数理解析研究所講究録 1241 巻 2001 年 66-74



# Scaling Algorithms for M-convex Function Minimization

東工大・情報理工 森口 聡子 (Satoko Moriguchi)
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology
京大・数理解析研究所 室田 一雄 (Kazuo Murota)
Research Institute for Mathematical Sciences, Kyoto University
東北大・情報科学 塩浦 昭義 (Akiyoshi Shioura)
Graduate School of Information Sciences, Tohoku University

**Abstract:** M-convex functions have various desirable properties as convexity in discrete optimization. We can find a global minimum of an M-convex function by a greedy algorithm, i.e., so-called descent algorithms work for the minimization. In this paper, we apply a scaling technique to a greedy algorithm and propose an efficient algorithm for the minimization of an M-convex function. Computational results are also reported.
**Keywords:** matroid, convex function, scaling algorithm, discrete optimization.

## 1 Introduction

The concept of convexity for sets and functions plays a central role in continuous optimization (or nonlinear programming with continuous variable). It has various applications in the areas of mathematical economics, engineering, operations research, etc. [2, 11, 13]. The importance of convexity relies on the fact that a local minimum of a convex function is also a global minimum. Due to this property, we can find a global minimum of a convex function by iteratively moving in descent directions, i.e., so-called descent algorithms work for the convex function minimization.

In the area of discrete optimization, on the other hand, discrete analogues of convexity, or "discrete convexity" for short, have been considered, with a view to identifying the discrete structure that guarantees the success of descent methods, so-called "greedy algorithms." Examples of discrete convexity are "discretely-convex functions" by Miller [5], "integrally-convex functions" by Favati-Tardella [3]. It would be natural to expect that discrete convexity yields a theory of "discrete convex analysis," which covers discrete analogues of the fundamental concepts such as conjugacy, subgradients, duality, and separation theorems. Unfortunately, neither "discretely-convex functions" nor "integrally-convex functions" seem to be fully suitable for such a theory. This suggests that we must identify a more restrictive class of well-behaved "discrete convex functions."

The concept of M-convex functions was proposed by Murota [6, 7] in 1996 as a natural extension of the concept of valuated matroids. Let $V$ be a finite set. 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 $\chi_w \in \{0,1\}^V$ is the characteristic vector of $w \in V$ and

$$\begin{aligned} \mathrm{dom}\, f &= \{x \in \mathbf{Z}^V \mid f(x) < +\infty\}, \\ \mathrm{supp}^+(x-y) &= \{w \in V \mid x(w) > y(w)\}, \\ \mathrm{supp}^-(x-y) &= \{w \in V \mid x(w) < y(w)\}. \end{aligned}$$

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 problem with separable-convex cost functions. M-convex functions have various desirable properties as discrete convexity:

(i) local minimality leads to global minimality for M-convex functions,
(ii) M-convex functions can be extended to ordinary convex functions,
(iii) various duality theorems hold,
(iv) M-convex functions are conjugate to L-convex functions.

In particular, the property (i) shows that greedy algorithms (descent algorithms) work for the minimization of an M-convex function. A theory of "discrete convex analysis" [7, 8, 9] has been developed with the use of M- and L-convex functions.

In this paper, we consider the problem of minimizing an M-convex function. Although an M-convex function can be minimized by a descent algorithm, it may require exponential time. A steepest descent algorithm, a faster version of a descent algorithm, terminates in pseudo-polynomial time. The domain reduction-type polynomial time algorithm of Shioura [12] has the time complexity $\mathrm{O}(n^4(\log L)^2)$, where

$$n = |V|, \qquad L = \max\{\|x - y\|_\infty \mid x, y \in \mathrm{dom}\, f\}.$$

Although the domain reduction-type algorithm has polynomial time complexity, our numerical experiments show that it does not run fast in practice.

The objective of this paper is to propose faster polynomial time algorithms for the minimization of an M-convex function by using a scaling technique. Scaling is a fundamental technique used extensively in polynomial time algorithms for combinatorial optimization problems. Indeed, scaling-based algorithms achieve better time complexities for the resource allocation problem [4], the minimum cost flow problem [1], etc.

