Control and Cybernetics

vol. **35** (2006) No. 2

Adaptive control of system entropy

by

Tadeusz Banek[1,2] **and Edward Kozłowski**[2]

[1] Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa

[2] Chair of Mathematics and Informatics, Lublin University of Technology
ul. Nadbystrzycka 38, 20-618 Lublin

Abstract: Controlling entropy of a system with unknown parameters is treated here as an adaptive control problem. Necessary conditions for optimality and an algorithm for computing extremal controls in the spirit of R. Rishel are obtained.

Keywords: adaptive control, active learning, entropy.

1. Introduction

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

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

where u_i are controls, w_i are the system disturbances, and ξ represents the unknown parameters of the system. The control actions u_i at time i can only base on observing the previous states of the system, i.e. $y_1, \ldots, y_i$, 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 $y_1, \ldots, y_i$, depends, however, on control actions undertaken before time i, i.e. on $(u_0, \ldots, u_{i-1})$, 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. However, one should remember that optimization of the performance criteria is a primary task, i.e., learning the unknown parameters, even though necessary, is a secondary task, and it is always dominated by the fundamental goal. In this paper we harmonize these two goals by introducing a unique one, called self-learning.

This is done by considering the control problems with the so-called joint system entropy $H(\xi, y_0, \ldots, y_N)$, entering explicitly in the performance criteria. The resulting trajectories say a lot about ξ, and at the same time, their entropy $H(y_0, \ldots, y_N)$ is of moderate size.

Application of the entropy concept in stochastic control is not new. The most known are the studies of Saridis and his followers (see Saridis, 1988, 1995, and the included literature). They consider systems with known dynamics but control affected by noise. This introduces uncertainty in system behavior. Notion of entropy is used to state the game optimization problem which allows to identify the "worst" noise and optimal control in the minmax sense. Generally, problems of this kind are not the self-learning problems considered in this paper. Stochastic systems with unknown parameters, which we are forced to learn can be evaluated by many and very different criteria. For instance, in financial stochastic models appearing in Banek and Kulikowski (2003) the Fisher measure of information was used. The choice of entropy, which was done in this paper, has some advantages. For technical systems such formulation of the problem is very natural and follows from the security requirements for instance, i.e., predictability of their behavior is just as important as the learning process itself. We show that this problem and its generalization can be treated as an optimal adaptive control problem, and solved by using Rishel's approach (see Harris, Rishel, 1986; Rishel, 1985, 1986). This approach incudes the following steps: first - Gatoux's differentiations combined with conditional expectation properties lead to necessary conditions for optimality, second - application of backward inductions to the necessary conditions leads to the Rishel's algorithm. In Section 2 we extend Rishel approach in two aspects: in the state equation $\sigma(\xi, y_i)$ is not necessarily equal to the identity matrix and the loss functional $\sum g(\xi, y_i, y_{i+1}, u_i)$ is allowed to depend additionally on ξ, y_{i+1}. This is necessary for our purpose. Indeed, in Section 3 it turns out that a required expression for the joint entropy includes these variables. In Section 4 we pose the problem and obtain necessary conditions of optimality by using the results of Section 2. It turns out that the resulting expression for the joint entropy is not necessarily a quadratic function of the trajectory, unless the system is linear. Thus, minimum error energy formulation (often combined with entropy concept) is generally not possible. At this point the reader is referred to papers by Saridis (1988) and others that follow his work. In Section 5 an algorithm for computing extremal controls is presented. We use here an idea of Rishel, which consists in application of backward induction to the necessary conditions. This is done in several steps, the most important being the introduction of a value function (which is not the Bellman function !) and using it in the manner similar as in dynamic programming. At the present time we test the algorithm on simple examples. In Section 6 we show such an example. More details will be presented in the next paper.

