IJMMS 26:6 (2001) 371–383
PII. S0161171201005476
http://ijmms.hindawi.com

OPTIMIZATION PROBLEMS FOR SET FUNCTIONS

SLAWOMIR DOROSIEWICZ

(Received 16 June 2000)

Abstract. This paper gives the formal definition of a class of optimization problems, that is, problems of finding conditional extrema of given set-measurable functions. It also formulates the generalization of Lyapunov convexity theorem which is used in the proof of first-order optimality conditions for the mentioned class of optimization problems.

2000 Mathematics Subject Classification. 90C48.

1. Introduction. This paper concentrates on the analysis of certain class of optimization problems, namely the problems of finding conditional extrema of set functions. This type of problems has been shown to arise in interval and pointwise estimation of probability distribution, testing hypothesis (in particular during the construction of the strongest tests)—proof of Neyman-Pearson lemma (cf. [3, 4, 10]). These problems also appear in analysis of some transportation problems, especially computations of traffic assignment in transportation networks including finding optimal systems of regular lines (see [7]). In future, one can expect the widening possibilities of applications of this class of optimization problems.

The considerations are divided into four parts. The first part consists of general remarks on set functions and optimization problems for these functions. Sections 3 and 4 have an auxiliary character. They include some special properties of bounded measures (the generalization of Lapunov convexity theorem (Lemma 3.1 and Theorem 3.2) and the definition of differentiability of set functions). The most important results are presented in Section 5. There are theorems and some corollaries which formulate the properties of solutions of optimization problems for set functions. These necessary conditions for optimality seem to be the generalization of the results formulated in [5, 6, 9, 11, 12]. The mentioned theorems are similar to the well-known Kuhner-Tucker first-order optimality conditions. The main difficulty to obtain these results is caused by the structure of the union of feasible solutions—this family does not usually have the useful structures: compactness, convexity, the structure of a Banach space.

2. Optimization problems for set functions. Let (X,M,μ) be a measurable space with bounded measure $\mu : M \to \mathbb{R}$.

Definition 2.1. Every map (for a given $k \in NM^k$ denotes the cartesian product $\underbrace{M\times\cdots\times M}_{n \text{ factors}}$) $F : M^k \to \mathbb{R}$ is called a function of measurable set or simply a set function.

Example 2.2. We consider the μ-integrable functions $f_i : X \to \mathbb{R}$ $(i = 1,\dots,k)$. An important family consists of set functions

$$F : M^k \longrightarrow \mathbb{R}, \qquad F(S) = \sum_{i=1}^{k} \int_{S_i} f_i \, d\mu. \tag{2.1}$$

The obvious generalization is the following:

$$F : M^k \longrightarrow \mathbb{R}, \qquad F(S) = u\left(\int_{S_1} f_1 \, d\mu, \dots, \int_{S_k} f_k \, d\mu \right), \tag{2.2}$$

where $u : \mathbb{R}^k \to \mathbb{R}$ is given.

We formulate the general form of optimization problem for a set function. This problem relies on finding extrema (without loss of generality—minima) of a given set function $F_0 : M^k \to \mathbb{R}$:

$$F_0(S) \longrightarrow \min, \tag{2.3}$$

under some constraints (conditions). They will have two main forms:

- constraints defined by other set functions $F_i : M^k \to \mathbb{R}$, where $i = 1,\dots,s$

$$F_i(S) \le 0 \quad (i = 1,\dots,s); \tag{2.4}$$

- constraints imposed directly on the sets, or equivalently on characteristic functions of these sets

$$\chi_S(x) \in v(x) \quad \text{for } \mu\text{-almost all (a.a.) } x \in X, \tag{2.5}$$

  where $v : X \to 2^{\{0,1\}^k}$ is a given function taking values in the family of all subsets of $\{0,1\}^k$, χ_S denotes the vector of characteristic functions of the element $S \in M^k$, if $S = (S_1,\dots,S_k)$, then $\chi_S = (\chi_{S_1},\dots,\chi_{S_k})$ and

$$\chi_{S_j}(x) = \begin{cases} 1 & \text{if } x \in S_j, \\ 0 & \text{if } x \in X - S_j. \end{cases} \tag{2.6}$$

The notions: feasible solution and optimal one can be defined in a standard way. The feasible solution is an element $S \in M^k$ satisfying conditions (2.4) and (2.5); an optimal solution is feasible which minimizes the value of (2.3) among other feasible solutions.

To investigate the properties of optimal solutions of (2.3), (2.4), and (2.5) we need some special properties of measures and set functions.

3. Some properties of bounded measure. This part provides the proof of some properties of bounded measures. It follows that a wide class of optimization problems with set functions is equivalent to the problems of convex mathematical programming in Euclidean space and can be examined using appropriate methods of convex analysis.

Let (X, M, μ) be a space with bounded measure. For $S = (S_1,\dots,S_k) \in M^k$ and $U = (U_1,\dots,U_m) \in M^m$, the symbols χ_S, χ_U denote the vector $[\chi_{S_1},\dots,\chi_{S_k}]^T$ and $[\chi_{U_1},\dots,\chi_{U_m}]^T$, respectively. The aim is to prove the following lemma.

Lemma 3.1. *Suppose that*
(1) *μ is bounded and nonatomic;*
(2) *C is a (totally) unimodular matrix;*
(3) *for all $x \in X$ the system of equations*

$$Cy = \chi_U(x) \tag{3.1}$$

is consistent.

The set

$$W = \{(\mu_{S_1}, \ldots, \mu_{S_k}) : S \in M^k,\ C\chi_S = \chi_U\} \tag{3.2}$$

is compact and convex. If μ is bounded (but not necessarily nonatomic), then the set W is compact (but it may not be connected).

Proof. Suppose first that μ is nonatomic. This part of the proof is similar to the proof of Lapunov convexity theorem and uses the ideas of Lindenstrauss (see [8]).

