Colloquium du CERMICS

École des Ponts
ParisTech

The role of convex analysis in optimization

Terry Rockafellar (University of Washington, Seattle)

25 octobre 2018

THE ROLE OF CONVEX ANALYSIS IN OPTIMIZATION

Terry Rockafellar
University of Washington, Seattle

CERMICS
École de Ponts Paris Tech
26 Oct 2018

New (???) Mathematics

Is mathematics something that is "discovered" or "created"?

- mathematics doesn't just consist of facts and rules, but also of concepts and their relationships
- mathematics provides an organized language for modeling complicated situations, posing questions and finding answers
- new challenges require completely new ideas in mathematics

physics ⇒ differential calculus, statistics ⇒ random variables

Optimization as a part of mathematics:

a fast-developing subject, hardly imagined in the classical past

Geometry as a part of mathematics:

not just physical, but a supportive way of thinking abstractly

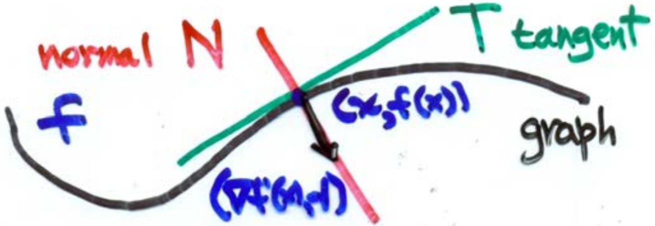
The Geometric Mindset of Classical Calculus

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 <u>classically</u> by systems of equations:

the associated geometry is that of "smooth manifolds"

solution parameterics $\longleftrightarrow$ implicit function theorem

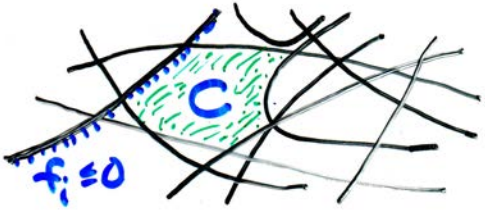
Optimization and Why it Requires Something More

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

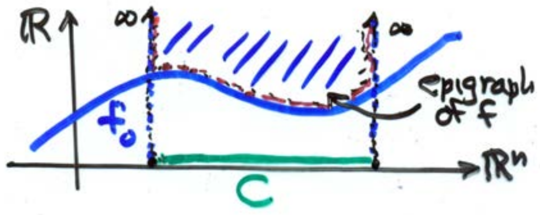
From Graphs to Epigraphs

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

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

δ_C is the "indicator function" for C

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



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

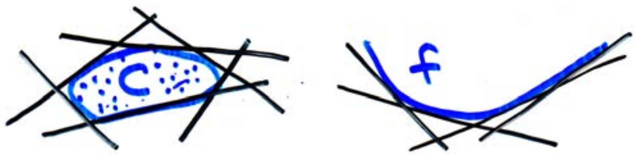
- every locally optimal solution is a globally optimal solution
- "strict" convexity precludes more than one optimal solution

- 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 $\longleftrightarrow$ linearly constrained epigraphs

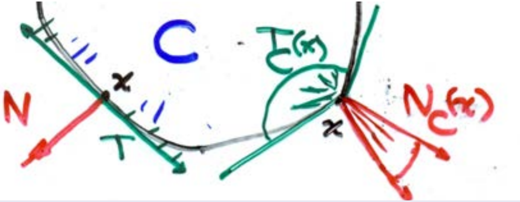
Tangents and Normals Via Convexity

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

$$N_C(x) = \{ v \,|\, v \cdot (x' - x) \le 0 \text{ for all } x' \in C \}$$

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

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



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

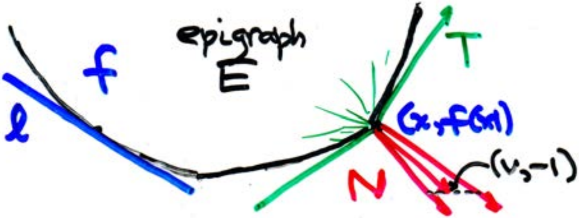
$$T_C(x) = \{ w \,|\, v \cdot w \le 0,\ \forall v \in N_C(x) \}$$
$$N_C(x) = \{ v \,|\, v \cdot w \le 0,\ \forall w \in T_C(x) \}$$

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

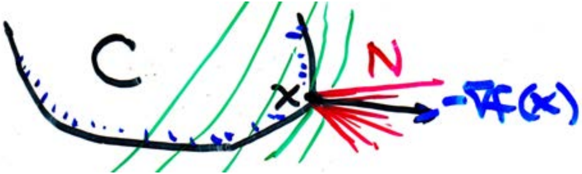
$$\text{minimum of } f \text{ occurs at } x \iff 0 \in \partial f(x)$$

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

"Generalized Equations" / "Variational Inequalities"

extending the classical idea 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 $\mathcal{C}^1$,

determine $x \in C$ such that $-F(x) \in N_C(x)$

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: www.math.washington.edu/∼rtr/mypage.html