VARIATIONAL ANALYSIS AND OPTIMIZATION

Terry Rockafellar, Prof. Emeritus
University of Washington, Seattle

UW/PIMS Colloquium
15 January 2021

Orientation

Branches of math have historically emerged out of need

- Arithmetic, classical geometry and trigonometry
- Differential calculus
- Probability

and they were then developed into powerful methodologies

Variational analysis fits this pattern

- classical version: calculus of variations
- modern version: underpinnings to optimization

linked with the advent of computers, starting around 1950

My personal role

- initial research in optimization in late 1950s
- developed "convex analysis" in 1960s to early 1970s
- worked then on bridging that with "classical analysis"

this is an ongoing effort with many new applications

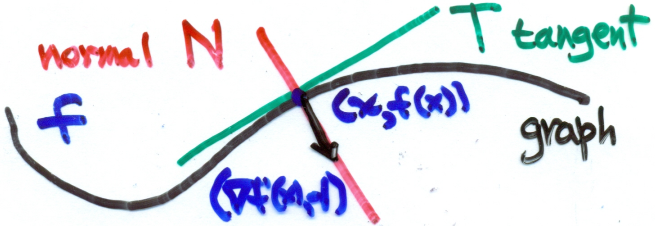
The Geometric Mindset of Classical Calculus

Functions seen through the geometry of their graphs:

as curves, surfaces and hypersurfaces

Differentiation corresponds to tangential linearization:

- tangent space T at $(x, f(x)) \longleftrightarrow$ graph of $u \mapsto \nabla f(x) \cdot u$
- normal space N at $(x, f(x)) \longleftrightarrow$ gradient vector $\nabla f(x)$



Modeling is dominated by systems of equations:

the associated geometry is that of "smooth manifolds"

solution parameterics $\longleftrightarrow$ implicit function theorem

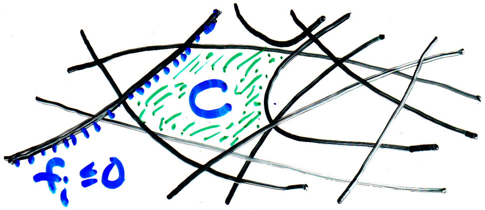
Optimization And Why It Requires Something More

Optimization problem: in finite dimensions here

minimize $f_0(x)$ over $x \in C$ for $C \subset R^n$, $x = (x_1, \ldots, x_n)$

Constraints: C = set of "**feasible solutions**", e.g.,

$$C = \left\{ x \,\middle|\, f_i(x) \le 0 \text{ for } i = 1, \ldots, m \right\}$$



Why inequalities? prescriptive versus descriptive mathematics

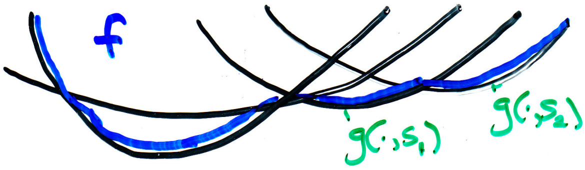
upper or lower bounds must be enforced on various quantities

there can be millions or billions of such constraints!

Max and Min Can Disrupt Differentiability

Max operations: $f(x) = \max_{s \in S} g(x, s)$ for s in some set S

Min operations: $f(x) = \min_{s \in S} g(x, s)$ for s in some set S



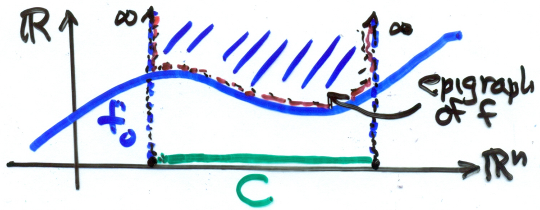
From Graphs to Epigraphs

Infinite penalties: in minimizing $f_0(x)$ over $x \in C \subset \boldsymbol{R}^n$

let $f = f_0 + \delta_C$, where $\delta_C(x) = \begin{cases} 0 \text{ if } x \in C \\ \infty \text{ if } x \notin C \end{cases}$

δ_C is the "indicator function" for C

minimizing f_0 over C $\longleftrightarrow$ minimizing f over $\boldsymbol{R}^n$



Geometry for the max and min operations just viewed:

max $\longleftrightarrow$ $\cap$ epigraphs, min $\longleftrightarrow$ $\cup$ epigraphs

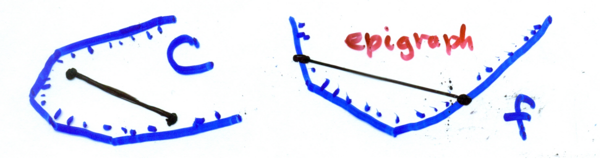
Convexity and Its Basic Consequences in Optimization

