J. Austral. Math. Soc. (Series B) 24 (1982), 130–137

SADDLE POINT AND DUALITY IN THE OPTIMIZATION THEORY OF CONVEX SET FUNCTIONS

HANG-CHIN LAI and SHU-SHIH YANG[1]

(Received 13 April 1981; revised 18 January 1982)

Abstract

For a set function G on an atomless finite measure space $(X, \mathfrak{S}, m)$, we define the subgradient, conjugate set of $\mathfrak{S}$ and conjugate functional of G. It is proved that a minimization problem of set function G has an optimal solution if and only if the Lagrangian on $\mathfrak{S} \times L_1(X, \mathfrak{S}, m)$ has a saddle point (Ω_0, f_0) such that

$$G(\Omega_0) = \inf_{\Omega \in \mathfrak{S}} G(\Omega) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0)$$

where f_0 is an element of the conjugate set $\mathfrak{S}^*$ (for the definition, see the later context).

1. Introduction

The mathematical programming of a set function was first studied by Morris [5], [6]. The authors investigated the minimization problem for a set function in [3] and proved that the Fenchel duality theorem holds for set functions, where we have defined the conjugate set of a σ-algebra and the conjugate functional of a convex set function. In this note we ask what relations hold between the original set function and the conjugate functional in mathematical programming. This question has been investigated by Scott and Jefferson [8–10] for several functionals. In this note our main result will investigate a convex set function in mathematical programming. It is related to convex integral functions on L_∞, see Rockafellar [7].

Let $(X, \mathfrak{S}, m)$ be a finite atomless measure space and G be a convex set function from $\mathfrak{S}$ to $\mathbf{R}$, the real numbers. We consider an optimization problem as follows

$$\underset{\Omega \in \mathfrak{S}}{\text{Minimize}}\, G(\Omega). \tag{1.1}$$

[1]Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, Republic of China.

https://doi.org/10.1017/S0334270000003635 Published online by Cambridge University Press

The existence of a solution of (1.1) is essentially related to its conjugate functional G^* defined on the conjugate set $\mathfrak{S}^*$. In this note, we prove that the minimal point Ω_0 of (1.1) satisfies $G(\Omega_0) = \inf_{\Omega\in\mathfrak{S}} L(\Omega; f_0)$ for some $f_0 \in \mathfrak{S}^*$ if and only if (Ω_0, f_0) is a saddle point of the Lagrangian $L(\Omega; f)$. For this purpose, we begin with some definitions about the subdifferential, conjugate set and conjugate functional of a convex set in Section 2. Section 3 is the main part of this note.

2. Conjugate functionals and subdifferential

Throughout this note, we assume that $(X, \mathfrak{S}, m)$ is a finite atomless measure space and G is a convex set function from a σ-algebra $\mathfrak{S}$ to $\mathbf{R}$ (for the definition, see [3]). We define a subgradient of the set function G as follows.

DEFINITION 1. *An element $f \in L_1(X, \mathfrak{S}, m)$ is said to be a* subgradient *of the convex set function G at a point $\Omega_0 \in \mathfrak{S}$ if it satisfies the inequality*

$$G(\Omega) \geqslant G(\Omega_0) + \langle f, \chi_\Omega - \chi_{\Omega_0}\rangle \quad \text{for all } \Omega \in \mathfrak{S}.$$

For a set function G, its subgradient at a point Ω_0 is not unique, it is a set of the following form:

$$\partial G(\Omega_0) = \{ f \in L_1(X, \mathfrak{S}, m) \mid G(\Omega) \geqslant G(\Omega_0) + \langle f, \chi_\Omega - \chi_{\Omega_0}\rangle \text{ for all } \Omega \in \mathfrak{S}\}. \tag{2.1}$$

We call this set a subdifferential *of G at Ω_0. If $\partial G(\Omega_0) \neq \phi$, then G is said to be* subdifferentiable *at Ω_0.*

If the set function G is convex and differentiable at Ω_0 (see [3]), then

$$\partial G(\Omega_0) = \{ f_{\Omega_0} = DG(\Omega_0)\}$$

where $f_{\Omega_0} = DG(\Omega_0)$ denotes the derivative of G at Ω_0. As G is differentiable at a point, $\Omega_0 \in \mathfrak{S}$, then Ω_0 is a minimal of G on $\mathfrak{S}$ if and only if for any $\Omega \in \mathfrak{S}$,

$$\langle DG(\Omega_0), \chi_{\Omega_0}\rangle \leqslant \langle DG(\Omega_0), \chi_\Omega\rangle \tag{2.2}$$

From the definition of subgradient at Ω_0, it is evident that Ω_0 is the minimal of the functional $G(\Omega) - \langle f, \chi_\Omega\rangle$.

