Analysing necessary optimality conditions for time optimal control problems in Wasserstein spaces.

Fernanda Urrea[1] Hasnaa Zidani[1] Cristopher Hermosilla[2]

[1]INSA Rouen Normandie

[2]Universidad Técnica Federico Santa María

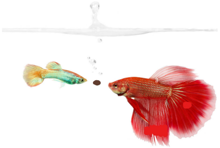
January 12, 2024

Example: Competitive Lotka-Volterra

Imagine we have a tank with 2 species of fishes that are fed with the same pellet forcing them to compete. The growth rate of the first specie can be modelled as

$$\dot{x}_1(t) = x_1(t)[\underbrace{r(t, u(t))}_{\text{inherit growth rate}} - \underbrace{(a(t, \boldsymbol{\mu}_t))}_{\text{effect of specie 2}} x_2(t)] \quad \text{s.t} \quad x(0) = \mu_1^0$$

where $u(t)$ is the amount of food at time t. We would like to to reach (in average) a certain amount of each specie in minimal time but also spending as less as possible in food.



Formally, the previous problem can be modelled as the following

$$\min_{u \in \mathcal{U}, T>0} T + \int_0^T u(t) \quad \text{s.t} \begin{cases} \partial_t \boldsymbol{\mu}_t + \text{div}_{x_1,x_2}(f(t, u_t, \boldsymbol{\mu}_t, \cdot)\boldsymbol{\mu}_t) = 0 \\ \boldsymbol{\mu}(0) = \mu_1^0 \otimes \mu_2^0, \end{cases}$$

$$\underbrace{\int_{\mathbb{R}} x_i \boldsymbol{\mu}_T(\cdot, x_i) \geq n_i}_{\text{Expected population i at final time}} \quad \forall i = 1, 2, \text{ and } \underbrace{\int_{\mathbb{R}} (x_1 + x_2)\boldsymbol{\mu}_T(x_1, x_2) \leq N}_{\text{capacity of the tank}}$$

But this problem enters in a more general setting...

Bolza free time problem

$$
\begin{cases}
\displaystyle\min_{u(\cdot)\in\mathcal{U},T>0} \quad T+\Lambda(\mu(T))+\int_0^T L(t,u(t))\mathrm{d}t \\
\text{s.t.} \begin{cases} \partial_t\mu(t)+\operatorname{div}(f(t,\mu(t),u(t),\cdot)\mu(t))=0, \\ \mu(0)=\mu^0, \\ \mu(T)\in\mathcal{Q}_T, \end{cases}
\end{cases}
\qquad (\mathcal{P}_{\mathrm{BP}})
$$

where $\Lambda : \mathscr{P}_c(\mathbb{R}^d)\to\mathbb{R}$, $\mu^0\in\mathscr{P}_c(\mathbb{R}^d)$ and the minimisation is taken over

$$
\mathcal{U} := \left\{ u : [0,T]\to U \text{ s.t. } u(\cdot) \text{ is } \mathscr{L}^1\text{-measurable } \right\},
$$

where (U, d_U) is a compact metric space, and the set of final-point constraints $\mathcal{Q}_T$ is defined by functional inequalities of the form

$$
\mathcal{Q}_T := \left\{ \mu\in\mathscr{P}_2\left(\mathbb{R}^d\right) \text{ s.t. } \Psi_i(\mu)\leq 0 \text{ for all } i\in\{1,\ldots,n\} \right\}
$$

where every $\Psi_i : \mathscr{P}_c\left(\mathbb{R}^d\right)\to\mathbb{R}$.

A bit about existence for the continuity equation in $[0, T]$

Hypotheses (CE): For every compact set $K \subset \mathbb{R}^d$:

(i) f is $\mathscr{L}^1$-measurable in time and for $\mathscr{L}^1$-almost every t and any $(\mu, x) \in \mathscr{P}_c\left(\mathbb{R}^d\right) \times \mathbb{R}^d$,

$$|f(t, \mu, x)| \leq m(t)\left(1 + |x| + \mathcal{M}_1(\mu)\right) \quad \text{for some } m(\cdot) \in L^1\left([0, T], \mathbb{R}_+\right)$$

(ii) f is Lipschitz in x and also in μ with Lipschitz constants in $L^1\left([0, T], \mathbb{R}_+\right)$:

$$|f(t, \mu, x) - f(t, \mu, y)| \leq l_K(t)|x - y| \text{ and}$$
$$|f(t, \mu, x) - f(t, \nu, x)| \leq L_K(t) W_p(\mu, \nu).$$

With this hypotheses on the dynamic and when $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ we have existence and uniqueness of the dynamical system (For instance in [BonnetFrankowska2021] but many others).

