Title	Hierarchical multi-objective decision systems and power-decentralized systems for general resource allocation problem
Sub Title	
Author	志水, 清孝(Shimizu, Kiyotaka) 相吉, 英太郎(Aiyoshi, Eitaro)
Publisher	慶応義塾大学工学部
Publication year	1980
Jtitle	Keio engineering reports Vol.33, No.2 (1980. 3) ,p.13- 29
JaLC DOI	
Abstract	We study optimization methods for hierarchical power-decentralized systems composed of a coordinating central system and plural semi-autonomous local systems in the lowerlevel, each of which possesses a decision-making unit. Such a decentralized system where both central and local systems possess their own objective function and decision variables is hence a multi-objective system. The basic principle of planning is that the central system allocates resources so as to optimize its own objective, while the local ones optimize their own objectives using the given resources. The lower-level composes a multi-objective programming problem, where local decision-makers minimize a vector objective function in cooperation. Thus, the lower level generates a set of noninferior solutions being parametric w. r. t. the given resources. The central decision-maker, then, chooses an optimal resource allocation and the best noninferior solution corresponding to it from among a set of resource-parametric noninferior solutions. Several theorems and computational methods are obtained based on parametric nonlinear mathematical programming using directional derivatives in order to treat nondifferentiable parametric functions. Note that, essentially, this paper is concerned with a combined theory for multi-objective decision problem and general resource allocation problem.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00330002-0013

KEIO ENGINEERING REPORTS
VOL. 33, NO. 2, pp. 13-29, 1980

HIERARCHICAL MULTI-OBJECTIVE DECISION SYSTEMS AND POWER-DECENTRALIZED SYSTEMS FOR GENERAL RESOURCE ALLOCATION PROBLEM

KIYOTAKA SHIMIZU* AND EITARO AIYOSHI*

Dept. of Instrumental Engineering, Keio University, Hiyoshi Kohoku-ku, Yokohama, 223, Japan

(Received August 30, 1979)

ABSTRACT

We study optimization methods for hierarchical power-decentralized systems composed of a coordinating central system and plural semi-autonomous local systems in the lower-level, each of which possesses a decision-making unit. Such a decentralized system where both central and local systems possess their own objective function and decision variables is hence a multi-objective system. The basic principle of planning is that the central system allocates resources so as to optimize its own objective, while the local ones optimize their own objectives using the given resources. The lower-level composes a multi-objective programming problem, where local decision-makers minimize a vector objective function in cooperation. Thus, the lower level generates a set of noninferior solutions being parametric w.r.t. the given resources. The central decision-maker, then, chooses an optimal resource allocation and the best noninferior solution corresponding to it from among a set of resource-parametric noninferior solutions. Several theorems and computational methods are obtained based on parametric nonlinear mathematical programming using directional derivatives in order to treat nondifferentiable parametric functions. Note that, essentially, this paper is concerned with a combined theory for multi-objective decision problem and general resource allocation problem.

1. Introduction

In decentralized decision-making system, which is composed of plural subsystems, the over-all optimization is achieved from a central point of view, while each subsystem is invested with considerable power. Such a system is characterized by plural decision-makers of the subsystems, each of which can exercise decentralized decisive power to pursue its own goal. Accordingly, the power-decentralized system implies plural objective functions. It is hence a multi-objective system.

When the decisive powers are decentralized, if objectives, powers and constraints are completely separated in correspondence to each decision-maker, in other words if subsystems are noninteracting, then overall system is merely a collection of independent usual problems. Thus, the decentralized system premises that the subsystems have different objectives and mutual interactions. Those interactions are brought in with system equations, constraints etc.

Decentralized systems have a hierarchical structure in most cases. In this paper, we study optimization methods for hierarchical multi-objective decision systems composed of a coordinating central system in the upper level and plural semi-autonomous local systems (subsystems) in the lower one, each of which possesses a decision-making unit. A central decision-maker determines values of decision variables peculiar to the central system (for instance, policy coefficients defining objectives and constraints in the lower level problems), based on the central objective for the over-all system. On the other hand, local decision-makers determine values of decision variables peculiar to the local systems, based on their own objectives, under the restriction of the given parameters from the central. It is noted that a two-level system is such that both central and local systems possess independent decision-making units.

A typical example of the system is a hierarchical planning for resource allocation problems. In economic activities, for instance, the central system distributes its available resources to the local ones so as to optimize its objective consisting of values of products of the local systems and cost of the resources, while the local systems perform optimal production activities utilizing the given resources.

The above mentioned problem is formulated as two-level planning problem. Henceforth, a term "hierarchical multi-objective system" is used in a case when the lower level composes a multi-objective programming problem due to interactions among the subsystems, and a term "hierarchical decentralized system" in a case when the lower level is completely separated into a set of scalar-objective programming problems with the given parameters (allocated resources).

The decentralized two-level planning problems were studied by Geoffrion (1972) and Shimizu (1975, 1976, 1977) for general resource allocation problems. Among studies related to ours exist primal decomposition methods by resource allocation proposed by Kornai (1965) for LP and Geoffrion (1970), Silverman (1972) for convex program. Theories on Multi-objective programming problems were studied in References by DaCunha (1967), Shimizu (1975) and Yu (1974).

2. Hierarchical Multi-objective Systems

In this section, we are concerned with hierarchical systems such that the central decision-maker decides the optimal resources allocation and the local systems try to optimize their own objective functions under the given resources. Thus, the local systems consist of a multi-objective problem.

The lower level consists of N local systems each of which possesses its own decision variable vector $\boldsymbol{x}_n$, an objective function f_n and a constraint vector function $\boldsymbol{g} \leqq \boldsymbol{0}$. And each local decision-maker desires the vector $\boldsymbol{x}_n$ such that its own objective f_n is minimized under the parameter $\boldsymbol{\alpha}_n$ assigned by the center. But there exist mutual interactions among the local systems, because f_n and/or $\boldsymbol{g}_n$ have not only $\boldsymbol{x}_n$ but $\boldsymbol{x}=(\boldsymbol{x}_1^T, \cdots, \boldsymbol{x}_N^T)^T$ as its argument generally. Therefore, the lower level is composed of a group of local systems having mutual interactions and different goals. So, for the lower level problem, we consider a case such that the local decision-makers carry out minimization of the vector objective function $\boldsymbol{f}=(f_1, \cdots, f_N)^T$ with respect to $\boldsymbol{x}$ cooperating together. The lower level decides the best value of the vector $\boldsymbol{x}$ which is dependent on the assigned parameters $\{\boldsymbol{\alpha}_n\}$, thus generates a set of noninferior (Pareto optimal) solutions $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ being parametric with respect to $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1^T, \cdots, \boldsymbol{\alpha}_N^T)^T$. Accordingly, one can not get a unique solution even if $\boldsymbol{\alpha}$ is fixed.

On the other hand, the central decision-maker in the upper level determines an optimal parameter $\boldsymbol{\alpha}^\circ=(\boldsymbol{\alpha}_1^{\circ T}, \cdots, \boldsymbol{\alpha}_N^{\circ T})^T$, based on the central objective function Φ. At the same time, it choose the best noninferior solution $\hat{\boldsymbol{x}}^\circ$ corresponding to $\boldsymbol{\alpha}^\circ$.

