TU WIEN SWM ORCOS Operations Research and Control Systems

Regularity and approximations of generalized equations; applications in optimal control

Vladimir M. Veliov

(Based on joint works with A. Dontchev, M. Krastanov, J. Preininger, T. Rockafellar, T. Scarinci, P. Vuong)

Wirtschafts Universität, Wien

Vienna, Nov. 25, 2017

About ORCOS

4 permanent TU faculty members + 1 emeritus (G. Feichtinger) + 4-8 project assistants

Main research direction:

Optimization and estimation of distributed (heterogeneous) systems – 15%

Numerical methods for optimal control – 20%

Variational inequalities, stability and computation of equilibria – 20%

Model predictive control – 15 %

Optimal control of ODE-systems and applications – 15%

Energy markets, smart grids, utilization of batteries – 15%

More information: ORCOS website: https://orcos.tuwien.ac.at/home/

Plan of the talk

1. “Coercive” problems.
2. “Affine” problems.

Generalized equations

$$0 \in G(x),$$

where $G : X \rightrightarrows Y$, X, Y – metric (Banach) spaces.

Examples:

1. For $X = I\!R^n$, $K \subset X$ – closed, $f : X \to I\!R$ – Fréchet-differentiable

$$\min_{x \in K} f(x) \quad \longrightarrow \quad 0 \in \nabla f(x) + N_K(x).$$

2. Robinson (1980): $0 \in f(x) + F(x)$, with $F(x)$ – set-valued mapping.

3. Differential variational inequalities (e.g. Pang and Steward, 2008):

$$\begin{aligned} \dot{x}(t) &= g(x(t), u(t)), \\ 0 &\in h(x(t), u(t)) + N_K(u(t)), \\ 0 &= \Gamma(x(0), x(T)). \end{aligned}$$

$x : [0, T] \to I\!R^n$, $u[0, T] \to I\!R^m$.

$$\text{minimize} \int_0^T l(y(t), u(t))\,\mathrm{d}t$$

$$\dot{y}(t) = g(y(t), u(t)), \quad y(0) = y_0, \quad u(t) \in U \quad t \in [0, T].$$

Hamiltonian: $H(y, p, u) = l(y, u) + p^T g(y, u)$

Optimality conditions:

$$\begin{cases} \dot{y}(t) & = \partial_p H(y(t), p(t), u(t)), & y(0) = y_0, \\ \dot{p}(t) & = -\partial_y H(y(t), p(t), u(t)), & p(T) = 0, \\ 0 & \in \partial_u H(y(t), p(t), u(t)) + N_U(u(t)), & \end{cases}$$

Usual spaces: $u \in L^\infty([0, T]; \mathbb{R}^m)$, $x = (y, p) \in {W_0}^{1,\infty}([0, T]; \mathbb{R}^{2n})$.

Reformulation: Differential Generalized Equation (DGE):

$$\begin{aligned} \dot{x} &= g(x, u), \\ 0 &\in f(x, u) + F(u), \end{aligned}$$

Differential Generalized Equation (DGE):

$$u \in L^\infty([0,T];\mathbb{R}^m), \qquad x = (y,p) \in {W_0}^{1,\infty}([0,T];\mathbb{R}^{2n}).$$

$$\begin{aligned} \dot{x} &= g(x,u), \\ 0 &\in f(x,u) + F(u), \end{aligned}$$

where

$$f(x,u) = \partial_u H(y,p,u), \qquad F(u) = N_{\mathcal{U}}(u),$$

with $\mathcal{U} = \{u \in L^\infty : \ u(t) \in U\}$, and for $u \in L^\infty$

$$N_{\mathcal{U}}(u) = \{w \in L^\infty \mid w(t) \in N_U(u(t)) \text{ for a.e. } t \in [0,T]\}.$$

$N_{\mathcal{U}}(u)$ is not the normal cone to $\mathcal{U}$!

$$f(x,u)(t) = f(x(t),u(t)), \quad F(u)(t) = F(u(t)).$$

A concept of (Lipschitz) regularity

$G : X \rightrightarrows Y$, X, Y – metric spaces.

