Optimal epidemic suppression under an ICU constraint

Laurent Miclo

Toulouse School of Economics
Institut de Mathématiques de Toulouse

Joint work with Daniel Spiro (Uppsala University)
and Jörgen Weibull (Stockholm School of Economics)

Plan of the talk

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

Plan

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

The SIR model (1)

The Covid-19 epidemic has popularized the Susceptible-Infectious-Recovered model [Kermack and McKendrick, 1927]. Denote by x, y and z the proportions of susceptible, infectious and recovered (via immunization or death...) people in a population. Their dynamic follows a simple system of ordinary differential equations:

$$\begin{cases} \dot{x}(t) &= -\beta y(t)x(t) \\ \dot{y}(t) &= \beta y(t)x(t) - \alpha y(t) \\ \dot{z}(t) &= \alpha y(t) \end{cases} \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ are respectively the recovery and transmission rates.
The famous basic reproduction number is given by $R_0 = \beta/\alpha$.

The SIR model (2)

Since $x + y + z = 1$, it is sufficient to consider the couple (x, y), taking values in the simplex

$$\triangle \quad := \quad \{(u, v) \,:\, u \geqslant 0,\ \ v \geqslant 0 \text{ and } u + v \leqslant 1\}$$

We will assume an initial condition $(x(0), y(0)) \in \triangle$ is given, typically $(1 - \epsilon, \epsilon)$, with $\epsilon \ll 1$.

The ICU constraint

At any time $t \geqslant 0$, there is a proportion of $y(t)$ which requires an Intensive Care Unit, which is in limited supply. As a consequence, when $y(t)$ is too large the health system is overwhelmed, a situation we would like to avoid. It leads to a constraint

$$\forall\ t \geqslant 0, \qquad y(t) \leqslant \gamma \tag{2}$$

where $\gamma > 0$ is directly linked to the proportion of severe cases among infectious people and to the ICU capacity.

Policies

Some policies can be implemented so that the constraint (2) is satisfied. The government can try to modify α or β, here we will consider policies acting on the transmission rate β: washing hands, social distancing, wearing masks, lockdowns, etc.
It amounts to replacing, at any time $t \geqslant 0$, β by some $b(t) \geqslant 0$ and we are led to the time-inhomogeneous SIR equations:

$$\begin{cases} \dot{x}(t) & = & -b(t)y(t)x(t) \\ \dot{y}(t) & = & b(t)y(t)x(t) - \alpha y(t) \end{cases} \tag{3}$$

We are only interested in policies $b := (b(t))_{t \geqslant 0}$ such that (2) is satisfied. Denote by $\mathcal{B}_\gamma$ the set of such policies, assumed to be right-continuous with left-limits, with a finite number of jumps, and a finite number of connected components for $\{b = 0\}$.

Cost

Each measure reducing the sociability parameter β has an economic cost, assumed to be linear. The total cost of a policy $b \in \mathcal{B}_\gamma$ is

$$C(b) = \int_0^\infty [\beta - b(t)]_+ \, dt \tag{4}$$

(so that it is cost-free to increase β).
We are interested in

$$c^*(\gamma) = \inf_{b \in \mathcal{B}_\gamma} C(b) \tag{5}$$

and would like to find the optimal policies b where this minimal is attained.

Plan

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

Preliminaries on the laissez-faire policy (1)

The laissez-faire policy consists in letting $b(t) = \beta$ for all $t \geqslant 0$. On one hand, integrating the first equation of (1), we get

$$\forall\ t \geqslant 0, \qquad \ln(x(t)) \quad = \quad \ln(x(0)) - \beta \int_0^t y(s)ds$$

On the other hand, integrating the sum of the first and second equations of (1), it appears

$$\forall\ t \geqslant 0, \qquad x(t) + y(t) \quad = \quad x(0) + y(0) - \alpha \int_0^t y(s)ds$$

so we deduce the orbit in $\triangle$, namely $y(t)$ as a function of $x(t)$:

$$\forall\ t \geqslant 0, \qquad y(t) \quad = \quad y(0) + \frac{\alpha}{\beta} \ln\left[\frac{x(t)}{x(0)}\right] - x(t) + x(0) \quad (6)$$

