Twin Estimates for Slopes

by
Humberto Muñoz
R. Baker Kearfott, `rbk@louisiana.edu`
Department of Mathematics, University of Louisiana at Lafayette

Note: The ideas for this talk are contained in the first author's Ph.D. dissertation *Interval Slopes and Twin Slope Arithmetic in Nonsmooth Optimization*, University of Louisiana at Lafayette, 2001.

What is Twin Arithmetic

- Twin arithmetic operates on pairs of intervals;
- The first element of the pair gives inner bounds on the exact answer.
- The second element of the pair gives outer bounds on the exact answer.
- Twin arithmetic is based on Kaucher arithmetic.
  - Kaucher arithmetic is in E. W. Kaucher, *Computing (Suppl.)* **2** (2), pp. 33–49, 1980.
  - Twin arithmetic is introduced in E. Gardeñes and A. Trepat, *Computing* **24**, pp. 161–179, 1980.
  - Twin arithmetic is clarified in V. Kreinovich, V. M. Nesterov and N. A. Zheludeva, *Reliable Computing* **2** (2), pp. 119–124, 1996.

Why Do We Consider Inner Bounds?

- To obtain bounds on the sharpness of outer estimates.
- To automatically compute elements guaranteed to be in certain sets.
  - There is much development of algorithms for non-smooth optimization based on *generalized gradients*.
  - Such algorithms require an element of the generalized gradient.
  - We have analyzed relationships between interval slopes and generalized gradients.
  - We compute inner slope bounds with twin arithmetic, to get elements of the generalized gradient.

What Are Interval Slopes?

- A *slope set* $\boldsymbol{A} = \boldsymbol{S}(F, \boldsymbol{x}, \check{x})$ for F at $\check{x}$ is any set with the property that $\forall x \in \boldsymbol{x}$, $\exists A \in \boldsymbol{A}$ such that

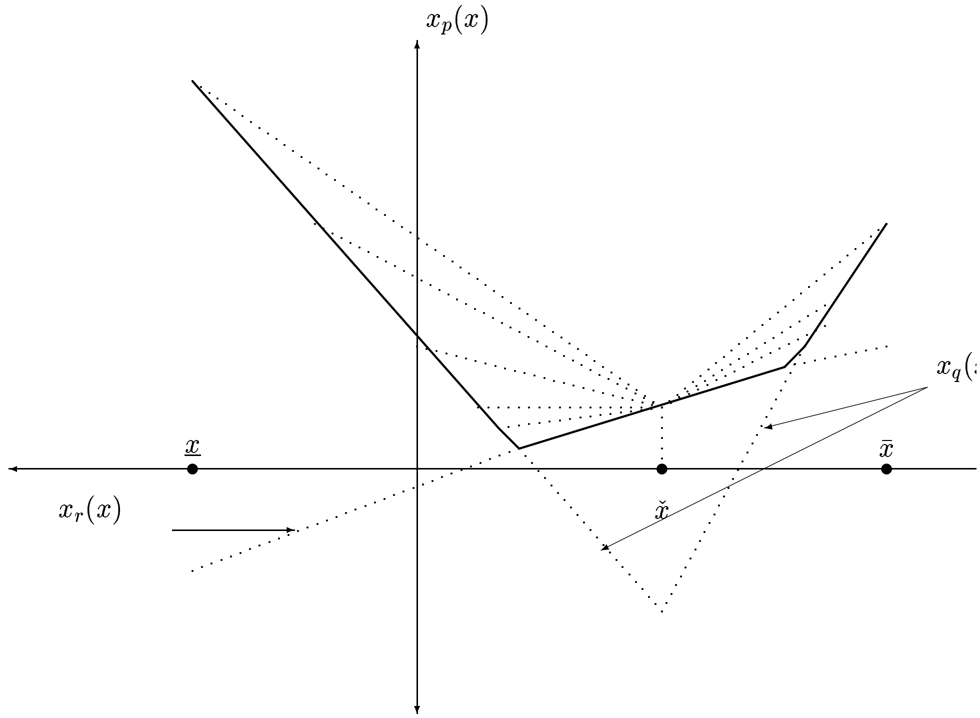
$$F(x) - F(\check{x}) = A(x - \check{x}).$$

- We denote any smallest such set by $\boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x})$.
- Outer bounds on $\boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x})$ can be computed via automatic differentiation. (See Kearfott's 1996 book or Ratz' dissertation.)

Alefeld (1980) and Krawczyk and Neumaier (1985) introduced the slope concept, while Shen and Wolfe (1989, etc.) furthered the ideas. Hansen (1993), one of the authors (1996), Ratz (1997), Oliviera (1996) and others have used the concept in optimization.

