RIMS Preprint No. 1324, Kyoto University, June 2001

Coordinatewise Domain Scaling Algorithm for M-convex Function Minimization†

Akihisa Tamura

Abstract

The concept of M-convex functions plays a central role in "discrete convex analysis," a unified framework of discrete optimization introduced by Murota. We develop a polynomial time scaling algorithm for M-convex function minimization problem. The algorithm is applicable also to quasi M-convex function minimization.

Akihisa Tamura
Research Institute for Mathematical Sciences,
Kyoto University, Kyoto 606-8502, Japan.
phone: +81-75-753-7236, facsimile: +81-75-753-7272
e-mail: tamura@kurims.kyoto-u.ac.jp.

† This work is supported by a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

1 Introduction

The concept of M-convexity was introduced by Murota [14, 15] in the context of "discrete convex analysis" which is a unified framework of discrete optimization with reference to existing studies on submodular functions [4, 8, 11], generalized polymatroids [5, 6, 8], valuated matroids [2, 3, 16] and convex analysis [20]. M-convex functions have applications in the areas of mathematical economics, engineering and so on [1, 17, 19]. This paper addresses a fundamental problem, namely, M-convex function minimization problem, which is an extension of separable discrete convex function minimization over an integral base polyhedron [7, 9, 10].

Let V be a nonempty finite set of cardinality n, and $\mathbf{R}$, $\mathbf{Z}$ and $\mathbf{Z}_{++}$ be the sets of reals, integers and positive integers, respectively. Given a vector $z = (z(v) : v \in V) \in \mathbf{Z}^V$, its positive support and negative support are defined by

$$\mathrm{supp}^+(z) = \{v \in V \mid z(v) > 0\} \quad \text{and} \quad \mathrm{supp}^-(z) = \{v \in V \mid z(v) < 0\}.$$

The characteristic vector (unit vector) of $v \in V$ is denoted by χ_v. For a function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$, the *effective domain* of f is defined by

$$\mathrm{dom}\, f = \{x \in \mathbf{Z}^V \mid f(x) < +\infty\}.$$

A function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, f \neq \emptyset$ is called *M-convex* [14, 15] if it satisfies

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

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

Several algorithms for minimizing an M-convex function f are known. Since a locally minimal solution of f is globally minimal, a descent algorithm finds an optimal solution, but it does not terminate in polynomial time. Shioura [21] showed a key property which guarantees that any nonoptimal solution in the effective domain can be efficiently separated from a minimizer, and exploited this property to develop the domain reduction algorithm, the first polynomial time algorithm for M-convex function minimization. Time complexity of the domain reduction algorithm is $O(Fn^4(\log L)^2)$, where F is an upper bound of the time to evaluate f and $L = \max\{|x(u) - y(u)| \mid u \in V,\ x, y \in \mathrm{dom}\, f\}$. Moriguchi, Murota and Shioura [12, 13] showed the proximity property and proposed a scaling technique. They showed that the technique gives an $O(Fn^3 \log(L/n))$-time algorithm for a subclass of M-convex functions f such that "scaled" function

$$f_\alpha(x) = f(x^0 + \alpha x) \qquad (x \in \mathbf{Z}^V)$$

is also M-convex for any $x^0 \in \operatorname{dom} f$ and $\alpha \in \mathbf{Z}_{++}$. Although their algorithm, being a descent algorithm, works for any M-convex function, it does not terminate in polynomial time in general. This is because M-convexity is not inherited in the scaling operation, that is, f_α is not necessarily M-convex.

We develop an $O(Fn^3 \log(L/n))$-time scaling algorithm applicable to any M-convex function, and also show that the algorithm works for a more general optimization problem involving quasi M-convex functions.

Section 2 explains fundamental properties of an M-convex function. Section 3 describes our scaling algorithm and shows its time complexity. Section 4 discusses its extension to quasi M-convex functions.

2 Properties of an M-convex Function

Here we review known results which are utilized in the next section.

Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function and $\arg\min f$ denote the set of all minimizers of f, that is,

$$\arg\min f = \{x \in \mathbf{Z}^V \mid f(x) \leq f(y)\ (\forall y \in \mathbf{Z}^V)\}.$$

By (M-EXC), $\operatorname{dom} f$ lies on a hyperplane whose normal vector is the vector of all ones, that is, for any $x, y \in \operatorname{dom} f$, $x(V) = y(V)$ holds, where $x(V) = \sum_{v \in V} x(v)$. It is also known that $\arg\min f$ and $\operatorname{dom} f$ for an M-convex function f are an M-convex set, where a nonempty set $B \subseteq \mathbf{Z}^V$ is said to be an *M-convex set* if it satisfies

(B-EXC) $\forall x, y \in B$, $\forall u \in \operatorname{supp}^+(x - y)$, $\exists v \in \operatorname{supp}^-(x - y)$:

$$x - \chi_u + \chi_v \in B, \quad y + \chi_u - \chi_v \in B.$$

An M-convex set is a synonym of the set of integer points of the base polyhedron of an integral submodular system (see [8] for submodular systems).

Lemma 2.1 ([14, 15]) *For an M-convex function f, $\operatorname{dom} f$ and $\arg\min f$ are M-convex sets.*

The following property is a direct consequence of Lemma 2.1.

Proposition 2.2 *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function. For $u, v \in V$ and $\gamma_1, \gamma_2 \in \mathbf{Z}$, suppose that there exist $x_1, x_2 \in \arg\min f$ with $x_1(u) \leq \gamma_1$ and $x_2(v) \geq \gamma_2$. Then, there exists $x_* \in \arg\min f$ with $x_*(u) \leq \gamma_1$ and $x_*(v) \geq \gamma_2$.*

M-convex functions have various features of convex functions. For example, a locally minimal solution of an M-convex function is globally minimal.

Theorem 2.3 ([14, 15]) *For an M-convex function f and $x \in \operatorname{dom} f$, $f(x) \leq f(y)$ for any $y \in \mathbf{Z}^V$ if and only if $f(x) \leq f(x - \chi_u + \chi_v)$ for all $u, v \in V$.*

The following properties, Theorems 2.4 and 2.5, play important roles in our algorithm. The former is due to Shioura [21] and the latter is implicit in [12, 13].

