MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

On Infimal Convolution of M-Convex Functions

Kazuo MUROTA

METR 2004–12 March 2004

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

On Infimal Convolution of M-Convex Functions

Kazuo MUROTA*

Abstract

The infimal convolution of M-convex functions is M-convex. This is a fundamental fact in discrete convex analysis that is often useful in its application to mathematical economics and game theory. M-convexity and its variant called M$^\natural$-convexity are closely related to gross substitutability, and the infimal convolution operation corresponds to an aggregation. This note provides a succinct description of the present knowledge about the infimal convolution of M-convex functions.

1 Definitions

Let V be a nonempty finite set, and let $\mathbf{Z}$ and $\mathbf{R}$ be the sets of integers and reals, respectively. We denote by $\mathbf{Z}^V$ the set of integral vectors indexed by V, and by $\mathbf{R}^V$ the set of real vectors indexed by V. For a vector $x = (x(v) \mid v \in V) \in \mathbf{Z}^V$, where $x(v)$ is the vth component of x, we define the positive support $\mathrm{supp}^+(x)$ and the negative support $\mathrm{supp}^-(x)$ by

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

We use notation $x(S) = \sum_{v \in S} x(v)$ for a subset S of V. For each $S \subseteq V$, we denote by χ_S the characteristic vector of S defined by: $\chi_S(v) = 1$ if $v \in S$ and $\chi_S(v) = 0$ otherwise, and write χ_v for $\chi_{\{v\}}$ for $v \in V$. For a vector $p = (p(v) \mid v \in V) \in \mathbf{R}^V$ and a function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$, we define functions $\langle p, x\rangle$ and $f[p](x)$ in $x \in \mathbf{Z}^V$ by

$$\langle p, x\rangle = \sum_{v \in V} p(v)x(v), \quad f[p](x) = f(x) + \langle p, x\rangle.$$

We also denote the set of minimizers of f and the effective domain of f by

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

We say that a function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom} f \neq \emptyset$ is *M-convex* if it satisfies the *exchange axiom*:

*Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan
E-mail: murota@mist.i.u-tokyo.ac.jp

(M-EXC) For $x, y \in \mathrm{dom} f$ and $u \in \mathrm{supp}^+(x-y)$, there exists $v \in \mathrm{supp}^-(x-y)$ such that

$$f(x) + f(y) \geq f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v). \tag{1}$$

The inequality (1) implicitly imposes the condition that $x - \chi_u + \chi_v \in \mathrm{dom} f$ and $y + \chi_u - \chi_v \in \mathrm{dom} f$ for the finiteness of the right-hand side. A function f is said to be *M-concave* if $-f$ is M-convex.

As a consequence of (M-EXC), the effective domain of an M-convex function f lies on a hyperplane $\{x \in \mathbf{R}^V \mid x(V) = r\}$ for some integer r, and accordingly, we may consider the projection of f along a coordinate axis. This means that, instead of the function f in $n = |V|$ variables, we may consider a function f' in $n-1$ variables defined by

$$f'(x') = f(x_0, x') \quad \text{with } x_0 = r - x'(V'), \tag{2}$$

where $V' = V \setminus \{v_0\}$ for an arbitrarily fixed element $v_0 \in V$, and a vector $x \in \mathbf{Z}^V$ is represented as $x = (x_0, x')$ with $x_0 = x(v_0) \in \mathbf{Z}$ and $x' \in \mathbf{Z}^{V'}$. Note that the effective domain $\mathrm{dom} f'$ of f' is the projection of $\mathrm{dom} f$ along the chosen coordinate axis v_0. A function f' derived from an M-convex function by such projection is called an M$^\natural$-convex[1)] function.

