TERRY FEST 2015, Limoges 18-22 May 2015

Calm and Locally Upper Lipschitz Multifunctions: Intersection Mappings and Applications in Optimization

Diethard Klatte, University Zurich

Bernd Kummer, Humboldt University Berlin

Based on:

[KK15] D. Klatte, B. Kummer, On calmness of the argmin mapping in parametric optimization problems, *J. Optim. Theory Appl.* (2015) 165: 708-719.

[KK09] D. Klatte, B. Kummer, Optimization methods and stability of inclusions in Banach spaces, *Math. Program. Ser. B* 117 (2009) 305-330.

[KK02] D. Klatte, B. Kummer, Constrained minima and Lipschitzian penalties in metric spaces, *SIAM J. Optim.* 13 (2002) 619-633. See also D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization*, Kluwer 2002.

Contents:

1. Motivation

Let us start with the parametric optimization problem

$$f(x,t) \to \min_x \quad \text{s.t.} \quad x \in M(t)\,, \; t \text{ varies near } t^0, \tag{1}$$

where T is a normed linear space, $M : T \rightrightarrows \mathbb{R}^n$, $f : \mathbb{R}^n \times T \to \mathbb{R}$.

For (1), define the **infimum value function** φ by

$$\varphi(t) := \inf_x \{f(x,t) \mid x \in M(t)\}\,, \; t \in T$$

and the **argmin mapping** Ψ by

$$\Psi(t) \; := \; \operatorname*{argmin}_x \{f(x,t) \mid x \in M(t)\}\,, \; t \in T\,. \tag{2}$$

Let

$$\boldsymbol{(t^0, x^0) \in \operatorname{gph}\Psi} \text{ be a given reference point.}$$

Inspired by Cánovas et al '14, we give conditions for **calmness** of the

argmin mapping $\boxed{t \mapsto \Psi(t) = \{x \in M(t) \mid f(x,t) \le \varphi(t)\},}$

for t near t^0, by relating this to calmness of the **auxiliary mappings**

$$\begin{aligned} (t,\mu) &\mapsto L(t,\mu) &= \{x \in M(t) \mid f(x,t^0) \le \mu\}, \\ \mu &\mapsto L(t^0,\mu) &= \{x \in M(t^0) \mid f(x,t^0) \le \mu\}. \end{aligned} \tag{3}$$

If $M(t)$ is described by inequalities, then $L(t,\mu)$ is so, too, and moreover, $L(t^0,\mu)$ is given by inequalities perturbed only at the right-hand side.

Main purposes of the paper:

- To show under suitable conditions and for a large class of problems

$$L(t^0,\cdot) \text{ calm at } (\varphi(t^0),x^0) \;\Rightarrow\; \Psi \text{ calm at } (t^0,x^0),$$

- to recall an essential tool: calm intersections of multifunctions,
- to discuss some consequences for special parametric programs.

2. Concepts of upper Lipschitz (u.L.) continuity

Let (X, d_X), (T, d_T) be metric spaces, and $S : T \rightrightarrows X$ be a multifunction. Let $B(x^0, \varepsilon) := \{x \in X \,|\, d_X(x, x^0) \le \varepsilon\}$, similarly $B(t^0, \varepsilon)$.

Given $t^0 \in T$ and $x^0 \in S(t^0)$ or $\emptyset \ne X^0 \subset S(t^0)$,

S is called **calm** at (t^0, x^0) (with rank $L > 0$) if there is some $\varepsilon > 0$ such that for all $t \in B(t^0, \varepsilon)$,

$$x \in S(t) \cap B(x^0, \varepsilon) \quad \Rightarrow \quad \text{dist}(x, S(t^0)) \le L d_T(t, t^0)\,, \tag{4}$$

S is called **locally u.L.** at (t^0, X^0) (with rank $L > 0$) if there is some $\varepsilon > 0$ such that for all $t \in B(t^0, \varepsilon)$, with $V = B(X^0, \varepsilon)$,

$$x \in S(t) \cap V \quad \Rightarrow \quad \text{dist}(x, X^0) \le L d_T(t, t^0)\,. \tag{5}$$

In particular, S is also called **isolated calm** at (t^0, x^0) if $X^0 = \{x^0\}$, and S is called **upper Lipschitz** at t^0 if $X^0 = S(t^0)$ and $V = X$.

Remarks:

1. For calmness of inequality systems, many verifiable conditions are known, cf. e.g. Henrion-Outrata '05, Ioffe-Outrata '08, [KK'09], Gfrerer '11 and the references therein.

