No. 364

Linear and Nonlinear Optimization Problems with Submodular Constraints

by

Satoru Fujishige

May 1988

Linear and Nonlinear Optimization Problems with Submodular Constraints

SATORU FUJISHIGE

INSTITUTE OF SOCIO-ECONOMIC PLANNING
UNIVERSITY OF TSUKUBA
TSUKUBA, IBARAKI 305, JAPAN

Abstract. The author introduced the concepts of submodular/supermodular system and its associated base polyhedron, which give a useful mathematical framework for treating combinatorial optimization problems related to submodular/supermodular functions. From the point of view of this framework we survey a class of linear/nonlinear and continuous/discrete optimization problems with constraints described by submodular functions.

1. Introduction

The theory of matroids has successfully been applied to many practical engineering problems, where a fundamental role is played by submodular functions (see, e.g., [24], [25], [30]). Submodular and supermodular functions arise typically in matroids as rank functions [4], in network flows as cut functions [8], in the Shannon information theory as entropy functions [13], in convex games as characteristic functions [33], and in many other combinatorial systems.

The author introduced the concepts of submodular/supermodular system and its associated base polyhedron ([12], [16]), which give a useful mathematical framework for treating combinatorial optimization problems related to submodular/supermodular functions. From the point of view of this framework we survey a class of linear/nonlinear and continuous/discrete optimization problems with constraints described by submodular functions.

For general information on submodular and supermodular functions readers should be referred, e.g., to [4], [6], [10], [16], and [28].

Propositions without any references in this paper seem to be well known to (poly-)matroid theorists or are easy corollaries.

2. Submodular/Supermodular Systems and Base Polyhedra

We give the definitions of submodular and supermodular systems, base polyhedron etc. and show their fundamental properties ([12], [16]).

2.1. Definitions

Let E be a finite nonempty set and $\mathcal{D}$ be a collection of subsets of E closed with respect to set union and intersection. Such $\mathcal{D}$ is a *distributive lattice* with set union and intersection as the lattice operations, join and meet.

Let R be the set of reals. Throughout the present paper, unless otherwise explicitly stated, R can be any totally ordered additive group, such as the set Z of integers, the set Q of rationals etc.

A function f from $\mathcal{D}$ to R is called a *submodular function* on $\mathcal{D}$ if for each pair of $X, Y \in \mathcal{D}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y). \tag{2.1}$$

We assume throughout the present paper that $\emptyset, E \in \mathcal{D}$ and $f(\emptyset) = 0$. We call the pair $(\mathcal{D}, f)$ a *submodular system* on E and f the *rank function* of $(\mathcal{D}, f)$.

For a submodular system $(\mathcal{D}, f)$ we define the polyhedron

$$P(f) = \{\, x \mid x \in R^E, \;\; \forall X \in \mathcal{D} : x(X) \leq f(X) \,\}, \tag{2.2}$$

where R^E is the set of all the vectors $x = (x(e) : e \in E)$ with coordinates indexed by E and $x(e) \in R$ $(e \in E)$ and for each $x \in R^E$ and $X \in \mathcal{D}$ $x(X)$ is defined by

$$x(X) = \sum_{e \in X} x(e). \tag{2.3}$$

We call $P(f)$ the *submodular polyhedron* associated with submodular system $(\mathcal{D}, f)$. We also define the polyhedron

$$B(f) = \{\, x \mid x \in P(f), \quad x(E) = f(E) \,\} \tag{2.4}$$

which is called the *base polyhedron* associated with $(\mathcal{D}, f)$. The base polyhedron $B(f)$ consists of all the maximal vectors in $P(f)$, where the order among vectors is the one in the ordinary vector lattice, i.e., $x \leq y$ if and only if $x(e) \leq y(e)$ for each $e \in E$. A vector in $B(f)$ is called a *base* of $(\mathcal{D}, f)$.

Let $\leq^*$ be the dual order of the ordinary order $\leq$ in R, i.e., $a \leq^* b$ if and only if $b \leq a$, and consider the dual totally ordered additive group $(R, \leq^*)$ in stead of $(R, \leq)$ in the above definitions of submodular function, submodular system, submodular polyhedron and base polyhedron. Then replacing f by g, we have a submodular function $g : \mathcal{D} \to H$, a submodular system $(\mathcal{D}, g)$, and the submodular polyhedron $P(g)$ and the base polyhedron $B(g)$ associated with the submodular system $(\mathcal{D}, g)$ with respect to the totally ordered additive group $(R, \leq^*)$ and they are, respectively, called, with respect to the original underlying totally ordered additive group $(R, \leq)$, a *supermodular function*, a *supermodular system*, and the *supermodular polyhedron* and the *base polyhedron associated with the supermodular system* $(\mathcal{D}, g)$.

Let $(\mathcal{D}, f)$ be a submodular system on E and define

$$\bar{\mathcal{D}} = \{ E - X \mid X \in \mathcal{D} \}, \tag{2.5}$$

$$f^{\#}(E - X) = f(E) - f(X) \qquad (X \in \mathcal{D}). \tag{2.6}$$

Here, $\bar{\mathcal{D}}$ is the dual distributive lattice of $\mathcal{D}$. We call $f^{\#} : \bar{\mathcal{D}} \to R$ the *dual supermodular function* of $f : \mathcal{D} \to R$ and $(\bar{\mathcal{D}}, f^{\#})$ the *dual supermodular sytem* of $(\mathcal{D}, f)$ ([12], [34]). Similarly we define the *dual submodular function* $g^{\#} : \bar{\mathcal{D}} \to R$ of a supermodular function $g : \mathcal{D} \to R$ and the *dual submodular system* $(\bar{\mathcal{D}}, g^{\#})$ of $(\mathcal{D}, g)$.

The following duality between submodular functions and supermodular functions is fundamental.

Proposition 2.1 ([12], [34]): For a submodular function $f : \mathcal{D} \to R$ and a supermodular function $g : \mathcal{D} \to R$ we have

$$(f^{\#})^{\#} = f, \qquad (g^{\#})^{\#} = g, \tag{2.7}$$

$$B(f) = B(f^{\#}), \qquad B(g) = B(g^{\#}). \tag{2.8}$$

□

It should be noted that Proposition 2.1 holds for any set functions f and g if we formally adopt the definitions of $B(f)$ and $B(g)$. The duality of (2.7) and (2.8) may also be useful for a more general class of set functions.

We say subsets X and Y of E *intersect* if $X \cap Y$ is nonempty. A family $\mathcal{F}_1$ of subsets of E is called an *intersecting family* if for each intersecting pair of $X, Y \in \mathcal{F}_1$ we have $X \cup Y$, $X \cap Y \in \mathcal{F}_1$. A function $f : \mathcal{F}_1 \to R$ is called a *submodular function on the intersecting family* $\mathcal{F}_1$ if for each intersecting pair of $X, Y \in \mathcal{F}_1$ we have

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \tag{2.9}$$

(cf.[4], [9]).

Moreover, we say subsets X and Y of E *cross* if the four sets $X \cap Y$, $(E-X) \cap Y$, $X \cap (E - Y)$ and $(E - X) \cap (E - Y)$ are nonempty. A family $\mathcal{F}_2$ of subsets of E is called a *crossing family* if for each crossing pair of $X, Y \in \mathcal{F}_2$ we have $X \cup Y$, $X \cap Y \in \mathcal{F}_2$. Also a function $f : \mathcal{F}_2 \to R$ is called a *submodular function on the* *crossing family* $\mathcal{F}_2$ if for each crossing pair of X, $Y \in \mathcal{F}_2$ we have the inequality (2.9) ([5]).

It should be noted that if $\mathcal{F}_2$ is a cross-free family of subsets of E, then $\mathcal{F}_2$ is a crossing family and any function f: $\mathcal{F}_2 \to R$ is a submodular function on the crossing family $\mathcal{F}_2$.

