Minimization of an M-convex Function*

Akiyoshi SHIOURA†

Abstract

We study the minimization of an M-convex function introduced by Murota. It is shown that any vector in the domain can be easily separated from a minimizer of the function. Based on this property, we develop a polynomial time algorithm.

Keywords: matroid, base polyhedron, convex function, minimization.

1 Introduction

M-convex function, recently introduced by Murota [8, 9, 10], is an extension of valuated matroid due to Dress and Wenzel [1, 2] as well as a quantitative generalization of (the integral points of) the base polyhedron of an integral submodular system [4]. M-convexity is quite a natural concept appearing in many situations; linear and separable-convex functions are M-convex, and more general M-convex functions arise from the minimum cost flow problems with separable-convex cost functions. M-convex function enjoys several nice properties which persuade us to regard it as "convexity" in combinatorial optimization. Let V be a finite set with cardinality n. A function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is said to be M-convex if it satisfies

(M-EXC) $\forall x, y \in \mathrm{dom}\, f$, $\forall u \in \mathrm{supp}^+(x-y)$, $\exists v \in \mathrm{supp}^-(x-y)$ such that

$$f(x) + f(y) \geq f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v),$$

where $\mathrm{dom}\, f = \{x \in \mathbf{Z}^V \mid f(x) < +\infty\}$, $\mathrm{supp}^+(x-y) = \{w \in V \mid x(w) > y(w)\}$, $\mathrm{supp}^-(x-y) = \{w \in V \mid x(w) < y(w)\}$, and $\chi_w \in \{0,1\}^V$ is the characteristic vector of $w \in V$. For an M-convex function f with $\mathrm{dom}\, f \subseteq \{0,1\}^V$, $-f$ is a valuation on a matroid in the sense of [1, 2]. The property (M-EXC) implies that $\mathrm{dom}\, f$ is a base polyhedron.

In this paper, we consider the problem of minimizing an M-convex function. While the concept of M-convexity is quite new and no efficient algorithm is known yet, several polynomial

*Research Report No. 20, Department of Mechanical Engineering, Sophia University, 1997. (version: 17/11/97)
†Department of Mechanical Engineering, Sophia University, 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102, Japan, shioura@keka.me.sophia.ac.jp.

time algorithms are proposed for special cases of M-convex functions. It is well-known that a linear function can be easily minimized over a base polyhedron by a simple greedy algorithm (see [4]). A strongly-polynomial time algorithm was proposed by Fujishige [3] for a separable-convex quadratic function, and weakly-polynomial time algorithms were given by Groenevelt [6] and Hochbaum [7] for a general separable-convex function. It was reported that there is no strongly-polynomial time algorithm for a general separable-convex function [7].

The aim of this paper is to develop an efficient algorithm for minimizing an M-convex function. Since the local optimality is equal to the global optimality, an optimal solution can be found by a descent method, which does not necessarily terminate in polynomial time. Instead, we propose a different approach based on the property that any vector in the domain can be efficiently separated from a minimizer of the function, which is shown later. Each iteration finds a certain vector in the current domain, and divides the domain so that the vector and an optimal solution are separated. By a clever choice of the vector, the size of the domain reduces in a certain ratio iteratively, which leads to a weakly-polynomial time algorithm.

2 Theorems

Throughout the paper we suppose $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is an M-convex function with bounded domain. The global minimality of an M-convex function is characterized by the local minimality.

Theorem 2.1 ([8, 10]) *For any $x \in \mathrm{dom}\, f$, $f(x) \le f(y)$ $(\forall y \in \mathbf{Z}^V)$ if and only if $f(x) \le f(x - \chi_u + \chi_v)$ $(\forall u, v \in V)$.* ∎

Any vector in $\mathrm{dom}\, f$ can be easily separated from some minimizer of f.

Theorem 2.2 *(i) For $x \in \mathrm{dom}\, f$ and $v \in V$, let $u \in V$ satisfy $f(x-\chi_u+\chi_v) = \min_{s \in V}\{f(x-\chi_s+\chi_v)\}$. Set $x' = x - \chi_u + \chi_v$. Then, there exists $x^* \in \arg\min f$ with $x^*(u) \le x'(u)$.*
(ii) For $x \in \mathrm{dom}\, f$ and $u \in V$, let $v \in V$ satisfy $f(x - \chi_u + \chi_v) = \min_{t \in V}\{f(x - \chi_u + \chi_t)\}$. Set $x' = x - \chi_u + \chi_v$. Then, there exists $x^* \in \arg\min f$ with $x^*(v) \ge x'(v)$.

