STABILITY OF CONVERGENT CONTINUOUS DESCENT METHODS

SERGIU AIZICOVICI, SIMEON REICH, ALEXANDER J. ZASLAVSKI

Abstract. We consider continuous descent methods for the minimization of convex functions defined on a general Banach space. In our previous work we showed that most of them (in the sense of Baire category) converged. In the present paper we show that convergent continuous descent methods are stable under small perturbations.

1. Introduction

The study of discrete and continuous descent methods is an important topic in optimization theory and in dynamical systems. See, for example, [1, 2, 4, 6, 7, 8, 9, 10, 11]. Given a continuous convex function f on a Banach space X, we associate with f a complete metric space of vector fields $V : X \to X$ such that $f^0(x, Vx) \leq 0$ for all $x \in X$. Here $f^0(x, u)$ is the right-hand derivative of f at x in the direction of $u \in X$. To each such vector field there correspond two gradient-like iterative processes. In the papers [8, 9], it is shown that for most of these vector fields, both iterative processes generate sequences $\{x_n\}_{n=1}^\infty$ such that the sequences $\{f(x_n)\}_{n=1}^\infty$ tend to $\inf(f)$ as $n \to \infty$. Here by "most" we mean an everywhere dense G_δ subset of the space of vector fields (cf., for example, [5, 8, 12]). In the recent paper [2], we studied the convergence of the trajectories of an analogous continuous dynamical system governed by such vector fields to the point where the function f attains its infimum. The first attempt to examine such continuous descent methods was made in [11]. However, it is assumed there that the convex function f is Lipschitz on all bounded subsets of X. No such assumption was made in [2]. We remark in passing that continuous descent methods for the minimization of Lipschitz (not necessarily convex) functions are studied in [1].

In the present paper, we show that convergent continuous descent methods are stable under small perturbations (see Theorem 1.3 below).

More precisely, let $(X, \|\cdot\|)$ be a Banach space and let $f : X \to R^1$ be a convex continuous function which satisfies the following conditions:

C(i) $\lim_{\|x\| \to \infty} f(x) = \infty$;
C(ii) there is a point $\bar{x} \in X$ such that $f(\bar{x}) \leq f(x)$ for all $x \in X$;

C(iii) if $\{x_n\}_{n=1}^{\infty} \subset X$ and $\lim_{n\to\infty} f(x_n) = f(\bar{x})$, then $\lim_{n\to\infty} x_n = \bar{x}$.

By C(iii), the point $\bar{x}$, where the minimum of f is attained, is unique.

For each $x \in X$, let

$$f^0(x, u) = \lim_{t \to 0^+} [f(x + tu) - f(x)]/t, \quad u \in X. \tag{1.1}$$

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

$$B(x, r) = \{z \in X : \|z - x\| \le r\}, \quad B(r) = B(0, r). \tag{1.2}$$

For each mapping $A : X \to X$ and each $r > 0$, put

$$\mathrm{Lip}(A, r) = \sup\{\|Ax - Ay\|/\|x - y\| : \ x, y \in B(r), \text{ and } x \ne y\}. \tag{1.3}$$

Denote by $\mathcal{A}_l$ the set of all mappings $V : X \to X$ such that $\mathrm{Lip}(V, r) < \infty$ for each positive r (this means that the restriction of V to any bounded subset of X is Lipschitz) and $f^0(x, Vx) \le 0$ for all $x \in X$.

For the set $\mathcal{A}_l$ we consider the uniformity determined by the base

$$\begin{aligned} E_s(n, \epsilon) = \big\{&(V_1, V_2) \in \mathcal{A}_l \times \mathcal{A}_l : \ \mathrm{Lip}(V_1 - V_2, n) \le \epsilon \\ &\text{and } \|V_1 x - V_2 x\| \le \epsilon \text{ for all } x \in B(n)\big\}. \end{aligned} \tag{1.4}$$

Clearly, this uniform space $\mathcal{A}_l$ is metrizable and complete. The topology induced by this uniformity in $\mathcal{A}_l$ will be called the strong topology.