More formally, an M$^\natural$-convex function is defined as follows. Let "0" denote a new element not in V and put $\tilde{V} = \{0\} \cup V$. A function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is called *M*$^\natural$*-convex* if the function $\tilde{f} : \mathbf{Z}^{\tilde{V}} \to \mathbf{R} \cup \{+\infty\}$ defined by

$$\tilde{f}(x_0, x) = \begin{cases} f(x) & \text{if } x_0 = -x(V) \\ +\infty & \text{otherwise} \end{cases} \qquad (x_0 \in \mathbf{Z}, x \in \mathbf{Z}^V) \tag{3}$$

is an M-convex function. It is known (see [4, Theorem 6.2]) that an M$^\natural$-convex function f can be characterized by a similar exchange property:

(M$^\natural$-EXC) For $x, y \in \mathrm{dom} f$ and $u \in \mathrm{supp}^+(x-y)$,

$$\begin{aligned} f(x) + f(y) \quad \geq \quad \min \Bigg[& f(x - \chi_u) + f(y + \chi_u), \\ & \min_{v \in \mathrm{supp}^-(x-y)} \{ f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v) \} \Bigg], \end{aligned} \tag{4}$$

where the minimum over an empty set is $+\infty$ by convention. A function f is said to be *M*$^\natural$*-concave* if $-f$ is M$^\natural$-convex.

Whereas M$^\natural$-convex functions are conceptually equivalent to M-convex functions, the class of M$^\natural$-convex functions is strictly larger than that of M-convex functions. This follows from the implication: (M-EXC) $\Rightarrow$ (M$^\natural$-EXC). The simplest example of an M$^\natural$-convex function that is not M-convex is a one-dimensional (univariate) discrete convex function, depicted in Fig. 1.

[1)] "M$^\natural$-convex" should be read "M-natural-convex."

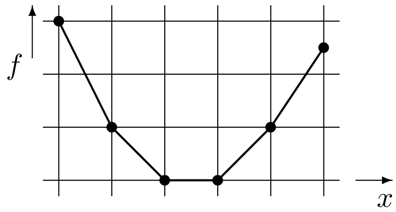

Figure 1: Univariate discrete convex function

Proposition 1 ([4, Theorem 6.3]). *An M-convex function is* $\mathrm{M}^\natural$*-convex. Conversely, an* $\mathrm{M}^\natural$*-convex function is M-convex if and only if the effective domain is contained in a hyperplane* $\{x \in \mathbf{Z}^V \mid x(V) = r\}$ *for some* $r \in \mathbf{Z}$*.*

$\mathrm{M}^\natural$-convex functions enjoy a number of nice properties that are expected of "discrete convex functions." Furthermore, $\mathrm{M}^\natural$-concave functions provide with a natural model of utility functions (see [4, §11.3] and [5]). In particular, it is known that $\mathrm{M}^\natural$-concavity is equivalent to gross substitutes property, and that $\mathrm{M}^\natural$-concavity implies submodularity, which is the discrete version of decreasing marginal returns.

It follows from (M-EXC) that the effective domain of an M-convex function f satisfies the exchange axiom:

> **(B-EXC)** For $x, y \in B$ and $u \in \mathrm{supp}^+(x - y)$, there exists $v \in \mathrm{supp}^-(x - y)$ such that $x - \chi_u + \chi_v \in B$ and $y + \chi_u - \chi_v \in B$,

since $x - \chi_u + \chi_v \in \mathrm{dom} f$ and $y + \chi_u - \chi_v \in \mathrm{dom} f$ for $x, y \in \mathrm{dom} f$ in (1). A nonempty set B of integer points satisfying (B-EXC) is referred to as an *M-convex set*.

2 Convolution Theorem

For a pair of functions $f_1, f_2 : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$, the *integer infimal convolution* is a function $f_1 \Box_{\mathbf{Z}} f_2 : \mathbf{Z}^V \to \mathbf{R} \cup \{\pm\infty\}$ defined by

$$(f_1 \Box_{\mathbf{Z}} f_2)(x) = \inf\{f_1(x_1) + f_2(x_2) \mid x = x_1 + x_2,\ x_1, x_2 \in \mathbf{Z}^V\} \quad (x \in \mathbf{Z}^V). \tag{5}$$

Provided that $f_1 \Box_{\mathbf{Z}} f_2$ is away from the value of $-\infty$, we have

$$\mathrm{dom}(f_1 \Box_{\mathbf{Z}} f_2) = \mathrm{dom} f_1 + \mathrm{dom} f_2, \tag{6}$$

where the right-hand side means the Minkowski sum of the effective domains.

The convolution theorem reads as follows.

