# Optimal control for diffusion processes with Markovian switching modelled as a differential game against nature

Beatris Adriana Escobedo-Trujillo

Carmen Geraldi Higuera-Chan

José Daniel López-Barrientos

July the 28th, 2020

STONY BROOK CENTER FOR GAME THEORY

Anáhuac México

# About the speaker

**José Daniel López-Barrientos**

- BSc in Actuarial Sciences, MSc and PhD in Mathematics
- Researcher at Universidad Anáhuac México; invited researcher at HEC, Montréal; Saint Petersburg State University; Institut Interafricain de Formation en Assurance et en Gestion des Entreprises, Dakar, Senegal.

**Facultad de Ciencias Actuariales/Universidad Anáhuac México**

- Only Faculty of Actuarial Sciences in Latin America.
- In the top 3 of best universities in Mexico (https://www.topuniversities.com/university-rankings/world-university-rankings/2020)

Anáhuac México

# Contents

# 1. Presentation

In general, references deal with the case of completely observable stochastic control with a stochastic integral equation of the form

$$x(t) = x_0 + \int_0^t b(x(s), \theta(s), u(s))\mathrm{d}s + \int_0^t \sigma(x(s), \theta(s))\mathrm{d}W(s),$$

and a continuous-time Markov chain $\theta(t)$ with finite state space $E = \{1, 2, \dots, N\}$, whose transition rule is as follows:

$$\mathbb{P}(\theta(t + \Delta t) = j | \theta(t) = i, (x(s), \theta(s)), s \leq t) = q_{ij}\Delta t + o(\Delta t), \; i \neq j; \tag{1.1}$$

for $t \geq 0$, $\theta(0) = \theta_0$, and $\sum_{j=1}^N q_{ij} = 0$, where $Q = (q_{ij})_{i,j \in E}$ is the *rate matrix* of the process $\theta$.

We work with the particular class of controlled diffusion processes on $\mathbb{R}^n$ with infinite horizon studied in [1,2], whose dynamics has the form

$$dx(t) = b(x(t), \theta(t), u(t), \alpha(t))dt + \sigma(x(t), \theta(t))dW(t); \quad x(0) = x_0, \tag{1.2}$$

along with the transition rule (1.1); where $b$ and $\sigma$ are given functions, but the drift coefficient $b$ depends on an unknown and possibly non-observable parameter $\alpha$.

# 2. Optimal pollution control with average payoff

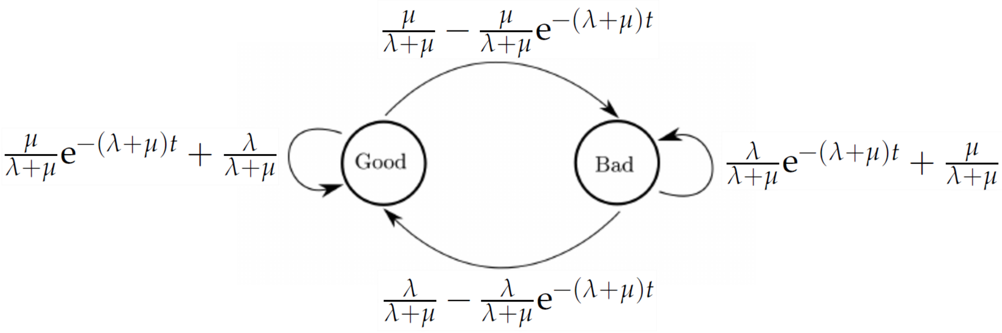
Consider the **pollution process** (cf. [5,7]) defined by the controlled diffusion

$$dx(t) = [u(t) - \alpha(t)x(t)]dt + k\,dW(t), x(0) = x > 0, \tag{2.1}$$

where

- $u(t)$ is the consumption flow at time $t \geq 0$,
- $0 \leq u(t) \leq \eta$ in "Bad state",
- $\eta$ is a consumption restriction imposed by local government,
- $\eta \leq u(t) \leq \gamma$ in "Good state",
- $\gamma$ is a consumption restriction imposed by worldwide protocols.

