postać

$$\bar{y} = \theta^T \varphi(u) = \theta^{(1)} + \theta^{(2)} u, \quad (2.163)$$

a jego optymalne parametry można wyznaczyć za pomocą wzoru

$$\theta_N^* = \begin{bmatrix} \theta_N^{*(1)} \\ \theta_N^{*(2)} \end{bmatrix} = \begin{bmatrix} \frac{\left(\sum_{n=1}^N u_n^2 \right) \left(\sum_{n=1}^N y_n \right) - \left(\sum_{n=1}^N u_n \right) \left(\sum_{n=1}^N y_n u_n \right)}{N \sum_{n=1}^N u_n^2 - \left(\sum_{n=1}^N u_n \right)^2} \\ \frac{N \sum_{n=1}^N y_n u_n - \left(\sum_{n=1}^N u_n \right) \left(\sum_{n=1}^N y_n \right)}{N \sum_{n=1}^N u_n^2 - \left(\sum_{n=1}^N u_n \right)^2} \end{bmatrix}. \quad (2.164)$$

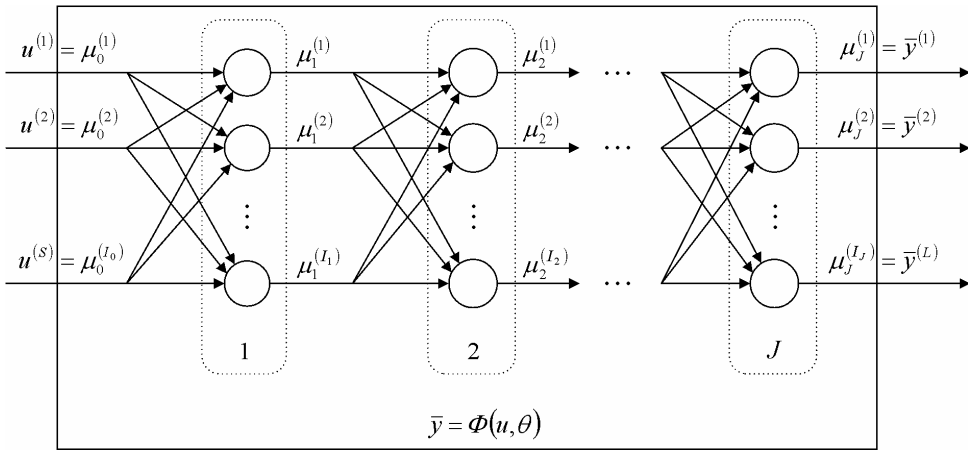


Rys. 2.19. Zestaw pomiarów i model liniowy



2.3.3. Zadanie wyboru optymalnego modelu z wykorzystaniem sieci neuronowej

Jednym z często stosowanych opisów jest model (2.121) w postaci sieci neuronowej [60, 63, 82, 136]. Do modelowania charakterystyk statycznych stosowany jest opis w postaci wielowarstwowej jednokierunkowej sieci neuronowej. Sieci te są popularnymi, uniwersalnymi „aproxymatorami”. Dobór struktury sieci, czyli liczby warstw oraz liczby neuronów w warstwie, stanowi odrębny problem. Zauważmy, że sieć z jedną warstwą ukrytą daje szeroką klasę modeli. Na rysunku 2.20 zaproponowano schemat J -warstwowej sieci dla obiektu statycznego z S -wymiarowym wejściem u oraz L -wymiarowym wyjściem y .



Rys. 2.20. Wielowarstwowa jednokierunkowa sieć neuronowa

W j -tej warstwie znajduje się I_j elementów zwanych neuronami. Oznaczmy przez E_{ij} i -ty neuron w j -tej warstwie, $i = 1, 2, \dots, I_j$, $j = 1, 2, \dots, J$. Na rysunku 2.21 przedstawiono schemat neuronu E_{ij} .

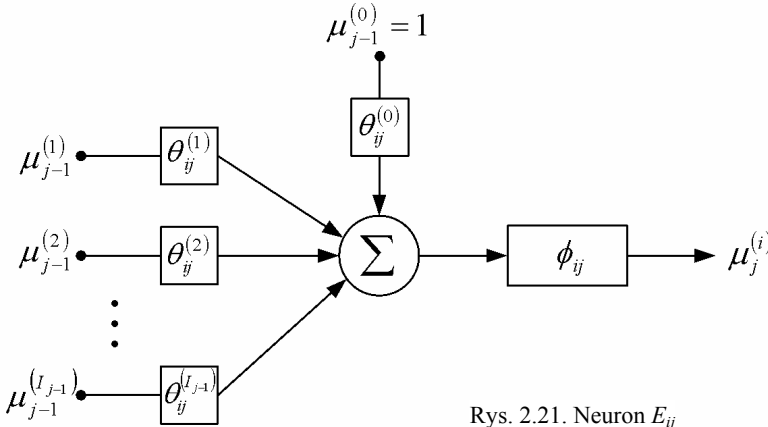
Dla elementu E_{ij} z wejściami

$$\mu_{j-1} = \begin{bmatrix} \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(I_{j-1})} \end{bmatrix}, \quad (2.165)$$

oraz wyjściem $\mu_j^{(i)}$ proponujemy opis postaci

$$\mu_j^{(i)} = \phi_{ij} \left(\sum_{s=1}^{I_{j-1}} \mu_{j-1}^{(s)} \theta_{ij}^{(s)} + \theta_{ij}^{(0)} \right), \quad (2.166)$$

gdzie: $\theta_{ij}^{(s)}$, $s = 0, 1, \dots, I_{j-1}$, jest zestawem współczynników wagowych, a ϕ_{ij} funkcją aktywacji.



Rys. 2.21. Neuron E_{ij}

Po oznaczeniu przez

$$\bar{\mu}_{j-1} \stackrel{\text{df}}{=} \begin{bmatrix} \mu_{j-1}^{(0)} \\ \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(I_{j-1})} \end{bmatrix} = \begin{bmatrix} \mu_{j-1}^{(0)} \\ \mu_{j-1}^{(1)} \\ \mu_{j-1}^{(2)} \\ \vdots \\ \mu_{j-1}^{(I_{j-1})} \end{bmatrix}, \quad (2.167)$$

gdzie $\mu_{j-1}^{(0)} = 1$, oraz zdefiniowaniu wektora współczynników wagowych

$$\theta_{ij} \stackrel{\text{df}}{=} \begin{bmatrix} \theta_{ij}^{(0)} \\ \theta_{ij}^{(1)} \\ \vdots \\ \theta_{ij}^{(I_{j-1})} \end{bmatrix}, \quad (2.168)$$

model (2.165) neuronu E_{ij} możemy przedstawić w zwartej postaci

$$\mu_j^{(i)} = \phi_{ij}(\bar{\mu}_{j-1}^T \theta_{ij}). \quad (2.169)$$

Zestawienie neuronów E_{ij} , $i = 1, 2, \dots, I_j$, daje j -tą warstwę. Zakładamy, że wejścia każdego neuronu w j -tej warstwie są takie same i określone przez (2.165). W konsekwencji opis j -tej warstwy przyjmuje postać

$$\mu_j = \begin{bmatrix} \mu_j^{(1)} \\ \mu_j^{(2)} \\ \vdots \\ \mu_j^{(I_j)} \end{bmatrix} = \begin{bmatrix} \phi_{1j}(\bar{\mu}_{j-1}^T \theta_{1j}) \\ \phi_{2j}(\bar{\mu}_{j-1}^T \theta_{2j}) \\ \vdots \\ \phi_{I_j j}(\bar{\mu}_{j-1}^T \theta_{I_j j}) \end{bmatrix} \stackrel{\text{df}}{=} \phi_j(\bar{\mu}_{j-1}, \theta_j), \quad (2.170)$$

gdzie ϕ_j jest wektorem funkcji charakteryzującej j -tą warstwę z wejściami μ_{j-1} oraz macierzą parametrów

$$\theta_j \stackrel{\text{df}}{=} [\theta_{1j} \quad \theta_{2j} \quad \dots \quad \theta_{I_j j}] = \begin{bmatrix} \theta_{1j}^{(0)} & \theta_{2j}^{(0)} & \dots & \theta_{I_j j}^{(0)} \\ \theta_{1j}^{(1)} & \theta_{2j}^{(1)} & \dots & \theta_{I_j j}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{1j}^{(I_j-1)} & \theta_{2j}^{(I_j-1)} & \dots & \theta_{I_j j}^{(I_j-1)} \end{bmatrix}. \quad (2.171)$$

Połączenie kolejnych warstw daje sieć neuronową, której rekurencyjny opis przyjmuje postać

$$\mu_j = \phi_j(\bar{\mu}_{j-1}, \theta_j), \quad j = 1, 2, \dots, J. \quad (2.172)$$

W sieciach tych należy wyróżnić warstwę wejściową ($j=1$), warstwy ukryte ($j=2, 3, \dots, J-1$) oraz warstwę wyjściową ($j=J$). Powracając do zadania modelowania charakterystyki statycznej (2.121) za pomocą sieci neuronowej, przyjmujemy, że wejściami warstwy pierwszej są wejścia obiektu, czyli $\mu_0 = u$ oraz $I_0 = S$, a zatem

$$\mu_1 = \phi_1(\bar{\mu}_0, \theta_1) = \phi_1(u, \theta_1). \quad (2.173)$$

Dla warstw ukrytych obowiązuje rekurencyjna zależność

$$\mu_j = \phi_j(\bar{\mu}_{j-1}, \theta_j), \quad j = 2, 3, \dots, J-1. \quad (2.174)$$

Wyjścia warstwy J -tej są wyjściami modelu, tj. $\mu_J = \bar{y}$ oraz $I_J = L$, czyli

$$\bar{y} = \mu_J = \phi_J(\bar{\mu}_{J-1}, \theta_J). \quad (2.175)$$

Złożenie (2.173), (2.174) oraz (2.175) daje charakterystykę (2.121) postaci

$$\bar{y} = \begin{bmatrix} \bar{y}^{(1)} \\ \bar{y}^{(2)} \\ \vdots \\ \bar{y}^{(L)} \end{bmatrix} = \begin{bmatrix} \mu_J^{(1)} \\ \mu_J^{(2)} \\ \vdots \\ \mu_J^{(I_J)} \end{bmatrix} = \phi_J(\phi_{J-1}(\dots \phi_1(u, \theta_1), \dots, \theta_{J-1}), \theta_J) \stackrel{\text{df}}{=} \begin{bmatrix} \Phi^{(1)}(u, \theta) \\ \Phi^{(2)}(u, \theta) \\ \vdots \\ \Phi^{(L)}(u, \theta) \end{bmatrix} \stackrel{\text{df}}{=} \Phi(u, \theta), \quad (2.176)$$

gdzie θ jest zestawem parametrów złożonych z macierzy (2.171) parametrów kolejnych warstw, czyli

$$\theta \stackrel{\text{df}}{=} \{\theta_1, \theta_2, \dots, \theta_J\}. \quad (2.177)$$

Wyznaczenie optymalnych parametrów modelu postaci (2.176) jest zwane „uczeniem sieci”. Problem ten, dla zadanej serii identyfikującej wraz z wynikami eksperymentu (2.149), sprowadza się do rozwiązania zadania optymalizacji (2.157), w którym w kryterium jakości przybliżenia (2.151) w miejsce $\Phi(u_n, \theta)$, $n = 1, 2, \dots, N$, wstawimy model postaci (2.176). Ze względu na złożoną postać modelu (2.176) oraz różne kryteria (2.151) z reguły natrafimy na problem analitycznego wyznaczenia rozwiązania optymalnego, dlatego też do rozwiązania (2.157) są stosowane numeryczne metody optymalizacji [47, 85]. Zastosowanie różnych metod daje liczne, znane w literaturze [63, 82, 136], algorytmy uczenia. Tutaj ograniczymy się do przytoczenia tzw. algorytmu wstecznej propagacji błędu. Konstrukcja tego algorytmu jest oparta na metodzie gradientu prostego. Przyjmuje się kwadratowy wskaźnik jakości przybliżenia, czyli w kryterium (2.153) ocena różnicy pomiędzy zmierzoną wartością wyjścia obiektu a wartością wyjścia modelu dla zadanego wejścia ma postać (2.123). Wskaźnik (2.151)

$$Q_N(\theta) = \sum_{n=1}^N [y_n - \Phi(u_n, \theta)]^T [y_n - \Phi(u_n, \theta)] = \sum_{n=1}^N \sum_{l=1}^L (y_n^{(l)} - \Phi^{(l)}(u_n, \theta))^2. \quad (2.178)$$

Korzystając z numerycznej metody optymalizacji do rozwiązania zadania (2.157) z kryterium (2.178), w kolejnych krokach algorytmu optymalizacji otrzymujemy ciąg wartości $\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_N$, takich że $Q_N(\tilde{\theta}_1) > Q_N(\tilde{\theta}_2) > \dots > Q_N(\tilde{\theta}_N)$, które po N krokach są przybliżeniem optymalnych wartości parametrów θ_N^* modelu (2.176), czyli $\tilde{\theta}_N \approx \theta_N^*$ oraz $Q_N(\tilde{\theta}_N) \approx Q_N(\theta_N^*)$. Zastosowanie metody gradientu prostego daje następujący algorytm

$$\tilde{\theta}_{ij,n+1}^{(s)} = \tilde{\theta}_{ij,n}^{(s)} - \eta_n \frac{\partial q_n(\theta)}{\partial \theta_{ij}^{(s)}} \Big|_{\theta_{ij}^{(s)} = \tilde{\theta}_{ij,n}^{(s)}}, \quad s = 0, 1, \dots, I_{j-1}, i = 1, 2, \dots, I_j, j = 1, 2, \dots, J, \quad (2.179)$$

gdzie η_n jest parametrem procedury optymalizacji, zwanym współczynnikiem uczenia, natomiast

$$q_n(\theta) \stackrel{\text{df}}{=} [y_n - \Phi(u_n, \theta)]^T [y_n - \Phi(u_n, \theta)] = \sum_{l=1}^L (y_n^{(l)} - \Phi^{(l)}(u_n, \theta))^2 = \sum_{l=1}^L \varepsilon_n^{(l)2} \quad (2.180)$$

jest n -tym składnikiem sumy (2.178), który odpowiada n -temu wynikowi pomiarów, a

$$\varepsilon_n^{(l)} \stackrel{\text{df}}{=} y_n^{(l)} - \bar{y}_n^{(l)} = y_n^{(l)} - \Phi^{(l)}(u_n, \theta), \quad l = 1, 2, \dots, L \quad (2.181)$$

jest różnicą pomiędzy wyjściem l -tym obiektu a odpowiednim wyjściem sieci.

Po zróżniczkowaniu wyrażenia (2.180) względem poszczególnych parametrów, z uwzględnieniem modelu neuronu (2.166), otrzymujemy

$$\left. \frac{\partial q_n(\theta)}{\partial \theta_{ij}^{(s)}} \right|_{\theta_{ij}^{(s)} = \tilde{\theta}_{ij,n}^{(s)}} = -2\delta_{jn}^{(i)} \mu_{j-1,n}^{(s)}, \quad (2.182)$$

gdzie $\mu_{j-1,n}^{(0)} = 1$, a $\mu_{j-1,n}^{(s)}$, $s = 1, 2, \dots, I_{j-1}$ są wartościami wejść w j -tej warstwie, w n -tym takcie procedury uczenia

$$\varepsilon_{jn}^{(i)} \stackrel{\text{df}}{=} \begin{cases} \varepsilon_n^{(i)} & \text{dla } j = J \\ \sum_{p=1}^{I_{j+1}} \delta_{j+1,n}^{(p)} \tilde{\theta}_{pj+1,n}^{(i)} & \text{dla } j = J-1, J-2, \dots, 1 \end{cases} \quad (2.183)$$

oraz

$$\delta_{jn}^{(i)} \stackrel{\text{df}}{=} \varepsilon_{jn}^{(i)} \left. \frac{d\phi_{ij}(\kappa_j^{(i)})}{d\kappa_j^{(i)}} \right|_{\theta_{ij}^{(s)} = \tilde{\theta}_{ij,n}^{(s)}}, \quad (2.184)$$

gdzie

$$\kappa_j^{(i)} \stackrel{\text{df}}{=} \sum_{s=1}^{I_{j-1}} \mu_{j-1,n}^{(s)} \theta_{ij}^{(s)} + \theta_{ij}^{(0)}.$$

Ostatecznie, po podstawieniu powyższego do (2.178), algorytm wstecznej propagacji błędów przyjmuje postać

$$\tilde{\theta}_{ij,n+1}^{(s)} = \tilde{\theta}_{ij,n}^{(s)} + 2\eta_n \delta_{jn}^{(i)} \mu_{j-1,n}^{(s)}, \quad s = 0, 1, \dots, I_{j-1}, \quad i = 1, 2, \dots, I_j, \quad j = 1, 2, \dots, J. \quad (2.185)$$

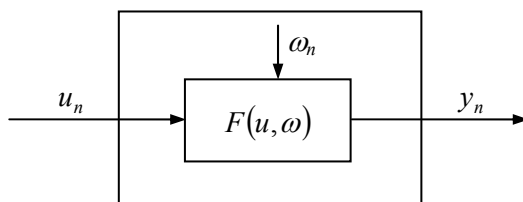
Zwróćmy uwagę na dwie fazy działania algorytmu, zaczynającego się od podania na wejście sieci n -tego wzorca uczącego (n -tego elementu z serii identyfikującej). Zostaje on przetworzony w kolejnych warstwach sieci według wzorów (2.173)–(2.175), co zapisano w zwartej postaci (2.176). W wyniku tych obliczeń, dla n -tego elementu z serii identyfikującej, otrzymujemy wartości wyjść kolejnych warstw sieci włącznie z ostatnią warstwą. Teraz, znając wartość składowych wektora wyjść obliczonych z modelu (2.176), tj. wartości wyjść z ostatniej warstwy sieci, obliczamy według (2.181) błąd odpowiadający n -temu pomiarowi wartości składowych wyjścia. Wykorzystując regułę delta (2.185), z uwzględnieniem (2.183) i (2.184), możemy zmodyfikować wagi

neuronów ostatniej warstwy. Zauważmy, że ze względu na rekurencyjną zależność (2.176), która jest konsekwencją zależności (2.173)–(2.175), minimalizowana funkcja (2.180) zależy nie tylko od wartości wag ostatniej warstwy, ale także od wartości wag ukrytych. Dlatego też błąd warstwy wyjściowej propagowany jest wstecz (od warstwy wyjściowej do warstwy wejściowej), co pokazuje druga część wzoru (2.183). Teraz, wykorzystując regułę delta (2.185), z uwzględnieniem (2.184), możemy zmodyfikować wagi neuronów warstw ukrytych.

Ze stosowaniem sieci neuronowych wiąże się wiele dodatkowych problemów podejmowanych w pracach [74, 75, 136], które nie są tutaj omawiane. Należą do nich: wybór struktury sieci, wybór początkowych wartości wag, który ma istotny wpływ na uzyskanie rozwiązania optymalnego, oraz dobór współczynnika uczenia η_n , który jest związany z zapewnieniem zbieżności przedstawionej procedury.

2.4. Zadanie wyboru optymalnego modelu w warunkach losowych

W tym podrozdziale rozważymy przypadek, w którym zarówno wektor wejść, jak i wektor wyjść są wartościami zmiennych losowych. Są dwie zasadnicze przyczyny losowości. Pierwsza z nich tkwi w obiekcie. Podając wielokrotnie tę samą wartość u_n na wejście obiektu, na wyjściu otrzymujemy różne losowe wartości y_n . Jest tak wtedy, gdy w badanym obiekcie pewne wielkości niemierzalne ω zmieniają się w sposób losowy (rys. 2.22).



Rys. 2.22. Obiekt identyfikacji z losowo zmienną niemierzalną wielkością ω

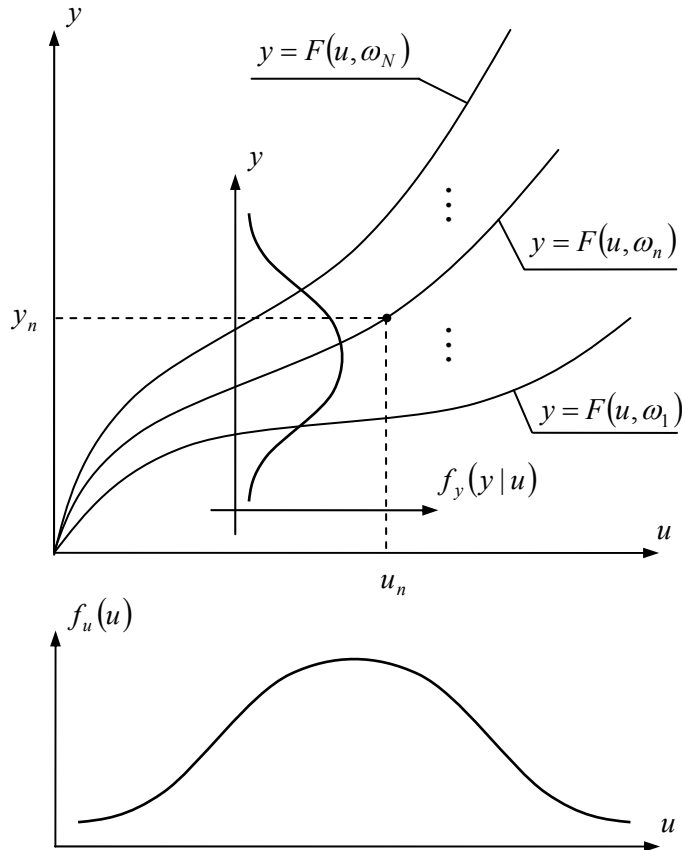
W konsekwencji możemy przyjąć, że badany obiekt jest opisany rodziną charakterystyk indeksowaną wielkością losową ω (rys. 2.23). Podając wartość u_n na wejście obiektu, na wyjściu otrzymujemy różne wartości y_n z charakterystyki, w której losowa wielkość przyjęła wartość ω_n . Opiszmy precyzyjnie aktualny obiekt identyfikacji. Dla badanego obiektu istnieje charakterystyka statyczna postaci

$$y = F(u, \omega). \quad (2.186)$$

Na obiekt działają pewne niemierzalne wielkości losowe ω z przestrzeni $\Omega \subseteq \mathbb{R}^L$. O losowo zmiennym wektorze ω zakładamy, że w n -tym pomiarze jest on równy ω_n

i jest wartością L -wymiarowej zmiennej losowej $\underline{\omega}$, dla której istnieje funkcja gęstości rozkładu prawdopodobieństwa $f_{\omega}(\omega)$. Dla ustalonego wejścia u na wyjściu obiektu obserwujemy więc wartość L -wymiarowej zmiennej losowej \underline{y} , która jest wynikiem przekształcenia zmiennej losowej $\underline{\omega}$ zgodnie z (2.186), czyli

$$\underline{y} = F(u, \underline{\omega}). \quad (2.187)$$



Rys. 2.23. Charakterystyka statyczna jednowymiarowego obiektu identyfikacji ($L = S = 1$) z losowo zmienną, niemierzalną wielkością ω

Przy założeniu, że funkcja F jest wzajemnie jednoznaczna względem ω i istnieje funkcja odwrotna względem ω , czyli

$$\underline{\omega} = F_{\omega}^{-1}(u, \underline{y}), \quad (2.188)$$

możemy wyznaczyć warunkową gęstość rozkładu prawdopodobieństwa wyjścia \underline{y} , pod warunkiem, że wejście przyjęło określoną wartość ($\underline{u} = u$), czyli

$$f_y(y|u) = f_\omega(F_\omega^{-1}(u, y)) |J_F|, \quad (2.189)$$

gdzie

$$J_F = \frac{\partial F_\omega^{-1}(u, y)}{\partial y}. \quad (2.190)$$

Można osłabić założenia dotyczące wzajemnej jednoznaczności funkcji F względem ω . W tym celu należy podzielić zbiór Ω na rozłączne podzbiory, w których funkcja F jest równowartościowa, i skorzystać z przekształcenia przedstawionego w Dodatku.

W ten sposób pokazaliśmy, że występująca w obiekcie losowość może być opisana przez warunkową gęstość rozkładu prawdopodobieństwa wyjścia przy ustalonym wejściu. W dalszej części będziemy zakładać, że opis (2.186) nie jest znany. Przedstawione rozważania mają na celu uzasadnienie przyczyn losowości w obiekcie identyfikacji.

Druga możliwa przyczyna losowości to losowe wejście. Dalej założymy, że wejście jest wartością zmiennej losowej \underline{u} i istnieje funkcja gęstości rozkładu prawdopodobieństwa $f_u(u)$ określona na przestrzeni wejść obiektu (rys. 2.23).

Podsumowując dotychczasowe rozważania, stwierdzamy, że wejścia u i wyjścia y są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$. Zmienne losowe przyjmują wartości ze zbiorów wejść i wyjść obiektu, tj. $u \in \mathbf{U} \subseteq \mathbf{R}^S$, \mathcal{U} – przestrzeń wejść, $y \in \mathbf{Y} \subseteq \mathbf{R}^L$, \mathcal{Y} – przestrzeń wyjść. Istnieje łączny rozkład prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ określony na $\mathbf{U} \times \mathbf{Y}$. Funkcję gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ oznaczmy przez $f(u, y)$. Warunkową funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , pod warunkiem, że zmienna losowa \underline{u} przyjmie wartość u , oznaczmy przez $f_y(y|u)$, a funkcję brzegowego rozkładu prawdopodobieństwa zmiennej losowej \underline{u} oznaczmy przez $f_u(u)$.

Możliwe są dwa następujące przypadki:

1. Funkcja gęstości rozkładu prawdopodobieństwa $f(u, y)$ lub funkcja gęstości warunkowego rozkładu prawdopodobieństwa wyjścia, pod warunkiem, że wejście przyjęło określoną wartość $f_y(y|u)$, oraz funkcja gęstości rozkładu prawdopodobieństwa wejścia $f_u(u)$ są znane. Przypadek ten nazwiemy „pełną informacją probabilistyczną”.
2. Rozkład prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ istnieje, ale nie jest znany. W zamian dysponujemy zbiorem obserwacji pary zmiennych loso-

wych $(\underline{u}, \underline{y})$, czyli zbiorem: (u_n, y_n) , $n=1, 2, \dots, N$. Przypadek ten nazwiemy „niepełną informacją probabilistyczną”.

2.4.1. Pełna informacja probabilistyczna

Rozważymy przypadek, w którym wektory wejść u i wyjść y obiektu identyfikacji są wartościami pary S -, L -wymiarowych zmiennych losowych $(\underline{u}, \underline{y})$. Zmienne losowe przyjmują wartości ze zbiorów wejść i wyjść obiektu, tj. u – wektor wejść obiektu $u \in \mathbf{U} \subseteq \mathbb{R}^S$, \mathcal{U} – przestrzeń wejść, y – wektor wyjść obiektu, $y \in \mathbf{Y} \subseteq \mathbb{R}^L$, \mathcal{Y} – przestrzeń wyjść. Znana jest funkcja gęstości łącznego rozkładu prawdopodobieństwa $f(u, y)$ pary zmiennych losowych $(\underline{u}, \underline{y})$, określona na $\mathbf{U} \times \mathbf{Y}$, lub funkcja gęstości warunkowego rozkładu prawdopodobieństwa wyjścia $f_y(y|u)$, określona na \mathcal{Y} dla każdego ustalonego $u \in \mathbf{U}$, oraz funkcja gęstości rozkładu prawdopodobieństwa wejścia $f_u(u)$ określona na \mathcal{U} , bowiem

$$f(u, y) = f_y(y|u)f_u(u). \quad (2.191)$$

Zaproponujemy dwa zadania wyboru optymalnego modelu.

Model nieparametryczny

W pierwszym przypadku do opisu badanego obiektu zaproponujemy model postaci

$$\bar{y} = \Phi(u), \quad (2.192)$$

gdzie Φ jest nieznaną funkcją, opisującą zależność wyjścia modelu \bar{y} od wejścia u obiektu. Wektory u oraz \bar{y} są elementami – odpowiednio – przestrzeni wejść – \mathcal{U} oraz wyjść modelu \mathcal{Y} , czyli: $u \in \mathbf{U} \subseteq \mathbb{R}^S$, a $\bar{y} \in \mathbf{Y} \subseteq \mathbb{R}^L$. Jest to tak zwany przypadek nieparametryczny.

Podobnie jak w punkcie 2.3.1, należy wprowadzić ocenę różnicy pomiędzy badanym wyjściem obiektu a wyjściem proponowanego modelu. Biorąc pod uwagę losowy charakter badanego obiektu, adekwatną oceną modelu będzie wartość oczekiwana oceny różnicy pomiędzy wyjściem obiektu i proponowanego modelu, czyli

$$Q(\Phi) = E_{\underline{u}, \underline{y}} [q(\underline{y}, \Phi(\underline{u}))] = \int_{\mathbf{U}} \int_{\mathbf{Y}} q(y, \Phi(u)) f(u, y) dy du, \quad (2.193)$$

gdzie wartość funkcji q jest oceną różnicy pomiędzy wyjściem obiektu i modelu dla ustalonego $u \in \mathbf{U}$ i jest zdefiniowana tak, jak w podrozdziale 2.3 (2.122).

Funkcjonał (2.193) jest miarą jakości przybliżenia i zależy od postaci funkcji Φ . Funkcjonał $Q(\Phi)$ stanowi kryterium wyboru modelu, a konkretnie kryterium wyboru

funkcji opisującej zależność pomiędzy wejściem obiektu i wyjściem modelu. Problem wyznaczenia optymalnego modelu sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką funkcję Φ^* , dla której miara jakości przybliżenia $Q(\Phi)$ (2.193) przyjmuje wartość minimalną względem Φ , tj.

$$\Phi^* : Q(\Phi^*) = \min_{\Phi} Q(\Phi), \quad (2.194)$$

gdzie Φ^* jest optymalną postacią funkcji, natomiast

$$\bar{y} = \Phi^*(u) \quad (2.195)$$

jest optymalnym modelem.

Po uwzględnieniu zależności (2.191) w (2.193) otrzymujemy

$$Q(\Phi) = E_{\underline{u}} \left[E_{\underline{y}} [q(\underline{y}, \Phi(\underline{u})) | \underline{u} = u] \right] = \int_{\underline{u}} \left[\int_{\underline{y}} q(y, \Phi(u)) f_y(y|u) dy \right] f_u(u) du. \quad (2.196)$$

Korzystając z postaci (2.196) kryterium wyboru optymalnego modelu, zauważmy, że zadanie (2.194) – optymalizacji funkcjonału po funkcjach Φ – możemy zastąpić zadaniem optymalizacji fragmentu funkcjonału (2.196) – warunkowej wartości oczekiwanej przy ustalonym u , względem wartości funkcji $\Phi(u)$. Przyjmujemy oznaczenie

$$Q_u(\Phi(u)) = E_{\underline{y}} [q(\underline{y}, \Phi(\underline{u})) | \underline{u} = u] = \int_{\underline{y}} q(y, \Phi(u)) f_y(y|u) dy. \quad (2.197)$$

Obecnie problem wyznaczenia optymalnego modelu sprowadza się do następującego zadania optymalizacji: dla ustalonego u należy wyznaczyć optymalną wartość funkcji $\Phi^*(u)$, dla której $Q_u(\Phi(u))$ (2.197) przyjmuje wartość minimalną względem wartości funkcji $\Phi(u)$, czyli

$$\Phi^*(u) : Q_u(\Phi^*(u)) = \min_{\Phi(u)} Q_u(\Phi(u)). \quad (2.198)$$

Analityczny wynik rozwiązania zadania (2.198) uzyskamy dla kryterium kwadratowego, czyli po przyjęciu

$$q(y, \bar{y}) = (y - \bar{y})^T (y - \bar{y}), \quad (2.199)$$

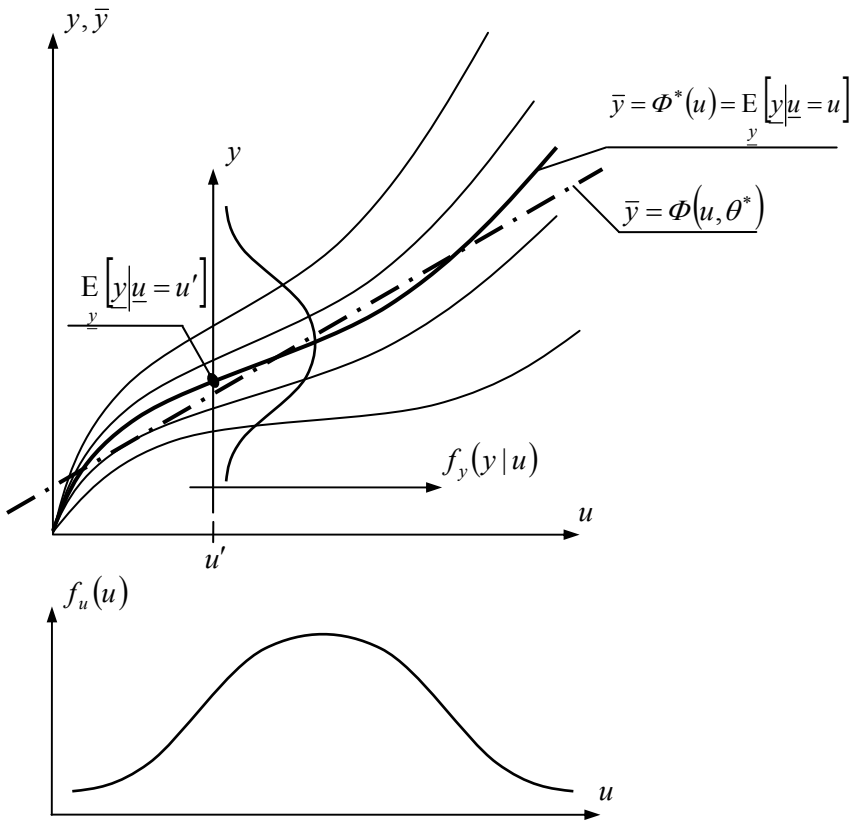
kryterium (2.197) ma postać

$$Q_u(\Phi(u)) = E_{\underline{y}} \left[(y - \Phi(u))^T (y - \Phi(u)) | \underline{u} = u \right] = \int_{\underline{y}} (y - \Phi(u))^T (y - \Phi(u)) f_y(y|u) dy, \quad (2.200)$$

a optymalny model (2.195)

$$\bar{y} = \Phi^*(u) = E_{\underline{y}} [y | \underline{u} = u] = \int_{\underline{y}} y f(y|u) dy. \quad (2.201)$$

Zależność (2.201) nazywamy **regresją pierwszego rodzaju** (rys. 2.24).



Rys. 2.24. Funkcje regresji dla jednowymiarowego obiektu ($L = S = 1$)

Przykład 2.8. Rozważmy przypadek, w którym wejścia u i wyjścia y obiektu identyfikacji są wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$. Zmienne losowe przyjmują wartości – odpowiednio – z S - oraz L -wymiarowych przestrzeni liczb rzeczywistych, tj. $u \in \mathbb{R}^S$, $y \in \mathbb{R}^L$. Para zmiennych losowych ma $L+S$ -wymiarowy rozkład normalny, a funkcja gęstości łącznego rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ ma postać

$$f(u, y) = (2\pi)^{-\frac{L+S}{2}} |\Sigma^{-1}|^{\frac{1}{2}} \exp \left[-\frac{1}{2} \left(\begin{bmatrix} u \\ y \end{bmatrix} - \begin{bmatrix} m_u \\ m_y \end{bmatrix} \right)^T \Sigma^{-1} \left(\begin{bmatrix} u \\ y \end{bmatrix} - \begin{bmatrix} m_u \\ m_y \end{bmatrix} \right) \right], \quad (2.202)$$

gdzie: m_u oraz m_y są wartościami oczekiwanymi zmiennych losowych – odpowiednio – \underline{u} oraz \underline{y} , natomiast Σ jest macierzą kowariancji postaci

$$\Sigma = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uy} \\ \Sigma_{yu} & \Sigma_{yy} \end{bmatrix}, \quad (2.203)$$

gdzie: Σ_{uu} jest macierzą kowariancji wektora zmiennej losowej \underline{u} , Σ_{yy} jest macierzą kowariancji wektora zmiennej losowej \underline{y} , Σ_{uy} jest macierzą kowariancji wektorów zmiennych losowych \underline{u} oraz \underline{y} , a Σ_{yu} jest macierzą kowariancji wektorów zmiennych losowych \underline{y} oraz \underline{u} .

W tym przypadku warunkowy rozkład prawdopodobieństwa zmiennej losowej \underline{y} pod warunkiem, że zmienna losowa \underline{u} przyjęła wartość u , jest L -wymiarowym rozkładem normalnym, a funkcja gęstości warunkowego rozkładu prawdopodobieństwa wyjścia dla ustalonego wejścia $f_y(y|u)$ ma postać

$$f(y|u) = (2\pi)^{-\frac{L}{2}} |\Sigma_{y|u}^{-1}|^{\frac{1}{2}} \exp \left[-\frac{1}{2} (y - m_{y|u})^T \Sigma_{y|u}^{-1} (y - m_{y|u}) \right], \quad (2.204)$$

gdzie macierz kowariancji $\Sigma_{y|u}$ wyraża się zależnością [79]

$$\Sigma_{y|u} = \Sigma_{yy} - \Sigma_{yu} \Sigma_{uu}^{-1} \Sigma_{uy}, \quad (2.205)$$

a wartość oczekiwana $m_{y|u}$ ma postać

$$m_{y|u} = m_y + \Sigma_{yu} \Sigma_{uu}^{-1} (u - m_u). \quad (2.206)$$

W konsekwencji otrzymamy optymalny model (2.201)

$$\bar{y} = \Phi^*(u) = E \left[\underset{y}{y} | \underline{u} = u \right] = m_{y|u} = m_y + \Sigma_{yu} \Sigma_{uu}^{-1} (u - m_u), \quad (2.207)$$

który po prostych przekształceniach algebraicznych przyjmuje postać

$$\bar{y} = \Phi^*(u) = Au + b, \quad (2.208)$$

gdzie: macierz A o wymiarach $L \times S$ dana jest zależnością

$$A = \Sigma_{yu} \Sigma_{uu}^{-1}, \quad (2.209)$$

a wektor b

$$b = m_y - \Sigma_{yu} \Sigma_{uu}^{-1} m_u. \quad (2.210)$$

Przykładem tym pokazaliśmy, że jeżeli para zmiennych losowych $(\underline{u}, \underline{y})$, opisujących wejście i wyjście obiektu, ma rozkład normalny, to optymalny model ma postać (2.208).



Model parametryczny

W drugim przypadku, w zadaniu wyboru modelu, do opisu badanego obiektu proponujemy model parametryczny postaci

$$\bar{y} = \Phi(u, \theta), \quad (2.211)$$

gdzie: Φ jest zaproponowaną funkcją opisującą zależność wyjścia modelu \bar{y} od wejścia u , natomiast θ jest wektorem parametrów modelu. Wielkości u , \bar{y} i θ są elementami – odpowiednio – przestrzeni wejść \mathcal{U} , wyjść modelu \mathcal{Y} oraz parametrów Θ , czyli: $u \in \mathcal{U} \subseteq \mathbb{R}^S$, $\bar{y} \in \mathcal{Y} \subseteq \mathbb{R}^L$ oraz $\theta \in \Theta \subseteq \mathbb{R}^R$.

Podobnie jak w podrozdziale 2.3, należy wprowadzić ocenę różnicy pomiędzy wyjściem badanego obiektu a wyjściem proponowanego modelu. Tak jak w przypadku nieparametrycznym (2.193), uwzględniając losowy charakter badanego obiektu, do oceny modelu wyznaczmy wartość oczekiwaną oceny różnicy pomiędzy wyjściem obiektu i proponowanego modelu, czyli

$$Q(\theta) = E_{\underline{u}, \underline{y}} [q(\underline{y}, \Phi(\underline{u}, \theta))] = \int_{\mathcal{U}} \int_{\mathcal{Y}} q(y, \Phi(u, \theta)) f(u, y) dy du. \quad (2.212)$$

Funkcja $Q(\theta)$ (2.212) jest miarą jakości przybliżenia i zależy od wartości składowych wektora parametrów θ . Funkcja $Q(\theta)$ stanowi kryterium wyboru optymalnego modelu, a konkretnie wyboru optymalnych wartości składowych wektora parametrów θ w funkcji (2.211), opisującej zależność pomiędzy wejściem obiektu i wyjściem modelu. Wyznaczenie optymalnego modelu sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość składowych wektora parametrów θ^* , dla której miara jakości przybliżenia $Q(\theta)$ (2.212) przyjmuje wartość minimalną względem $\theta \in \Theta$, tj.

$$\theta^* : (\theta^*) = \min_{\theta \in \Theta} Q(\theta), \quad (2.213)$$

gdzie θ^* jest optymalną wartością składowych wektora parametrów modelu, a funkcja

$$\bar{y} = \Phi(u, \theta^*) \quad (2.214)$$

jest optymalnym modelem parametrycznym z danej klasy modeli.

Często stosowana funkcja q w kryterium (2.212) ma postać (2.199), wówczas mówimy o kryterium kwadratowym, a model (2.214) z optymalnym wektorem parametrów θ^* , wyznaczonym dla tej postaci funkcji q , nazywany jest **regresją drugiego rodzaju** (rys. 2.24).

Zbadajmy relacje pomiędzy modelem (2.195) a (2.214) dla kryterium kwadratowego, tj. zależność pomiędzy regresją pierwszego i drugiego rodzaju.

Nie zmniejszając ogólności rozważań, w tym miejscu ograniczymy się do jednowymiarowego wyjścia ($L = 1$), wówczas kryterium (2.212) z (2.190) przyjmuje postać

$$Q(\theta) = E \left[\int_{\underline{u}, \underline{y}} (y - \Phi(u, \theta))^2 \right] = \int_U \int_Y (y - \Phi(u, \theta))^2 f(u, y) dy du. \quad (2.215)$$

Po uwzględnieniu (2.191) otrzymujemy

$$Q(\theta) = \int_U \int_Y \left((y - \Phi^*(u)) + (\Phi^*(u) - \Phi(u, \theta)) \right)^2 f_y(y|u) f_u(u) dy du. \quad (2.216)$$

Po podniesieniu do kwadratu wyrażenia pod całką otrzymujemy

$$\begin{aligned} Q(\theta) &= \int_U \int_Y (y - \Phi^*(u))^2 f_y(y|u) f_u(u) dy du \\ &\quad - 2 \int_U \left[\int_Y (y - \Phi^*(u)) (\Phi^*(u) - \Phi(u, \theta)) f_y(y|u) dy \right] f_u(u) du \\ &\quad + \int_U \int_Y \left((\Phi^*(u) - \Phi(u, \theta)) \right)^2 f_y(y|u) f_u(u) dy du. \end{aligned} \quad (2.217)$$

Biorąc pod uwagę (2.201), widzimy, że drugi składnik sumy (2.217) przyjmuje wartość równą zero. W konsekwencji, uwzględniając (2.191) w pierwszym składniku oraz całkując względem y w składniku drugim, otrzymujemy

$$Q(\theta) = \int_U \int_Y (y - \Phi^*(u))^2 f(u, y) dy du + \int_U (\Phi^*(u) - \Phi(u, \theta))^2 f_u(u) du. \quad (2.218)$$

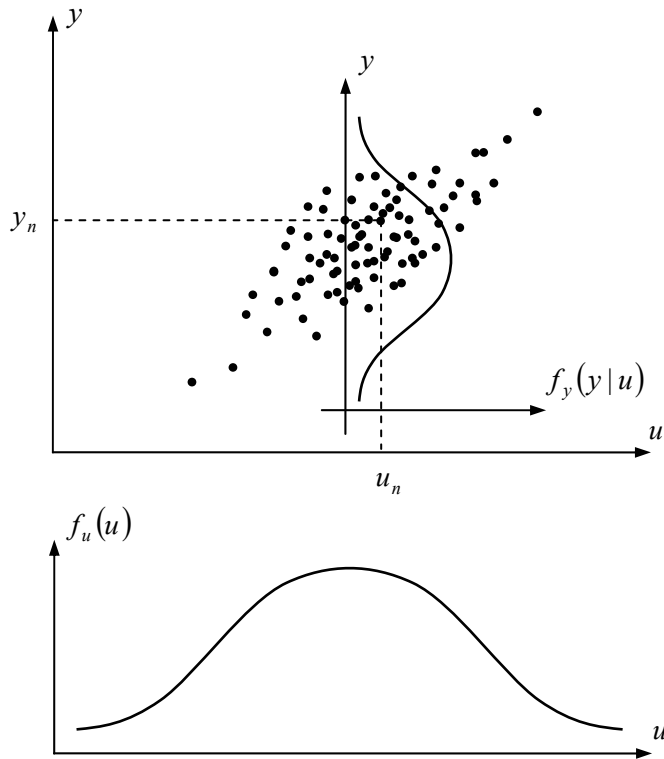
Zwróćmy uwagę, że pierwszy składnik sumy (2.218) odpowiadający optymalnej wartości funkcjonau (2.193) dla modelu (2.201) jest wartością stałą. Ostatecznie rozwiązanie zadania (2.213) sprowadza się do minimalizacji względem $\theta \in \Theta$ drugiego składnika sumy (2.218), stąd wynika wniosek: Funkcja regresji drugiego rodzaju $\Phi(u, \theta^*)$ jest optymalnym przybliżeniem funkcji regresji pierwszego rodzaju $\Phi^*(u)$ przy kwadratowym wskaźniku jakości z funkcją wagi jakości przybliżenia $f_u(u)$.

2.4.2. Niepełna informacja probabilistyczna

Teraz zajmiemy się przypadkiem, gdy rozkład prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ istnieje, ale nie jest znany. W zamian dysponujemy zbiorem

$$(u_n, y_n), \quad n = 1, 2, \dots, N, \quad (2.219)$$

pomiarów wartości składowych wektorów wejścia i wyjścia obiektu (rys. 2.25). Zakładamy, że wyniki pomiarów są niezależnymi wartościami pary zmiennych losowych $(\underline{u}, \underline{y})$.



Rys. 2.25. Pomiary wejścia i wyjścia jednowymiarowego obiektu ($L = S = 1$)

Zebrane, w wyniku eksperymentu, dane pomiarowe umożliwiają zastosowanie jednego z dwóch następujących podejść, tj. wykorzystanie empirycznych oszacowań wartości oczekiwanych (2.197) oraz (2.212) lub wykorzystanie empirycznych rozkładów prawdopodobieństwa.

Empiryczne oszacowania wskaźników jakości

Przystępując do empirycznego oszacowania wskaźnika jakości (2.197), zauważmy, że jest on warunkową wartością oczekiwaną dla ustalonego wejścia u . Empiryczne oszacowanie warunkowej wartości oczekiwanej wymaga odpowiedniej organizacji eksperymentu, tj. dla ustalonego wejścia u_n należy powtórzyć pomiar wartości składowych wyjścia obiektu y_{nm} . W konsekwencji otrzymujemy następujące wyniki pomiarów:

$$(u_n, y_{nm}), \quad m = 1, 2, \dots, M_n, \quad n = 1, 2, \dots, \bar{N}, \quad (2.220)$$

gdzie: M_n jest liczbą powtórzeń pomiarów wyjścia dla ustalonego wejścia u_n , \bar{N} – liczbą różnych wartości składowych wektora wejść, a $N = \sum_{n=1}^{\bar{N}} M_n$ – liczbą wszystkich pomiarów.

Empiryczne oszacowanie wskaźnika (2.197) dla ustalonego u_n na podstawie M_n powtórzonych pomiarów wyjścia ma postać

$$Q_{uM_n}(\Phi(u_n)) = \sum_{m=1}^{M_n} q(y_{nm}, \Phi(u_n)). \quad (2.221)$$

Optymalną wartość $\Phi_{M_n}^*(u_n)$ oszacowania wartości funkcji (2.195) dla ustalonego u_n wyznaczamy w wyniku minimalizacji (2.221) względem wartości funkcji $\Phi(u_n)$, czyli

$$\Phi_{M_n}^*(u_n) : Q_{uM_n}(\Phi_{M_n}^*(u_n)) = \min_{\Phi(u_n)} Q_{uM_n}(\Phi(u_n)). \quad (2.222)$$

Powtarzając zadanie (2.222) dla różnych wartości u_n , otrzymamy ciąg oszacowań $\Phi_{M_n}^*(u_n)$, $n = 1, 2, \dots, \bar{N}$. W szczególnym przypadku, przyjmując kwadratowy wskaźnik postaci (2.199), otrzymamy oczywisty wynik

$$\bar{y}_n = \frac{1}{M_n} \sum_{m=1}^{M_n} y_{nm}, \quad n = 1, 2, \dots, \bar{N}, \quad (2.223)$$

który jest oszacowaniem regresji pierwszego rodzaju.

Empiryczne oszacowanie wskaźnika (2.212) nie wymaga specjalnej organizacji eksperymentu. Oszacowanie to, wyznaczone na podstawie zbioru N pomiarów (2.219), ma postać

$$Q_N(\theta) = \frac{1}{N} \sum_{n=1}^N q(y_n, \Phi(u_n \theta)), \quad (2.224)$$

a na podstawie pomiarów (2.220)

$$Q_N(\theta) = \frac{1}{N} \sum_{n=1}^{\bar{N}} \sum_{m=1}^{M_n} q(y_{nm}, \Phi(u_n, \theta)). \quad (2.225)$$

Optymalną wartość wektora parametrów modelu θ_N^* wyznaczamy, minimalizując empiryczną wartość oczekiwaną (2.224) lub (2.225) względem $\theta \in \Theta$, czyli

$$\theta_N^* : Q_N(\theta_N^*) = \min_{\theta \in \Theta} Q_N(\theta). \quad (2.226)$$

Empiryczne oszacowania gęstości rozkładów prawdopodobieństwa

Kolejne podejście sprowadza się do zastąpienia gęstości rozkładu prawdopodobieństwa zmiennych losowych $(\underline{u}, \underline{y})$ estymatorami uzyskanymi na podstawie zbioru pomiarów (2.219). Wyróżnimy dwa następujące przypadki nieznanności gęstości

rozkładu prawdopodobieństwa. Pierwszy to taki, kiedy znana jest postać funkcji gęstości rozkładu prawdopodobieństwa zmiennych losowych $(\underline{u}, \underline{y})$, nie są natomiast znane parametry tego rozkładu. W drugim przypadku zakładamy, że nie jest znana postać funkcji gęstości rozkładu prawdopodobieństwa zmiennych losowych $(\underline{u}, \underline{y})$ i korzystamy z estymatorów nieparametrycznych.

W pierwszym przypadku zakładamy, że funkcja gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$ ma postać

$$f(u, y) = f_\alpha(u, y, \alpha), \quad (2.227)$$

gdzie: f_α jest znaną funkcją, natomiast $\alpha \in \mathbf{A}$ jest wektorem nieznanych parametrów, a warunkową funkcję gęstości rozkładu prawdopodobieństwa wyjścia przy ustalonym wejściu wyznaczamy z zależności

$$f_y(y|u) = \frac{f_\alpha(u, y, \alpha)}{\int_{\mathbf{Y}} f_\alpha(u, y, \alpha) dy}. \quad (2.228)$$

W tym przypadku, na podstawie pomiarów (2.219), oszacowanie wartości wektora parametrów α otrzymamy metodą maksymalnej wiarygodności. Funkcja wiarygodności ma postać

$$L_N(U_N, Y_N, \alpha) = \prod_{n=1}^N f_\alpha(u_n, y_n, \alpha). \quad (2.229)$$

Oszacowanie $\hat{\alpha}_N = \Psi_N(U_N, Y_N)$ wektora parametrów rozkładu uzyskamy w wyniku maksymalizacji względem $\alpha \in \mathbf{A}$ funkcji wiarygodności (2.229), czyli

$$\hat{\alpha}_N = \Psi_N(U_N, Y_N) : L_N(U_N, Y_N, \hat{\alpha}_N) = \max_{\alpha \in \mathbf{A}} L_N(U_N, Y_N, \alpha). \quad (2.230)$$

Po podstawieniu do (2.227) oraz (2.228) oszacowania $\hat{\alpha}_N$ w miejsce α możemy wyznaczyć odpowiednie wskaźniki zdefiniowane w punkcie 2.4.1. W wyniku podstawienia w kryterium (2.198) funkcji gęstości (2.228) z parametrami $\hat{\alpha}_N$ otrzymujemy

$$Q_{u\hat{\alpha}_N}(\Phi(u), \hat{\alpha}_N) = \int_{\mathbf{Y}} q(y, \Phi(u)) \frac{f_\alpha(u, y, \hat{\alpha}_N)}{\int_{\mathbf{Y}} f_\alpha(u, y, \hat{\alpha}_N) dy} dy. \quad (2.231)$$

Wskaźnik (2.231) zależy od $\hat{\alpha}_N$. Optymalny model uzyskamy w wyniku minimalizacji wyrażenia (2.231) względem $\Phi(u)$, czyli

$$\bar{y} = \Phi_{\hat{\alpha}_N}^*(u, \hat{\alpha}_N) : Q_{u\hat{\alpha}_N}(\Phi_{\hat{\alpha}_N}^*(u, \hat{\alpha}_N), \hat{\alpha}_N) = \min_{\Phi(u)} Q_{u\hat{\alpha}_N}(\Phi(u), \hat{\alpha}_N). \quad (2.232)$$

Po przyjęciu kwadratowego wskaźnika jakości (2.199) otrzymamy regresję pierwszego rodzaju postaci

$$\bar{y} = \Phi_{\hat{\alpha}_N}^*(u, \hat{\alpha}_N) = \int_Y y \frac{f_\alpha(u, y, \hat{\alpha}_N)}{\int_Y f_\alpha(u, y, \hat{\alpha}_N) dy} dy. \quad (2.233)$$

Podobnie dla kryterium (2.212), po podstawieniu funkcji gęstości rozkładu prawdopodobieństwa (2.227) z parametrami $\hat{\alpha}_N$, otrzymujemy

$$Q_{\hat{\alpha}_N}(\theta, \hat{\alpha}_N) = \int_U \int_Y q(y, \Phi(u, \theta)) f_\alpha(u, y, \hat{\alpha}_N) dy du. \quad (2.234)$$

Teraz optymalne wartości wektora parametrów modelu $\theta_{\hat{\alpha}_N}^*$ zależą od wektora parametrów funkcji gęstości rozkładu prawdopodobieństwa $\hat{\alpha}_N$. Wartość wektora $\theta_{\hat{\alpha}_N}^*$ uzyskamy w wyniku minimalizacji względem $\theta \in \Theta$ wyrażenia (2.234), czyli

$$\theta_{\hat{\alpha}_N}^* : Q_{\hat{\alpha}_N}(\theta_{\hat{\alpha}_N}^*, \hat{\alpha}_N) = \min_{\theta \in \Theta} Q_{\hat{\alpha}_N}(\theta, \hat{\alpha}_N). \quad (2.235)$$

Kolejny przypadek to taki, w którym rozkład prawdopodobieństwa istnieje, ale nie jest znany. Możemy wówczas skorzystać z nieparametrycznych estymatorów gęstości rozkładów prawdopodobieństwa [51].

W dalszych rozważaniach skorzystamy z estymatora Parzena. Istota estymatora Parzena jest następująca:

Dla P -wymiarowej zmiennej losowej \underline{x} istnieje funkcja gęstości rozkładu prawdopodobieństwa $f_x(x)$, ale nie jest znana. Dysponujemy zbiorem N niezależnych obserwacji zmiennej losowej \underline{x} , tj.

$$[x_1 \ x_2 \ \cdots \ x_N] = X_N. \quad (2.236)$$

Estymator Parzena nieznannej gęstości $f_x(x)$ ma postać (rys. 2.26)

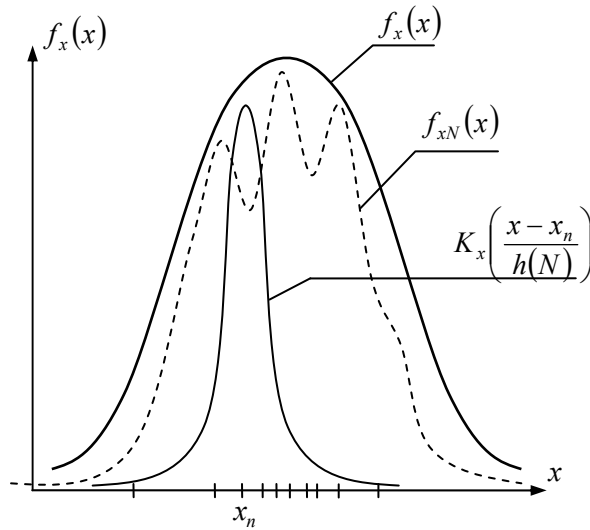
$$\hat{f}_{xN}(x; X_N) = \frac{1}{Nh^P(N)} \sum_{n=1}^N K_x \left(\frac{x - x_n}{h(N)} \right), \quad (2.237)$$

gdzie $h(N)$ jest ciągiem liczbowym, spełniającym warunki:

$$h(N) > 0, \quad \lim_{N \rightarrow \infty} h(N) = 0, \quad \lim_{N \rightarrow \infty} h^P(N) = \infty, \quad (2.238)$$

natomiast K_x – jądro estymatora – jest funkcją mierzalną, określoną na P -wymiarowej przestrzeni obserwacji, która spełnia warunki:

$$\sup_{x \in X} |K_x(x)| < \infty, \quad \int_X K_x(x) dx = 1, \quad \int_X |K_x(x)| dx < \infty, \quad \lim_{\|x\| \rightarrow \infty} |K_x(x)| = 0. \quad (2.239)$$



Rys. 2.26. Estymacja funkcji gęstości rozkładu prawdopodobieństwa ($P = 1$)

Jako przykład ciągu $h(N)$ spełniającego warunki (2.238) można podać

$$h(N) = cN^{-\alpha}, \quad (2.240)$$

gdzie: c oraz α są liczbami rzeczywistymi spełniającymi warunki:

$$c > 0, \quad 0 < \alpha < \frac{1}{P}. \quad (2.241)$$

Jako przykłady funkcji K_x , które spełniają warunki (2.239), można podać:

$$K_x(x) = \begin{cases} 1 & \text{dla } \|x\| \leq 1 \\ 0 & \text{dla } \|x\| > 1 \end{cases}, \quad K_x(x) = (2\pi)^{-\frac{P}{2}} e^{-\frac{1}{2}\|x\|^2}, \quad K_x(x) = \pi^{-P} \prod_{p=1}^P \sqrt{1 + |x^{(p)}|}. \quad (2.242)$$

Korzystając z estymatora Parzena, funkcję gęstości łącznego rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{u}, \underline{y})$, tj. funkcję $f(u, y)$, przybliżamy estymatorem

$$\hat{f}_N(u, y; U_N, Y_N) = \frac{1}{Nh^{L+S}(N)} \sum_{n=1}^N K_u\left(\frac{u - u_n}{h(N)}\right) K_y\left(\frac{y - y_n}{h(N)}\right), \quad (2.243)$$

gdzie: $h(N)$ jest ciągiem liczbowym spełniającym warunki (2.238), natomiast $K_u(u)$ oraz $K_y(y)$ są jądrami estymatora, skojarzonymi – odpowiednio – ze zmienną losową \underline{u} oraz \underline{y} , spełniającymi warunki (2.239).

Podobnie funkcję gęstości rozkładu prawdopodobieństwa wejścia $f_u(u)$ przybliżamy estymatorem

$$\hat{f}_{uN}(u; U_N) = \frac{1}{N h^S(N)} \sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right), \quad (2.244)$$

a estymator gęstości warunkowego rozkładu prawdopodobieństwa wyjścia $f_y(y|u)$ wyznaczamy

$$\begin{aligned} \hat{f}_{yN}(y|u; U_N, Y_N) &= \frac{\hat{f}_N(u, y; U_N, Y_N)}{\hat{f}_{uN}(u; U_N)} \\ &= \frac{1}{h^L(N)} \frac{\sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right) K_y \left(\frac{y - y_n}{h(N)} \right)}{\sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right)}. \end{aligned} \quad (2.245)$$

Po podstawieniu w kryterium (2.197) gęstości (2.245) otrzymujemy

$$Q_{uN}(\Phi(u); U_N, Y_N) = \int_Y q(y, \Phi(u)) \hat{f}_{yN}(y|u; U_N, Y_N) dy. \quad (2.246)$$

Zauważmy, że wskaźnik (2.246) zależy od wyników eksperymentu U_N, Y_N , dlatego optymalny model, uzyskany w wyniku minimalizacji wyrażenia (2.246) względem $\Phi(u)$, zależy od U_N oraz Y_N , czyli

$$\bar{y} = \Phi_N^*(u; U_N, Y_N) \rightarrow Q_{uN}(\Phi_N^*(u, U_N, Y_N), U_N, Y_N) = \min_{\Phi(u)} Q_{uN}(\Phi(u), U_N, Y_N). \quad (2.247)$$

Po przyjęciu kwadratowego wskaźnika jakości (2.200) otrzymamy funkcję regresji pierwszego rodzaju postaci

$$\begin{aligned} \bar{y} = \Phi_N^*(u; U_N, Y_N) &= \int_Y y \hat{f}_{yN}(y|u; U_N, Y_N) dy \\ &= \frac{1}{h^L(N)} \frac{\sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right) \int_Y y K_y \left(\frac{y - y_n}{h(N)} \right) dy}{\sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right)}. \end{aligned} \quad (2.248)$$

Przy dodatkowym założeniu, że funkcja $K_y(y)$ jest symetryczna, $K_y(y) = K_y(-y)$, mamy

$$\frac{1}{h^L(N)} \int_Y y K_y \left(\frac{y - y_n}{h(N)} \right) dy = y_n, \quad (2.249)$$

a zatem funkcja regresji pierwszego rodzaju (2.248) upraszcza się do postaci

$$\bar{y} = \Phi_N^*(u; U_N, Y_N) = \frac{\sum_{n=1}^N y_n K_u \left(\frac{u - u_n}{h(N)} \right)}{\sum_{n=1}^N K_u \left(\frac{u - u_n}{h(N)} \right)}. \quad (2.250)$$

Podobnie dla kryterium (2.212), po podstawieniu gęstości (2.243) otrzymujemy

$$Q_N(\theta; U_N, Y_N) = \int_U \int_Y q(y, \Phi(u, \theta)) \hat{f}_N(u, y; U_N, Y_N) dy du. \quad (2.251)$$

Teraz wskaźnik (2.251) zależy od wyników eksperymentu i optymalna wartość wektora parametrów modelu, którą uzyskamy po minimalizacji względem $\theta \in \Theta$ wyrażenia (2.251), zależy od U_N oraz Y_N , czyli

$$\theta_N^* = \Psi_N(U_N, Y_N) \rightarrow Q_N(\theta_N^*; U_N, Y_N) = \min_{\theta \in \Theta} Q_N(\theta; U_N, Y_N), \quad (2.252)$$

gdzie Ψ_N jest algorytmem identyfikacji.

3. Identyfikacja obiektów złożonych przy ograniczonych możliwościach pomiarowych

Identyfikacja statycznego systemu złożonego przy ograniczonych możliwościach pomiarowych dotyczy obiektu, w którym wyróżniono elementy składowe. Znana jest struktura systemu, tj. znane są powiązania pomiędzy poszczególnymi elementami (1.19). Charakterystyki statyczne poszczególnych elementów (1.13) są znane z dokładnością do parametrów. Mamy zatem wejściowo-wyjściowy system złożony, w którym wyróżniono M elementów O_1, O_2, \dots, O_M . Niech

$$y_m = F_m(u_m, \theta_m) \quad (3.1)$$

oznacza charakterystykę statyczną elementu O_m o wektorach wejść u_m i wyjść y_m , gdzie F_m jest znaną funkcją, a θ_m nieznanym wektorem parametrów charakterystyki m -tego elementu. Jest to idealny przypadek, gdy charakterystyka statyczna m -tego elementu należy do klasy modeli.

Wektor wejść i wyjść oraz wektor parametrów charakterystyki statycznej m -tego elementu są wektorami z odpowiednich przestrzeni, tj.

$$u_m = \begin{bmatrix} u_m^{(1)} \\ u_m^{(2)} \\ \vdots \\ u_m^{(S_m)} \end{bmatrix} \in U_m \subseteq \mathbb{R}^{S_m}, y_m = \begin{bmatrix} y_m^{(1)} \\ y_m^{(2)} \\ \vdots \\ y_m^{(L_m)} \end{bmatrix} \in Y_m \subseteq \mathbb{R}^{L_m}, \theta_m = \begin{bmatrix} \theta_m^{(1)} \\ \theta_m^{(2)} \\ \vdots \\ \theta_m^{(R_m)} \end{bmatrix} \in \Theta_m \subseteq \mathbb{R}^{R_m}, \quad (3.2)$$

gdzie: S_m, L_m oraz R_m są – odpowiednio – wymiarami przestrzeni wejść, wyjść i parametrów, $m = 1, 2, \dots, M$.

Przyjmujemy oznaczenie (1.15)

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}, \quad (3.3)$$

gdzie wektor wszystkich wejść u jest elementem przestrzeni

$$U = U_1 \times U_2 \times \cdots \times U_M \subseteq \mathbb{R}^S, \quad S = \sum_{m=1}^M S_m, \quad (3.4)$$

a wektor wszystkich wyjść y jest elementem przestrzeni

$$Y = Y_1 \times Y_2 \times \cdots \times Y_M \subseteq \mathbb{R}^L, \quad L = \sum_{m=1}^M L_m. \quad (3.5)$$

Struktura systemu złożonego jest zadana równaniem

$$u = Ay + Bx, \quad (3.6)$$

w którym: x jest \tilde{S} -wymiarowym wektorem wejść zewnętrznych $x \in X \subseteq U$ oraz $X \subseteq \mathbb{R}^{\tilde{S}}$. Macierz A jest $S \times L$ -wymiarową zero-jedynkową macierzą (1.20), która definiuje powiązania pomiędzy elementami systemu, natomiast zero-jedynkowa $S \times \tilde{S}$ -wymiarowa macierz B wskazuje wejścia zewnętrzne (1.21).

W dalszych rozważaniach zakładamy, że wszystkie wejścia zewnętrzne x oraz wyróżnione wyjścia v są dostępne pomiarowo. Wyróżnione wyjścia v są wskazane przez macierz zero-jedynkową C (1.24), o wymiarze $\tilde{L} \times L$, gdzie \tilde{L} jest liczbą wskazanych, dostępnych wyjść spośród wszystkich L wyjść, tj.

$$v = Cy, \quad (3.7)$$

gdzie mierzone wyjścia

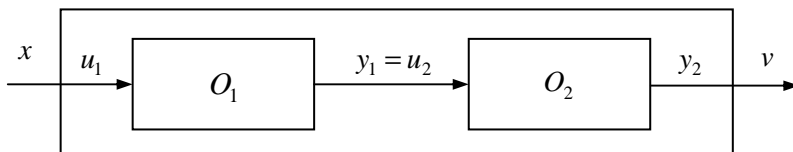
$$v \in V = \{v : \forall y \in Y, v = Cy\} \subseteq \mathbb{R}^{\tilde{L}}. \quad (3.8)$$

Struktura systemu złożonego, opisana zależnością (3.6), wraz z sytuacją pomiarową daną zależnością (3.7), określa nowy obiekt identyfikacji – system złożony jako całość z wejściem zewnętrznym x oraz wyróżnionymi, pomiarowo dostępnymi wyjściami v .

Pojawia się pytanie: Czy na podstawie danej sytuacji pomiarowej możliwe jest jednoznaczne wyznaczenie parametrów opisów poszczególnych elementów?

Podane dalej przykłady ilustrują ograniczone możliwości pomiarowe.

Rozważmy system złożony z dwóch szeregowo połączonych jednowymiarowych elementów (rys. 3.1).



Rys. 3.1. Szeregowe połączenie dwóch elementów

Założymy, że dostępne do pomiaru jest wejście zewnętrzne, którym jest wejście pierwszego elementu, oraz wyjście drugiego elementu, które jest wyjściem całego systemu. W tym przypadku zależność (3.6), która opisuje strukturę systemu, ma postać

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} x = \begin{bmatrix} x \\ y_2 \end{bmatrix}, \quad (3.9)$$

natomiast zależność (3.7)

$$v = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y_2. \quad (3.10)$$

Przykład 3.1. Charakterystyka statyczna pierwszego elementu ma postać

$$y_1 = u_1^{\theta_1}, \quad (3.11)$$

a drugiego

$$y_2 = \theta_2 u_2. \quad (3.12)$$

Po uwzględnieniu zależności (3.9) oraz (3.10) charakterystyka statyczna systemu złożonego (1.31) przyjmuje postać

$$v = \theta_2 x^{\theta_1} = e^{\theta_1 \ln x + \ln \theta_2}, \quad (3.13)$$

gdzie $\theta^T = [\theta_1 \quad \theta_2]$ jest wektorem nieznanymi parametrów.

W celu wyznaczenia wektora parametrów θ wykonano dwa pomiary ($N = 2$), przy założeniach, że: $x_1 > 0$, $x_2 > 0$ oraz $x_1 \neq x_2$.

Z układu równań

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \theta_2 x_1^{\theta_1} \\ \theta_2 x_2^{\theta_1} \end{bmatrix} = \begin{bmatrix} e^{\theta_1 \ln x_1 + \ln \theta_2} \\ e^{\theta_1 \ln x_2 + \ln \theta_2} \end{bmatrix} \quad (3.14)$$

otrzymujemy

$$\begin{bmatrix} \theta_1 \\ \ln \theta_2 \end{bmatrix} = \begin{bmatrix} \frac{\ln v_2 - \ln v_1}{\ln x_2 - \ln x_1} \\ \frac{\ln v_1 \ln x_2 - \ln v_2 \ln x_1}{\ln x_2 - \ln x_1} \end{bmatrix}. \quad (3.15)$$



Przykład 3.2. Załóżmy, że oba elementy są liniowe, czyli

$$y_1 = \theta_1 u_1 \quad (3.16)$$

oraz

$$y_2 = \theta_2 u_2. \quad (3.17)$$

Po uwzględnieniu (3.9) oraz (3.10) charakterystyka statyczna systemu złożonego (1.31) przyjmuje postać

$$v = \theta_1 \theta_2 x, \quad (3.18)$$

gdzie $\theta^T = [\theta_1 \quad \theta_2]$ jest wektorem nieznanymi parametrów.

W celu wyznaczenia wektora parametrów θ wykonano dwa pomiary ($N = 2$) przy założeniu, że $x_1 \neq x_2 \neq 0$.

Z układu równań

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \theta_1 \theta_2 x_1 \\ \theta_1 \theta_2 x_2 \end{bmatrix} \quad (3.19)$$

nie jest możliwe wyznaczenie wartości składowych wektora parametrów elementów systemu $\theta^T = [\theta_1 \quad \theta_2]$. Możliwe jest natomiast wyznaczenie iloczynu parametrów, tj.

$$\theta_1 \theta_2 = \frac{v_n}{x_n}, \quad n = 1, 2. \quad (3.20)$$



Podane przykłady skłaniają do dokładniejszego zbadania odpowiedzi na pytanie: Czy i kiedy na podstawie ograniczonych możliwości pomiarowych można jednoznacznie wyznaczyć parametry charakterystyk statycznych poszczególnych elementów systemu złożonego? Pojawia się problem separowalności, tj. możliwości jednoznacznego wyznaczenia wartości parametrów charakterystyk poszczególnych elementów systemu złożonego. Prowadzi to do tzw. problemu separowalności [25, 27].