Theorem 2.4 ([21]) *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function with $\arg\min f \neq \emptyset$. For $x \in \operatorname{dom} f$ and $v \in V$, the following statements hold.*

(a) *If $f(x - \chi_v + \chi_u) = \min_{s \in V} f(x - \chi_v + \chi_s)$, then there exists $x_* \in \arg\min f$ such that $x(u) - \chi_v(u) + 1 \leq x_*(u)$.*

(b) *If $f(x + \chi_v - \chi_u) = \min_{s \in V} f(x + \chi_v - \chi_s)$, then there exists $x_* \in \arg\min f$ such that $x_*(u) \leq x(u) + \chi_v(u) - 1$.*

Theorem 2.5 *Let f be an M-convex function with $\arg\min f \neq \emptyset$. For $x \in \operatorname{dom} f$, $v \in V$ and $\alpha \in \mathbf{Z}_{++}$, the following statements hold.*

(a) *If $f(x) = \min_{s \in V} f(x - \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x(v) - (n-1)(\alpha - 1) \leq x_*(v)$.*

(b) *If $f(x) = \min_{s \in V} f(x + \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x_*(v) \leq x(v) + (n-1)(\alpha - 1)$.*

Here we show a new property, Theorem 2.6 below, as a common generalization of Theorems 2.4 and 2.5. A special case of Theorem 2.6 with $\alpha = 1$ is equivalent to Theorem 2.4. Another special case of Theorem 2.6 with $u = v$ is identical with Theorem 2.5.

Theorem 2.6 *Let f be an M-convex function with $\arg\min f \neq \emptyset$. For $x \in \operatorname{dom} f$, $v \in V$ and $\alpha \in \mathbf{Z}_{++}$, the following statements hold.*

(a) *If $f(x - \alpha(\chi_v - \chi_u)) = \min_{s \in V} f(x - \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x(u) - \alpha(\chi_v(u) - 1) - (n-1)(\alpha - 1) \leq x_*(u)$.*

(b) *If $f(x + \alpha(\chi_v - \chi_u)) = \min_{s \in V} f(x + \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x_*(u) \leq x(u) + \alpha(\chi_v(u) - 1) + (n-1)(\alpha - 1)$.*

This theorem is proven in Section 4 as a special case of Theorem 4.4.

3 Proposed Algorithm

In this section, we describe a scaling algorithm of time complexity $O(Fn^3 \log(L/n))$ for the M-convex function minimization. It is assumed that the effective domain of a given M-convex function f is bounded and that a vector $x^0 \in \mathrm{dom}\, f$ is given.

We preliminarily show that $L = \max\{|x(u) - y(u)| \mid u \in V,\ x, y \in \mathrm{dom}\, f\}$ can be computed in $O(Fn^2 \log L)$ time. For $x \in \mathrm{dom}\, f$ and $u, v \in V$, the exchange capacity associated with x, u and v is defined as

$$\tilde{c}_f(x, v, u) = \max\{\alpha \mid x + \alpha(\chi_v - \chi_u) \in \mathrm{dom}\, f\},$$

which can be computed in $O(F \log L)$ time by the binary search because $0 \leq \tilde{c}_f(x, v, u) \leq L$. For each $w \in V$, define

$$l_f(w) = \min\{x(w) \mid x \in \mathrm{dom}\, f\}, \quad u_f(w) = \max\{x(w) \mid x \in \mathrm{dom}\, f\}.$$

The values $l_f(w)$ and $u_f(w)$ can be calculated by the following algorithm in $O(Fn \log L)$ time.

function CALCULATE_BOUND(f, x, w)
input: f : M-convex function, $x \in \mathrm{dom}\, f$, $w \in V$;
output: $(l_f(w), u_f(w))$;

$\ell 1$: number $V \setminus \{w\}$ from v_2 to v_n ;
$\ell 2$: $y := z := x$;
$\ell 3$: **for** $i := 2$ **to** n **do** { $y := y + \tilde{c}(y, v_i, w)$, $z := z + \tilde{c}(z, w, v_i)$ } ;
$\ell 4$: **return** $(y(w), z(w))$.

The correctness of the above algorithm can be verified easily from the fact that $\mathrm{dom}\, f$ satisfies (B-EXC) (see also [8, 21]).

Lemma 3.1 *Values $l_f(w)$ and $u_f(w)$ for a fixed $w \in V$ can be computed in $O(Fn \log L)$ time, and L in $O(Fn^2 \log L)$ time.*

For any two vectors $a, b \in \mathbf{Z}^V$, let $[a, b]$ denote the set $\{x \in \mathbf{Z}^V \mid a \leq x \leq b\}$ and f_a^b be defined by

$$f_a^b(x) = \begin{cases} f(x) & \text{if } x \in [a, b] \\ +\infty & \text{otherwise.} \end{cases}$$

Condition (M-EXC) directly yields the next property.

Proposition 3.2 *For an M-convex function f and $a, b \in \mathbf{Z}^V$, if $\mathrm{dom}\, f_a^b \neq \emptyset$ then f_a^b is also M-convex.*

We go on to the main topic of describing our algorithm for minimizing f. The novel idea of the algorithm is to use an individual scaling factor $\alpha_v \in \mathbf{Z}_{++}$ for each $v \in V$. Besides the factors α_v ($v \in V$), the algorithm maintains a current vector $x \in \operatorname{dom} f$, two vectors $a, b \in \mathbf{Z}^V$ and a subset $V' \subseteq V$, and it preserves the following four conditions:

(c1) $x \in \operatorname{dom} f \cap [a, b] = \operatorname{dom} f_a^b$,

(c2) $b(v) - a(v) \leq 2n\alpha_v$ for $v \in V$,

(c3) $\arg\min f \cap [a, b] \neq \emptyset$, (i.e., $\arg\min f_a^b \subseteq \arg\min f$),

(c4) there exists $x_* \in \arg\min f$ such that $x_*(v) = x(v)$ for all $v \in V \setminus V'$.

