CARPATHIAN J. MATH. Online version at https://www.carpathian.cunbm.utcluj.ro
36 (2020), No. 1, 141 - 146 Print Edition: ISSN 1584 - 2851 Online Edition: ISSN 1843 - 4401

Dedicated to Prof. Hong-Kun Xu on the occasion of his 60^{th} anniversary

Inexact descent methods for convex minimization problems in Banach spaces

SIMEON REICH and ALEXANDER J. ZASLAVSKI

ABSTRACT. Given a Lipschitz and convex objective function of an unconstrained optimization problem, defined on a Banach space, we revisit the class of regular vector fields which was introduced in our previous work on descent methods. We study, in particular, the asymptotic behavior of the sequence of values of the objective function for a certain inexact process generated by a regular vector field when the sequence of computational errors converges to zero and show that this sequence of values converges to the infimum of the given objective function of the unconstrained optimization problem.

1. INTRODUCTION

Given a Lipschitz and convex objective function of an unconstrained optimization problem, defined on a Banach space, we consider a certain complete metric space of vector fields, which are self-mappings of the Banach space, with the topology of uniform convergence on bounded subsets. With each such vector field, we associate a certain iterative process. In our previous work [13, 14] we introduced the class of regular vector fields and showed, using the generic approach and the porosity notion, that a typical vector field is regular and that for a regular vector field, the sequence of values of the objective function evaluated at the points generated by our process tends to the infimum of the given objective function of the unconstrained optimization problem. In the present paper we study the behavior of the values of the objective function for an *inexact* process generated by a regular vector field and show that the sequence of values of the given objective function still converges to the infimum of this objective function.

Assume that $(X, \|\cdot\|)$ is a Banach space with norm $\|\cdot\|$, $(X^*, \|\cdot\|_*)$ is its dual space with the corresponding dual norm $\|\cdot\|_*$ and that $f : X \to R^1$ is a convex continuous function which is bounded from below. Recall that for each pair of sets $A, B \subset X^*$,

$$H(A,B) = \max\{\sup_{x\in A}\inf_{y\in B}\|x-y\|_*,\ \sup_{y\in B}\inf_{x\in A}\|x-y\|_*\}$$

is the Pompeiu-Hausdorff distance between A and B.

For each point $x \in X$, let

$$\partial f(x) = \{l \in X^* :\ f(y) - f(x) \geq l(y-x) \text{ for all } y \in X\}$$

be the subdifferential of f at x [8, 11, 19]. It is well known that the set $\partial f(x)$ is a nonempty and bounded subset of $(X^*, \|\cdot\|_*)$.

Set

$$\inf(f) := \inf\{f(x) :\ x \in X\}.$$

Received: 22.10.2019. In revised form: 12.01.2020. Accepted: 19.01.2020

2010 *Mathematics Subject Classification.* 49M37, 90C25, 90C30.

Key words and phrases. *Banach space, complete metric space, convex function, descent method, inexact iterate, iterative process.*

Corresponding author: Simeon Reich; sreich@technion.ac.il

Denote by $\mathcal{A}$ the set of all mappings $V : X \to X$ such that V is bounded on every bounded subset of X (that is, for each $K_0 > 0$, there is $K_1 > 0$ such that $\|Vx\| \leq K_1$ if $\|x\| \leq K_0$), and for each $x \in X$ and each $l \in \partial f(x)$, $l(Vx) \leq 0$. We denote by $\mathcal{A}_c$ the set of all continuous $V \in \mathcal{A}$, by $\mathcal{A}_u$ the set of all $V \in \mathcal{A}$ which are uniformly continuous on each bounded subset of X, and by $\mathcal{A}_{au}$ the set of all $V \in \mathcal{A}$ which are uniformly continuous on the subsets

$$\{x \in X : \|x\| \leq n \text{ and } f(x) \geq \inf(f) + 1/n\}$$

for each integer $n \geq 1$. Finally, we let $\mathcal{A}_{auc} = \mathcal{A}_{au} \cap \mathcal{A}_c$.