3.1. Separowalność deterministyczna

Zwróćmy uwagę, że struktura systemu oraz sytuacja pomiarowa określa nowy obiekt identyfikacji widziany jako całość systemu złożonego, z wektorem wejść zewnętrznych x oraz wyróżnionym, pomiarowo dostępnym, wektorem wyjść v . Przyjmujemy oznaczenie

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} F_1(u_1, \theta_1) \\ F_2(u_2, \theta_2) \\ \vdots \\ F_M(u_M, \theta_M) \end{bmatrix} \stackrel{\text{df}}{=} \bar{F}(u, \theta), \quad (3.21)$$

gdzie θ jest wektorem parametrów, złożonym z wektorów parametrów poszczególnych elementów, tj.

$$\theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_M \end{bmatrix}, \quad (3.22)$$

$$\theta \in \Theta = \Theta_1 \times \Theta_2 \times \dots \times \Theta_M \subseteq \mathbf{R}^R, \quad R = \sum_{m=1}^M R_m. \quad (3.23)$$

Po podstawieniu w miejsce u w równaniu (3.21) zależności (3.6) otrzymujemy

$$y = \bar{F}(Ay + Bx, \theta). \quad (3.24)$$

Rozwiązanie równania (3.24) względem y daje

$$y = \bar{F}_y(x, \theta; A, B), \quad (3.25)$$

przy założeniu, że \bar{F}_y istnieje.

Po podstawieniu równania (3.25) do wzoru (3.7) otrzymujemy

$$v = C\bar{F}_y(x, \theta; A, B) \stackrel{\text{df}}{=} F(x, \theta). \quad (3.26)$$

Zależność (3.26) jest charakterystyką statyczną systemu złożonego, z wektorem wejść zewnętrznych x i wyróżnionym, pomiarowo dostępnym wektorem wyjść v .

Przykład 3.3. Rozważymy system złożony z M elementów. Charakterystyka statyczna m -tego elementu (3.1) ma postać

$$y_m = \bar{\Xi}_m u_m, \quad m = 1, 2, \dots, M, \quad (3.27)$$

gdzie $\bar{\Xi}_m$ jest $L_m \times S_m$ macierzą parametrów, tj.

$$\bar{\Xi}_m = \left[\theta_m^{(l,s)} \right]_{\substack{l=1,2,\dots,L_m \\ s=1,2,\dots,S_m}}. \quad (3.28)$$

Dla rozpatrywanego przykładu zależność (3.21) przyjmuje postać

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} \bar{\Xi}_1 & O & \cdots & O \\ O & \bar{\Xi}_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \bar{\Xi}_M \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}, \quad (3.29)$$

a po przyjęciu oznaczenia

$$\bar{\Xi} \stackrel{\text{df}}{=} \begin{bmatrix} \bar{\Xi}_1 & O & \cdots & O \\ O & \bar{\Xi}_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \bar{\Xi}_M \end{bmatrix}, \quad (3.30)$$

mamy

$$y = \bar{\Xi} u. \quad (3.31)$$

Po podstawieniu zależności (3.6) do równania (3.31) otrzymujemy

$$y = \Xi(Ay + Bx), \quad (3.32)$$

a rozwikłanie (3.32) względem y daje

$$y = (I - \Xi A)^{-1} \Xi Bx, \quad (3.33)$$

przy założeniu, że $(I - \Xi A)$ jest macierzą nieosobliwą.

Ostatecznie charakterystyka statyczna systemu złożonego (3.26), z wektorem wejść zewnętrznym x i pomiarowo dostępnym wektorem wyjść v , przyjmuje postać

$$v = C(I - \Xi A)^{-1} \Xi Bx. \quad (3.34)$$

Ten przykład pokazuje, że system złożony z elementów liniowych daje obiekt liniowy, tj.

$$v = \tilde{\Xi} x, \quad (3.35)$$

gdzie

$$\tilde{\Xi} = C(I - \Xi A)^{-1} \Xi B. \quad (3.36)$$



Zauważmy, że sytuacja pomiarowa, dana zależnością (3.7), dla systemu złożonego o strukturze (3.6), z elementami o charakterystykach statycznych (3.1), określa nowy obiekt identyfikacji o wektorze wejść $x \in X \subseteq \mathbb{R}^{\tilde{s}}$ i wektorze wyjść $v \in V \subseteq \mathbb{R}^{\tilde{l}}$, którego charakterystyka statyczna ma postać

$$v = F(x, \theta), \quad (3.37)$$

gdzie θ jest wektorem parametrów charakterystyki (3.37), a jego składowymi są wektory parametrów charakterystyk statycznych poszczególnych elementów systemu złożonego (3.22), tj. $\theta \in \Theta \subseteq \mathbb{R}^R$ (3.23).

Teraz powraca pytanie: Jakie są możliwości jednoznacznego wyznaczenia parametrów charakterystyk statycznych poszczególnych elementów systemu złożonego przy ograniczonych możliwościach pomiarowych? Charakterystyka statyczna (3.37) obiektu o wektorze wejść zewnętrznych x i wektorze dostępnych do pomiaru wyjść v zależy od wektora θ parametrów wszystkich charakterystyk elementów. Charakterystykę (3.37) otrzymujemy w wyniku przekształcenia (3.26). Przekształcenie to zależy od struktury systemu, charakterystyk poszczególnych elementów oraz możliwości pomiarowych. W konsekwencji wektor parametrów θ ulega pewnemu przekształceniu.

Identyfikowalność obiektu z wejściami zewnętrznymi x oraz dostępnymi do pomiaru wyjściami v nie zawsze jest równoważna możliwości wyznaczenia parametrów poszczególnych elementów. Często na podstawie ograniczonych możliwości pomiarowych możemy wyznaczyć parametry zastępcze, które są funkcją parametrów θ . Dla systemu złożonego pojawia się wtedy potrzeba wprowadzenia dodatkowego pojęcia

określającego możliwości wyznaczenia parametrów poszczególnych elementów, analogicznego do pojęcia identyfikowalności obiektu.

Nowo określony obiekt identyfikacji z charakterystyką statyczną (3.37) umożliwia wprowadzenie pojęcia separowalności systemu złożonego, która jest rozumiana jako możliwość jednoznacznego wyznaczenia parametrów charakterystyk statycznych poszczególnych elementów.

Definicja 3.1. System złożony o zadanej strukturze z elementami o charakterystykach statycznych, znanych z dokładnością do parametrów, nazywamy separowalnym, jeżeli obiekt identyfikacji, określony przez strukturę systemu, jego elementy oraz sytuację pomiarową, jest identyfikowalny.

Korzystając z definicji identyfikowalności (definicja 2.1), możemy stwierdzić, że system złożony o strukturze (3.6) z elementami o charakterystykach statycznych (3.1), przy dostępnych do pomiaru wyjściach wskazanych przez (3.7), jest separowalny, jeśli istnieje taka seria identyfikująca

$$X_N = [x_1 \quad x_2 \quad \cdots \quad x_N], \quad (3.38)$$

która wraz z macierzą wyników pomiarów wartości składowych wektorów wyróżnionych wyjść

$$V_N = [v_1 \quad v_2 \quad \cdots \quad v_N] \quad (3.39)$$

jednoznacznie określa wartość składowych wektora parametrów charakterystyki obiektu (3.27).

Innymi słowy – system złożony jest separowalny, jeżeli istnieje taka seria identyfikująca X_N , która wraz z macierzą wyników pomiarów wyjścia V_N tworzy układ równań

$$v_n = F(x_n, \theta), \quad n = 1, 2, \dots, N, \quad (3.40)$$

dla którego istnieje jednoznaczne rozwiązanie względem θ .

Zauważmy, że ze względu na ograniczone możliwości pomiarowe, identyfikowalność każdego elementu z osobna nie daje identyfikowalności nowo określonego obiektu identyfikacji – systemu złożonego jako całości, tj. obiektu z wektorem wejść zewnętrznym x i wektorem wyjść v .

Wektor parametrów θ charakterystyki statycznej systemu złożonego (3.37) ulega pewnemu przekształceniu. Charakterystykę (3.26) możemy przedstawić w postaci $v = C\bar{F}(x, \theta; A, B) = F(x, \theta) \stackrel{\text{df}}{=} \tilde{F}(x, \tilde{\theta})$, a ostatecznie

$$v = \tilde{F}(x, \tilde{\theta}), \quad (3.41)$$

gdzie: wektor parametrów charakterystyki statycznej (3.41) $\tilde{\theta}$ dany jest zależnością

$$\tilde{\theta} \stackrel{\text{df}}{=} \Gamma(\theta), \quad (3.42)$$

a Γ jest znaną funkcją, taką że $\Gamma: \Theta \rightarrow \tilde{\Theta}$, $\tilde{\Theta} = \{\tilde{\theta} : \forall \theta \in \Theta, \tilde{\theta} = \Gamma(\theta)\} \subseteq \mathbb{R}^{\tilde{R}}$, \tilde{R} jest wymiarowością wektora parametrów charakterystyki statycznej (3.41), \tilde{F} jest natomiast znaną funkcją taką, że $\tilde{F}: X \times \tilde{\Theta} \rightarrow V$. Funkcje \tilde{F} oraz Γ zależą od charakterystyk statycznych poszczególnych elementów (3.1), struktury systemu (3.6) oraz sytuacji pomiarowej (3.7).

Wracając do przykładów, charakterystykę statyczną (3.13) z przykładu 3.1 możemy przedstawić w postaci

$$v = \theta_2 x^{\theta_1} = e^{\theta_1 \ln x + \ln \theta_2} = e^{\tilde{\theta}_1 \ln x + \ln \tilde{\theta}_2}, \quad (3.43)$$

gdzie

$$\tilde{\theta} = \begin{bmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \ln \theta_2 \end{bmatrix}, \quad (3.44)$$

natomiast charakterystyka statyczna (3.18) z przykładu 3.2 ma postać

$$v = \theta_1 \theta_2 x = \tilde{\theta} x, \quad (3.45)$$

gdzie

$$\tilde{\theta} = \theta_1 \theta_2. \quad (3.46)$$

Zależności (3.44) oraz (3.46) przedstawiają przykładowe funkcje Γ , natomiast (3.43) oraz (3.45) – funkcje \tilde{F} . Podobnie zależność (3.36) w przykładzie 3.3. Funkcja Γ ma istotne znaczenie podczas badania separowalności systemu złożonego. Zgodnie z przyjętą definicją 3.1, system złożony z przykładu 3.1 jest separowalny, natomiast system z przykładu 3.2 nie jest separowalny.

Twierdzenie 3.1. Jeżeli obiekt identyfikacji, określony przez system złożony o strukturze (3.6) z elementami o charakterystykach statycznych (3.1) oraz sytuację pomiarową daną przez (3.7), ma charakterystykę statyczną (3.41), jest identyfikowalny oraz funkcja Γ (3.42) jest wzajemnie jednoznaczna względem θ , to system złożony jest separowalny.

Dowód: Ponieważ obiekt identyfikacji o charakterystyce statycznej (3.41) jest identyfikowalny, istnieje taka seria identyfikująca (3.38), która wraz z macierzą wyników pomiarów wyróżnionych wyjść (3.39) jednoznacznie określa parametry charakterystyki statycznej (3.41). Innymi słowy – istnieje taka seria identyfikująca X_N , która wraz z macierzą wyników pomiarów V_N tworzy układ równań

$$v_n = \tilde{F}(x_n, \tilde{\theta}), \quad n = 1, 2, \dots, N, \quad (3.47)$$

dla którego istnieje jednoznaczne rozwiązanie względem $\tilde{\theta}$. Układ równań (3.47) możemy zapisać

$$[v_1 \quad v_2 \quad \dots \quad v_N] = [\tilde{F}(x_1, \tilde{\theta}) \quad \tilde{F}(x_2, \tilde{\theta}) \quad \dots \quad \tilde{F}(x_N, \tilde{\theta})]. \quad (3.48)$$

Po przyjęciu oznaczenia

$$[\tilde{F}(x_1, \theta) \quad \tilde{F}(x_2, \theta) \quad \dots \quad \tilde{F}(x_N, \theta)] \stackrel{\text{df}}{=} \tilde{F}(X_N, \theta), \quad (3.49)$$

układ równań (3.47) ma zwartą postać

$$V_N = \tilde{F}(X_N, \tilde{\theta}), \quad (3.50)$$

a jego rozwiązanie względem wektora parametrów $\tilde{\theta}$ daje algorytm identyfikacji, tj.

$$\tilde{\theta} = \tilde{F}_\theta^{-1}(X_N, V_N) \stackrel{\text{df}}{=} \tilde{\Psi}_N(X_N, V_N), \quad (3.51)$$

gdzie: \tilde{F}_θ^{-1} oznacza rozwiązanie równania (3.50) względem $\tilde{\theta}$, natomiast $\tilde{\Psi}_N$ jest algorytmem identyfikacji.

Ponieważ funkcja Γ jest wzajemnie jednoznaczna, z zależności (3.42) możemy wyznaczyć θ , czyli

$$\theta = \Gamma^{-1}(\tilde{\theta}), \quad (3.52)$$

gdzie Γ^{-1} jest funkcją odwrotną funkcji Γ względem θ .

Po uwzględnieniu algorytmu identyfikacji (3.51) w (3.52) otrzymujemy algorytm identyfikacji

$$\theta = \Gamma^{-1}(\tilde{\Psi}_N(X_N, Y_N)) = \Psi_N(X_N, Y_N) \quad (3.53)$$

dla systemu złożonego z wejściami zewnętrznymi x oraz dostępnymi do pomiaru wyjściami v i charakterystyką statyczną (3.26). Tym samym pokazano, że obiekt identyfikacji, określony przez sytuację pomiarową daną (3.7) oraz system złożony o strukturze (3.6) z elementami o charakterystykach statycznych (3.1), jest identyfikowalny, a tym samym, zgodnie z definicją 3.1, system złożony jest separowalny.

Q.E.D.

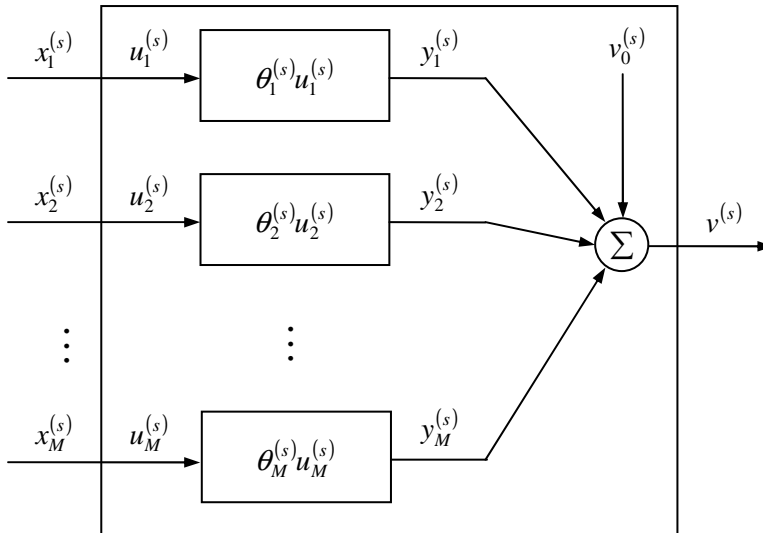
Przykład 3.4. Problem dotyczy badań związanych z ochroną środowiska [139]. W wybranych punktach terenu badane są zanieczyszczenia poprzez monitorowanie stężenia wybranych pierwiastków oraz związków chemicznych. Dalej wybrane, szkodliwe pierwiastki oraz związki chemiczne będziemy nazywać składnikami zanieczyszczeń. Zanieczyszczenia te są emitowane przez różne źródła, m.in. zakłady przemysłowe ulokowane w terenie. Wynik pomiaru jest sumą zanieczyszczeń pochodzących z poszczególnych źródeł (rys. 3.2). Pojawia się pytanie: Jaka jest wielkość emisji poszczególnych źródeł?

Dla przedstawionego problemu przyjmujemy następujący uproszczony opis

$$y_{mn}^{(s)} = \theta_m^{(s)} u_{mn}^{(s)}, \quad s = 1, 2, \dots, S, \quad m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N, \quad (3.54)$$

gdzie: S oznacza liczbę monitorowanych składników zanieczyszczeń, M – liczbę źródeł zanieczyszczeń, N – liczbę punktów pomiarowych, $\theta_m^{(s)}$ – wielkość emisji s -tego składnika zanieczyszczeń przez m -te źródło, $u_{mn}^{(s)}$ – współczynnik propagacji s -tego składnika

zanieczyszczeń z m -tego źródła, który zależy od położenia źródła w stosunku do punktu pomiarowego, ukształtowania terenu, „róży wiatrów”, pogody itp. (zakładamy, że te wielkości charakterystyczne dla badanego terenu są znane), $y_{mn}^{(s)}$ – udział s -tego składnika zanieczyszczeń emitowanego przez m -te źródło w n -tym punkcie pomiarowym.



Rys. 3.2. Struktura systemu monitorowania zanieczyszczeń

Pomiar wielkości udziału zanieczyszczeń pochodzących z poszczególnych źródeł nie jest możliwy. Wynik pomiaru jest sumą zanieczyszczeń pochodzących ze wszystkich źródeł (rys. 3.2), a zatem

$$v_n^{(s)} = \sum_{m=1}^M y_{mn}^{(s)} = \sum_{m=1}^M \theta_m^{(s)} x_{mn}^{(s)} + v_{0n}^{(s)}, \quad s = 1, 2, \dots, S, \quad n = 1, 2, \dots, N, \quad (3.55)$$

gdzie: $v_n^{(s)}$ oznacza stężenie s -tego składnika zanieczyszczeń w n -tym punkcie pomiarowym, $v_{0n}^{(s)}$ – znane stałe tło s -tego składnika zanieczyszczeń w n -tym punkcie pomiarowym, a $x_{mn}^{(s)} = u_{mn}^{(s)}$ – wejścia zewnętrzne systemu, które są współczynnikami propagacji.

Ponownie pojawia się pytanie: Czy na podstawie tak określonych warunków pomiarowych możliwe jest jednoznaczne wyznaczenie wielkości emisji składników zanieczyszczeń przez poszczególne źródła? Odpowiedź na to pytanie zależy od liczby punktów pomiarowych. W układzie (3.55) mamy NS równań, a $R = MS$ niewiadomych $\theta_m^{(s)}$. Przy odpowiednio dużej liczbie punktów pomiarowych, czyli $N = M$, można zapewnić istnienie rozwiązania. Zwykle jednak, ze względu na koszty instalacji i utrzymania stacji monitorującej, liczba punktów pomiarowych jest znacznie mniejsza od liczby źródeł

zanieczyszczeń, tj. $N < M$, a najczęściej w regionie mamy jedną stację monitorującą, czyli $N = 1$. W tym przypadku układ równań (3.55) nie ma jednoznacznego rozwiązania względem θ . Obiekt określony przez system złożony oraz sytuację pomiarową nie jest identyfikowalny, czyli system złożony nie jest separowalny. W takiej sytuacji pomiarowej nie możemy jednoznacznie wyznaczyć wielkości składników zanieczyszczeń emitowanych przez poszczególne źródła. Czy do końca jesteśmy bezradni?

Bardzo często znana jest charakterystyka źródła zanieczyszczeń. Wiadomo, że emitowane zanieczyszczenia mają specyficzny skład i dla każdego źródła znane są proporcje pomiędzy poszczególnymi składnikami zanieczyszczeń, czyli znana jest zależność

$$\theta_m^{(1)} : \theta_m^{(2)} : \dots : \theta_m^{(S)} = \bar{\theta}_m^{(1)} : \bar{\theta}_m^{(2)} : \dots : \bar{\theta}_m^{(S)}, m = 1, 2, \dots, M, \quad (3.56)$$

lub

$$\frac{\theta_m^{(s)}}{\sum_{s=1}^S \theta_m^{(s)}} = \frac{\bar{\theta}_m^{(s)}}{\sum_{s=1}^S \bar{\theta}_m^{(s)}}, s = 1, 2, \dots, S, m = 1, 2, \dots, M, \quad (3.57)$$

w której: $\bar{\theta}_m^{(1)}, \bar{\theta}_m^{(2)}, \dots, \bar{\theta}_m^{(S)}$, $m = 1, 2, \dots, M$, $s = 1, 2, \dots, S$ są znanymi wartościami.

Ta dodatkowa informacja aprioryczna pozwala na jednoznaczne wyznaczenie wielkości emisji składników zanieczyszczeń, czyli na uzyskanie separowalności. Zależność (3.56) zawiera $(S-1)M$ równań i nawet przy jednym pomiarze ($N = 1$), wraz układem równań (3.55), istnieje jednoznaczne rozwiązanie, przy założeniu, że liczba monitorowanych składników zanieczyszczeń jest równa liczbie źródeł, czyli $S = M$. Ogólnie liczba równań z układów (3.55) i (3.56) powinna zapewnić wyznaczenie MS parametrów, a zatem $MS \leq (NS) + (S-1)M$, co daje warunek

$$M \leq NS \quad (3.58)$$

wiążący liczbę punktów pomiarowych z liczbą źródeł zanieczyszczeń i liczbą monitorowanych składników zanieczyszczeń pozwalających uzyskać separowalność rozpatrywanego systemu złożonego. ◆

Podany przykład pokazuje, że dodatkowa informacja o badanym systemie pozwala uzyskać separowalność nawet wtedy, gdy sytuacja pomiarowa dla danego systemu złożonego czyni system nieseparowalnym. Precyzyjniej – gdy funkcja Γ w zależności (3.42) nie jest wzajemnie jednoznaczna, a tak jest wówczas, gdy $\tilde{R} < R$, dodatkowa informacja aprioryczna postaci

$$\bar{\theta} \stackrel{\text{df}}{=} \bar{\Gamma}(\theta), \quad (3.59)$$

gdzie $\bar{\Gamma}$ jest znaną funkcją, taką że $\bar{\Gamma} : \Theta \rightarrow \bar{\Theta}$, $\bar{\Theta} = \{\bar{\theta} : \forall \theta \in \Theta, \bar{\theta} = \bar{\Gamma}(\theta)\} \subseteq \mathbb{R}^{R-\tilde{R}}$, pozwala uzyskać separowalność systemu złożonego.

Wprowadzamy oznaczenie

$$\tilde{\theta} \stackrel{\text{df}}{=} \begin{bmatrix} \tilde{\theta} \\ \bar{\theta} \end{bmatrix} = \begin{bmatrix} \Gamma(\theta) \\ \bar{\Gamma}(\theta) \end{bmatrix} \stackrel{\text{df}}{=} \tilde{\Gamma}(\theta). \quad (3.60)$$

Twierdzenie 3.2. Jeżeli obiekt identyfikacji, określony przez system złożony o strukturze (3.6) z elementami o charakterystykach statycznych (3.1) oraz sytuację pomiarową daną zależnością (3.7), ma charakterystykę statyczną (3.41), jest identyfikowalny, a funkcja $\tilde{\Gamma}$ (3.60) jest wzajemnie jednoznaczna względem θ , to system złożony jest separowalny.

Dowód: Podobnie jak w twierdzeniu 3.1, zauważamy, że jeżeli obiekt identyfikacji o charakterystyce statycznej (3.41) jest identyfikowalny, to istnieje taka seria identyfikująca (3.38), która wraz z wynikami pomiarów (3.39) jednoznacznie określa parametry charakterystyki statycznej (3.41). Innymi słowy – istnieje algorytm identyfikacji (3.51) umożliwiający wyznaczenie wektora parametrów $\tilde{\theta}$, czyli

$$\tilde{\theta} = \tilde{\Psi}_N(X_N, V_N), \quad (3.61)$$

gdzie $\tilde{\Psi}_N$ jest algorytmem identyfikacji.

Ponieważ funkcja $\tilde{\Gamma}$ jest wzajemnie jednoznaczna, z zależności (3.60) możemy wyznaczyć θ , czyli

$$\theta = \tilde{\Gamma}^{-1}(\tilde{\theta}) = \tilde{\Gamma}^{-1}\left(\begin{bmatrix} \tilde{\theta} \\ \bar{\theta} \end{bmatrix}\right), \quad (3.62)$$

gdzie $\tilde{\Gamma}^{-1}$ jest funkcją odwrotną $\tilde{\Gamma}$.

Po uwzględnieniu algorytmu identyfikacji (3.61) w (3.62) otrzymujemy algorytm identyfikacji

$$\theta = \tilde{\Gamma}^{-1}\left(\begin{bmatrix} \tilde{\Psi}_N(X_N, V_N) \\ \bar{\theta} \end{bmatrix}\right) \stackrel{\text{df}}{=} \tilde{\Psi}_N(X_N, V_N, \bar{\theta}) \quad (3.63)$$

dla obiektu identyfikacji z wejściami zewnętrznymi x , pomiarowo dostępnymi wyjściami v i charakterystyką statyczną (3.26) oraz informacją aprioryczną (3.59). Tym samym pokazano, że obiekt określony przez sytuację pomiarową daną (3.7) oraz system złożony o strukturze (3.6), z elementami o charakterystykach statycznych (3.1), jest identyfikowalny, a zatem – zgodnie z definicją 3.1 – system złożony jest separowalny.

Q.E.D.

W podsumowaniu należy zwrócić uwagę na dwa przypadki separowalności systemów złożonych. Pierwszy – w którym wyłącznie wyniki pomiarów zapewniają sepa-

rowalność (twierdzenie 3.1), i drugi – w którym pomiary wraz z informacją aprioryczną umożliwiają jednoznaczne określenie parametrów elementów charakterystyk statycznych systemu złożonego (twierdzenie 3.2).

3.2. Separowalność probabilistyczna

Teraz rozważymy problem separowalności w warunkach probabilistycznych. Omówimy przypadki estymacji parametrów charakterystyk statycznych poszczególnych elementów systemu złożonego oraz charakterystyki statycznej obiektu określonego przez system złożony oraz sytuację pomiarową. Są to zadania analogiczne do tych, które przedstawiliśmy w podrozdziale 2.2, dotyczących estymacji parametrów obiektu, z uwzględnieniem zakłóceń pomiarowych oraz wielkości losowych, występujących w obiekcie. Zanim przejdziemy do omówienia wspomnianych przypadków, wprowadzimy odpowiednie pojęcia związane z separowalnością probabilistyczną [25–27, 92, 95], które korespondują z estymowalnością obiektu (definicje 2.2 i 2.3).

Definicja 3.2. System złożony o zadanej strukturze z elementami o charakterystykach statycznych znanych z dokładnością do parametrów nazywamy probabilistycznie separowalnym w sensie określonej własności (określoną metodą), jeżeli obiekt identyfikacji określony przez strukturę systemu oraz sytuację pomiarową jest estymowalny w sensie określonej własności (określoną metodą).

W przedstawionych dalej rozważaniach zwrócimy uwagę na probabilistyczną separowalność systemu złożonego. W zależności od informacji, jaką posiadamy o zidentyfikowanym obiekcie i systemie pomiarowym, oraz od warunków, jakie spełniają zakłócenia, możemy stosować różne metody estymacji. W dalszych punktach skupimy się na metodzie maksymalnej wiarygodności oraz metodach Bayesa, co koresponduje z rozważaniami przedstawionymi w podrozdziale 3.1, gdzie do wyznaczenia parametrów systemu złożonego wykorzystaliśmy wyłącznie informację pomiarową, a następnie uzupełniliśmy ją o informację aprioryczną.

3.2.1. Estymacja parametrów charakterystyk statycznych systemu złożonego na podstawie zakłóconych pomiarów wyróżnionych wyjść

Rozważymy przypadek, kiedy wyróżnione wyjścia systemu złożonego o zadanej strukturze są mierzone z zakłóceniami pomiarowymi. Innymi słowy – rozpatrujemy wejściowo-wyjściowy system złożony, w którym wyróżniono M elementów O_1, O_2, \dots, O_M , dla których znana jest charakterystyka statyczna postaci (3.1) z dokładnością do parametrów. Struktura systemu dana jest zależnością (3.6). W celu wyznaczenia nieznanymi parametrów charakterystyk statycznych elementów systemu złożonego dla zadanej serii wartości wejść zewnętrznych zmierzono z zakłóceniami wartości wyróżnionych wyjść. Wyróżnione wyjścia są opisane równaniem (3.7). W konsekwencji

charakterystyka statyczna systemu złożonego, z wektorem wejść zewnętrznych x i wektorem wyróżnionych, pomiarowo dostępnych wyjść v , dana jest zależnością (3.26). Zakładamy, tak jak w punkcie 2.2.1, że znany jest opis systemu pomiarowego i w wyniku pomiaru wyjść v z zakłóceniami \tilde{z} otrzymujemy

$$\tilde{w} = \tilde{h}(v, \tilde{z}), \quad (3.64)$$

gdzie: \tilde{h} jest znaną funkcją, taką że $\tilde{h}: V \times Z^{\sim} \rightarrow W^{\sim}$, $\tilde{z} \in Z^{\sim}$, $\tilde{w} \in W^{\sim}$, Z^{\sim} jest przestrzenią zakłóceń, a W^{\sim} – przestrzenią wyniku zakłóconego pomiaru.

W dalszych rozważaniach będziemy zakładać, że wymiar wektora y oraz zakłóceń z jest taki sam i równy L ($\dim v = \dim \tilde{z} = L$). Funkcja \tilde{h} jest ponadto wzajemnie jednoznaczna względem \tilde{z} . Znaczy to, że istnieje funkcja odwrotna \tilde{h}_z^{-1} względem \tilde{z} , czyli

$$\tilde{z} = \tilde{h}_z^{-1}(v, \tilde{w}). \quad (3.65)$$

Tak jak w punkcie 2.2.1, założenia dotyczące zarówno wymiarowości wektora wyjść oraz zakłóceń, jak i różnowartościowości funkcji \tilde{h} można osłabić. Wymaga to dodatkowych przekształceń przedstawionych w Dodatku.

W celu wyznaczenia nieznanymi parametrów charakterystyk statycznych elementów systemu złożonego wykonaliśmy eksperyment. W n -tym pomiarze, dla zadanych wartości wejść zewnętrznych x_n , dokonaliśmy pomiaru wartości wyróżnionych wyjść v_n . Na mierzoną wielkość nałożyły się przypadkowe zakłócenia i w wyniku pomiaru otrzymujemy wartość \tilde{w}_n . O zakłóceniach \tilde{z}_n w n -tym pomiarze założymy, że są niezależnymi wartościami zmiennej losowej \tilde{z} . W dalszych rozważaniach przyjmujemy, że zmienna losowa \tilde{z} przyjmuje wartości ze zbioru ciągłego Z^{\sim} , a funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \tilde{z} oznaczymy przez $f_{\tilde{z}}(\tilde{z})$. W wyniku eksperymentu, dla zadanej serii identyfikującej o długości N , uzyskujemy wyniki pomiarów, zebrane w macierzach:

$$X_N = [x_1 \quad x_2 \quad \cdots \quad x_N], \quad \tilde{W}_N = [\tilde{w}_1 \quad \tilde{w}_2 \quad \cdots \quad \tilde{w}_N]. \quad (3.66)$$

Zadanie identyfikacji sprowadza się do wyznaczenia oszacowania wektora nieznanymi parametrów θ (3.22), którego składowymi są wektory parametrów charakterystyk statycznych poszczególnych elementów systemu złożonego. Poszukujemy algorytmu estymacji

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N), \quad (3.67)$$

który dla danych pomiarowych X_N, \tilde{W}_N pozwoli wyznaczyć wartość oszacowania $\hat{\theta}_N$ wektora θ , gdzie Ψ_N oznacza algorytm estymacji dla serii pomiarowych o długości N .

Metoda maksymalnej wiarygodności

Korzystając z metody maksymalnej wiarygodności (jak w punkcie 2.2.1 – zadanie (2.36)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, dla której funkcja wiarygodności L_N , wyznaczona na podstawie macierzy pomiarów wyjścia \tilde{W}_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) : L_N(\tilde{W}_N, \hat{\theta}_N; X_N) = \max_{\theta \in \Theta} L_N(\tilde{W}_N, \theta; X_N), \quad (3.68)$$

gdzie funkcja wiarygodności ma postać

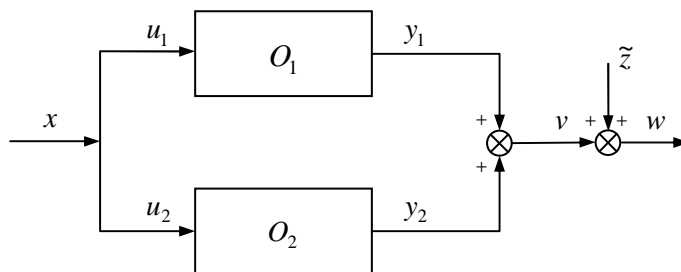
$$L_N(\tilde{W}_N, \theta; X_N) = \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(F(x_n, \theta), \tilde{w}_n)) |J_{\tilde{h}}|, \quad (3.69)$$

a jacobian przekształcenia odwrotnego (3.65) wyraża się zależnością

$$J_{\tilde{h}} = \frac{\partial \tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w})}{\partial \tilde{w}}. \quad (3.70)$$

Przykład 3.5. Rozważmy system złożony z równolegle połączonych dwóch liniowych, jednowymiarowych elementów O_1 i O_2 z nieznanymi parametrami θ_1 oraz θ_2 (rys. 3.3). Załóżmy, że pomiarowo dostępne jest wejście zewnętrzne x oraz wartość wyjścia v , która jest sumą wartości wyjść poszczególnych elementów. W wyniku pomiaru wyróżnionego wyjścia v z addytywnymi zakłóceniami otrzymujemy

$$\tilde{w} = h(v, \tilde{z}) = v + \tilde{z}. \quad (3.71)$$



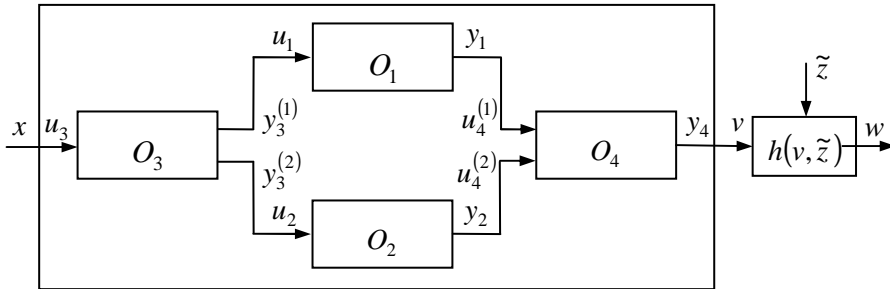
Rys. 3.3. Równoległe połączenie dwóch elementów

W kolejnych pomiarach zakłócenia są niezależnymi wartościami zmiennej losowej \tilde{z} o rozkładzie normalnym z wartością oczekiwaną $m_{\tilde{z}}$ i odchyleniem standardowym $\sigma_{\tilde{z}}$, czyli

$$f_{\tilde{z}}(\tilde{z}) = \frac{1}{\sigma_{\tilde{z}} \sqrt{2\pi}} \exp \left[-\frac{(\tilde{z} - m_{\tilde{z}})^2}{2\sigma_{\tilde{z}}^2} \right]. \quad (3.72)$$

W wyniku eksperymentu, dla zadanej serii identyfikującej X_N o długości N , uzyskaliśmy zakłócone wyniki pomiarów wyjść, zebrane w macierzy \tilde{W}_N (3.66).

W celu zachowania konwencji opisu struktury systemu złożonego postaci (3.6) wprowadzimy dwa dodatkowe elementy (rys. 3.4). Pierwszy (O_3), na wejściu systemu złożonego, który powiela wejście zewnętrzne, oraz drugi (O_4), który jest sumatorem.



Rys. 3.4. Równoległe połączenie dwóch elementów

W rozpatrywanym przypadku zależność (3.6), która opisuje strukturę systemu, ma postać

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4^{(1)} \\ u_4^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3^{(1)} \\ y_3^{(2)} \\ y_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} x = \begin{bmatrix} y_3^{(1)} \\ y_3^{(2)} \\ x \\ y_1 \\ y_2 \end{bmatrix}, \quad (3.73)$$

a pomiarowo dostępne, wyróżnione wyjścia dane są równaniem (3.7) postaci

$$v = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3^{(1)} \\ y_3^{(2)} \\ y_4 \end{bmatrix}. \quad (3.74)$$

Po uwzględnieniu charakterystyk statycznych poszczególnych elementów systemu złożonego równanie (3.21), podające zależność wszystkich wyjść systemu w funkcji wszystkich wejść, przyjmuje postać

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3^{(1)} \\ y_3^{(2)} \\ y_4 \end{bmatrix} = \begin{bmatrix} \theta_1 u_1 \\ \theta_2 u_2 \\ u_1 \\ u_2 \\ u_4^{(1)} + u_4^{(2)} \end{bmatrix}. \quad (3.75)$$

Ostatecznie charakterystyka statyczna systemu złożonego (3.26), z wejściem zewnętrznym x oraz wyróżnionym, dostępnym do pomiaru, wyjściem v , przyjmuje postać

$$v = (\theta_1 + \theta_2)x. \quad (3.76)$$

W rozpatrywanym przykładzie funkcja odwrotna \tilde{h}_z^{-1} , z uwzględnieniem charakterystyki statycznej systemu złożonego (3.26) postaci (3.76), dana jest zależnością

$$\tilde{z} = \tilde{h}_z^{-1}(F(x, \theta), w) = w - (\theta_1 + \theta_2)x. \quad (3.77)$$

Jakobian (3.70) przekształcenia odwrotnego (3.71)

$$J_{\tilde{h}} = \frac{\partial \tilde{h}_z^{-1}(F(x, \theta), w)}{\partial w} = \frac{d}{dw}(w - (\theta_1 + \theta_2)x) = 1. \quad (3.78)$$

Funkcja wiarygodności (3.69) w rozpatrywanym przykładzie ma postać

$$\tilde{L}_N(\tilde{W}_N, \theta; X_N) = \prod_{n=1}^N \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(w_n - (\theta_1 + \theta_2)x_n - m_z)^2}{2\sigma_z^2} \right], \quad (3.79)$$

a po przekształceniach

$$\tilde{L}_N(\tilde{W}_N, \theta; X_N) = \left(\frac{1}{\sigma_z \sqrt{2\pi}} \right)^N \exp \left[\sum_{n=1}^N -\frac{(w_n - (\theta_1 + \theta_2)x_n - m_z)^2}{2\sigma_z^2} \right]. \quad (3.80)$$

Maksymalizacja funkcji wiarygodności w tym przypadku sprowadza się do optymalizacji wykładnika wyrażenia (3.80). Przyrównanie gradientu wykładnika wyrażenia (3.80) względem θ do zera daje układ równań

$$\text{grad}_{\theta} L_N(W_N, \theta; X_N) = \begin{bmatrix} \sum_{n=1}^N \frac{(w_n - (\theta_1 + \theta_2)x_n - m_z)x_n}{\sigma_z^2} \\ \sum_{n=1}^N \frac{(w_n - (\theta_1 + \theta_2)x_n - m_z)x_n}{\sigma_z^2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (3.81)$$

Układ równań (3.81) nie ma jednoznacznego rozwiązania względem θ_1 oraz θ_2 . W zależności (3.76) możemy wprowadzić parametr zastępczy, który jest sumą po-

szczególnych parametrów, i wówczas – w wyniku rozwiązania równania (3.81) – uzyskamy wartość oszacowania $\hat{\theta}_N$ parametru zastępczego

$$\tilde{\theta} = \theta_1 + \theta_2, \quad (3.82)$$

a algorytm estymacji wyraża się wzorem

$$\hat{\theta}_N = \frac{\sum_{n=1}^N (w_n - m_{\tilde{z}}) x_n}{\sum_{n=1}^N x_n^2}. \quad (3.83)$$

◆

W rozpatrywanym przykładzie system nie jest separowalny probabilistycznie metodą maksymalnej wiarygodności.

Przez analogię do sytuacji deterministycznej, przedstawionej w podrozdziale 3.1, rozważymy przypadek, kiedy charakterystykę statyczną (3.26), a tym samym (3.37), systemu złożonego, z wektorem wejść zewnętrznych x i pomiarowo dostępnym wektorem wyróżnionych wyjść v , możemy przedstawić w postaci (3.41), z wektorem parametrów zastępczych $\tilde{\theta}$, określonym zależnością (3.42), czyli

$$v = C\bar{F}_y(x, \theta; A, B) = F(x, \theta) \stackrel{\text{df}}{=} \tilde{F}(x, \tilde{\theta}), \quad (3.84)$$

gdzie

$$\tilde{\theta} \stackrel{\text{df}}{=} \Gamma(\theta). \quad (3.85)$$

Postać funkcji Γ ma istotne znaczenie dla badania separowalności. W przykładzie 3.5 system złożony z charakterystyką statyczną (3.26), czyli obiekt identyfikacji z wektorem wejść zewnętrznych x oraz pomiarowo dostępnym wektorem wyróżnionych wyjść v , przyjmuje postać (3.76). W tym przykładzie możemy określić parametr zastępczy $\tilde{\theta}$, który jest sumą parametrów poszczególnych elementów. Zależność (3.42) ma postać (3.82). Funkcja ta nie jest wzajemnie jednoznaczna względem θ_1 i θ_2 . Spostrzeżenie z przykładu 3.5 łatwo można uogólnić. Po uwzględnieniu charakterystyki statycznej systemu złożonego (3.84) funkcja wiarygodności (3.69) ma postać

$$\begin{aligned} L_N(\tilde{W}_N, \theta; X_N) &= \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(\tilde{F}(x_n, \Gamma(\theta)), \tilde{w}_n)) |J_{\tilde{h}}| \\ &= \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(\tilde{F}(x_n, \tilde{\theta}), \tilde{w}_n)) |J_{\tilde{h}}| \stackrel{\text{df}}{=} \tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N). \end{aligned} \quad (3.86)$$

Oszacowanie wektora parametrów zastępczych $\tilde{\theta}$ uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$, dla której funkcja wiarygodności \tilde{L}_N (3.86), wyznaczona na podstawie zbioru N ob-

serwacji wyjścia \tilde{W}_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem $\tilde{\theta}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, czyli

$$\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) : \tilde{L}_N(\tilde{W}_N, \hat{\theta}_N; X_N) = \max_{\tilde{\theta} \in \tilde{\Theta}} \tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N). \quad (3.87)$$

Jeżeli funkcja Γ jest wzajemnie jednoznaczna, to znaczy, że istnieje funkcja odwrotna Γ^{-1} , przy czym Γ^{-1} jest ciągła, to z istnienia estymatora zgodnego dla $\tilde{\theta}$ wynika probabilistyczna separowalność w sensie zgodności dla systemu złożonego. Rzeczywiście, jeśli $\hat{\theta}_N$ (estymator dla $\tilde{\theta}$) zdąży stochastycznie do $\tilde{\theta}$, to $\hat{\theta}_N = \Gamma^{-1}(\hat{\theta}_N)$ zdąży stochastycznie do $\theta = \Gamma^{-1}(\tilde{\theta})$. Podobnie, z istnienia funkcji odwrotnej Γ^{-1} oraz estymowalności obiektu z charakterystyka statyczną (3.26) metodą maksymalnej wiarygodności wynika probabilistyczna separowalność systemu złożonego tą metodą. Jeśli bowiem $\tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N)$ jest funkcją wiarygodności dla $\tilde{\theta}$ (dla ustalonej serii identyfikującej X_N), to funkcją wiarygodności dla θ jest $L_N(\tilde{W}_N, \theta; X_N)$ i funkcje te, zgodnie z (3.86), są sobie równe, czyli

$$L_N(\tilde{W}_N, \theta; X_N) = \tilde{L}_N(\tilde{W}_N, \Gamma(\theta); X_N). \quad (3.88)$$

Łatwo więc zauważyć, że jeśli $\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$ maksymalizujące \tilde{L}_N (3.87) jest jednoznacznie określone przez macierze pomiarów X_N oraz \tilde{W}_N , to $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$ maksymalizujące L_N (3.68) jest także jednoznacznie określone przez X_N oraz \tilde{W}_N i zachodzi równość

$$\hat{\theta}_N = \Gamma^{-1}(\hat{\theta}_N). \quad (3.89)$$

Twierdzenie 3.3. Jeżeli obiekt identyfikacji, określony przez system złożony o strukturze (3.6), z elementami o charakterystykach statycznych (3.1), oraz sytuację pomiarową daną przez (3.7) i (3.64), ma charakterystykę statyczną (3.41) i funkcja Γ (3.42) nie jest wzajemnie jednoznaczna względem θ , to system złożony nie jest probabilistycznie separowalny metodą maksymalnej wiarygodności.

Dowód: Jeżeli obiekt identyfikacji określony przez system złożony oraz sytuację pomiarową jest estymowalny metodą maksymalnej wiarygodności, to zgodnie z (3.88) dla ustalonej wartości $\hat{\theta}_N = \Gamma(\theta)$ maksymalizującej $\tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N)$ względem $\tilde{\theta}$ istnieją różne wartości $\hat{\theta}_N$ maksymalizujące $L_N(\tilde{W}_N, \theta; X_N)$ względem θ , gdyż funkcja Γ nie jest wzajemnie jednoznaczna. Jeśli obiekt określony przez system złożony oraz sytuację pomiarową nie jest estymowalny metodą maksymalnej wiarygodności, to nie istnieje

je jedna wartość $\hat{\theta}_N$ maksymalizująca $L_N(\tilde{W}_N, \theta; X_N)$ względem θ , bo wówczas istniałaby jedna wartość $\hat{\tilde{\theta}}_N = \Gamma(\theta)$ maksymalizująca $\tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N)$ względem $\tilde{\theta}$.

Metody Bayesa

Przez analogię do przypadku deterministycznego, po uzupełnieniu opisu systemu złożonego o dodatkową informację możemy uzyskać separowalność. W przypadku probabilistycznym taką dodatkową informacją (analogiczną do (3.59)) może być założenie, że obserwowany system złożony został wylosowany z pewnej populacji i znana jest funkcja gęstości rozkładu prawdopodobieństwa parametrów systemu. Takie założenie prowadzi do bayesowskich metod estymacji. W konsekwencji może się okazać, że stosując bayesowskie metody estymacji, uzyskamy separowalność, mimo że system złożony nie jest probabilistycznie separowalny innymi metodami.

Teraz, dodatkowo w stosunku do poprzednich rozważań, założymy, że wektor parametrów obiektu θ jest realizacją ciągłej zmiennej losowej $\underline{\theta}$, która przyjmuje wartości ze zbioru $\Theta \subseteq \mathbb{R}^R$. Znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$ a priori zmiennej $\underline{\theta}$, określona na Θ . Korzystając z metod Bayesa (tak jak w punkcie 2.2.1 – zadanie (2.65)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, dla której ryzyko warunkowe, wyznaczone na podstawie macierzy pomiarów \tilde{W}_N przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) : r(\hat{\theta}_N, \tilde{W}_N; X_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, \tilde{W}_N; X_N). \quad (3.90)$$

gdzie ryzyko warunkowe ma postać

$$r(\bar{\theta}, \tilde{W}_N; X_N) \stackrel{\text{df}}{=} E_{\underline{\theta}} [L(\underline{\theta}, \bar{\theta}) | \tilde{W}_N; X_N] = \int_{\Theta} L(\theta, \bar{\theta}) f'(\theta | \tilde{W}_N; X_N) d\theta, \quad (3.91)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori wyraża się wzorem

$$f'(\theta | \tilde{W}_N; X_N) = \frac{f_\theta(\theta) \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(F(x_n, \theta), \tilde{w}_n)) | J_{\tilde{h}} |}{\int_{\Theta} f_\theta(\theta) \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(F(x_n, \theta), \tilde{w}_n)) | J_{\tilde{h}} | d\theta}. \quad (3.92)$$

Przykład 3.6. Rozważmy system złożony jak w przykładzie 3.5. Wzbogacamy wiedzę o systemie o informację aprioryczną. Niech parametry θ_1 oraz θ_2 będą wartościami niezależnych zmiennych losowych $\underline{\theta}_1$ i $\underline{\theta}_2$ o rozkładach normalnych z wartościami oczekiwanymi – odpowiednio – m_1 i m_2 oraz wariancjami σ_1^2 i σ_2^2 .

Przyjmując funkcję strat postaci (2.70), otrzymujemy metodę maksymalnego prawdopodobieństwa a posteriori, a wyznaczenie oszacowania wektora parametrów sprowadza się do maksymalizacji gęstości rozkładu prawdopodobieństwa względem θ . W naszym przykładzie funkcja gęstości rozkładu prawdopodobieństwa a posteriori wyraża się zależnością

$$f'(\theta | \tilde{W}_N; X_N) = \frac{\frac{1}{2\pi\sigma_1\sigma_2} \left(\frac{1}{\sqrt{2\pi}}\right)^N \exp[-G(\theta_1, \theta_2, X_N, \tilde{W}_N)]}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \left(\frac{1}{\sqrt{2\pi}}\right)^N \exp[-G(\theta_1, \theta_2, X_N, \tilde{W}_N)] d\theta_1 d\theta_2}, \quad (3.93)$$

w której

$$G(\theta_1, \theta_2, X_N, \tilde{W}_N) \stackrel{\text{df}}{=} \frac{(\theta_1 - m_1)^2}{2\sigma_1^2} + \frac{(\theta_2 - m_2)^2}{2\sigma_2^2} + \sum_{n=1}^N \frac{(w_n - (\theta_1 + \theta_2)x_n - m_z)^2}{2\sigma_z^2}. \quad (3.94)$$

Oszacowanie nieznanymi parametrów otrzymujemy w wyniku maksymalizacji funkcji gęstości rozkładu (3.93), co sprowadza się do minimalizacji wyrażenia (3.94) względem parametrów θ_1 oraz θ_2 . Przyrównując pochodne funkcji (3.94) względem θ_1 oraz θ_2 do zera, otrzymujemy układ równań

$$\begin{aligned} & \left[\begin{array}{c} \frac{\partial G(\theta_1, \theta_2, X_N, \tilde{W}_N)}{\partial \theta_1} \\ \frac{\partial G(\theta_1, \theta_2, X_N, \tilde{W}_N)}{\partial \theta_2} \end{array} \right]_{\substack{\theta_1 = \hat{\theta}_{1N} \\ \theta_2 = \hat{\theta}_{2N}}} \\ & = \left[\begin{array}{c} \frac{\hat{\theta}_{1N} - m_1}{\sigma_1^2} - \sum_{n=1}^N \frac{(w_n - (\hat{\theta}_{1N} + \hat{\theta}_{2N})x_n - m_z)x_n}{\sigma_z^2} \\ \frac{\hat{\theta}_{2N} - m_2}{\sigma_2^2} - \sum_{n=1}^N \frac{(w_n - (\hat{\theta}_{1N} + \hat{\theta}_{2N})x_n - m_z)x_n}{\sigma_z^2} \end{array} \right] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned} \quad (3.95)$$

z którego wyznaczamy oszacowanie $\hat{\theta}_{1N}$ i $\hat{\theta}_{2N}$

$$\begin{bmatrix} \hat{\theta}_{1N} \\ \hat{\theta}_{2N} \end{bmatrix} = \begin{bmatrix} m_1 + \sigma_1^2 \frac{\sum_{n=1}^N (w_n - m_z)x_n - (m_1 + m_2) \sum_{n=1}^N x_n^2}{\sigma_z^2 + (\sigma_1^2 + \sigma_2^2) \sum_{n=1}^N x_n^2} \\ m_2 + \sigma_2^2 \frac{\sum_{n=1}^N (w_n - m_z)x_n - (m_1 + m_2) \sum_{n=1}^N x_n^2}{\sigma_z^2 + (\sigma_1^2 + \sigma_2^2) \sum_{n=1}^N x_n^2} \end{bmatrix}. \quad (3.96)$$

Teraz otrzymujemy jednoznaczne wartości oszacowania $\hat{\theta}_{1N}$ i $\hat{\theta}_{2N}$, których nie otrzymaliśmy metodą maksymalnej wiarygodności. ◆

Zwróćmy uwagę na możliwość stosowania różnych podejść do estymacji parametrów charakterystyk statycznych systemu złożonego. Wynik $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, minimalizujący ryzyko (3.91), oznacza wartość oszacowania wektora parametrów charakterystyk statycznych wszystkich elementów systemu złożonego. Biorąc pod uwagę obiekt określony przez system złożony i sytuację pomiarową z opisem postaci $v = \tilde{F}(x, \tilde{\theta})$ (3.84), możemy sformułować zadanie estymacji wektora parametrów zastępczych $\tilde{\theta} = \Gamma(\theta)$. Oszacowanie wektora parametrów zastępczych $\tilde{\theta}$ charakterystyki statycznej systemu złożonego (3.41) uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$, dla której ryzyko warunkowe, dla uzyskanej macierzy pomiarów \tilde{W}_N , przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\tilde{\theta}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, tj.

$$\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) : \tilde{r}(\hat{\tilde{\theta}}_N, \tilde{W}_N; X_N) = \min_{\tilde{\theta} \in \tilde{\Theta}} \tilde{r}(\tilde{\theta}, \tilde{W}_N; X_N), \quad (3.97)$$

gdzie ryzyko warunkowe ma postać

$$\tilde{r}(\tilde{\theta}, \tilde{W}_N; X_N) \stackrel{\text{df}}{=} \mathbb{E}_{\tilde{\theta}} [L(\tilde{\theta}, \tilde{\theta}) | \tilde{W}_N; X_N] = \int_{\tilde{\Theta}} L(\tilde{\theta}, \tilde{\theta}) \tilde{f}'(\tilde{\theta} = \Gamma(\theta) | \tilde{W}_N; X_N) d\tilde{\theta}, \quad (3.98)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori wyraża się wzorem

$$\tilde{f}'(\tilde{\theta} = \Gamma(\theta) | \tilde{W}_N; X_N) = \frac{f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(\tilde{F}(x_n, \tilde{\theta}), \tilde{w}_n)) | J_{\tilde{h}} |}{\int_{\tilde{\Theta}} f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(\tilde{F}(x_n, \tilde{\theta}), \tilde{w}_n)) | J_{\tilde{h}} | d\tilde{\theta}} \quad (3.99)$$

oraz $f_{\tilde{\theta}}(\tilde{\theta})$ – funkcja gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej $\tilde{\theta}$ jest określona przez funkcję Γ i $f_{\theta}(\theta)$ – funkcję gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej θ . Jeśli przekształcenie Γ jest wzajemnie jednoznaczne, czyli istnieje przekształcenie odwrotne względem θ , tj. $\theta = \Gamma^{-1}(\tilde{\theta})$, to

$$f_{\tilde{\theta}}(\tilde{\theta}) = f_{\theta}(\Gamma^{-1}(\tilde{\theta})) | J_{\Gamma} |, \quad (3.100)$$

gdzie jacobian przekształcenia odwrotnego ma postać

$$J_r = \frac{\partial \Gamma^{-1}(\tilde{\theta})}{\partial \tilde{\theta}}. \tag{3.101}$$

Gdy funkcja Γ w opisie (3.42) nie jest wzajemnie jednoznaczna, a tak jest wtedy, gdy $\dim \tilde{\theta} = \tilde{R} < \dim \theta = R$, uzupełnienie o przekształcenie $\bar{\Gamma}$, postaci

$$\bar{\theta} \stackrel{\text{df}}{=} \bar{\Gamma}(\theta), \tag{3.102}$$

jeżeli takie istnieje, pozwala uzyskać wzajemnie jednoznaczne przekształcenie zmiennych losowych (zob. Dodatek). Teraz funkcje Γ i $\bar{\Gamma}$ tworzą przekształcenie $\tilde{\Gamma}$

$$\tilde{\theta} \stackrel{\text{df}}{=} \begin{bmatrix} \tilde{\theta} \\ \bar{\theta} \end{bmatrix} = \begin{bmatrix} \Gamma(\theta) \\ \bar{\Gamma}(\theta) \end{bmatrix} \stackrel{\text{df}}{=} \tilde{\Gamma}(\theta), \tag{3.103}$$

które jest wzajemnie jednoznaczne względem θ , czyli

$$\theta = \tilde{\Gamma}^{-1}(\tilde{\theta}) = \tilde{\Gamma}^{-1} \left(\begin{bmatrix} \tilde{\theta} \\ \bar{\theta} \end{bmatrix} \right), \tag{3.104}$$

gdzie $\tilde{\Gamma}^{-1}$ jest funkcją odwrotną funkcji $\tilde{\Gamma}$.

Ostatecznie funkcję $f_{\tilde{\theta}}(\tilde{\theta})$ gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej $\tilde{\theta}$ wyznaczamy ze wzoru

$$f_{\tilde{\theta}}(\tilde{\theta}) = \int_{\bar{\theta}} f_{\theta} \left(\tilde{\Gamma}^{-1} \left(\begin{bmatrix} \tilde{\theta} \\ \bar{\theta} \end{bmatrix} \right) \right) \left| J_{\tilde{\Gamma}} \right| d\bar{\theta}, \tag{3.105}$$

w którym jacobian przekształcenia odwrotnego (3.104) ma postać

$$J_{\tilde{\Gamma}} = - \frac{\partial \tilde{\Gamma}^{-1}(\tilde{\theta})}{\partial \tilde{\theta}}. \tag{3.106}$$

Przypadek, gdy $\dim \tilde{\theta} = \dim \theta$, czyli $\tilde{R} = R$, a funkcja Γ nie jest różnowartościowa, wymaga dalszych przekształceń, polegających na rozbiściu zbioru \mathcal{O} na rozłączne podzbiory, w których Γ jest różnowartościowa, i zastosowania przekształcenia analogicznego do (3.100) w tych podzbiorych.

Prześledźmy te różne podejścia do estymacji na przykładzie.

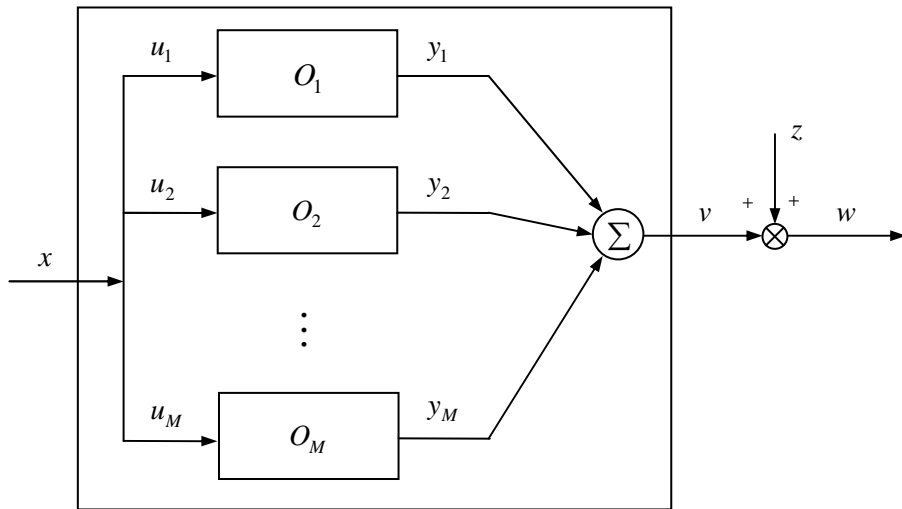
Przykład 3.7. Rozważmy system złożony z M równolegle połączonych liniowych jednowymiarowych elementów O_1, O_2, \dots, O_M z nieznanymi parametrami $\theta_1, \theta_2, \dots, \theta_M$ (rys. 3.5). Załóżmy, że do pomiaru jest dostępne wejście zewnętrzne x oraz wyróż-

nione wyjście v , które jest sumą wyjść poszczególnych elementów. Wyjście v mierzymy z addytywnymi zakłóceniami, czyli

$$\tilde{w} = h(v, \tilde{z}) = v + \tilde{z}. \quad (3.107)$$

W kolejnych pomiarach zakłócenia są niezależnymi wartościami zmiennej losowej \tilde{z} o rozkładzie normalnym z zerową wartością oczekiwaną i odchyleniem standardowym $\sigma_{\tilde{z}}$, czyli

$$f_{\tilde{z}}(\tilde{z}) = \frac{1}{\sigma_{\tilde{z}} \sqrt{2\pi}} \exp\left[-\frac{(\tilde{z})^2}{2\sigma_{\tilde{z}}^2}\right]. \quad (3.108)$$



Rys. 3.5. System złożony o strukturze równoległej

W wyniku eksperymentu, dla zadanej serii identyfikującej X_N o długości N , uzyskaliśmy zakłócone wyniki pomiarów wyjść, zebrane w macierzy \tilde{W}_N (3.66). Dodatkowo zakładamy, że wektor parametrów systemu θ jest wartością zmiennej losowej $\underline{\theta}$ o wielowymiarowym rozkładzie normalnym z M -wymiarowym wektorem wartości oczekiwanych m_θ i macierzą kowariancji Σ_θ , tj.

$$f_\theta(\theta) = (2\pi)^{-\frac{M}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\theta - m_\theta)^T \Sigma_\theta^{-1} (\theta - m_\theta)\right]. \quad (3.109)$$

Charakterystyka statyczna (3.26) rozpatrywanego systemu złożonego, z wejściem zewnętrznym x oraz pomiarowo dostępnym wyjściem v , przyjmuje postać

$$v = F(x, \theta) = \left(\sum_{m=1}^M \theta_m\right) x = \mathbf{1}_M^T \theta x, \quad (3.110)$$

gdzie $\mathbf{1}_M$ oznacza M -wymiarowy, kolumnowy wektor, w którym każda składowa przyjmuje wartość jeden, czyli

$$\mathbf{1}_M = \left. \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} M. \quad (3.111)$$

W **pierwszym podejściu** dla rozpatrywanego systemu złożonego wyznaczmy oszacowanie wektora parametrów charakterystyk statycznych poszczególnych elementów. Po przyjęciu w równaniu (3.91) funkcji strat postaci (2.70) otrzymujemy metodę maksymalnego prawdopodobieństwa a posteriori, a wyznaczenie oszacowania sprowadza się do maksymalizacji gęstości rozkładu prawdopodobieństwa a posteriori (3.92) względem θ . W naszym przykładzie funkcja gęstości rozkładu prawdopodobieństwa a posteriori wyraża się zależnością

$$f'(\theta | \tilde{W}_N; X_N) = \frac{(2\pi)^{-\frac{M+N}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp[-G(\theta, X_N, \tilde{W}_N)]}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (2\pi)^{-\frac{M+N}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp[-G(\theta, X_N, \tilde{W}_N)] d\theta}, \quad (3.112)$$

w której

$$G(\theta, X_N, \tilde{W}_N) \stackrel{\text{df}}{=} \frac{1}{2} (\theta - m_\theta)^T \Sigma_\theta^{-1} (\theta - m_\theta) + \frac{1}{2\sigma_z^2} \sum_{n=1}^N (w_n - \mathbf{1}_M^T \theta x_n)^2. \quad (3.113)$$

Oszacowanie wektora nieznanych parametrów otrzymujemy w wyniku maksymalizacji funkcji gęstości rozkładu prawdopodobieństwa (3.112), co sprowadza się do minimalizacji wyrażenia (3.113) względem wektora parametrów θ . Po przyrównaniu gradientu funkcji (3.113) względem θ do M -wymiarowego wektora zerowego $\mathbf{0}_M$ otrzymujemy układ równań

$$\text{grad}_\theta G(\theta, X_N, \tilde{W}_N) \Big|_{\theta=\hat{\theta}_N} = \Sigma_\theta^{-1} (\hat{\theta}_N - m_\theta) - \frac{1}{\sigma_z^2} \sum_{n=1}^N (w_n - \mathbf{1}_M^T \hat{\theta}_N x_n) \mathbf{1}_M x_n = \mathbf{0}_M, \quad (3.114)$$

a rozwiązanie względem $\hat{\theta}_N$ daje algorytm estymacji

$$\hat{\theta}_N = m_\theta + \frac{\sum_{n=1}^N x_n \tilde{w}_n - \mathbf{1}_M^T m_\theta \sum_{n=1}^N x_n^2}{\sigma_z^2 + \mathbf{1}_M^T \Sigma_\theta \mathbf{1}_M \sum_{n=1}^N x_n^2} \Sigma_\theta \mathbf{1}_M. \quad (3.115)$$

W rozpatrywanym przykładzie charakterystykę statyczną (3.110) systemu złożonego możemy sprowadzić do postaci (3.41), czyli

$$v = \tilde{F}(x, \tilde{\theta}) = \tilde{\theta} x, \quad (3.116)$$

gdzie zależność (3.42), która określa parametr zastępczy $\tilde{\theta}$, zapisujemy

$$\tilde{\theta} = \Gamma(\theta) = \mathbf{1}_M^T \theta = \sum_{m=1}^M \theta_m. \quad (3.117)$$

Teraz, stosując **drugie podejście**, wyznaczmy oszacowanie parametru zastępczego charakterystyki statycznej (3.116). Jak widać, przekształcenie (3.117) daje parametr zastępczy w postaci sumy parametrów poszczególnych elementów.

Przyjmując w równaniu (3.98) funkcję strat postaci (2.70), otrzymujemy metodę maksymalnego prawdopodobieństwa a posteriori, a wyznaczenie estymatora sprowadza się do maksymalizacji gęstości rozkładu prawdopodobieństwa względem $\tilde{\theta}$. W naszym przykładzie $f_{\tilde{\theta}}(\tilde{\theta})$ – funkcja gęstości rozkładu prawdopodobieństwa a priori, określona przez funkcję Γ (3.117) oraz funkcję gęstości $f_{\theta}(\theta)$ (3.109), przybiera postać

$$f_{\tilde{\theta}}(\tilde{\theta}) = \frac{1}{\sqrt{2\pi \mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M}} \exp \left[-\frac{1}{2} \frac{(\tilde{\theta} - \mathbf{1}_M^T m_{\theta})^2}{\mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M} \right],$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori (3.99) wyraża się zależnością

$$\tilde{f}'(\tilde{\theta} | \tilde{W}_N; X_N) = \frac{\frac{1}{\sqrt{2\pi \mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M}} \left(\frac{1}{\sqrt{2\pi \sigma_z^2}} \right)^N \exp[-G(\tilde{\theta}, X_N, \tilde{W}_N)]}{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M}} \left(\frac{1}{\sqrt{2\pi \sigma_z^2}} \right)^N \exp[-G(\tilde{\theta}, X_N, \tilde{W}_N)] d\tilde{\theta}}, \quad (3.118)$$

w której

$$G(\tilde{\theta}, X_N, \tilde{W}_N) \stackrel{\text{df}}{=} \frac{1}{2} \frac{(\tilde{\theta} - \mathbf{1}_M^T m_{\theta})^2}{\mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M} + \frac{1}{2\sigma_z^2} \sum_{n=1}^N (w_n - \tilde{\theta} x_n)^2. \quad (3.119)$$

Oszacowanie nieznanego parametru otrzymujemy w wyniku maksymalizacji funkcji gęstości rozkładu (3.118), co sprowadza się do minimalizacji wyrażenia (3.119) względem parametru $\tilde{\theta}$. Po przyrównaniu pochodnej względem $\tilde{\theta}$ funkcji (3.119) do zera otrzymujemy równanie

$$\left. \frac{dG(\tilde{\theta}, X_N, \tilde{W}_N)}{d\tilde{\theta}} \right|_{\tilde{\theta} = \hat{\tilde{\theta}}_N} = \frac{(\hat{\tilde{\theta}}_N - \mathbf{1}_M^T m_{\theta})}{\mathbf{1}_M^T \Sigma_{\theta} \mathbf{1}_M} - \frac{1}{\sigma_z^2} \sum_{n=1}^N (w_n - \hat{\tilde{\theta}}_N x_n) x_n = 0, \quad (3.120)$$

a rozwiązanie względem $\hat{\tilde{\theta}}_N$ daje algorytm estymacji

$$\hat{\theta}_N = \frac{\mathbf{1}_M^T m_\theta \sigma_z^2 + \mathbf{1}_M^T \Sigma_\theta \mathbf{1}_M \sum_{n=1}^N x_n \tilde{w}_n}{\sigma_z^2 + \mathbf{1}_M^T \Sigma_\theta \mathbf{1}_M \sum_{n=1}^N x_n^2}. \quad (3.121)$$



Dwa kroki w zadaniu estymacji

Takie dwoiste spojrzenie na zadanie estymacji parametrów charakterystyk statycznych poszczególnych elementów systemu oraz estymacji parametrów charakterystyki statycznej obiektu identyfikacji określonego przez strukturę i sytuację pomiarową pozwala na wyznaczenie oszacowania parametrów charakterystyk poszczególnych elementów systemu złożonego w dwóch krokach.

Naturalne dwa kroki są widoczne wtedy, gdy system jest probabilistycznie separowalny w sensie określonej metody, a funkcja Γ (3.42) jest wzajemnie jednoznaczna, wówczas w **kroku pierwszym** wyznaczamy oszacowanie

$$\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) \quad (3.122)$$

wektora parametrów $\tilde{\theta}$ charakterystyki statycznej (3.41) systemu złożonego. W metodzie maksymalnej wiarygodności odpowiada to rozwiązaniu zadania (3.87), a w przypadku metod Bayesa zadaniu (3.97).

Krok drugi polega na wyznaczeniu oszacowania $\hat{\theta}_N$ wektora parametrów charakterystyk statycznych poszczególnych elementów systemu θ , z wykorzystaniem przekształcenia odwrotnego Γ^{-1} , czyli

$$\hat{\theta}_N = \Gamma^{-1}(\hat{\theta}_N). \quad (3.123)$$

Po uwzględnieniu zależności (3.122) we wzorze (3.123) ostatecznie otrzymujemy

$$\hat{\theta}_N = \Gamma^{-1}(\tilde{\Psi}_N(X_N, \tilde{W}_N)) \stackrel{\text{df}}{=} \Psi_N(X_N, \tilde{W}_N). \quad (3.124)$$

Zadanie staje się ciekawsze, gdy funkcja Γ (3.42) nie jest wzajemnie jednoznaczna i system złożony nie jest separowalny (twierdzenie 3.3). Inaczej mówiąc – wyniki pomiarów nie wystarczają do jednoznacznego wyznaczenia oszacowania $\hat{\theta}_N$ wektora parametrów θ . Sięgnięcie po informację aprioryczną pozwala na uzyskanie separowalności. Możliwe są dwa przypadki. Pierwszy, analogiczny do przypadku deterministycznego (twierdzenie 3.2), w którym dodatkowa informacja o badanym systemie złożonym, w postaci funkcji $\bar{\Gamma}$ (3.59), pozwala uzyskać separowalność systemu złożonego.

W tym przypadku w **kroku pierwszym**, korzystając z metody maksymalnej wiarygodności (3.87), wyznaczamy oszacowanie $\hat{\theta}_N$ (3.122) wektora parametrów zastępczych $\tilde{\theta}$ charakterystyki statycznej (3.41). Następnie, w **kroku drugim**, korzystając z informacji opisanej funkcją $\bar{\Gamma}$, tworzymy układ równań

$$\hat{\theta}_N \stackrel{\text{df}}{=} \begin{bmatrix} \hat{\theta}_N \\ \bar{\theta} \end{bmatrix} = \begin{bmatrix} \Gamma(\hat{\theta}_N) \\ \bar{\Gamma}(\hat{\theta}_N) \end{bmatrix} \stackrel{\text{df}}{=} \tilde{\Gamma}(\hat{\theta}_N). \quad (3.125)$$

Teraz funkcja $\bar{\Gamma}$ łącznie z Γ tworzą przekształcenie $\tilde{\Gamma}$, które jest wzajemnie jednoznaczne względem $\hat{\theta}_N$. Z zależności (3.125) wyznaczamy $\hat{\theta}_N$, czyli

$$\hat{\theta}_N = \tilde{\Gamma}^{-1} \left(\begin{bmatrix} \hat{\theta}_N \\ \bar{\theta} \end{bmatrix} \right) = \tilde{\Gamma}^{-1} \left(\begin{bmatrix} \hat{\theta}_N \\ \bar{\theta} \end{bmatrix} \right), \quad (3.126)$$

gdzie $\tilde{\Gamma}^{-1}$ jest funkcją odwrotną $\tilde{\Gamma}$.

