Double Smoothing Algorithm for a class of Infinite-dimensional Optimization Problems

O. Devolder (F.R.S.-FNRS Research Fellow),
F. Glineur and Y. Nesterov

CORE-INMA, UCL

Blankenberge, March 12, 2010

Outline

1. Studied problem class with application in optimal control.
2. Dual Approach.
3. Double Regularization.
4. Solving the dual problem.
5. Reconstruction of a primal solution.
6. Conclusion and Further Research.

Outline

1. Studied problem class with application in optimal control.
2. Dual Approach.
3. Double Regularization.
4. Solving the dual problem.
5. Reconstruction of a primal solution.
6. Conclusion and Further Research.

Linear optimal control problem with scalar control

Differential system:

$$\mathbf{x}'(t) = A(t)\mathbf{x}(t) + B(t)u(t)$$

with $\mathbf{x}(t) = (x_1(t), ..., x_m(t))^T$, $A(t) \in \mathbb{R}^{m\times m}$ and $B(t) \in \mathbb{R}^{m\times 1}$.

Initial and final conditions for the state

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad \mathbf{x}(t_f) = \mathbf{x}_f.$$

Control constraints

$$0 \leq u(t) \leq 1 \quad \forall t \in [t_0, t_f].$$

Performance Measure

$$\min \int_{t_0}^{t_f} < \mathbf{c}(t), \mathbf{x}(t) > + < \mathbf{d}(t), \mathbf{x}'(t) > + e(t)u(t)dt.$$

State equation of the linear system

The state generated by the linear system for a control $u(t)$ is given by:

$$\mathbf{x}(t) = X(t, t_0)\mathbf{x}(t_0) + X(t, t_0)\left(\int_{t_0}^{t} X^{-1}(\tau, t_0)B(\tau)u(\tau)d\tau\right)$$

where $X(t, t_0)$ is the state transition matrix of the system i.e the unique solution of the matricial Cauchy problem:

$$\frac{d}{dt}X(t, t_0) = A(t)X(t, t_0)$$

$$X(t_0, t_0) = I.$$

Consequence: We can express all only in terms of the control

- Replacing $\mathbf{x}(t)$ by this expression, we have a performance measure of the form:

$$\int_{t_0}^{t_f} f(t)u(t)dt.$$

- Final conditions is satisfied by the state iff

$$\int_{t_0}^{t_f} X^{-1}(\tau, t_0)B(\tau)u(\tau)d\tau = X^{-1}(t_f, t_0)\mathbf{x}_f - \mathbf{x}_0.$$

  Using the explicit expression for $X(t, t_0)$ and $B(t)$, we have m linear equality constraints:

$$\int_{t_0}^{t_f} a_i(t)u(t)dt = b_i \quad i = 1, \ldots, m.$$

New formulation

Our initial optimal control problem can be expressed as an **infinite-dimensional linear programming** problem:

$$\inf_u \int_{t_0}^{t_f} f(t)u(t)dt$$

$$\int_{t_0}^{t_f} a_i(t)u(t)dt = b_i \quad \forall i = 1, ..., m$$

$$0 \leq u(t) \leq 1 \quad \forall t \in [t_0, t_f].$$

Studied problem class: Linear Programming in $L^\infty([\alpha, \beta])$.

$$P^* = \inf_{x \in L^\infty([\alpha,\beta])} \int_\alpha^\beta f(t)x(t)dt$$

$$\int_\alpha^\beta a_i(t)x(t)dt = b_i \quad \forall i = 1, ..., m$$

$$M_1 \leq x(t) \leq M_2 \quad \text{a.e. in } [\alpha, \beta]$$

where

- $b \in \mathbb{R}^m$
- $f \in L^1([\alpha, \beta])$
- $a_i \in L^1([\alpha, \beta])$ for all $i = 1...m$.

Outline

Primal Problem

$$P^* = \inf_{x \in L^\infty([\alpha,\beta])} \int_\alpha^\beta f(t)x(t)dt$$

$$\int_\alpha^\beta a_i(t)x(t)dt = b_i \quad \forall i = 1, \ldots, m$$

$$M_1 \leq x(t) \leq M_2 \quad \text{a.e. in } [\alpha, \beta].$$

If we dualize the linear equality constraints, we obtain the equivalent formulation:

$$P^* = \inf_{x \in \overline{X}} \sup_{y \in \mathbb{R}^m} \int_\alpha^\beta f(t)x(t)dt + \sum_{i=1}^m y_i \left(\int_\alpha^\beta a_i(t)x(t)dt - b_i \right)$$

where $\overline{X} = \{x \in L^\infty([\alpha, \beta]) : M_1 \leq x(t) \leq M_2 \text{ a.e. in } [\alpha, \beta]\}$.

