

A minimum principle for stochastic control problems with output feedback

Huibert KWAKERNAAK

Department of Applied Mathematics, Twente University of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands

Received 5 December 1980

A minimum principle for stochastic control problems with output feedback is derived by applying Bismut's minimum principle for stochastic control problems with full information about the past to the Kushner–Stratonovitch equation describing the controlled evolution of the conditional density of the state. The well-known solution of the linear-quadratic Gaussian problem is obtained from the principle.

Keywords: Stochastic differential equations, Minimum principle, Output feedback, Stochastic optimal control.

1. Introduction

Consider the stochastic system described by the scalar stochastic differential equation

$$dx_t = f(x_t, u_t, t)dt + g(x_t, t)dv_t, \quad 0 \leq t \leq T, \quad (1)$$

and the output equation

$$dy_t = h(x_t, t)dt + dw_t, \quad 0 \leq t \leq T. \quad (2)$$

Extension to the vector case presents no problems. The terminal time T is fixed and known. The processes v and w are independent Brownian motions, independent of the initial state x_0 . We study the problem of minimizing a criterion of the form

$$E \left\{ \int_0^T \gamma(x_t, u_t, t) dt + \eta(x_T) \right\}. \quad (3)$$

The sigma algebra generated by y_s , $0 \leq s \leq t$, will be denoted as Y_t . Since we consider output feedback, for each t the input u_t is required to be Y_t -measurable. We do not detail the conditions under which the problem formulation as described indeed makes sense.

2. Kushner–Stratonovitch equation

Let p_t denote the conditional density of the state x_t given Y_t . Under conditions studied by

Pardoux [6] this density exists and satisfies the Kushner–Stratonovitch equation [3]

$$dp_t(x) = L_t(u_t)p_t(x)dt + [h(x, t) - \widehat{h}(x_t, t)]p_t(x)[dy_t - \widehat{h}(x_t, t)dt]. \quad (4)$$

Here $L_t(u_t)$ is a partial differential operator defined by

$$L_t(u_t)p_t(x) := -\frac{\partial}{\partial x} [f(x, u_t, t)p_t(x)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [g^2(x, t)p_t(x)]. \quad (5)$$

Furthermore,

$$\widehat{h}(x_t, t) := \int h(x, t)p_t(x)dx. \quad (6)$$

In terms of the innovations process v_t , $0 \leq t \leq T$, defined by

$$dv_t = dy_t - \widehat{h}(x_t, t)dt, \quad (7)$$

the conditional density p_t evolves according to

$$dp_t(x) = L_t(u_t)p_t(x)dt + [h(x, t) - \widehat{h}(x_t, t)]p_t(x)dv_t, \quad 0 \leq t \leq T. \quad (8)$$

It is widely believed and under certain conditions it has been rigorously proved (Fujisaki, Kallianpur and Kunita [2], Lipster and Shirayev [4]) that the innovations process v is a standard Brownian motion. We shall continue the analysis under the assumption that this is indeed the case.

The criterion (3) can be expressed in the conditional density p_y as

$$E \left[\int_0^T dt \int \gamma(x, u_t, t)p_t(x)dx + \int \eta(x)p_T(x)dx \right]. \quad (9)$$

Therefore, the problem of controlling (1)–(2) optimally is equivalent to controlling the evolution of the conditional density as described by (8) through the input u such that the criterion (9) is minimized.

This was recognized by Mortensen [5] many years ago. Mortensen continued from this point with a dynamic programming approach. We choose a different route and consider the application of the stochastic minimum principle.

3. Bismut's minimum principle

Bismut's version [1] of the stochastic minimum principle applies to the minimization of the criterion (3) for the system (1) with full information about the past. Defining F_t as the sigma algebra generated by x_0 and $v_s, 0 \leq s \leq t$, the input u_t is required to be F_t -measurable for each t . Consider the Hamiltonian

$$H(t, x, u, q, r) := qf(x, u, t) + rg(x, t) + \gamma(x, u, t). \tag{10}$$

Suppose that the control $u_t, 0 \leq t \leq T$, minimizes (3), and let $x_t, 0 \leq t \leq T$, be the corresponding solution of (1). Suppose also that $q_t, 0 \leq t \leq T$, and $r_t, 0 \leq t \leq T$, are F -adapted processes (i.e. for each t both q_t and r_t are F_t -measurable), satisfying the adjoint equation

$$-dq_t = \frac{\partial H}{\partial x}(t, x_t, q_t, r_t) dt + \frac{\partial H}{\partial g}(t, x_t, q_t, r_t) dv_t, \tag{11}$$

$$q_T = \frac{d\eta}{dx}(x_T).$$

Then u_t minimizes $H(t, x_t, v, q_t, r_t)$ with respect to v for each t .