These parameters are initially put as $x := x^0$, $V' := V$, $a(v) := l_f(v)$, $b(v) := u_f(v)$ and $\alpha_v := 2^{\lceil \log_2 \{(u_f(v) - l_f(v))/n\} \rceil}/2$ for $v \in V$. Thus, conditions (c1) to (c4) are initially satisfied. At each iteration of the algorithm, interval $[a, b]$ is strictly reduced and V' is not increased. By (c4) and the fact that $y(V)$ is a constant for any $y \in \operatorname{dom} f$, the algorithm stops if $|V'| \leq 1$; then the current x is a minimizer of f. The algorithm terminates in $O(n^2 \log(L/n))$ iterations and requires $O(Fn)$ time at each iteration. Hence, the total time complexity of the algorithm is $O(Fn^3 \log(L/n))$.

Before entering into a precise description of the algorithm, we briefly explain its typical behaviour. First, take $v \in V'$ arbitrarily, and find $u \in V'$ minimizing $f_a^b(x - \chi_u + \chi_v)$. Here we explain what the algorithm does in the case where $u \neq v$. Then, there exists $x_1 \in \arg\min f_a^b$ with $x_1(u) \leq x(u) - 1$ by (b) of Theorem 2.4. This inequality suggests that an optimal solution can be found by decreasing $x(u)$. Next find $w \in V'$ minimizing $f_a^b(x - \alpha_u(\chi_u - \chi_w))$. When u can be chosen as w attaining the minimum value, (a) of Theorem 2.5 guarantees that there exists $x_2 \in \arg\min f_a^b$ with $x(u) - (n-1)(\alpha_u - 1) \leq x_2(u)$. By Proposition 2.2, there exists $x_* \in \arg\min f_a^b$ with $x(u) - (n-1)(\alpha_u - 1) \leq x_*(u) \leq x(u) - 1 < x(u)$. Thus, we can put $a(u) := \max[a(u), x(u) - (n-1)(\alpha_u - 1)]$ and $b(u) := x(u)$. Since $b(u) - a(u) \leq n\alpha_u$, the scaling factor α_u can be divided by 2 without violating (c2). In the other case with $w \neq u$, we update a, b and x as $b(u) := x(u) - 1$, $x := x - \alpha_u(\chi_u - \chi_w)$ and $a(w) := \max[a(w), x(w) - (n-1)(\alpha_u - 1)]$, where the update of a is justified by (a) of Theorem 2.6. This is a part of our algorithm described below (see CASE2).

A complete description of our algorithm is now given.

algorithm COORDINATEWISE_SCALING(f, V, x^0)
input: f : M-convex function with bounded $\operatorname{dom} f \subset \mathbf{Z}^V$, $x^0 \in \operatorname{dom} f$;
output: a minimizer of f ;

$\ell 01$: $n := |V|, \quad V' := V, \quad x := x^0$;
$\ell 02$: **for each** $v \in V$ **do** $(a(v), b(v)) :=$ CALCULATE_BOUND(f, x, v) ;
$\ell 03$: **for each** $v \in V$ **do** $\alpha_v := 2^{\lceil \log_2 \{(b(v)-a(v))/n\} \rceil}/2$;
$\ell 04$: **while** $|V'| \geq 2$ **do** {
$\ell 05$: take $v \in V'$;
$\ell 06$: find $u_1 \in V' : f_a^b(x + \chi_v - \chi_{u_1}) = \min_{s \in V' \setminus \{v\}} f_a^b(x + \chi_v - \chi_s)$;
$\ell 07$: find $u_2 \in V' : f_a^b(x - \chi_v + \chi_{u_2}) = \min_{s \in V' \setminus \{v\}} f_a^b(x - \chi_v + \chi_s)$;
$\ell 08$: **if** $f_a^b(x) \leq f_a^b(x + \chi_v - \chi_{u_1})$ **and** $f_a^b(x) \leq f_a^b(x - \chi_v + \chi_{u_2})$ **then**
$\ell 09$: $\{ a(v) := b(v) := x(v), \quad V' := V' \setminus \{v\} \}$;
$\ell 10$: **else if** $f_a^b(x) > f_a^b(x + \chi_v - \chi_{u_1})$ **then** CASE2(u_1) ;
$\ell 11$: **else** $(f_a^b(x) > f_a^b(x - \chi_v + \chi_{u_2}))$ **then** CASE3(u_2) ;
$\ell 12$: } ;
$\ell 13$: **return** (x) ;

function CASE2(u) $\quad (\exists x_1 \in \arg\min f$ with $x_1(u) \leq x(u) - 1)$
$\ell 1$: $W_u := \arg\min_{s \in V'} f_a^b(x - \alpha_u(\chi_u - \chi_s))$;
$\ell 2$: **if** $u \in W_u$ **then**
$\ell 3$: $\{ a(u) := \max[a(u), x(u) - (n-1)(\alpha_u - 1)], \quad b(u) := x(u) \}$;
$\ell 4$: **else** $(u \notin W_u)$ {
$\ell 5$: take $w \in W_u, \quad b(u) := x(u) - 1, \quad x := x - \alpha_u(\chi_u - \chi_w)$;
$\ell 6$: $a(w) := \max[a(w), x(w) - (n-1)(\alpha_u - 1)]$;
$\ell 7$: UPDATE_FACTOR(w) } ;
$\ell 8$: UPDATE_FACTOR(u) ;
$\ell 9$: **return** ;

function CASE3(u) $\quad (\exists x_2 \in \arg\min f$ with $x_2(u) \geq x(u) + 1)$
$\ell 1$: $W_u := \arg\min_{s \in V'} f_a^b(x + \alpha_u(\chi_u - \chi_s))$;
$\ell 2$: **if** $u \in W_u$ **then**
$\ell 3$: $\{ a(u) := x(u), \quad b(u) := \min[b(u), x(u) + (n-1)(\alpha_u - 1)] \}$;
$\ell 4$: **else** $(u \notin W_u)$ {
$\ell 5$: take $w \in W_u, \quad a(u) := x(u) + 1, \quad x := x + \alpha_u(\chi_u - \chi_w)$;
$\ell 6$: $b(w) := \min[b(w), x(w) + (n-1)(\alpha_u - 1)]$;
$\ell 7$: UPDATE_FACTOR(w) } ;
$\ell 8$: UPDATE_FACTOR(u) ;
$\ell 9$: **return** ;

function UPDATE_FACTOR(s)
$\ell 1$: **while** $\alpha_s > 1$ **and** $b(s) - a(s) \leq n\alpha_s$ **do** $\alpha_s := \alpha_s/2$;
$\ell 2$: **if** $a(s) = b(s)$ **then** $V' := V' \setminus \{s\}$;
$\ell 3$: **return** .