As a typical problem of the above mentioned two-level multi-objective decision problem, let us consider the resource allocation problem mentioned in Page 2. Let $\boldsymbol{x}_n$ be production activity of the local system n, and $\boldsymbol{\alpha}_n$ be the resources allocated to it. In this paper, we consider the case when the super objective function Φ is a function of $\boldsymbol{f}$ and $\boldsymbol{\alpha}$. Then, the resource allocation problem is formulated as follows.

$$\min_{\boldsymbol{\alpha}, \hat{\boldsymbol{x}}(\boldsymbol{\alpha})} \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})), \boldsymbol{\alpha}) \tag{1.a}$$

$$\text{subj. to } \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \qquad (\boldsymbol{b}\text{; total resources}) \tag{1.b}$$

$$\begin{pmatrix} f_1(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \\ \vdots \\ f_N(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \end{pmatrix} = \min_{\boldsymbol{x}} \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_N(\boldsymbol{x}) \end{pmatrix} \tag{1.c}$$

$$\begin{aligned} \text{subj. to } & \boldsymbol{g}_1(\boldsymbol{x}) \leqq \boldsymbol{\alpha}_1 \\ & \vdots \\ & \boldsymbol{g}_N(\boldsymbol{x}) \leqq \boldsymbol{\alpha}_N \end{aligned} \tag{1.d}$$

$$\boldsymbol{x} \in X \tag{1.e}$$

where $\boldsymbol{x}=(\boldsymbol{x}_1^T, \cdots, \boldsymbol{x}_N^T)^T$, $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1^T, \cdots, \boldsymbol{\alpha}_N^T)^T$ and $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ represents a parametric noninferior solution.

The constraint (1.b) bounds a total amount of resources, (1.d) represents an upper bound of the resources that are available to the local system n for production activity and (1.e) is a technological restriction not related to the resources. A practical example of $\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}), \boldsymbol{\alpha}))$ is often given as $\Phi=F(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})))+H(\boldsymbol{\alpha})$, where F is a preference function and H is a cost function concerned with resources.

After all, the problem (1) is to find the optimal resource allocation $\boldsymbol{\alpha}^0$ and the best noninferior solution $\hat{\boldsymbol{x}}^\circ=\hat{\boldsymbol{x}}_{\text{best}}(\boldsymbol{\alpha}^\circ)$ corresponding to $\boldsymbol{\alpha}^\circ$ so as to minimize Φ under the constraints (1.b)$\sim$(1.e). Therefore, Φ must provide both roles of determining the optimal parameter and at the same time choosing the best solution from among the noninferior solution set. Note that a value of Φ varies with $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ chosen even if $\boldsymbol{\alpha}$ is fixed.

The problem (1) is regarded as the hierarchical multi-objective decision system, in which there exist N mutually independent local objective function $\{f_n\}$ and a super objective function Φ. Further, it is such a problem that parameterized constrained optimization problems of the lower level are contained in a part of the constraints of the upper level. Therefore, it is not of a type of usual mathematical programming problem.

For simplicity, let $\boldsymbol{f}=(f_1, \cdots, f_N)^T$ and $\boldsymbol{g}=(\boldsymbol{g}_1^T, \cdots, \boldsymbol{g}_N^T)^T$. The problem (1. c) (1. d) (1. e) is then written as follows.

$$\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))=\min_{\boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}) \tag{1. c)$'$}$$

$$\text{subj. to } \boldsymbol{g}(\boldsymbol{x}) \leqq \boldsymbol{\alpha} \tag{1. d)$'$}$$

$$\boldsymbol{x} \in X \tag{1. e)$'$}$$

Now, let us consider the following scalarization problem, termed "$\boldsymbol{\varepsilon}$-constraint problem" (Haimes 1975, Lin 1977), where $(N-1)$ objectives are replaced with $(N-1)$ constraints in relation to the multi-objective programming problem (1. c)~(1. e).

$$\begin{aligned} &\min_{\boldsymbol{x}} f_p(\boldsymbol{x}) \\ &\text{subj. to } \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}) \leqq \boldsymbol{\varepsilon} \\ &\qquad \boldsymbol{g}(\boldsymbol{x}) \leqq \boldsymbol{\alpha} \\ &\qquad \boldsymbol{x} \in X, \end{aligned} \tag{2}$$

where $\boldsymbol{f}_{\bar{p}}(\boldsymbol{x})=(f_1(\boldsymbol{x}), \cdots, f_{p-1}(\boldsymbol{x}), f_{p+1}(\boldsymbol{x}), \cdots, f_N(\boldsymbol{x}))^T$, $\boldsymbol{\varepsilon}=(\varepsilon_1, \cdots, \varepsilon_{p-1}, \varepsilon_{p+1}, \cdots, \varepsilon_N)^T$.

Here, it is necessary for feasibility of the problem (2) that the parameter $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ satisfies the following condition.

$$\begin{aligned} (\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W = \{(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \mid & \text{ There exists } \boldsymbol{x} \in X \text{ such that } \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}) \leqq \boldsymbol{\varepsilon} \\ & \text{and } \boldsymbol{g}(\boldsymbol{x}) \leqq \boldsymbol{\alpha}\} \end{aligned} \tag{3}$$

Next, we consider a domain of the parameter $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ such that an optimal solution $\boldsymbol{x}^\circ(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ to the problem (2) makes all $\boldsymbol{\varepsilon}$-constraint $\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}) \leqq \boldsymbol{\varepsilon}$ binding, that is,

$$W_a = \{(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \mid (\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W,\ \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^\circ(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}))=\boldsymbol{\varepsilon}\}.$$

Then the following theorem holds.

Theorem 1. Suppose $\boldsymbol{\alpha}$ be fixed.

(i) A noninferior solution $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ to the problem (1. c)~(1. e) given $\boldsymbol{\alpha}$ solves the $\boldsymbol{\varepsilon}$-constraint problem (2) for some $\boldsymbol{\varepsilon}$ such that $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W_a$.

(ii) Assume that an optimal solution $\boldsymbol{x}^\circ(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ to the $\boldsymbol{\varepsilon}$-constraint problem (2) for any fixed $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is unique. Then $\boldsymbol{x}^\circ(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is a noninferior solution to the problem (1. c)~(1. e) given $\boldsymbol{\alpha}$.

(Proof) (i) Let $\boldsymbol{x}^\circ(\boldsymbol{\alpha})$ solve the $\boldsymbol{\varepsilon}$-constraint problem for $\boldsymbol{\varepsilon}=\boldsymbol{f}_{\bar{p}}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$. Then it holds that $(\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^\circ(\boldsymbol{\alpha})), f_p(\boldsymbol{x}^\circ(\boldsymbol{\alpha}))) \leqq (\boldsymbol{f}_{\bar{p}}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})), f_p(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})))$. If this relation holds except for equality, it contradicts to the noninferiority of $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$, since $\boldsymbol{g}(\boldsymbol{x}^\circ(\boldsymbol{\alpha})) \leqq \boldsymbol{\alpha}$ and $\boldsymbol{x}^\circ(\boldsymbol{\alpha}) \in X$. Therefore $\boldsymbol{f}(\boldsymbol{x}^\circ(\boldsymbol{\alpha}))=\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$. This indicates that $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ solves the $\boldsymbol{\varepsilon}$-constraint problem (2) for $\boldsymbol{\varepsilon}=\boldsymbol{f}_{\bar{p}}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$ and that $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W_a$.

