Second Order Variational Analysis in Optimal Control

Hélène Frankowska

CNRS and Sorbonne Université (ex-UPMC, ex Paris 6)

19th French-German-Swiss conference on Optimization

September 17-20, 2019, Nice

Outline

The aim of this talk is to discuss the role of tangent vectors in optimality conditions.

Second-Order Tangents (J.-P. Aubin, HF & 1990)

X- a normed vector space.

Adjacent tangent to $K \subset X$ at $x \in K$ is

$$T_K(x) := \left\{ u \in X \mid \lim_{\varepsilon \to 0+} \frac{\mathrm{dist}(x + \varepsilon u, K)}{\varepsilon} = 0 \right\}$$

Second-order tangent to K at x relative to $u \in X$

$$T_K^{(2)}(x; u) := \left\{ v \in X \mid \lim_{\varepsilon \to 0+} \frac{\mathrm{dist}(x + \varepsilon u + \varepsilon^2 v, K)}{\varepsilon^2} = 0 \right\}$$

Clarke tangent cone to $K \subset X$ at $x \in K$ is

$$C_K(x) := \{ u \in X \mid \lim_{y \to_K x,\, \varepsilon \to 0+} \frac{\mathrm{dist}(y + \varepsilon u, K)}{\varepsilon} = 0 \}$$

$$T_K^{(2)}(x; u) + C_K(x) = T_K^{(2)}(x; u) \qquad N_K(x) := C_K(x)^-$$

Example: Second-Order Tangents

Let $x \in M = \cap_{i=0}^{r} M_i$, where for $h \in C^2(\mathbb{R}^n; \mathbb{R}^s)$, $g_i \in C^2(\mathbb{R}^n; \mathbb{R})$

$$M_0 = \{y \ : \ h(y) = 0\}, \quad M_i = \{y \ : \ g_i(y) \leq 0\} \ \forall \ i = 1, ..., r$$

and $h'(x)$ is surjective. Assume that there exists u_0 such that

$$h'(x)u_0 = 0, \quad \langle \nabla g_i(x), u_0 \rangle < 0 \quad \forall \ i \in I_a(x) = \{i \mid g_i(x) = 0\}$$

(the Mangasarian-Fromowitz condition). Then

$$T_M(x) = \{u \in \mathbb{R}^n \ : \ h'(x)u = 0, \ \langle \nabla g_i(x), u \rangle \leq 0 \ \ \forall \ i \in I_a(x)\}$$

Moreover $\forall \ u \in T_M(x)$, a vector $v \in T_M^{(2)}(x; u)$ if and only if
$\forall \ 1 \leq j \leq s, \quad \forall \ i \in I = \{i \in I_a(x) \mid \langle \nabla g_i(x), u \rangle = 0\}$

$$\langle \nabla h_j(x), v \rangle + \frac{1}{2} \langle h_j''(x)u, u \rangle = 0, \ \langle \nabla g_i(x), v \rangle + \frac{1}{2} \langle g_i''(x)u, u \rangle \leq 0$$

Standard Minimization Problem

$\phi : X \to \mathbb{R}$, $\phi \in C^2$, $K \subset X$.

Let $\bar{x} \in K$ be a local minimizer of the problem

$$\min_{x \in K} \phi(x)$$

What can be said when $\bar{x} \in \partial K$? (**boundary** of K)

Primal first-order necessary condition (generalized Fermat rule) :

$$\langle \phi'(\bar{x}), u \rangle \geq 0 \quad \forall\, u \in T_K(\bar{x})$$

Define the cone of **critical directions** at $\bar{x}$

$$C(\bar{x}) = \{u \in T_K(\bar{x}) \;|\; \langle \phi'(\bar{x}), u \rangle = 0\}$$

Primal second-order necessary condition:

$$\langle \phi'(\bar{x}), v \rangle + \frac{1}{2}\phi''(\bar{x})(u, u) \geq 0 \quad \forall\, u \in C(\bar{x}),\ \forall\, v \in T_K^{(2)}(\bar{x}; u)$$

Set-Valued Inverse Mapping Theorem (HF 1989)

Let (Z, d) be a complet metric space, Y be a Banach space with the norm Gateaux differentiable (away from 0), $G : Z \to Y$ be continuous. Define

$$G^{(1)}(z) := \left\{ v \in Y \mid \liminf_{\varepsilon \to 0+} \frac{\mathrm{dist}\,(G(z) + \varepsilon v, G(B_\varepsilon(z)))}{\varepsilon} = 0 \right\}$$

If $\exists\, \rho > 0$ such that $\forall\, z$ near $\bar{z}$,

$$\rho B \subset \overline{co}\, G^{(1)}(z)$$

then G^{-1} is **pseudo-Lipschitz:**

$$\mathrm{dist}(z, G^{-1}(y)) \leq \frac{1}{\rho}|G(z) - y| \qquad \forall\, (z, y) \text{ near } (\bar{z}, G(\bar{z}))$$

Similar result is valid also when G is a set-valued map.

