International Journal of Pure and Applied Mathematics
Volume 32 No. 1 2006, 65-69

A NOTE ON THE MINIMIZATION OF CONVEX FUNCTIONS

Marina Arav[1], Simeon Reich[2 §], Alexander J. Zaslavski[3]

[1]Department of Mathematics and Statistics
Georgia State University
Atlanta, GA 30303, USA
e-mail: matmxa@langate.gsu.edu

[2,3]Department of Mathematics
The Technion-Israel Institute of Technology
Haifa, 32000, ISRAEL
[2]e-mail: sreichtx.technion.ac.il
[3]e-mail: ajzasltx.technion.ac.il

Abstract: We consider a complete metric space of Lipschitz mappings, acting on a bounded, closed and convex subset of a Banach space, which share a common convex Lyapunov function f. We show that the iterates of a generic element taken from this space converge (at an exponential rate) to the point where f attains its minimum.

AMS Subject Classification: Banach space, complete metric space, convex function, generic property, Lipschitz mapping, Lyapunov function, sharp minimum
Key Words: 49M99, 54E35, 54E52, 90C25, 90C48

1. Introduction

The study of minimization methods for convex functions is an important topic in optimization theory. In this note, we are given a convex, Lipschitz function f, defined on a bounded, closed and convex subset K of a Banach space X, which possesses a sharp minimum. A minimization algorithm is a self-mapping

Received: September 21, 2006

§Correspondence author

$A : K \to K$ such that $f(Ax) \le f(x)$ for all $x \in K$. We show that for most of these algorithms A, the sequences $\{A^n x\}_{n=1}^{\infty}$ tend to the point where f attains its sharp minimum (at an exponential rate) for all initial values $x \in K$.

When we say here that most of the elements of a complete metric space Y enjoy a certain property, we mean that the set of points which have this property contains an open, everywhere dense subset of Y. In particular, this property holds generically. Such an approach, when a certain property is investigated for the whole space Y and not just for a single point in Y, has already been successfully applied in many areas of analysis. Specifically, this approach has already been applied to the minimization of convex functions in [2, 6-9]. In those papers we established the existence of G_δ everywhere dense subsets of certain complete metric spaces of algorithms such that for any algorithm A in these subsets, the sequences $\{f(A^n x)\}_{n=1}^{\infty}$ tend to the infimum of f for all initial values $x \in K$. In the present paper we improve upon these results when the function f is Lipschitz and has a sharp minimum. As we have already mentioned, in this case we show (see the theorem below) that the sequences $\{A^n x\}_{n=1}^{\infty}$ themselves tend to the point where f attains its sharp minimum, at an exponential rate, for all values $x \in K$.

Let $K \subset X$ be a nonempty, bounded, closed and convex subset of a Banach space X. For each $A : K \to X$, set

$$\mathrm{Lip}(A) = \sup\{||Ax - Ay||/||x - y|| : \ x, y \in K \text{ such that } x \neq y\}. \tag{1.1}$$

Assume that $f : K \to R^1$ is a convex, Lipschitz function such that $\mathrm{Lip}(f) > 0$. We have

$$|f(x) - f(y)| \le \mathrm{Lip}(f)||x - y|| \text{ for all } x, y \in K.$$

Assume further that there exist a point $x_* \in K$ and a number $c_0 > 0$ such that

$$\inf(f) := \inf\{f(x) : \ x \in K\} = f(x_*)$$

and

$$f(x) \ge f(x_*) + c_0||x - x_*|| \text{ for all } x \in K. \tag{1.2}$$

In other words, we assume that the function f possesses a sharp minimum (cf. [1, 5]).

Denote by $\mathcal{A}$ the set of all self-mappings $A : K \to K$ such that $\mathrm{Lip}(A) < \infty$ and

$$f(Ax) \le f(x) \text{ for all } x \in K. \tag{1.3}$$

From the point of view of the theory of dynamical systems, each element of $\mathcal{A}$ describes a stationary dynamical system with a Lyapunov function f. Also, some optimization procedures in Hilbert and Banach spaces can be represented by elements of $\mathcal{A}$ (see the first example in [9, Section 3] and [3, 4]).

We equip the set $\mathcal{A}$ with the uniformity determined by the base

$$\mathcal{E}(\epsilon) = \{(A, B) \in \mathcal{A} \times \mathcal{A} : \ ||Ax - Bx|| \le \epsilon \ \text{ for all } x \in K \text{ and } \mathrm{Lip}(A - B) \le \epsilon\},$$

where $\epsilon > 0$. Clearly, the uniform space $\mathcal{A}$ is metrizable and complete.

Theorem. *There exists an open, everywhere dense subset $\mathcal{B} \subset \mathcal{A}$ such that for each $B \in \mathcal{B}$ there exist an open neighborhood $\mathcal{U}$ of B in $\mathcal{A}$ and a number $\lambda_0 \in (0, 1)$ such that for each $C \in \mathcal{U}$, each $x \in K$, and each natural number n,*

$$||C^n x - x_*|| \le c_0^{-1} \lambda^n (f(x) - f(x_*)).$$

2. Proof of Theorem

Let $\gamma \in (0, 1)$ and $A \in \mathcal{A}$. Set

$$A_\gamma x = (1 - \gamma)Ax + \gamma x_*, \ x \in K. \tag{2.1}$$

Clearly, for all $x \in K$,

$$f(A_\gamma x) \le (1 - \gamma) f(Ax) + \gamma f(x_*) \tag{2.2}$$

and

$$A_\gamma \in \mathcal{A}. \tag{2.3}$$

Next, we prove the following lemma.