We see from the definitions that a distributive lattice $\mathcal{D} \subseteq 2^E$ is an intersecting family and that an intersecting family $\mathcal{F}_1 \subseteq 2^E$ is a crossing family. Therefore, the degree of generality increases from submodular functions on distributive lattices to those on intersecting families and from submodular functions on intersecting families to those on crossing families. Howerever, we have the following

Proposition 2.2 [15]:

(1) Let f be a submodular function on an intersecting family $\mathcal{F}_1$ with $\emptyset$, $E \in \mathcal{F}_1$ and $f(\emptyset) = 0$ and define

$$P(f) = \{\, x \mid x \in R^E,\ \forall X \in \mathcal{F}_1 : x(X) \le f(X)\,\}. \tag{2.10}$$

Then there exists a unique submodular system $(\mathcal{D}_1, f_1)$ on E such that

$$P(f_1) = P(f). \tag{2.11}$$

(Hence $B(f_1) = B(f)$ if $B(f) \neq \emptyset$.) Moreover, if f is integer-valued, so is f_1.

(2) Let f be a submodular function on a crossing family $\mathcal{F}_2$ with $\emptyset$, $E \in \mathcal{F}_2$ and define

$$B(f) = \{\, x \mid x \in R^E, \forall X \in \mathcal{F}_2 : x(X) \le f(X), x(E) = f(E)\,\}. \tag{2.12}$$

If $B(f) \neq \emptyset$, then there exists a unique submodular system $(\mathcal{D}_2, f_2)$ on E such that

$$B(f_2) = B(f). \tag{2.13}$$

Moreover, if f is integer-valued, so is f_2. □

We see from this proposition that considering submodular functions f on intersecting or crossing families does not extend the class of associated polyhedra $P(f)$ and $B(f)$. It should, however, be noted that for a submodular function f on a crossing family and the submodular system $(\mathcal{D}_2, \mathcal{F}_2)$ in (2) of Proposition 2.2, we do not have $P(f) = P(f_2)$ in general.

2.2. Fundamental operations on submodular systems

Consider a submodular system $(\mathcal{D}, f)$. For any vector $x \in P(f^\#)$ the polyhedron

$$B(f)^x = \{\, y \mid y \in B(f), \forall e \in E : y(e) \le x(e)\,\} \tag{2.14}$$

is the base polyhedron $B(f^x)$ associated with a submodular system $(2^E, f^x)$, where the rank function f^x is given by

$$f^x(X) = \min\{\, f(Y) - x(X - Y) \mid X \supseteq Y \in \mathcal{D}\,\} \tag{2.15}$$

for each $X \subseteq E$. The submodular system $(2^E, f^x)$ is called the *reduction* of $(\mathcal{D}, f)$ by vector x. Note that if f is integer-valued and x is integral, then f^x is integer-valued. Moreover, for any vector $x \in P(f)$ the polyhedron

$$B(f)_x = \{\, y \mid y \in B(f),\ \forall e \in E : y(e) \geq x(e) \,\} \tag{2.16}$$

is the base polyhedron $B(f_x)$ associated with a submodular system $(2^E, f_x)$, where the rank function f_x is given by

$$f_x(X) = \min\{\, f(Y) - x(Y - X) \mid X \subseteq Y \in \mathcal{D} \,\} \tag{2.17}$$

for each $X \subseteq E$. The submodular system $(2^E, f_x)$ is called the *contraction* of $(\mathcal{D}, f)$ by x. Note that if f is integer-valued and x is integral, then f_x is integer-valued. A submodular system obtained by repeated reductions and/or contractions is called a *(vector) minor* of $(\mathcal{D}, f)$.

Proposition 2.3: If vectors $x, y \in R^E$ satisfy (i) $x \leq y$, (ii) $B(f)_x \neq \emptyset$ and (iii) $B(f)^y \neq \emptyset$, then we have $B(f)_x^y \ (= (B(f)_x)^y = (B(f)^y)_x) \neq \emptyset$. □

We see from (2.14), (2.16) and the duality (2.8) that the contraction corresponds to the reduction of the dual supermodular system $(\bar{\mathcal{D}}, f^{\#})$.

For any $A \in \mathcal{D}$ define

$$\mathcal{D}^A = \{\, X \mid X \subseteq A, \quad X \in \mathcal{D} \,\} \tag{2.18}$$

and a submodular function $f^A : \mathcal{D}^A \to R$ by

$$f^A(X) = f(X) \qquad (X \in \mathcal{D}^A). \tag{2.19}$$

The submodular system $(\mathcal{D}^A, f^A)$ on A is called the *reduction* (or *restriction*) *of* $(\mathcal{D}, f)$ *to* A and denoted by $(\mathcal{D}, f) \cdot A$ or $(\mathcal{D}, f) - (E - A)$. Also define

$$\mathcal{D}_A = \{\, X - A \mid X \supseteq A, \quad X \in \mathcal{D} \,\} \tag{2.20}$$

and a submodular function $f_A : \mathcal{D}_A \to R$ by

$$f_A(X) = f(X \cup A) - f(A) \qquad (X \in \mathcal{D}_A). \tag{2.21}$$

The submodular system $(\mathcal{D}_A, f_A)$ on $E - A$ is called the *contraction of* $(\mathcal{D}, f)$ *by* A and denoted by $(\mathcal{D}, f)/A$ or $(\mathcal{D}, f) \times (E - A)$.

Any vector $x \in R^E$ can be considered as a modular function on 2^E through (2.3). The submodular system $(\mathcal{D}, f + x)$ is called the *translation* of $(\mathcal{D}, f)$ by $x \in R^E$. We have

$$P(f + x) = P(f) + \{x\}, \qquad B(f + x) = B(f) + \{x\}, \tag{2.22}$$

where the sum in the right-hand side of each equation of (2.22) denotes the vector sum. It should be noted that the combinatorial structures of the submodular polyhedron and the base polyhedron are invariant under translation, whereas the monotonicity of the rank function is not invariant and any rank function can be made monotone increasing by an appropriate translation. Therefore, translation-invariant results in polymatroid theory such as the polymatroid intersection theorem [4] can easily be extended to submodular system (see [16]).

Consider a submodular system $(\mathcal{D}, f)$. Recall that $\mathcal{D} \subseteq 2^E$ is a distributive lattice with $\emptyset$, $E \in \mathcal{D}$. There uniquely exist a partition $\Pi = \{A_1, A_2, \ldots, A_k\}$ of E and a partial order $\preceq$ on Π such that $X \in \mathcal{D}$ if and only if X is expressed as

$$X = \bigcup \{A_i \mid A_i \in \mathcal{I}\} \tag{2.23}$$

for some ideal $\mathcal{I}$ of the partially ordered set (poset) $(\Pi, \preceq)$ ([1]). (An ideal $\mathcal{I}$ of the poset $(\Pi, \preceq)$ is a subset of Π such that for any A_i, $A_j \in \Pi$ with $A_i \preceq A_j \in \mathcal{I}$ we have $A_i \in \mathcal{I}$.) Let E' be a k-element set $\{a_1, a_2, \ldots, a_k\}$ and make a_i correspond to A_i for each $i = 1, 2, \ldots, k$. By this correspondance we obtain a distributive lattice $\mathcal{D}' \subseteq 2^{E'}$ from $\mathcal{D}$ and a submodular function $f' : \mathcal{D}' \to R$ from f. We call the pair $(\mathcal{D}', f')$ the *simplification* of $(\mathcal{D}, f)$. When the partition Π consists of singletons A_i, i.e., $|A_i| = 1$ $(i = 1, 2, \ldots, k)$, we call $\mathcal{D}$ and $(\mathcal{D}, f)$ *simple*. A simple distributive lattice $\mathcal{D} \subseteq 2^E$ (with $\emptyset, E \in \mathcal{D}$) is formed by the set of ideals of a poset $\mathcal{P} = (E, \preceq)$ and is denoted by $\mathbf{2}^{\mathcal{P}}$.

