Annales UMCS Informatica AI 3 (2005) 5-14

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI

http://www.annales.umcs.lublin.pl/

Adaptive control with random horizon

Tadeusz Banek[a,b*], Edward Kozłowski[b]

[a]*System Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warszawa, Poland*
[b]*KMI, Lublin University of Technology, Nadbystrzycka 38, 20-618 Lublin, Poland*

Abstract

The classical adaptive control problem approach by Rishel [1,2] for the 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.

1. Introduction

We consider the optimal control problem for a discrete time stochastic system

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

w_i are the system disturbances, and ξ represents the unknown parameters of the system. Control actions u_i at time i can only base on observing the previous states of the system, i.e. $y_0,\dots, y_i$, and on the knowledge of the a priori distributions $P(dy_0)$ and $P(d\xi)$. However, controlling and observing the states of the system can increase information about the parameters ξ. The a posteriori distribution at time i, characterizing the knowledge about ξ obtained from the observations $y_0,\dots, y_i$, depends, however, on control actions undertaken before time i, i.e. on $(u_0,\dots,u_{i-1}$, because they influence the states being observed. To fulfill the purpose of control, which is usually to optimize performance criteria depending on the states of the system and the controls, an optimal control process must have a dual nature – it should yield both fast increase of information and optimization. Balancing these two distinct but interdependent tasks is the core of adaptive control. Due to its importance for applications, adaptive control problems have attracted attention for a long time. The first publications appeared half a century ago and are connected with the names of N. Wiener [3], A.A. Feldbaum [4,5], R. Bellman [6], R. Kulikowski [7], R. Rishel [1,2,8], V.E. Beneš, I. Karatzas and R.Rishel [9]. The relevant

[*]Corresponding author: *e-mail address*: t.banek@pollub.pl

literature is very extensive (see the references in the cited papers). The practical aspects are described in hundreds of books and articles; some of them are listed in [10].
In classical setting the discrete time $i = 0,...,N$, where N is fixed but arbitrary (see R.Rishel [1,2] for instance) often called horizon, and the quality index is of the form

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

where g_i represents costs of control, h estimates the final state, E is an operator of averaging with respect to a probability measure. However, there are problems in which assuming a fixed – and independent of control results horizon N, does not lead to an adequate model of situation. In problems of the self-learning type, or in problems of artificial intelligence a horizon of activity dependent on obtained results is just assumed and the process is stopped at the first moment of obtaining satisfactory results. We come across such cases, for example, when training a neural network, in image recognition problems, in data mining or when obtaining an a'posteriori distribution of parameter ξ with entropy (or amount of Fisher information) on satisfactory level – in self-learning problems (see eg. T.Banek, E.Kozłowski [11]).

However, the optimal stopping of stochastic processes is a complicated mathematical problem being itself a challenge for the scientist. The approach proposed in mathematical monographs (discrete time) is based usually on construction of so-called Snell envelope, i.e., the smallest submartingale which dominates the stopped process. It is proved that the optimal stopping moment is the first moment when the dominating and stopped processes become equal. However, in the considered problem we stop not one process but a family of processes (indexed by the control). It means construction of not one Snell envelope only, but an entire family of them – of which one should only choose such one so that the stopping time determined by it would result in the optimal value of the performance criteria. So, the classical, purely probabilistic method is not practical in the considered problem. In this paper we use a simple trick coming from control theory ideas: *stopping the process by using control*. We extend the control vector u_i with the additional coordinates θ_i and modify the performance criteria by using the properties of the hyperbolic tangent function; for large, positive values of control $\theta_0,..,\theta_i$, the function

$$\psi_i(\theta_0,...,\theta_i) \stackrel{\Delta}{=} \prod_{j=0}^{i} th(\theta_j),$$

is close to one. If $\theta_{i+1} \le 0$, then $\psi_i(\theta_0,...,\theta_k) = 0$, for all k greater than or equal to $i+1$. So, the modification of the performance criteria consisting in multiplication of each function g_i by function ψ_i, and considering, for such modified performance criteria, a classical problem with a fixed horizon N, leads to the problem with adaptive control and stopping at a random moment (not greater than N). For such a problem we will present necessary conditions of optimality, and an algorithm leading to determine the optimal control and the stopping moment.

