JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 127, 193–205 (1987)

Optimization Theory for n-Set Functions

H. W. CORLEY

Department of Industrial Engineering, The University of Texas, Arlington, Texas 76019

Submitted by V. Lakshmikantham

Received February 15, 1986

An optimization theory is developed for functions of n sets. Optimality conditions are established, and a Lagrangian duality is obtained. © 1987 Academic Press, Inc.

1. INTRODUCTION

The concept of optimizing set functions (i.e., functions of sets) arises in various mathematical areas. For example, an early result was the Neyman–Pearson lemma of statistics [6, 11], which is simply the statement of a sufficient condition for maximizing an integral over a single set. The necessity of this condition, as well as the existence of a solution, was later established in [5]. These results were subsequently generalized to n sets and a duality theory was developed in [3, 4]. However, all these results were for special cases of set functions involving integrals. The first general theory for optimizing set functions was developed by Morris [10], who obtained for functions of a single set the analogs to standard mathematical programming results. Subsequent work [2, 7, 8, 14] on duality and multiple objective optimization has remained confined to functions of a single set.

In this paper previous work is generalized by minimizing n-set functions, i.e., functions of n sets. In Section 2 some preliminary matters are considered and the problem is formally stated. In Section 3 differential necessary conditions for local minima are developed. These conditions are shown to be sufficient under convexity assumptions in Section 4. Saddlepoint optimality conditions and a Lagrangian duality are obtained in Section 5.

0022-247X/87 $3.00

2. Preliminaries

Throughout this paper let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space. We will be concerned with functions on $\mathscr{A}^n = \{(S_1, ..., S_n)\colon S_i \in \mathscr{A}, i = 1, ..., n\}$. The fact that $\mathscr{A}^n$ is only a semialgebra, not a σ-algebra, does not limit the analysis. $\mathscr{A}^n$ is a pseudometric space under the pseudometric d defined by

$$d[(R_1, ..., R_n), (S_1, ..., S_n)] = \left[\sum_{i=1}^{n} \mu^2(R_i \Delta S_i)\right]^{1/2}, \qquad R_i, S_i \in \mathscr{A},\ i = 1, ..., n,$$

where Δ denotes symmetric difference. Essentially $(R_1, ..., R_n)$ and $(S_1, ..., S_n)$ will be regarded as equivalent if R_i and S_i, $i = 1, ..., n$, differ only by zero measure. This practice presents no difficulty in an optimization theory over the class of pseudocontinuous n-set functions $F\colon \mathscr{A}^n \to R^1$, as noted below. This paper is therefore restricted to such functions. The proof of Result 2.2 is an immediate consequence of Definition 2.1.

Definition 2.1. The n-set function F is said to be d-pseudocontinuous at $(R_1, ..., R_n)$ on the pseudometric space $(\mathscr{A}^n, d)$ if given $\varepsilon > 0$ there exists $\delta > 0$ for which $|F(R_1, ..., R_n) - F(S_1, ..., S_n)| < \varepsilon$ whenever $d[(R_1, ..., R_n), (S_1, ..., S_n)] < \delta$.

Result 2.2. If $F\colon \mathscr{A}^n \to R^1$ is d-pseudocontinuous on the pseudometric space $(\mathscr{A}^n, d)$ and $d[(R_1, ..., R_n), (S_1, ..., S_n)] = 0$, then $F(R_1, ..., R_n) = F(S_1, ..., S_n)$.

The counterparts of the usual definitions of local and global minima are next stated for $(\mathscr{A}^n, d)$.

Definition 2.3. Let $F\colon \mathscr{A}^n \to R^1$ and $\mathscr{B} \subset \mathscr{A}^n$. Then $(S_1^*, ..., S_n^*) \in \mathscr{B}$ is a (global) minimum of F on $\mathscr{B}$ if $F(S_1^*, ..., S_n^*) \leqslant F(S_1, ..., S_n)$ for all $(S_1, ..., S_n) \in \mathscr{B}$. $(S_1^*, ..., S_n^*)$ is a local minimum of F on $\mathscr{B}$ if there exists $\delta > 0$ such that $F(S_1^*, ..., S_n^*) \leqslant F(S_1, ..., S_n)$ for all $(S_1, ..., S_n) \in \mathscr{B}$ satisfying $d[(S_1, ..., S_n), (S_1^*, ..., S_n^*)] < \delta$.

For $F, G_1, ..., G_m\colon \mathscr{A}^n \to R^1$ the problem to be analyzed here is to find minima of F on $\mathscr{B} = \{(S_1, ..., S_n) : G_j(S_1, ..., S_n) \leqslant 0, j = 1, ..., m\}$, i.e., to

$$\begin{aligned} &\text{minimize } F(S_1, ..., S_n) \qquad \text{subject to} \\ &G_j(S_1, ..., S_n) \leqslant 0, \qquad j = 1, ..., m. \end{aligned} \tag{1}$$

3. Local Differential Theory

For $h \in L_1(X, \mathscr{A}, \mu)$ and $S \in \mathscr{A}$ with characteristic function $\chi_S \in L_\infty(X, \mathscr{A}, \mu)$, the integral $\int_S h \, d\mu$ will be denoted by $\langle h, \chi_S \rangle$. Definition 3.1 is due to Morris [10].

