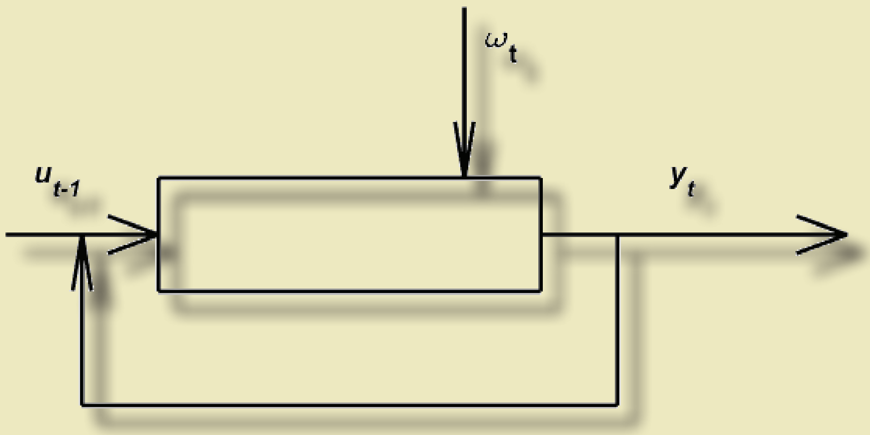
POLITECHNIKA LUBELSKA

MONOGRAFIE

Edward Kozłowski

Optymalne sterowanie dyskretnymi systemami stochastycznymi



Lublin 2018

Optymalne sterowanie dyskretnymi systemami stochastycznymi

Monografie – Politechnika Lubelska

Politechnika Lubelska
Wydział Zarządzania
ul. Nadbystrzycka 38
20-618 Lublin

Edward Kozłowski

Optymalne sterowanie dyskretnymi systemami stochastycznymi

Wydawnictwo Politechniki Lubelskiej
Lublin 2018

Recenzenci:
prof. dr hab. Jerzy Kozicki, Instytut Matematyki UMCS

Publikacja wydana za zgodą Rektora Politechniki Lubelskiej

ISBN: 978-83-7947-316-8

Wydawca: Wydawnictwo Politechniki Lubelskiej
www.biblioteka.pollub.pl/wydawnictwa
Realizacja: Biblioteka Politechniki Lubelskiej
Ośrodek ds. Wydawnictw i Biblioteki Cyfrowej
ul. Nadbystrzycka 36A, 20-618 Lublin
tel. (81) 538-46-59, email: wydawca@pollub.pl
www.biblioteka.pollub.pl
Druk: TOP Agencja Reklamowa Agnieszka Łuczak
www.agencjatop.pl

Elektroniczna wersja książki dostępna w Bibliotece Cyfrowej PL www.bc.pollub.pl
Nakład: 100 egz.

Spis treści

Wstęp

Teoria sterowania jest to nauka o zasadach i metodach sterowania systemami (obiektami, urządzeniami, procesami). Zadania z teorii sterowania polegają na analizie i modelowaniu matematycznym pewnych systemów oraz konstrukcji algorytmu (regulatora) dla realizacji określonego celu. Do modelowania systemu wykorzystujemy wiedzę z różnych dziedzin nauki: fizyki, techniki, chemii, medycyny, socjologii itp., natomiast do konstrukcji optymalnych sterowań wykorzystywane są narzędzia z matematyki, statystyki, rachunku prawdopodobieństwa.

Dość często na zachowanie systemów wpływają zaburzenia zewnętrzne, które powodują, że nie możemy dokładnie przewidzieć reakcji systemu na sterowanie. Ponadto nieznajomość parametrów systemu także uniemożliwia określenie dokładnych sterowań. Zarówno nieznajomość parametrów oraz występowanie zaburzeń zewnętrznych wpływają na wielkość sterowań oraz wielkość wskaźnika jakości. W wielu przypadkach identyfikacja parametrów jest dokonywana poprzez sterowanie systemem podczas realizacji celu. W szczególności stosuje się koncepcję adaptacji, która bazuje na dostrajaniu regulatora w pętli oraz idei działania w celu samouczenia. Wyznaczenie sterowań polega na optymalizacji wskaźnika jakości, który określa cel sterowania. Wskaźnik jakości zależy zarówno od stanów systemu, jak i od sterowań. W przypadku nieznajomości parametrów systemu optymalne sterowanie musi mieć dualną naturę: z jednej strony powinno odpowiednio szybko umożliwiać gromadzenie wiedzy o parametrach, z drugiej zaś realizować cel. Realizacja tych dwóch odrębnych, a jednocześnie współzależnych zadań stanowi istotę problemu adaptacyjnego sterowania. Umiejętne modelowanie systemów technicznych oraz odpowiednie sterowanie nimi zasadniczo jest dużym wyzwaniem dla naukowców, analityków, inżynierów. Ze względu na ogromne potencjalne znaczenie dla zastosowań problem sterowania adaptacyjnego przyciągał uwagę badaczy od dawna. Pierwsze publikacje pojawiły sie pół wieku temu i są znaczone takimi nazwiskami, jak Wiener N., Astrom K., Feldbaum A.A., Bellman R., Kulikowski R., Fleming W.H., Rishel R., Beneš V.E., Karatzas I. Literatura przedmiotu jest ogromna, aspekty praktyczne przedstawiane są w setkach książek i artykułów.

Okazuje się, że nieznajomość horyzontu sterowania również wywiera dość duży wpływ na wartość wskaźnika jakości oraz wielkości sterowań. W niektórych zadaniach przyjęcie ustalonego i niezależnego horyzontu sterowania nie prowadzi do adekwatnego modelu sytuacji. Na przykład w zadaniach typu selflearning, w problemach sztucznej inteligencji, zarządzaniu portfelem inwestycyjnym przyjmuje się, że horyzont działania zależny jest od otrzymanych wyników. Wskazane jest, aby zatrzymać proces w pierwszym momencie uzyskania satysfakcjonujących efektów. Widzimy, że w każdym momencie, oprócz wyznaczenia sterowań, dodatkowo należy podjąć decyzję odnośnie kontynuacji bądź zatrzymania sterowania systemem. W innych przypadkach nie mamy w ogóle żadnego wpływu na horyzont sterowania. Przykładowo czas "życia" produktu, czas bezawaryjnej pracy urządzenia technicznego czy też liczba strat poniesionych w procesie produkcji nie jest znana odgórnie. Powstaje pytanie: jak sterować systemem, jeżeli horyzont nie jest sprecyzowany.

Monografia przedstawia matematyczne aspekty dotyczące wyznaczenia optymalnego sterowania systemami stochastycznymi z czasem dyskretnym. W rozdziale pierwszym przedstawiono podstawowe wiadomości z teorii miary, teorii rachunku prawdopodobieństwa, procesów stochastycznych, teorii filtracji oraz zasad optymalnego zatrzymania procesów losowych. W rozdziale drugim omówiono problematykę adaptacyjnego sterowania systemami stochastycznymi, która polega na gromadzeniu wiedzy o parametrach systemu i jednocześnie realizacji celu. Pokazano również różnice pomiędzy sterowaniami w warunkach pełnej i niepełnej informacji o parametrach systemu. Dodatkowo zwrócono uwagę na wpływ horyzontu sterowania na wskaźnik jakości. W kolejnym rozdziale przedstawiono zagadnienie optymalnego sterowania systemem dla przypadku, gdy horyzont sterowania jest losowy i nie zależy od stanów systemu. W rozdziale czwartym zaprezentowano problem optymalnego zatrzymania dla systemów sterowanych. Zostały zaproponowane dwa sposoby rozwiązania tego problemu. Jeden z nich polega na rozszerzeniu wektora sterowań o dodatkową zmienną decyzyjną dotyczącą zatrzymania systemu (stopowanie przy pomocy sterowania). Drugi ze sposobów bazuje na zastosowaniu optymalnych zasad zatrzymania procesów losowych. W ostatnim rozdziale niniejszego opracowania problem wyznaczenia optymalnego sterowania zastąpiono problemem równoważnym, który polega na wyznaczeniu optymalnej trajektorii dla systemu sterowanego. Pod pojęciem trajektorii rozumie się uporządkowany ciąg stanów systemu. System, który porusza się po optymalnej trajektorii, realizuje ten sam cel jak dla problemu optymalnego sterowania. Wobec powyższego dla każdego z momentów znając optymalne stany systemu możemy określić sterowanie. Naśladując bezpośrednio te stany system realizuje cel sterowania. W każdym z omówionych wyżej zagadnień do wyznaczenia optymalnych sterowań bądź też optymalnej trajektorii wykorzystano programowanie dynamiczne wstecz. Dodatkowo podano algorytmy wyznaczenia optymalnych sterowań i optymalnej trajektorii.

Materiał zawarty w pracy może być pomocny studentom, pracownikom naukowym oraz osobom, które są zainteresowane teorią optymalnego sterowania oraz metodami konstrukcji optymalnych regulatorów dla systemów stochastycznych z czasem dyskretnym.

Rozdział 1

Podstawowe wiadomości

Istnieje dużo znakomitych książek, które szczegółowo omawiają rożnorodne zagadnienia z teorii miary, teorii rachunku prawdopodobieństwa, procesów stochastycznych (np.[26], [27], [29], [47], [50], [75], [79], [84], [88], [90], [95], [96], [112], [113], [115]–[118], [120], [121], [123], [124]). Poniżej krótko przedstawione zostaną tylko niezbędne wiadomości. W celu pogłębienia wiedzy i uzyskania więcej informacji z analizy funkcjonalnej i rachunku prawdopodobieństwa kieruję czytelnika do książek. Poniżej przyjmujemy $\mathbb{N}$ – zbiór liczb naturalnych, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $\mathbb{R}$ – zbiór liczb rzeczywistych, $\mathbb{C}$ – zbiór liczb zespolonych, natomiast $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$.

1.1 Miara. Podstawowe własności miary

Definicja 1.1 *Niech X będzie pewną przestrzenią. Rodzinę $\mathcal{M}$ nazywamy algebrą (ciałem) jeżeli:*

1. $\emptyset, X \in \mathcal{M}$;
2. *jeżeli* $A, B \in \mathcal{M}$, *to* $A \cup B \in \mathcal{M}$, $A \cap B \in \mathcal{M}$;
3. *jeżeli* $A \in \mathcal{M}$, *to* $\bar{A} = X \backslash A \in \mathcal{M}$.

Definicja 1.2 *Pewną rodzinę $\mathcal{F}$ przestrzeni X nazywamy $\sigma-$ algebrą ($\sigma-$ ciałem) jeżeli:*

1. $X \in \mathcal{F}$;
2. *jeżeli* $A \in \mathcal{F}$, *to* $\bar{A} \in \mathcal{F}$;

3. jeżeli $A_1, A_2, ... \in \mathcal{F}$*, to* $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

Ciało $\mathcal{M}$ jest klasą podzbiorów przestrzeni X zamkniętą ze względu na operacje skończonego dodawania, mnożenia i odejmowania zbiorów. $\sigma-$ciało $\mathcal{F}$ jest pewną klasą podzbiorów przestrzeni X zamkniętą ze względu na operacje przeliczalnego dodawania, mnożenia i odejmowania zbiorów.

Na tej samej przestrzeni X mogą być zadane różne $\sigma-$ ciała. Najmniejsze $\sigma-$ ciało $\mathcal{F}_0 = \{\emptyset, X\}$ złożone ze zbioru pustego i przestrzeni X nazywa się trywialnym, natomiast rodzina wszystkich możliwych podzbiorów przestrzeni X postaci $\mathcal{F} = \{A : A \subset X\}$ jest największym $\sigma-$ ciałem.

Uwaga 1.1 *Niech D będzie dowolną rodziną podzbiorów zbioru X. Istnieje wówczas jednoznacznie określone $\sigma-$ ciało zawierające wszystkie możliwe podzbiory zbioru D. Najmniejsze $\sigma-$ ciało zawierające rodzinę D (czyli, generowane przez D) oznaczamy przez $\sigma(D)$.*

Uwaga 1.2 *Niech $\{\mathcal{F}_t\}$ będzie ciągiem $\sigma-$ ciał określonych na przestrzeni X, wtedy $\mathcal{F} = \bigcap\limits_t \mathcal{F}_t$ jest również $\sigma-$ ciałem.*

Definicja 1.3 *Niech X będzie przestrzenią topologiczną. $\sigma-$ ciało zawierające wszystkie zbiory otwarte w przestrzeni X nazywamy $\sigma-$ ciałem zbiorów borelowskich w X i oznaczamy $\mathcal{B}(X)$.*

Definicja 1.4 *Funkcję μ_Z określoną na klasie wszystkich podzbiorów zbioru X o wartościach $\mathbb{R}_+$ spełniającą warunki:*

1. $\mu_Z(\emptyset) \geq 0$;
2. *jeżeli* $A \subset B$*, to* $\mu_Z(A) \leq \mu_Z(B)$;
3. *dla dowolnego ciągu* $\{A_n\}_{n\in\mathbb{N}}$ *podzbiorów zbioru* X

$$\mu_Z\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu_Z(A_n)$$

nazywamy miarą zewnętrzną.

Definicja 1.5 *Każdą funkcję μ określoną na $\mathcal{F}$ o wartościach w $\mathbb{R}_+$ spełniającą warunki:*

1. *dla każdego $A \in \mathcal{F} \quad \mu(A) \geq 0$;*
2. *$\mu(A) < \infty$ przynajmniej dla jednego $A \in \mathcal{F}$;*
3. *jeżeli $A_1, A_2, ... \in \mathcal{F}$ oraz $A_i \cap A_j = \emptyset$ dla dowolnych $i \neq j$, to*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

nazywamy miarą.

Uwaga 1.3 *Jeżeli $\mu(X) < \infty$, to miarę nazywamy miarą skończoną.*

Miara μ określona na $\mathcal{F}$ ma następujące własności:

1. $\mu(\emptyset) = 0$;
2. jeżeli $A \subset B$, to $\mu(A) \leq \mu(B)$;
3. jeżeli $A, B \in \mathcal{F}$, to $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$
4. jeżeli $A_1 \subset A_2 \subset ... \in \mathcal{F}$, to

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n);$$

5. jeżeli $... \subset A_2 \subset A_1 \in \mathcal{F}$, to

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n).$$

Definicja 1.6 *Parę $(X, \mathcal{F})$ nazywamy przestrzenią mierzalną, natomiast trójkę $(X, \mathcal{F}, \mu)$ nazywamy przestrzenią z miarą.*

Definicja 1.7 (Warunek Caratheodory'ego) *Niech μ_Z będzie miarą zewnętrzną na X. Mówimy, że zbiór A spełnia warunek Caratheodory'ego, wtedy i tylko wtedy gdy*

$$\mu_Z(C) = \mu_Z(C \cap A) + \mu_Z(C \backslash A)$$

dla dowolnego $C \subset X$.

Oczywiście każde $\sigma-$ ciało $\mathcal{F}$ można rozszerzyć na zbiory typu $\tilde{A} = A \cup B$, gdzie $A \in \mathcal{F}$, natomiast $B \subset A_0 \in \mathcal{F}$ oraz $\mu_Z(A_0) = 0$. Rodzina $\tilde{\mathcal{F}}$ zawierająca zbiory $\tilde{A}$ jest $\sigma-$ ciałem.

Twierdzenie 1.1 (Twierdzenie Caratheodory'ego) *Jeżeli μ_Z jest miarą zewnętrzną w zbiorze X oraz $\mathcal{F}$ oznacza rodzinę wszystkich podzbiorów zbioru X spełniających warunek Caratheodory'ego, to*

1. *$\mathcal{F}$ jest $\sigma-$ ciałem w X;*
2. *jeżeli $\mu_Z(A) = 0$, to $A \in \mathcal{F}$;*
3. *miara zewnętrzna μ_Z zawężona do $\mathcal{F}$ jest miarą (zupełną).*

Twierdzenie Caratheodory'ego ułatwia konstrukcję różnych miar. Widzimy, że wystarczy skonstruować miarę zewnętrzną μ_Z (warunki w definicji są słabsze) na przestrzeni X, a następnie zawęzić dziedzinę miary zewnętrznej μ_Z do zbiorów A, które spełniają warunek Caratheodory'ego. Poniżej przedstawione zostaną klasyczne sposoby określania miar.

Dyskretna przestrzeń mierzalna.

Niech zbiór $X = \{x_1, x_2, ...\}$ będzie co najwyżej zbiorem przeliczalnym, natomiast $\mathcal{F}$ $\sigma-$ ciałem zbioru X. Dowolna miara μ na $(X, \mathcal{F})$ jest zadana poprzez ciąg $\{\mu_n\}_{n\geq 1}$ postaci $\mu_n = \mu(\{x_n\}) \geq 0$. Ponadto dla dowolnego $A \in \mathcal{F}$

$$\mu(A) = \sum_{x_i \in A} \mu_i.$$

Jeżeli

$$\mu(X) = \mu\left(\bigcup_{n=1}^{\infty} \{x_n\}\right) = \sum_{n=1}^{\infty} \mu(\{x_n\}) = \sum_{n=1}^{\infty} \mu_n < \infty,$$

to miara μ jest miarą skończoną.

Przestrzeń mierzalna $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Niech $\mathbb{R}$ oznacza zbiór liczb rzeczywistych (prostą), natomiast $\langle a, b\rangle$ dowolny odcinek postaci (a, b), $[a, b]$, $(a, b]$, $[a, b)$, gdzie $-\infty \leq a \leq b \leq \infty$. Niech $\mathcal{A}$ oznacza rodzinę zbiorów postaci

$$A = \bigcup_{i=1}^{n} \langle a_i, b_i\rangle, \tag{1.1}$$

gdzie $n < \infty$ oraz $\langle a_i, b_i \rangle \cap \langle a_j, b_j \rangle = \emptyset$ dla $i \neq j$. Tak zdefiniowana rodzina $\mathcal{A}$ jest ciałem (natomiast nie jest $\sigma-$ ciałem).

Definicja 1.8 *$\sigma-$ ciało generowane poprzez rodzinę zbiorów $\mathcal{A}$ nazywamy $\sigma-$ ciałem zbiorów borelowskich oraz oznaczamy $\mathcal{B}(\mathbb{R})$. Elementy $\mathcal{B}(\mathbb{R})$ nazywamy zbiorami borelowskimi.*

W sposób podobny definiujemy przestrzeń mierzalną $([a,b], \mathcal{B}([a,b]))$. Niech $\sigma-$ ciało $\mathcal{B}([a,b])$ składa się ze zbiorów $A \subset [a,b]$ oraz $A \in \mathcal{B}(\mathbb{R})$. Rodzinę $\mathcal{B}([a,b])$ nazywamy $\sigma-$ ciałem zbiorów borelowskich na $[a,b]$.

Definicja 1.9 *Dla dowolnego $-\infty \leq a \leq b \leq \infty$ miarę*

$$\mu_\lambda(\langle a, b \rangle) = b - a$$

nazywamy miarą Lebesgue'a określoną na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Oczywiście miara Lebesgue'a określona na prostej $\mathbb{R}$ jest miarą zewnętrzną, natomiast z twierdzenia Caratheodory'ego wynika, że jest ona miarą zupełną. Ponieważ $\mu_\lambda(\mathbb{R}) = \infty$, zatem miara Lebesgue'a na $\mathbb{R}$ nie jest skończona .

Definicja 1.10 *Funkcję $F(x) = \mu((-\infty, x]), x \in \mathbb{R}$ spełniającą warunki:*

1. *$F(x)$ jest funkcją niemalejącą;*
2. *$F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$, $F(\infty) = \lim\limits_{x \to \infty} F(x) < \infty$;*
3. *$F(x)$ prawostronnie ciągła (dla $x_0 \in \mathbb{R}$ $\lim\limits_{x \to x_0^+} F(x) = F(x_0)$) oraz ma lewostronne granice ($\lim\limits_{x \to x_0^-} F(x) = F(x_0 - 0)$)*

nazywamy dystrybuantą.

Między dystrybuantą F a miarą μ istnieje wzajemnie jednoznaczna zależność, co oznacza, że mając miarę można określić dystrybuantę i odwrotnie. Z definicji powyżej wynika, że dla każdej miary skończonej μ (miary przypisującej skończoną wartość przestrzeni mierzalnej) określonej na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ istnieje jednoznacznie określona dystrybuanta F oraz $F(\infty) = \lim\limits_{x \to \infty} F(x) = \mu(\mathbb{R})$.

Jeżeli F jest dystrybuantą określoną na $\mathbb{R}$, to istnieje dokładnie jedna miara μ określona na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ taka, że dla dowolnych $a, b \in \mathbb{R}$, $a \leq b$

$$\mu((a,b]) = F(b) - F(a).$$

W sposób identyczny definiujemy $\mu([a,b]) = F(b) - F(a-0)$, $\mu((a,b)) = F(b-0) - F(a)$, $\mu([a,b)) = F(b-0) - F(a-0)$. Jeżeli zbiór A jest określony za pomocą (1.1), to

$$\mu\left(\bigcup_{i=1}^{n} \langle a_i, b_i\rangle\right) = \sum_{i=1}^{n} \mu(\langle a_i, b_i\rangle),$$

gdzie $n < \infty$ oraz $\langle a_i, b_i\rangle \cap \langle a_j, b_j\rangle = \emptyset$ dla $i \neq j$. Zatem dla dowolnego zbioru $A \in \mathcal{A}$ (A jest zbiorem postaci (1.1)) mamy $\mu(A) < \infty$.

Przestrzeń mierzalna $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Niech $\mathbb{R}^n$ oznacza zbiór wektorów $x = (x_1, ..., x_n) \in \mathbb{R}^n$, elementami którego są liczby rzeczywiste. Na tej przestrzeni definiujemy rodzinę $\mathcal{A}^n$, która składa się ze zbiorów postaci

$$A = A_1 \times ... \times A_n = \prod_{i=1}^{n} A_k,$$

gdzie $A_k \in \mathcal{A}$ dla dowolnego $1 \leq k \leq n$, natomiast $\mathcal{A}$ oznacza rodzinę zbiorów postaci (1.1). Tak zdefiniowana rodzina $\mathcal{A}^n$ jest ciałem.

Definicja 1.11 *$\sigma-$ ciało generowane poprzez rodzinę zbiorów $\mathcal{A}^n$ nazywamy $\sigma-$ ciałem zbiorów borelowskich oraz oznaczamy $\mathcal{B}(\mathbb{R}^n)$. Elementy $\mathcal{B}(\mathbb{R}^n)$ nazywamy zbiorami borelowskimi.*

Definicja 1.12 *Miarę*

$$\mu_\lambda\left(\prod_{i=1}^{n} \langle a_i, b_i\rangle\right) = \prod_{i=1}^{n} (b_i - a_i),$$

gdzie $-\infty \leq a_i \leq b_i \leq \infty$ dla każdego $1 \leq i \leq n$, nazywamy miarą Lebesgue'a określoną na $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Oczywiście miara Lebesgue'a na $\mathbb{R}^n$ jest miarą zewnętrzną, natomiast z twierdzenia Caratheodory'ego wynika, że jest ona miarą zupełną. Miara Lebesgue'a na $\mathbb{R}^n$ nie jest skończona, $\mu_\lambda(\mathbb{R}^n) = \infty$. Miarę skończoną na $\sigma-$ ciele zbiorów borelowskich $\mathcal{B}(\mathbb{R}^n)$ określamy za pomocą dystrybuanty $n-$wymiarowej.

Definicja 1.13 *Funkcję $F(x_1, x_2, ..., x_n)$, $x_1, x_2, ..., x_n \in \mathbb{R}$ spełniającą następujące warunki:*

1. *$F(x_1, x_2, ..., x_n)$ jest funkcją niemalejącą względem każdej współrzędnej, tzn.*
$$\triangle_1 ... \triangle_n F(x_1, x_2, ..., x_n) \geq 0,$$
gdzie $\triangle_i$ operator różnicowania względem zmiennej x_i jest dany wzorem
$$\begin{aligned}\triangle_i F(x_1, x_2, ..., x_n) =& F(x_1, ..., x_{i-1}, x_i + h_i, .x_{i+1}, ..., x_n) \\ &- F(x_1, ..., x_{i-1}, x_i, .x_{i+1}, ..., x_n)\end{aligned}$$
oraz $h_1 \geq 0$, ...,$h_n \geq 0$;
2. *jeżeli istnieje co najmniej jedna zmienna $x_i \to -\infty$, to*
$$F(x_1, ..., x_{i-1}, -\infty, .x_{i+1}, ..., x_n) = \lim_{x_i \to -\infty} F(x_1, x_2, ..., x_n) = 0,$$
jeżeli natomiast $x_i \to \infty$ dla każdego $i = 1, 2, ..., n$, to
$$F(\infty, ..., \infty) = \lim_{x_1 \to \infty, ..., x_n \to \infty} F(x_1, ..., x_n) < \infty;$$
3. *$F(x_1, x_2, ..., x_n)$ jest prawostronnie ciągła oraz ma lewostronne granice względem każdej współrzednej x_i, $1 \leq i \leq n$*

nazywamy $n-$wymiarową dystrybuantą.

Z twierdzenia Caratheodory'ego wynika, że jeżeli $F(x_1, x_2, ..., x_n)$ jest $n-$wymiarową dystrybuantą, to istnieje dokładnie jedna miara na $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, taka że
$$F(x_1, x_2, ..., x_n) = \mu\left(\prod_{i=1}^{n} (-\infty, x_i]\right). \tag{1.2}$$

Odwrotnie, jeżeli μ jest miarą skończoną na $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, to funkcja $F(x_1, x_2, ..., x_n)$ dana wzorem (1.2) jest $n-$wymiarową dystrybuantą oraz $\mu(\mathbb{R}^n) = F(\infty, ..., \infty)$. Zatem mając miarę skończoną można określić $n-$wymiarową dystrybuantę i odwrotnie.

Funkcje mierzalne.

Niech $(X, \mathcal{A})$ będzie przestrzenią mierzalną.

Definicja 1.14 *Funkcję rzeczywistą $f(x)$, $x \in X$ nazywamy $\mathcal{A}$ mierzalną, jeżeli dla dowolnego zbioru borelowskiego $B \in \mathcal{B}(\mathbb{R})$*

$$f^{-1}(B) \in \mathcal{A},$$

gdzie $f^{-1}(B) = \{x \in X : f(x) \in B\}$ oznacza przeciwobraz zbioru B.

Definicja funkcji mierzalnej oznacza, że przeciwobraz dowolnego podzbioru borelowskiego $\mathbb{R}$ jest zbiorem mierzalnym w X. Do weryfikacji mierzalności funkcji wykorzystujemy następujące twierdzenie.

Twierdzenie 1.2 *Funkcja $f(x)$ określona na przestrzeni mierzalnej $(X, \mathcal{A})$ jest $\mathcal{A}$ mierzalna, jeżeli dla dowolnego $a \in \mathbb{R}$ zbiór*

$$\{x \in X : f(x) \leq a\}$$

jest zbiorem mierzalnym.

Każdą funkcję $g(y)$, $y \in \mathbb{R}$ nazywamy funkcją borelowską, jeżeli jest ona $\mathcal{B}(\mathbb{R})$ mierzalna.

Twierdzenie 1.3 *Niech $(X, \mathcal{A})$ będzie przestrzenią mierzalną. Funkcja złożona $h(x) = g(f(x))$, $x \in X$ jest $\mathcal{A}$ mierzalna, jeżeli $f(x)$, $x \in X$ jest $\mathcal{A}$ mierzalna, natomiast $g(y)$, $y \in \mathbb{R}$ jest funkcją borelowską.*

Twierdzenie 1.4 *Niech $(X, \mathcal{A})$ będzie przestrzenią mierzalną oraz funkcje $f_1, f_2, ...$ określone na X są $\mathcal{A}$ mierzalne. Funkcje $f_1(x) + f_2(x)$, $f_1(x) f_2(x)$, $\frac{1}{f_1(x)}$ (pod warunkiem $f_1(x) \neq 0$), $|f_1(x)|$, $\max\{f_1(x), f_2(x)\}$, $\min\{f_1(x), f_2(x)\}$, $\sup_n f_n(x)$, $\inf_n f_n(x)$ są również funkcjami mierzalnymi.*

Definicja 1.15 *Funkcję mierzalną $f(x)$ określoną na $(X, \mathcal{A})$ nazywamy funkcją prostą, jeżeli przyjmuje skończenie wiele wartości.*

Funkcję prostą $f(x)$ możemy przedstawić w postaci

$$f(x) = \sum_{i=1}^{m} a_k I_{A_k}(x), \quad x \in X,$$

gdzie $a_k \in \mathbb{R}$, $A_k \in \mathcal{A}$ dla $k = 1, ..., m$, $m < \infty$ oraz $A_i \cap A_j = \emptyset$ dla $i \neq j$.

Definicja 1.16 *Niech $f(x)$ będzie funkcją mierzalną określoną na przestrzeni mierzalnej $(X, \mathcal{A})$. Rodzinę*

$$\mathcal{A}^f = \left\{ f^{-1}(B) : B \in \mathcal{B}(\mathbb{R}) \right\}$$

nazywamy $\sigma-$ciałem generowanym przez funkcję $f(x)$.

Łatwo udowodnić, że $\mathcal{A}^f$ jest $\sigma-$ciałem, dodatkowo $\mathcal{A}^f \subset \mathcal{A}$. Twierdzenie poniżej przedstawia sposób wyznaczenia klasy wszystkich funkcji mieralnych względem $\sigma-$ciała $\mathcal{A}^f$.

Twierdzenie 1.5 *Funkcja $h(x)$, $x \in X$ jest funkcją mierzalną względem $\sigma-$ciała $\mathcal{A}^f$, wtedy i tylko wtedy, jeżeli istnieje funkcja borelowska $g(y)$, $y \in \mathbb{R}$, taka że*

$$h(x) = g(f(x)) \quad \textit{dla dowolnego } x \in X.$$

1.2 Elementy rachunku prawdopodobieństwa

1.2.1 Przestrzeń probabilistyczna.

Rachunek prawdopodobieństwa jest działem matematyki zajmującym się badaniem i wykrywaniem praw w zakresie zjawisk losowych, np. losowy charakter napięcia prądu, czas poprawnej pracy elementów, stopy zwrotu z inwestycji w poszczególne aktywa. W każdej teorii istnieją pojęcia pierwotne: w arytmetyce - liczba, w geometrii - punkt, prosta, płaszczyzna.

W rachunku prawdopodobieństwa pojęciem pierwotnym jest zdarzenie elementarne ω. Przez przestrzeń zdarzeń elementarnych Ω rozumiemy zbiór tych zdarzeń, które są elementarnymi, niepodzielnymi wynikami doświadczenia czy też obserwacji. Zdarzeniem losowym nazywamy dowolny podzbiór przestrzeni zdarzeń elementarnych. Oczywiście zdarzeniem pewnym jest podzbiór zawierający wszystkie elementy tej przestrzeni (zdarzeniem pewnym jest Ω). Zdarzeniem niemożliwym jest zbiór pusty $\emptyset$.

Niech $\mathcal{F}$ oznacza $\sigma-$ ciało przestrzeni zdarzeń elementarnych Ω. Zgodnie z powyższą definicją elementy $\sigma-$ ciała $\mathcal{F}$ są zdarzeniami losowymi. Jeżeli Ω jest zbiorem przeliczalnym lub skończonym, to każde $\sigma-$ ciało $\mathcal{F}$ jest zbiorem wszystkich podzbiorów zbioru Ω.

Definicja 1.17 *Każdą funkcję P określoną na $\mathcal{F}$ o wartościach w $[0,1]$ spełniającą warunki:*

1. $P(\Omega) = 1$;
2. *dla każdego* $A \in \mathcal{F} \quad 0 \leq P(A) \leq 1$;
3. *jeżeli* $A_1, A_2, ... \in \mathcal{F}$ *oraz* $A_i \cap A_j = \emptyset$ *dla dowolnych* $i \neq j$, *to*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

nazywamy miarą prawdopodobieństwa.

Trójkę $(\Omega, \mathcal{F}, P)$ (przestrzeń Ω z miarą P określoną na $\mathcal{F}$) nazywamy przestrzenią probabilistyczną (zupełną przestrzenią probabilistyczną). Wielkość $P(A)$ oznacza prawdopodobieństwo zdarzenia A. Jeżeli $(\Omega, \mathcal{F}, P)$ jest przestrzenią probabilistyczną, to spełnione są następujące własności:

1. $P(\emptyset) = 0$;
2. dla dowolnego $A \in \mathcal{F} \quad P\left(\bar{A}\right) = 1 - P(A)$;
3. jeżeli $A \subset B$, to $P(A) \leq P(B)$;
4. jeżeli $A, B \in \mathcal{F}$, to $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
5. jeżeli $A_1 \subset A_2 \subset ... \in \mathcal{F}$, to

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n);$$

6. jeżeli $... \subset A_2 \subset A_1 \in \mathcal{F}$, to

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n).$$

Definicja 1.18 *Zdarzenia A i B są niezależne (niezależne względem miary P lub niezależnie stochastycznie), jeżeli $P(A \cap B) = P(A) P(B)$.*

Mówimy, że zdarzenia losowe $A_1, A_2, ..., A_m$ są niezależne, jeżeli dla dowolnych wskaźników $1 \leq i_1 < i_2 < ... < i_m \leq m$, $2 \leq m \leq n$ spełniona jest równość

$$P(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}) = P(A_{i_1}) P(A_{i_2}) ... P(A_{i_m}).$$

Definicja 1.19 *Prawdopodobieństwem warunkowym zdarzenia A pod warunkiem, że zaszło zdarzenie B takie, że $P(B) > 0$, nazywamy wielkość*

$$P(A|\, B) = \frac{P(AB)}{P(B)}.$$

Jeżeli zdarzenia A i B są niezależne oraz $P(A) > 0$ i $P(B) > 0$, to $P(A|\, B) = P(A)$ i $P(B|\, A) = P(B)$.

Twierdzenie 1.6 *Załóżmy, że zdarzenia $A_1, A_2, ... \in \mathcal{F}$ spełniają warunki:*

1. *$P(A_i) > 0$ dla $i \geq 1$;*
2. *$P(A_i \cap A_j) = 0$ dla $i \neq j$;*
3. *$\bigcup_{i=1}^{\infty} A_i = \Omega$.*

Dla dowolnego $B \in \mathcal{F}$ zachodzi równość

$$P(B) = \sum_{i=1}^{\infty} P(B|\, A_i) P(A_i). \tag{1.3}$$

Wzór (1.3) nazywamy wzorem na prawdopodobieństwo całkowite.

1.2.2 Zmienne losowe. Wektory losowe

Definicja 1.20 *Zmienną losową określoną na* $(\Omega, \mathcal{F}, P)$ *nazywamy każdą funkcję* $\xi : \Omega \to \mathbb{R}$ *mierzalną względem* $\sigma-$ *ciała* $\mathcal{F}$.

Zatem dla dowolnego zbioru borelowskiego $B \in \mathcal{B}(\mathbb{R})$ przeciwobraz

$$\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\}$$

jest zdarzeniem losowym. Bez zmiejszenia ogólności poniżej przy określeniu zmiennej losowej ξ argument ω będziemy pomijać. Z twierdzenia 1.4 wynika że suma, różnica, iloczyn, iloraz (pod warunkiem, że mianownik jest różny od zera) zmiennych losowych również jest zmienną losową. Z twierdzenia 1.5 wynika, że jeżeli ξ jest zmienną losową oraz $g(x), x \in \mathbb{R}$ jest funkcją borelowską, to $\eta = g(\xi)$ również jest zmienną losową, ponieważ funkcja $\eta(\omega) = g(\xi(\omega)), \omega \in \Omega$ jest $\mathcal{F}$ mierzalna.

Definicja 1.21 *Niech* $\xi : \Omega \to \mathbb{R}$ *będzie zmienną losową.* $\sigma-$ *ciało zawierające zdarzenia*

$$\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\}$$

dla dowolnego $B \in \mathcal{B}(\mathbb{R})$ *nazywamy* $\sigma-$ *ciałem generowanym przez zmienną losową* ξ *oraz oznaczamy* $\mathcal{F}^{\xi}$ *lub* $\sigma(\xi)$.

Z definicji wynika, że $\mathcal{F}^{\xi} \subset \mathcal{F}$, tzn. $\sigma-$ ciało $\mathcal{F}^{\xi}$ zawiera wszystkie zdarzenia losowe $A \in \mathcal{F}$, zajście (wystąpienie) których obserwujemy na podstawie realizacji zmiennej losowej ξ.

Twierdzenie 1.7 *Niech* $\xi : \Omega \to \mathbb{R}$ *będzie zmienną losową natomiast* $g(x)$, $x \in \mathbb{R}$ *dowolną funkcją borelowską. Zmienna losowa* $\eta = g(\xi)$ *jest* $\mathcal{F}^{\xi}$ *mierzalna. Odwrotnie, jeżeli zmienna losowa* η *jest* $\mathcal{F}^{\xi}$ *mierzalna, to istnieje funkcja borelowska* $g(x)$ *taka, że* $\eta = g(\xi)$.

Z każdą zmienną losową ξ jest powiązana miara P_{ξ} określona na $\mathcal{B}(\mathbb{R})$, która jest dana wzorem

$$P_{\xi}(B) = P\left(\xi^{-1}(B)\right) \text{ dla dowolnego } B \in \mathcal{B}(\mathbb{R}).$$

Miarę P_ξ nazywamy rozkładem zmiennej losowej ξ. Dla każdej miary na $\mathcal{B}(\mathbb{R})$ istnieje jednoznacznie określona dystrybuanta.

Definicja 1.22 *Funkcję* $F_\xi(x) = P_\xi((-\infty, x]) = P(\omega \in \Omega : \xi(\omega) \leq x)$, $x \in \mathbb{R}$ *nazywamy dystrybuantą zmiennej losowej* ξ.

Definicja 1.23 *Zmienna losowa ξ jest zmienną losową dyskretną (zmienną losową o rozkładzie dyskretnym), jeżeli przyjmuje ona wartości z co najwyżej przeliczalnego zbioru* $\{a_1, a_2, ...\}$ *z prawdopodobieństwami* $\{p_1, p_2, ...\}$, *gdzie* $p_n \geq 0$ *oraz* $\sum\limits_n p_n = 1$.

Zmienną losową ξ możemy przedstawić w postaci

$$\xi(\omega) = \sum_n a_n I_{A_n}(\omega),$$

gdzie $A_n \in \mathcal{F}$, $\bigcup\limits_n A_n = \Omega$ oraz $A_i \cap A_j = \emptyset$ dla $i \neq j$, $p_n = P(A_n)$, natomiast

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Funkcję $I_A : \Omega \to \{0, 1\}$ nazywamy indykatorem zbioru A. Dystrybuanta zmiennej losowej ξ o rozkładzie dyskretnym jest dana wzorem

$$F_\xi(x) = \sum_{n: a_n \leq n} p_n \text{ dla } x \in \mathbb{R}.$$

Punkty $a_1, a_2, ...$ są punktami nieciągłości pierwszego rodzaju dystrybuanty zmiennej losowej ξ o rozkładzie dyskretnym, $F_\xi(a_n) - F_\xi(a_n - 0) = p_n$. Dla dowolnego $B \in \mathcal{B}(\mathbb{R})$

$$P_\xi(B) = \sum_{n: a_n \in B} p_n.$$

Definicja 1.24 *Zmienna losowa ξ jest zmienną losową ciągłą (zmienną losową o rozkładzie ciągłym), jeżeli istnieje nieujemna funkcja* $p_\xi(x) \geq 0$,

$x \in \mathbb{R}$, że dla każdego zbioru borelowskiego $B \in \mathcal{B}(\mathbb{R})$

$$P_\xi(B) = \int_B p_\xi(s)\,ds \text{ oraz } \int_{-\infty}^{\infty} p_\xi(s)\,ds = 1.$$

Funkcję $p_\xi(x)$ nazywamy funkcją gęstości zmiennej losowej ξ. Dystrybuanta zmiennej losowej ξ o rozkładzie ciągłym

$$F_\xi(x) = \int_{-\infty}^{x} p_\xi(s)\,ds \text{ dla } x \in \mathbb{R}$$

jest funkcją ciągłą i różniczkowalną na $\mathbb{R}$ oraz $\frac{dF_\xi(x)}{dx} = p_\xi(x)$. Wobec powyższego prawdopodobieństwo zdarzenia $\{\xi \in B\}, B \in \mathcal{B}(\mathbb{R})$ wyznaczamy za pomocą całki Lebesgue'a–Stieltjesa

$$P_\xi(B) = P(\{\xi \in B\}) = \int_B dF_\xi(x).$$

Załóżmy, że zmienne losowe $\xi_1, ..., \xi_n$ są określone na $(\Omega, \mathcal{F}, P)$.

Definicja 1.25 *Wektorem losowym $\xi = (\xi_1, ..., \xi_n)$ określonym na $(\Omega, \mathcal{F}, P)$ nazywamy każdą funkcję $\xi : \Omega \to \mathbb{R}^n$ mierzalną względem $\sigma-$ ciała $\mathcal{F}$.*

Definicja 1.26 *$\sigma-$ ciało zawierające zdarzenia*

$$\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\}$$

dla dowolnego $B \in \mathcal{B}(\mathbb{R}^n)$ nazywamy $\sigma-$ ciałem generowanym przez wektor losowy $\xi : \Omega \to \mathbb{R}^n$ oraz oznaczamy $\mathcal{F}^\xi$ lub $\sigma(\xi)$.

Definicja 1.27 *Funkcję*

$$F_\xi(x_1, ..., x_n) = P_\xi(\xi_1 \le x_1, ..., \xi_n \le x_n) = P(\omega \in \Omega : \xi(\omega) \le x),$$

$x = (x_1, ..., x_n) \in \mathbb{R}^n$ nazywamy dystrybuantą wektora losowego ξ.

Dla dowolnego zbioru borelowskiego $B \in \mathcal{B}(\mathbb{R}^n)$ prawdopodobieństwo zdarzenia $\{\xi \in B\}$ wyznaczamy za pomocą całki Lebesgue'a–Stieltjesa

$$P_\xi(B) = P(\{\xi \in B\}) = \int_B dF_\xi(x_1, ..., x_n).$$

Jeżeli $p_\xi(x_1, ..., x_n)$ jest funkcją gęstości wektora losowego ξ, to dla dowolnego $B \in \mathcal{B}(\mathbb{R}^n)$ oraz $(x_1, ..., x_n) \in \mathbb{R}^n$

$$P_\xi(B) = \int_B p_\xi(s_1, ..., s_n)\, ds_1 ... ds_n,$$

$$F_\xi(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} p_\xi(s_1, ..., s_n)\, ds_1 ... ds_n.$$

Definicja 1.28 *Mówimy, że ciąg* $\{\xi_n\}_{1 \leq i \leq n}$ *jest ciągiem niezależnych zmiennych losowych, jeżeli dla dowolnych zbiorów borelowskich* $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$ *zdarzenia losowe* $\{\xi_1 \in B_1\}, ..., \{\xi_n \in B_n\}$ *są niezależne.*

Twierdzenie 1.8 *Zmienne losowe* $\xi_1, ..., \xi_n$ *są niezależne wtedy i tylko wtedy, gdy dla* $x_1, ..., x_n \in \mathbb{R}$

$$F_\xi(x_1, ..., x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i).$$

Niezależność zmiennych losowych jest powiązana z niezależnością $\sigma-$ciał generowanych poprzez te zmienne losowe. Mówimy, że $\sigma-$ciała $\mathcal{F}, \mathcal{G}$ są niezależne, jeżeli niezależne są dowolne zdarzenia losowe $A \in \mathcal{F}, B \in \mathcal{G}$, zatem $P(A \cap B) = P(A)P(B)$.

Zmienne losowe $\xi_1, ..., \xi_n$ są niezależne, wtedy i tylko wtedy gdy niezależne są $\sigma-$ciała $\mathcal{F}^{\xi_1}, ..., \mathcal{F}^{\xi_n}$ generowane poprzez te zmienne losowe. Dla dowolnych zbiorów borelowskich $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$ mamy

$$P(\xi_1 \in B_1, ..., \xi_n \in B_n) = P_{\xi_1, ..., \xi_n}(B_1 \times ... \times B_n) = \prod_{i=1}^{n} P_{\xi_i}(B_i).$$

Pojęcia niezależności zmiennych losowych oraz odpowiadających im $\sigma-$ciał będą wykorzystywane przy wyznaczeniu warunkowych wartości oczekiwanych. O ile nie będzie konieczne poniżej wskaźniki przy funkcjach rozkładu i dystrybuantach zmiennej losowej będą pomijane.

1.2.3 Wartość oczekiwana. Warunkowa wartość oczekiwana

Definicja 1.29 *Wartością oczekiwaną zmiennej losowej ξ określonej na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ nazywamy wielkość*

$$E\xi = m_\xi = \int_\Omega \xi(\omega) P(d\omega).$$

Widzimy, ze wartość oczekiwana zmiennej losowej ξ jest całką Lebesgue'a funkcji $\xi(\omega)$ względem miary P. Jeżeli miara P_ξ i funkcja F_ξ są rozkładem i dystrybuantą zmiennej losowej ξ na $\mathcal{B}(\mathbb{R})$, to

$$E\xi = m_\xi = \int_{\mathbb{R}} sP_\xi(ds) = \int_{-\infty}^{\infty} sdF_\xi(s).$$

Własności wartości oczekiwanej:

1. jeżeli istnieją wartości oczekiwane $E\xi$ i $E|\xi|$ oraz $E\xi < \infty$, $E|\xi| < \infty$, to

$$|E\xi| \leq E|\xi|;$$

2.

$$EI_A = \int_\Omega I_A(\omega) P(d\omega) = P(A),$$

gdzie I_A jest indykatorem zbioru A;

3. jeżeli $E\xi$ istnieje, to dla dowolnych $a, b \in \mathbb{R}$

$$E(a\xi + b) = aE\xi + b\ ;$$

4. jeżeli istnieją wartości oczekiwane $E\xi$ i $E\eta$ oraz $E\xi < \infty$, $E\eta < \infty$, to

$$E(\xi + \eta) = E\xi + E\eta \ ;$$

5. jeżeli $\xi \leq \eta$, to $E\xi \leq E\eta$;
6. jeżeli $f : \mathbb{R} \to \mathbb{R}$ jest dowolną funkcją borelowską, to

$$Ef(\xi) = \int_{-\infty}^{\infty} f(s)\, dF_{\xi}(s) \, ;$$

7. jeżeli $f : \mathbb{R} \to \mathbb{R}$ jest funkcją wypukłą oraz całkowalną, to spełniona jest nierówność

$$f(E\xi) \leq Ef(\xi) \, .$$

Powyższą nierówność nazywamy nierównością Jensena. W przypadku gdy $f : \mathbb{R} \to \mathbb{R}$ jest funkcją wklęsłą, to

$$Ef(\xi) \leq f(E\xi) \, .$$

Dla dowolnego ustalonego $a \in \mathbb{R}$ funkcja $f(x) = \min(x, a)$ jest funkcją wklęsłą, zatem

$$E(\min(\xi, a)) \leq \min(E\xi, a) \, .$$

Definicja 1.30 *Wariancją zmiennej losowej ξ nazywamy wielkość*

$$var\xi = E(\xi - m_{\xi})^2 = \int_{-\infty}^{\infty} (s - m_{\xi})^2 \, dF_{\xi}(s) \, .$$

Własności wariancji:

1. $var\xi = E\xi^2 - m_{\xi}^2 \geq 0$;
2. dla dowolnych $a, b \in \mathbb{R}$

$$var(a\xi + b) = a^2 Var\xi \ ;$$

3. jeżeli $\{\xi_i\}_{1\leq i\leq n}$ jest ciągiem niezależnych zmiennych losowych oraz $var\xi_i < \infty$ dla dowolnego $1 \leq i \leq n$, to

$$var\left(\sum_{i=1}^{n}\xi_i\right) = \sum_{i=1}^{n} Var\xi_i \ .$$

Definicja 1.31 *Kowariancją zmiennych losowych $\xi, \eta : \Omega \to \mathbb{R}$ nazywamy wielkość*

$$cov\,(\xi, \eta) = E\,(\xi - m_\xi)\,(\eta - m_\eta)\,.$$

Własności kowariancji:

1. jeżeli $E\xi^2 < \infty$ oraz $E\eta^2 < \infty$, to

$$cov\,(\xi, \eta) \leq \sqrt{var\xi}\sqrt{var\eta}\ ;$$

2. jeżeli zmienne losowe ξ, η są niezależne, to $cov\,(\xi, \eta) = 0$;
3. $var\xi = cov\,(\xi, \xi)$;
4. $var\,(\xi + \eta) = var\xi + var\eta + 2cov\,(\xi, \eta)$.

Niech $\xi = (\xi_1, ..., \xi_n)$ będzie $n-$wymiarowym wektorem losowym. Macierz $\Sigma_\xi = \{cov\,(\xi_i, \xi_j)\}_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$ nazywamy macierzą kowariancji oraz $\Sigma_\xi = E\xi\xi^T - m_\xi m_\xi^T$.

Jeżeli $P\left(\{\omega : \xi\,(\omega) = \eta\,(\omega)\}\right) = 1$, to mówimy że zmienne losowe ξ i η są równe prawie pewnie względem miary P lub też równe z prawdopodobieństwem 1 oraz oznaczamy $\xi = \eta$ (pp.).

Definicja 1.32 *Warunkową wartością oczekiwaną zmiennej losowej ξ określonej na $(\Omega, \mathcal{F}, P)$ względem $\sigma-$ciała $\mathcal{G} \subset \mathcal{F}$ nazywamy zmienną losową $E\,(\xi\,|\mathcal{G}\,)$, która spełnia następujące warunki:*

1. *$E\,(\xi\,|\mathcal{G}\,)$ jest $\mathcal{G}$ mierzalna;*
2. *dla dowolnego $A \in \mathcal{G}$ spełniona jest nierówność*

$$\int\limits_A \xi\,(\omega)\,P\,(d\omega) = \int\limits_A E\,(\xi\,|\mathcal{G}\,)\,(\omega)\,P\,(d\omega)\,.$$

Poniżej zakładamy, że wartości oczekiwane zmiennych losowych istnieją oraz są skończone. Własności warunkowej wartości oczekiwanej:

1. jeżeli $\xi = C = const$ (pp.), to $E(\xi\,|\mathcal{G}) = C$ (pp.);
2. jeżeli $P(\xi \leq \eta) = 1$, to $E(\xi\,|\mathcal{G}) \leq E(\eta\,|\mathcal{G})$ (pp.);
3. jeżeli $a, b \in \mathbb{R}$ oraz $E|\xi| < \infty$, $E|\eta| < \infty$, to
$$E(a\xi + b\eta\,|\mathcal{G}) = aE(\xi\,|\mathcal{G}) + bE(\eta\,|\mathcal{G}) \ \text{(pp.)};$$
4. jeżeli $E|\xi| < \infty$ oraz $\mathcal{G} = \{\emptyset, \Omega\}$, to $E(\xi\,|\mathcal{G}) = E\xi$ (pp.);
5. jeżeli $E|\xi| < \infty$, to $|E(\xi\,|\mathcal{G})| \leq E(|\xi|\,|\mathcal{G})$ (pp.);
6. jeżeli $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$, to $E(E(\xi\,|\mathcal{G}_2)\,|\mathcal{G}_1) = E(E(\xi\,|\mathcal{G}_1)\,|\mathcal{G}_2) = E(\xi\,|\mathcal{G}_1)$ (pp.);
7. jeżeli zmienna losowa ξ jest $\mathcal{G}$ mierzalna, to $E(\xi\,|\mathcal{G}) = \xi$ (pp.);
8. $E(E(\xi\,|\mathcal{G})) = E\xi$ (pp.);
9. jeżeli zmienna losowa ξ nie zależy od $\mathcal{G}$ (tzn. $\sigma-$ciała $\mathcal{F}^\xi$ i $\mathcal{G}$ są niezależne), to $E(\xi\,|\mathcal{G}) = E\xi$ (pp.);
10. jeżeli $E|\xi\eta| < \infty$ oraz zmienna losowa η jest $\mathcal{G}$ mierzalna, to $E(\xi\eta\,|\mathcal{G}) = \eta E(\xi\,|\mathcal{G})$ (pp.).

Zmienną losową $E(\xi\,|\mathcal{F}^\eta) = E(\xi\,|\eta)$, gdzie $\mathcal{F}^\eta = \sigma\{\eta\}$, nazywamy warunkową wartością oczekiwaną zmiennej losowej ξ względem zmiennej losowej η.

W pracy wykorzystywane będą wektory losowe o rozkładzie normalnym (rozkładzie Gaussa). Niech $\xi : \Omega \to \mathbb{R}^n$ będzie wektorem losowym o rozkładzie normalnym $\boldsymbol{N}(m, \Sigma)$, gdzie $m \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{n\times n}$ oraz

$$m = E\xi, \quad \Sigma = E(\xi - m)(\xi - m)^T.$$

Zakładamy, że macierz Σ jest macierzą nieosobliwą ($det\Sigma \neq 0$), więc także dodatnio określoną. Funkcja gęstości rozkładu normalnego $\boldsymbol{N}(m, \Sigma)$ jest określona wzorem

$$\gamma(x, m, \Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left(-\frac{1}{2}(x - m)^T \Sigma^{-1}(x - m)\right). \tag{1.4}$$

natomiast dla dowolnego $z \in \mathbb{R}^n$ funkcja charakterystyczna wektora losowego ξ jest dana wzorem

$$\varphi_\xi(z) = E\exp\left(iz^T\xi\right) = \exp\left(iz^T m - \frac{1}{2}z^T\Sigma z\right). \tag{1.5}$$

Własności wektorów losowych o rozkładzie normalnym:

1. Jeżeli wektor losowy $\xi : \Omega \to \mathbb{R}^n$ ma rozkład normalny $\boldsymbol{N}(m,\Sigma)$, macierz $A \in \mathbb{R}^{n\times n}$ jest macierzą deterministyczną, $a \in \mathbb{R}^n$ – wektorem deterministycznym, to wektor losowy $\mu = A\xi + a$ ma rozkład normalny $\boldsymbol{N}\left(Am + a, A\Sigma A^T\right)$, gdzie wektor wartości oczekiwanych wynosi

$$E\mu = E(A\xi + a) = Am + a,$$

natomiast macierz kowariancji

$$\begin{aligned} cov(\mu,\mu) &= E(\mu - E\mu)(\mu - E\mu)^T \\ &= E(A\xi + a - Am - a)(A\xi + a - Am - a)^T \\ &= E(A(\xi - m))(A(\xi - m))^T = A\Sigma A^T. \end{aligned}$$

Funkcja charakterystyczna wektora losowego μ jest równa

$$\varphi_\mu(z) = E\exp\left(iz^T\mu\right) = \exp\left(iz^T(Am + a) - \frac{1}{2}z^T A\Sigma A^T z\right). \tag{1.6}$$

2. Niech wektor losowy $\varepsilon : \Omega \to \mathbb{R}^n$ ma rozkład normalny $\boldsymbol{N}(\bar{0}, I)$, gdzie $\bar{0} \in \mathbb{R}^n$ jest wektorem o współrzędnych zerowych oraz macierz $I \in \mathbb{R}^{n\times n}$ jest macierzą jednostkową. Zatem $E\varepsilon_j = 0$ dla dowolnego $1 \le j \le n$ oraz

$$E\varepsilon_i\varepsilon_j = \begin{cases} 1, & \text{dla } i = j, \\ 0, & \text{dla } i \ne j, \end{cases}$$

ponieważ składowe wektora losowego ε są niezależne. Jeżeli wektor losowy ξ ma rozkład normalny $\boldsymbol{N}(m,\Sigma)$, to z własności 1 wynika, że wektor ξ możemy przedstawić w postaci

$$\xi = \Sigma^{\frac{1}{2}}\varepsilon + m.$$

3. Niech $\theta : \Omega \to \mathbb{R}^k$ i $\xi : \Omega \to \mathbb{R}^n$ będą wektorami losowymi o rozkładach normalnych $\boldsymbol{N}(m_\theta, \Sigma_\theta)$ i $\boldsymbol{N}(m_\xi, \Sigma_\xi)$ odpowiednio. Jeżeli wektory losowe θ i ξ są niezależne, to macierz kowariancji wektorów θ i ξ jest macierzą zerową oraz dla dowolnych $x \in \mathbb{R}^k$, $z \in \mathbb{R}^n$ funkcja charakterystyczna wektora $(\theta, \xi) \in \mathbb{R}^{k+n}$ jest dana wzorem

$$\begin{aligned}\varphi_{(\theta,\xi)}(x,z) &= E\exp\left(i\left(x^T\theta + z^T\xi\right)\right) = \varphi_{(\theta)}(x)\,\varphi_{(\xi)}(z)\\ &= \exp\left(ix^T m_\theta - \frac{1}{2}x^T\Sigma_\theta x\right)\exp\left(iz^T m_\xi - \frac{1}{2}z^T\Sigma_\xi z\right).\end{aligned}$$

1.3 Procesy stochastyczne

Zachowanie wiele zjawisk technicznych, fizycznych, ekonomicznych itp. opisujemy za pomocą ciągu zmiennych losowych. W pracy do matematycznego opisu tych zjawisk wykorzystane będą funkcje losowe zależne od czasu, które nazywamy procesami stochastycznymi. Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną.

Definicja 1.33 *Rodzinę $\{\xi_t\}_{t\in T}$ zmiennych losowych $\xi_t = \xi_t(\omega) : \Omega \to \mathbb{R}$ zależnych od parametru $t \in T = [0, \infty)$ nazywamy procesem stochastycznym (losowym) z czasem ciągłym. Jeżeli parametr $t \in T = \mathbb{N}_0$, to rodzinę $\{\xi_t\}_{t\in T}$ nazywamy szeregiem czasowym lub procesem stochastycznym z czasem dyskretnym.*

W zależności od potrzeb będziemy przyjmować $T = [0, \infty)$ lub $T = \mathbb{N}_0$. W przypadku, gdy $\xi_t = \xi_t(\omega) : \Omega \to \mathbb{R}^n$ dla $t \in T$, to rodzinę $\{\xi_t\}_{t\in T}$ nazywamy $n-$wymiarowym procesem stochastycznym. Zatem proces stochastyczny $\{\xi_t\}_{t\in T}$ jest funkcją, która odwzorowuje $T\times\Omega$ w przestrzeń mierzalną $\mathbb{R}$ (lub $\mathbb{R}^n$ w przypadku procesu $n-$wymiarowego). Dla ustalonego $\omega \in \Omega$ funkcję $\xi_t : T \to \mathbb{R}$ (lub $\xi_t : T \to \mathbb{R}^n$) nazywamy trajektorią lub realizacją procesu stochastycznego.

