Linear and Nonlinear Analysis
Yokohama Publishers
ISSN 2188-8167

Volume 4, Number 3, 2018, 451–457

DESCENT METHODS IN THE PRESENCE OF COMPUTATIONAL ERRORS IN BANACH SPACES

SIMEON REICH* AND ALEXANDER J. ZASLAVSKI

Abstract. Given a convex objective function on a Banach space, which is Lipschitz on bounded sets and satisfies a coercivity growth condition, we continue to study the class of regular vector fields introduced in our earlier work on descent methods. We analyze, in particular, the behavior of the values of the objective function for an iterative process generated by a regular vector field in the presence of computational errors and show that if the computational errors are small enough, then the values of the objective function approach its infimum.

1. Introduction

Given a Lipschitz convex and coercive objective function on a Banach space, we consider a 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 [12, 13] 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 values of the objective function at the points generated by our process tend to its infimum. As a matter of fact, in [12, 13] we considered *two* iterative processes generated by a regular vector field. In the present paper, taking into account computational errors, we study the behavior of the values of the objective function for one of these processes and show that if the computational errors are small enough, then the values of the objective function approach its infimum (see Theorem 2.1 below). We remark in passing that an analogous result for the second iterative process has already been obtained in [16].

Assume that $(X, \|\cdot\|)$ is a Banach space endowed with the norm $\|\cdot\|$, $(X^*, \|\cdot\|_*)$ is its dual space with the norm $\|\cdot\|_*$ and $f : X \to R^1$ 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 Hausdorff distance between A and B.

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

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

*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.

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

Set

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

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, 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

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

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

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

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 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 certain gradient-like iterative process.

At this point we recall that the study of minimization methods for convex functions is a central topic in optimization theory. See, for example, [1–6, 8, 9, 11, 17, 19] and the references mentioned therein. Note, in particular, that the counterexample studied in Section 2.2 of Chapter VIII of [7] shows that, even for two-dimensional problems, the simplest choice for a descent direction, namely the normalized steepest descent direction,

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

may produce sequences the functional values of which fail to converge to the infimum of f. This vector field V belongs to $\mathcal{A}$ and the Lipschitz function f attains its infimum. The steepest descent scheme (Algorithm 1.1.7) presented in Section 1.1 of Chapter VIII of [7] 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 [12] (under certain assumptions on the 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 values of f tend to its infimum for the process associated with V.

In [13] 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 Baire 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 [13] 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), the function f also satisfies assumption A(iii). Note that the results of [13] are also presented in Chapter 8 of the book [15], which contains many other generic and porosity results. For more applications of the generic approach and the porosity notion in optimization theory, see also [18].

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 $l \in \partial f(x)$, we have

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

In this connection, see also [14]. We denote by $\mathcal{F}$ the set of all regular vector fields $V \in \mathcal{A}$.

It is not difficult to verify the following property of regular vector fields. It means, in particular, 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 $\mathcal{G}$.

Proposition 1.1. *Assume that $V_1, V_2 \in \mathcal{A}$, V_1 is regular, $\phi : X \to [0,1]$, and that for each integer $n \geq 1$,*

$$\inf\{\phi(x) : \ x \in X \text{ and } \|x\| \leq n\} > 0.$$

Then the mapping $x \mapsto \phi(x)V_1x + (1-\phi(x))V_2x$, $x \in X$, also belongs to $\mathcal{F}$.

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 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. The main result

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

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

Let $W \in \mathcal{A}$ and let a sequence $\{a_i\}_{i=0}^{\infty} \subset (0,1]$ satisfy

$$\lim_{i\to\infty} a_i = 0, \ \sum_{i=1}^{\infty} a_i = \infty. \tag{2.1}$$

We associate with W the following iterative process. For each initial point $x_0 \in X$, we construct a sequence $\{x_i\}_{i=0}^{\infty} \subset X$ according to the following rule:

$$x_{i+1} = x_i + a_i W(x_i) \text{ if } f(x_i + a_i W(x_i)) < f(x_i),$$

$$x_{i+1} = x_i \text{ otherwise},$$

where $i = 0, 1, \dots$. This process and its convergence were studied in [12, 13]. In particular, in [13] it is shown that if W is regular, then $\lim_{n\to\infty} f(x_n) = \inf(f)$. More precisely, it is shown in [13] that if $V \in \mathcal{A}$ is regular, $\epsilon > 0$ and $W \in \mathcal{A}$ belongs to a sufficiently small neighborhood of V, then $f(x_n) \le \inf(f) + \epsilon$ for all sufficiently large natural numbers n.

