Discrete-time hybrid control models

Héctor Jasso-Fuentes

Department of Mathematics
CINVESTAV-IPN
Mexico City

May 9, 2018

Join work with
J.L. Menaldi (Wayne State University-USA), **T. Prieto-Rumeau** (UNED-Spain), and **M. Robin** (École Polytechnique-France)

What does “hybrid” mean?

- The idea behind a hybrid system is a so-called event-driven evolution.
- Under normal circumstances a standard-type sub-dynamic is a good description of the real phenomenon, but some events may occur (due to internal or exogenous causes) and the model becomes invalid, which forces the modeler to reconsider the data of the problem.
- The dynamic may undergo structural modifications, i.e., from time to time the law of evolution may suffer deep changes.

Example: A pollution accumulation problem

- Suppose that an economy consumes a specific product and that, as a byproduct of this consumption, it generates pollution.
- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{1}$$

  with initial stock $x_0 = x$.
- $x^f_t \in [0, \infty)$ = the stock of pollution at time t.
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{2}$$

with initial stock $x_0 = x$.

- $x^f_t \in [0, \infty)$ = the stock of pollution at time t.
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{3}$$

  with initial stock $x_0 = x$.
- $x^f_t \in [0, \infty)$ = the stock of pollution at time t.
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{4}$$

with initial stock $x_0 = x$.

- $x^f_t \in [0, \infty)$ = the stock of pollution at time t
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{5}$$

with initial stock $x_0 = x$.

- $x^f_t \in [0, \infty)$ = the stock of pollution at time t
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{6}$$

with initial stock $x_0 = x$.

- $x^f_t \in [0, \infty) =$ the stock of pollution at time t
- $x^s_t \in \{1, \cdots, M\} =$ the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0 =$ the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t) =$ the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{7}$$

  with initial stock $x_0 = x$.
- $x^f_t \in [0, \infty)$ = the stock of pollution at time t
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- We assume that the stock of pollution x_t is gradually degraded and is represented by the system

$$x^f_{t+1} = p(c_t) - g(x^s_t)x^f_t + \xi_t, \quad t \geq 1, \tag{8}$$

  with initial stock $x_0 = x$.
- $x^f_t \in [0, \infty)$ = the stock of pollution at time t
- $x^s_t \in \{1, \cdots, M\}$ = the levels (or modes) of environmental contingency decided by the government.
- $c_t \geq 0$ = the consumption rate at time t with range in $[0, \gamma(x^f_t, x^s_t)]$.
- $p(c_t)$ = the amount of pollution derived to consume the quantity c_t.
- For each $i \in \{1, \cdots, M\}$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$.
- $\{\xi_t\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

Example: A pollution accumulation problem

- The government (controller) is able to apply two actions at each period of time t:
  - To implement the consumption $c_t \mapsto$ producing a disutility $D(x_t^f, x_t^s)$, or
  - To change among the different modes of environmental contingency $\mapsto$ produces a cost $\ell(x_t^s, x_{t+1}^s)$ of switching from mode x_t^s to mode x_{t+1}^s. This action is instantaneous in time!
- The complete dynamic is given by

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = \left(p(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s\right)}_{\text{standard sub-dynamic}} \text{ if } \ a_t \in [0, \gamma(x_t^f, x_t^s)],$$

or

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_k^f, a_k)}_{\text{special sub-dynamic}} \quad \text{if } \ a_t \in \{0, 1, \cdots, M\},$$

- The objective is to find a consumption-mode policy $\{a_t\}_t$ of type $a_t = c_t$ (quantities to consume) or $a_t = x_t^f$ (change of modes) that minimizes certain cost functional.

Example: A pollution accumulation problem

- The government (controller) is able to apply two actions at each period of time t:
  - To implement the consumption $c_t \mapsto$ producing a disutility $D(x_t^f, x_t^s)$, or
  - To change among the different modes of environmental contingency $\mapsto$ produces a cost $\ell(x_t^s, x_{t+1}^s)$ of switching from mode x_t^s to mode x_{t+1}^s. This action is instantaneous in time!
- The complete dynamic is given by

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = \left(p(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s\right)}_{\text{standard sub-dynamic}} \text{ if } \ a_t \in [0, \gamma(x_t^f, x_t^s)],$$

or

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_k^f, a_k)}_{\text{special sub-dynamic}} \quad \text{if } \ a_t \in \{0, 1, \cdots, M\},$$

- The objective is to find a consumption-mode policy $\{a_t\}_t$ of type $a_t = c_t$ (quantities to consume) or $a_t = x_t^f$ (change of modes) that minimizes certain cost functional.

Example: A pollution accumulation problem

- The government (controller) is able to apply two actions at each period of time t:
  - To implement the consumption $c_t \mapsto$ producing a disutility $D(x_t^f, x_t^s)$, or
  - To change among the different modes of environmental contingency $\mapsto$ produces a cost $\ell(x_t^s, x_{t+1}^s)$ of switching from mode x_t^s to mode x_{t+1}^s. This action is instantaneous in time!
- The complete dynamic is given by

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (p(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s)}_{\text{standard sub-dynamic}} \text{ if } \; a_t \in [0, \gamma(x_t^f, x_t^s)],$$

or

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_k^f, a_k)}_{\text{special sub-dynamic}} \quad \text{if } \; a_t \in \{0, 1, \cdots, M\},$$

- The objective is to find a consumption-mode policy $\{a_t\}_t$ of type $a_t = c_t$ (quantities to consume) or $a_t = x_t^f$ (change of modes) that minimizes certain cost functional.

Example: A pollution accumulation problem

- The government (controller) is able to apply two actions at each period of time t:
  - To implement the consumption $c_t \mapsto$ producing a disutility $D(x_t^f, x_t^s)$, or
  - To change among the different modes of environmental contingency $\mapsto$ produces a cost $\ell(x_t^s, x_{t+1}^s)$ of switching from mode x_t^s to mode x_{t+1}^s. This action is instantaneous in time!
- The complete dynamic is given by

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (p(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s)}_{\text{standard sub-dynamic}} \text{ if } a_t \in [0, \gamma(x_t^f, x_t^s)],$$

or

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_k^f, a_k)}_{\text{special sub-dynamic}} \quad \text{if } a_t \in \{0, 1, \cdots, M\},$$

- The objective is to find a consumption-mode policy $\{a_t\}_t$ such that $a_t = c_t$ (quantities to consume) or $a_t = x_t^f$ (change of modes) that minimizes certain cost functional.