Niech $\mathcal{B}$ oznacza rodzinę zbiorów borelowskich w $\mathbb{R}$ (lub $\mathbb{R}^n$ w przypadku $n-$wymiarowym). Z każdym procesem stochastycznym $\{\xi_t\}_{t\in T}$ powiązane są $\sigma-$ciała $\mathcal{F}_t^\xi = \sigma\{\omega : \xi_s, s \le t\}$. Przez $\mathcal{F}_t^\xi$ rozumiemy zdarzenia, które zaszły do momentu t. Dla każdego $t \in T$ $\mathcal{F}_t^\xi$ oznacza $\sigma-$ciało generowane poprzez rodzinę zbiorów $\{\omega : \xi_s \in B, s \le t\}$, gdzie $B \in \mathcal{B}$, tzn. $\sigma-$ciało względem którego mierzalne są zmienne losowe $\xi_s, s \le t$.

Definicja 1.34 *Proces stochastyczny $\{\xi_t\}_{t\in T}$ nazywamy procesem mierzalnym, jeżeli dla dowolnego $B \in \mathcal{B}$*

$$\{(\omega, t) : \xi_s \in B\} \in \mathcal{F} \times \mathcal{B}(T),$$

gdzie $\mathcal{B}(T)$ jest $\sigma-$ciałem zbiorów borelowskich na T.

Niech $\{\mathcal{F}_t\}_{t\in T}$ będzie niemalejącą rodziną $\sigma-$ciał, tzn. $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ dla $s \le t$. Mówimy, że mierzalny proces stochastyczny $\{\xi_t\}_{t\in T}$ jest zgodny z rodziną $\sigma-$ciał $\{\mathcal{F}_t\}_{t\in T}$, jeżeli dla dowolnego $t \in T$ zmienne losowe ξ_t są $\mathcal{F}_t$ mierzalne. Oznacza to, że dla $t \in T$ wielkość ξ_t jest całkowicie wyznaczona poprzez $\mathcal{F}_t$. Ogólnie przyjmujemy $\mathcal{F}_t = \mathcal{F}_t^\xi$.

Niektóre procesy posiadają pewną własność, którą obrazowo można scharakteryzować jako: przyszłość nie zależy od przyszłości, gdy znana jest teraźniejszość. Takie procesy stanowią klasę procesów Markowa.

Definicja 1.35 *Proces stochastyczny $\{\xi_t\}_{t\in T}$ nazywamy procesem Markowa, jeżeli dla dowolnego $k \in \mathbb{N}$ oraz $0 \le t_1 < t_2 < ... < t_k$ oraz dowolnego $B \in \mathcal{B}$ spełniona jest równość*

$$P\left(\xi_{t_k} \in B \,\middle|\, \xi_{t_1}, ..., \xi_{t_{k-1}}\right) = P\left(\xi_{t_k} \in B \,\middle|\, \xi_{t_{k-1}}\right). \tag{1.7}$$

W lewej części równania (1.7) prawdopodobieństwo warunkowe wyznaczamy względem $\sigma-$ciała generowanego poprzez ciąg $\xi_{t_1}, ..., \xi_{t_{k-1}}$, natomiast w prawej części tego równania prawdopodobieństwo warunkowe wyznaczamy względem $\sigma-$ciała generowanego poprzez $\xi_{t_{k-1}}$. Widzimy, że dla procesów Markowa warunkowy rozkład ξ_{t_k}, gdy znane są wartości $\xi_{t_1}, ..., \xi_{t_{k-1}}$, zależy tylko od $\xi_{t_{k-1}}$. Z definicji wynika, że jeżeli proces stochastyczny $\{\xi_t\}_{t\in T}$ jest procesem Markowa, to dla dowolnych $0 \leq s \leq t$ oraz $B \in \mathcal{B}$

$$P\left(\xi_t \in B \left| \mathcal{F}_s \right.\right) = P\left(\xi_t \in B \left| \xi_s \right.\right).$$

Twierdzenie 1.9 *Proces stochastyczny $\{\xi_t\}_{t\in T}$ jest procesem Markowa wtedy i tylko wtedy, gdy dla dowolnej ograniczonej funkcji $f()$ oraz dowolnych $0 \leq t_1 < t_2 < ... < t_k$*

$$E\left(f\left(\xi_{t_k}\right) \left| \xi_{t_1}, ..., \xi_{t_{k-1}} \right.\right) = E\left(f\left(\xi_{t_k}\right) \left| \xi_{t_{k-1}} \right.\right).$$

Szczególnym przypadkiem procesów Markowa są procesy o przyrostach niezależnych. Mówimy, że proces stochastyczny $\{\xi_t\}_{t\in T}$ jest procesem o przyrostach niezależnych, jeżeli dla dowolnego $k \in \mathbb{N}$ oraz $0 \leq t_1 < t_2 < ... < t_k$ różnice $\xi_{t_2} - \xi_{t_1}, ..., \xi_{t_k} - \xi_{t_{k-1}}$ są niezależne.

Definicja 1.36 *Proces stochastyczny $\{\xi_t\}_{t\in T}$ względem rodziny $\mathcal{F} = \{\mathcal{F}_t\}_{t\in T}$ spełniający warunek $E(|\xi_t|) < \infty$ dla $t \in T$ nazywamy:*

1. *martyngałem, jeżeli $E\left(\xi_t \left| \mathcal{F}_s \right.\right) = \xi_s$ (pp.) dla $t \geq s$;*
2. *nadmartyngałem, jeżeli $E\left(\xi_t \left| \mathcal{F}_s \right.\right) \leq \xi_s$ (pp.) dla $t \geq s$;*
3. *podmartyngałem, jeżeli $E\left(\xi_t \left| \mathcal{F}_s \right.\right) \geq \xi_s$ (pp.) dla $t \geq s$.*

Bardzo ważną klasę procesów stochastycznych stanowią procesy gaussowskie.

Definicja 1.37 *Proces stochastyczny $\{\xi_t\}_{t\in T}$ nazywamy procesem Gaussa (gaussowskim, normalnym) jeżeli dla dowolnego $k \in \mathbb{N}$ oraz $t_1, t_2, ..., t_k \in T$ rozkład łączny $\xi_{t_1}, ..., \xi_{t_k}$ jest rozkładem normalnym.*

Proces Gaussa jest całkowicie określony poprzez wartość średnią $m_\xi(t) = E\xi_t$ oraz macierz kowariancji $\Sigma_\xi(t,s) = cov(\xi_t, \xi_s)$ dla $t, s \in T$. Jeżeli $\{\xi_t\}_{t\in T}$ jest procesem Gaussa, to liniowe przekształcenie tego procesu jest również procesem Gaussa.

1.4 Optymalne zatrzymanie procesów losowych

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną, natomiast $\{\mathcal{F}_t\}_{t\in T}$, $T = [0, \infty)$ oznacza niemalejącą rodzinę $\sigma-$ciał, tzn. $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ dla $s \leq t$. W teorii optymalnego zatrzymania procesów losowych decydujące znaczenie mają momenty Markowa.

Zmienną losową $\tau : \Omega \to T \cup \{\infty\}$ nazywamy momentem Markowa względem rodziny $\{\mathcal{F}_t\}_{t\in T}$, jeżeli dla każdego $t \in T$

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

Momenty Markowa są to zmienne losowe, które nie zależą od przyszłości.

Jeżeli $P(\tau < \infty) = 1$, to moment Markowa τ nazywamy momentem zatrzymania. W definicji powyżej wyrażenie $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ oznacza, że odpowiedź na pytanie, czy moment zatrzymania τ nie przekracza t, zależy od obserwacji zdarzeń do momentu t (zależy od $\mathcal{F}_t$).

Poniżej podane zostaną podstawowe własności momentów Markowa, które bezpośrednio są wykorzystywane podczas konstrukcji zasad zatrzymania procesów stochastycznych.

1. Jeżeli τ jest momentem Markowa względem rodziny $\{\mathcal{F}_t\}_{t\in T}$, to $\{\omega : \tau(\omega) < t\} \in \mathcal{F}_t$ oraz $\{\omega : \tau(\omega) = t\} \in \mathcal{F}_t$ dla $t \in T$.
2. Jeżeli τ_1 i τ_2 są momentami Markowa, to $\min(\tau_1, \tau_2)$, $\max(\tau_1, \tau_2)$ oraz $\tau_1 + \tau_2$ są momentami Markowa.
3. Niech $\tau_1, \tau_2, ...$ będzie ciągiem momentów Markowa, wtedy moment $\sup \tau_n$ również jest momentem Markowa.
4. Moment Markowa τ względem rodziny $\{\mathcal{F}_t\}_{t\in T}$ jest zmienną losową $\mathcal{F}_\tau$ mierzalną.
5. Jeżeli τ i ς są momentami Markowa oraz $\tau \leq \varsigma$ (pp.), to $\mathcal{F}_\tau \subset \mathcal{F}_\varsigma$.
6. Niech $\{\xi_t\}_{t\in T}$ będzie procesem stochastycznym względem rodziny $\{\mathcal{F}_t\}_{t\in T}$ oraz $\tau = \tau(\omega)$ będzie momentem Markowa takim, że $P(\tau < \infty) = 1$. Funkcja $\xi_\tau = \xi_{\tau(\omega)}(\omega)$ jest $\mathcal{F}_\tau$ mierzalna.
7. Niech $\{\xi_t\}_{t\in T}$ będzie procesem stochastycznym względem rodziny

$\{\mathcal{F}_t\}_{t\in T}$, natomiast D – zbiorem domkniętym. Moment

$$\tau = \inf\{t \geq 0 : \xi_t \in D\}$$

jest momentem Markowa względem rodziny $\{\mathcal{F}_t\}_{t\in T}$.

Dowody tych własności można znaleźć np. w [79], [122].

Poniżej przedstawione zostaną optymalne zasady zatrzymania procesów stochastycznych oraz sposób wyznaczenia momentu zatrzymania. Niech $\{\xi_t\}_{0\leq t\leq N}$, $\xi_t = \xi_t(\omega) : \Omega \to \mathbb{R}^n$ dla $0 \leq t \leq N$ będzie procesem stochastycznym z czasem dyskretnym względem rodziny $\mathcal{F} = \{\mathcal{F}_t\}_{0\leq t\leq N}$ spełniający własność $E(|\xi_t|) < \infty$ dla $0 \leq t \leq N$ oraz $g : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją mierzalną, taką że $Eg^-(\xi_t) < \infty$ dla każdego $0 \leq t \leq N$, gdzie $g^-(x) = -\min\{g(x), 0\}$. Wartość funkcji $g(x)$ nazywamy wygraną (zyskiem) dla realizacji x.

Niech τ będzie momentem Markowa τ o realizacjach w zbiorze $\{0, 1, ..., N\}$, gdzie horyzont $N < \infty$. Dla dowolnego $0 \leq j \leq N$ klasę momentów Markowa o relizacjach w zbiorze $\{j, j+1, ..., N\}$ oznaczamy przez $\mathcal{T}(j, N) = \{\tau : j \leq \tau(\omega) \leq N\}$ oraz przyjmujemy $\mathcal{T} = \mathcal{T}(0, N)$. Wyznaczenie optymalnego momentu zatrzymania τ^* dla procesu stochastycznego $\{\xi_t\}_{0\leq t\leq N}$ sprowadza się do rozwiązaniu zadania

$$V_0^N = \sup_{\tau\in\mathcal{T}} Eg(\xi_\tau). \tag{1.8}$$

Wielkość $V_0^N = Eg(\xi_{\tau^*})$ oznacza oczekiwaną wygraną (oczekiwany zysk), jeżeli proces $\{\xi_t\}_{0\leq t\leq N}$ zostanie zatrzymany w momencie τ^*.

Aby wyznaczyć optymalny moment zatrzymania skonstruujemy ciąg $\left\{\upsilon_t^N\right\}_{0\leq t\leq N}$ postaci

$$\upsilon_t^N = \max\left\{g(\xi_t), E\left(\upsilon_{t+1}^N |\mathcal{F}_t\right)\right\} \tag{1.9}$$

z warunkiem początkowym $\upsilon_N^N = \xi_N$ oraz definiujemy moment

$$\tau_0^N = \min\left\{0 \leq t \leq N : g(\xi_t) = \upsilon_t^N\right\}. \tag{1.10}$$

Ciąg $\left\{ v_t^N \right\}_{0 \le t \le N}$ nazywamy kopertą (otoczką) Snell'a. Twierdzenie poniżej przedstawia podstawową zasadę dotyczącą optymalnego zatrzymania procesów losowych na przedziale czasowym $0 \le t \le N$. Więcej na temat teorii optymalnych zasad zatrzymania można znaleźć np. w [33], [122].

Twierdzenie 1.10 *Niech ciąg $\left\{ v_t^N \right\}_{0 \le t \le N}$ spełnia równanie (1.9) oraz moment τ_0^N będzie określony za pomocą (1.10). Wówczas zachodzą następujące własności:*

1. $\tau_0^N \in \mathcal{T}$;
2. $E\left(g\left(\xi_{\tau_0^N}\right)|\mathcal{F}_0\right) = v_0^N$;
3. $E\left(g\left(\xi_\tau\right)|\mathcal{F}_0\right) \le E\left(g\left(\xi_{\tau_0^N}\right)|\mathcal{F}_0\right)$ *dla dowolnego* $\tau \in \mathcal{T}$;
4. $V_0^N = Ev_0^N$.

Dowód powyższego twierdzenia można znaleźć np. w [122], [124].

Wniosek 1.1 *Ciąg $\left\{ v_t^N \right\}_{0 \le t \le N}$ jest najmniejszym nadmartyngałem dominującym ciąg $\left\{ g\left(\xi_t\right) \right\}_{0 \le t \le N}$.*

Wniosek 1.2 *Moment $\tau^* = \tau_0^N$ jest optymalnym momentem zatrzymania w klasie momentów Markowa $\mathcal{T}$ dla procesu stochastycznego $\left\{ g\left(\xi_t\right) \right\}_{0 \le t \le N}$ oraz największa oczekiwana wygrana wynosi $V_0^N = Eg\left(\xi_{\tau_0^N}\right)$.*

Z momentami decyzyjnymi $t = 0, 1, ..., N$ są powiązane pewne podzbiory w $\mathbb{R}^n$. Niech dla dowolnego $0 \le t \le N$

$$G_t^N = \left\{ x \in \mathbb{R}^n : g\left(x\right) = v_t^N \right\}$$

oraz $S_t^N = \mathbb{R}^n \backslash G_t^N$. Ciągi zbiorów $\left\{ G_t^N \right\}_{0 \le t \le N}$ i $\left\{ S_t^N \right\}_{0 \le t \le N}$ określają odpowiednio obszary zatrzymania i kontynuacji obserwacji procesu $\left\{ \xi_t \right\}_{0 \le t \le N}$. W związku z powyższym, żeby określić optymalną decyzję w momencie t, należy sprawdzić do którego ze zbiorów należy ξ_t, tzn. jeżeli $\xi_t \in G_t^N$, to w tym momencie zatrzymujemy proces $\left\{ \xi_t \right\}_{0 \le t \le N}$, w przeciwnym razie kontynuujemy obserwację tego procesu.

Wniosek 1.3 *Optymalny moment zatrzymania procesu* $\{\xi_t\}_{0\leq t\leq N}$ *wynosi*

$$\tau_0^N = \min\left\{0 \leq t \leq N : \xi_t \in G_t^N\right\}.$$

Uwaga 1.4 *Ciągi zbiorów zatrzymania* $\left\{G_t^N\right\}_{0\leq t\leq N}$ *i kontynuacji* $\left\{S_t^N\right\}_{0\leq t\leq N}$ *spełniają własności*

$$G_0^N \subset G_1^N \subset ... \subset G_N^N = \mathbb{R}^n$$

oraz

$$\emptyset = S_N^N \subset S_{N-1}^N \subset ... \subset S_0^N.$$

Rozwiązanie zadania (1.8) dla przypadku gdy horyzont $N = \infty$ można znaleźć w [33], [122], [124].

1.5 Teoria filtracji

1.5.1 Wyznaczenie warunkowego rozkładu

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną. Poniżej rozważamy proces stochastyczny z czasem dyskretnym $\xi = \{\xi_t\}_{t\in\mathbb{N}_0}$, elementy którego są określone za pomocą równania

$$\xi_{t+1} = f(t, \xi, \theta) + \sigma(t, \xi)\,\varepsilon_{t+1}, \tag{1.11}$$

gdzie $f : \mathbb{N}\times\mathbb{R}^n\times\mathbb{R}^k \to \mathbb{R}^n$, $\sigma : \mathbb{N}\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$ oraz $f(t, \xi, \theta) = f(\xi_0, ..., \xi_t, \theta)$ i $\sigma(t, \xi) = \sigma(\xi_0, ..., \xi_t)$, natomiast $\{\varepsilon_t\}_{t\in\mathbb{N}}$ jest ciągiem niezależnych $m-$wymiarowych wektorów losowych o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I)$, $\bar{0} \in \mathbb{R}^m$ jest wektorem o współrzędnych zerowych, $I \in \mathbb{R}^{m\times m}$ jest macierzą jednostkową. Zakładamy, że wektor losowy $\theta : \Omega \to \mathbb{R}^k$ ma rozkład $p(\theta)$ oraz jest stochastycznie niezależny od elementów ciągu $\{\varepsilon_t\}_{t\in\mathbb{N}}$.

Zagadnienie filtracji polega na wyznaczeniu warunkowego rozkładu wektora θ na podstawie realizacji $\{\xi_s\}_{0\leq s\leq t}$ (patrz np. [19], , [21], [48], [79], [86], [89], [98], [99], [101]). Niech $\mathcal{F}_t^\xi = \sigma\{\xi_0, \xi_1, ..., \xi_t\}$ i $\mathcal{F}_t^{\xi\theta} = \mathcal{F}_t^\xi \vee \sigma\{\theta\}$, $t \in \mathbb{N}$ oznaczają niemalejące rodziny $\sigma-$ciał na przestrzeni $(\Omega, \mathcal{F}, P)$. Jeżeli ξ_0 jest wektorem deterministycznym, to wystarczy przyjąć $\mathcal{F}_0^\xi = \{\emptyset, \Omega\}$.

Twierdzenie 1.11 *Niech dla dowolnego $t \in \mathbb{N}$ funkcje $f(t,\xi,\theta) \in \mathbb{R}^n$, $\sigma(t,\xi) \in \mathbb{R}^{n\times m}$ są mierzalne ze względu na $\sigma-$ciało $\mathcal{F}_t^{\xi}$. Wówczas warunkowy rozkład zmiennej losowej ξ ze względu na $\sigma-$ciało $\mathcal{F}_t^{\xi}$, $t \in \mathbb{N}$ jest równy*

$$P\left(d\theta\middle|\mathcal{F}_t^{\xi}\right) = \frac{p(\theta)\prod\limits_{j=1}^{t}\gamma\left(\xi_j, f(j-1,\xi,\theta), \Sigma(j-1,\xi)\right)}{\int p(x)\prod\limits_{j=1}^{t}\gamma\left(\xi_j, f(j-1,\xi,x), \Sigma(j-1,\xi)\right)dx}d\theta. \tag{1.12}$$

gdzie $\Sigma(t,\xi) = \sigma(t,\xi)\sigma^T(t,\xi)$, natomiast $\gamma(x,m,\Sigma)$ jest funkcją gęstości rozkładu $\boldsymbol{N}(m,\Sigma)$ daną wzorem (1.4).

Dowód. Niech $p(\cdot)$ i $p_0(\cdot)$ – oznaczają gęstość apioryczną wektora losowego θ i wektora stanu ξ_0 odpowiednio. Z niezależności θ i ξ_0 wynika, że rozkład łączny jest równy

$$\mu_0(\theta,\xi_0) = p(\theta)p_0(\xi_0).$$

Zauważmy, że dla dowolnego $t \in \mathbb{N}$ i dowolnej ograniczonej borelowskiej funkcji h mamy

$$E\left(h\left(\theta,\xi_0,...,\xi_t,\xi_{t+1}\right)\right) = E\left(E\left(h\left(\theta,\xi_0,...,\xi_t,\xi_{t+1}\right)\middle|\mathcal{F}_t^{\xi\theta}\right)\right).$$

Wobec powyższego

$$Eh\left(\theta,\xi_0,...,\xi_{t+1}\right) = \int h\left(x,\xi_0,...,\xi_{t+1}\right)\mu_{t+1}(x,\xi_0,...,\xi_{t+1})dxd\xi_0...d\xi_{t+1}. \tag{1.13}$$

Z drugiej strony

$$\begin{aligned} &Eh\left(\theta,\xi_0,...,\xi_{t+1}\right) = Eh\left(\theta,\xi_0,...,\xi_t, f\left(t,\xi,\theta\right)+\sigma\left(t,\xi\right)\varepsilon_{t+1}\right) \\ &= \int h\left(x,\xi_0,...,\xi_t, f\left(t,\xi,x\right)+\sigma\left(t,\xi\right)w\right)\mu_t\left(x,\xi_0,...,\xi_t\right)\gamma\left(w,\tilde{0},I_m\right)dxd\xi_0...d\xi_t dw \\ &= \int h\left(x,\xi_0,...,\xi_t,\xi_{t+1}\right)\mu_t\left(x,\xi_0,...,\xi_t\right)\gamma\left(\xi_{t+1}, f\left(t,\xi,x\right),\Sigma\left(t,\xi\right)\right)dxd\xi_0...d\xi_{t+1}, \end{aligned} \tag{1.14}$$

gdzie $\gamma\left(w, \tilde{0}, I_m\right)$ jest funkcją gęstości rozkładu $\boldsymbol{N}\left(\tilde{0}, I_m\right)$ daną wzorem (1.4), $\tilde{0} \in \mathbb{R}^m$ jest wektorem zerowym, $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Porównując (1.13) i (1.14) otrzymujemy

$$\mu_{t+1}(\theta, \xi_0, ..., \xi_{t+1}) = \mu_t\left(\theta, \xi_0, ..., \xi_t\right) \gamma\left(\xi_{t+1}, f\left(t, \xi, \theta\right), \Sigma\left(t, \xi\right)\right)). \qquad (1.15)$$

Zatem dla systemu (1.11) rozkład łączny wektorów $\theta, \xi_0, ..., \xi_t$ jest równy

$$\mu_t(\theta, \xi_0, ..., \xi_t) = p(\theta) p_0(y_0) \prod_{j=1}^{t} \gamma\left(\xi_j, f\left(j-1, \xi, \theta\right), \Sigma\left(j-1, \xi\right)\right)).$$

Dla dowolnego $t \in \mathbb{N}$ warunkowy rozkład wektora losowego θ ze względu na $\sigma-$ciało $\mathcal{F}_t^\xi$ jest dany wzorem

$$P\left(d\theta \middle| \mathcal{F}_t^\xi\right) = \mu_t(\theta\,|\xi_0, ..., \xi_t) d\theta = \frac{\mu_t(\theta, \xi_0, ..., \xi_t)}{\int \mu_t(x, \xi_0, ..., \xi_t) dx} d\theta,$$

który sprowadzamy do postaci (1.12). ∎

1.5.2 Filtracja ciągów warunkowo normalnych

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną, natomiast elementy procesu stochastycznego z czasem dyskretnym $\xi = \{\xi_t\}_{t\in\mathbb{N}_0}$ będą określone za pomocą równania

$$\xi_{t+1} = f_0\left(t, \xi\right) + f_1\left(t, \xi\right) \theta + \sigma\left(t, \xi\right) \varepsilon_{t+1}, \qquad (1.16)$$

gdzie $f_0 : \mathbb{N}\times\mathbb{R}^n \to \mathbb{R}^n$, $f_1 : \mathbb{N}\times\mathbb{R}^n \to \mathbb{R}^{n\times k}$, $\sigma : \mathbb{N}\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$ oraz $f_0\left(t, \xi\right) = f_0\left(\xi_0, ..., \xi_t\right)$, $f_1\left(t, \xi\right) = f_1\left(\xi_0, ..., \xi_t\right)$ i $\sigma\left(t, \xi\right) = \sigma\left(\xi_0, ..., \xi_t\right)$, natomiast $\{\varepsilon_t\}_{t\in\mathbb{N}}$ jest ciągiem niezależnych wektorów losowych o rozkładzie normalnym $\boldsymbol{N}\left(\bar{0}, I\right)$, natomiast $\bar{0} \in \mathbb{R}^m$ jest wektorem o współrzędnych zerowych, macierz $I \in \mathbb{R}^{m\times m}$ jest macierzą jednostkową. Zakładamy, że wektor losowy $\theta : \Omega \to \mathbb{R}^k$ ma rozkład normalny $\boldsymbol{N}\left(m_0, \gamma_0\right)$ oraz jest stochastycznie niezależny od elementów szeregu $\{\varepsilon_t\}_{t\in\mathbb{N}}$. Niech $\mathcal{F}_t^\xi = \mathcal{F}_0^\xi \vee \sigma\left\{\xi_1, \xi_2, ..., \xi_t\right\}$, $\mathcal{F}_t^{\xi\theta} = \mathcal{F}_t^\xi \vee \sigma\left\{\theta\right\}$, $t \in \mathbb{N}$ oznacza niemalejące rodziny $\sigma-$ciał na $(\Omega, \mathcal{F}, P)$.

Jak zobaczymy poniżej warunkowy rozkład wektora θ na podstawie realizacji $\{\xi_s\}_{0\leq s\leq t}$ jest rozkładem normalnym, dodatkowo optymalny estymator w sensie średniokwadratowym wektora nieznanych parametrów θ ("filtr Kalmana - Bucy") pokrywa się z warunkową wartością oczekiwaną

$$m_t = E\left(\theta \middle| \mathcal{F}_t^{\xi}\right),$$

natomiast błąd estymacji (filtracji) jest równy warunkowej macierzy kowariancji

$$Q_t = E\left((\theta - m_t)(\theta - m_t)^T \middle| \mathcal{F}_t^{\xi}\right).$$

Wielkość błędu estymacji określamy jako ślad z macierzy warunkowej kowariancji. W rozważanym przypadku wektor m_t i macierz γ_t, $t \in \mathbb{N}$ zostaną wyprowadzone z twierdzenia o korelacji normalnej oraz zostaną przedstawione w postaci równań rekurencyjnych, co jest bardzo użyteczne od strony praktycznej. Więcej na temat filtracji można znaleć w [48], [79], [86], [98], [99], [101]. Dowody twierdzeń przedstawionych poniżej dla rozszerzonego modelu obserwacyjnego znajdują się w [79].

Twierdzenie 1.12 (Twierdzenie o korelacji normalnej) *Niech $\theta : \Omega \longrightarrow \mathbb{R}^k$ i $\xi : \Omega \longrightarrow \mathbb{R}^n$ będą wektorami o rozkładach normalnych odpowiednio $\boldsymbol{N}(m_\theta, Q_\theta)$ i $\boldsymbol{N}(m_\xi, Q_\xi)$, gdzie*

$$m_\theta \in \mathbb{R}^k, m_\xi \in \mathbb{R}^n, Q_\theta \in \mathbb{R}^{k\times k}, Q_\xi \in \mathbb{R}^{n\times n}.$$

Przyjmijmy $Q_{\theta\xi} = E(\theta - m_\theta)(\xi - m_\xi)^T$. Wtedy $Q_{\theta\xi} \in \mathbb{R}^{k\times n}$, a wektor warunkowej wartości oczekiwanej $E(\theta\,|\xi)$ oraz macierz warunkowej kowariancji

$$cov(\theta, \theta\,|\xi) = E\left((\theta - E(\theta\,|\xi))(\theta - E(\theta\,|\xi))^T\,|\xi\right)$$

dane są wzorami

$$E(\theta\,|\xi) = m_\theta + Q_{\theta\xi}Q_\xi^{-1}(\xi - m_\xi), \tag{1.17}$$

$$cov(\theta, \theta\,|\xi) = Q_\theta - Q_{\theta\xi}Q_\xi^{-1}Q_{\theta\xi}^T. \tag{1.18}$$

Dowód. Rozważmy wektor losowy μ dany wzorem

$$\mu = (\theta - m_\theta) - Q_{\theta\xi} Q_\xi^{-1} (\xi - m_\xi) . \tag{1.19}$$

Niech wektor $\bar{0} \in \mathbb{R}^k$ będzie wektorem o współrzędnych zerowych (wektor zerowy) oraz macierz $\emptyset \in \mathbb{R}^{k \times n}$ będzie macierzą o elementach zerowych (macierz zerowa). Wartość oczekiwana wektora μ jest równa

$$E\mu = E\theta - m_\theta - Q_{\theta\xi} Q_\xi^{-1} (E\xi - m_\xi) = \bar{0},$$

natomiast macierz kowariancji wektorów μ i $(\xi - m_\xi)$ jest równa

$$\begin{aligned} E\mu (\xi - m_\xi)^T &= E (\theta - m_\theta) (\xi - m_\xi)^T - Q_{\theta\xi} Q_\xi^{-1} E (\xi - m_\xi) (\xi - m_\xi)^T \\ &= Q_{\theta\xi} - Q_{\theta\xi} Q_\xi^{-1} Q_\xi = \emptyset. \end{aligned} \tag{1.20}$$

Ze wzoru (1.20) wynika, że wektory μ i ξ są niezależne. Zatem

$$E (\mu \,|\xi\,) = E\mu = \bar{0}.$$

Wobec powyższego z równania (1.19), biorąc obustronnie operator warunkowej wartości oczekiwanej ze względu na ξ, otrzymujemy

$$\bar{0} = E (\theta \,|\xi\,) - m_\theta - Q_{\theta\xi} Q_\xi^{-1} (\xi - m_\xi) .$$

Tym samym stwierdzamy prawdziwość wzoru (1.17). Ze wzoru (1.17) wynika

$$\theta - E (\theta \,|\xi\,) = \theta - m_\theta - Q_{\theta\xi} Q_\xi^{-1} (\xi - m_\xi) = \mu. \tag{1.21}$$

Z równania (1.21) oraz niezależności μ i ξ otrzymujemy

$$\begin{aligned} cov (\theta, \theta \,|\xi\,) &= E \Big((\theta - E (\theta \,|\xi\,)) (\theta - E (\theta \,|\xi\,))^T \,|\xi \Big) \\ &= E \left(\mu\mu^T \,|\xi\, \right) = E\mu\mu^T, \end{aligned} \tag{1.22}$$

natomiast z równania (1.19) mamy

$$\begin{aligned} E\mu\mu^T &= Q_\theta + Q_{\theta\xi}Q_\xi^{-1}Q_\xi Q_\xi^{-1}Q_{\theta\xi}^T - 2Q_{\theta\xi}Q_\xi^{-1}Q_{\theta\xi}^T \\ &= Q_\theta - Q_{\theta\xi}Q_\xi^{-1}Q_{\theta\xi}^T. \end{aligned} \tag{1.23}$$

Podstawiając (1.23) do (1.22) otrzymujemy wzór (1.18). ■

Wniosek 1.4 *Jeżeli θ i ξ są zmiennymi losowymi o rozkładach normalnych $\boldsymbol{N}\left(m_\theta,\sigma_\theta^2\right)$ i $\boldsymbol{N}\left(m_\xi,\sigma_\xi^2\right)$ odpowiednio oraz kowariancja zmiennych losowych θ i ξ wynosi $cov\left(\theta,\xi\right) = r\sigma_\theta\sigma_\xi$, gdzie r jest współczynnikiem korelacji zmiennych θ i ξ oraz $\sigma_\xi^2 > 0$, to z twierdzenia o korelacji normalnej otrzymujemy*

$$\begin{aligned} E\left(\theta\,|\xi\,\right) &= m_\theta + \frac{cov\left(\theta,\xi\right)}{\sigma_\xi^2}\left(\xi - m_\xi\right) = m_\theta + r\frac{\sigma_\theta}{\sigma_\xi}\left(\xi - m_\xi\right), \\ cov\left(\theta,\theta\,|\xi\,\right) &= \sigma_\theta^2 - \frac{cov^2\left(\theta,\xi\right)}{\sigma_\xi^2} = \sigma_\theta^2\left(1 - r^2\right). \end{aligned}$$

Wniosek 1.5 *Jeżeli θ i $\xi_1,...,\xi_k$ są zmiennymi losowymi o rozkładach normalnych $\boldsymbol{N}\left(m_\theta,\sigma_\theta^2\right)$ i $\boldsymbol{N}\left(m_1,\sigma_1^2\right),...,\boldsymbol{N}\left(m_k,\sigma_k^2\right)$ odpowiednio oraz zmienne losowe $\xi_1,...,\xi_k$ są niezależne, to warunkowa wartość oczekiwana zmiennej losowej θ ze względu na $\xi_1,...,\xi_k$ wynosi*

$$E\left(\theta\,|\xi_1,...,\xi_k\,\right) = m_\theta + \sum_{i=1}^{k}\frac{cov\left(\theta,\xi_i\right)}{\sigma_i^2}\left(\xi_i - m_i\right).$$

Poniżej zakładamy, że funkcje $f_0\left(t,\xi\right)$, $f_1\left(t,\xi\right)$, $\sigma\left(t,\xi\right)$ w równaniu (1.16) spełniają następujące założenia:

1. dla dowolnego $t \in \mathbb{N}$ wektor $f_0\left(t,\xi\right) \in \mathbb{R}^n$ oraz macierze $f_1\left(t,\xi\right) \in \mathbb{R}^{n\times k}$, $\sigma\left(t,\xi\right) \in \mathbb{R}^{n\times m}$ są mierzalne ze względu na $\sigma-$ciało $\mathcal{F}_t^\xi$;

2. z prawdopodobieństwem jeden elementy wektora $f_0\left(t,\xi\right)$ i macierzy $f_1\left(t,\xi\right),\sigma\left(t,\xi\right)$ są ograniczone, tzn. dla $t \in \mathbb{N}_0$

$$\begin{aligned} &P\left(\left|f_{0\;i}\left(t,\xi\right)\right| \le c\right) = 1,\; P\left(\left|f_{1\;ij}\left(t,\xi\right)\right| \le c\right) = 1, \\ &P\left(\left|\sigma_{is}\left(t,\xi\right)\right| \le c\right) = 1 \end{aligned}$$

dla $1 \leq i \leq n$, $1 \leq j \leq k$, $1 \leq s \leq m$, $c < \infty$ oraz

$$E\left|f_0(t,\xi)\right|^2 < \infty,\ E\left|f_1(t,\xi)\right|^2 < \infty,\ E\left|\sigma(t,\xi)\right|^2 < \infty;$$

3. warunkowy rozkład wektora losowego $\theta : \Omega \to \mathbb{R}^k$ ze względu na $\mathcal{F}_0^\xi$ jest rozkładem normalnym $\boldsymbol{N}(m_0, Q_0)$.

Procesy stochastyczne z czasem dyskretnym określone za pomocą równania (1.16) mają własność warunkowej normalności. Twierdzenie poniżej podaje sposób wyznaczenia warunkowego rozkładu wektora losowego θ na podstawie realizacji $\{\xi_s\}_{0\leq s\leq t}$ dla dowolnego $t \in \mathbb{N}_0$. Wartości estymatorów wektora losowego θ wyznaczone zaostaną za pomocą twierdzenia o korelacji normalnej.

Twierdzenie 1.13 *Jeżeli spełnione są założenia (1)-(3), to warunkowy rozkład wektora losowego θ ze względu na $\sigma-$ciało $\mathcal{F}_t^\xi$ jest rozkładem normalnym $\boldsymbol{N}(m_t, Q_t)$, gdzie wektor warunkowych wartości oczekiwanych m_t oraz macierz błędu estymacji γ_t określone są za pomocą równań rekurencyjnych*

$$\begin{aligned} m_{t+1} &= m_t + \gamma_t f_1^T(t,\xi)\left(f_1(t,\xi)\gamma_t f_1^T(t,\xi) + \sigma(t,\xi)\sigma^T(t,\xi)\right)^{-1} \\ &\quad \times\left(\xi_{t+1} - f_0(t,\xi) - f_1(t,\xi)m_t\right), \qquad (1.24) \\ Q_{t+1} &= \gamma_t - \gamma_t f_1^T(t,\xi) \\ &\quad \times\left(f_1(t,\xi)\gamma_t f_1^T(t,\xi) + \sigma(t,\xi)\sigma^T(t,\xi)\right)^{-1} A_1(t,\xi)\gamma_t \qquad (1.25) \end{aligned}$$

z warunkami początkowymi m_0 i Q_0 odpowiednio.

Dowód. Prawdziwość twierdzenia udowodniona zostanie za pomocą indukcji matematycznej. Załóżmy, że dla dowolnego $t \geq 1$ warunkowy rozkład $P\left(\theta \leq x \middle| \mathcal{F}_t^\xi\right)$ jest rozkładem normalnym $\boldsymbol{N}(m_t, Q_t)$, gdzie $m_t = E\left(\theta \middle| \mathcal{F}_t^\xi\right)$ oraz $Q_t = cov\left(\theta, \theta \middle| \mathcal{F}_t^\xi\right)$. Z równania (1.16) wynika, że warunkowa wartość oczekiwana wektora ξ_{t+1} ze względu na $\sigma-$ciało $\mathcal{F}_t^\xi$ jest równa

$$E\left(\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right) = f_0\left(t,\xi\right) + f_1\left(t,\xi\right)m_t, \tag{1.26}$$

natomiast macierz warunkowej kowariancji jest równa

$$\begin{aligned} Q_t^{\xi} &= cov\left(\xi_{t+1},\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right) \\ &= E\left(\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right)\right)\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right)\right)^T\middle|\mathcal{F}_t^{\xi}\right) \\ &= f_1\left(t,\xi\right)Q_t f_1^T\left(t,\xi\right) + \sigma\left(t,\xi\right)\sigma^T\left(t,\xi\right). \end{aligned} \tag{1.27}$$

Macierz warunkowej kowariancji wektorów θ i ξ_{t+1} jest dana wzorem

$$\begin{aligned} Q_t^{\theta\xi} &= cov\left(\theta,\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right) \\ &= E\left(\left(\theta - m_t\right)\left(\left(f_1\left(t,\xi\right)\left(\theta - m_t\right) + \sigma\left(t,\xi\right)\varepsilon_{t+1}\right)\right)^T\middle|\mathcal{F}_t^{\xi}\right) \\ &= Q_t A_1^T\left(t,\xi\right). \end{aligned} \tag{1.28}$$

Korzystając z twierdzenia o korelacji normalnej, otrzymujemy

$$E\left(\theta\middle|\mathcal{F}_{t+1}^{\xi}\right) = E\left(\theta\middle|\mathcal{F}_t^{\xi}\right) + Q_t^{\theta\xi}\left(Q_t^{\xi}\right)^{-1}\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right)\right). \tag{1.29}$$

Zdefiniujmy wektor losowy μ_t postaci

$$\begin{aligned} \mu_t = \theta - E\left(\theta\middle|\mathcal{F}_{t+1}^{\xi}\right) &= \theta - E\left(\theta\middle|\mathcal{F}_t^{\xi}\right) \\ &- Q_t^{\theta\xi}\left(Q_t^{\xi}\right)^{-1}\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^{\xi}\right)\right). \end{aligned} \tag{1.30}$$

Ze wzoru (1.30) otrzymujemy

$$\begin{aligned} E\left(\mu_t\middle|\mathcal{F}_t^{\xi}\right) &= \bar{0}, \\ cov\left(\mu_t,\mu_t\middle|\mathcal{F}_t^{\xi}\right) &= Q_t - Q_t^{\theta\xi}\left(Q_t^{\xi}\right)^{-1}\left(Q_t^{\theta\xi}\right)^T. \end{aligned}$$

Warunkowa macierz kowariancji wektorów μ_t i $\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^\xi\right)$ jest równa

$$E\left(\mu_t\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^\xi\right)\right)^T\middle|\mathcal{F}_t^\xi\right) = Q_t^{\theta\xi} - Q_t^{\theta\xi}\left(Q_t^\xi\right)^{-1}Q_t^\xi = \emptyset, \quad (1.31)$$

gdzie wektor $\bar{0} \in \mathbb{R}^k$ jest wektorem o współrzędnych zerowych (wektor zerowy) oraz macierz $\emptyset \in \mathbb{R}^{k\times n}$ jest macierzą o elementach zerowych (macierz zerowa). Zatem μ_t i $\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^\xi\right)$ jak również μ_t i ξ_{t+1} są stochastycznie niezależne. Z równania (1.30) wynika, że funkcja charakterystyczna wektora losowego θ ze względu na $\sigma-$ciało $\mathcal{F}_{t+1}^\xi$ jest równa

$$\begin{aligned}
&E\left(\exp\left(iz^T\theta\right)\middle|\mathcal{F}_{t+1}^\xi\right)\\
&= \exp\left(iz^T\left(E\left(\theta\middle|\mathcal{F}_t^\xi\right) + Q_t^{\theta\xi}\left(Q_t^\xi\right)^{-1}\left(\xi_{t+1} - E\left(\xi_{t+1}\middle|\mathcal{F}_t^\xi\right)\right)\right)\right)\\
&\times E\left(\exp\left(iz^T\mu_t\right)\middle|\mathcal{F}_{t+1}^\xi\right). \qquad (1.32)
\end{aligned}$$

Z niezależności μ_t i ξ_{t+1} otrzymujemy

$$\begin{aligned}
E\left(\exp\left(iz^T\mu_t\right)\middle|\mathcal{F}_{t+1}^\xi\right) &= E\left(\exp\left(iz^T\mu_t\right)\middle|\mathcal{F}_t^\xi\right)\\
&\exp\left(-\frac{1}{2}z^T\left(Q_t - Q_t^{\theta\xi}\left(Q_t^\xi\right)^{-1}\left(Q_t^{\theta\xi}\right)^T\right)z\right). \qquad (1.33)
\end{aligned}$$

Podstawiając (1.26)-(1.28) oraz (1.33) do wzoru (1.32) otrzymujemy

$$E\left(\exp\left(iz^T\theta\right)\middle|\mathcal{F}_{t+1}^\xi\right) = \exp\left(iz^Tm_{t+1} - \frac{1}{2}z^TQ_{t+1}z\right),$$

gdzie m_{t+1} i Q_{t+1} są określone za pomocą wzorów (1.24)-(1.25). Stąd warunkowy rozkład wektora losowego θ ze względu na $\sigma-$ciało $\mathcal{F}_{t+1}^\xi$ jest rozkładem normalnym $\boldsymbol{N}\left(m_{t+1}, Q_{t+1}\right)$. ■

Twierdzenie poniżej przedstawia efektywne rozwiązania równań rekurencyjnych (1.24)-(1.25), które reprezentują wektor warunkowych wartości oczekiwanych m_t oraz macierz błędu estymacji Q_t wektora losowego θ.

Twierdzenie 1.14 *Jeżeli dla $t \in \mathbb{N}_0$ macierze $B(t,\xi)\, B^T(t,\xi)$ są nieosobliwe, to rozwiązania równań (1.24)-(1.25) dane są wzorami*

$$
\begin{aligned}
m_{t+1} = {} & \left(I + Q_0 \sum_{i=0}^{t} f_1^T(i,\xi) \left(\sigma(i,\xi)\, \sigma^T(i,\xi) \right)^{-1} f_1(i,\xi) \right)^{-1} \\
& \times \left(m_0 + Q_0 \sum_{i=0}^{t} f_1^T(i,\xi) \left(\sigma(i,\xi)\, \sigma^T(i,\xi) \right)^{-1} \left(\xi_{i+1} - f_0(i,\xi) \right) \right), \qquad (1.34)
\end{aligned}
$$

$$
Q_{t+1} = \left(I + Q_0 \sum_{i=0}^{t} f_1^T(i,\xi) \left(\sigma(i,\xi)\, \sigma^T(i,\xi) \right)^{-1} f_1(i,\xi) \right)^{-1} Q_0, \qquad (1.35)
$$

gdzie $I \in \mathbb{R}^{k \times k}$ jest macierzą jednostkową.

Dowód twierdzenia można znaleć w [68], [79].

Rozdział 2

Optymalne sterowanie adaptacyjne dla horyzontu deterministycznego

Problemy optymalnego sterowania systemami deterministycznymi i stochastycznymi są dobrze znane w literaturze przedmiotu zarówno dla czasu dyskretnego, jak i dla czasu ciągłego (patrz np. [1], [3], [4], [23], [36], [38], [85], [91], [92], [97], [98], [108], [114]). Nie mniej jednak wpływ czynnika losowego na system powoduje istotną różnicę w sterowaniu obiektem stochastycznym, w odróżnieniu od sterowania obiektem deterministycznym. Czynnik losowy może być spowodowany zarówno nieznajomością parametrów systemu, jak i zaburzeniem zewnętrznym występującym w modelu. Jak zobaczymy poniżej, horyzont także wpływa na realizację celu sterowania.

Niech $(\Omega, \mathcal{F}, P)$ będzie zupełną przestrzenią probabilistyczną, $U \subset \mathbb{R}^l$ (lub $U \subset \mathbb{C}^l$), $l \in \mathbb{N}$ będzie zbiorem dopuszczalnych sterowań, natomiast $\mathcal{Y} \subset \mathbb{R}^n$ (lub $\mathcal{Y} \subset \mathbb{C}^n$), $n \in \mathbb{N}$ – przestrzenią stanów (sygnałów wyjściowych).

Definicja 2.1 *Sterowany system stochastyczny jest to system, w którym sygnałowi wejściowemu $u \in U$ odpowiada wektor losowy $y : \Omega \times U \to \mathcal{Y}$ o rozkładzie prawdopodobieństwa P_y w przestrzeni mierzalnej sygnałów wyjściowych $(\mathcal{Y}, \mathcal{B})$.*

Poniżej pojęcia system stochastyczny oraz sterowany system stochastyczny będziemy używać zamiennie. Funkcja P_y jest miarą określoną na σ–ciele $\mathcal{B}$ podzbiorów zbioru $\mathcal{Y}$ oraz dla każdego $S \in \mathcal{B}$ jest funkcją zmienniej u, tzn. $P_y(S\,|u)$. Niech $T \subset \mathbb{R}_+$ będzie zbiorem momentów decyzyjnych, w którym sterujemy systemem. Zachowanie systemu stochastycznego w dowolnym momencie $t \in T$ modelujemy za pomocą funkcji $f : T \times U \times \Omega \to \mathcal{Y}$.

Definicja 2.2 *Dyskretnym sterowanym systemem stochastycznym nazywamy proces stochastyczny $\{y_t\}_{t\in T}$ z czasem dyskretnym, gdzie $y_t \overset{def}{=} f(t, u_{t-s}, \omega)$ dla każdego momentu decyzyjnego $t \in T$, natomiast $s \in \mathbb{R}_+$ oznacza czas odpowiedzi systemu na sygnał wejściowy u_{t-s}.*

Zatem dla ustalonego momentu $t \in T$, czasu odpowiedzi $s > 0$ i sterowania $u_{t-s} \in U$ sygnał wyjściowy $y_t(\omega) = f(t, u_{t-s}, \omega)$ jest wektorem losowym (zmienną losową w przypadku jednowymiarowym)

$$y_t : \Omega \to \mathcal{Y}.$$

Ciąg sterowań $\{u_t\}_{t\in T}$ nazywamy regulatorem. Przy ustalonym $\omega \in \Omega$ oraz ustalonym regulatorze $\{u_t\}_{t\in T}$ ciąg $\{y_t\}_{t\in T}$ jest ciągiem wektorów o wartościach w przestrzeni stanów, który nazywamy realizacją (trajektorią) systemu stochastycznego $\{y_t\}_{t\in T}$.

Definicja systemu stochastycznego opiera się na pojęciu wektora losowego y_t (w przypadku jednowymiarowym - zmiennej losowej), który zależy od parametru t (czasu) i sterowania u_{t-s}. Wobec powyższego mamy do czynienia ze sparametryzowaną rodziną wektorów (lub zmiennych) losowych. Rozkłady tych zmiennych losowych, a w szczególności ich pierwsze i drugie momenty również mogą zależeć od czasu t i sterowania u_{t-s}. Zazwyczaj zakłada się, że odstępy pomiędzy momentami, w których obserwujemy stany systemu, są jednakowe $t_{i+1} - t_i = s > 0$ dla $i \in \mathbb{N}_0$ (stany systemu są obserwowane w momentach $0, s, 2s, 3s, ...$). Dla uproszczenia przyjmujemy, że czas odpowiedzi systemu na sygnał jest równy jeden ($s = 1$). W zależności od potrzeb jako zbiór momentów decyzyjnych będziemy przyjmować $T = \mathbb{N}_0$ lub $T = \{0, 1, ..., N\} \subset \mathbb{N}_0$.

Głównym celem teorii sterowania jest wyznaczenie regulatora $\{u_t\}_{t\in T}$, który spełnia określone własności. W zależności od typu własności rozwiązywane są różnorodne zagadnienia szczegółowe, np. zagadnienia dotyczące sterowałności, stabilności, optymalności (patrz np. [30], [49], [55], [109]). W teorii optymalnego sterowania zadanie polega na wyznaczeniu ciągu sterowań $\{u_t^*\}_{t\in T}$, dla którego wskanik jakości osiąga wartość najmniejszą (lub największą). Wskaźnik jakości w wielu przypadkach zależy od stanów systemu i sterowań oraz definiujemy $J\left(\{y_t\}_{t\in T}, \{u_t\}_{t\in T}\right)$.

Zasadniczo zachowanie każdego systemu, którym sterujemy, zależy od parametrów. Sterując systemem stochastycznym, nie zawsze posiadamy informację o wartościach tych parametrów. W przypadku gdy parametry systemu są znane sterującemu, mamy do czynienia z kompletną informacją o systemie. Jeżeli parametry systemu nie są znane, to w trakcie realizacji głównego celu sterowania dodatkowo musimy wzbogacić naszą wiedzę o nich. Przyjmujemy, że działanie sterujące podejmowane w chwili $i-$ tej może bazować tylko na obserwacjach poprzednich stanów systemu, tj. $y_0, y_1, ..., y_i$ oraz na znajomości rozkładów a priori parametrów systemu. Jednocześnie sterując systemem i obserwując jego stany, wzbogacamy naszą wiedzę odnośnie funkcjonowania systemu, co powoduje, że możemy dokładniej nim sterować. Zatem rozkład a posteriori w chwili $i-$tej charakteryzuje stan wiedzy o parametrach systemu zawartej w obserwacji $y_1, ..., y_i$, tym samym zależy od działań sterujących podjętych w chwilach poprzedzających, które bezpośrednio wpływają na obserwowane stany. Stąd widzimy, że optymalne sterownie systemem stochastycznym z nieznanymi parametrami ma podwójne zadanie: z jednej strony działanie systemu polega na gromadzeniu wiedzy o parametrach, z drugiej zaś na optymalnej realizacji celu.

Definicja 2.3 *Sterowanie adaptacyjne jest to metoda sterowania systemami z nieznanymi parametrami lub parametrami zmieniającymi się w czasie, gdzie te parametry są identyfikowane (dopasowywane, dostrajane) poprzez regulator.*

Wobec powyższego metodę sterowania, w której występują działania sterowania i uczenia się jednocześnie, nazywamy sterowaniem adaptacyjnym (patrz np. [4], [6]–[8], [10]–[12], [16]–[18], [34], [35], [38], [57], [72], [78], [85], [87], [91], [92], [97], [98], [103], [108], [110], [118]). Jak zobaczymy poniżej, koszty sterowania systemami z pełną i niepełną informacją są różne. Różnicę tych kosztów nazywamy kosztem uczenia się.

W dalszej kolejności omówiony zostanie problem wyboru optymalnego horyzontu. Mianowicie realizując to samo zadnie dla różnych horyzontów czasowych uzyskamy różne efekty: różne prawa sterowania oraz różne wielkości kosztów.

2.1 Optymalne sterowanie systemami stochastycznymi z czasem dyskretnym

2.1.1 Postawienie zadania

Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy ciąg $\{w_i\}_{1\le i\le N}$ niezależnych wektorów losowych $w_i : \Omega \to \mathbb{R}^m$ o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, gdzie $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m\times m}$– macierzą jednostkową. Niech $\xi : \Omega \to \mathbb{R}^k$ będzie wektorem losowym o rozkładzie a priori $P\,(d\xi)$, natomiast $y_0 : \Omega \to \mathbb{R}^n$ – stanem początkowym o rozkładzie $P\,(dy_0)$. Jeżeli stan początkowy y_0 dla sterowanego systemu stochastycznego jest znany, to jako rozkład stanu początkowego przyjmujemy delta-miarę Diraca zcentrowaną w punkcie y_0, tzn.

$$P\,(y) = \begin{cases} 0, & \text{gdy} \quad y \neq y_0, \\ 1, & \text{gdy} \quad y = y_0. \end{cases}$$

Zakładamy stochastyczną niezależność wszystkich wymienionych obiektów. Definiujemy niemalejące rodziny $\sigma-$ciał $\{\mathcal{F}_i\}_{0\le i\le N}$ oraz $\left\{\mathcal{F}_i^{\xi}\right\}_{0\le i\le N}$, gdzie $\mathcal{F}_i = \sigma\{y_0\} \vee \sigma\{w_s : s = 1, 2, ..., i\}$, $\mathcal{F}_i^{\xi} = \mathcal{F}_i \vee \sigma\{\xi\}$ oraz przyjmujemy $\mathcal{F} = \mathcal{F}_N^{\xi}$. Poniżej rozważamy system stochastyczny o równaniu stanu

$$y_{i+1} = f(\xi, y_i, u_i) + \sigma(\xi, y_i) w_{i+1}, \tag{2.1}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ oraz $\sigma : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{M}(n, m)$, gdzie $\mathbb{M}(n, m)$ jest zbiorem macierzy o $n-$ wierszach i $k-$ kolumnach. O funkcjach f, σ zakładamy, że są ciagłe względem wszystkich argumentów. Sterując systemem stochastycznym (2.1) z czasem dyskretnym podejmujemy działania sterujące w chwilach $0, 1, ..., N-1$. Wielkość N oznacza horyzont sterowania. Dla dowolnego $0 \leq i \leq N-1$ wektor $u_i \in \mathbb{R}^l$ mierzalny ze względu na σ-ciało $\mathcal{F}_i$ nazywamy działaniem sterującym.

Dla ustalonego horyzontu czasowego N sterując systemem (2.1) w momentach $i = 0, 1, ..., N-1$ podejmujemy ciąg działań $\{u_i\}_{0 \leq i \leq N-1}$. Wobec powyższego sterowanie $u = (u_0,\ u_1, ..., u_{N-1})$ nazywamy sterowaniem dopuszczalnym. Na ciąg działań sterujących $\{u_i\}_{0 \leq i \leq N-1}$ nie nakładamy dodatkowych ograniczeń. Klasę sterowań dopuszczalnych oznaczamy przez U.

Aby sformułować cel sterowania wprowadzimy funkcję strat $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ oraz funkcję dziedziczenia $h : \mathbb{R}^n \longrightarrow \mathbb{R}$. Zakładamy, że $g(y, u)$ oraz $h(y)$ sa ciągłe i ograniczone. Wskaźnik jakości definiujemy jako sumę strat w chwilach $j = 0, 1, ..., N-1$ oraz dziedziczenia na końcu horyzontu czasowego N. Funkcja celu reprezentuje stratę całkowitą oraz jest określona wzorem

$$J^N(u) = E\left(\sum_{i=0}^{N-1} g(y_i, u_i) + h(y_N)\right). \tag{2.2}$$

Problem sterowania adaptacyjnego systemem stochastycznym (2.1) polega na minimalizacji strat całkowitych. Aby wyznaczyć optymalne sterowanie systemem (2.1) należy rozwiązać zadanie

$$\inf_{u \in U} J^N(u), \tag{2.3}$$

tzn. należy wyznaczyć takie sterowanie dopuszczalne $u^* = \left(u_0^*, ..., u_{N-1}^*\right)$ dla którego infimum jest osiągnięte.

Dynamika systemu (2.1) zależy od parametrów ξ. Jeżeli dokładna wartość tych parametrów nie jest znana sterującemu, to realizując główny cel sterowania, który polega na minimalizacji wskaźnika jakości (2.2), dodatkowo obserwujemy trajektorię systemu oraz analizujemy jego zachowanie, tym samym wzbogacamy naszą wiedzę o działaniu tego systemu, dokładniej o parametrach ξ. Sterowanie adaptacyjne polega na odpowiednio szybkim gromadzeniu wiedzy o systemie (2.1) oraz optymalizacji wskaźnika jakości (2.2).

Do rozwiązania zadania (2.3) zostanie wykorzystane podejście R.Rishela (patrz np. [91]–[94]). Dla systemów z niepełną informacją o parametrach systemu (2.1) dodatkowo wykorzystamy zaganienie filtracji, które polega na oszacowaniu warunkowego rozkładu nieznanych parametrów ξ.

Twierdzenie 2.1 *Jeżeli funkcje $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ i $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ są ciągłe, wypukłe i ograniczone, funkcje $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ są funkcjami klasy C^1 ze względu na zmienną sterującą $u \in \mathbb{R}^l$ oraz $\det(\Sigma(\xi, y)) \neq 0$ dla $(\xi, y) \in \mathbb{R}^k \times \mathbb{R}^n$, gdzie $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$, to spełnienie równania*

$$\bigtriangledown_u g(y_j, u_j^*) + E\left(\left(\sum_{i=j+1}^{N-1} g(y_i, u_i^*) + h(y_N)\right)\big(y_{j+1} - f(\xi, y_j, u_j^*)\big)^T\right.$$
$$\left. \times\Sigma^{-1}(\xi, y_i) \bigtriangledown_u f(\xi, y_j, u_j^*) \,|\mathcal{F}_j\right) = \mathbf{0} \qquad (2.4)$$

dla wszystkich $j \in (0, 1, , ..., N-1)$ jest warunkiem koniecznym optymalnego sterowania $u^* = \left(u_0^*, u_1^*, ..., u_{N-1}^*\right)$, gdzie $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^l$.

Dowód. W równaniu stanu (2.1) wektor losowy $w_{j+1} : \Omega \to \mathbb{R}^m$ ma rozkład normalny $\boldsymbol{N}(\bar{0}, I_m)$. Zatem dla dowolnego $0 \leq j \leq N-1$ warunkowy rozkład stanu systemu jest dany wzorem

$$P\left(dy_{j+1}\middle|\mathcal{F}_j^\xi\right) = \gamma(y_{j+1}, f(\xi, y_j, u_j), \Sigma(\xi, y_j))dy_{j+1},$$

gdzie funkcja gęstości $\gamma()$ jest określona za pomocą (1.4).