Next, we endow the set $\mathcal{A}$ with a metric ρ: For each $V_1, V_2 \in \mathcal{A}$ and each integer $i \geq 1$, we first set

$$\rho_i(V_1, V_2) := \sup\{\|V_1x - V_2x\| : x \in X \text{ and } \|x\| \leq i\},$$

and then define

$$\rho(V_1, V_2) := \sum_{i=1}^{\infty} 2^{-i}[\rho_i(V_1, V_2)(1 + \rho_i(V_1, V_2))^{-1}].$$

Clearly $(\mathcal{A}, \rho)$ is a complete metric space. It is also not difficult to see that the collection of the sets

$$E(N, \epsilon) = \{(V_1, V_2) \in \mathcal{A} \times \mathcal{A} : \|V_1x - V_2x\| \leq \epsilon, \ x \in X, \ \|x\| \leq N\},$$

where $N, \epsilon > 0$, is a basis for the uniformity generated by the metric ρ. Evidently $\mathcal{A}_c$, $\mathcal{A}_u$, $\mathcal{A}_{au}$ and $\mathcal{A}_{auc}$ are all closed subsets of the metric space $(\mathcal{A}, \rho)$. In the sequel we assign to all these spaces the same metric ρ.

In order to compute $\inf(f)$, we associate in Section 2 with each vector field $W \in \mathcal{A}$ a gradient-like iterative process (see below).

At this point we recall that the study of minimization methods for convex functions is a central topic in optimization theory and its applications. See, for example, [1, 2, 3, 4, 5, 6, 9, 10, 12, 20, 22] and the references mentioned therein. Note, in particular, that the counterexample studied in Section 2.2 of Chapter VIII of [8] shows that, even for two-dimensional problems, the simplest choice for a descent direction, namely the normalized steepest descent direction,

$$V(x) = \text{argmin}\left\{\max_{l \in \partial f(x)} \langle l, d\rangle : \|d\| = 1\right\},$$

may produce sequences the functional values of which fail to converge to the infimum of the objective function f considered in [8], which attains its minimum. This vector field V, defined above and discussed in Section 2.2 of Chapter VIII of [8], belongs to our space of vector fields. The steepest descent scheme (Algorithm 1.1.7) presented in Section 1.1 of Chapter VIII of [8] corresponds to the iterative process we consider below.

In infinite-dimensional settings the problem is even more difficult and less understood. Moreover, positive results usually require special assumptions on the space and on the functions. However, in [13] (under certain assumptions on the objective function f), for an arbitrary Banach space X we established the existence of a set $\mathcal{F}$ which is a countable intersection of open and everywhere dense subsets of $\mathcal{A}$ such that for any $V \in \mathcal{F}$, the sequence of values of f tends to its infimum for the iterative process associated with V.

In [14] we introduced the class of regular vector fields $V \in \mathcal{A}$ and showed (under the two mild assumptions A(i) and A(ii) on f stated below) that the complement of the set of regular vector fields is not only of the first category, but also σ-porous in each of the spaces $\mathcal{A}$, $\mathcal{A}_c$, $\mathcal{A}_u$, $\mathcal{A}_{au}$ and $\mathcal{A}_{auc}$. We then showed in [14] that for any regular vector field $V \in \mathcal{A}_{au}$, the values of f tend to its infimum for the process associated with V. If, in addition to A(i) and A(ii), f also satisfies assumption A(iii), then this convergence result is valid for any regular $V \in \mathcal{A}$. Note that the results of [14] are also presented in Chapter

8 of the book [16], which contains many other generic and porosity results. For more applications of the generic approach and the porosity notion in optimization theory, see also [21].

Our results are established in any Banach space and for those convex functions which satisfy the following two assumptions.

A(i) There exists a norm-bounded set $X_0 \subset X$ such that

$$\inf(f) = \inf\{f(x) : \ x \in X\} = \inf\{f(x) : \ x \in X_0\};$$

A(ii) for each $r > 0$, the function f is Lipschitz on the ball $\{x \in X : \ \|x\| \leq r\}$.

We may assume that the set X_0 in A(i) is closed and convex.

It is clear that assumption A(i) holds if $\lim_{\|x\| \to \infty} f(x) = \infty$.

We say that a mapping $V \in \mathcal{A}$ is *regular* if for any natural number n, there exists a positive number $\delta(n)$ such that for each point $x \in X$ satisfying

$$\|x\| \leq n \text{ and } f(x) \geq \inf(f) + 1/n,$$

and each bounded linear functional $l \in \partial f(x)$, we have

$$l(Vx) \leq -\delta(n).$$

In this connection, see also [15].