Example: A pollution accumulation problem

- The government (controller) is able to apply two actions at each period of time t:
  - To implement the consumption $c_t \mapsto$ producing a disutility $D(x_t^f, x_t^s)$, or
  - To change among the different modes of environmental contingency $\mapsto$ produces a cost $\ell(x_t^s, x_{t+1}^s)$ of switching from mode x_t^s to mode x_{t+1}^s. This action is instantaneous in time!
- The complete dynamic is given by

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = \big(p(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s\big)}_{\text{standard sub-dynamic}} \text{ if } a_t \in [0, \gamma(x_t^f, x_t^s)],$$

or

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_k^f, a_k)}_{\text{special sub-dynamic}} \quad \text{if } a_t \in \{0, 1, \cdots, M\},$$

- The objective is to find a consumption-mode policy $\{a_t\}_t$ of type $a_t = c_t$ (quantities to consume) or $a_t = x_t^f$ (change of modes) that minimizes certain cost functional.

Example: A pollution accumulation problem

- Additional restrictions:
  - $D^\wedge := [k_1, \infty) \times \{1, \cdots, M\}$ and $D^\vee := [k_2, \infty) \times \{1, \cdots, M\}$ with $k_2 < k_1$.
- The rule is:
  - Outside $D^\vee$ (i.e., on $[0, k_2) \times \{1, \cdots, M\}$) (low level of pollution), the rule suggests to keep consuming the quantity c_t and to keep the same mode.
  - On $D^\wedge$ (i.e., on $[k_1, \infty) \times \{1, \cdots, M\}$) (high level of pollution), the rule suggests to change immediately of regime (mode),
  - On $D^\vee \setminus D^\wedge$, the controller has the option to either consume or change the mode.
- Every change also depends on certain set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution.

Example: A pollution accumulation problem

- Additional restrictions:
  - $D^{\wedge} := [k_1, \infty) \times \{1, \cdots, M\}$ and $D^{\vee} := [k_2, \infty) \times \{1, \cdots, M\}$ with $k_2 < k_1$.
- The rule is:
  - Outside $D^{\vee}$ (i.e., on $[0, k_2) \times \{1, \cdots, M\}$) (low level of pollution), the rule suggests to keep consuming the quantity c_t and to keep the same mode.
  - On $D^{\wedge}$ (i.e., on $[k_1, \infty) \times \{1, \cdots, M\}$) (high level of pollution), the rule suggests to change immediately of regime (mode),
  - On $D^{\vee} \setminus D^{\wedge}$, the controller has the option to either consume or change the mode.
- Every change also depends on certain set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution.

Example: A pollution accumulation problem

- Additional restrictions:
  - $D^{\wedge} := [k_1, \infty) \times \{1, \cdots, M\}$ and $D^{\vee} := [k_2, \infty) \times \{1, \cdots, M\}$ with $k_2 < k_1$.
- The rule is:
  - Outside $D^{\vee}$ (i.e., on $[0, k_2) \times \{1, \cdots, M\}$) (low level of pollution), the rule suggests to keep consuming the quantity c_t and to keep the same mode.
  - On $D^{\wedge}$ (i.e., on $[k_1, \infty) \times \{1, \cdots, M\}$) (high level of pollution), the rule suggests to change immediately of regime (mode),
  - On $D^{\vee} \setminus D^{\wedge}$, the controller has the option to either consume or change the mode.
- Every change also depends on certain set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution

Example: A pollution accumulation problem

- Additional restrictions:
  - $D^{\wedge} := [k_1, \infty) \times \{1, \cdots, M\}$ and $D^{\vee} := [k_2, \infty) \times \{1, \cdots, M\}$ with $k_2 < k_1$.
- The rule is:
  - Outside $D^{\vee}$ (i.e., on $[0, k_2) \times \{1, \cdots, M\}$) (low level of pollution), the rule suggests to keep consuming the quantity c_t and to keep the same mode.
  - On $D^{\wedge}$ (i.e., on $[k_1, \infty) \times \{1, \cdots, M\}$) (high level of pollution), the rule suggests to change immediately of regime (mode),
  - On $D^{\vee} \setminus D^{\wedge}$, the controller has the option to either consume or change the mode.
- Every change also depends on certain set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution

Example: A pollution accumulation problem

- Additional restrictions:
  - $D^{\wedge} := [k_1, \infty) \times \{1, \cdots, M\}$ and $D^{\vee} := [k_2, \infty) \times \{1, \cdots, M\}$ with $k_2 < k_1$.
- The rule is:
  - Outside $D^{\vee}$ (i.e., on $[0, k_2) \times \{1, \cdots, M\}$) (low level of pollution), the rule suggests to keep consuming the quantity c_t and to keep the same mode.
  - On $D^{\wedge}$ (i.e., on $[k_1, \infty) \times \{1, \cdots, M\}$) (high level of pollution), the rule suggests to change immediately of regime (mode),
  - On $D^{\vee} \setminus D^{\wedge}$, the controller has the option to either consume or change the mode.
- Every change also depends on certain set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution.

Model definition

- The *state space* X is the product $X = X^f \times X^s$ of two Borel spaces, where the components $x^f \in X^f$ and $x^s \in X^s$ are called the fast and slow states, respectively.
- The *action space* A is a Borel space and it is the union of two disjoint measurable subsets: $A = V^f \cup V^s$. The sets V^f and V^s are referred to as the fast and the slow action sets, respectively.
- The set of feasible state-action pairs is given by a measurable set $\mathbb{K} \subseteq X \times A$ with nonempty X-sections, which are denoted by $(x^f, x^s) \mapsto A(x^f, x^s) \subseteq A$ for each $(x^f, x^s) \in X$.

Model definition

- The *state space* X is the product $X = X^f \times X^s$ of two Borel spaces, where the components $x^f \in X^f$ and $x^s \in X^s$ are called the fast and slow states, respectively.
- The *action space* A is a Borel space and it is the union of two disjoint measurable subsets: $A = V^f \cup V^s$. The sets V^f and V^s are referred to as the fast and the slow action sets, respectively.
- The set of feasible state-action pairs is given by a measurable set $\mathbb{K} \subseteq X \times A$ with nonempty X-sections, which are denoted by $(x^f, x^s) \mapsto A(x^f, x^s) \subseteq A$ for each $(x^f, x^s) \in X$.

Model definition

- The *state space* X is the product $X = X^f \times X^s$ of two Borel spaces, where the components $x^f \in X^f$ and $x^s \in X^s$ are called the fast and slow states, respectively.
- The *action space* A is a Borel space and it is the union of two disjoint measurable subsets: $A = V^f \cup V^s$. The sets V^f and V^s are referred to as the fast and the slow action sets, respectively.
- The set of feasible state-action pairs is given by a measurable set $\mathbb{K} \subseteq X \times A$ with nonempty X-sections, which are denoted by $(x^f, x^s) \mapsto A(x^f, x^s) \subseteq A$ for each $(x^f, x^s) \in X$.

