

Akademia Górniczo-Hutnicza im. Stanisława Staszica  
Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki  
Katedra Automatyki

---

# ODTWARZANIE STANU SYSTEMÓW DYNAMICZNYCH Z DYSKRETNÝCH DANYCH POMIAROWYCH

---

*Autor:*  
JERZY BARANOWSKI

Promotor:  
prof. dr hab. inż. Wojciech Mitkowski

Marzec 2010



*Mojej cudownej Żonie*

## Podziękowania

Składam serdeczne podziękowania Prof. Wojciechowi Mitkowskiemu za cenne wskazówki, poprawki, opiekę naukową oraz przede wszystkim za zainteresowanie mnie teorią sterowania i umożliwienie mi rozwoju naukowego. Dziękuję również kolegom z Katedry Automatyki za współpracę, liczne dyskusje i podsuwanie nowych pomysłów.

Dziękuję również Ministerstwu Nauki i Szkolnictwa Wyższego za sfinansowanie moich badań w ramach projektu badawczego promotorskiego, nr N N 514 417734.

Dziękuję moim Rodzicom za ciągłą pomoc i zachętę podczas okresu studiów doktoranckich oraz moim Teściom, dzięki gościnności których mogłem w spokoju napisać większość tej pracy. Na sam koniec chciałem podziękować mojej Żonie Marcie, która jest dla mnie wsparciem każdego dnia.

*(Kraków-Dębowiec Mały, Marzec 2010)*

# Przyjęte oznaczenia

$\mathbb{R}$  Zbiór liczb rzeczywistych

$\mathbb{R}^n$  Rzeczywista przestrzeń euklidesowa  $n$  wymiarowa z normą

$$|a| = \sqrt{\sum_{i=1}^n a_i^2}$$

przyjmujemy, że elementy  $\mathbb{R}^n$  są wektorami pionowymi, a iloczyn skalarny wektorów  $a$  i  $b$  oznaczamy  $(a, b) = a^T b$  (str. 42).

$\mathbb{R}^{m \times n}$  Zbiór macierzy rzeczywistych  $m$  na  $n$  wymiarowych z normą spektralną (str. 31)

$$\|A\| = \max_{1 \leq k \leq n} \sqrt{\lambda_k(A^T A)}$$

$C^0([a, b], \mathbb{R}^n)$  Przestrzeń Banacha funkcji ciągłych na odcinku  $[a, b]$  o wartościach w zbiorze  $\mathbb{R}^n$  z normą

$$\|x\|_C = \sup_{t \in [a, b]} |x(t)|$$

$L^2([a, b], \mathbb{R}^n)$  Przestrzeń Hilberta funkcji całkowalnych z kwadratem na odcinku  $[a, b]$  o wartościach w zbiorze  $\mathbb{R}^n$  z iloczynem skalarnym

$$(x, y)_2 = \int_a^b x(t)^T y(t) dt$$

i normą

$$\|x\|_2 = \sqrt{(x, x)_2}$$

$L^\infty([a, b], \mathbb{R}^n)$  Przestrzeń Banacha funkcji istotnie ograniczonych na odcinku  $[a, b]$  o wartościach w zbiorze  $\mathbb{R}^n$  z normą

$$\|x\|_\infty = \text{ess sup } |x| = \{\inf C \geq 0 : |x(t)| \leq C \text{ dla prawie wszystkich } t\}$$

$$C^0[a, b] = C^0([a, b], \mathbb{R})$$

$\lambda(A)$  Widmo macierzy  $A$

$\lambda_i(A)$   $i$ -ta wartość własna macierzy  $A$

$W(\lambda)$  Wielomian charakterystyczny (str. 44)

$A \geq 0$  Macierz  $A$  jest dodatnio określona, t.j.  $\forall x, x^T A x \geq 0$

$A > 0$  Macierz  $A$  jest ściśle dodatnio określona, t.j.  $\forall x \neq 0, x^T A x > 0$

$\text{Ker}(A)$  Jądro macierzy  $A$

$\hat{x}(t)$  Estymata wartości stanu  $x$  w chwili  $t$

$\varepsilon(t) = x(t) - \hat{x}(t)$  Błąd estymacji w chwili  $t$

$\bar{\varepsilon}_k$  Norma błędu estymacji traktowanego jako funkcja ciągła na odcinku  $[t_k, t_{k+1}]$  (str. 15)

$\nabla_x F(x, u)$  Gradient funkcjonału  $F$  po zmiennych z wektora  $x$

$$\nabla_x F(x, u) = \begin{bmatrix} \frac{\partial F(x, u)}{\partial x_1} \\ \frac{\partial F(x, u)}{\partial x_2} \\ \vdots \\ \frac{\partial F(x, u)}{\partial x_r} \end{bmatrix}$$

gdzie  $x \in \mathbb{R}^r$ . Jeżeli indeks będzie pominięty, oznacza to że różniczkowanie następuje po wszystkich zmiennych funkcjonału (zarówno  $x$  i  $u$ ) (str. 67)

$\nabla^2 F(x)$  Hesjan funkcjonału  $F$

$$\nabla^2 F(x) = \begin{bmatrix} \frac{\partial^2 F(x)}{\partial x_1^2} & \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 F(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 F(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F(x)}{\partial x_1 \partial x_n} & \frac{\partial^2 F(x)}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 F(x)}{\partial x_n^2} \end{bmatrix}$$

(str. 75). W sytuacji, gdy będziemy liczyć Hesjan funkcji złożonej  $F(x)$  gdzie  $x = f(y)$  będziemy pisać  $\nabla_y^2 F(x)$ .

$J(x) = \frac{d}{dx} f(x, u)$  Macierz Jakobiego funkcji wektorowej  $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  takiej, że

$$f(x, u) = \begin{bmatrix} f_1(x, u) \\ f_2(x, u) \\ \vdots \\ f_n(x, u) \end{bmatrix}$$

zaś sama macierz jest dana wzorem

$$J(x) = \frac{d}{dx} f(x, u) = \begin{bmatrix} \frac{\partial f_1(x, u)}{\partial x_1} & \frac{\partial f_1(x, u)}{\partial x_2} & \cdots & \frac{\partial f_1(x, u)}{\partial x_n} \\ \frac{\partial f_2(x, u)}{\partial x_1} & \frac{\partial f_2(x, u)}{\partial x_2} & \cdots & \frac{\partial f_2(x, u)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x, u)}{\partial x_1} & \frac{\partial f_n(x, u)}{\partial x_2} & \cdots & \frac{\partial f_n(x, u)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla_x f_1(x, u)^\top \\ \nabla_x f_2(x, u)^\top \\ \vdots \\ \nabla_x f_n(x, u)^\top \end{bmatrix}$$

(str. 67)

$|\alpha|_A^2 = (\alpha, A\alpha) = \alpha^\top A\alpha$  Alternatywny zapis formy kwadratowej z macierzą  $A$  (str. 72)

$\mathcal{K}$  Klasa funkcji, (str. 19)

$\mathcal{K}^\infty$  Klasa funkcji, (str. 19)

$\mathcal{KL}$  Klasa funkcji, (str. 101)

$\text{sign}(x)$  Funkcja signum, definiowana jako

$$\text{sign} x = \begin{cases} 1 & \text{dla } x > 0 \\ 0 & \text{dla } x = 0 \\ -1 & \text{dla } x < 0 \end{cases}$$

# Spis treści

|                                                                       |           |
|-----------------------------------------------------------------------|-----------|
| <b>Podziękowania</b>                                                  | <b>3</b>  |
| <b>Przyjęte oznaczenia</b>                                            | <b>4</b>  |
| <b>1. Wstęp i tezy pracy</b>                                          | <b>8</b>  |
| 1.1. Wprowadzenie . . . . .                                           | 8         |
| 1.2. Tezy pracy . . . . .                                             | 9         |
| 1.3. Zawartość pracy . . . . .                                        | 9         |
| <b>2. Odtwarzanie ciągłego stanu z dyskretnych danych pomiarowych</b> | <b>12</b> |
| 2.1. Algorytm . . . . .                                               | 13        |
| 2.1.1. Predykcja . . . . .                                            | 15        |
| 2.1.2. Korekcja . . . . .                                             | 15        |
| 2.2. Stabilność . . . . .                                             | 15        |
| 2.3. Obserwowalność . . . . .                                         | 19        |
| 2.4. Zasada rozdzielności . . . . .                                   | 21        |
| 2.4.1. Zastosowanie stabilności wejście-stan . . . . .                | 21        |
| 2.5. Okna pomiarowe . . . . .                                         | 24        |
| <b>3. Predykcja</b>                                                   | <b>27</b> |
| 3.1. Generacja ciągłych estymat . . . . .                             | 27        |
| 3.1.1. Metoda RK3/8 . . . . .                                         | 29        |
| 3.2. Błędy przejściowe predykcji . . . . .                            | 31        |
| 3.2.1. Szacowanie błędu predykcji w układach liniowych . . . . .      | 33        |
| 3.2.2. Szacowanie błędu predykcji w układach nieliniowych . . . . .   | 37        |
| 3.3. Błędy predykcji wywołane zakłóceniami wejściowymi . . . . .      | 38        |
| <b>4. Korekcja w układach liniowych</b>                               | <b>40</b> |
| 4.1. Wprowadzenie . . . . .                                           | 40        |
| 4.2. Estymacja bez zakłóceń . . . . .                                 | 42        |
| 4.2.1. Obserwatory dyskretno-ciągłe . . . . .                         | 43        |
| 4.2.2. Dyskretny obserwator dokładny . . . . .                        | 44        |
| 4.3. Estymacja z zakłóceniami wyjścia . . . . .                       | 46        |
| 4.3.1. Obserwator najmniejszych kwadratów . . . . .                   | 46        |
| 4.3.2. Obserwacja $C^0$ -optymalna . . . . .                          | 47        |
| 4.4. Estymacja z zakłóceniami wyjścia i wejścia . . . . .             | 52        |
| 4.4.1. Dyskretno-ciągły filtr Kalmana . . . . .                       | 53        |
| 4.4.2. Inne podejścia . . . . .                                       | 53        |
| 4.5. Przykłady zastosowań . . . . .                                   | 54        |
| <b>5. Korekcja w układach nieliniowych</b>                            | <b>61</b> |
| 5.1. Wprowadzenie . . . . .                                           | 61        |
| 5.2. Estymacja bez zakłóceń . . . . .                                 | 62        |
| 5.2.1. Nieliniowy obserwator dyskretno-ciągły . . . . .               | 62        |
| 5.2.2. Hybrydowy obserwator Newtona . . . . .                         | 65        |

## Spis treści

|                                                                                                 |            |
|-------------------------------------------------------------------------------------------------|------------|
| 5.3. Estymacja z zakłóceniami . . . . .                                                         | 70         |
| 5.3.1. Dyskretno-ciągły rozszerzony filtr Kalmana . . . . .                                     | 70         |
| 5.3.2. Estymacja z ruchomym horyzontem ( <i>ang. moving horizon estimation, MHE</i> ) . . . . . | 71         |
| 5.4. Przykłady zastosowań . . . . .                                                             | 79         |
| <b>6. Metody przybliżone korekcji</b>                                                           | <b>85</b>  |
| 6.1. Przybliżanie dyskretnego wyjścia funkcją ciągłą . . . . .                                  | 85         |
| 6.1.1. Interpolacja . . . . .                                                                   | 85         |
| 6.1.2. Kolokacje wyjścia . . . . .                                                              | 86         |
| 6.1.3. Adaptacyjne kolokacje wyjścia . . . . .                                                  | 89         |
| 6.2. Metody przybliżone dla układów liniowych . . . . .                                         | 90         |
| 6.2.1. Ciągłe obserwatory asymptotyczne . . . . .                                               | 90         |
| 6.2.2. Estymacja $L^2$ -optymalna . . . . .                                                     | 90         |
| 6.3. Metody przybliżone dla układów nieliniowych . . . . .                                      | 91         |
| 6.3.1. Ciągłe nieliniowe obserwatory asymptotyczne . . . . .                                    | 91         |
| 6.4. Przykłady . . . . .                                                                        | 93         |
| <b>7. Podsumowanie</b>                                                                          | <b>97</b>  |
| <b>A. Istotne twierdzenia</b>                                                                   | <b>98</b>  |
| A.1. Lematy i twierdzenia pomocnicze . . . . .                                                  | 98         |
| A.2. Stabilizacja regulatorem liniowym . . . . .                                                | 99         |
| <b>B. Stabilność wejście-stan (Input-to-state stability)</b>                                    | <b>101</b> |
| <b>Bibliografia</b>                                                                             | <b>103</b> |
| <b>Skorowidz</b>                                                                                | <b>111</b> |
| <b>English abstract</b>                                                                         | <b>113</b> |
| <b>Streszczenie</b>                                                                             | <b>117</b> |

# Rozdział 1.

## Wstęp i tezy pracy

### 1.1. Wprowadzenie

Problem odtwarzania stanu towarzyszy ludzkości od samego początku. Człowiek pierwotny widząc ogień chciał wiedzieć skąd się on wzięł, co go zapoczątkowało. Starożytni rzymianie mówili „*Oculus animi index*” - oczy są zwierciadłem duszy, czyli patrząc w oczy możemy poznać naturę drugiego człowieka. Jeżeli potraktujemy oczy jako wyjście (czy może lepiej okno na świat), to zdobycie pełnej wiedzy poprzez obserwację wyjścia jest w swojej istocie właśnie odtwarzaniem stanu! Wszyscy znamy siłę pierwszego wrażenia. Mówi się „Wystarczy raz spojrzeć, by się na kimś poznać”. Skoro oczy są wyjściem, a pojedyncze spojrzenie pomiarem to właśnie uzyskaliśmy odtwarzanie stanu z dyskretnych (w tym przypadku jednorazowych) danych pomiarowych. Jest to temat poniższej pracy, dotyczy ona jednak nie kwestii psychologiczno społecznych a zastosowań w automatyce.

Rozważmy następujący przykład. W technologii chemicznej często mamy do czynienia z produkcją dużych ilości substancji. Znane są parametry procesu, takie jak stężenia substratów, dostarczana energia, parametry mieszania i inne. Możemy przyjąć, że znamy ich wartości przez cały czas. Analiza powstającego produktu nie jest jednak rzeczą prostą. Jest ona zazwyczaj czasochłonna (trzeba dokonać analizy chemicznej) i wiąże się z pewnymi, często niebagatelnymi kosztami (odczynniki, czasu pracy specjalistów). Praktyka i rachunek ekonomiczny wymuszają więc wykonywanie takiej analizy tylko co pewien czas (im rzadziej tym lepiej). Skład produktu jest jednak jedyną informacją o przebiegu procesu. To, że badamy go tylko w dyskretnych chwilach czasu, nie zmienia faktu, że wewnątrz linii produkcyjnej, bez jakichkolwiek pomiarów, przemiany chemiczne zachodzą w sposób ciągły, często gwałtowny. Stąd pojawia się konieczność odpowiedzi na pytanie - jak odtworzyć pełną informację o przebiegu procesu, przez cały czas jego trwania?

Przykłady gdzie pomiar jest dyskretny można mnożyć. W układach mechatronicznych bardzo często mamy do czynienia z dyskretnymi pomiarami pozycji (a prędkości zazwyczaj w ogóle nie mierzymy wprost). Jest to wywołane stosowaniem dyskretnych czujników (zob. np. Górecki, 1997) lub też wykorzystaniem systemów wizyjnych (zob. np. Papanikolopoulos, Khosla i Kanade, 1993). Szczególnie w tym ostatnim przypadku odstępy między pomiarami są wydłużane przez czas potrzebny na przetwarzanie obrazu. Oczywistym jest jednak, że dysponując pełną informacją o stanie systemu (prędkości i położenia) jesteśmy w stanie opracować dużo lepsze sterowanie.

Jeszcze inna klasa problemów, w których pomiary są z natury dyskretnie to sytuacje, w których automatyka styka się z biologią, a w szczególności medycyną. Oczywistym jest, że pacjenta nie można badać w sposób ciągły. Jego krew można pobierać tylko w ograniczonych ilościach. Zdjęć rentgenowskich nie powinno się dokonywać częściej niż raz do roku. Podobnie ograniczony czasowo jest szereg innych badań (endoskopia, punkcje itp.), nie mówiąc już o ekonomicznych czynnikach przeprowadzania konkretnych elementów diagnostyki. W sytuacji gdy interesujący jest np. poziom cukru i insuliny we krwi pacjenta metody estymacji stanu okazują się nad wyraz pomocne (np. Chee i Fernando, 2007).

Dlatego też problematykę odtwarzania stanu systemów dynamicznych z dyskretnych danych pomiarowych można uznać za ważną zarówno poznawczo jak też i ze względu na potencjalne zastosowania. W tej pracy będziemy się skupiać na sytuacjach, gdy dyskretnego charakteru pomiarów nie można zaniedbać i do odtwarzania stanu konieczne jest uwzględnianie konkretnych czynników.

*Uwaga 1.1.* W ramach niniejszej pracy pojęcia *estymacja stanu* i *odtworzenie stanu* będą traktowane jako synonimy. Tematyką niniejszej pracy nie jest statystyczne lub stochastyczne ujęcie problemu, dlatego też słowo estymacja będzie rozumiane w kontekście deterministycznym. Podobnie za równoważne będziemy uważać pojęcia *estymator* i *obserwator*. Nie oznacza to jednak, że znacznej części metod opisanych w tej pra-



cy nie można nadać interpretacji stochastycznej. W szczególności interpretacje te łatwo znaleźć dla filtrów Kalmana (liniowego i rozszerzonego) oraz metody MHE. Pojawiające się przy ich opisie sformułowania o charakterze probabilistycznym mają na celu jedynie jaśniejsze sformułowanie problemów, zaś w przypadku filtrów Kalmana mają dodatkowo znaczenie historyczne. ■

## 1.2. Tezy pracy

**Teza 1. Możliwe jest opracowanie metod odtwarzania ciągłego stanu z dyskretnych danych pomiarowych w liniowych oraz skończenie wymiarowych nieliniowych systemach dynamicznych.**

Znajomość modelu matematycznego systemu dynamicznego liniowego lub nieliniowego pozwala na różne możliwości jego analizy. Wykorzystanie dostępnego w teorii sterowania aparatu matematycznego umożliwia skonstruowanie sposobów odtwarzania wartości stanu w konkretnych punktach, oraz przewidywanie na jej podstawie zachowania się układu w przyszłości. W oparciu o dostępną wiedzę można zaproponować algorytm postępowania pozwalający na cykliczne uzyskiwanie ciągłych estymat stanu na przedziałach pomiędzy pomiarami. Możliwe jest też konstruowanie obserwatorów stanu o działaniu asymptotycznym, dokładnym lub optymalnym w obecności zakłóceń. Co więcej, można proponować sposoby doboru takich chwil pomiaru, aby uzyskana z ich użyciem estymata stanu była optymalna. Dodatkowo, w układach nieliniowych osiągalne jest uzyskanie efektywnych metod numerycznych, pozwalających na rozwiązanie problemu odtwarzania stanu.

**Teza 2. Z wykorzystaniem metod interpolacji i aproksymacji możliwe jest opracowanie efektywnych obliczeniowo algorytmów przybliżonego odtwarzania stanu**

Dla układów ciągłych pojęcia estymacji są rozwijane od dawna i dysponując ciągłym wyjściem można uzyskać bardzo efektywne obserwatory. Istotnym jest więc poszukiwanie sposobów aproksymacji dyskretnego wyjścia funkcją ciągłą i wykorzystanie metodyki ciągłego odtwarzania stanu tak, aby uzyskać przybliżone metody odtwarzania stanu. Jest to ważne zwłaszcza dla układów nieliniowych, dla których bardzo trudno (lub jest to w ogóle niemożliwe) skonstruować dokładny model dyskretny systemu. Dodatkową zaletą takiego rozwiązania jest to, że ciągle obserwatory stanu są łatwe w implementacji i strojeniu, co więcej, ta prostota pozwala na różnego rodzaju usprawnienia pracy algorytmów, jak np. obliczenia równoległe.

## 1.3. Zawartość pracy

Merytoryczną zawartość pracy można podzielić na trzy zasadnicze obszary. Pierwszy z nich, zawarty w rozdziale 2, obejmuje zaproponowany ogólny algorytm odtwarzania ciągłego stanu z dyskretnych danych pomiarowych. Opiera się on na dwóch etapach: predykcji i korekcji. Predykcja to przewidywanie zachowania się systemu w oparciu o posiadaną wiedzę, zaś korekcja to aktualizacja tej wiedzy wraz z pojawieniem się nowych pomiarów.

Opisano rozważaną klasę systemów dynamicznych, w ogólności nieliniowych, o pewnych konkretnych własnościach prawej strony. Rozdział ten zawiera również, opracowane na potrzeby pracy, definicje stabilności, globalnej atraktywności, obserwowalności i rozdzielności. Dla tych pojęć udowodniono twierdzenia pozwalające na ich praktyczne stosowanie, lub też ich związki z dotychczas znanymi terminami (np. obserwowalnością parą macierzy lub stabilnością wejście-stan). W szczególności omówiono też stabilność algorytmu i jej związek ze stabilnością procedury korekcji.

Drugi obszar, zawarty w rozdziale 3, poświęcony jest pewnym konkretnym zagadnieniom związanym z predykcją. Opisuje sposoby generowania ciągłych rozwiązań równań różniczkowych oraz pogłębia analizę błędu predykcji, omówioną w rozdziale wcześniejszym. W szczególności omówiono zmiennokrokowe metody numeryczne rozwiązywania równań różniczkowych o tzw. gęstym wyjściu, podając dla ilustracji metodę RK 3/8 (zmiennokrokową metodę typu Rungego-Kutty). Podano także pewne własności pozwalające na szacowanie błędu przejściowego predykcji w układach liniowych i nieliniowych, a także metodę sprawdzenia czy predykcja nie będzie wprowadzać błędów (dla układów liniowych). Udowodniono również twierdzenie pozwalające oszacować błąd predykcji wywołany zakłóceniami.

Trzeci obszar omówiony w pracy został zawarty w rozdziałach 4, 5 i 6, opisujących różne podejścia do problemu korekcji (dla układów liniowych, nieliniowych oraz metody przybliżone).

Rozdział 4 zawiera metody korekcji dla układów liniowych. Omówiono w nim m.in. obserwatory dyskretno-ciągłe, dyskretny obserwator dokładny, obserwator najmniejszych kwadratów, obserwator  $C^0$ -optymalny, hybrydowy filtr Kalmana. W rozdziale omówiona jest procedura autonomizacji. Podany jest również związek między problemem odtwarzania stanu liniowych systemów dynamicznych a rozwiązaniem równania liniowego w odpowiedniej przestrzeni. Rozdział kończą dwa przykłady zastosowania procedur korekcji do estymacji ciągłego stanu z dyskretnych danych pomiarowych w systemie parabolicznym oraz w modelu dynamiki populacji osób chorych na cukrzycę.

Rozdział 5 zawiera metody korekcji dla układów nieliniowych. W szczególności omówiono nieliniowy obserwator dyskretno ciągły, hybrydowy obserwator Newtona, rozszerzony hybrydowy filtr Kalmana oraz estymację z ruchomym Horyzontem (MHE). Zawarte są w nim opracowane przez autora metody obliczania stosownych pochodnych po warunkach początkowych dla obserwatora Newtona i MHE. Rozdział ten posiada zestaw przykładów, wykorzystujących symulację oraz pomiary rzeczywiste układu zbiorników wodnych, znajdującego się w Katedrze Automatyki AGH, a także symulację procesu metabolizmu glukozy u pacjenta oddziału OIOM.

Rozdział 6 zawiera przybliżone algorytmy korekcji dla systemów liniowych i nieliniowych. Algorytmy te bazują na połączeniu aproksymacji dyskretnego wyjścia systemu funkcją ciągłą oraz ciągłych metod odtwarzania stanu. Omówiono w nim zastosowanie interpolacji wielomianowej, a także zaproponowanych przez autora, metod kolokacji wyjścia i adaptacyjnych kolokacji wyjścia. Jako procedury korekcji dla układów liniowych rozważa się zastosowanie ciągłego obserwatora asymptotycznego oraz obserwatora  $L^2$  optymalnego, zaś dla układów nieliniowych zastosowania obserwatorów nieliniowych. Rozdział ten również zawiera przykłady zastosowań dla układu drabinkowego  $RC$  oraz dla modelu zbiorników wodnych.

Pracę kończy podsumowanie, zawierające wyszczególnienie tych wyników zawartych w pracy, które autor uważa za oryginalne oraz dwa dodatki zawierające twierdzenia i lematy wykorzystywane w pracy, ale nie będące integralną częścią rozumowania.

*Uwaga 1.2* (O uwagach i dowodach). W niniejszej pracy, pod koniec każdej uwagi umieszczany będzie symbol ■. Pod koniec każdego dowodu umieszczany będzie zaś symbol □.

*Uwaga 1.3.* Wszystkie rysunki w pracy za wyjątkiem dwóch zostały wykonane przez autora. Rysunki innych autorów to:

- Rysunek 5.1 na stronie 80, przedstawiający schemat zbiorników wodnych, jest autorstwa doktora Andrzeja Tutaja i jest wykorzystywany za jego pełną wiedzą i zgodą.
- Rysunek 5.5 na stronie 83, przedstawiający zdjęcie stanowiska laboratoryjnego zbiorników wodnych znajdującego się w Katedrze Automatyki AGH, jest autorstwa doktora Pawła Piątka i jest wykorzystywany za jego pełną wiedzą i zgodą.

*Uwaga 1.4* (Inne podejścia łączące dyskretnne pomiary i ciągły stan). Warto wspomnieć, że w literaturze można spotkać dwa zagadnienia dotyczące analizy zachowania układów ciągłych z dyskretnymi pomiarami. Są to zmodyfikowana transformata  $\mathcal{Z}$  oraz wykorzystanie różnych typów ekstrapolatorów. Zagadnienia te nie są omawiane w pracy, dlatego też wspomnimy o nich tutaj.

Zmodyfikowana transformata  $\mathcal{Z}$  jest to metoda rozszerzenia transformaty  $\mathcal{Z}$  próbkowanego sygnału ciągłego o informacje na temat wartości sygnału między próbkami. Metodę tą przedstawiono w pracy (Barker, 1954; Gibson, 1968, str. 122) i usprawniano w pracach (Jury, 1956a,b; Gibson, 1968, str. 122). Trochę inny wariant pod nazwą „transformata sygnału przesuniętego” rozważał (Cypkin, 1965, str. 100). Idea zmodyfikowanej transformaty  $\mathcal{Z}$  jest następująca. Rozpatrzmy sygnał ciągły

$$x(t) = \begin{cases} e^{-t} & \text{dla } t \geq 0 \\ 0 & \text{dla } t < 0 \end{cases} \quad (1.1)$$

jeżeli jest on próbkowany równomiernie z krokiem  $h > 0$  to jego transformata  $\mathcal{Z}$  ma postać

$$\mathcal{Z}\{x(t)\} = \sum_{k=0}^{\infty} e^{-kh} z^{-k} = \frac{1}{1 - \frac{e^{-h}}{z}} = \frac{z}{z - e^{-h}} \quad (1.2)$$

Rozważając zmodyfikowaną transformatę  $\mathcal{Z}$  rozważa się dodatkowo parametr  $\epsilon \in [0, 1]$  a zmodyfikowaną transformatę  $\mathcal{Z}_{\epsilon}\{x(t)\}$  oblicza się w następujący sposób (zob. Kaczorek, 1977, str. 254)

$$\mathcal{Z}_{\epsilon}\{x(t)\} = \sum_{k=0}^{\infty} e^{-(kh+\epsilon h)} z^{-k} = e^{-\epsilon h} \sum_{k=0}^{\infty} e^{-kh} z^{-k} = e^{-\epsilon h} \frac{1}{1 - \frac{e^{-h}}{z}} = \frac{e^{-\epsilon h} z}{z - e^{-h}} \quad (1.3)$$

Łatwo zauważyć, że zmodyfikowana transformata  $\mathcal{Z}$  może być traktowana jako transformata wszystkich ciągów dyskretnych postaci  $x(kh + \epsilon h)$  jednocześnie. Ciąg o konkretnym przesunięciu uzyskujemy ustalając wartość  $\epsilon$  (dla  $\epsilon = 0$  uzyskujemy wzór (1.2)). Praktyka pokazała, że zmodyfikowana transformata  $\mathcal{Z}$  wnosi pewne trudności w zastosowaniach przy umiarkowanych korzyściach. Podstawowe problemy powstają przy analizie zmodyfikowanej transmitancji. Dopuszcza ona wyłącznie wejścia dyskretne, więc nie jest możliwe uzyskanie zmodyfikowanej transmitancji przez podzielenie zmodyfikowanych transformat sygnałów wejściowych i wyjściowych. Zasadniczo transformata zmodyfikowana najlepiej sprawdza się w układach dla których była pomyślana, tzn. dla układów ciągłych, wśród których w sprzężeniach zwrotnych występują impulsatory. Warto jeszcze zaznaczyć (zob. Gibson, 1968, str. 122), że zmodyfikowana transformata  $\mathcal{Z}$  pozwala analizować zachowanie układów ciągłych, które byłyby próbkowane z taką częstotliwością, że utraciłyby sterowalność i obserwowalność (zob. tw. A.3 na stronie 99). Więcej o zmodyfikowanej transformacie zob. (Górecki, 1971, str. 257).

Koncepcja ekstrapolatorów jest znana w automatyce od dawna. Opiera się ona na określaniu wartości sygnału pomiędzy próbkami w oparciu o wartości próbkowanego sygnału. Urządzenia realizujące zadanie ekstrapolacji bywają również nazywane członami podtrzymującymi. Rozważmy sygnał dyskretny  $x(kh)$ ,  $h > 0$ ,  $k = 0, 1, 2, 3, \dots$  i sygnał ekstrapolowany  $\tilde{x}$ . Najpopularniejszy w zastosowaniach jest ekstrapolator rzędu zerowego

$$\tilde{x}(t) = x(kh), \quad \forall t \in [kh, (k+1)h) \quad (1.4)$$

który zamienia sygnał dyskretny na funkcję schodkową. Inne interesujące ekstrapolatory to ekstrapolator rzędu pierwszego (zob. Tutaj, 2008, str. 160)

$$\tilde{x}(t) = x(kh) + \frac{x(kh) - x((k-1)h)}{h}(t - kh), \quad \forall t \in [kh, (k+1)h) \quad (1.5)$$

oraz ekstrapolator rzędu ułamkowego (zob. Fajarewicz i Mazer, 2006)

$$\tilde{x}(t) = x(kh) + \beta \frac{x(kh) - x((k-1)h)}{h}(t - kh), \quad \forall t \in [kh, (k+1)h), \quad \beta \in \mathbb{R} \quad (1.6)$$

Warto zauważyć, że dla  $\beta = 0$  (1.6) przechodzi w (1.4) zaś dla  $\beta = 1$  (1.6) przechodzi w (1.5). Różne typy ekstrapolatorów porównywał (Skruch, 2002). Zdaniem autora ekstrapolatory nie nadają się do odtwarzania stanu z dyskretnych danych pomiarowych, ponieważ sama ich struktura powoduje wprowadzanie błędów, których nie da się wyeliminować. ■

## Rozdział 2.

### Odtwarzanie ciągłego stanu z dyskretnych danych pomiarowych

W ramach niniejszej pracy rozważane będą systemy dynamiczne opisane następującym równaniem różniczkowym

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.1)$$

przy czym

- $f: \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$
- $x(t) \in \mathbb{R}^n$ , będzie nazywane stanem,
- $u(t) \in \mathbb{R}^r$ , będzie nazywane sterowaniem,
- $t \geq 0$
- $x(0) = x_0$
- $f(a, b)$  spełnia globalny warunek Lipschitza ze względu na pierwszy argument, t.j.

$$\begin{aligned} |f(x_1, b) - f(x_2, b)| &\leq L|x_1 - x_2| \\ \forall x_1, x_2 \in X \quad \forall b \in \mathbb{R}^r \end{aligned} \quad (2.2)$$

ze stałą Lipschitza  $L > 0$  i jest ciągła ze względu na obydwa argumenty.

Do tej klasy układów dynamicznych zaliczyć można wiele układów nieliniowych oraz układy liniowe

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (2.3)$$

gdzie  $t \geq 0$ ,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times r}$ , dla których warunek (B.11) jest spełniony ze stałą Lipschitza  $L = \|A\|$ . Układy liniowe oraz nieliniowe będą w ramach pracy rozpatrywane oddzielnie.

Rozważamy wyjście systemu (2.1) w formie

$$y(t) = h(x(t)) \quad (2.4)$$

gdzie,  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , w układach liniowych równanie (2.4) ma postać

$$y(t) = Cx(t) \quad (2.5)$$

przy czym  $C \in \mathbb{R}^{m \times n}$ . W pewnych szczególnych przypadkach będziemy uwzględniać występowanie zakłóceń, zarówno addytywnych pomiaru, jak też i bezpośrednio wchodzących na system, wtedy równania (2.1) i (2.4) przyjmą postać

$$\dot{x}(t) = f(x(t), u(t), w(t)) \quad (2.6)$$

$$y(t) = h(x(t), v(t)) \quad (2.7)$$

gdzie  $f: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ ,  $h: \mathbb{R}^n \times \mathbb{R}^s \rightarrow \mathbb{R}^m$ . Dla systemów liniowych (2.3) i (2.5) będzie to postać

$$\dot{x}(t) = Ax(t) + Bu(t) + Lw(t) \quad (2.8)$$

$$y(t) = Cx(t) + Mv(t) \quad (2.9)$$

gdzie  $L \in \mathbb{R}^{n \times p}$ ,  $M \in \mathbb{R}^{m \times s}$ .

Przyjmujemy, że wartości wyjścia znane są w dyskretnych chwilach czasu  $t_i$ , gdzie

$$t_i = t_{i-1} + h_i \quad (2.10)$$

$$t_0 \geq 0 \quad (2.11)$$

$$h_i > 0 \quad (2.12)$$

$$i = 0, 1, 2, 3, \dots \quad (2.13)$$

przy czym

$$h_i \geq \tau_{max} \quad (2.14)$$

$$\tau_{max} = \max_i (t_{i+1} - t_i) < \infty \quad (2.15)$$

Pomiary mogą, ale nie muszą, być równoodległe. W takim przypadku mamy

$$h_i = h = \tau_{max} \quad \forall i \quad (2.16)$$

$$t_i = t_0 + ih \quad (2.17)$$

$h$  będziemy nazywać okresem próbkowania (czasem lub krokiem dyskretyzacji). Tak określone wyjście systemu dynamicznego jest często spotykane w zastosowaniach. Można je interpretować jako sytuację, w której dostępna jest niepełna informacja o stanie systemu, co więcej, informacja ta dotyczy tylko konkretnych chwil czasu.

System dynamiczny (2.1) jest ciągły w czasie (opisany jest równaniem różniczkowym), zatem klasyczne metody odtwarzania ciągłego stanu z wyjścia systemu nie mogą zostać bezpośrednio zastosowane. Stwarza to konieczność opracowania metody odtwarzania ciągłego stanu w oparciu o dyskretnie dane pomiarowe. W niniejszej pracy proponowany jest algorytm rozwiązujący ten problem. Rozważany algorytm składa się z dwóch cyklicznie powtarzających się etapów - predykcji i korekcji. W ramach tego rozdziału omawiane są pewne podstawowe własności algorytmu, w dalszej części pracy omawiane są poszczególne jego etapy.

## 2.1. Algorytm

Bazując na krokach **predykcji i korekcji** można zaproponować ogólny schemat algorytmu, którego podstawowe kroki mogą być dostosowywane w zależności od potrzeb. Schemat ten ma postać:

**Algorytm 1.** Odtwarzanie ciągłego stanu z dyskretnych danych pomiarowych

1. *START*

Ustalić  $\hat{x}(t_0)$  jako początkową wartość estymaty,  $i := 0$

2. *PREDYKCJA*

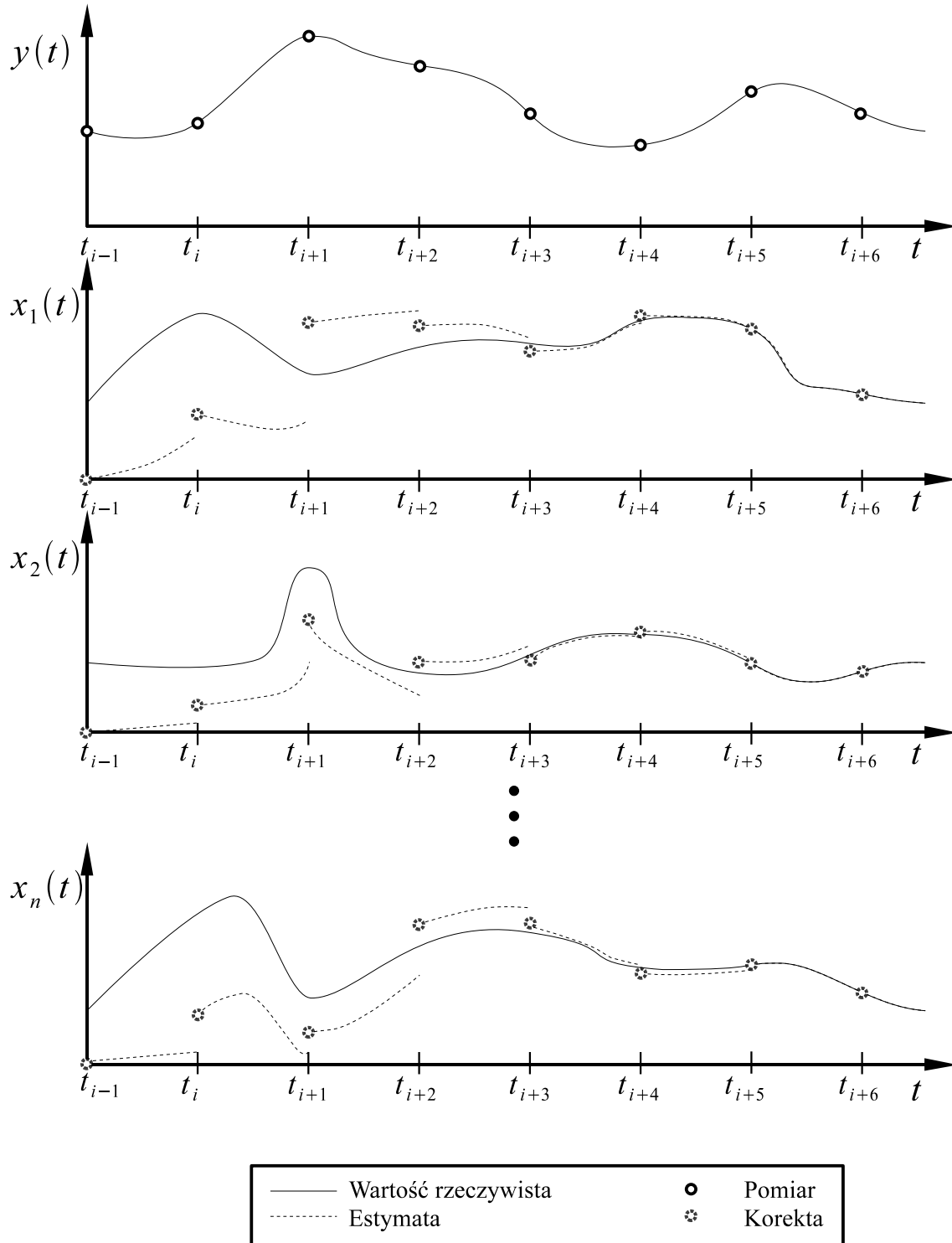
Rozwiązać równanie różniczkowe  $\dot{\hat{x}}(t) = f(\hat{x}(t), u(t))$  (równanie systemu (2.1)) dla  $t \in [t_i, t_{i+1}]$  z warunkiem początkowym  $\hat{x}(t_i)$ .

3. *KOREKCJA*

Za pomocą wartości  $y(t_j)$  dla  $j \leq i + 1$  ( $j = i + 1, i, i - 1, \dots, 0$ ) oraz przebiegu  $u(t)$  dla  $t \leq t_{i+1}$  wyznaczyć estymatę  $\hat{x}(t_{i+1})$ .

4.  $i := i + 1$ , przejść do kroku 2.

Ideę algorytmu ilustruje rysunek 2.1. Ciągłe przebiegi zmiennych stanu oraz wyjścia na wykresach rysunku oznaczone są liniami ciągłymi. Pomiary wyjścia  $y(t)$  systemu (2.1), przedstawionego na pierwszym wykresie rysunku, oznaczone okręgami o ciągłym brzegu, służą do wyznaczania estymat poszczególnych zmiennych stanu od  $x_1$  do  $x_n$  w kolejnych chwilach czasu. Estymaty te, oznaczone okręgami o przerywanym brzegu,



Rysunek 2.1.: Objasnienie do CSEDOM

służą jako warunki początkowe dla przewidywanych w kroku predykcji ciągłych estymat stanu, oznaczonych linią przerywaną. Jak można zauważyć, w momencie pojawienia się nowego pomiaru wyjścia estymaty są aktualizowane (następuje korekcja), i w przebiegu estymaty obserwowany jest skok jej wartości. W ten oto sposób algorytm generuje przedziałami ciągłą estymatę stanu systemu (2.1). Warto zauważyć, że im bliżej trajektorii stanu znajduje się wartość estymaty, tym skoki są mniejsze, zaś estymata staje się coraz bliższa ciągłej. W sytuacji, gdy estymata będzie dokładnie równa wartości stanu, a wyjście systemu będzie

niezakłócone, korekcja nie będzie już wносить zmian.

Omówmy teraz dokładniej poszczególne kroki algorytmu.

### 2.1.1. Predykcja

Wiadomo, że dla systemów postaci (2.1), które spełniają warunek (B.11) (przynajmniej lokalnie), znając warunek początkowy  $x(t_0)$  w chwili  $t_0$  oraz przebieg sterowania  $u(t)$  na odcinku czasu  $[t_0, T]$ , można jednoznacznie określić trajektorię systemu na tym samym odcinku czasu, przy czym, jeżeli warunek (B.11) jest spełniony globalnie (np. dla prawej strony (2.1) o ograniczonej pochodnej)  $T$  może być dowolnie duże (nazywane jest to nieograniczonym przedłużaniem rozwiązań). Wynika to wprost z twierdzenia o istnieniu i jednoznaczności rozwiązania równania różniczkowego (zob. np. Kudrewicz, 1976, str. 47, Mitkowski, 2007b, str. 191).

Analogicznie, znając wartość stanu  $x(t_i)$  w chwili  $t_i$  oraz przebieg sterowania na odcinku  $[t_i, T]$ , można jednoznacznie uzyskać przy przyjętych założeniach trajektorię systemu na odcinku czasu  $[t_i, T]$ . Niech  $T = t_{i+1}$  wtedy znając wartości stanu w chwilach  $t_i$  dla kolejnych wartości  $i$ , można sekwencyjnie przewidywać ciągle trajektorie stanu systemu dla  $[t_i, t_{i+1}]$ . Postępowanie takie nazywać będziemy **predykcją** (*ang. prediction*). Szczegółom predykcji poświęcony jest rozdział 3.

### 2.1.2. Korekcja

Na przykład, rozważmy chwilę  $t_{i+3}$  z rysunku 2.1. Dysponujemy dyskretnymi pomiarami wyjścia z chwil  $t_j$ ,  $j = 0, 1, \dots, i+3$ , oraz przebiegiem sterowania na przedziale  $t \in [t_0, t_{i+3}]$ . Istnieje wiele metod estymacji stanu, które wykorzystują albo wszystkie te dane albo pewien ich podzbiór. Zależy to od wymagań na szybkość obliczeń, dokładność, oraz od tego, czy system jest liniowy czy nieliniowy. Przykładowo, dla systemów nieliniowych rzędu  $n$  o wyjściu skalarnym można sformułować tzw. hybrydowy obserwator Newtona (zob. rozdział 5, str. 65), który do estymacji stanu będzie wykorzystywał pomiary  $y(t_k)$  dla  $k = i+4-n, i+4-n, \dots, i+2, i+3$  oraz  $u(t)$  dla  $t \in [t_{i+4-n}, i+3]$ .

Oczywiście jest to tylko jeden z możliwych sposobów odtworzenia stanu. Poprzez pojęcie **korekcji** (*ang. correction*) będziemy rozumieć wykorzystanie dostępnej do chwili obecnej informacji na temat wyjścia i sterowania w celu odtworzenia wartości stanu systemu w tej chwili. Różnym metodom estymacji stanu, które mogą być wykorzystane do korekcji poświęcone będą rozdziały 4, 5 i 6, dotyczące odpowiednio korekcji w układach liniowych, korekcji w układach nieliniowych oraz przybliżonym metodom korekcji.

Algorytm 1 pod angielską nazwą *Continuous State Estimation from Discrete Output Measurements* lub skrótem CSEDOM wprowadzony został przez autora w pracy (Mitkowski i Baranowski, 2007) i był rozwijany później w pracach (Baranowski, 2008c; Baranowski i Tutaj, 2007, 2008b,a; Baranowski, 2007a,c,b,d; Baranowski i Mitkowski, 2008; Baranowski, 2008a,b; Baranowski i Mitkowski, 2009; Baranowski, 2009b,a).

## 2.2. Stabilność

Przy analizie algorytmu 1 konieczne jest określenie jak zachowuje się błąd ciągłej estymaty wraz z upływem czasu względem prawdziwego stanu. Oznaczmy błąd estymacji jako

$$\varepsilon(t) = x(t) - \hat{x}(t) \quad (2.18)$$

Udowodnimy teraz następujący

**Lemat 2.1.** *Dany jest system (2.1) z dyskretnym wyjściem (2.4). Do estymacji stanu systemu stosowany jest algorytm 1. Niech*

$$\bar{\varepsilon}_k = \sup_{t \in [t_k, t_{k+1}]} |\varepsilon(t)| \quad (2.19)$$

*oznacza normę błędu estymacji traktowanego jako funkcja ciągła na odcinku  $[t_k, t_{k+1}]$ . Dla każdego  $k$  zachodzi następujące oszacowanie*

$$|\varepsilon(t_k)| \leq \bar{\varepsilon}_k \leq \alpha |\varepsilon(t_k)| \quad (2.20)$$

## Rozdział 2. Odtwarzanie ciągłego stanu z dyskretnych danych pomiarowych

przy czym  $\alpha = e^{L\tau_{max}}$  gdzie  $L$  to stała Lipschitza (zob. (B.11)), zaś  $\tau_{max} = \max_k (t_{k+1}, t_k)$ ,  $k = 0, 1, 2, \dots$

*Dowód.* Zaczniemy od oszacowania górnego. Niech stan systemu (2.1) pomiędzy chwilami  $t_k$  i  $t_{k+1}$  ewoluuje zgodnie z równaniem różniczkowym

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.21)$$

$$x(t_k) = x_k \quad (2.22)$$

zaś estymata zgodnie z

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (2.23)$$

$$\hat{x}(t_k) = \hat{x}_k \quad (2.24)$$

Błąd estymacji ewoluuje więc zgodnie z równaniem

$$\dot{\varepsilon}(t) = f(x(t), u(t)) - f(\hat{x}(t), u(t)) \quad (2.25)$$

$$\varepsilon(t_k) = x_k - \hat{x}_k \quad (2.26)$$

Oznacza to, że błąd estymacji dla  $t \in [t_k, t_{k+1}]$  wynosi

$$\varepsilon(t) = \varepsilon(t_k) + \int_{t_k}^{t_{k+1}} (f(x(t), u(t)) - f(\hat{x}(t), u(t))) dt \quad (2.27)$$

Możemy więc oszacować normę wartości błędu estymacji jako

$$|\varepsilon(t)| \leq |\varepsilon(t_k)| + \int_{t_k}^t |f(x(t), u(t)) - f(\hat{x}(t), u(t))| dt \quad (2.28)$$

Z warunku Lipschitza (B.11) mamy

$$|\varepsilon(t)| \leq |\varepsilon(t_k)| + \int_{t_k}^t L |x(t) - \hat{x}(t)| dt \quad (2.29)$$

$$= |\varepsilon(t_k)| + \int_{t_k}^t L |\varepsilon(t)| dt \quad (2.30)$$

Co pozwala zastosować lemat Gronwalla-Bellmana (zobacz lemat A.1 na stronie 98)

$$|\varepsilon(t)| \leq e^{L(t-t_k)} |\varepsilon(t_k)| \quad (2.31)$$

Wykorzystując monotoniczność funkcji eksponencjalnej oraz przyjęte założenia otrzymujemy oszacowanie:

$$|\varepsilon(t)| \leq e^{L\tau_{max}} |\varepsilon(t_k)| = \alpha |\varepsilon(t_k)| \quad \forall t \in [t_k, t_{k+1}] \quad (2.32)$$

Co daje

$$\bar{\varepsilon}_k = \sup_{t \in [t_k, t_{k+1}]} |\varepsilon(t)| \leq \alpha |\varepsilon(t_k)| \quad (2.33)$$

Dla dolnego oszacowania wystarczy zauważyć, że ponieważ estymata na odcinku  $[t_k, t_{k+1}]$  jest funkcją ciągłą to

$$\bar{\varepsilon}_k = \sup_{t \in [t_k, t_{k+1}]} |\varepsilon(t)| = \max \left( \varepsilon(t_k), \sup_{t \in (t_k, t_{k+1}]} |\varepsilon(t)| \right) \quad (2.34)$$

Wystarczy więc rozpatrzyć dwa przypadki. Jeżeli pierwszy z argumentów funkcji maksimum jest większy lub są one sobie równe, to  $\bar{\varepsilon}_k = \varepsilon(t_k)$ . Jeżeli natomiast drugi argument jest większy, to wtedy  $\bar{\varepsilon}_k > \varepsilon(t_k)$ . Razem warunki te dają dolne oszacowanie

$$|\varepsilon(t_k)| \leq \bar{\varepsilon}_k \quad (2.35)$$



□

Możemy teraz wprowadzić następujące definicje

**Definicja 2.1.** Algorytm 1 będziemy nazywać **stabilnym**, jeżeli dla każdego  $\epsilon > 0$  oraz  $i \in \mathbb{N}$  istnieje  $\delta > 0$  takie że:

$$|\varepsilon(t_i)| < \delta \Rightarrow \bar{\varepsilon}_k < \epsilon \quad \forall k \geq i$$

**Definicja 2.2.** Algorytm 1 będziemy nazywać **globalnie atraktywnym**, jeżeli dla każdego  $\varepsilon(t_0)$

$$\bar{\varepsilon}_k \rightarrow 0 \quad \text{gdy} \quad k \rightarrow \infty$$

**Definicja 2.3.** Algorytm 1 będziemy nazywać **globalnie asymptotycznie stabilnym**, jeżeli będzie on stabilny oraz globalnie atraktywny.

Należy zwrócić uwagę, że stabilność w sensie definicji 2.1 nie powinna być jednoznacznie kojarzona ze stabilnością systemu (2.25). Stabilność w sensie definicji 2.1 należy interpretować w następujący sposób. Niech ewolucja błędu estymacji algorytmu 1 będzie rozpatrywana jako system dynamiczny, na który składają się zarówno równanie różniczkowe rozwiązywane w kroku predykcji jak też i estymację warunków początkowych w kroku korekcji. Stabilność zerowego rozwiązania takiego systemu w sensie Lapunowa, na mocy definicji 2.1, będzie oznaczać stabilność algorytmu. Skupiamy się jedynie na zerowym rozwiązaniu, ponieważ jest to jedyne pożądane rozwiązanie - oznacza ono sytuację, gdy stan systemu i jego estymata są sobie równe. Jeżeli więc zerowe rozwiązanie jest stabilne, to jeżeli estymata raz znajdzie się blisko stanu pozostanie ona już zawsze blisko niego.

Z tego też powodu, bardzo ważną własnością algorytmu 1 jest jego globalna atraktywność. To znaczy, że dla dowolnego błędu początkowego, błąd ten będzie zmierzać do zera z kolejnymi iteracjami. Jak okaże się przy analizie rozdzielenia sterowania i estymacji, algorytm nie musi być stabilny w podanym sensie, aby móc ustabilizować system za pomocą generowanej przez niego estymaty stanu – pod warunkiem, że będzie on globalnie atraktywny. Podana definicja globalnej asymptotycznej stabilności ma jedynie charakter porządkowy i jest połączeniem definicji 2.1 i 2.2.

