Ulam Quarterly — Volume 1, Number 2, 1992

On Stochastic Adaptive Control*

Bożenna Pasik-Duncan
Department of Mathematics
University of Kansas
Lawrence, Kansas 66045

Abstract

An adaptive control problem for some continuous-time linear stochastic systems and its solution are presented in this paper. The solution includes showing the strong consistency of a family of maximum likelihood (or, equivalently, least squares) estimates of the unknown parameters and the convergence of the average quadratic costs with control based on these estimates to the optimal cost. The stochastic systems are described by linear stochastic differential equations where the unknown parameters occur affinely in the drift term. We shall also consider adaptive control with regard to the discounted cost criterion and with the general noise being a homogeneous stochastically continuous process with independent increments.

1. Introduction

Stochastic adaptive control has become an important area of activity in control theory. With the maturity of this area of research there are a number of books and general surveys (e.g., [1, 2, 4, 10, 11, 13, 14]). The adaptive control problem is to identify the unknown system and simultaneously control it. The models for the stochastic adaptive control problems that will be described here are assumed to evolve in continuous time rather than discrete time because this assumption is natural for many models and it is important for the study of discrete time models when the sampling rates are large and for the analysis of numerical round-off errors. The families of linear systems that are considered here include finite dimensional systems, delay time systems and infinite dimensional systems that are described by infinitesimal generators of C_0 semigroups [5, 6, 7, 8, 12, 15].

* Reserch partially supported by NSF Grants ECS-8718026, ECS-9102714 and ECS-9113029.

The general approach to adaptive control that is described here exhibits a separation of the problems of identification of the unknown parameters and adaptive control. Maximum likelihood (or, equivalently, least squares) estimates are used for the identification of unknown constant parameters [6, 12]. These estimates are given recursively and are shown to be strongly consistent. The adaptive control is usually constructed by the so-called certainty equivalence principle, that is, the optimal stationary control is computed replacing unknown true parameter values by the current estimates of these values [6]. Since the optimal stationary controls can be shown to be continuous functions of the unknown parameters, the self-tuning property is verified. It is shown that the family of average costs using the control from the certainty equivalence principle converges to the optimal average cost [6, 8]. This verifies the self-optimizing property. We shall also consider the adaptive control of linear diffusion processes with regard to the discounted cost criterion [3]. It is shown that the certainty-equivalence type of control is asymptotically discount optimal in the sense of Schäl [16]. We shall announce in this paper that the solution of the adaptive control problem can be obtained also for the systems with the general noise that is a homogeneous, stochastically continuous process with independent increments having a finite variance.

2. The Adaptive Control Scheme

The control as feedback gains will be obtained from the solution of infinite time, quadratic cost control problem by replacing the correct values of parameters in this solution by the estimates of parameters at time t to obtain the feedback gain at time t. The estimates of the parameters are the maximum likelihood estimates. The maximum likelihood estimate based on the observations of the state until time t is used as the correct value of the parameter to solve the infinite time control problem by solving the algebraic Riccati equation. This method gives a feedback gain at each time t and therefore a control policy.

3. Finite Dimensional Linear Systems

The stochastic system for the adaptive control problem that we shall consider first is

$$dX_t = \left(F_0 X_t + \sum_{i=1}^{p} \alpha_i F_i X_i + BU_t \right) dt + dW_t \tag{3.1}$$

where $\alpha_i \in I_i \subset R$ is an unknown parameter and I_i is bounded, open interval for $i = 1, 2, \ldots, p$. $X_t \in R^n$, $U_t \in R^m$, $F_i \in L(R^n, R^n)$ $i = 0, 1, \ldots, p$, $B \in L(R^m, R^m)$, $(W_t,\, t \geq 0)$ is a standard Wiener process and

$X_0 \equiv x \in R^n$. It will be assumed that $(F_i,\, i = 1, 2, \ldots, p)$ are linearly independent.

The probability space $(\Omega, \mathcal{F}, P)$ can be chosen such that Ω is the Frèchet space of R^n-valued continuous functions on $R_+ = [0, +\infty)$ with the seminorms of local uniform convergence. P is the Wiener measure on Ω, and $\mathcal{F}$ is the P-completion of the Borel σ-algebra of Ω.