Note that a regular vector field was constructed in Section 8.2 of [16]. Conditions for regularity of vector fields can be found in Section 8.14 of [16].

We denote by $\mathcal{F}$ the set of all regular vector fields $V \in \mathcal{A}$.

It was also shown in [14] that $\mathcal{G} := \mathcal{A} \setminus \mathcal{F}$ is a face of the convex cone $\mathcal{A}$ in the sense that if a non-trivial convex combination of two vector fields in $\mathcal{A}$ belongs to $\mathcal{G}$, then both of them must belong to $\mathcal{G}$.

In the sequel we also make use of the following assumption:

A(iii) For each integer $n \geq 1$, there exists $\delta > 0$ such that for each $x_1, x_2 \in X$ satisfying

$$\|x_1\|, \|x_2\| \leq n, \ f(x_i) \geq \inf(f) + 1/n, \ i = 1, 2, \text{ and } \|x_1 - x_2\| \leq \delta,$$

the following inequality holds:

$$H(\partial f(x_1), \partial f(x_2)) \leq 1/n.$$

This assumption is certainly satisfied if the function f is differentiable and its derivative is uniformly continuous on those bounded subsets of X over which the infimum of f is larger than $\inf(f)$.

2. Main result

For each point $x \in X$ and each number $r > 0$, set

$$B(x, r) := \{y \in X : \ \|x - y\| \leq r\}.$$

Let $W \in \mathcal{A}$. We associate with W the following iterative process.

For each point $x \in X$, denote by $Q_W(x)$ the set of all points

$$y \in \{x + \alpha W x : \ \alpha \in [0, 1]\}$$

such that

$$f(y) = \inf\{f(x + \beta W x) : \ \beta \in [0, 1]\}.$$

Given any initial point $x_0 \in X$, one can construct a sequence $\{x_i\}_{i=0}^{\infty} \subset X$ such that for all nonnegative integers i,

$$x_{i+1} \in Q_W(x_i).$$

This process and its convergence were studied in [13, 14]. In particular, in [14] it is shown that if W is regular, then $\lim_{n \to \infty} f(x_n) = \inf(f)$. In the present paper we establish an extension of this result for *inexact* iterates of the algorithm.

Let $x \in X$ and $\delta \geq 0$. Denote by $Q_{W,\delta}(x)$ the set of all $z \in X$ for which there exist a number $\lambda \in [0,1]$ and a point $y \in X$ such that

$$\|y - Wx\| \leq \delta,$$

$$f(x + \lambda y) \leq f(x + \beta y) + \delta \text{ for all } \beta \in [0,1],$$

$$f(x + \lambda y) \leq f(x)$$

and

$$z = x + \lambda y.$$

In other words,

$$Q_{W,\delta}(x) := \{z \in X : \text{ there exist } \lambda \in [0,1] \text{ and } y \in B(Wx, \delta) \text{ such that}$$

$$z = x + \lambda y \text{ and } f(z) \leq f(x + \beta y) + \delta \text{ for all } \beta \in [0,1] \text{ and } f(z) \leq f(x)\}. \tag{2.1}$$

Note that the set $Q_{W,\delta}(x)$ may be empty.

In Section 3 we establish the following result.

Theorem 2.1. *Assume that a vector field $V \in \mathcal{A}$ is regular, assumptions A(i) and A(ii) are valid and that at least one of the following conditions holds: 1. $V \in \mathcal{A}_{au}$; 2. A(iii) is valid.*

Let a sequence of nonnegative numbers $\{\delta_i\}_{i=0}^{\infty}$ satisfy

$$\lim_{i \to \infty} \delta_i = 0 \tag{2.2}$$

and let a sequence $\{x_i\}_{i=0}^{\infty} \subset X$ be such that

$$\liminf_{i \to \infty} \|x_i\| < \infty$$

and for each integer $i \geq 0$,

$$\text{if } Q_{V,\delta_i}(x_i) \neq \emptyset, \text{ then } x_{i+1} \in Q_{V,\delta_i}(x_i); \tag{2.3}$$

otherwise

$$x_{i+1} = x_i. \tag{2.4}$$

Then

$$\lim_{i \to \infty} f(x_i) = \inf(f).$$

Note that Theorem 2.1 is an extension of Theorem 2.1 of [18], which was established in the case where the sequence $\{x_i\}_{i=0}^{\infty}$ is *bounded*.