Wprowadzone zostaną teraz pewne definicje dotyczące procedury korekcji algorytmu 1.

**Definicja 2.4.** Procedurę korekcji algorytmu 1 będziemy nazywać **stabilną**, jeżeli dla każdego  $\epsilon > 0$  oraz  $i \in \mathbb{N}$  istnieje  $\delta > 0$  takie że:

$$|\varepsilon(t_i)| < \delta \Rightarrow |\varepsilon(t_k)| < \epsilon \quad \forall k \geq i$$

**Definicja 2.5.** Procedurę korekcji algorytmu 1 będziemy nazywać **globalnie atraktywną**, jeżeli dla każdego  $\varepsilon(t_0)$

$$\varepsilon(t_k) \rightarrow 0 \quad \text{gdy} \quad k \rightarrow \infty$$

**Definicja 2.6.** Procedurę korekcji algorytmu 1 będziemy nazywać **globalnie asymptotycznie stabilną**, jeżeli będzie ona stabilna oraz globalnie atraktywna.

Wprowadzone definicje dotyczące procedury korekcji, są analogicznymi własnościami całego algorytmu przeniesionymi na własności procedury korekcji. W przypadku definicji stabilności interpretujemy procedurę korekcji jako dyskretny system dynamiczny. Wykorzystując je możemy sformułować następujące twierdzenia.

**Twierdzenie 2.2.** *Algorytm 1 jest globalnie atraktywny wtedy i tylko wtedy, gdy jego procedura korekcji jest globalnie atraktywna.*

*Dowód.* Ponieważ procedura korekcji jest globalnie atraktywna, więc dla  $k \rightarrow \infty$  mamy  $\varepsilon(t_k) \rightarrow 0$ , w związku z tym mamy  $|\varepsilon(t_k)| \rightarrow 0$ . Na mocy lematu 2.1 mamy, że

$$0 \leq \bar{\varepsilon}_k \leq \alpha |\varepsilon(t_k)|$$

co z twierdzenia o trzech ciągach daje tezę twierdzenia.  $\square$

**Twierdzenie 2.3.** *Algorytm 1 jest stabilny wtedy i tylko wtedy, gdy jego procedura korekcji jest stabilna.*

*Dowód.* Na początek dowiedzimy warunku dostatecznego. Załóżmy, że procedura korekcji jest stabilna zgodnie z definicją 2.4, t.j. dla każdego  $\epsilon$  oraz  $i \in \mathbb{N}$  istnieje  $\delta > 0$

$$|\varepsilon(t_i)| < \delta \Rightarrow |\varepsilon(t_k)| < \epsilon \quad \forall k \geq i \quad (2.36)$$

Na mocy lematu 2.1 mamy zaś

$$\varepsilon(t_k) \geq |\bar{\varepsilon}_k|/\alpha \quad (2.37)$$

Z (2.36) i (2.37) wynika następująca implikacja

$$\forall \epsilon > 0, i \in \mathbb{N}, \exists \delta > 0 : |\varepsilon(t_i)| < \delta \Rightarrow \bar{\varepsilon}_k < \epsilon \alpha \quad \forall k \geq i \quad (2.38)$$

co dowodzi warunku dostatecznego. Warunku koniecznego dowiedzimy nie wprost. Załóżmy, że algorytm 1 jest stabilny, t.j. dla każdego  $\epsilon > 0$  oraz  $i \in \mathbb{N}$  istnieje  $\delta > 0$  takie że:

$$|\varepsilon(t_i)| < \delta \Rightarrow \bar{\varepsilon}_k < \epsilon \quad \forall k \geq i$$

oraz, że procedura korekcji jest niestabilna. Z niestabilności procedury korekcji wynika, że dla każdego  $\delta > 0$  istnieją takie,  $i^* \in \mathbb{N}$ ,  $\epsilon^* > 0$ ,  $\mathbb{N} \ni k^* > i^*$  oraz trajektoria błędu (interpretowana jako rozwiązanie systemu dyskretnego)  $\varepsilon^*$ , że

$$|\varepsilon^*(t_{i^*})| < \delta \quad \text{oraz} \quad |\varepsilon^*(t_{k^*})| \geq \epsilon^* \quad (2.39)$$

Rozważmy teraz ciągłą estymatę związaną z trajektorią  $\varepsilon^*$  (mowa tu o estymacie wygenerowanej za pomocą procedury predykcji bazującej na kolejnych wartościach korekcji z błędami wyrażonymi przez  $\varepsilon^*$ ). Analogicznie do (2.19) sformułujemy normę błędu takiej estymaty  $\bar{\varepsilon}_k^*$ . Ponieważ zakładamy, że algorytm 1 jest stabilny, więc w szczególności zachodzi implikacja (dla pewnego  $\delta > 0$ )

$$|\varepsilon^*(t_{i^*})| < \delta \Rightarrow \bar{\varepsilon}_k^* < \epsilon^* \quad \forall k \geq i$$

Z lematu 2.1 wiemy, że dla każdego  $k \in \mathbb{N}$  zachodzi oszacowanie

$$\bar{\varepsilon}_k^* \geq |\varepsilon^*(t_k)|$$

Procedura korekcji jest niestabilna, więc można znaleźć takie  $i^*$  oraz  $k^*$ , że zachodzi (2.39), oznacza to, że nie istnieje takie  $\delta$ , dla którego  $\bar{\varepsilon}_k^*$  pozostanie mniejsze od  $\epsilon^*$  dla wszystkich  $k > i^*$  co oznacza, że jeżeli procedura korekcji jest niestabilna to algorytm 1 nie może być stabilny.  $\square$

**Twierdzenie 2.4.** *Algorytm 1 jest globalnie asymptotycznie stabilny wtedy i tylko wtedy, gdy jego procedura korekcji jest globalnie asymptotycznie stabilna.*

*Dowód.* Jest to konsekwencja twierdzeń 2.2 i 2.3.  $\square$

Wnioskiem z twierdzeń 2.2-2.4 jest to, że do stabilnej, globalnie atraktywnej lub globalnie asymptotycznie stabilnej pracy algorytmu wystarczy zagwarantować stosowne własności procedury korekcji. Krok predykcji nie wpływa na własność stabilności algorytmu.

### 2.3. Obserwowalność

Ponieważ estymacja stanu następuje jedynie w kroku korekcji konkretne wymagania dotyczące obserwowalności systemu z dyskretnym wyjściem będą formułowane w zależności od potrzeb. Na tym etapie będziemy rozpatrywać jedynie pewną ogólną definicję. Na początek zdefiniujemy dwie klasy funkcji

- $\gamma: [0, +\infty) \rightarrow [0, +\infty)$  należy do klasy  $\mathcal{K}$ , jeżeli jest ściśle rosnąca, ciągła i  $\gamma(0) = 0$
- $\gamma \in \mathcal{K}^\infty$  jeżeli  $\gamma \in \mathcal{K}$  oraz  $\gamma(s) \rightarrow \infty$  gdy  $s \rightarrow \infty$

Obserwowalność zdefiniujemy w następujący sposób.

**Definicja 2.7.** System

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.40)$$

$$y(t) = h(x(t)) \quad (2.41)$$

$$t \geq 0 \quad (2.42)$$

będziemy uważać za **obserwowalny**, jeżeli będą istnieć takie:  $N$ , chwile czasu  $t_0, t_1, \dots, t_N$  oraz funkcja  $\varphi \in \mathcal{K}^\infty$ , że

$$|x(t_0) - \hat{x}(t_0)| \leq \varphi \left( \sum_{i=0}^N |y(t_i) - \hat{y}(t_i)|^2 \right) \quad (2.43)$$

gdzie

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (2.44)$$

$$\hat{y}(t) = h(\hat{x}(t)) \quad (2.45)$$

dla dowolnych  $x(t_0)$  i  $\hat{x}(t_0)$ .

Ciągłą wersję takiego sformułowania obserwowalności można znaleźć w pracy (Alamir, 1999). Interpretacja tej definicji jest następująca. System uważamy za obserwowalny, jeżeli mając dwa rozwiązania równania (2.40) o różnych warunkach początkowych, normę różnicy warunków początkowych  $|x(t_0) - \hat{x}(t_0)|$  (odległość między punktami początkowymi) można oszacować porównując odpowiednie pomiary wyjścia systemu (2.41) wygenerowane przez te warunki początkowe. Oznacza to, że im bliższe sobie będą wyjścia dla różnych warunków początkowych, tym mniejsza jest różnica między warunkami początkowymi.

Dla układów liniowych definicja 2.7 jest równoważna definicji klasycznej (zob. np. Mitkowski, 1991, str. 86). Dowodzi tego poniższe

**Twierdzenie 2.5.** Dany jest system

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \quad (2.46)$$

gdzie  $A \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{m \times n}$ . Następujące pojęcia są równoważne:

1. Para  $(C; A)$  jest obserwowalna;
2. System (2.46) jest obserwowalny w sensie definicji 2.7.

*Dowód.* Najpierw wykazemy, że z obserwowalności w sensie definicji 2.7 wynika obserwowalność pary macierzy  $(C; A)$ . Załóżmy, że para  $(C; A)$  nie jest obserwowalna, ale spełniona jest zależność (2.43). Z twierdzenia A.2 (str. 98) wynika, że ze względu na brak obserwowalności  $(C; A)$  istnieje taki niezerowy wektor  $\xi \in \mathbb{R}^n$ , że

$$Ce^{At}\xi = 0 \quad \text{dla } t \geq 0 \quad (2.47)$$

weźmy więc takie  $x(t_0)$  i  $\hat{x}(t_0)$ , że  $x(t_0) - \hat{x}(t_0) = \xi$ . Wtedy

$$\varphi \left( \sum_{i=0}^N |y(t_i) - \hat{y}(t_i)|^2 \right) = \varphi \left( \sum_{i=0}^N |Ce^{At_i}x(t_0) - Ce^{At_i}\hat{x}(t_0)|^2 \right) = \varphi \left( \sum_{i=0}^N |Ce^{At_i}\xi|^2 \right) = \varphi(0) = 0 \quad (2.48)$$

niezależnie od  $N$  oraz  $t_0, t_1, \dots, t_N$ . Jednocześnie mamy

$$|x(t_0) - \hat{x}(t_0)| = |\xi| > 0 \quad (2.49)$$

Z założenia o spełnieniu warunku (2.43) wynika więc

$$0 < |\xi| \leq 0 \quad (2.50)$$

czyli sprzeczność. Oznacza to, że z obserwowalności w sensie definicji 2.7 wynika obserwowalność pary macierzy  $(C; A)$ .

Pokażemy teraz, że zachodzi zależność odwrotna, t.j. z obserwowalności pary macierzy  $(C; A)$  wynika obserwowalność w sensie definicji 2.7. Przyjmijmy, że  $x(t_0) - \hat{x}(t_0) = \varepsilon_0$ . Mamy, więc

$$\sum_{i=0}^N |y(t_i) - \hat{y}(t_i)|^2 = \sum_{i=0}^N |Ce^{At_i}x(t_0) - Ce^{At_i}\hat{x}(t_0)|^2 = \sum_{i=0}^N |Ce^{At_i}\varepsilon_0|^2 = \sum_{i=0}^N \varepsilon_0^T e^{A^T t_i} C^T C e^{At_i} \varepsilon_0 \quad (2.51)$$

Przyjmijmy, że  $N = n - 1$  oraz  $t_i = t_0 + i\tau$ , przy czym  $\tau \operatorname{Im}(\lambda_i - \lambda_j) \neq 2\pi$  dla  $\operatorname{Re}(\lambda_i - \lambda_j) = 0$ , gdzie  $\lambda_i, \lambda_j \in \lambda(A)$ ,  $i \neq j$  zaś  $= \pm 1, \pm 2, \dots$ . Bez straty ogólności możemy przyjąć  $t_0 = 0$ . Oznaczmy  $A_d = e^{A\tau}$ . Na mocy twierdzenia A.3 (str. 99), ponieważ para  $(C; A)$  jest obserwowalna to para  $(C; A_d)$  również jest obserwowalna. Możemy teraz zapisać równość (2.51) jako

$$\sum_{i=0}^N \varepsilon_0^T (A_d^i)^T C^T C A_d^i \varepsilon_0 = \varepsilon_0^T \begin{bmatrix} C^T & A_d^T C^T & \dots & (A_d^{n-1})^T C^T \end{bmatrix} \begin{bmatrix} C \\ CA_d \\ \vdots \\ CA_d^{n-1} \end{bmatrix} \varepsilon_0 = \varepsilon_0^T S_d^T S_d \varepsilon_0 \quad (2.52)$$

Oznaczmy wiersze macierzy  $C$  jako  $C_j$ . Macierz  $S_d$  ze względu na obserwowalność  $(C; A_d)$  ma rząd równy  $n$ , w związku z tym, za (Kaczorek, 1977, str. 43) ma ona  $n$  liniowo niezależnych wierszy postaci

$$C_{j_1} A_d^{i_1}, C_{j_2} A_d^{i_2}, \dots, C_{j_n} A_d^{i_n}$$

gdzie  $j_k \in \{1, 2, \dots, m\}$ ,  $i_k \in \{0, 1, 2, \dots, n-1\}$ ,  $k = 1, 2, \dots, n$ . Utwórzmy z tych wierszy macierz

$$G_d = \begin{bmatrix} C_{j_1} A_d^{i_1} \\ C_{j_2} A_d^{i_2} \\ \vdots \\ C_{j_n} A_d^{i_n} \end{bmatrix} \quad (2.53)$$

przy czym,  $\det(G_d) \neq 0$ . Łatwo zauważyć, że

$$\varepsilon_0^T S_d^T S_d \varepsilon_0 \geq \varepsilon_0^T G_d^T G_d \varepsilon_0 \quad (2.54)$$

Macierz  $G_d^T G_d$  jest macierzą ściśle dodatnio określoną (dla ułatwienia czytania czytelnikowi zob. lemat A.4 str. 99), więc zachodzi następujące oszacowanie

$$\varepsilon_0^T G_d^T G_d \varepsilon_0 \geq \lambda_{\min}(G_d^T G_d) |\varepsilon_0|^2 \quad (2.55)$$

Przy czym  $\lambda_{\min}(G_d^T G_d)$  oznacza najmniejszą wartość własną macierzy  $(G_d^T G_d)$ . Dowód kończy wybranie takiej funkcji  $\varphi \in \mathcal{K}^\infty$  aby spełniony był warunek (2.43). Taką funkcją jest np.

$$\varphi(a) = \sqrt{\frac{a}{\lambda_{\min}(G_d^T G_d)}} \quad (2.56)$$

W ten sposób pokazaliśmy, że jeżeli para  $(C; A)$  jest obserwowalna, to system jest obserwowalny w sensie

definicji 2.7. □

## 2.4. Zasada rozdzielności

Bardzo ważnym problemem, przy stosowaniu wszelkiego rodzaju estymacji stanu, jest problem rozdzielności sterowania i estymacji (filtracji). Problem ten polega na tym, czy niezależnie zaprojektowane stabilizujące sprzężenie zwrotne od stanu oraz asymptotycznie stabilny estymator, razem stworzą asymptotycznie stabilny system. W układach liniowych ciągłych i dyskretnych zachodzi tzw. zasada rozdzielności (ang. *separation principle*), która pozwala na wspomniane podejście do projektowania. W układach nieliniowych sprawa jest bardziej skomplikowana. Rozważmy następujący przykład

**Przykład 2.1.** Dany jest system dynamiczny postaci

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_2 + x_2^3 + u\end{aligned}$$

System ten posiada trzy punkty równowagi, z czego dwa niestabilne. Globalnie stabilizujące sprzężenie zwrotne od stanu może mieć postać:

$$u = -x_2^3 - 1/4x_1$$

System zamknięty będzie miał wtedy postać

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1/4 & -1 \end{bmatrix} x$$

czyli będzie to system liniowy o podwójnej wartości własnej  $\lambda_{1,2} = -1/2$ . Załóżmy, że pomiar drugiej zmiennej stanu nie jest dostępny, i zamiast niej wykorzystana jest jej estymata, t.j.

$$u = -\hat{x}_2^3 - 1/4x_1$$

Założmy dodatkowo, że błąd estymacji zanika wykładniczo i ma postać:

$$\varepsilon_2(t) = e^{-t}$$

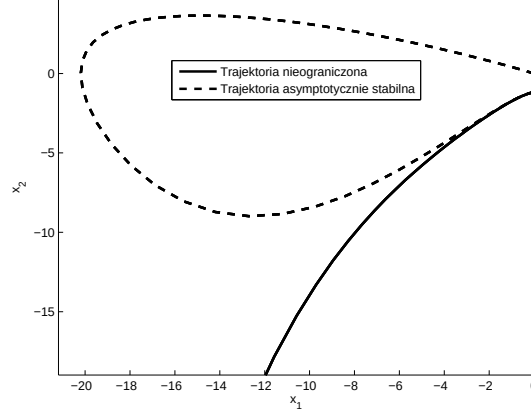
czyli zanika, ze współczynnikiem tłumienia 2 razy większym niż docelowy system. Okazuje się, że np. dla warunku początkowego  $x(0) = [0 \ -1.162]^\top$  system posiada trajektorię nieograniczoną. Dla porównania na rysunku 2.2 pokazane są trajektorie asymptotycznie stabilne ( $x(0) = [0 \ -1.16]^\top$ ), oraz właśnie trajektorie nieograniczone.

Przykład ten pokazuje, że w ogólności zasada rozdzielności nie zachodzi w systemach nieliniowych. Dla szczególnych klas układów problem ten rozważany był w licznych pracach (zob. np. Arcak, 2002; Gauthier i Kupka, 1992; Golubev, Krishchenko i Tkachev, 2002, 2001; Atassi i Khalil, 2000; Lin, 1994; Loria i Panteley, 1999). Zasadniczo, w literaturze, rozważa się rozdzielność regulatora i obserwatora asymptotycznego. W przypadku problemu odtwarzania ciągłego stanu z dyskretnych danych pomiarowych problem spełnienia zasady rozdzielności wymaga odrębnego potraktowania. Pewne wyniki, przy założeniach dotyczących ograniczoności błędów estymacji, można znaleźć w (Bania, 2008), dotyczą one jednak regulacji predykcyjnej.

### 2.4.1. Zastosowanie stabilności wejście-stan

Szczególnie interesujące wydają się być prace (np. Tsiniias, 1993) związane z pojęciem stabilności “wejście-stan” (ang. *input-to-state stability*, *ISS*, zobacz dodatek B). Poniżej przedstawione zostaną przedstawione pewne zastosowania własności ISS do analizy rozdzielności w kontekście odtwarzania ciągłego stanu z dyskretnych danych pomiarowych, dla pewnej klasy układów nieliniowych.

Na początek wprowadzimy formalną definicję rozdzielności:



Rysunek 2.2.: Przykład na nieprawdziwość zasady rozdzielności w systemach nieliniowych

**Definicja 2.8.** O układzie sterowania

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.57)$$

$$u(t) = g(x(t)) \quad (2.58)$$

gdzie  $f(0, 0) = 0$  i  $g(0) = 0$ , mówimy, że posiada własność rozdzielności gdy zachodzi następująca implikacja. Jeżeli system zamknięty

$$\dot{x}(t) = f(x(t), g(x(t))) \quad (2.59)$$

jest globalnie asymptotycznie stabilny w sensie Lapunowa (zob. Mitkowski, 1991, str. 32) oraz dysponujemy estymatą stanu  $\hat{x}$  taką, że  $\hat{x}(t) \rightarrow x(t)$  przy  $t \rightarrow \infty$ , to dla trajektorii systemu

$$\dot{x}(t) = f(x(t), g(\hat{x}(t))) \quad (2.60)$$

zachodzi

$$|x(t)| \rightarrow 0 \quad \text{gdy} \quad t \rightarrow \infty, \quad \forall x(0) \quad (2.61)$$

Z definicji tej można wyciągnąć następujący wniosek – korzystając z definicji błędu estymacji: (2.18) możemy zależność (2.60) przeformułować jako

$$\dot{x}(t) = f(x(t), g(x(t) + \varepsilon(t))) \quad (2.62)$$

przy założeniu

$$|\varepsilon(t)| \rightarrow 0 \quad \text{gdy} \quad t \rightarrow \infty \quad (2.63)$$

W takiej sytuacji żądanie spełnienia (2.61) (czyli posiadania własności rozdzielności) jest równoważne wymaganiu posiadania przez system własności “zbieżne wejście-zbieżny stan” (ang. *Convergent Input-Convergent State*, *CICS*) przy założeniu, że błąd estymacji jest niezależnym wejściem. Zachodzi następujący

**Lemat 2.6.** Jeżeli system dynamiczny

$$\dot{x}(t) = f(x(t), v(t)) \quad (2.64)$$

posiada własność ISS (zob. B.2) to zachodzi następująca implikacja

$$|v(t)| \rightarrow 0 \quad \text{gdy} \quad t \rightarrow \infty \Rightarrow |x(t)| \rightarrow 0 \quad \text{gdy} \quad t \rightarrow \infty \quad (2.65)$$

Oznacza to, że własność ISS jest silniejsza od CICS (zob. Sontag, 1989).

*Dowód.* Lemat wynika z własności (B.7) w twierdzeniu B.1. □

Ponieważ przy globalnie atraktywnej procedurze korekcji algorytm 1 generuje błędy estymacji zanikające do zera (patrz poprzednia sekcja), więc tak rozumiana własność rozdzielności gwarantuje też rozdzielność w kontekście odtwarzania ciągłego stanu z dyskretnych danych pomiarowych. Poniżej kilka wyników pozwalających na określenie rozdzielności rozważanych systemów. Na początek, bez dowodu, pewien oczywisty wynik o charakterze jakościowym.

**Twierdzenie 2.7.** *Jeżeli system*

$$\dot{x} = F(x, u) = f(x, g(x + u))$$

*posiada własność ISS, to system zamknięty (2.60) posiada własność rozdzielności.*

Powyższe twierdzenie ma mało konstruktywny charakter, zwłaszcza, że bezpośrednia analiza posiadania przez system własności ISS jest trudna. Poniżej podamy dwa twierdzenia, które przy prostszych do sprawdzenia założeniach, pozwalają na zagwarantowanie rozdzielności.

**Twierdzenie 2.8.** *Dany jest system*

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.66)$$

$$u(t) = g(x(t)) \quad (2.67)$$

*gdzie  $f(0, 0) = 0$  i  $g(0) = 0$  oraz*

$$|f(x, u_1) - f(x, u_2)| \leq L_1 |u_1 - u_2| \quad \forall u_1, u_2 \in \mathbb{R}^r \quad \forall x \in \mathbb{R}^n \quad (2.68)$$

$$|g(x_1) - g(x_2)| \leq L_2 |x_1 - x_2| \quad \forall x_1, x_2 \in \mathbb{R}^n \quad (2.69)$$

*przy czym zerowy punkt równowagi systemu*

$$\dot{x}(t) = f(x(t), g(x(t))) \quad (2.70)$$

*jest globalnie wykładniczo stabilny. System ten posiada własność rozdzielności.*

*Dowód.* Oznaczmy prawą stronę systemu (2.70) w sytuacji gdy nie ma błędu estymacji jako  $F(x, 0)$  zaś gdy błąd ten występuje przez

$$\dot{x}(t) = f(x(t), g(x(t) + \varepsilon(t))) = F(x(t), \varepsilon(t)) \quad (2.71)$$

Łatwo zauważyć, że  $F(v, w)$  spełnia globalnie warunek Lipschitza ze względu na drugi argument ze stałą  $L_u = L_1 L_2$ . Z globalnej wykładniczej stabilności systemu (2.70) mamy, że istnieje dla niego funkcjonal Lapunowa  $V(x)$ , a co więcej

$$\dot{V}(x) = \nabla V(x) \leq -\gamma |x|^2$$

oraz

$$|\nabla V(x)| \leq \eta |x|$$

gdzie  $\gamma$  i  $\eta$  są pewnymi dodatnimi stałymi (zob. Sastry i Bodson, 1994, str. 28). Zastosujmy ten sam funkcjonal Lapunowa do analizy sytemu (2.71). Otrzymujemy

$$\dot{V}(x) = \nabla V(x) F(x, \varepsilon) = \nabla V(x) f(x, 0) + \nabla V(x) (F(x, \varepsilon) - F(x, 0)) \leq -\gamma |x|^2 + \eta |x| L_u |\varepsilon| \quad (2.72)$$

Łatwo zauważyć, że jest to funkcjonal ISS-Lapunowa (patrz definicja B.2) z funkcjami

$$\chi(|u|) = \frac{\eta L_u}{\delta} |u| \quad (2.73)$$

$$\alpha_3(|x|) = (\gamma - \delta) |x|^2 \quad (2.74)$$

przy czym  $0 < \delta < \gamma$ . Czyli zgodnie z twierdzeniem B.1 posiada on własność ISS, a zgodnie z twierdzeniem 2.7 własność rozdzielności.  $\square$

Powyższe twierdzenie zakłada istnienie wykładniczo stabilizującego regulatora dla systemu (2.1). Jak wiadomo, znalezienie takiego regulatora nie jest zadaniem trywialnym. Przy dodatkowym zawężeniu klasy systemu można podać dodatkowo postać regulatora

**Twierdzenie 2.9.** *Dany jest system dynamiczny*

$$\dot{x}(t) = Ax(t) + Bu(t) + f(x(t)) \quad (2.75)$$

przy czym  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^r$ ,  $B \in \mathbb{R}^{n \times r}$ ,  $A \in \mathbb{R}^{n \times n}$ . Dodatkowo w stosunku do operatora  $f$  zakłada się, że  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $f(0) = 0$ ,  $\|f(x) - f(y)\| \leq L\|x - y\|$ ,  $\forall x, y \in \mathbb{R}^n$ . Istnieją także  $K \in \mathbb{R}^{r \times n}$  i ściśle dodatnio określona macierz  $Q = Q^\top \in \mathbb{R}^{n \times n}$ , dla których  $\operatorname{Re} \lambda(A + BK) < 0$  oraz spełniona jest nierówność

$$\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(H)} > L \quad (2.76)$$

gdzie  $H$  jest ściśle dodatnio określonym rozwiązaniem równania Lapunowa

$$(A + BK)^\top H + H(A + BK) = -Q \quad (2.77)$$

Sterowanie ma postać.

$$u(t) = Kx(t) \quad (2.78)$$

Taki system posiada własność rozdzielności.

*Dowód.* Przyjmijmy, że w regulatorze (2.78) została zastosowana estymata  $\hat{x} = x - \varepsilon$ , przy czym  $\varepsilon(t) \rightarrow 0$  gdy  $t \rightarrow \infty$ . Uzyskujemy w ten sposób system zamknięty

$$\dot{x}(t) = Ax(t) + BK\hat{x}(t) + f(x(t)) = (A + BK)x(t) - BK\varepsilon(t) + f(x(t))$$

Wykładniczą stabilność systemu zamkniętego wykazano w dowodzie twierdzenia A.6 (str. 99) za pomocą funkcjonału Lapunowa, jakim była forma kwadratowa. Korzystając z własności funkcjonału dla systemów wykładniczo stabilnych (zob. Sastry i Bodson, 1994, str. 28), wiemy że pochodna funkcjonału dla systemu bez błędu estymacji spełnia oszacowanie

$$\dot{V}(x) \leq -\gamma|x|^2$$

gdzie,  $\gamma = \lambda_{\min}(Q) - 2\lambda_{\max}(H)L > 0$ . Dla systemu z błędem estymacji ( $u = K(x + \varepsilon)$ ) zastosujemy ten sam funkcjonał Lapunowa, jego pochodna przyjmie więc postać:

$$\dot{V}(x) \leq -\gamma|x|^2 + 2x^\top H B K \varepsilon(t) \leq -\gamma|x|^2 + 2|x|\|H B K\|\|\varepsilon\| \quad (2.79)$$

Można zwrócić uwagę, że oszacowanie (2.79) również jest funkcjonałem ISS-Lapunowa (definicja B.2), przyjmując następujące funkcje

$$\chi(|u|) = \frac{2\|H B K\|}{\delta}|u| \quad (2.80)$$

$$\alpha_3(|x|) = (\gamma - \delta)|x|^2 \quad (2.81)$$

przy czym  $0 < \delta < \gamma$ . Stąd wynika, że system (2.75) z regulatorem (2.78) wyznaczonym zgodnie z A.6 posiada własność rozdzielności.  $\square$

## 2.5. Okna pomiarowe

W pracy rozważamy pomiary dyskretnie w czasie. Dlatego też istotnym jest, aby omówić pojęcie okna pomiarowego.

**Definicja 2.9.** Przez pojęcie okna pomiarowego będziemy rozumieć ciąg pomiarów wyjścia

$$y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m-1}} \quad (2.82)$$



przy czym  $m > 0$  będziemy nazywać **szerokością okna**. Konkretnie okno pomiarowe, o danej szerokości, będziemy identyfikować przez pierwszą i ostatnią chwilę pomiaru np. będziemy mówić o oknie od  $t_{k-3}$  do  $t_{k-4+m}$ .

Okno pomiarowe jest zmieniane wraz z pojawianiem się kolejnych pomiarów (kolejnymi iteracjami algorytmu). Naturalnym jest rozróżnienie dwóch rodzajów zmian okien pomiarowych – **przesuwanego** oraz **rozszerzanego**.

**Definicja 2.10.** Przez okno przesuwane, będziemy rozumieć okno, dla którego z każdą kolejną iteracją  $k$  jest zwiększane o 1, tzn:

$$\begin{aligned} & y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m-1}} \\ & y_{t_{k+1}}, y_{t_{k+2}}, \dots, y_{t_{k+m}} \\ & y_{t_{k+2}}, y_{t_{k+3}}, \dots, y_{t_{k+1+m}} \\ & \vdots \end{aligned}$$

Innymi słowy jest to okno, które ma stałą szerokość natomiast przesuwają się jego początek (rozpatrujemy zawsze  $m$  pomiarów).

**Definicja 2.11.** Przez okno rozszerzane, będziemy rozumieć okno, dla którego z każdą kolejną iteracją  $m$  jest zwiększane o 1, tzn:

$$\begin{aligned} & y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m-1}} \\ & y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m}} \\ & y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m+1}} \\ & \vdots \end{aligned}$$

Innymi słowy, jest to okno, które ma ustalony początek, natomiast zwiększa się jego szerokość (do okna z każdą iteracją dodajemy kolejny pomiar).

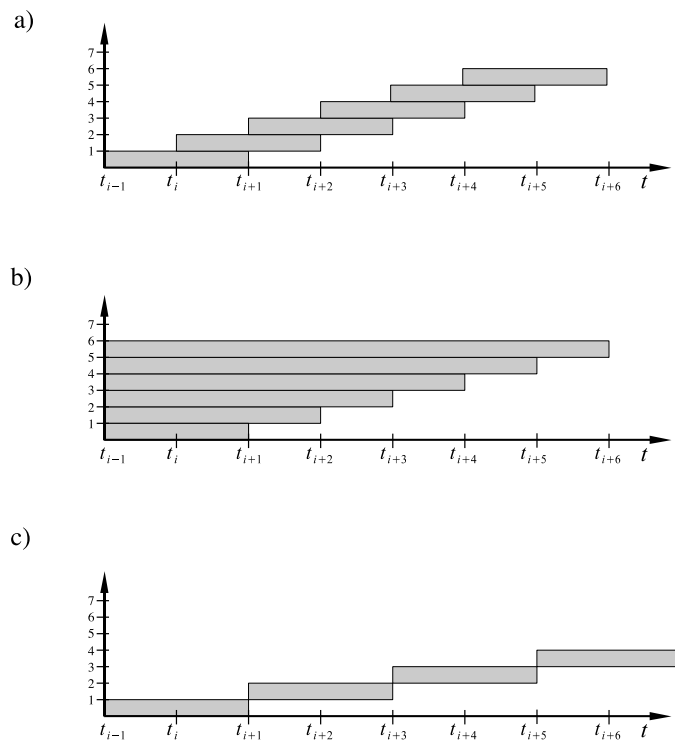
W przypadku obserwacji  $C^0$  optymalnej (zob. str. 47) będziemy rozważać jeszcze jeden rodzaj okna pomiarowego – **okno rozłączne**.

**Definicja 2.12.** Przez okno rozłączne będziemy rozumieć okno, dla którego z każdą kolejną iteracją  $k$  zwiększamy o  $m - 1$ , tzn:

$$\begin{aligned} & y_{t_k}, y_{t_{k+1}}, \dots, y_{t_{k+m-1}} \\ & y_{t_{k+m-1}}, y_{t_{k+m}}, \dots, y_{t_{k+2m-2}} \\ & y_{t_{k+2m-1}}, y_{t_{k+2m}}, \dots, y_{t_{k+3+m}} \\ & \vdots \end{aligned}$$

Innymi słowy, jest to okno o stałej szerokości, dla którego pomiędzy poszczególnymi iteracjami wspólny jest tylko jeden punkt.

Graficznie okna pomiarowe ilustruje rysunek 2.3. Na osi pionowej oznaczony jest numer iteracji, na osi poziomej są chwile pomiarów.



Rysunek 2.3.: Okna pomiarowe: a) okno przesuwane o szerokości  $m = 3$ , b) okno rozszerzane, c) okno rozłączne o szerokości  $m = 3$

## Rozdział 3.

### Predykcja

W ramach niniejszego rozdziału omówione zostaną pewne własności szczególnie związane z krokiem predykcji algorytmu 1. Rozpatrywane będą problemy generacji ciągłych estymat, powstawania błędów przejściowych predykcji oraz ograniczoności błędu predykcji wywołanego zakłóceniami wejściowymi.

#### 3.1. Generacja ciągłych estymat

Jak napisano wcześniej, w kroku predykcji konieczne jest rozwiązanie równania różniczkowego

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (3.1)$$

$$t \in [t_i, t_{i+1}] \quad (3.2)$$

$$i = 0, 1, 2, \dots \quad (3.3)$$

Jeżeli równanie to jest równaniem liniowym, t.j.  $f(x, u) = Ax + Bu$  to ciągle rozwiązanie dla danych  $t$  możemy otrzymać analitycznie ze wzoru

$$\hat{x}(t) = e^{tA}\hat{x}(t_i) + \int_{t_i}^t e^{(t-\tau)A}Bu(\tau)d\tau \quad (3.4)$$

Jeżeli jednak równanie jest nieliniowe (czyli ogólnie w postaci (3.1)), rozwiązania analityczne jest zazwyczaj niedostępne. Co więcej, znajomość ogólnej formuły analitycznej nie przydaje się w efektywnym wyznaczaniu rozwiązania. Na przykład całka spłotowa występująca w równości (3.4), dla pewnych funkcji sterujących  $u$  może okazać się bardzo trudna do analitycznego rozwiązania. Nie dziwi więc konieczność stosowania numerycznych metod umożliwiających rozwiązywanie równań różniczkowych.

Można uznać, że numeryczne metody rozwiązywania równań różniczkowych wywodzą się od Eulera (1768 praca „Institutiones Calculi Integralis”), gdzie zaproponowana została metoda znana obecnie jako metoda Eulera „w przód” (inaczej otwarta lub bezpośrednia, nazywana także metodą prostokątów). Dla ustalenia uwagi, dla równania  $\dot{x}(t) = f(t, x(t))$ , wartość rozwiązania w chwili odległej o  $h$  od chwili początkowej wyraża się wzorem

$$x(t_0 + h) \approx x(t_0) + hf(t_0, x(t_0))$$

Można wykazać, że metoda ta jest metodą pierwszego rzędu, t.j.

$$x(t_0 + h) = x(t_0) + hf(t_0, x(t_0)) + \mathcal{O}(h^2)$$

Metoda ta posłużyła do stworzenia bardziej złożonych metod o wyższych rzędach, które po pracach Rungego (1885) i Kutty (1901) znane są pod ogólną nazwą metod Rungego-Kutty. Przyjmując  $x_0$  jako warunek początkowy i  $x_1$  jako przybliżenie rozwiązania równania w chwili odległej o  $h$  od początkowej, możemy dla

równania  $\dot{x}(t) = f(t, x(t))$  podać ogólny schemat tych metod

$$\begin{aligned}
 k_1 &= f(t_0, x_0) \\
 k_2 &= f(t_0 + c_2 h, x_0 + h a_{21} k_1) \\
 k_3 &= f(t_0 + c_3 h, x_0 + h(a_{31} k_1 + a_{32} k_2)) \\
 &\vdots \\
 k_s &= f(t_0 + c_s h, x_0 + h(a_{s1} k_1 + \dots + a_{s,s-1} k_{s-1})) \\
 x_1 &= x_0 + h(b_1 k_1 + \dots + b_s k_s)
 \end{aligned} \tag{3.5}$$

Schemat ten opisuje  $s$ -etapową, bezpośrednią metodę Rungego-Kutty. Mianem etapów określamy  $k_i$ , które odpowiadają wyliczeniu prawej strony równania różniczkowego, w pewnym konkretnym punkcie. Innymi słowy, metoda  $s$ -etapowa potrzebuje  $s$ -krotnego obliczenia wartości prawej strony równania różniczkowego. Współczynniki  $a_{ij}$ ,  $b_i$  oraz  $c_i$  wygodnie jest zapisać w postaci wektorowo-macierzowej, często wyrażanej w postaci tabelki:

$$\begin{array}{c|c}
 c & \mathcal{A} \\
 \hline
 x_1 & b
 \end{array} = \begin{array}{c|ccc}
 0 & & & \\
 c_2 & a_{21} & & \\
 c_3 & a_{31} & a_{32} & \\
 \vdots & \vdots & \vdots & \ddots \\
 c_s & a_{s1} & a_{s1} & \dots & a_{s1} \\
 \hline
 x_1 & b_1 & b_2 & \dots & b_{s-1} & b_s
 \end{array} \tag{3.6}$$

Jak można zauważyć,  $\mathcal{A}$  jest macierzą trójkątną dolną z zerową diagonalą. Możliwe są metody, które mają niezerową przekątną, lub dla których  $\mathcal{A}$  nie jest macierzą trójkątną, wtedy jednak mamy do czynienia z metodami odwrotnymi, lub też uwikłanymi, które do obliczenia każdego etapu wymagają rozwiązania układu równań algebraicznych, w ogólności nieliniowych.

Metody Rungego-Kutty, stanowią obecnie trzon metod służących do rozwiązywania równań różniczkowych, zwłaszcza od czasu, gdy okazały się być bardziej wydajne obliczeniowo w nowoczesnych systemach komputerowych od tzw. metod wielokrokowych. Należy jednak podkreślić, że metody te, oferują przybliżone wartości rozwiązania, z dokładnością do pewnego rzędu jedynie w konkretnych chwilach czasu. W 1983 roku Horn zaproponował metody o tzw. „gęstym” wyjściu. Metody te pozwalają na otrzymanie ekonomicznej obliczeniowo aproksymacji rozwiązania w  $x(t_0 + \theta h)$ , dla  $0 \leq \theta \leq 1$ .

Konstrukcję takich aproksymacji można przedstawić w następujący sposób. Rozważmy  $s$ -etapową metodę Rungego-Kutty, zadaną przez pewne macierze  $\mathcal{A}$ ,  $b$  i  $c$ , o rzędzie równym  $p$ . Metodę tą można wyposażać w dodatkowe  $(s^* - s)$  etapów i zaproponować aproksymację rozwiązania postaci

$$u(\theta) = x_0 + h \sum_{i=1}^{s^*} b_i(\theta) k_i \tag{3.7}$$

gdzie  $k_i$  zadane są według (3.5), zaś  $b_i(\theta)$  są wielomianami dobranymi tak, że

$$u(\theta) - x(t_0 + \theta h) = \mathcal{O}(h^{p^*+1}) \quad \theta \in [0, 1] \tag{3.8}$$

$p^*$  oznacza w tym przypadku rząd aproksymacji (gęstego wyjścia). Zazwyczaj  $s^* \geq s + 1$  zaś  $p^* \leq p$ .

**Przykład 3.1.** Rozważmy, klasyczną i bardzo popularną znaną metodę Rungego-Kutty czwartego rzędu

$$\begin{array}{c|ccc}
 0 & & & \\
 1/2 & 1/2 & & \\
 1/2 & 0 & 1/2 & \\
 1 & 0 & 0 & 1 \\
 \hline
 x_1 & 1/6 & 2/6 & 2/6 & 1/6
 \end{array} \tag{3.9}$$

Dla tej metody, przyjmując  $s^* = s = 4$  oraz  $p^* = 3$  (sama metoda ma rząd  $p = 4$ ), można otrzymać gęste wyjście za pomocą wielomianów

$$b_1(\theta) = \theta - \frac{3\theta^2}{2} + \frac{2\theta^3}{3} \quad (3.10)$$

$$b_2(\theta) = b_3(\theta) = \theta^2 - \frac{2\theta^3}{3} \quad (3.11)$$

$$b_4(\theta) = -\frac{\theta^2}{2} + \frac{2\theta^3}{3} \quad (3.12)$$

Warto zauważyć, że dla  $\theta = 0$  wielomiany te zerują się (otrzymujemy więc warunek początkowy), zaś dla  $\theta = 1$  przechodzą we współczynniki macierzy  $b$ . Należy dodatkowo zaznaczyć, że takie gęste wyjście jest globalnie klasy  $C^0$ , ponieważ funkcja powstała ze sklejania aproksymacji w kolejnych krokach metody Rungego-Kutty nie jest różniczkowalna w punktach styku.

Jednym z typowych, i za razem prostych sposobów wyznaczania ciągłych i różniczkowalnych aproksymacji jest interpolacja Hermite’a. Należy jednak zauważyć, że aby wyznaczyć wielomian interpolujący potrzebna jest znajomość pochodnej rozwiązania w obydwu końcach przedziału pomiędzy krokami. W tym celu należy wywołać prawą stronę równania różniczkowego w punkcie  $x_1$ . Jest to równoważne zwiększeniu liczby etapów o 1 ( $s^* = s+1$ ), jednakże można wykorzystać to wywołanie prawej strony jako pierwszy etap następnego kroku, czyli koszt numeryczny pozostaje niewielki. Wielomian interpolujący, służący za aproksymację rozwiązania, ma postać

$$u(\theta) = (1 - \theta)x_0 + \theta x_1 + \theta(1 - \theta) \left( (1 - 2\theta) \cdot (x_1 - x_0) + (\theta - 1)hf_0 + \theta hf_1 \right) \quad (3.13)$$

przy czym  $f_0 = f(t_0, x_0)$  i  $f_1 = f(t_0 + h, x_1)$ . Takie gęste wyjście jest aproksymacją trzeciego rzędu, która posiada ciągłą pochodną także w punktach poszczególnych kroków.

### 3.1.1. Metoda RK3/8

Powszechnie wiadomo, że stosując metody typu Rungego-Kutty, dużą rolę odgrywa dobrze dobrany krok całkowania. Właściwa długość kroku zależy przede wszystkim od dynamiki rozwiązywanego równania różniczkowego. Z tego powodu, atrakcyjne są metody, które są w stanie krok dobrać samodzielnie w oparciu o dostępne dla nich wyliczenia prawej strony równania. Okazuje się, że odpowiednio skonstruowana metoda *zmiennokrokowa*, pozwala na znaczne poprawienie dokładności obliczeń przy jednocześnie zmniejszonym koszcie ich wykonywania. Bierze się to stąd, że w tych odcinkach osi czasu, w których dynamika zmienia się gwałtownie wykonywane są małe kroki, zaś w okresach kiedy dynamika jest „spokojniejsza” można sobie pozwolić na kroki dłuższe, bez utraty dokładności obliczeń.

Na potrzeby tej pracy opracowano metodę numeryczną typu zmiennokrokowego posiadającą gęste wyjście. Jako bazę tej metody zastosowano jedną z metod opracowaną przez Kuttę na początku XX wieku - tak zwaną metodą trzech ósmych (3/8). Metoda ta jest mniej popularna niż „Metoda Rungego-Kutty” (3.9), ale jest nieco bardziej dokładna (zob. Hairer, Nørsett i Wanner, 2000, str. 138). Tabela (3.6) dla tej metody ma postać

|       |      |     |     |     |
|-------|------|-----|-----|-----|
| 0     |      |     |     |     |
| 1/3   | 1/3  |     |     |     |
| 2/3   | -1/3 | 1   |     |     |
| 1     | 1    | -1  | 1   |     |
| $x_1$ | 1/8  | 3/8 | 3/8 | 1/8 |

(3.14)

Jest to 4 etapowa metoda czwartego rzędu. Aby metoda ta pozwalała na automatyczny dobór kroku dyskretyzacji, należy uzupełnić ją o tzw. *metodę wbudowaną* (ang. *embedded*) (zob. Butcher, 2003, str. 188).

### Rozdział 3. Predykcja

Konstrukcja metod wbudowanych polega na znalezieniu pewnego zestawu współczynników  $\tilde{b}_i$  zastępującego współczynniki  $b_i$  w ostatniej równości (3.5), tak aby uzyskać aproksymację rozwiązania  $\tilde{x}_1$ , która by miała rząd dokładności większy lub mniejszy o jeden od metody podstawowej. W ten sposób można tanim kosztem numerycznym otrzymać dodatkowe rozwiązanie, które można wykorzystać jako estymator błędu. Zwyczajowo metody wbudowane umieszcza się wraz z podstawową w tabeli jak niżej

$$\begin{array}{c|cccc}
 0 & & & & \\
 c_2 & a_{21} & & & \\
 c_3 & a_{31} & a_{32} & & \\
 \vdots & \vdots & \vdots & \ddots & \\
 c_s & a_{s1} & a_{s1} & \dots & a_{s1} \\
 \hline
 x_1 & b_1 & b_2 & \dots & b_{s-1} & b_s \\
 \hline
 \tilde{x}_1 & \tilde{b}_1 & \tilde{b}_2 & \dots & \tilde{b}_{s-1} & \tilde{b}_s
 \end{array} \tag{3.15}$$

Zostało dowiedzione (zob. Hairer et al., 2000, str. 173), że nie istnieją metody 4 ani 5 etapowe piątego rzędu. Co więcej dowiedziono też, że dla czteroetapowej metody czwartego rzędu nie istnieje metoda wbudowana rzędu trzeciego. Konieczne jest dodanie dodatkowego etapu. Istnieje jednak pewien techniczny zabieg nazywany *pierwszy taki jak ostatni* (ang. *FSAL, first same as last*), za pomocą którego można uzyskać następującą metodę

$$\begin{array}{c|cccc}
 0 & & & & \\
 1/3 & 1/3 & & & \\
 2/3 & -1/3 & 1 & & \\
 1 & 1 & -1 & 1 & \\
 1 & 1/8 & 3/8 & 3/8 & 1/8 \\
 \hline
 x_1 & 1/8 & 3/8 & 3/8 & 1/8 & 0 \\
 \hline
 \tilde{x}_1 & 1/12 & 1/2 & 1/4 & 0 & 1/6
 \end{array} \tag{3.16}$$

Na pierwszy rzut oka, można uznać że metoda ta jest pięcioetapowa, dokładniejsza analiza pozwala jednak zauważyć, że piąty etap jest wyliczeniem prawej strony równania różniczkowego (3.1) w punkcie  $x_1$  (czyli rozwiązania metody podstawowej) t.j.  $f(t_0 + h, x_1)$ . Ten piąty etap, w przypadku stosowania samej metody podstawowej byłby pierwszym etapem w następnym kroku. Oznacza to, że jeżeli dobrano właściwie długość kroku, efektywna złożoność obliczeniowa jest równa metodzie czteroetapowej. Wykorzystane z użyciem powstałych pięciu etapów rozwiązanie  $\tilde{x}_1$  jest rozwiązaniem trzeciego rzędu. Rozwiązanie to jest porównywane z rozwiązaniem czwartego rzędu  $x_1$ , i w oparciu o normę ich różnicy dobiera się długość kroku, spełniający przyjęte założenia o dokładności. Zazwyczaj wymaga się, aby norma różnicy rozwiązań, przyjęta jako pewna miara błędu metody (ale nie w sensie teorii miary), spełniała wymagania dotyczące błędu bezwzględnego i względnego.

*Uwaga 3.1.* Mogłoby się wydawać, że bardziej zachowawcze, byłoby używanie metody rzędu wyższego tylko do estymacji błędu, zaś jako rozwiązanie stosować „pewniejsze” przybliżenie rzędu trzeciego. Praktyka jednak (zob. Hairer et al., 2000) sugeruje, żeby wykorzystywać wyższy rząd, co jest nazywane lokalną ekstrapolacją. Uzasadnieniem podstawowym jest fakt, że dla równań nie będących równaniami sztywnymi (ang. *stiff*, trudnymi do rozwiązania numerycznie) podejście to się bardzo dobrze sprawdza, zaś dla równań sztywnych i tak trzeba stosować specjalne metody, jak np. metody Rungego-Kutty typu odwrotnego oparte na pewnego rodzaju kwadraturach Gaussa (Radau IIA i inne) albo wielokrokowe bazujące na różnicach wstecznych, jakimi są metody Geara oraz ich mutacje. ■

Metodę (3.16) nazwaną na potrzeby tej pracy metodą RK 3/8, można wyposażyć w gęste wyjście zadane równością (3.13). Jest to o tyle korzystne, że na potrzeby metody wbudowanej wyliczana jest już pochodna w prawym końcu przedziału (5 etap). Podsumowując uzyskano metodę czwartego rzędu, z estymatorem błędu rzędu trzeciego, generującą ciągle rozwiązania rzędu trzeciego (pomiędzy punktami dyskretyzacji). Ciągłe rozwiązanie to zaś posiada ciągłą pochodną (jest klasy  $C^1$ ).

Do symulacji wykonywanych na potrzeby tej pracy korzystano głównie z metody RK 3/8. W sytuacjach,

gdy potrzebna była większa dokładność, korzystano z metody Dormanda-Prince’a 5(4) (zob. Hairer et al., 2000, str. 178), która jest metodą piątego rzędu, z estymatorem czwartego rzędu i ciągłym różniczkowalnym wyjściem czwartego rzędu. Metoda Dormanda-Prince’a jest zaimplementowana w Matlabie pod nazwą `ode45` (Shampine, Gładwell i Thompson, 2003, str. 53). Metoda ta jest metodą ośmioetapową, więc wiąże się z oczywistym, dużo większym nakładem obliczeń. W sytuacji gdy napotymano równania sztywne, stosowano metodę wielokrokową Klopfensteina-Shampinea (zob. Shampine i Reichelt, 1997; Shampine et al., 2003, str. 67), która jest mutacją metody Geara o nieco mniejszej stabilności (w sensie stabilności metod numerycznych) ale lepszych własnościach pod względem dokładności. Metoda ta również posiada gęste wyjście, ale oparte o nieco inne podejście niż wyżej opisane.

### 3.2. Błędy przejściowe predykcji

W sytuacji gdy system, którego stan odtwarzamy jest niestabilny, oczywistym jest, że błędy korekcji będą wzmacniane w kroku predykcji. Okazuje się jednak, że jeżeli system jest asymptotycznie stabilny, a nawet wykładniczo stabilny błędy korekcji również mogą być wzmacniane przez predykcję, zwłaszcza jeżeli krok dyskretyzacji (więc czas trwania predykcji) będzie nieodpowiednio dobrany. W ramach niniejszego podrozdziału, będziemy rozpatrywać pojedynczy krok predykcji, i dla uproszczenia oraz zwiększenia czytelności za punkt początkowy przyjmujemy  $t_0 = 0$ . Dla ustalenia uwagi rozważmy następujący przykład.

**Przykład 3.2.** Dany jest liniowy system dynamiczny

$$\dot{x}(t) = Ax(t) \quad (3.17)$$

gdzie

$$A = \begin{bmatrix} -1 & 20 \\ 0 & -2 \end{bmatrix}$$

Niech  $t_0 = 0$  oraz

$$x(t_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

zaś krok korekcji poprzedniej iteracji algorytmu 1 wygenerował estymatę

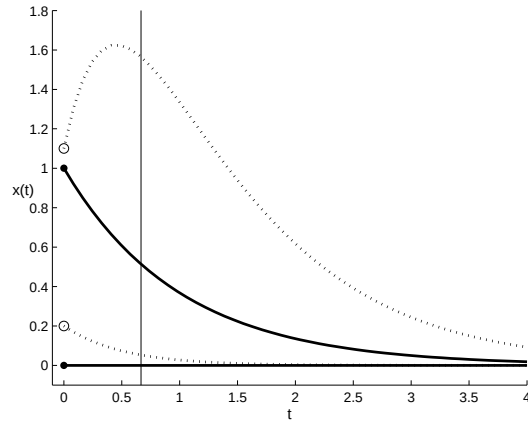
$$\hat{x}(t_0) = \begin{bmatrix} 1.1 \\ 0.2 \end{bmatrix}$$

Jak można zauważyć, norma błędu estymacji wynosi  $\sqrt{5}/10$ . Przeanalizujemy teraz jak zachowuje się estymata wygenerowana w kroku predykcji. Ilustruje to rysunek 3.1.