We propose efficient minimization algorithms for functions in the class of M-convex functions closed under the scaling operation. We apply the scaling technique to a steepest descent algorithm to obtain faster algorithms. A minimizer of an M-convex function $f$ can be found by the scaling algorithms proposed in this paper. Moreover, if $f$ is in the class of M-convex functions closed under the scaling operation, the time complexity of each scaling algorithms is bounded by a polynomial in $n$ and $L$. Some fundamental classes of M-convex functions such as separable convex functions and quadratic M-convex functions are closed under the scaling operation, although this is not the case with general M-convex functions.

In order to compare the performance of our new scaling algorithms to those of the previously proposed algorithms, we make numerical experiments with randomly generated test problems. It is observed from numerical results that our new scaling algorithms are much faster than the previously proposed algorithms from the viewpoint of both theory and practice.

## 2 Scaling of M-convex Functions

For $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$, a positive integer $\alpha$ and a vector $b \in \mathbf{Z}^V$, define a function $f^{\alpha,b} : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ by

$$f^{\alpha,b}(x) = f(\alpha x + b) \qquad (x \in \mathbf{Z}^V).$$

This operation is called *scaling*. Even if $f$ is an M-convex function, $f^{\alpha,b}$ is not necessarily M-convex in general. We can still identify a number of subclasses of M-convex functions that are closed under the scaling operation.

**Example 2.1 (Separable convex functions) :** For a family of convex functions $f_i : \mathbf{Z} \to \mathbf{R}$ indexed by $i \in V$ and an integer $\beta$, the (separable convex) function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ defined by

$$f(x) = \begin{cases} \displaystyle\sum_{i=1}^{n} f_i(x_i) & \text{if } x(V) = \beta, \\ +\infty & \text{otherwise} \end{cases}$$

is M-convex.

Since $f^{\alpha,b}(x) = \sum_{i=1}^{n} f_i(\alpha x_i + b_i)$ is also a separable convex function, the class of separable convex functions is closed under the scaling operation.

**Example 2.2 (Quadratic M-convex functions) :** Let $A = (a_{ij}) \in \mathbf{R}^{n\times n}$ be a symmetric matrix. A quadratic function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ given by

$$f(x) = \begin{cases} \frac{1}{2}x^{\mathrm{T}}Ax & \text{if } x(V) = 0, \\ +\infty & \text{otherwise} \end{cases}$$

is M-convex if and only if

$$\forall i, j, k, l \in V \text{ with } \{i, j\} \cap \{k, l\} = \emptyset,\ a_{ij} + a_{kl} \geq \min\{a_{ik} + a_{jl}, a_{il} + a_{jk}\}$$

(see [9, 10]). For the quadratic M-convex function $f$, the function $f^{\alpha,b}$ is written as

$$f^{\alpha,b}(x) = \frac{1}{2}(\alpha x + b)^{\mathrm{T}}A(\alpha x + b) = \frac{\alpha^2}{2}x^{\mathrm{T}}Ax + \alpha b^{\mathrm{T}}Ax + \frac{1}{2}b^{\mathrm{T}}Ab.$$

This expression shows that the function $f^{\alpha,b}$ is M-convex. Therefore, the class of quadratic M-convex functions is closed under the scaling operation.

**Example 2.3 (Laminar convex functions) :** A nonempty family $\mathcal{T}$ of subsets of $V$ is called a *laminar family* if it satisfies the following property:

$$\forall X, Y \in \mathcal{T} : X \cap Y = \emptyset \text{ or } X \subseteq Y \text{ or } X \supseteq Y.$$

Given a laminar family $\mathcal{T}$ and a family of convex functions $f_X : \mathbf{Z} \to \mathbf{R}$ indexed by $X \in \mathcal{T}$ as well as an integer $\beta$, define a function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ by

$$f(x) = \begin{cases} \sum_{X\in\mathcal{T}} f_X(x(X)) & \text{if } x(V) = \beta, \\ +\infty & \text{otherwise.} \end{cases}$$

This is called a *laminar convex function.* We show that laminar convex functions constitute a class of M-convex functions closed under the scaling operation.

