Exact Bounds for Steepest Descent Algorithms of L-convex Function Minimization

Kazuo Murota

Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan

Akiyoshi Shioura

Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan

---

Abstract

We analyze minimization algorithms for L$^\natural$-convex functions in discrete convex analysis, and establish exact bounds for the number of iterations required by the steepest descent algorithm and its variants.

Keywords: discrete convex function, analysis of algorithm, discrete optimization, steepest descent algorithm

---

1. Results

In this paper, we discuss minimization algorithms for discrete convex functions defined on integer lattice points called L$^\natural$-convex functions. A function $g : \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be *L$^\natural$-convex* [14] if for every $p, q \in \mathrm{dom}\, g$ and every nonnegative $\lambda \in \mathbb{Z}_+$, it holds that

$$g(p) + g(q) \geq g((p + \lambda \mathbf{1}) \wedge q) + g(p \vee (q - \lambda \mathbf{1})), \tag{1}$$

where $\mathrm{dom}\, g = \{p \in \mathbb{Z}^n \mid g(p) < +\infty\}$, $\mathbf{1} = (1, 1, \ldots, 1)$, and for $p, q \in \mathbb{Z}^n$ the vectors $p \wedge q$ and $p \vee q$ denote, respectively, the vectors of component-wise minimum and maximum of p and q. The concept of L$^\natural$-convex function plays a primary role in the theory of discrete convex analysis [14], and there exist many examples of L$^\natural$-convex functions arising from applications in various research areas such as discrete optimization, iterative auctions, and

Email addresses: `murota@mist.i.u-tokyo.ac.jp` (Kazuo Murota), `shioura@dais.is.tohoku.ac.jp` (Akiyoshi Shioura)

computer vision (see Section 2; see also [14]). Some applications of L$^\natural$-convex functions in inventory systems can also be found in [11, 19].

We consider minimization of an L$^\natural$-convex function $g : \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ with $\arg\min g \neq \emptyset$. It is known that this problem can be solved by the following steepest descent algorithm [15], where an initial vector $p^\circ \in \mathrm{dom}\, g$ is assumed to be given. Let $N = \{1, 2, \ldots, n\}$ and denote by $e_X \in \{0, 1\}^n$ the characteristic vector of $X \subseteq N$, i.e., $e_X(i) = 1$ if $i \in X$ and $e_X(i) = 0$ if $i \in N \setminus X$.

Algorithm SteepestDescent
Step 0: Set $p := p^\circ$.
Step 1: Find $\sigma \in \{+1, -1\}$ and $X \subseteq N$ that minimize $g(p + \sigma e_X)$.
Step 2: If $g(p + \sigma e_X) = g(p)$, then output p and stop.
Step 3: Set $p := p + \sigma e_X$ and go to Step 1.

Theorem 1.1 ([15]). *Suppose that* $\mathrm{dom}\, g$ *is bounded. Then, the algorithm* SteepestDescent *outputs a minimizer of* g *in* $\mathrm{O}(nK_\infty)$ *iterations, where* $K_\infty = \max\{\|p - q\|_\infty \mid p, q \in \mathrm{dom}\, g\}$.

The bound $\mathrm{O}(nK_\infty)$ for the number of iterations is later improved to $2K_\infty+1$ [13, Theorem 2.8].

The main aim of this paper is to give a refined analysis of this algorithm in terms of the "distance" between the initial vector and a minimizer of g. For a vector $q \in \mathbb{Z}^n$, denote

$$\|q\|_\infty^+ = \max_{i \in N} \max(0, q(i)), \qquad \|q\|_\infty^- = \max_{i \in N} \max(0, -q(i)).$$

Note that

$$\|q\|_\infty = \max(\|q\|_\infty^+, \|q\|_\infty^-)$$

holds, and $\|q\|_\infty^+ + \|q\|_\infty^-$ serves as a norm of q (satisfying the axioms of norms). Accordingly, the value $\|p^* - p\|_\infty^+ + \|p^* - p\|_\infty^-$ represents a distance between two vectors p^* and p. For $p \in \mathbb{Z}^n$, we define

$$\mu(p) = \min\{\|p^* - p\|_\infty^+ + \|p^* - p\|_\infty^- \mid p^* \in \arg\min g\},$$

which measures the distance between the vector p and the set of minimizers of g.

It is easy to see that $\mu(p)$ decreases by at most one if p is updated by adding or subtracting a 0-1 vector, i.e., $\mu(p + \sigma e_X) \geq \mu(p) - 1$ for $\sigma \in \{+1, -1\}$ and $X \subseteq N$. This implies that $\mu(p) + 1$ is a lower bound for the number of iterations in SteepestDescent. This is also an upper bound as follows.

Theorem 1.2. *The algorithm* SteepestDescent *terminates exactly in* $\mu(p^\circ) + 1$ *iterations.*

In some applications (see Section 2), the following variant of the steepest descent algorithm can be found, where the vector p is always incremented.

Algorithm SteepestDescentUp
Step0: Set $p := p^\circ$, where $p^\circ \in \mathbb{Z}^n$ is a lower bound of some $p^* \in \arg\min g$.
Step1: Find $X \subseteq N$ that minimizes $g(p + e_X)$.
Step2: If $g(p + e_X) = g(p)$, then output p and stop.
Step3: Set $p := p + e_X$ and go to Step 1.

