Annales UMCS Informatica AI 5 (2006) 343-356

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI

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

Comparison of stochastic optimal controls with different level of self-learning

Tomasz Hordjewicz*, Edward Kozłowski

Department of Quantitative Methods, Lublin University of Technology, Nadbystrzycka 38, 20-618 Lublin, Poland

Abstract

In the article the problem of control of a linear dynamic system with a square coefficient of quality was introduced. Depending on knowledge we have about the parameters of system, the principles of control of the system under conditions of full and incomplete information as well as the differences in values of coefficients of quality were presented. Potential usage of the above mentioned principles may pertain to physical systems for which the supreme aim of control is minimization of energetic losses as well as optimal control of a linear system in real-time, and also to computer systems where the pricing of value of information is required.

1. Introduction

In this paper we consider two types of controls of a linear stochastic system with discrete time: in the first case, the parameters of system as well as equation of state are well known – this is the classic optimal control problem with complete information about the system. The second case concerns the problem with incomplete information about the system, the equation of state is well known, however the parameters of system are unknown (in literature it is represented as a problem of adaptive control). The task of stochastic control (first case) has only element of acting, but the task of adaptive control (second case) has two elements: acting and learning simultaneously. To control accurately, we must possess indispensable knowledge about this fragment of reality which concerns working. Often we do not even have the comfort to possess such knowledge before working. The methodology of adaptive control proposes more active and decided attitude in diagnostics of reality even in comparison to self – tuning control, where we replace unknown parameters by their estimators. The same adaptive control is composed of elements of self-learning and working. The aim of control usually depends on the optimization of

*Corresponding author: *e-mail address*: t.hordjewicz@pollub.pl

performance criteria, which depends on the states of system and the controls. Therefore optimal control has dual nature – it yields fast increase of information about the parameters of system and realizes the purpose. Yet, we should remember that the optimization of performance criteria is the most important aim though knowing unknown parameters, it is necessary to fulfill the aim of control well.

The subject matter of adaptive control have promoted the researchers' interest for a long time with regard to a very wide use in many fields of present science. The relevant literature is very extensive, see e.g. Rishel [1], Harris and Rishel [2], Feldbaum [4], Bellman [5], Kulikowski [6], Beneš, Karatzas and Rishel [7]. The problem of self – tuning control was published in Zabczyk [8], the comparison of adaptive and self – tuning controls as well as the discussion of active accumulation of knowledge were described in Banek and Kozłowski [3]. The application of entropy concept to adaptive control can be found in Saridis [11], Banek, Kozłowski [3].

The aim of this paper is applying the theory of adaptive control to the tasks where the state of system is described by a linear equation, however, the coefficient of quality is square. We use backward induction to determine the optimal control. In the article the principles of control were given in the case when noises as well as parameters of system were modeled by the gaussian distributions. The received principles of control for the linear system can be also used in the case of nongaussian noises. Some technical remarks relating to filtering for expotential noise models we can find in Porosiński, Szajowski and Trybuła [9], Manton, Krishnamurty, Elliot [10]. The ways of determining control applying the Rishel's methodology (first – Gateaux's differentiations combined with conditional expectation properties lead to necessary conditions for optimality, second – application of backward indictions to the necessary conditions leads to the Rishel's algorithm) for the linear system with gaussian noises and different square coefficient of quality are described in Harris and Rishel [2].

The structure of paper is following: in section 2 the expression of linear quadratic control problem was introduced, in sections 3, 4 the principles of optimal control with complete and incomplete information about the system suitably were given. In section 5 the controls in the case of gaussian noises and the differences of value of performance criteria were given. In section 6 an example was presented.

2. The linear quadratic stochastic control

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Suppose that $w_1,..., w_N$ are independent n-dimensional random vectors on this space, with normal $N(0, I_n)$

distribution, let ξ be a k-dimensional vector with a priori distribution $P(d\xi)$. We assume that all the above objects are stochastically independent. On $(\Omega, \mathcal{F}, P)$ we define the family of sub $-\sigma-$ fields $\mathbb{Y}_k = \sigma\{y_i : i = 0,\ 1,...,\ k\}$ and the family of sub $-\sigma-$ fields $\mathcal{F}_k \triangleq \sigma(\xi) \vee \mathbb{Y}_k$. For brevity, we will write $\mathcal{F} = \mathcal{F}_N$.

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

$$y_{i+1} = y_i + B\xi - Cu_i + d + w_{i+1}, \tag{1}$$

where $i = 0,...,N-1$, $y_i \in \mathbb{R}^n$, $d \in \mathbb{R}^n$, $B \in \mathbb{R}^{n\times k}$, $C \in \mathbb{R}^{n\times l}$. A $\mathcal{F}_j$ – measurable vector $u_j \in \mathbb{R}^l$ will be called a control action with complete information, $u = (u_0, u_1, ..., u_{N-1})$ an admissible control and the class of admissible controls with complete information is denoted by U. A $\mathbb{Y}_j$– measurable vector $\tilde{u}_j \in \mathbb{R}^l$ will be called a control action with incomplete information, $\tilde{u} = (\tilde{u}_0, \tilde{u}_1, ..., \tilde{u}_{N-1})$ an admissible control and the class of admissible controls with incomplete information is denoted by $\tilde{U}$.

