On Optimal Control Problems with Controls Appearing Nonlinearly in an Elliptic State Equation

E. Casas and F. Tröltzsch

Universidad de Cantabria and Technische Universität Berlin

Webinar "Control in Time of Crisis"

11 March 2021

Outline

Optimal control problem

$$\min J(u) = \int_\Omega L(x, y_u(x))\,dx$$

subject to

$$\begin{aligned} Ay &= f(x,y,u) && \text{in } \Omega \\ \partial_{n_A} y &= 0 && \text{in } \Gamma, \end{aligned}$$

$$\alpha \le u(x) \le \beta \qquad \text{for a.a. } x \in \Omega.$$

Data

- $\Omega \subset \mathbb{R}^N$, $N \in \{2,3\}$, bounded Lipschitz domain;
- $Ay = -\sum_{i,j=1}^n \partial_{x_j}(a_{ij}(x)\partial_{x_i} y) + a_0 y$, (uniformly elliptic, bounded coefficients)
- ∂_{n_A} : outward co-normal derivative
- $-\infty < \alpha \le \beta < \infty$.

Optimal control problem

$$\min J(u) = \int_\Omega L(x, y_u(x))\,dx$$

subject to

$$\begin{aligned} A\,y &= f(x,y,u) && \text{in } \Omega \\ \partial_{n_A} y &= 0 && \text{in } \Gamma, \end{aligned}$$

$$\alpha \le u(x) \le \beta \qquad \text{for a.a. } x \in \Omega.$$

Data

- $\Omega \subset \mathbb{R}^N$, $N \in \{2,3\}$, bounded Lipschitz domain;
- $A\,y = -\sum_{i,j=1}^n \partial_{x_j}(a_{ij}(x)\partial_{x_i} y) + a_0 y$, (uniformly elliptic, bounded coefficients)
- ∂_{n_A} : outward co-normal derivative
- $-\infty < \alpha \le \beta < \infty$.

State: $y \in H^1(\Omega) \cap L^\infty(\Omega)$ **Control:** $u \in L^\infty(\Omega)$.

Main difficulties

- How to show existence of an optimal control?

  The standard proof via minimizing sequences of controls and weak convergence (direct method of calculus of variations) does not work in view of the nonlinear appearance of u in the state equation.

$$u_k \rightharpoonup u \quad \not\Rightarrow \quad f(\cdot, y, u_k) \rightharpoonup f(\cdot, y, u)$$

Main difficulties

- How to show existence of an optimal control?

  The standard proof via minimizing sequences of controls and weak convergence (direct method of calculus of variations) does not work in view of the nonlinear appearance of u in the state equation.

$$u_k \rightharpoonup u \quad \not\Rightarrow \quad f(\cdot, y, u_k) \rightharpoonup f(\cdot, y, u)$$

- Can we expect that a stationary control obeys the coercivity assumptions for second-order sufficiency?

  If f is linear w.r. to u, then this is known to be impossible.

Main difficulties

- How to show existence of an optimal control?

  The standard proof via minimizing sequences of controls and weak convergence (direct method of calculus of variations) does not work in view of the nonlinear appearance of u in the state equation.

$$u_k \rightharpoonup u \quad \not\Rightarrow \quad f(\cdot, y, u_k) \rightharpoonup f(\cdot, y, u)$$

- Can we expect that a stationary control obeys the coercivity assumptions for second-order sufficiency?

  If f is linear w.r.t. to u, then this is known to be impossible.

- What about convergence of numerical approximations?

Our main results

- Proof of existence via the Filippov selection theorem

Our main results

- Proof of existence via the Filippov selection theorem
- Confirmation that second-order sufficiency can be induced by the nonlinearity f in the equation

Our main results

- Proof of existence via the Filippov selection theorem
- Confirmation that second-order sufficiency can be induced by the nonlinearity f in the equation
- Sufficient optimality conditions for bang-bang controls

Our main results

- Proof of existence via the Filippov selection theorem
- Confirmation that second-order sufficiency can be induced by the nonlinearity f in the equation
- Sufficient optimality conditions for bang-bang controls
- New versions of second-order sufficient conditions

Our main results

- Proof of existence via the Filippov selection theorem
- Confirmation that second-order sufficiency can be induced by the nonlinearity f in the equation
- Sufficient optimality conditions for bang-bang controls
- New versions of second-order sufficient conditions
- A convergence result for numerical approximations via the Filippov theorem

Assumptions on f and L

- $f : \Omega \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to (y, u). It satisfies

$$\frac{\partial f}{\partial y}(x, y, u) \leq 0 \quad \text{a.e. in } \Omega, \quad \forall (y, u) \in \mathbb{R}^2$$

$$f(\cdot, 0, 0) \in L^{\bar{p}}(\Omega) \text{ with some } \bar{p} > n/2.$$

Assumptions on f and L

- $f : \Omega \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to (y, u). It satisfies

$$\frac{\partial f}{\partial y}(x, y, u) \leq 0 \quad \text{a.e. in } \Omega, \quad \forall (y, u) \in \mathbb{R}^2$$
$$f(\cdot, 0, 0) \in L^{\bar{p}}(\Omega) \text{ with some } \bar{p} > n/2.$$

- $L : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second component with $L(\cdot, 0) \in L^1(\Omega)$.

Assumptions on f and L

- $f : \Omega \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to (y, u). It satisfies

$$\frac{\partial f}{\partial y}(x, y, u) \le 0 \quad \text{a.e. in } \Omega, \quad \forall (y, u) \in \mathbb{R}^2$$
$$f(\cdot, 0, 0) \in L^{\bar{p}}(\Omega) \text{ with some } \bar{p} > n/2.$$

- $L : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second component with $L(\cdot, 0) \in L^1(\Omega)$.
- The first- and second-order derivatives of f and L obey natural conditions of boundedness or integrability w.r. to x and local Lipschitz conditions w.r. to (y, u) or y, respectively.

Outline

Well-posedness of the state equation

Theorem

For every $u \in L^\infty(\Omega)$*, the state equation has a unique solution* $y_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ *for some* $\mu \in (0, 1]$ *independent of* u*.*

Well-posedness of the state equation

Theorem

For every $u \in L^\infty(\Omega)$, the state equation has a unique solution $y_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ for some $\mu \in (0,1]$ independent of u. There exists a constant $C_{\alpha,\beta}$ such that

$$\|y_u\|_{H^1(\Omega)} + \|y_u\|_{C^{0,\mu}(\overline{\Omega})} \le C_{\alpha,\beta}\Big(\|f(\cdot,0,0)\|_{L^{\bar p}(\Omega)} + 1\Big) \quad \forall u \in \mathcal{U}_{ad}.$$

Well-posedness of the state equation

Theorem

For every $u \in L^\infty(\Omega)$*, the state equation has a unique solution* $y_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ *for some* $\mu \in (0,1]$ *independent of* u*. There exists a constant* $C_{\alpha,\beta}$ *such that*

$$\|y_u\|_{H^1(\Omega)} + \|y_u\|_{C^{0,\mu}(\overline{\Omega})} \le C_{\alpha,\beta}\Big(\|f(\cdot,0,0)\|_{L^{\bar p}(\Omega)} + 1\Big) \quad \forall u \in \mathcal{U}_{ad}.$$

