J. Appl. Numer. Optim. 1 (2019), No. 1, pp. 53-61
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.1.2019.1.05

THREE CONVERGENCE RESULTS FOR CONTINUOUS DESCENT METHODS WITH A CONVEX OBJECTIVE FUNCTION

ALEXANDER J. ZASLAVSKI

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

Abstract. We study continuous descent methods for the minimization of a convex objective function and establish three convergence results for those methods which are generated by regular vector fields. These results improve a known convergence result in the literature.

Keywords. Convex function; Descent method; Metric space; Regular vector field.

2010 Mathematics Subject Classification. 47J35, 49M37, 90C25, 90C30.

1. Introduction

Given a Lipschitz and convex 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 [14, 15] it was introduced the class of regular vector fields and 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 our process tend to its infimum. The convergence of the values of the function f to its infimum along the trajectories of an analogous continuous dynamical system governed by such vector fields was studied in [16]. In the present paper we prove three convergence results which improve one of the main results of [16]. It should be mentioned that the case of nonconvex objective functions was studied in [1, 2].

Assume that $(X, \|\cdot\|)$ is a Banach space with norm $\|\cdot\|$, $(X^*, \|\cdot\|_*)$ is its dual space with the corresponding dual 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 [12]. It is well known that $\partial f(x)$ is a nonempty and bounded subset of $(X^*, \|\cdot\|_*)$.

E-mail address: ajzasl@technion.ac.il.

Received March 7, 2019; Accepted March 19, 2019.

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 $\|Vx\| \le K_1$ if $\|x\| \le K_0$), and for each $x \in X$ and each $l \in \partial f(x)$, $l(Vx) \le 0$. 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

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

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

$$E(N,\varepsilon) = \{(V_1,V_2) \in \mathscr{A} \times \mathscr{A} :\ \|V_1x - V_2x\| \le \varepsilon,\ x \in X,\ \|x\| \le N\},$$

where $N, \varepsilon > 0$, is a basis for the uniformity generated by the metric ρ.

The study of minimization methods for convex functions is a central topic in optimization theory; see, for example, [3, 4, 5, 6, 7, 8, 10, 11, 13, 19, 21, 22] and the references cited therein. Note, in particular, that the counterexample studied in Section 2.2 of Chapter VIII of [9] shows that, even for two-dimensional problems, the simplest choice for a descent direction, may produce sequences the functional values of which fail to converge to the infimum of the objected function f considered in [9], which attains its minimum.

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

In [15] we introduced the class of regular vector fields $V \in \mathscr{A}$ 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 the space $\mathscr{A}$. We then showed in [15] that, for any regular vector field V, the values of f tend to its infimum for the process associated with V. Note that the results of [15] are also presented in Chapter 8 of the book [18], which contains many other generic and porosity results. For more applications of the generic approach and the porosity notion in optimization theory, see also [20].

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 $\{x \in X :\ \|x\| \le r\}$.

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

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

In this connection, see also [17]. We denote by $\mathscr{F}$ the set of all regular vector fields $V \in \mathscr{A}$.

It is known (Theorem 8.2 of [18]) that in a very strong sense most of the vector fields in $\mathscr{A}$ are regular.

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

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

2. Continuous descent methods

Let $T > 0$, $x_0 \in X$ and let $u : [0,T] \to X$ be a Bochner integrable function. Set

$$x(t) = x_0 + \int_0^t u(s)ds, \; t \in [0,T].$$

Then $x : [0,T] \to X$ is differentiable and $x'(t) = u(t)$ for almost every $t \in [0,T]$. Recall that the function $f : X \to R^1$ is assumed to be convex and continuous, and therefore it is, in fact, locally Lipschitzian. It follows that its restriction to the set $\{x(t) : \; t \in [0,T]\}$ is Lipschitzian (see Section 8.11 of [18]). Hence the function

$$(f \cdot x)(t) := f(x(t)), \quad t \in [0,T],$$

is absolutely continuous. It follows that for almost every $t \in [0,T]$, both the derivatives $x'(t)$ and $(f \cdot x)'(t)$ exist:

$$x'(t) = \lim_{h \to 0} h^{-1}[x(t+h) - x(t)],$$

$$(f \cdot x)'(t) = \lim_{h \to 0} h^{-1}[f(x(t+h)) - f(x(t))].$$

We continue with the following fact.