We introduce a performance criterion

$$J(u) = E\left\{\sum_{i=0}^{N-1} u_i^T R_i u_i + (y_N - a)^T Q_N (y_N - a)\right\}, \tag{2}$$

where $R_i \in \mathbb{R}^{l\times l}$ for $i = 0,...,N-1$ and $Q_N \in \mathbb{R}^{n\times n}$. The functional $u_i^T R_i u_i$ represents energetic losses caused by control u_j but the functional of heredity $(y_N - a)^T Q_N (y_N - a)$ means losses caused by missing to point a.

The task to find is:

$$\inf_{u\in U} J(u) \quad \text{and} \quad \inf_{\tilde{u}\in\tilde{U}} J(\tilde{u}) \tag{3}$$

and to determine an admissible control $u^* = (u_0^*, ..., u_{N-1}^*)$ and $\tilde{u}^* = (\tilde{u}_0^*, ..., \tilde{u}_{N-1}^*)$ for which *inf* are attained.

3. The control with complete information

The control action u_j undertaken at time j in condition of complete information about the parameters of linear system (1) is based on observing the previous states of the system $y_0,\dots,y_j$ and on the knowledge of value vector ξ, i.e.

$$u_j \triangleq u_j(\xi, y_0, ..., y_j).$$

We introduce some notations (Bellmans' functions). For $j = 0,...,N-1$ the coefficient of quality at time j is

$$V_j(\xi, y_j) = \inf_{u_j} E\left\{u_j^T R_j u_j + V_{j+1}(\xi,\ y_{j+1}) \middle| \mathcal{F}_j\right\} \tag{4}$$

with the terminal condition

$$V_N(\xi, y_N) = (y_N - a)^T Q_N (y_N - a).$$

At time j we condition on $\sigma -$field $\mathcal{F}_j$ because we know the parameters of system ξ and states of system $y_0,...,y_j$.

Therefore task (3) we can present

$$\inf_{u \in U} J(u, \xi) = V_0(\xi, y_0).$$

The theorem below gives the solution of problem (3) under condition of complete information.

Theorem 1. *If the matrices* $R_i \in R^{l \times l}$ *for* $i = 0,...,N-1$ *and* $Q_N \in R^{n \times n}$ *are symmetrically and positively determined, then the solution of task (3) under condition of complete information is:*
a) the optimal control of stochastic system (1)

$$u_j^* = \left[R_j + C^T Q_{j+1} C\right]^{-1} C^T Q_{j+1} \left[y_j + (N-j)(B\xi + d) - a\right] \tag{5}$$

b)

$$\begin{aligned} V_j(\xi, y_j) = &\left[y_j + (N-j)(B\xi + d) - a\right]^T Q_j \left[y_j + (N-j)(B\xi + d) - a\right] \\ &+ tr\left[\sum_{i=j+1}^{N} Q_i\right] \end{aligned} \tag{6}$$

where the matrices Q_j *we define as*

$$Q_j = Q_{j+1} - Q_{j+1} C\left[R_j + C^T Q_{j+1} C\right]^{-1} C^T Q_{j+1} \tag{7}$$

for all $j = 1,...,N$.

Proof. We will prove the above theorem by applying dynamic programming (backwards induction). Let us check first for $N-1$.

$$V_{N-1}(\xi, y_{N-1}) = \inf_{u_{N-1}} E\left\{u_{N-1}^T R_{N-1} u_{N-1} + V_N(\xi,\ y_N) \middle| F_{N-1}\right\} = \inf_{u_{N-1}} E\left\{u_{N-1}^T R_{N-1} u_{N-1} + \right.$$

$$\left.\left[y_{N-1} + B\xi + Cu_{N-1} + d + w_N - a\right]^T Q_N \left[y_{N-1} + B\xi + Cu_{N-1} + d + w_N - a\right] \middle| \mathcal{F}_{N-1}\right\}.$$

By the properties of conditional expectation (see e.g. Liptser, Shiryaev [12]) we have

$$E\left\{w_N^T Q_N w_N \middle| F_{N-1}\right\} = tr\left[Q_N\right]$$

and

$$V_{N-1}(\xi, y_{N-1}) = \inf_{u_{N-1}} \Big\{ u_{N-1}^T \left[R_{N-1} + C^T Q_N C \right] u_{N-1} + tr\left[Q_N \right] - 2u_{N-1}^T C^T Q_N \\ \left[y_{N-1} + B\xi + d - a \right] + \left[y_{N-1} + B\xi + d - a \right]^T Q_N \left[y_{N-1} + B\xi + d - a \right] \Big\}. \tag{8}$$

Hence, the optimal control is

$$u_{N-1}^* = \left[R_{N-1} + C^T Q_N C \right]^{-1} C^T Q_N \left[y_{N-1} + B\xi + d - a \right]. \tag{9}$$

