

Stochastic control for medical treatment optimization

Benoîte de Saporta

Institut Montpelliérain Alexander Grothendieck
Univ. Montpellier



Equipe EPS

Composition de l'équipe

- ▶ 8 PR, 17 MCF, 3 CR
- ▶ 2 ATER, 7 doctorant-es

Structuration

- ▶ **Responsable** : Jean-Noël Bacro
- ▶ thèmes **transversaux** : Apprentissage, Sélection/choix de modèles, Statistique en grande dimension + 3 axes
 - ▶ **Megadonc** statistiques modélisation en grande dimension, données fonctionnelles, approches machine learning et méthodes du noyau de convolution
 - ▶ **Memo** processus stochastiques, statistique bayésienne, statistique des processus
 - ▶ **Modest** modélisation de structures cachées, régularisation, fouille de données, extrêmes en dimension finie et infinie

Equipe EPS

Applications

- ▶ **Megadonc** médecine, agronomie, finance, génomique
- ▶ **Memo** écologie, environnement, épidémiologie, fiabilité
- ▶ **Modest** chimiométrie, risques environnementaux, sécurité, médico-social

Principales collaborations

- ▶ **A Montpellier** CIRAD, CHU, INRAE-SupAgro, Inria
- ▶ **Entreprises** Balea, Sanofi, Schlumberer, Thales

Collaborations pour l'optimisation de traitement médical

A Montpellier IMAG



Alice
Cleyen
(CR CNRS)



Orlane
Rossini
(Doctorante)



Amélie
Vernay
(Doctorante)

A Toulouse Inrae MIAT



Régis
Sabbadin
(DR)



Merixell Vinyals
Salgado (CR)

A Paris



Aymar Thierry d'Argenlieu
(Stagiaire M2, Institut Polytechnique de Paris)

Outline

Motivation

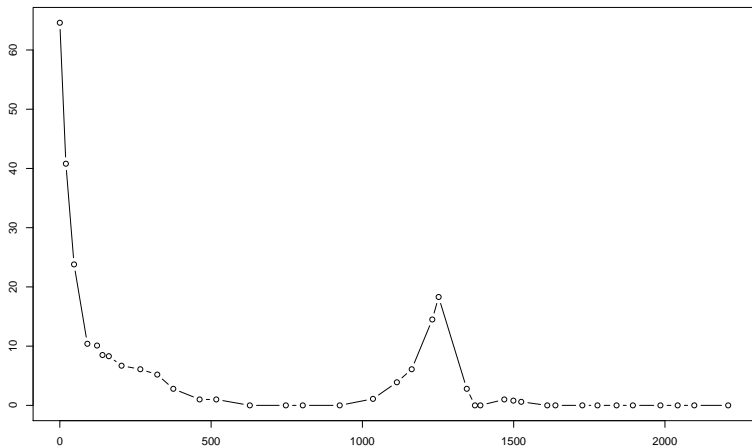
Modeling the control problem

Parameter calibration

Numerical experiments

Motivation : Multiple myeloma

Clinical trial : Intergroupe Francophone du Myélome 2009



Measurement of **monoclonal immunoglobulin** as a function of time
for one patient

Source : Centre de Recherche en Cancérologie de Toulouse

Medical treatment optimization

Use past and present monoclonal immunoglobulin **measurements** to sequentially **choose**

- ▶ the appropriate **treatment** to be applied until the next visit to the medical center
- ▶ the **date** of the next visit to the medical center

to **improve** the patient health status: maintain low levels of monoclonal immunoglobulin while minimizing **undesirable side effects** and treatment **constraints/costs**

Difficulties

- ▶ relapse date detection / relapse type detection: the overall health status is hidden and random
 - ↪ use monoclonal immunoglobulin as a marker of the disease

Difficulties

- ▶ relapse date detection / relapse type detection: the overall health status is hidden and random
~> use monoclonal immunoglobulin as a marker of the disease
- ▶ low frequency, noisy observations: continuous evolution of the marker, but discrete observations dates with low frequency and observations possibly corrupted by noise

Difficulties

- ▶ relapse date detection / relapse type detection: the overall health status is hidden and random
~> use monoclonal immunoglobulin as a marker of the disease
- ▶ low frequency, noisy observations: continuous evolution of the marker, but discrete observations dates with low frequency and observations possibly corrupted by noise
- ▶ treatment choice: several treatment types with strong influence on the patient health