Strategy for obtaining optimality conditions

The idea (as in more classical settings) is to work in a fixed time setting by introducing a new state variable (in our case φ) and a new control (for us v) by doing the following:

Change of variables

Let $v \in \mathcal{V} := \mathscr{L}^\infty([0,1], V)$ such that $\int_0^1 v(t)dt = T$ and we introduce $t = \varphi(s) = \int_0^s v(\tau)d\tau$.

So our steps will be:

1. We check that both formulations are equivalent
2. We write optimality conditions for the new fixed time problem
3. We go back to our original problem.

Step 1: Fixed time problem

After the change of variables we get:

$$\left\{\begin{array}{l} \inf\limits_{\substack{u\in\mathcal{U}\\ v\in\mathcal{V}}} \displaystyle\int_0^1 v(s)(1+L(\varphi(s),u(s)))ds + \Lambda(\gamma_1) \\ \text{s.t.} \left\{\begin{array}{l} \partial_s\gamma_s + \mathrm{div}_x\left(v(s)f(\varphi(s),\cdot,\gamma_s,u_s)\gamma_s\right) = 0 \quad s\in[0,1], \\ \gamma(0)=\mu^0 \\ \dot{\varphi}(s) = v(s) \quad s\in[0,1] \ \text{ and } \varphi(0)=0. \\ \gamma(1)\in\mathcal{Q}, \end{array}\right. \end{array}\right. \qquad \textbf{(BT)}$$

Equivalence

Admissible solutions for both problems are equivalent.

Step 1: Fixed time problem

After the change of variables we get:

$$\begin{cases} \displaystyle\inf_{\substack{u\in\mathcal{U}\\ v\in\mathcal{V}}} \int_0^1 v(s)(1+L(\varphi(s),u(s)))ds + \Lambda(\gamma_1) \\ \text{s.t.} \begin{cases} \partial_s\gamma_s + \mathrm{div}_x\left(v(s)f(\varphi(s),\cdot,\gamma_s,u_s)\gamma_s\right) = 0 \quad s\in[0,1], \\ \gamma(0) = \mu^0 \\ \dot{\varphi}(s) = v(s) \quad s\in[0,1] \text{ and } \varphi(0) = \delta_0. \\ \gamma(1) \in \mathcal{Q}, \end{cases} \end{cases} \tag{\textbf{BT}}$$

Step 1: Fixed time problem

After the change of variables we get:

$$\begin{cases} \inf\limits_{\substack{u\in\mathcal{U}\\ v\in\mathcal{V}}} \displaystyle\int_0^1 v(s)(1+L(\varphi(s),u(s)))ds + \Lambda(\gamma_1) \\ \text{s.t.} \begin{cases} \partial_s\gamma_s + \operatorname{div}_x\left(v(s)f(\varphi(s),\cdot,\gamma_s,u_s)\gamma_s\right) = 0 \quad s\in[0,1], \\ \gamma(0)=\mu^0 \\ \dot\varphi(s) = v(s) \quad s\in[0,1] \ \text{ and } \ \varphi(0)=\delta_0. \\ \gamma(1)\in\mathcal{Q}, \end{cases} \end{cases} \qquad \textbf{(BT)}$$

An admissible $(\gamma^*(\cdot), u^*(\cdot), v^*(\cdot))$ is a **strong local minimiser** for **(BT)** if exists $\varepsilon > 0$ that

$$\int_0^T v^*(t)(1+L(\varphi^*(t),u^*(t)))dt + \varphi\left(\gamma^*(1)\right) \leq \int_0^T v(t)(1+L(\varphi(t),u(t)))dt + \varphi(\gamma(1))$$

for every other admissible tuple $(\gamma(\cdot), u(\cdot), v(\cdot))$ satisfying $\sup_{t\in[0,1]} W_1\left(\gamma^*(t),\, \gamma(t)\right) \leq \varepsilon$.

Step 2: Optimality conditions for the fixed time problem

Now that we have transform our problem we can apply the optimality conditions in the form of a Pontryagin maximum principle for Bolza problems in [BonnetFrankowska2021-PMP].

But before moving on to step 2 we need to say something about the **differentiability with respect to measures...**