In the sequel of this section, we show the correctness and the time complexity of algorithm COORDINATEWISE_SCALING.

Lemma 3.3 COORDINATEWISE_SCALING *preserves conditions* (c1) *to* (c4).

Proof. As we mentioned above, conditions (c1) to (c4) are satisfied just after the execution of line $\ell 03$. We note that Proposition 3.2 guarantees that f_a^b is M-convex. The while loop, which is lines $\ell 04$ to $\ell 12$, consists of three cases.

The first case at line $\ell 08$ implies that there exists $x_* \in \arg\min f$ with $x_*(v) = x(v)$ by condition (c3), Proposition 2.2 and Theorem 2.4 for f_a^b. Trivially, conditions (c1) to (c4) are satisfied after the execution of line $\ell 09$.

We next consider the second case at line $\ell 10$. By (c3) and (b) of Theorem 2.4 for f_a^b, there exists $x_1 \in \arg\min f_a^b \subseteq \arg\min f$ with $x_1(u_1) \leq x(u_1) - 1$. Let us consider function CASE2, in which we have either $u \in W_u$ or $u \notin W_u$. Assume first that $u \in W_u$. By (c3) and (a) of Theorem 2.5 for f_a^b, there exists $x_2 \in \arg\min f$ with $a(u) \leq x_2(u)$, where a is the updated one. We note that the updated a and b have $a(u) \leq b(u)$. By Proposition 2.2, there exists $x_* \in \arg\min f$ with $a(u) \leq x_*(u) \leq b(u)$. Thus, (c1) to (c4) are satisfied. Assume next that $u \notin W_u$. By (c3) and (a) of Theorem 2.6 for f_a^b, there exists $x_2 \in \arg\min f$ with $a(w) \leq x_2(w)$, where a is the updated one. By Proposition 2.2, there exists $x_* \in \arg\min f$ with $a(w) \leq x_*(w)$ and $x_*(u) \leq b(u)$ for the updated a and b, and hence, (c3) holds. Obviously, the updated a, b and x violate no other conditions at the end of line $\ell 6$. At lines $\ell 7$ and $\ell 8$, UPDATE_FACTOR reduces α_w and α_u respectively, preserving (c1) to (c4).

Similarly, the third case at $\ell 11$ also preserves (c1) to (c4). ∎

Lemma 3.4 *The while loop in lines* $\ell 04$ *to* $\ell 12$ *of* COORDINATEWISE_SCALING *terminates in* $O(n^2 \log(L/n))$ *iterations.*

Proof. We divide the iterations of the while loop into three cases.

CLAIM A: The case at line $\ell 08$ occurs at most $n - 1$ times.
[Proof of Claim A] Every time this case occurs, the set V' is reduced by one element. Thus, this case occurs at most $n - 1$ times.

CLAIM B: The case of $u \in W_u$ in CASE2 or CASE3 occurs $O(n \log(L/n))$ times.
[Proof of Claim B] In this case, α_u is greater than 1 from the hypothesis at line $\ell 10$ or $\ell 11$ of COORDINATEWISE_SCALING. Hence, α_u must be updated at line $\ell 8$ of CASE2 or CASE3 at least once because $b(u) - a(u) < n\alpha_u$ for the updated a and b. Thus, this case occurs at most $\lceil \log(L/n) \rceil$ times for a fixed $u \in V$.

CLAIM C: The case of $u \notin W_u$ in CASE2 or CASE3 occurs $O(n^2 \log(L/n))$ times.

By Claims A, B and C, the while loop terminates in $O(n^2 \log(L/n))$ iterations. In the following, we prove Claim C.

[Proof of Claim C] We assume that for each $s \in V$, the s-axis is marked by "$\triangle$" at α_s intervals so that $x(s)$ lies on a mark (see Figure 1). To prove Claim C, we adopt two auxiliary vectors $a', b' \in \mathbf{Z}^V$ such that

$$b'(s) - x(s) \text{ and } x(s) - a'(s) \text{ are divisible by } \alpha_s, \tag{1}$$

$$b'(s) - a'(s) + 2\alpha_s \geq \alpha_s \lfloor (b(s) - a(s))/\alpha_s \rfloor, \tag{2}$$

$$a'(s) - 2\alpha_s < a(s) \leq a'(s) \leq x(s) \leq b'(s) \leq b(s) < b'(s) + 2\alpha_s \tag{3}$$

for any $s \in V$. For instance, the situation of Figure 1 satisfies conditions (1) to (3). We evaluate the number of the occurrences of the case in Claim C by using an integral number ψ defined by

$$\psi = \sum_{s \in V} \frac{b'(s) - a'(s)}{\alpha_s}. \tag{4}$$

By (c2) and (3), $(b'(s) - a'(s))/\alpha_s \leq 2n$ holds. We will show that ψ does not increase if no scaling factors are updated, and that either ψ or a certain scaling factor is strictly reduced except in one case, which does not occur consecutively in a sense.

