A Unified Approach to State Constrained Optimal Control, Based on Distance Estimates

Richard B. Vinter, Imperial College

'State Constrained Dynamical Systems' Conference

Dip. di Matematica "Tullio Levi-Civita", Univ. di Padova

September 25-29, 2017

Outline of the talk

- Overview of themes in optimal control
- The role of distance estimates in state constrained optimal control
- Linear and superlinear estimates : summary of known results and counter-examples
- An application from civil engineering design
- Final remarks

The Classical Optimal Control

$$\begin{cases} \text{Minimize } g(x(1)) \\ \text{over functions } u(.) : [0,1] \to \mathbb{R}^m \,, \\ \text{and trajectories } x(.) \text{ s.t.} \\ \dot{x}(t) \;=\; f(t, x(t), u(t)) \quad \text{for a.e. } t \in [0,1] \\ u(t) \in U \subset \mathbb{R}^m \quad \text{for a.e. } t \in [0,1] \\ \text{and } x(0) = x_0, \;\; x(1) \in C \end{cases}$$

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

Application Areas

1. Aerospace: flight trajectories for planetary exploration
2. Resource economics: optimal harvesting
3. Chemical engineering: optimize yield, purity etc.
4. Feedback Design: solution of optimal control problems for MPC schemes

Methodologies for Optimal Control

1. **Dynamic Programming** (Sufficient conditions for optimality): *'Analyse minimizers via solutions (the value function) to the Hamilton Jacobi equation'* (**R. Bellman**).

2. **Maximum Principle** (Necessary conditions for optimality): *'Analyse minimizers via solutions to a system which involves state and adjoint variables'* (**L.S. Pontryagin**)

3. **Higher Order Sufficient Conditions** (Sufficient conditions for local optimality): *'Confirm local optimality of extremals '*

'Analyse minimizers via solutions to the Hamilton Jacobi equation' (**R. Bellman**)

(Assume $C = \mathbb{R}^n$)

$$P(0, x_0) \begin{cases} \text{Minimize } g(x(1)) \\ \text{over trajectories } x(.) \text{ s.t. } x(0) = x_0 . \end{cases}$$

Embed in family of problems, parameterized by initial data

$$P(\tau, \xi) \begin{cases} \text{Minimize } g(x(1)) \\ \text{over trajectories } x(.) \text{ s.t. } x(\tau) = \xi\,. \end{cases}$$

Define $\boxed{V(\tau, \xi) = \text{Inf}(P(\tau, \xi))}$ **Value Function**

Hamilton Jacobi Methods (Dynamic Programming)

$$P(\tau,\xi)\begin{cases}\text{Minimize } g(x(1)) \\ \text{over trajectories } x(.) \text{ s.t. } x(\tau)=\xi\end{cases}$$

PDE of Dynamic Programming: $V(.,.)$ is a solution to

$$\begin{cases} V_t(t,x) + \min_{u\in U} V_x(t,x)\cdot f(t,x,u) \;=\; 0 \quad \forall (t,x)\in(0,1)\times\mathbb{R}^n \\ V(1,x) \;=\; g(x) \quad \forall x\in\mathbb{R}^n\,. \end{cases}$$

(HJE)

Modern methods of nonlinear analysis yield characterization:

'The value function is the unique generalized solution (appropriately defined) to (HJE) '

(non-smooth analysis, viability theory, viscosity solns. theory)

First Order Necessary Conditions

(**L.S. Pontryagin**,...)
Take an optimal pair $(\bar{x}(.), \bar{u}(.))$. Define

$$H(t, x, p, u) \;=\; p \cdot f(t, x, u) \quad \text{(The Hamiltonian)}\;.$$

Maximum Principle: There exist an arc $p(.)$ (adjoint variable) and $\lambda \geq 0$, s.t.

$$\begin{aligned}
(p(.), \lambda) \;&\neq\; 0 \\
-\dot{p}(t) \;&=\; p(t) \cdot f_x(t, \bar{x}(t), \bar{u}(t)) \\
H(t, \bar{x}(t), p(t), \bar{u}(t)) \;&=\; \max_{u \in U} H(t, \bar{x}(t), p(t), u) \\
-p(1) \;&=\; \lambda\, g_x(\bar{x}(1)) + \xi, \;\; \text{for some } \xi \in N_C(\bar{x}(1))
\end{aligned}$$

Widely used to solve optimal control problems, either directly or via numerical methods it inspires (Shooting Methods).

