CONTROLLED DIFFUSION PROCESSES WITH COST CONSTRAINTS*

Onésimo Hernández–Lerma

Mathematics Department

CINVESTAV–IPN

México City

<ohernand@math.cinvestav.mx>

*Joint work with A.F. Mendoza–Pérez and H. Jasso–Fuentes.

Abstract.

We consider an n–dimensional controlled diffusion process with cost constraints. Under suitable assumptions, the *existence* of optimal controls is a well–known fact. In our work we go a bit further and our goal is to introduce a technique to *compute* optimal controls. To this end, we follow the Lagrange multipliers approach.

Typical unconstrained optimal control problem (OCP) : We are given

- A controlled system with state space X, and time–horizon $\tau := [0, T]$, with $T < \infty$, $T = \infty$, or T random,
- A family $\mathcal{U}$ of admissible controls, and
- A performance index (say, a "reward" function) $J_0(x, \boldsymbol{u})$.

Then the OCP is : Find $\boldsymbol{u}^* \in \mathcal{U}$ such that

$$J_0(x, \boldsymbol{u}^*) = \max_{\boldsymbol{u} \in \mathcal{U}} J_0(x, \boldsymbol{u}) \quad \forall\, x(0) = x \in X. \tag{1}$$

Control problem with cost constraints :

In addition to the above consider cost functionals

$$J_i(x, \boldsymbol{u}) \quad \forall\, x \in X,\ \boldsymbol{u} \in \mathcal{U},\ i = 1, \ldots, N,$$

and constraint constants, $\theta_1, \ldots, \theta_N$. The cost–constrained problem is

maximize $J_0(x, \boldsymbol{u})$

subject to : $J_i(x, \boldsymbol{u}) \leq \theta_i \quad \forall\, i = 1, \ldots, N,\ x \in X,\ \boldsymbol{u} \in \mathcal{U}$.

Here : We consider the n–dimensional controlled system

$$dx(t) = b(x(t), u(t))dt + \sigma(x(t))dB(t) \quad \forall\, t \geq 0, \tag{2}$$

reward functional (=long–run average reward)

$$J_0(x, \boldsymbol{u}) := \liminf_{T \to \infty} \frac{1}{T} E_x^{\boldsymbol{u}} \left[\int_0^T r(x(t), \boldsymbol{u}(t))dt \right] \tag{3}$$

and $(N = 1)$ cost functional (=long–rung average cost)

$$J_1(x, \boldsymbol{u}) := \limsup_{T\to\infty} \frac{1}{T} E_x^{\boldsymbol{u}} \left[\int_0^T c(x(t), \boldsymbol{u}(t)) dt \right], \tag{4}$$

with cost constraint $\theta_1 \equiv \theta \in \mathbb{R}$.

References

A. Arapostathis, V.S. Borkar, M.K. Ghosh (2011). *Ergodic Control of Diffusion Processes*. Cambridge University Press.

V.S. Borkar, M.K. Ghosh (1990). Controlled diffusions with constraints. *J. Math. Anal. Appl.* **152**, 88–108.

F. Dufour, R.H. Stockbridge (2011). On the existence of strict optimal controls for constrained, controlled Markov in continuous–time. *Stochastics*, to appear.

A.F. Mendoza–Pérez, O. Hernández–Lerma (2010). Markov control processes with pathwise constraints. *Math. Methods Oper. Res.* **71**.

T. Prieto–Rumeau, O. Hernández–Lerma (2008). Ergodic control of continuous–time Markov chains with pathwise constraints. *SIAM J. Control Optim.* **47**, 1888–1908.

Remark.

For comparison, consider the OCP (1) and let $\mathcal{X} \subset X$ be a set of constraints. Then we obtain a **control problem with state constraints :**

$$\text{maximize } J_0(x, \boldsymbol{u})$$

over all $\boldsymbol{u} \in \mathcal{U}$ such that the state process $x(t; x, \boldsymbol{u})$ is in $\mathcal{X}$ for all $x(0) = x \in \mathcal{X}$ and all $t \in \tau := [0, T]$.

References

V.S. Borkar, A. Budhiraja (2004). Ergodic control of constrained diffusions : characterization using HJB equations. *SIAM J. Control Optim.* **43**, 1463–1492. [Here, $\mathcal{X} \subset \mathbb{R}^n$ is a polyhedral cone.]