We set $a'(s)$ and $b'(s)$ to be

$$\begin{aligned} a'(s) &:= x(s) - \alpha_s \lfloor (x(s) - a(s))/\alpha_s \rfloor, \\ b'(s) &:= x(s) + \alpha_s \lfloor (b(s) - x(s))/\alpha_s \rfloor, \end{aligned} \tag{5}$$

just after an update of α_s. Obviously, these $a'(s)$, $b'(s)$ satisfy conditions (1) to (3). We always put $a'(s) := a(s)$ and $b'(s) := b(s)$ when $\alpha_s = 1$, i.e.,

$$\alpha_s = 1 \implies a'(s) = a(s) \text{ and } b'(s) = b(s). \tag{6}$$

For a fixed element $s \in V$, we describe how to modify $a'(s)$ and $b'(s)$ in Case2. The modification in Case3 is similar. Values $a'(s)$ and $b'(s)$ are modified when s is selected as either u or w in functions Case2 or Case3. In the sequel, we assume that updated values are denoted by $a(s)$, $b(s)$, $a'(s)$, $b'(s)$ and $x(s)$, and that old valued are denoted by (old) $a(s)$, (old) $b(s)$, etc.

We first suppose that s is selected as u (e.g., see the modification from Figure 1 to Figure 2). We put $b'(s) := x(s)$ and do not change $a'(s)$. Obviously, (1) is preserved and

$$a'(s) - 2\alpha_s < a(s) \tag{7}$$

$$b(s) - b'(s) = \alpha_s - 1 \tag{8}$$

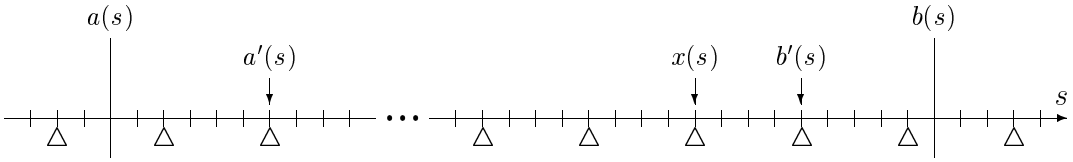

Figure 1: $\alpha_s = 4$.

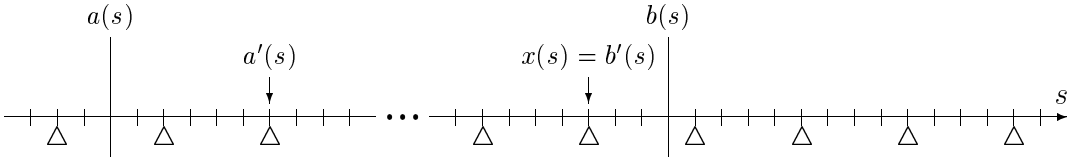

Figure 2: $s = u$ and $\alpha_s = 4$.

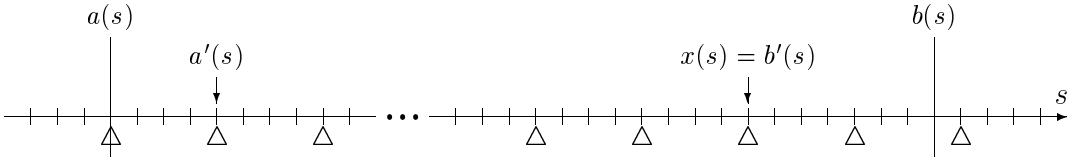

Figure 3: $s = w$, $\alpha_s = 4$ and $\alpha_u = 2$.

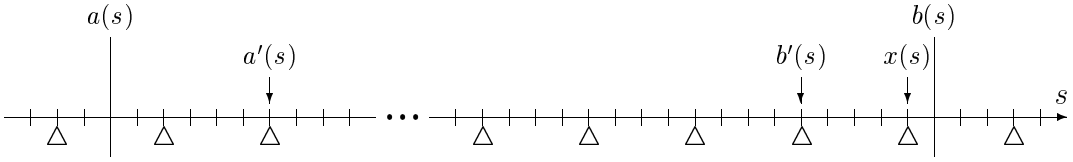

Figure 4: $s = w$, $\alpha_s = 4$ and $\alpha_u = 8$ (inadmissible case).

hold. By (7) and (8), we have

$$b'(s) - a'(s) + 2\alpha_s > b(s) - a(s) - \alpha_s + 1 = (q-1)\alpha_s + r + 1,$$

where we assume that $q = \lfloor (b(s) - a(s))/\alpha_s \rfloor$ and $r = (b(s) - a(s)) - q\alpha_s$. Since $b'(s) - a'(s) + 2\alpha_s$ is divisible by α_s, $a'(s)$ and $b'(s)$ satisfy (2). All inequalities in (3) other than $a'(s) \leq x(s)$ are trivially satisfied. If $x(s) < a'(s)$ then $b(s) - a(s) < 2\alpha_s$ holds by (7), (8) and the equation $x(s) = b'(s) = a'(s) - \alpha_s$, and hence, α_s must be updated at least once at line $\ell 8$ (and $a'(s)$ and $b'(s)$ are updated by (5)); otherwise (3) is satisfied. Moreover, the integral number $(b'(s) - a'(s))/\alpha_s$ is reduced at least by one when α_s remains the same, because of $b'(s) + \alpha_s = [(\text{old})\ x(s)] \leq [(\text{old})\ b'(s)]$. We note that the number of marks in $[a(s), b(s)] \setminus [a'(s), b'(s)]$ is at most one after this modification.

We next suppose that s is selected as w and $\alpha_u \leq \alpha_w$ (e.g., see the modification from Figure 1 to Figure 3). We modify $a'(s)$ and $b'(s)$ as follows: $a'(s)$ is set to

