RIMS Preprint No. 1622

# Minimizing Discrete Convex Functions with Linear Inequality Constraints

Satoru FUJISHIGE,[*] Takumi HAYASHI,[†] and Kiyohito NAGANO[‡]

March 6, 2008

**Abstract**

A class of discrete convex functions that can efficiently be minimized has been considered by Murota. Among them are L$^\natural$-convex functions, which are natural extensions of submodular set functions. We first consider the problem of minimizing an L$^\natural$-convex function with a linear inequality constraint having a positive normal vector. We propose a polynomial algorithm to solve it based on a binary search for an optimal Lagrange multiplier, where use is made of algorithms for minimum-ratio and maximum-ratio problems that are, respectively, associated with submodular and supermodular set functions. We also examine an extension of the problem to that with a linear inequality constraint having a not necessarily positive normal vector and adapt it to the problem of minimizing an M$^\natural$-convex function, the convex conjugate of an L$^\natural$-convex function, with a linear inequality constraint. The former extension can be solved in polynomial time by using a binary search for an optimal Lagrange multiplier and by adopting Nagano's algorithm for the intersection of line and a base polyhedron. The latter can also be solved in polynomial time by an approach similar to that for L$^\natural$-convex functions, based on a geometric characterization of M$^\natural$-convex functions.

## 1. Introduction

Murota [12, 13] considered a class of discrete convex functions that have nice min-max relations and can efficiently be minimized. Among them are L-convex (L$^\natural$-convex) func-

[*]Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan. E-mail: fujishig@kurims.kyoto-u.ac.jp. Research supported by a Grant-in-Aid of the Ministry of Education, Culture, Sports and Technology of Japan.

[†]Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan. E-mail: thayashi@kurims.kyoto-u.ac.jp

[‡]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. E-mail: kiyohito_nagano@mist.i.u-tokyo.ac.jp

tions and M-convex (M$^\natural$-convex) functions; the former are the conjugates (or the Legendre transforms) of the latter and vice versa. L$^\natural$-convex functions were also treated earlier by Favati and Tardella [1] and called submodular integrally convex functions (also see [6] for L$^\natural$-convex functions and [15] for M$^\natural$-convex functions). A weakly polynomial algorithm for minimizing L$^\natural$-convex functions was proposed by Murota [14].

We consider the problem of minimizing an L$^\natural$-convex function with a linear inequality constraint having a positive normal vector (or weight vector) such as a budget constraint. The effective domains of L$^\natural$-convex functions have special discrete structures (see [5, 13]), so that such an additional linear constraint does not fit into the effective domains of L$^\natural$-convex functions and we need a new algorithm to the problem other than those for L$^\natural$-convex functions without any additional constraints. Moreover, we examine possible extension and adaptation of the algorithm for other problems and give polynomial algorithms.

The present paper is organized as follows. Section 2 gives definitions and describes the optimization problem for an L$^\natural$-convex function with a linear inequality constraint. Section 3 shows a characterization of optimal solutions of the problem. In Section 4 we propose an algorithm to solve the problem based on a binary search for an optimal Lagrange multiplier, which finds an optimal solution in weakly polynomial time. In Section 5 we also examine an extension of the problem to the one having a not necessarily positive weight vector and adapt the algorithm for an L$^\natural$-convex function to minimizing an M$^\natural$-convex function with a linear inequality constraint. We propose polynomial algorithms for solving these problems.

## 2. Definitions and Problem Description

Let $E$ be a finite nonempty set, and $\mathbf{R}$ and $\mathbf{Z}$ be the sets of reals and of integers, respectively. We consider an L$^\natural$-convex function $f : \mathbf{Z}^E \to \mathbf{R} \cup \{+\infty\}$ on the integer lattice $\mathbf{Z}^E$ such that its effective domain $\mathrm{dom} f \equiv \{x \in \mathbf{Z}^E \mid f(x) < +\infty\}$ is nonempty and full-dimensional. See [13] for the precise definition. We give a characterization of an L$^\natural$-convex function in the following.

Denote by $\mathrm{Conv}$ the convex hull operator in $\mathbf{R}^E$. For any $z \in \mathbf{Z}^E$ and any linear ordering (or permutation) $\sigma$ of $E$ define a simplex

$$\Delta_z^\sigma = \mathrm{Conv}(\{z + \chi_{S_i} \mid i = 1, \cdots, m,\ S_i \text{ is the set of the first } i \text{ elements of } \sigma\}). \quad (2.1)$$

The collection of all such simplices $\Delta_z^\sigma$ for all points $z \in \mathbf{Z}^E$ and linear orderings $\sigma$ of $E$ forms a simplicial division of $\mathbf{R}^E$, which is called the *Freudentahl simplicial division*. We also call each $\Delta_z^\sigma$ a *Freudentahl cell*.