R. Buckdahn, D. Goreac, M. Quincampoix (2011). Stochastic optimal control and linear programming approach. *Appl. Math. Optim.* **63**, 257–276. [Here, $\mathcal{X} \subset \mathbb{R}^n$ is an arbitrary compact set.]

Remark : proof techniques.

For either cost–constrained or state–constrained problems the proof techniques are combinations of

- The direct method,
- Linear programming
- Convex Analysis,
- Dynamic programming,
- Lagrange multipliers.

Cost-constrained controlled diffusion (2)–(4)

. Consider the n–dimensional controlled diffusion (2) :

$$dx(t) = b(x(t), u(t))dt + \sigma(x(t))dB(t) \quad \forall\, t \geq 0, \tag{5}$$

with $x(0) = x$, and $B(\cdot)$ a d–dimensional Brownian motion ; coefficients

$$b(\cdot,\cdot) : \mathbb{R}^n \times U \to \mathbb{R}^n, \sigma(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n\times d},$$

where $U \subset \mathbb{R}^m$ is a compact *control set*.

Control policies : Let $\mathcal{P}(U)$ be the space of probability measures on U endowed with the topology of weak convergence, and let $\mathbb{F}$ be the family of measurable functions $f : \mathbb{R}^n \to U$

(a)

$\pi(du|x)$ is a *randomized stationary policy* (a.k.a. *relaxed* stationary control) if $\pi(\cdot|x)$ is in $\mathcal{P}(U)$ for every $x \in \mathbb{R}^n$, and $\pi(A|\cdot)$ is a measurable function on x for every Borel set $A \subset U$. We denote by Π the family of randomized stationary policies.

(b) We say that $\pi \in \Pi$ is a *deterministic stationary policy* (a.k.a. as a *strict* or *exact* control) if there exists $f \in \mathbb{F}$ such that $\pi(\cdot|x)$ is the Dirac measure concentrated at $f(x) \in U$ for all $x \in \mathbb{R}^n$. We identify $\mathbb{F}$ with the family of deterministic stationary policies. Note that $\mathbb{F} \subset \Pi$.

Remark : With a suitable topology, Π is a compact convex set, and

Remark.

We need conditions ensuring that, for every $\pi \in \Pi$, the SDE (5) has a unique *strong solution* [see Assumption A] which is also *uniformly exponentially ergodic* [see Assumption B].

Assumption A. (Uniform Ito conditions + uniform ellipticity).

(a) $b(x, u)$ is continuous on $\mathbb{R}^n \times U$, and there exists K such that

$$\sup_{u \in U} |b(x, u) - b(y, u)| \leq K|x - y| \quad \forall\, x, y \in \mathbb{R}^n.$$

(b) There exist $K > 0$ such that

$$|\sigma(x) - \sigma(y)| \leq K|x - y| \quad \forall\, x, y \in \mathbb{R}^n.$$

(c) *Uniform ellipticity :* There exists $\gamma > 0$ for which the so–called *diffusion matrix* $a(\cdot) := \sigma(\cdot)\sigma(\cdot)'$ satisfies that

$$xa(y)x' \geq \gamma |x|^2 \quad \forall\, x, y \in \mathbb{R}^n.$$

Assumption B. (Lyapunov condition). There exists a function $W \geq 1$ in $C^2(\mathbb{R}^n)$ and constants $\beta \geq \alpha > 0$ such that

(a) $\lim_{|x|\to\infty} W(x) = +\infty$, and

(b) $L^u W(x) \leq -\alpha W(x) + \beta \quad \forall\, x \in \mathbb{R}^n,\ u \in U,$

where, for every $h \in C^2(\mathbb{R}^n),\ u \in U$, and $x \in \mathbb{R}^n$:

$$L^u h(x) := \sum_{i=1}^{n} h_{x_i}(x) b_i(x,u) + \frac{1}{2} \sum_{i,j} h_{x_i,x_j}(x) a_{ij}(x) \qquad (6)$$

Remark.

(a)

For $\pi \in \Pi$, $i = 1, \ldots, n$, and $x \in \mathbb{R}^n$, let

$$b_i(x, \pi) := \int_U b_i(x, u)\pi(du|x).$$

Then $L^\pi h(x)$ is defined as in (6) with $b_i(x, \pi)$ in lieu of $b_i(x, u)$.