Control System

$$\begin{cases} x'(t) = f(x(t),u(t)), \quad u(t) \in U \quad \text{a.e. in } [0,\tau] \\ x(0) = x_0 \end{cases}$$

$$f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \quad x_0 \in \mathbb{R}^n, \quad U \subset \mathbb{R}^m$$

Controls are Lebesgue measurable functions $u(\cdot) : [0,\tau] \to U$

A **trajectory** $x(\cdot)$ of control system is an absolutely continuous function satisfying $x(0) = x_0$ and for some control $u(\cdot)$

$$x'(t) = f(x(t),u(t)) \text{ almost everywhere in } [0,\tau]$$

Denote by $\mathcal{S} \subset C([0,\tau];\mathbb{R}^n)$ the set of all such trajectories. We assume that $f \in C^2$, U is compact and $\exists\, k > 0$ such that $\max_{u\in U} |f(x,u)| \le k(|x|+1)$.

System with the End-Point Equality Constraints

$\mathcal{U} := L^1([0,\tau];U)$ - is a metric space.
Let $h = (h_1, ..., h_k) : \mathbb{R}^n \to \mathbb{R}^k,\ h \in C^2,\ x_0 \in \mathbb{R}^n$

$$(*) \qquad x'(t) = f(x(t), u(t)),\ x(0) = x_0,\ u(t) \in U,\ h(x(\tau)) = 0$$

Let $\bar{u} \in \mathcal{U}$ and the corresponding $\bar{x}(\cdot)$ satisfies $h(\bar{x}(\tau)) = 0$.
Set $[t] := (\bar{x}(t), \bar{u}(t))$ and consider the **variational system**

$$\begin{cases} y'(t) = f_x[t]y(t) + v(t), \quad v(t) \in f(\bar{x}(t), U) - f[t] \\ y(0) = 0 \end{cases} \qquad (1)$$

Its **reachable set** at time τ is

$$R^L = \{y(\tau) :\ y(\cdot) \ \text{is a trajectory of (1)}\} \subset \overline{co}\, G^{(1)}(\bar{u})$$

with $G(u) = x(\tau; u)$, $\forall\, u \in \mathcal{U}$ (trajectory corresponding to u).

A Metric Regularity Result

Theorem

If $0 \in \operatorname{Int} h'(\bar{x}(\tau))(R^L)$, *then there exist* $\varepsilon > 0$, $C > 0$ *such that for every trajectory-control pair* (x, u) *with* $\|u - \bar{u}\|_{L^1} < \varepsilon$ *we can find a trajectory-control pair* $(\tilde{x}, \tilde{u})$ *satisfying*

$$h(\tilde{x}(\tau)) = 0, \quad \|\tilde{u} - u\|_{L^1} \leq C|h(x(\tau))|$$

Consider the **linearized** control system at $(\bar{x}, \bar{u})$:

$$(L) \qquad \begin{cases} y'(t) = f_x[t]y(t) + f_u[t]u(t), & u(t) \in T_U(\bar{u}(t)) \\ y(0) = 0 \end{cases}$$

Let (y, u) be a trajectory-control pair of (L). Then $y \in T_S(\bar{x})$.

Second-Order Linearization

Assume

$$(B) \quad \begin{cases} \exists\, \delta_0 > 0, \quad \exists\, c \in L^1([0,\tau];\mathbb{R}_+) \quad \text{such that } \forall\, \delta \in [0,\delta_0] \\ \mathrm{dist}_U(\bar{u}(t) + \delta u(t)) \le c(t)\delta^2 \quad \text{for a.e. } \ t \in [0,\tau]. \end{cases}$$

Set $\xi(t) = (y(t), u(t))$ and consider the second-order linearization

$$(L2) \qquad \begin{cases} w'(t) = f_x[t]w(t) + f_u[t]v(t) + \frac{1}{2}\,\xi(t)^* f''[t]\xi(t) \\ v(t) \in T_U^{(2)}(\bar{u}(t); u(t)) \ \text{a.e.} \\ w(0) = 0 \end{cases}$$

If (w, v) is a trajectory-control pair of (L2), then $w \in T_{\mathcal{S}}^{(2)}(\bar{x}; y)$.
Denote by $\mathcal{R}^2$ its **reachable set** at time τ.

Second-Order Tangents to Trajectories of $(^*)$

Let (y, u) be a trajectory-control pair of (L) satisfying (B).