The Gaussian noise assumption looks very restrictive. In fact, the necessary conditions for optimality in the spirit of Rishel's can be obtained for any non-Gaussian noises. However, the resulting algorithms will depend on the entropy expressions for the noises and generally they can be obtained if these expressions are explicit. For instance, noises with distribution functions; binomial $prob\,(\xi = a) = 1 - prob\,(\xi = b) = \theta$, for any a, b in appropriate space, and $\theta \in [0, 1]$ are the cases. The reader is also referred to the interesting paper by Porosiński, Szajowski, Trybuła (1985) where non-Gaussian noise and random horizon are taken into account in stochastic control with unknown noise parameters.

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 Wiener (1948), Feldbaum (1960, 1961, 1965), Bellman (1961), Kulikowski (1965), Aoki (1967), Rishel (1986), Beneš and Karatzas (1991). The relevant literature is very extensive (see e.g. Liptser, Runggaldier and Taksar, 1996; Zabczyk, 1996; Dai Pra, Rudari and Runggaldier 1997; Saridis, 1995). The practical aspects are described in hundreds of books and articles; some of them are listed by Runggaldier and Zaccaria (2000).

2. Adaptive control

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Let $w_1, \ldots, w_N$ be a sequence of independent m-dimensional random vectors on Ω with normal distribution $N(0, I_m)$, let ξ be k-dimensional random vector with a priori distribution $P\,(d\xi)$, and let y_0 be an initial state with distribution $P\,(dy_0)$. All these objects are assumed to be stochastically independent. Define $\mathcal{F}_k \stackrel{\triangle}{=} \sigma\,(y_0) \vee \sigma\,(\xi) \vee \sigma\,\{w_i : i = 1, 2, \ldots, k\}$ and set $\mathcal{F} = \mathcal{F}_N$.

We will consider the adaptive control problem for the system with state equation

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

where $i = 0, \ldots, N-1$, $y_i \in \mathbb{R}^n$, $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n$ and $\sigma : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{M}\,(n, m)$, where $\mathbb{M}\,(n, m)$ is the set of $n \times m$ matrices. The functionsf, σ are assumed to be continuous in all their variables.

On $(\Omega, \mathcal{F}, \mathbb{P})$ we define a family of σ-subfields $\mathbb{Y}_j = \sigma\,\{y_i : i = 0, 1, \ldots, j\}$. A vector $u_j \in \mathbb{R}^l$ measurable with respect to $\mathbb{Y}_j$ is called a control action, and $u = (u_0,\, u_1, \ldots, u_{N-1})$ an admissible control. The class of admissible controls is denoted by U.

To specify the aim of control, we introduce loss functions g_i, $i = 0, 1, \ldots, N–1$. We assume that $g_i : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}$ are continuous and bounded. The task is to find

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

where

$$J(u) = E\left[\sum_{i=0}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i)\right]. \tag{3}$$

THEOREM 2.1 *Suppose that the functions g_j, $j = 0, 1, \ldots, N-1$, are continuous and bounded, f and g_j, $j = 0, 1, \ldots, N-1$, are continuously differentiable in u, and $\det \Sigma(\xi, y) \neq 0$ for $(\xi, y) \in \mathbb{R}^k \times \mathbb{R}^n$, where $\Sigma(\xi, y) = \sigma(\xi, y)\sigma^T(\xi, y)$. If u^* is an optimal control, then*

$$E\Bigg\{\nabla_u g_j(\xi, y_j, y_{j+1}, u_j^*) + \tag{4}$$

$$\left(\sum_{i=j}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i^*)\right)\left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_i)\nabla_u f(\xi, y_j, u_j^*)\Bigg|\mathbb{Y}_j\Bigg\} = 0$$

for all $j \in \{0, 1, \ldots, N-1\}$.

