Adaptive Optimization in Stochastic Systems via Fiducial Approach

Nicholas A. Nechval and Konstantin N. Nechval

Department of Applied Mathematics, Aviation University of Riga
Lomonosov Street 1, LV-1019 Riga, Latvia
fax: +371 7 100660, e-mail: rau@rau.lv

Abstract

In this paper, the problem of determining the optimal control law for discrete-time stochastic linear systems with respect to a quadratic performance criterion is considered. It is assumed that the system is subject to additive system noise and that the state variables are measured with additive measurement noise, without specifying the specific characteristics of random variables. It is shown that the problem of stochastic optimal control can be reduced to two independent problems, one of equivalent deterministic optimal control and the other of stochastic estimation of underlying uncertainties. This holds even if the system noise, the measurement noise and/or the initial state of the system are non-Gaussian, mutually and time-wise dependent. The aim of the present paper is to show how the invariant embedding technique and fiducial approach may be used to solve the problem of adaptive cautious controlling a discrete-time stochastic linear system in which the state transition matrix and the control driven matrix are unknown. This is the case when the certainty equivalence principle does not yield the admissible adaptive control laws for the present problem. The proposed approach does not require the arbitrary selection of priors as in the Bayesian approach. It makes it possible to simplify the problem of adaptive optimization of stochastic systems and, if the system noise and/or the measurement noise are Gaussian, to carry out the algorithm in closed form. The examples are given to illustrate the suggested methodology.

Keywords: Stochastic System, Adaptive Optimization, Fiducial Approach.

1 Introduction

This paper is concerned with the optimization problem of discrete-time stochastic linear systems with respect to a quadratic performance criterion. It is assumed that the system is subject to additive system noise and that the state variables are measured with additive measurement noise. It is known (see, for example, Aoki, 1967; Yakowitz, 1969; Åström, 1970; Åström and Eykhoff, 1971; Bertsekas, 1976; Nechval, 1984) that the aforementioned problem belongs to the class of problems for which the certainty equivalence principle holds. This principle holds if it is possible to first solve the deterministic problem with known parameters and then obtain the optimal controller for unknown parameters by substituting the true parameter values with the estimated values. It is shown that in the case considered the certainty equivalence principle holds even if the system noise, the measurement noise and/or the initial state of the system are non-Gaussian, mutually and time-wise dependent.

In adaptive control there are very few cases where the certainty equivalence principle is applicable. One exception is when the unknown parameters are stochastic variables that are independent between different sampling intervals. However, certainty equivalence principle has been successfully used as an ad hoc design principle. It is important to consider the information, $\mathscr{I}$, which is transformed from the estimator to the controller. The information, $\mathscr{I}$, might be the estimates of the unknown parameters, $\{\hat{\theta}\}$, or the estimates and the uncertainties of the estimates, $\{\hat{\theta}, P\}$, etc. The controllers obtained by enforcing the certainty equivalence principle, i.e. $\mathscr{I} = \{\hat{\theta}\}$, are called certainty equivalence controllers. These controllers do not take into consideration the fact that the estimated parameters are not equal to the true ones but are inaccurate. If the above fact is taken into consideration, i.e. that the information pattern is changed to $\mathscr{I} = \{\hat{\theta}, P\}$ and the separation principle is applied, then the controller will be called cautious. In this case the controller is aware of the errors in the estimates and takes a more cautious control action.

If the performance index only takes into account the previous measurements and does not assume that further information will be available then the resulting controller in Feldbaum's terminology (Feldbaum, 1965) will be called non-dual. On the other hand, the performance index can also be dependent on the future observations and this will result in a dual controller. The controller must, of course, be causal and the dependence of the future observations will be given as the probability distributions of the future observations given information up to the actual time. In a dual controller there is interaction between the identification and the control in the sense that the controller must compromise between a control action and a probing action. The interaction is obtained by considering that the future uncertainties of the parameters are functions of the control signals applied to the system. The loss function, which has to be minimized with respect to the control signal, thus contains some information of the future observations through the statistics of the observations given the present information. According to the above discussion the minimization of a loss function one step ahead will give a non-dual controller, while a minimization several steps ahead will give a dual controller. In the first case it is useless for the controller to take a control action in order to increase the accuracy of the unknown parameters. In the second case it might be worthwhile for the controller to take some control actions in order to improve the estimates of the unknown parameters. The dual controller must thus ensure good control and good estimation. However, these two tasks are, in general, contradictory, since good estimation might require large control signals while good control might require that the control signals are small. A dual controller thus must compromise between these two tasks. The formal solution of the dual control problem has been known for a long time (see for instance Feldbaum, 1965; Aoki, 1967; Nechval, 1982, 1984, 1999; Nechval et al., 1997). The solution leads, however, to a functional equation, which in most cases is difficult to solve.

This paper deals with the non-dual adaptive controllers that can be divided into two classes, certainty equivalence and cautious controllers. The first class includes methods where enforced certainty equivalence has been used as an ad hoc design method. The second class contains methods obtained by using the separation principle. It is shown that the cautious controllers are better than the certainty equivalence controllers.

In the present paper we consider the problem of determining the optimal control law for discrete-time linear systems, subject to additive system noise, measured under additive measurement noise, with respect to a quadratic performance criterion. It is shown how the invariant embedding technique and fiducial approach may be used to solve the problem of cautious adaptive controlling a discrete-time stochastic linear system in which the state transition matrix and the control driven matrix are unknown. In this case the certainty equivalence principle does not yield the admissible adaptive control laws for the present problem. The proposed approach does not require the arbitrary selection of priors as in the Bayesian approach. It makes it possible to simplify the problem of adaptive optimization of stochastic systems and, if the system noise and/or the measurement noise are Gaussian, to carry out the algorithm in closed form. The examples are given to illustrate the suggested methodology.

The outline of the paper is as follows. A formulation of the problem is presented in Section 2. A technique for determination of control laws is found in Section 3. In order to illustrate the proposed technique, examples are given in Section 4.