Let $f$ be a function on the integer lattice $\mathbf{Z}^E$ such that

(A1) $\quad \mathrm{Conv}(\mathrm{dom} f) \cap \mathbf{Z}^E = \mathrm{dom} f.$

(A2) The convex hull $\mathrm{Conv}(\mathrm{dom}f)$ of the effective domain of $f$ is full-dimensional and is the union of some Freudentahl cells.

The assumption of the full dimensionality is not essential but we assume it here for simplicity. Under Assumptions (A1) and (A2) we can uniquely construct a piecewise linear extension $\hat{f}$ of $f$ by means of the Freudentahl simplicial division as follows. For any $x \in \Delta_z^\sigma$ we have a unique expression of $x$ as a convex combination of extreme points of the cell $\Delta_z^\sigma$ as

$$x = \sum_{i=1}^{m} \alpha_i (z + \chi_{S_i}), \tag{2.2}$$

where $S_i$ is the set of the first $i$ elements of $\sigma$. According to the expression (2.2) we define

$$\hat{f}(x) = \sum_{i=1}^{m} \alpha_i f(z + \chi_{S_i}). \tag{2.3}$$

For all $x$ outside $\mathrm{Conv}(\mathrm{dom}f)$ we put $\hat{f}(x) = +\infty$. Note that $\hat{f}$ is well defined. Such a piecewise-linear extension has been considered in fixed-point algorithms as a piecewise-linear approximation of a nonlinear function on $\mathbf{R}^E$ being sampled on $\mathbf{Z}^E$ (see [21, 24]). It should also be noted that when $\mathrm{dom}f = \{\chi_X \mid X \subseteq E\}$, $\hat{f}$ is called the *Lovász extension* ([5, 11]).

Then, $f$ is an *L$^\natural$-convex function* if and only if the following (A3) holds.

(A3) The piecewise linear extension $\hat{f} : \mathbf{R}^E \to \mathbf{R} \cup \{+\infty\}$ of $f$ by (2.2) and (2.3) is a convex function on $\mathbf{R}^E$

(see [6, 12, 13]). The original definition of an L$^\natural$-convex function on $\mathbf{Z}^E$ is different, but see [5, Chapter VII] for the proof of their equivalence. Hence Conditions (A1), (A2), and (A3) imply

(S) $f$ is submodular on $\mathrm{dom}f$, i.e.,

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \qquad (x, y \in \mathrm{dom}f), \tag{2.4}$$

where $(x \vee y)(e) = \max\{x(e), y(e)\}$ and $(x \wedge y)(e) = \min\{x(e), y(e)\}$ for $e \in E$.

It follows that the effective domain $\mathrm{dom}f$ is a (distributive) lattice with lattice operations $\vee$ and $\wedge$. For any $z \in \mathbf{Z}^E$ define a set function $f_z$ on $2^E$ by

$$f_z(X) = f(z + \chi_X) \qquad (X \subseteq E). \tag{2.5}$$

Then (2.4) implies that the function $f_z$ is a submodular set function for every $z \in \mathbf{Z}^E$ whenever $\mathrm{dom}f_z \neq \emptyset$.

We also call the extension $\hat{f}$ on $\mathbf{R}^E$ an *L$^\natural$-convex function*. We consider L$^\natural$-convex functions on $\mathbf{Z}^E$ and $\mathbf{R}^E$ and their extensions in the above sense in the sequel.

For any L$^\natural$-convex function $f : \mathbf{Z}^E \to \mathbf{R} \cup \{+\infty\}$, any positive weight vector $w : E \to \mathbf{R}$, and any $\beta \in \mathbf{R}$ consider the following problem.

$$\begin{array}{lll} (P^\circ): & \text{Minimize} & \hat{f}(x) \\ & \text{subject to} & \langle w, x\rangle \le \beta, \\ & & x \in \mathbf{R}^E, \end{array} \tag{2.6}$$

where $\langle w, x\rangle = \sum_{e\in E} w(e)x(e)$. Here it should be noted that the objective function $\hat{f}$ is defined on $\mathbf{R}^E$ and is not restricted on the integer lattice $\mathbf{Z}^E$. This makes the problem substantially easier but the existence of a linear constraint makes it nontrivial.

# 3. The Lagrangian Function and the Optimality

For Problem $(P^\circ)$ given by (2.6) consider the associated Lagrangian function as follows.

$$L(x, \lambda) = \hat{f}(x) + \lambda(\beta - \langle w, x\rangle). \tag{3.1}$$

Then the optimal solutions of Problem $(P^\circ)$ are characterized by the following, which is well-known in (continuous) convex optimization.

