Journal of Convex Analysis
Volume 27 (2020), No. 2, 557–564

Asymptotic Behavior of Continuous Descent Methods with a Convex Objective Function

Simeon Reich
Dept. of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel
sreich@technion.ac.il

Alexander J. Zaslavski
Dept. of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel
ajzasl@technion.ac.il

Dedicated to Alexander Ioffe on the occasion of his 80th birthday.

Received: February 10, 2019
Accepted: July 31, 2019

Given a Lipschitz and convex objective function on a Banach space, we revisit the class of regular vector fields introduced in our previous work on descent methods. We consider continuous descent methods for minimizing our objective function and establish two convergence results for those methods which are generated by regular vector fields. These results improve upon a convergence result which we obtained in our previous work.

Keywords: Banach space, complete metric space, convex function, descent method, regular vector field.

2010 Mathematics Subject Classification: 37L99, 47J35, 49M37, 90C25, 90C30.

1. Introduction

Given a Lipschitz and convex objective function on a Banach space, we consider a certain complete metric space of vector fields. These vector fields are self-mappings of the Banach space and we are concerned 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 values of the objective function at the points generated by our process tend to its infimum. In [15] we studied 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. In the present paper we prove two results which improve one of the main results of [15].

Assume that $(X, \|\cdot\|)$ is a Banach space with *norm* $\|\cdot\|$ and let $(X^*, \|\cdot\|_*)$ be its dual space with the corresponding *dual norm* $\|\cdot\|_*$. Assume further that

ISSN 0944-6532 / $ 2.50

$f\colon X \to R^1$ 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 [11]. 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 $V\colon X \to X$ 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

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

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 closed subsets of the metric space $(\mathcal{A},\rho)$. We assign to all these spaces the same metric ρ.

The study of minimization methods for convex functions is a central topic in convex analysis and optimization theory. See, for example, [1, 3, 4, 5, 6, 7, 9, 10, 12, 18, 19, 21] 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 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.

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 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 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 (A1) and (A2) 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 [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. Note that the results of [14] 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 convex analysis and optimization theory, see also [2] and [20].

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

(A1) There exists a norm bounded set $X_0 \subset X$ such that

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

(A2) 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 (A1) is closed and convex.

It is clear that assumption (A1) 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 [16]. We denote by $\mathcal{F}$ the set of all regular vector fields $V \in \mathcal{A}$. It is known (Theorem 8.2 of [17]) that in a very strong sense most of the vector fields in $\mathcal{A}$ 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 $u \colon [0, T] \to X$ be a Bochner integrable function. Set

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

Then $x \colon [0, T] \to X$ is differentiable and $x'(t) = u(t)$ for almost every $t \in [0, T]$. Recall that the function $f \colon X \to R^1$ is assumed to be convex and continuous, and therefore it is, in fact, locally Lipschitz. It follows that its restriction to the set $\{x(t) : \ t \in [0,T]\}$ is Lipschitz (see Section 8.11 of [17]). Hence the function $(f \cdot x)(t) := f(x(t))$, $t \in [0,T]$, is absolutely continuous. It follows that for almost every $t \in [0,T]$, both the derivatives $x'(t)$ and $(f \cdot x)'(t)$ exist:

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

We continue with the following fact.

Proposition 2.1. (Proposition 8.13 of [17]) *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 [15] (see also Theorem 8.14 of [17]).

Theorem 2.2. *Let $V \in \mathcal{A}$ be regular, let $x\colon [0,\infty) \to X$ be differentiable and suppose that $x'(t) = V(x(t))$ for almost every $t \in [0,\infty)$. Assume that there exists a positive number r such that $\mu(\{t \in [0,T] : ||x(t)|| \le r\}) \to \infty$ as $T \to \infty$. Then $\lim_{t\to\infty} f(x(t)) = \inf(f)$.*

In the present paper we obtain two new versions of this result (see Theorems 3.1 and 4.1 below). They are established under more practical conditions.

3. The first result

Theorem 3.1. *Let $V \in \mathcal{A}$ be regular, $x\colon [0,\infty) \to X$ be differentiable, and assume that $x'(t) = V(x(t))$ for almost every $t \in [0,\infty)$ and $\liminf_{t\to\infty} \|x(t)\| < \infty$. Then $\lim_{t\to\infty} f(x(t)) = \inf(f)$.*

Proof. Choose
$$r > \liminf_{t\to\infty} \|x(t)\|. \tag{3.1}$$

There exists $K > 0$ such that
$$\|Vx\| \le K \text{ for all } x \in B(0,2r). \tag{3.2}$$