2 Problem Statement

We consider the discrete-time system with the dynamics described by the difference equation

$$\mathbf{x}_{t+1}(\omega) = \mathbf{A}_t \mathbf{x}_t(\omega) + \mathbf{B}_t \mathbf{u}_t(\omega) + \mathbf{w}_t(\omega) \tag{1}$$

and the measurement data generated according to

$$\mathbf{y}_t(\omega) = \mathbf{C}_t \mathbf{x}_t(\omega) + \mathbf{v}_t(\omega), \tag{2}$$

where t is the discrete time integer, t=0, 1, ... ,T-1; ω is an element in the probability space (Ω, D, P) consisting of the sure event Ω, the σ-field D of events, and the probability P on D; $\mathbf{x}_t(\omega)$ is a k-dimensional state vector, the initial state $\mathbf{x}_0(\omega)$ being a random variable; $\mathbf{u}_t(\omega)$ is an m-dimensional control vector; $\mathbf{y}_t(\omega)$ is an n-dimensional measurement vector; $\mathbf{w}_t(\omega)$ is a k-dimensional random vector representing system noise, $\mathbf{v}_t(\omega)$ is an n-dimensional random vector representing measurement noise; $\mathbf{A}_t$ is a known k $\times$ k state transition matrix; $\mathbf{B}_t$ is a known k $\times$ m control driven matrix, and $\mathbf{C}_t$ is a known n $\times$ k measurement matrix. For simplicity, the notation of the dependence of the variables on ω will be suppressed in the sequel. If d is a family of random variables, $\sigma\{d\}\subset D$ denotes the smallest σ-field of ω-sets relative to which d is measurable. It is assumed that the prior probability distributions of all random variables are known, and that each of these has a finite covariance matrix.

The performance criterion for the system of equation (1) is chosen to be of quadratic form:

$$J = E\left\{\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{G}_t\mathbf{x}_t + \mathbf{u}_{t-1}'\mathbf{H}_{t-1}\mathbf{u}_{t-1})\right\}, \tag{3}$$

where $\mathbf{G}_t$, t=1,2, ... ,T, is a symmetric non-negative definite matrix; $\mathbf{H}_t$, t=0,1, ... ,T-1, is a symmetric positive definite matrix, the symbol $E\{\cdot\}$ denotes the expectation and prime ($'$) denotes transpose.

The control sequence $\{\mathbf{u}_t, t=0,1, \ldots ,T-1\}$ is to be chosen as a function of the information available, in such a way as to minimize the value of the performance criterion (3). The information available at time t is the set consisting of all the past and present measurement data $\{\mathbf{y}_0, \mathbf{y}_1, \ldots ,\mathbf{y}_t\}$, all the past controls $\{\mathbf{u}_0, \mathbf{u}_1, \ldots ,\mathbf{u}_{t-1}\}$, and all the prior information. Let us denote the control sequence by $\mathbf{U}_t=\{\mathbf{u}_0, \mathbf{u}_1, \ldots ,\mathbf{u}_t\}$ and the sequence of the measurement data when the past control sequence $\mathbf{U}_{t-1}$ has been applied to the system by $\mathbf{Y}_t(\mathbf{U}_{t-1})=\{\mathbf{y}_0,\mathbf{y}_1(\mathbf{U}_0), \ldots ,\mathbf{y}_t(\mathbf{U}_{t-1})\}$. Then the control is of the form $\mathbf{u}_t=\phi(\mathbf{Y}_t(\mathbf{U}_{t-1}), t)$. Here, $\phi(\cdot\ , t)$ is a measurable mapping function from $\sigma\{\mathbf{Y}_t(\mathbf{U}_{t-1})\}$ to m-dimensional Euclidian space. In the following, we shall denote $\sigma\{\mathbf{Y}_t(\mathbf{U}_{t-1})\}$ by D(t).

3 Determining Control Laws

Theorem 1 (***Optimal Control Law***). For the above stochastic optimal control problem, the optimal control law is given by

$$\mathbf{u}_t^* = -[\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t + \mathbf{H}_t]^{-1}\mathbf{B}_t'[\mathbf{R}_{t+1}\mathbf{A}_t E\{\mathbf{x}_t;D(t)\} + E\{\mathbf{z}_t;D(t)\}], \tag{4}$$

where $\mathbf{R}_t$ satisfies the Riccati equation

$$\mathbf{R}_t = \mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{A}_t + \mathbf{G}_t - \mathbf{Q}_t, \tag{5}$$

$$\mathbf{Q}_t = \mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{B}_t\left[\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t + \mathbf{H}_t\right]^{-1}\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{A}_t \tag{6}$$

with

$$\mathbf{R}_T = \mathbf{G}_T, \tag{7}$$

and $\mathbf{z}_t$ is a random variable defined by recurrence relation

$$\mathbf{z}_t = \mathbf{R}_{t+1}\mathbf{w}_t + \mathbf{M}_{t+1}\mathbf{z}_{t+1} \tag{8}$$

with

$$\mathbf{z}_{T-1} = \mathbf{R}_T\mathbf{w}_{T-1}, \tag{9}$$

where

$$\mathbf{M}_t = \mathbf{A}_t' - \mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{B}_t\left[\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t + \mathbf{H}_t\right]^{-1}\mathbf{B}_t'. \tag{10}$$

Proof. We will use the dynamic programming technique for the proof. Let us introduce the optimal cost-to-go from time t to the final time

$$\mathrm{J}^*(\mathrm{D(t)},\mathrm{t}) = \min_{\mathbf{u}_t \cdots \mathbf{u}_{T-1}} \mathrm{E}\left\{\sum_{i=t+1}^{T}\left[\mathbf{x}_i'\mathbf{G}_i\mathbf{x}_i + \mathbf{u}_{i-1}'\mathbf{H}_{i-1}\mathbf{u}_{i-1}\right];\mathrm{D(t)}\right\}. \tag{11}$$

By the principle of optimality, $\mathrm{J}^*(\mathrm{D(t)},\mathrm{t})$ satisfies the following functional equation:

$$\mathrm{J}^*(\mathrm{D(t)},\mathrm{t}) = \min_{\mathbf{u}_t} \mathrm{E}\left\{[\mathbf{x}_{t+1}'\mathbf{G}_{t+1}\mathbf{x}_{t+1} + \mathbf{u}_t'\mathbf{H}_t\mathbf{u}_t + \mathrm{J}^*(\mathrm{D(t+1)},\mathrm{t+1})];\mathrm{D(t)}\right\} \tag{12}$$

with the initial condition

$$\mathrm{J}^*(\mathrm{D(T)},\mathrm{T}) = 0. \tag{13}$$

Thus the optimal control for the final time, $\mathbf{u}_{T-1}^*$, minimizes

$$\mathrm{J}(\mathrm{D(T-1)},\mathrm{T-1}) = \mathrm{E}\left\{\left[\mathbf{x}_T'\mathbf{G}_T\mathbf{x}_T + \mathbf{u}_{T-1}'\mathbf{H}_{T-1}\mathbf{u}_{T-1}\right];\mathrm{D(T-1)}\right\} \tag{14}$$

