Weak sharp minima on Riemannian manifolds[1]

Chong Li

Department of Mathematics
Zhejiang University
Hangzhou, 310027, P R China
cli@zju.edu.cn

April. 2010

Outline

Weak sharp minima on Banach spaces

- Some notations:

  X a normed linear space
  X^* the dual space of X
  $C \subset X$ a convex closed subset
  $\mathbb{B}$ denotes the closed unit ball of X
  $\mathbb{B}^*$ denotes the closed unit ball of X^*
  $T_C(x)$ the tangent cone to C at x
  $N_C(x)$ the normal cone to C at x
  $d_C(x) = \inf_{\bar{x} \in C} \|x - \bar{x}\|$
  $P(x|C)$ the projection

Weak sharp minima on Banach spaces

$f : X \to R \cup \{+\infty\}$ a proper lower semi-continuous convex function

- Consider the optimization problem

$$\mathcal{P} : \quad \begin{array}{l} \text{minimize } f(x) \\ \text{subject to } x \in X. \end{array}$$

Let $\bar{S}$ be the solution set of $\mathcal{P}$, that is,

$$\bar{S} = \{x \in X : f(x) = \min_{y \in X} f(y)\}.$$

Weak sharp minima on Banach spaces

solution set $\bar{S}$ is

- weak sharp minima :

$$f(x) \geq f(\bar{x}) + \alpha d_{\bar{S}}(x) \quad \forall \bar{x} \in \bar{S}, \ \forall x \in X. \tag{1}$$

- local weak sharp minima (α and ε depend on $\bar{x}$)

$$f(x) \geq f(\bar{x}) + \alpha d_{\bar{S}}(x) \quad \forall \bar{x} \in \bar{S}, \ \forall x \in \mathbf{B}(\bar{x}, \varepsilon).$$

Weak sharp minima on Banach spaces

- The notion of sharp minima, or strongly unique local minima emerged in Cromme (1978) and Polyak (1979) as an important tool in
  - the analysis of the perturbation behavior of certain classes of optimization problems
  - the convergence analysis of algorithms designed to solve these optimization problems.
- Later, Ferris (1988) introduced the term weak sharp minima to describe the extension of the notion of sharp minima to include the possibility of a non-unique solution set.

Weak sharp minima on Banach spaces

Applications of weak sharp minima:
sensitivity analysis and convergence analysis of optimization algorithms
((Burke (2000,2001), Jourani (2000), Lewis and Pang (1996), Ye, Zhu and Zhu (1997), Burke and Ferris (1993, 1995), Burke and More (1988), Feris (1990, 1991), Li and Wang (2002), Burke and Deng (2002) and references therein.)

Weak sharp minima on Banach spaces

Relationship between weak sharp minima, error bound and metrically regular:
$h : X \to R \cup \{+\infty\}$ a proper lower semi-continuous convex function

- consider the convexity inequality

$$h(x) \leq 0 \tag{2}$$

Let C denote the solution set of (2), that is,

$$C = \{x \in X : h(x) \leq 0\}.$$

Weak sharp minima on Banach spaces

- A global error bound for (2) means that there exists a constant $\tau > 0$ such that

$$d_C(x) \leq \tau[h(x)]_+ \quad \forall x \in X. \tag{3}$$

- $x_0 \in \partial C$, the topological boundary of C, (2) is said to be metrically regular if there exists $\tau, \delta \in (0, +\infty)$ such that

$$d_C(x) \leq \tau[h(x)]_+ \quad \forall x \in \mathbf{B}(x_0, \delta) \tag{4}$$

The study of error bounds and metrically regular has drawn much attention during recent years and there are more than 100 published papers

Weak sharp minima on Banach spaces

Let $f : X \to R \cup \{+\infty\}$ be given by

$$f(x) = [h(x)]_+.$$

Then

- error bound (resp. metrically regular) for $h(x) \leq 0 \iff C$ is weak sharp minima (resp. local weak sharp minima) for f.