2. If $T = \mathbb{R}^m$, $X = \mathbb{R}^n$, and gph S is the union of finitely many convex polyhedral sets, then S is upper Lipschitz at each $t^0 \in T$ (with uniform rank) and hence calm on gph S. (Robinson '81)

3. Calmness is implied by the **Aubin property** of S at $(t^0, x^0) \in \text{gph}\, S$,

   $\exists L, \varepsilon\ \forall t, t' \in B(t^0, \varepsilon):\ \ x \in S(t) \cap B(x^0, \varepsilon) \Rightarrow \text{dist}(x, S(t')) \leq L d_T(t, t')\,.$

4. **Obviously,** if $x^0 \in X^0$, then

$$S \text{ locally u.L. at } (t^0, X^0) \Rightarrow S \text{ calm at } (t^0, x^0),$$

   since $X^0 \subset S(t^0)$, while the opposite implication is not true.

Characterization of locally u.L. behavior by describing functions

Let X, T be metric spaces, $S : T \rightrightarrows X$, $t^0 \in T$, and $\emptyset \neq X^0 \subset S(t^0)$. We call p **Lipschitzian increasing** near X^0 if $p \equiv 0$ on X^0 and

$$\exists c, \delta > 0 : \ p(x) \geq c \, \text{dist}(x, X^0) \text{ whenever } \text{dist}(x, X^0) < \delta. \qquad (6)$$

We say that p is a **describing function** for S near (t^0, X^0) if

$$S \text{ is locally u.L. at } (y^0, X^0) \ \Leftrightarrow \ p \text{ is Lipschitzian increasing near } X^0.$$

Examples of describing functions

(i) $p_S(x) = \text{dist}((t^0, x), \text{gph}\, S)$ $(\leq \text{dist}(x, X^0))$, cf. e.g. [KK02].

(ii) For $S(t_1, t_2) = \{x \in \mathbb{R}^n \mid g(x) \leq t_1,\ h(x) = t_2\}$ and locally Lipschitz $(g, h) : \mathbb{R}^n \to \mathbb{R}^{m+k}$ with $t^0 = (0, 0)$ and $X^0 = S(0, 0)$, a classical example is the locally Lipschitz function

$$p(x) = \|h(x)\| + \max_i\{0, g_i(x)\}.$$

Recall: Optimality conditions and exact penalization

Already in the 1970-1980ies, concepts of this type were used in various settings to derive optimality conditions or exact penalization schemes (by Ioffe, Rockafellar, Clarke, Robinson, Dolecki, Rolewicz, Burke, Mangasarian, Penot, Thibault and ...); for a survey of that time see Burke '91.

In our abstract framework, one has (cf. e.g. [KK02])

Proposition 1: Assume $f : X \to \mathbb{R}$ is Lipschitz around $x^0 \in X^0$ and S is locally u.L at (t^0, X^0), p is any describing fct for S near (t^0, X^0), or, alternatively, S is calm at (t^0, x^0), $p(x) = p_S(x)$ and $X^0 = S(t^0)$.

If x^0 is a local minimizer of f on X^0, then, provided that α is large enough, x^0 is a free local minimizer of

$$P(x) = f(x) + \alpha p(x).$$

3. Calmness of the argmin map via calm intersections

Consider again the parametric optimization problem (1),

$$f(x,t) \to \min_x \quad \text{s.t.} \quad x \in M(t)\,,\ t \text{ varies near } t^0,$$

and assume, with T is normed linear, Ψ = argmin mapping,

$$\begin{array}{l} M : T \rightrightarrows \mathbb{R}^n \text{ is closed multifunction,} \\ f \text{ is locally Lipschitz, and } (t^0, x^0) \in \operatorname{gph} \Psi \text{ is given.} \end{array} \tag{7}$$

As announced above, the argmin map will be related to the auxiliary map

$$L(t,\mu) = M(t) \cap \{x \mid f(x,t^0) \le \mu\} \quad \text{(intersection map)}.$$

Define for given $V \subset \mathbb{R}^n$,

$$\begin{array}{rcll} \Psi_V(t) & := & \operatorname{argmin}_x \{f(x,t) \mid x \in M(t) \cap V\}, & t \in T, \\ \varphi_V(t) & := & \inf_x \{f(x,t) \mid x \in M(t) \cap V\}. & t \in T. \end{array} \tag{8}$$

Theorem 1. [KK15, Thm. 3.1]

Consider the problem (1) under the assumptions (7). Suppose that

(i) the feasible set map M is calm at (t^0, x^0) and satisfies, for some $\varrho > 0$, $\text{dist}(x^0, M(t)) \leq \varrho \|t - t^0\|$ for t near t^0 *(Lipschitz l.s.c.)*.

(ii) $L(t, \mu) = \{x \in M(t) \mid f(x, t^0) \leq \mu\}$ is calm at $((t^0, \varphi(t^0)), x^0)$.

