LOCAL MONOTONICITY OF SUBGRADIENT MAPPINGS

Terry Rockafellar
University of Washington, Seattle

Workshop on Optimization and Operator Theory
The Technion, Haifa, Israel (virtual)
15–17 October 2021

Global Monotonicity of Set-Valued Mappings

$\mathcal{H}$ = Hilbert space, $T : \mathcal{H} \rightrightarrows \mathcal{H}$ set-valued mapping, operator

$$\mathrm{gph}\, T = \{(x, y) \,|\, y \in T(x)\}, \quad \mathrm{dom}\, T = \{x \,|\, T(x) \neq \emptyset\}$$

Monotonicity: $y_0 \in T(x_0),\ y_1 \in T(x_1) \implies \langle x_1 - x_0, y_1 - y_0 \rangle \geq 0$

Maximality: $\nexists$ monotone T' with $\mathrm{gph}\, T'$ strictly $\supset \mathrm{gph}\, T$

Basic problem to solve for a given T mapping

Find x such that $0 \in T(x)$

Resolvants: $(I + cT)^{-1}$, $c > 0$ nonexpansive from monotonicity

Proximal point algorithm: $x^{k+1} = (I + c_k T)^{-1}(x^k)$, $c_k \nearrow c_\infty$

under max mononicity, $\{x^k\}$ converges to a solution (if $\exists$)

Subgradients Beyond Convex Analysis

function $f : \mathcal{H} \to (-\infty, \infty]$, lower semicontinuous, $\not\equiv \infty$

Regular ("Frêchet") subgradients: $y \in \widehat{\partial} f(x)$

$$f(x') \geq f(x) + \langle y, x' - x \rangle + o\big(||x' - x||\big)$$

General ("limiting") subgradients: $y \in \partial f(x)$

$\exists \{x^\nu\}, \{y^\nu\}$, with $y^\nu \in \widehat{\partial} f(x^\nu)$, such that

$$y^\nu \underset{w}{\to} y, \quad x^\nu \to x, \quad f(x^\nu) \to f(x) < \infty$$

Monotonicity characterization of the mapping $\partial f : \mathcal{H} \rightrightarrows \mathcal{H}$

∂f max monotone $\iff$ ∂f monotone $\iff$ f convex

Proximal point algorithm: for solving $0 \in \partial f(\bar{x})$ (optimization)

$$x^{k+1} = (I + c_k \partial f)^{-1}(x^k) \iff x^{k+1} = \mathrm{argmin}_x\, f^k$$

for the function $f^k(x) := f(x) + \frac{1}{2c_k}||x - x^k||^2$

application to dual problem underlies ALM in convex programming

Localization for Nonconvex Functions?

Local monotonicity: of $T : \mathcal{H} \rightrightarrows \mathcal{H}$ at $(x^*, y^*) \in \text{gph}\, T$

the monotonicity holds in neighborhood $\mathcal{X} \times \mathcal{Y}$ of (x^*, y^*)

likewise defined: local max monotonicity

Localized algorithm: for finding $\bar{x}$ with $0 \in T(\bar{x})$

proximal point iterations succeed if started near a solution!

Local minimization: $0 \in \partial f(\bar{x})$ is necessary for local optimality

find such $\bar{x}$ by local execution of the proximal point algorithm?

x^{k+1} = a "subcritical point" of f^k: $0 \in \partial f^k(x^{k+1})$

Challenges

- articulating the algorithm with local minimization steps
- implications for f of merely local monotonicity of ∂f **?????**
- trouble from the weak convergence in the definition of ∂f

$\Longrightarrow$ henceforth assume the Hilbert space $\mathcal{H}$ is finite-dimensional

Local Monotonicity From Variational Convexity

function $f : \mathcal{H} \to (-\infty, \infty]$, lower semicontinuous, $\not\equiv \infty$
for simplicity, let $f(x)$ depend continuously on $(x, y) \in \operatorname{gph} \partial f$