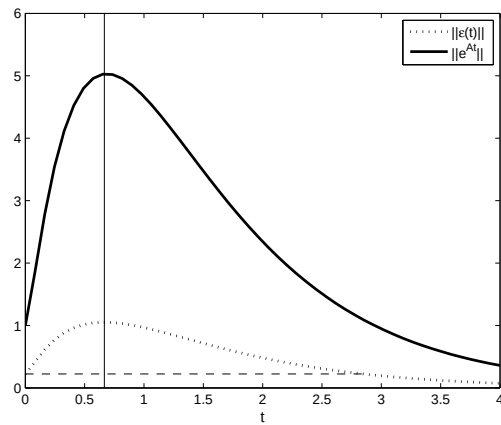
Analiza symulacji pozwala zauważyć, że początkowo mały błąd jest znacznie wzmacniany wraz z upływem czasu. W sytuacji krańcowej, norma błędu predykcji w chwili  $t_{kr} = 0.6659$  jest prawie 5 razy większa od początkowej. Jest to szczególnie istotne, jeżeli w procesie korekcji będziemy wykorzystywać wartości uzyskane w kroku predykcji (jest to wykorzystywane np. w obserwatorze dyskretno-ciągłym (zob. str. 43)). Jeżeli krok dyskretyzacji dobrano by na równy  $t_{kr}$  powstałoby ryzyko wprowadzenia dużych błędów estymacji do układu.

Przeanalizujmy jak zachowuje się norma błędu. Ilustruje to rysunek 3.2. Wykreślona na nim została norma euklidesowa wartości chwilowych błędu estymacji, oraz dla porównania znajduje się na nim również norma macierzy fundamentalnej układu (3.17). Zwróćmy uwagę, że błąd przejściowy predykcji - wynikający z dynamiki systemu - zostaje skompensowany dopiero po chwili  $t_{kom} = 2.8508$ , dopiero po tym czasie z pozoru prosty, asymptotycznie stabilny system zaczyna być dobrze tłumiony. Nie narusza to stabilności systemu (por. twierdzenie 2.3), jednak może znacznie pogorszyć jakość estymacji.

W dalszej części tego podrozdziału, omówione zostaną pewne metody szacowania błędu przejściowego, które w pewnych przypadkach mogą posłużyć jako pomoc w określaniu odstępów między pomiarami. We wszystkich oszacowaniach przez normę macierzy będziemy rozumieć normę spektralną. Można stosować



Rysunek 3.1.: Predykcja stanu w układzie (3.17). Linia ciągła - stan, linia kropkowana - estymata, o - początkowa estymata, • początkowy stan. Pionowa linia zaznacza chwilę, w której norma euklidesowa błędu jest największa



Rysunek 3.2.: Analiza norm w przykładzie 3.2. Linia kropkowana - norma euklidesowa błędu estymacji  $\varepsilon$  w funkcji czasu, linia ciągła - norma spektralna macierzy  $\|e^{At}\|$  układu (3.17), pionowa linia zaznacza chwilę, w której osiągane jest ekstremum obydwu norm, linia pozioma przerywana pokazuje kiedy norma błędu osiągnie wartość początkową



inne normy macierzowe (zob. Turowicz, 1995, str. 206), jednak norma spektralna prowadzi zazwyczaj do najlepszych oszacowań.

### 3.2.1. Szacowanie błędu predykcji w układach liniowych

Warto zwrócić uwagę, że zachowanie się błędu estymacji w przykładzie 3.2 jest silnie związane z normą macierzy  $\|e^{At}\|$ . Zarówno ekstremum, jak też i chwila w której wartość normy spada poniżej początkowej wartości pokrywają się. Jest to uzasadnione, i wynika z dynamiki błędu predykcji.

Rozważmy system

$$\dot{x} = Ax + Bu \quad (3.18)$$

oraz rozwiązywany w kroku predykcji model

$$\dot{\hat{x}} = A\hat{x} + Bu \quad (3.19)$$

błąd estymacji dany jest przez  $\varepsilon = x - \hat{x}$ . Zachowanie się błędu estymacji w kroku predykcji można opisać równaniem różniczkowym, t.j.

$$\dot{\varepsilon} = \dot{x} - \dot{\hat{x}} = \quad (3.20)$$

$$= Ax + Bu - A\hat{x} - Bu = \quad (3.21)$$

$$= A(x - \hat{x}) = A\varepsilon \quad (3.22)$$

Oznacza to, że za ewolucję błędu w kroku predykcji odpowiada macierz  $e^{At}$ , więc normę błędu w kroku predykcji, dla układów liniowych można oszacować w następujący sposób

$$\|\varepsilon(t)\| \leq \|e^{At}\| \cdot \|\varepsilon(0)\| \quad (3.23)$$

Nie znamy początkowej wartości błędu estymacji więc, jedyne co możemy analizować, to oszacowania normy macierzy  $\|e^{At}\|$ . Będziemy tu omawiać oszacowania, ponieważ w przypadkach układów wysokiego rzędu bezpośrednie obliczenie macierzy fundamentalnej, zwłaszcza dla wielu różnych wartości  $t$ , jest kłopotliwe. Poniżej podamy dwa oszacowania, z których można wyciągnąć pewne wnioski na temat zachowania się układu. Celowo nie uwzględniamy tu naturalnego oszacowania wynikającego ze stałej Lipschitza (por. lemat 2.1), ponieważ w przypadku układów liniowych jest ono zbyt zachowawcze.

#### Oszacowanie z wykorzystaniem równania Lapunowa

Dana jest macierz  $A$  o widmie w lewej otwartej półpłaszczyźnie zespolonej. Rozważa się równanie Lapunowa

$$A^T H + H A = -2I \quad (3.24)$$

Prawdziwe jest następujące oszacowanie

$$\|e^{At}\| \leq \sqrt{\|H\| \cdot \|H^{-1}\|} e^{-t/\|H\|} \quad (3.25)$$

Rozpatrując normę spektralną, można zauważyć, że z symetrii  $H$  wynika równoważne oszacowanie

$$\|e^{At}\| \leq \sqrt{\frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}} e^{-t/\lambda_{\max}(H)} \quad (3.26)$$

Dogłębna dyskusja tego rodzaju oszacowania, wraz z przypadkami abstrakcyjnymi oraz licznymi usprawnieniami znajduje się np. w (Veselić, 2003). Szczególnie interesującym wnioskiem jest pewien szczególny przypadek, mianowicie jeżeli  $A \in \mathbb{R}^{n \times n}$  jest macierzą normalną ( $A^T A = A A^T$ ) to zachodzi

$$\|e^{At}\| \leq e^{-t/\lambda_{\max}(H)} \quad (3.27)$$

**Oszacowanie z wykorzystaniem tzw. „normy logarytmicznej”**

Inne podejście do oszacowywania macierzy fundamentalnej opiera się na tzw. „normie logarytmicznej” (ang. *logarithmic norm*), nazywanej również współczynnikiem tłumienia. Z definicji:

$$\mu(A) = \lim_{\Delta \rightarrow 0+} \frac{\|I + \Delta A\| - 1}{\Delta} \quad (3.28)$$

W zależności od tego jaka norma zostanie zastosowana we wzorze (3.28) można uzyskać różne końcowe zależności, dla normy spektralnej

$$\mu(A) = \lambda_{\max} \left( \frac{A + A^T}{2} \right) \quad (3.29)$$

Co jest istotne „norma logarytmiczna” nie jest normą, ponieważ może przyjmować wartości ujemne. Za jej pomocą można uzyskać następujące oszacowanie

$$\|e^{At}\| \leq e^{\mu(A)t} \quad (3.30)$$

Omówienie tego rodzaju oszacowania, z przypadkami innych norm niż spektralna (zob. Guang-Da i Mingzhu, 2004).

**Inne oszacowania**

Dla normy spektralnej można podać jeszcze inny rodzaj oszacowania  $\|e^{At}\|$  (zob. Demidowicz, 1972; Mitkowski, 2007b, str. 133)

$$\forall \epsilon > 0, \quad \exists c(\epsilon) > 0: \quad \|e^{At}\| \leq c(\epsilon)e^{(\alpha+\epsilon)t}, \quad 0 \leq t < \infty \quad (3.31)$$

przy czym  $\alpha = \max_i \operatorname{Re} \lambda_i(A)$ . Jeżeli wartości własne, mające największe części rzeczywiste równe  $\alpha$ , są pojedynczymi pierwiastkami wielomianu minimalnego macierzy  $A$ , to

$$\|e^{At}\| \leq ce^{\alpha t}, \quad 0 \leq t < \infty \quad (3.32)$$

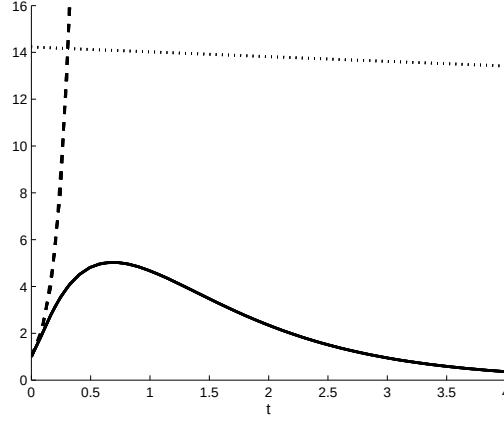
Metoda ta niestety nie jest zbyt praktyczna w zastosowaniach, pomimo bardzo z pozoru atrakcyjnego sposobu wyznaczania szybkości tłumienia trajektorii. Podstawową wadą jest brak jakiejkolwiek informacji o  $c$  lub  $c(\epsilon)$  co sprawia, że mogą być one potencjalnie dowolnie duże, oszacowanie to nie nadaje się więc do analizy ilościowej błędów przejściowych predykcji.

Zastosujmy pierwsze dwa oszacowanie dla systemu (3.17) z przykładu 3.2. Ilustruje to rysunek 3.3. Jak można zauważyć obydwa oszacowania są dość konserwatywne. Oszacowanie (3.26) jest eksponentą o ujemnym wykładniku, jednakże bardzo bliskim zeru, co więcej stała  $\sqrt{\frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}}$  jest stosunkowo duża, więc oszacowanie będzie bliskie dokładnej wartości normy dopiero po bardzo długim czasie. Oszacowanie (3.30) jest natomiast eksponentą o wykładniku dodatnim, więc nie oddaje asymptotyki estymowanej normy. Warto jednak zauważyć, że w początkowym odcinku przybliża ona normę  $\|e^{At}\|$  bardzo dokładnie. Niezależnie obydwa te oszacowania wnoszą stosunkowo niewiele informacji, warto jednak zauważyć że dają one pewną intuicję na temat jak szybko błąd przejściowy może narastać, oraz jakie wielkości nie przekroczy.

Analiza oszacowań, w szczególności zaś (3.27) i (3.30) sugeruje postawienie następującego pytania - kiedy norma  $\|e^{At}\|$  nie posiada ekstremum - czyli w konsekwencji kiedy predykcja w układzie liniowym nie będzie na pewno wносить błędu? Następuje to w sytuacji, gdy  $e^{At}$  jest półgrupą kontrakcji (półgrupą zwiężającą) (zob. np. Mitkowski, 1991, str. 170), tzn.

$$\|e^{At}\| \leq 1, \quad t \geq 0 \quad (3.33)$$

Jeżeli zachodzi (3.33) to mamy gwarancję, że norma błędu predykcji w układzie liniowym nie będzie większa niż norma błędu estymacji przeprowadzonej w kroku korekcji. Jeżeli dodatkowo system będzie asymptotycznie stabilny ( $\operatorname{Re} \lambda(A) < 0$ ) to błąd będzie wykładniczo zanikał w trakcie predykcji. W oparciu o podane wcześniej oszacowania można podać następujące kryteria, aby  $e^{At}$  było półgrupą kontrakcji (równoważnie, żeby macierz



Rysunek 3.3.: Analiza oszacowań normy macierzy  $\|e^{At}\|$  z przykładu 3.2. Linia ciągła - norma macierzy  $e^{At}$ , linia przerywana - oszacowanie (3.30), linia kropkowana - oszacowanie (3.26)

$A$  generowała półgrupę kontrakcji):

1.  $A \in \mathbb{R}^n$  jest macierzą normalną oraz  $\operatorname{Re} \lambda(A) < 0$  wtedy

$$\|e^{At}\| \leq e^{-t/\lambda_{\max}(H)} \leq 1 \quad t \geq 0 \quad (3.34)$$

gdzie  $H = H^T > 0$  jest rozwiązaniem równania (3.24).

2. Współczynnik tłumienia („norma logarytmiczna”)  $\mu(A) \leq 0$  wtedy

$$\|e^{At}\| \leq e^{\mu(A)t} \leq 1 \quad t \geq 0 \quad (3.35)$$

Warunek pierwszy jest wymaganiem bardzo restrykcyjnym, które jest szczególnie ograniczające, jeżeli chcemy wykorzystać je do doboru np. parametrów regulatora lub obserwatora (co nie jest zagadnieniem poruszonym w tej pracy). Drugie z wymagań jest łatwiejsze do spełnienia, w czym pomagają również poniższe

**Twierdzenie 3.1.** *Dana jest macierz  $A \in \mathbb{R}^{n \times n}$  taka, że*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

*oraz dla każdego  $1 \leq i \leq n$  zachodzą poniższe nierówności*

$$a_{ii} < 0 \quad (3.36)$$

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}| \quad (3.37)$$

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ji}| \quad (3.38)$$

wtedy  $e^{At}$  jest półgrupą kontrakcji oraz dodatkowo jest oszacowana przez

$$\|e^{At}\| \leq e^{\alpha t} \quad (3.39)$$

$$\alpha = \sup_{1 \leq i \leq n} \left( a_{ii} + \sum_{j=1, j \neq i}^n \left| \frac{1}{2}(a_{ij} + a_{ji}) \right| \right) \quad (3.40)$$

*Dowód.* Dowód bazuje na twierdzeniu Gerszgorina (zob. twierdzenie A.5) oraz na pojęciu „normy logarytmicznej” i oszacowaniu (3.30). Na początek należy zwrócić uwagę, że koła Gerszgorina dla macierzy  $A$  oraz  $A^T$  znajdują się ściśle w lewej otwartej półpłaszczyźnie zespolonej - wynika to z założeń (3.37) i (3.38). Przeanalizujemy teraz wyrażenie opisujące „normę logarytmiczną” dla macierzy  $A$

$$\mu(A) = \lambda_{\max}(B) \quad (3.41)$$

$$B = \frac{A + A^T}{2} \quad (3.42)$$

Współczynniki macierzy  $B$  dane są więc wzorami

$$b_{ii} = a_{ii} \quad (3.43)$$

$$b_{ij} = \frac{1}{2}(a_{ij} + a_{ji}) \quad (3.44)$$

Pokażemy teraz, że koła Gerszgorina dla macierzy  $B$  również znajdują się ściśle w lewej otwartej półpłaszczyźnie zespolonej. Po pierwsze - ponieważ diagonale (3.36) i (3.43) są sobie równe, więc środki kół Gerszgorina będą się znajdować na ujemnej części osi rzeczywistej. Przeanalizujemy teraz promienie kół. Dla ustalonego  $i$  mamy promień  $r_i$  równy

$$r_i = \sum_{j=1, j \neq i}^n |b_{ij}| = \sum_{j=1, j \neq i}^n \left| \frac{1}{2}(a_{ij} + a_{ji}) \right| \leq \quad (3.45)$$

$$\leq \frac{1}{2} \left( \sum_{j=1, j \neq i}^n |a_{ij}| + \sum_{j=1, j \neq i}^n |a_{ji}| \right) < \quad (3.46)$$

$$< \frac{1}{2}(|a_{ii}| + |a_{ii}|) = |a_{ii}| \quad (3.47)$$

Oznacza to, że dla każdego ustalonego  $i$ , promień danego koła Gerszgorina jest mniejszy od odległości środka tego koła od osi urojonej. Dlatego też, zgodnie z twierdzeniem Gerszgorina, oraz ponieważ macierz  $B$  jest symetryczna, wszystkie jej wartości własne leżą w ujemnej części osi rzeczywistej. Z równości (3.41) wynika więc, że „norma logarytmiczna” macierzy  $A$  jest ujemna. To zaś w konsekwencji oznacza, że zgodnie z oszacowaniem (3.30) norma  $e^{At}$  jest oszacowana przez zanikającą funkcję wykładniczą. Pozostaje jeszcze podać oszacowanie na  $\mu(A)$ . Ponieważ wszystkie wartości własne macierzy  $B$  leżą w pewnym przedziale domkniętym (lub sumie przedziałów) na ujemnej części osi liczb rzeczywistych, oszacowaniem górnym największej wartości własnej (wszystkie są ujemne) będzie koniec przedziału najbliższy osi urojonej. Wartość  $\alpha$  tego wierzchołka określa zaś wyrażenie (3.40).  $\square$

*Uwaga 3.2.* Warto porównać powyższe wyniki z twierdzeniem Lumera-Philipsa (zob. Engel i Nagel, 2000, str. 83). Twierdzenie to w pewnym uproszczeniu mówi, że jeżeli  $A$  jest liniowym, domkniętym operatorem dysypatywnym o gęstej dziedzinie oraz operator  $(\lambda I - A)$  jest surjekcją dla  $\lambda > 0$  to generuje on półgrupę kontrakcji. W przypadku, gdy  $A$  jest macierzą rzeczywistą, a przestrzenią stanu jest  $\mathbb{R}^n$ , to widmo  $A$  ma być rozłączne ze półosią  $(0, +\infty)$ , oraz spełniony musi być warunek dysypatywności operatora, t.j. (zob. Engel i Nagel, 2000, str. 82)

$$\|(\lambda I - A)x\| \geq \lambda \|x\|, \quad \lambda > 0 \text{ i } x \in \mathbb{R}^n \quad (3.48)$$

lub równoważnie (zob. Engel i Nagel, 2000, str. 88)

$$\forall x \in \mathbb{R}^n \quad (Ax, x) \leq 0 \quad (3.49)$$

Warto zauważyć, że warunek (3.49) prowadzi do takiego samego wniosku jak wykorzystanie współczynnika tłumienia („normy logarytmicznej”) ponieważ

$$(Ax, x) = x^T A^T x = x^T \left( \frac{A^T + A}{2} \right) x \quad (3.50)$$

zaś forma kwadratowa (3.50) jest niedodatnio określona, wtedy i tylko wtedy gdy  $\lambda(A^T + A) \leq 0$  (wartości własne macierzy  $(A^T + A)$  są liczbami rzeczywistymi), czyli gdy  $\mu(A) \leq 0$ . Dysypatywne operatory rozważał również (Skruch, 2005, str. 90). ■

### 3.2.2. Szacowanie błędu predykcji w układach nieliniowych

W przypadku układów nieliniowych ciężko podać konstruktywne oszacowania błędu predykcji. Zasadniczo jedynym pełnym, aczkolwiek bardzo konserwatywnym sposobem jest ten oparty o stałą Lipschitza, stosowany w lemacie 2.1. Warto jednak zwrócić uwagę na pewną metody szacowania trajektorii dla nieliniowych systemów asymptotycznie stabilnych. Znane jest następujące

**Twierdzenie 3.2.** *Dany jest system*

$$\dot{x}(t) = f(x(t)) \quad f(0) = 0$$

gdzie  $x(t) \in \mathbb{R}^n$ . Niech istnieje dla niego funkcjonal Lapunowa  $v(z) > 0$ ,  $z \in \mathbb{R}^n$ ,  $z \neq 0$ , taki, że  $\dot{v}(z) < 0$  dla  $z \in \mathbb{R}^n$ ,  $z \neq 0$ . Niech

$$\gamma = \max_z -\frac{v(z)}{\dot{v}(z)}, \quad z \in \mathbb{R}^n, \quad z \neq 0$$

dla  $t \in T = [0, \infty)$  mamy

$$-\frac{v(z)}{\dot{v}(z)} \leq \gamma$$

więc

$$v(x(t)) \leq v(x(0))e^{-\frac{t}{\gamma}}$$

(zob. np. Mitkowski, 1996, str. 90)

Wykorzystując powyższe twierdzenie można wyprowadzić oszacowanie (3.26). Dla układów nieliniowych nie daje ono już tak dużo informacji. Najlepsze wyniki są do osiągnięcia w układach wykładniczo stabilnych, gdzie funkcjonal Lapunowa można oszacować zarówno z dołu jak i z góry przez normę  $z$ . Wtedy można uzyskać dobre oszacowanie trajektorii. W pozostałych przypadkach informacje o trajektorii uzyskane z twierdzenia 3.2 mogą być dużo bardziej uwikłane. Informacje te jednak mogą posłużyć do sformułowania podobnej analizy jak na rysunku 3.3, dając pewne przybliżenie na temat możliwego położenia faktycznego stanu systemu, oraz ewentualnie potencjalnych okresów wzmocnienia błędu (jak przy układach liniowych). W pewnych sytuacjach, szczególnie jeżeli system nieliniowy znajduje się blisko położenia równowagi można posłużyć się metodą linearyzacji oraz wynikami z sekcji 3.2.1, uzyskane wyniki, będą mieć jednak wtedy jedynie charakter lokalny.

### 3.3. Błędy predykcji wywołane zakłóceniami wejściowymi

Jak dotąd w pracy rozważano systemy bez zakłóceń, zarówno w pomiarze jak i w równaniach systemu. Rozważmy system z zakłóceniami postaci

$$\dot{x}(t) = f(x(t), u(t), w(t)) \quad (3.51)$$

$$y(t) = h(x(t), v(t)) \quad (3.52)$$

gdzie  $f: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ ,  $h: \mathbb{R}^n \times \mathbb{R}^s \rightarrow \mathbb{R}^m$ . W tej sekcji zajmujemy się analizą wpływu zakłóceń na ciągłą estymatę powstającą w kroku predykcji. Można wyróżnić pośredni oraz bezpośredni wpływ zakłóceń. Przez wpływ pośredni będziemy rozumieć błąd estymacji w kroku korekcji wywołany obecnością zakłóceń (zarówno pomiarów wyjścia  $v$  jak też i działających bezpośrednio na system  $w$ ). W ten sposób wartość estymaty na początku predykcji będzie zaburzona, co będzie przenosić się na błąd ciągłej estymaty. Przez błąd bezpośredni będziemy rozumieć różnicę między ciągłą estymatą a faktycznym stanem systemu, wywołaną tym, że na system rzeczywisty działają nieznane zakłócenia, które nie są obecne w modelu. Przykładem mogą być np. wahania momentu obciążającego silnika. Jak łatwo zauważyć, wpływ bezpośredni ma jedynie zakłócenie  $w$  działające na system.

Poniżej przedstawiono twierdzenie, podające jak można oszacować błąd wywołany bezpośrednim wpływem zakłóceń. Dla uproszczenia, ograniczymy się do zakłóceń o wartościach skalarnych.

**Twierdzenie 3.3.** *Dany jest system*

$$\dot{x}(t) = f(x(t), u(t), w(t)), \quad x(t_i) = x_0 \quad (3.53)$$

*Oraz ciągła estymata, dana równaniem*

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t), 0), \quad \hat{x}(t_i) = \hat{x}_0 \quad (3.54)$$

*Przy czym  $t \in [t_i, t_{i+1}]$  oraz*

$$|f(x_1, u, w) - f(x_2, u, w)| \leq L_x |x_1 - x_2| \quad \forall x_1, x_2 \in \mathbb{R}^n \quad \forall u \in \mathbb{R}^r \quad \forall w \in \mathbb{R} \quad (3.55)$$

$$|f(x, u, w_1) - f(x, u, w_2)| \leq L_w |w_1 - w_2| \quad \forall x \in \mathbb{R}^n \quad \forall u \in \mathbb{R}^r \quad \forall w_1, w_2 \in \mathbb{R} \quad (3.56)$$

$$w \in L^\infty([t_i, t_{i+1}], \mathbb{R}) \quad (3.57)$$

*wtedy norma błędu estymacji  $\varepsilon(t)$  spełnia oszacowanie*

$$|\varepsilon(t)| \leq e^{L_x(t_{i+1}-t_i)} (|\varepsilon_0| + L_w(t_{i+1} - t_i) \|w\|_\infty) \quad \text{dla } t \in [t_i, t_{i+1}] \quad (3.58)$$

*przy czym  $\varepsilon_0 = x_0 - \hat{x}_0$ .*

*Dowód.* W dowodzie pominiemy argument czasowy tam, gdzie nie będzie to stwarzać wątpliwości. Zgodnie z definicją błędu estymacji (2.18) możemy opisać jego ewolucję równaniem

$$\dot{\varepsilon} = f(x, u, w) - f(\hat{x}, u, 0) \quad (3.59)$$

$$\varepsilon(t_i) = \varepsilon_0 = x_0 - \hat{x}_0 \quad (3.60)$$

Rozwiązanie równania (3.59) dla  $t \in [t_i, t_{i+1}]$  wyrazić można wzorem

$$\varepsilon(t) = \varepsilon_0 + \int_{t_i}^t (f(x, u, w) - f(\hat{x}, u, 0)) dt \quad (3.61)$$

stąd mamy

$$|\varepsilon(t)| \leq |\varepsilon_0| + \left| \int_{t_i}^t (f(x, u, w) - f(\hat{x}, u, 0) + f(x, u, 0) - f(x, u, 0)) dt \right| \leq \quad (3.62)$$

$$\leq |\varepsilon_0| + \int_{t_i}^t (|f(x, u, w) - f(x, u, 0)| + |f(x, u, 0) - f(\hat{x}, u, 0)|) dt \leq \quad (3.63)$$

$$\leq |\varepsilon_0| + \int_{t_i}^t L_w |w| dt + \int_{t_i}^t L_x |x - \hat{x}| dt \leq \quad (3.64)$$

$$\leq |\varepsilon_0| + (t_{i+1} - t_i) L_w \|w\|_\infty + \int_{t_i}^t L_x |x - \hat{x}| dt \quad (3.65)$$

Do ostatniej nierówności możemy zaś zastosować lemat Gronwalla-Bellmana (lemat A.1 na stronie 98), co daje

$$|\varepsilon(t)| \leq e^{L_x(t_{i+1}-t_i)} (|\varepsilon_0| + L_w(t_{i+1} - t_i) \|w\|_\infty) \quad \text{dla } t \in [t_i, t_{i+1}] \quad (3.66)$$

□

Twierdzenie 3.3 pokazuje, że pomimo występowania zakłóceń w prawej stronie równania różniczkowego (3.51) błąd predykcji pozostaje ograniczony. Co więcej, warto zauważyć, że przy globalnie atraktywnej procedurze korekcji, błąd predykcji będzie zależny od normy zakłóceń, więc wyeliminowanie go bez ich znajomości raczej nie jest możliwe. Pewne metody pozwalają na estymowanie zakłóceń, lecz wiąże się to zazwyczaj z ogromnymi nakładami obliczeniowymi. Należy jednak zwrócić uwagę, że jeżeli zakłócenia będą miały zerową wartość średnią i wysoko częstotliwościowy charakter, zaś system dynamiczny będzie posiadał pewne właściwości filtrujące to ich efekt będzie niewielki, a oszacowanie (3.58) bardzo zachowawcze. Jeżeli natomiast ich wartość średnia jest różna od zera, to najprawdopodobniej, należy uwzględnić w modelu pewne dodatkowe czynniki.

## Rozdział 4.

### Korekcja w układach liniowych

#### 4.1. Wprowadzenie

W ramach niniejszego rozdziału, rozważane będą systemy liniowe (4.1) z wyjściem liniowym (4.2)

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.1)$$

$$y(t) = Cx(t) \quad (4.2)$$

gdzie  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^r$ ,  $y(t) \in \mathbb{R}^m$ ,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times r}$  oraz  $C \in \mathbb{R}^{m \times n}$ . Dla ustalenia uwagi przyjmujemy, że wartości wyjścia znane są w dyskretnych chwilach czasu  $t_i$ , gdzie

$$t_i = t_{i-1} + h_i \quad (4.3)$$

$$t_0 \geq 0 \quad (4.4)$$

$$h_i > 0 \quad (4.5)$$

$$i = 0, 1, 2, 3, \dots \quad (4.6)$$

przy czym

$$h_i \geq \tau_{max} \quad (4.7)$$

$$\tau_{max} = \max_i (t_{i+1} - t_i) < \infty \quad (4.8)$$

Pomiary mogą, ale nie muszą być równoodległe. W takim przypadku mamy

$$h_i = h = \tau_{max} \quad \forall i \quad (4.9)$$

$$t_i = t_0 + ih \quad (4.10)$$

Będziemy zakładać, że para  $(C; A)$  jest obserwowalna, oraz w przypadku, gdy rozpatrywane będą pomiary równoodległe, że dla danych pary macierzy  $(C; A)$  i kroku dyskretyzacji  $h$  zachodzi twierdzenie A.3. W następnych podrozdziałach rozważana będzie również postać modelu (4.1) oraz wyjścia (4.2) z zakłóceniami. Będą one omawiane we właściwym czasie.

*Uwaga 4.1 (Autonomizacja).* Podany wyżej system liniowy, można bez straty informacji poddać tzw. autonomizacji. Podobne podejście, pod nazwą zmodyfikowana odpowiedź układu występuje w (Kaczorek, 1977, str. 43). Autonomizacja polega to na oddzielnym rozpatrywaniu wpływu warunków początkowych (zasadnicze zadanie estymacji stanu) oraz sterowania

$$x(t) = x_p(t) + x_u(t) \quad (4.11)$$

gdzie

$$x_p(t) = e^{A(t-t_0)}x(t_0) \quad (4.12)$$

$$x_u(t) = \int_{t_0}^t \exp(A(t-\tau))Bu(\tau)d\tau \quad (4.13)$$

Ponieważ  $x_u$  nie zależy od nieznanych warunków początkowych a jedynie od znanego sterowania, może zostać wyliczone niezależnie. Zadanie estymacji stanu systemu (4.1) więc można sprowadzić do estymacji stanu w



systemie autonomicznym

$$\dot{x}_p(t) = \dot{x}(t) - \dot{x}_u(t) = A(x(t) - x_u(t)) + Bu(t) - Bu(t) = Ax_p(t) \quad (4.14)$$

$$y_p(t) = Cx_p(t) = y(t) - Cx_u(t) \quad (4.15)$$

Dlatego też poniżej, tam gdzie to wygodniejsze rozważać będziemy estymację stanu w systemach autonomicznych

$$\dot{x}(t) = Ax(t) \quad (4.16)$$

■

*Uwaga 4.2* (Wyliczanie  $x_u$ ). Teoretycznie  $x_u(t)$  może zostać wyliczone z wykorzystaniem wzoru analitycznego (4.13). Takie podejście jest niestety niepraktyczne. Abstrahując od wyliczania wartości całki spłotowej analitycznie dla arbitralnie przyjętego  $u(t)$ , również numeryczne procedury kwadratur nie są najlepszą drogą do rozwiązania problemu. Jest to spowodowane pewnymi trudnościami związanymi z wyliczeniem macierzy  $e^{At}$ , szczególnie dla macierzy  $A$  o dużym wymiarze, nie będących macierzami o szczególnej strukturze (np. macierzami normalnymi). Jak można znaleźć w literaturze (Molery i Van Loan, 2003) najbardziej ogólne i efektywne metody wyliczania macierzy fundamentalnej oparte są o algorytmy skalowania i kwadratowania z wykorzystaniem aproksymacji Padé. Wyliczanie całek z iloczynów takich macierzy staje się jednak kłopotliwe, zwłaszcza w sytuacji, gdy np.  $\|e^{At}\|$  ma wiele ekstremów. Warto jednak zauważyć, że wyliczanie (4.13) jest równoważne rozwiązywaniu równania różniczkowego

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.17)$$

$$x(t_0) = 0 \quad (4.18)$$

$$t \in [t_0, t] \quad (4.19)$$

które można wyznaczyć za pomocą względnie dowolnej procedury numerycznego całkowania równań różniczkowych. Warto jednak zaznaczyć, że wyznaczając takie rozwiązanie należy uwzględnić ewentualne nieregularności funkcji  $u(t)$ , tzn. podzielić przedział całkowania równania w punktach jej nieciągłości lub nieróżniczkowalności (zob. np. Hairer et al., 2000, str. 198).

■

*Uwaga 4.3* (Problem estymacji stanu jako równanie liniowe). Rozważmy następujący system

$$\dot{x}(t) = Ax(t) \quad (4.20)$$

$$y(t) = Cx(t) \quad (4.21)$$

$$x(0) = x_0 \in \mathbb{R}^n \quad (4.22)$$

przy czym para  $(C; A)$  jest obserwowalna.

Oznaczmy przestrzeń warunków początkowych przez  $X_0 = \mathbb{R}^n$ , przestrzeń stanu przez  $X$  i przestrzeń wyjść przez  $Y$ . Można zdefiniować następujące operacje liniowe

$$\mathcal{A}: X_0 \rightarrow X \quad (4.23)$$

$$\mathcal{C}: X \rightarrow Y \quad (4.24)$$

z ich pomocą, system (4.20) - (4.21) można zapisać jako

$$y = \mathcal{C}\mathcal{A}x_0 = \mathcal{G}x_0 \quad (4.25)$$

gdzie  $y \in Y$  a  $\mathcal{G}: \mathbb{R}^n \rightarrow Y$ . Wynika stąd, że problem odtwarzania stanu można sprowadzić do rozwiązywania równania liniowego (4.25) względem  $x_0$ . Para  $(C; A)$  jest obserwowalna, a przestrzeń  $Y$  została odpowiednio dobrana, więc jeżeli równanie (4.25) posiada rozwiązanie to jest ono jednoznaczne (brak rozwiązania może być wywołany przez obecność zakłóceń wyjścia). W zależności jednak od tego jak zostaną wybrane przestrzenie  $Y$  rozwiązanie może mieć różną postać.

Szczególnie dogodnie, jeżeli  $Y$  jest przestrzenią Hilberta (zob. np. Mlak, 1970, str. 56). Należy wtedy rozpatrzyć dwa przypadki (zob. np. Mitkowski, 2008).

#### Przypadek 1

Jeżeli  $\mathcal{G}: \mathbb{R}^n \rightarrow Y$  jest odwzorowaniem różnowartościowym oraz odwzorowuje  $\mathbb{R}^n$  na  $Y$  to istnieje operacja odwrotna  $\mathcal{G}^{-1}$ . W tym przypadku istnieje dokładnie jedno rozwiązanie równania (4.25) postaci

$$x_0 = \mathcal{G}^{-1}y \quad (4.26)$$

#### Przypadek 2

W przypadku gdy  $\|\mathcal{G}x - y\| > 0$  - rozwiązanie dokładne równania (4.25) nie istnieje - można poszukiwać tak zwanego rozwiązania przybliżonego optymalnie minimalizującego normę  $\|\mathcal{G}x - y\|$ . Brak istnienia rozwiązania dokładnego w przypadku problemu odtwarzania stanu w systemie (4.20) może być wywołany przez obecność zakłóceń.

Przy ustalonym  $y$  wektor  $x_0$  minimalizuje normę  $\|\mathcal{G}x - y\|$  wtedy i tylko wtedy gdy (zob. np. Mitkowski, 2007a)

$$\mathcal{G}^*\mathcal{G}x_0 = \mathcal{G}^*y \quad (4.27)$$

gdzie  $\mathcal{G}^*$  to operator sprzężony do  $\mathcal{G}$  taki, że

$$\mathcal{G}^*: Y \rightarrow \mathbb{R}^n, \quad (\mathcal{G}v, w)_Y = (v, \mathcal{G}^*w), \quad \forall v \in \mathbb{R}^n, \quad \forall w \in Y \quad (4.28)$$

przy czym  $(\cdot, \cdot)_Y$  oznacza iloczyn skalarny w przestrzeni wyjść  $Y$ . W przypadku szczególnym, gdy  $\mathcal{G}^*\mathcal{G}$  jest operatorem odwracalnym, to

$$x_0 = (\mathcal{G}^*\mathcal{G})^{-1} \mathcal{G}^*y \quad (4.29)$$

przy czym równość  $\mathcal{G}^\# = (\mathcal{G}^*\mathcal{G})^{-1} \mathcal{G}^*$  określa nam operator pseudoodwrotny do  $\mathcal{G}$ .

*Przykład 4.1* ( $Y = L^2[0, T, \mathbb{R}^m]$ ). Dla tak określonej przestrzeni wyjścia mamy (zob. np. Mitkowski, 2008)

$$\mathcal{G}x_0 = Ce^{At}x_0, \quad \mathcal{G}: \mathbb{R}^n \rightarrow L^2[0, T, \mathbb{R}^m] \quad (4.30)$$

$$\mathcal{G}^*y = \int_0^T e^{\tau A^\top} C^\top y(\tau) d\tau, \quad \mathcal{G}^*: L^2[0, T, \mathbb{R}^m] \rightarrow \mathbb{R}^n \quad (4.31)$$

Operacja  $\mathcal{G}^*\mathcal{G}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  jest macierzą i ma postać

$$\mathcal{G}^*\mathcal{G} = \int_0^T e^{\tau A^\top} C^\top Ce^{A\tau} d\tau \quad (4.32)$$

i jest nieosobliwa (por. tw. A.2 na str. 98). Efektywny wzór na  $x_0$  można wyrazić jako (zob. np. Mitkowski, 2007a)

$$x_0 = \left( \int_0^T e^{\tau A^\top} C^\top Ce^{A\tau} d\tau \right)^{-1} \int_0^T e^{\tau A^\top} C^\top y(\tau) d\tau \quad (4.33)$$

W ten sposób, dobierając odpowiednią przestrzeń wyjść  $Y$  można uzyskać różne sposoby estymacji warunku początkowego (lub końcowego), są nimi dyskretny obserwator dokładny (str. 44) i obserwator najmniejszych kwadratów (str. 46). Inna sytuacja zachodzi w przypadku obserwacji  $C^0$  optymalnej (str. 47) gdzie przestrzenią wyjść jest przestrzeń  $C^0[0, T]$  (ew. ogólniej  $C^0[0, T, \mathbb{R}^n]$ ), która jest przestrzenią Banacha ale nie jest przestrzenią Hilberta. ■

## 4.2. Estymacja bez zakłóceń

W ramach tego podrozdziału rozważać będziemy systemy (4.1) z wyjściem (4.2). Zaproponowane zostaną dwa podejścia - asymptotyczne obserwatory dyskretno-ciągłe, ze szczególnym uwzględnieniem obserwatora dead-beat, oraz dyskretną estymację dokładną, opartą o rozwiązywanie układów algebraicznych równań liniowych.

Dla obydwu tych podejść zaprezentowane zostaną stosowne algorytmy procedury korekcji do wykorzystania w algorytmie 1.

#### 4.2.1. Obserwatory dyskretno-ciągłe

Dla liniowych ciągłych systemów dynamicznych (4.1) z wyjściem (4.2) znana jest koncepcja obserwatora Luenbergera, t.j.

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + G(y(t) - C\hat{x}(t)) \quad (4.34)$$

(zob. np. Mitkowski, 1991, str. 89, Byrski, 2007, str. 359). Jak można zauważyć, do wykorzystania obserwatora (4.34) potrzebna jest ciągła znajomość wyjścia systemu. Z drugiej strony znany jest również dyskretny wariant obserwatora (dla równoodległych pomiarów) w postaci

$$\hat{x}(h \cdot (k + 1)) = A_d \hat{x}(h \cdot k) + B_d u(h \cdot k) + G_d (y(h \cdot k) - C \hat{x}(h \cdot k)) \quad (4.35)$$

(zob. np. Mitkowski, 1991, str. 141; Baranowski, Hajduk, Korytowski, Mitkowski (red.) i Tutaj, 2007, str. 173). Przy czym

$$A_d = \exp(Ah) \quad (4.36)$$

$$B_d = \int_0^h \exp(As) B ds \quad (4.37)$$

Podobnie jak w przypadku obserwatora (4.34) w obserwatorze (4.35) pojawia się pewna niedogodność. Mianowicie, zakłada się, że sterowanie  $u(t)$  dla  $t \in [kh, (k + 1)h]$  przyjmuje wartość stałą. Problem ten można obejść za pomocą tzw. obserwatora dyskretno-ciągłego (zob. np. Astorga, Othman, Othman, Hammouri i McKenna, 2002; Angeles, Astorga, Alvarado i Vela, 2004; Baranowski, 2007c; Mitkowski i Baranowski, 2007; Baranowski i Tutaj, 2007).

Obserwator ten można sformułować w następujący sposób. Rozważmy system dyskretny

$$x(h \cdot (k + 1)) = A_d x(h \cdot k) + B_u(h \cdot k) \quad (4.38)$$

przy czym  $A_d$  dana jest wzorem (4.36), zaś

$$B_u(h \cdot k) = \int_{h \cdot k}^{h \cdot (k+1)} \exp(A(t - \tau)) B u(\tau) d\tau \quad (4.39)$$

Jak łatwo zauważyć, postać systemu dyskretnego (4.38) wynika bezpośrednio z rozkładu rozwiązania (4.11) systemu liniowego (4.1) na człony od warunku początkowego i od sterowania. Prawa strona (4.38) jest po prostu rozwiązaniem (4.11), na odcinku  $t \in [kh, (k + 1)h]$ . Dla systemu dyskretnego (4.38) możemy sformułować obserwator

$$\hat{x}(h \cdot (k + 1)) = A_d \hat{x}(h \cdot k) + B_u(h \cdot k) + G_d (y(h \cdot k) - C \hat{x}(h \cdot k)) \quad (4.40)$$

Przeanalizujemy dynamikę błędu estymacji, niech

$$\varepsilon(t) = x(t) - \hat{x}(t) \quad (4.41)$$

wtedy

$$\begin{aligned} \varepsilon(h \cdot (k + 1)) &= x(h \cdot (k + 1)) - \hat{x}(h \cdot (k + 1)) = \\ &= A_d x(h \cdot k) + B_u(h \cdot k) - A_d \hat{x}(h \cdot k) - B_u(h \cdot k) - G_d (y(h \cdot k) - C \hat{x}(h \cdot k)) = \\ &= A_d x(h \cdot k) - A_d \hat{x}(h \cdot k) - G_d C (x(h \cdot k) - \hat{x}(h \cdot k)) = \\ &= (A_d - G_d C) (x(h \cdot k) - \hat{x}(h \cdot k)) = (A_d - G_d C) \varepsilon(h \cdot k) \end{aligned} \quad (4.42)$$

Dynamika błędu estymacji takiego obserwatora jest opisana równaniem różnicowym, w którym sterowanie

nie występuje. Można więc zastosować do niego klasyczne techniki projektowania dla układów liniowych dyskretnych.

Jedną z takich technik jest obserwator dead-beat (zob. np. Baranowski i Mitkowski, 2006; Baranowski, Długosz i Mitkowski, 2008a). Koncepcja takiego obserwatora jest konsekwencją twierdzenia Hamiltona-Cayleya zob. np. Mitkowski, 2007b, 70; Turowicz, 1995, str. 120).

Zwróćmy uwagę, że dla  $k = n - 1$  możemy wyrazić błąd obserwatora jako

$$\begin{aligned}\varepsilon(h \cdot n) &= (A_d - G_d C) \varepsilon(h \cdot (n - 1)) = (A_d - G_d C)^2 \varepsilon(h \cdot (n - 2)) = \\ &= \dots = (A_d - G_d C)^n \varepsilon(0)\end{aligned}\tag{4.43}$$

Jedną z konsekwencji obserwowalności pary  $(C; A_d)$  (zob. np. Mitkowski, 1991, str. 87), jest fakt, że dla każdego zestawu rzeczywistych współczynników wielomianu charakterystycznego macierzy rzeczywistej o wymiarach  $n \times n$ , istnieje takie  $G_d$ , że macierz  $(A_d - G_d C)$  ma wielomian charakterystyczny o takich właśnie współczynnikach. Jeżeli macierz o wymiarach  $n \times n$  ma wszystkie wartości własne równe 0, to wtedy jej wielomian charakterystyczny ma postać

$$W(\lambda) = \lambda^n\tag{4.44}$$

Jeżeli macierz  $G_d$  zostanie dobrana tak aby wielomian charakterystyczny macierzy  $(A_d - G_d C)$  był równy (4.44) to z tw. Hamiltona-Cayleya mamy

$$(A_d - G_d C)^n = 0\tag{4.45}$$

Jeżeli zaś podstawimy (4.45) do (4.43) uzyskamy

$$\varepsilon(h \cdot n) = 0, \quad \forall \varepsilon(0)\tag{4.46}$$

Oznacza to, że obserwator dyskretno ciągły (4.40) zaprojektowany jako dead-beat po  $n$  krokach (pomiarach) dokładnie odtwarza stan systemu ciągłego (4.1) z dyskretnie mierzonym wyjściem (4.2).

Jak się wydaje, po raz pierwszy dyskretny obserwator dyskretno-ciągły do ciągłego odtwarzania stanu z dyskretnych danych pomiarowych zastosowano w pracy (Mitkowski i Baranowski, 2007).

#### 4.2.2. Dyskretny obserwator dokładny

Inne podejście do zadania estymacji stanu opiera się na rozwiązywaniu algebraicznych równań liniowych (zob. Byrski, 1993; Byrski, 2007, str. 418). Idea takiego podejścia jest następująca, dla ustalenia uwagi niech  $C \in \mathbb{R}^{1 \times n}$ . Rozważmy rozwiązanie równania (4.1) na odcinku  $t \in [s, T]$ , przy czym  $T$  jest aktualną chwilą, w której interesuje nas wartość stanu

$$x(T) = e^{A(T-s)}x(s) + \int_s^T e^{A(T-\tau)}Bu(\tau)d\tau\tag{4.47}$$

Dokonamy teraz pewnych przekształceń równości (4.47)

$$e^{A(s-T)}x(T) = x(s) + \int_s^T e^{A(s-T)}e^{A(T-\tau)}Bu(\tau)d\tau\tag{4.48}$$

$$e^{A(s-T)}x(T) = x(s) + \int_s^T e^{A(s-\tau)}Bu(\tau)d\tau\tag{4.49}$$

$$Ce^{A(s-T)}x(T) = Cx(s) + C \int_s^T e^{A(s-\tau)}Bu(\tau)d\tau\tag{4.50}$$

$$Ce^{A(s-T)}x(T) = y(s) + C \int_s^T e^{A(s-\tau)}Bu(\tau)d\tau\tag{4.51}$$

Należy zwrócić uwagę, że równość (4.51) wiąże nam wartość stanu w chwili  $T$  z wartością wyjścia w chwili  $s$  oraz sterowania na odcinku  $[s, T]$ . Z samej tej równości nie można wyznaczyć pożądanej estymaty. Rozważmy

teraz  $n$  pomiarów wyjścia w chwilach

$$0 \leq t_1 < t_2 < t_3 < \dots < t_n = T$$

i podstawmy je do wzoru (4.51) tworząc układ równań

$$\begin{aligned} Ce^{A(t_1-T)}x(T) &= y(t_1) + C \int_{t_1}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ Ce^{A(t_2-T)}x(T) &= y(t_2) + C \int_{t_2}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ Ce^{A(t_3-T)}x(T) &= y(t_3) + C \int_{t_3}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ &\vdots \\ Cx(T) &= y(T) \end{aligned} \quad (4.52)$$

Układ równań (4.52) można zapisać macierzowo (formułując go jednocześnie jako obserwator), t.j.

$$\begin{bmatrix} Ce^{A(t_1-T)} \\ Ce^{A(t_2-T)} \\ Ce^{A(t_3-T)} \\ \vdots \\ C \end{bmatrix} \hat{x}(T) = \begin{bmatrix} y(t_1) \\ y(t_2) \\ y(t_3) \\ \vdots \\ y(T) \end{bmatrix} + \begin{bmatrix} C \int_{t_1}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ C \int_{t_2}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ C \int_{t_3}^T e^{A(s-\tau)}Bu(\tau)d\tau \\ \vdots \\ 0 \end{bmatrix} \quad (4.53)$$

Rozwiązanie (4.53) określa nam obserwator dokładny, który w oparciu o pomiary wyjścia układu oraz sterowania pozwala nam wyliczyć dokładną wartość stanu w chwili  $T$ .

*Uwaga 4.4.* Warto zaznaczyć, że wyliczanie wpływu sterowania można nie wymaga wyliczania całek występujących w prawej stronie (4.53). Po pierwsze, można tego uniknąć dokonując autonomizacji zgodnie z (4.13). Wtedy obserwator redukuje się do postaci

$$\begin{bmatrix} Ce^{A(t_1-T)} \\ Ce^{A(t_2-T)} \\ Ce^{A(t_3-T)} \\ \vdots \\ C \end{bmatrix} \hat{x}_p(T) = \begin{bmatrix} y_p(t_1) \\ y_p(t_2) \\ y_p(t_3) \\ \vdots \\ y_p(T) \end{bmatrix} \quad (4.54)$$

$$\hat{x}(T) = \hat{x}_p(T) + x_u(T) \quad (4.55)$$

Innym sposobem uproszczenia postaci obserwatora (4.53) jest przeformułowanie go jako obserwatora warunku początkowego tj.  $x(t_1)$ . Wtedy dużo prościej powyliczać występujące całki, np. korzystając z informacji uzyskanej w kroku predykcji. ■

*Uwaga 4.5.* Warto zauważyć, że jeżeli  $t_i = i \cdot h$  to obliczenie macierzy  $\mathcal{M}$  można przeprowadzić raz, ponieważ wtedy

$$\mathcal{M} = \begin{bmatrix} CA_d^{n-1} \\ CA_d^{n-2} \\ \vdots \\ CA_d \\ C \end{bmatrix} \\ A_d = e^{Ah}$$

co więcej rozwiązywanie układu równań można zastąpić wtedy mnożeniem przez macierz odwrotną  $\mathcal{M}^{-1}$  (co

normalnie jest niezbyt opłacalne numerycznie), którą również można wyliczyć a priori. ■

Po raz pierwszy dyskretny obserwator dokładny do ciągłego odtwarzania stanu z dyskretnych danych pomiarowych zastosowano w pracy (Mitkowski i Baranowski, 2007). Zastosowanie do estymacji w systemie parabolicznym zobacz (Baranowski, 2007c) oraz przykład 4.3.

### 4.3. Estymacja z zakłóceniami wyjścia

W ramach niniejszego podrozdziału będziemy poszukiwać metod uodpornienia metod dokładnej estymacji stanu na występowanie zakłóceń w pomiarach wartości wyjścia systemu. Oznacza to, że zamiast wyjścia  $y(t)$  mierzymy

$$\tilde{y}(t) = y(t) + \psi(t) \quad (4.56)$$

W zależności od potrzeb wobec  $\psi$  będą postawione pewne założenia, np. do jakiej klasy funkcji będzie ono należeć.