W przypadku gdy parametry ξ są znane, to dla dowolnego $Y \subset \mathbb{R}^n$ wielkość $P\left(y_{j+1} \in Y \middle| \mathcal{F}_j^{\xi}\right) = \int_Y P\left(ds \middle| \mathcal{F}_j^{\xi}\right)$ oznacza prawdopodobieństwo, że w chwili $j+1$ realizacja systemu (2.1) będzie należała do zbioru Y, jeżeli do momentu j ten system pokonał trajektorię $\{y_i : 0 \leq i \leq j\}$. Z niezależności zaburzeń zewnętrznych $w_1, ..., w_N$ wynika, że łączny rozkład zmiennych $(\xi, y_0, y_1, ..., y_j)$, dla $1 \leq j \leq N$ jest dany wzorem

$$P(d\xi, dy_0, \ldots, dy_i) = P(d\xi)\, P(dy_0)\, P(dy_1, \ldots, dy_j) \tag{2.5}$$

oraz dla $1 \leq i \leq j \leq N$

$$P(dy_i, \ldots, dy_j) = \prod_{k=i}^{j} P\left(dy_k \middle| \mathcal{F}_{k-1}\right). \tag{2.6}$$

Zatem dla dowolnego $0 \leq j \leq N-1$ wskaźnik jakości (2.2) możemy przedstawić w postaci

$$J^N(u) = E\left(\sum_{i=0}^{j} g(y_i, u_i) + E\left(\sum_{i=j+1}^{N-1} g(y_i, u_i) + h(y_N) \middle| \mathcal{F}_j^{\xi}\right)\right)$$

lub

$$J^N(u) = \int \left(\sum_{i=1}^{j} g(y_i, u_i)\right) P(d\xi, dy_0, dy_1 ... dy_j) +$$
$$\int \left(\int \left(\sum_{i=j+1}^{N-1} g(y_i, u_i) + h(y_N)\right) P(dy_{j+1}, ..., dy_N)\right) P(d\xi, dy_0, dy_1 ... dy_j). \tag{2.7}$$

Ustalmy liczbę $j \in \{0, \ldots, N-1\}$. Niech $u = u^* + \epsilon v_j$, gdzie $u^* = \left(u_0^*, u_1^*, ..., u_{N-1}^*\right) \in \mathbb{R}^{l \times N}$ oznacza sterowanie optymalne systemem (2.1), $\epsilon-$ skalar, natomiast $v_j : \mathbb{R}^{n \times (j+1)} \longrightarrow \mathbb{R}^{l \times N}$ jest dane wzorem

$v_j = \left(c_0^j, c_1^j, ..., c_{N-1}^j\right)$, gdzie

$$c_i^j = \begin{cases} \operatorname{col}(\kappa_1, ..., \kappa_l), & \text{dla } \; i = j, \\ \operatorname{col}(0, ..., 0), & \text{dla } \; i \neq j \end{cases}$$

oraz $\kappa_s(y_0, ..., y_j) : \mathbb{R}^{n \times (j+1)} \longrightarrow \mathbb{R}$ dla $1 \leq s \leq l$ jest dowolną funkcją borelowską.

W momencie j dla sterowania u_j^* realizującego infimum funkcjonału J zdefiniowanego wzorem (2.2) wariacja powinna być równa zero, innymi słowy, jeżeli sterowanie u_j^* jest optymalne, to znika wariacja funkcjonału J (patrz np. [40], [111], [119]). Zatem warunek konieczny istnienia optymalnego sterowania w momencie j możemy przedstawić za pomocą równania

$$\left.\frac{\partial}{\partial \epsilon} J^N(u^* + \epsilon v_j)\right|_{\epsilon \to 0} = 0. \tag{2.8}$$

Korzystając z rozwinięcia (2.7) otrzymujemy

$$\frac{\partial}{\partial \epsilon} J^N(u^* + \epsilon v_j) = \int \left\langle \left(\nabla_u g(y_j, u_j^* + \epsilon v_j) + \int \left(\sum_{i=j+1}^{N-1} g(y_i, u_i^* + \epsilon v_j) + h(y_N) \right) \right.\right.$$
$$\left.\left. \times \nabla_u P(dy_{j+1}, ..., dy_N) \right), c_j^j \right\rangle P(d\xi, dy_0, dy_1 ... dy_j). \tag{2.9}$$

Ze wzorów (2.5)-(2.6) mamy

$$\nabla_u P_{j+1,N}(dy_{j+1}, ..., dy_N) = \left(y_{j+1} - f(\xi, y_j, u_j^* + \epsilon v_j)\right)^T \Sigma^{-1}(\xi, y_j)$$
$$\times \nabla_u f(\xi, y_j, u_j^*) P(dy_{j+1}, ..., dy_N). \tag{2.10}$$

Podstawiając (2.10) do (2.9) oraz korzystając z warunku (2.8) otrzymujemy równanie

$$\int \left\langle \left(\nabla_u g(y_j, u_j^*) + \int \left(\sum_{i=j+1}^{N-1} g(y_i, u_i^*) + h(y_N) \right) \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \right.\right.$$
$$\left.\left. \times \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j^*) P(dy_{j+1}, ..., dy_N) \right), c_j^j \right\rangle P(d\xi, dy_0, dy_1 ... dy_j) = 0. \tag{2.11}$$

Ponieważ warunek (2.11) musi być spełniony dla dowolngo $c_j^j = \mathrm{col}\,(\kappa_1, ..., \kappa_l)$, gdzie κ_s, $1 \le s \le l$ są funkcjami borelowskimi mierzalnymi względem $\sigma-$ ciała $\mathcal{F}_j$, zatem otrzymujemy tezę. ■

Uwaga 2.1 *Jeżeli we wzorze (2.2) funkcjonał strat $g\,()$ dodatkowo zależy od parametrów systemu ξ (tzn. $g(\xi, y_j, u_j)$) lub dodatkowo od przyszłych stanów (np. $g(\xi, y_j, y_{j+1,} u_j)$ lub $g(y_j, y_{j+1,} u_j)$), to warunek konieczny optymalnego sterowania (2.4) jest podobny, z tym że w wyżej wymienionych przypadkach gradient funkcjonału strat $\bigtriangledown_u g(\cdot)$ w równaniu (2.4) wchodzi pod znak operatora warunkowej wartości oczekiwanej. Dowody przebiegają w sposób identyczny oraz można je znaleźć np. w pracach [6], [11], [16], [15].*

2.1.2 Wyznaczenie optymalnego sterowania

Rozwiązanie zadania (2.3) polega na wyznaczeniu ciągu optymalnych sterowań $\{u_i^*\}_{0 \le i \le N-1}$ systemem (2.1). Do wyznaczenia optymalnych sterowań zostanie wykorzystane programowanie dynamiczne. Dla dowolnego $0 \le j \le N$ definiujemy funkcje Bellmana (patrz [23]) w sposób następujący

$$V_j^N(\xi, y_0, ..., y_j) = E\left(\sum_{i=j}^{N-1} g\,(y_i, u_i^*) + h\,(y_N) \middle| \mathcal{F}_j^\xi \right), \tag{2.12}$$

natomiast dla momentu N przyjmujemy

$$V_N^N(\xi, y_0, ..., y_N) = h(y_N). \tag{2.13}$$

Funkcja Bellmana $V_j^N(\xi, y_0, ..., y_j)$ reprezentuje oczekiwane koszty sterowania i dziedziczenia od momentu j do końca horyzontu N dla przypadku gdy znane są parametry $\xi \in \mathbb{R}^k$ systemu (2.1). Korzystając z własności warunkowej wartości oczekiwanej otrzymujemy

$$\begin{aligned} V_j^N(\xi, y_0, ..., y_j) &= \inf_{u_j} \left(g(y_j, u_j) + E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1}) \middle| \mathcal{F}_j^\xi \right) \right) \\ &= g(y_j, u_j^*) + E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1}) \middle| \mathcal{F}_j^\xi \right), \end{aligned} \tag{2.14}$$

gdzie u_j^* oznacza sterowanie optymalne w chwili j. Widzimy, że funkcje Bellmana mają kształt teleskopowy.

Ze wzoru (2.14) wynika, że trajektorii $y_0, y_1, ..., y_j$ oraz parametrom ξ odpowiada sterowanie u_j^* dla $0 \leq j < N-1$. Oczywiście pokonując trajektorię $y_0, y_1, ..., y_j$ na system (2.1) oddziaływał ciąg sterowań $u_0, u_1, ..., u_{j-1}$. Dla uproszczenia przyjmujemy $V_j^N(\xi, y_0, ..., y_j) = V_j^N(\xi, y_0, ..., y_j, u_0, ..., u_{j-1})$.

Jeżeli parametry $\xi \in \mathbb{R}^k$ są nieznane, to wielkość

$$\begin{aligned}\bar{V}_j^N(y_0, ..., y_j) &= E\left(V_j^N(\xi, y_0, ..., y_{j+1})\,|\mathcal{F}_j\right) \\ &= \inf_{u_j}\left(g\left(y_j, u_j\right) + E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1})\,|\mathcal{F}_j\right)\right) \qquad (2.15)\end{aligned}$$

reprezentuje oczekiwane koszty sterowania i dziedziczenia od momentu j do końca horyzontu N.

Uwaga 2.2 *Funkcja Bellmana określona wzorami (2.13)–(2.14) ma następujące własności:*

1. *dla ustalonego parametru $\xi \in \mathbb{R}^k$ proces stochastyczny $\{y_t\}_{1 \leq t \leq N}$ dany wzorem (2.1) ma własność Markowa oraz wartość funkcji Bellmana w momencie j zależy od stanu systemu w tym momencie, tzn.*

$$V_j^N(\xi, y_0, ..., y_j) = V_j^N(\xi, y_j) \qquad (2.16)$$

dla $0 \leq j \leq N$, natomiast na potrzeby konstrukcji algorytmu wyznaczenia optymalnych sterowań oraz oszacowania kosztów sterowań w warunkach niepełnej informacji o systemie w dalszej części pracy wykorzystujemy oznaczenia $V_j^N(\xi, y_0, ..., y_j)$, $0 \leq j \leq N$;

2. *zgodnie ze wzorem (2.16) dla dowolnego $\xi \in \mathbb{R}^k$ i $y \in \mathbb{R}^n$ mamy*

$$V_j^N(\xi, y) = V_0^{N-j}(\xi, y),$$

zatem wartość funkcji Bellmana $V_j^N(\xi, y_0, ..., y_j)$ w momencie $j \in \{0, 1, ..., N\}$ zależy od parametru ξ, stanu systemu y i długości horyzontu $N-j$ (dokładniej od pozostałego czasu sterowania).

Podobnie jak w pracach [7]–[16], [38], [44], [91], [92] przedstawiona zostanie procedura wyznaczenia optymalnego ciągu sterowań $\left\{u_j^*\right\}_{0\leq j\leq N-1}$ dla systemu (2.1). Ze wzorów (2.12) oraz (2.14) otrzymujemy

$$\begin{aligned}
&\bigtriangledown_u g(y_j, u_j^*) + E\left(\left(\sum_{i=j+1}^{N-1} g(y_i, u_i^*) + h(y_N)\right)\right. \\
&\qquad \left.\times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_j) \bigtriangledown_u f(\xi, y_j, u_j^*) \middle| \mathcal{F}_j\right) \\
&\quad = \bigtriangledown_u g(y_j, u_j^*) + E\left(V_{j+1}^N\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j) w_{j+1}\right)\right. \\
&\qquad \left.\times \left(\sigma(\xi, y_j) w_{j+1}\right)^T \Sigma^{-1}(\xi, y_j) \bigtriangledown_u f(\xi, y_j, u_j^*) \middle| \mathcal{F}_j\right).
\end{aligned}$$

Uwaga 2.3 *Dla dowolnego $0 \leq j \leq N-1$ warunek konieczny (2.4) istnienia optymalnego sterowania u_j^* możemy przedstawić za pomocą równania*

$$\begin{aligned}
\bigtriangledown_u g(y_j, u_j^*) + \int V_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j) x) x^T \sigma^T(\xi, y_j) \\
\times \Sigma^{-1}(\xi, y_j) \bigtriangledown_u f(\xi, y_j, u_j^*) \gamma\left(x, \bar{0}, I_m\right) P\left(d\xi \left| \mathcal{F}_j\right.\right) dx = \mathbf{0},
\end{aligned}$$

gdzie $P\left(d\xi \left| \mathcal{F}_j\right.\right)$ oznacza warunkowy rozkład zmiennej losowej ξ względem $\sigma-$ciała $\mathcal{F}_j$.

Dla dowolnych systemów wyznaczenie warunkowego rozkładu $P\left(d\xi \left| \mathcal{F}_j\right.\right)$, $1 \leq j \leq N-1$ zostało omówione w poprzednim rozdziale. W przypadku gdy rozkład a priori wektora nieznanych parametrów ξ jest rozkładem normalnym oraz funkcja $f\left(\right)$ w równaniu (2.1) jest funkcją liniową ze względu na ξ, to warunkowy rozkład $P\left(d\xi \left| \mathcal{F}_j\right.\right)$, $1 \leq j \leq N-1$ jest rozkładem normalnym o statystykach generowanych przez filtr Kalmana – Bucy, wartości których wyznaczamy korzystając z twierdzeń 1.13 - 1.14. Więcej na temat filtracji ciągów warunkowo normalnych można znaleźć np. w [79], [99], [101].

Uwaga 2.4 *Sterując systemem stochastycznym w warunkach pełnej informacji (parametry systemu ξ są znane sterującemu) od momentu j do momentu N ponosimy koszty sterowania wielkości $V_j^N(\xi, y_0, ..., y_j)$ dla $0 \leq j \leq N-1$. W warunkach niepełnej informacji o systemie (parametry systemu są modelowane za pomocą wektora losowego $\xi : \Omega \to \mathbb{R}^k$ oraz nie* *są znane sterującemu) oczekiwany koszt realizacji celu* $\bar{V}_j^N(y_0,...,y_j)$ *(koszty sterowania i uczenia się jednocześnie) od momentu* j *do momentu* N *jest dany wzorem (2.15) oraz zależy bezpośrednio od rozkładu a priori wektora losowego* ξ, *trajektorii* $y_0,...,y_j$ *oraz sterowań* $u_0,...,u_{j-1}$ *(pośrednio), natomiast w przypadku pełnej informacji (patrz wzór (2.16)) zależy od parametrów* ξ *i położenia* y_j. *Porównując koszty sterowania systemami z pełną i niepełną informacją wynika, że dla* $0 \le j \le N-1$ *dodatkowy koszt związany z nieznajomością parametrów* ξ *wynosi*

$$C_j = \bar{V}_j^N(y_0,...,y_j) - V_j^N(\xi, y_0,...,y_j).$$

Koszt C_0 *oznacza oczekiwany całkowity koszt uczenia się podczas realizacji celu sterowania.*

Uwaga 2.5 *Do wyznaczenia optymalnego sterowania w warunkach pełnej i niepełnej informacji o systemie korzystamy z równania (2.4). W warunkach pełnej informacji sterowanie* u_j, $0 \le j \le N-1$ *wyznaczamy na podstawie trajektorii* $(y_0,...,y_j)$ *i wektora* ξ, *zatem przyjmujemy*

$$u_j \overset{def}{=} u_j\left(\xi, y_0,...,y_j\right). \tag{2.17}$$

W przypadku niepełnej informacji sterowanie adaptacyjne u_j, $0 \le j \le N-1$ *wyznaczamy na podstawie trajektorii* $(y_0,...,y_j)$ *i warunkowego rozkładu* $P\left(d\xi\left|\mathcal{F}_j\right.\right)$, *wobec powyższego przyjmujemy*

$$\bar{u}_j \overset{def}{=} \bar{u}_j\left(y_0,...,y_j\right). \tag{2.18}$$

Uwaga 2.6 *W przypadku niepełnej informacji o systemie w zastosowaniach inżynierskich wykorzystuje się sterowanie "dostrojone", które polega na wykorzystaniu sterowań jak w przypadku pełnej informacji o systemie, z tym że nieznany parametr zostaje zastąpiony estymatorem. Ogólnie, sterowanie adaptacyjne różni się od sterowania "dostrojonego", tzn. dla* $0 \le j \le N-1$

$$\bar{u}_j\left(y_0,...,y_j\right) \neq u_j\left(E\left(\xi\left|\mathcal{F}_j\right.\right), y_0,...,y_j\right).$$

Poniżej przedstawiony zostanie algorytm wyznaczenia ciągu optymalnych sterowań $\left\{u_j^*\right\}_{0\leq j\leq N-1}$ w warunkach niepełnej informacji. Podobnie jak w pracach [7]–[16], [57]–[66] przedstawiony poniżej algorytm wykorzystuje metodę programowania dynamicznego wstecz (patrz np. [23], [38], [44], [92], [109]), gdzie za pomocą techniki teleskopowej konstruowane są funkcje Bellmana (2.14).

Algorytm wyznaczania optymalnego sterowania.

1. Definiujemy
$$V_N^N(\xi, y_0, ..., y_N) = h(y_N) \quad \text{oraz kładziemy} \quad j = N.$$

2. Kładziemy
$$j = j - 1.$$

3. Definiujemy
$$\tilde{V}_{j+1}^N(\xi, y_0, ..., y_j, u_j, w_{j+1}) = V_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)w_{j+1}).$$

4. Wyznaczamy
$$\begin{aligned} &Z_j^N(y_0, ..., y_j, u_j) = \\ &E\left(\tilde{V}_{j+1}^N(\xi, y_0, ..., y_j, u_j, w_{j+1})w_{j+1}^T\sigma^T(\xi, y_j)\Sigma^{-1}(\xi, y_j) \bigtriangledown_u f(\xi, y_j, u_j)\Big| \mathcal{F}_j\right) \\ &= \int \tilde{V}_{j+1}^N(\xi, y_0, ..., y_j, u_j, x)x^T\sigma^T(\xi, y_j)\Sigma^{-1}(\xi, y_j) \\ &\times \bigtriangledown_u f(\xi, y_j, u_j)\gamma(x, I_m)\, dx P\left(d\xi\,|\mathcal{F}_j\right). \end{aligned}$$

5. Szukamy optymalnego sterowania u_j^* spełniającego warunek
$$\bigtriangledown_u g(y_j, u_j^*) + Z_j^N(y_0, ..., y_j, u_j^*) = \mathbf{0}.$$

6. Wyznaczamy wartości funkcji Bellmana dla momentu j

$$V_j^N(\xi, y_0, ..., y_j) = g(y_j, u_j^*) +$$
$$\int V_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x)\gamma\,(x, I_m)\,dx.$$

7. Jeżeli $j = 0$ to zatrzymujemy, w przeciwnym razie wracamy do punktu 2.

Uwaga 2.7 *W przypadku pełnej informacji o systemie algorytm wyznaczenia optymalnego sterowania jest podobny. W tym celu wystarczy warunkowy rozkład* $P\left(d\xi\,|\mathcal{F}_j\right)$ *zastąpić rozkładem punktowym.*

2.2 Optymalne sterowanie liniowo-kwadratowe

Problem sterowania liniowo-kwadratowego (*Linear Quadratic Control Problem*) jest dobrze znany w literaturze przedmiotu (patrz np. [13], [36], [38], [44], [57], [71], [92], [104], [108]). Zachowanie obiektu (systemu) jest określone za pomocą równania liniowego ze względu na sterowanie, natomiast wskaźnik jakości jest formą kwadratową. W tej części rozdziału przedstawiony zostanie problem sterowania liniowym systemem stochastycznym z losowym przesunięciem. Ten problem pojawia się, gdy sterujemy obiektem (bądź identycznymi obiektami) w różnych warunkach, natomiast dokładnej informacji o otoczeniu nie posiadamy (np. na dynamikę statku morskiego wpływają prądy wodne oraz warunki atmosferyczne zmieniające się w czasie, na dynamikę statku powietrznego wpływają warunki atmosferyczne). Zatem oprócz sterowania obiektem należy dodatkowo wzbogacać wiedzę o funkcjonowaniu systemu (np. wpływu otoczenia na system).

Dla liniowego systemu stochastycznego wyznaczone zostaną prawa sterowania. Podane zostaną również prawa sterowania systemem liniowym dla przypadków pełnej i niepełnej informacji o systemie. Dodatkowo pokazana zostanie różnica wartości funkcjonałów jakości dla wyżej wymienionych przypadków. Jeżeli parametry systemu nie są znane, to sterujący ponosi dodatkowe koszty, które określamy jako koszty uczenia się. Na koniec rozdziału porównane zostaną sterowania liniowo-kwadratowe dla systemów stochastycznego i deterministycznego (parametry systemu są znane oraz nie występują zaburzenia zewnętrzne w systemie). Jak zobaczymy poniżej nie mniej istotnym czynnikiem wpływającym na wskaźnik jakości jest horyzont sterowania, a mianowicie pokazany zostanie wpływ horyzontu na koszty sterowania liniowymi systemami stochastycznymi i deterministycznymi.

2.2.1 Problem sterowania liniowym obiektem z losowym przesunięciem

Niech $(\Omega, \mathcal{F}, P)$ będzie zupełną przestrzenią probabilistyczną, na której definiujemy ciąg $\{w_i\}_{1 \le i \le N}$ niezależnych wymiarowych wektorów losowych $w_i : \Omega \to \mathbb{R}^m$ o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m \times m}$ – macierzą jednostkową. Niech $\xi : \Omega \to \mathbb{R}^k$ będzie wektorem losowym o rozkładzie normalnym $\boldsymbol{N}(m_0, \Gamma_0)$, gdzie $m_0 \in \mathbb{R}^k$, $\Gamma_0 \in \mathbb{R}^{k \times k}$. Poniżej rozważamy problem sterowania stochastycznym systemem liniowym o równaniu stanu

$$y_{i+1} = Ay_i + B\xi - Cu_i + \sigma w_{i+1}, \tag{2.19}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times k}$, $C \in \mathbb{R}^{n \times l}$, $\sigma \in \mathbb{R}^{n \times m}$. Zakładamy, że spełnione są następujące nierówności $\|A\| < \infty$, $\|B\| < \infty$, $\|C\| < \infty$, $\|\sigma\| < \infty$, gdzie $\|\cdot\|$ oznacza normę macierzy jako $\|A\| = \max_{\|x\| \le 1} \|Ax\|$. Macierz A nazywana jest macierzą transformacji stanów, natomiast k– wymiarowy wektor ξ reprezentuje parametry systemu (np. wpływ otoczenia w którym działa system). Składnik $B\xi$ reprezentuje dodatkowe przesunięcie (wywołane np. losowymi czynnikami zewnętrznymi) systemu (2.19) występujące w przedziale czasowym $(i, i+1]$. Przyjmujemy, że w przypadku pełnej informacji o systemie (2.19) wektor $\xi \in R^k$ jest deterministyczny (np. wpływ otoczenia jest znany sterującemu), zatem przesunięcie $B\xi$ jest stałe. W przypadku niepełnej informacji o systemie $\xi : \Omega \to R^k$ jest wektorem losowym o rozkładzie normalnym, wtedy przesunięcie $B\xi$

również jest wektorem losowym o rozkładzie $\boldsymbol{N}(Bm_0, B\Gamma_0 B^T)$. Zakładamy, że stan początkowy y_0 systemu (2.19) ma rozkład $P(dy_0)$. Jeżeli stan początkowy jest znany (y_0 jest wektorem deterministycznym) to wystarczy przyjąć, że y_0 ma rozkład punktowy. Zakładamy stochastyczną niezależność $w_1, ..., w_N, \xi, y_0$ oraz definiujemy $\sigma-$ ciała $\mathcal{F}_k = \sigma\{y_0\} \vee \sigma\{w_i : i = 1, 2, ..., k\}$ oraz $\mathcal{F}_k^\xi = \mathcal{F}_k \vee \sigma\{\xi\}$, gdzie $\mathcal{F} = \mathcal{F}_N^\xi$. Dla $j = 0, 1, ..., N-1$ wektor $u_j \in \mathbb{R}^l$ mierzalny względem $\sigma-$ ciała $\mathcal{F}_j$ nazywamy działaniem sterującym w chwili j, natomiast ciąg $\{u_i\}_{0 \leq j \leq N-1}$ sterowaniem dopuszczalnym. Klasę wszystkich sterowań dopuszczalnych oznaczamy przez U.

Rozważmy zadanie sterowania stochastycznym systemem liniowym(2.19), gdzie w kolejnych N krokach należy przeprowadzić ten system ze stanu początkowego y_0 do stanu (celu) $a \in \mathbb{R}^n$. Dodatkowo koszty związane ze sterowaniem (nakłady energetyczne) oraz straty spowodowane nietrafieniem do celu były jak najmniejsze. Przyjmujemy, że macierze $R_i \in \mathbb{R}^{l \times l}$ dla $i = 0, ..., N-1$ oraz $Q_N \in \mathbb{R}^{n \times n}$ są dodatnio określone.

Dla momentów $i = 0, 1, ..., N-1$ nakłady energetyczne (koszty związane ze sterowaniem systemu (2.19)) są równe $u_i^T R_i u_i$, natomiast straty związane z nietrafieniem do celu (stanu a) wynoszą $(y_N - a)^T Q_N (y_N - a)$. Widzimy, że zarówno funkcjonały kosztów, jak i funkcjonał dziedziczenia są określone za pomocą form kwadratowych. Wskaźnik jakości reprezentuje koszt całkowity, który jest sumą kosztów sterowania i strat końcowych, oraz jest dany wzorem

$$J^N(u) = E\left(\sum_{i=0}^{N-1} u_i^T R_i u_i + (y_N - a)^T Q_N (y_N - a)\right). \tag{2.20}$$

Cel sterowania polega na wyznaczeniu takiego prawa sterowania $u^* = \left(u_0^*, ..., u_{N-1}^*\right)$, dla którego łączne koszty (nakłady energetyczne wraz z końcowymi stratami) są najmniejsze. Rozwiązując zadanie (2.3) wyznaczymy optymalne sterowanie systemem (2.19).

Uwaga 2.8 *Dla zadania (2.3) ze wskaźnikiem jakości (2.20) system liniowy jest przeprowadzany ze stanu początkowego y_0 do celu a, natomiast dokonu-*

jąc przesunięcia układu współrzędnych o wektor a, przeprowadzamy system w nowym układzie ze stanu $y_0 - a$ do początku układu współrzędnych.

2.2.2 Sterowanie systemem liniowym z pełną informacją

Najpierw wyznaczymy optymalne sterowanie dla liniowego systemu stochastycznego (2.19) z pełną informacją. W tym przypadku przyjmujemy, że wartości parametrów systemu ξ są znane sterującemu. Wyznaczając optymalne sterowanie w momentach $j = 0, 1, ..., N-1$ wykorzystujemy znajomość zarówno bieżącego, jak i poprzednich stanów systemu $y_0, ..., y_j$ oraz wiedzę o parametrach systemu ξ, tzn.

$$u_j \overset{def}{=} u_j\left(\xi, y_0, ..., y_j\right).$$

Aby rozwiązać zadanie (2.3) najpierw konstruujemy funkcje Bellmana (funkcje ze strukturą teleskopową). Zgodnie z uwagą 2.2 dla dowolnego momentu $j \in \{0, ..., N-1\}$ funkcje Bellmana definiujemy wzorem

$$V_j^N\left(\xi, y_0, ..., y_j\right) = \inf_{u_j} E\left(u_j^T R_j u_j + V_{j+1}^N\left(\xi, y_{j+1}\right) \middle| \mathcal{F}_j^{\xi}\right) \tag{2.21}$$

z warunkiem początkowym

$$V_N^N\left(\xi, y_0, ..., y_N\right) = \left(y_N - a\right)^T Q_N \left(y_N - a\right). \tag{2.22}$$

Stosownie do powyższych oznaczeń najmniejsza wartość wskaźnika jakości jest równa wartości funkcji Bellmana dla momentu $j = 0$, tzn.

$$\inf_{u \in U} J^N\left(u, \xi\right) = V_0^N\left(\xi, y_0\right). \tag{2.23}$$

Sposób wyznaczenia ciągu optymalnych sterowań $\left\{u_j^*\right\}_{0 \le j \le N-1}$ dla systemów (2.19) ze znanymi parametrami $\xi \in R^k$ podaje twierdzenie poniżej.

Twierdzenie 2.2 *Niech macierze $R_i \in \mathbb{R}^{l \times l}$ dla $0 \le i \le N-1$ oraz $Q_j^N \in \mathbb{R}^{n \times n}$, $0 \le j \le N$ są dodatnio określone oraz ciąg macierzy $\left\{Q_j^N\right\}_{0 \le j \le N-1}$*

spełnia równanie

$$Q_j^N = Q_{j+1}^N - \left(Q_{j+1}^N\right)^T A^{N-j-1}C \times \left(\bar{R}_j + \left(A^{N-j-1}C\right)^T \bar{Q}_{j+1}^N A^{N-j-1}C\right)^{-1} \left(A^{N-j-1}C\right)^T Q_{j+1}^N, \quad (2.24)$$

gdzie

$$\bar{R}_j = \frac{1}{2}\left(R_j + R_j^T\right), \quad \bar{Q}_j^N = \frac{1}{2}\left(Q_j^N + \left(Q_j^N\right)^T\right) \quad (2.25)$$

z warunkiem początkowym $Q_N^N = Q_N$. *Optymalnym rozwiązaniem zadania (2.23) dla systemu liniowego (2.19) w warunkach pełnej informacji jest:*

1. *optymalne sterowanie* u_j^* *systemem (2.19) w momentach* $0 \leq j \leq N-1$ *jest równe*

$$u_j^* = \left(\bar{R}_j + \left(A^{N-j-1}C\right)^T \bar{Q}_{j+1}^N A^{N-j-1}C\right)^{-1} \times \left(A^{N-j-1}C\right)^T Q_{j+1}^N \left(A^{N-j}y_j + \psi_{N-j}(\xi) - a\right); \quad (2.26)$$

2. *wartość funkcji Bellmana dla momentu* $j = 0, ..., N-1$ *jest równa*

$$V_j^N(\xi, y_0, ..., y_j) = \left(A^{N-j}y_j + \psi_{N-j}(\xi) - a\right)^T Q_j^N \times \left(A^{N-j}y_j + \psi_{N-j}(\xi) - a\right) + \varphi_j^N, \quad (2.27)$$

gdzie funkcje φ_j^N, $\psi_j(\cdot)$ *są dane wzorami*

$$\varphi_j^N = \varphi_{j+1}^N + tr\left(\left(A^{N-j-1}\sigma\right)^T Q_{j+1}^N A^{N-j-1}\sigma\right) \quad (2.28)$$

z warunkiem początkowym $\varphi_N^N = 0$ *oraz*

$$\psi_j(\xi) = \Lambda_j B\xi, \quad \Lambda_j = I + A + ... + A^{\max(0,j-1)}, \quad (2.29)$$

natomiast $I \in \mathbb{R}^{n\times n}$ *jest macierzą jednostkową.*

Dowód. Teza udowodniona zostanie za pomocą indukcji matematycznej. Najpierw wyznaczymy wartość funkcji Bellmana dla momentu $N-1$. Dla systemu (2.19) ze wzorów (2.21)–(2.22) otrzymujemy

$$\begin{aligned} V_{N-1}^N(\xi, y_0, ..., y_{N-1}) = \inf_{u_{N-1}} E \big(& u_{N-1}^T R_{N-1} u_{N-1}^T \\ & + (Ay_{N-1} + B\xi - Cu_{N-1} + \sigma w_N - a) Q_N^N \\ & \times (Ay_{N-1} + B\xi - Cu_{N-1} + \sigma w_N - a) \Big| \mathcal{F}_{N-1}^{\xi} \Big) . \end{aligned}$$

Z własności warunkowej wartości oczekiwanej mamy

$$E\left(w_N^T \sigma^T Q_N \sigma w_N \Big| \mathcal{F}_{N-1}^{\xi} \right) = tr\left(\sigma^T Q_N^N \sigma\right) = \varphi_{N-1}^N,$$

zatem

$$\begin{aligned} V_{N-1}^N(\xi, y_0, ..., y_{N-1}) = \inf_{u_{N-1}} \big(& u_{N-1}^T \left(R_{N-1} + C^T Q_N^N C\right) u_{N-1} \\ & - 2u_{N-1}^T C^T Q_N^N (Ay_{N-1} + B\xi - a) + \varphi_{N-1}^N \\ & + (Ay_{N-1} + B\xi - a)^T Q_N^N (Ay_{N-1} + B\xi - a) \Big) . \end{aligned} \tag{2.30}$$

Optymalne sterowanie dla momentu $N-1$ jest równe

$$u_{N-1}^* = \left(\bar{R}_{N-1} + C^T \bar{Q}_N^N C\right)^{-1} C^T Q_N^N (Ay_{N-1} + \psi_1(\xi) - a) . \tag{2.31}$$

Podstawiając optymalne sterowanie (2.31) do wzoru (2.30) otrzymujemy, że wartość funkcji Bellmana w chwili $N-1$ wynosi

$$\begin{aligned} V_{N-1}^N(\xi, y_0, ..., y_{N-1}) = & (Ay_{N-1} + \psi_1(\xi) - a)^T Q_{N-1}^N \\ & \times (Ay_{N-1} + \psi_1(\xi) - a) + \varphi_{N-1}^N, \end{aligned} \tag{2.32}$$

gdzie $Q_{N-1}^N = Q_N^N - \left(Q_N^N\right)^T C \left(\bar{R}_{N-1} + C^T \bar{Q}_N^N C\right)^{-1} C^T Q_N^N$.
Zakładamy, że wzory (2.26)–(2.27) są prawdziwe dla dowolnego momentu $j = 1, 2, ..., N-1$ oraz sprawdzamy je dla momentu $j-1$. Ze wzoru (2.21)

wartość funkcji Bellmana w chwili $j-1$ jest równa

$$V_{j-1}^N(\xi, y_0, ..., y_{j-1}) = \inf_{u_{j-1}} E\left(u_{j-1}^T R_{j-1} u_{j-1} + \varphi_j^N\right.$$
$$\left. + \left(A^{N-j} y_j + \psi_{N-j}(\xi) - a\right)^T Q_j^N \left(A^{N-j} y_j + \psi_{N-j}(\xi) - a\right) \middle| \mathcal{F}_{j-1}^\xi\right).$$

Z równania (2.19) otrzymujemy

$$\begin{aligned} V_{j-1}^N(\xi, y_0, ..., y_{j-1}) &= \inf_{u_{j-1}} E\left\{u_{j-1}^T R_{j-1} u_{j-1} + \varphi_j^N\right. \\ &+ \left(A^{N-j}(Ay_{j-1} + B\xi - Cu_{j-1} + \sigma w_j) + \psi_{N-j}(\xi) - a\right)^T Q_j^N \\ &\times \left.\left(A^{N-j}(Ay_{j-1} + B\xi - Cu_{j-1} + \sigma w_j) + \psi_{N-j}(\xi) - a\right) \middle| \mathcal{F}_{j-1}^\xi\right) \\ &= \inf_{u_{j-1}} \left(u_{j-1}^T \left(R_{j-1} + \left(A^{N-j}C\right)^T Q_j^N A^{N-j} C\right) u_{j-1} + tr\left(\left(A^{N-j}\sigma\right)^T Q_j^N A^{N-j}\sigma\right)\right. \\ &+ \varphi_j^N - 2u_{j-1}^T \left(A^{N-j}C\right)^T Q_j^N \left(A^{N-j+1} y_{j-1} + \psi_{N-j+1}(\xi) - a\right) \\ &+ \left.\left(A^{N-j+1} y_{j-1} + \psi_{j-1}(\xi) - a\right)^T Q_j^N \left(A^{N-j+1} y_{j-1} + \psi_{N-j+1}(\xi) - a\right)\right). \end{aligned} \tag{2.33}$$

Optymalne sterowanie dla momentu $j-1$ jest równe

$$\begin{aligned} u_{j-1}^* &= \left(\bar{R}_{j-1} + \left(A^{N-j}C\right)^T \bar{Q}_j^N A^{N-j} C\right)^{-1} \\ &\times \left(A^{N-j}C\right)^T Q_j^N \left(A^{N-j+1} y_{j-1} + \psi_{N-j+1}(\xi) - a\right). \end{aligned} \tag{2.34}$$

Podstawiając optymalne sterowanie (2.34) do (2.33) otrzymujemy

$$\begin{aligned} V_{j-1}(\xi, y_0, ..., y_{j-1}) &= \left(A^{N-j+1} y_{j-1} + \psi_{N-j+1}(\xi) - a\right)^T Q_{j-1}^N \\ &\times \left(A^{N-j+1} y_{j-1} + \psi_{N-j+1}(\xi) - a\right) + \varphi_{j-1}^N, \end{aligned} \tag{2.35}$$

gdzie Q_{j-1}^N i φ_{j-1}^N są zdefiniowane za pomocą wzorów (2.24) i (2.28). ∎

Uwaga 2.9 *Z powyższego twierdzenia wynika, że dla sterowania liniowo–kwadratowego zarówno optymalne sterowanie w chwili* $0 \le j \le N-1$, *jak i funkcja Bellmana (2.21) zależą od wartości parametrów* ξ *i stanu* y_j *(nie zależą od stanów poprzednich* $y_0, ..., y_{j-1}$*), tzn.* $V_j^N(\xi, y_0, ..., y_j) = V_j^N(\xi, y_j)$.

Uwaga 2.10 *Jeżeli $R_i = R$ dla $0 \le i \le N-1$, to ze wzorów (2.24) i (2.28) wynikają następujące równości*

$$Q_j^N = Q_{j+s}^{N+s} \quad \text{oraz} \quad \varphi_j^N = \varphi_{j+s}^{N+s}$$

dla $0 \le j \le N$, $j, s, N \in \mathbb{N}$. Zgodnie z uwagą powyżej oraz korzystając ze wzoru (2.27) otrzymujemy

$$V_j^N(\xi, y) = V_0^{N-j}(\xi, y)$$

dla dowolnego $\xi \in \mathbb{R}^k$ i $y \in \mathbb{R}^n$ oraz $0 \le j \le N$, $j, N \in \mathbb{N}$. Zatem wartość funkcji Bellmana $V_j^N(\xi, y)$ (suma oczekiwanych kosztów sterowania systemem (2.19) i dziedziczenia) zależy od czasu pozostałego do końca horyzontu sterowania $N-j$.

2.2.3 Sterowanie systemem liniowym z niepełną informacją

W przypadku niepełnej informacji o parametrach systemu stochastycznego zakładamy, że sterujący zna dynamikę zachowania liniowego systemu (2.19) (znane jest równanie stanu) oraz nie zna dokładnych wartości parametrów ξ, natomiast posiada wiedzę odnośnie rozkładu a priori wektora losowego ξ. W związku z powyższym sterujący posiada tylko częściową wiedzę o działaniu (funkcjonowaniu, zachowaniu) systemu stochastycznego (2.19). Poniżej zakładamy, że wektor losowy $\xi : \Omega \to \mathbb{R}^k$ ma rozkład normalny $\boldsymbol{N}(m_0, \Gamma_0)$. Podejmując sterowanie w chwili $0 \le j \le N-1$ wykorzystujemy znajomość (zaobserwowane wartości) bieżącego i poprzednich stanów systemu $y_0, ..., y_j$ oraz sterowań $u_0, ..., u_{j-1}$, zatem przyjmujemy

$$\bar{u}_j \stackrel{def}{=} \bar{u}_j(y_0, ..., y_j, u_0, ..., u_{j-1}).$$

Do rozwiązania zadania (2.3) w warunkach niepełnej informacji o systemie (2.19) wykorzystane zostaną funkcje Bellmanna $V_j^N(\xi, y_0, .., y_j)$, $0 \le j \le N$ dane wzrorami (2.21)-(2.22). Dla dowolnego $j \in \{0, ..., N-1\}$

definiujemy

$$\bar{V}_j^N(y_0,..,y_j) = \inf_{\bar{u}_j} E\left(\bar{u}_j^T R_j \bar{u}_j + V_{j+1}^N(\xi, y_{j+1})\,|\mathcal{F}_j\right) \tag{2.36}$$

oraz

$$\bar{V}_N^N(y_0,..,y_N) = (y_N - a)^T Q_N (y_N - a)\,. \tag{2.37}$$

Dla momentu $j \in \{0,1,...,N-1\}$ wielkość $\bar{V}_j^N(y_0,..,y_j)$ reprezentuje najmniejszy oczekiwany koszt sterowania systemem (2.19) w warunkach niepełnej informacji od momentu j do końca horyzontu czasowego N. Zatem najmniejsza wartość wskaźnika jakości (2.20) jest równa

$$\inf_{\bar{u}\in U} J^N(\bar{u}) = \bar{V}_j^N(y_0)\,.$$

Sposób wyznaczenia ciągu optymalnych sterowań $\left\{\bar{u}_j^*\right\}_{0\le j\le N-1}$ w warunkach niepełnej informacji o liniowym systemie stochastycznym (2.19) przedstawiony zostanie w twierdzeniu. Do dowodu twierdzenia wykorzystane zostaną lematy przedstawione poniżej.

Lemat 2.1 *Niech $\xi : \Omega \to \mathbb{R}^k$ będzie wektorem losowym, $\Psi \in R^{k\times k}$ będzie macierzą deterministyczną, $\{\mathcal{F}_j\}_{j\ge 0}$ określa niemalejącą rodzinę $\sigma-$ciał na zupełnej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ oraz Γ_j oznacza drugi moment centralny wektora losowego ξ warunkowany na $\sigma-$ciele $\mathcal{F}_j$*

$$\Gamma_j = E\left[\left[\xi - E(\xi|\,\mathcal{F}_j)\right]\left[\xi - E(\xi|\,\mathcal{F}_j)\right]^T\Big|\,\mathcal{F}_j\right].$$

Dla dowolnego $0 \le i < j$ mamy

$$\begin{aligned} E\left[\xi^T \Psi E(\xi|\,\mathcal{F}_j)\big|\,\mathcal{F}_i\right] &= E\left[E(\xi|\,\mathcal{F}_j)^T \Psi E(\xi|\,\mathcal{F}_j)\Big|\,\mathcal{F}_i\right] = \\ &= E(\xi|\,\mathcal{F}_i)^T \Psi E(\xi|\,\mathcal{F}_i) + tr\left[\Psi(\Gamma_i - \Gamma_j)\right] \end{aligned} \tag{2.38}$$

oraz

$$E\left[\xi^T \Psi \xi\big|\,\mathcal{F}_i\right] = E(\xi|\,\mathcal{F}_i)^T \Psi E(\xi|\,\mathcal{F}_i) + tr(\Psi\Gamma_i)\,. \tag{2.39}$$

Dowód. Dla dowolnego $0 \leq i < j$ korzystając z własności warunkowej wartości oczekiwanej otrzymujemy

$$\begin{aligned}
&E\left[\xi^T \Psi E\left(\xi | \mathcal{F}_j\right) \middle| \mathcal{F}_i\right] = E\left[E\left[\xi^T \Psi E\left(\xi | \mathcal{F}_j\right) \middle| \mathcal{F}_j\right] \middle| \mathcal{F}_i\right] \\
&\quad = E\left[E\left(\xi | \mathcal{F}_j\right)^T \Psi E\left(\xi | \mathcal{F}_j\right) \middle| \mathcal{F}_i\right] = tr\left(E\left[\Psi E\left(\xi | \mathcal{F}_j\right) E\left(\xi | \mathcal{F}_j\right)^T \middle| \mathcal{F}_i\right]\right) \\
&\quad = tr\left\{\Psi E\left[E\left(\xi\xi^T \middle| \mathcal{F}_j\right) - \Gamma_j \middle| \mathcal{F}_i\right]\right\} = tr\left\{\Psi E\left(\xi\xi^T \middle| \mathcal{F}_i\right)\right\} - tr\left\{\Psi \Gamma_j\right\} \\
&\quad = tr\left(\Psi\left[E\left(\xi | \mathcal{F}_i\right) E\left(\xi | \mathcal{F}_i\right)^T + \Gamma_i\right]\right) - tr\left\{\Psi \Gamma_j\right\} \\
&\quad = E\left(\xi | \mathcal{F}_i\right)^T \Psi E\left(\xi | \mathcal{F}_i\right) + tr\left[\Psi\left(\Gamma_i - \Gamma_j\right)\right].
\end{aligned}$$

Ponadto mamy

$$\begin{aligned}
&E\left[\xi^T \Psi \xi \middle| \mathcal{F}_i\right] = tr\left\{\Psi E\left(\xi\xi^T \middle| \mathcal{F}_i\right)\right\} \\
&\quad = tr\left(\Psi\left[E\left(\xi | \mathcal{F}_i\right) E\left(\xi | \mathcal{F}_i\right)^T + \Gamma_i\right]\right) = E\left(\xi | \mathcal{F}_i\right)^T \Psi E\left(\xi | \mathcal{F}_i\right) + tr\left(\Psi \Gamma_i\right).
\end{aligned}$$

■

Dla systemu liniowego (2.19) w dowolnym momencie $j \in \{0, 1, ..., N-1\}$ wyznaczamy warunkowy rozkład $P(d\xi \,|\mathcal{F}_j)$ na podstawie informacji dotyczącej stanów $y_0, ..., y_j$ i sterowań $u_0, ..., u_{j-1}$. Wyznaczenie optymalnego estymatora (filtra Kalmana–Bucy) wektora losowego ξ polega na oszacowaniu wektora warunkowych wartości oczekiwanych $m_j = E(\xi \,|\mathcal{F}_j)$ oraz wyznaczeniu macierzy warunkowej kowariancji (macierzy błędu estymacji) $\Gamma_j = E\left((\xi - m_j)(\xi - m_j)^T |\mathcal{F}_j\right)$. Zagadnienie filtracji ciągów warunkowo normalnych przedstawione zostało w poprzednim rozdziale. Więcej na temat estymacji parametrycznej dla systemów liniowych można znaleć w pracach [24], [79], [98], [99], [101]. Przykłady konstrukcji optymalnych filtrów dla systemów liniowych z nieznanymi parametrami zostały omówione w [17], [18], [20]. Lemat poniżej przedstawia własności warunkowego rozkładu wektora nieznanych parametrów ξ dla systemu liniowego (2.19).

Lemat 2.2 *Zakładamy, że liniowy system stochastyczny jest określony za pomocą równania (2.19), rozkład a priori (ze względu na σ–ciało $\mathcal{F}_0$) wektora losowego ξ jest rozkładem normalnym $\boldsymbol{N}(m_0, \Gamma_0)$, z prawdopodobieństwem* *jeden elementy wektora* $Ay_j - Cu_j$*,* $j \in \mathbb{N}_0$ *i elementy macierzy* B*,* σ *są ograniczone, tzn.*

$$P\left(\left|(Ay_j - Cu_j)_i\right| \leq c\right) = 1,\ P\left(|B_{il}| \leq c\right) = 1,$$
$$P\left(|\sigma_{is}| \leq c\right) = 1$$

dla $1 \leq i \leq n$*,* $1 \leq l \leq k$*,* $1 \leq s \leq m$*,* $c < \infty$ *oraz*

$$E\left|Ay_j - Cu_j\right|^2 < \infty,\ \ |B|^2 < \infty,\ \ |\sigma|^2 < \infty.$$

Dla dowolnego $j \in \mathbb{N}_0$ *warunkowy rozkład wektora losowego* ξ *względem* σ*–ciała* $\mathcal{F}_j$ *jest rozkładem normalnym* $\boldsymbol{N}(m_j, \Gamma_j)$*, gdzie warunkowa wartość oczekiwana wektora losowego* m_j *oraz macierz warunkowej kowariancji* Γ_j *są dane wzorami*

$$m_j = \left(I + j\Gamma_0 B^T\left(\sigma\sigma^T\right)^{-1} B\right)^{-1} \times \left(m + \Gamma_0 \sum_{i=0}^{j-1} B^T\left(\sigma\sigma^T\right)^{-1}\left[y_{i+1} - Ay_i + Cu_i\right]\right) \tag{2.40}$$

oraz

$$\Gamma_j = \left(I + j\Gamma_0 B^T\left(\sigma\sigma^T\right)^{-1} B\right)^{-1}\Gamma_0. \tag{2.41}$$

Teza lematu wynika bezpośrednio z twierdzenia 1.14.

Twierdzenie 2.3 *Niech liniowy system stochastyczny będzie określony równaniem (2.19), wektor losowy* $\xi : \Omega \to \mathbb{R}^k$ *ma rozkład normalny* $\boldsymbol{N}(m_0, \Gamma_0)$ *względem* σ*–ciała* $\mathcal{F}_0$*, ciąg macierzy* $\left\{Q_j^N\right\}_{0 \leq j \leq N}$ *spełnia równanie (2.24) oraz macierze* $R_i \in R^{l \times l}$*,* $0 \leq i \leq N-1$ *oraz* $Q_j^N \in R^{n \times n}$*,* $0 \leq j \leq N$ *są dodatnio określone. Rozwiązaniem zadania (2.3) w przypadku niepełnej informacji o systemie jest:*

1. *dla momentów* $0 \leq j \leq N-1$ *optymalne sterowanie systemem liniowym (2.19) jest równe*

$$\bar{u}_j^* = \left(\bar{R}_j + \left(A^{N-j+1}C\right)^T \bar{Q}_{j+1}^N A^{N-j+1}C\right)^{-1}$$
$$\times \left(A^{N-j+1}C\right)^T Q_{j+1}^N \left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right), \quad (2.42)$$

gdzie funkcjonał $\psi_j\left(\right)$ *jest dany wzorem (2.29), estymator wektora losowego* ξ *wzorem (2.40), natomiast macierze* $\bar{R}_j$ *i* $\bar{Q}_j^N$ *są określone za pomocą (2.25);*

2. *oczekiwany koszt sterowań od momentu* $j = 0, ..., N-1$ *do końca horyzontu* N *wynosi*

$$\bar{V}_j^N\left(y_0, ..., y_j\right) = \left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right)^T Q_j^N$$
$$\times \left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right) + \varphi_j^N + \Psi_j^N, \quad (2.43)$$

gdzie

$$\Psi_j^N = tr\left(\left(\Lambda_{N-j}B\right)^T Q_{j+1}^N \Lambda_{N-j}B\Gamma_j\right), \quad (2.44)$$

natomiast dla $0 \leq j \leq N-1$ *wartość* φ_j^N *jest dana wzorem (2.28), macierz* I *jest macierzą jednostkową,* $\Lambda_j = I + A + ... + A^{\max(0,j-1)}$ *oraz macierze warunkowej kowariancji* Γ_j *dane są wzorami (2.41).*

Dowód. Dla dowolnego momentu $j \in \{0, 1, 2, ..., N-1\}$ korzystając ze wzoru (2.36) oraz funkcji Bellmana $V_j^N\left(\xi, y_0, ..., y_j\right)$ określonej za pomocą równań (2.21)-(2.22) szukamy optymalnego sterowania $\bar{u}_j^*$. W warunkach niepełnej informacji o systemie wartość funkcji Bellmana jest równa

$$\bar{V}_j^N\left(y_0, ..., y_j\right) = \inf_{\bar{u}_j} E\left(\bar{u}_j^T R_j \bar{u}_j + \left(A^{N-j-1}y_{j+1} + \psi_{N-j-1}\left(\xi\right) - a\right)^T Q_{j+1}^N\right.$$
$$\left.\times \left(A^{N-j-1}y_{j+1} + \psi_{N-j-1}\left(\xi\right) - a\right) + \varphi_{j+1}^N \right| \mathcal{F}_j\Big).$$

Z własności warunkowej wartości oczekiwanej otrzymujemy

$$\bar{V}_j^N\left(y_0, ..., y_j\right) = \inf_{\bar{u}_j} u_j^T \left(R_j + \left(A^{N-j-1}C\right)^T Q_{j+1}^N A^{N-j-1}C\right) u_j$$

$$
\begin{aligned}
&+ E\left(\left(A^{N-j}y_j + A^{N-j-1}B\xi + \psi_{N-j-1}(\xi) - a\right)^T Q_{j+1}^N\right.\\
&\times \left(A^{N-j}y_j + A^{N-j-1}B\xi + \psi_{N-j-1}(\xi) - a\right) - 2u_j^T\left(A^{N-j-1}C\right)^T\\
&\times Q_{j+1}^N\left(A^{N-j}y_{j-1} + A^{N-j-1}B\xi + \psi_{N-j-1}(\xi) - a\right)|\mathcal{F}_j)\\
&+ tr\left(\left(A^{N-j-1}\sigma\right)^T Q_{j+1}^N A^{N-j-1}\sigma\right) + \varphi_{j+1}^N. \qquad (2.45)
\end{aligned}
$$

Ze wzoru (2.29) mamy

$$\psi_{N-j}(\xi) = A^{N-j-1}B\xi + \psi_{N-j-1}(\xi).$$

Zatem optymalne sterowanie w chwili j jest równe

$$
\begin{aligned}
\bar{u}_j^* = &\left(\bar{R}_j + \left(A^{N-j-1}C\right)^T \bar{Q}_{j+1}^N A^{N-j-1}C\right)^{-1}\left(A^{N-j-1}C\right)^T\\
&\times Q_{j+1}^N\left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right). \qquad (2.46)
\end{aligned}
$$

Z lematu 2.1 wynika, że

$$
\begin{aligned}
&E\left(\left(A^{N-j}y_{j-1} + \psi_{N-j}(\xi) - a\right)^T Q_{j+1}^N\left(A^{N-j}y_{j-1} + \psi_{N-j}(\xi) - a\right)|\mathcal{F}_j\right)\\
&= \left(A^{N-j}y_{j-1} + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right)^T Q_{j+1}^N\\
&\quad \times \left(A^{N-j}y_{j-1} + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right)\\
&\quad + tr\left(\left(\Lambda_{N-j}B\right)^T Q_{j+1}^N \Lambda_{N-j}B\Gamma_j\right). \qquad (2.47)
\end{aligned}
$$

Ostatecznie, podstawiając (2.46) i (2.47) do (2.45) otrzymujemy

$$
\begin{aligned}
\bar{V}_j^N(y_0, ..., y_j) = &\left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right)^T Q_j^N\\
&\times\left(A^{N-j}y_j + \psi_{N-j}\left(E\left(\xi\,|\mathcal{F}_j\right)\right) - a\right) + \varphi_j^N + \Psi_j^N,
\end{aligned}
$$

gdzie Q_j^N, Ψ_j^N, φ_j^N spełnia równania (2.24), (2.44), (2.28) odpowiednio. ■

Uwaga 2.11 *Jeżeli* $R_i = R$ *dla* $0 \leq i \leq N-1$, *to ze wzorów (2.24) oraz (2.28) wynika, że*

$$Q_j^N = Q_{j+s}^{N+s}, \quad \textit{oraz} \quad \varphi_j^N = \varphi_{j+s}^{N+s},$$

natomiast ze wzorów (2.41) i (2.44)

$$\Psi_j^N \neq \Psi_{j+s}^{N+s}$$

dla $0 \leq j \leq N$, $j, s, N \in \mathbb{N}$. *Przyjmijmy* $\bar{V}_j^N(y_j) = \bar{V}_j^N(y_0, ..., y_j)$ *dla* $0 \leq j \leq N$. *Wobec powyższego dla dowolnego* $y \in \mathbb{R}^n$

$$\bar{V}_j^N(y) \neq \bar{V}_{j+s}^{N+s}(y)$$

oraz

$$\bar{V}_j^N(y) \neq \bar{V}_0^{N-j}(y)$$

dla $0 \leq j \leq N$, $j, s, N \in \mathbb{N}$. *Zatem, sterując systemem (2.19) z niepełną informacją, funkcja* $\bar{V}_j^N(y_0, ..., y_j)$ *reprezentująca oczekiwane koszty sterowania i dziedziczenia jest dana wzorem (2.43) oraz zależy od momentu, w którym znajduje się system (porównaj z uwagą 2.10).*

2.2.4 Wpływ informacji o systemie liniowym na koszty sterowania

W tej części rozdziału przedstawione zostaną różnice i podobieństwa sterowań stochastycznym systemem liniowym (2.19) w warunkach pełnej i niepełnej informacji. Również oszacowane zostaną koszty, które należy ponieść dodatkowo w przypadku nieznajomości parametrów systemu. Ze wzorów (2.26) i (2.42) wynika, że dla dowolnego $j \in \{0, 1, ..., N-1\}$ w warunkach pełnej informacji optymalne sterowanie systemem (2.19) jest równe

$$u_j^* = H_j\left(A^{N-j}y_j + \Lambda_j B\xi - a\right), \tag{2.48}$$

natomiast w warunkach niepełnej informacji

$$\bar{u}_j^* = H_j\left(A^{N-j}y_j + \Lambda_j B m_j - a\right), \tag{2.49}$$

gdzie

$$H_j = \left(\bar{R}_j + \left(A^{N-j+1}C\right)^T \bar{Q}_{j+1}^N A^{N-j+1}C\right)^{-1} \left(A^{N-j+1}C\right)^T Q_{j+1}^N,$$
$$\Lambda_j = I + A + ... + A^{j-1}$$

oraz $m_j = E\left(\xi \left| \mathcal{F}_j \right.\right)$ jest najlepszym estymatorem w sensie średnio-kwadratowym wektora losowego ξ, który wyznaczamy za pomocą (2.40), natomiast ciąg $\left\{Q_j^N\right\}_{0 \leq j \leq N}$ spełnia równanie (2.24).

Uwaga 2.12 *Do wyznaczenia optymalnego sterowania liniowo–kwadratowego w chwili $j \in \{0, 1, ..., N-1\}$ w przypadku pełnej informacji o systemie wykorzystujemy znajomość bieżącego stanu y_j i wartości parametrów ξ, natomiast w przypadku niepełnej informacji wartości $y_0, ..., y_j$. Ze wzorów (2.48) oraz (2.49) (zgodnie z oznaczeniami (2.17)-(2.18)) wynika*

$$\bar{u}_j\left(y_0, ..., y_j\right) = u_j\left(E\left(\xi \left| \mathcal{F}_j \right.\right), y_0, ..., y_j\right),$$

zatem dla systemu (2.19) ze wskaźnikiem jakości (2.20) konstrukcja optymalnych sterowań w warunkach pełnej i niepełnej informacji jest podobna. W przypadku niepełnej informacji wystarczy w miejsce nieznanego parametru podstawić jego estymator.

Porównując koszty sterowania $\bar{V}_j^N\left(y_0, ..., y_j\right)$ i $V_j^N\left(\xi, y_0, ..., y_j\right)$ zdefiniowane wzorami (2.27) i (2.43) widzimy, że nieznajomość parametrów ξ systemu (2.19) (parametrów odpowiadających za przesunięcie) generuje dodatkowe koszty. Dla dowolnego $j = 0, 1, ..., N-1$ oraz $y \in \mathbb{R}^n$ różnica oczekiwanych kosztów sterowania od momentu j do końca horyzontu N jest równa

$$\bar{V}_j^N\left(y_0, ..., y_j\right) - V_j^N\left(E\left(\xi \left| \mathcal{F}_j \right.\right), y_0, ..., y_j\right) = \Psi_j^N,$$

gdzie Ψ_j^N jest dane wzorem (2.44). Wielkość Ψ_j^N oznacza wartość informacji o parametrach ξ. Jeżeli w chwili $j = 0$ przyjmiemy, że $\xi = m_0$ oraz system znajduje się w punkcie y_0, to wartość informacji o nieznanych parametrach ξ przy przeprowadzeniu systemu (2.19) w N krokach ze stanu y_0 do punktu

a jest równa

$$\Psi_0^N = tr\left(B^T\left(I + A + ... + A^{N-1}\right)^T Q_1^N\left(I + A + ... + A^{N-1}\right)B\Gamma_0\right).$$

Wielkość Ψ_0^N oznacza dodatkowy oczekiwany koszt, który zostanie poniesiony przez sterującego podczas realizacji celu w N krokach.