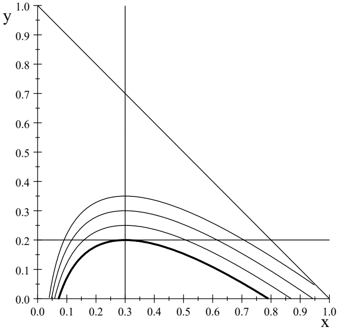
Preliminaries on the laissez-faire policy (2)

From the second equation of (1), when y attains its maximum, we have $x = \alpha/\beta$ (independently from the initial condition), and from (6), the maximum value of y is

$$\begin{aligned}\max_{t\geqslant 0} y(t) &= y(0) + x(0) + \frac{\alpha}{\beta}\ln\left[\frac{\alpha}{\beta x(0)}\right] - \frac{\alpha}{\beta} \\ &= 1 + \frac{\alpha}{\beta}\ln\left[\frac{\alpha}{\beta(1-\epsilon)}\right] - \frac{\alpha}{\beta}\end{aligned}$$

If this value is less than or equal to the ICU constraint γ, the laissez-faire is an optimal policy and $c^*(\gamma) = 0$.
From now on, assume we are in the "interesting" case is where this bound does not hold.

Preliminaries on the laissez-faire policy (3)



Several orbits in $\triangle$, with $\alpha/\beta = 0.3$, $\gamma = 0.2$.

Preliminaries on the laissez-faire policy (4)

Define two times: first

$$\begin{aligned}\tau_1 &:= \min\{t \geqslant 0 \,:\, y(t) = \gamma\}\\ &= \min\left\{t \geqslant 0 \,:\, x(t) = 1-\gamma+\frac{\alpha}{\beta}\ln\left(\frac{x(t)}{1-\varepsilon}\right)\right\}\end{aligned}$$

(it is possible to give a formula for τ_1 in terms of the data, via the dilogarithm function Li_2). The value $x(\tau_1)$ is the largest of the two solutions of the equation in x:

$$x = \frac{\alpha}{\beta}\ln\left(\frac{x}{1-\epsilon}\right)+1-\gamma \tag{7}$$

Consider next

$$\tau_2 := \tau_1+\frac{1}{\alpha\gamma}\left[x(\tau_1)-\frac{\alpha}{\beta}\right]$$

Optimal policies

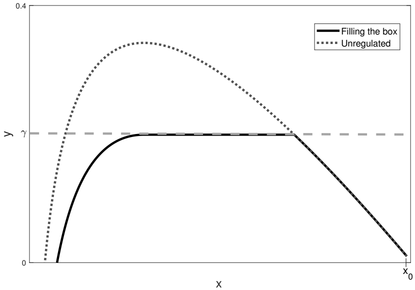
Theorem 1

There exist optimal solutions to the minimization of the functional C on $\mathcal{B}_\gamma$. One of them is the policy $b^* \in \mathcal{B}_\gamma$ defined by

$$b^*(t) = \begin{cases} \beta & \text{for } t \leqslant \tau_1 \\ \beta / \left[1 + \beta\gamma(\tau_2 - t)\right] & \text{for } \tau_1 < t \leqslant \tau_2 \\ \beta & \text{for } t > \tau_2 \end{cases} \tag{8}$$

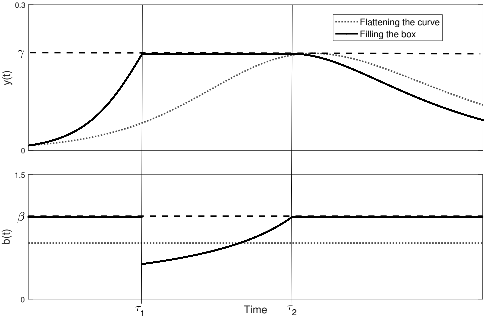
Every optimal policy $b \in \mathcal{B}_\gamma$ agrees with b^* on $[0, \tau_2]$ and satisfies $b(t) \geqslant \beta$ for all $t > \tau_2$.

b^* in $\triangle$



The orbit (solid) in $\triangle$ under the optimal policy b^* and the unregulated orbit (dotted) under the laissez-faire policy. Parameter values used: $\alpha = 0.3$, $\beta = 1$, $\gamma = 0.2$, and $\epsilon = 0.01$.