For the analysis of SteepestDescentUp, we define

$$\hat{\mu}(p) = \min\{\|p^* - p\|_\infty \mid p^* \in \arg\min g,\ p^* \geq p\} \quad (p \in \mathbb{Z}^n).$$

Theorem 1.3. *Suppose that the initial vector* $p^\circ \in \mathrm{dom}\, g$ *in the algorithm* SteepestDescentUp *is a lower bound of some minimizer of* g*. Then, the algorithm outputs a minimizer of* g *and terminates exactly in* $\hat{\mu}(p^\circ) + 1$ *iterations.*

Similarly to SteepestDescentUp, we can consider an algorithm SteepestDescentDown, where the vector is decreased by a vector $e_X \in \{0, 1\}^n$ that minimizes $g(p - e_X)$.

Theorem 1.4. *Suppose that the initial vector* $p^\circ \in \mathrm{dom}\, g$ *in the algorithm* SteepestDescentDown *is an upper bound of some minimizer of* g*. Then, the algorithm outputs a minimizer of* g *and terminates exactly in* $\check{\mu}(p^\circ) + 1$ *iterations, where*

$$\check{\mu}(p) = \min\{\|p^* - p\|_\infty \mid p^* \in \arg\min g,\ p^* \leq p\} \quad (p \in \mathbb{Z}^n).$$

Theorems 1.2, 1.3, and 1.4 show that the trajectory of a vector p generated by the steepest descent algorithms is the "shortest" path between the initial vector and a minimizer of g. This reveals an additional advantage of the steepest descent algorithms, which is important in applications such as iterative auction and computer vision. The proofs of Theorems 1.2 and 1.3 are given in Section 3. The proof of Theorem 1.4 is essentially the same as that for Theorem 1.3 and omitted.

Remark 1.5. The algorithms SteepestDescentUp and SteepestDescentDown are originally proposed for a subclass of L$^\natural$-convex functions called L-convex functions (see, e.g., [14, Section 10.3.1]); a function g :

$\mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ is called *L-convex* if it is an L$^\natural$-convex function satisfying an additional property that

$$g(p + \mathbf{1}) = g(p) + r \qquad (\forall p \in \mathbb{Z}^n) \tag{2}$$

for some real number r. While the class of L-convex functions is a subclass of L$^\natural$-convex functions, they are equivalent concepts in the sense that a function $g : \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ is L$^\natural$-convex if and only if a function $\tilde{g} : \mathbb{Z} \times \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ given by

$$\tilde{g}(p_0, p) = g(p - p_0\mathbf{1}) \qquad ((p_0, p) \in \mathbb{Z} \times \mathbb{Z}^n)$$

is L-convex. Note that for an L-convex function g with $\arg\min g \neq \emptyset$, every vector $p \in \mathrm{dom}\, g$ is lower and upper bounds of some $p^* \in \arg\min g$ by the property (2). Hence, the minimization of an L-convex function can be solved by SteepestDescentUp and SteepestDescentDown with an arbitrarily chosen initial vector.

The minimization of an L-convex function often appears in applications, and some existing algorithms used in such applications can be regarded as special cases of SteepestDescentUp and SteepestDescentDown (see Section 2). □

An Example. We illustrate the behavior of the algorithms SteepestDescent and SteepestDescentUp for an L$^\natural$-convex function $g : \mathbb{Z}^2 \to \mathbb{R} \cup \{+\infty\}$ given by

$$\begin{aligned} \mathrm{dom}\, g &= \{(p(1), p(2)) \in \mathbb{Z}^2 \mid 0 \leq p(i) \leq 4 \ (i = 1, 2)\}, \\ g(p(1), p(2)) &= \max(0, -p(1) + 2, -p(2) + 1, \\ &\qquad -p(1) + p(2) - 1, p(1) - p(2) - 2) \quad ((p(1), p(2)) \in \mathrm{dom}\, g) \end{aligned}$$

(see Figure 1). L$^\natural$-convexity of g can be confirmed by checking the inequality (1) for every pair of vectors $p, q \in \mathrm{dom}\, g$; alternatively, we can use various characterizations of L$^\natural$-convex functions (see [14]).

If we apply SteepestDescent with the initial vector $p^\circ = (1, 4)$, then the trajectory of vector p is one of the three paths depicted by dotted arrows in Figure 1 and the minimizer found by the algorithm is either $(3, 4)$ or $(2, 3)$. Note that $\mu(p^\circ) = 2$ with

$$\|p^* - p^\circ\|_\infty^+ + \|p^* - p^\circ\|_\infty^- = \begin{cases} 2 + 0 & \text{for } p^* = (3, 4), \\ 1 + 1 & \text{for } p^* = (2, 3). \end{cases}$$

For another initial vector $p^\circ = (0, 0)$, the trajectory of vector p in SteepestDescent (or SteepestDescentUp) is one of the three paths depicted