Proposition 2.1. *(Proposition 8.13 of [18]) Assume that $t \in [0,T]$ and that both the derivatives $x'(t)$ and $(f \cdot x)'(t)$ exist. Then*

$$(f \cdot x)'(t) = \lim_{h \to 0} h^{-1}[f(x(t) + hx'(t)) - f(x(t))].$$

In the sequel we denote by $\mu(E)$ the Lebesgue measure of $E \subset R^1$.

The following theorem is one of the two main results obtained in [16] (see also Theorem 8.14 of [18]).

Theorem 2.1. *Let $V \in \mathscr{A}$ be regular and let $x : [0,\infty) \to X$ be differentiable. Suppose that*

$$x'(t) = V(x(t)) \text{ for almost every } t \in [0,\infty).$$

Assume that there exists a positive number r such that

$$\mu(\{t \in [0,T] : \; ||x(t)|| \leq r\}) \to \infty \text{ as } T \to \infty.$$

Then $\lim_{t \to \infty} f(x(t)) = \inf(f)$.

In the present paper we obtain three extensions of this result.

3. The first result

Theorem 3.1. *Let $V \in \mathscr{A}$ be regular and let $r, \varepsilon > 0$. Then there exists a neighborhood $\mathscr{U}$ of V in $\mathscr{A}$ such that for each $W \in \mathscr{U}$ and each differentiable function $x : [0,\infty) \to X$ which satisfies*

$$x'(t) = Wx(t) \text{ for almost every } t \in [0,\infty) \tag{3.1}$$

and

$$\mu(\{t \in [0,T] : \ \|x(t)\| \le r\}) \to \infty \text{ as } T \to \infty \tag{3.2}$$

the inequality

$$\lim_{t\to\infty} f(x(t)) \le \inf(f) + \varepsilon$$

holds.

Proof. Since V is a regular vector field, there exists a positive number $\delta < 1$ such that the following property holds:

(a) for each point $x \in B(0, r+1)$ satisfying

$$f(x) > \inf(f) + \varepsilon,$$

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

$$l(Vx) \le -\delta.$$

Since the function f is Lipschitz on bounded sets, there exists $L > 0$ such that

$$|f(z_1) - f(z_2)| \le L\|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(0, r+2). \tag{3.3}$$

There exists an open neighborhood $\mathscr{U}$ of V in $\mathscr{A}$ such that, for each $W \in \mathscr{U}$ and each $x \in B(0, r+2)$,

$$\|Wx - Vx\| \le 4^{-1}\delta(L+1)^{-1}. \tag{3.4}$$

Assume that

$$W \in \mathscr{U}, \tag{3.5}$$

a function $x : [0,\infty) \to X$ is differentiable and that (3.1) and (3.2) hold. Clearly, the function

$$t \to f(x(t)), \ t \in [0,\infty)$$

is decreasing. We need to show that

$$\lim_{t\to\infty} f(x(t)) \le \inf(f) + \varepsilon.$$

Assume the contrary. It follows that

$$f(x(t)) > \inf(f) + \varepsilon \text{ for all } t \in [0,\infty). \tag{3.6}$$

Let $T > 0$ and

$$\Omega_T = \{t \in [0,T] : \ \|x(t)\| \le r\}. \tag{3.7}$$

Proposition 2.1 and (3.1) imply that

$$\begin{aligned} f(x(T)) - f(x(0)) &= \int_0^T (f\cdot x)'(t)dt \\ &= \int_0^T f^0(x(t),x'(t))dt \\ &= \int_0^T f^0(x(t),Wx(t))dt \\ &\le \int_{\Omega_T} f^0(x(t),Wx(t))dt \\ &\le \int_{\Omega_T} f^0(x(t),Vx(t))dt + \int_{\Omega_T} f^0(x(t),Wx(t)-Vx(t))dt. \end{aligned} \tag{3.8}$$