In the present paper, taking into account computational errors, we study the behavior of the values of the objective function for another process generated by a regular vector field and show that if the computational errors are small enough, then the values of the objective functions approach its infimum.

Let $x \in X$, $\delta \ge 0$ and let $i \ge 0$ be an integer. Define

$$P_{W,\delta,i}(x) := \{y \in X : \ \text{there exists } z \in B(Wx,\delta) \text{ such that}$$

$$y = x + a_i z \text{ and } f(y) \le f(x)\}. \tag{2.2}$$

Note that this set may well be empty.

We are now ready to state our main result. It is established in Section 4. Section 3 contains an auxiliary result.

Theorem 2.1. *Assume that $f(x) \to \infty$ as $\|x\| \to \infty$, the vector field $V \in \mathcal{A}$ is regular, assumption A(ii) is valid and that at least one of the following conditions holds: 1. $V \in \mathcal{A}_{au}$; 2. A(iii) is valid.*

Let $K, \epsilon > 0$ be given. Then there exist $\delta > 0$ and a natural number N_0 such that for each sequence $\{x_i\}_{i=0}^{\infty} \subset X$ which satisfies

$$\|x_0\| \le K$$

and for each $i = 0, 1, \dots,$

$$x_{i+1} \in P_{V,\delta,i}(x_i) \text{ if } P_{V,\delta,i}(x_i) \ne \emptyset,$$

$$x_{i+1} = x_i \text{ otherwise},$$

the inequality $f(x_i) \le \inf(f) + \epsilon$ holds for all integers $i \ge N_0$.

3. An auxiliary result

In the proof of Theorem 2.1 we use the following lemma which was proved in [16].

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

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

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

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

4. Proof of Theorem 2.1

We may assume without any loss of generality that $\epsilon < 1$, $K > 2$ and that

$$\{x \in X :\ f(x) \le \inf(f) + 4\} \subset B(0, K-2). \tag{4.1}$$

Let

$$K_0 > \sup\{f(x) :\ x \in B(0, K+1)\} \tag{4.2}$$

and set

$$E_0 := \{x \in X :\ f(x) \le K_0 + 1\}. \tag{4.3}$$

It is clear that the set E_0 is bounded and closed. Choose

$$K_1 > \sup\{\|x\| :\ x \in E_0\} + 1 + K. \tag{4.4}$$

Lemma 3.1 implies that there exist positive numbers $\bar{\alpha} \in (0,1)$, γ and δ such that the following property holds:

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

$$\|x\| \le K_1,\ f(x) \ge \inf(f) + \epsilon/4,$$

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

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

In view of (2.1), there exists a natural number N_1 such that

$$a_i < \bar{\alpha} \text{ for all integers } i \ge N_1. \tag{4.5}$$

Choose a natural number $N_0 > N_1 + 2$ such that

$$\sum_{i=N_1}^{N_0-1} a_i > \gamma^{-1}(K_0 - \inf(f)). \tag{4.6}$$

Assume that a sequence $\{x_i\}_{i=0}^{\infty} \subset X$ satisfies

$$\|x_0\| \le K \tag{4.7}$$

and for each $i = 0, 1, \dots,$

$$x_{i+1} \in P_{V,\delta,i}(x_i) \text{ if } P_{V,\delta,i}(x_i) \ne \emptyset, \tag{4.8}$$

$$x_{i+1} = x_i \text{ otherwise.} \tag{4.9}$$

By (2.2), (4.8) and (4.9),

$$f(x_{i+1}) \leq f(x_i),\ i = 0, 1, \ldots. \tag{4.10}$$

It follows from (4.2), (4.4), (4.7) and (4.10) that for all integers $i \geq 0$,

$$f(x_i) \leq K_0,\ \|x_i\| < K_1. \tag{4.11}$$

In order to complete the proof of the theorem it is sufficient to show that

$$f(x_{N_0}) \leq \inf(f) + \epsilon.$$

Suppose to the contrary that this does not hold. Then for all integers $i = 0, \ldots, N_0$, we have