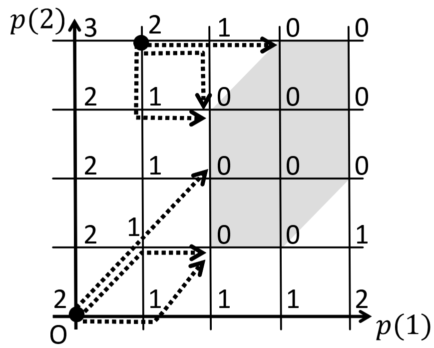

Figure 1: Behavior of steepest descent algorithms. The number associated with each integral lattice point shows the function value of g at that point. The shaded region shows the set of minimizers of g.

by dotted arrows and the minimizer found by the algorithm is either $(2,1)$ or $(2,2)$. We have $\mu(p^\circ) = 2$ with

$$\|p^* - p^\circ\|_\infty^+ + \|p^* - p^\circ\|_\infty^- = 2 + 0 \quad \text{for } p^* = (2,1), (2,2),$$

and

$$\hat{\mu}(p^\circ) = \|p^* - p^\circ\|_\infty = 2 \quad \text{for } p^* = (2,1), (2,2).$$

2. Examples of L$^\natural$-convex Functions and Minimization Algorithms

In this section we show some examples of L$^\natural$-convex functions arising from applications. We also point out that the minimization algorithms used in those applications can be regarded as special cases of the steepest descent algorithms for L$^\natural$-convex functions.

2.1. Hassin's Algorithm for Minimum Cost Flow Problem

For a directed graph $G = (V, E)$, nonnegative edge capacity $c(e)$, and edge cost $\gamma(e) \in \mathbb{R}$ for $e \in E$, the minimum cost flow problem is formulated as:

$$\begin{aligned}
&\text{Minimize} && \sum_{(u,v)\in E} \gamma(u,v)x(u,v) \\
&\text{subject to} && \sum_{v:(u,v)\in E} x(u,v) - \sum_{v:(v,u)\in E} x(v,u) = 0 \quad (u \in V), \\
& && 0 \le x(u,v) \le c(u,v) \quad ((u,v) \in E).
\end{aligned}$$

The dual problem is given as:

$$\begin{array}{ll} \text{Maximize} & g_{\mathrm{H}}(p) \equiv \displaystyle\sum_{(u,v)\in E} c(u,v)\min\{0, p(u)-p(v)+\gamma(u,v)\} \\ \text{subject to} & p(v) \in \mathbb{R} \quad (v \in V). \end{array}$$

We here assume that edge cost $\gamma(u,v)$ is integer-valued. Then, there exists an integral optimal solution to the dual problem, and we may assume that $p(v) \in \mathbb{Z}$ $(v \in V)$ in the dual problem.

It is known that g_{H} is an L$^\natural$-concave function (i.e., $-g_{\mathrm{H}}$ is L$^\natural$-convex) if we regard g_{H} as a function in integer vectors (see [14]). In fact, g_{H} is an L-concave function since $g_{\mathrm{H}}(p+\mathbf{1}) = g_{\mathrm{H}}(p)$ $(\forall p \in \mathbb{Z}^V)$ (see Remark 1.5 for the definition of L-convex function).

Hassin's algorithm in [10] can be seen as an application of algorithm SteepestDescentUp to the L-convex function $-g_{\mathrm{H}}$; see [16] for details. We also mention that the algorithm by Chung and Tcha [5] for the minimum-cost submodular flow problem, which is a generalization of Hassin's algorithm, can also be seen as a special implementation of algorithm SteepestDescentUp.

2.2. Discrete Optimization Approach in Computer Vision

Given an undirected graph $G = (V, E)$ and univariate convex functions $\varphi_u : \mathbb{Z} \to \mathbb{R} \cup \{+\infty\}$ $(u \in V)$ and $\psi_{uv} : \mathbb{Z} \to \mathbb{R} \cup \{+\infty\}$ $((u,v) \in E)$, consider the following optimization problem:

$$\begin{array}{ll} \text{(P): Minimize} & g_{\mathrm{CV}}(p) \equiv \displaystyle\sum_{u\in V} \varphi_u(p(u)) + \sum_{(u,v)\in E} \psi_{uv}(p(v)-p(u)) \\ \text{subject to} & p \in \mathbb{Z}^V. \end{array}$$

It is known that the objective function g_{CV} of this problem is an L$^\natural$-convex function; moreover, if g_{CV} does not contain the term $\sum_{u\in V} \varphi_u(p(u))$, then it is an L-convex function (see [13, 14]).

The problem (P) arises in many applications in computer vision such as panoramic image stitching [20], image restoration [4], minimization of total variation [7], and phase unwrapping in SAR images [2]. In such applications, the node set V of the undirected graph $G = (V, E)$ usually corresponds to the set of pixels in a given image, and variable $p(u)$ is the "label" of pixel $u \in V$ that represents disparity, intensity, etc. Functions φ_u encode unary data penalty functions, and ψ_{uv} are pairwise interaction potentials. The objective function of (P) is often derived in the context of Markov random fields [8]; a minimizer of the function g_{CV} corresponds to a maximum a-posteriori labeling.

There have been proposed many algorithms for (P) in computer vision (see, e.g., [2, 13]). Among them, the primal algorithm of Kolmogorov and Shioura [13] can be seen as an application of algorithm SteepestDescent to (P), while the algorithm of Bioucas-Dias and Valadão [2] is an application of SteepestDescentUp to a special case of (P) where the objective function does not contain the term $\sum_{u \in V} \varphi_u(p(u))$.

