ELSEVIER

Systems & Control Letters 25 (1995) 151–157

SYSTEMS & CONTROL LETTERS

Finite-dimensional quasi-linear risk-sensitive control

Lakhdar Aggoun [a], Alain Bensoussan [b], Robert J. Elliott [c,*], John B. Moore [d]

[a] *Department of Statistics, University of Auckland, private Bag 92019, Auckland, New Zealand*
[b] *I.N.R.I.A. Rocquencourt, 78153 Le Chesnay Cedex, France*
[c] *Department of Mathematical Sciences, University of Alberta, Edmonton, AB, Canada*
[d] *Department of Systems Engineering and Cooperative Research Centre for Robust and Adaptive Systems, Research School of Physical Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia*

Received 8 October 1993; revised 10 July 1994

Abstract

A discrete-time partially observed stochastic control problem with exponential running cost is considered. The dynamics are linear and the running cost quadratic in the state variable, but the control may enter nonlinearly. Explicit solutions for a forward Zakai equation and a backward adjoint equation are derived in terms of finite-dimensional dynamics. This enables the partially observed problem to be expressed in finite-dimensional terms and a separation principle applied.

Keywords: Risk-sensitive partially observed stochastic control; Finite-dimensional information states

1. Introduction

Recent interest in risk-sensitive stochastic control problems is due in part to connections with H_∞ or robust control problems and dynamic games. The solution of a risk-sensitive problem leads to a conservative optimal policy, corresponding to the controller's aversion to risk.

For linear/quadratic risk-sensitive problems with full state information, Jacobson [6] established the connection with dynamic games. The analogous nonlinear problem was studied recently, and a dynamic game is obtained as a small noise limit [3, 5, 7, 10]. A risk-neutral stochastic control problem obtains as a small risk limit [3, 7].

Whittle [9] solved the discrete-time linear/quadratic risk-sensitive stochastic control problem with incomplete state information, and characterized the solution in terms of a certainty equivalence principle. The analogous continuous-time problem was solved by Bensoussan and van Schuppen [2], where the problem was converted to an equivalent one with full state information. The general nonlinear exponential cost with incomplete information in discrete time is treated in [8].

Recently Bensoussan and Elliott [1] discussed a finite-dimensional risk sensitive control problem with quasi-linear dynamics and quadratic cost; the results of this paper are a discrete-time version of [1].

2. Dynamics

Suppose $\{\Omega, \mathscr{F}, P\}$ is a probability space with a complete filtration $\{\mathscr{F}_k\}$, $k \in \mathbb{N}$, on which are given two sequences of independent random variables x_k and y_k, having normal densities $\psi_k = \mathrm{N}(0, \Sigma_k)$ and $\phi_k = \mathrm{N}(0, \Gamma_k)$ where Σ_k and Γ_k are $n \times n$ and $m \times m$ positive definite matrices for all $k \in \mathbb{N}$.

* Corresponding author.

Elsevier Science B.V.
SSDI 0167-6911(94)00073-5

U is a nonempty subset of $\mathbb{R}^p$ and $\mathscr{Y}_k = \sigma\{y_l,\ l \leqslant k\}$ is the complete filtration generated by y. The admissible controls u are the set of U-valued $\{\mathscr{Y}_k\}$ adapted processes.

Write $U_{k,l}$ for the set of such control processes defined for $k, \ldots, l$.

Consider the following functions

$$A_k(u) : U \times \mathbb{N} \to L(\mathbb{R}^n, \mathbb{R}^n), \quad \text{(the space of } n \times n \text{ matrices)},$$

$$G_k(u) : U \times \mathbb{N} \to \mathbb{R}^n,$$

$$C_k(u) : U \times \mathbb{N} \to L(\mathbb{R}^m, \mathbb{R}^m).$$

For any control u, define $\Lambda_{0,0} = 1$ and

$$\Lambda_{0,k+1} = \prod_{l=0}^{k} \frac{\psi_{l+1}(x_{l+1} - A_l(u)x_l - G_k(u))\phi_{l+1}(y_{l+1} - C_l(u)x_{l+1}(u))}{\psi_{l+1}(x_{l+1})\phi_{l+1}(y_{l+1})}.$$

Then $\Lambda_{0,k}$ is an $\mathscr{F}_k$ martingale and $E[\Lambda_{0,k}] = 1$. A new probability measure can be defined by putting $\mathrm{d}\hat{P}/\mathrm{d}P|_{\mathscr{F}_k} = \Lambda_{0,k}$.