Enter State Constraints

Consider *state constrained* control system

$$\begin{cases} \text{Minimize } g(x(1)) \\ \text{over functions } u(.) : [0,1] \to \mathbb{R}^m \ , \\ \text{and trajectories } x(.) \text{ s.t.} \\ \dot{x}(t) \ = \ f(t, x(t), u(t)) \quad \text{for a.e. } t \in [0,1] \\ u(t) \in U \subset \mathbb{R}^m \quad \text{for a.e. } t \in [0,1] \\ x(t) \in A \quad \text{for all } t \in [0,1] \qquad (\text{state constraint}) \\ \text{and } x(0) = x_0 . \end{cases}$$

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

Special case: A has a *functional inequality representation*

$$A = \{x \in \mathbb{R}^n \,|\, h_j(x) \leq 0, j = 1, \ldots, r\}$$

for some C^1 functions $h_j(.) : \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, r$.

Standing Hypotheses

Assume that for some $c > 0$ and $k_f(.) \in L^1$

- $f(.,x,.)$ is $\mathcal{L} \times \mathcal{B}^m$ (Lebesgue-Borel) meas. for each x; $U(.)$ has Borel-meas. graph; $f(t,x,U)$ is closed, for each t, x
- $|f(t,x,u)| \le c(1+|x|)$ for all $(t,x) \in [0,1] \times \mathbb{R}^n$, $u \in U(t)$
- $|f(t,x,u) - f(t,x',u)| \le k_f(t)|x - x'|$ for all $t \in [0,1]$, $x, x' \in \mathbb{R}^n$ and $u \in U$.

(summarized as 'f is meas., integr. Lip., with linear growth')

- $g(.)$ Lipschitz, C closed.

Dynamic Programming for State Constrained Problems

$$\begin{cases} \text{Minimize } g(x(1)) \\ \text{over trajectories } x(.) \text{ s.t.} \\ x(t) \in A \\ x(0) = x_0. \end{cases}$$

How does state constraint affect optimality conditions?

Now, value function $V(.,.) : [0,1] \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lsc solution to

$$\begin{cases} V_t(t,x) + \min_{v \in U} V_x(t,x) \cdot f(t,x,u) = 0 \quad \forall (t,x) \in (0,1) \times \text{int } A \\ V(1,x) = g(x) \quad \forall x \in A \end{cases}$$

(the unique lsc solution if certain distance estimates are satisfied)

State Constrained Maximum Principle

Take an optimal pair $(\bar{x}(.), \bar{u}(.))$.

There exist an arc $p(.)$, 'bounded variation' multipliers $\mu_j \geq 0$, $j = 1, \ldots, r$ and $\lambda \geq 0$, s.t.

$$
\begin{aligned}
(p(.), \mu, \lambda) &\neq 0 \\
\text{supp}\{d\mu_j\} &\in \{t | h_j(\bar{x}(t)) = 0\} \\
-\dot{d}p(t) &= p(t) \cdot f_x(\bar{x}(t), \bar{u}(t))dt - \sum_j h_{xj}(\bar{x}(t))d\mu_j \\
H(\bar{x}(t), p(t), \bar{u}(t)) &= \max_{u \in U} H(\bar{x}(t), p(t), u) \\
-p(1) &= \lambda g_x(\bar{x}(1)) .
\end{aligned}
$$

(Formally obtained by inserting into cost the 'penalty' term $+K \sum_j \int_0^1 h_j(x(t))d\mu_j$*).)*

(Gamkrelidze, Neustadt, Warga, Milyutin . .)

Abstract Optimization Problem

Consider the optimization problem

$$\begin{cases} \text{Minimize } g(x) \\ \text{over } x \in \mathcal{X} \\ \text{ s.t.} \\ F(x) \subset D \end{cases}$$

Data: Metric Spaces $(\mathcal{X}, d_{\mathcal{X}}(.))$ and $(\mathcal{Y}, d_{\mathcal{Y}}(.))$, function $g : \mathcal{X} \to R$, multifunction $F : \mathcal{X} \rightsquigarrow \mathcal{Y}$.