The structure of the paper is the following. In section 2 we formulate the problem and present the necessary conditions of optimality obtained by making weak variations of the functional equal to zero. In section 3 we present a formula for evolution of the conditional distribution. In section 3 we use Ray Rishel's idea of joining the necessary conditions and backward induction to obtain a computational algorithm. Interpretation of the conditions obtained in the algorithm allows to determine an extremal stopping moment.

2. Adaptive control

Let (Ω, F, P) be a fixed probability space. On this space an n–dimensional random vector y_0, a sequence of m–dimensional random vectors $w_1,...,w_N$ with the normal distribution $N(0, I_m)$ and k–dimensional random vector ξ are defined. We assume stochastic independence of the object mentioned above: $y_0, w_1,..., w_N, \xi$. We define $F_k = \sigma\{w_i : i = 1,...,k\} \vee \sigma(\xi)$ and put $F = \sigma(y_0) \vee F_N$. We will consider a problem of adaptive control of the system with the equation of state

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

with the random initial condition y_0, where $i = 0,..., N$, $y_i \in R^n$, $f : R^k \times R^n \times R^l \to R^n$ and $\sigma : R^k \times R^n \to M(n,m)$, where $M(n,m)$ is a set of matrices with n–rows and m–columns. We assume the functions f and σ to be continuous with respect to all their arguments. On (Ω, F, P) we define a family of σ–subfields $Y_j = \sigma\{y_i : i = 0,..., j\}$. We name Y_j measurable processes $u_j \in R^l$, $\theta_j \in R$ control activities, and

$$\mathbf{u} = \begin{bmatrix} \theta \\ u \end{bmatrix}$$

$$\theta = (\theta_0, \theta_1, ..., \theta_{N-1})$$

$$u = (u_0, u_1, ..., u_{N-1})$$

a feasible control. A class of feasible controls is denoted by U. In order to formulate the aim of control we introduce functions $g_i, i = 0,..., N-1$. We assume that $g_i : R^n \times R^l \to R$ are continuous and bounded. The problem consists in determining

$$\sup_{\mathbf{u}\in U} J(\mathbf{u}), \tag{2}$$

where

$$J(\mathbf{u}) = E\left[\sum_{i=0}^{N-1} g_i(y_i, u_i)\prod_{l=0}^{i} th(\theta_l)\right], \tag{3}$$

and in determining such a control

$$\mathbf{u}^* = \left(\begin{bmatrix}\theta_0^*\\ u_0^*\end{bmatrix}, \begin{bmatrix}\theta_1^*\\ u_1^*\end{bmatrix}, \dots, \begin{bmatrix}\theta_{N-1}^*\\ u_{N-1}^*\end{bmatrix}\right) = \begin{bmatrix}\theta^*\\ u^*\end{bmatrix},$$

for which this supremum is attained. Weak differentiation and properties of the conditional expectation lead to necessary conditions presented in the theorem below.

Theorem 1 If functions g_i are continuous and bounded and functions g_i, f have continuous derivatives with respect to the variable u and $\det\left(\sigma(\xi,y)\sigma^T(\xi,y)\right) \neq 0$ for $(\xi,y)\in R^k \times R^n$, then the necessary condition of control optimality u^* and θ^* is given below:

$$\frac{1}{ch^2(\theta_j^*)}\left[g_j(y_j,u_j^*) + E\left\{\sum_{i=j+1}^{N-1} g_i(y_i,u_i^*)\prod_{l=j+1}^{i} th(\theta_l^*)\middle| Y_j\right\}\right] = 0 \tag{4}$$

$$\begin{aligned} th(\theta_j^*)\Bigg[\nabla_u g_j(y_j,u_j^*) + E\Bigg\{\Bigg(\sum_{i=j+1}^{N-1} g_i(y_i,u_i^*)\prod_{l=j+1}^{i} th(\theta_l^*)\Bigg) \\ \times\left(y_{j+1} - f(\xi,y_j,u_j^*)\right)^T\left(\sigma(\xi,y_j)\sigma^T(\xi,y_j)\right)^{-1} f_u'(\xi,y_j,u_j^*)\Bigg| Y_j\Bigg\}\Bigg] = 0 \end{aligned} \tag{5}$$

for $j\in\{0,1,\dots,N-1\}$, *where the symbol* $E\{\cdot|Y_j\}$ *denotes a conditional expected value (CEV).*