Dual Problem

$$
\begin{aligned}
D^* &= \sup_{y\in\mathbb{R}^m} \inf_{x\in\overline{X}} \int_\alpha^\beta f(t)x(t)dt + \sum_{i=1}^m y_i \left(\int_\alpha^\beta a_i(t)x(t)dt - b_i \right) \\
&= \sup_{y\in\mathbb{R}^m} -\sum_{i=1}^m y_i b_i - \phi(y)
\end{aligned}
$$

where

$$
\phi(y) = \sup_{x\in\overline{X}} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t) \right) x(t)dt.
$$

Why a Dual Approach ?

Advantages of the dual problem:

- The dual is an unconstrained optimization problem in finite-dimension ($y \in \mathbb{R}^m$).
- The infinite-dimensional problem defining $\phi(y) = \sup_{M_1 \leq x(t) \leq M_2} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t)\right) x(t) dt$ is easy:

$$x_y(t) = \begin{cases} M_1, & \text{when } f(t) + \sum_{i=1}^m y_i a_i(t) > 0, \\ M_2, & \text{when } f(t) + \sum_{i=1}^m y_i a_i(t) < 0 \\ \frac{M_1+M_2}{2}, & \text{when } f(t) + \sum_{i=1}^m y_i a_i(t) = 0 \end{cases}$$

  is an optimal solution.
- There is Strong Duality i.e : $P^* = D^*$ when the dual is solvable.

But θ can be non-differentiable...

If we rewrite the dual problem as a minimization problem:

$$-D^* = \min_{y \in \mathbb{R}^m} \theta(y) = \min_{y \in \mathbb{R}^m} b^T y + \phi(y)$$

where $\phi(y) = \sup_{x \in \overline{X}} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t)\right) x(t) dt$,

θ can be **non-differentiable**:

$$\partial\theta(y) = \left\{ b - A\tilde{x} = \left(b_1 - \int_\alpha^\beta a_1(t)\tilde{x}(t)dt, ..., b_m - \int_\alpha^\beta a_m(t)\tilde{x}(t)dt \right)^T \right.$$

$$\left. \text{for any optimal solution } \tilde{x} \text{ of the problem defining } \phi(y) \right\}$$

and the optimization problem defining $\phi(y)$ can have multiple optimal solutions.

Conclusion:

We have to solve a non-smooth convex optimization problem.

How to solve a non-smooth convex problem ?

- **The classical approach: subgradient-type scheme.**
  Advantage : Can be applied directly on the dual objective function without any regularization
  Disadvantage: Slow Convergence
  $$\theta(y_k) \to \theta^* \text{ in } O\left(\frac{1}{\epsilon^2}\right).$$
- **The smoothing approach.**
  We modify the dual objective function in order to be able to apply more efficient scheme of smooth convex optimization.
  Advantage : Faster convergence, we will obtain a scheme such that
  $$\theta(y_k) \to \theta^* \text{ in } O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right).$$
  Disadvantage : We have to modify the dual objective function with two regularizations.

Outline

1. Studied problem class with application in optimal control.
2. Dual Approach.
3. Double Regularization.
4. Solving the dual problem.
5. Reconstruction of a primal solution.
6. Conclusion and Further Research.

Double Regularization of the dual objective function

In order to

- be able to solve efficiently the dual problem
- to be able to obtain a nearly optimal and feasible primal solution from a nearly optimal dual solution

we will modify the dual objective function with two regularizations:

1. A first regularization on the infinite-dimensional side of the problem
2. A second regularization on the finite-dimensional side of the problem.

First regularization

Why?
In order to obtain a smooth objective function for the dual problem with gradient Lipschitz-continuous i.e.:

$$\|\nabla g(y) - \nabla g(\bar{y})\| \leq L \|y - \bar{y}\| \quad \forall y, \bar{y} \text{ with } L < +\infty$$

and therefore be able to apply efficient schemes of smooth convex optimization.