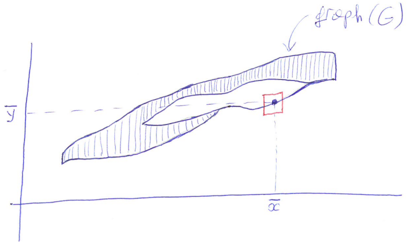
Definition. G is *strongly metrically regular* (SMR) at $\bar{x}$ for $\bar{y} \in G(\bar{x})$ if there are balls $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$, $a, b > 0$ such that the mapping

$$\mathbb{B}_b(\bar{y}) \ni y \to G^{-1}(y) \cap \mathbb{B}_a(\bar{x})$$

is single-valued and Lipschitz continuous (with Lipschitz constant κ).

Here $G^{-1}(y) := \{x : \ y \in G(x)\}$.

SMR means that G^{-1} has a *Lipschitz localization:*



The weaker property of "metric regularity" will not be discussed here.

A Ljusternik-Graves-type theorem (e.g. Dontchev and Rockafellar - 2013)

Theorem

Let a, b, and κ be positive scalars such that G is strongly metrically regular at $\bar{x}$ for $\bar{y}$ with neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$ and constant κ. Let $\mu > 0$ be such that $\kappa\mu < 1$ and let $\kappa' > \kappa/(1-\kappa\mu)$. Then for every positive α and β such that

$$\alpha \leq a/2, \quad 2\mu\alpha + 2\beta \leq b \quad \text{and} \quad 2\kappa'\beta \leq \alpha$$

and for every function $\gamma : X \to Y$ satisfying

$$\|\gamma(\bar{x})\| \leq \beta \;\; \text{and} \;\; \|\gamma(x) - \gamma(x')\| \leq \mu\|x - x'\| \quad \forall\, x, x' \in \mathbb{B}_{2\alpha}(\bar{x}),$$

the mapping $y \mapsto (\gamma + G)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$ is a Lipschitz continuous function on $\mathbb{B}_\beta(\bar{y})$ with Lipschitz constant κ'. (Hence $\gamma + G$ is SMR at $\bar{x}$ for $\bar{y}$.)

Qualitative consequences in the case of DGE

R. Cibulka, A. Dontchev, M. Krastanov, V.V., SIAM J. Contr. Opt., (2017(8))

Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be a solution of the DGE

$$\begin{aligned} \dot{x}(t) &= g(x(t), u(t)), \\ 0 &\in f(x(t), u) + F(u(t)). \end{aligned}$$

Assumption (*): $\forall\, (t, u) \in \operatorname{cl} \operatorname{gr} \bar{u}$ the mapping

$$I\!R^m \ni v \mapsto \mathcal{W}_{t,u}(v) := f(\bar{x}(t), u) + \partial_u f(\bar{x}(t), u)(v - u) + F(v)$$

is SMR at u for 0.

Theorem

$\exists a, b, \kappa > 0$: $\forall (t, u) \in \operatorname{cl} \operatorname{gr} \bar{u}$ the mapping $\mathcal{W}_{t,u}(\cdot)$ is SMR at u for 0 with parameters a, b, κ. That is, the mapping $I\!B_b(0) \ni z \mapsto \mathcal{W}^{-1}_{t,u}(z) \cap I\!B_a(u)$ is single-valued and Lipschitz with constant κ.

Theorem

If Assumption (*) is fulfilled then the mapping

$$(x, u) \mapsto \begin{pmatrix} \dot{x} - g(x, u) \\ f(x, u) \end{pmatrix} + \begin{pmatrix} 0 \\ F(u) \end{pmatrix}$$

is SMR at $(\hat{x}, \hat{u})$ *for* 0.

Recall: $u \in L^\infty([0, T]; \mathbb{R}^m), \quad x = (y, p) \in {W_0}^{1,\infty}([0, T]; \mathbb{R}^{2n})$

Other consequences:

Conditions for Lipschitz continuity of $\bar{u}$ ...

Convergence of discrete approximations and “path-following” methods ...
(more detailed analysis in

A. Dontchev, M. Krastanov, R.T. Rockafellar, V.V., SIAM J. Contr. Optim., 2013.)

Extensions for non-differentiable Lipschitz functions f (in terms of the strict prederivative of f): R. Cibulka, A. Dontchev, V.V., SIAM J. Contr. Optim., 2016.