Comparison with the "flattening the curve" strategy



Upper panel: The share of infected over time under the optimal policy b^* (solid) and flattening the curve (dotted). The horizontal dashed line represents the ICU constraint γ. Lower panel: Optimal suppression (solid) and flattening-the-curve suppression (dotted). The horizontal dashed line represents the baseline spread β. Parameter values used: $\alpha = 0.3$, $\beta = 1$, $\gamma = 0.2$, and $\epsilon = 0.01$.

Minimal cost

From the minimal policy b^*, we can deduce the minimal cost, we get:

$$c^*(\gamma) = \frac{1}{\gamma}\left[\frac{\beta-\alpha}{\alpha} - \ln\frac{\beta(1-\epsilon)}{\alpha}\right] - \frac{\beta}{\alpha}$$

In the previous continuous population model, the limit $\epsilon \to 0_+$ is relevant and we get a cost only depending on the "structural data":

$$c^*(\gamma) = \frac{1}{\gamma}\left[\frac{\beta-\alpha}{\alpha} - \ln\frac{\beta}{\alpha}\right] - \frac{\beta}{\alpha}$$

Numerical illustration (1)

In the previous simple model, all the parameters are easily available, at least for a back-of-the-envelope calculation.
For Covid-19, we have, very approximatively, $R_0 = 2.5$ and with a week as time unit, $\alpha = 0.5$, corresponding to an infectious period of 2 weeks. So $\beta = \alpha R_0 = 1.25$.
To evaluate γ, assume 0.5% of those who are ill need intensive care. At least in 2013, the density of critical care beds in France was 11.6/100,000 inhabitants. Let N the population of France, the constraint on the ICU writes

$$\forall\ t \geqslant 0, \qquad y(t)\frac{0.5}{100}N \quad \leqslant \quad \frac{11.6}{100000}N$$

i.e. $\gamma = 11.6/(0.5 \times 1000) = 0.0232$.

Numerical illustration (2)

To get the minimal cost, note that $52 \times \beta$ is proportional to the French GDP. It follows that $c^*(\gamma)/(52 \times \beta)$ is the cost expressed in terms of the French GDP. We get

$$\left(\frac{1}{52 \times \gamma}\left[\frac{1}{\alpha} - \frac{1}{\beta} - \frac{1}{\beta}\ln\frac{\beta}{\alpha}\right] - \frac{1}{52 \times \alpha}\right)\mathrm{GDP}$$

$$\approx \quad 0.35\,\mathrm{GDP}$$

Strategy of the proof

Our strategy consists of the following steps:

- The optimization problem is written in the phase space $\triangle$.
- The new formulation admits a natural extension on a signed-measure space.
- Topological properties of this measure space and of the functional C imply the existence of a global minimizer.
- A priori a global minimizer is a general signed measure, but it turns out to be an absolutely continuous, bringing us back to a functional setting.
- Calculus of variation arguments show that the minimizer is uniquely determined until the time when x reaches the level α/β, and this leads to Theorem 1.

Plan

1 Introduction

2 The main result

3 The phase space $\triangle$

4 Extension to measures

5 Characterization of the minimizer

6 Future works?

Lagrangian formulation

In (3), x is always non-increasing and thus admits a limit for large time, say $x(\infty)$. First assume that $b > 0$. Then x is decreasing, so we can write

$$\forall\ t \geqslant 0, \qquad y(t) \quad = \quad \varphi(x(t))$$

where $\varphi : (x(\infty), x(0)] \to (0,1)$ is piecewise $\mathcal{C}^1$, its left and right derivatives exist everywhere and $\varphi'(r) > -1$, where φ' stands for the right derivative. Furthermore, the cost can be written under a Lagrangian form:

$$C(b) \quad = \quad \mathcal{J}(\varphi) \ := \int_{x(\infty)}^{x(0)} L(\xi, \varphi(\xi), \varphi'(\xi))\, d\xi$$

where for any $(\xi, \chi, \chi') \in (x(\infty), x(0)] \times (0,1) \times (-1, +\infty)$,

$$L(\xi, \chi, \chi') \quad := \quad \frac{\beta}{\alpha}\left(\frac{1+\chi'}{\chi} - \frac{\alpha}{\beta\xi\chi}\right)_+$$