Substituting optimal control (9) to (8) the coefficient of quality at time $N-1$ can be presented as

$$V_{N-1}(\xi, y_{N-1}) = \left[y_{N-1} + B\xi + d - a \right]^T \left[Q_N - Q_N C \left[R_{N-1} + C^T Q_N C \right]^{-1} C^T Q_N \right] \times \\ \times \left[y_{N-1} + B\xi + d - a \right] + tr\left[Q_N \right].$$

We define $Q_{N-1} = Q_N - Q_N C \left[R_{N-1} + C^T Q_N C \right]^{-1} C^T Q_N$. The coefficient of quality at time $N-1$ can be presented as

$$V_{N-1}(\xi, y_{N-1}) = \left[y_{N-1} + B\xi + d - a \right]^T Q_{N-1} \left[y_{N-1} + B\xi + d - a \right] + tr\left[Q_N \right]. \tag{10}$$

We suppose that formulas (5)-(6) are true for any $j = 1, 2, \ldots, N-1$ and we check (5)-(6) are true for $(j-1)$.

$$V_{j-1}(\xi, y_{j-1}) = \inf_{u_{N-j-1}} E\left\{ u_{j-1}^T R_{j-1} u_{j-1} + V_j(\xi, y_j) \middle| \mathcal{F}_{j-1} \right\} \\ = \inf_{u_{j-1}} E\left\{ u_{j-1}^T R_{j-1} u_{j-1} + tr\left[\sum_{i=j+1}^{N} Q_i \right] + \left[y_j + (N-j)(B\xi + d) - a \right]^T \right. \\ \left. Q_j \left[y_j + (N-j)(B\xi + d) - a \right] \middle| \mathcal{F}_{j-1} \right\}.$$

From (1)

$$V_{j-1}(\xi, y_{j-1}) = \inf_{u_{j-1}} E\left\{ u_{j-1}^T R_{j-1} u_{j-1} + tr\left[\sum_{i=j+1}^{N} Q_i \right] \right. \\ + \left[y_{j-1} + (N-j+1)(B\xi + d) - Cu_{j-1} + w_j - a \right]^T Q_j \\ \left. \times \left[y_{j-1} + (N-j+1)(B\xi + d) - Cu_{j-1} + w_j - a \right] \middle| \mathcal{F}_{j-1} \right\}.$$

By the properties of conditional expectation (see e.g. Liptser, Shiryaev [12]) we have

$$\begin{aligned} V_{j-1}\left(\xi, y_{j-1}\right) = \inf_{u_{j-1}} \Bigg\{ & u_{j-1}^T \left[R_{j-1} + C^T Q_j C \right] \ u_{j-1} + tr\left[\sum_{i=j}^{N} Q_i \right] \\ & - 2u_{j-1}^T C^T Q_j \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right] \\ & + \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right]^T \\ & Q_j \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right] \Bigg\}. \end{aligned} \tag{11}$$

Hence, the optimal control is

$$u_{j-1}^* = \left[R_{j-1} + C^T Q_j C \right]^{-1} C^T Q_j \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right]. \tag{12}$$

Substituting optimal control (12) to (11) we have

$$\begin{aligned} V_{j-1}\left(\xi, y_{j-1}\right) = tr\left[\sum_{i=j}^{N} Q_i \right] + & \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right]^T \\ & Q_{j-1} \left[y_{j-1} + \left(N - j + 1 \right)\left(B\xi + d \right) - a \right] \end{aligned} \tag{13}$$

where

$$Q_{j-1} = Q_j - Q_j C \left[R_{j-1} + C^T Q_j C \right]^{-1} C^T Q_j .$$

4. The control with incomplete information

We have partial knowledge about the system in conditions of incomplete information, and namely we know only the dynamics of behavior of the system, which is described by equation (1). We do not know the exact values of parameters of system (the vector ξ is responsible for them), however we possess the knowledge about *a piori* distribution of random vector ξ. Thus, the control action u_j undertaken at time j under condition of incomplete information about the parameters of linear system (1) is based on observing the previous states of the system $y_0, ..., y_j$, i.e.

$$\tilde{u}_j \triangleq \tilde{u}_j \left(y_0, ..., y_j \right).$$

We introduce some notations (Bellmans' functions). For $j = 0, ..., N-1$ the coefficient of quality at time j is

$$W_j\left(y_j\right) = \inf_{\tilde{u}_j} E\left\{ \tilde{u}_j^T R_j \tilde{u}_j + V_{j+1}\left(\xi, y_{j+1}\right) \middle| \mathbb{Y}_j \right\} \tag{14}$$

with the terminal condition

$$W_N\left(y_N\right) = \left(y_N - a\right)^T Q_N \left(y_N - a\right).$$

At time j we condition on $\sigma -$field $\mathbb{Y}_j$ because we know only states of system $y_0, ..., y_j$.

Therefore task (3) we can present as

$$\inf_{\tilde{u}\in U} J(u) = W_0(y_0).$$

The solution of problem (3) under condition of incomplete information gives theorem 2.