Proof. We prove the first claim only. Let $x^* \in \arg\min f$ with the minimum value of $x^*(u)$, and to the contrary suppose $x^*(u) > x'(u)$. By (M-EXC), there exists $w \in \mathrm{supp}^-(x^* - x')$ such that $f(x^*) + f(x') \ge f(x^* - \chi_u + \chi_w) + f(x + \chi_v - \chi_w)$. The assumptions for x^* and x' imply $x^* - \chi_u + \chi_w \in \arg\min f$, a contradiction. ∎

Corollary 2.3 *Let $x \in \mathrm{dom}\, f$ with $x \notin \arg\min f$, and $u, v \in V$ satisfy $f(x - \chi_u + \chi_v) = \min_{s,t \in V}\{f(x - \chi_s + \chi_t)\}$. Then, there exists $x^* \in \arg\min f$ with $x^*(u) \le x(u) - 1$, $x^*(v) \ge x(v) + 1$.*

Let $B \subseteq \mathbf{Z}^V$ be a base polyhedron, i.e., B satisfies the next property:

> **(B-EXC)** $\forall x, y \in \mathrm{dom}\, f$, $\forall u \in \mathrm{supp}^+(x - y)$, $\exists v \in \mathrm{supp}^-(x - y)$ such that $x - \chi_u + \chi_v,\ y + \chi_u - \chi_v \in B$.

Assume B is bounded. We define the *narrowed base polyhedron* $N_B(\subseteq B)$ of B as follows. For each $w \in V$, define

$$l_B(w) = \min_{y \in B}\{y(w)\}, \quad u_B(w) = \max_{y \in B}\{y(w)\}, \tag{1}$$

$$l'_B(w) = \lfloor (1-1/n)l_B(w) + (1/n)u_B(w) \rfloor,\ u'_B(w) = \lceil (1/n)l_B(w) + (1-1/n)u_B(w) \rceil. \tag{2}$$

Then, N_B is defined as $N_B = \{y \in B \mid l'_B(w) \le y(w) \le u'_B(w)\ (\forall w \in V)\}$. We see from definition that N_B is a base polyhedron if it is not empty.

Theorem 2.4 $N_B \neq \emptyset$.

Proof. Let $\rho : 2^V \to \mathbf{Z}$ be the submodular function with $\rho(\emptyset) = 0$ and $B = \{y \in \mathbf{Z}^V \mid y(X) \le \rho(X)\ (\forall X \subseteq V), y(V) = \rho(V)\}$. Note that $l_B(w) = \rho(V) - \rho(V - w)$, $u_B(w) = \rho(w)\ (\forall w \in V)$. It suffices to show the following (see [4, Theorem 3.8]):

$$\text{(i) } l'_B(X) \le \rho(X)\ (\forall X \subseteq V), \qquad \text{(ii) } u'_B(X) \ge \rho(V) - \rho(V{-}X)\ (\forall X \subseteq V).$$

Since (ii) can be shown similarly, we prove (i) only. Let $X \subseteq V$ with cardinality k. We claim

$$k\rho(X) + k\sum_{v \in X}\{\rho(V-v) - \rho(V)\} \ge \sum_{v \in X}\{\rho(v) + \rho(V-v) - \rho(V)\}. \tag{3}$$

Indeed, we have

$$\begin{aligned}
\text{LHS} &= k\rho(X) + \sum_{v \in X}\sum_{w \in X-v}\{\rho(V-w) - \rho(V)\} + \sum_{v \in X}\{\rho(V-v) - \rho(V)\} \\
&\ge k\rho(X) + \sum_{v \in X}\{\rho(V-(X-v)) - \rho(V)\} + \sum_{v \in X}\{\rho(V-v) - \rho(V)\} \ \ge\ \text{RHS},
\end{aligned}$$

where the inequalities are by the submodularity of ρ. Since the LHS is nonnegative, k in (3) can be replaced by $n(\ge k)$. Thus,

$$\rho(X) \ge (1-1/n)\sum_{v \in X}\{\rho(V) - \rho(V-v)\} + (1/n)\sum_{v \in X}\rho(v) \ge l'_B(X). \qquad \blacksquare$$

For $x \in B$ and $u, v \in V$, define

$$\tilde{c}_B(x, v, u) = \max\{\alpha \mid \alpha \in \mathbf{Z}, x + \alpha(\chi_v - \chi_u) \in B\}\ (\ge 0),$$

which is called the exchange capacity associated with x, v and u. For any α with $0 \le \alpha \le \tilde{c}_B(x, v, u)$, we have $x + \alpha(\chi_v - \chi_u) \in B$. The next theorem shows that a vector in N_B can be computed efficiently by using the exchange capacity.