Theorem 2 ([4, Theorem 6.13]). *For M-convex functions* f_1 *and* f_2*, the integer infimal convolution* $f = f_1 \Box_{\mathbf{Z}} f_2$ *is M-convex, provided* $f > -\infty$*.*

A proof of this theorem is given in Section 3, whereas the M$^\natural$-version below is an immediate corollary.

Corollary 3 ([4, Theorem 6.15]). *For M$^\natural$-convex functions f_1 and f_2, the integer infimal convolution $f = f_1 \Box_{\mathbf{Z}} f_2$ is M$^\natural$-convex, provided $f > -\infty$.*

Proof. Let $\tilde{f}_1$ and $\tilde{f}_2$ be the M-convex functions associated with the M$^\natural$-convex functions f_1 and f_2 as in (3). For $x_0 \in \mathbf{Z}$, $x \in \mathbf{Z}^V$ we have

$$
\begin{aligned}
&(\tilde{f}_1 \Box_{\mathbf{Z}} \tilde{f}_2)(x_0, x) \\
&= \inf\{\tilde{f}_1(y_0, y) + \tilde{f}_2(z_0, z) \mid x = y + z, x_0 = y_0 + z_0\} \\
&= \inf\{f_1(y) + f_2(z) \mid x = y + z, x_0 = y_0 + z_0, y_0 = -y(V), z_0 = -z(V)\} \\
&= \inf\{f_1(y) + f_2(z) \mid x = y + z, x_0 = -x(V)\} \\
&= \begin{cases} (f_1 \Box_{\mathbf{Z}} f_2)(x) & \text{if } x_0 = -x(V) \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}
$$

This shows $\tilde{f}_1 \Box_{\mathbf{Z}} \tilde{f}_2 = (f_1 \Box_{\mathbf{Z}} f_2)\tilde{}$ in the notation of (3), whereas $\tilde{f}_1 \Box_{\mathbf{Z}} \tilde{f}_2$ is M-convex by Theorem 2 applied to $\tilde{f}_1$ and $\tilde{f}_2$. Therefore, $f_1 \Box_{\mathbf{Z}} f_2$ is M$^\natural$-convex. □

Remark 1. The convolution theorem (Theorem 2) originates in [1, Theorem 6.10], and is described in [2, p. 80, Theorem 2.44 (5)], [3, p. 118, Theorem 4.8 (8)], and [4, p. 143, Theorem 6.13 (8)]. The M$^\natural$-version (Corollary 3) is also stated in [2, p. 83], [3, p. 119, Theorem 4.10], and [4, p. 144, Theorem 6.15 (1)]. An application of this fact to the aggregation of utility functions can be found in [3, p. 275, Proposition 9.13] and [4, p. 337, Theorem 11.12]. In particular, the convolution theorem implies that if the individual utility functions enjoy gross substitutes property, so does the aggregated utility function. ■

3 Proof

The proof of Theorem 2 given here relies on two fundamental facts stated in the lemmas below. The first shows that the class of M-convex sets is closed under Minkowski addition, and the second gives a characterization of an M-convex function in terms of M-convex sets.

Lemma 4 ([4, Theorem 4.23]). *The Minkowski sum of two M-convex sets is M-convex.*

Lemma 5 ([4, Theorem 6.30]). *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be a function with a bounded nonempty effective domain. Then, f is M-convex if and only if $\arg\min f[-p]$ is an M-convex set for each $p \in \mathbf{R}^V$.*

Let f_1 and f_2 be M-convex functions, and put $f = f_1 \Box_{\mathbf{Z}} f_2$. First we treat the case where $\mathrm{dom} f_1$ and $\mathrm{dom} f_2$ are bounded. The expression (6) shows that $\mathrm{dom} f$ is bounded.

For each $p \in \mathbf{R}^V$ we have

$$f[-p] = (f_1[-p]) \Box_{\mathbf{Z}} (f_2[-p]),$$

from which follows

$$\arg\min f[-p] = \arg\min f_1[-p] + \arg\min f_2[-p]$$

by (5). In this expression, both $\arg\min f_1[-p]$ and $\arg\min f_2[-p]$ are M-convex sets by Lemma 5 (only if part), and therefore, their Minkowski sum (the right-hand side) is M-convex by Lemma 4. This means that $\arg\min f[-p]$ is M-convex for each $p \in \mathbf{R}^V$, which implies the M-convexity of f by Lemma 5 (if part).