Without loss of generality, we assume that μ is scalar and nonnegative. Let

$$\begin{aligned} \mathcal{L} &= \prod_{i=1}^{k} L^\infty(X, M, \mu), \\ K &= \{g = (g_1, \ldots, g_k) \in \mathcal{L} : 0 \le g_i \le 1,\ Cg = \chi_U\}, \\ J &: \mathcal{L} \longrightarrow \mathbb{R}^k, \qquad J(g) = \left(\int_X g_1\, d\mu, \ldots, \int_X g_k\, d\mu\right). \end{aligned} \tag{3.3}$$

The set $\mathcal{L}$ with the standard norm $\|g\| = \sum_{i=1}^{k} \|g_i\|_\infty$, where $\|\cdot\|_\infty$ is the norm in $L^\infty(X, M, \mu)$ such that $\|h\|_\infty = \inf_{A:\mu(A)=0} \sup_{x \in X - A} |h(x)|$ is a Banach space.

The set K is an intersection of cartesian product of k sets $\{h \in L^\infty(X, M, \mu) : 0 \le h \le 1\}$ and the set of functions satisfying the equation $Cy = \chi_U$. From Alaoglu theorem (cf. [1, 2]) and Tichonov theorem, we obtain that

$$\prod_{i=1}^{k} \{g \in L^\infty(X, M, \mu) : 0 \le g \le 1\} \tag{3.4}$$

is compact in $*$-weak topology of $\mathcal{L}$. Because the set of solutions of (3.1) is closed, the set K is compact.

Let J_i for $i = 1, \ldots, k$ denotes the ith components of the map J. The functions J_i are linear and continuous in the topology of $L^\infty(X, M, \mu)$ defined by the norm $\|\cdot\|_\infty$. Dunford-Schwartz theorem implies that J_i are continuous in the $*$-weak topology of $L^\infty(X, M, \mu)$. Therefore the map J is continuous in the weak $*$-topology of $\mathcal{L}$. The set $J(K) \subset \mathbb{R}^k$ is the image of a compact and convex set in the linear map, therefore it is also compact and convex.

We show that K is nonempty. We denote by $\mathcal{R}_x$, for any given $x \in X$, the set of nonnegative solutions of the system (3.1). Under the assumptions, for μ-a.a. $x \in X$ we have $\mathcal{R}_x \neq \emptyset$. Because $\{(x, z) : x \in X,\ z \in \mathcal{R}_x\} \in M \otimes \beta(R^k)$, then there exists a measurable selector v_0 of the family $\{\mathcal{R}_x : x \in X\}$ (i.e., the map $v_0 : X \to \mathbb{R}^k$, such that for μ-a.a. $x \in X$, $v_0(x) \in \mathcal{R}_x$ holds). Definition of K implies that $v_0 \in K$, thus K is nonempty. Of course W is nonempty too. We prove that $J(K) = W$. Because $\bigcup_{a \in \mathbb{R}^k} J^{-1}(\{a\}) \cap K = K$

and $W \subset J(K)$, then it suffices to show that either the set $J^{-1}(\{a\}) \cap K$ is empty or it includes the map χ_S for any $S \in M^k$.

Fix $a \in J(K)$. The set $J^{-1}(\{a\}) \cap K$ is nonempty, compact, and convex. Krein-Millman theorem follows that this set has at least one extreme point. Let $f = (f_1, \dots, f_m)$ be such a point. We will show that each component of f takes only values $0, 1$ (μ-a.e. on X).

Let X_C denote the set $\{y : Cy = 0\}$. If $\dim X_C = 0$, then for every $x \in X$, $\mathcal{R}_x$ contains one element. This unique element is obviously the extreme point of $\mathcal{R}_x$. As the matrix C is unimodular we have that every component of $r(x) \in \mathcal{R}_x$ is either 0 or 1. The selector v_0 of the family $\{\mathcal{R}_x : x \in X\}$ (the map $v_0(x) = r(x)$ for $x \in X$) can be written as $v_0 = \chi_S$ for any $S \in M^k$. This means that Lemma 3.1 holds in this case.

Suppose that for any $i \in \{1, \dots, k\}$, f_i is not a characteristic function of any measurable set. In this case, $\dim X_C > 0$ and there exist $\epsilon > 0$ and the set $E_0 \in M$ such that $\mu(E_0) > 0$ and $\epsilon < f_i(x) < 1 - \epsilon$ for every $x \in E_0$. The unimodularity of C implies that every component of the extreme point of $\mathcal{R}_x$ is equal to 0 or 1. Therefore, for every $x \in E_0$, the vector $f(x)$ is not an extreme point of $\mathcal{R}_x$. Hence there exist a set $E_0' \subset E_0$, a vector $c \in X_C$, $c \neq 0$, and a number $\delta > 0$, such that

$$\mu(E_0') > 0, \quad \forall x \in E_0', \quad (\lambda \in [0, \delta] \Longrightarrow f_i(x) + \lambda c \in +\mathcal{R}_x - \mathcal{R}_{x,ex}), \tag{3.5}$$

where $\mathcal{R}_{x,ex}$ denotes the set of extreme points of $\mathcal{R}_x$.