The cost functional C_t is defined as

$$C_t = \int_0^t \langle QX_s, X_s\rangle + \langle PU_s, U_s\rangle \, ds \tag{3.2}$$

where $Q \in L(R^n, R^n)$ and $P \in L(R^m, R^m)$ are symmetric and positive definite.

The control $(U_t,\, t \geq 0)$ will be given in a feedback form

$$U_t = K_t X_t, \qquad t \geq 0.$$

Since $(K_t,\, t \geq 0)$ will be a stochastic process that is adapted to $\sigma(X_u,\, u \leq t)$, it is necessary to define the solution of (3.1).

We compute K_t from the estimate at $t - \delta$, where $\delta > 0$ is fixed. K_t is adapted to $\sigma(X_u,\, u \leq t - \delta)$ and it becomes a known stochastic process for (3.1) at time t.

Let $R = BP^{-1}B'$ and

$$A = A(\alpha) = F_0 + \sum_{i=1}^{p} \alpha_i F_i.$$

It is well known that if (A, B) is controllable then the deterministic system

$$\frac{dx}{dt} = F_0 x + \sum_{i=1}^{p} \alpha_i F_i x + Bu \tag{3.3}$$

is stabilizable by state feedback.

A stabilizing feedback gain can be obtained from the unique symmetric positive definite solution of the algebraic Riccati equation

$$0 = \Pi A + A' \Pi - \Pi R \Pi + Q. \tag{3.4}$$

The optimal feedback control u_t^* for the infinite time, deterministic optimal control problem with state equation (3.3) and the cost function C_∞ from (3.2) is

$$-P^{-1}B'\Pi x(t).$$

Then $(K_t,\, t \geq 0)$ is said to be a stable control policy or $(X_t,\, t \geq 0)$ is said to be stable if

$$\limsup_{t\to\infty} \frac{1}{t} \int_0^t E\,\langle X_s, X_s\rangle \, ds \leq C < \infty. \tag{3.5}$$

Let $(X_t,\, t \geq 0)$ be the solution of (3.1) with the feedback control policy $(K_t,\, t \geq 0)$. Define $(a_{ij}(t))$ $i, j = 1, \ldots, p$ as

$$a_{ij}(t) := \int_0^t \langle F_i X_s, F_j X_s \rangle \, ds.$$

Let

$$\tilde{A}(t) = \{\tilde{a}_{ij}(t)\} \quad \text{where} \quad \tilde{a}_{ij}(t) = \frac{a_{ij}(t)}{a_{ii}(t)}.$$

For the computation of the maximum likelihood estimate $\hat{\alpha}_t$, it will be assumed that

$$\liminf_{t\to\infty} |\det \tilde{A}(t)| > 0 \qquad \text{a.s.} \tag{3.6}$$

In specific cases it is often easy to verify (3.6). For example, let $(f_k^i,\ k = 1, \ldots, n)$ be the columns of F_i for $i = 1, 2, \ldots, p$. If $\langle f_k^i, f_\ell^j \rangle = 0$ for all $i \neq j$ and $k, \ell \in \{1, s, \ldots, n\}$, then the condition (3.6) is trivially satisfied. For an elementary example of this let $n = 3$ and $p = 2$ where $F_1 = E_{11} + E_{22}$, $F_2 = E_{12} + E_{31}$ and E_{ij} is the elementary 3×3 matrix with 1 in the (i, j) position and zeroes elsewhere.