Define the processes $v_k = v_k^u$ and $w_k = w_k^u$,

$$v_{k+1} = x_{k+1} - A_k(u)x_k - G_k(u), \qquad w_k = y_k - C_k(u)x_k.$$

Then under, $\hat{P}, v_k$ and w_k are two sequences of independent normally distributed random variables with densities ϕ_k and ψ_k, respectively. Therefore under $\hat{P}$,

$$x_{k+1} = A_k(u)x_k + G_k(u) + v_{k+1}, \tag{1}$$

$$y_k = C_k(u)x_k + w_k. \tag{2}$$

3. Cost

Consider the following mappings

$$M(\cdot) : U \to \mathbb{R}, \qquad m(\cdot) : U \to \mathbb{R}^n,$$

$$N(\cdot) : U \to \mathbb{R}, \qquad \Phi(\cdot) : \mathbb{R}^n \to \mathbb{R}.$$

The results extend immediately to the case where M, m, and N are time dependent.

For any admissible control u and real number θ we consider the expected exponential risk sensitive cost

$$\begin{aligned} J(u) &= \theta\hat{E}\left[\exp\theta\left\{\sum_{l=0}^{T-1}[\langle M(u)x_l, x_l\rangle + \langle m(u), x_l\rangle + N(u)] + \Phi(x_T)\right\}\right] \\ &= \theta E\left[\Lambda_{0,T}\exp\theta\left\{\sum_{l=0}^{T-1}[\langle M(u)x_l, x_l\rangle + \langle m(u), x_l\rangle + N(u)] + \Phi(x_T)\right\}\right]. \end{aligned} \tag{3}$$

Write $D_{0,0} = 1$,

$$D_{0,k}^u = D_k^u = \exp\theta\left(\sum_{l=0}^{k-1}[\langle M(u)x_l, x_l\rangle + \langle m(u), x_l\rangle + N(u)]\right). \tag{4}$$

4. Finite-dimensional information states

Notation 4.1. For any admissible control u consider the measure

$$\alpha_k^u(x)\,\mathrm{d}x := E[\Lambda_k^u D_k^u I(x_k \in \mathrm{d}x) | \mathscr{Y}_k]. \tag{5}$$

Note $\alpha_0^u(x) = \psi_0(x)$.

Then α_k^u satisfies the following forward recursion

$$\alpha_{k+1}^u(x) = \frac{\phi_{k+1}(y_{k+1} - C_{k+1}(u)x)}{\phi_{k+1}(y_{k+1})} \exp\theta(\langle M(u)x,x\rangle + \langle m(u),x\rangle + N(u)) \times \int_{\mathbb{R}^n} \psi_{k+1}(x - A_{k+1}(u)z - G_{k+1}(u))\alpha_k(z)\,\mathrm{d}z. \tag{6}$$

The linearity of the dynamics and the fact that v_k and w_k are independent and normally distributed implies that $\alpha_k^u(x)$ is an unnormalized normal density which we write as

$$\alpha_k^u(x) = Z_k(u)\exp -\tfrac{1}{2}(x - \mu_k(u))'R_k^{-1}(u)(x - \mu_k(u)). \tag{7}$$

