Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

Akiyoshi Shioura[1]

Note on L$^\natural$-convex Function Minimization Algorithms: Comparison of Murota's and Kolmogorov's Algorithms

[1]Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustr. 7, 14195 Berlin, Germany. On leave from Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan. e-mail: shioura@dais.is.tohoku.ac.jp

ZIB-Report 06-03 (Januar 2006)

Note on L$^\natural$-convex Function Minimization Algorithms: Comparison of Murota's and Kolmogorov's Algorithms

Akiyoshi Shioura[†]

Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7, 14195 Berlin, Germany
shioura@dais.is.tohoku.ac.jp

January, 2006

Abstract

The concept of L$^\natural$-convexity is introduced by Fujishige–Murota (2000) as a discrete convexity for functions defined over the integer lattice. The main aim of this note is to understand the difference of the two algorithms for L$^\natural$-convex function minimization: Murota's steepest descent algorithm (2003) and Kolmogorov's primal algorithm (2005).

1 Introduction

The concept of L$^\natural$-convexity is introduced by Fujishige–Murota [2] as a discrete convexity for functions defined over the integer lattice. This is a variant of L-convexity due to Murota [5], and later turned out to be equivalent to integral convexity by Favati–Tardella [1]. See [6] for details.

The main aim of this note is to understand the difference of the two algorithms for L$^\natural$-convex function minimization: Murota's steepest descent algorithm [7] and Kolmogorov's primal algorithm [4].

1.1 L$^\natural$-convex Functions

Let V be a nonempty finite set. A function $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, g \neq \emptyset$ is called L-convex if it satisfies the following properties:

(LF1) $g(p) + g(q) \geq g(p \wedge q) + g(p \vee q) \qquad (\forall p, q \in \mathrm{dom}\, g)$,
(LF2) $\exists r \in \mathbf{R}$ such that $g(p+\lambda \mathbf{1}) = g(p)+\lambda r$ $(\forall p \in \mathrm{dom}\, g,\ \forall \lambda \in \mathbf{Z})$,

[†]On leave from Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan.

where $\mathrm{dom}\, g = \{p \in \mathbf{Z}^V \mid g(p) < +\infty\}$, the vectors $p \wedge q, p \vee q \in \mathbf{Z}^V$ are defined by

$$(p \wedge q)(v) = \min\{p(v), q(v)\}, \quad (p \vee q)(v) = \max\{p(v), q(v)\} \quad (v \in V),$$

and $\mathbf{1} \in \mathbf{Z}^V$ is the vector with all components equal to one. A function $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, g \neq \emptyset$ is called L$^\natural$-convex if the function $\tilde{g} : \mathbf{Z} \times \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ defined by

$$\tilde{g}(p_0, p) = g(p - p_0\mathbf{1}) \quad (p_0 \in \mathbf{Z}, p \in \mathbf{Z}^V) \tag{1}$$

is L-convex. The class of L$^\natural$-convex functions contains that of L-convex functions as a proper subclass.

L$^\natural$-convex functions can be characterized by the following property:

Theorem 1 ([6]). *A function $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, g \neq \emptyset$ is L$^\natural$-convex if and only if for all $p, q \in \mathbf{Z}^V$ with $\mathrm{supp}^+(p - q) \neq \emptyset$, we have*

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

where $Z = \arg\max\{p(v) - q(v) \mid v \in V\}$.

An L$^\natural$-convex function restricted to the integer interval has the unique minimal and maximal minimizers.

Proposition 2. *Let $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an L$^\natural$-convex function, and $a, b \in \mathbf{Z}^V$ be vectors with $\{p \in \mathrm{dom}\, g \mid a(v) \leq p(v) \leq b(v)\ (v \in V)\} \neq \emptyset$. Then, the set $\arg\min\{g(p) \mid a(v) \leq p(v) \leq b(v)\ (v \in V)\}$ contains the unique minimal and maximal minimizers.*

See [6] for more accounts on L$^\natural$-convex functions.

1.2 Murota's and Kolmogorov's Algorithms

To the end of this note, we assume that $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is an L$^\natural$-convex function with $\arg\min g \neq \emptyset$. Murota's steepest descent algorithm [7] is described as follows:

> **Murota's steepest descent algorithm**
> S0: Find a vector $p \in \mathrm{dom}\, g$.
> S1: Set $\varepsilon \in \{1, -1\}$ and $X \subseteq V$ as follows.
> S1-1: Let X^+ be the minimal minimizer of $\rho_p^+(X) = g(p+\chi_X) - g(p)$.
> S1-2: Let X^- be the maximal minimizer of $\rho_p^-(X) = g(p-\chi_X) - g(p)$.
> S1-3: If $\min \rho_p^+ \leq \min \rho_p^-$ then set $(\varepsilon, X) = (1, X^+)$;
> otherwise set $(\varepsilon, X) = (-1, X^-)$.
> S2: If $g(p) \leq g(p + \varepsilon\chi_X)$, then stop ($p$ is a minimizer of g).
> S3: Set $p := p + \varepsilon\chi_X$ and go to S1.