The likelihood function is obtained from the mutual absolute continuity of the probability measure of (3.1) and the Wiener measure for $(W_t,\, t \geq 0)$ as

$$L_t(\alpha) = \exp\Bigg(\int_0^t \langle A(\alpha) X_s + BU_s, dX_s \rangle - \frac{1}{2} \int_0^t \langle A(\alpha) X_s + BU_s, A(\alpha) X_s + BU_s \rangle \, ds\Bigg). \tag{3.7}$$

To maximize L_t it is necessary (and in fact sufficient) that

$$DL_t = 0 \tag{3.8}$$

$$\frac{\partial L_t}{\partial \alpha_j} = \int_0^t \langle F_j X_s, dX_s \rangle - \int_0^t \langle F_j X_s, F_0 X_s + \sum_{i=1}^p \hat{\alpha}_k F_k X_s + BU_s \rangle ds = 0, \qquad j = 1, 2, \ldots, p.$$

Now use

$$dX_t = A(\alpha_0) X_t dt + BU_t dt + dW_t$$

and rewrite (3.8) to obtain

$$\hat{\alpha}_j \int_0^t \langle F_j X_s, F_j X_s \rangle \, ds = \alpha_{0j} \int_0^t \langle F_j X_s, F_j X_s \rangle \, ds + \int_0^t \langle F_j X_s, dW_s \rangle$$
$$+ \int_0^t \langle \alpha_{0k} F_j X_s, \sum_{\substack{k \\ k \neq j}} \alpha_{0k} F_k X_s \rangle ds - \int_0^t \langle F_j X_s, \sum_{\substack{k \\ k \neq j}} \hat{\alpha}_k F_k X_s \rangle ds. \quad (3.9)$$

Define

$$a_{jk}(t) := \int_0^t \langle F_j X_s, F_k X_s \rangle \, ds$$
$$b_j(t) := \int_0^t \langle F_j X_s, dW_s \rangle$$
$$\tilde{a}_{jk}(t) := \frac{a_{jk}(t)}{a_{jj}(t)}$$
$$\tilde{b}_j(t) := \frac{b_j(t)}{a_{jj}(t)}.$$

Thus the family of linear equations (3.9) can be rewritten as

$$\tilde{A}(t)\hat{\alpha}_t = \tilde{A}(t)\alpha_0 + \tilde{b}(t). \quad (3.10)$$

From this we get

$$\hat{\alpha}_t = \alpha_0 + \tilde{A}^{-1}(t)\tilde{b}(t). \quad (3.11)$$

Strong Consistency Property.

Theorem 3.1. *Let* $(K_t,\, t \geq 0)$ *be a control policy that is uniformly bounded almost surely. Assume that* (3.6) *is satisfied. Then*

$$P\left(\lim_{t \to \infty} \hat{\alpha}_t = \alpha_0\right) = 1$$

Proof: To verify this theorem the following result is important. If $(W_t, \mathcal{F}_t)$ is a Wiener process on $(\Omega, \mathcal{F}, P)$ and (f_t) is $\mathcal{F}_t$-adapted s.t.

$$\int_0^t \langle f_s, f_s \rangle \, ds < \infty \quad \text{a.s. for} \quad 0 \leq t < \infty$$

and

$$\int_0^\infty \langle f_s, f_s \rangle \, ds = +\infty \quad \text{a.s.}$$

then

$$\lim_{t\to\infty} \frac{\int_0^t \langle f_s, dW_s\rangle}{\int_0^t \langle f_s, f_s\rangle\, ds} = 0 \quad \text{a.s.}$$

We shall show that our estimates are recursive estimates of the unknown parameters

$$A(t)\hat{\alpha}_t = A(t)\alpha_0 + b(t)$$

where

$$b(t) = (b_1(t), \ldots, b_p(t))'.$$

Let $\langle \tilde{F}X, Y\rangle$ be the vector whose i^{th} component is $\langle F_i X, Y\rangle$

$$\hat{\alpha}_t = A^{-1}(t) \int_0^t \left\langle \tilde{F}X_s, dX_s - F_0 X_s ds - BU_s ds \right\rangle \tag{3.12}$$

$A^{-1}(t)$ satisfies $dA^{-1}(t) = -A^{-1}(t)dA(t)A^{-1}(t)$ so

$$d\hat{\alpha}_t = A^{-1}(t) \left\langle \tilde{F}X_t, dX_t - A(\hat{\alpha}_t)X_t dt - BU_t dt \right\rangle.$$

The Self-Optimizing Property.

