XV Krajowa Konferencja Automatyki

Tom I

KKA 2005

Redaktorzy:
Zdzisław Bubnicki
Roman Kulikowski
Janusz Kacprzyk

XV Krajowa Konferencja Automatyki
Tom I

Redaktorzy:
Zdzisław BUBNICKI
Roman KULIKOWSKI
Janusz KACPRZYK

ORGANIZATOR
Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk

WSPÓŁORGANIZATORZY
Politechnika Warszawska
Przemysłowy Instytut Automatyki i Pomiarów
Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

ORGANIZATOR

Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk

WSPÓŁORGANIZATORZY

Politechnika Warszawska
Przemysłowy Instytut Automatyki i Pomiarów
Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

KOMITET PROGRAMOWY

Przewodniczący	Zdzisław BUBNICKI
Zastępca Przewodniczącego	Roman KULIKOWSKI

CZŁONKOWIE

Stanisław BAŃKA
Mikołaj BUSŁOWICZ
Ryszard GESSING
Jakub GUTENBAUM
Stanisław KACZANOWSKI
Janusz KACPRZYK
Józef KORBICZ
Krzysztof KOZŁOWSKI
Krzysztof KUŹMIŃSKI
Krzysztof MALINOWSKI
Antoni NIEDERLIŃSKI
Tadeusz PUCHAŁKA
Stanisław SKOCZOWSKI
Jerzy ŚWIĄTEK
Ryszard TADEUSIEWICZ
Krzysztof TCHOŃ
Jan WĘGLARZ
Michał BIAŁKO
Władysław FINDEISEN
Henryk GÓRECKI
Jerzy JÓZEFCZYK
Tadeusz KACZOREK
Jerzy KLAMKA
Zbigniew KOWALSKI
Juliusz L. KULIKOWSKI
Kazimierz MALANOWSKI
Wojciech MITKOWSKI
Władysław PEŁCZEWSKI
Leszek RUTKOWSKI
Roman SŁOWIŃSKI
Andrzej ŚWIERNIAK
Piotr TATJEWSKI
Leszek TRYBUS
Andrzej P. WIERZBICKI

KOMITET ORGANIZACYJNY

Przewodniczący	Roman KULIKOWSKI
Zastępcy Przewodniczącego	Janusz KACPRZYK
	Stanisław KACZANOWSKI
	Tadeusz KACZOREK
	Krzysztof MALINOWSKI
Członkowie	Roman OSTROWSKI
	Tadeusz PUCHAŁKA
	Dariusz WAGNER
Sekretarze naukowi	Jan STUDZIŃSKI
	Jan W. OWSIŃSKI

ISBN 83-89475-00-6

Druk: ARGRAF, Warszawa

STEROWANIE OPTYMALNE I ADAPTACYJNE

ADAPTACYJNE STEROWANIE TERMINALNE Z LOSOWYM HORYZONTEM

Tadeusz BANEK*+, Edward KOZŁOWSKI****

*Polska Akademia Nauk, Instytut Badan Systemowych
ul. Newelska 6, 01-447 Warszawa, e-mail:tabank@ramzes.umcs.lublin.pl
**Politechnika Lubelska, Wydział Zarządzania i Podstaw Techniki
ul. Nadbystrzycka 38,20-618 Lublin, e-mail:e.kozlovski@pollub.pl

Streszczenie: Uogólniono klasyczne wyniki adaptacyjnego sterowania dla zadania z ustalonym horyzontem otrzymane przez Rishela [6, 7] na przypadek horyzontu losowego. Otrzymano warunki konieczne optymalności oraz algorytm do obliczania ekstremalnego sterowania oraz momentu stopu. Wskazano na potencjalne zastosowania w sztucznej inteligencji, w rozpoznawaniu obrazów, w problematyce data mining, oraz w wycenie informacji. Zastosowano otrzymane wyniki do zadania minimalizacji entropii warunkowej, co można interpretować jako zadanie optymalnego samo-uczenia.

