TOP Discussion on the article:
Generalized Derivatives and Nonsmooth Optimization
A Finite Dimensional Tour
(by J. Dutta)

Discussion comment: Aris DANIILIDIS

It is commonly accepted that nonsmoothness arises naturally in optimization: even if one considers a smooth data model, several operations associated with control or optimization destroy the initial differentiability and lead to the need of employing nonsmooth techniques. For example, considering the minimization of a nonsmooth function, it has been observed that, in general, the minimum occurs at a point of nondifferentiability. This having said, even the mere formulation of optimality conditions needs to be revised in the light of nonsmooth analysis.

The survey starts its tour from the class of convex functions and moves progressively to the classes of locally Lipschitz and lower semicontinuous functions. It mainly deals with topics as subdifferential calculus, optimality conditions, regularity and semismoothness, set regularity, graphical derivatives, extreme principle and second order nonsmooth theory. It also devotes a paragraph to the interesting theory of quasi-differentiability developed by A. Rubinov and V. Demyanov for the class of DC functions.

Convex functions and subdifferential theory have their origins in the seminar work of J.-J. Moreau in nonregular mechanics, followed by a systematic treatment of T. Rockafellar in the early 60's. These functions enjoy good differentiable properties[1] and form a first class of nonsmooth functions for which everything works good. In particular, convex functions admit a natural definition of subdifferential and of generalized derivative (as a support function) and enjoy robust calculus rules. These rules have been served as a model for the theory of maximal monotone operators and led to the definitions of the variational sum or of the extended sum of maximal monotone operators (see [13], for example). Furthermore, convex functions can be completely determined by their subdifferentials: a classical result of T. Rockafellar ([22]) asserts that equality of subdifferentials (that is, $\partial f = \partial g$) implies equality of functions up to a constant (that is, $f = g + c$), provided at least one of f and g is a priori assumed to be convex. For an extension of this result we quote [24]. Let us also mention the remarkable nonsmooth integration result of Rockafellar asserting that a given multivalued operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the subdifferential of a convex function f if and only if it is maximal cyclically monotone[2]. On the other hand, besides the almost everywhere differentiability of locally Lipschitz functions (Rademacher theorem) and their the differentiability on a dense set (Preiss theorem, see [21]), the Clarke subdifferential does not determine in a unique manner the function

[1] Convex continuous functions on reflexive Banach spaces are generically differentiable [4].
[2] This result is valid in a general normed space.

(see [7]), unless the function is subdifferentially regular. This having said, it is not surprising that the majority of results concerning locally Lipschitz functions are naturally restricted to subclass of subdifferentially regular functions, as for example the important class of lower-C^1 functions introduced by Spingarn in [23]. Lower-C^k functions, that is, functions that are locally representable as the maximum of a continuously compactly parameterized family of C^k functions defined on the same open set, are a natural extension of both convex continuous functions and of C^k functions and enjoy good stability and subdifferentiability properties. Integration of multivalued operators for operators satisfying a weak cyclic monotonicity in the spirit of [22] and giving rise to these functions are established in [15], [11]. Nonsmooth integration results for operators associated to the class of essentially smooth functions are given in [6]. The use of normal cones relates naturally nonsmooth functions to their epigraphs and leads to several classes of set regularity (see [3] for a recent survey).

The subject is vast. As observed by the author, it is practically impossible to deal with or even to mention all possible aspects or developments. The author gives at the end a list of relevant topics that are not treated for practical reasons. Let me add some more topics in that list: the development and implementation of the bundle method for the minimization of nondifferentiable functions ([16]), the important applications of the monotone (or hypomonotone) multivalued operators to the proximal algorithm (see [10] and references therein), the theory of differential inclusions for such operators (see [5]) and its applications to the non-regular mechanics (see [1], for example) or to the Moreau sweeping process (see [9], for example).

Another issue is the overall presentation, which of course can neither be unique nor take into consideration all historical or further developments. Let us mention that an alternative development could have been based on the Ekeland Variational Principle [12], followed by the Borwein-Preiss smooth variational principle [8], and/or the nonsmooth Mean Value Theorem of Lebourg for locally Lipschitz functions or of Zagrodny for lower semicontinuous functions (see [25] or [2]).