Lemma. *Let $A \in \mathcal{A}$, $\gamma \in (0, 1)$ and $B \in \mathcal{A}$. Then for each $x \in K$,*

$$f(Bx) - f(x_*) \le [(1 - \gamma) + \mathrm{Lip}(f)\mathrm{Lip}(B - A_\gamma)c_0^{-1}](f(x) - f(x_*)).$$

Proof. Let $x \in K$. By (2.2), the relations $A_\gamma x_* = Bx_* = x_*$ and (1.2),

$$\begin{aligned} f(Bx) - f(x_*) &= f(A_\gamma x) - f(x_*) + f(Bx) - f(A_\gamma x) \\ &\le (1 - \gamma)(f(x) - f(x_*)) + \mathrm{Lip}(f)||Bx - A_\gamma x|| \\ &\le (1 - \gamma)(f(x) - f(x_*)) + \mathrm{Lip}(f)\mathrm{Lip}(B - A_\gamma)||x - x_*|| \\ &\le (1 - \gamma)(f(x) - f(x_*)) + \mathrm{Lip}(f)\mathrm{Lip}(B - A_\gamma)c_0^{-1}(f(x) - f(x_*)) \end{aligned}$$

$$\leq [(1-\gamma) + \mathrm{Lip}(f)\mathrm{Lip}(B - A_\gamma)c_0^{-1}](f(x) - f(x_*)).$$

The lemma is proved. □

Completion of Proof of Theorem. Let $A \in \mathcal{A}$ and $\gamma \in (0,1)$. Choose $r(\gamma) > 0$ such that

$$\lambda_\gamma := (1-\gamma) + \mathrm{Lip}(f) r(\gamma) c_0^{-1} < 1. \tag{2.4}$$

Denote by $\mathcal{U}(A,\gamma)$ an open neighborhood of A_γ in $\mathcal{A}$ such that

$$\mathcal{U}(A,\gamma) \subset \{B \in \mathcal{A} : \ (A_\gamma, B) \in \mathcal{E}(r(\gamma))\}. \tag{2.5}$$

Set

$$\mathcal{B} = \cup\{\mathcal{U}(A,\gamma) : \ A \in \mathcal{A}, \ \gamma \in (0,1)\}. \tag{2.6}$$

Clearly, we have for each $A \in \mathcal{A}$,

$$A_\gamma \to A \text{ as } \gamma \to 0^+.$$

Therefore $\mathcal{B}$ is an everywhere dense, open subset of $\mathcal{A}$. Let $B \in \mathcal{A}$. There are $A \in \mathcal{A}$ and $\gamma \in (0,1)$ such that

$$B \in \mathcal{U}(A,\gamma). \tag{2.7}$$

Assume that

$$C \in \mathcal{U}(A,\gamma) \text{ and } x \in K. \tag{2.8}$$

By the lemma, (2.8), (2.5) and (2.4),

$$f(Cx) - f(x_*) \leq [(1-\gamma) + \mathrm{Lip}(f)\mathrm{Lip}(C - A_\gamma)c_0^{-1}](f(x) - f(x_*))$$
$$\leq \lambda_\gamma (f(x) - f(x_*)).$$

This implies that for each $x \in K$ and each natural number n,

$$f(C^n x) - f(x_*) \leq \lambda_\gamma^n (f(x) - f(x_*)).$$

When combined with (1.2), this last inequality implies, in its turn, that for each $x \in K$ and each integer $n \geq 1$,

$$||C^n x - x_*|| \leq c_0^{-1}(f(C^n x) - f(x_*)) \leq c_0^{-1}\lambda_\gamma^n (f(x) - f(x_*)).$$

This completes the proof of Theorem. □

Acknowledgments

The work of the second author was partially supported by the Fund for the Promotion of Research at the Technion and by the Technion VPR Fund - B. and G. Greenberg Research Fund (Ottawa).

References

[1] J.V. Burke, M.C. Ferris, Weak sharp minima in mathematical programming, *SIAM J. Control Optim.*, **31** (1993), 1340-1359.

[2] M. Gabour, S. Reich, A.J. Zaslavski, A class of dynamical systems with a convex Lyapunov function, *Constructive, Experimental, and Nonlinear Analysis, Canadian Mathematical Society Conference Proceedings*, **27** (2000), 83-91.

[3] Yu.I. Lyubich, G.D. Maistrovskii, On the stability of relaxation processes, *Soviet Math. Dokl.*, **11** (1970), 311-313.

[4] Yu.I. Lyubich, G.D. Maistrovskii, The general theory of relaxation processes for convex functionals, *Russian Mathematical Surveys*, **25** (1970), 57-117.

[5] B.T. Polyak, Sharp minima, *Institute of Control Sciences Lecture Notes*, Moscow (1979).

[6] S. Reich, A.J. Zaslavski, On the minimization of convex functionals, *Calculus of Variations and Differential Equations*, Chapman and Hall/CRC Research Notes in Mathematics Series, **410**, CRC Press, Boca Raton, FL (1999), 200-209.

[7] S. Reich, A.J. Zaslavski, Asymptotic behavior of dynamical systems with a convex Lyapunov function, *J. Nonlinear and Convex Analysis*, **1** (2000), 107-113.

[8] S. Reich, A.J. Zaslavski, Generic convergence for a class of dynamical systems, *Nonlinear Analysis and Applications*, Kluwer, Dordrecht (2003), 851-859.

[9] S. Reich, A.J. Zaslavski, Convergence of dynamical systems with convex Lyapunov functions, *Nonlinear Analysis and Convex Analysis*, RIMS Proceedings Series, **1415**, Kyoto University, Kyoto (2005), 127-131.