subject to equation (1). The optimal control $\mathbf{u}_{T-1}^*$ is found to be

$$\mathbf{u}_{T-1}^* = -\left[\mathbf{B}_{T-1}'\mathbf{R}_T\mathbf{B}_{T-1} + \mathbf{H}_{T-1}\right]^{-1}\mathbf{B}_{T-1}'\left[\mathbf{R}_T\mathbf{A}_{T-1}\mathrm{E}\{\mathbf{x}_{T-1};\mathrm{D(T-1)}\} + \mathrm{E}\{\mathbf{z}_{T-1};\mathrm{D(T-1)}\}\right], \tag{15}$$

where $\mathbf{R}_T$ and $\mathbf{z}_{T-1}$ are given by equations (5)-(7) and (8)-(9), respectively. It is found without difficulty that the minimum cost can be written as

$$J^*(D(T-1),T-1) = \min_{\mathbf{u}_{T-1}} J(D(T-1),T-1)$$

$$= E\{[\mathbf{x}'_{T-1}(\mathbf{A}'_{T-1}\mathbf{R}_T\mathbf{A}_{T-1} - \mathbf{Q}_{T-1})\mathbf{x}_{T-1} + 2\mathbf{x}'_{T-1}\mathbf{M}_{T-1}\mathbf{z}_{T-1}$$

$$+ \operatorname{trace}(\mathbf{R}_T\mathbf{w}_{T-1}\mathbf{w}'_{T-1} - B_{T-1}[\mathbf{B}'_{T-1}\mathbf{R}_T\mathbf{B}_{T-1} + \mathbf{H}_{T-1}]^{-1} \tag{16}$$

$$\bullet \mathbf{B}'_{T-1}[\mathbf{z}_{T-1}\mathbf{z}'_{T-1} + (\mathbf{R}_T\mathbf{A}_{T-1}\tilde{\mathbf{x}}_{T-1} + \tilde{\mathbf{z}}_{T-1})(\mathbf{R}_T\mathbf{A}_{T-1}\tilde{\mathbf{x}}_{T-1} + \tilde{\mathbf{z}}_{T-1})'])]; D(T-1)\},$$

where $\mathbf{Q}_{T-1}$ and $\mathbf{M}_{T-1}$ are given by equations (6) and (10), respectively; $\tilde{\mathbf{x}}_{T-1}$ and $\tilde{\mathbf{z}}_{T-1}$ are defined respectively by

$$\tilde{\mathbf{x}}_t = \mathbf{x}_t - E\{\mathbf{x}_t; D(t)\} \tag{17}$$

and

$$\tilde{\mathbf{z}}_t = \mathbf{z}_t - E\{\mathbf{z}_t; D(t)\}. \tag{18}$$

Now we will show that $\tilde{\mathbf{x}}_{T-1}$ and $\tilde{\mathbf{z}}_{T-1}$ are not affected by the control $\mathbf{U}_{T-2}$. The procedure follows Wonham (1968). Let us break $\mathbf{x}_t$ into two vectors

$$\mathbf{x}_t = \mathbf{x}_t^{\bullet} + \mathbf{x}_t^{\circ}, \tag{19}$$

where $\mathbf{x}_t^{\bullet}$ and $\mathbf{x}_t^{\circ}$ are defined respectively by

$$\mathbf{x}_{t+1}^{\bullet} = \mathbf{A}_t\mathbf{x}_t^{\bullet} + \mathbf{w}_t, \quad \mathbf{x}_0^{\bullet} = \mathbf{x}_0 \tag{20}$$

and

$$\mathbf{x}_{t+1}^{\circ} = \mathbf{A}_t\mathbf{x}_t^{\circ} + \mathbf{B}_t\mathbf{u}_t, \quad \mathbf{x}_0^{\circ} = \mathbf{0}. \tag{21}$$

Since $\mathbf{x}_t^{\circ}$ is D(t)-measurable, it follows

$$E\{\mathbf{x}_t; D(t)\} = E\{\mathbf{x}_t^{\bullet}; D(t)\} + \mathbf{x}_t^{\circ}. \tag{22}$$

Let us define

$$\mathbf{y}_t^{\bullet} = \mathbf{y}_t - \mathbf{C}_t\mathbf{x}_t^{\circ} = \mathbf{C}_t\mathbf{x}_t^{\bullet} + \mathbf{v}_t. \tag{23}$$

By the result in Wonham (1968), we have

$$D(t) = D^{\bullet}(t) \qquad (24)$$

where

$$D^{\bullet}(t) = \sigma\{\mathbf{y}_0^{\bullet}, \mathbf{y}_1^{\bullet}, \ldots, \mathbf{y}_t^{\bullet}\} \qquad (25)$$

In view of equations (19), (20) and (22) to (24), we have

$$\mathbf{x}_t - E\{\mathbf{x}_t; D(t)\} = \mathbf{x}_t^{\bullet} - E\{\mathbf{x}_t^{\bullet}; D^{\bullet}(t)\}, \qquad (26)$$

$$\mathbf{w}_i - E\{\mathbf{w}_i; D(t)\} = \mathbf{w}_i - E\{\mathbf{w}_i; D^{\bullet}(t)\}, \quad \text{for } i \geq t. \qquad (27)$$

It follows from equations (26) and (27) that the minimum variance estimation error of the state, $\mathbf{x}_t$-E{$\mathbf{x}_t$;D(t)}, and the minimum variance prediction error of the system noise, $\mathbf{w}_i$-E{$\mathbf{w}_i$;D(t)}, for $i \geq t$, are independent of the control sequence $\mathbf{U}_{t-1}$. Thus, $\widetilde{\mathbf{x}}_{T-1}$ and $\widetilde{\mathbf{z}}_{T-1}$ are not affected by the control $\mathbf{U}_{T-2}$. Therefore, the third term in the right-hand side of equation (16), trace (...), can be ignored in determining the control $\mathbf{U}_{T-2}$. When this fact is found, the remainder of the proof becomes obvious and hence will be omitted. □

Corollary 1.1. The optimal control laws for the stochastic control problem of equations (1)-(3) can be obtained by considering the optimal control laws for the related deterministic systems where the random variables are replaced by their expected values, i.e., the above problem is certainty equivalent.

Proof. When $\mathbf{w}_t$ is a known deterministic variable and the full information concerning the state $\mathbf{x}_t$ is assumed available to the decision maker, it is easy to find that the optimal control law of the deterministic equivalent problem can be written as

$$\mathbf{u}_t^* = -[\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t + \mathbf{H}_t]^{-1}\mathbf{B}_t'[\mathbf{R}_{t+1}\mathbf{A}_t\mathbf{x}_t + \mathbf{z}_t], \qquad (28)$$

where $\mathbf{R}_t$ and $\mathbf{z}_t$ are given by equations (5)-(7) and (8)-(10), respectively. Comparing equation (4) with (28), we see that the stochastic problem is certainty equivalent. □

Corollary 1 shows that the statistical characteristics of the initial state, the system noise and the measurement noise have not any influence upon the basic form of the optimal control law.