Model definition

- We assume further the existence of two measurable sets

$$D^{\wedge} \subseteq D^{\vee} \subseteq X$$

such that

(a) When the state of the system is in $D^{\wedge}$, the controller must necessarily choose an action in V^{s} (a slow action); and

(b) When the state of the system is outside $D^{\vee}$, the controller must necessarily choose an action in V^{f} (a fast action).

Model definition

- We assume further the existence of two measurable sets

$$D^{\wedge} \subseteq D^{\vee} \subseteq X$$

such that

(a) When the state of the system is in $D^{\wedge}$, the controller must necessarily choose an action in V^s (a slow action); and

(b) When the state of the system is outside $D^{\vee}$, the controller must necessarily choose an action in V^f (a fast action).

Model definition

- We assume further the existence of two measurable sets

$$D^{\wedge} \subseteq D^{\vee} \subseteq X$$

such that

(a) When the state of the system is in $D^{\wedge}$, the controller must necessarily choose an action in V^s (a slow action); and

(b) When the state of the system is outside $D^{\vee}$, the controller must necessarily choose an action in V^f (a fast action).

Dynamic of the system

The dynamic is composed by two sub-dynamics:

- One sub-dynamic is of a standard type that only affects the fast states $x^f \in X^f$ through the stochastic transition kernel

$$Q^f : \mathcal{B}(X^f) \times \big(\mathbb{K} \cap (X \times V^f)\big) \mapsto [0,1],$$

- One sub-dynamic is of a special type and it produces a transition of both components $(x^f, x^s) \in X$ following the stochastic kernel

$$Q^s : \mathcal{B}(X) \times \big(\mathbb{K} \cap (X \times V^s)\big) \mapsto [0,1].$$

- Summarizing, the whole dynamic is given by

$$\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a) = \begin{cases} Q^f(dy^f | x^f, x^s, a)\delta_{x^s}(dy^s) & \text{if } a \in V^f, \\ Q^s(dy^f \times dy^s | x^f, x^s, a) & \text{if } a \in V^s. \end{cases} \tag{9}$$

Dynamic of the system

The dynamic is composed by two sub-dynamics:

- One sub-dynamic is of a standard type that only affects the fast states $x^f \in X^f$ through the stochastic transition kernel

$$Q^f : \mathcal{B}(X^f) \times \big(\mathbb{K} \cap (X \times V^f)\big) \mapsto [0,1],$$

- One sub-dynamic is of a special type and it produces a transition of both components $(x^f, x^s) \in X$ following the stochastic kernel

$$Q^s : \mathcal{B}(X) \times \big(\mathbb{K} \cap (X \times V^s)\big) \mapsto [0,1].$$

- Summarizing, the whole dynamic is given by

$$\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a) = \begin{cases} Q^f(dy^f | x^f, x^s, a)\delta_{x^s}(dy^s) & \text{if } a \in V^f, \\ Q^s(dy^f \times dy^s | x^f, x^s, a) & \text{if } a \in V^s. \end{cases} \tag{10}$$

Dynamic of the system

The dynamic is composed by two sub-dynamics:

- One sub-dynamic is of a standard type that only affects the fast states $x^f \in X^f$ through the stochastic transition kernel

$$Q^f : \mathcal{B}(X^f) \times \big(\mathbb{K} \cap (X \times V^f)\big) \mapsto [0,1],$$

- One sub-dynamic is of a special type and it produces a transition of both components $(x^f, x^s) \in X$ following the stochastic kernel

$$Q^s : \mathcal{B}(X) \times \big(\mathbb{K} \cap (X \times V^s)\big) \mapsto [0,1].$$

- Summarizing, the whole dynamic is given by

$$\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a) = \begin{cases} Q^f(dy^f | x^f, x^s, a)\delta_{x^s}(dy^s) & \text{if } a \in V^f, \\ Q^s(dy^f \times dy^s | x^f, x^s, a) & \text{if } a \in V^s. \end{cases} \tag{11}$$

Dynamic of the system

The dynamic is composed by two sub-dynamics:

- One sub-dynamic is of a standard type that only affects the fast states $x^f \in X^f$ through the stochastic transition kernel

$$Q^f : \mathcal{B}(X^f) \times \big(\mathbb{K} \cap (X \times V^f)\big) \mapsto [0,1],$$

- One sub-dynamic is of a special type and it produces a transition of both components $(x^f, x^s) \in X$ following the stochastic kernel

$$Q^s : \mathcal{B}(X) \times \big(\mathbb{K} \cap (X \times V^s)\big) \mapsto [0,1].$$

- Summarizing, the whole dynamic is given by

$$\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a) = \begin{cases} Q^f(dy^f | x^f, x^s, a)\delta_{x^s}(dy^s) & \text{if } a \in V^f, \\ Q^s(dy^f \times dy^s | x^f, x^s, a) & \text{if } a \in V^s. \end{cases} \tag{12}$$

Remark

The dynamic system can be also also given by means of two measurable functions $F : X \times V^f \times S \to X^f$ and $G : X \times V^s \times S \to X$, with S a Borel space, where

$$\underbrace{\left(x_{k+1}^f, x_{k+1}^s\right) = \left(F\left(x_k^f, x_k^s, a_k, w_k\right), x_k^s\right)}_{\text{standard sub-dynamic}} \quad \text{if} \quad a_k \in V^f, \tag{13}$$

or

$$\underbrace{\left(x_{k+1}^f, x_{k+1}^s\right) = G\left(x_k^f, x_k^s, a_k, w_k\right)}_{\text{special sub-dynamic}} \quad \text{if} \quad a_k \in V^s, \tag{14}$$

and where $\{w_k\}$ is a sequence of i.i.d. random variables on S.

Control policies

- Define $H_0 = X$ and $H_k = \mathbb{K}^k \times X$ for $k \geq 1$, and let $H_\infty = \mathbb{K}^\infty$, all endowed with the corresponding product σ-algebras. The history up to step k is

$$h_k = \left(x_0^f, x_0^s, a_0, \cdots, x_{k-1}^f, x_{k-1}^s, a_{k-1}, x_k^f, x_k^s\right) \in H_k.$$

- A control policy is a sequence $\{\nu_k\}_{k\geq 0}$ of transition probability measures on A given H_k such that $\nu_k\left(A(x_k^f, x_k^s)|h_k\right) = 1$ for all $h_k \in H_k$.
- In particular, we necessarily have

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^s|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in D^\wedge, \text{ and}$$

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^f|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in X \setminus D^\vee.$$

- We denote by Π the set of admissible control policies.