From (a), (3.6) and (3.7), we obtain that, for all $t \in \Omega_T$,

$$f^0(x(t),Vx(t))dt \le -\delta.$$

This implies that

$$\int_{\Omega_T} f^0(x(t),Vx(t))dt \le -\delta\mu(\Omega_T). \tag{3.9}$$

Let $t \in \Omega_T$. In view of (3.7), one has

$$\|x(t)\| \le r. \tag{3.10}$$

It follows from (3.3) and (3.10) that

$$f^0(x(t),Wx(t)-Vx(t)) \le L\|Wx(t)-Vx(t)\|. \tag{3.11}$$

By (3.4), (3.5) and (3.10), we have

$$\|Wx(t)-Vx(t)\| \le 4^{-1}\delta(L+1)^{-1}. \tag{3.12}$$

Relations (3.11) and (3.12) imply that

$$f^0(x(t),Wx(t)-Vx(t))dt \le 4^{-1}\delta.$$

This implies that

$$\int_{\Omega_T} f^0(x(t),Wx(t)-Vx(t))dt \le 4^{-1}\delta\mu(\Omega_T).$$

This together with (3.8) and (3.9) implies that

$$\inf(f) - f(x(0)) \le f(x(T)) - f(x(0)) \le -\delta\mu(\Omega_T) + 4^{-1}\delta\mu(\Omega_T) \le -2^{-1}\mu(\Omega_T) \to -\infty$$

as $T \to \infty$. The contradiction we have reached completes the proof of Theorem 3.1. □

4. THE SECOND RESULT

Theorem 4.1. *Let $V \in \mathscr{A}$ be regular, $S > \inf(f)$ and let $r, \varepsilon > 0$. Then there exist a neighborhood $\mathscr{U}$ of V in $\mathscr{A}$ and $N_0 > 0$ such that, for each $W \in \mathscr{U}$, each $T > N_0$ and each differentiable function $x : [0,T] \to X$ which satisfies*

$$f(x(0)) \le S, \tag{4.1}$$

$$x'(t) = Wx(t) \text{ for almost every } t \in [0,T] \tag{4.2}$$

and

$$\mu(\{t \in [0,T] :\ \|x(t)\| \le r\}) \ge N_0 \tag{4.3}$$

the inequality

$$f(x(T)) \le \inf(f) + \varepsilon$$

holds.

Proof. Since V is a regular vector field, there exists a positive number $\delta < 1$ such that the following property holds:

(b) for each point $x \in B(0, r+2)$ satisfying

$$f(x) > \inf(f) + \varepsilon,$$

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

$$l(Vx) \le -\delta.$$

Since f is Lipschitz on bounded sets there exists $L > 0$ such that

$$|f(z_1) - f(z_2)| \le L\|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(0, r+2). \tag{4.4}$$

There exists an open neighborhood $\mathscr{U}$ of V in $\mathscr{A}$ such that, for each $W \in \mathscr{U}$ and each $x \in B(0, r+2)$,

$$\|Wx - Vx\| \le 4^{-1}\delta(L+1)^{-1}. \tag{4.5}$$

Choose a number $N_0 > 0$ such that

$$N_0 > 1 + 2(S - \inf(f))\delta^{-1}. \tag{4.6}$$

Assume that

$$W \in \mathscr{U}, \tag{4.7}$$

$T > N_0$, $x : [0, T] \to X$ is differentiable function and that (4.1)-(4.3) hold.