Slopes for Non-Smooth Functions

An Illustration



$\boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x})$, where
$F(x) = \max\{x_q(x), x_r(x)\}$.

What About Generalized Gradients?

- Generalized gradients also are designed to describe the variation of non-smooth functions F.
- The basic definition of generalized gradient is more involved than the definition of slope.
- Generalized gradients are obtained as limits at a point
- To our knowledge, general techniques of automatically computing generalized gradients have not been developed in the literature.
- Nonetheless, there is much literature on algorithms involving generalized gradients.

Generalized Gradients

Some Intuition

Theorem 1 (Clarke's book, p. 63) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x, and suppose S is any set of Lebesgue measure 0 in $\mathbb{R}^n$ and $\Omega_f \cup S$ is the set of points at which f is not differentiable. Then the generalized gradient $\partial f(x)$ obeys*

$$\partial f(x) = \\ co\left\{\lim \nabla f(x_i) : x_i \to x, \quad x_i \notin S \cup \Omega_f\right\},$$

where co *represents the convex hull of all limits of convergent sequences $\nabla f(x_i)$ and x_i is any sequence in $\mathbb{R}^n \setminus (\Omega_f \cup S)$ converging to x.*

Generalized Gradients

Some References

The theoretical development: F. H. Clarke, *Optimization and Nonsmooth Analysis*, Classics in Applied Mathematics, SIAM, Philadelphia, 1990 (originally Wiley, 1983)

Some Algorithms: Krzysztof C. Kiwiel, *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1985.

Coordinate Projections

Generalized Gradients

- Neither Clarke's generalized gradient nor interval slopes are rectangularly-shaped sets (i.e. interval vectors).
- We have found comparisons of the coordinate-wise projections of these sets to be the most meaningful comparisons.
- Coordinate-wise projections of the generalized gradient (from Clarke's book):

Definition 1 *The i-th projection, $\pi_i \partial f(x_1, \ldots, x_n)$, of the generalized gradient ∂f is the set*

$$\pi_i \partial f(x_1, \ldots, x_n) = \{z_i \in \mathbb{R} : \text{ for some } z_j \in \mathbb{R}, j = 1, \ldots, n, j \neq i, (z_1, \ldots, z_{i-1}, z_i, z_{i+1}, \ldots, z_n) \in \partial f(x)\}.$$

Relationships between Generalized Gradients and Slopes

Slopes at a Point?

- For non-smooth functions, the limiting value of the slope as we make the domain smaller is a non-degenerate set similar to the generalized gradient.
- Formally, we define this limiting set as

$$\lim_{\mathrm{w}(\boldsymbol{x})\to 0} \square \boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x}) =$$

$$\left[\liminf_{x\to\check{x}} \tfrac{f(x)-f(\check{x})}{x-\check{x}},\ \limsup_{x\to\check{x}} \tfrac{f(x)-f(\check{x})}{x-\check{x}}\right].$$

Coordinate Projections

Interval Slopes and Comparisons

Definition 2 *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near $\check{x}$. For $i = 1, \ldots, n$, the* ***ith projection*** $\pi_i \lim_{w(\boldsymbol{x})\to 0} \boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x})$ *is defined as the set*

$$\lim_{w(\boldsymbol{x})\to 0} cl\{s_i \in \mathbb{R} : \exists s \in \mathbb{R}^n, \\ f(x) - f(\check{x}) = s \cdot (x - \check{x}), x \in \boldsymbol{x}, x \neq \check{x}\}.$$

Theorem 2 *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be Lipschitz of rank K near $\check{x}$. Then*

$$\lim_{w(\boldsymbol{x})\to 0} \pi_i \boldsymbol{S}^{\#}(F, \boldsymbol{x}, \check{x}) \subseteq \pi_i \partial f(\check{x}).$$

- Thus, if we compute inner bounds on the coordinates of the slopes, we obtain elements of the generalized gradient.

How Do We Automatically Compute Inner Interval Slopes?