Because μ is nonatomic, there exists a measurable set $E_1 \subset E_0'$ such that $0 < \mu(E_1) < \mu(E_0')$. Choose the measurable sets $G_1 \subset E_1$ and $G_2 \subset E_0' - E_1$ with positive measures and the numbers $s_1, s_2 \in R(s_1 + s_2 > 0)$, such that

$$\begin{gathered} \max(|s_1|, |s_2|) < \delta; \\ s_1(\mu(E_1) - 2\mu(G_1)) + s_2(\mu(E_0' - E_1) - 2\mu(G_2)) = 0. \end{gathered} \tag{3.6}$$

Consider the function

$$h = s_1(\chi_{E_1} - 2\chi_{G_1}) + s_2(\chi_{E_0' - E_1} - 2\chi_{G_2}). \tag{3.7}$$

It is easy to check that $\int_X h \, d\mu = 0$ and for sufficiently small s_1, s_2 functions $f^+ = f + hc$, $f^- = f - hc$ belong to $J^{-1}(\{a\}) \cap K$. Because $f = 1/2(f^+ + f^-)$, then f is not an extreme point of $J^{-1}(\{a\}) \cap K$. This contradicts the definition of f. For $i = 1, \dots, k$ and μ-a.a. $x \in X$, we have $f_i(x) \in \{0, 1\}$, which implies the existence of an element $S \in M^k$ such that $\chi_S = f$. This finishes the proof of the first part of Lemma 3.1.

Suppose that the measure μ is only bounded (not necessarily nonatomic). It is easy to see that the following decomposition holds:

$$\mu = \mu_{na} + \mu_a, \tag{3.8}$$

where the measures μ_{na}, μ_a are singular and the first of them is nonatomic. Of course μ_a is concentrated on the family of atoms of μ. Let A_μ denote the family of μ's atoms and a_μ the sum of sets of this family. For $S = (S_1,\dots,S_k) \in M^k$, we denote $S - a_\mu = (S_1 - a_\mu,\dots,S_k - a_\mu)$ and $S \cap a_\mu = (S_1 \cap a_\mu,\dots,S_k \cap a_\mu)$. For every $A \in M$ we have an obvious decomposition $A = (A - a_\mu) \cup (A \cap a_\mu)$, so $\chi_U = \chi_{U-a_\mu} + \chi_{U\cap a_\mu}$ and $\chi_S = \chi_{S-a_\mu} + \chi_{S\cap a_\mu}$. This means that

$$W = W_{na} W_a, \tag{3.9}$$

where

$$\begin{aligned} W_{na} &= \{(\mu_{na}(S_1),\dots,\mu_{na}(S_k)) : S \in M^k,\ C\chi_S = \chi_{U-a_\mu}\}, \\ W_a &= \{(\mu_a(S_1),\dots,\mu_a(S_k)) : S \in M^k,\ C\chi_S = \chi_{U\cap a_\mu}\}. \end{aligned} \tag{3.10}$$

From the first part of Lemma 3.1, it follows that W_{na} is compact. It remains to prove that W_a also has this property. Note that W_a can be rewritten as follows:

$$\begin{aligned} W_a &= \left\{ \sum_{b\in A_\mu} (\mu(S_1\cap b),\dots,\mu(S_k\cap b)) : S \in M^k,\ C\chi_S = \chi_{U\cap a_\mu} \right\} \\ &= \left\{ \sum_{b\in A_\mu} (\mu(b)x_{1b},\dots,\mu(b)x_{bk}) : (x_{b1},\dots,x_{bk}) \in \{0,1\}^k \cap \mathcal{R}_b,\ b \in A_\mu \right\}, \end{aligned} \tag{3.11}$$

where $\mathcal{R}_b$, for $b \in A_\mu$, satisfies the condition

$$\mu(\{x \in b : \mathcal{R}_x = \mathcal{R}_b\}) = \mu(b). \tag{3.12}$$

The set W_a is equal $\phi(U)$, where ($\prod$ denotes the cartesian product of set)

$$\begin{aligned} &U = \prod_{b\in A_\mu} (\{0,1\}^k \cap \mathcal{R}_b), \\ &\phi : [0,1]^{k|A_\mu|} \longrightarrow \mathbb{R}^k, \qquad \phi((x_{bi})_{b\in A_\mu, i=1,\dots,k}) = \sum_{b\in A_\mu} (\mu(b)x_{b1},\dots,\mu(b)x_{bk}). \end{aligned} \tag{3.13}$$

Continuity of ϕ and the compactness of U imply that W_a is compact. This completes the proof. □

Note that in Lemma 3.1, the system $C\chi_S = \chi_U$ was rewritten as the condition $\chi_S(x) \in \mathcal{R}(x)$ for μ-a.a. $x \in X$. Unimodularity of matrix C and other assumptions guarantee that the sets $\mathcal{R}(x)$ were nonempty subsets of $\{0,1\}^k$. Taking the measurable map $v : X \to 2^{\{0,1\}^k} - \{\varnothing\}$ (the symbol $2^{\{0,1\}^k}$ denotes the family of all subsets of $\{0,1\}^k$), measurability of v means that $v^{-1}(B) \in M$ for every $B \subset \{0,1\}^k$, and replacing (3.1) by (the map χ_S satisfying such condition will be called the selection of v)

$$\chi_S(x) \in v(x) \quad \text{for } \mu\text{-a.a. } x \in X, \tag{3.14}$$

it is easy to prove the following theorem.

Theorem 3.2. *Let μ be the bounded, nonatomic measure and let $v : X \to 2^{\{0,1\}^k}$ be the measurable map, such that for μ-a.a. $x \in X v(x) \neq \emptyset$. Then the set*

$$W = \{(\mu(S_1),\dots,\mu(S_k)) : S = (S_1,\dots,S_k) \in M^k,\ \chi_S \text{ is selection of } v\} \tag{3.15}$$

is compact and convex.

Proof. The collection $\{v^{-1}(\{B\}) : B \subset \{0,1\}^k\}$ is the partition of X. It follows that $W = \sum_{B\subset\{0,1\}^k} W_B$, where

$$\begin{aligned} W = \Big\{&(\mu(S_1),\dots,\mu(S_k)) : S = (S_1,\dots,S_k) \in M^k, \\ &S_j \subset v^{-1}(B),\ \chi_S|_{v^{-1}(B)} \text{ is a selection of } v|_{v^{-1}(B)}\Big\}, \end{aligned} \tag{3.16}$$

see Lemma 3.1. The set W is nonempty and compact since any W_B has these properties (see Lemma 3.1). This completes the proof. □

4. Differentiability of set functions. The definition presented below seems to be a generalization (on the case of bounded measure with any family of atoms) of the notion of differentiability introduced in [5, 9].