DEFINITION 3.1. A set function $H: \mathscr{A} \to R^1$ is differentiable at $S^* \in \mathscr{A}$ if there exists $h_{S^*} \in L_1(X, \mathscr{A}, \mu)$, the derivative of H at S^*, such that

$$H(S) = H(S^*) + \langle h_{S^*}, \chi_S - \chi_{S^*} \rangle + E_H(S^*, S),$$

where $E_H(S^*, S)$ is $o[d(S^*, S)]$, i.e., $\lim_{d(S^*,S) \to 0} E_H(S^*, S)/d(S^*, S) = 0$.

Differentiation for n-set functions is next defined.

DEFINITION 3.2. Let $F: \mathscr{A}^n \to R^1$ and $(S_1^*, ..., S_n^*) \in \mathscr{A}^n$. Then F is said to have a partial derivative at $(S_1^*, ..., S_n^*)$ with respect to S_i if the set function $H(S_i) = F(S_1^*, ..., S_{i-1}^*, S_i, S_{i+1}^*, ..., S_n^*)$ has derivative $h_{S_i^*}$ at S_i^*. In that case we define the ith partial derivative of F at $(S_1^*, ..., S_n^*)$ to be $f^i_{S_1^*, ..., S_n^*} = h_{S_i^*}$.

DEFINITION 3.3. Let $F: \mathscr{A}^n \to R^1$ and $(S_1^*, ..., S_n^*) \in \mathscr{A}^n$. Then F is said to be differentiable at $(S_1^*, ..., S_n^*)$ if all the partials $f^i_{S_1^*, ..., S_n^*}$, $i = 1, ..., n$, exist and satisfy

$$F(S_1, ..., S_n) = F(S_1^*, ..., S_n^*) + \sum_{i=1}^{n} \langle f^i_{S_1^*, ..., S_n^*}, \chi_{S_i} - \chi_{S_i^*} \rangle + W_F[(S_1^*, ..., S_n^*), (S_1, ..., S_n)], \tag{2}$$

where $W_F[(S_1^*, ..., S_n^*), (S_1, ..., S_n)]$ is $o\{d[(S_1^*, ..., S_n^*), (S_1, ..., S_n)]\}$.

As in Proposition 2.2 of [10], if $F: \mathscr{A}^n \to R^1$ is differentiable, its partial derivatives are unique. An example of a differentiable n-set function is as follows.

EXAMPLE 3.4. Define $F(S_1, ..., S_n) = u(\langle v_1, \chi_{S_1} \rangle, ..., \langle v_n, \chi_{S_n} \rangle)$, where $u: R^n \to R^1$ is differentiable and $v_1, ..., v_n \in L_\infty(X, \mathscr{A}, \mu)$. Then F is differentiable and

$$f^i_{S_1^*, ..., S_n^*} = u^{(i)}(\langle v_i, \chi_{S_1^*} \rangle, ..., \langle v_n, \chi_{S_n^*} \rangle)\, v_i, \qquad i = 1, ..., n,$$

where $u^{(i)}$ denotes the ith partial derivative of u.

Differential necessary conditions for a local (and hence global) minimum to (1) are next established. Result 3.5, which follows readily from elementary properties of integration and (2), justifies such a development in view of Result 2.2 and the previous restriction to pseudocontinuous functions. The result also illustrates similarities to differentiability of real-valued function of n-variables. In that setting (see [12]) differentiability implies continuity, and the continuity of the partial derivatives implies that the gradient exists and is continuous.

RESULT 3.5. If $F: \mathscr{A}^n \to R^1$ is differentiable at $(S_1^*, ..., S_n^*)$, then F is pseudocontinuous a $(S_1^*, ..., S_n^*)$ and $\sum_{i=1}^n \langle f^i_{S_1^*,...,S_n^*}, \chi_{S_i} - \chi_{S_i^*}\rangle$ is a pseudocontinuous function of $(S_1, ..., S_n)$.

Theorem 3.7 below is the analog of the Fritz John conditions of mathematical programming [1]. A well-known lemma of Liapunov [13] is needed.

LEMMA 3.6. *Let $h_i: X \to R^1$, $i = 1, ..., p$, be integrable functions on the atomless measure space $(X, \mathscr{A}, \mu)$ and $S \in \mathscr{A}$. Then the range of the vector measure $(\langle h_1, \chi_S\rangle, ..., \langle h_p, \chi_S\rangle)$ is convex and compact.*

THEOREM 3.7. *Let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space and let $F, G_1, ..., G_m: \mathscr{A}^n \to R^1$ be differentiable at $(S_1^*, ..., S_n^*)$. If $(S_1^*, ..., S_n^*)$ is a local minimum for (1), then there exist scalars $\lambda_0^*, \lambda_1^*, ..., \lambda_m^*$ such that*

$$\left\langle \lambda_0 f^i_{S_1^*,...,S_m^*} + \sum_{j=1}^m \lambda_j^* g^{ij}_{S_1^*,...,S_n^*}, \chi_{S_i} - \chi_{S_i^*}\right\rangle \geqslant 0 \qquad \textit{for all} \quad S_i \in \mathscr{A}, i = 1, ..., n \tag{3}$$

$$\lambda_j^* G_j(S_1^*, ..., S_n^*) = 0, \qquad j = 1, ..., m \tag{4}$$

$$\lambda_0^*, \lambda_1^*, ..., \lambda_m^* \geqslant 0 \tag{5}$$

$$G_j(S_1^*, ..., S_n^*) \leqslant 0, \qquad j = 1, ..., m \tag{6}$$

$$\lambda_j^* \textit{ not all zero,} \tag{7}$$

where $g^{ij}_{S_1^*,...,S_n^*}$ is the ith partial of G_j at $(S_1^*, ..., S_n^*)$.