Beyond theory of Necessary Conditions, early interest shown in

- Non-degeneracy of optimality conditions
- Sensitivity/continuous dependence
- Stability of solutions to generalized equations to parameter variation
- Rates of convergence for computational schemes

(Robinson, Rockafellar, Mordukhovich, Aubin, Bonnans etc. $\geq$ 1970's)

Key concept: Metric Regularity

Metric Regularity

Take metric spaces $(X, d_X).,.)$ and $(Y, d_Y).,.)$ and $H : X \rightsquigarrow Y$.

Definition. H is metrically regular at $(\bar{x}, \bar{y})$ if there exist $\kappa \geq 0$ and neighbourhoods $\mathcal{V}$ and $\mathcal{W}$ of $\bar{x}$ and $\bar{y}$ such that

$$d_X(H^{-1}(y)|x) \;\leq\; \kappa\, d_Y(H(x)|y) \quad \text{for all } (x, y) \in \mathcal{V} \times \mathcal{W}\,.$$

where $d_X(S|x) = \inf_{x' \in S}\{d_X(x, x')\}$, etc.

Metric regularity is an unrestrictive hyp. ensuring these 'good' properties.

For example:

- Interest in verifiable sufficient conditions of metric regularity, e.g. if '$H(x) \subset D$' $\equiv$ '$\psi_i(x) \leq 0$, $\forall i$, '$\phi_j(x) \leq 0$, $\forall j$'

'positive linear independence' $\implies$ *'metric regularity'.*

Return to Control . .

Control system:

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \ \ \textit{and} \ \ u(t) \in U \\ h_j(x(t)) \leq 0 \quad \textit{for } j = 1, \ldots, r \,. \end{cases}$$

- 'metric regularity' replaced by 'linear distance estimates'
- verifiable sufficient conditions replace by

  Inward pointing condition:

  for each $t \in [0, 1]$ *and* $x \in \partial A$

  $$\limsup_{t' \to t} \nabla_x h_j(x) \cdot f(t', x, u) < 0 \ \forall j \in I(x)$$

  $$(I(x) := \text{ 'active' indices at } x \ )$$

  More generally:

  $$(\limsup_{(t',x') \to t} f(t', x', U)) \cap \mathrm{int} T_A(x) \neq \emptyset, \ \forall \ t \in [0, 1], \ x \in \partial A$$

  $T_A(x)$ is (Clarke) tangent cone.

Distance Estimates

For an arc $x(.)$ define

$$h^+(x(.)) \ := \ \max_{t\in[0,1]} d_A(x(t))$$

in which

$$d_A(x) := \inf_{y\in A} |x - y|$$

(Euclidean distance of x from A).

$h^+(x(.))$ is the 'constraint violation index' of an arc $x(.)$:

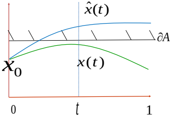
- $h^+(x(.)) = 0$ iff $x(.)$ is 'feasible',
  (i.e. $x(.)$ satisfies the state constraint)
- *$h^+(x(.))$ quantifies the state constraint violation.*

Linear Distance Estimates

A typical (linear) distance estimate asserts:

Given a non-feasible state trajectory $\hat{x}(.)$ with $\hat{x}(0) \in A$, there exists a feasible state trajectory $x(.)$ s.t. $x(0) = \hat{x}(0)$ and

$$||x(.) - \hat{x}(.)|| \leq K \times h^{+}(\hat{x}(.)),$$



where K is a positive constant that does not depend on $\hat{x}(.)$.

($||.||$ is some norm defined on the set of trajectories, for instance L^{∞} or $W^{1,1}$.)