An example of a base polyhedron is given as follows. Suppose that we are given a real c and vectors $\ell, u \in R^E$ such that $\ell \leq u$ and $\ell(E) \leq c \leq u(E)$. Then the polyhedron $B_0 = \{ x \mid x \in R^E, \; x(E) = c, \; \ell \leq x \leq u \}$ is a base polyhedron. For B_0 is the minor $B(f)^u_\ell$ of $B(f)$ such that $\mathcal{D} = \{\emptyset, E\}$ and $f(\emptyset) = 0$, $f(E) = c$. If c is an integer and ℓ, u are integral vectors, B_0 is an integral base polyhedron. Polyhedron B_0 is a typical set of feasible solutions of a resource allocation problem (see [22]).

2.3. Fundamental properties of base polyhedra

The following propositions show fundamental structural properties of base polyhedra.

Proposition 2.4 [16]: For a submodular system $(\mathcal{D}, f)$ we have:
(1) The base polyhedron $B(f)$ is pointed if and only if $\mathcal{D}$ is simple, i.e., $\mathcal{D} = \mathbf{2}^{\mathcal{P}}$ for some poset $\mathcal{P} = (E, \preceq)$.
(2) The base polyhedron $B(f)$ is bounded if and only if $\mathcal{D}$ is simple and complemented, i.e., $\mathcal{D} = 2^E$. □

Proposition 2.5 (The extreme point theorem) [19] (also see [4], [33], [28], for $\mathcal{D} = 2^E$): For a simple submodular system $(\mathcal{D}, f)$ a base $x \in B(f)$ is an extreme point of $B(f)$ if and only if for a maximal chain

$$\mathcal{C} : \quad \emptyset = S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_n = E \tag{2.24}$$

in $\mathcal{D}$ we have

$$x(S_i - S_{i-1}) = f(S_i) - f(S_{i-1}) \qquad (i = 1, 2, \ldots, n). \tag{2.25}$$

□

For a simple submodular system $(\mathcal{D}, f)$ the base polyhedron $B(f)$ is expressed as the vector sum of the convex hull $Q(f)$ of the extreme points of $B(f)$ and the recession cone (or the characteristic cone) $C(f)$ of $B(f)$:

$$B(f) = Q(f) + C(f), \tag{2.26}$$

where

$$C(f) = \{\, x \mid x \in R^E, \ \forall X \in \mathcal{D} : x(X) \leq 0, \ x(E) = 0 \,\}. \tag{2.27}$$

Let $\mathcal{P} = (E, \preceq)$ be the poset such that $\mathcal{D} = \mathbf{2}^{\mathcal{P}}$. Also let $G = (E, A(\mathcal{P}))$ be the (directed) graph representing the Hasse diagram of $\mathcal{P} = (E, \preceq)$, i.e., E is the vertex set of G and $A(\mathcal{P})$ is the arc set of G such that (e, e') is an arc in $A(\mathcal{P})$ if and only if $e' \prec e$ and there is no $e'' \in E$ such that $e' \prec e'' \prec e$.

Proposition 2.6 (The extreme ray theorem) [36]: The set of all the extreme rays of the cone $C(f)$ in (2.26) is given by

$$\{\, \chi_e - \chi_{e'} \mid (e, e') \in A(\mathcal{P}) \,\}, \tag{2.28}$$

where $A(\mathcal{P})$ is the arc set of the Hasse diagram $G = (E, A(\mathcal{P}))$ of the poset $\mathcal{P} = (E, \preceq)$, and χ_e is the unit vector in R^E such that $\chi_e(e) = 1$. □

For any base $x \in B(f)$ and $e \in E$ we define

$$\mathrm{dep}(x, e) = \bigcap \{\, X \mid e \in X \in \mathcal{D}, \quad x(X) = f(X) \,\}. \tag{2.29}$$

The function dep: $B(f) \times E \to \mathcal{D}$ is called the *dependence function* associated with $(\mathcal{D}, f)$ [11]. For any $e' \in \mathrm{dep}(x, e) - \{e\}$ we also define

$$\tilde{c}(x, e, e') = \min\{\, f(X) - x(X) \mid e \in X \in \mathcal{D}, \ e' \notin X \,\}. \tag{2.30}$$

We can easily see from the definition that $\tilde{c}(x, e, e') > 0$ for $e' \in \mathrm{dep}(x, e) - \{e\}$, and $\tilde{c}(x, e, e')$ is called the *exchange capacity* associated with $x \in B(f)$, $e \in E$ and $e' \in \mathrm{dep}(x, e) - \{e\}$. It should be noted that for any $d \in R$ with $0 \leq d \leq \tilde{c}(x, e, e')$ we have

$$x + d(\chi_e - \chi_{e'}) \in B(f). \tag{2.31}$$

The transformation of $x \in B(f)$ into $x + d(\chi_e - \chi_{e'}) \in B(f)$ is called an *elementary transformation* of base x.

Proposition 2.7: Given any two bases $x,\ y \in B(f)$, x can be transformed into y by (at most $\lceil \frac{1}{4}|E|^2 \rceil$) repeated elementary transformations so that $x(e)$ with $x(e) > y(e)$ monotonically decreases and $x(e)$ with $x(e) < y(e)$ monotonically increases. □

For a base $x \in B(f)$ define a graph $G_x = (E, A_x)$ with vertex set E and arc set A_x as follows.

$$A_x = \{\, (e, e') \mid e \in E, \quad e' \in \mathrm{dep}(x, e) - \{e\} \,\}. \tag{2.32}$$

We call G_x the *exchangeability graph* associated with base x. Note that graph G_x with selfloops appropriately attached to vertices is transitively closed.

For any base $x \in B(f)$ the *tangent cone* of $B(f)$ at x, denoted by $\mathrm{TC}(B(f), x)$, is defined by

$$\mathrm{TC}(B(f), x) = \{\, \lambda y \mid \lambda \geq 0, \ \ y \in R^E, \ \ x + y \in B(f) \,\} \tag{2.33}$$

The following proposition is essentially subsumed in [11, Lemma 9].

Proposition 2.8: The tangent cone $\mathrm{TC}(B(f), x)$ is generated by the set of the following vectors:

$$\chi_e - \chi_{e'} \qquad\qquad ((e, e') \in A_x), \tag{2.34}$$

where A_x is the arc set of the exchangeability graph G_x. In other words, for any vector $y \in \mathrm{TC}(B(f), x)$ there exist some nonnegative coefficients $\lambda(e, e')$ $((e, e') \in A_x)$ such that

$$y = \sum \{\, \lambda(e, e')(\chi_e - \chi_{e'}) \mid (e, e') \in A_x \,\}. \tag{2.35}$$

□

The above characterization of the tangent cone plays an important role in the optimality conditions for a certain class of optimization problems on base polyhedra, which will be discussed in the subsequent sections.

Let $G_x^0 = (E, A_x^0)$ be the graph whose arc set A_x^0 is minimal with the property that G_x is the transitive closure of G_x^0 with possible selfloops deleted. Proposition 2.8 is apparently strengthened by replacing A_x appearing in Proposition 2.8 by A_x^0. When $(\mathcal{D}, f)$ is simple and x is an extreme point of $B(f)$, G_x is acyclic and this strengthened version of Proposition 2.8 implies that $\chi_e - \chi_{e'}$ $((e, e') \in A_x^0)$ are exactly the extreme rays of the tangent cone $\mathrm{TC}(B(f), x)$ (cf. [2], [36], [37]).

3. Linear Optimization

Let $(\mathcal{D}, f)$ be a simple submodular system with $\mathcal{D} = 2^{\mathcal{P}}$ and $\mathcal{P} = (E, \preceq)$ and let $w : E \to R$ be an arbitrary function on E. Consider the following linear optimization problem

$$P_0: \quad \text{Minimize} \quad \sum_{e \in E} w(e)x(e)$$
$$\text{subject to} \quad x \in B(f). \tag{3.1}$$

Proposition 3.1 [19]: Problem P_0 has a finite optimal solution if and only if w is a monotone nondecreasing function from $\mathcal{P} = (E, \preceq)$ to $(R, \leq)$, where $\mathcal{P}$ is the poset such that $\mathcal{D} = \mathbf{2}^{\mathcal{P}}$. □

Suppose that Problem P_0 has a finite optimal solution and let the distinct values of $w(e)$ $(e \in E)$ be given by

$$w_1 < w_2 < \cdots < w_p. \tag{3.2}$$

Also define

$$A_i = \{\, e \mid e \in E, \quad w(e) \leq w_i \,\} \qquad (i = 1, 2, \ldots, p), \tag{3.3}$$
$$A_0 = \emptyset. \tag{3.4}$$

Note that from the assumption and Proposition 3.1 we have $A_i \in \mathcal{D}$ for each $i = 0, 1, \ldots, p$.