A minimizer of a submodular set function can be found in strongly polynomial time by the existing algorithms [3, 8]. In particular, a maximal/minimal element in the set of minimizers can be found without extra running time by Iwata–Fleischer–Fujishige's algorithm [3].

On the other hand, Kolmogorov's primal algorithm [4] is obtained by replacing Step S1 of Murota's algorithm with the following:

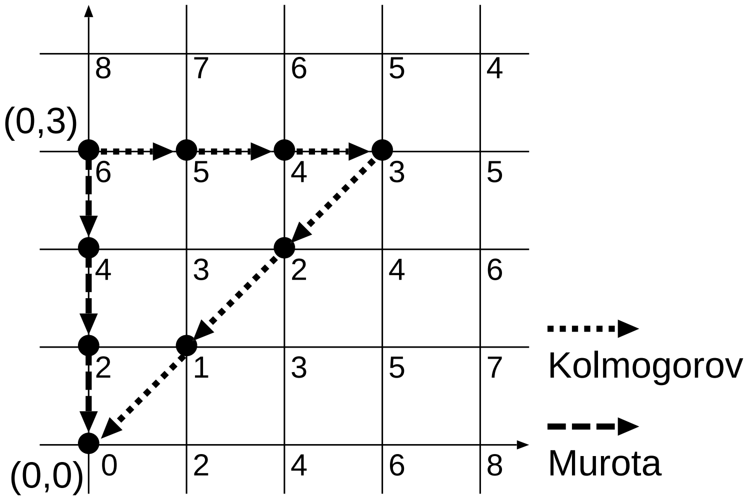

Figure 1: Behavior of Kolmogorov's and Murota's algorithms for g_2 with the initial vector $(0, 3)$. Each value associated with each integral lattice point shows the function value of g_2 at that point.

Kolmogorov's primal algorithm
S1: Set $\varepsilon \in \{1, -1\}$ and $X \subseteq V$ as follows.
S1-1: Let X^+ be any minimizer of $\rho_p^+(X)$.
S1-2: Let X^- be any minimizer of $\rho_p^-(X)$.
S1-3: If $\rho_p^+(X^+) = 0$ then set $(\varepsilon, X) = (-1, X^-)$;
if $\rho_p^-(X^-) = 0$ then set $(\varepsilon, X) = (1, X^+)$;
otherwise choose either of $(1, X^+)$ and $(-1, X^-)$ arbitrarily as (ε, X).

Kolmogorov's algorithm has more flexibility in the choice of a next step (ε, X) than Murota's algorithm, and therefore Murota's algorithm can be seen as a specialized implementation of Kolmogorov's algorithm.

Kolmogorov [4] has shown that the number of iterations required by his algorithm (and hence Murota's) is bounded by $2K_g^\infty$, where

$$K_g^\infty = \max\{||p - q||_\infty \mid p, q \in \mathrm{dom}\, g\}.$$

Kolmogorov's algorithm, however, may require more iterations than Murota's, as shown in the following example.

Let $g_2 : \mathbf{Z}^2 \to \mathbf{R} \cup \{+\infty\}$ be an L$^\natural$-convex function defined as

$$g_2(p_1, p_2) = \begin{cases} \max\{2p_1 - p_2, -p_1 + 2p_2\} & ((p_1, p_2) \in \mathbf{Z}_+^2), \\ +\infty & (\text{otherwise}). \end{cases}$$

Note that $(0, 0)$ is the unique minimizer of g_2. Let $(0, k)$ be the initial vector of the algorithms. Then, Kolmogorov's algorithm may possibly generate the following sequence of vectors with length $2k + 1$:

(p_1, p_2)	$(0, k)$	$(1, k)$	$\cdots$	$(k-1, k)$	(k, k)	$(k-1, k-1)$	$\cdots$	$(1, 1)$	$(0, 0)$
$g_2(p_1, p_2)$	$2k$	$2k-1$	$\cdots$	$k+1$	k	$k-1$	$\cdots$	1	0

On the other hand, Murota's algorithm generates the following sequence of length $k+1$:

(p_1, p_2)	$(0,k)$	$(0,k-1)$	$\cdots$	$(0,1)$	$(0,0)$
$g_2(p_1,p_2)$	$2k$	$2(k-1)$	$\cdots$	2	0

which is shorter than the one by Kolmogorov's algorithm (see Figure 1).

More generally, we consider an L$^\natural$-convex function $g_n : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined as

$$g_n(p) = \begin{cases} (n+1)\max\{p_i \mid i = 1,\dots,n\} - \displaystyle\sum_{j=1}^{n} p_j & (p \in \mathbf{Z}^n_+), \\ +\infty & (\text{otherwise}) \end{cases}$$

where n is a positive integer with $n \geq 2$. If we apply the two algorithms to g_n with the initial vector $(0,\dots,0,k)$, then Kolmogorov's algorithm may generate the following sequence of length $2k+1$:

$$(0,\dots,0,k), (1,\dots,1,k), (2,\dots,2,k), \dots, (k-1,\dots,k-1,k), (k,\dots,k,k),$$
$$(k-1,\dots,k-1,k-1), (k-2,\dots,k-2,k-2), \dots, (1,\dots,1,1), (0,\dots,0,0),$$

while Murota's algorithm generates the following sequence of length $k+1$:

$$(0,\dots,0,k), (0,\dots,0,k-1), (0,\dots,0,k-2), \dots, (0,\dots,0,1), (0,\dots,0,0).$$

It should be noted that both algorithms require $2K^\infty_g$ iterations in the worst case (see [4] for such an example), i.e., the order of the worst-case bound is the same.

