International Journal of Pure and Applied Mathematics
Volume 20 No. 4 2005, 527-535

A NOTE ON THE MAXIMIZATION OF STRONGLY CONVEX FUNCTIONS

Altannar Chinchuluun[1 §], Enkhbat Rentsen[2], Panos M. Pardalos[3]
[1,3]Department of Industrial and Systems Engineering
University of Florida
303 Weil Hall, Gainesville, FL, 32611, USA
[1]e-mail: altannar@ufl.edu
[3]e-mail: pardalos@ufl.eud
[2]School of Mathematics and Computer Science
National University of Mongolia
Ulaanbaatar, MONGOLIA
e-mail: renkhbat@ses.edu.mn

Abstract: In this paper we consider the problem of maximizing a strongly convex function over an arbitrary set. We present an algorithm for solving this problem based on global optimality conditions. The proposed algorithm is shown to be globally convergent.

AMS Subject Classification: 52A41
Key Words: marginal functions, nonconvex programming

1. Introduction

In this paper we consider the problem of maximizing a convex function $f(x)$ over an arbitrary set, which may be nonconvex, $D \subset R^n$:

$$\begin{array}{ll} \max & f(x) \\ s.t. & x \in D. \end{array} \tag{1.1}$$

There are many iterative algorithms (Pardalos and Rosen [4, 5], and Tuy [7]) for solving the convex maximization problem on a polyhedral set by resorting

Received: April 17, 2005
§Correspondence author

to cutting plane and/or branch and bound techniques. The main purpose of this paper is to present a new algorithm with global convergence for solving the above problem with arbitrary feasible regions D. The proposed algorithm relies on global optimality conditions using properties of the marginal functions.

2. Global Maximality Conditions

Consider the problem of maximizing a strongly convex and continuously differentiable function $f: R^n \to R$ over an arbitrary compact set $D \subset R^n$ as a special case of problem (1.1).

$$\begin{aligned} \max \quad & f(x) \\ s.t. \quad & x \in D. \end{aligned} \tag{2.1}$$

In this case, Strekalovsky's global optimality conditions (Strekalovsky [6]) were simplified as follows in Enkhbat and Ibaraki [2].

Theorem 1. *A point $z \in D$ with $\nabla f(z) \neq 0$ is a solution of problem (2.1) if and only if the following conditions hold:*

$$(x-y)^T \nabla f(y) \leq 0 \text{ for all } y \in E_{f(z)}(f) \text{ and } x \in D, \tag{2.2}$$

where and $E_c(f) = \{y \in R^n \mid f(y) = c\}$.

3. A Maximization Algorithm and its Convergence

We now describe an algorithm, based on the optimality condition (2.2), for solving problem (2.1). For this purpose, we need to reformulate Theorem 1 in an equivalent form by introducing the following marginal functions:

$$\varphi(x,z) = \max_{y \in E_{f(z)}(f)} (x-y)^T \nabla f(y)$$

for all $x \in R^n$ and $z \in D$, and

$$\theta(z) = \max_{x \in D} \varphi(x,z).$$

There are useful properties of the marginal functions $\varphi(x,z)$ and $\theta(z)$.

Lemma 1. *$\varphi(x,z)$ is continuous with respect to x on R^n.*

Proof. Let Δx be a small increment at a point x. Then we have the following inequalities:

$$|\varphi(x+\Delta x, z) - \varphi(x, z)| = \Big| \max_{y \in E_{f(z)}(f)} (x + \Delta x - y)^T \nabla f(y) -$$
$$\max_{y \in E_{f(z)}(f)} (x - y)^T \nabla f(y) \Big| \leq \max_{y \in E_{f(z)}(f)} |(\Delta x)^T \nabla f(y)|$$
$$\leq \|\Delta x\| \max_{y \in E_{f(z)}(f)} \|\nabla f(y)\| .$$

Since $f(x)$ is continuously differentiable, and the set $E_{f(z)}(f)$ is compact, the function $\|\nabla f(y)\|$ is continuous and achieves its maximum, i.e., $\max_{y \in E_{f(z)}(f)} \|\nabla f(y)\|$ is bounded. Therefore, we have

$$\lim_{\Delta x \to 0} |\varphi(x+\Delta x, z) - \varphi(x, z)| = 0,$$

proving the lemma. □

Next using the function $\theta(z)$, we can reformulate Theorem 1 as follows.