Theorem 3.2. *Let $(K_t,\, t \geq 0)$ be the feedback gains determined from the algebraic Riccati equation using the maximum likelihood esitmates $(\hat{\alpha}_t)$ as the parameter values. Assume that* (3.6) *is satisfied for the control policy $(K_t,\, t \geq 0)$ and that (A, B) is controllable. The feedback gains $(K_t,\, t \geq 0)$ form a stable control policy and*

$$\lim_{t\to\infty} \frac{1}{t} C_t = tr\, V \quad a.s. \tag{3.13}$$

where V is the unique, symmetric, positive definite solution of Riccati equation using α_0, and C_t is the cost at time t using the feedback gains $(K_t,\, t \geq 0)$.

The solution of the adaptive control can be also obtained for the following linear systems with delays [5, 15].

$$dX_t = \sum_{i=1}^{k} F_0^i X_{t-\tau_i} + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \alpha_{ij} F_j^i X_{t-\tau_i} + BU_t dt + dW_t$$

where $\alpha_{ij} \in I_{ij} \subset R$, $\tau_i \geq 0$, $x_t \in R^n$, $U_t \in R^m$, $F_j^i \in L(R^n, R^n)$, $B \in L(R^m, R^m)$, $(W(t),\, t \geq 0)$ is a standard Wiener process and $\tilde{X}_0 \equiv h$ is a fixed function, $(F_j^i,\ j = 1, 2, \dots, n_i)$ are linearly independent. The cost functional is

$$C_t = \int_0^t \langle QX_s, X_s\rangle + \langle PU_s, U_s\rangle\, ds$$

where $Q \in L(R^n, R^n)$ and $P \in L(R^m, R^m)$ are symmetric and positive definite.

4. Infinite Dimensional Linear systems

The model for the adaptive control problem is described by the following infinite dimensional stochastic differential equation

$$dX_t = A(\alpha)X_t dt + BU_t dt + dW_t \tag{4.1}$$

where

$$A(\alpha) = F_0 + \sum_{i=1}^{p} \alpha_i F_i,$$

$X_t \in H$, H is a real, separable, infinite dimensional Hilbert space, $X_0 = x$ is an element of H, $(W_t,\ t \geq 0)$ is an H-valued Wiener process s.t. W_1 has the nuclear covariance Q_W, $A(\alpha)$ is the generator of a strongly continuous semigroup $(G_t^\alpha,\ t \geq 0)$ of bounded linear operators on H, $B \in L(H_1, H)$ where H_1 is a real, separable Hilbert space and $t \geq 0$. The controls U_t take values in H_1.

The inner product and norm will be denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$ and the operator norm by $\|\cdot\|$.

For adaptive control problem the control policies $(U_t,\ t \geq 0)$ that are considered are linear feedback controls

$$U_t = K_t X_t,$$

where $(K_t,\ t \geq 0)$ is an $L(H, H_1)$-valued adapted process that is uniformly bounded a.s.

The cost functional that is to be minimized is

$$C_t(T) = \int_0^t \langle QX_s, X_s\rangle + \langle PU_s, U_s\rangle ds \tag{4.2}$$

over all $U \in L^2(0, T; H_1)$ where $Q \in L(H, H)$ and $P \in L(H_1, H_1)$ are bounded, symmetric positive definite and bounded, symmetric, positive definite, respectively, and

$$\exists\, \varepsilon > 0\ \forall\, x \quad \langle Px, x\rangle \geq \varepsilon |x|^2.$$

The probability space is $(\Omega, \mathcal{F}, P)$, where $\Omega = C([0,\infty), H)$, P is the distribution of a Wiener process $(W_t,\, t \geq 0)$ s.t. $\mathrm{Cov}(W(1), W(1)) = Q_W$ and $\mathcal{F}$ is the P-completion of the Borel σ-algebra on Ω.

The mild solution of (4.1) is

$$X_t = G_t X_t + \int_0^t G_{t-s} B U_s ds + \int_0^t G_{t-s} dW_s \tag{4.3}$$

where $(G_t,\, t \geq 0)$ is the semigroup $(e^{tA},\, t \geq 0)$.