Po uwzględnieniu algorytmu estymacji (3.122), uzyskanego w kroku pierwszym, otrzymujemy oszacowanie wektora parametrów charakterystyk statycznych poszczególnych elementów systemu θ , czyli

$$\hat{\theta}_N = \tilde{\Gamma}^{-1} \left(\begin{bmatrix} \tilde{\Psi}_N(X_N, \tilde{W}_N) \\ \bar{\theta} \end{bmatrix} \right) \stackrel{\text{df}}{=} \tilde{\Psi}_N(X_N, \tilde{W}_N, \bar{\theta}). \quad (3.127)$$

Drugi przypadek to informacja aprioryczna w postaci funkcji gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej θ , tj. funkcji $f_\theta(\theta)$. W **kroku pierwszym**, w wyniku rozwiązania zadania (3.87) lub (3.97), otrzymujemy oszacowanie $\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$ wektora parametrów zastępczych $\tilde{\theta}$ charakterystyki statycznej (3.41) systemu złożonego. W **kroku drugim** oszacowanie wektora parametrów charakterystyk statycznych poszczególnych elementów systemu θ uzyskamy w wyniku minimalizacji ryzyka

$$R(\bar{\theta}) = \mathbb{E} \left[L(\underline{\theta}, \bar{\theta}) \right] = \int_{\underline{\theta}} L(\underline{\theta}, \bar{\theta}) f_\theta(\underline{\theta}) d\underline{\theta}, \quad (3.128)$$

względem $\bar{\theta}$ przy ograniczeniu

$$\hat{\theta}_N = \Gamma(\bar{\theta}). \quad (3.129)$$

Oznaczamy przez $\bar{\Theta}_N$ zbiór dopuszczalnych decyzji $\bar{\theta}$

$$\bar{\Theta}_N = \left\{ \bar{\theta} \in \Theta \subseteq \mathbb{R}^R : \hat{\theta}_N = \Gamma(\bar{\theta}) \right\}. \quad (3.130)$$

Ostatecznie wyznaczenie oszacowania wektora parametrów θ sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, dla której

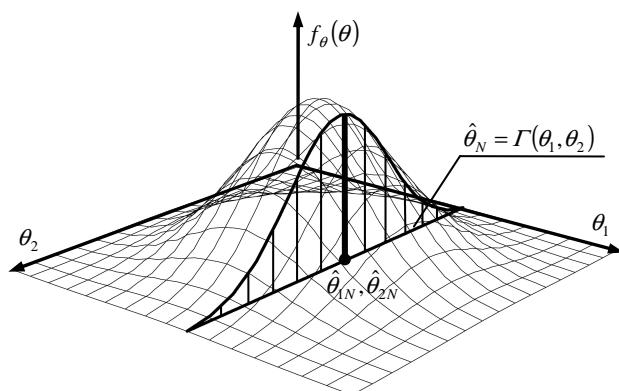
$$R(\hat{\theta}_N) = \min_{\theta \in \bar{\Theta}_N} R(\bar{\theta}). \quad (3.131)$$

W szczególnym przypadku, po przyjęciu funkcji strat w postaci (2.70), ryzyko

$$R(\bar{\theta}) = -f_\theta(\bar{\theta}), \quad (3.132)$$

a zadanie (3.131) sprowadza się do maksymalizacji funkcji gęstości rozkładu prawdopodobieństwa a priori $f_\theta(\theta)$ przy ograniczeniach (3.129), (rys. 3.6), a zatem

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) : f_\theta(\hat{\theta}_N) = \max_{\theta \in \bar{\Theta}_N} f_\theta(\bar{\theta}). \quad (3.133)$$



Rys. 3.6. Estymacja parametrów systemu złożonego w dwóch krokach

W kroku pierwszym wykorzystujemy wiedzę zdobytą w wyniku eksperymentu. Gdy ta wiedza nie pozwala na jednoznaczne wyznaczenie oszacowania parametrów, w kroku drugim korzystamy z informacji apriorycznej, która prowadzi do separowalności.

Przykład 3.8. Ponownie rozważymy system równolegle połączonych elementów opisanych tak jak w przykładzie 3.7. Zastosujemy dwukrokowe podejście do rozwiązania tego zadania. W kroku pierwszym skorzystamy z metody maksymalnej wiarygodności w celu wyznaczenia oszacowania parametru zastępczego $\tilde{\theta}$, a w drugim – opierając się na informacji apriorycznej – wyznaczmy oszacowanie wektora parametrów θ , przy założeniu, że funkcja strat ma postać (2.70).

Krok pierwszy. W rozpatrywanym przykładzie funkcja wiarygodności (3.86) ma postać

$$\tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N) = \left(\frac{1}{\sigma_z \sqrt{2\pi}} \right)^N \exp \left[\sum_{n=1}^N - \frac{(w_n - \tilde{\theta} x_n)^2}{2\sigma_z^2} \right]. \quad (3.134)$$

Oszacowanie zastępczego parametru otrzymujemy w wyniku maksymalizacji funkcji wiarygodności (3.134), co sprowadza się do maksymalizacji wyrażenia w wykładniku względem parametru $\tilde{\theta}$. Po przyrównaniu pochodnej wykładnika wyrażenia (3.134) względem $\tilde{\theta}$ do zera otrzymujemy równanie

$$\frac{d}{d\tilde{\theta}} \left[\sum_{n=1}^N - \frac{(w_n - \tilde{\theta} x_n)^2}{2\sigma_z^2} \right] \Big|_{\tilde{\theta}=\hat{\theta}_N} = \frac{1}{\sigma_z^2} \sum_{n=1}^N (w_n - \hat{\theta}_N x_n) x_n = 0, \quad (3.135)$$

którego rozwiązanie względem $\hat{\theta}_N$ daje algorytm estymacji

$$\hat{\theta}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) = \frac{\sum_{n=1}^N x_n \tilde{w}_n}{\sum_{n=1}^N x_n^2}. \quad (3.136)$$

Krok drugi. Oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego otrzymamy w wyniku maksymalizacji funkcji gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej $\underline{\theta}$

$$f_{\theta}(\bar{\theta}) = (2\pi)^{-\frac{M}{2}} |\Sigma_{\theta}|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\bar{\theta} - m_{\theta})^T \Sigma_{\theta}^{-1} (\bar{\theta} - m_{\theta}) \right] \quad (3.137)$$

względem $\bar{\theta}$, przy ograniczeniu

$$\hat{\theta}_N = \Gamma(\bar{\theta}) = \mathbf{1}_M^T \bar{\theta}, \quad (3.138)$$

co jest równoważne maksymalizacji wykładnika wyrażenia (3.137) przy ograniczeniu (3.138).

W rozpatrywanym zadaniu optymalizacji z ograniczeniem równościowym (3.138) funkcja Lagrange'a ma postać

$$\text{Lag}(\bar{\theta}, \lambda) \stackrel{\text{df}}{=} -\frac{1}{2} (\bar{\theta} - m_{\theta})^T \Sigma_{\theta}^{-1} (\bar{\theta} - m_{\theta}) + \lambda \left(\mathbf{1}_M^T \bar{\theta} - \hat{\theta}_N \right), \quad (3.139)$$

gdzie λ jest współczynnikiem Lagrange'a.

Korzystając z warunków koniecznych, otrzymujemy układ równań:

$$\text{grad}_{\bar{\theta}} \text{Lag}(\bar{\theta}, \lambda) \Big|_{\bar{\theta}=\hat{\theta}_N} = -\Sigma_{\theta}^{-1} (\hat{\theta}_N - m_{\theta}) + \lambda \mathbf{1}_M = \mathbf{0}_M, \quad (3.140)$$

$$\frac{d}{d\lambda} \text{Lag}(\bar{\theta}, \lambda) \Big|_{\bar{\theta}=\hat{\theta}_N} = \mathbf{1}_M^T \hat{\theta}_N - \hat{\theta}_N = 0, \quad (3.141)$$

z którego wyznaczamy

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) = m_\theta + \frac{\hat{\theta}_N - \mathbf{1}_M^T m_\theta}{\mathbf{1}_M^T \Sigma \mathbf{1}_M} \Sigma \mathbf{1}_M, \quad (3.142)$$

a po podstawieniu algorytmu estymacji (3.136) w miejsce $\hat{\theta}_N$ oraz prostych przekształceniach otrzymujemy algorytm estymacji parametrów elementów systemu złożonego

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) = m_\theta + \frac{\sum_{n=1}^N x_n \tilde{w}_n - \mathbf{1}_M^T m_\theta \sum_{n=1}^N x_n^2}{\mathbf{1}_M^T \Sigma \mathbf{1}_M \sum_{n=1}^N x_n^2} \Sigma \mathbf{1}_M. \quad (3.143)$$

Zauważmy, że wynik uzyskany w podejściu dwukrokowym nie zawsze jest równoważny z wynikiem uzyskanym w podejściu bezpośrednim metodą bayesowską, przedstawioną w zadaniu (3.90). Wynik (3.143) w przykładzie 3.8, uzyskany w podejściu dwukrokowym, różni się od wyniku (3.115), uzyskanego w podejściu bezpośrednim w przykładzie 3.7. W dwuetapowym podejściu w kroku pierwszym możemy skorzystać z różnych metod estymacji, w zależności od posiadanej informacji. Gdybyśmy w ostatnio analizowanym przypadku w kroku pierwszym posłużyli się metodą maksymalnego prawdopodobieństwa a posteriori i w miejsce $\hat{\theta}_N$ w (3.142) podstawili algorytm estymacji (3.121), otrzymamy

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) = m_\theta + \frac{\frac{\mathbf{1}_M^T m_\theta \sigma_z^2 + \mathbf{1}_M^T \Sigma_\theta \mathbf{1}_M \sum_{n=1}^N x_n \tilde{w}_n}{\sigma_z^2 + \mathbf{1}_M^T \Sigma_\theta \mathbf{1}_M \sum_{n=1}^N x_n^2} - \mathbf{1}_M^T m_\theta}{\mathbf{1}_M^T \Sigma \mathbf{1}_M} \Sigma \mathbf{1}_M, \quad (3.144)$$

a po prostych przekształceniach widzimy, że jest on identyczny z algorytmem (3.115).

Problem badania równoważności podejścia bezpośredniego i dwukrokowego jest otwarty i wymaga dalszych prac. Niektóre wyniki przedstawiono w publikacjach [25– 27, 92, 95].

3.2.2. Estymacja parametrów charakterystyk statycznych systemu złożonego ze zmiennymi niemierzalnymi wielkościami losowymi

Rozważmy przypadek, kiedy w elementach systemu złożonego wskazano niemierzalne zmienne wielkości losowe. Innymi słowy – rozpatrujemy wejściowo-wyjściowy system złożony, w którym wyróżniono M elementów O_1, O_2, \dots, O_M , dla których

charakterystyka statyczna m -tego obiektu, o wejściu $u_m \in U_m \subseteq \mathbb{R}^{S_m}$ i wyjściu $y_m \in Y_m \subseteq \mathbb{R}^{L_m}$, jest znana z dokładnością do wektora parametrów $\theta_m \in \Theta_m \subseteq \mathbb{R}^{R_m}$. Na obiekt działają zmienne niemierzalne wielkości losowe

$$\omega_m = \begin{bmatrix} \omega_m^{(1)} \\ \omega_m^{(2)} \\ \vdots \\ \omega_m^{(L_m)} \end{bmatrix} \in \Omega_m \subseteq \mathbb{R}^{L_m}. \quad (3.145)$$

Charakterystyka statyczna elementu jest uzupełniona o wielkość losową ω_m , czyli charakterystyka statyczna elementu O_m dana jest zależnością

$$y_m = F_m(u_m, \theta_m, \omega_m), \quad m = 1, 2, \dots, M. \quad (3.146)$$

W dalszych rozważaniach będziemy zakładać, że wymiar wektora y_m oraz wielkości przypadkowej ω_m jest taki sam i równy L_m ($\dim y_m = \dim \omega_m = L_m$). Funkcja F_m jest wzajemnie jednoznaczna względem ω_m , czyli istnieje funkcja odwrotna $F_{\omega_m}^{-1}$

$$\omega_m = F_{\omega_m}^{-1}(u_m, \theta_m, y_m). \quad (3.147)$$

O losowo zmiennym wektorze parametrów zakładamy, że jest wartością L_m -wymiarowej zmiennej losowej $\underline{\omega}_m$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_{\omega_m}(\omega_m)$. Zakładamy ponadto, że zmienne losowe $\underline{\omega}_1, \underline{\omega}_2, \dots, \underline{\omega}_M$ są niezależne. Struktura systemu złożonego jest dana zależnością (3.6), a wyróżnione, mierzalne wyjścia są opisane równaniem (3.7). Przyjmujemy oznaczenie

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} F_1(u_1, \theta_1, \omega_1) \\ F_2(u_2, \theta_2, \omega_2) \\ \vdots \\ F_M(u_M, \theta_M, \omega_M) \end{bmatrix} \stackrel{\text{df}}{=} \bar{F}(u, \theta, \omega), \quad (3.148)$$

gdzie ω jest wektorem wartości losowych złożonym z wektorów wielkości losowych poszczególnych elementów, tj.

$$\omega = \begin{bmatrix} \omega^{(1)} \\ \omega^{(2)} \\ \vdots \\ \omega^{(L)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_M \end{bmatrix}. \quad (3.149)$$

$$\omega \in \Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_M \subseteq \mathbb{R}^L, \quad L = \sum_{m=1}^M L_m. \quad (3.150)$$

Wektor ω jest wartością zmiennej losowej $\underline{\omega}$, dla której funkcja gęstości rozkładu prawdopodobieństwa jest

$$f_{\omega}(\omega) = \prod_{m=1}^M f_{\omega_m}(\omega_m). \quad (3.151)$$

Po podstawieniu w miejsce u w równaniu (3.148) zależności (3.6) otrzymujemy

$$y = \bar{F}(Ay + Bx, \theta, \omega). \quad (3.152)$$

Rozwikłanie (3.152) względem y daje

$$y = \bar{F}_y(x, \theta, \omega; A, B), \quad (3.153)$$

a po podstawieniu zależności (3.153) do (3.7) otrzymujemy

$$v = C\bar{F}_y(x, \theta, \omega; A, B) \stackrel{\text{df}}{=} F(x, \theta, \omega). \quad (3.154)$$

Zależność (3.154) jest charakterystyką statyczną systemu złożonego z wektorem wejść zewnętrznych x i wektorem wyróżnionych, dostępnych do pomiaru, wyjść v .

W celu wyznaczenia wektora parametrów charakterystyk statycznych elementów systemu złożonego dla zadanej serii wartości wejść zewnętrznych zmierzono wartości wyróżnionych wyjść. Dla zadanej serii identyfikującej o długości N uzyskano wyniki pomiarów wartości wyróżnionych wyjść systemu złożonego. Wyniki eksperymentu zebrano w macierzach:

$$X_N = [x_1 \quad x_2 \quad \cdots \quad x_N], \quad V_N = [v_1 \quad v_2 \quad \cdots \quad v_N]. \quad (3.155)$$

Zadanie identyfikacji sprowadza się do wyznaczenia oszacowania wektora parametrów θ charakterystyki statycznej (3.154). Poszukujemy algorytmu estymacji

$$\hat{\theta}_N = \Psi_N(X_N, V_N), \quad (3.156)$$

który dla danych pomiarowych X_N, V_N umożliwi wyznaczenie wartości oszacowania $\hat{\theta}_N$ wektora θ .

W zależności od informacji, jakie posiadamy o zmiennej losowej $\underline{\omega}$, charakterystyce statycznej (3.154) oraz jej parametrach θ , możemy stosować różne metody estymacji.

Korzystając z **metody maksymalnej wiarogodności** (tak jak w punkcie 2.2.2 – zadanie (2.95)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymaliza-

cji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, V_N)$, dla której funkcja wiarygodności L_N (3.158), wyznaczona na podstawie macierzy pomiarów V_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, V_N) : L_N(V_N, \hat{\theta}_N; X_N) = \max_{\theta \in \Theta} L_N(V_N, \theta; X_N), \quad (3.157)$$

gdzie funkcja wiarygodności ma postać

$$L_N(V_N, \theta; X_N) = \prod_{n=1}^N f_{\omega}(F_{\omega}^{-1}(x_n, \theta, v_n)) |J_F|, \quad (3.158)$$

gdzie: F_{ω}^{-1} jest funkcją odwrotną funkcji (3.154) względem ω , a jacobian przekształcenia odwrotnego wyraża się zależnością

$$J_F = \frac{\partial F_{\omega}^{-1}(x, \theta, v)}{\partial v}. \quad (3.159)$$

Podobnie jak w poprzednich przypadkach, parametry występujące w charakterystyce statycznej systemu złożonego są wynikiem przekształcenia parametrów charakterystyk statycznych poszczególnych elementów systemu, a charakterystyka (3.154) przyjmuje postać

$$v = C\bar{F}_y(x, \theta, \omega; A, B) = F(x, \theta, \omega) \stackrel{\text{df}}{=} \tilde{F}(x, \tilde{\theta}, \omega), \quad (3.160)$$

gdzie

$$\tilde{\theta} \stackrel{\text{df}}{=} \Gamma(\theta). \quad (3.161)$$

Po uwzględnieniu charakterystyki statycznej postaci (3.160) funkcja wiarygodności przyjmie postać

$$\begin{aligned} L_N(V_N, \theta; X_N) &= \prod_{n=1}^N f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \Gamma(\theta), v_n)) |J_{\tilde{F}}| \\ &= \prod_{n=1}^N f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \tilde{\theta}, v_n)) |J_{\tilde{F}}| \stackrel{\text{df}}{=} \tilde{L}_N(V_N, \tilde{\theta}; X_N), \end{aligned} \quad (3.162)$$

gdzie: \tilde{F}_{ω}^{-1} jest funkcją odwrotną funkcji (3.160) względem ω , a jacobian przekształcenia odwrotnego jest dany zależnością

$$J_{\tilde{F}} = \frac{\partial \tilde{F}_{\omega}^{-1}(x, \tilde{\theta}, v)}{\partial v}. \quad (3.163)$$

Oszacowanie wektora parametrów $\tilde{\theta}$ charakterystyki statycznej systemu złożonego (3.160) uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, V_N)$, dla której funkcja wiarygodności \tilde{L}_N (3.162),

wyznaczona na podstawie macierzy pomiarów V_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem $\tilde{\theta}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, tj.

$$\hat{\theta}_N = \tilde{\Psi}_N(X_N, V_N) : \tilde{L}_N(V_N, \hat{\theta}_N; X_N) = \max_{\tilde{\theta} \in \tilde{\Theta}} \tilde{L}_N(V_N, \tilde{\theta}; X_N). \quad (3.164)$$

Zauważmy, że istotną rolę w uzyskaniu separowalności probabilistycznej ma postać funkcji Γ . Można przeprowadzić dyskusję zagadnienia separowalności analogicznie do przedstawionej w punkcie 3.2.1 (twierdzenie 3.3).

Założymy, dodatkowo w stosunku do poprzednich rozważań, że badany system złożony jest wylosowany z pewnej populacji. Znaczy to, że wektor parametrów θ jest wartością ciągłej zmiennej losowej $\underline{\theta}$, o znanej funkcji gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$. Korzystając z **metod Bayesa** (jak w punkcie 2.2.2 – zadanie (2.99)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, V_N)$, dla której ryzyko warunkowe, wyznaczone na podstawie macierzy pomiarów V_N przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, V_N) : r(\hat{\theta}_N, V_N; X_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, V_N; X_N), \quad (3.165)$$

gdzie ryzyko warunkowe ma postać

$$\begin{aligned} r(\bar{\theta}, V_N; X_N) &\stackrel{\text{df}}{=} \mathbb{E}_{\underline{\theta}} [L(\underline{\theta}, \bar{\theta}) | V_N; X_N] \\ &= \int_{\Theta} L(\theta, \bar{\theta}) f'(\theta | V_N; X_N) d\theta, \end{aligned} \quad (3.166)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori $f'(\theta | V_N; X_N)$ wyraża się wzorem

$$f'(\theta | V_N; U_N) = \frac{f_\theta(\theta) \prod_{n=1}^N f_\omega(F_\omega^{-1}(x_n, \theta, v_n)) |J_F|}{\int_{\Theta} f_\theta(\theta) \prod_{n=1}^N f_\omega(F_\omega^{-1}(x_n, \theta, v_n)) |J_F| d\theta}. \quad (3.167)$$

Biorąc pod uwagę charakterystykę statyczną postaci (3.160) dla systemu złożonego, oszacowanie wektora zastępczych parametrów $\tilde{\theta} = \Gamma(\theta)$ uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \tilde{\Psi}_N(X_N, V_N)$, dla której ryzyko warunkowe, wyznaczone na podstawie macierzy pomiarów V_N przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, tj.

$$\hat{\theta}_N = \tilde{\Psi}_N(X_N, V_N) : \tilde{r}\left(\hat{\theta}_N, V_N; X_N\right) = \min_{\theta \in \tilde{\Theta}} \tilde{r}(\theta, V_N; X_N), \quad (3.168)$$

gdzie ryzyko warunkowe ma postać

$$\begin{aligned} \tilde{r}(\bar{\theta}, V_N; X_N) &\stackrel{\text{df}}{=} \mathbb{E} \left[L(\tilde{\theta}, \bar{\theta}) | V_N; X_N \right] \\ &= \int_{\tilde{\Theta}} L(\tilde{\theta}, \bar{\theta}) f'(\tilde{\theta} | V_N; X_N) d\tilde{\theta}, \end{aligned} \quad (3.169)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori $\tilde{f}'(\tilde{\theta} | V_N; X_N)$ wyraża się wzorem

$$\tilde{f}'(\tilde{\theta} | V_N; X_N) = \frac{f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \tilde{\theta}, v_n)) | J_{\tilde{F}}|}{\int_{\tilde{\Theta}} f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \tilde{\theta}, v_n)) | J_{\tilde{F}}| d\tilde{\theta}}. \quad (3.170)$$

Funkcja $f_{\tilde{\theta}}(\tilde{\theta})$ gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej $\tilde{\theta}$ jest określona przez funkcję Γ oraz $f_{\theta}(\theta)$ funkcję gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej θ i dana jest wzorem (3.100) lub (3.105), w zależności od funkcji Γ .

Zauważmy, że istotną rolę w uzyskaniu separowalności probabilistycznej gra postać funkcji Γ . Można przeprowadzić dyskusję zagadnienia separowalności analogicznie do przedstawionej w punkcie 3.2.1 (twierdzenie 3.3). Podobnie jak w punkcie 3.2.1, można zastosować dwukrokowe podejście do zadania estymacji parametrów, uwzględniając wzajemną jednoznaczność przekształcenia Γ w przypadku systemu probabilistycznie separowalnego. Uzupełnienie zadania dodatkową informacją aprioryczną w postaci dodatkowych relacji $\bar{\Gamma}$ pomiędzy parametrami charakterystyk statycznych elementów systemu złożonego (3.125) lub funkcją gęstości rozkładu prawdopodobieństwa $f_{\theta}(\theta)$ zmiennej losowej θ pozwala na uzyskanie separowalności.

3.2.3. Estymacja parametrów charakterystyk statycznych systemu złożonego ze zmiennymi niemierzalnymi wielkościami losowymi na podstawie zakłóconych pomiarów wyróżnionych wyjść systemu złożonego

Rozważmy zadanie estymacji parametrów systemu złożonego z losowo zmiennymi wielkościami w elementach wchodzących w skład systemu, przy założeniu, że pomiar wyróżnionych wyjść systemu złożonego jest obciążony błędem pomiarowym. Innymi

słowy – rozpatrujemy wejściowo-wyjściowy system złożony, w którym wyróżniono M elementów O_1, O_2, \dots, O_M , dla których charakterystyki statyczne mają postać (3.146). Strukturę systemu opisuje równanie (3.6), a wyróżnione, pomiarowo dostępne, wyjścia są opisane równaniem (3.7). W konsekwencji charakterystyka statyczna systemu złożonego, z wejściami zewnętrznymi x i wyróżnionymi, dostępnymi do pomiaru, wyjściami v , dana jest zależnością (3.154). W zależności tej występuje wektor nieznanych parametrów θ oraz losowo zmienny wektor wartości ω , dla którego znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\omega(\omega)$. Tak jak w punkcie 3.2.1, wektor wyróżnionych wyjść v mierzony jest z zakłóceniami pomiarowymi \tilde{z} i opis systemu pomiarowego dany jest zależnością (3.64). W celu wyznaczenia nieznanych parametrów charakterystyk elementów systemu złożonego dla danego ciągu wartości wejść zewnętrznych zmierzono wartości wyróżnionych wyjść z zakłóceniami. Dla zadanej serii identyfikującej o długości N uzyskano wyniki pomiarów wartości wyróżnionych wyjść systemu złożonego. Wyniki eksperymentu zapisano w macierzach:

$$X_N = [x_1 \quad x_2 \quad \dots \quad x_N], \quad \tilde{W}_N = [\tilde{w}_1 \quad \tilde{w}_2 \quad \dots \quad \tilde{w}_N]. \quad (3.171)$$

Zadanie identyfikacji sprowadza się do wyznaczenia oszacowania wektora nieznanych parametrów θ charakterystyki (3.154). Poszukujemy algorytmu estymacji

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N), \quad (3.172)$$

który dla danych pomiarowych X_N, \tilde{W}_N umożliwi wyznaczenie wartości oszacowania $\hat{\theta}_N$ wektora θ .

W zależności od informacji, jakie posiadamy o systemie pomiarowym, zakłóceniach pomiarowych \tilde{z} , zmiennej losowej ω , charakterystyce statycznej (3.154) oraz jej parametrach θ , możemy stosować różne metody estymacji.

Korzystając z **metody maksymalnej wiarygodności** (jak w punkcie 2.2.3 – zadanie (2.114)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, dla której funkcja wiarygodności L_N , wyznaczona na podstawie macierzy zakłóconych wyników pomiarów wartości wyróżnionych wyjść \tilde{W}_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem θ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) : L_N(\tilde{W}_N, \hat{\theta}_N; X_N) = \max_{\theta \in \Theta} L_N(\tilde{W}_N, \theta; X_N), \quad (3.173)$$

gdzie funkcja wiarygodności ma postać

$$L_N(\tilde{W}_N, \theta; X_N) = \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w}_n)) f_\omega(F_\omega^{-1}(x_n, \theta, v)) |J_F| |J_{\tilde{h}}| dv, \quad (3.174)$$

jakobian przekształcenia odwrotnego J_F wyraża się zależnością

$$J_F = \frac{\partial F_\omega^{-1}(x, \theta, v)}{\partial v}, \quad (3.175)$$

a jakobian przekształcenia odwrotnego $J_{\tilde{h}}$

$$J_{\tilde{h}} = \frac{\partial \tilde{h}_z^{-1}(v, \tilde{w})}{\partial v}. \quad (3.176)$$

Podobnie jak w poprzednich przypadkach, parametry charakterystyki statycznej (3.154) systemu złożonego są wynikiem przekształcenia parametrów charakterystyk statycznych poszczególnych elementów, a charakterystyka statyczna przyjmuje postać

$$v = C\bar{F}_y(x, \theta, \omega; A, B) = F(x, \theta, \omega) \stackrel{\text{df}}{=} \tilde{F}(x, \tilde{\theta}, \omega), \quad (3.177)$$

gdzie

$$\tilde{\theta} \stackrel{\text{df}}{=} \Gamma(\theta). \quad (3.178)$$

Po uwzględnieniu charakterystyki statycznej postaci (3.160) funkcję wiarygodności zapiszemy

$$\begin{aligned} L_N(\tilde{W}_N, \theta; X_N) &= \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_z^{-1}(v, \tilde{w}_n)) f_\omega(\tilde{F}_\omega^{-1}(x_n, \Gamma(\theta), v)) |J_{\tilde{F}}| |J_{\tilde{h}}| dv \\ &= \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_z^{-1}(v, \tilde{w}_n)) f_\omega(\tilde{F}_\omega^{-1}(x_n, \tilde{\theta}, v)) |J_{\tilde{F}}| |J_{\tilde{h}}| dv = \tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N), \end{aligned} \quad (3.179)$$

a oszacowanie wektora parametrów zastępczych $\tilde{\theta}$ charakterystyki statycznej (3.177) systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$, dla której funkcja wiarygodności \tilde{L}_N (3.179), wyznaczona na podstawie macierzy pomiarów \tilde{W}_N dla zadanej serii identyfikującej X_N , przyjmuje wartość maksymalną względem $\tilde{\theta}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, tj.

$$\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) \rightarrow \tilde{L}_N(\tilde{W}_N, \hat{\tilde{\theta}}_N; X_N) = \max_{\tilde{\theta} \in \tilde{\Theta}} \tilde{L}_N(\tilde{W}_N, \tilde{\theta}; X_N). \quad (3.180)$$

Dodatkowo, w stosunku do poprzednich rozważań, zakładamy, że badany system złożony jest wylosowany z pewnej populacji. Znaczący to, że wektor parametrów θ jest wartością zmiennej losowej $\underline{\theta}$, dla której znana jest funkcja gęstości rozkładu prawdopodobieństwa $f_\theta(\theta)$. Korzystając z **metod Bayesa** (jak w punkcie 2.2.3

– zadanie (2.119)), oszacowanie wektora parametrów charakterystyk statycznych elementów systemu złożonego uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N)$, dla której ryzyko warunkowe, wyznaczone na podstawie macierzy pomiarów \tilde{W}_N przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\bar{\theta}$ ze zbioru dopuszczalnego Θ , tj.

$$\hat{\theta}_N = \Psi_N(X_N, \tilde{W}_N) : r(\hat{\theta}_N, \tilde{W}_N; X_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, \tilde{W}_N; X_N), \quad (3.181)$$

gdzie ryzyko warunkowe ma postać

$$r(\bar{\theta}, \tilde{W}_N; X_N) \stackrel{\text{df}}{=} E_{\bar{\theta}} [L(\bar{\theta}, \bar{\theta}) | \tilde{W}_N; X_N] = \int_{\Theta} L(\theta, \bar{\theta}) f'(\theta | \tilde{W}_N; X_N) d\theta, \quad (3.182)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori $f'(\theta | \tilde{W}_N; X_N)$ wyraża się wzorem

$$f'(\theta | \tilde{W}_N; X_N) = \frac{f_{\theta}(\theta) \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w}_n)) f_{\omega}(F_{\omega}^{-1}(x_n, \theta, v)) |J_F| |J_{\tilde{h}}| dv}{\int_{\Theta} f_{\theta}(\theta) \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w}_n)) f_{\omega}(F_{\omega}^{-1}(x_n, \theta, v)) |J_F| |J_{\tilde{h}}| dv d\theta}. \quad (3.183)$$

Biorąc pod uwagę postać (3.160) charakterystyki statycznej systemu złożonego, oszacowanie wektora parametrów zastępczych $\tilde{\theta} = \Gamma(\theta)$ uzyskamy po rozwiązaniu następującego zadania optymalizacji: wyznaczyć taką wartość oszacowania $\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N)$, dla której ryzyko warunkowe, wyznaczone na podstawie macierzy pomiarów \tilde{W}_N przy zadanej serii identyfikującej X_N , przyjmuje wartość minimalną względem $\bar{\tilde{\theta}}$ ze zbioru dopuszczalnego $\tilde{\Theta}$, tj.

$$\hat{\tilde{\theta}}_N = \tilde{\Psi}_N(X_N, \tilde{W}_N) : \tilde{r}(\hat{\tilde{\theta}}_N, \tilde{W}_N; X_N) = \min_{\bar{\tilde{\theta}} \in \tilde{\Theta}} \tilde{r}(\bar{\tilde{\theta}}, \tilde{W}_N; X_N), \quad (3.184)$$

gdzie ryzyko warunkowe ma postać

$$\tilde{r}(\bar{\tilde{\theta}}, \tilde{W}_N; X_N) \stackrel{\text{df}}{=} E_{\bar{\tilde{\theta}}} [L(\bar{\tilde{\theta}}, \bar{\tilde{\theta}}) | \tilde{W}_N; X_N] = \int_{\tilde{\Theta}} L(\tilde{\theta}, \bar{\tilde{\theta}}) \tilde{f}'(\tilde{\theta} | \tilde{W}_N; X_N) d\tilde{\theta}, \quad (3.185)$$

a funkcja gęstości rozkładu prawdopodobieństwa a posteriori $\tilde{f}'(\tilde{\theta} | \tilde{W}_N; X_N)$ wyraża się wzorem

$$\tilde{f}'(\tilde{\theta}|\tilde{W}_N; X_N) = \frac{f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w}_n)) f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \tilde{\theta}, v)) |J_{\tilde{F}}| |J_{\tilde{h}}| dv}{\int_{\tilde{\Theta}} f_{\tilde{\theta}}(\tilde{\theta}) \prod_{n=1}^N \int_{\mathbb{V}} f_{\tilde{z}}(\tilde{h}_{\tilde{z}}^{-1}(v, \tilde{w}_n)) f_{\omega}(\tilde{F}_{\omega}^{-1}(x_n, \tilde{\theta}, v)) |J_{\tilde{F}}| |J_{\tilde{h}}| dv d\tilde{\theta}}, \quad (3.186)$$

w którym jacobiany przekształceń odwrotnych wyrażają się zależnościami (3.170) oraz (3.176).

Funkcja gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej $\tilde{\theta}$ $f_{\tilde{\theta}}(\tilde{\theta})$ jest określona przez funkcję Γ oraz funkcję gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej θ $f_{\theta}(\theta)$ i dana jest, w zależności od funkcji Γ , wzorem (3.100) lub (3.105).

Zauważmy, że istotną rolę w uzyskaniu separowalności probabilistycznej ma postać funkcji Γ . Można przeprowadzić dyskusję zagadnienia separowalności analogicznie do przedstawionej w punkcie 3.2.1 (twierdzenie 3.3). Podobnie jak w punkcie 3.2.1, można zastosować dwukrokowe podejście do zadania estymacji parametrów, uwzględniając wzajemną jednoznaczność przekształcenia Γ w przypadku systemu probabilistycznie separowalnego. Uzupełnienie zadania informacją aprioryczną w postaci dodatkowych relacji $\bar{\Gamma}$ pomiędzy parametrami charakterystyk statycznych elementów systemu złożonego (3.125) lub funkcją gęstości rozkładu prawdopodobieństwa a priori zmiennej losowej θ $f_{\theta}(\theta)$ pozwala na uzyskanie separowalności.

4. Wybór optymalnego modelu systemu złożonego

Przedstawimy teraz zadanie wyboru optymalnego modelu dla systemu złożonego. Rozważymy wejściowo-wyjściowy system złożony, w którym jest wyróżnionych M elementów O_1, O_2, \dots, O_M . Zakładamy, że struktura systemu jest znana, tzn. znane są powiązania pomiędzy poszczególnymi elementami (1.19). Opisy poszczególnych elementów (1.13) nie są znane. Dla m -tego elementu o wektorach wejść **Błąd! Nie zdefiniowano zakładki.** u_m oraz wyjść y_m przyjmujemy model

$$\bar{y}_m = \Phi_m(u_m, \theta_m), \quad (4.1)$$

w którym: \bar{y}_m jest wektorem wyjść modelu, Φ_m – znaną, zadaną funkcją, a θ_m – nieznanym wektorem parametrów modelu m -tego elementu.

Wektory wejść, wyjść obiektu oraz modelu, a także wektor parametrów modelu m -tego elementu są wektorami z odpowiednich przestrzeni, tj.

$$u_m = \begin{bmatrix} u_m^{(1)} \\ u_m^{(2)} \\ \vdots \\ u_m^{(S_m)} \end{bmatrix}, \quad y_m = \begin{bmatrix} y_m^{(1)} \\ y_m^{(2)} \\ \vdots \\ y_m^{(L_m)} \end{bmatrix}, \quad \bar{y}_m = \begin{bmatrix} \bar{y}_m^{(1)} \\ \bar{y}_m^{(2)} \\ \vdots \\ \bar{y}_m^{(L_m)} \end{bmatrix}, \quad \theta_m = \begin{bmatrix} \theta_m^{(1)} \\ \theta_m^{(2)} \\ \vdots \\ \theta_m^{(R_m)} \end{bmatrix}, \quad (4.2)$$

$$u_m \in \mathbf{U}_m \subseteq \mathbf{R}^{S_m}, \quad y_m, \bar{y}_m \in \mathbf{Y}_m \subseteq \mathbf{R}^{L_m}, \quad \theta_m \in \Theta_m \subseteq \mathbf{R}^{R_m},$$

gdzie: S_m, L_m oraz R_m są – odpowiednio – wymiarami przestrzeni wejść, wyjść obiektu i modelu oraz parametrów, $m = 1, 2, \dots, M$.

Przyjmujemy oznaczenia

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(S)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(L)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}, \quad \bar{y} = \begin{bmatrix} \bar{y}^{(1)} \\ \bar{y}^{(2)} \\ \vdots \\ \bar{y}^{(L)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_M \end{bmatrix}, \quad (4.3)$$

gdzie: wektor wszystkich wejść u jest elementem przestrzeni

$$\mathbf{U} = \mathbf{U}_1 \times \mathbf{U}_2 \times \cdots \times \mathbf{U}_M \subseteq \mathbf{R}^S, \quad S = \sum_{m=1}^M S_m, \quad (4.4)$$

a wektory wszystkich wyjść obiektu y i modelu \bar{y} są elementami przestrzeni

$$\mathbf{Y} = \mathbf{Y}_1 \times \mathbf{Y}_2 \times \cdots \times \mathbf{Y}_M \subseteq \mathbf{R}^L, \quad L = \sum_{m=1}^M L_m. \quad (4.5)$$

Struktura systemu jest dana równaniem

$$u = Ay + Bx, \quad (4.6)$$

w którym: x jest \tilde{S} -wymiarowym wektorem wejść zewnętrznych $x \in \mathbf{X}$, $\mathbf{X} \subseteq \mathbf{R}^{\tilde{S}}$. Macierz A jest $S \times L$ -wymiarową macierzą zero-jedynkową postaci (1.20). Macierz ta definiuje powiązania pomiędzy elementami systemu, natomiast zero-jedynkowa $S \times \tilde{S}$ -wymiarowa macierz B wskazuje wejścia zewnętrzne (1.21).

W dalszych rozważaniach zakładamy, że ze względu na przewidywane wykorzystanie modelu wyróżniono pewne wyjścia v , wskazane przez macierz zero-jedynkową C (1.24), o wymiarze $\tilde{L} \times L$ (\tilde{L} jest liczbą wyróżnionych wyjść spośród wszystkich L wyjść), tj.

$$v = Cy, \quad (4.7)$$

gdzie wyróżnione wyjścia

$$v \in \mathbf{V} \stackrel{\text{df}}{=} \{v : \forall y \in \mathbf{Y}, v = Cy\} \subseteq \mathbf{R}^{\tilde{L}}. \quad (4.8)$$

Struktura systemu dana równaniem (4.6) oraz wyróżnione równaniem (4.7) wyjścia określają nowy obiekt jako całość, z wektorem wejść zewnętrznych x oraz wyróżnionych wyjść v . Dla tak określonego obiektu identyfikacji możemy wyznaczyć model oparty na modelach poszczególnych elementów (4.1) oraz strukturze systemu danej równaniami (4.6) i (4.7), czyli struktura modelu jest określona zależnościami:

$$u = A\bar{y} + Bx, \quad (4.9)$$

$$\bar{v} = C\bar{y}, \quad (4.10)$$

gdzie wyróżnione wyjścia modelu

$$\bar{v} \in \mathbf{V} \stackrel{\text{df}}{=} \{\bar{v} : \forall \bar{y} \in \mathbf{Y}, \bar{v} = C\bar{y}\} \subseteq \mathbf{R}^{\tilde{L}}. \quad (4.11)$$

Przyjmujemy oznaczenie

$$\bar{y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_M \end{bmatrix} = \begin{bmatrix} \Phi_1(u_1, \theta_1) \\ \Phi_2(u_2, \theta_2) \\ \vdots \\ \Phi_M(u_M, \theta_M) \end{bmatrix} \stackrel{\text{df}}{=} \bar{\Phi}(u, \theta), \quad (4.12)$$

gdzie θ jest wektorem parametrów, którego składowymi są wektory parametrów modeli poszczególnych elementów, tj.

$$\theta = \begin{bmatrix} \theta^{(1)} \\ \theta^{(2)} \\ \vdots \\ \theta^{(R)} \end{bmatrix} \stackrel{\text{df}}{=} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_M \end{bmatrix}, \quad (4.13)$$

$$\theta \in \Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_M \subseteq \mathbf{R}^R, \quad R = \sum_{m=1}^M R_m. \quad (4.14)$$

Po podstawieniu w miejsce u w równaniu (4.12) zależności (4.9) otrzymujemy

$$\bar{y} = \bar{\Phi}(A\bar{y} + Bx, \theta). \quad (4.15)$$

Rozwikłanie równania (4.15) względem \bar{y} daje

$$\bar{y} = \bar{\Phi}_y(x, \theta; A, B), \quad (4.16)$$

gdzie $\bar{\Phi}_y$ jest wynikiem rozwikłania,

a po podstawieniu (4.16) do (4.10) mamy

$$\bar{v} = C\bar{\Phi}_y(x, \theta; A, B) \stackrel{\text{df}}{=} \Phi(x, \theta). \quad (4.17)$$

Zależność (4.17) jest modelem systemu złożonego jako całości, z wejściami zewnętrznymi x i mierzonymi wyjściami \bar{v} .

W obecnej sytuacji możliwe są dwa różne sformułowania zadania wyboru optymalnego modelu. Pierwsze polega na tym, że wyznaczamy modele optymalne postaci (4.1) dla poszczególnych elementów, niezależnie od struktury systemu. Następnie, uwzględniając połączenia pomiędzy elementami, wyznaczamy model systemu złożonego. Taki model nazwiemy lokalnie optymalnym. Drugie podejście polega na wyznaczeniu optymalnego modelu z uwzględnieniem jego struktury i opisu systemu złożonego jako całości postaci (4.17). Ten model nazwiemy globalnie optymalnym.

4.1. Lokalnie optymalny model systemu złożonego

Założymy, że każdy z elementów systemu złożonego jest obserwowany niezależnie. W przypadku m -tego elementu, w wyniku eksperymentu, dla zadanej serii identyfikującej o długości N_m , dokonano niezakłóconego pomiaru wartości składowych wektora wyjść. Wyniki zebrano w macierzach:

$$U_{mN_m} \stackrel{\text{df}}{=} [u_{m1} \quad u_{m2} \quad \cdots \quad u_{mN_m}], \quad Y_{mN_m} \stackrel{\text{df}}{=} [y_{m1} \quad y_{m2} \quad \cdots \quad y_{mN_m}]. \quad (4.18)$$

Dla m -tego elementu proponujemy model (4.1). Tak jak w punkcie 2.3.2, jako miarę porównania macierzy pomiarów wartości składowych wektora wyjść obiektu Y_{mN_m} , dla zadanej serii identyfikującej U_{mN_m} z macierzą wartości składowych wektora wyjść modelu $\bar{Y}_{mN_m}(\theta_m)$ wyznaczonych dla U_{mN_m} , wprowadzimy

$$Q_{mN_m}(\theta_m) \stackrel{\text{df}}{=} \left\| Y_{mN_m} - \bar{Y}_{mN_m}(\theta_m) \right\|_{U_{mN_m}}, \quad (4.19)$$

gdzie

$$\bar{Y}_{mN_m}(\theta_m) \stackrel{\text{df}}{=} \left[\Phi_m(u_{m1}, \theta_m) \quad \Phi_m(u_{m2}, \theta_m) \quad \cdots \quad \Phi_m(u_{mN_m}, \theta_m) \right]. \quad (4.20)$$

Jest to lokalne kryterium jakości przybliżenia, zwane lokalnym kryterium jakości identyfikacji. Przykładowe funkcje $Q_{N_m m}(\theta_m)$ mają postać, odpowiednio:

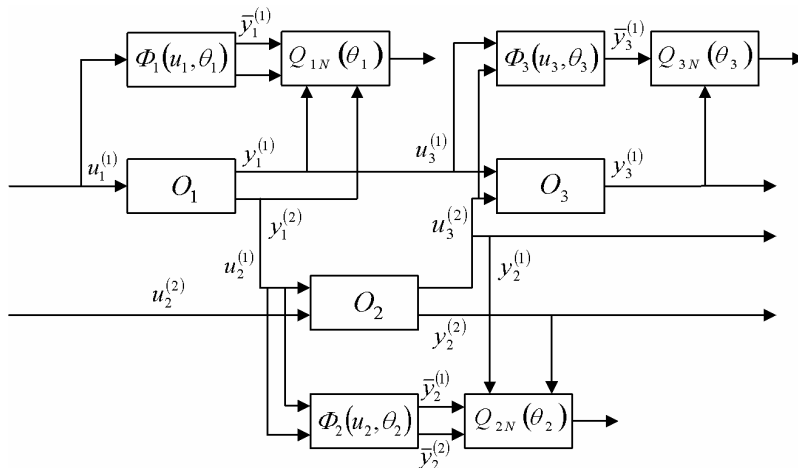
$$Q_{N_m m}(\theta_m) = \sum_{n=1}^{N_m} q_m(y_{mn}, \bar{y}_{mn}) = \sum_{n=1}^{N_m} q_m(y_{mn}, \Phi_m(u_{mn}, \theta_m)) \quad (4.21)$$

lub

$$Q_{mN_m}(\theta_m) = \max_{1 \leq n \leq N_m} \{q_m(y_{mn}, \bar{y}_{mn})\} = \max_{1 \leq n \leq N_m} \{q_m(y_{mn}, \Phi_m(u_{mn}, \theta_m))\}. \quad (4.22)$$

gdzie q_m jest funkcją, której wartość ocenia różnicę pomiędzy zmierzoną wartością składowych wektora wyjść m -tego elementu y_{mn} a wartością składowych wektora wyjść modelu m -tego elementu \bar{y}_{mn} , wyznaczoną dla wektora wejść u_{mn} , czyli $\bar{y}_{mn} = \Phi(u_{mn}, \theta_m)$.

Funkcja $Q_{mN_m}(\theta_m)$ stanowi kryterium wyboru optymalnego modelu dla m -tego elementu, a konkretnie kryterium wyboru wektora optymalnych wartości jego parametrów.



Rys. 4.1. Lokalnie optymalne modele systemu złożonego

Problem wyznaczenia optymalnego modelu sprowadza się do rozwiązania następującego zadania optymalizacji: wyznaczyć taką wartość wektora parametrów $\theta_{mN_m}^*$, dla której miara jakości przybliżenia $Q_{mN_m}(\theta_m)$ przyjmuje wartość minimalną ze względu na θ_m ze zbioru dopuszczalnego \mathcal{O}_m , czyli

$$\theta_{mN_m}^* = \Psi_{mN_m}(U_{mN_m}, Y_{mN_m}) : Q_{mN_m}(\theta_{mN_m}^*) = \min_{\theta_m \in \mathcal{O}_m} Q_{mN_m}(\theta_m), \quad (4.23)$$

gdzie: $\theta_{mN_m}^*$ jest optymalnym wektorem parametrów m -tego modelu, Ψ_{mN_m} – algorytmem identyfikacji, a funkcja (4.1) z wektorem parametrów $\theta_{mN_m}^*$, tj.

$$\bar{y}_m = \Phi_m(u_m, \theta_{mN_m}^*), \quad (4.24)$$

jest nazywana lokalnie optymalnym modelem m -tego elementu z danej klasy modeli.

Zadanie identyfikacji powtarzamy kolejno dla poszczególnych elementów, tzn. $m = 1, 2, \dots, M$.

Teraz, korzystając z równań (4.9) i (4.10), uzyskamy model systemu złożonego danego zależnością (4.17). Wektor wszystkich optymalnych wektorów parametrów modeli lokalnych oznaczymy przez

$$\theta_N^* \stackrel{\text{df}}{=} \begin{bmatrix} \theta_{1N_1}^* \\ \theta_{2N_2}^* \\ \vdots \\ \theta_{MN_M}^* \end{bmatrix}, \quad (4.25)$$

gdzie $N = \sum_{m=1}^M N_m$.

Model (4.17) z parametrami (4.25), tj.

$$\bar{v} = C \bar{\Phi}_y(x, \theta_N^*; A, B) \stackrel{\text{df}}{=} \Phi(x, \theta_N^*), \quad (4.26)$$

nazywamy lokalnie optymalnym modelem systemu złożonego z danej klasy modeli.

4.2. Globalnie optymalny model systemu złożonego

Inne spojrzenie na model wejściowo-wyjściowy systemu złożonego, czyli porównanie tylko wyróżnionych, wskazanych przez projektanta modelu, wyjść systemu złożonego z wyjściami modelu systemu złożonego, danego zależnością (4.17), prowadzi do określenia modelu globalnie optymalnego (rys. 4.2). Teraz założymy, że dla zadanej serii wejść zewnętrznych x o długości N dokonano pomiaru wartości wyróżnionych wyjść systemu złożonego v . Pomiaru zebrano w następujących macierzach:

$$X_N = [x_1 \ x_2 \ \dots \ x_N], \quad V_N = [v_1 \ v_2 \ \dots \ v_N], \quad (4.27)$$

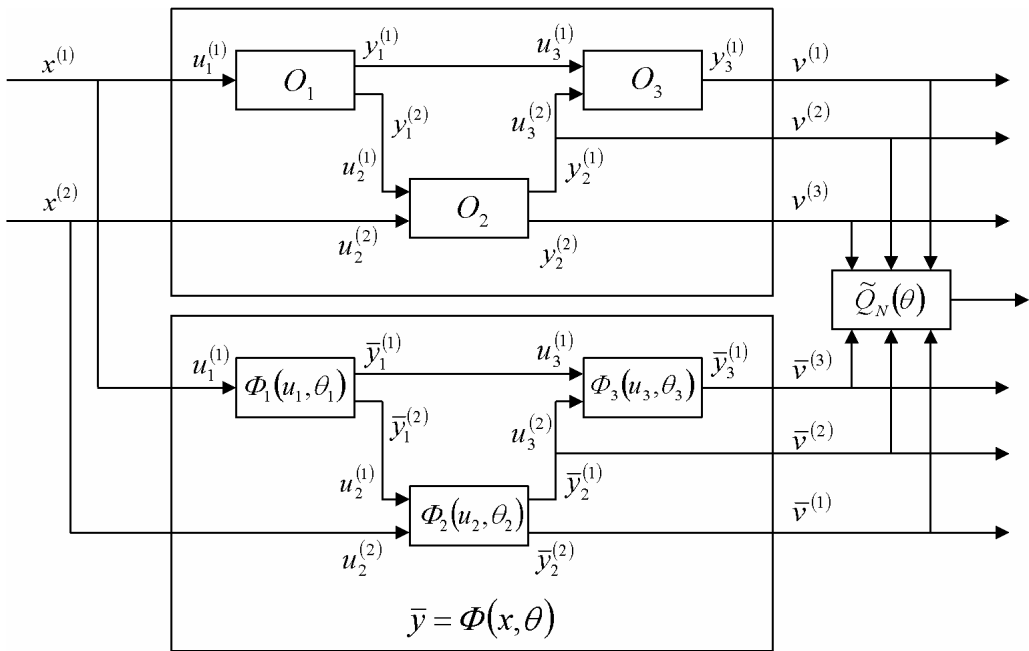
gdzie: X_N jest serią identyfikującą, a V_N jest wynikiem eksperymentu.

Tak jak w punkcie 2.3.2, jako miarę porównania macierzy pomiarów wartości wskazanych wyjść systemu złożonego V_N , dla zadanej serii identyfikującej X_N , z macierzą wartości wskazanych wyjść modelu systemu złożonego $\bar{V}_N(\theta)$, przyjmujemy

$$\tilde{Q}_N(\theta) = \|V_N - \bar{V}_N(\theta)\|_{X_N}, \quad (4.28)$$

gdzie

$$\bar{V}_N(\theta) \stackrel{\text{df}}{=} [\Phi(x_1, \theta) \ \Phi(x_2, \theta) \ \dots \ \Phi(x_N, \theta)]. \quad (4.29)$$



Rys. 4.2. System złożony z modelem globalnie optymalnym

Funkcja $\tilde{Q}_N(\theta)$ (4.28) stanowi globalne kryterium jakości identyfikacji. Przykładowe funkcje $\tilde{Q}_N(\theta)$ mają postać:

$$\tilde{Q}_N(\theta) = \sum_{n=1}^N q(v_n, \bar{v}_n) = \sum_{n=1}^N q(v_n, \Phi(x_n, \theta)) \quad (4.30)$$

lub

$$\tilde{Q}_N(\theta) = \max_{1 \leq n \leq N} \{q(v_n, \bar{v}_n)\} = \max_{1 \leq n \leq N} \{q(v_n, \Phi(x_n, \theta))\}. \quad (4.31)$$

Teraz wartość funkcji q jest miarą różnicy pomiędzy, zaobserwowanymi w n -tym pomiarze, wyróżnionymi wyjściami systemu złożonego v_n a wyróżnionymi wyjściami modelu \bar{v}_n danego zależnością (4.17) dla zadanego wejścia x_n , czyli $\bar{v}_n = \Phi(x_n, \theta)$. Globalnie optymalny wektor parametrów modelu systemu złożonego otrzymujemy w wyniku minimalizacji kryterium (4.28) względem θ , tj.

$$\tilde{\theta}_N^* = \tilde{\Psi}_N(X_N, V_N) : \tilde{Q}_N(\tilde{\theta}_N^*) = \min_{\theta \in \Theta} \tilde{Q}_N(\theta), \quad (4.32)$$

gdzie: $\tilde{\theta}_N^*$ jest optymalnym wektorem parametrów systemu złożonego, $\tilde{\Psi}_N$ – algorytmem identyfikacji, a model (4.17) z parametrami $\tilde{\theta}_N^*$, czyli

$$\bar{v} = C \bar{\Phi}_y(x, \tilde{\theta}_N^*; A, B) \stackrel{\text{df}}{=} \Phi(x, \tilde{\theta}_N^*), \quad (4.33)$$

jest modelem globalnie optymalnym systemu złożonego z danej klasy modeli.

4.3. Model globalnie optymalny z uwzględnieniem jakości modelu lokalnego

W poprzednich podrozdziałach przedstawiono dwa różne modele systemu złożonego, tj. modele optymalne lokalnie i globalnie. Pojawia się pytanie: Który z proponowanych modeli jest lepszy? Odpowiedź: zależy to od projektanta modelu, a konkretnie – od celu, jaki projektant zamierza osiągnąć, budując model systemu złożonego. Do wyznaczania lokalnych decyzji celowe jest budowanie modeli lokalnie optymalnych. W innym przypadku podstawą do projektowania zarządzania nadrzędnego będzie model globalnie optymalny. W dwupoziomowym systemie zarządzania istotna jest zarówno jakość modelu globalnego w zarządzaniu nadrzędnym, jak i jakość modeli lokalnych w podejmowaniu decyzji na dolnym poziomie. Takie postawienie problemu prowadzi do wielokryterialnej oceny modelu systemu złożonego. Z jednej strony stosuje się ocenę lokalną (4.19), osobno dla poszczególnych elementów systemu, a z drugiej strony kryterium (4.28), które ocenia model systemu złożonego jako całość. Korzystając z podejścia wielokryterialnego [85], możemy zaproponować nowe sformułowania zadania wyznaczania optymalnych modeli systemów złożonych.

Jedno z możliwych sformułowań polega na wyznaczeniu globalnie optymalnych parametrów modelu, dla których syntetyczny wskaźnik jakości identyfikacji przyjmuje wartość minimalną. Syntetyczny wskaźnik jakości identyfikacji jest definiowany jako funkcja poszczególnych wskaźników, czyli

$$\bar{Q}_N(\theta) \stackrel{\text{df}}{=} H(\tilde{Q}_N(\theta), Q_{1N}(\theta_1), Q_{2N}(\theta_2), \dots, Q_{MN}(\theta_M)), \quad (4.34)$$

gdzie H jest funkcją $M + 1$ zmiennych, wypukłą ze względu na każdą ze składowych.

Najprostszym przykładem funkcji H jest funkcja liniowa, wówczas syntetyczny wskaźnik jakości identyfikacji określamy zależnością

$$\bar{Q}_N(\theta) = \alpha_0 \tilde{Q}_N(\theta) + \sum_{m=1}^M \alpha_m Q_{mN}(\theta_m), \quad (4.35)$$

w której $\alpha_0, \alpha_1, \dots, \alpha_M$ jest zbiorem nieujemnych współczynników wagowych, spełniających warunek

$$\sum_{m=0}^M \alpha_m = 1, \quad \alpha_m \geq 0, \quad m = 0, 1, \dots, M. \quad (4.36)$$

W tym przypadku współczynniki wagowe $\alpha_0, \alpha_1, \dots, \alpha_M$ przedstawiają udział globalnego wskaźnika oraz poszczególnych, lokalnych wskaźników w syntetycznym wskaźniku jakości identyfikacji. Optymalny wektor parametrów modelu systemu złożonego otrzymamy w wyniku minimalizacji syntetycznego wskaźnika jakości identyfikacji (3.34) względem θ , tj.

$$\bar{\theta}_N^* = \bar{\Psi}_N(X_N, V_N) : \bar{Q}_N(\bar{\theta}_N^*) = \min_{\theta \in \Theta} \bar{Q}_N(\theta), \quad (4.37)$$

gdzie: $\bar{\theta}_N^*$ jest optymalną wartością wektora parametrów modelu systemu złożonego, $\bar{\Psi}_N$ jest algorytmem identyfikacji, a model (4.17) z wektorem parametrów $\bar{\theta}_N^*$, czyli

$$\bar{v} = C \bar{\Phi}_y(x, \bar{\theta}_N^*; A, B) \stackrel{\text{df}}{=} \Phi(x, \bar{\theta}_N^*), \quad (4.38)$$

jest globalnie optymalnym modelem systemu z syntetycznym wskaźnikiem jakości modelu dla zadanej klasy modeli.

Inny sposób uwzględnienia jakości modeli lokalnych polega na wyznaczeniu globalnie optymalnych parametrów modelu przy założeniu, że modele lokalne są zadowalająco dokładne. Znaczy to, że wektory parametrów poszczególnych modeli θ_m spełniają warunek

$$Q_{mN}(\theta_m) \leq \beta_m, \quad \beta_m > Q_{mN}(\theta_{mN}^*), \quad m = 1, 2, \dots, M, \quad (4.39)$$

gdzie β_m jest „miarą zadowolenia” – spodziewaną dokładnością dla m -tego lokalnego modelu.

Wartość ta jest większa od optymalnej wartości lokalnego wskaźnika jakości optymalizacji $Q_{mN}(\theta_{mN}^*)$. Tak sformułowane zadanie wyznaczenia optymalnego modelu wymaga modyfikacji zbioru dopuszczalnych wartości wektora parametrów modelu θ . W zbiorze tym dodatkowo należy uwzględnić ograniczenia (4.39), które zapewniają wymaganą dokładność modeli lokalnych. Zmodyfikowany zbiór dopuszczalnych wartości parametrów modelu systemu złożonego ma więc postać

$$\Theta^* \stackrel{\text{df}}{=} \left\{ \theta \in \Theta \subseteq \mathbb{R}^R : Q_{mN}(\theta_m) \leq \beta_m, \quad \beta_m > Q_{mN}(\theta_{mN}^*), \quad m = 1, 2, \dots, M \right\}. \quad (4.40)$$

Teraz optymalną wartość wektora parametrów modelu systemu złożonego z zadowalającymi wartościami wektorów parametrów modeli lokalnych otrzymamy w wyniku minimalizacji globalnego wskaźnika jakości identyfikacji (4.28) względem θ , przy ograniczeniach (4.40), czyli

$$\tilde{\theta}_N^{**} = \tilde{\Psi}_N^*(X_N, V_N) : \tilde{Q}_N(\tilde{\theta}_N^{**}) = \min_{\theta \in \Theta^*} \tilde{Q}_N(\theta), \quad (4.41)$$

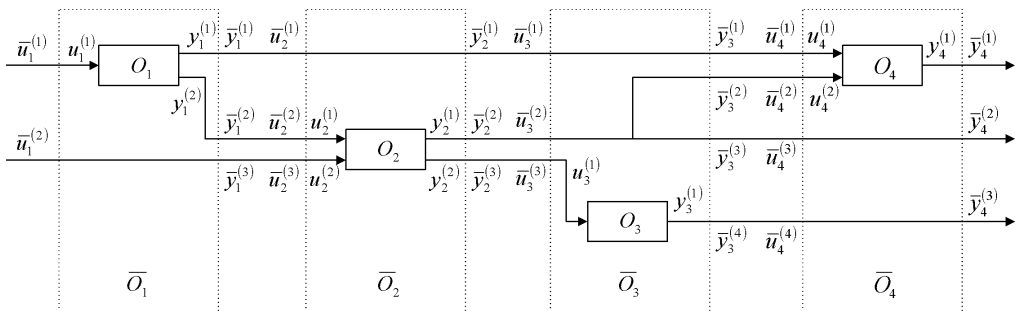
gdzie: $\tilde{\theta}_N^{**}$ jest optymalną wartością wektora parametrów systemu złożonego, $\tilde{\Psi}_N^*$ – algorytmem identyfikacji, a model (4.17) z parametrami $\tilde{\theta}_N^{**}$, czyli

$$\bar{v} = C \bar{\Phi}_y(x, \tilde{\theta}_N^{**}; A, B) \stackrel{\text{df}}{=} \Phi(x, \tilde{\theta}_N^{**}), \quad (4.42)$$

jest globalnie optymalnym modelem systemu złożonego z zadowalająco dobrymi modelami lokalnymi.

4.4. Algorytm identyfikacji dla systemu złożonego o strukturze szeregowej

Bardzo ważnym przypadkiem obiektu złożonego jest system o strukturze szeregowej. System złożony, który nie zawiera sprzężeń zwrotnych (rys. 4.3), można sprowadzić do systemu o strukturze szeregowej (rys. 4.4).



Rys. 4.3. System złożony

Przyjmijmy modele elementów O_1 , O_2 oraz O_3 systemu złożonego z rysunku 4.3, odpowiednio:

$$y_1 = \begin{bmatrix} y_1^{(1)} \\ y_1^{(2)} \end{bmatrix} = \begin{bmatrix} \Phi_1^{(1)}(u_1^{(1)}, \theta_1) \\ \Phi_1^{(2)}(u_1^{(1)}, \theta_1) \end{bmatrix} = \Phi_1(u_1, \theta_1), \quad (4.43)$$

$$y_2 = \begin{bmatrix} y_2^{(1)} \\ y_2^{(2)} \end{bmatrix} = \begin{bmatrix} \Phi_2^{(1)}(u_2^{(1)}, u_2^{(2)}, \theta_2) \\ \Phi_2^{(2)}(u_2^{(1)}, u_2^{(2)}, \theta_2) \end{bmatrix} = \Phi_2(u_2, \theta_2), \quad (4.44)$$

$$y_3 = [y_3^{(1)}] = [\Phi_3^{(1)}(u_3^{(1)}, \theta_3)] = \Phi_3(u_3, \theta_3), \quad (4.45)$$

$$y_4 = [y_4^{(1)}] = [\Phi_4^{(1)}(u_4^{(1)}, u_4^{(2)}, \theta_4)] = \Phi_4(u_4, \theta_4). \quad (4.46)$$

Zdefiniujemy nowe, pomocnicze wejścia i wyjścia w następujący sposób:

$$\bar{u}_1 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{u}_1^{(1)} \\ \bar{u}_1^{(2)} \\ \bar{u}_1^{(3)} \end{bmatrix}, \quad \bar{u}_2 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{u}_2^{(1)} \\ \bar{u}_2^{(2)} \\ \bar{u}_2^{(3)} \end{bmatrix}, \quad \bar{u}_3 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{u}_3^{(1)} \\ \bar{u}_3^{(2)} \\ \bar{u}_3^{(3)} \end{bmatrix}, \quad \bar{u}_4 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{u}_4^{(1)} \\ \bar{u}_4^{(2)} \\ \bar{u}_4^{(3)} \\ \bar{u}_4^{(4)} \end{bmatrix}, \quad (4.47)$$

$$\bar{y}_1 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{y}_1^{(1)} \\ \bar{y}_1^{(2)} \\ \bar{y}_1^{(3)} \end{bmatrix}, \quad \bar{y}_2 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{y}_2^{(1)} \\ \bar{y}_2^{(2)} \\ \bar{y}_2^{(3)} \end{bmatrix}, \quad \bar{y}_3 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{y}_3^{(1)} \\ \bar{y}_3^{(2)} \\ \bar{y}_3^{(3)} \\ \bar{y}_3^{(4)} \end{bmatrix}, \quad \bar{y}_4 \stackrel{\text{df}}{=} \begin{bmatrix} \bar{y}_4^{(1)} \\ \bar{y}_4^{(2)} \\ \bar{y}_4^{(3)} \end{bmatrix}.$$

Teraz możemy zdefiniować nowe elementy $\bar{O}_1, \bar{O}_2, \bar{O}_3$ oraz \bar{O}_4 , dla których proponujemy modele:

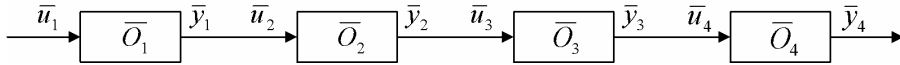
$$\bar{y}_1 = \begin{bmatrix} \bar{y}_1^{(1)} \\ \bar{y}_1^{(2)} \\ \bar{y}_1^{(3)} \end{bmatrix} = \begin{bmatrix} \Phi_1^{(1)}(\bar{u}_1^{(1)}, \theta_1) \\ \Phi_1^{(2)}(\bar{u}_1^{(1)}, \theta_1) \\ \bar{u}_1^{(2)} \end{bmatrix} \stackrel{\text{df}}{=} \bar{\Phi}_1(\bar{u}_1, \bar{\theta}_1), \quad (4.48)$$

$$\bar{y}_2 = \begin{bmatrix} \bar{y}_2^{(1)} \\ \bar{y}_2^{(2)} \\ \bar{y}_2^{(3)} \end{bmatrix} = \begin{bmatrix} \bar{u}_2^{(1)} \\ \Phi_2^{(1)}(\bar{u}_2^{(1)}, \bar{u}_2^{(2)}, \theta_2) \\ \Phi_2^{(2)}(\bar{u}_2^{(1)}, \bar{u}_2^{(2)}, \theta_2) \end{bmatrix} \stackrel{\text{df}}{=} \bar{\Phi}_2(\bar{u}_2, \bar{\theta}_2), \quad (4.49)$$

$$\bar{y}_3 = \begin{bmatrix} \bar{y}_3^{(1)} \\ \bar{y}_3^{(2)} \\ \bar{y}_3^{(3)} \\ \bar{y}_3^{(4)} \end{bmatrix} = \begin{bmatrix} \bar{u}_3^{(1)} \\ \bar{u}_3^{(2)} \\ \bar{u}_3^{(2)} \\ \Phi_3^{(1)}(\bar{u}_3^{(3)}, \theta_3) \end{bmatrix} \stackrel{\text{df}}{=} \bar{\Phi}_3(\bar{u}_3, \bar{\theta}_3), \quad (4.50)$$

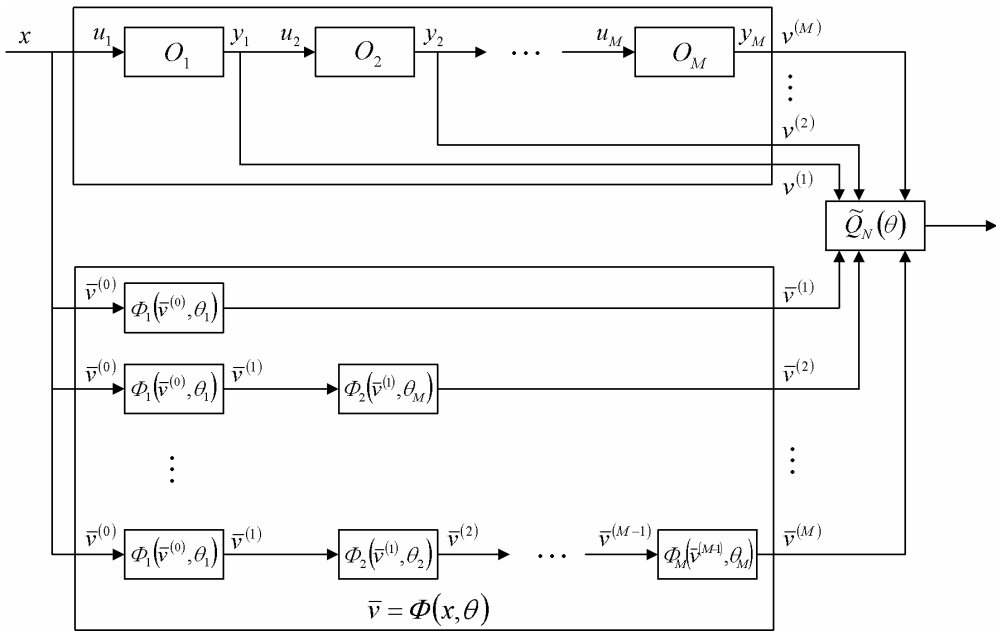
$$\bar{y}_4 = \begin{bmatrix} \bar{y}_4^{(1)} \\ \bar{y}_4^{(2)} \\ \bar{y}_4^{(3)} \end{bmatrix} = \begin{bmatrix} \Phi_4^{(1)}(\bar{u}_4^{(1)}, \bar{u}_4^{(2)}, \theta_4) \\ \bar{u}_4^{(3)} \\ \bar{u}_4^{(4)} \end{bmatrix} \stackrel{\text{df}}{=} \bar{\Phi}_4(\bar{u}_4, \bar{\theta}_4). \quad (4.51)$$

Przy tak zdefiniowanych elementach system złożony o strukturze przedstawionej na rysunku 4.3 sprowadza się do systemu o strukturze szeregowej (rys. 4.4).



Rys. 4.4. System złożony o strukturze szeregowej

Specyfika systemu o strukturze szeregowej umożliwia zastosowanie metody programowania dynamicznego do wyznaczenia algorytmu identyfikacji. Przyjmujemy, że system o strukturze szeregowej zawiera M elementów, dla których zakładamy modele postaci (4.1).



