On Infimal Convolution of M-Convex Functions (Applications of Discrete Convex Analysis to Game Theory and Mathematical Economics)

Author(s)
Murota, Kazuo

Citation
数理解析研究所講究録 (2004), 1371: 20-26

Issue Date
2004-04

URL
http://hdl.handle.net/2433/25472

Type
Departmental Bulletin Paper

Textversion
publisher

Kyoto University
On Infimal Convolution of M-Convex Functions

Graduate School of Information Science and Technology, University of Tokyo

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'-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 \( v \)th component of \( x \), we define the positive support \( \text{supp}^+(x) \) and the negative support \( \text{supp}^-(x) \) by

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

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 \( \chi_S \) the characteristic vector of \( S \) defined by: \( \chi_S(v) = 1 \) if \( v \in S \) and \( \chi_S(v) = 0 \) otherwise, and write \( \chi_v \) for \( \chi_{\{v\}} \) for \( v \in V \). For a vector \( p = (p(v) \mid v \in V) \in \mathbf{R}^V \) and a function \( f : \mathbf{Z}^V \rightarrow \mathbf{R} \cup \{+\infty\} \), 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

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

We say that a function \( f : \mathbf{Z}^V \rightarrow \mathbf{R} \cup \{+\infty\} \) with \( \text{dom } f \neq \emptyset \) is \textit{M-convex} if it satisfies the exchange axiom:

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

\[
f(x) + \langle p, x \rangle = f(y) + \langle p, y \rangle,
\]

for every \( p \in \mathbf{R}^V \) with \( p(u) > 0 \) and \( p(v) < 0 \).
\[ f(x) + f(y) \geq f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v). \]  

(1)

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

As a consequence of (M-EXC), the effective domain of an \( M\)-convex function \( f \) lies on a hyperplane \( \{ x \in \mathbb{R}^V \mid x(V) = r \} \) 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

\[ f'(x') = f(x_0, x') \quad \text{with} \quad x_0 = r - x'(V'), \]

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

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

\[ \tilde{f}(x_0, x) = \begin{cases} f(x) & \text{if } x_0 = -x(V) \\ +\infty & \text{otherwise} \end{cases} \quad (x_0 \in \mathbb{Z}, x \in \mathbb{Z}^{\tilde{V}}) \]

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

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

\[ f(x) + f(y) \geq \min \left[ f(x - \chi_u) + f(y + \chi_u), \right. \]

\[ \left. \min_{v \in \text{supp}^+(x - y)} \{ f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v) \} \right], \]

(4) where the minimum over an empty set is \(+\infty\) by convention. A function \( f \) is said to be \( M^a\)-concave if \(-f\) is \( M^a\)-convex.

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

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

\(^1\)"\( M^a\)-convex" should be read "\( M\)-natural-convex."
$\vec{x}$

Figure 1: Univariate discrete convex function

$M^h$-convex functions enjoy a number of nice properties that are expected of "discrete convex functions." Furthermore, $M^h$-concave functions provide with a natural model of utility functions (see [4, §11.3] and [5]). In particular, it is known that $M^h$-concavity is equivalent to gross substitutes property, and that $M^h$-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 $u \in \text{supp}^+(x - y)$, there exists $v \in \text{supp}^-(x - y)$ such that $x - \chi_u + \chi_v \in B$ and $y + \chi_u - \chi_v \in B$,

since $x - \chi_u + \chi_v \in \text{dom} f$ and $y + \chi_u - \chi_v \in \text{dom} f$ for $x, y \in \text{dom} f$ 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 $f_1, f_2 : Z^V \to R \cup \{+\infty\}$, the integer infimal convolution is a function $f_1 \square_Z f_2 : Z^V \to R \cup \{+\infty\}$ defined by

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

(5)

Provided that $f_1 \square_Z f_2$ is away from the value of $-\infty$, we have

$$\text{dom}(f_1 \square_Z f_2) = \text{dom} f_1 + \text{dom} f_2,$$

(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 $f = f_1 \square_Z f_2$ is $M$-convex, provided $f > -\infty$.