#### 4.3.1. Obserwator najmniejszych kwadratów

Najbardziej naturalnym uodpornieniem obserwatora (4.54) na zakłócenia wyjścia jest obserwator najmniejszych kwadratów. Oczywiście jest, że posiadanie większej ilości informacji o wyjściu systemu pozwala na dokładniejszą charakteryzację jego stanu. Rozważmy więc system (4.16) z wyjściem (4.56) oraz  $N \geq n$  pomiarów wyjścia. Zauważmy, że związek pomiędzy zakłóconym wyjściem a stanem w każdej chwili pomiarowej wyraża się jako

$$\begin{bmatrix} \tilde{y}(t_1) \\ \tilde{y}(t_2) \\ \tilde{y}(t_3) \\ \vdots \\ \tilde{y}(t_N) \end{bmatrix} = \begin{bmatrix} Ce^{A(t_1-t_N)} \\ Ce^{A(t_2-t_N)} \\ Ce^{A(t_3-t_N)} \\ \vdots \\ C \end{bmatrix} x(t_N) + \begin{bmatrix} \psi(t_1) \\ \psi(t_2) \\ \psi(t_3) \\ \vdots \\ \psi(t_N) \end{bmatrix} \quad (4.57)$$

Dla uproszczenia zapiszmy powyższy układ równań jako

$$Y = Mx(t_N) + \Psi \quad (4.58)$$

Oznaczmy poszukiwaną estymatę jako  $\hat{x}$  oraz wprowadźmy residuum

$$\epsilon = Y - M\hat{x}(t_N) \quad (4.59)$$

Już K.F. Gauss zauważył (zob. Gauss, 1809; Simon, 2006, str. 80), że dla problemów tego typu najlepiej poszukiwać rozwiązania minimalizującego wyrażenie

$$J = \epsilon^\top \epsilon$$

Śledząc ten tok rozumowania (zob. np. Simon, 2006, str. 80 oraz Mitkowski, 2007b, str. 163) dostajemy

$$\begin{aligned} J &= (Y - M\hat{x}(t_N))^\top (Y - M\hat{x}(t_N)) = \\ &= Y^\top Y - \hat{x}(t_N)^\top M Y - Y^\top M \hat{x}(t_N) + \hat{x}(t_N)^\top M^\top M \hat{x}(t_N) \end{aligned}$$

W celu wyliczenia minimum bierzemy gradient

$$\nabla_{\hat{x}(t_N)} J = -2M^\top Y + 2M^\top M \hat{x}(t_N)$$

i przyrównujemy go do zera (ponieważ w problemie nie występują ograniczenia)

$$M^T Y = M^T M \hat{x}(t_N) \quad (4.60)$$

$$\hat{x}(t_N) = (M^T M)^{-1} M^T Y = M^L Y \quad (4.61)$$

Równość (4.61) określa nam obserwator najmniejszych kwadratów, przy czym  $M^L$  oznacza lewą pseudoodwrotność macierzy  $M$ . Warto zaznaczyć, że często numerycznie dogodniejsze od wyliczania odwrotności macierzy  $M^T M$  jest wyznaczanie estymaty stanu wprost z układu równań (4.60). Obserwator (4.61) w literaturze funkcjonuje też pod nazwą uproszczony obserwator minimalno-normowy (zob. Byrski, 2007, str. 419)

*Uwaga 4.6.* Jeżeli zastąpimy  $M^L$  przez macierz pseudoodwrotną  $M^\#$  (zob. Mitkowski, 2007b, str. 106, 114) obserwator (4.61) będzie funkcjonował również dla  $N < n$ , wtedy jednak układ równań (4.57), będzie układem niedostatecznie określonym i uzyskane  $\hat{x}(t_N)$  takim rozwiązaniem układu równań (4.57), którego norma euklidesowa jest najmniejsza. ■

*Uwaga 4.7.* Warto zaznaczyć, że podana wersja obserwatora najmniejszych kwadratów jest wersją najprostszą. Charakterystyki obserwatora można poprawić posiadając informacje na temat zakłócenia, np. jego wariancji. Można wtedy sformułować ważony problem najmniejszych kwadratów. Dodatkowo kwestia ilości pomiarów  $N$  pozostaje kwestią otwartą, teoretycznie nieskończona liczba pomiarów dałaby najlepszy wynik, jednak z oczywistych przyczyn nie jest to rozwiązanie praktyczne. ■

*Uwaga 4.8.* Obserwator najmniejszych kwadratów może być wyrażony również w sposób rekurencyjny (zob. np. Stoica i Söderström, 1997 oraz Simon, 2006, str. 84). Obserwator taki wykorzystuje wszystkie dostępne pomiary (por. obserwatory o rozszerzanym oknie np. Byrski, 2007, str. 410) jednak zamiast rozwiązywać coraz większe układy równań aktualizuje estymatę wraz z pojawieniem się nowych pomiarów. Warto zaznaczyć, że taki obserwator nadaje się zdecydowanie lepiej do odtwarzania warunku początkowego, ponieważ wtedy dużo łatwiej wykorzystywać pomocnicze obliczenia z wcześniejszych iteracji, np. autonomizację. Nie będzie on tutaj szczegółowo opisywany, ponieważ po pierwsze jest to algorytm powszechnie znany, a po drugie można go sprowadzić do szczególnego przypadku filtru Kalmana (zob. Stoica i Söderström, 1997, str. 373). ■

### 4.3.2. Obserwacja $C^0$ -optymalna

Rozważmy następujący system

$$\dot{x}(t) = Ax(t) \quad (4.62)$$

$$y(t) = Cx(t) \quad (4.63)$$

$$x(0) = x_0 \in \mathbb{R}^n \quad (4.64)$$

$$y(t) \in \mathbb{R} \quad (4.65)$$

przy czym para  $(C; A)$  jest obserwowalna.

W tym podrozdziale poszukujemy optymalnej metody estymacji stanu, zakładając, że system (4.62) z wyjściem (4.63) można interpretować jako

$$y = \mathcal{G}x_0 \quad (4.66)$$

przy czym

$$y \in Y = C^0[0, T] \quad (4.67)$$

$$\mathcal{G}: \mathbb{R}^n \rightarrow C^0[0, T] \quad (4.68)$$

Innymi słowy przyjmujemy, że wyjście systemu  $y$  jest ciągłą, skalarną funkcją czasu. Założenie to jest uzasadnione, ponieważ jest to bardzo ogólna klasa funkcji, do której mogą należeć również pewne zakłócenia pomiaru wyjścia. Nasze rozważania opieramy o następujące

**Twierdzenie 4.1** (Rolewicz). *Dany jest układ liniowy*

$$y = \mathcal{G}x_0 \quad (4.69)$$

$$x_0 \in \mathbb{R}^n \quad (4.70)$$

$$y \in Y = C^0[0, T, \mathbb{R}^m] \quad (4.71)$$

$$\mathcal{G}: \mathbb{R}^n \rightarrow C^0[0, T, \mathbb{R}^m] \quad (4.72)$$

Niech  $f$  będzie wybranym funkcjonałem liniowym ciągłym określonym na przestrzeni  $\mathbb{R}^n$ . Istnieje wtedy funkcjonał  $\varphi \in Y^*$  o minimalnej normie taki, że

$$fx_0 = \varphi \mathcal{G}x_0 \quad (4.73)$$

lub równoważnie

$$f = \mathcal{G}^* \varphi \quad (4.74)$$

Funkcjonał ten ma postać

$$\varphi(y) = \sum_{i=1}^n (v_i, y(t_i)) \quad (4.75)$$

przy czym,  $t_1, \dots, t_n$  są układem punktów odcinka  $[0, T]$  oraz  $v_1, \dots, v_n \in \mathbb{R}^m$ .  $Y^*$  jest przestrzenią wszystkich funkcjonałów liniowych ciągłych określonych na przestrzeni  $Y$  natomiast  $\mathcal{G}^*$  to operator dualny do  $\mathcal{G}$  zawężony do przestrzeni  $Y^*$ .

*Uwaga 4.9.* Wykorzystywane pojęcie operatora dualnego należy rozumieć w sensie definicji z pracy (Rolewicz, 1974, str. 133). Operator dualny do  $\mathcal{G}$ , oznaczany  $\mathcal{G}'$ , jest określony na przestrzeni  $Y'$  dualnej do  $Y$ , t.j. przestrzeni wszystkich funkcjonałów liniowych na przestrzeni  $Y$ . Przestrzeń funkcjonałów liniowych ciągłych  $Y^*$ , będąca podprzestrzenią przestrzeni dualnej  $Y'$ , bywa nazywana przestrzenią sprzężoną. Zawężenie operatora dualnego  $\mathcal{G}'$  do przestrzeni  $Y^*$ , oznaczane jako  $\mathcal{G}^*$ , bywa nazywane operatorem sprzężonym. Tego pojęcia nie należy jednak mylić z operatorem sprzężonym w przestrzeni Hilberta. ■

*Dowód.* Ponieważ dowód twierdzenia jest stosunkowo skomplikowany i znany w literaturze (Rolewicz, 1974, Twierdzenie VI.2.3, str. 277). Podamy tylko pewien szkic postępowania. Dobiera się liczbę  $a$  taką, że

$$a = \inf \{ \|\varphi\| : \varphi \in Y^*, f = \mathcal{G}^* \varphi \} \quad (4.76)$$

Poszukuje się punktów ekstremalnych  $\gamma_1, \dots, \gamma_n$  zbioru  $\Gamma_a \subset \mathbb{R}^n$  będącego odwzorowaniem kuli  $K_a^*$  o promieniu  $a$  w przestrzeni  $Y^*$  przez operator  $\mathcal{G}^*$ . Wykazuje się, że funkcjonał  $f$  można przedstawić jako kombinację punktów  $\gamma_1, \dots, \gamma_n$  z dodatnimi współczynnikami  $\alpha_i$  sumującymi się do jedności. Następnie pokazuje się, że można znaleźć układ funkcjonałów  $\varphi_1, \dots, \varphi_n$  w przestrzeni  $Y^*$ , będących punktami ekstremalnymi kuli  $K_a^*$ , które operator  $\mathcal{G}^*$  odwzorowuje w punkty  $\gamma_1, \dots, \gamma_n$ . Funkcjonał  $\varphi$  konstruuje się przez kombinację  $\varphi_1, \dots, \varphi_n$  ze współczynnikami  $\alpha_i$ .

Postać funkcjonału wynika z pewnych własności funkcjonałów w przestrzeniach funkcji o wahanii ograniczonym, które określają postać punktów ekstremalnych w przestrzeni  $Y^*$ . Konkretnie

$$\varphi_i(x) = \int_0^T (x(t), dg(t)) = (x(t_i), v_i) \quad (4.77)$$

przy czym  $x \in Y$ ,  $g_i$  jest stała na przedziałach  $[0, t_i]$  i  $(t_i, T]$  i ma skok w punkcie  $t_i$  równy  $v_i$ . □

Twierdzenie to można wykorzystać, aby skonstruować optymalny obserwator dla systemu (4.62) - (4.63). Dla skalarne go wyjścia mamy

$$\varphi(y) = \sum_{i=1}^n a_i^J y(t_i^J) \quad (4.78)$$



gdzie  $a_i^J$  dla  $i = 1, 2, 3, \dots, n$  są rzeczywistymi współczynnikami a  $t_i^J$  są chwilami czasu  $0 \leq t_1^J < t_2^J < \dots < t_n^J \leq T$ . Współczynniki  $a_i^J$  oraz chwile  $t_i^J$  należy interpretować jako  $i$ -te składowe wektorów  $a^J$  i  $t^J$ . Warto zaznaczyć, że wartości współczynników  $a_i^J$  są zależne od chwil  $t^J$ , pomijamy jednak to w zapisie, żeby nie zmniejszać czytelności.

Zauważmy, że

$$\varphi(y) = f x_0 \quad (4.79)$$

Funkcjonał  $\varphi$  służy więc za obserwator wielkości  $f \hat{x}(0)$ . Ponieważ  $\hat{x}(0) \in \mathbb{R}^n$ , to możemy zapisać

$$f \hat{x}(0) = (w, \hat{x}(0)) \quad (4.80)$$

gdzie  $w \in \mathbb{R}^n$  jest pewnym wektorem. W szczególności można wybrać taki funkcyjonał, że

$$f \hat{x}(0) = (e^J, \hat{x}(0)) = \hat{x}_J(0) \quad (4.81)$$

gdzie  $x_J$  oznacza  $J$ -tą współrzędną wektora stanu zaś  $e^J$  to  $J$ -ty wektor jednostkowy (wszystkie jego elementy poza  $J$ -tym są równe zero a  $J$ -ty jest równy 1). W tej interpretacji twierdzenia 4.1 dla optymalnej estymacji warunku początkowego jednej ze zmiennych stanu potrzeba  $n$  pomiarów. Aby wyznaczyć estymatę całego warunku początkowego w sposób optymalny należy dokonać  $n^2$  pomiarów. Innymi słowy możemy uzyskać obserwator postaci:

$$\hat{x}(0) = \begin{bmatrix} \hat{x}_1(0) \\ \hat{x}_2(0) \\ \vdots \\ \hat{x}_J(0) \\ \vdots \\ \hat{x}_n(0) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_i^1 y(t_i^1) \\ \sum_{i=1}^n a_i^2 y(t_i^2) \\ \vdots \\ \sum_{i=1}^n a_i^J y(t_i^J) \\ \vdots \\ \sum_{i=1}^n a_i^n y(t_i^n) \end{bmatrix} \quad (4.82)$$

*Uwaga 4.10.* Poszukiwanie funkcyjonału estymującego (4.78) o minimalnej normie jest istotne ze względu na potencjalną obecność zakłóceń w pomiarach. Rozważmy mierzony sygnał  $\tilde{y} \in C^0[0, T]$  taki, że

$$\tilde{y} = y + \psi \quad (4.83)$$

gdzie  $\psi \in C^0[0, T]$  jest zakłóceniem. Stosując funkcyjonał  $\varphi$  do  $\tilde{y}$  uzyskujemy

$$\varphi(\tilde{y}) = \varphi(y) + \varphi(\psi) \quad (4.84)$$

Występuje więc błąd estymacji  $\varphi(\psi)$ . Normę tego błędu można oszacować,

$$\|\varphi(\psi)\| \leq \|\varphi\| \cdot \|\psi\| \quad (4.85)$$

Z oszacowania (4.85) wynika, że im mniejsza jest norma  $\|\varphi\|$  tym mniejszy jest wpływ zakłóceń na estymację. Oczywiście jest, że przy braku zakłóceń estymacja jest dokładna. ■

Aby wyznaczyć funkcyjonał postaci (4.78) o minimalnej normie należy tę normę wyznaczyć. Posłuży do tego poniższy

**Lemat 4.2** (O normie funkcyjonału estymującego). *Rozważmy funkcyjonał liniowy  $f: C^0[0, T] \rightarrow \mathbb{R}$ , taki że*

$$f y = \sum_{i=1}^n a_i y(t_i) \quad (4.86)$$

wtedy

$$\|f\| = \sum_{i=1}^n |a_i| \quad (4.87)$$

*Dowód.* Wiadomo, że norma operacji wyraża się wzorem (zob. Rolewicz, 1974, str. 94, Kołodziej, 1982, str. 87)

$$\|f\| = \sup_{\|x\|_C=1} |fx| \quad (4.88)$$

Podstawiając (4.86) uzyskujemy

$$\|f\| = \sup_{\|x\|_C=1} \left| \sum_{i=1}^n a_i x(t_i) \right| \quad (4.89)$$

Norma w przestrzeni  $C^0[0, T]$  wyraża się wzorem

$$\|x\|_C = \sup_{t \in [0, T]} |x(t)| \quad (4.90)$$

Rozpatrzmy funkcję  $u$  taką, że

$$u(t) = \begin{cases} \text{sign}(a_i) & \text{for } t = t_i \\ \frac{(t - t_i)\text{sign}(a_{i+1}) - (t_{i+1} - t)\text{sign}(a_i)}{t_{i+1} - t_i} & \text{dla } t \in (t_i, t_{i+1}) \\ \text{sign}(a_1) & \text{dla } t \in [0, t_1) \\ \text{sign}(a_n) & \text{dla } t \in (t_n, T] \end{cases} \quad (4.91)$$

Funkcja ta należy do przestrzeni  $C^0[0, T]$  (jej wykresem jest łamana), zaś jej norma  $\|u\| = 1$ . Zauważmy, że w punktach  $t_i$  jest ona równa  $\text{sign} a_i$ , co oznacza, że dla  $u$  norma (4.88) przyjmuje wartość maksymalną, t.j.

$$\|f\| = \left| \sum_{i=1}^n |a_i| \right| \quad (4.92)$$

Co jest równoważne (4.87). □

Znając normę (4.78) należy określić jakie są współczynniki  $a_i^J$  funkcjonału estymującego o minimalnej normie. Poniższy lemat podaje ich wartości w zależności od systemu (4.62), wyjścia (4.63) oraz chwil pomiarów  $t_i^J$ .

**Lemat 4.3** (O współczynnikach funkcjonału estymującego). *Oznaczmy  $e^J$  jako  $J$ -ty wektor jednostkowy (wszystkie jego elementy poza  $J$ -tym są równe zero a  $J$ -ty jest równy 1),  $t^J$  oznacza wektor chwil czasu  $t_i^J$  zaś  $a^J$  jest wektorem kolumnowym współczynników  $a_i^J$  wtedy*

$$a^J = \Xi^{-1}(t^J)e^J \quad (4.93)$$

gdzie

$$\Xi(t^J) = \begin{bmatrix} C \exp(At_1^J) \\ \vdots \\ C \exp(At_n^J) \end{bmatrix}^T \quad (4.94)$$

Oznacza to, że dla danego systemu (4.62)-(4.63) współczynniki  $a_i^J$ , a więc i norma (4.87) funkcjonału (4.78) zależy od  $t^J$ . Poprzez właściwy dobór chwil czasu, w których dokonywane są pomiary wyjścia, można więc otrzymać funkcjonał o minimalnej normie. W ten sposób będziemy optymalizować dokonywane pomiary. Minimalizacja normy (4.78) w prostych przypadkach (np. system opisany równaniem  $\ddot{x}(t) = 0$ ) może zostać przeprowadzona analitycznie (zob. Rolewicz, 1974, str. 282; Mitkowski, 1996, str. 122) jednak w ogólnym przypadku jest to nieliniowy, niegładki problem optymalizacji z ograniczeniami bez gwarancji wypukłości.

## Optymalizacja

Norma (4.87) jest sumą wartości bezwzględnych. Jej minimalizacja jest to problem nieróżniczkowalny (nieładki), dlatego też, zamiast minimalizować normę wprost, będziemy rozwiązywać ciąg problemów optymalizacyjnych:

$$\min_{t^J \in \Omega} Q^\epsilon = \sum_i^n v_\epsilon(a_i^J) \quad (4.95)$$

$$\epsilon \rightarrow 0 \quad (4.96)$$

gdzie  $v_\epsilon(\xi)$  jest różniczkowalna,  $\forall \epsilon \forall \mathbb{R} \ni \xi, \neq 0 v_\epsilon(\xi) > 0$ ,  $v_\epsilon(0) = 0$  i gdy  $\epsilon \rightarrow 0$  to  $\forall \xi v_\epsilon(\xi) \rightarrow |\xi|$ . Jedną z funkcji o tej własności jest

$$v_\epsilon(\xi) = \begin{cases} \xi & \text{for } \xi \geq \epsilon \\ \xi^2 \left( \frac{2}{\epsilon} - \frac{2}{\epsilon^2 \xi} \right) & \text{for } \xi \in (0, \epsilon) \\ \xi^2 \left( \frac{2}{\epsilon} + \frac{2}{\epsilon^2 \xi} \right) & \text{for } \xi \in (-\epsilon, 0) \\ -\xi & \text{for } \xi \leq -\epsilon \end{cases} \quad (4.97)$$

Funkcja (4.97) oprócz własności wymienionych powyżej, posiada jeszcze tą cechę, że ma ściśle minimum w zerze. Dla uproszczenia, gdy będzie to potrzebne  $v_\epsilon$  z argumentem wektorowym oznaczać będzie wektor kolumnowy, którego elementami będą wartości  $v_\epsilon$  od odpowiednich elementów wektora argumentów. Kształt funkcji (4.97) dla różnych  $\epsilon$  ilustruje rysunek 4.1. Wskaźniki jakości w ciągu (4.95) są różniczkowalne, a ich gradienty względem  $t^J$  można wyznaczyć analitycznie, co ilustruje poniższy

**Lemat 4.4.** *Rozważmy wskaźnik jakości  $Q^\epsilon$  dla ustalonego  $\epsilon$  wtedy*

$$\nabla_{t^J} Q^\epsilon = \mathcal{J}(a^J)^\top v'_\epsilon(a^J) \quad (4.98)$$

gdzie

$$\mathcal{J}(a^J)^\top = -\text{diag}(a^J) \Xi(t^J)^\top A \left( \Xi(t^J)^{-1} \right)^\top \quad (4.99)$$

Przy czym  $\text{diag}(a^J)$  oznacza macierz diagonalną, której diagonalę tworzą elementy wektora  $a^J$  zaś  $v'_\epsilon$  oznacza pochodną  $v_\epsilon$ .

Wykorzystując analityczną formułę na gradient  $\nabla_{t^J} Q^\epsilon$  można łatwo zastosować różne algorytmy gradientowe. Należy jednak zanać, że wszystkie problemy w ciągu (4.95) posiadają ograniczenia tworzące wspólny dla wszystkich zbiór dopuszczalny

$$\Omega = \left\{ t^J \in \mathbb{R}^n : 0 \leq t_1^J < t_2^J < \dots < t_n^J \leq T \right\} \quad (4.100)$$

Warto zauważyć, że zazwyczaj w problemach optymalizacyjnych rozważamy ograniczenia w postaci nierówności słabych, zaś w (4.100) wszystkie prócz dwóch są silne. Można je jednak przeformułować jako

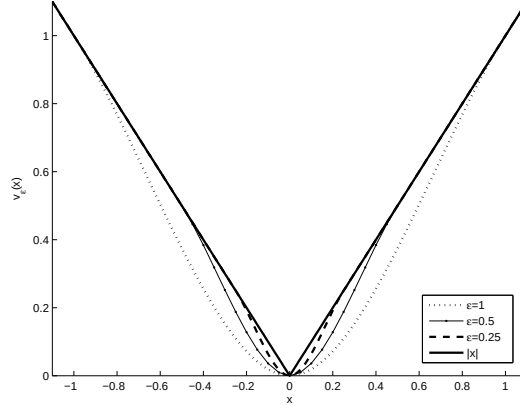
$$h_1(t^J) = t_1^J \geq 0 \quad (4.101)$$

$$h_2(t^J) = t_2^J - t_1^J \geq \kappa_2 \quad (4.102)$$

⋮

$$h_n(t^J) = t_n^J - t_{n-1}^J \geq \kappa_n \quad (4.103)$$

$$h_{n+1}(t^J) = T - t_n^J \geq 0 \quad (4.104)$$

Rysunek 4.1.: Przykłady  $v_\epsilon(x)$  dla różnych  $\epsilon$ 

gdzie  $\kappa_i > 0$  dla  $i = 2, 3, \dots, n$  są a priori przyjętymi, małymi parametrami. Analiza (4.95) oraz (4.98) pokazuje, że nie prowadzi to do utraty optymalności, ponieważ gdy ograniczenia stają się aktywne, a którekolwiek  $\kappa_i = 0$  współczynniki (4.93) nie istnieją zaś  $\|\nabla_{t^J} Q^\epsilon\| \rightarrow \infty$  gdy  $\kappa_i \rightarrow 0$ . Co więcej, eksperymenty numeryczne pokazują, że jeżeli któreś z ograniczeń  $h_i(t^J)$  for  $i \in \{2, 3, \dots, n\}$  staje się aktywne, jest to zazwyczaj minimum lokalne, które po zwiększeniu odpowiedniego  $\kappa_i$  jest opuszczane przez algorytm.

#### Zastosowanie obserwatora $C^0$ -optymalnego jako algorytmu korekcji

Warto zaznaczyć, że w pewnych sytuacjach optymalne pomiary, dla wszystkich zmiennych stanu, następują do chwili  $\tau = \max_J t_n^J$ , która znacznie poprzedza  $T$ . Można to zaobserwować w przykładzie (4.2). Dobrym rozwiązaniem jest wtedy przyjęcie  $T = \tau$  (skrócenie rozpatrywanego przedziału obserwacji).

W przypadku zastosowania obserwatora  $C^0$ -optymalnego do odtwarzania ciągłego stanu należy zwrócić uwagę, że uzyskane optymalne chwile dokonywania pomiarów niezwykle rzadko dają się „zazębiać” na kolejne okna (odstęp między pomiarami mogą być niewspółmierne). Oznacza to, że raz wykonane  $n^2$  pomiarów zazwyczaj nie nadaje się do powtórnego wykorzystania. Dlatego też należy rozpatrywać okna rozłączne (zob. definicja 2.12 na str. 25) mające ze sobą wspólne tylko jeden punkt pomiaru (punkt styku okien).

Dalszym ważnym aspektem jest to, że ponieważ obserwator projektowany jest dla systemów liniowych stacjonarnych optymalizację chwil pomiarów wystarczy przeprowadzić raz, a w stosunku do kolejnych okien dokonać jedynie przesunięcia ich w osi czasu o czas między pierwszym i ostatnim pomiarem. Co więcej, współczynniki obserwatora pozostaną te same. Ostatnim aspektem na jaki należy zwrócić uwagę to to, że dla pewnych systemów optymalne momenty dokonywania pomiarów mogą być identyczne dla różnych zmiennych stanu. Oznacza to, że liczba  $n^2$  pomiarów jest jedynie oszacowaniem górnym szerokości okna, zaś minimalna liczba potrzebnych pomiarów to  $n$ . Informację o ilości koniecznych pomiarów uzyskuje się poprzez przeprowadzenie optymalizacji.

#### 4.4. Estymacja z zakłóceniami wyjścia i wejścia

W ramach niniejszego podrozdziału rozważane są metody estymacji stanu uwzględniające występowanie zakłóceń zarówno w pomiarach wyjścia jak, i bezpośrednio wchodzących na system (przykładem takiego zakłócenia jest np. moment obciążający wał silnika). Podstawowym systemem tu rozpatrywanym jest

$$\dot{x}(t) = Ax(t) + Bu(t) + Lw(t) \quad (4.105)$$

$$y(t) = Cx(t) + M\psi(t) \quad (4.106)$$

Przy czym  $w$  i  $\psi$  są odpowiednimi zakłóceniami,  $L \in \mathbb{R}^{n \times p}$ ,  $M \in \mathbb{R}^{m \times s}$ , wyjście  $y$  i  $\psi$  określone są w równoodległych chwilach czasu  $t_i = ih$ .

#### 4.4.1. Dyskretno-ciągły filtr Kalmana

Dyskretno-ciągły filtr Kalmana (nazywany także hybrydowym filtrem Kalmana (zob. Simon, 2006, str. 403)) można traktować jako pewien wariant obserwatora dyskretno-ciągłego. Podobnie jak w nim, poszczególne elementy algorytmu można łączyć z krokami predykcji i korekcji.

Ze względu na stochastyczne uzasadnienie filtru, w ramach tego podrozdziału będziemy bazować na następującej charakterystyce zakłóceń. Niech  $w(t)$  jest ciągłym białym szumem o macierzy kowariancji  $Q$ , zaś  $\psi(kh)$  jest dyskretnym białym szumem o kowariancji  $R_k$ . Dodatkowo oprócz równania (4.105) rozpatrywać będziemy równanie propagacji

$$\dot{P}(t) = AP(t) + P(t)A^T + LQL^T \quad (4.107)$$

Równanie (4.105) oraz (4.107) rozwiązuje się na przedziałach  $[(k-1)h, h]$  z warunkami początkowymi  $\hat{x}^+((k-1)h)$  i  $P^+((k-1)h)$  (przyjmuje się  $w(t) \equiv 0$ ) zaś ich rozwiązania w chwili  $t = kh$  oznaczają się jako  $\hat{x}^-(kh)$  oraz  $P^-(kh)$ . Obliczenia te można wykorzystać w kroku predykcji, zwłaszcza że wtedy rozwiązuje się równanie (4.105).

W kroku korekcji należy dokonać następujących obliczeń

$$K(kh) = P^-(kh)C^T(CP^-(kh)C^T + MR_kM^T)^{-1} \quad (4.108)$$

$$\hat{x}^+(kh) = \hat{x}^-(kh) + K(kh)(y(kh) - C\hat{x}^-(kh)) \quad (4.109)$$

$$P^+(kh) = (I - K(kh)C)P^-(kh)(I - K(kh)C)^T + K(kh)MR_kM^TK(kh)^T \quad (4.110)$$

Uzasadnienia takiej konstrukcji filtru należy szukać w połączeniu ciągłego i dyskretnego filtru Kalmana ze skończonym horyzontem. Dyskusja optymalności filtru w sensie stochastycznym wykracza poza planowany zakres tej pracy, a co więcej można ją znaleźć w literaturze (zob. np. Simon, 2006, str. 123; Kwakernaak i Sivan, 1972; Mitkowski, 1996, str. 155).

#### 4.4.2. Inne podejścia

Bardzo interesujące podejście do estymacji stanu z dyskretnych pomiarów można znaleźć w (Byrski, 2007, str. 421). Analizuje się tam systemy z zakłóceniami w pomiarach wyjścia oraz w sterowaniu. Do tego celu stosuje się opis deterministyczny poszukując obserwatora o minimalnej normie operacji z przestrzeni wyjść i przestrzeni sterowań do przestrzeni stanów końcowych. Generalnie problem można sprowadzić do ogólnej teorii estymacji stanu w przestrzeniach Hilberta (Byrski, 2007, str. 392), przy czym przestrzeń wyjść to  $Y = \mathbb{R}^{N+1}$ , przestrzeń stanów końcowych  $X = \mathbb{R}^n$  zaś przestrzenią sterowań jest  $U = L^2([0, T], \mathbb{R})$ . W takiej sytuacji można podać formalnie formuły na optymalny obserwator, przy czym można dodatkowo uwzględnić wagę względem tego, które zakłócenia chcemy tłumić bardziej. Niestety zastosowanie tych formuł jest dość kłopotliwe ze względu na konieczność znajdowania operacji sprzężonych i odwrotnych do pewnych, dość skomplikowanych operacji splotowych działających  $L^2([0, T], \mathbb{R}) \rightarrow \mathbb{R}^{n+1}$ . Jeżeli jednak można założyć, że sterowanie jest stałe pomiędzy pomiarami, zaś pomiary są równoodległe (z krokiem  $h$ ) można zastąpić przestrzeń sterowań przez  $U = \mathbb{R}^{N+1}$ , i uzyskać bardzo praktyczne formuły na obserwator.

Konstrukcja obserwatora opiera się o zależność

$$y = H_1x(Nh) + H_2u \quad (4.111)$$

$$H_1 = \begin{bmatrix} Ce^{-ANh} \\ \vdots \\ Ce^{-A(N-i)h} \\ \vdots \\ C \end{bmatrix} \quad H_2 = \begin{bmatrix} -CB_d & -Ce^{-Ah}B_d & \dots & -Ce^{-A(N-1)h}B_d \\ 0 & -CB_d & \dots & -Ce^{-A(N-2)h}B_d \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -CB_d \end{bmatrix} \quad (4.112)$$

przy czym

$$y = \begin{bmatrix} y(0) & \dots & y(ih) & \dots & y(Nh) \end{bmatrix}^T \quad (4.113)$$

$$u = \begin{bmatrix} u(0) & \dots & u(ih) & \dots & u(Nh) \end{bmatrix}^T \quad (4.114)$$

$$B_d = \int_0^h e^{-A\tau} B d\tau \quad (4.115)$$

Zakładamy tutaj, że sterowanie jest prawostronnie ciągłą funkcją czasu. Obserwator optymalny, który dla tak zdefiniowanych przestrzeni i jednakowych wag dla obu rodzajów zakłóceń posiada minimalną normę ma postać

$$\hat{x}(Nh) = G_1 y + G_2 u \quad (4.116)$$

$$G_1 = \left( H_1^T F^{-1} H_1 \right)^{-1} H_1^T F^{-1} \quad (4.117)$$

$$G_2 = -G_1 H_2 \quad (4.118)$$

$$F = I + H_2 H_2^T \quad (4.119)$$

W przypadku sytuacji, w których spełnione są założenia co do sterowania, a pomiary są równoodległe obserwator ten jest bardzo dogodnym narzędziem do estymacji stanu, ponieważ wszystkie bardziej skomplikowane obliczenia wykonuje się a priori, co redukuje problem odtwarzania stanu do mnożenia znanych macierzy z wektorami pomiarów.

## 4.5. Przykłady zastosowań

Poniżej prezentujemy przykłady zastosowań algorytmu 1 do odtwarzania ciągłego stanu z dyskretnych danych pomiarowych w dwóch systemach - modelu populacyjnym cukrzycy oraz w systemie parabolicznym. Przykłady te pochodzą odpowiednio z prac (Baranowski, 2009a) i (Baranowski, 2007c).

**Przykład 4.2** (Model populacyjny cukrzycy - estymacja  $C^0$  optymalna). Konkluzją wielu prac badawczych, między innymi wspólnego programu WHO i Międzynarodowej Federacji Diabetologicznej (IDF), jest wniosek, że świat stoi przed rosnącą epidemią cukrzycy o potencjalnie zagrażających proporcjach (zob. Boutayeb i Chetouani, 2007). Szacuje się, że w skali światowej, liczba osób dotkniętych cukrzycą wynosiła ponad 171 milionów w 2000 roku, zaś do 2030 wyniesie 366 milionów, z czego duża część wystąpi w krajach rozwijających się. Cukrzyca jest znacznym obciążeniem dla systemów ochrony zdrowia na całym świecie. Przewiduje się, że w nadchodzących dekadach odbije się ona na osobach w wieku produkcyjnym, zwłaszcza w krajach o niskim i średnim przychodzie. Wspomniany program podkreśla, że ok. 3.2 miliona zgonów rocznie można przypisać powikłaniom wywołanym przez cukrzycę. Umieszcza to pod względem śmiertelności w jednym szeregu z HIV/AIDS. Niezależnie od tych niepokojących informacji problem ten jest generalnie mało znany. Wiadomo, że wielu powikłaniom cukrzycy można zapobiec lub kontrolować je o ile wystąpią, poprzez regularne badanie poziomu cukru i eliminację czynników ryzyka takich jak otyłość, palenie, brak ruchu lub nieprawidłowa dieta. Ponadto - co wydaje się nawet bardziej interesujące, wystąpienia cukrzycy można uniknąć, lub chociaż je spowolnić jeżeli ludzie w stanie prediabetycznym będą dobrze doinformowani. W 2003 roku liczba 314 milionów została podana, jako estymata liczby prediabetyków, z których co najmniej 30 % zapadnie na cukrzycę w przeciągu najbliższych 10 lat.

Jest to powodem wzrostu zainteresowania modelowaniem populacji cukrzyków. Model matematyczny można wykorzystać jako metodę analizy zjawisk zachodzących w tej populacji. Do takich i zbliżonych celów powstała specjalna gałąź biomatematyki o nazwie dynamika populacyjna (zob. Rudnicki, praca w przygotowaniu, materiał roboczy dostępny on-line). Ważne wyniki na temat modelowania populacji cukrzyków znaleźć można w pracach (Boutayeb, Twizell, Achouayb i Chetouani, 2004; Boutayeb i Chetouani, 2007). W ramach tego przykładu będziemy rozważać zastosowanie modelu Boutayeba, do estymacji ewolucji ilości ludzi w stanie prediabetycznym oraz cukrzyków bez powikłań w oparciu o liczbę chorych na cukrzycę, u których

rozwinęły się powikłania. Jest to naturalne podejście, ponieważ tacy właśnie ludzie zazwyczaj zgłaszają się do różnych placówek medycznych na własną rękę, więc określenie ich liczby będzie najprostsze.

Oznaczmy liczby prediabetyków, ludzi z cukrzycą bez powikłań oraz ludzi z powikłaniami przez odpowiednio  $x_1$ ,  $x_2$  i  $x_3$ . wtedy dla  $x(t) = [x_1(t) \ x_2(t) \ x_3(t)]^T$  mamy następujący model

$$\dot{x}(t) = Ax(t) + B \cdot I(t) \quad (4.120)$$

$$A = \begin{bmatrix} -\mu - \lambda_3 - \lambda_1 & 0 & 0 \\ \lambda_1 & -\mu - \lambda_2 & \gamma \\ \lambda_3 & \lambda_2 & -\theta \end{bmatrix} \quad (4.121)$$

$$B = [1 \ 0 \ 0]^T \quad (4.122)$$

Gdzie:

- $I(t)$  – częstotliwość występowania cukrzycy
- $\mu$  – naturalny współczynnik śmiertelności
- $\lambda_1$  – prawdopodobieństwo wystąpienia cukrzycy
- $\lambda_2$  – prawdopodobieństwo wystąpienia powikłań u osoby z cukrzycą
- $\lambda_3$  – prawdopodobieństwo wystąpienia cukrzycy od razu w stadium powikłań
- $\gamma$  – tempo leczenia powikłań
- $\nu$  – tempo w jakim pacjenci z powikłaniami poważnie podupadają na zdrowiu
- $\delta$  – śmiertelność wywołana powikłaniami
- $\theta$  –  $\gamma + \mu + \nu + \delta$

Model pochodzi z pracy (Boutayeb i Chetouani, 2007) i konstruuje się go jako tzw. model pojemnikowy (ang. compartmental model). Zakładamy, że w przewidywanym horyzoncie czasowym częstotliwość występowania cukrzycy nie będzie się zmieniać t.j  $I(t) = I = \text{const}$ . Przyjęto również, że współczynniki  $\nu$  i  $\delta$  odpowiadają równie prawdopodobnym kierunkom rozwoju choroby, więc przyjęto ich wartość na 5%. Tempo leczenia powikłań  $\gamma$  przyjęto na 8%, natomiast naturalny współczynnik śmiertelności  $\mu$  jako 2%. Zgodnie z założeniem o znajomości jedynie liczby pacjentów z powikłaniami przyjmujemy równanie wyjścia.

$$y(t) = Cx(t) = [0 \ 0 \ 1]x(t) \quad (4.123)$$

Ponieważ system (4.120) jest liniowy i ma jedno wyjście, jego obserwowalność można sprawdzić za pomocą standardowego kryterium obserwowalności, tzn

$$\det \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} \neq 0 \quad (4.124)$$

w tym przypadku, dla  $A$  i  $C$  danych odpowiednio przez (4.121) i (4.123) mamy

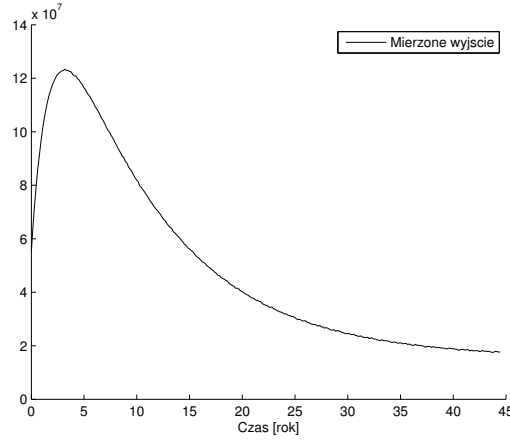
$$\det \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = -\lambda_3 \lambda_2^2 + \lambda_2 \lambda_3^2 + \lambda_2 \lambda_3 \lambda_1 - \lambda_2^2 \lambda_1 = (\lambda_3 - \lambda_2)(\lambda_3 + \lambda_1)\lambda_2 \quad (4.125)$$

Taka analiza prowadzi do konkluzji, że stan systemu (4.120) z wyjściem (4.123) jest obserwowalny wtedy i tylko wtedy, gdy  $\lambda_2 \neq \lambda_3$ . Interpretacja tego warunku jest taka, że prawdopodobieństwo wystąpienia powikłań u chorego na cukrzycę i wystąpienia cukrzycy od razu w stadium powikłań są różne, co można uznać za całkiem logiczne.

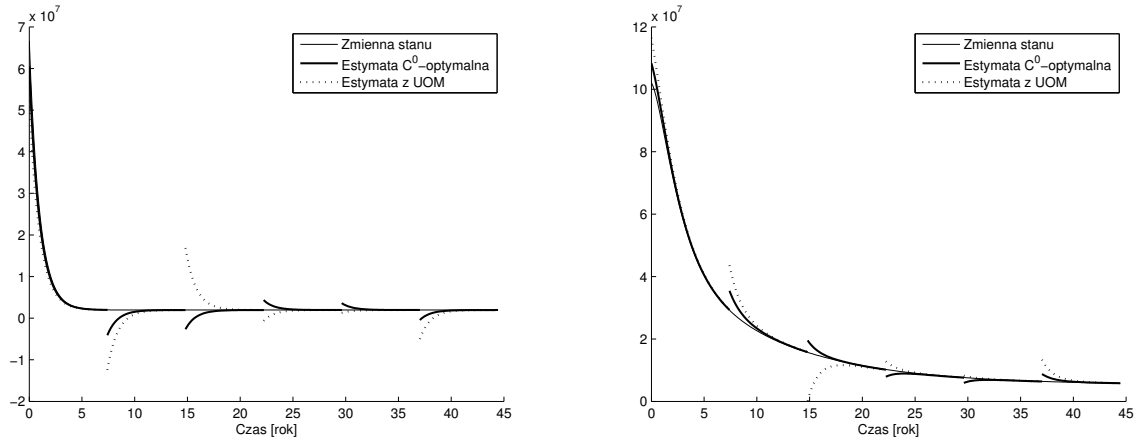
Do kroku korekcji podczas estymacji ciągłego stanu wykorzystano obserwator  $C^0$  optymalny. W tym celu przeprowadzono optymalizację czasu pomiarów zgodnie z (4.95). Zaobserwowano, że dla różnych  $T$  optymalne chwile pomiarów potrzebne do estymacji poszczególnych zmiennych stanu systemu (4.120) są sobie równe, oznacza to, że potrzebne są jedynie trzy pomiary na okno pomiarowe. Dodatkowo zaobserwowano, że zwiększenie  $T$  ponad 7.4069 roku nie prowadzi do zmian w chwilach pomiarów. Sugeruje to

więc optymalną wielkość okna pomiarowego. Optymalne chwile pomiarów dla okna o takiej długości to  $t_1^J = 0$ ,  $t_2^J = 1.1994$ ,  $t_3^J = 7.4069$  (dla  $J = 1, 2, 3$ ). Celem sprawdzenia efektywności obserwatora wyjście systemu zaburzano ciągłym sygnałem o amplitudzie odpowiadającej 1% wartości wyjścia. Mierzone wyjście ilustruje rysunek 4.2.

Ciągłe estymaty  $x_1$  i  $x_2$  zaprezentowano na rysunkach 4.3. Co jest interesujące, dla tego właśnie rozmiaru okna, obserwator  $C^0$  optymalny wprowadza zdecydowanie mniejsze błędy niż uproszczony obserwator minimalnonormowy (UOM, inaczej obserwator najmniejszych kwadratów). Zmieniając rozmiar okna obserwujemy zmianę tendencji, i przy krótkich oknach UOM sprawuje się lepiej. Przy bardzo krótkich oknach, tendencja ponownie się odwraca, jednak różnice są marginalne. Można wnioskować, że to względne podobieństwo bierze się stąd, że dla krótkich okien norma nawet optymalnego operatora obserwacji jest duża.



Rysunek 4.2.: Populacja diabetyków z powikłaniami



Rysunek 4.3.: Populacja ludzi w stanie prediabetycznym (z lewej) i diabetyków (z prawej) oraz ich estymaty

**Przykład 4.3** (Estymacja stanu w systemie parabolicznym). W tym przykładzie rozważamy system dynamiczny opisany jednowymiarowym równaniem przepływu ciepła

$$\frac{\delta T(t, z)}{\delta t} = \pi^{-2} \frac{\delta^2 T(t, z)}{\delta z^2} + b(z)u(t) \quad (4.126)$$



gdzie  $t \geq 0$ ,  $z \in (0, 1)$  z warunkami brzegowymi Neumanna

$$\left. \frac{\delta T(t, z)}{\delta z} \right|_{z=0} = \left. \frac{\delta T(t, z)}{\delta z} \right|_{z=1} = 0 \quad (4.127)$$

spełniające warunek początkowy

$$T(0, z) = x^0(z) \quad (4.128)$$

Jako wyjście z systemu rozważamy następujący funkcjonal

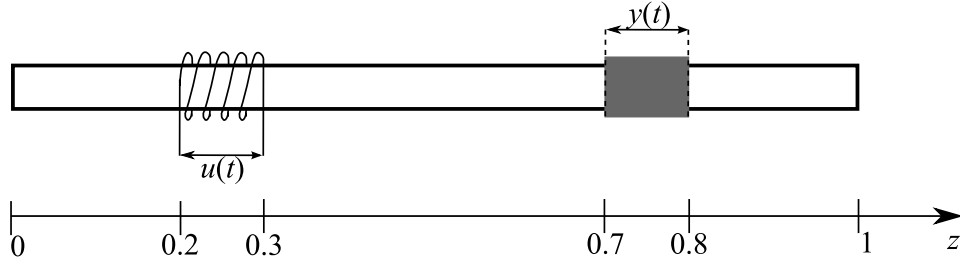
$$y_c(t) = \int_0^1 c(z)T(t, z)dz \quad (4.129)$$

Rozpatrujemy przypadek z rozłożonym sterowaniem i obserwacją, przy czym  $z \in [0, 1]$  oraz

$$b(z) = \begin{cases} \sqrt{10} & \text{dla } z \in [0.2, 0.3] \\ 0 & \text{dla } z \notin [0.2, 0.3] \end{cases} \quad (4.130)$$

$$c(z) = \begin{cases} \sqrt{10} & \text{dla } z \in [0.7, 0.8] \\ 0 & \text{dla } z \notin [0.7, 0.8] \end{cases} \quad (4.131)$$

Tak opisany system dynamiczny jest modelem systemu z rysunku 4.4. Systemy tego typu były rozpatrywane między innymi w (Mitkowski, 1991, str. 237) oraz (Oprzędkiewicz, 1997, 2008, str. 88), pomiar i sterowanie punktowe zobacz (Grabowski, 1997).



Rysunek 4.4.: Ilustracja rozważanego systemu parabolicznego

System taki można zamodelować również za pomocą następującego równania różniczkowego

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y_c(t) &= Cx(t) \end{aligned} \quad (4.132)$$

gdzie  $A$  jest generatorem infinitesimalnym półgrupy  $C^0$  w przestrzeni Hilberta  $X$ ,  $B$  jest liniowym operatorem ograniczonym  $\mathbb{R} \rightarrow X$ , zaś  $C$  jest liniowym operatorem ograniczonym  $X \rightarrow \mathbb{R}$ . Przestrzeń  $X$  przyjmujemy jako  $X = L^2([0, 1], \mathbb{R})$  zaś operatory możemy interpretować jako nieskończone macierze następującej postaci

$$\begin{aligned} A &= \text{diag}(\lambda_0, \lambda_1, \lambda_2, \dots) \\ B &= \begin{bmatrix} b_0 & b_1 & b_2 & \dots \end{bmatrix}^T \\ C &= \begin{bmatrix} c_0 & c_1 & c_2 & \dots \end{bmatrix} \end{aligned} \quad (4.133)$$

przy czym  $\lambda_i = -i^2$ ,  $b_i = (b, h_i)_2$ ,  $c_i = (c, h_i)_2$ ,  $i = 0, 1, 2, \dots$ , zaś  $h_i$  są wektorami (funkcjami) własnymi operatora  $A$  i są równe (zob. np. Mitkowski, 1991, str. 176 i 177)

$$h_i = \begin{cases} 1 & \text{dla } i = 0 \\ \sqrt{2} \cos(i\pi z) & \text{dla } i = 1, 2, \dots \end{cases} \quad (4.134)$$

Funkcje własne (4.134) tworzą ortonormalną bazę w przestrzeni  $X$ , zaś przyjmując  $x(t)$  jako nieskończony wektor, rozwiązanie równania (4.126) (lub równoważnie (4.132)) - czyli rozkład temperatury w funkcji czasu - przyjmuje postać

$$T(t, z) = \sum_{i=0}^{\infty} x_i(t) h_i(z) \quad (4.135)$$

Warto zwrócić uwagę, że można aproksymować rozkład temperatury systemu (4.126) zastępując sumę (4.135) sumą skończoną pierwszych kilku elementów - tworzymy w ten sposób aproksymację skończenie wymiarową. Taka aproksymacja może być uzasadniona jedynie w układach tłumionych, ponieważ wartości własne odpowiadające składowym wektora stanu o wysokich indeksach gwarantują bardzo silne tłumienie, co pozwala na ich zaniedbanie. Innymi słowy - w procesie estymacji będziemy rozpatrywać tylko pierwsze kilka równań różniczkowych systemu (4.132) z odpowiednio dobranym wyjściem. Formalna obserwowalność systemu (4.126) nie jest tutaj rozważana (zob. np. Kobayashi, 1980; Mitkowski, 1991, str. 197), natomiast skończenie wymiarowa aproksymacja jest obserwowalna. Przyjęto, że pomiary wyjścia  $y_c$  mają charakter dyskretny i są odległe o krok dyskretyzacji  $h = 0.5$ .

Do przeprowadzenia estymacji ciągłego stanu z dyskretnych danych pomiarowych wykorzystano procedurę korekcji obserwatora dyskretno-ciągłego (typu dead-beat) oraz dyskretnego obserwatora dokładnego. Przyjęto, że początkowy rozkład temperatury jest stały i równy zero, zaś układ jest pobudzany wejściem o charakterze impulsu prostokątnego postaci

$$u(t) = \begin{cases} 10 & \text{dla } t \in [1, 2] \\ 0 & \text{dla } t \notin [1, 2] \end{cases}$$

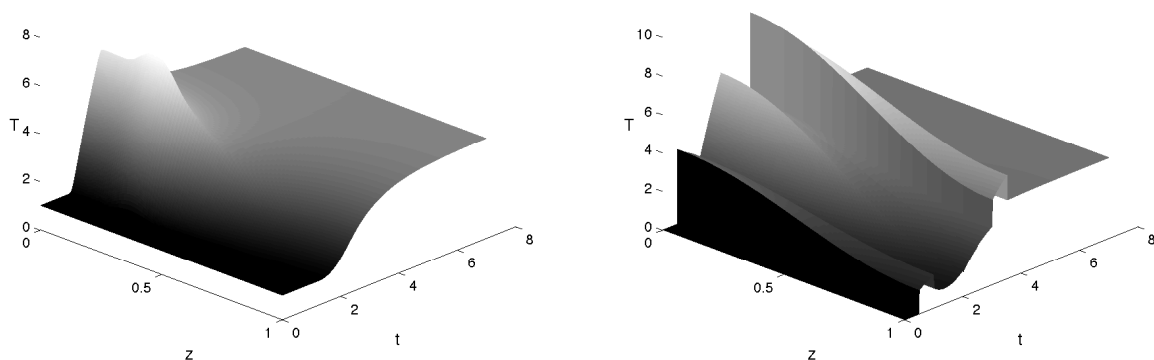
Jako sygnał pomiarowy wykorzystano wyjście z aproksymacji skończenie wymiarowej systemu (4.126) rzędu  $n = 10$ , natomiast estymator skonstruowano w oparciu o aproksymację rzędu  $n = 2$ . Rząd aproksymacji 10 wybrano ze względu na szybkość obliczeń. Co więcej dokonano następującego porównania, niech  $Y_{10}$ ,  $Y_{20}$  i  $Y_{40}$  oznaczają wektor pomiarów wyjścia na odcinku  $t \in [0, 15]$  uzyskany odpowiednio z aproksymacji rzędu 10, 20 i 40. Różnice między tymi aproksymacjami są następujące

$$|Y_{10} - Y_{20}| = 1.2 \cdot 10^{-3} \quad |Y_{40} - Y_{20}| = 4.1691 \cdot 10^{-6}$$

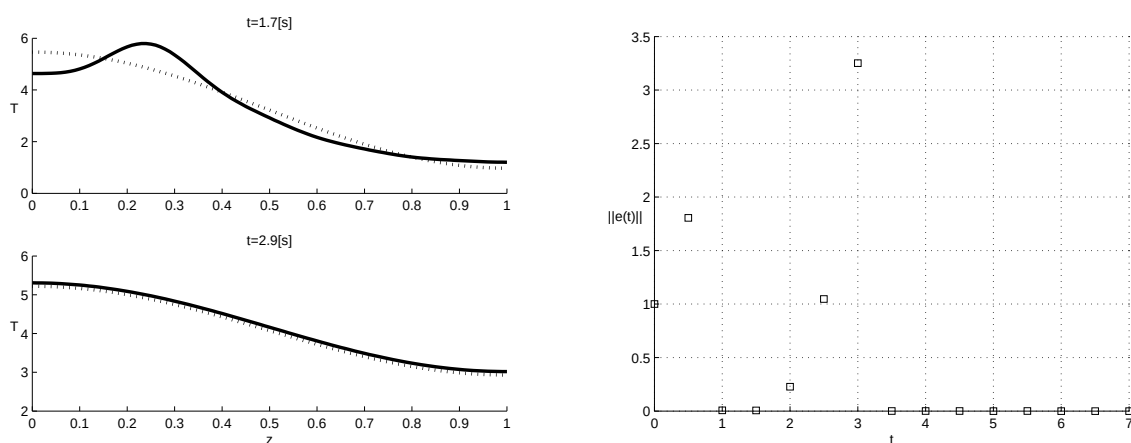
uznano więc, że aproksymacja rzędu 10 nie wprowadza istotnych błędów. Na rysunku (4.5) przedstawiono trójwymiarowe wykresy przedstawiające ewolucję rozkładu temperatury zachodzącej w systemie (aproksymacji wyższego rzędu), oraz uzyskane przez estymator. Jak można zauważyć, w początkowej fazie różnią się one dość znacznie. Wywołane jest to nieciągłością estymaty uzyskanej z obserwatora dyskretno-ciągłego. Jest to typowe dla zachowania obserwatorów typu dead-beat, w których przed całkowitym zanikiem błędu można zaobserwować duże oscylacje. Na rysunku 4.6 można zobaczyć porównania rozkładów rzeczywistego i estymowanego w dwóch reprezentatywnych chwilach czasu. Razem z nimi znajduje się wykres prezentujący zachowanie normy błędu. Jak widać w kilku krokach od zakończenia impulsu błąd estymacji spada praktycznie do zera. Brak całkowitego zaniku wywołany jest przez różnice między procesem a modelem wykorzystanym w obserwatorze.