Słowa kluczowe: sterowanie adaptacyjne, optymalne stopowanie, entropia, samo-uczenie

1. WSTĘP

Rozważamy problem sterowania systemem stochastycznym z czasem dyskretnym

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

gdzie u_i– oznacza działanie sterujące, w_i– zewnętrzne zakłócenia, ξ– reprezentuje nieznane parametry systemu. Działanie sterujące podejmowane w chwili i– tej może bazować tylko na obserwacjach poprzednich stanów systemu, tj. $y_1, ..., y_i$, oraz na znajomości rozkładów a'piori $P(dy_0)$, $P(d\xi)$. Jednakże sterując i obserwując stany systemu, można wzbogacać wiedzę o parametrach ξ. Rozkład a'posteriori w chwili i– tej, charakteryzujący stan wiedzy o ξ otrzymanej z obserwacji $y_1, ..., y_i$ zależy jednak od działań sterujących podjętych do chwili i, tzn. od $(u_0, ..., u_{i-1})$, ponieważ te wpływają na obserwowane stany. Aby wypełnić cel sterowania jakim jest zwykle optymalizacja jakiegoś wskaźnika jakości zależnego od stanów systemu i sterowań, optymalne sterowanie musi mieć dualną naturę – umożliwiać odpowiednio szybkie gromadzenie wiedzy o ξ, oraz realizować cel. Właściwe proporcje w realizacji tych dwóch odrębnych, a jednocześnie współzależnych zadań stanowią istotę problemu adaptacyjnego sterowania. Ze względu na ogrome, 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 N.Wiener [1], A.A.Feldbaum [2, 3], R.Bellman [4], R.Kulikowski [5], R.Rishel [6, 7, 8], V.E.Beneš, I.Karatzas, R.Rishel [9]. Literatura przedmioru jest ogromna, patrz spis literatury zamieszczony w cytowanych publikacjach. Aspekty praktyczne przedstawiane są setkach książek i artykułów. Niektóre wymienione są w [10].

W klasycznym sformułowaniu czas dyskretny i przebiega wartości $i = 0, ..., i = N$, gdzie N jest zadaną liczbą naturalną (patrz np. R.Rishel [6, 7]) zwaną horyzontem, a wskaźnik jakości jest postaci

$$E\left[h(y_N) + \sum_{i=0}^{N-1} g_i(y_i, u_i)\right]$$

gdzie g_i reprezentują koszty sterowania, h ocenia stan końcowy, E jest operatorem uśredniania względem miary probabilistycznej. Bywają jednak zadania w których przyjęcie ustalonego – i niezależnego od wyników sterowania horyzontu N, nie prowadzi do adekwatnego modelu sytuacji. W zadaniach typu self–learning, czy w problemach sztucznej inteligencji właśnie przyjmuje się horyzont działania zależny od otrzymanych wyników i stopuje się proces w pierwszym momencie uzyskania satysfakcjonujących rezultatów. Z takimi przypadkami spotykamy się, na przykład, przy trenowaniu sieci neuronowej, w zadaniach rozpoznawania obrazów, w data minning lub przy otrzymywaniu rozkładu a'posteriori parametru ξ o entropii (lub ilości informacji Fishera) na zadowalającym poziomie – w problemach samo–uczenia (patrz np. T.Banek, E.Kozłowski [14, 15]).

Jednak optymalne stopowanie procesów stochastycznych jest skomplikowanym zadaniem matematycznym stanowiącym dla badacza wyzwanie samo w sobie. Podejście proponowane w matematycznych monografiach (czas dyskretny) bazuje zwykle na konstrukcji tzw. koperty Snella, czyli najmniejszego semimartyngału dominującego zatrzymywany proces. Dowodzi się, że optymalny moment stopu jest pierwszym momentem zrównania się procesów – dominującego (zdominowanego) i stopowanego Jednak w rozważanym zadaniu stopujemy nie jeden proces, lecz rodzinę procesów (indeksowanych sterowaniem). Oznacza to konstrukcję nie jednej koperty Snella, lecz całej ich rodziny - z której dopiero należy wybrać taką, aby wyznaczony przez nią czas zatrzymania dał optymalną wartość wskaźnika jakości. Klasyczna, czysto probabilistyczna metoda nie jest więc praktyczna w rozważanym zadaniu. Dla zadań postaci

$$\sup_{\tau \le N} E\left[\sum_{i=0}^{\tau-1} g_i(y_i, u_i)\right]$$

gdzie τ jest zmienną losową przyjmująca wartości $1,...,N$, podaliśmy w pracy [16] metodę wykorzystującą ideę sterowania do *stopowania procesu przy pomocy sterowania*. W tej pracy uogólniamy naszą metodę tak aby rozważyć zadanie ogólniejsze

$$\sup_{\tau \le N} E\left[h\left(y_\tau\right) + \sum_{i=0}^{\tau-1} g_i(y_i, u_i)\right]. \qquad (1)$$