Theorem 2.5 (cf. [4, Theorem 3.27]) *A vector in N_B can be obtained by evaluating the exchange capacity associated with B at most n^2 times, provided a vector in B is given.*

Proof. Suppose we are given a vector $x_0 \in B$ with either $x_0(u) < l'_B(u)$ or $x_0(u) > u'_B(u)$ for some $u \in V$. It suffices to show that the following algorithm finds $x \in B$ such that

$$l'_B(w) \leq x(w) \leq u'_B(w) \text{ if } l'_B(w) \leq x_0(w) \leq u'_B(w) \ (\forall w \in V - u), \quad l'_B(u) \leq x(u) \leq u'_B(u)$$

by evaluating the exchange capacity at most n times. Assume w.l.o.g. that $x_0(u) > u'_B(u)$, $n \geq 2$ and $V = \{u, v_1, v_2, \cdots, v_{n-1}\}$.

Step 0: Set $x := x_0$, $i := 1$.
Step 1: If $x(v_i) < u'_B(v_i)$, set $\alpha := \min\{\tilde{c}_B(x, v_i, u), x(u) - u'_B(u), u'_B(v_i) - x(v_i)\}$,
$x := x + \alpha(\chi_{v_i} - \chi_u)$.
Step 2: If $i = n - 1$ or $x(u) = u'_B(u)$ then stop; otherwise $i := i + 1$ and go to Step 1.

To the contrary assume $x(u) > u'_B(u)$ for the vector x obtained by the algorithm. Let x_* be any vector in N_B. Since $x(u) > u'_B(u) \geq x_*(u)$, (B-EXC) implies that the existence of $v_i \in V - u$ with $x' = x - \chi_u + \chi_{v_i} \in B$ holds for some i with $x(v_i) < x^*(v_i)$ $(\leq u'_B(v_i))$. Let x_i be the vector x after Step 1 of the i-th iteration. Then, it holds $x'(u) < x_i(u)$, $x'(w) \geq x_i(w)$ $(\forall w \in V - u)$ and $x'(v_i) > x_i(v_i)$. Hence, $\tilde{c}_B(x_i, v_i, u) = 0$. On the other hand, we have $x_i + \chi_{v_i} - \chi_u \in B$ by applying (B-EXC) to x', x_i and v_i, a contradiction. ∎

The values $l_B(w)$ and $u_B(w)$ defined by (1) can be computed in the similar way.

Theorem 2.6 *For any $w \in V$, the values $l_B(w)$ and $u_B(w)$ can be computed by evaluating the exchange capacity associated with B at most n times, provided a vector in B is given.*

3 Algorithms

Theorem 2.1 immediately leads to the following algorithm.

Algorithm STEEPEST_DESCENT
Step 0: Let x be any vector in $\mathrm{dom}\, f$.
Step 1: If $f(x) = \min_{s,t \in V}\{f(x - \chi_s + \chi_t)\}$ then stop. x is a minimizer.
Step 2: Find $u, v \in V$ with $f(x - \chi_u + \chi_v) = \min_{s,t \in V}\{f(x - \chi_s + \chi_t)\}$.
Step 3: Set $x := x - \chi_u + \chi_v$. Go to Step 1. □

This algorithm always terminates since the function value of x decreases strictly in each iteration. However, there is no guarantee for the polynomiality of the number of iterations.

The next algorithm maintains a set B $(\subseteq \mathrm{dom}\, f)$ which is a base polyhedron containing a minimizer of f. It reduces B iteratively by exploiting Corollary 2.3 and finally finds a minimizer.

Algorithm DOMAIN_REDUCTION
Step 0: Set $B := \mathrm{dom}\, f$.

Step 1: Find a vector $x \in N_B$.

Step 2: If $f(x) = \min_{s,t \in V}\{f(x - \chi_s + \chi_t)\}$ then stop.

Step 3: Find $u, v \in V$ with $f(x - \chi_u + \chi_v) = \min_{s,t \in V}\{f(x - \chi_s + \chi_t)\}$.

Step 4: Set $B := B \cap \{y \in \mathbf{Z}^V \mid y(u) \le x(u) - 1,\ y(v) \ge x(v) + 1\}$. Go to Step 1. □

We analyze the number of iterations of the algorithm. Denote by B_i the set B in the i-th iteration, and let $l_i(w) = l_{B_i}(w)$, $u_i(w) = u_{B_i}(w)$ for each $w \in V$. It is clear that $u_i(w) - l_i(w)$ is monotonically nonincreasing w.r.t. i. Furthermore, we have the following:

Lemma 3.1 *$u_{i+1}(w) - l_{i+1}(w) < (1 - 1/n)\{u_i(w) - l_i(w)\}$ for $w \in \{u, v\}$, where $u, v \in V$ are the elements found in Step 3.*

Proof. We show the case $w = u$. Let $x \in N_{B_i}$ be the vector chosen in Step 1. Then,

$$\begin{aligned} u_{i+1}(u) - l_{i+1}(u) &\le x(u) - 1 - l_i(u) \\ &\le \lceil (1/n) l_i(u) + (1 - 1/n) u_i(u) \rceil - 1 - l_i(u) < (1 - 1/n)\{u_i(u) - l_i(u)\}. \end{aligned}$$

The proof for the case $w = v$ is similar and omitted. ■

Let $L = \max_{w \in V}\{u_1(w) - l_1(w)\}$.