Theorem 2 (***Adaptive Cautious Control Law***). Under conditions of Theorem 1, when the k × k state transition matrix **A** (**A**=$\mathbf{A}_t$, ∀t=0(1)T-1) and the k × m control driven matrix **B** (**B**=$\mathbf{B}_t$, ∀t=0(1)T-1) are unknown, the adaptive cautious control law for the above control problem is given by

$$\mathbf{u}_t^* = -\left[E\{\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t;D(\tau)\} + \mathbf{H}_t\right]^{-1} E\{\mathbf{B}_t'[\mathbf{R}_{t+1}\mathbf{A}_t\mathbf{x}_t + \mathbf{z}_t];D(\tau)\}, \quad \text{for } t \geq \tau, \tag{29}$$

where $\mathbf{R}_t$ satisfies the Riccati equation

$$\mathbf{R}_t = E\{\mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{A}_t\} + \mathbf{G}_t - \mathbf{Q}_t, \tag{30}$$

$$\mathbf{Q}_t = E\left\{\mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{B}_t\left[E\{\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t\} + \mathbf{H}_t\right]^{-1}\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{A}_t\right\} \tag{31}$$

with

$$\mathbf{R}_T = \mathbf{G}_T, \tag{32}$$

and $\mathbf{z}_t$ is a random variable defined by recurrence relation

$$\mathbf{z}_t = \mathbf{R}_{t+1}\mathbf{w}_t + E\{\mathbf{M}_{t+1}\mathbf{z}_{t+1}\} \tag{33}$$

with

$$\mathbf{z}_{T-1} = \mathbf{R}_T\mathbf{w}_{T-1}, \tag{34}$$

where

$$\mathbf{M}_t = \mathbf{A}_t' - \mathbf{A}_t'\mathbf{R}_{t+1}\mathbf{B}_t\left[E\{\mathbf{B}_t'\mathbf{R}_{t+1}\mathbf{B}_t\} + \mathbf{H}_t\right]^{-1}\mathbf{B}_t'. \tag{35}$$

Proof. The proof is similar to that of Theorem 1 and so it is omitted here. □

Here, the following corollaries clearly hold.

Corollary 2.1. Under conditions of Theorem 2, the certainty equivalence principle does not yield the admissible adaptive control laws for the above problem.

Corollary 2.2. For the final value control problem (t=T), the adaptive cautious control law is optimal.

4 Examples

Example 4.1. Consider the scalar linear stochastic system with state equation

$$x_{t+1} = x_t + bu_t + w_t, \tag{36}$$

where b is an unknown constant, $\{w_t\}$ is a sequence of independent Gaussian random variables with zero mean and variance σ^2. Let the observation equation be described by

$$y_t = x_t, \tag{37}$$

where y_t is the observation.

The control sequence $\{u_t, t=0,1, \dots ,T\text{-}1\}$ is to be chosen as a function of the information available, in such a way as to minimize the value of the performance criterion

$$J(U_{T\text{-}1}) = E\left\{\sum_{t=1}^{T} x_t^2\right\}, \tag{38}$$

where $U_{T\text{-}1}=\{u_t, t=0,1, \dots ,T\text{-}1\}$. The information available at time t is the set consisting of all the past and present measurement data $\{\mathbf{y}_0, \mathbf{y}_1, \dots ,\mathbf{y}_t\}$, all the past controls $\{\mathbf{u}_0, \mathbf{u}_1, \dots ,\mathbf{u}_{t\text{-}1}\}$, and all the prior information.

The unknown parameter b in (36) can be estimated, using the invariant embedding technique (see Nechval, 1982, 1984, 1988), as

$$\hat{b}_t = \frac{\sum_{j=1}^{t} u_{j-1}(x_j - x_{j-1})}{\sum_{j=1}^{t} u_{j-1}^2}, \quad \forall t = 1(1)T\text{-}1. \tag{39}$$

The invariant embedding technique allows one easily to find that the statistic $\hat{b}_t, \forall t\text{=}1(1)T\text{-}1$, follows the normal distribution with a probability density function (pdf)

$$f(\hat{b}_t) = \frac{1}{(2\pi\hat{\sigma}_{\hat{b}_t}^2)^{1/2}} \exp\left(-\frac{(\hat{b}_t - b)^2}{2\hat{\sigma}_{\hat{b}_t}^2}\right), \quad \hat{b}_t \in (-\infty, \infty), \tag{40}$$

where

$$\hat{\sigma}_{\hat{b}_t}^2 = \sigma^2\left[\sum_{j=1}^{t} u_{j-1}^2\right]^{-1}. \tag{41}$$

In terms of fiducial approach (Fisher, 1948; Fraser, 1961; Nechval, 1982, 1984), it follows from (40) that a fiducial pdf of b is given by

$$f(b) = \frac{1}{(2\pi\hat{\sigma}_{\hat{b}_t}^2)^{1/2}} \exp\left(-\frac{(b - \hat{b}_t)^2}{2\hat{\sigma}_{\hat{b}_t}^2}\right), \quad b \in (-\infty, \infty), \tag{42}$$

If b were known then the optimal expected loss is given as

$$J(u_t^\circ, \dots, u_{T-1}^\circ) = (T-t)\sigma^2. \tag{43}$$

The optimal control law is (see Theorem 1)

$$u_t^\circ = -x_t / b, \quad \forall t = 0(1)T-1. \tag{44}$$

If the estimated values, $\hat{b}_t$, are used in (44) instead of the true value we get

$$\hat{u}_t^\circ = -x_t / \hat{b}_t, \quad \forall t = 1(1)T-1. \tag{45}$$

i.e., we have assumed that the certainty equivalence principle can be used. In terms of a fiducial approach, the expected loss when using (45) will be

$$J(\hat{u}_t^\circ, \dots, \hat{u}_{T-1}^\circ) = x_t^2 \sum_{j=1}^{T-t} \left(\frac{\hat{\sigma}_{\hat{b}_t}^2}{\hat{b}_t^2} \right)^j + \sigma^2 \left[(T-t-1) + \left(1 + \frac{\hat{\sigma}_{\hat{b}_t}^2}{\hat{b}_t^2} \right)^{T-t-1} \right], \tag{46}$$

where $t \in \{1, \dots, T-1\}$, $T \geq 2$. If the adaptive cautious control law, given by Theorem 2,

$$\hat{u}_t^* = -x_t \, \hat{b}_t \left(\hat{b}_t^2 + \hat{\sigma}_{\hat{b}_t}^2 \right)^{-1}, \quad \forall t = 1(1)T-1, \tag{47}$$

is used then the expected loss (in terms of a fiducial approach) will be

$$J(\hat{u}_t^*, \dots, \hat{u}_{T-1}^*) = x_t^2 \sum_{j=1}^{T-t} \left(\frac{\hat{\sigma}_{\hat{b}_t}^2}{\hat{b}_t^2 + \hat{\sigma}_{\hat{b}_t}^2} \right)^j + \sigma^2 \left[(T-t-1) + \left(1 + \frac{\hat{\sigma}_{\hat{b}_t}^2}{\hat{b}_t^2 + \hat{\sigma}_{\hat{b}_t}^2} \right)^{T-t-1} \right]. \tag{48}$$

The loss in equation (48) is less than in (46) since $\hat{\sigma}_{\hat{b}_t}^2 \geq 0$, $\forall t = 1(1)T-1$. The adaptive control law (47) is cautious since it considers the inaccuracy of the estimates of b.