If $\{u_k\}_{k=1}^\infty$ *is a bounded sequence in* $L^\infty(\Omega)$ *converging to* u *in* $L^1(\Omega)$*, then*

$$\lim_{k\to\infty}\Big(\|y_{u_k} - y_u\|_{H^1(\Omega)} + \|y_{u_k} - y_u\|_{C^{0,\mu}(\overline{\Omega})}\Big) = 0.$$

Well-posedness of the state equation

Theorem

For every $u \in L^\infty(\Omega)$, the state equation has a unique solution $y_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ for some $\mu \in (0,1]$ independent of u. There exists a constant $C_{\alpha,\beta}$ such that

$$\|y_u\|_{H^1(\Omega)} + \|y_u\|_{C^{0,\mu}(\overline{\Omega})} \le C_{\alpha,\beta}\Big(\|f(\cdot,0,0)\|_{L^{\bar p}(\Omega)} + 1\Big) \quad \forall u \in \mathcal{U}_{ad}.$$

If $\{u_k\}_{k=1}^\infty$ is a bounded sequence in $L^\infty(\Omega)$ converging to u in $L^1(\Omega)$, then

$$\lim_{k\to\infty}\Big(\|y_{u_k} - y_u\|_{H^1(\Omega)} + \|y_{u_k} - y_u\|_{C^{0,\mu}(\overline{\Omega})}\Big) = 0.$$

Control-to-state mapping: $G : L^\infty(\Omega) \longrightarrow H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega}),\ \ G : u \mapsto y_u$.

Theorem

The mapping G is of class C^2.

Derivatives of J

Theorem

Under our assumptions, the mapping $J : L^\infty(\Omega) \longrightarrow \mathbb{R}$ is of class C^2 and we have

$$J'(u)v = \int_\Omega \varphi_u \frac{\partial f}{\partial u}(x, y_u, u)v\,dx \quad \forall u, v \in L^\infty(\Omega),$$

Derivatives of J

Theorem

Under our assumptions, the mapping $J : L^\infty(\Omega) \longrightarrow \mathbb{R}$ *is of class* C^2 *and we have*

$$J'(u)v = \int_\Omega \varphi_u \frac{\partial f}{\partial u}(x, y_u, u)v\,dx \quad \forall u, v \in L^\infty(\Omega),$$

$$J''(u)(v_1, v_2) = \int_\Omega \frac{\partial^2 L}{\partial y^2}(x, y_u) z_{v_1} z_{v_2}\,dx$$

$$+ \int_\Omega \varphi_u \Big[\frac{\partial^2 f}{\partial y^2}(x, y_u, u) z_{v_1} z_{v_2} + \frac{\partial^2 f}{\partial y \partial u}(x, y_u, u)(v_1 z_{v_2} + v_2 z_{v_1}) + \frac{\partial^2 f}{\partial u^2}(x, y_u, u) v_1 v_2\Big]\,dx$$

for all $u, v, v_1, v_2 \in L^\infty(\Omega)$, *where* $z_v = G'(u)v$ *and* $z_{v_i} = G'(u)v_i$, $i = 1, 2$, *and* $\varphi_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ *is the adjoint state associated with u, defined as the solution of the adjoint equation*

Derivatives of J

Theorem

Under our assumptions, the mapping $J : L^\infty(\Omega) \longrightarrow \mathbb{R}$ is of class C^2 and we have

$$J'(u)v = \int_\Omega \varphi_u \frac{\partial f}{\partial u}(x, y_u, u)v\,dx \quad \forall u, v \in L^\infty(\Omega),$$

$$J''(u)(v_1, v_2) = \int_\Omega \frac{\partial^2 L}{\partial y^2}(x, y_u) z_{v_1} z_{v_2}\,dx$$

$$+ \int_\Omega \varphi_u \Big[\frac{\partial^2 f}{\partial y^2}(x, y_u, u) z_{v_1} z_{v_2} + \frac{\partial^2 f}{\partial y \partial u}(x, y_u, u)(v_1 z_{v_2} + v_2 z_{v_1}) + \frac{\partial^2 f}{\partial u^2}(x, y_u, u) v_1 v_2\Big]\,dx$$

for all $u, v, v_1, v_2 \in L^\infty(\Omega)$, where $z_v = G'(u)v$ and $z_{v_i} = G'(u)v_i$, $i = 1, 2$, and $\varphi_u \in H^1(\Omega) \cap C^{0,\mu}(\overline{\Omega})$ is the adjoint state associated with u, defined as the solution of the adjoint equation

$$\begin{aligned} A^*\varphi &= \frac{\partial f}{\partial y}(x, y_u, u)\varphi + \frac{\partial L}{\partial y}(x, y_u) && \text{in } \Omega, \\ \partial_{n_{A^*}}\varphi &= 0 && \text{on } \Gamma. \end{aligned}$$

Outline

Simple counter example

An effect related to the nonlinear appearance of u.

$$\begin{aligned} -\Delta y(x) + y(x) &= u(x)^2 \quad \text{in } \Omega \\ \partial_n y &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Simple counter example

An effect related to the nonlinear appearance of u.

$$\begin{aligned} -\Delta y(x) + y(x) &= u(x)^2 && \text{in } \Omega \\ \partial_n y &= 0 && \text{on } \Gamma. \end{aligned}$$

Select a sequence of bang-bang control functions $u_k \in L^\infty(\Omega)$ such that

$$u_k(x) \in \{-1, 1\} \text{ a.e. in } \Omega$$

and $u_k \rightharpoonup 0$ (weakly) in $L^2(\Omega)$.

Simple counter example

An effect related to the nonlinear appearance of u.

$$\begin{aligned} -\Delta y(x) + y(x) &= u(x)^2 \quad \text{in } \Omega \\ \partial_n y &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Select a sequence of bang-bang control functions $u_k \in L^\infty(\Omega)$ such that

$$u_k(x) \in \{-1, 1\} \text{ a.e. in } \Omega$$

and $u_k \rightharpoonup 0$ (weakly) in $L^2(\Omega)$. We have $y_k = y_{u_k} = 1$,

$$\underbrace{-\Delta y_k(x)}_{0} + \underbrace{y_k(x)}_{1} = 1 \quad \text{in } \Omega$$

$y_k \to y \equiv 1$, but y is not the state associated with the weak limit control $\tilde{u} = 0$.

Less naive example

Consider

$$-\Delta y + y \quad = \quad e^{-y} u^2 + u \quad \text{in } \Omega, \qquad \partial_n y = 0.$$

Insert the sequence u_k of the last example.

Less naive example

Consider

$$-\Delta y + y \;=\; e^{-y}u^2 + u \quad \text{in } \Omega, \qquad \partial_n y = 0.$$

Insert the sequence u_k of the last example. The states y_k solve

$$-\Delta y_k + y_k \;=\; e^{-y_k} + u_k \quad \text{in } \Omega$$

and converge in $H^1(\Omega) \cap L^\infty(\Omega)$ to the solution y of

$$-\Delta y + y \;=\; e^{-y} \quad \text{in } \Omega$$

but not to the state $y_{\tilde{u}} = 0$ that is associated with weak limit control $\tilde{u} = 0$.

Filippov's Lemma on measurable selection

Lemma (Filippov 1959)

Let $\Omega \subset \mathbb{R}^n$ be measurable, $K \subset \mathbb{R}^m$ be compact, $f = f(x,u) : \Omega \times K \to \mathbb{R}^\ell$ be a Carathèodory function.