(ii) We suppose that $\boldsymbol{x}^\circ(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is not a noninferior solution to the problem (1. c)~(1. e) given $\boldsymbol{\alpha}$, then there exists $\tilde{\boldsymbol{x}} \in X$ such that $\boldsymbol{g}(\tilde{\boldsymbol{x}}) \leqq \boldsymbol{\alpha}$ and that

$$\boldsymbol{f}(\tilde{x}) \leqq \boldsymbol{f}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})) \text{ and } f_j(\tilde{x}) < f_j(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})) \text{ for some } j=1, \cdots, N.$$

a) If $f_p(\tilde{x}) < f_p(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}))$, since $\boldsymbol{f}_{\bar{p}}(\tilde{x}) \leqq \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})) \leqq \boldsymbol{\varepsilon}$, $\boldsymbol{g}(\tilde{x}) \leqq \boldsymbol{\alpha}$ and $\tilde{x} \in X$, $\tilde{x}$ becomes feasible for the $\boldsymbol{\varepsilon}$-constraint problem (2). This contradicts to the fact that $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ solves the $\boldsymbol{\varepsilon}$-constraint problem (2).

b) If $f_p(\tilde{x}) = f_p(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}))$, by uniqueness of $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$, we must have $\tilde{x} \equiv \boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. On the other hand, it must also hold

$$f_j(\tilde{x}) < f_j(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})) \qquad \text{for some } j \neq p.$$

But the above relation never holds as $\tilde{x} \equiv \boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$.

Let us assume:

Assumption (i) An optimal solution to the $\boldsymbol{\varepsilon}$-constraint problem for any fixed $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is unique.

Then by use of Theorem 1, under the above assumption it turns out that the problem (1) may be equivalently represented as follows using the $\boldsymbol{\varepsilon}$-constraint problem (2) and the condition (3):

$$\min_{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}} \Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})), \boldsymbol{\alpha}) \tag{4.a}$$

$$\text{subj. to } \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{4.b}$$

$$(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W_a. \tag{4.c}$$

$$f_p(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})) = \min_{\boldsymbol{x}} f_p(\boldsymbol{x}) \tag{4.d}$$

$$\text{subj. to } \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}) \leqq \boldsymbol{\varepsilon} \tag{4.e}$$

$$\boldsymbol{g}(\boldsymbol{x}) \leqq \boldsymbol{\alpha} \tag{4.f}$$

$$\boldsymbol{x} \in X, \tag{4.g}$$

where the problem (4.d)~(4.g) is a parametric optimization problem having the constraints with the right hand side parameter $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is a parametric optimal solution to the scalarization problem (4.d)~(4.g), and is assumed to be uniquely determined when $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is fixed.

For development of computational methods one can regard $\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}$ as the parameters of the same kind mathematically, although in practice $\boldsymbol{\alpha}$ is a resource variable and $\boldsymbol{\varepsilon}$ is a parameter of the $\boldsymbol{\varepsilon}$-constraint problem. A difficulty to solve the problem is due to the fact that it is almost impossible to obtain a parametric optimal solution $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ for the lower level problem (4.d)~(4.g) in explicit form. But if the parameter $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is fixed, one can attempt to solve the problem by means of some nonlinear programming technique. Therefore, we consider an iterative algorithm to search for a better value than the current $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ based on the information of the solution for the lower level problem (4.d)~(4.g) with current $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. Namely, the iterative procedure consists of the center asking the local systems what would happen if the parameter vector were set at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$, to which the lower level responds by giving some local information concerning the optimal solution for the lower level problem. The center then uses this information in a prescribed manner to determine a revised trial setting for $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$.

By the way, the parametric solution $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ may not always be differentiable with respect to $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$, hence, neither is Φ. This is the principal difficulty in solving the problem. But, we can apply a feasible direction method using a directional derivative of the function

$$v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=f_p(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})), \tag{5}$$

in the case when Φ is a function of $\boldsymbol{f}$ as the eqn. (4.a).

In order to make further theoretical progress, we impose the following conditions.

Assumption (ii) Φ is convex, continuous and differentiable with respect to $(\boldsymbol{f}, \boldsymbol{\alpha})$ and monotone increasing with respect to f_p.

(iii) $f_n, \boldsymbol{g}_n$, $n=1, \cdots, N$ are convex, continuous and differentiable in $\boldsymbol{x}$.

(iv) X is a nonempty, compact and convex set.

(v) The problem (1) is feasible.

Convexity assumptions on $f_n, \boldsymbol{g}_n, X$ ensures that the set W, which is the domain of the function v, is convex and that the function v is convex over W. Furthermore, by taking account of the set $W_a=\{(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})|\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}))=\boldsymbol{\varepsilon}\}$ introduced in the eqn. (4.c), Φ can be transformed over W_a as follows.

$$\begin{aligned}&\Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})), \boldsymbol{\alpha})\\ &=\Phi(\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})), f_p(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})\\ &=\Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha}) \qquad (6)\end{aligned}$$

Therefore, the two-level planning problem (4) is equivalent to

$$\min_{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}} \tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=\Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha}) \tag{7.a}$$

$$\text{subj. to } \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{7.b}$$

$$(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W_a \tag{7.c}$$

Here, from the assumptions (ii), (iii) we can easily prove that the functions $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ and $\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=\Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})$ are convex in $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in W$. By the convexity of $v, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is continuous in $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \in \operatorname{int} W$, since convex function defined on a convex set is continuous in the interior of the set (Rockafellar 1970). So by continuity of v and Φ and the assumption (v), there always exists the optimal solution $(\boldsymbol{\alpha}^{\circ}, \boldsymbol{\varepsilon}^{\circ})$ for the problem (7), if the set of $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ satisfying eqns. (7.b), (7.c) is compact.

A feasible direction method using directional derivatives was first applied by Geoffrion (1970) and Silverman (1972) to separable mathematical programs. Although our problem is not the same that they studied, we can still apply the similar technique as far as calculation is concerned.

To make further theoretical progress, we extend the function $\tilde{\Phi}$ to all of the $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$-space by defining $\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=+\infty$ for $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \notin W_a$. Thus $\tilde{\Phi}$ becomes a proper convex function with effective domain $\operatorname{dom} \tilde{\Phi}=W_a$. Then, the directional derivative $D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\,; \boldsymbol{y}, \boldsymbol{s})^{\dagger}$ of $\Phi(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ has a finite value at an interior point of the set W_a. It can be obtained as a function of $(\boldsymbol{y}, \boldsymbol{s})$ by solving the lower level problem (4.d) $\sim$(4.g) given $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. Hence, we can construct a direction-finding problem as follows to seek a direction $(\boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ})$ making locally the best improvements of the value of $\tilde{\Phi}$ at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ by minimizing the directional derivative $D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\,; \boldsymbol{y}, \boldsymbol{s})$ with respect to the direction $(\boldsymbol{y}, \boldsymbol{s})$.

$$\min_{(\boldsymbol{y}, \boldsymbol{s})} D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\,; \boldsymbol{y}, \boldsymbol{s}) \tag{8.a}$$

$$\text{subj. to } (\boldsymbol{y}, \boldsymbol{s}) \in R(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) \tag{8.b}$$

$$\|(\boldsymbol{y}, \boldsymbol{s})\| \leqq 1 \quad \text{(normalization condition)}, \tag{8.c}$$

where the eqn. (8.b) is needed for local feasibility with respect to a direction $(\boldsymbol{y}, \boldsymbol{s})$ at a current point $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. Thus,

$$R(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}) = \{(\boldsymbol{y}, \boldsymbol{s}) \mid \text{There exists } \tilde{\delta} > 0 \text{ such that}$$

$$\sum_{n=1}^{N} (\boldsymbol{\alpha}_n + \delta \boldsymbol{y}_n) \leqq \boldsymbol{b} \text{ and } (\boldsymbol{\alpha} + \delta \boldsymbol{y}, \boldsymbol{\varepsilon} + \delta \boldsymbol{s}) \in W_a \text{ for all } 0 < \delta \leqq \tilde{\delta}\} \tag{9}$$

If a parameter $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is not optimal, a step is taken in the direction $(\boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ})$, a new parameter is determined and the process is repeated. So we assume the following.