In order to induce the Lagrangian of a set function G, we have to define the conjugate set and conjugate functional with respect to G.

DEFINITION 2. *A subset of $L_1(X, \mathfrak{S}, m)$ which is defined by*

$$\mathfrak{S}^* = \left\{ f \in L_1(X, \mathfrak{S}, m) \mid \sup_{\Omega\in\mathfrak{S}} [\langle f, \chi_\Omega\rangle - G(\Omega)] < \infty \right\}, \tag{2.3}$$

https://doi.org/10.1017/S0334270000003635 Published online by Cambridge University Press

is called the conjugate set *of* $\mathfrak{S}$. *The functional* G^* *on* $\mathfrak{S}^*$ *defined by*

$$G^*(f) = \sup_{\Omega \in \mathfrak{S}} [\langle f, \chi_\Omega \rangle - G(\Omega)], \tag{2.4}$$

for $f \in \mathfrak{S}^*$ *is called the* conjugate functional *of* G.

Evidently, G^* is a convex function (see [3]) and for any $\Omega \in \mathfrak{S}$,

$$G(\Omega) = \sup_{f \in \mathfrak{S}^*} [\langle f, \chi_\Omega \rangle - G^*(f)]. \tag{2.5}$$

DEFINITION 3. *The subdifferential of a conjugate functional* G^* *at a point* $f_0 \in \mathfrak{S}^*$ *is defined to be a subfamily of measurable subset in* $\mathfrak{S}$:

$$\partial G^*(f_0) = \{\Omega \in \mathfrak{S} \mid G^*(f) \geq G^*(f_0) + \langle f - f_0, \chi_\Omega \rangle, \textit{for all } f \in \mathfrak{S}^*\}. \tag{2.6}$$

Each element of $\partial G^*(f_0)$ *is called a* subgradient *of* G^* *at the point* $f_0 \in \mathfrak{S}^*$.

Note that if $(X, \mathfrak{S}, m)$ is a finite atomless measure space then the conjugate transform for the set exists (cf. [3]). Throughout this paper we assume that $(X, \mathfrak{S}, m)$ is a finite atomless measure space.

By the definition of the conjugate functional, we have Young's inequality:

$$G^*(f) + G(\Omega) \geq \langle f, \chi_\Omega \rangle \tag{2.7}$$

for any $\Omega \in \mathfrak{S}$ and $f \in \mathfrak{S}^*$. The question arises whether or not the equality in (2.7) holds. We would give the answer as follows.

PROPOSITION 4. *If* G *is a convex set function on* $\mathfrak{S}$ *with its conjugate functional* G^* *on the conjugate set* $\mathfrak{S}^*$, *then*

(i) $f \in \partial G(\Omega_0)$ *if and only if* $G(\Omega_0) + G^*(f) = \langle f, \chi_{\Omega_0} \rangle$ *whenever* $\Omega_0 \in \mathfrak{S}$,

(ii) $\Omega \in \partial G^*(f_0)$ *if and only if* $G(\Omega) + G^*(f_0) = \langle f_0, \chi_\Omega \rangle$ *whenever* $f_0 \in L_1(X, \mathfrak{S}, m)$.

PROOF. (i) If $f \in \partial G(\Omega_0)$, then by definition, for any $\Omega \in \mathfrak{S}$, we have

$$G(\Omega) \geq G(\Omega_0) + \langle f, \chi_\Omega - \chi_{\Omega_0} \rangle.$$

This implies that

$$\langle f, \chi_{\Omega_0} \rangle \geq G(\Omega_0) + \langle f, \chi_\Omega \rangle - G(\Omega)$$

for all $\Omega \in \mathfrak{S}$, and

$$\langle f, \chi_{\Omega_0} \rangle \geq G(\Omega_0) + \sup_{\Omega \in \mathfrak{S}} [\langle f, \chi_\Omega \rangle - G(\Omega)].$$

https://doi.org/10.1017/S0334270000003635 Published online by Cambridge University Press

It follows that

$$\langle f, \chi_{\Omega_0}\rangle \geqslant G(\Omega_0) + G^*(f),$$

and by Young's inequality, it would imply that

$$G(\Omega_0) + G^*(f) = \langle f, \chi_{\Omega_0}\rangle.$$

Conversely, if $G(\Omega_0) + G^*(f) = \langle f, \chi_{\Omega_0}\rangle$, then by the definition of conjugate functional, we have

$$\langle f, \chi_{\Omega_0}\rangle = G(\Omega_0) + G^*(f) \geqslant G(\Omega_0) + \langle f, \chi_\Omega\rangle - G(\Omega),$$

or

$$G(\Omega) \geqslant G(\Omega_0) + \langle f, \chi_\Omega - \chi_{\Omega_0}\rangle \quad \text{for all } \Omega \in \mathfrak{S}.$$

This implies that $f \in \partial G(\Omega_0)$.