Proposition 3.2 (Greedy algorithm) [19] (also see [4], [28] for $\mathcal{D} = 2^E$): Let

$$\mathcal{C}: \quad S_0 = \emptyset \subsetneq S_1 \subsetneq \cdots \subsetneq S_n = E \tag{3.5}$$

be a maximal chain of $\mathcal{D}$ which contains A_i $(i = 0, 1, \ldots, p)$ in it, and define a vector $x \in R^E$ by

$$x(S_i - S_{i-1}) = f(S_i) - f(S_{i-1}) \qquad (i = 1, 2, \ldots, n). \tag{3.6}$$

Then x is an optimal solution of Problem P_0. □

It should be noted that the optimal solution x given in Proposition 3.2 is an extreme point of base polyhedron $B(f)$ and that any optimal extreme-point solution is given by (3.5) and (3.6) by appropriately choosing a maximal chain $\mathcal{C}$ containing A_i $(i = 0, 1, \ldots, p)$. When f is integer-valued, the optimal solution x given in Proposition 3.2 is also integral. Therefore, the greedy algorithm finds an optimal solution of both continuous and discrete linear optimization problems on base polyhedra. It should, however, be noted that if f is real- or rational-valued and x is restricted to integral vectors in (3.1), then the problem becomes difficult. The structure of the set of integral points in a real or rational base polyhedron has not been elucidated at all. The difficulty is that f does not remain submodular by rounding. The rounding problems for submodular functions and base polyhedra are left for future research.

4. Continuous Nonlinear Optimization

In this section we consider a submodular system $(\mathcal{D}, f)$ with a real-valued (or rational-valued) submodular function f. The underlying totally ordered additive group is assumed to be the set R of reals (or the set Q of rationals). The problems to be considered in Sections 4 and 5 include as special cases most of the so-called *resource allocation problems* hitherto investigated in the literature. The readers should be referred to the book [22] by T. Ibaraki and N. Katoh for resource allocation problems and related topics.

4.1. Separable convex optimization

For each $e \in E$ let $w_e : R \to R$ be a real-valued convex function on R, and consider the following problem

$$P_1: \quad \text{Minimize} \quad \sum_{e \in E} w_e(x(e))$$
$$\text{subject to} \quad x \in B(f). \tag{4.1}$$

The author [14] considered Problem P_1 where for each $e \in E$ $w_e(x(e))$ is a quadratic function given by $x(e)^2/w(e)$ with a positive real weight $w(e)$ and f is a polymatroid rank function. H. Groenevelt [20] also considered Problem P_1 where f is a rank function of a polymatroid. It is almost straightforward to generalize the result of [14] and [20] to Problem P_1 for a general submodular system.

Proposition 4.1 ([20] (also see [14])): A base $x \in B(f)$ is an optimal solution of Problem P_1 if and only if for each $e \in E$ and $e' \in \mathrm{dep}(x, e) - \{e\}$ we have

$$w_e{}^+(x(e)) \geq w_{e'}{}^-(x(e')), \tag{4.2}$$

where $w_e{}^+$ denotes the right derivative of w_e and $w_{e'}{}^-$ the left derivative of $w_{e'}$.
(Proof) *"If" part :* Suppose that (4.2) holds for each $e \in E$ and $e' \in \mathrm{dep}(x, e) - \{e\}$. From Proposition 2.8, for any base $z \in B(f)$ there exist some nonnegative coefficients $\lambda(e, e')$ ($(e, e') \in A_x$) such that

$$z = x + \sum \{ \lambda(e, e')(\chi_e - \chi_{e'}) \mid (e, e') \in A_x \}. \tag{4.3}$$

For each $e \in E$ define

$$\bar{w}_e(x(e)) = \max\{ w_{e'}{}^-(x(e')) \mid e' \in \mathrm{dep}(x, e) \}. \tag{4.4}$$

We see from (4.2) and (4.4) that

$$w_e{}^-(x(e)) \leq \bar{w}_e(x(e)) \leq w_e{}^+(x(e)) \qquad (e \in E), \tag{4.5}$$
$$\bar{w}_e(x(e)) \geq \bar{w}_{e'}(x(e')) \qquad ((e, e') \in A_x). \tag{4.6}$$

From (4.3) $\sim$ (4.6) and the convexity of w_e $(e \in E)$ we have

$$\begin{aligned}\sum_{e\in E} w_e(z(e)) &= \sum_{e\in E} w_e(x(e) + \partial\lambda(e)) \\ &\geq \sum_{e\in E}\{w_e(x(e)) + \partial\lambda(e)\bar{w}_e(x(e))\} \\ &= \sum_{e\in E} w_e(x(e)) + \sum_{(e,e')\in A_x} \lambda(e,e')(\bar{w}_e(x(e)) - \bar{w}_{e'}(x(e'))) \\ &\geq \sum_{e\in E} w_e(x(e)),\end{aligned} \tag{4.7}$$

where $\partial\lambda : E \to R$ is defined by

$$\partial\lambda(e) = \sum_{(e,e')\in A_x} \lambda(e,e') - \sum_{(e',e)\in A_x} \lambda(e',e) \qquad (e \in E). \tag{4.8}$$

This implies the optimality of x.

"Only if" part : Suppose that for a base $x \in B(f)$ there exist $e \in E$ and $e' \in \mathrm{dep}(x,e) - \{e\}$ such that

$$w_e^+(x(e)) < w_{e'}^-(x(e')). \tag{4.9}$$

Then for a sufficiently small positive number ε we have

$$w_e(x(e)) + w_{e'}(x(e')) > w_e(x(e) + \varepsilon) + w_{e'}(x(e') - \varepsilon), \tag{4.10}$$

$$x + \varepsilon(\chi_e - \chi_{e'}) \in B(f). \tag{4.11}$$

Therefore, x is not an optimal solution. □

For each $e \in E$ and $\xi \in R$ define the interval

$$J_e(\xi) = [w_e^-(\xi), w_e^+(\xi)]. \tag{4.12}$$

$J_e(\xi)$ is the subdifferential of w_e at ξ. Conversely, for each $e \in E$ and $\eta \in R$ define

$$I_e(\eta) = \{\, \xi \mid \xi \in R, \quad \eta \in J_e(\xi)\,\}. \tag{4.13}$$

Because of the convexity of w_e, $I_e(\eta)$, if nonempty, is an interval in R and we express it as

$$I_e(\eta) = [i_e^-(\eta), i_e^+(\eta)]. \tag{4.14}$$

In the following we assume for simplicity that $I_e(\eta) \neq \emptyset$ for every $\eta \in R$, which garantees the existence of an optimal solution even if $B(f)$ is unbounded. When $B(f)$ is bounded, there is no loss of generality with this assumption.

By adapting the algorithm in [14] and [20] an efficient algorithm for solving Problem P_1 is given as follows, where x^* is the output vector giving an optimal solution.

Algorithm A1 by Decomposition

Step 1 : Choose $\eta \in R$ such that

$$\sum_{e\in E} i_e{}^-(\eta) \le f(E) \le \sum_{e\in E} i_e{}^+(\eta). \tag{4.15}$$

Step 2 : Find a base $x \in B(f)$ such that for each e, $e' \in E$

(1) if $w_e{}^+(x(e)) < \eta$ and $w_{e'}{}^-(x(e')) > \eta$, then we have $e' \notin \mathrm{dep}(x, e)$,

(2) if $w_e{}^+(x(e)) < \eta$, $w_{e'}{}^-(x(e')) = \eta$ and $e' \in \mathrm{dep}(x, e)$, then for any $\varepsilon > 0$ we have $w_{e'}{}^-(x(e') - \varepsilon) < \eta$, i.e., $x(e') = i_{e'}{}^-(\eta)$,

(3) if $w_e{}^+(x(e)) = \eta$, $w_{e'}{}^-(x(e')) > \eta$ and $e' \in \mathrm{dep}(x, e)$, then for any $\varepsilon > 0$ we have $w_e{}^+(x(e) + \varepsilon) > \eta$, i.e., $x(e) = i_e{}^+(\eta)$.