We show that

$$f(x(T)) \le \inf(f) + \varepsilon.$$

Assume the contrary. Then

$$f(x(t)) > \inf(f) + \varepsilon \text{ for all } t \in [0, T]. \tag{4.8}$$

Set

$$\Omega = \{t \in [0, T] : \ \|x(t)\| \le r\}. \tag{4.9}$$

Proposition 2.1, (4.2) and (4.9) imply that

$$\begin{aligned}
f(x(T)) - f(x(0)) &= \int_0^T (f \cdot x)'(t)dt \\
&= \int_0^T f^0(x(t), x'(t))dt \\
&= \int_0^T f^0(x(t), Wx(t))dt \\
&\le \int_\Omega f^0(x(t), Wx(t))dt \\
&\le \int_\Omega f^0(x(t), Vx(t))dt + \int_\Omega f^0(x(t), Wx(t) - Vx(t))dt.
\end{aligned} \tag{4.10}$$

Using (b), (4.8) and (4.9), we obtain that, for all $t \in \Omega$,

$$f^0(x(t), Vx(t))dt \le -\delta.$$

This implies that

$$\int_{\Omega} f^0(x(t), Vx(t))dt \le -\delta\mu(\Omega). \tag{4.11}$$

Let $t \in \Omega$. In view of (4.9), one has

$$\|x(t)\| \le r. \tag{4.12}$$

It follows from (4.4) and (4.12) that

$$f^0(x(t), Wx(t) - Vx(t)) \le L\|Wx(t) - Vx(t)\|. \tag{4.13}$$

By (4.5), (4.7) and (4.12), one has

$$\|Wx(t) - Vx(t)\| \le 4^{-1}\delta(L+1)^{-1}.$$

This together with (4.13) implies that

$$f^0(x(t), Wx(t) - Vx(t))dt \le 4^{-1}\delta.$$

Hence,

$$\int_{\Omega} f^0(x(t), Wx(t) - Vx(t))dt \le 4^{-1}\delta\mu(\Omega).$$

This together with (4.1), (4.3), (4.10) and (4.11) implies that

$$\inf(f) - S \le f(x(T)) - f(x(0)) \le -\delta\mu(\Omega) + 4^{-1}\delta\mu(\Omega) \le -2^{-1}\delta\mu(\Omega) \le -2^{-1}N_0$$

and

$$N_0 \le 2\delta^{-1}(S - \inf(f)).$$

This contradicts (4.6). The contradiction we have reached completes the proof of Theorem 4.1. □

5. THE THIRD RESULT

Theorem 5.1. *Let $r > 0$, $V \in \mathscr{A}$ be regular and let $\gamma : [0, \infty) \to (0, \infty)$ satisfy*

$$\lim_{t\to\infty} \gamma(t) = 0. \tag{5.1}$$

Assume that a differentiable function $x : [0, \infty) \to X$ satisfies for almost every $t \ge 0$,

$$f^0(x(t), x'(t)) \le 0, \tag{5.2}$$

$$\|x'(t) - Vx(t))\| \le \gamma(t) \tag{5.3}$$

and that

$$\mu(\{t \in [0, T] : \|x(t)\| \le r\}) \to \infty \text{ as } T \to \infty. \tag{5.4}$$

Then

$$\lim_{t\to\infty} f(x(t)) = \inf(f).$$

Proof. Proposition 2.1 implies that, for almost every $t \in [0, T]$,

$$(f \circ x)'(t) = f^0(x(t), x'(t)) \le 0. \tag{5.5}$$

Thus the function $f \circ x$ is decreasing.

Let $\varepsilon > 0$. It is sufficient to show that there exists $t > 0$ such that $f(x(t)) \le \inf(f) + \varepsilon$. Since V is regular there exists $\delta \in (0, \varepsilon)$ such that the following property holds:

(c) for each $x \in B(0, r)$ satisfying $f(x) > \inf(f) + \varepsilon$ and each $l \in \partial f(x)$,

$$l(Vx) \le -\delta.$$

Since the function f is Lipschitz on bounded sets, there exists $L > 0$ such that

$$|f(z_1) - f(z_2)| \leq L\|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(0, r+2). \tag{5.6}$$