Variational convexity: of f at x^* for $y^* \in \partial f(x^*)$

$\exists$ open convex neighborhood $\mathcal{X} \times \mathcal{Y}$ of (x^*, y^*) and some **convex** lsc function $\widehat{f}$ on $\mathcal{X}$ such that, for $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$(x, y) \in \operatorname{gph} \partial f \iff (x, y) \in \operatorname{gph} \partial \widehat{f} \implies f(x) = \widehat{f}(x)$$

Consequences of variational convexity

(a) the mapping $T = \partial f$ is maximal monotone in $\mathcal{X} \times \mathcal{Y}$
(b) $f(x') \geq f(x) + \langle y, x' - x \rangle$ for all $x' \in \mathcal{X}$,
when $(x, y) \in \operatorname{gph} \partial f$ in $\mathcal{X} \times \mathcal{Y}$

moreover **equivalence** holds when $y^* \in \widehat{\partial} f(x^*)$

Strong Monotonicity and Variational Strong Convexity

Strong monotonicity: of $T:\mathcal{H}\rightrightarrows\mathcal{H}$ in $\mathcal{X}\times\mathcal{Y}$ at level $s>0$

$\langle x_1-x_0, y_1-y_0\rangle \geq s||x_1-x_0||^2$ for all $(x_i,y_i)\in[\mathcal{X}\times\mathcal{Y}]\cap\operatorname{gph}T$

maximal: if $\operatorname{gph}T$ can't be enlarged in $\mathcal{X}\times\mathcal{Y}$ with this maintained

Variational strong convexity: of f at level $s>0$

same as earlier except with $\widehat{f}$ strongly convex at level s on $\mathcal{X}$

this corresponds to $f(x)-\frac{s}{2}||x||^2$ being convex on $\mathcal{X}$

Consequences of variational strong convexity

(a) $T=\partial f$ is strongly maximal monotone at level s in $X\times Y$

(b) $f(x')\geq f(x)+\langle y, x'-x\rangle+\frac{s}{2}||x'-x||^2$ for all $x'\in X$, when $(x,y)\in\operatorname{gph}\partial f$ in $X\times Y$

moreover **equivalence** holds when $y^*\in\widehat{\partial}f(x^*)$

Variational Convexity Versus Convexity

Local convexity from variational convexity?

variational convexity of f for $\mathcal{X} \times \mathcal{Y} \implies$ convexity of f on $\mathcal{X}$

when $\mathcal{Y} = \mathcal{H}$ or when $f = f_0 + \delta_C$ for C closed convex, $f \in \mathcal{C}^1$

Example 1: $f(x) = \max\{1 - e^x, 1 - e^{-x}\}$ for $x \in \mathbb{R}$

graph = two "concave wings" rising from the origin of $\mathbb{R} \times \mathbb{R}$

f is variationally convex at $x^* = 0$ for $y^* = 0$

Example 2: $f(x_1, x_2) = x_2 + \delta_P(x_1, x_2)$, $P = \{(x_1, x_2) \mid x_2 = x_1^2\}$

graph = a "tilted parabola" floating in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$

f is variationally strongly convex <u>everywhere</u>, yet $\operatorname{dom} f$ <u>nonconvex</u>

References

[1] R.T. Rockafellar (2019) "Variational convexity and local monotonicity of subgradient mappings," *Vietnam Journal of Mathematics* (issue in honor of A.D. Ioffe's 80th birthday)

[2] R.T. Rockafellar (2019) "Progressive decoupling of linkages in optimization and variational inequalities with elicitable convexity or monotonicity," *Set-valued and Variational Analysis*

[3] R.T. Rockafellar (2020) "Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality," *Mathematical Programming*

[4] R.T. Rockafellar (2021) "Convergence of augmented Lagrangian methods in extensions beyond nonlinear programming," *Mathematical Programming*

downloads: sites.washington.edu/∼rtr/mypage.html