(vi) A point $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is an interior point of the set W_a.

Note that we can take a sufficiently small step in an arbitrary direction because of this assumption.

Optimality of $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ and usable feasibility of $(\boldsymbol{y}, \boldsymbol{s})$

We have the following theorem for optimality test for $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$.

Theorem 2. Let $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ be feasible for the problem (7). If $(\boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ}) = \boldsymbol{0}$ is optimal for the direction-finding problem (8), then $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ solves the problem (7).

(**Proof**) Proved in a similar way to theorem 1 in Section 9.3 of Lasdon (1970).

If the above optimality condition is satisfied at a feasible point $(\boldsymbol{\alpha}^{\circ}, \boldsymbol{\varepsilon}^{\circ})$ for the problem (7), then $\boldsymbol{\alpha}^{\circ}$ is an optimal resource allocation and a parametric optimal solution $\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}^{\circ}, \boldsymbol{\varepsilon}^{\circ})$ to the problem (4.d)$\sim$(4.g) is the best noninferior solution.

As is well known, the above optimality condition of $(\boldsymbol{\alpha}^{\circ}, \boldsymbol{\varepsilon}^{\circ})$ is equivalently stated as follows (See Appendix 2 of Lasdon (1970)).

$$D\tilde{\Phi}(\boldsymbol{\alpha}^{\circ}, \boldsymbol{\varepsilon}^{\circ}\,; \boldsymbol{y}, \boldsymbol{s}) \geqq 0 \qquad \text{for all } (\boldsymbol{y}, \boldsymbol{s})$$

Thus, if optimality test of Theorem 2 is not passed, then a direction $(\boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ})$ is produced such that

$$D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\,; \boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ}) < 0.$$

† A directional derivative of f in $\boldsymbol{s}$ direction with respect to $\boldsymbol{x}$ is defined:

$$Df(\boldsymbol{x}\,; \boldsymbol{s}) = \lim_{B \to 0^{+}} \frac{f(\boldsymbol{x} + \beta \boldsymbol{s}) - f(\boldsymbol{x})}{\beta}.$$

In a case when f is differentiable at $\boldsymbol{x}$, $Df(\boldsymbol{x}\,; \boldsymbol{s}) = \dfrac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{s}$.

Such a direction $(\boldsymbol{y}^{\circ}, \boldsymbol{s}^{\circ})$ is a usable feasible direction at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ (Proved in a similar way to theorem 2 in Section 9.3 of Lasdon (1970)).

Taking into account that the constraints (7.b) is linear in $\boldsymbol{\alpha}$ and that the assumption (vi) enables us to take a sufficiently small step in any direction from $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$, we can write the local feasibility condition for $(\boldsymbol{y}, \boldsymbol{s})$ concretely as follows.

$$R(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=\{(\boldsymbol{y}, \boldsymbol{s}) \mid \sum_{n=1}^{N} y_{ni} \leqq 0,\ i \in B=\{i \mid b_i-\sum_{n=1}^{N} \alpha_{ni}=0\},$$
$$\boldsymbol{s}\ ;\ \text{arbitrary}\} \tag{10}$$

An expression of $D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s})$ and a direction-finding problem

Since $\partial \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}))/\partial \boldsymbol{\varepsilon}=I_{p-1}((p-1)\times(p-1)$ unit matrix) from the eqn. (4.c), the directional derivative of $\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=\Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})$ is given as follows.

$$\begin{aligned} D\tilde{\Phi}&(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s}) \\ &=\frac{\partial \Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \boldsymbol{y}+\frac{\partial \Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})}{\partial \boldsymbol{f}_{\bar{p}}} \boldsymbol{s} \\ &\qquad +\frac{\partial \Phi(\boldsymbol{\varepsilon}, v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}), \boldsymbol{\alpha})}{\partial f_p} Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s}) \end{aligned} \tag{11}$$

where $Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s})$ is a directional derivative of $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ in a direction $(\boldsymbol{y}, \boldsymbol{s})$ at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. So, we must get an expression of $Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s})$ at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$. It is well known that a subgradient of $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ and $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is a gradient vector of a supporting hyperplane at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ for a graph of v, and that the following theorem holds (Lasdon 1970).

Theorem 3. Let $\boldsymbol{x}^{\circ}$ solve the lower level problem (4.d)~(4.g) given $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ and define the Lagrangian function L_p related with the problem (4.d)~(4.g) as

$$L_p(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda})=f_p(\boldsymbol{x})+\boldsymbol{\mu}^T(\boldsymbol{f}_{\bar{p}}(\boldsymbol{x})-\boldsymbol{\varepsilon})+\boldsymbol{\lambda}^T(\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{\alpha}). \tag{12}$$

Then, $-(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})$ is a subgradient of $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$, if and only if $(\boldsymbol{x}^{\circ}, \boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})$ is a suddle point for $L_p(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda})$.

From this theorem we see that a subgradient $(\boldsymbol{\eta}, \boldsymbol{\xi})$ of $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is the optimal Lagrange multiplier vector whose sign is opposite. Denoting a set of subgradients of v at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ by $\partial v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ and a set of optimal Lagrange multiplier vectors by M, we can represent the directional derivative $Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s})$ as

$$\begin{aligned} Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\ ; \boldsymbol{y}, \boldsymbol{s}) &= \max_{(\boldsymbol{\eta}, \boldsymbol{\xi}) \in \partial v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})} (\boldsymbol{\eta}^T \boldsymbol{y}+\boldsymbol{\xi}^T \boldsymbol{s}) \\ &= \max_{(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ}) \in M} (-\boldsymbol{\lambda}^{\circ T} \boldsymbol{y}-\boldsymbol{\mu}^{\circ T} \boldsymbol{s}) \end{aligned} \tag{13}$$

(Refer to Theorem 16 in Appendix 2 of Ref. [6]). The existence of optimal Lagrange multipliers $(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})$ is guaranteed by Slater's constraint qualification (Mangasarian 1969) such that there exists $\boldsymbol{x} \in X$ satisfying $\boldsymbol{f}_{\bar{p}}(\boldsymbol{x})<\boldsymbol{\varepsilon}$ and $\boldsymbol{g}(\boldsymbol{x})<\boldsymbol{\alpha}$ from the assumption (vi).