**Proposition 3.1**: *A vector $x^* \in \mathbf{R}^E$ is an optimal solution of Problem $(P^\circ)$ if and only if there exists a $\lambda^* \le 0$ such that the following* (a), (b), *and* (c) *hold.*

(a) $x^*$ *is a minimizer of $L(x, \lambda^*)$ in $x$.*

(b) $\langle w, x^*\rangle \le \beta$.

(c) $\lambda^*(\beta - \langle w, x^*\rangle) = 0$.

*Condition* (a) *is rewritten as*

$$\mathbf{0} \in \partial L(x^*, \lambda^*) = \partial \hat{f}(x^*) - \lambda^* w \tag{3.2}$$

*or $\lambda^* w \in \partial \hat{f}(x^*)$, where $\partial L(x^*, \lambda^*)$ (resp. $\partial \hat{f}(x^*)$) denotes the subdifferential of $L(x, \lambda^*)$ (resp. $\hat{f}(x)$) in $x$ at $x = x^*$.* □

It should be noted that Proposition 3.1 holds for any convex function. For the L$^\natural$-convex function $f$ the subdiferentials $\partial L(x^*, \lambda^*)$ and $\partial \hat{f}(x^*)$ are *generalized polymatroids* of Frank [3] because of the L$^\natural$-convexity of $L(x, \lambda^*)$ and $\hat{f}(x)$ (see [5, 13]). This will play a crucial rôle in constructing efficient algorithms. Note that for any $x \in \mathrm{dom} f$ we have

$$\partial \hat{f}(x) = \{p \in \mathbf{R}^E \mid \forall X \subseteq E : f(x) - f(x - \chi_X) \le p(X) \le f(x + \chi_X) - f(x)\} \tag{3.3}$$

(see [5, 13]).

Define $\mathcal{Z}(\lambda)$ to be the set of minimizers of $L(x,\lambda)$ or $f(x) - \lambda\langle w, x\rangle$ in $x \in \mathbf{Z}^E$. We see from its submodularity that $\mathcal{Z}(\lambda)$ is a distributive lattice with respective to $\vee$ and $\wedge$.

**Theorem 3.2**: *Suppose that for a parameter $\lambda^* < 0$ there exist $x, y \in \mathcal{Z}(\lambda^*)$ such that*

$$\langle w, x\rangle \leq \beta, \qquad \langle w, y\rangle \geq \beta. \tag{3.4}$$

*Then a vector $x^*$ lying on the line segment between $x$ and $y$ and satisfying $\langle w, x^*\rangle = \beta$ is an optimal solution of Problem $(P^\circ)$.*

(Proof) For any feasible solution $z$ of Problem $(P^\circ)$,

$$\begin{aligned} \hat{f}(z) &\geq \hat{f}(z) + \lambda^*(\beta - \langle w, z\rangle) \\ &\geq \min\{f(z') + \lambda^*(\beta - \langle w, z'\rangle) \mid z' \in \mathrm{dom} f\} \\ &= \hat{f}(x^*) + \lambda^*(\beta - \langle w, x^*\rangle) \\ &= \hat{f}(x^*), \end{aligned} \tag{3.5}$$

where note that $f(x) - \lambda^*\langle w, x\rangle = f(y) - \lambda^*\langle w, y\rangle = \hat{f}(x^*) - \lambda^*\langle w, x^*\rangle$ since $x, y \in \mathcal{Z}(\lambda^*)$. It follows from Proposition 3.1 and the assumption that $x^*$ is an optimal solution of $(P^\circ)$. $\square$

When $\mathrm{dom} f$ is bounded, we can apply Murota's weakly polynomial algorithm [14] for minimizing L$^\natural$-convex functions to find a vector in $\mathcal{Z}(\lambda)$ for each $\lambda$. We can perform a binary search to find an optimal Lagrange multiplier $\lambda^*$ by making use of algorithms for the minimum-ratio (and maximum-ratio) problems to be described in the next section.

It should also be noted that the structure of parametric minimizers $\mathcal{Z}(\lambda)$ of $L(x,\lambda)$ or $f(x) - \lambda\langle w, x\rangle$ in $x \in \mathbf{Z}^E$ is closely related to the theory of principal partitions (see [5, 7]) and also to the monotonicity results on parametric minimization of submodular functions on $\mathbf{Z}^E$ (see [22, 23]).

# 4. A Polynomial Algorithm

In this section we propose algorithms for solving Problem $(P^\circ)$. In the following we suppose

(i) The effective domain $\mathrm{dom} f$ is bounded and we are given the maximum element $\bar{u}$ and the minimum element $\underline{u}$ of $\mathrm{dom} f$ as a lattice.