Analogiczne wyniki dla dokładnego obserwatora dyskretnego ilustrują rysunki 4.7 i 4.8. Tutaj estymata jest gładzsza, zostaje zaburzona jedynie przy wprowadzeniu impulsu. Przez pierwsze dwa pomiary obserwator nie pracuje, ponieważ nie ma potrzebnych danych do odtworzenia stanu, i przyjęta estymata równa zero zgadza się z początkowym stanem systemu. Zaburzenia po wprowadzeniu impulsu wynikają ze wspomnianych wcześniej różnic między modelem wykorzystanym w estymacji a rozważanym procesem.

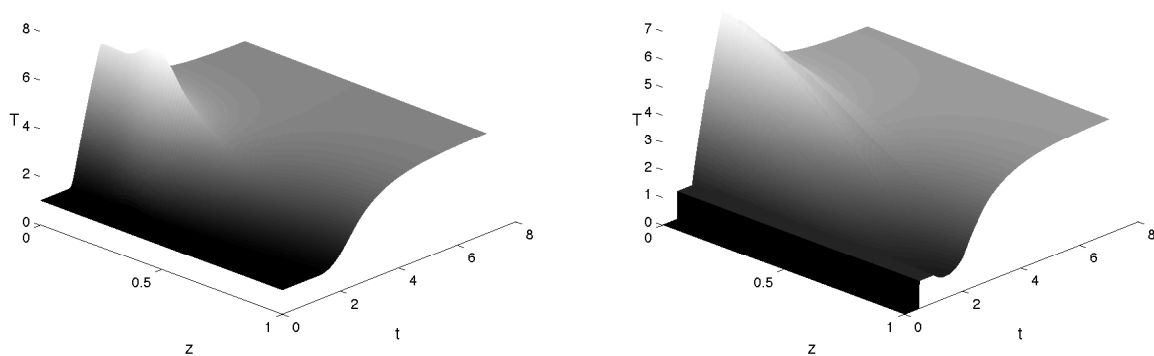
Wnioskiem wynikającym z przedstawionego przykładu jest fakt, że przy istnieniu dobrej aproksymacji można wykorzystać liniowe metody korekcji do estymacji ciągłego stanu również w przypadku procesów liniowych nieskończenie wymiarowych.



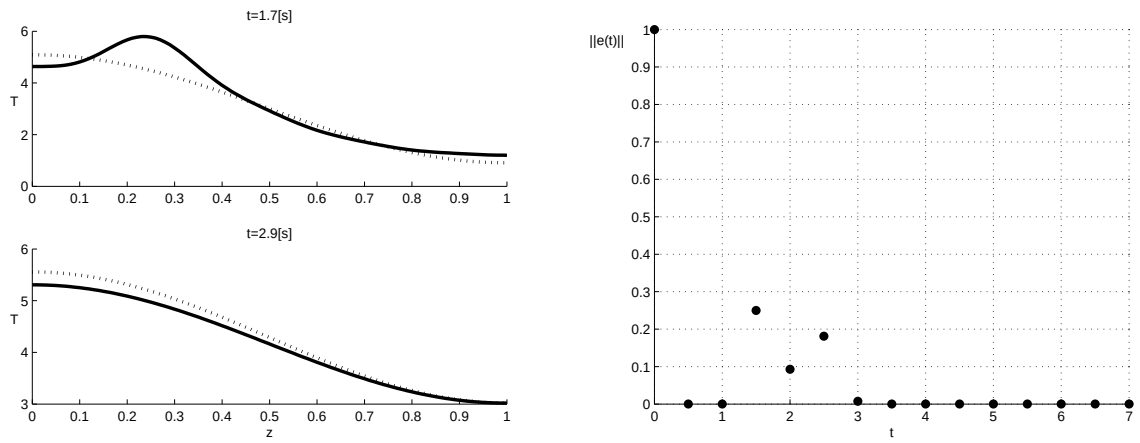
Rysunek 4.5.: Porównanie ewolucji temperatury (z lewej) oraz jej estymaty z obserwatora dyskretno-ciągłego (z prawej)



Rysunek 4.6.: Porównanie wartości chwilowych rozkładu temperatury i jego estymaty z obserwatora dyskretno-ciągłego (z lewej) oraz norma błędu estymacji w poszczególnych krokach korekcji (z prawej)



Rysunek 4.7.: Porównanie ewolucji temperatury (z lewej) oraz jej estymaty z dokładnego obserwatora dyskretnego (z prawej)



Rysunek 4.8.: Porównanie wartości chwilowych rozkładu temperatury i jego estymaty z dokładnego obserwatora dyskretnego (z lewej) oraz norma błędu estymacji w poszczególnych krokach korekcji (z prawej)

## Rozdział 5.

### Korekcja w układach nieliniowych

#### 5.1. Wprowadzenie

W ramach niniejszego podrozdziału rozważane będą systemy

$$\dot{x}(t) = f(x(t), u(t)) \quad (5.1)$$

z wyjściem

$$y(t) = h(x(t)) \quad (5.2)$$

W stosunku do (5.1) i (5.2) przyjmujemy te same założeniaco do (2.1) i (2.4) (por. str. 12). Dodatkowo zakładamy, że prawe strony (5.1) i (5.2) są różniczkowalne w sposób ciągły ze względu na swoje argumenty. Przyjmujemy, że wartości wyjścia znane są w dyskretnych chwilach czasu  $t_i$ , gdzie

$$t_i = t_{i-1} + h_i \quad (5.3)$$

$$t_0 \geq 0 \quad (5.4)$$

$$h_i > 0 \quad (5.5)$$

$$i = 0, 1, 2, 3, \dots \quad (5.6)$$

przy czym

$$h_i \geq \tau_{max} \quad (5.7)$$

$$\tau_{max} = \max_i (t_{i+1} - t_i) < \infty \quad (5.8)$$

Pomiary mogą, ale nie muszą być równoodległe. W takim przypadku mamy

$$h_i = h = \tau_{max} \quad \forall i \quad (5.9)$$

$$t_i = t_0 + ih \quad (5.10)$$

Ponieważ rozważane systemy są nieliniowe, ewentualne rozważania dotyczące obserwowalności będą dyskutowane w kontekście poszczególnych metod korekcji. Ogólne zakładamy jednak, że są one obserwowalne w sensie definicji 2.7 (zob. str. 19).

Podobnie jak w przypadku systemów liniowych wprowadzono rozdział algorytmów na te uwzględniające zakłócenia pomiarowe jak też i te bez zakłóceń. Ze względu na nieliniowość systemu, nie można rozważać autonomizacji, więc wszystkie wyniki sformułowane będą z uwzględnieniem sterowania.

*Uwaga 5.1* (Układy dodatnie). Ciekawą klasę układów nieliniowych stanowią tak zwane układy dodatnie. Same równania są liniowe, natomiast odpowiednie przestrzenie zmiennych nie są przestrzeniami liniowymi (zob. np. Mitkowski, 2000, str. 81). Rozważmy system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \quad (5.11)$$

gdzie  $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbb{R}^n$ ,  $u(t) = [u_1(t), u_2(t), \dots, u_r(t)]^T \in \mathbb{R}^r$  oraz  $y(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T \in \mathbb{R}^m$  zaś  $A$ ,  $B$  i  $C$  są rzeczywistymi macierzami o odpowiednich wymiarach.

System (5.11) jest układem dodatnim, jeżeli zachodzi następująca implikacja

$$\left\{ \begin{array}{l} x_i(0) \geq 0 \\ u_j(t) \geq 0 \\ i = 1, 2, \dots, n \\ j = 1, 2, \dots, r \\ t \geq 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x_i(t) \geq 0 \\ y_s(t) \geq 0 \\ i = 1, 2, \dots, n \\ s = 1, 2, \dots, m \\ t \geq 0 \end{array} \right. \quad (5.12)$$

System ciągły (5.11) jest dodatni wtedy i tylko wtedy, gdy  $A$  jest macierzą Metzlera (jej elementy poza diagonalnymi są nieujemne), a macierze  $B$  i  $C$  mają nieujemne elementy. Układy dodatnie mają wiele zastosowań, np. w obwodach elektrycznych, modelach ewolucji rozkładów prawdopodobieństwa lub dynamice populacyjnej (układ (4.120) z przykładu 4.2 na str. 54 jest układem dodatnim). Analogicznie formułuje się dyskretne układy dodatnie opisywane równaniami różnicowymi

$$\begin{aligned} x_{k+1} &= A_d x(k) + B_d u(k) \\ y(k) &= C x(k) \end{aligned} \quad (5.13)$$

Układ (5.13) jest dodatni wtedy, i tylko wtedy gdy  $A$ ,  $B$  i  $C$  mają nieujemne elementy. W dziedzinie układów dodatnich (również z opóźnieniami) istnieje wiele wyników dotyczących stabilności, osiągalności, sterowania minimalno-energetycznego oraz teorii realizacji (zob. np. Kaczorek, 2007, str. 425, 437, 442, 446) warto również zwrócić uwagę na wyniki dotyczące sterowalności (Busłowicz, 2005) oraz odpornej stabilności układów dodatnich z niepewnościami (Busłowicz, 2007).

Estymacją stanu w układach dodatnich zajmował się (Kociszewski, 2009, str. 144). W jego pracy, można znaleźć wyniki dotyczące konstrukcji dodatnich obserwatorów pełnego i zredukowanego rzędu dla układów dodatnich (także z opóźnieniami). Zaproponowana tam metodyka opiera się na stosowaniu liniowych nierówności macierzowych (ang. *LMI, linear matrix inequalities*), w oparciu o które sformułowano warunki istnienia oraz algorytm efektywnego wyznaczania wzmocnień obserwatora. Zdaniem autora, wyniki te mogą być z powodzeniem wykorzystane do estymacji ciągłego stanu z dyskretnych danych pomiarowych za pomocą obserwatora dyskretno-ciągłego. ■

## 5.2. Estymacja bez zakłóceń

W tym podrozdziale zaprezentujemy dwa podejścia, w których zakłócenia nie są uwzględniane podczas konstrukcji obserwatora. Warto jednak zaznaczyć, że wcale nie oznacza to tego, że nie mają one żadnych własności filtrujących. Omówiony zostanie nieliniowy obserwator dyskretno-ciągły, łączący pomysły stojące za ciągłym obserwatorem typu High-Gain oraz za dyskretno-ciągłym filtrem Kalmana. Drugim omówionym algorytmem odtwarzania stanu jest hybrydowy obserwator Newtona, w którym problem estymacji sprowadza się do rozwiązywania pewnego układu równań nieliniowych.

### 5.2.1. Nieliniowy obserwator dyskretno-ciągły

Rozważany tu algorytm oparty jest na pracach (Astorga et al., 2002; Angeles et al., 2004), w których rozważano odtwarzanie stanu w pewnych konkretnych procesach chemicznych. Proponowane tam podejście może być jednak stosowane także w układach innego rodzaju. Na początek przyjmujemy, że pomiary są równoodległe z krokiem dyskretyzacji  $h > 0$ .

Pomysł za nieliniowym obserwatorem dyskretno-ciągłym opiera się o zastosowanie metodyki pochodnych Liego. Podstawowym założeniem wobec systemu, jest to aby był on lokalnie obserwowalny w całej przestrzeni stanu w sensie poniższego twierdzenia.

**Twierdzenie 5.1** (zob. Nijmeijer i van der Schaft, 1991, Keller, 1987, Dalla Mora, Germani i Manes, 1997).  
Dany jest system (5.1) z wyjściem (5.2). Rozważana jest funkcja wektorowa:

$$\Phi(x) = \begin{bmatrix} h(x) \\ \mathcal{L}_f h(x) \\ \vdots \\ \mathcal{L}_f^{n-1} h(x) \end{bmatrix} \quad (5.14)$$

gdzie  $\mathcal{L}_f h(x)$  oznacza pochodną Liego funkcji  $h$  w kierunku pola wektorowego  $f$  zdefiniowaną jako

$$\mathcal{L}_f^0 h(x) = h(x) \quad (5.15)$$

$$\mathcal{L}_f h(x) = \mathcal{L}_f^1 h(x) = \nabla h(x)^\top f(x) \quad (5.16)$$

$$\mathcal{L}_f^s h(x) = \mathcal{L}_f (\mathcal{L}_f^{s-1} h(x)) \quad (5.17)$$

Dla każdego  $x^*$  takiego, że

$$\text{rank} Q(x^*) = n$$

System system (5.1) z wyjściem (5.2) jest lokalnie obserwowalny w punkcie  $x^*$ . Przy czym

$$Q(x) = \frac{d\Phi(x)}{dx} \quad (5.18)$$

W kontekście nieliniowego obserwatora dyskretno-ciągłego rozważać będziemy system

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), & f: \mathbb{R}^n \times \mathbb{R}^r &\rightarrow \mathbb{R}^n \\ y(t) &= h(x(t)), & h: \mathbb{R}^n &\rightarrow \mathbb{R} \end{aligned} \quad (5.19)$$

dla którego spełniony jest warunek lokalnej obserwowalności w każdym punkcie przestrzeni stanu. Dla takiego systemu macierz  $Q(x)$  jest kwadratową macierzą nieosobliwą. Do konstrukcji nieliniowego obserwatora dyskretno-ciągłego stosujemy podobny schemat postępowania jak przy obserwatorze typu High Gain (zob. Hou, 2002; Baranowski i Piątek, 2008). Zauważmy, że po dokonaniu następującej zamiany zmiennych w systemie (5.19)

$$z(t) = \Phi(x(t)) \quad (5.20)$$

uzyskujemy system

$$\begin{aligned} \dot{z}(t) &= Q(\Phi^{-1}(z(t))) f(\Phi^{-1}(z(t)), u(t)) = A_b z(t) + B_b \xi(z(t), u(t)) \\ y(t) &= C_b z(t) \end{aligned} \quad (5.21)$$

przy czym macierze  $A_b$ ,  $B_b$  oraz  $C_b$  (wykorzystywana dalej) są tak zwanymi macierzami Brunowskiego

$$A_b = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad (5.22)$$

$$B_b = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}^\top \quad (5.23)$$

$$C_b = \begin{bmatrix} 1 & \cdots & 0 & 0 \end{bmatrix} \quad (5.24)$$

zaś funkcja  $\xi$  jest pewną nieliniową skalarną funkcją transformowanego stanu  $z$  oraz sterowania i ewentualnie jego pochodnych. Konstrukcja obserwatora typu High-Gain bazuje na transformacji (5.20) i powstałym systemie (5.21), dla której konstrukcja obserwatora jest stosunkowo prosta. Idea obserwatora High-Gain polega na tym, że przy pomocy odpowiednio dużego liniowego wzmocnienia można będzie skompensować

nieliniowość  $\xi$  (w systemach z globalnym warunkiem Lipschitza, jest to uzasadnione, por. tw. A.6 na stronie 99). Obserwator w układzie transformowanym ma postać

$$\dot{\hat{z}}(t) = A_b \hat{z}(t) + B_b \xi(\hat{z}(t), u(t)) + G(y(t) - C_b \hat{z}(t)) \quad (5.25)$$

przy czym  $G$  dobrano tak, aby  $\text{Re } \lambda(A_b - GC_b) < 0$ . Co więcej dokonywanie transformacji explicite nie jest potrzebne, ponieważ obserwator (5.25) po powrocie do oryginalnych współrzędnych przyjmuje postać

$$\dot{\hat{x}} = f(\hat{x}, u) + Q^{-1}(\hat{x})G(y - h(\hat{x})) \quad (5.26)$$

Jak można zauważyć do działania takiego obserwatora niezbędna jest lokalna obserwowalność w podanym powyżej sensie (macierz  $Q$  musi być odwracalna).

*Uwaga 5.2.* Warto zauważyć, że warunek obserwowalności z tw. 5.1 zastosowany w układzie liniowym przechodzi w standardowy warunek obserwowalności pary  $(C; A)$  (zob. tw. A.2, punkt 3, str. 98). Natomiast zastosowanie zamiany zmiennych (5.20) prowadzi do transformacji do postaci kanonicznej Frobeniusa (zob. Ackermann, 1976, str. 71 oraz Mitkowski, 2007b, str. 88). ■

Opisany powyżej schemat postępowania wykorzystuje się w konstrukcji nieliniowego obserwatora dyskretno-ciągłego. Podobnie, jak w opisanym w poprzednim rozdziale filtrze Kalmana mamy bezpośrednią implementację predykcji i korekcji.

W kroku predykcji dla  $t \in [kh, (k+1)h)$  rozwiązuje się następujące równania różniczkowe

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (5.27)$$

$$\dot{S}(t) = -\theta S(t) - A_b^T S(t) - S(t)A_b \quad (5.28)$$

z warunkami początkowymi  $\hat{x}^+(kh)$  i  $S^+(kh)$ , przy czym  $A_b$  jest dane przez (5.22), zaś  $\theta > 0$  jest parametrem strojącym obserwatora. Rozwiązania równań (5.27) i (5.28) w chwili  $t = (k+1)h$  oznacza się jako  $\hat{x}^-((k+1)h)$  oraz  $S^-((k+1)h)$ . Pierwsze równanie jest równaniem modelu, zaś drugie można uznać za pewną modyfikację równania propagacji (4.107). Ponieważ zakładamy brak zakłóceń, człon za nie odpowiadający znika, zaś ponieważ wykorzystujemy transformację zmiennych, dostajemy bardzo prostą macierz Brunowskiego  $A_b$ . Ze względu na to, że nie są one asymptotycznie stabilne dodaje się dodatkowo człon  $\theta S$ , który z jednej strony ma zapewnić stabilność równania, a z drugiej strony wprowadza dodatkowy parametr strojący  $\theta$  umożliwiający poprawienie działania obserwatora w oparciu o eksperymenty symulacyjne dla konkretnej dynamiki procesu.

W kroku korekcji przeprowadza się następujące operacje algebraiczne

$$S^+((k+1)h) = S^-((k+1)h) + hC_b^T C_b \quad (5.29)$$

$$\hat{x}^+((k+1)h) = \hat{x}^-((k+1)h) + \kappa Q \left( \hat{x}^-((k+1)h) \right)^{-1} S^+(t_{k+1})^{-1} C^T \left( y((k+1)h) - h \left( \hat{x}^-((k+1)h) \right) \right) \quad (5.30)$$

gdzie  $\kappa > 0$  jest drugim parametrem strojącym a  $C_b$  jest dane wzorem (5.24). Tutaj można zauważyć większe różnice w stosunku do filtru Kalmana, aktualizacja macierzy  $S$  jest względnie nieskomplikowana, jednak równanie (5.30) wymaga pewnego komentarza. Można tu zauważyć pewną analogię z obserwatorem High-Gain, występuje tutaj błąd wyjścia a także macierz  $Q^{-1}$  pojawiająca się przy przejściu od zmiennych transformowanych do pierwotnych (por. (5.26)). Macierz wzmocnienia natomiast ma postać  $S^+(t_{k+1})^{-1} C_b^T$  co oznacza, że potencjalnie zmienia się ona z każdym krokiem. Dodatkową zmianę wprowadza tu parametr  $\kappa$ , który podobnie jak  $\theta$ , można dobrać eksperymentalnie do dynamiki systemu, tak aby poprawić działanie obserwatora.

Nieliniowy obserwator dyskretno-ciągły był stosowany do estymacji stanu układu zbiorników wodnych w pracach (Baranowski i Tutaj, 2007; Baranowski, 2007d).



### 5.2.2. Hybrydowy obserwator Newtona

Pod nazwą obserwator Newtona, rozumie się metodę odtwarzania warunku początkowego równania różniczkowego na drodze rozwiązania układu równań algebraicznych metodą Newtona. Koncepcja obserwatora Newtona została zaproponowana przez (Moraal i Grizzle, 1995) dla systemów dyskretnych postaci

$$\begin{aligned} x_{k+1} &= f(x_k, u_k), & f: \mathbb{R}^n \times \mathbb{R}^r &\rightarrow \mathbb{R}^n \\ y_k &= h(x_k), & h: \mathbb{R}^n &\rightarrow \mathbb{R} \\ k &= 0, 1, 2, 3, \dots \end{aligned} \quad (5.31)$$

Dla ustalenia uwagi zdefiniujemy

$$\begin{aligned} \mathfrak{F}^1(x_k) &= f(x_k, u_k) \\ \mathfrak{F}^n(x_k) &= f(\mathfrak{F}^{n-1}(x_k), u_{k+n-1}) \end{aligned}$$

możemy teraz zapisać

$$\begin{aligned} y_0 &= h(x_0) \\ y_1 &= h(\mathfrak{F}^1(x_0)) \\ &\vdots \\ y_{n-1} &= h(\mathfrak{F}^{n-1}(x_0)) \end{aligned} \quad (5.32)$$

Niech

$$\mathfrak{Y} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \quad \text{oraz} \quad \mathfrak{H}(x_0) = \begin{bmatrix} h(x_0) \\ h(\mathfrak{F}^1(x_0)) \\ \vdots \\ h(\mathfrak{F}^{n-1}(x_0)) \end{bmatrix} \quad (5.33)$$

Jak można zauważyć (5.32) tworzy układ  $n$  nieliniowych równań algebraicznych ze względu na  $x_0$ . Po przeniesieniu wyrazów (5.32) na lewą stronę i stosując zapis (5.33) mamy

$$\mathfrak{Y} - \mathfrak{H}(x_0) = 0 \quad (5.34)$$

Taki układ równań można rozwiązać za pomocą metody Newtona dla układów równań nieliniowych (zob. np. Stoer, 1979, str. 183)

$$\begin{aligned} x_0^{i+1} &= x_0^i + s_i \\ \frac{d}{dx_0} \mathfrak{H}(x_0^i) s_i &= \mathfrak{Y} - \mathfrak{H}(x_0^i) \\ i &= 0, 1, \dots, d \end{aligned} \quad (5.35)$$

gdzie  $x_0^i$  oznacza  $i$ -te przybliżenie  $x_0$  zaś  $d$  to liczba iteracji. Liczbę iteracji zazwyczaj dobiera się na drodze symulacyjnej. Taki sposób wyznaczania stanu początkowego został nazwany obserwatorem Newtona. Jest to metoda efektywna i jedyną faktyczną komplikacją jest wyznaczenie macierzy Jakobiego wielokrotnie złożonej funkcji wektorowej  $\mathfrak{H}(x_0)$ . O ile jednak w przypadku systemów dyskretnych jest możliwe sformułowanie układu równań algebraicznych, tak w przypadku systemów ciągłych nie ma takiej możliwości. Autorzy pracy (Biyik i Arcak, 2006b) zaproponowali, aby stosowny układ równań wyznaczyć za pomocą symulacji a następnie zastosować do niego metodę Newtona - tak powstał hybrydowy obserwator Newtona (HNO).

Dla ustalenia uwagi przyjmijmy

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t)) \end{aligned} \quad (5.36)$$

Przy czym  $t \in [t_k, t_{k+n-1}]$ ,  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$ ,  $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $h: \mathbb{R}^n \rightarrow \mathbb{R}$ . Dodatkowo zakładamy, że  $f$  i  $h$  są przynajmniej jednokrotnie różniczkowalne w sposób ciągły ze względu na każdą ze składowych  $x$ . Koncepcja HNO opiera się na oczywistym fakcie, że rozwiązanie (5.36), dla warunku początkowego  $x(t_k) =$

$w_k$ , przy spełnieniu założeń twierdzeń o istnieniu i jednoznaczności rozwiązań, ma ogólnie postać

$$x(t) = F(w_k, t) \quad (5.37)$$

gdzie  $F(\cdot, \cdot)$  jest pewną funkcją  $\mathbb{R}^n \times [t_k, t_{k+n-1}] \rightarrow \mathbb{R}^n$  spełniającą równanie różniczkowe systemu (5.36) (w zależności od regularności sterowania wszędzie, lub prawie wszędzie). Wyjście z tego systemu ma zatem postać

$$y(t) = h(F(w_k, t)) \quad (5.38)$$

Znając wartości wyjścia systemu, możemy zbudować układ równań analogiczny do (5.32)

$$\begin{aligned} y(t_k) &= h(F(w_k, t_k)) \\ y(t_{k+1}) &= h(F(w_k, t_{k+1})) \\ &\vdots \\ y(t_{k+n-1}) &= h(F(w_k, t_{k+n-1})) \end{aligned} \quad (5.39)$$

wprowadźmy oznaczenia

$$Y_k = \begin{bmatrix} y(t_k) \\ y(t_{k+1}) \\ \vdots \\ y(t_{k+n-1}) \end{bmatrix} \quad \text{oraz} \quad H(w_k) = \begin{bmatrix} h(w_k) \\ h(F(w_k, t_{k+1})) \\ \vdots \\ h(F(w_k, t_{k+n-1})) \end{bmatrix} \quad (5.40)$$

Wektor  $Y_k$  znany z pomiarów, natomiast dla danego  $w_k$  możemy wyliczyć wektor  $H(w_k)$  za pomocą numerycznego rozwiązania systemu (5.36). Można zastosować metodę Newtona

$$\begin{aligned} w_k^{i+1} &= w_k^i + s_i \\ \frac{d}{dw_k} H(w_k^i) s_i &= Y_k - H(w_k) \\ i &= 0, 1, \dots, d \end{aligned} \quad (5.41)$$

Analogicznie  $w_k^i$  to  $i$ -te przybliżenie stanu systemu w chwili  $t_k$ , zaś  $d$  to liczba iteracji. Przyjmuje się ustaloną ich liczbę zamiast iteracji do zbieżności, ponieważ w stosunku do obserwatorów często stawiane są wymagania co do czasu obliczeń, którego bez górnego ograniczenia nie dałoby się zapewnić. Podobnie jak w przypadku dyskretnym konieczne jest dobranie jej eksperymentalnie w zależności od systemu.

Podstawowa różnica pomiędzy problemem dyskretnym (5.32) i ciągłym (5.39) to sposób wyliczania macierzy Jakobiego potrzebnej do wyznaczania  $s^i$  (por. 5.41). W dalszej części omawiamy trzy takie podejścia.

### Wyliczanie macierzy Jakobiego

W ogólności, aby wyliczyć macierz Jakobiego należy wyznaczyć gradienty pewnych  $n$  skalnych funkcji rozwiązania równania różniczkowego (elementów macierzy  $H(w_k)$  względem warunku początkowego. Poniżej omawiamy trzy sposoby jej obliczania (zob. Baranowski i Mitkowski, 2009; Baranowski, 2009b)

- Aproksymację różnicową,
- Zastosowanie równania wariacyjnego,
- Wariację funkcjonału.

### Aproksymacja różnicowa

Najprostszy sposób wyliczania macierzy Jakobiego, jednocześnie tym proponowanym przez autorów metody (zob. Bıyık i Arca, 2006b,a) jest aproksymacja za pomocą różnic skończonych. Polega to na zastąpieniu

różniczkowania po każdej ze składowych warunku początkowego przez odpowiedni iloraz różnicowy

$$\frac{d}{dw_k} H(w_k) \approx \begin{bmatrix} \left( \frac{H(w_k^i + e_1 \rho) - H(w_k^i)}{\rho} \right)^T \\ \left( \frac{H(w_k^i + e_2 \rho) - H(w_k^i)}{\rho} \right)^T \\ \vdots \\ \left( \frac{H(w_k^i + e_n \rho) - H(w_k^i)}{\rho} \right)^T \end{bmatrix}^T \quad (5.42)$$

gdzie  $\rho$  jest małym parametrem określającym przyrost, zaś  $e_j$  jest  $j$ -tym wektorem jednostkowym. Wyliczanie poszczególnych kolumn macierzy Jakobiego polega więc na wyliczeniu wektorów  $H(w_k^i + e_j \rho)$  poprzez symulacje z zaburzonym warunkiem początkowym, a następnie odjęciu od nich wektora dla warunku niezaburzonego (który musiał być i tak wyliczony, aby określić czy spełnia on układ równań (5.39)). W teorii na drodze przejścia granicznego  $\rho \rightarrow 0$  powinniśmy uzyskać wynik dokładny. W praktyce jednak wymaga to postawienia bardzo dużych wymagań wobec błędu metody numerycznej wyliczającej rozwiązanie równania różniczkowego, i w konsekwencji bardzo drobnego kroku. To z kolei wydłuża czas obliczeń, a dodatkowo przy bardzo małych długościach kroku metody numerycznej pojawiają się inne źródła błędów numerycznych. Aproksymacja różnicowa macierzy Jakobiego jest efektywna tylko do uzyskiwania przybliżeń, a nie dokładnych wyników.

### Zastosowanie równania wariacyjnego

Dużo bardziej interesujące, wydaje się wyliczenie macierzy Jakobiego w sposób dokładny (przez dokładny rozumiemy tu uzyskanie na drodze całkowania numerycznego). Jednym ze sposobów uzyskania dokładnego rozwiązania jest zastosowanie tzw. *równania wariacyjnego* (zob. Hairer et al., 2000, str. 95).

Równanie wariacyjne ma postać

$$\dot{\Phi}(t) = J(x(t))\Phi(t) \quad (5.43)$$

$$\Phi(t_k) = I \in \mathbb{R}^{n \times n} \quad (5.44)$$

$$x(t) = F(w_k, t) \quad (5.45)$$

$$t \in [t_k, t_{k+n-1}] \quad (5.46)$$

przy czym  $J(x(t))$  jest macierzą Jakobiego prawej strony równania różniczkowego systemu (5.36). Rozwiązanie równania wariacyjnego ma następującą własność

$$\frac{dF(w_k, \tau)}{dw_k} = \Phi(\tau) \quad \forall \tau \in [t_k, t_{k+n-1}] \quad (5.47)$$

Innymi słowy - rozwiązanie równania wariacyjnego w chwili  $\tau$  określa nam pochodną rozwiązania równania różniczkowego względem warunku początkowego w danej chwili. Warto zauważyć, że w systemach liniowych pochodną tą jest macierz  $e^{At}$ .

Rozważmy jedno z równań układu (5.39)

$$h(F(w_k, t_{k+l})) - y(t_{k+l}) = 0$$

Pochodna jego lewej strony względem warunku początkowego, w oparciu o różniczkowanie funkcji złożonych, wyraża się przez

$$\begin{aligned} \nabla_{w_k}(h(F(w_k, t_{k+l})) - y(t_{k+l})) &= \nabla_{w_k}(h(F(w_k, t_{k+l}))) = \\ &= \left( \frac{\partial F(w_k, t_{k+l})}{\partial w_k} \right)^T \nabla_x h(F(w_k, t_{k+l})) \end{aligned} \quad (5.48)$$

Aby wykorzystać wzór (5.48) konieczna jest znajomość pochodnej (5.37) względem warunków początkowych. Do tego celu zaś należy rozwiązać (5.43).

Wykorzystując rozwiązanie równania wariacyjnego (5.43) i wzór (5.48) dostajemy wyrażenia na kolejne wiersze macierzy Jakobiego

$$\nabla_{w_k}(h(F(w_k, t_{k+l})))^\top = \nabla_x h(F(w_k, t_{k+l}))^\top \Phi(t_{k+l})$$

Macierz Jakobiego ma więc postać

$$\frac{d}{dw_k} H(w_k) = \begin{bmatrix} \nabla_x h(F(w_k, t_{k+1}))^\top \Phi(t_{k+1}) \\ \nabla_x h(F(w_k, t_{k+2}))^\top \Phi(t_{k+2}) \\ \vdots \\ \nabla_x h(F(w_k, t_{k+n-1}))^\top \Phi(t_{k+n-1}) \end{bmatrix} \quad (5.49)$$

### Wariacja funkcjonału

Inne podejście, pozwalające na wyliczenie macierzy Jakobiego opiera się o wariację funkcjonału, zainspirowane zaś zostało poprzez dowód zasady maksimum (Bołtiański, 1971, str. 232). Sama zasada maksimum nie jest jednak wykorzystywana. Dla zwiększenia czytelności, w wyprowadzeniu zamiast  $w_k$  będziemy korzystać z  $x_0$ , a zamiast przedziału  $[t_k, t_{k+l}]$  będziemy rozpatrywać przedział  $[0, T]$ . Rozważmy pewien funkcjonal  $Q$  na trajektoriach systemu dynamicznego (5.36), oraz jego przyrost wygenerowany zmianą warunku początkowego. Zdefiniujmy funkcjonal oraz przyrost wywołany zmianą warunków początkowych z  $x^*(0)$  na  $x(0)$

$$Q(x_0) = q(x(T)) \quad (5.50)$$

$$q(x(T)) = h(F(x_0, T)) \quad (5.51)$$

$$\dot{x}^*(t) = f(x^*(t), u(t)) \quad (5.52)$$

$$t \in [0, T] \quad (5.53)$$

$$\Delta x(t) = x^*(t) - x(t) \quad (5.54)$$

$$(5.55)$$

Tak zdefiniowany funkcjonal odpowiada wartości wyjścia w jednym punkcie, czyli jednemu z równań układu (5.39). Wprowadzimy również Hamiltonian i równania sprzężone

$$\mathcal{H}(\psi, x, u) = \psi^\top f(x, u) \quad (5.56)$$

$$\dot{\psi}(t) = -\nabla_x \mathcal{H}(\psi(t), x(t), u(t)) \quad (5.57)$$

$$\psi(T) = -\nabla q(x(T)) \quad (5.58)$$

Możemy teraz przeanalizować przyrost funkcjonału (5.50)

$$\Delta Q = q(x^*(T)) - q(x(T)) = \nabla q(x(T)) \Delta x(T) + o(\Delta x(T)) = -\psi(T) \Delta x(T) + o(\Delta x(T)) \quad (5.59)$$

Wykorzystamy teraz następującą zależność

$$\psi(T) \Delta x(T) = \psi(0) \Delta x(0) + \int_0^T \dot{\psi}(t)^\top \Delta x(t) dt + \int_0^T \psi(t)^\top \Delta \dot{x}(t) dt$$

którą możemy podstawić do

$$\Delta Q = -\psi(0) \Delta x(0) - \int_0^T \dot{\psi}^\top \Delta x dt - \int_0^T \psi^\top \Delta \dot{x} dt + o(\Delta x(T)) = \quad (5.60)$$

$$= -\int_0^T -\nabla_x \mathcal{H}(\psi, x, u)^\top \Delta x dt - \int_0^T \psi^\top (f(x^*) + g(x^*)u - f(x) + g(x)u) dt - \psi(0) \Delta x(0) + o(\Delta x(T)) \quad (5.61)$$

W powyższych przekształceniach skorzystano z definicji równań sprzężonych (5.57). Dalej wykorzystujemy definicję Hamiltonianu (5.56) i pochodnej Frechet’a (zob. Bronsztejn, Siemiendajew, Musiol i Mühlig, 2004, str. 689)

$$\Delta Q = -\psi(0)\Delta x(0) - \int_0^T -\nabla_x \mathcal{H}(\psi, x, u)^\top \Delta x dt - \int_0^T (\mathcal{H}(\psi, x + \Delta x, u) - \mathcal{H}(\psi, x, u)) dt + o(\Delta x(T)) =$$
(5.62)

$$= -\psi(0)\Delta x(0) - \int_0^T o(\Delta x(t)) dt + o(\Delta x(T))$$
(5.63)

co daje nam formułę na przyrost

$$\Delta Q = -\psi(0)\Delta x(0) + o(\|\Delta x\|)$$
(5.64)

Co po przejściu granicznym daje wyrażenie na pochodną  $Q$  względem warunku początkowego  $x_0$  (gradient)

$$\nabla_{x_0} Q = -\psi(0)$$
(5.65)

Ponownie gradienty te można wykorzystać jako wiersze macierzy Jakobiego (powracając do wcześniejszych oznaczeń)

$$\frac{d}{dw_k} H(w_k) = - \begin{bmatrix} \psi^{t_k}(t_k) & \psi^{t_{k+1}}(t_k) & \dots & \psi^{t_{k+n-1}}(t_k) \end{bmatrix}^\top$$
(5.66)

gdzie  $\psi^\tau$  oznacza zmienną sprzężoną dla problemu (5.50) na przedziale  $[t_k, \tau]$ .

### Porównanie podejść

- Aproksymacja różnicowa wymaga rozwiązania układu  $n$  równań różniczkowych  $n$  razy na całym przedziale  $[t_k, t_{k+n-1}]$ . Jej efektywność zależy od parametru  $\rho$ , przy czym zbyt duże i zbyt małe jego wartości mogą prowadzić do błędów. Jest to metoda przybliżona, zalecana jedynie w sytuacji, gdy obliczenia analityczne potrzebne do zastosowania pozostałych podejść są zbyt skomplikowane.
- Zastosowanie równania wariacyjnego wymaga rozwiązania układu  $n^2$  równań jednocześnie na przedziale  $[t_k, t_{k+n-1}]$ . Jest to metoda dokładna i zarazem najłatwiejsza w implementacji ze wszystkich trzech - równanie wariacyjne można rozwiązywać równolegle z wyliczaniem wektora  $H(w_k)$ .
- Podejście trzecie, wymaga rozwiązania  $n$  układów  $n$  równań różniczkowych wstecz ale za to na różnych przedziałach całkowania (coraz krótszych). Jest to również metoda dokładna i potencjalnie najmniej kosztowna numerycznie. Wymaga ona jednakże dodatkowych nakładów poświęconych na implementację. Jej numeryczna przewaga może być również bardzo zależna od wymiaru i struktury macierzy Jakobiego systemu (5.36).

*Uwaga 5.3.* Autorzy hybrydowego obserwatora Newtona rozwijali go w innym kierunku. W ich badaniach wykorzystywane są aproksymacje macierzy Jakobiego ze wskazaniem na przybliżenie różnicowe (Biyik i Arcak, 2006a). Proponowane przez nich (Biyik i Arcak, 2006b) inne przybliżenia wywodzące się z tzw. metod „typu Newtona” jak np. metoda Broydena (zob. np. Łukasik, 2007, str. 19) zdaniem autora nie zdają egzaminu. Jest to uzasadnione, ponieważ rozwiązania równań różniczkowych nie muszą podlegać tym samym regułom co równania algebraiczne. Obszar badań autorów oryginalnej metody dotyczy bardziej modyfikacji metody Newtona-Rhapsona, w której to pierwszy wzór metody Newtona (5.41) zastępuje się przez

$$w_k^{i+1} = w_k^i + \theta s_i$$

gdzie parametr  $\theta \in (0, 1]$  jest dobierany tak aby zmniejszyć wartość wyrażenia

$$(Y_k - H(w_k))^\top Y_k - H(w_k)$$

Zdaniem autora, przy rozpatrywaniu wskaźników jakości tego typu dużo lepszym rozwiązaniem jest podejście optymalizacyjne typu MHE (opisane w dalszej części rozdziału), w którym to minimalizuje się wskaźnik

jakości, potencjalnie za pomocą metody Newtona rozwiązywania problemów optymalizacji (zob. np. Nocedal i Wright, 2006, str. 22), z wykorzystaniem zarówno informacji o gradiencie jak też i o Hessianie. ■

*Uwaga 5.4.* Należy zaznaczyć, że algorytm hybrydowego obserwatora Newtona podobnie jak sama metoda Newtona ma lokalny charakter zbieżności. Można próbować poprawiać obszary zbieżności za pomocą np. stosowania tzw. niedokładnej metody Newtona lub przez poszukiwanie na kierunku. Warto jednak zauważyć, że informacja dotycząca macierzy Jakobiego, zwłaszcza wyliczonej dokładnie może służyć za interesujące kryterium lokalnej obserwowalności systemu nieliniowego, w postaci rzędu macierzy Jakobiego (istnieją analogiczne kryteria obserwowalności dla układów liniowych). Warto jednak zauważyć, że nie jest to warunek rozstrzygający, ponieważ pomimo nieosobliwej macierzy Jakobiego, obserwator Newtona może nie być zbieżny. ■

Hybrydowy obserwator Newtona był rozważany we wcześniejszych pracach autora (Baranowski i Tutaj, 2007, 2008a) z wykorzystaniem przybliżonej macierzy Jakobiego, zaś z wykorzystaniem dokładnej w pracach (Baranowski i Mitkowski, 2009; Baranowski, 2009b).

## 5.3. Estymacja z zakłóceniami

### 5.3.1. Dyskretno-ciągły rozszerzony filtr Kalmana

Rozszerzony filtr Kalmana (*ang. Extended Kalman Filter, EKF*) jest obecnie najczęściej stosowanym estymatorem stanu w zastosowaniach przemysłowych. Można go znaleźć w przemyśle chemicznym, ciepłownictwie ale i również jako obserwator prędkości samochodu w sterownikach układu ABS. Najbardziej popularnym wariantem (ze względu na prostotę implementacji) jest dyskretny rozszerzony filtr Kalmana (zob. Byrski, 2007, str. 372). W niniejszym podrozdziale zaprezentowany zostanie dyskretno-ciągły (hybrydowy) EKF (zob. Simon, 2006, str. 403).

W kontekście EKF rozważamy równoodległe pomiary wyjścia (5.10) z  $t_0 = 0$  i następujące systemy dynamiczne

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t), w(t), t) \\ y(kh) &= h(x(kh), v_k, kh) \\ k &= 0, 1, 2, \dots\end{aligned}\tag{5.67}$$

Przy czym  $w(t)$  jest ciągłym białym szumem o macierzy kowariancji  $Q$ , zaś  $v_k$  jest dyskretnym białym szumem o macierzy kowariancji  $R_k$ .

U podstaw EKF leży metoda linearyzacji. Jednak zamiast linearyzować system (5.67) wokół stałego punktu równowagi, dokonujemy linearyzacji równań systemu wokół estymowanej trajektorii (zob. np. Mitkowski, 1996, str. 127 i 128). W tym celu wprowadzamy macierze

$$A(t) = \left. \frac{\partial f}{\partial x} \right|_{(\hat{x}(t), u(t), 0, t)} \quad L(t) = \left. \frac{\partial f}{\partial w} \right|_{(\hat{x}(t), u(t), 0, t)}\tag{5.68}$$

$$H_k = \left. \frac{\partial h}{\partial x} \right|_{(\hat{x}(kh), 0, kh)} \quad M_k = \left. \frac{\partial h}{\partial v} \right|_{(\hat{x}(kh), 0, kh)}\tag{5.69}$$

Z wykorzystaniem tych macierzy możemy, analogicznie jak w liniowym hybrydowym filtrze Kalmana, wprowadzić od razu podział na predykcję i korekcję. Tak też w kroku predykcji rozwiązujemy równanie systemu i równanie propagacji dla  $t \in [(k-1)h, kh]$

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t), 0, t)\tag{5.70}$$

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)^\top + L(t)QL(t)^\top\tag{5.71}$$

przy czym, analogicznie jak w przypadku liniowym mamy  $P((k-1)h) = P_{k-1}^+$  i  $\hat{x}((k-1)h) = \hat{x}_{k-1}^+$  zaś końcowe wartości rozwiązań oznaczmy jako  $P_k^- = P(kh)$  oraz  $\hat{x}_k^- = \hat{x}(kh)$ . Posiadając te informacje możemy

skonstruować, podobnie analogicznie jak w przypadku liniowym krok korekcji. Mamy więc

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + M_k R_k M_k^T)^{-1} \quad (5.72)$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y(kh) - h(\hat{x}_k^-, 0, kh)) \quad (5.73)$$

$$P_k^+ = (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k M_k R_k M_k^T K_k^T \quad (5.74)$$

*Uwaga 5.5.* Z hybrydowym rozszerzonym filtrem Kalmana związane są pewne problemy. Po pierwsze nie ma dobrych kryteriów doboru wartości początkowych.

W teorii, asymptotycznie stabilny obserwator powinien dobrze działać dla dowolnych warunków początkowych. W praktyce okazuje się jednak, że rozszerzony filtr Kalmana w pewnych przypadkach wykazuje wolną zbieżność błędu estymacji do zera. Zachodzi to wtedy, gdy nie został on prawidłowo zainicjowany. Szczególnie dla poprawnej pracy filtru istotny jest dobór początkowej macierzy kowariancji  $P_0^+$ .

Optymalnie wartości inicjalizujące filtr powinny być równe

$$\begin{aligned} \hat{x}_0^+ &= E[x_0] \\ P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \end{aligned}$$

gdzie  $E[\cdot]$  oznacza wartość oczekiwaną. Niestety praktyczna stosowalność powyższych wzorów jest niewielka, i początkowe wartości trzeba zgadnąć. W zależności od tego jak zostaną one dobrane zbieżność estymaty do prawdziwego stanu może być diametralnie różna. Drugim problemem związanym z filtrem Kalmana, istotnym zwłaszcza w zastosowaniach przemysłowych jest to, że nie można w nim uwzględnić ograniczeń zmiennych stanu. Z tego powodu, na podobnej zasadzie jak opisano w podrozdziale 3.2, mogą wystąpić duże błędy przejściowe, które mogą uczynić estymatę bezużyteczną. W pewnym stopniu z tymi problemami radzi sobie opisana niżej estymacja z ruchomym horyzontem (MHE) (Tenny i Rawlings, 2002).

Jeszcze inną słabością EKF jest jego zależność od linearyzacji systemu. Jeżeli system zlinearyzowany w pewnych obszarach przestrzeni stanu nie jest obserwowalny filtr może przestać działać. Przykładem systemu tracącego lokalną obserwowalność jest system (5.142) z przykładu 5.2 (str. 81), który nie jest obserwowalny w punkcie równowagi i jego otoczeniu.

■

### 5.3.2. Estymacja z ruchomym horyzontem (*ang. moving horizon estimation, MHE*)

MHE jest obecnie jedną z najbardziej rozwijanych metod estymacji stanu oraz parametrów układów dynamicznych. Literatura dotycząca tego zagadnienia jest olbrzymia i obejmuje procesy liniowe, nieliniowe, dyskretnie, ciągłe i hybrydowe, z czego ważniejsze pozycje dotyczące teorii to (Albuquerque i Biegler, 1994; Haseltine i Rawlings, 2008; Ferrari-Trecate, Mignone i Morari, 2002; Rao i Rawlings, 2002; Rao, Rawlings i Lee, 1999; Rao, Rawlings i Mayne, 2003; Tenny i Rawlings, 2002; Rawlings i Bakshi, 2006; Wang, Liu, Ban i Pei, 2003; Zavala, Laird i Biegler, 2007). Sama koncepcja nie jest jednak nowa, można np. znaleźć prezentację metody w (Eykhoff, 1980, str. 560), zaś bibliografia tamże kieruje do prac z wczesnych lat 60-tych. Można podejrzewać, że rozkwit zainteresowania metodą związany jest z rozwojem możliwości komputerów, pozwalających na rozwiązywanie występujących w MHE problemów optymalizacyjnych.

Poniżej przedstawiona zostanie wersja dla układów ciągłych z dyskretnymi pomiarami, z zakłóceniami jedynie pomiarów wyjścia estymująca warunek początkowy systemu. Wersje dopuszczające estymację parametrów oraz uwzględniające zakłócenia wchodzące na system są jak najbardziej możliwe, jednak stopień ich komplikacji znacząco rośnie.

Rozważmy system dynamiczny postaci

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t)) + v(t) \end{aligned} \quad (5.75)$$

Gdzie  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^r$ ,  $y(t) \in \mathbb{R}$  i  $v(t) \in \mathbb{R}$  jest zakłóceniem. Rozpatrujemy  $N+1 > n$  pomiarów wyjścia w pewnych chwilach  $t_k, t_{k+1}, \dots, t_N$ . Innymi słowy będziemy rozpatrywać przesuwane okno pomiarowe o długości  $m = N + 1$  (stąd też określenie ruchomy horyzont). Poszukujemy estymaty wartości stanu na początku okna  $\hat{x}(t_k)$ . Zakładamy, że system (5.75) jest obserwowalny w sensie definicji 2.7.

Metodę MHE można wywodzić z estymacji największej wiarygodności Bayesa (Rao i Rawlings, 2002). Oznacza to, że szukamy takiego  $\hat{x}(t_k)$ , które maksymalizuje prawdopodobieństwo warunkowe, że jest on stanem początkowym dla zmierzonego wyjścia t.j.

$$\hat{x}(t_k) = \arg \max_{x(t_k)} p(x(t_k) | y(t_k), y(t_{k+1}), \dots, y(t_{k+N})) \quad (5.76)$$

Przyjmijmy, że szum pomiarowy  $v(t_i)$  dla  $i = k, k+1, \dots, N$  jest nieskorelowany. Wprowadźmy oznaczenia

$$p(x(t_k)) = p_{x(t_k)}(x(t_k)) \quad (5.77)$$

$$p(y(t_k), y(t_{k+1}), \dots, y(t_{k+N}) | x(t_k)) = \prod_{i=k}^{k+N} p_{v(t_i)}(y(t_i) - h(x(t_i))) \quad (5.78)$$

gdzie  $p_{x(t_k)}(x(t_k))$  oznacza naszą wiedzę a priori o stanie na początku okna,  $p_{v(t_i)}(\cdot)$  to gęstość prawdopodobieństwa szumu pomiarowego, zaś  $x(t_i)$ , dla  $i > k$  oznacza rozwiązanie (5.75) z warunkiem początkowym  $x(t_k)$ . Z twierdzenia Bayesa (zob. Bronsztejn et al., 2004, str. 817) mamy

$$p(x(t_k) | y(t_k), y(t_{k+1}), \dots, y(t_{k+N})) = \frac{p_{x(t_k)}(x(t_k)) \prod_{i=k}^{k+N} p_{v(t_i)}(y(t_i) - h(x(t_i)))}{p(y(t_k), y(t_{k+1}), \dots, y(t_{k+N}))} \quad (5.79)$$

Mianownik (5.79) nie zależy od  $x(t_k)$ , więc nie wpływa on na położenie maksimum. Rozważmy więc funkcję wiarygodności

$$\mathcal{J}(x(t_k)) = p_{x(t_k)}(x(t_k)) \prod_{i=k}^{k+N} p_{v(t_i)}(y(t_i) - h(x(t_i))) \quad (5.80)$$

dla której

$$\hat{x}(t_k) = \arg \max \mathcal{J}(x(t_k)) \quad (5.81)$$

Wiadomo, że funkcja wiarygodności i jej logarytm osiągają maksimum w tym samym punkcie, więc

$$\hat{x}(t_k) = \arg \max_{x(t_k)} \ln \mathcal{J}(x(t_k)) \quad (5.82)$$

co po podstawieniu (5.80) daje

$$\hat{x}(t_k) = \arg \max_{x(t_k)} \ln \left( p_{x(t_k)}(x(t_k)) \prod_{i=k}^{k+N} p_{v(t_i)}(y(t_i) - h(x(t_i))) \right) = \quad (5.83)$$

$$= \arg \max_{x(t_k)} \left( \ln p_{x(t_k)}(x(t_k)) + \sum_{i=k}^{k+N} \ln (p_{v(t_i)}(y(t_i) - h(x(t_i)))) \right) \quad (5.84)$$

Wprowadźmy teraz następujące oznaczenie (celem skrócenia zapisu)

$$|\alpha|_A^2 = (\alpha, A\alpha) = \alpha^\top A \alpha \quad (5.85)$$

Założmy, że gęstość  $p_{x(t_k)}(x(t_k))$  jest Gaussowska z wartością średnią  $\bar{x}$  oraz kowariancją  $\Pi_k^{-1}$  a szum pomia-



rowy  $v$  jest szumem Gaussowskim o zerowej średniej i kowariancji  $R$  otrzymujemy

$$\hat{x}(t_k) = \arg \max_{x(t_k)} \left( -|x(t_k) - \bar{x}|_{\Pi_k}^2 + \sum_{i=k}^{k+N} \left( -|y(t_i) - h(x(t_i))|_{R^{-1}}^2 \right) \right) = \quad (5.86)$$

$$= \arg \min_{x(t_k)} \left( |x(t_k) - \bar{x}|_{\Pi_k}^2 + \sum_{i=k}^{k+N} |y(t_i) - h(x(t_i))|_{R^{-1}}^2 \right) \quad (5.87)$$

Uzyskaliśmy w ten sposób wygodną do wykorzystania postać estymacji stanu w formie problemu optymalizacyjnego. W ramach tej pracy będziemy rozwiązywać problem (5.87), jednakże będziemy traktować go deterministycznie - analogicznie do obserwatora najmniejszych kwadratów.