Newton-type methods

R. Cibulka, A. Dontchev, J. Preininger, T. Roubdal, V.V., Journal of Convex Analysis (2018)
X and Y – Banach spaces. Consider the equation $f(x) = 0$, $f : X \to Y$ with a Fréchet-differentiable f.

Newton method: Generate $\{x_k\}$ such that

$$f(x_k) + \partial f(x_k)(x_{k+1} - x_k) = 0, \; x_0 \text{ – given.}$$

Assumption for (quadratic) convergence: a solution $\bar{x}$ exists, $\partial f(\bar{x})$ is invertible, and $\|x_0 - \hat{x}\|$ is small enough.

Kantorovich version: two differences:
(i) the invertibility assumption is posed for $\partial f(x_0)$, some "checkable" assumptions are posed. Then: a solution $\bar{x}$ exists and the convergence is quadratic.

(ii) One can modify the iterations as

$$f(x_k) + \partial f(x_0)(x_{k+1} - x_k) = 0, \; x_0 \text{ – given.}$$

Then the convergence is linear: $\|x_k - \hat{x}\| \leq \alpha^k \|x_0 - \hat{x}\|$, $\alpha \in (0, 1)$.

Further extensions:

– Bartle (1955): $f(x_k) + \partial f(z_k)(x_{k+1} - x_k) = 0$, x_0 – given. Any z_k – ...

– Qi and Sun (1993): f can be only Lipschitz; take $A_k \in \hat{\partial} f(x_k)$ – the Clarke generalized Jacobian ...

Our problem: $0 \in f(x) + F(x)$, where $f : X \to Y$, $F : X \rightrightarrows Y$, X, Y – Banach spaces.

Newton-Kantorovich iterations:

$$f(x_k) + A_k(x_{k+1} - x_k) + F(x_{k+1}) \ni 0,$$

where $A_k = A_k(x_0, \ldots, x_k) \in \mathcal{L}(X, Y)$, together with some $y_0 \in f(x_0) + F(x_0)$ have the following properties:

(i) for very k the mapping

$$x \mapsto f(x_0) + A_k(x - x_0) + F(x)$$

is SMR at x_0 for y_0 with a constant κ and neighborhoods $\mathbb{B}_a(x_0)$, $\mathbb{B}_b(y_0)$;

(ii) $\|f(x) - f(x_k) - A_k(x - x_k)\| \leq \omega(\|x - x_k\|) \, \|x - x_k\| \quad \forall \, x \in \mathbb{B}_a(x_0)$, where $\omega : [0, a] \to [0, \delta]$, $\delta > 0$.

Theorem

Assume that $\kappa\delta < 1$ *and* $\|y_0\| < (1-\kappa\delta)\min\{\frac{a}{\kappa}, b\}$. *Then the Newton-Kantorovich method generates a unique sequence in* $\mathbb{B}_a(x_0)$, *and it* *linearly* *converges to a solution* $\bar{x}$:

$$\|x_k - \bar{x}\| < (\kappa\delta)^k a. \tag{1}$$

If $\lim_{\xi\to 0}\omega(\xi) = 0$, *then the sequence* $\{x_k\}$ *is* *superlinearly* *convergent: there exist sequences of positive numbers* $\{\varepsilon_k\}$ *and* $\{\eta_k\}$ *such that* $\|x_k - \bar{x}\| \le \varepsilon_k$ *and* $\varepsilon_{k+1} \le \eta_k\varepsilon_k$ *for all sufficiently large* k, *and* $\eta_k \to 0$.

If there exists a constant $L > 0$ *such that* $\omega(\xi) \le \min\{\delta, L\xi\}$ *for each* $\xi \in [0, a]$, *then the convergence of* $\{x_k\}$ *is* *quadratic*: *there exists a sequence of positive numbers* $\{\varepsilon_k\}$ *such that* $\|x_k - \bar{x}\| \le \varepsilon_k$ *and* $\varepsilon_{k+1} \le \frac{\alpha L}{\delta}\varepsilon_k^2$ *for all sufficiently large* k.