(ii) For $\Omega \in \partial G^*(f_0)$ we have

$$G^*(f) \geqslant G^*(f_0) + \langle f - f_0, \chi_\Omega\rangle,$$

or

$$\langle f_0, \chi_\Omega\rangle \geqslant G^*(f_0) + \langle f, \chi_\Omega\rangle - G^*(f) \quad \text{for all } f \in \mathfrak{S}^*.$$

It follows that

$$\langle f_0, \chi_\Omega\rangle \geqslant G^*(f_0) + \sup_{f\in\mathfrak{S}^*}\left[\langle f, \chi_\Omega\rangle - G^*(f)\right].$$

That is

$$\langle f_0, \chi_\Omega\rangle \geqslant G^*(f_0) + G(\Omega).$$

Hence, by Young's inequality, we obtain

$$G(\Omega) + G^*(f_0) = \langle f_0, \chi_\Omega\rangle.$$

Conversely, if $G(\Omega) + G^*(f_0) = \langle f_0, \chi_\Omega\rangle$, then

$$\langle f_0, \chi_\Omega\rangle = G(\Omega) + G^*(f_0) \geqslant \langle f, \chi_\Omega\rangle - G^*(f) + G^*(f_0),$$

or

$$G^*(f) \geqslant G^*(f_0) + \langle f - f_0, \chi_\Omega\rangle.$$

This means that $\Omega \in \partial G^*(f_0)$.

3. Characterization for the optimality of a set function

We define a function on $\mathfrak{S} \times L_1(X, \mathfrak{S}, m)$ by

$$L(\Omega; f) = \langle f, \chi_\Omega\rangle - G^*(f). \tag{3.1}$$

This function $L(\Omega; f)$ may be called *Lagrangian*. We would show that $L(\Omega; f)$ has a saddle-point property. The following theorem is essential in this paper.

https://doi.org/10.1017/S0334270000003635 Published online by Cambridge University Press

THEOREM 5. *If $\Omega_0 \in \mathfrak{S}$ and $f_0 \in \mathfrak{S}^*$, then the following statements are equivalent.*

(i) *Ω_0 minimizes the problem* (1.1) *so that $G(\Omega_0) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0)$.*

(ii) *(Ω_0, f_0) is a* saddle point *of the Lagrangian $L(\Omega; f)$, that is,*

$$L(\Omega_0; f) \leqslant L(\Omega_0; f_0) \leqslant L(\Omega; f_0)$$

for all $f \in \mathfrak{S}^*$ and $\Omega \in \mathfrak{S}$. Consequently, $L(\Omega_0; f_0) = G(\Omega_0)$.

PROOF. Suppose (i) holds, then, by definition,

$$\begin{aligned} L(\Omega_0; f_0) &= \langle f_0, \chi_{\Omega_0} \rangle - G^*(f_0) \\ &\leqslant \sup_{f \in \mathfrak{S}^*} \left[\langle f, \chi_{\Omega_0} \rangle - G^*(f) \right] \\ &= G(\Omega_0) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0) \leqslant L(\Omega; f_0) \quad \text{for all } \Omega \in \mathfrak{S}. \end{aligned}$$

Since $G(\Omega_0) = \inf L(\Omega; f_0) \leqslant \langle f_0, \chi_{\Omega_0} \rangle - G^*(f_0)$, we have

$$G(\Omega_0) + G^*(f_0) \leqslant \langle f_0, \chi_{\Omega_0} \rangle.$$

Thus, by the Young's inequality, we obtain

$$G(\Omega_0) + G^*(f_0) = \langle f_0, \chi_{\Omega_0} \rangle,$$

and, by Proposition 4(ii), we have $\Omega_0 \in \partial G^*(f_0)$. This implies

$$\langle f_0, \chi_{\Omega_0} \rangle - G^*(f_0) \geqslant \langle f, \chi_{\Omega_0} \rangle - G^*(f) \quad \text{for all } f \in \mathfrak{S}^*.$$