Before defining a derivative of set functions, we note a few facts. Note that any set function can be equivalently defined on the family of characteristic functions of measurable subsets of X. We denote $\hat{M} = \{\chi_S : S \in M\}$. Any element $S \in M^k$ corresponds to k characteristic functions $\chi_S \in L^\infty(X,M,\mu,R^k)$ (or $\chi_S \in L^\infty(X,M,\mu,R^k)$, where $L^\infty(X,M,\mu,R^k)$ denotes the space of essentially bounded functions), this implies equivalence between set functions and maps defined on $\hat{M}^k$, set function F corresponds to $\hat{F} : \hat{M}^k \to R$, $\hat{F}(\chi_A) = F(A)$. Additionally, if we can identify the sets whose symmetric difference has measure zero, then M^k can be viewed as a metric space—the distance can be defined as follows: if $A = (A_1,\dots,A_k) \in M^k$, $B = (B_1,\dots,B_k) \in M^k$, then

$$\rho : M^k \times M^k \longrightarrow R, \qquad \rho(A,B) = \sum_{i=1}^{k} \mu(A_i \triangle B_i), \tag{4.1}$$

where $\triangle$ denotes the symmetric difference of sets.

Of course, the derivative of the function can be easily defined, if its domain has the structure of a linear and metric space. Unfortunately, M^k (or equivalently $\hat{M}^k$) has no "natural" linear structure. Additionally $\hat{M}^k$ is not a convex nor a closed subset of the space of integrable or essentially bounded functions. These facts do not make impossible defining differentiability, but they require making the appropriate modifications.

Definition 4.1. We say that a set function F is differentiable at $S^0 \in M^k$ if for any $S \in M^k$ there exist an integrable function $f_{S^0,S\cap a_\mu} : X - a_\mu \to \mathbb{R}^k$ and the map $\phi_{S^0} : M^k \to \mathbb{R}$, such that the following decomposition holds:

$$\begin{aligned} F(S) - F(S^0) &= \int_X f_{S^0,S\cap a_\mu}(\chi_S - \chi_{S^0})\,d\mu + \phi_{S^0}(S\cap a_\mu) + R_{S^0}(S), \\ R_{S^0}(S) &= o(\rho(S-a_\mu, S^0 - a_\mu)). \end{aligned} \tag{4.2}$$

The pair $((f_{S^0,S\cap a_\mu})_{S\in M}, \phi_{S^0})$ is called the derivative of F at S^0.

Example 4.2. The set functions (2.1) are differentiable. Moreover,

$$f_{S^0,S\cap a_\mu} = f|_{X-a_\mu}, \qquad \phi_{S^0}(S) = \int_X f(\chi_{S\cap a_\mu} - \chi_{S^0\cap a_\mu})\,d\mu. \tag{4.3}$$

The notion of differentiability has some "good" properties. For example, the uniqueness property: any set function $F : M^k \to \mathbb{R}$ has, at given $S^0 \in M^k$, no more than one derivative. The proof can be found in [7].

In the special case when μ is nonatomic, we have $a_\mu = \emptyset$ and for every $S \in M^k S \cap a_\mu = \emptyset \overset{\text{def}}{=} (\emptyset,\dots,\emptyset) \in M^k$ which means that $f_{S^0,S\cap a_\mu} = f_{S^0,\emptyset}$ and $\phi_{S^0} \equiv 0$ for $S \in M^k$.

5. Properties of solutions of optimization problems with set functions. In this section, we formulate the necessary conditions for optimality in the problems (2.3), (2.4), and (2.5) with differentiable set functions F_i $(i = 0,1,\dots,s)$. The derivatives of these functions are denoted by $(f_{S,i},0)$, respectively, the components of $f_{S,i}$ are denoted by $f_{S,ij}$, where $j = 1,\dots,k$.

We begin by considering the case with nonatomic measure.

Theorem 5.1. *Let $S^* \in M^k$ be an optimal solution of (2.3), (2.4), and (2.5).*

(1) *Suppose that*

(a) *μ is bounded and nonatomic;*

(b) *F_i $(i = 0,1,\dots,s)$ are differentiable at S^* in the sense of Definition 4.1;*

(c) *the map $v : X \to 2^{\{0,1\}^k} - \{\emptyset\}$ is measurable; then there exist the nonnegative numbers $\lambda_0^*,\lambda_1^*,\dots,\lambda_s^*$, not simultaneously equal to zero, such that for every $S \in M^k$ satisfying $\chi_S(x) \in v(x)$ for $x \in X$, the following inequality holds:*

$$\sum_{i=0}^{s} \lambda_i^* \int_X f_{S^*,i}(\chi_S - \chi_{S^*})\,d\mu \ge 0. \tag{5.1}$$

Moreover, for $i = 1,\dots,s$,

$$\lambda_i^* F_i(S^*) = 0. \tag{5.2}$$

(2) *If there exists $\hat{S} \in M^k$, such that $\chi_{\hat{S}}$ is measurable selection of v satisfying, for $i = 1,\dots,s$, the inequality*

$$F_i(\hat{S}) + \int_X f_{S^*,i}(\chi_{\hat{S}} - \chi_{S^*})\,d\mu < 0, \tag{5.3}$$

then in (5.1) we may put $\lambda_0^* = 1$.

Proof. The presented conditions are similar to the Lusternik theorem describing necessary conditions for conditional extrema of functionals in Banach spaces. The proof is divided into a sequence of steps.