2 Analysis of the Number of Iterations

In this section we analyze the number of iterations required by Kolmogorov's and Murota's algorithms.

2.1 Analysis of Kolmogorov's Algorithm

The analysis given in this section is essentially the same as the one in [4]. We present the result in [4] in a way consistent with the analysis for Murota's algorithm given in Section 2.2.

To analyze the number of iterations required by Kolmogorov's algorithm, we define values $\beta^+(p)$ and $\beta^-(p)$ for each vector $p \in \mathrm{dom}\, g$ as follows:

$$\begin{aligned} \beta^+(p) &= \min\{||q-p||_\infty \mid q \in \arg\min\{g(q') \mid q' \geq p\}\}, \\ \beta^-(p) &= \min\{||q-p||_\infty \mid q \in \arg\min\{g(q') \mid q' \leq p\}\}. \end{aligned}$$

The value $\beta^+(p)$ is the distance between p and the unique minimal minimizer of g in the region $\{q' \in \mathbf{Z}^V \mid q' \geq p\}$; $\beta^-(p)$ is the distance between p and the unique maximal minimizer of g in the region $\{q' \in \mathbf{Z}^V \mid q' \leq p\}$.

Proposition 3 ([4]). *For $p \in \mathrm{dom}\, g$, if $\beta^+(p) = \beta^-(p) = 0$ then $p \in \arg\min g$.*

Proof. If $\beta^+(p) = \beta^-(p) = 0$, then we have $g(p) \leq g(p+\varepsilon\chi_X)$ for all $\varepsilon \in \{1, -1\}$ and $X \subseteq V$. Hence, p is a minimizer of g since g is an $\mathrm{L}^\natural$-convex function (see, e.g., [6]). $\square$

Note that $\beta^+(p) = 0$ (resp. $\beta^-(p) = 0$) alone implies $p \in \arg\min\{g(p') \mid p' \geq p\}$ (resp. $p \in \arg\min\{g(p') \mid p' \leq p\}$), but does not imply $p \in \arg\min g$ in general.

Each iteration of Kolmogorov's algorithm increases neither of $\beta^+(p)$ nor $\beta^-(p)$ and decreases strictly at least one of $\beta^+(p)$ and $\beta^-(p)$.

Proposition 4 ([4])**.** *In Step* S1*, we have the following:*
(i) *If* $\beta^+(p) > 0$*, then* $\beta^+(p + \chi_{X^+}) = \beta^+(p) - 1$ *and* $\beta^-(p + \chi_{X^+}) \leq \beta^-(p)$*.*
(ii) *If* $\beta^-(p) > 0$*, then* $\beta^-(p - \chi_{X^-}) = \beta^-(p) - 1$ *and* $\beta^+(p - \chi_{X^-}) \leq \beta^+(p)$*.*

Proof. We prove (i) only; the claim (ii) can be shown in the same way.

[Proof of "$\beta^+(p+\chi_{X^+}) = \beta^+(p) - 1$"] We denote by $\hat{p}, \hat{q} \in \mathbf{Z}^V$ the unique minimal vectors in $\arg\min\{g(p') \mid p' \geq p\}$ and in $\arg\min\{g(p') \mid p' \geq p+\chi_{X^+}\}$, respectively. We will show that

$$\hat{q} = \hat{p} \vee (p + \chi_{X^+}), \tag{2}$$

$$Y^+ \subseteq X^+, \quad \text{where } Y^+ = \arg\max\{\hat{p}(v) - p(v) \mid v \in V\}. \tag{3}$$

Then, we have

$$\begin{aligned} \beta^+(p + \chi_{X^+}) &= ||(\hat{p} \vee (p + \chi_{X^+})) - (p + \chi_{X^+})||_\infty \\ &= ||\hat{p} - p||_\infty - 1 \ = \ \beta^+(p) - 1, \end{aligned}$$

where the first equality is by (2) and the second by (3).