Convexity of sets: $C \subset R^n$

C is convex $\iff$ it includes all its joining line segments

Convexity of functions: $f : R^n \to (-\infty, \infty]$

f is convex $\iff$ its epigraph is a convex set



Minimizing a convex function

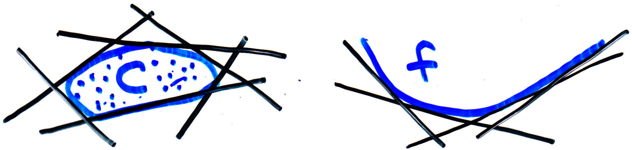
- every locally optimal solution is a globally optimal solution
- "strict" convexity precludes more than one optimal solution

- f is lower semicontinuous (lsc) $\iff$ its epigraph is a closed set

Convexity as the Next Stage Beyond Linearity

Dual charactization of convexity

- C is a closed convex set $\Longleftrightarrow$ C is some $\cap$ of closed half-spaces
- f is a lsc convex function $\Longleftrightarrow$ f is some sup of affine functions



Constraint interpretation

- convex sets $\longleftrightarrow$ systems of linear constraints
- lsc convex functions $\longleftrightarrow$ linear constrained epigraphs

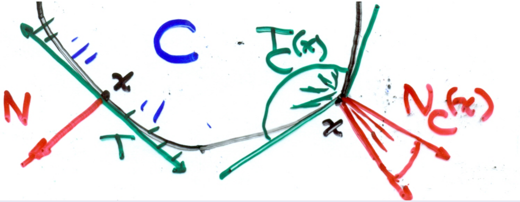
Tangents and Normals Via Convexity

Normal cone: to C at $x \in C$

$$N_C(x) = \{ v \mid v{\cdot}(x' - x) \leq 0 \text{ for all } x' \in C \}$$

Tangent cone: to C at $x \in C$

$$T_C(x) = \mathrm{cl}\,\{ w \mid x + \varepsilon w \in C \text{ for some } \varepsilon > 0 \}$$



$T_C(x)$ and $N_C(x)$ are closed convex cones polar to each other

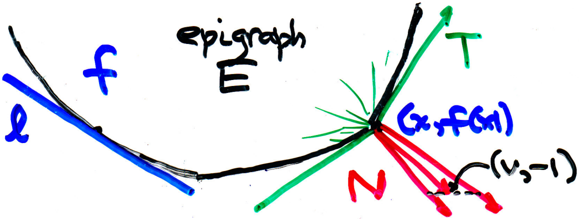
$$T_C(x) = \{ w \mid v{\cdot}w \leq 0,\ \forall v \in N_C(x) \}$$
$$N_C(x) = \{ v \mid v{\cdot}w \leq 0,\ \forall w \in T_C(x) \}$$

Cones: sets that are comprised of 0 and rays emanating from 0

polar cones generalize orthogonal subspaces!

Application to Convex Epigraphs

consider a function $f : R^n \to (-\infty, \infty]$ that is convex, lsc



Subgradient vectors: $v \in \partial f(x) \iff (v, -1) \in N_E(x, f(x))$
$\iff f(x') \geq f(x) + v \cdot (x' - x)$ for all x'

- $\partial f(x)$ is a closed, convex set [$\emptyset$ when $f(x) = \infty$]
- $\partial f(x)$ reduces to $\nabla f(x)$ if f is differentiable at x
- $\partial (f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$ if f_1 is continuous at x
- $\partial \delta_C(x) = N_C(x)$ for an indicator function δ_C

$T_E(x)$ = epigraph of associated directional derivative function

Subgradients in Convex Optimization

Optimization problem: minimize $f(x)$ over all $x \in \boldsymbol{R}^n$

for a function $f : \boldsymbol{R}^n \to (-\infty, \infty]$ that is convex, lsc, $\not\equiv \infty$

Characterization of optimality

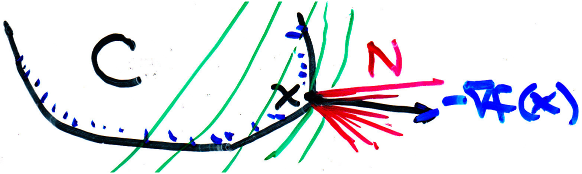
minimum of f occurs at $x \iff 0 \in \partial f(x)$

Example with a constraint set: $f = f_0 + \delta_C$

Let f_0 be differentiable convex and C closed convex $\neq \emptyset$. Then

$$\partial(f_0 + \delta_C)(x) = \partial f_0(x) + \partial\delta_C(x) = \nabla f_0(x) + N_C(x)$$

$$0 \in \partial(f_0 + \delta_C)(x) \iff -\nabla f_0(x) \in N_C(x)$$



function constraints representing $C \longrightarrow$ Lagrange multiplier rules

Tangent Vector Concepts Beyond Convex Analysis

consider $C \subset R^n$ not necessarily convex, and some $x \in C$

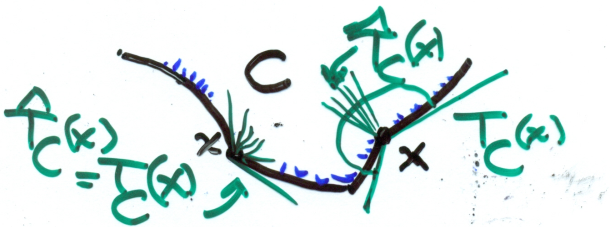
Tangent vectors, general kind: forming a cone $T_C(x)$

$$w \in T_C(x) \text{ if } \exists \tau_k \searrow 0, \exists w_k \to w : \ x + \tau_k w_k \in C$$