Rys. 4.5. System złożony o strukturze szeregowej z modelem globalnie optymalnym gdy wyróżniono wszystkie wyjścia

W najogólniejszym przypadku jako wyróżnione wyjścia systemu złożonego wybierzmy wszystkie wyjścia (rys. 4.5). Macierz C w równaniu (4.7) jest macierzą jednostkową o wymiarze $L \times L$, tj. $C = I_{L \times L}$, czyli równanie (4.7) ma postać

$$v = y. \quad (4.52)$$

Po uwzględnieniu proponowanych modeli (4.1) dla poszczególnych elementów struktury systemu złożonego oraz równania (4.52), model systemu złożonego (4.17) przyjmuje postać

$$\bar{v} = \begin{bmatrix} \bar{v}^{(1)} \\ \bar{v}^{(2)} \\ \vdots \\ \bar{v}^{(M)} \end{bmatrix} = \begin{bmatrix} \Phi_1(x, \theta_1) \\ \Phi_2(\bar{v}^{(1)}, \theta_2) \\ \vdots \\ \Phi_M(\bar{v}^{(M-1)}, \theta_M) \end{bmatrix}, \quad (4.53)$$

a po podstawieniu modeli (4.1), $m = 1, 2, \dots, M$, do (4.53) otrzymujemy

$$\bar{v} = \begin{bmatrix} \bar{v}^{(1)} \\ \bar{v}^{(2)} \\ \vdots \\ \bar{v}^{(M)} \end{bmatrix} = \begin{bmatrix} \Phi_1(x, \theta_1) \\ \Phi_2(\Phi_1(x, \theta_1), \theta_2) \\ \vdots \\ \Phi_M(\Phi_{M-1}(\dots \Phi_1(x, \theta_1), \dots \theta_{M-1}), \theta_M) \end{bmatrix}. \quad (4.54)$$

Zauważmy, że elementy wektora \bar{v} (4.53), odpowiadające wyjściom poszczególnych elementów modelu, możemy przedstawić w postaci rekurencyjnej, tj.

$$\bar{v}^{(m+1)} = \Phi_m(\bar{v}^{(m)}, \theta_m), \quad m = 0, 1, \dots, M, \quad (4.55)$$

gdzie

$$\bar{v}^{(0)} \stackrel{\text{df}}{=} x. \quad (4.56)$$

W dalszych rozważaniach ograniczymy się do zadania wyboru modelu globalnie optymalnego, przedstawionego w punkcie 4.2.3, przyjmując globalny wskaźnik jakości identyfikacji (4.28) postaci (4.30), czyli

$$\tilde{Q}_N(\theta) = \sum_{n=1}^N q(v_n, \bar{v}_n) = \sum_{n=1}^N q(v_n, \Phi(x_n, \theta)). \quad (4.57)$$

Przypomnijmy, że zgodnie z (4.52) wektor wyróżnionych wyjść v jest wektorem wszystkich wyjść poszczególnych elementów. Zaproponujemy funkcję q w kryterium (4.57), która pozwoli na ocenę różnicy pomiędzy – zaobserwowanymi w n -tym pomiarze – wartościami składowych wyjść poszczególnych elementów systemu złożonego a wartościami składowych wyjść ich modeli. Niech

$$\begin{aligned} q(v_n, \bar{v}_n) &= \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \bar{v}_n^{(m)}) \\ &= \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \Phi_m(\Phi_{m-1}(\dots \Phi_1(x_n, \theta_1) \dots, \theta_{m-1}), \theta_m)), \end{aligned} \quad (4.58)$$

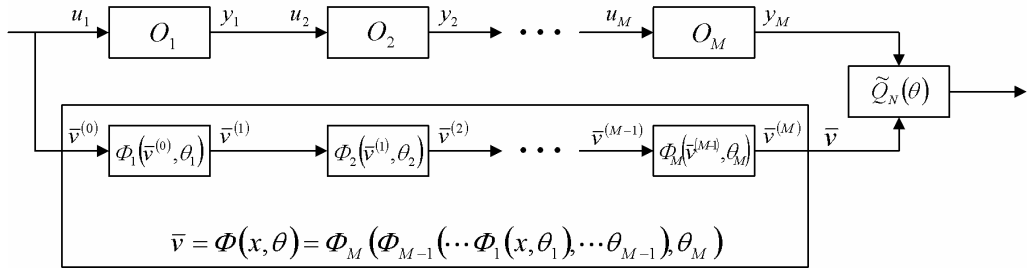
gdzie: wartość funkcji q_m jest miarą różnicy pomiędzy – zaobserwowanymi w n -tym pomiarze – wartościami składowych wektora wyjść m -tego elementu systemu złożonego $v_n^{(m)}$ a wartościami składowych wektora wyjść modelu m -tego elementu $\bar{v}_n^{(m)}$ (4.54) dla zadanej wartości składowych wektora wejść x_n , natomiast λ_m , $m = 1, 2, \dots, M$, jest

zestawem współczynników wagowych umożliwiającą swobodne uwzględnienie modeli poszczególnych elementów w modelu globalnym.

Zestaw współczynników spełnia warunek

$$0 \leq \lambda_m \leq 1, \quad m = 1, 2, \dots, M-1, \quad \lambda_M \neq 0. \quad (4.59)$$

W szczególnym przypadku, gdy $\lambda_m = 1, m = 1, 2, \dots, M$, model globalny przyjmuje strukturę jak na rysunku 4.5, jeżeli natomiast $\lambda_M = 1$ oraz $\lambda_m = 0, m = 1, 2, \dots, M-1$, model globalny przyjmuje strukturę jak na rysunku 4.6, czyli oceniamy wyłącznie wektor wyjść ostatniego elementu.



Rys. 4.6. Globalnie optymalny model o strukturze szeregowej gdy wyróżnione wyjście jest wyjściem ostatniego elementu

Po uwzględnieniu rekurencyjnej postaci modelu globalnego (4.55) w funkcji (4.58) otrzymujemy

$$q(v_n, \bar{v}_n) = \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \bar{v}_n^{(m)}) = \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)). \quad (4.60)$$

Ostatecznie globalny wskaźnik jakości identyfikacji (4.57) możemy zapisać

$$\tilde{Q}_N(\theta) = \sum_{n=1}^N \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \bar{v}_n^{(m)}) = \sum_{n=1}^N \sum_{m=1}^M \lambda_m q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)), \quad (4.61)$$

a po zamianie kolejności sumowania otrzymujemy

$$\tilde{Q}_N(\theta) = \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \bar{v}_n^{(m)}) = \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)). \quad (4.62)$$

4.4.1. Globalnie optymalny model systemu o strukturze szeregowej

Globalny wskaźnik jakości postaci (4.62) oraz rekurencyjny model (4.55) umożliwiają zastosowanie metody programowania dynamicznego do rozwiązania zadania (4.32) [126, 127]. Po wyodrębnieniu wektorów parametrów modeli poszczególnych elementów zadanie (4.32) sprowadza się do wyznaczenia takiego zestawu wektorów

parametrów modeli elementów $\tilde{\theta}_{1N}^*, \tilde{\theta}_{2N}^*, \dots, \tilde{\theta}_{MN}^*$, który minimalizuje wskaźnik (4.62), czyli

$$\begin{aligned} & \tilde{\theta}_{1N}^*, \tilde{\theta}_{2N}^*, \dots, \tilde{\theta}_{MN}^* : \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \tilde{\theta}_{mn}^*)) \\ & = \min_{\theta_1, \theta_2, \dots, \theta_M \in \Theta_1 \times \Theta_2 \times \dots \times \Theta_M} \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)). \end{aligned} \quad (4.63)$$

Zastosowanie metody programowania dynamicznego daje następujący algorytm identyfikacji przy globalnym wskaźniku jakości (4.62) dla struktury szeregowej:

Krok 1. Oznaczamy ostatni M -ty składnik sumy w kryterium (4.62) przez

$$\tilde{Q}_{MN}(\theta_M) \stackrel{\text{df}}{=} \lambda_M \sum_{n=1}^N q_M(v_n^{(M)}, \Phi_M(\bar{v}_n^{(M-1)}, \theta_M)). \quad (4.64)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{MN}^*$, który minimalizuje wskaźnik (4.64). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\tilde{\Psi}_{MN}$

$$\tilde{\theta}_{MN}^* = \tilde{\Psi}_{MN}(V_N^{(M)}, \bar{V}_N^{(M-1)}), \quad (4.65)$$

który minimalizuje wskaźnik (4.64) względem θ_M , czyli

$$\tilde{\theta}_{MN}^* = \tilde{\Psi}_{MN}(V_N^{(M)}, \bar{V}_N^{(M-1)}) : \tilde{Q}_{MN}(\tilde{\theta}_{MN}^*) = \min_{\theta_M \in \Theta_M} \tilde{Q}_{MN}(\theta_M), \quad (4.66)$$

gdzie $V_N^{(M)}$ jest macierzą pomiarów wartości składowych wektora wyjść M -tego elementu, tj.

$$V_N^{(M)} = [v_1^{(M)} \ v_2^{(M)} \ \dots \ v_N^{(M)}], \quad (4.67)$$

natomiast $\bar{V}_N^{(M-1)}$ jest macierzą wartości składowych wektora wejść modelu M -tego elementu, która jest również macierzą wartości składowych wektora wyjść modelu $M-1$. elementu, czyli

$$\bar{V}_N^{(M-1)} = [\bar{v}_1^{(M-1)} \ \bar{v}_2^{(M-1)} \ \dots \ \bar{v}_N^{(M-1)}]. \quad (4.68)$$

W algorytmie (4.65) występuje macierz (4.68), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść poprzedniego $M-1$. elementu.

Po podstawieniu do (4.64) algorytmu wyznaczania optymalnego wektora parametrów $\tilde{\theta}_{MN}^*$ modelu M -tego elementu, uzyskanego w wyniku rozwiązania zadania (4.66), otrzymujemy optymalną postać kryterium, zależną od macierzy $\bar{V}_N^{(M-1)}$ (4.68)

$$\tilde{Q}_{MN}(\tilde{\theta}_{MN}^*) = \tilde{Q}_{MN}(\Psi_{MN}(V_N^{(M)}, \bar{V}_N^{(M-1)})) \stackrel{\text{df}}{=} \tilde{Q}_{MN}^*(V_N^{(M)}, \bar{V}_N^{(M-1)}). \quad (4.69)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.68) mamy

$$\begin{aligned} \bar{V}_N^{(M-1)} &= \left[\Phi_{M-1}(\bar{v}_1^{(M-2)}, \theta_{M-1}) \Phi_{M-1}(\bar{v}_2^{(M-2)}, \theta_{M-1}) \cdots \Phi_{M-1}(\bar{v}_N^{(M-2)}, \theta_{M-1}) \right] \\ &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1}), \end{aligned} \quad (4.70)$$

a po podstawieniu (4.70) do (4.69) otrzymujemy

$$\tilde{Q}_{MN}^*(V_N^{(M)}, \bar{V}_N^{(M-1)}) = \tilde{Q}_{MN}^*(V_N^{(M)}, \tilde{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1})). \quad (4.71)$$

Krok 2. Oznaczamy sumę przedostatniego $M - 1$. składnika w kryterium (4.62) i \tilde{Q}_{MN}^* – optymalnej postaci (4.71) składowej kryterium, wyznaczonej w kroku 1., przez

$$\begin{aligned} \tilde{Q}_{M-1N}(\theta_{M-1}) &\stackrel{\text{df}}{=} \lambda_{M-1} \sum_{n=1}^N q_{M-1}(v_n^{(M-1)}, \Phi_{M-1}(\bar{v}_n^{(M-2)}, \theta_{M-1})) \\ &+ \tilde{Q}_{MN}^*(V_N^{(M)}, \tilde{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1})). \end{aligned} \quad (4.72)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{M-1N}^*$, który minimalizuje wskaźnik (4.72). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\tilde{\Psi}_{M-1N}$

$$\tilde{\theta}_{M-1N}^* = \tilde{\Psi}_{M-1N}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-2)}), \quad (4.73)$$

który minimalizuje wskaźnik (4.72) względem θ_{M-1} , czyli

$$\begin{aligned} \tilde{\theta}_{M-1N}^* &= \tilde{\Psi}_{M-1N}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-2)}): \tilde{Q}_{M-1N}(\tilde{\theta}_{M-1N}^*) \\ &= \min_{\theta_{M-1} \in \Theta_{M-1}} \tilde{Q}_{M-1N}(\theta_{M-1}), \end{aligned} \quad (4.74)$$

gdzie $V_N^{(M-1)}$ jest macierzą pomiarów wartości składowych wektora wyjść $M - 1$. elementu, tj.

$$V_N^{(M-1)} = [v_1^{(M-1)} \ v_2^{(M-1)} \ \dots \ v_N^{(M-1)}], \quad (4.75)$$

natomiast $\bar{V}_N^{(M-2)}$ jest macierzą wartości składowych wektora wejść modelu $M - 1$. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu $M - 2$. elementu, czyli

$$\bar{V}_N^{(M-2)} = [\bar{v}_1^{(M-2)} \ \bar{v}_2^{(M-2)} \ \dots \ \bar{v}_N^{(M-2)}]. \quad (4.76)$$

W algorytmie (4.73) występuje macierz (4.76), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść poprzedniego $M - 2$. elementu.

Po podstawieniu do (4.72) algorytmu wyznaczania optymalnego wektora parametrów $\tilde{\theta}_{M-1N}^*$ modelu $M - 1$. elementu, uzyskanego w wyniku rozwiązania zadania (4.74), otrzymujemy optymalną postać kryterium, zależną od macierzy \bar{V}_N^{M-2} (4.76)

$$\begin{aligned}\tilde{Q}_{M-1N}(\tilde{\theta}_{M-1N}^*) &= \tilde{Q}_{M-1N}(\tilde{\Psi}_{M-1N}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-2)})) \\ &\stackrel{\text{df}}{=} \tilde{Q}_{M-1N}^*(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-2)}).\end{aligned}\quad (4.77)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.76) mamy

$$\begin{aligned}\bar{V}_N^{(M-2)} &= [\Phi_{M-2}(\bar{v}_1^{(M-3)}, \theta_{M-2}) \Phi_{M-2}(\bar{v}_2^{(M-3)}, \theta_{M-2}) \cdots \Phi_{M-2}(\bar{v}_N^{(M-3)}, \theta_{M-2})] \\ &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-2}(\bar{V}_N^{(M-3)}, \theta_{M-2}),\end{aligned}\quad (4.78)$$

a po podstawieniu (4.78) do (4.77) otrzymujemy

$$\begin{aligned}\tilde{Q}_{M-1N}^*(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-2)}) &= \tilde{Q}_{M-1N}^*(V_N^{(M)}, V_N^{(M-1)}, \tilde{\Phi}_{M-2}(\bar{V}_N^{(M-3)}, \theta_{M-2})). \\ &\vdots\end{aligned}\quad (4.79)$$

Krok m . Oznaczamy sumę $M - m + 1$. składnika w kryterium (4.62) i \tilde{Q}_{M-m+2N}^* – optymalnej postaci składowej kryterium wyznaczonej w $m - 1$. kroku – przez

$$\begin{aligned}\tilde{Q}_{M-m+1N}(\theta_{M-m+1}) &\stackrel{\text{df}}{=} \lambda_{M-m+1} \sum_{n=1}^N q_{M-m+1}(v_n^{(M-m+1)}, \Phi_{M-m+1}(\bar{v}^{(M-m)}, \theta_{M-m+1})) \\ &+ \tilde{Q}_{M-m+2N}^*(V_N^{(M)}, \dots, V_N^{(M-m+2)}, \tilde{\Phi}_{M-m+1}(\bar{V}_N^{(M-m)}, \theta_{M-m+1})).\end{aligned}\quad (4.80)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{M-m+1N}^*$, który minimalizuje wskaźnik (4.80). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\tilde{\Psi}_{M-m+1N}$

$$\tilde{\theta}_{M-m+1N}^* = \tilde{\Psi}_{M-m+1N}(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)}), \quad (4.81)$$

który minimalizuje wskaźnik (4.80) względem θ_{M-m+1} , czyli

$$\begin{aligned}\tilde{\theta}_{M-m+1N}^* &= \tilde{\Psi}_{M-m+1N}(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)}) : \tilde{Q}_{M-m+1N}(\tilde{\theta}_{M-m+1N}^*) \\ &= \min_{\theta_{M-m+1} \in \Theta_{M-m+1}} \tilde{Q}_{M-m+1N}(\theta_{M-m+1}),\end{aligned}\quad (4.82)$$

gdzie $V_N^{(M-m+1)}$ jest macierzą pomiarów wartości składowych wektora wyjść $M - m + 1$. elementu, tj.

$$V_N^{(M-m+1)} = [v_1^{(M-m+1)} \ v_2^{(M-m+1)} \ \dots \ v_N^{(M-m+1)}], \quad (4.83)$$

natomiast $\bar{V}_N^{(M-m)}$ jest macierzą wartości składowych wektora wejść modelu $M - m + 1$. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu $M - m$ -tego elementu, czyli

$$\bar{V}_N^{(M-m)} \stackrel{\text{df}}{=} [\bar{v}_1^{(M-m)} \ \bar{v}_2^{(M-m)} \ \dots \ \bar{v}_N^{(M-m)}]. \quad (4.84)$$

W algorytmie (4.81) występuje macierz (4.84), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść poprzedniego $M - m$ -tego elementu.

Po podstawieniu do (4.80) algorytmu wyznaczania optymalnego wektora parametrów $\tilde{\theta}_{M-m+1N}^*$ modelu $M - m + 1$. elementu, uzyskanego w wyniku rozwiązania zadania (4.82), otrzymujemy optymalną postać kryterium

$$\begin{aligned} \tilde{Q}_{M-m+1N}(\tilde{\theta}_{M-m+1N}^*) &= Q_{M-m+1N}(\tilde{\Psi}_{M-m+1N}(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)})) \\ &\stackrel{\text{df}}{=} \tilde{Q}_{M-m+1N}^*(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)}). \end{aligned} \quad (4.85)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.84) mamy

$$\begin{aligned} \bar{V}_N^{(M-m)} &= [\Phi_{M-m}(\bar{v}_1^{(M-m-1)}, \theta_{M-m}) \Phi_{M-m}(\bar{v}_2^{(M-m-1)}, \theta_{M-m}) \dots \Phi_{M-m}(\bar{v}_N^{(M-m-1)}, \theta_{M-m})] \\ &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-m}(\bar{V}_N^{(M-m-1)}, \theta_{M-m}), \end{aligned} \quad (4.86)$$

a po podstawieniu (4.86) do (4.85) otrzymujemy

$$\begin{aligned} &\tilde{Q}_{M-m+1N}^*(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)}) \\ &= \tilde{Q}_{M-m+1N}^*\left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \tilde{\Phi}_{M-m}(\bar{V}_N^{(M-m-1)}, \theta_{M-m})\right). \end{aligned} \quad (4.87)$$

⋮

Krok $M - 1$. Oznaczamy sumę 2. składnika w kryterium (4.62) i \tilde{Q}_{3N}^* – optymalnej postaci składowej kryterium wyznaczonej w $M - 2$. kroku – przez

$$\begin{aligned} \tilde{Q}_{2N}(\theta_2) &\stackrel{\text{df}}{=} \lambda_2 \sum_{n=1}^N q_2(v_n^{(2)}, \Phi_{M-m+1}(\bar{v}^{(1)}, \theta_2)) \\ &+ \tilde{Q}_{3N}^*(V_N^{(M)}, \dots, V_N^{(3)}, \tilde{\Phi}_2(\bar{V}_N^{(1)}, \theta_2)). \end{aligned} \quad (4.88)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{2N}^*$, który minimalizuje wskaźnik (4.88). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\tilde{\Psi}_{2N}$

$$\tilde{\theta}_{2N}^* = \tilde{\Psi}_{2N}(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{(1)}), \quad (4.89)$$

który minimalizuje wskaźnik (4.88) względem θ_2 , czyli

$$\tilde{\theta}_{2N}^* = \tilde{\Psi}_{2N}(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{(1)}) : \tilde{Q}_{2N}(\tilde{\theta}_{2N}^*) = \min_{\theta_2 \in \Theta_2} \tilde{Q}_{2N}(\theta_2), \quad (4.90)$$

gdzie: $V_N^{(2)}$ jest macierzą pomiarów wartości składowych wektora wyjść 2. elementu, tj.

$$V_N^{(2)} = [v_1^{(2)} \ v_2^{(2)} \ \dots \ v_N^{(2)}], \quad (4.91)$$

natomiast $\bar{V}_N^{(1)}$ jest macierzą wartości składowych wektora wejść modelu 2. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu 1. elementu, czyli

$$\bar{V}_N^{(1)} = [\bar{v}_1^{(1)} \ \bar{v}_2^{(1)} \ \dots \ \bar{v}_N^{(1)}]. \quad (4.92)$$

W algorytmie (4.89) występuje macierz (4.92), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść poprzedniego – pierwszego elementu.

Po podstawieniu do (4.88) algorytmu wyznaczania optymalnego wektora parametrów modelu 2. elementu $\tilde{\theta}_{2N}^*$, uzyskanego w wyniku rozwiązania zadania (4.90), otrzymujemy optymalną postać kryterium

$$\tilde{Q}_{2N}(\tilde{\theta}_{2N}^*) = \tilde{Q}_{2N}(\tilde{\Psi}_{2N}(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{(1)})) \stackrel{\text{df}}{=} \tilde{Q}_{M-m+1N}^*(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{(1)}). \quad (4.93)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.92) mamy

$$\bar{V}_N^{(1)} = [\Phi_{M-2}(\bar{v}_1^{(0)}, \theta_1) \ \Phi_{M-m}(\bar{v}_2^{(0)}, \theta_1) \ \dots \ \Phi_{M-2}(\bar{v}_N^{(0)}, \theta_1)] \stackrel{\text{df}}{=} \tilde{\Phi}_1(\bar{V}_N^{(0)}, \theta_1), \quad (4.94)$$

a po podstawieniu (4.94) do (4.93) otrzymujemy

$$\tilde{Q}_{2N}^*(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{(1)}) = \tilde{Q}_{2N}^*(V_N^{(M)}, \dots, V_N^{(2)}, \tilde{\Phi}_1(\bar{V}_N^{(0)}, \theta_1)). \quad (4.95)$$

Krok M. Oznaczamy sumę 1. składnika w kryterium (4.62) i \tilde{Q}_{2N}^* – optymalnej postaci składowej kryterium wyznaczonej w $M-1$. kroku – przez

$$\tilde{Q}_{1N}(\theta_1) \stackrel{\text{df}}{=} \lambda_1 \sum_{n=1}^N q_1(v_n^{(1)}, \Phi_1(\bar{v}_n^{(0)}, \theta_1)) + \tilde{Q}_{2N}^*(V_N^{(M)}, \dots, V_N^{(2)}, \tilde{\Phi}_1(\bar{V}_N^{(0)}, \theta_1)). \quad (4.96)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{1N}^*$, który minimalizuje wskaźnik (4.96). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\tilde{\Psi}_{1N}$

$$\tilde{\theta}_{1N}^* = \tilde{\Psi}_{1N}(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(0)}), \quad (4.97)$$

który minimalizuje wskaźnik (4.96) względem θ_1 , czyli

$$\tilde{\theta}_{1N}^* = \tilde{\Psi}_{1N}(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(0)}): \tilde{Q}_{1N}(\tilde{\theta}_{1N}^*) = \min_{\theta_1 \in \Theta_1} \tilde{Q}_{1N}(\theta_1), \quad (4.98)$$

gdzie $V_N^{(1)}$ jest macierzą pomiarów wartości składowych wektora wyjść 1. elementu, tj.

$$V_N^{(1)} = [v_1^{(1)} \ v_2^{(1)} \ \dots \ v_N^{(1)}], \quad (4.99)$$

natomiast $\bar{V}_N^{(0)}$, zgodnie z (4.56), jest macierzą wartości składowych wektora wejść zewnętrznych, czyli

$$\bar{V}_N^{(0)} = [\bar{v}_1^{(0)} \bar{v}_2^{(0)} \dots \bar{v}_N^{(0)}] = [x_1 \ x_2 \ \dots \ x_N] \stackrel{\text{df}}{=} X_N. \quad (4.100)$$

Dopiero teraz wszystkie występujące w algorytmie (4.97) macierze, z uwzględnieniem (4.100), są znane jako macierze pomiarów. Możliwe jest wyznaczenie wartości składowych wektora parametrów modelu 1. elementu.

Optymalne wartości składowych wektorów parametrów $\tilde{\theta}_{MN}^*$, $\tilde{\theta}_{M-1N}^*$, ..., $\tilde{\theta}_{2N}^*$, wyznaczone w krokach 1., 2., ..., $M-1$., oprócz wyników pomiarów wyjść

$$V_N^{(M)}, V_N^{(M-1)}, \dots, V_N^{(1)}, \quad (4.101)$$

zależą od macierzy wartości składowych wektorów wyjść modeli

$$\bar{V}_N^{(M-1)}, \bar{V}_N^{(M-1)}, \dots, \bar{V}_N^{(1)}. \quad (4.102)$$

Dopiero wektor $\tilde{\theta}_{1N}^*$ zależy od macierzy pomiarów wartości składowych wektora wejść zewnętrznych X_N oraz macierzy pomiarów wartości składowych wszystkich wektorów wyjść $V_N^{(1)}, V_N^{(2)}, \dots, V_N^{(M)}$.

Powracając do algorytmów identyfikacji określonych w kolejnych krokach, możemy wyznaczyć optymalne wartości parametrów poszczególnych elementów systemu. I tak:

Krok $M+1$. Po uwzględnieniu w (4.97) macierzy pomiarów wejść zewnętrznych (4.100) otrzymujemy algorytm

$$\tilde{\theta}_{1N}^* = \tilde{\Psi}_{1N}(V_N^{(M)}, \dots, V_N^{(1)}, X_N), \quad (4.103)$$

który na podstawie macierzy pomiarów wartości składowych wektora wejść zewnętrznych, tj. serii identyfikującej X_N , oraz macierzy pomiarów wartości składowych wszystkich wektorów wyjść $V_N^{(1)}, V_N^{(2)}, \dots, V_N^{(M)}$ umożliwia wyznaczenie wartości optymalnego wektora parametrów modelu 1. elementu $\tilde{\theta}_{1N}^*$.

Krok $M+2$. Teraz, po podstawieniu do (4.94) w miejsce θ_1 optymalnych wartości $\tilde{\theta}_{1N}^*$, określonych przez (4.103), otrzymujemy macierz wartości składowych wektora wyjść modelu 1. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , tj.

$$\bar{V}_N^{*(1)} \stackrel{\text{df}}{=} \tilde{\Phi}_1(X_N, \tilde{\theta}_{1N}^*) = \tilde{\Phi}_1(X_N, \tilde{\Psi}_{1N}(V_N^{(M)}, \dots, V_N^{(1)}, X_N)). \quad (4.104)$$

Po podstawieniu w zależności (4.89) w miejsce $\bar{V}_N^{(1)}$ macierzy wartości składowych wektora wyjść modelu optymalnego, określonego przez (4.104), otrzymujemy

$$\tilde{\theta}_{2N}^* = \tilde{\Psi}_{2N}(V_N^{(M)}, \dots, V_N^{(2)}, \bar{V}_N^{*(1)}). \quad (4.105)$$

⋮

Krok $M + m$. W kroku $M + m - 1$. wyznaczyliśmy macierz wartości składowych wektora wyjść $\bar{V}_N^{*(m-2)}$ modelu $m - 2$. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , która jest jednocześnie macierzą wartości składowych wektora wejść do $m - 1$. elementu modelu. Umożliwia to wyznaczenie macierzy wartości składowych wektora wyjść modelu $m - 1$. elementu $\bar{V}_N^{*(m-1)}$ z optymalnym wektorem parametrów $\tilde{\theta}_{m-1N}^*$, tj.

$$\begin{aligned}\bar{V}_N^{*(m-1)} &\stackrel{\text{df}}{=} \tilde{\Phi}_{m-1}(\bar{V}_N^{*(m-2)}, \tilde{\theta}_{m-1N}^*) \\ &= \tilde{\Phi}_{m-1}(\bar{V}_N^{*(m-2)}, \tilde{\Psi}_{m-1N}(V_N^{(M)}, \dots, V_N^{(m-1)}, \bar{V}_N^{*(m-2)})).\end{aligned}\quad (4.106)$$

Wartość optymalnego wektora parametrów modelu m -tego elementu $\tilde{\theta}_{mN}^*$ wyznaczamy według zależności

$$\begin{aligned}\tilde{\theta}_{mN}^* &= \tilde{\Psi}_{mN}(V_N^{(M)}, \dots, V_N^{(m)}, \bar{V}_N^{*(m-1)}). \\ &\vdots\end{aligned}\quad (4.107)$$

Krok $2M$. W kroku $2M - 1$. wyznaczyliśmy macierz wartości składowych wektora wyjść $\bar{V}_N^{*(M-2)}$ modelu $M - 2$. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , która jest jednocześnie macierzą wartości składowych wektora wejść do $M - 1$. elementu modelu. Umożliwia to wyznaczenie macierzy wartości składowych wektora wyjść modelu $M - 1$. elementu $\bar{V}_N^{*(M-1)}$ z optymalnym wektorem parametrów $\tilde{\theta}_{M-1N}^*$, tj.

$$\begin{aligned}\bar{V}_N^{*(M-1)} &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-1}(\bar{V}_N^{*(M-2)}, \tilde{\theta}_{M-1N}^*) \\ &= \tilde{\Phi}_{M-1}(\bar{V}_N^{*(M-2)}, \tilde{\Psi}_{M-1N}(V_N^{(M)}, \dots, V_N^{(M-1)}, \bar{V}_N^{*(M-2)})).\end{aligned}\quad (4.108)$$

Ostatecznie, po podstawieniu do algorytmu (4.65) macierzy $\bar{V}_N^{*(M-1)}$, określonej przez (4.108), otrzymujemy

$$\tilde{\theta}_{MN}^* = \tilde{\Psi}_{MN}(V_N^{(M)}, \bar{V}_N^{*(M-1)}),\quad (4.109)$$

wartość optymalnego wektora $\tilde{\theta}_{MN}^*$ parametrów modelu M -tego elementu.

Przykład 4.1. Rozważmy wejściowo-wyjściowy system złożony o strukturze szeregowej, w którym wyróżniono trzy jednowymiarowe elementy ($M = 3$, $S_m = L_m = 1$, $m = 1, 2, 3$). Dla poszczególnych elementów o wejściu u_m i wyjściu y_m przyjmujemy modele liniowe, czyli model (4.1) ma postać

$$\bar{y}_m = \Phi_m(u_m, \theta_m) = \theta_m u_m, \quad m = 1, 2, 3,\quad (4.110)$$

gdzie: \bar{y}_m jest wyjściem, a θ_m – nieznanym parametrem modelu.

Wejściem zewnętrznym x systemu złożonego jest wejście pierwszego elementu, a wyróżnionymi wyjściami wszystkie wyjścia (rys. 4.5 dla $M = 3$), czyli

$$x = u_1, v^{(1)} = y_1, v^{(2)} = y_2, v^{(3)} = y_3.\quad (4.111)$$

W rozpatrywanym przykładzie model systemu złożonego (4.53) przyjmuje postać

$$\bar{v} = \begin{bmatrix} \bar{v}^{(1)} \\ \bar{v}^{(2)} \\ \bar{v}^{(3)} \end{bmatrix} = \begin{bmatrix} \Phi_1(x, \theta_1) \\ \Phi_2(\bar{v}^{(1)}, \theta_2) \\ \Phi_3(\bar{v}^{(2)}, \theta_3) \end{bmatrix} = \begin{bmatrix} \theta_1 x \\ \theta_2 \bar{v}^{(1)} \\ \theta_3 \bar{v}^{(2)} \end{bmatrix}, \quad (4.112)$$

a w rekurencyjnej postaci (4.55) zapisujemy

$$\bar{v}^{(m+1)} = \Phi_m(\bar{v}^{(m)}, \theta_m) = \theta_m \bar{v}^{(m)}, \quad m = 0, 1, 2, \quad (4.113)$$

gdzie

$$\bar{v}^{(0)} \stackrel{\text{df}}{=} x. \quad (4.114)$$

Przyjmujemy kwadratowy wskaźnik jakości identyfikacji, czyli funkcja q_m w kryterium (4.62) ma postać

$$q_m(v_n^{(m)}, \bar{v}_n^{(m)}) = (v_n^{(m)} - \bar{v}_n^{(m)})^2. \quad (4.115)$$

Przyjmujemy ponadto wartości współczynników $\lambda_m = 1$, $m = 1, 2, 3$, co znaczy, że w globalnym wskaźniku jakości z taką samą wagą uwzględniamy wszystkie wyróżnione wyjścia systemu złożonego. Ostatecznie, po uwzględnieniu (4.113), (4.114) oraz (4.115), globalny wskaźnik jakości identyfikacji (4.62) możemy zapisać

$$\tilde{Q}_N(\theta) = \sum_{m=1}^3 \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)) = \sum_{m=1}^3 \sum_{n=1}^N (v_n^{(m)} - \theta_m \bar{v}_n^{(m-1)})^2. \quad (4.116)$$

Zadanie identyfikacji sprowadza się do wyznaczenia optymalnych wartości parametrów modeli elementów $\tilde{\theta}_{1N}^*$, $\tilde{\theta}_{2N}^*$, $\tilde{\theta}_{3N}^*$, dla których wskaźnik (4.116) przyjmuje wartość minimalną, czyli

$$\tilde{\theta}_{1N}^*, \tilde{\theta}_{2N}^*, \tilde{\theta}_{3N}^* : \sum_{m=1}^3 \sum_{n=1}^N (v_n^{(m)} - \tilde{\theta}_{mn}^* \bar{v}_n^{(m-1)})^2 = \min_{\theta_1, \theta_2, \theta_3} \sum_{m=1}^3 \sum_{n=1}^N (v_n^{(m)} - \tilde{\theta}_{mn}^* \bar{v}_n^{(m-1)})^2. \quad (4.117)$$

Zastosowanie metody programowania dynamicznego dla rozpatrywanego przykładu daje następujący algorytm identyfikacji:

Krok 1. Oznaczamy ostatni – 3. składnik sumy w kryterium (4.116) przez

$$\tilde{Q}_{3N}(\theta_3) \stackrel{\text{df}}{=} \sum_{n=1}^N (v_n^{(3)} - \theta_3 \bar{v}_n^{(2)})^2. \quad (4.118)$$

Minimalizacja wyrażenia (4.118) względem θ_3 daje algorytm identyfikacji $\tilde{\Psi}_{3N}$

$$\tilde{\theta}_{3N}^* = \tilde{\Psi}_{3N}(V_N^{(3)}, \bar{V}_N^{(2)}) = \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(2)}}{\sum_{n=1}^N (\bar{v}_n^{(2)})^2}, \quad (4.119)$$

gdzie: $V_N^{(3)}$ jest macierzą pomiarów wartości wyjść 3. elementu, tj.

$$V_N^{(3)} = [v_1^{(3)} \ v_2^{(3)} \ \dots \ v_N^{(3)}], \quad (4.120)$$

natomiast $\bar{V}_N^{(2)}$ jest macierzą wartości wejść modelu 3. elementu, która jest również macierzą wartości wyjść modelu 2. elementu, czyli

$$\bar{V}_N^{(2)} = [\bar{v}_1^{(2)} \ \bar{v}_2^{(2)} \ \dots \ \bar{v}_N^{(2)}]. \quad (4.121)$$

Po podstawieniu algorytmu (4.119) w wyrażeniu (4.118) w miejsce θ_3 otrzymujemy

$$\tilde{Q}_{3N}(\tilde{\theta}_{3N}^*) = \tilde{Q}_{3N}(\Psi_{3N}(V_N^{(3)}, \bar{V}_N^{(3)})) \stackrel{\text{df}}{=} \tilde{Q}_{3N}^*(V_N^{(3)}, \bar{V}_N^{(3)}) = \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(2)}}{\sum_{n=1}^N (\bar{v}_n^{(2)})^2} \bar{v}_n^{(2)} \right)^2. \quad (4.122)$$

Po uwzględnieniu zależności (4.55), tj.

$$\bar{v}_n^{(2)} = \Phi_2(\bar{v}_n^{(1)}, \theta_2) = \theta_2 \bar{v}_n^{(1)}, \quad (4.123)$$

w macierzy (4.121) mamy

$$\bar{V}_N^{(2)} = [\theta_2 \bar{v}_1^{(1)} \ \theta_2 \bar{v}_2^{(1)} \ \dots \ \theta_2 \bar{v}_N^{(1)}] \stackrel{\text{df}}{=} \tilde{\Phi}_2(\bar{V}_N^{(1)}, \theta_2), \quad (4.124)$$

a po podstawieniu (4.124) do (4.122), po prostych przekształceniach, otrzymujemy

$$\tilde{Q}_{3N}^*(V_N^{(3)}, \bar{V}_N^{(3)}) = \tilde{Q}_{3N}^*(V_N^{(3)}, \tilde{\Phi}_2(\bar{V}_N^{(1)}, \theta_2)) = \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(1)}}{\sum_{n=1}^N (\bar{v}_n^{(1)})^2} \bar{v}_n^{(1)} \right)^2. \quad (4.125)$$

Krok 2. Oznaczamy sumę przedostatniego – 2. składnika w kryterium (4.116) i \tilde{Q}_{3N}^* – optymalnej wartości (4.125) składowej kryterium, wyznaczonej w kroku 1. – przez

$$\tilde{Q}_{2N}(\theta_2) \stackrel{\text{df}}{=} \sum_{n=1}^N (v_n^{(2)} - \theta_2 \bar{v}_n^{(1)})^2 + \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(1)}}{\sum_{n=1}^N (\bar{v}_n^{(1)})^2} \bar{v}_n^{(1)} \right)^2. \quad (4.126)$$

Zwróćmy uwagę, że 2. składnik wyrażenia (4.126) nie zależy od θ_2 . Minimalizacja (4.126) względem θ_2 daje algorytm identyfikacji $\tilde{\Psi}_{2N}$,

$$\tilde{\theta}_{2N}^* = \tilde{\Psi}_{2N}(V_N^{(3)}, V_N^{(2)}, \bar{V}_N^{(1)}) = \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{(1)}}{\sum_{n=1}^N (\bar{v}_n^{(1)})^2}, \quad (4.127)$$

gdzie: $V_N^{(2)}$ jest macierzą pomiarów wartości wyjść 2. elementu, tj.

$$V_N^{(2)} = [v_1^{(2)} \ v_2^{(2)} \ \dots \ v_N^{(2)}], \quad (4.128)$$

natomiast $\bar{V}_N^{(1)}$ jest macierzą wartości wejść modelu 2. elementu, która jest jednocześnie macierzą wartości wyjść modelu 1. elementu, czyli

$$\bar{V}_N^{(1)} = [\bar{v}_1^{(1)} \ \bar{v}_2^{(1)} \ \dots \ \bar{v}_N^{(1)}]. \quad (4.129)$$

Algorytm (4.127) nie zależy od macierzy pomiarów $V_N^{(3)}$. Jest to uzasadnione, gdyż 2. składnik wyrażenia (4.126) nie zależy od θ_2 .

Po podstawieniu algorytmu (4.127) w wyrażeniu (4.126) w miejsce θ_2 otrzymujemy

$$\begin{aligned} \tilde{Q}_{2N}(\tilde{\theta}_{2N}^*) &= \tilde{Q}_{2N}(\Psi_{2N}(V_N^{(3)}, V_N^{(2)}, \bar{V}_N^{(1)})) \stackrel{\text{df}}{=} \tilde{Q}_{2N}^*(V_N^{(3)}, V_N^{(2)}, \bar{V}_N^{(1)}) \\ &= \sum_{n=1}^N \left(v_n^{(2)} - \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{(1)}}{\sum_{n=1}^N (\bar{v}_n^{(1)})^2} \bar{v}_n^{(1)} \right)^2 + \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(1)}}{\sum_{n=1}^N (\bar{v}_n^{(1)})^2} \bar{v}_n^{(1)} \right)^2. \end{aligned} \quad (4.130)$$

Po uwzględnieniu zależności (4.55), tj.

$$\bar{v}_n^{(1)} = \Phi_1(\bar{v}_n^{(0)}, \theta_1) = \theta_1 \bar{v}_n^{(0)}, \quad (4.131)$$

w macierzy (4.129) mamy

$$\bar{V}_N^{(1)} = [\theta_1 \bar{v}_1^{(0)} \ \theta_1 \bar{v}_2^{(0)} \ \dots \ \theta_1 \bar{v}_N^{(0)}] \stackrel{\text{df}}{=} \tilde{\Phi}_1(\bar{V}_N^{(0)}, \theta_1), \quad (4.132)$$

a po podstawieniu (4.132) do (4.130), po prostych przekształceniach, otrzymujemy

$$\begin{aligned} \tilde{Q}_{2N}^*(V_N^{(3)}, V_N^{(2)}, \bar{V}_N^{(2)}) &= \tilde{Q}_{2N}^*(V_N^{(3)}, V_N^{(2)}, \tilde{\Phi}_1(\bar{V}_N^{(0)}, \theta_1)) \\ &= \sum_{n=1}^N \left(v_n^{(2)} - \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{(0)}}{\sum_{n=1}^N (\bar{v}_n^{(0)})^2} \bar{v}_n^{(0)} \right)^2 + \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(0)}}{\sum_{n=1}^N (\bar{v}_n^{(0)})^2} \bar{v}_n^{(0)} \right)^2. \end{aligned} \quad (4.133)$$

Krok 3. Oznaczamy sumę 1. składnika w kryterium (4.166) i \tilde{Q}_{2N}^* – optymalnej wartości składowej kryterium wyznaczonej w kroku 2. – przez

$$\begin{aligned} \tilde{Q}_{1N}(\theta_1) \stackrel{\text{df}}{=} & \sum_{n=1}^N (v_n^{(1)} - \theta_1 \bar{v}_n^{(0)})^2 + \sum_{n=1}^N \left(v_n^{(2)} - \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{(0)}}{\sum_{n=1}^N (\bar{v}_n^{(0)})^2} \bar{v}_n^{(0)} \right)^2 \\ & + \sum_{n=1}^N \left(v_n^{(3)} - \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{(0)}}{\sum_{n=1}^N (\bar{v}_n^{(0)})^2} \bar{v}_n^{(0)} \right)^2. \end{aligned} \quad (4.134)$$

Zwróćmy uwagę, że 2. i 3. składnik wyrażenia (4.134) nie zależą od θ_1 . Minimalizacja (4.134) względem θ_1 daje algorytm identyfikacji $\tilde{\Psi}_{1N}$,

$$\tilde{\theta}_{1N}^* = \tilde{\Psi}_{1N}(V_N^{(3)}, V_N^{(2)}, V_N^{(1)}, \bar{V}_N^{(0)}) = \frac{\sum_{n=1}^N v_n^{(1)} \bar{v}_n^{(0)}}{\sum_{n=1}^N (\bar{v}_n^{(0)})^2}, \quad (4.135)$$

gdzie: $V_N^{(l)}$ jest macierzą pomiarów wartości wyjść 1. elementu, tj.

$$V_N^{(1)} = [v_1^{(1)} \ v_2^{(1)} \ \dots \ v_N^{(1)}], \quad (4.136)$$

natomiast $\bar{V}_N^{(0)}$, zgodnie z (4.114), jest macierzą wartości wejść zewnętrznych, czyli

$$\bar{V}_N^{(0)} = [\bar{v}_1^{(0)} \ \bar{v}_2^{(0)} \ \dots \ \bar{v}_N^{(0)}] = [x_1 \ x_2 \ \dots \ x_N] \stackrel{\text{df}}{=} X_N. \quad (4.137)$$

Algorytm (4.135) nie zależy od macierzy pomiarów $V_N^{(3)}$ oraz $V_N^{(2)}$. Jest to uzasadnione, gdyż składniki 2. i 3. wyrażenia (4.134) nie zależą od θ_1 .

Dopiero teraz występujące w algorytmie (4.135) wszystkie macierze, z uwzględnieniem (4.137), są znane jako wyniki pomiarów. Optymalne wartości parametrów $\tilde{\theta}_{3N}^*$ oraz $\tilde{\theta}_{2N}^*$, wyznaczone w krokach 1. i 2., oprócz macierzy pomiarów wyjść $V_N^{(3)}, V_N^{(2)}, V_N^{(1)}$, zależą od macierzy wartości wyjść modeli $\bar{V}_N^{(2)}, \bar{V}_N^{(1)}$. Jedyne wartości optymalnego parametru $\tilde{\theta}_{1N}^*$ zależą od serii identyfikującej X_N oraz macierzy pomiarów wszystkich wyjść $V_N^{(3)}, V_N^{(2)}, V_N^{(1)}$.

Powracając do algorytmów identyfikacji wyznaczonych w kolejnych krokach, możemy wyznaczyć optymalne wartości parametrów poszczególnych elementów systemu. I tak:

Krok 4. Po uwzględnieniu w (4.135) serii identyfikującej (4.137) otrzymujemy algorytm

$$\tilde{\theta}_{1N}^* = \tilde{\Psi}_{1N} \left(V_N^{(3)}, V_N^{(2)}, V_N^{(1)}, X_N \right) = \frac{\sum_{n=1}^N v_n^{(1)} x_n}{\sum_{n=1}^N (x_n)^2}, \quad (4.138)$$

który pozwala wyznaczyć optymalną wartość $\tilde{\theta}_{1N}^*$ – parametru modelu 1. elementu.

Krok 5. Po podstawieniu w macierzy (4.132) w miejsce θ_1 optymalnej wartości $\tilde{\theta}_{1N}^*$, określonej przez (4.138), otrzymujemy macierz wartości wyjść modelu 1. elementu systemu złożonego dla zadanej serii identyfikującej X_N , tj.

$$\bar{V}_N^{*(1)} \stackrel{\text{df}}{=} \tilde{\Phi}_1 \left(X_N, \tilde{\theta}_{1N}^* \right) = \begin{bmatrix} \bar{v}_1^{*(1)} & \bar{v}_2^{*(1)} & \dots & \bar{v}_N^{*(1)} \end{bmatrix} = \begin{bmatrix} \tilde{\theta}_{1N}^* x_1 & \tilde{\theta}_{1N}^* x_2 & \dots & \tilde{\theta}_{1N}^* x_N \end{bmatrix}. \quad (4.139)$$

Po podstawieniu w algorytmie (4.127) w miejsce $\bar{V}_N^{(1)}$ macierzy wartości wyjść optymalnego modelu (4.139) otrzymujemy algorytm

$$\tilde{\theta}_{2N}^* = \tilde{\Psi}_{2N} \left(V_N^{(3)}, V_N^{(2)}, \bar{V}_N^{*(1)} \right) = \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{*(1)}}{\sum_{n=1}^N (\bar{v}_n^{*(1)})^2} = \frac{\sum_{n=1}^N v_n^{(2)} \tilde{\theta}_{1N}^* x_n}{\sum_{n=1}^N (\tilde{\theta}_{1N}^* x_n)^2} = \frac{\sum_{n=1}^N v_n^{(2)} x_n}{\tilde{\theta}_{1N}^* \sum_{n=1}^N (x_n)^2}, \quad (4.140)$$

a biorąc pod uwagę algorytm (4.138), po prostych przekształceniach otrzymujemy algorytm

$$\tilde{\theta}_{2N}^* = \frac{\sum_{n=1}^N v_n^{(2)} x_n}{\frac{\sum_{n=1}^N v_n^{(1)} x_n}{\sum_{n=1}^N (x_n)^2}} = \frac{\sum_{n=1}^N v_n^{(2)} x_n}{\sum_{n=1}^N v_n^{(1)} x_n}, \quad (4.141)$$

umożliwiający wyznaczenie optymalnej wartości parametru modelu 2. elementu systemu złożonego.

Krok 6. Podobnie jak w kroku poprzednim, po podstawieniu w macierzy (4.124) w miejsce θ_2 optymalnej wartości $\tilde{\theta}_{2N}^*$, określonej przez (4.141), otrzymujemy macierz wartości wyjść optymalnego modelu 2. elementu systemu złożonego, tj.

$$\bar{V}_N^{*(2)} \stackrel{\text{df}}{=} \tilde{\Phi}_2(X_N, \tilde{\theta}_{2N}^*) = [\bar{v}_1^{*(2)} \ \bar{v}_2^{*(2)} \ \dots \ \bar{v}_N^{*(2)}] = [\tilde{\theta}_{2N}^* \bar{v}_1^{*(1)} \ \tilde{\theta}_{2N}^* \bar{v}_2^{*(1)} \ \dots \ \tilde{\theta}_{2N}^* \bar{v}_N^{*(1)}]. \quad (4.142)$$

Po podstawieniu w algorytmie (4.119) w miejsce $\bar{V}_N^{(2)}$ macierzy (4.142) otrzymujemy

$$\tilde{\theta}_{3N}^* = \tilde{\Psi}_{3N}(V_N^{(3)}, \bar{V}_N^{*(2)}) = \frac{\sum_{n=1}^N v_n^{(3)} \bar{v}_n^{*(2)}}{\sum_{n=1}^N (\bar{v}_n^{*(2)})^2} = \frac{\sum_{n=1}^N v_n^{(3)} \tilde{\theta}_{2N}^* \bar{v}_n^{*(1)}}{\sum_{n=1}^N (\tilde{\theta}_{2N}^* \bar{v}_n^{*(1)})^2} = \frac{\sum_{n=1}^N v_n^{(2)} \bar{v}_n^{*(1)}}{\tilde{\theta}_{2N}^* \sum_{n=1}^N (\bar{v}_n^{*(1)})^2}, \quad (4.143)$$

a biorąc pod uwagę (4.139) oraz algorytm (4.141), po prostych przekształceniach otrzymujemy algorytm

$$\tilde{\theta}_{3N}^* = \frac{\sum_{n=1}^N v_n^{(3)} x_n}{\tilde{\theta}_{2N}^* \tilde{\theta}_{1N}^* \sum_{n=1}^N (x_n)^2} = \frac{\sum_{n=1}^N v_n^{(3)} x_n}{\frac{\sum_{n=1}^N v_n^{(2)} x_n}{\sum_{n=1}^N v_n^{(1)} x_n} \sum_{n=1}^N (x_n)^2} = \frac{\sum_{n=1}^N v_n^{(3)} x_n}{\sum_{n=1}^N v_n^{(2)} x_n}, \quad (4.144)$$

umożliwiający wyznaczenie optymalnej wartości parametru 3. elementu systemu złożonego. ◆

4.4.2. Globalnie optymalny model z uwzględnieniem jakości modelu lokalnego dla systemu o strukturze szeregowej

Analogiczne algorytmy identyfikacji systemu złożonego o strukturze szeregowej (rys. 4.7), oparte na metodzie programowania dynamicznego, można zaproponować dla zadań wyboru globalnie optymalnego modelu z uwzględnieniem jakości modeli lokalnych (podrozdział 4.3).

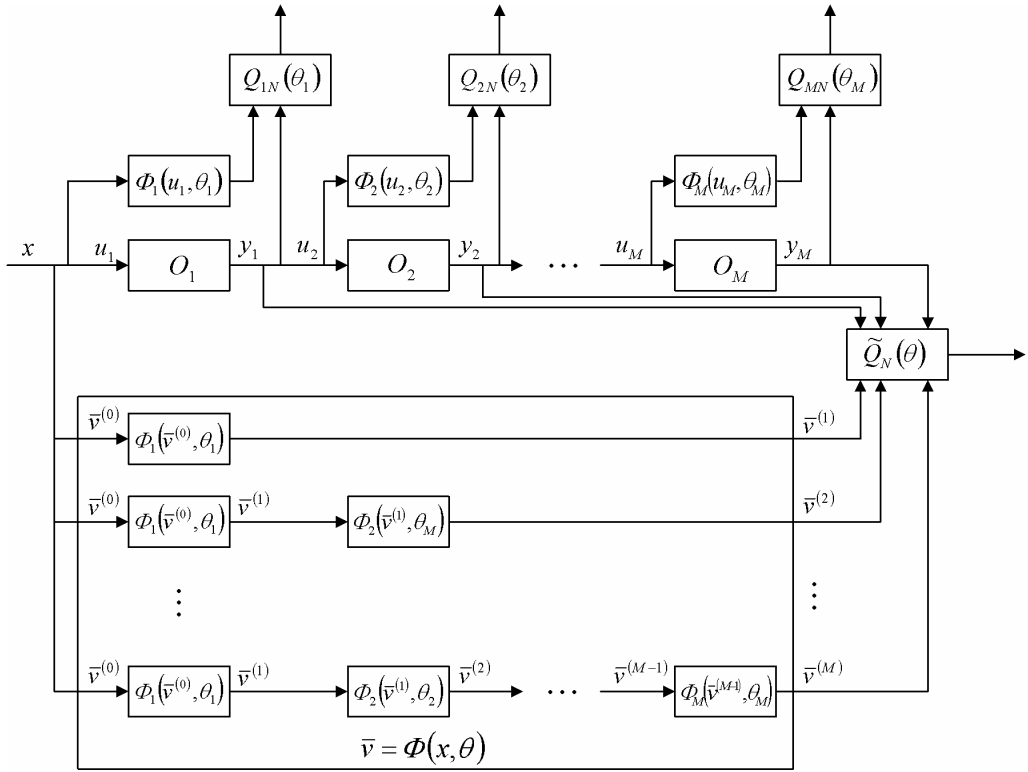
W przypadku systemu o strukturze szeregowej zależności pomiędzy wejściami i wyjściami systemu złożonego (przy założeniu (4.52)) są następujące:

$$x = u_1 = v^{(0)}, \quad u_2 = y_1 = v^{(1)}, \quad \dots, \quad u_M = y_{M-1} = v^{(M-1)}.$$

W wyniku eksperymentu, dla zadanej serii identyfikującej o długości N , dokonano niezakłóconego pomiaru wartości wszystkich wyjść systemu złożonego. Wyniki zebrano w macierzach

$$V_N^{(0)} = X_N = [x_1 \ x_2 \ \dots \ x_N], \quad V_N^{(m)} = [v_1^{(m)} \ v_2^{(m)} \ \dots \ v_N^{(m)}], \quad m = 1, 2, \dots, M, \quad (4.145)$$

gdzie: X_N jest serią identyfikującą o długości N , a $V_N^{(m)}$, $m = 1, 2, \dots, M$, są macierzami pomiarów wartości składowych wektorów wyjść poszczególnych elementów systemu złożonego.



Rys. 4.7. Globalnie optymalny model systemu złożonego o strukturze szeregowej z uwzględnieniem jakości modelu lokalnego

Teraz lokalny wskaźnik jakości identyfikacji (4.19) dla m -tego elementu ($m = 1, 2, \dots, M$) ma postać

$$Q_{mN}(\theta_m) = \left\| V_N^{(m)} - \bar{V}_N^{(m)}(\theta_m) \right\|_{V_N^{(m-1)}}, \quad (4.146)$$

gdzie

$$\bar{V}_N^{(m)}(\theta_m) \stackrel{\text{df}}{=} \left[\Phi_m(v_1^{(m-1)}, \theta_m) \Phi_m(v_2^{(m-1)}, \theta_m) \dots \Phi_m(v_N^{(m-1)}, \theta_m) \right], \quad (4.147)$$

a po przyjęciu lokalnego kryterium jakości identyfikacji postaci (4.21) otrzymujemy

$$Q_{mN}(\theta_m) = \sum_{n=1}^M q_m(v_n^{(m)}, \bar{v}_n^{(m)}) = \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(v_n^{(m-1)}, \theta_m)). \quad (4.148)$$

Syntetyczny wskaźnik jakości identyfikacji

Dla globalnego wskaźnika (4.62) oraz lokalnych wskaźników jakości identyfikacji postaci (4.148) syntetyczny wskaźnik (4.35) jest określony zależnością

$$\begin{aligned} \bar{Q}_N(\theta) = & \alpha_0 \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)) \\ & + \sum_{m=1}^M \alpha_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(v_n^{(m-1)}, \theta_m)) \end{aligned} \quad (4.149)$$

Syntetyczny wskaźnik postaci (4.149) oraz rekurencyjny model (4.55) ponownie pozwalają zastosować metodę programowania dynamicznego do rozwiązania zadania wyznaczenia globalnie optymalnego wektora parametrów modelu systemu złożonego z uwzględnieniem jakości modeli lokalnych (4.37). Po wyodrębnieniu wektorów parametrów modeli poszczególnych elementów zadanie (4.37) z kryterium (4.149) sprowadza się do wyznaczenia takiego zestawu wektorów parametrów modeli elementów $\bar{\theta}_{1N}^*, \bar{\theta}_{2N}^*, \dots, \bar{\theta}_{MN}^*$, dla których wskaźnik (4.149) przyjmuje wartość minimalną, czyli

$$\begin{aligned} \bar{\theta}_{1N}^*, \bar{\theta}_{2N}^*, \dots, \bar{\theta}_{MN}^* : & \alpha_0 \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \bar{\theta}_{mN}^*)) \\ & + \sum_{m=1}^M \alpha_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(v_n^{(m-1)}, \bar{\theta}_{mN}^*)) \\ = & \min_{\theta_1, \theta_2, \dots, \theta_M \in \Theta_1 \times \Theta_2 \times \dots \times \Theta_M} \left\{ \alpha_0 \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)) \right. \\ & \left. + \sum_{m=1}^M \alpha_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(v_n^{(m-1)}, \theta_m)) \right\}. \end{aligned} \quad (4.150)$$

Zastosowanie metody programowania dynamicznego daje następujący algorytm identyfikacji przy syntetycznym wskaźniku jakości identyfikacji (4.149) dla systemu o strukturze szeregowej:

Krok 1. Oznaczamy ostatni M -ty składnik sumy w kryterium (4.149) przez

$$\begin{aligned} \bar{Q}_{MN}(\theta_M) \stackrel{\text{df}}{=} & \alpha_0 \lambda_M \sum_{n=1}^N q_M(v_n^{(M)}, \Phi_M(\bar{v}_n^{(M-1)}, \theta_M)) \\ & + \alpha_M \sum_{n=1}^N q_M(v_n^{(M)}, \Phi_M(v_n^{(M-1)}, \theta_M)). \end{aligned} \quad (4.151)$$

Określamy optymalny wektor parametrów $\bar{\theta}_{MN}^*$, który minimalizuje wskaźnik (4.151). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\bar{\Psi}_{MN}$

$$\bar{\theta}_{MN}^* = \bar{\Psi}_{MN}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-1)}), \quad (4.152)$$

który minimalizuje wskaźnik (4.151) względem θ_M , czyli

$$\bar{\theta}_{MN}^* = \bar{\Psi}_{MN}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-1)}) : \bar{Q}_{MN}(\bar{\theta}_{MN}^*) = \min_{\theta_M \in \Theta_M} \bar{Q}_{MN}(\theta_M), \quad (4.153)$$

gdzie: $V_N^{(M-1)}, V_N^{(M)}$ są macierzami pomiarów (4.145) dla $m = M - 1, M$, natomiast $\bar{V}_N^{(M-1)}$ jest macierzą wartości składowych wektora wejść modelu M -tego elementu, która jest również macierzą wartości składowych wektora wyjść modelu $M - 1$. elementu, czyli

$$\bar{V}_N^{(M-1)} = [\bar{v}_1^{(M-1)} \bar{v}_2^{(M-1)} \dots \bar{v}_N^{(M-1)}]. \quad (4.154)$$

W algorytmie (4.152) występuje macierz (4.154), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść modelu poprzedniego $M - 2$. elementu.

Po podstawieniu do (4.151) algorytmu wyznaczania optymalnego wektora parametrów $\bar{\theta}_{MN}^*$ modelu M -tego elementu, uzyskanego w wyniku rozwiązania zadania (4.153), otrzymujemy optymalną postać kryterium zależną od macierzy $\bar{V}_N^{(M-1)}$ (4.154)

$$\bar{Q}_{MN}(\bar{\theta}_{MN}^*) = \bar{Q}_{MN}(\bar{\Psi}_M(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-1)})) \stackrel{\text{df}}{=} \bar{Q}_{MN}^*(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-1)}). \quad (4.155)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.154) mamy

$$\begin{aligned} \bar{V}_N^{(M-1)} &= [\Phi_{M-1}(\bar{v}_1^{(M-2)}, \theta_{M-1}) \Phi_{M-1}(\bar{v}_2^{(M-2)}, \theta_{M-1}) \dots \Phi_{M-1}(\bar{v}_N^{(M-2)}, \theta_{M-1})] \\ &\stackrel{\text{df}}{=} \bar{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1}), \end{aligned} \quad (4.156)$$

a po podstawieniu (4.156) do (4.155) otrzymujemy

$$\bar{Q}_{MN}^*(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{(M-1)}) = \bar{Q}_{MN}^*(V_N^{(M)}, V_N^{(M-1)}, \bar{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1})). \quad (4.157)$$

Krok 2. Oznaczamy sumę przedostatniego składnika w kryterium (4.149) i \bar{Q}_{MN}^* – optymalnej postaci (4.157) składowej kryterium, wyznaczonej w kroku 1. – przez:

$$\begin{aligned} \bar{Q}_{M-1N}(\theta_{M-1}) &\stackrel{\text{df}}{=} \alpha_0 \lambda_{M-1} \sum_{n=1}^N q_{M-1}(v_n^{(M-1)}, \Phi_{M-1}(v_n^{(M-2)}, \theta_{M-1})) \\ &\quad + \alpha_{M-1} \sum_{n=1}^N q_{M-1}(v_n^{(M-1)}, \Phi_{M-1}(v_n^{(M-2)}, \theta_{M-1})) \\ &\quad + \bar{Q}_{MN}^*(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}) \tilde{\Phi}_{M-1}(\bar{V}_N^{(M-2)}, \theta_{M-1}). \end{aligned} \quad (4.158)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{M-1N}^*$, który minimalizuje wskaźnik (4.158). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\bar{\Psi}_{M-1N}$

$$\bar{\theta}_{M-1N}^* = \bar{\Psi}_{M-1N} \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{V}_N^{(M-2)} \right), \quad (4.159)$$

który minimalizuje wskaźnik (4.158) względem θ_{M-1} , czyli

$$\begin{aligned} \bar{\theta}_{M-1N}^* &= \bar{\Psi}_{M-1N} \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{V}_N^{(M-2)} \right) : \bar{Q}_{M-1N} \left(\bar{\theta}_{M-1N}^* \right) \\ &= \min_{\theta_{M-1} \in \Theta_{M-1}} \bar{Q}_{M-1N} \left(\theta_{M-1} \right), \end{aligned} \quad (4.160)$$

gdzie: $V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}$ są macierzami pomiarów (4.145) dla $m = M-2, M-1, M$, natomiast $\bar{V}_N^{(M-2)}$ jest macierzą wartości składowych wektora wejść modelu $M-1$. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu $M-2$. elementu, czyli

$$\bar{V}_N^{(M-2)} = \left[\bar{v}_1^{(M-2)} \bar{v}_2^{(M-2)} \dots \bar{v}_N^{(M-2)} \right]. \quad (4.161)$$

W algorytmie (4.159) występuje macierz (4.161), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść modelu poprzedniego $M-3$. elementu.

Po podstawieniu do (4.158) algorytmu wyznaczania optymalnego wektora parametrów $\bar{\theta}_{M-1N}^*$ modelu $M-1$. elementu, uzyskanego w wyniku rozwiązania zadania (4.160), otrzymujemy optymalną postać kryterium zależną od macierzy $\bar{V}_N^{(M-2)}$ (4.161)

$$\begin{aligned} \bar{Q}_{M-1N} \left(\bar{\theta}_{M-1N}^* \right) &= \bar{Q}_{M-1N} \left(\bar{\Psi}_{M-1N} \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{V}_N^{(M-2)} \right) \right) \\ &\stackrel{\text{df}}{=} \bar{Q}_{M-1N}^* \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{V}_N^{(M-2)} \right). \end{aligned} \quad (4.162)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.161) mamy

$$\begin{aligned} \bar{V}_N^{(M-2)} &= \left[\Phi_{M-2} \left(\bar{v}_1^{(M-3)}, \theta_{M-2} \right) \Phi_{M-2} \left(\bar{v}_2^{(M-3)}, \theta_{M-2} \right) \dots \Phi_{M-2} \left(\bar{v}_N^{(M-3)}, \theta_{M-2} \right) \right] \\ &\stackrel{\text{df}}{=} \bar{\Phi}_{M-2} \left(\bar{V}_N^{(M-3)}, \theta_{M-2} \right), \end{aligned} \quad (4.163)$$

a po podstawieniu (4.161) do (4.162) otrzymujemy

$$\begin{aligned} &\bar{Q}_{M-1N}^* \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{V}_N^{(M-2)} \right) \\ &= \bar{Q}_{M-1N}^* \left(V_N^{(M)}, V_N^{(M-1)}, V_N^{(M-2)}, \bar{\Phi}_{M-2} \left(\bar{V}_N^{(M-3)}, \theta_{M-2} \right) \right). \end{aligned} \quad (4.164)$$

⋮

Krok m . Oznaczamy sumę $M-m+1$. składnika w kryterium (4.149) i \bar{Q}_{M-m+2N}^* – optymalnej postaci składowej kryterium wyznaczonej w $m-1$. kroku – przez

$$\begin{aligned}
\bar{Q}_{M-m+1N}(\theta_{M-m+1}) &\stackrel{\text{df}}{=} \alpha_0 \lambda_{M-m+1} \sum_{n=1}^N q_{M-m+1} \left(v_n^{(M-m+1)}, \Phi_{M-m+1} \left(\bar{v}^{(M-m)}, \theta_{M-m+1} \right) \right) \\
&+ \alpha_{M-m-1} \sum_{n=1}^N q_{M-m-1} \left(v_n^{(M-m-1)}, \Phi_{M-m-1} \left(v_n^{(M-m)}, \theta_{M-m-1} \right) \right) \\
&+ \bar{Q}_{M-m+2N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{\Phi}_{M-m+1} \left(\bar{V}_N^{(M-m)}, \theta_{M-m+1} \right) \right).
\end{aligned} \quad (4.165)$$

Określamy optymalny wektor parametrów $\bar{\theta}_{M-m+1N}^*$, który minimalizuje wskaźnik (4.165). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\bar{\Psi}_{M-m+1N}$

$$\bar{\theta}_{M-m+1N}^* = \bar{\Psi}_{M-m+1N} \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{V}_N^{(M-m)} \right), \quad (4.166)$$

który minimalizuje wskaźnik (4.165) względem θ_{M-m+1} , czyli

$$\begin{aligned}
\bar{\theta}_{M-m+1N}^* &= \bar{\Psi}_{M-m+1N} \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{V}_N^{(M-m)} \right) : \bar{Q}_{M-m+1N} \left(\bar{\theta}_{M-m+1N}^* \right) \\
&= \min_{\theta_{M-m+1} \in \Theta_{M-m+1}} \bar{Q}_{M-m+1N} \left(\theta_{M-m+1} \right),
\end{aligned} \quad (4.167)$$

gdzie: $V_N^{(M)}, \dots, V_N^{(M-m)}$ są macierzami pomiarów (4.145) dla $m = M-m, M-m+1, \dots, M$, a $\bar{V}_N^{(M-m)}$ jest macierzą wartości składowych wektora wejść modelu $M-m+1$. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu $M-m$ -tego elementu, czyli

$$\bar{V}_N^{(M-m)} \stackrel{\text{df}}{=} \left[\bar{v}_1^{(M-m)} \bar{v}_2^{(M-m)} \dots \bar{v}_N^{(M-m)} \right]. \quad (4.168)$$

W algorytmie (4.166) występuje macierz (4.168), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść modelu poprzedniego $M-m-1$. elementu.

Po podstawieniu do (4.165) algorytmu wyznaczania optymalnego wektora parametrów modelu $M-m+1$. elementu $\bar{\theta}_{M-m+1N}^*$, uzyskanego w wyniku rozwiązania zadania (4.167), otrzymujemy optymalną postać kryterium

$$\begin{aligned}
\bar{Q}_{M-m+1N} \left(\bar{\theta}_{M-m+1N}^* \right) &= \bar{Q}_{M-m+1N} \left(\bar{\Psi}_{M-m+1N} \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{V}_N^{(M-m)} \right) \right) \\
&\stackrel{\text{df}}{=} \bar{Q}_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{V}_N^{(M-m)} \right).
\end{aligned} \quad (4.169)$$

Po uwzględnieniu zależności (4.55) w $\bar{\theta}_{M-m+1N}^*$ (4.168) mamy

$$\begin{aligned}
\bar{V}_N^{(M-m)} &= \left[\Phi_{M-m} \left(\bar{v}_1^{(M-m-1)}, \theta_{M-m} \right) \Phi_{M-m} \left(\bar{v}_2^{(M-m-1)}, \theta_{M-m} \right) \dots \Phi_{M-m} \left(\bar{v}_N^{(M-m-1)}, \theta_{M-m} \right) \right] \\
&\stackrel{\text{df}}{=} \bar{\Phi}_{M-m} \left(\bar{V}_N^{(M-m-1)}, \theta_{M-m} \right),
\end{aligned} \quad (4.170)$$

a po podstawieniu (4.170) do (4.169) otrzymujemy

$$\begin{aligned} & \bar{Q}_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{V}_N^{(M-m)} \right) \\ &= \bar{Q}_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m)}, \bar{\Phi}_{M-m} \left(\bar{V}_N^{(M-m-1)}, \theta_{M-m} \right) \right). \end{aligned} \quad (4.171)$$

⋮

Krok M – 1. Oznaczamy sumę 2. składnika w kryterium (4.149) i \bar{Q}_{3N}^* – optymalnej postaci składowej kryterium wyznaczonej w $M - 2$. kroku – przez

$$\begin{aligned} \bar{Q}_{2N}(\theta_2) &\stackrel{\text{df}}{=} \alpha_0 \lambda_2 \sum_{n=1}^N q_2 \left(v_n^{(2)}, \Phi_{M-m+1} \left(\bar{v}_N^{(1)}, \theta_2 \right) \right) + \alpha_2 \sum_{n=1}^N q_2 \left(v_n^{(2)}, \Phi_2 \left(v_n^{(1)}, \theta_2 \right) \right) \\ &+ \tilde{Q}_{3N}^* \left(V_N^{(M)}, \dots, V_N^{(1)}, \tilde{\Phi}_2 \left(\bar{V}_N^{(1)}, \theta_2 \right) \right). \end{aligned} \quad (4.172)$$

Określamy optymalny wektor parametrów $\bar{\theta}_{2N}^*$, który minimalizuje wskaźnik (4.172). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\bar{\Psi}_{2N}$

$$\bar{\theta}_{2N}^* = \bar{\Psi}_{2N} \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(1)} \right), \quad (4.173)$$

który minimalizuje wskaźnik (4.137) względem θ_2 , czyli

$$\bar{\theta}_{2N}^* = \bar{\Psi}_{2N} \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(1)} \right) : \bar{Q}_{2N} \left(\bar{\theta}_{2N}^* \right) = \min_{\theta_2 \in \Theta_2} \bar{Q}_{2N}(\theta_2), \quad (4.174)$$

gdzie: $V_N^{(M)}, \dots, V_N^{(1)}$ są macierzami pomiarów (4.145) dla $m = 1, 2, \dots, M$, natomiast $\bar{V}_N^{(1)}$ jest macierzą wartości składowych wektora wejść modelu 2. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu 1. elementu, czyli

$$\bar{V}_N^{(1)} \stackrel{\text{df}}{=} \left[\bar{v}_1^{(1)} \quad \bar{v}_2^{(1)} \quad \dots \quad \bar{v}_N^{(1)} \right]. \quad (4.175)$$

W algorytmie (4.173) występuje macierz (4.175), która nie jest znana i zależy od macierzy wartości składowych wektora wyjść modelu poprzedniego – pierwszego elementu.