The mild solution is equivalent to the following two other notions of solution:

(i) $X_t = X_0 + A \int_0^t X_s ds + \int_0^t BU_s\, ds + W_t$ and

(ii) for each $y \in D(A^*)$

$$\langle y, X_t \rangle = \langle y, X_0 \rangle + \int_0^t \langle A^* y, X_s \rangle ds + \int_0^t \langle y, BU_s \rangle ds + \langle y, W_t \rangle.$$

Assumptions:

(A1) $\alpha_i \in I_i \subset R$ is an unknown parameter where I_i is a bounded, open interval for $i = 1, 2, \ldots, p$ and $\mathcal{A} = I_1 \times \ldots \times I_p \subset \mathcal{K}$ where $\mathcal{K}$ is a compact set.

(A2) For each $\alpha \in \mathcal{K}$, $A(\alpha)$ is the infinitesimal generator of C_0 semigroup.

(A3) For each $\alpha \in \mathcal{K}$, generators $A(\alpha)$ are uniformly exponentially stable, that is

$$\exists\, M > 0 \quad \text{and} \quad \omega > 0 \text{ s.t.} |e^{tA(\alpha)}| \leq M e^{-\omega t}$$

for all $\alpha \in \mathcal{K}$.

(A4) The map $\alpha \to e^{tA(\alpha)}$ is continuous for $t > 0$ and $\alpha \in \mathcal{K}$ in the uniform operator topology.

(A5) $\bigcap_{i=0}^{p} D(F_i)$ is dense in H.

(A6) There is a finite dimensional projection $\tilde{P}$ with range in $\bigcap_{i=0}^{p} D(F_i^*)$ that is dense in H s.t. the family $(\tilde{P} F_i,\, i = 1, \ldots, p)$ is linearly independent on $\bigcap_{i=0}^{p} D(F_i)$.

Define

$$a_{mn}(t) := \int_0^t \left\langle \tilde{P} F_m X_s, \tilde{P} F_n X_s \right\rangle ds$$

and

$$A(t) = (a_{mn}(t)), \qquad m, n \in \{1, \dots, p\}.$$

Let

$$\tilde{A}(t) = (\tilde{a}_{ij}(t))$$

where

$$\tilde{a}_{ij}(t) = \frac{a_{ij}(t)}{a_{ii}(t)}$$

for $i, j \in \{1, \dots, p\}$.

(A7) $\liminf_{t\to\infty} |\det \tilde{A}(t)| > 0 \qquad$ a.s.

Let $\delta > 0$ be fixed. It is assumed that the $L(H, H_1)$ valued process $(K_t,\, t \geq 0)$ has the property that K_t is adapted to $\sigma(X_u,\, u \leq t - \delta)$ for each $t \geq \delta$, and it is assumed that $(K_t,\, t \in (0, \delta))$ is a bounded deterministic operator. For the adaptive control problem this measurability is accomplished by computing K_t from the parameter estimate at time $t - \delta$.

If $(K_t,\, t \geq 0)$ satisfies this measurability condition and the boundedness condition, then $\exists!$ mild solution of

$$dX_t = (A(\alpha) + BK_t)X_t dt + dW_t. \tag{4.9}$$

Let $(X_t,\, t \geq 0)$ be the mild solution of (4.9).

The Strong Consistency Property.

Theorem 4.1. *Let $(K_t,\, t \geq 0)$ be a feedback control policy s.t. K_t is $\sigma(X_s,\, s \leq t - \delta)$ measurable for each $t \geq \delta$ and*

$$\sup_{t\geq 0} \|K(t)\| \leq M \qquad a.s.$$

where $\delta > 0$ is fixed, $M \in R_+$. Assume (A1–A3) *and* (A5–A7) *are satisfied. Then the family of least squares estimates $(\hat{a}_t,\, t \geq 0)$ from the process $(X_t,\, t \geq 0)$ is strongly consistent, that is*

$$P\left(\lim_{t\to\infty} \hat{a}_t = \alpha_0\right) = 1.$$

The Continuity Property.