Metoda składa się z następujących etapów; rozszerzamy wektor sterowania u_i o dodatkową współrzędną θ_i, oraz modyfikujemy wskaźnik mnożąc wyrażenia pod znakiem wartości oczekiwanej, przez

$$\psi_i\left(\theta_1, ..., \theta_i\right) \triangleq \prod_{j=0}^{i} \theta_j.$$

Wprowadzamy oznaczenia: wektor stanów i sterowań

$$\begin{aligned} Y_j &= \left(y_1, ..., y_j\right) \text{ oraz } Y_0 = y_0 \\ U_j &= \left(u_1, ..., u_j\right) \text{ oraz } U_0 = u_0 \end{aligned}$$

a także funkcję

$$\phi_i\left(\theta_i, Y_i, U_i\right) = g_i(y_i, u_i)\theta_i + h_i\left(Y_i, U_{i-1}\right)\left(1 - \theta_i\right) \qquad (2)$$

przy pomocy której zapisujemy pierwotny funkcjonał jakości (1) w postaci

$$E\left[\sum_{i=0}^{N-1} \phi_i\left(\theta_i, Y_i, U_i\right)\psi_{i-1}\left(\theta_1, ..., \theta_{i-1}\right) + h_N\left(Y_N, U_{N-1}\right)\psi_{N-1}\left(\theta_1, ..., \theta_{N-1}\right)\right]$$

Okazuje się, że modyfikacja wskaźnika jakości polegająca na wymnożeniu każdej funkcji ϕ_i oraz h_i przez funkcję ψ_{i-1}, a następnie rozważenie dla tak zmodyfikowanego wskaźnika, klasycznego zadania z ustalonym horyzontem N, prowadzi do zadania ze sterowaniem adaptacyjnym i stopowaniem w momencie losowym (nie większym niż N). Dla takiego zadania podamy warunki konieczne optymalności, oraz algorytm prowadzący do wyznaczenia optymalnego sterowania i momentu stopu.

Struktura pracy jest następująca. W punkcie drugim formułujemy problem i podajemy warunki konieczne optymalności otrzymane z przyrównania słabych wariacji funkcjonału do zera. W punkcie trzecim wykorzystujemy ideę Rishela połączenia warunków koniecznych oraz indukcji wstecz do otrzymania algorytmu obliczeniowego. Interpretacja otrzymanych w algorytmie warunków pozwala wyznaczyć optymalny moment stopu. W ostatniej części pracy stosujemy otrzymane wcześniej wyniki do zadania optymalnego samo-uczenia. Jako miarę efektywności procesu samo-uczenia przyjmujemy entropię rozkładu warunkowego $P\left(d\xi \left| y_1, ..., y_\tau\right.\right)$.

Ze względów edytorskich prezentowane w tej pracy wyniki podane są bez dowodów. Pełna wersja pracy zostanie opublikowana wkrótce.

2. STEROWANIE ADAPTACYJNE

Niech $(\Omega, \mathcal{F}, P)$ będzie ustaloną przestrzenią probabilistyczną Na tej przestrzeni zdefiniowany jest ciąg niezależnych, m−wymiarowych wektorów losowych $w_1, w_2, ..., w_N$, o rozkładzie normalnym $N(0, I_m)$, oraz k−wymiarowy wektor ξ stochastycznie niezależny od $w_1, w_2, ..., w_N$. Definiujemy $\mathcal{F}_k \triangleq \sigma\left(\xi\right) \vee \sigma\{w_i : i = 1, 2, ..., k\}$ i kładziemy $\mathcal{F} = \mathcal{F}_N$. Będziemy rozważać problem sterowania adaptacyjnego systemem o równaniu stanu

$$y_{i+1} = f(\xi, y_i, u_i) + \sigma(\xi, y_i)w_{i+1} \qquad (3)$$

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 $\sigma : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{M}\left(n, m\right)$. $\mathbb{M}\left(n, m\right)$ jest zbiorem macierzy o n− wierszach i m− kolumnach. Na $(\Omega, \mathcal{F}, P)$ definiujemy rodzinę σ−podciał $\mathbf{Y}_j = \sigma\left\{y_i : i = 0, 1, ..., j\right\}$. Aby sformułować cel sterowania wprowadzimy funkcje $g_i : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$, oraz $h_i : \mathbb{R}^{n\times(i+1)} \times \mathbb{R}^{l\times i} \longrightarrow \mathbb{R}$, $i = 0, ..., N-1$. **Terminalne zadanie sterowania adaptacyjnego z losowym horyzontem** ma postać