The general case without the boundedness assumption on effective domains can be treated via limiting procedure as follows. For $i = 1, 2$ and $k = 1, 2, \ldots$, define $f_i^{(k)} : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ by

$$f_i^{(k)}(x) = \begin{cases} f_i(x) & \text{if } ||x||_\infty \le k \\ +\infty & \text{otherwise} \end{cases} \qquad (x \in \mathbf{Z}^V),$$

which is an M-convex function with a bounded effective domain, provided that k is large enough for $\mathrm{dom} f_i^{(k)} \neq \emptyset$. For each k, the infimal convolution $f^{(k)} = f_1^{(k)} \Box_{\mathbf{Z}} f_2^{(k)}$ is M-convex by the above argument, and moreover, $\lim_{k\to\infty} f^{(k)}(x) = f(x)$ for each x. It remains to demonstrate the property (M-EXC) for f. Take $x, y \in \mathrm{dom} f$ and $u \in \mathrm{supp}^+(x-y)$. There exists $k_0 = k_0(x, y)$, depending on x and y, such that $x, y \in \mathrm{dom} f^{(k)}$ for every $k \ge k_0$. Since $f^{(k)}$ is M-convex, there exists $v_k \in \mathrm{supp}^-(x-y)$ such that

$$f^{(k)}(x) + f^{(k)}(y) \ge f^{(k)}(x - \chi_u + \chi_{v_k}) + f^{(k)}(y + \chi_u - \chi_{v_k}).$$

Since $\mathrm{supp}^-(x-y)$ is a finite set, at least one element of $\mathrm{supp}^-(x-y)$ appears infinitely many times in the sequence $v_1, v_2, \ldots$. More precisely, there exists $v \in \mathrm{supp}^-(x-y)$ and an increasing subsequence $k_1 < k_2 < \cdots$ such that $v_{k_j} = v$ for $j = 1, 2, \ldots$. By letting $k \to \infty$ along this subsequence in the above inequality we obtain

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

Thus f satisfies (M-EXC). This completes the proof of Theorem 2.

Remark 2. Here is an example to demonstrate the necessity of the limiting argument in the above proof. For M-convex functions $f_1, f_2 : \mathbf{Z}^2 \to \mathbf{R}$ defined by

$$f_1(x) = \begin{cases} \exp(-x(1)) & \text{if } x(1) + x(2) = 0, \\ +\infty & \text{otherwise,} \end{cases} \qquad f_2(x) = \begin{cases} \exp(x(1)) & \text{if } x(1) + x(2) = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

we have

$$f(x) = (f_1 \Box_{\mathbf{Z}} f_2)(x) = \inf\{\exp(-t) + \exp(x(1) - t) \mid t \in \mathbf{Z}\} = 0$$

for all $x \in \mathbf{Z}^2$ with $x(1) + x(2) = 0$. The infimum is not attained by any finite t, and consequently, $f^{(k)}(x)$ is not equal to $f(x)$ for any finite k. This is why we need the limiting argument in the proof. ■

Remark 3. The infimal convolution operation of M-convex functions can be formulated as a special case of the transformation of an M-convex function by a network, and the convolution theorem (Theorem 2) can be understood as a special case of a theorem on network transformation.

The general framework of the network transformation is as follows. Let $G = (V, A; S, T)$ be a directed graph with vertex set V, arc set A, entrance set S and exit set T, where S and T are disjoint subsets of V. We consider an integer-valued flow $\xi = (\xi(a) \mid a \in A) \in \mathbf{Z}^A$. For each $a \in A$, the cost of the flow $\xi(a)$ through arc a is represented by a function $f_a : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$. Given a function $f : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ associated with the entrance set S, we define another function $\widehat{f} : \mathbf{Z}^T \to \mathbf{R} \cup \{\pm\infty\}$ on the exit set T by

