Relaxation Techniques in Optimization and Control: an Overview of the Recently Published Elsevier Book

Vadim Azhmyakov

∗ Department of Basic Sciences, Universidad de Medelln, Republic of Colombia,
∗∗ Division for Automation and Robotics, Tomsk Polytechnic University, Tomsk, Russian Federation

Medellin, Antioquia, March 2018

Outline

Publishing with Elsevier

Optimization technique is nowadays not only a mathematical technique but also a ”technology”.

https://www.elsevier.com/books/a-relaxation-based-approach-to-optimal-control-of-hybrid-and-switched-systems/azhmyakov/978-0-12-814788-7

Elsevier publishing process

Step 1. a book proposal ⇒ Step 2. some 5 internationally recognized Reviewers ⇒ Step 3. decision of the general Elsevier Committee ⇒ Step 4. if positive ⇒ Step 5. definition of the time scheduling for the Chapters-by-Chapters delivery ⇒ Step 6. book design definition / production

Main Concepts

A HS is a 7-tuple $\{\mathscr{Q}, M, U, F, \mathscr{U}, I, \mathscr{S}\}$, where

- $\mathscr{Q}$ is a finite set of discrete states (called *locations*);
- $M = \{M_q\}_{q\in\mathscr{Q}} \subset \mathbb{R}^n$ is a family of smooth manifolds;
- $U \subseteq \mathbb{R}^m$ is a set of admissible control input values (called *control set*);
- $F = \{f_q\}$, $q \in \mathscr{Q}$ is a family of maps $f_q : [0,t_f] \times M_q \times U \to TM_q$, where TM_q is the tangent bundle of M_q (see e.g., [19,27]);
- $\mathscr{U}$ is the set of all admissible control functions;
- $I = \{I_q\}$ is a family of adjoint subintervals of $[0,t_f]$ such that $\sum_{q\in Q} |I_q| = t_f$;
- $\mathscr{S}$ is a subset of Ξ, where $\Xi := \{(q,x,q',x') \ : \ q,q' \in \mathscr{Q}, x \in M_q, x' \in M_{q'}\}$

Main Concepts

Let $u(\cdot) \in \mathscr{U}$ be an admissible control for a HS. Then a "continuous" trajectory of HS is an absolutely continuous function $x : [0, t_f] \to \bigcup_{q \in \mathscr{Q}} M_q$ such that $x(0) = x_0 \in M_{q_1}$ and

- $\dot{x}(t) = f_{q_i}(t, x(t), u(t))$ for almost all $t \in [t_{i-1}, t_i]$ and all $i = 1, ..., r+1$;
- the switching condition $(x(t_i), x(t_{i+1})) \in S_{q_i, q_{i+1}}$ holds if $i = 1, ..., r$.

The vector $\mathscr{R}_{r+1} := (q_1, ... q_{r+1})$ is called a "discrete trajectory" of the hybrid control system. Let HS be defined above. For an admissible control $u(\cdot) \in \mathscr{U}$, the triplet $\mathscr{X}^u := (\tau, x(\cdot), \mathscr{R})$, where τ is the set of the corresponding switching times $\{t_i\}$, $x(\cdot)$ and $\mathscr{R}$ are the corresponding continuous and discrete trajectories, is called a hybrid trajectory of HS.

Main Problem

$$\text{minimize } \phi(x(t_f))$$
$$\text{subject to } \dot{x}(t) = f_{q_i}(t, x(t), u(t)) \text{ a.e. on}[t_{i-1}, t_i]$$
$$q_i \in \mathcal{Q}\ i = 1, \ldots, r+1,\ x(0) = x_0 \in M_{q_1},\ u(\cdot) \in \mathcal{U}.$$

Problem Formulation

smooth system:

$$\dot{x}(t) = f(t, x(t), u(t)),\ t \in [t_0, t_f],\ x(t_0) = x_0, \tag{1}$$

for each component $u_k(\cdot)$ of $u(\cdot) = [u_1(\cdot), \dots, u_m(\cdot)]^T \Rightarrow$
$\mathcal{Q}^{(k)} := \left\{ q_j^{(k)} \in \mathbb{R}, j = 1, \dots, M_k \right\}$, $M_k \in \mathbb{N}$, $k = 1, \dots, m$,
$q_1^{(k)} < q_2^{(k)} < \dots < q_{M_k}^{(k)}$.