When $b = 0$

When $b = 0$, say between the times $t_1 < t_2$, the lockdown is complete during this period, it is an attempt to suppress the disease. For $t \in (t_1, t_2)$,

$$\begin{cases} \dot{x}(t) &= 0 \\ \dot{y}(t) &= -\alpha y(t) \end{cases}$$

namely

$$\begin{cases} x(t) &= x(t_1) \\ y(t) &= y(t_1) \exp(-\alpha(t - t_1)) \end{cases}$$

To circumvent the difficulty that x remains constant while y is changing, we allow φ to jump at $x(t_1)$, taking

$$\begin{aligned} \varphi(x(t_1)) &:= y(t_1) \\ \varphi(x(t_1)-) &:= y(t_2) = y(t_1) \exp(-\alpha(t_2 - t_1)) \end{aligned}$$

The functional $\mathcal{J}$

The contribution of the period $(t_1, t_2]$ to the cost $C(b)$ is

$$\int_{t_1}^{t_2} (\beta - 0)_+ \, dt \quad = \quad \beta(t_2 - t_1) = \frac{\beta}{\alpha} \ln\left(\frac{y(t_1)}{y(t_2)}\right) = \frac{\beta}{\alpha} \ln\left(\frac{\varphi(x(t_1))}{\varphi(x(t_1)-)}\right)$$

It follows that in general,

$$\begin{aligned} C(b) \quad &= \quad \mathcal{J}(\varphi) \; := \int_{(x(\infty),x(0))} L(\xi, \varphi(\xi), \varphi'(\xi))\, d\xi \\ &\quad + \frac{\beta}{\alpha} \sum_{u \in (x(\infty),x(0)] \,:\, \varphi(u) \neq \varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right) \end{aligned}$$

and we are led to minimize $\mathcal{J}$ under appropriate conditions on φ, among which, $\varphi(u) > \varphi(u-)$ at any discontinuity point $u \in (x(\infty), x(0)]$ and $\varphi' > -1$ where this right derivative is defined. We can also restrict our attention to function φ such that $\varphi(\alpha/\beta) = \gamma$.

Plan

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

The measure μ

To any function φ as in the previous section, associate the (signed) measure μ on $I := [\alpha/\beta, x(0)]$ defined by

$$\mu(dx) \quad := \quad \frac{\varphi'(x)}{\varphi(x)}\, dx + \sum_{u\in(\alpha/\beta,x(0)]\,:\,\varphi(u)\neq\varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right)\delta_u(dx)$$

Conversely, we recover φ from μ via

$$\forall\ x \in \mathcal{I}, \qquad \varphi(x) \quad = \quad \gamma \exp(F_\mu(x))$$

where F_μ is the repartition function of μ. Another relation between μ and φ is:

$$\|\mu\|_{\mathrm{tv}} \quad \leqslant \quad \frac{2\exp(\beta\mathcal{J}(\varphi)/\alpha)}{y_0}(x_0 - \alpha/\beta) + \ln(y_0/\gamma) \qquad (9)$$

Two more measures

Define two non-negative measures on I via

$$\psi_\mu(dx) := \frac{\exp(-F_\mu(x))}{\gamma} dx$$

$$\nu_\mu(dx) := \left(1 - \frac{\alpha}{\beta x}\right) \frac{\exp(-F_\mu(x))}{\gamma} dx$$

In interest of the latter is

$$\mathcal{J}(\varphi) = \frac{\beta}{\alpha}(\mu + \nu_\mu)_+(I) =: \mathcal{K}(\mu)$$

Furthermore the conditions on φ (coming from $\mathcal{B}_\gamma$) can be written as

$$\mu(I) = \ln(y(0)/\gamma), \qquad F_\mu \leqslant 0, \qquad \mu + \psi_\mu \geqslant 0$$

Call $\mathcal{M}_\gamma$ the corresponding set of measures on I.