We first prove (2). By the submodularity of g, we have

$$g(\hat{p}) + g(p + \chi_{X^+}) \geq g(\hat{p} \vee (p + \chi_{X^+})) + g(\hat{p} \wedge (p + \chi_{X^+})). \tag{4}$$

Since $p \leq \hat{p} \wedge (p+\chi_{X^+}) \leq p+\chi_{X^+}$ and $X^+ \in \arg\min \rho_p^+$, we have $g(p+\chi_{X^+}) \leq g(\hat{p} \wedge (p + \chi_{X^+}))$, which, together with (4), implies

$$g(\hat{p}) \geq g(\hat{p} \vee (p + \chi_{X^+})). \tag{5}$$

Since $\hat{q} \geq p + \chi_{X^+} \geq p$ and $\hat{p} \in \arg\min\{g(p') \mid p' \geq p\}$, we have

$$g(\hat{q}) \geq g(\hat{p}). \tag{6}$$

Similarly, we have

$$g(\hat{p} \vee (p + \chi_{X^+})) \geq g(\hat{q}) \tag{7}$$

since $\hat{p} \vee (p+\chi_{X^+}) \geq p+\chi_{X^+}$ and $\hat{q} \in \arg\min\{g(p') \mid p' \geq p+\chi_{X^+}\}$. It follows from (5), (6), and (7) that $g(\hat{p}) = g(\hat{q}) = g(\hat{p} \vee (p+\chi_{X^+}))$, which in turn implies

$$\hat{q} \in \arg\min\{g(p') \mid p' \geq p\}, \qquad \hat{p} \vee (p + \chi_{X^+}) \in \arg\min\{g(p') \mid p' \geq p + \chi_{X^+}\}.$$

It follows from the choices of $\hat{p}$ and $\hat{q}$ that $\hat{p} \leq \hat{q}$ and $\hat{q} \leq \hat{p} \vee (p + \chi_{X^+})$. These inequalities and $p + \chi_{X^+} \leq \hat{q}$ imply (2).

We then prove (3). Assume, to the contrary, that $Y^+ \setminus X^+ \neq \emptyset$. Put

$$Z^+ = \arg\max\{\hat{p}(v) - p(v) - \chi_{X^+}(v) \mid v \in V\} = Y^+ \setminus X^+.$$

Theorem 1 implies

$$g(\hat{p}) + g(p + \chi_{X^+}) \quad \geq \quad g(\hat{p} - \chi_{Z^+}) + g(p + \chi_{X^+} + \chi_{Z^+}).$$

Since $\chi_{X^+} + \chi_{Z^+} = \chi_{X^+ \cup Y^+}$, we have $g(p + \chi_{X^+} + \chi_{Z^+}) = g(p + \chi_{X^+ \cup Y^+}) \geq g(p + \chi_{X^+})$, where the inequality is by $X^+ \in \arg\min \rho_p^+$. Hence, we have $g(\hat{p}) \geq g(\hat{p} - \chi_{Z^+})$, a contradiction to the fact that $\hat{p}$ is the minimal minimizer in $\{p' \in \mathbf{Z}^V \mid p' \geq p\}$ since $\hat{p} - \chi_{Z^+} \geq p$.

[Proof of "$\beta^-(p + \chi_{X^+}) \leq \beta^-(p)$"] We denote by $\check{p}, \check{q} \in \mathbf{Z}^V$ the unique maximal vectors in $\arg\min\{g(p') \mid p' \leq p\}$ and in $\arg\min\{g(p') \mid p' \leq p + \chi_{X^+}\}$, respectively.

We first show $\check{q} \geq \check{p}$. By the submodularity of g, we have $g(\check{p}) + g(\check{q}) \geq g(\check{p} \vee \check{q}) + g(\check{p} \wedge \check{q})$. Since $\check{p} \in \arg\min\{g(p') \mid p' \leq p\}$ and $\check{p} \wedge \check{q} \leq \check{p}$, we have $g(\check{p}) \leq g(\check{p} \wedge \check{q})$. Therefore, $g(\check{q}) \geq g(\check{p} \vee \check{q})$ holds. Since $\check{q} \leq \check{p} \vee \check{q} \leq p + \chi_{X^+}$ and $\check{q}$ is the maximal vector in $\arg\min\{g(p') \mid p' \leq p + \chi_{X^+}\}$, we have $\check{q} = \check{p} \vee \check{q}$, i.e., $\check{q} \geq \check{p}$.

If $\beta^-(p) = 0$, i.e., $\check{p} = p$, then we have $\beta^-(p + \chi_{X^+}) = 0$ and $\check{q} = p + \chi_{X^+}$ since $p + \chi_{X^+}$ is a minimizer of g in the set $\{p' \in \mathbf{Z}^V \mid \check{p} = p \leq p' \leq p + \chi_{X^+}\}$. Hence, we assume $\beta^-(p) > 0$ in the following.

We then show $X^+ \cap Y^- = \emptyset$, where $Y^- = \arg\max\{p(v) - \check{p}(v) \mid v \in V\}$. Assume, to the contrary, that $X^+ \cap Y^- \neq \emptyset$. Then, Theorem 1 implies

$$g(p + \chi_{X^+}) + g(\check{p}) \geq g(p + \chi_{X^+ \setminus Y^-}) + g(\check{p} + \chi_{X^+ \cap Y^-})$$

since $X^+ \cap Y^- = \arg\max\{p(v) + \chi_{X^+}(v) - \check{p}(v) \mid v \in V\}$. Since $X^+ \in \arg\min \rho_p^+$, we have $g(p + \chi_{X^+}) \leq g(p + \chi_{X^+ \setminus Y^-})$, which implies $g(\check{p}) \geq g(\check{p} + \chi_{X^+ \cap Y^-})$. Since $\beta^-(p) > 0$, we have $\check{p} \leq \check{p} + \chi_{X^+ \cap Y^-} \leq p$, a contradiction to the fact that $\check{p}$ is the maximal minimizer of g in $\{p' \in \mathbf{Z}^V \mid p' \leq p\}$.