Proof:

Let us denote by

$$\gamma(x-m,Q) = \frac{1}{\sqrt{(2\pi)^n|Q|}}\exp\left(-\frac{1}{2}[x-m]^T Q^{-1}[x-m]\right),$$

densities of the normal distributions $N(m,Q)$. Let us notice that

$$\gamma\left(y_k - f(\xi,y_{k-1},u_{k-1}),\Sigma(\xi,y_{k-1})\right),$$

is the density of conditional probability of passing from state y_{k-1}, at the moment k-1, to state y_k, at the moment k, where F_{k-1} and u_{k-1} are known and $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$, which we denote as

$$p^{u_{k-1}}(k-1, y_{k-1}; k, y_k) = \gamma\left(y_k - f(\xi, y_{k-1}, u_{k-1}), \Sigma(\xi, y_{k-1})\right)$$

$$P\left(dy_k \middle| F_{k-1}\right) = p^{u_{k-1}}(k-1, y_{k-1}; k, y_k)\,dy_k$$

For $0 \le j \le i \le N$, we define conditional measures and joint measures

$$P\left(dy_j, ..., dy_i\right) = \prod_{k=j}^{i} P\left(dy_k \middle| F_{k-1}\right), \tag{6}$$

$$P\left(d\xi, dy_0, ..., dy_j\right) = P(d\xi)P(dy_0)P\left(dy_1, ..., dy_j\right). \tag{7}$$

Later, let us notice that functional (4) can be written in the form

$$J(\mathbf{u}) = J(\theta, u) = E\left[\sum_{i=0}^{j} g_i(y_i, u_i)\prod_{l=0}^{i} th(\theta_l) + E\left(\sum_{i=j+1}^{N-1} g_i(y_i, u_i)\prod_{l=0}^{i} th(\theta_l)\middle| F_j\right)\right]$$

$$= \int\left(\sum_{i=0}^{j-1} g_i(y_i, u_i)\prod_{l=0}^{i} th(\theta_l) + g_j(y_j, u_j)\prod_{l=0}^{j} th(\theta_l)\right)P\left(d\xi, dy_0, ..., dy_j\right)$$

$$+\int\left(\prod_{l=0}^{j} th(\theta_l)\int\left(\sum_{i=j+1}^{N-1} g_i(y_i, u_i)\prod_{l=j+1}^{i} th(\theta_l)\right)P\left(dy_{j+1}, ..., dy_N\right)\right)P\left(d\xi, dy_0, ..., dy_j\right). \tag{8}$$

We can see that the control u_{k-1} acts directly on passing the system from state y_{k-1} to state y_k, and, in the indirect way, to later states $y_{k+1}, ..., y_N$.

Let us fix the number $j \in \{0, ..., N-1\}$. Let $u = u^* + \varepsilon v$ where u^* – optimal control of the system, $\varepsilon > 0$, instead $v : R^{n\times(j+1)} \to R^{l\times N}$, $v = col\left(\tilde{0}, ..., \tilde{0}, \tilde{v}_j, \tilde{0}, ..., \tilde{0}\right)$, $\tilde{0} = col(0, ..., 0)$ where $\tilde{v}_j : R^{n\times(j+1)} \to R^l$, $\tilde{v}_j = col\left(v_j, ..., v_j\right)$ and $v_j = v_j\left(y_0, ..., y_j\right)$ is an arbitrary Borel function. Also, let $\theta = \theta^* + \varepsilon s$ where θ^* – optimal control of the stopping, $s : R^{n\times(j+1)} \to R^N$, $s = \left(0, ..., 0, s_j, 0, ..., 0\right)$, where $s_j = s_j\left(y_0, ..., y_j\right)$ is an arbitrary Borel function.

