Unconstrained and Constrained Optimal Control of Piecewise Deterministic Markov Processes

Oswaldo Costa, François Dufour, Alexey Piunovskiy

Universidade de Sao Paulo

Institut de Mathématiques de Bordeaux
INRIA Bordeaux Sud-Ouest

University of Liverpool

This work has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the "Investments for the future" Programme IdEx Bordeaux - CPU (ANR-10-IDEX-03-02)

Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the controlled process
   - Admissible strategies and conditional distribution
2. Optimization problems
   - Unconstrained and constrained problems
   - Assumptions and preliminary result
3. The unconstrained problem and the dynamic programming approach
4. The constrained problem and the linear programming approach

Introduction

Davis (80's)

General class of non-diffusion stochastic hybrid models:

Family of processes with discrete/continuous state space and deterministic/stochastic jumps.

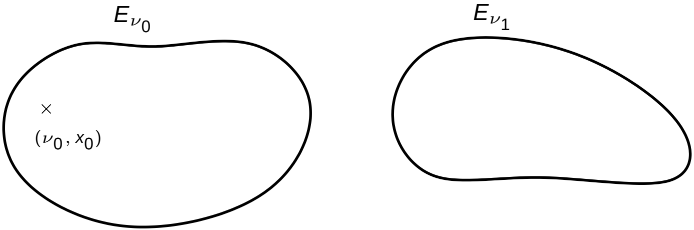
Applications

Engineering systems, biology, operations research, management science, economics, dependability and safety, ...

Uncontrolled process

Definition of a PDMP

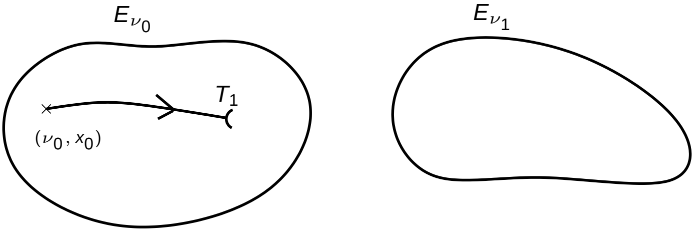
Parameters: flow ϕ, intensity of the jumps λ, transition kernel Q



Uncontrolled process

Definition of a PDMP

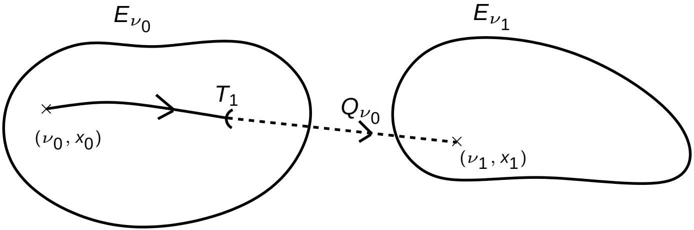
Parameters: flow ϕ, intensity of the jumps λ, transition kernel Q



Uncontrolled process

Definition of a PDMP

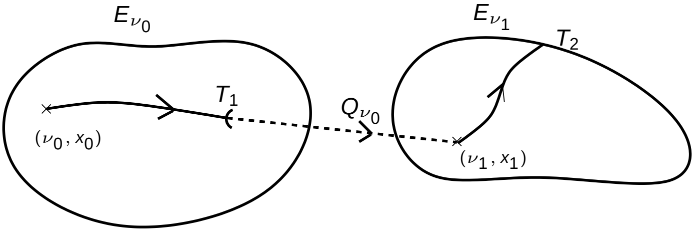
Parameters: flow ϕ, intensity of the jumps λ, transition kernel Q



Uncontrolled process

Definition of a PDMP

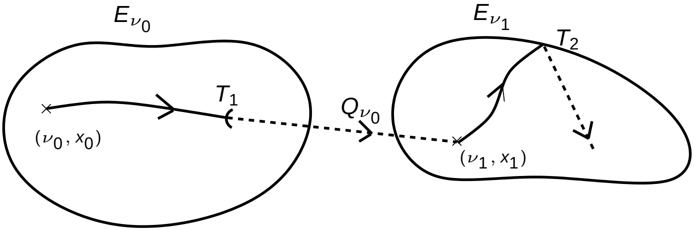
Parameters: flow ϕ, intensity of the jumps λ, transition kernel Q