Then the argmin map Ψ is calm at (t^0, x^0).

Note.

Under Lipschitz l.s.c. of M, the proof of Thm. 3.1 in [KK15] can be modified to obtain similar statements for

$$M \text{ and } L \text{ locally u.L.} \Rightarrow \Psi \text{ locally u.L.,}$$

$$M \text{ and } L \text{ Hölder calm} \Rightarrow \Psi \text{ Hölder calm.}$$

The proof of Theorem 1

first gives that for some nbhd V of x^0 and t near t^0,

$$|\varphi_V(t) - \varphi_V(t^0)| \text{ has a Lipschitz estimate and } \Psi_V(t) \neq \emptyset$$

(using M Lipschitz l.s.c.). Further, one has

$$\Psi(t) \cap V \neq \emptyset \;\Rightarrow\; \Psi_V(t) = \Psi(t) \cap V \quad \boxed{\text{(hence, } \varphi_V(t)) = \varphi(t))}$$

for given $t \in T$ and $V \subset \mathbb{R}^n$, and one uses

$$\boxed{\Psi(t) = L(t, \mu(x,t))} \quad \text{with } \mu(x,t) := \varphi(t) + f(x,t^0) - f(x,t).$$

The rest is straightforward application of the assumptions.

The idea of proof combines standard tools from parametric optimization in the 1980ies, cf. e.g. Alt '83, Cornet '83, Robinson '83, Kl '84, '85.

Intersection Theorem (KK02, Thm. 3.6). Consider closed mappings $G : Y \rightrightarrows X$, $\Gamma : Z \rightrightarrows X$, X, Y, Z metric spaces, such that

- G, Γ and $z \mapsto G(y^0) \cap \Gamma(z)$ are calm at (y^0, x^0) resp. (z^0, x^0),
- Γ^{-1} has the Aubin property at (x^0, z^0),

then the intersection map $(y, z) \mapsto G(y) \cap \Gamma(z)$ is calm at $((y^0, z^0), x^0)$.

Applying this to the current setting (1) under (7),

$$f(x,t) \to \min_x \quad \text{s.t.} \quad x \in M(t)\,, \; t \text{ varies near } t^0,$$

we consider either $G = M$, $\Gamma = F$ or $G = F$, $\Gamma = M$ for

$$L(t,\mu) = M(t) \cap F(\mu), \quad \text{where } F(\mu) := \{x \mid f(x,t^0) \leq \mu\}.$$

Note: $F^{-1}(x)$ has the Aubin property since f is locally Lipschitz.

Apply the setting $G = M$, $\Gamma = F$ of the intersection theorem:

Theorem 2. [KK15]

Suppose the assumptions of Theorem 1, but replace the assumption

(ii) $L(t,\mu) = \{x \in M(t) \mid f(x,t^0) \le \mu\}$ is calm at $((t^0,\varphi(t^0)),x^0)$.

by the assumption that both

(ii)' the level set map $\boxed{F(\mu) := \{x \,|\, f(x,t^0) \le \mu\}}$ is calm at $(\varphi(t^0),x^0)$,

(ii)" and $\boxed{\mu \mapsto L(t^0,\mu) = M(t^0) \cap F(\mu)}$ is calm at $(\varphi(t^0),x^0)$,

Then the argmin map Ψ is calm at (t^0,x^0). *)

*) where (similarly in Theorem 1) $\Psi(t) \neq \emptyset$ for t near t^0 if x^0 is isolated calm.

Question: Is the opposite direction of Theorem 2 true under the Aubin poperty on M? **No!** "Ψ calm" does not imply "L calm".

Example 1: see [KK15]. Consider

$$\min\ y - c_1 x - c_2 y \quad \text{s.t. } x^2 - y \le b, \quad (c_1, c_2, b) \text{ close to } \underline{o} = (0,0,0).$$

Its argmin mapping Ψ is Lipschitz near $\underline{o}$, and hence calm at $(\underline{o}, (0,0))$:

$$\Psi(c_1, c_2, b) = \left\{ \left(\tfrac{c_1}{2(1-c_2)}, \tfrac{c_1^2}{4(1-c_2)^2} - b \right) \right\}.$$

However, $\boxed{L(0,\mu) = \{(x,y) \mid y \le \mu,\ x^2 \le y\}}$ is not calm at the origin.

Hence, the opposite direction of Theorem 2 (and Theorem 1) is not true even for a program with linear objective and convex quadratic constraint(s).

