On the optimal control of linear complementarity systems

Alexandre Vieira[1] - Bernard Brogliato[1] - Christophe Prieur[2]

[1] Univ. Grenoble Alpes, INRIA Grenoble - [2] Univ. Grenoble Alpes, GIPSA-Lab

26th September 2017

alexandre.vieira@inria.fr, bernard.brogliato@inria.fr, christophe.prieur@gipsa-lab.fr.

Introduction

Problem:

$$C(u) = \int_0^T \left(x(t)^\intercal Q x(t) + u(t)^\intercal U u(t)\right) dt \rightarrow \min$$

such that:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bv(t) + Fu(t) \\ 0 &\leq v(t) \perp Cx(t) + Dv(t) + Eu(t) \geq 0 \\ x(0) &= x_0, \ x(T) \text{ free} \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$, $B, F \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $D, E \in \mathbb{R}^{m\times m}$, $T > 0$, $x : [0, T] \rightarrow \mathbb{R}^n$ and $u, v : [0, T] \rightarrow \mathbb{R}^m$, Q and U matrices of according dimensions, supposed symmetric positive definite.

Hypothesis : D is a P-Matrix.

Motivation: Mechanics, Electronic Circuits, Chemical reactions

A difficult problem

- Existence of optimal solution not proved (classical Fillipov theory does not apply here due to lack of convexity). Cesari (2012), Theorem 9.2i and onwards
- Special cases arise when $E = 0$: switching modes are activated when the state reaches some threshold defined by the complementarity conditions. Georgescu et al. (2012), Passenberg et al. (2013)
- Since u is also involved $\implies$ mixed constraints; makes use of non-smooth analysis. Clarke and De Pinho (2010)

Direct method

Alexandre Vieira - Bernard Brogliato - Christophe Prieur

First way to compute numerical approximation: direct method.

$$\begin{aligned} \min \ & \sum_{k=0}^{N-1} \left(x_k^\intercal Q x_k + u_k^\intercal U u_k \right) \\ \text{s.t. } & \frac{x_{k+1} - x_k}{h} = A x_{k+1} + B v_k + F u_k, \forall k \in 0, ..., N-1 \\ & 0 \leq v_k \perp C x_k + D v_k + E u_k \geq 0 \end{aligned}$$

$\implies$ Mathematical Program with Equilibrium Constraints (MPEC)

Direct method

Such optimality problems are hard to tackle. Complementarity constraints:

$$\begin{aligned} &v \geq 0 \\ &Cx + Dv + Eu \geq 0 \\ &v^{\mathsf{T}}(Cx + Dv + Eu) = 0 \end{aligned}$$

violate usual constraint qualifications.

Need to redefine usual qualification for this problem, and associated stationarity properties.

Direct method

Denote λ^H, λ^G multipliers associated to $0 \leq v \perp Cx + Dv + Eu \geq 0$.

- Weak stationarity: $\lambda_i^G = 0$ if $v_i > 0 = (Cx + Dv + Eu)_i$ and $\lambda_i^H = 0$ if $v_i = 0 < (Cx + Dv + Eu)_i$
- Strong stationarity: Weak stationarity + $\lambda_i^G, \lambda_i^H \geq 0$ if $v_i = 0 = (Cx + Dv + Eu)_i$

Property

If (x^*, u^*, v^*) is a minimum for MPEC, it is weak stationary.
Here, if we suppose E invertible, the optimal solution is strong stationary.

Direct method

Suppose E invertible $\implies$ algorithm converging to a strong stationary point.

- [1] : Algorithm relaxing smartly the complementarity constraint, adding a parameter that continuously converge to 0.
- [2] : Complementarity added in the cost, creating a barrier problem solved with interior point method.

Under some conditions, both converge to strong stationary points.

---

[1] C. Kanzow and A. Schwartz. A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. SIAM Journal on Optimization, 23(2):770–798, 2013.