$R_k^{-1}(u)$, $\mu_k(u)$ and $Z_k(u)$ are given by the following algebraic recursions:

$$\begin{aligned} R_{k+1}^{-1}(u) &= -2\theta M(u) + C_{k+1}'(u)\Gamma_{k+1}^{-1}C_{k+1}(u) + \Sigma_{k+1}^{-1}\{I - A_k(u)a^{-1}A_k'\Sigma_{k+1}^{-1}\}, \\ \tilde{\mu}_{k+1}(u) &= \theta m(u) + C_{k+1}'(u)\Gamma_{k+1}^{-1}y_{k+1} + \Sigma_{k+1}^{-1}A_k'(u)a^{-1}\tilde{\mu}_k(u) + (I + \Sigma_{k+1}^{-1}A_k(u)a^{-1}A_k'(u))\Sigma_{k+1}^{-1}G_k(u), \\ Z_{k+1}(u) &= Z_k(u)(2\pi)^{-n/2}|\Sigma_{k+1}|^{-1/2}\exp -\tfrac{1}{2}\{b - \tilde{\mu}_{k+1}'(u)\tilde{\mu}_{k+1}(u)\}, \end{aligned} \tag{8}$$

where

$$\begin{aligned} a &= A_k'(u)\Sigma_{k+1}^{-1}A_k(u) + R_k^{-1}, \\ b &= -2\theta N(u) + G_k'(u)\Sigma_{k+1}^{-1}\{I - A_k(u)a^{-1}A_k'(u)\Sigma_{k+1}^{-1}\}G_k(u) \\ &\quad + \tilde{\mu}_k'(u)\{\tilde{\mu}_k(u) - a^{-1}\tilde{\mu}_k(u)\} - 2G_k'(u)\Sigma_{k+1}^{-1}A_k(u)a^{-1}\tilde{\mu}_k(u) \end{aligned}$$

and $\tilde{\mu}_k(u) = R^{-1}\mu_k(u)$.

The recursions for $\tilde{\mu}$, R and Z are algebraic and involve no integration.

For this partially observed stochastic control problem the information state $\alpha_k^u(x)$ (which is, in general, a measure valued process), is determined by the three finite-dimensional parameters $\mu_k(u)$, $R_k(u)$ and $Z_k(u)$. These parameters can be considered as the state ξ of the process:

$$\xi_k^u = (\mu_k(u), R_k(u), Z_k(u))$$

and we can write

$$\alpha_k^u(x) = \alpha_k(\xi_k^u, x) = Z_k(u)\exp -\tfrac{1}{2}(x - \mu_k(u))'R_k^{-1}(u)(x - \mu_k(u)).$$

For integrable $f(x)$,

$$\langle \alpha_k(\xi_k^u), f\rangle = \int_{\mathbb{R}^n} \alpha_k(\xi_k^u, x)f(x)\,\mathrm{d}x,$$

which in fact equals $E[\Lambda_{0,k}^u D_{0,k}^u f(x_k)|\mathscr{Y}_k]$.

5. A separation principle

For any admissible control u, we saw that the expected total cost is

$$\begin{aligned} J(u) &= \theta E\left[\Lambda_{0,T}^u \exp\theta\left\{\sum_{l=1}^{T}[\langle M(u)x_l,x_l\rangle + \langle m(u),x_l\rangle + N(u)] + \Phi(x_T)\right\}\right] \\ &= \theta E[\Lambda_{0,T}^u D_{0,T}^u \exp\theta\Phi(x_T)] = \theta E[E[\Lambda_{0,T}^u D_{0,T}^u \exp\theta\Phi(x_T)|\mathscr{Y}_T]] \end{aligned}$$

$$= \theta E\left[\int_{\mathbb{R}^n} \exp \theta\Phi(x)\alpha_T(x)\,\mathrm{d}x\right] = \theta E[\langle \alpha_T(\xi_T^u), \exp \theta\Phi\rangle],$$

where Φ is the terminal cost as in Eq. (3).