Theorem

Assume that $\kappa\delta < 1$ *and* $\|y_0\| < (1-\kappa\delta)\min\{\frac{a}{\kappa}, b\}$*.*
Then the Newton-Kantorovich method generates a unique sequence in $\mathbb{B}_a(x_0)$*, and it* *linearly* *converges to a solution* $\bar{x}$*:*

$$\|x_k - \bar{x}\| < (\kappa\delta)^k a. \tag{1}$$

If $\lim_{\xi\to 0}\omega(\xi) = 0$*, then the sequence* $\{x_k\}$ *is* *superlinearly* *convergent: there exist sequences of positive numbers* $\{\varepsilon_k\}$ *and* $\{\eta_k\}$ *such that* $\|x_k - \bar{x}\| \le \varepsilon_k$ *and* $\varepsilon_{k+1} \le \eta_k\varepsilon_k$ *for all sufficiently large* k*, and* $\eta_k \to 0$*.*

If there exists a constant $L > 0$ *such that* $\omega(\xi) \le \min\{\delta, L\xi\}$ *for each* $\xi \in [0, a]$*, then the convergence of* $\{x_k\}$ *is* *quadratic**: there exists a sequence of positive numbers* $\{\varepsilon_k\}$ *such that* $\|x_k - \bar{x}\| \le \varepsilon_k$ *and* $\varepsilon_{k+1} \le \frac{\alpha L}{\delta}\varepsilon_k^2$ *for all sufficiently large* k*.*

Special cases: $A_k = \partial f(x_0)$ – Kantorovich
$A_k = \partial f(x_k)$ – Newton
Other choices of A_k – extended Bartle.

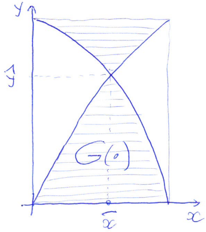
Strong Metric Sub-Regularity (SMs-R)

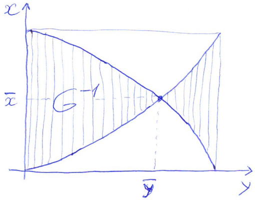
(Cibulka, Dontchev, Kruger (2017(8)))

$G : X \rightrightarrows Y$, X, Y – metric spaces.

Definition. G is *strongly metrically sub-regular* (SMs-R) at $\bar{x}$ for $\bar{y} \in G(\bar{x})$ if there are $\kappa > 0$ and balls $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$, $a, b > 0$, such that

$$G^{-1}(y) \cap \mathbb{B}_a(\bar{x}) \subset \mathbb{B}_{\kappa\, \mathrm{dist}(y,\bar{y})}(\bar{x}) \quad \forall\, y \in \mathbb{B}_b(\bar{y}).$$





This property is enough for many contexts: error analysis of approximations; Newton method.

Newton method for $0 \in f(x) + F(x)$, where $f : X \to Y$, $F : X \rightrightarrows Y$, X, Y – Banach spaces, f has Lipschith Fréchet derivative.

Newton iterations:

$$f(x_k) + \partial f(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0.$$

Theorem

Assume that linearized mapping $x \to f(\bar{x}) + \partial f(\bar{x})(x - \bar{x}) + F(x)$ is SMs-R at $\bar{x}$ for 0. Then there exists a neighborhood O of $\bar{x}$ such that if a sequence $\{x_k\}$ generated by the Newton method has a tail in O, then x_k is quadratically convergent to $\bar{x}$.

Existence of such a Newton sequence is not granted!

IMPORTANT: When the general results involving SMR or SMs-R are used for

$$\text{minimize} \int_0^T l(y(t), u(t))\, \mathrm{d}$$

$$\dot{y}(t) = g(y(t), u(t)), \quad y(0) = y_0, \quad u(t) \in U \quad t \in [0, T],$$

hence for the optimality conditions

$$\begin{cases} \dot{y}(t) & = \partial_p H(y(t), p(t), u(t)), & y(0) = y_0, \\ \dot{p}(t) & = -\partial_y H(y(t), p(t), u(t)), & p(T) = 0, \\ 0 & \in \partial_u H(y(t), p(t), u(t)) + N_U(u(t)), & \end{cases}$$

the space specifications are always $u \in L^\infty([0, T]; \mathbb{R}^m)$,

$x = (y, p) \in {W_0}^{1,\infty}([0, T]; \mathbb{R}^{2n})$.

The conditions for SMR and SMs-R involve coercivity!

This spaces are not appropriate for problems with discontinuous optimal controls.