Without loss of generality, assume $V \in \mathcal{T}$. Otherwise, we can add $V$ to $\mathcal{T}$ and put $f_V(\alpha) = 0$ $(\forall \alpha \in \mathbf{Z})$. We denote by $\mathcal{T}(X)$ the family of all maximal proper subsets of $X$ in $\mathcal{T}$. For any $x \in \mathbf{Z}^V$ we have

$$x(X) = \sum\{x(Y) \mid Y \in \mathcal{T}(X)\} + \sum\{x(v) \mid v \in X \setminus \bigcup_{Y\in\mathcal{T}(X)} Y\}. \tag{1}$$

Take any $x, y \in \operatorname{dom} f$ and $u \in \operatorname{supp}^+(x - y)$. To prove (M-EXC), it suffices to show that there exists some $v \in \operatorname{supp}^-(x - y)$ satisfying

$$u \in X,\ v \notin X,\ X \in \mathcal{T} \implies x(X) > y(X) \tag{2}$$

and

$$u \notin X,\ v \in X,\ X \in \mathcal{T} \implies x(X) < y(X). \tag{3}$$

Let $X_0$ be the unique minimal set in $\mathcal{T}$ satisfying $u \in X$ and $x(X) \leq y(X)$. By the minimality of $X_0$ and (1), there are two cases:

(i) $\exists v \in X_0 \setminus \bigcup_{Y\in\mathcal{T}(X_0)} Y : x(v) < y(v)$,

(ii) $\exists X_1 \in \mathcal{T}(X_0) : x(X_1) < y(X_1)$.

In case of (i), this $v$ satisfies (2) and (3). In case of (ii), from (1) follows

(i) $\exists v \in X_1 \setminus \bigcup_{Y\in\mathcal{T}(X_1)} Y : x(v) < y(v)$, or

(ii) $\exists X_2 \in \mathcal{T}(X_1) : x(X_2) < y(X_2)$.

Repeating this argument, we reach the case (i). Therefore, a laminar convex function is M-convex.

Moreover,

$$f^{\alpha,b}(x) = \sum_{X\in\mathcal{T}} f_X(\alpha x(X) + b(X))$$

is a laminar convex function. Therefore the class of laminar convex functions is closed under the scaling operation.

## 3 Theorems on the Minimizers of M-convex Functions

Global minimality of an M-convex function is characterized by local minimality.

**Theorem 3.1 ([6, 7])** : *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function with* (M-EXC). *For $x \in \mathrm{dom}\, f$, $f(x) \leq f(y)$ $(\forall y \in \mathbf{Z}^V)$ if and only if $f(x) \leq f(x - \chi_u + \chi_v)$ $(\forall u, v \in V)$.* ∎

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

**Theorem 3.2 ([12])** : *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function with* (M-EXC). *Assume $\arg\min f \neq \emptyset$.*
(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 $x^*(u) \leq x'(u)$.*
(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 $x^*(v) \geq x'(v)$.* ∎

**Corollary 3.3 ([12])** : *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 $x^*(u) \leq x(u) - 1$, $x^*(v) \geq x(v) + 1$.* ∎

Let $\alpha$ be a positive integer, and $x_\alpha \in \mathrm{dom}\, f$. We call $x_\alpha$ an $\alpha$-local minimum of $f$ if it satisfies

$$f(x_\alpha) \leq f(x_\alpha + \alpha(\chi_v - \chi_u)) \ (\forall u, v \in V).$$

The following is a "proximity theorem," showing that a global minimizer of an M-convex function exists in the neighborhood of an $\alpha$-local minimum.

**Theorem 3.4.** *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function and $\alpha$ be any positive integer. Suppose that $x_\alpha \in \mathrm{dom}\, f$ satisfies $f(x_\alpha) \leq f(x_\alpha + \alpha(\chi_v - \chi_u))$ for all $u, v \in V$. Then, $\arg\min f \neq \emptyset$ and there exists some $x_* \in \arg\min f$ such that*

$$|x_\alpha(v) - x_*(v)| \leq (n-1)(\alpha - 1) \qquad (v \in V). \tag{4}$$

∎

**Proof.** It suffices to show that for any $\gamma > \inf f$ there exists some $x_* \in \mathrm{dom}\, f$ satisfying $f(x_*) \leq \gamma$ and (4).

Let $x_* \in \mathrm{dom}\, f$ satisfy $f(x_*) \leq \gamma$, and suppose that $x_*$ minimizes the value $\|x_* - x_\alpha\|_1$ among all such vectors. In the following, we fix $v \in V$ and prove $x_\alpha(v) - x_*(v) \leq (n-1)(\alpha-1)$. The inequality $x_*(v) - x_\alpha(v) \leq (n-1)(\alpha-1)$ can be shown similarly.

We may assume $x_\alpha(v) > x_*(v)$. We first prove the following two claims. Let $k = x_\alpha(v) - x_*(v)$.