Warto zwrócić uwagę, że do zbliżonej postaci można by dojść deterministycznie. Powyższe rozważania zostały tu przedstawione, ze względu na stosunkowo intuicyjną dla czytelnika interpretację, bazującą na najprostszych pojęciach. Wyprowadzenie deterministyczne wymaga bardziej skomplikowanego aparatu matematycznego (np. pojęcia funkcji wartości (ang. *value function*)).

W ramach ujęcia deterministycznego będziemy zakładać, że problem ma postać problemu najmniejszych kwadratów. Dlatego też będziemy rozważać  $R = I$ . Kwestię deterministycznej interpretacji członu z macierzą  $\Pi_k$  omawia zaś uwaga 5.6. W sformułowaniu deterministycznym mamy następujący problem optymalizacyjny

$$\min_{\hat{x}(t_k)} L(\hat{x}(t_k)) \quad (5.88)$$

$$L(\hat{x}) = |\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2 + \sum_{i=k}^{k+N} |y(t_i) - h(\hat{x}(t_i))|^2 \quad (5.89)$$

p.o.

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (5.90)$$

przy czym  $\bar{x}_{k-1}$  oznacza estymatę stanu w chwili  $t_k$  uzyskaną z okna pomiarowego  $[t_{k-1}, t_{k+N-2}]$ , zaś macierz  $\Pi_k$  zmienia się z okna na okno (patrz uwaga 5.6). Problem optymalizacji często uzupełnia się o dodatkowe ograniczenia na zmienne stanu - wynikające z charakteru rzeczywistego procesu (np. dodatniość stężeń, zakresy temperatur, licznosc populacji itp.), lub też powstałe w oparciu o wiedzę a priori (np. przy zmianie punktu pracy instalacji przemysłowej możemy określić obszar przestrzeni stanu, do której należy warunek początkowy), mające na celu ułatwienie znalezienia minimum globalnego.

*Uwaga 5.6* (O wyznaczaniu macierzy kosztu początkowego). Macierz  $\Pi_k$  jest konieczna aby metoda MHE gwarantowała własności filtrujące. W przypadku gdyby wskaźnik składał się z samej sumy mielibyśmy efektywnie do czynienia z rozwiązywaniem pewnego układu równań za pomocą metody najmniejszych kwadratów (podobnie jak w metodzie Newtona-Raphsona). Podana już wyżej interpretacja probabilistyczna przedstawia macierz  $\Pi_k^{-1}$  jako kowariancję rozkładu określającego dotychczasową informację o stanie systemu. Poniżej spróbujemy przedstawić interpretację deterministyczną.

Oczywistym jest, że odtwarzanie stanu w oparciu o okno rozszerzane (od  $t_0$  do ostatniego dostępnego pomiaru  $t_{k+N-1}$ ) zawiera maksymalną dostępną ilość informacji dotyczącą ewolucji stanu systemu dynamicznego (5.75). Niestety, wraz z wzrastającą liczbą pomiarów rozwiązanie w ogólności nieliniowego problemu najmniejszych kwadratów z więzami w formie równań różniczkowych jest numerycznie coraz mniej praktyczne. Dlatego też rozsądnym jest aproksymowanie wskaźnika jakości

$$L_{roz}(\hat{x}(t_0)) = \sum_{j=0}^{k+N} |h(\hat{x}(t_j)) - y(t_j)|^2 \quad (5.91)$$

za pomocą

$$L(\hat{x}(t_k)) = P(\hat{x}(t_k), k) + \sum_{j=k}^{k+N} |h(\hat{x}(t_j)) - y(t_j)|^2 \quad (5.92)$$

tak, aby minimum obydwu wskaźników było takie samo.  $P(\hat{x}(t_k), k)$  nazywane jest *kosztem początkowym*

(ang. arrival cost).

Najlepiej, jeżeli wartość  $P(\hat{x}(t_k), k)$  byłaby równa zmianie wartości sumy

$$\sum_{j=0}^k |h(\bar{x}(t_j)) - y(t_j)|^2 \quad (5.93)$$

gdzie

$$\begin{aligned} \dot{\bar{x}}(t) &= f(\bar{x}(t), u(t)) \\ t &\in [t_0, t_k] \\ \bar{x}(t_k) &= \bar{x}_{k-1} \end{aligned} \quad (5.94)$$

wywołanej zmianą  $\bar{x}_{k-1}$  o  $\xi = \bar{x}_{k-1} - \hat{x}(t_k)$ . W ten sposób określałby on jak zmiana estymaty w danym kroku względem estymaty poprzedzającej odbija się na dokładności estymacji. Oznaczmy funkcję, która posiada taką własność jako  $V(\xi)$ . Oczywiście jest, że  $V(\xi)$  nie jest funkcją, którą można w praktyczny sposób wyznaczyć (jeżeli chcemy uniknąć analizy wszystkich dotychczasowych pomiarów).

W praktyce, chcielibyśmy aby koszt początkowy spełniał następującą zależność

$$\bar{Q}_k \leq P(\hat{x}(t_k), k) \leq V(\xi) \quad (5.95)$$

gdzie  $\bar{Q}_k$  jest minimalną wartością sumy (5.93) po wartościach  $\hat{x}(t_k)$ . Jeżeli w systemie (5.75) nie występują zakłócenia, to oczywiście jest, że  $\bar{Q}_k = 0$ ,  $\forall k$ , jeżeli natomiast zakłócenia występują, to  $Q_k$  jest funkcją nie-  
malejącą względem  $k$ . Co więcej, można wykazać (wynika to z warunku Lipschitza), że

$$V(\xi) - \bar{Q}_k \leq \gamma(|\xi - \hat{x}|) \quad (5.96)$$

gdzie  $\gamma \in \mathcal{K}$ .

Konsekwencją powyższych faktów jest to, że koszt początkowy również musi być podobnie ograniczony i niemalejący. Dwa nasuwające się naturalnie sposoby aproksymacji kosztu początkowego to

$$P(\hat{x}(t_k), k) = \bar{Q}_k = \text{const.} \quad (5.97)$$

$$P(\hat{x}(t_k), k) = V(\xi) \quad (5.98)$$

Pierwszy z nich nie zmienia wyniku optymalizacji, zaś drugiego nie da się wyznaczyć. Dlatego też popularnym rozwiązaniem jest aproksymacja  $P$  za pomocą formy kwadratowej, która lokalnie powinna przybliżać funkcję  $V$ , czyli:

$$P(\hat{x}(t_k), k) = \bar{Q}_k + |\xi - \hat{x}|_{\Pi_k}^2 \quad (5.99)$$

gdzie  $\Pi_k$  jest macierzą dodatnio określoną.

Wybór tej macierzy jej jest osobnym problemem. Oczywiście jest, że powinna się ona zmieniać wraz z  $k$ . Większość autorów sugeruje, aby korzystać z interpretacji probabilistycznej i przyjąć, że  $\Pi_k^{-1}$  jest macierzą kowariancji otrzymaną z rozwiązania rozszerzonego filtru Kalmana. Z doświadczenia autora niestety wynika, że zastosowanie EKF bez dobrego przybliżenia początkowego macierzy kowariancji prowadzi do wolnej zbieżności, co tylko utrudnia działanie MHE. Inną metodą to zastosowanie Hesjanu pewnego pomocniczego wskaźnika jakości, określającego jak zmiana warunku na końcu poprzedniego okna pogarsza jakość estymaty stanu. W końcowym wskaźniku jakości pomijamy  $\bar{Q}_k$ , ponieważ stała nie wpływa na optymalizację.

Doświadczenie autora, poparte eksperymentami na obiekcie rzeczywistym (zob. przykład 5.3 na stronie 83) sugeruje, że dobrym wyborem  $\Pi_k$  jest następująca macierz diagonalna z wykładniczymi elementami

$$\Pi_k = \text{diag}(1 - e^{-\alpha_1 k}, \dots, 1 - e^{-\alpha_n k}) \quad (5.100)$$

gdzie  $\alpha_i$  są pewnymi dodatnimi parametrami, dobranymi eksperymentalnie. Uzasadnienie tej właśnie postaci ma charakter bardziej praktyczny. Można mianowicie uważać  $|\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2$  jako pewien środek zabezpieczający przed pogorszeniem estymaty przez zakłócenia wyjścia. Założeniem podstawowym dla tego podejścia jest, że jeżeli uzyskana z poprzedniego okna estymata była dobra (czyli wskaźnik jakości był mały, a system

jest obserwowalny w sensie definicji 2.7), to znacznie gorsza wartość wskaźnika na kolejnym oknie jest oznaką działania zakłóceń. Na początku optymalizacji macierz  $\Pi_k$  jest bliska zeru, ponieważ nie mamy zbyt dużej wiedzy na temat stanu, im dalej jednak się posuwamy, tym bardziej wiarygodne stają się informacje uzyskane do tej pory. ■

*Uwaga 5.7.* Zwróćmy uwagę, że we wskaźniku jakości (5.87) występuje macierz  $R$ , odpowiadająca za kowariancję szumu pomiarowego. Jeżeli pomiary są skalarne, to zazwyczaj można ją zaniedbać, co i tak robi się w przypadku podejścia deterministycznego. Jeżeli jednak dysponujemy informacją o rozkładzie prawdopodobieństwa zakłóceń (np. z estymatora jądrowego, zob. Kulczycki, 2005, str. 55), i różni się on znacznie od normalnego, dogodniej jest pozostawić wskaźnik jakości w formie (5.84), ponieważ może to znacznie poprawić jakość estymacji. Dobór kosztu początkowego pozostaje jednak wtedy całkowicie otwartym problemem. ■

## Rozwiązywanie problemu optymalizacji

W niniejszym podrozdziale omówimy możliwość zastosowania optymalizacyjnej metody Newtona-Raphsona (zob. Nocedal i Wright, 2006, str. 22) do rozwiązywania problemu optymalizacji (5.89). Dla przypomnienia, metodę tą możemy zapisać w następujący sposób. Dany jest problem minimalizacji

$$\min_x F(x), \quad F: \mathbb{R}^n \rightarrow \mathbb{R} \quad (5.101)$$

Poszukujemy przybliżenia rozwiązania optymalnego  $u_{i+1}$  problemu (5.101) w następujący sposób

$$\nabla^2 F(u_i) d_i = -\nabla F(u_i) \quad (5.102)$$

$$v_\theta = u_i + d_i \theta, \quad \theta \in \mathbb{R} \quad (5.103)$$

$$u_{i+1} = \arg \min_{v_\theta} F(v_\theta) \quad (5.104)$$

gdzie  $\nabla^2 F$  oznacza Hesjan (macierz drugich pochodnych). Aby móc zastosować metodę Newtona do minimalizacji wskaźnika jakości  $L(\hat{x}(t_k))$  konieczne jest wyliczenie jego gradientu i Hesjanu. W tym celu podamy dwa sposoby dokładnego wyliczenia gradientów wskaźnika jakości  $L(\hat{x}(t_k))$  oraz jeden sposób wyliczania jego Hesjanu.

Dla celu zachowania prostoty obliczeń rozważać będziemy systemy, w których operacja określająca wyjście z systemu ma postać  $h(x(t)) = Cx(t)$ . Możliwe jest przeprowadzenie poniższych rozważań dla postaci ogólniejszej, jednak wiąże się to ze zwiększeniem ilości obliczeń bez szczególnych różnic merytorycznych.

Poniższą analizę opieramy na strukturze wskaźnika jakości (5.89). Ponieważ operacja różniczkowania jest operacją liniową możemy wyliczyć gradienty oddzielnie dla każdego ze składników wskaźnika jakości

$$L(\hat{x}(t_k)) = |\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2 + \sum_{i=k}^{k+N} |y(t_i) - C\hat{x}(t_i)|^2$$

Ponieważ pierwszy człon jest formą kwadratową niezależną od rozwiązania równań różniczkowych systemu, więc jego gradient to  $\nabla_{\hat{x}(t_k)} |\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2 = 2\Pi_k(\hat{x}(t_k) - \bar{x}_{k-1})$  a Hesjan jest równy  $\nabla_{\hat{x}(t_k)}^2 |\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2 = 2\Pi_k$ .

W przypadku pozostałych składników sumy obliczenia są bardziej skomplikowane i dwa sposoby ich przeprowadzenia podajemy poniżej.

## Zastosowanie równania wariacyjnego

Rozważmy więc jeden ze składników

$$\nabla_{\hat{x}(t_k)} |y(t_i) - C\hat{x}(t_i)|^2 = \nabla_{\hat{x}(t_k)} \left( (y(t_i) - C\hat{x}(t_i))^T (y(t_i) - C\hat{x}(t_i)) \right) = \quad (5.105)$$

$$= -\nabla_{\hat{x}(t_k)} (C\hat{x}(t_i)) 2(y(t_i) - C\hat{x}(t_i)) = \quad (5.106)$$

$$= -\Phi(t_i)^T C^T 2(y(t_i) - C\hat{x}(t_i)) \quad (5.107)$$

gdzie  $\Phi(t)$  jest rozwiązaniem następującego równania wariacyjnego

$$\begin{aligned}\dot{\Phi}(t) &= J(\hat{x}(t))\Phi(t) \\ \Phi(t_k) &= I \\ t &\in [t_k, t_{k+N}]\end{aligned}\tag{5.108}$$

przy czym  $\hat{x}(t)$  jest rozwiązaniem równania różniczkowego (5.75) z warunkiem początkowym  $\hat{x}(t_k)$  dla  $t \in [t_k, t_{k+N}]$ , natomiast  $J(x)$  jest macierzą Jakobiego prawej strony równania różniczkowego systemu (5.75).

W ten sposób gradient wskaźnika jakości (5.89) ma postać

$$\nabla_{\hat{x}(t_k)} L(\hat{x}(t_k)) = 2\Pi_k(\hat{x}(t_k) - \bar{x}_{k-1}) - \sum_{i=k}^{k+N} \left( \Phi(t_i)^\top C^\top 2(y(t_i) - C\hat{x}(t_i)) \right)\tag{5.109}$$

Koszt numeryczny wyliczenia gradientu, to rozwiązanie równania wariacyjnego (5.108) i wyliczenie wartości jego rozwiązania w punktach  $t_k, \dots, t_{k+N}$ .

### Zastosowanie wariacji funkcjonału

Celem wyliczenia jednego ze składników sumy występującej we wskaźniku (5.89) można wykorzystać rozważania podane wcześniej przy wyliczaniu obserwatora Newtona. Podobnie jak tam, dla zwiększenia czytelności, rozpatrujemy przyrost pewnego funkcjonału jakości na rozwiązaniach równania różniczkowego dla  $t \in [0, T]$ , wywołany przyrostem warunku początkowego  $x_0$ . Jeżeli przyjmiemy

$$q(x(T)) = |y(T) - Cx(T)|^2\tag{5.110}$$

to rozważania we wzorach (5.50)-(5.65) można w całości powtórzyć. Oznacza to, że po powrocie do starych oznaczeń

$$\nabla_{\hat{x}(t_k)} |y(t_i) - C\hat{x}(t_i)|^2 = -\psi(t_k)\tag{5.111}$$

przy czym  $\psi$  jest rozwiązaniem równania (5.57) na przedziale  $t \in [t_k, t_i]$  z warunkiem końcowym

$$\psi(t_i) = C^\top 2(y(t_i) - C\hat{x}(t_i))\tag{5.112}$$

Aby wykorzystać (5.112) do wyliczenia gradientu sumy we wskaźniku jakości (5.89) można zastosować podobne podejście jak w wyważeniu (5.109). Dogodniej jednak, z punktu widzenia ilości obliczeń rozważyć następującą formułę

$$\nabla_{\hat{x}(t_k)} L(\hat{x}(t_k)) = 2\Pi_k(\hat{x}(t_k) - \bar{x}_{k-1}) - \psi^k(t_k)\tag{5.113}$$

przy czym  $\psi^k$  jest wyrażone w sposób rekurencyjny:

$$\dot{\psi}^i(t) = -\nabla_x \mathcal{H}(\hat{x}(t), u(t), \psi^i(t))\tag{5.114}$$

$$t \in [t_i, t_{i+1}]\tag{5.115}$$

$$\psi^i(t_{i+1}) = \begin{cases} C^\top 2(y(t_i) - C\hat{x}(t_i)) + \psi^{i+1}(t_{i+1}) & \text{dla } i = k, \dots, k+N-1 \\ C^\top 2(y(t_i) - C\hat{x}(t_i)) & \text{dla } i = k+N \end{cases}\tag{5.116}$$

gdzie  $\mathcal{H}(x, u, \psi) = \psi^\top f(x, u)$ . Koncepcja takiego wyliczania gradientu opiera się na tym, że równanie sprzężone (5.57) można interpretować jako równanie liniowe względem  $\psi$  o zmiennych współczynnikach. W ten sposób operację dodawania jak w sumie (5.109) został przeniesiony na dodawanie warunków końcowych. Dlatego zamiast rozwiązywać układ równań sprzężonych  $N+1$  razy na przedziałach o zmiennej szerokości od  $t_{k+N}-t_k$  do  $t_{k+1}-t_k$  rozwiązujemy je na krótszych przedziałach o długości  $t_{k+i+1}-t_{k+i}$  dla  $i=0, \dots, N-1$ .

Aby wykorzystać metodę Newtona oprócz gradientu potrzebny jest również Hesjan, poniżej prezentowana jest metoda wyliczenia Hesjanu z wykorzystaniem równania wariacyjnego.

### Wyliczanie Hesjanu

Ponownie tok rozumowania przeprowadzimy dla pojedynczego składnika sumy występujące w (5.89). Gradient tego składnika jest dany wzorem (5.107). Dla zwięzłości zapisu oznaczmy

$$H_i = \nabla_{\hat{x}(t_k)}^2 |y(t_i) - C\hat{x}(t_i)|^2 \quad i = k, k+1, \dots, k+N \quad (5.117)$$

Ponieważ Hesjan jest macierzą drugich pochodnych to generalnie możemy sformułować Hesjan jako

$$H_i = \frac{d}{d\hat{x}(t_k)} \left( \nabla_{\hat{x}(t_k)} |y(t_i) - C\hat{x}(t_i)|^2 \right) = \frac{d}{d\hat{x}(t_k)} \left( -\Phi(t_i)^\top C^\top 2(y(t_i) - C\hat{x}(t_i)) \right) \quad (5.118)$$

gdzie  $\Phi$  jest rozwiązaniem równania wariacyjnego (5.108). Po pewnych elementarnych przekształceniach uzyskujemy

$$H_i = -2 \frac{d}{d\hat{x}(t_k)} \left( \Phi(t_i) \right)^\top C^\top (y(t_i) - C\hat{x}(t_i)) - 2\Phi(t_i)^\top C^\top \frac{d}{d\hat{x}(t_k)} (y(t_i) - C\hat{x}(t_i)) = \quad (5.119)$$

$$= -2 \frac{d}{d\hat{x}(t_k)} \left( \Phi(t_i) \right)^\top C^\top (y(t_i) - C\hat{x}(t_i)) + 2\Phi(t_i)^\top C^\top \frac{d}{d\hat{x}(t_k)} C\hat{x}(t_i) = \quad (5.120)$$

$$= -2 \frac{d}{d\hat{x}(t_k)} \left( \Phi(t_i) \right)^\top C^\top (y(t_i) - C\hat{x}(t_i)) + 2\Phi(t_i)^\top C^\top C\Phi(t_i) \quad (5.121)$$

We wzorze (5.121) występuje wyrażenie

$$\frac{d}{d\hat{x}(t_k)} \left( \Phi(t_i) \right)$$

które jest pochodną funkcji macierzowej po argumentie wektorowym, czyli wyraża się tensorem czwartego rzędu. Jest to kłopotliwe do zastosowania, dlatego też proponujemy podejście pozwalające na ominięcie rachunku tensorów. Wprowadźmy następujące oznaczenia

$$\Phi(t) = \begin{bmatrix} \phi_{11}(t) & \dots & \phi_{1n}(t) \\ \vdots & \ddots & \vdots \\ \phi_{n1}(t) & \dots & \phi_{nn}(t) \end{bmatrix} \quad (5.122)$$

$$\hat{x}(t) = [\hat{x}_1(t) \quad \dots \quad \hat{x}_n(t)]^\top \quad (5.123)$$

Rozważmy teraz alternatywne sformułowanie Hesjanu

$$H_i = -2V(t_i) + 2\Phi(t_i)^\top C^\top C\Phi(t_i) \quad (5.124)$$

gdzie

$$V(t_i) = \begin{bmatrix} V_1(t_i)^\top C^\top (y(t_i) - C\hat{x}(t_i)) & V_2(t_i)^\top C^\top (y(t_i) - C\hat{x}(t_i)) & \dots & V_n(t_i)^\top C^\top (y(t_i) - C\hat{x}(t_i)) \end{bmatrix} \quad (5.125)$$

$$V_j(t) = \frac{\partial}{\partial \hat{x}_j(t_k)} \Phi(t) \quad (5.126)$$

W tak przedstawionym sformułowaniu Hesjanu występuje macierz  $V(t)$  o wymiarach  $n \times n$ , do której wyliczenia potrzeba  $n$  macierzy  $V_j(t)$  o wymiarach  $n \times n$ . Aby wyliczyć macierz  $V_j(t)$  zróżniczkujemy równość

(5.126) po czasie. Mamy

$$\dot{V}_j(t) = \frac{d}{dt} \frac{\partial}{\partial \hat{x}_j(t_k)} \Phi(t) = \frac{\partial}{\partial \hat{x}_j(t_k)} \dot{\Phi}(t) = \frac{\partial}{\partial \hat{x}_j(t_k)} \left( J(\hat{x}(t)) \Phi(t) \right) = \quad (5.127)$$

$$= \frac{\partial}{\partial \hat{x}_j(t_k)} \left( J(\hat{x}(t)) \right) \Phi(t) + J(\hat{x}(t)) \frac{\partial}{\partial \hat{x}_j(t_k)} \Phi(t) = \quad (5.128)$$

$$= \sum_{p=1}^n \left( \phi_{pj}(t) \frac{\partial}{\partial z_p} J(z) \Big|_{z=\hat{x}(t)} \right) \Phi(t) + J(\hat{x}(t)) V_j(t) \quad (5.129)$$

Pierwszy składnik końcowej sumy uzyskano w oparciu o rachunek pochodnych funkcji złożonych (zob. np. Fichtenholtz, 1972, tom 1, str. 340), zaś drugi składnik wynika wprost ze wzoru (5.126). Uzyskujemy w ten sposób układ  $n^2$  równań różniczkowych, pozwalający na wyliczenie macierzy  $V_j$ . Pozostaje jeszcze kwestia warunku początkowego dla równania (5.129). W tym celu rozważmy Hesjan składnika sumy odpowiadającego za początek okna (w chwili  $t_k$ ), tzn.

$$H_k = \frac{d}{d\hat{x}(t_k)} \nabla_{\hat{x}(t_k)} |y(t_k) - C\hat{x}(t_k)|^2 = \frac{d}{d\hat{x}(t_k)} \nabla_{\hat{x}(t_k)} (y(t_k) - C\hat{x}(t_k))^T (y(t_k) - C\hat{x}(t_k)) = \quad (5.130)$$

$$= \frac{d}{d\hat{x}(t_k)} \nabla_{\hat{x}(t_k)} \left( y(t_k)^T y(t_k) - 2y(t_k)^T C\hat{x}(t_k) + \hat{x}(t_k)^T C^T C\hat{x}(t_k) \right) = \quad (5.131)$$

$$= \frac{d}{d\hat{x}(t_k)} \left( -2C^T y(t_k) + 2C^T C\hat{x}(t_k) \right) = 2C^T C \quad (5.132)$$

Z drugiej strony, bazując na wzorze (5.124) mamy

$$H_k = -2V(t_k) + 2\Phi(t_k)^T C^T C\Phi(t_k) = -2V(t_k) + 2C^T C \quad (5.133)$$

(skorzystano tu z warunku początkowego równania wariacyjnego (5.108)). Porównując wyrażenia (5.132) i (5.133) mamy

$$2C^T C = -2V(t_k) + 2C^T C \quad (5.134)$$

$$0 = -2V(t_k) \quad (5.135)$$

$$V(t_k) = 0 \in \mathbb{R}^{n \times n} \quad (5.136)$$

Analizując postać macierzy  $V$  (5.125), łatwo zauważyć, że z równości (5.136) wynika

$$V_j(t_k) = 0 \in \mathbb{R}^{n \times n} \quad j = 1, \dots, n \quad (5.137)$$

Podsumowując, Hesjan wskaźnika jakości  $L(\hat{x}(t_k))$  danego przez (5.89) wyraża się następującymi wzorami:

$$\begin{aligned}
 \nabla_{\hat{x}(t_k)}^2 L(\hat{x}(t_k)) &= 2\Pi_k + \sum_{i=k}^{N+k} H_i \\
 H_i &= -2V(t_i) + 2\Phi(t_i)^\top C^\top C \Phi(t_i) \\
 V(t_i) &= \begin{bmatrix} V_1(t_i)^\top C^\top (y(t_i) - C\hat{x}(t_i)) & \dots & V_n(t_i)^\top C^\top (y(t_i) - C\hat{x}(t_i)) \end{bmatrix} \\
 \dot{\Phi}(t) &= J(\hat{x}(t))\Phi(t) \\
 \dot{V}_j(t) &= \sum_{p=1}^n \left( \phi_{pj}(t) \frac{\partial}{\partial z_p} J(z) \Big|_{z=\hat{x}(t)} \right) \Phi(t) + J(\hat{x}(t))V_j(t) \\
 \Phi(t_k) &= I \in \mathbb{R}^{n \times n} \\
 V_j(t_k) &= 0 \in \mathbb{R}^{n \times n} \\
 j &= 1, \dots, n \\
 t &\in [t_k, t_{k+N}]
 \end{aligned} \tag{5.138}$$

Zakładając, że równanie wariacyjne (5.108) zostało rozwiązane na potrzeby liczenia gradientu, to dodatkowy koszt numeryczny wyliczenia Hesjanu to rozwiązanie  $n^3$  równań różniczkowych.

## 5.4. Przykłady zastosowań

**Przykład 5.1** (Zastosowanie nieliniowego obserwatora dyskretno ciągłego do estymacji wysokości cieczy w układzie zbiorników wodnych). Rozważany jest układ zbiorników wodnych jak na rysunku (5.1). Poniższy model matematyczny został opracowany podczas próby identyfikacji stanowiska zbiorników wodnych znajdującego się w Katedrze Automatyki AGH (zdjęcie stanowiska jest na stronie 83) na potrzeby pracy (Baranowski i Tutaj, 2007). Przedstawione w tym przykładzie wyniki mają charakter symulacyjny i pochodzą z cytowanej pracy.

System ten, opisany jest równaniami

$$\frac{dh_1}{dt} = \frac{1}{S_1} q(u) - \frac{1}{S_1} q_1(h_1) \tag{5.139}$$

$$\frac{dh_2}{dt} = \frac{1}{S_2(h_2)} q(u) - \frac{1}{S_2(h_2)} q_2(h_2) \tag{5.140}$$

$$\frac{dh_3}{dt} = \frac{1}{S_3(h_3)} q(u) - \frac{1}{S_3(h_3)} q_3(h_3) \tag{5.141}$$

gdzie  $h_i$  oznacza poziom cieczy w  $i$ -tym zbiorniku, zaś

$$S_1 = a d = \text{const}$$

$$S_2 = d \left( n + \frac{m-n}{c} h_2 \right)$$

$$S_3 = d \sqrt{2 r h_3 - h_3^2}$$

oznaczają powierzchnie swobodne cieczy w zbiornikach. Wynikają one z kształtu poszczególnych zbiorników. Przepływy cieczy przez poszczególne zbiorniki charakteryzowane są przez

$$q_1 = w_1(h_1) = w_{10} + w_{11} h_2 + w_{12} h_2^2 + w_{13} h_2^3$$

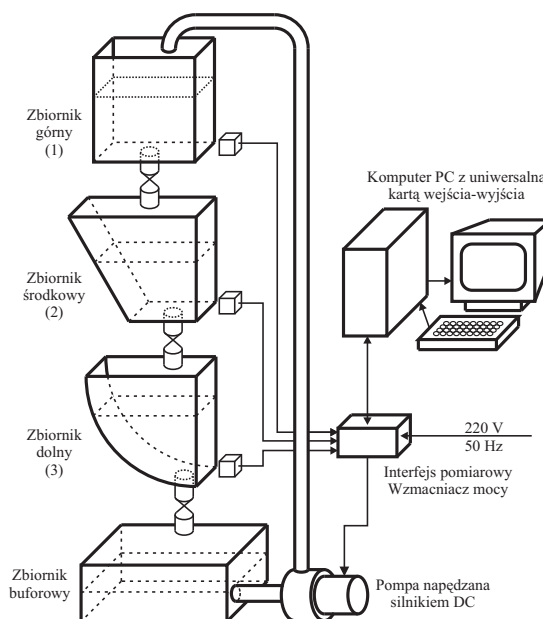
$$q_2 = w_2(h_2) = w_{20} + w_{21} h_3 + w_{22} h_3^2 + w_{23} h_3^3$$

$$q_3 = w_3(h_3) = w_{30} + w_{31} h_3 + w_{32} h_3^2 + w_{33} h_3^3$$

Zaś dopływ wody do górnego zbiornika z pompy opisuje (dla  $u \in [0, 1]$ ):

$$q = w_0(u) = w_{00} + w_{01} u + w_{02} u^2 + w_{03} u^3$$

Charakterystyki tu opisane mają postać wielomianów 3 stopnia i były próbą zamodelowania pewnych zjawisk zachodzących w obiekcie rzeczywistym, nie modelowanych przez prawo Toricellego. Parametry modelu, pochodzące z identyfikacji obiektu rzeczywistego zbiera tabela 5.1.



Rysunek 5.1.: Schemat systemu trzech zbiorników wodnych

Tabela 5.1.: Parametry modelu zbiorników

| $i$ | $w_{i0}$ | $w_{i1}$   | $w_{i2}$  | $w_{i3}$ |
|-----|----------|------------|-----------|----------|
| 0   | 625.1548 | -1325.7819 | 1140.4683 | -143.835 |
| 1   | 0.0006   | 0.044      | 4.6819    | 76.9685  |
| 2   | 0.0023   | -0.1331    | 6.5529    | 77.829   |
| 3   | 0.0052   | -0.2308    | 7.2811    | 67.2805  |

Jako cel postawiono estymację ciągłego stanu wysokości cieczy w zbiornikach górnym i środkowym w oparciu o pomiary wysokości cieczy w zbiorniku środkowym z okresem próbkowania równym 2s. Wykorzystano w tym celu nieliniowy obserwator dyskretno ciągły. Macierz  $Q$  obserwatora wyliczono symbolicznie przy użyciu programu Maple<sup>TM</sup>. Wraz z wyliczeniem macierzy wyliczono jej wyznacznik, który określa obserwowalność systemu:

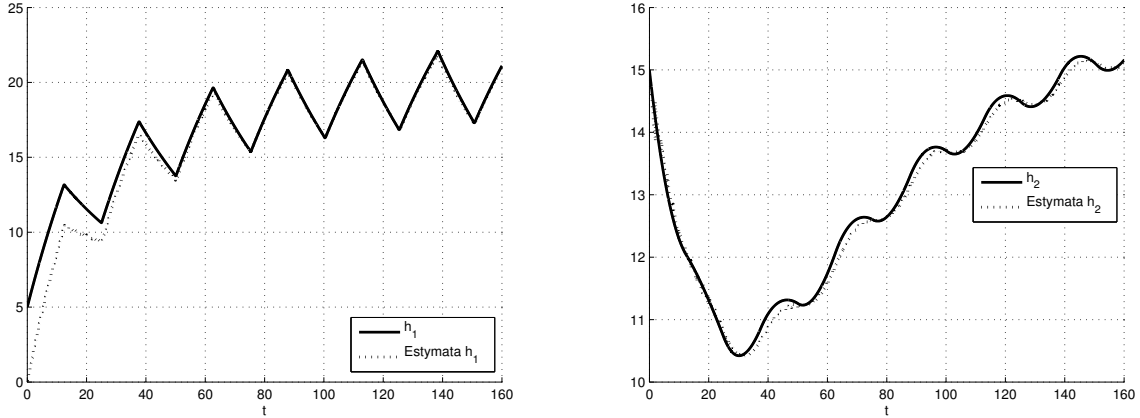
$$\det(Q) = -\frac{w_{11} + 2w_{12}h_1 + 3w_{13}h_1^2}{d\left(n + \frac{(m-n)h_2}{c}\right)}$$

Zeruje się on jedynie w sytuacji, w której charakterystyka zaworu przyjmuje wartość ekstremalną, co zachodzi dopiero poza zakresem stosowalności modelu.

Symulacyjne wyniki pracy obserwatora ilustruje rysunek 5.2. Ponieważ początkowa estymata stanu systemu była znacznie rozbieżna od stanu rzeczywistego obserwator wykazywał się powolną zbieżnością. W przypadku, gdy wiedza ta jest dokładniejsza, obserwator ten sprawuje się dużo lepiej (zob. Baranowski i Tutaj, 2007).



Dodatkowym problemem dla obserwatora był skokowo zmieniający się sygnał sterujący, który spowalniał zbieżność.



Rysunek 5.2.: Estymata wysokości cieczy w górnym (z lewej) i środkowym (z prawej) zbiorniku wodnym za pomocą nieliniowego obserwatora dyskretno ciągłego

Dokładniejszy opis zastosowania obserwatora, także do rzeczywistych danych pomiarowych zawierają prace (Baranowski i Tutaj, 2007; Baranowski, 2007d).

**Przykład 5.2** (Zastosowanie hybrydowego obserwatora Newtona do estymacji stanu pacjenta oddziału intensywnej opieki medycznej (OIOM)). Jednym z bardziej istotnych problemów w oddziałach OIOM jest nadzorowanie poziomu glukozy we krwi pacjentów. Wiele przypadłości, takich jak zawał serca lub uszkodzenia wielonarządowe powoduje nagły skok poziomu cukru. Jest to bardzo niebezpieczne, ponieważ wiąże się to ze zwiększonym ryzykiem infekcji, zaburza krzepnięcie krwi oraz destabilizuje równowagę metaboliczną. Badania pokazały, że rygorystyczne kontrolowanie poziomu cukru we krwi znacząco zmniejsza śmiertelność pacjentów na oddziałach. Obecnie, w celu uniknięcia hiperglikemii, bazuje się na częstych pomiarach poziomu cukru połączonych z podawaniem insuliny, lub też stosuje się automatyczne pompy insulinowe. Dotyczy to nie tylko cukrzyków, ale też ludzi nie mających wcześniej historii problemów z metabolizmem cukru.

Oczywistym jest, że aby prawidłowo podawać insulinę należy wiedzieć jak najwięcej o stanie pacjenta. Ponieważ jedyną wielkością łatwą do zmierzenia jest poziom cukru we krwi, istotnym staje się estymowanie pozostałych informacji, najlepiej uzyskując ciągłą informację, tak jak to jest rozważane w niniejszej pracy.

Model dynamiki cukru we krwi pacjenta OIOM opisuje poniższy model (zob. Van Herpe, Espinoza, Haverbeke, De Moor i Van den Berghe, 2007)

$$\begin{aligned}
 \dot{G}(t) &= (P_1 - X(t))G(t) - P_1G_b + \frac{F_G}{V_G} \\
 \dot{X}(t) &= P_2X(t) + P_3(I_1 - I_b) \\
 \dot{I}_1(t) &= \alpha \max(0, I_2) - n(I_1 - I_b) + \frac{F_I}{V_I} \\
 \dot{I}_2(t) &= \beta\gamma(G(t) - h) - nI_2(t)
 \end{aligned} \tag{5.142}$$

gdzie  $G$  i  $I_1$  to odpowiednio zawartość cukru i insuliny we krwi. Zmienna  $X$  odpowiada za efektywność insuliny zaś zmienna  $I_2$  nie ma bezpośredniej interpretacji klinicznej - wprowadzono ją po to, aby w pewien sposób symulować działanie trzustki u pacjenta nie będącego diabetikiem. Parametry  $G_b$  i  $I_b$  oznaczają bazową wartość glukozy i insuliny we krwi. Model posiada dwa wejścia - dopływ podawanej zewnętrznie insuliny -  $F_I$ , oraz dopływ węglowodanów (glukozy) -  $F_G$ . Obydwie te wielkości symulują podawanie dożylnie. Przestrzeń rozprzeczania glukozy i insuliny to odpowiednio  $V_G$  i  $V_I$ . Współczynnik  $P_1$  odpowiada za efektywność glukozy (ułamkowe oczyszczenie glukozy) gdy insulina pozostaje na poziomie bazowym.  $P_2$  i  $P_3$  są współczynnikami rozpadu insuliny oraz jej wzrostu zależnego od jej własnego poziomu. Własna insulina

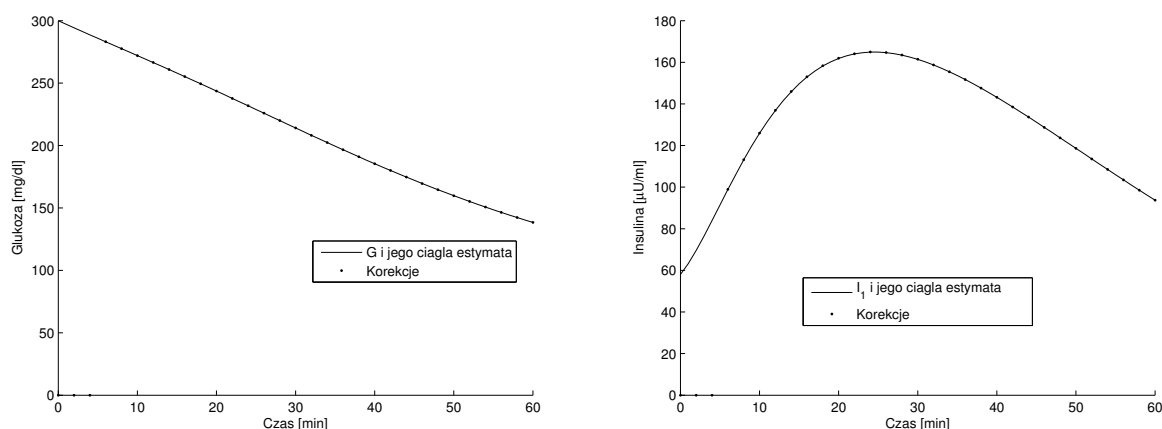
organizmu reprezentowana jest jako dopływ insuliny, wyzwalany proporcjonalnie do poziomu odniesienia  $h$  odpowiadający normalnemu poziomowi cukru czyli 7.5 mmol/dl. Stała czasowa rozpadu insuliny to  $n$  zaś współczynnik  $\beta = 1$  min odpowiada za zachowanie właściwych jednostek. Współczynnik  $\alpha$  jest współczynnikiem skalującym drugiej zmiennej insulinowej  $I_2$ . Więcej informacji można znaleźć w (Van Herpe et al., 2007), z której to pracy pochodzą parametry modelu.

Model ten jest rozszerzeniem tzw. modelu minimalnego (MM) opracowanego przez Bergmana i innych w 1981 r., w celu modelowania metabolizmu glukozy po dożylnym teście tolerancji glukozy (ang. intravenous glucose tolerance test (IVGTT)). Model ten następnie był dostosowywany przez Furlera i innych w 1985 r., aby modelować stan diabetyka. Wersja podana tutaj, została opracowana na potrzeby analizy pacjentów OIOM przez (Van Herpe et al., 2007) i nosi nazwę ICU-MM (Intensive Care Unit Minimal Model - Model minimalny dla pacjentów OIOM). Więcej szczegółów dotyczących rozwoju modelu minimalnego można znaleźć w (Makroglou, Li i Kuang, 2006; Kovács, 2006) oraz (Chee i Fernando, 2007, str. 73). Stabilizację poziomu cukru we krwi w modelu ICU-MM rozważali (Haverbeke, Van Herpe, Diehl, Van den Berghe i De Moor, 2008) oraz (Bauer, Płuciennik, Baranowski i Mitkowski, 2009; Baranowski i Mitkowski, 2009).

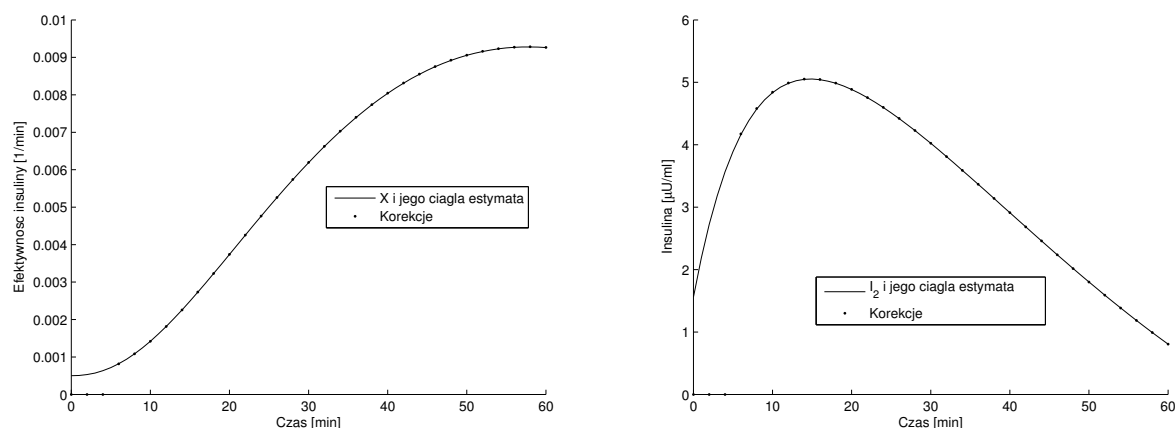
Do estymacji stanu modelu wykorzystano hybrydowy obserwator Newtona. Przyjęto, że pomiary glukozy dokonywane są co 15 minut, zaś system nie jest pobudzany z zewnątrz - nie ma dopływu ani insuliny ani glukozy. Analiza modelu pokazała, że nie jest on obserwowalny w otoczeniu punktu równowagi. Zastosowany obserwator Newtona, pozwolił jednak stwierdzić obserwowalność w części przejściowej dynamiki. Ze względu na to, że system jest asymptotycznie stabilny, postanowiono wykorzystać fakt, że obserwator dobrze znajduje estymatę na początku - w obserwowalnej części przestrzeni stanu i wykorzystać normę różnicy estymowanego i mierzonego wyjścia jako kryterium korzystania z obserwatora. Jeżeli błąd estymacji był mały, zdawano się na asymptotycznie stabilną dynamikę do utrzymania go na niskim poziomie. Przykład iteracji obserwatora Newtona ilustruje tabela 5.2, zaś wyniki symulacji przedstawione są na rysunkach 5.3 i 5.4. Jak można zauważyć, po uzyskaniu czterech pomiarów obserwator zaczął odtwarzać stan dokładnie. Zaprezentowane wyniki pochodzą z prac (Baranowski, 2009b; Baranowski i Mitkowski, 2009).

Tabela 5.2.: Przykład iteracji hybrydowego obserwatora Newtona

| $x_0$  | $w_0^1$ | $w_0^2$  | $w_0^3$ | $w_0^4$ | $w_0^5$ |
|--------|---------|----------|---------|---------|---------|
| 300    | 200     | 300      | 300     | 300     | 300     |
| 0.0005 | 0.01    | -0.0033  | 0.0004  | 0.0005  | 0.0005  |
| 58     | 56      | -57.7166 | 60.617  | 58.0021 | 58      |
| 1.5600 | -1.8966 | 20.6963  | 1.6996  | 1.5600  | 1.5600  |



Rysunek 5.3.: Estymata poziomu glukozy (z lewej) i insuliny (z prawej) u pacjenta OIOM uzyskana za pomocą hybrydowego obserwatora Newtona



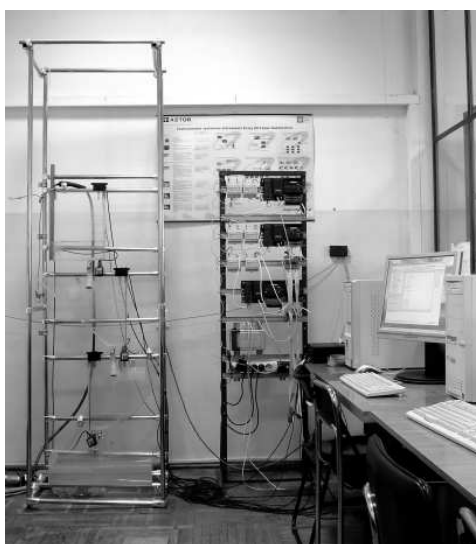
Rysunek 5.4.: Estymata efektywności insuliny (z lewej) oraz zmiennej pomocniczej  $I_2$  (z prawej) u pacjenta OIOM uzyskana za pomocą hybrydowego obserwatora Newtona

**Przykład 5.3** (Odtwarzanie stanu układu z rzeczywistych pomiarów przy użyciu estymacji z ruchomym horyzontem (MHE)). Rozważane jest rzeczywiste stanowisko laboratoryjne znajdujące się w Katedrze Automatyki AGH przedstawione na zdjęciu 5.5. System ten odpowiada układowi z przykładu 5.1. W toku badań nad układem, zdecydowano się zmienić modele przepływów z zaworów i pompy (zob. Baranowski i Tutaj, 2008b) na

$$q(u) = w_5 u^5 + w_4 u^4 + w_3 u^3 + w_2 u^2 + w_1 u + w_0 \quad (5.143)$$

oraz

$$q_i(h_i) = C_i \sqrt{D_i + h_i}, \quad i \in \{1, 2, 3\} \quad (5.144)$$



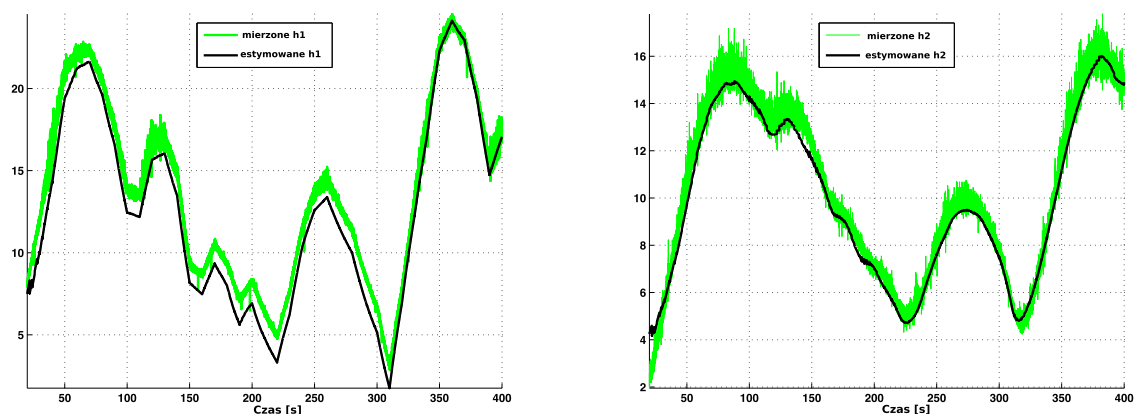
Rysunek 5.5.: Zdjęcie stanowiska laboratoryjnego trzech zbiorników wodnych znajdującego się w Katedrze Automatyki AGH

Wykorzystując taki model układu zastosowano metodę MHE, przy czym:

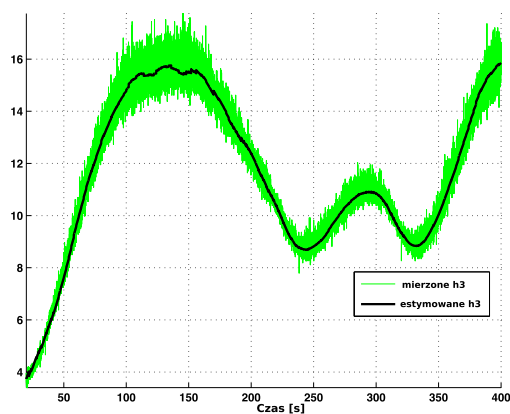
- Odstęp pomiędzy pomiarami wynosił  $h = 2$  s.
- Wykorzystano macierz kosztu początkowego typu (5.100)
- Szerokość okna pomiarowego wynosiła  $m = 10$

- Ze względu na niski rząd systemu, do optymalizacji stosowano algorytm Levenberga-Marquardta do rozwiązywania nieliniowych problemów najmniejszych kwadratów zaimplementowany w Matlabie.

Wyniki eksperymentów można zaobserwować na rysunkach 5.6 i 5.7. Jak można zauważyć metoda MHE ma bardzo dobre własności filtrujące, odtwarzając stan bardzo dokładnie. Wyniki pochodzą z pracy (Baranowski i Tutaj, 2008b).



Rysunek 5.6.: Estymata poziomu wody w górnym (z lewej) i środkowym (z prawej) zbiorniku uzyskana z MHE i wielkość mierzona - dane eksperymentalne



Rysunek 5.7.: Estymata poziomu wody w dolnym zbiorniku uzyskana z MHE i wielkość mierzona - dane eksperymentalne

## Rozdział 6.

### Metody przybliżone korekcji

#### 6.1. Przybliżanie dyskretnego wyjścia funkcją ciągłą

Rozważane będą systemy postaci

$$\begin{aligned}\dot{x} &= f(x, u) \\ y &= h(x)\end{aligned}\tag{6.1}$$

W ramach niniejszego rozdziału omawiane będą przybliżone metody korekcji. Pod tym pojęciem będziemy rozumieć metody wykorzystujące aproksymację dyskretnego wyjścia systemu przez pewną ciągłą funkcję czasu. Uzasadnienie tego podejścia jest bardzo proste, w stosunku do systemów dynamicznych, opisanych równaniami różniczkowymi jest bardzo dobrze opracowana metodyka odtwarzania stanu o ile dostępne jest ciągłe wyjście. Nasuwa się więc naturalnie pomysł, aby wykorzystując dyskretny pomiar wyjścia (2.4),  $y(t_i)$ ,  $i = k, k + 1, \dots, k + m$  wygenerować funkcję czasu  $\hat{y}(t)$ , określoną dla  $t \in [t_k, t_{k+m}]$  a następnie wykorzystać, którąś z metod ciągłej estymacji stanu na tym przedziale, aby uzyskać estymatę stanu w chwili  $t_{k+m}$ . Oczywiście musimy pogodzić się z faktem, że taka aproksymacja wyjścia wprowadza pewien błąd

$$\hat{y}(t) = y(t) + \chi(t)\tag{6.2}$$

jednakże można traktować  $\chi$  jako zakłócenie i dobrać taki algorytm estymacji, który będzie je kompensował. Co więcej, można dobierać metodę aproksymacji wyjścia w taki sposób, aby błąd  $\chi(t)$  z pomiaru na pomiar powodował coraz mniejsze błędy. Poniżej opisane zostaną metody przybliżania wyjścia za pomocą interpolacji oraz dwóch wariantów opracowanych przez autora kolokacji wyjścia - podstawowego i adaptacyjnego.