(ii) $f$ is integer-valued on $\mathrm{dom} f$, $w$ is a positive integral vector, and $\beta$ is an integer.

It should be noted that although we assume in (i) the boundedness of $\mathrm{dom} f$, the inequality constraint and the existence of the minimum element $\underline{u}$ of $\mathrm{dom} f$ imply that the feasible region of the problem is bounded from above because of the positiveness of the weight vector $w$.

## 4.1. A binary search algorithm

We construct an algorithm for finding an optimal solution of Problem $(P^\circ)$ based on a binary search for an optimal Lagrange multiplier $\lambda^*$.

If we have a minimizer $x^*$ of $\hat{f}$ that satisfies $\langle w, x^* \rangle \le \beta$, then $x^*$ is an optimal solution of $(P^\circ)$. Hence let us suppose $\langle w, x^* \rangle > \beta$. Then there exists an optimal solution lying on the hyperplane $\langle w, x \rangle = \beta$. We try to find such an optimal solution.

Let $x \in \mathrm{dom} f$ satisfy $x \in \mathcal{Z}(\alpha)$ for some $\alpha \in \mathbf{R}$. For such an $x$ define

$$\hat{\lambda}_x = \inf\{\alpha \mid \alpha w \in \partial \hat{f}(x)\}, \tag{4.1}$$

$$\hat{\mu}_x = \sup\{\alpha \mid \alpha w \in \partial \hat{f}(x)\}. \tag{4.2}$$

It follows from (3.3) and the positiveness of $w$ that

$$\hat{\lambda}_x = \inf\{\alpha \mid \forall X \subseteq E : \alpha w(X) \ge f(x) - f(x - \chi_X)\}, \tag{4.3}$$

$$\hat{\mu}_x = \sup\{\alpha \mid \forall X \subseteq E : \alpha w(X) \le f(x + \chi_X) - f(x)\}. \tag{4.4}$$

Since $w$ is a positive vector, the problem of determining the value of (4.4) is reduced to the minimum-ratio problem

$$\text{Minimize} \quad \frac{f(x + \chi_X) - f(x)}{w(X)} \quad \text{subject to} \quad \emptyset \ne X \subseteq E \tag{4.5}$$

(see [5, Sec. 7.2]). The minimum value is equal to $\hat{\mu}_x$. Dually, the maximum-ratio problem

$$\text{Maximize} \quad \frac{f(x) - f(x - \chi_X)}{w(X)} \quad \text{subject to} \quad \emptyset \ne X \subseteq E \tag{4.6}$$

gives the value of $\hat{\lambda}_x$ of (4.3). Note that $f_x(X) = f(x + \chi_X) - f(x)$ in $X$ is a submodular set function and $f(x) - f(x - \chi_X)$ is a supermodular set function in $X$. Strongly polynomial algorithms for the minimum-ratio and maximum-ratio problems are obtained by [2, 17].

**Theorem 4.1**: *Suppose that we are given $x, y \in \mathrm{dom} f$ that define $\hat{\mu}_x, \hat{\lambda}_y \in \mathbf{R}$ such that*

$$\hat{\mu}_x < \hat{\lambda}_y \tag{4.7}$$

*and*

$$\langle w, x \rangle < \beta, \qquad \langle w, y \rangle > \beta. \tag{4.8}$$

*Put $\alpha = (\hat{\mu}_x + \hat{\lambda}_y)/2$ and let $z \in \mathcal{Z}(\alpha)$. Then we have*

$$\langle w, x \rangle < \langle w, z \rangle < \langle w, y \rangle. \tag{4.9}$$

*Moreover,*

$$\hat{\mu}_x \leq \hat{\lambda}_z \leq \hat{\mu}_z \leq \hat{\lambda}_y. \tag{4.10}$$

(Proof) It follows from the definitions of $\hat{\mu}_x$ and $\alpha$ that $\langle w, x\rangle < \langle w, z\rangle$. Similarly, from the definitions of $\hat{\lambda}_y$ and $\alpha$ we get $\langle w, z\rangle < \langle w, y\rangle$. Moreover, (4.10) follows from the definitions of $\alpha$, $\hat{\lambda}_z$, and $\hat{\mu}_z$. □

Based on Theorem 4.1 we propose an algorithm as follows.

**Algorithm** BS

**Input**: The minimum element $\underline{u}$ of $\mathrm{dom} f$.

**Output**: An optimal solution $x^*$ and an optimal Lagrange multiplier $\alpha^*$.

**Step 1**: Compute a global minimizer $x^* \in \mathbf{Z}^E$ of $f$.
If $x^*$ is a feasible solution of $(P^\circ)$, then return $x^*$ and $\alpha^* = 0$.
Else put $x \leftarrow \underline{u}$ and $y \leftarrow x^*$, compute $\hat{\mu}_x$ and $\hat{\lambda}_y$, and go to Step 2.