Theorem 2. *Let $z \in D$ and $\nabla f(z) \neq 0$. If $\theta(z) = 0$ then the point z is a solution to problem (2.1).*

Proof. The proof is immediate from the following inequalities:

$$(x - y)^T f(y) \leq \max_{y \in E_{f(z)}(f)} \varphi(x, z) \leq \theta(z) \leq 0,$$

which hold for all $x \in D$ and $y \in E_{f(z)}(f)$. □

Now we are ready to present an algorithm based on the above results, in which we assume that the problem $\max_{x \in D} \varphi(x, z)$ can be solved exactly for any given $z \in D$.

Algorithm MAX

Input: A strongly convex function f and a compact set D.

Output: A solution x to problem (2.1); i.e., a global maximizer of f over D.

Step 1. Choose x^0 such that $\nabla f(x^0) \neq 0$. Set $k := 0$.

Step 2. Solve the following problem

$$\begin{aligned} \max \quad & \varphi(x, x^k) \\ s.t. \quad & x \in D. \end{aligned}$$

Let x^{k+1} be a solution of this problem, i.e.,

$$\theta(x^k) = \varphi(x^{k+1}, x^k) = \max_{x \in D} \varphi(x, x^k).$$

Step 3. If $\theta(x^k) = 0$, then output $x := x^k$ and terminate. Otherwise, set $k := k + 1$ and return to Step 2.

In order to ensure the convergence of Algorithm MAX, we use the following auxiliary result.

Theorem 3. (see Enkhbat [1]) *Suppose that $f : R^n \to R$ is strongly convex and differentiable. Let a sequence $\{x^k,\ k = 0, 1, ...\} \subset R^n$ be such that*

$$x^0 \neq \arg \min_{x \in R^n} f(x) \quad \textit{and} \quad f(x^0) < f(x^{k-1}) < ... < f(x^k).$$

Then there exists a positive constant $\delta > 0$ satisfying

$$\|\nabla f(x^k)\| \geq \delta, \quad k = 0, 1, ...$$

Then convergence of Algorithm MAX is given by the following theorem.

Theorem 4. *Assume that $f : R^n \to R$ be strongly convex and continuously differentiable, and D be a compact set. Then the sequence $\{x^k, k = 0, 1, ...\}$ generated by Algorithm MAX is a maximizing sequence for problem (2.1), that is,*

$$\lim_{k \to \infty} f(x^k) = \max_{x \in D} f(x),$$

and an accumulation point of the sequence is a global maximizer of (2.1).

Proof. Note that Algorithm MAX generates points $x^k \in D$ and $f(x^k) \leq f^*$, where $f^* = f(x^*) = \max_{x \in D} f(x)$. Clearly, for all $y \in E_{f(x^k)}(f)$ and $x \in D$ we have

$$(x - y)^T \nabla f(y) \leq \varphi(x, x^k) \leq \theta(x^k).$$

If there exists a k such that $\theta(x^k) = 0$ then, by Theorem 2, x^k is a solution to problem (2.1) and, consequently, the desired result follows. Therefore, we suppose $\theta(x^k) > 0$ for all $k = 0, 1, ...,$ and prove the theorem by contradiction. In this case, since $\{x^k\}$ is not a maximizing sequence, we have

$$\limsup_{k \to \infty}\ f(x^k) < f^*. \tag{3.1}$$

First, we show that the sequence $\{f(x^k)\}$ is strictly increasing. By assumption, we have

$$\theta(x^k) = \varphi(x^{k+1}, x^k) = (x^{k+1} - y^k)^T \nabla f(y^k) > 0,$$

where $y^k \in E_{f(x^k)}(f)$. Then the convexity of f implies that

$$f(x^{k+1}) - f(x^k) = f(x^{k+1}) - f(y^k) \geq (x^{k+1} - y^k)^T \nabla f(y^k) > 0. \quad (3.2)$$

Therefore, $f(x^{k+1}) > f(x^k)$ for all k. Furthermore, as the sequence $\{f(x^k)\}$ is bounded from above by f^*, it has a limit:

$$\lim_{k \to \infty} f(x^k) = A < +\infty,$$

and hence we have

$$\lim_{k \to \infty} (f(x^{k+1}) - f(x^k)) = 0. \quad (3.3)$$

Then from (3.2) and (3.3), we can conclude that

$$\lim_{k \to \infty} \theta(x^k) = 0. \quad (3.4)$$

Now introduce the following sets which are closed and convex.