Proof. In the proof we write f^i_* for $f^i_{S_1^*,...,S_n^*}$, g^{ij}_* for $g^{ij}_{S_1^*,...,S_n^*}$, and f^i, g^{ij} for arbitrary $(S_1, ..., S_n)$. Define

$$A = \left\{ \begin{array}{l} (v_0, v_1, ..., v_m)\text{: there exists } (S_1, ..., S_n) \in \mathscr{A}^n \text{ such that} \\ v_0 \geqslant \sum_{i=1}^n \langle f^i_*, \chi_{S_i} - \chi_{S_i^*}\rangle, \\ v_j \geqslant G_j(S_1^*, ..., S_m^*) + \sum_{i=1}^n \langle g^{ij}_*, \chi_{S_i} - \chi_{S_i^*}\rangle, \qquad j = 1, ..., m \end{array} \right\}$$

and

$$B=\{(v_0, v_1, ..., v_m): v_j<0, j=0, 1, ..., m\}.$$

The set B is clearly convex. To prove that A is convex it suffices to prove that the translate of A,

$$A_1=\left\{\begin{array}{l}(v_0, v_1, ..., v_m)\text{: there exists } (S_1, ..., S_n)\in \mathscr{A}^n \text{ such that}\\ v_0\geqslant \sum_{i=1}^{n} \langle f^i_*, \chi_{S_i}\rangle,\\ v_j\geqslant \sum_{i=1}^{n} \langle g^{ij}_*, \chi_{S_i}\rangle, j=1, ..., m\end{array}\right\}$$

is convex. Consider

$$A_2=\left\{\begin{array}{l}(v_0, v_1, ..., v_m)\text{: there exists } (S_1, ..., S_n)\in \mathscr{A}^n \text{ such that}\\ v_0=\sum_{i=1}^{n} \langle f^i_*, \chi_{S_i}\rangle, v_j=\sum_{i=1}^{n} \langle g^{ij}_*, \chi_{S_i}\rangle,\\ \qquad j=1, ..., m\end{array}\right\}$$

and

$$C_i=\left\{\begin{array}{l}(u_0, u_1, ..., u_m)\text{: there exists } S_i\in \mathscr{A} \text{ such that}\\ u_0=\langle f^i_*, \chi_{S_i}\rangle, u_j=\langle g^{ij}_*, \chi_{S_i}\rangle, j=1, ..., m\end{array}\right\},$$

$i=1, ..., n$. Each C_i is convex from Lemma 3.6, so $A_2=C_1+\cdots+C_n=\{\sum_{i=1}^n c_i : c_i\in C_i,\ i=1, ..., n\}$ is convex. Similarly $A_1=A_2+R^+_{m+1}$ is convex, and the convexity of A follows.

A and B are next shown to be disjoint. Assume the contrary, i.e., that there exists $(S_1, ..., S_n)\in \mathscr{A}^n$ for which

$$\sum_{i=1}^{n} \langle f^i_*, \chi_{S_i}-\chi_{S_i^*}\rangle<0 \quad\text{and}\quad G_j(S_1^*, ..., S_m^*)+\sum_{i=1}^{n} \langle g^{ij}_*, \chi_{S_i}-\chi_{S_i^*}\rangle$$
$$<0, \quad j=1, ..., m.$$

For $i=1, ..., n$, let $S_i^+=S_i\backslash S_i^*$ and $S_i^-=S_i^*\backslash S_i$ so that $\chi_{S_i}-\chi_{S_i^*}=\chi_{S_i^+}-\chi_{S_i^-}$. Then from Lemma 3.6 there exist families $S_i^+(\alpha)\subset S_i^+$ and $S_i^-(\alpha)\subset S_i^-$ satisfying

$$\int_{S_i^\pm(\alpha)} (1, f, g_1, ..., g_m)\, d\mu=\alpha \int_{S_i^\pm} (1, f, g_1, ..., g_m)\, d\mu$$

for $\alpha \in [0, 1]$. Also let $S_i(\alpha) = [S_i^+(\alpha) \cup S_i^*] \backslash S_i^-(\alpha)$. Since $d[S_i^*, S_i(\alpha)] = \alpha d[S_i^*, S_i]$, it can be deduced from (2) that

$$F[S_1(\alpha), ..., S_n(\alpha)] - F[S_1^*, ..., S_n^*] = \alpha \sum_{i=1}^{n} \langle f_*^i, \chi_{S_i} - \chi_{S_i^*} \rangle + o(\alpha). \tag{8}$$