be as large as possible in such a way that $a'(s) - 2\alpha_s < a(s) \leq a'(s) \leq x(s)$ and $x(s) - a'(s)$ is divisible by α_s; $b'(s)$ is next set to be as large as possible so that $b'(s) \leq b(s)$, $b'(s) - x(s)$ is divisible by α_s, and $b'(s) - a'(s)$ does not increase. If $b(s) < b'(s) + \alpha_s$, then we have $b(s) - a(s) < b'(s) - a'(s) + 3\alpha_s$ by $a'(s) - 2\alpha_s < a(s)$; otherwise, $b'(s) - a'(s) =$ (old) $(b'(s) - a'(s))$ must hold by the definition of $b'(s)$. In either case, $a'(s)$ and $b'(s)$ satisfy (2). Suppose to the contrary that $b'(s) + 2\alpha_s \leq b(s)$ holds. This inequality and (2) imply $a'(s) - \alpha_s < a(s)$. By $a'(s) - \alpha_s < a(s)$ and the definition of $a'(s)$, we have $a'(s) = x(s)$. Since (old) $a'(s) \leq$ (old) $x(s) < x(s) = a'(s)$ and $b'(s) - a'(s) =$ (old) $(b'(s) - a'(s))$, we have (old) $b'(s) < b'(s) \leq b(s) - 2\alpha_s$. This contradicts the fact that (old) $a'(s)$ and (old) $b'(s)$ satisfy (3). Thus, $a'(s)$ and $b'(s)$ satisfy (1), (2) and inequalities in (3) except for $x(s) \leq b'(s)$, and $b'(s) - a'(s)$ does not increase. If $a'(s)$ and $b'(s)$ can satisfy an additional requirement $x(s) \leq b'(s)$, then (3) can be satisfied. In the other case, we have $\alpha_s > 1$ and $b'(s) + \alpha_s = x(s)$ because (1), (6) and $b'(s) < x(s) \leq b(s) < b'(s) + 2\alpha_s$ hold. Inequalities $x(s) - (n-1)(\alpha_u - 1) \leq a(s)$, $b(s) < b'(s) + 2\alpha_s$, $b'(s) + \alpha_s = x(s)$ and $\alpha_u \leq \alpha_s$ yield $b(s) - a(s) < n\alpha_s$, and hence, α_s must be updated at least once ($a'(s)$ and $b'(s)$ are updated by (5)).

We finally suppose that s is selected as w and $\alpha_u > \alpha_w$ (e.g., see the modification from Figure 1 to Figure 4). Since α_u is divisible by α_w, marks on the s-axis remain the same. In the same way as the case where $s = w$ and $\alpha_u > \alpha_w$, we can modify $a'(s)$ and $b'(s)$ so that $b'(s) - a'(s)$ does not increase and that (1), (2) and inequalities in (3) except for $x(s) \leq b'(s)$ hold. If $a'(s)$ and $b'(s)$ can be chosen to satisfy an additional requirement $x(s) \leq b'(s)$, then (3) can be satisfied. Here we call the other case an *inadmissible case*. In the inadmissible case, we have $x(s) = b'(s) + \alpha_s \leq b(s)$ because $b'(s) < x(s) \leq b(s) < b'(s) + 2\alpha_s$ and $b'(s) - x(s)$ is divisible by α_s. Moreover, we have $a(s) \leq a'(s) - \alpha_s$, since $\alpha_u > \alpha_s$ holds. That is, the number of marks in $[a(s), b(s)] \setminus [a'(s), b'(s)]$ is precisely two in the inadmissible case. To satisfy (3), it suffices to increase $b'(s)$ by α_s. The number of marks in $[a(s), b(s)] \setminus [a'(s), b'(s)]$ is reduced to one after this modification.

By the above discussion, the integral number ψ defined in (4) does not increase if no scaling factors are updated. In particular, either ψ or a certain scaling factor is strictly reduced except in the inadmissible case. We now attach label w to each iteration. We claim that, for a fixed $s \in V$, the inadmissible case with the same scaling factor α_s does not occur consecutively in the subsequence of iterations labeled s. Hence, the case in Claim C occurs $O(n^2 \log(L/n))$ times.

We prove the last claim. Suppose, to the contrary, that the inadmissible case with the same scaling factor α_s occurs at iterations i and j $(i < j)$ labeled s and that s is not selected as w at iterations $i+1$ to $j-1$. Although s may be selected as u at iterations $i+1$ to $j-1$ in CASE2 or CASE3, marks on the s-axis remain

the same. After the modification at iteration i, $[a(s), b(s)] \setminus [a'(s), b'(s)]$ contains at most one mark and this must be preserved just before iteration j. However, this contradicts the fact that the inadmissible case occurs at iteration j. ∎

By Lemmas 3.3 and 3.4, our main result is obtained.

Theorem 3.5 *Suppose that $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is an M-convex function with a bounded effective domain and that $x^0 \in \operatorname{dom} f$ is given. Algorithm* COORDINATEWISE_SCALING *finds a minimizer of f in $O(Fn^3 \log(L/n))$ time, where F is an upper bound of time to evaluate f and $L = \max\{|x(u) - y(u)| \mid u \in V,\ x, y \in \operatorname{dom} f\}$.*

Proof. By Lemma 3.3, if COORDINATEWISE_SCALING terminates then it finds a minimizer of f. Lines ℓ01 to ℓ03 terminate in $O(Fn^2 \log L)$ time from Lemma 3.1. Lemma 3.4 says that the while loop of COORDINATEWISE_SCALING terminates in $O(n^2 \log(L/n))$ iterations. Since the value $f_a^b(x)$ is evaluated in $O(F)$ time for any $x \in \mathbf{Z}^V$ and function f_a^b is referred to at most $3n$ times per iteration, the total time complexity is $O(Fn^3 \log(L/n))$. ∎

Remark 3.6 For the sake of simplicity, both u_1 and u_2 are computed before line ℓ08, and hence, f_a^b is evaluated $3|V'|$ times in each iteration. We can easily reduce the number of evaluations of f_a^b from $3|V'|$ to $2|V'|$ per iteration. We can also replace constant n with variable $|V'|$. These revision may improve execution time of our algorithm.

4 Extension to Quasi M-convex Functions

Murota and Shioura [18] introduced the concept of quasi M-convex functions and generalized results on M-convexity to quasi M-convexity. In particular, they showed that Shioura's domain reduction algorithm can be extended to the class of functions $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ such that $\operatorname{dom} f$ is a bounded M-convex set and f satisfies

(SSQM$^{\neq}$) $\forall x, y \in \operatorname{dom} f$ with $f(x) \neq f(y)$, $\forall u \in \operatorname{supp}^+(x - y)$, $\exists v \in \operatorname{supp}^-(x - y)$:

$$\begin{aligned} &\text{(i)} \quad f(x - \chi_u + \chi_v) - f(x) \geq 0 \ \Rightarrow\ f(y + \chi_u - \chi_v) - f(y) \leq 0, \quad \text{and} \\ &\text{(ii)} \quad f(y + \chi_u - \chi_v) - f(y) \geq 0 \ \Rightarrow\ f(x - \chi_u + \chi_v) - f(x) \leq 0. \end{aligned}$$

Obviously, an M-convex function satisfies (SSQM$^{\neq}$). This section shows that our algorithm works also for quasi M-convex functions.