Proof. From the properties of conditional expectation it follows that for every $j \in \{0, 1, \ldots, N-1\}$ the functional (3) can be represented as

$$J(u) = E\Bigg[\sum_{i=0}^{j-1} g_i(\xi, y_i, y_{i+1}, u_i) + E\Bigg(\sum_{i=j}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i)\Bigg|\mathcal{F}_j\Bigg)\Bigg] \tag{5}$$

$$= \int \left(\sum_{i=0}^{j-1} g_i(\xi, y_i, y_{i+1}, u_i)\right) P(d\xi, dy_0, \ldots, dy_j)$$

$$+ \int \left(\int \left[\sum_{i=j}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i)\right] P_{j+1,N}(dy_{j+1}, \ldots, dy_N)\right) P(d\xi, dy_0, \ldots, dy_j),$$

where

$$P_{ji}(dy_j, \ldots, dy_i) = \prod_{k=j}^{i} \mathbb{P}(dy_k | \mathcal{F}_{k-1}), \tag{6}$$

$$P(d\xi, dy_0, \ldots, dy_j) = P(d\xi)\, P(dy_0)\, P_{1j}(dy_1, \ldots, dy_j), \tag{7}$$

for $0 \leq j < i \leq N$. Note that $\mathbb{P}(dy_k | \mathcal{F}_{k-1})$ is the transition probability for the process $\{y_i; 0 \leq i \leq N\}$ defined by (1); we write it in the form

$$\mathbb{P}(dy_k | \mathcal{F}_{k-1}) = p^{u_{k-1}}(k-1, y_{k-1}; k, y_k)\, dy_k. \tag{8}$$

Here $p^u(s, x; t, y)$ is the probability of transition from state x at time s to state y at time t under control u. Note that

$$p^{u_{k-1}}(k-1, y_{k-1}; k, y_k) = \gamma\left(y_k - f(\xi, y_{k-1}, u_{k-1}), \Sigma(\xi, y_{k-1})\right), \tag{9}$$

where

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

is the density of the normal distribution. We see that the control u_{k-1} affects directly the transition from state y_{k-1} to state y_k, and indirectly the transition to the later states $y_{k+1},\ldots,y_N$.

Fix $j\in\{0,\ldots,N-1\}$. Let $u=u^*+\epsilon v$, where u^* is an optimal control and ϵ a scalar, and let $v:\mathbb{R}^{n\times(j+1)}\longrightarrow\mathbb{R}^{l\times N}$, $v=\left(\tilde{0},...,\tilde{0},\tilde{v}_j,\tilde{0},...,\tilde{0}\right)$, $\tilde{0}=col\,(0,...,0)$ where $\tilde{v}_j:\mathbb{R}^{n\times(j+1)}\longrightarrow\mathbb{R}^l$, $\tilde{v}_j=col\,(v_j,...,v_j)$, and $v_j=v_j\,(y_0,...,y_j)$ is any Borel function. From (5) we compute

$$\frac{\partial}{\partial\epsilon}J(u^*+\epsilon v)=\int\Bigg[\int\nabla_u g_j(\xi,y_j,y_{j+1},u_j^*)P_{j+1,N}(dy_{j+1},\ldots,dy_N)+\int\left(\sum_{i=j}^{N-1}g_i(\xi,y_i,y_{i+1},u_i^*)\right)\nabla_u P_{j+1,N}(dy_{j+1},\ldots,dy_N)\Bigg]v_jP(d\xi,dy_0,\ldots,dy_j). \tag{10}$$

From (6), (8), (9) we have

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

Substituting (11) to (10) and equating to zero we obtain

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

which proves the assertion, because condition (12) has to be satisfied by any $\mathbb{Y}_j$-measurable Borel function. ■

3. Entropy

Consider the following situation. We wish to control an object and simultaneously learn its properties as precisely as possible. We are allowed to make N

tests. The problem of active learning is to find a control $u^* = \left(u_0^*, \ldots, u_{N-1}^*\right)$ with the smallest possible entropy $H(\xi, y_0, \ldots, y_N)$, that is, one that minimizes the uncertainty concerning the object being controlled. Accordingly, the task is to minimize the joint entropy, i.e., to find

$$\inf_{u \in U} H(\xi, y_0, \ldots, y_N). \tag{13}$$

Let $p(\cdot)$ and $p_0(\cdot)$ be the a priori distributions of the random vector ξ and the state vector y_0 respectively, and suppose that the density of the joint distribution of (ξ, y_0) is

