Partially observed risk-sensitive control of path-dependent discrete-time systems

Hidehiro Kaise
with S. Tomiyama (Aioi Nissay Dowa Insurance)

Graduate School of Engineering Science and MMDS, Osaka University, Japan

IMA, Minneapolis, May 7, 2018

Background and aim

Recently there is large attention to control problems which cannot be directly treated in a classical framework based on Markovian structures. One of such problems is *path-dependent (PD) system*:

- continuous-time PD deterministic system: Lukoyanov '00, '07 (classical)
- continuous-time PD stochastic system: Hamadène & Lepeltier '95 (classical), El-Karoui & Hamadène '03 (risk-sensitive) cf. Ren, Touzi & Zhang '14, Zhang '17 (PD PDEs of 2nd order)

Aim: Focusing on discrete-time settings, we will seek for possibility on studying partially observed risk-sensitive control with path-dependent dynamics and costs.

cf. Partially observed risk-sensitive control for Markovian systems

- continuous-time: Bensoussan & van Schuppen '85
- discrete-time: James, Baras & Elliott '94

Outline

Key ideas for Markovian case (cf. Bensoussan & van-Schuppen '85, James, Baras & Elliott '94):

- Measure change by Girsanov transformation
- Reduction to control of unnormalized conditional density (information state)

We will extend arguments of James-Baras-Elliott '94 to path-dependent systems under dynamic programming with time-varying state spaces.

1. Problem formulation
2. Discrete-time Girsanov transformation and unnormalized conditional distribution of state paths
3. Reduction to completely observed stochastic control by information state: stochastic control with time-varying infinite-dimensional spaces
4. Small noise limit: dynamic game under partial observation (if we have time)

Path-dependent stochastic system

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Consider the following *path-dependent stochastic system*: Let initial distribution of X_0 be given. For $k = 0, 1, \ldots, M-1$

$$X_{k+1} = F_k(X_{0,k}, U_k) + W_{k+1}$$
$$Y_{k+1} = H_k(X_{0,k}) + V_{k+1}$$

X_k; $\mathbb{R}^n$-valued r.v. (state) Y_k; $\mathbb{R}^m$-valued r.v. (observation)
W_k; $\mathbb{R}^n$-valued r.v. (system noise) V_k; $\mathbb{R}^m$-valued r.v. (observation noise)
U_k; U-valued r.v. (control) (U; Borel subset of Euclidean space)

Notations: $X_{0,k} := (X_0, X_1, \ldots, X_k)$, $Y_{0,k} := (Y_0, Y_1, \ldots, Y_k)$
$F_k : (\mathbb{R}^n)^{k+1} \times U \to \mathbb{R}^n$, $H_k : (\mathbb{R}^m)^{k+1} \to \mathbb{R}^m$
Remark: (i) F_k and H_k depend on path $X_{0,k} = (X_1, X_2, \ldots, X_k)$.
(ii) [Markovian case] $F_k(X_{0,k}, U_k) = f_k(X_k, U_k)$, $H_k(X_{0,k}) = h_k(X_k)$.

Path-dependent RS control under partial observations

We consider path-dependent risk-sensitive criterion: For $u = \{U_k\}_{k=0}^{M-1}$,

$$J(u) = E^P\left[e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\right].$$

$L_k : (\mathbb{R}^n)^{k+1} \times U \to \mathbb{R}$ (running cost), $\Phi : (\mathbb{R}^n)^{M+1} \to \mathbb{R}$ (terminal cost)
$\mu \in \mathbb{R}\backslash\{0\}$; risk-sensitive parameter

- We suppose X_k cannot be directly observed but $Y_1, Y_2, \ldots, Y_k$ are available. Thus $u = \{U_k\}_{k=0}^{M-1}$ is adapted to $\mathcal{Y}_k = \sigma(Y_1, Y_2, \ldots, Y_k)$.

$\mathcal{U}_{0,M-1}$ denotes the set of U-valued $\mathcal{Y}_k$-adapted processes.

Remark: Note that U_k is measurable with respect to $\sigma(Y_1, Y_2, \ldots, Y_k)$ if and only if $U_k = h_k(Y_{1,k})$ for some Borel measurable function $h_k : (\mathbb{R}^m)^{k+1} \to U$. We often identify $u = \{U_k\}_{k=0}^{M-1}$ with $h = \{h_k\}_{k=0}^{M-1}$.