Suppose that $g(\cdot) : \Omega \to \mathbb{R}^\ell$ is a bounded and measurable function with

$$g(x) \in f(x, K)$$

for a.a. $x \in \Omega$. Then there is a measurable function $u(\cdot) : \Omega \to K$ such that

$$g(x) = f(x, u(x)) \text{ for a.a. } x \in \Omega.$$

The function u is called *measurable selection.*

Filippov's Lemma on measurable selection

Lemma (Filippov 1959)

Let $\Omega \subset \mathbb{R}^n$ *be measurable,* $K \subset \mathbb{R}^m$ *be compact,* $f = f(x,u) : \Omega \times K \to \mathbb{R}^\ell$ *be a Carathèodory function.*

Suppose that $g(\cdot) : \Omega \to \mathbb{R}^\ell$ *is a bounded and measurable function with*

$$g(x) \in f(x, K)$$

for a.a. $x \in \Omega$*. Then there is a measurable function* $u(\cdot) : \Omega \to K$ *such that*

$$g(x) = f(x, u(x)) \text{ for a.a. } x \in \Omega.$$

The function u is called *measurable selection.*

Application of Filippov's Lemma in optimal control:

Control of ODEs: Macki/Strauss 1982, Cesari 1983, ...

Filippov's Lemma on measurable selection

Lemma (Filippov 1959)

Let $\Omega \subset \mathbb{R}^n$ be measurable, $K \subset \mathbb{R}^m$ be compact, $f = f(x,u) : \Omega \times K \to \mathbb{R}^\ell$ be a Carathèodory function.

Suppose that $g(\cdot) : \Omega \to \mathbb{R}^\ell$ is a bounded and measurable function with

$$g(x) \in f(x, K)$$

for a.a. $x \in \Omega$. Then there is a measurable function $u(\cdot) : \Omega \to K$ such that

$$g(x) = f(x, u(x)) \text{ for a.a. } x \in \Omega.$$

The function u is called *measurable selection.*

Application of Filippov's Lemma in optimal control:

Control of ODEs: Macki/Strauss 1982, Cesari 1983, ...

Control of PDEs: Papageorgiou 1991, Eppler 1993, Eppler and Unger 1997,

Recent trends in set-valued analysis: Frankowska, Rampazzo, ...

Existence theorem

Theorem

The optimal control problem (P) *has at least one solution.*

Existence theorem

Theorem

The optimal control problem (P) *has at least one solution.*

Main steps of the proof

Existence theorem

Theorem

The optimal control problem (P) *has at least one solution.*

Main steps of the proof

- Select a minimizing sequence $\{u_k\}_{k=1}^{\infty} \subset \mathcal{U}_{ad}$: $J(u_k) \searrow \inf$ (P).
- A subsequence of the states, say $\{y_k\}$ converges weakly in $H^1(\Omega)$ and strongly in $C(\bar{\Omega})$ to some $\bar{y} \in H^1(\Omega) \cap C(\bar{\Omega})$.

Existence theorem

Theorem

The optimal control problem (P) *has at least one solution.*

Main steps of the proof

- Select a minimizing sequence $\{u_k\}_{k=1}^{\infty} \subset \mathcal{U}_{ad}$: $J(u_k) \searrow \inf(\mathrm{P})$.
- A subsequence of the states, say $\{y_k\}$ converges weakly in $H^1(\Omega)$ and strongly in $C(\overline{\Omega})$ to some $\bar{y} \in H^1(\Omega) \cap C(\overline{\Omega})$.
- $$J(u_k) = \int_\Omega L(x, y_k(x))\, dx \to \int_\Omega L(x, \bar{y}(x))\, dx = \inf(\mathrm{P}).$$

Existence theorem

Theorem

The optimal control problem (P) *has at least one solution.*

Main steps of the proof

- Select a minimizing sequence $\{u_k\}_{k=1}^{\infty} \subset \mathcal{U}_{ad}$: $J(u_k) \searrow \inf(\mathrm{P})$.
- A subsequence of the states, say $\{y_k\}$ converges weakly in $H^1(\Omega)$ and strongly in $C(\overline{\Omega})$ to some $\bar{y} \in H^1(\Omega) \cap C(\overline{\Omega})$.
- $$J(u_k) = \int_\Omega L(x, y_k(x))\, dx \to \int_\Omega L(x, \bar{y}(x))\, dx = \inf(\mathrm{P}).$$
- It remains to show that $\bar{y}$ is the state for some control $\bar{u} \in \mathcal{U}_{ad}$. Then $\bar{u}$ is an optimal control.

- Consider the multifunction

$$F(x) = f(x, \bar{y}(x), \underbrace{[\alpha, \beta]}_{K}), \quad x \in \Omega.$$

For a.a. $x \in \Omega$, $F(x)$ is a compact interval of $\mathbb{R}$, since f is continuous in u.

- Consider the multifunction

$$F(x) = f(x, \bar{y}(x), \underbrace{[\alpha, \beta]}_{K}), \quad x \in \Omega.$$

For a.a. $x \in \Omega$, $F(x)$ is a compact interval of $\mathbb{R}$, since f is continuous in u.

- Let

$$g_k(x) = f(x, \bar{y}(x), u_k(x)).$$

We show that $g_k \rightharpoonup \bar{g}$ in some space $L^p(\Omega)$ and

$$\bar{g}(x) \in f(x, \bar{y}(x), [\alpha, \beta]) \text{ for a.a. } x \in \Omega.$$

- Consider the multifunction

$$F(x) = f(x, \bar{y}(x), \underbrace{[\alpha, \beta]}_{K}), \quad x \in \Omega.$$

For a.a. $x \in \Omega$, $F(x)$ is a compact interval of $\mathbb{R}$, since f is continuous in u.

- Let

$$g_k(x) = f(x, \bar{y}(x), u_k(x)).$$

We show that $g_k \rightharpoonup \bar{g}$ in some space $L^p(\Omega)$ and

$$\bar{g}(x) \in f(x, \bar{y}(x), [\alpha, \beta]) \text{ for a.a. } x \in \Omega.$$

- Filippov Lemma $\Rightarrow$: There is a measurable $\bar{u} : \Omega \longrightarrow [\alpha, \beta]$, such that

$$\bar{g}(x) = f(x, \bar{y}(x), \bar{u}(x)) \qquad \text{a.e. in } \Omega.$$

Obviously, $\bar{u} \in \mathcal{U}_{ad}$.

- Finally, we show $\bar{y} = y_{\bar{u}}$.

