Minimizers that are not also Relaxed Minimizers

Richard B. Vinter

Imperial College London

SADCO Summer School, Sept. 2012

Outline of the Talk

- Relaxation
- The occurence of an infimum gap
- Condition for existence of an infimum gap
- Implications for necessary conditions of optimality
- Concluding Remarks

(Joint work with Michele Palladino, ESR)

The Optimal Control Problem

Consider

$$(P)\begin{cases} \text{Minimize } g(x(0),x(1)) \\ \text{over absolutely continuous functions } x(.) : [0,1] \to \mathbb{R}^n \\ \text{ satisfying} \\ \\ \dot{x}(t) \in F(t,x(t)) \text{ a.e.,} \\ x(0) = x_0 \quad \text{and} \quad (x(0),x(1)) \in C \,, \end{cases}$$

Data: $g : \mathbb{R}^n \to \mathbb{R}$**,** $x_0 \in \mathbb{R}$**, a closed set** $C \subset \mathbb{R}^n \times \mathbb{R}^n$ **and a multifunction** $F(.,.) : [0,1] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$**.**

State trajectory $x(.)$: $W^{1,1}$ function s.t. $\dot{x}(t) \in F(t,x(t))$, a.e., $x(0) = x_0$

$x(.)$ is feasible if $x(1) \in C$.

Minimizers

$\bar{x}(.)$ is a minimizer if

$$g(x(0), x(1)) \;\geq\; g((\bar{x}(0), \bar{x}(1)))$$

for all state trajectories $x(.)$.

$\bar{x}(.)$ is a strong local minimizer if, for some $\epsilon > 0$,

$$g(x(0), x(1)) \;\geq\; g((\bar{x}(0), \bar{x}(1)))$$

for all state trajectories $x(.)$ such that $||x(.) - \bar{x}(.)||_{L^\infty} \leq \epsilon$.

Standing Hypotheses

$(H1)$: $F(t,x)$ is closed for all (t,x) and $F(.,x)$ is measurable.

$(H2)$: For given $R>0$, there exist $\varepsilon>0$, $k(.)\in L^1$ and $c(.)\in L^1$ such that

$$F(t,x)\subset F(t,x')+k(t)|x-x'|\mathbb{B} \quad \text{and} \quad F(t,x)\subset c(t)\mathbb{B}$$

for all $x,x'\in R\mathbb{B}$, a.e. $t\in[0,1]$.

Fact: Assume (H1) and (H2). Suppose the set of feasible state trajectories is non-empty and bounded, and

(C): $F(t,x)$ is convex for all (t,x) .

Then there exists a minimizer.

Relaxation

Relaxation:

'Enlarge the space of state trajectories to guarantee existence of minimizers'

Relaxed Problem:

$$(R) \begin{cases} \text{Minimize } g(x(1)) \\ \text{over } x(.) : [0,1] \in W^{1,1} \text{ s.t.} \\ \dot{x}(t) \in \text{co}\, F(t, x(t)) \text{ a.e.,} \\ x(0) = x_0 \quad \text{and} \quad x(1)) \in C \ , \end{cases}$$

(Refer to **relaxed state trajectories, relaxed minimizers**, etc.).

Fact: Assume (H1) and (H2). Suppose the set of feasible relaxed state trajectories is non-empty and bounded.

Then there exists a relaxed minimizer.

(Relaxed problem automatically has a convex velocity set co $F(t, x)$).

Relaxation Theorem

Relaxation Theorem:

Take any relaxed state trajectory $\bar{x}(.)$ and $\epsilon > 0$.
Then there exist a state trajectory $x(.)$ such that

$$||x(.) - \bar{x}(.)||_{L^\infty} \leq \epsilon$$

Caution: $x(.)$ and $\bar{x}(.)$ close, but their velocities can be very different!