Here we have a *linear* estimate w.r.t. the constraint violation index $h^{+}(\hat{x}(.))$

$$||x(.)||_{L^{\infty}} = \sup_{t \in [0,1]} |x(t)|, \quad ||x(.)||_{W^{1,1}} = |x(0)| + \int_{[0,1]} |\dot{x}(t)| \, dt$$

More General Estimates

More generally, we can consider the following estimate

$$m((x(.), u(.)), (\hat{x}(.)\hat{u}(.))) \leq \theta(h^+(\hat{x}(.))) \ ,$$

where

- $m(.,.)$ *is a metric on the set of processes*
- $\theta(.) : \mathbb{R}^+ \to \mathbb{R}^+$ *is a rate of convergence modulus*, i.e. a function satisfying $\lim_{h \downarrow 0} \theta(h) = 0$.

The stronger the metric $m(.,.)$ and greater the rate at which $\theta(h)$ tends to zero as $h \to 0$, the more the information that is conveyed by the estimates. A variety of estimates has been considered, distinguished by the choice of $m(.,.)$ and $\theta(.)$.

Soner's Linear L^∞ Estimate

First significant distance estimate:

Theorem *(Soner '86, improved Frankowska/Rampazzo '00)*
Assume standing hyps. and

- *$F(.,.)$ is Lipschitz continuous*
- *For all $x \in \partial A$,*

$$\text{int}\, T_A(x) \cap f(t,x,U) \neq \emptyset \quad (\textit{\textbf{'inward pointing' condition}})$$

Then, for any pair $(\hat{x}(.), \hat{u}(.))$ s.t. $\hat{x}(0) \in A$, there exists a feasible pair $(x(.), u(.))$ such that $x(0) = \hat{x}(0)$ and

$$||\hat{x}(.) - x(.)||_{L^\infty} \leq K \times h(\hat{x}(.))$$

(K does not depend on $\hat{x}(.)$)

application: value function regularity

Significance of Distance Estimates

Distance estimates constitute a common set of analytical tools which can be used to resolve a number of important questions in *state constrained* optimal control.
Some applications are

- non-degeneracy and normality of the *maximum principle* (which provides necessary conditions for optimality);
- existence, characterization and regularity of the *value function* for Hamilton-Jacobi-Bellman and Hamilton-Jacobi-Isaacs equations;
- *sensitivity conditions*: adjoint variables in the Maximum Principle can be interpreted as 'gradients' of the value function;
- Minimizer regularity;
- identify possible *ill-conditioning* of numerical schemes

Contributions to This Area (Partial List)

H. M. Soner, "Optimal Control Problems with State-Space Constraints 1 & 2", *SIAM J. Control Optim.*, 24, 1986.

F. Rampazzo and R. B. Vinter, "A Theorem on Existence of Neighbouring Trajectories Satisfying a State Constraint, with Applications to Optimal Control", *IMA J. Math. Control Inform*, 16, 1999.

F. Forcellini and F. Rampazzo, "On Non-convex Differential Inclusions whose State is Constrained in the closure of an Open Set", *J. Differential Integral Equations*, 12, 1999.

H. Frankowska and F. Rampazzo, "Filippov's and Filippov-Wazewski's Theorems on Closed Domains", *JDE*, 2000.

H. Frankowska and R. B. Vinter, "Existence of Neighbouring Feasible Trajectories: Applications to Dynamic Programming for State Constrained Optimal Control Problems", *JOTA*, 2000.

F. Rampazzo and R. B. Vinter, "Degenerate Optimal Control Problems with State Constraints", *SIAM J. Control Optim.*, 39, 2000.

F. H. Clarke, L. Rifford and R.J. Stern, "Feedback in State Constrained Optimal Control", *ESAIM: COCV*, 7, 2002.

P. Bettiol, A. Bressan and R. B. Vinter, "On trajectories satisfying a state constraint: $W^{1,1}$ estimates and counter-examples", *SIAM J. Control and Optimization*, Vol. 48, No. 7, 2010, pp. 4664–4679.

P. Bettiol, A. Bressan and R. B. Vinter, "Estimates for trajectories confined to a cone in $\mathbb{R}^n$", *SIAM J. Control and Optimization*, Vol. 49, No. 1, 2011, pp. 21-42.

P. Bettiol, A. Bressan and R. B. Vinter, "On trajectories satisfying a state constraint: $W^{1,1}$ estimates and counter-examples", *SIAM J. Control and Optimization*, Vol. 48, No. 7, 2010, pp. 4664–4679.