Difficulties

- ▶ relapse date detection / relapse type detection: the overall health status is hidden and random
↪ use monoclonal immunoglobulin as a marker of the disease
- ▶ low frequency, noisy observations: continuous evolution of the marker, but discrete observations dates with low frequency and observations possibly corrupted by noise
- ▶ treatment choice: several treatment types with strong influence on the patient health
- ▶ choice of the next visit date / treatment duration: non trivial compromise between
 - ▶ too early heavy treatments with severe side effects
 - ▶ too late increased risk of death if the disease is not treated in time

Our approach

- ▶ propose a **continuous-time stochastic model** for the joint evolution of the health status and the marker
 - ↪ some parameters may be **unknown**
- ▶ **formulate** the treatment optimization problem as a **stochastic control** problem
- ▶ propose **numerical** methods to
 - ▶ construct an **explicit** a policy close to optimality
 - ▶ (close to) optimally control a **new patient** in real time given their medical history
- ▶ study the performance of these policies on **simulated** patients with parameters **calibrated** from the clinical trial data or **simultaneously learned** from the patient history

Outline

Motivation

Modeling the control problem

Continuous-time dynamics of health status and marker

Optimization problem

Parameter calibration

Numerical experiments

Variables of interest

- ▶ **mode m** : overall health status
 - ▶ $m = 0$: healthy / remission,
 - ▶ $m = 1$: disease / type 1 relapse
 - ▶ $m = 2$: disease / type 2 relapse
 - ▶ $m = 3$: death

- ▶ **marker ζ** : monoclonal immunoglobulin $\zeta \in [\zeta_0, D]$
 - ▶ $\zeta = \zeta_0$: nominal value in healthy mode
 - ▶ $\zeta = D$: death threshold

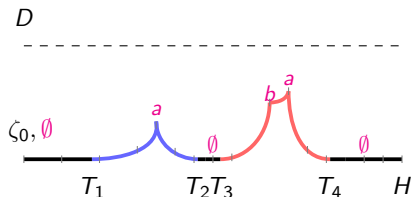
Treatments

The dynamics of the marker depend on the overall health status m and on the treatment chosen

$$\zeta(t+s) = \zeta(t) \exp(v_m^\ell s)$$

Possible treatments ℓ

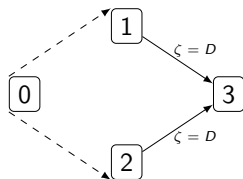
- ▶ $\ell = \emptyset$: no treatment
- ▶ $\ell = a$: treatment a
 - ▶ efficient on type 1 disease
 - ▶ slows the evolution of type 2 disease
- ▶ $\ell = b$: treatment b
 - ▶ efficient on type 2 disease
 - ▶ slows the evolution of type 1 disease



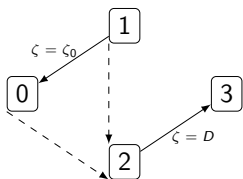
Changes of overall health status

The health status is **piecewise constant**, it changes at **deterministic** (solid lines) or **random** (dashes) dates with an **intensity** depending on

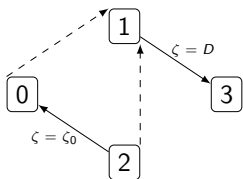
- ▶ the **marker** value ζ and / or the **time** spent in the current mode m (additional variable u required to keep a Markov process)
- ▶ and the **treatment** applied ℓ



Possibles transitions
without treatment



Possibles transitions
with treatment a



Possibles transitions
with treatment b

Observations dates = visits dates

δ minimum time between two observations

- ▶ δ is not small: typically 15 days for multiple myeloma (control horizon $H = 2400$ days)
- ▶ choice of the time r until the next visit restricted to some multiples of δ : $r \in \{15, 30, 60\}$ days

Observations dates = visits dates

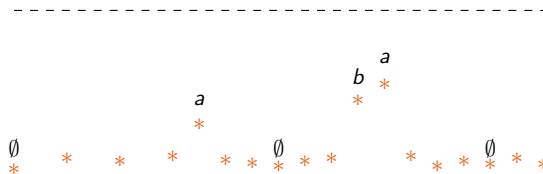
δ minimum time between two observations

- ▶ δ is not small: typically 15 days for multiple myeloma (control horizon $H = 2400$ days)
- ▶ choice of the time r until the next visit restricted to some multiples of δ : $r \in \{15, 30, 60\}$ days