control switching times: $\mathcal{T}^{(k)} := \left\{ t_i^{(k)} \in \mathbb{R}_+, i = 0, \dots, N_k \right\}$,
where $N_k \in \mathbb{N}$, $\quad k = 1, \dots, m$.. Moreover, $t_0^{(k)} < t_1^{(k)} < \dots < t_{N_k}^{(k)}$
and for each $\mathcal{T}^{(k)}$, $t_{N_1}^{(1)} = \dots = t_{N_m}^{(m)} = t_f$

Problem Formulation

space of admissible (piecewise constant) controls

$$\mathscr{S} = \mathscr{S}_1 \times \ldots \times \mathscr{S}_m,$$

where $\mathscr{S}_k := \left\{ v : [t_0, t_f] \to \mathbb{R} \mid v(t) = \sum_{i=1}^{N_k} I_{[t_{i-1}^{(k)}, t_i^{(k)})}(t) q_{j_i}^{(k)} \right\}$,
$q_{j_i}^{(k)} \in \mathscr{Q}^{(k)}$, $j_i \in \mathbb{Z}[1, M_k]$, $t_i^{(k)} \in \mathscr{T}^{(k)}$.

main OCP

$$\text{minimize } J(u(\cdot)) = \frac{1}{2} \int_{t_0}^{t_f} (\langle Q(t)x(t), x(t)\rangle + \langle R(t)u(t), u(t)\rangle)\, dt + \frac{1}{2}\langle Gx(t_f), x(t_f)\rangle, \qquad (2)$$

$$\text{subject to (1)}, \; u(\cdot) \in \mathscr{S}$$

Classic (Excessive) Relaxations and PGs

admissible control space convexification

$$\text{conv}(\mathscr{S}) := \left\{ v(\cdot) \mid v(t) = \sum_{s=1}^{|\mathscr{S}|} \lambda_s u_s(t), \ \sum_{s=1}^{|\mathscr{S}|} \lambda_s = 1 \right\},$$

$$\lambda_s \geq 0, u_s(\cdot) \in \mathscr{S}, s = 1, \ldots, |\mathscr{S}|.$$

fully relaxed OCP

$$\text{minimize } \overline{co}\{J(u(\cdot))\}, \text{ subject to } (1),\ u(\cdot) \in \text{conv}(\mathscr{S}). \quad (3)$$

! a convex minimization problem in $\mathbb{L}^2\{[t_0, t_f]; \mathbb{R}^m\}$!

Classic (Excessive) Relaxations and PGs

the gradient method (GM)

$$u_{(l+1)}(\cdot) = \gamma_l \mathscr{P}_{\mathrm{conv}(\mathscr{S})}\left[u_{(l)}(\cdot) - \alpha_l \nabla \overline{co}\{J(u_{(l)}(\cdot))\}\right] + (1-\gamma_l) u_{(l)}(\cdot),$$

where $l \in \mathbb{N}$, $\nabla \overline{co}\{J(u(\cdot))\}(t) = -\partial H(t, x(t), u(t), p(t), p_{n+1})/\partial u$,

$$\frac{d\tilde{p}(t)}{dt} = -\frac{\partial H(t, x(t), u(t), p(t), p_{n+1})}{\partial \tilde{x}},$$

$$\tilde{p}(t_f) = -\frac{\partial(\overline{co}\{\phi(\tilde{x}(t_f))\})}{\partial \tilde{x}}, \; \tilde{x}(t_0) = (x_0^T, 0)^T, \; x := (x, x_{n+1})^T,$$

$$\frac{d\tilde{x}(t)}{dt} = \frac{\partial H(t, x(t), u(t), p(t), p_{n+1})}{\partial \tilde{p}}, \; \tilde{p}(t) := (p(t), p_{n+1})^T,$$

$$H(t, x, u, p, p_{n+1}) = \langle p, f(t, x, u)\rangle + \frac{1}{2} p_{n+1}\left(\langle Q(t)x, x\rangle + \langle R(t)u, u\rangle\right).$$