From formula (8) we compute

$$\frac{\partial}{\partial\varepsilon} J\left(\theta^* + \varepsilon s, u^*\right) = \int\prod_{l=0}^{j-1} th\left(\theta_l^*\right)$$

$$\left(\frac{1}{ch^2\left(\theta_j^*\right)}\times\left[g_j\left(y_j, u_j^*\right) + \int\left(\sum_{i=j+1}^{N-1} g_i\left(y_i, u_i^*\right)\prod_{l=j+1}^{i} th\left(\theta_l^*\right)\right)P\left(dy_{j+1}, ..., dy_N\right)\right]\right) \tag{9}$$

$$s_j P\left(d\xi, dy_0, ..., dy_j\right)$$

and

$$\frac{\partial}{\partial\varepsilon}J\left(\theta^*,u^*+\varepsilon v\right)=\int\prod_{l=0}^{j-1}th\left(\theta_l^*\right)\Big[\nabla_u g_j\left(y_j,u_j^*\right)th\left(\theta_j^*\right)$$
$$+th\left(\theta_j^*\right)\int\left(\sum_{i=j+1}^{N-1}g_i\left(y_i,u_i^*\right)\prod_{l=j+1}^{i}th\left(\theta_l^*\right)\right)\nabla_u P\left(dy_{j+1},...,dy_N\right)\Big]v_jP\left(d\xi,dy_0,...,dy_j\right) \tag{10}$$

Differentiation of the conditional measure results in the formula

$$\nabla_u P\left(dy_{j+1},...,dy_N\right)=\left(y_{j+1}-f\left(\xi,y_j,u_j\right)\right)$$
$$\Sigma^{-1}\left(\xi,y_j\right)\nabla_u f\left(\xi,y_j,u_j\right)P\left(dy_{j+1},...,dy_N\right) \tag{11}$$

which, after substituting to the expressions for weak variances (9)-(10) and making them equal to zero

$$\int\prod_{l=0}^{j-1}th\left(\theta_l^*\right)\Big[\nabla_u g_j\left(y_j,u_j^*\right)th\left(\theta_j^*\right)$$
$$+th\left(\theta_j^*\right)\int\left(\sum_{i=j+1}^{N-1}g_i\left(y_i,u_i^*\right)\prod_{l=j+1}^{i}th\left(\theta_l^*\right)\right)$$
$$\left(y_{j+1}-f\left(\xi,y_j,u_j^*\right)\right)\Sigma^{-1}\left(\xi,y_j\right)\nabla_u f\left(\xi,y_j,u_j^*\right)$$
$$\times\prod_{i=j}^{N-1}p^{u_i^*}\left(i,y_i;i+1,y_{i+1}\right)dy_{j+1}...dy_N\Big]v_jP\left(d\xi,dy_0,...,dy_j\right)=0 \tag{12}$$

and

$$\int\prod_{l=0}^{j-1}th\left(\theta_l^*\right)\left(\frac{1}{ch^2\left(\theta_j^*\right)}\left[g_j\left(y_j,u_j^*\right)+\int\sum_{i=j+1}^{N-1}g_i\left(y_i,u_i^*\right)\prod_{l=j+1}^{i}th\left(\theta_l^*\right)\right.\right.$$
$$\left.\left.\times P\left(dy_{j+1},...,dy_N\right)\right]\right)s_jP\left(d\xi,dy_0,...,dy_j\right)=0 \tag{13}$$

Because the conditions (12) and (13) must be satisfied for arbitrary $\sigma-$ field Y_j measurable Borel function v_j and s_j, so the standard argument from the measure theory implies that CEV occurring in the thesis of the theorem must disappear.

□

3. Determining the optimal control

In this section we will present Ray Rishel's method of using backward induction and iterative values of CEV to determine the optimal control based on necessary conditions (4), (5).