## 2.1. Dynamics of the “Doomsday pendulum”

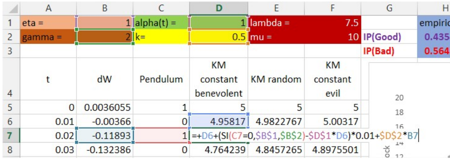



We readily know that, for $i = \text{Bad}, \text{Good}$, the limiting probabilities are given by:

$$\mathbb{P}^*(\text{Bad}) := \lim_{t\to\infty} P_{i,\text{Bad}}(t) = \frac{\mu}{\lambda+\mu}, \quad (2.2)$$

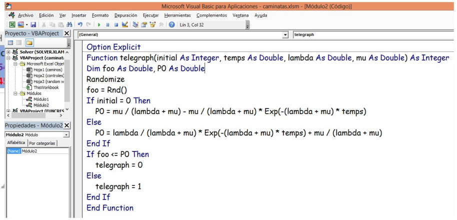
$$\mathbb{P}^*(\text{Good}) := \lim_{t\to\infty} P_{i,\text{Good}}(t) = \frac{\lambda}{\lambda+\mu}. \quad (2.3)$$

# 2.2. Three comparisons

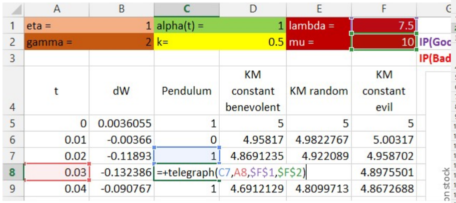


| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | eta = | 1 | alpha(t) = | 1 | lambda = | 7.5 | | empiric |
| 2 | gamma = | 2 | k= | 0.5 | mu = | 10 | IP(Good) | 0.435 |
| 3 | | | | | | | IP(Bad) | 0.564 |
| 4 | t | dW | Pendulum | KM constant benevolent | KM random | KM constant evil | | |
| 5 | 0 | 0.0036055 | 1 | 5 | 5 | 5 | | |
| 6 | 0.01 | -0.00366 | 0 | 4.95817 | 4.9822767 | 5.00317 | | |
| 7 | 0.02 | -0.11893 | 1 | =+D6+(SI(C7=0,$B$1,$B$2)-$D$1*D6)*0.01+$D$2*B7 | | | | |
| 8 | 0.03 | -0.132386 | 0 | 4.764239 | 4.8457265 | 4.8975501 | | |

Euler-Maruyama's method (cf. [3])

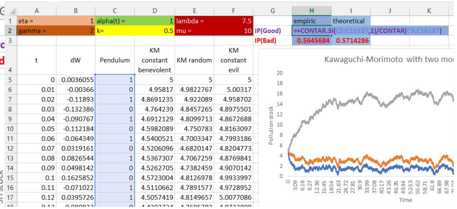



A random telegraphic signal



| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | eta = | 1 | alpha(t) = | 1 | lambda = | 7.5 | |
| 2 | gamma = | 2 | k= | 0.5 | mu = | 10 | IP(Goc |
| 3 | | | | | | | IP(Bad |
| 4 | t | dW | Pendulum | KM constant benevolent | KM random | KM constant evil | |
| 5 | 0 | 0.0036055 | 1 | 5 | 5 | 5 | |
| 6 | 0.01 | -0.00366 | 0 | 4.95817 | 4.9822767 | 5.00317 | |
| 7 | 0.02 | -0.11893 | 1 | 4.8691235 | 4.922089 | 4.958702 | |
| 8 | 0.03 | -0.132386 | =+telegraph(C7,A8,$F$1,$F$2) | | | 4.8975501 | |
| 9 | 0.04 | -0.090767 | 1 | 4.6912129 | 4.8099713 | 4.8672688 | |