2.3. Iterative Auction in Mathematical Economics

In an auction, we want to find "good" prices for items to be allocated to bidders. An algorithm for computing such "good" prices, called ascending auction [1], can be seen as a special implementation of algorithm SteepestDescentUp.

We consider an auction market with n types of items or goods, denoted by $N = \{1, 2, \ldots, n\}$, and m bidders, denoted by $M = \{1, 2, \ldots, m\}$. Each bidder $j \in M$ has his valuation function $f_j : 2^N \to \mathbb{R}$ with the value $f_j(X)$ representing the degree of satisfaction for an item set $X \subseteq N$. We assume that each f_j satisfies the so-called "gross-substitutes" condition, which is a natural assumption for valuation functions (see [1, 9, 12] for the precise definition). We also assume that each f_j is an integer-valued function. An allocation of items is defined as a family of item sets $X_1, X_2, \ldots, X_m$ satisfying $X_j \cap X_k = \emptyset$ if $j \neq k$ and $\bigcup_{j \in M} X_j = N$.

Given a price vector $p \in \mathbb{R}^n$, each bidder $j \in M$ wants to have an item set X which maximizes the value $f_j(X) - p(X)$, where $p(X) = \sum_{i \in X} p(i)$. On the other hand, the auctioneer wants to find a price vector under which all items are sold completely. Hence, all of the auctioneer and bidders are happy if we can find a pair of a price vector p^* and an allocation $X_1^*, X_2^*, \ldots, X_m^*$ satisfying the condition

$$X_j^* \in \arg\max\{f_j(X) - p(X) \mid X \subseteq N\} \qquad (j \in M).$$

Such a pair is called a Walrasian equilibrium (see, e.g., [3, 6]).

In the auction literature an algorithm called the iterative auction (or dynamic auction, Walrasian tâtonnement process, etc.) is often used to find an equilibrium [3, 6]. An iterative auction finds an equilibrium price vector by iteratively updating a current price vector p. The most natural and popular iterative auction is the ascending auction, in which the current price vector is increased monotonically. The ascending auction is a natural generalization of the classical English auction for a single item, and known to have various nice properties (see, e.g., [3, 6]); in particular, it is natural from the economic point of view, and easy to understand and implement.

The ascending auction presented in Ausubel [1] uses a function defined by

$$L(p) = \sum_{j=1}^{m} \max\{f_j(X) - p(X) \mid X \subseteq N\} + p(N) \quad (p \in \mathbb{R}^n),$$

which is called the Lyapunov function. Under the assumption that each f_j satisfies the gross-substitutes condition, p^* is an equilibrium price vector if and only if it is a minimizer of the Lyapunov function, and there exists an integral minimizer $p^* \in \mathbb{Z}^n$ of the Lyapunov function. Based on this fact, the ascending auction in [1] tries to find a minimizer of the Lyapunov function.

It can be shown that the Lyapunov function L is an $\mathrm{L}^\natural$-concave function if it is regarded as a function in integer vectors, which follows from the conjugacy results in discrete convex analysis and the assumption that each f_j satisfies the gross-substitutes condition (see, e.g., [14, 18]). Moreover, it is observed in [18] that the ascending auction in [1] can be seen as an application of algorithm SteepestDescentUp to the function $-L$.

Similarly to the ascending auction, an algorithm called the descending auction is proposed in [1], in which the current price vector is decreased monotonically. The descending auction is a natural generalization of the well-known Dutch auction for a single item. It is also observed in [18] that the descending auction in [1] can be seen as an application of algorithm SteepestDescentDown to the function $-L$.

3. Proofs

In this section, we prove Theorems 1.2 and 1.3. The key fact used in our proofs is the following property of $\mathrm{L}^\natural$-convex functions. For $p \in \mathbb{Z}^n$, we denote $\mathrm{supp}^+(p) = \{i \in N \mid p(i) > 0\}$.

Lemma 3.1 ([14, Theorem 7.7]). *Let $g : \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ be an $L^\natural$-convex function. For every $p, q \in \mathrm{dom}\, g$ with $\mathrm{supp}^+(p - q) \neq \emptyset$, it holds that*

$$g(p) + g(q) \geq g(p - e_Y) + g(q + e_Y),$$

where $Y = \arg\max_{i \in N}\{p(i) - q(i)\}$.

Proof. We reproduce the proof of this lemma in [14, Theorem 7.7] for its fundamental importance. Let $p, q \in \mathrm{dom}\, g$, and suppose that $\mathrm{supp}^+(p-q) \neq \emptyset$. By (1), we have

$$g(q) + g(p) \geq g((q + \lambda\mathbf{1}) \wedge p) + g(q \vee (p - \lambda\mathbf{1})) \tag{3}$$

for every nonnegative $\lambda \in \mathbb{Z}_+$. Putting $\lambda = \max_{i \in N}\{p(i) - q(i)\} - 1$, we have

$$\lambda \geq 0, \quad (q + \lambda \mathbf{1}) \wedge p = p - e_Y, \quad q \vee (p - \lambda \mathbf{1}) = q + e_Y.$$

Hence, the inequality $g(p)+g(q) \geq g(p-e_Y)+g(q+e_Y)$ follows from (3). □

3.1. Proof of Theorem 1.2

The bound $\mu(p^\circ)+1$ for the number of iterations in algorithm Steepest-Descent can be obtained by repeated application of the following lemma.