Similarly,

$$G_j[S_1(\alpha), ..., S_n(\alpha)] - G_j[S_1^*, ..., S_n^*] = \alpha \sum_{i=1}^{n} \langle g_*^{ij}, \chi_{S_i} - \chi_{S_i^*} \rangle + o(\alpha), \qquad j = 1, ..., m. \tag{9}$$

From (8) and (9) there exists $\delta \in (0, 1)$ such that for $\alpha \in (0, \delta)$ both $F[S_1(\alpha), ..., S_n(\alpha)] - F[S_1^*, ..., S_n^*] < 0$ and $G_j[S_1(\alpha), ..., S_n(\alpha)] - G_j[S_1^*, ..., S_n^*] < 0$, $j = 1, ..., m$, contradicting that $(S_1^*, ..., S_n^*)$ is local minimum.

A and B are thus disjoint convex sets and can be separated by a hyperplane. Hence there exist scalars $\lambda_0^*, \lambda_1^*, ..., \lambda_m^*$, not all zero, and ξ for which $\sum_{j=0}^{m} \lambda_j^* v_j \geqslant \xi$ if $(v_0, v_1, ..., v_m) \in A$ and $\sum_{j=0}^{m} \lambda_j^* v_j \leqslant \xi$ if $(v_0, v_1, ..., v_m) \in B$. As usual, it can be shown that $\xi = 0$. Hence

$$\lambda_0^* \sum_{i=1}^{n} \langle f_*^i, \chi_{S_i} - \chi_{S_i^*} \rangle + \sum_{j=1}^{m} \lambda_j^* G_j(S_1^*, ..., S_n^*) + \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j^* \langle g_*^{ij}, \chi_{S_i} - \chi_{S_i^*} \rangle \geqslant 0 \qquad \text{for all } (S_1, ..., S_n) \in \mathscr{A}^n. \tag{10}$$

Setting $S_i = S_i^*$ in (10) yields (4), and (10) becomes

$$\lambda_0^* \sum_{i=1}^{n} \langle f_*^i, \chi_{S_i} - \chi_{S_i^*} \rangle + \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j^* \langle g_*^{ij}, \chi_{S_i} - \chi_{S_i^*} \rangle \geqslant 0 \qquad \text{for all } (S_1, ..., S_n) \in \mathscr{A}^n. \tag{11}$$

Letting $S_k = S_k^*$, $k \neq i$, in (11) establishes (3). Since (5)–(7) are immediate, the proof is complete. ∎

Kuhn–Tucker conditions for (1) can be stated under the additional assumption that $(S_1^*, ..., S_n^*)$ is regular as defined in Definition 3.8. Corollary 3.9 then follows from Theorem 3.7 by standard arguments as in [9].

DEFINITION 3.8. $(S_1^*, ..., S_n^*)$ is said to be regular if there exists $(\hat{S}_1, ..., \hat{S}_n)$ for which

$$G_j(S_1^*, ..., S_n^*) + \sum_{i=1}^{n} \langle g_*^{ij}, \chi_{\hat{S}_i} - \chi_{S_i^*} \rangle < 0, \qquad j = 1, ..., m.$$

COROLLARY 3.9. *In addition to the hypotheses of the previous theorem, suppose that* $(S_1^*, ..., S_n^*)$ *is regular. Then there exist* $\lambda_1^*, ..., \lambda_m^*$ *which, together with* $\lambda_0^* = 1$, *satisfy* (3)–(6).

4. CONVEXITY AND GLOBAL DIFFERENTIAL THEORY

The convexity of n-set functions is now defined, some properties related to convexity are established, and the conditions of Corollary 3.9 are shown to be sufficient for a global minimum if $F, G_1, ..., G_m$ are convex.

DEFINITION 4.1. Let $F: \mathscr{A}^n \to R^1$. F is said to be convex if for each $\lambda \in [0, 1]$ and $(R_1, ..., R_n), (S_1, ..., S_n) \in \mathscr{A}^n$

$$\overline{\lim_{k \to \infty}} F[R_1^k \cup S_1^k \cup (R_1 \cap S_1), ..., R_n^k \cup S_n^k \cup (R_n \cap S_n)]$$
$$\leqslant \lambda F(R_1, ..., R_n) + (1 - \lambda) F(S_1, ..., S_n)$$

for any sequences of sets $R_i^k \subset R_i \backslash S_i$ and $S_i^k \subset S_i \backslash R_i$, $k = 1, 2, ...$, satisfying $\chi_{R_i^k} \to^{w^*} \lambda \chi_{R_i \backslash S_i}$ and $\chi_{S_i^k} \to^{w^*} (1 - \lambda) \chi_{S_i \backslash R_i}$, $i = 1, ..., n$.

EXAMPLE 4.2. The n-set function $F(S_1, ..., S_n) = u(\langle v_1, \chi_{S_1} \rangle, ..., \langle v_n, \chi_{S_n} \rangle)$, where $u: R^n \to R^1$ is convex and $v_1, ..., v_n \in L_1(X, \mathscr{A}, \mu)$, is convex.

The following two lemmas are proved in [10].