P. Bettiol, A. Bressan and R. B. Vinter, "Estimates for trajectories confined to a cone in $\mathbb{R}^n$", *SIAM J. Control and Optimization*, Vol. 49, No. 1, 2011, pp. 21-42.

A. Bressan and G.Facchi, "Trajectories of Differential Inclusions with State Constraints", JDE, 2011.

Application to Non-Degenerate Necessary Conditions

Take $(\bar{x}(.), \bar{u}(.))$ a minimizing process.

Necessary conditions yield Lagrange multiplier set $(p(.), \mu(.), \lambda)$ satisfying

costate eqn. + Weierstrass cond. + transversality cond. + . . .

When are the necessary conditions valid with $\lambda > 0$)?

Theorem

- inward pointing condition is satisfied
- $C = \mathbb{R}^n$

Then necessary conditions are valid with $\lambda > 0$.

(Extensive Russian literature (Arutyunov, Aseev), Rampazzo, Vinter)

Proof Based on Distance Estimates

Distance estimate is valid, since inward pointing condition is satisfied.

Step 1 From distance estimate and '$C = \mathbb{R}^n$':
$(\bar{x}(.), \bar{y}(.) \equiv h^+(x(.)), \bar{u}(.))$ also is minimizer for

$$\begin{cases} \text{Minimize } g(x(1)) + Ky(1) \\ \dot{x}(t) = f(t, x(t), u(t)),\ \dot{y}(t) = 0 \\ u(t) \in U \\ h(x(t)) \vee 0 - y(t) \leq 0 \\ x(0) = x_0 \end{cases}$$

Step 2 from transversality condition for $y(.)$ deduce

$$\int_0^1 d\mu(t) \leq K\lambda$$

Step 3 So, if $\lambda = 0$, $\mu = 0$. This implies $p(.) = 0$, Then

$$(p(.), \mu, \lambda) = 0 \text{ contradiction!}$$

Linear L^∞ Estimates

L^∞ linear estimates the most frequently applied.

$$||x(.) - \hat{x}(.)||_{L^\infty} \leq K \times h^+(\hat{x}(.)),$$

Write $F(t,x) = f(t,x,U)$ (velocity set)

Assume

- the inward pointing condition is satisfied
- A is a closed set

L^∞ linear estimates have be proved, when:

1. $t \rightsquigarrow F(t,x)$ is Lipschitz continous and A has smooth boundary *(Soner, '86)*
2. $t \rightsquigarrow F(t,x)$ is absolutely continous *(Bettiol, Frankowska, RBV, 2012)*
3. $t \rightsquigarrow F(t,x)$ has bounded variation *(Bettiol, RBV, 2016)*

Linear L^∞ Estimates, cont.

Regularity of $t \rightsquigarrow F(t,x)$ crucial for L^∞ linear distance estimates

A recent counter-example *(Bettiol + RBV, based on custruction of Bressan)*:

there exists $F(.,.)$ and closed set A such that

- $F(.,.)$ satisfying the inward pointing condition
- $F(.,x)$ is continuous

BUT: for any continuity modulus $\theta(.)$ and $K > 0$,

there exists a non-feasible trajectory $\hat{x}(.)$ such that

$$||x(.) - \hat{x}(.)||_{L^\infty} > K \times \theta(h^+(\hat{x}(.))),$$

for all feasible trajectories $x(.)$.

Fundamental discontinuity phenomena even for continuous $F(.,x)$.

$W^{1,1}$ Linear Estimates

Assume *standing hyps.* (allows meas. time dependence) and

- $F(., x)$ is $\mathcal{L}$ measurable
- uniform 'inward pointing' condition

Then

1. If A has smooth boundary:

$$||\hat{x}(.) - x(.)||_{W^{1,1}} \leq K \times h(\hat{x}(.))$$

   (Bettiol, Bressan, RBV 2010, Rampazzo, RBV 1990)

2. If A has non-smooth boundary:

   ($W^{1,1}$ linear estimate) is not in general valid

   (Counter-example: Bettiol, Bressan, RBV 2012)