Step 1. We show that

$$\begin{aligned} V = \Big\{&(x_0,\dots,v_s) \in \mathbb{R}^{s+1} : \exists S \in M^k,\ \chi_S \text{ selection of } v, \\ &\forall i = 1,\dots,s,\ x_i \ge \int_X f_{S^*,i}(\chi_S - \chi_{S^*})\,d\mu + (1-\delta_{i0})F_i(S^*)\Big\}, \end{aligned} \tag{5.4}$$

where δ denotes Kronecker delta, is convex. Fix $p^1,p^2 \in V$, $a \in [0,1]$. There exist $S^1,S^2 \in M^k$ such that, for $m = 1,2$, $i = 0,1,\dots,s$,

$$p_i^m \ge \int_X f_{S^*,i}(\chi_{S^m} - \chi_{S^*})\,d\mu + (1-\delta_{i,0})F_i(S^*) \tag{5.5}$$

and $\chi_{S^m}(x) \in v(x)$ for μ-a.a. $x \in X$. Lemma 3.1 applied to the measure

$$M \ni A \longmapsto \nu(A) = \left(\int_X f_{S^*,i}\chi_A \, d\mu\right)_{i=0,1,\dots,s} \in \mathbb{R}^{s+1} \tag{5.6}$$

shows that the set

$$\left\{\left(\int_X f_{S^*,0}\chi_S \, d\mu, \dots, \int_X f_{S^*,s}\chi_S \, d\mu\right) \in \mathbb{R}^{s+1} : S \in M^k,\ \chi_S(x) \in v(x) \text{ for } x \in X\right\} \tag{5.7}$$

is convex. For any $a \in [0,1]$ there exists an element $S^a \in M^k$ such that

$$\begin{aligned}\left(\int_X f_{S^a,0}\chi_{S^a} \, d\mu, \dots, \int_X f_{S^a,s}\chi_{S^a} \, d\mu\right) &= a\left(\int_X f_{S^1,0}\chi_{S^1} \, d\mu, \dots, \int_X f_{S^1,s}\chi_{S^1} \, d\mu\right) \\ &\quad + (1-a)\left(\int_X f_{S^2,0}\chi_{S^2} \, d\mu, \dots, \int_X f_{S^2,s}\chi_{S^2} \, d\mu\right)\end{aligned} \tag{5.8}$$

and for μ-a.a. $x \in X$, $\chi_{S^a}(x) \in v(x)$.

This implies that

$$ap^1 + (1-a)p^2 \geq \left(\int_X f_{S^a,0}(\chi_{S^a} - \chi_{S^*}) \, d\mu, \dots, \int_X f_{S^a,s}(\chi_{S^a} - \chi_{S^*}) \, d\mu\right) + (1-\delta_{i0})F_i(S^*), \tag{5.9}$$

which means that $ap^1 + (1-a)p^2 \in V$. This finishes the proof of convexity.

Step 2. We show that V is disjoint with $]-\infty, 0[^{s+1}$. If for μ-a.a. $x \in X$ the set $v(x)$ has only one element, then for $S \in M^k$ satisfying the condition $\chi_S(x) \in v(x)$ we have $S = S^*$. The set V is disjoint with $]-\infty, 0[^{s+1}$ and the inequality (5.1) obviously holds.

Assume that there exists $S^0 \neq S^*$ such that $\chi_{S^0}(x) \in v(x)$ for μ-a.a. $x \in X$. Suppose that $V \cap]-\infty, 0[^{s+1} \neq \varnothing$. Hence there exists $S^0 \in M^k$, such that for all $i = 0, 1, \dots, s$ the following inequality holds:

$$\int_X f_{S^*,i}(\chi_{S^0} - \chi_{S^*}) \, d\mu + (1-\delta_{i0})F_i(S^*) \leq 0. \tag{5.10}$$

Lapunov convexity theorem applied to the measure

$$M \ni A \longmapsto \left(\int_A |\chi_{S^0} - \chi_{S^*}| \, d\mu, \int_A f_{S^*,0}(\chi_{S^0} - \chi_{S^*}) \, d\mu, \dots, \int_A f_{S^*,s}(\chi_{S^0} - \chi_{S^*}) \, d\mu\right) \in \mathbb{R}^{s+2}, \tag{5.11}$$

where $|\chi_{S^0} - \chi_{S^*}| = \sum_{j=1}^k |\chi_{S^0_j} - \chi_{S^*_j}|$, implies that for any $\alpha \in [0,1]$ there exists a μ-measurable set

$$\Omega^\alpha \subset \{x \in X : \chi_{S^*}(x) \neq \chi_{S^0}(x)\} \tag{5.12}$$

such that

$$\alpha \int_X |\chi_{S^0} - \chi_{S^*}| \, d\mu = \int_{\Omega^\alpha} |\chi_{S^0} - \chi_{S^*}| \, d\mu, \tag{5.13}$$

and for every $i = 0, \dots, s$,

$$\alpha \int_X f_{S^*,i}(\chi_{S^0} - \chi_{S^*}) \, d\mu = \int_{\Omega^\alpha} f_{S^*,i}(\chi_{S^0} - \chi_{S^*}) \, d\mu. \tag{5.14}$$

Consider for fixed $\alpha \in [0,1]$, the map

$$u^\alpha : X \longrightarrow \mathbb{R}^k, \qquad u^\alpha(x) = \chi_{S^*}(x) + (\chi_{S^0}(x) - \chi_{S^*}(x))\chi_{\Omega^\alpha}(x), \tag{5.15}$$

and we denote by S^α the element M^k which corresponds to u^α, that is, $\chi_{S^\alpha} = u^\alpha$. It is easy to check that $\chi_{S^\alpha}(x) \in v(x)$ for μ-a.a. $x \in X$. For sufficiently small $\alpha > 0$ the element S^α is closed to S^* with respect to the metric ρ (see (4.1)):

$$\rho(S^\alpha, S^*) = \int_X |\chi_{S^\alpha} - \chi_{S^*}|\, d\mu = \int_{\Omega^\alpha} |\chi_{S^\alpha} - \chi_{S^*}|\, d\mu = \alpha \int_X |\chi_{S^0} - \chi_{S^*}|\, d\mu, \tag{5.16}$$

therefore $\rho(S^\alpha, S^*) \to 0$ if $\alpha \to 0+$.