LEMMA 4.3. *Let* $(X, \mathscr{A}, u)$ *be a finite atomless measure space with* $L_1(X, \mathscr{A}, \mu)$ *separable. Then* $\lambda \chi_S$ *is in the weak* closure of* $\chi = \{\chi_R : R \in \mathscr{A}\} \subset L_\infty(X, \mathscr{A}, \mu)$ *for all* $S \in \mathscr{A}$ *and* $\lambda \in [0, 1]$.

LEMMA 4.4. *Let* $R, S \in \mathscr{A}$ *and* $\lambda \in [0, 1]$. *If* R_k, S_k, $k = 1, 2, ...,$ *are sequences of sets in* $\mathscr{A}$ *such that the* $L_\infty(X, \mathscr{A}, \mu)$ *sequences*

$$\chi_{R_k} \xrightarrow{w^*} \lambda \chi_{R \backslash S}, \chi_{S_k} \xrightarrow{w^*} (1 - \lambda) \chi_{S \backslash R},$$

then

$$\chi_{R_k \cup S_k \cup (R \cap S)} \xrightarrow{w^*} \lambda \chi_R + (1 - \lambda) \chi_S.$$

THEOREM 4.5. *Let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space and let $F: \mathscr{A}^n \to R^1$ be differentiable on $\mathscr{A}^n$. If $L_1(X, \mathscr{A}, \mu)$ is separable and F is convex, then for all $(R_1, ..., R_n)$, $(S_1, ..., S_n) \in \mathscr{A}^n$*

$$\sum_{i=1}^{n} \langle f^i_{R_1, ..., R_n}, \chi_{S_i} - \chi_{R_i} \rangle \leqslant F(S_1, ..., S_n) - F(R_1, ..., R_n). \tag{12}$$

Proof. Fix $0 < \lambda \leqslant 1$. It follows from Lemma 4.3 that there exist sequences $R_i^k(\lambda) \subset R_i \setminus S_i$ and $S_i^k(\lambda) \subset S_i \setminus R_i$ for which

$$\chi_{R_i^k(\lambda)} \xrightarrow{w^*} (1 - \lambda) \chi_{R_i \setminus S_i}, \qquad \chi_{S_i^k(\lambda)} \xrightarrow{w^*} \lambda \chi_{S_i \setminus R_i}.$$

Set $T_i^k(\lambda) = R_i^k(\lambda) \cup S_i^k(\lambda) \cup (R_i \cap S_i)$, $i = 1, ..., n$. Then from Definition 4.1,

$$\overline{\lim_{k \to \infty}}\; F[T_1^k(\lambda), ..., T_n^K(\lambda)] \leqslant (1 - \lambda) F(R_1, ..., R_n) + \lambda F(S_1, ..., S_n),$$

so

$$\overline{\lim_{k \to \infty}}\; \{F[T_1^k(\lambda), ..., T_n^k(\lambda)] - F(R_1, ..., R_n\}/\lambda$$
$$\leqslant F(S_1, ..., S_n) - F(R_1, ..., R_n). \tag{13}$$

Applying (2) to (12) and invoking Lemma 4.4 give

$$\left\langle \sum_{i=1}^{n} f^i_{R_1, ..., R_n}, \chi_{S_i} - \chi_{R_i} \right\rangle + \overline{\lim_{k \to \infty}}\; W_F[(T_1^k(\lambda), ..., T_n^k(\lambda)), (R_1, ..., R_n)]/\lambda$$
$$\leqslant F(S_1, ..., S_n) - F(R_1, ..., R_n). \tag{14}$$

But using the fact that $R_i^k(\lambda) \subset R_i \setminus S_i$ and $S_i^k(\lambda) \subset R_i \setminus S_i$, it can be shown that

$$\lim_{\lambda \to 0} \overline{\lim_{k \to \infty}}\; W_F[(T_1^k(\lambda), ..., T_n^k(\lambda)), (R_1, ..., R_n)]/\lambda = 0.$$

Letting $\lambda \to 0$ in (14) now yields (12). ∎

THEOREM 4.6. *Let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space and let $F: \mathscr{A}^n \to R^1$ be differentiable on $\mathscr{A}^n$. If (12) is satisfied for all $(R_1, ..., R_n)$, $(S_1, ..., S_n) \in \mathscr{A}^n$, then F is convex.*

Proof. For $\lambda \in [0, 1]$ and $(R_1, ..., R_n)$, $(S_1, ..., S_n) \in \mathscr{A}^n$, let $R_i^k \subset R_i \setminus S_i$ and $S_i^k \subset S_i \setminus R_i$ be such that $\chi_{R_i^k} \to^{w^*} \lambda \chi_{R_i/S_i}$ and $\chi_{S_i^k} \to^{w^*} (1 - \lambda) \chi_{S_i/R_i}$. Setting $T_i^k = R_i^k \cup S_i^k \cup (R_i \cap S_i)$ for each k we have by hypothesis that

$$\sum_{i=1}^{n} \langle f^i_{T_1^k, ..., T_n^k}, \chi_{S_i} - \chi_{T_i^k} \rangle \leqslant F(S_1, ..., S_n) - F(T_1^k, ..., T_n^k). \tag{15}$$