3. Proof of Theorem 2.1

We first recall the following lemma, which was proved in [17].

Lemma 3.1. *Assume that $W \in \mathcal{A}$ is regular, A(i), A(ii) are valid and that at least one of the following conditions holds: 1. $W \in \mathcal{A}_{au}$; 2. A(iii) is valid.*

Let $\bar{K}$ and $\bar{\epsilon}$ be positive. Then there exist positive numbers $\bar{\alpha}, \gamma$ and δ such that for each point $x \in X$ satisfying

$$\|x\| \leq \bar{K}, \; f(x) \geq \inf(f) + \bar{\epsilon},$$

each number $\beta \in (0, \bar{\alpha}]$, and each point $y \in B(Wx, \delta)$, we have

$$f(x) - f(x + \beta y) \geq \beta\gamma.$$

.

Proof. By assumption, there exist a strictly increasing sequence of natural numbers $\{n_p\}_{p=1}^{\infty}$ and a number $K > 0$ such that

$$\|x_{n_p}\| \le K,\ p = 1, 2, \ldots. \tag{3.5}$$

By (2.1), (2.3) and (2.4), for all integers $i \ge 0$, we have

$$f(x_{i+1}) \le f(x_i) \le f(x_0). \tag{3.6}$$

Let $\epsilon \in (0, 1)$. In view of (3.6), in order to complete the proof of Theorem 2.1, it is sufficient to show that there exists an integer $n \ge 0$ such that

$$f(x_n) \le \inf(f) + \epsilon.$$

Lemma 3.1 implies that there exist numbers

$$\alpha \in (0, 1), \gamma > 0 \text{ and } \delta > 0$$

such that the following property holds:

(a) for each point $x \in X$ satisfying

$$\|x\| \le K,\ f(x) \ge \inf(f) + \epsilon,$$

each number $\beta \in (0, \alpha]$ and each point $y \in B(Vx, \delta)$,

$$f(x) - f(x + \beta y) \ge \beta\gamma.$$

In view of (2.2), there exists a natural number p_1 such that for all integers $i \ge n_{p_1}$, we have

$$\delta_i \le \min\{2^{-1}\delta,\ 2^{-1}\alpha\gamma\}. \tag{3.7}$$

Fix a natural number

$$p_2 > p_1 + 4 + 2(\alpha\gamma)^{-1}(f(x_0) - \inf(f)). \tag{3.8}$$

We claim that there exists a natural number $i \in [n_{p_1}, n_{p_2}]$ such that

$$f(x_i) \le \inf(f) + \epsilon.$$

Suppose to the contrary that

$$f(x_i) > \inf(f) + \epsilon,\ i = n_{p_1}, \ldots, n_{p_2}. \tag{3.9}$$

Let $p \in \{p_1, \ldots, p_2\}$. Property (a), (2.1), (2.3), (3.5), (3.7) and (3.9) imply that

$$Q_{V,\delta_{n_p}}(x_{n_p}) \ne \emptyset, \qquad x_{n_p+1} \in Q_{V,\delta_{n_p}}(x_{n_p})$$

and that there exists

$$y_p \in B(Vx_{n_p}, \delta_{n_p})$$

such that

$$f(x_{n_p+1}) \le f(x_{n_p} + \alpha y_p) + \delta_{n_p} \le f(x_{n_p}) - \alpha\gamma + \delta_{n_p} \le f(x_{n_p}) - \alpha\gamma/2. \tag{3.10}$$

It follows from (3.6) and (3.10) that

$$f(x_{n_{p+1}}) \le f(x_{n_p+1}) \le f(x_{n_p}) - \alpha\gamma/2. \tag{3.11}$$

In view of (3.6) and (3.11), we have

$$f(x_0) - \inf(f) \ge f(x_{n_{p_1}}) - f(x_{n_{p_2}}) = \sum_{p=p_1}^{p_2-1} (f(x_{n_p}) - f(x_{n_{p+1}})) \ge (p_2 - p_1)\alpha\gamma/2$$

and

$$p_2 \le p_1 + 2(\alpha\gamma)^{-1}(f(x_0) - \inf(f)).$$

This, however, contradicts (3.8). The contradiction we have reached yields our claim and completes the proof of Theorem 2.1. □

4. Conclusions

We have considered an unconstrained optimization problem with a Lipschitz and convex objective function on a general Banach space, and a certain inexact process generated by a regular vector field associated with the given objective function. We have shown that if the sequence of computational errors converges to zero, then the sequence of values of the objective function for this inexact process converges to the infimum of the given objective function.