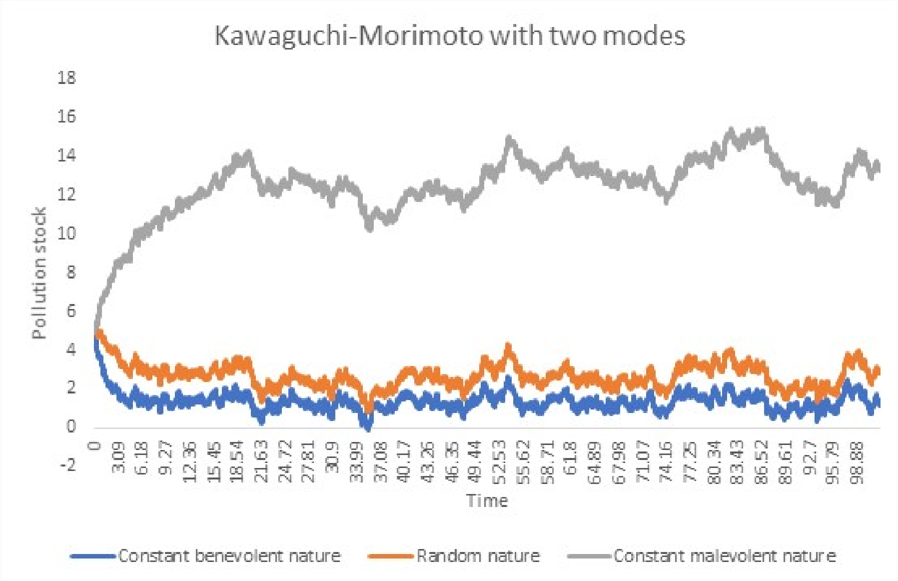
A telegram



| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | eta = | 1 | alpha(t) = | 1 | lambda = | 7.5 | | empiric | theoretical |
| 2 | gamma = | 2 | k= | 0.5 | mu = | 10 | IP(Good) | =+CONTAR.SI(C5:C10187,1)/CONTAR(C5:C10187) | |
| 3 | | | | | | | IP(Bad) | 0.5645684 | 0.5714286 |
| 4 | t | dW | Pendulum | KM constant benevolent | KM random | KM constant evil | | | |
| 5 | 0 | 0.0036055 | 1 | 5 | 5 | 5 | | | |
| 6 | 0.01 | -0.00366 | 0 | 4.95817 | 4.9822767 | 5.00317 | | | |
| 7 | 0.02 | -0.11893 | 1 | 4.8691235 | 4.922089 | 4.958702 | | | |
| 8 | 0.03 | -0.132386 | 0 | 4.764239 | 4.8457265 | 4.8975501 | | | |
| 9 | 0.04 | -0.090767 | 1 | 4.6912129 | 4.8099713 | 4.8672688 | | | |
| 10 | 0.05 | -0.112184 | 0 | 4.5982089 | 4.750783 | 4.8163097 | | | |
| 11 | 0.06 | -0.064349 | 1 | 4.5400521 | 4.7003347 | 4.7993186 | | | |
| 12 | 0.07 | 0.0319161 | 0 | 4.5206096 | 4.6820147 | 4.8204773 | | | |
| 13 | 0.08 | 0.0826544 | 1 | 4.5367307 | 4.7067259 | 4.8769841 | | | |
| 14 | 0.09 | 0.0498142 | 0 | 4.5262705 | 4.7382459 | 4.9070142 | | | |
| 15 | 0.1 | 0.1625852 | 0 | 4.5723004 | 4.8126978 | 4.9933997 | | | |
| 16 | 0.11 | -0.071022 | 1 | 4.5110662 | 4.7891577 | 4.9728952 | | | |
| 17 | 0.12 | 0.0395726 | 1 | 4.5057419 | 4.8149657 | 5.0077086 | | | |



Empiric corroboration of (2.2)-(2.3)


Kawaguchi-Morimoto with two modes
Pollution stock
18
16
14
12
10
8
6
4
2
0
-2
0
3.09
6.18
9.27
12.36
15.45
18.54
21.63
24.72
27.81
30.9
33.99
37.08
40.17
43.26
46.35
49.44
52.53
55.62
58.71
61.8
64.89
67.98
71.07
74.16
77.25
80.34
83.43
86.52
89.61
92.7
95.79
98.88
Time
Constant benevolent nature
Random nature
Constant malevolent nature

## 2.3. Utility *vs.* Disutility

**Social welfare:** $r(x,i,u) := F(u) - D(x,i)$,

- where $F \in \mathcal{C}^2[0;\infty[$ is the *social utility of the consumption* $u$, and
- $D \in \mathcal{C}([0;\infty[\times\{\text{Good},\text{Bad}\})$ is the *social disutility of the pollution stock* $x$.

$$\begin{cases} F' \geq 0, & F'' \leq 0, \\ F'(\infty) = F(0) = 0, & F'(0+) = F(\infty) = \infty, \\ D(x,i) \geq 0 & \text{convex and locally Lipschitz for } i \in \{\text{Good},\text{Bad}\}. \end{cases}$$

We look for a consumption policy $u$ that **maximizes** the *worst* long-run average welfare $J(x,i,f,\alpha)$ when the unknown process takes the value $\alpha(t) \in [0;a]$:

$$J(x,i,u,\alpha) := \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}^{u,\alpha}_{x,i} \left[ \int_0^T [F(u) - D(x,i)]\mathrm{d}t \right]. \tag{2.4}$$