Example 4.2. Consider the scalar linear stochastic system with state equation

$$x_{t+1} = ax_t + bu_t + w_t, \tag{49}$$

where a and b are unknown constants, $\{w_t\}$ is a sequence of independent Gaussian random variables with zero mean and variance σ^2. Let the observation equation be described by

$$y_t = x_t, \tag{50}$$

where y_t is the observation. We want to select the control u_{T-1}, based on all available observed data $Y_{T-1}=\{y_0, y_1, \ldots, y_{T-1}\}$ and $U_{T-2}=\{u_0, u_1, \ldots, u_{T-2}\}$, in order to minimize the loss function (final value control problem)

$$J(u_{T-1}) = E\{x_T^2\}. \tag{51}$$

The unknown parameters a and b in (49) can be estimated, using the invariant embedding technique (see Nechval, 1982, 1984, 1988 and Appendix), as

$$\hat{a}_{T-1} = \frac{\left(\sum_{t=1}^{T-1} x_t x_{t-1}\right)\left(\sum_{t=1}^{T-1} u_{t-1}^2\right) - \left(\sum_{t=1}^{T-1} x_t u_{t-1}\right)\left(\sum_{t=1}^{T-1} x_{t-1} u_{t-1}\right)}{\left(\sum_{t=1}^{T-1} x_{t-1}^2\right)\left(\sum_{t=1}^{T-1} u_{t-1}^2\right) - \left(\sum_{t=1}^{T-1} x_{t-1} u_{t-1}\right)^2} \tag{52}$$

and

$$\hat{b}_{T-1} = \frac{\left(\sum_{t=1}^{T-1} x_{t-1}^2\right)\left(\sum_{t=1}^{T-1} x_t u_{t-1}\right) - \left(\sum_{t=1}^{T-1} x_{t-1} u_{t-1}\right)\left(\sum_{t=1}^{T-1} x_t x_{t-1}\right)}{\left(\sum_{t=1}^{T-1} x_{t-1}^2\right)\left(\sum_{t=1}^{T-1} u_{t-1}^2\right) - \left(\sum_{t=1}^{T-1} x_{t-1} u_{t-1}\right)^2}, \tag{53}$$

respectively. The invariant embedding technique allows one easily to find that the statistic $(\hat{a}_{T-1}, \hat{b}_{T-1})$ follows the bivariate normal distribution with a probability density function (pdf)

$$f(\hat{a}, \hat{b}) = \frac{1}{2\pi\sigma_{\hat{a}}\sigma_{\hat{b}}\sqrt{1-\rho_{\hat{a}\hat{b}}^2}} \exp\left(-\frac{1}{2(1-\rho_{\hat{a}\hat{b}}^2)}\left[\frac{(\hat{a}-a)^2}{\sigma_{\hat{a}}^2} + \frac{(\hat{b}-b)^2}{\sigma_{\hat{b}}^2} - 2\rho_{\hat{a}\hat{b}}\frac{(\hat{a}-a)(\hat{b}-b)}{\sigma_{\hat{a}}\sigma_{\hat{b}}}\right]\right),$$

$$\hat{a}, \hat{b} \in (-\infty, \infty), \tag{54}$$

where

$$\hat{\sigma}_{\hat{a}}^2 = \frac{\sum_{t=1}^{T-1} u_{t-1}^2}{\left(\sum_{t=1}^{T-1} x_{t-1}^2\right)\left(\sum_{t=1}^{T-1} u_{t-1}^2\right) - \left(\sum_{t=1}^{T-1} x_{t-1}u_{t-1}\right)^2}\sigma^2, \tag{55}$$

$$\hat{\sigma}_{\hat{b}}^2 = \frac{\sum_{t=1}^{T-1} x_{t-1}^2}{\left(\sum_{t=1}^{T-1} x_{t-1}^2\right)\left(\sum_{t=1}^{T-1} u_{t-1}^2\right) - \left(\sum_{t=1}^{T-1} x_{t-1}u_{t-1}\right)^2}\sigma^2, \tag{56}$$

and

$$\hat{\rho}_{\hat{a}\hat{b}} = -\frac{\sum_{t=1}^{T-1} x_{t-1}u_{t-1}}{\sqrt{\sum_{t=1}^{T-1} x_{t-1}^2}\sqrt{\sum_{t=1}^{T-1} u_{t-1}^2}}. \tag{57}$$

In terms of fiducial approach (Fisher, 1948; Fraser, 1961; Nechval, 1982, 1984), it follows from (54) that a fiducial pdf of (a,b) is given by

$$f^\bullet(a,b) = \frac{1}{2\pi\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}}\sqrt{1-\hat{\rho}_{\hat{a}\hat{b}}^2}}\exp\left(-\frac{1}{2(1-\hat{\rho}_{\hat{a}\hat{b}}^2)}\left[\frac{(a-\hat{a})^2}{\hat{\sigma}_{\hat{a}}^2} + \frac{(b-\hat{b})^2}{\hat{\sigma}_{\hat{b}}^2} - 2\hat{\rho}_{\hat{a}\hat{b}}\frac{(a-\hat{a})(b-\hat{b})}{\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}}}\right]\right),$$

$$a, b \in (-\infty, \infty). \tag{58}$$

If a and b were known then the optimal loss is given as (see Theorem 1)

$$\min_{u_{T-1}} J(u_{T-1}) = \min_{u_{T-1}} E\left\{(ax_{T-1} + bu_{T-1} + w_{T-1})^2\right\} = \min_{u_{T-1}} \left\{(ax_{T-1} + bu_{T-i})^2 + \sigma^2\right\} = \sigma^2. \tag{59}$$