Let $\mathcal{C}$ be the class of functions f such that f satisfies (SSQM$^{\neq}$) and $\arg\min f$ is an M-convex set. It is noted that the class $\mathcal{C}$ includes the semistrictly quasi M-convex functions due to [18], which are defined by the property, labeled (SSQM) in [18], obtained from (SSQM$^{\neq}$) by eliminating hypothesis $f(x) \neq f(y)$.

The key properties of M-convex functions stated in Propositions 2.2, 3.2 and Theorems 2.3 and 2.6 are shared by quasi M-convex functions as follows. Since the set of the minimizers of any function in $\mathcal{C}$ is an M-convex set, Proposition 2.2 is trivially extended to $\mathcal{C}$. For any $f \in \mathcal{C}$ and $a, b \in \mathbf{Z}^V$, if $\arg\min f \cap [a, b] \neq \emptyset$, then $\arg\min f_a^b$ is an M-convex set and f_a^b belongs to $\mathcal{C}$, and hence, Proposition 3.2 is extended to $\mathcal{C}$. Moreover, Theorems 2.3, 2.4 and 2.5 are extended to the class of functions satisfying ($\mathrm{SSQM}^{\neq}$) as follows.

Theorem 4.1 ([18]) *For a function f satisfying ($\mathrm{SSQM}^{\neq}$) and $x \in \mathrm{dom}\, f$, $f(x) \leq f(y)$ for any $y \in \mathbf{Z}^V$ if and only if $f(x) \leq f(x - \chi_u + \chi_v)$ for all $u, v \in V$.*

Theorem 4.2 ([18]) *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function with ($\mathrm{SSQM}^{\neq}$) and $\arg\min f \neq \emptyset$. For $x \in \mathrm{dom}\, f$ and $v \in V$, the following statements hold.*

(a) *If $f(x - \chi_v + \chi_u) = \min_{s \in V} f(x - \chi_v + \chi_s)$, then there exists $x_* \in \arg\min f$ such that $x(u) - \chi_v(u) + 1 \leq x_*(u)$.*

(b) *If $f(x + \chi_v - \chi_u) = \min_{s \in V} f(x + \chi_v - \chi_s)$, then there exists $x_* \in \arg\min f$ such that $x_*(u) \leq x(u) + \chi_v(u) - 1$.*

Theorem 4.3 *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function satisfying ($\mathrm{SSQM}^{\neq}$) and $\arg\min f \neq \emptyset$. For $x \in \mathrm{dom}\, f$, $v \in V$ and $\alpha \in \mathbf{Z}_{++}$, the following statements hold.*

(a) *If $f(x) = \min_{s \in V} f(x - \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x(v) - (n-1)(\alpha - 1) \leq x_*(v)$.*

(b) *If $f(x) = \min_{s \in V} f(x + \alpha(\chi_u - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x_*(v) \leq x(v) + (n-1)(\alpha - 1)$.*

Theorem 4.3 is implicit in [18] and Theorem 2.6 is generalized to quasi M-convex functions by the argument there.

Theorem 4.4 *Let f be a function satisfying ($\mathrm{SSQM}^{\neq}$) and $\arg\min f \neq \emptyset$. For $x \in \mathrm{dom}\, f$, $v \in V$ and $\alpha \in \mathbf{Z}_{++}$, the following statements hold.*

(a) *If $f(x - \alpha(\chi_v - \chi_u)) = \min_{s \in V} f(x - \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x(u) - \alpha(\chi_v(u) - 1) - (n-1)(\alpha - 1) \leq x_*(u)$.*

(b) *If $f(x + \alpha(\chi_v - \chi_u)) = \min_{s \in V} f(x + \alpha(\chi_v - \chi_s))$, then there exists $x_* \in \arg\min f$ such that $x_*(u) \leq x(u) - \alpha(\chi_v(u) - 1) + (n-1)(\alpha - 1)$.*

Proof. Here we prove the assertion (a) because we can similarly prove (b). Let $\hat{x} = x - \alpha(\chi_v - \chi_u)$. It is sufficient to consider the case where $\max\{x'(u) \mid x' \in$

$\arg\min f\} < \widehat{x}(u)$. Let $x_* \in \arg\max\{x'(u) \mid x' \in \arg\min f\}$, and suppose that x_* minimizes the ℓ_1-norm $\| x_* - \widehat{x} \|_1$ among all such vectors.

Let

$$S = \left\{ \widehat{x} - \lambda\chi_u + \sum_{w \in \mathrm{supp}^-(\widehat{x}-x_*)} \mu_w \chi_w \in \mathrm{dom}\, f \;\middle|\; \begin{array}{l} 0 \le \lambda \le \widehat{x}(u) - x_*(u), \\ 0 \le \mu_w \le x_*(w) - \widehat{x}(w) \\ \quad (w \in \mathrm{supp}^-(\widehat{x} - x_*)), \\ \lambda = \sum \mu_w \end{array} \right\}.$$

Claim A: For any $y \in \arg\min\{f(y') \mid y' \in S\}$, we have $y(u) = x_*(u)$.

[Proof of Claim A] Suppose to the contrary that $y(u) > x_*(u)$. From the definition of x_*, we have $f(y) > f(x_*)$. By (SSQM$^{\neq}$) for y, x_* and u, there exists $w \in \mathrm{supp}^-(y - x_*) \subseteq \mathrm{supp}^-(\widehat{x} - x_*)$ such that if $f(x_* + \chi_u - \chi_w) - f(x_*) > 0$ then $f(y - \chi_u + \chi_w) - f(y) < 0$. By the choice of x_*, we have $f(x_* + \chi_u - \chi_w) > f(x_*)$, and hence, $f(y - \chi_u + \chi_w) < f(y)$. However, this contradicts the minimality of $f(y)$ because $y - \chi_u + \chi_w \in S$.