Uncontrolled process

Definition of a PDMP

Parameters: flow ϕ, intensity of the jumps λ, transition kernel Q



Parameters of the model

- the state space: $\mathbf{X}$ open subset of $\mathbb{R}^d$ (boundary $\partial\mathbf{X}$).
- the flow: $\phi(x,t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ satisfying
  $\phi(x, t+s) = \phi(\phi(x,s), t)$ for all $x \in \mathbb{R}^d$ and $(t,s) \in \mathbb{R}^2$.
  $\to$ active boundary:
  $\Delta = \{z \in \partial\mathbf{X} : z = \phi(x,t) \text{ for some } x \in \mathbf{X} \text{ and } t \in \mathbb{R}_+^*\}$.
  For $x \in \overline{\mathbf{X}} \doteq \mathbf{X} \cup \Delta$,

  $$t^*(x) = \inf\{t \in \mathbb{R}_+ : \ \phi(x,t) \in \Delta\}.$$

  The flow is not controlled.
- $\mathbf{A}$ is the action space, assumed to be a Borel space. The set of *feasible* actions in state $x \in \overline{\mathbf{X}}$ is $\mathbf{A}(x) \subset \mathbf{A}$.
  The set of feasible pairs

  $$\mathbf{K} = \{(x,a) \in \overline{\mathbf{X}} \times \mathbf{A} : a \in \mathbf{A}(x)\}$$

Parameters of the model

- The jumps intensity λ is a $\mathbb{R}_+$-valued measurable function defined on $\mathbf{K}$.
- The stochastic kernel Q on $\mathbf{X}$ given $\mathbf{K}$ satisfies $Q(\mathbf{X} \setminus \{x\}|x, a) = 1$ for any $(x, a) \in \mathbf{K}$.
  In state x and choosing the action $a \in \mathbf{A}(x)$, the distribution of the next state is given by $Q(\cdot|x, a) = 1$ at a time of jump.

The canonical space

$$\Omega = \Big(\bigcup_{n=1}^{\infty} \Omega_n\Big) \bigcup \big(\mathbf{X} \times (\mathbb{R}_+^* \times \mathbf{X})^{\infty}\big)$$

with $\Omega_n = \mathbf{X} \times (\mathbb{R}_+^* \times \mathbf{X})^n \times (\{\infty\} \times \{x_\infty\})^{\infty}$.

Interpretation of Ω:

Consider $\omega = (x_0, \theta_1, x_1, \theta_2, x_2, \ldots) \in \Omega$

- $x_0 \in \mathbf{X}$ is the initial state.
- Given $n \geq 0$, if $x_n \in \mathbf{X}$ then
  - either $0 < \theta_{n+1} < \infty$, and we interpret θ_{n+1} as the sojourn time in state $x_n \in \mathbf{X}$, while $x_{n+1} \in \mathbf{X}$ is the post-jump location of the process;
  - or $\theta_{n+1} = \infty$; this means that the system has been absorbed by x_n. In this case we set $x_m = x_\infty$ and $\theta_m = \infty$ for all $m > n$. Such sample paths belong to Ω_n.

Construction of the controlled process

Introduce the mappings $X_n : \ \Omega \to \mathbf{X}_\infty = \mathbf{X} \cup \{x_\infty\}$ by $X_n(\omega) = x_n$ and $\Theta_n : \ \Omega \to \overline{\mathbb{R}}_+^*$ by $\Theta_n(\omega) = \theta_n$; $\Theta_0(\omega) = 0$ where

$$\omega = (x_0, \theta_1, x_1, \theta_2, x_2, \dots) \in \Omega.$$

In addition $T_n(\omega) = \sum_{i=1}^{n} \Theta_i(\omega) = \sum_{i=1}^{n} \theta_i$ with $T_\infty(\omega) = \lim_{n\to\infty} T_n(\omega)$.