To obtain this we have used that w_T is independent of a, b, $x_0, \ldots, x_{T-1}$, $u_0, \ldots, u_{T-2}$. The optimal control law is

$$u_{T-1}^\circ = -a\,x_{T-1}/b. \tag{60}$$

If the estimated values, $\hat{a}$ and $\hat{b}$, are used in (60) instead of the true values we get

$$\hat{u}^{\circ}_{T-1} = -\,\hat{a}x_{T-1}/\hat{b}, \tag{61}$$

i.e., we have assumed that the certainty equivalence principle can be used. The loss when using (61) will be

$$J(\hat{u}^{\circ}_{T-1}) = E\left\{\left(ax_{T-1} - b\frac{\hat{a}}{\hat{b}}x_{T-1} + w_{T-1}\right)^2\right\} = \frac{\hat{a}^2\hat{\sigma}^2_{\hat{b}} - 2\hat{\rho}_{\hat{a}\hat{b}}\hat{a}\hat{b}\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}} + \hat{b}^2\hat{\sigma}^2_{\hat{a}}}{\hat{b}^2}x^2_{T-1} + \sigma^2. \tag{62}$$

To get the last equality the formulae

$$E\{a^2\} = \hat{a}^2 + \hat{\sigma}^2_{\hat{a}}, \tag{63}$$

$$E\{b^2\} = \hat{b}^2 + \hat{\sigma}^2_{\hat{b}}, \tag{64}$$

and

$$E\{ab\} = \hat{a}\hat{b} + \hat{\rho}_{\hat{a}\hat{b}}\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}} \tag{65}$$

have been used. The loss has increased with the term

$$\frac{\hat{a}^2\hat{\sigma}^2_{\hat{b}} - 2\hat{\rho}_{\hat{a}\hat{b}}\hat{a}\hat{b}\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}} + \hat{b}^2\hat{\sigma}^2_{\hat{a}}}{\hat{b}^2}x^2_{T-1} \tag{66}$$

compared with the optimal loss when a and b were known. The control law (61) does not minimize (51) (see Corollary 2.1) because

$$\min_{u_{T-1}} J(u_{T-1}) = \min_{u_{T-1}} E\left\{(ax_{T-1} + bu_{T-1} + w_{T-1})^2\right\} = \frac{\hat{a}^2\hat{\sigma}^2_{\hat{b}} - 2\hat{\rho}_{\hat{a}\hat{b}}\hat{a}\hat{b}\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}} + \hat{b}^2\hat{\sigma}^2_{\hat{a}}}{\hat{b}^2 + \hat{\sigma}^2_{\hat{b}}}x^2_{T-1} + \sigma^2 \tag{67}$$

and the minimum is assumed for the control law (see Theorem 2)

$$\hat{u}^{*}_{T-1} = -\frac{\hat{a}\hat{b} + \hat{\rho}_{\hat{a}\hat{b}}\hat{\sigma}_{\hat{a}}\hat{\sigma}_{\hat{b}}}{\hat{b}^2 + \hat{\sigma}^2_{\hat{b}}}x_{T-1}. \tag{68}$$

The loss in eqn. (67) is less than in (62) since $\hat{\sigma}^2_{\hat{b}} \geq 0$. The first term in (67) is the loss due to the uncertainty of the parameters and the second term is due to the process noise w_{T-1}.

The controller (68) is cautious since it considers the inaccuracy of the estimates of a and b. It follows from (67) (see also Corollary 2.2) that this controller is optimal. If $\hat{\sigma}_{\hat{a}} \to 0$ and $\hat{\sigma}_{\hat{b}} \to 0$ then (61) and (68) will be the same and the loss approaches the optimal loss for known a and b, (59).

It will be noted that the above results can be also obtained by using a variational approach (see Nechval et al., 1998, 1999), but the technique proposed here simplifies the problem of adaptive optimization in stochastic systems.

5 Conclusion

The authors hope that this work will stimulate further investigation using the approach on specific applications to see whether obtained results with it are feasible for realistic applications and may be extended to provide existence results for nonlinear stochastic control problems.

Appendix: Invariant Embedding Technique

Here we present an invariant embedding technique based on the constructive use of invariance principle in mathematical statistics. This technique is a special case of more general considerations applicable whenever the statistical problem is invariant under a group of transformations, which acts transitively on the parameter space. It allows one to solve many problems of the theory of statistical inferences in a simple way.

Preliminaries

Our underlying structure consists of a class of probability models ($\mathcal{Y}$, $\mathcal{A}$, $\mathcal{P}$), a one-one mapping ψ taking $\mathcal{P}$ onto an index set Θ, a measurable space of actions ($\mathcal{D}$, $\mathcal{B}$), and a real-valued loss function r defined on $\Theta \times \mathcal{D}$. We assume that a group G of one-one $\mathcal{A}$ - measurable transformations acts on $\mathcal{Y}$ and that it leaves the class of models ($\mathcal{Y}$, $\mathcal{A}$, $\mathcal{P}$) invariant. We further assume that homomorphic images $\bar{G}$ and $\tilde{G}$ of G act on Θ and $\mathcal{D}$, respectively. ($\bar{G}$ may be induced on Θ through ψ; $\tilde{G}$ may be induced on $\mathcal{D}$ through r). We shall say that r is invariant if for every $(\theta,d) \in \Theta \times \mathcal{D}$

$$r(\bar{g}\theta, \tilde{g}d) = r(\theta,d), \quad g \in G. \tag{1}$$

Given the structure described above there are aesthetic and sometimes admissibility grounds for restricting attention to estimators (decision rules) φ: $\mathcal{X} \to \mathcal{D}$ which are $(G,\tilde{G})$ equivariant in the sense that

$$\varphi(gy) = \tilde{g}\varphi(y), \ \ y \in \mathscr{Y}, \ \ g \in G. \tag{2}$$

If $\overline{G}$ is trivial and (1), (2) hold, we say φ is G-invariant, or simply invariant.

Invariant Embedding

We begin by noting that r is invariant in the sense of (1) if and only if r is a $G^{\bullet}$-invariant function, where $G^{\bullet}$ is defined on $\Theta \times \mathscr{D}$ as follows: to each $g \in G$, with homomorphic images $\overline{g}$, $\tilde{g}$ in $\overline{G}$, $\tilde{G}$ respectively, let $g^{\bullet}(\theta,d)=(\overline{g}\theta, \tilde{g}d)$, $(\theta,d)\in(\Theta \times \mathscr{D})$. It is assumed that $\tilde{G}$ is a homomorphic image of $\overline{G}$.

