PARABOLIC REGULARITY IN VARIATIONAL ANALYSIS AND OPTIMIZATION

BORIS MORDUKHOVICH

Wayne State University, USA

Lecture 3 at the **Forum on Developments and Origins of Operations Research**

Shenzhen, China, November 2021

ORGANIZERS: **Operations Research Society of China & Southern University of Science and Technology**

TOOLS OF FIRST-ORDER ANALYSIS

Tangent cone to a set $\Omega \subset I\!R^n$ at $\bar{x} \in \Omega$

$$T_\Omega(\bar{x}) := \Big\{ w \in I\!R^n \Big| \, \exists \, t_k \downarrow 0, \ \ w_k \to w \ \text{ as } \ k \to \infty \ \text{ with } \ \bar{x} + t_k w_k \in \Omega \Big\}$$

Normal cone to $\Omega \subset I\!R^n$ at $\bar{x} \in \Omega$

$$N_\Omega(\bar{x}) := \Big\{ v \in I\!R^n \Big| \, \exists \, x_k \stackrel{\Omega}{\to} \bar{x}, \ v_k \to v \ \text{ with } \ v_k \in T^*_\Omega(x_k) \Big\}$$

Graphical derivative of a mapping $F\colon I\!R^n \rightrightarrows I\!R^m$ at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$

$$DF(\bar{x}, \bar{y})(u) := \Big\{ v \in I\!R^m \Big| \, (w, v) \in T_{\operatorname{gph} F}(\bar{x}, \bar{y}) \Big\}, \ u \in I\!R^n$$

Coderivative of $F\colon I\!R^n \rightrightarrows I\!R^m$ at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$

$$D^*F(\bar{x}, \bar{y})(v) := \Big\{ u \in I\!R^n \Big| (u, -v) \in N_{\operatorname{gph} F}(\bar{x}, \bar{y}) \Big\}, \ v \in I\!R^m$$

Subderivative of $\varphi\colon I\!R^n \to \overline{I\!R} := (-\infty, \infty]$ at $\bar{x} \in \operatorname{dom}\varphi$

$$\mathrm{d}\varphi(\bar{x})(\bar{w}) := \liminf_{t\downarrow 0,\, w\to\bar{w}} \frac{\varphi(\bar{x}+tw)-\varphi(\bar{x})}{t},\ \ \bar{w} \in I\!R^n$$

Subdifferential of $\varphi\colon I\!R^n \to \overline{I\!R}$ at $\bar{x} \in \operatorname{dom}\varphi$

$$\partial\varphi(\bar{x}) := \Big\{ v \in I\!R^n \Big|\ (v,-1) \in N_{\operatorname{epi}\varphi}\big(\bar{x}, \varphi(\bar{x})\big) \Big\}$$

SECOND SUBDERIVATIVE AND EPI-DIFFERENTIABILITY

Second-order difference quotients for a function $\varphi\colon I\!R^n \to \overline{I\!R}$ at $(\bar{x},\bar{v}) \in (\operatorname{dom}\varphi)\times I\!R^n$

$$\Delta_t^2\varphi(\bar{x},\bar{v})(u) := \frac{\varphi(\bar{x}+tu)-\varphi(\bar{x})-t\langle\bar{v},u\rangle}{\frac{1}{2}t^2} \quad \text{as} \quad u\in I\!R^n,\ t>0$$

Second subderivative of φ at $\bar{x}$ for $\bar{v}$

$$d^2\varphi(\bar{x},\bar{v})(w) := \liminf_{t\downarrow 0,\, u\to w} \Delta_t^2\varphi(\bar{x},\bar{v})(u),\ w\in I\!R^n$$

Function φ is said to be twice epi-differentiable at $\bar{x}$ for $\bar{v}$ if

$$\operatorname{epi}\Delta_t^2\varphi(\bar{x},\bar{v}) \longrightarrow \operatorname{epi} d^2\varphi(\bar{x},\bar{v}) \quad \text{as} \quad t\downarrow 0$$

which is equivalent to say that for every $w\in I\!R^n$ and every $t_k\downarrow 0$ there exists $w_k\to w$ as $k\to\infty$ such that

$$\Delta_{t_k}^2\varphi(\bar{x},\bar{v})(w_k)\to d^2\varphi(\bar{x},\bar{v})(w)$$

PARABOLIC REGULARITY

We associate with a set $\Omega \subset I\!R^n$ its indicator function $\delta_\Omega \colon I\!R^n \to \overline{I\!R}$ equal to 0 for $x \in \Omega$ and ∞ otherwise.