Put

$$E_- = \bigcup\{\, \mathrm{dep}(x, e) \mid e \in E, \quad w_e{}^+(x(e)) < \eta \,\}, \tag{4.16}$$

$$E_+ = \bigcup\{\, \mathrm{dep}^{\#}(x, e) \mid e \in E, \quad w_e{}^-(x(e)) > \eta \,\}, \tag{4.17}$$

$$E_0 = E - (E_+ \cup E_-), \tag{4.18}$$

where

$$\mathrm{dep}^{\#}(x, e) = \bigcap\{\, X \mid e \in X \in \bar{\mathcal{D}}, \quad x(X) = f^{\#}(X) \,\}. \tag{4.19}$$

Put $x^*(e) = x(e)$ for each $e \in E_0$.

Step 3: If $E_- \neq \emptyset$, then apply the present algorithm recursively to the problem with E and f, respectively, replaced by E_- and f^{E_-} and the base polyhedron associated with the reduction $(\mathcal{D}, f) \cdot E_-$. Also, if $E_+ \neq \emptyset$, then apply the present algorithm recursively to the problem with E and f, respectively, replaced by E_+ and f_{E_+} and the base polyhedron associated with the contraction $(\mathcal{D}, f) \times E_+$. (End)

The validity of the algorithm easily follows from Proposition 4.1.

It should be noted that if w_e is strictly convex for each $e \in E$, then conditions (2) and (3) in Step 2 are always satisfied, so that we have only to consider condition (1).

The above algorithm lays a basis for the algorithms for the other problems to be considered in the subsequent sections.

4.2. Lexicographically optimal base

For each $e \in E$ let h_e be a continuous and monotone increasing function from R onto R. For any vector $x \in R^E$ we denote by $T(x)$ the sequence of the components $x(e)$ $(e \in E)$ of x arranged in order of increasing magnitude, i.e., $T(x) = (x(e_1), x(e_2), \ldots, x(e_n))$ with $x(e_1) \leq x(e_2) \leq \cdots \leq x(e_n)$, where $\{e_1, e_2, \ldots, e_n\} = E$ and $|E| = n$.

Consider the following problem

$$P_2: \quad \text{Lexicographically maximize} \quad T((h_e(x(e)) : e \in E)) \\ \text{subject to} \quad x \in B(f). \tag{4.20}$$

We call an optimal solution of Problem P_2 a *lexicographically optimal base* of $(\mathcal{D}, f)$ *with respect to functions* h_e $(e \in E)$. When $h_e(x(e))$ is a linear function expressed as $x(e)/w(e)$ with $w(e) > 0$ for each $e \in E$, such a lexicographically optimal base is called a *lexicographically optimal base with respect to the weight vector* $w = (w(e) : e \in E)$ [14]. It is a generalization of the concept of (lexicographically) optimal flow introduced by N. Megiddo [29] concerning multiple-source multiple-sink networks.

Proposition 4.2 (cf. [14]): Let x be a base in $B(f)$. Define a vector $\eta \in R^E$ by

$$\eta(e) = h_e(x(e)) \qquad (e \in E) \tag{4.21}$$

and let the distinct numbers of $\eta(e)$ $(e \in E)$ be given by

$$\eta_1 < \eta_2 < \cdots < \eta_p. \tag{4.22}$$

Also, define

$$S_i = \{\, e \mid e \in E, \quad \eta(e) \leq \eta_i \,\} \qquad (i = 1, 2, \ldots, p). \tag{4.23}$$

Then the following are equivalent:

(i) x is a lexicographically optimal base of $(\mathcal{D}, f)$ with respect to h_e $(e \in E)$.

(ii) $S_i \in \mathcal{D}$ and $x(S_i) = f(S_i)$ $(i = 1, 2, \ldots, p)$.

(iii) $\mathrm{dep}(x, e) \subseteq S_i$ $(e \in S_i,\ i = 1, 2, \ldots, p)$.

(iv) x is an optimal solution of Problem P_1 where for each $e \in E$ the derivative of w_e coincides with h_e.

(Proof) The equivalence, (ii) $\Longleftrightarrow$ (iii), can be proved by the direct adaptation of the proof in [14]. Also, the equivalence, (ii), (iii) $\Longleftrightarrow$ (iv), follows from Proposition 4.1. We show the equivalence, (i) $\Longleftrightarrow$ (ii) $\sim$ (iv).

(i) $\Longrightarrow$ (iii): Suppose (i). If for some $i \in \{1, 2, \ldots, p\}$ and $e \in S_i$ there exists $e' \in \mathrm{dep}(x, e) - S_i$, then for a sufficiently small positive number ε the vector

$$y = x + \varepsilon(\chi_e - \chi_{e'}) \tag{4.24}$$

is a base in $B(f)$ and $T((h_e(y(e)) : e \in E))$ is lexicographically greater than $T((h_e(x(e)) : e \in E))$. This contradicts (i). So, (iii) holds.

(ii), (iii) $\Longrightarrow$ (i): Suppose (ii) (and (iii)). Let $\bar{x}$ be an arbitrary base such that $T((h_e(\bar{x}(e)) : e \in E))$ is lexicographically greater than or equal to $T((h_e(x(e)) : e \in E))$. Define a vector $\bar{\eta} \in R^E$ by

$$\bar{\eta}(e) = h_e(\bar{x}(e)) \qquad (e \in E). \tag{4.25}$$

Also define $S_0 = \emptyset$. We show by induction on i that

$$x(e) = \bar{x}(e) \qquad (e \in S_i) \tag{4.26}$$

for $i = 0, 1, \ldots, p$, from which the optimality of x follows. For $i = 0$ (4.26) trivially holds. So, suppose that (4.26) holds for some $i = i_0 < p$. Since $T(\bar{\eta})$ is lexicographically greater than or equal to $T(\eta)$, we have from (4.22) and (4.23)

$$\bar{\eta}(e) \geq \eta(e) = \eta_{i_0+1} \qquad (e \in S_{i_0+1} - S_{i_0}). \tag{4.27}$$

From (4.27) and the monotone increasingness of h_e $(e \in E)$,

$$\bar{x}(e) \geq x(e) \qquad (e \in S_{i_0+1} - S_{i_0}). \tag{4.28}$$

Since $\bar{x} \in B(f)$, it follows from (4.26) with $i = i_0$, (4.28) and assumption (ii) that

$$f(S_{i_0+1}) \geq \bar{x}(S_{i_0+1}) \geq x(S_{i_0+1}) = f(S_{i_0+1}). \tag{4.29}$$

From (4.28) and (4.29) we have $\bar{x}(e) = x(e)$ $(e \in S_{i_0+1})$. □

From Proposition 4.2 we can find a lexicographically optimal base by using Algorithm A1 given in Section 4.1.

For a vector $x \in R^E$ define $T^*(x)$ to be the sequence of the components $x(e)$ $(e \in E)$ of x arranged in order of decreasing magnitude. We call a base $x \in B(f)$ which lexicographically minimizes $T^*((h_e(x(e)) : e \in E))$ a *co-lexicographically optimal base of* $(\mathcal{D}, f)$ with respect to h_e $(e \in E)$.