Let the set X be expressed as

$$X=\{\boldsymbol{x}|\boldsymbol{q}(\boldsymbol{x})\leqq\boldsymbol{0}\}, \tag{14}$$

where each component of $\boldsymbol{q}$ is assumed to be convex and differentiable, and $\boldsymbol{\gamma}^{\circ}$ be an optimal Lagrange multiplier vectors associated with the constraint $\boldsymbol{q}(\boldsymbol{x})\leqq\boldsymbol{0}$. Then, $(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})\in M$ if and only if $\boldsymbol{x}^{\circ}$, $\boldsymbol{\mu}^{\circ}$, $\boldsymbol{\lambda}^{\circ}$ and $\boldsymbol{\gamma}^{\circ}$ satisfy the Kuhn-Tucker conditions for the lower level problem (4. d)$\sim$(4. g) as follows.

$$\frac{\partial \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\boldsymbol{\mu}^{\circ}+\frac{\partial \boldsymbol{g}(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\boldsymbol{\lambda}^{0}+\frac{\partial \boldsymbol{q}(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\boldsymbol{\gamma}^{\circ}=-\frac{\partial f_p(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}} \tag{15.a}$$

$$\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})-\boldsymbol{\varepsilon}\leqq\boldsymbol{0},\ \boldsymbol{\mu}^{\circ T}(\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})-\boldsymbol{\varepsilon})=0,\ \boldsymbol{\mu}^{\circ}\geqq\boldsymbol{0} \tag{15.b}$$

$$\boldsymbol{g}(\boldsymbol{x}^{\circ})-\boldsymbol{\alpha}\leqq\boldsymbol{0},\ \boldsymbol{\lambda}^{\circ T}(\boldsymbol{g}(\boldsymbol{x}^{\circ})-\boldsymbol{\alpha})=0,\ \boldsymbol{\lambda}^{\circ}\geqq\boldsymbol{0} \tag{15.c}$$

$$\boldsymbol{q}(\boldsymbol{x}^{\circ})\leqq\boldsymbol{0},\ \boldsymbol{\gamma}^{\circ T}\boldsymbol{q}(\boldsymbol{x}^{\circ})=0,\ \boldsymbol{\gamma}^{\circ}\geqq\boldsymbol{0} \tag{15.d}$$

Taking into account that $\boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})=\boldsymbol{\varepsilon}$ for fixed $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})\in W_a$ from uniqueness of $\boldsymbol{x}^{\circ}$ (Assumption (i)) and eqn. (7. c) and that $\lambda_{ni}{}^{\circ}=0$ for $i\in\{i|g_{ni}(\boldsymbol{x}^{\circ})<\alpha_{ni}\}$ from the eqn. (15. c) etc., we find that $Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}; \boldsymbol{y}, \boldsymbol{s})$ is the maximum value of the objective function of the following LP.

$$\begin{aligned}
&\max_{(\mu^{\circ}, \{\lambda_{ni}{}^{\circ}\})}\ -\sum_{i\in C_n}\lambda_{ni}{}^{\circ}y_{ni}-\boldsymbol{\mu}^{\circ T}\boldsymbol{s}\\
&\text{subj. to}\ \frac{\partial \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\boldsymbol{\mu}^{\circ}+\sum_{n=1}^{N}\sum_{i\in C_n}\frac{\partial \boldsymbol{g}_{ni}(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\lambda_{ni}{}^{\circ}\\
&\qquad\qquad +\sum_{i\in D}\frac{\partial \boldsymbol{q}_i(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}}\gamma_i{}^{\circ}=-\frac{\partial f_p(\boldsymbol{x}^{\circ})^T}{\partial \boldsymbol{x}}\\
&\boldsymbol{\mu}^{\circ}\geqq\boldsymbol{0}\\
&\lambda_{ni}{}^{\circ}\geqq 0,\ i\in C_n=\{i|g_{ni}(\boldsymbol{x}^{\circ})-\alpha_{ni}=0\},\ n=1,\cdots,N\\
&\gamma_i{}^{\circ}\geqq 0,\ i\in D=\{i|q_i(\boldsymbol{x}^{\circ})=0\}
\end{aligned} \tag{16}$$

Moreover, as the direction $(\boldsymbol{y}, \boldsymbol{s})$ is given, it coincides with the minimum value of the objective function of the dual problem:

$$\begin{aligned}
&\min_{\boldsymbol{z}}\ \frac{\partial f_p(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}}\boldsymbol{z}\\
&\text{subj. to}\ \frac{\partial \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}}\boldsymbol{z}\leqq\boldsymbol{s}\\
&\qquad\frac{\partial \boldsymbol{g}_{ni}(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}}\boldsymbol{z}\leqq y_{ni},\ i\in C_n,\ n=1,\cdots,N\\
&\qquad\frac{\partial \boldsymbol{q}_i(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}}\boldsymbol{z}\leqq 0,\ i\in D
\end{aligned} \tag{17}$$

After all, substituting the problem (17) for $Dv(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}; \boldsymbol{y}, \boldsymbol{s})$ in the eqn. (11), and taking into account that $\partial\Phi/\partial f_p\geqq 0$ from the assumption (i), we can represent the directional derivative $D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}; \boldsymbol{y}, \boldsymbol{s})$ explicitly, and have the direction-finding problem in a final form:

$$\begin{aligned}
&\min_{(\boldsymbol{y},\boldsymbol{s},\boldsymbol{z})} \frac{\partial \Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \boldsymbol{y} + \frac{\partial \Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial \boldsymbol{f}_{\bar{p}}} \boldsymbol{s} \\
&\qquad + \frac{\partial \Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial f_p} \frac{\partial f_p(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}} \boldsymbol{z} \\
&\text{subj. to } \sum_{n=1}^{N} \boldsymbol{y}_{ni} \leqq 0, \ i \in B = \{i \mid b_i - \sum_{n=1}^{N} \alpha_{ni} = 0\} \\
&\qquad \frac{\partial \boldsymbol{f}_{\bar{p}}(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}} \boldsymbol{z} \leqq \boldsymbol{s} \\
&\qquad \frac{\partial \boldsymbol{g}_{ni}(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}} \boldsymbol{z} \leqq \boldsymbol{y}_{ni}, \ i \in C_n = \{i \mid \boldsymbol{g}_{ni}(\boldsymbol{x}^{\circ}) - \alpha_{ni} = 0\} \\
&\qquad\qquad n = 1, \cdots, N \\
&\qquad \frac{\partial \boldsymbol{q}_i(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}} \boldsymbol{z} \leqq 0, \ i \in D = \{i \mid \boldsymbol{q}_i(\boldsymbol{x}^{\circ}) = 0\} \\
&\qquad -1 \leqq \boldsymbol{y}_{ni} \leqq 1, \ i = 1, \cdots, \dim \boldsymbol{\alpha}_n, \ n = 1, \cdots, N \\
&\qquad -1 \leqq s_i \leqq 1, \ i = 1, \cdots, p-1, \ p+1, \cdots, N \\
&\qquad -1 \leqq z_i \leqq 1, \ i = 1, \cdots, \dim \boldsymbol{x}
\end{aligned} \tag{18}$$

The locally best direction $(\boldsymbol{y}, \boldsymbol{s})$ is obtained by solving the above LP. Given this usable feasible direction $(\boldsymbol{y}^k, \boldsymbol{s}^k)$ in the k-th interation, a change of $f_p = v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ corresponding to the displacement $(\boldsymbol{y}^k, \boldsymbol{s}^k)$ from the current point $(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$ can be approximated as

$$s_p{}^k = Dv(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k; \boldsymbol{y}^k, \boldsymbol{s}^k) = \frac{\partial f_p(\boldsymbol{x}^{\circ})}{\partial \boldsymbol{x}} z^k \tag{19}$$

from the eqns. (13), (17). A new point $(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\varepsilon}^{k+1}) = (\boldsymbol{\alpha}^k + \delta^k \boldsymbol{y}^k, \boldsymbol{\varepsilon}^k + \delta^k \boldsymbol{s}^k)$ is generated by solving a one-dimensional convex program in the $(\boldsymbol{f}, \boldsymbol{\alpha})$-space to determine a step size δ^k.

$$\begin{aligned}
&\min_{\delta} \{\Phi(\boldsymbol{\varepsilon}^k + \delta \boldsymbol{s}^k, f_p{}^k + \delta s_p{}^k, \boldsymbol{\alpha}^k + \delta \boldsymbol{y}^k) \mid (\boldsymbol{\alpha}^k + \delta \boldsymbol{y}^k, \boldsymbol{\varepsilon}^k + \delta \boldsymbol{s}^k) \\
&\qquad \text{satisfies (7. b), (7. c)}\},
\end{aligned} \tag{20}$$

where $f_p{}^k \equiv v(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$. Then, it holds that $\Phi(\boldsymbol{\varepsilon}^{k+1}, f_p{}^k + \delta^k s_p{}^k, \boldsymbol{\alpha}^{k+1}) < \Phi(\boldsymbol{\varepsilon}^k, v(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k), \boldsymbol{\alpha}^k)$. However, it does not necessarily hold that

$$\Phi(\boldsymbol{\varepsilon}^{k+1}, v(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\varepsilon}^{k+1}), \boldsymbol{\alpha}^{k+1}) < \Phi(\boldsymbol{\varepsilon}^k, v(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k), \boldsymbol{\alpha}^k). \tag{21}$$

If the above relation (21) does not hold, the step size δ^k is reduced by half until the relation (21) holds.