Finally, we have

$$\beta^-(p + \chi_{X^+}) = ||(p + \chi_{X^+}) - \check{q}||_\infty \leq ||(p + \chi_{X^+}) - \check{p}||_\infty = ||p - \check{p}||_\infty = \beta^-(p),$$

where the inequality follows from $\check{q} \geq \check{p}$ and the second equality by $X^+ \cap Y^- = \emptyset$. □

We note that a slightly weaker statement is shown in [4][Theorem 1], where the equalities "=" in "$\beta^+(p + \chi_{X^+}) = \beta^+(p) - 1$" and "$\beta^-(p - \chi_{X^-}) = \beta^-(p) - 1$" in the statement of Proposition 4 are replaced with inequalities "$\leq$".

Proposition 5 ([4]). *The number of iterations of Kolmogorov's primal algorithm for $L^\natural$-convex function g is bounded by $\beta^+(p^\circ) + \beta^-(p^\circ)$, which is further bounded by $2K_g^\infty$.*

2.2 Analysis of Murota's Algorithm

The results in this section except for the last proposition (Proposition 12) are based on the unpublished memorandum [9].

In [7], Murota firstly proposes a steepest descent algorithm for L-convex functions, which is then adapted to $L^\natural$-convex functions through the relation (1). For the simplicity of the proof, we firstly analyze the number of iterations required by the algorithm for L-convex functions, and then restate the result in terms of $L^\natural$-convex functions.

2.2.1 Analysis of Steepest Descent Algorithm for L-convex Functions

Murota's steepest descent algorithm for L-convex functions is described as follows, where $\tilde{V} = \{0\} \cup V$ and $\tilde{g} : \mathbf{Z}^{\tilde{V}} \to \mathbf{R} \cup \{+\infty\}$ is an L-convex function with $\arg\min \tilde{g} \neq \emptyset$.

Murota's steepest descent algorithm for L-convex functions
S0: Find a vector $\tilde{q} \in \mathrm{dom}\,\tilde{g}$.
S1: Let $\tilde{X}$ be the minimal minimizer of $\rho_{\tilde{q}}(\tilde{X}) = \tilde{g}(\tilde{q} + \chi_{\tilde{X}}) - \tilde{g}(\tilde{q})$.
S2: If $\tilde{g}(\tilde{q}) \leq \tilde{g}(\tilde{q} + \chi_{\tilde{X}})$, then stop ($\tilde{q}$ is a minimizer of $\tilde{g}$).
S3: Set $\tilde{q} := \tilde{q} + \chi_{\tilde{X}}$ and go to S1.

We analyze the number of iterations required by the algorithm. Let $\tilde{q}^\circ$ be the initial vector found in Step S0, and denote by $\tilde{q}^*$ the smallest of minimizers of $\tilde{g}$ with $\tilde{q}^* \geq \tilde{q}^\circ$. It is shown in [7] that the number of iterations of the algorithm is bounded by

$$\hat{K}_{\tilde{g}} = \max\{||\tilde{q} - \tilde{q}'||_1 \mid \tilde{q}, \tilde{q}' \in \mathrm{dom}\,\tilde{g},\ \tilde{q}(v) = \tilde{q}'(v) \text{ for some } v \in \tilde{V}\}.$$

We show that the number of iterations is bounded by

$$\hat{K}^\infty_{\tilde{g}} = \max\{||\tilde{q} - \tilde{q}'||_\infty \mid \tilde{q}, \tilde{q}' \in \mathrm{dom}\,\tilde{g},\ \tilde{q}(v) = \tilde{q}'(v) \text{ for some } v \in \tilde{V}\},$$

which is smaller than $\hat{K}_{\tilde{g}}$.

Lemma 6. *In Step* S1*, $\tilde{q} \leq \tilde{q}^*$ and $\tilde{q} \neq \tilde{q}^*$ imply*

$$\tilde{X} \cap \{v \in \tilde{V} \mid \tilde{q}^*(v) - \tilde{q}(v) = 0\} = \emptyset \quad \textit{and} \quad \arg\max\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \subseteq \tilde{X},$$

and hence, $\tilde{q} + \chi_{\tilde{X}} \leq \tilde{q}^*$ and $||(\tilde{q} + \chi_{\tilde{X}}) - \tilde{q}^*||_\infty = ||\tilde{q} - \tilde{q}^*||_\infty - 1$, in particular.