$$f(x_i) > \inf(f) + \epsilon. \tag{4.12}$$

By (4.5), (4.11) and (4.12), for all integers $i \in [N_1, N_0)$, there exists a point

$$y \in B(Vx_i, \delta)$$

such that

$$x_{i+1} = x_i + a_i y$$

and

$$f(x_{i+1}) \leq f(x_i) - a_i\gamma. \tag{4.13}$$

It follows from (4.11) and (4.13) that

$$K_0 - \inf(f) \geq f(x_{N_1}) - f(x_{N_0}) = \sum_{i=N_1}^{N_0-1} (f(x_i) - f(x_{i+1})) \geq \gamma \sum_{i=N_1}^{N_0-1} a_i$$

and

$$\sum_{i=N_1}^{N_0-1} a_i \leq \gamma^{-1}(K_0 - \inf(f)).$$

This, however, contradicts (4.6). The contradiction we have reached completes the proof of Theorem 2.1.

References

[1] M. Bačák, *A variational approach to stochastic minimization of convex functionals*, Pure and Applied Functional Analysis **3** (2018), 287–295.

[2] A. Cegielski, *A method of projection onto an acute cone with level control in convex minimization*, Mathematical Programming **85** (1999), 469–490.

[3] H. B. Curry, *The method of steepest descent for nonlinear minimization problems*, Quarterly Appl. Math. **2** (1944), 258–261.

[4] A. Gibali, K.-H. Küfer, D. Reem and P. Süss, *A generalized projection-based scheme for solving convex constrained optimization problems*, Comput. Optim. Appl. **70** (2018), 737–762.

[5] S.-M. Grad and E.-L. Pop, *Vector duality for convex vector optimization problems by means of the quasi-interior of the ordering cone*, Optimization **63** (2014), 21–37.

[6] I. Griva, *Convergence analysis of augmented Lagrangian – fast projected gradient method for convex quadratic problems*, Pure and Applied Functional Analysis **3** (2018), 417–428.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms*, Springer, Berlin, 1993.

[8] Y. Hu, C. Li and X. Yang, *On convergence rates of linearized proximal algorithms for convex composite optimization with applications*, SIAM J. Optim. **26** (2016), 1207–1235.

[9] I. Konnov, *On convergence properties of a subgradient method*, Optim. Methods Softw. **18** (2003), 53–62.

[10] B. S. Mordukhovich, *Variational analysis and generalized differentiation. I: Basic theory*, Springer, Berlin, 2006.

[11] R. Polyak, *Exterior distance function*, Pure and Applied Functional Analysis **2** (2017), 369–394.

[12] S. Reich and A. J. Zaslavski, *Generic convergence of descent methods in Banach spaces*, Math. Oper. Research **25** (2000), 231–242.

[13] S. Reich and A. J. Zaslavski, *The set of divergent descent methods in a Banach space is σ-porous*, SIAM J. Optim. **11** (2001), 1003–1018.

[14] S. Reich and A. J. Zaslavski, *Regular vector-fields in Banach spaces*, Taiwanese J. Math. **12** (2008), 1165–1176.

[15] S. Reich and A. J. Zaslavski, *Genericity in nonlinear analysis*, Developments in Mathematics, **34**, Springer, New York, 2014.

[16] S. Reich and A. J. Zaslavski, *Descent methods with computational errors in Banach spaces*, submitted.

[17] C. Zălinescu, *On Lagrange multipliers in convex entropy minimization*, Pure and Applied Functional Analysis **3** (2018), 393–402.

[18] A. J. Zaslavski, *Optimization on metric and normed spaces*, Springer, New York, 2010.

[19] A. J. Zaslavski, *Numerical optimization with computational errors*, Springer, New York, 2016.

Manuscript received 28 October 2018
revised 22 November 2018

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

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