Outline

First order necessary optimality conditions

The Hamiltonian of (P) is given by

$$H : \Omega \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad H(x, y, \varphi, u) = L(x, y) + \varphi f(x, y, u).$$

First order necessary optimality conditions

The Hamiltonian of (P) is given by

$$H : \Omega \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad H(x, y, \varphi, u) = L(x, y) + \varphi f(x, y, u).$$

Theorem

If $\bar{u}$ is a local minimum of (P) in the $L^p(\Omega)$-sense ($1 \le p \le \infty$), then

$$\int_\Omega \frac{\partial H}{\partial u}(x, \bar{y}, \bar{\varphi}, \bar{u})(u - \bar{u})\, dx \ge 0 \quad \forall u \in \mathcal{U}_{ad}.$$

First order necessary optimality conditions

The Hamiltonian of (P) is given by

$$H : \Omega \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad H(x, y, \varphi, u) = L(x, y) + \varphi f(x, y, u).$$

Theorem

If $\bar{u}$ is a local minimum of (P) in the $L^p(\Omega)$-sense ($1 \le p \le \infty$), then

$$\int_\Omega \frac{\partial H}{\partial u}(x, \bar{y}, \bar{\varphi}, \bar{u})(u - \bar{u})\, dx \ge 0 \quad \forall u \in \mathcal{U}_{ad}.$$

If $p < \infty$, then the Pontryagin principle holds

$$H(x, \bar{y}(x), \bar{\varphi}(x), \bar{u}(x)) = \min_{s \in [\alpha, \beta]} H(x, \bar{y}(x), \bar{\varphi}(x), s) \quad \text{for a.a. } x \in \Omega.$$

Second order necessary condition

Critical cone:

$$C_{\bar u} = \{v \in L^2(\Omega) : J'(\bar u)v = 0 \text{ and } v \text{ satisfies the sign conditions below}\},$$

$$v(x) \begin{cases} \geq 0 & \text{if } \bar u(x) = \alpha, \\ \leq 0 & \text{if } \bar u(x) = \beta. \end{cases}$$

Theorem (Casas/T. 2012)

If $\bar u$ is a local solution of (P) *in the $L^p(\Omega)$ sense, $1 \leq p \leq \infty$, then*

$$J''(\bar u)v^2 \geq 0 \quad \forall v \in C_{\bar u}.$$

Outline

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points (in exercises ...)

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points (in exercises ...)
- showing the convergence of optimal solutions of discretized nonlinear control problems as $h \to 0$

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points (in exercises ...)
- showing the convergence of optimal solutions of discretized nonlinear control problems as $h \to 0$
- ensuring their local uniqueness

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points (in exercises ...)
- showing the convergence of optimal solutions of discretized nonlinear control problems as $h \to 0$
- ensuring their local uniqueness
- proving convergence of numerical algorithms such as the semismooth Newton or SQP method.

The role of sufficient optimality conditions

Sufficient optimality conditions play a similar role as constraint qualifications in constrained optimization: They are needed in the analysis but in general they depend on the unknown optimal solution and hence they are difficult to confirm.

In optimal control, second-order sufficient conditions are an analysis tool for:

- verifying local optimality of stationary points (in exercises ...)
- showing the convergence of optimal solutions of discretized nonlinear control problems as $h \to 0$
- ensuring their local uniqueness
- proving convergence of numerical algorithms such as the semismooth Newton or SQP method.

Can sufficient conditions be confirmed numerically? NOT in general.

Structural assumption

Structural Assumption: There are $K > 0$, $\varepsilon_0 > 0$ and $\gamma \in (0, +\infty]$:

$$meas\left\{x \in \Omega : \left|\bar{\varphi}(x)\frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| < \varepsilon\right\} \le K\varepsilon^{\gamma} \quad \forall \varepsilon \le \varepsilon_0.$$

Structural assumption

Structural Assumption: There are $K > 0$, $\varepsilon_0 > 0$ and $\gamma \in (0, +\infty]$:

$$meas\left\{x \in \Omega : \left|\bar{\varphi}(x)\frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| < \varepsilon\right\} \le K\varepsilon^{\gamma} \quad \forall \varepsilon \le \varepsilon_0.$$

Notice that $\gamma = \infty$ can be taken, if the measure is zero for some $\varepsilon > 0$.

Structural assumption

Structural Assumption: There are $K > 0$, $\varepsilon_0 > 0$ and $\gamma \in (0, +\infty]$:

$$meas\left\{x \in \Omega : \left|\bar{\varphi}(x)\frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| < \varepsilon\right\} \leq K\varepsilon^\gamma \quad \forall \varepsilon \leq \varepsilon_0.$$

Notice that $\gamma = \infty$ can be taken, if the measure is zero for some $\varepsilon > 0$.

Remark: If $\bar{u}$ satisfies the structural assumption, it is bang-bang. It is fulfilled with $\gamma = +\infty$ and some $\varepsilon_0 > 0$ iff there exists $\sigma > 0$ such that

$$meas\left\{x \in \Omega : \left|\bar{\varphi}(x)\frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| < \sigma\right\} = 0.$$

Structural assumption

Structural Assumption: There are $K > 0$, $\varepsilon_0 > 0$ and $\gamma \in (0, +\infty]$:

$$meas\left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x, \bar y(x), \bar u(x))\right| < \varepsilon\right\} \le K\varepsilon^\gamma \quad \forall \varepsilon \le \varepsilon_0.$$

Notice that $\gamma = \infty$ can be taken, if the measure is zero for some $\varepsilon > 0$.

Remark: If $\bar u$ satisfies the structural assumption, it is bang-bang. It is fulfilled with $\gamma = +\infty$ and some $\varepsilon_0 > 0$ iff there exists $\sigma > 0$ such that

$$meas\left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x, \bar y(x), \bar u(x))\right| < \sigma\right\} = 0.$$

Theorem

Let the Structural Assumption hold and $\bar u \in \mathcal{U}_{ad}$ satisfy the first order necessary conditions. Then,

$$J'(\bar u)(u - \bar u) \ge \kappa \|u - \bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} \quad \forall u \in \mathcal{U}_{ad}$$

holds with $\kappa = \dfrac{1}{2[2(\beta-\alpha)K]^{\frac{1}{\gamma}}}$.

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x, \bar y(x), \bar u(x))\right| \ge \varepsilon\right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \le K\varepsilon^\gamma$.

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{ x \in \Omega : \left| \bar{\varphi}(x) \frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x)) \right| \geq \varepsilon \right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \leq K\varepsilon^\gamma$.

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{ x \in \Omega : \left| \bar\varphi(x) \frac{\partial f}{\partial u}(x, \bar y(x), \bar u(x)) \right| \ge \varepsilon \right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \le K\varepsilon^\gamma$.

$$J'(\bar u)(u - \bar u) = \int_\Omega \underbrace{\bar\varphi(x) \frac{\partial f}{\partial u}(x, \bar y(x), \bar u(x))(u(x) - \bar u(x))}_{\ge 0} \, dx$$

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{ x \in \Omega : \left| \bar{\varphi}(x) \frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x)) \right| \geq \varepsilon \right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \leq K\varepsilon^\gamma$.

$$J'(\bar{u})(u - \bar{u}) = \int_\Omega \underbrace{\bar{\varphi}(x) \frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))(u(x) - \bar{u}(x))}_{\geq 0} \, dx$$

$$\geq \int_{\Omega_\varepsilon} \underbrace{|\bar{\varphi}(x) \frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))|}_{\geq \varepsilon} \, |u(x) - \bar{u}(x)| \, dx \geq \varepsilon \, \|u - \bar{u}\|_{L^1(\Omega_\varepsilon)}$$

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))\right| \geq \varepsilon\right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \leq K\varepsilon^\gamma$.