In view of (5.1), there exists $T_0 > 0$ such that

$$L\gamma(t) \leq \delta/2 \text{ for all } t \in [T_0, \infty). \tag{5.7}$$

Choose a number $N_0 > 1$ such that

$$f(x(0)) - 2^{-1}\delta N_0 < \inf(f). \tag{5.8}$$

By (5.4), there exists $T > T_0$ such that

$$\mu(\{t \in [T_0, T] : \|x(t)\| \leq r\}) > N_0. \tag{5.9}$$

We show that there exists $t \in [T_0, T]$ such that

$$f(x(t)) \leq \inf(f) + \varepsilon.$$

Assume the contrary. Then

$$f(x(t)) > \inf(f) + \varepsilon \text{ for all } t \in [T_0, T]. \tag{5.10}$$

Set

$$\Omega = \{t \in [T_0, T] : \|x(t)\| \leq r\}. \tag{5.11}$$

(c), (5.10) and (5.11) imply that, for all $t \in \Omega$,

$$f^0(x(t), Vx(t)) \leq -\delta. \tag{5.12}$$

Using Proposition 2.1, (5.5) and (5.11), we obtain that

$$\begin{aligned} f(x(T)) &= f(x(T_0)) + f(x(T)) - f(x(T_0)) \\ &\leq f(x(0)) + \int_{T_0}^{T} (f \cdot x)'(t)dt \\ &\leq f(x(0)) + \int_{\Omega} (f \cdot x)'(t)dt \\ &= f(x(0)) + \int_{\Omega} f^0(x(t), x'(t))dt \\ &\leq f(x(0)) + \int_{\Omega} f^0(x(t), Vx(t))dt + \int_{\Omega} f^0(x(t), x'(t) - Vx(t))dt. \end{aligned} \tag{5.13}$$

By (5.3), (5.6) and (5.11), for almost every $t \in \Omega$, we have

$$f^0(x(t), x'(t) - Vx(t)) \leq L\|x'(t) - Vx(t)\| \leq L\gamma(t). \tag{5.14}$$

It follows from (5.7)-(5.9) and (5.12)-(5.14) that

$$\begin{aligned} f(x(T)) &\leq f(x(0)) - \delta\mu(\Omega) + L\mu(\Omega)\sup\{\gamma(t) : t \in [T_0, T]\} \\ &\leq f(x(0)) - 2^{-1}\delta\mu(\Omega) \leq f(x(0)) - 2^{-1}\delta N_0 < \inf(f). \end{aligned}$$

The contradiction we have reached proves that

$$f(x(t)) \leq \inf(f) + \varepsilon$$

and complete the proof of Theorem 5.1. □

REFERENCES

[1] S. Aizicovici, S. Reich, A.J. Zaslavski, Convergence theorems for continuous descent methods, J. Evol. Equ. 4 (2004), 139-156.

[2] S. Aizicovici, S. Reich, A. J. Zaslavski, Convergence results for a class of abstract continuous descent methods, Electron. J. Differential Equations 2004 (2004), Article ID 45.

[3] M. Bačák, A variational approach to stochastic minimization of convex functionals, Pure Appl. Funct. Anal. 3 (2018), 287–295.

[4] A. Cegielski, A method of projection onto an acute cone with level control in convex minimization, Math. Program. 85 (1999), 469–490.

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

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

[7] 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.

[8] I. Griva, Convergence analysis of augmented Lagrangian – fast projected gradient method for convex quadratic problems, Pure Appl. Funct. Anal. 3 (2018), 417-428.

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

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

[13] R. Polyak, Exterior distance function, Pure Appl. Funct. Anal. 2 (2017), 369-394.

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

[15] 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.

[16] S. Reich, A. J. Zaslavski, Two convergence results for continuous descent methods, Electron. J. Differential Equations 2003 (2003), 1-11.

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

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

[19] Y. Shehu, A convergence analysis result for constrained convex minimization problem, Optimization 64 (2015), 2587-2597.

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

[21] A.J. Zaslavski, Numerical optimization with computational errors, Springer, 2016.

[22] C. Zalinescu, On Lagrange multipliers in convex entropy minimization, Pure Appl. Funct. Anal. 3 (2018), 393-402.