Letting $k \to \infty$ in (15) and applying Lemma 4.4 give

$$\varliminf_{k \to \infty} \sum_{i=1}^{n} \langle f_{T_1^k, \ldots, T_n^k}, \lambda(\chi_{S_i} - \chi_{R_i}) \rangle \leqslant F(S_1, \ldots, S_n) - \varlimsup_{k \to \infty} F(T_1^k, \ldots, T_n^k). \tag{16}$$

Replacing $(S_1, \ldots, S_n)$ by $(R_1, \ldots, R_n)$ in (15) similarly yields

$$\varliminf_{k \to \infty} \sum_{i=1}^{n} \langle f_{T_1^k, \ldots, T_n^k}, (1-\lambda)(\chi_{R_i} - \chi_{S_i}) \rangle \leqslant F(R_1, \ldots, R_n) - \varlimsup_{k \to \infty} F(T_1^k, \ldots, T_n^k). \tag{17}$$

Upon multiplying (16) by $(1-\lambda)$, (17) by λ, and then adding, it follows that $\varlimsup F(T_1^k, \ldots, T_n^k) \leqslant \lambda F(R_1, \ldots, R_n) + (1-\lambda) F(S_1, \ldots, S_n)$ to establish the convexity of F. ∎

It should be noted that the separability of $L_1(X, \mathscr{A}, \mu)$ is not required in Theorem 4.6.

THEOREM 4.7. *Let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space with $L_1(X, \mathscr{A}, \mu)$ separable. Suppose that $F, G_1, \ldots, G_m \colon \mathscr{A}^n \to R^1$ are differentiable at $(S_1^*, \ldots, S_n^*)$, as well as convex. If there exist scalars $\lambda_1^*, \ldots, \lambda_m^*$ such that $(S_1^*, \ldots, S_n^*)$ and $\lambda_1^*, \ldots, \lambda_m^*$, together with $\lambda_0^* = 1$, satisfy (3)–(6), then $(S_1^*, \ldots, S_n^*)$ is a (global) minimum for (1).*

Proof. Suppose that $(S_1, \ldots, S_n) \in \mathscr{A}^n$ satisfies $G_j(S_1, \ldots, S_n) \leqslant 0$, $j = 1, \ldots, m$. Then from the convexity of F and Theorem 4.5

$$F(S_1, \ldots, S_n) \geqslant F(S_1^*, \ldots, S_n^*) + \sum_{i=1}^{n} \langle f^i{}_*, \chi_{S_i} - \chi_{S_i^*} \rangle, \tag{18}$$

where the notation for derivatives in the proof of Theorem 3.7 is used. Summing (3) over i with $\lambda_0^* = 1$ and substituting into (18) then give

$$F(S_1, \ldots, S_n) \geqslant F(S_1^*, \ldots, S_n^*) - \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_j^* \langle g_*^{ij}, \chi_{S_i} - \chi_{S_i^*} \rangle. \tag{19}$$

Upon next using in (19) the convexity of the G_j and (4), it follows that

$$F(S_1, \ldots, S_n) \geqslant F(S_1^*, \ldots, S_n^*) - \sum_{j=1}^{m} \lambda_j^* G_j(S_1, \ldots, S_n). \tag{20}$$

But (5) and the fact that $G_j(S_1, \ldots, S_n) \leqslant 0$, $j = 1, \ldots, m$, establish that $F(S_1, \ldots, S_n) \geqslant F(S_1^*, \ldots, S_n^*)$ to complete the proof. ∎

5. Saddlepoints and Duality

Saddlepoint optimality conditions are next obtained for (1) and a Lagrangian duality theory is presented.

Definition 5.1. Let $L(S_1, ..., S_n; \lambda_1, ..., \lambda_m) = F(S_1, ..., S_n) + \sum_{j=1}^{m} \lambda_j G_j(S_1, ..., S_n)$. Then $(S_1^*, ..., S_n^*)$ and $(\lambda_1^*, ..., \lambda_m^*)$ form a saddlepoint for (1) if

$$L(S_1^*, ..., S_n^*; \lambda_1, ..., \lambda_m) \leqslant L(S_1^*, ..., S_n^*; \lambda_1^*, ..., \lambda_m^*) \leqslant L(S_1, ..., S_n; \lambda_1^*, ..., \lambda_m^*) \quad \text{for all } (S_1, ..., S_n) \in \mathscr{A}^n \text{ and } \lambda_1, ..., \lambda_m \geqslant 0. \tag{21}$$

Theorem 5.2. *Let $(X, \mathscr{A}, \mu)$ be a finite atomless measure space with $L_1(X, \mathscr{A}, \mu)$ separable and let F, $G_1, ..., G_m$ be convex. Suppose that there exists $(R_1, ..., R_n) \in \mathscr{A}^n$ such that $G_j(R_1, ..., R_n) < 0$, $j = 1, ..., m$. Then f $(S_1^*, ..., S_n^*)$ is a minimum for (1), there exist $\lambda_1^*, ..., \lambda_m^* \geqslant 0$ such that $(S_1^*, ..., S_n^*)$ and $(\lambda_1^*, ..., \lambda_m^*)$ form a saddlepoint for (1).*