Control policies

- Define $H_0 = X$ and $H_k = \mathbb{K}^k \times X$ for $k \geq 1$, and let $H_\infty = \mathbb{K}^\infty$, all endowed with the corresponding product σ-algebras. The history up to step k is

$$h_k = \left(x_0^f, x_0^s, a_0, \cdots, x_{k-1}^f, x_{k-1}^s, a_{k-1}, x_k^f, x_k^s\right) \in H_k.$$

- A control policy is a sequence $\{\nu_k\}_{k\geq 0}$ of transition probability measures on A given H_k such that $\nu_k\left(A(x_k^f, x_k^s)|h_k\right) = 1$ for all $h_k \in H_k$.
- In particular, we necessarily have

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^s|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in D^\wedge, \text{ and}$$

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^f|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in X \setminus D^\vee.$$

- We denote by Π the set of admissible control policies.

Control policies

- Define $H_0 = X$ and $H_k = \mathbb{K}^k \times X$ for $k \geq 1$, and let $H_\infty = \mathbb{K}^\infty$, all endowed with the corresponding product σ-algebras. The history up to step k is

$$h_k = \left(x_0^f, x_0^s, a_0, \cdots, x_{k-1}^f, x_{k-1}^s, a_{k-1}, x_k^f, x_k^s\right) \in H_k.$$

- A control policy is a sequence $\{\nu_k\}_{k\geq 0}$ of transition probability measures on A given H_k such that $\nu_k\left(A(x_k^f, x_k^s)|h_k\right) = 1$ for all $h_k \in H_k$.
- In particular, we necessarily have

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^s|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in D^\wedge, \text{ and}$$

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^f|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in X \setminus D^\vee.$$

- We denote by Π the set of admissible control policies.

Control policies

- Define $H_0 = X$ and $H_k = \mathbb{K}^k \times X$ for $k \geq 1$, and let $H_\infty = \mathbb{K}^\infty$, all endowed with the corresponding product σ-algebras. The history up to step k is

$$h_k = \left(x_0^f, x_0^s, a_0, \cdots, x_{k-1}^f, x_{k-1}^s, a_{k-1}, x_k^f, x_k^s\right) \in H_k.$$

- A control policy is a sequence $\{\nu_k\}_{k\geq 0}$ of transition probability measures on A given H_k such that $\nu_k\left(A(x_k^f, x_k^s)|h_k\right) = 1$ for all $h_k \in H_k$.
- In particular, we necessarily have

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^s|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in D^\wedge, \text{ and}$$

$$\nu_k\left(A(x_k^f, x_k^s) \cap V^f|h_k\right) = 1 \quad \text{if } (x_k^f, x_k^s) \in X \setminus D^\vee.$$

- We denote by Π the set of admissible control policies.

Control policies

- For any initial state $x = (x^f, x^s) \in X$ and any policy $\nu \in \Pi$ there exists a unique probability measure on H_∞, denoted by P_x^ν, which models the controlled dynamic system under ν. Its expectation operator is denoted by E_x^ν.
- If there is some $\mathtt{f} \in \mathbb{F}$ such that the policy $\nu \in \Pi$ satisfies $\nu_k(\cdot|h_k) = \delta_{\mathtt{f}(x_k^f, x_k^s)}(\cdot)$ for any $h_k \in H_k$ and $k \geq 0$, then we say that ν is a deterministic stationary policy.
- In what follows, we will identify the set of such policies with $\mathbb{F}$. Hence, we have $\mathbb{F} \subseteq \Pi$.

Control policies

- For any initial state $x = (x^f, x^s) \in X$ and any policy $\nu \in \Pi$ there exists a unique probability measure on H_∞, denoted by P_x^ν, which models the controlled dynamic system under ν. Its expectation operator is denoted by E_x^ν.
- If there is some $\mathtt{f} \in \mathbb{F}$ such that the policy $\nu \in \Pi$ satisfies $\nu_k(\cdot|h_k) = \delta_{\mathtt{f}(x_k^f, x_k^s)}(\cdot)$ for any $h_k \in H_k$ and $k \geq 0$, then we say that ν is a deterministic stationary policy.
- In what follows, we will identify the set of such policies with $\mathbb{F}$. Hence, we have $\mathbb{F} \subseteq \Pi$.

Control policies

- For any initial state $x = (x^f, x^s) \in X$ and any policy $\nu \in \Pi$ there exists a unique probability measure on H_∞, denoted by P_x^ν, which models the controlled dynamic system under ν. Its expectation operator is denoted by E_x^ν.
- If there is some $\mathtt{f} \in \mathbb{F}$ such that the policy $\nu \in \Pi$ satisfies $\nu_k(\cdot|h_k) = \delta_{\mathtt{f}(x_k^f, x_k^s)}(\cdot)$ for any $h_k \in H_k$ and $k \geq 0$, then we say that ν is a deterministic stationary policy.
- In what follows, we will identify the set of such policies with $\mathbb{F}$. Hence, we have $\mathbb{F} \subseteq \Pi$.

Time component

- Consider a sequence $\{\mathtt{t}_k\}$ with values in $\mathbb{N}$, such that $\mathtt{t}_k$ = the number of times that, previous to k, an action in V^f has been taken.
- Use the notation

$$\omega = \left(x_0^f, x_0^s, a_0, \cdots, x_k^f, x_k^s, a_k, \cdots\right)$$

  for an element of $H_\infty = \mathbb{K}^\infty$.
- Thus, given an arbitrary $\omega \in H_\infty$, we put $\mathtt{t}_0(\omega) = 0$ and, for each $k \geq 1$, we let

$$\mathtt{t}_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{V^f}(a_j). \tag{15}$$

Time component

- Consider a sequence $\{\mathtt{t}_k\}$ with values in $\mathbb{N}$, such that $\mathtt{t}_k =$ the number of times that, previous to k, an action in V^f has been taken.
- Use the notation

$$\omega = (x_0^f, x_0^s, a_0, \cdots, x_k^f, x_k^s, a_k, \cdots)$$

  for an element of $H_\infty = \mathbb{K}^\infty$.
- Thus, given an arbitrary $\omega \in H_\infty$, we put $\mathtt{t}_0(\omega) = 0$ and, for each $k \geq 1$, we let

$$\mathtt{t}_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{V^f}(a_j). \tag{16}$$

Time component

- Consider a sequence $\{\mathtt{t}_k\}$ with values in $\mathbb{N}$, such that $\mathtt{t}_k$ = the number of times that, previous to k, an action in V^f has been taken.
- Use the notation

$$\omega = \left(x_0^f, x_0^s, a_0, \cdots, x_k^f, x_k^s, a_k, \cdots\right)$$

for an element of $H_\infty = \mathbb{K}^\infty$.