$$J'(\bar u)(u-\bar u) = \int_\Omega \underbrace{\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))(u(x)-\bar u(x))}_{\geq 0}\,dx$$

$$\geq \int_{\Omega_\varepsilon} \underbrace{\left|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))\right|}_{\geq \varepsilon} |u(x)-\bar u(x)|\,dx \geq \varepsilon\,\|u-\bar u\|_{L^1(\Omega_\varepsilon)}$$

$$= \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - \varepsilon\,\|u-\bar u\|_{L^1(\Omega\setminus\Omega_\varepsilon)}$$

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))\right| \ge \varepsilon\right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \le K\varepsilon^\gamma$.

$$J'(\bar u)(u-\bar u) = \int_\Omega \underbrace{\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))(u(x)-\bar u(x))}_{\ge 0}\,dx$$

$$\ge \int_{\Omega_\varepsilon} \underbrace{\left|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))\right|}_{\ge\varepsilon} |u(x)-\bar u(x)|\,dx \ge \varepsilon\,\|u-\bar u\|_{L^1(\Omega_\varepsilon)}$$

$$= \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - \varepsilon\,\|u-\bar u\|_{L^1(\Omega\setminus\Omega_\varepsilon)}$$

$$\ge \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - \varepsilon\,(\beta-\alpha)\underbrace{|\Omega\setminus\Omega_\varepsilon|}_{\le K\varepsilon^\gamma} \ge \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - (\beta-\alpha)K\varepsilon^{1+\gamma}.$$

Steps of the proof

Proof.

$$\Omega_\varepsilon := \left\{x \in \Omega : \left|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))\right| \geq \varepsilon\right\}.$$

We have $|\Omega \setminus \Omega_\varepsilon| \leq K\varepsilon^\gamma$.

$$J'(\bar u)(u-\bar u) = \int_\Omega \underbrace{\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))(u(x)-\bar u(x))}_{\geq 0}\,dx$$

$$\geq \int_{\Omega_\varepsilon} \underbrace{|\bar\varphi(x)\frac{\partial f}{\partial u}(x,\bar y(x),\bar u(x))|}_{\geq\varepsilon}\,|u(x)-\bar u(x)|\,dx \geq \varepsilon\,\|u-\bar u\|_{L^1(\Omega_\varepsilon)}$$

$$= \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - \varepsilon\,\|u-\bar u\|_{L^1(\Omega\setminus\Omega_\varepsilon)}$$

$$\geq \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - \varepsilon\,(\beta-\alpha)\underbrace{|\Omega\setminus\Omega_\varepsilon|}_{\leq K\varepsilon^\gamma} \geq \varepsilon\,\|u-\bar u\|_{L^1(\Omega)} - (\beta-\alpha)K\varepsilon^{1+\gamma}.$$

Define $\varepsilon = [2(\beta-\alpha)K]^{-\frac{1}{\gamma}}\,\|u-\bar u\|_{L^1(\Omega)}^{\frac{1}{\gamma}}$. Then the statement of the Theorem is obtained. □

Sufficiency theorem

Theorem

Let the Structural Assumption and $\gamma = +\infty$ be fulfilled. If $\bar{u} \in \mathcal{U}_{ad}$ satisfies the variational inequality, then $\varepsilon > 0$ exists such that

$$J(u) \geq J(\bar{u}) + \frac{\kappa}{2}\|u - \bar{u}\|_{L^1(\Omega)} \quad \forall u \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u}),$$

where $\bar{B}_\varepsilon(\bar{u})$ is the closed ball in $L^\infty(\Omega)$ and κ is taken from the last theorem. Therefore, $\bar{u}$ is strictly locally optimal in the sense of $L^\infty(\Omega)$.

Proof

Take $u \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u})$ with $\varepsilon > 0$ to be fixed later.

Proof

Take $u \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u})$ with $\varepsilon > 0$ to be fixed later.

$$J(u) = J(\bar{u}) + J'(\bar{u})(u - \bar{u}) + \frac{1}{2} J''(\bar{u} + \theta(u - \bar{u}))(u - \bar{u})^2, \quad 0 < \theta < 1,$$

Proof

Take $u \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u})$ with $\varepsilon > 0$ to be fixed later.

$$J(u) = J(\bar{u}) + J'(\bar{u})(u - \bar{u}) + \frac{1}{2}J''(\bar{u} + \theta(u - \bar{u}))(u - \bar{u})^2, \quad 0 < \theta < 1,$$
$$\geq J(\bar{u}) + \underbrace{\kappa\, \|u - \bar{u}\|_{L^{1+0}(\Omega)}}_{\text{by } \gamma=\infty} - C_{\alpha,\beta}\|u - \bar{u}\|^2_{L^2(\Omega)}$$

Proof

Take $u \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u})$ with $\varepsilon > 0$ to be fixed later.

$$\begin{aligned}
J(u) &= J(\bar{u}) + J'(\bar{u})(u-\bar{u}) + \frac{1}{2}J''(\bar{u}+\theta(u-\bar{u}))(u-\bar{u})^2, \quad 0<\theta<1,\\
&\geq J(\bar{u}) + \underbrace{\kappa\,\|u-\bar{u}\|_{L^{1+0}(\Omega)}}_{\text{by } \gamma=\infty} - C_{\alpha,\beta}\|u-\bar{u}\|^2_{L^2(\Omega)}\\
&\geq J(\bar{u}) + \Big[\kappa - C_{\alpha,\beta}\underbrace{\|u-\bar{u}\|_{L^\infty(\Omega)}}_{\leq\varepsilon}\Big]\|u-\bar{u}\|_{L^1(\Omega)}\\
&\geq J(\bar{u}) + [\kappa - C_{\alpha,\beta}\varepsilon]\|u-\bar{u}\|_{L^1(\Omega)}
\end{aligned}$$

holds with a constant $C_{\alpha,\beta}$ independent of u.

Take $\varepsilon \leq \frac{\kappa}{2C_{\alpha,\beta}}$ to obtain the result.

Example - A local solution in the sense of $L^\infty(\Omega)$

$$\min \int_\Omega (y_u - y_d)^2 \, dx,$$

$$-\Delta y + y = e^{-y} u^2 + 2\, e^{u-2} + e_\Omega \quad \text{in } \Omega, \qquad \partial_n y = 0,$$

$$-1 \le u(x) \le 1 \quad \text{a.e. in } \Omega = (0, 2\pi) \times (0, 2\pi).$$

Example - A local solution in the sense of $L^\infty(\Omega)$

$$\min \int_\Omega (y_u - y_d)^2 \, dx,$$

$$-\Delta y + y = e^{-y} u^2 + 2\, e^{u-2} + e_\Omega \quad \text{in } \Omega, \qquad \partial_n y = 0,$$

$$-1 \le u(x) \le 1 \quad \text{a.e. in } \Omega = (0, 2\pi) \times (0, 2\pi).$$

We select a triplet $(\bar\varphi, \bar y, \bar u)$ as desired candidate for optimality:

$$\bar y \in H^2(\Omega) \text{ with } \partial_n \bar y = 0 \;\text{ and }\; \|\bar y\|_{C(\bar\Omega)} < 3,$$

e.g. $\bar y \equiv 1$ or $\bar y(x) = \cos(x_1)\cos(x_2)$, $\quad \bar\varphi \equiv -1$ as adjoint state, and

$$\bar u(x) = \begin{cases} +1 & \text{in } (0,\pi) \times (0,\pi) \cup (\pi, 2\pi) \times (\pi, 2\pi), \\ -1 & \text{in } (\pi, 2\pi) \times (0,\pi) \cup (0,\pi) \times (\pi, 2\pi). \end{cases}$$