Partially observed risk-sensitive control of path-dependent system

Minimize $J(u)$ on $\mathcal{U}_{0,M-1}$ and design filter to calculate optimal control

Assumptions

(A1) X_0 has a density $\rho : \mathbb{R}^n \to [0, \infty)$.

(A2) X_0 and $\{(W_k, V_k)\}_{k=1}^M$ are independent.

(A3) $\{(W_k, V_k)\}_{k=1}^M$ are independently and identically distributed. (W_1, V_1) has bounded uniformly continuous density $\phi : \mathbb{R}^n \times \mathbb{R}^m \to (0, \infty)$.

(A4) U is compact .

(A5) $F_k : (\mathbb{R}^n)^{k+1} \times U \to \mathbb{R}^n$, $H_k : (\mathbb{R}^m)^{k+1} \to \mathbb{R}^m$ are bounded uniformly continuous.

(A6) $L_k : (\mathbb{R}^n)^{k+1} \times U \to \mathbb{R}$, $\Phi : (\mathbb{R}^n)^{M+1} \to \mathbb{R}$ are bounded uniformly continuous non-negative functions.

Remark: W_k and V_k can be correlated.

Discrete-time Girsanov transformation

We define sub σ-fields for the subsequent arguments:

$$\mathcal{G}_0 := \sigma(X_0), \mathcal{G}_k := \sigma(X_0, W_{0,k}, V_{0,k}) \ (k = 1, 2, \ldots, M),$$
$$\mathcal{Y}_k = \sigma(Y_{1,k}) \ (k = 1, 2, \ldots, M).$$

Suppose $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$ is given. Let $\{\Gamma_k\}_{k=0}^{M}$ be given by

$$\Gamma_{k+1} = \Gamma_k \gamma_{k+1} \ (k = 0, 1, \ldots, M-1), \ \Gamma_0 = 1,$$
$$\text{where } \gamma_k = \phi(W_k, Y_k)/\phi(W_k, V_k).$$

We define measure Q on $\mathcal{G}_M$ by

$$\left.\frac{dQ}{dP}\right|_{\mathcal{G}_M} = \Gamma_M.$$

Discrete-time Girsanov transformation (cont'd)

Proposition (cf. Elliott-Moore '93 for Markovian system under uncorrelated noises)

(i) $\{\Gamma_k\}_{k=0}^M$ *is a* $\{\mathcal{G}_k\}$*-martingale under* P.
(ii) Q *is a probability measure.*
(iii) $Q(X_0 \in dx) = \rho(x)dx$.
(iv) (W_{k+1}, Y_{k+1}) *is independent of* $\mathcal{G}_k$ *under* Q. *Also,* (W_1, Y_1) *has density* ϕ *under* Q.

Remark: If W_k and V_k are independent and V_k is normal distribution with mean 0 and covariance matrix $\hat{\Sigma}$, we have

$$\left.\frac{dQ}{dP}\right|_{\mathcal{G}_M} = e^{-\sum_{k=0}^M H_k(X_{0,k})^T \hat{\Sigma}^{-1} Y_k + \sum_{k=0}^M H_k(X_{0,k})^T \hat{\Sigma}^{-1} H_k(X_{0,k})},$$

Q is a probability measure, Y_k is independent of $\mathcal{G}_k$ under Q and Y_k is normal distribution with mean 0 and covariance matrix $\hat{\Sigma}$ under Q.

Risk-sensitive criterion under Q

For $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$, recall the risk-sensitive criterion

$$J(h) = E^P\left[e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\right]$$

where $U_k = h_k(Y_{1,k})$. Note that

$$J(h) = E^P\left[E^P\left[e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\middle|\mathcal{Y}_M\right]\right]. \qquad (*)$$

By Bayes' rule, we have

$$\begin{aligned} &E^P\left[e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\middle|\mathcal{Y}_M\right] \\ &= E^Q\left[\Gamma_M^\dagger e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\middle|\mathcal{Y}_M\right]/E^Q\left[\Gamma_M^\dagger\middle|\mathcal{Y}_M\right], \end{aligned}$$

where $\Gamma_M^\dagger = \Gamma_M^{-1}$. Plugging this back in $(*)$, we have

$$J(h) = E^Q\left[E^Q\left[\Gamma_M^\dagger e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\middle|\mathcal{Y}_M\right]\right].$$

Risk-sensitive criterion under Q (cont'd)

Risk-sensitive criterion under Q

$$J(h) = E^Q\big[E^Q\big[\Gamma^\dagger_M e^{\mu\{\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)+\Phi(X_{0,M})\}}\big|\mathcal{Y}_M\big]\big]$$
$$= E^Q\big[E^Q\big[\Gamma^\dagger_M e^{\mu\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)}e^{\mu\Phi(X_{0,M})}\big|\mathcal{Y}_M\big]\big].$$

Note that $E^Q[\cdots|\mathcal{Y}_M]$ can be calculated if $Y_0, Y_1, \cdots Y_M$ are observed.