Classic (Excessive) Relaxations and PGs

Theorem

Let $p_{n+1} \neq 0$. Consider $\{u_{(l)}(\cdot)\}$ generated by GM with a constant step size α. Then for $u_{(0)}(\cdot) \in conv(\mathscr{S})$ the resulting sequence $\{u_{(l)}(\cdot)\}$ is a minimizing sequence for (3), i.e., $\lim_{l\to\infty} \overline{co}\{J(u_{(l)}(\cdot))\} = \overline{co}\{J(u^*(\cdot))\}$. Additionally assume that that $\partial f(t,x,u)/\partial u$ is Lipschitz with respect to (x,u) and $\alpha \in (0,2/L)$, where $L := (L_x l + L_u) + \lambda$,

$$l := \max_{t\in[t_0,t_f]} \{l_t(t)\}, \quad \lambda := \max_{t\in[t_0,t_f]} \{\lambda^R_{\max}(t)\},$$

$l_t(t)$ are Lipschitz constants of $x^u(t)$, for $t \in [t_0,t_f]$ and $\lambda^R_{\max}(t)$ is the maximal eigenvalue of $R(t)$. Then $\{u_{(l)}(\cdot)\}$ converges $\mathbb{L}^2\{[t_0,t_f];\mathbb{R}^m\}$ - weakly to a solution $u^*(\cdot)$ of (3).

Some Comments

- **future modifications: Armijo step sizes (Armijo line search), Exogenous step size, others**
- **from the computational point of view the fully convexified OCP (3) is related with a mathematically sophisticated procedure, namely, with the calculation of a convex envelope of a composite functional in Hilbert space**

Infimal Convolution Based Relaxation and PGs

two interesting concepts:

Definition

$J(u(\cdot))$ is locally para-convex around $u(\cdot) \in \mathbb{L}^2\{[t_0,t_f];\mathbb{R}^m\}$ if the infimal (prox) convolution $J_\lambda(u(\cdot))$ is convex and continuous on a δ-ball $\mathscr{B}_\delta(u(\cdot))$ around $u(\cdot)$ for some $\delta > 0$, $\lambda > 0$.

Definition

$J(u(\cdot))$ is prox-regular at $\hat{u}(\cdot) \in \mathbb{L}^2\{[t_0,t_f];\mathbb{R}^m\}$ if $\exists \varepsilon > 0$, $r > 0$ such that $J(u_1(\cdot)) > J(u_2(\cdot)) +$ $\langle \nabla J(\hat{u}(\cdot)), u_1(\cdot) - u_2(\cdot)\rangle_{\mathbb{L}^2\{[t_0,t_f];\mathbb{R}^m\}} - \frac{r}{2}||u_1(\cdot) - u_2(\cdot)||^2_{\mathbb{L}^2\{[t_0,t_f];\mathbb{R}^m\}}$ $\forall u_1(\cdot)$ from a ε-ball $\mathscr{B}_\varepsilon(\hat{u}(\cdot))$ around $\hat{u}(\cdot)$ whenever $u_2(\cdot) \in \mathscr{B}_\varepsilon(\hat{u}(\cdot))$ and $|J(u_1(\cdot)) - J(\hat{u}(\cdot))| < \varepsilon$.

Infimal Convolution Based Relaxation and PGs

infimal prox convolution for the original OCP (2)

$$J_\lambda(u(\cdot)) = \frac{1}{2}\int_{t_0}^{t_f} \left(\langle Q(t)x(t), x(t)\rangle + \langle (R(t)+\lambda I)u(t), u(t)\rangle\right) dt + \frac{1}{2}\langle Gx(t_f), x(t_f)\rangle$$

infimal convolution based OCP

$$\text{minimize } J_\lambda(u(\cdot)), \text{ subject to } (1), \; u(\cdot) \in \text{conv}(\mathscr{S}), \tag{4}$$

assume that (4) possesses an optimal solution $u_\lambda^{opt}(\cdot)$.