- Thus, given an arbitrary $\omega \in H_\infty$, we put $\mathtt{t}_0(\omega) = 0$ and, for each $k \geq 1$, we let

$$\mathtt{t}_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{V^f}(a_j). \tag{17}$$

Optimality criterion

- We will consider a running cost function $\mathbf{c} : \mathbb{K} \to [0, \infty)$ which will be written

$$\mathbf{c}(x^f, x^s, a) = c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a),$$

with

$$c : \mathbb{K} \cap (X \times V^f) \mapsto [0, \infty) \quad \text{and} \quad \ell : \mathbb{K} \cap (X \times V^s) \mapsto [0, \infty)$$

interpreted as running cost functions for the standard and the special sub dynamics, respectively.

Optimality criterion

- For an initial state $(x^f, x^s) \in X$ and a control policy $\nu \in \Pi$, let

$$J(x^f, x^s, \nu) = \limsup_{n\to\infty} \frac{E^{\nu}_{x^f,x^s}\left[\sum_{k=0}^{n-1} \mathbf{c}(x^f_k, x^s_k, a_k)\right]}{E^{\nu}_{x^f,x^s}[\mathrm{t}_n]}, \qquad (18)$$

- The optimal ratio-average cost function is then defined as

$$J^*(x^f, x^s) = \inf_{\nu\in\Pi} J(x^f, x^s, \nu) \quad \text{for } (x^f, x^s) \in X, \qquad (19)$$

- We will say that $\nu^* \in \Pi$ is ratio-average optimal when

$$J(x^f, x^s, \nu^*) = J^*(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X$$

.

Optimality criterion

- For an initial state $(x^f, x^s) \in X$ and a control policy $\nu \in \Pi$, let

$$J(x^f, x^s, \nu) = \limsup_{n\to\infty} \frac{E^{\nu}_{x^f,x^s}\left[\sum_{k=0}^{n-1} \mathbf{c}(x^f_k, x^s_k, a_k)\right]}{E^{\nu}_{x^f,x^s}[\mathrm{t}_n]}, \qquad (20)$$

- The optimal ratio-average cost function is then defined as

$$J^*(x^f, x^s) = \inf_{\nu\in\Pi} J(x^f, x^s, \nu) \quad \text{for } (x^f, x^s) \in X, \qquad (21)$$

- We will say that $\nu^* \in \Pi$ is ratio-average optimal when

$$J(x^f, x^s, \nu^*) = J^*(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X$$

.

Optimality criterion

- For an initial state $(x^f, x^s) \in X$ and a control policy $\nu \in \Pi$, let

$$J(x^f, x^s, \nu) = \limsup_{n\to\infty} \frac{E^{\nu}_{x^f, x^s}\left[\sum_{k=0}^{n-1} \mathbf{c}(x^f_k, x^s_k, a_k)\right]}{E^{\nu}_{x^f, x^s}[\mathrm{t}_n]}, \qquad (22)$$

- The optimal ratio-average cost function is then defined as

$$J^*(x^f, x^s) = \inf_{\nu\in\Pi} J(x^f, x^s, \nu) \quad \text{for } (x^f, x^s) \in X, \qquad (23)$$

- We will say that $\nu^* \in \Pi$ is ratio-average optimal when

$$J(x^f, x^s, \nu^*) = J^*(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X$$

.

Assumptions

Assumption

For each $(x^f, x^s) \in X$ we have:

(i) *The action set $A(x^f, x^s)$ is compact.*

(ii) *Given a bounded and measurable function $u : X \to \mathbb{R}$, the function*

$$a \mapsto \int_X u(y^f, y^s) \mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$.

(iii) *The cost function $\mathbf{c}$ is in $\mathbb{B}^+(\mathbb{K})$ and the function $a \mapsto \mathbf{c}(x^f, x^s, a)$ is continuous on $A(x^f, x^s)$ for each fixed $(x^f, x^s) \in X$.*

(iv) *There exists a constant $\ell_0 > 0$ with $\ell(x^f, x^s, a) \geq \ell_0$ for all $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$.*

Assumptions

Assumption

For each $(x^f, x^s) \in X$ *we have:*

(i) *The action set* $A(x^f, x^s)$ *is compact.*

(ii) *Given a bounded and measurable function* $u : X \to \mathbb{R}$*, the function*

$$a \mapsto \int_X u(y^f, y^s)\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$.

(iii) *The cost function* $\mathbf{c}$ *is in* $\mathbb{B}^+(\mathbb{K})$ *and the function* $a \mapsto \mathbf{c}(x^f, x^s, a)$ *is continuous on* $A(x^f, x^s)$ *for each fixed* $(x^f, x^s) \in X$.

(iv) *There exists a constant* $\ell_0 > 0$ *with* $\ell(x^f, x^s, a) \geq \ell_0$ *for all* $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$.

Assumptions

Assumption

For each $(x^f, x^s) \in X$ *we have:*

(i) *The action set* $A(x^f, x^s)$ *is compact.*

(ii) *Given a bounded and measurable function* $u : X \to \mathbb{R}$, *the function*

$$a \mapsto \int_X u(y^f, y^s)\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$.

(iii) *The cost function* $\mathbf{c}$ *is in* $\mathbb{B}^+(\mathbb{K})$ *and the function* $a \mapsto \mathbf{c}(x^f, x^s, a)$ *is continuous on* $A(x^f, x^s)$ *for each fixed* $(x^f, x^s) \in X$.

(iv) *There exists a constant* $\ell_0 > 0$ *with* $\ell(x^f, x^s, a) \geq \ell_0$ *for all* $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$.

Assumptions

Assumption

For each $(x^f, x^s) \in X$ *we have:*

(i) *The action set* $A(x^f, x^s)$ *is compact.*

(ii) *Given a bounded and measurable function* $u : X \to \mathbb{R}$*, the function*

$$a \mapsto \int_X u(y^f, y^s)\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$.

(iii) *The cost function* $\mathbf{c}$ *is in* $\mathbb{B}^+(\mathbb{K})$ *and the function* $a \mapsto \mathbf{c}(x^f, x^s, a)$ *is continuous on* $A(x^f, x^s)$ *for each fixed* $(x^f, x^s) \in X$.

(iv) *There exists a constant* $\ell_0 > 0$ *with* $\ell(x^f, x^s, a) \geq \ell_0$ *for all* $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$.