Corollary

Assume $h'(\bar{x}(\tau))y(\tau) = 0$ and let w be a trajectory of (L2) such that

$$h_j{}'(\bar{x}(\tau))w(\tau) + \frac{1}{2}\, y(\tau)^* h_j{}''(\bar{x}(\tau))y(\tau) = 0 \quad \forall j$$

Then w is in the second order tangent at $(\bar{x}, y)$ to the set of solutions to

$$(^*) \qquad x'(t) = f(x(t), u(t)),\; x(0) = x_0,\; u \in \mathcal{U},\; h(x(\tau)) = 0$$

Optimal Control with Mixed End-Point Constraints

$$\text{Minimize} \quad \varphi(x(\tau))$$

$$g_i(x(\tau)) \leq 0, \quad i = 1, \dots, r$$

$$h_j(x(\tau)) = 0, \quad j = 1, \dots, k$$

$$x'(t) = f(x(t), u(t)), \quad x(0) = x_0, \quad u(t) \in U \text{ a.e. } t,$$

where U is an arbitrary nonempty compact subset of $\mathbb{R}^m$, $x_0 \in \mathbb{R}^n$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $h_j : \mathbb{R}^n \to \mathbb{R}$, $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are C^2.
Any control $u \in \mathcal{U}$ such that the corresponding x satisfies end-point constraints is **admissible**.

An admissible $(\bar{x}, \bar{u})$ is a *weak local minimizer* (in L^1) if for some $\varepsilon > 0$ we have $\varphi(x(\tau)) \geq \varphi(\bar{x}(\tau))$ for any admissible (x, u) such that $\|u - \bar{u}\|_{L^1} < \varepsilon$.

First Order Necessary Conditions

The ***Hamiltonian*** and the ***terminal Lagrange function*** are defined by $\mathcal{H}(x,u,p) = \langle p, f(x,u)\rangle$ and

$$\ell(x,\lambda,\mu) = \lambda_0\varphi(x) + \sum_{i=1}^{r}\lambda_i g_i(x) + \sum_{j=1}^{k}\mu_j h_j(x),$$

Theorem If $(\bar{x},\bar{u})$ is a weak local minimizer, then $\exists\, \mu = (\mu_1, ..., \mu_k) \in \mathbb{R}^k$ and $\lambda = (\lambda_0,\ldots,\lambda_r) \in \mathbb{R}_+^{r+1}$, not all $= 0$, satisfying $\lambda_i g_i(\bar{x}(\tau)) = 0$ such that for the solution p of

$$-p'(t) = \mathcal{H}_x(\bar{x}(t),\bar{u}(t),p(t)) \quad p(\tau) = \ell_x(\bar{x}(\tau),\lambda,\mu)$$

we have

$$-\mathcal{H}_u(\bar{x}(t),\bar{u}(t),p(t)) \in N_U(\bar{u}(t)) \quad \text{a.e. } t$$

Denote by $\Lambda(\bar{x},\bar{u})$ the set of all such $(\lambda,\mu,p) \neq 0$.

Strategy of Derivation of Second-Order Conditions

Take a critical solution (y, u) of (L) satisfying (B), i.e. $\langle \varphi'(\bar{x}(\tau)), y(\tau) \rangle = 0$ and y is tangent to constraints.

1. Consider the **reachable set** $\mathcal{R}^2$ of (L2) (it is convex) and **second-order tangent** $\mathcal{Q}$ to the set defined by the **end-point equality constraints**.
2. Let $\mathcal{Q}^2$ be the interior of the **second-order tangent** to the set defined by the **end-point inequality constraints**.
3. Show that $\mathcal{R}^2 \cap \mathcal{Q} \cap \mathcal{Q}^2 \cap \mathcal{L}^2 = \emptyset$, where

$$\mathcal{L}^2 := \{z \mid \langle \varphi'(\bar{x}(\tau)), z \rangle + \frac{1}{2}\, y(\tau)^* \varphi''(\bar{x}(\tau)) y(\tau) < 0\}$$

4. Apply a separation theorem.

Same strategy works for optimal control problems involving PDEs and **stochastic control** systems.

Critical Cone

$\mathcal{C}(\bar{x},\bar{u})$ is the set of all $\xi = (y,u)$ solving (L), satisfying (B) with

$$\varphi'(\bar{x}(\tau))y(\tau) = 0,\ h'(\bar{x}(\tau))y(\tau) = 0,\ g_i'(\bar{x}(\tau))y(\tau) \le 0 \quad \forall\, i \in I_{active}$$

Let $I := \{i \in I_{active} :\ \langle g_i'(\bar{x}(\tau)), y(\tau)\rangle = 0\}$.
For $(\lambda,\mu,p) \in \Lambda(\bar{x},\bar{u})$ and $t \in [0,\tau]$, define $[t] = (\bar{x}(t),\bar{u}(t),p(t))$,

$$\Upsilon(u(t),p(t)) := \inf\left\{\mathcal{H}_u[t]v \ :\ v \in T_U^{(2)}(\bar{u}(t);u(t))\right\}$$

$$\Omega(\xi,\lambda,\mu,p) := y(\tau)^*\ell_{xx}(\bar{x}(\tau),\lambda,\mu)y(\tau) + \int_0^\tau \xi(t)^*\mathcal{H}''[t]\xi(t)dt,$$

where $\mathcal{H}''[t]$ is the Hessian of $\mathcal{H}(\cdot,\cdot,p(t))$ at $(\bar{x}(t),\bar{u}(t))$.