Proposition 4.3: x is a lexicographically optimal base of $(\mathcal{D}, f)$ with respect to h_e $(e \in E)$ if and only if it is a co-lexicographically optimal base of $(\mathcal{D}, f)$ with respect to h_e $(e \in E)$.
(Proof) Using $\eta(e)$ $(e \in E)$ and η_i $(i = 1, 2, \ldots, p)$ appearing in Proposition 4.2, define

$$S_i^* = \{\, e \mid e \in E, \quad \eta(e) \geq \eta_{p-i+1} \,\} \qquad (i = 1, 2, \ldots, p). \tag{4.30}$$

Also define $S_0 = \emptyset = S_0^*$. Since $S_i^* = E - S_{p-i}$ $(i = 0, 1, \ldots, p)$ and $x(E) = f(E)$, we can easily see that for a base $x \in B(f)$ x satisfies (ii) of Proposition 4.2 if and only if x satisfies

(ii*) $S_i^* \in \bar{\mathcal{D}}$ and $x(S_i^*) = f^{\#}(S_i^*)$ $(i = 1, 2, \ldots, p)$.

Consequently, the present proposition follows from Proposition 4.2 and the duality shown in Proposition 2.1. □

4.3. Weighted max-min/min-max problems

For each $e \in E$ let $h_e : R \to R$ be a right-continuous and monotone nondecreasing function such that $\lim_{\xi \to +\infty} h_e(\xi) = +\infty$ and $\lim_{\xi \to -\infty} h_e(\xi) = -\infty$.

Consider the following max-min problem with nonlinear weight functions h_e ($e \in E$):

$$P_* : \quad \text{Maximize} \quad \min_{e \in E} h_e(x(e)) \qquad \text{subject to} \quad x \in B(f). \tag{4.31}$$

For each $e \in E$ let $w_e : R \to R$ be a convex function whose right derivative ${w_e}^+$ is given by h_e.

Proposition 4.4: Consider Problem P_1 of (4.1) with w_e ($e \in E$) defined as above. Let x be an optimal solution of Problem P_1. Then x is an optimal solution of Problem P_*.
(Proof) Define

$$\eta_1 = \min\{ h_e(x(e)) \mid e \in E \}, \tag{4.32}$$

$$S_1 = \{ e \mid e \in E, \quad h_e(x(e)) = \eta_1 \}, \tag{4.33}$$

$$S_1^* = \bigcup\{ \text{dep}(x, e) \mid e \in S_1 \}. \tag{4.34}$$

We have from (4.34)

$$x(S_1^*) = f(S_1^*). \tag{4.35}$$

It follows from Proposition 4.1 that

$${w_e}^-(x(e)) \le \eta_1 \qquad (e \in S_1^*). \tag{4.36}$$

If there were a base $y \in B(f)$ such that

$$\eta_1 < \min\{ h_e(y(e)) \mid e \in E \}. \tag{4.37}$$

Then, from (4.32) $\sim$ (4.37) we would have

$$x(e) < y(e) \quad (e \in S_1), \qquad x(e) \le y(e) \quad (e \in S_1^* - S_1), \tag{4.38}$$

since $h_e = {w_e}^+$. Hence, from (4.35) and (4.38), $f(S_1^*) = x(S_1^*) < y(S_1^*)$, which contradicts the fact that $y \in B(f)$. □

We see from the above proof of Proposition 4.4 that Algorithm A1 can be simplified for solving Problem P_* as follows. We may put $x^*(e) = x(e)$ for $e \in E_0 \cup E_+$ in Step 2 and apply Algorithm A1 recursively to the problem on E_- but not to the one on E_+ in Step 3(cf. [23]).

Moreover, consider the following min-max problem

$$P^* : \quad \text{Minimize} \quad \max_{e \in E} h_e(x(e)) \qquad \text{subject to} \quad x \in B(f). \tag{4.39}$$

We see from Proposition 4.4 and the duality shown in Proposition 2.1 that an optimal solution of Problem P_1 obtained by Algorithm A1, where the left derivative $w_e{}^-$ of w is given by h_e, is also an optimal solution of Problem P^*.

Problems P_* and P^* are sometimes called the *sharing problems* in the literature ([3], [23]). An algebraic generalization of the shairing problem is given by U. Zimmermann [38].

4.4. The continuous fair resource allocation problem

Let $g : R^2 \to R$ be a function such that $g(u, v)$ is monotone nondecreasing in u and monotone nonincreasing in v. Also, for each $e \in E$ let h_e be a continuous monotone nondecreasing function from R onto R. Consider

$$\begin{aligned} P_3: \quad & \text{Minimize} \quad g(\max_{e\in E} h_e(x(e)), \min_{e\in E} h_e(x(e))) \\ & \text{subject to} \quad x \in B(f). \end{aligned} \tag{4.40}$$

We call Problem P_3 the *continuous fair resource allocation problem with submodular constraints.*

Using the same functions h_e $(e \in E)$ appearing in (4.40), let us consider Problem P_* and P^* described by (4.31) and (4.39), respectively. Denote the optimal values of the objective functions of Problems P_* and P^* by v_* and v^*, respectively, and define vectors ℓ, $u \in (R \cup \{-\infty, +\infty\})^E$ by

$$\ell(e) = \inf\{\, \alpha \mid \alpha \in R, \quad h_e(\alpha) \geq v_*\} \qquad (e \in E), \tag{4.41}$$

$$u(e) = \sup\{\, \alpha \mid \alpha \in R, \quad h_e(\alpha) \leq v^*\} \qquad (e \in E). \tag{4.42}$$

Proposition 4.5: Suppose values v_* and v^* and vectors ℓ and u are defined as above. Then we have $v_* \leq v^*$ and $\ell \leq u$. Moreover, $B(f)_\ell^u$ is nonempty and any $x \in B(f)_\ell^u$ is an optimal solution of Problem P_3.
(Proof) Let x^* and x_* be, respectively, optimal solutions of Problem P^* and P_*. If $v_* > v^*$, then $x^*(e) \leq u(e) < \ell(e) \leq x_*(e)$ $(e \in E)$, which contradicts the fact that $x^*(E) = f(E) = x_*(E)$. Therefore, we have $v_* \leq v^*$. This implies $\ell \leq u$. Moreover, since $x_* \in B(f)_\ell$, $x^* \in B(f)^u$ and $\ell \leq u$, from Proposition 2.3 we have $B(f)_\ell^u \neq \emptyset$. For any $x \in B(f)_\ell^u$ and $y \in B(f)$ we have

$$\begin{aligned} & g(\max_{e\in E} h_e(y(e)), \min_{e\in E}(y(e))) \\ & \geq g(v^*, v_*) \\ & = g(\max_{e\in E} h_e(x(e)), \min_{e\in E} h_e(x(e))), \end{aligned} \tag{4.43}$$

due to the monotonicity of g. This shows that any $x \in B(f)_\ell^u$ is an optimal solution of Problem P_3. □

The continuous fair resource allocation problem P_3 can thus be solved by using Algorithm A1 given in Section 4.1.

5. Discrete Nonlinear Optimization

In this section we consider the optimization problems treated in the preceding section in the case where the values of the variables are restricted to integers. We suppose that the underlying totally ordered additive group is the set Z of integers, the rank function f of submodular system $(\mathcal{D}, f)$ is integer-valued, and the base polyhedron $B(f)$ is given by

$$B(f) = \{\, x \mid x \in Z^E, \quad \forall X \in \mathcal{D} : x(X) \leq f(X), x(E) = f(E) \,\}. \tag{5.1}$$

We denote the base polyhedron (5.1) by $B_Z(f)$, emphasizing that the underlying totally ordered additive group is Z. We also define for the set R of reals

$$B_R(f) = \{\, x \mid x \in R^E, \quad \forall X \in \mathcal{D} : x(X) \leq f(X), x(E) = f(E) \,\}. \tag{5.2}$$

Here, $B_R(f)$ is the base polyhedron associated with $(\mathcal{D}, f)$ when we consider R as the underlying totally ordered additive group. It can be seen from Propositions 2.5 and 2.6 that polyhedron $B_R(f)$ is the convex hull, in R^E, of $B_Z(f)$.