Po podstawieniu do (4.172) algorytmu wyznaczania optymalnego wektora parametrów modelu 2. elementu $\bar{\theta}_{2N}^*$, uzyskanego w wyniku rozwiązania zadania (4.174), otrzymujemy optymalną postać kryterium

$$\bar{Q}_{2N} \left(\bar{\theta}_{2N}^* \right) = \bar{Q}_{2N} \left(\bar{\Psi}_{2N} \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(1)} \right) \right) \stackrel{\text{df}}{=} \bar{Q}_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(1)} \right). \quad (4.176)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.175) mamy

$$\bar{V}_N^{(1)} = \left[\Phi_{M-2} \left(\bar{v}_1^{(0)}, \theta_1 \right) \quad \Phi_{M-m} \left(\bar{v}_2^{(0)}, \theta_1 \right) \quad \dots \quad \Phi_{M-2} \left(\bar{v}_N^{(0)}, \theta_1 \right) \right] \stackrel{\text{df}}{=} \bar{\Phi}_1 \left(\bar{V}_N^{(0)}, \theta_1 \right), \quad (4.177)$$

a po podstawieniu (4.177) do (4.176) otrzymujemy

$$\bar{Q}_{2N}^* \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{(1)} \right) = \bar{Q}_{2N}^* \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{\Phi}_1 \left(\bar{V}_N^{(0)}, \theta \right) \right). \quad (4.178)$$

Krok M. Oznaczamy sumę 1. składnika w kryterium (4.149) i \bar{Q}_{2N}^* – optymalnej postaci składowej kryterium wyznaczonej w $M - 1$. kroku – przez

$$\tilde{Q}_{1N}(\theta_1) \stackrel{\text{df}}{=} \alpha_0 \lambda_1 \sum_{n=1}^N q_1 \left(v_n^{(1)}, \Phi_1 \left(\bar{v}^{(0)}, \theta_1 \right) \right) + \bar{Q}_{2N}^* \left(V_N^{(M)}, \dots, V_N^{(1)}, \tilde{\Phi}_1 \left(\bar{V}_N^{(0)}, \theta_1 \right) \right). \quad (4.179)$$

Określamy optymalny wektor parametrów $\bar{\theta}_{1N}^*$, który minimalizuje wskaźnik (4.179). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji $\bar{\Psi}_{1N}$

$$\bar{\theta}_{1N}^* = \bar{\Psi}_{1N} \left(V_N^{(M)}, \dots, V_N^{(0)}, \bar{V}_N^{(0)} \right), \quad (4.180)$$

który minimalizuje wskaźnik (4.179) względem θ_1 , czyli

$$\bar{\theta}_{1N}^* = \bar{\Psi}_{1N} \left(V_N^{(M)}, \dots, V_N^{(0)}, \bar{V}_N^{(0)} \right) : \bar{Q}_{1N} \left(\bar{\theta}_{1N}^* \right) = \min_{\theta_1 \in \Theta_1} \bar{Q}_{1N} \left(\theta_1 \right), \quad (4.181)$$

gdzie: $V_N^{(M)}, \dots, V_N^{(0)}$ są macierzami pomiarów (4.145) dla $m = 0, 1, \dots, M$.

Zauważmy, że dopiero teraz występujące w algorytmie (4.180), z uwzględnieniem (4.100), wszystkie macierze pomiarów są znane i możliwe jest wyznaczenie optymalnego wektora parametrów modelu pierwszego elementu.

Optymalne wartości składowych wektorów parametrów $\bar{\theta}_{MN}^*, \bar{\theta}_{M-1N}^*, \dots, \bar{\theta}_{2N}^*$, wyznaczone w krokach 1., 2., ..., $M - 1$., oprócz macierzy pomiarów wartości składowych wektora wyjść poszczególnych elementów

$$V_N^{(M)}, V_N^{(M-1)}, \dots, V_N^{(0)}, \quad (4.182)$$

zależą od macierzy wartości składowych wektora wyjść modeli

$$\bar{V}_N^{(M-1)}, \bar{V}_N^{(M-1)}, \dots, \bar{V}_N^{(1)}. \quad (4.183)$$

Dopiero wektor $\bar{\theta}_{1N}^*$ zależy od serii identyfikującej $X_N = V_N^{(0)} = \bar{V}_N^{(0)}$ oraz macierzy pomiarów wartości składowych wektorów wszystkich wyjść $V_N^{(1)}, V_N^{(2)}, \dots, V_N^{(M)}$. Powracając do algorytmów identyfikacji określonych w kolejnych krokach, możemy wyznaczyć optymalne wartości składowych wektorów parametrów modeli poszczególnych elementów systemu. I tak:

Krok M + 1. Zgodnie z (4.56) oraz (4.100)

$$x_n = \bar{v}_n^{(0)} = v_n^{(0)}, \quad (4.184)$$

a zatem macierz pomiarów wartości składowych wektora wejść zewnętrznych spełnia

$$X_N = V_N^{(0)} = \bar{V}_N^{(0)}. \quad (4.185)$$

Po uwzględnieniu (4.185) w (4.180) otrzymujemy algorytm

$$\bar{\theta}_{1N}^* = \bar{\Psi}_{1N} \left(V_N^{(M)}, \dots, V_N^{(1)}, X_N \right), \quad (4.186)$$

który na podstawie macierzy pomiarów wartości składowych wektora wejść zewnętrznych X_N oraz macierzy pomiarów wartości składowych wszystkich wektorów wyjść $V_N^{(1)}, V_N^{(2)}, \dots, V_N^{(M)}$ umożliwia wyznaczenie wartości optymalnego wektora parametrów modelu 1. elementu $\bar{\theta}_{1N}^*$.

Krok $M + 2$. Teraz, po podstawieniu do (4.177) w miejsce θ_1 optymalnego wektora $\bar{\theta}_{1N}^*$, określonego przez (4.186), otrzymujemy macierz wartości składowych wektora wyjść modelu 1. elementu systemu z optymalnymi parametrami dla zadanej serii identyfikującej X_N , tj.

$$\bar{V}_N^{*(1)} \stackrel{\text{df}}{=} \tilde{\Phi}_1 \left(X_N, \bar{\theta}_{1N}^* \right) = \tilde{\Phi}_1 \left(X_N, \bar{\Psi}_{1N} \left(V_N^{(M)}, \dots, V_N^{(1)}, X_N \right) \right). \quad (4.187)$$

Po podstawieniu w zależności (4.173) w miejsce $\bar{V}_N^{(1)}$ macierzy wartości składowych wektora wyjść modelu optymalnego, określonego przez (4.187), otrzymujemy

$$\bar{\theta}_{2N}^* = \bar{\Psi}_{2N} \left(V_N^{(M)}, \dots, V_N^{(1)}, \bar{V}_N^{*(1)} \right). \quad (4.188)$$

$$\vdots$$

Krok $M + m$. W kroku $M + m - 1$. wyznaczyliśmy macierz wartości składowych wektora wyjść $\bar{V}_N^{*(m-2)}$ modelu $m - 2$. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , która jest jednocześnie macierzą wartości składowych wektora wejść do $m - 1$. elementu modelu. Umożliwia to wyznaczenie macierzy wartości składowych wektora wyjść modelu $m - 1$. elementu $\bar{V}_N^{*(m-1)}$ z optymalnym wektorem parametrów $\bar{\theta}_{m-1N}^*$, tj.

$$\begin{aligned} \bar{V}_N^{*(m-1)} &\stackrel{\text{df}}{=} \tilde{\Phi}_{m-1} \left(\bar{V}_N^{*(m-2)}, \bar{\theta}_{m-1N}^* \right) \\ &= \tilde{\Phi}_{m-1} \left(\bar{V}_N^{*(m-2)}, \bar{\Psi}_{m-1N} \left(V_N^{(M)}, \dots, V_N^{(m-1)}, \bar{V}_N^{*(m-2)} \right) \right). \end{aligned} \quad (4.189)$$

Wartość optymalnego wektora parametrów modelu m -tego elementu $\bar{\theta}_{mN}^*$ wyznaczamy według zależności

$$\bar{\theta}_{mN}^* = \bar{\Psi}_{mN} \left(V_N^{(M)}, \dots, V_N^{(m-1)}, \bar{V}_N^{*(m-1)} \right). \quad (4.190)$$

$$\vdots$$

Krok 2M. W kroku $2M - 1$. wyznaczyliśmy macierz wartości składowych wektora wyjść $\bar{V}_N^{*(M-2)}$ modelu $M - 2$. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , która jest jednocześnie macierzą wartości składowych wektora wejść do modelu $M - 1$. elementu. Umożliwia to wyznaczenie macierzy wartości składowych wektora wyjść modelu $M - 1$. elementu $\bar{V}_N^{*(M-1)}$ z optymalnym wektorem parametrów $\bar{\theta}_{M-1N}^*$, tj.

$$\begin{aligned} \bar{V}_N^{*(M-1)} &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-1}(\bar{V}_N^{*(M-2)}, \bar{\theta}_{M-1N}^*) \\ &= \tilde{\Phi}_{M-1}(\bar{V}_N^{*(M-2)}, \bar{\Psi}_{M-1N}(V_N^{(M)}, \dots, V_N^{(M-1)}, \bar{V}_N^{*(M-2)})). \end{aligned} \quad (4.191)$$

Ostatecznie, po podstawieniu do algorytmu (4.157) macierzy $\bar{V}_N^{*(M-1)}$, określonej przez (4.191), otrzymujemy

$$\tilde{\theta}_{MN}^* = \tilde{\Psi}_{MN}(V_N^{(M)}, V_N^{(M-1)}, \bar{V}_N^{*(M-1)}), \quad (4.192)$$

wartość optymalnego wektora parametrów modelu M -tego elementu $\bar{\theta}_{MN}^*$.

Globalny wskaźnik jakości identyfikacji z zadowalającym modelem lokalnym

Zastosowanie metody programowania dynamicznego wymaga drobnej modyfikacji poprzednio przedstawionego algorytmu identyfikacji, przy globalnym wskaźniku jakości (4.62) oraz zadowalającym modelem lokalnym (4.41) dla struktury szeregowej.

Po uwzględnieniu (4.62) oraz wyodrębnieniu wektorów parametrów modeli poszczególnych elementów zadanie (4.41) sprowadza się do wyznaczenia takiego zestawu wektorów parametrów modeli elementów $\tilde{\theta}_{1N}^{**}, \tilde{\theta}_{2N}^{**}, \dots, \tilde{\theta}_{MN}^{**}$, który minimalizuje wskaźnik (4.62) względem $\theta_1, \theta_2, \dots, \theta_M$, czyli

$$\begin{aligned} \tilde{\theta}_{1N}^{**}, \tilde{\theta}_{2N}^{**}, \dots, \tilde{\theta}_{MN}^{**} &: \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \tilde{\theta}_{mN}^{**})) \\ &= \min_{\theta_1, \theta_2, \dots, \theta_M \in \Theta_1^* \times \Theta_2^* \times \dots \times \Theta_M^*} \sum_{m=1}^M \lambda_m \sum_{n=1}^N q_m(v_n^{(m)}, \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)), \end{aligned} \quad (4.193)$$

przy czym

$$\Theta_m^* \stackrel{\text{df}}{=} \left\{ \theta_m \in \Theta_m \subseteq \mathbf{R}^{R_m} : Q_{mN}(\theta_m) \leq \beta_m, \beta_m > Q_{mN}(\theta_{mN}^*), m = 1, 2, \dots, M \right\}. \quad (4.194)$$

W tym przypadku algorytm identyfikacji jest analogiczny do algorytmu przedstawionego w punkcie 4.4.1 dla globalnie optymalnego modelu, z tą różnicą, że w kolejnych krokach procedury w zadaniach optymalizacji należy uwzględnić ograniczenia (4.194). W konsekwencji m -ty oraz $M + m$ -ty krok procedury przyjmie postać:

Krok m. Oznaczamy sumę $M - m + 1$. składnika w kryterium (4.62) oraz \tilde{Q}_{M-m+2N}^{**} – optymalnej postaci składowej kryterium wyznaczonej w $m - 1$. kroku – przez

$$\begin{aligned} \tilde{Q}_{M-m+1N}(\theta_{M-m+1}) &\stackrel{\text{df}}{=} \lambda_{M-m+1} \sum_{n=1}^N q_{M-m+1} \left(v_n^{(M-m+1)}, \Phi_{M-m+1} \left(\bar{v}^{(M-m)}, \theta_{M-m+1} \right) \right) \\ &+ \tilde{Q}_{M-m+2N}^{**} \left(V_N^{(M)}, \dots, V_N^{(M-m+2)}, \tilde{\Phi}_{M-m+1} \left(\bar{V}_N^{(M-m)}, \theta_{M-m+1} \right) \right). \end{aligned} \quad (4.195)$$

Określamy optymalny wektor parametrów $\tilde{\theta}_{M-m+1N}^*$, który minimalizuje wskaźnik (4.195). Zadanie sprowadza się do wyznaczenia takiego algorytmu identyfikacji Ψ_{M-m+1N}^*

$$\tilde{\theta}_{M-m+1N}^{**} = \Psi_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)} \right), \quad (4.196)$$

który minimalizuje wskaźnik (4.195) względem θ_{M-m+1} , czyli

$$\begin{aligned} \tilde{\theta}_{M-m+1N}^{**} &= \Psi_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)} \right) : \tilde{Q}_{M-m+1N} \left(\tilde{\theta}_{M-m+1N}^{**} \right) \\ &= \min_{\theta_{M-m+1} \in \Theta_{M-m+1}^*} \tilde{Q}_{M-m+1N} \left(\theta_{M-m+1} \right), \end{aligned} \quad (4.197)$$

gdzie: $V_N^{(M-m+1)}$ jest macierzą pomiarów wartości składowych wektora wyjść $M - m + 1$. elementu, tj.

$$V_N^{(M-m+1)} = \left[v_1^{(M-m+1)} \ v_2^{(M-m+1)} \ \dots \ v_N^{(M-m+1)} \right], \quad (4.198)$$

natomiast $\bar{V}_N^{(M-m)}$ jest macierzą wartości składowych wektora wejść modelu $M - m + 1$. elementu, która jest jednocześnie macierzą wartości składowych wektora wyjść modelu $M - m$ -tego elementu, czyli

$$\bar{V}_N^{(M-m)} = \left[\bar{v}_1^{(M-m)} \ \bar{v}_2^{(M-m)} \ \dots \ \bar{v}_N^{(M-m)} \right]. \quad (4.199)$$

Po podstawieniu do (4.195) algorytmu wyznaczania optymalnego wektora parametrów modelu $M - m + 1$. elementu $\tilde{\theta}_{M-m+1N}^{**}$, uzyskanego w wyniku rozwiązania zadania (4.197), otrzymujemy optymalną postać kryterium

$$\begin{aligned} \tilde{Q}_{M-m+1N} \left(\tilde{\theta}_{M-m+1N}^{**} \right) &= \tilde{Q}_{M-m+1N} \left(\Psi_{M-m+1N}^* \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)} \right) \right) \\ &\stackrel{\text{df}}{=} \tilde{Q}_{M-m+1N}^{**} \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)} \right). \end{aligned} \quad (4.200)$$

Po uwzględnieniu zależności (4.55) w macierzy (4.199) mamy

$$\begin{aligned} \bar{V}_N^{(M-m)} &= \left[\Phi_{M-m} \left(\bar{v}_1^{(M-m-1)}, \theta_{M-m} \right) \Phi_{M-m} \left(\bar{v}_2^{(M-m-1)}, \theta_{M-m} \right) \dots \Phi_{M-m} \left(\bar{v}_N^{(M-m-1)}, \theta_{M-m} \right) \right] \\ &\stackrel{\text{df}}{=} \tilde{\Phi}_{M-m} \left(\bar{V}_N^{(M-m-1)}, \theta_{M-m} \right), \end{aligned} \quad (4.201)$$

a po podstawieniu (4.201) do (4.200) otrzymujemy

$$\begin{aligned} & \tilde{Q}_{M-m+1N}^{**} \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \bar{V}_N^{(M-m)} \right) \\ &= \tilde{Q}_{M-m+1N}^{**} \left(V_N^{(M)}, \dots, V_N^{(M-m+1)}, \tilde{\Phi}_{M-m} \left(\bar{V}_N^{(M-m-1)}, \theta_{M-m} \right) \right). \end{aligned} \quad (4.202)$$

Analogicznie, w kroku $M + m$ -tym należy skorzystać z optymalnych wartości parametrów $\tilde{\theta}_{m-1N}^{**}$ oraz algorytmów Ψ_{m-1}^* i Ψ_m^* , czyli:

Krok $M + m$. W kroku $M + m - 1$. wyznaczyliśmy macierz wartości składowych wektora wyjść $\bar{V}_N^{** (m-2)}$ modelu $m - 2$. elementu systemu z optymalnym wektorem parametrów dla zadanej serii identyfikującej X_N , która jest jednocześnie macierzą wartości składowych wektora wejść modelu $m - 1$. elementu modelu. Umożliwia to wyznaczenie macierzy wartości wektora wyjść modelu $m - 1$. elementu $\bar{V}_N^{** (m-1)}$ z optymalnym wektorem parametrów $\tilde{\theta}_{m-1N}^{**}$, tj.

$$\begin{aligned} \bar{V}_N^{** (m-1)} &\stackrel{\text{df}}{=} \tilde{\Phi}_{m-1} \left(\bar{V}_N^{** (m-2)}, \tilde{\theta}_{m-1N}^{**} \right) \\ &= \tilde{\Phi}_{m-1} \left(\bar{V}_N^{** (m-2)}, \Psi_{m-1N}^* \left(V_N^{(M)}, \dots, V_N^{(m-1)}, \bar{V}_N^{** (m-2)} \right) \right). \end{aligned} \quad (4.203)$$

Ostatecznie wartość optymalnego wektora parametrów modelu m -tego elementu $\tilde{\theta}_{mN}^*$ wyznaczamy według zależności

$$\tilde{\theta}_{mN}^{**} = \Psi_{mN}^* \left(V_N^{(M)}, \dots, V_N^{(m)}, \bar{V}_N^{** (m-1)} \right). \quad (4.204)$$

4.5. Modelowanie systemów złożonych za pomocą sieci neuronowych

Jak wspomniano w punkcie 2.3.3, jednym z często stosowanych opisów systemów jest model w postaci sieci neuronowej. Teraz założymy, że dla każdego elementu systemu złożonego proponujemy model w postaci wielowarstwowej sieci neuronowej (rys. 4.8).

Tak jak w punkcie 2.3.2, w j -tej warstwie sieci, która modeluje m -ty element systemu złożonego, znajduje się I_{j_m} neuronów. Przez E_{mij} oznaczmy i -ty neuron w j -tej warstwie, $i = 1, 2, \dots, I_{j_m}$, $j_m = 1, 2, \dots, J_m$. Schemat neuronu E_{mij} przedstawiono na rysunku 4.9.

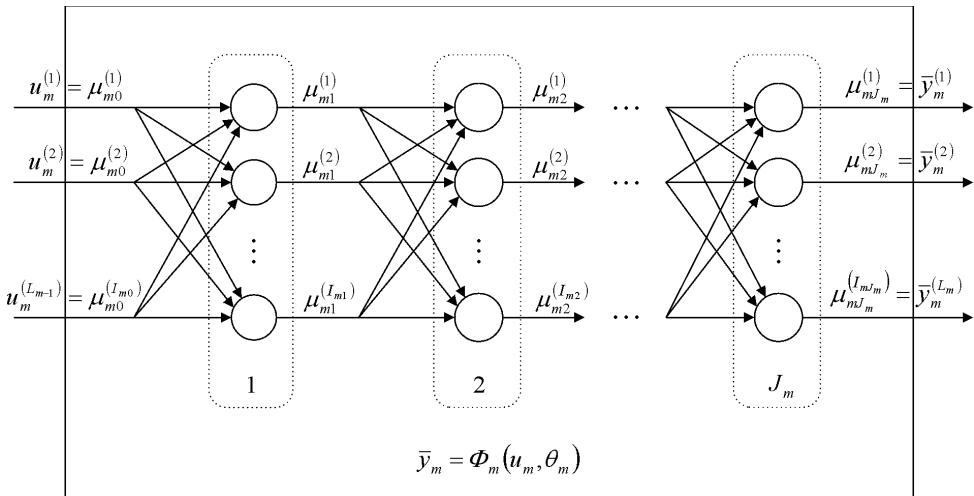
Dla neuronu E_{mij} z wejściami

$$\mu_{mj-1} = \begin{bmatrix} \mu_{mj-1}^{(1)} \\ \mu_{mj-1}^{(2)} \\ \vdots \\ \mu_{mj-1}^{(I_{j_m-1})} \end{bmatrix} \quad (4.205)$$

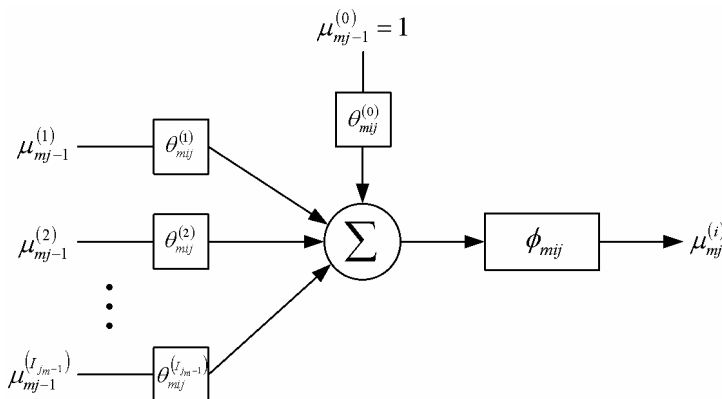
oraz wyjściem $\mu_{mj}^{(i)}$ proponujemy model

$$\mu_{mj}^{(i)} = \phi_{mij} \left(\sum_{s=1}^{I_{jm}-1} \mu_{mj-1}^{(s)} \theta_{mij}^{(s)} + \theta_{mij}^{(0)} \right), \quad (4.206)$$

gdzie: $\theta_{mij}^{(s)}, s = 0, 1, \dots, I_{jm}-1$ jest zestawem współczynników wagowych, a ϕ_{mij} – funkcją aktywacji.



Rys. 4.8. Wielowarstwowa jednokierunkowa sieć neuronowa dla m -tego elementu



Rys. 4.9. Neuron E_{mij}

Po oznaczeniu przez

$$\bar{\mu}_{mj-1} \stackrel{\text{df}}{=} \begin{bmatrix} \mu_{mj-1}^{(0)} \\ \mu_{mj-1}^{(1)} \\ \mu_{mj-1}^{(2)} \\ \vdots \\ \mu_{mj-1}^{(I_{jm-1})} \end{bmatrix}, \quad (4.207)$$

przy czym $\mu_{mj-1}^{(0)} = 1$,

oraz uwzględnieniu wektora współczynników wagowych

$$\theta_{mij} \stackrel{\text{df}}{=} \begin{bmatrix} \theta_{mij}^{(0)} \\ \theta_{mij}^{(1)} \\ \vdots \\ \theta_{mij}^{(I_{jm-1})} \end{bmatrix} \quad (4.208)$$

model (4.206) neuronu E_{mij} możemy przedstawić w zwartej postaci

$$\mu_{mj}^{(i)} = \phi_{mij}(\bar{\mu}_{mj-1}^T \theta_{mij}). \quad (4.209)$$

Model j -tej warstwy przyjmuje postać

$$\mu_{mj} = \begin{bmatrix} \mu_{mj}^{(1)} \\ \mu_{mj}^{(2)} \\ \vdots \\ \mu_{mj}^{(I_j)} \end{bmatrix} = \begin{bmatrix} \phi_{1j}(\bar{\mu}_{mj-1}^T \theta_{1j}) \\ \phi_{2j}(\bar{\mu}_{mj-1}^T \theta_{2j}) \\ \vdots \\ \phi_{I_{jm}j}(\bar{\mu}_{mj-1}^T \theta_{I_{jm}j}) \end{bmatrix} \stackrel{\text{df}}{=} \phi_{mj}(\bar{\mu}_{j-1}, \theta_{mj}), \quad (4.210)$$

gdzie ϕ_{mj} jest wektorem funkcji charakteryzującej j -tą warstwę m -tej sieci z wejściami μ_{mj-1} oraz macierzą parametrów

$$\theta_{mj} \stackrel{\text{df}}{=} \begin{bmatrix} \theta_{m1j} & \theta_{m2j} & \cdots & \theta_{mI_{jm}j} \end{bmatrix} = \begin{bmatrix} \theta_{m1j}^{(0)} & \theta_{m2j}^{(0)} & \cdots & \theta_{mI_{jm}j}^{(0)} \\ \theta_{m1j}^{(1)} & \theta_{m2j}^{(1)} & \cdots & \theta_{mI_{jm}j}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{m1j}^{(I_{jm-1})} & \theta_{m2j}^{(I_{jm-1})} & \cdots & \theta_{mI_{jm}j}^{(I_{jm-1})} \end{bmatrix}. \quad (4.211)$$

Połączenie kolejnych warstw daje sieć neuronową, której rekurencyjny model przyjmuje postać

$$\mu_{mj} = \phi_{mj}(\bar{\mu}_{mj-1}, \theta_{mj}), \quad j = 1, 2, \dots, J_m. \quad (4.212)$$

Wejściami do warstwy wejściowej są wejścia m -tego elementu systemu złożonego, czyli $\mu_{m0} = u_m$ oraz $I_{m0} = S_m$, a zatem

$$\mu_{m1} = \phi_{m1}(\bar{\mu}_{m0}, \theta_{m1}) = \phi_{m1}(u_m, \theta_{m1}). \quad (4.213)$$

Dla warstw ukrytych mamy rekurencyjną zależność

$$\mu_{mj} = \phi_{mj}(\bar{\mu}_{mj-1}, \theta_{mj}), \quad j = 2, 3, \dots, J_m - 1. \quad (4.214)$$

Wyjścia warstwy wyjściowej są wyjściami modelu m -tego elementu systemu złożonego, tj. $\mu_{mJ_m} = \bar{y}_m$ oraz $I_{mJ_m} = L_m$, czyli

$$\bar{y}_m = \mu_{mJ_m} = \phi_{mJ_m}(\bar{\mu}_{mJ_m-1}, \theta_{mJ_m}). \quad (4.215)$$

Złożenie (4.213), (4.214) oraz (4.215) daje charakterystykę (4.1) postaci

$$\begin{aligned} \bar{y}_m &= \phi_{mJ_m} \left(\phi_{mJ_m-1} \left(\dots \phi_{m1} (u_m, \theta_{m1}), \dots, \theta_{mJ_m-1} \right), \theta_{mJ_m} \right) \\ &\stackrel{\text{df}}{=} \begin{bmatrix} \Phi_m^{(1)}(u_m, \theta_m) \\ \Phi_m^{(2)}(u_m, \theta_m) \\ \vdots \\ \Phi_m^{(L_m)}(u_m, \theta_m) \end{bmatrix} \stackrel{\text{df}}{=} \Phi_m(u_m, \theta_m), \end{aligned} \quad (4.216)$$

gdzie θ_m jest zestawem parametrów złożonych z macierzy (4.211) parametrów kolejnych warstw, czyli

$$\theta_m \stackrel{\text{df}}{=} \{ \theta_{m1}, \theta_{m2}, \dots, \theta_{mJ_m} \}. \quad (4.217)$$

Teraz, biorąc pod uwagę rozważania przedstawione w podrozdziałach 4.1 oraz 4.2, możemy zaproponować lokalnie optymalny lub globalnie optymalny algorytm uczenia sieci dla systemu złożonego.

4.5.1. Algorytm uczenia sieci neuronowej dla systemu złożonego z lokalnym wskaźnikiem jakości identyfikacji

Wyznaczenie lokalnie optymalnych parametrów modelu postaci (4.217) dla zadanej serii identyfikującej wraz z wynikami eksperymentu (4.18), czyli dla tak zwanego zbioru uczącego postaci (4.18), sprowadza się do rozwiązania zadania optymalizacji (4.23), w którym w kryterium jakości przybliżenia (4.19) w miejsce $\Phi_m(u_{mn}, \theta_m)$, $n = 1, 2, \dots, N$, wstawimy model postaci (4.216).

Podobnie jak w punkcie 2.3.2, ograniczymy się do algorytmu wstecznej propagacji błędów, przy kwadratowym wskaźniku jakości przybliżenia. Wskaźnik (4.19) ma zatem postać

$$\begin{aligned}
Q_{mN_m}(\theta_m) &= \sum_{n=1}^{N_m} [y_{mn} - \Phi_m(u_{mn}, \theta_m)]^T [y_{mn} - \Phi_m(u_{mn}, \theta_m)] \\
&= \sum_{n=1}^{N_m} \sum_{l=1}^{L_m} (y_{mn}^{(l)} - \Phi_m^{(l)}(u_{mn}, \theta))^2.
\end{aligned} \tag{4.218}$$

Korzystając z numerycznej metody optymalizacji do rozwiązania zadania (4.23) z kryterium (4.218), w kolejnych krokach algorytmu otrzymujemy zbiór wartości $\tilde{\theta}_{m1}, \tilde{\theta}_{m2}, \dots, \tilde{\theta}_{mN_m}$ takich, że $Q_{mN_m}(\tilde{\theta}_{m1}) > Q_{mN_m}(\tilde{\theta}_{m2}) > \dots > Q_{mN_m}(\tilde{\theta}_{mN_m})$, które po N_m krokach są przybliżeniem optymalnych wartości parametrów $\theta_{mN_m}^*$ modelu (4.217), czyli $\tilde{\theta}_{mN_m} \approx \theta_{mN_m}^*$ oraz $Q_{mN_m}(\tilde{\theta}_{mN_m}) \approx Q_{mN_m}(\theta_{mN_m}^*)$. Zastosowanie metody gradientu prostego daje algorytm wstecznej propagacji błędów przy lokalnym wskaźniku jakości identyfikacji postaci

$$\begin{aligned}
\tilde{\theta}_{mij,n+1}^{(s)} &= \tilde{\theta}_{mij,n}^{(s)} - \eta_{mn} \left. \frac{\partial q_{mn}(\theta_m)}{\partial \theta_{mij}^{(s)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij,n}^{(s)}}, \\
s &= 0, 1, \dots, I_{J_m-1}, \quad i = 1, 2, \dots, I_{J_m}, \quad j = 1, 2, \dots, J_m,
\end{aligned} \tag{4.219}$$

gdzie: η_{mn} jest współczynnikiem uczenia, natomiast

$$\begin{aligned}
q_{mn}(\theta_m) &\stackrel{\text{df}}{=} [y_{mn} - \Phi_m(u_{mn}, \theta_m)]^T [y_{mn} - \Phi_m(u_{mn}, \theta_m)] \\
&= \sum_{l=1}^{L_m} (y_{mn}^{(l)} - \Phi_m^{(l)}(u_n, \theta_m))^2 = \sum_{l=1}^{L_m} \varepsilon_{mn}^{(l)2}
\end{aligned} \tag{4.220}$$

jest n -tym składnikiem sumy (4.218), który odpowiada n -temu elementowi ciągu uczącego, a

$$\varepsilon_{mn}^{(l)} \stackrel{\text{df}}{=} y_{mn}^{(l)} - \bar{y}_{mn}^{(l)} = y_{mn}^{(l)} - \Phi_m^{(l)}(u_{mn}, \theta_m), \quad l = 1, 2, \dots, L_m \tag{4.221}$$

jest różnicą pomiędzy l -tym wyjściem obiektu a odpowiednim wyjściem sieci.

Różniczkując wyrażenie (4.220) po poszczególnych parametrach, z uwzględnieniem modelu neuronu (4.206), otrzymujemy

$$\left. \frac{\partial q_{mn}(\theta_m)}{\partial \theta_{mij}^{(s)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij,n}^{(s)}} = -2\delta_{mjn}^{(i)} \mu_{mj-1,n}^{(s)}, \tag{4.222}$$

gdzie: $\mu_{mj-1,n}^{(0)} = 1$, a $\mu_{mj-1,n}^{(s)}$, $s = 1, 2, \dots, I_{j_m-1}$ są wartościami wejść w j -tej warstwie, w n -tym kroku procedury uczenia

$$\varepsilon_{mjn}^{(i)} \stackrel{\text{df}}{=} \begin{cases} \varepsilon_{mn}^{(i)} & \text{dla } j = J_m, \\ \sum_{p=1}^{I_{j_m+1}} \delta_{mj+1,n}^{(p)} \tilde{\theta}_{mpj+1,n}^{(i)} & \text{dla } j = J_m - 1, J_m - 2, \dots, 1, \end{cases} \quad (4.223)$$

$$\delta_{mjn}^{(i)} \stackrel{\text{df}}{=} \varepsilon_{mjn}^{(i)} \left. \frac{d\phi_{mij}(\kappa_{mj}^{(i)})}{d\kappa_{mj}^{(i)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij,n}^{(s)}} \quad (4.224)$$

oraz

$$\kappa_{mj}^{(i)} \stackrel{\text{df}}{=} \sum_{s=1}^{I_{j_m-1}} \mu_{mj-1}^{(s)} \theta_{mij}^{(s)} + \theta_{mij}^{(0)}.$$

Ostatecznie, po podstawieniu powyższego do (4.219), algorytm wstecznej propagacji błędów z lokalnym wskaźnikiem jakości identyfikacji przyjmuje postać

$$\begin{aligned} \tilde{\theta}_{mij,n+1}^{(s)} &= \tilde{\theta}_{mij,n}^{(s)} + 2\eta_{mn} \delta_{mjn}^{(i)} \mu_{mj-1,n}^{(s)}, \\ s &= 0, 1, \dots, I_{j_m-1}, \quad i = 1, 2, \dots, I_{J_m}, \quad j = 1, 2, \dots, J_m. \end{aligned} \quad (4.225)$$

Obliczenia wykonujemy niezależnie dla każdego elementu systemu złożonego, czyli dla $m = 1, 2, \dots, M$.

4.5.2. Algorytm uczenia sieci neuronowej dla systemu złożonego z globalnym wskaźnikiem jakości identyfikacji

Wyznaczenie globalnie optymalnego modelu systemu złożonego z zastosowaniem sieci neuronowych do opisu poszczególnych elementów wymaga opracowania odpowiedniego algorytmu uczenia sieci [31, 32, 34, 128]. Bazując na modelu w postaci sieci neuronowej (4.216), dla każdego z elementów systemu złożonego oraz po uwzględnieniu struktury modelu zadanego równaniami (4.9) i (4.10) otrzymamy model (4.17). Wyznaczenie globalnie optymalnych parametrów modelu postaci (4.17), dla zadanej serii identyfikującej wraz z wynikami eksperymentu (4.27), sprowadza się do rozwiązania zadania optymalizacji (4.32), w którym – w kryterium jakości przybliżenia (4.28) – w miejsce $\Phi(x_n, \theta)$, $n = 1, 2, \dots, N$, wstawimy model postaci (4.17), z uwzględnieniem (4.216).

Dla dowolnej struktury systemu złożonego, różnych typów sieci z dowolnymi postaciami modeli neuronów oraz różnych kryteriów jakości przybliżenia otrzymujemy złożone zadanie optymalizacji. Stosując wybraną numeryczną metodę optymalizacji do rozwiązania zadania (4.32), otrzymamy globalnie optymalne algorytmy uczenia sieci dla systemu złożonego.

Po przyjęciu dodatkowych założeń dotyczących struktury systemu, rodzaju sieci oraz metody optymalizacji można uzyskać konkretne algorytmy uczenia. Podobnie jak

w punkcie 2.3.2, ograniczymy się do algorytmu wstecznej propagacji błędu, przy kwadratowym wskaźniku jakości przybliżenia oraz dla systemu o strukturze szeregowej (rys. 4.5). Teraz model (4.53) z uwzględnieniem modeli (4.216) tworzy model systemu złożonego w postaci sieci neuronowej, przy czym w modelu (4.216) przyjmujemy: $\bar{y}_m = \bar{v}^{(m)}$, $u_m = \bar{v}^{(m-1)}$, $m = 1, 2, \dots, M$, a $\bar{v}^{(0)} = x$.

Połączenie kolejnych warstw daje sieć neuronową modelującą system złożony, której rekurencyjny model przyjmuje postać

$$\begin{aligned} \mu_{mj} &= \phi_{mj}(\mu_{mj-1}, \theta_{mj}), \\ j &= 1, 2, \dots, J_m, \quad m = 1, 2, \dots, M. \end{aligned} \quad (4.226)$$

Teraz – oprócz warstw: wejściowej, ukrytych oraz wyjściowej – pojawiają się dodatkowe warstwy ukryte, które odpowiadają połączeniu poszczególnych elementów systemu złożonego. Wejścia zewnętrzne x systemu złożonego pokrywają się z wejściami warstwy wejściowej, czyli $\mu_{10} = u_1 = x$ oraz $I_{10} = S$, a zatem

$$\mu_{11} = \phi_{11}(\mu_{10}, \theta_{11}) = \phi_{11}(x, \theta_{11}). \quad (4.227)$$

Dla warstw ukrytych mamy rekurencyjną zależność

$$\begin{aligned} \mu_{mj} &= \phi_{mj}(\mu_{mj-1}, \theta_{mj}), \\ j &= 2, 3, \dots, J_m - 1, \quad m = 1, 2, \dots, M - 1. \end{aligned} \quad (4.228)$$

Wyjścia dodatkowych warstw są wyjściami modelu m -tego elementu systemu złożonego, tj.: $\mu_{mJ_m} = \bar{y}_m \stackrel{\text{df}}{=} \bar{v}^{(m)}$ oraz $I_{J_m} = L_m$, czyli

$$\begin{aligned} \bar{v}^{(m)} = \bar{y}_m &= \mu_{mJ_m} = \phi_{mJ_m}(\mu_{mJ_m-1}, \theta_{mJ_m}), \\ m &= 1, 2, \dots, M. \end{aligned} \quad (4.229)$$

Złożenie (4.227), (4.228) oraz (4.229) daje

$$\bar{v}^{(m)} = \begin{bmatrix} \bar{v}^{(m1)} \\ \bar{v}^{(m2)} \\ \vdots \\ \bar{v}^{(mL_m)} \end{bmatrix} = \phi_{mJ_m} \left(\phi_{mJ_m-1} \left(\dots \phi_{m1} \left(\bar{v}^{(m-1)}, \theta_{m1} \right), \dots, \theta_{mJ_m-1} \right), \theta_{mJ_m} \right) \quad (4.230)$$

$$\stackrel{\text{df}}{=} \begin{bmatrix} \Phi_m^{(1)}(\bar{v}^{(m-1)}, \theta_m) \\ \Phi_m^{(2)}(\bar{v}^{(m-1)}, \theta_m) \\ \vdots \\ \Phi_m^{(L_m)}(\bar{v}^{(m-1)}, \theta_m) \end{bmatrix} \stackrel{\text{df}}{=} \Phi_m(\bar{v}^{(m-1)}, \theta_m), 1, 2, \dots, M,$$

gdzie: θ_m jest zestawem parametrów (4.217), natomiast $\bar{v}^{(0)} = x$.

Ostatecznie, po podstawieniu (4.230) do (4.53), otrzymamy model systemu złożonego w postaci sieci, gdzie

$$\theta \stackrel{\text{df}}{=} \{\theta_1, \theta_2, \dots, \theta_M\}, \quad (4.231)$$

jest zestawem parametrów sieci neuronowej modelującej system złożony, którego składowymi są parametry poszczególnych elementów sieci określone przez (4.217).

Kryterium jakości przybliżenia (4.28) dla struktury szeregowej i kwadratowego wskaźnika jakości przyjmuje postać

$$\begin{aligned} \tilde{Q}_N(\theta) &= \sum_{n=1}^N \sum_{m=1}^M \lambda_m [v_n^{(m)} - \bar{v}_n^{(m)}]^T [v_n^{(m)} - \bar{v}_n^{(m)}] \\ &= \sum_{n=1}^N \sum_{m=1}^M \lambda_m [v_n^{(m)} - \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)]^T [v_n^{(m)} - \Phi_m(\bar{v}_n^{(m-1)}, \theta_m)], \end{aligned} \quad (4.232)$$

a po uwzględnieniu (4.230) otrzymujemy

$$\tilde{Q}_N(\theta) = \sum_{n=1}^N \sum_{m=1}^M \lambda_m \sum_{l=1}^{L_m} (v_n^{(ml)} - \Phi_m^{(l)}(\bar{v}_n^{(m-1)}, \theta_m))^2. \quad (4.233)$$

Korzystając z numerycznej metody optymalizacji do rozwiązania zadania (4.32) z kryterium (4.233), w kolejnych krokach algorytmu otrzymujemy ciąg wartości $\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_N$, takich że $\tilde{Q}_N(\tilde{\theta}_1) > \tilde{Q}_N(\tilde{\theta}_2) > \dots > \tilde{Q}_N(\tilde{\theta}_N)$, które po N krokach są przybliżeniem optymalnych wartości wektora parametrów $\tilde{\theta}_N^*$ modelu (4.217), czyli $\tilde{\theta}_N \approx \tilde{\theta}_N^*$ oraz $\tilde{Q}_N(\tilde{\theta}_N) \approx \tilde{Q}_N(\tilde{\theta}_N^*)$.

Zastosowanie metody gradientu prostego daje algorytm wstecznej propagacji błędu przy globalnym wskaźniku jakości identyfikacji postaci

$$\tilde{\theta}_{mij,n+1}^{(s)} = \tilde{\theta}_{mij,n}^{(s)} - \eta_{mn} \left. \frac{\partial \tilde{q}_n(\theta)}{\partial \theta_{mij}^{(s)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij,n}^{(s)}}, \quad (4.234)$$

$$s = 0, 1, \dots, I_{J_m-1}, \quad i = 1, 2, \dots, I_{J_m}, \quad j = 1, 2, \dots, J_m, \quad m = 1, 2, \dots, M,$$

gdzie: η_{mn} jest współczynnikiem uczenia, natomiast

$$\tilde{q}_n(\theta) = \sum_{m=1}^M \lambda_m \sum_{l=1}^{L_m} (v_n^{(ml)} - \Phi_m^{(l)}(\bar{v}_n^{(m-1)}, \theta_m))^2 = \sum_{m=1}^M \lambda_m \sum_{l=1}^{L_m} \tilde{\varepsilon}_{mn}^{(l)2} \quad (4.235)$$

jest n -tym składnikiem sumy (4.218), który odpowiada n -temu elementowi zbioru uczącego, a

$$\tilde{\varepsilon}_{mn}^{(l)} \stackrel{\text{df}}{=} v_n^{(ml)} - \Phi_m^{(l)}(\bar{v}_n^{(m-1)}, \theta_m), \quad l = 1, 2, \dots, L_m. \quad (4.236)$$

W wyniku zróżniczkowania wyrażenia (4.200) po poszczególnych parametrach, z uwzględnieniem modelu neuronu (4.206), otrzymujemy

$$\left. \frac{\partial \tilde{q}_n(\theta)}{\partial \theta_{mij}^{(s)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij}^{(s)}} = -2\tilde{\delta}_{mijn}^{(i)} \mu_{mj-1,n}^{(s)}, \quad (4.237)$$

gdzie: $\mu_{mj-1,n}^{(0)} = 1$, a $\mu_{mj-1,n}^{(s)}$, $s = 1, 2, \dots, I_{j_m-1}$ są wartościami wejść w j -tej warstwie, w n -tym takcie procedury uczenia,

$$\tilde{\varepsilon}_{mijn}^{(i)} \stackrel{\text{df}}{=} \begin{cases} \lambda_M \tilde{\varepsilon}_{Mn}^{(i)} & \text{dla } j = J_M, \\ \sum_{p=1}^{I_{j_m+1}} \left(\tilde{\delta}_{mj+1,n}^{(p)} \tilde{\theta}_{mpj+1,n}^{(i)} + \lambda_m \tilde{\varepsilon}_{mn}^{(i)} \right) & \text{dla } j = J_m, m = M-1, M-2, \dots, 1, \\ \sum_{p=1}^{I_{j_m+1}} \tilde{\delta}_{mj+1,n}^{(p)} \tilde{\theta}_{mpj+1,n}^{(i)} & \text{dla } j = J_m - 1, J_m - 2, \dots, 1, m = 1, 2, \dots, M, \end{cases} \quad (4.238)$$

$$\tilde{\delta}_{mijn}^{(i)} \stackrel{\text{df}}{=} \tilde{\varepsilon}_{mijn}^{(i)} \left. \frac{d\phi_{mij}(\kappa_{mij}^{(i)})}{d\kappa_{mij}^{(i)}} \right|_{\theta_{mij}^{(s)} = \tilde{\theta}_{mij}^{(s)}}, \quad (4.239)$$

oraz

$$\kappa_{mj}^{(i)} \stackrel{\text{df}}{=} \sum_{s=1}^{I_{j_m-1}} \mu_{mj-1}^{(s)} \theta_{mij}^{(s)} + \theta_{mij}^{(0)}.$$

Ostatecznie, po podstawieniu (4.238) do (4.237), globalnie optymalny algorytm wstecznej propagacji błędów przyjmuje postać

$$\tilde{\theta}_{mij,n+1}^{(s)} = \tilde{\theta}_{mij,n}^{(s)} + 2\eta_{mn} \delta_{mijn}^{(i)} \mu_{mj-1,n}^{(s)}, \quad (4.240)$$

$$s = 0, 1, \dots, I_{j_m-1}, \quad i = 1, 2, \dots, I_{J_m}, \quad j = 1, 2, \dots, J_m \quad m = 1, 2, \dots, M.$$

Zadanie wyznaczenia globalnie optymalnych parametrów sieci dla systemu złożonego z uwzględnieniem jakości modeli lokalnych przedstawiono w pracach [33, 36, 37, 124, 125, 129]. Problem identyfikacji złożonych systemów dynamicznych z wykorzystaniem sieci neuronowych podjęto także w publikacjach [38–42].

5. Identyfikacja dwustopniowa

Koncepcję identyfikacji wielostopniowej wprowadzono wstępnie w pracach [16, 64], a następnie rozwinięto w publikacjach [96, 100–102, 108–110].

W niniejszym rozdziale ograniczymy się do przedstawienia wybranych zadań identyfikacji dwustopniowej, wskazując na możliwość ich uogólnienia na przypadek wielu stopni.

Istota koncepcji dla dwóch stopni jest następująca: W przypadku pewnego obiektu (procesu, zjawiska) na pierwszym stopniu (rys. 5.1) badamy zależność wektora wyjść y od wektora wejść u_1 i wyznaczamy wektor parametrów θ_1 tej zależności. Jest to zadanie identyfikacji na pierwszym stopniu. Następnie zwracamy uwagę na pewien wektor u_2 , którego wartość w czasie identyfikacji na pierwszym stopniu była stała. Zmieniamy wartość tego wektora i powtarzamy identyfikację na stopniu pierwszym, uzyskując na ogół inną wartość θ_1 . Powtarzając wielokrotnie identyfikację na pierwszym stopniu dla różnych wartości u_2 , możemy zbadać zależność θ_1 od u_2 , a dokładniej – wyznaczyć wartość wektora parametrów θ_2 w opisie tej zależności. Jest to identyfikacja na drugim stopniu.

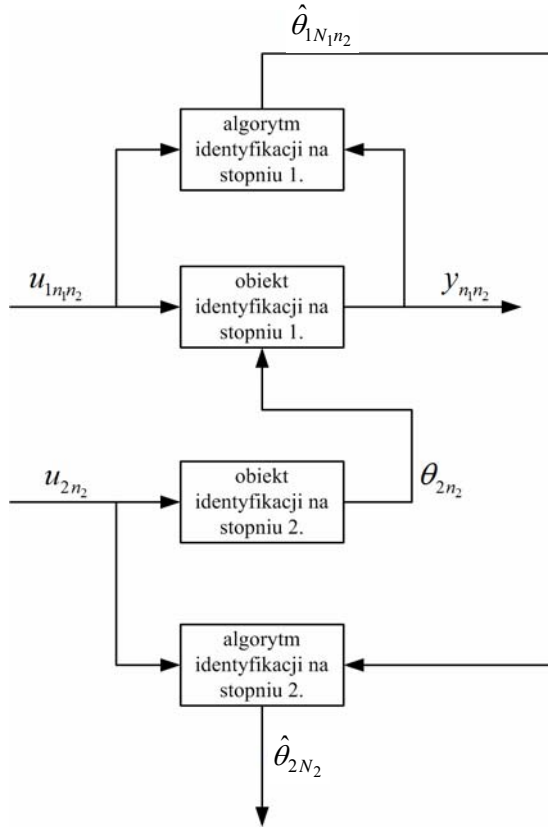
Implikuje to określony sposób organizacji eksperymentu (rys. 5.1), mianowicie: dla stałej wartości wejścia na drugim stopniu $u_2 = u_{2n_2}$ uzyskamy macierze wyników pomiarów wartości składowych wektora wejść oraz wyjść i na pierwszym stopniu

$$U_{1N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{11n_2} & u_{12n_2} & \cdots & u_{1N_1n_2} \end{bmatrix}, \quad Y_{N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} y_{1n_2} & y_{2n_2} & \cdots & y_{N_1n_2} \end{bmatrix}, \quad (5.1)$$

gdzie N_1 jest liczbą pomiarów na pierwszym stopniu.

Dla uproszczenia zapisu zakładamy, że w kolejnych pomiarach na drugim stopniu liczba ta jest stała. (Można przyjąć, że dla n_2 -tego pomiaru na drugim stopniu liczba pomiarów na pierwszym stopniu jest różna i wynosi N_{1n_2} , wówczas we wszystkich dalszych zależnościach w miejsce N_1 należy wstawić N_{1n_2}). Stosując odpowiedni algorytm identyfikacji dla pierwszego stopnia, wyznaczamy wektor parametrów charakterystyki na pierwszym stopniu θ_{1n_2} , odpowiadający stałemu wektorowi wejść na drugim stopniu $u_2 = u_{2n_2}$

$$\hat{\theta}_{1N_1n_2} = \Psi_{1N_1} \left(U_{1N_1n_2}, Y_{N_1n_2} \right). \quad (5.2)$$



Rys. 5.1. Koncepcja identyfikacji dwustopniowej

Zwróćmy uwagę, że wyjście na drugim stopniu nie jest mierzone, lecz jest wynikiem obliczeń, zgodnie z algorytmem identyfikacji na stopniu pierwszym. Powtarzanie zadania identyfikacji na pierwszym stopniu dla różnych wartości składowych wektora wejść na stopniu drugim, czyli $u_2 = u_{2n_2}$, $n_2 = 1, 2, \dots, N_2$, jest eksperymentem na drugim stopniu, gdzie N_2 jest liczbą pomiarów na drugim stopniu. Dla zadanej serii identyfikującej na drugim stopniu, tj.

$$U_{2N_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{21} & u_{22} & \cdots & u_{2N_2} \end{bmatrix}, \quad (5.3)$$

w wyniku rozwiązania zadania identyfikacji na pierwszym stopniu dla każdego elementu serii (5.3) wyznaczamy wektor parametrów na pierwszym stopniu. Znaczy to, że dla ustalonego $u_2 = u_{2n_2}$ wyznaczamy $\hat{\theta}_{1N_2}$ zgodnie z (5.2). Wartości te traktujemy jak pomiar wyjścia na drugim stopniu i dla zadanej serii identyfikującej (5.3) tworzą one macierz

$$\hat{\Xi}_{1N_1N_2} \stackrel{\text{df}}{=} \begin{bmatrix} \hat{\theta}_{1N_1} & \hat{\theta}_{1N_2} & \cdots & \hat{\theta}_{1N_1N_2} \end{bmatrix}. \quad (5.4)$$

Teraz na drugim stopniu, na podstawie serii identyfikującej (5.3) i macierzy wartości składowych wektora parametrów (5.4), wyznaczonych na pierwszym stopniu, wyznaczamy wartość składowych wektora parametrów θ_2 , korzystając z algorytmu identyfikacji

$$\hat{\theta}_{2N_2} = \Psi_{2N_2}(U_{2N_2}, \hat{\Sigma}_{1N_1N_2}). \quad (5.5)$$

W niektórych zadaniach praktycznych identyfikacja na pierwszym stopniu ma charakter pomocniczy, a celem jest wyznaczenie wektora parametrów θ_2 obiektu o wektorach wejść u_1 oraz u_2 i wyjść y , którego opis jest złożeniem opisów z pierwszego i drugiego stopnia, a wyniki eksperymentu są zebrane w macierzach (5.1), dla $n_2 = 1, 2, \dots, N_2$, i (5.3).

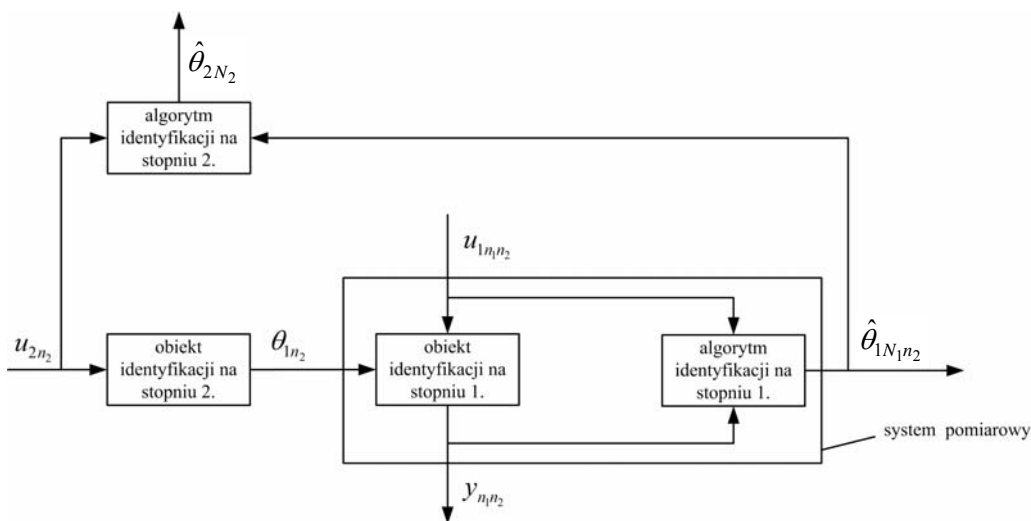
Zastosowanie koncepcji dwustopniowej, czyli rozróżnianie dwóch stopni, oznacza pewnego rodzaju dekompozycję zadania identyfikacji obiektu o wyróżnionych wektorach wejść u_1 i u_2 oraz wektora wyjść y . Można wymienić powody takiego postępowania, które są często ze sobą związane i mogą występować łącznie, uzasadniając stosowanie podejścia dwustopniowego.

Pierwszy z nich dotyczy przypadku, gdy badamy oddzielne realne obiekty o tej samej naturze (urządzenia, procesy), charakteryzujące się stałymi u_2 . Jest to częsty przypadek w badaniach doświadczalnych. Poszczególne obiekty mogą być badane w różnych laboratoriach i w każdym z nich badana jest zależność pomiędzy tymi samymi wejściami u_1 i wyjściami y , lecz obiekty te różnią się stałymi u_2 . Przykładem zastosowania takiego podejścia jest wspomniane w rozdziale 1. badanie destylacyjnej kolumny wypełnionej z pulsacją fazy parowej. Identyfikacja na pierwszym stopniu może dotyczyć badania kolumny z konkretnym ustalonym wypełnieniem. Drugi stopień natomiast to uwzględnienie parametrów wypełnienia w opisie. Zwróćmy uwagę, że zmiana wypełnienia to przebudowa kolumny. Podobnie w badaniach populacyjnych można ustalić zależność pomiędzy wejściami u_1 i wyjściami y dla różnych osobników. Celem identyfikacji na drugim stopniu jest uwzględnienie w ostatecznym opisie wyników uzyskanych z różnych laboratoriów lub badań przeprowadzonych dla pewnej populacji. Ogólnie omawiany przypadek można nazwać dekompozycją przestrzenną, w której celem identyfikacji na drugim stopniu jest zbadanie zbioru realnych obiektów identyfikowanych na pierwszym stopniu.

Kolejny przypadek dotyczy organizacji eksperymentu na obu stopniach. Równoczesna zmiana u_1 oraz u_2 jest niewygodna lub niemożliwa nawet wtedy, gdy oba stopnie dotyczą tego samego obiektu. Na pierwszym stopniu badany jest obiekt, a w trakcie eksperymentu utrzymywana jest stała wartość drugiego wejścia. W kolejnych etapach badań powtarzany jest eksperyment na pierwszym stopniu dla różnych wartości drugiego wejścia. Przykładem takiego dwustopniowego zadania jest badanie wpływu zmian temperatury na wartość parametrów obiektu. W takim przypadku na pierwszym stopniu badamy obiekt w ustalonej temperaturze i wyznaczamy parametry obiektu jej odpowiadające. Następnie zmieniamy temperaturę i powtarzamy zadanie identyfikacji na pierwszym stopniu, w wyniku którego otrzymamy parametry na ogół inne, odpowiadające tej temperaturze. Powtarzając zadanie identyfikacji na pierwszym stopniu dla zadanej serii warto-

ści temperatury, otrzymamy odpowiednią serię wartości parametrów. Teraz możemy badać wpływ temperatury na zmianę parametrów obiektu na pierwszym stopniu. Jest to zadanie identyfikacji na drugim stopniu. Przypadek ten można określić jako dekompozycję czasową, w której z konieczności, ze względu na różny charakter wejść lub dla wygody, eksperyment zorganizowany jest tak, że utrzymywana jest stała wartość u_2 w kolejnych jego etapach, w których zmieniają się u_1 i y .

Następny przypadek wyodrębnienia pierwszego stopnia jest związany z sytuacją, gdy pewne wielkości θ_1 nie mogą być mierzone bezpośrednio, ale można je wyznaczyć jako parametry innej zależności. Podstawowym celem identyfikacji jest więc zbadanie zależności pomiędzy θ_1 a u_2 , a identyfikacja na pierwszym stopniu odgrywa rolę pomocniczą jako pewnego rodzaju układ pomiarowy, w którym mierzy się pomocnicze wielkości u_1 oraz y i na tej podstawie oblicza się niedostępną do pomiaru wielkość θ_1 .



Rys. 5.2. Pierwszy stopień jako system pomiarowy

Wyodrębnienie pierwszego stopnia może być także podyktowane względami obliczeniowymi. Dekompozycja bezpośredniego zadania identyfikacji i wyodrębnienie pierwszego stopnia może się wiązać z rozproszonym przetwarzaniem wyników eksperymentu na pierwszym stopniu, a następnie ich koordynacją na drugim stopniu. W tym przypadku aktualny jest problem porównania podejścia bezpośredniego z podejściem dwustopniowym.

Przedstawioną koncepcję identyfikacji dwustopniowej łatwo jest uogólnić na przypadek identyfikacji wielostopniowej, dodając w kolejnych stopniach badanie zależności pomiędzy parametrami opisu poprzedniego stopnia a nowym, kolejnym wejściem. Rozważmy obiekt statyczny o M stopniach (rys. 1.19), opisany zależnościami:

$$\begin{aligned}
y &= F_1(u_1, \theta_1), \\
\theta_1 &= F_2(u_2, \theta_2), \\
\theta_2 &= F_3(u_3, \theta_3), \\
&\vdots \\
\theta_{M-1} &= F_M(u_M, \theta_M),
\end{aligned} \tag{5.6}$$

czyli na m -tym stopniu

$$\theta_{m-1} = F_m(u_m, \theta_m), \text{ dla } m=1, 2, \dots, M, \theta_0 = y. \tag{5.7}$$

Jeden pomiar do identyfikacji na $m + 1$. stopniu wymaga określonej serii pomiarów na potrzeby identyfikacji na m -tym stopniu, w czasie której mierzymy (lub nastawiamy) kolejne wartości u_m i obliczamy odpowiadający im wektor parametrów θ_{m-1} na stopniu $m - 1$. Jedynie dla pierwszego stopnia wyjście, tj. y , mierzymy bezpośrednio. Jeśli oznaczymy długość serii na m -tym stopniu przez N_m , to jeden pomiar do identyfikacji na $m + 1$. stopniu wymaga jednej serii, czyli N_m pomiarów na m -tym stopniu, co z kolei wymaga N_{m-1} pomiarów, czyli $N_m N_{m-1}$ pomiarów na $m - 1$. stopniu itd. Wymagana liczba pomiarów na pierwszym stopniu jest zatem równa $N_1 N_2 \dots N_M$. Ma to określone implikacje dla formy i organizacji zbioru danych.

Po wprowadzeniu oznaczeń poszczególnych pomiarów możemy dla m -tego stopnia i k -tego pomiaru na $m + 1$. stopniu wprowadzić oznaczenia serii wartości wejść i wyjść w postaci macierzy, podobnie jak (5.3) i (5.4):

$$[u_{m1,k} \quad u_{m2,k} \quad \dots \quad u_{mN_m,k}] \stackrel{\text{df}}{=} U_{mN_m,k}, \tag{5.8}$$

$$[\hat{\theta}_{m-1,1,k} \quad \hat{\theta}_{m-1,2,k} \quad \dots \quad \hat{\theta}_{m-1,N_m,k}] \stackrel{\text{df}}{=} \hat{\Sigma}_{m-1,N_m,k}, \tag{5.9}$$

gdzie $\hat{\theta}_{m-1,n,k}$ oznacza wynik identyfikacji na $m - 1$. stopniu dla wartości $u_{mn,k}$.

Dla uproszczenia pominięto tu indeks wskazujący na to, że chodzi o wyznaczenie wartości θ_{m-1} dla N_{m-1} pomiarów. Dokładniej powinno być $\hat{\theta}_{m-1,N_{m-1},n,k}$. Dla m -tego stopnia

$$\hat{\theta}_{m,N_m,k} = \Psi_{m,N_m}(U_{mN_m,k}; \hat{\Sigma}_{m-1,N_m,k}), \tag{5.10}$$

gdzie: $\hat{\theta}_{m,N_m,k}$ jest oszacowaniem wektora parametrów $\theta_{m,k}$ dla stałej wartości $u_{m+1,k}$ na $m + 1$. stopniu, a Ψ_{m,N_m} oznacza algorytm identyfikacji.

W dalszej części ograniczymy się do dwustopniowego zadania estymacji parametrów i wyboru optymalnego modelu w warunkach losowych.

5.1. Dwustopniowe algorytmy estymacji parametrów obiektu

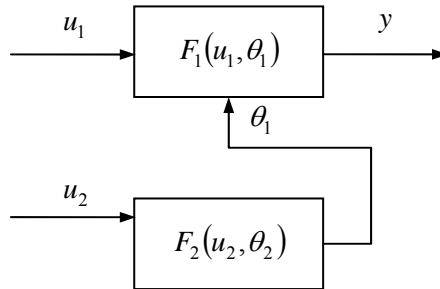
Rozważmy dwustopniowy obiekt statyczny (rys. 5.3), dla którego charakterystyka na pierwszym i drugim stopniu dana jest zależnościami:

$$y = F_1(u_1, \theta_1), \quad (5.11)$$

$$\theta_1 = F_2(u_2, \theta_2), \quad (5.12)$$

w których: $u_1 \in U_1 \subseteq \mathbb{R}^{S_1}$ jest wektorem wejść na pierwszym stopniu o S_1 składowych; $u_2 \in U_2 \subseteq \mathbb{R}^{S_2}$ jest wektorem wejść na drugim stopniu o S_2 składowych; $\theta_1 \in \Theta_1 \subseteq \mathbb{R}^{R_1}$ jest R_1 -wymiarowym wektorem parametrów charakterystyki statycznej na pierwszym stopniu; $\theta_2 \in \Theta_2 \subseteq \mathbb{R}^{R_2}$ jest R_2 -wymiarowym wektorem parametrów charakterystyki statycznej na drugim stopniu; $y \in Y \subseteq \mathbb{R}^L$ jest L -wymiarowym wektorem wyjść na pierwszym stopniu; F_1, F_2 są znanymi funkcjami, takimi że:

$$F_1 : U_1 \times \Theta_1 \rightarrow Y, \quad F_2 : U_2 \times \Theta_2 \rightarrow \Theta_1.$$



Rys. 5.3. Obiekt statyczny o dwóch stopniach

Funkcje F_1 i F_2 są znane. Zadanie identyfikacji na pierwszym stopniu sprowadza się zatem do wyznaczenia oszacowania wektora parametrów θ_1 na podstawie wyników zakłóconych pomiarów wartości składowych wektora wyjść dla zadanej serii wartości składowych wektora wejść u_1 . Na drugim stopniu należy wyznaczyć oszacowanie wektora parametrów θ_2 na podstawie wyników estymacji na pierwszym stopniu dla zadanej serii wartości składowych wektora wejść u_2 .

Zauważmy, że wartość składowych wektora wyjść dla drugiego stopnia, tj. θ_1 , nie jest mierzona bezpośrednio, lecz jest wynikiem estymacji na pierwszym stopniu. Jeden pomiar do identyfikacji na drugim stopniu wymaga całego procesu estymacji, czyli określonej serii pomiarów na pierwszym stopniu, w czasie której u_2 jest stałe. Do oznaczenia wyników eksperymentu wprowadzimy następującą notację:

u_{2n_2}, θ_{1n_2} są – odpowiednio – wartościami składowych wektorów u_2 i θ_1 w n_2 -tym pomiarze na drugim stopniu, $n_2 = 1, 2, \dots, N_2$,

$u_{1n_1n_2}, y_{n_1n_2}$ są – odpowiednio – wartościami składowych wektorów u_1 oraz y w n_1 -tym pomiarze na pierwszym stopniu dla stałej wartości wektora u_{2n_2} , $n_1 = 1, 2, \dots, N_1$,

$$U_{1N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{11n_2} & u_{12n_2} & \cdots & u_{1N_1n_2} \end{bmatrix} \quad (5.13)$$

jest serią identyfikującą na pierwszym stopniu w n_2 -tym pomiarze na drugim stopniu, tj. dla $u_2 = u_{2n_2}$,

$$w_{n_1n_2} = h(y_{n_1n_2}, z_{n_1n_2}) \quad (5.14)$$

jest wynikiem pomiaru wartości składowych wektora wyjść $y_{n_1n_2}$ z zakłóceniem $z_{n_1n_2} \in Z \subseteq \mathbf{R}^L$, h jest znaną funkcją (2.20), mającą własność (2.23),

$$W_{N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} w_{1n_2} & w_{2n_2} & \cdots & w_{N_1n_2} \end{bmatrix} \quad (5.15)$$

jest macierzą pomiarów na pierwszym stopniu w n_2 -tym pomiarze na drugim stopniu,

$$\bar{U}_{1N_1N_2} \stackrel{\text{df}}{=} \begin{bmatrix} U_{1N_11} & U_{1N_12} & \cdots & U_{1N_1N_2} \end{bmatrix}, \quad (5.16)$$

$$\bar{W}_{N_1N_2} \stackrel{\text{df}}{=} \begin{bmatrix} W_{N_11} & W_{N_12} & \cdots & W_{N_1N_2} \end{bmatrix}, \quad (5.17)$$

$$U_{2N_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{21} & u_{22} & \cdots & u_{2N_2} \end{bmatrix} \quad (5.18)$$

są macierzami wyników pomiarów dwustopniowego eksperymentu.

Dla poszczególnych pomiarów zależności (5.11) i (5.12) przyjmują postać

$$y_{n_1n_2} = F_1(u_{1n_1n_2}, \theta_{1n_1}), \quad (5.19)$$

$$\theta_{1n_1} = F_2(u_{2n_2}, \theta_2). \quad (5.20)$$

W dalszych rozważaniach założymy, że $z_{n_1n_2}$ jest realizacją ciągłej zmiennej losowej $z_{n_1n_2}$, przy czym zmienne losowe $z_{n_1n_2}$ są niezależne dla kolejnych pomiarów i mają tę samą gęstość rozkładu prawdopodobieństwa $f_z(z)$. Na pierwszym stopniu dla ustalonego $u_2 = u_{2n_2}$ na podstawie serii pomiarów (5.13) oraz (5.15) wyznaczamy

$$\hat{\theta}_{1N_1n_2} = \Psi_{1N_1}(U_{1N_1n_2}, W_{N_1n_2}), \quad (5.21)$$

gdzie: $\hat{\theta}_{1N_1n_2}$ jest oszacowaniem nieznanego wektora parametrów θ_{1n_2} , a Ψ_{1N_1} oznacza algorytm estymacji na pierwszym stopniu.

Estymację na pierwszym stopniu zilustrowano na rysunku 5.4.

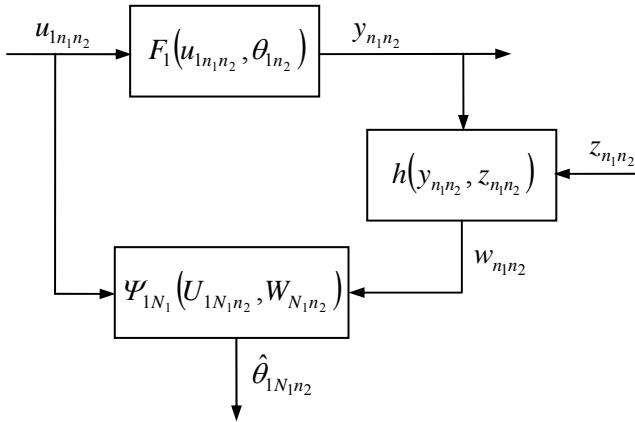
Na drugim stopniu, dla zadanej serii (5.18) wyznaczamy

$$\hat{\theta}_{2N_2} = \Psi_{2N_2}(U_{2N_2}, \hat{\Xi}_{1N_1N_2}), \quad (5.22)$$

gdzie: $\hat{\theta}_{2N_2}$ jest oszacowaniem nieznanego wektora parametrów θ_2 , Ψ_{2N_2} oznacza algorytm estymacji na drugim stopniu, a

$$\hat{\Xi}_{1N_1N_2} \stackrel{\text{df}}{=} \begin{bmatrix} \hat{\theta}_{1N_11} & \hat{\theta}_{1N_12} & \cdots & \hat{\theta}_{1N_1N_2} \end{bmatrix} \quad (5.23)$$

jest macierzą wyników estymacji na pierwszym stopniu dla $n_2 = 1, 2, \dots, N_2$.



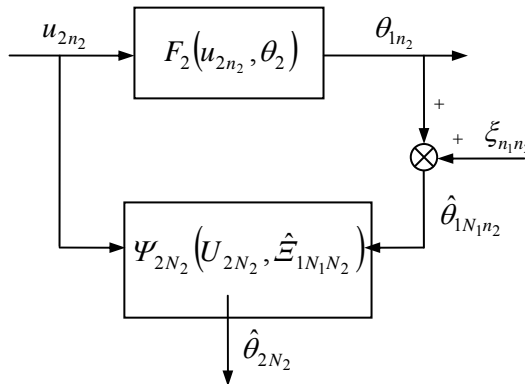
Rys. 5.4. Estymacja parametrów na 1. stopniu

Oszacowanie $\hat{\theta}_{1N_1n_2}$ jest traktowane jak wynik pomiaru wartości składowych wektora wyjścia w n_2 -tym kroku eksperymentu na drugim stopniu, z błędem estymacji

$$\xi_{n_2} = \hat{\theta}_{1N_1n_2} - \theta_{1n_2}, \quad (5.24)$$

analogicznym do zakłócenia $z_{n_1n_2}$ na pierwszym stopniu.

Estymację na drugim stopniu zilustrowano na rysunku 5.5.