**Claim 1.** There exist $w_1, w_2, \cdots, w_k \in V \setminus \{v\}$ and $y_0(= x_\alpha), y_1, \cdots, y_k \in \mathrm{dom}\, f$ such that

$$y_i = y_{i-1} - \chi_v + \chi_{w_i}, \ f(y_i) < f(y_{i-1}) \ (i = 1, \cdots, k).$$

[Proof of Claim 1] We show the claim by induction on $i$. Suppose $y_{i-1} \in \mathrm{dom}\, f$. By (M-EXC) applied to $y_{i-1}, x_*$, and $v \in \mathrm{supp}^+(y_{i-1} - x_*)$, we have some $w_i \in \mathrm{supp}^-(y_{i-1} - x_*) \subseteq \mathrm{supp}^-(x_\alpha - x_*) \subseteq V \setminus \{v\}$ such that $f(x_*) + f(y_{i-1}) \geq f(x_* - \chi_{w_i} + \chi_v) + f(y_{i-1} + \chi_{w_i} - \chi_v)$. By the choice of $x_*$, we have $f(x_* + \chi_v - \chi_{w_i}) > f(x^*)$ since $\|(x_* + \chi_v - \chi_{w_i}) - x_\alpha\|_1 < \|x_* - x_\alpha\|_1$. Therefore, $f(y_i) = f(y_{i-1} - \chi_v + \chi_{w_i}) < f(y_{i-1})$. [End of Proof for Claim 1]

**Claim 2.** For any $w \in V \setminus \{v\}$ with $y_k(w) > x_\alpha(w)$ and $\mu \in [0, y_k(w) - x_\alpha(w) - 1]$, we have

$$f(x_\alpha - (\mu + 1)(\chi_v - \chi_w)) < f(x_\alpha - \mu(\chi_v - \chi_w)). \tag{5}$$

[Proof of Claim 2] We prove (5) by induction on $\mu$. Put $x' = x_\alpha - \mu(\chi_v - \chi_w)$ for $\mu \in [0, y_k(w) - x_\alpha(w) - 1]$, and suppose $x' \in \mathrm{dom}\, f$. Let $j_*$ $(1 \leq j_* \leq k)$ be the largest index such that $w_{j_*} = w$. Then, $y_{j_*}(w) = y_k(w) > x'(w)$ and $\mathrm{supp}^-(y_{j_*} - x') = \{v\}$. (M-EXC) implies that $f(x') + f(y_{j_*}) \geq f(x' - \chi_v + \chi_w) + f(y_{j_*} + \chi_v - \chi_w)$. By Claim 1, we have $f(y_{j_*} + \chi_v - \chi_w) > f(y_{j_*})$. Hence, (5) follows. [End of Proof for Claim 2]

The $\alpha$-local minimality of $x_\alpha$ implies $f(x_\alpha - \alpha(\chi_v - \chi_w)) \geq f(x_\alpha)$, which, combined with Claim 2, implies $y_k(w) - x_\alpha(w) \leq \alpha - 1$ for all $w \in V \setminus \{v\}$. Thus,

$$x_\alpha(v) - x_*(v) \;=\; x_\alpha(v) - y_k(v) = \sum_{w \in V \setminus \{v\}} \{y_k(w) - x_\alpha(w)\} \leq (n-1)(\alpha - 1),$$

where the second equality is by $x(V) = y(V)$ $(\forall x, y \in \mathrm{dom}\, f)$. □

# 4 Minimization Algorithms of an M-convex Function

## 4.1 Previous Algorithms

Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function such that $\mathrm{dom}\, f$ is a nonempty bounded set, and put $L = \max\{\|x - y\|_\infty \mid x, y \in \mathrm{dom}\, f\}$. Assume (M-EXC) for $f$. Then, Theorem 3.1 and Corollary 3.3 immediately lead to the following algorithm.