Zarówno dla przypadków pełnej i niepełnej informacji o parametrach systemu wielkość

$$\varphi_j^N = \sum_{i=j+1}^{N} tr\left(\left(A^{N-i}\sigma\right)^T Q_i^N A^{N-i}\sigma\right)$$

oznacza dodatkowy oczekiwany koszt spowodowany zaburzeniami zewnętrznymi od momentu j do końca horyzontu N. Wielkość φ_0^N reprezentuje oczekiwany koszt spowodowany ciągiem zaburzeń zewnętrznych $\{w_i\}_{1\leq i\leq N}$ podczas sterowania systemem (2.19).

Uwaga 2.13 *Dla liniowego systemu stochastycznego (2.19) zarówno Ψ_j^N (wartość informacji o parametrze ξ), jak i φ_j^N, $0 \leq j \leq N-1$ (koszt spowodowany zaburzeniami zewnętrznymi) zależą od horyzontu czasowego N.*

Dla dowolnego $j \in \{0, 1, ..., N-1\}$ definiujemy

$$\tilde{V}_j^N\left(y_0, ..., y_j\right) = \inf_{\tilde{u}_j} E\left(\bar{u}_j^T R_j \bar{u}_j + \bar{V}_{j+1}^N\left(y_0, ..., y_{j+1}\right)|\mathcal{F}_j\right) \tag{2.50}$$

(porównaj z (2.36)). Z własności warunkowej wartości oczekiwanej mamy

$$\begin{aligned}
&\tilde{V}_j^N\left(y_0, ..., y_j\right) = \inf_{\bar{u}_j} u_j^T\left(R_j + \left(A^{N-j-1}C\right)^T Q_{j+1}^N A^{N-j-1}C\right)u_j \\
&+ E\left(\left(A^{N-j}y_j + A^{N-j-1}B\xi + \psi_{N-j-1}\left(E\left(\xi\,|\mathcal{F}_{j+1}\right)\right) - a\right)^T\right. \\
&\times Q_{j+1}^N\left(A^{N-j}y_j + A^{N-j-1}B\xi + \psi_{N-j-1}\left(E\left(\xi\,|\mathcal{F}_{j+1}\right)\right) - a\right) - 2u_j^T\left(A^{N-j-1}C\right)^T \\
&\left.\times Q_{j+1}^N\left(A^{N-j}y_{j-1} + A^{N-j-1}B\xi + \psi_{N-j-1}\left(E\left(\xi\,|\mathcal{F}_{j+1}\right)\right) - a\right)|\mathcal{F}_j\right) \\
&+ tr\left(\left(A^{N-j-1}\sigma\right)^T Q_{j+1}^N A^{N-j-1}\sigma\right) + \varphi_{j+1}^N + \Psi_{j+1}^N.
\end{aligned}$$

Zatem optymalne sterowanie w chwili j jest dane wzorem (2.42), natomiast z lematu 1 otrzymujemy

$$\tilde{V}_j^N (y_0, ..., y_j) = V_j^N (E(\xi | \mathcal{F}_j), y_j) + \Psi_{j+1}^N + \Upsilon_j^N,$$

gdzie

$$\Upsilon_j^N = tr\left(B^T \left(2A^{N-j-1} + \Lambda_{N-j-1}\right)^T Q_{j+1}^N \Lambda_{N-j-1} B (\Gamma_j - \Gamma_{j+1})\right) \\ + tr\left(\left(A^{N-j-1}B\right)^T Q_{j+1}^N A^{N-j-1} B\Gamma_j\right),$$

natomiast Ψ_j^N jest dane wzorem (2.44). Wielkość

$$\tilde{V}_j^N (y_0, ..., y_j) - \bar{V}_j^N (y_0, ..., y_j) = \Upsilon_j^N + \Psi_{j+1}^N - \Psi_j^N$$

oznacza koszt uczenia się na przedziale czasowym $(j, j+1]$.

Przykład 2.1 *Rozważamy problem sterowania liniowym systemem stochastycznym o równaniu stanu (2.19), gdzie*

$$A = \begin{bmatrix} 0.8 & 0.1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1.2 & 0.5 \\ 0.05 & 0.8 \end{bmatrix},$$
$$C = \begin{bmatrix} 1.5 & -0.1 \\ 0.04 & 1.1 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 0.35 & 0.03 \\ -0.02 & 0.4 \end{bmatrix}.$$

Zadanie polega na przeprowadzeniu systemu przy najmniejszych kosztach (suma nakładów energetycznych wraz ze stratami końcowymi muszą być najmniejsze) ze stanu $y_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ *do stanu* $a = \begin{bmatrix} 60 \\ 35 \end{bmatrix}$ *np. w 15 krokach* $(N = 15)$*. Macierze odpowiadające za koszty sterowania i dziedziczenia dla wskaźnika jakości (2.20) są równe*

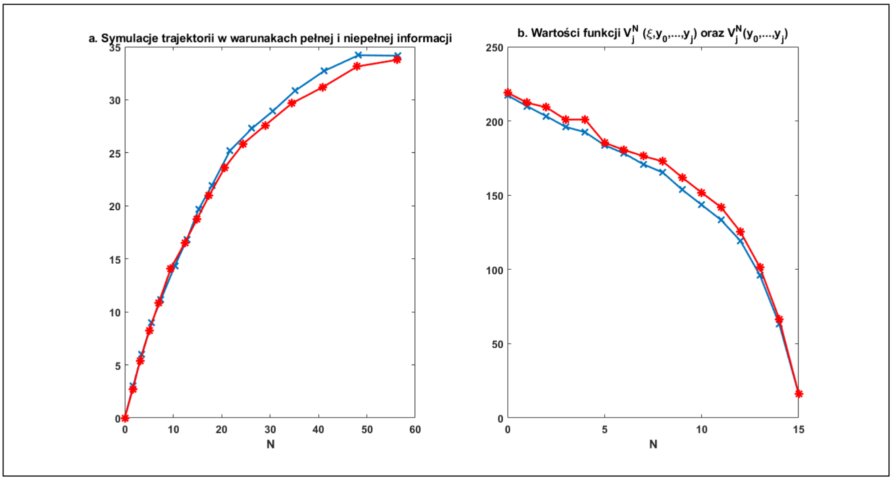
$$R_j = \begin{bmatrix} 0.64 & 0.05 \\ 0.01 & 0.95 \end{bmatrix}, \quad dla\ j = 0, 1, ..., 14, \quad Q_N = \begin{bmatrix} 1.25 & 0 \\ 0 & 0.75 \end{bmatrix}$$

odpowiednio. Zakładamy, że w warunkach niepełnej informacji wektor para-

metrów ξ odpowiadających za przesunięcie jest wektorem losowym o rozkładzie normalnym $\boldsymbol{N}(m_0, \Gamma_0)$, gdzie

$$m_0 = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} 0.4 & -0.02 \\ -0.02 & 0.1 \end{bmatrix},$$

natomiast w warunkach pełnej informacji przyjmujemy $\xi = m_0$.



Rysunek 2.1: Symulacje trajektorii systemu (2.19) oraz wartości funkcji Bellmana.

Wartości optymalnych sterowań w warunkach pełnej i niepełnej informacji o systemie (2.19) wyznaczamy ze wzorów (2.48) i (2.49) odpowiednio, natomiast w warunkach niepełnej informacji dla wektora losowego ξ wartości parametrów rozkładu a posteriori $\boldsymbol{N}(m_j, \Gamma_j)$, $j = 1, \ldots, N-1$ wyznaczamy ze wzorów (2.40) i (2.41). Rysunek 2.1a prezentuje przykładowe symulacje zachowania systemu (2.19). Krzywa granatowa przedstawia przykładową trajektorię dla przypadku pełnej informacji o systemie, natomiast krzywa czerwona – dla przypadku niepełnej informacji. Zachowanie funkcji Bellmana $V_j^N(\xi, y_0, \ldots, y_j)$, $\bar{V}_j^N(y_0, \ldots, y_j)$ dla $j = 0, \ldots, N$ są pokazane na rysunku 2.1b.

Tabela 2.1 zawiera symulację systemu (2.19) $\{y_i\}_{0 \le i \le N}$, które są oznaczone na rysunku (2.1a) znacznikiem " $\times$ ", wielkości optymalnych sterowań $\{u_i^*\}_{0 \le i \le N-1}$ w warunkach pełnej informacji, wartości funkcji Bellmana

Tabela 2.1: Symulacja systemu (2.19) w warunkach pełnej informacji, sterowania optymalne, wartości funkcji Bellmana, koszt zaburzeń zewnętrznych.

i	y_i	u_i^*	$V_i^N(\xi, y_0, ..., y_i)$	φ_j^N
0	$(0,0)$	$(-0.389,-2.246)$	217.263	0.612
1	$(1.757,3.046)$	$(-0.494,-2.206)$	209.981	0.603
2	$(3.504,5.984)$	$(-0.624,-2.164)$	203.203	0.593
3	$(5.539,8.985)$	$(-0.987,-2.094)$	195.999	0.583
4	$(7.479,11.185)$	$(-0.987,-2.094)$	192.433	0.571
5	$(10.353,14.359)$	$(-1.229,-1.987)$	183.675	0.559
6	$(12.824,16.814)$	$(-1.536,-1.922)$	178.183	0.545
7	$(15.246,19.672)$	$(-1.913,-1.792)$	170.830	0.53
8	$(18.009,21.931)$	$(-2.391,-1.694)$	165.293	0.513
9	$(21.716,25.196)$	$(-2.952,-1.428)$	153.708	0.493
10	$(26.25,27.292)$	$(-3.657,-1.253)$	143.542	0.47
11	$(30.588,28.934)$	$(-4.567,-1.103)$	133.345	0.442
12	$(35.174,30.850)$	$(-5.731,-0.832)$	119.262	0.405
13	$(41.209,32.725)$	$(-7.142,-0.411)$	96.235	0.355
14	$(48.268,34.204)$	$(-8.961,0.25)$	63.258	0.275
15	$(56.454,34.156)$		16.251	0

$\left\{V_j^N(\xi, y_0, ..., y_j)\right\}_{0\le j\le N}$ oraz $\left\{\varphi_j^N\right\}_{0\le j\le N}$ oczekiwane koszty spowodowane zaburzeniami zewnętrznymi.

Tabela 2.2 natomiast przedstawia stany systemu (2.19) $\{y_i\}_{0\le i\le N}$, które są oznaczone na rysunku (2.1a) znacznikiem " $*$ ", wielkości optymalnych sterowań $\{\tilde{u}_i^*\}_{0\le i\le N-1}$ w warunkach niepełnej informacji o systemie, wartości estymatorów $\{m_i\}_{0\le i\le N}$ wektora nieznanych parametrów ξ, wartości funkcji Bellmana $\left\{\bar{V}_j^N(y_0, ..., y_j)\right\}_{0\le j\le N}$. Dodatkowo w tabeli 2.2 podane są wielkości kosztów $\left\{\Psi_j^N\right\}_{0\le j\le N}$ spowodowanych nieznajomością parametrów systemu, które są ponoszone przez sterującego od momentu j do końca horyzontu sterowania N. Widzimy, ze zarówno ciągi kosztów $\left\{\varphi_j^N\right\}_{0\le j\le N}$ i $\left\{\Psi_j^N\right\}_{0\le j\le N}$ są ciągami nierosnącymi.

2.2.5 Wpływ horyzontu na koszty sterowania

W wielu zadaniach sterowania obiektami stochastycznymi bądź deterministycznymi horyzont (czas sterowania obiektem) jest ustalany odgórnie. Wobec powyższego jeżeli znany jest przedział czasowy w jakim sterujemy

Tabela 2.2: Symulacja systemu (2.19) w warunkach niepełnej informacji, wartości optymalnych sterowań, wartości estymatorów nieznanych parametrów, wartości funkcji Bellmana, wartości kosztów nieznajomości parametrów systemu.

i	y_i	$\tilde{u}_i^*$	m_i	$\bar{V}_j^N(y_0,...,y_j)$	Ψ_i^N
0	$(0,0)$	$(-0.389,-2.246)$	$(1,0.5)$	219.167	1.904
1	$(1.663,2.735)$	$(-0.459,-2.185)$	$(0.951,0.538)$	212.295	0.626
2	$(3.164,5.377)$	$(-0.630,-2.147)$	$(0.898,0.542)$	209.2	0.409
3	$(5.076,8.245)$	$(-0.789,-2.068)$	$(0.906,0.568)$	200.920	0.297
4	$(7.004,10.826)$	$(-1.001,-2.144)$	$(0.919,0.444)$	198.995	0.225
5	$(9.513,14.104)$	$(-1.228,-1.978)$	$(0.995,0.514)$	185.305	0.174
6	$(12.397,16.501)$	$(-1.536,-1.935)$	$(1.011,0.478)$	180.505	0.136
7	$(14.896,18.720)$	$(-1.919,-1.787)$	$(0.967,0.512)$	176.329	0.105
8	$(17.403,20.951)$	$(-2.407,-1.711)$	$(0.953,0.483)$	172.779	0.081
9	$(20.654,23.615)$	$(-2.957,-1.382)$	$(0.942,0.578)$	161.941	0.061
10	$(24.475,25.855)$	$(-3.643,-1.212)$	$(1.001,0.566)$	151.748	0.043
11	$(29.014,27.597)$	$(-4.552,-1.084)$	$(1.017,0.529)$	141.921	0.029
12	$(34.508,29.668)$	$(-5.729,-0.814)$	$(0.986,0.535)$	125.342	0.016
13	$(40.910,31.191)$	$(-7.139,-0.385)$	$(0.979,0.555)$	101.49	0.006
14	$(48.023,33.141)$	$(-8.962,0.285)$	$(0.965,0.575)$	66.585	0.005
15	$(56.282,33.771)$			16.252	0

obiektem, to stosując twierdzenia 2.2 i 2.3 dla systemów z czasem dyskretnym w warunkach pełnej i niepełnej informacji wyznaczamy prawa sterowania. Niestety koszt sterowania zależy liczby sterowań, dokładniej od długości przedziału czasowego w jakim sterujemy obiektem. Aby wyznaczyć optymalny horyzont należy rozwiązać zadanie czasooptymalne

$$\inf_{N\in\mathbb{N}}\inf_{u\in U} J^N(u). \tag{2.51}$$

Zadania czasooptymalne są omawiane w literaturze przedmiotu, patrz np. [11], [105], [109]. Poniżej zobaczymy, że horyzont ma dość duży wpływ na wartość wskaźnika jakości.

Przykład 2.2 *Załóżmy, że zachowanie systemu liniowego jest określone za pomocą równania*

$$y_{i+1} = Ay_i - Cu_i + \sigma w_{i+1} \tag{2.52}$$

dla $i \in \{0,1,...,N-1\}$*, gdzie* $\{w_i\}_{1\leq i\leq N}$ *jest ciągiem niezależnych dwuwymiarowych wektorów losowych o rozkładzie normalnym* $\boldsymbol{N}(\bar{0},I)$*,* $\bar{0}=(0,0)^T$,

I – macierzą jednostkową. Macierze odpowiadające za koszty sterowania i dziedziczenia są równe

$$R_i = \begin{bmatrix} 0.64 & 0.05 \\ 0.01 & 0.95 \end{bmatrix}, \quad Q_N = \begin{bmatrix} 2.2 & 0 \\ 0 & 1.5 \end{bmatrix}$$

dla $i = 0, 1, ..., N-1$, natomiast macierze transformacji stanów i sterowań

$$A = \begin{bmatrix} 0.9 & 0.2 \\ 0.6 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 1.5 & -0.1 \\ 0.04 & 1.1 \end{bmatrix}.$$

Dla systemu stochastycznego przyjmujemy

$$\sigma = \begin{bmatrix} 2.8 & 0.7 \\ 0.2 & 1.4 \end{bmatrix},$$

natomiast dla systemu deterministycznego macierz σ jest macierzą zerową. Zadanie polega na przeprowadzeniu systemów deterministycznego i stochastycznego (2.52) ze stanu $y_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ do stanu $a = \begin{bmatrix} 60 \\ 35 \end{bmatrix}$ przy najmniejszych kosztach, z tym że horyzont nie może przekroczyć np. maksymalnego czasu sterowania równego 20 ($N \leq 20$).

Korzystając ze wzorów (2.23) i (2.27) dla ustalonego horyzontu N oczekiwany całkowity koszt sterowania jest równy

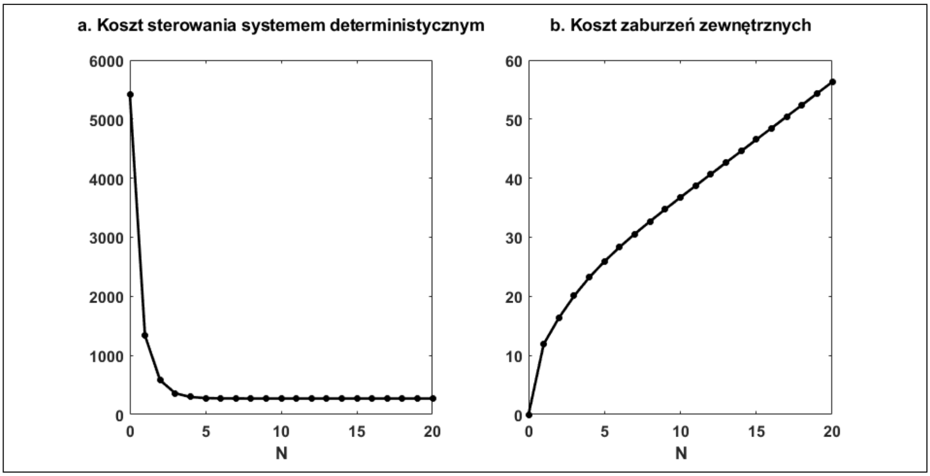
$$\inf_{u \in U} J^N(u) = V_0^N(y_0) = \chi_0^N(y_0) + \varphi_0^N,$$

gdzie $\chi_0^N(y_0) = \left[A^N y_0 - a\right]^T Q_0^N \left[A^N y_0 - a\right]$. W przypadku systemu stochastycznego dla każdego momentu $j \in \{0, 1, ..., N-1\}$ koszt eliminacji przyszłych zaburzeń zewnętrznych $\{w_i\}_{j+1 \leq i \leq N}$ wynosi $\varphi_j^N = \varphi_{j+1}^N + tr\left[\left(A^{N-j-1}\sigma\right)^T Q_{j+1} A^{N-j-1}\sigma\right]$ z warunkiem początkowym $\varphi_N^N = 0$. Dla systemu deterministycznego $\varphi_0^N = \varphi_1^N = ... = \varphi_N^N = 0$. Rozwiązując osobno dla systemów deterministycznego i stochastycznego zadanie

$$\inf_{0 \leq N \leq 20} \inf_{u \in U} J^N(u) = \inf_{0 \leq N \leq 20} V_0^N(y_0)$$

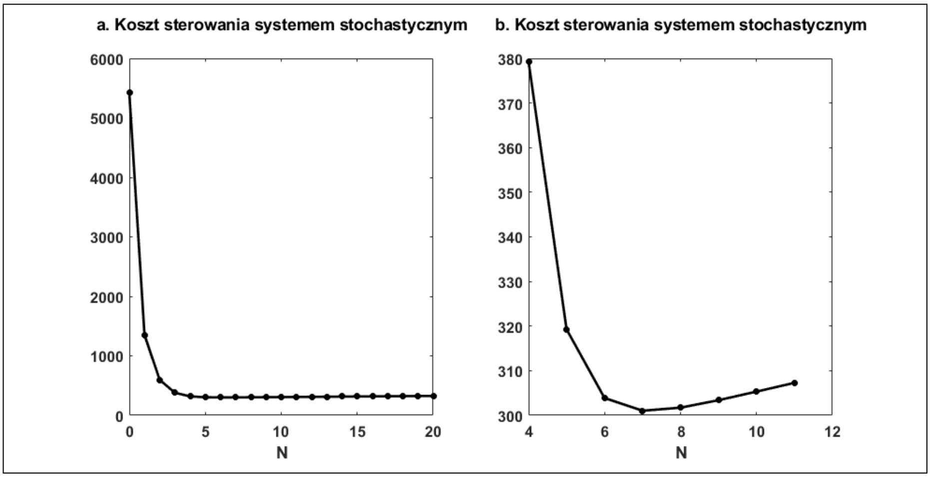
wyznaczymy optymalne horyzonty sterowań dla tych systemów.

Rysunek 2.2a przedstawia koszty sterowania systemem deterministycznym (σ jest macierzą zerową) $\chi_0^N(y_0)$ dla $N = 0, 1, 2, ..., 20$. Z rysunku widać, że ciąg kosztów sterowań $\left\{\chi_0^N(y_0)\right\}_{1 \le N \le 20}$ jest ciągiem nierosnącym (wraz ze wzrostem horyzontu N koszty realizacji celu sterowania nie rosną). Na rysunku 2.2b przedstawione są koszty związane z wpływem zaburzeń zewnętrznych φ_0^N podczas sterowania systemem (2.52). Z rysunku widać, że ciąg kosztów sterowań $\left\{\varphi_0^N\right\}_{1 \le N \le 20}$ jest ciągiem niemalejącym. Zatem wydłużając horyzont sterowania N dla systemu stochastycznego ponosimy coraz większe koszty spowodowane wpływem zaburzeń zewnętrznych.



Rysunek 2.2: Koszty sterowania systemem deterministycznym oraz koszty eliminacji zaburzeń zewnętrznych dla systemu (2.52).

Rysunek 2.3a przedstawia oczekiwane koszty sterowania $V_0^N(y_0) = \chi_0^N(y_0) + \varphi_0^N$ systemem stochastycznym (2.52) dla horyzontu $N = 0, 1, 2, ..., 20$, natomiast rysunek 2.3b pokazuje wartości $V_0^N(y_0)$ dla $N \in [4, 11]$. Widzimy, że dla $N = 7$ oczekiwane koszty sterowania systemem stochastycznym (2.52) są najmniejsze. Zatem, sterując systemem deterministycznym (σ jest macierzą zerową) należy ustalić horyzont $N = 20$, natomiast sterując systemem stochastycznym należy przyjąć $N = 7$.



Rysunek 2.3: Koszty sterowania systemem stochastycznym (2.52).

Dla systemu deterministycznego wydłużając horyzont sterowania N ponosimy mniejsze koszty, ponieważ istnieje możliwość rozłożenia nakładów energetycznych w czasie. W przypadku sterowania systemem stochastycznym dodatkowo należy uwzględnić koszty φ_0^N związane z wpływem zaburzeń zewnętrznych. Nadmierne wydłużanie horyzontu sterowania N może nie być opłacalne (φ_0^N jest funkcją niemalejącą ze względu na N). Tak jak ilustruje przykład powyżej, ponieważ horyzonty sterowania dla systemów stochastyczeego i deterministycznego są różne, to optymalne sterowania dla tych systemów również są różne.

Wniosek 2.1 *Prawa sterowania pośrednio zależą od długości przedziału czasowego, w którym sterujemy obiektem. Prosty przykład pokazuje, że horyzont sterowania ma wpływ na koszty realizacji celu $J^N(u)$ (wielkość wskaźnika jakości (2.2), (2.20)).*

Wniosek 2.2 *Jeżeli horyzont sterowania N jest ustalony, to prawa sterowania systemami (2.1), (2.19) wyznaczamy korzystając z twierdzeń 2.1–3. Jeżeli horyzont sterowania nie jest określony, to najpierw należy rozwiązać zadanie (2.51) określając przedział czasowy (dokładniej horyzont N), w którym będzie realizowany cel sterowania (2.3).*

Rozdział 3

Optymalne sterowanie adaptacyjne dla horyzontu losowego niezależnego od stanów systemu

Większość zadań z teorii sterowania jest zdefiniowana dla ustalonego horyzontu (patrz np. [2], [6], [7], [10], [13], [15], [16], [38], [44], [71], [91], [92], [98], [108]). Przedział, w którym sterujemy systemem, jest ustalany odgórnie lub poprzez rozwiązanie zadania czasooptymalnego. W poprzednim rozdziale pokazano, że optymalne sterowanie zależy nie tylko od stanu systemu i momentu w którym sterujemy, ale również od horyzontu. Dodatkowy problem sterowania obiektem powstaje, gdy horyzont nie jest znany (nie jest ustalony odgórnie). Jak w takim przypadku skonstruować prawa sterowania?

Zadania z losowym horyzontem są mniej widoczne w literaturze przedmiotu i można je podzielić na dwa typy. Pierwszy typ zadań optymalnego sterowania z nieznanym horyzontem polega na uwzględnieniu możliwości zatrzymania systemu (dotyczy optymalnego zatrzymania procesu sterowania obiektem, zobacz np. [8], [11], [15], [25], [28], [51], [52], [61], [63]).

Optymalny moment zatrzymania sterowania systemem zależy od stanu tego systemu. Zadania tego typu zostaną omówione w kolejnym rozdziale.

Drugi typ zadań optymalnego sterowania dotyczy przypadku, gdy horyzont sterowania nie zależy od zachowania systemu (stan systemu nie ma wpływu na czas działania, sterowania systemem), patrz np. [57], [58], [59], [60]. Z taką sytuacją mamy do czynienia na przykład w procesie produkcyjnym, gdy nie mamy gwarancji, że w określonym przedziale czasowym maszyny produkcyjne będą działać bezawaryjnie, będą wykonywać zaprogramowane czynności prawidłowo (czas bezawaryjnej pracy urządzeń nie jest znany, np. [41]). Horyzont sterowania opisujący np. liczbę możliwych strat, liczbę elementów wymienianych podczas procesu decyzyjnego, czas prawidłowej pracy maszyny, czas "życia" produktu itp. modelujemy za pomocą zmiennej losowej niezależnej od stanów systemu. Zadania tego typu zostaną omówione w bieżącym rozdziale. Rozwiązanie polega na konstrukcji zadania zastępczego, dokładniej na modyfikacji wskaźnika jakości.

3.1 Postawienie zadania

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną. Na tej przestrzeni definiujemy ciąg niezależnych wektorów losowych $\{w_i\}_{i\geq 1}$ o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, gdzie $w_i : \Omega \to \mathbb{R}^m$ dla $i \geq 1$, $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Niech $\xi : \Omega \to \mathbb{R}^k$ będzie wektorem losowym o rozkładzie a priori $P\,(d\xi)$, który reprezentuje nieznane parametry systemu. Wektor y_0 oznacza stan początkowy systemu. Ogólnie zakładamy, że wektor $y_0 : \Omega \to \mathbb{R}^n$ ma rozkład $P\,(dy_0)$. W przypadku, gdy y_0 jest wektorem deterministycznym, to wystarczy przyjąć, że y_0 ma rozkład punktowy. Zakładamy, że wszystkie wymienione obiekty są stochastycznie niezależne. Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy niemalejące rodziny $\sigma-$ciał $\{\mathcal{F}_i\}_{i\geq 0}$ oraz $\left\{\mathcal{F}_i^\xi\right\}_{i\geq 0}$, gdzie $\mathcal{F}_i = \sigma\,(y_0) \vee \sigma\,\{w_s : s = 1, 2, ..., i\}$, $\mathcal{F}_i^\xi = \mathcal{F}_i \vee \sigma\,\{\xi\}$.

Poniżej rozważamy problem sterowania systemem stochastycznym o równaniu stanu

$$y_{i+1} = f(\xi, y_i, u_i) + \sigma(\xi, y_i) w_{i+1}, \tag{3.1}$$

gdzie $i \in N_0$, $y_i \in \mathbb{R}^n$, $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ oraz $\sigma : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n \times m}$. Zakładamy, że funkcje f, σ są ciągłe względem wszystkich argumentów. Wektor $u_j \in \mathbb{R}^l$ mierzalny ze względu na σ–ciało $\mathcal{F}_j$ nazywamy działaniem sterującym w chwili $j \in \mathbb{N}_0$, natomiast $u = (u_0, u_1, ...)$ – sterowaniem dopuszczalnym. Klasę wszystkich sterowań oznaczamy przez U.

Niech $\tau : \Omega \to \mathbb{N}_0$ będzie zmienną losową o rozkładzie dyskretnym, tzn. $P(\tau = k) = p_k$, $0 \le p_k \le 1$ dla $k \in \mathbb{N}_0$ oraz $\sum_{i=0}^{\infty} p_i = 1$. Zmienna losowa τ oznacza horyzont sterowania systemem (3.1). Zakładamy, że horyzont sterowania τ nie zależy od stanów systemu. Przyjmujemy następujące oznaczenia: $\mathcal{F}^\tau = \sigma\{\tau\}$, $\mathcal{F}_j^{\xi\tau} = \mathcal{F}_j^\xi \vee \sigma\{\tau > j\}$ oraz $\mathcal{F}_j^\tau = \mathcal{F}_j \vee \sigma\{\tau > j\}$. W przypadku gdy $\tau : \Omega \to \{0, 1, 2, ..., N\}$ przyjmujemy $\mathcal{F} = \mathcal{F}_N^{\xi\tau}$, natomiast dla przypadku gdy $\tau : \Omega \to \mathbb{N}_0$, to przyjmujemy $\mathcal{F} = \lim_{t\to\infty} \mathcal{F}_t^{\xi\tau}$.

Niech funkcja $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ oznacza wartość strat, natomiast funkcja $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ wielkość dziedziczenia na końcu horyzontu sterowania τ. Zakładamy, że funkcje g i h są ciągłe względem swoich argumentów oraz ograniczone. Dla każdego momentu $j < \tau$ dokonujemy sterowania u_j oraz ponosimy koszty w wysokości $g(y_j, u_j)$. W chwili τ nie dokonujemy sterowania systemem (3.1), natomiast ponosimy koszty dziedziczenia $h(y_\tau)$ (koszty związanie ze stanem systemu y_τ). Jeżeli zmienna losowa τ przyjmie wartość 0, to w takim przypadku system nie będzie sterowany, zatem nie ponosimy żadnych kosztów sterowania, natomiast ponosimy tylko koszt dziedziczenia wielkości $h(y_0)$. Wskaźnik jakości definiujemy jako sumę strat (kosztów, nakładów energetycznych) dla momentów $0, 1, ..., \tau - 1$ oraz dziedziczenia dla horyzontu sterowania τ. Całkowity koszt sterownia $J^\tau(u)$ definiujemy wzorem

$$J^\tau(u) = E\left(\sum_{i=-1}^{\tau-1} g(y_i, u_i) + h(y_\tau) \middle| \mathcal{F}^\tau \right), \tag{3.2}$$

gdzie $g(y_{-1}, u_{-1}) = 0$. We wzorze (3.2) bierzemy operator wartości oczekiwanej względem σ–ciała generowanego przez zmienną losową τ. Wskaźnik jakości $J^{\tau}(u)$ zależy od τ, a zatem jest zmienną losową.

Głównym celem sterowania adaptacyjnego jest minimalizacja całkowitych kosztów sterowania (3.2). Zatem, sterując systemem (3.1) z losowym horyzontem niezależnym od stanów, należy rozwiązać zadanie

$$\inf_{u \in U} EJ^{\tau}(u) \tag{3.3}$$

oraz wyznaczyć sterowanie u^* dla którego infimum jest osiągnięte. Tak skonstruowane zadanie (3.3) jest zadaniem typu planistycznego. Operator E jest operatorem wartości oczekiwanej po zmiennej losowej τ.

Uwaga 3.1 *Jeżeli horyzont sterowania zostanie zastąpiony wartością deterministyczną $[E\tau]$ (w przypadku sterowania systemem (3.1) z losowym horyzontem τ niezależnym od stanów systemu), gdzie $[\cdot]$ oznacza część całkowitą z liczby rzeczywistej (lub np. zaokrąglenie do liczby naturalnej), to do rozwiązania zadania*

$$\inf_{u \in U} J^{[E\tau]}(u), \tag{3.4}$$

stosujemy bezpośrednio twierdzenie 2.1. Warunki konieczne optymalności sterowań dla zadań (3.3) i (3.4) są różne.

Zbiór realizacji zmiennej losowej τ reprezentującej horyzont sterowania może być skończony bądź przeliczalny. Poniżej zostaną przedstawione rozwiązania zadania (3.3) dla dwóch przypadków. W rozdziale 3.2 będzie przedstawiony problem optymalnego sterowania z losowym horyzontem τ o skończonej liczbie realizacji. Zadanie (3.3) zostanie przekształcone do zadania optymalnego sterowania z ustalonym horyzontem. W rozdziale 3.3 będzie rozważany problem optymalnego sterowania z losowym horyzontem τ, gdzie zbiór realizacji jest zbiorem przeliczalnym. W tym przypadku zadanie (3.3) zostanie zmodyfikowane do zadania optymalnego sterowania z nieskończonym horyzontem.

3.2 Sterowanie dla horyzontu losowego o skończonej liczbie realizacji

3.2.1 Optymalne sterowanie dla zadania planistycznego

Niech system stochastyczny będzie określony za pomocą równania stanu (3.1). Zakładamy również, że horyzont sterowania we wskaźniku jakości (3.2) jest losowy i nie zależy od stanów systemu (3.1). Przyjmujemy, że losowy horyzont τ ma skończoną liczbę realizacji, tzn. $\tau : \Omega \rightarrow \{0, 1, 2, ..., N\}$, o rozkładzie

$$P(\tau = k) = p_k$$

dla $k \in \{0, 1, 2, ..., N\}$, gdzie $p_k \geq 0$ oraz $\sum\limits_{k=0}^{N} p_k = 1$. Wielkość p_k, $0 \leq k \leq N$ oznacza prawdopodobieństwo zatrzymania systemu w momencie k. Rozwiązując zadanie (3.3) wyznaczymy optymalne sterowanie.

Ze wzoru na prawdopodobieństwo całkowite funkcjonał (3.2) możemy przedstawić w postaci

$$\begin{aligned} EJ^{\tau}(u) &= \sum_{j=0}^{N} E\left(J^{\tau}(u) \,|\tau = j\right) P(\tau = j) = P(\tau = 0)\, E\left(h(y_0) \,|\tau = 0\right) \\ &+ \sum_{j=1}^{N} P(\tau = j)\, E\left(\sum_{i=0}^{j-1} g(y_i, u_i) + h(y_j) \,|\tau = j\right). \end{aligned}$$

Stąd otrzymujemy

$$\begin{aligned} EJ^{\tau}(u) &= P(\tau = 0)\, h(y_0) \\ &+ E\left(\sum_{j=0}^{N-1} g(y_j, u_j) \sum_{i=j+1}^{N} P(\tau = i) + \sum_{j=1}^{N} h(y_j)\, P(\tau = j)\right). \end{aligned}$$

Zatem wartość oczekiwana wskaźnika jakości (3.2) jest równa

$$\bar{J}^{N}(u) = EJ^{\tau}(u) = E\left(\sum_{j=0}^{N} \phi_j(y_j, u_j)\right), \tag{3.5}$$

gdzie dla dowolnego $0 \leq j \leq N$

$$\begin{aligned} \phi_j(y_j, u_j) &= g(y_j, u_j) \sum_{i=j+1}^{N} p_i + h(y_j) p_j \\ &= g(y_j, u_j) s_j + h(y_j) p_j, \end{aligned} \tag{3.6}$$

natomiast $s_j = P(\tau > j) = \sum_{i=j+1}^{N} p_i$ oznacza prawdopodobieństwo przeżycia systemu momentu j, tzn. prawdopodobieństwo, że horyzont jest większy od j. Ponieważ $s_N = P(\tau > N) = 0$, zatem

$$\phi_N(y_N, u_N) = h(y_N) p_N. \tag{3.7}$$

Dla momentu N (patrz wzór (3.7)) koszt sterowania jest równy iloczynowi kosztu dziedziczenia $h(y_N)$ i prawdopodobieństwa zatrzymania systemu. W chwili N nie podejmujemy żadnego sterowania oraz przyjmujemy $u_N = \mathrm{col}(0, 0, ..., 0) \in \mathbb{R}^l$.

Ze wzoru (3.6) wynika, że w chwili $j = 0, 1, ..., N-1$ oczekiwany koszt sterowania $\phi_j(y_j, u_j)$ systemem stochastycznym (3.1) jest równy sumie iloczynu nakładów energetycznych $g(y_j, u_j)$ i prawdopodobieństwa przeżycia momentu j (prawdopodobieństwa, że system zostanie zatrzymany w momencie późniejszym niż j) oraz iloczynu kosztu dziedziczenia $h(y_j)$ i prawdopodobieństwa, że horyzont wynosi j.

Wobec powyższego zadanie optymalnego sterowania (3.3) z losowym horyzontem τ o skończonej liczbie realizacji zastępujemy zadaniem optymalnego sterowania ze skończonym horyzontem N. Aby wyznaczyć optymalne prawa sterowania systemem (3.1) należy rozwiązać zadanie

$$\inf_{u \in U} \bar{J}^N(u), \tag{3.8}$$

gdzie wskaźnik jakości $\bar{J}^N(u)$ jest dany wzorem (3.5).

Uwaga 3.2 *Wartość oczekiwaną wskaźnika jakości (3.2) również możemy przedstawić w postaci*

$$\bar{J}^N(u) = E\left(\sum_{j=0}^{N-1}\left(g(y_j,u_j)\,I_{\tau>j} + h(y_j)\,I_{\tau=j}\right) + h(y_N)\,I_{\tau=N}\right), \tag{3.9}$$

gdzie $I_{\tau>j} = I_{\{\omega:\tau(\omega)>j\}}$, $I_{\tau=j} = I_{\{\omega:\tau(\omega)=j\}}$ *oraz* $I_A = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$

Definicja 3.1 *Sterowanie* $\left(u_0^*, u_1^*, ..., u_{\tau-1}^*\right) \subset \left(u_0^*, u_1^*, ..., u_{N-1}^*\right)$ *nazywamy optymalnym sterowaniem dla systemu (3.1) z losowym horyzontem* $\tau : \Omega \to \{0, 1, 2, ..., N\}$ *o skończonej liczbie realizacji, gdzie* $\left(u_0^*, u_1^*, ..., u_{N-1}^*\right)$ *jest rozwiązaniem zadania (3.8).*

Twierdzenie poniżej przedstawia warunki konieczne wyznaczenia optymalnego sterowania dla zadania (3.8).

Twierdzenie 3.1 *Załóżmy, że funkcje* $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ *i* $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ *są ciągłe, wypukłe i ograniczone, funkcje* $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ *i* $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ *są funkcjami klasy* C^1 *ze względu na zmienną sterującą* $u \in \mathbb{R}^l$ *oraz* $\det(\Sigma(\xi, y)) \neq 0$ *dla* $(\xi, y) \in \mathbb{R}^k \times \mathbb{R}^n$, *gdzie* $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$. *Jeżeli* u^* *jest optymalnym sterowaniem dla zadania (3.8), to dla dowolnego* $j \in \{0, 1, 2, ..., N-1\}$ *spełniona jest równość*

$$E\left(\left(\sum_{i=j+1}^{N}\phi_i(y_i,u_i^*)\right)\left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_i)\nabla_u f(\xi, y_j, u_j^*)\middle|\,\mathcal{F}_j\right)$$
$$+P(\tau > j)\,\nabla_{u_j} g\left(y_j, u_j^*\right) = \mathbf{0}, \tag{3.10}$$

gdzie $\phi_j(\cdot)$, $0 \leq j \leq N$ *są dane wzorami (3.6)-(3.7), natomiast* $\mathbf{0} = \mathrm{col}(0, ..., 0) \in \mathbb{R}^l$.

Dowód. Stosując bezpośrednio twierdzenie 2.1 otrzymujemy tezę. ∎

Uwaga 3.3 *Za pomocą twierdzenia powyżej wyznaczamy sterowanie optymalne* $u^* = \left(u_0^*, u_1^*, ..., u_{N-1}^*\right)$ *dla systemu (3.1), natomiast sterując tym* *systemem z losowym horyzontem* $\tau : \Omega \to \{0, 1, ..., N\}$ *(czas sterowania jest losowy) należy zastosować sterowanie* $\left(u_0^*, u_1^*, ..., u_{\tau-1}^*\right) \subset u^*$.

Przykład poniżej pokazuje różnice w sterowaniach bardzo prostego systemu liniowego dla przypadków z ustalonym i losowym horyzontem.

Przykład 3.1 *Niech liniowy system stochastyczny będzie określony za pomocą równania*

$$y_{j+1} = \xi u_j + w_{j+1} \tag{3.11}$$

dla $j \in \mathbb{N}_0$. *Przyjmujemy, że parametr* $\xi \in \mathbb{R}$ *jest znany oraz* $\{w_i\}_{i \geq 1}$ *jest ciągiem niezależnych zmiennych losowych o rozkładzie normalnym* $\boldsymbol{N}(0, \sigma^2)$. *System przeprowadzamy z punktu* $y_0 \in \mathbb{R}$ *do punktu* $a \in \mathbb{R}$. *Jako funkcjonał kosztów przyjmujemy* $g(y, u) = u^2$, *natomiast jako funkcjonał dziedziczenia* $h(y) = (y - a)^2$. *Należy zauważyć, że na stan bieżący systemu (3.11) nie wpływa wartość poprzednich stanów (system jest sterowany w pętli otwartej). Horyzont sterowania* τ *nie jest znany i nie zależy od stanów systemu oraz przyjmujemy, że jest zmienną losową o skończonej liczbie realizacji. Prawdopodobieństwo, że system zostanie zatrzymany w momencie* $k \in \{0, 1, 2, ..., N\}$ *jest równe*

$$P(\tau = k) = p_k,$$

gdzie $p_k \geq 0$ *oraz* $\sum\limits_{k=0}^{N} p_k = 1$. *Żadnych ograniczeń na wielkość sterowania nie nakładamy.*
Aby wyznaczyć optymalne sterowanie należy rozwiązać zadanie

$$\inf_u \bar{J}^N(u), \tag{3.12}$$

gdzie wskaźnik jakości

$$\bar{J}^N(u) = E\left(\sum_{j=0}^{N} s_j u_j^2 + p_j (y_j - a)^2\right) \tag{3.13}$$

zależy od sterowania $u = (u_0, u_1, ..., u_{N-1})$, *natomiast prawdopodobieństwo przeżycia* $s_j = P(\tau > j)$. *System maksymalnie może być sterowany do* *momentu $N-1$, natomiast w chwili N ponosimy tylko koszt dziedziczenia $p_N\left(y_N-a\right)^2$ oraz $s_N=0$.*

Do wyznaczenia optymalnych sterowań definiujemy funkcje Bellmana

$$V_j\left(y_j\right)=\inf_{u_j} E\left(s_j u_j^2+p_j\left(y_j-a\right)^2+V_{j+1}\left(y_{j+1}\right)|\mathcal{F}_j\right) \tag{3.14}$$

dla $0\le j\le N-1$. Dla momentu N wartość funkcji Bellmana jest równa

$$V_N\left(y_N\right)=p_N\left(y_N-a\right)^2.$$

Dla dowolnego $j=0,1,...,N-1$ optymalne sterowanie jest równe

$$u_j^*=\frac{ap_{j+1}\xi}{s_j+p_{j+1}\xi^2}, \tag{3.15}$$

natomiast wartość funkcji Bellmana dla momentów $0\le j\le N$ wynosi

$$V_j\left(y_j\right)=p_j\left(y_j-a\right)^2+s_j\left(\sigma^2+a^2\right)-\varphi_j, \tag{3.16}$$

gdzie

$$\varphi_j=\varphi_{j+1}+a^2\xi^2\frac{p_{j+1}^2}{s_j+p_{j+1}\xi^2} \tag{3.17}$$

z warunkiem początkowym $\varphi_N=0$.

Wzór (3.16) udowodnimy za pomocą indukcji matematycznej. Zakładamy, że jest on prawdziwy dla dowolnego $j+1$ oraz sprawdzamy dla momentu j. Ze wzoru (3.14) wartość funkcji Bellmana dla systemu (3.11) w chwili j jest równa

$$\begin{aligned}V_j\left(y_j\right)=\inf_{u_j} E\Big(&s_j u_j^2+p_j\left(y_j-a\right)^2+s_{j+1}\left(\sigma^2+a^2\right)\\&+p_{j+1}\left(\xi u_j+w_{j+1}-a\right)^2-\varphi_{j+1}|\mathcal{F}_j\Big).\end{aligned}$$

Z własności warunkowej wartości oczekiwanej

$$V_j(y_j) = \inf_{u_j} \left(s_j + p_{j+1}\xi^2\right)^2 u_j^2 - 2a\xi p_{j+1}u_j$$
$$+ p_j(y_j - a)^2 + (s_{j+1} + p_{j+1})\left(\sigma^2 + a^2\right) - \varphi_{j+1}. \tag{3.18}$$

Zatem optymalne sterowanie w momencie j jest dane wzorem (3.15). Podstawiając (3.15) do (3.18) mamy

$$V_j(y_j) = p_j(y_j - a)^2 + (s_{j+1} + p_{j+1})\left(\sigma^2 + a^2\right) - \varphi_{j+1} - a^2\xi^2 \frac{p_{j+1}^2}{s_j + p_{j+1}\xi^2}.$$

Ponieważ $s_j = s_{j+1} + p_{j+1}$ oraz korzystając ze wzoru (3.17) otrzymujemy (3.16). Oczekiwany koszt sterowania systemem (3.11) wynosi

$$\inf_u \bar{J}^N(u) = V_0(y_0) = p_0(y_0 - a)^2 + s_0\left(\sigma^2 + a^2\right) - a^2\xi^2 \sum_{k=0}^{N-1} \frac{p_{k+1}^2}{s_k + p_{k+1}\xi^2}.$$

Uwaga 3.4 *Otrzymane wyniki (optymalne sterowanie (3.15) oraz wyznaczenie wartości funkcji Bellmana (3.14)) można wykorzystać również dla sterowania systemem dla przypadku z ustalonym horyzontem N. W tym celu wystarczy przyjąć $P(\tau = N) = p_N = 1$ oraz $P(\tau = j) = p_j = 0$ dla $j = 0, 1, ..., N-1$. Optymalne sterowanie jest dane wzorem*

$$u_j^* = \begin{cases} 0, & j = 0, 1, ..., N-1, \\ \frac{a\xi}{1+\xi^2}, & j = N, \end{cases} \tag{3.19}$$

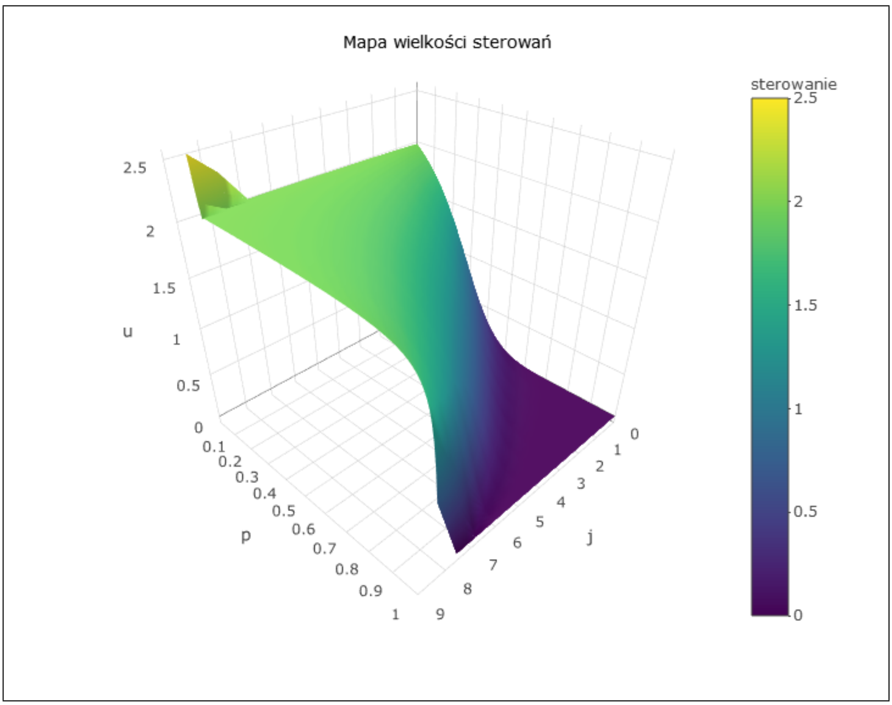
natomiast wartość funkcji Bellmana

$$V_j(y_j) = \begin{cases} \sigma^2 + \frac{a^2}{1+\xi^2}, & j = 0, 1, ..., N-1, \\ (y_N - a)^2, & j = N. \end{cases} \tag{3.20}$$

Uwaga 3.5 *Optymalne sterowania systemem stochastycznym (3.11) (system bez pamięci poprzednich stanów) dla przypadków z horyzontem deterministycznym i z horyzontem losowym są całkowicie różne. Jeżeli horyzont jest ustalony, to system pozostaje w bezczynności w momentach $j = 0, 1, ..., N-2$,* *natomiast działanie sterujące podejmujemy tylko w momencie* $j = N - 1$. *Jeżeli horyzont sterowania jest losowy (nie wiemy ile razy będziemy sterować systemem), to zgodnie ze wzorem (3.15) działania sterujące należy podejmować w każdej chwili, która nie jest momentem końcowym.*

Uwaga 3.6 *Jeżeli zadanie (3.12) zastępujemy zadaniem (3.4), to, zgodnie z uwagą powyżej, działania sterujące dla systemu (3.11) powinny być podejmowane tylko w momencie* $j = [E\tau] - 1$.

Uwaga 3.7 *Dla systemu (3.11) (systemu sterowanego w pętli otwartej) widzimy, że w przypadku ustalonego horyzontu oczekiwany koszt sterowania* $V_0(y_0)$ *nie zależy od stanu początkowego* y_0, *natomiast w przypadku losowego horyzontu oczekiwany koszt sterowania* $V_0(y_0)$ *zależy od stanu poczatkowego* y_0 *(wystarczy porównać (3.20) i (3.16)).*



Rysunek 3.1: Wartości optymalnych sterowań dla systemu stochastycznego (3.11) z losowym horyzontem τ o rozkładzie dwumianowym.

Przykład 3.2 *Niech liniowy system stochastyczny będzie określony za pomocą równania (3.11), natomiast wskaźnik jakości za pomocą (3.13). Poniżej zakładamy, że horyzont sterowania* τ *ma rozkład dwumianowy*

$$P(\tau = j) = \binom{N}{j} p^j (1-p)^{N-j},$$

gdzie $0 \leq p \leq 1$ *oraz* $0 \leq j \leq N$. *Dodatkowo przyjmujemy, że* $N = 10$, $a = 5$, $\xi = 2$. *Rysunek 3.1 przedstawia mapę wielkości sterowań dla* $0 < p < 1$. *W przypadku gdy* $p = 1$, *to optymalne sterowanie jest dane wzorem (3.19). W przypadku gdy* $p \to 0$, *to* $\lim\limits_{p\to 0} P(\tau = 0) = 1$, *natomiast* $\lim\limits_{p\to 0} P(\tau = j) = 0$ *dla* $1 \leq j \leq N$. *Zatem* $\lim\limits_{p\to 0} s_j = 0$ *dla* $0 \leq j \leq N-1$. *Ze wzoru (3.15) otrzymujemy*

$$\lim_{p\to 0} u_j^* = \frac{a}{\xi}$$

dla dowolnego momentu $j = 0, 1, ..., N-1$.

3.2.2 Optymalne sterowanie dla zadania operacyjnego

Twierdzenie 3.1 podaje warunki konieczne optymalnego sterowania dla zadania planistycznego (zadania w którym wykorzystujemy tylko informację a priori dotyczącą horyzontu losowego τ). Sterując obiektem (3.1) w każdym momencie możemy również wykorzystać informację a posteriori dotyczącą horyzontu τ. Jeżeli system znajduje się w momencie j oraz ten moment nie jest horyzontem (tzn. system powinien być sterowany w tym momencie), oznacza to, że system był sterowany w chwilach $0, 1, ..., j-1$ oraz może być zatrzymany w chwilach $j+1, ..., N$. Poniżej odpowiemy na pytanie w jaki sposób dodatkowa informacja o horyzoncie τ (dokładniej informacja o warunkowym rozkładzie zmiennej τ) w momencie $j < \tau$ wpływa na sterowanie obiektem.

Wskaźnik jakości $\bar{J}^N(u)$ określony wzorem (3.5) możemy przedstawić również za pomocą rozwinięcia

$$\begin{aligned}\bar{J}^N(u) = E\Big(&h(y_0)P(\tau=0)+P(\tau>0)\Big(g(y_0,u_0)\\ &+\frac{1}{P(\tau>0)}\Big(h(y_1)P(\tau=1)+P(\tau>1)\Big(g(y_1,u_1)\\ &+\frac{1}{P(\tau>1)}\Big(h(y_2)P(\tau=2)+...+P(\tau>N-1)\\ &\times\Big(g(y_{N-1},u_{N-1})+h(y_N)\frac{P(\tau=N)}{P(\tau>N-1)}\Big)...\Big)\Big)\Big)\Big)\Big).\end{aligned}$$

Ponieważ zmienna losowa $\tau:\Omega\to\{0,1,2,...,N\}$, zatem $P(\tau=N)=P(\tau>N-1)$. Korzystając ze wzoru na prawdopodobieństwo warunkowe mamy

$$\begin{aligned}\bar{J}^N(u) = E\Big(&h(y_0)P(\tau=0)+P(\tau>0)\Big(g(y_0,u_0)+h(y_1)P(\tau=1\,|\tau>0)\\ &+P(\tau>1\,|\tau>0)\Big(g(y_1,u_1)+h(y_2)P(\tau=2\,|\tau>1)+...\\ &+P(\tau>N-1\,|\tau>N-2)\left(g(y_{N-1},u_{N-1})+h(y_N)\right)\Big)\Big)\Big).\end{aligned} \tag{3.21}$$

Dla dowolnego $0\le j\le N-1$ wskaźnik jakości (3.21) możemy również przedstawić w postaci

$$\begin{aligned}\bar{J}^N(u) = E\Bigg(&\sum_{k=0}^{j-1}\left(g(y_k,u_k)P(\tau>k)+h(y_k)P(\tau=k)\right)+h(y_j)P(\tau=j)\\ &+P(\tau>j)\Bigg(g(y_j,u_j)+\sum_{s=j+1}^{N-1}\Big(g(y_s,u_s)P(\tau>s\,|\tau>j)\\ &+h(y_s)P(\tau=s\,|\tau>j)\Big)+h(y_N)P(\tau=N\,|\tau>j)\Bigg)\Bigg).\end{aligned} \tag{3.22}$$

Warunki konieczne optymalnego sterowania dla zadania operacyjnego podaje twierdzenie poniżej.

Twierdzenie 3.2 *Załóżmy, że funkcje* $g:\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}$ *i* $h:\mathbb{R}^n\longrightarrow\mathbb{R}$ *są ciągłe, wypukłe i ograniczone, funkcje* $f:\mathbb{R}^k\times\ \mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}^n$

i $g:\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}^n$ *są funkcjami klasy* C^1 *ze względu na zmienną sterującą* $u\in\mathbb{R}^l$ *oraz* $\det(\Sigma(\xi,y))\neq 0$ *dla* $(\xi,y)\in\mathbb{R}^k\times\mathbb{R}^n$, *gdzie* $\Sigma(\xi,y)=\sigma(\xi,y)\sigma^T(\xi,y)$. *Jeżeli moment* $j\in\{0,1,2,...,N-1\}$ *nie jest horyzontem sterowania (w momencie* j *system (3.1) nie zostanie zatrzymany), to optymalne sterowanie* u_j^* *dla zadania (3.8) spełnia równanie*

$$E\left(\left(\sum_{i=j+1}^{N}\phi_{i|j}(y_i,u_i^*)\right)\left(y_{j+1}-f(\xi,y_j,u_j^*)\right)^T\Sigma^{-1}(\xi,y_i)\nabla_u f(\xi,y_j,u_j^*)\middle|\,\mathcal{F}_j^{\tau}\right)$$

$$+\nabla_{u_j}g\left(y_j,u_j\right)=\mathbf{0}, \tag{3.23}$$

gdzie $\mathbf{0}=\mathrm{col}\,(0,...,0)\in\mathbb{R}^l$, $\phi_{i|j}(\cdot)$ *dla* $0\le j<i\le N-1$ *jest dane wzorem*

$$\phi_{i|j}\left(y_i,u_i\right)=g\left(y_i,u_i\right)P\left(\tau>i\,|\tau>j\,\right)+h\left(y_i\right)P\left(\tau=i\,|\tau>j\,\right), \tag{3.24}$$

natomiast

$$\phi_{N|j}\left(y_i,u_i\right)=h\left(y_i\right)P\left(\tau=N\,|\tau>j\,\right). \tag{3.25}$$

Dowód. Z równania (3.1) wynika, że dla dowolnego $1\le i\le N$ warunkowy rozkład wektora y_i jest dany wzorem

$$P\left(dy_i\,\middle|\,\mathcal{F}_{i-1}^{\xi}\right)=\gamma(y_i,f(\xi,y_{i-1},u_{i-1}),\Sigma(\xi,y_{i-1}))dy_i,$$

gdzie $\gamma()$ jest funkcją gęstości rozkładu normalnego określoną wzorem (1.4). Ponieważ zaburzenia zewnętrzne $w_1,...,w_N$ oraz wektor losowy ξ są wzajemnie niezależne, to rozkłady łączne $P\left(d\xi,dy_0,\ldots,dy_j\right)$ oraz $P\left(dy_j,\ldots,dy_i\right)$ dla dla $1\le j\le i\le N$ są określone za pomocą wzorów (2.5) i (2.6) odpowiednio.

Korzystając z własności warunkowej wartości oczekiwanej dla dowolnego $j\in\{0,1,...,N-1\}$ funkcjonał (3.5) możemy przedstawić w postaci

$$\bar{J}^N(u)=E\left(\sum_{i=0}^{j}\phi_i(y_i,u_i)+I_{\tau>j}E\left(\sum_{i=j+1}^{N}\phi_i(y_i,u_i)\middle|\,\mathcal{F}_j^{\xi\tau}\right)\right)$$

$$= \int \left(\sum_{i=0}^{j-1} \phi_i(y_i, u_i) + h\left(y_j\right) P\left(\tau = j\right) + P\left(\tau > j\right) \left(g\left(y_j, u_j\right)\right.\right.$$
$$\left.\left. + \int \sum_{i=j+1}^{N} \frac{\phi_i(y_i, u_i)}{P\left(\tau > j\right)} P_{j+1,N}\left(dy_{j+1}, \ldots, dy_N\right)\right) P\left(d\xi, dy_0, \ldots, dy_j\right). \quad (3.26)$$

Dla dowolnych $0 \leq j < i \leq N$ korzystając ze wzoru na prawdopodobieństwo warunkowe oraz wzoru (3.6) otrzymujemy

$$\phi_{i|j}\left(y_i, u_i\right) = \frac{\phi_i(y_i, u_i)}{P\left(\tau > j\right)} = g\left(y_i, u_i\right) P\left(\tau > i \,|\tau > j\,\right)$$
$$+ h\left(y_i\right) P\left(\tau > i \,|\tau > j\,\right). \quad (3.27)$$

Niech u^* oznacza optymalne sterowanie oraz przyjmujemy $u = u^* + \epsilon v$, gdzie $\epsilon \in \mathbb{R}$. Ustalamy liczbę $j \in \{0, \ldots, N-1\}$ oraz definiujemy $v : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^{l\times N}$, $v = (\mathbf{0}, ..., \mathbf{0}, \tilde{v}_j, \mathbf{0}, ..., \mathbf{0})$, $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^l$, gdzie $\tilde{v}_j : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$, $\tilde{v}_j = \mathrm{col}\,(v_1, ..., v_l)$ oraz $v_s = v_s\,(y_0, ..., y_j) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \leq s \leq l$ jest dowolną funkcją borelowską.