DEFINITION (Rockafellar-Wets, 1998) $\Omega \subset I\!R^n$ is parabolically regular at $\bar{x} \in \Omega$ for $\bar{v} \in I\!R^n$ if for any $w \in I\!R^n$ with $\mathrm{d}^2\delta_\Omega(\bar{x}, \bar{v})(w) < \infty$ there exist, among all the sequences t_k satisfying the condition

$$\Delta^2_{t_k}\delta_\Omega(\bar{x}, \bar{v})(w_k) \to \mathrm{d}^2\delta_\Omega(\bar{x}, \bar{v})(w) \quad \text{as} \quad k \to \infty$$

those with the additional property

$$\limsup_{k\to\infty} \frac{\|w_k - w\|}{t_k} < \infty$$

The property of parabolic regularity is stable (preserved) under various compositions while being more general than other known second-order regularities.

TWICE EPI-DIFFERENTIABILITY FROM PARAB. REGULARITY

THEOREM Let $(\bar{x}, \bar{v}) \in \text{gph}\, N_\Omega$, where Ω is locally closed and parabolically derivable at $\bar{x} \in \Omega$ for every $w \in K_\Omega(\bar{x}, \bar{v})$. If Ω is parabolically regular at $\bar{x}$ for $\bar{v}$, then it is properly twice epi-differentiable at $\bar{x}$ for this normal vector.

The obtained result seems to be the first one in the literature to lay down a systematic approach to verify twice epi-differentiability via parabolic regularity. It allows us to justify twice epi-differentiability for various classes of nonconvex and not fully amenable sets that frequently appear in the framework of constrained optimization.

METRIC REGULARITY AND SUBREGULARITY

A set-valued mapping $F\colon I\!R^n \rightrightarrows I\!R^m$ is metrically regular around $(\bar{x},\bar{y}) \in \operatorname{gph} F$ if there exist $\ell \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that we have the distance estimate

$$\text{dist}\Big(x; F^{-1}(y)\Big) \leq \ell\,\text{dist}\Big(y; F(x)\Big) \quad \text{for all} \quad (x,y) \in U \times V$$

The coderivative/Mordukhovich criterion gives us the complete characterization of metric regularity of closed-graph mappings

$$\ker D^*F(\bar{x},\bar{y}) := \Big\{v \in I\!R^m \Big|\ 0 \in D^*F(\bar{x},\bar{y})(v)\Big\} = \{0\}$$

Metric subregularity of F at $(\bar{x},\bar{y})$ reads as

$$\text{dist}\Big(x; F^{-1}(\bar{y})\Big) \leq \ell\,\text{dist}\Big(\bar{y}; F(x)\Big) \quad \text{for all} \quad x \in U$$

It is less investigated but leads to much stronger results. The inverse property to subregularity is calmness.

CONSTRAINT SYSTEMS (CS)

refer to sets $\Omega \subset I\!R^n$ admitting the representation: there exist a neighborhood $\mathcal{O}$ of $\bar{x}$, a single-valued mapping $f\colon I\!R^n \to I\!R^m$ twice differentiable at $\bar{x}$, and a closed subset Θ of $I\!R^m$ with

$$\Omega \cap \mathcal{O} = \Big\{ x \in \mathcal{O} \Big|\ f(x) \in \Theta \Big\}$$

Various aspects of constrained optimization (including numerical ones) have been conventionally studied under metric regularity of the mapping $x \mapsto f(x) - \Theta$ around $(\bar{x}, 0)$. The coderivative criterion amounts in this setting to

$$N_\Theta\Big(f(\bar{x})\Big) \cap \ker \nabla f(\bar{x})^* = \{0\}$$

which reduces to basic CQs for particular classes (MFCQ for NLPs, Robinson CQ for conic programs, etc.)