**Algorithm** STEEPEST_DESCENT (SD)
**S0:** Let $x$ be any vector in dom $f$. Set $B := \mathrm{dom}\, f$.
**S1:** If $f(x) = \min_{s,t \in V} f(x - \chi_s + \chi_t)$ then stop [$x$ is a minimizer of $f$].
**S2:** Find $u, v \in V$ with $x - \chi_u + \chi_v \in B$ satisfying

$$f(x - \chi_u + \chi_v) = \min\{f(x - \chi_s + \chi_t) \mid s, t \in V, x - \chi_s + \chi_t \in B\}.$$

**S3:** Set $x := x - \chi_u + \chi_v$ and $B := B \cap \{y \in \mathbf{Z}^V \mid y(u) \leq x(u) - 1,\ y(v) \geq x(v) + 1\}$.
Go to S1. □

By Corollary 3.3, the set $B$ always contains a minimizer of $f$. Hence, Algorithm SD finds a minimizer of $f$. To analyze the number of iterations, we consider the value

$$\sum_{w \in V} \{\max_{y \in B} y(w) - \min_{y \in B} y(w)\}.$$

This value is bounded by $nL$ and decreases at least by two in each iteration. Therefore, SD terminates in $\mathrm{O}(nL)$ iterations. Each iteration can be done in $\mathrm{O}(n^2)$ time. Therefore, Algorithm SD finds a minimizer of $f$ in $\mathrm{O}(n^3 L)$ time, i.e., SD is a pseudo-polynomial time algorithm. In particular, if $\mathrm{dom}\, f \subseteq \{0,1\}^V$ then the number of iterations is $\mathrm{O}(n^2)$.

We propose the following modified version of Algorithm SD, where we exploit Theorem 3.2.

**Algorithm** MODIFIED_STEEPEST_DESCENT (MSD)
**S0:** Let $x$ be any vector in dom $f$. Set $B := \mathrm{dom}\, f$.
**S1:** Choose $u \in V$ such that $\exists v \in V$ with $x - \chi_u + \chi_v \in B$. If there is no such $u$ then stop [$x$ is a minimizer of $f$].
**S2:** For $u$, find $v \in V$ with $x - \chi_u + \chi_v \in B$ satisfying

$$f(x - \chi_u + \chi_v) = \min\{f(x - \chi_u + \chi_t) \mid t \in V,\ x - \chi_u + \chi_t \in B\}.$$

**S3:** Set $x := x - \chi_u + \chi_v$ and $B := B \cap \{y \in \mathbf{Z}^V \mid y(v) \geq x(v) + 1\}$. Go to S1. □

Although the number of iterations of Algorithm MSD is equal to that of Algorithm SD, each iteration of MSD can be done in $O(n)$ time, while each iteration of SD can be done in $O(n^2)$ time. MSD is also a pseudo-polynomial time algorithm.

It is shown in [12] that the minimization of an M-convex function can be done in polynomial time by the domain reduction method explained below.

Given a bounded M-convex set $B$, the set $N_B \subseteq B$ is defined as follows. For $w \in V$, define

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

$$l'_B(w) = \left\lfloor (1 - \frac{1}{n}) l_B(w) + \frac{1}{n} u_B(w) \right\rfloor, \qquad u'_B(w) = \left\lceil \frac{1}{n} l_B(w) + (1 - \frac{1}{n}) u_B(w) \right\rceil.$$

Then, $N_B$ is defined as

$$N_B = \{y \in B \mid l'_B(w) \leq y(w) \leq u'_B(w) \ (\forall w \in V)\}.$$

**Theorem 4.1 ([12]) :** *$N_B$ is a (nonempty) M-convex set.* ∎

The next algorithm maintains a set $B \subseteq \operatorname{dom} f$ which is an M-convex set containing a minimizer of $f$. It reduces $B$ iteratively by exploiting Corollary 3.3 and finally finds a minimizer.

**Algorithm** DOMAIN_REDUCTION (DR)
**S0:** Set $B := \operatorname{dom} f$.
**S1:** Find a vector $x \in N_B$.
**S2:** If $f(x) = \min_{s,t \in V} f(x - \chi_s + \chi_t)$ then stop [$x$ is a minimizer of $f$].
**S3:** Find $u, v \in V$ with $x - \chi_u + \chi_v \in B$ satisfying

$$f(x - \chi_u + \chi_v) = \min\{f(x - \chi_s + \chi_t) \mid s, t \in V,\ x - \chi_s + \chi_t \in B\}.$$

**S4:** Set $B := B \cap \{y \in \mathbf{Z}^V \mid y(u) \leq x(u) - 1,\ y(v) \geq x(v) + 1\}$. Go to S1. □

**Theorem 4.2 ([12]) :** *If a vector in* dom $f$ *and the value $L$ are given, Algorithm* DR *finds a minimizer of $f$ in $O(n^4 (\log L)^2)$ time.* ∎

## 4.2 Scaling Algorithms

We apply a scaling technique to Algorithm SD to obtain a faster algorithm.