(b) Under Assumptions A and B, for every $\pi \in \Pi$, the corresponding solution $x(\cdot) \equiv x^\pi(\cdot)$ of (5) is a Markov process, which is positive recurrent with a unique invariant probability measure μ_π such that

$$\mu_\pi(W) := \int_{\mathbb{R}^n} W(y)\mu_\pi(dy) < \infty.$$

(c) Let $B_W(\mathbb{R}^n)$ be the normed linear space of measurable functions $v : \mathbb{R}^n \to \mathbb{R}$ with finite *W−norm* defined as

Let

$$\mu(v) := \int_{\mathbb{R}^n} v(y)\mu(dy).$$

(d) Under Assumptions A and B, for every $\pi \in \Pi$, the state process $x(\cdot) \equiv x^\pi(\cdot)$ is *uniformly W–exponentially ergodic*, which means that there exist positive constants C and δ such that

$$\sup_{\pi\in\Pi} |E_x^\pi v(x(t)) - \mu_\pi(v)| \leq Ce^{-\delta t}||v||_W W(x) \tag{7}$$

for all $x \in \mathbb{R}^n$, $v \in B_W(\mathbb{R}^n)$, and $t \geq 0$.

For every $\pi \in \Pi$ and $x \in \mathbb{R}^n$, consider the long–run average reward $J_0(x,\pi)$ in (3), i.e.,

$$J_0(x,\pi) := \liminf_{T\to\infty} \frac{1}{T}\int_0^T E_x^\pi[r(x(t),\pi)]dt,$$

where

$$r(x,\pi) := \int_U r(x,u)\pi(du|x),$$

and $r(x,f) = r(x,f(x))$ if $\pi \equiv f$ is in $\mathbb{F}$. Then, by the exponential ergodicity (7), it is evident that $J_0(x,\pi)$ is a *constant* $\bar{r}(\pi)$ independent of $x \in \mathbb{R}^n$, where

$$\bar{r}(\pi) = \int_{\mathbb{R}^n} r(y,\pi)\mu_\pi(dy),$$

i.e.

$$J_0(x,\pi) \equiv \bar{r}(\pi) \quad \forall\, \pi \in \Pi,\ x \in \mathbb{R}^n. \tag{8}$$

Similarly, the cost functional in (4), that is

$$J_1(x,\pi) := \limsup_{T\to\infty} \frac{1}{T}\int_0^T E_x^\pi[c(x(t),\pi)]dt,$$

is such that

$$J_1(x,\pi) \equiv \bar{c}(\pi) \quad \forall\, \pi \in \Pi,\ x \in \mathbb{R}^n. \tag{9}$$

with $\overline{c}(\pi) := \int c(y,\pi)\mu_\pi(dy)$, and $c(y,\pi) := \int_U c(y,u)\pi(du|y)$.

Consider the **constrained problem** CP_θ :

maximize $J_0(x,\pi)$

subject to : $J_1(x,\pi) \leq \theta \quad \forall\, x \in \mathbb{R}^n, \pi \in \Pi$.

By (7)–(8), we can express CP_θ as

maximize $\overline{r}(\pi)$

subject to : $\overline{c}(\pi) \leq \theta,\ \pi \in \Pi$.

Assumption C.

The constraint constant θ is in $(\theta_{\min}, \theta_{\max})$, where

$$\theta_{\min} := \inf_{\pi \in \Pi} \overline{c}(\pi) \ \text{ and } \ \theta_{\max} := \sup_{\pi \in \Pi} \overline{c}(\pi).$$

Let $V(\theta)$ be the optimal value of CP_θ, i,e.,

$$V(\theta) := \sup\{\overline{r}(\pi) : \overline{c}(\pi) \leq \theta, \ \pi \in \Pi\}$$

The Lagrange multipliers approach

For every $\Lambda \leq 0$ consider the reward rate function

$$r_\Lambda(x,u) := r(x,u) + (c(x,u) - \theta) \cdot \Lambda.$$

The corresponding long-run average reward

$$J_{r_\Lambda}(x,\pi) := \liminf_{T\to\infty} \frac{1}{T} E_x^\pi \left[\int_0^T r_\Lambda(x(t),\pi)dt \right]$$

satisfies that

$$J_{r_\Lambda}(x,u) = J_0(x,\pi) + (J_1(x,\pi) - \theta) \cdot \Lambda.$$

Assumption D.