Affine problems

$$\min \left\{ \int_0^T [g_0(x(t)) + g(x(t))u(t))] \, \mathrm{d}t \, + \, \Phi(x(T)) \right\}.$$

$$\dot{x} = f_0(x) + u\, f(x), \quad x(0) \; \text{– given}, \qquad u(t) \in U = [0,1].$$

Optimality system:

$$\begin{aligned} 0 &= \dot{x} - f_0(x) + u\, f(x), \\ 0 &= \dot{p} + p^T \partial_x (f_0(x) + u\, f(x)) + \partial_x (g_0(x) + u\, g(x)), \\ 0 &\in g(x(t)) + p(t)^\top f(x(t)) + N_U(u(t)), \\ 0 &= p(T) - \partial \Phi(x(T)). \end{aligned}$$

What are the appropriate spaces?
Under what conditions we have SMR or SMs-R?
Can we apply the Newton method?

Consider the linearized problem:

$$\begin{array}{ll} \text{minimize} & J(x,u) \\ \text{subject to} & \dot{x}(t) = A(t)x(t) + B(t)u(t) + d(t), \quad x(0) = x_0, \\ & u(t) \in U := [-1,1], \end{array}$$

where

$$J(x,u) := \Phi(x(T)) + \int_0^T \left(\frac{1}{2} x(t)^\top W(t)x(t) + x(t)^\top S(t)u(t) \right) \mathrm{d}t.$$

Optimality system:

$$0 \in G(x,p,u) := \begin{pmatrix} \dot{x} - Ax - Bu - d \\ \dot{p} + A^\top p + Wx + Su \\ B^\top p + S^\top x + N_{\mathcal{U}}(u) \\ p(T) - \partial\Phi(x(T)) \end{pmatrix},$$

$$N_{\mathcal{U}}(u) = \{ w \in L^\infty \mid w(t) \in N_U(u(t)), \;\; t \in [0,T] \}.$$

Spaces:

$$\mathcal{X} := W^{1,1}_{x_0} \times W^{1,1} \times L^1, \qquad \mathcal{Y} := L^1 \times L^1 \times L^\infty \times \mathbf{R}^n$$

Sufficient conditions for SMs-R (J. Preininger, T. Scarinci, V.V., 2017(?))

(A1) Continuous differentiability of the data; $W(t)$ symmetric; Φ – differentiable with Lipschitz derivative.

(A2) The functional $J(x, u)$ is convex on the set of admissible control-trajectory pairs.

(A3) For a given reference solution $(\hat{x}, \hat{p}, \hat{u})$ there are numbers $\alpha, \tau > 0$ such that at every zero s of the function

$$H_u(\hat{x}(t), \hat{p}(t), \hat{u}(t)) = \hat{\sigma}(t) = B(t)^\top \hat{p}(t) + S(t)^\top \hat{x}(t)$$

it holds that

$$|\hat{\sigma}(t)| \geq \alpha |t - s| \quad \forall\, t \in [s - \tau, s + \tau] \cap [0, T].$$

Theorem

$\exists c > 0$ such that $\forall y \in \mathcal{Y}$ there exists a solution $(x, p, u) \in \mathcal{X}$ of $y \in G(x, p, u)$ and for every such (x, p, u)

$$\|x - \hat{x}\|_{1,1} + \|p - \hat{p}\|_{1,1} + \|u - \hat{u}\|_1 \leq c\|y\|.$$

(W. Alt, U.Felgenhauer, M. Seidenschwanz, 2016-17)

A (surprising) consequence: (J. Preininger, T. Scarinci, V.V., 2017)

Theorem

Under conditions a bit stronger than (A1)–(A3) for the linearized problem at the solution point $(\hat{x}, \hat{p}, \hat{u})$*, the sequence of any Newton iterates starting from any initial point* (x_0, p_0, u_0) *sufficiently close to* $(\hat{x}, \hat{p}, \hat{u})$ *converges quadratically to* $(\hat{x}, \hat{p}, \hat{u})$*.*

A similar theorem under a number of more restrictive conditions - in [U.Felgenhauer (2017)].

A (surprising) consequence: (J. Preininger, T. Scarinci, V.V., 2017)

Theorem

Under conditions a bit stronger than (A1)–(A3) for the linearized problem at the solution point $(\hat{x}, \hat{p}, \hat{u})$, the sequence of any Newton iterates starting from any initial point (x_0, p_0, u_0) sufficiently close to $(\hat{x}, \hat{p}, \hat{u})$ converges quadratically to $(\hat{x}, \hat{p}, \hat{u})$.