Jeżeli $j \in \{0, \ldots, N-1\}$ nie jest momentem zatrzymania systemu (3.1), to z równania (3.26) otrzymujemy

$$\frac{\partial}{\partial \epsilon} \bar{J}^N(u^* + \epsilon v) = \int \cdots \int P\left(\tau > j\right) \Big\langle \nabla_{u_j} g_j(y_j, u_j) +$$
$$\int \cdots \int \left(\sum_{i=j+1}^{N} \phi_{i|j}\left(y_i, u_i\right) \right) \nabla_{u_j} P_{j+1,N}\left(dy_{j+1}, \ldots, dy_N\right), \tilde{v}_j \Big\rangle$$
$$\times P\left(d\xi, dy_0, \ldots, dy_j\right). \quad (3.28)$$

Ze wzorów (2.6) oraz (1.4) mamy

$$\nabla_{u_j} P_{j+1,N}\left(dy_{j+1}, \ldots, dy_N\right) = \left(y_{j+1} - f(\xi, y_j, u_j)\right) \Sigma^{-1}(\xi, y_j)$$
$$\times \nabla_{u_j} f(\xi, y_j, u_j) P_{j+1,N}\left(dy_{j+1}, \ldots, dy_N\right). \quad (3.29)$$

Warunek konieczny optymalnego sterowania w momencie $j \in \{0, \ldots, N-1\}$ możemy przedstawić za pomocą równania

$$\frac{\partial}{\partial \epsilon} \bar{J}^N(u^* + \epsilon v)\bigg|_{\epsilon=0} = 0. \tag{3.30}$$

Podstawiając (3.29) do (3.28) oraz z warunku (3.30) otrzymujemy

$$\begin{aligned}\int \cdots \int P(\tau > j) \Bigg\langle \nabla_{u_j} g_j(y_j, u_j^*) + \int \cdots \int \Bigg(\sum_{i=j+1}^{N} \phi_{i|j}(y_i, u_i^*) \Bigg) \\ \times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right) \Sigma^{-1}(\xi, y_j) \nabla_{u_j} f(\xi, y_j, u_j^*) \\ \times P\left(dy_{j+1} ... dy_N\right), \tilde{v}_j \Bigg\rangle P\left(d\xi, dy_0, \ldots, dy_j\right) = 0.\end{aligned} \tag{3.31}$$

Ponieważ warunek (3.31) powinien być spełniony dla dowolnego $\tilde{v}_j = \mathrm{col}\,(v_1, ..., v_l)$, gdzie $v_s = v_s\,(y_0, ..., y_j) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \leq s \leq l$ są funkcjami borelowskimi mierzalnymi względem σ–ciała $\mathcal{F}_j$, zatem otrzymujemy tezę. ■

Uwaga 3.8 *Mnożąc obustronnie równanie (3.23) poprzez $P(\tau > j)$ oraz korzystając ze wzoru (3.27) otrzymujemy (3.10). Warunek konieczny optymalnego sterowania możemy również przedstawić za pomocą równania*

$$\begin{aligned}E\left(\left(\sum_{i=j+1}^{N} \phi_i(y_i, u_i^*)\right) \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_i) \nabla_u f(\xi, y_j, u_j^*) \middle| \mathcal{F}_j^\tau\right) \\ + \nabla_{u_j} g\left(y_j, u_j\right) = \mathbf{0},\end{aligned} \tag{3.32}$$

gdzie $\phi_i()$, $0 \leq i \leq N$ są określone za pomocą (3.6)–(3.7).

Wniosek 3.1 *Optymalne sterowanie systemem (3.1) z losowym horyzontem niezależnym od stanów systemu (zarówno dla regulatora planistycznego, jak i operacyjnego) wyznaczamy ze wzoru (3.10) lub ze wzoru (3.23). Do konstrukcji algorytmu wyznaczenia optymalnego sterowania wykorzystany zostanie warunek (3.10).*

Wniosek 3.2 *Warunki konieczne (3.10) i (3.23) optymalnego sterowania systemem (3.1) (przedstawione w twierdzeniach 3.1 i 3.2) również możemy wykorzystać do zadania (2.3) z ustalonym horyzontem N. W tym celu wystarczy przyjąć, że rozkład zmiennej losowej τ jest rozkładem punktowym $P(\tau = N) = 1$ oraz $P(\tau = i) = 0$ dla $i = 0, 1, 2, ..., N-1$.*

3.2.3 Wyznaczenie optymalnego sterowania

Rozwiązanie zadania optymalnego sterowania systemem (3.1) z losowym horyzontem $\tau : \Omega \to \{0, 1, ..., N\}$ polega na wyznaczeniu ciągu sterowań $\left\{u_j^*\right\}_{0 \le j \le N-1}$, natomiast bezpośrednio sterując tym systemem stosujemy sterowanie $\left\{u_j^*\right\}_{0 \le j \le \tau-1} \subset \left\{u_j^*\right\}_{0 \le j \le N-1}$. Poniżej przedstawiony zostanie algorytm wyznaczenia optymalnego sterowania systemem (3.1). Do konstrukcji ciągu optymalnych sterowań (regulatora) wykorzystane zostanie programowanie dynamiczne wstecz. Dla dowolnego $0 \le j \le N-1$ definiujemy funkcje Bellmana (patrz np. [23], [38], [44])

$$\begin{aligned} V_j^N(\xi, y_0, ..., y_j) &= \inf_{u_j} \left(\phi_j(y_j, u_j) + E\left(\sum_{i=j+1}^{N} \phi_i(y_i, u_i^*) \middle| \mathcal{F}_j^\xi \right) \right) \\ &= \phi_j(y_j, u_j^*) + E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1}) \middle| \mathcal{F}_j^\xi \right) \end{aligned} \quad (3.33)$$

z warunkiem początkowym

$$V_N^N(\xi, y_0, ..., y_N) = h(y_N)\, P(\tau = N),$$

gdzie $\phi_j(y, u)$ są dane wzorem (3.6). Jeżeli moment $j \in \{0, 1, ..., N-1\}$ nie jest horyzontem sterowania (system nie został zatrzymany w chwili j), to zgodnie z rozwinięciem (3.21) wielkość

$$E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1}) \left| \mathcal{F}_j^\tau \right.\right) = \frac{1}{P(\tau > j)} E\left(V_{j+1}^N(\xi, y_0, ..., y_{j+1}) \left| \mathcal{F}_j \right.\right)$$

reprezentuje oczekiwane koszty sterowania i dziedziczenia od momentu $j+1$ do horyzontu losowego τ.

Ze wzoru (3.33) warunek konieczny optymalności sterowania (3.10) możemy przedstawić za pomocą równania

$$\nabla_{u_j} g\left(y_j, u_j^*\right) P\left(\tau > j\right) + \iint V_{j+1}(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x)$$
$$\times x^T \sigma^T(\xi, y_j) \Sigma^{-1}(\xi, y_j) \nabla_{u_j} f(\xi, y_j, u_j^*) \gamma\left(x, \bar{0}, I_m\right) dx P\left(d\xi | \mathcal{F}_j\right) = \mathbf{0},$$

gdzie $P\left(d\xi | \mathcal{F}_j\right)$ oznacza rozkład wektora nieznanych parametrów ξ względem $\sigma-$ ciała $\mathcal{F}_j$.

Algorytm wyznaczenia optymalnego sterowania.

1. Definiujemy $V_N(\xi, y_0, ..., y_N) = h\left(y_N\right) P\left(\tau = N\right)$ oraz przyjmujemy $j = N$.

2. Kładziemy $j = j - 1$.

3. Definiujemy

$$\tilde{V}_{j+1}(\xi, y_0, ..., y_j, u_j, x) = V_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)x).$$

4. Wyznaczamy

$$Z_j(y_0, ..., y_j, u_j) = P\left(\tau > j\right) \nabla_{u_j} g\left(y_j, u_j\right) + \iint \tilde{V}_{j+1}(\xi, y_0, ..., y_j, u_j, x)$$
$$\times x^T \Sigma^{-\frac{1}{2}}(\xi, y_j) \nabla_{u_j} f(\xi, y_j, u_j) \gamma\left(x, I_m\right) dx P\left(d\xi | \mathcal{F}_j\right).$$

5. Wyznaczamy optymalne sterowanie u_j^* spełniające warunek

$$Z_j(y_0, ..., y_j, u_j^*) = \mathbf{0}.$$

6. Wyznaczamy wartości funkcji Bellmana dla momentu j

$$V_j(\xi, y_0, ..., y_j) = h\left(y_j\right) P\left(\tau = j\right) + g\left(y_j, u_j^*\right) P\left(\tau > j\right)$$
$$+ \int V_{j+1}\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x\right) \gamma\left(x, I_m\right) dx.$$

7. Jeżeli $j = 0$ to zatrzymujemy, w przeciwnym razie wracamy do punktu 2.

Uwaga 3.9 *Algorytm powyżej podaje sposób wyznaczenia optymalnego sterowania* $\{u_0^*, u_1^*, ..., u_{N-1}^*\}$ *tak jak dla przypadku gdy* $\tau(\omega) = N$. *Oczywiście sterując systemem (3.1) nie znamy wartości horyzontu (zmiennej losowej* τ*). Bezpośrednio sterując systemem (3.1) z losowym horyzontem* τ *stosujemy sterowanie* $\{u_0^*, u_1^*, ..., u_{\tau-1}^*\} \subset \{u_0^*, u_1^*, ..., u_{N-1}^*\}$.

3.3 Sterowanie dla horyzontu losowego o przeliczalnej liczbie realizacji

3.3.1 Postawienie zadania

W poprzedniej części rozdziału przedstawione zostało rozwiązanie zadania (3.3) dotyczącego sterowania systemem (3.1) z losowym horyzontem $\tau : \Omega \to \{0, 1, 2, ..., N\}$ niezależnym od stanów systemu, gdzie zbiór realizacji zmiennej losowej opisującej horyzont sterowania jest zbiorem skończonym. W tym przypadku zadanie optymalnego sterowania z losowym horyzontem zostało zastąpione zadaniem optymalnego sterowania z ustalonym horyzontem. Algorytm wyznaczenia optymalnego sterowania wykorzystuje zasadę programowania dynamicznego wstecz (tzn. polega na zastosowaniu techniki teleskopowej podczas kontrukcji funkcji Bellmana). Funkcje Bellmana zaczynamy konstruować od momentu N, gdzie wartość tej funkcji jest równa iloczynowi prawdopodobieństwa, że system zostanie zatrzymany w momencie N oraz wartości funkcji dziedziczenia. Natomiast powstaje pytanie jak wyznaczyć optymalne sterowania, jeżeli zbiór realizacji zmiennej losowej opisującej horyzont sterowania jest zbiorem przeliczalnym (patrz np. [32], [54]).

Poniżej rozważamy problem sterowania systemem stochastycznym (3.1) z losowym horyzontem $\tau : \Omega \to \mathbb{N}_0$ niezależnym od stanów systemu. Zbiór realizacji zmiennej losowej τ jest zbiorem przeliczalnym.

Wartość oczekiwana funkcjonału jakości (3.2) jest równa

$$\begin{aligned} EJ^{\tau}(u) &= E\left(P(\tau=0)\,h(y_0)+P(\tau>0)\,E\left(\sum_{i=0}^{\tau-1} g(y_i,u_i)+h(y_\tau)\middle|\,\mathcal{F}_0^{\tau}\right)\right)\\ &= E\Big(P(\tau=0)\,h(y_0)+P(\tau>0)\,E\Big(g(y_0,u_0)+P(\tau=1|\,\tau>0)\,h(y_1)\\ &\quad +P(\tau=1|\,\tau>0)\,E\left(\sum_{i=1}^{\tau-1} g(y_i,u_i)+h(y_\tau)\middle|\,\mathcal{F}_1^{\tau}\right)\Bigg|\,\mathcal{F}_0^{\tau}\Big)\Big)=\ldots \end{aligned} \tag{3.34}$$

Przyjmujemy

$$J_j^{\tau}(u)=E\left(\sum_{i=j}^{\tau-1} g(y_i,u_i)+h(y_\tau)\middle|\,\mathcal{F}_j^{\tau}\right), \tag{3.35}$$

zatem otrzymujemy

$$\begin{aligned} J_j^{\tau}(u) &= g(y_j,u_j)+E\left(P(\tau=j+1|\,\tau>j)\,h(y_{j+1})\right.\\ &\quad \left.+P(\tau>j+1|\,\tau>j)\,J_{j+1}(u)\,\middle|\mathcal{F}_j^{\tau}\right). \end{aligned} \tag{3.36}$$

Wielkość $J_j^{\tau}(u)$, $j\in\mathbb{N}_0$ oznacza oczekiwane koszty sterowania i dziedziczenia od momentu j do momentu losowego τ. Zgodnie z powyższym wartość wskaźnika jakości (3.34) jest równa

$$EJ^{\tau}(u)=E\left(P(\tau=0)\,h(y_0)+P(\tau>0)\,J_0^{\tau}(u)\right). \tag{3.37}$$

Z drugiej strony, korzystając z definicji prawdopodobieństwa warunkowego, wzór (3.34) możemy przedstawić w postaci

$$\begin{aligned} EJ^{\tau}(u) &= E\left(P(\tau=0)\,h(y_0)+P(\tau>0)\,E\left(\sum_{i=0}^{\tau-1} g(y_i,u_i)+h(y_\tau)\middle|\,\mathcal{F}_0\right)\right)\\ &= E\left(P(\tau=0)\,h(y_0)+P(\tau>0)\,g(y_0,u_0)\right.\\ &\quad +E\left(P(\tau=1)\,h(y_1)+P(\tau>1)\,E\left(\sum_{i=1}^{\tau-1} g(y_i,u_i)+h(y_\tau)\middle|\,\mathcal{F}_1\right)\middle|\,\mathcal{F}_0\right)\Bigg)\ldots \end{aligned}$$

Oczekiwany koszt sterowania $EJ^{\tau}(u)$ jest równy

$$\bar{J}^{\infty}(u) = EJ^{\tau}(u) = E\sum_{j=0}^{\infty}\phi_j(y_j, u_j), \tag{3.38}$$

gdzie

$$\phi_j(y, u) = g(y, u)\,P(\tau > j) + h(y)\,P(\tau = j). \tag{3.39}$$

Zadanie optymalnego sterowania (3.3) z losowym horyzontem τ o nieskończonej liczbie realizacji zastępujemy zadaniem optymalnego sterowania z nieskończonym horyzontem. Aby wyznaczyć optymalne sterowanie systemem (3.1) należy rozwiązać zadanie

$$\inf_{u\in U}\bar{J}^{\infty}(u). \tag{3.40}$$

Definicja 3.2 *Sterowanie* $\left(u_0^*, u_1^*, ..., u_{\tau-1}^*\right) \subset u^* = (u_0^*, u_1^*, ...)$ *nazywamy optymalnym sterowaniem dla systemu (3.1) z losowym horyzontem* $\tau : \Omega \to \mathbb{N}_0$ *o przeliczalnej liczbie realizacji, gdzie* $(u_0^*, u_1^*, ...)$ *jest rozwiązaniem zadania (3.40).*

W przypadku gdy sterujemy systemem deterministycznym bądź stochastycznym z nieskończonym horyzontem czasowym należy dodatkowo uwzględnić problem dotyczący zbieżności wskaźnika jakości. Lemat poniżej podaje ograniczenie górne dotyczące wskaźnika jakości (3.38), natomiast ograniczenie dolne szacowane jest w sposób podobny.

Lemat 3.1 *Niech fukcjonały* $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ *będą ograniczone oraz spełnione są nierówności* $g(\cdot) \le U_c < \infty$ *i* $h(\cdot) \le U_h < \infty$. *Jeżeli* τ *ma rozkład dyskretny* $p_i = P(\tau = i)$ *dla* $i \in \mathbb{N}_0$ *oraz* $0 \le p_i \le 1$, $\sum\limits_{i=0}^{\infty} p_i = 1$, *to*

$$\bar{J}^{\infty}(u) = E\sum_{j=0}^{\infty}\phi_j(y_j, u_j) \le U_c E\tau + U_h,$$

gdzie $\phi_j(y, u)$ *jest dane wzorem (3.39).*

Dowód. Jeżeli zmienna losowa $\tau:\Omega\to\mathbb{N}_0$ ma rozkład dyskretny o przeliczalnej liczbie realizacji, gdzie $p_i = P(\tau = i)\,, i\in\mathbb{N}_0$ oraz $0\le p_i\le 1$, $\sum\limits_{i=0}^{\infty} p_i = 1$, to zachodzi następująca równość

$$\begin{aligned}\sum_{j=0}^{\infty} P(\tau > j) &= \sum_{j=0}^{\infty}\sum_{i=j+1}^{\infty} p_i = \sum_{i=1}^{\infty} p_i + \sum_{i=2}^{\infty} p_i + ...\\ &= \sum_{i=1}^{\infty} ip_i = \sum_{i=0}^{\infty} ip_i = E\tau. \qquad (3.41)\end{aligned}$$

Wzór (3.41) jest również prawdziwy, w przypadku gdy zmienna losowa τ ma rozkład dyskretny o skończonej liczbie realizacji. Korzystając ze wzoru (3.41) otrzymujemy

$$\begin{aligned}\bar{J}^{\infty}(u) &= \sum_{j=0}^{\infty}\phi_j(y_j,u_j) = \sum_{j=0}^{\infty}\left(g(y_j,u_j)P(\tau>j) + h(y_j)P(\tau=j)\right)\\ &\le U_c\sum_{j=0}^{\infty}P(\tau>j) + U_h\sum_{j=0}^{\infty}P(\tau=j) = U_cE\tau + U_h.\end{aligned}$$

■

Twierdzenie pożniej podaje warunki konieczne istnienia optymalnego sterowania systemem (3.1) dla zadania (3.40).

Twierdzenie 3.3 *Załóżmy, że funkcje* $g:\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}$ *i* $h:\mathbb{R}^n\longrightarrow\mathbb{R}$ *są ciągłe, wypukłe i ograniczone, funkcje* $f:\mathbb{R}^k\times\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}^n$ *i* $g:\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}$ *są funkcjami klasy* C^1 *ze względu na zmienną sterującą* $u\in\mathbb{R}^l$ *oraz* $\det(\Sigma(\xi,y))\neq 0$ *dla* $(\xi,y)\in\mathbb{R}^k\times\mathbb{R}^n$, *gdzie* $\Sigma(\xi,y)=\sigma(\xi,y)\sigma^T(\xi,y)$. *Jeżeli* $u^*=(u_0^*,u_1^*,...)$ *jest optymalnym sterowaniem dla zadania (3.40), to dla dowolnego* $j\in\mathbb{N}_0$ *spełniona jest równość*

$$\begin{aligned}E\left(\left(\sum_{i=j+1}^{\infty}\phi_i(y_i,u_i^*)\right)\left(y_{j+1}-f(\xi,y_j,u_j^*)\right)^T\Sigma^{-1}(\xi,y_i)\nabla_u f(\xi,y_j,u_j^*)\right|\left.\mathcal{F}_j\right)\\ +\nabla_{u_j}g\left(y_j,u_j^*\right)P(\tau>j) = \mathbf{0}, \qquad (3.42)\end{aligned}$$

gdzie $\phi_j(\cdot)$ *jest dane wzorem (3.39) oraz* $\mathbf{0}=\mathrm{col}\,(0,...,0)\in\mathbb{R}^l$.

Dowód. Korzystając z własności warunkowej wartości oczekiwanej dla dowolnego $j \in \mathbb{N}_0$ funkcjonał (3.38) możemy przedstawić w postaci

$$\begin{aligned}\bar{J}^{\infty}(u) &= E\left[\sum_{i=0}^{j}\phi_i(y_i,u_i) + E\left(\sum_{i=j+1}^{\infty}\phi_i(y_i,u_i)\middle|\mathcal{F}_j^{\xi}\right)\right] \\ &= \int\left(\sum_{i=0}^{j-1}\phi_i(y_i,u_i) + h\left(y_j\right)P\left(\tau=j\right) + P\left(\tau>j\right)g\left(y_j,u_j\right)\right. \\ &\left.+\int\sum_{i=j+1}^{\infty}\phi_i(y_i,u_i)P_{j+1,\infty}\left(dy_{j+1},dy_{j+2},\ldots\right)\right)P\left(d\xi,dy_0,\ldots,dy_j\right), \quad (3.43)\end{aligned}$$

gdzie rozkłady łączne $P\left(d\xi,dy_0,\ldots,dy_j\right)$ i $P_{j+1,\infty}\left(dy_{j+1},dy_{j+2},\ldots\right)$ są dane wzorami (2.5)–(2.6).
Niech u^* oznacza optymalne sterowanie oraz przyjmujemy $u = u^* + \epsilon v$, gdzie ϵ – skalar. Ustalmy $j \in \mathbb{N}_0$ oraz definiujemy $v : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^{l\times\mathbb{N}}$, $v = (\mathbf{0},...,\mathbf{0},\tilde{v}_j,\mathbf{0},...)$, $\mathbf{0} = \mathrm{col}\,(0,...,0) \in \mathbb{R}^l$, gdzie $\tilde{v}_j : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$, $\tilde{v}_j = \mathrm{col}\,(v_1,...,v_l)$ oraz $v_s = v_s\left(y_0,...,y_j\right) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \le s \le l$ jest dowolną funkcją borelowską. Jeżeli $j \in \mathbb{N}_0$ nie jest momentem zatrzymania systemu (3.1), to z równania (3.43) otrzymujemy

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\bar{J}^{\infty}(u^*+\epsilon v) = \int\cdots\int\left\langle P\left(\tau>j\right)\nabla_{u_j}g_j(y_j,u_j) + \int\cdots\int\left(\sum_{i=j+1}^{\infty}\phi_i\left(y_i,u_i\right)\right)\right. \\ \left.\times\nabla_{u_j}P_{j+1,\infty}\left(dy_{j+1},dy_{j+2},\ldots\right),\tilde{v}_j\right\rangle P\left(d\xi,dy_0,\ldots,dy_j\right). \quad (3.44)\end{aligned}$$

Ze wzorów (2.6) oraz (1.4) mamy

$$\begin{aligned}\nabla_{u_j}P_{j+1,N}\left(dy_{j+1},\ldots,dy_N\right) = \left(y_{j+1}-f(\xi,y_j,u_j)\right)\Sigma^{-1}(\xi,y_j) \\ \times\nabla_{u_j}f(\xi,y_j,u_j)P_{j+1,\infty}\left(dy_{j+1},dy_{j+2},\ldots\right). \quad (3.45)\end{aligned}$$

Warunek konieczny istnienia optymalnego sterowania w momencie $j \in \mathbb{N}_0$ możemy przedstawić za pomocą równania

$$\left.\frac{\partial}{\partial\epsilon}\bar{J}^{\infty}(u^*+\epsilon v)\right|_{\epsilon=0} = 0. \quad (3.46)$$

Podstawiając (3.45) do (3.44) oraz korzystając z warunku (3.46) otrzymujemy

$$\int \cdots \int \Big\langle P(\tau > j) \nabla_{u_j} g_j(y_j, u_j^*) + \int \cdots \int \left(\sum_{i=j+1}^{\infty} \phi_i(y_i, u_i^*) \right) \times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right) \Sigma^{-1}(\xi, y_j) \nabla_{u_j} f(\xi, y_j, u_j^*) \times P_{j+1,\infty}(dy_{j+1}, dy_{j+2}, \ldots), \tilde{v}_j \Big\rangle P(d\xi, dy_0, \ldots, dy_j) = 0. \quad (3.47)$$

Ponieważ warunek (3.47) powinien być spełniony dla dowolnego $\tilde{v}_j = col\,(v_1, ..., v_l)$, gdzie $v_s = v_s\,(y_0, ..., y_j) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \leq s \leq l$ są funkcjami borelowskimi mierzalnymi względem σ–ciała $\mathcal{F}_j$, zatem otrzymujemy tezę. ■

Twierdzenie powyżej podaje warunki konieczne istnienia sterowania optymalnego $u^* = (u_0^*, u_1^*, ...)$, natomiast sterując systemem (3.1) z losowym horyzontem $\tau : \Omega \to \mathbb{N}_0$ należy zastosować sterowanie $(u_0^*, u_1^*, ..., u_\tau^*) \subset u^*$.

Uwaga 3.10 *Wyznaczenie optymalnego sterowania u^* polega na zastosowaniu idei programowania dynamicznego wstecz. Od strony praktycznej konstrukcja funkcji Bellmana z warunkiem początkowym dla horyzontu ∞ nie jest możliwa. Dlatego poniżej podana zostanie konstrukcja sterowania dla zadania z horyzontem skończonym, gdzie różnica wartości wskaźników jakości dla zadań ze skończonym i nieskończonym horyzontem sterowania nie przekracza $\varepsilon > 0$.*

Niech $u|_k = (u_0, u_1, ..., u_{k-1}, \mathbf{0}, \mathbf{0}, ...,)$, gdzie $u_j \in \mathbb{R}^l$ dla $0 \leq j \leq k-1$ oraz $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^l$. Bezpośrednio rozwiązanie zadania (3.40) z horyzontem nieskończonym polegające na wyznaczeniu optymalnego sterowania $u^* = (u_0^*, u_1^*, ...)$ jest nierealistyczne. Zadanie (3.40) zastępujemy zadaniem z horyzontem skończonym

$$\inf_{u|_N} \bar{J}^N\left(u|_N\right), \quad (3.48)$$

gdzie

$$\bar{J}^N\left(u|_N\right) = \sum_{j=0}^{N-1} \phi_j\left(y_j, u_j\right) + P\left(\tau = N\right) h\left(y_N\right), \quad (3.49)$$

z tym że dla ustalonego $\varepsilon > 0$ moment N, do którego podejmujemy działania sterujące, wyznaczamy w sposób następujący

$$N \geq N_\varepsilon = \min \left\{ k \in \mathbb{N}_0 : \sum_{j=k}^{\infty} P(\tau > j) \left| g(y_j, u_j) \right| + \sum_{j=k}^{\infty} P(\tau = j) \left| h(y_j) \right| < \varepsilon \right\}. \tag{3.50}$$

Rozwiązując zadanie (3.48) wyznaczymy ciąg optymalnych sterowań $\{u_i^*\}_{0 \leq i \leq N-1}$. Dodatkowo przyjmujemy, że od momentu N system (3.1) zachowuje się w sposób bierny (nie sterujemy tym systemem), tzn. $u_j = \operatorname{col}(0, ..., 0) \in \mathbb{R}^l$ dla $j \geq N$.

Zatem dla momentu N zdefiniowanego za pomocą wzoru (3.50) spełniona jest nierówność

$$P(\tau = N) \left| h(y_N) \right| < \varepsilon.$$

Definicja 3.3 *Sterowanie* $\left(u_0^*, u_1^*, ..., u_{\tau-1}^*\right) \subset u^*|_N = \left(u_0^*, u_1^*, ..., u_{N-1}^*, \mathbf{0}, ...\right)$ *nazywamy* $\varepsilon-$*optymalnym sterowaniem dla systemu (3.1) z losowym horyzontem* $\tau : \Omega \to \mathbb{N}_0$ *o przeliczalnej liczbie realizacji.*

3.3.2 Wyznaczenie $\varepsilon-$optymalnego sterowania

Rozwiązanie zadania optymalnego sterowania systemem stochastycznym (3.1) z losowym horyzontem $\tau : \Omega \to \mathbb{N}_0$ polega na wyznaczeniu ciągu sterowań $\left\{u_j^*\right\}_{j \in \mathbb{N}_0}$, natomiast sterując tym systemem stosujemy sterowanie $\left\{u_j^*\right\}_{0 \leq j \leq \tau-1} \subset \left\{u_j^*\right\}_{j \in \mathbb{N}_0}$. Do konstrukcji algorytmu wyznaczenia optymalnego sterowania systemem (3.1) wykorzystana zostanie zasada programowania dynamicznego wstecz. Dla dowolnego $j \in \mathbb{N}_0$ definiujemy funkcję Bellmana (patrz np. [23], [38])

$$\begin{aligned} V_j(\xi, y_0, ..., y_j) &= \inf_{u_j} \phi_j(y_j, u_j) + E\left(\sum_{i=j+1}^{\infty} \phi_i(y_i, u_i^*) \middle| \mathcal{F}_j^\xi \right) = \\ &= \phi_j(y_j, u_j^*) + E\left(V_{j+1}(\xi, y_0, ..., y_{j+1}) \middle| \mathcal{F}_j^\xi \right). \end{aligned} \tag{3.51}$$

Wielkość $V_j(\xi, y_0, ..., y_j)$ reprezentuje oczekiwane przyszłe koszty sterowania systemem (3.1) pod warunkiem, że znane są parametry systemu ξ. Warunek konieczny optymalnego sterowania (3.42) możemy przedstawić w postaci

$$\nabla_u \phi_j(y_j, u_j^*) + \iint V_{j+1}(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x)$$
$$\times x^T \sigma^T(\xi, y_j)\Sigma^{-1}(\xi, y_j)\nabla_u f(\xi, y_j, u_j^*)\gamma(x, I_m)\, dx P(d\xi|\, \mathcal{F}_j) = \mathbf{0}, \tag{3.52}$$

gdzie $P(d\xi|\, \mathcal{F}_j)$ oznacza rozkład wektora nieznanych parametrów ξ warunkowany na $\sigma-$ ciele $\mathcal{F}_j$.

Ze wzoru (3.51) mamy

$$|V_N(\xi, y_0, ..., y_N)| = \left| \bar{J}^\infty(u^*) - \sum_{j=0}^{N-1} \phi_j(y_j, u_j^*) \right|. \tag{3.53}$$

Dla ustalonego $\varepsilon > 0$ konstruując optymalne sterowanie dla zadania zastępczego (3.40) z nieskończonym horyzontem najpierw, zgodnie z uwagą 3.10, należy wyznaczyć najmniejszy horyzont N, dla którego spełniona jest nierówność

$$|V_N(\xi, y_0, ..., y_N)| \le \sum_{j=N}^{\infty} P(\tau > j)\, |g(y_j, u_j)| + \sum_{j=N}^{\infty} P(\tau = j)\, |h(y_j)| < \varepsilon. \tag{3.54}$$

Powyższa nierówność oznacza, że sterując systemem (3.1) od momentu $N \in \mathbb{N}_0$ oczekiwany przyszły koszt sterowania nie przekracza $\varepsilon > 0$. Zgodnie z powyższym jeżeli od momentu N nie podejmujemy działań sterujących (tzn. $u_j = \mathbf{0} = \mathrm{col}(0, ..., 0) \in \mathbb{R}^l$ dla $j \ge N$), to również suma przyszłych oczekiwanych kosztów nie przekroczy $\varepsilon > 0$. Zatem w chwili N przewidywany koszt dziedziczenia $P(\tau = N)\, h(y_N)$ również nie przekroczy ε.

Pojawia się problem wyznaczenia momentu $N \in \mathbb{N}_0$. Lematy podane poniżej są pomocne podczas oszacowania horyzontu N, dla którego spełniona jest nierówność (3.54).

Lemat 3.2 *Niech zmienna losowa τ ma rozkład dyskretny*

$$P(\tau = i) = p_i,$$

gdzie $0 \le p_i \le 1$ *dla* $i \in \mathbb{N}_0$ *oraz* $\sum_{i=0}^{\infty} p_i = 1$. *Dla dowolnego* $s \ge 0$ *zachodzi równość*

$$\sum_{j=s}^{\infty} \sum_{i=j+1}^{\infty} p_i = E\tau - s + \sum_{j=0}^{s-1} \sum_{i=0}^{j} p_i. \tag{3.55}$$

Dowód. *Z definicji rozkładu dyskretnego mamy*

$$\sum_{j=s}^{\infty} \sum_{i=j+1}^{\infty} p_i = \sum_{j=s}^{\infty} P(\tau > j) = \sum_{j=0}^{\infty} P(\tau > j) - \sum_{j=0}^{s-1} P(\tau > j).$$

Ze wzoru (3.48) otrzymujemy

$$\sum_{j=s}^{\infty} P(\tau > j) = E\tau - \sum_{j=0}^{s-1} [1 - P(\tau \le j)] = E\tau - s + \sum_{j=0}^{s-1} P(\tau \le j).$$

■

Lemat 3.3 *Załóżmy, że zmienna losowa τ ma rozkład dyskretny* $0 \le p_i \le 1$ *dla* $i \ge 0$ *oraz* $\sum_{i=0}^{\infty} p_i = 1$. *Jeżeli funkcje* $|g(\cdot)| \le U_c < \infty$ *i* $|h(\cdot)| \le U_h < \infty$, *to dla dowolnego* $s \in \mathbb{N}_0$ *spełniona jest nierówność*

$$|V_s(\xi, y_0, ..., y_s)| \le U_c \left(E\tau - s + \sum_{j=0}^{s-1} P(\tau \le j) \right) + U_h P(\tau \ge s).$$

Dowód. Bezpośrednio korzystając ze wzorów (3.39) i (3.53) otrzymujemy

$$|V_s(\xi, y_0, ..., y_s)| \le U_c \sum_{i=s}^{\infty} P(\tau > i) + U_h \sum_{i=s}^{\infty} P(\tau = i).$$

Korzystając z lematu 3.2 otrzymujemy

$$|V_s(\xi, y_0, ..., y_s)| \le U_c \left(E\tau - s + \sum_{j=0}^{s-1} P(\tau \le j) \right) + U_h \sum_{i=s}^{\infty} P(\tau = i),$$

co dowodzi tezy.[4] ■

Od strony praktycznej ograniczenia górne $U_c < \infty$ i $U_h < \infty$ dla funkcjonałów kosztów i dziedziczenia obieramy w sposób następujący:

– maksymalnie możliwy koszt dziedziczenia jest równy $h(y_0)$ oraz otrzymujemy w przypadku, gdy horyzont sterowania $\tau(\omega) = 0$, zatem przyjmujemy $h(y_0) \leq U_h < \infty$;

– zgodnie z wnioskiem 2.1 największy koszt sterowania poniesiemy w przypadku gdy $\tau(\omega) = 1$, zatem przyjmujemy $g(y_0, u_0^*) \leq U_c < \infty$, gdzie u_0^* jest rozwiązaniem zadania

$$\inf_{u_0} E\left(g(y_0, u_0) + h(y_1)\right).$$

Algorytm wyznaczenia optymalnego sterowania.

1. Ustalamy $\varepsilon > 0$ (dopuszczalny błąd dla wskaźnika jakości (3.38)).

2. Wyznaczamy horyzont sterowania N spełniający warunek (3.50) jako

$$N = \min\left\{ s \in \mathbb{N}_0 : U_c\left(E\tau - s + \sum_{j=0}^{s-1} P(\tau \leq j)\right) + U_h P(\tau \geq s) \leq \varepsilon \right\}.$$

3. Definiujemy $V_N(\xi, y_0, ..., y_N) = P(\tau = N)\, h(y_N)$ oraz przyjmujemy $j = N$.

4. Kładziemy $j = j - 1$.

5. Definiujemy

$$\tilde{V}_{j+1}(\xi, y_0, ..., y_j, u_j, x) = V_{j+1}(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)x).$$

6. Wyznaczamy

$$Z_j(y_0,...,y_j,u_j) = \nabla_u \phi_j(y_j,u_j) + \iint \tilde{V}_{j+1}(\xi, y_0,...,y_j,u_j,x)$$
$$\times x^T \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j) \gamma(x, I_m)\, dx P\left(d\xi \middle| \mathcal{F}_j\right).$$

7. Szukamy optymalnego sterowania u_j^* spełniającego warunek

$$Z_j(y_0,...,y_j,u_j^*) = \mathbf{0}.$$

8. Wyznaczamy

$$V_j(\xi, y_0,...,y_j) = \phi_j(y_j, u_j^*)$$
$$+ \int V_{j+1}\left(\xi, y_0,...,y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x\right) \gamma(x, I_m)\, dx.$$

9. Jeżeli $j = 0$, to zatrzymujemy, w przeciwnym razie wracamy do punktu 4.

Uwaga 3.11 *Porównując prawa sterowań dla systemów z losowym horyzontem o skończonej i przeliczalnej liczbie realizacji widzimy, że są one konstruowane w sposób podobny (wystarczy porównać warunki konieczne (3.23) i (3.42)). W przypadku, gdy horyzont jest zmienną losową o przeliczalnej liczbie realizacji, to zadanie optymalnego sterowania z nieskończonym horyzontem (3.40) zastępujemy zadaniem ze skończonym horyzontem (3.48), gdzie dla ustalonego $\varepsilon > 0$ dodatkowo spełniony jest warunek (3.50).*

3.4 Optymalne sterowanie liniowo-kwadratowe z losowym horyzontem

W rozdziale 2.2 przedstawiony został problem sterowania liniowo-kwadratowego w warunkach pełnej i niepełnej informacji o systemie dla ustalonego horyzontu. Dodatkowo pokazano, że horyzont ma dość duży wpływ na wartość wskaźnika jakości. W tej części rozdziału również przedstawiona zostanie problematyka sterowania liniowo-kwadratowego, z tym że główny akcent zostanie skierowany na wyznaczenie praw sterowania stochastycznym systemem liniowym z kwadaratowym wskaźnikiem jakości. Poniżej przyjmujemy, że horyzont jest losowy oraz nie zależy od stanów systemu. Dodatkowo podane zostaną różnice w sterowaniu dla przypadków, gdy horyzont jest znany i losowy (nieznany).

Niech stochastyczny system liniowy będzie określony za pomocą równania

$$y_{j+1} = Ay_j - Cu_j + d + \sigma w_{j+1} \tag{3.56}$$

dla $j \in \mathbb{N}_0$, gdzie $y_j, d \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{n\times l}$, $\sigma \in \mathbb{R}^{n\times m}$ oraz $\{w_i\}_{i\geq 1}$ oznacza ciąg niezależnych wektorów losowych o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy rodzinę σ–ciał $\mathcal{F}_j = \sigma\{y_i : i = 0, 1, ..., j\}$. Poniżej zakładamy, że $\|A\| < \infty$, $\|C\| < \infty$, $\|\sigma\| < \infty$, gdzie $\|\cdot\|$ oznacza normę macierzy $\|A\| = \max\limits_{\|x\|\leq 1} \|Ax\|$, oraz $\langle d, d\rangle < \infty$.

Klasyczne zadanie sterowania liniowo–kwadratowego polega na przeprowadzeniu w N krokach systemu (3.56) z punktu y_0 do celu $a \in \mathbb{R}^n$. Dla zadania sterowania liniowo–kwadratowego z losowym horyzontem przyjmujemy, że horyzont sterowania jest zmienną losową $\tau : \Omega \to \mathbb{N}_0$ niezależną od stanów systemu oraz $\mathcal{F}^\tau = \sigma\{\tau\}$. Dla momentów $i = 0, 1, ..., \tau - 1$ koszty sterowania systemem (3.56) są równe $u_i^T R u_i$, natomiast wartość funkcjonału dziedziczenia w chwili τ jest równa $(y_\tau - a)^T Q (y_\tau - a)$. Poniżej zakładamy, że macierze R i Q są dodatnio określone. Wskaźnik jakości jest zdefiniowany jako łączne koszty sterowania i koszty strat w momencie zatrzymania oraz jest dany wzorem

$$J^\tau(u) = E\left(\sum_{i=-1}^{\tau-1} u_i^T R u_i + (y_\tau - a)^T Q (y_\tau - a) \middle| \mathcal{F}^\tau\right), \tag{3.57}$$

gdzie $R \in \mathbb{R}^{l\times l}$, $Q \in \mathbb{R}^{n\times n}$ oraz $u_{-1} = \mathrm{col}(0, 0, ..., 0)$. Jeżeli system (3.56) zostanie zatrzymany w momencie zero (tzn. $\tau(\omega) = 0$), to koszt sterowania wynosi $u_{-1}^T R u_{-1} = 0$, natomiast koszt dziedziczenia jest równy $(y_0 - a)^T Q (y_0 - a)$. Cel sterowania polega na wyznaczeniu prawa sterowania $u^* = \left(u_0^*, ..., u_{\tau-1}^*\right)$, dla którego łączne koszty sterowania (nakłady energetyczne) wraz z końcowymi stratami (straty związane z nietrafieniem do celu) byłyby jak najmniejsze. Rozwiązać zadanie (3.3) wyznaczymy prawa sterowania $u^* = \left(u_0^*, ..., u_{\tau-1}^*\right)$.

3.4.1 Optymalne sterowanie dla losowego horyzontu o skończonej liczbie realizacji

Niech horyzont sterowania systemem (3.56) będzie określony za pomocą zmiennej losowej τ o skończonej liczbie realizacji, tzn. $\tau : \Omega \to \{0, 1, 2, ..., N\}$ oraz $P(\tau = k) = p_k$, $0 \le p_k \le 1$ dla $0 \le k \le N$, $\sum_{i=0}^{N} p_i = 1$. Najpierw oszacujemy wartość oczekiwaną wskaźnika jakości (3.57) oraz definiujemy $\bar{J}^N(u) = EJ^{\tau}(u)$. Ze wzorów (3.5)–(3.7) mamy

$$\bar{J}^N(u) = E\left(\sum_{i=0}^{N-1}\left(u_i^T R_i u_i + (y_i - a)^T Q_i (y_i - a)\right) + (y_N - a)^T Q_N (y_N - a)\right), \tag{3.58}$$

gdzie

$$R_i = RP(\tau > i) = R\sum_{i=j+1}^{N} p_i, \tag{3.59}$$

$$Q_i = P(\tau = i)\, Q = p_i Q \tag{3.60}$$

dla $0 \le i \le N$. Z (3.58) wynika, że przewidywany koszt sterowania w chwili $0 \le j \le N-1$ jest równy sumie kosztów $P(\tau > j)\, u_j^T R u_j$ (iloczyn kosztu sterowania $u_j^T R u_j$ i prawdopodobieństwa $P(\tau > j)$, że system w momencie j będzie sterowany, tzn. system zostanie zatrzymany w momencie późniejszym) oraz $P(\tau = j)\,(y_j - a)^T Q (y_j - a)$ (iloczyn kosztu dziedziczenia $(y_j - a)^T Q (y_j - a)$ i prawdopodobieństwa $P(\tau = j)$, że system zostanie zatrzymany w momencie j).

W celu wyznaczenia optymalnych praw sterowania systemem (3.56) należy rozwiązać zadanie zastępcze (3.8), gdzie wskaźnik jakości $\bar{J}^N(u)$ jest dany wzorem (3.58).

Uwaga 3.12 *W przypadku gdy horyzont sterowania jest ustalony oraz przyjmujemy $\tau = N$, to $Q_i = 0$, $R_i = R$ dla $0 \leq i \leq N-1$ oraz $Q_N = Q$. Rozwiązanie zadania z ustalonym horyzontem przedstawia twierdzenie 2.2.*

Dla dowolnego momentu $j = 0, ..., N-1$ definiujemy funkcję Bellmana

$$V_j^N(y_j) = \inf_{u_j} u_j^T R_j u_j + (y_j - a)^T Q_j (y_j - a) + E\left(V_{j+1}^N(y_{j+1}) \,|\mathcal{F}_j\right) \quad (3.61)$$

z warunkiem początkowym

$$V_N^N(y_N) = (y_N - a)^T Q_N (y_N - a). \quad (3.62)$$

Z równania (3.58) wynika, że

$$\inf_{u \in U} \bar{J}^N(u) = V_0^N(y_0). \quad (3.63)$$

Sterując systemem (3.56) z losowym horyzontem τ należy zastosować sterowanie $\left(u_0^*, ..., u_{\tau-1}^*\right) \subset u^* = \left(u_0^*, ..., u_{N-1}^*\right)$, gdzie $\bar{J}^N(u^*) = \inf_{u \in U} \bar{J}^N(u)$. Sposób wyznaczenia optymalnego sterowania u^* przedstawia twierdzenie poniżej.

Twierdzenie 3.4 *Załóżmy, że macierze $R \in \mathbb{R}^{l \times l}$ i $Q \in \mathbb{R}^{n \times n}$ są dodatnio określone, macierze R_j, $0 \leq j \leq N-1$ i Q_i, $0 \leq i \leq N$ są dane wzorami (3.59)-(3.60) odpowiednio, natomiast dla $0 \leq i \leq N-1$*

$$K_i = Q_i + A^T \left(K_{i+1} - K_{i+1}^T C \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T K_{i+1}\right) A, \quad (3.64)$$

$$\begin{aligned} L_i = A^T \big(&L_{i+1} + 2K_{i+1}d - 2Q_i a \\ &-K_{i+1}^T C \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T \left(2K_{i+1}d + L_{i+1}\right)\big), \end{aligned} \quad (3.65)$$

$$M_i = a^T Q_i a + d^T K_{i+1} d + tr\left(K_{i+1}\sigma\sigma^T\right) + d^T L_{i+1} + M_{i+1}$$
$$- \left(K_{i+1}d + \frac{1}{2}L_{i+1}\right)^T C\left(R_i + C^T K_{i+1} C\right)^{-1} C^T \left(K_{i+1}d + \frac{1}{2}L_{i+1}\right), \tag{3.66}$$

gdzie $K_N = Q_N$, $L_N = -2Q_N a$, $M_N = a^T Q_N a$ *oraz* $\bar{R}_i = \frac{1}{2}\left(R_i + R_i^T\right)$, $\bar{K}_{i+1} = \frac{1}{2}\left(K_{i+1} + K_{i+1}^T\right)$. *Jeżeli* $\det\left(\bar{R}_i + C^T \bar{K}_{i+1} C\right) \neq 0$ *dla* $0 \leq i \leq N-1$ *to optymalne sterowanie jest równe*

$$u_i^* = \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T \left(K_{i+1}\left(Ay_i + d\right) + \frac{1}{2}L_{i+1}\right). \tag{3.67}$$

Wartość funkcji Bellmana (3.61) jest dana wzorem

$$V_i^N\left(y_i\right) = y_i^T K_i y_i + y_i^T L_i + M_i \tag{3.68}$$

dla $0 \leq i \leq N$ *oraz najmniejszy oczekiwany koszt sterowania wynosi*

$$\inf_{u \in U} \bar{J}^N(u) = V_0^N\left(y_0\right).$$

Dowód. Dla momentu N wartość funkcji Bellmana wyznaczamy ze wzoru (3.62), zatem otrzymujemy

$$V_N^N\left(y_N\right) = y_N^T K_N y_N + y_N^T L_N + M_N.$$

Następnie dowiedziemy prawdziwości wzoru (3.68) dla dowolnego momentu $i \in \{0, 1, ..., N-1\}$. Zakładamy, że wzór (3.68) jest prawdziwy dla momentu $i+1$. Ze wzoru (3.61) oraz własności warunkowej wartości oczekiwanej dla momentu i otrzymujemy, że wartość funkcji Bellmana jest równa

$$V_i^N\left(y_i\right) = \inf_{u_i}\Big\{u_i^T\left(R_i + C^T K_{i+1} C\right)u_i + \left(Ay_i + d\right)^T L_{i+1} + M_{i+1}$$
$$- 2u_i^T C^T \left(K_{i+1}\left(Ay_i + d\right) + \frac{1}{2}L_{i+1}\right) + \left(y_i - a\right)^T Q_i\left(y_i - a\right)$$
$$+ \left(Ay_i + d\right)^T K\left(Ay_i + d\right) + tr\left(K_{i+1}\sigma\sigma^T\right)\Big\}. \tag{3.69}$$

Zatem, optymalne sterowanie w chwili i jest równe

$$u_i^* = \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T \left(K_{i+1} \left(Ay_i + d\right) + \frac{1}{2} L_{i+1}\right). \tag{3.70}$$

Podstawiając (3.70) do (3.69) otrzymujemy

$$\begin{aligned} V_i^N \left(y_i\right) &= a^T Q_i a + d^T K_{i+1} d + d^T L_{i+1} + M_{i+1} + tr\left(K_{i+1} \sigma \sigma^T\right) \\ &- \left(K_{i+1} d + \frac{1}{2} L_{i+1}\right)^T C \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T \left(K_{i+1} d + \frac{1}{2} L_{i+1}\right) \\ &+ y_i^T \left(Q_i + A^T K_{i+1} A - A^T K_{i+1}^T C \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T K_{i+1} A\right) y_i \\ &+ y_i^T \left(A^T \left(2K_{i+1} d + L_{i+1}\right) - 2Q_i a\right. \\ &\left. - A^T K_{i+1}^T C \left(\bar{R}_i + C^T \bar{K}_{i+1} C\right)^{-1} C^T \left(2K_{i+1} d + L_{i+1}\right)\right) \\ &= y_i^T K_i y_i + y_i^T L_i + M_i, \end{aligned}$$

gdzie K_i, L_i, M_i są dane wzorami (3.64)–(3.66). ■

Uwaga 3.13 *Twierdzenie 3.4 podaje sposób konstrukcji optymalnego sterowania dla systemu liniowego (3.56) z losowym horyzontem niezależnym od stanów systemu. Przedstawione rozwiązanie można wykorzystać dla zadania z horyzontem ustalonym. Zarówno dla horyzontu deterministycznego* N*, jak i dla horyzontu losowego* $\tau : \Omega \to \{0, 1, 2, ..., N\}$ *optymalne sterowanie* u_i^**,* $0 \le i \le N-1$ *systemem (3.56) jest dane wzorem (3.67), oczekiwany koszt sterowania wynosi* $V_0^N \left(y_0\right) = y_0^T K_0 y_0 + y_0^T L_0 + M_0$*. Do wyznaczenia ciągów* $\{K_i\}_{0 \le i \le N}$*,* $\{L_i\}_{0 \le i \le N}$*,* $\{M_i\}_{0 \le i \le N}$ *(patrz wzory (3.64)–(3.66)) dla zadania z losowym horyzontem przyjmujemy*

$$\begin{aligned} Q_i &= P\left(\tau = i\right) Q \ \text{dla}\ 0 \le i \le N, \\ R_i &= P\left(\tau > i\right) R \ \text{dla}\ 0 \le i \le N-1. \end{aligned}$$

Dla zadania z horyzontem ustalonym N *przyjmujemy* $P\left(\tau = N\right) = 1$*, zatem* $Q_N = Q$*, natomiast dla* $0 \le i \le N-1$ *macierz* Q_i *jest macierzą zerową oraz* $R_i = R$*.*

Przykład 3.3 *Niech liniowy system stochastyczny będzie określony za pomocą równania*

$$y_{i+1} = y_i - Cu_i + \sigma w_{i+1}, \tag{3.71}$$

gdzie $y_i, u_i \in \mathbb{R}^2$ *dla* $i \in \mathbb{N}_0$ *oraz* $\{w_i\}_{i\geq 1}$ *będzie ciągiem niezależnych wektorów losowych* $w_i : \Omega \to \mathbb{R}^2$ *o rozkładzie normalnym* $\boldsymbol{N}(\bar{0}, I_2)$, *gdzie* $\bar{0} = \mathrm{col}\,(0,0) \in \mathbb{R}^2$ *wektorem zerowym,* $I_2 \in \mathbb{R}^{2\times 2}$ *– macierzą jednostkową. Zadanie polega na przeprowadzeniu systemu (3.71) ze stanu* y_0 *do celu* a *w czasie losowym* $\tau : \Omega \to \{0, 1, ..., N\}$. *Poniżej zakładamy, że horyzont sterowania* τ *jest zmienną losową o rozkładzie dwumianowym z prawdopodobieństwem sukcesu* $0 \leq s \leq 1$ *oraz* $N = 10$. *Rozwiązując zadanie (3.3) dla wskaźnika jakości (3.58) reprezentującego całkowity koszt sterowania wyznaczymy optymalne sterowanie systemem (3.71).*

Z twierdzenia 3.4 wynika, że jeżeli det $\left(\bar{R}_i + C^T\bar{K}_{i+1}C\right) \neq 0$ *dla* $i = 0, 1, ..., N-1$, *to optymalne sterowanie dla momentów* $0 \leq i \leq N-1$ *jest dane wzorem*

$$u_i^* = \left(\bar{R}_i + C^T\bar{K}_{i+1}C\right)^{-1} C^T K_{i+1}\left(y_i - a\right), \tag{3.72}$$

gdzie

$$K_i = Q_i + K_{i+1} - K_{i+1}^T C \left(\bar{R}_i + C^T\bar{K}_{i+1}C\right)^{-1} C^T K_{i+1} \tag{3.73}$$

oraz $K_N = Q_N$, *natomiast macierze* R_i *i* Q_i *są dane wzorami (3.59)–(3.60),* $\bar{R}_i = \frac{1}{2}\left(R_i + R_i^T\right)$, $\bar{K}_{i+1} = \frac{1}{2}\left(K_{i+1} + K_{i+1}^T\right)$ *oraz*

$$p_i = \binom{N}{i} s^i \left(1-s\right)^{N-i}.$$

Zarówno macierze $R_i = R_i\left(s\right)$, $Q_i = Q_i\left(s\right)$, $K_i = K_i\left(s\right)$, *jak i optymalne sterowanie* $u_i^* = u_i^*\left(s\right)$ *zależą od prawdopodobieństwa sukcesu* $0 \leq s \leq 1$. *Wartości funkcji Bellmana są równe*

$$V_i^N\left(y_i, s\right) = \left(y_i - a\right)^T K_i\left(y_i - a\right) + \sum_{j=i+1}^{N} tr\left(K_j\sigma\sigma^T\right) \tag{3.74}$$

dla $0 \leq i \leq N-1$ *oraz*

$$V_N^N\left(y_N, s\right) = \left(y_N - a\right)^T K_N \left(y_N - a\right).$$

Najmniejszy oczekiwany koszt sterowania wynosi

$$\inf_{u \in U} \bar{J}^N(u) = V_0^N\left(y_0, s\right) = \left(y_0 - a\right)^T K_0 \left(y_0 - a\right) + \sum_{j=1}^{N} tr\left(K_j \sigma \sigma^T\right). \quad (3.75)$$

Wielkość $\sum_{j=1}^{N} tr\left(K_j \sigma \sigma^T\right)$ *reprezentuje dodatkowy koszt związany z występowaniem zaburzeń zewnętrznych w systemie (3.71). Poniżej przyjmujemy macierze* $R = \begin{bmatrix} 0.7 & 0.02 \\ 0.02 & 0.5 \end{bmatrix}$, $Q = \begin{bmatrix} 0.9 & 0.03 \\ 0.03 & 1.4 \end{bmatrix}$, $C = \begin{bmatrix} 1.3 & 0.7 \\ 0.2 & 2 \end{bmatrix}$, $\sigma = \begin{bmatrix} 0.5 & 0.02 \\ -0.03 & 0.9 \end{bmatrix}$, *stan początkowy oraz stan docelowy są równe* $y_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $a = \begin{bmatrix} 120 \\ 75 \end{bmatrix}$ *odpowiednio. Rysunek 3.2 przedstawia symulacje trajektorii systemu (3.71) przy zastosowaniu sterowania (3.72) dla prawdopodobieństw sukcesu* $s = 0.1, 0.2, ..., 1$. *Dla* $s = 1$ *mamy klasyczne zadanie z ustalonym horyzontem* N. *Z rysunku 3.2 widzimy, że system równomiernie pokonuje ścieżkę z punktu* y_0 *do punktu* a *(realizacje ciągu* $\{y_i\}_{0 \leq i \leq N}$ *są równomiernie rozłożone na drodze od* y_0 *do celu* a*). Jeżeli* $0 < s < 1$, *to system (3.71) większe odległości pokonuje na początku, a następnie jest dosterowywany do celu* a. *Im mniejsze prawdopodobieństwo sukcesu* s, *tym większe są nakłady energetyczne związane ze sterowaniem w momentach początkowych. Dla ustalonego* y_0 *oraz celu* a *wartości oczekiwanych kosztów sterowania* $V_0^N\left(y_0, s\right)$ *(najmniejsze wartości dla wskaźnika jakości (3.75)) dla różnych wielkości prawdopodobieństwa sukcesu* s *przedstawia rysunek 3.3.*

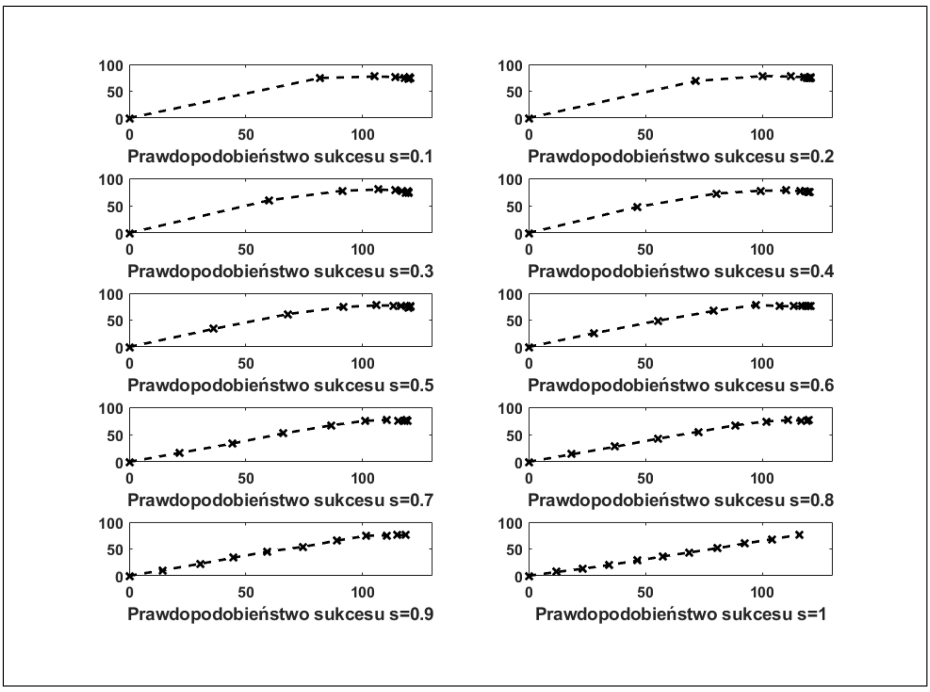
Uwaga 3.14 *Analizując zachowanie prostego systemu liniowego możemy stwierdzić:*

– *nie można bezpośrednio wykorzystać wartości sterowań wyznaczonych dla zadań z horyzontem deterministycznym do zadań z losowym hory-*

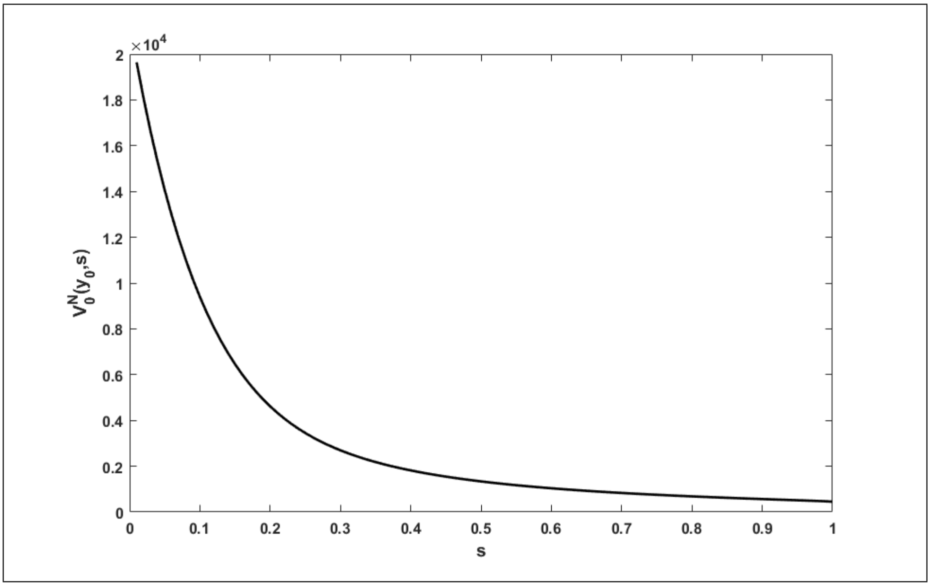
zontem (z uwagi 3.12 wynika, że prawa sterowania są konstruowane w sposób podobny, natomiast wartości sterowań są różne!);

- *wartości sterowań (podczas transformacji systemu ze stanu y_i do y_{i+1}, $0 \leq i \leq N$) dla systemów z horyzontem losowym są zdecydowanie większe w początkowych momentach sterowań niż na końcu horyzontu, natomiast dla systemów z horyzontem deterministycznym wartości sterowań są równomiernie rozłożone w czasie na całej trajektorii;*

- *dla systemów z losowym horyzontem wartości sterowań są większe do momentu $[E\tau]$ ($[\cdot]$ oznacza część całkowitą z liczby) oraz do tego momentu system w znacznym stopniu realizuje cel sterowania (prawie trafia do punktu a), natomiast po momencie $[E\tau]$ wartości sterowań są mniejsze oraz koszty są ponoszone głównie z powodu zaburzeń występujących w systemie (po przekroczeniu momentu $[E\tau]$ system praktycznie oscyluje dookoła punktu docelowego a).*



Rysunek 3.2: Symulacje zachowań systemu liniowego z losowym horyzontem o rozkładzie dwumianowym.