Assumptions

Assumption

For each $(x^f, x^s) \in X$ *we have:*

(i) *The action set* $A(x^f, x^s)$ *is compact.*

(ii) *Given a bounded and measurable function* $u : X \to \mathbb{R}$, *the function*

$$a \mapsto \int_X u(y^f, y^s)\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$.

(iii) *The cost function* $\mathbf{c}$ *is in* $\mathbb{B}^+(\mathbb{K})$ *and the function* $a \mapsto \mathbf{c}(x^f, x^s, a)$ *is continuous on* $A(x^f, x^s)$ *for each fixed* $(x^f, x^s) \in X$.

(iv) *There exists a constant* $\ell_0 > 0$ *with* $\ell(x^f, x^s, a) \geq \ell_0$ *for all* $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$.

Assumptions

Assumption

There exists a control policy $\mathtt{f} \in \mathbb{F}$, *and constants* $q \in \mathbb{N}$, $\epsilon > 0$ *such that*

$$\mathbf{Q}^q(D^c | x^f, x^s, \mathtt{f}) \geq \epsilon \quad \textit{for all } (x^f, x^s) \in X,$$

where $D = \{(x^f, x^s) \in X : \mathtt{f}(x^f, x^s) \in V^s\}$.

Preliminary results

Proposition (Jasso-Fuentes et. al. (2018))

- *If* $\mathtt{f} \in \mathbb{F}$ *is the control policy defined in Assumption 6 then*

$$\liminf_{n\to\infty} \frac{E_x^{\mathtt{f}}[\mathtt{t}_n]}{n} > 0 \quad \textit{for all } x = (x^f, x^s) \in X.$$

- *Moreover, the criterion (22) is uniformly bounded if Assumption 5 is also satisfied; in fact,*

$$J(x, \mathtt{f}) \leq \frac{q||\mathbf{c}||}{\epsilon},$$

and so $J(x, \mathtt{f})$ *is finite for every* $x \in X$*, and* $J^*(x) \leq \frac{q||\mathbf{c}||}{\epsilon}$ *for all* $x \in X$*.*

Preliminary results

Proposition (Jasso-Fuentes et. al. (2018))

- *If* $\mathtt{f} \in \mathbb{F}$ *is the control policy defined in Assumption 6 then*

$$\liminf_{n\to\infty} \frac{E_x^{\mathtt{f}}[\mathtt{t}_n]}{n} > 0 \quad \text{for all } x = (x^f, x^s) \in X.$$

- *Moreover, the criterion (22) is uniformly bounded if Assumption 5 is also satisfied; in fact,*

$$J(x, \mathtt{f}) \leq \frac{q||\mathbf{c}||}{\epsilon},$$

and so $J(x,\mathtt{f})$ *is finite for every* $x \in X$, *and* $J^*(x) \leq \frac{q||\mathbf{c}||}{\epsilon}$ *for all* $x \in X$.

Preliminary results

- For a general model $(\mathbf{X}, \mathbf{A}, \mathbf{A}(x), x \in \mathbf{X}, \mathbf{Q}, \mathbf{c})$
- Average payoff

$$\mathbf{J}(x, \nu) = \limsup_{n \to \infty} \frac{1}{n} E_x^{\nu} \Big[\sum_{k=0}^{n-1} \mathbf{c}(x_k, a_k) \Big]$$

- The associated DPE to this criterion is

$$\lambda + u(x) = \inf_{a \in A(x)} \Big\{ \mathbf{c}(x, a) + \int_{\mathbf{X}} u(y) \mathbf{Q}(dt | x, a) \Big\}.$$

- Iterating, we have for all ν

$$n\lambda + u(x) \leq E_x^{\nu} \Big[\sum_{k=0}^{n-1} \mathbf{c}(x_k, a_k) \Big] + E_x^{\nu} [u(x_n)].$$

Preliminary results

- Going back to the DPE:

$$\lambda + u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}.$$

If and only if

$$u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) - \lambda + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}$$

Preliminary results

- Going back to the DPE:

$$\lambda + u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) + \int_{\hat{X}} u(y)\mathbf{Q}(dt|x,a) \right\}.$$

If and only if

$$u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) - \lambda + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}$$

Preliminary results

- Going back to the DPE:

$$\lambda + u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}.$$

If and only if

$$u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) - \lambda + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}$$

- In our case, the good equation becomes

$$u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}(x,a) - \lambda \mathbf{1}_{V^f}(a) + \int_{\mathbf{X}} u(y)\mathbf{Q}(dt|x,a) \right\}$$

Preliminary results

- Going back to the DPE:

$$\lambda + u(x) = \inf_{a \in A(x)} \Big\{ \mathbf{c}(x,a) + \int_{\mathbf{X}} u(y) \mathbf{Q}(dt|x,a) \Big\}.$$

If and only if

$$u(x) = \inf_{a \in A(x)} \Big\{ \mathbf{c}(x,a) - \lambda + \int_{\mathbf{X}} u(y) \mathbf{Q}(dt|x,a) \Big\}$$

- In our case, the good equation becomes

$$u(x) = \inf_{a \in A(x)} \Big\{ \mathbf{c}(x,a) - \lambda \mathbf{1}_{V^f}(a) + \int_{\mathbf{X}} u(y) \mathbf{Q}(dt|x,a) \Big\}$$

- Because iterating, we have for all $\nu \in \Pi$

$$\lambda E_x^{\nu}[\mathrm{t}_n] + u(x) \leq E_x^{\nu}\Big[\sum_{k=0}^{n-1} \mathbf{c}(x_k, a_k)\Big] + E_x^{\nu}[u(x_n)].$$

Dynamic programming equations

- We define the operators

$$\mathcal{M}u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^s} \left\{ \ell(x^f, x^s, a) + \int_X u(y^f, y^s) Q^s(dy^f \times dy^s \mid x^f, x^s, a) \right\} \quad (24)$$

  defined on $D^\vee$,

- and

$$\mathcal{H}_\lambda u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^f} \left\{ c(x^f, x^s, a) - \lambda + \int_{X^f} u(y^f, x^s) Q^f(dy^f | x^f, x^s, a) \right\} \quad (25)$$

  defined on $X \setminus D^\wedge$.

Dynamic programming equations

- We define the operators

$$\mathcal{M}u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^s} \left\{ \ell(x^f, x^s, a) + \int_X u(y^f, y^s) Q^s(dy^f \times dy^s \mid x^f, x^s, a) \right\} \tag{26}$$

  defined on $D^\vee$,

- and

$$\mathcal{H}_\lambda u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^f} \left\{ c(x^f, x^s, a) - \lambda + \int_{X^f} u(y^f, x^s) Q^f(dy^f | x^f, x^s, a) \right\} \tag{27}$$

  defined on $X \setminus D^\wedge$.