At n -th visit / observation date

- ▶ hidden value of the process $X_n = (m_n, \zeta_n)$ health status, marker + additional variables
- ▶ observation $Y_n = \zeta_n + \epsilon_n$ with ϵ_n random noise + additional variables

D



Stochastic impulse control problem

Sequential decision-making

- ▶ which **treatment**? / for how long?

to keep the marker as close as possible to the nominal value ζ_0 , while

- ▶ the mode is **hidden**
- ▶ the jump dates are **hidden**
- ▶ the marker is observed through **noise** and with **low frequency**
- ▶ some parameters must be **learnt**

Value function

$$V(x, y) = \inf_{\pi \in \Pi} \mathbb{E}_{(x, y)}^{\pi} \left[\sum_{n=0}^{N-1} c(X_n, Y_n, A_n) + C(X_N, Y_N) \right]$$

over all sequences π of admissible decisions (A_n)

Outline

Motivation

Modeling the control problem

Parameter calibration

- Off-line calibration of the jump intensity

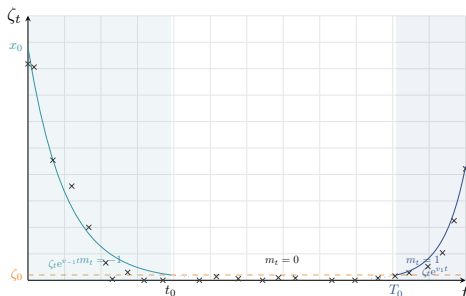
- On-line calibration of the slopes

Numerical experiments

Calibrating the jump intensity

Work in progress: PhD of Amélie Vernay

- ▶ simultaneously detect changes and fit the slopes
- ▶ estimate the intensity in mode m_0 with reconstructed start and stop dates (+ possible censoring)



Belief space

Work in progress: PhD of Orlane Rossini

- ▶ if the mode is **known**, on-line calibration of slopes is easy through **regression** as the motion is deterministic between jumps
- ▶ when the modes are **hidden**, one must estimate the **belief** ie the **probability** to be in each mode (or state) given the past and present observations to know which observations to use or not to update the estimation

Belief space

Computing beliefs is **challenging**

- ▶ **untractable** analytical recursive form
- ▶ possibility to obtain **numerical approximation** through state space discretizations \rightsquigarrow highly problem dependent

Outline

Motivation

Modeling the control problem

Parameter calibration

Numerical experiments

With a known model

With unknown slope parameters

Solutions tried

State space + belief space discretization + dynamic programming

[CdS 23]

- ▶ highly problem specific grids, numerically intensive
- ▶ cannot be extended to more complex problems

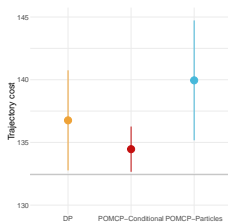
Solutions tried

State space + belief space discretization + dynamic programming
[CdS 23]

- ▶ highly problem specific grids, numerically intensive
- ▶ cannot be extended to more complex problems

Monte Carlo Planning [dSTdASC 24]

- ▶ adaptation of a simulation-based algorithm made for discrete space models
- ▶ can be extended to more complex problems
- ▶ combination with the discretized belief yields the best results



Unknown slope parameters : work program

Work in progress: PhD of Orlane Rossini

Use Deep Reinforcement Learning neural networks (DQN, R2D2, RNN)

- ▶ meant for discrete state-space problems \rightsquigarrow requires adaptation to continuous spaces
- ▶ observation (simulation) - based \rightsquigarrow try to enrich with some model knowledge in a Bayesian framework

Thank you !

References

[CdS 18] A. Cleynen, B. de Saporta *Change-point detection for PDMPs*

[CdS 23] A. Cleynen, B. de Saporta *Numerical method to solve impulse control problems for partially observed piecewise deterministic Markov processes*

<https://github.com/acleynen/PDMP-control>

[dSTdASC 24] B. de Saporta, A. Thierry d'Argenlieu, R. Sabbadin, A. Cleynen *A Monte-Carlo planning strategy for medical follow-up optimization: Illustration on multiple myeloma data*