Lemma 1. *Let* ξ *be k-dimensional random vector,* $\Psi \in R^{k\times k}$ *be any deterministic matrix, however,* $\{\mathbb{Y}_j\}_{j\geq 0}$ *represented the undecreasing family of* $\sigma-$*fields on* $(\Omega, \mathcal{F}, P)$. *If* Σ_j *represented the second central moment of random vector* ξ *conditioned on the* $\sigma-$*field* $\mathbb{Y}_j$

$$\Sigma_j = E\Big[\big[\xi - E(\xi|\mathbb{Y}_j)\big]\ \big[\xi - E(\xi|\mathbb{Y}_j)\big]^T \Big| \mathbb{Y}_j\Big],$$

then for any $0 \leq i < j$ *we have:*

$$\begin{aligned} E\Big[\xi^T \Psi E(\xi|\mathbb{Y}_j)\Big|\mathbb{Y}_i\Big] &= E\Big[E(\xi|\mathbb{Y}_j)^T \Psi E(\xi|\mathbb{Y}_j)\Big|\mathbb{Y}_i\Big] = \\ &= E(\xi|\mathbb{Y}_i)^T \Psi E(\xi|\mathbb{Y}_i) + tr\big[\Psi(\Sigma_i - \Sigma_j)\big] \end{aligned} \tag{15}$$

and

$$E\Big[\xi^T \Psi \xi \Big| \mathbb{Y}_i\Big] = E(\xi|\mathbb{Y}_i)^T \Psi E(\xi|\mathbb{Y}_i) + tr(\Psi\Sigma_i). \tag{16}$$

Proof. By the properties of conditional expectation for any $0 \leq i < j$ we have

$$\begin{aligned} &E\Big[\xi^T \Psi E(\xi|\mathbb{Y}_j)\Big|\mathbb{Y}_i\Big] = E\Big[E\Big[\xi^T \Psi E(\xi|\mathbb{Y}_j)\Big|\mathbb{Y}_j\Big]\Big|\mathbb{Y}_i\Big] \\ &= E\Big[E(\xi|\mathbb{Y}_j)^T \Psi E(\xi|\mathbb{Y}_j)\Big|\mathbb{Y}_i\Big] = tr\Big(E\Big[\Psi E(\xi|\mathbb{Y}_j) E(\xi|\mathbb{Y}_j)^T \Big|\mathbb{Y}_i\Big]\Big) \\ &= tr\Big\{\Psi E\Big[E(\xi\xi^T|\mathbb{Y}_j) - \Sigma_j\Big|\mathbb{Y}_i\Big]\Big\} = tr\Big\{\Psi E(\xi\xi^T|\mathbb{Y}_i)\Big\} - tr\{\Psi\Sigma_j\} \\ &= tr\Big(\Psi\Big[E(\xi|\mathbb{Y}_i)E(\xi|\mathbb{Y}_i)^T + \Sigma_i\Big]\Big) - tr\{\Psi\Sigma_j\} \\ &= E(\xi|\mathbb{Y}_i)^T \Psi E(\xi|\mathbb{Y}_i) + tr\big[\Psi(\Sigma_i - \Sigma_j)\big]. \end{aligned}$$

However

$$\begin{aligned} E\Big[\xi^T \Psi \xi\Big|\mathbb{Y}_i\Big] &= tr\Big\{\Psi E(\xi\xi^T|\mathbb{Y}_i)\Big\} \\ &= tr\Big(\Psi\Big[E(\xi|\mathbb{Y}_i)E(\xi|\mathbb{Y}_i)^T + \Sigma_i\Big]\Big) \\ &= E(\xi|\mathbb{Y}_i)^T \Psi E(\xi|\mathbb{Y}_i) + tr(\Psi\Sigma_i). \end{aligned}$$

Theorem 2. *If the matrices* $R_i \in R^{l\times l}$ *for* $i = 0,...,N-1$ *and* $Q_N \in R^{n\times n}$ *are symmetrically and positively determined, then the solution of task (3) under condition of incomplete information is:*
a) the optimal control of stochastic system (1)

$$\tilde{u}_j^* = \left[R_j + C^T Q_{j+1} C\right]^{-1} C^T Q_{j+1}\left[y_j + (N-j)\ \left(BE\left[\xi \middle| \mathbb{Y}_j\right] + d\right) - a\right] \tag{17}$$

b)

$$\begin{aligned} W_j\left(y_j\right) = &\left[y_j + (N-j)\left(BE\left[\xi \middle| \mathbb{Y}_j\right] + d\right) - a\right]^T \\ &Q_j\left[y_j + (N-j)\left(BE\left[\xi \middle| \mathbb{Y}_j\right] + d\right) - a\right] \\ &+ tr\left[\sum_{k=j}^{N-2} B^T Q_{k+1} B\left[(N-k)^2 \Sigma_k - \left((N-k)^2 - 1\right)\Sigma_{k+1}\right]\right. \\ &\left. + B^T Q_N B \Sigma_{N-1} + \sum_{i=j+1}^{N} Q_k\right], \end{aligned} \tag{18}$$

where $\Sigma_j = E\left(\left[\xi - m_j\right]\left[\xi - m_j\right]^T \middle| \mathbb{Y}_j\right)$ *and matrices* Q_j *are defined by (7) for any* $j = 1,...,N$.