Proof. The first claim is already shown in [7][Lemma 3.2]; hence we prove below the second claim. Put $\tilde{Y} = \arg\max\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\}$, and assume to the contrary that $\tilde{Y} \setminus \tilde{X} \neq \emptyset$ holds. Put

$$\tilde{Z} = \arg\max\{\tilde{q}^*(v) - \tilde{q}(v) - \chi_{\tilde{X}}(v) \mid v \in \tilde{V}\} = \tilde{Y} \setminus \tilde{X}.$$

Theorem 1 implies

$$g(\tilde{q}^*) + g(\tilde{q} + \chi_{\tilde{X}}) \geq g(\tilde{q}^* - \chi_{\tilde{Z}}) + g(\tilde{q} + \chi_{\tilde{X}} + \chi_{\tilde{Z}}).$$

Since $\chi_{\tilde{X}} + \chi_{\tilde{Z}} = \chi_{\tilde{X} \cup \tilde{Y}}$, we have $g(\tilde{q} + \chi_{\tilde{X}} + \chi_{\tilde{Z}}) = g(\tilde{q} + \chi_{\tilde{X} \cup \tilde{Y}}) \geq g(\tilde{q} + \chi_{\tilde{X}})$, where the inequality is by $\tilde{X} \in \arg\min \rho_{\tilde{q}}$. Hence, we have $g(\tilde{q}^*) \geq g(\tilde{q}^* - \chi_{\tilde{Z}})$, a contradiction to the fact that $\tilde{q}^*$ is the minimal minimizer of $\tilde{g}$ with $\tilde{q}^* \geq \tilde{q}$ since $\tilde{q}^* - \chi_{\tilde{Z}} \geq \tilde{q}$. □

Proposition 7. *The number of iterations of the steepest descent algorithm for L-convex function $\tilde{g}$ is equal to $||\tilde{q}^\circ - \tilde{q}^*||_\infty$, which is bounded by $\hat{K}^\infty_{\tilde{g}}$.*

The following lemma is used in the analysis of the steepest descent algorithm for L$^\natural$-convex functions.

Lemma 8. *In each iteration it holds that*

$$\arg\min\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} = \{v \in \tilde{V} \mid \tilde{q}^*(v) - \tilde{q}(v) = 0\}.$$

Proof. The property (LF1) for $\tilde{g}$ implies $\tilde{q}^*(v) = \tilde{q}^\circ(v)$ for some $v \in \tilde{V}$. Hence, the set $\{v \in \tilde{V} \mid \tilde{q}^*(v) - \tilde{q}(v) = 0\}$ is nonempty at the beginning of the algorithm. Then, Lemma 6 implies that this set is nonempty during the following iterations. Since $\tilde{q}^* \geq \tilde{q}$ holds by Lemma 6, we have the claim. □

2.2.2 Analysis of Steepest Descent Algorithm for L$^\natural$-convex Functions

We analyze the number of iterations required by the steepest descent algorithm for L$^\natural$-convex functions.

The behavior of the steepest descent algorithm for g with the initial vector $p^\circ \in \mathbf{Z}^V$ is essentially the same as that of the steepest descent algorithm for the L-convex function $\tilde{g}$ defined by (1) with the initial vector $\tilde{q}^\circ = (0, p^\circ) \in \mathbf{Z} \times \mathbf{Z}^V$. The correspondence between the two steepest descent algorithms is as follows (see [7]):

L$^\natural$-convex g		L-convex $\tilde{g}$
$p \to p + \chi_X$	$\Longleftrightarrow$	$\tilde{q} \to \tilde{q} + (0, \chi_X)$
$p \to p - \chi_X$	$\Longleftrightarrow$	$\tilde{q} \to \tilde{q} + (1, \chi_{V \setminus X})$

where $\tilde{q} = (p_0, p + p_0\mathbf{1})$ and p_0 is a nonnegative integer representing the number of iterations with $(\varepsilon, X) = (-1, X^-)$ so far.

To analyze the number of iterations required by the algorithm for L$^\natural$-convex functions, we define values $\alpha^+(p)$ and $\alpha^-(p)$ for each vector $p \in \operatorname{dom} g$ as follows. For all $p, p' \in \mathbf{Z}^V$ we define

$$\begin{aligned} d_\infty^+(p, p') &= \max\big[0, \max_{v \in \mathrm{supp}^+(p - p')} |p(v) - p'(v)|\big], \\ d_\infty^-(p, p') &= \max\big[0, \max_{v \in \mathrm{supp}^-(p - p')} |p(v) - p'(v)|\big]. \end{aligned}$$

Let $\tilde{q}^* = (q_0^*, q^*) \in \mathbf{Z} \times \mathbf{Z}^V$ be the unique minimal vector in the set

$$\{(q_0, q) \in \mathbf{Z} \times \mathbf{Z}^V \mid q - q_0\mathbf{1} \in \arg\min g,\ (q_0, q) \geq (0, p^\circ)\}.$$

Note that $p^* = q^* - q_0^*\mathbf{1}$ is the minimizer of g found by the algorithm for L$^\natural$-convex functions. Then, $\alpha^+(p)$ and $\alpha^-(p)$ are defined as

$$\alpha^+(p) = d_\infty^+(p^*, p), \qquad \alpha^-(p) = d_\infty^-(p^*, p).$$

Since

$$\alpha^+(p^\circ) + \alpha^-(p^\circ) = d_\infty^+(p^*, p^\circ) + d_\infty^-(p^*, p^\circ) = ||(q_0^*, q^*) - (0, p^\circ)||_\infty,$$

the number of iterations is equal to $\alpha^+(p^\circ) + \alpha^-(p^\circ)$ by Proposition 7. In particular, we can prove the following property.