$$\mu_0(\xi, y_0) = p(\xi) p_0(y_0).$$

By induction it is easy to obtain (see, e.g., Banek, Kozłowski, 2004) the following recurrence formula for the density of the joint distribution of $\mu_i(\xi, y_0, y_1, \ldots, y_i)$:

$$\mu_i(\xi, y_0, y_1, \ldots, y_i) = \\ = \mu_{i-1}(\xi, y_0, y_1, \ldots, y_{i-1}) \gamma\left(y_j - f(\xi, y_{j-1}, u_{j-1}), \Sigma(\xi, y_{j-1})\right),$$

where

$$\Sigma(\xi, y) = \sigma(\xi, y) \sigma^T(\xi, y)$$

and

$$\mu_N(\xi, y_0, y_1, \ldots, y_N) = p(\xi) p_0(y_0) \prod_{j=0}^{N-1} \gamma\left(y_{j+1} - f(\xi, y_j, u_j), \Sigma(\xi, y_j)\right). \tag{14}$$

Hence the entropy of the entire system is

$$\begin{aligned} H(\xi, y_0, \ldots, y_N) &= E\left[-\ln \mu_N(\xi, y_0, y_1, \ldots, y_N)\right] \qquad (15) \\ &= E\Bigg[\frac{1}{2} \sum_{j=0}^{N-1} \Big(\left[y_{j+1} - f(\xi, y_j, u_j)\right]^T \Sigma^{-1}(\xi, y_j) \left[y_{j+1} - f(\xi, y_j, u_j)\right] \\ &\quad + n \ln 2\pi + \ln\left|\det \Sigma(\xi, y_j)\right| \Big) - \ln p(\xi) - \ln p_0(y_0) \Bigg]. \end{aligned}$$

We introduce the following notation:

$$h(\xi, x, y, u) = \left[y - f(\xi, x, u)\right]^T \Sigma^{-1}(\xi, x) \left[y - f(\xi, x, u)\right] + \ln\left|\det \Sigma(\xi, x)\right|. \tag{16}$$

Hence (15) can be rewritten in the form

$$H(\xi, y_0, \ldots, y_N) = \frac{1}{2} E\left[\sum_{j=0}^{N-1} h(\xi, y_i, y_{i+1}, u_i)\right] + H(\xi) + H(y_0) + \frac{nN}{2} \ln 2\pi. \tag{17}$$

4. Minimal cost control

Problem (13) concerns optimal control of active learning, but with no costs taken into account. However, every learning process involves some costs. To find some characteristics of an object, or identify it, we have to find a control minimizing not only the uncertainty after N tests, but also the costs incurred. Therefore the problem may be stated as

$$\inf_{u\in U} E\left[\sum_{i=0}^{N-1} g_i\left(\xi, y_i, y_{i+1}, u_i\right)\right], \tag{18}$$

where

$$g_i\left(\xi, y_i, y_{i+1}, u_i\right) = g\left(h\left(\xi, y_i, y_{i+1}, u_i\right), y_i, u_i\right)$$

for some continuous function $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \mathbb{R}^l \longrightarrow \mathbb{R}$, where $g\left(h, x, y, u\right)$ is the cost of obtaining an a posteriori distribution with the given entropy $h\left(\cdot\right)$ in state x under control u. By the chain rule, we have

$$\begin{aligned}&\nabla_u g_i\left(\xi, x, y, u\right) = \nabla_h g\left(h\left(\xi, x, y, u\right), x, u\right)\nabla_u h\left(\xi, x, y, u\right)\\&+\nabla_u g\left(h\left(\xi, x, y, u\right), x, u\right),\end{aligned}$$

so (16) yields

$$\begin{aligned}&\nabla_u g_i\left(\xi, x, y, u\right) =\\&-2\nabla_h g\left(h\left(\xi, x, y, u\right), x, u\right)\left[y - f(\xi, x, u)\right]^T \Sigma^{-1}(\xi, x)\nabla_u f(\xi, x, u)\\&+\nabla_u g\left(h\left(\xi, x, y, u\right), x, u\right).\end{aligned}$$