Weak sharp minima on Banach spaces

The two key tools to characterize weak sharp minima:

$$\partial d_{\bar{S}}(x) = N_{\bar{S}}(x) \cap \mathbb{B}^* \quad \forall x \in \bar{S}$$

and

$$y \in P(x|\bar{S}) \Longleftrightarrow \langle x - y, z - y \rangle \leq 0 \quad \forall z \in \bar{S}.$$

Weak sharp minima on Banach spaces

Characterization of weak sharp minima are

In terms of normal cone:

Theorem A (Burke and Deng (2002)) The following statements are equivalent:

1. $\bar{S}$ is a set of weak sharp minima for f with modulus $\alpha > 0$
2. $\alpha\mathbb{B}^* \cap N_{\bar{S}}(x) \subset \partial f(x)$ holds for all $x \in \bar{S}$

In terms of directional derivative:

Theorem B (Burke and Deng (2002)) The following statements are equivalent:

1. $\bar{S}$ is a set of weak sharp minima for f with modulus $\alpha > 0$
2. $f'(x, u) \geq \alpha d_{T_{\bar{S}}(x)}(u)$ holds for all $x \in \bar{S}$ and $u \in X$
3. $f'(z, y - z) \geq \alpha d_{\bar{S}}(y)$ holds for all $y \in X$ and $z \in P(y|\bar{S})$

Extensions of some results for optimization problems

Recently, the interests are focused on solving optimization problem on Riemannian manifolds, extending non-smooth analysis, monotone operator, proximal point algorithm and so on to Riemannian manifolds.

- Optimization problems posed on manifolds arise in many natural contexts. See for example Smith (1994), Edelman, Arias and Smith (1998), Mahony (1996) $\cdots\cdots$

Extensions of some results for optimization problems

- Smith (1994): Newton's method and conjugate gradient method to solve optimization problems on Riemannian manifold.
- Edelman, Arias and Smith (1998): Newton's method and conjugate gradient method to solve optimization problem on the Stiefel and Grassmann manifolds.
- Ferreira and Svaiter (2002): the well-known Kantorovich theorem for Newton's method to Riemannian manifolds
- Dedieu, Priouret and Malajovich (2003): Smale's α-theory and γ-theory to analytic vector fields on Riemannian manifolds,
- Udriste (book, 1994): convex functions and optimization methods on Riemannian manifolds.

Extensions of some results for optimization problems

Extensions of non-smooth analysis, monotone operator, proximal point algorithm and so on to manifolds are also quite fruitful. For example:

- Ledyaev and Zhu (2007): subdifferential calculus for l.s.c. functions on manifolds to study constrained optimization problems, non-classical problems of calculus of variations, and generalized solutions of first-order partial differential equations on manifolds.
- Azagra, Ferrera and López-Mesas (2005): subdifferential on manifolds to solve the Hamilton-Jacobi equations defined on Riemannian manifolds.

Extensions of some results for optimization problems

- Németh (2003): the variational inequalities on Hadamard manifolds.
- Azagra nad Ferrera (2005): proximal subdifferential for non-smooth functions defined on Riemannian manifolds.
- Li, López and Martín-Márqnez (2009): extended monotone vector fields and the proximal point algorithm to Hadamard manifolds
- Ferreira and Oliveira (2002): proximal point algorithm on Riemannian manifolds
- Azagra, Ferrera and Sanz (2009): fixed points and zeros for set valued mappings on Riemannian manifolds

Weak sharp minima on Riemannian manifolds

M a complete connected m-dimensional Riemannian manifold
T_xM the tangent space of M at x
$\mathbb{B}_x = \{v \in T_xM : \|v\| \leq 1\}$
$\exp_x : T_xM \to M$, the exponential map at x
$\mathbf{B}(x,r)$ the open metric ball at x with radius r
$\overline{\mathbf{B}(x,r)}$ the closed metric ball at x with radius r
$C \subset X$, $\delta_C(x)$ the indicator function and $\sigma_C(x)$ the support function

Weak sharp minima on Riemannian manifolds

- $D \subset M$, D is said to be totally convex if D contains every geodesic γ of M whose endpoints x and y are in D.
- $f : M \to R \cup \{+\infty\}$ is said to be convex if for each $x, y \in M$ and $\gamma \in \Gamma_{xy}$, the composite $f \circ \gamma : R \to R \cup \{+\infty\}$ is a convex function.