6. Adjoint process

Now for any k, $0 < k < T$,

$$J(u) = \theta E[\Lambda_{0,k}^u \Lambda_{k+1,T}^u D_{0,k}^u D_{k+1,T}^u \exp \theta\Phi(x_T)]$$
$$= \theta E[\Lambda_{0,k}^u D_{0,k}^u E[\Lambda_{k+1,T}^u D_{k+1,T}^u \exp \theta\Phi(x_T) \mid x_0, \ldots, x_k, \mathcal{Y}_T]].$$

Write $\beta_k^u(x_k) = E[\Lambda_{k+1,T}^u D_{k+1,T}^u \exp \theta\Phi(x_T) | x_k, \mathcal{Y}_T]$ where, using the Markov property of x, the conditioning involves only x_k. Note that $\beta_T(x_T) = \exp \Phi(x_T)$. Therefore

$$J(u) = \theta E[\Lambda_{0,k}^u D_{0,k}^u \beta_k^u(x_k)]$$
$$= \theta E[E[\Lambda_{0,k}^u D_{0,k}^u \beta_k^u(x_k) | \mathcal{Y}_k]]$$
$$= \theta E[\langle \alpha_k(\xi_k^u), \beta_k^u\rangle].$$

Note this decomposition is independent of k, so

$$J(u) = \theta E[\langle \psi_0, \beta_0^u\rangle] = \theta E[\langle \alpha_T(\xi_T^u), \exp \theta\Phi\rangle].$$

Lemma 6.1. *We have the following backward recursion for the process* β:

$$\beta_k^u(x_k) = \int_{\mathbb{R}^n} \frac{\phi_{k+1}(y_{k+1} - C_{k+1}(u)x)}{\phi_{k+1}(y_{k+1})} \psi_{k+1}(x - A_k(u)x_k - G_k(u))$$
$$\times \exp \theta\{x'M(u)x + \langle m(u), x\rangle + N(u)\}\beta_{k+1}^u(x)\,\mathrm{d}x. \tag{9}$$

Proof.

$$\beta_k^u(x_k) = E[\Lambda_{k+1,T}^u D_{k+1,T}^u \exp \theta\Phi(x_T) | x_k, \mathcal{Y}_T]$$
$$= E\left[E\left[\frac{\phi_{k+1}(y_{k+1} - C_{k+1}(u)x_{k+1})\psi_{k+1}(x_{k+1} - A_k(u)x_k - G_k(u))}{\phi_{k+1}(y_{k+1})\psi_{k+1}(x_{k+1})}\right.\right.$$
$$\left.\left.\times \exp \theta\{x_{k+1}'M(u)x_{k+1} + \langle m(u), x_{k+1}\rangle + N(u)\}\Lambda_{k+2,T}^u D_{k+2,T}^u \exp \theta\Phi(x_T) \mid x_k, x_{k+1}, \mathcal{Y}_T\right] |x_k, \mathcal{Y}_T\right]$$
$$= E\left[\frac{\phi_{k+1}(y_{k+1} - C_{k+1}(u)x_{k+1})\psi_{k+1}(x_{k+1} - A_k(u)x_k - G_k(u))}{\phi_{k+1}(y_{k+1})\psi_{k+1}(x_{k+1})}\right.$$
$$\left.\times \exp \theta\{x_{k+1}'M(u)x_{k+1} + \langle m(u), x_{k+1}\rangle + N(u)\}\beta_{k+1}^u(x_{k+1}) | x_k, \mathcal{Y}_T\right].$$

Integrating with respect to the density of x_{k+1} and using the independence assumption under P gives the result. □

Consider the unnormalized Gaussian densities $\tilde{\beta}_k^u(x, z)$ given by

$$\tilde{\beta}_k^u(x, z) = \tilde{Z}_k^u \exp -\tfrac{1}{2}(x - \gamma_k(z))'S_k^{-1}((x - \gamma_k(z)).$$

We put $\tilde{\beta}^u_T(x,z) = \delta(x-z)$, $S_T = 0$ and $\gamma_T = z$. Then γ_k, S_k^{-1} and $\tilde{Z}_k$ are given by the following backward recursions.

$$\begin{aligned} S_k^{-1} &= A_k'[\Sigma_{k+1}^{-1} - \Sigma_{k+1}^{-1}a^{-1}\Sigma_{k+1}^{-1}]A_k \\ \tilde{\gamma}_k &= A_k'\Sigma_{k+1}^{-1}a^{-1}[C_{k+1}'\Gamma_{k+1}^{-1}y_{k+1} + \Sigma_{k+1}^{-1}G_k + \theta m + \tilde{\gamma}_{k+1} - aG_k] \\ \tilde{Z}_k &= \tilde{Z}_{k+1}|\Sigma_{k+1}|^{-1/2}|a|^{1/2}\exp -\tfrac{1}{2}\{b - \tilde{\gamma}_k'\gamma_k\}. \end{aligned} \tag{10}$$