Key: Under path-dependent systems, it may be reasonable to infer not only state X_M but also path of states $X_{0,M} = (X_0, \ldots, X_M)$.

- We want to find unnormalized conditional density $\sigma^h_M(x_{0,M}) = \sigma^h_M(x_{0,M}, Y_{1,M})$ of $X_{0,M}$ given $Y_{1,M}$: For any *test function* $\varphi : (\mathbb{R}^n)^{M+1} \to \mathbb{R}$,

$$E^Q\big[\Gamma^\dagger_M e^{\mu\sum_{k=0}^{M-1} L_k(X_{0,k},U_k)}\varphi(X_{0,M})\big|\mathcal{Y}_M\big] = \int \varphi(x_{0,M})\sigma^h_M(x_{0,M}, Y_{1,M})dx_{0,M}$$

where $x_{0,M} = (x_0, x_1, \ldots, x_M) \in (\mathbb{R}^n)^{M+1}$, $dx_{0,M} = dx_0 dx_1 \cdots dx_M$.

Dynamics of unnormalized conditional density

Let $k = 0, 1, \ldots, M-1$. Given $u \in U$ and $y \in \mathbb{R}^m$, define operator $\Sigma_k[u, y] : L^1((\mathbb{R}^n)^{k+1}) \to L^1((\mathbb{R}^{k+2}))$ as follows: for $\sigma \in L^1((\mathbb{R}^n)^{k+1})$,

$$(\Sigma_k[u, y]\sigma)(x_{0,k+1}) = e^{\mu L_k(x_{0,k},u)} \frac{\phi(x_{k+1} - F_k(x_{0,k}, u), y - H_k(x_{0,k}))}{\hat{\phi}(y)} \sigma(x_{0,k})$$

with $x_{0,k+1} = (x_{0,k}, x_{k+1}) \in (\mathbb{R}^n)^{k+1} \times \mathbb{R}^n$. Here $\hat{\phi}(y) = \int \phi(w, y) dw$.

Given control $h = \{h_k\}_{k=0}^{M-1}$, let $Y_1, Y_2, \ldots, Y_M$ be the observation process corresponding to h. We define $\sigma_k^h(\omega) = \sigma_k^h(\cdot, Y_{1,k}(\omega)) \in L^1((\mathbb{R}^n)^{k+1})$ by

$$\begin{aligned} \sigma_{k+1}^h(\omega) &= \Sigma_k[U_k(\omega), Y_{k+1}(\omega)]\sigma_k^h(\omega) \ (k = 0, 1, \ldots, M-1), \\ \sigma_0^h(\omega) &= \rho \end{aligned} \quad \text{(I)}$$

where $U_k = h_k(Y_{1,k})$.

Remark: σ_k^h is $L^1((\mathbb{R}^n)^{k+1})$-valued random variable.

Dynamics of unnormalized conditional density (cont'd)

Notation: $\langle \sigma, \varphi \rangle_k := \int \sigma(x_{0,k})\varphi(x_{0,k})dx_{0,k}$ for $\sigma \in L^1((\mathbb{R}^n)^{k+1})$, $\varphi \in L^\infty((\mathbb{R}^n)^{k+1})$.

Theorem

Let $k = 0, 1, \ldots, M$. For any $\varphi \in L^\infty((\mathbb{R}^n)^{k+1})$,

$$E^Q\left[\Gamma_k^\dagger e^{\mu \sum_{i=0}^{k-1} L_i(X_{0,i}, h_i(Y_{1,i}))}\varphi(X_{0,k})\middle|\mathcal{Y}_k\right] = \langle \sigma_k^h, \varphi \rangle_k,$$

where $\sigma_k^h(\omega) = \sigma_k^h(x_{0,k}, Y_{1,k}(\omega))$ is given by (I).