**Algorithm** SCALING_STEEPEST_DESCENT (SSD)
**S0:** Put $\alpha := 2^{\lceil \log(L/4n) \rceil}$, $B := \operatorname{dom} f$. Let $x_{2\alpha}$ be any vector in $\operatorname{dom} f$.
**S1:** [$\alpha$-scaling phase] Define $\tilde{f} : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ by

$$\tilde{f}(y) = \begin{cases} f(x_{2\alpha} + \alpha y) & \text{if } x_{2\alpha} + \alpha y \in B, \\ +\infty & \text{if } x_{2\alpha} + \alpha y \notin B. \end{cases}$$

Compute a minimizer $y_*$ of $\tilde{f}$ by applying Algorithm STEEPEST_DESCENT.
Set $x_\alpha = x_{2\alpha} + \alpha y_*$.
**S2:** If $\alpha = 1$ then stop [$x_\alpha$ is a minimizer of $f$].
**S3:** Put

$$B := B \cap \{y \in \mathbf{Z}^V \mid x_\alpha(w) - (n-1)(\alpha - 1) \leq y(w) \leq x_\alpha(w) + (n-1)(\alpha - 1) \ (\forall w \in V)\}$$

and $\alpha := \alpha/2$. Go to S1. □

Although this algorithm works for any M-convex function, it does not terminate in polynomial time in general. This algorithm terminates in polynomial time for a function in the class of M-convex functions closed under the scaling operation. We analyze the time complexity of Algorithm SSD for a function closed under the scaling operation. The number of scaling phases is $\lceil \log(L/4n) \rceil$. Since the number of iterations in each scaling phase is $(4n\alpha \times n)/\alpha$, each scaling phase terminates in $O((4n\alpha \times n)/\alpha \times n^2) = O(n^4)$ time. We can compute the value $L$ in $O(n^2 \log L)$ time. Here, we have the following theorem.

**Theorem 4.3.** *Suppose that* $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ *satisfies* (M-EXC) *and is closed under the scaling operation. If a vector in* $\mathrm{dom}\, f$ *is given, Algorithm* SSD *finds a minimizer of* $f$ *in* $O(n^4 \log(L/n))$ *time.*

Algorithm SSD above can be improved further by using MSD in place of SD in each scaling phase. We refer to the algorithm resulting from this modification as SCALING_MODIFIED_STEEPEST_DESCENT (SMSD). Each scaling phase of SMSD terminates in $O(n^3)$ time, and therefore, its overall time complexity for finding a minimizer of $f$ is $O(n^3 \log(L/n))$. Thus the replacement of SD by MSD results in an $O(n)$ improvement upon SSD.

**Theorem 4.4.** *Suppose that* $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ *satisfies* (M-EXC) *and is closed under the scaling operation. If a vector in* $\mathrm{dom}\, f$ *is given, Algorithm* SMSD *finds a minimizer of* $f$ *in* $O(n^3 \log(L/n))$ *time.* ∎

# 5 Numerical Experiments

## 5.1 Test Problems and Implementation

As test problems we consider the minimization of a quadratic laminar convex function of the following form:

$$\begin{array}{lll} \text{TREE}: & \text{minimize} & \displaystyle\sum_{X \in \mathcal{T}} \{a_X x(X)^2 + b_X x(X) + c_X\} \\ & \text{subject to} & \displaystyle\sum_{i=1}^{n} x(i) = L, \\ & & x_i \geq 0, \ \text{integer}, \ i = 1, \ldots, n. \end{array}$$

For each $n$ and $L$ fixed (dimension of the variable $x$ and the sum of $x(i)$, respectively), we generated ten test problems with randomly chosen real variables $0 \leq a_X,\ b_X,\ c_X \leq 1000$ ($X \in \mathcal{T}$) and laminar families $\mathcal{T}$. The C language function `random()` is used to generate these pseudo-random numbers. We measure the execution time and present average execution times of ten generated test problems for each size. The two main parameters $n$ and $L$ have a strong influence on the execution time. We make experiments with test problems of various sizes by changing $n$ and $L$. For comparison of the performance of four algorithms, we implemented SD, DR, SSD and SMSD.

In our implementation, we tailored DR for the minimization of a laminar convex function, in which the following algorithm is used to find a vector $x$ in $N_B$.