Proposition 9.
(i) *If* $(\varepsilon, X) = (1, X^+)$ *in Step* S1*, then* $\alpha^+(p + \varepsilon\chi_X) = \alpha^+(p) - 1$ *and* $\alpha^-(p + \varepsilon\chi_X) = \alpha^-(p)$.
(ii) *If* $(\varepsilon, X) = (-1, X^-)$ *in Step* S1*, then* $\alpha^+(p + \varepsilon\chi_X) = \alpha^+(p)$ *and* $\alpha^-(p + \varepsilon\chi_X) = \alpha^-(p) - 1$.

To prove Proposition 9, we restate Lemma 6 in terms of L$^\natural$-convex functions by using the correspondence between the two steepest descent algorithms.

Lemma 10.
(i) *Suppose that* $(\varepsilon, X) = (1, X^+)$ *holds in Step* S1*. Then, we have the following:*

> (i-1) $\mathrm{supp}^+(p^* - p) \neq \emptyset$ *and* $\{v \in \mathrm{supp}^+(p^* - p) \mid p^*(v) - p(v) = \alpha^+(p)\} \subseteq X$.
> (i-2) *if* $V \setminus \mathrm{supp}^+(p^* - p) \neq \emptyset$*, then* $X \cap \{v \in V \setminus \mathrm{supp}^+(p^* - p) \mid p^*(v) - p(v) = -\alpha^-(p)\} = \emptyset$.

(ii) *Suppose that* $(\varepsilon, X) = (-1, X^-)$ *holds in Step* S1*. Then, we have the following:*

(ii-1) $\mathrm{supp}^-(p^* - p) \neq \emptyset$ *and* $\{v \in \mathrm{supp}^-(p^* - p) \mid p^*(v) - p(v) = -\alpha^-(p)\} \subseteq X$.
(ii-2) *if* $V \setminus \mathrm{supp}^-(p^* - p) \neq \emptyset$*, then* $X \cap \{v \in V \setminus \mathrm{supp}^-(p^* - p) \mid p^*(v) - p(v) = \alpha^+(p)\} = \emptyset$.

Proof. We first note that for all $v \in V$,

$$p^*(v) - p(v) = \tilde{q}^*(v) - \tilde{q}(v) - q_0^* + p_0 = \{\tilde{q}^*(v) - \tilde{q}(v)\} - \{\tilde{q}^*(0) - \tilde{q}(0)\}.$$

Hence, we have the following equivalences for all $u \in V$ and $v \in V \cup \{0\}$, where $p^*(0) - p(0) = 0$ for convenience.

$$p^*(u) - p(u) < p^*(v) - p(v) \iff \tilde{q}^*(u) - \tilde{q}(u) < \tilde{q}^*(v) - \tilde{q}(v), \tag{8}$$
$$p^*(u) - p(u) = p^*(v) - p(v) \iff \tilde{q}^*(u) - \tilde{q}(u) = \tilde{q}^*(v) - \tilde{q}(v), \tag{9}$$
$$p^*(u) - p(u) > p^*(v) - p(v) \iff \tilde{q}^*(u) - \tilde{q}(u) > \tilde{q}^*(v) - \tilde{q}(v). \tag{10}$$

[Proof of (i-1)] By Lemma 6 we have

$$\arg\max\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \subseteq X. \tag{11}$$

This implies $\tilde{q}^*(u) - \tilde{q}(u) > \tilde{q}^*(0) - \tilde{q}(0)$ for some $u \in X$ $(\subseteq V)$, which in turn implies $\mathrm{supp}^+(p^* - p) \neq \emptyset$ by (10) with $v = 0$. Therefore, we have

$$\begin{aligned} &\arg\max\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \cap V \\ &\quad = \{v \in \mathrm{supp}^+(p^* - p) \mid p^*(v) - p(v) = \alpha^+(p)\}, \end{aligned}$$

which, together with (11), implies the latter claim.

[Proof of (i-2)] By (8) and (9), we have $\tilde{q}^*(v) - \tilde{q}(v) \leq \tilde{q}^*(0) - \tilde{q}(0)$ for all $v \in V \setminus \mathrm{supp}^+(p^* - p)$. Therefore, if $V \setminus \mathrm{supp}^+(p^* - p) \neq \emptyset$ then

$$\begin{aligned} &\arg\min\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \cap V \\ &\quad = \{v \in V \setminus \mathrm{supp}^+(p^* - p) \mid p^*(v) - p(v) = -\alpha^-(p)\}. \end{aligned}$$

Hence, the claim follows immediately from Lemmas 6 and 8.