First, let us introduce necessary notations. Let

$$V_j\left(\xi,y_0,...,y_j\right)=E\left[\sum_{i=j}^{N-1}g_i\left(y_i,u_i\right)\prod_{l=j}^{i}th\left(\theta_l\right)\middle|F_j\right]. \tag{14}$$

By using properties of CEV we obtain

$$V_j\left(\xi, y_0,\dots,y_j\right) = \left[g_j\left(y_j,u_j\right) + E\left(V_{j+1}\left(\xi,y_0,\dots,y_{j+1}\right)\middle|F_j\right)\right]th\left(\theta_j\right) \qquad (15)$$

and

$$V_N\left(\xi, y_0,\dots,y_N\right) = 0 .$$

So, by formulas (14), (15) we have

$$E\left\{g'_u\left(y_j,u_j\right) + \left(\sum_{i=j+1}^{N-1} g_i\left(y_i,u_i\right)\prod_{l=j+1}^{i} th\left(\theta_l\right)\right)\left(y_{j+1} - f\left(\xi,y_j,u_j\right)\right)^T \Sigma^{-1}\left(\xi,y_j\right) f_u'\left(\xi,y_j,u_j\right)\middle|Y_j\right\} th\left(\theta_j\right)$$

$$= \left[E\left\{E\left\{V_{j+1}\left(\xi,y_0,\dots,y_j,f\left(\xi,y_j,u_j\right)+\sigma\left(\xi,y_j\right)w_{j+1}\right)\left(\sigma\left(\xi,y_j\right)w_{j+1}\right)^T \Sigma^{-1}\left(\xi,y_j\right) f_u'\left(\xi,y_j,u_j\right)\middle|F_j\right\}\middle|Y_j\right\}\right.$$

$$\left. + g'_u\left(y_j,u_j\right)\right] th\left(\theta_j\right) = th\left(\theta_j\right)\left[g'_u\left(y_j,u_j\right) + \int\dots\int V_{j+1}\left(\xi,y_0,\dots,y_j,f\left(\xi,y_j,u_j\right)+\sigma\left(\xi,y_j\right)x\right)\right.$$

$$\left.\times x^T\sigma^T\left(\xi,y_j\right)\Sigma^{-1}\left(\xi,y_j\right) f_u'\left(\xi,y_j,u_j\right)\gamma\left(x,I_m\right)dxP\left(d\xi\middle|Y_j\right)\right]$$

where $P\left(d\xi\middle|Y_j\right)$ denotes a conditional probability of the random variable ξ with respect to filtration Y_j. We will compute the conditional distribution $P\left(d\xi\middle|Y_j\right)$ in the next section.

3.1. Determining the conditional distribution of the random variable ξ

Let $p(\cdot)$ and $p_0(\cdot)$ – denote an a priori density of the random vector ξ and the state vector of y_0, respectively. Obviously

$$\mu_0\left(\xi, y_0\right) = p\left(\xi\right)p_0\left(y_0\right)$$

and, because densities μ_n of the joint distribution $\left(\xi, y_0,\dots,y_n\right)$ are expressed by μ_{n-1} with the formula

$$\mu_n\left(\xi,y_0,\dots,y_n\right) = \mu_{n-1}\left(\xi,y_0,\dots,y_{n-1}\right)\gamma\left(y_n - f\left(\xi,y_{n-1},u_{n-1}\right),\Sigma\left(\xi,y_{n-1}\right)\right),$$

so

$$\mu_n\left(\xi,y_0,\dots,y_n\right) = p\left(\xi\right)p_0\left(y_0\right)\prod_{i=1}^{n}\gamma\left(y_i - f\left(\xi,y_{i-1},u_{i-1}\right),\Sigma\left(\xi,y_{i-1}\right)\right).$$

By the Bayes formula we have a conditional distribution of the random variable ξ with respect to σ-field Y_n

$$P\left(d\xi\middle|Y_n\right) = \mu_n\left(\xi\middle|y_0,\dots,y_n\right)d\xi = \frac{\mu_n\left(\xi,y_0,\dots,y_n\right)}{\int\mu_n\left(x,y_0,\dots,y_n\right)dx}d\xi ,$$

and finally

$$P\left(d\xi \middle| Y_n\right) = \frac{p(\xi)\prod_{i=1}^{n}\gamma\left(y_i - f\left(\xi, y_{i-1}, u_{i-1}\right), \Sigma\left(\xi, y_{i-1}\right)\right)}{\int p(x)\prod_{i=1}^{n}\gamma\left(y_i - f\left(x, y_{i-1}, u_{i-1}\right), \Sigma\left(x, y_{i-1}\right)\right)dx} d\xi . \tag{16}$$