Infimal Convolution Based Relaxation and PGs

GM applied to (4)

$$u_{(l+1)}(\cdot) = \gamma_l \mathscr{P}_{\text{conv}(\mathscr{S})} \left[u_{(l)}(\cdot) - \alpha_l \nabla J_\lambda(u_{(l)}(\cdot)) \right] + (1-\gamma_l) u_{(l)}(\cdot), \; l \in \mathbb{N}$$

Theorem

Let $p_{n+1} \neq 0$ and $u_0^{opt}(\cdot) \in \text{int}\{conv(\mathscr{S})\}$. Consider $\{u_{(l)}(\cdot)\}$ generated by GM with a constant step size α. Then there exists $u_{(0)}(\cdot) \in conv(\mathscr{S})$ such that

$$\lim_{\lambda \to 0} \lim_{l \to \infty} J_\lambda(u_{(l)}(\cdot)) = \min_{conv(\mathscr{S})} J(u(\cdot)) = J(u_0^{opt}(\cdot)).$$

Numerical Treatment of the Initial OCP

numerical example

$$\begin{aligned}\dot{x}_1(t) &= u_1(t)\cos(x_3(t)),\\ \dot{x}_2(t) &= u_1(t)\sin(x_3(t)),\\ \dot{x}_3(t) &= u_2(t),\\ x(0) &= \begin{bmatrix} 15 & 15 & 180 \end{bmatrix}^T.\end{aligned}$$

$J(u(\cdot)) = \frac{1}{2}\int_0^1 \left(x_1^2(t) + x_2^2(t) + x_3^2(t)\right) dt$ and
$\mathcal{Q} = \{-50, -49, -48, \dots, 48, 49, 50\}$.

Numerical Treatment of the Initial OCP

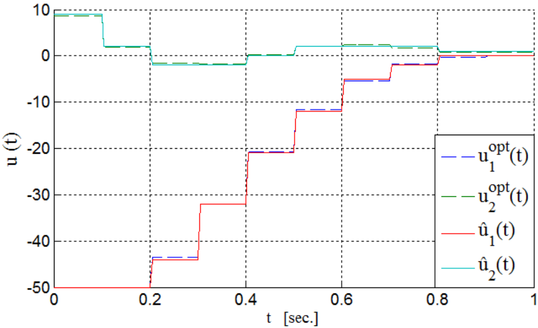


Figure: Optimal controls for the original and weakly relaxed OCPs

Numerical Treatment of the Initial OCP

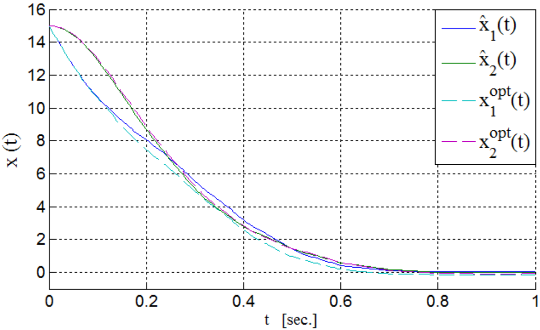
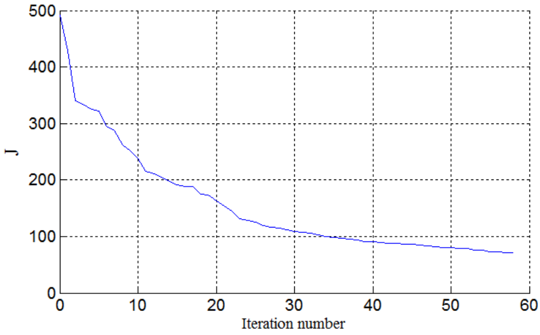


Figure: Optimal trajectories for the original and weakly relaxed OCPs

Numerical Treatment of the Initial OCP

numerical evaluation of the cost functional $J(\hat{u}(\cdot))$



THANKS!