4. Specializations of Theorem 2

Model 1: Consider the standard parametric NLP

$$\min_x f(x,p,c) = h(x,p) + c^{\mathsf{T}} x \quad \text{s.t.} \quad x \in M(p,b),$$

$t = (p,c,b)$ varies near $t^0 = (p^0, c^0, b^0) \in T = \mathbb{R}^{q+n+m}$,

where

- $h, g_i : \mathbb{R}^n \to \mathbb{R} \quad \in C^1$
- $M(t) = M(p,b) = \{x \in \mathbb{R}^n \,|\, g_i(x,p) \le b_i\,,\ i = 1,\dots,m\}$,
- $F(\mu) = \{x \,|\, f(x,p^0,c^0) \le \mu\}$.

Verify assumptions of Theorem 2:

M is Lipschitz l.s.c. at $((p^0,b^0),x^0)$ $\Leftrightarrow$ MFCQ holds for $M(p^0,b^0)$ at x^0 (cf. e.g. [KK09]), this is equivalent to the Aubin prop. (Robinson '76).

Conditions for calmness of F and $M(t^0) \cap F$ (finite C^1 inequality system, RHS perturbations) are discussed e.g. in Henrion-Jourani-Outrata '02, Henrion-Outrata '05, Ioffe-Outrata '08, [KK09], Kummer '09, Gfrerer '11.

Model 2: Consider the canonically perturbed program

$$\min_x f(x,c) = h(x) + c^{\mathsf{T}} x \quad \text{s.t.} \quad g_i(x) \le b_i \ \ \forall i \in I,$$

$t = (c,b) \in \mathbb{R}^n \times C(I,\mathbb{R})$ varies near $t^0 = (c^0, b^0) \in \operatorname{gph} \Psi$, and

- I compact metric space (including finite I),
- $h, g_i : \mathbb{R}^n \to \mathbb{R}$ convex, $\in C^1$, $(i,x) \mapsto g_i(x)$ continuous,
- $M(b) = \{x | \, g_i(x) \le b_i \, \forall i \in I\}$, $F(\mu) = \{x \,|\, f(x,c^0) \le \mu\}$,
- the Slater CQ at $M(b^0)$ be satisfied.

For h, g_i linear, these are setting + assumptions in Cánovas et al '14. They prove in their special case: Theorem 2 even holds as "if-and-only-if".

Verify assumptions of Theorem 2:

Slater CQ $\Rightarrow$ M has Aubin property at (b^0, x^0), while criteria for calmness of F and $M(b^0) \cap F$ can be found for I finite e.g. in Li '97, Pang '97, Henrion-Jourani '02, Zheng-Ng '08, or including I infinite e.g. in Henrion-Outrata '05, [KK09].

5. Final remarks

1. The presented approach can be helpful also in determining the calmness modulus for argmin mappings. Recently, Cánovas, Kruger, López, Parra, Théra '14 demonstrates this for linear SIPs.

2. Calmness looks like a rather weak Lipschitz stability concept for the argmin mapping. However, it is useful as a kind of minimal requirement for the lower level in bi-level problems (CQ).

3. We have shown: Calmness of $L^0(\mu) = M(t^0) \cap \{x \,|\, f(x,t^0) \le \mu\}$ is essential for checking calmness of the argmin map Ψ. Note: If L^0 is calm at $(\varphi(t^0), x^0)$ **for each** $x^0 \in \Psi(t^0)$ (if $\Psi(t^0)$ is compact) then $\Psi(t^0)$ is a weak sharp minimizing set of the problem $f(x,t^0) \to \min_x$ s.t. $x \in M(t^0)$ (Henrion-Jourani-Outrata '02).

4. Our Theorem 1 also applies to complementarity or equilibrium constraints M. It would be of interest to see interrelations to recent calmness results for MPECs (including Gfrerer-Kl '15).

5. Concerning Hölder type calmness properties for inequality systems cf. e.g. Kummer '09, [KK09], Gfrerer '11, Kl-Kruger-Kummer '12.

6. The calm intersection theorem used in the proof of Theorem 2 is a powerful tool also in other situations, see recent papers by Henrion, Outrata, Surowiec and the authors.

Some closely related references

M.J. Cánovas, A. Hantoute, J. Parra, F.J. Toledo: Calmness of the argmin mapping in linear semi-infinite optimization. JOTA **160**, 111–126 (2014)

M.J. Cánovas, A.Y. Kruger, M.A. López, J. Parra, M. Théra: Calmness modulus of linear semi-infinite programs. SIOPT **24**, 29–48 (2014)

R. Henrion, A. Jourani. J. Outrata: On the calmness of a class of multifunctions. SIOPT **13**, 603–618 (2002)

R. Henrion, J. Outrata: Calmness of constraint systems with applications. Math. Progr. B **104**, 437–464 (2005)