Second-Order Necessary Optimality Conditions

Theorem

Let $(\bar{x}, \bar{u})$ *be a weak local minimizer and* $\xi = (y, u) \in \mathcal{C}(\bar{x}, \bar{u})$. *If* $0 \in \mathrm{Int}\, h'(\bar{x}(\tau))(R^L)$ *and* $\exists\, v \in L^\infty([0,\tau];\mathbb{R}^m)$ *such that*

$$v(t) \in T_U^{(2)}(\bar{u}(t); u(t)) \quad \text{a.e.},$$

then for some $(\lambda, \mu, p) \in \Lambda(\bar{x}, \bar{u})$ *with* $\lambda_i = 0$ *for* $i \notin I$, *the function* $\Upsilon(u(\cdot), p(\cdot))$ *is integrable and*

$$\frac{1}{2}\Omega(\xi, \lambda, \mu, p) + \int_0^\tau \Upsilon(u(t), p(t))dt \geq 0$$

Proof: Variational Inequality and Separation Thm

For every $i \in I$ define

$$Q_i = \left\{\eta \in \mathbb{R}^n : \; g_i'(\bar{x}(\tau))\eta + \frac{1}{2}\, y(\tau)^* g_i''(\bar{x}(\tau))y(\tau) < 0\right\}$$

$$\mathcal{Q} = \{\omega \in \mathbb{R}^n : h_j'(\bar{x}(\tau))\omega + \frac{1}{2}\, y(\tau)^* h_j''(\bar{x}(\tau))y(\tau) = 0 \;\forall j\}$$

$$\mathcal{L}^2 = \left\{\eta \in \mathbb{R}^n : \; \varphi'(\bar{x}(\tau))\eta + \frac{1}{2}\, y(\tau)^* \varphi''(\bar{x}(\tau))y(\tau) < 0\right\}$$

Let $I = \{i_1, ..., i_\gamma\}$. Then for every $w(\tau) \in (\cap_{i \in I} Q_i) \cap \mathcal{Q} \cap \mathcal{R}^2$, the following variational inequality holds true:

$$\varphi'(\bar{x}(\tau))w(\tau) + \frac{\tau}{2}\, y(\tau)^* \varphi''(\bar{x}(\tau))y(\tau) \geq 0.$$

Thus zero does not belong to the convex set

$$\{(q_0 - \kappa, q_{i_1} - \kappa, ..., q_{i_\gamma} - \kappa, \theta - \kappa) \; : \; \kappa \in \mathcal{R}^2, \\ \theta \in \mathcal{Q},\; q_0 \in \mathcal{L}^2,\; q_{i_j} \in Q_{i_j}, j = 1, ..., \gamma\}.$$

Stochastic Optimal Control Problem

Stochastic optimal control problem of the Mayer type:

$$\text{Minimize } \mathbb{E}\ \varphi(x(\tau))$$

over process-control pairs of the stochastic control system

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t) \\ x(0) = x_0 \end{cases}$$

with the end-points constraints

$$x_0 \in K_0, \quad \mathbb{E}\ g^i(x(\tau)) \leq 0,\ \forall\ i = 1, \cdots, k, \tag{2}$$

where $W(\cdot)$ is a d-dimensional Wiener process, $\tau > 0$,
K_0 is a closed subset of $\mathbb{R}^n$, $u(\cdot)$ is the control variable,
$x(\cdot)$ solves the above system. $\varphi : \mathbb{R}^n \times \Omega \to \mathbb{R}$,
$b : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$, $g^i : \mathbb{R}^n \times \Omega \to \mathbb{R}$, $i = 1, \cdots, k$.
$\sigma = (\sigma^1, \cdots, \sigma^d) : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^{n\times d}$,

Framework

$(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a complete filtered probability space, $\mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ is the natural filtration generated by $W(\cdot)$ (augmented by all the P-null sets) ;

$U \subset \mathbb{R}^m$ is closed and nonempty ;

$u(\cdot)$ is a control taking values in U, while $x(\cdot)$ is the corresponding solution to the stochastic control system :

$$x(t) = x_0 + \int_0^t b(s, x(s), u(s))ds + \int_0^t \sigma(s, x(s), u(s))dW(s)$$

We impose the usual measurability/regularity/boundedness assumptions on b, σ, φ, g_i and their derivatives.