Take any $\tilde{y}$ from $\arg\min\{f(y') \mid y' \in S\}$ and represent

$$\tilde{y} = \widehat{x} - \tilde{\lambda}\chi_u + \sum_{w \in \mathrm{supp}^-(\widehat{x}-x_*)} \tilde{\mu}_w \chi_w.$$

Claim A guarantees that $\tilde{\lambda} = \widehat{x}(u) - x_*(u)$.

Claim B: For any $w \in \mathrm{supp}^-(\widehat{x} - x_*)$ with $\tilde{\mu}_w > 0$ and $\mu \in [0, \tilde{\mu}_w - 1]$, we have $f(\widehat{x} - (\mu+1)(\chi_u - \chi_w)) < f(\widehat{x} - \mu(\chi_u - \chi_w))$.

[Proof of Claim B] We prove the claim by induction on μ. For $\mu \in [0, \tilde{\mu}_w - 1]$, put $x' = \widehat{x} - \mu(\chi_u - \chi_w)$ and assume $x' \in \mathrm{dom}\, f$. We note that $x' \in S$ and $x'(u) > x_*(u)$. Therefore, Claim A yields $f(x') > f(\tilde{y})$. Since $\mathrm{supp}^-(\tilde{y} - x') = \{u\}$, (SSQM$^{\neq}$) applied to $\tilde{y}$, x' and $w (\in \mathrm{supp}^+(\tilde{y} - x'))$ implies that if $f(\tilde{y} + \chi_u - \chi_w) - f(\tilde{y}) > 0$ then $f(x' - \chi_u + \chi_w) - f(x') < 0$. By Claim A, we have $f(\tilde{y} + \chi_u - \chi_w) > f(\tilde{y})$, and hence, $f(\widehat{x} - (\mu+1)(\chi_u - \chi_w)) = f(x' - \chi_u + \chi_w) < f(x')$.

The hypothesis of (a) and Claim B imply

$$f(\widehat{x} - \tilde{\mu}_w(\chi_u - \chi_w)) < \cdots < f(\widehat{x} - (\chi_u - \chi_w)) < f(\widehat{x}) \le f(\widehat{x} - \alpha(\chi_u - \chi_w)),$$

for any w with $\tilde{\mu}_w > 0$. This says $\tilde{\mu}_w \le \alpha - 1$ for any $w \in \mathrm{supp}^-(\widehat{x} - x_*)$. Thus, we have

$$\widehat{x}(u) - x_*(u) = \widehat{x}(u) - \tilde{y}(u) = \tilde{\lambda} = \sum_{w \in \mathrm{supp}^-(\widehat{x}-x_*)} \tilde{\mu}_w \le (n-1)(\alpha - 1),$$

where the third equality follows from the definition of S. ∎

Since our algorithm, COORDINATEWISE_SCALING, relies solely on the above properties and the fact that l and u with $\mathrm{dom}\, f \subseteq [l, u]$ are efficiently computable, it also works with the same time complexity for a subclass of $\mathcal{C}$ such that l and u are efficiently computable.

Theorem 4.5 *Suppose that a function f satisfies* ($\mathrm{SSQM}^{\neq}$), $\arg\min f$ *is an M-convex set, and vectors $x^0, l, u \in \mathbf{Z}^V$ with $x^0 \in \mathrm{dom}\, f \subseteq [l, u]$ are given. Then,* COORDINATEWISE_SCALING *finds a minimizer of f in $O(Fn^3 \log(L/n))$ time, where F is an upper bound of time to evaluate f and $L = \max\{u(v) - l(v) \mid v \in V\}$.*

Acknowledgements

The author thanks Kazuo Murota and Akiyoshi Shioura for their helpful comments and suggestions.

References

[1] V. Danilov, G. Koshevoy, and K. Murota, Discrete convexity and equilibria in economies with indivisible goods and money, *Math. Social Sci.* **41** (2001), 251–273.

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

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

[4] J. Edmonds, Submodular functions, matroids and certain polyhedra, in "Combinatorial Structures and Their Applications" (R. Guy, H. Hanani, N. Sauer and J. Schönheim, Eds.), pp. 69–87, Gordon and Breach, New York, 1970.

[5] A. Frank, Generalized polymatroids, in "Finite and Infinite Sets, I" (A. Hajnal, L. Lovász and V. T. Sós, Eds.), pp. 285–294, North-Holland, Amsterdam, 1984.

[6] A. Frank and É. Tardos, Generalized polymatroids and submodular flows, *Math. Programming* **42** (1988), 489–563.

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

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

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

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

[11] L. Lovász, Submodular functions and convexity, in "Mathematical Programming — The State of the Art" (A. Bachem, M. Grötschel and B. Korte, Eds.), pp. 235–257, Springer-Verlag, Berlin, 1983.

[12] S. Moriguchi, K. Murota and A. Shioura, Minimization of an M-convex function with a scaling technique (in Japanese), *IPSJ SIG Notes* 2001–AL–76 (2001), 27–34.

[13] S. Moriguchi, K. Murota and A. Shioura, Scaling algorithms for M-convex function minimization, Department of Mathematical and Computing Sciences, Technical Report B-371, Tokyo Institute of Technology, 2001.

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

[15] K. Murota, Discrete convex analysis, *Math. Programming* **83** (1998), 313–371.

[16] K. Murota, "Matrices and Matroids for Systems Analysis," Springer-Verlag, Berlin, 2000.

[17] K. Murota, "Discrete Convex Analysis (in Japanese)," Kyoritsu-Shuppan, Tokyo, 2001, to appear.

[18] K. Murota and A. Shioura, Quasi M-convex and L-convex functions: Quasi-convexity in discrete optimization, RIMS Preprint 1306, Kyoto University, 2000.

[19] K. Murota and A. Tamura, New characterizations of M-convex functions and their applications to economic equilibrium models with indivisibilities, *Discrete Appl. Math.* to appear.

[20] R. T. Rockafellar, "Convex Analysis," Princeton University Press, Princeton, 1970.

[21] A. Shioura, Minimization of an M-convex function, *Discrete Appl. Math.* **84** (1998), 215–220.