Proof. Let

$$A = \left\{ \begin{array}{l} (v_0, v_1, ..., v_m)\text{: there exists } (S_1, ..., S_n) \in \mathscr{A}^n \text{ such that} \\ v_0 \geqslant F(S_1, ..., S_n),\ v_j \geqslant G_j(S_1, ..., S_n),\ j = 1, ..., m \end{array} \right\}$$

and

$$B = \{(v_0, v_1, ..., v_m): v_0 \leqslant F(S_1^*, ..., S_n^*),\ v_j < 0,\ j = 1, ..., m\}.$$

B is obviously convex, and we next show the closure $\bar{A}$ to be convex. Assume $\bar{A} \neq \varnothing$; otherwise $\bar{A}$ is trivially convex. Fix $\varepsilon > 0$, $\lambda \in [0, 1]$, and $a_1, a_2 \in \bar{A}$. Then there exist $(R_1^1, ..., R_n^1)$, $(R_1^2, ..., R_n^2) \in \mathscr{A}^n$ for which

$$a_p \geqslant \left[F(R_1^p, ..., R_n^p) - \frac{\varepsilon}{2}, G_1(R_1^p, ..., R_n^p) - \frac{\varepsilon}{2}, ..., G_m(R_1^p, ..., R_n^p) - \frac{\varepsilon}{2} \right], \quad p = 1, 2,$$

where the inequality is meant componentwise here. Since $L_1(X, \mathscr{A}, \mu)$ is separable, from Lemma 4.3 there exist sequences $R_i^{1k} \subset R_i^1 \setminus R_i^2$, $R_i^{2k} \subset R_i^2 \setminus R_i^1$ for which

$$\chi_{R_i^{1k}} \xrightarrow{w^*} \lambda \chi_{R_i^1 \setminus R_i^2}, \quad \chi_{R_i^{2k}} \xrightarrow{w^*} (1 - \lambda)\, \chi_{R_i^2 \setminus R_i^1}, \qquad i = 1, ..., n.$$

It follows from the convexity of F, $G_1, ..., G_m$ that there exist k and $R_i = R_i^{1k} \cup R_i^{2k} \cup (R_i^1 \cap R_i^2) \in \mathscr{A}$, $i = 1, ..., n$, such that

$$F(R_1, ..., R_n) \leqslant \lambda F(R_1^1, ..., R_n^1) + (1 - \lambda)\, F(R_1^2, ..., R_n^2) + \frac{\varepsilon}{2}$$

$$G_j(R_1, ..., R_n) \leqslant \lambda G_j(R_1^1, ..., R_n^1) + (1 - \lambda)\, G(R_1^2, ..., R_n^2) + \frac{\varepsilon}{2}, \qquad j = 1, ..., m.$$

Thus $\lambda a_1 + (1-\lambda) a_2 \geqslant [F(R_1, ..., R_n) - \varepsilon,\ G_1(R_1, ..., R_n) - \varepsilon, ..., G_m(R_1, ..., R_n) - \varepsilon]$. But ε is arbitrary, so $\lambda a_1 + (1-\lambda) a_2 \in \bar{A}$ and hence $\bar{A}$ is convex.

Since $\bar{A}$ and B are clearly disjoint, there exists a hyperplane determined by scalars $\lambda_1^*, ..., \lambda_m^*$ separating $\bar{A}$ and B, and therefore A and B. The remainder of the proof follows standard arguments as in [9, p. 218]. ∎

COROLLARY 5.3. *Under the hypotheses of Theorem* 5.2, *let* $(S_1^*, ..., S_n^*)$ *be a minimum for* (1). *Then there exist* $\lambda_1^*, ..., \lambda_m^* \geqslant 0$ *such that*

$$(S_1^*, ..., S_n^*) \text{ minimizes } \left[F(S_1, ..., S_n) + \sum_{j=1}^{m} \lambda_j^* G_j(S_1, ..., S_n) \right]$$

$$\text{subject to } (S_1, ..., S_n) \in \mathscr{A}^n \tag{22}$$

$$G_j(S_1^*, ..., S_n^*) \leqslant 0, \qquad j = 1, ..., m \tag{23}$$

$$\sum_{j=1}^{m} \lambda_j^* G_j(S_1^*, ..., S_n^*) = 0. \tag{24}$$

The proof of the next theorem again follows standard arguments.

THEOREM 5.4. *If* $(S_1^*, ..., S_n^*)$ *and* $(\lambda_1^*, ..., \lambda_m^*)$ *form a saddlepoint for* (1), *then* $(S_1^*, ..., S_n^*)$ *is a minimum for* (1).

COROLLARY 5.5. *If there exist* $(S_1^*, ..., S_n^*)$ *and* $\lambda_1^*, ..., \lambda_m^* \geqslant 0$ *satisfying* (22)–(24), *then* $(S_1^*, ..., S_n^*)$ *is a minimum for* (1).

A Lagrangian dual for problem (1) is next defined, and the duality relations between it and (1) are presented. Theorem 5.7 is the statement of weak duality, while Theorem 5.9 is the strong dual relationship. Proofs are omitted because they are similar to standard mathematical programming arguments. For example, Theorem 5.9 follows from Theorem 5.2 as in [9].