[2] S. Leyffer, G. López-Calva, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. SIAM Journal on Optimization, 17(1):52–77, 2006.

Why do we bother ?

This method works well... For small precision. Possible pseudominima?

$$\implies \text{Indirect method.}$$

Suppose an optimal solution exists $\implies$ Search for necessary conditions.
Two reasons for that:

- Useful for analyzing the solution (continuity, sensitivity...)
- Indirect method needs a good initial guess: direct method used for that.

Really general necessary conditions were obtained in [1]. But as such, they are not really practical (complicated hypothesis, really general equations...).

[1] L. Guo and J. J. Ye. Necessary optimality conditions for optimal control problems with equilibrium constraints (2016).

Weak stationarity

Define $S = \{(x, u, v) | 0 \leq v \perp Cx + Dv + Eu \geq 0\}$ and the partition of $\{0, ..., m\}$:

$$I_t^{0+}(x, u, v) = \{i \mid v_i(t) = 0 < (Cx(t) + Dv(t) + Eu(t))_i\}$$

$$I_t^{+0}(x, u, v) = \{i \mid v_i(t) > 0 = (Cx(t) + Dv(t) + Eu(t))_i\}$$

$$I_t^{00}(x, u, v) = \{i \mid v_i(t) = 0 = (Cx(t) + Dv(t) + Eu(t))_i\}$$

Theorem

Let (x^*, u^*, v^*) be a local minimizer of radius $R(\cdot)$. Suppose $\mathrm{Im}(C) \subseteq \mathrm{Im}(E)$. Then there exist an arc p and measurable functions $\lambda^G : \mathbb{R} \to \mathbb{R}^m$, $\lambda^H : \mathbb{R} \to \mathbb{R}^m$ such that the following conditions hold:

1. the transversality condition: $p(T) = 0$

Weak stationarity

Theorem

2. the Weierstrass condition for radius R: for almost every $t \in [t_0, t_1]$,

$$(x^*(t), u, v) \in S, \left\| \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u^*(t) \\ v^*(t) \end{pmatrix} \right\| < R(t)$$

$$\implies \langle p(t), Ax^*(t) + Bv + Fu) \rangle - \frac{1}{2}(x^*(t)^\intercal Q x^*(t) + u^\intercal U u)$$

$$\leq \langle p(t), Ax^*(t) + Bv^*(t) + Fu^*(t)) \rangle - \frac{1}{2}(x^*(t)^\intercal Q x^*(t) + u^*(t)^\intercal U u^*(t))$$

Weak stationarity

Theorem

3. the Euler adjoint equation: for almost every $t \in [0, T]$,

$$\begin{aligned}
\dot{p}(t) &= -A^\intercal p(t) + Qx^*(t) - C^\intercal \lambda^H(t) \\
0 &= F^\intercal p(t) - Uu^*(t) + E^\intercal \lambda^H(t) \\
0 &= B^\intercal p(t) + \lambda^G + D^\intercal \lambda^H(t) \\
0 &= \lambda_i^G(t), \ \forall i \in I_t^{+0}(x^*(t), u^*(t), v^*(t)) \\
0 &= \lambda_i^H(t), \ \forall i \in I_t^{0+}(x^*(t), u^*(t), v^*(t))
\end{aligned}$$

Euler equation

How can we solve the following BVP?

$$
\begin{aligned}
\dot{x} &= Ax + Bv + Fu \\
\dot{p} &= -A^\intercal p + Qx - C^\intercal \lambda^H \\
0 &= F^\intercal p - Uu + E^\intercal \lambda^H \\
0 &= B^\intercal p + \lambda^G + D^\intercal \lambda^H \\
0 &= \lambda_i^G(t),\ \forall i \in I_t^{+0}(x(t), u(t), v(t)) \\
0 &= \lambda_i^H(t),\ \forall i \in I_t^{0+}(x(t), u(t), v(t)) \\
x_0 &= x(0), \\
0 &= p(T)
\end{aligned}
$$