$$C_k = \{x \in R^n : f(x) \leq f(x^k)\}, \quad k = 0, 1, 2, \ldots.$$

It is clear that $x^* \notin C_k$. On the other hand, from the construction of Algorithm MAX, we have $\nabla f(x^0) \neq 0$, consequently, $x^0 \neq \arg\min_{x \in R^n} f(x)$. Hence, *int* $C_k \neq \emptyset$. Then take the projection $u^k \in C_k$ of the point x^* on C_k such that

$$\|u^k - x^*\|^2 = \min_{x \in C_k} \|x - x^*\|^2. \quad (3.5)$$

Note that

$$\|u^k - x^*\| > 0 \quad (3.6)$$

holds because $x^* \notin C_k$. The optimality conditions at the solution u^k for the convex minimization problem (3.5) is given as follows,

$$\begin{cases} u^k - x^* + \lambda_k \nabla f(u^k) = 0, \\ f(u^k) = f(x^k), \end{cases} \quad (3.7)$$

where λ_k is the Lagrange multiplier. Hence, we have

$$\lambda_k = \frac{\|u^k - x^*\|}{\|\nabla f(u^k)\|}. \quad (3.8)$$

Then condition $f(u^k) = f(x^k)$ of (3.7) and $\theta(x^k)$ imply

$$\theta(x^k) = \max_{x \in D} \varphi(x, x^k) \geq \max_{y \in E_{f(x^k)}(f)} (x^* - y)^T \nabla f(y)$$
$$\geq (x^* - u^k)^T \nabla f(u^k). \quad (3.9)$$

Using (3.7), (3.8) and (3.9), we have

$$(x^* - u^k)^T \nabla f(u^k) = \|\nabla f(u^k)\| \|x^* - u^k\| \leq \theta(x^k). \tag{3.10}$$

By Theorem 3, we conclude that

$$0 \leq \delta \|x^* - u^k\| \leq \theta(x^k).$$

Taking into account the inequality (3.4), we have

$$\lim_{k \to \infty} u^k = x^*.$$

Hence, by continuity of f on R^n,

$$\lim_{k \to \infty} f(x^k) = \lim_{k \to \infty} f(u^k) = f(x^*), \tag{3.11}$$

which yields a contradiction to (3.1).

Consequently, $\{x^k\} \subset D$ is a maximizing sequence for problem (2.1). Since D is compact, we can always select a convergent subsequence $\{x^{k_l}\}$ from $\{x^k\}$ such that

$$\lim_{l \to \infty} x^{k_l} = \bar{x} \in D.$$

Then together with (3.11), we obtain

$$\lim_{l \to \infty} f(x^{k_l}) = f(\bar{x}) = f^*,$$

which completes the proof. □

Remark 1. An Algorithm for solving problem (2.1) based on the so called gap function (Hearn [3])

$$g(y) = \max_{x \in D} (x - y)^T \nabla f(y)$$

was given in Enkhbat [1].

Remark 2. It is hard to numerically implement the above algorithm because the problem

$$\begin{array}{ll} \max & \varphi(x, x^k) \\ s.t. & x \in D \end{array}$$

may not be simpler than the original problem. That is why the algorithm might be referred to as theoretical. However, we do not need to solve the problem exactly at iteration k. If we have $\varphi(w, x^k) > 0$ for an admissible $w \in D$, it implies that there exists a y^k such that

$$\varphi(w, x^k) = \max_{y \in E_{f(x^k)}(f)} (w - y)^T \nabla f(y) \geq (w - y^k)^T \nabla f(y^k) > 0,$$

i.e., $f(w) > f(x^k)$, and thus we can improve the current value of $f(x^k)$ in Algorithm MAX. Furthermore, if the second order derivatives of f exist. Then the problem

$$\begin{array}{ll} \max & (w-y)^T \nabla f(y) \\ s.t. & y \in E_{f(x^k)}(f) \end{array}$$

can be solved by the Lagrangian method. In this case, we write down the optimality condition for this problem and obtain the following system of equations:

$$\begin{cases} \nabla^2 f(y)(w - y) + (\lambda - 1)\nabla f(y) = 0, \\ f(y) = f(x^k). \end{cases}$$

On the other hand, in Algorithm MAX, we do not need to solve problem

$$\begin{array}{ll} \max & \varphi(x, x^k) \\ s.t. & x \in D \end{array}$$

exactly at each iteration. We can find only ε approximate solutions to these problems. Then we find an approximate solution to the original problem (2.1) with a given accuracy. To make Algorithm MAX numerically implementable for this purpose, we need adapt it as follows.

Algorithm ϵ-MAX

Input: A strongly convex function f and a simple set D, and a sequence $\{\varepsilon_k\}$ such that $\varepsilon_k > 0$ for all k and $\sum_{k=0}^{\infty} \varepsilon_k < +\infty$.

Output: A ε approximate global maximizer x of f over D.