Admissible Pairs for Stochastic System

$\mathcal{B}([0,\tau])$ denotes the Borel σ-field on $[0,\tau]$,
$\mathcal{U}$ is the set of $\mathcal{B}([0,\tau]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes with values in U such that
$\|u\|_2 := \left(\mathbb{E}\int_0^\tau |u(t)|^2 dt\right)^{1/2} < \infty$.

Denote by x the stochastic process corresponding to a control $u \in \mathcal{U}$ and initial condition x_0. (x,u) is called an admissible pair if in addition $x_0 \in K_0$, $\mathbb{E}g_i(x(\tau)) \leq 0, i = 1, ..., k$. Set

$$J(x,u) := \mathbb{E}\ \varphi(x(\tau))$$

An admissible pair $(\bar{x},\bar{u})$ is a weak local minimizer if $\exists\ \delta > 0$ such that for any admissible (x,u) with $\|u-\bar{u}\|_2 + |x(0) - \bar{x}(0)| < \delta$

$$J(x,u) \geq J(\bar{x},\bar{u})$$

Linearized System

Let $(\bar{x}, \bar{u})$ be a weak local minimizer.
For $\phi = b_x, b_u, \sigma_x, \sigma_u$, denote $\phi[t] := \phi(t, \bar{x}(t), \bar{u}(t))$.
Consider the **linearized stochastic control system**:

$$(L) \quad \begin{cases} dy(t) &= \left(b_x[t]y(t) + b_u[t]v(t)\right)dt \\ & \quad + \sum_{j=1}^{d} \left(\sigma_x^j[t]y(t) + \sigma_u^j[t]v(t)\right)dW^j(t) \\ y(0) &= \nu_0 \end{cases}$$

$v \in L^2_{\mathbb{F}}(\Omega; L^2(0, \tau; \mathbb{R}^m))$ satisfies $v \in T_{\mathcal{U}}(\bar{u})$, $\nu_0 \in T_{K_0}(\bar{x}(0))$.

$L^2_{\mathbb{F}}(\Omega; C([0, \tau]; \mathbb{R}^n))$ - $\mathcal{B}([0, \tau]) \otimes \mathcal{F}$-measurable, $\mathbb{F}$-adapted continuous processes $\zeta : [0, \tau] \times \Omega \to \mathbb{R}^n$ such that

$$\|\zeta\|_{\infty,2} := \left[\mathbb{E}\left(\sup_{t\in[0,\tau]} |\zeta(t,\omega)|^2\right)\right]^{\frac{1}{2}} < +\infty$$

Variational Equation

Consider $\nu_0^\varepsilon \in \mathbb{R}^n$ and $v_\varepsilon \in L^2_{\mathbb{F}}(\Omega; L^2(0,\tau; \mathbb{R}^m))$ such that

$$\bar{x}(0) + \varepsilon \nu_0^\varepsilon \in K_0, \qquad \bar{u} + \varepsilon v_\varepsilon \in \mathcal{U}$$

$$\nu_0^\varepsilon \to \nu_0, \quad v_\varepsilon \to v \quad \text{as} \quad \varepsilon \to 0^+$$

Let x^ε be the state corresponding to the control $u^\varepsilon := \bar{u} + \varepsilon v_\varepsilon$ and the initial datum $x_0^\varepsilon := \bar{x}(0) + \varepsilon \nu_0^\varepsilon$.

Lemma

$$\left\| \frac{x^\varepsilon - \bar{x}}{\varepsilon} - y \right\|_{\infty,2} \to 0 \quad as \quad \varepsilon \to 0+$$

Reachable Set

We restrict our attention to controls $v \in L^2_{\mathbb{F}}(\Omega; L^2(0,\tau;\mathbb{R}^m))$ satisfying $v(t) \in C_U(\bar{u}(t))$ a.s., for a.e. $t \in [0,\tau]$.
Define the **reachable set of the linearized system** by:

$$\mathcal{R}^{(1)} := \{y(\tau) \in L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^n) \mid y \text{ solves (L)}, \; v \text{ as above}, \; \nu_0 \in C_{K_0}(\bar{x}(0)\}$$

Consider a **linearization of end-point constraints**

$$\mathcal{Q}^{(1)} := \{z \in L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^n) \mid \mathbb{E}\left\langle g^i_x(\bar{x}(\tau)), z\right\rangle < 0 \;\; \forall\, i \in I_a)\}$$

where

$$I_a := \{i \mid \mathbb{E}\, g^i(\bar{x}(\tau)) = 0\}$$

Nonintersection of Convex Sets

Define $\mathcal{L}^{(1)} := \{z \in L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^n) \mid \mathbb{E}\langle \varphi_x(\bar{x}(\tau)), z\rangle < 0\}$
Show using the variational equation that $\mathcal{R}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{L}^{(1)} = \emptyset$.
Apply separation theorems to get first order necessary optimality conditions.