Rysunek 3.3: Zależność kosztu sterowania z losowym horyzontem o rozkładzie dwumianowym od prawdopodobieństwa sukcesu s.

3.4.2 Optymalne sterowanie dla losowego horyzontu o przeliczalnej liczbie realizacji

Załóżmy, że horyzont sterowania systemem (3.56) jest określony za pomocą zmiennej losowej τ o przeliczalnej liczbie realizacji, tzn. $\tau : \Omega \to \mathbb{N}_0$ oraz $P(\tau = k) = p_k$, $0 \leq p_k \leq 1$ dla $0 \leq k \leq N$, $\sum_{i=0}^{\infty} p_i = 1$. Na początek oszacujemy wartość oczekiwaną wskaźnika jakości (3.57) oraz definiujemy $\bar{J}^{\infty}(u) = EJ^{\tau}(u)$. Ze wzorów (3.38)–(3.39) mamy

$$\bar{J}^{\infty}(u) = E\left(\sum_{i=0}^{\infty}\left(u_i^T R_i u_i + (y_i - a)^T Q_i (y_i - a)\right)\right), \tag{3.76}$$

gdzie R_i, Q_i dla $0 \leq i \leq N$ są dane wzorami (3.59)–(3.60). Ze wzoru (3.76) wynika, że oczekiwany koszt sterowania w chwili $0 \leq j < \infty$ jest równy sumie kosztów $P(\tau > j)\, u_j^T R u_j$ oraz $P(\tau = j)\,(y_j - a)^T Q\,(y_j - a)$.

Rozwiązując zadanie (3.40) dla wskaźnika jakości, który jest określony wzorem (3.76), wyznaczamy prawa optymalnego sterowania systemem

(3.56). Dla dowolnego momentu $j \in \mathbb{N}_0$ definiujemy funkcję Bellmana

$$V_j^{\infty}(y_j) = \inf_{u_j} E\left(u_j^T R_j u_j + (y_j - a)^T Q_j (y_j - a) + V_{j+1}(y_{j+1}) | \mathcal{F}_j\right). \tag{3.77}$$

Zgodnie z powyższym otrzymujemy

$$\inf_{u \in U} \bar{J}^{\infty}(u) = V_0^{\infty}(y_0).$$

Sterując systemem (3.56) z losowym horyzontem $\tau : \Omega \to \mathbb{N}_0$ należy zastosować sterowanie $\left(u_0^*, u_1^* ..., u_{\tau-1}^*\right) \subset u^* = (u_0^*, u_1^*, ...)$, gdzie u^* jest rozwiązaniem zadania (3.40), zatem $\bar{J}^{\infty}(u^*) = \inf_{u \in U} \bar{J}^{\infty}(u)$.

Bezpośrednio rozwiązanie zadania (3.40) z horyzontem nieskończonym jest nierealistyczne. Zadanie(3.40) zastępujemy zadaniem z horyzontem skończonym (3.48), rozwiązując które wyznaczymy ε–optymalne sterowania. Wskaźnik jakości definiujemy w sposób następujący

$$\begin{aligned} \bar{J}^N(u|_{N-1}) =& E\left(\sum_{i=0}^{N-1}\left(u_i^T R_i u_i + (y_i - a)^T Q_i (y_i - a)\right)\right. \\ &\left. + (y_N - a)^T Q_N (y_N - a)\right). \end{aligned} \tag{3.78}$$

Aby rozwiązać zadanie (3.48) dla wskaźnika jakości (3.78) najpierw należy ustalić odpowiedni horyzont N, po osiągnięciu którego system biernie ewaluuje (nie sterujemy dalej obiektem) oraz od momentu N oczekiwany koszt sterowania $V_N^{\infty}(y_N)$ nie przekroczy ustalonego $\varepsilon > 0$. Powstaje zatem problem doboru odpowiedniego horyzontu dla zadania (3.48). Ze wzoru (3.77) otrzymujemy, że dla dowolnego $u \in U$ spełniona jest nierówność

$$\begin{aligned} (y_N - a)^T Q_N (y_N - a) &< V_N^{\infty}(y_N) \\ &\le E\left(\sum_{i=N}^{\infty}\left(u_i^T R_i u_i + (y_i - a)^T Q_i (y_i - a)\right) | \mathcal{F}_N\right). \end{aligned}$$

Dla dowolnego $\varepsilon > 0$ wyznaczamy horyzont sterowania N, dla kórego spełniona jest nierówność

$$E\left(\sum_{i=N}^{\infty}\left(u_i^T R_i u_i + (y_i - a)^T Q_i (y_i - a)\right)|\mathcal{F}_N\right) < \varepsilon.$$

Nierówność powyżej spełniona jest również i dla optymalnego sterowania u^*.

Uwaga 3.15 *Największy koszt sterowania systemem (3.56) ponosimy dla zadania z horyzontem $N = 1$,*

$$\inf_{u\in U} E\left(u^T R u + (y_1 - a)^T Q (y_1 - a)\right), \tag{3.79}$$

natomiast największy koszt dziedziczenia ponosimy dla zadania z horyzontem $N = 0$. Dla zadania (3.79) optymalne sterowanie systemem (3.56) jest równe

$$\acute{u} = \left(\bar{R} + C^T Q C\right)^{-1} C^T Q \left(A y_0 + d - a\right).$$

Niech dla zadania (3.40) ze wskaźnikiem jakości (3.76) wielkości U_c i U_h oznaczają górne ograniczenia funkcjonałów kosztów sterowania $u_i^T R_i u_i$ i dziedziczenia $(y_i - a)^T Q_i (y_i - a)$, $i \in \mathbb{N}_0$ odpowiednio . Zgodnie z uwagą powyżej wartości U_c i U_h spełniają nierówności

$$\acute{u}^T R \acute{u} \leq U_c < \infty, \tag{3.80}$$

$$(y_0 - a)^T Q (y_0 - a) \leq U_k < \infty. \tag{3.81}$$

Korzystając z lematu 3.3 dla dowolnego $\varepsilon > 0$ ustalamy horyzont sterowania $N \geq N_\varepsilon$, gdzie

$$N_\varepsilon = \min\left\{s \in \mathbb{N}_0 : U_c\left(E\tau - s + \sum_{j=0}^{s-1} P(\tau \leq j)\right) + U_h P(\tau \geq s) \leq \varepsilon\right\}. \tag{3.82}$$

Zatem dla dowolnego $N \geq N_\varepsilon$ i optymalnego sterowania u^* zachodzi nierówność

$$\bar{J}^\infty(u^*) - \bar{J}^{N-1}(u^*|_{N-1}) \leq V_N^\infty(y_N) < \varepsilon.$$

Dla ustalonego N rozwiązując zadanie zastępcze (3.48), gdzie wskaźnik jakości jest dany wzorem (3.78), otrzymujemy ciąg sterowań $\left\{u_j^*\right\}_{0\leq j\leq N-1}$. Sposób wyznaczenia ciągu sterowań optymalnych podaje twierdzenie 3.4, natomiast jako $\varepsilon-$ optymalne sterowanie dla zadania (3.40) przyjmujemy $u_\varepsilon^* = \left(u_0^*, u_1^*, ..., u_{N-1}^*, \mathbf{0}, \mathbf{0}, ...\right)$, gdzie $\mathbf{0} = \operatorname{col}(0, ..., 0) \in \mathbb{R}^l$.

Uwaga 3.16 *Dla horyzontu N spełniającego warunek (3.82) zachodzi nierówność $V_0^\infty(y_0) - V_0^N(y_0) < \varepsilon$, gdzie $V_0^\infty(y_0)$ jest dane wzorem (3.76), natomiast $V_0^N(y_0)$ spełnia (3.61).*

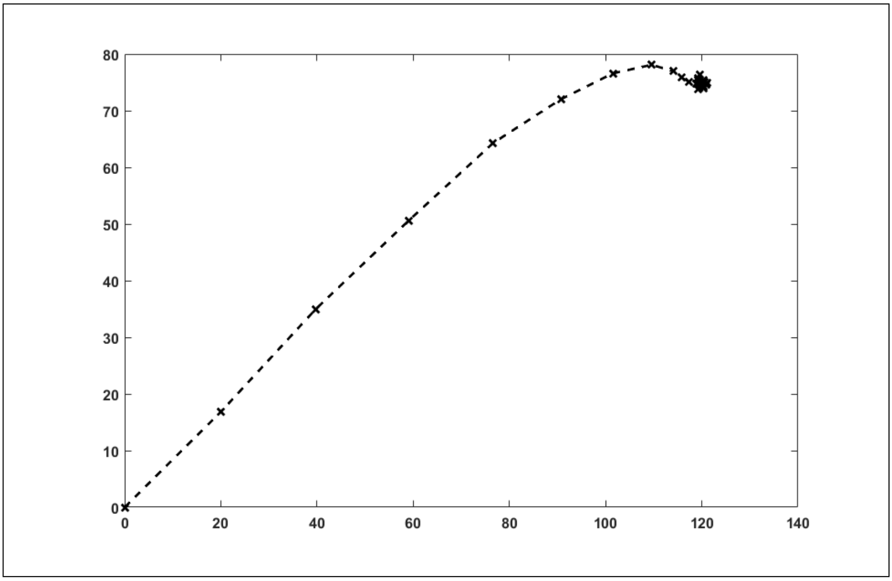
Uwaga 3.17 *Dla systemu liniowego (3.56) z losowym horyzontem o przeliczalnej liczbie realizacji $\varepsilon-$ optymalne sterowania dla momentów $0, 1, ..., N-1$ wyznaczamy jak dla zadania z losowym horyzontem $\tau : \Omega \to \{0, 1, ..., N\}$ o skończonej liczbie realizacji, gdzie dla horyzontu N spełniony jest warunek (3.82).*

Przykład 3.4 *Rozważamy problem sterowania liniowym systemem stochastycznym, który jest określony za pomocą równania (3.71). Zadanie polega na przeprowadzeniu systemu (3.71) ze stanu y_0 do celu a w czasie losowym $\tau : \Omega \to \mathbb{N}_0$, gdzie horyzont sterowania τ jest zmienną losową o rozkładzie Poissone'a z intensywnością $\lambda = 10$. Przyjmujemy, że macierze R, Q, C, σ, stan początkowy y_0 oraz cel a są identyczne jak dla przykładu 3.3. Stosownie do postawionego problemu wyznaczymy $\varepsilon-$optymalne sterowanie systemem (3.71) z losowym horyzontem o rozkładzie Poissone'a. W tym celu rozwiązujemy zadanie (3.48) oraz zakładamy, że różnica pomiędzy wskaźnikami jakości (3.76) i (3.78) nie przekracza $\varepsilon > 0$.*

Korzystając ze wzorów (3.80)–(3.81) szacujemy ograniczenia dla funkcjonału kosztu sterowania oraz funkcjonału dziedziczenia, które ustalamy na poziomie $U_c = 2407.2$ i $U_h = 21375$ odpowiednio. Następnie określamy horyzont $N \geq N_\varepsilon$, moment do którego co najwyżej podejmujemy działania sterujące, natomiast od momentu N system zachowuje się w sposób bierny ($u_j = \operatorname{col}(0, 0)$ dla $j \geq N$). Wielkość N_ε szacujemy korzystając ze wzoru (3.82). Tabela 3.1 podaje wartości N_ε w zależności od ε (maksymalna dopuszczalna różnica pomiędzy wskaźnikami jakości dla horyzontu nieskończonego *i skończonego (3.76) i (3.78) odpowiednio). Twierdzenie 3.4 podaje sposób wyznaczenia ciągu optymalnych sterowań* $\{u_i^*\}_{0\le i\le N-1}$ *dla ustalonego horyzontu* N.

Tabela 3.1: Wartości horyzontu N dla ε–optymalnych sterowań.

ε	0.1	0.01	0.001	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
N	27	29	31	33	35	37	38	40



Rysunek 3.4: Symulacje zachowań systemu liniowego (3.71) z losowym horyzontem o rozkładzie Poissone'a z intensywnością $\lambda = 10$ oraz dokładnością wskaźnika jakości $\varepsilon = 0.001$.

Poniżej przyjmujemy $\varepsilon = 0.001$ *oraz rozwiązujemy zadanie (3.48) dla* $N = 31$ *(patrz tabela 3.1). Korzystając ze wzorów (3.72)–(3.73) wyznaczamy* $\varepsilon-$ *optymalne sterowanie* $u_\varepsilon^* = (u_0^*, u_1^*, ..., u_{30}^*, \mathbf{0}, \mathbf{0}, ...)$. *Macierze* R_j *i* Q_j *dla* $0 \le j \le N$ *są określone za pomocą (3.59)–(3.60) odpowiednio oraz*

$$p_j = e^{-\lambda}\frac{\lambda^j}{j!},$$

natomiast ciąg macierzy $\{K_i\}_{0\le i\le N-1}$ *spełnia równanie (3.73) z warunkiem* *początkowym* $K_N = Q_N$. *Rysunek 3.4 przedstawia symulację trajektorii systemu (3.71).*

Identycznie jak dla problemu opisanego w przykładzie 3.3 możemy stwierdzić, że sterowania systemem (3.71) nie są równomiernie rozłożone w czasie. Dla systemów z horyzontem losowym odległości pomiędzy stanami y_i *i* y_{i+1} *są zdecydowanie większe w początkowych momentach sterowań, natomiast z upływem czasu maleją. Dla systemów z losowym horyzontem wartości sterowań są większe do momentu* $[E\tau]$ *(*$[\cdot]$ *oznacza część całkowitą z liczby) oraz po momencie* $[E\tau]$ *system praktycznie oscyluje dookoła punktu docelowego* a.

Rozdział 4

Optymalne zatrzymanie systemów sterowanych

W zadaniach optymalnego sterowania przyjęcie ustalonego horyzontu N (horyzontu niezależnego od wyników sterowania) nie prowadzi do adekwatnego modelowania sytuacji, czasami uniemożliwia osiągnięcie lepszych efektów. Problem wyznaczenia optymalnego horyzontu dla systemu sterowanego w literaturze jest nazywany zadaniem optymalnoczasowym (patrz np. [14], [105]). W rozdziale drugim niniejszej pracy (patrz rozdział 2.2.5) pokazany został wpływ horyzontu sterowania na wartość wskaźnika jakości. Horyzont sterownia został zdefiniowany jako zmienna deterministyczna. Wartość tej zmiennej wyznaczamy (wybieramy ze zbioru możliwych momentów) przed rozpoczęciem sterowania obiektem. W tym przypadku horyzont sterowania zależy od stanu początkowego.

Wydaje się, że najbardziej racjonalnym podejściem jest uzależnienie czasu sterowania systemem (obiektem) od bieżących stanów systemu. Na przykład w pracach [76], [77] horyzont sterowania jest określony jako moment, dla którego stan systemu znajduje się w sąsiedztwie pewnego celu (punktu, powierzchni), lub moment, dla którego stan systemu spełnia określone warunki. Od strony praktycznej jest to rozsądne i uzasadnione. W pracach typu self-learning, data mining, rozpoznawania obrazów, trenowania sieci neuronowych czy też w problemach sztucznej inteligencji przyjmuje się, że czas uczenia się systemu zależy od otrzymanych wyników. Sterując systemem ponosimy koszty, ale z drugiej strony obserwując reakcję systemu na impuls sterujący poznajemy jego działanie. Oczywiście im dłużej sterujemy, tym dokładniej możemy poznać zachowanie systemu, ale również ponosimy większe koszty. Powstaje zatem pytanie: jak długo powinniśmy sterować systemem? Wskazane jest aby zatrzymać system w pierwszym momencie uzyskania satysfakcjonujących rezultatów (patrz np. [8], [9], [10], [15]). Zatem żeby lepiej zrealizować cel w zadaniach optymalnego sterowania należy uwzględnić możliwość zatrzymania systemu (procesu sterowania systemem). Wobec powyższego w każdym momencie oprócz wyznaczenia optymalnego sterowania należy również podjąć decyzję odnośnie kontynuowania bądź zatrzymania sterowania tym systemem. Zadania tego typu nazywamy zadaniami optymalnego sterowania ze stopowaniem systemu stochastycznego. Oprócz zadania optymalnego sterowania równolegle należy rozwiązywać zadanie optymalnego stopowania. Podobnie jak w rozdziale 3, również i w tym przypadku horyzont sterowania nie jest znany (losowy), nie mniej jednak w rozważanym przypadku horyzont sterowania zależy od stanów systemu.

Problem optymalnego zatrzymania procesów losowych zarówno z czasem ciągłym, jak i z czasem dyskretnym jest dobrze znany w probabilistyce, finansach, teorii gier, teorii sterowania np. [22], [31], [33], [45], [53], [56], [70], [102], [122]. Problem optymalnego stopowania występuje również w zadaniach sterowania systemami stochastycznymi np. [39], [42], [43], [46]. Zadanie polega na wyznaczeniu momentu zatrzymania systemu, dla którego wskaźnik jakości osiąga wartość ekstremalną. Zatem w każdym momencie należy odpowiedzieć na pytanie: czy system powinien być dalej sterowany? W przypadku odpowiedzi twierdzącej należy dodatkowo podać wielkość sterowania systemem. W tym rozdziale przedstawione zostaną dwa podejścia rozwiązania wyżej omówionego problemu. Jedno z nich polega na możliwości stopowania (zatrzymania) procesu przy pomocy sterowania. To podejście zostało zaproponowane w pracach [8], [9], [61]. Podejście drugie oparte jest na konstrukcji koperty Snell'a (patrz np. [51], [52], [64], [65], [122], [123]).

4.1 Postawienie zadania

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną. Na przestrzeni probabilistycznej definiujemy ciąg niezależnych wektorów losowych $\{w_i\}_{i\geq 1}$ o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, gdzie $w_i : \Omega \to \mathbb{R}^m$ dla $i \geq 1$, $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Niech $\xi : \Omega \to \mathbb{R}^k$ oraz $y_0 : \Omega \to \mathbb{R}^n$ będą wektorami losowymi o rozkładach $P\,(d\xi)$ i $P\,(dy_0)$ odpowiednio. Wektor ξ oznacza nieznane parametry systemu, natomiast y_0 oznacza stan początkowy systemu (jeżeli y_0 jest wektorem deterministycznym, to wystarczy przyjąć, że y_0 ma rozkład punktowy). Zakładamy, że wszystkie wymienione objekty są stochastycznie niezależne.

Rozważamy problem sterowania systemem stochastycznym o równaniu stanu

$$y_{i+1} = f(\xi, y_i, u_i) + \sigma(\xi, y_i)w_{i+1}, \tag{4.1}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $f : \mathbb{R}^k\times\ \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ oraz $\sigma : \mathbb{R}^k\times$ $\mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$. Zakładamy, że funkcje f, σ są ciągłe względem wszystkich argumentów. Wektor $u_j \in \mathbb{R}^l$ mierzalny względem σ–ciała $\mathcal{F}_j$ oznacza działanie sterujące w chwili $j \in \mathbb{N}_0$, natomiast ciąg $u = \{u_t\}_{t\geq 0}$ – sterowanie dopuszczalne. Klasę sterowań dopuszczalnych oznaczamy przez U. Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy niemalejące rodziny $\sigma-$ciał $\{\mathcal{F}_i\}_{i\geq 0}$ oraz $\left\{\mathcal{F}_i^\xi\right\}_{i\geq 0}$, gdzie $\mathcal{F}_i = \sigma\,(y_0) \vee \sigma\,\{w_s : s = 1, 2, ..., i\}$, $\mathcal{F}_i^\xi = \mathcal{F}_i \vee \sigma\,(\xi)\,.$

Niech $\tau : \Omega \to \{0, 1, ..., N\}$ będzie momentem Markowa względem rodziny σ–ciał $\mathcal{F}_j$, $0 \leq j \leq N$ oraz $P\,(\tau \leq N) = 1$. Zmienna losowa τ oznacza moment zatrzymania systemu (4.1). Zatem horyzont sterowania systemem jest losowy oraz nie przekracza czasu końcowego N. W dalszej części monografii pojęcie zatrzymanie systemu traktujemy jako zakończenie procesu sterowania tym systemem. W rozważanym przypadku zatrzymujemy proces stochastyczny $\{y_j\}_{0\leq j\leq N}$, realizacje którego zależą nie tylko od parametrów ξ, ale i od sterowań (tzn. $y_{j+1} = y_{j+1}\,(\xi, u_0, ..., u_j)$ dla $j \neq 0$). Dla dowolnego $0 \leq j \leq N$ klasę momentów Markowa o realizacjach w zbiorze $\{j, j+1, ..., N\}$ oznaczamy przez $\mathcal{T}\,(j, N)$. Klasę wszystkich momentów

Markowa oznaczamy przez $\mathcal{T} = \mathcal{T}(0, N)$.

Niech funkcje $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ i $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ oznaczają wartości strat (kosztów) i dziedziczenia odpowiednio. Zakładamy, że są one ciągłe względem swoich argumentów oraz ograniczone. Dla dowolnego horyzontu $t \in \mathbb{N}_0$ wskaźnik jakości definiujemy jako sumę strat (lub kosztów) dla momentów $0 \leq j \leq t-1$ (sterując system (4.1) ponosimy straty (koszty) wielkości $g(y_j, u_j)$ w chwilach $0 \leq j \leq t-1$) oraz dziedziczenia $h(y_t)$ dla momentu t. Całkowity koszt sterownia $J(u,t)$ systemem (4.1) jest równy

$$J(u,t) = E\left(\sum_{i=0}^{t-1} g(y_i, u_i) + h(y_t)\right). \tag{4.2}$$

Głównym celem sterowania jest minimalizacja całkowitych kosztów sterowania (4.2). System (4.1) może być zatrzymany w dowolnym momencie nie przekraczającym czasu końcowego N. Aby optymalnie sterować tym systemem należy rozwiązać zadanie

$$\inf_{(u,\tau)\in U\times\mathcal{T}} J(u,\tau). \tag{4.3}$$

Rozwiązanie zadania (4.3) polega na wyznaczeniu optymalnego momentu stopu τ oraz sterowania $u^* = \left(u_0^*, ..., u_{\tau-1}^*\right)$, dla którego infimum jest osiągnięte. Problem optymalnego sterowania ze stopowaniem systemu stochastycznego jest złożony, wymaga rozwiązywania dwóch zadań jednocześnie: zadania optymalnego sterowania i zadania optymalnego stopowania. Zatem w każdym momencie decyzyjnym należy odpowiedzieć na pytania: czy system powinien być sterowany oraz jakie jest optymalne sterowanie.

Poniżej przedstawione zostaną dwa sposoby rozwiązania zadania (4.3). Jeden z nich polega na modyfikacji wskaźnika jakości poprzez zastąpienie zmiennej określającej moment stopu zmienną sterującą. Jest to podejście zaproponowane przez T. Banka i E. Kozłowskiego (patrz [8], [9], [15]). Rozwiązanie tak zmodyfikowanego zadania nie wymaga wykorzystania narzędzi dotyczących optymalnego zatrzymania procesów losowych (klasyczna, czysto probabilistyczna metoda nie jest rozważana w tym przypadku).

Inny sposób rozwiązania zadania (4.3) polega na wykorzystaniu zagadnienia optymalnego stopowania procesów stochastycznych. Jest to klasyczne podejście proponowane w matematycznych monografiach. Podejście to wymaga konstrukcji koperty (otoczki) Snell'a jako najmniejszego nadmartyngału (największego podmartyngału) dominującego zatrzymywany proces. Dowodzi się, że optymalny moment stopu jest pierwszym momentem zrównania się procesów – dominującego (zdominowanego) i stopowanego. W odróżnieniu od klasycznej teorii optymalnego stopowania procesu stochastycznego w zadaniu optymalnego sterowania ze stopowaniem (4.3) należy stopować nie jeden proces, lecz rodzinę procesów (indeksowanych sterowaniem). Oznacza to konstrukcję nie jednej koperty Snell'a, lecz całej ich rodziny. Dopiero z tej rodziny należy wybrać otoczkę. Czas optymalnego zatrzymania systemu (4.1) jest to moment przecięcia wskaźnika jakości (4.2) z wybraną otoczką.

4.2 Optymalne zatrzymanie systemu przy pomocy sterowania

4.2.1 Zamiana stopowania na sterowanie

Poniżej wprowadzamy ideę zamiany stopowania na sterowanie oraz sformułujemy problem równoważny dla zadania (4.3). Metoda stopowania procesu przy pomocy sterowania polega na rozszerzeniu wektora sterowania o zmienną binarną. Modyfikujemy pojęcie sterowania jak następuje. $\mathcal{F}_j$ mierzalny proces $\mathbf{u} = (\theta, u)$ nazywamy rozszerzonym sterowaniem dopuszczalnym, gdzie $\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$, dla $0 \le i \le N-1$ oraz $u = (u_0, u_1, ..., u_{N-1})$. W każdym momencie i decyzja $\theta_i\left(y_0, y_{1,...,}y_i\right) : \mathbb{R}^{n\times(i+1)} \longrightarrow \{0,1\}$ o zatrzymaniu systemu (4.1) zależy od trajektorii $y_0, y_{1,...,}y_i$ tego systemu do momentu i. Wektor $\theta : \Omega \to \{0,1\}^N$ oznacza sterowanie decyzyjne dotyczące zatrzymania systemu (4.1), natomiast $u-$ sterowanie dopuszczalne. Jeżeli w momencie i kontynuujemy sterowanie systemem (4.1), to przyjmujemy $\theta_i = 1$. Jeżeli w momencie i zatrzymujemy system (4.1), to przyjmujemy

$\theta_i = 0$ oraz $\theta_j = 0$ dla $i < j \leq N-1$. Klasę sterowań dopuszczalnych (rozszerzonych) oznaczamy przez $\mathbf{U}$.

Zmienną opisującą moment stopu τ w zadaniu (4.3) zastępujemy wektorem sterującym θ w sposób następujący: jeżeli $\tau(\omega) = k$ to przyjmujemy $\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$, gdzie

$$\theta_i = \begin{cases} 1, & \text{dla } 0 \leq i \leq k-1 \\ 0, & i \geq k. \end{cases}$$

Niech

$$\psi_j^i(\theta) \triangleq \begin{cases} \prod\limits_{s=j}^{i} \theta_s, & \text{dla } i \geq j, \\ 1, & \text{dla } i < j \end{cases} \tag{4.4}$$

dla $0 \leq i \leq j < N$.

Uwaga 4.1 *Jeżeli system (4.1) jest sterowany do momentu* $0 \leq j < N-1$ *oraz zatrzymany w momencie* $j+1$*, to*

$$\theta_i = \psi_0^i(\theta) = \begin{cases} 1, & \textit{dla } 0 \leq i \leq j, \\ 0, & \textit{dla } j < i < N. \end{cases}$$

Koszt sterowania w momencie $i \in \{0, 1, ..., N-1\}$ jest równy

$$\phi_i(\theta_i, y_i, u_i) = \theta_i g(y_i, u_i) + (1-\theta_i) h(y_i). \tag{4.5}$$

Oznacza to, że jeżeli w momencie $0 \leq i < N$ zmienna $\theta_i = 1$, to ponosimy koszt sterowania wielkości $g(y_i, u_i)$, w przeciwnym razie ($\theta_i = 1$) ponosimy tylko koszt dziedziczenia $h(y_i)$. Analizując koszt sterowania systemem (4.1) definiujemy wskaźnik jakości dla całego horyzontu N. Wobec powyższego oczekiwany całkowity koszt sterowania jest równy

$$J(\mathbf{u}) = E\left(\sum_{i=0}^{N-1} \phi_i(\theta_i, y_i, u_i) \psi_0^{i-1}(\theta) + h(y_N) \psi_0^{N-1}(\theta)\right), \tag{4.6}$$

gdzie z (4.4) mamy $\psi_0^{-1}(\theta) = 1$. W rozważanym przypadku horyzont losowy

τ w zadaniu (4.3) zostaje zastąpiony poprzez wektor sterowań decyzyjnych $\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$. Zmianie ulega również i wskaźnik jakości (4.2). Modyfikacja polega na wymnożeniu w każdym momencie $0 \leq i < N$ kosztu sterowania $\phi_i\,()$ (zależnego od zmiennej decyzyjnej θ_i) danego wzorem (4.5) przez $\psi_0^{i-1}(\theta)$ oraz funkcji dziedziczenia h poprzez $\psi_0^{N-1}(\theta)$. Po dokonanej modyfikacji wskaźnika (4.6) rozważamy klasyczne zadanie z ustalonym horyzontem N. Aby wyznaczyść optymalne prawa sterowania dla systemu (4.1), gdzie w każdym momencie $0 \leq i < N$ dodatkowo podejmujemy decyzję odnośnie kontynuacji bądź zaprzestania sterowania tym systemem, należy rozwiązać zadanie

$$\inf_{\mathbf{u} \in \mathbf{U}} J(\mathbf{u}). \tag{4.7}$$

4.2.2 Wyznaczenie optymalnego sterowania i momentu zatrzymania

Zadanie (4.7) jest to zadanie sterowania adaptacyjnego, które jest połączone z problemem stopowania w momencie losowym nie większym niż N. Jeżeli system zostanie zatrzymany w momencie N, wtedy wektor sterowania decyzyjnego jest równy $\theta = (1, 1, ..., 1)$ (tzn. $\theta_i = 1$ dla każdego $i \in \{0, 1, ..., N-1\}$ oraz $\min\{i : \theta_i = 0,\ 0 \leq i \leq N-1\} = \emptyset$). Zgodnie z konstrukcją wskaźnika jakości (4.6) moment zatrzymania systemu (4.1) jest równy $\tilde{\tau} = \min(\min\{i : \theta_i = 0,\ 0 \leq i \leq N-1\}, N)$. Dla zadania (4.7) podane zostaną warunki konieczne optymalności oraz algorytm prowadzący do wyznaczenia optymalnego sterowania i momentu stopu.

Uwaga 4.2 *Zastępując problem (4.3) zadaniem (4.7) dokonujemy zamiany problemu sterowania adaptacyjnego z losowym horyzontem na równoważny problem sterowania adaptacyjnego z ustalonym horyzontem (ale dla zmodyfikowanego wskaźnika jakości), z tym że sterowanie dopuszczalne rozszerzamy o zmienną decyzyjną dotyczącą zatrzymania systemu.*

Twierdzenie 4.1 *Niech funkcje $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ i $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ są ciągłe, wypukłe i ograniczone, funkcje $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ i $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ są funkcjami klasy C^1 ze względu na zmienną sterującą*

$u \in \mathbb{R}^l$ *oraz* $\det(\Sigma(\xi, y)) \neq 0$ *dla* $(\xi, y) \in \mathbb{R}^k \times \mathbb{R}^n$, *gdzie* $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$. *Jeżeli* $\mathbf{u}^* = (\theta^*, u^*)$ *jest rozszerzonym optymalnym sterowaniem systemem (4.1) dla zadania (4.7), to dla dowolnego* $j \in (0, 1, ..., N-1)$:

1. *warunekiem koniecznym optymalnego sterowania w momencie* j *jest spełnie równania*

$$\nabla_{u_j} g_j(y_j, u_j^*) + E\left\{\left(\sum_{i=j+1}^{N-1} \phi_i\left(\theta_i^*, y_i, u_i^*\right)\psi_{j+1}^{i-1}\left(\theta^*\right) + h\left(y_N\right)\psi_{j+1}^{N-1}\left(\theta^*\right)\right)\right.$$
$$\left.\times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)\Sigma^{-1}(\xi, y_j)\nabla_u f\left(\xi, y_j, u_j^*\right)|\mathcal{F}_j\right\} = \mathbf{0}, \quad (4.8)$$

gdzie $\mathbf{0} = \operatorname{col}(0, ..., 0) \in \mathbb{R}^l$ *oraz* $\psi_j^i(\theta)$ *dla* $0 \leq i \leq j < N$ *jest dane wzorem (4.4);*

2. *wartość zmiennej decyzyjnej* θ_j^* *jest równa*

$$\theta_j^* = \begin{cases} 1, & \text{dla } Z_j^N(y_0, ..., y_j) + g_j(y_j, u_j^*) \leq h(y_j), \\ 0, & \text{w przeciwnym razie,} \end{cases} \quad (4.9)$$

gdzie

$$Z_j^N(y_0, ..., y_j) = E\left(\sum_{i=j+1}^{N-1} \phi_i\left(\theta_i^*, y_i, u_i^*\right)\psi_{j+1}^{i-1}\left(\theta^*\right) + h\left(y_N\right)\psi_{j+1}^{i-1}\left(\theta^*\right)\middle|\mathcal{F}_j\right). \quad (4.10)$$

Dowód. Korzystając z własności warunkowej wartości oczekiwanej dla dowolnego $j \in \{0, 1, ..., N-1\}$ funkcjonał (4.6) możemy przedstawić w postaci

$$J(\mathbf{u}) = E\left(\sum_{i=0}^{j} \phi_i\left(\theta_i, y_i, u_i\right)\psi_0^{i-1}(\theta) + \psi_0^j(\theta)\right.$$
$$\left.\times E\left(\sum_{i=j+1}^{N-1} \phi_i\left(\theta_i, y_i, u_i\right)\psi_j^{i-1}(\theta) + h\left(y_N\right)\psi_j^{N-1}(\theta)\middle|\mathcal{F}_j\right)\right). \quad (4.11)$$

Zatem

$$
\begin{aligned}
J(\mathbf{u}) = & \int \left(\sum_{i=0}^{j} \phi_i \left(\theta_i, y_i, u_i\right) \prod_{l=0}^{i-1} \theta_l \right) P\left(d\xi, dy_0, .., dy_j\right) \\
& + \int \left(\prod_{l=0}^{j} \theta_l \int \left(\sum_{i=j+1}^{N-1} \phi_i \left(\theta_i, y_i, u_i\right) \prod_{l=j+1}^{i-1} \theta_l + h\left(y_N\right) \prod_{l=0}^{N-1} \theta_l \right) \right. \\
& \left. \times P\left(dy_{j+1}, .., dy_N\right) \right) P\left(d\xi, dy_0, .., dy_j\right), \qquad (4.12)
\end{aligned}
$$

gdzie rozkłady łączne $P\left(d\xi, dy_0, \ldots, dy_j\right)$ i $P_{j+1,N}\left(dy_{j+1}, \ldots, dy_N\right)$ są dane wzorami (2.5)–(2.6).

Niech $\mathbf{u}^* = (\theta^*, u^*)$ oznacza sterowanie optymalne, gdzie u^* – bezpośrednie sterowanie systemem, θ^*- optymalne sterowanie decyzyjne dotyczące zatrzymania systemu (4.1). Ustalamy liczbę $0 \leq j < N-1$ reprezentującą moment w którym podejmujemy decyzję. Przyjmujemy $u = u^* + \epsilon v$, gdzie $\epsilon > 0$, natomiast $v : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^{l\times N}$, $v = (\mathbf{0}, ..., \mathbf{0}, \tilde{v}_j, \mathbf{0}, ..., \mathbf{0})$, $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^l$, gdzie $\tilde{v}_j : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$, $\tilde{v}_j = \mathrm{col}\,(v_1, ..., v_l)$ oraz $v_s = v_s\left(y_0, ..., y_j\right) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \leq s \leq l$ są dowolną funkcją borelowską. Warunek konieczny istnienia optymalnego sterowania systemem (4.1) dla momentu $j \in \{0, 1, , ..., N-1\}$ jest dany wzorem

$$
\frac{\partial}{\partial \epsilon} J(\theta^*, u^* + \epsilon v) \bigg|_{\epsilon=0} = 0. \qquad (4.13)
$$

Korzystając z rozwinięcia (4.12) otrzymujemy

$$
\begin{aligned}
\frac{\partial}{\partial \epsilon} J(\theta^*, u^* + \epsilon v) = & \int \prod_{l=0}^{j-1} \theta_l^* \left\langle \theta_j^* \nabla_{u_j} g_j(y_j, u_j^*) + \theta_j^* \int \left(\sum_{i=j+1}^{N-1} \phi_i \left(\theta_i^*, y_i, u_i^*\right) \right. \right. \\
& \left. \left. \times \prod_{l=j+1}^{i-1} \theta_l^* + h\left(y_N\right) \prod_{l=j+1}^{N-1} \theta_l^* \right) \nabla_u P\left(dy_{j+1}, .., dy_N\right), \tilde{v}_j \right\rangle P\left(d\xi, dy_0, .., dy_j\right).
\end{aligned}
\qquad (4.14)
$$

Ze wzorów (2.5)-(2.6) mamy

$$\nabla_u P\left(dy_{j+1},..,dy_N\right) = \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)\Sigma^{-1}(\xi, y_j) \\ \times \nabla_u f(\xi, y_j, u_j^*) P\left(dy_{j+1},..,dy_N\right). \tag{4.15}$$

Podstawiając (4.15) do (4.14) oraz korzystając z warunku (4.13) otrzymujemy

$$\int \Big\langle \nabla_{u_j} g_j(y_j, u_j^*) + \int \left(\sum_{i=j+1}^{N-1} \phi_i\left(\theta_i^*, y_i, u_i^*\right)\psi_{j+1}^{i-1}\left(\theta^*\right) + h_N\left(y_N\right)\psi_{j+1}^{N-1}\left(\theta^*\right)\right) \\ \times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)\Sigma^{-1}(\xi, y_j)\nabla_u f(\xi, y_j, u_j^*) \\ \times \theta_j^* P\left(dy_{j+1},..,dy_N\right), \tilde{v}_j\Big\rangle \psi_0^{i-1}\left(\theta^*\right) P\left(d\xi, dy_0,..,dy_j\right) = 0. \tag{4.16}$$

Warunek (4.16) powinien być spełniony dla dowolnego $\tilde{v}_j = col\left(v_1, ..., v_l\right)$, gdzie $v_s = v_s\left(y_0, ..., y_j\right) : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^l$ dla $1 \le s \le l$ oraz θ_j^* są funkcjami borelowskimi mierzalnymi względem σ–ciała $\mathcal{F}_j$. Wobec powyższego, jeżeli w momencie $j \in \{0, 1, ..., N-1\}$ kontynuujemy sterowanie systemem (4.1) oraz $\theta_j^* = 1$, to optymalne sterowanie u_j^*, $0 \le j \le N-1$ spełnia warunek (4.8). Jeżeli dla momentu $j \in \{0, 1, ..., N-1\}$ spełniony jest warunek

$$J\left(\left(1, ..., 1, 1, \theta_{j+1}^*, ..., \theta_{N-1}^*\right), u^*\right) \ge J\left(\left(1, ..., 1, 0, \theta_{j+1}^*, ..., \theta_{N-1}^*\right), u^*\right), \tag{4.17}$$

to zatrzymujemy system (4.1) oraz przyjmujemy $\theta_j^* = 0$, ponieważ w przypadku kontynuacji sterowania zostaną poniesione większe koszty niż w przypadku zatrzymania systemu. Nierówność (4.17) oznacza, że w momencie j nie ma szans na poprawę wskaźnika jakości w przyszłości. Jeżeli nierówność (4.17) nie jest spełniona, to przyjmujemy $\theta_j^* = 1$ oraz kontynuujemy sterowanie systemem (4.1). Warunek (4.9) bezpośrednio wynika z (4.17). ■

Uwaga 4.3 *Od strony praktycznej dla każdego momentu* $0 \le j < N$ *wyznaczamy najpierw sterowanie* u_j^**, które spełnia warunek (4.8), a następnie ze wzoru (4.9) wyznaczamy* θ_j^**. Jeżeli w momencie* j *okazuje się że* $\theta_j^* = 1$*, to stosujemy sterowanie* u_j^**, w przeciwnym razie nie sterujemy obiektem.*

4.2.3 Algorytm wyznaczenia optymalnego sterowania i momentu zatrzymania

Rozwiązanie zadania (4.7) optymalnego sterowania z zatrzymaniem systemu (4.1) polega na wyznaczeniu ciągu rozszerzonych sterowań $\{(\theta_i^*, u_i^*)\}_{0 \le i < N}$, gdzie optymalne sterowanie u_i^* spełnia równanie (4.8), natomiast wartość zmiennej decyzyjnej θ_i^* odnośnie kontynuacji sterowania jest dana wzorem (4.9). Poniżej przedstawiony zostanie algorytm wyznaczenia optymalnych sterowań systemem (4.1) oraz decyzji dotyczących zatrzymania tego systemu. Do wyznaczenia sterowań i momentu zatrzymania wykorzystana zostanie zasada programowania dynamicznego wstecz. Dla dowolnego $j \in \{0, ..., N-1\}$ funkcjonał (4.6) możemy przedstawić w postaci

$$J(\theta, u) = E\left(\sum_{i=0}^{j-1} \phi_i(\theta_i, y_i, u_i)\, \psi_0^{i-1}(\theta) + \psi_0^{j-1}(\theta)\Big(h(y_j)(1-\theta_j) + \theta_j\right.$$
$$\left.\times\left(g(y_j, u_j) + E\left(\sum_{i=j+1}^{N-1} \phi_i(\theta_i, y_i, u_i)\, \psi_{j+1}^{i-1}(\theta) + h(y_N)\, \psi_{j+1}^{N-1}(\theta) \,\middle|\, \mathcal{F}_j\right)\right)\Big)\right).$$

Zgodnie z twierdzeniem 4.1 rozszerzone sterowanie $\mathbf{u}^* = \left\{\left(\theta_j^*, u_j^*\right)\right\}_{0 \le j < N-1}$ jest optymalne, jeżeli dla dowolnego $0 \le j \le N-1$ sterowanie u_j^* i decyzja θ_j^* spełniają warunki (4.8)–(4.9). Dla dowolnego $0 \le j \le N-1$ definiujemy funkcję Bellmana

$$W_j^N(\xi, y_0, ..., y_j) = E\left(\sum_{i=j}^{N-1} \phi_i(\theta_i^*, y_i, u_i^*)\, \psi_{j+1}^{i-1}(\theta^*) + h_N(y_N)\, \psi_{j+1}^{N-1}(\theta^*) \,\middle|\, \mathcal{F}_j^\xi\right), \tag{4.18}$$

która reprezentuje oczekiwane koszty sterowania systemem (4.1) od momentu j, w przypadku jeżeli znane są parametry systemu. Korzystając z własności warunkowej wartości oczekiwanej dla $0 \le j \le N-1$ mamy

$$W_j^N(\xi, y_0, ..., y_j) = \inf_{\substack{u_j \\ \theta_j \in \{0,1\}}} \phi_j(\theta_j, y_j, u_j) + \theta_j E\left(W_{j+1}^N(\xi, y_0, ..., y_{j+1}) \,\middle|\, \mathcal{F}_j^\xi\right), \tag{4.19}$$

gdzie

$$W_N^N(\xi, y_0, ..., y_N) = h(y_N). \tag{4.20}$$

Uwaga 4.4 *Niech* $\left\{V_j^k(\xi, y_0, ..., y_j)\right\}_{0\le j\le k}$ *będzie ciągiem funkcji Bellmana dla zadania (2.3) z ustalonym horyzontem k, które są zdefiniowane za pomocą równań (2.14)–(2.13). Wielkość* $V_j^k(\xi, y_0, ..., y_j)$ *zależy od momentu* j *do końca horyzontu sterowań* k *oraz dla* $0 \le j \le N-1$ *przyjmujemy* $V_j^j(\xi, y_0, ..., y_j) = h(y_j)$. *Dla dowolnego momentu* $j \in \{0, 1, ..., N\}$ *wartość funkcji Bellmana* $W_j^N(\xi, y_0, ..., y_j)$ *jest równa*

$$W_j^N(\xi, y_0, ..., y_j) = \min\Big(V_j^j(\xi, y_0, ..., y_j), ..., V_j^N(\xi, y_0, ..., y_N)\Big). \tag{4.21}$$

Uwaga 4.5 *Wielkość*

$$\bar{W}_j^N(y_0, ..., y_j) = E\left(W_j^N(\xi, y_0, ..., y_j)\middle|\, \mathcal{F}_j\right)$$

dla $0 \le j \le N$ *reprezentuje najmniejsze oczekiwane koszty sterowania od momentu* j *do zatrzymania systemu (4.1). Ciąg* $\left\{\bar{W}_j^N(y_0, ..., y_j)\right\}_{0\le j\le N}$ *określa (wyznacza) pewną otoczkę najmniejszych kosztów. Zgodnie z twierdzeniem 4.1 moment zatrzymania systemu (4.1) jest równy*

$$\tilde{\tau} = \min\left\{0 \le j \le N: \ \bar{W}_j^N(y_0, ..., y_j) = h(y_j)\right\}.$$

Algorytm wyznaczenia sterowania i momentu zatrzymania

1. Definiujemy
$$W_N^N(\xi, y_0, ..., y_N) = h(y_N).$$
oraz przyjmujemy $j = N$.

2. Kładziemy
$$j = j - 1.$$

3. Definiujemy
$$\tilde{W}_{j+1}^N(\xi, y_0, ..., y_j, u_j, x) = W_{j+1}(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)x).$$

4. Wyznaczamy

$$
\begin{aligned}
Z_j^N(y_0, ..., y_j, u_j) = \int \tilde{W}_{j+1}^N(\xi, y_0, ..., y_j, u_j, x) x^T \sigma^T(\xi, y_j) \\
\times \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j) \gamma\left(x, \bar{0}_k, I_k\right) dx P\left(d\xi \left| \mathcal{F}_j\right.\right),
\end{aligned}
$$

gdzie $\bar{0}_k \in \mathbb{R}^k$ jest wektorem zerowym, natomiast $I_k \in \mathbb{R}^{k \times k}$ macierzą jednostkową.

5. Szukamy optymalne sterowanie u_j^*, dla którego spełniony jest warunek

$$
\nabla_{u_j} g(y_j, u_j) + Z_j^N(y_0, ..., y_j, u_j) = \mathbf{0},
$$

gdzie $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^l$.

6. Obliczamy

$$
\begin{aligned}
Z_j^N(y_0, ..., y_j) &= E\left(\tilde{W}_{j+1}^N(\xi, y_0, ..., y_j, u_j^*, w_{j+1}) \left| \mathcal{F}_j\right.\right) \\
&= \int \tilde{W}_{j+1}^N\left(\xi, y_0, ..., y_j, u_j^*, w_{j+1}\right) \gamma\left(x, \bar{0}_k, I_k\right) dx P\left(d\xi \left| \mathcal{F}_j\right.\right).
\end{aligned}
$$

7. Jeżeli

$$
g(y_j, u_j^*) + Z_j^N\left(y_0, ..., y_j\right) \leq h\left(y_j\right),
$$

to zmienna decyzyjna $\theta_j^* = 1$, w przeciwnym razie $\theta_j^* = 0$ oraz zatrzymujemy proces sterowania systemem.

8. Wyznaczamy wartość funkcji Bellmana

$$
\begin{aligned}
W_j^N(\xi, y_0, ..., y_j) = \phi_j\left(\theta_j^*, y_j, u_j^*\right) \\
+ \theta_j^* \int W_{j+1}^N\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j) x\right) \gamma\left(x, \bar{0}_k, I_k\right) dx.
\end{aligned}
$$

9. Jeżeli $j = 0$ to zatrzymujemy, w przeciwnym razie wracamy do punktu 2.

4.3 Optymalne zatrzymanie sterowanych systemów stochastycznych

Rozwiązanie zadania (4.3) polega na wyznaczeniu optymalnego momentu zatrzymania $\tau : \Omega \to \{0, 1, ..., N\}$ dla systemu (4.1) oraz optymalnego sterowania $u_\tau^* = (u_0^*, u_1^*, ..., u_\tau^*)$ jednocześnie. Aby rozwiązać zadanie (4.3) należy zastosować zasadę optymalnego zatrzymania procesów stochastycznych z czasem dyskretnym (patrz np. [33], [86], [122], [124]). Zgodnie z klasyczną zasadą zatrzymania procesów losowych wyznaczenie optymalnego momentu stopu dla zadania (4.3) polega na konstrukcji koperty Snell'a, która jest określona jako największy podmartyngał dominujący zatrzymywany proces. Pierwszy moment, dla którego wartość procesu stochastycznego i otoczki Snell'a są równe, jest momentem zatrzymania tego procesu.

Złożoność zatrzymania systemu (4.1) dla zadania (4.3) polega na tym, że zamiast jednego procesu stochastycznego należy rozważać rodzinę procesów indeksowanych sterowaniem. W tym celu należy skonstruować otoczkę dla całej rodziny kopert Snell'a. Jak zobaczymy poniżej moment zatrzymania systemu (4.1) spełniającego cel sterowania (4.3) jest to moment zrównania się kosztu dziedziczenia (jako procesu stochastycznego $\{h(y_j)\}_{0 \le j \le N}$) i otoczki dla rodziny kopert Snell'a.

Do wyznaczenia otoczki Snell'a wykorzystamy technikę programowania dynamicznego wstecz. Dla dowolnego $j \in \{0, 1, ..., N-1\}$ konstruujemy funkcję Bellmana

$$S_j^N(\xi, y_0, ..., y_j) = \inf_{u_j} \left\{ h(y_j), g(y_j, u_j) + E\left(S_{j+1}(\xi, y_0, ..., y_{j+1}) \middle| \mathcal{F}_j^\xi \right) \right\} \tag{4.22}$$

z warunkiem początkowym

$$S_N^N(\xi, y_0, ..., y_N) = h(y_N). \tag{4.23}$$

Otoczkę Snell'a $\left\{\bar{S}_j(y_0, ..., y_j)\right\}_{0 \le j \le N}$ dla systemu sterowanego (4.1) definiujemy w sposób następujący. Dla dowolnego momentu $j \in \{0, 1, ..., N-1\}$

przyjmujemy

$$\bar{S}_j^N(y_0, ..., y_j) = \inf_{u_j} \left\{ h\left(y_j\right), g\left(y_j, u_j\right) + E\left(S_{j+1}^N\left(\xi, y_0, ..., y_{j+1}\right) | \mathcal{F}_j\right) \right\}, \tag{4.24}$$

natomiast $\bar{S}_N^N(y_0, ..., y_N) = h\left(y_N\right)$.

Uwaga 4.6 *Wartość funkcji Bellmana $S_j^N(\xi, y_0, ..., y_j)$ określa najmniejszy możliwy do osiągnięcia całkowity koszt sterowania od momentu j dla przypadku, gdy znane są parametry $\xi \in \mathbb{R}^k$. Jeżeli $S_j^N(\xi, y_0, ..., y_j)$ jest wypukła względem ξ, to z nierówności Jensena dla dowolnego $0 \leq j \leq N$ mamy $\bar{S}_j^N(y_0, ..., y_j) \leq E\left(S_j^N\left(\xi, y_0, ..., y_j\right) | \mathcal{F}_j\right)$.*

W zadaniu (4.3) moment zatrzymania τ jest momentem Markowa względem rodziny σ–ciał $\mathcal{F}_j$, $0 \leq j \leq N$ oraz $P\left(\tau \leq N\right) = 1$. Twierdzenie poniżej podaje sposób rozwiązania tego zadania, dokładniej sposób wyznaczenia momentu zatrzymania systemu (4.1) oraz warunek konieczny istnienia optymalnego sterowania.

Twierdzenie 4.2 *Niech funkcje $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ i $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ są ciągłe, wypukłe i ograniczone, funkcje $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ i $g : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ są funkcjami klasy C^1 ze względu na zmienną sterującą $u \in \mathbb{R}^l$ oraz $\det\left(\Sigma\left(\xi, y\right)\right) \neq 0$ dla $(\xi, y) \in \mathbb{R}^k \times \mathbb{R}^n$, gdzie $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$. Rozwiązaniem zadania (4.3) jest:*

1. *optymalny moment zatrzymania systemu (4.1)*

$$\tau = \min\left\{0 \leq j \leq N : h\left(y_j\right) = \bar{S}_j^N\left(y_0, ..., y_j\right)\right\}, \tag{4.25}$$

   gdzie $\bar{S}_j^N\left(y_0, ..., y_j\right)$ dla $0 \leq j \leq N$ jest dane wzorem (4.24);

2. *jeżeli $j \in \{0, 1, ..., N-1\}$ nie jest momentem zatrzymania, to optymalne sterowanie u_j^* systemem (4.1) spełnia równanie*

$$\begin{aligned} &\nabla_u\, g(y_j, u_j^*) + E\left(S_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)w_{j+1})\right. \\ &\left.\times \left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_j)\, \nabla_u\, f(\xi, y_j, u_j^*) \,|\mathcal{F}_j\right) = \mathbf{0}, \end{aligned} \tag{4.26}$$

   gdzie $\mathbf{0} = col\,(0, ..., 0) \in \mathbb{R}^l$.

Dowód. Rozważmy najpierw zadanie optymalnego stopowania w klasie $\mathcal{T}(N,N)$. Optymalny moment zatrzymania wynosi $\tau^0 = N$ oraz koszt dziedziczenia w momencie N jest równy

$$\bar{S}_N^N(y_0,...,y_N) = E\left(S_N^N(\xi, y_0,...,y_N)\,|\mathcal{F}_N\right) = h(y_N).$$

W klasie $\mathcal{T}(N-1,N)$ możemy zatrzymać system (4.1) w chwili $N-1$ lub sterować tym systemem do momentu N. Jeżeli zatrzymamy system (4.1), to ponosimy tylko koszt dziedziczenia $h(y_{N-1})$. Jeżeli sterujemy tym systemem w momencie $N-1$ oraz zatrzymujemy go w chwili N, to ponosimy koszty sterowania wielkości $g(y_{N-1},u_{N-1})$ (podczas przeprowadzenia systemu ze stanu y_{N-1} do stanu y_N) i dziedziczenia $h(y_N)$.

Optymalny koszt dla klasy $\mathcal{T}(N-1,N)$ jest równy

$$\begin{aligned}\bar{S}_{N-1}^N(y_0,...,y_{N-1}) = \inf_{u_{N-1}} \{&h(y_{N-1}), g(y_{N-1},u_{N-1})\\ &+ E\left(S_N^N(\xi,y_0,...,y_N)\,|\mathcal{F}_{N-1}\right)\}\end{aligned}$$

oraz optymalny moment zatrzymania wynosi

$$\tau^1 = \begin{cases} N-1, & \text{jeżeli } \bar{S}_{N-1}^N(y_0,...,y_{N-1}) = h(y_{N-1}), \\ \tau^0, & \text{jeżeli } \bar{S}_{N-1}^N(y_0,...,y_{N-1}) < h(y_{N-1}). \end{cases}$$

Jeżeli moment $N-1$ (tzn. $\tau^1 = \tau^0$) nie jest momentem zatrzymania, to należy dalej sterować systemem (4.1). Zgodnie z twierdzeniem 2.1 warunek konieczny optymalnego sterowania w momencie $N-1$ jest dany wzorem

$$\begin{aligned}\nabla_u g(y_{N-1},u_{N-1}^*) + E\Big(h(y_N)\left(y_N - f(\xi,y_{N-1},u_{N-1}^*)\right)^T\\ \times\Sigma^{-1}(\xi,y_{N-1})\,\nabla_u f(\xi,y_{N-1},u_{N-1}^*)\big|\,\mathcal{F}_{N-1}\Big) = \mathbf{0}.\end{aligned}$$

W podobny sposób wyznaczamy momenty zatrzymania w klasach $\mathcal{T}(N-2,N),...,\mathcal{T}(0,N) = \mathcal{T}$. Dla klasy $\mathcal{T}(j,N)$, $0 \le j \le N-1$ możemy zatrzymać system (4.1) w momencie j lub optymalnie sterować systemem w chwili j do momentów $j+1,...,N$. Jeżeli system (4.1) zostanie zatrzymany w momencie j, to ponosimy tylko koszt dziedziczenia wielkości $h(y_j)$. Jeżeli w chwili j nadal sterujemy systemem (4.1), to ponosimy koszt sterowania wielkości $g(y_j, u_j)$ (podczas transformacji systemu ze stanu y_j do stanu y_{j+1}) oraz oczekiwane koszty sterowania w chwilach późniejszych wielkości $E\left(S_{j+1}^N(\xi, y_0, ..., y_{j+1}) | \mathcal{F}_j\right)$. Wobec powyższego optymalny koszt w klasie $\mathcal{T}(j, N)$ jest równy

$$\bar{S}_j^N(y_0, ..., y_j) = \inf_{u_j} \left\{h(y_j), g(y_j, u_j) + E\left(S_{j+1}^N(\xi, y_0, ..., y_{j+1}) | \mathcal{F}_j\right)\right\},$$

natomiast optymalny moment zatrzymania w klasie $\mathcal{T}(j, N)$ wynosi

$$\tau^j = \begin{cases} j, & \text{jeżeli } \bar{S}_j^N(y_0, ..., y_j) = h(y_j), \\ \tau^{j-1}, & \text{jeżeli } \bar{S}_j^N(y_0, ..., y_j) < h(y_j). \end{cases}$$

Jeżeli j nie jest momentem zatrzymania, to nadal sterujemy systemem (4.1). Wyznaczenie optymalnego sterowania w momencie j polega na rozwiązaniu zadania

$$\inf_{u_j} g(y_j, u_j) + E\left(S_{j+1}^N(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)w_{j+1}) | \mathcal{F}_j\right). \quad (4.27)$$

Jeżeli u_j^* jest optymalnym sterowaniem w momencie j, to oczekiwany najmniejszy koszt sterowania od momentu j wynosi

$$\begin{aligned} \bar{S}_j^N(y_0, ..., y_j) &= g\left(y_j, u_j^*\right) \\ &\quad + E\left(S_{j+1}^N\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)w_{j+1}\right) | \mathcal{F}_j\right). \end{aligned}$$

Warunek konieczny optymalnego sterowania zadania (4.27) jest dany wzorem

$$\begin{aligned} &\nabla_u g(y_j, u_j^*) + E\Big(S_{j+1}^N(\xi, y_0, ..., y_{j+1})\left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \\ &\qquad \times \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j^*) | \mathcal{F}_k\Big) = \mathbf{0}. \end{aligned}$$

■

Poniżej przedstawiony zostanie algorytm, który podaje sposób wyznaczenia ciągu optymalnych sterowań $\left\{u_j^*\right\}_{0\leq j<N}$ oraz konstrukcję momentu zatrzymania systemu (4.1). Algorytm ten wykorzystuje zasadę programowania dynamicznego wstecz oraz wyniki zawarte w twierdzeniu 4.2.