Proof. We will prove the above theorem by applying backwards induction. Let us check first for $j = N-1$.

$$W_{N-1}\left(y_{N-1}\right) = \inf_{\tilde{u}_{N-1}} E\left\{\tilde{u}_{N-1}^T R_{N-1}\tilde{u}_{N-1} + W_N\left(y_N\right) \middle| \mathbb{Y}_{N-1}\right\} = \inf_{\tilde{u}_{N-1}} E\left\{\tilde{u}_{N-1}^T R_{N-1}\tilde{u}_{N-1} + \right.$$
$$\left.\left(y_{N-1} + B\xi + C\tilde{u}_{N-1} + d + w_N - a\right)^T Q_N\left(y_{N-1} + B\xi + C\tilde{u}_{N-1} + d + w_N - a\right) \middle| \mathbb{Y}_{N-1}\right\}.$$

By the properties of conditional expectation (see e.g. Liptser, Shiryaev [12]) we have

$$E\left\{w_N^T Q_N w_N \middle| \mathbb{Y}_{N-1}\right\} = tr\left[Q_N\right]$$

and

$$\begin{aligned} W_{N-1}\left(y_{N-1}\right) = \inf_{\tilde{u}_{N-1}} E&\left\{\tilde{u}_{N-1}^T\left[R_{N-1} + C^T Q_N C\right]\ \tilde{u}_{N-1}\right. \\ &- 2\tilde{u}_{N-1}^T C^T Q_N\left(y_{N-1} + B\xi + d - a\right) \\ &\left. + \left(y_{N-1} + B\xi + d - a\right)^T Q_N\left(y_{N-1} + B\xi + d - a\right) \middle| \mathbb{Y}_{N-1}\right\} \\ &+ tr\left[Q_N\right]. \end{aligned} \tag{19}$$

Hence, the optimal control is

$$\tilde{u}_{N-1}^* = \left[R_{N-1} + C^T Q_N C\right]^{-1} C^T Q_N\left(y_{N-1} + BE\left[\xi \middle| \mathbb{Y}_{N-1}\right] + D - a\right). \tag{20}$$

Substituting optimal control (20) to (19) the coefficient of quality at time $N-1$ can be presented as

$$\begin{aligned} W_{N-1}\left(y_{N-1}\right) = &-\left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right)^T \\ &\times Q_N C\left[R_{N-1}+C^T Q_N C\right]^{-1} C^T Q_N \left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right) \\ &+E\left[\left(y_{N-1}+B\xi+d-a\right)^T Q_N \left(y_{N-1}+B\xi+d-a\right)\middle| \mathbb{Y}_{N-1}\right]+tr\left[Q_N\right]. \end{aligned}$$

From lemma 1

$$\begin{aligned} &E\left[\left(y_{N-1}+B\xi+d-a\right)^T Q_N \left(y_{N-1}+B\xi+d-a\right)\middle| \mathbb{Y}_{N-1}\right]=tr\left[B^T Q_N B\Sigma_{N-1}\right] \\ &+\left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right)^T Q_N \left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right). \end{aligned}$$

We define $Q_{N-1}=Q_N-Q_N C\left[R_{N-1}+C^T Q_N C\right]^{-1} C^T Q_N$. The coefficient of quality at time $N-1$ can be presented as

$$\begin{aligned} W_{N-1}\left(y_{N-1}\right) = &\left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right)^T \\ &Q_{N-1}\left(y_{N-1}+BE\left[\xi \middle| \mathbb{Y}_{N-1}\right]+d-a\right) \\ &+tr\left[Q_N+B^T Q_N B\Sigma_{N-1}\right]. \end{aligned} \tag{21}$$

We suppose that formulas (17)-(18) are true for any $j=1,\ 2,\ldots,N-1$ and we check (17)-(18) are true for $\left(j-1\right)$. Thus

$$\begin{aligned} W_{j-1}\left(y_{j-1}\right) = \inf_{\tilde{u}_{j-1}} E\left\{\tilde{u}_{j-1}^T R_{j-1}\tilde{u}_{j-1}+W_j\left(y_j\right)\middle| \mathbb{Y}_{j-1}\right\} = \inf_{\tilde{u}_{j-1}} E\Bigg\{\tilde{u}_{j-1}^T R_{j-1}\tilde{u}_{j-1} \\ +\left[y_j+\left(N-j\right)\left(BE\left[\xi \middle| \mathbb{Y}_j\right]+d\right)-a\right]^T \\ Q_j\left[y_j+\left(N-j\right)\left(BE\left[\xi \middle| \mathbb{Y}_j\right]+d\right)-a\right] \\ +tr\Bigg[\sum_{k=j}^{N-2} B^T Q_{k+1} B\left(k^2\Sigma_k-\left(k^2-1\right)\ \Sigma_{k+1}\right) \\ +B^T Q_N B\Sigma_{N-1}+\sum_{k=j+1}^{N} Q_k\Bigg]\Bigg\| \mathbb{Y}_{j-1}\Bigg\}. \end{aligned}$$