DEFINITION 5.6. The Lagrangian dual for the primal problem (1) is the problem

$$\underset{\lambda_1, ..., \lambda_n \geqslant 0}{\text{maximize}}\ h(\lambda_1, ..., \lambda_m), \tag{25}$$

where $h(\lambda_1, ..., \lambda_m) = \inf\{F(S_1, ..., S_n) + \sum_{j=1}^{m} \lambda_j G_j(S_1, ..., S_n)$: $(S_1, ..., S_n) \in \mathscr{A}^n\}$.

THEOREM 5.7. *If* $(S_1, ..., S_n)$ *is feasible to* (1) *and* $(\lambda_1, ..., \lambda_m)$ *is feasible to* (25), *then* $F(S_1, ..., S_n) \geqslant h(\lambda_1, ..., \lambda_m)$.

COROLLARY 5.8. *If* $(S_1^*, ..., S_n^*)$ *is feasible to the primal,* $\lambda_1^*, ..., \lambda_m^*)$ *is feasible to the dual, and* $F(S_1^*, ..., S_n^*) \leqslant h(\lambda_1^*, ..., \lambda_m^*)$, *then* $(S_1^*, ..., S_n^*)$ *and* $(\lambda_1^*, ..., \lambda_m^*)$ *are optimal to* (1) *and* (25), *respectively. If* (25) *is unbounded, then* (1) *has no feasible* $(S_1, ..., S_n)$; *and if* (1) *is unbounded, then* (25) *has no feasible* $(\lambda_1, ..., \lambda_m)$.

THEOREM 5.9. *Under the hypotheses of Theorem* 5.2, *then*

$$\begin{aligned}\inf\{F(S_1, ..., S_n)&: G_j(S_1, ..., S_n) \leqslant 0, j = 1, ..., m\}\\ &= \sup\{h(\lambda_1, ..., \lambda_m): \lambda_1, ..., \lambda_m \geqslant 0\},\end{aligned} \tag{26}$$

and the rhs of (26) *is attained by* $\lambda_1^*, ..., \lambda_m^* \geqslant 0$. *Furthermore, if the lhs of* (26) *is attained by* $(s_1^*, ..., S_n^*) \in A^n$ (*i.e.,* $(S_1^*, ..., S_n^*)$ *minimizes* (1)), *then* $\sum_{j=1}^m \lambda_j^* G_j(S_1^*, ..., S_n^*) = 0$.

6. REMARKS

There are two types of constraints that have not been considered here. First, inequality constraints were not included in (1) because there is no suitable version for set functions of the inverse function theorem to use in proofs. Second, constraints involving the set operations $\cup$ or $\cap$ (such as $S_i \cap S_j = \phi$, $i \neq j$, and $\bigcup_{i=1}^n S_i = X$) were not considered. Since a theory involving both of the above types of constraints was developed for the special case in [3, 4], it is conceivable that a general theory including such constraints can be established. Subsequent work will be directed at doing so.

REFERENCES

1. M. S. BAZARAA AND C. M. SHETTY, "Nonlinear Programming: Theory and Algorithms," Wiley, New York, 1979.
2. J. H. CHOU, W. S. HSIA, AND T. Y. LEE, On multiple objective programming problems with set functions, *J. Math. Anal. Appl.* **105** (1985), 383–394.
3. H. W. CORLEY AND S. D. ROBERTS, A partitioning problem with applications in regional design, *Oper. Res.* **20** (1972), 1010–1019.
4. H. W. CORLEY AND S. D. ROBERTS, Duality relationships for a partitioning problem, *SIAM J. Appl. Math.* **23** (1972), 490–494.
5. G. DANTZIG AND A. WALD, On the fundamental lemma of Neyman and Pearson, *Ann. Math. Statist.* **22** (1951), 87–93.
6. R. V. HOGG AND A. T. CRAIG, "Introduction to Mathematical Statistics," Macmillan Co., New York, 1978.
7. H. C. LAI AND S. S. YANG, Saddlepoint and duality in the optimization theory of convex-set functions, *J. Austral. Math. Soc. Ser. B* **24** (1982), 130–137.

8. H. C. LAI, S. S. YANG, AND G. R. HWANG, Duality in mathematical programming of set functions: On Fenchel duality theorem, *J. Math. Anal. Appl.* **95** (1983), 223–234.
9. D. G. LUENBERGER, "Optimization by Vector Space Methods," Wiley, New York, 1969.
10. R. J. T. MORRIS, Optimal constrained selection of a measurable subset, *J. Math. Anal. Appl.* **70** (1979), 546–562.
11. J. NEYMAN AND E. S. PEARSON, On the problem of the most efficient tests of statistical hypotheses, *Philos. Trans. Roy. Soc. London Ser. A* **231** (1933), 289–337.
12. W. RUDIN, "Principles of Mathematical Analysis," McGraw-Hill, New York, 1964.
13. W. RUDIN, "Functional Analysis," McGraw-Hill, New York, 1973.
14. K. TANAKA AND Y. MARUYAMA, The multiobjective optimization problem of set function, *J. Inform. Optim. Sci.* **5** (1984), 293–306.