Dynamic programming equations

- Therefore, the dynamic programming operator $\mathcal{T}_\lambda$ can be written, for $u \in \mathbb{B}(X)$ and $(x^f, x^s) \in X$, as

$$\mathcal{T}_\lambda u(x^f, x^s) = \begin{cases} \mathcal{M}u(x^f, x^s), & \text{if } (x^f, x^s) \in D^\wedge, \\ \min\left\{\mathcal{M}u(x^f, x^s), \mathcal{H}_\lambda u(x^f, x^s)\right\}, & \text{if } (x^f, x^s) \in D^\vee \setminus D^\wedge, \\ \mathcal{H}_\lambda u(x^f, x^s), & \text{if } (x^f, x^s) \in X \setminus D^\vee. \end{cases} \tag{28}$$

- A function $u \in \mathbb{B}(X)$ is said to be a solution of the dynamic programming equation (DPE) when

$$u(x^f, x^s) = \mathcal{T}_\lambda u(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X, \tag{29}$$

which will be written, in short, as $u = \mathcal{T}_\lambda u$.

Dynamic programming equations

- Therefore, the dynamic programming operator $\mathcal{T}_\lambda$ can be written, for $u \in \mathbb{B}(X)$ and $(x^f, x^s) \in X$, as

$$\mathcal{T}_\lambda u(x^f, x^s) = \begin{cases} \mathcal{M}u(x^f, x^s), & \text{if } (x^f, x^s) \in D^\wedge, \\ \min\left\{\mathcal{M}u(x^f, x^s), \mathcal{H}_\lambda u(x^f, x^s)\right\}, & \text{if } (x^f, x^s) \in D^\vee \setminus D^\wedge, \\ \mathcal{H}_\lambda u(x^f, x^s), & \text{if } (x^f, x^s) \in X \setminus D^\vee. \end{cases} \tag{30}$$

- A function $u \in \mathbb{B}(X)$ is said to be a solution of the dynamic programming equation (DPE) when

$$u(x^f, x^s) = \mathcal{T}_\lambda u(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X, \tag{31}$$

which will be written, in short, as $u = \mathcal{T}_\lambda u$.

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda\mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$ $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X turns out to be:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda\mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$ $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X turns out to be:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda \mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$
  $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X turns out to be:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda \mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$ $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X turns out to be:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda\mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$ $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X is:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Dynamic programming equations

For any $(x^f, x^s) \in X$, $a \in A(x^f, x^s)$, we let

- $\mathbf{c}(x^f, x^s, a) := c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a)$,
- $\mathbf{c}_\lambda(x^f, x^s, a) := \mathbf{c}(x^f, x^s, a) - \lambda\mathbf{1}_{V^f}(a)$,
- $\mathbf{Q}(\cdot|x^f, x^s, a) :=$
  $Q^f(dy^f|x^f, x^s, a)\delta_{x^s}(dy^s)\mathbf{1}_{V^f}(a) + Q^s(\cdot|x^f, x^s, a)\mathbf{1}_{V^s}(a)$

Given a function $u \in \mathbb{B}(X)$ and using the notation $x = (x^f, x^s)$, the dynamic programming operator $\mathcal{T}_\lambda u$ on X turns out to be:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x, a) + \int_X u(y)\mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

together with the DPE

$$u(x) = \mathcal{T}_\lambda u(x) \quad \text{for all } x \in X,$$

Existence of solutions

Assumption

There exists $0 < \beta < 1$ *such that*

$$\sup_{(x,a),(x',a')\in\mathbb{K}} ||\mathbf{Q}(\cdot|x,a) - \mathbf{Q}(\cdot|x',a')||_{TV} \leq 2\beta,$$

where the norm $\|\cdot\|_{TV}$ *is the total variation norm.*

Existence of solutions

Lemma (Hernández-Lerma (1989) and Jasso-Fuentes et. al. (2018))

Suppose that Assumptions5, 6, and 7 are satisfied. Then the following assertions hold:

(i) *For each* $\lambda \geq 0$, *there exists a pair* $(\rho_\lambda, u_\lambda) \in \mathbb{R} \times \mathbb{B}(X)$ *that satisfies the DPE* (31); *i.e.,* $\rho_\lambda + u_\lambda = \mathcal{T}_\lambda u_\lambda$. *Furthermore, the constant* ρ_λ *satisfies:*

$$\rho_\lambda = \inf_{\nu \in \Pi} \limsup_{n \to \infty} \frac{1}{n} E_x^\nu \left[\sum_{k=0}^{n-1} [\mathbf{c}(x_k, a_k) - \lambda \mathbf{1}_{V^f}(a_k)] \right].$$

(ii) *There exists* $\lambda^* \geq 0$ *such that* $\rho_{\lambda^*} = 0$, *for which the DPE* (31) *becomes*

$$u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}. \tag{32}$$

Existence of solutions

Lemma (Hernández-Lerma (1989) and Jasso-Fuentes et. al. (2018))

Suppose that Assumptions5, 6, and 7 are satisfied. Then the following assertions hold:

(i) *For each $\lambda \geq 0$, there exists a pair $(\rho_\lambda, u_\lambda) \in \mathbb{R} \times \mathbb{B}(X)$ that satisfies the DPE (31); i.e., $\rho_\lambda + u_\lambda = \mathcal{T}_\lambda u_\lambda$. Furthermore, the constant ρ_λ satisfies:*

$$\rho_\lambda = \inf_{\nu \in \Pi} \limsup_{n \to \infty} \frac{1}{n} E_x^\nu \left[\sum_{k=0}^{n-1} [\mathbf{c}(x_k, a_k) - \lambda \mathbf{1}_{V^f}(a_k)] \right].$$

(ii) *There exists $\lambda^* \geq 0$ such that $\rho_{\lambda^*} = 0$, for which the DPE (31) becomes*

$$u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}. \tag{33}$$

Existence of solutions

Lemma (Hernández-Lerma (1989) and Jasso-Fuentes et. al. (2018))

Suppose that Assumptions5, 6, and 7 are satisfied. Then the following assertions hold:

(i) *For each $\lambda \geq 0$, there exists a pair $(\rho_\lambda, u_\lambda) \in \mathbb{R} \times \mathbb{B}(X)$ that satisfies the DPE (31); i.e., $\rho_\lambda + u_\lambda = \mathcal{T}_\lambda u_\lambda$. Furthermore, the constant ρ_λ satisfies:*

$$\rho_\lambda = \inf_{\nu \in \Pi} \limsup_{n \to \infty} \frac{1}{n} E_x^\nu \left[\sum_{k=0}^{n-1} [\mathbf{c}(x_k, a_k) - \lambda \mathbf{1}_{V^f}(a_k)] \right].$$

(ii) *There exists $\lambda^* \geq 0$ such that $\rho_{\lambda^*} = 0$, for which the DPE (31) becomes*

$$u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}. \tag{34}$$

Existence of solutions

Definition

Any policy $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (34) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called canonical.*

Theorem (Jasso-Fuentes et. al. (2018))

Suppose that Assumptions 6, 5, and 7 hold. Then the following statements hold true.