Theorem 4.2. *Assume that (A1–A4) are satisfied. Then the mild solution $(V(\alpha),\, \alpha \in \mathcal{K})$ of the stationary Riccati equation is a continuous function of $\alpha \in \mathcal{K}$ in the uniform operator topology.*

For the proof see [8].

The Self-Optimizing Property.

Theorem 4.3. *Assume* (A1–A7) *are satisfied. Let* $(K_t,\, t \geq 0)$ *be the family of feedback gains determined from the stationary Riccati equation using the* (*truncated*) *least squares estimates* $(\hat{a}_{t-\delta},\, t \geq 0)$. *Then*

$$\lim_{t\to\infty} K(t) = k_0 \qquad a.s.$$

in the uniform operator topology and

$$\lim_{t\to\infty} \frac{1}{t} C_t = tr(VQ_W) \qquad a.s.$$

where V *is the unique bounded, nonnegative symmetric solution of the stationary Riccati equation using the true parameter value* α_0, C_t *is the cost at time* t *using* $(K_t,\, t \geq 0)$ *and* $Q_W = Cov(W_1, W_1)$.

Examples:

1. A very simple example of the model (4.1) is a stochastic heat equation on the unit cube Γ in R^n with zero temperature on the boundary and unknown thermal conductivity, where it is assumed that the specific heat c and the mass density ρ are known and $c\rho = 1$. Let Δ be the Laplacian on R^n.
$F := \Delta$ with the domain $D(F) = H^2(\Gamma) \cap H_0^1(\Gamma)$ generates a compact semigroup $(G_t,\, t > 0)$ of bounded linear operators on $H = L^2(\Gamma)$ and this semigroup is exponentially stable. Let$A(\alpha) = \alpha F$ where $\alpha \in I \subset [a, b]$ and $a > 0$. $D(A(\alpha)) = D(F)$ for $\alpha \in I$ and $A(\alpha)$ generates the semigroup $(G_t^\alpha,\, t \geq 0)$ where $G_t^\alpha = G_{\alpha t}$. $\alpha \to G_{\alpha t}$ is continuous in the uniform operator topology.

2. $A(\alpha) = \alpha_1 F_1 + \alpha_2 F_2 \qquad \alpha_1 > 0$
where

$$F_1 := \frac{\partial^2}{\partial x^2} \qquad \text{and} \qquad F_2 = \frac{\partial}{\partial x}$$

$(\alpha_1, \alpha_2) \in I_1 \times I_2$
the semigroup G_t^α generated by $\alpha_1 F_1 + \alpha_2 F_2$ is $\forall\, t$ continuous in $\alpha = (\alpha_1, \alpha_2)$ in the uniform operator topology.

5. Adaptive Control With Discounted Cost Criterion

We consider the stochastic controlled system (3.1). The cost criterion is a discounted one, namely, is given in terms of the process C_t^β defined as

$$C_t^\beta(u) = \int_0^t e^{-\beta s} \left(\langle QX_s, X_s\rangle + \langle PU_s, U_s\rangle\right) ds, \quad t \geq 0 \tag{5.1}$$

where Q, P are symmetric, positive definite and $\beta > 0$ is a discount factor.

The Strong Consistency Property holds in this case [3]. It can be shown that the certainty-equivalence control is asymptotically discount optimal in the sense of Schäl [16], which means

$$\lim_{t\to\infty}\left|E_x V(X_t,\beta) - E_x \int_t^{+\infty} e^{-\beta(s-t)} F(X_s,U_s)ds\right| = 0,\ x \in R^n \tag{5.2}$$

and

$$\lim_{t\to\infty}\left|V(X_t,\beta) - \int_t^{\infty} e^{-\beta(s-t)} F(X_s,U_s)ds\right| = 0 \qquad \text{a.s.} \tag{5.3}$$

where $(U_t,\, t \geq 0)$ is given by

$$U_t = \begin{cases} 0, & 0 \leq t < \delta \\ K(\hat{\alpha}_{t-\delta,\beta})X_t, & \delta \leq t \end{cases}$$

for some fixed $\delta > 0$, $F(X,U) = \langle QX, X\rangle + \langle PU, U\rangle$ and V is the Bellman function (for $\alpha = \alpha_0$) [3]. E_x is the conditional expectation with respect to $P(\,\cdot\,|X_0 = x)$.