Here

$$\begin{aligned} a &= C_{k+1}'\Gamma_{k+1}^{-1}C + \Sigma_{k+1}^{-1} + S_{k+1}^{-1} \\ b &= 2y_{k+1}'\Gamma_{k+1}^{-1}C_{k+1}a^{-1}[C_{k+1}'\Gamma_{k+1}^{-1}y_{k+1} + \Sigma_{k+1}^{-1}G_k + \tilde{\gamma}_{k+1} + \theta m] \\ &\quad +2\tilde{\gamma}_k'a^{-1}[\Sigma_{k+1}^{-1}G_k + \tilde{\gamma}_{k+1} + \theta m] + G_k'[\Sigma_{k+1}^{-1} + \Sigma_{k+1}^{-1}a^{-1}\Sigma_{k+1}^{-1}]G_k \\ &\quad +\theta^2 m'a^{-1}m - 2\theta N, \\ \tilde{\gamma}_k &= S_k^{-1}\gamma_k. \end{aligned}$$

Furthermore,

$$\beta_k^u(x) = \int_{\mathbb{R}^n} \tilde{\beta}_k^u(x,z)\exp\theta\Phi(z)\,dz.$$

Remark 6.2. β_k^u is the adjoint process and again it is determined by the finite-dimensional parameters γ, S and Z which satisfy the reverse-time, algebraic, recursions (10).

7. Dynamic programming

We have noted that the information state $\alpha_k(x)$ is determined by the finite-dimensional parameters

$$\xi_k^u = (\mu_k(u), R_k(u), Z_k(u)).$$

Given ξ_k^u a control u_k and the new observation y_{k+1}, Eqs. (8) determine the next value

$$\xi_{k+1}^u = \xi_{k+1}(\xi_k^u, u_k, y_{k+1}).$$

Suppose at some intermediate time k, $0 < k < T$, the information state ξ_k is $\xi = (\mu, R, Z)$.

The value function for this control problem is

$$V(\xi,k) = \inf_{u\in U_{k,T-1}} E[\langle\alpha_k^u, \beta_k^u\rangle \,|\, \alpha_k = \alpha_k(\xi)].$$

Theorem 7.1. *The value function satisfies the recursion*:

$$V(\xi,k) = \inf_{u\in U_{k,k}} E[V(\xi_{k+1}(\xi,u,y_{k+1}),k+1)] \tag{11}$$

and $V(\xi,T) = \langle\alpha_T(\xi), \exp\theta\Phi\rangle$.

Proof. Now

$$V(\xi,k) = \inf_{u\in U_{k,T-1}} E[\langle\alpha_k^u(\xi), \beta_k^u\rangle \,|\, \xi_k = \xi].$$

From (10) β_k is given by a backward recursion from β_{k+1}, that is we can write $\beta_k = \beta_k^u(\beta_{k+1}^v)$. Therefore

$$\begin{aligned} V(\xi,k) &= \inf_{u\in U_{k,k}} \inf_{v\in U_{k+1,T-1}} E[\langle \alpha_k^u(\xi), \beta_k^u(\beta_{k+1}^v)\rangle \mid \xi_k = \xi] \\ &= \inf_{u\in U_{k,k}} \inf_{v\in U_{k+1,T-1}} E[E[\langle \alpha_{k+1}(\xi_{k+1}^u), \beta_{k+1}^v\rangle \mid \mathcal{Y}_{k+1}, \xi_k = \xi] \mid \xi_k = \xi] \\ &= \inf_{u\in U_{k,k}} E\left[\inf_{v\in U_{k+1,T-1}} E[\langle \alpha_{k+1}(\xi_{k+1}^u), \beta_{k+1}^v\rangle \mid \mathcal{Y}_{k+1}, \xi_k = \xi] \mid \xi_k = \xi \right] \\ &= \inf_{u\in U_{k,k}} E[V(\xi_{k+1}(\xi, u, y_{k+1}), k+1)]. \end{aligned}$$

The interchange of conditional expectation and minimization is justified by application of the lattice property of the controls (see [4]). □

Write $U_{k,l}^s$ for the set of control processes on the time interval $k,\dots,l$ which are adapted to the filtration $\sigma\{\xi_j : k \leqslant j \leqslant l\}$. We call such controls separated.