Tangent vectors, regular kind: forming a cone $\widehat{T}_C(x)$

$$w \in \widehat{T}_C(x) \text{ if } \forall x_k \to_C x, \forall \tau_k \searrow 0, \exists w_k \to w : \ x_k + \tau_k w_k \in C$$

These are **closed** cones with $\widehat{T}_C(x) \subset T_C(x)$ and $\widehat{T}_C(x)$ **convex**



Normal Vector Concepts Beyond Convex Analysis

consider $C \subset R^n$ not necessarily convex, and some $x \in C$

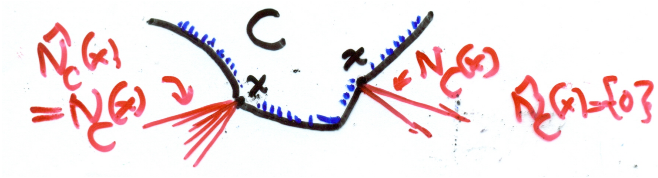
Normal vectors, regular kind: forming a cone $\widehat{N}_C(x)$

$v \in \widehat{N}_C(x)$ if $v \cdot (x' - x) \leq o(x' - x)$ for $x' \in C$

Normal vectors, general kind: forming a cone $N_C(x)$

$w \in N_C(x)$ if $\exists x_k \to_C x, \exists v_k \in \widehat{N}_C(x_k) : \; v_k \to v$

These are **closed** cones with $\widehat{N}_C(x) \subset N_C(x)$ and $\widehat{N}_C(x)$ **convex**



Regular Points of Sets in Variational Geometry

consider $C \subset R^n$ not necessarily convex, and some $x \in C$

Regularity property equivalences

$$\widehat{T}_C(x) = T_C(x) \iff T_C(x) \text{ is convex} \iff$$
$$\widehat{N}_C(x) = N_C(x) \iff N_C(x) \text{ is convex} \iff$$
$$w \in T_C(x),\ v \in N_C(x) \implies w \cdot v \leq 0 \iff$$
$$T_C(x) \text{ and } N_C(x) \text{ are convex cones polar to each other}$$

x then is called a **regular point** of C, in the variational sense

Basic examples of such regularity:

- all points x of a closed convex set C
- all points x of a "smooth manifold" C
- all points x of a set C well specified by "smooth constraints"

$\Longrightarrow$ general variational geometry reduces to previous theory in the setting of either convex analysis or classical analysis

Subgradients in Variational Analysis Beyond Convexity

Subgradients, regular kind: forming a set $\widehat{\partial} f(x)$

$$v \in \widehat{\partial} f(x) \text{ if } f(x') \geq f(x) + v{\cdot}(x' - x) + o(x' - x)$$

Subgradients, general kind: forming a set $\partial f(x)$

$$v \in \partial f(x) \text{ if } \exists x_k \to x, \exists v_k \in \widehat{\partial} f(x_k): \; v_k \to v, \; f(x_k) \to f(x)$$

Epigraphical characterization: with $E =$ epigraph of f

$$v \in \partial f(x) \iff (v, -1) \in N_E(x, f(x))$$
$$v \in \widehat{\partial} f(x) \iff (v, -1) \in \widehat{N}_E(x, f(x))$$

Definition: f is called **subdifferentially regular** at x when its epigraph E is regular at $(x, f(x))$ in the variational sense

Consequences of the geometric regularity equivalences

- these notions $\longrightarrow$ previous ones in classical or convex settings
- more generally, subdifferential regularity $\longrightarrow$ a full duality between subgradients in $\partial f(x)$ and "subderivatives" of f at x

"Generalized Equations" / "Variational Inequalities"

extending the classical paradigm of solving a system of equations

Variational inequality problem with respect to C and F

For $C \subset \boldsymbol{R}^n$ nonempty closed convex and $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ of class $\mathcal{C}^1$,

$$\text{determine } x \in C \text{ such that } -F(x) \in N_C(x)$$
$$\text{i.e., } F(x)\cdot(x' - x) \geq 0 \ \forall x' \in C$$

Reduction to equation case: $N_C(x) = \{0\}$ when $x \in \operatorname{int} C$

$\implies$ in case of $C = \boldsymbol{R}^n$, $\quad -F(x) \in N_C(x) \iff F(x) = 0$

Modeling territory: optimality conditions, equilibrium conditions

Parametric version: $-F(p,x) \in N_C(x)$, solution(s) $x \in S(p)$

$\longrightarrow$ corresponding extensions of the implicit function theorem

Some References

[1] R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press

[2] R. T. Rockafellar, R. J-B Wets (1998), *Variational Analysis*, Springer-Verlag

[3] A. L. Dontchev, R. T. Rockafellar (2009), *Implicit Functions and Solution Mappings: A View From Variational Analysis*, Springer-Verlag, **second edition: 2014**

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