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x, \nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J^* *in* (23) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Existence of solutions

Definition

Any policy $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (34) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called* *canonical*.

Theorem (Jasso-Fuentes et. al. (2018))

Suppose that Assumptions 6, 5, and 7 hold. Then the following statements hold true.

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x, \nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J^* *in* (23) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Existence of solutions

Definition

Any policy $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (34) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called canonical.*

Theorem (Jasso-Fuentes et. al. (2018))

Suppose that Assumptions 6, 5, and 7 hold. Then the following statements hold true.

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x, \nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J^* *in* (23) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Existence of solutions

Definition

Any policy $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (34) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called* *canonical*.

Theorem (Jasso-Fuentes et. al. (2018))

Suppose that Assumptions 6, 5, and 7 hold. Then the following statements hold true.

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x, \nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J^* *in* (23) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Existence of solutions

Definition

Any policy $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (34) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called* *canonical*.

Theorem (Jasso-Fuentes et. al. (2018))

Suppose that Assumptions 6, 5, and 7 hold. Then the following statements hold true.

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x, \nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J^* *in* (23) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Existence of solutions

The following is a direct consequence of both Theorem 5 and the definition of a canonical policy.

Corollary

Let $\mathtt{f} \in \mathbb{F}$ *be a canonical policy. On the set* $D^{\vee} \setminus D^{\wedge}$*, the following holds:*

(a). *If* $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^f*.*

(b). *If* $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^s*.*

Existence of solutions

The following is a direct consequence of both Theorem 5 and the definition of a canonical policy.

Corollary

Let $\mathtt{f} \in \mathbb{F}$ *be a canonical policy. On the set* $D^{\vee} \setminus D^{\wedge}$, *the following holds:*

(a). *If* $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*} u^*(x) = u^*(x)$, *then the optimal action* $\mathtt{f}(x)$ *is in* V^f.

(b). *If* $\mathcal{H}_{\lambda^*} u^*(x) > \mathcal{M}u^*(x) = u^*(x)$, *then the optimal action* $\mathtt{f}(x)$ *is in* V^s.

Existence of solutions

The following is a direct consequence of both Theorem 5 and the definition of a canonical policy.

Corollary

Let $\mathtt{f} \in \mathbb{F}$ *be a canonical policy. On the set* $D^{\vee} \setminus D^{\wedge}$*, the following holds:*

(a). *If* $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^f*.*

(b). *If* $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^s*.*

Existence of solutions

The following is a direct consequence of both Theorem 5 and the definition of a canonical policy.

Corollary

Let $\mathtt{f} \in \mathbb{F}$ *be a canonical policy. On the set* $D^{\vee} \setminus D^{\wedge}$*, the following holds:*

(a). *If* $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^f*.*

(b). *If* $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^s*.*

Existence of solutions

Corollary

Let $\mathtt{f} \in \mathbb{F}$ *be a canonical policy. On the set* $D^{\vee} \setminus D^{\wedge}$*, the following holds:*

(a). *If* $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^f*.*

(b). *If* $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$*, then the optimal action* $\mathtt{f}(x)$ *is in* V^s*.*

Remark

(i) *When* $\mathcal{M}u^*(x) = \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$*, then the controller can run either the usual or the special sub-dynamic.*

(ii) *Obviously, when the state is in* $D^{\wedge}$ *or in* $X \setminus D^{\vee}$*, then optimal actions are necessarily in* V^s *or in* V^f*, respectively.*

Existence of solutions

Corollary

Let $\mathtt{f} \in \mathbb{F}$ be a canonical policy. On the set $D^{\vee} \setminus D^{\wedge}$, the following holds:

(a). *If $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$, then the optimal action $\mathtt{f}(x)$ is in V^f.*

(b). *If $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$, then the optimal action $\mathtt{f}(x)$ is in V^s.*

Remark

(i) *When $\mathcal{M}u^*(x) = \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$, then the controller can run either the usual or the special sub-dynamic.*

(ii) *Obviously, when the state is in $D^{\wedge}$ or in $X \setminus D^{\vee}$, then optimal actions are necessarily in V^s or in V^f, respectively.*

Contact set and continuation region

We introduce the set D^* so-named **contact set** defined as:

$$D^* = \{x \in D^\vee : u^*(x) = \mathcal{M}u^*(x)\},$$

with $x := (x^f, x^s)$.

Proposition

An optimal rule outside D^* must necessarily be in V^f, whereas inside D^* the optimal rule can be taken in V^s.

Contact set and continuation region

We introduce the set D^* so-named **contact set** defined as:

$$D^* = \{x \in D^\vee : u^*(x) = \mathcal{M}u^*(x)\},$$

with $x := (x^f, x^s)$.

Proposition

An optimal rule outside D^* must necessarily be in V^f, whereas inside D^* the optimal rule can be taken in V^s.

Earliest references

- Branicky, M. S.: Studies in hybrid systems: Modeling, analysis, and control, Thesis MIT, LIDS-TH-2304 (1995), Cambridge, USA.
- Bensoussan, A., Menaldi J.L.: (1997) Hybrid control and dynamic programming. *Dynamics and continuous, discrete and impulsive systems* **3**, pp. 395–442.
- Branicky, M.S; Borkar, V.S; Mitter, S.K.: (1998) A unified framework of hybrid control: model and optimal control theory *IEEE Trans. Autom. Control* **43**, pp. 31–45.
- Borkar, V.S; Ghosh, M.K., Sahay, P.: (1999) Optimal control of a stochastic hybrid system with discounted cost *J. Optim. Theory Appl* **101**, pp. 557–580.
- Bensoussan, A., Menaldi J.L.: (2000) Stochastic hybrid control *J. Math. Anal. Appl.* **249**, pp. 261–288.

Main reference

Jasso-Fuentes, H.; Menaldi, J.L.; Prieto-Rumeau; T., Robin, M. (2018) *Discrete-time hybrid control in Borel spaces: average cost optimality criterion.* J. Math. Anal. Appl. **462**, pp. 1695-1713.

Thank you!