Definition 1 (*Transitivity*). A transformation group $\overline{G}$ acting on a set Θ is called (uniquely) transitive if for every $\theta, \vartheta \in \Theta$ there exists a (unique) $\overline{g} \in \overline{G}$ such that $\overline{g}\theta=\vartheta$. When $\overline{G}$ is transitive on Θ we may index $\overline{G}$ by Θ: fix an arbitrary point $\theta \in \Theta$ and define $\overline{g}_{\theta_1}$ to be the unique $\overline{g} \in \overline{G}$ satisfying $\overline{g}\theta=\theta_1$. The identity of $\overline{G}$ clearly corresponds to θ. An immediate consequence is Lemma 1.

Lemma 1 (*Transformation*). Let $\overline{G}$ be transitive on Θ. Fix $\theta \in \Theta$ and define $\overline{g}_{\theta_1}$ as above. Then $\overline{g}_{\overline{q}\theta_1} = \overline{q}\,\overline{g}_{\theta_1}$ for $\theta \in \Theta$, $\overline{q} \in \overline{G}$.

Proof. The identity $\overline{g}_{\overline{q}\theta_1}\theta = \overline{q}\theta_1 = \overline{q}\overline{g}_{\theta_1}\theta$ shows that $\overline{g}_{\overline{q}\theta_1}$ and $\overline{q}\,\overline{g}_{\theta_1}$ both take θ into $\overline{q}\theta_1$, and the lemma follows by unique transitivity. □

Theorem 1 (*Maximal Invariant*). Let $\overline{G}$ be transitive on Θ. Fix a reference point $\theta_0 \in \Theta$ and index $\overline{G}$ by Θ. A maximal invariant M with respect to $G^{\bullet}$ acting on $\Theta \times \mathscr{D}$ is defined by

$$M(\theta,d) = \tilde{g}_{\theta}^{-1}d, \ \ (\theta,d) \in \Theta \times \mathscr{D}. \tag{3}$$

Proof. For each $(\theta,d)\in(\Theta \times \mathscr{D})$ and $\overline{g} \in \overline{G}$

$$M(\overline{g}\theta, \tilde{g}d) = (\tilde{g}_{\overline{g}\theta}^{-1})\tilde{g}d = (\tilde{g}\tilde{g}_{\theta})^{-1}\tilde{g}d = \tilde{g}_{\theta}^{-1}\tilde{g}^{-1}\tilde{g}d = \tilde{g}_{\theta}^{-1}d = M(\theta,d) \tag{4}$$

by Lemma 1 and the structure preserving properties of homomorphisms. Thus M is $G^{\bullet}$-invariant. To see that M is maximal, let $M(\theta_1,d_1)=M(\theta_2,d_2)$. Then $\tilde{g}_{\theta_1}^{-1}d_1 = \tilde{g}_{\theta_2}^{-1}d_2$ or $d_1=\tilde{g}\,d_2$ where $\tilde{g} = \tilde{g}_{\theta_1}\tilde{g}_{\theta_2}^{-1}$. Since $\theta_1=\overline{g}_{\theta_1}\theta_0 = \overline{g}_{\theta_1}\overline{g}_{\theta_2}^{-1}\theta_2 = \overline{g}\theta_2$, $(\theta_1,d_1)=g^{\bullet}(\theta_2,d_2)$ for some $g^{\bullet} \in G^{\bullet}$, and the proof is complete. □

Corollary 1.1 (*Invariant Embedding*). An invariant loss function, $r(\theta,d)$, can be transformed as follows:

$$r(\theta,d) = r(\bar{g}_{\hat{\theta}}^{-1}\theta, \tilde{g}_{\hat{\theta}}^{-1}d) = \ddot{r}(v,\eta), \tag{5}$$

where $v=v(\theta,\hat{\theta})$ is a function (it is called a pivotal quantity) such that the distribution of v does not depend on θ; $\eta=\eta(d,\hat{\theta})$ is an ancillary factor; $\hat{\theta}$ is the maximum likelihood estimator of θ (or the sufficient statistic for θ).

Corollary 1.2 (*Best Invariant Estimator*). The best invariant estimator (decision rule) is given by

$$\varphi^*(y) = d^* = \eta^{-1}(\eta^*,\hat{\theta}), \tag{6}$$

where

$$\eta^* = \arg\inf_{\eta} E_{\eta}\{\ddot{r}(v,\eta)\}. \tag{7}$$

Corollary 1.3 (*Risk*). A risk function

$$R(\theta,\varphi(y)) = E_{\theta}\{r(\theta,\varphi(y))\} = E_{\eta_\circ}\{\ddot{r}(v_\circ,\eta_\circ)\} \tag{8}$$

is constant on orbits when an invariant estimator (decision rule) $\varphi(y)$ is used, where $v_\circ = v_\circ(\theta,y)$ is a function (pivotal quantity) whose distribution does not depend on θ; $\eta_\circ = \eta_\circ(d,y)$ is an ancillary factor.

Consider, for instance, the problem of estimating the location-scale parameter of a distribution belonging to a family generated by a continuous cdf F: $\mathscr{P}=\{P_\theta: F((y-\mu)/\sigma), y\in R, \theta\in\Theta\}$, $\Theta=\{(\mu,\sigma): \mu,\sigma\in R, \sigma>0\}=\mathscr{D}$. The group G of location and scale changes leaves the class of models invariant. Since $\overline{G}$ induced on Θ by $P_\theta \to \theta$ is uniquely transitive, we may apply Theorem 1 and obtain invariant loss functions of the form

$$r(\theta,\varphi(y)) = r[(\varphi_1(y)-\mu)/\sigma, \varphi_2(y)/\sigma] \tag{9}$$

if $\theta=(\mu,\sigma)$ and $\varphi(y)=(\varphi_1(y), \varphi_2(y))$. Let $\hat{\theta}=(\hat{\theta}_1,\hat{\theta}_2)$, $d=(d_1,d_2)$, then

$$r(\theta,d) = r[(d_1-\mu)/\sigma, d_2/\sigma] = r(v_1+\eta_1 v_2, \eta_2 v_2) = \ddot{r}(v,\eta), \tag{10}$$

where $v=(v_1,v_2)$, $v_1=(\hat{\theta}_1-\mu)/\sigma$, $v_2=\hat{\theta}_2/\sigma$; $\eta=(\eta_1,\eta_2)$, $\eta_1=(d_1-\hat{\theta}_1)/\hat{\theta}_2$, $\eta_2=d_2/\hat{\theta}_2$.