Lemma 3.2. *Let $p \in \mathbb{Z}^n$ be a vector with $\mu(p) > 0$. Suppose that $\sigma \in \{+1, -1\}$ and $X \subseteq N$ minimize the value $g(p + \sigma e_X)$. Then, $\mu(p + \sigma e_X) = \mu(p) - 1$.*

Below we give a proof of Lemma 3.2. We consider the case with $\sigma = +1$ since the other case with $\sigma = -1$ can be dealt with similarly.

We first show the inequality $\mu(p+e_X) \geq \mu(p) - 1$. For every $d \in \mathbb{Z}^n$ and $Y \subseteq N$, we have

$$\|d - e_Y\|^+_\infty \geq \|d\|^+_\infty - 1, \qquad \|d - e_Y\|^-_\infty \geq \|d\|^-_\infty.$$

Hence, it holds that

$$\begin{aligned} \mu(p + e_X) &= \min\{\|q - (p + e_X)\|^+_\infty + \|q - (p + e_X)\|^-_\infty \mid q \in \arg\min g\} \\ &\geq \min\{\|q - p\|^+_\infty + \|q - p\|^-_\infty \mid q \in \arg\min g\} - 1 \\ &= \mu(p) - 1. \end{aligned}$$

In the following, we prove the reverse inequality

$$\mu(p + e_X) \leq \mu(p) - 1. \tag{4}$$

The outline of the proof is as follows. We denote

$$\begin{aligned} &S = \{q \in \arg\min g \mid \|q - p\|^+_\infty + \|q - p\|^-_\infty = \mu(p)\}, \\ &\xi = \max\{\|q - p\|^+_\infty \mid q \in S\}. \end{aligned}$$

Let q^* be a vector in S with $\|q^* - p\|^+_\infty = \xi$, and assume that q^* is a minimal vector among all such vectors. We first show that

$$\xi = \|q^* - p\|^+_\infty > 0. \tag{5}$$

Note that this condition is equivalent to $\mathrm{supp}^+(q^* - p) \neq \emptyset$. Using this, we then prove that

$$\arg\max_{i \in N}\{q^*(i) - p(i)\} \subseteq X. \tag{6}$$

By using (5) and (6), we finally derive the inequality (4).

[Proof of (5)] Assume, to the contrary, that $\xi = \|q^* - p\|^+_\infty = 0$, i.e., $q^* \leq p$ holds. This assumption implies $\|q^* - p\|^-_\infty > 0$ since $\mu(p) > 0$.

By L$^\natural$-convexity of g in (1), we have

$$g(p + e_X) + g(q^*) \geq g((p + e_X - \mathbf{1}) \vee q^*) + g((p + e_X) \wedge (q^* + \mathbf{1})). \quad (7)$$

Let $Y = \{i \in N \mid q^*(i) - p(i) = 0\}$, which may be the empty set. Since $q^* \leq p$, we have

$$(p + e_X - \mathbf{1}) \vee q^* = p - e_{N\setminus(X\cup Y)}, \quad (p + e_X) \wedge (q^* + \mathbf{1}) = q^* + e_{(N\setminus Y)\cup X},$$

which, together with (7), implies

$$g(p + e_X) + g(q^*) \geq g(p - e_{N\setminus(X\cup Y)}) + g(q^* + e_{(N\setminus Y)\cup X}). \quad (8)$$

By the choice of $\sigma = +1$ and X, we have $g(p + e_X) \leq g(p - e_{N\setminus(X\cup Y)})$. From this and (8) follows that $g(q^*) \geq g(q^* + e_{(N\setminus Y)\cup X})$, implying that $q^* + e_{(N\setminus Y)\cup X} \in \arg\min g$. By $q^* \leq p$ and the definition of Y, we have

$$\|(q^* + e_{(N\setminus Y)\cup X}) - p\|^-_\infty = \max_{i\in N\setminus Y}\{p(i) - (q^*(i) + 1)\} = \|q^* - p\|^-_\infty - 1 \quad (9)$$

since $\|q^* - p\|^-_\infty > 0$. We also have

$$\|(q^* + e_{(N\setminus Y)\cup X}) - p\|^+_\infty \leq 1 \ = \ \|q^* - p\|^+_\infty + 1 \quad (10)$$

since $\|q^* - p\|^+_\infty = 0$. From (9) and (10) follows that

$$\begin{aligned} \mu(p) &\leq \|(q^* + e_{(N\setminus Y)\cup X}) - p\|^+_\infty + \|(q^* + e_{(N\setminus Y)\cup X}) - p\|^-_\infty \\ &\leq \|q^* - p\|^+_\infty + \|q^* - p\|^-_\infty = \mu(p), \end{aligned} \quad (11)$$

where the first inequality is by the definition of $\mu(p)$. Hence, the inequality (10) and the first inequality in (11) must hold with equality, i.e., we have $q^* + e_{(N\setminus Y)\cup X} \in S$ and

$$\|(q^* + e_{(N\setminus Y)\cup X}) - p\|^+_\infty = \|q^* - p\|^+_\infty + 1 > \|q^* - p\|^+_\infty = \xi.$$

This, however, is a contradiction to the definition of ξ. Hence, (5) holds.