Problem 1 *Szukamy*

$$\inf_{(u,\tau)\in\Gamma\times\Upsilon} E\left[\sum_{i=0}^{\tau-1} g_i(y_i, u_i) + h_\tau\left(Y_\tau, U_{\tau-1}\right)\right] \qquad (4)$$

gdzie Y_j mierzalne procesy $u_j \in R^l$ nazywamy działaniami sterującymi, $u = (u_0, u_1, ..., u_{\tau-1})$ nazywamy sterowaniem dopuszczalnym, Γ oznacza klasę sterowań dopuszczalnych. Zmienna losowa τ jest elementem z Υ, gdzie Υ oznacza klasę wszystkich momentów Markowa względem F_j, $j = 0, ..., N$, określonych na Ω o wartościach w $\{0, 1, ..., N\}$.

Wprowadzimy teraz ideę *stopowania procesu przy pomocy sterowania* i sformułujemy problem równoważny. Modyfikujemy pojęcie sterowania jak następuje. $\mathbf{Y}_j$ mierzalny proces $\mathbf{u} = (\theta, u)$, gdzie $\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$, $\theta_i \in \{0, 1\}$, $u = (u_0, u_1, ..., u_{N-1})$, nazywamy sterowaniem dopuszczalnym. Klasę sterowań dopuszczalnych oznaczamy przez $\mathbf{U}$. Stawiamy problem

Problem 2 *Szukamy*

$$\inf_{\mathbf{u}\in\mathbf{U}} J\left(\mathbf{u}\right) \qquad (5)$$

gdzie dla ϕ_i z (2)

$$J\left(\mathbf{u}\right) = \qquad (6)$$

$$E\left[\sum_{i=0}^{N-1} \phi_i\left(\theta_i, Y_i, U_i\right)\prod_{j=0}^{i-1}\theta_j + h_N\left(Y_N, U_{N-1}\right)\prod_{j=0}^{N-1}\theta_j\right],$$

oraz sterowania $u^* = (\theta^*, u^*)$ dla którego to supremum jest osiągnięte.

Widać, że zadanie postawione w (5) zamienia problem sterownaia adaptacyjnego z losowym horyzontem postawiony w (4), na równoważny problem sterowania adaptacynego z ustalonym horyzontem (ale dla zmodyfikowanego wskaźnika jakości). Słabe różniczkowanie, oraz własności warunkowej wartości oczekiwanej prowadzą do warunków koniecznych podanych w poniższym twierdzeniu.

Twierdzenie 3 *Jeżeli funkcje g_i i h_i są ciągłe i ograniczone oraz funkcje g_i, f mają ciągłe pochodne ze względu na zmienną u oraz $\det\Sigma(\xi,y) \neq 0$ dla $(\xi,y) \in R^k \times R^n$, gdzie $\Sigma(\xi,y) = \sigma(\xi,y)\sigma^T(\xi,y)$, to dla $j \in (0,1,...,N-1)$ warunek konieczny optymalności sterowania u^* w zadaniu (5) jest postaci:*

$$\nabla_{u_j} g_j(y_j,u_j^*)+E\Bigg\{\sum_{i=j+1}^{N-1}(1-\theta_i^*)\nabla_{u_j}h_i\left(Y_i,U_{i-1}^*\right)\Xi_{j+1}^{i-1} + \nabla_{u_j}h_N\left(Y_N,U_{N-1}^*\right)\Xi_{j+1}^{N-1}+\Bigg(\sum_{i=j+1}^{N-1}\phi_i\left(\theta_i^*,Y_i,U_i^*\right)\Xi_{j+1}^{i-1} + h_N\left(Y_N,U_{N-1}^*\right)\Xi_{j+1}^{N-1}\Bigg)\Gamma\left(\xi,y_j,y_{j+1},u_j^*\right)|\mathbb{Y}_j\Bigg\}=0 \tag{7}$$