We are thus led to the minimization of $\mathcal{K}$ on $\mathcal{M}_\gamma$ and more precisely, due to (9), on $\mathcal{M}_\gamma \cap \{\mu : \|\mu\|_{\mathrm{tv}} \leqslant M\}$, for an appropriate $M \geqslant 0$ such that $\mathcal{M}_\gamma \cap \{\mu : \|\mu\|_{\mathrm{tv}} \leqslant M\} \neq \varnothing$. Standard arguments on the weak topology on measures defined on I enable to get:

- The set $\mathcal{M}_\gamma \cap \{\mu : \|\mu\|_{\mathrm{tv}} \leqslant M\}$ is compact.
- The mapping $\mathcal{K} : \mathcal{M}_\gamma \cap \{\mu : \|\mu\|_{\mathrm{tv}} \leqslant M\} \to \mathbb{R}$ is lower semi-continuous.

It follows that $\mathcal{K}$ admits a minimizer on $\mathcal{M}_\gamma$.

Plan

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

Reduction to absolutely continuous measures

Denote μ^* a minimizer of $\mathcal{K}$ on $\mathcal{M}_\gamma$.
Decompose any measure μ on I into a sum $\mu_{\mathrm{a}} + \mu_{\mathrm{s}} + \mu_{\mathrm{c}}$, where μ_{a} is atomic, μ_{s} is diffuse and singular with respect to λ and μ_{c} is absolutely continuous with respect to λ, i.e. admits a (signed) density $f : I \to \mathbb{R}$ with respect to λ, $\mu_{\mathrm{c}} = f \cdot \lambda$.
The cost functional $\mathcal{K}$ can be written

$$\mathcal{K}(\mu) \quad = \quad \frac{\beta}{\alpha}\left(\mu_{\mathrm{a}}(I) + \mu_{\mathrm{s}}(I) + \int_I (f + \nu_\mu)_+ \, d\lambda\right)$$

A first step consists in proving μ^* is absolutely continuous with respect to λ, the Lebesgue measure on I, i.e. $\mu^*_{\mathrm{a}} = \mu^*_{\mathrm{s}} = 0$.
This is done by contradiction: if it was not true, some mass of μ^*_{a} or μ^*_{s} could be pushed to the left under an absolutely continuous form to get a measure with smaller $\mathcal{K}$.

Minimal f^*

We are thus led to a minimization problem over measures μ of the form $f \cdot \lambda$, namely over a functional space of (signed) densities f. Define x^* as the unique solution belonging to $[\alpha/\beta, x(0)]$ of the equation

$$x^* - \frac{\alpha}{\beta}\ln(x^*) \quad = \quad 1 - \gamma - \frac{\alpha}{\beta}\ln(x(0)) \tag{10}$$

The value x^* coincides with $x(\tau_1)$ solution of (7). By separating the analysis on $[\alpha/\beta, x^*]$ and on $[x^*, x(0)]$, we end up by finding that a.e.

$$f^*(x) \quad := \quad \begin{cases} 0 & , \text{ if } x \leqslant x^* \\ -\left(\gamma - \left(x - x^* - \frac{\alpha}{\beta}\ln(x/x^*)\right)\right)^{-1}\left(1 - \frac{\alpha}{\beta x}\right) & , \text{ if } x > x^* \end{cases}$$

Proof of Theorem 1

From the explicit form of f^*, we compute

$$\min_{\mathcal{M}_\gamma} \mathcal{K} = \mathcal{K}(f^* \cdot \lambda) \quad = \quad \frac{1}{\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1 + \frac{\beta}{\alpha} - \ln(x(0))\right) - \frac{\beta}{\alpha}$$

This is also the value of $C(b^*)$ and even only of

$$\int_0^{\tau_2} [\beta - b^*(t)]_+ \, dt$$

Theorem 1 follows.

Plan

1. Introduction
2. The main result
3. The phase space $\triangle$
4. Extension to measures
5. Characterization of the minimizer
6. Future works?