Rys. 5.5. Estymacja parametrów na 2. stopniu

Wektor parametrów θ_2 może być też estymowany bezpośrednio po przyjęciu opisu

$$y = F_1[u_1, F_2(u_2, \theta_2)] \stackrel{\text{df}}{=} F(u_1, u_2, \theta_2), \quad (5.25)$$

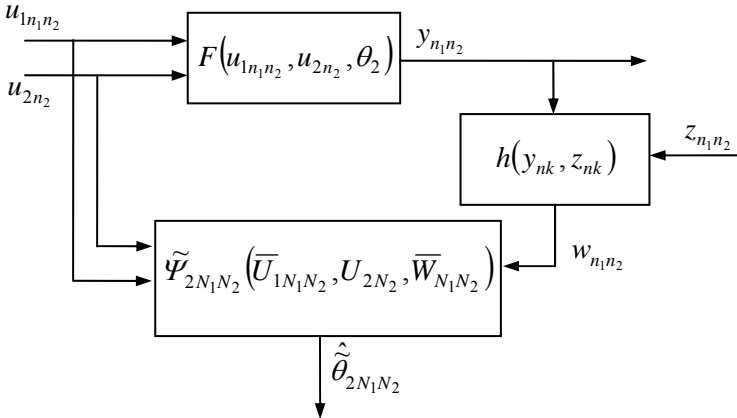
czyli

$$y_{n_1 n_2} = F(u_{1n_1 n_2}, u_{2n_2}, \theta_2), \quad (5.26)$$

z wykorzystaniem wyników pomiarów wartości u_1 , u_2 , y . Wynik estymacji oznaczmy przez

$$\hat{\theta}_{2N_1 N_2} = \tilde{\Psi}_{2N_1 N_2}(\bar{U}_{1N_1 N_2}, U_{2N_2}, \bar{W}_{N_1 N_2}), \quad (5.27)$$

gdzie: $\hat{\theta}_{2N_1 N_2}$ jest oszacowaniem wektora parametrów θ_2 przy podejściu bezpośrednim, a $\tilde{\Psi}_{2N_1 N_2}$ oznacza algorytm estymacji bezpośredniej (rys. 5.6).



Rys. 5.6. Bezpośrednia estymacja parametrów

W dalszych punktach przedstawimy metodykę postępowania z zastosowaniem dwóch metod estymacji, przedstawionych w punkcie 2.2.1, tj. metody maksymalnej wiarygodności oraz podejścia bayesowskiego. Dla typowych przypadków podane zostaną algorytmy estymacji.

Istotne jest także porównanie wyniku estymacji bezpośredniej z wynikiem estymacji dwustopniowej, którą można traktować jako zastosowanie pewnego rodzaju dekompozycji, czyli porównanie wartości $\hat{\theta}_{2N_2}$ (5.22) i $\hat{\theta}_{2N_1 N_2}$ (5.27) za pomocą tej samej metody estymacji. Algorytm $\tilde{\Psi}_{2N_1 N_2}$ należy porównać z kompozycją algorytmów (5.21) i (5.22). Po wprowadzeniu oznaczeń

$$\hat{\Xi}_{1N_1 N_2} = [\Psi_{1N_1}(U_{1N_1}, W_{N_1}) \cdots \Psi_{1N_1}(U_{1N_1 N_2}, W_{N_1 N_2})] \stackrel{\text{df}}{=} \bar{\Psi}_{1N_1 N_2}(\bar{U}_{1N_1 N_2}, \bar{W}_{N_1 N_2}) \quad (5.28)$$

i wstawieniu (5.28) do (5.22) otrzymujemy

$$\hat{\theta}_{2N_2} = \Psi_{2N_2} \left[U_{2N_2}, \bar{\Psi}_{1N_1N_2} \left(\bar{U}_{1N_1N_2}, \bar{W}_{N_1N_2} \right) \right] \stackrel{\text{df}}{=} \tilde{\Psi}_{2N_1N_2} \left(\bar{U}_{1N_1N_2}, U_{2N_2}, \bar{W}_{N_1N_2} \right). \quad (5.29)$$

Podjęcia bezpośrednie i dwustopniowe są równoważne, jeśli algorytmy (5.27) i (5.29) są identyczne, tzn. dla tych samych wyników pomiarów otrzymujemy identyczne oszacowania $\hat{\theta}_{2N_1N_2}$ i $\hat{\theta}_{2N_2}$.

5.1.1. Metoda maksymalnej wiarygodności

W wyniku zastosowania metody maksymalnej wiarygodności należy na obu stopniach wyznaczyć i maksymalizować funkcje wiarygodności. Postępowanie na pierwszym stopniu sprowadza się do znanej metody estymacji parametrów obiektu statycznego, przedstawionej w punkcie 2.2.1. Zadanie dla drugiego stopnia jest nieco bardziej złożone, wymaga bowiem wyznaczenia funkcji gęstości rozkładu prawdopodobieństwa nie wyniku obserwacji, lecz wyniku estymacji na pierwszym stopniu dla kolejnych pomiarów na drugim stopniu. Procedura wyznaczania algorytmów estymacji jest zatem następująca:

Krok 1. Dla pierwszego stopnia, gdy $u_2 = u_{2n_2}$, wyznaczamy funkcję wiarygodności

$$L_{1N_1} \left(W_{N_1n_2}, \theta_{1n_2}; U_{1N_1n_2} \right) = \prod_{n_1=1}^{N_1} f_{wn_2} \left(w_{n_1n_2}, \theta_{1n_2}; u_{1n_1n_2} \right), \quad (5.30)$$

gdzie funkcja gęstości rozkładu prawdopodobieństwa f_{wn_2} – obserwowanej zmiennej losowej $w_{n_1n_2}$ z wektorem parametrów θ_{1n_2} i zadaniem $u_{n_1n_2}$ – dana jest zależnością

$$f_{wn_2} \left(w_{n_1n_2}, \theta_{1n_2}; u_{1n_1n_2} \right) = f_z \left(h_z^{-1} \left(F_1 \left(u_{1n_1n_2}, \theta_{1n_2} \right), w_{n_1n_2} \right) \right) |J_h|, \quad (5.31)$$

w której: h_z^{-1} jest funkcją odwrotną do (5.14) względem $z_{n_1n_2}$, $|J_h|$ jest jacobianem przekształcenia h_z^{-1} z argumentem $w_{n_1n_2}$, przy założeniu, że przekształcenie h jest ciągle i wzajemnie jednoznaczne.

Krok 2. Maksymalizujemy (5.30) względem θ_{1n_2} i otrzymujemy algorytm estymacji (5.21), gdzie $\hat{\theta}_{1N_1n_2}$ oznacza wynik maksymalizacji.

Krok 3. Na podstawie zależności (5.21), po podstawieniu (5.20) do (5.14), wyznaczamy funkcję gęstości rozkładu prawdopodobieństwa

$$f_{\theta_1} \left(\hat{\theta}_{1N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2} \right) \quad (5.32)$$

dla zmiennej losowej $\hat{\theta}_{1N_1n_2}$. Gęstość tę możemy też wyznaczyć z zależności (5.21), po podstawieniu (5.26) do (5.14) oraz skorzystaniu ze znajomości $f_z(z)$.

Krok 4. Dla drugiego stopnia wyznaczamy funkcję wiarygodności

$$L_{2N_2}(\hat{\Xi}_{1N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) = \prod_{n_1=1}^{N_1} f_{\theta_1}(\hat{\theta}_{1N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2}), \quad (5.33)$$

bowiem z niezależności $z_{n_1n_2}$ dla kolejnych pomiarów wynika niezależność $\theta_{1N_1n_2}$ dla różnych n_2 .

Krok 5. Maksymalizujemy (5.33) względem θ_2 i otrzymujemy algorytm estymacji (5.22), gdzie $\hat{\theta}_{2N_2}$ oznacza wynik maksymalizacji.

W **podjęciu bezpośrednim** funkcja wiarygodności ma postać

$$L_{N_2N_1}(\bar{W}_{N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) = \prod_{n_2=1}^{N_2} \prod_{n_1=1}^{N_1} f_w(w_{n_1n_2}, \theta_2; u_{1n_1n_2}, u_{2n_2}), \quad (5.34)$$

przy czym funkcję gęstości rozkładu prawdopodobieństwa f_w – obserwowanej zmiennej losowej $w_{n_1n_2}$ z wektorem parametrów θ_2 i zadanymi $u_{1n_1n_2}$ oraz u_{2n_2} – otrzymamy po podstawieniu w zależności (5.31) w miejsce θ_{1n_2} zależności (5.20), czyli

$$f_w(w_{n_1n_2}, \theta_2; u_{1n_1n_2}, u_{2n_2}) = f_z(h_z^{-1}(F_1(u_{1n_1n_2}, F_2(u_{2n_2}, \theta_2)), w_{n_1n_2})) |J_h|, \quad (5.35)$$

a po uwzględnieniu (5.25) przyjmuje ona postać

$$f_w(w_{n_1n_2}, \theta_2; u_{1n_1n_2}, u_{2n_2}) = f_z(h_z^{-1}(F(u_{1n_1n_2}, u_{2n_2}, \theta_2), w_{n_1n_2})) |J_h|. \quad (5.36)$$

W wyniku maksymalizacji funkcji wiarygodności (5.34) względem θ_2 otrzymujemy algorytm estymacji bezpośredniej (5.27), gdzie $\hat{\theta}_{2N_1N_2}$ oznacza wynik maksymalizacji.

Rozpatrzmy w charakterze przykładów typowe przypadki, kiedy można analitycznie wyznaczyć algorytmy estymacji.

Przykład 5.1. Niech $S_1 = L = R_1 = 1$, $S_2 = R_2$, zależności (5.19), (5.20) i (5.14) są – odpowiednio – postaci:

$$y_{n_1n_2} = \theta_{1n_2} u_{1n_1n_2}, \quad (5.37)$$

$$\theta_{1n_2} = \theta_2^T u_{2n_2}, \quad (5.38)$$

$$w_{n_1n_2} = y_{n_1n_2} + z_{n_1n_2}. \quad (5.39)$$

Ostatecznie zależność (5.26) przyjmuje postać

$$y_{n_1 n_2} = \theta_2^T u_{2n_2} u_{1n_1 n_2}, \quad (5.40)$$

a zatem

$$w_{n_1 n_2} = \theta_2^T u_{2n_2} u_{1n_1 n_2} + z_{n_1 n_2}. \quad (5.41)$$

Niech zakłócenie ma rozkład normalny

$$f_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left(\frac{-z^2}{2\sigma_z^2}\right). \quad (5.42)$$

Łatwo w tym przypadku (tak jak w przykładzie 2.3) otrzymać analityczny wynik dla pierwszego stopnia. Funkcja gęstości rozkładu prawdopodobieństwa (5.31) ma postać

$$f_{w_{n_2}}(w_{n_1 n_2}, \theta_{n_2}; u_{1n_1 n_2}) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left[-\frac{(w_{n_1 n_2} - \theta_{n_2} u_{1n_1 n_2})^2}{2\sigma_z^2}\right], \quad (5.43)$$

a funkcja wiarygodności (5.30)

$$L_{1N_1}(W_{N_1 n_2}, \theta_{n_2}; U_{1N_1 n_2}) = \frac{1}{(\sigma_z \sqrt{2\pi})^{N_1}} \exp\left[-\frac{1}{2\sigma_z^2} \sum_{n_1=1}^{N_1} (w_{n_1 n_2} - \theta_{n_2} u_{1n_1 n_2})^2\right]. \quad (5.44)$$

Maksymalizacja funkcji (5.44) względem θ_{n_2} daje algorytm estymacji na pierwszym stopniu (5.21) postaci

$$\hat{\theta}_{1N_1 n_2} = \frac{\sum_{n_1=1}^{N_1} u_{1n_1 n_2} w_{n_1 n_2}}{\sum_{n_1=1}^{N_1} u_{1n_1 n_2}^2} = \frac{U_{1N_1 n_2} W_{N_1 n_2}^T}{U_{1N_1 n_2} U_{N_1 n_2}^T}. \quad (5.45)$$

Po wstawieniu zależności (5.41) do (5.45) otrzymujemy

$$\hat{\theta}_{1N_1 n_2} = \theta_2^T u_{2n_2} + \frac{\sum_{n_1=1}^{N_1} u_{1n_1 n_2} z_{n_1 n_2}}{\sum_{n_1=1}^{N_1} u_{1n_1 n_2}^2}. \quad (5.46)$$

Zmienna losowa $\hat{\theta}_{1N_1 n_2}$, która jest wynikiem przekształcenia (5.46) zmiennej losowej z , ma również rozkład normalny z wartością oczekiwaną $E(\hat{\theta}_{1N_1 n_2}) = \theta_2^T u_{2n_2}$ i wariancją

$$\text{Var}(\hat{\theta}_{1N_1n_2}) = \frac{\sigma_z^2 \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2}{\left(\sum_{n_1=1}^{N_1} u_{1n_1n_2}^2 \right)^2} = \frac{\sigma_z^2}{\sum_{n_1=1}^{N_1} u_{1n_1n_2}^2} \stackrel{\text{df}}{=} \sigma_{n_2}^2. \quad (5.47)$$

Funkcja gęstości rozkładu prawdopodobieństwa (5.32) zmiennej losowej $\hat{\theta}_{1N_1n_2}$ jest zatem następująca

$$f_{\theta_1}(\hat{\theta}_{1N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2}) = \frac{1}{\sigma_{n_2} \sqrt{2\pi}} \exp \left[-\frac{(\theta_{1n_2} - \theta_2^T u_{2n_2})^2}{2\sigma_{n_2}^2} \right], \quad (5.48)$$

a po uwzględnieniu (5.47)

$$f_{\theta_1}(\hat{\theta}_{1N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2}) = \frac{\sqrt{\sum_{n_1=1}^{N_1} u_{1n_1n_2}^2}}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{1}{2\sigma_z^2} (\theta_{1n_2} - \theta_2^T u_{2n_2})^2 \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2 \right]. \quad (5.49)$$

W konsekwencji funkcja wiarygodności (5.33) ma postać

$$L_{2N_2}(\hat{\Xi}_{1N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) \quad (5.50)$$

$$= \frac{\prod_{n_2=1}^{N_2} \sqrt{\sum_{n_1=1}^{N_1} u_{1n_1n_2}^2}}{(\sigma_z \sqrt{2\pi})^{N_2}} \exp \left[-\frac{\sum_{n_2=1}^{N_2} (\theta_{1n_2} - \theta_2^T u_{2n_2})^2 \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2}{2\sigma_z^2} \right].$$

W wyniku maksymalizacji funkcji wiarygodności (5.50) względem θ_2 otrzymujemy algorytm estymacji (5.22) na drugim stopniu postaci

$$\hat{\theta}_{2N_2} = \left[\sum_{n_2=1}^{N_2} u_{2n_2} u_{2n_2}^T \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2 \right]^{-1} \sum_{k=1}^K \hat{\theta}_{1N_1n_2} u_{2n_2} \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2, \quad (5.51)$$

czyli

$$\hat{\theta}_{2N_2} = (U_{2N_2} D_{N_2} U_{2N_2}^T)^{-1} U_{2N_2} D_{N_2} \hat{\Xi}_{1N_1N_2}^T, \quad (5.52)$$

gdzie D_{N_2} jest macierzą diagonalną

$$\text{diag } D_{N_2} \stackrel{\text{df}}{=} [U_{1N_11} U_{1N_11}^T, U_{1N_12} U_{1N_12}^T, \dots, U_{1N_1N_2} U_{1N_1N_2}^T].$$

Macierz D_{N_2} występuje jako parametr algorytmu (5.22) postaci (5.52), bowiem wariancja $\sigma_{n_2}^2$, w odróżnieniu od stałej wariancji σ_z^2 , na pierwszym stopniu zależy od n_2 , a dokładniej od $U_{1N_1n_2}$.

W rozpatrywanym przykładzie w **podejściu bezpośrednim** funkcja gęstości rozkładu prawdopodobieństwa (5.35) ma postać

$$f_w(w_{n_1n_2}, \theta_2; u_{1n_1n_2}, u_{2n_2}) = \frac{1}{\sigma_z \sqrt{2\pi}} \left[-\frac{(w_{n_1n_2} - \theta_2^T u_{n_2} u_{n_1n_2})^2}{2\sigma_z^2} \right], \quad (5.53)$$

a funkcja wiarygodności (5.34)

$$\begin{aligned} & L_{N_2N_1}(\bar{W}_{N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) \\ &= \frac{1}{(\sigma_z \sqrt{2\pi})^{N_1N_2}} \left[-\frac{\sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} (w_{n_1n_2} - \theta_2^T u_{n_2} u_{n_1n_2})^2}{2\sigma_z^2} \right]. \end{aligned} \quad (5.54)$$

W wyniku maksymalizacji funkcji wiarygodności (5.54) względem θ_2 otrzymujemy algorytm estymacji bezpośredniej (5.27) postaci

$$\hat{\theta}_{2N_1N_2} = \left[\sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{2n_2} u_{2n_2}^T u_{1n_1n_2}^2 \right]^{-1} \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{2n_2} u_{1n_1n_2} w_{n_1n_2}, \quad (5.55)$$

czyli

$$\hat{\theta}_{2N_1N_2} = (U_{2N_2} D_{N_2} U_{2N_2}^T)^{-1} U_{2N_2} B_{N_2}, \quad (5.56)$$

gdzie B_{N_2} jest wektorem

$$B_K^T \stackrel{\text{df}}{=} [U_{1N_11} W_{N_11}^T \quad U_{1N_22} W_{N_22}^T \quad \cdots \quad U_{1N_KK} W_{N_KK}^T]. \quad (5.57)$$

Łatwo zauważyć, że po wstawieniu (5.45) do (5.51) lub (5.52) w macierzy (5.23) i po prostych przekształceniach otrzymujemy $\hat{\theta}_{2N_2} = \hat{\theta}_{2N_1N_2}$, a zatem w rozpatrywanym przypadku dwustopniowa i bezpośrednia estymacja θ_2 dają ten sam wynik. ◆

Przykład 5.2. Rozpatrzmy ponownie przypadek z przykładu 5.1 dla jednowymiarowej zależności na drugim stopniu, tzn. $S_1 = S_2 = L = R_2 = 1$, czyli zależności (5.19) i (5.20) przyjmują – odpowiednio – postać:

$$y_{n_1 n_2} = \theta_{1n_2} u_{1n_1 n_2}, \quad (5.58)$$

$$\theta_{1n_2} = \theta_2 u_{2n_2}. \quad (5.59)$$

Załóżmy ponadto, że zakłócenia są multiplikatywne, tj. zależność (5.14) jest postaci

$$w_{n_1 n_2} = y_{n_1 n_2} z_{n_1 n_2}. \quad (5.60)$$

Ostatecznie zależność (5.26) przyjmuje postać

$$y_{n_1 n_2} = \theta_2 u_{2n_2} u_{1n_1 n_2}, \quad (5.61)$$

a zatem

$$w_{n_1 n_2} = \theta_2 u_{2n_2} u_{1n_1 n_2} z_{n_1 n_2}. \quad (5.62)$$

Załóżmy równomierny rozkład prawdopodobieństwa zakłóceń w przedziale $[\alpha, \beta]$, $\beta > \alpha > 0$, co oznacza, że funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej z ma postać

$$f_z(z) = \begin{cases} \frac{1}{\beta - \alpha} & \text{dla } z \in [\alpha, \beta] \\ 0 & \text{dla } z \notin [\alpha, \beta], \end{cases} \quad (5.63)$$

Obecnie (tak jak w przykładzie 2.4) funkcja gęstości rozkładu prawdopodobieństwa (5.31) ma postać

$$f_{w_{n_2}}(w_{n_1 n_2}, \theta_{1n_2}; u_{1n_1 n_2}) = \begin{cases} \frac{1}{(\beta - \alpha) \theta_{1n_2} u_{1n_1 n_2}} & \text{dla } \frac{w_{n_1 n_2}}{u_{1n_1 n_2}} \in [\alpha \theta_{1n_2}, \beta \theta_{1n_2}] \\ 0 & \text{dla } \frac{w_{n_1 n_2}}{u_{1n_1 n_2}} \notin [\alpha \theta_{1n_2}, \beta \theta_{1n_2}], \end{cases} \quad (5.64)$$

a po przekształceniu

$$\begin{aligned} & f_{w_{n_2}}(w_{n_1 n_2}, \theta_{1n_2}; u_{1n_1 n_2}) \\ &= \begin{cases} \frac{1}{(\beta - \alpha) \theta_{1n_2} u_{1n_1 n_2}} & \text{dla } w_{n_1 n_2} \in [\alpha \theta_{1n_2} u_{1n_1 n_2}, \beta \theta_{1n_2} u_{1n_1 n_2}] \\ 0 & \text{dla } w_{n_1 n_2} \notin [\alpha \theta_{1n_2} u_{1n_1 n_2}, \beta \theta_{1n_2} u_{1n_1 n_2}]. \end{cases} \end{aligned} \quad (5.65)$$

Funkcja wiarygodności (5.30) przyjmuje postać

$$\begin{aligned} & L_{1N_1}(W_{N_1 n_2}, \theta_{1n_2}; U_{1N_1 n_2}) \\ &= \begin{cases} \frac{1}{((\beta - \alpha) \theta_{1n_2})^{N_1} \prod_{n_1=1}^{N_1} u_{1n_1 n_2}} & \text{dla } \theta_{1n_2} \in \left[\frac{1}{\beta} \max_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}}, \frac{1}{\alpha} \min_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}} \right] \\ 0 & \text{dla } \theta_{1n_2} \notin \left[\frac{1}{\beta} \max_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}}, \frac{1}{\alpha} \min_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}} \right]. \end{cases} \end{aligned} \quad (5.66)$$

Ponieważ funkcja (5.66) jest funkcją monotoniczną, malejącą względem θ_{1n_2} w przedziale

$$\left[\frac{1}{\beta} \max_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}}, \frac{1}{\alpha} \min_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}} \right], \quad (5.67)$$

maksymalizacja tej funkcji względem θ_{1n_2} daje zatem następujący algorytm estymacji (5.21) na pierwszym stopniu

$$\hat{\theta}_{1N_1 n_2} = \frac{1}{\beta} \max_{1 \leq n_1 \leq N_1} \frac{w_{n_1 n_2}}{u_{1n_1 n_2}}. \quad (5.68)$$

Na podstawie znanego sposobu wyznaczania rozkładu zmiennej losowej $w_{n_1 n_2}$, przekształconej według zależności (5.68) [30, 48], oraz z (5.64) po podstawieniu $\theta_{1n_2} = \theta_2 u_{2n_2}$ otrzymujemy funkcję (5.32) gęstości prawdopodobieństwa zmiennej losowej $\hat{\theta}_{1N_1 n_2}$

$$f_{\theta_1}(\hat{\theta}_{1N_1 n_2}, \theta_2; U_{1N_1 n_2}, u_{2n_2}) = \begin{cases} \frac{\beta^{N_1} N_1 \left(\hat{\theta}_{1N_1 n_2} - \frac{\alpha}{\beta} \theta_2 u_{2n_2} \right)^{N_1 - 1}}{(\beta - \alpha)^{N_1} \theta_2^{N_1} u_{2k}^{N_1}} & \text{dla } \hat{\theta}_{1N_1 n_2} \in \left[\frac{\alpha}{\beta} \theta_2 u_{2n_2}, \theta_2 u_{2n_2} \right] \\ 0 & \text{dla } \hat{\theta}_{1N_1 n_2} \notin \left[\frac{\alpha}{\beta} \theta_2 u_{2n_2}, \theta_2 u_{2n_2} \right]. \end{cases} \quad (5.69)$$

Ostatecznie funkcja wiarygodności (5.33) ma postać

$$L_{2N_2}(\hat{\Xi}_{1N_1 N_2}, \theta_2; \bar{U}_{1N_1 N_2}, U_{2N_2}) = \begin{cases} \frac{(\beta^{N_1} K)^{N_2} \prod_{n_2=1}^{N_2} \left(\hat{\theta}_{1N_1 n_2} - \frac{\alpha}{\beta} \theta_2 u_{2n_2} \right)^{N_1 - 1}}{(\beta - \alpha)^{N_1 N_2} \theta_2^{N_1 N_2} \left(\prod_{n_2=1}^{N_2} u_{2n_2} \right)^{N_1}} & \text{dla } \theta_2 \in \left[\max_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}}, \frac{\beta}{\alpha} \min_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}} \right] \\ 0 & \text{dla } \theta_2 \notin \left[\max_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}}, \frac{\beta}{\alpha} \min_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}} \right]. \end{cases} \quad (5.70)$$

Jest to monotoniczna, malejąca funkcja θ_2 w przedziale

$$\left[\max_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}}, \frac{\beta}{\alpha} \min_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}} \right]. \quad (5.71)$$

W wyniku jej maksymalizacji względem θ_2 otrzymujemy następujący algorytm estymacji (5.22) na drugim stopniu

$$\hat{\theta}_{2N_2} = \max_{1 \leq n_2 \leq N_2} \frac{\hat{\theta}_{1N_1 n_2}}{u_{2n_2}}. \quad (5.72)$$

W rozpatrywanym przykładzie w **podejściu bezpośrednim** funkcja gęstości rozkładu prawdopodobieństwa (5.35) ma postać

$$f_w(w_{n_1 n_2}, \theta_2; u_{1n_1 n_2}, u_{2n_2}) \quad (5.73)$$

$$= \begin{cases} \frac{1}{(\beta - \alpha)\theta_2 u_{1n_1 n_2} u_{2n_2}} & \text{dla } \theta_2 \in \left[\frac{1}{\beta} \max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}}, \frac{1}{\alpha} \min_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}} \right] \\ 0 & \text{dla } \theta_2 \notin \left[\frac{1}{\beta} \max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}}, \frac{1}{\alpha} \min_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}} \right], \end{cases}$$

a funkcja wiarygodności (5.34)

$$L_{N_2 N_1}(\bar{w}_{N_1 N_2}, \theta_2; \bar{u}_{1N_1 N_2}, U_{2N_2}) \quad (5.74)$$

$$= \begin{cases} \frac{1}{((\beta - \alpha)\theta_2)^{N_1 N_2} \prod_{n_2=1}^{N_2} u_{2n_2} \prod_{n_1=1}^{N_1} u_{1n_1 n_2}} & \text{dla } \theta_2 \in \left[\max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{\beta u_{1n_1 n_2} u_{2n_2}}, \min_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{\alpha u_{1n_1 n_2} u_{2n_2}} \right] \\ 0 & \text{dla } \theta_2 \notin \left[\max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{\beta u_{1n_1 n_2} u_{2n_2}}, \min_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{\alpha u_{1n_1 n_2} u_{2n_2}} \right]. \end{cases}$$

Funkcja (5.74) jest monotoniczną, malejącą funkcją θ_2 w przedziale

$$\left[\frac{1}{\beta} \max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}}, \frac{1}{\alpha} \min_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1 n_2}}{u_{1n_1 n_2} u_{2n_2}} \right]. \quad (5.75)$$

W wyniku jej maksymalizacji względem θ_2 otrzymujemy następujący algorytm estymacji (5.22) na drugim stopniu

$$\hat{\theta}_{2N_1N_2} = \frac{1}{\beta} \max_{\substack{1 \leq n_1 \leq N_1 \\ 1 \leq n_2 \leq N_2}} \frac{w_{n_1n_2}}{u_{1n_1n_2} u_{2n_2}}. \quad (5.76)$$

Łatwo zauważyć, że po wstawieniu (5.68) do (5.72) otrzymujemy $\hat{\theta}_{2N_2} = \hat{\theta}_{2N_1N_2}$, a zatem także w tym przypadku dwustopniowa i bezpośrednia estymacja θ_2 dają ten sam wynik. ◆

Analiza algorytmów estymacji przedstawionych w przykładach 5.1 i 5.2 skłania do postawienia pytania: Czy zawsze estymacja bezpośrednia i dwustopniowa dają ten sam wynik?

Teraz sformułujemy ogólny warunek wystarczający równoważności dwustopniowej i bezpośredniej estymacji wektora parametrów θ_2 z zastosowaniem metody maksymalnej wiarygodności:

Twierdzenie 5.1. Jeśli estymator (5.21) na pierwszym stopniu jest dostatecznym estymatorem wektora parametrów θ_{1n_2} , a estymatory na drugim stopniu (5.22) i bezpośredni (5.27) zostały wyznaczone metodą maksymalnej wiarygodności, to estymatory dwustopniowy i bezpośredni są sobie równoważne, czyli oszacowania uzyskane metodami dwustopniową oraz bezpośrednią dla tych samych danych pomiarowych są sobie równe.

Dowód: Wprowadzimy funkcję

$$b_{N_1n_2} \stackrel{\text{df}}{=} \Gamma_{N_1}(W_{N_1n_2}), \quad b_{N_1n_2} \in \mathbf{B} \subseteq \mathbb{R}^{L-R_1} \quad (5.77)$$

taką, że dla funkcji

$$\bar{\Gamma}_{N_1}(W_{N_1n_2}, U_{1N_1n_2}) \stackrel{\text{df}}{=} \begin{bmatrix} \hat{\theta}_{1N_1n_2} \\ b_{N_1n_2} \end{bmatrix} = \begin{bmatrix} \Psi_{1N_1}(U_{1N_1n_2}, W_{N_1n_2}) \\ \Gamma_{N_1}(W_{N_1n_2}) \end{bmatrix} \quad (5.78)$$

istnieje funkcja odwrotna względem $W_{N_1n_2}$

$$W_{N_1n_2} = \bar{\Gamma}_{WN_1}^{-1} \left[\left(\hat{\theta}_{1N_1n_2}, b_{N_1n_2} \right), U_{1N_1n_2} \right], \quad (5.79)$$

gdzie $\bar{\Gamma}_{WN_1}^{-1}$ jest funkcją odwrotną funkcji (5.78) względem $W_{N_1n_2}$.

Funkcję gęstości rozkładu prawdopodobieństwa (5.32) można przedstawić następująco:

$$f_{\theta_1}(\hat{\theta}_{1N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2}) = \frac{f_{\theta_1 b}(\hat{\theta}_{1N_1n_2}, b_{N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2})}{f_{b|\theta_1}(b_{N_1n_2}, \theta_2 | \hat{\theta}_{1N_1n_2}; U_{1N_1n_2}, u_{2n_2})}, \quad (5.80)$$

gdzie: $f_{\theta b}$ jest funkcją gęstości łącznego rozkładu prawdopodobieństwa pary zmiennych losowych $(\hat{\theta}_{1N_1n_2}, \underline{b}_{N_1n_2})$, natomiast $f_{b|\theta_1}$ jest warunkową gęstością $\underline{b}_{N_1n_2} | \hat{\theta}_{1N_1n_2}$ po wstawieniu (5.20), a funkcja wiarygodności (5.33)

$$L_{2N_2}(\hat{\Xi}_{1N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) = \prod_{n_1=1}^{N_1} \frac{f_{\theta b}(\hat{\theta}_{1N_1n_2}, \underline{b}_{N_1n_2}, \theta_2; U_{1N_1n_2}, u_{2n_2})}{f_{b|\theta_1}(\underline{b}_{N_1n_2}, \theta_2 | \hat{\theta}_{1N_1n_2}; U_{1N_1n_2}, u_{2n_2})}. \quad (5.81)$$

Określony przez (5.21) wektor $\hat{\theta}_{1N_1n_2}$ jest estymatorem dostatecznym wektora parametrów θ_{1n_2} [30, 48], warunkowa gęstość prawdopodobieństwa $f_{b|\theta_1}$ nie zależy więc od θ_{1n_2} , a w konsekwencji – po podstawieniu (5.20) – nie zależy od θ_2 , a zatem $\hat{\theta}_{2N_2}$, które jest wynikiem maksymalizacji funkcji wiarygodności (5.81), nie zależy od mianownika wyrażenia (5.80). Z drugiej strony, korzystając z przekształcenia (5.78), iloczyn funkcji gęstości prawdopodobieństwa (5.36), występujący w funkcji wiarygodności (5.34) w podejściu bezpośrednim, możemy wyrazić przez mianownik wyrażenia (5.80), tj.

$$\prod_{n_1=1}^{N_1} f_w(w_{n_1n_2}, \theta_2; u_{1n_1n_2}, u_{2n_2}) = f_{\theta b}(\bar{\Gamma}_{N_1}(W_{N_1n_2}, U_{1N_1n_2}), \theta_2; U_{1N_1n_2}, u_{2n_2}) |J_{\Gamma}|, \quad (5.82)$$

gdzie J_{Γ} jest jacobianem przekształcenia (5.78), który nie zależy od θ_2 .

W tym przypadku funkcja wiarygodności (5.34) przyjmuje postać

$$\begin{aligned} & L_{N_2N_1}(\bar{W}_{N_1N_2}, \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) \\ &= \prod_{n_2=1}^{N_2} f_{\theta b}(\bar{\Gamma}_{N_1}(W_{N_1n_2}, U_{1N_1n_2}), \theta_2; U_{1N_1n_2}, u_{2n_2}) |J_{\Gamma}|, \end{aligned} \quad (5.83)$$

stąd otrzymane w podejściu bezpośrednim $\hat{\theta}_{2N_1N_2}$, tj. θ_2 maksymalizująca (5.83), jest równe $\hat{\theta}_{2N_2}$ maksymalizującej (5.82) względem θ_2 dla tych samych macierzy pomiarów $\bar{U}_{1N_1N_2}$, U_{2N_2} , $\bar{W}_{N_1N_2}$, czyli algorytmy $\tilde{\Psi}_{2N_1N_2}$ (5.27) i $\tilde{\tilde{\Psi}}_{2N_1N_2}$ (5.29) są identyczne.

Q.E.D.

W rozpatrywanych przykładach estymatory na pierwszym stopniu były dostateczne. Autor w pracy [99] podaje przykład, w którym estymacje dwustopniowa i bezpośrednia były równoważne, choć estymator na pierwszym stopniu nie był dostateczny. Nie jest to zatem warunek konieczny równoważności.

5.1.2. Metoda maksymalnego prawdopodobieństwa

Podejście bayesowskie do estymacji parametrów dla obiektu statycznego przedstawiliśmy w punkcie 2.2.1. Stosując różne postacie funkcji strat, otrzymamy różne metody estymacji. Metoda maksymalnej gęstości prawdopodobieństwa a posteriori – zwana dalej w skrócie metodą maksymalnego prawdopodobieństwa – jest szczególnym, najczęściej występującym przypadkiem podejścia ogólnego, w którym funkcja strat ma postać (2.71). Dla tej metody rozpatrzmy dalej typowy przypadek naszego problemu oraz kwestię równoważności.

Zacniemy jednak od przypadku ogólnego. Teraz założymy, że θ_2 jest wartością ciągłej zmiennej losowej $\underline{\theta}_2$ o znanej funkcji gęstości rozkładu prawdopodobieństwa $f_{\theta_2}(\theta_2)$. Zgodnie z zależnością (5.20) także θ_{1n_2} jest wartością odpowiedniej zmiennej losowej $\underline{\theta}_{1n_2}$ i – znając $f_{\theta_2}(\theta_2)$ oraz funkcję F_2 w (5.20) – możemy wyznaczyć funkcję gęstości rozkładu prawdopodobieństwa $f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2})$. Gdy stosujemy metodę minimalnego ryzyka, należy na obu stopniach minimalizować ryzyko średnie dla określonej funkcji strat, co – jak wiadomo – można sprowadzić do minimalizacji odpowiedniego ryzyka warunkowego. Na pierwszym stopniu zadanie identyfikacji sprowadza się do zastosowania znanej metody estymacji bayesowskiej dla obiektu statycznego, przedstawionego w punkcie 2.2.1. Zadanie na drugim stopniu, podobnie jak w metodzie maksymalnej wiarygodności, jest bardziej złożone, wymaga bowiem wyznaczenia odpowiedniego warunkowego rozkładu prawdopodobieństwa. Procedura wyznaczania algorytmów estymacji jest następująca:

Krok 1. Dla danej funkcji gęstości rozkładu prawdopodobieństwa $f_2(\theta_2)$ wyznaczamy na podstawie (5.20) funkcję gęstości

$$f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) = f_{\theta_2}(F_{\theta_2 2}^{-1}(u_{2n_2}, \theta_{1n_2})) |J_{F_2}|, \quad (5.84)$$

przy czym $F_{\theta_2 2}^{-1}$ jest przekształceniem odwrotnym (5.20) względem θ_2 , przy założeniu wzajemnej jednoznaczności funkcji F_2 , czyli $\theta_2 = F_{\theta_2 2}^{-1}(u_{2n_2}, \theta_{1n_2})$.

Krok 2. Przez analogię do (2.60) oraz (2.62) wyznaczamy funkcje gęstości warunkowych rozkładów prawdopodobieństwa

$$f_{W_{N_1}|\theta_1}(W_{N_1 n_2} | \theta_{1n_2}; U_{N_1 n_2}) = \prod_{n_1=1}^{N_1} f_z(h_z^{-1}(F_1(u_{1n_1 n_2}, \theta_{1n_2}), w_{n_1 n_2})) |J_h| \quad (5.85)$$

oraz

$$f_1'(\theta_{1n_2} | W_{N_1 n_2}; U_{N_1 n_2}, u_{2n_2}) = \frac{f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1}|\theta_1}(W_{N_1 n_2} | \theta_{1n_2}; U_{N_1 n_2})}{\int_{\Theta_1} f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1}|\theta_1}(W_{N_1 n_2} | \theta_{1n_2}; U_{N_1 n_2}) d\theta_{1n_2}}. \quad (5.86)$$

Krok 3. Na pierwszym stopniu, przy danej funkcji strat $L_1(\theta_{1n_2}, \bar{\theta}_{1n_2})$, minimalizujemy względem $\bar{\theta}_{1n_2}$ ryzyko warunkowe

$$r_1(\bar{\theta}_{1n_2}, W_{N_1n_2}; U_{N_1n_2}) \stackrel{\text{df}}{=} E_{\theta_{1n_2}} \left[L_1(\theta_{1n_2}, \bar{\theta}_{1n_2}) \mid W_{N_1n_2} \right], \quad (5.87)$$

a korzystając z (5.86), otrzymujemy

$$\begin{aligned} & r_1(\bar{\theta}_{1n_2}, W_{N_1n_2}; U_{N_1n_2}, u_{2n_2}) \\ &= \int_{\theta_1} L_1(\theta_{1n_2}, \bar{\theta}_{1n_2}) \frac{f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1} | \theta_1}(W_{N_1n_2} | \theta_{1n_2}; U_{N_1n_2})}{\int_{\theta_1} f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1} | \theta_1}(W_{N_1n_2} | \theta_{1n_2}; U_{N_1n_2}) d\theta_{1n_2}} d\theta_{1n_2}. \end{aligned} \quad (5.88)$$

Ponieważ licznik wyrażenia (5.86) dla ustalonych pomiarów jest wartością stałą i nie zależy od θ_{1n_2} , minimalizacja wyrażenia (5.88) względem $\bar{\theta}_{1n_2}$ sprowadza się do minimalizacji wyrażenia

$$\begin{aligned} & \bar{r}_1(\bar{\theta}_{1n_2}, W_{N_1n_2}; U_{N_1n_2}, u_{2n_2}) \\ & \stackrel{\text{df}}{=} \int_{\theta_1} L_1(\theta_{1n_2}, \bar{\theta}_{1n_2}) f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1} | \theta_1}(W_{N_1n_2} | \theta_{1n_2}; U_{N_1n_2}) d\theta_{1n_2} \end{aligned} \quad (5.89)$$

względem $\bar{\theta}_{1n_2}$. W wyniku minimalizacji (5.89) względem $\bar{\theta}_{1n_2}$ otrzymujemy algorytm estymacji (5.21), przy czym $\hat{\theta}_{1N_1n_2}$ oznacza wynik minimalizacji. Warto zauważyć, że u_{2n_2} występuje tu jako parametr algorytmu (5.21).

Krok 4. Na podstawie zależności (5.21) oraz po podstawieniu (5.20) do (5.14) wyznaczamy funkcję gęstości rozkładu prawdopodobieństwa

$$f_{\theta_1 | \theta_2}(\hat{\theta}_{1N_1n_2} | \theta_2; U_{1N_1n_2}, u_{2n_2}) \quad (5.90)$$

dla zmiennej losowej $\hat{\theta}_{1N_1n_2}$. Funkcję tę możemy także wyznaczyć z zależności (5.21), po podstawieniu (5.26) do (5.14) oraz skorzystaniu ze znajomości $f_z(z)$. W konsekwencji warunkowa funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej $\hat{\underline{\theta}}_{1N_1N_2} | \underline{\theta}_2$ ma postać

$$f_{\hat{\underline{\theta}}_1 | \underline{\theta}_2}(\hat{\underline{\theta}}_{1N_1N_2} | \underline{\theta}_2; \bar{U}_{1N_1N_2}, U_{2N_2}) = \prod_{n_2=1}^{N_2} f_{\theta_1 | \theta_2}(\hat{\theta}_{1N_1n_2} | \theta_2; U_{1N_1n_2}, u_{2n_2}). \quad (5.91)$$

Krok 5. Dla danej funkcji strat $L_2(\theta_2, \bar{\theta}_2)$ minimalizujemy względem $\bar{\theta}_2$ ryzyko warunkowe, tzn.

$$r_2(\theta_2, \hat{\Xi}_{1N_1N_2}; U_{2N_2}) \stackrel{\text{df}}{=} E_{\theta_2} \left[L_2(\theta_2, \bar{\theta}_2) \middle| \hat{\Xi}_{1N_1N_2} \right] \quad (5.92)$$

co, podobnie jak w kroku 3., sprowadza się do minimalizacji względem $\bar{\theta}_2$ wyrażenia

$$\bar{r}_2(\theta_2, \hat{\Xi}_{1N_1N_2}; U_{2N_2}) \stackrel{\text{df}}{=} \int_{\theta_2} L(\theta_2, \bar{\theta}_2) f_{\theta_2}(\theta_2) f_{\hat{\Xi}_1 | \theta_2}(\hat{\Xi}_{1N_1N_2} | \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) d\theta_2, \quad (5.93)$$

stąd otrzymujemy algorytm estymacji (5.22), przy czym $\hat{\theta}_{2N_2}$ oznacza $\bar{\theta}_2$, który minimalizuje (5.93).

W **podjęciu bezpośrednim** minimalizujemy względem $\bar{\theta}_2$ ryzyko warunkowe

$$r(\bar{\theta}_2, \bar{W}_{N_1N_2}; \bar{U}_{1N_1N_2}, U_{1N_2}) \stackrel{\text{df}}{=} E_{\bar{\theta}_2} \left[L_2(\bar{\theta}_2, \bar{\theta}_2) \middle| \bar{W}_{N_1N_2} \right], \quad (5.94)$$

co sprowadza się do minimalizacji względem $\bar{\theta}_2$ wyrażenia

$$\bar{r}(\bar{\theta}_2, \bar{W}_{NK}; \bar{U}_{1N_1N_2}, U_{1N_2}) \stackrel{\text{df}}{=} \int_{\theta_2} L_2(\theta_2, \bar{\theta}_2) f_2(\theta_2) f_{W_{N_1N_2}}(\bar{W}_{N_1N_2} | \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) d\theta_2, \quad (5.95)$$

gdzie funkcję gęstości $f_{W_{N_1N_2}}$ wyznaczamy według zależności

$$f_{W_{N_1N_2}}(\bar{W}_{N_1N_2} | \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) = \prod_{n_2=1}^{N_2} \prod_{n_1=1}^{N_1} f_z(h^{-1}(F(u_{1n_1n_2}, u_{2n_2}, \theta_2), w_{n_1n_2}))). \quad (5.96)$$

W wyniku otrzymujemy algorytm estymacji (5.27), przy czym $\hat{\theta}_{2N_1N_2}$ oznacza wektor $\bar{\theta}_2$, który minimalizuje (5.95).

Jak wiadomo, po odpowiednim doborze funkcji strat, minimalizacja ryzyka sprowadza się do maksymalizacji funkcji gęstości rozkładu prawdopodobieństwa a posteriori dla nieznanego wektora parametrów. Po pominięciu mianowników niezależnych od zmiennych maksymalizujących $\hat{\theta}_{1N_1n_2}$ jest wektorem θ_{1n_2} , który maksymalizuje iloczyn gęstości w zależności (5.89), tj.

$$f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1} | \theta_1}(W_{N_1n_2} | \theta_{1n_2}; U_{N_1n_2}). \quad (5.97)$$

Podobnie $\hat{\theta}_{2N_2}$ i $\hat{\theta}_{2N_1N_2}$ otrzymamy w wyniku maksymalizacji względem θ_2 , odpowiednio:

$$f_{\theta_2}(\theta_2) f_{\hat{\Xi}_1 | \theta_2}(\hat{\Xi}_{1N_1N_2} | \theta_2; \bar{U}_{1N_1N_2}, U_{2N_2}) \quad (5.98)$$

oraz

$$f_{\theta_2}(\theta_2) f_{W_{N_1 N_2}}(\bar{W}_{N_1 N_2} | \theta_2; \bar{U}_{1 N_1 N_2}, U_{2 N_2}). \quad (5.99)$$

Rozpatrzmy w charakterze przykładu typowy przypadek, kiedy można analitycznie wyznaczyć algorytm estymacji metodą maksymalnego prawdopodobieństwa.

Przykład 5.3. Rozpatrzmy dwustopniowy obiekt, którego wartości wyjścia i wejścia są mierzone z zakłóceniami. Zależności (5.19), (5.20) i (5.26) są takie same jak w przykładzie 5.1, tzn. mają postać (5.37), (5.38) i (5.40). Zakłócenia pomiarowe są addytywne i mają rozkład normalny (5.42), a funkcja h ma postać (5.39). Zakładamy ponadto rozkład normalny wektora $\underline{\theta}_2$, tj.

$$f_{\theta_2}(\theta_2) = \frac{(2\pi)^{-\frac{R_2}{2}}}{\sqrt{|\Sigma_2|}} \exp \left[-\frac{1}{2} (\theta_2 - m_{\theta_2})^T \Sigma_2^{-1} (\theta_2 - m_{\theta_2}) \right], \quad (5.100)$$

gdzie: m_{θ_2} i Σ_2 oznaczają – odpowiednio – wektor wartości oczekiwanych i macierz kowariancji zmiennej losowej $\underline{\theta}_2$.

Po uwzględnieniu zależności (5.38) oraz rozkładu (5.100) zmienna losowa $\underline{\theta}_{1n_2}$ ma rozkład normalny z wartością oczekiwaną i wariancją, odpowiednio:

$$E(\underline{\theta}_{1n_2}) = m_{\theta_2}^T u_{2n_2}, \quad \text{Var}(\underline{\theta}_{1n_2}) = u_{2n_2}^T \Sigma_2 u_{2n_2}, \quad (5.101)$$

czyli funkcja gęstości rozkładu prawdopodobieństwa (5.84) przyjmuje postać

$$f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) = (2\pi u_{2n_2}^T \Sigma_2 u_{2n_2})^{-\frac{1}{2}} \exp \left[-\frac{(\theta_{1n_2} - m_{\theta_2}^T u_{2n_2})^2}{2u_{2n_2}^T \Sigma_2 u_{2n_2}} \right], \quad (5.102)$$

a wyrażenie (5.97)

$$\begin{aligned} & f_{\theta_{1n_2}}(\theta_{1n_2}; u_{2n_2}) f_{W_{N_1} | \theta_1}(W_{N_1 n_2} | \theta_{1n_2}; U_{N_1 n_2}) \\ &= \frac{(2\pi)^{-\frac{N_1+1}{2}}}{\sigma_z \sqrt{u_{2n_2}^T \Sigma_2 u_{2n_2}}} \exp \left[-\frac{(\theta_{1n_2} - m_{\theta_2}^T u_{2n_2})^2}{2u_{2n_2}^T \Sigma_2 u_{2n_2}} - \frac{\sum_{n_1=1}^{N_1} (w_{n_1 n_2} - \theta_{1n_2} u_{1n_1 n_2})^2}{2\sigma_z^2} \right]. \end{aligned} \quad (5.103)$$

W wyniku maksymalizacji funkcji (5.103) względem θ_{1n_2} otrzymujemy algorytm estymacji dla pierwszego stopnia

$$\hat{\theta}_{1N_1n_2} = \frac{m_{\theta_2}^T u_{2n_2} + \frac{u_{2n_2}^T \Sigma_2 u_{2n_2}}{\sigma_z^2} U_{1N_1n_2} W_{N_1n_2}^T}{1 + \frac{u_{2n_2}^T \Sigma_2 u_{2n_2}}{\sigma_z^2} U_{1N_1n_2} U_{1N_1n_2}^T}. \quad (5.104)$$

Na podstawie (5.41) i (5.104) łatwo zauważyć, że rozkład warunkowy $\hat{\theta}_{1N_1n_2}$ względem θ_2 jest rozkładem normalnym, z warunkową wartością oczekiwaną i warunkową wariancją równym, odpowiednio:

$$E(\hat{\theta}_{1N_1n_2} | \theta_2) = \frac{\sigma_z^2 m_{\theta_2}^T u_{2n_2} + u_{2n_2}^T \Sigma_2 u_{2n_2} U_{1N_1n_2} U_{1N_1n_2}^T m_{\theta_2}^T u_{2n_2}}{\sigma_z^2 + u_{2n_2}^T \Sigma_2 u_{2n_2} U_{1N_1n_2} U_{1N_1n_2}^T} \stackrel{\text{df}}{=} m_{\theta_{1n_2}}, \quad (5.105)$$

$$\text{Var}(\hat{\theta}_{1N_1n_2} | \theta_2) = \frac{(u_{2n_2}^T \Sigma_2 u_{2n_2})^2 U_{1N_1n_2} U_{1N_1n_2}^T}{(\sigma_z^2 + u_{2n_2}^T \Sigma_2 u_{2n_2})^2} \stackrel{\text{df}}{=} \sigma_{1n_2}^2, \quad (5.106)$$

czyli funkcja gęstości rozkładu prawdopodobieństwa (5.90) przyjmuje postać

$$f_{\theta_1 | \theta_2}(\hat{\theta}_{1N_1n_2} | \theta_2; U_{1N_1n_2}, u_{2n_2}) = \frac{1}{\sigma_{1n_2} \sqrt{2\pi}} \exp \left[-\frac{(\hat{\theta}_{1N_1n_2} - m_{\theta_{1n_2}})^2}{2\sigma_{1n_2}^2} \right], \quad (5.107)$$

a wyrażenie (5.98) zapisujemy

$$\begin{aligned} & f_{\theta_2}(\theta_2) f_{\hat{\theta}_{1N_1n_2} | \theta_2}(\hat{\theta}_{1N_1n_2} | \theta_2; \bar{U}_{1N_1n_2}, U_{2N_2}) \\ &= \frac{(2\pi)^{\frac{R_2+1}{2}}}{\sigma_{1n_2} \sqrt{|\Sigma_2|}} \exp \left[-\frac{(\theta_2 - m_{\theta_2})^T \Sigma_2^{-1} (\theta_2 - m_{\theta_2})}{2} - \frac{(\hat{\theta}_{1N_1n_2} - m_{\theta_{1n_2}})^2}{2\sigma_{1n_2}^2} \right]. \end{aligned} \quad (5.108)$$

W wyniku minimalizacji wyrażenia (5.108) otrzymujemy algorytm (5.21) estymacji na drugim stopniu postaci

$$\begin{aligned} \hat{\theta}_{2N_2} &= \left[I + \frac{\Sigma_2}{\sigma_z^2} \sum_{n_2=1}^{N_1} U_{1N_1n_2} U_{1N_1n_2}^T u_{2n_2} u_{2n_2}^T \right]^{-1} \\ &\times \left[\left(I - \Sigma_2 \sum_{n_2=1}^{N_1} \frac{U_{1N_1n_2} U_{1N_1n_2}^T u_{2n_2} u_{2n_2}^T}{u_{2n_2}^T \Sigma_2 u_{2n_2}} \right) m_{\theta_2} + \frac{\Sigma_2}{\sigma_z^2} \sum_{n_2=1}^{N_1} \left(\frac{\sigma_z^2}{u_{2n_2}^T \Sigma_2 u_{2n_2}} + U_{1N_1n_2} U_{1N_1n_2}^T \right) \hat{\theta}_{1N_1n_2} u_{2n_2} \right], \end{aligned} \quad (5.109)$$

gdzie I oznacza macierz jednostkową odpowiedniego wymiaru.

W podejściu bezpośrednim wyrażenie (5.99) przyjmuje postać

$$f_{\theta_2}(\theta_2) f_{W_{N_1 N_2}}(\bar{W}_{N_1 N_2} | \theta_2; \bar{U}_{1 N_1 N_2}, U_{2 N_2}) \quad (5.110)$$

$$= \frac{(2\pi)^{\frac{R_2 + N_1 N_2}{2}}}{\sqrt{|\Sigma_2|}} \exp \left[-\frac{1}{2} (\theta_2 - m_{\theta_2})^T \Sigma_2^{-1} (\theta_2 - m_{\theta_2}) - \frac{1}{2\sigma_z^2} \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} (w_{n_1 n_2} - \theta_2^T u_{n_2} u_{n_1 n_2})^2 \right].$$

Maksymalizacja wyrażenia (5.110) względem θ_2 daje następujący bezpośredni algorytm estymacji (5.27)

$$\hat{\theta}_{2, N_1 N_2} = \left[I + \frac{1}{\sigma_z^2} \Sigma_2 \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{2 n_2} u_{2 n_2}^T u_{1 n_1 n_2}^2 \right]^{-1} \left[m_{\theta_2} + \frac{1}{\sigma_z^2} \Sigma_2 \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{2 n_2} u_{1 n_1 n_2} w_{n_1 n_2} \right], \quad (5.111)$$

Po wstawieniu (5.104) do (5.109) i po przekształceniach można sprawdzić, że $\hat{\theta}_{2, N_2} = \hat{\theta}_{2, N_1 N_2}$, a zatem w rozpatrywanym przypadku dwustopniowa i bezpośrednia estymacja θ_2 są równoważne. ◆

Ogólny warunek wystarczający dla równoważności estymacji dwustopniowej i bezpośredniej θ_2 jest tu identyczny z warunkiem dla metody maksymalnej wiarygodności (twierdzenie 5.1).

Twierdzenie 5.2. Jeśli estymator (5.21) na pierwszym stopniu jest dostatecznym estymatorem wektora parametrów $\theta_{1 n_2}$, a estymatory na drugim stopniu (5.22) i bezpośredni (5.27) zostały wyznaczone metodą maksymalnego prawdopodobieństwa a posteriori, to estymatory dwustopniowy i bezpośredni są sobie równoważne, czyli oszacowania uzyskane metodami dwustopniową oraz bezpośrednią dla tych samych danych pomiarowych są sobie równe.

Dowód: Dowód jest analogiczny do dowodu twierdzenia 5.1. Po wprowadzeniu $b_{N_1 n_2}$ (5.77) i funkcji Γ_{N_1} (5.78) funkcję gęstości rozkładu prawdopodobieństwa (5.90) można przedstawić następująco:

$$f_{\theta_1 | \theta_2}(\hat{\theta}_{1 N_1 n_2} | \theta_2; U_{1 N_1 n_2}, u_{2 n_2}) = \frac{f_{\theta_1 b | \theta_2}(\hat{\theta}_{1 N_1 n_2}, b_{N_1 n_2} | \theta_2; U_{1 N_1 n_2}, u_{2 n_2})}{f_{b | \theta_1 \theta_2}(b_{N_1 n_2} | \theta_2, \hat{\theta}_{1 N_1 n_2}; U_{1 N_1 n_2}, u_{2 n_2})}, \quad (5.112)$$

gdzie: $f_{\theta_1 b | \theta_2}$ jest warunkową funkcją gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $\hat{\theta}_{1 N_1 n_2}, b_{N_1 n_2} | \theta_2$, a $f_{b | \theta_1 \theta_2}$ jest warunkową funkcją gęstości rozkładu prawdopodobieństwa zmiennej $b_{N_1 n_2} | \hat{\theta}_{1 N_1 n_2}, \theta_2$.

Po wstawieniu (5.112) do (5.91) otrzymujemy

$$f_{\hat{\Xi}_1|\theta_2}(\hat{\Xi}_1|_{N_1N_2}|\theta_2;\bar{U}_{1N_1N_2},U_{2N_2})=\prod_{n_2=1}^{N_2}\frac{f_{\theta_1b|\theta_2}(\hat{\theta}_{1N_1n_2},b_{N_1n_2}|\theta_2;U_{1N_1n_2},u_{2n_2})}{f_{b|\theta_1\theta_2}(b_{N_1n_2}|\theta_2,\hat{\theta}_{1N_1n_2};U_{1N_1n_2},u_{2n_2})}. \quad (5.113)$$

Określony przez (5.21) wektor $\hat{\theta}_{1N_1n_2}$ jest estymatorem dostatecznym wektora parametrów θ_{1n_2} [30, 48], warunkowa funkcja gęstości rozkładu prawdopodobieństwa $f_{b|\theta_1\theta_2}$ nie zależy więc od θ_{1n_2} , a w konsekwencji, po podstawieniu (5.20), nie zależy od θ_2 , zatem $\hat{\theta}_{2N_2}$ – które jest wynikiem maksymalizacji wyrażenia (5.98), z uwzględnieniem (5.113) – nie zależy od mianownika wyrażenia (5.113). Korzystając jednak z przekształcenia (5.78), iloczyn funkcji gęstości prawdopodobieństwa (5.36), występujący w funkcji (5.99) w podejściu bezpośrednim, czyli funkcję gęstości rozkładu prawdopodobieństwa możemy wyrazić

$$\begin{aligned} & f_{W_{N_1N_2}}(\bar{W}_{N_1N_2}|\theta_2;\bar{U}_{1N_1N_2},U_{2N_2}) \\ &= \prod_{n_2=1}^{N_2} f_{\theta_1b|\theta_2}(\bar{r}_{N_1}(W_{N_1n_2},U_{1N_1n_2})|\theta_2;U_{1N_1n_2},u_{2n_2})|J_F|, \end{aligned} \quad (5.114)$$

gdzie J_F jest jakobianem przekształcenia (5.78), który nie zależy od θ_2 .

Stąd otrzymane w podejściu bezpośrednim $\hat{\theta}_{2N_1N_2}$, tj. wartość θ_2 maksymalizująca (5.99) z uwzględnieniem (5.114), jest równe wartości $\hat{\theta}_{2N_2}$ maksymalizującej wyrażenie (5.98) względem θ_2 dla tych samych macierzy pomiarów $\bar{U}_{1N_1N_2}$, U_{2N_2} , $\bar{W}_{N_1N_2}$, czyli algorytmy $\tilde{\Psi}_{2N_1N_2}$ (5.27) i $\tilde{\Psi}_{2N_1N_2}$ (5.29) są identyczne. *Q.E.D.*

Łatwo zauważyć, że twierdzenie 5.2 można uogólnić na metodę minimalnego ryzyka z dowolnymi funkcjami strat.

Twierdzenie 5.3. Jeśli estymator (2.6) na pierwszym stopniu jest dostatecznym estymatorem parametrów θ_{1n} , a estymatory na drugim stopniu (2.7) i bezpośredni (2.10) zostały wyznaczone metodą minimalnego ryzyka z tą samą funkcją strat L_2 , to estymatory dwustopniowy i bezpośredni są sobie równoważne, czyli oszacowania uzyskane metodami dwustopniową oraz bezpośrednią dla tych samych danych pomiarowych są sobie równe.

Dowód: Wystarczy skorzystać z dowodu poprzedniego twierdzenia i zauważyć, że minimalizacja ryzyka (5.93) względem $\bar{\theta}_2$ sprowadza się do minimalizacji wyrażenia (5.94). Podobnie, minimalizacja ryzyka (5.94) względem $\bar{\theta}_2$ sprowadza się do minimalizacji (5.95) po wstawieniu (5.114), a zatem zależności $\hat{\theta}_{2N_2}$ i $\hat{\theta}_{2N_1N_2}$ od $\bar{U}_{1N_1N_2}$, U_{2N_2} , $\bar{W}_{N_1N_2}$ są identyczne. *Q.E.D.*

5.2. Dwustopniowe zadanie wyboru optymalnego modelu w warunkach losowych

Rozważmy dwustopniowy obiekt statyczny o wektorach wejść u_1 i u_2 oraz wyjść y , takich jak w podrozdziale 5.1. W dalszym ciągu, podobnie jak w punkcie 2.4.1, będziemy zakładać, że (u_1, u_2, y) są wartościami zmiennych losowych $(\underline{u}_1, \underline{u}_2, \underline{y})$, dla których istnieje funkcja gęstości rozkładu prawdopodobieństwa $f(u_1, u_2, y)$. Charakterystyki statyczne na pierwszym i drugim stopniu nie są znane. Proponujemy modele na pierwszym i drugim stopniu – odpowiednio – postaci:

$$\bar{y} = \Phi_1(u_1, \theta_1), \quad (5.115)$$

$$\bar{\theta}_1 = \Phi_2(u_2, \theta_2), \quad (5.116)$$

gdzie: $u_1 \in U_1 \subseteq \mathbb{R}^{S_1}$ jest S_1 -wymiarowym wektorem wejść na pierwszym stopniu; $u_2 \in U_2 \subseteq \mathbb{R}^{S_2}$ jest S_2 -wymiarowym wektorem wejść na drugim stopniu; $\theta_1 \in \Theta_1 \subseteq \mathbb{R}^{R_1}$ jest R_1 -wymiarowym wektorem parametrów modelu na pierwszym stopniu; $\bar{\theta}_1 \in \Theta_1 \subseteq \mathbb{R}^{R_1}$ jest R_1 -wymiarowym wektorem wyjść modelu na drugim stopniu; $\theta_2 \in \Theta_2 \subseteq \mathbb{R}^{R_2}$ jest R_2 -wymiarowym wektorem parametrów modelu na drugim stopniu; $y, \bar{y} \in Y \subseteq \mathbb{R}^L$ są L -wymiarowymi wektorami wyjściowymi – odpowiednio – obiektu i modelu na pierwszym stopniu; Φ_1, Φ_2 są zadanymi funkcjami takimi, że

$$\Phi_1 : U_1 \times \Theta_1 \rightarrow Y, \quad \Phi_2 : U_2 \times \Theta_2 \rightarrow \Theta_1.$$

Wprowadzimy definicje precyzujące, w jakim sensie będą optymalne rozpatrywane modele.

Definicja 5.1. Model na pierwszym stopniu (5.115) nazwiemy modelem optymalnym dla danego u_2 , jeśli $\theta_1 = \theta_1^*$, gdzie θ_1^* jest wartością θ_1 , minimalizującą kryterium jakości identyfikacji

$$\begin{aligned} Q_1(\theta_1, u_2) &= \underset{\underline{u}_1, \underline{y}}{\text{E}} \left[q_1(y, \bar{y} = \Phi_1(\underline{u}_1, \theta_1)) \middle| u_2 \right] \\ &= \int_Y \int_{U_1} q_1(y, \Phi_1(u_1, \theta_1)) f_1(u_1, y | u_2) du_1 dy \end{aligned} \quad (5.117)$$

czyli

$$Q_1(\theta_1^*, u_2) = \min_{\theta_1 \in \Theta_1} Q_1(\theta_1, u_2), \quad (5.118)$$

gdzie: $q_1(y, \bar{y})$ jest zadaną funkcją, której wartość ocenia różnicę pomiędzy wyjściem obiektu i wyjściem modelu na pierwszym stopniu, $q_1 : Y \times Y \rightarrow \mathbb{R}$, $q_1(y, \bar{y}) \geq 0$, $q_1(y, \bar{y}) = 0 \Leftrightarrow y = \bar{y}$, a f_1 oznacza warunkową funkcję gęstości roz-

kładu prawdopodobieństwa zmiennych losowych $\underline{u}_1, \underline{y}$ pod warunkiem, że zmienna losowa \underline{u}_2 przyjęła wartość u_2 .

W wyniku minimalizacji (5.117) względem θ_1 dla danego u_2 otrzymamy zależność θ_1^* od u_2

$$\theta_1^* \stackrel{\text{df}}{=} G(u_2). \quad (5.119)$$

Jest to model dla drugiego stopnia, wyznaczony w wyniku optymalizacji na stopniu pierwszym. Wartości θ_1^* można porównywać z wartościami $\bar{\theta}_1$, wyznaczonymi z modelu (5.116) dla tych samych wartości u_2 , i wyznaczyć optymalny wektor parametrów modelu (5.116), tj. najlepszą aproksymację zależności (5.119) za pomocą wzoru postaci (5.116).

Definicja 5.2. Model na drugim stopniu (5.116) nazwiemy modelem optymalnym dla $\theta_2 = \theta_2^*$, gdy θ_2^* jest wartością θ_2 , minimalizującą kryterium jakości identyfikacji

$$Q_2(\theta_2) \stackrel{\text{df}}{=} E_{u_2} [q_2(\theta_1^*, \bar{\theta}_1 = \Phi_2(\underline{u}_2, \theta_2))] = \int_{U_2} q_2(G(u_2), \Phi_2(u_2, \theta_2)) f_2(u_2) du_2, \quad (5.120)$$

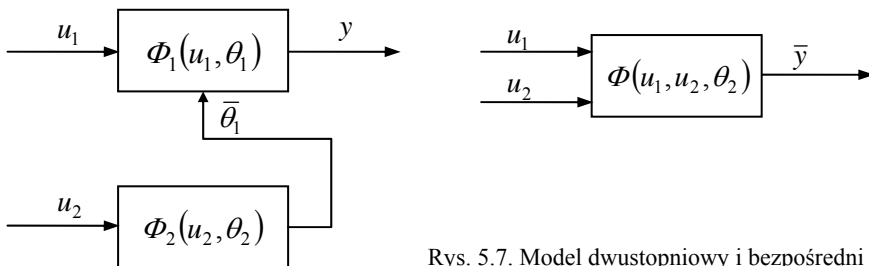
czyli

$$Q_2(\theta_2^*) = \min_{\theta_2 \in \Theta_2} Q_2(\theta_2), \quad (5.121)$$

gdzie: $q_2(\theta_1^*, \bar{\theta}_1)$ jest zadaną funkcją ($q_2: \mathbb{R}^{R_1} \times \mathbb{R}^{R_1} \rightarrow \mathbb{R}$ o własnościach takich jak q_1) porównującą wartość θ_1^* , wyznaczoną na pierwszym stopniu, z wartością $\bar{\theta}_1$, wyznaczoną z modelu (5.116) na drugim stopniu, a f_2 oznacza brzegową funkcję gęstości rozkładu prawdopodobieństwa dla \underline{u}_2 .

Gdy zastosujemy podejście bezpośrednie, wyznaczamy optymalną wartość θ_2 dla modelu bezpośredniego o zadanej postaci $\bar{y} = \Phi(u_1, u_2, \theta_2)$. W dalszym ciągu będziemy rozpatrywać model bezpośredni (rys. 5.7), otrzymany w wyniku kompozycji modeli (5.115) i (5.116), tzn.

$$\bar{y} = \Phi(u_1, u_2, \theta_2) \stackrel{\text{df}}{=} \Phi_1(u_1, \Phi_2(u_2, \theta_2)). \quad (5.122)$$



Rys. 5.7. Model dwustopniowy i bezpośredni

Oznacza to, że pierwotnie mamy określone modele na obu stopniach, tzn. określenie postaci modeli ma charakter dwustopniowy, natomiast wybór optymalnej wartości θ_2 może być dwustopniowy lub bezpośredni.

Definicja 5.3. Model bezpośredni (5.122) nazwiemy modelem optymalnym dla $\theta_2 = \tilde{\theta}_2^*$, minimalizującego kryterium jakości identyfikacji

$$\begin{aligned} Q(\theta_2) &\stackrel{\text{df}}{=} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[q_1(\underline{y}, \bar{y} = \Phi(\underline{u}_1, \underline{u}_2, \theta_2)) \right] \\ &= \int_Y \int_{U_2} \int_{U_1} q_1(y, \Phi(u_1, u_2, \theta_2)) f(u_1, u_2, y) du_1 du_2 dy, \end{aligned} \quad (5.123)$$

czyli

$$Q(\tilde{\theta}_2^*) = \min_{\theta_2 \in \Theta_2} Q(\theta_2). \quad (5.124)$$

Podójście dwustopniowe i bezpośrednie nazwiemy podejściem równoważnym, jeśli $\theta_2^* = \tilde{\theta}_2^*$.

W przypadku pełnej informacji probabilistycznej, tj. przy znajomości funkcji gęstości rozkładu prawdopodobieństwa $f(u_1, u_2, y)$, problem dwustopniowego wyboru optymalnego modelu polega na wyznaczeniu zależności (5.119), minimalizującej (5.117), oraz wartości θ_2 , minimalizującej (5.120), dla danych funkcji Φ_1 , Φ_2 , q_1 , q_2 . Procedura dwustopniowego wyboru najlepszego modelu jest zatem następująca:

Krok 1. Wyznaczamy funkcje gęstości rozkładu prawdopodobieństwa f_2 i f_1 :

$$f_2(u_2) = \int_Y \int_{U_1} f(u_1, u_2, y) du_1 dy, \quad (5.125)$$

$$f_1(u_1, y | u_2) = \frac{f(u_1, u_2, y)}{f_2(u_2)}. \quad (5.126)$$

Krok 2. W wyniku obliczenia całki (5.117) wyznaczamy funkcję $Q_1(\theta_1, u_2)$.

Krok 3. Minimalizujemy funkcję $Q_1(\theta_1, u_2)$ względem θ_1 przy ograniczeniu $\theta_1 \in \Theta_1$ i otrzymujemy zależność (5.119).

Krok 4. W wyniku obliczenia całki (5.120) dla wyznaczonej funkcji (5.119) wyznaczamy funkcję $Q_2(\theta_2)$.

Krok 5. Minimalizujemy $Q_2(\theta_2)$ względem θ_2 przy ograniczeniu $\theta_2 \in \Theta_2$.

Nawet bez ograniczeń narzuconych na wybór θ_1 i θ_2 (tzn. $\Theta_1 \equiv \mathbb{R}^{R_1}$, $\Theta_2 \equiv \mathbb{R}^{R_2}$) wyznaczenie funkcji Q_1 i Q_2 oraz ich minimalizacja mogą być zadaniami trudnymi lub niemożliwymi do rozwiązania analitycznego i konieczne jest stosowanie odpowiednich procedur numerycznych. Wyniki analityczne można uzyskać dla przypadku liniowo-kwadratowego, przedstawionego w punkcie 5.2.1. Podobnie jak w estymacji dwustopniowej, do wyboru optymalnego modelu na drugim stopniu nie można zastosować wprost podejść i wzorów przedstawionych w podrozdziale 2.4. Wyjście na drugim stopniu, tj. θ_1^* , nie jest wprost realizowane i ewentualnie obserwowane, lecz jest wyznaczane w wyniku identyfikacji na pierwszym stopniu. Dlatego też nie można wprost posługiwać się łączną gęstością $f(u_2, \theta_1)$ jako daną.

Gdy brak jest pełnej informacji probabilistycznej, a więc w praktycznej sytuacji, należy posługiwać się danymi pomiarowymi, uzyskiwanymi tak jak w punkcie 2.4.2, czyli:

$$U_{1N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{11n_2} & u_{12n_2} & \cdots & u_{1N_1n_2} \end{bmatrix}, \quad (5.127)$$

$$Y_{N_1n_2} \stackrel{\text{df}}{=} \begin{bmatrix} y_{1n_2} & y_{2n_2} & \cdots & y_{N_1n_2} \end{bmatrix}, \quad (5.128)$$

gdzie: $U_{1N_1n_2}$, $Y_{N_1n_2}$ są macierzami pomiarów – odpowiednio – wartości składowych wektora wejść i wyjść na pierwszym stopniu dla ustalonego $u_2 = u_{2n_2}$, a na drugim stopniu

$$U_{2N_2} \stackrel{\text{df}}{=} \begin{bmatrix} u_{21} & u_{22} & \cdots & u_{2N_2} \end{bmatrix}, \quad (2.129)$$

U_{2N_2} jest serią identyfikującą na drugim stopniu.