In a case when $v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is differentiable

The above arguments are simplified at a point $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ at which the function

$v(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is differentiable. In such a case, the subgradient of v at $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})$ is composed with a single element, and there exists a unique optimal Lagrange multiplier vector $(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})$ satisfying the linear system (15). When the $(\boldsymbol{\mu}^{\circ}, \boldsymbol{\lambda}^{\circ})$ is obtained uniquely, the directional derivative of v is simply represented:

$$D\tilde{\Phi}(\boldsymbol{\alpha}, \boldsymbol{\varepsilon}; \boldsymbol{y}, \boldsymbol{s})$$
$$=\left(\frac{\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial\boldsymbol{\alpha}}-\frac{\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial f_p}\boldsymbol{\lambda}^{\circ T}\right)\boldsymbol{y}$$
$$+\left(\frac{\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial\boldsymbol{f}_{\bar{p}}}-\frac{\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^{\circ}), \boldsymbol{\alpha})}{\partial f_p}\boldsymbol{\mu}^{\circ T}\right)\boldsymbol{s} \tag{22}$$

Then, the direction-finding problem is the problem (8) in which the objective function (8.a) is replaced with the eqn. (22) and the constraint with the eqn. (10).

Let us state the proposed computational method in algorithmic form. By the way, since we cannot represent W_a explicitly, in practice, we cannot help ignoring temporarily the constraint $(\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k, \boldsymbol{\varepsilon}^k+\delta\boldsymbol{s}^k)\in W_a$, and determining δ^k by one-dimensional program $\min_{\delta}\Phi(\boldsymbol{\varepsilon}^k+\delta\boldsymbol{s}^k, f_p{}^k+\delta s_p{}^k, \boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k)$ subject to $(\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k, \boldsymbol{\varepsilon}^k+\delta\boldsymbol{s}^k)$ satisfying only (7.b). After that, if the solution to the $\boldsymbol{\varepsilon}$-constraint problem by setting $\boldsymbol{\alpha}=\boldsymbol{\alpha}^k+\delta^k\boldsymbol{y}^k, \boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}^k+\delta^k\boldsymbol{s}^k$ does not exist or does not make the $\boldsymbol{\varepsilon}$-constraint binding, the step size δ^k is reduced until a solution exists and makes it binding.

Taking the matters into account, we can given a computational procedure as below.

Step 1. Choose any initial point $(\boldsymbol{\alpha}^1, \boldsymbol{\varepsilon}^1)\in W_a$ and set $k=1$.

Step 2. Set $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$, solve the $\boldsymbol{\varepsilon}$-constraint problem (2) and obtain a noninferior solution $\boldsymbol{x}^k\equiv\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$ and a noninferior value $\boldsymbol{f}^k=(\boldsymbol{\varepsilon}^k, f_p(\boldsymbol{x}^k))$ corresponding to $\boldsymbol{\alpha}^k$.

Step 3. Calculate $\partial\boldsymbol{f}(\boldsymbol{x}^k)/\partial\boldsymbol{x}$, $\partial g_{ni}(\boldsymbol{x}^k)/\partial\boldsymbol{x}$, $i\in C_n=\{i|g_{ni}(\boldsymbol{x}^k)-\alpha_{ni}=0\}$, $n=1, \cdots, N$, $\partial q_i(\boldsymbol{x}^k)/\partial\boldsymbol{x}$, $i\in D=\{i|q_i(\boldsymbol{x}^k)=0\}$, $\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^k), \boldsymbol{\alpha}^k)/\partial\boldsymbol{f}$ and $\partial\Phi(\boldsymbol{f}(\boldsymbol{x}^k), \boldsymbol{\alpha}^k)/\partial\boldsymbol{\alpha}$, compose a direction-finding problem (18) and find the optimal direction $(\boldsymbol{y}^k, \boldsymbol{s}^k)$ at the point $(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$.

Step 4. If $\|(\boldsymbol{y}^k, \boldsymbol{s}^k)\|\leqq\sigma$, stop, where the tolerance σ is a sufficiently small positive number. Let $(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$ be the optimal resource allocation and the optimal $\boldsymbol{\varepsilon}$-parameter. Then, $\boldsymbol{x}^k$ is the best noninferior solution corresponding to $\boldsymbol{\alpha}^k$. Otherwise go to Step 5.

Step 5. Find δ^k minimizing $\Phi(\boldsymbol{\varepsilon}^k+\delta s^k, f_p{}^k+\delta s_p{}^k, \boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k)$ subject to $(\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k, \boldsymbol{\varepsilon}^k+\delta\boldsymbol{s}^k)$ satisfying (7.b), where $f_p=v(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)$ and $s_p{}^k=(\partial f_p(\boldsymbol{x}^k)/\partial\boldsymbol{x})\boldsymbol{z}^k$.

Step 6. A point $(\boldsymbol{\varepsilon}^k+\delta^k\boldsymbol{s}^k, f_p{}^k+\delta^k s_p{}^k)$ is not necessarily located on the noninferior surface. So set $(\boldsymbol{\alpha}, \boldsymbol{\varepsilon})=(\boldsymbol{\alpha}^k+\delta^k\boldsymbol{y}^k, \boldsymbol{\varepsilon}^k+\delta^k\boldsymbol{s}^k)$ and solve the $\boldsymbol{\varepsilon}$-constraint problem (2).

Step 7. If the problem is feasible and the solution makes the $\boldsymbol{\varepsilon}$-costraint binding, then set $\boldsymbol{\alpha}^{k+1}=\boldsymbol{\alpha}^k+\delta^k\boldsymbol{y}^k$, $\boldsymbol{\varepsilon}^{k+1}=\boldsymbol{\varepsilon}^k+\delta^k\boldsymbol{s}^k$. But at this time take a step-size in plactice slightly shorter than δ^k in the step-size choosing problem (Step 5).† Then obtain a noninferior solution $\boldsymbol{x}^{k+1}\equiv\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\varepsilon}^{k+1})$ and a noninferior value

† Our direction-finding problem is based on the condition that $(\boldsymbol{\alpha}^k, \boldsymbol{\varepsilon}^k)\in\text{int}\, W_a$.

$f^{k+1}=(\varepsilon^{k+1}, v(\alpha^{k+1}, \varepsilon^{k+1}))$ and go to Step 8. Otherwise go to Step 9.

Step 8. If $\Phi(f^{k+1}, \alpha^{k+1}) \geqq \Phi(f^k, \alpha^k)$, then go to Step 9. Otherwise set $k:=k+1$ and go back to Step 2.

