Sensitivity analysis for relaxed optimal control problems with final-state constraints

J. Frédéric Bonnans, [1] Laurent Pfeiffer,[1] and Oana Silvia Serea[2]

[1]Inria-Saclay, CMAP, Ecole Polytechnique

[2]Université de Perpignan

Abstract

We consider a family of optimal control problems with final-state constraints parameterized by a nonnegative variable $\theta \geq 0$. The value function is denoted by $V(\theta)$. We consider bounded strong solutions to these problems, *ie*, optimal solutions in a small neighborhood in L^∞ for trajectories and a large bounded neighborhood for the controls. Our aim is to obtain a second-order expansion of $V(\theta)$ near 0. By introducing relaxed controls, we are able to deal with a wide class of perturbations and we obtain sharp estimates.

1 Formulation of the problem

1.1 Setting

For a control u in $L^\infty([0,T],\mathbb{R}^m)$ and $\theta \geq 0$, consider the trajectory $y[u,\theta]$ solution of the following differential system:

$$\begin{cases} \dot{y}_t = f(u_t, y_t, \theta), & \text{for a. a. } t \text{ in } [0,T], \\ y_0 = y^0(\theta). \end{cases}$$

We set $K = \{0_{n_E}\} \times \mathbb{R}_+^{n_I}$. The family of optimal control problems that we consider is the following:

$$\underset{u \in L^\infty([0,T],\mathbb{R}^n)}{\text{Min}} \phi(y_T[u,\theta]), \quad \text{s.t. } \Phi(y_T[u],\theta) \in K.$$

A control $\overline{u}$ is said to be a **bounded strong solution** for the reference problem (with $\theta = 0$) if for all $R > ||u||_\infty$, there exists $\eta > 0$ such that $\overline{u}$ is solution to the localized problem

$$\underset{u \in L([0,T],B_R)}{\text{Min}} \phi(y_T[u,0]), \quad \text{s.t. } \Phi(y_T[u],0) \in K, \ ||y[u,0] - \overline{y}||_\infty \leq \eta, \qquad (\mathcal{P})$$

where B_R is the ball of radius R and $\overline{y} = y[\overline{u},0]$. Now, we fix $\overline{u}$, R, and η.

1.2 Relaxation

Let X be a closed subset of $\mathbb{R}^m$, we denote by $\mathcal{P}(X)$ the set of probabilities on X. The space of **Young measures** $\mathcal{M}^Y(X)$ is the set of measurable mapping from $[0,T]$ to $\mathcal{P}(X)$ [5]. We equip this space with:

- ▷ the weak-$*$ topology,
- ▷ the narrow topology,
- ▷ the usual L^p-distance of transportation theory, denoted by d_p.

For example, a sequence of controls oscillating increasingly fast between to values a and b converges weakly-$*$ to $\mu_t = (\delta_a + \delta_b)/2$. We denote by $\overline{\mu}$ the Young measure such that $\overline{\mu}_t = \delta_{\overline{u}_t}$. For μ in $\mathcal{M}^Y(B_R)$, we denote by $y[\mu,\theta]$ the solution to

$$\begin{cases} \dot{y}_t = \int_{B_R} f(u, y_t, \theta)\, \mathrm{d}\mu_t(u), & \text{for a. a. } t \text{ in } [0,T], \\ y_0 = y^0(\theta). \end{cases}$$

We consider the family of relaxed problems with value function $V(\theta) =$

$$\underset{\mu \in \mathcal{M}^Y(B_R)}{\text{Min}} \phi(y_T[\mu,\theta]), \quad \text{s.t. } \Phi(y_T[\mu],\theta) \in K, \ ||y[\mu,\theta] - \overline{y}||_\infty \leq \eta. \qquad (\mathcal{P}^Y)$$