**Algorithm 1:** Itô's integral

**Data:** $x_0$, $\text{Pendulum}_0$, $\lambda$, $\mu$, $\mathrm{d}t$, $T < \infty$, $\lambda$, $\mu$, $k$

**Result:** The integral inside the expectation operator (2.4)

$x \leftarrow x_0$; $\text{Pendulum} \leftarrow \text{Pendulum}_0$;

$r \leftarrow F(u(x, \text{Pendulum})) - D(x, \text{Pendulum})$; $I \leftarrow r$; $j \leftarrow 0$;

**while** $j \leq T$ **do**

  $\text{Pendulum} \leftarrow \texttt{telegraph}(\text{Pendulum}, j, \lambda, \mu)$; $\mathrm{d}W \leftarrow N^{-1}(0, \mathrm{d}t)$;

  $x \leftarrow x + (u(x, \text{Pendulum}) - \alpha(x) \cdot x)\mathrm{d}t + k \cdot \mathrm{d}W$;

  $r \leftarrow F(u(x, \text{Pendulum})) - D(x, \text{Pendulum})$; $I \leftarrow I + r$;

  $j \leftarrow j + \mathrm{d}t$;

**end**

$I \leftarrow I \cdot \mathrm{d}t$;

**return** $I$;

**Algorithm 2:** Monte Carlo algorithm

**Data:** $x_0$, Pendulum$_0$, $\lambda$, $\mu$, d$t$, $T < \infty$, $\lambda$, $\mu$, $k$, $N$

**Result:** An approximate of (2.4), that is, the average of $N$ iterations of Algorithm 1 divided by $T$

MC $\leftarrow$ 0;

**for** $i \leftarrow 0$ **to** $N$ **do**

$\quad$ MC $\leftarrow \frac{(i-1)\cdot \text{MC}+\texttt{Integral}(x_0,\text{Pendulum}_0,\lambda,\mu,\text{d}t,T,\lambda,\mu,k)}{i}$

**end**

**return** $\frac{\text{MC}}{T}$;

I2 | =+mc($F$5,$C$5,0,$F$1,$F$2,0.01,1000,10)/1000

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | eta = | 1 | a0 | 1 | lambda = | 15 | Monte Carlo Expectations | | | | empiric | theoretical |
| 2 | gamma = | 2 | a1 | 0.1 | mu = | 5 | 0.1765727 | 0.1746823 | -284.2196 | IP(Good) | 0.7399588 | 0.75 |
| 3 | k= | 0.5 | | | | one Itô's integral | 0.7228799 | 0.1486899 | -4.086426 | IP(Bad) | 0.2600412 | 0.25 |
| 4 | t | dW | Pendulum | KM constant benevolent | KM random | KM constant evil | Reward (constant benevolent nature) | Reward (random nature) | Reward (constant malevolent nature) | | | |
| 5 | 0 | 0.0470604 | 1 | 5 | 5 | 5 | 0.9142136 | 0.9142136 | 0.9142136 | | | |
| 6 | 0.01 | 0.0870142 | 0 | 5.0035071 | 5.0416492 | 5.0485071 | -4.003507 | -4.041649 | -4.048507 | | | |
| 7 | 0.02 | 0.0942122 | 0 | 5.0105781 | 5.0816895 | 5.1005647 | -4.010578 | -4.081689 | -4.100565 | | | |
| 8 | 0.03 | -0.046811 | 0 | 4.947067 | 5.0444031 | 5.0820588 | -3.947067 | -4.044403 | -4.082059 | | | |

## 2.4. Standard Dynamic Programming tools

The infinitesimal generator of (2.1) for a function $v \in \mathcal{C}^2(\mathbb{R} \times \{\text{Good}, \text{Bad}\})$ is

$$\begin{aligned}
\mathbb{L}^{u,\alpha} v(x, \text{Bad}) &= (u - \alpha x) v'(x, \text{Bad}) + \frac{1}{2} k^2 v''(x, \text{Bad}) \\
&\quad -\lambda (v(x, \text{Bad}) - v(x, \text{Good})), \\
\mathbb{L}^{u,\alpha} v(x, \text{Good}) &= (u - \alpha x) v'(x, \text{Good}) + \frac{1}{2} k^2 v''(x, \text{Good}) \\
&\quad +\mu (v(x, \text{Bad}) - v(x, \text{Good})).
\end{aligned}$$

