

SEMI-ALGEBRAIC IDEAS IN NONSMOOTH OPTIMIZATION

Adrian Lewis

School of ORIE, Cornell University, USA

Keywords: semi-algebraic, variational analysis, nonsmooth optimization, semismooth, Lojasiewicz inequality, critical point, Sard's theorem, pseudospectra

Modern variational analysis provides a sophisticated unification of convex and smooth optimization theory, achieving striking generality but at the expense of possible pathology. The general theory must handle highly irregular and oscillatory functions and sets, and yet, on the other hand, a rich family of concrete instances involve no such pathology. In particular, from a variety of variational-analytic perspectives, semi-algebraic sets—finite unions of sets defined by finitely many polynomial inequalities—are well behaved.

In joint work with J. Bolte, A. Daniilidis, A. Ioffe, C.H.J. Pang, and M. Shiota, the author illustrates a variety of situations where semi-algebraic techniques resolve variational-analytic challenges. Most of the following examples extend to sets and functions that are “subanalytic” or, more generally, “tame”.

We begin with an algorithmic application. Consider a locally Lipschitz function $F: \mathbf{R}^n \rightarrow \mathbf{R}^n$. Superlinear convergence of nonsmooth Newton methods for the equation $F(x) = 0$, depend on “semismoothness” of F . Semi-algebraic locally Lipschitz functions are always semismooth (Bolte et al, 2007b).

A more classical example involves the famous Lojasiewicz inequality, which states that, for any critical point a of a real-analytic function $f: \mathbf{R}^n \rightarrow \mathbf{R}$, there is an exponent $\theta \in [0, 1)$ such that the function

$$\frac{|f - f(a)|^\theta}{\|\nabla f\|}$$

remains bounded around a . An analogous inequality holds for semi-algebraic functions (Bolte et al, 2007a), leading to proofs of the finite length of trajectories of the associated subgradient dy-

namical system

$$\frac{dx}{dt} \in -\partial f(x).$$

Critical point theory furnishes another example where semi-algebraic assumptions lead to elegantly simple characterizations. Among notions of critical points available for nonsmooth functions, the approach via the “weak slope” has considerable theoretical appeal. Unfortunately, such critical points seem hard to recognize in general. However, for semi-algebraic functions on \mathbf{R}^2 , a simple topological characterization suffices (Ioffe and Lewis, 2007).

The graph of any semi-algebraic function admits a Whitney stratification. This technique allows us to relate a classical smooth notion, the size of gradients of the function restricted to its various strata, to a fundamental nonsmooth idea, the size of Clarke subgradients (Bolte et al, 2007c).

Stratification provides one route to a nonsmooth Morse-Sard theorem. The classical version asserts that any sufficiently smooth function $f: \mathbf{R}^n \rightarrow \mathbf{R}$ is constant along any arc of critical points. Analogous nonsmooth versions hold for semi-algebraic f (Bolte et al, 2005, 2006). Consequently, such functions can have only finitely many critical values—a nonsmooth variant of the classical theorem of Sard, and a precursor of recent very general set-valued versions of Sard's theorem due to Ioffe. Such results have striking implications for variational analysis: given any semi-algebraic generalized equation (such as a semi-algebraic system of inequalities, for example), small perturbations almost surely render the system “metrically regular”. Metric regularity is a central notion both in variational theory and computational practice, guaranteeing that approximate solutions to the system, as measured

by the *a posteriori* error, are close to exact solutions.

We end with a more concrete example. The eigenvalues of a nonsymmetric matrix A may be very sensitive to slight perturbations to the matrix, due to eigenvalue coalescence. Furthermore, it is well known that eigenvalues may be misleading as practical modeling tools. For example, the spectral radius of A predicts the asymptotic stability of the dynamical system $x_{k+1} = Ax_k$, but is insensitive to transient peaks. A more predictive modeling tool is the *pseudospectrum*

$$\Lambda_\epsilon(A) = \left\{ z \in \mathbf{C} : \|(A - zI)^{-1}\| \geq \frac{1}{\epsilon} \right\},$$

for some small constant $\epsilon > 0$. The pseudospectrum consists of all eigenvalues of small perturbations to A . Since the “resolvent norm” $z \mapsto \|(A - zI)^{-1}\|$ is semi-algebraic, it has at most finitely many critical values. Consequently, unlike the spectrum, the set-valued mapping Λ_ϵ is Lipschitz around A for all small $\epsilon > 0$ (Lewis and Pang, 2006).

REFERENCES

- Bolte, J., A. Daniilidis, A.S. Lewis and M. Shiota (2005): Clarke critical values of subanalytic Lipschitz continuous functions. *Annales Polonici Mathematici*, vol. 87, 13–25.
- Bolte, J., A. Daniilidis and A.S. Lewis (2006): The Morse-Sard theorem for non-differentiable subanalytic functions. *Journal of Mathematical Analysis and Applications*, vol. 321, 729–740.
- Ioffe, A.D. and A.S. Lewis (2007): Critical points of simple functions. *Optimization*, to appear.
- Bolte, J., A. Daniilidis and A.S. Lewis (2007a): The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *SIAM Journal on Optimization*, vol. 17, 1205–1223.
- Bolte, J., A. Daniilidis and A.S. Lewis (2007b): Tame mappings are semismooth. *Mathematical Programming, Series B*, to appear.
- Bolte, J., A. Daniilidis, A.S. Lewis and M. Shiota (2007c): Clarke subgradients of stratifiable functions. *SIAM Journal on Optimization*, to appear.
- Lewis, A.S., and C.H.J. Pang (2006): Variational analysis of pseudospectra, submitted.