$$\text{gdzie} \quad \Xi_j^k = \prod_{l=j}^{k}\theta_l^* \tag{8}$$

$$\Gamma(\xi,x,y,u) = (y-f(\xi,y,u))^T\Sigma^{-1}(\xi,y)\nabla_u f(\xi,y,u) \tag{9}$$

natomiast optymalny stop θ^* wyznaczamy z warunku

$$\theta_j^* = \begin{cases} 1, & \text{gdy } G(Y_j,U_j)+g_j(y_j,u_j^*) \le h_j\left(Y_j,U_{j-1}^*\right) \\ 0, & \text{w przeciwnym razie} \end{cases} \tag{10}$$

oraz

$$G(Y_j,U_j) = E\left(\sum_{i=j+1}^{N-1}\phi_i\left(\theta_i^*,Y_i,U_i^*\right)\prod_{l=j+1}^{i-1}\theta_l^* + h_N\left(Y_N,U_{N-1}^*\right)\prod_{l=j+1}^{N-1}\theta_l^*\,\middle|\,\mathbb{Y}_j\right) \tag{11}$$

3. WYZNACZENIE EKSTREMALNEGO STEROWANIA

Dla zadania (5) zgodnie z twierdzeniem 1 sterowanie $\mathbf{u} = (\theta,u)$, gdzie $u = \{u_0,u_1,..,u_{N-1}\}$ i $\theta = \{\theta_0,\theta_1,..,\theta_{N-1}\}$ jest optymalne jeżeli dla dowolnego $j \in \{0,1,,...,N-1\}$ spełniony jest układ równań (7)–(10). Poniżej przedstawimy wyznaczenie optymalnego sterowania wykorzystując warunek (7)–(10). Wprowadźmy najpierw niezbędne oznaczenia. Niech

$$V_j(\xi,Y_j,U_j) = E\left[\sum_{i=j}^{N-1}\phi_i(\theta_i,Y_i,U_i)\prod_{l=j}^{i-1}\theta_l + h_N(Y_N,U_{N-1})\prod_{l=j}^{N-1}\theta_j\,|\mathcal{F}_j\right]. \tag{12}$$

Korzystając z własności warunkowej wartości oczekiwanej mamy

$$V_j(\xi,Y_j,U_j) = \phi_j(\theta_j,Y_j,Y_j) + +\theta_j E\left[V_{j+1}(\xi,Y_{j+1},U_{j+1})\,|\mathcal{F}_j\right]. \tag{13}$$

Z postaci wskaźnika jakości wynika, że

$$V_N(\xi,Y_N,U_N) = h_N(Y_N,U_{N-1}).$$

3.1. Algorytm wyznaczenia sterowania

Za pomocą indukcji wstecz wyznaczymy sterowanie ekstremalne $\mathbf{u}^*$. Poniżej podajemy algorytm wyznaczenia $\mathbf{u}^*$.

1. Kładziemy

$$V_N(\xi,Y_0,U_N) = h_N(Y_N,U_{N-1}) \text{ oraz } j = N$$

2. Definiujemy

$$j = j-1$$

3. Kładziemy

$$\tilde{V}_{j+1}(\xi,Y_j,w_{j+1},U_{j+1}) = V_{j+1}(\xi,Y_j,f(\xi,y_j,u_j)+\sigma(\xi,y_j)w_{j+1},U_{j+1})$$

oraz

$$\tilde{\Phi}_{j+1}(\xi,Y_j,w_{j+1},U_{j+1}) = \nabla_{u_j}\tilde{V}_{j+1}(\xi,Y_j,w_{j+1},U_{j+1})$$

4. Obliczamy

$$Z(Y_j,U_j) \triangleq \int\left\{\tilde{\Phi}_{j+1}(\xi,Y_j,x,U_{j+1}^*)+\tilde{V}_{j+1}(\xi,Y_j,x,U_{j+1}^*)x^T\times \sigma^T(\xi,y_j)\Sigma^{-1}(\xi,y_j)f_u'(\xi,y_j,u_j)\right\}\gamma(x)\,dxP(d\xi\,|\mathbb{Y}_j)$$

gdzie

$$\gamma(x) = \frac{1}{\sqrt{(2\pi)^m}}\exp\left(-\frac{1}{2}x^Tx\right)$$

5. Szukamy optymalnego sterowania u_j^* takiego, dla którego spełniony jest układ równań (7)

$$\nabla_{u_j}g_j(y_j,u_j)+Z(Y_j,U_j) = 0$$

6. Obliczamy

$$W(Y_j,U_j) \triangleq E\left[\tilde{V}_{j+1}(\xi,Y_j,w_{j+1},U_{j+1})\,|\mathbb{Y}_j\right] = \int\tilde{V}_{j+1}(\xi,Y_j,x,U_{j+1})\gamma(x,I_m)\,dxP(d\xi\,|\mathbb{Y}_j)$$