The HJB equations for maximizing (2.4) subject to (2.1) are:

$$J = \sup_{u \in [0,\eta]} \inf_{\alpha \in [0,a]} \left( \mathbb{L}^{u,\alpha} v(x, \text{Bad}) + r(x, \text{Bad}, u) \right), \tag{2.5}$$

$$J = \sup_{u \in [\eta,\gamma]} \inf_{\alpha \in [0,a]} \left( \mathbb{L}^{u,\alpha} v(x, \text{Good}) + r(x, \text{Good}, u) \right). \tag{2.6}$$

**Theorem 2.1.** *(See Theorem 1 in [4].) Suppose that*

(a) *the system (2.1) meets Itô's conditions and the uniform ellipticity condition,*

(b) *there exists a unique invariant probability measure $\mu_{u,\alpha}(\mathrm{d}x, i)$ for (2.1),*

(c) *the process (2.1) is exponentially ergodic with respect to a Lyapunov function $w$.*

*Then,*

**(i)** *There is a unique solution $(J, v)$ for (2.5)-(2.6), with $v \in \mathcal{C}^2(\mathbb{R} \times \{\mathrm{Bad}, \mathrm{Good}\})$.*

**(ii)** *The scalar $J$ in (2.5)-(2.6) coincides with the maximal worst payoff*

$$J^* := \sup_{u \in [0,\gamma]} \inf_{\alpha \in [0,a]} \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}_{x,i}^{u,\alpha} \left[ \int_0^T [F(u) - D(x,i)] \mathrm{d}t \right].$$

STONY BROOK CENTER FOR GAME THEORY

Home Page

Title Page

Contents



Go Back

Full Screen

Close

Quit

Anáhuac México

*Sketch of the proof.* **(i)** The existence of a constant $J$ and a function $v$ such that (2.5)-(2.6) hold, is based on the *vanishing discount technique*. Cf. [6, Theorem 5.5] for further reference.

**(ii)** An application of Dynkin's formula for controlled Markov-modulated diffusions to $v$ (see [7, p.48, Theorem 1.45]) yields:

$$\begin{aligned}\mathbb{E}_{x,i}^{u,\alpha_u} v(x(t),\theta(t)) &= v(x,i) + \mathbb{E}_{x,i}^{u,\alpha_u}\left(\int_0^t \mathbb{L}^{u,\alpha_u} v(x(s),\theta(s))\mathrm{d}s\right)\\ &\leq v(x,i) + Jt - \mathbb{E}_{x,i}^{u,\alpha_u}\left(\int_0^t r(x(s),\theta(s),u,\alpha_u)\mathrm{d}s\right).\end{aligned}$$

Thus, multiplying by $t^{-1}$, we have

$$t^{-1}J_t(x,i,f,u,\alpha_u) \leq J + t^{-1}h(x,i) - t^{-1}\mathbb{E}_{x,i}^{u,\alpha_u} h(x(t),\theta(t)). \tag{2.7}$$

Now, by **(c)**, we get

$$\left|\mathbb{E}_{x,i}^{u,\alpha_u}[v(x(t),\theta(t))]\right| \leq \|v\|_w \left[\mathrm{e}^{-c_1 t} w(x,i) + \frac{d_1}{c_1}\left(1-\mathrm{e}^{-c_1 t}\right)\right]. \tag{2.8}$$

Let $t \to \infty$ in (2.7) and use (2.8) to obtain

$$J \geq J(u, \alpha_u), \ \ \forall u \in [0, \gamma].$$

Hence, for each $u \in [0, \gamma]$:

$$J \geq \inf_{\alpha \in [0,\alpha]} J(u, \alpha) \quad \text{implies that } J \geq \sup_{u \in [0,\gamma]} \inf_{\alpha \in [0,a]} J(u, \alpha). \tag{2.9}$$

To obtain the inverse inequality, observe that by **(i)**, we can assert the existence of a strategy $u^* \in [0, \gamma]$ satisfying

$$\begin{aligned} J &= \inf_{\alpha \in [0,a]} \{r(x, i, u^*, \alpha) + \mathbb{L}^{u^*,\alpha} v(x, i)\}, \ \ \forall (x, i) \in \mathbb{R} \times \{\text{Good}, \text{Bad}\}, \\ &\leq r(x, i, u^*, \alpha) + \mathbb{L}^{u^*,\alpha} v(x, i), \ \ \forall\, \alpha \in [0, a], \ (x, i) \in \mathbb{R} \times \{\text{Good}, \text{Bad}\}. \end{aligned}$$