**Step 2**: While $\hat{\mu}_x < \hat{\lambda}_y$ do (*):

(*) Put $\alpha \leftarrow (\hat{\mu}_x + \hat{\lambda}_y)/2$ and find $z \in \mathcal{Z}(\alpha)$.
If $z$ is feasible, then put $x \leftarrow z$ and compute $\hat{\mu}_x$,
else put $y \leftarrow z$ and compute $\hat{\lambda}_y$.

**Step 3**: Find a point $x^*$ in the intersection of the line segment between $x$ and $y$ and the hyperplane $H = \{z \in \mathbf{R}^E \mid \langle w, z\rangle = \beta\}$. Return $x^*$ and $\alpha^* = \hat{\mu}_x$.

(End)

It should be noted that computing $\hat{\mu}_x$ and $\hat{\lambda}_y$ plays a crucial rôle in achieving the polynomial complexity of the proposed algorithm, which will be seen in the sequel.

## 4.2. Validity of the algorithm

Let $B = \max\{\max\{|f(z)| \mid z \in \mathrm{dom} f\}, \max\{w(e) \mid e \in E\}, |\beta|\}$ and $K = \max\{\bar{u}(e) - \underline{u}(e) \mid e \in E\}$. We assume that $B, K \geq 2$.

**Theorem 4.2**: *Algorithm* BS *computes an optimal solution of* $(P^\circ)$ *in* $\mathrm{O}((\mathbf{L}^\natural + \mathbf{SFM}) \log B)$ *time, where* **SFM** *denotes the complexity of the submodular function minimization algorithm* (*of Orlin* [18]) *and* $\mathbf{L}^\natural$ *denotes the complexity of minimizing an* $L^\natural$*-convex function* (*which is* $\mathrm{O}(\mathbf{SFM} \log K)$ *by Murota's algorithm* [14]). *Consequently, it runs in* $\mathrm{O}(\mathbf{SFM} \log K \log B)$ *time.*

(Proof) Since we have

$$\hat{\mu}_{\underline{u}} = \min\left\{ \frac{f(\underline{u} + \chi_X) - f(\underline{u})}{w(X)} \;\middle|\; \emptyset \neq X \subseteq E \right\} \geq -2B, \tag{4.11}$$

the initial difference $\hat{\lambda}_y - \hat{\mu}_x$ is bounded by $-\min\{0, \hat{\mu}_{\underline{u}}\} \leq 2B$, and any nonzero difference $\hat{\lambda}_y - \hat{\mu}_x$ is not less than $1/(B^2)$. Since the difference $\hat{\lambda}_y - \hat{\mu}_x$ is cut in half in each execution of (*) of Step 2, it follows that the number of the executions of (*) of Step 2 is $\mathrm{O}(\log B)$. Each computation of $\hat{\mu}_x$ and $\hat{\lambda}_y$ requires $\mathrm{O}(\textbf{SFM})$ (see [2, 17]). Also we can apply Murota's algorithm for minimizing L$^\natural$-convex functions to compute a minimizer $z$ of $L(\cdot, \alpha)$ (i.e., $z \in \mathcal{Z}(\alpha)$) in each execution of (*), which requires $\mathrm{O}(\textbf{SFM} \log K)$ time (see [10] for the complexity estimation of Murota's algorithm). It follows that the total running time of Algorithm BS is $\mathrm{O}((\mathbf{L}^\natural + \textbf{SFM}) \log B)$ or $\mathrm{O}(\textbf{SFM} \log K \log B)$. □

In Algorithm BS we can use the submodular function minimization (SFM) algorithms that allow parametric minimization of a strong-map sequence of submodular set functions such as those in [2, 8, 9, 18]. The complexity of currently the best SFM algorithm is $\mathrm{O}(|E|^5(|E| + \mathrm{EO}) \log |E|)$ due to Orlin, where EO denotes the time required for a function evaluation of $f$. The applicability of Orlin's algorithm [18] to the minimum-ratio problem is shown by Nagano [17].

# 5. Extensions and Related Algorithms

In this section we discuss possible extension and adaptation of the algorithm for Problem $(P^\circ)$ to other problems.

## 5.1. General weight vectors

We have assumed that the weight vector $w : E \to \mathbf{R}$ is positive. This assumption leads us to the minimum-ratio (resp. maximum-ratio) problem for determining the value $\hat{\mu}_x$ (resp. $\hat{\lambda}_y$) for which we have efficient algorithms. Although Problem $(P^\circ)$ becomes more difficult, we can solve Problem$(P^\circ)$ with a not necessarily positive but nonzero weight vector $w$ as follows.