Corollary

For any control $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$

$$J(h) = E^Q[\langle \sigma_M^h, e^{\mu\Phi} \rangle_M].$$

Remark: $J(h)$ is a criterion of stochastic control with terminal cost $\langle \cdot, e^{\mu\Phi} \rangle_M : L^1((\mathbb{R}^n)^{M+1}) \to \mathbb{R}$ under new state dynamics (I).
We call $\sigma_k^h (k = 0, 1, \ldots, M)$ *information state* for risk-sensitive control (cf. James-Baras-Elliott '94).

Reduction to completely observable stochastic control

As seen in the previous slides, we can rewrite the risk-sensitive control under Q: Let $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$.

Dynamics of information state

$$\begin{aligned} \sigma_{k+1}^h(\omega) &= \Sigma_k[U_k(\omega), Y_{k+1}(\omega)]\sigma_k^h(\omega) \ (k = 0, 1, \dots, M-1), \\ \sigma_0^h(\omega) &= \rho \end{aligned} \tag{I}$$

where $U_k = h_k(Y_{1,k})$.

Criterion

$$J(h) = E^Q[\langle \sigma_M^h, e^{\mu\Phi} \rangle_M].$$

Remark: (i) Recall that (W_{k+1}, Y_{k+1}) is independent of $\mathcal{G}_k = \sigma(X_0, W_{1,k}, Y_{1,k})$ and (W_k, Y_k) has density ϕ under Q. Thus $\{Y_k\}_{k=1}^M$ may serve as driving noise in (I).
(ii) $u = \{U_k\}_{k=0}^{M-1}$ is adapted to $\mathcal{Y}_k$ generated by $Y_1, Y_2, \dots, Y_k$.

Reformulation by completely observable problem

Let $(\Omega, \mathcal{F}, Q)$ be a probability space. $\{Y_k\}_{k=1}^M$ is a sequence of i.i.d. random variables with density $\hat{\phi}(y) = \int \phi(w, y) dw$.

Let $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$ be given.

Dynamics of new state: Define $L^1((\mathbb{R}^n)^{k+1})$-valued random variable σ_k^h by

$$\sigma_{k+1}^h(\omega) = \Sigma_k[U_k(\omega), Y_{k+1}(\omega)]\sigma_k^h(\omega) \ (k = 0, 1, \ldots, M-1),$$
$$\sigma_0^h(\omega) = \rho,$$

where $U_k(\omega) = h_k(Y_{1,k}(\omega))$.

Criterion

$$J(h) = E^Q[\langle \sigma_M^h, e^{\mu\Phi} \rangle_M].$$

Reformulation to completely observable problem

Minimize $J(h)$ over $h = \{h_k\}_{k=0}^{M-1} \in \mathcal{U}_{0,M-1}$.

Remark: Note that σ_k^h is $L^1((\mathbb{R}^n)^{k+1})$-valued r.v.. It can be considered a completely observable stochastic control with time-varying state space.

Problem formulation for dynamic programming

Let $(\Omega, \mathcal{F}, Q)$ be a probability space and $\{Y_k\}_{k=1}^M$ be i.i.d. random variables with $Q(Y_1 \in dy) = \hat{\phi}(y)dy$ where $\hat{\phi}(y) = \int \phi(w,y)dw$

Given $k = 0, 1, \ldots, M-1$, consider a problem starting at time k.

Admissible controls: U-valued process $u = \{U_i\}_{i=k}^{M-1}$ is *admissible* if U_i is $\sigma(Y_{k+1,i})$-measurable where $Y_{k+1,i} = (Y_{k+1}, Y_{k+2}, \ldots, Y_i)$. We understand $\sigma(Y_{k+1,k}) = \{\emptyset, \Omega\}$.

$\mathcal{U}_{k,M-1}$ denotes the set of admissible controls.

Remark: We often identify $u = \{U_i\}_{i=k}^{M-1}$ with $h = \{h_i\}_{i=k}^{M-1}$ where $h_i : (\mathbb{R}^m)^{i-k} \to U$ is Borel measurable via relation $U_i = h_i(Y_{k+1,i})$.