##### 6.1.1. Interpolacja

Najprostszym sposobem przybliżania wyjścia systemu jest interpolacja wielomianowa. Weźmy  $m + 1$  pomiarów wyjścia. Wielomian interpolacyjny przybliżający wyjście będzie miał postać

$$\hat{y}(t) = p_m t^m + \dots + p_1 t + p_0\tag{6.3}$$

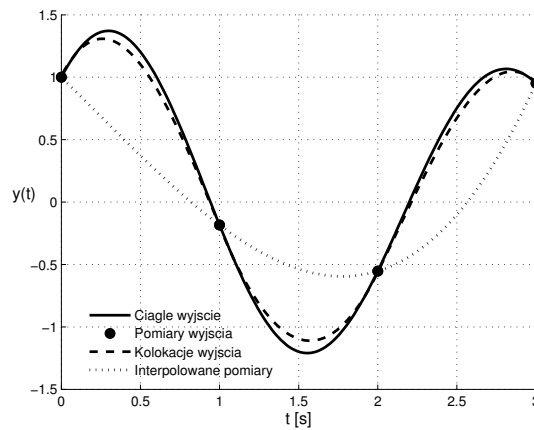
przy czym  $\hat{y}(t_i) = y(t_i)$  dla  $i = k, k + 1, \dots, k + m$  (zob. np. Mitkowski, 2003, str. 19 oraz Mitkowski, 2007b, str. 170). Dobór stopnia wielomianu jest decyzją arbitralną i zasadniczo zależy od tego jak zachowuje się wyjście systemu. Popularnym wyborem jest wielomian trzeciego stopnia, ponieważ stanowi sensowny kompromis pomiędzy elastycznością a prostotą obliczeń. Praktyka pokazuje, że zasadniczo jedyną zaletą przybliżenia wielomianem interpolacyjnym jest jego prostota. Za wyjątkiem pewnych bardzo prostych systemów taki wielomian może wprowadzać stosunkowo duże błędy. Ogólnie można zakładać, że wielomian będzie przybliżał wyjście względnie dobrze, jeżeli przy równoodległych pomiarach spełnione jest twierdzenie o próbkowaniu (Kotelnikowa-Shanona, (zob. np. Baranowski et al., 2007, str. 110)).

Zasadniczo lepszym sposobem interpolacji jest wykorzystanie funkcji sklepanych (zazwyczaj trzeciego stopnia), które pozwalają na większą elastyczność i zapobiegają efektowi Rungego. Uciąganie wyjścia za pomocą funkcji sklepanych (zob. np. Mitkowski, 2003, str. 23, 24 oraz Mitkowski, 2007b, str. 165). stosował np. (Nowak, 2007) do identyfikacji systemów. Także i tutaj ważne jest gęste rozmieszczenie punktów pomiarowych, co ogranicza błąd interpolacji.

### 6.1.2. Kolokacje wyjścia

Pewnym szczególnym podejściem do przybliżania wyjścia jest zaproponowana przez autora w pracy (Baranowski, 2008c) metoda kolokacji wyjścia. Co jest istotne, metoda ta jest bezpośrednio związana z ciągłym odtwarzaniem stanu z dyskretnych danych pomiarowych, albowiem do tego celu została opracowana. Poniżej prezentowany wariant ma parametry dobrane arbitralnie, modyfikacje jednak dla wymagań konkretnych problemów są łatwe do wprowadzenia.

Podstawową wadą interpolacji jest fakt, że nie bazowała ona w żaden sposób na tym, że wyjście systemu jest pewną funkcją rozwiązania równania różniczkowego. Co więcej, forma tego rozwiązania w żaden sposób nie wpływała na kształt funkcji aproksymującej. Problem powstający przy aproksymacji wyjścia wielomianem ilustruje rysunek 6.1.



Rysunek 6.1.: Porównanie aproksymacji wyjścia za pomocą wielomianów i za pomocą kolokacji wyjścia.

Dla ustalenia uwagi rozważmy następującą formułę określającą pochodną wyjścia systemu (6.1)

$$\dot{y}(t) = (\nabla h(x(t)))^T f(x(t), u(t)) \quad (6.4)$$

Gdybyśmy znali wartość tej pochodnej, moglibyśmy sformułować pewien rodzaj interpolacji Hermite'a. Trudność polega na tym, że pochodna (6.4) wymaga wiedzy na temat  $x(t)$ , co samo w sobie jest celem estymacji stanu. Możemy jednak wykorzystać estymatę  $\hat{x}(t)$  wyliczoną w kroku predykcji. Konsekwencją takiego działania jest jednak wprowadzenie błędu estymacji do procesu aproksymacji wyjścia.

Podsumowując można wprowadzić dodatkową informację o aproksymowanej funkcji - jej pochodną - ale tym samym wprowadza się błąd estymacji. Co więcej, należy zwrócić uwagę, że pomiary wyjścia będą zazwyczaj zaburzone pewnymi zakłóceniami, więc będą się różnić od faktycznego wyjścia z systemu. Dodatkowo należy rozważyć jaki rodzaj regularności jest wymagany od aproksymującej funkcji. Zazwyczaj wystarczającym wymaganiem będzie jej ciągłość. Jest ona konieczna, bo tylko wtedy metody estymacji dla układów ciągłych będą niosły ze sobą sensowne rezultaty. Z drugiej strony potrzebna jest pewna swoboda w kształtowaniu aproksymacji. Taką swobodę, kosztem regularności, oferują funkcje sklepane. Poszukujemy aproksymacji wyjścia o następujących cechach (jak wspomniano wcześniej określonych arbitralnie):

- jest funkcją sklepaną trzeciego stopnia na czterech węzłach,
- jest ciągła,
- wpasowuje się w pomiary w sensie najmniejszych kwadratów,
- jej pochodne (odpowiednio jednostronne) wpasowują się w estymowane pochodne wyjścia w sensie najmniejszych kwadratów w punktach pomiarów, oraz w punktach leżących w połowie przedziału między pomiarami.

Wymienione warunki tworzą problem programowania kwadratowego z ograniczeniami. Można udowodnić poniższy

**Lemat 6.1.** *Niech pomiary wyjścia systemu 6.1  $y(t_j)$  są znane dla  $j = 0, 1, 2, 3$ , a pochodne wyjścia  $\hat{y}'(t_k)$  są estymowane dla  $k = 0, 1/2, 1, \dots, 3$ . Istnieje wtedy funkcja aproksymująca na przedziale  $t \in [t_0, t_3]$  taka, że:*

- *jest funkcją sklejaną trzeciego stopnia na węzłach  $t_j$  for  $j = 0, 1, 2, 3$ ,*
- *jest ciągła,*
- *wpasowuje się w pomiary w sensie najmniejszych kwadratów,*
- *jej pochodne (odpowiednio jednostronne wpasowują się w estymowane pochodne wyjścia w sensie najmniejszych kwadratów w punktach  $t_k$  gdzie  $k = 0, 1/2, 1, \dots, 3$ .*

*Co więcej funkcja ta jest jedyna*

Warto zaznaczyć, że  $\hat{y}'$  nie oznacza pochodnej funkcji aproksymującej, a jedynie estymatę pochodnej wyjścia systemu w oparciu o informacje z kroku predykcji.

*Dowód.* Do dowodu sformułujemy metodę kolokacji wyjścia jako problem programowania kwadratowego z ograniczeniami i pokażemy jego wypukłość.

W związku z tym, że będą potrzebne trzy wielomiany ich współczynniki zostaną oznaczone jako  $a_i$  gdzie  $i = 1, 2, \dots, 12$ . Rozważamy dopasowanie funkcji aproksymującej do pomiarów

$$\begin{aligned} a_1 t_0^3 + a_2 t_0^2 + a_3 t_0 + a_4 - y(t_0) &= 0 \\ a_1 t_1^3 + a_2 t_1^2 + a_3 t_1 + a_4 - y(t_1) &= 0 \\ a_9 t_2^3 + a_{10} t_2^2 + a_{11} t_2 + a_{12} - y(t_2) &= 0 \\ a_9 t_3^3 + a_{10} t_3^2 + a_{11} t_3 + a_{12} - y(t_3) &= 0 \end{aligned} \tag{6.5}$$

oraz dopasowanie pochodnych

$$\begin{aligned} 3a_1 t_0^2 + 2a_2 t_0 + a_3 - \hat{y}'(t_0) &= 0 \\ 3a_1 t_{1/2}^2 + 2a_2 t_{1/2} + a_3 - \hat{y}'(t_{1/2}) &= 0 \\ 3a_1 t_1^2 + 2a_2 t_1 + a_3 - \hat{y}'(t_1) &= 0 \\ 3a_5 t_1^2 + 2a_6 t_1 + a_7 - \hat{y}'(t_1) &= 0 \\ 3a_5 t_{3/2}^2 + 2a_6 t_{3/2} + a_7 - \hat{y}'(t_{3/2}) &= 0 \\ 3a_5 t_2^2 + 2a_6 t_2 + a_7 - \hat{y}'(t_2) &= 0 \\ 3a_9 t_2^2 + 2a_{10} t_2 + a_{11} - \hat{y}'(t_2) &= 0 \\ 3a_9 t_{5/2}^2 + 2a_{10} t_{5/2} + a_{11} - \hat{y}'(t_{5/2}) &= 0 \\ 3a_9 t_3^2 + 2a_{10} t_3 + a_{11} - \hat{y}'(t_3) &= 0 \end{aligned} \tag{6.6}$$

Dodatkowo nakładamy nienaruszalne więzy ciągłości

$$\begin{aligned} a_1 t_1^3 + a_2 t_1^2 + a_3 t_1 + a_4 - a_5 t_1^3 - a_6 t_1^2 - a_7 t_1 - a_8 &= 0 \\ a_5 t_2^3 + a_6 t_2^2 + a_7 t_2 + a_8 - a_9 t_2^3 - a_{10} t_2^2 - a_{11} t_2 - a_{12} &= 0 \end{aligned} \tag{6.7}$$

Chcemy spełnić (6.5) i (6.6) w sensie najmniejszych kwadratów. Dlatego możemy przedstawić równania (6.5) jako problemy minimalizacji (poniżej przykład dla pierwszego równania (6.5))

$$\min_{a_{[1..4]}} a_{[1..4]}^\top B_1(t_0) a_{[1..4]} - 2y(t_0)c_1(t_0)a_{[1..4]} + y(t_0)^2 \tag{6.8}$$

gdzie:

$$B_1(\alpha) = \begin{bmatrix} \alpha^6 & \alpha^5 & \alpha^4 & \alpha^3 \\ \alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 \\ \alpha^4 & \alpha^3 & \alpha^2 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix} \quad (6.9)$$

oraz

$$c_1(\alpha) = \begin{bmatrix} \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix} \quad (6.10)$$

Równania w (6.6) przeformułowane są w sposób analogiczny (poniżej pierwsze równanie (6.6))

$$\min_{a_{[1..4]}} a_{[1..4]}^\top B_2(t_0) a_{[1..4]} - 2y(t_0)c_2(t_0)a_{[1..4]} + \hat{y}'(t_0)^2 \quad (6.11)$$

gdzie:

$$B_2(\alpha) = \begin{bmatrix} 9\alpha^4 & 6\alpha^3 & 3\alpha^2 & 0 \\ 6\alpha^3 & 4\alpha^2 & 2\alpha & 0 \\ 3\alpha^2 & 2\alpha & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (6.12)$$

oraz

$$c_2(\alpha) = \begin{bmatrix} 3\alpha^2 & 2\alpha & 1 & 0 \end{bmatrix} \quad (6.13)$$

Wszystkie problemy minimalizacji takie jak (6.8) i (6.11) można zebrać (z pominięciem stałych)

$$\begin{aligned} \min_a \quad & a^\top \mathcal{B}a + \mathcal{C}a \\ \text{przy} \quad & \mathcal{G}a = 0 \end{aligned} \quad (6.14)$$

przy czym:

$$\mathcal{B} = \begin{bmatrix} B^* & 0 & 0 \\ 0 & B_2(t_1) + B_2(t_{3/2}) + B_2(t_2) & 0 \\ 0 & 0 & B^{**} \end{bmatrix} \quad (6.15)$$

$$B^* = B_1(t_0) + B_1(t_1) + B_2(t_0) + B_2(t_{1/2}) + B_2(t_1)$$

$$B^{**} = B_1(t_2) + B_1(t_3) + B_2(t_2) + B_2(t_{5/2}) + B_2(t_3)$$

$$\mathcal{C} = -2 \begin{bmatrix} C^* & \hat{y}'(t_1)c_2(t_1) + \hat{y}'(t_{3/2})c_2(t_{3/2}) + \hat{y}'(t_2)c_2(t_2) & C^{**} \end{bmatrix} \quad (6.16)$$

$$C^* = y(t_0)c_1(t_0) + y(t_1)c_1(t_1) + \hat{y}'(t_0)c_2(t_0) + \hat{y}'(t_{1/2})c_2(t_{1/2}) + \hat{y}'(t_1)c_2(t_1)$$

$$C^{**} = y(t_2)c_1(t_2) + y(t_3)c_1(t_3) + \hat{y}'(t_2)c_2(t_2) + \hat{y}'(t_{5/2})c_2(t_{5/2}) + \hat{y}'(t_3)c_2(t_3)$$

$$\mathcal{G} = \begin{bmatrix} c_1(t_1) & -c_1(t_1) & 0 \\ 0 & c_1(t_2) & -c_1(t_2) \end{bmatrix} \quad (6.17)$$

Łatwo sprawdzić, że macierze  $B_1 \geq 0$  i  $B_2 \geq 0$ , zaś macierz (6.15) jest liniową kombinacją o dodatnich współczynnikach macierzy dodatnio określonych. Prowadzi to do wniosku, że  $\mathcal{B}$  jest macierzą dodatnio określoną, a problem minimalizacji (6.14) jest wypukły (zob. Wierzbicki, Szymanowski i Findeisen, 1980; Nocedal i Wright, 2006).  $\square$

*Uwaga 6.1.* Warto zauważyć, że sformułowanie (6.14) problemu kolokacji wyjścia prowadzi do konstruktywnego sposobu wyznaczania jego rozwiązania. Można je uzyskać rozwiązując poniższy układ równań liniowych

$$\begin{bmatrix} \mathcal{B} & \mathcal{G}^\top \\ \mathcal{G} & 0 \end{bmatrix} \begin{bmatrix} a \\ \mu \end{bmatrix} = \begin{bmatrix} -\mathcal{C} \\ 0 \end{bmatrix} \quad (6.18)$$

gdzie  $\mu$  jest wektorem mnożników Lagrange'a związanych z ograniczeniami równościowymi. Ze względu na lemat 6.1 układ (6.18) posiada jednoznaczne rozwiązanie.  $\blacksquare$



*Uwaga 6.2.* Warto zaznaczyć, że dla równoodległych pomiarów dogodniej jest wyznaczać współczynniki funkcji sklepanych dla małych wartości  $t_i$  (na początku osi czasu), a później dokonać stosownej translacji. W przeciwnym razie macierz układu (6.18) staje się bardzo szybko źle uwarunkowana. Metoda kolokacji wyjścia została opublikowana w pracy (Baranowski, 2008c). ■

### 6.1.3. Adaptacyjne kolokacje wyjścia

Metoda kolokacji wyjścia nie jest pozbawiona wad. Dla ustalenia uwagi, przez siatkę będziemy rozumieć ciąg pomiarów służących za węzły metody kolokacji.

Po pierwsze - siatka węzłów kolokacji jest ustalona i wybrana a priori, co stwarza konieczność osobnych rozważań dla każdego systemu z osobna. Co więcej, nawet jeżeli  $\hat{x}$  zbiega do  $x$  i nie ma zakłóceń, nie można zagwarantować zbieżności aproksymacji wyjścia do prawdziwego wyjścia. Co więcej waga dopasowania pomiarów i estymat pochodnych jest taka sama nie weryfikując poprawności estymaty.

Problem siatki węzłów można rozwiązać bazując na zmiennokrokowych metodach numerycznych o gęstym wyjściu takich jak RK 3/8 (3.16). Opierając się na takim sposobie generacji siatki można dokonać pewnych usprawnień metody. Po pierwsze należy rozważyć problem poprawności estymaty. Oczywiście jest, że jeżeli estymata z poprzedniego kroku jest błędna, powinno się bardziej bazować na pomiarach do konstrukcji aproksymacji. Załóżmy, że  $\hat{x}(t_0)$  jest estymatą stanu na początku okna uzyskaną z poprzedniej iteracji algorytmu 1. Wykorzystując ją jako warunek początkowy w metodzie (3.16) można określić rozwiązanie na całym oknie. Z tego rozwiązania, uzyskać zaś można ciąg kroków (siatkę), oraz wartości stanu w nich określone. Ponieważ w każdym kroku wyliczane są wartości prawej strony równania różniczkowego należy je zachować. Można je wykorzystać razem z równaniem wyjścia i (6.4) do obliczenia stosownych wartości estymowanego wyjścia i jego pochodnej.

Za pomocą gęstego wyjścia, można wyliczyć wartości wyjścia w punktach pomiaru. Z ich pomocą można sformułować wskaźnik poprawności estymacji

$$err = \frac{\sqrt{\sum_{i=1}^N (y(t_i) - \bar{y}(t_i))^2}}{\sqrt{\sum_{i=1}^N (y(t_i))^2}} \quad (6.19)$$

Gdzie  $\bar{y}$  jest estymowanym wyjściem z poprzedniej iteracji,  $t_i$  są chwilami pomiarów,  $N$  jest liczbą pomiarów na okno. Z użyciem  $err$  można sformułować współczynnik wagowy

$$W = \begin{cases} err & \text{dla } err \leq 1 \\ 1 & \text{w przeciwnym przypadku} \end{cases} \quad (6.20)$$

Oznaczmy węzły siatki uzyskane z wykorzystania (3.16) jako  $\tau_j$ , niech  $\hat{y}$  będzie funkcją aproksymującą,  $\bar{y}$  i  $\dot{\bar{y}}$  to odpowiednio wartości wyjścia i wartości pochodnych wyjścia uzyskane z poprzedniej iteracji.  $M$  to liczba punktów w siatce. Można teraz szukać  $\hat{y}$  w formie funkcji sklepanej trzeciego stopnia minimalizującej funkcjonal

$$\begin{aligned} & (1 - W) \sum_{j=1}^M \left( (\bar{y}(\tau_j) - \hat{y}(\tau_j))^2 + (\dot{\bar{y}}(\tau_j) - \dot{\hat{y}}(\tau_j))^2 \right) + \\ & + W \sum_{i=1}^N (y(t_i) - \hat{y}(t_i))^2 \end{aligned} \quad (6.21)$$

przy  $\forall 2 \leq i \leq N - 1$

$$\begin{aligned} \hat{y}(\tau_i) &= \hat{y}(\tau_{i+1}) \\ \dot{\hat{y}}(\tau_i) &= \dot{\hat{y}}(\tau_{i+1}) \end{aligned}$$

Jak można zauważyć, waga  $W$  oraz zmieniająca się siatka dostosowują problem optymalizacji do dynamiki systemu i poprawności estymaty, stąd też nazwa **adaptacyjne kolokacje wyjścia**. Jak można zauważyć  $\hat{y}$  będzie ciągłą i różniczkowalną funkcją na całym oknie. Co więcej, gdy estymata staje się bliższa stanowi,

aproksymacja staje się coraz bliższa prawdziwemu wyjściu, ponieważ staje się ona gęstym wyjściem dla wyjścia systemu. Z drugiej strony, gdy błąd względny estymacji jest zbyt duży, część odpowiadająca za estymatę jest całkowicie usunięta z wyznaczania aproksymacji wyjścia.

*Uwaga 6.3.* Warto zwrócić uwagę, że niejako w domyśle założono, że albo sygnał sterujący jest ciągły, albo też pochodną wyjścia da się przedstawić w formie

$$\dot{y}(t) = g(x(t))$$

Zapewnia to, że wyjście jest klasy  $C^1$ . Nie jest to całkowicie konieczne, ale brak tego założenia wymaga pewnych dodatkowych działań. Rozwiązanie musi być wyznaczane na przedziałach ciągłości sterowania, zaś w punktach nieciągłości sterowania muszą wystąpić stosowne skoki w pochodnej wyjścia. Jeżeli sterowanie w (6.4) występuje liniowo skoki można dokładnie wyliczyć, jeżeli jednak sterowanie występuje w bardziej uwikłany sposób również wartości skoków muszą być estymowane. ■

Metoda adaptacyjnych kolokacji wyjścia została opublikowana w pracy (Baranowski, 2008a).

## 6.2. Metody przybliżone dla układów liniowych

W ramach tej sekcji będziemy rozważać systemy postaci

$$\begin{aligned}\dot{x} &= Ax \\ y &= Cx\end{aligned}\tag{6.22}$$

przy czym para  $(C; A)$  jest obserwowalna. Wyniki z tej sekcji pochodzą z prac (Baranowski, 2007a,b).

### 6.2.1. Ciągłe obserwatory asymptotyczne

Dla ciągłych liniowych systemów dynamicznych bardzo dobrze znany jest obserwator Luenbergera (np. Kaczorek, 1977, str. 157 oraz Baranowski et al., 2007, str. 141. Dla systemów postaci (6.22) przyjmuje on postać

$$\dot{\hat{x}}(t) = A\hat{x}(t) + G(y(t) - C\hat{x}(t))\tag{6.23}$$

Przy czym  $G$  dobieramy tak aby  $\operatorname{Re} \lambda(A - GC) < 0$  (warto zaznaczyć, że takie  $G$  istnieje, jeżeli  $(C, A)$  jest wykrywalna (zob. Mitkowski, 1991, str. 90)). W takim przypadku  $\hat{x}(t) \rightarrow x(t)$  gdy  $t \rightarrow \infty$ . Jeżeli dysponujemy aproksymacją wyjścia  $\hat{y}(t)$  dla  $t \in [t_k, t_{k+m}]$ , możemy wprowadzić ją do obserwatora zamiast wyjścia.

$$\dot{\hat{x}}(t) = A\hat{x}(t) + G(\hat{y}(t) - C\hat{x}(t))\tag{6.24}$$

Rozwiązując równanie obserwatora na stosownym przedziale (albo na całym oknie albo tylko na odcinku od poprzedniego pomiaru do obecnego) uzyskujemy wartość stanu  $\hat{x}(t_{k+m})$  do wykorzystania w nowym kroku predykcji.

### 6.2.2. Estymacja $L^2$ -optymalna

Poniższa metoda jest zastosowaniem sposobu odtwarzania stanu opisanego w przykładzie 4.1 na stronie 42. Rozważmy ciągłe wyjście  $y$  na przedziale aproksymacji  $t \in [t_k, t_{k+m}]$ . Ponieważ funkcja ta jest ciągła i ograniczona, możemy stwierdzić, że  $y \in L^2([t_k, t_{k+m}], \mathbb{R}^r)$ . Co więcej, możemy zauważyć, że na tym samym przedziale

$$\hat{y}(t) = y(t) + \chi(t)\tag{6.25}$$

gdzie  $\chi$  odpowiada za błąd aproksymacji, przy czym  $\hat{y}, \chi \in L^2([t_k, t_{k+m}], \mathbb{R}^r)$ . Chcemy wyznaczyć  $\hat{x}(t_k)$ . Rozwiązaniem optymalnie przybliżonym jest

$$\hat{x}(t_k) = M^{-1} \int_{t_k}^{t_{k+m}} e^{At^T} C^T \hat{y}(t) dt \quad (6.26)$$

$$M = \int_{t_k}^{t_{k+m}} e^{At^T} C^T C e^{At} dt \quad (6.27)$$

przy czym  $M$ , czyli Gramian obserwowalności, jest odwracalny wtedy i tylko wtedy, gdy system jest obserwowalny.

*Uwaga 6.4.* Metoda ta okazuje się szczególnie korzystna, jeżeli aproksymacja wyjścia ma postać (6.3). Powoduje to, że

$$\hat{x}(0) = M^{-1}(p_0 I_0 + \dots + p_m I_m) \quad (6.28)$$

przy czym  $I_m$  można wyrazić rekurencyjnie przez całkowanie przez części. Jeżeli  $A$  jest macierzą nieosobliwą postać obserwatora upraszcza się jeszcze bardziej ponieważ (dla  $m = 3$ )

$$\begin{aligned} I_0 &= A^{-1}(e^{Ah_k} - I) \\ I_1 &= h_k A^{-1} e^{Ah_k} + A^{-1} I_0 \\ I_2 &= h_k^2 A^{-1} e^{Ah_k} + 2A^{-1} I_1 \\ I_3 &= h_k^3 A^{-1} e^{Ah_k} + 3A^{-1} I_2 \end{aligned}$$

gdzie  $h_k = t_{k+m} - t_k$ . ■

## 6.3. Metody przybliżone dla układów nieliniowych

### 6.3.1. Ciągłe nieliniowe obserwatory asymptotyczne

Ponieważ dynamika nieliniowa jest dużo bogatsza od liniowej, tak też sposobów projektowania ciągłych obserwatorów jest dużo więcej. Generalnie można przyjąć, że obserwator nieliniowy dla systemu

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \quad (6.29)$$

ma postać

$$\dot{\hat{x}} = f(\hat{x}, u) + g(\hat{x}, u, y) \quad (6.30)$$

Zazwyczaj postać funkcji  $g$  jest stosunkowo prosta i zawiera iloczyn wyrażenia  $y - h(\hat{x})$  ze stałą macierzą, lub ew. pewną funkcją  $\hat{x}$ .

Aproksymację wyjścia wykorzystuje się tu analogicznie jak w układach liniowych, tj. wstawiając  $\hat{y}$  w miejsce  $y$  w równaniach obserwatora. Poniżej przedstawiono kilka przykładów obserwatorów nieliniowych.

#### Obserwator Thau

Najprostszym i jednocześnie historycznie pierwszym obserwatorem dla układów nieliniowych jest obserwator Thau (Thau, 1973) (nazywany także obserwatorem lipshitzowskim) (zob. także Corless, Walcott i Żak, 1986; Yu, 2003; Baranowski et al., 2007, str. 142). Rozważany jest system postaci:

$$\begin{aligned} \dot{x} &= Ax + f(x) + Bu \\ y &= Cx \end{aligned}$$

gdzie para  $(C; A)$  jest obserwowalna, zaś  $f(x)$  spełnia warunek Lipschitza ze stałą  $L$  oraz  $f(0) = 0$ . Proponowany jest obserwator postaci:

$$\dot{\hat{x}} = (A - GC)\hat{x} + f(\hat{x}) + Gy + Bu \quad (6.31)$$

Przy czym dobierając  $G$  stosuje się poniższe

**Twierdzenie 6.2.**  $G$  dobrane tak, żeby spełnić równanie Lapunowa:

$$(A - GC)^T P + P(A - GC) = -2Q$$

gdzie  $P$  symetryczna dodatnio określona,  $Q$  symetryczna dodatnio określona oraz:

$$\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} > L$$

gwarantuje asymptotyczne zanikanie błędu estymacji obserwatora (6.31).

*Dowód.* Dowód przeprowadza się analogicznie do dowodu tw. A.6 na str. 99. Zobacz także (Corless et al., 1986).  $\square$

### Obserwator o dużym wzmocnieniu (High-Gain)

Konstrukcję obserwatora typu High-Gain omówiono dokładniej na stronie 64. Wyraża się on wzorem

$$\dot{\hat{x}} = f(\hat{x}, u) + Q(\hat{x})^{-1} G_{\epsilon}(y - h(\hat{x}))$$

gdzie  $Q(x)$  jest dana przez (5.18). Tutaj jedynie uzupełnimy, że zazwyczaj wzmocnienie obserwatora dobiera się w sposób następujący

$$G_{\epsilon} = \begin{bmatrix} \frac{g_1}{\epsilon} \\ \vdots \\ \frac{g_n}{\epsilon^n} \end{bmatrix}$$

gdzie  $0 < \epsilon < 1$ , natomiast współczynniki  $g_1, \dots, g_n$  dobrane są tak, aby wielomian

$$W(s) = s^n + g_1 s^{n-1} + \dots + g_n$$

był hurwitzowski (zwróćmy uwagę, że jest to wielomian charakterystyczny macierzy  $A_b - GC_b$  gdzie  $G = [g_1, \dots, g_n]^T$  natomiast  $A_b$  i  $C_b$  są odpowiednimi macierzami Brunowskiego (5.22) i (5.24)). Aby zapewnić globalną stabilność obserwatora High-Gain zakłada się zazwyczaj globalny warunek Lipschitza i korzysta się z rozumowania takiego jak w dowodzie twierdzenia A.6. Innym sposobem jest gwarantowanie tzw. semi-globalnej stabilności, czyli możliwości dowolnego powiększania obszaru atrakcji za pomocą zmiany parametrów obserwatora (w szczególności  $\epsilon$ ). (Przykład zastosowania obserwatora typu High-Gain zob. Baranowski i Piątek, 2008; Baranowski, Piątek i Piłat, 2008b).

### Obserwator z nieliniowością wyjścia (Output injection)

Bardzo korzystną sytuacją, przy projektowaniu nieliniowego obserwatora jest sytuacja, gdy rozpatrywany system nieliniowy ma postać:

$$\begin{aligned} \dot{x} &= Ax + f(Cx, u) \\ y &= Cx \end{aligned} \tag{6.32}$$

gdzie para  $(C; A)$  jest obserwowalna. Dla takiego systemu można skonstruować obserwator następującej postaci:

$$\dot{\hat{x}} = A\hat{x} + f(y, u) + G(y - C\hat{x}) \tag{6.33}$$

W takiej sytuacji równanie dynamiki błędu estymacji ma postać:

$$\dot{\epsilon} = (A - GC)\epsilon$$

czyli jest liniowe. W ten sposób uzyskuje się bardzo łatwo nieliniowy obserwator o wykładniczym zanikaniu błędu do zera. Ze względu na tę właściwość podejmowane były próby opracowania algorytmu umożliwiającego

sprowadzenie systemu nieliniowego do postaci 6.32. W pracach (Corless et al., 1986) i (Keller, 1987) można znaleźć takie algorytmy, obydwie o dużej komplikacji obliczeń. Należy jednak zwrócić uwagę, że algorytm z pracy (Corless et al., 1986) ma bardzo ograniczoną przydatność ze względu na konieczność rozwiązywania układów nieliniowych równań różniczkowych. Przykład zastosowania metody Kellera znaleźć można w pracy (Baranowski, 2006a) dobór parametrów obserwatora zob. (Baranowski et al., 2008a).

### Obserwator „ślizgowy” (Sliding mode)

Interesującym sposobem konstrukcji obserwatora są tzw. obserwatory „ślizgowe”, które oprócz wzmocnienia błędu estymacji, zawierają jeszcze dodatkowy człon o charakterze przekaźnikowym, który odpowiada za sprowadzanie systemu na zadaną tzw. *powierzchnię ślizgową*. Obserwator taki można sformułować (zob. np. Alessandri, 2000) dla systemu:

$$\begin{aligned}\dot{x} &= Ax + f(x, u) \\ y &= Cx\end{aligned}$$

przy czym  $f$  spełnia warunek Lipschitza ze stałą  $L$ , para  $(C; A)$  jest obserwowalna. Proponowany jest obserwator o następującej strukturze:

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}, u) + G(y - C\hat{x}) + S_{obs}(\hat{x}, y) \quad (6.34)$$

$G$  jest stałą macierzą taką, że  $\lambda(A - GC) < 0$ , zaś

$$S_{obs}(\hat{x}, y) = \begin{cases} \frac{P^{-1}C^T(y - C\hat{x})}{\|y - C\hat{x}\|} & \text{dla } \|y - C\hat{x}\| > \varepsilon \\ \frac{P^{-1}C^T(y - C\hat{x})}{\varepsilon} & \text{dla } \|y - C\hat{x}\| \leq \varepsilon \end{cases}$$

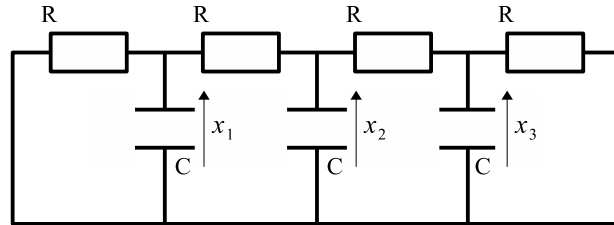
gdzie  $\varepsilon$  to amplituda warstwy brzegowej, dla której chcemy uniknąć gwałtownego przełączania się członu przekaźnikowego.

**Twierdzenie 6.3** (Alessandri, 2000). *Jeżeli istnieją  $G$  takie, że  $\lambda(A - GC) < 0$ , macierz  $Q = Q^T > 0$  oraz  $P = P^T > 0$  będące rozwiązaniem równania Riccatiego:*

$$(A - GC)^T P + P(A - GC) + LPP + I = -Q$$

*to błąd estymacji  $\varepsilon = x - \hat{x}$  obserwatora (6.34) zbiega wykładniczo do zera.*

## 6.4. Przykłady



Rysunek 6.2.: Układ drabinkowy z  $n = 3$

**Przykład 6.1** (Przybliżona estymacja  $L^2$  optymalna dla układu RC). Do ilustracji działania metod przybliżonej estymacji ciągłego stanu w układach liniowych wykorzystano układ łańcuchowy typu RC z  $n = 3$

(rysunek 6.2, przyjmujemy  $RC=1$ ) opisany równaniami (6.22) z macierzami

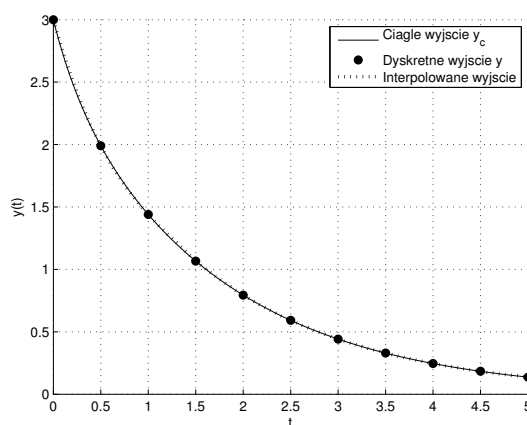
$$A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad (6.35)$$

$$C = [1 \ 0 \ 0] \quad (6.36)$$

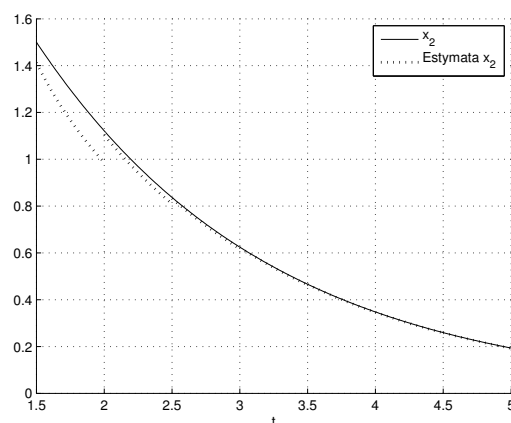
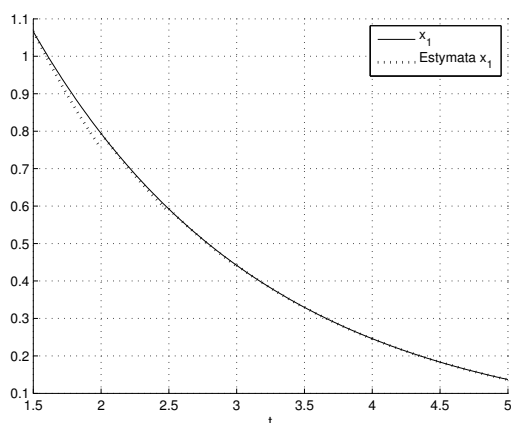
System ten jest asymptotycznie stabilny, a jego wartościami własnymi są

$$\lambda_i = -2 \left( 1 - \cos \left( \frac{i\pi}{4} \right) \right)$$

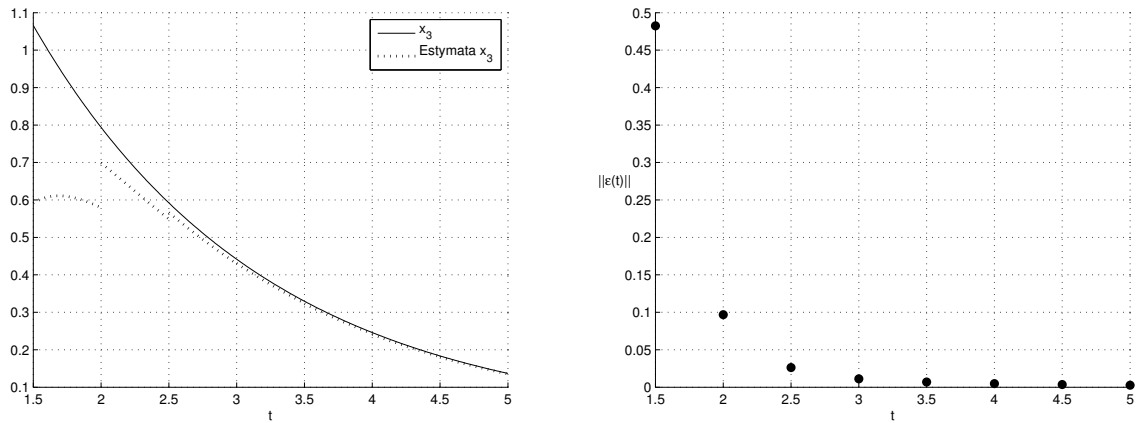
dla  $i = 1, 2, 3$  oraz obserwowalny. System ten analizowano między innymi w Mitkowski (1976); Baranowski (2006b). Do aproksymacji wyjścia wykorzystano wielomian (6.3) stopnia trzeciego. Aproksymację wyjścia ilustruje rysunek 6.3. Działanie metod przybliżonego odtwarzania stanu ilustrują rysunki 6.4-6.7.



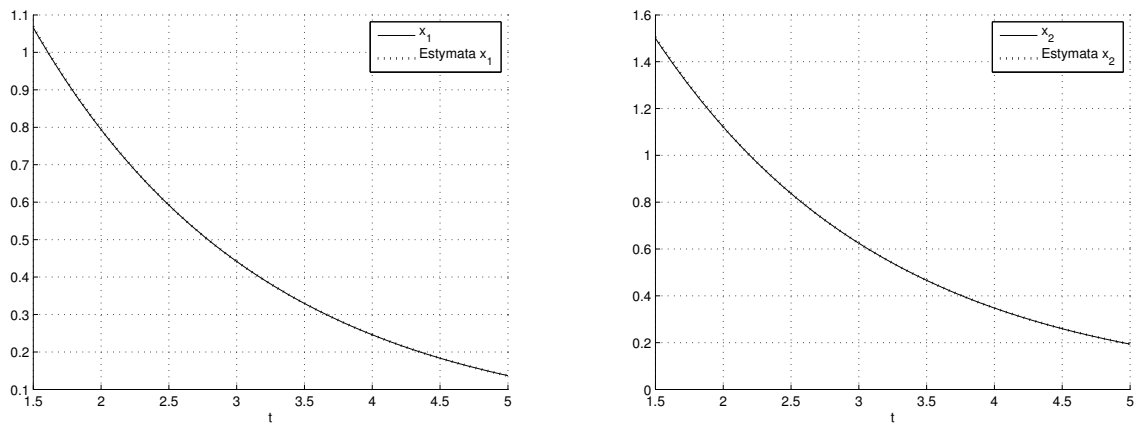
Rysunek 6.3.: Aproksymacja wyjścia w układzie drabinkowym



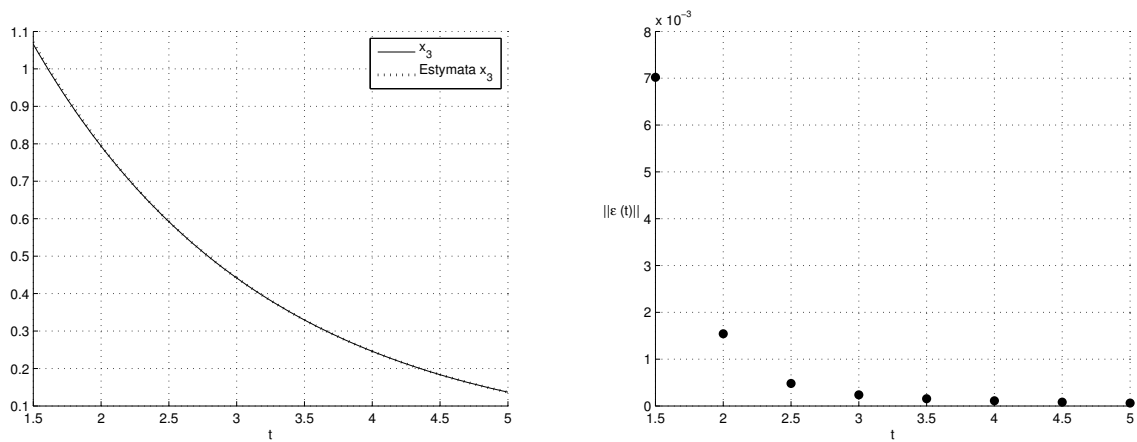
Rysunek 6.4.: Przybliżone odtwarzanie pierwszej (z lewej) i drugiej (z prawej) zmiennej stanu w układzie drabinkowym - obserwator asymptotyczny



Rysunek 6.5.: Przybliżone odtwarzanie trzeciej zmiennej stanu (z lewej) oraz norma błędu estymacji (z prawej) w układzie drabinkowym - obserwator asymptotyczny



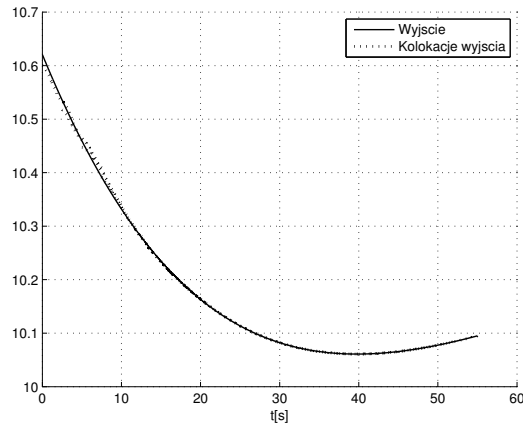
Rysunek 6.6.: Przybliżone odtwarzanie pierwszej (z lewej) i drugiej (z prawej) zmiennej stanu w układzie drabinkowym - obserwator  $L^2$  optymalny



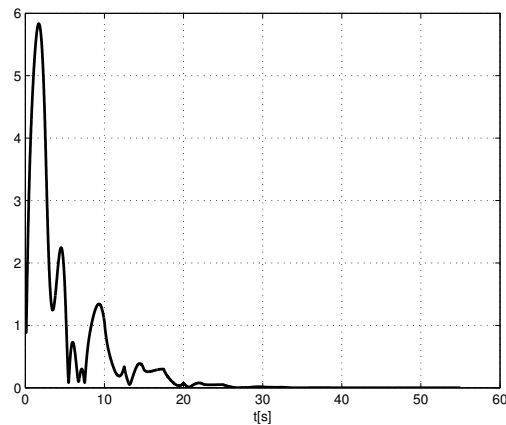
Rysunek 6.7.: Przybliżone odtwarzanie trzeciej zmiennej stanu (z lewej) oraz norma błędu estymacji (z prawej) w układzie drabinkowym - obserwator  $L^2$  optymalny

**Przykład 6.2** (Zastosowanie kolokacji wyjścia do estymacji stanu w układzie zbiorników wodnych). Do

ilustracji metody kolokacji wyjścia wykorzystano model układu zbiorników opisany wcześniej w przykładach 5.1 i 5.3. Do estymacji stanu w oparciu o aproksymowane wyjście wykorzystano obserwator typu Thau. Celowo zastosowano tak prosty obserwator, aby unaocznić podatność obserwacji stanu na błędy przejściowe (rysunek 6.9), nawet przy dobrze przybliżonym wyjściu (rysunek 6.8). Przykład pochodzi z pracy (Baranowski i Mitkowski, 2008).



Rysunek 6.8.: Aproksymacja wyjścia z dolnego zbiornika przy pomocy metody kolokacji wyjścia



Rysunek 6.9.: Norma euklidesowa błędu estymacji stanu w układzie zbiorników wodnych wykonana przy pomocy obserwatora typu Thau i kolokacji wyjścia



## Rozdział 7.

### Podsumowanie

Rozprawa doktorska stanowi efekt pracy badawczej autora nad zagadnieniami odtwarzania ciągłego stanu systemów dynamicznych z dyskretnych danych pomiarowych. Niektóre rezultaty zawarte w rozprawie były przedstawiane przez autora na konferencjach naukowych bądź też publikowane w czasopismach naukowych (zob. prace Baranowski, 2008c; Baranowski i Tutaj, 2007, 2008b; Baranowski i Piątek, 2008; Baranowski i Tutaj, 2008a; Baranowski, 2007a,c, 2008a,b, 2007b,d; Mitkowski i Baranowski, 2007; Baranowski i Mitkowski, 2008, 2009; Baranowski, 2009b,a) Znaczna część tych wyników została rozszerzona i uzupełniona w niniejszej pracy. Według autora najważniejsze oryginalne wyniki pracy to:

- Opracowanie ogólnego algorytmu odtwarzania ciągłego stanu z dyskretnych danych pomiarowych (str. 13)
- Udowodnienie lematu 2.1 dającego oszacowania błędu estymacji (str. 15).
- Udowodnienie twierdzenia 2.3 precyzującego stabilność algorytmu (str. 18)
- Przeprowadzenie analizy własności rozdzielnosci algorytmu i udowodnienie koniecznych twierdzeń (str. 21)
- Sformułowanie definicji obserwowalności (definicja 2.7 str. 19), oraz pokazanie jej równoważności z definicją klasyczną dla układów liniowych stacjonarnych (twierdzenie 2.5 str 19).
- Opracowanie metody RK3/8 - zmiennokrokowej metody czwartego rzędu, z estymatorem błędu rzędu trzeciego o gęstym wyjściu (str. 29).
- Podanie oszacowań błędu przejściowego predykcji oraz udowodnienie twierdzenia 3.1 (str. 35).
- Udowodnienie twierdzenia 3.3 o wpływie zakłóceń na błąd predykcji (str. 38).
- Opracowanie efektywnych wzorów i algorytmu wyliczania parametrów obserwatora  $C^0$  optymalnego (str. 47)
- zastosowanie metody obserwacji  $C^0$  optymalnej do analizy dynamiki populacyjnej (str. 54)
- zastosowanie metod odtwarzania ciągłego stanu w układzie parabolicznym (str. 56)
- Opracowanie sposobu wyliczania macierzy Jakobiego występującej w obserwatorze Newtona za pomocą równania wariacyjnego (str. 68) oraz w oparciu o wariację funkcjonału (str. 69)
- wyliczenie gradientów wskaźnika jakości (5.89) na dwa sposoby (str. 75 i 76)
- Wyliczenie Hesjanu wskaźnika jakości (5.89) z wykorzystaniem równania wariacyjnego (str. 79)
- Zastosowanie metodyki odtwarzania stanu do modelu metabolizmu cukru pacjenta OIOM (str. 81)
- Zastosowanie metody MHE do odtwarzania stanu w obiekcie rzeczywistym (str. 83)
- Opracowanie metody kolokacji wyjścia (str. 87) i adaptacyjnych kolokacji wyjścia (str. 89)

## Dodatek A.

### Istotne twierdzenia

#### A.1. Lematy i twierdzenia pomocnicze

**Lemat A.1** (Gronwalla-Bellmana). *Niech  $k$  będzie funkcją mierzalną, nieujemną, ograniczoną na przedziale  $[t_0, t_1]$ , a  $l$  funkcją niemalejącą. Jeżeli  $v$  jest funkcją rzeczywistą, całkowalną na  $[t_0, t_1]$  i taką, że dla prawie każdego  $t \in [t_0, t_1]$*

$$v(t) \leq l(t) + \int_{t_0}^t k(s)v(s)ds \quad (\text{A.1})$$

*to dla  $t \in [t_0, t_1]$  spełniających (A.1) zachodzi*

$$v(t) \leq \exp\left(\int_{t_0}^t k(s)ds\right) l(t) \quad (\text{A.2})$$

*W szczególności (A.2) zachodzi prawie wszędzie na  $[t_0, t_1]$ . (zob. np. Bellman i Cooke, 1963 str. 59, Zabczyk, 1991, str. 92)*

**Twierdzenie A.2.** *Następujące pojęcia są sobie równoważne:*

1. System

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t), \quad t \geq 0 \quad x(t) \in \mathbb{R}^n$$

*jest obserwowalny.*

2. Para  $(C; A)$  jest obserwowalna.

3. Rząd

$$\text{Rank}(S) = \text{Rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n$$

4. Para  $(A^\top; C^\top)$  jest sterowalna.

5. Dla  $t > 0$

$$M_t = \int_0^t e^{-sA^\top} C^\top C e^{-sA} ds > 0$$

6.

$$X_{no} = \{h \in \mathbb{R}^n : Ce^{At}h = 0 \text{ dla } t \geq 0\} = \text{Ker}(S) = \text{Ker}(M_t) = \bigcap_{i=1}^n \text{Ker}(CA^{i-1}) = 0$$

7. Jeżeli  $Ce^{At}x^0 = 0$  dla  $t \geq 0$ , to  $x^0 = 0$ .

(zob. Wonham, 1979; Mitkowski, 1991, str.87)

**Twierdzenie A.3** (O obserwowalności układu zdyskretyzowanego). *Dane są macierze  $A \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{m \times n}$  oraz okres próbkowania  $h > 0$ . Niech para  $(C; A)$  jest obserwowalna oraz  $\lambda_i, \lambda_j \in \lambda(A)$ . Jeżeli*

$$\forall i \neq j: \text{Re}(\lambda_i - \lambda_j) = 0 \quad \Rightarrow \quad h \text{Im}(\lambda_i - \lambda_j) \neq 2\pi k \quad k = \pm 1, \pm 2, \dots$$

to para  $(C; A_d)$  jest obserwowalna, przy czym  $A_d = \exp(Ah)$ . (zob. np. Mitkowski, 1991, str. 141)

**Lemat A.4.** Niech  $P \in \mathbb{R}^{n \times n}$  jest macierzą nieosobliwą ( $\det P \neq 0$ ). Macierz  $Q = P^T P$  jest ściśle dodatnio określona ( $Q > 0$ ).

*Dowód.* Ponieważ macierz  $P$  jest nieosobliwa, więc jej kolumny tworzą układ  $n$  wektorów liniowo niezależnych

$$P = \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} \quad (\text{A.3})$$

Macierz  $Q$  można więc zapisać jako

$$Q = \begin{bmatrix} p_1^T \\ p_2^T \\ \vdots \\ p_n^T \end{bmatrix} \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} = \begin{bmatrix} (p_1, p_1) & (p_1, p_2) & \dots & (p_1, p_n) \\ (p_2, p_1) & (p_2, p_2) & \dots & (p_2, p_n) \\ \vdots & \vdots & \ddots & \vdots \\ (p_n, p_1) & (p_n, p_2) & \dots & (p_n, p_n) \end{bmatrix} \quad (\text{A.4})$$

Z (A.4) wynika, że  $Q$  jest macierzą Grama. Macierz Grama jest zawsze dodatnio określona. Wynika to z tego, że dla macierzy Grama  $\mathbb{R}^{n \times n} \ni G = HH^T$  mamy  $x^T G x = (H^T x, H^T x) = |H^T x|^2 \geq 0$  dla każdego  $x \in \mathbb{R}^n$  (zob. np. Levine, 1996, str. 44). Wektory  $p_1, p_2, \dots, p_n$  są liniowo niezależne więc  $\det(Q) \neq 0$  (zob. np. Turowicz, 1995, str. 94). Macierz dodatnio określona ma nieujemne wartości własne, natomiast macierz nieosobliwa ma wartości własne różne od zera. Wynika stąd, że macierz  $Q$  ma dodatnie wartości własne, więc jest macierzą ściśle dodatnio określoną.  $\square$

**Twierdzenie A.5** (Gerszgorina). Niech  $A = [a_{ij}]$  będzie macierzą o elementach zespolonych  $a_{ij}$ . Oznaczmy przez  $K_i$  koło leżące na płaszczyźnie zespolonej o środku  $a_{ii}$  oraz o promieniu

$$r_i = \sum_{j=1, j \neq i}^n |a_{ij}|$$

Każda wartość własna macierzy  $A$  leży wewnątrz lub na okręgu któregoś koła  $K_1, K_2, \dots, K_n$ . (zob. np. Turowicz, 1967, str. 50).