Acknowledgments. The first author was partially supported by the Israel Science Foundation (Grant No. 820/17), by the Fund for the Promotion of Research at the Technion and by the Technion General Research Fund. Both authors are grateful to the two referees for their useful comments and helpful suggestions.

References

[1] Bačák, M., *A variational approach to stochastic minimization of convex functionals*, Pure Appl. Funct. Anal., **3** (2018), 287–295
[2] Cegielski, A., *A method of projection onto an acute cone with level control in convex minimization*, Math. Program., **85** (1999), 469–490
[3] Curry, H. B., *The method of steepest descent for nonlinear minimization problems*, Quarterly Appl. Math., **2** (1944), 258–261
[4] Gibali, A., Küfer, K.-H., Reem, D. and Süss, P., *A generalized projection-based scheme for solving convex constrained optimization problems*, Comput. Optim. Appl., **70** (2018), 737–762
[5] Grad, S.-M. and Pop, E.-L., *Vector duality for convex vector optimization problems by means of the quasi-interior of the ordering cone*, Optimization, **63** (2014), 21–37
[6] Griva, I., *Convergence analysis of augmented Lagrangian-fast projected gradient method for convex quadratic problems*, Pure Appl. Funct. Anal., **3** (2018), 417–428
[7] He, H., Hou, L. and Xu, H.-K., *A partially isochronous splitting algorithm for three-block separable convex minimization problems*, Adv. Comput. Math., **44** (2018), 1091–1115
[8] Hiriart-Urruty, J.-B. and Lemaréchal, C., *Convex analysis and minimization algorithms*, Springer, Berlin, 1993
[9] Hu, Y., Li, C. and Yang, X., *On convergence rates of linearized proximal algorithms for convex composite optimization with applications*, SIAM J. Optim., **26** (2016), 1207–1235
[10] Konnov, I., *On convergence properties of a subgradient method*, Optim. Methods Softw., **18** (2003), 53–62
[11] Mordukhovich, B. S., *Variational analysis and generalized differentiation. I: Basic theory*, Springer, Berlin, 2006
[12] Polyak, R., *Exterior distance function*, Pure Appl. Funct. Anal., **2** (2017), 369–394
[13] Reich, S. and Zaslavski, A. J., *Generic convergence of descent methods in Banach spaces*, Math. Oper. Research, **25** (2000), 231–242
[14] Reich, S. and Zaslavski, A. J., *The set of divergent descent methods in a Banach space is σ-porous*, SIAM J. Optim., **11** (2001), 1003–1018
[15] Reich, S. and Zaslavski, A. J., *Regular vector-fields in Banach spaces*, Taiwanese J. Math., **12** (2008), 1165–1176
[16] Reich, S. and Zaslavski, A. J., *Genericity in nonlinear analysis*, Developments in Mathematics, **34**, Springer, New York, 2014
[17] Reich, S. and Zaslavski, A. J., *Descent methods with computational errors in Banach spaces*, Optimization, **14** (2019), 41–50
[18] Reich, S. and Zaslavski, A. J., *Inexact descent methods with convex objective functions in Banach spaces*, J. Nonlinear Convex Anal., **20** (2019), 1875–1882
[19] Zălinescu, C., *Convex analysis in general vector spaces*, World Scientific, New Jersey- London-Singapore-Hong Kong, 2002
[20] Zălinescu, C., *On Lagrange multipliers in convex entropy minimization*, Pure Appl. Funct. Anal., **3** (2018), 393–402
[21] Zaslavski, A. J., *Optimization on metric and normed spaces*, Springer, New York, 2010
[22] Zaslavski, A. J., *Numerical optimization with computational errors*, Springer, Cham, 2016

Department of Mathematics
The Technion – Israel Institute of Technology
32000 Haifa, Israel
E-mail address: sreich@technion.ac.il
E-mail address: ajzasl@technion.ac.il