4. Minimum principle for output feedback

We apply a generalized version of Bismuth's minimum principle to the problem of controlling p_t as described by the stochastic partial differential equation

$$dp_t = L_t(u_t)p_t dt + \sigma_t dv_t, \quad 0 \leq t \leq T, \tag{12}$$

such that

$$E \left[\int_0^T \langle \gamma, p_t \rangle dt + \langle \eta, p_T \rangle \right] \tag{13}$$

is minimal. Here

$$\sigma_t(x) := [h(x, t) - \widehat{h(x_t, t)}] p_t(x), \tag{14}$$

and $\langle \cdot, \cdot \rangle$ denotes

$$\langle a, b \rangle := \int a(x)b(x) dx. \tag{15}$$

We form the generalized Hamiltonian

$$H(t, p, u, q, r) := \langle q, L_t p \rangle + \langle r, \sigma \rangle + \langle \gamma, p \rangle. \tag{16}$$

The adjoint equation takes the form

$$-dq_t = \frac{\delta H}{\delta p}(t, p_t, u_t, q_t, r_t) dt + \frac{\delta H}{\delta \sigma}(t, p_t, u_t, q_t, r_t) dv_t, \tag{17}$$

$$q_T = \frac{\delta}{\delta p} \langle \eta, p \rangle_{t=T},$$

with $\delta \cdot / \delta \cdot$ denoting a Fréchet derivative. It remains to specify precisely on which space this is defined. then if u_t is optimal, it minimizes $H(t, p_t, v, q_t, r_t)$ with respect to v .

This can be worked out in more detail. We have $\delta H / \delta \sigma = r$. Writing

$$H = \langle q, L_t p \rangle + \langle r, (h - \langle h, p \rangle) p \rangle + \langle \gamma, p \rangle = \langle L_t^* q, p \rangle + \langle r, hp \rangle - \langle h, p \rangle \langle r, p \rangle + \langle \gamma, p \rangle \tag{18}$$

it follows that

$$\frac{\delta H}{\delta p} = L_t^* q + hr - \langle h, p \rangle r - \langle r, p \rangle h + \gamma. \tag{19}$$

Here L_t^* is the adjoint of L_t , defined by

$$L_t^*(u_t)q(x) = f(x, u_t) \frac{\partial q}{\partial x}(x) + \frac{1}{2} g^2(x, t) \frac{\partial^2}{\partial x^2} q(x). \tag{20}$$

We thus obtain for the adjoint equation

$$-dq_t(x) = L_t^*(u_t)q_t(x) dt + r_t(x)h(x, t) dt - \widehat{h(x_t, t)}r_t(x) dt - \widehat{r_t(x_t)}\widehat{h(x, t)} dt + \gamma(x, u_t, t) dt + r_t(x) dv_t, \tag{21}$$

with the terminal condition

$$q_T(x) = \eta(x). \tag{22}$$

These equations are to be satisfied by Y -adapted processes q_t and r_t , taking their values in function space. Omitting the terms from the Hamiltonian that do not depend on the control it follows that for each t in $[0, T]$ the optimal u_t minimizes

$$\langle L_t^*(v)q_t, p_t \rangle + \langle \gamma(\cdot, v, t), p_t \rangle \quad (23)$$

with respect to v .

5. Application to the LQG problem

As an application we consider in this section the linear system

$$\begin{aligned} dx_t &= \alpha x_t dt + \beta u_t dt + \lambda dv_t, \\ dy_t &= \gamma x_t dt + dw_t, \end{aligned} \quad (24)$$

where for simplicity we take α , β , λ and γ constants. The criterion is quadratic and takes the form

$$\frac{1}{2} E \left[\int_0^T (x_t^2 + \rho u_t^2) dt + P_1 x_T^2 \right], \quad (25)$$

with ρ a positive and P_1 a nonnegative constant. The adjoint operator L_t^* now takes the form

$$L_t^* = (\alpha x + \beta u_t) \frac{\partial}{\partial x} + \frac{1}{2} \lambda^2 \frac{\partial^2}{\partial x^2}. \quad (26)$$

Substitution into (23) shows that the control-dependent terms of the Hamiltonian are

$$\beta v \int \frac{\partial q_t(x)}{\partial x} p_t(x) dx + \frac{1}{2} \rho v^2. \quad (27)$$

