J. Appl. Numer. Optim. 3 (2021), No. 1, pp. 231-238
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.3.2021.1.14

CONVERGENCE OF AN ITERATIVE PROCESS GENERATED BY A REGULAR VECTOR FIELD

ALEXANDER J. ZASLAVSKI

Department of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel

Abstract. Given a convex objective function on a Banach space which is Lipschitz on bounded sets and satisfies a coercivity growth condition, we consider an iterative process generated by a regular vector field, under the presence of computational errors. We show that if the computational errors are small enough, then the values of the objective function become close to its infimum.

Keywords. Banach space; Complete metric space, Convex function, Descent method; Iterative process.

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. The class of regular vector fields was introduced in [1, 2], where it was shown, 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 some iterative process tend to its infimum. Taking into account computational errors, we study in the present paper the behavior of the values of the objective function for an iterative processes generated by a regular vector field and show that if the computational errors are small enough, then the values of the objective functions become close to its infimum.

Assume that $(X, \|\cdot\|)$ is a Banach space with 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.

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

E-mail address: ajzasl@technion.ac.il.
Received August 8, 2020; Accepted November 12, 2020.
©2021 Journal of Applied and Numerical Optimization

Set

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

Denote by $\mathscr{A}$ 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 $\|V(x)\| \le K_1$ if $\|x\| \le K_0$), and for each $x \in X$ and each $l \in \partial f(x)$, $l(V(x)) \le 0$. We denote by $\mathscr{A}_c$ the set of all continuous $V \in \mathscr{A}$, by $\mathscr{A}_u$ the set of all $V \in \mathscr{A}$ which are uniformly continuous on each bounded subset of X, and by $\mathscr{A}_{au}$ the set of all $V \in \mathscr{A}$ which are uniformly continuous on the subsets

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

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

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

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

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

$$E(N, \varepsilon) = \{(V_1, V_2) \in \mathscr{A} \times \mathscr{A} : \|V_1(x) - V_2(x)\| \le \varepsilon,\ x \in X,\ \|x\| \le N\},$$

where $N, \varepsilon > 0$, is a basis for the uniformity generated by the metric ρ. Evidently, $\mathscr{A}_c$, $\mathscr{A}_u$, $\mathscr{A}_{au}$ and $\mathscr{A}_{auc}$ are closed subsets of the metric space $(\mathscr{A}, \rho)$. 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 $W \in \mathscr{A}$ a gradient-like iterative process.

The study of minimization methods for convex functions is a central topic in optimization theory. See, for example, [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] and the references mentioned therein. Note, in particular, that the counterexample studied in Section 2.2 of Chapter VIII of [16] 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.

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 [1] (under certain assumptions on the function f), for an arbitrary Banach space X, it was established the existence of a set $\mathscr{F}$, which is a countable intersection of open everywhere dense subsets of $\mathscr{A}$ such that, for any $V \in \mathscr{F}$, the values of f tend to its infimum for some iterative process associated with V.

In [2], it was introduced the class of regular vector fields $V \in \mathscr{A}$ and it was shown (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 $\mathscr{A}$, $\mathscr{A}_c$, $\mathscr{A}_u$, $\mathscr{A}_{au}$ and $\mathscr{A}_{auc}$. It was shown in [2] that, for any regular vector field $V \in \mathscr{A}_{au}$, the values of f tend to its infimum for some iterative process associated with V if, in addition to

A(i) and A(ii), f also satisfies assumption A(iii). Note that the results of [2] are also presented in Chapter 8 of the book [17], 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].

These results established in any Banach space and the convex functions satisfying the following two assumptions are also used in this paper.

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 $V \in \mathscr{A}$ 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

$$l(V(x)) \leq -\delta(n).$$

In this connection, we refer to [19].

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

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

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

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

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 $i = 0, 1, \ldots$. This process and its convergence were studied in [1, 2]. In particular, in [2], it was shown that if W is regular, then $\lim_{n\to\infty} f(x_n) = \inf(f)$. More precisely, it was shown in [2] that if $V \in \mathscr{A}$ is regular, $\varepsilon > 0$ and $W \in \mathscr{A}$ belongs to a sufficiently small neighborhood of V, then $f(x_n) \leq \inf(f) + \varepsilon$ for all sufficiently large natural numbers n. Taking into account computational errors, in [20], it was studied the behavior of the values of the objective function for this process generated by a regular vector field and it was shown that if the computational errors are small enough, then the values of the objective functions become close to its infimum.