Step 9. Reset $\delta^k:=\delta^k/2$ and go back to Step 6.

Our algorithm is a feasible direction method in the (α, ε)-space. Convergence of feasible direction algorithms for nonlinear programming was studied in detail by D.M. Topkis and A.F. Veinott (1967) Through their theorems, our feasible direction method does converge to an optimal parameter $(\alpha^\circ, \varepsilon^\circ)$.

3. Hierarchical Decentralized Systems

The hierarchical multi-objective system becomes a hierarchical decentralized system when an objective function f_n and a constraint $g_n \leqq 0$ of each local system contain only its own decision variable vector x_n and a resource allocation α_n given from the central system. Accordingly, the each local system is separated with respect to x each other. Then, the two-level resource allocation problem (1) becomes as follows.

$$\min_{\alpha} \Phi(f(x^\circ(\alpha)), \alpha) \tag{23.a}$$

$$\text{subj. to } \sum_{n=1}^{N} \alpha_n \leqq b \quad (b\,;\text{ total resources}) \tag{23.b}$$

$$f_n(x_n{}^\circ(\alpha_n)) = \min_{x_n} f_n(x_n) \tag{23.c}$$

$$\text{subj. to } g_n(x_n) \leqq \alpha_n \tag{23.d}$$

$$x_n \in X_n \tag{23.e}$$

$$n=1, \cdots, N$$

where $\alpha=(\alpha_1^T, \cdots, \alpha_N^T)^T$, $x^\circ(\alpha)=(x_1{}^\circ(\alpha_1)^T, \cdots, x_N^\circ(\alpha_N)^T)^T$. The central decision-maker allocates the resource vector b to the local systems such that the central objective Φ is optimized. The lower level of the problem (23) consists of a set of N separated optimization subproblems, each of which is usual parametric optimization problem with a scalar objective function. They are mutually interacting only through the restriction on the total amount of resources (23.b). A vector $x_n{}^\circ(\alpha_n)$ means a usual parametric optimal solution. This is a special case of the two-level multi-objective system. It is noted that the eqn. (23.a) achieves minimization with respect to only α in contrast to the eqn. (1.a).

We try to solve the problem (23) iteratively by choosing a feasible allocation testing it for optimality and improving it if it is not optimal. Then, given the allocation α_n, the problem of the local system n (we shall term it Local Problem $P_n(\alpha_n)$) is represented as follows.

$$P_n(\alpha_n)\,;\ \min_{x_n} f_n(x_n) \tag{24.a}$$

$$\text{subj. to } g_n(x_n) \leqq \alpha_n \tag{24.b}$$

$$\boldsymbol{x}_n \in X_n \tag{24.c}$$

We can apply the feasible direction method mentioned in the last section in order to solve the problem (23). But it becomes much simpler.

Denote the minimal value of the objective f_n for the problem (24) by $w_n(\boldsymbol{\alpha}_n)$. Then, the two-level decentralized planning problem (23) is equivalent to the following problem.

$$\min_{\{\boldsymbol{\alpha}_n\}} \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) \tag{25.a}$$

$$\text{subj. to } \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{25.b}$$

$$\boldsymbol{\alpha}_n \in V_n, \; n=1, \cdots, N, \tag{25.c}$$

where $\boldsymbol{w}(\boldsymbol{\alpha})=(\boldsymbol{w}_1(\boldsymbol{\alpha}), \cdots, w_n(\boldsymbol{\alpha}_n))^T$. The set V_n is defined as

$$V_n=\{\boldsymbol{\alpha}_n | \text{ There exists } \boldsymbol{x}_n \in X_n \text{ satisfying } \boldsymbol{g}_n(\boldsymbol{x}_n) \leqq \boldsymbol{\alpha}_n\},$$

which is necessary for feasibility of $P_n(\boldsymbol{\alpha}_n)$.

To make further theoretical progress, we assume the following.

(vii) $\Phi(\boldsymbol{f}, \boldsymbol{\alpha})$ is convex and differentiable with respect to $(\boldsymbol{f}, \boldsymbol{\alpha})$ and monotone increasing with respect to $\boldsymbol{f}$.

(viii) $f_n(\boldsymbol{x}_n)$, $\boldsymbol{g}_n(\boldsymbol{x}_n)$, $n=1, \cdots, N$ are convex and differentiable in $\boldsymbol{x}_n$.

(ix) X_n, $n=1, \cdots, N$, is a nonempty, compact and convex set.

(x) The problem (23) is feasible.

(xi) A point $\boldsymbol{\alpha}_n$ is an interior point of the set V_n.

The convexity assumption on f_n, $\boldsymbol{g}_n$ and X_n ensures that V_n is convex and that w_n is convex with respect to $\boldsymbol{\alpha}_n$ over V_n, so the problem (25) is a convex program.

A direction-finding problem can be constructed in analogous manner to the last section to seek a usable feasible direction $\boldsymbol{y}^{\circ}$ making improvement of Φ than the current resource allocation $\boldsymbol{\alpha}$, by minimizing the directional derivative $D\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}; \boldsymbol{y})$ with respect to $\boldsymbol{y}$. The directional derivative $D\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}; \boldsymbol{y})$ in a direction $\boldsymbol{y}=(\boldsymbol{y}_1^T, \cdots, \boldsymbol{y}_N^T)^T$ with respect to $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1^T, \cdots, \boldsymbol{\alpha}_N^T)^T$ is given as

$$D\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}; \boldsymbol{y}) = \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial f_n} Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{y}_n) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{y}_n \right], \tag{26}$$

where $Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{y}_n)$ is a directional derivative of $w_n(\boldsymbol{\alpha}_n)$. Therefore, the direction-finding problem is represented as

$$\begin{aligned} &\min_{\{\boldsymbol{y}_n\}} \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial f_n} Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{y}_n) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{y}_n \right] \\ &\text{subj. to } \sum_{n=1}^{N} y_{ni} \leqq 0, \; i \in B=\{i | b_i - \sum_{n=1}^{N} \alpha_{ni}=0\} \\ &\qquad -1 \leqq y_{ni} \leqq 1, \; i=1, \cdots, \dim \boldsymbol{\alpha}_n, \; n=1, \cdots, N \end{aligned} \tag{27}$$

Then the following properties hold.

(A) Let $\boldsymbol{\alpha}$ be feasible for the problem (25). If $\boldsymbol{y}=\boldsymbol{0}$ is optimal for the direction-finding problem, then $\boldsymbol{\alpha}$ solves the problem (25).

(B) Let $\partial w_n(\boldsymbol{\alpha}_n)$ be a set of subgradients of w_n at $\boldsymbol{\alpha}_n$. Then the directional derivative $Dw_n(\boldsymbol{\alpha}_n ; \boldsymbol{y}_n)$ is expressed as

$$Dw_n(\boldsymbol{\alpha}_n ; \boldsymbol{y}_n)=\max_{\boldsymbol{\xi}_n\in\partial w_n(\boldsymbol{\alpha}_n)} \boldsymbol{\xi}_n^T \boldsymbol{y}_n=\max_{\boldsymbol{\lambda}_n^{\circ}\in\Lambda_n}(-\boldsymbol{\lambda}_n^{\circ T}\boldsymbol{y}_n),$$

where Λ_n is a set of all optimal Lagrange multiplier vectors $\boldsymbol{\lambda}_n^{\circ}$ for the constraint (24. b) in $P_n(\boldsymbol{\alpha}_n)$.