Algorytm wyznaczenia optymalnego sterowania i momentu zatrzymania

1. Definiujemy
$$S_N^N(\xi, y_0, ..., y_N) = h\left(y_N\right) \text{ oraz } \ j = N.$$

2. Kładziemy
$$j = j - 1.$$

3. Definiujemy
$$\tilde{S}_{j+1}^N(\xi, y_0, ..., y_j, u_j, w) = S_{j+1}^N\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j)w\right).$$

4. Wyznaczamy
$$\begin{aligned} Z_j^N(y_0, ..., y_j, u_j) = \int & \tilde{S}_{j+1}^N(\xi, y_0, ..., y_j, u_j, x) x^T \sigma^T(\xi, y_j) \\ & \times \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j) \gamma\left(x, \bar{0}_k, I_k\right) dx P\left(d\xi \left| \mathcal{F}_j\right.\right), \end{aligned}$$
gdzie $\bar{0}_k \in \mathbb{R}^k$ jest wektorem zerowym, natomiast $I_k \in \mathbb{R}^{k\times k}$ macierzą jednostkową.

5. Szukamy optymalne sterowanie u_j^*, dla którego spełniony jest warunek
$$\nabla_{u_j} g_j(y_j, u_j) + Z_j^N(y_0, ..., y_j, u_j) = \mathbf{0}.$$

6. Wyznaczamy
$$\begin{aligned} \bar{S}_{j+1}^N(y_0, ..., y_j) &= E\left(\tilde{S}_{j+1}^N(\xi, y_0, ..., y_j, u_{j+1}^*, w_{j+1}) \left| \mathcal{F}_j\right.\right) \\ &= \int \tilde{S}_{j+1}^N\left(\xi, y_0, ..., y_j, u_{j+1}^*, w_{j+1}\right) \gamma\left(x, \bar{0}_k, I_k\right) dx P\left(d\xi \left| \mathcal{F}_j\right.\right) \end{aligned}$$

7. Jeżeli
$$g(y_j, u_j^*) + \bar{S}_{j+1}^N(y_0, ..., y_j) \geq h(y_j),$$
to zatrzymujemy proces sterowania systemem (4.1) oraz $\tau = j$, w przeciwnym razie kontynuujemy sterowanie.
8. Wyznaczamy wartość funkcji Bellmana
$$S_j^N(\xi, y_0, ..., y_j) = g_j\left(y_j, u_j^*\right) + \int S_{j+1}^N\left(\xi, y_0, ..., y_j, f(\xi, y_j, u_j^*) + \sigma(\xi, y_j)x\right)\gamma\left(x, \bar{0}_k, I_k\right)dx.$$
9. Jeżeli $j = 0$, to zatrzymujemy proces sterowania oraz $\tau^* = 0$, w przeciwnym razie wracamy do punktu 2.

4.4 Różnice i podobieństwa zadań optymalnego zatrzymania

Porównując przedstawione wyżej sposoby rozwiązania zadania (4.3) i zadania (4.7), które wykorzystują odpowiednio zasadę optymalnego zatrzymania procesów losowych i zamianę stopowania na sterowanie, widzimy, że występują zarówno podobieństwa i różnice. Do wyznaczenia ciągów optymalnych sterowań zarówno dla zadania (4.3), jak i dla zadania zastępczego (4.7) wykorzystana została zasada programowania dynamicznego wstecz. Wydawać by się mogło, że wprowadzenie zmiennej sterującej θ_j w równaniu (4.19) powinno dawać podobny efekt jak równanie (4.22), ponieważ dla dowolnych f_1 i f_2 spełniona jest równość
$$\min\left(f_1, f_2\right) = \min_{\theta \in \{0,1\}} \theta f_1 + (1 - \theta) f_2.$$
Nie mniej jednak konstrukcje ciągu funkcji Bellmana $\left\{W_j^N(\xi, y_0, ..., y_j)\right\}_{0 \leq j \leq N}$ i ciągu funkcji Bellmana $\left\{S_j^N(\xi, y_0, ..., y_j)\right\}_{0 \leq j \leq N}$ zdefiniowanych odpowiednio wzorami (4.19)–(4.20) i (4.22)–(4.23) są zupełnie inne.

Zgodnie z uwagą 4.4 dla zadania zastępczego (4.7) polegającego na zamianie problemu stopowania na sterowanie, wartość $W_j^N(\xi, y_0, ..., y_j)$, $0 \leq j < N$ możemy oszacować jako najmniejszą wartość z funkcji Bellmana $V_j^j(\xi, y_0, ..., y_j), ..., V_j^N(\xi, y_0, ..., y_N)$ dla zadania (2.3) z ustalonym horyzontem N. Konstrukcja zmiennej sterującej $\theta = (\theta_0, ..., \theta_{N-1})$ umożliwia spełnienie równości (4.21). Natomiast dla zadania (4.3) bezpośrednie wyznaczenie funkcji Bellmana $S_j^N(\xi, y_0, ..., y_j)$, $0 \leq j < N$ za pomocą $V_j^j(\xi, y_0, ..., y_j), ..., V_j^N(\xi, y_0, ..., y_N)$ nie jest możliwe.

Wniosek 4.1 *Dla dowolnych $\xi \in \mathbb{R}^k$, $y_i \in \mathbb{R}^k$, $0 \leq i \leq j \leq N$ spełniona jest nierówność*

$$S_j^N(\xi, y_0, ..., y_j) \leq W_j^N(\xi, y_0, ..., y_j) \leq V_j^N(\xi, y_0, ..., y_j),$$

gdzie funkcje Bellmana $V_j^N(\xi, y_0, ..., y_j), W_j^N(\xi, y_0, ..., y_j), S_j^N(\xi, y_0, ..., y_j)$ są dane wzorami (2.14), (4.19), (4.22) odpowiednio. Zatem oczekiwane koszty sterowania od momentu j dla zadań (2.3), (4.7), (4.3) odpowiednio spełniają nierówność

$$\bar{S}_j^N(y_0, ..., y_j) \leq \bar{W}_j^N(y_0, ..., y_j) \leq \bar{V}_j^N(y_0, ..., y_j).$$

Dla momentu N natomiast zachodzi równość

$$S_N^N(\xi, y_0, ..., y_N) = W_N^N(\xi, y_0, ..., y_N) = V_N^N(\xi, y_0, ..., y_N) = h(y_N).$$

Uwaga 4.7 *W warunkach pełnej informacji o systemie (4.1) (wektor $\xi \in \mathbb{R}^k$ jest znany) wartości funkcji $\left\{\bar{W}_j^N(y_0, ..., y_j)\right\}_{0\leq j\leq N}$ i $\left\{\bar{S}_j^N(y_0, ..., y_j)\right\}_{0\leq j\leq N}$ są równe wartościom funkcji Bellmana danymi wzorami (4.19) i (4.22) odpowiedno*

$$\bar{S}_j^N(y_0, ..., y_j) = S_j^N(\xi, y_0, ..., y_j) = S_j^N(\xi, y_j),$$
$$\bar{W}_j^N(y_0, ..., y_j) = W_j^N(\xi, y_0, ..., y_j) = W_j^N(\xi, y_j).$$

Wniosek 4.2 *Optymalny moment zatrzymania systemu (4.1) dla zadania (4.7) (zadania wykorzystującego zamianę stopowania na sterowanie) jest równy*

$$\tilde{\tau} = \min \left\{ 0 \le j \le N : \ \bar{W}_j^N (y_0, ..., y_j) = h\left(y_j\right) \right\}$$

oraz nie przekracza momentu zatrzymania dla zadania (4.3) (zadania wykorzystującego zasadę optymalnego zatrzymania)

$$\tau = \min \left\{ 0 \le j \le N : \ \bar{S}_j^N (y_0, ..., y_j) = h\left(y_j\right) \right\}.$$

4.5 Zbiory optymalnego zatrzymania

Decyzja czy moment $j \in \{0, 1, ..., N\}$ jest momentem zatrzymania systemu stochastycznego (4.1) zależy głównie od trajektorii $\{y_i\}_{0 \le i \le j}$ tego systemu. Zatem z momentami decyzyjnymi są powiązane pewne podzbiory w $\mathbb{R}^n$. Te podzbiory nie są stałe, zależą zarówno od momentu decyzyjnego, jak i od dostępnej informacji o parametrach systemu. Dla dowolnego momentu $0 \le j \le N$ przestrzeń $\mathbb{R}^n$ możemy podzielić na dwa rozłączne zbiory: zbiór zatrzymania i zbiór kontynuacji sterowania.

Definicja 4.1 *Dla momentu* $0 \le j \le N$ *zbiór*

$$G_{jN}^S \left(y_0, ..., y_{j-1}\right) = \left\{ y \in \mathbb{R}^n : \bar{S}_j^N (y_0, ..., y_{j-1}, y) = h\left(y\right) \right\} \tag{4.28}$$

nazywamy zbiorem optymalnego zatrzymania dla zadania (4.3).

Definicja 4.2 *Dla momentu* $0 \le j \le N$ *zbiór*

$$G_{jN}^W \left(y_0, ..., y_{j-1}\right) = \left\{ y \in \mathbb{R}^n : \bar{W}_j^N (y_0, ..., y_{j-1}, y) = h\left(y\right) \right\} \tag{4.29}$$

nazywamy zbiorem optymalnego zatrzymania dla zadania (4.7).

Uwaga 4.8 *Zgodnie z uwagą 4.7 zbiory optymalnego zatrzymania dla zadań z pełną i niepełną informacją są różne. W przypadku pełnej informacji o systemie zbiór zatrzymania zależy od parametrów systemu* ξ *i momentu*

j, natomiast w warunkach niepełnej informacji o systemie zbiór zatrzymania zależy od rozkładu a priori wektora losowego ξ, trajektorii $\{y_i\}_{0\leq i\leq j}$ oraz momentu decyzyjnego j.

Uwaga 4.9 *Z uwagi 4.7 wynika, że w przypadku pełnej informacji o systemie (4.1) dla momentu $0 \leq j \leq N$ zbiory optymalnych zatrzymań dla zadań (4.3) i (4.7) zależą od parametrów systemu*

$$G^S_{jN}(\xi) = \left\{y \in \mathbb{R}^n : S^N_j(\xi, y) = h(y)\right\}, \tag{4.30}$$

$$G^W_{jN}(\xi) = \left\{y \in \mathbb{R}^n : W^N_j(\xi, y) = h(y)\right\}. \tag{4.31}$$

Ponieważ optymalne stopowanie w klasie $\mathcal{T}(N, N)$ polega na zatrzymaniu systemu, zatem

$$G^S_{NN}(y_0, ..., y_{N-1}) = G^W_{NN}(y_0, ..., y_{N-1}) = G^S_{NN}(\xi) = G^W_{NN}(\xi) = \mathbb{R}^n.$$

Dla zadań (4.3) i (4.7) w momencie $0 \leq j \leq N-1$ zbiory optymalnego zatrzymania $G^S_{jN}(y_0, ..., y_{j-1})$ i $G^W_{jN}(y_0, ..., y_{j-1})$ odpowiednio są zbiorami stanów systemu, dla których koszt dziedziczenia nie przekracza oczekiwanych kosztów związanych z kontynuacją sterowania (innymi słowy jest to zbiór stanów systemu dla których nie mamy szans na osiągnięcie lepszego efektu, jeżeli będziemy dalej sterować systemem). Zatem dla dowolnego momentu $0 \leq j \leq N-1$ zbiory optymalnego zatrzymania dla zadań (4.3) i (4.7) możemy przedstawić w postaci

$$\begin{aligned} G^S_{jN}(y_0, ..., y_{j-1}) = \Big\{ & y \in \mathbb{R}^n : h(y) \leq \inf_{u_j} g(y, u_j) \\ & + E\left(S^N_{j+1}(\xi, y_0, ..., y_{j-1}, y, y_{j+1}) | \mathcal{F}_j\right)\Big\} \end{aligned} \tag{4.32}$$

i

$$\begin{aligned} G^W_{jN}(y_0, ..., y_{j-1}) = \Big\{ & y \in \mathbb{R}^n : h(y) \leq \inf_{u_j} g(y, u_j) \\ & + E\left(W^N_{j+1}(\xi, y_0, ..., y_{j-1}, y, y_{j+1}) | \mathcal{F}_j\right)\Big\}. \end{aligned} \tag{4.33}$$

Ze wzorów (4.32) – (4.33) widzimy, że dla dowolnego momentu $j \in \{0, 1, ..., N-1\}$ zbiór stanów systemu, w których system powinien być zatrzymany, zależy od kosztu sterowania $g(y_j, u_j)$ i wartości funkcji Bellmana $S_{j+1}^N(\xi, y_0, ..., y_{j+1})$ i $W_{j+1}^N(\xi, y_0, ..., y_{j+1})$. Zatem dla dowolnych momentów decyzyjnych zbiory otymalnego zatrzymania są różne. Dla zadania (4.3) ciąg $\left\{G_{jN}^S(y_0, ..., y_{j-1})\right\}_{0 \le j \le N}$ definiuje niemalejącą rodzinę zbiorów optymalnego zatrzymania, która odpowiada ciągowi funkcji Bellmana $\left\{S_j^N(\xi, y_0, ..., y_j)\right\}_{0 \le j \le N}$. Dla zadania (4.7) ciąg $\left\{G_{jN}^W(y_0, ..., y_{j-1})\right\}_{0 \le j \le N}$ definiuje niemalejącą rodzinę zbiorów optymalnego zatrzymania, która odpowiada ciągowi funkcji Bellmana $\left\{W_j^N(\xi, y_0, ..., y_j)\right\}_{0 \le j \le N}$.

Moment optymalnego zatrzymania τ systemu (4.1) dla zadania (4.3) jest to pierwszy moment, dla którego stan systemu należy do zbioru optymalnego zatrzymania

$$\tau = \min\left\{0 \le j \le N : y_j \in G_{jN}^S(y_0, ..., y_{j-1})\right\}.$$

Również dla zadania (4.7) moment optymalnego zatrzymania $\tilde{\tau}$ wyznaczamy jako

$$\tilde{\tau} = \min\left\{0 \le j \le N : y_j \in G_{jN}^W(y_0, ..., y_{j-1})\right\}.$$

Z wniosku 4.1 wynika, że dla dowolnego momentu $j \in \{0, 1, ..., N-1\}$ mamy $G_{jN}^S(y_0, ..., y_{j-1}) \subset G_{jN}^W(y_0, ..., y_{j-1})$, zatem $\tilde{\tau} \le \tau$.

Dla zadania (4.3) optymalna decyzja dotycząca zatrzymania systemu (4.1) w momencie $0 \le j \le N$ jest następująca: jeżeli $y_j \notin G_{jN}^S(y_0, ..., y_{j-1})$ to w momencie j system powinien być sterowany, w przeciwnym razie (tzn. $y_j \in G_{jN}^S(y_0, ..., y_{j-1})$) w momencie j system powinien być zatrzymany. W podobny sposób określamy decyzję o zatrzymaniu systemu (4.1) w momencie $0 \le j \le N$ dla zadania (4.7): jeżeli $y_j \notin G_{jN}^W(y_0, ..., y_{j-1})$ to w momencie j system powinien być sterowany, w przeciwnym razie (tzn. $y_j \in G_{jN}^W(y_0, ..., y_{j-1})$) w momencie j system powinien być zatrzymany.

Definicja 4.3 *W momencie $j \in \{0, 1, ..., N-1\}$ wartość*

$$L_{jN}^{S}(y_0, ..., y_{j-1}) = \max\left\{h(y) : y \in G_{jN}^{S}(y_0, ..., y_{j-1})\right\} \tag{4.34}$$

oznacza największy akceptowalny koszt dziedziczenia dla zadania (4.3) żeby zatrzymać system.

Definicja 4.4 *W momencie $j \in \{0, 1, ..., N-1\}$ wartość*

$$L_{jN}^{W}(y_0, ..., y_{j-1}) = \max\left\{h(y) : y \in G_{jN}^{W}(y_0, ..., y_{j-1})\right\} \tag{4.35}$$

oznacza największy akceptowalny koszt dziedziczenia dla zadania (4.7) żeby zatrzymać system.

Wniosek 4.3 *Ponieważ $G_{jN}^{S}(y_0, ..., y_{j-1}) \subset G_{jN}^{W}(y_0, ..., y_{j-1})$ dla dowolnego $j \in \{0, 1, ..., N-1\}$, zatem $L_{jN}^{S}(y_0, ..., y_{j-1}) \leq L_{jN}^{W}(y_0, ..., y_{j-1})$.*

4.6 Optymalne sterowanie i zatrzymanie dla zadania liniowo-kwadratowego

W rozdziale 2.2 przedstawiony został klasyczny problem sterowania liniowo-kwadratowego z ustalonym horyzontem, w którym to dodatkowo zwrócono uwagę, że horyzont ma dość duży wpływ na realizację celu. W tej części rozdziału również przedstawiona zostanie problematyka sterowania liniowo-kwadratowego, z tym że główny akcent zostanie skierowany na wyznaczenie optymalnego momentu zatrzymania podczas sterowania stochastycznym systemem liniowym z kwadaratowym wskaźnikiem jakości.

Niech system stochastyczny będzie określony za pomocą równania stanu

$$y_{i+1} = Ay_i - Cu_i + d + \sigma w_{i+1}, \tag{4.36}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{n \times l}$, $\sigma \in \mathbb{R}^{n \times m}$, $d \in \mathbb{R}^n$ oraz $\|A\| < \infty$, $\|C\| < \infty$, $\|\sigma\| < \infty$, $\langle d, d\rangle < \infty$, $\{w_i\}_{i \geq 1}$ oznacza ciąg niezależnych wektorów losowych o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, $\bar{0} \in \mathbb{R}^m$

jest wektorem zerowym, $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy niemalejącą rodzinę $\sigma-$ciał $\{\mathcal{F}_i\}_{i\geq 0}$, gdzie $\mathcal{F}_i = \sigma\{w_s : s = 1, 2, ..., i\}$.

Klasyczne zadanie sterowania liniowo–kwadratowego polega na przeprowadzeniu w N krokach systemu (4.36) z punktu y_0 do celu $a \in \mathbb{R}^n$. W odróżnieniu od klasycznego zadania rozważamy dodatkowo możliwość zatrzymania systemu jeszcze przed momentem N. W każdym momencie $j = 0, 1, ..., N$ możemy podjąć decyzję odnośnie sterowania bądź zatrzymania systemu. W momencie $0 \leq j < N$ ponosimy koszt sterowania wysokości $u_j^T R_j u_j$, jeżeli system jest sterowany, lub koszt dziedziczenia $(y_j - a)^T Q(y_j - a)$, jeżeli system zostanie zatrzymany. Celem sterowania jest minimalizacja całkowitych kosztów związanych ze sterowaniem i dziedziczeniem.

Zamiana stopowania na sterowanie dla liniowego systemu stochastycznego

Sterując systemem (4.36) w każdej chwili $0 \leq i < N$ podejmujemy decyzję odnośnie sterowania i zatrzymania systemu. Niech $\mathbf{u} = (\theta, u)$ oznacza rozszerzone sterowanie, gdzie decyzje dotyczące zatrzymania systemu są określone wektorem $\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$, $\theta_i \in \{0, 1\}$ dla $0 \leq i \leq N-1$, natomiast $u = (u_0, u_1, ..., u_{N-1})$ oznacza sterowanie dopuszczalne. Jeżeli moment $0 \leq i < N$ nie jest momentem zatrzymania, to $\theta_i = 1$ oraz koszt sterowania jest równy $u_i^T R_i u_i$. Natomiast jeżeli moment $0 \leq i \leq N$ jest momentem zatrzymania, to $\theta_i = 0$ oraz straty związane z nietrafieniem do celu $a \in \mathbb{R}^n$ są równe $(y_i - a)^T Q(y_i - a)$. Koszt sterowania w momencie i jest dany wzorem

$$\phi_i(\theta_i, y_i, u_i) = \theta_i u_i^T R_i u_i + (1 - \theta_i)(y_i - a)^T Q(y_i - a). \tag{4.37}$$

Całkowity koszt sterowania jest równy

$$J(\mathbf{u}) = E\left(\sum_{i=0}^{N-1} \phi_i(\theta_i, y_i, u_i)\,\psi_0^{i-1}(\theta) + (y_N - a)^T Q(y_N - a)\,\psi_0^{N-1}(\theta)\right), \tag{4.38}$$

gdzie $\psi_0^i(\theta)$, $0 \le i < N$ jest dane wzorem (4.4).

Do rozwiązania zadania (4.7) konstruujemy ciąg funkcji Bellmana. Dla momentu N przyjmujemy

$$W_N^N(y_N) = (y_N - a)^T Q (y_N - a), \tag{4.39}$$

natomiast dla momentów $j = 0, 1, 2, ..., N-1$

$$\begin{aligned} W_j^N(y_j) = \min_{u_j, \theta_j} \Big\{ &\theta_j u_j^T R_j u_j + (1-\theta_j)(y_j - a)^T Q (y_j - a) \\ &+ \theta_j E\left(W_{j+1}^N(y_{j+1}) \middle| \mathcal{F}_j \right) \Big\}. \end{aligned} \tag{4.40}$$

Dla dowolnego $0 \le k \le N$ zdefiniujemy ciąg macierzy $\left\{Q_j^k\right\}_{0 \le j \le k}$, który spełnia równanie

$$\begin{aligned} Q_j^k = {} & Q_{j+1}^k - \left(Q_{j+1}^k\right)^T A^{k-j-1} C \\ & \times \left(\bar{R}_j + \left(A^{k-j-1}C\right)^T \bar{Q}_{j+1}^k A^{k-j-1} C\right)^{-1} \left(A^{k-j-1}C\right)^T Q_{j+1}^k \end{aligned} \tag{4.41}$$

z warunkiem początkowym $Q_k^k = Q$, gdzie $\bar{R}_j$, $\bar{Q}_j^k$ są dane wzorem (2.25). Korzystając z twierdzenia 2.2 oraz uwagi 4.4 funkcje Bellmana zdefiniowane wzorami (4.39)–(4.40) możemy przedstawić w postaci

$$W_j^N(y) = \min\left\{ V_j^j(y), V_j^{j+1}(y), ..., V_j^N(y) \right\}, \tag{4.42}$$

gdzie dla $j \le k \le N$

$$V_j^k(y) = \left(A^{k-j} y + \Lambda_{k-j} d - a\right)^T Q_j^k \left(A^{k-j} y + \Lambda_{k-j} d - a\right) + \varphi_j^k, \tag{4.43}$$

natomiast

$$\varphi_j^k = \varphi_{j+1}^k + tr\left(\left(A^{k-j-1}\sigma\right)^T Q_{j+1}^k A^{k-j-1}\sigma\right) \tag{4.44}$$

z warunkiem początkowym $\varphi_k^k = 0$ oraz

$$\Lambda_s = I + A + ... + A^{\max(0,s-1)}, \tag{4.45}$$

$I \in \mathbb{R}^{n\times n}$ jest macierzą jednostkową.

Rozwiązanie zadania (4.7) zamiany stopowania na sterowanie dla wskaźnika jakości (4.38) przedstawia twierdzenie poniżej.

Twierdzenie 4.3 *Niech macierze $R_i \in \mathbb{R}^{l\times l}$ dla $0 \le i \le N-1$ oraz $Q \in \mathbb{R}^{n\times n}$ są dodatnio określone oraz dla dowolnego $0 \le k \le N$ ciąg macierzy $\left\{Q_j^k\right\}_{0\le j\le k}$ spełnia równanie (4.41). Optymalnym rozwiązaniem zadania (4.7) dla systemu liniowego (4.36) jest:*

1. *jeżeli moment $0 \le j \le N$ nie jest momentem zatrzymania, to optymalne sterowanie systemem (4.36) w tym momencie jest równe*

$$\begin{aligned} u_j^* = &\left(\bar{R}_j + \left(A^{k-j-1}C\right)^T \bar{Q}_{j+1}^k A^{k-j-1}C\right)^{-1} \\ &\times \left(A^{k-j-1}C\right)^T Q_{j+1}^k \left(A^{k-j}y_j + \Lambda_{k-j}d - a\right), \end{aligned} \tag{4.46}$$

   gdzie przewidywany horyzont k spełnia równanie $W_j^N(y) = V_j^k(y)$, $j \le k \le N$ oraz $W_j^N(y)$ i $V_j^k(y)$ dane są wzorami (4.42)–(4.45);

2. *wartość zmiennej decyzyjnej dotyczącej zatrzymania systemu w momencie $0 \le j \le N-1$ wyznaczamy ze wzoru*

$$\theta_j = \begin{cases} 0, & \text{jeżeli } W_j^N(y_j) = (y_j - a)^T Q (y_j - a), \\ 1, & \text{jeżeli } W_j^N(y_j) < (y_j - a)^T Q (y_j - a). \end{cases} \tag{4.47}$$

Dowód. Stosując bezpośrednio twierdzenie 4.1 oraz uwagę 4.4 otrzymujemy tezę. ■

Optymalne zatrzymanie liniowego systemu sterowanego

W przypadku gdy sterujemy systemem (4.36) do momentu $t \le N$, to ponosimy całkowity koszt (suma kosztów sterowań i dziedziczenia) wielkości

$$J(u,t) = E\left(\sum_{i=0}^{t-1} u_i^T R u_i + (y_t - a)^T Q (y_t - a)\right). \tag{4.48}$$

Głównym celem sterowania jest minimalizacja całkowitych kosztów (4.48). Aby optymalnie sterować tym systemem należy rozwiązać zadanie (4.3), które polega na wyznaczeniu zarówno optymalnego momentu $\tau : \Omega \to \{0, 1, ..., N\}$ zatrzymania systemu (4.36), jak również optymalnego sterowania $u^* = \left(u_0^*, ..., u_{\tau-1}^*\right)$. Do rozwiązania zadania (4.3) konstruujemy ciąg funkcji Bellmana. Dla momentu N przyjmujemy

$$S_N^N(y_N) = (y_N - a)^T Q (y_N - a), \tag{4.49}$$

natomiast dla momentów $j = 0, 1, 2, ..., N-1$

$$S_j^N(y_j) = \min_{u_j} \left\{ (y_j - a)^T Q (y_j - a), u_j^T R_j u_j + E\left(S_{j+1}^N(y_{j+1}) \middle| \mathcal{F}_j \right) \right\}. \tag{4.50}$$

Twierdzenie poniżej przedstawia warunki konieczne wyznaczenia optymalnego sterowania systemem (4.36) oraz sposób wyznaczenia optymalnego momentu zatrzymania.

Twierdzenie 4.4 *Niech macierze $R_i \in \mathbb{R}^{l\times l}$ dla $0 \le i \le N-1$ oraz $Q \in \mathbb{R}^{n\times n}$ są dodatnio określone. Rozwiązaniem zadania (4.3) jest:*

1. *optymalny moment zatrzymania systemu liniowego (4.36)*

$$\tau^* = \min\left\{ 0 \le j \le N : (y_j - a)^T Q (y_j - a) = S_j^N(y_j) \right\}, \tag{4.51}$$

gdzie $S_j^N(y_j)$ dla $0 \le j \le N$ jest dane wzorem (4.49)–(4.50);

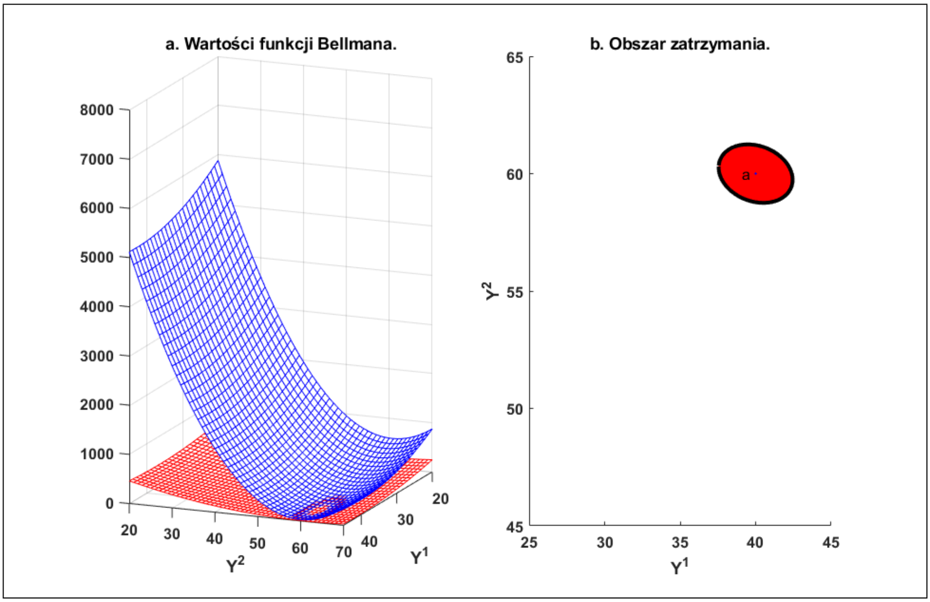
2. *jeżeli moment $j \in \{0, 1, ..., N-1\}$ nie jest momentem zatrzymania, to optymalne sterowanie u_j^* systemem (4.36) jest równe*

$$u_j^* = \left(R_j + R_j^T\right)^{-1} E\left(S_{j+1}^N(y_{j+1}) \left(y_{j+1} - Ay_j + Cu_j^* - d\right)^T \left(\sigma\sigma^T\right)^{-1} C \,|\mathcal{F}_j \right). \tag{4.52}$$

Dowód. Stosując bezpośrednio twierdzenie 4.2 otrzymujemy tezę. ■

Przykład 4.1 *Wyznaczymy zbiory optymalnego zatrzymania dla liniowego systemu stochastycznego, który jest określony za pomocą równania (4.36). Poniżej przyjmujemy macierze* $R = \begin{bmatrix} 1.25 & 0.2 \\ 0.2 & 1.8 \end{bmatrix}$, $Q = \begin{bmatrix} 1.5 & 0.1 \\ 0.1 & 3.2 \end{bmatrix}$,

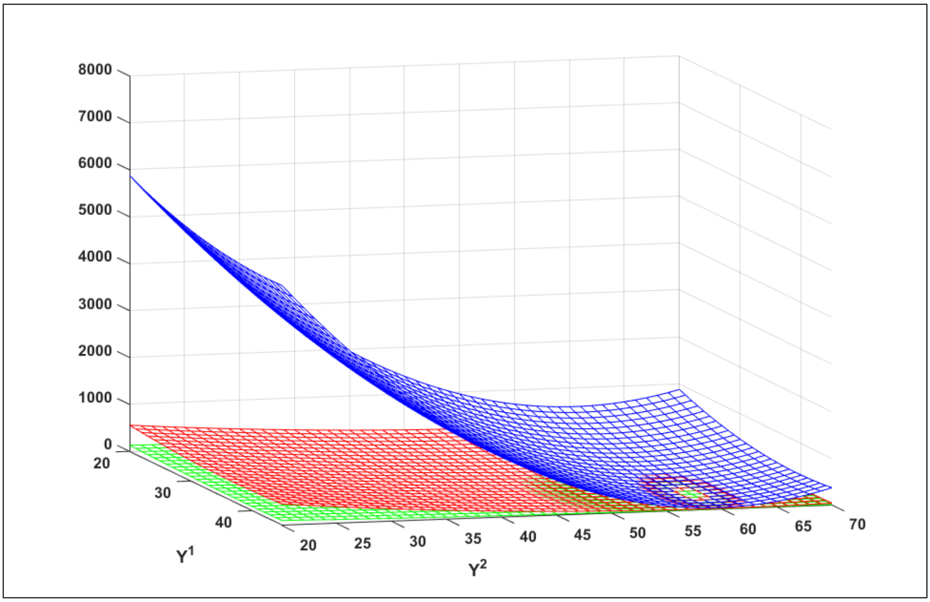
$A = \begin{bmatrix} 1.1 & 0 \\ 0 & 0.9 \end{bmatrix}$, $C = \begin{bmatrix} 1.3 & 0.1 \\ 0.5 & 2.5 \end{bmatrix}$, $\sigma = \begin{bmatrix} 0.95 & 0.2 \\ 0.1 & 1.2 \end{bmatrix}$, *stan początkowy* $y_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ *oraz* $d = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, *cel sterowania* $a = \begin{bmatrix} 40 \\ 60 \end{bmatrix}$, *natomiast horyzont sterowania* $N = 20$.



Rysunek 4.1: Wartości funkcji Bellmana $S_{20}^{20}(y)$ oraz $S_{19}^{20}(y)$, obszar zatrzymania dla momentu $j = 19$.

Bezpośrednio ze wzorów (4.49)-(4.50) wyznaczamy wartości funkcji Bellmana w momentach $j = 0, 1, ..., N$. *Na rysunku 4.1a powierzchnia zaznaczona kolorem granatowym przedstawia wartości funkcji Bellmana dla momentu* $j = 20$, *natomiast powierzchnia zaznaczona kolorem czerwonym przedstawia wartości funkcji Bellmana dla momentu* $j = 19$. *Również widzimy, że w pewnym otoczeniu punktu* $a \in \mathbb{R}^2$ *spełniona jest nierówność* $S_{20}^{20}(y) \leq S_{19}^{20}(y)$.

Rysunek 4.1b przedstawia $G_{19,20}^{S}$ *obszar zatrzymania systemu (4.36) w momencie* $j = 19$, *który wyznaczamy za pomocą wzoru (4.30). Rzut ortogonalny przecięcia powierzchni przedstawiających wartości funkcji Bellmana*

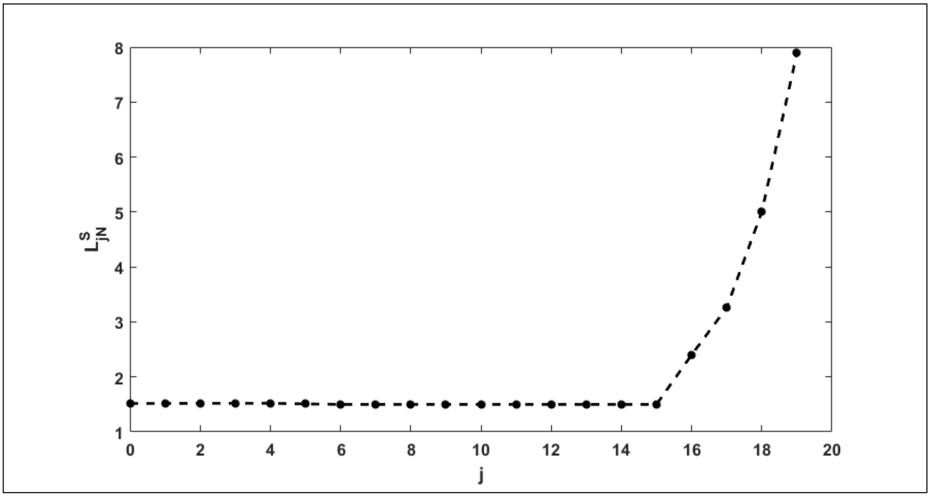


Rysunek 4.2: Wartości funkcji Bellmana $S_{20}^{20}(y)$, $S_{19}^{20}(y)$ oraz $S_{15}^{20}(y)$.

$S_{20}^{20}(y)$ i $S_{19}^{20}(y)$ na przestrzeń $\mathbb{R}^2$ jest to zbiór $\{y \in \mathbb{R}^2 : S_{20}^{20}(y) = S_{19}^{20}(y)\}$, który jest zaznaczony na rysunku 4.1b krzywą koloru czarnego, natomiast zbiór punktów $\{y \in \mathbb{R}^2 : S_{20}^{20}(y) < S_{19}^{20}(y)\}$ jest zaznaczony jako obszar koloru czerwonego. Dla momentów $j = 0, 1, 2, ..., 19$ zbiory optymalnego zatrzymania systemu (4.36) wyglądają w sposób podobny, z tym że $G_{0,20}^S \subset G_{1,20}^S \subset ... \subset G_{19,20}^S$.

Zgodnie z definicją funkcji Bellmana (4.49)-(4.50) dla dowolnych $0 \leq i \leq j \leq N-1$ spełniona jest nierówność $S_i^{20}(y) \leq S_j^{20}(y)$, $y \in \mathbb{R}^2$. Na rysunku 4.2 przedstawione są wartości funkcji Bellmana dla momentów $j = 20$ (powierzchnia zaznaczona kolorem granatowym), $j = 19$ (powierzchnia zaznaczona kolorem czerwonym) oraz $j = 15$ (powierzchnia zaznaczona kolorem zielonym).

Rysunek 4.3 przedstawia wartości największych akceptowalnych kosztów dziedziczenia $L_{j,20}^S$ dla momentów $0 \leq j \leq 19$, które szacujemy korzystając ze wzoru (4.34). Zgodnie z wnioskiem 4.3 wartości akceptowalnych kosztów dziedziczenia spełniają nierówność $L_{0,20}^S \leq L_{1,20}^S \leq ... \leq L_{19,20}^S$. Z rysunku widać, że dla momentów $j = 0, 1, ..., 15$ maksymalny akceptowalny *koszt dziedziczenia oscyluje w okolicach wartości* 1.5*, natomiast od momentu* $j = 16$ *zaczyna dynamicznie wzrastać. Również obszary zatrzymania* $G^S_{j,20}$ *dla momentów* $j = 0, 1, ..., 15$ *są zdecydowanie mniejsze niż dla momentów* $j \in \{16, 17, 18, 19\}$*. Oznacza to, że zbliżając się do horyzontu* $N = 20$ *wzrasta skłonność do zatrzymania systemu ponosząc przy tym większe koszty.*



Rysunek 4.3: Największy akceptowalny koszt dziedziczenia L^S_{jN} dla $j = 0, 1, ..., N - 1$.

Rozdział 5

Optymalna trajektoria

Wyznaczenie optymalnej (najkrótszej, najtańszej, najszybszej) trasy jest jednym z problemów, rozwiązanie którego znajduje zastosowanie w wielu gałęziach ludzkiej aktywności. Wśród wielu znanych przykładów możemy znaleźć problem najkrótszej trasy (shortest path problem SSP), problem komiwojażera (travelling salesman problem TSP), problem nawigacji (vehicle routing problem VRP) oraz różnorodne modyfikacje i rozszerzenia tych zagadnień (patrz np. [5], [80], [82], [107]). Powyższe zadania głównie polegają na odwiedzeniu niektórych, bądź wszystkich węzłów (punktów). Wśród wielu możliwych do przebycia tras należy wybrać taką, dla której łączny koszt (waga) połączeń (łuków między węzłami) jest najmniejszy. Większość problemów zaprezentowanych w literaturze przedmiotu dotyczy liniowych systemów deterministycznych, nie mniej jednak nawet przy tak restrykcyjnych założeniach mogą być trudne do rozwiązania.

W niektórych przypadkach dla sterowanych systemów dynamicznych lepiej określić optymalną drogę zamiast bezpośrednio wyznaczać prawa optymalnego sterowania. W tym przypadku należy wyznaczyć uporządkowany ciąg stanów systemu (zbiór punktów) $\{y_i\}_{0\leq i\leq N}$, który odpowiada znalezieniu pewnej drogi (ścieżki jako sekwencji węzłów, punktów), zwanej trajektorią. Zadanie polega na określeniu optymalnej trajektorii, poruszając się po której ponosimy najmniejsze koszty. Po wyznaczeniu węzłów (punktów) należy sterować systemem w taki sposób, aby podążać (naśladować) wyznaczoną ścieżką (trasą, drogą). Bezpośrednio naśladując określone stany systemu realizujemy główny cel sterowania w sposób optymalny. Dla każdego momentu i znając bieżący oraz przyszły stan systemu (system znajduje się w stanie y_i oraz musi być przeprowadzony do wyznaczonego wcześniej stanu y_{i+1}) możemy określić sterowanie systemem u_i.

Problematyka wyznaczenia optymalnej trajektorii dla liniowych systemów stochastycznych częściowo została przedstawiona np. w pracach [62], [63], [66], [69]. W niniejszym rozdziale monografii przedstawione zostanie zadanie oraz sposób wyznaczenia optymalnej trajektorii dla systemów stochastycznych z czasem dyskretnym. Dodatkowo pokazany zostanie wpływ horyzontu sterowania na optymalną trajektorię dla liniowych systemów stochastycznych z kwadratowym wskaźnikiem jakości.

5.1 Wyznaczenie trajektorii dla systemów stochastycznych z czasem dyskretnym

5.1.1 Postawienie zadania

Niech $(\Omega, \mathcal{F}, P)$ będzie zupełną przestrzenią probabilistyczną Na tej przestrzeni zdefiniujemy ciąg $\{w_i\}_{1\leq i\leq N}$ niezależnych wektorów losowych ($w_i : \Omega \to \mathbb{R}^m$ dla $i = 1, 2, ..., N$) o rozkładzie normalnym $\boldsymbol{N}(\bar{0}, I_m)$, gdzie $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m\times m}$ – macierzą jednostkową. Wielkość N oznacza horyzont sterowania. Niech ξ będzie $k-$ wymiarowym wektorem losowym o rozkładzie a priori $P\,(d\xi)$, natomiast y_0 – stanem poczatkowym o rozkładzie $P\,(dy_0)$. Zakładamy stochastyczną niezależność wszystkich wymienionych obiektów. Definiujemy ciąg σ–ciał $\mathcal{F}_j = \sigma\,(y_0) \vee \sigma\,\{w_i : i = 1, 2, ..., j\}$, $\mathcal{F}_j^\xi = \mathcal{F}_j \vee \sigma\,(\xi)$ dla $0 \leq j \leq N$ oraz przyjmujemy $\mathcal{F} = \mathcal{F}_N^\xi$.

Niech system stochastyczny będzie określony za pomocą równania

$$y_{i+1} = f(\xi, y_i, u_i) + \sigma(\xi, y_i)w_{i+1}, \tag{5.1}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ oraz

$\sigma : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{M}(n,m)$, natomiast $\mathbb{M}(n,m)$ jest zbiorem macierzy o $n-$ wierszach i $k-$ kolumnach. Zakładamy, że funkcje f, σ są ciagłe względem wszystkich argumentów oraz $f(\xi, y, u)$ jest różnowartościowa względem sterowania u. Sterując systemem stochastycznym z czasem dyskretnym (5.1) podejmujemy działania sterujące w chwilach $0, 1, ..., N-1$.

Jeżeli jest określona trajektoria, po której należy przeprowadzić system (5.1), to możemy wyznaczyć sterowanie. Przeprowadzając system ze stanu y_j do stanu y_{j+1} podejmujemy sterowanie wielkości

$$u_j = u_j\left(\xi, y_j, y_{j+1}, w_{j+1}\right) = f^{-1}\left(\xi, y_j, y_{j+1} - \sigma(\xi, y_j) w_{j+1}\right). \tag{5.2}$$

W momencie $i = 0$ system znajduje się w stanie y_0. Ciąg stanów $\{y_i\}_{1 \le i \le N}$ nazywamy trajektorią dopuszczalną oraz oznaczamy $y = (y_1, y_2, ..., y_N)$. Nie nakładamy dodatkowych ograniczeń na stany systemu (5.1). Klasę dopuszczalnych trajektorii oznaczamy przez Y. Zadanie polega na wyznaczeniu optymalnej trajektorii, która powinna być naśladowana przez system (5.1).

Aby sformułować problem wyznaczenia optymalnej trasy definiujemy funkcję kosztów $g : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ jako

$$g(\xi, y_i, y_{i+1}, w_{i+1}) = g\left(y_i, f^{-1}\left(\xi, y_j, y_{j+1} - \sigma(\xi, y_j) w_{j+1}\right)\right) \tag{5.3}$$

oraz funkcję dziedziczenia $h : \mathbb{R}^n \longrightarrow \mathbb{R}$. Zakładamy, że g oraz h są ciągłe i ograniczone. Dla momentów $i \in \{0, 1, ..., N-1\}$ funkcjonał g reprezentuje łączny koszt związany z transformacją systemu (5.1) ze stanu y_i do stanu y_{i+1} oraz skutkami spowodowanymi zaburzeniami zewnętrznymi w_{i+1}. Wskaźnik jakości definiujemy jako sumę kosztów poniesionych w momentach $i \in \{0, 1, ..., N-1\}$ oraz dziedziczenia na końcu horyzontu czasowego N

$$J^N(y) = E\left(\sum_{i=0}^{N-1} g(\xi, y_i, y_{i+1}, w_{i+1}) + h(y_N)\right). \tag{5.4}$$

Dla tak skonstruowanego wskaźnika jakości klasyczny problem optymalnego sterowania (2.3) systemem stochastycznym (5.1) zastępujemy zadaniem dotyczącym wyznaczenia optymalnej trajektorii, poruszając się po której osiągamy najmniejsze koszty całkowite. Aby wyznaczyć optymalną trajektorię należy rozwiązać zadanie

$$\inf_{y \in Y} J^N(y), \tag{5.5}$$

tzn. wyznaczyć taką trajektorię $y^* = (y_1^*, ..., y_N^*)$, dla której infimum jest osiągnięte.

Trajektoria systemu (5.1) zależy od parametrów ξ. Jeżeli te parametry nie są znane, to bezpośrednio realizując główny cel (5.5) oraz dodatkowo obserwując trajektorię analizujemy zachowanie systemu, tym samym wzbogacamy naszą wiedzę o parametrach ξ. Wyznaczając ciąg stanów $\left\{y_j^*\right\}_{1 \le j \le N}$ definiujemy trajektorię, na której odpowiednio szybko gromadzimy wiedzę o systemie (5.1) oraz realizujemy główny cel (5.5).

Twierdzenie 5.1 *Jeżeli funkcje* $g : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ *i* $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ *są ciągłe, wypukłe i ograniczone oraz są funkcjami klasy* C^1 *ze względu na zmienną* y*, to spełnienie równości*

$$E\Bigg(\bigtriangledown_{y_{j+1}} g(\xi, y_j^*, y_{j+1}^*, w_{j+1}) + \bigtriangledown_{y_{j+1}} g(\xi, y_{j+1}^*, y_{j+2}^*, w_{j+2}) + \sum_{i=j+2}^{N-1} g(\xi, y_i^*, y_{i+1}^*, w_{i+1}) + h(y_N^*) \Bigg| \mathcal{F}_j \Bigg) = \mathbf{0} \tag{5.6}$$

dla wszystkich $j \in (0, 1, ..., N-2)$ *oraz równości*

$$E\left(\bigtriangledown_{y_N} g(\xi, y_{N-1}^*, y_N^*, w_{i+1}) + \bigtriangledown_{y_N} h(y_N^*) \big| \mathcal{F}_j\right) = \mathbf{0} \tag{5.7}$$

jest warunkiem koniecznym wyznaczenia optymalnej trajektorii $(y_1^*, y_2^*, ..., y_N^*)$*, gdzie* $\mathbf{0} = \mathrm{col}\,(0, ..., 0) \in \mathbb{R}^n$.

Dowód. Dla dowolnego momentu $j = 1, 2, ..., N$ wskaźnik jakości (5.4) możemy przedstawić w postaci

$$J^N(u) = E\left(\sum_{i=0}^{j-1} g\left(\xi, y_i, y_{i+1}, w_{i+1}\right)\right.$$

$$+E\left(\sum_{i=j}^{N-1} g\left(\xi, y_i, y_{i+1}, w_{i+1}\right) + h\left(y_N\right)\middle| \mathcal{F}_j^{\xi}\right)\Bigg)$$

lub

$$J^N(u) = \int\left(\sum_{i=1}^{j-1} g\left(\xi, y_i, y_{i+1}, w_{i+1}\right)\right) P\left(d\xi, dy_0, dw_1, dw_2, ..., dw_j\right) +$$
$$\int\left(\int\left(\sum_{i=j}^{N-1} g\left(\xi, y_i, y_{i+1}, w_{i+1}\right) + h(y_N)\right) P\left(dw_{j+1}, ..., dw_N\right)\right)$$
$$\times P\left(d\xi, dy_0, dw_1, dw_2, ..., dw_j\right), \tag{5.8}$$

gdzie rozkłady łączne $d\xi, dy_0, dw_1, dw_2, ..., dw_j$ oraz $dw_{j+1}, ..., dw_N$ dane są wzorami

$$P\left(d\xi, dy_0, dw_1, \ldots, dw_i\right) = P\left(d\xi\right) P\left(dy_0\right) P\left(dw_1, \ldots, dw_j\right)$$

oraz dla $1 \leq i \leq j \leq N$

$$P\left(dw_i, \ldots, dw_j\right) = \prod_{k=i}^{j} \gamma\left(w_k, \bar{0}, I_m\right) dw_k,$$

natomiast $\gamma\left(x, \bar{0}, I_m\right)$ jest funkcją gęstości rozkładu normalnego $\boldsymbol{N}(\bar{0}, I_m)$ daną wzorem (1.4).

Ustalmy liczbę $j \in \{1, \ldots, N\}$. Niech $y = y^* + \epsilon v$, gdzie $y^* \in \mathbb{R}^{n\times N}$ oznacza optymalną trajektorię dla systemu (5.1), $\epsilon-$ skalar, $v : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^{n\times N}$, $v = (\mathbf{0}, ..., \mathbf{0}, \tilde{v}_{j+1}, \mathbf{0}, ..., \mathbf{0})$, gdzie $\mathbf{0} = \operatorname{col}(0, ..., 0) \in \mathbb{R}^n$ oraz $\tilde{v}_j : \mathbb{R}^{n\times(j+1)} \longrightarrow \mathbb{R}^n$, natomiast $\tilde{v}_{j+1} = \operatorname{col}(v_1, ..., v_n)$, gdzie $v_s = v_s(y_0, ..., y_j)$ dla $1 \leq s \leq n$ jest dowolną funkcją borelowską. Warunek konieczny istnienia optymalnego stanu w momencie j jest dany równaniem

$$\frac{\partial}{\partial\epsilon} J^N(y^* + \epsilon v)\bigg|_{\epsilon=0} = 0. \tag{5.9}$$

Korzystając z rozwinięcia (5.8) otrzymujemy

$$\frac{\partial}{\partial\epsilon}J^N(y^*+\epsilon v)\bigg|_{\epsilon=0} = \int \langle \nabla_{y_{j+1}} g(\xi, y_j^*, y_{j+1}^*, w_{j+1})$$
$$+ \nabla_{y_{j+1}} g(\xi, y_{j+1}^*, y_{j+2}^*, w_{j+2}) + \left(\sum_{i=j+2}^{N-1} g(\xi, y_i^*, y_{i+1}^*, w_{i+1}) + h(y_N) \right)$$
$$\times P\left(dw_{j+1}, ..., dw_N\right), v_{j+1} \rangle \, P\left(d\xi, dy_0, dw_1, dw_2, ..., dw_j\right). \tag{5.10}$$

Z warunku (5.9) otrzymujemy równanie

$$\int \Big\langle \nabla_{y_{j+1}} \left(g(\xi, y_j^*, y_{j+1}^*, w_{j+1}) + g(\xi, y_{j+1}^*, y_{j+2}^*, w_{j+2}) \right)$$
$$+ \left(\sum_{i=j+2}^{N-1} g(\xi, y_i^*, y_{i+1}^*, w_{i+1}) + h(y_N) \right) P\left(dw_{j+1}, ..., dw_N\right), v_{j+1} \Big\rangle$$
$$\times P\left(d\xi, dy_0, dw_1, dw_2, ..., dw_j\right) = 0. \tag{5.11}$$

Ponieważ warunek (5.11) musi być spełniony dla dowolnej funkcji borelowskiej v_{j+1}, gdzie v_{j+1} jest mierzalne względem $\sigma-$ ciała $\mathcal{F}_j$, zatem otrzymujemy tezę. ■

Uwaga 5.1 *Jeżeli znana jest optymalna trajektoria* $(y_1^*, y_2^*, ..., y_N^*)$ *dla systemu (5.1), to ze wzoru (5.2) w każdym momencie* $j \in (0, 1, ..., N-1)$ *możemy wyznaczyć optymalne sterowanie jako*

$$u_j^* = E\left(f^{-1}\left(\xi, y_j, y_{j+1}^* - \sigma(\xi, y_j) w_{j+1}\right) \middle| \mathcal{F}_j \right). \tag{5.12}$$

Uwaga 5.2 *Stosując wzór (5.12) klasyczny problem optymalnego sterowania (2.3) zastępujemy problemem śledzenia trajektorii (tracking problem), który polega na poruszaniu się systemu wzdłuż optymalnej trasy* $(y_1^*, y_2^*, ..., y_N^*)$.

5.1.2 Wyznaczenie optymalnej trajektorii

Do wyznaczenia optymalnej trajektorii systemu (5.1) również zostanie wykorzystana technika programowania dynamicznego wstecz. Dla każdego momentu $0 \leq j \leq N-1$ definiujemy funkcję Bellmana postaci

$$V_j^N(\xi, y_0, ..., y_j) = E\left(g\left(\xi, y_j, y_{j+1}^*, w_{j+1}\right)\right. \\ \left. + \sum_{i=j+1}^{N-1} g\left(\xi, y_i^*, y_{i+1}^*, w_{i+1}\right) + h\left(y_N^*\right) \middle| \mathcal{F}_j^\xi \right), \tag{5.13}$$

natomiast dla momentu N przyjmujemy

$$V_N^N(\xi, y_0, ..., y_N) = h\left(y_N\right). \tag{5.14}$$

Funkcja Bellmana $V_j^N(\xi, y_0, ..., y_j)$ reprezentuje oczekiwane koszty poruszania się systemu (5.1) wzdłuż trajektorii $y_{j+1}^*, y_{j+2}^*, ..., y_N^*$, jeżeli znane są parametry ξ. Ze wzoru (5.13) otrzymujemy

$$\begin{aligned} V_j^N(\xi, y_0, ..., y_j) &= \inf_{y_{j+1}} E\left(g\left(\xi, y_j, y_{j+1}, w_{j+1}\right) + V_{j+1}^N(\xi, y_0, ..., y_j, y_{j+1}) \middle| \mathcal{F}_j^\xi\right) \\ &= E\left(g\left(\xi, y_j, y_{j+1}^*, w_{j+1}\right) + V_{j+1}^N(\xi, y_0, ..., y_j, y_{j+1}^*) \middle| \mathcal{F}_j^\xi\right), \end{aligned}$$

gdzie y_{j+1}^* jest optymalnym stanem w chwili $j+1$, jeżeli system (5.1) pokonał trajektorię $y_0, ..., y_j$. Korzystając ze wzorów (5.6) oraz (5.13) dla dowolnego momentu $0 \leq j \leq N-1$ warunek konieczny wyznaczenia optymalnego stanu y_{j+1}^* (jeżeli system pokonał trasę $y_0, ..., y_j$) możemy przedstawić w postaci

$$E\left(\bigtriangledown_{y_{j+1}} g\left(\xi, y_j, y_{j+1}^*, w_{j+1}\right) + \bigtriangledown_{y_{j+1}} V_{j+1}^N(\xi, y_0, ..., y_j, y_{j+1}^*) \,|\mathcal{F}_j\right) = \mathbf{0}. \tag{5.15}$$

Zatem

$$\begin{aligned} &E\left(\bigtriangledown_{y_{j+1}} g\left(\xi, y_j, y_{j+1}, w_{j+1}\right) + \bigtriangledown_{y_{j+1}} V_{j+1}^N(\xi, y_0, ..., y_j, y_{j+1}) \,|\mathcal{F}_j\right) \\ &= \int\int \left(\bigtriangledown_{y_{j+1}} g\left(\xi, y_j, y_{j+1}, x\right) + \bigtriangledown_{y_{j+1}} V_{j+1}^N(\xi, y_0, ..., y_j, y_{j+1})\right) \\ &\times \gamma\left(x, \bar{0}, I_m\right) P\left(d\xi \,|\mathcal{F}_j\right) dx, \end{aligned}$$

gdzie $\gamma\left(x, \bar{0}, I_m\right)$ jest funkcją gęstości rozkładu normalnego $\boldsymbol{N}(\bar{0}, I_m)$, natomiast $P\left(d\xi \,|\mathcal{F}_j\right)$ oznacza warunkowy rozkład ξ względem $\sigma-$ciała $\mathcal{F}_j$.

Poniżej, tak jak w poprzednich rozdziałach tej książki oraz w pracach np. [7]–[13], [38], [91], [92] przedstawiony zostanie algorytm wyznaczenia optymalnej trajektorii $\left\{y_j^*\right\}_{1\leq j\leq N}$ dla systemu (5.1). Wykorzytuje on metodę programowania dynamicznego wstecz, gdzie za pomocą techniki teleskopowej konstruowane są funkcje Bellmana (5.13).

Algorytm wyznaczania optymalnego sterowania.

1. Definiujemy

$$V_N^N(\xi, y_0, ..., y_N) = h(y_N) \ \text{ oraz kładziemy } \ j = N.$$

2. Kładziemy

$$j = j - 1.$$

3. Wyznaczamy

$$Z_j^N(y_0, ..., y_j, y_{j+1}) = \int\int \Big(\bigtriangledown_{y_{j+1}} g\left(\xi, y_j, y_{j+1}, x\right) \\ + \bigtriangledown_{y_{j+1}} V_{j+1}^N\left(\xi, y_0, ..., y_j, y_{j+1}\right)\Big)\, \gamma\left(x, \bar{0}, I_m\right) P\left(d\xi \left|\mathcal{F}_j\right.\right) dx.$$

4. Szukamy optymalnego stanu y_{j+1}^* spełniającego równanie

$$Z_j^N(y_0, ..., y_j, y_{j+1}^*) = \mathbf{0}.$$

5. Następnie wyznaczamy wartości funkcji Bellmana

$$V_j^N(\xi, y_0, ..., y_j) = \int g\left(\xi, y_j, y_{j+1}^*, x\right) \gamma\left(x, \bar{0}, I_m\right) dx \\ + V_{j+1}^N\left(\xi, y_0, ..., y_j, y_{j+1}^*\right).$$

6. Jeżeli $j = 0$, to zatrzymujemy, w przeciwnym razie wracamy do punktu 2.

Dla przypadku pełnej informacji o systemie algorytm jest podobny, wystarczy tylko warunkowy rozkład $P\left(d\xi \left|\mathcal{F}_j\right.\right)$ zastąpić rozkładem punktowym.