1.3 Pontryagin linearization

For a given μ in $M^Y(B_R)$, we define the **Pontryagin linearization** $\xi[\mu]$ as follows:

$$\begin{cases} \dot{\xi}_t[\mu] = f_y(\overline{u}_t, \overline{y}_t)\xi_t[\mu] + \int_{U_R} f(u, \overline{y}_t) - f(\overline{u}_t, \overline{y}_t)\, \mathrm{d}\mu_t(u), \\ \xi_0[\mu] = 0. \end{cases}$$

We denote by ξ^θ the solution to

$$\begin{cases} \dot{\xi}_t^\theta = f_y[t]\xi_t^\theta + f_\theta[t], & \text{for a. a. } t \text{ in } [0,T], \\ \xi^\theta = y_\theta^0(0). \end{cases}$$

The following estimate holds

$$||y[\mu,\theta] - (\overline{y} + \xi[\mu] + \theta\xi^\theta)||_\infty = O(d_1(\mu,\overline{\mu})^2 + \theta^2).$$

1.4 Qualification

We set $\mathcal{R}_T = \{\xi_T[\mu], \mu \in \mathcal{M}^Y(B_R)\}$ and we denote by $\mathcal{C}(\mathcal{R}_T)$ the smallest closed cone containing $\mathcal{R}_T$. We assume that the following **qualification condition** holds: there exists $\varepsilon > 0$ such that

$$B_\varepsilon \subset \Phi(\overline{y}_T, 0) + \Phi_{y_T}(\overline{y}_T, 0)\mathcal{C}(\mathcal{R}_T) - K.$$

Theorem (Metric regularity). *There exist $\tilde{\theta} > 0$, $\delta > 0$ and $C > 0$ such that for all θ in $[0,\tilde{\theta}]$ and for all μ in $\mathcal{M}^Y(B_R)$ satisfying $d_1(\mu,\overline{\mu}) \leq \delta$, there exists a control μ' such that*

$$\Phi(y_T[\mu',\theta],\theta) \in K \quad \textit{and} \quad d_1(\mu,\mu') \leq C \cdot \mathit{dist}(\Phi(y_T[\mu,\theta],\theta), K).$$

1.5 Motivations for the relaxation

It can be checked that

- ▷ $\mathcal{M}^Y(B_R)$ is weakly-$*$ compact
- ▷ $L([0,T],B_R)$ is weakly-$*$ dense in $\mathcal{M}^Y(B_R)$
- ▷ $y[\mu,\theta]$ is weakly-$*$ continuous.

Therefore,

- ▷ the relaxed problems posseses optimal solutions
- ▷ if $\overline{\mu}$ is the only control μ such that $y[\mu] = \overline{y}$, then problems $(\mathcal{P})$ and $(\mathcal{P}^Y)$ have the same value.

2 Methodology of sensitivity analysis

Following [3], we describe the methodology used in an abstract framework:

$$V(\theta) = \underset{x \in H}{\text{Min}}\ f(x,\theta) \quad \text{s.t. } g(x,\theta) \in K, \qquad (P_\theta)$$

where H is a Hilbert space and K stands for inequalities and equalities. The Lagrangian is

$$L(x,\lambda,\theta) = f(x,\theta) + \langle \lambda, g(x,\theta) \rangle.$$

Let $\overline{x}$ be an optimal solution to (P_0) and Λ be the set of Lagrange multipliers associated.

2.1 First-order upper estimate

Let d in H be such that $g'(\overline{x},0)(d,1) \in T_K(g(\overline{x},0))$. With a regularity theorem, we construct a feasible sequence $x_\theta = \overline{x} + \theta d + o(\theta)$. Therefore, the linear problem

$$\underset{d \in H}{\text{Min}} f'(\overline{x},0)(d,1) \quad \text{s.t. } g'(\overline{x},0)(d,1) \in T_K(g(\overline{x},0)), \qquad (LP)$$

provides the upper estimate $V(x) \leq V(0) + \theta\, \mathrm{Val}(LP) + o(\theta)$. Moreover, the dual of (LP) is

$$\underset{\lambda \in \Lambda}{\text{Max}} L_\theta(\overline{x},\lambda,0). \qquad (LD)$$