$\mathbf{H}_n$ is the set of path up to n.
$H_n = (X_0, \Theta_1, X_1, \dots, \Theta_n, X_n)$ is the history of the process up to n.

Construction of the process

The controlled process $\left\{\xi_t\right\}_{t\in\mathbb{R}_+}$

$$\xi_t(\omega) = \begin{cases} \phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ x_\infty, & \text{if } T_\infty \leq t. \end{cases}$$

$\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ denote the filtration generated by the process $\left\{\xi_t\right\}_{t\in\mathbb{R}_+}$.

Admissible control strategy $u = (\pi_n)_{n\in\mathbb{N}}$

It is a sequence of stochastic kernels on $\mathbf{A}$ given $\mathbf{H}_n \times \mathbb{R}_+^*$ satisfying: $\pi_n(da|h_n, t) = 1$ for $t \in]0, t^*(x_n)] \cap \mathbb{R}_+^*$, where $h_n = (x_0, \theta_1, x_1, \ldots \theta_n, x_n) \in \mathbf{H}_n$.

The set of admissible control strategies is denoted by $\mathcal{U}$.

Admissible strategies and conditional distribution

Consider an admissible strategy $u \in \mathcal{U}$ and an initial state $x_0 \in \mathbf{X}$, there exists $\mathbb{P}^u_{x_0}$ on $(\Omega, \mathcal{F})$ [Jacod, 75, Multivariate point processes] satisfying

$$\mathbb{P}^u_{x_0}\Big((\Theta_{n+1}, X_{n+1}) = (+\infty, x_\infty)\big|\mathcal{F}_{T_n}\Big) = \begin{cases} 1 & \text{if } X_n = x_\infty, \\ e^{-\Lambda^u_n(H_n, +\infty)} & \text{if } X_n \in \mathbf{X} \text{ and } t^*(X_n) = +\infty, \\ 0 & \text{otherwise,} \end{cases}$$

where $\Lambda^u_n(H_n, t)$ is the rate of jumps *averaged over the action*

$$\Lambda^u_n(H_n, t) = \int_{]0,t]} \int_{\mathbf{A}} \lambda(\phi(X_n, s), a)\pi_n(da|H_n, s)ds$$

at step n.

Admissible strategies and conditional distribution

For $\Gamma_2 \in \mathcal{B}(\mathbf{X})$,

$$\mathbb{P}^u_{x_0}\Big(\Theta_{n+1} = t^*(X_n); X_{n+1} \in \Gamma_2 \big| \mathcal{F}_{T_n}\Big) = \begin{cases} Q^u_n(\Gamma_2|H_n)e^{-\Lambda^u_n(H_n, t^*(X_n))} & \text{if } X_n \in \mathbf{X} \text{ and } t^*(X_n) < +\infty, \\ 0 & \text{otherwise.} \end{cases}$$

where $Q^u_n(dx|H_n)$ is the distribution of the state after a *deterministic* jump

$$Q^u_n(dx|H_n) = \int_{\mathbf{A}} Q(dx|\phi(X_n, t^*(X_n)), a)\pi_n(da|H_n, t^*(X_n)).$$

Admissible strategies and conditional distribution

For $\Gamma_1 \in \mathcal{B}(\mathbb{R}_+^*)$, $\Gamma_2 \in \mathcal{B}(\mathbf{X})$

$$\mathbb{P}_{x_0}^{u}\Big(\Theta_{n+1} \in \Gamma_1; X_{n+1} \in \Gamma_2 \Big| \mathcal{F}_{T_n}\Big)$$
$$= \begin{cases} \displaystyle\int_{]0,t^*(X_n)[\cap\Gamma_1} Q_n^u(\Gamma_2|H_n,t)\lambda_n^u(H_n,t)e^{-\Lambda_n^u(H_n,t)}dt & \text{if } X_n \in \mathbf{X} \\ 0 & \text{otherwise.} \end{cases}$$

where Q_n^u is the distribution of the state after a *stochastic* jump

$$Q_n^u(dx|h_n,t)$$
$$= \frac{1}{\lambda_n^u(h_n,t)} \int_{\mathbf{A}} Q(dx|\phi(x_n,t),a)\lambda(\phi(x_n,t),a)\pi_n(da|h_n,t)$$

with $\lambda_n^u(h_n,t)$ being the corresponding intensity of jumps

$$\lambda_n^u(h_n,t) = \int_{\mathbf{A}} \lambda(\phi(x_n,t),a)\pi_n(da|h_n,t).$$

Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the controlled process
   - Admissible strategies and conditional distribution
2. Optimization problems
   - Unconstrained and constrained problems
   - Assumptions and preliminary result
3. The unconstrained problem and the dynamic programming approach
4. The constrained problem and the linear programming approach

Unconstrained and constrained problems

Cost functions

$\left(C_k\right)_{k\in\{0,1,\ldots,p\}}$ real-valued mapping defined on $\mathbf{K}$.

The associated infinite-horizon discounted criteria corresponding to $u \in \mathcal{U}$ are given for any $k \in \{0, 1, \ldots, p\}$ by

$$\mathcal{V}_k(u, x_0) = \mathbb{E}^u_{x_0}\left[\int_{]0,+\infty[} e^{-\alpha s} \int_{\mathbf{A}(\xi_s)} C_k(\xi_s, a)\pi(da|s)ds\right]$$

$$+ \mathbb{E}^u_{x_0}\left[\int_{]0,+\infty[} e^{-\alpha s} I_{\{\xi_{s-}\in\Delta\}} \int_{\mathbf{A}(\xi_{s-})} C_k(\xi_{s-}, a)\pi(da|s)\mu(ds)\right].$$

with

$$\pi(da|t) = \sum_{n\in\mathbb{N}} I_{\{T_n<t\leq T_{n+1}\}}\pi_n(da|H_n, t - T_n)$$

and

$\mu(ds)$ is the point process associated to $\{T_n\}_{n\in\mathbb{N}}$

Unconstrained and constrained problems

- The optimization problem without constraint consists in minimizing the performance criterion

$$\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0).$$

- The optimization problem with p constraints consists in minimizing the performance criterion

$$\inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0)$$

where $\mathcal{U}^f$ is the set of feasible controls, that is, $\mathcal{U}^f$ and such that the constraint criteria

$$\mathcal{V}_k(u, x_0) \leq B_k$$

are satisfied for any $k \in \{1, \ldots, p\}$, where $(B_k)_{k \in \{1,\ldots,p\}}$ are real numbers representing the constraint bounds.

Different classes of strategies

- *stationary*, if $u = (\pi_n)_{n\in\mathbb{N}}$ with $\pi_n(da|h_n, t) = \pi(da|\phi(x_n, t))$ for some stochastic kernel π on $\mathbf{A}$ given $\mathbf{X}$.
- *deterministic stationary*, if $\pi_n(\cdot|h_n, t) = \delta_{\varphi^s(\phi(x_n,t))}(\cdot)$, where $\varphi^s : \overline{\mathbf{X}} \to \mathbf{A}$ is a measurable mapping satisfying $\varphi^s(y) \in \mathbf{A}(y)$ for any $y \in \overline{\mathbf{X}}$.

Hypotheses

Assumption A. There are constants $K \geq 0, \varepsilon_1 > 0$ and $\varepsilon_2 \in [0, 1[$ such that

(A1) *For any* $(x, a) \in \mathbf{K}^g$, $\lambda(x, a) \leq K$

(A2) $\inf_{(z,b)\in\Delta\times\mathbf{A}} Q(A_{\varepsilon_1}|z, b) \geq 1 - \varepsilon_2$, *with* $A_{\varepsilon_1} = \{x \in \mathbf{X} : t^*(x) > \varepsilon_1\}$.

Assumption B.

(B1) *The set* $\mathbf{A}(y)$ *is compact for every* $y \in \overline{\mathbf{X}}$.

(B2) *The kernel* Q *is weakly continuous.*

(B3) *The function* λ *is continuous on* $\mathbf{K}$.

(B4) *The flow* ϕ *is continuous on* $\mathbb{R}_+ \times \mathbb{R}^p$.

(B5) *The function* t^* *is continuous on* $\overline{\mathbf{X}}$.