[Proof of (ii-1)] The proof is similar to that for (i-1). By Lemmas 6 and 8, we have

$$[(V \setminus X) \cup \{0\}] \cap \arg\min\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} = \emptyset. \tag{12}$$

This implies $\tilde{q}^*(0) - \tilde{q}(0) > \tilde{q}^*(u) - \tilde{q}(u)$ for some $u \in X$ $(\subseteq V)$, which in turn implies $\mathrm{supp}^-(p^* - p) \neq \emptyset$ by (8) with $v = 0$. Therefore, we have

$$\arg\min\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \cap V = \{v \in \mathrm{supp}^-(p^* - p) \mid p^*(v) - p(v) = -\alpha^-(p)\},$$

which, together with (12), implies the latter claim.

[Proof of (ii-2)] The proof is similar to that for (i-2). By (9) and (10), we have $\tilde{q}^*(v) - \tilde{q}(v) \geq \tilde{q}^*(0) - \tilde{q}(0)$ for all $v \in V \setminus \mathrm{supp}^-(p^* - p)$. Therefore, if $V \setminus \mathrm{supp}^-(p^* - p) \neq \emptyset$ then

$$\begin{aligned} &\arg\max\{\tilde{q}^*(v) - \tilde{q}(v) \mid v \in \tilde{V}\} \cap V \\ &\quad = \{v \in V \setminus \mathrm{supp}^-(p^* - p) \mid p^*(v) - p(v) = \alpha^+(p)\}, \end{aligned}$$

which, together with Lemma 6, implies

$$\{v \in V \setminus \mathrm{supp}^-(p^* - p) \mid p^*(v) - p(v) = \alpha^+(p)\} \subseteq V \setminus X.$$

Hence, the claim follows. □

Proposition 9 is an immediate consequence of Lemma 10. Hence, we obtain the following proposition.

Proposition 11. *The number of iterations of Murota's steepest descent algorithm for $L^\natural$-convex function g is equal to $\alpha^+(p^\circ) + \alpha^-(p^\circ)$, which is bounded by $2K_g^\infty$.*

Finally, we show that Murota's steepest descent algorithm can be seen as the best implementation of Kolmogorov's algorithm from the viewpoint of the number of iterations.

Proposition 12. *The number of iterations of Kolmogorov's primal algorithm for $L^\natural$-convex function g is at least $\alpha^+(p^\circ) + \alpha^-(p^\circ)$.*

Proof. Let $p \in \mathrm{dom}\, g$ be any minimizer of g which can be found by Kolmogorov's algorithm. Then, Kolmogorov's algorithm requires at least $d^+_\infty(p, p^\circ) + d^-_\infty(p, p^\circ)$ iterations. On the other hand, the minimizer $p = p^*$ found by Murota's algorithm attains the minimum value of $d^+_\infty(p, p^\circ) + d^-_\infty(p, p^\circ)$ among all minimizers of g, as shown below. This fact implies the claim of the proposition since $\alpha^+(p^\circ) + \alpha^-(p^\circ) = d^+_\infty(p^*, p^\circ) + d^-_\infty(p^*, p^\circ)$.

Assume, to the contrary, that there exists $p' \in \arg\min g$ such that

$$d^+_\infty(p', p^\circ) + d^-_\infty(p', p^\circ) < d^+_\infty(p^*, p^\circ) + d^-_\infty(p^*, p^\circ) = ||(q_0^*, q^*) - (0, p^\circ)||_\infty. \quad (13)$$

Put $p'_0 = d^-_\infty(p', p^\circ)$. Then, the vector $(p'_0, p' + q'_0\mathbf{1}) \in \mathbf{Z} \times \mathbf{Z}^V$ is contained in the set

$$S = \{(q_0, q) \in \mathbf{Z} \times \mathbf{Z}^V \mid q - q_0\mathbf{1} \in \arg\min g,\ (q_0, q) \geq (0, p^\circ)\}$$

and satisfies $||(p'_0, p' + p'_0\mathbf{1}) - (0, p^\circ)||_\infty = d^+_\infty(p', p^\circ) + d^-_\infty(p', p^\circ)$, which is a contradiction to the fact that the vector (q_0^*, q^*) is the unique minimal vector in S. □

Acknowledgement

The author thanks Vladimir Kolmogorov and Kazuo Murota for communicating the recent report [4]. This research is supported by Humboldt research fellowship of the Alexander von Humboldt Foundation.

References

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

[2] S. Fujishige and K. Murota, Notes on L-/M-convex functions and the separation theorems, Math. Programming 88 (2000), 129–146.

[3] S. Iwata, L. Fleischer, and S. Fujishige, A combinatorial, strongly polynomial-time algorithm for minimizing submodular functions, J. ACM 48 (2001), 761–777.

[4] V. Kolmogorov, Primal-dual algorithm for convex Markov random fields, Technical Report MSR-TR-2005-117, Microsoft Research, September, 2005.

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

[6] K. Murota, Discrete Convex Analysis, SIAM, 2003.

[7] K. Murota, On steepest descent algorithms for discrete convex functions, SIAM J. Optimization 14 (2003), 699-707.

[8] A. Schrijver, A combinatorial algorithm minimizing submodular functions in strongly polynomial time, J. Combin. Theory Ser. B 80 (2000), 346–355.

[9] A. Shioura, An unpublished memorandum, August 2003.