2.2 Second-order upper estimate

Let d be a solution to (LP). We define

$$\begin{cases} \underset{h \in H}{\text{Min}} & f_x(\overline{x},0)h + \frac{1}{2}f''(\overline{x},0)(d,1)^2 \\ \text{s.t.} & g_x(\overline{x},0)h + \frac{1}{2}g''(\overline{x},0)(d,1)^2 \in T_K^2(g(\overline{x},0), g'(\overline{x},0)d). \end{cases} \qquad (QP)$$

The dual of this problem is

$$\underset{\lambda \in S(LD)}{\text{Max}} L_{(x,\theta)^2}(\overline{x},\lambda,0)(d,1)^2. \qquad (QD)$$

Finally, we obtain the upper expansion of $V(\theta)$

$$V(0) + \theta\Big(\mathrm{Val}(LP)\Big) + \theta^2\Big(\underset{d \in S(LP_\theta)}{\text{Min}}\ \underset{\lambda \in S(LD)}{\text{Max}} L_{(x,\theta)^2}(\overline{x},\lambda,0)(d,1)^2\Big) + o(\theta^2).$$

2.3 Rate of convergence of solutions

We consider a **strong sufficient second-order condition:** there exists $\alpha > 0$ such that for all h in the critical cone,

$$\sup_{\lambda \in S(LD_\theta)} L_{xx}(\overline{x},\lambda,0)h^2 \geq \alpha|h|^2.$$

If this condition is satisfied, then the solutions x^θ to (P_θ) are such that

$$|x^\theta - \overline{x}| = O(\theta).$$

Moreover, the sequence $(x^\theta - \overline{x})/\theta$ has all its limit points in $S(LP)$.

2.4 Second-order lower estimate

A second order expansion follows from a Taylor expansion: for all λ in $S(LD)$,

$$\begin{aligned} V(\theta) - V(0) &= f(x^\theta) - f(\overline{x}) \\ &\geq L(x^\theta,\lambda,\theta) - L(\overline{x},\lambda,0) \\ &= \theta L_\theta(\overline{x},\lambda,0) + \frac{\theta^2}{2}\Big(L_{(x,\theta)^2}(\overline{x},\lambda,0)\big(\frac{x^\theta - \overline{x}}{\theta},1\big)^2\Big) + o(\theta^2) \\ &\geq \theta L_\theta(\overline{x},\lambda,0) + \frac{\theta^2}{2}\Big(\underset{d \in S(LP)}{\text{Min}} L_{(x,\theta)^2}(\overline{x},\lambda,0)(d,1)^2\Big) + o(\theta^2). \end{aligned}$$

3 Upper estimates

3.1 First-order upper estimate

For the optimal control problems, we consider perturbations of this form:

$$\mu^\theta = (1-\theta)\overline{\mu} + \theta\mu,$$

where the addition is **the addition of measures**. We have

$$y[\mu^\theta,\theta] = \overline{y} + \theta(\xi[\mu] + \xi^\theta) + o(\theta).$$

The equivalent of problem (LP) is now:

$$\underset{\xi \in \mathcal{C}(\mathcal{R}_T)}{\text{Min}} \phi'(\overline{y}_T,0)(\xi + \xi_T^\theta, 1) \quad \text{s.t. } \Phi'(\overline{y}_T,0)(\xi + \xi_T^\theta, 1) \in T_K(\Phi(\overline{y}_T,0)).$$

Let us define:

- ▷ the end-point Lagrangian, $\Phi[\lambda](y,\lambda,\theta) = \phi(y,\theta) + \lambda\Phi(y,\theta)$,
- ▷ the Hamiltonian, $H[p](u,y,\theta) = \langle p, f(u,y,\theta) \rangle$,
- ▷ the costate p^λ associated with λ in $N_K(\Phi(\overline{y}_T,0))$, the solution to

$$\begin{cases} \dot{p}_t = & -H_y[p_t](\overline{u}_t,\overline{y}_t) \\ p_T = & \Phi_{y_T}[\lambda](\overline{y}_T,\lambda,0). \end{cases}$$

- ▷ **Pontryagin multipliers** Λ^P, the set of λ in $N_K(\Phi(\overline{y}_T,0))$ such that for almost all t, $u \mapsto H[p_t^\lambda](u,\overline{y}_t,0)$ is minimized by $\overline{u}_t$.

The dual of problem (LP) is

$$\underset{\lambda \in \Lambda}{\text{Max}} \Big\{ p_0^\lambda y_\theta^0(0) + \int_0^T H_\theta[p_t^\lambda](\overline{u}_t,\overline{y}_t,0)\,\mathrm{d}t + \Phi_\theta(\overline{y}_T,0) \Big\}. \qquad (LD)$$

3.2 Second-order upper estimate

Unfortunately, problem (LP) does not have necessarily solutions. We consider the linearization associated with the perturbation $\overline{u} + \theta v$. We denote by $z[v]$ the solution of

$$\begin{cases} \dot{z}_t[v] &= f_{u,y}(\overline{u}_t,\overline{y}_t,0)(v_t, z_t[v]), \\ z_0[v] &= y_0^\theta(0), \end{cases}$$

and we set $z^1[v] = z[v] + \xi^\theta$. This definition extends to ν in $\mathcal{M}_2^Y$, the space of Young measures with a finite L^2-norm. The standard linearized problem is

$$\underset{\nu \in \mathcal{M}_2^Y}{\text{Min}} \phi'(\overline{y}_T,0)(z_T^1[v],1) \text{ s.t. } \Phi'(\overline{y}_T,0)(z_T^1[v],1) \in T_K(\Phi(\overline{y}_T,0)). \qquad (SLP)$$

Now, $\lambda \in N_K(\Phi(\overline{y}_T,0))$ is said to be a **Lagrange multiplier** if for almost all t,

$$H_u[p_t^\lambda](\overline{u}_t,\overline{y}_t) = 0.$$

The set of Lagrange multipliers is denoted by Λ^L. Note that $\Lambda^P \subset \Lambda^L$. The dual of (SLP) is:

$$\underset{\lambda \in \Lambda^L}{\text{Max}} L_\theta(\overline{u},\overline{y},\lambda,0). \qquad (SLD)$$

Now we assume that:

$$\mathrm{Val}(SLP) = \mathrm{Val}(LP).$$

Consider a solution ν to (SLP). Considering a perbutation of the form

$$\mu^\theta = (1-\theta^2)(\overline{u} + \theta v) + \theta^2\mu,$$

we obtain a second-order problem whose dual is the following:

$$\underset{\lambda \in S(LD)}{\text{Max}} \Omega^\theta[\lambda](v), \qquad (QD(\nu))$$

where Ω^θ is defined by

$$\begin{aligned} \Omega^\theta[\lambda](\nu) = {} & p_0^\lambda y_{\theta\theta}^0(0) + \Phi''[\lambda](\overline{y}_T,0)(z_T^1[\nu],1) \\ & + \int_0^T \int_{\mathbb{R}^m} H''[p_t^\lambda](\overline{u}_t,\overline{y}_t,0)(u, z_t^1[\nu],1)^2\,\mathrm{d}\mu_t(u)\,\mathrm{d}t. \end{aligned}$$

Theorem. *The following estimate holds:*

$$V(\theta) \leq V(0) + \theta\, \mathit{Val}(LP_\theta) + \frac{\theta^2}{2}\Big(\underset{\nu \in S(SLP)}{\mathit{Min}}\ \underset{\lambda \in S(LD)}{\mathit{Max}} \Omega[\lambda](\nu)\Big) + o(\theta^2).$$