In view of (3.1), there exists a strictly increasing sequence $\{t_k\}_{k=1}^{\infty} \subset (0,\infty)$ such that
$$t_{k+1} \ge t_k + 8 \text{ for all integers } k \ge 1 \tag{3.3}$$
and
$$\|x(t_k)\| \le r \text{ for all integers } k \ge 1. \tag{3.4}$$

Let $k \ge 1$ be an integer. If for all $t \ge t_k$, we have $\|x(t)\| \le 2r$, then our result is true in view of Theorem 2.2. Therefore we may assume without any loss of generality that there exists a number
$$t > t_k \ \text{ such that } \ \|x(t)\| > 2r. \tag{3.5}$$

Set
$$E := \{\tau > t_k : \ \|x(t)\| \le 2r \text{ for all } t \in [t_k, \tau]\}. \tag{3.6}$$

By (3.4), $E \neq \emptyset$. It follows from (3.5) that E is bounded. Now put

$$\tau := \sup E. \tag{3.7}$$

It is clear that
$$\tau \in E \tag{3.8}$$
and
$$\|x(t)\| \leq 2r, \ t \in [t_k, \tau]. \tag{3.9}$$

Since the function x is continuous, it follows from (3.7)–(3.9) that

$$\|x(\tau)\| = 2r. \tag{3.10}$$

It is clear that
$$x(\tau) - x(t_k) = \int_{t_k}^{\tau} V(x(t))dt. \tag{3.11}$$

By (3.2) and (3.9),
$$\|V(x(t))\| \leq K \text{ for all } t \in [t_k, \tau]. \tag{3.12}$$

It follows from (3.4) and (3.10)–(3.12) that

$$\begin{aligned} r &\leq \|x(\tau)\| - \|x(t_k)\| \leq \|x(\tau) - x(t_k)\| \\ &= \|\int_{t_k}^{\tau} V(x(t))dt\| \leq \int_{t_k}^{\tau} \|V(x(t))\|dt \leq K(\tau - t_k) \end{aligned}$$

and $\tau - t_k \geq rK^{-1}$. Since k is any natural number, we conclude that for all

$$t \in \cup_{k=1}^{\infty}[t_k, t_k + \min\{1, rK^{-1}\}],$$

we have $\|x(t)\| \leq 2r$. This implies that

$$\mu(\{t \in [0, T] : \ \|x(t)\| \leq 2r\}) \to \infty \ \text{ as } \ T \to \infty,$$

and Theorem 2.2 implies that $\lim_{t\to\infty} f(x(t)) = \inf(f)$, completing the proof. □

4. The second result

Theorem 4.1. *Let $V \in \mathcal{A}$ be regular and $S, r, \epsilon > 0$. Then there exists an integer $k_0 \geq 2$ such that for each number $T \geq k_0$ and for each differentiable function $x \colon [0, T] \to X$ such that*

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

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

and there exists a finite sequence $\{t_i\}_{i=1}^{k_0} \subset [0, T]$ for which

$$\|x(t_i)\| \leq r, \ i = 1, \ldots, k_0, \ t_{i+1} - t_1 \geq 1, \ i = 1, \ldots, k_0 - 1, \tag{4.3}$$

the inequality
$$f(x(T)) \leq \inf(f) + \epsilon \tag{4.4}$$
holds.

Proof. There exists $L > r + 1$ such that

$$\|Vx\| \le L \text{ for all } x \in B(0, 2r). \tag{4.5}$$

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

(a) for each point $x \in X$ satisfying $\|x\| \le 2r$ and $f(x) \ge \inf(f) + \epsilon$, and for each bounded linear functional $l \in \partial f(x)$, we have $l(Vx) \le -\delta$.

Choose a constant $k_0 > (\delta r)^{-1}(S - \inf(f))L + 2$.

Let $T \ge k_0$, let a differentiable function $x\colon [0, T] \to X$ satisfy (4.1) and (4.2), and let a finite sequence $\{t_i\}_{i=1}^{k_0} \subset [0, T]$ satisfy (4.3). We claim that (4.4) holds.