Example - A local solution in the sense of $L^\infty(\Omega)$

$$\min \int_\Omega (y_u - y_d)^2\,dx,$$

$$-\Delta y + y = e^{-y}u^2 + 2\,e^{u-2} + e_\Omega \quad \text{in } \Omega, \qquad \partial_n y = 0,$$

$$-1 \le u(x) \le 1 \quad \text{a.e. in } \Omega = (0,2\pi)\times(0,2\pi).$$

We select a triplet $(\bar\varphi, \bar y, \bar u)$ as desired candidate for optimality:

$$\bar y \in H^2(\Omega) \text{ with } \partial_n \bar y = 0 \;\text{ and }\; \|\bar y\|_{C(\bar\Omega)} < 3,$$

e.g. $\bar y \equiv 1$ or $\bar y(x) = \cos(x_1)\cos(x_2)$, $\quad \bar\varphi \equiv -1$ as adjoint state, and

$$\bar u(x) = \begin{cases} +1 & \text{in } (0,\pi)\times(0,\pi)\cup(\pi,2\pi)\times(\pi,2\pi), \\ -1 & \text{in } (\pi,2\pi)\times(0,\pi)\cup(0,\pi)\times(\pi,2\pi). \end{cases}$$

Next, y_d and e_Ω are fixed such that $(\bar y, \bar\varphi, \bar u)$ obeys the first-order conditions.

Example - continuation

State equation $\Rightarrow$

$$e_\Omega = -\Delta \bar{y} + \bar{y} - e^{-\bar{y}} - 2e^{\bar{u}-2}$$

Example - continuation

State equation $\Rightarrow$

$$e_\Omega = -\Delta\bar{y} + \bar{y} - e^{-\bar{y}} - 2e^{\bar{u}-2}$$

Adjoint equation $\Rightarrow$

$$y_d = 1 + \bar{y} + e^{-\bar{y}}.$$

Example - continuation

State equation $\Rightarrow$

$$e_\Omega = -\Delta\bar{y} + \bar{y} - e^{-\bar{y}} - 2e^{\bar{u}-2}$$

Adjoint equation $\Rightarrow$

$$y_d = 1 + \bar{y} + e^{-\bar{y}}.$$

We set

$$\sigma := 2[e^{-\|\bar{y}\|_{C(\bar{\Omega})}} - e^{-3}] > 0.$$

After quite lengthy computations and estimations, we find

$$meas\ \left\{x \in \Omega : \left|\bar{\varphi}(x)\frac{\partial f}{\partial u}(x, \bar{y}(x), \bar{u}(x))\right| < \sigma\right\} = 0.$$

Therefore, $\gamma = +\infty$ can be taken. The last theorem yields that $\bar{u}$ is a strict local minimum of (P) in the $L^\infty(\Omega)$-sense.

L^2-sufficiency theorem

Theorem

Let $\bar{u} \in \mathcal{U}_{ad}$ satisfy the first order optimality conditions with $\bar{y}$ and $\bar{\varphi}$ and assume

$$\exists \nu > 0 \text{ such that } \frac{\partial^2 H}{\partial u^2}(x, \bar{y}(x), \bar{\varphi}(x), s) \geq \nu \quad \text{for a.a. } x \in \Omega,\ \forall s \in [\alpha, \beta].$$

L^2-sufficiency theorem

Theorem

Let $\bar{u} \in \mathcal{U}_{ad}$ satisfy the first order optimality conditions with $\bar{y}$ and $\bar{\varphi}$ and assume

$$\exists \nu > 0 \text{ such that } \frac{\partial^2 H}{\partial u^2}(x, \bar{y}(x), \bar{\varphi}(x), s) \geq \nu \quad \text{for a.a. } x \in \Omega,\ \forall s \in [\alpha, \beta].$$

(i) If the structural assumption is fulfilled with $\gamma > 1$, there exists $\varepsilon > 0$ such that

$$J(u) \geq J(\bar{u}) + \frac{\kappa}{2}\|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} + \frac{\nu}{8}\|u - \bar{u}\|_{L^2(\Omega)}^2 \quad \forall u \in \mathcal{U}_{ad} \cap B_\varepsilon^2(\bar{u})$$

with κ is given as above.

L^2-sufficiency theorem

Theorem

Let $\bar{u} \in \mathcal{U}_{ad}$ satisfy the first order optimality conditions with $\bar{y}$ and $\bar{\varphi}$ and assume

$$\exists \nu > 0 \text{ such that } \frac{\partial^2 H}{\partial u^2}(x, \bar{y}(x), \bar{\varphi}(x), s) \geq \nu \quad \text{for a.a. } x \in \Omega,\ \forall s \in [\alpha, \beta].$$

(i) If the structural assumption is fulfilled with $\gamma > 1$, there exists $\varepsilon > 0$ such that

$$J(u) \geq J(\bar{u}) + \frac{\kappa}{2}\|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} + \frac{\nu}{8}\|u - \bar{u}\|_{L^2(\Omega)}^2 \quad \forall u \in \mathcal{U}_{ad} \cap B_\varepsilon^2(\bar{u})$$

with κ is given as above.

(ii) If

$$J''(\bar{u})v^2 > 0 \quad \forall v \in C_{\bar{u}} \setminus \{0\},$$

$\exists \varepsilon > 0$, $\exists \delta > 0$ *such that*

$$J(u) \geq J(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|_{L^2(\Omega)}^2 \quad \forall u \in \mathcal{U}_{ad} \cap B_\varepsilon^2(\bar{u}).$$

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2} \int_\Omega (y_u - y_d)^2 \, dx, \qquad -1 \leq u \leq 1,$$
$$-\Delta y + y = e^{-y} u^2 + e^{2\,u} + e_\Omega, \qquad \partial_n y = 0,$$

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2}\int_\Omega (y_u - y_d)^2\,dx, \qquad -1 \le u \le 1,$$
$$-\Delta y + y = e^{-y}u^2 + e^{2\,u} + e_\Omega, \qquad \partial_n y = 0,$$

Again, y_d and $e_\Omega \in L^2(\Omega)$ will be adapted to have a desired optimal solution.

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2} \int_\Omega (y_u - y_d)^2 \, dx, \qquad -1 \le u \le 1,$$
$$-\Delta y + y = e^{-y} u^2 + e^{2\,u} + e_\Omega, \qquad \partial_n y = 0,$$

Again, y_d and $e_\Omega \in L^2(\Omega)$ will be adapted to have a desired optimal solution.

Adjoint equation:

$$-\Delta \varphi + \varphi = -e^{-y} u^2 \varphi + y - y_d.$$

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2} \int_\Omega (y_u - y_d)^2\,dx, \qquad -1 \le u \le 1,$$
$$-\Delta y + y = e^{-y} u^2 + e^{2\,u} + e_\Omega, \qquad \partial_n y = 0,$$

Again, y_d and $e_\Omega \in L^2(\Omega)$ will be adapted to have a desired optimal solution.

Adjoint equation:

$$-\Delta\varphi + \varphi = -e^{-y} u^2 \varphi + y - y_d.$$

We wish that $\bar u \equiv -1$, $\bar y \equiv 2$, $\bar\varphi \equiv 1$ solve the optimality system.