Suppose that we are given a vector $z \in \mathcal{Z}(\alpha)$ for some $\alpha \in \mathbf{R}$. For such a vector $z$ we modify the definitions of $\hat{\lambda}_z$ and $\hat{\mu}_z$ as follows.

$$\hat{\lambda}_z = \inf\{\nu \mid \forall X \subseteq E : f(z) - f(z - \chi_X) \leq \nu w(X) \leq f(z + \chi_X) - f(z)\}, \tag{5.1}$$

$$\hat{\mu}_z = \sup\{\nu \mid \forall X \subseteq E : f(z) - f(z - \chi_X) \leq \nu w(X) \leq f(z + \chi_X) - f(z)\}. \tag{5.2}$$

Recall that the polyhedron represented by the system of inequalities in the right-hand side of (5.1) and (5.2) is a generalized polymatroid, which is a projection of a base polyhedron into the coordinate space of codimension one (see [4]). Hence the problem of computing $\hat{\lambda}_z$ and $\hat{\mu}_z$ defined as above is to find the end-points of the intersection of the generalized polymatroid and the line $L = \{\nu w \mid \nu \in \mathbf{R}\}$. We can adapt the strongly polynomial algorithm of Nagano [16] to find such end-points. When a computed global minimizer $x^*$

of $f$ is infeasible, we should assume that we are given a feasible solution $z$ and a negative value $\alpha$ such that $z \in \mathcal{Z}(\alpha)$. Then we can start from Step 2 of Algorithm BS by putting $x \leftarrow z$ and $y \leftarrow x^*$ to solve Problem $(P^\circ)$ with given $w$, using the above definitions of $\hat{\lambda}_z$ and $\hat{\mu}_z$.

## 5.2. M$^\natural$-convex functions

For any L$^\natural$-convex function $f : \mathbf{Z}^E \to \mathbf{Z}$ we have the associated M$^\natural$-convex function $f^\bullet : \mathbf{Z}^E \to \mathbf{Z} \cup \{+\infty\}$ as the convex conjugate or the Legendre transform of $f$, which is given by

$$f^\bullet(p) = \sup\{\langle p, x\rangle - f(x) \mid x \in \mathbf{Z}^E\} \qquad (p \in \mathbf{Z}^E). \tag{5.3}$$

Put $g = f^\bullet$. The convex extension $\hat{g}$ of $g(= f^\bullet)$ is expressed as

$$\hat{g}(p) = \sup\{\langle p, x\rangle - f(x) \mid x \in \mathbf{Z}^E\} \qquad (p \in \mathbf{R}^E). \tag{5.4}$$

We also call $\hat{g} : \mathbf{R}^E \to \mathbf{R} \cup \{+\infty\}$ an M$^\natural$-convex function. (M$^\natural$-convex functions from $\mathbf{R}^E$ to $\mathbf{R} \cup \{+\infty\}$ in a more general sense have been considered in [15, 13] (also see [5]).)

For the M$^\natural$-convex function $\hat{g} : \mathbf{R}^E \to \mathbf{R} \cup \{+\infty\}$, a nonzero integral vector $c : E \to \mathbf{Z}$, and an integer $\gamma$ consider the following problem.

$$\begin{aligned} (P^\bullet) : \text{Minimize} \quad & \hat{g}(p) \\ \text{subject to} \quad & \langle p, c\rangle \le \gamma, \\ & p \in \mathbf{R}^E. \end{aligned} \tag{5.5}$$

For any $\alpha \in \mathbf{R}$ denote by $\mathcal{Z}^\bullet(\alpha)$ the set of all minimizers of $g(p) - \alpha\langle p, c\rangle$ in $p \in \mathbf{Z}^E$.

**Theorem 5.1**: *Suppose that $p \in \mathcal{Z}^\bullet(\alpha)$ for some $\alpha \in \mathbf{R}$. Then,*

$$\begin{aligned} &\inf\{\alpha \mid p \in \mathcal{Z}^\bullet(\alpha)\} \\ &= \max\left\{\max\left\{\frac{\hat{g}(p + \chi_e - \chi_{e'}) - \hat{g}(p)}{c(e) - c(e')} \;\middle|\; e, e' \in E,\ c(e) < c(e')\right\},\right. \\ &\qquad\qquad \left.\max\left\{\frac{\hat{g}(p + \epsilon\chi_e) - \hat{g}(p)}{\epsilon c(e)} \;\middle|\; e \in E,\ \epsilon c(e) < 0,\ \epsilon = +, -\right\}\right\}, \end{aligned} \tag{5.6}$$