A similar theorem under a number of more restrictive conditions - in [U.Felgenhauer (2017)].

A numerical problem: how to solve the linear-quadratic problem

$$\text{minimize } \Phi(x(T)) + \int_0^T \left(\frac{1}{2} x(t)^\top W(t) x(t) + x(t)^\top S(t) u(t) \right) \mathrm{d}t,$$

$$\text{subject to} \quad \begin{array}{l} \dot{x}(t) = A(t)x(t) + B(t)u(t) + d(t), \quad x(0) = x_0, \\ u(t) \in U := [-1, 1], \quad \text{or } U := \{-1, 1\} \end{array}$$

A new discretization scheme

V.V., 1989

A. Pietrus, T. Scarinci, V.V. (SIAM J. CO, 2017(8))

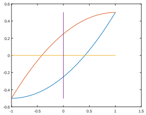
T. Scarinci and V.V. (Comput. Optim. and Appl., 2017)

Basic idea: $\{t_i\}_{i=0}^N$ a mesh with step h on $[\tau, T]$. Consider $w_i = (u_i, v_i)$,

$$u_i = \frac{1}{h}\int_{t_i}^{t_{i+1}} u(t)\,\mathrm{d}t, \qquad v_i = \frac{1}{h^2}\int_{t_i}^{t_{i+1}} (t - t_i)u(t)\,\mathrm{d}t$$

as discrete controls associated with $u(t) \in \{0, 1\}$. When $u(t) \in \{0, 1\}$ or $u(t) \in [0, 1]$, it holds that for $w_i = (u_i, v_i)$

$$w_i \in Z := \text{Aumann-}\int_0^1 \begin{pmatrix} 1 \\ s \end{pmatrix} [-1, 1] ds.$$



Explicit representation:

$$Z = \{(\alpha, \beta) : \alpha \in [-1, 1],\ \beta \in [\varphi_1(\alpha), \varphi_2(\alpha)]\},$$

where $\varphi_1(\alpha) := \frac{1}{4}\left(-1 + 2\alpha + \alpha^2\right)$ and $\varphi_2(\alpha) := \frac{1}{4}\left(1 + 2\alpha - \alpha^2\right)$.

Conversely, there is a mapping $\Phi^h : Z^N \to \{0,1\}$ such that $\forall w := (w_0, \dots, w_{N-1}) = ((u_0, v_0), \dots, (u_{N-1}, v_{N-1})) \in Z^N$

$$u_i = \frac{1}{h} \int_{t_i}^{t_{i+1}} \Phi^h(w)(t)\,\mathrm{d}t, \qquad v_i = \frac{1}{h^2} \int_{t_i}^{t_{i+1}} (t - t_i)\Phi^h(w)(t)\,\mathrm{d}t.$$

$\Phi^h(w)(t) \in \{0,1\}$ has 0, 1 or at most 2 jumps in every interval $[t_i, t_{i+1}]$.

Then we use the 2nd order Volterra-Fliess series to approximate the dynamics and the objective functional.

Under (A1)–(A3), for any solution w^h of the discrete problem it holds that $\|\Phi^h(w^h) - \hat{u}\|_1 \leq ch^2$.

Second order accuracy cannot be provided by any Runge-Kutta scheme! Schemes with second order accuracy (and still “nice” discretized problem) were not known so far.

Next numerical problem: How to solve the resulting mathematical programming problem?

The discretized problem has the general form

$$\min_{w \in K} f(w),$$

where f is a linear-quadratic function (not necessarily convex) and K is strongly convex.

The paper [V.V., P. Vuong, 2018(?)] presents linear convergence results for the GPM and the CGM for such problems in Hilbert spaces.

More specialized methods taking into account the structure of the constraints:

$$K = Z \times Z \ldots \times Z$$

and of the objective function – future work.