First order adjoint system

$$\begin{cases} dP_1(t) = -\big(b_x[t]^\star P_1(t) + \sum_{j=1}^{d} \sigma_x^j[t]^\star Q_1^j(t)\big)dt + \sum_{j=1}^{d} Q_1^j(t)dW^j(t) \\ P_1(\tau) = \xi \end{cases}$$

Under our assumptions, by El Karoui, Peng, Quenez (1997), it admits a unique strong solution
$(P_1(\cdot), Q_1(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C([0,\tau];\mathbb{R}^n)) \times \ L^2_{\mathbb{F}}(\Omega; L^2(0,\tau;\ \mathbb{R}^{n\times d}))$

Hamiltonian

Define the **Hamiltonian** by :

$$H(t,x,u,p,q,\omega) := \langle p, b(t,x,u,\omega)\rangle + \sum_{j=1}^{d}\langle q^j, \sigma^j(t,x,u,\omega)\rangle$$

for $(t,x,u,p,q,\omega) \in [0,\tau]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times\Omega$. Let

$$H[t] := H(t,\bar{x}(t),\bar{u}(t),P_1(t),Q_1(t)), \quad t\in[0,\tau],$$

$H_u[t]$, $H_{xx}[t]$, $H_{xu}[t]$ and $H_{uu}[t]$ are defined in a similar way.

The separation theorem and the Ito formulae imply

Theorem 1: First Order Pointwise Necessary Optimality Condition

(i) If $I_a = \emptyset$ or if $\mathcal{Q}^{(1)} \neq \emptyset$, then $\exists\ \lambda_0 \in \{0,1\}$, $\lambda_i \geq 0$ for $i \in I_a$ and a solution (P_1, Q_1) to the first order adjoint equation with $\lambda_0 + \mathbb{E}\,|P_1(\tau)| \neq 0$ such that

$$-H_u[t] \in N_U(\bar{u}(t)) \quad \text{a.s.,} \quad \text{for a.e. } t \in [0,\tau]$$

$$-P_1(0) \in N_{K_0}(\bar{x}_0),\ P_1(\tau) = \lambda_0 \varphi_x(\bar{x}(\tau)) + \sum_{i \in I_a} \lambda_i g_x^i(\bar{x}(\tau)).$$

(ii) If $I_a \neq \emptyset$ but $\mathcal{Q}^{(1)} = \emptyset$, then $\exists\ \lambda_i \geq 0$ for $i \in I_a$ such that $\sum_{i \in I_a} \lambda_i > 0$ and $\sum_{i \in I_a} \lambda_i g_x^i(\bar{x}(\tau)) = 0$. In particular, the relations from **(i)** hold true with $\lambda_0 = 0$ and $(P_1, Q_1) \equiv 0$.

Furthermore, $\lambda_0 = 1$ if $I_a = \emptyset$ or if $I_a \neq \emptyset$ and $\mathcal{R}^{(1)} \cap \mathcal{Q}^{(1)} \neq \emptyset$.

Needle variations versus variational approach

When b, σ, f and g are differentiable up to the second order with respect to x, then a stochastic maximum principle (which uses a second order adjoint process) was proved by Peng, 1990 for global minimizers using the classical spike (needle) variations of Boltyanski.

This principle implies the above result. However we have imposed weaker regularity requirements and do not need to consider second order adjoint system. Further, our result concerns weak minimizers.

Critical Variations of Controls

Assume in addition that $\bar{u} \in L^4_{\mathbb{F}}(\Omega; L^4(0,\tau;\mathbb{R}^m))$.
Let $v \in L^4_{\mathbb{F}}(\Omega; L^4(0,\tau;\mathbb{R}^m))$ be such that
$\exists\, \eta(\cdot) \in L^4_{\mathbb{F}}(\Omega; L^4(0,\tau;\mathbb{R}_+))$ and $\varepsilon_0 > 0$ satisfying the following **distance estimates** : $\forall\, \varepsilon \in [0,\varepsilon_0]$

$$\mathrm{dist}_U\left(\bar{u}(t,\omega) + \varepsilon v(t,\omega)\right) \leq \varepsilon^2 \eta(t,\omega), \text{ a.e. } (t,\omega) \in [0,\tau] \times \Omega$$

Let $\nu_0 \in T_{K_0}(\bar{x}(0))$ and suppose that the corresponding solution y of the linearized system is critical :

$$\mathbb{E}\left\langle g^i_x(\bar{x}(\tau)), y(\tau)\right\rangle \leq 0,\ \forall\, i \in I_a, \quad \mathbb{E}\left\langle \varphi_x(\bar{x}(\tau)), y(\tau)\right\rangle = 0$$

Now the usual regularity/boundedness assumptions are imposed also on the second derivatives of b, σ, φ, g_i.