Differentiability of F_i implies that

$$F_i(S^\alpha) - F_i(S^*) = \int_X f_{S^*,i}(\chi_{S^\alpha} - \chi_{S^*})\, d\mu + R_{i,S^*}(S^\alpha), \tag{5.17}$$

and hence

$$F_i(S^\alpha) - F_i(S^*) = \alpha \int_X f_{S^*,i}(\chi_{S^0} - \chi_{S^*})\, d\mu + R_{i,S^*}(S^\alpha), \tag{5.18}$$

where $R_{i,S^*}(S^\alpha) = o(\rho(S^\alpha, S^*)) = o(\alpha)$ if $\alpha \to 0$. Equation (5.10) gives that the first component in (5.18) is negative, therefore for sufficiently small $\alpha > 0$ its absolute value is greater than the second one. This implies that $F_0(S^\alpha) < F_0(S^*)$ and $F_i(S^\alpha) < F_i(S^*) \le 0$ for $i = 1, \dots, s$. Element S^α is the feasible solution of (2.3), (2.4), and (2.5); the value of the objective function for S^α is less than for S^*. This contradicts the definition of S^*.

Step 3. This step finishes the proof of inequality (5.1), the separation theorem implies that there exist nonnegative numbers $\lambda_0^*, \dots, \lambda_s^*$ not all zero such that, for all $(t_0, t_1, \dots, t_s) \in V$,

$$\sum_{i=0}^{s} \lambda_i^* t_i \ge 0. \tag{5.19}$$

For $S \in M^k$ satisfying $\chi_S(x) \in v(x)$ for μ-a.a. $x \in X$, we have

$$\sum_{i=0}^{s} \lambda_i^* \left(\int_X f_{S^*,i}(\chi_S - \chi_{S^*})\, d\mu + (1 - \delta_{0,i}) F_i(S^*) \right) \ge 0. \tag{5.20}$$

Letting $S = S^*$ we obtain $\int_{i=1}^{s} \lambda_i^* F_i(S^*) \ge 0$. Because $\lambda_i^* \ge 0$ and $F_i(S^*) \le 0$ for $i = 1, \dots, s$, then $\lambda_i^* F_i(S^*) = 0$. This finishes the first part of the proof.

Now suppose that there exists an element $\hat{S} \in M^k$, which satisfies the inequalities (5.3) with $\lambda_0^* = 0$. The inequality (5.20) gives

$$\sum_{i=1}^{s} \lambda_i^* \left(\int_X f_{S^*,i}(\chi_{\hat{S}} - \chi_{S^*})\, d\mu + F_i(S^*) \right) \ge 0. \tag{5.21}$$

Each component of this sum is nonpositive, then $\lambda_i^* = 0$ for $i = 1, \dots, s$. We obtain the contradiction with the fact that $\lambda_0^*, \lambda_1^*, \dots, \lambda_s^*$ are not all zero. Hence $\lambda_0^* > 0$. Dividing both sides of (5.1) by λ_0^*, we obtain a new sequence of numbers where $\lambda_0^* = 1$. The proof is complete. □

Corollary 5.2. *Suppose that $S^* = (S_1^*,\dots,S_k^*) \in M^k$ is an optimal solution of (2.3), (2.4), and (2.5) and this problem does not have the conditions (2.5). If the assumptions of Theorem 5.1 are satisfied, then there exist nonnegative numbers $\lambda_0^*,\lambda_1^*,\dots,\lambda_s^*$, not simultaneously equal to zero, such that for $j = 1,\dots,k$,*

$$\sum_{i=0}^{s} \lambda_i^* f_{S^*,i}(x) \begin{cases} \le 0 & \text{if } x \in S_j^*, \\ \ge 0 & \text{if } x \in X - S_j^*. \end{cases} \tag{5.22}$$

Proof. According to Theorem 5.1, for every $S \in M^k$ we have

$$\sum_{i=0}^{s} \lambda_i^* \int_X f_{S^*,i}(\chi_S - \chi_{S^*})\,d\mu \ge 0. \tag{5.23}$$

Fix $j \in \{1,\dots,k\}$. Inequality (5.23) must hold for every measurable $u = (u_1,\dots,u_k)$ such that $X \to \{0,1\}^k$, where $u_i = \chi_{S_i^*}$ for $i \ne j$. This gives (the symbol $f_{S^*,ij}$ denotes the jth component of $f_{S^*,i}$)

$$\sum_{i=0}^{s} \lambda_i^* \int_X f_{S^*,ij}\left(u_j - \chi_{S_j^*}\right) d\mu \ge 0, \tag{5.24}$$

which is equivalent to (5.22). This completes the proof. □

Corollary 5.3. *Let S^* be the optimal solution of (2.3), (2.4), and (2.5) in which the condition (2.5) can be written as $\chi_S(x) \in V$ for μ-a.a. $x \in X$, where V is a given subset of $\{0,1\}^k$. Theorem 5.1 shows that there exist nonnegative numbers λ_i^* $(i = 0,1,\dots,s)$, which do not vanish simultaneously, such that, for $\bar{v} \in V$ and $x \in \chi_{S^*}^{-1}(\bar{v})$,*

$$\forall v \in V \quad \sum_{i=0}^{s} \lambda_i^* f_{S^*,i}(v - \bar{v}) \ge 0. \tag{5.25}$$