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

Corollary 3 ([4, Theorem 6.15]). For $M^h$-convex functions $f_1$ and $f_2$, the integer infimal convolution $f = f_1 \square_Z f_2$ is $M^h$-convex, provided $f > -\infty$. 
Proof. Let $\tilde{f}_1$ and $\tilde{f}_2$ be the M-convex functions associated with the $M^\mathfrak{h}$-convex functions $f_1$ and $f_2$ as in (3). For $x_0 \in \mathbb{Z}$, $x \in \mathbb{Z}^V$ we have

$$(\tilde{f}_1 \square_\mathbb{Z} \tilde{f}_2)(x_0, x) = \inf \{\tilde{f}_1(y_0, y) + \tilde{f}_2(z_0, z) \mid x = y + z, x_0 = y_0 + z_0\}$$

$$= \inf \{f_1(y) + f_2(z) \mid x = y + z, x_0 = y_0 + z_0, y_0 = -y(V), z_0 = -z(V)\}$$

$$= \inf \{f_1(y) + f_2(z) \mid x = y + z, x_0 = -x(V)\}$$

$$= \left\{ \begin{array}{ll}
(f_1 \square_\mathbb{Z} f_2)(x) & \text{if } x_0 = -x(V) \\
+\infty & \text{otherwise.}
\end{array} \right.$$ 

This shows $\tilde{f}_1 \square_\mathbb{Z} \tilde{f}_2 = (f_1 \square_\mathbb{Z} f_2)^\sim$ in the notation of (3), whereas $\tilde{f}_1 \square_\mathbb{Z} \tilde{f}_2$ is M-convex by Theorem 2 applied to $\tilde{f}_1$ and $\tilde{f}_2$. Therefore, $f_1 \square_\mathbb{Z} f_2$ is $M^\mathfrak{h}$-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 $M^2$-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 $f : \mathbb{Z}^V \rightarrow \mathbb{R} \cup \{+\infty\}$ 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 $p \in \mathbb{R}^V$.

Let $f_1$ and $f_2$ be M-convex functions, and put $f = f_1 \square_\mathbb{Z} f_2$. First we treat the case where $\text{dom} f_1$ and $\text{dom} f_2$ are bounded. The expression (6) shows that $\text{dom} f$ is bounded. For each $p \in \mathbb{R}^V$ we have

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

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 $p \in \mathbb{R}^V$, 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 $f_i^{(k)} : \mathbb{Z}^V \to \mathbb{R} \cup \{\pm \infty\}$ by

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

which is an M-convex function with a bounded effective domain, provided that $k$ is large enough for $\text{dom} f_i^{(k)} \neq \emptyset$. For each $k$, the infimal convolution $f^{(k)} = f_1^{(k)} \square_{\mathbb{Z}} f_2^{(k)}$ 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 $x, y \in \text{dom } f$ and $u \in \text{supp}^+(x-y)$. There exists $k_0 = k_0(x,y)$, depending on $x$ and $y$, such that $x, y \in \text{dom } f^{(k)}$ for every $k \geq k_0$. Since $f^{(k)}$ is M-convex, there exists $v_k \in \text{supp}^-(x-y)$ such that

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

Since $\text{supp}^-(x-y)$ is a finite set, at least one element of $\text{supp}^-(x-y)$ appears infinitely many times in the sequence $v_1, v_2, \ldots$. More precisely, there exists $v \in \text{supp}^-(x-y)$ 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

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

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 $f_1, f_2 : \mathbb{Z}^2 \to \mathbb{R}$ defined by

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

we have

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

for all $x \in \mathbb{Z}^2$ 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 $\xi = (\xi(a) | a \in A) \in \mathbb{Z}^A$. For each $a \in A$, the cost of the flow $\xi(a)$ through arc $a$ is represented by a function $f_a : \mathbb{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$. Given a function $f : \mathbb{Z}^S \rightarrow \mathbb{R} \cup \{+\infty\}$ associated with the entrance set $S$, we define another function $\hat{f} : \mathbb{Z}^T \rightarrow \mathbb{R} \cup \{-\infty\}$ on the exit set $T$ by

$$
\hat{f}(y) = \inf_{\xi,x} \{f(x) + \sum_{a \in A} f_a(\xi(a)) | \partial \xi = (x, -y, 0), \xi \in \mathbb{Z}^A, (x, -y, 0) \in \mathbb{Z}^S \times \mathbb{Z}^T \times \mathbb{Z}^{V \setminus (S \cup T)} \} \quad (y \in \mathbb{Z}^T),
$$

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

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

We may think of $\hat{f}(y)$ 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 $\xi$ subject to the flow conservation constraint $\partial \xi = (x, -y, 0)$. We regard $\hat{f}$ 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 $\hat{f}$ is an M-convex function, provided that $\hat{f} > -\infty$ and $\hat{f} \neq +\infty$.

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 $A = \{(v_1, v) | v \in V_1\} \cup \{(v_2, v) | v \in V_2\}$, 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 $\hat{f}$ induced on $T$ coincides with the infimal convolution $f_1 \mathcal{D}^Z f_2$. In this case it is always true that $\hat{f} \neq +\infty$. 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 $f_1 \mathcal{D}^Z f_2$ 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.

References

Figure 2: Bipartite graph for infimal convolution


