Electronic Journal of Differential Equations, Vol. 2010(2010), No. 19, pp. 1–7.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

MINIMIZING CONVEX FUNCTIONS BY CONTINUOUS DESCENT METHODS

SERGIU AIZICOVICI, SIMEON REICH, ALEXANDER J. ZASLAVSKI

Abstract. We study continuous descent methods for minimizing convex functions, defined on general Banach spaces, which are associated with an appropriate complete metric space of vector fields. We show that there exists an everywhere dense open set in this space of vector fields such that each of its elements generates strongly convergent trajectories.

1. Introduction and preliminaries

Discrete and continuous descent methods are important topics in optimization theory and in the study of dynamical systems. See, for example, the references in this article. Given a real-valued 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 $f^0(x, Vx) \le 0$ for all $x \in X$. Here $f^0(x, u)$ is the directional derivative of f at x in the direction of $u \in X$. In the discrete case, to each such vector field there correspond two gradient-like iterative processes. In the papers [10, 11, 12] 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. In our recent papers [1, 2, 3, 4] 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.

In the present paper, we improve upon the results of our previous papers by showing that in a certain space of Lipschitz vector fields, equipped with an appropriate complete metric, there exists an *open* everywhere dense set such that each of its elements generates strongly convergent trajectories.

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 $\bar{x} \in X$ such that $f(\bar{x}) \le f(x)$ 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

$$\lim_{n\to\infty} \|x_n - \bar{x}\| = 0.$$

2000 *Mathematics Subject Classification.* 37L99, 47J35, 49M99, 54E35, 54E50, 54E52, 90C25.
Key words and phrases. Complete uniform space; convex function; descent method; initial value problem; minimization problem.

Submitted June 27, 2009. Published January 28, 2010.

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

For each $x \in X$, let

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

Let $(X^*, \|\cdot\|)$ be the dual space of $(X, \|\cdot\|)$. For each $x \in X$ and each $r > 0$, 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 $\mathcal{A}_1$ the set of all mappings $V : X \to X$ such that the restriction of V to any bounded subset of X is Lipschitz (namely, $\mathrm{Lip}(V,r) < \infty$ for any positive r) and $f^0(x, Vx) \le 0$ for all $x \in X$.

For the set $\mathcal{A}_1$, we consider the uniformity determined by the base

$$\begin{aligned} E(n,\epsilon) = \big\{&(V_1,V_2) \in \mathcal{A}_1 \times \mathcal{A}_1 : \mathrm{Lip}(V_1 - V_2, n) \le \epsilon, \\ &\|V_1x - V_2x\| \le \epsilon \text{ for all } x \in B(n)\big\}, \end{aligned} \tag{1.4}$$

where $n, \epsilon > 0$.

Clearly, this uniform space $\mathcal{A}_1$ is metrizable and complete. We equip it with the topology induced by this uniformity. The following existence result was proved in [2, Section 3].

Proposition 1.1. *Let $x_0 \in X$ and $V \in \mathcal{A}_1$. Then there exists a unique continuously differentiable mapping $x : [0,\infty) \to X$ such that*

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

Let $x \in W^{1,1}(0,T;X)$; i.e.,

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

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. $t \in [0,T]$.

It was shown in [2] that 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:

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

We now quote [1, Proposition 3.1].

Proposition 1.2. *Assume that $t \in [0,T]$ and that both the derivatives $x'(t)$ and $(f \circ x)'(t)$ exist. Then*

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

Our results are proved under two additional assumptions on the function f. First, we assume in addition that

C(iv) f is Lipschitz on bounded subsets of X.

The second assumption concerns the existence of a sharp minimum. Specifically, this hypothesis can be formulated as follows.

C(v) There is $\Delta > 0$ such that
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\| \quad \text{for all } x \in X.$$
Clearly, we may assume without any loss of generality that $\Delta < 1$.

We next show that seemingly different variants of this notion are, in fact, equivalent.

Proposition 1.3. *The following conditions are equivalent.*

(i) *There is $\Delta > 0$ such that*
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\| \quad \textit{for all } x \in X;$$
(ii) *there exist $\Delta > 0$ and $r > 0$ such that*
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\| \quad \textit{for all } x \in X \textit{ satisfying } f(x) \leq f(\bar{x}) + r;$$
(iii) *there exist $\Delta > 0$ and $r > 0$ such that*
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\| \quad \textit{for all } x \in B(\bar{x}, r).$$

Proof. Clearly, (i) implies (ii). By the continuity of f, (ii) implies (iii).