Admissible Second Order Variations

Consider convex subsets

$$\Psi(t,\omega) \subset T_U^{(2)}(\bar{u}(t,\omega); v(t,\omega)) \text{ a.e. } (t,\omega) \in [0,\tau] \times \Omega$$

and a convex subset $\mathcal{W}(\bar{x}(0), \nu_0)$ of $T_{K_0}^{(2)}(\bar{x}(0), \nu_0)$ and let

$$\mathcal{M}(\bar{u}, v) := \left\{ h(\cdot) \in L^4_{\mathbb{F}}(\Omega; L^4(0,\tau;\mathbb{R}^m)) \;\middle|\; h(t,\omega) \in \Psi(t,\omega) \quad \text{a.e.} \right\}$$

Define

$$I := \left\{ i \in I_a \;\middle|\; \mathbb{E} \left\langle g_x^i(\bar{x}(\tau)), y(\tau) \right\rangle = 0 \right\}$$

Same Strategy

1. Write the **second order linearization of control system** and prove a second order variational equation.
2. Define its **reachable set** $\mathcal{R}^2$ for admissible controls in $\mathcal{M}(\bar{u}, v)$ and initial conditions in $\mathcal{W}(\bar{x}(0), \nu_0)$.
3. Define **second order linearization of end-point constraints** :

$$\mathcal{Q}^2 := \Big\{ z \in L^2_{\mathcal{F}_\tau}(\Omega; \mathbb{R}^n) \;\big|\; \forall\, i \in I \quad \mathbb{E}\left\langle g^i_x(\bar{x}(\tau)), z \right\rangle + \frac{1}{2}\mathbb{E}\left\langle g^i_{xx}(\bar{x}(\tau))y(\tau), y(\tau) \right\rangle < 0 \Big\}.$$

4. Show that $\mathcal{R}^2 \cap \mathcal{Q}^2 \cap \mathcal{L}^2 = \emptyset$, where $\mathcal{L}^2 :=$

$$\{ z \in L^2_{\mathcal{F}_\tau}(\Omega; \mathbb{R}^n) \,|\, \mathbb{E}\left\langle \varphi_x(\bar{x}(\tau)), z \right\rangle + \frac{1}{2}\mathbb{E}\left\langle \varphi_{xx}(\bar{x}(\tau))y(\tau), y(\tau) \right\rangle < 0 \}$$

5. Apply a separation theorem.

Second Order Adjoint System

$$
\begin{cases}
dP_2(t) = -\Big(b_x[t]^\star P_2(t) + P_2(t)b_x[t] + \sum_{j=1}^{d} \sigma_x^j[t]^\star P_2(t)\sigma_x^j[t] \\
\qquad + \sum_{j=1}^{d} \sigma_x^j[t]^\star Q_2^j(t) + \sum_{j=1}^{d} Q_2^j(t)\sigma_x^j[t] + H_{xx}[t]\Big)dt \\
\qquad + \sum_{j=1}^{d} Q_2^j(t)dW^j(t) \\
P_2(\tau) = \zeta,
\end{cases}
$$

where $\zeta \in L^2_{\mathcal{F}_\tau}(\Omega; \mathbf{S}^n)$. It admits a unique strong solution $(P_2(\cdot), Q_2(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C([0,\tau]; \mathbf{S}^n)) \times \left(L^2_{\mathbb{F}}(\Omega; L^2(0,\tau; \mathbf{S}^n))\right)^d$,
where $H_{xx}[t] = H_{xx}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t))$.
cf. El Karoui, Peng, Quenez (1997).

Notation

To simplify the notation, we define

$$\mathbb{S}(t,x,u,y_1,z_1,y_2,z_2,\omega) := H_{xu}(t,x,u,y_1,z_1,\omega) + b_u(t,x,u,\omega)^\star y_2 + \sum_{j=1}^{d} \sigma_u^j(t,x,u,\omega)^\star z_2^j + \sum_{j=1}^{d} \sigma_u^j(t,x,u,\omega)^\star y_2 \sigma_x^j(t,x,u,\omega),$$

where $(t,x,u,y_1,z_1,y_2,z_2,\omega) \in [0,\tau] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times \mathbf{S}^n \times (\mathbf{S}^n)^d \times \Omega$. Write

$$\mathbb{S}[t] = \mathbb{S}(t,\bar{x}(t),\bar{u}(t),P_1(t),Q_1(t),P_2(t),Q_2(t)),$$

where $(P_1(\cdot),Q_1(\cdot))$ and $(P_2(\cdot),Q_2(\cdot))$ solve the first and second order adjoint systems, respectively.