Uwaga 5.3 *W momencie $j = 0, 1, ..., N-1$ system (5.1) znajduje się w stanie y_j oraz przebył trajektorię $y_0, ..., y_j$. Korzystając z warunku (5.15) dla momentu $j+1$ wyznaczamy optymalny stan y^*_{j+1}, na który następnie system (5.1) jest naprowadzany poprzez zastosowanie sterowania (5.12). W kolejnej chwili $j+1$ odczytujemy stan systemu y_{j+1} oraz wyznaczamy stan optymalny y^*_{j+2} itd.*

Do wyznaczenia optymalnej trajektorii w warunkach pełnej i niepełnej informacji o systemie (5.1) korzystamy z równań (5.6)–(5.7). W warunkach pełnej informacji wystarczy przyjąć, że wektor ξ ma rozkład punktowy. Dla dowolnego momentu $j = 0, 1, .., N-1$ optymalny stan y^*_{j+1} w chwili przyszłej wyznaczamy na podstawie przebytej trajektorii $(y_0, ..., y_j)$ i wartości wektora ξ oraz przyjmujemy

$$y^*_{j+1} \stackrel{def}{=} \phi_j\left(\xi, y_0, ..., y_j\right). \tag{5.16}$$

W przypadku niepełnej informacji o systemie (wektor parametrów $\xi : \Omega \to \mathbb{R}^k$ jest wektorem losowym oraz nie jest znany sterującemu) optymalny stan y^*_{j+1}, $0 \le j \le N-1$ wyznaczamy na postawie trajektorii $(y_0, ..., y_j)$ i warunkowego rozkładu $P\left(d\xi \left| \mathcal{F}_j \right.\right)$ jako

$$\bar{y}^*_{j+1} \stackrel{def}{=} \bar{\phi}_j\left(y_0, ..., y_j\right) = E\left(\phi_j\left(\xi, y_0, ..., y_j\right) \left| \mathcal{F}_j \right.\right). \tag{5.17}$$

W momencie $j = 0$ korzystając ze wzorów (5.16)–(5.17) możemy wyznaczyć ciągi stanów $\left\{y^*_j\right\}_{1 \le j \le N}$ lub $\left\{\bar{y}^*_j\right\}_{1 \le j \le N}$, które powinien naśladować system w warunkach pełnej lub niepełnej informacji odpowiednio.

Definicja 5.1 *Ciąg stanów $\{y^*_i\}_{1 \le i \le N}$ postaci*

$$y^*_{j+1} \stackrel{def}{=} \phi_j\left(\xi, y_0, y^*_1 ..., y^*_j\right)$$

dla $0 \le j \le N-1$ nazywamy optymalną trajektorią systemu (5.1) w warunkach pełnej informacji, natomiast ciąg $\{\bar{y}^*_i\}_{1 \le i \le N}$ postaci

$$\bar{y}^*_{j+1} \stackrel{def}{=} E\left(\phi_j\left(\xi, y_0, y^*_1, ..., y^*_j\right) \left| \mathcal{F}_0 \right.\right)$$

dla $0 \leq j \leq N-1$ *nazywamy oczekiwaną optymalną trajektorią systemu (5.1) w warunkach niepełnej informacji.*

W warunkach niepełnej informacji trajektoria $\{\bar{y}_i^*\}_{1 \leq i \leq N}$ wyznaczona za pomocą wzoru (5.17) różni się od trajektorii $\{\hat{y}_i^*\}_{1 \leq i \leq N}$, gdzie $\hat{y}_{j+1}^* \overset{def}{=} \phi_j\left(E\left(\xi\left|\mathcal{F}_j\right.\right), y_0, y_1, ..., y_j\right)$ dla $0 \leq j \leq N-1$, ponieważ $\bar{\phi}_j\left(y_0, ..., y_j\right) \neq \phi_j\left(E\left(\xi\left|\mathcal{F}_j\right.\right), y_0, ..., y_j\right)$ dla $0 \leq j \leq N-1$. Natomiast jak zobaczymy poniżej, dla niektórych systemów liniowych z kwadratowym wskaźnikiem jakości trajektorie $\left\{\bar{y}_j^*\right\}_{1 \leq j \leq N}$ i $\left\{\hat{y}_j^*\right\}_{1 \leq j \leq N}$ są identyczne.

Uwaga 5.4 *W warunkach pełnej informacji oczekiwany koszt sterowania systemem (5.1) wynosi* $V_0^N(\xi, y_0)$. *W warunkach niepełnej informacji o systemie oczekiwany koszt realizacji celu (koszty sterowania i uczenia się jednocześnie) jest równy*

$$\bar{V}_j^N(y_0) = E\left(V_0^N(\xi, y_0)\right| \mathcal{F}_0\left.\right).$$

W warunkach niepełnej informacji o systemie oczekiwany koszt uczenia na ścieżce $\left\{\bar{y}_j^*\right\}_{1 \leq j \leq N}$ *podczas realizacji celu wynosi*

$$C_0 = \bar{V}_0^N(y_0) - V_0^N(\xi, y_0).$$

5.2 Optymalna trajektoria dla zadania liniowo-kwadratowego

5.2.1 Optymalna trajektoria dla zadania z ustalonym horyzontem

Na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ definiujemy ciąg niezależnych wektorów losowych w_1,..., w_N o rozkładzie normalnym $\mathbf{N}(\bar{0}, I_m)$, gdzie $\bar{0} \in \mathbb{R}^m$ jest wektorem zerowym, natomiast $I_m \in \mathbb{R}^{m \times m}$ macierzą jednostkową. Również definiujemy ciąg σ–ciał $\mathcal{F}_j = \sigma\left(y_0\right) \vee \sigma\left\{w_i : i = 1, 2, ..., j\right\}$ dla $0 \leq j \leq N$ oraz przyjmujemy $\mathcal{F} = \mathcal{F}_N$.

Poniżej rozważamy problem sterowania stochastycznym systemem liniowym o równaniu stanu

$$y_{i+1} = Ay_i - Cu_i + d + \sigma w_{i+1}, \tag{5.18}$$

gdzie $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{n\times l}$, $\sigma \in \mathbb{R}^{n\times m}$, $d \in \mathbb{R}^n$ oraz $\|A\| < \infty$, $\|C\| < \infty$, $\|\sigma\| < \infty$, natomiast wektor $u_j \in \mathbb{R}^l$ oznacza działanie sterujące. Dodatkowo zakładamy, że $det\left(C^T C\right) \neq 0$.

Głównym celem jest przeprowadzenie systemu (5.18) ze stanu y_0 do stanu $a = \mathrm{col}\,(0, 0, ..., 0)$ – początku układu współrzędnych. Wyznaczenie optymalnego sterowania dla zadania (2.3) ze wskaźnikiem jakości (2.20) zostało przedstawione w rozdziale drugim. Rozwiązując zadanie (2.3) wyznaczamy prawa sterowania oraz realizujemy cel sterowania najtańszym kosztem.

Z drugiej strony, zamiast wyznaczać prawa sterowania możemy określić trajektorię, poruszając się wzdłuż której zostanie zrealizowany cel (2.3). Podobnie jak dla wskaźnika jakości (2.20), sterując systemem (5.18) w kolejnych N krokach przyjmujemy, że dla momentów $i = 0, 1, ..., N-1$ koszty sterowania wynoszą $u_i^T R_i u_i$, natomiast dla momentu N straty związane z nietrafieniem do początku układu współrzędnych wynoszą $y_N^T Q y_N$. Poniżej zakładamy, że macierze R_i, $i \in \{0, 1, ..., N-1\}$ oraz Q są dodatnio określone.

Jeżeli przeprowadzamy system liniowy (5.18) ze stanu y_i do stanu y_{i+1}, $i = 0, 1, ..., N-1$, to sterowanie jest równe

$$u_i = -\left(C^T C\right)^{-1} C^T \left(y_{i+1} - Ay_i - d - \sigma w_{i+1}\right). \tag{5.19}$$

Uwaga 5.5 *Ze wzoru (5.19) wynika, że dla dowolnego momentu $i = 0, 1, ..., N-1$ oraz ustalonego przyszłego stanu y_{i+1} sterowanie u_i jest wektorem losowym (zależy od zaburzeń zewnętrznych w_{i+1}). W celu przeprowadzenia systemu (5.18) ze stanu y_i w okolice (do otoczenia) stanu y_{i+1}*

należy zastosować sterowanie

$$\bar{u}_i = E\left(u_i\left|\mathcal{F}_i\right.\right) = -\left(C^T C\right)^{-1} C^T \left(y_{i+1} - Ay_i - d\right).$$

Niech $y = (y_1, y_2, ..., y_N)$ oznacza dopuszczalną trajektorię dla systemu (5.18). Klasę wszystkich dopuszczalnych trajektorii oznaczamy przez Y. Korzystając z (5.19) wskaźnik jakości (2.20) zastępujemy wskaźnikiem

$$\begin{aligned} J^N(y) = E\Bigg(&\sum_{i=0}^{N-1} \left(y_{i+1} - Ay_i - d - \sigma w_{i+1}\right)^T \\ &\times K_i\left(y_{i+1} - Ay_i - d - \sigma w_{i+1}\right) + y_N^T Q y_N\Bigg), \end{aligned} \tag{5.20}$$

gdzie dla $i = 0, 1, ..., N-1$

$$K_i = C\left(C^T C\right)^{-1} R_i \left(C^T C\right)^{-1} C^T. \tag{5.21}$$

Rozwiązując zadanie (5.5) wyznaczamy trajektorię, po której powinien poruszać się system (5.18). Wartość wskaźnika jakości (5.20) przedstawia całkowity koszt, który składa się z kosztów transformacji systemu (5.18) wzdłuż trajektorii $y_0, y_1, ..., y_N$ i strat związanych z nietrafieniem do początku układu współrzędnych (celu $a = \mathrm{col}\,(0, 0, ..., 0)$). Problem sterowania liniowym systemem stochastycznym (5.18) zastępujemy problemem wyznaczenia optymalnej trajektorii $y^* = (y_1^*, ..., y_N^*)$. Rozwiązaniem zadania (5.5) jest optymalna trajektoria, na której osiągamy najmniejsze koszty całkowite. Zatem realizując główny cel (przeprowadzenie ze stanu y_0 do stanu a) system powinien poruszać się wzdłuż trajekrorii $(y_1^*, ..., y_N^*)$.

Aby nie przysłaniać głównej problematyki związanej z wyznaczeniem optymalnej trajektorii dla systemu (5.18) poniżej przyjęto, że przesunięcie d jest stałe. W pracy [62] przedstawione zostały rozwiązania zadania, gdy przesunięcie jest wektorem losowym, natomiast w pracy [66] przedstawiono rozwiązania dla specjalnego przypadku, gdy cel sterowania jest wektorem losowym.

Do rozwiązania zadania (5.5) konstruujemy ciąg funkcji Bellmana. Dla dowolnego momentu $i = 0, 1, ..., N-1$ definiujemy

$$V_i^N(y_i) = \inf_{y_{i+1}} E\Big((y_{i+1} - y_i - C\xi - \sigma w_{i+1})^T K_i \times (y_{i+1} - y_i - C\xi - \sigma w_{i+1}) + V_{i+1}^N(y_{i+1}) \Big| \mathcal{F}_i \Big) \tag{5.22}$$

z warunkiem początkowym

$$V_N^N(y_N) = y_N^T Q y_N. \tag{5.23}$$

Całkowity oczekiwany koszt transformacji systemu (5.18) ze stanu y_0 do początku układu współrzędnych wynosi

$$\inf_{y \in Y} J^N(y) = V_0^N(y_0).$$

Twierdzenie poniżej przedstawia sposób wyznaczenia optymalnej trajektorii.

Twierdzenie 5.2 *Niech ciągi macierzy* $\left\{H_j^N\right\}_{0 \le j \le N}$, $\left\{L_j^N\right\}_{0 \le j \le N}$, $\left\{M_j^N\right\}_{0 \le j \le N}$ *oraz ciąg* $\left\{\varphi_j^N\right\}_{0 \le j \le N}$ *będą dane wzorami*

$$H_j^N = A^T \left(K_j - K_j^T \left(\bar{K}_j + \bar{H}_{j+1}^N \right)^{-1} K_j \right) A, \tag{5.24}$$

$$L_j^N = A^T \left(K_j - K_j^T \left(\bar{K}_j + \bar{H}_{j+1}^N \right)^{-1} \left(K_j - L_{j+1}^N \right) \right), \tag{5.25}$$

$$M_j^N = M_{j+1}^N + K_j - \left(K_j - L_{j+1}^N \right)^T \left(\bar{K}_j + \bar{H}_{j+1}^N \right)^{-1} \left(K_j - L_{j+1}^N \right), \tag{5.26}$$

$$\varphi_j^N = \varphi_{j+1}^N + tr\left(K_j \sigma \sigma^T \right), \tag{5.27}$$

gdzie $H_N^N = Q$, *natomiast* L_N^N, M_N^N *są macierzami zerowymi,* $\varphi_N^N = 0$, $\bar{K}_j = \frac{1}{2}\left(K_j + K_j^T\right)$, $\bar{H}_{j+1}^N = \frac{1}{2}\left(H_{j+1}^N + \left(H_{j+1}^N\right)^T\right)$ *dla* $0 \le j \le N-1$. *Jeżeli* $\det\left(\bar{K}_j + \bar{H}_{j+1}^N\right) \ne 0$ *dla* $j = 0, 1, ..., N-1$, *to optymalny stan systemu (5.18) dla momentu* $j+1$ *(na podstawie przebytej trajektorii* $\{y_s\}_{0 \le s \le j}$*) wynosi*

$$y_{j+1}^* = \left(\bar{K}_j + \bar{H}_{j+1}^N \right)^{-1} \left(K_j A y_j + \left(K_j - L_{j+1}^N \right) d \right). \tag{5.28}$$

Wartości funkcji Bellmana dla momentów $0 \le j \le N$ są równe

$$V_i^N(y_i) = y_j^T H_j^N y_j + 2y_j^T L_j^N d + d^T M_j^N d + \varphi_j^N. \tag{5.29}$$

Dowód. Najpierw wyznaczymy wartość funkcji Bellmana dla momentu $N-1$. Do momentu $N-1$ system (5.18) pokonał trajektorię $(y_0, y_1, ..., y_{N-1})$. Ze wzoru (5.22) otrzymujemy

$$\begin{aligned} V_{N-1}^N(y_{N-1}) &= \inf_{y_N}\left(y_N^T\left(K_{N-1} + H_{N-1}^N\right) y_N - 2y_N^T K_{N-1}(Ay_{N-1} + d)\right) \\ &+tr\left(K_{N-1}\sigma\sigma^T\right) + y_{N-1}^T A^T K_{N-1} A y_{N-1} + 2y_{N-1} A^T K_{N-1} d + d^T K_{N-1} d. \end{aligned} \tag{5.30}$$

Optymalny stan dla momentu N jest równy

$$y_N^* = \left(\bar{K}_{N-1} + \bar{H}_{N-1}^N\right)^{-1} K_{N-1}(Ay_{N-1} + d). \tag{5.31}$$

Podstawiając (5.31) do (5.30) otrzymujemy

$$\begin{aligned} V_{N-1}^N(y_{N-1}) = {}& y_{N-1}^T A^T \left(K_{N-1} - K_{N-1}^T\left(\bar{K}_{N-1} + \bar{H}_{N-1}^N\right)^{-1} K_{N-1}\right) A^T y_{N-1} \\ &+ 2y_{N-1}^T A^T \left(K_{N-1} - K_{N-1}^T\left(\bar{K}_{N-1} + \bar{H}_{N-1}^N\right)^{-1} K_{N-1}\right) d \\ &+ d^T \left(K_{N-1} - K_{N-1}^T\left(\bar{K}_{N-1} + \bar{H}_{N-1}^N\right)^{-1} K_{N-1}\right) d + tr\left(K\sigma\sigma^T\right) \\ = {}& y_{N-1}^T H_{N-1}^N y_{N-1} + 2y_{N-1}^T L_{N-1}^N d + d^T M_{N-1}^N d + \varphi_{N-1}^N. \end{aligned}$$

Zakładamy, że wzór (5.29) jest prawdziwy dla dowolnego momentu $j+1$, gdzie $0 \le j \le N-1$. Korzystając z własności warunkowej wartości oczekiwanej oraz (5.22) wyznaczamy wartość funkcji Bellmana dla momentu j

$$\begin{aligned} V_j^N(y_j) &= \inf_{y_{j+1}} E\left((y_{j+1} - Ay_j - d - \sigma w_{j+1})^T K_j (y_{j+1} - Ay_j - d - \sigma w_{j+1})\right. \\ &\left. +y_{j+1}^T H_{j+1}^N y_{j+1} + 2y_{j+1}^T L_{j+1}^N d + d^T M_{j+1}^N d + \varphi_{j+1}^N \right| \mathcal{F}_j\Big) \\ &= \inf_{y_{j+1}}\left(y_{j+1}^T\left(K_j + H_{j+1}^N\right) y_{j+1} - 2y_{j+1}^T\left(K_j A y_j + (K - L_{j+1}) d\right)\right) + \varphi_{j+1}^N \\ &+ d^T\left(M_{j+1}^N + K_j\right) d + y_j^T A^T K_j A y_j + 2y_j^T A^T K_j d + tr\left(K_j\sigma\sigma^T\right). \end{aligned} \tag{5.32}$$

Zatem optymalny stan systemu (5.18) dla momentu $j+1$ spełnia równanie (5.28). Podstawiając (5.28) do (5.32) otrzymujemy, że wartość funkcji Bellmana dla momentu j jest równa

$$\begin{aligned}V_j^N(y_j) &= y_j^T A^T \left(K_j - K_j^T \left(\bar{K}_j + \bar{H}_{j+1}^N\right)^{-1} K_j\right) A y_j + \varphi_{j+1}^N + tr\left(K\sigma\sigma^T\right)\\ &\quad + 2 y_j^T A^T \left(K_j - K_j^T \left(\bar{K}_j + \bar{H}_{j+1}^N\right)^{-1} \left(K - L_{j+1}^N\right)\right) d\\ &\quad + d^T \left(M_{j+1}^N + K_j - \left(K_j - L_{j+1}^N\right)^T \left(\bar{K}_j + \bar{H}_{j+1}^N\right)^{-1} \left(K_j - L_{j+1}^N\right)\right) d\\ &= y_j^T H_j^N y_j + 2 y_j^T L_j^N d + d^T M_j^N d + \varphi_j^N,\end{aligned}$$

co dowodzi prawdziwości twierdzenia. ■

Uwaga 5.6 *Z twierdzenia 5.2 wynika, że dla systemu liniowego (5.18) ze wskaźnikiem jakości (5.20) zarówno funkcja Bellmana (5.22), jak i optymalny stan systemu dla momentu kolejnego $j+1$ zależą od stanu y_j (nie zależą od stanów $y_0, y_1, ..., y_{j-1}$), $j = 0, 1, .., N-1$.*

Uwaga 5.7 *Załóżmy że ciągi macierzy $\left\{H_j^N\right\}_{0\le j\le N}, \left\{L_j^N\right\}_{0\le j\le N}$ spełniają równania (5.24) i (5.25) odpowiednio. W momencie $j=0$ za pomocą wzoru (5.28) możemy określić optymalną trajektorię, która powinna być naśladowana przez system (5.18). Jest to ciąg stanów $\left\{y_j^*\right\}_{1\le j\le N}$ postaci*

$$y_{j+1}^* = \left(\bar{K}_j + \bar{H}_{j+1}^N\right)^{-1} \left(K_j A y_j^* + \left(K_j - L_{j+1}^N\right) d\right)$$

dla $0 \le j \le N-1$ z warunkiem początkowym y_0.

Jeżeli $d = \mathrm{col}(0, 0, ..., 0) \in \mathbb{R}^n$ oraz $K_i = K$ ($R_i = R$), $0 \le i \le N-1$, to optymalna trajektoria $\left\{y_j^*\right\}_{1\le j\le N}$ jest dana wzorem

$$y_j^* = \prod_{i=0}^{j} \left(\left(\bar{K} + \bar{H}_{i+1}^N\right)^{-1} K A\right) y_0 \tag{5.33}$$

dla $1 \le j \le N$.

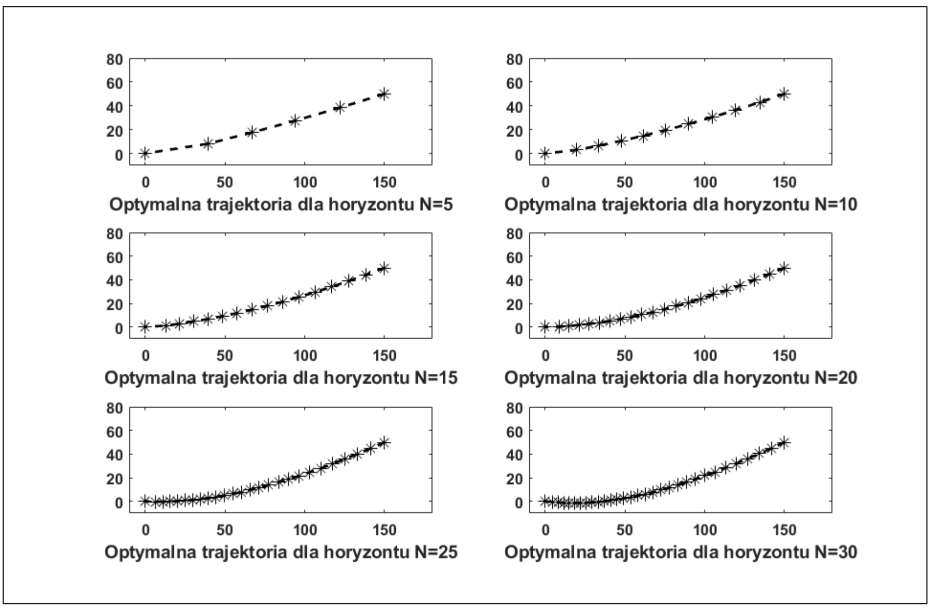
Uwaga 5.8 *Jeżeli $K_j = K$ (równoważne z warunkiem $R_j = R$) dla $0 \le j \le N-1$, to z (5.24) – (5.25) wynika, że $H_j^N = H_0^{N-j}$, $L_j^N = L_0^{N-j}$,*

$M_j^N = M_0^{N-j}$ *oraz* $\varphi_j^N = (N-j)\, tr\left(K\sigma\sigma^T\right)$. *Zatem ze wzoru (5.29) otrzymujemy*

$$V_j^N(y) = V_0^{N-j}(y)$$

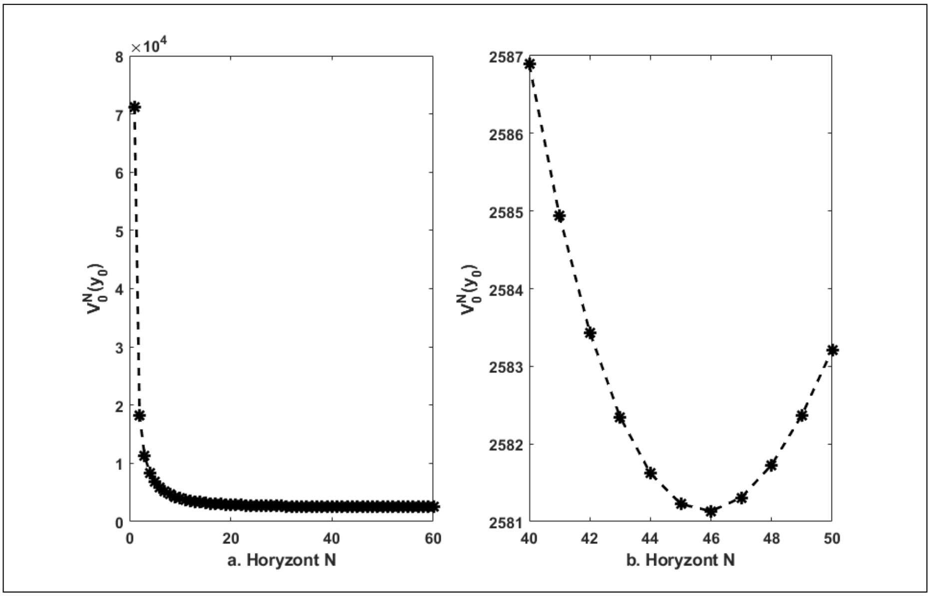
dla $y \in \mathbb{R}^n$.

Przykład 5.1 *Załóżmy, że liniowy system stochastyczny jest określony za pomocą równania (5.18). Wyznaczymy optymalną trajektorię, po której powinien się poruszać system. Poniżej przyjmujemy macierze* $R = \begin{bmatrix} 1.5 & 0.25 \\ 0.25 & 2 \end{bmatrix}$, $Q = \begin{bmatrix} 2.5 & 0.2 \\ 0.2 & 4.8 \end{bmatrix}$, $A = \begin{bmatrix} 1.05 & 0 \\ 0 & 0.98 \end{bmatrix}$, $C = \begin{bmatrix} 1.2 & 0.1 \\ 0.5 & 2 \end{bmatrix}$, $\sigma = \begin{bmatrix} 1.05 & 0 \\ 0 & 1.2 \end{bmatrix}$, *stan początkowy* $y_0 = \begin{bmatrix} 150 \\ 50 \end{bmatrix}$ *oraz* $d = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. *Bezpośrednio stosując wzór (5.33) w momencie* $j = 0$ *wyznaczymy ciąg stanów* $\left\{y_j^*\right\}_{1\le j\le N}$*, który reprezentuje optymalną trajektorię. Rysunek 5.1 przedstawia optymalne trajektorie dla horyzontów* $N = \{5, 10, 15, 20, 25, 30\}$.



Rysunek 5.1: Optymalna trajektoria dla systemu liniowego.

Rysunek 5.2a przedstawia wartości funkcji Bellmana $V_0^N(y_0)$ dla horyzontu $N = 1, 2, ..., 60$, które reprezentują oczekiwane koszty poruszania się systemu (5.18) wzdłuż optymalnej trajektorii. Z rysunku 5.2b widzimy, że wartość najmniejszą tych kosztów osiągamy dla horyzontu $N = 46$.



Rysunek 5.2: Zależność kosztu transformacji systemu wzdłuż optymalnej trajektorii od horyzontu N.

5.2.2 Optymalna trajektoria dla zadania z losowym horyzontem niezależnym od stanów systemu

W rozdziale 3 omówiony został problem wyznaczenia optymalnego sterowania systemem, dla przypadku gdy horyzont nie jest znany oraz nie zależy od stanów systemeu. Poniżej przedstawiona zostanie problematyka wyznaczenia optymalnej trajektorii dla zadania liniowo-kwadratowego, gdzie horyzont sterowania również nie jest znany sterującemu oraz nie zależy od stanów systemu. Przyjmujemy, że horyzont sterowania jest zmienną losową $\tau : \Omega \to \{0, 1, 2, ..., N\}$ niezależną od stanów systemu (5.18) o rozkładzie $P(\tau = k) = p_k$, $0 \le p_k \le 1$ dla $0 \le k \le N$, $\sum_{i=0}^{N} p_i = 1$ oraz $\mathcal{F}^\tau = \sigma\{\tau\}$.

Dla momentów $i = 0, 1, ..., \tau - 1$ koszty sterowania systemem (5.18) są równe $u_i^T R u_i$ oraz ze wzorów (5.19) i (5.21) otrzymujemy, że odpowiadające im koszty przeprowadzenia systemu ze stanu y_i do stanu y_{i+1} wynoszą

$$(y_{i+1} - Ay_i - d - \sigma w_{i+1})^T K (y_{i+1} - Ay_i - d - \sigma w_{i+1})^T,$$

gdzie $K = C\left(C^T C\right)^{-1} R \left(C^T C\right)^{-1} C^T$. Wartość funkcjonału dziedziczenia w momencie τ jest równa $(y_\tau - a)^T Q (y_\tau - a)$.

Niech $y = (y_1, y_2, ..., y_\tau)$ oznacza trajektorię, po której będzie poruszał się system (5.18). Łączny koszt przeprowadzenia systemu wzdłuż trajektorii y oraz dziedziczenia w momencie τ wynosi

$$\begin{aligned} J^\tau(y) = E\Bigg(\sum_{i=0}^{\tau-1} &(y_{i+1} - Ay_i - d - \sigma w_{i+1})^T \\ &\times K(y_{i+1} - Ay_i - d - \sigma w_{i+1}) + y_\tau^T Q y_\tau \Big| \mathcal{F}^\tau\Bigg). \end{aligned} \tag{5.34}$$

Uwaga 5.9 *Jeżeli system (5.18) zostanie zatrzymany w momencie zero (tzn. $\tau(\omega) = 0$), to ponosimy tylko koszt dziedziczenia $y_0^T Q y_0$.*

Wskaźnik jakości (5.34) jest zmienną losową. Zadanie wyznaczenia optymalnej trajektorii dla losowego horyzontu niezależnego od stanu systemu sprowadzamy do zadania z ustalonym horyzontem. W tym celu oszacujemy wartość oczekiwaną wskaźnika jakości (5.34) oraz przyjmujemy $\bar{J}^N(y) = EJ^\tau(y)$. Ze wzoru na prawdopodobieństwo całkowite mamy

$$\begin{aligned} \bar{J}^N(y) = E\Bigg(\sum_{i=0}^{\tau-1} \Big(&(y_{i+1} - Ay_i - d - \sigma w_{i+1})^T \\ &\times K_i(y_{i+1} - Ay_i - d - \sigma w_{i+1}) + y_i^T Q_i y_i\Big) + y_N^T Q_N y_N\Bigg), \end{aligned} \tag{5.35}$$

gdzie

$$K_i = KP(\tau > i) = K \sum_{i=j+1}^{N} p_i, \tag{5.36}$$

$$Q_i = P(\tau = i)\, Q = p_i Q. \tag{5.37}$$

Głównym celem jest wyznaczenie optymalnej trajektorii, poruszając się wzdłuż której osiągamy najmniejsze koszty łączne. Zatem należy rozwiązać zadanie

$$\inf_{y \in Y} \bar{J}^N(y), \tag{5.38}$$

tzn. wyznaczyć taką trajektorię $y^* = (y_1^*, ..., y_N^*)$, dla której infimum jest osiągnięte.

Dla dowolnego momentu $i = 0, 1, ..., N-1$ definiujemy funkcje Bellmana

$$\bar{V}_i^N(y_i) = \inf_{y_{i+1}} E\Big((y_{i+1} - y_i - C\xi - \sigma w_{i+1})^T K_i \\ \times (y_{i+1} - y_i - C\xi - \sigma w_{i+1}) + y_i^T Q_i y_i + \bar{V}_{i+1}^N(y_{i+1}) \Big| \mathcal{F}_i\Big) \tag{5.39}$$

z warunkiem początkowym

$$\bar{V}_N^N(y_N) = y_N^T Q_N y_N. \tag{5.40}$$

Oczekiwany koszt transformacji systemu (5.18) w czasie losowym ze stanu y_0 do początku układu współrzędnych wynosi $\bar{J}^N(y^*) = \bar{V}_0^N(y_0)$.

W przypadku jeżeli horyzont τ jest zmienną losową o rozkładzie dyskretnym, to system (5.18) powinien być przeprowadzany wzdłuż trajektorii $(y_1^*, ..., y_\tau^*) \subset y^* = (y_1^*, ..., y_N^*)$, gdzie y^* jest rozwiązaniem zadania (5.38). Sposób wyznaczenia optymalnej trajektorii przedstawia twierdzenie poniżej.

Twierdzenie 5.3 *Niech ciągi macierzy* $\left\{\dot{H}_j^N\right\}_{0 \le j \le N}$, $\left\{\dot{L}_j^N\right\}_{0 \le j \le N}$, $\left\{\dot{M}_j^N\right\}_{0 \le j \le N}$ *będą dane wzorami*

$$\dot{H}_j^N = Q_j + A^T\left(K_j - K_j^T\left(\bar{K}_j + \bar{\dot{H}}_{j+1}^N\right)^{-1} K_j\right)A, \tag{5.41}$$

$$\dot{L}_j^N = A^T\left(K_j - K_j^T\left(\bar{K}_j + \bar{\dot{H}}_{j+1}^N\right)^{-1}\left(K_j - \dot{L}_{j+1}^N\right)\right), \tag{5.42}$$

$$\dot{M}_j^N = \dot{M}_{j+1}^N + K_j - \left(K_j - \dot{L}_{j+1}^N\right)^T\left(\bar{K}_j + \bar{\dot{H}}_{j+1}^N\right)^{-1}\left(K_j - \dot{L}_{j+1}^N\right), \tag{5.43}$$

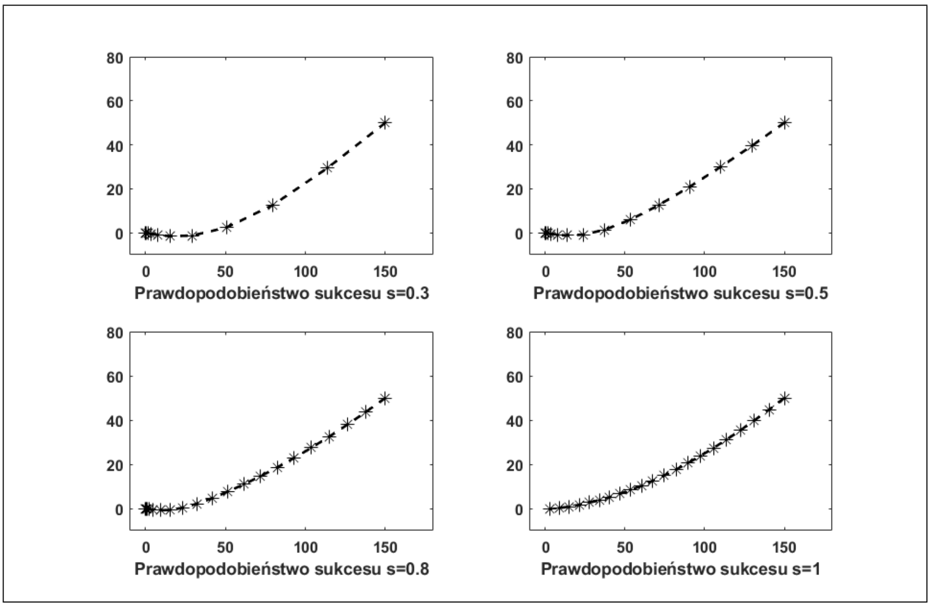
gdzie $\dot{H}_N^N = Q$, *natomiast* $\dot{L}_N^N$, $\dot{M}_N^N$ *są macierzami zerowymi, ciąg* $\left\{\varphi_j^N\right\}_{0\le j\le N}$ *spełnia równanie (5.27) oraz* $\bar{K}_j = \frac{1}{2}\left(K_j + K_j^T\right)$, $\bar{\dot{H}}_{j+1}^N = \frac{1}{2}\left(\dot{H}_{j+1}^N + \left(\dot{H}_{j+1}^N\right)^T\right)$ *dla* $0 \le j \le N-1$. *Jeżeli* $\det\left(\bar{K}_j + \bar{\dot{H}}_{j+1}^N\right) \neq 0$ *dla* $j = 0,1,...,N-1$, *to optymalny stan systemu (5.18) dla momentu* $j+1$ *wynosi*

$$y_{j+1}^* = \left(\bar{K}_j + \bar{\dot{H}}_{j+1}^N\right)^{-1}\left(K_jAy_j + \left(K_j - \dot{L}_{j+1}^N\right)d\right). \tag{5.44}$$

Wartości funkcji Bellmana dla momentów $0 \le j \le N$ *są równe*

$$\bar{V}_i^N(y_i) = y_j^T\dot{H}_j^N y_j + 2y_j^T\dot{L}_j^N d + d^T\dot{M}_j^N d + \varphi_j^N. \tag{5.45}$$

Dowód twierdzenia przebiega w podobny sposób jak dowód twierdzenia 5.2. Porównując funkcje Bellmana $V_i^N(y)$ i $\bar{V}_i^N(y)$ zdefiniowane za pomocą wzorów (5.22) i (5.39) otrzymujemy tezę twierdzenia.



Rysunek 5.3: Optymalna trajektoria dla liniowego systemu stochastycznego z losowym horyzontem.

Przykład 5.2 *Rozważamy problem sterowania liniowym systemem stochastycznym, który jest określony za pomocą równania (3.71). Zadanie polega na przeprowadzeniu systemu (5.18) ze stanu* y_0 *do początku układu współrzędnych w czasie losowym* $\tau : \Omega \to \mathbb{N}_0$*, gdzie horyzont sterowania* τ *jest zmienną losową o rozkładzie dwumianowym z prawdopodobieństwem sukcesu* $0 \leq s \leq 1$ *oraz* $N = 20$*. Przyjmujemy, że macierze* R*,* Q*,* A*,* C*,* σ *oraz stan* y_0 *są identyczne jak dla przykładu 5.1. Bezpośrednio stosując wzór (5.44) w momencie* $j = 0$ *wyznaczymy ciąg stanów* $\left\{y_j^*\right\}_{1 \leq j \leq N}$*, który reprezentuje optymalną trajektorię. Rysunek 5.3 przedstawia optymalne trajektorie dla prawdopodobieństw sukcesu* $s = 0.3, 0.5, 0.8, 1$*.*

Analizując zachowanie systemu (5.18) z losowym horyzontem o rozkładzie dwumianowym możemy stwierdzić, że im mniejsze jest prawdopodobieństwo sukcesu s tym odległości pomiędzy stanami y_i *i* y_{i+1} *są zdecydowanie większe w początkowych momentach oraz wraz z upływem czasu maleją. W przypadku gdy horyzont jest ustalony (prawdopodobieństwo* $s = 1$*), to system równomiernie pokonuje trajektorię ze stanu* y_0 *do początku układu.*

Uwaga 5.10 *Jeżeli* τ *jest zmienną losową o przeliczalnej liczbie realizacji, to dla ustalonego* $\varepsilon > 0$ *wyznaczamy maksymalnie możliwy horyzont* N_ε *określony wzorem (3.82). Optymalną trajektorię* $\{y_i^*\}_{1 \leq i \leq N_\varepsilon}$ *wyznaczamy rozwiązując zadanie (5.38) z ustalonym horyzontem* N_ε*. Jako* ε*–optymalną trajektorię przyjmujemy* $\left\{y_1^*, y_2^*, ..., y_{N_\varepsilon}^*, y_{N_\varepsilon}^*, ...\right\}$*.*

Bibliografia

[1] Aoki M., *Optimization of stochastic systems*, Academic Press, 1967.

[2] Abouzaid B., Achhab M.E., Wertz V., *Feedback stabilization of infinite-dimensional linear systems with constraints on control and its rate*, "European Journal of Control", 2011 vol. 17(2), p. 183–190.

[3] Astrom K. J., *Introduction to stochastic control theory*, Academic Press, New York 1970.

[4] Astrom K. J., Wittenmark B., *Adaptive control*, Addison-Wesley Publishing Company, 1995.

[5] Azuma S., Sakar, M.S., Pappas G.J., *Stochastic source seeking by mobile robots*, "IEEE Transactions on Automatic Control", 2012 vol. 57(9), p. 2308–2321.

[6] Banek T., Kozłowski E., *Adaptive control of system entropy*, "Control and Cybernetics", 2006 vol. 35(2), p. 279–289.

[7] Banek T., Kozłowski E., *Active and passive learning in control processes application of the entropy concept*, "Systems Sciences", 2005 vol. 31(2), p. 29–44.

[8] Banek T., Kozłowski E., *Adaptive control with random horizon*, "Annales Informatica", 2005 vol. 3, p. 5–14.

[9] Banek T., Kozłowski E., *Adaptacyjne sterowanie terminalne z losowym horyzontem*, [W:] *XV Krajowa Konferencja Automatyki*, pod red. Bubnicki Z., Kulikowski R., Kacprzyk J., t. 1, ARGRAF, Warszawa 2005, s. 257–261.

[10] Banek T., Kozłowski E., *Sterowanie adaptacyjne jako metoda samouczenia w zadaniach portfolio*, [W:] *Modelowanie preferencji a ryzyko'05*, pod red. Traskalik T., Ustroń 2005, s. 21–33.

[11] Banek T., Kozłowski E., *Self-learning jako zadanie sterowania dualnego*, [W:] *Systemowo-komputerowe wspomaganie zarządzania wiedzą*, pod red. Kulikowski R., Bubnicki Z., Kacprzyk J., Akademicka Oficyna Wydawnicza Exit, Warszawa 2006, s. 157–236.

[12] Banek T., Kozłowski E., *Active learning in discrete time stochastic systems*, [In:] *Knowledge-Based Intelligent System Advancements: Systemic and Cybernetic Approaches*, ed. by Jozefczyk J., Orski D., Information Science References, New York 2011, p. 350–371.

[13] Banek T., Kozłowski E., *LQG multi-jumps follower as a model of company reorganization*, [In:] *Process Control in Management*, ed. by Banek T., Toruń 2009, p. 13–26.

[14] Banek T., Kozłowski E., *Deterministic versus stochastic in the free horizon LQC problem*, [W:] *Aktualne problemy automatyki i robotyki*, pod red. Malinowski K., Józefczyk J., Świątek J., Wrocław 2014, s. 94–101.

[15] Banek T., Kowalik P., Kozłowski E., *Fisher information criterion in stopped self-learning of linear systems*, [In:] *Process Control in Management*, ed. by Banek T., Toruń 2009, p. 41–52.

[16] Banek T., Kulikowski R., *Adaptive control of Fisher information*, "Cybernetics and Systems Analysis", 2005 vol. 41(2), p. 224–232.

[17] Bania P., *Example for equivalence of dual and information-based optimal control*, "International Journal of Control", 2018, https://doi.org/10.1080/00207179.2018.1436775.

[18] Bania P., *Simple example of dual control problem with almost analytical solution*, [In:] *Proceedings of 19th Polish Control Conference*, Krakow, June 18-21, 2017, p. 55–64.

[19] Bania P., Baranowski J., *Bayesian estimator of a faulty state: Logarithmic odds approach*, [In:] *Proceedings of 22th International Conference on Methods and Models in Automation and Robotics, MMAR 2017*, Miedzyzdroje, 28-31 Aug. 2017, p. 253–257.

[20] Bania P., Baranowski J.,*Field Kalman Filter and its approximation*, [In:] *Proceedings of 55th IEEE Conference on Decision and Control*, Las Vegas, 12-14 December 2016, p. 2875–2880.

[21] Baranowski J., Bania P., Prasad I., Cong T., *Bayesian fault detection and isolation using Field Kalman Filter*, "EURASIP Journal on Advances in Signal Processing", 2017:79.

[22] Barbu V., Sritharan S.S., *Optimal stopping time problem for stochastic Navier-Stokes equations and infinite dimensional variational inequalities*, "Nonlinear Analysis: Theory, Methods and Applications", 2006 vol. 64(5), p. 1018–1024.

[23] Bellman R., *Adaptive Control Processes*, Princeton, 1961.

[24] Beneš V.E., Karatzas I., Rishel R., *The separation principle for a Bayesian adaptive control problem with no strict-sense optimal law*, [In:] *Applied stochastic analysis*, Stochastics Monographs vol. 5, ed. by. Davis M.H.A., Elliott R.J., 1991, p. 121–156.

[25] Benth F. E., Reikvam K., *A connection between singular stochastic control and optimal stopping*, "Applied Mathematics and Optimization", 2004 vol. 49, p. 27–41.

[26] Billingsley P., *Probability and measure*, John Wiley and Sons, 1995.

[27] Birkholc A., *Analiza matematyczna funkcji wielu zmiennych*, PWN, Warszawa 1986.

[28] Boetius F., Kohlmann M., *Connections between optimal stopping and singular stochastic control*, "Stochastic Processes and their Applications", 1998 vol. 77, p. 253–281.

[29] Borowkow A.A., *Rachunek prawdopodobieństwa*, PWN, Warszawa 1975.

[30] Bubnicki Z., *General approach to stability and stabilization for a class of uncertain discrete non-linear systems*, "International Journal of Control", 2000 vol. 73 (14), p. 1298–1306.

[31] Buonaguidi B., *A remark on optimal stopping problems*, "Journal of Applied Probability", 2015 vol. 52, p. 1187–1194.

[32] Chena Y., Edgarb T., Manousiouthakisa V., *On infinite-time nonlinear quadratic optimal control*, "Systems and Control Letters", 2004 vol.51 (3-4), p. 259–268.

[33] Chow Y.S., Robbins H., Siegmund D., *Great expectations: the theory of optimal stopping*, Houghton Mifflin 1971.

[34] Feldbaum A.A., *Dual control theory*, "Automation and Remote Control", 1960 vol. 21, p. 874–1033.

[35] Feldbaum A.A., *Dual control theory*, "Automation and Remote Control", 1961 vol. 22, p. 1–109.

[36] Feldbaum A.A., *Optimal Control Systems*, Academic Press 1965.

[37] Fichtenholtz G.M., *Rachunek różniczkowy i całkowy*, PWN, Warszawa 1985.

[38] Fleming W. H., Rishel R., *Deterministic and stochatic optimal control*, Springer-Verlag, Berlin 1975.

[39] Ford J.J, *Partially observed non-linear risk sensitive optimal stopping control for non-linear discrete time systems*, "Systems and Control letters", 2006 vol. 55, p. 770–776.

[40] Gelfand I. M., Fomin S. W., *Rachunek wariacyjny*, PWN, Warszawa 1979.

[41] Gola A., Kozłowski E., *Optimal production planning for a randon horizon*, "Applied Mechanics and Materials", 2015 vol. 791, p. 63–69.

[42] Guo D., Xu X., *Reflected BSDEs with random default time and related mixed optimal stopping-control problems*, "Acta Mathematicae Applicatae Sinica", 2013 vol. 29(1), p. 165–178.

[43] Guo X., Tomecek P., *Connection between singular control and optimal switching*, "SIAM Journal on Control and Optimization", 2008 vol. 47(1), p. 421 – 443.

[44] Harris L., Rishel R., *An algorithm for a solution of a stochastic adaptive linear quadratic optimal control problem*, "IEEE Transactions on Automatic Control", 1986 vol. 31(12), p. 1165–1170.

[45] Henderson V., Hobson D., *An explicit solution for an optimal stopping/optimal control problem with models an asset sale*, "The Annals of Applied Probability", 2008 vol. 18(5), p. 1681–1705.

[46] Huang J., Li X., Yong J., *A mixed linear quadratic optimal control problem with a controlled time horizon*, "Applied Mathematics and Optimization", 2014 vol. 70(1), p. 29–59.

[47] Ibragimov I.A., Rozanov Y.A., *Gaussian Random Processes*, Springer 1978.

[48] Jazwinski A.H., *Stochastic processes and filtering theory*, Dover Books on Electrical Engineering, 2008.

[49] Kaczorek T., Dzieliński A., Dąbrowski W., Łopatka R., *Podstawy teorii sterowania*, WNT 2009.

[50] Karatzas I., *Lectures on the mathematics of finance*, American Mathematical Society, 1997.

[51] Karatzas I., Shreve S.E., *Connections between optimal stopping and singular control I. Monotone follower problems*, "SIAM Journal Control and Optimization", 1984 vol. 22, p. 856–877.

[52] Karatzas I., Shreve S.E., *Connections between optimal stopping and singular control II. Reflected follower problems*, "SIAM Journal Control and Optimization", 1985 vol. 22, p. 433–451.

[53] Karatzas I., Zamfirescu I.M., *Martingale approach to stochastic differential games of control and stopping*, "The Annals of Probability", 2008 vol. 36(4), p. 1495–1527.

[54] Korytkowski A., *Przejście do horyzontu nieskończonego w problemie liniowo-kwadratowym z opónieniem*, [W:] *Prace z automatyki dedykowane profesorowi H. Góreckiemu z okazji 70-lecia urodzin*, pod red. Mitkowski W., Wydawnictwo AGH, 1997, s.155-162.

[55] Kozin F., *Stability of the linear stochastic system*, [In:] *Stability of stochastic dynamical systems. Lecture Notes in Mathematics*, ed. by Curtain R.F., 1972 vol. 294, s. 186-229.

[56] Kozłowski E., *Optymalizacja portfela inwestycyjnego z uwzględnieniem kosztu informacji*, rozprawa doktorska, IBS PAN 2002.

[57] Kozłowski E., *The linear quadratic stochastic optimal control problem with random horizon at finite number of events intependent of state system*, "Systems Science", 2010 vol. 36(3), p. 5–11.

[58] Kozłowski E., *Identyfication of linear system in random time*, "International Journal of Computer and Information Technology", 2012 vol. 1(2), p. 103–108.

[59] Kozłowski E., *Discrete time stochastic control with random finite horizon*, [In:] *Process Control in Management*, ed. by Banek T., Toruń 2009, p. 53–66.

[60] Kozłowski E., *Stabilization of linear system in random horizon via control*, "Control and Cybernetics", 2013 vol. 42(2), p. 527–541.

[61] Kozłowski E., *Optimal stopping in controlled linear systems with quadratic criterion*, [In:] *Robotics and Manufacturing systems*, ed. by Koukolova L., Świć A., 2014, p. 73–84.

[62] Kozłowski E., *Optimal route and control for the LQC problem*, [In:] *Probability in action*, ed. by Banek T., Kozłowski E., 2014, p. 41–50.

[63] Kozłowski E., *Optimal route determining for LQ problem with optimally stopped horizon*, [In:] *Proceedings of 20th International Conference on Methods and Models in Automation and Robotics, MMAR 2015*, Międzyzdroje, 24-27 August 2015, p. 553–557.

[64] Kozłowski E., *Optimal stopping of controlled linear stochastic systems*, [In:] *Proceedings of 21th International Conference on Methods and Models in Automation and Robotics, MMAR 2016*, Międzyzdroje, 29 Aug.-1 Sept. 2016, p. 272–277.

[65] Kozłowski E., *Optimal stopping areas for discrete time linear quadratic control problem*, [In:] *Proceedings of 22th International Conference on Methods and Models in Automation and Robotics, MMAR 2017*, Międzyzdroje, 28-31 August 2017, p. 503–507.

[66] Kozłowski E., *Differences between optimal routes for linear quadratic problems with fixed and optimally stopped horizon*, [In:] *Probability in action*, ed. by Łagodowski Z., 2015, p. 33–45.

[67] Kozłowski E., Gola A., Świć A., *Model of production control in Just-in-Time delivery system conditions*, "Advances in Manufacturing Science and Technology", 2014 vol. 38(1), p. 77–88.

[68] Kozłowski E., *Analiza i identyfikacja szeregów czasowych*, Politechnika Lubelska, Lublin 2015.

[69] Kozłowski E., Gola A., Świć A., Hajduk M., Terkaj W., *A predective model of Multi-stage production planning for fixe time orders*, "Management and Production Engineering Review", 2014 vol. 5(3), p. 23–32.

[70] Krasnosielska-Kobos A., *Construction of Nash equilibrium based on multiple stopping problem in multi-person game*, Mathematical Methods of Operational Research", 2016 vol. 83, p. 53–70.

[71] Kulikowski R., *Procesy optymalne i adaptacyjne w układach regulacji automatycznej*, PWN, Warszawa 1965.

[72] Kumar P.R., *A survey of some results in stochastic adaptive control*, "SIAM Journal on Control and Optimization", 1985 vol. 23(3), p. 329–380.

[73] Kuratowski K., *Rachunek różniczkowy i całkowy. Funkcje jednej zmiennej*, PWN, Warszawa 1975.

[74] Leja F., *Rachunek różniczkowy i całkowy*, PWN, Warszawa 1975.

[75] Lifshits M.A., *Gaussian Random Functions*, Springer, 1995.

[76] Lin Q., Loxton R., Teo K.L., Wu Y.H., *Optimal control problem with stopping constraints*, "Journal of Global Optimization", 2015 vol. 63(4), p. 835–861.

[77] Lin Q., Loxton R., Teo K.L., Wu Y.H., *Optimal control computation for nonlinear systems with state dependent stopping criteria*, "Automatica", 2012 vol. 48, p. 2116–2129.

[78] Liptser R.Sh., Runggaldier W.J., Taksar M., *Deterministic approximation for stochastic control problems*, "SIAM Journal of Control and Optimization", 1996 vol. 34, s. 161-178.

[79] Liptser, R.SH., Shiryaev, A.N., *Statistics of Stochastic Processes*, Springer-Verlag, New York 1978.

[80] Mainali M.K., Shimada K., Mabu S., Hirasawa K., *Optimal route of road networks by dynamic programming*, [In:] *Proceedings of IEEE International Joint Conference on Neural Networks*, Hong Kong 2008, p. 3416–3420.

[81] Magnus J.R., Neudecker H., *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 1999.

[82] Manzie C., Krstic M., *Extremum seeking with stochastic perturbation*, "IEEE Transactions on Automatic Control", 2009 vol. 54(3), p. 580–585.

[83] Musielak J., *Wstęp do analizy funkcjonalnej*, PWN, Warszawa 1989.

[84] Muszyński J., *Teoria całki: miara i całka*, PWN, Warszawa 1990.

[85] Niederliński A., Mościński J., Ogonowski Z., *Regulacja adaptacyjna*, PWN, Warszawa 1995.

[86] Oksendal B., *Stochastic differential equations. An introduction with applications*, Springer, 2000.

[87] Oksendal B., Sulem A., *Applied Stochastic Control of Jump Diffusions*, Springer, 2007.

[88] Plucińska A., Pluciński E., *Rachunek prawdopodobieństwa. Statystyka matematyczna. Procesy stochastyczne*, WNT, Warszawa 2006.

[89] Porosiński Z., Szajowski K., Trybuła S., *Bayes control for a multidimensional stochastic system*, "System Sciences", 1985 vol. 11, p. 51–64.

[90] Prochrow J.W., Rozanow J.A., *Rachunek prawdopodobieństwa*, PWN, Warszawa 1972.

[91] Rishel R., *A comment on a dual control problem*, "IEEE Transactions on Automatic Control", 1981 vol. 26(2), p. 606–609.

[92] Rishel R., *An exact formula for a linear quadratic adaptive stochastic optimal control law*, "SIAM Journal on Control and Optimization", 1986 vol. 24(4), p. 667–674.

[93] Rishel R., *A strong separation principle for stochastic control systems driven by a hidden Markov model*, "SIAM Journal on Control and Optimization", 1994 vol. 32(4), p. 1008–1020.

[94] Rishel R., *Dynamic programming and minimum principles for systems with jump Markov disturbances*, "SIAM Journal on Control", 1975 vol. 13(2), p. 338–371.

[95] Ross Sh. M., *Introduction to probability models*, Academic Press, 6 Ed., 1997.

[96] Rudin W., *Analiza funkcjonalna*, PWN, Warszawa 2001.

[97] Runggaldier W.J., *Concepts and methods for discrete and continuous time control under uncertainty*, "Insurance Mathematics and Economics", 1998 vol. 22, p. 25–39.

[98] Saridis G. N., *Stochastic processes, estimation and control: the entropy approach*, John Wiley and Sons, 1995.

[99] Sarkka S., *Bayesian filtering and smoothing*, Cambridge University Press, 2013.

[100] Seierstad A., *Stochastic control in discrete and continuous time*, Springer 2009.

[101] Stengel R., *Optimal Control and Estimation*, Dover Books on Advanced Mathematics, 1994.

[102] Tamaki M., *Optimal stopping rule for the full-information duration problem with random horizon*, "Advances in Applied Probability", 2016 vol. 48, p.52–68.

[103] Tenno R., *Dual adaptive control properties*, "IFAC Proceedings Volumes", 2011 vol. 44(1), p. 4125–4130.

[104] Tenno R., *Dual adaptive controls for linear system with unknown constant parameters*, "International Journal of Control", 2010, vol. 83(11), p. 2232–2240.

[105] Turnau A., *Sterowanie docelowe układami nieliniowymi w czasie rzeczywistym - algorytmy inteligentne i optymalno czasowe*, AGH Uczelniane wydawnictwa naukowo dydaktyczne, Kraków 2002.

[106] Wiener N., *Cybernetics*, Wiley, New York 1948.

[107] Yuan B., Orlowska M., Sadig S., *On the optimal robot routing problem in wireless sensor networks*, "IEEE Transactions on knowledge and data engineering", 2007 vol. 19(9), p. 1252–1261.

[108] Zabczyk J., *Chance and decision. Stochastic control in discrete time*, Scuola Normale Superiore, Pisa 1996.

[109] Zabczyk J., *Zarys matematycznej teorii sterowania*, PWN, Warszawa 1991.

[110] Zeitouni O., *A class of adaptive control problems solved via stochastic control*, "Systems and Control Letters", 1989 vol. 12(1), p. 57–62.

[111] Ванько В.И., Ермошина О.В., Кувыркин Г.Х., *Вариационное исчисление и оптимальное управление*, МГТУ им. Баумана, Москва 2001.

[112] Вентцель Е.С., Овчаров Л.А., *Теория вероятностей и ее инженерные приложения*, Академия, Москва, 2003.

[113] Вентцель Е.С., Овчаров Л.А., *Теория случайных процессов и ее инженерные приложения*, Академия, Москва, 2003.

[114] Ким Д.П., *Теория автоматического управления. Многомерные, нелинейные, оптимальные и адаптативные системы*, Физматлит Москва 2004

[115] Маланин В.В., Полосков И.Е., *Методы и практика анализа случайных процессов в динамических системах*, НИЦ Регулярная и хаотическая динамика, Москва 2005.

[116] Маланин В.В., Полосков И.Е., *Случайные процессы в нелинейных динамических системах. Аналитичесие и численные методы исследования*, НИЦ Регулярная и хаотическая динамика, Москва 2001.

[117] Миллер Б.М., Панков А.Р., *Теория случайных процессов*, Физматлит, Москва 2002.

[118] Пугачов В.Ц., Синицын И.Н., *Теория стохастических систем*, Логос, Москва 2000.

[119] Понтрягин Л.С., *Принцип максимума в оптималном управлении*, Едиториал УРСС, Москва 2004.

[120] Ширяев А. Н., *Вероятность 1. Элементарная теория вероятностей. Математические основания. Пределные теоремы*, МЦНМО, Москва 2004.

[121] Ширяев А. Н., *Вероятность 2. Суммы и последовательности случайных величин - стационарные, мартингалы, марковские цепи*, МЦНМО, Москва 2004.

[122] Ширяев А. Н., *Статистический последователный анализ. Оптимальные правила остановки*, Наука, Москва 1976.

[123] Ширяев А. Н., *Основы стохастической финансовой математики. Факты. Модели*, Фазис, Москва 1998.

[124] Ширяев А. Н., *Основы стохастической финансовой математики. Теория*, Фазис, Москва 1998.