Where?
On the infinite-dimensional side of the problem.

First regularization: How ?

How?
Modify the dual objective function:

$$\begin{aligned}\theta(y) &= b^T y + \phi(y)\\ &= b^T y + \sup_{x \in \overline{X}} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t) \right) x(t) dt\end{aligned}$$

in

$$\begin{aligned}\theta_\mu(y) &= b^T y + \phi_\mu(y)\\ &= b^T y + \sup_{x \in \overline{X}} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t) \right) x(t) - \frac{\mu}{2} x^2(t) dt\end{aligned}$$

with $\mu > 0$.

First regularization: θ_μ is differentiable

With the addition of the strongly concave function $-\frac{\mu}{2}\int_\alpha^\beta x^2(t)dt$, the optimization problem defining $\phi_\mu(y)$ has only one optimal solution given by:

$$x_{y,\mu}(t) = \begin{cases} M_1, & \text{if } f(t) + \sum_{i=1}^m y_i a_i(t) \geq -\mu M_1, \\ M_2, & \text{if } f(t) + \sum_{i=1}^m y_i a_i(t) \leq -\mu M_2 \\ \frac{-1}{\mu}(f(t) + \sum_{i=1}^m y_i a_i(t)), & \text{else.} \end{cases}$$

The function θ_μ is therefore differentiable.

First regularization: Further properties of θ_μ

θ_μ is gradient Lipschitz-continuous with constant

$$L_\mu = \frac{\|A\|^2}{\mu}$$

where

$$Ax = \left(\int_\alpha^\beta a_1(t)x(t)dt, \ldots, \int_\alpha^\beta a_m(t)x(t)dt \right)^T$$

and

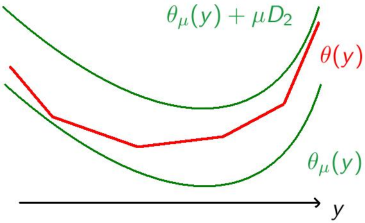
$$\|A\| = \sup_{x \in L^2([\alpha,\beta]), y \in \mathbb{R}^m} \{ < Ax, y >: \|x\|_2 = 1, \|y\|_2 = 1 \}.$$

First regularization: Further properties of θ_μ

θ_μ is a good approximation of θ with absolute accuracy bound depending on μ:

$$\theta_\mu(y) \leq \theta(y) \leq \theta_\mu(y) + \mu D_2 \quad \forall y \in \mathbb{R}^m$$

where $D_2 = \frac{1}{2}(\beta - \alpha) \max\{|M_1|, |M_2|\}^2$.



Second regularization

Why?

We want not only

- to solve the dual problem

but also

- to reconstruct from the obtained nearly optimal dual solution, a nearly optimal and feasible primal solution.

Reconstruction of a primal solution

For a given dual iterate y_k, if we consider the function

$$x_k(t) = x_{y_k,\mu}(t),$$

the unique optimal solution of the problem defining $\phi_\mu(y_k)$, we have:

- $$\left| \int_\alpha^\beta f(t)x_k(t)dt - P^* \right| \leq Cste \, |\theta(y_k) - \theta^*| + Cste \, \|\nabla\theta_\mu(y_k)\|$$
- $$\|Ax_k - b\| = \|\nabla\theta_\mu(y_k)\| \,.$$

The quality of this primal solution depends not only on the convergence rate of $\theta(y_k)$ to θ^* but also on the convergence rate of $\|\nabla\theta_\mu(y_k)\|$ to 0.

Convexity and Gradient Lipschitz-continuity are not enough

If the dual objective function is convex, gradient Lipschitz-continuous and if we apply the optimal scheme for $F_L^{1,1}(\mathbb{R}^m)$:

$$g(y_k) \to g^* \text{ in } O\left(\frac{1}{\sqrt{\epsilon}}\right)$$

but the convergence of the gradient is slower:

$$\|\nabla g(y_k)\| \to 0 \text{ in } O\left(\frac{1}{\epsilon}\right).$$

Convexity and Gradient Lipschitz-continuity are not enough

In our case, if we apply this scheme to our function $\theta_\mu \in F^{1,1}_{L_\mu}(\mathbb{R}^m)$ with a good choice for μ, we have

$$\theta(y_k) - \theta^* \to 0 \text{ in } O\left(\frac{1}{\epsilon}\right)$$

but

$$\|\nabla \theta_\mu(y_k)\| \to 0 \text{ in } O\left(\frac{1}{\epsilon^2}\right).$$

Therefore if the dual objective function is only convex and gradient Lipschitz-continuous, we have a convergence rate in $O\left(\frac{1}{\epsilon^2}\right)$ for the primal sequence.