Step 1. Choose a feasible solution $x^0 \in D$ and a sequence $\{\varepsilon_k\}$ such that $\nabla f(x^0) \neq 0$. Set $k := 0$.

Step 2. Find an ε_k solution $\bar{x}^k$ of the problem

$$\begin{array}{ll} \max & \varphi(x, x^k) \\ s.t. & x \in D \end{array}$$

that is, $\varphi(\bar{x}^k, x^k) = \max_{y \in E_{f(x^k)}(f)} (\bar{x}^k - y)^T \nabla f(y) \geq \max_{x \in D} \varphi(x, x^k) - \epsilon_k$.
Let $\bar{y}^k \in E_{f(x^k)}(f)$ be a solution satisfying $\varphi(\bar{x}^k, x^k) = (\bar{x}^k - \bar{y}^k)^T \nabla f(\bar{y}^k)$.

Step 3. If $\varphi(\bar{x}^k, x^k) \leq -\varepsilon_k$ then output $x := \bar{x}^k$ and terminate. Otherwise, set $x^k := \bar{x}^k, k := k + 1$ and return to Step 2.

In order to ensure the convergence of this algorithm, first we need to have the following theorems.

Theorem 5. (see Enkhbat and Ibaraki [2]) *Assume that $f : R^n \to R$ be strongly convex and continuously differentiable. Let sequences $\{x^k\} \subset R^n$ and $\{\varepsilon_k\} \subset R$ be such that*

$$f(x^k) \geq f(x^{k-1}) - \varepsilon_k, \quad \varepsilon_k > 0 \text{ for } k = 1, 2, \dots,$$
$$\sum_{k=1}^{\infty} \varepsilon_k + f_* < f(x^0),$$

where $f_* = \min_{x \in R^n} f(x)$. Then there exists a positive constant $\bar{\delta} > 0$ satisfying

$$\lim_{k \to \infty} \inf \ \|\nabla f(x^k)\| \geq \bar{\delta}.$$

Theorem 6. (see Enkhbat and Ibaraki [2]) *Assume that $f : D \to R$ be a continuous function defined on a compact set $D \subset R^n$. Let sequences $\{x^k\} \subset D$ and $\{\varepsilon_k\} \subset R$ be such that*

$$f(x^k) \geq f(x^{k-1}) - \varepsilon_k, \quad \varepsilon_k > 0 \text{ for } k = 1, 2, \dots,$$
$$\sum_{k=1}^{\infty} \varepsilon_k < +\infty.$$

Then there exists a finite limit of the sequence $\{f(x^k)\}$, that is, $\lim_{k \to \infty} f(x^k) = A < +\infty$.

Theorem 7. *Assume that $f : R^n \to R$ be strongly convex and continuously differentiable. Let a sequence $\{\varepsilon_k\} \subset R$ be such that $\varepsilon_k > 0$ for $k = 1, 2, \dots$ and*

$$\sum_{k=1}^{\infty} \varepsilon_k + f_* < f(x^0),$$

where $f_* = \min_{x \in R^n} f(x)$. Then the sequence $\{\bar{x}^k\}$ generated by Algorithm ϵ-MAX is a maximizing sequence for problem (2.1), that is,

$$\lim_{k \to \infty} f(\bar{x}^k) = \max_{x \in D} f(x),$$

and every accumulation point of the sequence $\{\bar{x}^k\}$ *is a global maximizer of (2.1).*

The proofs follow similarly from Theorems 4-7.

Acknowledgements

This research was partially supported by NSF and CRDF grants.

References

[1] R. Enkhbat, An algorithm for maximizing a convex function over a simple set, *Journal of Global Optimization*, **8** (1996), 379-391.

[2] R. Enkhbat, T. Ibaraki, On the maximization and minimization of quasiconvex function, *Journal of Nonlinear and Convex Analysis*, **4** (2003), 43-76.

[3] D.W. Hearn, The gap function of a convex program, *Operations Research Letters*, **1** (1982), 67-71.

[4] P.M. Pardalos, J.B. Rosen, *Constrained Global Optimization: Algorithms and Applications*, Lecture Notes in Computer Science, **268**, Springer, Berlin (1987).

[5] P.M. Pardalos, J.B. Rosen, Methods for global concave minimization: A bibliographic survey, *SIAM Review*, **28** (1986), 367-379.

[6] A.S. Strekalovsky, Global optimality conditions for nonconvex optimization, *Journal of Global Optimization*, **12** (1998), 415-434.

[7] H. Tuy, Normal conical algorithm for concave minimization over polytopes, *Mathematical Programming*, **51** (1991), 229-245.