## A.2. Stabilizacja regulatorem liniowym

Rozważane są systemy nieliniowe o następującej strukturze

$$\dot{x}(t) = Ax(t) + Bu(t) + f(x(t)) \quad (\text{A.5})$$

przy czym  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^r$ ,  $B \in \mathbb{R}^{n \times r}$ ,  $A \in \mathbb{R}^{n \times n}$ . Funkcja  $f$  jest ciągła oraz

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (\text{A.6})$$

$$f(0) = 0 \quad (\text{A.7})$$

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n \quad (\text{A.8})$$

Poniższe twierdzenie opracowane zostało w oparciu o wyniki z prac (Golubev, Krishchenko i Tkachev, 2005; Gibson, 1968; Mitkowski, 1996; Górecki, 1974; Ogata, 1974; Kaczorek, 1977; Bublik i Kiriczenko, 1975)

**Twierdzenie A.6.** System (A.5) jest globalnie wykładniczo stabilizowany przez regulator

$$u(t) = Kx(t) \quad (\text{A.9})$$

wtedy gdy istnieją  $K \in \mathbb{R}^{r \times n}$  i ściśle dodatnio określona macierz  $Q = Q^T \in \mathbb{R}^{n \times n}$ , dla których:

- $\operatorname{Re} \lambda(A + BK) < 0$

- spełniona jest nierówność

$$\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(H)} > L \quad (\text{A.10})$$

gdzie  $H$  jest ściśle dodatnio określonym rozwiązaniem równania Lapunowa

$$(A + BK)^T H + H(A + BK) = -Q \quad (\text{A.11})$$

*Dowód.* Dowód rozpoczniemy od wykazania asymptotycznej stabilności. Przyjmijmy funkcjonal Lapunowa w postaci

$$V(z) = z^T H z \quad (\text{A.12})$$

gdzie  $H$  jest rozwiązaniem (A.11). Pochodna funkcjonału Lapunowa po trajektoriach systemu zamkniętego złożonego z (A.5) i (A.9) przyjmie postać

$$\dot{V}(x(t)) = 2x(t)^T H \left( (A + BK)^T x(t) + f(x(t)) \right) = -x(t)^T Q x(t) + 2x(t)^T H f(x(t))$$

z (A.7) i (A.8) mamy

$$\|f(x)\| \leq L\|x\|$$

co pozwala na oszacowanie

$$\dot{V}(x(t)) \leq -\lambda_{\min}(Q)\|x(t)\|^2 + 2\lambda_{\max}(H)L\|x\|^2 \leq (2\lambda_{\max}(H)L - \lambda_{\min}(Q))\|x(t)\|^2$$

łatwo zauważyć, że jeżeli spełniona jest nierówność (A.10) to pochodna funkcjonału Lapunowa jest szacowana z góry przez iloczyn liczby ujemnej i normy  $x(t)$ . Ponieważ funkcjonal (A.12) jest dodatnio określony, oraz  $V(z) \rightarrow \infty$  gdy  $\|z\| \rightarrow \infty$  to z II metody Lapunowa wynika globalna asymptotyczna stabilność systemu (A.5) z regulatorem (A.9).

Do dowodu wykładniczej stabilności skorzystamy z oszacowania

$$\lambda_{\min}(H)\|z\| \leq V(z)$$

z którego wynika

$$\|z\| \leq \frac{1}{\lambda_{\min}(H)} V(z)$$

stąd mamy

$$\dot{V}(x(t)) \leq (2\lambda_{\max}(H)L - \lambda_{\min}(Q))\|x(t)\|^2 \leq \frac{2\lambda_{\max}(H)L - \lambda_{\min}(Q)}{\lambda_{\min}(H)} V(z)$$

co w oparciu o (Mitkowski, 1996, s. 90) pozwala na oszacowanie

$$V(x(t)) \leq V(x(0)) \exp \left( t \frac{2\lambda_{\max}(H)L - \lambda_{\min}(Q)}{\lambda_{\min}(H)} \right)$$

co w konsekwencji prowadzi do

$$\|x(t)\| \leq \sqrt{\frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}} \exp \left( t \frac{2\lambda_{\max}(H)L - \lambda_{\min}(Q)}{2\lambda_{\min}(H)} \right) \|x(0)\|$$

co przy spełnionym (A.10) oznacza wykładniczą stabilność. □

## Dodatek B.

### Stabilność wejście-stan (Input-to-state stability)

Wyniki zawarte w tym dodatku pochodzą z prac (Sontag, 1995; Sontag i Wang, 1995). W ramach tego dodatku rozważane będą systemy postaci

$$\dot{x}(t) = f(x(t), u(t)) \quad (\text{B.1})$$

gdzie  $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  jest różniczkowalna w sposób ciągły, oraz  $f(0, 0) = 0$ . Wejście do systemu  $u$  jest mierzalną i lokalnie istotnie ograniczoną funkcją czasu. Będziemy zakładać, że funkcje te należą do przestrzeni  $L^\infty([0, \infty), \mathbb{R}^m)$ . Dodatkowo oznaczymy  $x(t, \xi, u)$  jako trajektorię systemu (B.1) z warunkiem początkowym  $x(0) = \xi$  i wejściem  $u$ . Będziemy wykorzystywać także funkcje następujących klas

- $\gamma: [0, +\infty) \rightarrow [0, +\infty)$  należy do klasy  $\mathcal{K}$ , jeżeli jest ściśle rosnąca, ciągła i  $\gamma(0) = 0$
- $\gamma \in \mathcal{K}^\infty$  jeżeli  $\gamma \in \mathcal{K}$  oraz  $\gamma(s) \rightarrow \infty$  gdy  $s \rightarrow \infty$
- $\beta: [0, +\infty) \times [0, +\infty) \rightarrow [0, +\infty)$  należy do  $\mathcal{KL}$  jeżeli dla każdego ustalonego  $t \geq 0$ ,  $\beta(\cdot, t) \in \mathcal{K}$  oraz dla każdego ustalonego  $s \geq 0$ ,  $\beta(s, t)$  maleje do zera gdy  $t \rightarrow \infty$

**Definicja B.1.** System (B.1) jest stabilny wejście-stan (*input/state stable*, posiada własność ISS) jeżeli istnieje funkcja  $\beta \in \mathcal{KL}$  oraz funkcja  $\gamma \in \mathcal{K}$ , takie że dla każdego  $u \in L^\infty([0, \infty), \mathbb{R}^m)$  i każdego  $\xi \in \mathbb{R}^n$  zachodzi

$$|x(t, \xi, u)| \leq \beta(|\xi|, t) + \gamma(\|u\|_{L^\infty}) \quad (\text{B.2})$$

dla każdego  $t \geq 0$ .

Warto zauważyć, że ze względu na przyczynowość, ta sama definicja zachodzi gdy zastąpić  $\|u\|_{L^\infty}$  przez  $\|u_t\|_{L^\infty}$ , gdzie

$$u_t(s) = \begin{cases} u(s) & \text{dla } s \leq t \\ 0 & \text{dla } s > t \end{cases} \quad (\text{B.3})$$

Definicja B.1 jest nieliniowym uogólnieniem oszacowania

$$|x(t)| \leq be^{-\alpha t}|\xi| + c\|u\|_{L^\infty} \quad (\text{B.4})$$

które zachodzi dla systemów liniowych  $\dot{x} = Ax + Bu$  gdy  $\text{Re } \lambda(A) < 0$ . W analizie własności ISS, szczególną wagę posiadają tzw. funkcjonały ISS-Lapunowa

**Definicja B.2.** Gładka funkcja  $V: \mathbb{R}^n \rightarrow [0, +\infty)$  nazywana jest funkcjonałem ISS-Lapunowa, dla systemu (B.1) jeżeli istnieją funkcje  $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$  oraz  $\alpha_3, \chi \in \mathcal{K}$ , takie że dla każdego  $\xi \in \mathbb{R}^n$

$$\alpha_1(|\xi|) \leq V(\xi) \leq \alpha_2(|\xi|) \quad (\text{B.5})$$

oraz dla każdego  $\xi \in \mathbb{R}^n$  i każdego  $\mu \in \mathbb{R}^m$ , takiego że  $|\xi| \geq \chi(|\mu|)$

$$\nabla V(\xi) \cdot f(\xi, \mu) \leq -\alpha_3(|\xi|) \quad (\text{B.6})$$

*Uwaga B.1.* Warto zauważyć, że dla zerowego wejścia, funkcjonal ISS-Lapunowa jest funkcjonałem Lapunowa dla systemu  $\dot{x} = f(x, 0)$ . ■

Zachodzi następujące twierdzenie

**Twierdzenie B.1.** Dla systemu (B.1) następujące własności są równoważne

1. Posiada własność ISS
2. Posiada funkcjonal ISS-Lapunowa
3. Jest odpornie stabilny, tj. istnieją funkcja  $\rho \in \mathcal{K}_\infty$  oraz funkcja  $\beta \in \mathcal{KL}$ , że dla każdego sterowania  $u(t) = k(t, x(t))$ , dla którego

$$|k(t, \xi)| \leq \rho(|\xi|)$$

(dla prawie wszystkich  $t \geq 0$ ) wszystkie trajektorie systemu zamkniętego  $\dot{x}(t) = f(x(t), k(t, x(t)))$  spełniają

$$|x(t)| \leq \beta(|x(0)|) \quad \forall t \geq 0$$

4. Istnieją funkcje  $\beta_0, \beta_1 \in \mathcal{KL}$  oraz funkcja  $\gamma \in \mathcal{K}$  takie, że dla każdego  $\xi \in \mathbb{R}^n$  oraz wejścia  $u$  zachodzi

$$|x(t, \xi, u)| \leq \beta_0(|\xi|, t) + \beta_1(\|u_T\|_{L^\infty}, t - T) + \gamma(\|u^T\|_{L^\infty}) \quad (\text{B.7})$$

dla każdego  $0 \leq T \leq t$ , przy czym  $u_T$  jest dane przez (B.3) zaś  $u^T = u - u_T$ .

5. Istnieje funkcja  $\gamma \in \mathcal{K}$  taka, że dla każdego  $\varepsilon > 0$  zachodzi

a) Istnieje  $\delta > 0$  taka że

$$|x(t, \xi, u)| \leq \varepsilon \quad \forall t \geq 0 \quad (\text{B.8})$$

dla wszystkich wejść  $u$  oraz warunków początkowych  $\xi$  takich, że  $|\xi| < \delta$  i  $\|u\|_{L^\infty} < \delta$

b) dla każdego  $r > 0$  istnieje  $T > 0$  takie, że dla każdego wejścia  $u$ , gdy  $|\xi| < r$  oraz  $t \geq T$ , zachodzi

$$|x(t, \xi, u)| \leq \varepsilon + \gamma(\|u\|_{L^\infty}) \quad (\text{B.9})$$

6. Jest słabo odpornie stabilny (patrz Sontag i Wang, 1995).

## Bibliografia

- Ackermann, J. (1976). *Regulacja Impulsowa*. Wydawnictwa Naukowo Techniczne, Warszawa.
- Alamir, M. (1999). Optimization based non-linear observers revisited. *International Journal of Control* **72**(13), str. 1204–1217.
- Albuquerque, J. S. i Biegler, L. T. (1994). Decomposition algorithms for on-line estimation with nonlinear models. W *Proceedings of the American Control Conference*. Baltimore, Maryland, str. 1444–1448.
- Alessandri, A. (2000). Design of sliding mode observers and filters for nonlinear dynamic systems. W *Proceedings of 39th IEEE Conference on Decision and Control*. str. 3892–3897.
- Angeles, M., Astorga, C., Alvarado, V. i Vela, L. (2004). Nonlinear estimation via discrete and continuous-discrete observers. Application to CSTR's. W *XI Congreso Latinoamericano de Control Automático CLCA'04*. Ciudad de La Habana, Cuba.
- Arcak, M. (2002). A global separation theorem for a new class of nonlinear observers. W *Proc. 41st IEEE Conf. Decision Contr.*, Las Vegas, Nevada, str. 676–681.
- Astorga, C., Othman, N., Othman, S., Hammouri, H. i McKenna, T. (2002). Nonlinear continuous-discrete observers: application to emulsion polymerization reactors. *Control Engineering Practice* **10**(1), str. 3–13(11).
- Atassi, A. i Khalil, H. (2000). Separation results for the stabilisation of nonlinear systems using different high-gain observer designs. *Systems & Control Letters* **39**(3), str. 183–191.
- Bania, P. (2008). *Algorytmy sterowania optymalnego w nieliniowej regulacji predykcyjnej*. Rozprawa doktorska, Akademia Górniczo-Hutnicza im. Stanisława Staszica, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków. Promotor: A. Korytowski.
- Baranowski, J. (2006a). Projektowanie obserwatora dla silnika szeregowego prądu stałego. *Półrocznik AGH AUTOMATYKA* **10**(1), str. 33–52.
- Baranowski, J. (2006b). State estimation in linear multi output systems - design example and discussion of optimality. *Półrocznik AGH AUTOMATYKA* **10**(2), str. 119–131.
- Baranowski, J. (2007a). Approximate continuous state estimation in linear dynamical systems. W *Materiały IX Międzynarodowych Warsztatów Doktranych OWD 2007*, tom 1. Wisła, 20-23.10.2007, str. 11–16.
- Baranowski, J. (2007b). Approximate continuous state estimation in linear dynamical systems. W *PPEEm'2007 : materiały XII sympozjum „Podstawowe Problemy Energoelektroniki, Elektromechaniki i mechatroniki”*, tom 2. Wisła, 9-12.12.2007, str. 293–298.
- Baranowski, J. (2007c). Continuous state estimation from discrete output measurements in a linear parabolic system. W *Materiały IX Międzynarodowych Warsztatów Doktranych OWD 2006*. Wisła, 20-23.10.2007, str. 17–22.
- Baranowski, J. (2007d). Csedom for water tank system. W *PPEEm'2007 : materiały XII sympozjum „Podstawowe Problemy Energoelektroniki, Elektromechaniki i mechatroniki”*, tom 2. Wisła, 9-12.12.2007, str. 299–304.
- Baranowski, J. (2008a). Adaptive output collocations. W *Materiały X Międzynarodowych Warsztatów Doktranych OWD 2008*. Wisła, str. 519–524.

- Baranowski, J. (2008b). Continuous moving horizon estimation in a hydraulic system. W *Materiały X Międzynarodowych Warsztatów Doktranckich OWD 2008*. Wisła, str. 525–530.
- Baranowski, J. (2008c). Output collocation method for continuous state estimation from discrete output measurements in linear dynamical systems. *Automatyka* **12**.
- Baranowski, J. (2009a). Optimal continuous estimation of diabetes population dynamics. W *Proceedings of XI International PhD Workshop, OWD*. Wisła, str. 146–151.
- Baranowski, J. (2009b). Patient glycemic state estimation in intensive care units. W *Proceedings of XI International PhD Workshop, OWD*. Wisła, str. 152–157.
- Baranowski, J., Długosz, M. i Mitkowski, W. (2008a). Remarks about dc motor control. *Archives of Control Sciences* **18(LIV)**(3), str. 289–322.
- Baranowski, J., Hajduk, K., Korytowski, A., Mitkowski (red.), W. i Tutaj, A. (2007). *Teoria sterowania. Materiały pomocnicze do ćwiczeń laboratoryjnych*. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków.
- Baranowski, J. i Mitkowski, W. (2006). Optimal discretization step for second order siso systems. W *Materiały XXIX Międzynarodowej konferencji z podstaw elektrotechniki i teorii obwodów IC-SPETO*. str. 415–418.
- Baranowski, J. i Mitkowski, W. (2008). New output approximation method for continuous state estimation. W *Materiały XXXI Międzynarodowej konferencji z podstaw elektrotechniki i teorii obwodów IC-SPETO*. Ustroń, 28-31.05.2008, str. 95–96. Extended version on CD.
- Baranowski, J. i Mitkowski, W. (2009). Optimal glucose level regulation for intensive care unit patients: Csedom approach. W *Papers. 24th IFIP TC7 Conference on System Modelling and Optimization Optimal glucose level regulation for Intensive Care unit patients: CSEDOM approach*. Organized by: Herramientas Gerenciales, Palais Rouge, Buenos Aires, Argentina, str. 46–47.
- Baranowski, J. i Piątek, P. (2008). Nonlinear dynamical feedback for motion control of magnetic levitation system. W *Proceedings of 13th International Power Electronics and Motion Control Conference EPE-PEMC*. Poznań, 1-3 September 2008, str. 1469–1476.
- Baranowski, J., Piątek, P. i Piłat, A. (2008b). Nonlinear observer design for the magnetic levitation system. W Malinowski, K. i Rutkowski, L., red., *Recent advances in control and automation*, rozdz. 2. Akademicka Oficyna Wydawnicza EXIT, str. 73–82.
- Baranowski, J. i Tutaj, A. (2007). Continuous state estimation in water tank system. W *CMS'07 : Computer Methods and Systems, 21–23 November 2007, Kraków, Poland*. Kraków, 21-23.11.2007, str. 373–378.
- Baranowski, J. i Tutaj, A. (2008a). Comparison of continuous state estimation algorithms in a water tank system. W Malinowski, K. i Rutkowski, L., red., *Recent advances in control and automation*, rozdz. 2. Akademicka Oficyna Wydawnicza EXIT, str. 63–72.
- Baranowski, J. i Tutaj, A. (2008b). State and parameter estimation in a hydraulic system — moving horizon approach. W *Proceedings of 13th International Power Electronics and Motion Control Conference EPE-PEMC*. Poznań, 1-3 September 2008, str. 1455–1462.
- Barker, R. (1954). The Pulse Transfer Function and its Application to Sampling Servo Systems. *Trans. ASME* przedrukowane przez R. Oldenburg (red.) pt. Frequency Response. The MacMillan Company, New York 1956.
- Bauer, W., Pluciennik, P., Baranowski, J. i Mitkowski, W. (2009). LQ based glucose stabilisation for Intensive Care Unit patients. W Tadeusiewicz, R., Ligeza, A., Mitkowski, W. i Szymkat, M., red., *CMS'09: Computer Methods And Systems*. Kraków, Poland, str. 307–312.



- Bellman, R. i Cooke, K. L. (1963). *Differential-Difference Equations*, tom 6 z serii *Mathematics in Science and Engineering*. Academic Press, London.
- Bıyık, E. i Arcak, M. (2006a). Hybrid Newton Observer Design Using The Inexact Newton Method and GMRES. W *Proceedings of the 2006 American Control Conference*. Minneapolis, Minnesota, USA, str. 3334–3339.
- Bıyık, E. i Arcak, M. (2006b). A hybrid redesign of Newton observers in the absence of an exact discrete-time model. *Systems & Control Letters* **55**(6), str. 429–436.
- Bołtjański, W. (1971). *Metody matematyczne sterowania optymalnego*. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Boutayeb, A. i Chetouani, A. (2007). A population model of diabetes and pre-diabetes. *Int. J. of Comp. Math* **84**(1), str. 57–66.
- Boutayeb, A., Twizell, E., Achouayb, K. i Chetouani, A. (2004). A mathematical model for the burden of diabetes and its complications. *BioMed Eng OnLine*.
- Bronsztejn, I., Siemiendajew, K., Musiol, G. i Mühlig, H. (2004). *Nowoczesne Kompendium Matematyki*. PWN, Warszawa.
- Bublik, B. i Kiriczenko, H. (1975). *Osnowy teorii uprząwlenija*. Kijów, Szkoła Wyższa.
- Busłowicz, M. (2005). Osiągalność i sterowalność liniowych układów dyskretnych z opóźnieńiami zmiennych stanu. W *Mat. XV Krajowej Konferencji Automatyki*, tom 1. Warszawa, str. 209–212.
- Busłowicz, M. (2007). Robust stability of positive discrete-time linear systems with multiple delays with linear unity rank uncertainty structure or non-negative perturbation matrices. *Bulletin of the Polish Academy of Sciences Technical Sciences* **55**(1), str. 1–5.
- Butcher, J. C. (2003). *Numerical Methods for Ordinary Differential Equations*. John Wiley and Sons, Ltd.
- Byrski, W. (1993). Integral description of the optimal state observers. W Nieuwenhuis, J., Praagman, C. i Trentelman, H., red., *Proceedings of the second European Control Conference ECC'93, Groningen*, tom 4. str. 1832–1837.
- Byrski, W. (2007). *Obserwacja i sterowanie w systemach dynamicznych*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Kraków.
- Chee, F. i Fernando, T. (2007). *Closed-Loop Control of Blood Glucose*. Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin-Heidelberg.
- Corless, M., Walcott, B. i Żak, S. (1986). Comparative study of non-linear state-observation techniques. *Int. J. Control* **45**(6), str. 2109–2132.
- Cypkin, J. Z. (1965). *Teoria układów impulsowych*. PWN, Warszawa.
- Dalla Mora, M., Germani, A. i Manes, C. (1997). A state observer for nonlinear dynamical systems. *Nonlinear Analysis, Theory, Methods and Applications* **30**, str. 4485–4496.
- Demidowicz, B. (1972). *Matematyczna teoria stabilności*. WNT, Warszawa.
- Engel, K.-J. i Nagel, R. (2000). *One-parameter Semigroups for Linear Evolution Equations*. Nr 194 w Graduate Texts in Mathematics. Springer, New York.
- Eykhoff, P. (1980). *Identyfikacja w układach dynamicznych*. PWN, Warszawa.
- Ferrari-Trecate, G., Mignone, D. i Morari, M. (2002). Moving horizon estimation for hybrid systems. *IEEE Transactions on Automatic Control* **47**(10), str. 1663–1676.

- Fichtenholtz, G. M. (1972). *Rachunek różniczkowy i całkowy*. PWN, Warszawa.
- Fujarewicz, K. i Mazer, S. (2006). Strojzenie ekstrapolatora ułamkowego rzędu. *Zeszyty Naukowe. Automatyka / Politechnika Śląska* **145**, str. 69–74.
- Gauss, K. F. (1809). *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*. Dover, New York. Angielskie tłumaczenie, 2004 *Theory of Motion of the Heavenly Bodies Moving About the Sun in Conic Sections: A Translation of Theoria Motus*.
- Gauthier, J. i Kupka, I. (1992). A separation principle for bilinear systems with dissipative draft. *IEEE Transactions on Automatic Control* **37**(12), str. 1970–1974.
- Gibson, J. E. (1968). *Nieliniowe układy sterowania automatycznego*. WNT, Warszawa.
- Golubev, A., Krishchenko, A. i Tkachev, S. (2001). Separation principle for affine systems. *Diff. Uravn.* **37**(11), str. 1468–1475.
- Golubev, A., Krishchenko, A. i Tkachev, S. (2002). Separation principle for a class of nonlinear systems. W *Prepr. 15th IFAC World Congress*. Barcelona.
- Golubev, A., Krishchenko, A. i Tkachev, S. (2005). Stabilization of nonlinear dynamic systems using the system state estimates made by the asymptotic observer. *Automation and Remote Control* **66**(7), str. 1021–1058. URL <http://www.springerlink.com/index/T363716148553830.pdf>.
- Górecki, H. (1971). *Analiza i synteza układów regulacji z opóźnieniem*. Wydawnictwa Naukowo Techniczne, Warszawa.
- Górecki, H. (1974). *Teoria Sterowania. Część 1. Wykłady na studium doktoranckim w zakresie automatyki i elektrotechniki w roku 1970/71*. Uczelniane Wydawnictwa Naukowo Dydaktyczne AGH, Kraków, 2 wyd.
- Górecki, R. (1997). Uniwersalny dyskretny miernik pozycji. W Mitkowski, W., red., *Prace z Automatyki dedykowane profesorowi Henrykowi Góreckiemu z okazji 70-lecia urodzin*. Uczelniane Wydawnictwa Naukowo Dydaktyczne AGH, str. 145–155.
- Grabowski, P. (1997). Przykład identyfikacji obiektu parabolicznego. Modele z obserwacją/sterowaniem punktowym. W Mitkowski, W., red., *Prace z Automatyki dedykowane profesorowi Henrykowi Góreckiemu z okazji 70-lecia urodzin*. Uczelniane Wydawnictwa Naukowo Dydaktyczne AGH, str. 145–155.
- Guang-Da, H. i Mingzhu, L. (2004). The weighted logarithmic matrix norm and bounds of the matrix exponential. *Linear Algebra and its Applications* **390**, str. 145–154.
- Hairer, E., Nørsett, S. i Wanner, G. (2000). *Solving Ordinary Differential Equations: I Nonstiff problems*. Springer, 2 wyd.
- Haseltine, E. L. i Rawlings, J. B. (2008). A critical evaluation of extended kalman filtering and moving horizon estimation. Rap. tech. 2002-03, Texas-Wisconsin Modeling and Control Consortium.
- Haverbeke, N., Van Herpe, T., Diehl, M., Van den Berghe, G. i De Moor, B. (2008). Nonlinear model predictive control with moving horizon state and disturbance estimation - application to the normalization of blood glucose in the critically ill. W *Proceedings of the IFAC World Congress 2008 (IFAC08)*. Seoul, Korea, str. 9069–9074.
- Hou, M. (2002). Output injection and high-gain observer designs. <http://www.eng.hull.ac.uk/research/control/DAMADICS/3rdws/dso2mh.pdf>.
- Jury, E. I. (1956a). Hidden Oscillations in Sampled-data Control Systems. *Trans. AIEE* **75**, str. 391–395. Part II.

- Jury, E. I. (1956b). Synthesis and Critical Study of Sampled-data Control Systems. *Trans. AIEE* **75**, str. 141–151. Part II.
- Kaczorek, T. (1977). *Teoria układów automatycznej regulacji*. WNT, Warszawa.
- Kaczorek, T. (2007). *Polynomial and Rational Matrices. Applications in Dynamical Systems Theory*. Springer-Verlag, London.
- Keller, H. (1987). Non-linear observer design by transformation into a generalized observer canonical form. *Int J. Control* **46**(6), str. 1915–1930.
- Kobayashi, T. (1980). Discrete-time observability for distributed parameter systems. *International Journal of Control* **31**(1), str. 181–193.
- Kociszewski, R. (2009). *Sterowalność i obserwowalność liniowych stacjonarnych układów dodatnich dyskretnych z opóźnieniami*. Rozprawa doktorska, Politechnika Białostocka; Wydział Elektryczny. Promotor: M. Busłowicz.
- Kołodziej, W. (1982). *Wybrane rozdziały analizy matematycznej*, tom 36 z serii *Biblioteka Matematyczna*. PWN, II poprawione wyd.
- Kovács, L. (2006). Extension of the Bergman minimal model for the glucose-insulin interaction. *PERIODICA POLYTECHNICA SER. EL. ENG.* **50**(1-2), str. 23–32.
- Kudrewicz, J. (1976). *Analiza funkcjonalna dla automatyków i elektroników*. PWN, Warszawa.
- Kulczycki, P. (2005). *Estymatory jądrowe w analizie systemowej*. Wydawnictwa Naukowo Techniczne.
- Kwakernaak, H. i Sivan, R. (1972). *Linear Optimal Control Systems*. Wiley, New York.
- Levine, W. S., red. (1996). *The Control Handbook*. CRC Press and IEEE Press.
- Lin, W. (1994). Bounded smooth state feedback and a global separation principle for non-affine nonlinear systems. *Systems & Control Letters* **26**(1), str. 41–53.
- Loria, A. i Panteley, E. (1999). A separation principle for a class of euler-lagrange systems. W *New Directions in Nonlinear Observer Design*. Springer, London.
- Łukasik, E. (2007). *Iteracyjne metody dla osobliwych nieliniowych układów równań*. Rozprawa doktorska, Uniwersytet Marii Curie-Skłodowskiej; Wydział Matematyki, Fizyki i Informatyki. Promotor: Stanisław Grzegórski.
- Makroglou, A., Li, J. i Kuang, Y. (2006). Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview. *Applied Numerical Mathematics* **56**, str. 559–573.
- Mitkowski, W. (1976). Aproksymacja linii długiej układem drabinkowym. *Arch. Elektrotechniki* **25**(4), str. 943–953.
- Mitkowski, W. (1991). *Stabilizacja systemów dynamicznych*. WNT, Warszawa.
- Mitkowski, W. (1996). *Stabilizacja systemów dynamicznych*. Wydawnictwa AGH, Kraków.
- Mitkowski, W. (2000). *Systemy dynamiczne - Materiały uzupełniające do wykładu*. Wydawnictwa Wydziału Elektrotechniki, Automatyki, Informatyki i Elektroniki AGH, Kraków.
- Mitkowski, W. (2003). Zastosowania metod optymalizacji. W Piela, A., Grosman, F., Kusiak, J. i Pietrzyk, M., red., *Informatyka w technologii metali*, rozdz. 1. Wydawnictwo Politechniki Śląskiej, str. 11–72.
- Mitkowski, W. (2007a). Równania liniowe i ich zastosowania (linear equations and their applications). W Tadeusiewicz, R., red., *Prace Komisji Nauk Technicznych PAU*, tom 2. PAU, str. 253–265.

- Mitkowski, W. (2007b). *Równania macierzowe i ich zastosowania*. Wydawnictwa AGH, Kraków, 2 wyd.
- Mitkowski, W. (2008). Zastosowania równań liniowych w teorii sterowania. W Malinowski, K. i Rutkowski, L., red., *Sterowanie i Automatyzacja: Aktualne problemy i ich rozwiązywanie*, rozdz. 1. Akademicka Oficyna Wydawnicza EXIT, str. 11–22.
- Mitkowski, W. i Baranowski, J. (2007). Estimation of continuous state in linear dynamical systems from discrete output measurements. W *Materiały XXX Międzynarodowej konferencji z podstaw elektrotechniki i teorii obwodów IC-SPETO*. Ustroń, 23-26.05.2007, str. 193–194. Extended version on CD.
- Mlak, W. (1970). *Wstęp do teorii przestrzeni Hilberta*, tom 35 z serii *Biblioteka Matematyczna*. PWN, Warszawa, I wyd.
- Molery, C. i Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review* **45**(1).
- Moraal, P. i Grizzle, J. (1995). Observer design for nonlinear systems with discrete-time measurements. *IEEE Transactions on Automatic Control* **40**(3), str. 395–404.
- Nijmeijer, H. i van der Schaft, A. (1991). *Nonlinear Dynamical Control Systems*. Springer-Verlag.
- Nocedal, J. i Wright, S. (2006). *Numerical Optimization*. Springer, 2 wyd.
- Nowak, M. W. (2007). *Metody identyfikacji układów ciągłych z wykorzystaniem funkcji modulujących i sklejanych i ich zastosowanie w regulatorze adaptacyjnym*. Rozprawa doktorska, Akademia Górniczo-Hutnicza im. Stanisława Staszica, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków. Promotor: W. Byrski.
- Ogata, K. (1974). *Metody przestrzeni stanów w teorii sterowania*. Wydawnictwa Naukowo Techniczne, Warszawa.
- Oprędkiewicz, K. (1997). Przykład identyfikacji obiektu parabolicznego. *Zeszyty Naukowe AGH, Elektrotechnika* **16**(2), str. 99–106.
- Oprędkiewicz, K. (2008). *Praktyczne sterowanie systemami dynamicznymi z widmem punktowym i parametrami przedziałowymi*. Nr 186 w Rozprawy Monografie. Uczelniane Wydawnictwa Naukowo Dydaktyczne AGH, Kraków.
- Papanikolopoulos, N., Khosla, P. i Kanade, T. (1993). Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision. *Robotics and Automation, IEEE Transactions on* **9**(1), str. 14–35.
- Rao, C. V. i Rawlings, J. B. (2002). Constrained Process Monitoring: Moving Horizon Approach. *AIChE Journal* **48**(1), str. 97–109.
- Rao, C. V., Rawlings, J. B. i Lee, J. H. (1999). Stability of constrained linear moving horizon estimation. W *Proceedings of the American Control Conference*. San Diego, California, str. 3387–3391.
- Rao, C. V., Rawlings, J. B. i Mayne, D. Q. (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control* **48**(2), str. 246–258.
- Rawlings, J. B. i Bakshi, B. R. (2006). Particle filtering and moving horizon estimation. *Computers and Chemical Engineering* **30**, str. 1529–1541.
- Rolewicz, S. (1974). *Analiza funkcjonalna i teoria sterowania*. PWN, Warszawa.
- Rudnicki, R. (). *Dynamika populacyjna*. Unpublished. URL [http://www.mimuw.edu.pl/~biolmat/Dynam\\_po.pdf](http://www.mimuw.edu.pl/~biolmat/Dynam_po.pdf). Plik pdf manuskryptu udostępniany przez autora.

- Sastry, S. i Bodson, M. (1994). *Adaptive Control: Stability, Convergence, and Robustness*. Prentice-Hall Advanced Reference Series (Engineering).
- Shampine, L. F., Gladwell, I. i Thompson, S. (2003). *Solving ODEs with MATLAB*. Cambridge University Press.
- Shampine, L. F. i Reichelt, M. W. (1997). The MATLAB ODE Suite. *SIAM Journal on Scientific Computing* **18**(1).
- Simon, D. (2006). *Optimal State Estimation. Kalman,  $H_\infty$  and Nonlinear Approaches*. John Wiley and Sons, Ltd.
- Skruch, P. (2002). Odtwarzanie sygnału analogowego z dyskretnego ciągu próbek. *Automatyka* **6**(2), str. 201–217.
- Skruch, P. (2005). *Stabilizacja liniowych nieskończenie wymiarowych układów oscylacyjnych*. Rozprawa doktorska, Akademia Górniczo-Hutnicza im. Stanisława Staszica, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków. Promotor: W. Mitkowski.
- Sontag, E. (1989). Remarks on stabilisation and input-to-state stability. W *Proc. IEEE Conf Decision and Control*. IEEE Publications, Tampa, str. 1376–1378.
- Sontag, E. (1995). On the input-to-state stability property. *European Journal of Control* **1**, str. 24–36.
- Sontag, E. i Wang, Y. (1995). On characterizations of the input-to-state stability property. *Systems Control Lett.* **24**(5), str. 351–359.
- Stoer, J. (1979). *Wstęp do metod numerycznych*. PWN.
- Stoica, P. i Söderström, T. (1997). *Identyfikacja systemów*. PWN, Warszawa.
- Tenny, M. i Rawlings, J. (2002). Efficient moving horizon estimation and nonlinear model predictive control. *American Control Conference*, str. 4475–4480.
- Thau, F. (1973). Observing the states of nonlinear dynamical systems. *International Journal of Control* **17**(3), str. 471–479.
- Tsinias, J. (1993). Sontag’s ‘input to state stability condition’ and global stabilization using state detection. *Systems & Control Letters* **20**(3), str. 219–226.
- Turowicz, A. (1967). *Geometria zer wielomianów*. PWN, Warszawa.
- Turowicz, A. (1995). *Teoria Macierzy*. Uczelniane Wydawnictwa Naukowo Dydaktyczne AGH, Kraków, 5 wyd.
- Tutaj, A. (2008). *Odporne algorytmy sterowania rozproszonego*. Rozprawa doktorska, Akademia Górniczo-Hutnicza im. Stanisława Staszica, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków. Promotor W. Grega.
- Van Herpe, T., Espinoza, M., Haverbeke, N., De Moor, B. i Van den Berghe, G. (2007). Glycemia prediction in critically ill patients using an adaptive modeling approach. *Journal of Diabetes Science and Technology* **1**(3), str. 348–356.
- Veselić, K. (2003). Bounds for exponentially stable semigroups. *Linear Algebra and its Applications* **358**, str. 309–333.
- Wang, Z., Liu, Z., Ban, X. i Pei, R. (2003). A study on robust estimation via moving horizon strategy. W *Proceedings of the Fourth International Conference on Control and Automation (ICCA'03)*. Montreal, Canada, str. 747–751.

### *Bibliografia*

- Wierzbicki, A., Szymanowski, J. i Findeisen, W. (1980). *Teoria i metody obliczeniowe optymalizacji*. PWN, Warszawa.
- Wonham, W. (1979). *Linear multivariable control: a geometric approach*. Springer, New York, 2 wyd.
- Yu, W. (2003). System identification. [Http://www.ctrl.cinvestav.mx/~yuw/teach.htm](http://www.ctrl.cinvestav.mx/~yuw/teach.htm).
- Zabczyk, J. (1991). *Zarys matematycznej teorii sterowania*, tom 73 z serii *Biblioteka Matematyczna*. PWN, Warszawa.
- Zavala, V. M., Laird, C. D. i Biegler, L. T. (2007). A fast computational framework for large-scale moving horizon estimation. W *Proceedings of the 8th International Symposium on Dynamics and Control of Process Systems*. Cancun, Mexico.

# Skorowidz

- autonomizacja, 40, 45
- błąd estymacji, 15
- CSEDOM, 15
- filtr Kalmana, 53, 64
  - rozszerzony, 70
- funkcjonał
  - jakości, 68, 76
  - liniowy, 48, 49
- gęste wyjście, 28
- globalna asymptotyczna stabilność
  - algorytmu, 17, 18
  - procedury korekcji, 17, 18
- globalna atraktywność
  - algorytmu, 17
  - procedury korekcji, 17
- hiperglikemia pacjentów OIOM, 81
- koła Gerszgorina, 36, 99
- kolokacje wyjścia, 86, 96
  - adaptacyjne, 89
- korekcja, 15
- kowariancja, 53, 70, 72, 73
- MHE, 71, 83
- model populacji cukrzycy, 55
- norma
  - „logarytmiczna”, 34
- obserwator
  - $C^0$ -optymalny, 47
  - $L^2$  optymalny, 90
  - „ślizgowy”, 93
  - dokładny dyskretny, 44
  - dyskretno-ciągły, 43
  - dyskretno-ciągły, nieliniowy, 62, 80
  - High-Gain, 63, 92
  - hybrydowy Newtona, 65, 82
  - Luenbergera, 90
  - najmniejszych kwadratów, 46
  - Newtona, 65
  - output injection, 92
  - sliding mode, 93
  - Thau, 91
- obserwowalność, 19, 40, 64, 70, 98
  - lokalna, 62, 64
  - układu zdyskretyzowanego, 20, 40, 98
- okno pomiarowe, 24
  - przesuwane, 25
  - rozłączne, 25
  - rozszerzane, 25
- operator
  - dualny, 48
  - dysypatywny, 36
  - liniowy, 41, 57
  - ograniczony, 57
  - sprzężony, 42, 48
- półgrupa
  - $C^0$ , 57
  - kontrakcji, 34, 36
  - zweżająca, 34
- pochodna Liego, 63
- predykcja, 15
- przestrzeń
  - $C^0$ , 47, 49
  - $L^2$ , 42, 57, 90
  - $L^\infty$ , 38, 101
  - Banacha, 42
  - dualna, 48
  - Hilberta, 42, 53
  - sprzężona, 48
- równanie
  - Lapunowa, 33
  - Riccatiego, 93
  - wariacyjne, 67, 75, 78, 79
- reguła Bayesa, 72
- rozkład Gaussa, 53, 70, 72
- stała Lipschitza, 12, 15, 23
- stabilność
  - algorytmu, 17, 18
  - procedury korekcji, 17, 18
  - wejście-stan, 21, 101
- system
  - dynamiczny, 12

- liniowy, 12
- paraboliczny, 56
- szerokość okna, 25
  
- układ zbiorników wodnych, 79, 83
  
- własność ISS, 22, 23
- własność rozdzielności, 22–24
- własność zbieżne wejście-zbieżny stan, 22
- wariacja funkcjonału, 68, 76
- wartość oczekiwana, 71
- warunek Lipschitza, 12, 16, 23, 38, 74, 91, 93, 99
- współczynnik tłumienia, 34
  
- zasada rozdzielności, 21



## English abstract

The focus of this thesis is Continuous State Estimation from Discrete Output Measurements (CSEDOM). It is a problem which arises in many practical situations. The main objective of an industrial control system or algorithm is to stabilise a process and at the same time to obtaining the best or reasonably good control quality. For this purposes continuous information of a plant full state is desired as well as good knowledge of its mathematical model and parameters. In spite of these requirements, many limitations exist, especially in chemical, petrochemical, metallurgical industry, to name a few, which limit the amount of information. Instead of a full state information, only incomplete output data are often available. Reasons for that can be of a technological nature if the measurement of a particular state variable is difficult or impossible or else economical if needed sensors are exceptionally expensive. In medical situations, frequency of measurements (examinations) is also limited for health reasons, for instance multiple X-rays can be directly harmful to the patients. Moreover, some measurements cannot be performed continuously, especially when specialist laboratory chemical analyses are employed, and thus have discrete character. Although measurements are known only for discrete moments, a continuous control could usually better the quality of control and hence is highly recommended.

Our focus is on dynamical systems in the following form

$$\dot{x}(t) = f(x(t), u(t)) \quad (\text{B.10})$$

where

- $f: \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$
- $x(t) \in \mathbb{R}^n$ ,
- $u(t) \in \mathbb{R}^r$ ,
- $t \geq 0$
- $x(0) = x_0$
- $f(a, b)$  fulfills global Lipschitz condition for the first argument

$$\begin{aligned} |f(x_1, b) - f(x_2, b)| &\leq L|x_1 - x_2| \\ \forall x_1, x_2 \in X \quad \forall b \in \mathbb{R}^r \end{aligned} \quad (\text{B.11})$$

with Lipschitz constant  $L > 0$  and is continuous for both arguments.

We also consider system output

$$y(t) = h(x(t)) \quad (\text{B.12})$$

where  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ . We assume that output measurements are available only in discrete time instances  $t_i$ , where

$$t_i = t_{i-1} + h_i \quad (\text{B.13})$$

$$t_0 \geq 0 \quad (\text{B.14})$$

$$h_i > 0 \quad (\text{B.15})$$

$$i = 0, 1, 2, 3, \dots \quad (\text{B.16})$$

with

$$h_i \geq \tau_{max} \quad (B.17)$$

$$\tau_{max} = \max_i (t_{i+1} - t_i) < \infty \quad (B.18)$$

Measurements can, but might not be, evenly spaced. In that case

$$h_i = h = \tau_{max} \quad \forall i \quad (B.19)$$

$$t_i = t_0 + ih \quad (B.20)$$

where  $h$  is the sampling period.

For such systems it is possible to propose an algorithm, which presents a general solution to the CSEDOM problem.

1. START

Set  $\hat{x}(t_0)$  as an initial estimate,  $i := 0$

2. PREDICTION

Solve differential equation  $\dot{\hat{x}}(t) = f(\hat{x}(t), u(t))$  (equation of system (B.10)) for  $t \in [t_i, t_{i+1}]$  z warunkiem początkowym  $\hat{x}(t_i)$ .

3. CORRECTION

Using values  $y(t_j)$  for  $j \leq i+1$  ( $j = i+1, i, i-1, \dots, 0$ ) and the values of  $u(t)$  for  $t \leq t_{i+1}$  estimate the  $\hat{x}(t_{i+1})$ .

4.  $i := i+1$ , go to step 2.

The main idea of the algorithm is given in figure 2.1 on page 14. Different aspects of the algorithm are addressed in different parts of the thesis.

Chapter two address certain special general issues regarding the algorithm. Firstly, the stability and attractivity of the algorithm, the stability and attractivity of the correction procedure and their relations to one another are discussed. In particular a lemma giving bounds on continuous estimation error was proven (lemma 2.1 on page 15), and subsequently used for proofs of appropriate stability (theorem 2.3 on page 18) and attractivity (theorem 2.2 on page 17) theorems. Also the problem of observability from discrete output measurements was discussed. A definition was formulated especially for this situation (definition 2.7 on page 19). This definition for linear systems was proven to be equivalent to classical observability (theorem 2.5). Final general problem considering CSEDOM was the separation principle. An example of system not fulfilling the principle was presented and discussed. Some sufficient conditions (theorems 2.70 on page 23 and 2.9 on page 24) for verification of separation principle were proven. The base of these proofs was concept of “input-to-state” stability.

In the third chapter, the problems of prediction were discussed. Firstly a general theory of variable step, dense output Runge-Kutta formulas was presented. As an illustration a method called “RK 3/8” was constructed, giving a fourth order solver with third order error estimator and third order differentiable dense output.

In addition to that, the problem of prediction errors was discussed. It was shown, that even for asymptotically stable, linear systems error dynamics can generate large transitional errors. The analysis of bounds on trajectories was included. Especially interesting problems were considered for linear systems, where bounds on the norm of matrix exponential were discussed. Especially promising are:

- bound based on Lyapunov equation ((3.26) page 33),
- bound based on damping coefficient ((3.30) page 34).

Both of them can be used to estimate the possible maximum of matrix exponential norm (and in consequence maximal error). Moreover one more question was addressed – that is, when the error dynamics is a contraction

semigroup, i.e.

$$\|e^{At}\| \leq 1, \quad t \geq 0 \quad (\text{B.21})$$

A theorem giving a sufficient condition was proven (theorem 3.1 on page 35). The proof was based on the concept of damping coefficient and Gershgorin theorem.

Final part of the third chapter was the problem of disturbances influence on the error of continuous estimate. A bound on the error, using the  $L^\infty$  norm of disturbances was introduced (theorem 3.58 on page 38).

Chapter four was focused on the correction methods in linear systems. A concept of autonomisation was introduced. It was shown, that state estimation problem in systems

$$\begin{aligned} \dot{x} &= Ax + Bu & \text{and} & & \dot{x} &= Ax \\ y &= Cx & & & y &= Cx \end{aligned} \quad (\text{B.22})$$

are equivalent. Additionally it was discussed, that problem of state estimation in linear systems can be interpreted as a solution of a linear equation

$$y = \mathcal{G}x_0 \quad (\text{B.23})$$

where  $\mathcal{G}$  is a linear mapping from initial state space  $X_0$  to output space  $Y$ . The properties and form of the solution depend on the choice of output space  $Y$ .

The main part of the chapter were the algorithms of state estimation. Following algorithms were discussed:

- discrete-continuous observer (page 43),
- exact discrete observer (page 45),
- least squares observer (page 47),
- $C^0$ -optimal observer (page 49),
- hybrid Kalman filter (page 53).

What is especially interesting is the  $C^0$ -optimal observer. This method of estimation is based on choosing the optimal moments of output sampling. The choice of these moments is a consequence of minimisation of the norm of certain linear functional on  $C^0[a, b]$  space. Because this minimisation is a non-smooth, possibly non-convex, non-linear programming problem a form of relaxation is introduced. This relaxation takes form of sequence of sub-problems, all of which allow analytical formula for gradients. Chapter also includes two application examples –  $C^0$ -optimal observer in a population model of diabetes (example 4.2 on page 54) and application of discrete-continuous and exact discrete observers for the approximation of a parabolic system (example 4.3 on page 56).

Fifth chapter focuses on non-linear state estimation problems. Four algorithms are discussed:

- nonlinear discrete-continuous observer (page 64),
- hybrid Newton observer (page 66),
- Moving Horizon Estimation (MHE) (page 73),
- hybrid Extended Kalman Filter (page 71).

These algorithms are based on different approaches to state estimation in non-linear systems. For discrete-continuous observer, these are the concepts of local observability, Lie derivatives and High-gain observer. The main idea is to use the state variable coordinate change, obtained with Lie derivatives, in order to obtain easily tuneable observer. Time varying gain is used, obtained through a solution of “Lyapunov-like” equation.

Hybrid Newton observer is based on the simple concept, that output of the non-linear system can be expressed as a non-linear mapping from initial state space to output space

$$y(t) = h(F(x_0, t)) \quad (\text{B.24})$$

which can be used for construction of system of equations

$$\begin{aligned} y(t_k) &= h(F(w_k, t_k)) \\ y(t_{k+1}) &= h(F(w_k, t_{k+1})) \\ &\vdots \\ y(t_{k+n-1}) &= h(F(w_k, t_{k+n-1})) \end{aligned} \quad (\text{B.25})$$

let us denote

$$Y_k = \begin{bmatrix} y(t_k) \\ y(t_{k+1}) \\ \vdots \\ y(t_{k+n-1}) \end{bmatrix} \quad \text{oraz} \quad H(w_k) = \begin{bmatrix} h(w_k) \\ h(F(w_k, t_{k+1})) \\ \vdots \\ h(F(w_k, t_{k+n-1})) \end{bmatrix} \quad (\text{B.26})$$

We can now solve this system of equations with Newton's method

$$\begin{aligned} w_k^{i+1} &= w_k^i + s_i \\ \frac{d}{dw_k} H(w_k^i) s_i &= Y_k - H(w_k) \\ i &= 0, 1, \dots, d \end{aligned} \quad (\text{B.27})$$

Two methods for obtaining the Jacobian  $\frac{d}{dw_k} H(w_k^i) s_i$  needed for Newton's method were developed. First one is based on the application of variational equation (page 68) and the second one uses the variation of the functional (page 69).

Moving horizon estimation is an optimisation based observer. State is estimated through solving

$$\min_{\hat{x}(t_k)} L(\hat{x}(t_k)) \quad (\text{B.28})$$

$$L(\hat{x}) = |\hat{x}(t_k) - \bar{x}_{k-1}|_{\Pi_k}^2 + \sum_{i=k}^{k+N} |y(t_i) - h(\hat{x}(t_i))|^2 \quad (\text{B.29})$$

s.t.

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \quad (\text{B.30})$$

The form of the performance index can be justified via maximum likelihood method or by deterministic means. In order to improve the optimisation the formulas for gradients  $\nabla L$  and Hessian  $\nabla^2 L$  were determined. For gradients formulas are based on variational equation (page 76) and on the variation of functional (page 76). Hessian formula was obtained through application of variational equation (page 79).

Chapter also include three examples. First one is the application of discrete-continuous observer for estimation in water tank system (example 5.1 on page 79). Second one uses hybrid Newton observer for patient state estimation in model of ICU-patient glycemia (example 5.2 on page 81). The final example is the application of MHE for state estimation in real life laboratory model of water tank system. This example is especially interesting, because it shows remarkable filtering capabilities of MHE (example 5.3 on page 83).

Sixth chapter presents the concept of approximate CSEDOM. The main idea is to construct a continuous approximation of the discrete output. This continuous approximation is then used with classical continuous state estimation techniques. Choice of such approximation is discussed. Besides standard polynomial and spline interpolation two original methods of approximation are presented. Those methods are output collocations (page 87) and adaptive output collocations (page 89). Proposed solutions use the system model and previous state estimates in order to create the approximation, that corresponds well to system dynamics. Also different methods of continuous estimation are presented and discussed. Two examples are included – one linear (page 93) and one nonlinear (page 96) – presenting different issues in approximate CSEDOM.

The thesis concludes with a summary – listing author's original contributions – and appendices including important theorems.

## Streszczenie

W pracy rozważana jest problematyka odtwarzania ciągłego stanu systemów dynamicznych z dyskretnych danych pomiarowych. Jest to problem o dużym praktycznym znaczeniu. W zastosowaniach przemysłowych bardzo często mamy potrzebę stabilizacji przy jednoczesnej wysokiej jakości sterowania. Do tego celu potrzebna jest pełna informacja o stanie przez cały czas trwania procesu. Praktyka jednak pokazuje, że zazwyczaj pomiary mają charakter dyskretny i to tylko pewnych wielkości wyjściowych. Jest to uwarunkowane różnymi czynnikami. Często nie jest możliwe dokonywanie pomiarów wszystkich wielkości charakteryzujących system, ze względu na niemożność lub nieopłacalność instalacji czujnika. Co więcej częstość wykonywania pomiarów jest ograniczona. W przemyśle chemicznym koszt przeprowadzenia analizy produktu bywa niebagatelny (potrzebne są kosztowne odczynniki). W wykorzystaniu systemów wizyjnych potrzeba czasu na obróbkę obrazu. W medycynie ograniczenia narzuca ochrona zdrowia pacjenta, np. nie można za często wykonywać zdjęć rentgenowskich.

W ramach pracy omawiany jest sposób estymacji ciągłego stanu z dyskretnych danych pomiarowych, oparty o cykliczne dokonywanie predykcji stanu i korekcji estymaty. Omawiane są pojęcia stabilności, atraktywności, obserwowalności oraz rozdzielności wraz ze stosownymi twierdzeniami. Przeanalizowane są zagadnienia numerycznej generacji ciągłych rozwiązań, szacowania błędów przejściowych predykcji oraz błędów ciągłej estymaty wywołanych zakłóceniami. Trzy rozdziały poświęcone są metodom korekcji (estymacji stanu) w układach liniowych, nieliniowych oraz metodom przybliżonym dla układów liniowych i nieliniowych. Między innymi omówiono metodę obserwatora  $C^0$  optymalnego, obserwatora Newtona oraz estymacji z ruchomym horyzontem. Wraz z metodami podano różne usprawnienia, np. metody analitycznego wyznaczania pochodnych wskaźnika jakości po warunku początkowych. Działanie algorytmów zilustrowano przykładami różnych systemów liniowych i nieliniowych z różnych dziedzin: medycyny, dynamiki populacyjnej, mechatroniki i innych.