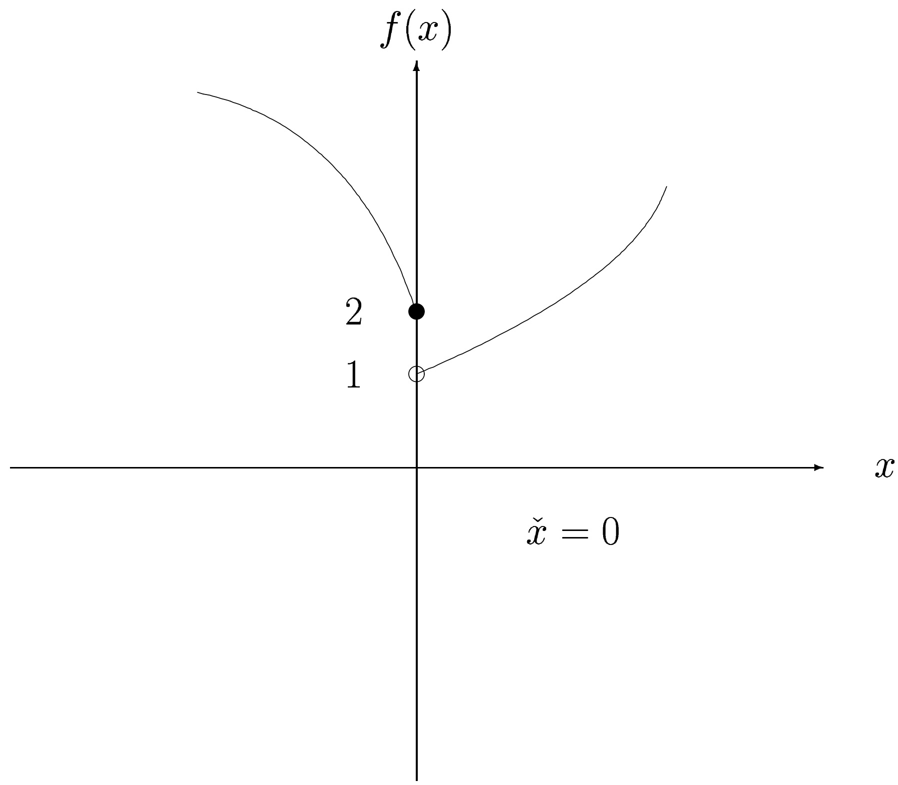
- We use twin arithmetic to obtain inner bounds for rational expressions.
- We have developed a special analysis to obtain sharp inner estimates for particular elementary functions such as $\max\{f_1, f_2\}$ and $\min\{f_1, f_2\}$.
- Non-sharp inner slope estimates for any function $f : \mathbb{R} \to \mathbb{R}$ can always be obtained by computing

$$\frac{f(x) - f(\check{x})}{x - \check{x}} \subseteq \boldsymbol{S}^{\#}(f, \boldsymbol{x}, \check{x})$$

  for a finite number of points $x \in \boldsymbol{x}$, $x \neq \check{x}$.
- This is all automated through usual automatic differentiation procedures, as used to compute partial derivatives or ordinary slopes.

Additional Results

Discontinuous Functions



- $\lim_{\mathrm{w}(\boldsymbol{x})\to 0} \boldsymbol{S}^{\#}(f, \boldsymbol{x}, \check{x}) = (-\infty, -2]$.
- $0 \notin \lim_{\mathrm{w}(\boldsymbol{x})\to 0} \boldsymbol{S}^{\#}(f, \boldsymbol{x}, \check{x})$.
- Nonetheless, f has an infimum at $\check{x}$.

Discontinuous Functions

- We want global optimization algorithms to detect critical points corresponding to extrema.
- Symmetric slopes are used for this purpose:

  Definition 3 *Let* $f : \mathbb{R}^n \to \mathbb{R}$*. The vector* $\mathbf{SS}$ *is said to be a* symmetric slope set *for* f *over* $\boldsymbol{x}$ *and centered on the interval vector* $\check{\mathbf{x}}$ *if, for each coordinate vector* e_i*,* $x \in \boldsymbol{x}$ *and* $\check{x} \in \check{\mathbf{x}}$*,* $x \neq \check{x}$*,*

  $$f(x) - \lim_{t \uparrow 0} f(\check{x} + te_i) \;=\; S_{i1} \cdot (x - \check{x}) \text{ and}$$
  $$f(x) - \lim_{t \downarrow 0} f(\check{x} + te_i) \;=\; S_{i2} \cdot (x - \check{x}),$$

  for some $S_{i1}, S_{i2} \in \mathbf{SS}, \quad i = 1, \ldots, n.$

- Symmetric slopes are related to the non-smooth analysis concept of *semigradient.*

What About the Details?

(Careful definitions, additional theory, formulas, etc.)

- Details can be found in Humberto Muñoz' Ph.D. dissertation (University of Louisiana at Lafayette, December, 2001)
- Results include many additional theoretical comparisons, including comparisons with slant slopes, etc.
- We are presently working on additional publications.
- We are presently finishing an algorithm for twin arithmetic and twin slopes, for possible publication in *ACM TOMS*.