7. Jeżeli

$$g_j(y_j,u_j^*) - h_j(Y_j,U_{j-1}) + W(Y_j,U_j) \le 0$$

to $\theta_j^* = 1$ w przeciwnym razie $\theta_j^* = 0$

8. Następnie korzystamy z równania (13) i wyznaczamy

$$V_j(\xi,Y_j,U_j) = \phi_j(\theta_j^*,y_j,u_j^*)+\theta_j^*\times \int V_{j+1}(\xi,Y_j,f(\xi,y_j,u_j^*)+\sigma(\xi,y_j)x,U_{j+1})\gamma(x,I_m)\,dx$$

9. Jeżeli $j = 0$ to zatrzymujemy, w przeciwnym razie wracamy do **kroku 2**

□

4. WARUNKOWA ENTROPIA SYSTEMU LINIOWEGO

Załóżmy że system jest opisywany równaniem

$$y_{j+1} = f_1(y_j, u_j) + f_2(y_j, u_j)\,\xi + \sigma(y_j)\,\varepsilon_{j+1}, \quad (14)$$

gdzie $j = 0, \ldots, N-1$, $y_i \in \mathbb{R}^n$, $f_1 : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$, $f_2 : \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{M}(n,k)$, gdzie $\mathbb{M}(n,k)$ jest zbiorem macierzy o wymiarach $n \times k$, oraz $\sigma : \mathbb{R}^n \longrightarrow \mathbb{M}(n,m)$. Zakładamy że funkcje f_1, f_2, σ są ciągłe i ograniczone. Również zakładamy że macierze $\Sigma(y) = \sigma(y)\,\sigma^T(y)$ są nieosobliwe. Wektor losowy ξ ma apioryczny rozkład $N(\mathrm{m}, \mathrm{Q})$, wtedy z teorii filtracji ciągów warunkowo normalnych (patrz [12] rozd. 13) mamy:

1. warunkowy rozkład

$$P(d\xi\,|\mathbb{Y}_j)$$

jest rozkładem normalnym $N(\mathrm{m}_j, \mathrm{Q}_j)$;

2. warunkowa wartość oczekiwana ξ,

$$\mathrm{m}_j = E(\xi\,|\mathbb{Y}_j)$$

oraz macierz warunkowej kowariancji

$$\mathrm{Q}_j = E\left([\xi - \mathrm{m}_j][\xi - \mathrm{m}_j]^T\,|\mathbb{Y}_j\right)$$

dane są wzorem

$$\mathrm{m}_j = \left[I + \mathrm{Q}\sum_{i=0}^{j-1} f_2^T(y_i,u_i)\,\Sigma^{-1}(y_i)\,f_2(y_i,u_i)\right]^{-1} \times \left[\mathrm{m} + \mathrm{Q}\sum_{i=0}^{j-1} f_2^T(y_i,u_i)\,\Sigma^{-1}(y_i)\,[y_{i+1} - f_1(y_i,u_i)]\right] \quad (15)$$

oraz

$$\mathrm{Q}_j = \left[I + \mathrm{Q}\sum_{i=0}^{j-1} f_2^T(y_i,u_i)\,\Sigma^{-1}(y_i)\,f_2(y_i,u_i)\right]^{-1}\mathrm{Q}. \quad (16)$$

Entropia rozkładu warunkowego przy sterowaniu $U_{j-1} = (u_0, \ldots, u_{j-1})$ wynosi

$$H_u(\xi\,|\mathbb{Y}_j) = -\int [\ln\gamma(\xi - \mathrm{m}_j, \mathrm{Q}_j)]\,\gamma(\xi - \mathrm{m}_j, \mathrm{Q}_j)\,d\xi = \frac{1}{2}\left\{\ln(2\pi)^k + \ln\det(\mathrm{Q}_j) + tr\left[\mathrm{Q}_j^{-\frac{1}{2}}\mathrm{Q}_j\mathrm{Q}_j^{-\frac{1}{2}}\right]\right\}.$$