Strong Metric Regularity of affine problems

$$\min \left\{ \int_0^T [g_0(x(t)) + g(x(t))u(t)] \,\mathrm{d}t \;+\; \Phi(x(T)) \right\}.$$

$$\dot{x} = f_0(x) + u\, f(x), \quad x(0) \text{ – given}, \quad u(t) \in U = [0,1].$$

Linearized optimality system:

$$0 \in G(x,p,u) := \begin{pmatrix} \dot{x} - Ax - Bu - d \\ \dot{p} + A^\top p + Wx + Su \\ B^\top p + S^\top x + N_{\mathcal{U}}(u) \\ p(T) - \partial\Phi(x(T)) \end{pmatrix},$$

$$N_{\mathcal{U}}(u) = \{ w \in L^\infty \mid w(t) \in N_U(u(t)), \;\; t \in [0,T] \}.$$

SMR in the spaces

$$\mathcal{X} := W^{1,1}_{x_0} \times W^{1,1} \times L^1, \qquad \mathcal{Y} := L^1 \times L^1 \times L^\infty \times \mathbf{R}^n$$

"never" holds!!!

Strong bi-Metric Regularity of affine problems (Sbi-MR

General: $G : X \rightrightarrows Y$, $\quad X, Y$ – metric spaces with metric d_X and d_Y.

Definition. G is *strongly metrically regular* (SMR) at $\bar{x}$ for $\bar{y} \in G(\bar{x})$ if there are balls $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$, $a, b > 0$ such that the mapping

$$\mathbb{B}_b(\bar{y}) \ni y \to G^{-1}(y) \cap \mathbb{B}_a(\bar{x})$$

is single-valued and Lipschitz continuous (with Lipschitz constant κ):

$$d_X(G^{-1}(y) \cap \mathbb{B}_a(\bar{x}), G^{-1}(y') \cap \mathbb{B}_a(\bar{x})) \leq \kappa d_Y(y, y') \quad \forall y, y' \in \mathbb{B}_b(\bar{y}).$$

The bi-metric modification:

M. Quincampoix and V.V., SIAM J. CO (2013)

J. Preininger, T. Scarinci, and V.V., (2018)(??)

Explanation for the two metrics

Consider $\dim(u) = 1$, $U = [-1, 1]$, $\hat{\sigma}(t) = -\frac{1}{2} + t$, $t \in [0, 1]$.

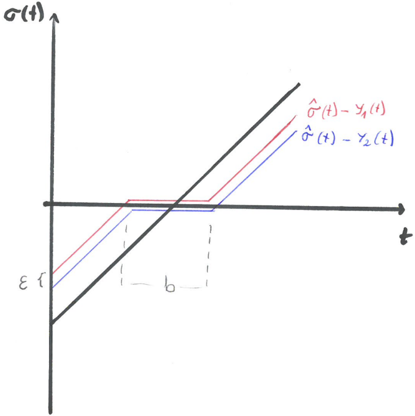
The solution of $y(t) \in \hat{\sigma}(t) + N_U(u(t))$ is

$u(t) = u[y](t) :=$ sgn$(\hat{\sigma}(t) - y(t))$ whenever $\hat{\sigma}(t) - y(t) \neq 0$,
$\hat{u}(t) = u[0](t)$.

When do we have (for some κ and $b > 0$)

$$\|u[y_1] - u[y_2]\|_1 \leq \kappa d_Y(y_1, y_2) \qquad \forall y_1,\ y_2 :\ d_Y(y_i, 0) \leq b.$$

What is the metric space $Y \subset L^\infty$?



Here

$$\|u[y_1] - u[y_2]\|_1 \approx 2b \gg \kappa\varepsilon = \kappa\|y_1 - y_2\|_\infty,$$

thus with $Y = L^\infty$ the mapping $u \to \hat\sigma + N_U(u)$ is not SMR at $\hat u$ for 0!

However, for $y_1,\, y_2 \in Y = W^{1,\infty}$ we have

$$\|u[y_1] - u[y_2]\|_1 \leq \frac{8}{3}\|y_1 - y_2\|_{1,\infty}$$

whenever $\|y_i\|_{1,\infty} \leq b := \frac{1}{4}$.

However, for $y_1, y_2 \in Y = W^{1,\infty}$ we have

$$\|u[y_1] - u[y_2]\|_1 \le \frac{8}{3}\|y_1 - y_2\|_{1,\infty}$$

whenever $\|y_i\|_{1,\infty} \le b := \frac{1}{4}$.