From (1) and properties of conditional expectation we can present the above equation as

$$\begin{aligned} W_{j-1}\left(y_{j-1}\right) = \inf_{\tilde{u}_{j-1}} \Bigg\{ & E\Big[\Big[y_{j-1} + B\Big(\xi + (N-j) E\big[\xi \big| \mathbb{Y}_j\big]\Big) + (N-j+1)\ \ d - a\Big]^T Q_j \\ & \times\Big[y_{j-1} + B\Big(\xi + (N-j)\ \ E\big[\xi \big| \mathbb{Y}_j\big]\Big) + (N-j+1)\ \ d - a\Big] \Big| \mathbb{Y}_{j-1}\Big] \\ & + \tilde{u}_{j-1}^T\Big[R_{j-1} + C^T Q_j C\Big]\ \ \tilde{u}_{j-1} \\ & - 2\tilde{u}_{j-1}^T C^T Q_j \Big[y_{j-1} + (N-j+1)\ \ \Big(BE\big[\xi \big| \mathbb{Y}_{j-1}\big] + d\Big) - a\Big] \\ & + tr\Bigg[\sum_{k=j}^{N-2} B^T Q_{k+1} B\Big[(N-k)^2\ \Sigma_k - \Big((N-k)^2 - 1\Big)\ \ \Sigma_{k+1}\Big] \\ & + B^T Q_N B \Sigma_{N-1} + \sum_{k=j}^{N} Q_k \Bigg]\Bigg\}. \end{aligned} \tag{22}$$

The optimal control at time $j-1$ is

$$\tilde{u}_{j-1}^* = \Big[R_{j-1} + C^T Q_j C\Big]^{-1} C^T Q_j \Big[y_{j-1} + (N-j+1)\Big(BE\big[\xi \big| \mathbb{Y}_{j-1}\big] + d\Big) - a\Big]. \tag{23}$$

By lemma 1 we have

$$\begin{aligned} & E\Big[\Big[B\Big(\xi + kE\big[\xi \big| \mathbb{Y}_j\big]\Big)\Big]^T Q_j \Big[B\Big(\xi + kE\big[\xi \big| \mathbb{Y}_j\big]\Big)\Big] \Big| \mathbb{Y}_{j-1}\Big] = \\ & (k+1)^2\ E\big[\xi \big| \mathbb{Y}_{j-1}\big] B^T Q_j B E\big[\xi \big| \mathbb{Y}_{j-1}\big] + tr\Big[(k+1)^2\ B^T Q_j B\Big(\Sigma_{j-1} - \Sigma_j\Big) + B^T Q_j B \Sigma_j\Big] \end{aligned}$$

for any $k \in \mathbb{R}$ and $j = 0,\ 1, \ldots, N$. Hence

$$\begin{aligned} & E\Big[\Big[y_{j-1} + B\Big(\xi + (N-j) E\big[\xi \big| \mathbb{Y}_j\big]\Big) + (N-j+1)\ \ d - a\Big]^T Q_j \\ & \times\Big[y_{j-1} + B\Big(\xi + (N-j) E\big[\xi \big| \mathbb{Y}_j\big]\Big) + (N-j+1)\ \ d - a\Big] \Big| \mathbb{Y}_{j-1}\Big] = \\ & \Big[y_{j-1} + (N-j+1)\Big(BE\big[\xi \big| \mathbb{Y}_{j-1}\big] + d\Big) - a\Big]^T \\ & Q_j \Big[y_{j-1} + (N-j+1)\Big(BE\big[\xi \big| \mathbb{Y}_{j-1}\big] + d\Big) - a\Big] \\ & + tr\Big[B^T Q_j B\Big((N-j+1)^2\ \Sigma_{j-1} - \Big((N-j+1)^2 - 1\Big)\Sigma_j\Big)\Big]. \end{aligned} \tag{24}$$

Finally, by substituting (23)-(24) to (22) we obtain

$$W_{j-1}\left(y_{j-1}\right)=\left[y_{j-1}+\left(N-j+1\right)\left(BE\left[\xi\middle|\mathbb{Y}_{j-1}\right]+d\right)-a\right]^{T}$$
$$\times Q_{j-1}\left[y_{j-1}+\left(N-j+1\right)\left(BE\left[\xi\middle|\mathbb{Y}_{j-1}\right]+d\right)-a\right]$$
$$+tr\left[\sum_{k=j-1}^{N-2}B^{T}Q_{k+1}B\left[\left(N-k\right)^{2}\Sigma_{k}-\left(\left(N-k\right)^{2}-1\right)\ \Sigma_{k+1}\right]\right.$$
$$\left.+B^{T}Q_{N}B\Sigma_{N-1}+\sum_{k=j}^{N}Q_{k}\right]$$

which proves the assertion.

5. Comparison of optimal controls with complete and incomplete information

We assume that *a piori* distribution of random vector ξ is a normal distribution $N\left(m_0,\Sigma_0\right)$.