Weak sharp minima on Riemannian manifolds

$x \in D$

$T_x D = \{v \in T_x M | \ c : [0, \varepsilon] \to D \quad c(0) = x, \ c'(0) = v\}$

$N_D(x) = \{u \in T_x M | \langle u, v \rangle \leq 0 \quad \forall v \in T_x D\}$

Weak sharp minima on Riemannian manifolds

- directional derivative $f'(x;v) = \lim_{t\to 0^+} \frac{f(\exp_x tv) - f(x)}{t}$
- Subdifferential
  $\partial f(x) = \{\omega \in T_x M : \ f(y) \geq f(x) + \langle \omega, \gamma'_{xy}(0)\rangle \quad \forall y \in M\}$
- Relationship between subdifferential and directional derivative
  - $\partial f(x) = \{\omega \in T_x M : \langle \omega, v\rangle \leq f'(x;v) \quad \forall v \in T_x M\}$.
  - $\sigma_{\partial f(x)}(\cdot)$ is the lower semicontinuous hull of the function $f'(x;\cdot)$

Weak sharp minima on Riemannian manifolds

Theorem

Let D be a closed subset of M and $y \in D$.

- *If $x \in M$ such that $y \in P(x \mid D)$ then*

$$\langle \gamma'_{xy}(1), \gamma'_{yz}(0) \rangle \geq 0 \quad \textit{for each } z \in D \textit{ and } \gamma_{yz} \subset \Gamma^D_{yz}, \quad (3.1)$$

  where $\gamma_{xy} \in \Gamma_{xy}$ is a minimizing geodesic.
- *Conversely, there exists $\varepsilon > 0$ such that for each $x \in \mathbf{B}(y, \varepsilon)$, if (3.1) holds, then $y \in P(x \mid D)$.*

Weak sharp minima on Riemannian manifolds

Let $f : M \to R \cup \{+\infty\}$ be a proper convex function and $S \subset M$ be a nonempty closed totally convex set.
Consider the optimization problem:

$$\mathcal{P}: \quad \text{minimize } f(x) \\ \text{subject to } x \in S.$$

Weak sharp minima on Riemannian manifolds

$f_0 : M \to R \cup \{+\infty\}$ is given by

$$f_0(x) = f(x) + \delta_S(x) = \begin{cases} f(x) & x \in S \\ +\infty & \text{otherwise.} \end{cases}$$

The optimization problem $\mathcal{P}$ can equivalently be stated as the unconstrained optimization problem

$$\mathcal{P}_0 : \quad \begin{array}{l} \text{minimize } f_0(x) \\ \text{subject to } x \in M, \end{array}$$

Following (Burke and Deng (2002)), below we will also assume that

$$\partial f_0(x) = \mathrm{cl}(\partial f(x) + N_S(x)). \tag{3.2}$$

Weak sharp minima on Riemannian manifolds

$\bar{S}$ the solution set of $\mathcal{P}$, i.e.,

$$\bar{S} = \mathrm{argmin}_S f = \{x \in S | \, f(x) = \min_{y \in S} f(y)\}$$

$\bar{x} \in \bar{S}$. $\bar{S}$ is a set of weak sharp minima at $\bar{x}$ for f over the set S with modulus $\alpha > 0$ if there exists $\varepsilon > 0$ such that

$$f(x) \geq f(\bar{x}) + \alpha d_{\bar{S}}(x) \tag{3.3}$$

holds for all $x \in S \cap \mathbf{B}(\bar{x}, \varepsilon)$.