Corollary 4.1 *If u^* is an optimal control, then*

$$\begin{aligned}&E\Bigg\{\nabla_u g\left(h\left(\xi, y_j, y_{j+1}, u_j^*\right), y_j, u_j^*\right)\\&+\left(\sum_{i=j}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i^*) - 2\nabla_h g\left(h\left(\xi, y_j, y_{j+1}, u_j^*\right), y_j, u_j^*\right)\right)\\&\times\left(y_{j+1} - f(\xi, y_j, u_j^*)\right)^T \Sigma^{-1}(\xi, y_i)\nabla_u f(\xi, y_j, u_j^*)\Bigg|\mathbb{Y}_j\Bigg\} = 0\end{aligned} \tag{19}$$

for all $j \in \{0, 1, \ldots, N-1\}$.

5. Determining the control

We present a procedure for finding an optimal control $\left\{u_0^*, u_1^*, \ldots, u_{N-1}^*\right\}$ for the stochastic system (1), applying the condition (19), based on the idea of dynamic programming. Set

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

and

$$\begin{aligned}V_j(\xi, y_0, \ldots, y_j) &= E\left[\sum_{i=j}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i)\middle|\mathcal{F}_j\right] \\ &= E\left[g_j(\xi, y_j, y_{j+1}, u_j) + E\left[\sum_{i=j+1}^{N-1} g_i(\xi, y_i, y_{i+1}, u_i)\middle|\mathcal{F}_{j+1}\right]\middle|\mathcal{F}_j\right] \\ &= E\left[g_j(\xi, y_j, y_{j+1}, u_j) + V_{j+1}(\xi, y_0, \ldots, y_{j+1})\right|\mathcal{F}_j] .\end{aligned}$$

By the properties of conditional expectation and the definition of $V_j(\xi, y_0, \ldots, y_j)$, the left hand side (LHS) of (19) can be represented as follows:

$$\begin{aligned}&LHS\ (19) = E\Big\{E\Big\{\nabla_u g\left(h\left(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) w_{j+1}, u_j\right), y_j, u_j\right) \\ &+ [g_j(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) w_{j+1}, u_j) \\ &+ V_{j+1}(\xi, y_0, \ldots, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) w_{j+1}) \\ &- 2\nabla_h g\left(h\left(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) w_{j+1}, u_j\right), y_j, u_j\right)] \cdot \\ &\cdot \left(\sigma(\xi, y_j) w_{j+1}\right)^T \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j)\Big| \mathcal{F}_j\Big\}\Big| \mathbb{Y}_j\Big\} = \\ &\int \Big[\nabla_u g\left(h\left(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) x, u_j\right), y_j, u_j\right) \\ &+ [g_j(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) x, u_j) \\ &+ V_{j+1}(\xi, y_0, \ldots, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) x) \\ &- 2\nabla_h g\left(h\left(\xi, y_j, f(\xi, y_j, u_j) + \sigma(\xi, y_j) x, u_j\right), y_j, u_j\right)] \cdot \\ &\cdot x^T \sigma^T(\xi, y_j) \Sigma^{-1}(\xi, y_j) \nabla_u f(\xi, y_j, u_j)\Big] \gamma\left(x, I_m\right) P\left(d\xi \left|\mathbb{Y}_j\right.\right) dx,\end{aligned}$$

where the conditional distribution $P\left(d\xi \left|\mathbb{Y}_j\right.\right)$ is determined from the Bayes formula:

$$P\left(d\xi \left|\mathbb{Y}_j\right.\right) = \frac{\mu_j(\xi, y_0, y_1, \ldots, y_j)}{\displaystyle\int \mu_j(x, y_0, y_1, \ldots, y_j) dx}$$

and $\mu_j(\xi, y_0, y_1, \ldots, y_j)$ is given by (14).