**Algorithm** FIND_VECTOR_IN_$N_B$
**S1:** For each $w \in V$, compute $l'_B(w)$ and $u'_B(w)$.
**S2:** For $w = 1, 2, \cdots, n$, put

$$x(w) = \begin{cases} u'_B(w) & \text{if } \displaystyle\sum_{i=1}^{w-1} x(w) + u'_B(w) + \sum_{i=w+1}^{n} l'_B(i) \leq L, \\ L - \displaystyle\sum_{i=1}^{w-1} x(i) - \sum_{i=w+1}^{n} l'_B(i) & \text{otherwise.} \end{cases}$$

□

Algorithm FIND_VECTOR_IN_$N_B$ finds a vector in $N_B$ in $O(n)$ time. The time complexity of the specialized DR is $O(n^4 \log L)$ while those of DR mentioned in Section 4 is $O(n^4 (\log L)^2)$.

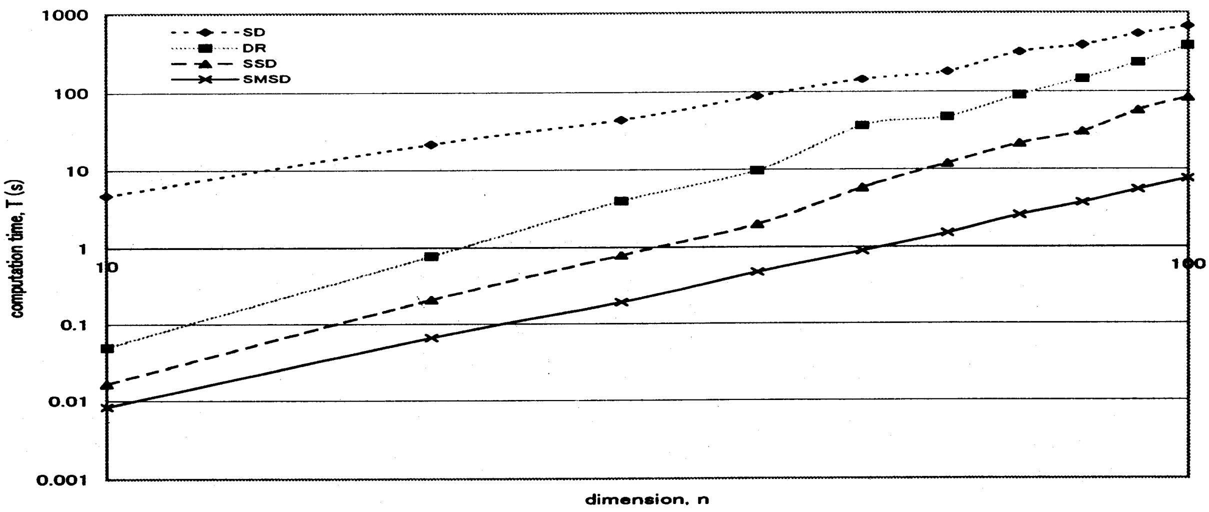


Figure 1: The execution time in the case $L = 50000$.