Ostatecznie

$$H_u(\xi\,|\mathbb{Y}_j) = \frac{1}{2}\left\{\ln\left[(2\pi e)^k\right] + \ln[\det(\mathrm{Q}_j)]\right\}. \quad (17)$$

5. OPTYMALNE SAMO-UCZENIE Z UWZGLĘDNIENIEM KOSZTÓW

Zadanie samo-uczenia realizujem w N krokach. Zadanie polega na minimalizacji zarówno warunkowej entropii, jak i kosztów samo-uczenia. Na ten koszt składają się koszty związane ze sterowaniem obiektem (operacyjne), oraz tzw. koszty ryzyka (ze popełnimy w przyszłości błąd wynikający z naszej niewiedzy). Nasza wiedzę o parametrach możemy oceniać wielkością entropii warunkowej. Koszt ryzyka jest oczywiście funkcją rosnącą ze względu na entropię warunkową. Zatem problem samo-uczenia można przedstawić w postaci; szukamy

$$\inf_{\mathrm{u}\in\mathrm{U}} J(\mathrm{u}) \quad (18)$$

gdzie $J(\mathrm{u})$ jest postaci (6) natomiast

$$\phi(\theta_i, Y_i, U_i) = g(y_i,u_i)\theta_i + (1-\theta_i)\,h_i(Y_i, U_{i-1}). \quad (19)$$

Występują tutaj koszty operacyjne $g(y_i,u_i)$, oraz koszty ryzyka $h_i(Y_i, U_{i-1})$ postaci

$$h_i(Y_i, U_{i-1}) \triangleq q(y_i, r(Y_{i-1}, U_{i-1})), \quad (20)$$

gdzie q jest rosnąca względem drugiego argumentu

$$r(Y_{i-1}, U_{i-1}) \triangleq H_u(\xi\,|\mathbb{Y}_j) = \frac{1}{2}\left\{\ln\left[(2\pi e)^k\right] + \ln[\det(\mathrm{Q}_j)]\right\}. \quad (21)$$

Rozwiązanie zadania (18) jest podane w twierdzeniu.

Twierdzenie 4. *Jeżeli funkcje g i h są ciągłe i ograniczone oraz funkcje g, f_2 mają ciągłe pochodne ze względu na u oraz $tr\{f_2^T(y,u)\,\Sigma^{-1}(y)\,f_2(y,u)\} \le C < \infty$, dla dowolnych $(y,u) \in R^n \times R^l$, to warunek konieczny optymalnego sterowania u^* w modelu (18) jest postaci*

$$\nabla_{u_j} g(y_j, u_j^*) + E\left\{\sum_{i=j+1}^{N-1}(1-\theta_i^*)\,\Psi_j(Y_i, U_{i-1}^*)\,\Xi_{j+1}^{i-1} + \Psi_j(Y_N, U_{N-1}^*)\,\Xi_{j+1}^{N-1} + \left(\sum_{i=j+1}^{N-1}\phi(\theta_i^*, Y_i, U_i^*)\,\Xi_{j+1}^{i-1} + h_N(Y_N, U_{N-1}^*)\,\Xi_{j+1}^{N-1}\right)\Gamma(\xi, y_j, y_{j+1}, u_j^*)\,|\mathbb{Y}_j\right\} = 0 \quad (22)$$

dla $j \in (0, 1, \ldots, N-1)$, gdzie Ξ_j^k jest określone wzorem (8),

$$\Gamma(\xi, x, y, u) = (y - f_1(y,u) - f_2(y,u)\,\xi)^T \times \Sigma^{-1}(y)\,[\nabla_u f_1(y,u) - \nabla_u f_2(y,u)\,\xi]$$

oraz

$$\Psi_j(Y_i, U_{i-1}) = -2\,(\nabla_{u_j} f_2(y_j,u_j))^T\,\Sigma(y_j)^{-1} \times f_2(y_j,u_j)\,\frac{\partial}{\partial r}h_i(y_i, r(Y_{i-1}, U_{i-1}))\,tr\{A_i^{-1}(U_{i-1})\}$$

$$A_i(U_{i-1}) = I + \mathrm{Q}\sum_{k=0}^{i-1} f_2^T(y_k,u_k)\,\Sigma(y_k)^{-1}\,f_2(y_k,u_k)$$

natomiast stop θ_j^* jest postaci

$$\theta_j^* = \begin{cases} 1, & \text{gdy } G(Y_j, U_j) + g_j(y_j, u_j^*) \le h_j(Y_j, U_{j-1}^*) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (23)$$

gdzie $G(Y_j, U_j)$ jest określony wzorem (11).