Subdifferential calculus in Wasserstein spaces

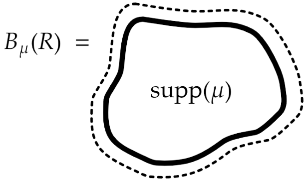
Let $\phi : \mathscr{P}_c\left(\mathbb{R}^d\right) \to \mathbb{R}$ and $\mu \in \mathscr{P}_c\left(\mathbb{R}^d\right)$, we define the localised sudifferential $\partial^-_{\mathrm{loc}}\phi(\mu)$ as the set of all $\xi \in L^2\left(\mathbb{R}^d, \mathbb{R}^d; \mu\right)$

$$\phi(\nu) - \phi(\mu) \geq \inf_{\gamma \in \Gamma_o(\mu,\nu)} \int_{\mathbb{R}^{2d}} \langle \xi(x), y - x \rangle \mathrm{d}\gamma(x,y) + o_R\left(W_2(\mu,\nu)\right),$$

for every $R > 0$ and any $\nu \in \mathscr{P}\left(B_\mu(R)\right)$

Analogously, $\xi \in L^2\left(\mathbb{R}^d, \mathbb{R}^d; \mu\right)$ belongs to the localised superdifferential $\partial^+_{\mathrm{loc}}\phi(\mu)$ if

$$(-\xi) \in \partial^-_{\mathrm{loc}}(-\phi)(\mu)$$



The space $\left(\mathscr{P}_2\left(\mathbb{R}^d\right), W_2\right)$ can be endowed with a pseudo-Riemannian structure. Given an element $\mu \in \mathscr{P}_2\left(\mathbb{R}^d\right)$, the analytical tangent space to $\mathscr{P}_2\left(\mathbb{R}^d\right)$ at μ is defined in this context as

$$\mathrm{Tan}_\mu \, \mathscr{P}_2\left(\mathbb{R}^d\right) := \overline{\left\{\nabla\xi(\cdot) \text{ s.t. } \xi \in C_c^\infty\left(\mathbb{R}^d\right)\right\}}^{L^2(\mu)}$$

It can be shown that $\partial^-_{\mathrm{loc}}\,\phi(\mu) \cap \partial^+_{\mathrm{loc}}\,\phi(\mu)$ contains at most one element, which also belongs to $\mathrm{Tan}_\mu \, \mathscr{P}_2\left(\mathbb{R}^d\right)$.

Definition: Locally differentiable functional

A functional $\phi : \mathscr{P}_c\left(\mathbb{R}^d\right) \to \mathbb{R}$ is locally differentiable at $\mu \in \mathscr{P}_c\left(\mathbb{R}^d\right)$ if there exists a map $\nabla\phi(\mu) \in \mathrm{Tan}_\mu \, \mathscr{P}_2\left(\mathbb{R}^d\right)$ - called the Wasserstein gradient of $\phi(\cdot)$ at μ-, such that

$$\partial^-_{\mathrm{loc}}\phi(\mu) \cap \partial^+_{\mathrm{loc}}\phi(\mu) = \{\nabla\phi(\mu)\}.$$

For example:

$$\mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} V(x)\mathrm{d}\mu(x) \quad \text{with } V \text{ continuously differentiable}$$

(back to) Step 2: some assumptions to be considered..

Hypotheses for the control u

- Hypotheses (**CE**) with constants independent of $u \in U$.
- The map $u \in U \mapsto f(t, \mu, u)(x) \in \mathbb{R}^d$ is continuous for $\mathscr{L}^1$-a.e t and any $(\mu, x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$ and also $u \in U \mapsto L(t, u)$ is continuous for $\mathscr{L}^1$-a.e t

Regularity assumptions for the cost and constraints

For every $R > 0$, assume that the following holds with $K := B(0, R)$.

(ii) The final cost $\Lambda(\cdot)$ and the constraint functionals $\{\Psi_i(\cdot)\}_{1 \leq i \leq n}$ are Lipschitz continuous in the W_1-metric over $\mathscr{P}(K)$ and locally differentiable over $\mathscr{P}_c(\mathbb{R}^d)$. Moreover, the maps

$$x \in \mathbb{R}^d \mapsto \nabla\Lambda(\mu)(x) \in \mathbb{R}^d \quad \text{and} \quad x \in \mathbb{R}^d \mapsto \nabla\Psi_i(\mu)(x) \in \mathbb{R}^d,$$

are continuous for every $i \in \{1, \dots, n\}$.

Conditions for the minimal time problem

There exists non-trivial multipliers $(\lambda_0, \lambda_1, \ldots, \lambda_n) \in \{0, 1\} \times \mathbb{R}^n_+$ and a curve of measures $\bar{\nu}^* \in \mathrm{AC}([0, T^*], \mathscr{P}_c(\mathbb{R}^{2d+1}))$ that solves

$$\partial_s \bar{\nu}^*_s + \mathrm{div}_{x,r,q}(\mathcal{V}(s, \bar{\nu}^*(s))\bar{\nu}^*(s)) = 0$$