3.2. The algorithm of determining the optimal control

We will apply the results of previous sections to the construction of the algorithm using the backward induction

1. We put

$$V_N\left(\xi, y_0, \ldots, y_N\right) = 0 \text{ and } j = N .$$

2. We define

$$j = j - 1 .$$

3. We put

$$\tilde{V}_{j+1}\left(\xi, y_0, \ldots, y_j, u_j, w_{j+1}\right) = V_{j+1}\left(\xi, y_0, \ldots, y_j, f\left(\xi, y_j, u_j\right) + \sigma\left(\xi, y_j\right) w_{j+1}\right).$$

4. We set

$$W\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right) = E\left[\tilde{V}_{j+1}\left(\xi, y_0, \ldots, y_j, u_j, w_{j+1}\right) \middle| Y_j\right]$$
$$= \int \ldots \int \tilde{V}_{j+1}\left(\xi, y_0, \ldots, y_j, u_j, x\right)\gamma\left(x, I_m\right) dx P\left(d\xi \middle| Y_j\right).$$

5. We compute

$$Z\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right) = E\left[\tilde{V}_{j+1}\left(\xi, y_0, \ldots, y_j, u_j, w_{j+1}\right) {w_{j+1}}^T \sigma^T\left(\xi, y_j\right)\Sigma^{-1}\left(\xi, y_j\right) f_u'\left(\xi, y_j, u_j\right) \middle| Y_j\right]$$
$$= \int \ldots \int \tilde{V}_{j+1}\left(\xi, y_0, \ldots, y_j, u_j, x\right) x^T \sigma^T\left(\xi, y_j\right)\Sigma^{-1}\left(\xi, y_j\right) f_u'\left(\xi, y_j, u_j\right)\gamma\left(x, I_m\right) dx P\left(d\xi \middle| Y_j\right)$$

6. We search for an optimal control $\begin{bmatrix} \theta_j^* \\ u_j^* \end{bmatrix}$ for which the system of equations (4), (5)

$$\begin{cases} \dfrac{1}{ch^2\left(\theta_j\right)}\left[g\left(y_j, u_j\right) + W\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right)\right] = 0 \\ th\left(\theta_j\right)\left[g_u'\left(y_j, u_j\right) + Z\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right)\right] = 0 \end{cases}$$

is satisfied.

7. Next, we use equation (15) and determine

$$V_j\left(\xi, y_0, \ldots, y_j\right) = th\left(\theta_j\right)\Big[g_j\left(y_j, u_j\right)$$
$$+ \int V_{j+1}\left(\xi, y_0, \ldots, y_j, f\left(\xi, y_j, u_j\right) + \sigma\left(\xi, y_j\right) x\right)\gamma\left(x, I_m\right) dx\Big].$$

8. If $j = 0$ then we stop computations, otherwise we return to step 2.

□

Remark 2. We consider the following three cases of solutions $\left(\theta^*, u^*\right)$ of the system of equations

$$\begin{cases} \dfrac{1}{ch^2(\theta)}\kappa(u) = 0 \\ th(\theta)\chi(u) = 0 \end{cases}$$

1. $\chi\left(u^*\right) = 0$, $\kappa\left(u^*\right) \neq 0$ *then* $\theta^* = \infty$,

2. $\chi\left(u^*\right) = \kappa\left(u^*\right) = 0$ *then* θ^* *can be arbitrary,*

3. $\chi\left(u^*\right) \neq 0$, $\kappa\left(u^*\right) = 0$ *then* $\theta^* = 0$.

Remark 3. If u_j^* satisfies equation

$$\chi\left(u_j\right) \overset{\Delta}{=} g_u'\left(y_j, u_j\right) + Z\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right) = 0. \tag{17}$$

And if $\theta_{j+1}^* = \theta_{j+2}^* = \ldots = \theta_{N-1}^* = \infty$ holds, then we also set $\theta_j^* = \infty$.