$$\begin{aligned} \widehat{f}(y) &= \inf_{\xi,x} \{ f(x) + \sum_{a \in A} f_a(\xi(a)) \mid \partial\xi = (x, -y, \mathbf{0}), \\ &\qquad \xi \in \mathbf{Z}^A, (x, -y, \mathbf{0}) \in \mathbf{Z}^S \times \mathbf{Z}^T \times \mathbf{Z}^{V \setminus (S \cup T)} \} \quad (y \in \mathbf{Z}^T), \end{aligned}$$

where $\partial\xi \in \mathbf{Z}^V$ denotes a vector defined by

$$\partial\xi(v) = \sum\{\xi(a) \mid \text{arc } a \text{ leaves vertex } v\} - \sum\{\xi(a) \mid \text{arc } a \text{ enters vertex } v\} \quad (v \in V).$$

We may think of $\widehat{f}(y)$ as the minimum cost of an integer-valued flow to meet a demand specification y at the exit, where the cost consists of two parts, the cost $f(x)$ of supply or production of x at the entrance and the cost $\sum_{a \in A} f_a(\xi(a))$ of transportation through arcs; the sum of these is to be minimized over varying supply x and flow ξ subject to the flow conservation constraint $\partial\xi = (x, -y, \mathbf{0})$. We regard $\widehat{f}$ as a transformation of f by the network.

It is known ([4, Theorem 9.27]) that if f_a is a univariate discrete convex function for each $a \in A$ and f is an M-convex function, then $\widehat{f}$ is an M-convex function, provided that $\widehat{f} > -\infty$ and $\widehat{f} \not\equiv +\infty$.

For the infimal convolution of functions f_1 and f_2, let V_1 and V_2 be copies of V and consider a bipartite graph $G = (S \cup T, A; S, T)$ (see Fig. 2) with $S = V_1 \cup V_2$, $T = V$ and $A = \{(v_1, v) \mid v \in V\} \cup \{(v_2, v) \mid v \in V\}$, where $v_i \in V_i$ is the copy of $v \in V$ for $i = 1, 2$. We regard f_i as being defined on V_i for $i = 1, 2$ and assume that the arc cost functions f_a $(a \in A)$ are identically zero. The function $\widehat{f}$ induced on T coincides with the infimal convolution $f_1 \Box_{\mathbf{Z}} f_2$. In this case it is always true that $\widehat{f} \not\equiv +\infty$. Thus the convolution theorem (Theorem 2) follows from [4, Theorem 9.27], as is explained in [4, Note 9.30].

The connection to network transformation also suggests that the infimal convolution $f_1 \Box_{\mathbf{Z}} f_2$ can be evaluated by solving an M-convex submodular flow problem; see [4, Section 9.2] for the definition of the problem and [4, Section 10.4] for algorithms. ■

Acknowledgement The author thanks Takuya Iimura and Akihisa Tamura for helpful comments.

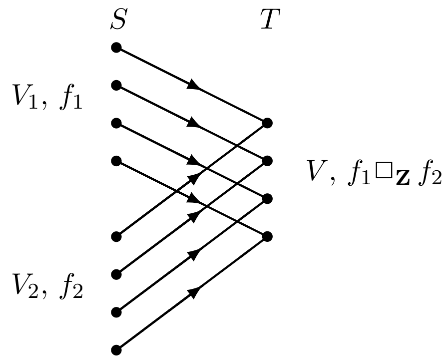

Figure 2: Bipartite graph for infimal convolution

References

[1] K. Murota: Convexity and Steinitz's exchange property, *Advances in Mathematics*, **124** (1996), 272–311.

[2] K. Murota: Discrete convex analysis (in Japanese), in: S. Fujishige, ed., *Discrete Structures and Algorithms, Vol. V*, Kindai-Kagakusha, Tokyo, 1998, Chapter 2, 51–100.

[3] K. Murota: *Discrete Convex Analysis—An Introduction* (in Japanese), Kyoritsu Publishing Co., Tokyo, 2001.

[4] K. Murota: *Discrete Convex Analysis*, SIAM Monographs on Discrete Mathematics and Applications, Vol. 10, Society for Industrial and Applied Mathematics, Philadelphia, 2003.

[5] A. Tamura: Applications of discrete convex analysis to mathematical economics, *Publications of Research Institute for Mathematical Sciences*, 2004, to appear.