5.1. Algorithm for determining u^*

1. Define $V_N(\xi, y_0, \ldots, y_N) = 0$ and set $j = N$.
2. Set $j = j - 1$.
3. Define

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

4. Compute

$$Z_j(y_0,\ldots,y_j,u_j) \triangleq$$
$$\int \Big[\nabla_u g\left(h\left(\xi, y_j, f(\xi,y_j,u_j)+\sigma(\xi,y_j)x, u_j\right), y_j, u_j\right)$$
$$+\Big[g_j(\xi,y_j,f(\xi,y_j,u_j)+\sigma(\xi,y_j)x,u_j)+\tilde{V}_{j+1}(\xi,y_0,\ldots,y_j,u_j,x)$$
$$-2\nabla_h g\left(h\left(\xi,y_j,f(\xi,y_j,u_j)+\sigma(\xi,y_j)x,u_j\right),y_j,u_j\right)\Big]\cdot$$
$$\cdot\, x^T\sigma^T(\xi,y_j)\Sigma^{-1}(\xi,y_j)\nabla_u f(\xi,y_j,u_j)\Big]\gamma\left(x,I_m\right)P\left(d\xi\,|\mathbb{Y}_j\right)dx.$$

5. Find u_j^* satisfying (19), i.e.,

$$Z_j(y_0,\ldots,y_j,u_j^*)=0.$$

6. Compute

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

7. If $j=0$ then stop; otherwise go to step 2.

6. Example

Consider the optimal control problem for the one-dimensional system

$$y_{i+1}=\xi-u_i+w_{i+1} \tag{20}$$

where the joint entropy of (20) is

$$H\left(\xi,y_0,\ldots,y_N\right)=\frac{1}{2}E\left[\sum_{j=0}^{N-1}\left(y_{j+1}-\xi+u_j\right)^2\right]+H\left(\xi\right)+H\left(y_0\right)+\frac{N}{2}\ln 2\pi. \tag{21}$$

The task is to minimize (21), i.e. to find

$$\inf_{u\in U}E\left[\sum_{j=0}^{N-1}\left(y_{j+1}-\xi+u_j\right)^2\right].$$

For simplicity let $N=3$. The necessary conditions for optimality are:
for the control u_0^*

$$E\Big\{\left(y_1-\xi+u_0^*\right)$$
$$\times\left[2-\left(y_1-\xi+u_0^*\right)^2-\left(y_2-\xi+u_1^*\right)^2-\left(y_3-\xi+u_2^*\right)^2\right]\Big|\,\mathbb{Y}_0\Big\}=0,$$

for the control u_1^*

$$E\left\{(y_2-\xi+u_1^*)\left[2-(y_2-\xi+u_1^*)^2-(y_3-\xi+u_2^*)^2\right]\middle|\,\mathbb{Y}_1\right\}=0,$$

and for the control u_2^*

$$E\left\{(y_3-\xi+u_2^*)\left[2-(y_3-\xi+u_2^*)^2\right]\middle|\,\mathbb{Y}_2\right\}=0.$$

The algorithm given in the previous section requires about five minutes for calculation of u_0^*, a few seconds for u_1^* and u_2^*. It should be stressed however, that calculations done by the algorithm given in the previous section do not make use explicit forms of sufficient statistics given by the Kalman - Bucy filter.

7. Conclusion

In this paper, the problem of controlling the joint entropy of a system with unknown parameters was stated and solved by applying Rishel's adaptive control methodology. A formal extension of the results of Rishel (1985) enabled us to obtain the necessary conditions for optimality and to construct an algorithm for finding the optimal control. Controlling the joint entropy $H(\xi, y_0, \ldots, y_N)$ is important from the practical point of view, because it models the situations where it is desirable that not only the distribution of ξ, but also the joint distribution of ξ and $(y_0, \ldots, y_N)$ be concentrated for the system being identified. This is of importance for technical systems, where a control intensifying the learning process only, i.e., minimizing the conditional entropy