[Proof of (6)] We denote

$$A = \arg\max_{i\in N}\{q^*(i) - p(i)\}.$$

Then, (6) is simply rewritten as $A \subseteq X$.

Assume, to the contrary, that $A \backslash X \neq \emptyset$ holds. We claim that $q^* - e_{A\backslash X} \in \arg\min g$. By (5), it holds that $\xi = \|q^* - p\|_\infty^+ > 0$. Therefore, we have $A \subseteq \mathrm{supp}^+(q^* - p)$, from which follows that

$$\mathrm{supp}^+(q^* - (p + e_X)) \supseteq A \setminus X \neq \emptyset.$$

Since $A \setminus X \neq \emptyset$, we also have

$$\arg\max_{i \in N} \{q^*(i) - (p + e_X)(i)\} = A \setminus X.$$

Hence, Lemma 3.1 implies that

$$\begin{aligned} g(q^*) + g(p + e_X) &\geq g(q^* - e_{A\backslash X}) + g(p + e_X + e_{A\backslash X}) \\ &= g(q^* - e_{A\backslash X}) + g(p + e_{X \cup A}). \end{aligned} \tag{12}$$

By the choice of X, we have $g(p + e_X) \leq g(p + e_{X \cup A})$, which, together with (12), implies that $g(q^*) \geq g(q^* - e_{A\backslash X})$, i.e., $q^* - e_{A\backslash X} \in \arg\min g$.

Since $A \setminus X \subseteq A \subseteq \mathrm{supp}^+(q^* - p)$, we have

$$\begin{aligned} &\|(q^* - e_{A\backslash X}) - p\|_\infty^+ \leq \|q^* - p\|_\infty^+ = \xi, \qquad (13) \\ &\|(q^* - e_{A\backslash X}) - p\|_\infty^- = \|q^* - p\|_\infty^-, \end{aligned}$$

from which follows that

$$\begin{aligned} \mu(p) &\leq \|(q^* - e_{A\backslash X}) - p\|_\infty^+ + \|(q^* - e_{A\backslash X}) - p\|_\infty^- \\ &\leq \|q^* - p\|_\infty^+ + \|q^* - p\|_\infty^- = \mu(p), \end{aligned} \tag{14}$$

where the first inequality is by the definition of $\mu(p)$. Hence, the inequality (13) and the first inequality in (14) must hold with equality. Hence, the vector $q^* - e_{A\backslash X}$ belongs to S with $\|(q^* - e_{A\backslash X}) - p\|_\infty^+ = \xi$, a contradiction to the minimality of q^*.

[Proof of (4)] To show the inequality (4), we first consider the case with $\min_{i \in N}\{q^*(i) - p(i)\} > 0$. Since $q^*(i) > p(i)$ for all $i \in N$, it holds that $q^* \geq p + e_X$. Therefore, we have

$$\|q^* - (p + e_X)\|_\infty^- = 0 = \|q^* - p\|_\infty^-.$$

By (6), it holds that

$$\|q^* - (p + e_X)\|_\infty^+ = \|q^* - p\|_\infty^+ - 1.$$

Therefore, it follows that

$$\begin{aligned}\mu(p+e_X) &\leq \|q^* - (p+e_X)\|^+_\infty + \|q^* - (p+e_X)\|^-_\infty \\ &= (\|q^* - p\|^+_\infty - 1) + \|q^* - p\|^-_\infty = \mu(p) - 1.\end{aligned}$$

We next consider the remaining case where $\min_{i\in N}\{q^*(i) - p(i)\} \leq 0$. We denote

$$B = \arg\min_{i\in N}\{q^*(i) - p(i)\}.$$

We claim that $q^* + e_{B\cap X} \in \arg\min g$ holds. If $B \cap X = \emptyset$, then $q^* + e_{B\cap X} = q^* \in \arg\min g$. Hence, we assume $B \cap X \neq \emptyset$. Since

$$\max_{i\in N}\{p(i) - q^*(i)\} \geq 0, \qquad B = \arg\max_{i\in N}\{p(i) - q^*(i)\},$$

it holds that

$$\mathrm{supp}^+((p+e_X) - q^*) \supseteq B \cap X \neq \emptyset, \quad \arg\max_{i\in N}\{(p+e_X)(i) - q^*(i)\} = B \cap X.$$

It follows from Lemma 3.1 that

$$\begin{aligned}g(p+e_X) + g(q^*) &\geq g(p + e_X - e_{B\cap X}) + g(q^* + e_{B\cap X}) \\ &= g(p + e_{X\setminus B}) + g(q^* + e_{B\cap X}). \qquad (15)\end{aligned}$$

By the choice of X, we have $g(p+e_X) \leq g(p+e_{X\setminus B})$, which, together with (15), implies that $g(q^*) \geq g(q^* + e_{B\cap X})$, i.e., $q^* + e_{B\cap X} \in \arg\min g$.