Example 5.4. Let (X,M,μ) be a probabilistic space with nonatomic, bounded measure μ, F_i $(i=0,1,\dots,s)$, the differentiable set functions on M^3. Consider the problem

$$F_0(S_1,S_2,S_3) \longrightarrow \min, \tag{5.26}$$

subject to

$$(S_1,S_2,S_3) \in M^3, \quad F_i(S_1,S_2,S_3) \le 0 \quad (i = 1,\dots,s), \quad S_1 \cup S_2 \subset S_3. \tag{5.27}$$

Note that the last constraint can be written, using the characteristic functions of S_1,S_2,S_3, in the following form:

$$\chi_{S_1 \cup S_2} = \max\left(\chi_{S_1},\chi_{S_2}\right) \le \chi_{S_3}. \tag{5.28}$$

This inequality has five solutions:

$$v_1=(0,0,0), \quad v_2=(0,0,1), \quad v_3=(1,0,1), \quad v_4=(0,1,1), \quad v_5=(1,1,1), \tag{5.29}$$

thus we have

$$(\chi_{S_1}(x),\chi_{S_2}(x),\chi_{S_3}(x)) \in V \quad \text{for } x \in X, \tag{5.30}$$

where $V = \{v_1,v_2,v_3,v_4,v_5\}$.

Suppose that F_i are differentiable. We denote the derivative of F_i at $S = (S_1, S_2, S_3)$ by $f_{S,i}$, and its components by $f_{S,i,1}$, $f_{S,i,2}$, $f_{S,i,3}$, respectively. Assume that the problem (5.26) and (5.27) has the optimal solution $S^* = (S_1^*, S_2^*, S_3^*)$. Corollary 5.3 implies that there exist nonnegative numbers λ_i^* $(i = 0, \dots, s)$ (which do not vanish simultaneously), such that for any $v \in V$,

$$\int_X \mathcal{F}_{S^*}(v - \chi_{S^*})\, d\mu \ge 0, \tag{5.31}$$

where $\mathcal{F}_{S^*} = \sum_{i=0}^{s} \lambda_i^* f_{S^*,i}$. The function $\mathcal{F}_{S^*}$, like $f_{S^*,i}$, has three components which correspond to S_1, S_2, S_3. We denote them by $\mathcal{F}_{S_1^*}, \mathcal{F}_{S_2^*}, \mathcal{F}_{S_3^*}$, respectively.

Consider the inequality (5.31). Putting the functions

$$v^{(k)}(x) = \begin{cases} \chi_{S^*}(x) & \text{for } x \in S_1^* \cup S_2^* \cup S_3^*, \\ v_k & \text{for } x \notin S_1^* \cup S_2^* \cup S_3^*, \end{cases} \tag{5.32}$$

where $k = 1, \dots, 5$, we obtain the system of inequalities which holds μ for a.a. $X - S_1^* \cup S_2^* \cup S_3^*$ (i.e., for x such that $\chi_{S^*}(x) = v_1$),

$$\begin{gathered} \mathcal{F}_{S_1^*}(x) + \mathcal{F}_{S_3^*}(x) \ge 0, \qquad \mathcal{F}_{S_2^*}(x) + \mathcal{F}_{S_3^*}(x) \ge 0, \\ \mathcal{F}_{S_1^*}(x) + \mathcal{F}_{S_2^*}(x) + \mathcal{F}_{S_3^*}(x) \ge 0. \end{gathered} \tag{5.33}$$

For $x \in S_3^* - (S_1^* \cup S_2^*)$ (if $\chi_{S^*}(x) = v_2$),

$$\mathcal{F}_{S_3^*}(x) \le 0, \qquad \mathcal{F}_{S_1^*}(x) \ge 0, \qquad \mathcal{F}_{S_2^*}(x) \ge 0. \tag{5.34}$$

If $x \in S_1^* - S_2^* (\chi_{S^*}(x) = v_3)$, then

$$\mathcal{F}_{S_1^*}(x) \le 0, \qquad \mathcal{F}_{S_2^*}(x) \ge 0, \qquad \mathcal{F}_{S_3^*}(x) \le 0. \tag{5.35}$$

If $x \in S_2^* - S_1^* (\chi_{S^*}(x) = v_4)$, then

$$\mathcal{F}_{S_1^*}(x) \ge 0, \qquad \mathcal{F}_{S_2^*}(x) \le 0, \qquad \mathcal{F}_{S_3^*}(x) \le 0. \tag{5.36}$$

In the case when $x \in S_1^* \cap S_2^* \cap S_3^* (\chi_{S^*}(x) = v_5)$, then

$$\mathcal{F}_{S_1^*}(x) \le 0, \qquad \mathcal{F}_{S_2^*}(x) \le 0, \qquad \mathcal{F}_{S_3^*}(x) \le 0. \tag{5.37}$$

Inequalities (5.33), (5.34), (5.35), (5.36), and (5.37) allow us to determine the optimal solution $(-s)S^* = (S_1^*, S_2^*, S_3^*)$ (if exist) of the problem (5.26) and (5.27).

The considerations presented so far concerned the special case when the measure μ was nonatomic. It is easy to prove some generalizations of Theorem 5.1. We have the following corollary.

Corollary 5.5. *Let S^* be an optimal solution of (2.3), (2.4), and (2.5), with differentiable set functions F_i $(i = 0, \dots, s)$. There exist the nonnegative numbers λ_i^* $(i = 0, \dots, s)$, which do not vanish simultaneously, such that any feasible solution S satisfies the following inequality:*

$$\sum_{i=0}^{s} \lambda_i^* \int_{X - a_\mu} f_{S^*, S^* \cap a_\mu, i}(\chi_S - \chi_{S^*})\, d\mu \ge 0, \tag{5.38}$$

where $((f_{S^*,S,i})_{S \in M^k}, \phi_{S^*,i})$ is the derivative of F_i.