6. General Noise Case

Let us consider the following stochastic system:

$$dX_t = \left(F_0 X_t + \sum_{i=1}^{p} \alpha_i F_i X_t + BU_t\right) dt + d\xi_t. \tag{6.1}$$

In this system standard Wiener process (see (3.1)) was replaced by the process $(\xi_t,\, t \geq 0)$ which is homogeneous, stochastically continuous with independent increments having a finite variance for each $t \geq 0$, that is, it can be represented as [9]

$$\xi_t = at + W_t + \int_{R^n} X(\nu(t,dx) - t\,\Pi(dx)) \tag{6.2}$$

where $a \in R$, $(W_t,\, t \geq 0)$ is a standard n-dimensional Wiener process, $\nu(dt,dx)$ is a Poisson random measure, $E\nu(t,A) = t\,\Pi(A)$ for each Borel set A and $\Pi = \int_{R^n} XX'\Pi(dx)$, ν and W are independent.

Remark. To simplify the subsequent formulas we set $a = 0$. This is simply centering ξ_t around 0.

We fix the canonical probability measure $(\Omega, \mathcal{F}, P)$ for $(\xi_t,\, t \geq 0)$ and E above denotes the expectation with respect to P. We will make use of the following structural assumption imposed on system (6.1):

(A1) $\alpha_i \in I_i \subset R$, where I_i is a bounded, open interval for $i = 1, 2, \ldots, p$;

(A2) The pair $\left(F_0 + \sum_{i=1}^{p} \alpha_i F_i, B\right)$ is controllable for each $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathop{X}_{i=1}^{p} I_i = I$.

We will consider the infinite horizon problems only with β-discounted and ergodic cost functionals defined via the processes $C_t(x, \beta, u)$ and $C_t(x, u)$, respectively,

$$C_t(x, \beta, u.) = E_x \int_0^t e^{-\beta s} F(x_s, u_s) ds, \tag{6.3}$$

$$C_t(x, u.) = \int_0^t F(x_s, u_s) dx, \tag{6.4}$$

where $\beta \in (0, +\infty)$ is a discount factor, $u_x = (U_t,\, t \geq 0)$, $F(x, u) = \langle QX, X\rangle + \langle RU, U\rangle$ for some $Q \in L(R^n, R^n)$, $R \in L(R^m, R^m)$ and $Q > 0$, $R > 0$. E_x denotes the expectation with respect to conditinal probability $P_x(\cdot) = P(\,\cdot\,|x_0 = x)$. The corresponding cost functionals are

$$\mathcal{I}(x, \beta, u.) = \lim_{t \to \infty} C_t(x, \beta, u.) \tag{6.5}$$

and

$$\mathcal{I}(x, u.) = \varliminf_{t \to \infty} \frac{1}{t} C_t(x, u.). \tag{6.6}$$

The existence of optimal control can be shown for both cost functionals.

Strong Consistency of the Least Squares Estimator of α_0***.***

We assume that α in (6.1) is unknown. Let α_0 denote the unknown true value of the vector of parameters in model (6.1). We can construct the least squares estimator $\hat{\alpha}_t$ of α and prove strong consistency of the estimator. Let us first rewrite model (6.1) in the form

$$\begin{aligned} &dX_t = F_0 X_t dt + H_t' \alpha dt + BU_t dt + d\xi_t \\ &X_0 = x, \qquad t \geq 0 \end{aligned}$$

where

$$H_t = [F_1 X_t, F_2 X_t, \ldots, F_p X_t]', \qquad t \geq 0.$$

Let, for $t \geq 0$,

$$A(t) = \int_0^t H_s H_s' ds + I$$

and

$$\tilde{A}(t) = \operatorname{diag}\left[\frac{1}{a_{11}(t)}, \dots, \frac{1}{a_{pp}(t)}\right] A(t),$$

where $a_{ii}(t)$ is the i-th diagonal element of $A(t)$, $i = 1, 2, \dots, p$. Also define

$$\tilde{N}_t = \operatorname{diag}\left[\frac{1}{a_{11}(t)}, \dots, \frac{1}{a_{pp}(t)}\right] \int_0^t H_s d\xi_s, \ t \geq 0.$$