Since $\min_{i\in N}\{q^*(i) - p(i)\} \leq 0 < \max_{i\in N}\{q^*(i) - p(i)\}$ by the assumption and (5), we have $A \cap B = \emptyset$, which, together with (6), implies $A \subseteq X \setminus B$. Hence, it holds that

$$\|(q^* + e_{B\cap X}) - (p+e_X)\|^+_\infty = \|q^* - p - e_{X\setminus B}\|^+_\infty = \|q^* - p\|^+_\infty - 1.$$

We also have

$$\|(q^* + e_{B\cap X}) - (p+e_X)\|^-_\infty = \|q^* - p - e_{X\setminus B}\|^-_\infty = \|q^* - p\|^-_\infty,$$

where the second equality follows from the definition of B. Hence, it holds that

$$\begin{aligned}\mu(p+e_X) &\leq \|(q^* + e_{B\cap X}) - (p+e_X)\|^+_\infty + \|(q^* + e_{B\cap X}) - (p+e_X)\|^-_\infty \\ &= (\|q^* - p\|^+_\infty - 1) + \|q^* - p\|^-_\infty \;=\; \mu(p) - 1.\end{aligned}$$

This concludes the proof of Lemma 3.2 (and also of Theorem 1.2).

3.2. Proof of Theorem 1.3

The proof of Theorem 1.3 is quite similar to and simpler than that of Theorem 1.2. Theorem 1.3 can be proved by using the following property repeatedly.

Lemma 3.3. *Let $p \in \mathbb{Z}^n$ be a vector with $\hat{\mu}(p) > 0$, and $X \subseteq N$ be a set that minimizes the value of $g(p + e_X)$. Then, $\hat{\mu}(p + e_X) = \hat{\mu}(p) - 1$.*

Below we give a proof of Lemma 3.3. The inequality $\hat{\mu}(p+e_X) \geq \hat{\mu}(p)-1$ can be shown as follows. By the triangle inequality, we have $\|q - (p + e_X)\|_\infty \geq \|q - p\|_\infty - 1$ for every $q \in \mathbb{Z}^n$. Taking the minimum over all $q \in \arg\min g$ with $q \geq p + e_X$, we obtain

$$\begin{aligned}\hat{\mu}(p + e_X) &\geq \min\{\|q - p\|_\infty \mid q \in \arg\min g,\ q \geq p + e_X\} - 1 \\ &\geq \min\{\|q - p\|_\infty \mid q \in \arg\min g,\ q \geq p\} - 1 = \hat{\mu}(p) - 1.\end{aligned}$$

In the following, we show the reverse inequality:

$$\hat{\mu}(p + e_X) \leq \hat{\mu}(p) - 1. \tag{16}$$

Let p^* be a vector such that $p^* \in \arg\min g$, $p^* \geq p$, and $\|p^* - p\|_\infty = \hat{\mu}(p)$, and assume that p^* is minimal among all such vectors. We denote

$$A = \arg\max_{i \in N}\{p^*(i) - p(i)\}.$$

We have $p^* \neq p$ and $\max_{i \in N}\{p^*(i) - p(i)\} > 0$ since $\|p^* - p\|_\infty = \hat{\mu}(p) > 0$ and $p^* \geq p$.

We claim that

$$A \subseteq X. \tag{17}$$

Assume, to the contrary, that $A \setminus X \neq \emptyset$ holds. Since $A \subseteq \mathrm{supp}^+(p^* - p)$, we have

$$\mathrm{supp}^+(p^* - (p + e_X)) \supseteq A \setminus X \neq \emptyset.$$

We also have

$$\arg\max_{i \in N}\{p^*(i) - (p + e_X)(i)\} = A \setminus X.$$

Hence, Lemma 3.1 implies that

$$\begin{aligned}g(p^*) + g(p + e_X) \quad &\geq \quad g(p^* - e_{A \setminus X}) + g(p + e_X + e_{A \setminus X}) \\ &= \quad g(p^* - e_{A \setminus X}) + g(p + e_{X \cup A}).\end{aligned} \tag{18}$$

By the choice of X, we have $g(p + e_X) \leq g(p + e_{X \cup A})$. This inequality, together with (18), implies that $g(p^*) \geq g(p^* - e_{A \setminus X})$, i.e., $p^* - e_{A \setminus X} \in \arg\min g$ holds. This, however, is a contradiction to the choice of p^* since

$$p^* \geq p^* - e_{A \setminus X} \geq p, \qquad \|(p^* - e_{A \setminus X}) - p\|_\infty \leq \|p^* - p\|_\infty = \hat{\mu}(p).$$

Hence, we have (17).

We now prove the inequality (16). Suppose first that the condition $p^* \geq p + e_X$ holds. Then, we have

$$\hat{\mu}(p + e_X) \leq \|p^* - (p + e_X)\|_\infty = \|p^* - p\|_\infty - 1 = \hat{\mu}(p) - 1,$$

where the first equality is by (17).