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

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

In this paper, 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:

$$x_{i+1} = x_i + a_i W(x_i),$$

where $i = 0,1,\ldots$. This process is simpler than the process considered in [1, 2, 20] but its iterates do not satisfy the inequality

$$g(x_{i+1}) \le g(x_i), \ i = 0,1,2\ldots.$$

This makes the analysis of its convergence more difficult. Here, we study its convergence taking into account computational errors.

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

$$Q_{W,\delta,i}(x) = \{y \in X : \text{ there exists } z \in B(W(x),\delta) \text{ such that } y = x + a_i z\}. \tag{2.2}$$

In Section 4, we prove the following result.

Theorem 2.1. *Assume that $f(x) \to \infty$ as $\|x\| \to \infty$, the sequence $\{a_i\}_{i=0}^{\infty} \subset (0,1]$ satisfies (2.1), the vector field $V \in \mathscr{A}$ is regular, assumption A(ii) is valid and that at least one of the following conditions holds: 1. $V \in \mathscr{A}_{au}$; 2. A(iii) is valid.*

Let $K, \varepsilon > 0$ be given. Then, there exists $\delta > 0$ such that, for each sequence $\{x_i\}_{i=0}^{\infty} \subset X$ which satisfies

$$\liminf_{i\to\infty} \|x_i\| < K \tag{2.3}$$

and

$$x_{i+1} \in Q_{V,\delta,i}(x_i) \tag{2.4}$$

for each $i = 0,1,\ldots$, the inequality $f(x_i) \le \inf(f) + \varepsilon$ holds for all sufficiently large natural numbers i.

This theorem is an extension of an analogous result of [21] obtained under the assumption that the sequence $\{x_i\}_{i=0}^{\infty}$ is bounded.

Section 3 contains an auxiliary result.

3. An Auxiliary Result

In the proof of Theorem 2.1, we use the following lemma, which was proved in [22].

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

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

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

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

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

4. Proof of Theorem 2.1

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

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

Let

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

Set

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

Clearly, the set E_0 is bounded and closed. Choose

$$K_1 > \max\{\sup\{\|x\| : \ x \in E_0\} + 1 + K, \ \sup\{\|V(x)\| : \ x \in E_0\} + 1\}. \tag{4.4}$$

There exists $L_0 \geq 1$ such that

$$\|f(z_1) - f(z_2)| \leq L_0 \|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(0, K + 2). \tag{4.5}$$

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

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

$$\|x\| \leq K_1, \ f(x) \geq \inf(f) + \varepsilon/4,$$

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

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

Assume that $\{x_i\}_{i=0}^{\infty} \subset X$ satisfies (2.3) and (2.4) for all integers $i \geq 0$. In view of (2.1), there exists a natural number N_1 such that

$$a_i < (4 L_0 K_1)^{-1} \bar{\alpha} \varepsilon \text{ for all integers } i \geq N_1. \tag{4.6}$$

By (2.3), there exists an integer $N_2 > N_1 + 2$ such that

$$\|x_{N_2}\| < K. \tag{4.7}$$

In view of (2.1), there exists a natural number $N_0 > N_2 + 2$ such that

$$\sum_{i=N_2+1}^{N_0 - 1} a_i > \gamma^{-1}(K_0 - \inf(f)). \tag{4.8}$$

In order to complete the proof of the theorem, it is sufficient to show that $f(x_i) \leq \inf(f) + \varepsilon$ for all integers $i \geq N_0$.