This is not better than with the subgradient scheme!

Is the smoothing approach useless ?

No!!!

If the dual objective function is also strongly convex, we can apply the optimal scheme for $S_{\sigma,L}^{1,1}(\mathbb{R}^m)$ for wich we have the same rate of convergence for $g(y_k) - g^*$ and $\|\nabla g(y_k)\|$ in

$$O\left(\exp(\left(-k\sqrt{\frac{\sigma}{L}}\right)\right).$$

Conclusion: We want a strongly convex objective function on the dual.

Second Regularization

Where?
On the finite-dimensional side of the problem.
How?
Let $\sigma > 0$ and $y_0 \in \mathbb{R}^m$, adding to the function θ_μ, the strongly convex function $\frac{\sigma}{2}\|y - y_0\|^2$, we obtain the function:

$$\theta_{\mu,\sigma}(y) = \theta_\mu(y) + \frac{\sigma}{2}\|y - y_0\|^2$$

wich is strongly convex with parameter σ .

Modified Dual Objective Function

$$\theta_{\mu,\sigma}(y) = b^T y + \phi_\mu(y) + \frac{\sigma}{2} \|y - y_0\|^2$$

where

- $\phi_\mu(y) = \sup_{x \in \overline{X}} \int_\alpha^\beta \left(-f(t) - \sum_{i=1}^m y_i a_i(t)\right) x(t) - \frac{\mu}{2} x^2(t) dt$
- y_0 is any element in $\mathbb{R}^m$
- $\sigma > 0$
- $\mu > 0$.

This function is:

- Strongly convex with parameter σ
- still Gradient Lipschitz-continuous now with constant $L_\mu + \sigma$.

Outline

Optimal Scheme for $S^{1,1}_{\sigma,L}(\mathbb{R}^m)$

Let $g : \mathbb{R}^m \to \mathbb{R}$ be

- strongly convex with parameter $\sigma > 0$
- gradient Lipschitz-continuous with constant $L > 0$.

Algorithm

1. Initialization
   Choose $w_0 = y_0 \in \mathbb{R}^m$.
2. Iteration $(k \geq 0)$
   Set

$$y_{k+1} = w_k - \frac{1}{L}\nabla g(w_k)$$

$$w_{k+1} = y_{k+1} + \frac{\sqrt{L} - \sqrt{\sigma}}{\sqrt{L} + \sqrt{\sigma}}(y_{k+1} - y_k).$$

Optimal Scheme for $S_{\sigma,L}^{1,1}(\mathbb{R}^m)$: Convergence Results

1. **Convergence of the objective values:**

$$g(y_k) - g^* \le \left(g(y_0) - g^* + \frac{\sigma}{2}\|y_0 - y^*\|^2\right) \exp\left(-k\sqrt{\frac{\sigma}{L}}\right)$$

2. **Convergence of the gradient:**

$$\|\nabla g(y_k)\| \le \sqrt{2L}\sqrt{g(y_0) - g^* + \frac{\sigma}{2}\|y_0 - y^*\|^2} \exp\left(-\frac{k}{2}\sqrt{\frac{\sigma}{L}}\right)$$

3. **Convergence of the iterates:**

$$\|y_k - y^*\| \le \sqrt{\frac{2}{\sigma}}\sqrt{g(y_0) - g^* + \frac{\sigma}{2}\|y_0 - y^*\|^2} \exp\left(-\frac{k}{2}\sqrt{\frac{\sigma}{L}}\right).$$

Solving the dual problem using the optimal scheme for $S^{1,1}_{\sigma,L_\mu+\sigma}(\mathbb{R}^m)$