We define the least squares estimate $\hat{\alpha}_t$ of α_0 by

$$\hat{\alpha}_t = A^{-1}(t) \int_0^t H_s(dX_s - BU_s ds - F_0 X_s ds)$$

for $t \geq 0$. Then

$$\hat{\alpha}_t = \alpha_0 + \tilde{A}^{-1}(t) \operatorname{diag}\left[\frac{1}{a_{11}(t)}, \dots, \frac{1}{a_{pp}(t)}\right] \alpha_0 + \tilde{A}^{-1}(t)\tilde{N}(t), \qquad t \geq 0.$$

We will need the following two assumptions:

$$|U_t|^2 \leq \text{ const}|X_t|^2, \qquad t \geq 0 \tag{A.3}$$

$$\liminf_{t \to \infty} |\det \tilde{A}(t)| > 0. \tag{A.4}$$

The strong consistency property and the self-optimizing property in the ergodic case or the asymptotically discount optimality in the sense of Schäl in the discounted case can be shown.

References

[1] A. Aloneftis, *Stochastic Adaptive Control,* Lecture Notes in Control and Inf. Sc. 98, Springer-Verlag, Berlin 1987.

[2] K. J. Åström and B. Wittenmark, *Adaptive control,* Addison-Wesley, Reading 1989.

[3] T. R. Bielecki, *A note on adaptive control of continuous-time linear stochastic systems with discounted cost criterion,* to appear in JOTA.

[4] P. E. Caines, *Linear Stochastic Systems,* John Wiley, New York, 1988.

[5] T. E. Duncan and B. Pasik-Duncan, *Adaptive control of continuous time linear delay time systems,* Stochastics 24 (1988), 45–74.

[6] T. E. Duncan and B. Pasik-Duncan, *Adaptive control of continuous time liner sochastic systems,* Mathematics of Control, Signals and Systems 3 (1990), 45–60.

[7] T. E. Duncan and B. Pasik-Duncan, *Some aspects of the adaptive control of stochastic evolution equations,* Proc. 28th conf. on Decision and control (1989), 732–735.

[8] T. E. Duncan, B., Goldys, and B. Pasik-Duncan, *Adaptive control of linear stochastic evolution systems,* Stochastics and Stochastics Reports, Vol. 36 (1991), 71–90.

[9] I. I. Gikhman and A. V. Skorokhad, *Stochastic differential equations,* springer-Verelag, New York, 1972.

[10] P. R. Kumar, *A survey of some results in stochastic adaptive control,* SIAM J. Control Optim. 23 (1985), 329–380.

[11] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control,* Prentice Hall, Englewood Cliffs, 1986.

[12] P. Mandl, T. E. Duncan and B. pasik-Duncan, *On the consistency of a least squares identifiction procedure,* Kybernetika 24 (1987), 3, 296–307.

[13] B. Pasik-Duncan, *On adaptive control,* Central School of Planning and Statistics Publishers, Warsaw, 1986.

[14] B. Pasik-Duncan, *On adaptive control of continuous time linear Sto chastic systems;* Lecture Notes in Control and Info. Sci. 136 (1989), 328–343.

[15] B. Pasik-Duncan, *On the consistency of a least squares identifiction procedure in linear evolution systems,* to appear in Stochastics.

[16] M. Schäl, *Estimation and Control in Discounted Stochastic Dynamic Programming,* Stochastics 20 (1987), 51–71.

This electronic publication and its contents are ©copyright 1992 by Ulam Quarterly. Permission is hereby granted to give away the journal and its contents, but no one may "own" it. Any and all financial interest is hereby assigned to the acknowledged authors of individual texts. This notification must accompany all distribution of Ulam Quarterly.