6. KONKLUZJE

Podaliśmy metodę wyznaczania sterowń ekstremalnych dla problemu terminalnego sterowania adaptacyjnego z losowym horyzontem nie dłuższym niż N, gdzie N jest ustaloną, dowolną liczbą naturalną. Proponowana metoda jest oparta na idei sterowania i ma charakter analityczny – w tym sensie, że zamienia problem konstrukcji kopert Snella dla oryginalnego zadania, na problem różniczkowania funkcjonału bardziej skomplikowanego. Otrzymane warunki są jednak czytelne i umożliwiają podanie algorytmu do obliczania optymalnych sterowań. Ten z kolei, zawiera wyrażenia łatwe do interpretacji. W szczególności, możliwe jest wyznaczenie optymalnego momentu stopu. Badanie przypadku gdy $N \to \infty$, wymaga podania warunków dla których moment stopu jest skończony z $P-$prawdopodobieństwem jeden. Wówczas wykorzystanie zaproponowanej metody jest natychmiastowe. Rozpoczynamy obliczenia startując z dowolnego horyzontu N_0. Kontynułujemy dla $N_1 = 1+N_0,...,N_{k+1} = 1+N_k$. Ponieważ z założenia moment stopu jest skończony, więc istnieje skończone K takie, że wskaźnik jakości przestanie wzrastać dla $k \geq K$.

TERMINAL ADAPTIVE CONTROL WITH RANDOM HORIZON

Abstract: The classical adaptive control problem approach by Rishel [6, 7] for problems with fixed time horizon is extended to random cases. Necessary conditions are obtained and an algorithm for extremal control and stopping time is presented. Potential applications in artificial intelligence, in pattern recognition, data mining, self-learning and in information pricing are mentioned. Moreover, these results are applied to the specific problem of minimization of the conditional entropy, which is an example of the self – learning type of problems.

Literatura

[1] Wiener N. (1948) Cybernetics. New York , Wiley.

[2] Feldbaum A.A. (1960-1961) Dual control theory I -IV. Automation and Remote Control, 21, 874-1033 and 22, 1-109.

[3] Feldbaum A.A. (1965) Optimal Control Systems. Academic Press.

[4] Bellman R. (1961) Adaptive control processes. Princeton.

[5] Kulikowski R. (1965) Procesy optymalne i adaptacyjne w układach regulacji automatycznej. PWN, Warszawa.

[6] Rishel R. A nonlinear discrete time stochastic adaptive control problem (working paper).

[7] Harris L., Rishel R. (1986) An algorithm for a solutions of a stochastic adaptive linear quadratic optimal control problem. IEEE Transactions on Automatic Control, 31.

[8] Rishel R. (1986) An exact formula for a linear quadratic adaptive stochastic optimal control law. SIAM J. Control and Optimization, 24, No 4.

[9] Beneš V.E., Karatzas I., Rishel R. (1989) The separation principle for a Bayesian adaptive control problem with no strict-sense oprimal law. Imperial College Workshop on Applied Stochastic Analysis.

[10] Zabczyk J. (1996) Chance and decision. Scuola Normale Superiore, Pisa.

[11] Niederliński A., Mościński J., Ogonowski Z. (1995) Regulacja adaptacyjna. PWN, Warszawa.

[12] Liptser R. S., Shiryaev A. N. (1981) Statistics of Stochastic Processes (in Polish). PWN, Warszawa.

[13] Magnus J.R., Neudecker H. (1999) Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley, New York.

[14] Banek T., Kozłowski E. Adaptive control of systems entropy (submitted for publication)

[15] Banek T., Kozłowski E. (2004) Active and passive learning in control processes. Proc. 15th Int. Conf. System Science, II, 38-48.

[16] Banek T., Kozłowski E. (2005) Adaptive control with random horizon. Konf. Badania i Zastosowania Informatyki, Kazimierz Dolny.

Instytut Badań Systemowych
Polskiej Akademii Nauk

ISBN 83-89475-02-2