Podobnie jak w punkcie 2.4.1, algorytmy identyfikacji na obu stopniach i identyfikacji bezpośredniej otrzymamy, posługując się empirycznymi rozkładami, uzyskanymi na podstawie wyników pomiarów, lub minimalizując empiryczne oszacowania wskaźników jakości identyfikacji. W tym drugim przypadku zamiast (5.117) minimalizujemy

$$Q_{1N_1n_2}(\theta_1) = \frac{1}{N_1} \sum_{n_1=1}^{N_1} q_1(y_{n_1n_2}, \Phi_1(u_{1n_1n_2}, \theta_1)) \quad (5.130)$$

względem θ_1 i otrzymujemy algorytm identyfikacji na pierwszym stopniu

$$\theta_{1N_1n_2}^* = \Psi_{1N_1}(U_{1N_1n_2}, Y_{N_1n_2}). \quad (5.131)$$

Zamiast (5.120) minimalizujemy

$$Q_{2N_2}(\theta_2) = \frac{1}{N_2} \sum_{n_2=1}^{N_2} q_2(\theta_{1N_1n_2}^*, \Phi_2(u_{2n_2}, \theta_2)) \quad (5.132)$$

względem θ_2 i otrzymujemy algorytm identyfikacji na drugim stopniu

$$\theta_{2N_2}^* = \Psi_{2N_2}(U_{2N_2}, \Xi_{1N_1N_2}^*), \quad (5.133)$$

gdzie

$$\Xi_{1N_1N_2}^* \stackrel{\text{df}}{=} [\theta_{1N_11}^* \quad \theta_{1N_12}^* \quad \cdots \quad \theta_{1N_1N_2}^*]. \quad (5.134)$$

Wreszcie, w wyniku minimalizacji

$$Q_{N_1N_2}(\theta_2) = \frac{1}{N_1N_2} \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} q_1(y_{n_1n_2}, \Phi(u_{1n_1n_2}, u_{2n_2}, \theta_2)) \quad (5.135)$$

względem θ_2 , otrzymujemy algorytm identyfikacji bezpośredniej

$$\tilde{\theta}_{2N_1N_2}^* = \tilde{\Psi}_{2N_1N_2}^*(\bar{U}_{1N_1N_2}, U_{2N_2}, \bar{Y}_{N_1N_2}), \quad (5.136)$$

gdzie:

$$\bar{U}_{1N_1N_2} \stackrel{\text{df}}{=} [U_{1N_11} \quad U_{1N_12} \quad \cdots \quad U_{1N_1N_2}], \quad \bar{Y}_{N_1N_2} \stackrel{\text{df}}{=} [Y_{N_11} \quad Y_{N_12} \quad \cdots \quad Y_{N_1N_2}]. \quad (5.137)$$

5.2.1. Przypadek liniowo-kwadratowy

Rozpatrzmy bliżej przypadek liniowo-kwadratowy, dla którego można uzyskać wyniki analityczne. Przyjmijemy dla obu stopni modele liniowe postaci:

$$\bar{y} = \Xi_1 u_1, \quad (5.138)$$

$$\bar{\theta}_{1l} = \Xi_{2l} u_2, \quad l=1, 2, \dots, L, \quad (5.139)$$

gdzie: θ_{1l}^T jest l -tym wierszem macierzy Ξ_1 , Ξ_1 jest $(L \times S_1)$ -macierzą parametrów na pierwszym stopniu, Ξ_{2l} są $(S_1 \times S_2)$ -macierzami parametrów na drugim stopniu.

W celu wprowadzenia wektorów parametrów θ_1 i θ_2 – jak w zapisie ogólnym (5.115) i (5.116) – zależności (5.138) i (5.139) można zapisać następująco:

$$\bar{y} = \Lambda_1(u_1) \theta_1, \quad (5.140)$$

$$\bar{\theta}_1 = \Lambda_2(u_2) \theta_2, \quad (5.141)$$

gdzie:

$$\theta_1^T \stackrel{\text{df}}{=} [\theta_{11}^T \quad \theta_{12}^T \quad \cdots \quad \theta_{1L}^T], \quad (R_1 = LS_1), \quad (5.142)$$

$$\theta_2^T \stackrel{\text{df}}{=} [\theta_{21}^T \quad \theta_{22}^T \quad \cdots \quad \theta_{2L}^T], \quad \theta_{2l}^T \stackrel{\text{df}}{=} [\theta_{2l1}^T \quad \theta_{2l2}^T \quad \cdots \quad \theta_{2lS_1}^T], \quad l=1, 2, \dots, L, \quad (5.143)$$

θ_{2lp}^T oznacza p -ty wiersz macierzy Ξ_{2l} , $p=1, 2, \dots, S_1$, a zatem ($R_2 = L S_1 S_2$), $A_1(u_1)$, $A_2(u_2)$ są macierzami – odpowiednio – o wymiarach ($L \times R_1$) i ($R_1 \times R_2$):

$$A_1(u_1) \stackrel{\text{df}}{=} \begin{bmatrix} u_1^T & 0_{S_1}^T & \cdots & 0_{S_1}^T \\ 0_{S_1}^T & u_1^T & \cdots & 0_{S_1}^T \\ \vdots & \vdots & \ddots & \vdots \\ 0_{S_1}^T & 0_{S_1}^T & \cdots & u_1^T \end{bmatrix}, \quad A_2(u_2) \stackrel{\text{df}}{=} \begin{bmatrix} u_2^T & 0_{S_2}^T & \cdots & 0_{S_2}^T \\ 0_{S_2}^T & u_2^T & \cdots & 0_{S_2}^T \\ \vdots & \vdots & \ddots & \vdots \\ 0_{S_2}^T & 0_{S_2}^T & \cdots & u_2^T \end{bmatrix}, \quad (5.144)$$

gdzie $0_{S_1}^T, 0_{S_2}^T$ są – odpowiednio – S_1 - oraz S_2 -wymiarowym, transponowanym wektorem zerowym, czyli

$$0_{S_1}^T = \overbrace{[0 \ 0 \ \cdots \ 0]}^{S_1}, \quad 0_{S_2}^T = \overbrace{[0 \ 0 \ \cdots \ 0]}^{S_2}.$$

Zgodnie z podanymi oznaczeniami, model bezpośredni (5.122) przyjmuje postać

$$\bar{y} = A_1(u_1) A_2(u_2) \theta_2. \quad (5.145)$$

Niech

$$q_1(y, \bar{y}) = (y - \bar{y})^T (y - \bar{y}), \quad (5.146)$$

$$q_2(\theta_1^*, \bar{\theta}_1) = (\theta_1^* - \bar{\theta}_1)^T (\theta_1^* - \bar{\theta}_1), \quad (5.147)$$

wówczas kryteria (5.117), (5.120) i (5.123) przyjmują – odpowiednio – postać:

$$Q_1(\theta_1) = E_{\underline{u}_1, \underline{y}} \left[\left(\underline{y} - A_1(\underline{u}_1) \theta_1 \right)^T \left(\underline{y} - A_1(\underline{u}_1) \theta_1 \right) \middle| u_2 \right], \quad (5.148)$$

$$Q_2(\theta_2) = E_{\underline{u}_2} \left[\left(\theta_1^* - A_2(\underline{u}_2) \theta_2 \right)^T \left(\theta_1^* - A_2(\underline{u}_2) \theta_2 \right) \right], \quad (5.149)$$

$$Q(\theta_2) = E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[\left(\underline{y} - A_1(\underline{u}_1) A_2(\underline{u}_2) \theta_2 \right)^T \left(\underline{y} - A_1(\underline{u}_1) A_2(\underline{u}_2) \theta_2 \right) \right]. \quad (5.150)$$

Przez przyrównanie do zera gradientów tych funkcji względem wektorów parametrów łatwo otrzymać następujące wyniki:

$$\theta_1^* = \left\{ E_{\underline{u}_1} \left[A_1^T(\underline{u}_1) A_1(\underline{u}_1) \middle| u_2 \right] \right\}^{-1} E_{\underline{u}_1, \underline{y}} \left[A_1^T(\underline{u}_1) \underline{y} \middle| u_2 \right], \quad (5.151)$$

$$\theta_2^* = \left\{ E_{\underline{u}_2} \left[A_2^T(\underline{u}_2) A_2(\underline{u}_2) \right] \right\}^{-1} E_{\underline{u}_2} \left[A_2^T(\underline{u}_2) \theta_1^* \right], \quad (5.152)$$

$$\tilde{\theta}_2^* = \left\{ E_{\underline{u}_1, \underline{u}_2} \left[A_2^T(\underline{u}_2) A_1^T(\underline{u}_1) A_1(\underline{u}_1) A_2(\underline{u}_2) \right] \right\}^{-1} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[A_2^T(\underline{u}_2) A_1^T(\underline{u}_1) \underline{y} \right]. \quad (5.153)$$

Ostateczny wynik dla zadania wyboru optymalnego modelu na drugim stopniu otrzymamy po wstawieniu (5.151) do (152). Zauważmy, że do wyznaczenia θ_1^* i $\tilde{\theta}_2^*$ wystarczy jedynie znajomość odpowiednich momentów rozkładu $f(u_1, u_2, y)$, natomiast ostateczne wyznaczenie wartości θ_2^* wymaga znajomości rozkładu brzegowego $f_2(u_2)$ i może być trudne lub niemożliwe do uzyskania analitycznie nawet dla najprostszych rozkładów, z powodu trudności z całkowaniem w drugim członie wyrażenia (5.152).

Zależności liniowe (5.138) i (5.139) dogodnie jest zapisywać z użyciem iloczynu Kroneckera macierzy [86]

$$A \otimes B \stackrel{\text{df}}{=} \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1N}B \\ a_{21}B & a_{22}B & \cdots & a_{2N}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1}B & a_{K2}B & \cdots & a_{KN}B \end{bmatrix},$$

gdzie $A = [a_{nk}]$ jest macierzą $(K \times N)$.

Wyniki (5.151), (5.153) oraz (5.154) można przepisać w innej formie, po podstawieniu:

$$A_1(u_1) = I_L \otimes u_1^T, \quad (5.154)$$

$$A_2(u_2) = I_L \otimes I_{S_1} \otimes u_2^T, \quad (5.155)$$

gdzie: I_L i I_{S_1} są – odpowiednio – macierzami jednostkowymi $(L \times L)$ oraz $(S_1 \times S_1)$.

Z zapisu tego skorzystamy w punkcie 5.2.2, w dowodzie twierdzenia o równoważności.

Rozpatrzmy szczególny przypadek $S_1 = L = R_1 = 1$, $S_2 = R_2$, tzn. zależności (5.138) i (5.139) mają postać:

$$\bar{y} = \theta_1 u_1, \quad (5.156)$$

$$\bar{\theta}_1 = \theta_2^T u_2, \quad (5.157)$$

wówczas wyniki (5.151), (5.152) oraz (5.153) są następujące:

$$\theta_1^* = \left\{ E \left[\underline{u}_1^2 \mid u_2 \right] \right\}^{-1} E \left[\underline{u}_1 \underline{y} \mid u_2 \right], \quad (5.158)$$

$$\theta_2^* = \left\{ E \left[\underline{u}_2 \underline{u}_2^T \right] \right\}^{-1} E \left[\underline{u}_2 \theta_1^* \right], \quad (5.159)$$

$$\tilde{\theta}_2^* = \left\{ E \left[\underline{u}_1^2 \underline{u}_2 \underline{u}_2^T \right] \right\}^{-1} E \left[\underline{u}_2 \underline{u}_1 \underline{y} \right]. \quad (5.160)$$

Aby ostatecznie wyznaczyć θ_2^* , należy obliczyć drugi człon (5.159), czyli

$$E \left(\underline{u}_2 \theta_1^* \right) = \int_{\underline{u}_2} \left\{ E \left[\underline{u}_1^2 \mid \underline{u}_2 \right] \right\}^{-1} E \left[\underline{u}_1 y \mid \underline{u}_2 \right] f_2(\underline{u}_2) d\underline{u}_2. \quad (5.161)$$

Trudności, o których była mowa, wiążą się z całkowaniem (5.161). Trudności takich nie ma natomiast w sytuacji empirycznej podczas wyznaczania algorytmu identyfikacji (5.133), bowiem obecnie dostępne są realizacje $\theta_{1N_1n_2}^*$ dla danego u_2 , czyli wyniki identyfikacji na pierwszym stopniu. Zgodnie z (5.151), (5.152) i (5.153) algorytm identyfikacji (5.131), (5.133) oraz (5.136) mają następującą postać, otrzymaną po wstawieniu momentów empirycznych:

$$\theta_{1N_1n_2}^* = \left[\sum_{n_1=1}^{N_1} A_1^T(u_{1n_1n_2}) A_1(u_{1n_1n_2}) \right]^{-1} \sum_{n_1=1}^{N_1} A_1^T(u_{1n_1n_2}) y_{n_1n_2}, \quad (5.162)$$

$$\theta_{2N_2}^* = \left[\sum_{n_2=1}^{N_2} A_2^T(u_{2n_2}) A_2(u_{2n_2}) \right]^{-1} \sum_{n_2=1}^{N_2} A_2^T(u_{2n_2}) \theta_{1N_1n_2}^*, \quad (5.163)$$

$$\begin{aligned} \tilde{\theta}_{2N_1N_2}^* &= \left[\sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} A_2^T(u_{2n_2}) A_1^T(u_{1n_1n_2}) A_1(u_{1n_1n_2}) A_2(u_{2n_2}) \right]^{-1} \\ &\quad \times \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} A_2^T(u_{2n_2}) A_1^T(u_{1n_1n_2}) y_{n_1n_2}. \end{aligned} \quad (5.164)$$

W szczególnym przypadku, gdy modele na pierwszym i drugim stopniu mają postać (5.156) i (5.157), algorytm identyfikacji są następujące:

$$\theta_{1N_1n_2}^* = \frac{\sum_{n_1=1}^{N_1} u_{1n_1n_2} y_{n_1n_2}}{\sum_{n_1=1}^{N_1} u_{1n_1n_2}^2}, \quad (5.165)$$

$$\theta_{2N_2}^* = \left(\sum_{n_2=1}^{N_2} u_{2n_2} u_{2n_2}^T \right)^{-1} \sum_{n_2=1}^{N_2} u_{2n_2} \theta_{1N_1n_2}^*, \quad (5.166)$$

$$\tilde{\theta}_{2N_1N_2}^* = \left(\sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{1n_1n_2}^2 u_{2n_2} u_{2n_2}^T \right)^{-1} \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} u_{2n_2} u_{1n_1n_2} y_{n_1n_2}. \quad (5.167)$$

Przykład 5.4. Niech w (5.156) oraz (5.157) $S_2 = R_2 = 1$, tzn.:

$$\bar{y} = \theta_1 u_1, \quad (5.168)$$

$$\bar{\theta}_1 = \theta_2 u_2. \quad (5.169)$$

Założymy rozkład normalny dla $(\underline{y}, \underline{u}_1, \underline{u}_2)$ z wektorem wartości oczekiwanych

$$E_{\underline{y}, \underline{u}_1, \underline{u}_2} [\underline{y}, \underline{u}_1, \underline{u}_2] = \mu = \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix} \quad (5.170)$$

oraz macierzą kowariancji

$$\text{cov}_{\underline{y}, \underline{u}_1, \underline{u}_2} [\underline{y}, \underline{u}_1, \underline{u}_2] = \Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{10} & \sigma_{20} \\ \sigma_{10} & \sigma_1^2 & \sigma_{12} \\ \sigma_{20} & \sigma_{12} & \sigma_2^2 \end{bmatrix}. \quad (5.171)$$

W celu wyznaczenia θ_1^* (5.158) należy ustalić odpowiednie momenty, które dla przyjętego rozkładu mają postać:

$$E_{u_1, y} [\underline{u}_1 y | u_2] = \sigma_0' \sigma_1' + \sigma_{10}' \quad (5.172)$$

oraz

$$E_{u_1} [u_1^2 | u_2] = (\mu_1')^2 + (\sigma_1')^2, \quad (5.173)$$

gdzie: μ_0' , μ_1' , σ_1' , σ_{10}' oznaczają – odpowiednio – momenty warunkowe, które można obliczyć ze znanych zależności dla rozkładu normalnego [79].

Po przekształceniach otrzymamy (5.172) postaci

$$E_{u_1, y} [\underline{u}_1 y | u_2] = \alpha_0 + \alpha_1 u_2 + \alpha_2 u_2^2, \quad (5.174)$$

gdzie:

$$\alpha_0 = \sigma_{10} - \frac{\sigma_{20} \sigma_{12}}{\sigma_2^2} + \mu_0 \mu_1 - \mu_2 \left(\mu_1 \frac{\sigma_{20}}{\sigma_2^2} + \mu_0 \frac{\sigma_{12}}{\sigma_2^2} \right) + \frac{\sigma_{20} \sigma_{12}}{\sigma_2^4} \mu_2^2, \quad (5.175)$$

$$\alpha_1 = \mu_1 \frac{\sigma_{20}}{\sigma_2^2} + \mu_0 \frac{\sigma_{12}}{\sigma_2^2} - 2 \mu_2 \frac{\sigma_{20} \sigma_{12}}{\sigma_2^4}, \quad (5.176)$$

$$\alpha_2 = \frac{\sigma_{20} \sigma_{12}}{\sigma_2^4}, \quad (5.177)$$

oraz (5.173)

$$E\left[u_1^2 \mid u_2\right] = \beta_0 + \beta_1 u_2 + \beta_2 u_2^2, \quad (5.178)$$

gdzie:

$$\beta_0 = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} + \left(\mu_1 - \frac{\sigma_{12}}{\sigma_2^2} \mu_2 \right)^2, \quad (5.179)$$

$$\beta_1 = 2 \left(\mu_1 - \frac{\sigma_{12}}{\sigma_2^2} \mu_2 \right) \frac{\sigma_{12}}{\sigma_2^2}, \quad (5.180)$$

$$\beta_2 = \frac{\sigma_{12}^2}{\sigma_2^4}. \quad (5.181)$$

Ostatecznie, zgodnie z (5.158),

$$\theta_1^* = \frac{\alpha_0 + \alpha_1 u_2 + \alpha_2 u_2^2}{\beta_0 + \beta_1 u_2 + \beta_2 u_2^2}. \quad (5.182)$$

Na podstawie (5.159)

$$\theta_2^* = \frac{1}{\sigma_2^2 + \mu_2^2} E \left[\frac{u_2}{\beta_0 + \beta_1 u_2 + \beta_2 u_2^2} \right]. \quad (5.183)$$

Wartości oczekiwanej względem u_2 nie można tu wyznaczyć analitycznie i należy stosować odpowiednią metodę całkowania numerycznego dla konkretnych wartości liczbowych momentów rozkładu normalnego.

W celu wyznaczenia $\tilde{\theta}_2^*$ (5.160) można skorzystać z poprzednich wyników

$$\tilde{\theta}_2^* = \frac{E \left[\frac{u_2}{u_2} E \left[\frac{u_1 y}{u_1} \mid u_2 \right] \right]}{E \left[\frac{u_2^2}{u_2} E \left[\frac{u_1^2}{u_1} \mid u_2 \right] \right]} = \frac{E \left[u_2 (\alpha_0 + \alpha_1 u_2 + \alpha_2 u_2^2) \right]}{E \left[u_2^2 (\beta_0 + \beta_1 u_2 + \beta_2 u_2^2) \right]}. \quad (5.184)$$

Po przekształceniach otrzymamy

$$\tilde{\theta}_2^* = \frac{\mu_0 \mu_1 \mu_2 + \mu_0 \sigma_{12} + \mu_1 \sigma_{20} + \mu_2 \sigma_{10}}{(\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2) + 4\mu_1 \mu_2 \sigma_{12} + 2\sigma_{12}^2}. \quad (5.185)$$

Zadanie znacznie się upraszcza, gdy $\sigma_{12} = 0$, wówczas $\alpha_2 = \beta_1 = \beta_2 = 0$ i θ_2^* , określone zależnością (5.183), można wyznaczyć analitycznie

$$\theta_2^* = \frac{\alpha_0 \mu_2 + \alpha_1 (\sigma_2^2 + \mu_2^2)}{(\sigma_2^2 + \mu_2^2) \beta_0} = \frac{\mu_0 \mu_1 \mu_2 + \mu_1 \sigma_{20} + \mu_2 \sigma_{10}}{(\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2)} = \tilde{\theta}_2^*. \quad (5.186)$$

Zwróćmy uwagę, że z przyjętym uproszczeniem, które oznacza, że wejścia na pierwszym i drugim stopniu są wartościami niezależnych zmiennych losowych, podejścia bezpośrednie i dwustopniowe są równoważne. \blacklozenge

Przykład 5.5. Niech $L = S_1 = S_2 = 1$, $R_1 = R_2 = 2$ oraz modele na pierwszym i drugim stopniu mają postać, odpowiednio:

$$\bar{y} = \theta_1^{(1)}u_1 + \theta_1^{(2)}, \quad (5.187)$$

$$\bar{\theta}_1^{(2)} = \theta_2^{(1)}u_2 + \theta_2^{(2)}, \quad (5.188)$$

z wektorami parametrów modeli:

$$\theta_1 = \begin{bmatrix} \theta_1^{(1)} \\ \theta_1^{(2)} \end{bmatrix}, \quad \theta_2 = \begin{bmatrix} \theta_2^{(1)} \\ \theta_2^{(2)} \end{bmatrix}. \quad (5.189)$$

Oznacza to, że u_2 wpływa tylko na wartość drugiej składowej wektora θ_1 . Zakładamy kryterium kwadratowe na obu stopniach oraz rozkład normalny dla $(\underline{y}, \underline{u}_1, \underline{u}_2)$, jak w przykładzie 5.4. Zadanie dla pierwszego stopnia jest znane [14] i jego wynik jest następujący:

$$\theta_1^{*(1)} = \frac{E_{\underline{u}_1, \underline{y}} [y \underline{u}_1 | u_2] - E_{\underline{y}} [y | u_2] E_{\underline{u}_1} [\underline{u}_1 | u_2]}{E_{\underline{u}_1} [\underline{u}_1^2 | u_2] - \left\{ E_{\underline{u}_1} [\underline{u}_1 | u_2] \right\}^2}, \quad (5.190)$$

$$\theta_1^{*(2)} = \frac{E_{\underline{y}} [y | u_2] E_{\underline{u}_1} [\underline{u}_1^2 | u_2] - E_{\underline{u}_1, \underline{y}} [y \underline{u}_1 | u_2] E_{\underline{u}_1} [\underline{u}_1 | u_2]}{E_{\underline{u}_1} [\underline{u}_1^2 | u_2] - \left\{ E_{\underline{u}_1} [\underline{u}_1 | u_2] \right\}^2}, \quad (5.191)$$

a po wstawieniu odpowiednich momentów i przekształceniach otrzymujemy:

$$\theta_1^{*(1)} = \frac{\sigma_{10}\sigma_2^2 - \sigma_{12}\sigma_{20}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}, \quad (5.192)$$

$$\theta_1^{*(2)} = \gamma_1 u_2 + \gamma_2 = G(u_2), \quad (5.193)$$

gdzie:

$$\gamma_1 = \frac{\sigma_{20}\sigma_1^2 - \sigma_{10}\sigma_{12}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}, \quad (5.194)$$

$$\gamma_2 = \mu_0 - \frac{(\sigma_{10}\sigma_2^2 - \sigma_{12}\sigma_{20})\mu_1 + (\sigma_{20}\sigma_1^2 - \sigma_{12}\sigma_{10})\mu_2}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}. \quad (5.195)$$

Porównując $G(u_2)$ (5.193) z przyjętym modelem $\Phi_2(u_2, \theta_2)$ postaci (5.188) na drugim stopniu, łatwo zauważyć, że $\theta_2^{*(1)} = \gamma_1$, $\theta_2^{*(2)} = \gamma_2$.

W podejściu bezpośrednim model (5.122) przyjmuje postać

$$\bar{y} = \theta_2^{(0)} u_1 + \theta_2^{(1)} u_2 + \theta_2^{(2)}, \quad (5.196)$$

gdzie $\theta_2^{(0)} = \theta_1^{(1)}$.

W wyniku minimalizacji względem $\theta_2^{(0)}$, $\theta_2^{(1)}$, $\theta_2^{(2)}$ funkcji

$$Q(\theta_2) = E_{u_1, u_2, y} \left[\left(y - \theta_2^{(0)} u_1 - \theta_2^{(1)} u_2 - \theta_2^{(2)} \right)^2 \right] \quad (5.197)$$

i po odpowiednich przekształceniach otrzymujemy

$$\tilde{\theta}_2^{*(0)} = \theta_1^{*(1)}, \quad \tilde{\theta}_2^{*(1)} = \gamma_1 = \theta_2^{*(1)}, \quad \tilde{\theta}_2^{*(2)} = \gamma_2 = \theta_2^{*(2)}. \quad (5.198)$$



Przykład 5.6. Dane są takie, jak w przykładzie 5.5, przy założeniu $\theta_2^{(2)} = 0$, tzn. modele na pierwszym i drugim stopniu mają postać, odpowiednio:

$$\bar{y} = \theta_1^{(1)} u_1 + \theta_1^{(2)}, \quad (5.199)$$

$$\bar{\theta}_1^{(2)} = \theta_2^{(1)} u_2, \quad (5.200)$$

a przy podejściu bezpośrednim

$$\bar{y} = \theta_2^{(0)} u_1 + \theta_2^{(1)} u_2, \quad (5.201)$$

gdzie $\theta_2^{(0)} = \theta_1^{(1)}$.

Wartości $\theta_1^{*(1)}$ i $\theta_1^{*(2)}$ są oczywiście takie jak w przykładzie poprzednim. Dla drugiego stopnia minimalizujemy względem $\theta_2^{(2)}$ funkcję

$$Q_2(\theta_2) = E_{u_2} \left[\left(\theta_1^{(2)} - \theta_2^{(1)} u_1 \right)^2 \right]. \quad (5.202)$$

Po wstawieniu $\theta_1^{*(2)}$ i przekształceniach otrzymujemy

$$\theta_2^{*(1)} = \gamma_1 + \gamma_2 \frac{\mu_2}{\mu_2 + \sigma_2^2}, \quad (5.203)$$

gdzie: γ_1 i γ_2 są takie jak w przykładzie 5.5

$$\gamma_2 = \mu_0 - \mu_1 \gamma_0 - \mu_2 \gamma_1, \quad \gamma_0 = \theta_1^{*(1)}. \quad (5.204)$$

W podejściu bezpośrednim minimalizujemy względem $\theta_2^{(0)}$ i $\theta_2^{(1)}$ funkcję

$$Q(\theta_2) = E_{u_1, u_2, y} \left[\left(y - \theta_2^{(0)} u_1 - \theta_2^{(1)} u_2 \right)^2 \right]. \quad (5.205)$$

Po przekształceniach otrzymujemy następujący wynik:

$$\tilde{\theta}_2^{*(0)} = \gamma_0 + \gamma_1 \frac{\mu_1(\sigma_2^2 + \mu_2^2) - \mu_2(\sigma_{12} + \mu_1\mu_2)}{(\mu_1^2 + \sigma_1^2)(\mu_2^2 + \sigma_2^2) - (\sigma_{12} + \mu_1\mu_2)^2}, \quad (5.206)$$

$$\tilde{\theta}_2^{*(1)} = \gamma_1 + \gamma_2 \frac{\mu_2(\sigma_1^2 + \mu_1^2) - \mu_1(\sigma_{12} + \mu_1\mu_2)}{(\mu_1^2 + \sigma_1^2)(\mu_2^2 + \sigma_2^2) - (\sigma_{12} + \mu_1\mu_2)^2}. \quad (5.207)$$



5.2.2. Porównanie podejścia dwustopniowego i bezpośredniego

Sformułujemy i udowodnimy twierdzenie podające warunek równoważności obu podejść. Sens tego twierdzenia polega na tym, że jeśli zależność (5.119), otrzymana w wyniku identyfikacji na pierwszym stopniu, ma tę samą postać co założony model (5.116) na drugim stopniu, to oba podejścia są równoważne.

Twierdzenie 5.4. Zakładamy, że rozwiązania zadań minimalizacji funkcji (5.117), (5.120) i (5.113), tj. – odpowiednio – $\theta_1^* = G(u_2)$, θ_2^* , $\tilde{\theta}_2^*$ są jednoznaczne. Jeśli istnieje taka wartość $\theta_2 \stackrel{\text{df}}{=} b \in \mathcal{O}_2$, że

$$\bigwedge_{u_2 \in \mathcal{U}_2} G(u_2) = \Phi_2(u_2, b), \quad (5.208)$$

to $\tilde{\theta}_2^* = \theta_2^* = b$.

Dowód: Widzimy, że jeśli θ_2^* jest rozwiązaniem jednoznacznym, to $\theta_2^* = b$, bo tylko wtedy \mathcal{Q}_2 przyjmuje minimalną wartość równą zeru. Wystarczy wykazać, że $\tilde{\theta}_2^* = b$. Zgodnie z (5.208)

$$\begin{aligned} \min_{\theta_1 \in \mathcal{O}_1} \mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \theta_1)) | u_2] &= \mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, G(u_2))) | u_2] \\ &= \mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \Phi_2(u_2, b))) | u_2] = \min_{\theta_2 \in \mathcal{O}_2} \mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \Phi_2(u_2, \theta_2))) | u_2]. \end{aligned} \quad (5.209)$$

Wynik minimalizacji względem θ_2 w ostatnim wyrażeniu równa się b i nie zależy od u_2 . Identyczny jest zatem wynik minimalizacji wartości oczekiwanej tego wyrażenia względem \underline{u}_2 , czyli

$$\begin{aligned} \min_{\theta_2 \in \mathcal{O}_2} \mathbb{E}_{\underline{u}_1, \underline{u}_2, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \Phi_2(\underline{u}_2, \theta_2)))] & \quad (5.210) \\ = \min_{\theta_2 \in \mathcal{O}_2} \mathbb{E}_{\underline{u}_2} \left[\mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \Phi_2(\underline{u}_2, \theta_2)))] | u_2 \right] &= \mathbb{E}_{\underline{u}_2} \left[\min_{\theta_2 \in \mathcal{O}_2} \mathbb{E}_{\underline{u}_1, \underline{y}} [q_1(\underline{y}, \Phi_1(\underline{u}_1, \Phi_2(\underline{u}_2, \theta_2)))] | u_2 \right]. \end{aligned}$$

Z porównania pierwszej minimalizacji w (5.210) z (5.124), po wstawieniu do (5.123) i uwzględnieniu jednoznaczności rozwiązania tego zadania θ_2^* , otrzymujemy $\tilde{\theta}_2^* = b = \theta_2^*$.

Q.E.D.

Zauważmy, że w przykładzie 5.5 spełniony był wystarczający warunek równoważności, tzn. $G(u_2)$ i $\Phi_2(u_2, \theta_2)$ miały tę samą postać i dlatego podejścia dwustopniowe i bezpośrednie dały te same wyniki.

W przypadku liniowo-kwadratowym łatwo pokazać, że równoważność występuje dla niezależnych wejść na obu stopniach.

Twierdzenie 5.5. Jeśli w przypadku liniowo-kwadratowym zmienne losowe u_1 i u_2 są niezależne, to $\theta_2^* = \tilde{\theta}_2^*$.

Dowód: W dowodzie skorzystamy z zapisu (5.154) i (5.155) oraz ze znanych własności iloczynu Kroneckera [86]

$$(A \otimes B)(C \otimes D) = AC \otimes BD, \quad (A \otimes B)e = Ae \otimes B, \quad (5.211)$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \quad (5.212)$$

gdzie: A, B, C, D są macierzami, e jest wektorem kolumnowym.

Po wstawieniu (5.151) do (5.152) i skorzystaniu z zapisu (5.154) i (5.155) oraz z założenia o niezależności otrzymujemy

$$\begin{aligned} \theta_2^* = & \left\{ E_{u_2} \left[\left(I_L \otimes I_{S_1} \otimes \underline{u}_2 \right) \left(I_L \otimes I_{S_1} \otimes \underline{u}_2^T \right) \right] \right\}^{-1} \\ & \times E_{u_2} \left\{ \left(I_L \otimes I_{S_1} \otimes \underline{u}_2 \right) \left\{ E_{u_1} \left[\left(I_L \otimes \underline{u}_1 \right) \left(I_L \otimes \underline{u}_1^T \right) \right] \right\}^{-1} E_{u_1, y} \left[\left(I_L \otimes \underline{u}_1 \right) y \mid u_2 \right] \right\}. \end{aligned} \quad (5.213)$$

Wyrażenie (5.213) przekształcamy, korzystając w odpowiednich miejscach z własności (5.211) i (5.212)

$$\begin{aligned} \theta_2^* = & \left\{ I_L \otimes I_{S_1} \otimes \left[E_{u_2} \left(\underline{u}_2 \underline{u}_2^T \right) \right]^{-1} \right\} \\ & \times E_{u_2} \left\{ \left(I_L \otimes I_{S_1} \otimes \underline{u}_2 \right) \left(I_L \otimes \left[E_{u_1} \left(\underline{u}_1 \underline{u}_1^T \right) \right]^{-1} \right) E_{u_1, y} \left[\left(y \otimes \underline{u}_1 \right) \mid u_2 \right] \right\} \\ = & \left\{ I_L \otimes I_{S_1} \otimes \left[E_{u_2} \left(\underline{u}_2 \underline{u}_2^T \right) \right]^{-1} \right\} E_{u_2} \left\{ \left(I_L \otimes \left[E_{u_1} \left(\underline{u}_1 \underline{u}_1^T \right) \right]^{-1} \otimes \underline{u}_2 \right) E_{u_1, y} \left[\left(y \otimes \underline{u}_1 \right) \mid u_2 \right] \right\}. \end{aligned} \quad (5.214)$$

W wyrażeniu tym możemy w odpowiednich miejscach dopisać macierze jednostkowe odpowiednich stopni, potrzebne do dalszych przekształceń, a następnie skorzystać z pierwszej własności (5.211), w której C i B są macierzami jednostkowymi odpowiednich rozmiarów

$$\begin{aligned} \theta_2^* &= \left(I_L \otimes I_{S_1} \otimes \left[E_{\underline{u}_2}(\underline{u}_2 \underline{u}_2^T) \right]^{-1} \right) \\ &\quad \times E_{\underline{u}_2} \left[\left(\left(I_L \otimes \left[E_{\underline{u}_1}(\underline{u}_1 \underline{u}_1^T) \right]^{-1} \right) (I_L \otimes I_{S_1}) \otimes I_{S_2} \underline{u}_2 \right) E_{\underline{u}_1, \underline{y}} \left[(\underline{y} \otimes \underline{u}_1) | \underline{u}_2 \right] \right] \quad (5.215) \\ &= \left(I_L \otimes I_{S_1} \otimes \left[E_{\underline{u}_2}(\underline{u}_2 \underline{u}_2^T) \right]^{-1} \right) \left(I_L \otimes \left[E_{\underline{u}_1}(\underline{u}_1 \underline{u}_1^T) \right]^{-1} \otimes I_{S_2} \right) E_{\underline{u}_2} \left\{ (I_L \otimes I_{S_1} \otimes \underline{u}_2) E_{\underline{u}_1, \underline{y}} \left[(\underline{y} \otimes \underline{u}_1) | \underline{u}_2 \right] \right\}. \end{aligned}$$

Korzystając z (5.211), wykonujemy dalsze przekształcenia

$$\theta_2^* = \left(I_L \otimes \left[E_{\underline{u}_1}(\underline{u}_1 \underline{u}_1^T) \right]^{-1} \otimes \left[E_{\underline{u}_2}(\underline{u}_2 \underline{u}_2^T) \right]^{-1} \right) E_{\underline{u}_1, \underline{u}_2, \underline{y}} (\underline{y} \otimes \underline{u}_1 \otimes \underline{u}_2). \quad (5.216)$$

Do wyrażenia (5.153) wstawiamy (5.154) oraz (5.155) i przekształcamy, korzystając w odpowiednich miejscach z własności (5.211) i (5.212)

$$\begin{aligned} \tilde{\theta}_2^* &= \left\{ E_{\underline{u}_1, \underline{u}_2} \left[(I_L \otimes I_{S_1} \otimes \underline{u}_2) (I_L \otimes \underline{u}_1) (I_L \otimes \underline{u}_1^T) (I_L \otimes I_{S_1} \otimes \underline{u}_2^T) \right] \right\}^{-1} \\ &\quad \times E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[(I_L \otimes I_{S_1} \otimes \underline{u}_2) (I_L \otimes \underline{u}_1) \underline{y} \right] \\ &= \left\{ E_{\underline{u}_1, \underline{u}_2} \left[(I_L \otimes \underline{u}_1 \otimes \underline{u}_2) (I_L \otimes \underline{u}_1^T \otimes \underline{u}_2^T) \right] \right\}^{-1} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[(I_L \otimes \underline{u}_1 \otimes \underline{u}_2) \underline{y} \right] \\ &= \left\{ E_{\underline{u}_1, \underline{u}_2} \left[I_L \otimes (\underline{u}_1 \otimes \underline{u}_2) (\underline{u}_1^T \otimes \underline{u}_2^T) \right] \right\}^{-1} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[(\underline{y} \otimes \underline{u}_1 \otimes \underline{u}_2) \right] \\ &= \left\{ I_L \otimes E_{\underline{u}_1}(\underline{u}_1 \underline{u}_1^T) \otimes E_{\underline{u}_2}(\underline{u}_2 \underline{u}_2^T) \right\}^{-1} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[(\underline{y} \otimes \underline{u}_1 \otimes \underline{u}_2) \right] \\ &= \left(I_L \otimes \left[E_{\underline{u}_1}(\underline{u}_1 \underline{u}_1^T) \right]^{-1} \otimes \left[E_{\underline{u}_2}(\underline{u}_2 \underline{u}_2^T) \right]^{-1} \right) E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left[(\underline{y} \otimes \underline{u}_1 \otimes \underline{u}_2) \right] = \theta_2^*. \quad (5.217) \end{aligned}$$

Q.E.D.

Równoważność łatwo sprawdzić wprost dla przypadku (5.156) i (5.157), wówczas po wstawieniu (5.158) do (5.159) otrzymujemy

$$\theta_2^* = \left\{ E_{\underline{u}_2} \left[\underline{u}_2 \underline{u}_2^T \right] \right\}^{-1} E_{\underline{u}_2} \left[\frac{E_{\underline{u}_1, \underline{y}} \left[\underline{u}_1 \underline{y} | \underline{u}_2 \right]}{E_{\underline{u}_1} \left[\underline{u}_1^2 | \underline{u}_2 \right]} \right]. \quad (5.218)$$

Zmienne losowe \underline{u}_1 i \underline{u}_2 są stochastycznie niezależne, zatem $E_{\underline{u}_1} \left[\underline{u}_1^2 | \underline{u}_2 \right] = E_{\underline{u}_1} \left[\underline{u}_1^2 \right]$.

W konsekwencji zależność (5.218) przyjmuje postać

$$\theta_2^* = \left\{ E_{\underline{u}_2} \left[\underline{u}_2 \underline{u}_2^T \right] \right\}^{-1} \frac{E_{\underline{u}_2} \left[\frac{E_{\underline{u}_1, \underline{y}} \left(\underline{u}_1 \underline{y} | \underline{u}_2 \right)}{E_{\underline{u}_1} \left(\underline{u}_1^2 \right)} \right]}{E_{\underline{u}_1} \left(\underline{u}_1^2 \right)}, \quad (5.219)$$

a po prostych przekształceniach

$$\theta_2^* = \frac{1}{E_{\underline{u}_1} \left[\underline{u}_1^2 \right]} \left\{ E_{\underline{u}_2} \left[\underline{u}_2 \underline{u}_2^T \right] \right\}^{-1} E_{\underline{u}_1, \underline{u}_2, \underline{y}} \left(\underline{u}_2, \underline{u}_1 \underline{y} \right) = \tilde{\theta}_2^*. \quad (5.220)$$

Przedstawione w podrozdziałach 5.1 oraz 5.2 dwustopniowe zadania identyfikacji obiektów statycznych można przenieść na dwustopniowe zadanie identyfikacji obiektów dynamicznych [109, 110, 118].

Zadanie identyfikacji obiektów dynamicznych z wykorzystaniem identyfikacji dwustopniowej rozwinięto w pracach [5, 66, 116, 121, 125]. Zwrócono tam uwagę na rekurencyjne algorytmy identyfikacji dwustopniowej, przydatne do projektowania adaptacyjnych algorytmów sterowania. W pracach [9, 10, 13, 119, 120] podjęto problemy, w których pierwszy stopień ma charakter pomocniczy i odgrywa rolę systemu pomiarowego. Uzyskane wyniki są przydatne do projektowania dwupoziomowych systemów wspomagania decyzji.

5.3. Identyfikacja destylacyjnej kolumny wypełnionej z pulsacją fazy parowej

W celu ilustracji zastosowania identyfikacji wielostopniowej do tworzenia modelu matematycznego wrócimy do przykładu wstępnie opisanego w podrozdziale 1.2, a dotyczącego badań destylacyjnych kolumn wypełnionych z pulsacją fazy parowej, przedstawionych w pracach [76, 77]. Analiza zjawisk towarzyszących przepływowi faz i wymianie masy pokazuje, że intensyfikację wymiany masy w układzie ciec–gaz osiąga się przede wszystkim przez zwiększenie powierzchni kontaktu faz i zmniejsze-

nie sumarycznego oporu przenikania masy, czyli zakłócenie stabilności warstw granicznych na powierzchni styku. Efekt ten można uzyskać w wyniku zwiększenia burzliwości naturalnego przepływu faz po ustawieniu przeszkód na drodze naturalnego przepływu, jak również przez wywołanie dodatkowego, zewnętrznego zaburzenia przepływu. Spostrzeżenia te są podstawą do projektowania kolumn destylacyjnych wypełnionych z pulsacją fazy parowej (rys. 1.6).

Często stosowaną miarą sprawności kolumny jest liczba pólk teoretycznych lub wartość objętościowego współczynnika przenikania masy. Wielkości te zależą przede wszystkim od parametrów geometrycznych kolumny oraz natężenia przepływu przez nią destylowanej mieszaniny. Sprawność kolumny można poprawić dzięki zastosowaniu odpowiednio dobranych wypełnień oraz przez wprowadzenie pulsacji fazy parowej. Cel badań prowadzonych w Instytucie Inżynierii Chemicznej Urzędzeń Ciepłych Politechniki Wrocławskiej [76] to określenie wpływu wypełnienia oraz parametrów pulsacji na sprawność kolumny destylacyjnej.

W rozpatrywanym przykładzie ograniczymy się do przedstawienia zastosowania identyfikacji wielostopniowej do badania względnego przyrostu wartości objętościowego współczynnika przenikania masy w wypełnionej kolumnie destylacyjnej z pulsacją fazy parowej. Wyniki badań względnego przyrostu liczby pólk teoretycznych kolumn destylacyjnych wypełnionych z pulsacją fazy parowej przedstawiono w pracach [76, 77]. Zastosowanie identyfikacji wielostopniowej do rozpatrywanego zadania wynika z organizacji eksperymentu. W pierwszej kolejności, dla zadanej serii wartości natężenia strumienia przepływu u_1 , mierzono skład fazy parowej i na tej podstawie określano wartość objętościowego współczynnika przenikania masy ρ . Pozwoliło to na określenie zależności sprawności kolumny z ustalonym wypełnieniem od strumienia przepływu. Następnie powtórzono badania z pulsacją fazy parowej. Pomiary prowadzono dla zadanej serii wartości częstotliwości pulsacji u_2 , przy ustalonej amplitudzie pulsacji. W kolejnym etapie eksperymentu badania powtórzono dla zadanej serii wartości amplitudy pulsacji u_3 . Badania dotyczyły kolumn z różnym wypełnieniem. Na podstawie uzyskanych danych pomiarowych opracowano odpowiednie modele matematyczne kolumn wypełnionych z pulsacją fazy parowej.

Przykład ten pokazuje możliwość zastosowania identyfikacji wielostopniowej do tworzenia modelu matematycznego obiektu. Szczegółowo przedstawimy zadanie identyfikacji dwustopniowej do stworzenia modelu kolumn destylacyjnych wypełnionych z pulsacją fazy parowej, a następnie omówimy rozbudowę modelu na kolejnych stopniach przez uwzględnienie kolejnych wyników badań.

Wprowadzamy oznaczenia:

u_1 – natężenie strumienia przepływu destylowanej mieszaniny przez kolumnę, $\text{m}^3 \cdot \text{h}^{-1}$,

u_2 – częstotliwość pulsacji, Hz,

ρ – objętościowy współczynnik przenikania masy w kolumnie destylacyjnej z pulsacją fazy parowej, $\text{kg} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$,

ρ_0 – objętościowy współczynnik przenikania masy w kolumnie destylacyjnej ze swobodnym przepływem (bez pulsacji), $\text{kg} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$,

y – względny przyrost wartości objętościowego współczynnika przenikania masy w kolumnie destylacyjnej z pulsacją fazy parowej, określony zależnością

$$y = \frac{\rho - \rho_0}{\rho_0}. \quad (5.221)$$

W tym przypadku zadanie identyfikacji polega na wyznaczeniu zależności względnego przyrostu wartości objętościowego współczynnika przenikania masy od natężenia przepływu destylowanej mieszaniny przez kolumnę oraz częstotliwości pulsacji, tj. należy ustalić zależność y od u_1 oraz u_2 .

5.3.1. Opis danych pomiarowych

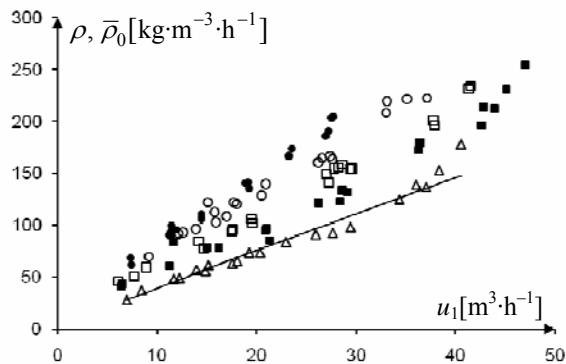
Badania eksperymentalne wykonano na kolumnie destylacyjnej, o wysokości 1 m i średnicy 0,15 m, z wypełnieniem pierścieniami Raschiga. Stosowano mieszaninę 80% n-heptanu i 20% toluenu. W pierwszym etapie badań kolumny ze swobodnym przepływem (bez pulsacji) dla różnych wartości natężenia strumienia przepływu zmierzono stężenie mieszaniny n-heptan–toluen na wyjściu kolumny. Na tej podstawie wyznaczono wartość objętościowego współczynnika przenikania masy dla różnych wartości natężenia strumienia przepływu. Następnie eksperyment powtórzono dla różnych wartości częstotliwości pulsacji. Pomiary wykonano w Instytucie Inżynierii Chemicznej i Urządzeń Ciepłych Politechniki Wrocławskiej. Przykładowe wyniki pomiarów dla ustalonej amplitudy pulsacji przedstawiono w tabeli 5.1 oraz na rysunku 5.8.

Dla pomiarów przedstawionych w tabeli 5.1 przyjęto następujące oznaczenia:

- $u_{1n_1,0}$ – natężenie strumienia przepływu destylowanej mieszaniny przez kolumnę bez pulsacji, $n_1 = 1, 2, \dots, N_1$,
- $u_{1n_1n_2}$ – natężenie strumienia przepływu destylowanej mieszaniny przez kolumnę dla zadanej wartości częstotliwości pulsacji u_{2n_2} , $n_1 = 1, 2, \dots, N_1$, $n_2 = 1, 2, \dots, N_2$,
- u_{2n_2} – częstotliwość pulsacji, $n_2 = 1, 2, \dots, N_2$,
- $\rho_{n_1,0}$ – objętościowy współczynnik przenikania masy dla zadanego natężenia strumienia przepływu $u_{1n_1,0}$ przez kolumnę bez pulsacji, $n_1 = 1, 2, \dots, N_1$,
- $\rho_{n_1n_2}$ – objętościowy współczynnik przenikania masy dla ustalonego natężenia strumienia przepływu $u_{1n_1n_2}$ oraz zadanej częstotliwości pulsacji u_{2n_2} , $n_1 = 1, 2, \dots, N_1$, $n_2 = 1, 2, \dots, N_2$,
- N_1 – liczba różnych wartości natężenia przepływu destylowanej mieszaniny (zakładamy, że jest ona taka sama w przypadku kolumny bez pulsacji, jak i kolumny z różnymi parametrami pulsatora),
- N_2 – liczba zadanych wartości częstotliwości pulsacji.

Tabela 5.1. Amplituda pulsacji $u_3 = 0,045$ m, $N_2 = 4$, $N_1 = 20$

n_2	Bez pulsacji (0)		$u_{21} = 5,33$ (1)		$u_{22} = 7,00$ (2)		$u_{23} = 8,67$ (3)		$u_{24} = 10,00$ (4)	
n_1	$u_{1n_1 0}$	$\rho_{n_1 0}$	$u_{1n_1 1}$	$\rho_{n_1 1}$	$u_{1n_1 2}$	$\rho_{n_1 2}$	$u_{1n_1 3}$	$\rho_{n_1 3}$	$u_{1n_1 4}$	$\rho_{n_1 4}$
1	6,9	27,4	6,4	40,7	6,1	45,5	9,2	69,4	7,3	68,3
2	8,4	36,8	6,5	43,3	7,7	50,2	12,2	91,2	7,4	61,9
3	12,2	48,5	11,2	60,0	8,9	58,9	12,6	92,8	11,3	92,4
4	11,6	47,5	11,6	83,5	14,2	84,0	13,9	95,8	11,2	90,5
5	13,9	56,3	15,0	78,5	14,7	77,7	15,8	112,3	11,4	99,3
6	15,1	60,8	16,2	77,9	17,5	93,2	15,9	102,4	11,4	89,7
7	14,8	55,3	20,9	94,2	17,6	95,7	17,0	108,4	11,4	94,2
8	17,5	62,5	21,0	96,2	19,5	105,5	17,7	121,6	11,9	95,6
9	18,0	64,5	21,3	85,1	19,6	101,2	18,0	120,4	14,4	110,7
10	19,2	72,9	26,2	121,4	27,0	149,4	15,1	121,6	14,4	105,4
11	20,3	73,1	28,4	122,8	27,3	141,0	20,5	128,2	14,4	107,7
12	22,9	83,2	28,6	132,8	27,8	154,6	20,9	139,5	18,8	140,6
13	25,9	90,4	29,1	131,5	28,2	156,0	26,2	159,9	19,1	141,5
14	27,6	92,1	36,4	179,3	28,6	158,4	26,6	164,5	19,2	135,1
15	29,4	97,5	36,3	171,7	29,4	154,0	27,4	166,4	23,5	173,5
16	34,3	124,4	42,8	213,6	29,6	154,6	27,6	163,3	23,2	166,1
17	36,0	138,3	42,6	195,5	37,8	201,0	33,1	219,2	26,9	185,4
18	37,0	136,1	43,9	212,1	37,9	196,3	33,0	208,3	27,2	190,4
19	38,3	152,4	45,1	231,0	41,3	231,5	35,1	221,3	27,5	203,1
20	40,5	177,4	47,0	253,8	41,5	234,0	37,1	221,9	27,7	204,2



Rys. 5.8. Wyniki eksperymentu – wartość objętościowego współczynnika przenikania masy dla ustalonego natężenia strumienia przepływu oraz ustalonych częstotliwości pulsacji: (■) – $u_{21} = 5,33$ Hz, (□) – $u_{22} = 7,00$ Hz, (○) – $u_{23} = 8,67$ Hz, (●) – $u_{24} = 10,00$ Hz oraz (Δ) – bez pulsacji. Linią ciągłą zaznaczono pomocniczą zależność (5.228)

Zwróćmy uwagę, że wartości objętościowego współczynnika przenikania masy spowodowane pulsacją są wyznaczone niekoniecznie dla tych samych wartości natężenia przepływu co wartości objętościowego współczynnika przenikania masy bez pulsacji (tj. $u_{1n_1,0} \neq u_{1n_1n_2}$, $n_2 = 1, 2, \dots, N_2$, tab. 5.1). Jest to spowodowane względami technicznymi prowadzenia eksperymentu. Nie jest możliwe zadanie dokładnej wartości natężenia strumienia przepływu destylowanej mieszaniny przez kolumnę. Wartość tę można zadać w przybliżeniu poprzez ustawienie wartości temperatury kotła i zmierzyć faktyczną wartość natężenia strumienia przepływu w stanie ustalonym. W celu wyznaczenia przyrostu wartości objętościowego współczynnika przenikania masy dla zmierzonych wartości natężenia przepływu należy rozwiązać pomocnicze zadanie identyfikacji, polegające na wyznaczeniu zależności między objętościowym współczynnikiem przenikania masy w kolumnie bez pulsacji a natężeniem strumienia przepływu, tj. należy wyznaczyć optymalny wektor parametrów modelu z klasy

$$\bar{\rho}_0 = \Phi_0(u_1, \theta_0) = \theta_0^{(2)} u_1 \theta_0^{(1)}, \quad (5.222)$$

gdzie: $\theta_0 = [\theta_0^{(1)} \quad \theta_0^{(2)}]^T$ jest wektorem parametrów, a $\bar{\rho}_0$ – wyjściem modelu. Postać funkcji Φ_0 ustalono na podstawie analizy danych pomiarowych przedstawionych w kolumnie (0) tabeli 5.1. Zależność ta będzie pomocna w dalszej części do wyznaczenia względnego przyrostu wartości objętościowego współczynnika przenikania masy dla dowolnej wartości natężenia strumienia przepływu.

Przyjmujemy kryterium jakości identyfikacji

$$Q_0(\theta_0) = \sum_{n_1=1}^{N_1} \left(\ln \rho_{n_1,0} - \ln \left(\theta_0^{(2)} u_{1n_1,0} \theta_0^{(1)} \right) \right)^2 = \sum_{n_1=1}^{N_1} \left(\ln \rho_{n_1,0} - \ln \theta_0^{(2)} - \theta_0^{(1)} \ln u_{1n_1,0} \right)^2. \quad (5.223)$$

Taka postać kryterium pozwala na analityczne wyznaczenie algorytmu identyfikacji. W wyniku minimalizacji kryterium (5.223) względem θ_0 otrzymujemy

$$\theta_{0N_1}^* = \begin{bmatrix} \theta_{0N_1}^{*(1)} \\ \theta_{0N_1}^{*(2)} \end{bmatrix} = \begin{bmatrix} \frac{A_{0N_1}^{(1)}}{B_{0N_1}} \\ \exp \left(\frac{A_{0N_1}^{(2)}}{B_{0N_1}} \right) \end{bmatrix}, \quad (5.224)$$

gdzie $\theta_{0N_1}^*$ to optymalny wektor parametrów modelu (5.222),

$$A_{0N_1}^{(1)} = \sum_{n_1=1}^{N_1} \ln \rho_{n_1,0} \ln u_{1n_1,0} - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln \rho_{n_1,0} \right) \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1,0} \right), \quad (5.225)$$

$$A_{0N_1}^{(2)} = \frac{1}{N_1} \sum_{n_1=1}^{N_1} (\ln u_{1n_1 0})^2 \sum_{n_1=1}^{N_1} \ln \rho_{n_1 0} - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1 0} \right) \left(\sum_{n_1=1}^{N_1} \ln \rho_{n_1 0} \ln u_{1n_1 0} \right), \quad (5.226)$$

$$B_{0N_1} = \sum_{n_1=1}^{N_1} (\ln u_{1n_1 0})^2 - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1 0} \right)^2. \quad (5.227)$$

Dla danych pomiarowych przedstawionych w kolumnie (0) tabeli 5.1 otrzymujemy $\theta_{0N_1}^{*(1)} = 0,94$, $\theta_{0N_2}^{*(2)} = 4,49$, a optymalny model z zadanej klasy (5.222) ma postać

$$\bar{\rho}_0 = \Phi_0(u_1, \theta_{0N_1}^*) = 4,49 u_1^{0,94}. \quad (5.228)$$

Wynik zilustrowano na rysunku 5.8.

Tabela 5.2. Amplituda pulsacji $u_3 = 0,045$ m, $N_2 = 4$, $N_1 = 20$

n_2	$u_{21} = 5,33$ (1)		$u_{22} = 7,00$ (2)		$u_{23} = 8,67$ (3)		$u_{24} = 10,00$ (4)	
n_1	$u_{1n_1 1}$	$y_{n_1 1}$	$u_{1n_1 2}$	$y_{n_1 2}$	$u_{1n_1 3}$	$y_{n_1 3}$	$u_{1n_1 4}$	$y_{n_1 4}$
1	6,4	0,572060	6,1	0,838889	9,2	0,903488	7,3	1,3301716
2	6,5	0,648202	7,7	0,628602	12,2	0,916698	7,4	1,0848920
3	11,2	0,366938	8,9	0,666820	12,6	0,891862	11,3	1,0875064
4	11,6	0,840378	14,2	0,529828	13,9	0,780235	11,2	1,0617987
5	15,0	0,357619	14,7	0,369640	15,8	0,849268	11,4	1,2248224
6	16,2	0,252894	17,5	0,393696	15,9	0,676236	11,4	1,0097338
7	20,9	0,191408	17,6	0,423408	17,0	0,665933	11,4	1,1105566
8	21,0	0,211237	19,5	0,424521	17,7	0,798994	11,9	1,0569201
9	21,3	0,057237	19,6	0,359882	18,0	0,753221	14,4	0,9896686
10	26,2	0,240598	27,0	0,484021	15,1	1,089871	14,4	0,8944089
11	28,4	0,162991	27,3	0,386058	20,5	0,651258	14,4	0,9357480
12	28,6	0,249399	27,8	0,493950	20,9	0,764347	18,8	0,9650770
13	29,1	0,217105	28,2	0,487298	26,2	0,634033	19,1	0,9483388
14	36,4	0,343625	28,6	0,490247	26,6	0,657183	19,2	0,8510747
15	36,3	0,290017	29,4	0,411630	27,4	0,630113	23,5	0,9645854
16	42,8	0,373851	29,6	0,408095	27,6	0,588806	23,2	0,9037284
17	42,6	0,263002	37,8	0,453555	33,1	0,796697	26,9	0,8480748
18	43,9	0,331933	37,9	0,416033	33,0	0,712234	27,2	0,8781611
19	45,1	0,414180	41,3	0,539947	35,1	0,716245	27,5	0,9828131
20	47,0	0,494438	41,5	0,549499	37,1	0,633244	27,7	0,9799704

Zależność (5.228) zastosujemy do wyznaczenia przyrostu wartości współczynnika (5.221), tj. do obliczenia względnego przyrostu wartości objętościowego współczyn-

nika przenikania masy $y_{n_1 n_2}$, spowodowanego pulsacją o częstotliwości u_{2n_2} dla zadanego natężenia strumienia przepływu $u_{1n_1 n_2}$, czyli

$$y_{n_1 n_2} = \frac{\rho_{n_1 n_2} - \rho_{n_1 n_2}^*}{\rho_{n_1 n_2}^*}, \quad (5.229)$$

gdzie

$$\rho_{n_1 n_2}^* \stackrel{\text{df}}{=} \Phi_0(u_{1n_1 n_2}, \theta_{0N_1}^*), \quad (5.230)$$

a zatem

$$y_{n_1 n_2} = \frac{\rho_{n_1 n_2} - \Phi_0(u_{1n_1}, \theta_{0N_1}^*)}{\Phi_0(u_{1n_1}, \theta_{0N_1}^*)}. \quad (5.231)$$

Tak obliczone wartości $y_{n_1 n_2}$, $n_1 = 1, 2, \dots, N_1$, $n_2 = 1, 2, \dots, N_2$, (tab. 5.2), w dalszych rozważaniach wykorzystamy do wyznaczenia wpływu wypełnienia oraz parametrów pulsacji na poprawę sprawności kolumny destylacyjnej.

5.3.2. Algorytmy identyfikacji i wyniki obliczeń

Organizacja eksperymentu umożliwia zastosowanie identyfikacji dwustopniowej. Przyjmijemy strumień natężenia przepływu destylowanej mieszaniny u_1 za wejście na 1. stopniu, a przyrost wartości objętościowego współczynnika przenikania masy y za wyjście badanego obiektu. Zadanie identyfikacji na 1. stopniu polega na zbadaniu zależności pomiędzy wielkościami u_1 a y , dla ustalonych wartości częstotliwości pulsacji u_2 . Dla przyjętej klasy – model Φ_1 na 1. stopniu – zadanie sprowadza się do wyznaczenia wartości optymalnych wektora parametrów modelu. Dla zadanej serii wartości częstotliwości pulsacji u_2 otrzymamy zbiór optymalnych wartości wektora parametrów θ_1 modelu na 1. stopniu. Teraz odpowiednie zadanie identyfikacji na 2. stopniu polega na zbadaniu zależności pomiędzy częstotliwością pulsacji u_2 a wektorem parametrów modelu θ_1 wyznaczonym na 1. stopniu. Częstotliwość pulsacji u_2 jest tu wejściem obiektu na 2. stopniu, a optymalna wartość wektora parametrów modelu na 1. stopniu θ_1 jest wyjściem obiektu na 2. stopniu. Dla przyjętej klasy – model Φ_2 na 2. stopniu – zadanie sprowadza się do wyznaczenia optymalnej wartości wektora parametrów θ_2 modelu na 2. stopniu.

Po złożeniu funkcji Φ_1 i Φ_2 otrzymamy poszukiwaną zależność pomiędzy natężeniem strumienia przepływu, częstotliwością pulsacji a przyrostem wartości objętościowego współczynnika przenikania masy.

Zależność tę możemy teraz badać metodą bezpośrednią, a zadanie identyfikacji na 1. i 2. stopniu potraktować jako zadanie pomocnicze podczas ustalania postaci funkcji Φ jako złożenia funkcji Φ_1 i Φ_2 , a optymalną wartość parametrów modelu wyzna-

czoną na 2. stopniu jako pierwsze przybliżenie wartości wektora parametrów modelu bezpośredniego Φ .

Dwustopniowe algorytmy identyfikacji

Na podstawie analizy zmian wartości przyrostu objętościowego współczynnika przenikania masy dla różnych wartości natężenia przepływu oraz ustalonej częstotliwości pulsacji przyjmujemy następującą postać funkcji Φ_1 w modelu na 1. stopniu

$$\bar{y} = \Phi_1(u_1, \theta) = \theta_1^{(2)} u_1^{\theta_1^{(1)}}, \quad (5.232)$$

gdzie $\theta_1 = [\theta_1^{(1)} \quad \theta_1^{(2)}]^T$ jest wektorem parametrów modelu.

Zadanie identyfikacji na 1. stopniu sprowadza się do wyznaczenia optymalnych wartości parametrów $\theta_1^{(1)}$ i $\theta_1^{(2)}$ dla ustalonej częstotliwości pulsacji $u_2 = u_{2n_2}$.

Przyjmujemy kryterium jakości identyfikacji na 1. stopniu

$$\begin{aligned} Q_{1N_1n_2}(\theta_1) &= \sum_{n_1=1}^{N_1} \left(\ln y_{n_1n_2} - \ln \left(\theta_1^{(2)} u_{1n_1n_2}^{\theta_1^{(1)}} \right) \right)^2 \\ &= \sum_{n_1=1}^{N_1} \left(\ln y_{n_1n_2} - \ln \theta_1^{(2)} - \theta_1^{(1)} \ln u_{1n_1n_2} \right)^2. \end{aligned} \quad (5.233)$$

Optymalną wartość wektora parametrów modelu z klasy (5.232), która minimalizuje kryterium (5.233) względem θ_1 , wyznaczamy ze wzoru

$$\theta_{1N_1n_2}^* = \begin{bmatrix} \theta_{1N_1n_2}^{*(1)} \\ \theta_{1N_1n_2}^{*(2)} \end{bmatrix} = \begin{bmatrix} \frac{A_{1N_1n_2}^{(1)}}{B_{1N_1n_2}} \\ \exp \left(\frac{A_{1N_1n_2}^{(2)}}{B_{1N_1n_2}} \right) \end{bmatrix}, \quad (5.234)$$

w którym $\theta_{1N_1n_2}^*$ jest optymalną wartością wektora parametrów modelu (5.232),

$$A_{1N_1n_2}^{(1)} = \sum_{n_1=1}^{N_1} \ln y_{n_1n_2} \ln u_{1n_1n_2} - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln y_{n_1n_2} \right) \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1n_2} \right), \quad (5.235)$$

$$A_{1N_1n_2}^{(2)} = \frac{1}{N_1} \sum_{n_1=1}^{N_1} (\ln u_{1n_1n_2})^2 \sum_{n_1=1}^{N_1} \ln y_{n_1n_2} - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1n_2} \right) \left(\sum_{n_1=1}^{N_1} \ln y_{n_1n_2} \ln u_{1n_1n_2} \right), \quad (5.236)$$

$$B_{1N_1n_2} = \sum_{n_1=1}^{N_1} (\ln u_{1n_1n_2})^2 - \frac{1}{N_1} \left(\sum_{n_1=1}^{N_1} \ln u_{1n_1n_2} \right)^2. \quad (5.237)$$

Jest to algorytm identyfikacji na 1. stopniu.

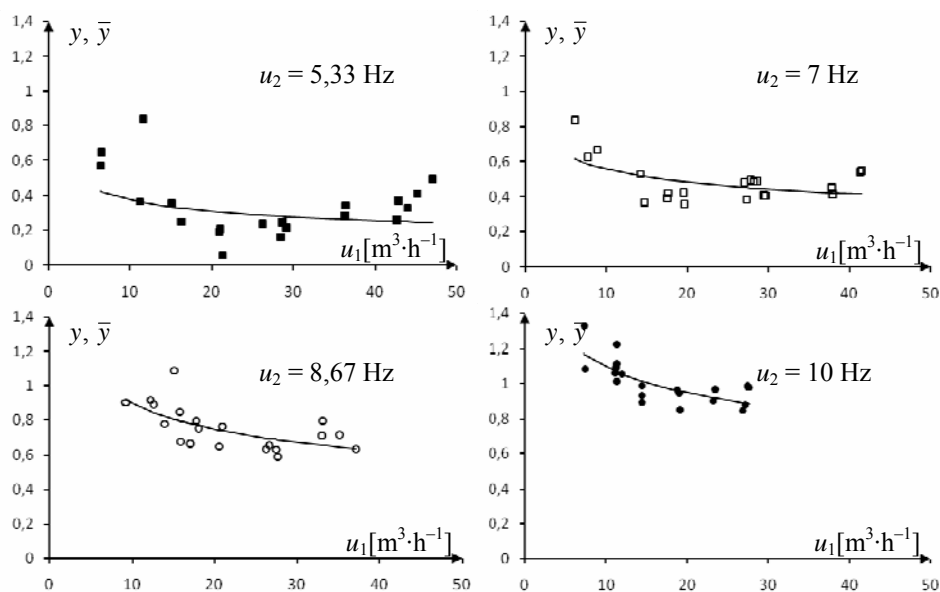
Dla danych pomiarowych przedstawionych w tabeli 5.1 i odpowiadających im wartości względnego przyrostu objętościowego współczynnika przenikania masy (tab. 5.2) wyznaczono optymalne wartości składowych wektora parametrów modelu (5.232) dla różnych wartości częstotliwości pulsacji. Wyniki przedstawiono w tabeli 5.3 oraz zilustrowano na rysunku 5.9.

Tabela 5.3. Optymalne wartości składowych wektora parametrów modelu (5.232) dla różnych wartości częstotliwości pulsacji

n_2	1	2	3	4
u_{2n_2}	5,33	7,00	8,67	10,0
$\theta_{1N_1n_2}^{*(1)}$	-0,274	-0,203	-0,260	-0,207
$\theta_{1N_1n_2}^{*(2)}$	0,707	0,886	1,635	1,767

Tym sposobem na 1. stopniu otrzymujemy zbiór optymalnych modeli dla różnych wartości częstotliwości pulsacji, czyli dla $u_2 = u_{2n_2}$

$$\bar{y} = \theta_{1N_1n_2}^{*(2)} u_1^{\theta_{1N_1n_2}^{*(1)}}, n_2 = 1, 2, \dots, N_2. \quad (5.238)$$



Rys. 5.9. Wartość względnego przyrostu objętościowego współczynnika przenikania masy dla zadanego natężenia strumienia przepływu, spowodowanego pulsacją o częstotliwości pulsacji:

(\blacksquare) – $u_{21} = 5,33 \text{ Hz}$, (\square) – $u_{22} = 7,00 \text{ Hz}$, (\circ) – $u_2 = 8,67 \text{ Hz}$, (\bullet) – $u_2 = 10,00 \text{ Hz}$.

Linia ciągłą zaznaczono zależność (5.228) z optymalnymi parametrami wyznaczonymi na 1. stopniu zawartymi w tabeli 5.3

Na podstawie analizy zmian wartości wektora optymalnych parametrów modelu (5.232) przy zmianie częstotliwości pulsacji (tab. 5.3) zaproponowano model na 2. stopniu

$$\bar{\theta}_1 = \Phi_2(u_2, \theta_2) = \begin{bmatrix} \theta_2^{(1)} \\ \theta_2^{(3)} u_2^{\theta_2^{(2)}} \end{bmatrix}, \quad (5.239)$$

gdzie $\theta_2 = [\theta_2^{(1)} \quad \theta_2^{(2)} \quad \theta_2^{(3)}]^T$ jest wektorem parametrów modelu.

Zadanie identyfikacji na 2. stopniu polega na wyznaczeniu optymalnych wartości wektora parametrów modelu (5.239). Przyjmujemy kryterium jakości identyfikacji

$$\begin{aligned} Q_{2N_2}(\theta_2) &= \sum_{n_2=1}^{N_2} \left(\left(\theta_{1N_1n_2}^{*(1)} - \theta_2^{(1)} \right)^2 + \left(\ln \theta_{1N_1n_2}^{*(2)} - \ln \left(\theta_2^{(3)} u_{2n_2}^{\theta_2^{(2)}} \right) \right)^2 \right) \\ &= \sum_{n_2=1}^{N_2} \left(\left(\theta_{1N_1n_2}^{*(1)} - \theta_2^{(1)} \right)^2 + \left(\ln \theta_{1N_1n_2}^{*(2)} - \ln \theta_2^{(3)} - \theta_2^{(2)} \ln u_{2n_2} \right)^2 \right). \end{aligned} \quad (5.240)$$

W wyniku minimalizacji kryterium (5.240) względem θ_2 otrzymujemy

$$\theta_{2N_2}^* = \begin{bmatrix} \theta_{2N_2}^{*(1)} \\ \theta_{2N_2}^{*(2)} \\ \theta_{2N_2}^{*(3)} \end{bmatrix} = \begin{bmatrix} \frac{1}{N_2} \sum_{n_2=1}^{N_2} \theta_{1N_1n_2}^{*(1)} \\ \frac{A_{2N_2}^{(1)}}{B_{2N_2}} \\ \exp \left(\frac{A_{2N_2}^{(2)}}{B_{2N_2}} \right) \end{bmatrix}, \quad (5.241)$$

gdzie θ_{2N_2} jest optymalną wartością wektora parametrów modelu z klasy (5.239),

$$A_{2N_2}^{(1)} = \sum_{n_2=1}^{N_2} \ln \theta_{1n_1n_2}^{*(2)} \ln u_{2n_2} - \frac{1}{N_2} \left(\sum_{n_2=1}^{N_2} \ln \theta_{1n_1n_2}^{*(2)} \right) \left(\sum_{n_2=1}^{N_2} \ln u_{2n_2} \right), \quad (5.242)$$

$$A_{2N_2}^{(2)} = \frac{1}{N_2} \left(\sum_{n_2=1}^{N_2} (\ln u_{2n_2})^2 \right) \left(\sum_{n_2=1}^{N_2} \ln \theta_{1n_1n_2}^{*(2)} \right) - \frac{1}{N_2} \left(\sum_{n_2=1}^{N_2} \ln u_{2n_2} \right) \left(\sum_{n_2=1}^{N_2} \ln \theta_{1n_1n_2}^{*(2)} \ln u_{2n_2} \right), \quad (5.243)$$

$$B_{2N_2} = \sum_{n_2=1}^{N_2} (\ln u_{2n_2})^2 - \frac{1}{N_2} \left(\sum_{n_2=1}^{N_2} \ln u_{2n_2} \right)^2. \quad (5.244)$$

Jest to algorytm identyfikacji na 2. stopniu.