Therefore, (Ω_0, f_0) is a saddle point of $L(\Omega; f)$. That is,

$$L(\Omega_0; f_0) \leqslant L(\Omega_0; f_0) \leqslant L(\Omega; f_0) \quad \text{for all } \Omega \in \mathfrak{S}, f \in \mathfrak{S}^*.$$

Conversely, suppose that (Ω_0, f_0) is a saddle point of the Lagrangian L, then

$$\begin{aligned} G(\Omega) &= \sup_{f \in \mathfrak{S}^*} \left[\langle f, \chi_{\Omega_0} \rangle - G^*(f) \right] \geqslant \langle f_0, \chi_{\Omega_0} \rangle - G^*(f_0) \\ &= L(\Omega; f_0) \geqslant L(\Omega_0; f_0) \geqslant L(\Omega_0; f) \\ &= \langle f, \chi_{\Omega_0} \rangle - G^*(f) \quad \text{for all } f \in \mathfrak{S}^*. \end{aligned}$$

Hence $G(\Omega) \geqslant \sup_{f \in \mathfrak{S}^*}[\langle f, \chi_{\Omega_0} \rangle - G^*(f)] = G(\Omega_0)$ holds for all $\Omega \in \mathfrak{S}$. Therefore, Ω_0 is the minimal point of (1.1). It remains to show that $G(\Omega_0) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0)$. Since (Ω_0, f_0) is a saddle point,

$$\inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0) \geqslant L(\Omega_0; f_0) \geqslant \sup_{f \in \mathfrak{S}^*} L(\Omega_0; f) = G(\Omega_0).$$

On the other hand,

$$G(\Omega_0) \geqslant \langle f_0, \chi_{\Omega_0} \rangle - G^*(f_0) = L(\Omega_0; f_0) \geqslant \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0).$$

It follows that

$$G(\Omega_0) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0).$$

COROLLARY 6. *In order that the supremum* $\sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) = L(\Omega_0; f_0)$ *is attained at a point* $f_0 \in \mathfrak{G}^* \subset L(X, \mathfrak{G}, m)$ *if and only if* $f_0 \in \partial G(\Omega_0)$.

PROOF. Suppose $\sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) = L(\Omega_0; f_0)$. Then

$$\begin{aligned} G(\Omega_0) &= \sup_{f\in\mathfrak{G}^*} \left[\langle f, \chi_{\Omega_0}\rangle - G^*(f)\right] \\ &= \sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) = L(\Omega_0; f_0) \\ &= \langle f_0, \chi_{\Omega_0}\rangle - G^*(f_0). \end{aligned}$$

That is, $G(\Omega_0) + G^*(f_0) = \langle f_0, \chi_{\Omega_0}\rangle$. Hence by Proposition 4, we see that $f_0 \in \partial G(\Omega_0)$.

Conversely, if $f_0 \in \partial G(\Omega_0)$, then, for any $\Omega \in \mathfrak{G}$,

$$\begin{aligned} L(\Omega_0; f) &\leqslant \sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) \\ &= G(\Omega_0) \leqslant G(\Omega) + \langle f_0, \chi_{\Omega_0} - \chi_\Omega\rangle \\ &= G(\Omega) - \langle f_0, \chi_\Omega\rangle + \langle f_0, \chi_{\Omega_0}\rangle. \end{aligned}$$

Hence,

$$\begin{aligned} L(\Omega_0; f) &\leqslant \inf_{\Omega\in\mathfrak{G}} \left[G(\Omega) - \langle f_0, \chi_\Omega\rangle\right] + \langle f_0, \chi_{\Omega_0}\rangle \\ &= -\sup_{f\in\mathfrak{G}^*} \left[\langle f_0, \chi_\Omega\rangle - G(\Omega)\right] + \langle f_0, \chi_{\Omega_0}\rangle \\ &= -G^*(f_0) + \langle f_0, \chi_{\Omega_0}\rangle \\ &= L(\Omega_0; f_0) \quad \text{for all } f \in \mathfrak{G}^*. \end{aligned}$$

It follows that

$$\sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) \leqslant L(\Omega_0; f_0).$$

And then,

$$\sup_{f\in\mathfrak{G}^*} L(\Omega_0; f) = L(\Omega_0; f_0).$$

By the above discussion, we could characterize the following equivalent statements for optimization of a set function.

THEOREM 7. *For* $\Omega_0 \in \mathfrak{G}$ *and* $f_0 \in \mathfrak{G}^* \subset L_1(X, \mathfrak{G}, m)$, *then the following statements are equivalent*:

(i) *A point* $\Omega_0 \in \mathfrak{G}$ *is the minimal of the problem* (1.1), *such that* $G(\Omega_0) = \inf_{\Omega\in\mathfrak{G}} L(\Omega; f_0)$.