DEFINITION We say that that metric subregularity constraint qualification (MSCQ) holds for the constraint system Ω at $\bar{x}$ if the mapping $x \mapsto f(x) - \Theta$ is metrically subregular at $(\bar{x}, 0)$.

SECOND SUBDERIVATIVES OF CONSTRAINT SYSTEMS

Fix $(\bar{x}, \bar{v}) \in \mathrm{gph}\, N_\Omega$ for the constraint system Ω and define the set of Lagrange multipliers

$$\Lambda(\bar{x}, \bar{v}) := \Big\{ \lambda \in N_\Theta\big(f(\bar{x})\big) \,\Big|\, \nabla f(\bar{x})^* \lambda = \bar{v} \Big\}$$

THEOREM Let MSCQ hold for (CS), and let Θ be parabolically regular at $f(\bar{x})$ for any $\lambda \in \Lambda(\bar{x}, \bar{v})$. Then Ω is parabolically regular at $\bar{x}$ for $\bar{v}$, and the second subderivative of δ_Ω is

$$\mathrm{d}^2\delta_\Omega(\bar{x}, \bar{v})(w) = \max_{\lambda \in \Lambda(\bar{x}, \bar{v})} \Big\{ \langle \lambda, \nabla^2 f(\bar{x})(w, w) \rangle + \mathrm{d}^2\delta_\Theta\big(f(\bar{x}), \lambda\big)\big(\nabla f(\bar{x}) w\big) \Big\}$$

PARABOLIC REGULARITY OF REDUCIBLE SETS

DEFINITION (Bonnans and Shapiro, 2000) A closed set $\Omega \subset \mathbb{R}^m$ is said to be $\mathcal{C}^2$-cone reducible at $\bar{y} \in \Omega$ to a closed convex cone $\Gamma \subset \mathbb{R}^s$ if there exist a neighborhood $\mathcal{U} \subset \mathbb{R}^m$ of $\bar{y}$ and a $\mathcal{C}^2$-smooth mapping $h\colon \mathbb{R}^m \to \mathbb{R}^s$ such that

$$\Omega \cap \mathcal{U} = \left\{ y \in \mathcal{U} \,\middle|\, h(y) \in \Gamma \right\}, \quad h(\bar{y}) = 0, \text{ and } \nabla h(\bar{y}) \text{ has full rank}$$

This includes major constraint systems in conic programming (polyhedral, SOCP, SDP, etc.)

THEOREM Let $\Omega \subset \mathbb{R}^m$ be $\mathcal{C}^2$-cone reducible at $\bar{y} \in \Omega$ to a closed convex cone $\Gamma \subset \mathbb{R}^s$, and let $(\bar{y}, \lambda) \in \operatorname{gph} N_\Omega$. Then Ω is parabolically derivable at $\bar{y}$ for any $w \in T_\Omega(\bar{y})$ and parabolically regular at $\bar{y}$ for λ. Consequently, its indicator function δ_Ω is properly twice epi-differentiable at $\bar{y}$ for λ with the second subderivative calculated by

$$\mathrm{d}^2\delta_\Omega(\bar y,\lambda)(w)=\begin{cases}\langle\mu,\nabla^2 h(\bar y)(w,w)\rangle & \text{if } \; w\in K_\Omega(\bar y,\lambda),\\ \infty & \text{otherwise,}\end{cases}$$

where $\mu\in I\!R^s$ is the unique solution to the system

$$\lambda=\nabla h(\bar y)^*\mu,\quad \mu\in N_\Gamma\Big(h(\bar y)\Big)$$

The converse implication

$$\mathcal{C}^2-\text{cone reducibility}\Longleftarrow\;\text{parabolic regularity}$$

fails even for simple examples

NO-GAP SECOND-ORDER OPTIMALITY CONDITIONS

Consider the constrained optimization problem (CO)

$$\min_{x \in I\!R^n} \varphi(x) \quad \text{subject to} \quad f(x) \in \Theta$$

where φ and f is twice differentiable at the references point and where Ω is convex.