Suppose to the contrary that this is not true. Then

$$f(x(t)) > \inf(f) + \epsilon, \quad t \in [0, T]. \tag{4.6}$$

Proposition 2.1, (4.2) and (4.3) imply that

$$\begin{aligned} f(x(T)) - f(x(0)) &= \int_0^T (f \cdot x)'(t)dt = \int_0^T f^0(x(t), x'(t))dt \\ &= \int_0^T f^0(x(t), Vx(t))dt \le \sum_{i=1}^{k_0-1} \int_{t_i}^{t_{i+1}} f^0(x(t), Vx(t))dt. \end{aligned} \tag{4.7}$$

Let $i \in [1, k_0)$ be an integer. In view of (4.3), $\|x(t_i)\| \le r$. (4.8)

There are two cases: $\|x(t)\| \le 2r$ for all $t \in [t_i, t_{i+1}]$ (4.9)

and $\max\{\|x(t)\| :\ t \in [t_i, t_{i+1}]\} > 2r$. (4.10)

Assume that (4.9) holds. By (4.6), (4.9) and property (a),

$$f^0(x(t), Vx(t)) \le -\delta,\ t \in [t_i, t_i + 1]. \tag{4.11}$$

In view of (4.11),

$$\int_{t_i}^{t_{i+1}} f^0(x(t), Vx(t))dt \le \int_{t_i}^{t_i+1} f^0(x(t), Vx(t))dt \le -\delta. \tag{4.12}$$

Assume that (4.10) holds. Set

$$E := \{\tau \in [t_i, t_{i+1}] :\ \|x(t)\| \le 2r \text{ for all } t \in [t_i, \tau]\}. \tag{4.13}$$

By (4.8), $E \ne \emptyset$. Set $\tau := \sup E$. (4.14)

Since the function x is continuous, we conclude that $\tau \in E$. (4.15)

By (4.5) and (4.13)–(4.15), $\|Vx(t)\| \le L,\ t \in [t_i, \tau]$. (4.16)

It follows from the continuity of x and (4.10) that $\tau < t_{i+1}$ and that

$$\|x(\tau)\| = 2r. \tag{4.17}$$

The relations (4.2), (4.8), (4.16) and (4.17) imply that

$$r \leq \|x(\tau)\| - \|x(t_i)\| \leq \|x(\tau) - x(t_i)\|$$

$$= \| \int_{t_i}^{\tau} x'(t)dt\| \leq \int_{t_i}^{\tau} \|Vx(t)\|dt \leq L(\tau - t_i)$$

and

$$\tau - t_i \geq L^{-1}r. \tag{4.18}$$

Property (a), (4.6) and (4.13)–(4.15) imply that for all $t \in [t_i, \tau]$,

$$f^0(x(t), Vx(t))dt \leq -\delta. \tag{4.19}$$

By (4.13)–(4.15), (4.18) and (4.19),

$$\int_{t_i}^{t_{i+1}} f^0(x(t), Vx(t))dt \leq \int_{t_i}^{\tau} f^0(x(t), Vx(t))dt \leq -\delta(\tau - t_i) \leq -\delta L^{-1}r.$$

Thus in view of (4.11) and the above relation, in both cases we have

$$\int_{t_i}^{t_{i+1}} f^0(x(t), Vx(t))dt \leq \max\{-\delta L^{-1}r, -\delta\} = -\delta L^{-1}r.$$

When combined with (4.1), this implies that

$$\inf(f) - S \leq f(x(T)) - f(x(0)) \leq (k_0 - 1)(-\delta r L^{-1})$$

and

$$k_0 \leq 1 + (S - \inf(f))(\delta r)^{-1}L.$$

This, however, contradicts the choice of k_0. The contradiction we have reached completes the proof of Theorem 4.1. □

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 an anonymous referee for several helpful comments and suggestions.

References

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

[2] D. Butnariu, S. Reich, A. J. Zaslavski: *There are many totally convex functions*, J. Convex Analysis 13 (2006) 323–332.

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

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

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

[6] 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.

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

[8] J.-B. Hiriart-Urruty, C. Lemaréchal: *Convex Analysis and Minimization Algorithms*, Springer, Berlin (1993).

[9] 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.

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

[11] B. S. Mordukhovich: *Variational Analysis and Generalized Differentiation. I: Basic Theory*, Springer, Berlin (2006).

[12] R. Polyak: *Exterior distance function*, Pure Appl. Functional Analysis 2 (2017) 369–394.

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

[14] 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.

[15] S. Reich, A. J. Zaslavski: *Two convergence results for continuous descent methods*, Electron. J. Diff. Eqns. 2003 (2003) 1–11.

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

[17] S. Reich, A. J. Zaslavski: *Genericity in Nonlinear Analysis*, Developments in Mathematics 34, Springer, New York (2014).

[18] Y. Shehu: *A convergence analysis result for constrained convex minimization problem*, Optimization 64 (2015) 2587–2597.

[19] C. Zalinescu: *On Lagrange multipliers in convex entropy minimization*, Pure Appl. Functional Analysis 3 (2018) 393–402.

[20] A. J. Zaslavski: *Optimization on Metric and Normed Spaces*, Springer, New York (2010).

[21] A. J. Zaslavski: *Numerical Optimization with Computational Errors*, Springer, Cham (2016).