Theorem 7.2 (Verification). *Suppose $u^* \in U_{0,T-1}^s$ is a control which for each $k = 0,\dots,T-1$, $u_k^*(\xi_k)$ achieves the minimum in* (11)*. Then $u^* \in U_{0,T-1}$ and is an optimal control.*

Proof. Write

$$\bar{V}(\xi,k,u) = E[\langle \alpha_k^u(\xi), \beta_k^u\rangle \mid \xi_k = \xi].$$

We shall show that

$$V(\xi,k) = \bar{V}(\xi,k,u^*), \quad \text{for each } k = 0,\dots,T. \tag{12}$$

For $k = T$ (12) is clearly satisfied. Suppose (12) is true for $k+1,\dots,T$. Then

$$\begin{aligned} \bar{V}(\xi,k,u^*) &= E[E[\langle \alpha_{k+1}(\xi_{k+1}^{u^*}), \beta_{k+1}^{u^*}\rangle \mid \xi_k = \xi, \mathcal{Y}_{k+1}] \mid \xi_k = \xi] \\ &= E[\bar{V}(\xi_{k+1}^{u^*}(\xi), k+1), u^* \in U_{k+1,T-1})] \\ &= E[V(\xi_{k+1}^{u^*}(\xi), k+1)] \\ &= V(\xi,k). \end{aligned}$$

This gives (12).

Putting $k = 0$ we see

$$\bar{V}(\xi,0,u^*) = V(\xi,0) \leqslant \bar{V}(\xi,0,u)$$

for any $u \in U_{0,T-1}$. That is, u^* is an optimal control.

Remark 7.3. This result shows the optimal policy u^* for the risk sensitive control problem is a separated policy, in that it is a function of the information state ξ.

Acknowledgments

The authors wish to acknowledge support of the Natural Sciences and Engineering Research Council of Canada, grant A7964, INRIA, Rocquencourt France and the hospitality of the Department of Systems Engineering at the Australian National University.

References

[1] A. Bensoussan and R.J. Elliott, A finite dimensional risk sensitive control problem, to appear in *SIAM J. Control Optim.*

[2] A. Bensoussan and J.H. van Schuppen, Optimal control of partially observable stochastic systems with an exponential-of-integral performance index, *SIAM J. Control Optim.* **23** (1985) 599–613.

[3] C. Campi and M.R. James, Discrete-time nonlinear risk-sensitive control, Preprint, 1992.

[4] R.J. Elliott, Stochastic calculus and applications, in: *Applications of Mathematics*, Vol. 18 (Springer, Berlin, 1982).

[5] W.H. Fleming and W.M. McEneaney, Risk sensitive optimal control and differential games, in: *Proc. Conf. on Adaptive and Stochastic Control*, Univ. of Kansas (Springer, Berlin, 1991).

[6] D.H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Trans. Automat. Control* **AC-18** (2) (1973) 124–131.

[7] M.R. James, Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games, *Math. Control Signals Systems* **5** (4) (1992) 401–417.

[8] M.R. James, J. Baras and R.J. Elliott, Risk sensitive control and dynamic games for partially observed discrete-time nonlinear systems, *IEEE Trans. Automat. Control* **AC-39** (1994) 399–407

[9] P. Whittle, Risk-sensitive linear/quadratic/Gaussian control, *Adv. Appl. Probab.* **13** (1981) 764–777.

[10] P. Whittle, A risk-sensitive maximum principle, *Systems Control Lett.* **15** (1990) 183–192.