Let us finally mention that nonsmoothness seldom occurs in a random manner, but instead often has an underlining structure which can be exploited in optimization. A first step towards this direction has been made by the INRIA team of numerical optimization, headed by C. Lemaréchal, with the introduction of the $\mathcal{UV}$-Lagrangian together with a first conceptual algorithm for convex functions (see [17], for details). This led to an implementable algorithm (called the *fast-track* algorithm), proposed by R. Mifflin and C. Sagastizábal (see [20], for example). Inspiring by similar considerations, A. Lewis introduced the notion of *partial smoothness* (see [18], [14]) in order to unify several *structured optimization* problems considered in previous works of J. Burke, M. Overton, F. Oustry and others. As a recent development it is worthmentioned the use of real algebraic geometry techniques in optimization (see [19], for example). Similar attempts are currently under development in other fields as game theory, optimal control and dynamical systems.

References

[1] Acary, V. Brogliato, B., Daniilidis, A. & Lemarechal, C., On the Equivalence Between Complementarity Systems, Projected Dynamical Systems and Unilateral Differential Inclusions, *Systems & Control Letters* (to appear)

[2] Aussel, D., Corvellec, J.-N. & Lassonde, M., Mean value property and subdifferential criteria for lower semicontinuous functions, *Trans. Amer. Math. Soc.* 347 (1995), no. 10, 4147–4161.

[3] Aussel, D., Daniilidis, A. & Thibault, L, Subsmooth sets: functional characterizations and related concepts, *Trans. Amer. Math. Soc.* **357** (2005), 1275–1301.

[4] Asplund, E., Fréchet differentiability of convex functions, *Acta Math.* **121** (1968), 31–47.

[5] Brézis, H., *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert,* North-Holland Mathematics Studies **5**, New York, 1973.

[6] Borwein, J., Moors, W. & Shao, Y., Subgradient representation of multifunctions, *J. Austral. Math. Soc.* **40** (1999), 301–313.

[7] Borwein, J. & Wang, X., Lipschitz functions with maximal Clarke subdifferentials are generic, *Proc. Amer. Math. Soc.* **128** (2000), 3221–3229.

[8] Borwein, J. & Preiss, D., A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, *Trans. Amer. Math. Soc.* **303** (1987), 517–527.

[9] Colombo, G. & Monteiro Marques, M., Sweeping by a continuous prox-regular set, *J. Differential Equations* **187** (2003), 46–62.

[10] Combettes, P. & Pennanen, T., Proximal methods for cohypomonotone operators, *SIAM J. Control Optim.* **43** (2004), 731–742.

[11] Daniilidis, A., Georgiev, P. & Penot, J.-P., Integration of multivalued operators and cyclic submonotonicity, *Trans. Amer. Math. Soc.* **355** (2003), 177–195.

[12] Ekeland, I., On the variational principle, *J. Math. Anal. Appl.* **47** (1974), 324–353.

[13] Garcia, Y., Lassonde, M. & Revalski, J., Extended sums and extended compositions of monotone operators, *J. Convex Analysis* (to appear).

[14] Hare, W. & Lewis, A., Identifying active constraints via partial smoothness and prox-regularity, *J. Convex Analysis* **11** (2004), 251-266.

[15] Janin, R., Sur des multiapplications qui sont des gradients généralisés, *C. R. Acad. Sci. Paris* **294** (1982), 115–117.

[16] Lemaréchal, C. & Sagastizábal, C., Variable metric bundle methods: from conceptual to implementable forms, *Math. Programming* **76** (1997), 393–410.

[17] Lemaréchal, C., Oustry, F. & Sagastizábal, C., The U-Lagrangian of a convex function, *Trans. Amer. Math. Soc.* **352** (2000), 711–729.

[18] Lewis, A., Active sets, nonsmoothness, and sensitivity, *SIAM J. Optim.* **13** (2002), 702–725.

[19] Lewis, A., *Robust regularization*, preprint SFU 2002 (availiable electonically at `http://www.orie.cornell.edu/~aslewis`).

[20] Mifflin, R. & Sagastizábal, C., On $\mathcal{VU}$-theory for functions with primal-dual gradient structure, *SIAM J. Optim.* **11** (2000), 547–571.

[21] Preiss, D., Differentiability of Lipschitz functions on Banach spaces, *J. Funct. Anal.* **91** (1990), 312–345.

[22] Rockafellar, R. T., On the maximal monotonicity of subdifferential mappings, *Pacific J. Math.* **33** (1970), 209–216.

[23] Spingarn, J., Submonotone subdifferentials of Lipschitz functions, *Trans. Amer. Math. Soc.* **264** (1981), 77–89.

[24] Thibault, L. & Zagrodny, D., Integration of subdifferentials of lower semicontinuous functions on Banach spaces, *J. Math. Anal. Appl.* **189** (1995), 33–58.

[25] Zagrodny, D., Approximate mean value theorem for upper subderivatives, *Nonlinear Anal.* **12** (1988), 1413–1428.