Assumption C.

(C1) *The multifunction* Ψ *from* $\overline{\mathbf{X}}$ *to* $\mathbf{A}$ *defined by* $\Psi(x) = \mathbf{A}(x)$ *is upper semicontinuous.*

(C2) *The cost function* C_0^g *(respectively,* C_0^i*) is bounded and lower semicontinuous on* $\mathbf{K}^g$ *(respectively,* $\mathbf{K}^i$*).*

Lemma

Suppose Assumption A is satisfied. Then there exists $M < \infty$ such that, for any control strategy $u \in \mathcal{U}$ and for any $x_0 \in \mathbf{X}$

$$\mathbb{E}_{x_0}^{u}\Big[\sum_{n \in \mathbb{N}^*} e^{-\alpha T_n}\Big] \leq M \text{ and } \mathbb{P}_{x_0}^{u}(T_\infty < +\infty) = 0.$$

Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the controlled process
   - Admissible strategies and conditional distribution
2. Optimization problems
   - Unconstrained and constrained problems
   - Assumptions and preliminary result
3. The unconstrained problem and the dynamic programming approach
4. The constrained problem and the linear programming approach

There are two approaches to deal with such problems:

- the associated discrete-stage Markov decision model:
  - A. Almudevar. A dynamic programming algorithm for the optimal control of piecewise deterministic Markov processes, 2001.
  - O.L.V Costa and F. Dufour. Continuous average control of piecewise deterministic Markov processes, 2013.
  - M.H.A. Davis. Control of piecewise-deterministic processes via discrete-time dynamic programming, 1986.
  - L. Forwick, M. Schal, and M. Schmitz. Piecewise deterministic Markov control processes with feedback controls and unbounded costs, 2004.
  - M. Schal. On piecewise deterministic Markov control processes: control of jumps and of risk processes in insurance, 1998.
  - A.A. Yushkevich. On reducing a jump controllable Markov model to a model with discrete time, 1980.
- the infinitesimal approach (HJB equation):
  - M.H.A. Davis. Markov models and optimization, volume 49 of Monographs on Statistics and Applied Probability, 1993.
  - M.A.H. Dempster and J.J. Ye. Necessary and sufficient optimality conditions for control of piecewise deterministic processes, 1992.
  - M.A.H. Dempster and J.J. Ye. Generalized Bellman-Hamilton-Jacob optimality conditions for a control problem with boundary conditions, 1996.
  - A.A. Yushkevich. Bellman inequalities in Markov decision deterministic drift processes. Stochastics, 1987

Notation:

- $\mathbb{A}(\overline{\mathbf{X}})$ is the set of functions $g \in \mathbb{B}(\overline{\mathbf{X}})$ such that for any $x \in \overline{\mathbf{X}}$, the function $g(\phi(x, \cdot))$ is absolutely continuous on $[0, t^*(x)] \cap \mathbb{R}_+$.
- Let $g \in \mathbb{A}(\overline{\mathbf{X}})$, there exists a real-valued measurable function $\mathcal{X}g$ defined on $\mathbf{X}$ satisfying for any $t \in [0, t^*(x)[$

$$g(\phi(x,t)) = g(x) + \int_{[0,t]} \mathcal{X}g(\phi(x,s))ds.$$

- $q(dy|x,a) \doteq \lambda(x,a)\big[Q(dy|x,a) - \delta_x(dy)\big]$

Theorem

Suppose assumptions A, B and C hold. Then there exist $W \in \mathbb{A}(\overline{\mathbf{X}})$ *and* $\mathcal{X}W \in \mathbb{B}(\mathbf{X})$ *satisfying for any* $x \in \mathbf{X}$*,*

$$-\alpha W(x) + \mathcal{X}W(x) + \inf_{a \in A(x)} \Big\{ C_0(x,a) + qW(x,a) \Big\} = 0,$$

and for any $z \in \Delta$

$$W(z) = \inf_{b \in A(z)} \Big\{ C_0(z,b) + QW(z,b) \Big\}.$$

Moreover, for any $x \in \mathbf{X}$

$$W(x) = \inf_{u \in \mathcal{U}} \mathcal{V}_0(u,x).$$

There exists an optimal stationary deterministic strategy $\widehat{\varphi}$ *satisfying* $\widehat{\varphi}(y) \in \mathbf{A}(y)$ *for any* $y \in \overline{\mathbf{X}}$

Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the controlled process
   - Admissible strategies and conditional distribution
2. Optimization problems
   - Unconstrained and constrained problems
   - Assumptions and preliminary result
3. The unconstrained problem and the dynamic programming approach
4. The constrained problem and the linear programming approach

The constrained problem and the linear programming approach

The method has been extensively studied in the literature

- Continuous and discrete time MDP:
  - Eitan Altman. Constrained Markov decision processes, 1999.
  - Vivek S. Borkar. A Convex Analytic Approach to Markov Decision Processes, 1988.
  - Vivek S. Borkar. Convex analytic methods in Markov decision processes, 2002.
  - Alexey B. Piunovskiy. Optimal control of random sequences in problems with constraints, 1997.

Occupation measure

For any admissible control strategy $u \in \mathcal{U}$, the occupation measure $\eta_u \in \mathcal{M}(\mathbf{K})$ associated with u is defined as follows

$$\begin{aligned}\eta_u(\Gamma) =& \mathbb{E}^u_{x_0}\Big[\int_\Gamma \int_{]0,\infty[} e^{-\alpha s}\delta_{\xi_s}(dx)\pi(da|s)ds\Big] \\ &+ \mathbb{E}^u_{x_0}\Big[\int_\Gamma \sum_{n\in\mathbb{N}^*} e^{-\alpha T_n} I_{\{\xi_{T_n -}\in\Delta\}}\delta_{\xi_{T_n -}}(dx)\pi(da|T_n)\Big].\end{aligned}$$

for any $\Gamma \in \mathcal{B}(\mathbf{K})$. The infinite-horizon discounted criteria can be rewritten as

$$\begin{aligned}\mathcal{V}_j(u, x_0) =& \mathbb{E}^u_{x_0}\Big[\int_{]0,+\infty[} e^{-\alpha s}\int_{\mathbf{A}(\xi_s)} C_j(\xi_s, a)\pi(da|s)ds\Big] \\ &+ \mathbb{E}^u_{x_0}\Big[\int_{]0,+\infty[} e^{-\alpha s} I_{\{\xi_{s-}\in\Delta\}}\int_{\mathbf{A}(\xi_{s-})} C_j(\xi_{s-}, a)\gamma(da|s)\mu(ds)\Big] \\ =& \int_{\mathbf{K}} C_j(x,a)\eta_u(dx,da) \doteq \eta_u(C_j)\end{aligned}$$

Linear programming approach

The constrained linear program is defined as

$$\inf_{\eta \in \mathbb{M}} \eta(C_0)$$

where $\mathbb{M}$ is the set of measures η in $\mathcal{M}(\mathbf{K})$ satisfying for any $(W, \mathcal{X}W) \in \mathbb{A}(\overline{\mathbf{X}}) \times \mathbb{B}(\mathbf{X})$

$$\int_{\overline{\mathbf{X}}} \Big[I_{\mathbf{X}}(x) \big[\alpha W(x) - \mathcal{X}W(x) \big] + I_{\Delta}(x) W(x) \Big] \widehat{\eta}(dx)$$

$$= W(x_0) + \int_{\mathbf{K}} \Big[I_{\mathbf{X}}(x) qW(x,a) + I_{\Delta}(x) QW(z,b) \Big] \eta(dx, da).$$

where $\widehat{\eta}$ denotes the marginal of η w.r.t. to $\overline{\mathbf{X}}$ and also the constraints

$$\eta(C_j) \leq B_j$$

for $j \geq 1$.

Linear programming approach

Theorem

Suppose Assumption A holds and the cost functions C_j are bounded from below for any $j \in \mathbb{N}_p$. Then the values of the constrained control problem and the linear program are equivalent:

$$\inf_{\eta \in \mathbb{M}} \eta(C_0) = \inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0).$$

Thank you for your attention.