Euler equation

How can we solve the following BVP?

$$\dot{x} = Ax + Bv + Fu$$
$$\dot{p} = -A^\intercal p + Qx - C^\intercal \lambda^H$$
$$0 = F^\intercal p - Uu + E^\intercal \lambda^H \rightarrow \text{ isolate } u$$
$$0 = B^\intercal p + \lambda^G + D^\intercal \lambda^H \rightarrow \text{ isolate } \lambda^G$$

$$\boxed{\begin{aligned} 0 &= \lambda_i^G(t),\ \forall i \in I_t^{+0}(x(t), u(t), v(t)) \\ 0 &= \lambda_i^H(t),\ \forall i \in I_t^{0+}(x(t), u(t), v(t)) \end{aligned}}$$

$$\boxed{\begin{aligned} x_0 &= x(0), \\ 0 &= p(T) \end{aligned}}$$

Strong stationarity

$$\boxed{\begin{array}{l} 0 = \lambda_i^G(t), \ \forall i \in I_t^{+0}(x(t), u(t), v(t)) \\ 0 = \lambda_i^H(t), \ \forall i \in I_t^{0+}(x(t), u(t), v(t)) \end{array}}$$

We miss a piece of information: what happens on I_t^{00} ?

Proposition

Let (x^*, u^*, v^*) be a local minimizer and suppose E invertible. Then (x^*, u^*, v^*) is strongly stationary, meaning:

$$\lambda_i^G(t) \geq 0, \lambda_i^H(t) \geq 0, \ \forall i \in I_t^{00}(x(t), u(t), v(t))$$

Strong stationarity

$$\boxed{\begin{array}{ll} 0 = \lambda_i^G(t), & \forall i \in I_t^{+0}(x(t), u(t), v(t)) \\ 0 = \lambda_i^H(t), & \forall i \in I_t^{0+}(x(t), u(t), v(t)) \\ \lambda_i^G(t) \geq 0, \lambda_i^H(t) \geq 0, & \forall i \in I_t^{00}(x(t), u(t), v(t)) \end{array}}$$

Almost like a linear complementarity problem!

Strong stationarity

Theorem

Let (x^*, u^*, v^*) be a local minimizer and suppose E invertible. Fix an arbitrary $r > 0$. Then there exist an arc p and measurable functions $\beta : [0, T] \to \mathbb{R}^m$, $\zeta : [0, T] \to \mathbb{R}$ such that, $u^*(t) = U^{-1}\left(F^\intercal p(t) + E^\intercal \beta(t) - (\zeta(t) + r)E^\intercal v^*(t)\right)$ and:

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \mathcal{A}_r(\zeta) \begin{pmatrix} x \\ p \end{pmatrix} + \mathcal{B}_r(\zeta) \begin{pmatrix} \beta \\ v^* \end{pmatrix}$$

$$\begin{cases} 0 \leq \begin{pmatrix} \beta \\ v^* \end{pmatrix} \perp \mathcal{D}_r(\zeta) \begin{pmatrix} \beta \\ v^* \end{pmatrix} + \mathcal{C}_r(\zeta) \begin{pmatrix} x \\ p \end{pmatrix} \geq 0 \\ \beta \geq r v^* \\ x(0) = x_0, \;\; p(T) = 0 \end{cases}$$

How to solve a BVP LCS

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \mathcal{A} \begin{pmatrix} x \\ p \end{pmatrix} + \mathcal{B} \begin{pmatrix} \beta \\ v \end{pmatrix}$$

$$\begin{cases} 0 \leq \begin{pmatrix} \beta \\ v \end{pmatrix} \perp \mathcal{D} \begin{pmatrix} \beta \\ v \end{pmatrix} + \mathcal{C} \begin{pmatrix} x \\ p \end{pmatrix} \geq 0 \\ \beta \geq rv \\ \boxed{x(0) = x_0,\ p(T) = 0} \end{cases}$$

Numerically, we usually do shooting: find the good $p(0) = p_0$ such that the computed solution $p(t; p_0)$ complies with $p(T; p_0) = 0$: nonsmooth Newton method.