On the cost function

Consider a cost of the form

$$\widetilde{C}(b) \quad := \quad \int_0^{+\infty} F(\beta - b(t))\, dt$$

for some mapping $F : \mathbb{R} \to \mathbb{R}_+$. Of course, if F coincides with $(\cdot)_+$ on $\mathbb{R}_+$, Theorem 1 still holds.
But it is more natural to assume that F is strictly convex on $\mathbb{R}_+$ (in the discutable assumption there is a continuum range of possible actions): the government should first apply the measures with the best ratio (reduction of R_0)/cost. In this case we think that Theorem 1 is no longer true: early measures should appears and the optimal orbit will be solution to some Euler-Lagrange equations. Nevertheless, once the level γ will be reached by y, the end of the orbit coincides with b^*, in particular the optimal policy will generically have a jump.

Cost of deaths

The cost C does not take into account the cost of deaths. It can be included as follows: among the total number of recovered people, which is $1 - x(\infty)$ times the total population, a certain proportion has died. So the cost of deaths can be modeled by an additional term to C of the form

$$\begin{aligned}\delta(1 - x(\infty)) &= \delta(1 - x(0)) + \delta(x(0) - x(\infty)) \\ &= \delta(1 - x(0)) - \delta \int_0^\infty \dot{x}(t)\, dt \\ &= \delta(1 - x(0)) + \delta \int_0^\infty b(t)x(t)y(t)\, dt\end{aligned}$$

where $\delta > 0$.
It leads to a Euler-Lagrange formulation. We believe that the optimal policy remains b^* if δ is small enough, but not for large δ.

Types (1)

The population can be partitioned in several types (e.g. young, worker and retired), indexed by a finite set $\mathcal{I}$. Each type $i \in \mathcal{I}$ has a proportion of susceptible (respectively infectious) people $x_i \in [0, 1]$ (resp. $y_i \in [0, 1 - x_i]$). The whole state is $(x, y) := (x_i, y_i)_{i \in \mathcal{I}} \in \triangle^{\mathcal{I}}$. The government can impose different policies to each type and the evolution of (x, y) is given by

$$\begin{cases} \dot{x}_i & = & -b_i(t)\left(\sum_{j \in \mathcal{I}} \mu_{i,j} y_j\right) x_i \\ \dot{y}_i & = & b_i(t)\left(\sum_{j \in \mathcal{I}} \mu_{i,j} y_j\right) x_i - \alpha_i y_i \end{cases}$$

where for all type $i \in \mathcal{I}$,

- $\alpha_i > 0$ is the "recovering" rate,
- $(b_i(t))_{t \geqslant 0}$ is the (non-negative) "sociability" policy,
- $(\mu_{i,j})_{i,j \in \mathcal{I}}$ is the Markov matrix of type meetings: for $i, j \in I$, $\mu_{i,j}$ is the probability that a type i meets a type j.

Types (2)

Under the constraint

$$\forall\ t \geqslant 0, \qquad \sum_{i\in\mathcal{I}} \chi_i y_i(t) \ \leqslant \ \gamma$$

where $\chi_i \geqslant 0$ is the impact of the infectious of type i on the health system and $\gamma > 0$ is the global limitation of the health system, we would like to find the optimal policy $b := (b_i)_{i\in\mathcal{I}}$ for costs $\mathcal{C}(b)$ of the form

$$\mathcal{C}(b) \ = \ \sum_{i\in\mathcal{I}} \int_0^{+\infty} \xi_i(b_i(t))\, dt + \delta_i(1 - x_i(\infty))$$

where for all $i \in \mathcal{I}$,

- $\xi_i : \mathbb{R}_+ \to \mathbb{R}_+$ gives the elementary cost of deviation of b_i from β_i, the natural sociability of i: typically ξ_i is convex and vanishes at β_i,
- $\delta_i \geqslant 0$ is the death cost.

In this framework, we don't expect to find a closed form for the optimal policies. Nevertheless, we would like to construct a stochastic algorithm finding them, by mixing Euler-Lagrange equations and simulated annealing.