Take an arbitrary $\alpha \in [0, a]$, and apply Dynkin's formula to $v$, to get

$$\mathbb{E}^{u^*,\alpha}_{x,i} v(x(t), \theta(t)) \ \leq \ v(x, i) + Jt - \mathbb{E}^{u^*,\alpha}_{x,i} \left( \int_0^t r(x(s), \theta(s), u^*, \alpha) \mathrm{d}s \right).$$

Analogously, we can show that $J \leq J(u^*, \alpha)$, *i.e.*:

$$J \leq \inf_{\alpha \in [0,a]} J(u^*, \alpha),$$

and consequently,

$$J \leq \sup_{u \in [0,\gamma]} \inf_{\alpha \in [0,a]} J(u, \alpha).$$

This inequality, together with (2.9) yields that $J = J^*$.

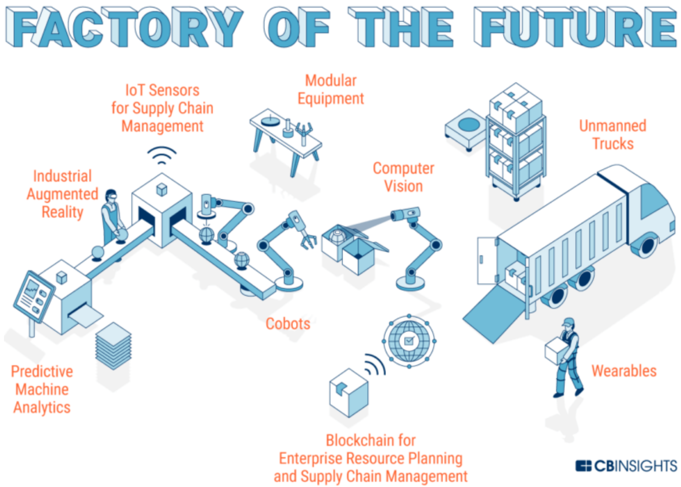
□

Work (2.5)-(2.6) by cases ($v'(x,\cdot) > 0$, $v'(x,\cdot) < 0$ and $v'(x^*,\cdot) = 0$) and apply Theorem 2.1 to see that

$$u^*(x, \text{Bad}) = \begin{cases} (F')^{-1}(-v'(x, \text{Bad})) & \text{if } F'(\eta) < -v'(x, \text{Bad}), \\ \eta & \text{if } F'(\eta) \geq -v'(x, \text{Bad}); \end{cases}$$

$$u^*(x, \text{Good}) = \begin{cases} (F')^{-1}(-v'(x, \text{Good})) & \text{if } F'(\gamma) < -v'(x, \text{Good}), \\ \gamma & \text{if } F'(\gamma) \geq -v'(x, \text{Good}). \end{cases}$$

## 3. Optimal control for a single/parallel machine

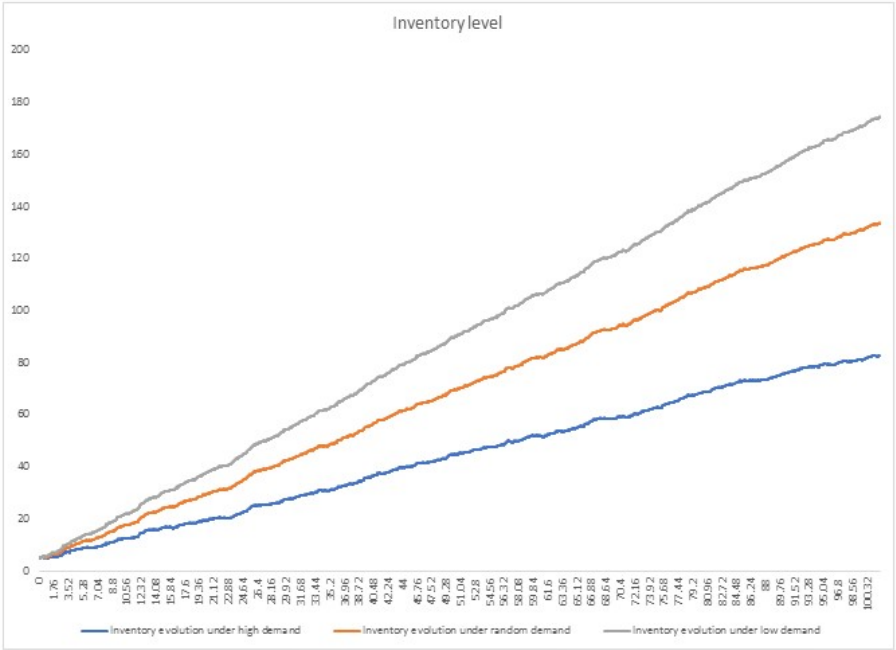


Consider the **single/parallel machine system process** studied in [8, Chapter 3], and defined by the controlled diffusion with Markovian switching