Proof. The element S^* is the optimal solution of (2.3), (2.4), and (2.5) if and only if $S^*-a_\mu=(S_1^*-a_\mu,\dots,S_k^*-a_\mu)$ is the optimal solution of the problem (with decision variables $A=(A_1,\dots,A_k)$)

$$F_0(A\cup S^*\cap a_\mu)\longrightarrow \min, \tag{5.39}$$

subject to

$$\begin{aligned} &A_j\in M,\quad A_j\subset X-a_\mu \quad \text{for } j=1,\dots,k;\\ &F_i(A\cup S^*\cap a_\mu)\le 0 \quad \text{for } i=1,\dots,s;\\ &\chi_A(x)\in v(x) \quad \text{for } \mu\text{-a.a. } x\in X-a_\mu. \end{aligned} \tag{5.40}$$

Measure μ restricted to the family of measurable subsets of $X-a_\mu$ is obviously nonatomic. Applying Theorem 5.1 to the problem (5.39) and (5.40) finishes the proof. □

Corollary 5.7 concentrates on ordinary (unconditional) extrema of set function. This kind of extreme can be viewed as the optimal solutions of (2.3), (2.4), and (2.5), in which the constraints (2.4) do not appear explicitly.

Definition 5.6. We say that the set function $F:M^k\to\mathbb{R}$ has in $S^*\in M^k$ minimum (local minimum), if S^* is the optimal solution of the problem

$$F(S)\longrightarrow \min, \tag{5.41}$$

subject to

$$S\in M^k(\text{there exists } \epsilon>0, \text{ such that } \rho(S^*,S)<\epsilon). \tag{5.42}$$

The necessary condition for optimality gives the following corollary.

Corollary 5.7. *If F has minimum (respectively, local minimum) in $S^*=(S_1^*,\dots,S_k^*)$ and $((f_{S^*,S\cap a_\mu})_{S\in M^k},\phi_{S^*})$ is the derivative of F in S^*, then for any $S\in M^k$ (respectively, there exists $\epsilon>0$, such that $\rho(S,S^*)<\epsilon$) and $j=1,\dots,k$,*

$$\begin{aligned} x\in S_j^*-a_\mu &\Longrightarrow f_{S^*,S^*\cap a_\mu,j}(x)\le 0;\\ x\notin S_j^*-a_\mu &\Longrightarrow f_{S^*,S^*\cap a_\mu,j}(x)\ge 0, \end{aligned} \tag{5.43}$$

$$\phi_{S^*}(S\cap a_\mu)\ge 0. \tag{5.44}$$

Proof. Without loss of generality, it is sufficient to consider the case of local minimum. Minimum of F corresponds to the situation with $\epsilon\ge\mu(X)$. Formulas (5.43) follows directly from Theorem 5.1. To finish the proof, we should show the inequality (5.44). Differentiability of F and optimality of S^* imply that if $\rho(S^*,S)<\epsilon$, then

$$\begin{aligned} 0\le F(S)-F(S^*)=\int_{X-a_\mu} & f_{S^*,S^*\cap a_\mu}(\chi_S-\chi_{S^*})\,d\mu\\ &+\phi_{S^*}(S\cap a_\mu)+o(\rho(S-a_\mu,S^*-a_\mu)). \end{aligned} \tag{5.45}$$

Putting in inequality (5.45) the element $S\in M^k$ satisfying the conditions $\rho(S,S^*)<\epsilon$ and $S-a_\mu=S^*-a_\mu$, we obtain (5.44). The proof is complete. □

References

[1] A. Alexiewicz, *Functional Analysis*, Monografie Matematyczne, vol. 49, Polish Scientific Publishers, Warsaw, 1968 (Polish).

[2] A. V. Balakrishnan, *Analiza Funkcjonalna Stosowana*, PWN, Warsaw, 1992.

[3] J.-R. Barra, *Notions Fondamentales de Statistique Mathématique*, Dunod, Paris, 1971 (French), Maîtrise de mathématiques et applications fondamentales. MR 53#6805. Zbl 257.62004.

[4] J. Bartoszewicz, *Wykłady ze Statystyki Matematycznej* [*Lectures on Mathematical Statistics*], Wydawnictwo Naukowe PWN, Warsaw, 1996 (Polish). CMP 1 624 321.

[5] H. W. Corley, *Optimization theory for n-set functions*, J. Math. Anal. Appl. **127** (1987), no. 1, 193–205. MR 88j:90185. Zbl 715.90096.

[6] H. W. Corley and S. D. Roberts, *A partitioning problem with applications in regional design*, Operations Res. **20** (1972), 1010–1019. MR 48#7924. Zbl 249.90021.

[7] S. Dorosiewicz, *Modelowanie sieci Transportowych. Wybrane Zagadnienia* [*Modelling of Transportation Networks. Selected Topics*], Monografie i Opracowania, vol. 437, Warsaw School of Economics, Warsaw, 1997 (Polish).

[8] J. Lindenstrauss, *A short proof of Liapounoff's convexity theorem*, J. Math. Mech. **15** (1966), 971–972. MR 34#7754. Zbl 152.24403.

[9] R. J. T. Morris, *Optimal constrained selection of a measurable subset*, J. Math. Anal. Appl. **70** (1979), no. 2, 546–562. MR 80h:49013. Zbl 417.49032.

[10] 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. Zbl 006.26804.

[11] P. K. C. Wang, *On a class of optimization problems involving domain variations*, New Trends in Systems Analysis (Proc. Internat. Sympos., Versailles, 1976), vol. 2, Springer-Verlag, Berlin, 1977, Lecture Notes in Control and Information Sciences, pp. 49–60. MR 58:7308. Zbl 372.93028.

[12] G. J. Zalmai, *Optimality conditions and duality for constrained measurable subset selection problems with minmax objective functions*, Optimization **20** (1989), no. 4, 377–395. MR 91b:90199. Zbl 677.90072.

Slawomir Dorosiewicz: Warsaw School of Economics, Warsaw, Poland
E-mail address: doro@sgh.waw.pl