Problem formulation for dynamic programming (cont'd)

Let $u = \{U_i\}_{i=k}^{M-1} \in \mathcal{U}_{k,M-1}$ be given.
System dynamics: Given $\sigma \in L^1((\mathbb{R}^n)^{k+1})$, we define $L^1((\mathbb{R}^n)^{i+1})$-valued random variable σ_i^u $(i = k, k+1, \dots, M)$ by

$$\sigma_{i+1}^u(\omega) = \Sigma_i[U_i(\omega), Y_{i+1}(\omega)]\sigma_i^u(\omega) \ (i = k, k+1, \dots, M-1),$$
$$\sigma_k^u(\omega) = \sigma.$$

Criterion:

$$J_k(\sigma; u) := E^Q[\langle \sigma_M^u, e^{\mu\Phi} \rangle_M].$$

Value function:

$$V_k(\sigma) := \inf_{u \in \mathcal{U}_{k,M-1}} J_k(\sigma; u), \ \sigma \in L^1((\mathbb{R}^n)^{k+1}).$$

Remark: Note that state space at k is $L^1((\mathbb{R}^n)^{k+1})$ which depends on k. The above problem can be discussed by the dynamic programming with time-varying state spaces.

Verification Theorem

Theorem

Define $W_k : L^1((\mathbb{R}^n)^{k+1}) \to \mathbb{R}$ $(k = 0, 1, \dots, M)$ *by*

$$W_M(\sigma) = \langle \sigma, e^{\mu\Phi} \rangle_M, \ \sigma \in L^1((\mathbb{R}^n)^{M+1}),$$

$$W_k(\sigma) = \inf_{u \in U} E^Q[W_{k+1}(\Sigma_k[u, Y_{k+1}]\sigma)], \ \sigma \in L^1((\mathbb{R}^n)^{k+1}),$$

$$k = M-1, M-2, \dots, 0.$$

Then (i)–(iii) *hold:*
(i) *For any* $u \in \mathcal{U}_{k,M-1}$, $J_k(\sigma; u) \geq W_k(\sigma), \forall \sigma \in L^1((\mathbb{R}^n)^{k+1})$.
(ii) *For each* $k = 0, 1, \dots, M-1$, *there exists a Borel measurable function* $h_k^* : L^1((\mathbb{R}^n)^{k+1}) \to U$ *such that*

$$W_k(\sigma) = E^Q[W_{k+1}(\Sigma_k(h_k^*(\sigma), Y_{k+1})\sigma)], \ \forall \sigma \in L^1((\mathbb{R}^n)^{k+1}).$$

Verification Theorem (cont'd)

Theorem (cont'd)

(iii) *Consider* $L^1((\mathbb{R}^n)^{k+1})$*-valued r.v.* σ_i^* $(i = k, k+1, \ldots, M-1)$ *by*

$$\begin{aligned} \sigma_{i+1}^*(\omega) &= \Sigma_i[h_i^*(\sigma_i^*(\omega)), Y_{i+1}(\omega)]\sigma_i^*(\omega) \ (i = k, k+1, \ldots, M-1), \\ \sigma_k^*(\omega) &= \sigma. \end{aligned} \tag{I*}$$

Set $u^* = \{U_i^*\}_{i=k}^{M-1}$ *with* $U_i^* = h_i^*(\sigma_i^*)$*. Then* $u^* \in \mathcal{U}_{k,M-1}$ *and* $J_k(\sigma; u^*) = W_k(\sigma)$*, i.e.* u^* *is an optimal control.*

Remark: $U_i^* = h_i(\sigma_i^*)$ is a feedback control of information state.
[Offline] a) Calculate $h^* = \{h_k^*\}_{k=0}^{M-1}$ given in Theorem (ii).
b) For any $y_{1,M} = (y_1, y_2, \ldots, y_M) \in (\mathbb{R}^m)^M$, calculate $\sigma_k^* = \sigma_k^*(\cdot, y_{1,k}) \in L^1((\mathbb{R}^n)^{k+1})$ $(k = 0, 1, \ldots, M-1)$ by

$$\sigma_{k+1}^* = \Sigma_k[h_k^*(\sigma_k^*), y_{k+1}]\sigma_k^* \ (i = 0, 1, \ldots, M-1), \ \sigma_0^* = \rho.$$

[Online] Under observation $Y_{1,k} = (Y_1, Y_2, \ldots, Y_k)$, calculate $u_k^* := \sigma_i^*(\cdot, Y_{1,i})$.