$$dx(t) = (u(t) - \alpha(t))\mathrm{d}t + k\mathrm{d}W(t), \tag{3.1}$$

where

- $x(t)$ is the stock level at time $t \geq 0$,
- $\alpha(t)$ is the demand rate at time $t$,
- $u(t)$ is the production rate at time $t$, which depends implicitly on $\alpha(x)$, and the machine capacity level $\theta(t)$.


Inventory level
200
180
160
140
120
100
80
60
40
20
0
Inventory evolution under high demand
Inventory evolution under random demand
Inventory evolution under low demand

**Assumption 3.1.** *Let* $h : \mathbb{R} \to \mathbb{R}$ *and* $c : \mathbb{R} \to \mathbb{R}$ *be the surplus cost and the production cost function, respectively. We suppose that* $h(x)$ *is a nonnegative, convex function and locally Lipschitz with* $h(0) = 0$*; whereas* $c(u)$ *is a nonnegative function,* $c(0) = 0$ *and* $c(u)$ *is twice differentiable. In addition,* $c(u)$ *is either strictly convex or linear.*

The objective is to find a production rate $\{u(t), t \geq 0\}$ that minimizes the long-run expected average cost

$$J(x, u, i, \alpha) := \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}^{u,\alpha}_{x,i} \left[ \int_0^T [h(x(t)) + c(u)] \mathrm{d}t \right].$$

**Assumption 3.2.** *The unknown demand rate* $\alpha(t)$ *is in* $[0; a]$ *with* $a < i$ *for all* $i \in E$.

# HOW TO CHOOSE THE STRATEGIES

An application of Theorem 2.1 yields the robust strategy for the controller:

$$u^*(x,i) = \begin{cases} (c')^{-1}(-v'(x,i)) & \text{if } c'(i) \le -v'(x,i), \\ i & \text{if } c'(i) > -v'(x,i). \end{cases}$$

```vba
Option Explicit
Const k = 0.5
Function ItoIntegral2(x0 As Double, intPendulum As Integer, intModNat As Integer, lambda As Double, mu As Double, dt As Double, T As Double) As Double
Dim dblProcess As Double, dblTime As Double, dblStrategy As Double, dblProd As Double, dblSurplus As Double, dblCost As Double, dW As Double, dblNature As Double
Randomize
dblProcess = x0: dblStrategy = intPendulum
Rem The following line accounts for our choice of dblProd(u)=sqrt(u) and dblSurplus(x)=x
dblProd = Sqr(dblStrategy): dblSurplus = dblProcess: dblCost = dblProd + dblSurplus: ItoIntegral2 = dblCost: dblTime = 0
While dblTime <= T
  intPendulum = telegraph2(intPendulum, dblTime, lambda, mu): dblStrategy = intPendulum
  dW = Application.Norm_Inv(Rnd(), 0, Sqr(dt))
  Select Case intModNat
    Case intModNat = 0 'low demand
      dblNature = 0.1
    Case intModNat = 1 'high demand
      dblNature = intModNat
    Case Else 'Random demand
      dblNature = Rnd()
  End Select
  dblProcess = dblProcess + (dblStrategy - dblNature * dblProcess) * dt + k * dW
  dblProd = Sqr(dblStrategy): dblSurplus = dblProcess: dblCost = dblProd + dblSurplus
  ItoIntegral2 = ItoIntegral2 + dblCost
  dblTime = dblTime + dt
Wend
ItoIntegral2 = ItoIntegral2 * dt
End Function
```

A variation of Algorithm 1 with $h(x) = x$ and $c(u) = \sqrt{u}$.

```vba
Function MonteCarlo2(x0 As Double, intPendulum As Integer, intModNat As Integer, lambda As Double, mu As Double, dt As Double, T As Double, iterations As Integer) As Double
Dim i As Integer
MonteCarlo2 = 0
For i = 1 To iterations
  MonteCarlo2 = ((i - 1) * MonteCarlo2 + ItoIntegral2(x0, intPendulum, intModNat, lambda, mu, dt, T)) / i
Next i
End Function
```

A variation of Algorithm 2.

