THEORY OF SUBMODULAR PROGRAMS:
A FENCHEL-TYPE MIN-MAX THEOREM AND
SUBGRADIENTS OF SUBMODULAR FUNCTIONS

Satoru FUJISHIGE

Institute of Socio-Economic Planning, University of Tsukuba, Sakura, Ibaraki 305, Japan

Received 26 January 1982
Revised manuscript received 12 July 1983

We consider submodular programs which are problems of minimizing submodular functions on distributive lattices with or without constraints. We define a convex (or concave) conjugate function of a submodular (or supermodular) function and show a Fenchel-type min-max theorem for submodular and supermodular functions. We also define a subgradient of a submodular function and derive a necessary and sufficient condition for a feasible solution of a submodular program to be optimal, which is a counterpart of the Karush-Kuhn-Tucker condition for convex programs.

Keywords: Submodular functions, Supermodular functions, Subgradients, Fenchel's Duality Theorem, Karush-Kuhn-Tucker Condition.

1. Introduction

Let $\mathcal{D}$ be a distributive lattice formed by subsets of a finite set $E$ with set union and intersection as the lattice operations and let $f$ be a function from $\mathcal{D}$ to the set $\mathbb{R}$ of reals such that

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \quad (1.1)$$

for any $X, Y \in \mathcal{D}$. Then $f$ is called a submodular function on $\mathcal{D}$. If $-g$ is a submodular function, then $g$ is called a supermodular function. If $f$ is a submodular and, at the same time, supermodular function, we call $f$ a modular function.

We shall develop a theory of submodular programs from the point of view of the duality in mathematical programs. Submodular (or supermodular) functions on distributive lattices share similar structures with convex (or concave) functions on convex sets. We shall define a convex (or concave) conjugate function of a submodular (or supermodular) function and show a Fenchel-type min-max theorem for submodular and supermodular functions. We also define a subgradient of a submodular function and consider the problem of minimizing a submodular function with or without constraints. We derive a necessary and sufficient condition for a feasible solution to be optimal, which is a counterpart of the Karush-Kuhn-Tucker condition for ordinary convex programs.

This work is supported by the Alexander von Humboldt fellowship (1982/83), West Germany.
2. Preliminaries

In this section we give propositions which are well known, or are immediate consequences of those well known, in polymatroid theory (see [1, 3-5, 9, 15]).

Let $E$ be a finite set, $2^E$ be the set of all the subsets of $E$ and $\mathbb{R}$ be the set of reals. Throughout the present paper we assume, for the sake of simplicity, that for every distributive lattice $\mathcal{D} \subseteq 2^E$ and for every function $f : \mathcal{D} \to \mathbb{R}$ we have $\emptyset \in \mathcal{D}$ and $f(\emptyset) = 0$ unless otherwise stated.

For a distributive lattice $\mathcal{D} \subseteq 2^E$ let $f : \mathcal{D} \to \mathbb{R}$ be a submodular function. Then we call the pair $(\mathcal{D}, f)$ a submodular system. Similarly, we define a supermodular system $(\mathcal{D}, g)$ for a supermodular function $g : \mathcal{D} \to \mathbb{R}$.

Denote by $\mathbb{R}^E$ the set of all $|E|$-vectors $x = (x(e) : e \in E)$ with $x(e) \in \mathbb{R}$ ($e \in E$) and define for any $X \subseteq E$

$$x(X) = \sum_{e \in X} x(e), \quad (2.1)$$

where by convention $x(\emptyset) = 0$. Each vector $x \in \mathbb{R}^E$ will also be regarded as a modular function $x : 2^E \to \mathbb{R}$ by (2.1).

For distributive lattices $\mathcal{D}_1, \mathcal{D}_2 \subseteq 2^E$, a submodular function $f : \mathcal{D}_1 \to \mathbb{R}$ and a supermodular function $g : \mathcal{D}_2 \to \mathbb{R}$, let us define polyhedra

$$P(f) = \{x \in \mathbb{R}^E, \forall X \in \mathcal{D}_1 : x(X) \leq f(X)\}, \quad (2.2)$$

$$P(g) = \{x \in \mathbb{R}^E, \forall X \in \mathcal{D}_2 : x(X) \geq g(X)\} \quad (2.3)$$

and, if $E \in \mathcal{D}_1$ and $E \in \mathcal{D}_2$,

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

$$B(g) = \{x \in P(g), x(E) = g(E)\}. \quad (2.5)$$

We call $P(f)$ the submodular polyhedron associated with the submodular system $(\mathcal{D}_1, f)$ and $B(f)$ the base polyhedron associated with $(\mathcal{D}_1, f)$. Similarly, we call $P(g)$ the supermodular polyhedron and $B(g)$ the base polyhedron associated with $(\mathcal{D}_2, g)$. A vector $x \in B(f)$ (or $B(g)$) is called a base of $(\mathcal{D}_1, f)$ (or $(\mathcal{D}_2, g)$).

**Proposition 2.1.** For any submodular system $(\mathcal{D}, f)$ with $E \in \mathcal{D}$, if $y \in P(f)$, then there exists a base $x \in B(f)$ such that $y \leq x$, i.e., $\forall e \in E : y(e) \leq x(e)$.

**Proposition 2.2.** For any submodular system $(\mathcal{D}, f)$ and $A \in \mathcal{D}$ there exists a vector $x \in P(f)$ such that $x(A) = f(A)$. Moreover, if $A \subseteq E$ and $A \not\in \mathcal{D}$, then for any $r \in \mathbb{R}$ there exists $x \in P(f)$ such that $x(A) > r$.

**Proposition 2.3.** Let $(\mathcal{D}_1, f_1)$ and $(\mathcal{D}_2, f_2)$ be submodular systems. Then,

$$P(f_1 + f_2) = P(f_1) + P(f_2), \quad (2.6)$$

where the domain of $f_1 + f_2$ is $\mathcal{D}_1 \cap \mathcal{D}_2$. 