Lemma 2 (the optimal filtration of conditionally normal sequences). *If the random vector ξ has a priori normal $N\left(m_0,\Sigma_0\right)$ distribution and the system is described by the linear state equation*

$$y_{i+1}=f_1\left(y_i\right)+f_2\left(y_i\right)\xi+\sigma\left(y_i\right)w_{i+1}$$

then we have:

1. the conditional distribution $P\left(\xi\middle|\mathbb{Y}_j\right)$ is normal $N\left(m_j,\Sigma_j\right)$;

2. the best estimator of ξ, $m_j=E\left(\xi\middle|\mathbb{Y}_j\right)$ and the conditional covariance matrix $\Sigma_j=E\left(\left[\xi-m_j\right]\left[\xi-m_j\right]^{T}\middle|\mathbb{Y}_j\right)$ are given by

$$m_j=\left[I+\Sigma_0\sum_{i=0}^{j-1}f_2^{T}\left(y_i\right)\left(\sigma\left(y_i\right)\ \sigma^{T}\left(y_i\right)\right)^{-1}f_2\left(y_i\right)\right]^{-1}$$
$$\times\left[m_0+\Sigma_0\sum_{i=0}^{j-1}f_2^{T}\left(y_i\right)\left(\sigma\left(y_i\right)\sigma^{T}\left(y_i\right)\right)^{-1}\left[y_{i+1}-f_1\left(y_i\right)\right]\right] \tag{25}$$

and

$$\Sigma_j=\left[I+\Sigma_0\sum_{i=0}^{j-1}f_2^{T}\left(y_i\right)\left(\sigma\left(y_i\right)\sigma^{T}\left(y_i\right)\right)^{-1}f_2\left(y_i\right)\right]^{-1}\Sigma_0\,. \tag{26}$$

The proof of the above lemma and more information of optimal filtration of conditionally normal sequences we can find in Liptser, Shiryaev [12], chap. 13.

Corollary 1. *If the stochastic linear system is represented by state equation (1) then the conditional expected value of random vector* ξ, $m_j = E\left(\xi \middle| \mathbb{Y}_j\right)$ *and the conditional covariance matrix* $\Sigma_j = E\left(\left[\xi - m_j\right]\left[\xi - m_j\right]^T \middle| \mathbb{Y}_j\right)$ *are described by*

$$m_j = \left[I + \Sigma_0 \sum_{i=0}^{j-1} B^T B\right]^{-1} \left[m + \Sigma_0 \sum_{i=0}^{j-1} B^T \left[y_{i+1} - y_i + Cu_i - d\right]\right] \tag{27}$$

and

$$\Sigma_j = \left[I + \Sigma_0 \sum_{i=0}^{j-1} B^T B\right]^{-1} \Sigma_0 . \tag{28}$$

Thus, for any $j = 0, 1, \ldots, N-1$ the optimal control under conditions of complete information is

$$u_j^* = \left[R_j + C^T Q_{j+1} C\right]^{-1} C^T Q_{j+1} \left[y_j + (N-j)\ (B\xi + d) - a\right] \tag{29}$$

but under conditions of incomplete information is

$$\tilde{u}_j^* = \left[R_j + C^T Q_{j+1} C\right]^{-1} C^T Q_{j+1} \left[y_j + (N-j)\left(Bm_j + d\right) - a\right]. \tag{30}$$

The surplus of losses due to the unknown parameters of system (1) is

$$\begin{aligned} W_j\left(y_j\right) - V_j\left(\xi,\ y_j\right) &= V_j\left(m_j,\ y_j\right) - V_j\left(\xi,\ y_j\right) + \\ &+ tr\left[\sum_{k=j}^{N-2} B^T Q_{k+1} B\left[(N-k)^2 \Sigma_k - \left((N-k)^2 - 1\right)\ \Sigma_{k+1}\right] + B^T Q_N B \Sigma_{N-1}\right], \end{aligned} \tag{31}$$

where Q_j is given by (7), however m_j and Σ_j are described by (27) and (28) suitably.

6. Example

We present the control problem for a system with the state equation

$$y_{i+1} = y_i + B\xi - Cu_i + d + \sigma w_{i+1}$$

for $N = 10$, $y_i \in \mathbb{R}^2$, $d \in \mathbb{R}^2$, $B \in \mathbb{R}^{2\times 2}$, $C \in \mathbb{R}^{2\times 2}$, $\sigma \in \mathbb{R}^{2\times 2}$. Let

$$R_j = \begin{bmatrix} 0.27 & 0.02 \\ 0.01 & 0.35 \end{bmatrix}, \quad \text{for } j = 0,\ 1, \ldots,\ 9,\ Q_N = \begin{bmatrix} 17 & 1.3 \\ 1.3 & 15 \end{bmatrix}, \quad a = \begin{bmatrix} 100 \\ 120 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.3 & -0.4 \\ -0.6 & 1.8 \end{bmatrix}, \quad C = \begin{bmatrix} 2.4 & -0.2 \\ -0.3 & 2.1 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 1.3 & -0.4 \\ -0.4 & 1.5 \end{bmatrix}, \quad d = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Under conditions of complete information of system (1) we assume that the vector of parameters of system is

$$\xi = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

but under conditions of incomplete information of system (1) we assume that the random vector ξ of parameters of system has a normal distribution $N\left(m_0, \Sigma_0\right)$, where

$$m_0 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}.$$

The table below presents y_i – the states of system (1) and optimal controls under conditions of complete and incomplete information. The optimal controls u_i^* and $\tilde{u}_j^*$ we determine from (29) and (30) suitably.