Remark 4. If u_j^* determined from condition (17) also satisfies equation

$$\kappa\left(u_j\right) \overset{\Delta}{=} g_j\left(y_j, u_j\right) + E\left[\sum_{i=j+1}^{N-1} g_i\left(y_i, u_i^*\right) \prod_{l=j+1}^{i} th\left(\theta_l^*\right) \middle| Y_j\right] = 0, \tag{18}$$

then expected increase of the functional $J(\cdot)$ *in later moments is equal to zero because*

$$\begin{aligned} J(\mathbf{u}) &= E\left[\sum_{i=0}^{j-1} g_i\left(y_i, u_i^*\right) \prod_{l=0}^{i} th\left(\theta_l^*\right) + \prod_{l=0}^{j} th\left(\theta_l^*\right)\left(g_j\left(y_j, u_j^*\right) + E\left(\sum_{i=j+1}^{N-1} g_i\left(y_i, u_i^*\right) \prod_{l=j+1}^{i} th\left(\theta_l^*\right) \middle| Y_j\right)\right)\right] \\ &= E\left[\sum_{i=0}^{j-1} g_i\left(y_i, u_i^*\right) \prod_{l=0}^{i} th\left(\theta_l^*\right)\right] \end{aligned}$$

So we set $\theta_j^* = 0$ *which means that we stop the process.*

Following the above remarks we modify the step (6) of the algorithm in the following way

6. We determine u_j^* from the condition

$$g_u'\left(y_j, u_j\right) + Z\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right) = 0.$$

If

$$g_u\left(y_j, u_j\right) + W\left(y_0, \ldots, y_j, u_0, \ldots, u_j\right) = 0,$$

then we put $\theta_j^* = 0$, $\theta_j^* = \infty$ otherwise.

4. Conclusions

We presented a solution of the adaptive control problem with random horizon not longer than N, where N is an arbitrary integer. The proposed method uses the idea of control and has the analytic character – in this sense that it transforms the problem of construction of the Snell envelope for the original problem to the problem of differentiation of a functional more complicated. The conditions obtained in this way are readable however and enable to present an algorithm to compute consecutive controls. The algorithm contains expressions easy for interpretation. In particular it is possible to determine the optimal stopping moment. Investigating the case when $N \to \infty$ requires to present conditions for which the stopping moment is finite with P-probability equal to 1. Then usage of the proposed method is immediate. We begin computations starting from an arbitrary horizon N_0. We continue for $N_1 = 1+N_0,\ldots,N_{k+1} = 1+N_k$. Because, by assumption, the stopping moment is finite, so there exists a finite K such that the quality index stops growing for $k \geq K$.

References

[1] Rishel R., *A nonlinear discrete time stochastic adaptive control problem* (working paper).
[2] Harris L., Rishel R., *An algorithm for a solution of a stochastic adaptive linear quadratic optimal control problem*, IEEE Transactions on Automatic Control, 31(N) (1986) 12.
[3] Wiener N., *Cybernetics*, New York, Wiley, (1948).
[4] Feldbaum A.A., *Dual control theory I-IV*, Automation and Remote Control, 21 (1960) 874 and 22 (1961) 1.
[5] Feldbaum A.A., *Optimal Control Systems*, Academic Press, (1965).
[6] Bellman R., *Adaptive control processes*, Princeton, (1961).
[7] Kulikowski R., *Procesy optymalne i adaptacyjne w układach regulacji automatycznej*, PWN, (1965).
[8] Rishel R., *An exact formula for a linear quadratic adaptive stochastic optimal control law*, SIAM J. Control and Optimization 24(4) July (1986).
[9] Beneš V.E., Karatzas I., Rishel R., *The separation principle for a Bayesian adaptive control problem with no strict-sense optimal law*, Imperial College Workshop on Applied Stochastic Analysis, April (1989).
[10] Zabczyk J., *Chance and decision*, Scuola Normale Superiore, Pisa, (1996).
[11] Banek T., Kozłowski E., *Active and passive learning in control processes*, Proc. 15th Int. Conf. System Science, II (2004) 38.