$$H(\xi\,|y_0,\ldots,y_N) = H(\xi, y_0,\ldots,y_N) - H(y_0,\ldots,y_N)$$

could result in trajectories with entropy $H(y_0,\ldots,y_N)$ too large to be tolerated by the system. Consider e.g. the process of learning the dynamical properties (hydrodynamic resistance coefficients) of a ship just launched. It is evident to any captain that to get as much information as possible one has to manoeuvre the ship with various speeds and under different weather conditions. However, the testing process cannot be completely unpredictable to the captain, that is, $H(y_0,\ldots,y_N)$ cannot be too large. An experienced captain gets to know the maneuvering properties of the ship gradually, slightly intensifying the parameters at each step and drawing conclusions from the previous steps. This means that the conditional entropy is minimized in a long series of experiments which minimize the joint entropy (then $H(y_0,\ldots,y_N)$ is also small) rather than the difference $H(\xi, y_0,\ldots,y_N) - H(y_0,\ldots,y_N)$. Such experiences are also shared by happy owners of new cars - or not necessarily new, but during the first snowfalls.

References

Aoki, M. (1967) *Optimization of Stochastic Systems.* Academic Press.

Banek, T. and Kozłowski, E. (2004) Active and passive learning in control processes. *XV Int. Conf. on System Science*, Wrocław, **II**, 38-48.

Banek, T. and Kulikowski, R. (2003) Information pricing for portfolio optimization. *Control and Cybernetics* **32**, 867-882.

Bellman, R.(1961) *Adaptive Control Processes.* Princeton.

Beneš, V.E., Karatzas, I. and Rishel R. (1991) The separation principle for a Bayesian adaptive control problem with no strict-sense oprimal law. *Stochastic Monographs* **5**, 121-156.

Dai Pra, P., Rudari, C., and Runggaldier, W.J. (1997) On dynamic programming for sequential decision problems under a general form of uncertainty. *ZOR - Mathematical Methods of Operations Research* **45**, 81-107.

Feldbaum, A.A. (1960) Dual control theory. *Automation and Remote Control* **21**, 874-1033.

Feldbaum, A.A. (1961) Dual control theory. *Automation and Remote Control* **22**, 1-109.

Feldbaum, A.A. (1965) *Optimal Control Systems.* Academic Press.

Harris, L. and Rishel, R. (1986) An algorithm for a solution of a stochastic adaptive linear quadratic optimal control problem. *IEEE Transactions on Automatic Control* **31**, 1165-1170.

Kulikowski, R. (1965) *Procesy optymalne i adaptacyjne w układach regulacji automatycznej* (in Polish). PWN, Warszawa.

Liptser, R.Sh., Runggaldier, W.J. and Taksar, M. (1996) Deterministic approximation for stochastic control problems. *SIAM J. Control and Optimization*, **34**, 161-178.

Porosiński, Z., Szajowski, K. and Trybuła, S. (1985) Bayes control for a multidimensional stochastic system. *System Sciences* **11**, 51-64.

Rishel, R. (1985) A nonlinear discrete time stochastic adaptive control problem. *Theory and applications of nonlinear control systems. Sel. Pap. 7th Int. Symp. Math. Theory Networks Systems*, 585-592.

Rishel, R. (1986) An exact formula for a linear quadratic adaptive stochastic optimal control law. *SIAM J. Control and Optimization* **24**, 667-674.

Runggaldier, W.J. and Zaccaria, A. (2000) A stochastic control approach to risk management under restricted information. *Mathematical Finance* **10**, 277-288.

Saridis, G.N. (1988) Entropy formulation of optimal and adaptive control. *IEEE Transactions on Automatic Control* **33**, 713-721.

Saridis, G.N. (1995) *Stochastic Processes, Estimation and Control: the Entropy Approach.* John Wiley & Sons.

Wiener, N. (1948) *Cybernetics.* John Wiley & Sons.

Zabczyk, J. (1996) *Chance and Decision.* Pisa, Scuola Normale Superiore.