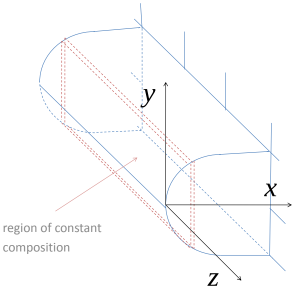
Example - beam design

The *objective*:

- *design a (cantilever) beam* with a smooth surface and having a constant cross-section in the direction of the z-axis.
- *maximize bending rigidity*

Composition of *two materials*:

- **A** is an *expensive material* which adds *stiffness* to the structure
- **B** is *less expensive* material to reduce the cost



A mathematical description of the problem

1) the cross-section of the beam orthogonal to the z axis is a parabola:

$$y = x^{1/2}, \ 0 \le x \le 1;$$

2) the free edge is located at $(x, y) = (0, 0)$;

3) $w(x) \in [0, 1]$ denotes the variation of the proportion of material A w.r.to x:

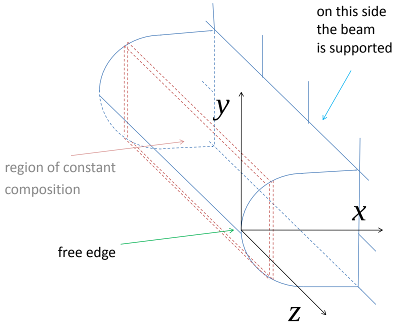
$$w(x) \ \in \ [0, 1] \qquad \text{for all } x \in [0, 1] \ ;$$

4) there is a bound V of the volume per unit length of material A in the beam (**isoperimetric constraint**):

$$\int_0^1 2w(x)|x|^{1/2} dx \le V \ ;$$

5) a restriction is placed on the rate of variation of the composition along the x axis:

$$|dw(x)/dx| \ \le \ k \qquad \text{for all } x \in [0, 1] \ .$$



The **cost function** will be a complicated function obtained by solving the 'beam equation'.

Example - beam design - optimal control

Set up as an **optimal control problem** in which x is a 'time-like variable' and $u(x) = dw(x)/dx$ is the **control**:

$$\begin{cases} \frac{dw}{dx}(x) = u(x) & \text{for a.e. } x \in [0,1] \\ u(x) \in [-k,k] & \text{for a.e. } x \in [0,1] \\ w(x) \in [0,1] & \text{for all } x \in [0,1] \quad \text{state constraint} \end{cases}$$

Replace the isoperimetric constraint with a differential equation for an **augmented state variable** $e(.)$ satisfying the differential equation

$$\frac{de}{dx} = 2w(x)\,|x|^{1/2}$$

Notice: the augmented dynamics above involve data exhibiting non-Lipschitz behavior w.r.to the time-like variable x.

Further Research of $W^{1,1}$ Linear Estimates

Assume *standing hyps.* (allows meas. time dependence) and

- $F(., x)$ is $\mathcal{L}$ measurable
- uniform 'inward pointing' condition

Special classes of $F(., .)$ and A non-smooth boundary have been identified such

1. $W^{1,1}$ linear distance estimates are valid

   or

2. $W^{1,1}$ superlinear distance estimates are valid

(e.g. with $\theta(h) = h \log h$ modulus)

(Bettiol. Bressan, RBV, Facchi)

Also:

- $W^{1,1}$ linear estimates are valid under a strengthened inward pointing condition

(Frankowska, Mazzola, 2013)

Concluding Remarks

Distance estimates have an important role in the derivation of optimality conditions for state constrained optimal control (first order necessary conditions and Hamilton Jacobi conditions).

- Linear (L^∞, $W^{1,1}$) estimates are valid for a smooth state constraint sets.
- *It is surprising that similar linear estimates are not valid in general, for non-smooth state constraint sets.*
- Under some assumptions, distance estimates can be established involving either a *linear* or a *superlinear* ($h|\log(h)|$) modulus.
- *Open questions*: what kind of estimates (linear, superlinear, Hölder?) are in general valid w.r.t. the $W^{1,1}$ or L^∞ norm, when there is a '*non smooth*' state constraint and the time dependence of the data is non-Lipschitz?