First, we show that there exists an integer $j \in [N_2, N_0]$ such that

$$f(x_j) \leq \inf(f) + \varepsilon/4.$$

Assume the contrary. It follows that

$$f(x_i) > \inf(f) + \varepsilon/4, \ i = N_2, \ldots, N_0. \tag{4.9}$$

Assume that

$$i \in \{N_2, \ldots, N_0\}, \ \|x_i\| \leq K_1. \tag{4.10}$$

From (2.2) and (2.4), we find that there exists

$$y_i \in B(V(x_i), \delta) \tag{4.11}$$

such that

$$x_{i+1} = x_i + a_i y_i. \tag{4.12}$$

It follows from property (a), (4.6) and (4.9)-(4.12) that

$$f(x_i) - f(x_{i+1}) = f(x_i) - f(x_i + a_i y_i) \geq a_i \gamma.$$

Thus, we obtain that the following property holds:

(b) If

$$i \in \{N_2, \dots, N_0\}, \ \|x_i\| \leq K_1,$$

then

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

We show that

$$\|x_i\| \leq K_1, \ i = N_2, \dots, N_0.$$

Assume the contrary. In view of (4.4) and (4.7), we find that there exists an integer

$$p \in (N_2, N_0]$$

such that

$$\|x_p\| > K_1, \tag{4.13}$$

$$\|x_i\| \leq K_1, \ i = N_2, \dots, p-1. \tag{4.14}$$

Property (b) and (4.14) imply that, for each $i \in \{N_2, \dots, p-1\}$,

$$f(x_{i+1}) \leq f(x_i) - a_i \gamma \leq f(x_i). \tag{4.15}$$

In view of (4.3), (4.7) and (4.15), we have, for all $i = N_2, \dots, p$,

$$f(x_i) \leq f(x_{N_2}) \leq K_0,$$

$$f(x_p) \leq K_0,$$

which together with (4.3) and (4.4) implies that

$$\|x_p\| \leq K_1.$$

This contradicts (4.13). The contradiction we have reached proves that

$$\|x_i\| \leq K_1, \ i = N_2, \dots, N_0. \tag{4.16}$$

Property (b) and (4.16) imply that, for all $i = N_2, \dots, N_0 - 1$,

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

Using (4.2) and (4.7) yields that

$$\begin{aligned} K_0 - \inf(f) &\geq f(x_{N_2}) - f(x_{N_0}) \\ &= \sum_{i=N_2}^{N_0-1} (f(x_i) - f(x_{i+1})) \\ &\geq \gamma \sum_{i=N_2}^{N_0-1} a_i, \end{aligned}$$

and

$$\sum_{i=N_2}^{N_0-1} a_i \leq \gamma^{-1}(K_0 - \inf(f)).$$

This contradicts (4.8). The contradiction we have reached proves that there exists

$$j \in \{N_2, \dots, N_0\} \tag{4.17}$$

such that

$$f(x_j) \leq \inf(f) + \varepsilon/4. \tag{4.18}$$

We show that, for all integers $i \geq j$,

$$f(x_i) \leq \inf(f) + \varepsilon.$$

Assume the contrary. Then there is an integer $k > j$ for which

$$f(x_k) > \inf(f) + \varepsilon. \tag{4.19}$$

We may assume without loss of generality that

$$f(x_i) \leq \inf(f) + \varepsilon \text{ for all } i = j, \dots, k-1. \tag{4.20}$$

From (4.1) and (4.20), we have

$$\|x_{k-1}\| \leq K - 2. \tag{4.21}$$

There are two cases:

$$f(x_{k-1}) > \inf(f) + \varepsilon/4; \tag{4.22}$$

$$f(x_{k-1}) \leq \inf(f) + \varepsilon/4. \tag{4.23}$$

From (2.2) and (2.4), we find that there is

$$y_{k-1} \in B(V(x_{k-1}), \delta) \tag{4.24}$$

such that

$$x_k = x_{k-1} + a_{k-1} y_{k-1}. \tag{4.25}$$

Assume that (4.22) is valid. It follows from property (a), (4.4), (4.6), (4.17), (4.20)-(4.22), (4.24) and (4.25) that

$$\inf(f) + \varepsilon \geq f(x_{k-1}) \geq f(x_k).$$

This contradicts (4.19). The contradiction we have reached proves that (4.23) is true. In view of (4.1) and (4.23), we have