Dla danych z tabeli 5.3 otrzymujemy

$$\theta_{2N_2}^{*(1)} = -0,236, \quad \theta_{2N_2}^{*(2)} = 1,624, \quad \theta_{2N_2}^{*(3)} = 0,043.$$

Optymalny model z klasy (5.238) ma zatem postać

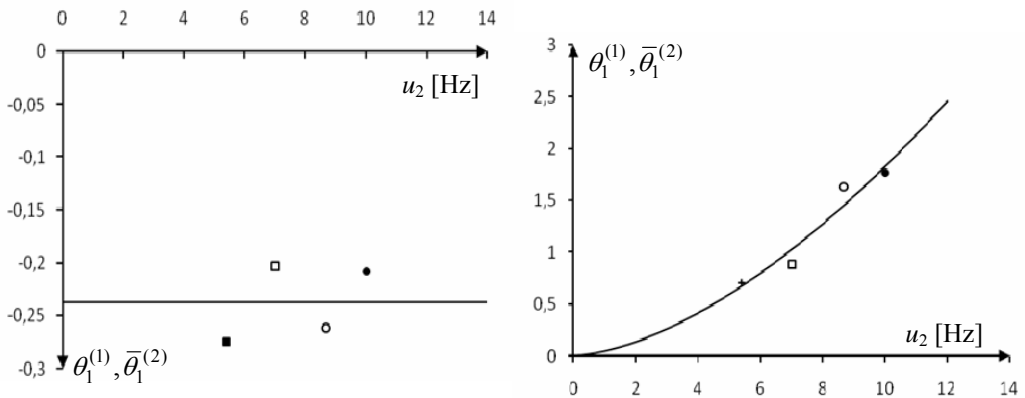
$$\bar{\theta}_1 = \Phi_2(u_2, \theta_{2N_2}^*) = \begin{bmatrix} \bar{\theta}_2^{(1)} \\ \bar{\theta}_2^{(2)} \end{bmatrix} = \begin{bmatrix} \theta_{2N_2}^{*(1)} \\ \theta_{2N_2}^{*(3)} u_2 \theta_{2N_2}^{*(2)} \end{bmatrix} = \begin{bmatrix} -0,236 \\ 0,043 u_2 \end{bmatrix}. \quad (5.245)$$

Wyniki identyfikacji na drugim stopniu zilustrowano na rysunku 5.10.

Ostatecznie, po podstawieniu (5.245) do (5.232), otrzymamy optymalny model przyrostu objętościowego współczynnika przenikania masy

$$\bar{y} = \Phi(u_1, u_2, \theta_{2N_2}^*) \equiv \Phi_1(u_1, \Phi_2(u_2, \theta_{2N_2}^*)) = 0,043 u_2^{1,624} u_1^{-0,236} \quad (5.246)$$

wyznaczony metodą dwustopniową.



Rys. 5.10. Optymalne parametry modelu na pierwszym stopniu dla różnych wartości częstotliwości pulsacji: (■) – $u_{21} = 5,33$ Hz, (□) – $u_{22} = 7,00$ Hz, (○) – $u_2 = 8,67$ Hz, (●) – $u_2 = 10,00$ Hz (tab. 5.3).
Linia ciągłą zaznaczono optymalny model na drugim stopniu – zależność (5.245)

Bezpośredni algorytm identyfikacji

Optymalne parametry modelu przyrostu objętościowego współczynnika przenikania masy spowodowanego pulsacją fazy parowej możemy wyznaczyć metodą bezpośrednią, proponując klasę modeli Φ , która jest złożeniem modeli Φ_1 (5.232) na 1. stopniu oraz Φ_2 (5.239) na 2. stopniu, czyli

$$\bar{y} = \Phi(u_1, u_2, \theta_2) = \theta_2^{(3)} u_2^{\theta_2^{(2)}} u_1^{\theta_2^{(1)}}, \quad (5.247)$$

gdzie $\theta_2 = [\theta_2^{(1)} \quad \theta_2^{(2)} \quad \theta_2^{(3)}]^T$ jest wektorem parametrów modelu.

Początkowo trudno było na podstawie analizy danych eksperymentalnych ustalić taki model. Dopiero wyniki identyfikacji na 1. stopniu oraz złożenie modeli (5.232) i (5.239) pozwoliły na określenie klasy modeli opisujących przyrost objętościowego współczynnika przenikania masy spowodowany pulsacją. W wyniku minimalizacji bezpośredniego kryterium

$$\begin{aligned} Q_{N_1 N_2}(\theta_2) &= \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} \left(\ln y_{n_1 n_2} - \ln \left(\theta_2^{(3)} u_{2n_2}^{\theta_2^{(2)}} u_{1n_1 n_2}^{\theta_2^{(1)}} \right) \right)^2 \\ &= \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} \left(\ln y_{n_1 n_2} - \ln \theta_2^{(3)} - \theta_2^{(2)} \ln u_{2n_2} - \theta_2^{(1)} \ln u_{1n_1 n_2} \right)^2 \end{aligned} \quad (5.248)$$

względem θ_2 otrzymujemy bezpośredni algorytm identyfikacji

$$\tilde{\theta}_{2N_1 N_2}^* = \begin{bmatrix} \tilde{\theta}_{2N_1 N_2}^{*(1)} \\ \tilde{\theta}_{2N_1 N_2}^{*(2)} \\ \tilde{\theta}_{2N_1 N_2}^{*(3)} \end{bmatrix} = \begin{bmatrix} A_{N_1 N_2}^{(1)} \\ A_{N_1 N_2}^{(2)} \\ \exp A_{N_1 N_2}^{(3)} \end{bmatrix}, \quad (5.249)$$

w którym: $\tilde{\theta}_{2N_1 N_2}^*$ jest optymalną wartością wektora parametrów modelu wyznaczoną metodą bezpośrednią, a

$$A_{N_1 N_2} \stackrel{\text{df}}{=} \begin{bmatrix} A_{N_1 N_2}^{(1)} \\ A_{N_1 N_2}^{(2)} \\ A_{N_1 N_2}^{(3)} \end{bmatrix} = M_{N_1 N_2}^{-1} b_{N_1 N_2}, \quad (5.250)$$

$$M_{N_1 N_2} = \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} \begin{bmatrix} \ln u_{1n_1 n_2} \\ \ln u_{2n_2} \\ 1 \end{bmatrix} \begin{bmatrix} \ln u_{1n_1 n_2} & \ln u_{2n_2} & 1 \end{bmatrix}, \quad (5.251)$$

$$b_{N_1 N_2} = \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} \begin{bmatrix} \ln u_{1n_1 n_2} \\ \ln u_{2n_2} \\ 1 \end{bmatrix} \ln y_{n_1 n_2}. \quad (5.252)$$

Korzystając z algorytmu (5.249), dla danych zawartych w tabeli 5.3 obliczamy $\tilde{\theta}_{2N_1 N_2}^{*(1)} = -0,237$, $\tilde{\theta}_{2N_1 N_2}^{*(2)} = 1,826$, $\tilde{\theta}_{2N_1 N_2}^{*(3)} = 0,029$. Optymalny model wyznaczony metodą bezpośrednią ma postać

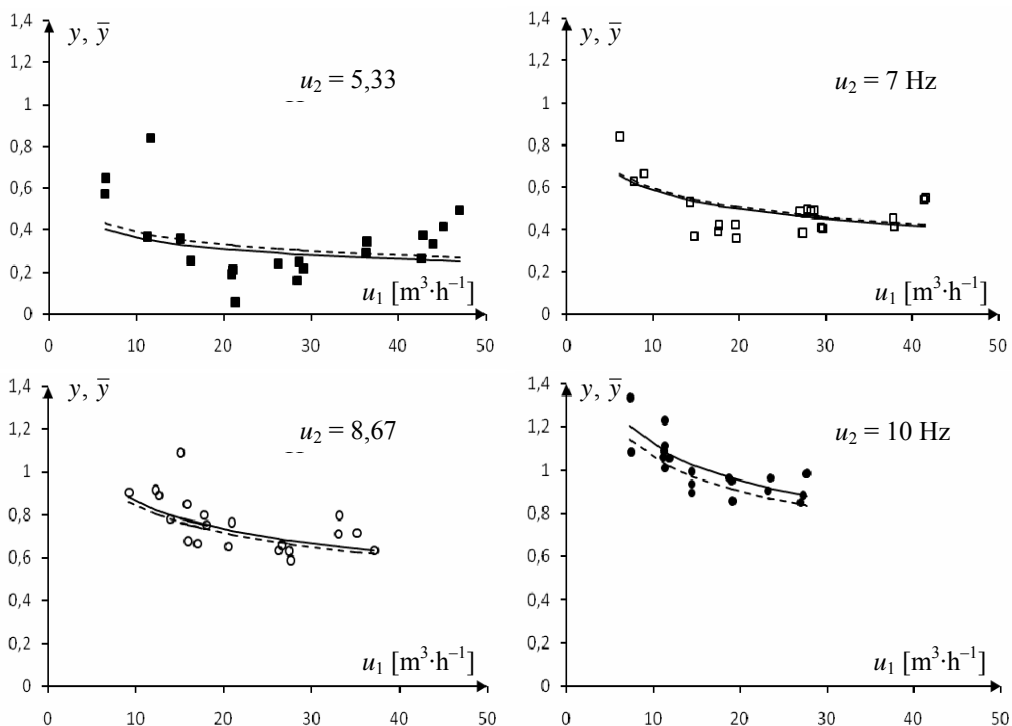
$$\bar{y} = \Phi(u_1, u_2, \tilde{\theta}_{2N_1 N_2}^*) = 0,029 u_2^{1,826} u_1^{-0,17}. \quad (5.253)$$

Wyniki uzyskane z wykorzystaniem podejścia bezpośredniego oraz identyfikacji dwustopniowej przedstawiono w tabeli 5.4 oraz na rysunku 5.11.

Tabela 5.4. Optymalne parametry modelu względnego przyrostu wartości objętościowego współczynnika przenikania masy uzyskane w podejściu dwustopniowym i bezpośrednim oraz ich ocena

Podejście	θ_2	$\theta_2^{(1)}$	$\theta_2^{(2)}$	$\theta_2^{(3)}$	$Q_{N_1, N_2}(\theta_2)$
Dwustopniowe	$\theta_2 = \theta_{2N_2}^*$	-0,236	1,624	0,043	1,053014
Bezpośrednie	$\theta_2 = \tilde{\theta}_{2N_1, N_2}^*$	-0,237	1,826	0,029	1,016943

Wyniki uzyskane w podejściu dwustopniowym oraz bezpośrednim nieznacznie się różnią. Spowodowane jest to organizacją eksperymentu. Ze względów technicznych nie jest możliwe wykonanie pomiarów dla tych samych wartości wejścia na 1. stopniu, tj. natężenia strumienia przepływu.



Rys. 5.11. Wartość względnego przyrostu objętościowego współczynnika przenikania masy dla zadanego natężenia strumienia przepływu, spowodowanego pulsacją o częstotliwości pulsacji: (■) – $u_2 = 5,33$ Hz, (□) – $u_2 = 7,00$ Hz, (○) – $u_2 = 8,67$ Hz, (●) – $u_2 = 10,00$ Hz. Linia ciągłą zaznaczono zależność (5.246) z optymalnymi parametrami wyznaczonymi metodą dwustopniową, natomiast linią przerywaną zależność (5.254) z optymalnymi parametrami wyznaczonymi z wykorzystaniem podejścia bezpośredniego

Porównując wartość kryterium dla modelu uzyskanego w podejściu dwustopniowym i bezpośrednim, widzimy, że podejście bezpośrednie daje lepszy wynik. Dwustopniowe zadanie identyfikacji pozwala jednak na usystematyzowanie badań i pokazuje procedurę tworzenia modelu z uwzględnieniem kolejnych nowych wyników badań, a podejście bezpośrednie jest ostatecznym podsumowaniem procesu identyfikacji.

5.3.3. Rozwinięcie badań z wykorzystaniem identyfikacji wielostopniowej

Innym parametrem pulsatora, który ma wpływ na przyrost wartości objętościowego współczynnika przenikania masy w wypełnionej kolumnie destylacyjnej z pulsacją fazy parowej jest amplituda pulsacji, którą oznaczymy przez u_3 . W celu zbadania jej wpływu na przyrost wartości objętościowego współczynnika przenikania masy można skorzystać z uogólnienia zadania identyfikacji dwustopniowej na zadanie identyfikacji wielostopniowej. Na 3. stopniu należy wówczas zbadać zależność między wektorem parametrów θ_2 modelu (5.239) a amplitudą pulsacji, tj. należy ustalić zależność

$$\bar{\theta}_2 = \Phi_3(u_3, \theta_3), \quad (5.254)$$

w której: $\bar{\theta}_2 = [\bar{\theta}_2^{(1)} \quad \bar{\theta}_2^{(2)} \quad \bar{\theta}_2^{(3)}]^T$ jest wyjściem modelu na 3. stopniu, a θ_3 – wektorem parametrów modelu.

W celu ustalenia zależności (5.254) powtórzono badania eksperymentalne opisane w p. 5.3.1 dla różnych wartości amplitudy pulsacji u_3 . Wymaga to powtórzenia zadania identyfikacji na 1. i 2. stopniu dla zadanej serii wartości amplitudy pulsacji – w konsekwencji wyznaczmy zbiór optymalnych wartości wektora parametrów modelu (5.247). Badania takie przedstawiono w pracy [76]. Tutaj ograniczymy się do przedstawienia algorytmu identyfikacji na 3. stopniu.

Wprowadzimy oznaczenie:

u_{3n_3} – n_3 -ta wartość amplitudy pulsacji,

$\theta_{2N_2n_3}^*$ – optymalna wartość wektora parametrów modelu (5.247) dla n_3 -tej wartości amplitudy pulsacji

$$\theta_{2N_2n_3}^* = [\theta_{2N_2n_3}^{*(1)} \quad \theta_{2N_2n_3}^{*(2)} \quad \theta_{2N_2n_3}^{*(3)}]^T, \quad n_3 = 1, 2, \dots, N_3,$$

N_3 – liczba zadanych wartości amplitudy pulsacji.

Na podstawie analizy wyników identyfikacji na 2. stopniu można zaproponować model (5.254) postaci

$$\bar{\theta}_2 = \Phi_3(u_3, \theta_3) = \begin{bmatrix} \theta_3^{(1)} \\ \theta_3^{(2)} \\ \theta_3^{(4)} u_3^{\theta_3^{(3)}} \end{bmatrix}, \quad (5.255)$$

gdzie $\theta_3 = [\theta_3^{(1)} \quad \theta_3^{(2)} \quad \theta_3^{(3)} \quad \theta_3^{(4)}]^T$ jest wektorem parametrów modelu na 3. stopniu.

Optymalny wektor parametrów modelu (5.255) wyznaczamy przez minimalizację kryterium

$$\begin{aligned} Q_{3N_3}(\theta_3) &= \sum_{n_3=1}^{N_3} \left[\left(\theta_{2N_2n_3}^{*(1)} - \theta_3^{(1)} \right)^2 + \left(\theta_{2N_2n_3}^{*(2)} - \theta_3^{(2)} \right)^2 + \left(\ln \theta_{2N_2n_3}^{*(3)} - \ln \left(\theta_3^{(4)} u_{3n_3} \theta_3^{(3)} \right) \right)^2 \right] \\ &= \sum_{n_3=1}^{N_3} \left[\left(\theta_{2N_2n_3}^{*(1)} - \theta_3^{(1)} \right)^2 + \left(\theta_{2N_2n_3}^{*(2)} - \theta_3^{(2)} \right)^2 + \left(\ln \theta_{2N_2n_3}^{*(3)} - \ln \theta_3^{(4)} - \theta_3^{(3)} \ln u_{3n_3} \right)^2 \right] \end{aligned} \quad (5.256)$$

względem θ_3 . W wyniku otrzymujemy algorytm identyfikacji na 3. stopniu

$$\theta_{3N_3}^* = \begin{bmatrix} \theta_{3N_3}^{*(1)} \\ \theta_{3N_3}^{*(2)} \\ \theta_{3N_3}^{*(3)} \\ \theta_{3N_3}^{*(4)} \end{bmatrix} = \begin{bmatrix} \frac{1}{N_3} \sum_{k=1}^{N_3} \theta_{2N_2n_3}^{*(1)} \\ \frac{1}{N_3} \sum_{k=1}^{N_3} \theta_{2N_2n_3}^{*(2)} \\ \frac{A_{3N_3}^{(1)}}{B_{3N_3}} \\ \exp \left(\frac{A_{3N_3}^{(2)}}{B_{3N_3}} \right) \end{bmatrix}, \quad (5.257)$$

gdzie $\theta_{3N_3}^*$ jest optymalną wartością wektora parametrów modelu (5.255),

$$A_{3N_3}^{(1)} = \sum_{n_3=1}^{N_3} \ln \theta_{2N_2n_3}^{*(3)} \ln u_{3n_3} - \frac{1}{N_3} \left(\sum_{n_3=1}^{N_3} \ln \theta_{2N_2n_3}^{*(3)} \right) \left(\sum_{n_3=1}^{N_3} \ln u_{3n_3} \right), \quad (5.258)$$

$$A_{3N_3}^{(2)} = \frac{1}{N_3} \sum_{n_3=1}^{N_3} (\ln u_{3n_3})^2 \sum_{n_3=1}^{N_3} \ln \theta_{2N_2n_3}^{*(3)} - \frac{1}{N_3} \left(\sum_{n_3=1}^{N_3} \ln u_{3k} \right) \left(\sum_{n_3=1}^{N_3} \ln \theta_{2N_2n_3}^{*(3)} \ln u_{3n_3} \right), \quad (5.259)$$

$$B_{3N_3} = \sum_{n_3=1}^{N_3} (\ln u_{3n_3})^2 - \frac{1}{N_3} \left(\sum_{n_3=1}^{N_3} \ln u_{3n_3} \right)^2. \quad (5.260)$$

Model przyrostu wartości objętościowego współczynnika przenikania masy, uwzględniający wpływ amplitudy pulsacji, uzyskamy w wyniku złożenia modeli wyznaczonych na 1., 2., i 3. stopniu, czyli

$$\bar{y} = \Phi(u_1, u_2, u_3, \theta_{3N_3}^*) = \Phi_1(u_1, \Phi_2(u_2, \Phi_3(u_3, \theta_{3N_3}^*))). \quad (5.261)$$

W pracy [76] przedstawiono badania destylacyjnych kolumn wypełnionych z pulsacją fazy parowej z wypełnieniami pierścieniami: Raschiga, PALL, Białeckiego oraz I-13-II. W tym przypadku wielostopniowe zadanie identyfikacji pozwala na rozbudowę modelu wypełnionej kolumny destylacyjnej z pulsacją fazy parowej dzięki kolejnym wynikom badań, a zwłaszcza na określenie wpływu wielkości charakteryzujących wypełnienie u_4 na sprawność kolumny. Można to uzyskać przez zastosowanie zadania identyfikacji wielostopniowej, a konkretnie na 4. stopniu należy zbadać wpływ wielkości charakteryzujących wypełnienie u_4 na parametry modelu (5.261).

Do przedstawionych zadań identyfikacji na 2., 3. i 4. stopniu można, oczywiście, stosować podejście bezpośrednie. Wielostopniowe zadanie identyfikacji pozwala jednak na usystematyzowanie badań i pokazuje procedurę tworzenia modelu z uwzględnieniem kolejnych, nowych wyników badań.

Dodatek. Przekształcenia zmiennych losowych

W rozdziałach dotyczących identyfikacji obiektów statycznych w warunkach losowych wielokrotnie, w różnych punktach, korzystaliśmy z przekształcenia wielowymiarowych zmiennych losowych. Dla uproszczenia zapisu zakładaliśmy tam, że przekształcenia są wzajemnie jednoznaczne i istnieją przekształcenia odwrotne. Założenia te można osłabić [48, 79]. Wymaga to jednak dodatkowych przekształceń. Ponieważ w różnych punktach przekształcenia dotyczyły różnych zmiennych losowych oraz różnych przekształceń, wprowadzone tutaj oznaczenia są niezależne od oznaczeń używanych we wspomnianych rozdziałach książki.

Niech \underline{x} będzie S -wymiarową ciągłą zmienną losową określoną na zbiorze \mathcal{X} , który jest podzbiorem S -wymiarowej przestrzeni liczb rzeczywistych. Wartość zmiennej losowej $x \in \mathcal{X} \subseteq \mathbb{R}^S$. Znana jest funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{x} , oznaczona przez $f_x(x)$. Ciągła L -wymiarowa zmienna losowa \underline{y} , określona na zbiorze \mathcal{Y} , który jest podzbiorem L -wymiarowej przestrzeni liczb rzeczywistych, jest wynikiem przekształcenia zmiennej losowej \underline{x} według zależności

$$\underline{y} = h(\underline{x}), \quad (\text{D.1})$$

gdzie h jest znaną funkcją taką, że $h: \mathcal{X} \rightarrow \mathcal{Y}$.

Wartość zmiennej losowej $y \in \mathcal{Y} \subseteq \mathbb{R}^L$.

Pojawia się pytanie: Jak wyznaczyć funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , czyli funkcję $f_y(y)$, na podstawie znajomości przekształcenia h oraz funkcji $f_x(x)$?

Sposób postępowania zależy od własności funkcji h w przekształceniu (D.1).

D.1. Załóżymy, że zmienne losowe \underline{x} i \underline{y} mają ten sam wymiar, czyli $S = L$, a ponadto że funkcja h jest wzajemnie jednoznaczna i istnieje funkcja odwrotna względem \underline{x} , tj.

$$\underline{x} = h^{-1}(\underline{y}), \quad (\text{D.2})$$

gdzie h^{-1} oznacza funkcję odwrotną.

Przy podanych założeniach funkcję gęstości rozkładu prawdopodobieństwa $f_y(y)$ wyznaczamy ze wzoru

$$f_y(y) = f_x(h^{-1}(y)) |J_h|, \quad (\text{D.3})$$

w którym J_h jest jacobianem przekształcenia odwrotnego (D.2), czyli

$$J_h = \frac{\partial h^{-1}(y)}{\partial y}. \quad (\text{D.4})$$

Przykład D.1. Niech S -wymiarowa zmienna losowa \underline{x} ma rozkład normalny o wartości oczekiwanej m_x i macierzy kowariancji Σ_x , czyli

$$f_x(x) = (2\pi)^S |\Sigma_x^{-1}|^{\frac{1}{2}} \exp\left[-\frac{1}{2}(x - m_x)^T \Sigma_x^{-1}(x - m_x)\right]. \quad (\text{D.5})$$

Zmienna losowa \underline{y} jest wynikiem przekształcenia zmiennej losowej \underline{x} (D.1) zgodnie z zależnością

$$\underline{y} = h(\underline{x}) = A\underline{x} + b, \quad (\text{D.6})$$

w której: A jest macierzą nieosobliwą o wymiarze $(S \times S)$, a b jest S -wymiarowym wektorem.

W wyniku przekształcenia otrzymujemy $L = S$ -wymiarową zmienną losową \underline{y} . Ponieważ macierz A jest nieosobliwa, przekształcenie (D.6) jest wzajemnie jednoznaczne, w konsekwencji przekształcenie odwrotne (D.2) ma postać

$$\underline{x} = h^{-1}(\underline{y}) = A^{-1}(\underline{y} - b). \quad (\text{D.7})$$

Jacobian (D.4) przekształcenia odwrotnego (D.7) ma postać

$$J_h = \frac{\partial h^{-1}(y)}{\partial y} = \frac{\partial(A^{-1}(y - b))}{\partial y} = A^{-1}. \quad (\text{D.8})$$

Zgodnie z (D.3) funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} wyznaczamy według wzoru

$$f_y(y) = (2\pi)^S |\Sigma_x^{-1}|^{\frac{1}{2}} \exp\left[-\frac{1}{2}(A^{-1}(y - b) - m_x)^T \Sigma_x^{-1}(A^{-1}(y - b) - m_x)\right] |A^{-1}|. \quad (\text{D.9})$$

Po prostych przekształceniach algebraicznych otrzymujemy

$$f_y(y) = (2\pi)^S |\Sigma_x^{-1}|^{\frac{1}{2}} |A^{-1}| \exp\left[-\frac{1}{2}(y - (Am_x + b))^T A^{-1} \Sigma_x^{-1} A^{-1}(y - (Am_x + b))\right]. \quad (\text{D.10})$$

Po przyjęciu oznaczeń: $m_y = Am_x + b$ i $\Sigma_y = A\Sigma_x A^T$ funkcja gęstości rozkładu prawdopodobieństw zmiennej losowej (D.7) przyjmuje postać

$$f_y(y) = (2\pi)^S |\Sigma_y^{-1}|^{\frac{1}{2}} \exp\left[-\frac{1}{2}(y - m_y)^T \Sigma_y^{-1}(y - m_y)\right]. \quad (\text{D.11})$$

W przekształceniach skorzystaliśmy z następujących własności rachunku macierzowego:

$$(A\Sigma_x A^T)^{-1} = A^{-T} \Sigma_x^{-1} A^{-1}, \quad |A| = |A^T| \quad \text{oraz} \quad |A\Sigma_x A^T|^{-1} = |A^{-T}| |\Sigma_x^{-1}| |A^{-1}| = |\Sigma_x^{-1}| |A^{-1}|^2.$$



D.2. Załóżmy, że zmienna losowa \underline{y} po przekształceniu (D.1) ma wymiar mniejszy od wymiaru zmiennej losowej \underline{x} , czyli $L < S$. W tym przypadku funkcja h nie jest wzajemnie jednoznaczna względem \underline{x} i nie możemy skorzystać z rozważań przedstawionych w punkcie **D.1**.

Zdefiniujemy pomocniczą $(S - L)$ -wymiarową ciągłą zmienną losową \underline{z} określoną na zbiorze $\bar{\mathcal{Z}}$, który jest podzbiorem $(S - L)$ -wymiarowej przestrzeni liczb rzeczywistych. Wartość zmiennej losowej $z \in \mathcal{Z} \subseteq \mathbb{R}^{S-L}$. Załóżmy dodatkowo, że istnieje przekształcenie h' zmiennej \underline{x} na zmienną losową \underline{z} , czyli

$$\underline{z} = h'(\underline{x}), \quad (\text{D.12})$$

które wraz z funkcją h tworzy przekształcenie wzajemnie jednoznaczne, czyli

$$\begin{bmatrix} \underline{y} \\ \underline{z} \end{bmatrix} = \begin{bmatrix} h(\underline{x}) \\ h'(\underline{x}) \end{bmatrix} \stackrel{\text{def}}{=} \bar{h}(\underline{x}), \quad (\text{D.13})$$

gdzie h' jest znaną funkcją taką, że $h': X \rightarrow \mathcal{Z}$.

Teraz funkcja \bar{h} jest wzajemnie jednoznaczna względem \underline{x} , a zatem istnieje funkcja odwrotna

$$\underline{x} = \bar{h}^{-1}(\underline{y}, \underline{z}), \quad (\text{D.14})$$

w której \bar{h}^{-1} oznacza funkcję odwrotną.

Teraz funkcję gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{y}, \underline{z})$, czyli $f_{y,z}(y, z)$, wyznaczamy ze wzoru

$$f_{y,z}(y, z) = f_x(\bar{h}^{-1}(y, z)) |J_{\bar{h}}|, \quad (\text{D.15})$$

w którym $J_{\bar{h}}$ jest jacobianem przekształcenia odwrotnego (D.14).

Interesującą nas funkcję gęstości prawdopodobieństwa zmiennej losowej \underline{y} wyznaczamy jako rozkład brzegowy (D.15), czyli

$$f_y(y) = \int_{\underline{z}} f_{y,z}(y,z) dz = \int_{\underline{z}} f_x(\bar{h}^{-1}(y,z)) |J_{\bar{h}}| dz. \quad (\text{D.16})$$

Przykład D.2. Rozważmy dwuwymiarową ($S = 2$) ciągłą zmienną losową \underline{x} , o składowych $\underline{x}^{(1)}$ i $\underline{x}^{(2)}$, określoną na dwuwymiarowej przestrzeni liczb rzeczywistych, czyli wartość zmiennej losowej $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} \in \mathbb{R}^2$. Niech $f_x(x^{(1)}, x^{(2)})$ oznacza funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{x} . Wyznamy funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , będącą iloczynem zmiennych $\underline{x}^{(1)}$ oraz $\underline{x}^{(2)}$. Przekształcenie (D.1) ma postać

$$\underline{y} = h(\underline{x}) = \underline{x}^{(1)} \underline{x}^{(2)}, \quad (\text{D.17})$$

gdzie h jest znaną funkcją taką, że $h: \mathbb{R}^2 \rightarrow \mathbb{R}^1$. Wartość zmiennej losowej $y \in \mathbb{R}^1$.

W rozpatrywanym przykładzie wymiar zmiennej losowej $\underline{y} - L = 1$, a zatem $L < S$. Przekształcenie (D.17) nie jest wzajemnie jednoznaczne. Zdefiniujemy pomocniczą $(S - L) = (2 - 1) = 1$ -wymiarową ciągłą zmienną losową \underline{z} , określoną na zbiorze liczb rzeczywistych. Wartość zmiennej losowej $z \in \mathbb{R}^1$. Przyjmujemy dodatkowe przekształcenie

$$\underline{z} = h'(\underline{x}) = \underline{x}^{(1)}, \quad (\text{D.18})$$

takie że $h': \mathbb{R}^2 \rightarrow \mathbb{R}^1$.

Teraz przekształcenia (D.17) oraz (D.18) tworzą

$$\begin{bmatrix} \underline{y} \\ \underline{z} \end{bmatrix} = \begin{bmatrix} h(\underline{x}) \\ h'(\underline{x}) \end{bmatrix} = \begin{bmatrix} \underline{x}^{(1)} \underline{x}^{(2)} \\ \underline{x}^{(1)} \end{bmatrix} \stackrel{\text{df}}{=} \bar{h}(\underline{x}), \quad (\text{D.19})$$

przekształcenie \bar{h} , takie że $\bar{h}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, które jest wzajemnie jednoznaczne względem dwuwymiarowej zmiennej losowej \underline{x} , a przekształcenie odwrotne (D.14) ma postać

$$\underline{x} = \begin{bmatrix} \underline{x}^{(1)} \\ \underline{x}^{(2)} \end{bmatrix} = \begin{bmatrix} \underline{z} \\ \frac{\underline{y}}{\underline{z}} \end{bmatrix} = \bar{h}^{-1}(\underline{y}, \underline{z}). \quad (\text{D.20})$$

Wyznacznik jacobianu przekształcenia odwrotnego (D.20) – $|J_{\bar{h}}|$ – ma postać

$$|J_{\bar{h}}| = \begin{vmatrix} 0 & 1 \\ \frac{1}{z} & -\frac{y}{z^2} \end{vmatrix} = \frac{1}{|z|} = \frac{1}{|z|}. \quad (\text{D.21})$$

W punkcie $z = 0$ jacobian nie jest określony.

Funkcję gęstości rozkładu prawdopodobieństwa pary zmiennych losowych $(\underline{y}, \underline{z})$, czyli $f_{y,z}(y, z)$ – (D.15), wyznaczamy ze wzoru

$$f_{y,z}(y, z) = f_x\left(z, \frac{y}{z}\right) \frac{1}{|z|}, \quad (\text{D.22})$$

przy założeniu $z \neq 0$, a poszukiwana funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , która jest iloczynem zmiennych $\underline{x}^{(1)}$ oraz $\underline{x}^{(2)}$, zgodnie z (D.16) ma postać

$$f_y(y) = \int_{-\infty}^{\infty} f_{y,z}(y, z) dz = \int_{-\infty}^{\infty} f_x\left(z, \frac{y}{z}\right) \frac{1}{|z|} dz. \quad (\text{D.23})$$



D.3. Załóżmy, że zmienne losowe \underline{x} i \underline{y} mają ten sam wymiar, czyli $S = L$, natomiast funkcja h w przekształceniu (D.1) nie jest wzajemnie jednoznaczna w obszarze określoności $X \subseteq \mathbb{R}^S$. W tym przypadku dokonamy podziału zbioru argumentów

X na skończoną liczbę K rozłącznych podzbiorów X_k , $k = 1, 2, \dots, K$, $X = \bigcup_{k=1}^K X_k$,

$X_k \cap X_n = \emptyset$ dla $k \neq n$, $k, n = 1, 2, \dots, K$ oraz – odpowiednio – zbiór wartości Y na (niekoniecznie rozłączne) podzbiory Y_k , $k = 1, 2, \dots, K$, $Y = \bigcup_{k=1}^K Y_k$ w taki sposób że

odwzorowanie X_k w Y_k (D.1) jest wzajemnie jednoznaczne. Oznaczmy przez h_k wzajemnie jednoznaczne przekształcenie określone na X_k , takie że $h_k : X_k \rightarrow Y_k$, czyli

$$\underline{y} = h_k(\underline{x}), \quad k = 1, 2, \dots, K, \quad (\text{D.24})$$

oraz przekształcenie odwrotne h_k^{-1} , takie że $h_k^{-1} : Y_k \rightarrow X_k$, czyli

$$\underline{x} = h_k^{-1}(\underline{y}), \quad k = 1, 2, \dots, K, \quad (\text{D.25})$$

wtedy funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} ma postać

$$f_y(y) = \sum_{k=1}^K \pi_k f_x(h_k^{-1}(y)) |J_{h_k}|, \quad (\text{D.26})$$

gdzie J_{h_k} , $k=1, 2, \dots, K$ są jacobianami przekształceń odwrotnych (D.25), natomiast

$$\pi_k = \begin{cases} 1 & \text{gdy } y \in Y_k \\ 0 & \text{gdy } y \notin Y_k \end{cases}, \quad k=1, 2, \dots, K. \quad (\text{D.27})$$

Przykład D.3. Rozważmy jednowymiarową ($S=1$) ciągłą zmienną losową \underline{x} , określoną na przestrzeni liczb rzeczywistych, czyli wartość zmiennej losowej $x \in X \equiv \mathbb{R}^1$. Niech $f_x(x)$ oznacza funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{x} . Wyznamy funkcję gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , będącą kwadratem zmiennej losowej \underline{x} . Przekształcenie (D.1) ma postać

$$\underline{y} = h(\underline{x}) = \underline{x}^2. \quad (\text{D.28})$$

Przekształcenie (D.28) nie jest wzajemnie jednoznaczne w przestrzeni $X \equiv \mathbb{R}^1$. Podzielimy zbiór $X \equiv \mathbb{R}^1$ na dwa rozłączne podzbiory $X_1 = \{x \in \mathbb{R}^1 : x < 0\}$ oraz $X_2 = \{x \in \mathbb{R}^1 : x > 0\}$. W przestrzeni Y zbiorom tym odpowiadają dwa zbiory $Y_1 = \{y \in \mathbb{R}^1 : y > 0\}$ oraz $Y_2 = \{y \in \mathbb{R}^1 : y > 0\}$. Parze X_1, Y_1 odpowiada przekształcenie $\underline{y} = h_1(\underline{x}) = \underline{x}^2$ (D.24), korespondujące z nim przekształcenie odwrotne (D.25) ma postać $\underline{x} = h_1^{-1}(\underline{y}) = -\sqrt{\underline{y}}$, a odpowiadający mu jacobian przekształcenia odwrotnego $J_{h_1} = -\frac{1}{2\sqrt{y}}$. Parze X_2, Y_2 odpowiada natomiast przekształcenie $\underline{y} = h_2(\underline{x}) = \underline{x}^2$ (D.24), korespondujące z nim przekształcenie odwrotne (D.25) ma postać $\underline{x} = h_2^{-1}(\underline{y}) = \sqrt{\underline{y}}$, a odpowiadający mu jacobian przekształcenia odwrotnego $J_{h_2} = \frac{1}{2\sqrt{y}}$. W punkcie $y=0$ jacobian przekształcenia odwrotnego nie jest określony. Zgodnie z (D.26) funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej \underline{y} , będącą kwadratem zmiennej losowej \underline{x} , dana jest wzorem

$$f_y(y) = f_x(h_1^{-1}(x))|J_{h_1}| + f_x(h_2^{-1}(x))|J_{h_2}| = (f_x(-\sqrt{y}) + f_x(\sqrt{y}))\frac{1}{2\sqrt{y}}, \quad y \neq 0. \quad (\text{D.29})$$



Bibliografia

- [1] ABRAMOWITZ M., STEGUN I. (eds.), *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, U.S. Government Printing Office, Washington 1964.
- [2] ARAFEH S.A., SAGE A.P., *Multilevel discrete time systems identification for large scale systems*, Int. J. on Systems Science, Vol. 5, 1974, 753–783.
- [3] ARAFEH S.A., SAGE A.P., *Hierarchical system identification of states and parameters in interconnected power systems*, Int. J. on Systems Science, Vol. 5, 1974, 817–864.
- [4] ABDEL GHANY A.M., ŚWIĄTEK J., *Two-stage dynamic identification of direct current motor*, Systems Science, Vol. 20, No. 3, 1994, 65–77.
- [5] ABDEL GHANY A.M., ŚWIĄTEK J., *Identification of the two-stage dynamic system using instrumental variable method*, Proc. of the 13th International Conference on Systems Science, Wrocław, Poland, 15–18 September 1998, Vol. 1, Bubnicki Z., Grzech A. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1998, 243–252.
- [6] BEKEY G.A., BENEKEN J.E., *Identification of biological systems: a survey*, Automatica, Vol. 14, 1978, 41–47.
- [7] BORZEMSKI L., GRZECH A., KASPRZAK A., KOSZAŁKA L., *Development of distributed applications for experiment and education support based on the NETEX local area network*, Proc. of International Symposium COMNET '85 Services Conveyed by Computer Networks, John von Neumann Soc. Comput. Sc., Budapest 1985, 4.38–4.47.
- [8] BOX G.E.P., HUNTER J.S., HUNTER W.G., *Statistics for experiments: design, Innovation and discovery* (second edition), Wiley, New Jersey 2005.
- [9] BRZOSTOWSKI K., ŚWIĄTEK J., *Adaptive control algorithm with two-stage identification*, Proc. 18th International Conference on Systems Engineering, Coventry, UK, 5–7 September 2006, Burnham K.J., Haas O.C.L. (eds.), Coventry University, Coventry 2006, 55–58.
- [10] BRZOSTOWSKI K., ŚWIĄTEK J., *On convergence an adaptive control algorithm based on pattern recognition*, Proc. 16th International Conference on Systems Science, Wrocław, Poland, 4–6 September 2007, Vol. 1, Grzech A. (ed.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2007, 341–350.
- [11] BRZOSTOWSKI K., ŚWIĄTEK J., *A pattern recognition algorithm used in knowledge-based adaptive control system to select strategy. Lecture Notes in Computer Science*, Lecture Notes in Artificial Intelligence, Vol. 4693, 2007, 247–254.
- [12] BRZOSTOWSKI K., ŚWIĄTEK J., *Adaptacyjny algorytm wyboru scenariusza oparty na wiedzy eksperta. Sterowanie i automatyzacja: aktualne problemy i ich rozwiązania*, Malinowski K., Rutkowski L. (red.), Akademicka Oficyna Wydawnicza EXIT, Warszawa 2008, 275–284.
- [13] BRZOSTOWSKI K., DRAPAŁA J., ŚWIĄTEK J., *How to replace inexact expert's knowledge by precise diagnostic system – assessment of internal state of human elbow neuromuscular system*, Knowledge processing and reasoning for information society, Ngoc Thanh Nguyen, Kołaczek G., Gabryś B. (eds.), Akademicka Oficyna Wydawnicza EXIT, Warszawa 2008, 293–303.

- [14] BUBNICKI Z., *Identification of Control Plants*, PWN–Elsevier, Warszawa–Amsterdam–Oxford–New York 1980.
- [15] BUBNICKI Z., *Global and local identification of complex systems with cascade structure*, Systems Science, Vol. 1, No. 1, 1975, 55–65.
- [16] BUBNICKI Z., *On the multistage identification*, Systems Science, Vol. 3, No. 2, 1977, 207–210.
- [17] BUBNICKI Z., *Optimization problems in large-scale systems modeling and identification*, in: A. Straszak (ed.), *Large Scale Systems: Theory and Applications 1983*, Pergamon Press, Oxford 1984, 411–416.
- [18] BUBNICKI Z., *Optimal models of complex operation systems*, 6^e Congress International de Cybernetique et de Systemique, College de Systemique de l’AFCET, Paris 1984, 871–876.
- [19] BUBNICKI Z., *Global modeling and identification of complex systems*, Proc. of 7th IFAC/IFORS Symp. York, UK, 1985, Identification and System Parameter Estimation, Pergamon Press, Oxford–New York–Toronto–Sydney–Frankfurt 1985, 261–263.
- [20] Bubnicki Z., *Wstęp do systemów ekspertowych*, Państwowe Wydawnictwo Naukowe, Warszawa 1990.
- [21] BUBNICKI Z., *Analysis and Decision Making in Uncertain Systems*, Springer, London 2004.
- [22] BUBNICKI Z., *Teoria i algorytmy sterowania*, WNT, Warszawa 2005.
- [23] BUBNICKI Z., KURZYŃSKI M., PUCHAŁA E., ŚWIĄTEK J., WILIMOWSKI M., ŻOŁNIEREK A., *Pakiet procedur statystycznych systemu BAMED*, cz. I – dokumentacja projektowa, cz. II – dokumentacja eksploatacyjna, Raporty ICT serii SPR 3,4/81, Politechnika Wrocławska, Wrocław 1982.
- [24] BUBNICKI Z., ŚWIĄTEK J., *On the Bayesian estimation of a complex dynamic system*, Systems Science, Vol. 6, No. 4, 1980, 305–315.
- [25] BUBNICKI Z., ŚWIĄTEK J., *Separability and estimation problem in the identification of complex system*, Proc. of 5th Polish–Italian Symposium: Application of Systems Theory to Economics, Management and Technology, Toruń 1980, PWN, Warszawa–Łódź 1980, 128–137.
- [26] BUBNICKI Z., ŚWIĄTEK J., *On the parameter estimation in the identification of complex static system*, Bulletin de l’Academie Polonaise des Sciences, Serie des sciences techniques, Vol. 29, No. 1–2, 1981, 35–44.
- [27] BUBNICKI Z., ŚWIĄTEK J., *Separowalność i estymacja parametrów w identyfikacji złożonych systemów statycznych*, Archiwum Automatyki i Telemekhaniki, t. 26, z. 3, 1981, 349–363.
- [28] BUBNICKI Z., ŚWIĄTEK J., *Optimization problems in the parameter estimation for complex input-output systems*, Proc. of X Internationaler Kongress über Anwendungen der Mathematik in den Ingenieurwissenschaften, Berichte 4, Weimar 1984, 29–32.
- [29] CHERKASSKY V., MULIER F., *Learning from data: concepts, theory, and methods* (second edition), Wiley, New Jersey 2007.
- [30] CRAMER H., *Mathematical Methods of Statistics*, Princeton, New York, 1946.
- [31] DRAŁUS G., ŚWIĄTEK J., *A modified back propagation algorithm for modeling static complex systems using neural network*, Proc. 5th International Conference Neural networks and soft computing, Rutkowski L., Tadeusiewicz R. (red.), Polish Neural Network Society, Częstochowa 2000, 463–468.
- [32] DRAŁUS G., ŚWIĄTEK J., *Static neural networks in global modeling of complex system*, Proc. 14th International Conference on Systems Engineering, September 12–14, 2000, Coventry University, Vol. 2, Coventry 2000, 547–551.
- [33] DRAŁUS G., ŚWIĄTEK J., *Globally optimal models of complex systems with regard to quality of local models*, Proc. 14th International Conference on Systems Science, 11–14 September 2001, Wrocław, Vol. 1, Bubnicki Z., Grzech A. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2001, 217–226.
- [34] DRAŁUS G., ŚWIĄTEK J., *Global modelling of complex systems by neural network*, Proc. 7th International Symposium on Artificial Life and Robotics (AROB ’02) for Information Technol-

- ogy, Beppu, Oita – Beppu, Japan, January 16–18, 2002 (AROB '02), Vol. 1, Masanori Sugisaka M., Tanaka H. (eds.), Oita 2004, 618–621.
- [35] DRAŁUS G., ŚWIĄTEK J., *Sieci neuronowe w modelowaniu złożonego obiektu chemicznego*, XIV Krajowa Konferencja Automatyki, Zielona Góra, 24–27 czerwca 2002, (KKA 2002), t. 2, Bubnicki Z., Korbacz J. (red.), Wydawnictwo Naukowo-Techniczne Uniwersytetu Zielonogórskiego, Zielona Góra 2002, 825–830.
- [36] DRAŁUS G., ŚWIĄTEK J., *Global network modeling of complex systems with respect to local models quality*, Proc. 15th International Conference on Systems Engineering, Las Vegas, Nevada, US, August 6–8, 2002, ICSE '02, Selvaraj H., Muthukumar V. (eds.), Las Vegas 2002, 218–226.
- [37] DRAŁUS G., ŚWIĄTEK J., *A penalty function to obtain satisfactory local models of complex systems*, Proc. 6th International Conference Neural networks and soft computing, Zakopane, June 11–15, 2002, Rutkowski L., Kacprzyk J. (eds.), Physica-Verlag, Heidelberg–New York 2003, 167–172.
- [38] DRAPAŁA J., ŚWIĄTEK J., *Identyfikacja dynamicznych systemów złożonych z wykorzystaniem sieci neuronowych. Inżynieria wiedzy i systemy ekspertowe*, t. 1, Grzech A. (red.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2006, 311–320.
- [39] DRAPAŁA J., ŚWIĄTEK J., *Modeling of dynamic complex systems by neural networks*, Proc. 18th International Conference on Systems Engineering, Coventry, UK, 5–7 September 2006, Burnham K.J., Haas O.C.L. (eds.), Coventry University, Coventry 2006, 109–112.
- [40] DRAPAŁA J., ŚWIĄTEK J., *Algorithm of recurrent multilayer perceptions learning for global modeling of complex systems*, Proc. 16th International Conference on Systems Science, Wrocław, Poland, 4–6 September 2007, Vol. 1, Grzech A. (ed.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2007, 351–358.
- [41] DRAPAŁA J., ŚWIĄTEK J., *Global and local approach to complex systems modeling using dynamic neural networks – analogy with multi-agent system*, Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence, Vol. 4693, 2007, 279–286.
- [42] DRAPAŁA J., ŚWIĄTEK J., *Dynamiczne sieci neuronowe jako globalnie optymalny model systemu złożonego – zbieżność algorytmów uczenia*, w: Sterowanie i automatyzacja: aktualne problemy i ich rozwiązania, Malinowski K., Rutkowski L. (red.), Akademicka Oficyna Wydawnicza EXIT, Warszawa 2008, 155–164.
- [43] EYKHOFF P., *System Identification, Parameter and State Estimation*, Wiley, New York–London–Sydney 1974.
- [44] EYKHOFF P., *Trends and Progress in System Identification*, Pergamon Press, Oxford–New York–Toronto–Sydney–Paris–Frankfurt 1981.
- [45] EYKHOFF P., *Biomedical identification: overview, problems and prospects*, Proc. of 7th IFAC/IFORS Symp., York, UK, 1985, Identification and System Parameter Estimation, Pergamon Press, Oxford–New York–Toronto–Sydney–Frankfurt 1985, 37–44.
- [46] FINDEISEN W., *Struktury sterowania dla złożonych systemów*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 1997.
- [47] FINDEISEN W., SZYMANOWSKI J., WIERZBICKI A., *Teoria i metody obliczeniowe optymalizacji*, PWN, Warszawa 1977.
- [48] FISZ M., *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, Warszawa 1969.
- [49] FRESEWINKEL T., UNBEHAUEN H., *On the identification of subsystems of decomposed large scale systems*, Proc. of 7th IFAC/IFORS Symp., York, UK, 1985, Identification and System Parameter Estimation, Pergamon Press, Oxford–New York–Toronto–Sydney–Frankfurt 1985, 267–272.
- [50] GIERACHA M., ŚWIĄTEK J., *Estimation error as a measure for on-line two-stage experiment design*, Proc. 14th International Conference on Systems Science, 11–14 September 2001, Wrocław, Poland, Vol. 1, Bubnicki Z., Grzech A. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2001, 227–233.

- [51] GREBLICKI W., *Asymptotycznie optymalne algorytmy rozpoznawania i identyfikacji w warunkach probabilistycznych*, Prace Naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej, seria: Monografie nr 3, Wrocław 1974.
- [52] GREBLICKI W., *Recursive identification of continuous-time Hammerstein systems*, International Journal on Systems Science, Vol. 33, No. 12, 2002, 969–977.
- [53] HASIEWICZ Z., STANKIEWICZ A., *On applicability of interconnection balance method to global identification of interconnected steady-state systems*, Tran Automatic Control, Vol. AC-31, No. 1, 1986, 77–80.
- [54] HASIEWICZ Z., *Identyfikacja sterowanych systemów o złożonej strukturze*, Prace Naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej, seria: Monografie nr 22, Wrocław 1993.
- [55] HASIEWICZ Z., *Non-parametric estimation of non-linear in a cascade time-series system by multiscale approximation*, Signal Processing, Vol. 81, No. 4, 2001, 791–807.
- [56] KACPRZYŃSKI B., *Planowanie eksperymentów, podstawy matematyczne*, WNT, Warszawa 1974.
- [57] KOZIÓŁ A., PIENIAŻEK W., *Zatrzymanie cieczy w kolumnach wypełnionych z pulsacją fazy parowej*, Inżynieria Chemiczna i Procesowa, t. 23, z. 4, 2002, 507–514.
- [58] KUCHARSKI S., GŁOGOWSKI J., *Podstawy obliczeń projektowych w technologii chemicznej*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2005.
- [59] KORBICZ J., WITCZAK M., *An evolutionary approach to identification of non-linear dynamic systems*, Proc. Int. Conf. Artificial Neural Nets and Genetic Algorithms, ICANNGA, Prague, Springer, Wien–New York 2001, 240–246.
- [60] KORBICZ J., MRUGALSKI M., ARINTON E., *Systems identification with GMDH neural networks; a multidimensional case*, Proc. Int. Conf. Artificial Neural Nets and Genetic Algorithms, ICANNGA, Roanne, Springer, Wien–New York 2003, 115–120.
- [61] KORBICZ J., JANCZAK A., *Identification of polynomial MIMO Wiener systems using instrumental variables method*, Proc. 12th IEEE Int. Conf. Methods and Models in Automation and Robotics, MMAR, Vol. 1, Międzyzdroje 2006, 926–934.
- [62] KUCZYŃSKI L., *Technologia leków*, WNT, Warszawa 1971.
- [63] KULCZYCKI P., HRYNIEWICZ O., KACPRZYK J., *Techniki informacyjne w badaniach systemowych*, WNT, Warszawa 2007.
- [64] KURZYŃSKI M., *Wielostopniowa identyfikacja obiektów statycznych w warunkach probabilistycznych*, Prace naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej, nr 46, Wrocław 1975, 47–58.
- [65] LJUNG L., *System identification toolbox for use with MATLAB. User's guide*, The MATH WORKS Inc. 1995.
- [66] LIUNG L., *System identification: Theory for the user* (second edition), Prince Hall, Englewood Cliffs, New Jersey 1999.
- [67] LUYBEN W.L., *Process modeling, simulation and control for chemical engineers* (second edition), McGraw-Hill, New York 1999.
- [68] MAŃCZAK K., NACHORSKI Z., *Komputerowa identyfikacja obiektów dynamicznych*, PWN, Warszawa 1983.
- [69] MAŃCZAK K., *Metody identyfikacji wielowymiarowych obiektów sterowania*, WNT, Warszawa 1979.
- [70] *MATLAB. User's guide*, The MATH WORKS Inc., 1992.
- [71] NIEDERLIŃSKI A., *Systemy komputerowe automatyki przemysłowej*, WNT, Warszawa 1984.
- [72] NIEDERLIŃSKI A., KASPRZYK J., FIGWER J., *Analizator wielowymiarowych sygnałów i obiektów*, Wydawnictwo Politechniki Śląskiej, Gliwice 1997.
- [73] NIEDERLIŃSKI A., *Regulowe systemy ekspertowe*, Wydawnictwo Pracowni Komputerowej Jacka Skalmierskiego, Gliwice 2000.
- [74] NELLES O., *Nonlinear system identification. From classical approaches to neural networks and fuzzy models*, Springer, Berlin, Heidelberg, New York 2001.
- [75] OSOWSKI S., *Sieci neuronowe do przetwarzania informacji*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2006.

- [76] PIENIAŻEK W., *Hydraulika i wymiana masy w destylacyjnych kolumnach wypełnionych z pulsacją fazy parowej* (rozprawa doktorska), Instytut Inżynierii Chemicznej i Urządzeń Ciepłych, Politechnika Wrocławska, Wrocław 2000.
- [77] PIENIAŻEK W., KOZIOŁ A., *Hydraulika i wymiana masy w destylacyjnych kolumnach wypełnionych z pulsacją fazy parowej*, Inżynieria Chemiczna i Procesowa, t. 22, z. 3D, 2001, 1115–1120.
- [78] PODSIADŁO P., ŚWIĄTEK J., *Adaptive control via two-stage identification*, Proc. 8th International Conference on Systems Engineering, 10–12 September 1991, Coventry, UK, Coventry Polytechnic, Coventry 1991, 127–134.
- [79] RAO C.R., *Linear Statistical Inference and its Applications*, Wiley, New York–London–Sydney–Toronto 1965.
- [80] RAO G.P., *Decomposition, decentralization and coordination of identification algorithms for large scale systems*, Proc. of 7th IFAC/IFORS Symp., York, UK, 1985, Identification and System Parameter Estimation, Pergamon Press, Oxford–New York–Toronto–Sydney–Frankfurt 1985, 279–302.
- [81] RAFAJŁOWICZ E., *Repeated least squares with inversion and its application in identifying linear distributed parameter systems*, International Journal of Systems Science, Vol. 31, No. 8, 2000, 1003–1010.
- [82] RUTKOWSKI L., *Metody i techniki sztucznej inteligencji, inteligencja obliczeniowa*, PWN, Warszawa 2006.
- [83] SAGE A.P., MELSA J.L., *Estimation Theory with Application to Communication and Control*, McGraw-Hill, New York 1972.
- [84] SAGE A.P., *Hierarchical estimation and identification method for large scale systems*, in: Singh M., Tittli A. (eds.), *Handbook of Large Scale Systems, Engineering Applications*, North-Holland, Amsterdam 1979.
- [85] SEIDLER J., BADACH A., MOLISZ W., *Metody rozwiązywania zadań optymalizacji*, WNT, Warszawa 1980.
- [86] SHAW R., *Linear Algebra and Group Representations*, Academic Press, New York 1983.
- [87] *SIMULINK Dynamic system simulation software*, The MATH WORKS Inc., 1994.
- [88] SÖDERSTRÖM T., STOICA P., *Identyfikacja systemów*, PWN, Warszawa 1997.
- [89] SZUCS B., ŚWIĄTEK J., *Two-stage approach to the modeling of the aorta*, Systems Science. Vol. 15, No. 1, 1989, 75–81.
- [90] ŚWIĄTEK J., *Model matematyczny i identyfikacja procesów hemodynamicznych w układzie krążenia* (rozprawa doktorska), Raport ICT serii PRE 84/79, Politechnika Wrocławska, Wrocław 1979.
- [91] ŚWIĄTEK J., *Some problems of mathematical models and identification of cardiovascular system*, Systems Science, Vol. 5, No. 1, 1979, 291–298.
- [92] ŚWIĄTEK J., *The Bayes approach to the parameter estimation in complex systems*, Foundations of Control Engineering, Vol. 5, No. 1, 1980, 45–52.
- [93] ŚWIĄTEK J., *Mathematical model and identification of haemodynamical process in cardiovascular system*, Proc. of International Conference on Systems Engineering, Coventry (Lanchester) Polytechnic, Coventry 1980, 216–228.
- [94] ŚWIĄTEK J., *Optimalnaja zadača ocenki parametrov v identifikacii složnych sistem*, Materiały VII Polsko-Bułagrskiego Simpozjuma: Optimalizacija i Upravljenje v Kinbernetičeskich Sistemach, Prace IBS, PAN, Warszawa, 1980, 58–65.
- [95] ŚWIĄTEK J., *Parameter identification of complex systems with limited measurement possibilities*, Proc. of International Conference on Systems Engineering, Coventry (Lanchester) Polytechnic, Coventry 1980, 249–255.
- [96] ŚWIĄTEK J., *Two stage parameter estimation in linear systems by maximum likelihood method*, Proc. 2nd International Conference on Systems Engineering, Coventry (Lanchester) Polytechnic, Coventry 1982, 15–24.
- [97] ŚWIĄTEK J., *Identification and sensitivity analysis for pulmonary circuit in the cardiovascular system*, in: Ricciardi L., Scott A. (eds.), *Biomathematics in 1980*, North-Holland, Amsterdam–New York–Oxford 1982, 265–279.

- [98] ŚWIĄTEK J., *Identification of haemodynamical process in cardiovascular system*, in: Tarpl B., Ricciardi L., Pask G. (eds.), *Progress in Cybernetics and Systems Research, Cybernetics in Biology and Medicine*, Vol. 9, Hemisphere Pub. Co. McGraw-Hill, Washington–New York–London 1982, 87–96.
- [99] ŚWIĄTEK J., *Maximum likelihood approach to the two-stage estimation in the identification of static systems*, *Systems Science*, Vol. 8, No. 1, 1982, 39–44.
- [100] ŚWIĄTEK J., *Nekotoraja problema mnogourovnevnych modelej i mnogošagovoj identifikacii, Sbornik Lekci i Soobščenija, VIII Nacionalna Škola za Mlodi Naučni Robotnici „Priloženje Matematikata v Technikata”*, Varna '82, Sofija 1983, 137–145.
- [101] ŚWIĄTEK J., *On two-stage parameter estimation in static system*, *IEEE Tran on System, Man and Cybernetics*, Vol. SMC-13, No. 1, 1983, 77–81.
- [102] ŚWIĄTEK J., *Multistage models and identification. Tagung Modellierung und Optimierung von Systemen*, Lichte 1983, *Wissenschaftliche Berichte der Technischen Hochschule Leipzig* 1984, 27–31.
- [103] ŚWIĄTEK J., *Application of the multistage identification to modeling of absorption process*, *Proc. of 3rd International Conference on Systems Engineering*, Wright State University, Dayton 1984.
- [104] ŚWIĄTEK J., *Analiza modelu układu krążenia dla potrzeb rozpoznawania wad serca*, mat. konf. „Przetwarzanie sygnałów w telekomunikacji, sterowaniu i kontroli”, Instytut Telekomunikacji i Elektrotechniki ATR, Ministerstwo Łączności, Bydgoszcz 1984, 147–150.
- [105] ŚWIĄTEK J., *Application of the two-stage identification to the modeling of packed distillation column with pulsation*, *Proc. 4th International Conference on Systems Engineering*, Coventry (Lanchester) Polytechnic, Coventry 1985, 251–258, przedruk: *Systems Science*, Vol. 11, No. 3–4, 1985, 125–134.
- [106] ŚWIĄTEK J., *Application of two-level pattern recognition system to the hormonal investigation*, *Proc. of the Colloquium on Application of Computational and Cybernetics Methods in Medicine and Biology*, Szeged 1985.
- [107] ŚWIĄTEK J., REYNOLDS D., *Two-stage approach to the modeling of the influence of drugs on the respiratory system*, *Proc. of 3rd International Conference on Systems Engineering*, Wright State University, Dayton 1984, 209–212.
- [108] ŚWIĄTEK J., *On two-stage optimal models and identification*, *Systems Science*, Vol. 11, No. 3–4, 17–30.
- [109] ŚWIĄTEK J., *Identyfikacja dwustopniowa oraz jej zastosowania techniczne i biomedyczne (praca habilitacyjna)*, *Prace Naukowe Instytutu Sterowania i Techniki Systemów Politechniki Wrocławskiej*, Monografie nr 2, Wrocław 1987.
- [110] ŚWIĄTEK J., *On the two – stage identification of dynamic systems*, *Proc. 5th International Conference on Systems Engineering*, September 9–11, 1987 Fairborn, Ohio, USA: IEEE, Dayton 1987, 49–52.
- [111] ŚWIĄTEK J., *Dwustopniowa identyfikacja kolumny destylacyjnej z wypełnieniem oraz przepływem pulsacyjnym. Dziś i jutro automatyki i robotyki w Polsce*, X Krajowa Konferencja Automatyki, t. 3, czerwiec 21–24 1988, Lublin, Politechnika Lubelska, Lublin 1988, 173–174.
- [112] ŚWIĄTEK J., *Wybrane problemy identyfikacji kompleksów operacji*, materiały VI krajowej konferencji „Automatyzacja dyskretnych procesów przemysłowych”, 02–05. 10. 1988, Kozubnik k. Porąbki, Sekcja 1: Podstawy teoretyczne, Wydawnictwo Politechniki Śląskiej, Gliwice 1988, 329–339.
- [113] ŚWIĄTEK J., *Some problems of identification for relational system*, *Systems Science*, Vol. 15, No. 1, 1989, 733–740.
- [114] ŚWIĄTEK J., ABDEL GHANY A. M., *Design of controller for two stage dynamic system. Computer applications in industry*, *Proc. 2nd IASTED International Conference*, May 5–7, Alexandria, Egipt 1992, Vol. 1, Dorrah H.T. (ed.), Acta Press, Zurich 1992, 68–71.
- [115] ŚWIĄTEK J., PLATA A., *Pakiet programów do modelowania i identyfikacji kompleksów operacji*, materiały VIII krajowej konferencji „Automatyzacja dyskretnych procesów przemysłowych”, 02–05. 10. 1988, Kozubnik k. Porąbki, Sekcja 1: Podstawy teoretyczne, Wydawnictwo Politechniki Śląskiej, Gliwice 1988, 329–339.

- słowych”. Sekcja 2: Modelowanie i sterowanie. Sekcja 3: Komputerowo zintegrowana produkcja i robotyka, 16–19. 09. 1992, Kozubnik k. Porąbki, Wydawnictwo Politechniki Śląskiej, Gliwice 1992, 125–134.
- [116] ŚWIĄTEK J., *Problems of complex systems identification for adaptive control*, Proc. 9th International Conference on Systems Engineering (ICSE '93), July 14–16, 1993, Las Vegas, Nevada State University, Las Vegas 1993, 280–284.
- [117] ŚWIĄTEK J., *Identyfikacja złożonych systemów relacyjnych*, materiały II krajowej konferencji naukowej „Inżynieria wiedzy i systemy ekspertowe”, t. 1, 15–17 czerwca 1993, Wrocław, Wydawnictwo Politechniki Wrocławskiej, Wrocław 1993, 281–288.
- [118] ŚWIĄTEK J., *On-line two-stage estimation algorithms for dynamic system*, 10th International Conference on Systems Engineering (ICSE '94) 6–8 September 1994, Coventry, UK, Vol. 2, Coventry University, Coventry 1994, 1206–1212.
- [119] ŚWIĄTEK J., *Some problems of self-tuning control via pattern recognition*, Proc. of the 12th International Conference on Systems Science, 12–15 September 1995, Wrocław, Poland, Vol. 1, Bubnicki Z., Grzech A. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1995, 533–538.
- [120] ŚWIĄTEK J., *Self-tuning control via pattern recognition for discrete type systems*, Proc. 11th International Conference on Systems Engineering (ICSE '96), July 9–11 1996, Las Vegas, USA, University of Nevada, Las Vegas 1996, 134–139.
- [121] ŚWIĄTEK J., *Two-stage estimation algorithms for adaptive control*, 12th International Conference on Systems Engineering. Proceedings, 9–11 September 1997, Vol. 2, Coventry, UK, Coventry University, Coventry 1997, 670–672.
- [122] ŚWIĄTEK J., *Identyfikacja, Problemy automatyki i informatyki*, Grzech A. (red.), Zakład Narodowy im. Ossolińskich Wydawnictwo Polskiej Akademii Nauk, Wrocław 1998, 29–44.
- [123] ŚWIĄTEK J., *Rekurencyjne algorytmy estymacji dwustopniowej*, XIII Krajowa Konferencja Automatyki, Opole, 21–24 września 1999, t. 1, Zdzisław Bubnicki Z., Józefczyk J. (red.), Oficyna Wydawnicza Politechniki Opolskiej, Opole 1999, 261–264,
- [124] ŚWIĄTEK J., *Identyfikacja globalna systemów złożonych z modelem lokalnie optymalnym*, w: Inżynieria wiedzy i systemy ekspertowe, t. 1, Bubnicki Z., Grzech A. (red.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2003, 367–374.
- [125] ŚWIĄTEK J., *Global and local modeling of complex input output system*, Proc. 16th International Conference on Systems Engineering, 9–11 September 2003, Coventry (ICSE 2003), Vol. 2, Burnham K.J., Haas O.C.L. (eds.), Coventry University, Coventry 2003, 669–671.
- [126] ŚWIĄTEK J., *Global identification of complex systems with cascade structure*, Proc. 7th International Conference Artificial Intelligence and Soft Computing (ICAISC 2004), Zakopane, June 7–11 2004, Rutkowski L. (ed.), Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence, Vol. 3070, Springer, Berlin 2004, 990–995.
- [127] ŚWIĄTEK J., *Globally optimal model of complex systems – identification algorithm for cascade structure*, Proc. 15th International Conference on Systems Science, 7–10 September 2004, Wrocław, Vol. 1, Bubnicki Z., Grzech A. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2004, 367–376.
- [128] ŚWIĄTEK J., *Modeling of complex systems by neural networks*, Information Systems Architecture and Technology (ISAT 2004), Proc. 25th International Scientific School, Information models, concepts, tools and applications, Grzech A., Wilimowska Z. (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2004, 17–24.
- [129] ŚWIĄTEK J., *Global identification of complex systems*, Proc. 9th International Symposium on Artificial Life and Robotics Oita – Beppu, Japan, January 28–30, 2004 (AROB '04), Vol. 1, Sugisaka M., Tanaka H. (eds.), Oita 2004, 291–294.
- [130] ŚWIĄTEK J., *Identyfikacja kompleksów operacji przy ograniczonych możliwościach pomiarowych*, materiały XV Krajowej Konferencji Automatyki (KKA 2005), Warszawa, 27–30 czerwca 2005, t. 1, Bubnicki Z., Kulikowski R., Kacprzyk J. (red.), Instytut Badań Systemowych PAN, Warszawa 2005, 333–336.

- [131] ŚWIĄTEK J., BRZOSTOWSKI K., *Algorytm wyboru scenariusza z wykorzystaniem identyfikacji dwustopniowej*, w: Inżynieria wiedzy i systemy ekspertowe, t. 1, Grzech A. (red.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2006, 231–236.
- [132] ŚWIĄTEK J., *Identyfikacja systemów złożonych*, materiały III konferencji naukowej „Informatyka w ekonomii”, Wyższa Szkoła Zarządzania i Administracji, Zamość 2006, 25–33.
- [133] ŚWIĄTEK J., *Parameter estimation of systems described by the relation by maximum likelihood method*, Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence, Vol. 4029, 2006, 1217–1222.
- [134] ŚWIĄTEK J., *Parameter estimation of systems described by the relation with noisy observation*, Journal of Universal Computer Science, Vol. 13, No. 2, 2007, 199–208.
- [135] ŚWIĄTEK J., *Selected problems of complex systems identification*, in: Knowledge processing and reasoning for information society, Ngoc Thanh Nguyen, Kołaczek G., Gabryś B. (eds.), Akademicka Oficyna Wydawnicza EXIT, Warszawa 2008, 201–229.
- [136] TADEUSIEWICZ R., *Sieci neuronowe*, Akademicka Oficyna Wydawnicza RM, Warszawa 1998.
- [137] TKACZYŃSKI T., TKACZYŃSKA D., *Synteza i technologia chemiczna leków: podręcznik dla studentów farmacji*, PZWL, Warszawa 1984.
- [138] WIERZBICKI A., *Modele i wrażliwość układów sterowania*, WNT, Warszawa 1997.
- [139] ZWOŹDZIAK J., *Metody identyfikacji źródeł emisji pyłów i oceny ich oddziaływania*, Prace Naukowe Instytutu Inżynierii Ochrony Środowiska Politechniki Wrocławskiej, seria: Monografie nr 24, Wrocław 1986.



Jerzy Świątek urodził się w 1953 r. w Jaśle. W roku 1977 uzyskał tytuł magistra inżyniera elektronika na Wydziale Elektroniki, a w 1979 r. stopień doktora nauk technicznych w Instytucie Cybernetyki Politechniki Wrocławskiej. Stopień doktora habilitowanego nauk technicznych z zakresu automatyki i robotyki nadała mu Rada Wydziału Elektroniki, Automatyki i Elektrotechniki Akademii Górniczo-Hutniczej w Krakowie w 1987 r. Jest profesorem Politechniki Wrocławskiej. Jego zainteresowania naukowe skupiają się wokół zagadnień identyfikacji, rozpoznawania oraz

wspomagania decyzji. Jest członkiem Komitetu Automatyki i Robotyki PAN oraz Wrocławskiego Towarzystwa Naukowego.

W latach 1993–1999 pełnił funkcję dziekana Wydziału Informatyki i Zarządzania, a w latach 1999–2005 prorektora ds. nauczania Politechniki Wrocławskiej. Od roku 2005 ponownie pełni funkcję dziekana Wydziału Informatyki i Zarządzania.

W książce przedstawiono wybrane problemy identyfikacji obiektów statycznych, czyli tworzenie modeli matematycznych obiektów na podstawie wyników eksperymentu oraz omówiono rolę modelu w badaniach systemowych. Wprowadzono pojęcia dotyczące zadania identyfikacji oraz wskazano aktualne problemy identyfikacji systemów złożonych. Sformułowano podstawowe zadania identyfikacji obiektów statycznych. Omówiono zadania wyznaczania parametrów charakterystyk statycznych oraz zadania wyboru optymalnego modelu w warunkach deterministycznych i losowych, to znaczy takich, gdy na obiekt działają pewne wielkości losowe, a wyniki eksperymentu są zniekształcone zakłóceniami pomiarowymi.

Zasadnicza część książki dotyczy identyfikacji statycznych systemów złożonych, czyli identyfikacji takich obiektów, w których wyróżniono elementy składowe i wskazano powiązania między nimi. Przedstawiono: identyfikację przy ograniczonych możliwościach pomiarowych, identyfikację globalną oraz identyfikację dwustopniową.

Publikacja jest przeznaczona dla osób zainteresowanych tworzeniem modeli matematycznych obiektów złożonych. Adresowana jest przede wszystkim do studentów i pracowników naukowych kierunków: automatyka i robotyka, ekonometria, informatyka, matematyka.

ISBN 978-83-7493-442-8

ISSN 1640-8969



9 788374 193442