The coefficienties of quality with complete and incomplete information about the parameters of system (1) we determine from (6) and (18)suitably and surplus of losses due to the unknown parameters of system we determine from (31).

	y_i	u_i^*	$\tilde{u}_j^*$	$V_j\left(\xi, y_j\right)$	$W_j\left(y_j\right)$	$W_j\left(y_j\right) - V_j\left(\xi, y_j\right)$
$i = 0$	(0;0)	-3.62;-4.89)	(-4.35;-6.65)	126.35	362.31	235.96
$i = 1$	(11.88;10.39)	(-3.5;-4.95)	(-3.87;-6.81)	113.94	337.5	223.56
$i = 2$	(20.95;22.4)	(-3.52;-4.96)	(-4.23;-6.64)	101.67	315.62	213.95
$i = 3$	(32.34;32.63)	(-3.38;-5.07)	(-3.99;-6.88)	90.18	298.46	208.28
$i = 4$	(41.24;47.25)	(-3.38;-4.89)	(-4.21;-6.36)	73.35	263.68	190.33
$i = 5$	(50.25;60.48)	(-3.38;-4.78)	(-4.41;-6.09)	58.97	239.24	180.27
$i = 6$	(61.78;69.82)	(-3.11;-5.03)	(-4.11;-6.51)	48.86	229.32	180.46
$i = 7$	(70.74;79.28)	(-3.08;-5.46)	(-4.22;-7.27)	41.42	220.12	178.7
$i = 8$	(79.82;92.78)	(-2.92;-5.39)	(-4.24;-7.41)	26.22	195.69	169.47
$i = 9$	(86.17;106.72)	(-3.59;-5.09)	(-5.56;-7.9)	13.1	174.43	161.33

Conclusions

Above are presented the principles of control when the noises and parameters of system are modeled by the gaussian distributions. However, the received principles of control can also be used in the case of nongaussian distributions, i.e. noises as well as *a piori* distributions of random vector ξ modeled by e.g. Poisson, binominal, gamma distributions which are very frequently considered in practice.

In the case of system with incomplete information, we have considered “passive” learning process, we use separation principle, e.g. we determine the control action for the system with complete information, and substitute the estimator in place of unknown parameters, or we condition the control action on the σ – field $\mathbb{Y}_j$.

$$\tilde{u}_j\left(y_0,\ldots,\ y_j\right)\triangleq u_j\left(E\left(\xi\middle|\mathbb{Y}_j\right),y_0,\ldots,\ y_j\right)\triangleq E\left(u_j\left(\xi,\ y_0,\ldots,\ y_j\right)\middle|\mathbb{Y}_j\right).$$

In the article we presented the value of surplus of losses due to the unknown parameters of system. It can be used to pricing of value of information about parameters of the system.

8. References

[1] Rishel R., *A nonlinear discrete time stochastic adaptive control problem*, Theory and applications of nonlinear control systems, Sel. Pap. 7th Int. Symp. Math. Theory Networks Systems, (1985) 585.
[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 (1986) 1165.
[3] Banek T., Kozłowski E., *Active and passive learning in control processes application of the entropy concept*, Systems Science, 31(2) (2005) 29.
[4] Feldbaum A.A., *Optimal Control Systems*, Academic Press, (1965).
[5] Bellman R., *Adaptive control processes*, Princeton, (1961).
[6] Kulikowski R., *Procesy optymalne i adaptacyjne w układach regulacji automatycznej*, Warszawa, PWN (1965), in Polish.
[7] Beneš V.E., Karatzas I., Rishel R., *The separation principle for a Bayesian adaptive control problem with no strict-sense oprimal law*, Stochastic Monographs, 5 (1991) 121.
[8] Zabczyk J., *Chance and decision*, Pisa, Scuola Normale Superiore (1996).
[9] Porosiński Z., Szajowski K., Trybuła S., *Bayes control for a multidimensional stochastic system*, Systems Science, 11 (1985) 51.
[10] Manton J. H., Krishnamurty V., Elliot R.J., *Discrete time filters for doubly stochastic poisson processes and other exponential noise models*, International Journal of Adaptive Control and Signal Processing, 13 (1999) 393.
[11] Saridis G., N., *Entropy formulation of optimal and adaptive control*, IEEE Transactions on Automatic Control, 33 (1988) 713.
[12] Liptser R.Sh., Shiryaev A.N., *Statistics of Stochastic Processes*, New York: Springer-Verlag, (1978).