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2} \int_\Omega (y_u - y_d)^2 \, dx, \qquad -1 \leq u \leq 1,$$
$$-\Delta y + y = e^{-y} u^2 + e^{2u} + e_\Omega, \qquad \partial_n y = 0,$$

Again, y_d and $e_\Omega \in L^2(\Omega)$ will be adapted to have a desired optimal solution.

Adjoint equation:

$$-\Delta \varphi + \varphi = -e^{-y} u^2 \varphi + y - y_d.$$

We wish that $\bar{u} \equiv -1$, $\bar{y} \equiv 2$, $\bar{\varphi} \equiv 1$ solve the optimality system.

Adjoint equation $\Rightarrow \quad y_d = 1 - e^{-2}$.

Example - local solution in the sense of $L^2(\Omega)$

$$\min \frac{1}{2}\int_\Omega (y_u - y_d)^2\,dx, \qquad -1 \le u \le 1,$$
$$-\Delta y + y = e^{-y}u^2 + e^{2u} + e_\Omega, \qquad \partial_n y = 0,$$

Again, y_d and $e_\Omega \in L^2(\Omega)$ will be adapted to have a desired optimal solution.

Adjoint equation:

$$-\Delta\varphi + \varphi = -e^{-y}u^2\varphi + y - y_d.$$

We wish that $\bar u \equiv -1$, $\bar y \equiv 2$, $\bar\varphi \equiv 1$ solve the optimality system.

Adjoint equation $\Rightarrow$ $y_d = 1 - e^{-2}$.

Variational inequality:

$$\frac{\partial H}{\partial u}(\cdot, \bar y, \bar\varphi, \bar u) = 2(e^{-\bar y}\bar u + e^{2\bar u}) = 2(-e^{-2} + e^{-2}) = 0,$$

the variational inequality is trivially satisfied.

Example, continuation

Inserting $\bar{y} = 2$ in the state equation, we find $e_\Omega = 2(1 - e^{-2})$.

Example, continuation

Inserting $\bar{y} = 2$ in the state equation, we find $e_\Omega = 2(1 - e^{-2})$.

$$J''(\bar{u})v^2 = \int_\Omega \{z_v^2 + \bar{\varphi}\,(e^{-\bar{y}}\bar{u}^2 z_v^2 - 4\,\bar{u}e^{-\bar{y}} v z_v + (2e^{-\bar{y}} + 4\,e^{2\,\bar{u}})v^2)\}\,dx$$

$$= \dots$$

$$= \int_\Omega \left\{z_v^2 + 2\,e^{-2}\Big[-\frac{1}{2}z_v^2 + (z_v + v)^2 + 2v^2\Big]\right\}\,dx$$

$$\geq 4\,e^{-2}\|v\|_{L^2(\Omega)}^2.$$

Example, continuation

Inserting $\bar{y} = 2$ in the state equation, we find $e_\Omega = 2(1 - e^{-2})$.

$$
\begin{aligned}
J''(\bar{u})v^2 &= \int_\Omega \{z_v^2 + \bar{\varphi}\left(e^{-\bar{y}}\bar{u}^2 z_v^2 - 4\,\bar{u}e^{-\bar{y}}vz_v + (2e^{-\bar{y}} + 4\,e^{2\,\bar{u}})v^2\right)\}\,dx \\
&= \dots \\
&= \int_\Omega \left\{z_v^2 + 2\,e^{-2}\left[-\frac{1}{2}z_v^2 + (z_v + v)^2 + 2v^2\right]\right\}\,dx \\
&\geq 4\,e^{-2}\|v\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Legendre-Clebsch condition

$$
\frac{\partial^2 H}{\partial u^2}(\cdot, \bar{y}, \bar{\varphi}, \bar{u}) = 6e^{-2} > 0.
$$

Therefore, the control $\bar{u} \equiv -1$ is locally optimal in the $L^2(\Omega)$-sense.

Outline

Additional assumptions for the discretization

Assume that:

- Ω is convex in $\mathbb{R}^n$ with $n \in \{2, 3\}$.

Additional assumptions for the discretization

Assume that:

- Ω is convex in $\mathbb{R}^n$ with $n \in \{2, 3\}$.
- Ω is polygonal ($n = 2$) or polyhedral ($n = 3$).

Additional assumptions for the discretization

Assume that:

- Ω is convex in $\mathbb{R}^n$ with $n \in \{2, 3\}$.
- Ω is polygonal ($n = 2$) or polyhedral ($n = 3$).
- $a_{ij} \in C^{0,1}(\overline{\Omega})$ for $1 \leq i, j \leq n$ and $f(\cdot, 0, 0) \in L^2(\Omega)$.

Additional assumptions for the discretization

Assume that:

- Ω is convex in $\mathbb{R}^n$ with $n \in \{2, 3\}$.
- Ω is polygonal ($n = 2$) or polyhedral ($n = 3$).
- $a_{ij} \in C^{0,1}(\overline{\Omega})$ for $1 \leq i, j \leq n$ and $f(\cdot, 0, 0) \in L^2(\Omega)$.

Under these conditions, the state belongs to $H^2(\Omega)$ and all possible state functions are bounded in the norm of $H^2(\Omega)$, independently of u.

Additional assumptions for the discretization

Assume that:

- Ω is convex in $\mathbb{R}^n$ with $n \in \{2, 3\}$.
- Ω is polygonal ($n = 2$) or polyhedral ($n = 3$).
- $a_{ij} \in C^{0,1}(\overline{\Omega})$ for $1 \leq i, j \leq n$ and $f(\cdot, 0, 0) \in L^2(\Omega)$.

Under these conditions, the state belongs to $H^2(\Omega)$ and all possible state functions are bounded in the norm of $H^2(\Omega)$, independently of u.

We consider a quasi-uniform family of triangulations $\{\mathcal{T}_h\}_{h>0}$ of $\overline{\Omega}$ with mesh size h according to Brenner and Scott (2008).

The discretized problem

Associated with the triangulations, we consider finite dimensional

Ansatz spaces

$$Y_h = \{y_h \in C(\bar{\Omega}) : \; y_{h|T} \in P_1(T) \; \forall T \in \mathcal{T}_h\}$$

$$\mathcal{U}_h = \{u_h \in L^\infty(\Omega) : \; u_{h|T} \in P_0(T) \; \forall T \in \mathcal{T}_h\}$$

where $P_i(T)$ denotes the space of polynomials in T of degree i, $i = 1, 2$.