$$\|x_{k-1}\| \leq K - 2. \tag{4.26}$$

By (4.2)-(4.4) and (4.24)-(4.26), we have

$$\|x_{k-1} - x_k\| = a_{k-1}\|y_{k-1}\| \leq a_{k-1} K_1. \tag{4.27}$$

In view of (4.6), (4.17), (4.26) and (4.27), we have

$$\|x_k\| \leq K - 1. \tag{4.28}$$

It follows from (4.5), (4.6), (4.17), (4.23), (4.26)-(4.28) that

$$\begin{aligned} f(x_k) \leq f(x_{k-1}) + |f(x_{k-1}) - f(x_k)| &\leq \inf(f) + \varepsilon/4 + L_0\|x_{k-1} - x_k\| \\ &\leq \inf(f) + \varepsilon/4 + L_0 a_{k-1} K_1 \\ &\leq \inf(f) + \varepsilon/2. \end{aligned}$$

This contradicts (4.19). The contradiction we have reached proves that

$$f(x_i) \leq \inf(f) + \varepsilon$$

for all integers $i \geq j$. This completes the proof of Theorem 2.1.

REFERENCES

[1] S. Reich, A. J. Zaslavski, Generic convergence of descent methods in Banach spaces, Math. Oper. Res. 25 (2000), 231-242.
[2] S. Reich, A. J. Zaslavski, The set of divergent descent methods in a Banach space is σ-porous, SIAM J. Optim. 11 (2001), 1003-1018.
[3] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation. I: Basic theory, Springer, Berlin, 2006.
[4] M. Bačák, A variational approach to stochastic minimization of convex functionals, Pure Appl. Funct. Anal. 3 (2018), 287-295.
[5] A. Cegielski, A method of projection onto an acute cone with level control in convex minimization, Math. Program. 85 (1999), 469-490.
[6] H. B. Curry, The method of steepest descent for nonlinear minimization problems, Quarterly Appl. Math. 2 (1944), 258-261.
[7] A. Gibali, K.-H. Kufer, D. Reem, P. Suss, A generalized projection-based scheme for solving convex constrained optimization problems, Comput. Optim. Appl. 70 (2018), 737-762.
[8] S.-M. Grad, 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.
[9] I. Griva, Convergence analysis of augmented Lagrangian – fast projected gradient method for convex quadratic problems, Pure Appl. Funct. Anal. 3 (2018), 417-428.
[10] Y. Hu, C. Li, X. Yang, On convergence rates of linearized proximal algorithms for convex composite optimization with applications, SIAM J. Optim. 26 (2016), 1207-1235.
[11] I. Konnov, On convergence properties of a subgradient method, Optim. Methods Softw. 18 (2003), 53-62.
[12] R. Polyak, Exterior distance function, Pure Appl. Funct. Anal. 2 (2017), 369-394.
[13] A. J. Zaslavski, Numerical Optimization with Computational Errors, Springer, 2016.
[14] A. J. Zaslavski, Convex Optimization with Computational Errors, Springer Optimization and Its Applications, Springer, 2020.
[15] C. Zalinescu, On Lagrange multipliers in convex entropy minimization, Pure Appl. Funct. Anal. 3 (2018), 393-402.
[16] J.-B. Hiriart-Urruty, C. Lemaréchal, Convex Analysis and Minimization Algorithms, Springer, Berlin, 1993.
[17] S. Reich, A. J. Zaslavski, Genericity in Nonlinear Analysis, Developments in Mathematics, 34, Springer, New York, 2014.
[18] A. J. Zaslavski, Optimization on Metric and Normed Spaces, Springer, New York, 2010.
[19] S. Reich, A. J. Zaslavski, Regular vector-fields in Banach spaces, Taiwanese J. Math. 12 (2008), 1165-1176.
[20] S. Reich, A. J. Zaslavski, Descent methods under the presence of computational errors in Banach space, Linear Nonlinear Anal. 4 (2018), 451-457.
[21] S. Reich, A. J. Zaslavski, Two iterative processes generated by regular vector fields in Banach spaces, Optimization, https://doi.org/10.1080/02331934.2020.1865348.
[22] S. Reich, A. J. Zaslavski, Descent methods with computational errors in Banach spaces, Optimization 14 (2019), 41-50.