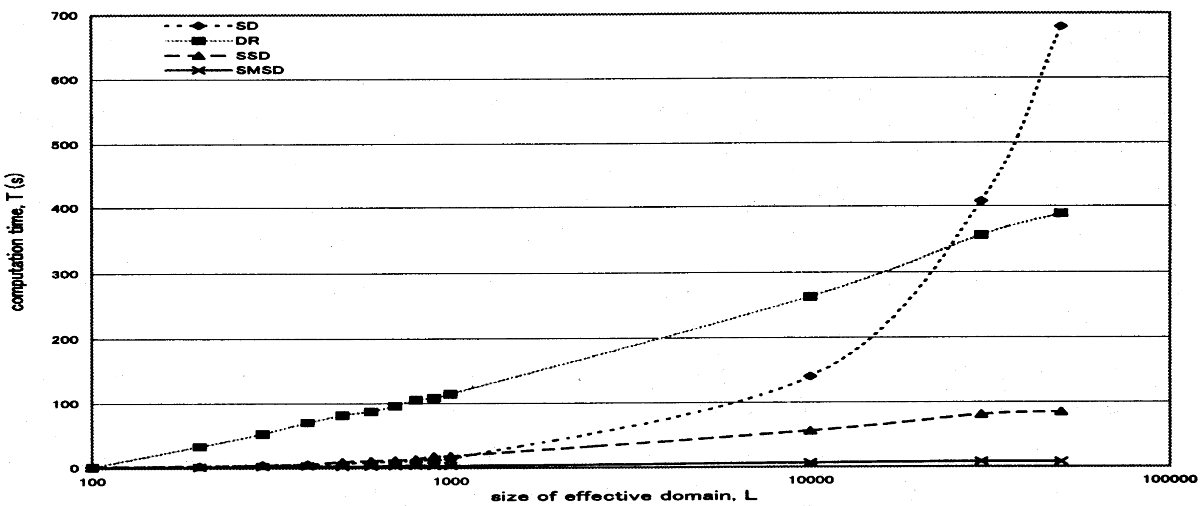


Figure 2: The execution time in the case $n = 100$.

Also, in our implementations of SD, DR, SSD and SMSD, it takes $O(n)$ time to evaluate the function value. Hence, the execution time in our numerical experiments is $O(n)$ times larger than the theoretical time complexity.

Each of SD, DR, SSD and SMSD is written in the C language, compiled under a personal computer with the CPU Pentium III 450MHz and 256 MB of memory under Vine Linux V1.1 using the compiler pgcc 2.95.2 with the option -mcpu=pentiumpro -march=pentiumpro -O9 -funroll-loops.

## 5.2 Computational Results

Our numerical results are summarized in Figures 1 and 2. Figure 1 shows the relationship between the computation time $T$ and the dimension $n$ for the case of $L = 50000$. In all the four algorithms the relationship is linear in $\log T$ and $\log n$, which implies $T = O(n^p)$ for some $p$. Our results show the following:

| Algorithm | SD | DR | SSD | SMSD |
|---|---|---|---|---|
| Time $T$ | $n^{2.16}$ | $n^{3.89}$ | $n^{3.70}$ | $n^{2.96}$ |

Figure 2 shows the relationship between the computation time $T$ and the size of the effective domain $L$ for the case of $n = 100$. $L$ is given in log scale whereas time $T$ is on linear scale in this graph. It is verified that $T = O(\log L)$ in SSD and SMSD, $T = O((\log L)^2)$ in DR and $T = O(L)$ in SD.

The table below shows the standard deviations of execution times in the case of $L = 50000$ and $n = 100$, which is the case of the biggest problems in our numerical experiments.

| Algorithm | SD | DR | SSD | SMSD |
|---|---|---|---|---|
| Stand. Dev. | 2.770 | 10.25 | 0.9075 | 0.2357 |

By numerical experiments with randomly generated test problems, we can conclude that our scaling algorithms are faster than the previously proposed algorithms. In particular, Algorithm SMSD is the fastest algorithm.

## Acknowledgement

The authors thank Yoshitsugu Yamamoto for a stimulating comment.

## References

[1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network Flows – Theory, Algorithms, and Applications, Prentice Hall, New Jersey, 1993.

[2] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, Nonlinear Programming: Theory and Algorithm (Second Edition), John Wiley and Sons, New York, 1993.

[3] P. Favati and F. Tardella, "Convexity in nonlinear integer programming," Ricerca Operativa, vol.53, pp.3–44, 1990.

[4] D.S. Hochbaum, "Lower and upper bounds for the allocation problem and other nonlinear optimization problems," Math. Oper. Res., vol.19, pp.390–409, 1994.

[5] B.L. Miller, "On minimizing nonseparable functions defined on the integers with an investory application," SIAM J. Appl. Math., vol.21, pp.166–185, 1971.

[6] K. Murota, "Convexity and Steinitz's exchange property," Adv. Math., vol.124, pp.272–311, 1996.

[7] K. Murota, "Discrete convex analysis," Math. Programming, vol.83, pp.313–371, 1998.

[8] K. Murota, "Discrete convex analysis — Exposition on conjugacy and duality," in Graph Theory and Combinatorial Biology, eds. L. Lovász et al., pp.253–278, The János Bolyai Mathematical Society, 1999.

[9] K. Murota, Discrete Convex Analysis, Kyoritsu-Shuppan, Tokyo, 2001. [In Japanese]

[10] K. Murota and A. Shioura, "Quadratic M-convex and L-convex functions," RIMS Preprint No. 1326, Kyoto University, 2001.

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

[12] A. Shioura, "Minimization of an M-convex function," Discrete Appl. Math., vol.84, pp.215–220, 1998.

[13] J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimension I, Springer-Verlag, Berlin, 1970.