If we apply the optimal scheme for $S^{1,1}_{\sigma,L_\mu+\sigma}(\mathbb{R}^m)$ to our modified dual function, we have:

$$\theta(y_k) - \theta^* \leq \mu D_2 + \frac{\sigma}{2}\|y^{**} - y_0\|^2$$

$$+2(\theta(y_0) - \theta^* + \mu D_2)\exp\left(-k\sqrt{\frac{\sigma}{L_\mu + \sigma}}\right)$$

$$+\sqrt{2\sigma}\sqrt{\|y^{**} - y_0\|^2 + \frac{2\mu}{\sigma}D_2}\sqrt{\theta(y_0) - \theta^* + \mu D_2}\exp\left(-\frac{k}{2}\sqrt{\frac{\sigma}{L_\mu + \sigma}}\right)$$

where y^{**} is an optimal solution of the original dual problem $\min_{y \in \mathbb{R}^m} \theta(y)$ and $\theta^* = -P^*$ is the optimal value of this problem. If we want an accuracy $\theta(y_k) - \theta^* \leq \epsilon$, we can choose μ, σ and k such that each of the four terms are $\leq \epsilon/4$.

Choice of μ and σ

- If we want

$$\mu D_2 \leq \frac{\epsilon}{4}$$

we choose

$$\mu(\epsilon) = \frac{1}{4D_2}\epsilon = C_1\epsilon.$$

- If we want

$$\frac{\sigma}{2}\|y^{**} - y_0\|^2 \leq \frac{\epsilon}{4}$$

we choose

$$\sigma(\epsilon) = \frac{\epsilon}{2\|y^{**} - y_0\|^2} = C_2\epsilon.$$

Number of iterations needed

- If we want $2(\theta(y_0) - \theta^* + \mu(\epsilon) D_2) \exp\left(-k\sqrt{\frac{\sigma(\epsilon)}{L_\mu(\epsilon)+\sigma(\epsilon)}}\right) \leq \frac{\epsilon}{4}$ we have to choose

$$k(\epsilon) \geq g_1(\epsilon) = O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right).$$

- If we want $\sqrt{2\sigma(\epsilon)}\sqrt{\|y^{**} - y_0\|^2 + \frac{2\mu(\epsilon)}{\sigma(\epsilon)} D_2}\sqrt{\theta(y_0) - \theta^* + \mu(\epsilon) D_2}$ $\exp\left(-\frac{k}{2}\sqrt{\frac{\sigma(\epsilon)}{L_\mu(\epsilon)+\sigma(\epsilon)}}\right) \leq \frac{\epsilon}{4}$ we have to choose

$$k(\epsilon) \geq g_2(\epsilon) = O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right).$$

Convergence of the gradient

In order to reconstruct from the nearly optimal dual solution, a nearly feasible and optimal primal solution with a given accuracy, we need the rate of convergence of $\|\nabla\theta_\mu(y_k)\|$.

We have:

$$\|\nabla\theta_\mu(y_k)\| \leq \sigma\sqrt{\|y^{**} - y_0\|^2 + \frac{2\mu}{\sigma}D_2}$$

$$+\left(\sqrt{2(L_\mu+\sigma)} + \sqrt{2\sigma}\right)\sqrt{\theta(y_0) - \theta^* + \mu D_2}\,\exp\left(-\frac{k}{2}\sqrt{\frac{\sigma}{L_\mu+\sigma}}\right).$$

Convergence of the gradient: Number of iterations

With $\mu = C_1\epsilon$ and $\sigma = C_2\epsilon$, we have:

$$\sigma\sqrt{\|y^{**}-y_0\|^2+\frac{2\mu}{\sigma}D_2} = \left(C_2\sqrt{\|y^{**}-y_0\|^2+\frac{2C_1}{C_2}D_2}\right)\epsilon = C_3\epsilon.$$

Furthermore, if we want

$$\left(\sqrt{2(L_\mu+\sigma)}+\sqrt{2\sigma}\right)\sqrt{\theta(y_0)-\theta^*+\mu D_2}\,\exp\left(-\frac{k}{2}\sqrt{\frac{\sigma}{L_\mu+\sigma}}\right) \leq \epsilon$$

we have to take

$$k(\epsilon) \geq g_3(\epsilon) = O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right).$$

Complexity

Let $\epsilon > 0$ and choose $\mu(\epsilon) = C_1\epsilon$, $\sigma(\epsilon) = C_2\epsilon$, after

$$k(\epsilon) = \max\{g_1(\epsilon), g_2(\epsilon), g_3(\epsilon)\} = O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right)$$

iterations, we have:

- $$\theta(y_{k(\epsilon)}) - \theta^* \leq \epsilon$$
- $$\left\|\nabla\theta_\mu(y_{k(\epsilon)})\right\| \leq (C_3 + 1)\epsilon.$$

Outline

A nearly feasible and optimal primal solution

Using the dual iterate $y_{k(\epsilon)}$, we can obtain a primal solution with the same order of accuracy.
Consider

$$x_{k(\epsilon)}(t) = x_{y_{k(\epsilon)},\mu(\epsilon)}(t)$$

the unique optimal solution of the problem defining $\phi_{\mu(\epsilon)}(y_{k(\epsilon)})$ that we can compute analytically.

A nearly feasible and optimal primal solution (2)

This function $x_{k(\epsilon)}$ is:

- In $\overline{X}$ by construction.
- Nearly optimal for the primal problem:

$$\left| \int_\alpha^\beta f(t) x_{k(\epsilon)}(t) dt - P^* \right| \leq ((C_3+1)C_4 + C_1 D_2 + \max\{1, C_1 D_2\})\epsilon.$$

where $C_4 = \sqrt{\|y^{**} - y_0\|^2 + \frac{2C_1}{C_2} D_2} + \|y_0\|$.

- Nearly feasible for the linear equality constraints:

$$\|\nabla \theta_\mu(y_k(\epsilon))\| = \left\| A x_{k(\epsilon)} - b \right\| \leq (C_3 + 1)\epsilon.$$

Outline

1. Studied problem class with application in optimal control.
2. Dual Approach.
3. Double Regularization.
4. Solving the dual problem.
5. Reconstruction of a primal solution.
6. Conclusion and Further Research.

Double Smoothing Algorithm: Conclusion

After $k(\epsilon) = \max\{g_1(\epsilon), g_2(\epsilon), g_3(\epsilon)\} = O\left(\frac{1}{\epsilon} \ln\left(\frac{1}{\epsilon}\right)\right)$ iterations, we have:

- A nearly optimal dual solution:

$$\theta(y_k(\epsilon)) - \theta^* \leq \epsilon$$

- A nearly optimal and feasible primal solution:

$$\left| \int_\alpha^\beta f(t) x_{k(\epsilon)}(t) dt - P^* \right| \leq ((C_3+1)C_4 + C_1 D_2 + \max\{1, C_1 D_2\})\epsilon$$

$$\left\| A x_{k(\epsilon)} - b \right\| \leq (C_3 + 1)\epsilon.$$

Come back to the Optimal Control Problem

Applying the double smoothing algorithm to the optimal control problem, we generate after $k(\epsilon) = O\left(\frac{1}{\epsilon}\ln\left(\frac{1}{\epsilon}\right)\right)$ iterations, a control $u_{k(\epsilon)}(t)$:

- Nearly optimal for the performance measure:

$$\left|\int_{t_0}^{t_f} < \mathbf{c}(t), \mathbf{x}(t) > + < \mathbf{d}(t), \mathbf{x}'(t) > + e(t)u_{k(\epsilon)}(t)dt - P^*\right|$$

$$\leq ((C_3 + 1)C_4 + C_1 D_2 + \max\{1, C_1 D_2\})\epsilon$$

- Satisfying nearly the initial and final conditions for the states:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad \mathbf{x}(t_f) = \mathbf{x}_f + \Delta$$

  where $\|\Delta\| \leq \|X(t_f, t_0)\| (1 + C_3)\epsilon$

- Satisfying the control constraints:

$$0 \leq u_{k(\epsilon)}(t) \leq 1 \quad \forall t \in [t_0, t_f].$$

Further Research

- Introduction of integral inequality constraints

$$\int_{\alpha}^{\beta} a_j(t)x(t) \leq b_j$$

- Generalization to vector-valued function $\mathbf{x}(t)$ in order to consider optimal control problem with multiple controls

$$\mathbf{u}(t) = (u_1(t), ..., u_n(t))^T$$

- Consider more general pointwise constraints:

$$\mathbf{x}(t) \in Q(t) \quad \forall t \in [\alpha, \beta]$$

  where $Q(t)$ is a bounded convex set for all $t \in [\alpha, \beta]$

- ...

Thanks for your attention !