Conditions for the minimal time problem

There exists non-trivial multipliers $(\lambda_0, \lambda_1, \ldots, \lambda_n) \in \{0,1\} \times \mathbb{R}_+^n$ and a curve of measures $\bar{\nu}^* \in \mathrm{AC}([0, T^*], \mathscr{P}_c(\mathbb{R}^{2d+1}))$ that solves

$$\partial_s \bar{\nu}_s^* + \mathrm{div}_{x,r,q}(\mathcal{V}(s, \bar{\nu}^*(s))\bar{\nu}^*(s)) = 0$$

with $\mathcal{V}(s, \bar{\nu}^*(s), u^*(s)) =$

$$\begin{pmatrix} f\left(s, u^*(s), \mu^*(s), x\right) \\ -\mathrm{D}_x f\left(s, u^*(s), \mu^*(s), x\right)^\top r - \int_{\mathbb{R}^{2d}} \mathrm{D}_\mu f\left(s, u^*(s), \mu^*(s), y\right)(x)^\top p \, \mathrm{d}\bar{\nu}^*(s)(y, p, \cdot) \\ -\frac{\mathrm{d}}{\mathrm{d}s} f\left(s, u^*(s), \mu^*(s), x\right)^\top r \end{pmatrix}$$

which implies that the $(\pi^1, \ldots, \pi^d)_\# \bar{\nu}^*(s) = \mu^*(s)$

Conditions for the minimal time problem

There exists non-trivial multipliers $(\lambda_0, \lambda_1, \dots, \lambda_n) \in \{0,1\} \times \mathbb{R}^n_+$ and a curve of measures $\bar{\nu}^* \in \mathrm{AC}([0, T^*], \mathscr{P}_c(\mathbb{R}^{2d+1}))$ that solves

$$\partial_s \bar{\nu}^*_s + \mathrm{div}_{x,r,q}(\mathcal{V}(s, \bar{\nu}^*(s))\bar{\nu}^*(s)) = 0$$

with $\mathcal{V}(s, \bar{\nu}^*(s), u^*(s)) =$

$$\begin{pmatrix} f\left(s, u^*(s), \mu^*(s), x\right) \\ -\mathrm{D}_x f\left(s, u^*(s), \mu^*(s), x\right)^\top r - \int_{\mathbb{R}^{2d}} \mathrm{D}_\mu f\left(s, u^*(s), \mu^*(s), y\right)(x)^\top p \, \mathrm{d}\bar{\nu}^*(s)(y, p, \cdot) \\ -\frac{\mathrm{d}}{\mathrm{d}s} f\left(s, u^*(s), \mu^*(s), x\right)^\top r \end{pmatrix}$$

which implies that the $(\pi^1, \dots, \pi^d)_{\#}\bar{\nu}^*(s) = \mu^*(s)$ and also we will have by the construction of $\bar{\nu}^*$ that

$$(\pi^{d+1}, \dots, \pi^{2d})_{\#}\bar{\nu}^*(T^*) = \left(-\lambda_0 \nabla \Lambda\left(\mu^*(T^*)\right) - \sum_{i=1}^{n} \lambda_i \nabla \Psi_i\left(\mu^*(T^*)\right)\right)_{\#} \mu^*(T^*) \qquad (1)$$

(ii) for every $i = 1, \dots, n$,

$$\lambda_i \int_{\mathbb{R}^d} \Psi_i(x) \mathrm{d}\mu^*(T^*)(x) = 0,$$

(iii) There exists a continuous function $\Xi : s \mapsto \int_{\mathbb{R}} Q(s,q) \mathrm{d}\delta_{T^*}(q)$ such that,

$$\overline{\mathbb{H}}(s, \bar{\nu}^*(s), u^*(s)) = \Xi(s) \quad \text{a.e } s \in [0, T^*],$$

where

$$\overline{\mathbb{H}}(s, \bar{\nu}^*(s), u^*(s)) := \int_{\mathbb{R}^{2d+1}} \langle r, f(s, u^*(s), \mu^*(s), x) \rangle, \ \mathrm{d}\bar{\nu}^*(s)(x, r, q) - \lambda_0 L(s, u^*(s))$$

and $Q(s,q)$ is the backward flow of

$$\left\{ \ \dot{q}(s) = -\partial_s f \left(s, \mu^*(s), u^*(s), \Phi^*_{(T^*,s)}(x) \right) q(s), \right. \tag{2}$$

and finally

$$\overline{\mathbb{H}}(s, \bar{\nu}^*(s), u^*(s)) = \sup_{u \in U} \overline{\mathbb{H}}(s, \bar{\nu}^*(s), u)$$

References I

[BonnetFrankowska2021] Benoît Bonnet and Hélène Frankowska. "Differential inclusions in Wasserstein spaces: The Cauchy-Lipschitz framework". In: *Journal of Differential Equations* 271 (2021), pp. 594–637.

[BonnetFrankowska2021-PMP] Benoît Bonnet and Hélène Frankowska. "Necessary optimality conditions for optimal control problems in Wasserstein spaces". In: *Applied Mathematics & Optimization* 84 (2021), pp. 1281–1330.