We next consider the case where the condition $p^* \geq p + e_X$ fails. Then, $B \cap X \neq \emptyset$ for $B = \{i \in N \mid p^*(i) = p(i)\}$. Since $\max_{i \in N}\{p^*(i) - p(i)\} > 0$, we have $A \cap B = \emptyset$, which, together with (17), implies $A \subseteq X \setminus B$. Since $p^* \geq p$, we have

$$p^*(i) = p(i) \quad (\forall i \in B), \qquad p^*(i) > p(i) \quad (\forall i \in N \setminus B), \tag{19}$$

from which $p^* + e_{B \cap X} \geq p + e_X$ follows. As shown below, we have $p^* + e_{B \cap X} \in \arg\min g$. Hence, it holds that

$$\begin{aligned} \hat{\mu}(p + e_X) \leq \|(p^* + e_{B \cap X}) - (p + e_X)\|_\infty &= \|p^* - p - e_{X \setminus B}\|_\infty \\ &= \|p^* - p\|_\infty - 1 = \hat{\mu}(p) - 1, \end{aligned}$$

where the second equality is by $A \subseteq X \setminus B$, (19), and the definition of A.

We now show that $p^* + e_{B \cap X} \in \arg\min g$ holds. The condition (19) implies

$$\mathrm{supp}^+((p + e_X) - p^*) = \arg\max_{i \in N}\{(p + e_X)(i) - p^*(i)\} = B \cap X.$$

Hence, it follows from Lemma 3.1 that

$$\begin{aligned} g(p + e_X) + g(p^*) &\geq g(p + e_X - e_{B \cap X}) + g(p^* + e_{B \cap X}) \\ &= g(p + e_{X \setminus B}) + g(p^* + e_{B \cap X}). \end{aligned} \tag{20}$$

By the choice of X, we have $g(p + e_X) \leq g(p + e_{X \setminus B})$, which, together with (20), implies that $g(p^*) \geq g(p^* + e_{B \cap X})$, i.e., $p^* + e_{B \cap X}$ is a minimizer of g.

This concludes the proof of Lemma 3.3 (and also of Theorem 1.3).

4. Conclusion

The concept of L$^\natural$-convexity is generalized to polyhedral convex functions. A steepest descent algorithm similar to SteepestDescent works for the minimization of a polyhedral L$^\natural$-convex function, and has a similar property as SteepestDescent; in particular, the trajectory of a vector p generated by the algorithm is the "shortest" path between the initial vector and a minimizer (see [17] for details).

Part of the results in this paper is presented (without proofs) in the extended abstract [18] in Proceedings of 24th International Symposium on Algorithms and Computation (ISAAC 2013). This research is supported by KAKENHI (21360045, 21740060, 24500002) and the Aihara Project, the FIRST program from JSPS.

References

[1] L. M. Ausubel, An efficient dynamic auction for heterogeneous commodities, Amer. Econ. Rev. 96 (2006) 602–629.

[2] J. M. Bioucas-Dias, G. Valadão, Phase unwrapping via graph cuts, IEEE Trans. Image Process. 16 (2007) 698–709.

[3] L. Blumrosen, N. Nisan, Combinatorial auction. in: N. Nisan, et al. (eds.) Algorithmic Game Theory, Cambridge Univ. Press, Cambridge, 2007, pp. 267–299.

[4] Y. B. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2001) 1222–1239.

[5] N.-k. Chung, D.-w. Tcha, A dual algorithm for submodular flow problems, Oper. Res. Lett. 10 (1991) 489–495.

[6] P. Cramton, Y. Shoham, R. Steinberg, Combinatorial Auctions. MIT Press, Cambridge, MA, 2006.

[7] J. Darbon, M. Sigelle, Image restoration with discrete constrained total variation, part I: fast and exact optimization, J. Math. Imaging Vision 26 (2006) 261–276.

[8] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. 6 (1984) 721–741.

[9] F. Gul, E. Stacchetti, The English auction with differentiated commodities, J. Economic Theory 92 (2000) 66–95.

[10] R. Hassin, The minimum cost flow problem: a unifying approach to dual algorithms and a new tree-search algorithm, Math. Program. 25 (1983) 228–239.

[11] W.T. Huh, G. Janakiraman, On the optimal policy structure in serial inventory systems with lost sales, Oper. Res. 58 (2010) 486–491.

[12] A. S. Kelso, V. P. Crawford, Job matching, coalition formation and gross substitutes, Econometrica 50 (1982) 1483–1504.

[13] V. Kolmogorov, A. Shioura, New algorithms for convex cost tension problem with application to computer vision, Discrete Optim. 6 (2009) 378–393.

[14] K. Murota, Discrete Convex Analysis, SIAM, Philadelphia, PA, 2003.

[15] K. Murota, On steepest descent algorithms for discrete convex functions, SIAM J. Optim. 14 (2003) 699–707.

[16] K. Murota, A. Shioura, Dijkstra's algorithm and L-concave function maximization, Math. Program. (2013) published online.

[17] K. Murota, A. Shioura, On some properties of polyhedral L-concave maximization algorithm, Technical report METR 2013-18, Department of Mathematical Informatics, University of Tokyo (2013).

[18] K. Murota, A. Shioura, Z. Yang, Computing a Walrasian equilibrium in iterative auctions with multiple differentiated items, in: Proc. 24th International Symposium on Algorithms and Computation (ISAAC 2013), to appear.

[19] P. Zipkin, On the structure of lost-sales inventory models, Oper. Res. 56 (2008) 937–944.

[20] A. Zomet, A. Levin, S. Peleg, Y. Weiss, Seamless image stitching by minimizing false edges, IEEE Trans. Image Process. 15 (2006) 969–977.