Small noise limit: Settings

Let $\epsilon > 0$ denote intensity parameter of noises. We assume (A3)' holds instead of (A3):

(A3)' $\{(W_k, V_k)\}_{k=1}^{M}$ are i.i.d. sequence of random variables. (W_1, V_1) is $n \times m$-dimensional normal distribution of mean 0 and covariance matrix $\epsilon\Sigma$ and V_1 is m-dimensional normal distribution of mean 0 and covariance matrix $\epsilon\hat{\Sigma}$.

We denote (W_k, V_k) by $(W_k^\epsilon, V_k^\epsilon)$ to specify the dependence on ϵ.

State and observation: Suppose $P(X_0 \in dx) = c_\epsilon e^{p(x)/\epsilon} dx$ (c_ϵ;normalizing constant). For $k = 0, 1, \dots, M-1$

$$X_{k+1}^\epsilon = F_k(X_{0,k}^\epsilon, U_k) + W_{k+1}^\epsilon$$

$$Y_{k+1}^\epsilon = H_k(X_{0,k}^\epsilon) + V_{k+1}^\epsilon,$$

where $u = \{U_k\}_{k=0}^{M-1}$ is U-valued process adapted to $\mathcal{Y}_k = \sigma(Y_{1,k}^\epsilon)$.

Criterion:

$$J^\epsilon(u) = E^P\big[e^{(\mu/\epsilon)\{\sum_{k=0}^{M-1} L_k(X_{0,k}^\epsilon, U_k) + \Phi(X_{0,M}^\epsilon)\}}\big].$$

[Q] Asymptotic behavior of $V^\epsilon := \inf_u J^\epsilon(u)$ as $\epsilon \to 0$?

Small noise limit: Notations

$G := \{\gamma = (\gamma_1, \gamma_2);\ \gamma_1 > 0, \gamma \geq 0\}$. For $\gamma = (\gamma_1, \gamma_2) \in G$,

$$\mathcal{D}^\gamma(\mathbb{R}^d) := \{p \in C(\mathbb{R}^d);\ p(x) \leq -\gamma_1 |x|^2 + \gamma_2\},$$
$$\mathcal{D}(\mathbb{R}^d) := \bigcup_{\gamma \in G} \mathcal{D}^\gamma(\mathbb{R}^d).$$

We suppose $\mathcal{D}(\mathbb{R}^d)$ has a topology induced by uniform convergence on each compact sets and denote by d a metric compatible with this topology.

Dynamic programing equation with small noise

To study the asypmptoics of V^ϵ, we use dynamic programming equation. We recall the reformulation of risk-sensitive control.

Let $(\Omega, \mathcal{F}, Q)$ be a probability space. $\{Y_k^\epsilon\}_{k=1}^M$ is a sequence of i.i.d. normal random variables with mean 0 and covariance $\epsilon\hat{\Sigma}$.

Let $u = \{U_i\}_{i=k}^{M-1} \in \mathcal{U}_{k,M-1}$ be given.

System dynamics: Given $\sigma \in L^1((\mathbb{R}^n)^{k+1})$, we define $L^1((\mathbb{R}^n)^{i+1})$-valued random variable σ_i^u $(i = k, k+1, \dots, M)$ by

$$\sigma_{i+1}^u(\omega) = \Sigma_i[U_i(\omega), Y_{i+1}^\epsilon(\omega)]\sigma_i^u(\omega) \ (i = k, k+1, \dots, M-1),$$
$$\sigma_k^u(\omega) = \sigma.$$

Criterion:

$$J_k^\epsilon(\sigma; u) := E^Q[\langle \sigma_M^u, e^{(\mu/\epsilon)\Phi} \rangle_M].$$

Value function:

$$V_k^\epsilon(\sigma) := \inf_{u \in \mathcal{U}_{k,M-1}} J_k^\epsilon(\sigma; u).$$

Dynamic programing equtation with small noise (cont'd)