- Need for an initial guess close enough
- How to compute a sensitivity matrix for $p(T; \cdot)$?

How to solve a BVP LCS

$$\boxed{\begin{aligned} &\dot{z} = \mathcal{A}z + \mathcal{B}\Lambda \\ &0 \leq \Lambda \perp \mathcal{D}\Lambda + \mathcal{C}z \geq 0 \end{aligned}}$$

Denote $\mathcal{T}_h(z)$ a linear Newton Approximation to the solution Λ of the LCP. Then, a linear Newton approximation for the solution map $z(T, \cdot)$ can be obtained by solving the DI in matrix function:

$$\dot{J}(t) \in \mathcal{A}J(t) + (\text{co } \mathcal{T}_h(z(t;\xi)))J(t), \; J(0) = I$$

JS Pang, D. Stewart, Solution dependence on initial conditions in differential variational inequalities (2009)

A 1D example

$$\int_0^T (x(t)^2 + u(t)^2)dt \to \min$$
$$\dot{x} = ax + bv + fu$$
$$0 \leq v \perp dv + eu \geq 0$$
$$x(0) = x_0$$

A 1D example

Alexandre Vieira - Bernard Brogliato - Christophe Prieur

We can show that the (strong) stationary solution in this case is given by:

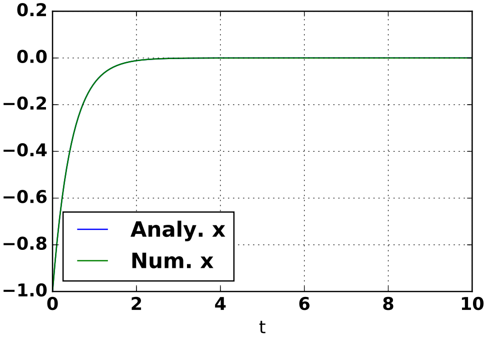
$$p(t) = \left[\cosh(\sqrt{\gamma}t) - \frac{a}{\sqrt{\gamma}}\sinh(\sqrt{\gamma}t)\right] p(0) + \frac{\sinh(\sqrt{\gamma}t)}{\sqrt{\gamma}}x(0)$$

$$p(0) = -\frac{\sinh(\sqrt{\gamma}T)}{\sqrt{\gamma}\cosh(\sqrt{\gamma}T) - a\sinh(\sqrt{\gamma}T)}x(0).$$

$$u(t) = \begin{cases} fp(t) & \text{if } efp(0) \geq 0, \\ \left(f - \frac{eb}{d}\right) p(t) & \text{if } efp(0) \leq 0. \end{cases}$$

$$x(t) = \dot{p}(t) + ap(t).$$

A 1D example



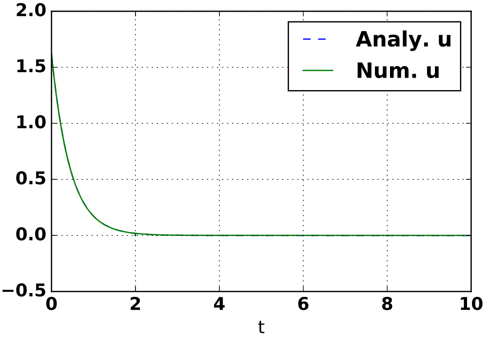


Figure: Solution via indirect method : state x and control u, on $[0, 10]$.
$a = 1, b = 0.5, d = 1, e = -2, f = 3, x(0) = -1$. Initial guess with direct method and 300 nodes. Indirect method with 10 000 nodes and 20 intervals of shooting. Obtained in 54s. (In order to have this same precision with the direct method : 453s.)