It remains to show that (iii) implies (i). To this end, assume that (iii) holds. There are then $\Delta > 0$ and $r > 0$ such that
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\| \quad \text{for all } x \in B(\bar{x}, r). \tag{1.5}$$
Assume that
$$x \in X \setminus B(\bar{x}, r). \tag{1.6}$$
Put
$$z = \bar{x} + r\|x - \bar{x}\|^{-1}(x - \bar{x}). \tag{1.7}$$
By (1.7),
$$z \in B(\bar{x}, r). \tag{1.8}$$
By (1.5), (1.7) and (1.8),
$$f(z) \geq f(\bar{x}) + \Delta\|z - \bar{x}\| = f(\bar{x}) + \Delta r. \tag{1.9}$$
Since f is convex, it follows from (1.7) and (1.6) that
$$\begin{aligned} f(z) &= f(r\|x - \bar{x}\|^{-1}x + (1 - r\|x - \bar{x}\|^{-1})\bar{x}) \\ &\leq r\|x - \bar{x}\|^{-1}f(x) + (1 - r\|x - \bar{x}\|^{-1})f(\bar{x}). \end{aligned}$$
When combined with (1.9), this implies that
$$\begin{gathered} f(\bar{x}) + \Delta r \leq r\|x - \bar{x}\|^{-1}f(x) + (1 - r\|x - \bar{x}\|^{-1})f(\bar{x}), \\ r\|x - \bar{x}\|^{-1}f(\bar{x}) + \Delta r \leq r\|x - \bar{x}\|^{-1}f(x), \end{gathered}$$
whence
$$f(x) \geq f(\bar{x}) + \Delta\|x - \bar{x}\|. \tag{1.10}$$
Thus (1.10) holds for all $x \in X$ satisfying (1.6). Together with (1.5) this implies that (i) holds. Proposition 1.3 is proved. □

Proposition 1.4. *For each $h \in X$, $f^0(\bar{x}, h) \geq \Delta\|h\|$.*

The above propositon is a direct consequence of (1.1) and C(v).

Proposition 1.5. *For any $V \in \mathcal{A}_1$, $V\bar{x} = 0$.*

Proof. Let $V \in \mathcal{A}_1$. By Proposition 1.1, there is a continuously differentiable mapping $x : [0,\infty) \to X$ such that

$$x(0) = \bar{x}, \quad x'(t) = Vx(t), \quad t \in [0,\infty). \tag{1.11}$$

In view of (1.11), the inclusion $V \in \mathcal{A}_1$ and Proposition 1.2, the composition $f \circ x$ is a decreasing function. By C(ii) and C(iii), $x(t) = \bar{x}$ for all $t \ge 0$. When combined with (1.11), this implies that for all $t \ge 0$,

$$0 = x'(t) = V\bar{x}.$$

The proof is complete. $\square$

We are finally ready to state our main result. It improves upon previous results because it guarantees the existence of a good *open* everywhere dense set, and not just a good G_δ dense set. Its proof is given in Section 3 while Section 2 is devoted to several auxiliary results.

Theorem 1.6. *Let $M_0 > 0$. Then there exists an open everywhere dense $\mathcal{F} \subset \mathcal{A}_1$ such that for each $U \in \mathcal{F}$, there exists a neighborhood $\mathcal{U}$ of U in $\mathcal{A}_1$ so that the following property holds:*

For each $\epsilon > 0$, there exists $T_\epsilon > 0$ such that for each $W \in \mathcal{U}$ and each $x \in C^1([0,\infty); X)$ satisfying

$$x'(t) = Wx(t), \quad t \in [0,\infty), \quad \|x(0\| \le M_0,$$

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

Corollary 1.7. *Let $M_0 > 0$ and $U \in \mathcal{F}$, where $\mathcal{F}$ is the set the existence of which is guaranteed by Theorem 1.6. Let $x \in C^1([0,\infty); X)$ satisfy $x'(t) = Ux(t)$, $t \in [0,\infty)$, $\|x(0\| \le M_0$. Then $\lim_{t\to\infty} x(t) = \bar{x}$.*

2. Auxiliary results

For each $V \in \mathcal{A}_1$ and each $\gamma \in (0,1)$, set

$$V_\gamma x = Vx + \gamma(\bar{x} - x), \ x \in X. \tag{2.1}$$

Lemma 2.1 ([2, Lemma 4.1])**.** *Let $V \in \mathcal{A}_1$ and $\gamma \in (0,1)$. Then $V_\gamma \in \mathcal{A}_1$.*

Lemma 2.2 ([2, Lemma 4.2])**.** *Let $V \in \mathcal{A}_1$. Then $\lim_{\gamma\to 0^+} V_\gamma = V$.*