By the Verification theorem, $V_k^\epsilon : L^1((\mathbb{R}^n)^{k+1}) \to \mathbb{R}$ $(k = 0, 1, \ldots, M)$ satisfy the dynamic programming equation:

$$V_M^\epsilon(\sigma) = \langle \sigma, e^{(\mu/\epsilon)\Phi} \rangle_M, \ \sigma \in L^1((\mathbb{R}^n)^{M+1}),$$

$$V_k^\epsilon(\sigma) = \inf_{u \in U} E^Q[V_{k+1}^\epsilon(\Sigma_k[u, Y_{k+1}^\epsilon]\sigma)], \ \sigma \in L^1((\mathbb{R}^n)^{k+1}),$$

$$k = M-1, M-2, \ldots, 0.$$

Remark: Under (A3)', $\Sigma_k[u, y]$ is given as follows:

$$(\Sigma_k[u, y]\sigma)(x_{0,k+1})$$
$$= c_\epsilon \exp\left(\frac{\mu}{\epsilon} \left\{ L_k(x_{0,k}, u) + \frac{1}{2\mu} |y|^2_{\Sigma_2^{-1}} - \frac{1}{2\mu} |z(x_{0,k+1}, u, y)|^2_{\Sigma^{-1}} \right\} \right) \sigma(x_{0,k}),$$

where $|y|^2_{\Sigma_2^{-1}} = y^T \Sigma_2^{-1} y$, $|z|^2_{\Sigma^{-1}} = z^T \Sigma^{-1} z$,
$z(x_{0,k+1}, u, y) := (w_{k+1}, v_{k+1}) = (x_{k+1} - F_k(x_{0,k}, u), y - H_k(x_{0,k}))$
We expect $(\epsilon/\mu) \log V_k^\epsilon(e^{(\mu/\epsilon)p})$ converges as $\epsilon \to 0$.

Small noise limit

We define $W_k : \mathcal{D}((\mathbb{R}^n)^{k+1}) \to \mathbb{R}$ $(k = 0, 1, \dots, M)$ by

$$W_M(p) = (p, \Phi)_M, \ p \in \mathcal{D}(\mathbb{R}^n)^{M+1},$$
$$W_k(p) = \inf_{u \in U} \sup_{y \in \mathbb{R}^m} \left\{ W_{k+1}(T_k[u, y]p) - \frac{1}{2\mu} |y|^2_{\Sigma_2^{-1}} \right\}, \qquad \text{(G)}$$
$$p \in \mathcal{D}((\mathbb{R}^n)^{k+1}); k = M-1, M-2, \dots, 0,$$

where $(p, \Phi)_M := \sup_{x \in (\mathbb{R}^n)^{M+1}} \{p(x) + \Phi(x)\}$ and operator $T_k[u, y] : \mathcal{D}((\mathbb{R}^n)^{k+1}) \to \mathcal{D}((\mathbb{R}^n)^{k+2})$ is given by

$$(T_k[u, y]p)(x_{0,k+1}) = L_k(x_{0,k}, u) + \frac{1}{2\mu}|y|^2_{\Sigma_2^{-1}} - \frac{1}{2\mu}|z(x_{0,k+1}, u, y)|^2_{\Sigma^{-1}} + p(x_{0,k})$$

Small noise limit (cont'd)

Theorem

Let $p^\epsilon, p \in \mathcal{D}^\gamma((\mathbb{R}^n)^{k+1})$ $(\epsilon > 0)$ for some $\gamma \in G$ and suppose p^ϵ converges to p in $\mathcal{D}((\mathbb{R}^n)^{k+1})$ as $\epsilon \to 0$. Then

$$\frac{\epsilon}{\mu} \log V_k^\epsilon(e^{(\mu/\epsilon)p^\epsilon}) \to W_k(p) \ (\epsilon \to 0).$$

In particular, we can have

$$\frac{\epsilon}{\mu} \log V_0^\epsilon(e^{(\mu/\epsilon)p^\epsilon}) \to W_0(p) \ (\epsilon \to 0).$$

Remark: (i) The above theorem is proved by Varadhan-Laplace lemma using dynamic programming equations (cf. James-Baras-Elliott '94).
(ii) Dynamic programming equation (G) of min-max type might arise in partially observed H^∞-control of path-dependent systems.

Concluding remarks

Conclusions

- Partially observed risk-sensitive control of path-dependent discrete-time systems was considered.
- Using discrete-time analogue of Girsanov transformation, we found a sufficient statistic, *i.e.*, "unnormalized conditional density of state trajectory (information state)" to describe the risk-sensitive criterion under observation.
- The problem was reduced to completely observable stochastic control of information state and then an optimal control was constructed by using dynamic programming with time-varying state spaces.
- Min-max game was derived by small noise limit.

Future researches

- Computations of value functions and optimal control
- Gaps between Markovian and path-dependent systems