Strategy for finding sub-optimal state trajectory:

Step 1: Solve relaxed problem (it has a solution)

Step 2: Approximate relaxed minimizer by a neigboring state trajectory (possible by relaxation theorem)

Geometric Interpretation

Define **Reachable Set:**

$\mathcal{R} := \{(x(0), x(1)) \,|\, x(.) \text{ is a feasible state trajectory }\},$

$\mathcal{R}_{relaxed} := \{(x(0), x(1)) \,|\, x(.) \text{ is a feasible relaxed state trajectory }\}$

From Relaxation Thm:

$$\mathcal{R}_{relaxed} = \overline{\mathcal{R}}\,.$$

We have

$$\inf(P) = \inf\{g(x_0, x_1) \,|\, (x_0, x_1) \in \mathcal{R} \cap C\}\,,$$

$$\inf(R) = \inf\{g(x_0, x_1) \,|\, (x_0, x_1) \in \overline{\mathcal{R}} \cap C\}$$

The Infimum Gap

in general

$$\inf(R) \leq \inf(P).$$

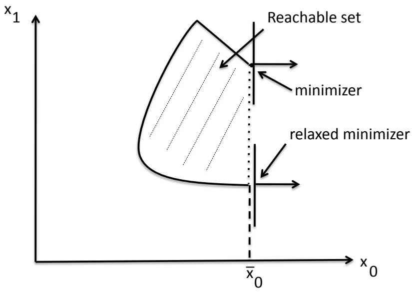
Important to identify situations when there is an infimum gap:

$$\inf(R) < \inf(P) \quad \text{(strict inequality!)}$$

because then

- Feasible trajectories cannot be closely approximated by feasible relaxed trajectories
- The dynamic programming methods yields only the relaxed infimum, not the 'true' infimum
- Numerical methods are often ill-conditioned when there is an infimum gap

The Infimum Gap



Occurence of an infimum gap

Link with Multiplier Abnormality

There is a relation between infimum gaps and abnormality of Lagrange multipliers.

Consider the finite dimensional optimization problem:

$$(P_1)\ \begin{cases} \text{Minimize } g(x) \\ \text{s.t. } \ h(x) < 0 \ \text{ and } \ x \in C \end{cases}$$

and its relaxation (replace $\{x|h(x) < 0\}$ by closed set $\{x|h(x) \leq 0\}$):

$$(R_1)\ \begin{cases} \text{Minimize } g(x) \\ \text{s.t. } \ h(x) \leq 0 \ \text{ and } \ x \in C \end{cases}$$

(Data: $h(.) : \mathbb{R}^n \to \mathbb{R}$, $g(.) : \mathbb{R}^n \to \mathbb{R}$, Lipschitz functions)

Finite Dimensional Case

Fact: Suppose there exists $\delta > 0$, $\epsilon > 0$ and $\bar{x}$ such that

$\bar{x}$ is feasible for (R_1) and

$$g(\bar{x}) \leq \inf\{g(x) \,|\, h(x) < 0 \text{ and } x \in C\}\,.$$

Then

$$0 \in 0\,.\,\partial g(\bar{x}) + \partial h(\bar{x}) + N_C(\bar{x})$$

i.e. infimum gap implies abnormal multiplier rule

Proof

'$\bar{x}$ has cost strictly less than infimum cost over unrelaxed points x' implies

(S): If $x_i \to \bar{x}$ and $h(x_i) < 0$ for all i, then

$$x_i \;\notin\; C \;\text{ for all } i \text{ sufficiently large.}$$

Take $\epsilon_i \to 0$ and consider:

$$(P_i) \text{ Minimize } \{J_i(x) \;:=\; (h(x+\epsilon_i))d_C(x) \,|\, x \in \mathbb{R}\} \;.$$

For each i, $\bar{x}$ is a minimizer. So by Ekeland's Theorem there exists x_i such that:

1) $\bar{x}$ minimizes $\to J_i(x) + \epsilon^{1/2}|x - x_i|$, and

2) $|x - x_i| \;\leq\; \epsilon^{1/2}$

But, by (S), $\quad$ '$h(x_i) + \epsilon_i \leq 0$' $\quad \Longrightarrow \quad$ '$d_C(x_i) > 0$'

for i sufficietly large. Use this to show

$$0 \;\in\; 0\,.\,\partial g(x_i) + \partial h(x_i) + N_C(\bar{x}) + \text{'error term'}$$

Pass to limit . .