(ii) (Ω_0, f_0) *is a saddle point of the Lagrangian* $L(\Omega; f)$, *that is*

$$L(\Omega_0; f) \leqslant L(\Omega_0; f_0) \leqslant L(\Omega; f_0)$$

for all $f \in \mathfrak{S}^*$ *and* $\Omega \in \mathfrak{S}$.

(iii) *A subgradient* $f_0 \in \partial G(\Omega_0)$ *such that* $L(\Omega_0; f_0) = \inf_{\Omega \in \mathfrak{S}} L(\Omega; f_0)$.

In fact, the equivalence of (i) and (ii) follows from Theorem 5, and the equivalence of (ii) and (iii) follows from Corollary 6. Therefore (i), (ii) and (iii) are equivalent.

4. Examples

Let X be an infinite compact subset of $\mathbf{R}^n$, $\mathfrak{S}$ a family of Lebesgue measurable subsets of X, and let m be the Lebesgue measure on $\mathbf{R}^n$. Then $(X, \mathfrak{S}, m)$ is an atomless finite measure space. We consider the following problem.

$$\underset{\Omega \in \mathfrak{S}}{\text{Minimize}}\, G(\Omega) = \int_\Omega g(x)\, dm$$

where g is an integrable function from $\mathbf{R}^n$ into $\mathbf{R}$. Then for any $f \in \mathfrak{S}^* = L_1(X, \mathfrak{S}, m)$, the conjugate functional of G is given by

$$G^*(f) = \sup_{f \in \mathfrak{S}^*} [\langle f, \chi_\Omega \rangle - G(\Omega)]$$

$$= \int_{\Omega_1 \cap X} [f(x) - g(x)]\, dm,$$

where

$$\Omega_1 = \{x \in R^n \mid f(x) \geqslant g(x)\}.$$

The Lagrangian, defined by equation (3.1), is

$$L(\Omega; f) = \langle f, \chi_\Omega \rangle - G^*(f)$$

$$= \int_\Omega g(x)\, dm - \int_{\Omega_1 \cap X} [f(x) - g(x)]\, dm.$$

Then by straightforward calculation, one sees that the equivalent relations of Theorem 7 hold.

5. Additional remark

There are a variety of interesting applications of the set function optimization problem. These include applications in fluid flow [1], electrical insulator design [2] and optimal plasma confinement [11].

Acknowledgements

The authors would like to thank the referees for their valuable suggestions. Hang-chin Lai was partly supported by the National Science Council, Taiwan, Republic of China.

References

[1] D. Begis and R. Glowinski, "Application de la méthode des éléments finis à l'approximation d'un problème de domaine optimal, Méthodes de résolution des problèmes approchés", *Appl. Math. Optim.* 2 (1975), 130–169.

[2] J. Cea, A. Gioan and J. Michel, "Quelques résultats sur l'identification de domaines", *Calcolo* 10 (1973), 133–145.

[3] 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.* (in press).

[4] D. G. Luenberger, *Optimization by vector space methods* (Wiley, New York, 1969).

[5] R. J. T. Morris, "Optimal constrained selection of a measurable subset", *J. Math. Anal. Appl.* 70 (1979), 546–562.

[6] R. J. T. Morris, "Optimization problems involving set functions", Ph.D. Dissertation, University of California, Los Angeles, March 1978; also available as *UCLA Tech. Report UCLA-ENG-*7815, March 1978.

[7] R. T. Rockafellar, "Convex integral functionals and duality", in *Contributions to nonlinear functional analysis* (ed. E. Zarantonello), (Academic, New York, 1971), 215–234.

[8] C. H. Scott and T. R. Jefferson, "Characterizations of optimality for continuous convex mathematical programs. Part 1. Linear constraints", *J. Austral. Math. Soc. B* 21 (1979), 37–44.

[9] C. H. Scott and T. R. Jefferson, "A generalization of geometric programming with an application to information theory", *Inform. Sci.* 12 (1977), 263–269.

[10] C. H. Scott and T. R. Jefferson, "Duality in infinite-dimensional mathematical programming: convex integral functionals", *J. Math. Anal. Appl.* 61 (1977), 251–261.

[11] P. K. C. Wang, "On a class of optimization problems involving domain variations", in *International symposium on new trends in system analysis*, Versailles, France, (Dec. 1976), *Lecture Notes in Control and Information Sciences, No.* 2, (Springer-Verlag, Berlin, 1977).