$$\begin{aligned} &\sup\{\alpha \mid p \in \mathcal{Z}^\bullet(\alpha)\} \\ &= \min\left\{\min\left\{\frac{\hat{g}(p + \chi_e - \chi_{e'}) - \hat{g}(p)}{c(e) - c(e')} \;\middle|\; e, e' \in E,\ c(e) > c(e')\right\},\right. \\ &\qquad\qquad \left.\min\left\{\frac{\hat{g}(p + \epsilon\chi_e) - \hat{g}(p)}{\epsilon c(e)} \;\middle|\; e \in E,\ \epsilon c(e) > 0,\ \epsilon = +, -\right\}\right\}. \end{aligned} \tag{5.7}$$

(Proof) Because of the conjugacy between L$^\natural$-convex functions and M$^\natural$-convex functions the tangent cone of the epigraph of $\hat{g}$ at $p$ is generated by those vectors from among $(\chi_e-\chi_{e'}, \hat{g}(p+\chi_e-\chi_{e'})-\hat{g}(p))$ $(e, e' \in E)$ and $(\epsilon\chi_e, \hat{g}(p+\epsilon\chi_e)-\hat{g}(p))$ $(e \in E,\ \epsilon = +, -)$ that belong to $\mathbf{R}^E \times \mathbf{R}$. (Note that L$^\natural$-convex functions are affine on every Freudentahl cell and the facets of Freudentahl cells are determined by hyperplanes having normal vectors from among $\chi_e - \chi_{e'}$ $(e, e' \in E)$ and $\epsilon\chi_e$ $(e \in E,\ \epsilon = +, -)$ (see [5, 13].) Hence we have (5.6) and (5.7). □

Let us denote the values of (5.6) and (5.7) by $\hat{l}_p$ and $\hat{m}_p$, respectively. Note that $\hat{l}_p$ and $\hat{m}_p$ can be computed in $\mathrm{O}(|E|^2\mathrm{EO}^\bullet)$ time, where $\mathrm{EO}^\bullet$ means the time for the function evaluation oracle for $g$.

Now, suppose that we are given vectors $p_1, p_2 \in \mathbf{R}^E$ such that

$$p_1 \in \mathcal{Z}^\bullet(\alpha_1), \qquad p_2 \in \mathcal{Z}^\bullet(\alpha_2) \tag{5.8}$$

and

$$\langle p_1, c\rangle \le \gamma \le \langle p_2, c\rangle. \tag{5.9}$$

If we have $\hat{m}_{p_1} \ge \hat{l}_{p_2}$, then a point $p^*$ in the intersection of the line segment between $p_1$ and $p_2$ and the hyperplane $H^\bullet = \{p \in \mathbf{R}^E \mid \langle p, c\rangle = \gamma\}$ is an optimal solution of Problem $(P^\bullet)$.

If $\hat{m}_{p_1} < \hat{l}_{p_2}$, then put $\alpha \leftarrow (\hat{m}_{p_1} + \hat{l}_{p_2})/2$ and find a point $q \in \mathcal{Z}^\bullet(\alpha)$. If $q$ is feasible, then put $p_1 \leftarrow q$ or else $p_2 \leftarrow q$. Repeat this process until $\hat{m}_{p_1} = \hat{l}_{p_2}$ and we obtain an optimal solution $p^*$ as above.

The validity of this algorithm can be proven by the same arguments as made for Algorithm BS.

Starting with vectors $p_1, p_2 \in \mathbf{R}^E$ satisfying (5.8) and (5.9), the above-mentioned algorithm computes an optimal solution of Problem $(P^\bullet)$ in $\mathrm{O}((\mathbf{M}^\natural + |E|^2\mathrm{EO}^\bullet)\log B^\bullet)$, where $\mathbf{M}^\natural$ denotes the complexity of minimizing an M$^\natural$-convex function, which is $\mathrm{O}((|E|^3 + |E|^2\log(K^\bullet/|E|))\mathrm{EO}^\bullet(\log(K^\bullet/|E|)/\log|E|))$ or $\mathrm{O}(|E|^3\mathrm{EO}^\bullet\log(K^\bullet/|E|))$ due to Shioura [19] and Tamura [20], and

$$B^\bullet = \max\{\max\{|g(p)| \mid p \in \mathrm{dom}\, g\},\ \max\{|c(e)| \mid e \in E,\ |\gamma|\}\}, \tag{5.10}$$

$$K^\bullet = \max\{|x(e) - y(e)| \mid x, y \in \mathrm{dom}\, g\}. \tag{5.11}$$

Here we assume that both $B^\bullet$ and $K^\bullet$ are finite values.