5.1. Separable convex optimization

For each $e \in E$ let $\hat{w}_e$ be a real-valued function on Z such that the piecewise linear extension, denoted by w_e, of $\hat{w}_e$ on R is a convex function, where $w_e(\xi) = \hat{w}_e(\xi)$ for $\xi \in Z$ and w_e restricted on each unit interval $[\xi, \xi+1]$ $(\xi \in Z)$ is a linear function. Let us consider

$$IP_1: \quad \text{Minimize} \quad \sum_{e \in E} \hat{w}_e(x(e)) \qquad \text{subject to} \quad x \in B_Z(f). \tag{5.3}$$

Also consider the continuous version of IP_1:

$$P_1: \quad \text{Minimize} \quad \sum_{e \in E} w_e(x(e)) \qquad \text{subject to} \quad x \in B_R(f). \tag{5.4}$$

Proposition 5.1(cf. [20]): There exists an integral optimal solution for Problem P_1 described by (5.4).
(Proof) Problem P_1 can be solved by Algorithm A1. Because of the definition of w_e $(e \in E)$ and the integrality of $B_R(f)$, an integral optimal solution of P_1 can be obtained through Algorithm A1 by choosing an integral base x in Step 2 of Algorithm A1. □

An integral optimal solution of P_1 is also an optimal solution of Problem IP_1 (see [7], [20]). An incremental algorithm is also given in [7].

5.2. Weighted max-min/min-max problems

For each $e \in E$ let $\hat{h}_e : Z \to R$ be a monotone nondecreasing function on Z such that $\lim_{\xi \to +\infty} \hat{h}_e(\xi) = +\infty$ and $\lim_{\xi \to -\infty} \hat{h}_e(\xi) = -\infty$ for each $e \in E$. Consider

$$IP_* : \quad \text{Minimize} \quad \min_{e \in E} \hat{h}_e(x(e)) \qquad \text{subject to} \quad x \in B_Z(f). \tag{5.5}$$

For each $e \in E$ let $w_e : R \to R$ be a piecewise-linear convex function such that its right derivative $w_e{}^+$ satisfies

$$w_e{}^+(\xi) = \hat{h}_e(\xi) \qquad (\xi \in Z) \tag{5.6}$$

and w_e is linear on each unit interval $[\xi, \xi + 1]$ $(\xi \in Z)$.

Proposition 5.2: Let x_* be an integral optimal solution of Problem P_1 with w_e $(e \in E)$ defined as above. Then x_* is an optimal solution of Problem IP_* of (5.5).
(Proof) For each $e \in E$ let $h_e : R \to R$ be a right-continuous piecewise-constant nondecreasing function such that $h_e(\xi) = \hat{h}_e(\xi)$ $(\xi \in Z)$ and $h_e(\eta) = \hat{h}_e(\xi)$ $(\eta \in [\xi, \xi + 1), \xi \in Z)$. It follows from Proposition 4.4 that an integral optimal solution of Problem P_1 with w_e $(e \in E)$ defined by (5.6) is an integral optimal solution of Problem P_* of (4.31) with h_e $(e \in E)$ defined as above. Therefore, x_* is an optimal solution of Problem IP_*. □

The reduction of Problem IP_* to Problem P_1 was also communicated by N. Katoh [26]. A direct algorithm for Problem IP_* is given in [18].

The weighted min-max problem

$$IP^* : \quad \text{Minimize} \quad \max_{e \in E} \hat{h}_e(x(e)) \qquad \text{subject to} \quad x \in B_Z(f) \tag{5.7}$$

can be solved similarly in a dual form.

5.3. The discrete fair resource allocation problem [18]

We consider the discrete version of the continuous fair resource allocation problem P_3 treated in Section 4.5.

Let $g : R^2 \to R$ be a function such that $g(u, v)$ is monotone nondecreasing in u and monotone nonincreasing in v. Also, for each $e \in E$ let $\hat{h}_e : Z \to R$ be a monotone nondecreasing function. We assume for simplicity $\lim_{\zeta \to +\infty} \hat{h}_e(\zeta) = +\infty$ and $\lim_{\zeta \to -\infty} \hat{h}_e(\zeta) = -\infty$.

Consider the problem

$$IP_3 : \quad \text{Minimize} \quad g(\max_{e \in E} \hat{h}_e(x(e)), \min_{e \in E} \hat{h}_e(x(e))) \qquad \text{subject to} \quad x \in B_Z(f). \tag{5.8}$$

Problem IP_3 is not so easy as its continuous version P_3 because of the integer constraints.

Using the same functions $\hat{h}_e$ $(e \in E)$, consider the weighted integral max-min problem IP_* and the weighted integral min-max problem IP^*, and let $\hat{v}_*$ and $\hat{v}^*$, respectively, be the optimal values of the objective functions of IP_* and IP^*. Define vectors $\hat{\ell}$, $\hat{u} \in (Z \cup \{-\infty, +\infty\})^E$ by

$$\hat{\ell}(e) = \inf\{\, \alpha \mid \alpha \in Z, \quad \hat{h}_e(\alpha) \geq \hat{v}_* \,\}, \tag{5.9}$$

$$\hat{u}(e) = \sup\{\, \alpha \mid \alpha \in Z, \quad \hat{h}_e(\alpha) \leq \hat{v}^* \,\}. \tag{5.10}$$

We have $\hat{v}_* \leq \hat{v}^*$ but, unlike the continuous version of the problem, we may not have $\hat{\ell} \leq \hat{u}$ in general. We have

$$\hat{\ell}(e) \leq \hat{u}(e) + 1 \qquad (e \in E). \tag{5.11}$$

Similarly as in the continuous version, we can show that if we have $\hat{\ell} \leq \hat{u}$, then $B_Z(f)_{\hat{\ell}}^{\hat{u}}$ is nonempty and any $x \in B_Z(f)_{\hat{\ell}}^{\hat{u}}$ is an optimal solution of IP_3.

So, let us suppose that we do not have $\hat{\ell} \leq \hat{u}$. Define

$$D = \{\, e \mid e \in E, \quad \hat{\ell}(e) > \hat{u}(e) \,\}. \tag{5.12}$$

It follows from (5.9) $\sim$ (5.12) that

$$\hat{\ell}(e) = \hat{u}(e) + 1 \qquad (e \in D), \tag{5.13}$$

$$\hat{\ell}(e) \leq \hat{u}(e) \qquad (e \in E - D), \tag{5.14}$$

$$\hat{h}_e(\hat{u}(e)) < \hat{v}_* \leq \hat{v}^* < \hat{h}_e(\hat{\ell}(e)) \qquad (e \in D), \tag{5.15}$$

$$\hat{v}_* \leq \hat{h}_e(\hat{\ell}(e)) \leq \hat{h}_e(\hat{u}(e)) \leq \hat{v}^* \qquad (e \in E - D). \tag{5.16}$$

Moreover, let the distinct values of $\hat{h}_e(\hat{u}(e))$ $(e \in D)$ be given by

$$d_1 < d_2 < \cdots < d_k \tag{5.17}$$

and define

$$A_i = \{\, e \mid e \in D, \quad \hat{h}_e(\hat{u}(e)) \leq d_i \,\} \qquad (i = 1, 2, \ldots, k). \tag{5.18}$$

Then we can show [18] that the minimum values of the objective function of Problem IP_3 is equal to the minimum of the following $k+1$ values

$$g(\hat{v}^*, d_1), \quad g(\max_{e \in A_i} \hat{h}_e(\hat{u}(e) + 1), d_{i+1}) \qquad (i = 1, 2, \ldots, k) \tag{5.19}$$

If $g(\hat{v}^*, d_1)$ is minimum, then define vectors ℓ^0, $u^0 \in R^E$ by $\ell^0 = \hat{\ell} \wedge \hat{u}$ and $u^0 = \hat{u}$, where $\hat{\ell} \wedge \hat{u} = (\min(\hat{\ell}(e), \hat{u}(e)) : e \in E)$. If $g(\max_{e \in A_{i^*}} \hat{h}_e(\hat{u}(e) + 1), d_{i^*+1})$ is minimum for some $i^* \in \{1, 2, \ldots, k\}$, then putting $w^* = \max_{e \in A_{i^*}} \hat{h}_e(\hat{u}(e) + 1)$, define vectors ℓ^0, $u^0 \in R^E$ by

$$\ell^0(e) = \inf\{\, \alpha \mid \alpha \in Z, \quad \hat{h}_e(\alpha) \geq d_{i^*+1} \,\}, \tag{5.20}$$

$$u^0(e) = \sup\{\, \alpha \mid \alpha \in Z, \quad \hat{h}_e(\alpha) \leq w^* \,\} \tag{5.21}$$

for $e \in E$. We can show that $B_Z(f)_{\ell^0}^{u^0}$ is nonempty and that any $x \in B_Z(f)_{\ell^0}^{u^0}$ is an optimal solution of IP_3.