The discretized problem

Associated with the triangulations, we consider finite dimensional

Ansatz spaces

$$Y_h = \{y_h \in C(\bar\Omega) : \ y_{h|T} \in P_1(T)\ \forall T \in \mathcal{T}_h\}$$
$$\mathcal{U}_h = \{u_h \in L^\infty(\Omega) : \ u_{h|T} \in P_0(T)\ \forall T \in \mathcal{T}_h\}$$

where $P_i(T)$ denotes the space of polynomials in T of degree i, $i = 1, 2$.
Define

$$U_{\mathrm{ad},h} = \mathcal{U}_h \cap \mathcal{U}_{ad}.$$

The discretized problem

Associated with the triangulations, we consider finite dimensional

Ansatz spaces

$$Y_h = \{y_h \in C(\bar\Omega) : \; y_{h|T} \in P_1(T) \; \forall T \in \mathcal{T}_h\}$$
$$\mathcal{U}_h = \{u_h \in L^\infty(\Omega) : \; u_{h|T} \in P_0(T) \; \forall T \in \mathcal{T}_h\}$$

where $P_i(T)$ denotes the space of polynomials in T of degree i, $i = 1, 2$.
Define

$$U_{\mathrm{ad},h} = \mathcal{U}_h \cap \mathcal{U}_{ad}.$$

Bilinear form $a : H^1(\Omega) \times H^1(\Omega) \longrightarrow \mathbb{R}$

$$a(y, v) = \int_\Omega \Big(\sum_{i,j=1}^n a_{ij}(x) \partial_{x_i} y \partial_{x_j} v + a_0(x) y \, v \Big) \, dx.$$

Discretization

Discretized state equation

Given $u \in L^\infty(\Omega)$, let $y_h(u) \in Y_h$ be the solution of the nonlinear system

$$a(y_h, z_h) = \int_\Omega f(x, y_h(x), u)\, z_h(x)\, dx \quad \forall z_h \in Y_h.$$

Discretization

Discretized state equation

Given $u \in L^\infty(\Omega)$, let $y_h(u) \in Y_h$ be the solution of the nonlinear system

$$a(y_h, z_h) = \int_\Omega f(x, y_h(x), u)\, z_h(x)\, dx \quad \forall z_h \in Y_h.$$

Existence of a solution is proved by Brouwer's fixed point theorem.
Uniqueness follows from monotonicity of f and coercivity of a.

Discretization

Discretized state equation

Given $u \in L^\infty(\Omega)$, let $y_h(u) \in Y_h$ be the solution of the nonlinear system

$$a(y_h, z_h) = \int_\Omega f(x, y_h(x), u)\, z_h(x)\, dx \;\; \forall z_h \in Y_h.$$

Existence of a solution is proved by Brouwer's fixed point theorem.
Uniqueness follows from monotonicity of f and coercivity of a.

Discretized optimal control problem

$$(\mathcal{P}_h) \qquad \min_{u_h \in U_{ad,h}} J_h(u_h) := \int_\Omega L(x, y_h(u_h)(x))\, dx.$$

Convergence of discretized solutions

Theorem

- *For every $h > 0$, the discrete control problem $(\mathcal{P}_h)$ has at least one solution $\bar{u}_h$.*

Convergence of discretized solutions

Theorem

- *For every* $h > 0$*, the discrete control problem* $(\mathcal{P}_h)$ *has at least one solution* $\bar{u}_h$*.*
- *If* $\{\bar{y}_h\}_{h>0}$ *is a family of discrete optimal states associated with optimal controls* $\{\bar{u}_h\}_{h>0}$*, then,* $\{\bar{y}_h\}_{h>0}$ *is bounded in* $H^1(\Omega) \cap C(\overline{\Omega})$*.*

Convergence of discretized solutions

Theorem

- *For every $h > 0$, the discrete control problem $(\mathcal{P}_h)$ has at least one solution $\bar{u}_h$.*
- *If $\{\bar{y}_h\}_{h>0}$ is a family of discrete optimal states associated with optimal controls $\{\bar{u}_h\}_{h>0}$, then, $\{\bar{y}_h\}_{h>0}$ is bounded in $H^1(\Omega) \cap C(\bar{\Omega})$.*
- *If $h_k \searrow 0$, $k \to \infty$, and $\{\bar{y}_{h_k}\}_{k=1}^{\infty} \rightharpoonup \bar{y}$ in $H^1(\Omega)$, then there is an optimal solution $\bar{u} \in \mathcal{U}_{ad}$ of (P), such that $\bar{y}$ is the associated state.*

Convergence of discretized solutions

Theorem

- *For every $h > 0$, the discrete control problem $(\mathcal{P}_h)$ has at least one solution $\bar{u}_h$.*
- *If $\{\bar{y}_h\}_{h>0}$ is a family of discrete optimal states associated with optimal controls $\{\bar{u}_h\}_{h>0}$, then, $\{\bar{y}_h\}_{h>0}$ is bounded in $H^1(\Omega) \cap C(\bar{\Omega})$.*
- *If $h_k \searrow 0$, $k \to \infty$, and $\{\bar{y}_{h_k}\}_{k=1}^{\infty} \rightharpoonup \bar{y}$ in $H^1(\Omega)$, then there is an optimal solution $\bar{u} \in \mathcal{U}_{ad}$ of (P), such that $\bar{y}$ is the associated state.*
- *$\{\bar{y}_{h_k}\}_{k=1}^{\infty}$ converges strongly in $H^1(\Omega)$ to $\bar{y}$.*

Proof of part 3:

We apply the Filippov theorem to show that $\bar{y}$ is a state associated with some control $\bar{u}$:

Again, we use the multifunction

$$F(x) = f(x, \bar{y}(x), [\alpha, \beta]) = \{f(x, \bar{y}(x), s) : s \in [\alpha, \beta]\}.$$

$$S := \{g \in L^2(\Omega) : g(x) \in F(x) \text{ for a.a. } x \in \Omega\}.$$

S is a convex and closed subset of $L^2(\Omega)$.

$$g_k(x) := f(x, \bar{y}(x), \bar{u}_{h_k}(x)), \quad k \geq 1.$$

All g_k belong to S; boundedness of $\{g_k\}$ in $L^2(\Omega)$; weak convergence, weak closedness of S:

$$g_k \rightharpoonup \bar{g} \in S.$$

$$g_k(x) := f(x, \bar{y}(x), \bar{u}_{h_k}(x)), \quad k \geq 1.$$

All g_k belong to S; boundedness of $\{g_k\}$ in $L^2(\Omega)$; weak convergence, weak closedness of S:

$$g_k \rightharpoonup \bar{g} \in S.$$

Filippov theorem $\Longrightarrow$ $\exists \bar{u} : \Omega \to [\alpha, \beta]$, measurable, such that

$$\bar{g}(x) = f(x, \bar{y}(x), \bar{u}(x)).$$

$$g_k(x) := f(x, \bar y(x), \bar u_{h_k}(x)), \quad k \ge 1.$$

All g_k belong to S; boundedness of $\{g_k\}$ in $L^2(\Omega)$; weak convergence, weak closedness of S:

$$g_k \rightharpoonup \bar g \in S.$$

Filippov theorem $\Longrightarrow$ $\exists \bar u : \Omega \to [\alpha, \beta]$, measurable, such that

$$\bar g(x) = f(x, \bar y(x), \bar u(x)).$$

We show that $\bar y$ is the state associated with $\bar u$. Standard estimates confirm that $\bar y$ attains the infimum of (P). Therefore, $\bar u$ is an optimal control.

References

E. Casas and F.T.
On optimal control problems with controls appearing nonlinearly in an elliptic state equation.
SIAM J. Control and Optimization 58 (4), 2020, 1961-1983.

E. Casas, D. Wachsmuth, and G. Wachsmuth.
Sufficient second-order conditions for bang-bang control problems.
SIAM J. Control Optim., 55(5):3066-3090, 2017.

E. Casas, D. Wachsmuth, and G. Wachsmuth.
Second order analysis and numerical approximation for bang-bang bilinear control problems.
SIAM J. Control Optim., 56(6):4203-4227, 2018.

Thank you