We will also equip the space $\mathcal{A}_l$ with the uniformity determined by the base

$$E_w(n, \epsilon) = \{(V_1, V_2) \in \mathcal{A}_l \times \mathcal{A}_l : \ \|V_1 x - V_2 x\| \le \epsilon \text{ for all } x \in B(n)\} \tag{1.5}$$

where $n, \epsilon > 0$. The topology induced by this uniformity will be called the weak topology.

The following existence result was proved in [2, Section 3].

Proposition 1.1. *Let $x_0 \in X$ and $V \in \mathcal{A}_l$. Then there exists a unique continuously differentiable mapping $x : [0, \infty) \to X$ such that*

$$x'(t) = Vx(t), \quad t \in [0, \infty),$$
$$x(0) = x_0.$$

We now recall the main result of [2].

Theorem 1.2. *There exists a set $\mathcal{F} \subset \mathcal{A}_l$ which is a countable intersection of open (in the weak topology) everywhere dense (in the strong topology) subsets of $\mathcal{A}_l$ such that for each $V \in \mathcal{F}$, the following property holds:*

For each $\epsilon > 0$ and each $n > 0$, there exist $T_{\epsilon,n} > 0$ and a neighborhood $\mathcal{U}$ of V in $\mathcal{A}_l$ with the weak topology such that for each $W \in \mathcal{U}$ and each differentiable mapping $y : [0, \infty) \to X$ satisfying

$$|f(y(0))| \le n \quad and \quad y'(t) = Wy(t) \text{ for all } t \ge 0,$$

the inequality $\|y(t) - \bar{x}\| \le \epsilon$ holds for all $t \ge T_{\epsilon,n}$.

Denote by $\mathcal{F}_*$ the set of all $V \in \mathcal{A}_l$ which have the following property:

(P1) For each $\epsilon > 0$ and each $n > 0$, there exists $T_{\epsilon,n} > 0$ such that for each differentiable mapping $y : [0, \infty) \to X$ satisfying

$$|f(y(0))| \le n \quad \text{and} \quad y'(t) = Vy(t) \quad \text{for all } t \ge 0,$$

the inequality $\|y(t) - \bar{x}\| \le \epsilon$ holds for some $t \in [0, T_{\epsilon,n}]$.

By Theorem 1.2, $\mathcal{F}_*$ contains a subset which is a countable intersection of open (in the weak topology) everywhere dense (in the strong topology) subsets of $\mathcal{A}_l$.

We denote by $\mathcal{A}$ the set of all mappings $V : X \to X$ which are bounded on bounded subsets of X and satisfy $f^0(x, Vx) \leq 0$ for all $x \in X$. Clearly, $\mathcal{A}_l \subset \mathcal{A}$.

For the set $\mathcal{A}$ we consider the uniformity determined by the base

$$G_w(n, \epsilon) = \{(V_1, V_2) \in \mathcal{A} \times \mathcal{A} : \|V_1 x - V_2 x\| \leq \epsilon \; \forall x \in B(n)\}, \qquad (1.6)$$

where $n, \epsilon > 0$. Clearly, the space $\mathcal{A}$ with this uniformity is metrizable. We are now ready to formulate the main result of the present paper.

Theorem 1.3. *Let $V \in \mathcal{F}_*$ and $n, \epsilon > 0$. Then there exist $T_{\epsilon,n} > 0$ and a neighborhood $\mathcal{U}$ of V in $\mathcal{A}$ such that for each $W \in \mathcal{U}$, each $T \geq T_{\epsilon,n}$ and each $x \in W^{1,1}(0, T; X)$ satisfying*

$$x'(t) = Wx(t), \quad t \in [0, T] \text{ a.e.}, \qquad (1.7)$$

and

$$|f(x(0))| \leq n, \qquad (1.8)$$

the inequality $\|x(t) - \bar{x}\| \leq \epsilon$ holds for all $t \in [T_{\epsilon,n}, T]$.