THEOREM Let $\bar{x}$ be a feasible solution to (CO) with $\bar{v} := -\nabla\varphi(\bar{x})$. Assume MSCQ and parabolic regularity of Θ at $f(\bar{x})$ for every $\lambda \in \Lambda(\bar{x}, \bar{v})$. The following hold:

(i) If $\bar{x}$ is a local minimizer of (CO), then

$$\max_{\lambda \in \Lambda(\bar{x},\bar{v})} \Big\{ \langle \nabla^2_{xx} L(\bar{x}, \lambda) w, w \rangle + \mathrm{d}^2\delta_\Theta\Big(f(\bar{x}), \lambda\Big)\Big(\nabla f(\bar{x}) w\Big) \Big\} \geq 0$$

for all $w \in K_\Omega(\bar{x}, \bar{v})$

(ii) The second-order sufficient condition

$$\max_{\lambda \in \Lambda(\bar{x},\bar{v})} \Big\{ \langle \nabla^2_{xx} L(\bar{x},\lambda)w, w\rangle + \mathrm{d}^2\delta_\Theta\Big(f(\bar{x}),\lambda\Big)\Big(\nabla f(\bar{x})w\Big)\Big\} > 0$$

being fulfilled for all $w \in K_\Omega(\bar{x},\bar{v}) \setminus \{0\}$ gives us the quadratic growth with some $\ell > 0$ and $\varepsilon > 0$:

$\psi(x) \geq \psi(\bar{x}) + \frac{\ell}{2}\|x-\bar{x}\|^2$ when $x \in I\!B_\varepsilon(\bar{x})$, where $\psi := \varphi + \delta_\Theta \circ f$

and hence ensures that $\bar{x}$ is a strict minimizer for (CO)

AUGMENTED LAGRANGIANS UNDER PARABOLIC REGULARITY

Augmented Lagrangian for the constraint system (CS)

$$\mathcal{L}(x,\lambda,\rho) := \varphi(x) + \frac{\rho}{2}\Big[\mathrm{dist}\Big(f(x)+\rho^{-1}\lambda;\Theta\Big)^2 - \|\rho^{-1}\lambda\|^2\Big],\ \rho>0$$

THEOREM Let $(\bar{x},\bar{\lambda})$ satisfy the KKT

$$\nabla_x L(\bar{x},\bar{\lambda}) = 0, \quad \bar{\lambda} \in N_\Theta\big(f(\bar{x})\big)$$

where Θ is parabolically regular at $\bar{x}$ for $\bar{\lambda}$. Then the following assertions are equivalent:

(i) There exists $\bar{\rho} > 0$ such that

$$\mathrm{d}_x^2\mathcal{L}\Big((\bar{x},\bar{\lambda},\rho),0\Big)(w) > 0 \ \text{ for all } \ w \in \mathbb{R}^n \setminus \{0\},\ \rho > \bar{\rho}$$

(ii) There are $\bar{\rho} > 0, \varepsilon, \ell > 0$ ensuring the quadratic growth

$$\mathcal{L}(x,\bar{\lambda},\rho) \geq \varphi(\bar{x}) + \frac{\ell}{2}\|x-\bar{x}\|^2 \ \text{ for all } \ x \in \mathbb{B}_\varepsilon(\bar{x}),\ \rho > \bar{\rho}$$

STRONG METRIC SUBREGULARITY OF SUBDIFFERENTIAL

Recall that $F\colon I\!R^n \rightrightarrows I\!R^m$ is strongly metrically subregular at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ if there exist $\kappa \geq 0$ and neigh. U of $\bar{x}$ with

$$\|x-\bar{x}\| \leq \kappa\, \operatorname{dist}\big(\bar{y}; F(x)\big) \quad \text{for all} \quad x \in U$$

THEOREM In the setting of (CO) assumed that MSCQ holds for Ω at $\bar{x}$ and that Θ is parabolically regular at $f(\bar{x})$ for every $\lambda \in \Lambda(\bar{x},\bar{v})$ with $\bar{v} := -\nabla\varphi(\bar{x})$. Then we have the equivalence

(i) The point $\bar{x}$ is a local minimizer of $\psi := \varphi + \delta_\Omega$, and the subgradient mapping $\partial\psi$ is strongly metrically subregular at $(\bar{x},0)$