Characterization of Estimators

For any statistical decision problem, an estimator (a decision rule) d_1 is said to be equivalent an estimator (a decision rule) d_2 if $R(\theta,d_1)=R(\theta,d_2)$ for all $\theta\in\Theta$, where $R(\bullet)$ is a risk function, Θ is a parameter space,. An estimator d_1 is said to be uniformly better than an estimator d_2 if $R(\theta,d_1)<R(\theta,d_2)$ for all $\theta\in\Theta$. An estimator d_1 is said to be as good as an estimator d_2 if $R(\theta,d_1)\leq R(\theta,d_2)$ for all $\theta\in\Theta$. However, it is also possible that we may have "d_1 and d_2 are incomparable", that is, $R(\theta,d_1)<R(\theta,d_2)$ for at least one $\theta\in\Theta$, and $R(\theta,d_1)>R(\theta,d_2)$ for at least one $\theta\in\Theta$. Therefore, this ordering gives a partial ordering of the set of estimators.

An estimator d is said to be uniformly non-dominated if there is no estimator uniformly better than d. The conditions that an estimator must satisfy in order that it be uniformly non-dominated are given by the following theorem.

Theorem 2 (*Uniformly non-dominated estimator*). Let (ξ_τ; $\tau=1,2,\ldots$) be a sequence of the prior distributions on the parameter space Θ. Suppose that (d_τ; $\tau=1,2,\ldots$) and ($Q(\xi_\tau,d_\tau)$; $\tau=1,2,\ldots$) are the sequences of Bayes estimators and prior risks, respectively. If there exists an estimator d* such that its risk function $R(\theta,d^*)$, $\theta\in\Theta$, satisfies the relationship

$$\lim_{\tau\to\infty}\left[Q(\xi_\tau,d^*)-Q(\xi_\tau,d_\tau)\right]=0, \tag{11}$$

where

$$Q(\xi_\tau,d)=\int_\Theta R(\theta,d)\xi_\tau(d\theta), \tag{12}$$

then d* is an uniformly non-dominated estimator.

Proof. Suppose d* is uniformly dominated. Then there exists an estimator d** such that $R(\theta,d^{**})< R(\theta,d^*)$ for all $\theta\in\Theta$. Let

$$\varepsilon=\inf_{\theta\in\Theta}\left[R(\theta,d^*)-R(\theta,d^{**})\right]>0. \tag{13}$$

Then

$$Q(\xi_\tau,d^*)-Q(\xi_\tau,d^{**})\geq\varepsilon. \tag{14}$$

Simultaneously,

$$Q(\xi_\tau,d^{**})-Q(\xi_\tau,d_\tau)\geq 0, \tag{15}$$

$\tau=1,2,\ldots$, and

$$\lim_{\tau\to\infty}\left[Q(\xi_\tau,d^{**})-Q(\xi_\tau,d_\tau)\right]\geq 0. \tag{16}$$

On the other hand,

$$\begin{aligned}Q(\xi_\tau,d^{**})-Q(\xi_\tau,d_\tau)&=\left[Q(\xi_\tau,d^*)-Q(\xi_\tau,d_\tau)\right]-\left[Q(\xi_\tau,d^*)-Q(\xi_\tau,d^{**})\right]\\&\leq\left[Q(\xi_\tau,d^*)-Q(\xi_\tau,d_\tau)\right]-\varepsilon\end{aligned} \tag{17}$$

and

$$\lim_{\tau\to\infty}\left[Q(\xi_\tau, d^{**}) - Q(\xi_\tau, d_\tau)\right] < 0. \tag{18}$$

This contradiction proves that d* is an uniformly non-dominated estimator. □

Comparisons of Estimators

In order to judge which estimator might be preferred for a given situation, a comparison based on some "closeness to the true value" criteria should be made. The following approach is commonly used (Nechval, 1982). Consider two estimators, say, d_1 and d_2 having risk function $R(\theta,d_1)$ and $R(\theta,d_2)$, respectively. Then the relative efficiency of d_1 relative to d_2 is given by

$$\text{rel.eff.}_R\{d_1,d_2;\theta\} = R(\theta,d_2)/R(\theta,d_1). \tag{19}$$

When $\text{rel.eff.}_R\{d_1,d_2;\theta_0\} < 1$ for some θ_0, we say that d_2 is more efficient than d_1 at θ_0. If $\text{rel.eff.}_R\{d_1,d_2;\theta\} \le 1$ for all θ with a strict inequality for some θ_0, then d_1 is inadmissible relative to d_2.

Acknowledgements

This work was supported in part by Grant GR-96.0213 from the Latvian Council of Sciences and the National Institute of Mathematics and Informatics of Latvia.

References

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

Åström, K.J. (1970). *Introduction to Stochastic Control Theory.* New York: Academic Press.

Åström, K.J. and Eykhoff, P. (1971). System identification: a survey. *Automatica,* 7, pp. 123-164.

Bertsekas, D.P. (1976). *Dynamic Programming and Stochastic Control.* New York: Academic Press.

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

Fisher, R.A. (1948). Conclusions fiduciares. *Ann. Inst. Henri Poincare*, 10, pp. 191-213.

Fraser, D.A.S. (1961). On fiducial inference. *Ann. Math. Stat.*, 32, pp. 661-676.

Nechval, N.A. (1982). *Modern Statistical Methods of Operations Research*. Riga: RCAEI.

Nechval, N.A. (1984). *Theory and Methods of Adaptive Control of Stochastic Processes*. Riga: RCAEI.

Nechval, N.A. (1988). A new efficient approach to constructing the minimum risk estimators of state of stochastic systems from the statistical data of small samples. In: *Preprint of the 8th IFAC Symposium on Identification and System Parameter Estimation*, pp. 71-76, Beijing, P.R. China.

Nechval, N.A., Nechval, K.N. and Heimanis, B.M. (1997). Optimal rescheduling of airplanes under uncertainty. In: *Proceedings of the Nordic Operations Research Conference (NOAS '97)*, pp. 213-224, University of Copenhagen, Denmark.

Nechval, N.A., Nechval, K.N. and Heimanis, B.M. (1998). Variational approach to adaptive optimization of stochastic systems. In: *Cybernetics and Systems '98* (edited by R. Trappl), Vol. 1, pp. 44-49, Austrian Society for Cybernetic Studies, Vienna, Austria.

Nechval, N.A., Nechval, K.N. and Heimanis, B.M. (1999). Adaptive optimization in stochastic systems via the variational technique. *International Journal of Computing Anticipatory Systems* (edited by Daniel M. Dubois), 3, pp. 317-328.

Nechval, K.N. and Nechval, N.A. (1999). Adaptive dual control of finding an optimal dosage of certain drugs. In: *Proceedings of the First International Symposium on Industrial Statistics*, 19-21 August 1999, Stockholm, Sweden, pp. 25-31.

Wonham, W.M. (1968). On the separation theorem of stochastic control. *SIAM J. Control*, 6, pp. 312-317.

Yakowitz, S.J. (1969). *Mathematics of Adaptive Control Processes*. New York: American Elsevier.