Necessary Conditions of Optimality

It is possible to reformulate (P) as a generalized problem in the Calculus of Variations:

$$\begin{cases} \text{Minimize } \int_0^1 L(t, x(t), \dot{x}(t))dt + g(x(0), x(1)) \\ \text{over } x(.) \in W^{1,1} \text{ s.t. } (x(0), x(1)) \in C\,, \end{cases}$$

where

$$L(t, x, v) := \begin{cases} 0 & \text{if } v \in F(t, x) \\ +\infty & \text{if } v \notin F(t, x)\,. \end{cases}$$

Write Hamiltonian:

$$H(t, x, p) = \sup\{e \cdot p \,|\, e \in F(t, x)\}$$

Nonsmooth analysis approaches have validated generalization of the classical necessary conditions:

$(p(t), \dot{p}(t)) = \nabla_{x,p} L(t, x(t), p(t))$ (The Euler Lagrange condition)

and

$(-\dot{p}(t), \bar{x}(t)) = \nabla_{x,p} H(t, \bar{x}(t), p(t))$ (Hamilton's condition)

Theorem (Euler Lagrange Inclusion, Ioffe/Rockafellar 1996)

Let $\bar{x}$ be a strong local minimizer. Assume (H1) and (H2).

Then there exists $p(.) \in W^{1,1}$ and $\lambda \geq 0$ such that

$$(p(.), \lambda) \neq 0$$

$$\dot{p}(t) \in \text{co}\,\{q \,|\, (q, p(t)) \in N_{\text{Gr}F(t,.)}(\bar{x}(t)), \dot{\bar{x}}(t), \} \text{ a.e.}$$

$$p(t) \cdot \dot{\bar{x}}(t) \;=\; \sup\{p(t) \cdot e \,|\, e \in F(t, \bar{x})(t)\} \text{ a.e.}$$

$$(p(0), -p(1)) \;\in\; \lambda \partial g(\bar{x}(0), \bar{x}(1))) + N_C(\bar{x}(0), \bar{x}(1))\,.$$

Generalized Hamiltian Inclusion

Theorem (Generalized Hamiltonian Inclusion, Clarke 1973)

Let $\bar{x}$ be a strong local minimizer. Assume (H1) and (H2) and

$$F(t, x) \text{ is convex for all } (t, x)$$

Then there exists $p(.) \in W^{1,1}$ and $\lambda \geq 0$ such that

$$(p(.), \lambda) \neq 0$$

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\, \partial H(t, \bar{x}(t)), p(t)(t), \} \text{ a.e.}$$

$$(p(0), -p(1)) \in \lambda \partial g(\bar{x}(0), \bar{x}(1))) + N_C(\bar{x}(0), \bar{x}(1)) .$$

The Generalized Hamiltonian Inclusion condition is valid also for non-convex $F(., .)$'s (Clarke, 2005)

Partially Convexified Hamiltonian Inclusion

Theorem (Euler Lagrange Inclusion, Loewen/Rockafellar 1996)

Let $\bar{x}$ be a strong local minimizer. Assume (H1) and (H2), and

(C) $f(t,x)$ is convex

Then there exists $p(.) \in W^{1,1}$ and $\lambda \geq 0$ such that

$$(p(.),\lambda) \neq 0$$

$$-\dot{p}(t) \in \text{co}\,\{q\,|\,(q,\dot{\bar{x}}(t)) \in \partial_{x,p}H(t,\bar{x}(t),\dot{\bar{x}}(t))\} \ \text{ a.e.}$$

$$(p(0),-p(1)) \in \lambda\partial g(\bar{x}(0),\bar{x}(1))) + N_C(\bar{x}(0),\bar{x}(1))\,.$$

Open question: True without hypothesis (C)?