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | k= | | empiric | theoretical | lambda = | 15 | Monte Carlo Expectations | | |
| 2 | 0.5 | IP(2) | 0.7477168 | 0.75 | mu = | 5 | 4.8067414 | 4.8075225 | 880.7919 |
| 3 | | IP(1) | 0.2522832 | 0.25 | | one Itô's integral | 45.814763 | 71.165586 | 91.638263 |
| 4 | t | dW | Pendulum | inventory under constant high demand | inventory under random demand | inventory under constant low demand | Cost (constant high demand) | Cost (random demand) | Cost (constant low demand) |
| 5 | 0 | 0.0435207 | 1 | 5 | 5 | 5 | 6 | 6 | 6 |
| 6 | 0.01 | -0.008002 | 1 | 4.9959988 | 4.9962002 | 5.0049988 | 5.9959988 | 5.9962002 | 6.0049988 |
| 7 | 0.02 | 0.0526218 | 1 | 5.0223097 | 5.0236778 | 5.0403097 | 6.0223097 | 6.0236778 | 6.0403097 |
| 8 | 0.03 | 0.0889028 | 1 | 5.0667611 | 5.0704047 | 5.0937611 | 6.0667611 | 6.0704047 | 6.0937611 |
| 9 | 0.04 | 0.1328352 | 1 | 5.1331787 | 5.1379944 | 5.1691787 | 6.1331787 | 6.1379944 | 6.1691787 |

Full implementation

# 4. Concluding remarks

- We work with a class of controlled diffusion processes with infinite horizon whose dynamics has the form (1.2) along with the transition rule (1.1); where the drift coefficient $b$ depends on an unknown and possibly non-observable parameter $\alpha$. Due to the lack of knowledge of such parameter, we reformulate the problem as an optimal control model under ambiguity, or *game against nature*, or *worst case optimal control*.

- These models can be applied in:

  – the problem of vaccine distribution,

  – the optimal allocation of renewable resources,

  – the problem of measuring effectiveness of molecular programs.

Anáhuac México

STONY BROOK CENTER FOR GAME THEORY

- Our work can be located within the field of optimal control models with Markovian switching under the average payoff criterion. We the impose general conditions on:
  - the sets where the actions and the unknown parameter take values,
  - the kind of continuity that the drift and diffusion coefficients satisfy,
  - the possibility of considering an unbounded reward rate function.

# References

[1] Ghosh, M. K., Arapostathis, A., and Marcus, S. I. Optimal control of switching diffusions with applications to flexible manufacturing systems. *SIAM J. Control Optim. 31*, 5 (1993), 1183–1204.

[2] Ghosh, M. K., Arapostathis, A., and Marcus, S. I. Ergodic control of switching diffusions. *SIAM J. Control Optim. 35* (1997), 1962–1988.

[3] Higham, D. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review 43(3)* (2001), 525–546.

[4] Higuera-Chan, C. G., Escobedo-Trujillo, B. A., and López-Barrientos, J. D. Robust optimal control for diffusion processes with markovian switch-

ing : the average criterion. *Submitted to Acta Applicandae Mathematicae* (2020). Available here.

[5] Kawaguchi, K., and Morimoto, H. Long-run average welfare in a pollution accumulation model. *J. Econom. Dyn. Control 31* (2007), 703–720.

[6] López-Barrientos, J. D., Jasso-Fuentes, H., and Escobedo-Trujillo, B. A. Discounted robust control for Markov diffusion processes. *TOP 23* (2015), 53–76. Available here.

[7] Mao, X., and Yuan, C. *Stochastic Differential Equations with Markovian Switching*. World Scientific Publishing Co., UK, 2006.

[8] Sethi, S., Zhang, H., and Zhang, Q. *Average-cost control of stochastic manufacturing systems*. Springer, New York, 2005.

# Thank you for your attention!

## Contact information

Facultad de Ciencias Actuariales, Universidad Anáhuac México.

Av. Universidad Anáhuac 46, Lomas Anáhuac, CP52786

Naucalpan de Juárez, México.

Tel. 01-52-55-56-27-0210 ext. 8506

daniel.lopez@anahuac.mx

https://www.researchgate.net/profile/Jose_Daniel_Lopez-Barrientos

STONY BROOK CENTER FOR GAME THEORY

Home Page

Title Page

Contents

Page 29 of 29

Go Back

Full Screen

Close

Quit

Anáhuac México