Theorem 2: Second Order Integral Condition

If $\mathcal{M}(\bar{u}, v) \neq \emptyset$, then $\exists\ \lambda_0 \in \{0, 1\}$, $\lambda_i \geq 0$ ($\forall\ i \in I$) not vanishing simultaneously, adjoint processes (P_1, Q_1), (P_2, Q_2)

$$P_1(\tau) = -\lambda_0 \varphi_x(\bar{x}(\tau)) - \sum_{i \in I} \lambda_i g^i_x(\bar{x}(\tau)),$$

$$P_2(\tau) = -\lambda_0 \varphi_{xx}(\bar{x}(\tau)) - \sum_{i \in I} \lambda_i g^i_{xx}(\bar{x}(\tau))$$

such that (P_1, Q_1) satisfies the **first order necessary conditions** and $\forall\ \varpi_0 \in \mathcal{W}(\bar{x}(0), \nu_0)$, $\forall\ h(\cdot) \in \mathcal{M}(\bar{u}, v)$,

$$\langle P_1(0), \varpi_0 \rangle + \frac{1}{2} \langle P_2(0)\nu_0, \nu_0 \rangle + \mathbb{E} \int_0^\tau (\langle H_u[t], h(t) \rangle + \langle \mathbb{S}[t] y(t), v(t) \rangle +$$

$$\frac{1}{2} \langle H_{uu}[t] v(t), v(t) \rangle + \frac{1}{2} \sum_{j=1}^d \langle \sigma^j_u[t]^\star P_2(t) \sigma^j_u[t] v(t), v(t) \rangle) dt \leq 0$$

Furthermore, $\lambda_0 = 1$ if $I_a = \emptyset$ or if $I_a \neq \emptyset$ and $\mathcal{R}^2 \cap \mathcal{Q}^2 \neq \emptyset$.

References for this talk:

H.F., Zhang H. & Zhang X. (2018) SIAM J. Control Optim.

H.F., Zhang H. & Zhang X. (2019) Trans. Amer. Math. Soc.

H.F. & Osmolovskii N. P. (2019) Appl. Math. Optim.

H.F. (2019) Proceedings of 58th IEEE CDC, Nice

H.F. (1989) Annales de l'Institut Henri Poincaré, Analyse Non Linéaire

Example

$$n = m = 2,\ T = 1,\ U = \{(u_1, u_2) \in \mathbb{R}^2 \mid |u_1 + 1|^2 + |u_2|^2 = 1\}.$$

$$T_U^{\flat}((0,0)) = \{0\} \times \mathbb{R}, \qquad T_U^{\flat(2)}((0,0);(0,1)) \ni (-\frac{1}{2}, 0).$$

$$\begin{cases} dx_1(t) & = & (x_2(t) - \frac{1}{2})dt + dW(t), & t \in [0,1], \\ dx_2(t) & = & u_1(t)dt + |u_2(t)|^4 dW(t), & t \in [0,1], \\ x_1(0) & = & 0, \ \ x_2(0) = 0 \end{cases}$$

$$J(u) = \mathbb{E}\Big[\frac{1}{2}|x_1(1) - W(1)|^2 + \int_0^1 |u_2(t)|^4 dt\Big].$$

We show that $(u_1(t), u_2(t)) \equiv (0,0)$ is not locally optimal.
The corresponding solution of the control system is

$$(x_1(t), x_2(t)) = (W(t) - \frac{t}{2}, 0)$$

Example

The first order adjoint equation is

$$\begin{cases} dP_1^1(t) & = & Q_1^1(t)dW(t) \\ dP_1^2(t) & = & -P_1^1(t)dt + Q_1^2(t)dW(t) \\ P_1^1(1) & = & -\frac{1}{2}, \quad P_1^2(1) = 0 \end{cases}$$

Then

$$(P_1^1(t), Q_1^1(t)) = (-\frac{1}{2}, 0), \ (P_1^2(t), Q_1^2(t)) = (\frac{t-1}{2}, 0) \quad a.s.$$

$$H_u[t] = (P_1^2(t), 4Q_1^2(u_2(t))^3 - 4(u_2(t))^3) = (\frac{t-1}{2}, 0).$$

Hence, the first order condition

$$\langle H_u[t], v\rangle \leq 0, \quad \forall \ v = (v_1, v_2) \in T_U^\flat((0,0))$$

is satisfied.

Example

Let $v = (0,1)$ and $\nu_0 = (0,0)$. Then $y_1(t) \equiv (0,0)$ and the second order necessary condition is

$$\mathbb{E} \int_0^1 \langle H_u[t], h \rangle dt \leq 0, \qquad \forall\, h \in T_U^{\flat(2)}((0,0),(0,1)).$$

For $h = (-\frac{1}{2}, 0)$, we have $\mathbb{E} \int_0^1 \langle H_u[t], h \rangle dt = \frac{1}{8} > 0$.

A contradiction. Thus $\bar{u} \equiv 0$ is not optimal.

Thank you for your attention !!!