Lemma 3.2 *The algorithm* DOMAIN_REDUCTION *terminates in $\mathrm{O}(n^2 \log L)$ iterations.*

Proof. Since the value $u_i(w) - l_i(w)$ $(w \in V)$ is a nonnegative integer, the algorithm stops if $u_i(w) - l_i(w) < 1$ for all $w \in V$. Let k be the minimum integer with $(1 - 1/n)^k\{u_1(w) - l_1(w)\} < 1$. Suppose $u_1(w) \ne l_1(w)$ and $n \ge 2$. Then,

$$k \le -\ln\{u_1(w) - l_1(w)\} / \ln(1 - 1/n) + 1 \le n \ln\{u_1(w) - l_1(w)\} + 1.$$

by a well-known inequality $\ln z \le z - 1$ $(\forall z > 0)$. Thus the claim follows. ■

In the following, we explain how to perform each step, especially how to find a vector in N_B. We assume that a vector $x_0 \in \mathrm{dom}\, f$ and the value L are given in advance.

We maintain the set B by using two vectors a, b with $-a(w), b(w) \in \mathbf{Z} \cup \{+\infty\}$ $(\forall w \in V)$ as $B = \mathrm{dom}\, f \cap \{y \in \mathbf{Z}^V \mid a(w) \le y(w) \le b(w)\ (\forall w \in V)\}$. Maintenance of a, b is easy: initially set $a(w) = -\infty, b(w) = +\infty$ $(\forall w \in V)$, and update only the values $a(v)$ and $b(u)$ to $x(v) + 1$, $x(u) - 1$, respectively in Step 4 of each iteration.

When finding a vector in N_B, we first compute the values $l_B(w), u_B(w)$ $(\forall w \in V)$ defined by (1), which can be done by $\mathrm{O}(n^2)$-time evaluation of the exchange capacity associated with B from Theorem 2.6. The exchange capacity can be computed in $\mathrm{O}(\log L)$ time by the binary search since $0 \le \tilde{c}_B(x, v, u) \le L$ $(\forall x \in B, \forall u, v \in V)$. Then, we compute $l'_B(w), u'_B(w)$ $(\forall w \in V)$ defined by (2) by using floor and ceiling operations. Note that floor and ceiling operations can be performed easily since n is the denominator of each value for which floor or ceiling is operated.

After computing the values $l'_B(w), u'_B(w)$ we can find $x \in N_B$ by $\mathrm{O}(n^2)$-time evaluation of the exchange capacity. Thus, Step 1 can be performed in $\mathrm{O}(n^2 \log L)$ time.

The other steps require $\mathrm{O}(n^2)$-time evaluation of f.

Theorem 3.3 *If a vector in* dom f *and the value* L *are given, the algorithm* DOMAIN_REDUCTION *finds a minimizer of* f *in* $\mathrm{O}(n^4 \log^2 L)$ *time.*

References

[1] A. W. M. Dress and W. Wenzel, Valuated matroid: A new look at the greedy algorithm, Appl. Math. Lett. 3 (1990) 33–35.

[2] A. W. M. Dress and W. Wenzel, Valuated matroids, Adv. Math. 93 (1992) 214–250.

[3] S. Fujishige, Lexicographically optimal base of a polymatroid with respect to a weight vector, Math. Oper. Res. 5 (1980) 186–196.

[4] S. Fujishige, Submodular Functions and Optimization (Annals of Discrete Mathematics 47, North-Holland, Amsterdam, 1991).

[5] M. Grötschel, L. Lovász and A. Schrijver, Geometric Algorithms and Combinatorial Optimization (Springer-Verlag, Berlin, 1988).

[6] H. Groenevelt, Two algorithms for maximizing a separable concave function over a polymatroid feasible region, Euro. J. Oper. Res. 54 (1991) 227–236.

[7] D. S. Hochbaum, Lower and upper bounds for the allocation problem and other nonlinear optimization problems, Math. Oper. Res. 19 (1994) 390–409.

[8] K. Murota, Submodular flow problem with a nonseparable cost function, Report No. 95843-OR, Forschungsinstitut für Diskrete Mathematik, Universität Bonn (1995).

[9] K. Murota, Convexity and Steinitz's exchange property, Adv. Math. 124 (1996) 272–311.

[10] K. Murota, Discrete convex analysis, RIMS preprint, No. 1065, Kyoto University (1996).