Our paper is organized as follows. An auxiliary result, Proposition 2.1, is presented in Section 2. Our main result, Theorem 1.3, is proved in Section 3.

2. An auxiliary result

Let the mapping x belong to the Sobolev space $W^{1,1}(0, T; X)$, i.e. (see, e.g., [3]),

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

where $T > 0$, $x_0 \in X$ and $u \in L^1(0, T; X)$. Then $x : [0, T] \to X$ is absolutely continuous and $x'(t) = u(t)$ for a.e. $t \in [0, T]$. Recall that the function $f : X \to R^1$ is assumed to be convex and continuous and therefore it is, in fact, locally Lipschitz. It follows that its restriction to the set $\{x(t) : t \in [0, T]\}$ is Lipschitz. Indeed, since this set is compact, the restriction of f to it is Lipschitz.

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

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

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

We now quote [11, Proposition 3.1].

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

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

3. Proof of Theorem 1.3

By C(i), there is $n_1 > n$ such that

$$\text{if } z \in X \text{ and } f(z) \leq n, \text{ then } \|z\| \leq n_1. \tag{3.1}$$

By C(iii), there is $\delta_1 > 0$ such that

$$\text{if } z \in X \text{ and } f(z) \leq f(\bar{x}) + \delta_1, \text{ then } \|z - \bar{x}\| \leq \epsilon. \tag{3.2}$$

Since f is continuous, there is $\epsilon_1 > 0$ such that

$$|f(\bar{x}) - f(z)| \leq \delta_1 \text{ for each } z \in X \text{ satisfying } \|z - \bar{x}\| \leq \epsilon_1. \tag{3.3}$$

Since $V \in \mathcal{A}_l$, there is $L > 1$ such that

$$\|Vz_1 - Vz_2\| \leq L\|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(n_1). \tag{3.4}$$

By (P1), there is $T_0 > 0$ such that the following property holds:

(P2) For each differentiable mapping $y : [0, \infty) \to X$ satisfying

$$|f(y(0))| \leq n \text{ and } y'(t) = Vy(t) \text{ for all } t \geq 0, \tag{3.5}$$

the inequality $\|y(t) - \bar{x}\| \leq \epsilon_1/4$ holds for some $t \in [0, T_0]$.

Choose a positive number Δ such that

$$\Delta(T_0 + 1)e^{LT_0} \leq \epsilon_1/4 \tag{3.6}$$

and set

$$\mathcal{U} = \{W \in \mathcal{A} : \|Wz - Vz\| \leq \Delta \text{ for all } z \in B(n_1)\}. \tag{3.7}$$

Assume that

$$W \in \mathcal{U}, \tag{3.8}$$

$T \geq T_0$ and that $x \in W^{1,1}(0, T; X)$ satisfies both (1.7) and (1.8).

We show that

$$\|x(t) - \bar{x}\| \leq \epsilon \text{ for all } t \in [T_0, T]. \tag{3.9}$$

By Proposition 1.1, there exists a differentiable mapping $y : [0, \infty) \to X$ which satisfies

$$y'(t) = Vy(t), \ t \in [0, \infty), \tag{3.10}$$

and

$$y(0) = x(0). \tag{3.11}$$

In view of Proposition 2.1 and the inclusions $V, W \in \mathcal{A}$, the function $f \circ x$ is decreasing on $[0, T]$ and the function $f \circ y$ is decreasing on $[0, \infty)$. Together with (1.8) and (3.11), this implies that

$$f(x(t)) \leq n, \ t \in [0, T], \text{ and } f(y(t)) \leq n, \ t \in [0, \infty). \tag{3.12}$$

When combined with (3.1), this inequality implies

$$\|x(t)\|, \|y(t)\| \leq n_1 \text{ for all } t \in [0, T]. \tag{3.13}$$

By property (P2), (1.8), (3.10) and (3.11), there is a number τ such that

$$\tau \in [0, T_0] \text{ and } \|y(\tau) - \bar{x}\| \leq \epsilon_1/4. \tag{3.14}$$