Lemma 2.3. *Let $M_0 > 0$. Then there exists $M > 0$ such that for each $V \in \mathcal{A}_1$ and each $x \in C^1([0,\infty); X)$ which satisfies*

$$x'(t) = Vx(t), \quad t \in [0,\infty), \quad \|x(0)\| \le M_0, \tag{2.2}$$

the following inequality holds:

$$\|x(t)\| \le M \quad \textit{for all } t \in [0,\infty). \tag{2.3}$$

Proof. By C(iv), there is $M_1 > 0$ such that

$$f(z) \le M_1 \text{ for all } x \in B(0, M_0). \tag{2.4}$$

By C(i), there is $M > 0$ such that

$$\|z\| \le M \quad \text{for all } z \in X \text{ satisfying } f(z) \le M_1. \tag{2.5}$$

Assume that $V \in \mathcal{A}_1$ and that $x \in C^1([0,\infty); X)$ satisfies (2.2). By (2.2), Proposition 1.2 and the inclusion $V \in \mathcal{A}_1$, the function $f \circ x$ is decreasing. Together with (2.2), (2.4) and (2.5) this implies that $f(x(t)) \le f(x(0)) \le M_1$ and $\|x(t)\| \le M$ for all $t \ge 0$. The proof is complete. $\square$

Lemma 2.4. *Let $V \in \mathcal{A}_1$, $\gamma \in (0,1)$ and $x \in X$. Then*

$$f^0(x, V_\gamma x) \le \gamma(f(\bar{x}) - f(x)) \le -\gamma\Delta\|x - \bar{x}\|. \tag{2.6}$$

Proof. By (2.1), the properties of the directional derivative of a convex function and the relation $V \in \mathcal{A}_1$,

$$\begin{aligned} f^0(x, V_\gamma x) &= f^0(x, Vx + \gamma(\bar{x} - x)) \\ &\le f^0(x, Vx) + \gamma f^0(x, \bar{x} - x) \\ &\le \gamma f^0(x, \bar{x} - x) \le \gamma(f(\bar{x}) - f(x)) \\ &\le -\gamma\Delta\|x - \bar{x}\|. \end{aligned}$$

The proof is complete. □

Lemma 2.5. *Let $V \in \mathcal{A}_1$, $\gamma \in (0,1)$ and $M > \|\bar{x}\|$. Then there exists a neighborhood $\mathcal{U}$ of V_γ in $\mathcal{A}_l$ such that for each $W \in \mathcal{U}$ and each $x \in B(0, M)$,*

$$f^0(x, Wx) \le (\gamma/2)(f(\bar{x}) - f(x)).$$

Proof. By C(iv), there exists $L_0 > 1$ such that

$$|f(z_1) - f(z_2)| \le L_0\|z_1 - z_2\| \quad \text{for all } z_1, z_2 \in B(0, M+1). \tag{2.7}$$

Choose

$$\delta \in (0, \gamma\bar{\Delta}(2L_0)^{-1}). \tag{2.8}$$

Put

$$\mathcal{U} = \{W \in \mathcal{A}_1 : (W, V_\gamma) \in E(M+1, \delta)\}. \tag{2.9}$$

Let

$$W \in \mathcal{U} \text{ and } x \in B(0, M). \tag{2.10}$$

By Lemma 2.4,

$$f^0(x, V_\gamma x) \le \gamma(f(\bar{x}) - f(x)). \tag{2.11}$$

By the properties of the directional derivative of a convex function, (2.11), (2.7), Proposition 1.5, (2.10), the inequality $M > \|\bar{x}\|$, (2.9), C(v) and (2.8),

$$\begin{aligned} f^0(x, Wx) &\le f^0(x, V_\gamma x) + f^0(x, Wx - V_\gamma x) \\ &\le \gamma(f(\bar{x}) - f(x)) + L_0\|Wx - V_\gamma x\| \\ &\le \gamma(f(\bar{x}) - f(x)) + L_0\|(W - V_\gamma)x - (W - V_\gamma)\bar{x}\| \\ &\le \gamma(f(\bar{x}) - f(x)) + L_0\delta\|x - \bar{x}\| \\ &\le (f(\bar{x}) - f(x))(\gamma - L_0\delta\bar{\Delta}^{-1}) \\ &\le (\gamma/2)(f(\bar{x}) - f(x)). \end{aligned}$$

The proof is complete. □

3. Completion of the proof of Theorem 1.6

Let $M_0 > 0$. By Lemma 2.3, there exists $M > \|\bar{x}\|$ such that the following property holds:

(P1) For each $V \in \mathcal{A}_1$ and each $x \in C^1([0,\infty); X)$ which satisfies

$$x'(t) = Vx(t), \quad t \in [0,\infty), \quad \|x(0)\| \le M_0,$$

the following inequality holds:

$$\|x(t)\| \le M, \quad t \in [0,\infty).$$

Let $V \in \mathcal{A}_1$ and $\gamma \in (0,1)$. By Lemma 2.5, there exists an open neighborhood $\mathcal{U}(V,\gamma)$ of V_γ in $\mathcal{A}_1$ such that following property holds:

(P2) For each $W \in \mathcal{U}(V,\gamma)$ and each $x \in B(0,M)$,
$$f^0(x, Wx) \le (\gamma/2)(f(\bar{x}) - f(x)).$$

Put
$$\mathcal{F} = \cup\{\mathcal{U}(V,\gamma) : V \in \mathcal{A}_1,\ \gamma \in (0,1)\}. \tag{3.1}$$
Clearly, $\mathcal{F}$ is an open everywhere dense subset of $\mathcal{A}_1$.