6. Some Extensions

The optimization problems considered in the previous sections have base polyhedron as their feasible regions. For these problems we may consider the intersection of two base polyhedron as each of their feasible regions. The linear optimization problems over the intersection of two base polyhedra is equivalent to the submodular flow problem [5], the independent flow problem [11] and the polymatroidal flow problem [21], [27]. The nonlinear optimization problem over the intersection of two base polyhedra with a separable convex objective function can be formulated, for example, as a submodular flow problem with a separable convex cost function. This problem can be solved by adapting the out-of-kilter method proposed in [17]. All the problems considered in Sections 4 and 5 have been reduced to separable convex optimization problems. However, when the feasible region is given as the intersection of two base polyhedra, such reductions may not be possible.

The concept of lexicographically optimal base is generalized by M. Nakamura [32] and N. Tomizawa [35]. Suppose we are given two (polymatroid) base polyhedra $B(f_i)$ $(i = 1, 2)$ such that every base of $B(f_i)$ $(i = 1, 2)$ consists of positive components. If b_1 is a lexicographically optimal base of $B(f_1)$ with respect to a weight vector $b_2 \in B(f_2)$ and b_2 is a lexicographically optimal base of $B(f_2)$ with respect to b_1, the pair (b_1, b_2) is a *universal pair of bases* (the orignal definition in [32], [35] is different from but equivalent to the present one). Some characterizations of universal pairs are given by K. Murota [31].

References

[1] G. Birkhoff: Rings of sets. *Duke Mathematical Journal* 3 (1937) 443–454.

[2] R. E. Bixby, W. H. Cunningham and D. M. Topkis: The partial order of a polymtroid extreme point. *Mathematics of Operations Research* 10 (1985) 367–378.

[3] R. Brown: The sharing problem. *Operations Research* 27 (1979) 324–340.

[4] J. Edmonds: Submodular functions, matroids, and certain polyhedra. *Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications* (eds., R. Guy, H. Hanani, N. Sauer and J. Schönheim and Gordon and Breach, New York, 1970), pp. 69–87.

[5] J. Edmonds and R. Giles: A min-max relation for submodular functions on graphs. *Annals of Discrete Mathematics* **1** (1977) 185–204.

[6] U. Faigle: Matroids in combinatorial optimization.*Combinatorial Geometries* (ed., N. White, Encyclopedia of Mathematics and Its Application **29**, Combridge University Press, 1987) pp. 161–210.

[7] A. Federgruen and H. Groenevelt: The greedy procedure for resource allocation problems – Necessary and sufficient conditions for optimality. *Operations Research* **34** (1986) 909–918.

[8] L. R. Ford, Jr. and D. R. Fulkerson: *Flows in Networks*. Princeton University Press, Prinseton, 1962.

[9] A. Frank: An algorithm for submodular functions on graph. *Annals of Discrete Mathematics* **16** (1982) 189–212.

[10] A. Frank and É. Tardos: Generalized polymatroids and submodular folws. Report No.85389, Institut für Ökonometrie und Operations Research, Universität Bonn, West Germany, August 1986.

[11] S. Fujishige: Algorithms for solving the independent-flow problems. *Journal of the Operations Research Society of Japan* **21** (1978) 189–204.

[12] S. Fujishige: The independent-flow problems and submodular functions (in Japanese). *Journal of the Faculty of Engineering, University of Tokyo* **A-16** (1978) 42–43.

[13] S. Fujishige: Polymatroidal dependence structure of a set of random variables. *Information and Control* **39** (1987) 55–72.

[14] S. Fujishige: Lexicographically optimal base of a polymatroid with respect to a weight vector. *Mathematics of Operations Research* **5** (1980) 186–196.

[15] S. Fujishige: Structures of polyhedra determined by submodular functions on crossing families. *Mathematical Programming* **29** (1984) 125–141.

[16] S. Fujishige: Submodular systems and related topics. *Mathematical Programming Study* **22** (1984) 113–131.

[17] S. Fujishige: An out-of-kilter method for submodular flows. *Discrete Applied Mathematics* **17** (1987) 3–16.

[18] S. Fujishige, N. Katoh and T. Ichimori: The fair resource allocation problem with submodular constraints. *Mathematics of Operations Research* **13** (1988) 164–173.

[19] S. Fujishige and N. Tomizawa: A note on submodular functions on distributive lattices. *Journal of the Operations Research Society of Japan* **26** (1983) 309–318.

[20] H. Groenevelt: Two algorithms for maximizing a separable concave function over a polymatroid feasible region. Working Paper, The Graduate School of Management, The University of Rochester, August 1985.

[21] R, Hassin: Minimum cost flow with set-constraints. *Networks* **12** (1982) 1–21.

[22] T. Ibaraki and N. Katoh: *Resource Allocation Problems – Algorithmic Approaches*. Foundation of Computing Series, MIT Press, 1988.

[23] T. Ichimori, H. Ishii and T. Nishida: Optimal sharing. *Mathematical Programming* **23** (1982) 341–348.

[24] M. Iri: Applications of matroid theory. *Mathematical Programming – The state of the Art* (eds., A. Bachem, M. Grötschel and B. Korte, Springer, Berlin, 1983), pp. 158–201.

[25] M. Iri and S. Fujishige: Use of matroid theory in operations research, circuits and system theory. *International Journal of Systems Science* **12** (1981) 27–54.

[26] N. Katoh: Private communication, 1985.

[27] E. L. Lawler and C. U. Martel: Computing maximal polymatroidal network flows. *Mathematics of Operations Research* **7** (1982) 334–347.

[28] L. Lovász: Submodular functions and convexity. *Mathematical Programming – The state of the Art* (eds., A. Bachem, M. Grötschel and B. Korte, Springer, Berlin, 1983), pp. 235–257.

[29] N. Megiddo: Optimal flows in networks with multiple sources and sinks. *Mathematical Programming* **7** (1974) 97–107.

[30] K. Murota: *Systems Analysis by Graphs and Matroids – Structural Solvability and Controllability*. Algorithms and Combinatorics **3**, Springer, Berlin, 1987.

[31] K. Murota: Note on the universal bases of a pair of polymatroids. Research Mamorandum, RMI88-05, Department of Mathematical Engineering and Information Physics, University of Tokyo, February 1988.

[32] M. Nakamura: Mathematical Anlysis of Discrete Systems and Its Applications (in Japanese). Dissertation, Department of Mathematical Engineering and Instrumentation Physics, Faculty of Engineering, University of Tokyo, 1983.

[33] L. S. Shapley: Cores of convex games. *International Journal of Game Theory* **1** (1971) 11–26.

[34] N. Tomizawa: Theory of hyperspaces (I) – supermodular functions and generalization of concept of 'base' (in Japaneae). Papers of the Technical Group on Circuits and Systems, Institute of Electronics and Communication Engineers of Japan, CAS 80-72 (1980).

[35] N. Tomizawa: Theory of hyperspaces (IV) – principal partitions of hypermatroids (in Japanese). Papers of the Technical Group on Circuits and Systems, Institute of Electronics and Communication Engineers of Japan, CAS 80-85 (1980).

[36] N. Tomizawa: Theory of hyperspaces (XVI) – on the structures of hedrons (in Japanese). Papers of the Technical Group on Circuits and Systems, Institute of Electronics and Communication Engineers of Japan, CAS 82-174 (1983).

[37] D. M. Topkis: Adjacency on polymatroids. *Mathematical Programming* **30** (1984) 229–237.

[38] U. Zimmermann: Linear and combinatorial sharing problems. *Discrete Applied Mathematics* **15** (1986) 85–105.