Weak sharp minima on Riemannian manifolds

$\bar{S}$ is

- weak sharp minima for f if (3.3) holds for all $\bar{x} \in \bar{S}$ and $x \in S$.
- local weak sharp minima for f if for each point $\bar{x} \in \bar{S}$, $\bar{S}$ is a set of weak sharp minima at $\bar{x}$ for f over S.
- boundedly weak sharp minima for f if for every bounded set $W \subset M$ with $W \cap \bar{S} \neq \emptyset$ there exists $\alpha_W > 0$ such that

$$f(x) \geq f(\bar{x}) + \alpha_W d_{\bar{S}}(x)$$

holds for all $x \in S \cap W$ and $\bar{x} \in \bar{S}$.

Weak sharp minima on Riemannian manifolds

Theorem

Assume that (3.2) *holds for all* $\bar{x} \in \bar{S}$. *Then the following statements are equivalent:*

(i). $\bar{S}$ *is a set of weak sharp minima for* f *over* $S \subset M$ *with modulus* $\alpha > 0$.

(ii). *For each* $\bar{x} \in \bar{S}$, $\alpha \mathbb{B}^{\circ}_{\bar{x}} \subseteq \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}) + T_{\bar{x}}\bar{S})$.

(iii). *For each* $\bar{x} \in \bar{S}$, $\alpha \mathbb{B}^{\circ}_{\bar{x}} \cap N_{\bar{S}}(\bar{x}) \subseteq \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}) + T_{\bar{x}}\bar{S})$.

(iv). *For each* $\bar{x} \in \bar{S}$,

$\hat{\alpha} \mathbb{B}^{\circ}_{\bar{x}} \subseteq \partial f(\bar{x}) + (T_{\bar{x}} S \cap N_{\bar{S}}(\bar{x}))^{\circ}$ *for each* $\hat{\alpha} \in (0, \alpha)$.

Weak sharp minima on Riemannian manifolds

Theorem

Assume that (3.2) holds for all $\bar{x} \in \bar{S}$. Then the following statements are equivalent:

(i). *$\bar{S}$ is a set of weak sharp minima for f over $S \subset M$ with modulus $\alpha > 0$.*

(ii). *For each $\bar{x} \in \bar{S}$, $f'(\bar{x}; v) \geq \alpha \|v\| \quad$ for each $v \in T_{\bar{x}}S \cap N_{\bar{S}}(\bar{x})$.*

(iii). *Let $x \in S$ and $\bar{x} \in P(x \mid \bar{S})$. Let $\gamma \in \Gamma_{\bar{x}x}$ be a minimizing geodesic. Then,*

$$f'(\bar{x}; \gamma'(0)) \geq \alpha d_{\bar{S}}(x).$$

Weak sharp minima on Riemannian manifolds

Corollary

Assume that (3.2) *holds at each point of* $\bar{S}$. *The following statements are equivalent:*

(i) $\bar{S}$ *is the set of local weak sharp minima for* f *over* S.

(ii) *For every* $z \in M$ *and* $r > 0$ *satisfying* $\overline{\mathbf{B}(z,r)} \cap \bar{S} \neq \emptyset$, *there exists* $\alpha(r) > 0$ *such that*

$$\alpha(r)\mathbb{B}^{o}_{\bar{x}} \cap N_{\bar{S}}(\bar{x}) \subseteq \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}) + T_{\bar{x}}\bar{S}) \quad \forall \bar{x} \in \bar{S} \cap \overline{\mathbf{B}(z,r)}.$$

Weak sharp minima on Riemannian manifolds

Theorem

$\bar{S}$ is the set of boundedly weak sharp minima for f over S if and only if $\bar{S}$ is the set of local weak sharp minima for f over S.

Weak sharp minima on Hadamard manifolds

- M a Hadamard manifold, that is, a complete connected m-dimensional Riemannian manifold with non-positive sectional curvature

Proposition

Sakai (1996) (Comparison theorem for triangles).