(a) $r(x,u)$ and $c(x,u)$ are continuous on $\mathbb{R}^n \times U$, and locally Lipschitz in x, uniformly in $u \in U$; that is, for each $R > 0$, there exists a constant $K(R)$ such that

$$\sup_{u \in U} |r(x,u) - r(y,u)| \leq K(R)|x - y| \quad \forall\, |x|, |y| \leq R,$$

and similarly for $c(x,u)$.

(b) The squared functions $r(x,u)^2$ and $c(x,u)^2$ are in $B_W(\mathbb{R}^n)$ uniformly in U; that is, there exists $M > 0$ such that

$$\sup_{u \in U} r(x,u)^2 \leq MW(x) \quad \text{and} \quad \sup_{u \in U} c(x,u)^2 \leq MW(x)$$

for all $x \in \mathbb{R}^n$.

Theorem.

Suppose that the Assumptions A, B, C, D are satisfied.

(a) For each $\Lambda \leq 0$, there exists a solution $(\rho(\Lambda), h_\Lambda)$, with $\rho(\Lambda) \in \mathbb{R}$ and $h_\Lambda \in C^2(\mathbb{R}^n) \cap B_W(\mathbb{R}^n)$, of the HJB equation

$$\rho(\Lambda) = \max_{u \in U}[r_\Lambda(x, a) + L^u h_\Lambda(x)] \quad \forall\, x \in \mathbb{R}^n. \tag{10}$$

(b) $V(\theta) = \min_{\Lambda \leq 0} \rho(\Lambda) = \rho(\Lambda_0)$ for some $\Lambda_0 \leq 0$.

(c) Suppose that there exists $\Lambda \leq 0$ and $\hat{\pi} \in \Pi$ satisfying

$$\bar{c}(\hat{\pi}) = \theta \quad \text{and} \quad \bar{r}_\Lambda(\hat{\pi}) = \rho(\Lambda),$$

with $\rho(\Lambda)$ as in (a). Then $\hat{\pi}$ is an optimal policy for CP_θ, i.e.

$$\bar{r}(\hat{\pi}) = V(\theta).$$

Further, if $f_\Lambda \in \mathbb{F}$ attains the maximum in the r.h.s. of (10) and $\overline{c}(f_\Lambda) = \theta$, then f_Λ is a deterministic (or exact or strict) optimal policy for CP_θ.

(d) Let $\rho(\Lambda)$ and f_Λ be as above, and suppose that $\Lambda \mapsto \rho(\Lambda)$ is differentiable at some $\Lambda < 0$. Then

$$\rho'(\Lambda) = \overline{c}(f_\Lambda) - \theta.$$

In particular, if $\Lambda < 0$ is a critical point of $\rho(\cdot)$, then f_Λ is an optimal policy for CP_θ, and part (b) holds with $\Lambda_0 = \Lambda$.

(e) Summarizing : If $\rho(\cdot)$ is differentiable at some $\Lambda < 0$, then the following statements are equivalent.

(1) f_Λ is an optimal policy for CP_θ and $\rho(\Lambda) = V(\theta)$;

(2) $\overline{c}(f_\Lambda) = \theta$;

(3) Λ is a critical point of $\rho(\cdot)$.

(f) In addition, assume that the mapping $\Lambda \mapsto \overline{c}(f_\Lambda)$ is continuous on the interval $(-\infty, 0)$. Then the function $\rho(\cdot)$ is differentiable on $(-\infty, 0)$.

(g) [What happens at $\Lambda = 0$?] If $\overline{c}(f_0) \leq \theta$, then f_0 is an optimal policy for CP_θ, and (b) holds at $\Lambda_0 = 0$, i.e.,

$$V(\theta) = \min_{\Lambda \leq 0} \rho(\Lambda) = \rho(0).$$

For an example and further details see :

- A.F. Mendoza–Pérez, H. Jasso–Fuentes, O. Hernández–Lerma, *The Lagrange approach to ergodic control of diffusions with cost constraints.* Submitted to *Optimization.*

THANK YOU FOR YOUR ATTENTION !