This expression is minimized by

$$u_t = -\frac{\beta}{\rho} \int p_t(x) \frac{\partial q_t(x)}{\partial x} dx. \quad (28)$$

Written out in full for the case at hand, the adjoint equation takes the form

$$\begin{aligned} -dq_t(x) &= (\alpha x + \beta u_t) \frac{\partial q_t(x)}{\partial x} dt \\ &\quad + \frac{1}{2} \lambda^2 \frac{\partial^2 q_t(x)}{\partial x^2} dt \\ &\quad + \gamma(x - \hat{x}_t) r_t(x) dt \\ &\quad - \gamma x \widehat{r_t(x)} dt + \frac{1}{2} (x^2 + \rho u_t^2) dt \\ &\quad + r_t(x) dv_t, \end{aligned} \quad (29)$$

$$q_T(x) = \frac{1}{2} P_1 x^2.$$

These equations may be satisfied by a solution of the form

$$\begin{aligned} q_t(x) &= \frac{1}{2} P(t) x^2 + \frac{1}{2} A(t) (x - \hat{x}_t)^2 \\ &\quad + \pi(t), \quad 0 \leq t \leq T. \end{aligned} \quad (30)$$

P , A and π are scalar functions, of which P satisfies the Riccati equation

$$\begin{aligned} -\dot{P}(t) &= 2\alpha P(t) + 1 - \frac{\beta^2}{\rho} P^2(t), \\ P(T) &= P_1. \end{aligned} \quad (31)$$

We note that q_t , as given by (30) is Y -adapted, as required. Evaluation of (28) with the aid of (30) shows that

$$u_t = -\frac{\beta}{\rho} P(t) \hat{x}_t, \quad (32)$$

which is of course the optimal control. It remains to verify (29) and to determine the functions A and π . The terminal condition on q_t is satisfied if

$$A(T) = 0, \quad \pi(T) = 0. \quad (33)$$

We know in the present case the conditional expectation \hat{x}_t of x_t given Y_t satisfies

$$\begin{aligned} d\hat{x}_t &= \alpha \hat{x}_t dt + \beta u_t dt + k(t) dv_t, \\ &\quad 0 \leq t \leq T, \end{aligned} \quad (34)$$

where k is the Kalman–Bucy filter gain. The adjoint equation (29) may be verified by substituting the proposed solution (30) and using stochastic calculus with the aid of (34). Identification of the martingale terms of the resulting expression shows that

$$A(t)(x - \hat{x}_t)k(t) dv_t = r_t(x) dv_t. \quad (35)$$

From this we obtain that the Y -adapted process r_t is given by

$$r_t(x) = A(t)(x - \hat{x}_t)k(t). \quad (36)$$

It follows immediately that $\widehat{r_t(x_t)} = 0$. By identification of the remaining terms it may be seen that the adjoint equation is satisfied provided A and π are solved from the differential equations

$$\begin{aligned} -\dot{A}(t) &= 2[\alpha + \gamma k(t)] A(t) + \frac{\beta^2}{\rho} P^2(t), \\ A(T) &= 0, \end{aligned} \quad (37)$$

$$-\dot{\pi}(t) = \frac{1}{2} [k^2(t) + \lambda^2] A(t),$$
$$\pi(T) = 0.$$

6. Concluding remarks

The minimum principle for output feedback control problems obtained in this note shows that the optimal control u_t minimizes (23) with respect to v , where the adjoint variable q_t takes its values in function space and satisfies the adjoint equation (21)–(22). This adjoint equation is a stochastic partial differential equation and is in fact the adjoint of the Kushner–Stratonovitch equation.

The application of this minimum principle to other problems than the LQG problem still remains to be tackled. Also, since the derivation of

the minimum principle given here is formal, its range of validity remains to be established.

References

- [1] J.-M. Bismut, An introductory approach to duality in optimal stochastic control, *SIAM Review* 20 (1978) 62–78.
- [2] M. Fujisaki, G. Kallianpur and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, *Osaka J. Math.* 9 (1972).
- [3] H.J. Kushner, Dynamical equations for optimal nonlinear filtering. *J. Differential Equations* 3 (1967) 179–190.
- [4] R.S. Lipster and A.N. Shiriyayev, *Statistics of Random Processes, Vol. II* (Springer, Berlin, 1977).
- [5] R.E. Mortensen, Stochastic optimal control with noisy observations, *Internat. J. Control* 4 (1966) 455–464.
- [6] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* 3 (1979) 127–167.