- $\Delta(x_1x_2x_3)$ *a geodesic triangle.*
- $\gamma_i : [0, l_i] \to M$ *geodesic segments joining* x_i *to* x_{i+1},
- $l_i := L(\gamma_i)$ *the length of* γ_i
- $\alpha_i := \angle(\gamma_i'(0), -\gamma_{i-1}'(l_{i-1}))$, *where* $i = 1, 2, 3$ *(mod 3).*

Then

- $\alpha_1 + \alpha_2 + \alpha_3 \leq \pi$,
- $l_i^2 + l_{i+1}^2 - 2l_i l_{i+1} cos\alpha_{i+1} \leq l_{i-1}^2$

Weak sharp minima on Hadamard manifolds

Theorem

If D is a totally convex subset of M and $x \in D$. Then

$$\partial d_D(x) = \mathbb{B}_x^{\circ} \cap N_D(x).$$

Weak sharp minima on Hadamard manifolds

Theorem

Assume that $\partial f_0(\bar{x}) = \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}))$ *holds for all* $\bar{x} \in \bar{S}$. *Then the following statements are equivalent:*
(i) $\bar{S}$ *is the set of weak sharp minima for* f *over the* S *with modulus* $\alpha > 0$.
(ii) *For each* $\bar{x} \in \bar{S}$, $\alpha\mathbb{B}^{o}_{\bar{x}} \cap N_{\bar{S}}(\bar{x}) \subseteq \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}))$.

Weak sharp minima on Hadamard manifolds

Theorem

Assume that $\partial f_0(\bar{x}) = \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}))$ *holds for all* $\bar{x} \in \bar{S}$. *Then the following statements are equivalent:*

(i). $\bar{S}$ *is a set of weak sharp minima for the function* f *over the* S *with modulus* α.

(ii). *For each* $\bar{x} \in \bar{S}$ *and* $v \in T_{\bar{x}}S$, $f'(\bar{x}; v) \geq \alpha d_{T_{\bar{x}}\bar{S}}(v)$.

(iii). *For each* $\bar{x} \in \bar{S}$ *and* $v \in T_{\bar{x}}S \cap N_{\bar{S}}(\bar{x})$, $f'(\bar{x}; v) \geq \alpha \|v\|$.

(iv). *For each* $x \in S$, $\bar{x} = P(x \mid \bar{S})$ *and* $v = \exp_{\bar{x}}^{-1} x$, $f'(\bar{x}; v) \geq \alpha d_{\bar{S}}(x)$.

Weak sharp minima on Hadamard manifolds

Corollary

Assume that $\partial f_0(\bar{x}) = \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x}))$ *holds for all* $\bar{x} \in \bar{S}$. *Then the following statements are equivalent:*

(i) $\bar{S}$ *is the set of local weak sharp minima for* f *over* S.

(ii) *For every* $z \in M$ *and* $r > 0$ *satisfying* $\overline{\mathbf{B}(z,r)} \cap \bar{S} \neq \emptyset$, *there exists* $\alpha(r) > 0$ *such that*

$$\alpha(r)\mathbb{B}^{\circ}_{\bar{x}} \cap N_{\bar{S}}(\bar{x}) \subseteq \mathrm{cl}(\partial f(\bar{x}) + N_S(\bar{x})) \quad \forall \bar{x} \in \bar{S} \cap \overline{\mathbf{B}(z,r)}.$$

- J. M. Borwein and A. S. Lewis, Convex analysis and nonlinear optimization, Springer Science+Business Media, Inc, 2006.
- J. V. Burke and S. Deng, Weak sharp minima revisited, part I: basic theory, Control and Cybernetics, **31**, 439-469 (2002).
- J. V. Burke and S. Deng, Weak sharp minima revisited, part II: application to linear regularity and error bounds, Math. Program., Ser. B, 104: 235-261, 2005.
- J. V. Burke, A. Lewis and M. Overton, Optimizing matrix stability, Proc. of the American Math. Soc., 129: 1635-1642, 2000.
- J. V. Burke, A. Lewis and M. Overton, Optimal stability and eigenvalue multiplicity, Found. Comput. Math., 1:205-225, 2001.