4 Lower estimate

The critical cone C is the set of ν in $M_Y^2(\mathbb{R}^m)$ such that

$$\begin{aligned} & \phi_{y_T}(\overline{y}_T,0)z_T[\nu] \leq 0, \\ & \Phi_{y_T}(\overline{y}_T,0)z_T[\nu] \in T_K(\Phi(\overline{y}_T,0)). \end{aligned}$$

We also define the quadratic form $\Omega[\lambda](\nu)$ by

$$\begin{aligned} \Omega[\lambda](\nu) = {} & \Phi_{(y_T)^2}[\lambda](\overline{y}_T,0)(z_T[\nu]) \\ & + \int_0^T \int_{\mathbb{R}^m} H_{(u,y)^2}[p_t^\lambda](\overline{u}_t,\overline{y}_t,0)(u, z_t[\nu])^2\,\mathrm{d}\mu_t(u)\,\mathrm{d}t. \end{aligned}$$

The **strong second-order sufficient condition** is: $\exists \alpha > 0$ such that

1. For all μ in $\mathcal{M}^Y(B_R)$,

$$\sup_{\lambda \in S(LD_\theta)} \int_0^T H[p_t^\lambda](u,\overline{y}_t,0) - H[p_t^\lambda](\overline{u}_t,\overline{y}_t,0)\,\mathrm{d}\mu_t(u)\,\mathrm{d}t \geq \alpha d_2(\overline{\mu},\mu).$$

2. For all ν in C, $\sup_{\lambda \in S(LD_\theta)} \Omega[\lambda](\nu) \geq \alpha||\nu||_2^2$.

We consider solutions μ^θ to the perturbed problem.

Theorem. *For all sequence $\theta_k \downarrow 0$, the sequence $\frac{\mu^{\theta_k} \ominus \overline{u}}{\theta_k}$ has a limit point for the narrow topology in $S(SLP)$. Moreover,*

$$V(\theta) \geq V(0) + \theta\, \mathit{Val}(LP_\theta) + \frac{\theta^2}{2}\Big(\underset{\nu \in S(SLP)}{\mathit{Min}}\ \underset{\lambda \in S(LD_\theta)}{\mathit{Max}} \Omega[\lambda](\nu)\Big) + o(\theta^2).$$

Sketch of the proof. Following [1], we decompose a solution μ^k into two controls:

- ▷ $\mu^{A,k}$, accounting for the small variations in $L^\infty-$norm of the control,
- ▷ $\mu^{B,k}$, accounting for the large variations in $L^\infty-$norm of the control, but on a small subset of $[0,T] \times B_R$.

For all λ in $S(LD)$, we have

$$\phi(y_T[\mu^\theta,\theta],\theta) - \phi(\overline{y}_T,0) \geq \Phi[\lambda](y_T[\mu^\theta,\theta],\theta) - \Phi[\lambda](\overline{y}_T,0).$$

Then, we expand the r.h.s. and we neglect the part due to $\mu^{B,k}$.

References

[1] J.F. Bonnans & N. Osmolovskiĭ. Second-order analysis of optimal control problems with control and initial-final state constraints. *Journal of Convex Analysis*, 2010.

[2] J.F. Bonnans, L. Pfeiffer & O.S. Serea. Sensitivity analysis for relaxed optimal control problems with final-state constraints. *Submitted, available as the Inria Research Report 7977, May 2012.*

[3] J.F. Bonnans & A. Shapiro. *Perturbation analysis of optimization problems.* Springer-Verlag, New York, 2000.

[4] M. Valadier. Young measures. *Methods of nonconvex analysis*, 1990.

[5] L.C. Young. *Lectures on the calculus of variations and optimal control theory.* Foreword by Wendell H. Fleming. W. B. Saunders Co., Philadelphia, 1969.