Let $U \in \mathcal{F}$. By (3.1) there are $V \in \mathcal{A}_1$ and $\gamma \in (0,1)$ such that
$$U \in \mathcal{U}(V,\gamma). \tag{3.2}$$
Let $\epsilon > 0$. By C(ii), there is $\epsilon_1 \in (0,\epsilon)$ such that
$$\|x - \bar{x}\| \le \epsilon \quad \text{for all } x \in X \text{ satisfying } f(x) \le f(\bar{x}) + \epsilon_1. \tag{3.3}$$
Choose a number
$$T_\epsilon > 1 + 4(\epsilon_1\gamma)^{-1}[\sup\{|f(z)| : z \in B(0,M_0)\} - f(\bar{x})]. \tag{3.4}$$
Assume that
$$W \in \mathcal{U}(V,\gamma), \quad x \in C^1([0,\infty); X) \tag{3.5}$$
and that
$$x'(t) = Wx(t), \quad t \in [0,\infty), \quad \|x(0)\| \le M_0.$$
To complete the proof it is sufficient to show that
$$\|x(t) - \bar{x}\| \le \epsilon, \quad t \ge T_\epsilon. \tag{3.6}$$
By (3.3), it suffices to show that
$$f(x(t)) \le f(\bar{x}) + \epsilon_1, \quad t \in [T_\epsilon, \infty). \tag{3.7}$$
Since $(f \circ x)$ is decreasing, (3.7) would follow once we prove that
$$\min\{f(x(t)) : t \in [0, T_\epsilon]\} \le f(\bar{x}) + \epsilon_1. \tag{3.8}$$
Assume the contrary. Then
$$f(x(t)) > f(\bar{x}) + \epsilon_1, \quad t \in [0, T_\epsilon]. \tag{3.9}$$
By (3.5) and (P1), we have
$$\|x(t)\| \le M, \quad t \in [0,\infty). \tag{3.10}$$
By (3.10), (3.9), (P2) and (3.5), for all $t \in [0, T_\epsilon]$,
$$f^0(x(t), Wx(t)) \le (\gamma/2)(f(\bar{x}) - f(x(t)) \le -\epsilon_1\gamma/2. \tag{3.11}$$
By Proposition 1.2, (3.5) and (3.11),
$$f(x(T_\epsilon)) - f(x(0)) = \int_0^{T_0} (f \circ x)'(t)dt = \int_0^{T_\epsilon} f^0(x(t), x'(t))dt < -T_\epsilon\epsilon_1\gamma/2$$
and (*cf.* (3.5) and (3.9))
$$T_\epsilon\epsilon_1\gamma/2 \le f(x(0)) - f(x(T_\epsilon)) \le \sup\{f(z) : z \in B(0,M_0)\} - f(\bar{x}).$$
This contradicts (3.4). The contradiction we have reached shows that (3.8) holds and therefore the proof of Theorem 1.6 is complete.

Acknowledgments. The second author was partially supported by the Fund for the Promotion of Research at the Technion and by the Technion President's Research Fund.

References

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

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

[3] S. Aizicovici, S. Reich and A. J. Zaslavski; *Stability of convergent continuous descent methods*, Electron. J. Differential Equations, **2007** (2007), No. 32, 1-6.

[4] S. Aizicovici, S. Reich and A. J. Zaslavski; *Asymptotic behavior of approximate solutions to evolution equations in Banach spaces*, Z. Anal. Anwend., **28** (2009), 295-303.

[5] R. Choudhary; *A stability theorem for convergence of a Lyapunov function along trajectories of nonexpansive semigroups*, Electron. J. Differential Equations, **2006** (2006), No. 120, 1-10.

[6] R. Choudhary; *Generic convergence of a convex Lyapunov function along trajectories of nonexpansive semigroups in Hilbert space*, J. Nonlinear Convex Anal., **7** (2006), 245-268.

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

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

[9] J. W. Neuberger; *Sobolev Gradients and Differential Equations*, Lecture Notes in Math. **1670**, Second Edition, Springer, Berlin, 2009.

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

[11] 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.

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

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

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`