- J. V. Burke and M. C. Ferris, Weak sharp minima in mathematical programming, SIAM J. Control and Optim., 31: 1340-1359, 1993.
- J. V. Burke and M. C. Ferris, A Gauss-Newton method for convex composite optimization, Mathematical Programming, 71: 179-194, 1995.
- J. V. Burke and J. J. More, On the identification of active constraints, SIAM J. Numer. Anal., 25: 1197-1211, 1988.
- F. H. Clarke, Optimization and nonsmooth analysis, Wiley, New York, 1983.
- J. Cheeger and D. Gromoll, On the structure of complete manifolds of nonnegative curvature, Ann. of Math. 96, 413-443, (1972).

- L. Cromme, Strong Uniqueness, Numerische Mathematik, 29, 179-193 (1978).
- M. P. DoCarmo, Riemannian Geometry, Birkhauser, Boston, 1992.
- M. C. Ferris, Weak Sharp Minima and Penalty Functions in Mathematical Programming, PhD thesis, University of Cambridge, 1988.
- M. C. Ferris, Iterative linear programming solution of convex programs, Journal of Optimization Theorey and Applications, 65: 53-65, 1990.
- M. C. Ferris, Finite termination of the proximal point algorithm, Mathematical Programming, 50: 359-366, 1991.

M. S. Gowda, An analysis of zero set and global error bound properties of a piecewise affine function via its recession function, SIAM Journal on Matrix Analysis, 17: 594-609, 1996.

S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, Inc, New York, 1978.

R. Henrion and J. Outrata, A subdifferential condition for calmness of multifunctions, Journal of Mathematical Analysis and Applications, 258: 110-130, 2001.

A. Jourani, Hoffman's error bounds, local controllability, and sensitivity analysis, SIAM J. Control and Optim., 38: 947-970, 2000.

A. S. Lewis and J. P. Pang, Error bounds for convex inequality systems, (in J. P. Crouzeix, ed.) Proceedings of the fifth Symposium on Generalized Convexity, Luminy-Marseille, 1996.

- B. T. Polyak, Sharp Minima, Institute of Control Sciences Lecture Notes, Moscow, USSR, 1979; Presented at the IIASA Workshop on Generalized Lagrangians and Their Applications, IIASA, Laxenburg, Austria, 1979.
- C. Udriste, Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and Its Applications, Vol. 297, Kluwer Academic, Dordrecht, 1994.
- T. Sakai, Riemannian Geometry, Translations of Mathematical Monographs 149, American Mathematical Society, 1996.
- R. Walter, (1974), On the metric projections onto convex sets in Riemannian spaces, Arch. Math. 25, 91-98.
- J. J. Ye, New uniform parametric error bounds, Journal of Optimization Theory and Applications, 98: 197-219, 1998.

J. J. Ye and D. L. Zhu, Optimality conditions for bilevel programming problems, Optimization, 33: 9-27, 1995.

J. J. Ye and D. L. Zhu, A note on optimality conditions for bilevel programming problems, Optimization, 39: 361-366, 1997.

J. J. Ye, D. L. Zhu and Q. J. Zhu, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, SIAM Journal on Optimization, 7: 481-507, 1997.

R. T. Rockafellar, Conjugate duality and optimization, SIAM, Philadelphia, Pa., 1974.

J. M. Lee, Riemannian manifolds: an introduction to curvature, Springer-Verlag New York, Inc.

- R. E. Mahony. The constrained Newton method on a Lie group and the symmetric eigenvalue problem, Linear Algebra Appl. 248, 67-89 (1996).
- O.P. Ferreira and P.R. Oliveira, proximal point algorithm on Riemannian manifolds, *Optimization*, 51(2), 257-270, 2002.

Thank you