A 2D example

$$\min \int_0^1 \left(\|x(t)\|_2^2 + u(t)^2\right) dt$$

$$\text{s.t. } \dot{x} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} x + \begin{pmatrix} -1 \\ 1 \end{pmatrix} v + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$

$$0 \leq v \perp \begin{pmatrix} 3 & -1 \end{pmatrix} x + v + 2u \geq 0$$

$$x(0) = \begin{pmatrix} -0.5 \\ -1 \end{pmatrix}$$

A 2D example

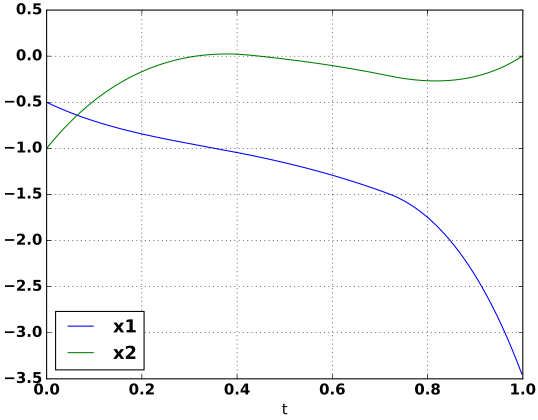
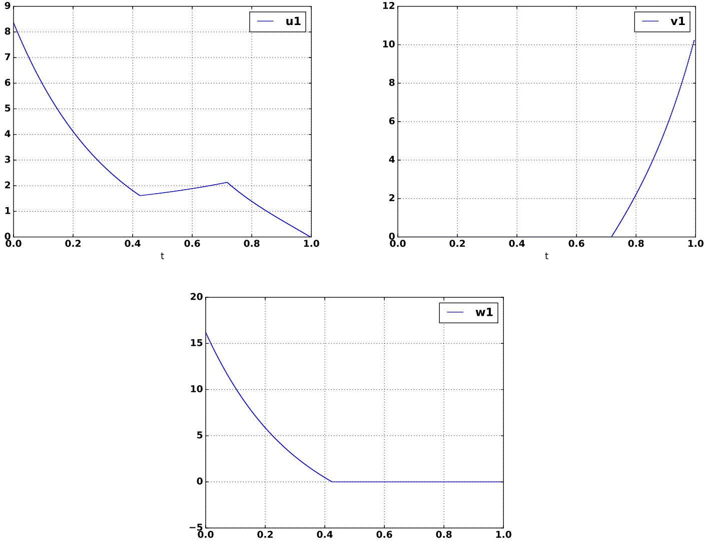


Figure: Solution via direct method for previous example : state

A 2D example



Conclusion

- First stationarity results, that we can use analytically and numerically.
- Numerical algorithms working fast, even with high precision.

What is left to be done:

- The stationarity LCS, even in this case, still is not entirely analysed.
- Drop some assumptions (E invertible, D P-matrix...).

(For those interested: the whole code will be soon on `https://gitlab.inria.fr/avieira/optLCS`)

Strong stationarity

Define, for two scalars ζ and r:

$$\mathcal{A}_r(\zeta) = \begin{pmatrix} A & FU^{-1}F^\intercal \\ Q & -A^\intercal \end{pmatrix}$$

$$\mathcal{B}_r(\zeta) = \begin{pmatrix} FU^{-1}E^\intercal & B-(\zeta+r)FU^{-1}E^\intercal \\ -C^\intercal & (\zeta+r)C^\intercal \end{pmatrix}$$

$$\mathcal{C}_r(\zeta) = \begin{pmatrix} C & EU^{-1}F^\intercal \\ \zeta C & \zeta EU^{-1}F^\intercal - B^\intercal \end{pmatrix}$$

$$\mathcal{D}_r(\zeta) = \begin{pmatrix} EU^{-1}E^\intercal & D-(\zeta+r)EU^{-1}E^\intercal \\ \zeta EU^{-1}E^\intercal - D^\intercal & \zeta D + (\zeta+r)\left(D^\intercal - \zeta EU^{-1}E^\intercal\right) \end{pmatrix}$$