(ii) The second-order sufficient optimality condition

$$\max_{\lambda \in \Lambda(\bar{x},\bar{v})} \Big\{\langle \nabla^2_{xx} L(\bar{x},\lambda)w, w\rangle + \mathrm{d}^2\delta_\Theta\big(f(\bar{x}),\lambda\big)\big(\nabla f(\bar{x})w\big)\Big\} > 0$$

is satisfied for all $w \in K_\Omega(\bar{x},\bar{v}) \setminus \{0\}$

SUBGRADIENT GRAPHICAL DERIVATIVES

THEOREM Let N_Ω be the normal cone mapping associated with the constraint system Ω in (CS) under MSCQ and parabolic regularity of Θ. Then the graphical derivative of N_Ω is

$$DN_\Omega(\bar{x},\bar{v})(w) = \bigcup_{\lambda\in\Lambda(\bar{x},\bar{v},w)} \nabla^2\langle\lambda,f\rangle(\bar{x})w + \partial_w\Big(\tfrac{1}{2}\mathrm{d}^2\delta_\Theta(f(\bar{x}),\lambda)\Big)\Big(\nabla f(\bar{x})\cdot\Big)$$

for all $w \in K_\Omega(\bar{x},\bar{v})$, where $\Lambda(\bar{x},\bar{v},w)$ stands for the set of optimal solutions to the dual second-order problem

$$\max_{\lambda\in I\!R^m} \langle\lambda,\nabla^2 f(\bar{x})(w,w)\rangle - \sigma_{T^2_\Theta(f(\bar{x}),\nabla f(\bar{x})w)}(\lambda) \quad \text{subject to} \quad \lambda\in\Lambda(\bar{x},\bar{v})$$

SEQUENTIAL QUADRATIC PROGRAMMING METHOD

Consider the constrained optimization problem

$$\text{minimize } \varphi(x) \text{ subject to } \Phi(x) \in \Theta$$

ALGORITHM (basic SQP method) Choose any $(x_k, \lambda_k) \in \mathbb{R}^n \times \mathbb{R}^m$ and set $k = 0$

- If (x_k, λ_k) satisfies the KKT system, then stop
- Compute (x_{k+1}, λ_{k+1}) as a solution to the KKT system (with $H_k = \nabla^2 L(x_k, \lambda_k)$)

$$\begin{aligned} &\min \varphi(x_k) + \langle \nabla\varphi(x_k), x - x_k\rangle + \tfrac{1}{2}\langle H(x_k)(x - x_k), x - x_k\rangle \\ &\text{subject to } \Phi(x_k) + \nabla\Phi(x_k)(x - x_k) \in \Theta \end{aligned}$$

- Increase k by 1 and then go back to Step 1

SUPERLINEAR CONVERGENCE OF SQP METHOD

THEOREM Assume that

- $\bar{x}$ is a local minimizer and $\Lambda(\bar{x}) = \{\bar{\lambda}\}$
- Θ is parabolically regular at $\Phi(\bar{x})$
- 2nd-order sufficient optimality condition holds at $\bar{x}$
- The multiplier mapping $M_{\bar{x}}$ is calm at $((0,0), \bar{\lambda})$

Then the SQP subproblems are solvable, and for any starting point (x_0, λ_0) sufficiently close to $(\bar{x}, \bar{\lambda})$, the SQP iterates (x_k, λ_k) superlinearly converges to $(\bar{x}, \bar{\lambda})$

REFERENCES

• R. T. Rockafellar and R. J-B. Wets, Variational Analysis, Springer, 1998.

• J. F. Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems, Springer, 2000.

• B. S. Mordukhovich, Variational Analysis and Applications, Springer, 2018.

• A. Mohammadi, B. S. Mordukhovich and M. E. Sarabi, Parabolic regularity in geometric variational analysis, Trans. Amer. Math. Soc. **374** (2021), 1711–1763.

• A. Mohammadi, B. S. Mordukhovich and M. E. Sarabi, Stability of KKT systems and superlinear convergence of the SQP method under parabolic regularity, J. Optim. Theory Appl. **186** (2020), 731–758.