Conditions for an Infimum Gap (1)

Theorem A, Palladino/Vinter

Let $\bar{x}(.)$ be a strong local minimizer. Assume (H1) and (H2). Assume also that, for each $\epsilon > 0$, there exists a feasible relaxed trajectory $x(.)$ such that

$$g(x(0), x(1)) \; < \; g(\bar{x}(0), \bar{x}(1)) \text{ and } ||x(.) - \bar{x}(.)||_{L^\infty} \leq \epsilon$$

Then there exists $p(.) \in W^{1,1}$ and such that

$$p(.) \;\neq\; 0$$

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\, \partial H(t, \bar{x}(t)), p(t)(t), \} \text{ a.e.}$$

$$(p(0), -p(1)) \;\in\; 0 \,.\, \partial g(\bar{x}(0), \bar{x}(1)) + N_C(\bar{x}(0), \bar{x}(1)) \,.$$

"a strong local minimizer which is not also a strong local relaxed minimizer is an abnormal extremal (w.r.t. the Hamiltonian Inclusion)"

(Ioffe 1996 proved related theorem for a $W^{1,1}$ infimum gap - strongr hypotheses.)

Conditions for an Infimum Gap (2)

Theorem B, Palladino/Vinter

Let $\bar{x}(.)$ be a feasible relaxed trajectory. Assume (H1) and (H2).

Suppose that there exist $\delta > 0$ and $\epsilon > 0$ such that

$$g(\bar{x}(0), \bar{x}(1)) \; < \; \inf\{g(x(0), x(1)) \,|\, x(.) \text{ is feasible and } ||x(.) - \bar{x}(.)||_{L^\infty} \leq \epsilon$$

Then there exists $p(.) \in W^{1,1}$ and such that

$$p(.) \; \neq \; 0$$

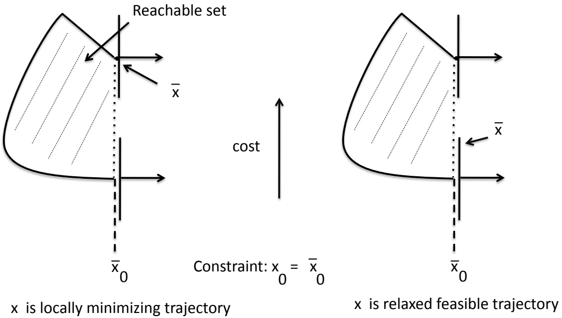
$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\, \partial H(t, \bar{x}(t)), p(t)(t), \} \;\text{ a.e.}$$

$$(p(0), -p(1)) \; \in \; 0 \,.\, \partial g(\bar{x}(0), \bar{x}(1)) + N_C(\bar{x}(0), \bar{x}(1)) \,.$$

"a feasible relaxed trajectory which has cost strictly less that that of any L^∞-neighbouring trajectory is an abnormal extremal (w.r.t. Hamiltonian Inclusion)"

(Warga 1972 proved a related theorem for controlled differential equations)

Type A and Type B Theorems



Type A Theorem

Type B Theorem

Concluding Comments/Open Questions

Links between abnormality and infimum gap forst investigated by Warga

Take a minimizer $\bar{x}$

Ioffe:

'$\bar{x}$ s not a $W^{1,1}$ local relaxed minimizer' $\Longrightarrow$ 'E-L conditions are abnormal'

Now we know:

'$\bar{x}$ s not a strong local relaxed minimizer' $\Longrightarrow$ 'Hamiltonian inclusion abnormal'

-Theory readily adapts to allow for state constraints

Open questions:

- Conditions for all multipliers to be abnormal item[-] Counter-examples distiguishing differ conditions
- Relaxation theorems with extra constraint $||x(.) - \bar{x}(.)||_{W^{1,1}} \leq \epsilon$