Now we will estimate $\|y(\tau) - x(\tau)\|$. It follows from (1.7), (3.10) and (3.11) that for each $s \in [0, \tau]$,

$$
\begin{aligned}
\|y(s) - x(s)\| &= \|y(0) + \int_0^s Vy(t)dt - (x(0) + \int_0^s Wx(t)dt)\| \\
&= \| \int_0^s (Vy(t) - Wx(t))dt\| \\
&\leq \int_0^s \|Vy(t) - Wx(t)\|dt \\
&\leq \int_0^s \|Vy(t) - Vx(t)\|dt + \int_0^s \|Vx(t) - Wx(t)\|dt.
\end{aligned}
\tag{3.15}
$$

By (3.13), (3.7) and (3.8), for each $s \in (0, \tau]$, we have

$$
\int_0^s \|Vx(t) - Wx(t)\|dt \leq \int_0^s \Delta dt \leq \Delta s \leq \Delta \tau.
\tag{3.16}
$$

By (3.15), (3.16), (3.13) and (3.4), for each $s \in [0, \tau]$,

$$
\|y(s) - x(s)\| \leq \Delta \tau + \int_0^s \|Vy(t) - Vx(t)\|dt \leq \Delta \tau + L \int_0^s \|y(t) - x(t)\|dt.
$$

Applying Gronwall's inequality, we obtain that

$$
\|y(\tau) - x(\tau)\| \leq \Delta \tau e^{\int_0^\tau L dt} = \Delta \tau e^{L\tau}.
$$

Since $\tau \leq T_0$, we conclude, by using (3.6), that

$$
\|y(\tau) - x(\tau)\| \leq \Delta T_0 e^{LT_0} < \epsilon_1/4.
$$

Together with (3.14), this inequality implies that $\|x(\tau) - \bar{x}\| \leq \epsilon_1/2$. It follows from this last inequality and (3.3) that $f(x(\tau)) \leq f(\bar{x}) + \delta_1$. Since the function $f \circ x$ is decreasing, we infer that

$$
f(x(t)) \leq f(\bar{x}) + \delta_1 \quad \text{for all } t \in [\tau, T].
$$

When combined with (3.2), this inequality implies that

$$
\|x(t) - \bar{x}\| \leq \epsilon \text{ for all } t \in [\tau, T].
$$

The proof of Theorem 1.3 is complete.

References

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

[2] S. Aizicovici, S. Reich and A. J. Zaslavski; *Most continuous descent methods converge*, Arch. Math. (Basel), Vol. 85 (2005), 268-277.

[3] H. Brezis; *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North Holland, Amsterdam, 1973.

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

[5] F. S. De Blasi and J. Myjak; *Sur la convergence des approximations successives pour les contractions non linéaires dans un espace de Banach*, C. R. Acad. Sci. Paris, Vol. 283 (1976), 185-187.

[6] J.-B. Hiriart-Urruty and C. Lemaréchal; *Convex Analysis and Minimization Algorithms*, Springer, Berlin, 1993.

[7] J. W. Neuberger; *Sobolev Gradients and Differential Equations.* Lecture Notes in Math. No. 1670, Springer, Berlin, 1997.

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

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

[10] S. Reich and A. J. Zaslavski; *Porosity of the set of divergent descent methods*, Nonlinear Anal., Vol. 47 (2001), 3247-3258.

[11] S. Reich and A. J. Zaslavski; *Two convergence results for continuous descent methods*, Electron. J. Differential Equations, Vol. 2003 (2003), No. 24, 1-11.

[12] G. Vidossich; *Most of the successive approximations do converge*, J. Math. Anal. Appl., Vol. 45 (1974), 127-131.

Sergiu Aizicovici

Department of Mathematics, Ohio University, Athens, OH 45701-2979, USA

E-mail address: aizicovi@math.ohiou.edu

Simeon Reich

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

E-mail address: sreich@tx.technion.ac.il

Alexander J. Zaslavski

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

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