# 6. Concluding Remarks

We have proposed polynomial algorithms for minimizing a discrete convex function, such as an L$^\natural$-convex function and an M$^\natural$-convex function, with a linear inequality constraint.

We can adapt the proposed algorithm to minimizing those extensions of discrete convex functions for which the end-points of the intersection of a line and a subdifferential of the function at every point can efficiently be computed.

In the present paper we have considered the minimization problem with a single inequality constraint. The simple binary search approach adopted here does not work for the problem with multiple inequality constraints. Related research on minimization of submodular functions with multiple parameters has been made in the theory of principal partitions [7] (also see [5, Sec. 7.2]). We will investigate the problem elsewhere.

# References


[1] P. Favati and F. Tardella: Convexity in nonlinear integer programming. *Ricerca Operativa* **53** (1990) 3–44.

[2] L. Fleischer and S. Iwata: A push-relabel framework for submodular function minimization and applications to parametric optimization. *Discrete Applied Mathematics* **131** (2003) 311–322.

[3] A. Frank: Generalized polymatroids. *Proceedings of the Sixth Hungarian Combinatorial Colloquium* (Eger, 1981); also, *Finite and Infinite Sets, I* (A. Hajnal, L. Lovász and V. T. Sós, eds., Colloquia Mathematica Societatis János Bolyai **37**, North-Holland, Amsterdam, 1984), pp. 285–294.

[4] S. Fujishige: A note on Frank's generalized polymatroids. *Discrete Applied Mathematics* **7** (1984) 105–109.

[5] S. Fujishige: *Submodular Functions and Optimization* (Second Edition) (Annals of Discrete Mathematics **58**) (Elsevier, 2005).

[6] S. Fujishige and K. Murota: Notes on L-/M-convex functions and the separation theorems. *Mathematical Programming*, Ser. A **88** (2000) 129–146.

[7] M. Iri: Structural theory for the combinatorial systems characterized by submodular functions. In: *Progress in Combinatorial Optimization* (W. R. Pulleyblank, ed., Academic Press, Toronto, 1984), pp. 197–219.

[8] S. Iwata: A faster scaling algorithm for minimizing submodular functions. *SIAM Journal on Computing* **32** (2003) 833–840.

[9] S. Iwata: Submodular function minimization. *Mathematical Programming*, Ser. B **112** (2008) 45–64.

[10] V. Kolmogorov and A. Shioura: New algorithms for the dual of the convex cost network flow problem with application to computer vision. Technical Report, June 4, 2007. http://www.adastral.ucl.ac.uk/˜vladkolm/papers/CONVEX_MRF.html

[11] L. Lovász: Submodular functions and convexity. In: *Mathematical Programming—The State of the Art* (A. Bachem, M. Grötschel and B. Korte, eds., Springer, Berlin, 1983), pp. 235–257.

[12] K. Murota: Discrete convex analysis. *Mathematical Programming* **83** (1998) 313–371.

[13] K. Murota: *Discrete Convex Analysis* (SIAM Monographs on Discrete Mathematics and Applications **10**, SIAM, 2003).

[14] K. Murota: On steepest descent algorithms for discrete convex functions. *SIAM Journal on Optimization* **14** (2003) 699–707.

[15] K. Murota and A. Shioura: M-convex function on generalized polymatroid. *Mathematics of Operations Research* **24** (1999) 95–105.

[16] K. Nagano: A strongly polynomial algorithm for line search in submodular polyhedra. *Discrete Optimization* **4** (2007) 349–359.

[17] K. Nagano: A faster parametric submodular function minimization algorithm and applications. Mathematical Engineering Technical Reports, METR 2007-43, Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, July 2007.

[18] J. B. Orlin: A faster strongly polynomial time algorithm for submodular function minimization. *Proceedings of the 12th IPCO*, LNCS **4513** (2007), pp. 240–251; *Mathematical Programming* (to appear).

[19] A. Shioura: Fast scaling algorithms for M-convex function minimization with application to the resource allocation problem. *Discrete Applied Mathematics* **134** (2004) 303–316.

[20] A. Tamura: Coordinatewise domain scaling algorithm for M-convex function minimization. *Mathematical Programming* **102** (2005) 339–354.

[21] M. J. Todd: *The Computation of Fixed Points and Applications* (Lecture Notes in Economics and Mathematical Systems **124**, Springer, 1976).

[22] D. M. Topkis: Minimizing a submodular function on a lattice. *Operations Research* **26** (1978) 305–321.

[23] D. M. Topkis: *Supermodularity and Complementarity* (Princeton University Press, Princeton, N.J., 1998).

[24] Z. Yang: *Computing Equilibria and Fixed Points* (Kluwer, Boston, MA, 1999).