Even more, for $y_1, y_2 \in W^{1,\infty}$

$$\|u[y_1] - u[y_2]\|_1 \le \frac{8}{3}\|y_1 - y_2\|_{\infty}.$$

Thus the Lipschitz property is with respect to the L^∞-norm for y, but the disturbances y should be close to the reference point $\hat{y} = 0$ in the larger norm of $W^{1,\infty}$.

This explains the necessity of using two norms for the disturbances.

(X, d_X), (Y, d_Y) and $(\widetilde{Y}, \widetilde{d}_Y)$ – metric spaces, with $\widetilde{Y} \subset Y$ and $d_Y \leq \widetilde{d}_Y$ on $\widetilde{Y}$.

Definition

The map $G : X \rightrightarrows Y$ is strongly bi-metrically regular (relative to $\widetilde{Y} \subset Y$) at $\bar{x} \in X$ for $\bar{y} \in \widetilde{Y}$ with constants $\varsigma \geq 0$, $a > 0$ and $b > 0$ if $(\bar{x}, \bar{y}) \in \mathrm{graph}(\Phi)$ and the following properties are fulfilled:

1. the mapping $B_{\widetilde{Y}}(\bar{y}; b) \ni y \mapsto G^{-1}(y) \cap B_X(\bar{x}; a)$ is single-valued
2. for all $y, y' \in B_{\widetilde{Y}}(\bar{y}; b)$,

$$d_X(G^{-1}(y) \cap B_X(\bar{x}; a), G^{-1}(y') \cap B_X(\bar{x}; a)) \leq \varsigma d_Y(y, y').$$

Lyusternik-Graves-type theorem (J. Preininger, T. Scarinci, V.V., 2017(?))

Theorem

Let the metric space X be complete, let Y be a subset of a linear space and let both metrices d_Y and $\widetilde{d}_Y$ in Y and $\widetilde{Y} \subset Y$, respectively, be shift-invariant. Let $G : X \rightrightarrows Y$ be strongly bi-metrically regular at $\bar{x}$ for $\bar{y}$ with constants κ, a, b. Let $\mu > 0$ and κ' be such that $\kappa\mu < 1$ and $\kappa' \geq \kappa/(1 - \kappa\mu)$. Then for every positive a', b', and γ such that

$$a' \leq a, \quad b' + \gamma \leq b, \quad \kappa b' \leq (1 - \kappa\mu)a',$$

and for every function $\varphi : X \to \widetilde{Y}$ such that

$$d_Y(g(\bar{x}), 0) \leq b', \qquad \widetilde{d}_Y(g(x), 0) \leq \gamma \qquad \forall x \in B_X(\bar{x}, a'),$$

and

$$d_Y(g(x), g(x')) \leq \mu\, d_X(x, x') \quad \forall x, x' \in B_X(\bar{x}, a'),$$

the mapping $B_{\widetilde{Y}}(\bar{y} + g(\bar{x}); b') \ni y \mapsto (g + G)^{-1}(y) \cap B_X(\bar{x}, a')$ is single-valued and Lipschitz continuous with constant κ' with respect to the metric d_Y. This implies strong bi-metric regularity of $g + G$ ...

X, Y, $\tilde{Y}$ – convex subsets of linear normed spaces, X – complete.

(to be) Theorem. (M. Quincampoix, T. Scarinci, V.V., 2018(?))
Let $f : X \to \tilde{Y}$ be Fréchet differentiable at $\bar{x}$ in the norm of $\tilde{Y}$, and be differentiable in a neighborhood of $\bar{x}$ in the norm of Y, with uniformly continuous (in Y) derivative. Then the mapping $G = f + F$ is strongly bi-metrically regular at $\bar{x}$ for $\bar{y}$ if and only if the mapping $x \mapsto f(\bar{x}) + \partial f(\bar{x})(x - \bar{x}) + F(x)$ is such.

Consequence: Sbi-MR of the affine differential variational inequality is equivalent to that of the linearized one. (A1)–(A3) are sufficient for that.

Conclusions

1. SMR and SMs-R are key concepts of Lipschitz stability: they are themselves stable, enable Newton-Kantorovich methods, analysis of approximations, etc.
2. for DGEs the concepts have been developed and applied in the "coercive" case
3. for "affine" DGEs – recent developments: Newton, new discretization, new problems in mathematical programming
4. A lot more work needed: presence of singular arcs, extensions of the "new discretization", ...

Thank You!