Let each set X_n be expressed as

$$X_n=\{\boldsymbol{x}_n|\boldsymbol{q}_n(\boldsymbol{x}_n)\leqq\boldsymbol{0}\}.$$

We assume that each component of $\boldsymbol{q}_n$ is convex and differentiable. Let $\boldsymbol{x}_n^{\circ}$ is an optimal solution to $P_n(\boldsymbol{\alpha}_n)$ and $\boldsymbol{\gamma}_n^{\circ}$ is the Lagrange multiplier vector associated with $\boldsymbol{q}_n\leqq\boldsymbol{0}$. Then $\boldsymbol{\lambda}_n^{\circ}\in\Lambda_n$ if and only if $\boldsymbol{x}_n^{\circ}$, $\boldsymbol{\lambda}_n^{\circ}$ and some $\boldsymbol{\gamma}_n^{\circ}$ satisfy the Kuhn-Tucker conditions for $P_n(\boldsymbol{\alpha}_n)$. Accordingly, under the assumption that Φ be monotone increasing function with respect to $\boldsymbol{f}$, the direction-finding problem (27) becomes as follows in a final form.

$$\min_{\boldsymbol{z},\,\boldsymbol{y}} \sum_{n=1}^{N}\left[\frac{\partial\Phi(f(\boldsymbol{x}_n^{\circ}),\boldsymbol{\alpha})}{\partial f_n}\frac{\partial f_n(\boldsymbol{x}_n^{\circ})}{\partial \boldsymbol{x}_n}\boldsymbol{z}_n+\frac{\partial\Phi(\boldsymbol{f}(\boldsymbol{x}_n^{\circ}),\boldsymbol{\alpha})}{\partial\boldsymbol{\alpha}_n}\boldsymbol{y}_n\right]$$

$$\text{subj. to } \sum_{n=1}^{N} y_{ni}\leqq 0,\ i\in B=\{i|b_i-\sum_{n=1}^{N}\alpha_{ni}=0\}$$

$$\frac{\partial \boldsymbol{g}_{ni}(\boldsymbol{x}^{\circ})}{\partial\boldsymbol{x}_n}\boldsymbol{z}_n\leqq\boldsymbol{y}_{ni},\ i\in C_n=\{i|\boldsymbol{g}_{ni}(\boldsymbol{x}_n^{\circ})-\alpha_{ni}=0\}\,,\ n=1,\cdots,N \qquad (28)$$

$$\frac{\partial \boldsymbol{q}_{ni}(\boldsymbol{x}_n^{\circ})}{\partial\boldsymbol{x}_n}\boldsymbol{z}_n\leqq 0,\ i\in D_n=\{i|\boldsymbol{q}_{ni}(\boldsymbol{x}_n^{\circ})=0\}\,,\ n=1,\cdots,N$$

$$-1\leqq y_{ni}\leqq 1,\ i=1,\cdots,\dim\boldsymbol{\alpha}_n,\ n=1,\cdots,N$$

$$-1\leqq \boldsymbol{z}_{ni}\leqq 1,\ i=1,\cdots,\dim\boldsymbol{x}_n,\ n=1,\cdots,N$$

We obtain the locally best $\boldsymbol{y}$ by solving the above LP. Given a usable feasible direction $\boldsymbol{y}^k$, a new point $\boldsymbol{\alpha}^{k+1}$ is generated by solving a one-dimensional convex problem to determine a step size:

$$\min_{\delta}\{\Phi(\boldsymbol{w}(\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k),\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k)|\boldsymbol{\alpha}^k+\delta\boldsymbol{y}^k \text{ satisfies (25b, c)}\},$$

where k is an iteration number.

4. Numerical Example

Let us consider the following numerical example of the hierarchical multi-objective system.

$$\min_{\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha})} \Phi=(f_1(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))+10)^2+20f_2(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))+(\alpha_1-10)^2+(\alpha_2-5)^2$$

$$\text{subj. to}\quad \alpha_1+\alpha_2\leqq 15$$

$$\alpha_1,\ \alpha_2\geqq 0$$

$$\begin{pmatrix} f_1(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \\ f_2(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \end{pmatrix} = \min_{\boldsymbol{x}} \begin{pmatrix} -2x_1+x_2+15 \\ \frac{1}{20}(x_1^2+x_2^2-50x_2+425) \end{pmatrix} \tag{29}$$

$$\text{subj. to}\quad x_1\leqq\alpha_1$$

$$x_2\leqq\alpha_2$$

$$x_1,\ x_2\geqq 0,$$

where $f_1(\boldsymbol{x})$, x_1 and α_1 are an objective function, a decision variable and resource allocation for the local system 1, respectively and $f_2(\boldsymbol{x})$, $\boldsymbol{x}_2$ and α_2 are for local system 2, respectively. The true values for the above problem are known as follows.

Optimal resource allocation $\boldsymbol{\alpha}^{\circ}=(\alpha_1^{\circ}, \alpha_2^{\circ})=(10.00,\ 5.00)$

Best noninferior solution $\hat{\boldsymbol{x}}^{\circ}(\boldsymbol{\alpha}^{\circ})=(\hat{\boldsymbol{x}}_1^{\circ}(\boldsymbol{\alpha}^{\circ}), \hat{\boldsymbol{x}}_2^{\circ}(\boldsymbol{\alpha}^{\circ}))=(10.00,\ 5.00)$

Best noninferior local objective value

$$\boldsymbol{f}(\hat{\boldsymbol{x}}^{\circ}(\boldsymbol{\alpha}^{\circ}))=(f_1(\hat{\boldsymbol{x}}^{\circ}(\boldsymbol{\alpha}^{\circ})),\ f_2(\hat{\boldsymbol{x}}^{\circ}(\boldsymbol{\alpha}^{\circ})))=(0.00,\ 15.00)$$

Optimal central objective value $\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^{\circ}(\boldsymbol{\alpha}^{\circ})), \boldsymbol{\alpha}^{\circ})=400.0$

Those values can be easily obtained by analytical and geometrical investigations.

We apply the feasible direction method to solve the above problem. So, transforming the problem (29) into the problem (4), we have the following.

The upper level problem $\min_{\boldsymbol{\alpha},\varepsilon_1} \tilde{\Phi}(\boldsymbol{\alpha}, \varepsilon_1)=(\varepsilon_1+10)^2+20f_2(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \varepsilon_1))$

$$+(\alpha_1-10)^2+(\alpha_2-5)^2$$

$$\text{subj. to}\quad \alpha_1+\alpha_2\leqq 15,\ \alpha_1,\ \alpha_2\geqq 0$$

$$(\alpha_1,\ \alpha_2,\ \varepsilon_1)\in W_a$$

The lower level problem $f_2(\boldsymbol{x}^{\circ}(\boldsymbol{\alpha}, \varepsilon_1))=\min_{\boldsymbol{x}} \frac{1}{20}(x_1^2+x_2^2-50x_2+425)$

($\boldsymbol{\varepsilon}$-constraint problem)

$$\text{subj. to}\quad -2x_1+x_2+15\leqq\varepsilon_1$$

$$x_1\leqq\alpha_1,\ \ x_2\leqq\alpha_2$$

$$x_1,\ x_2\geqq 0$$

At first, an initial trial point $(\alpha_1^1, \alpha_2^1, \varepsilon_1^1)$ is set at (0.00, 0.00, 15.00) by the upper level, to which the lower level responds by obtaining the optimal solution

$\boldsymbol{x}^{o1}=(x_1^{o1}, x_2^{o1})=(0.00, 0.00)$ with some nonlinear programming code. Then the lower level objective values $(f_1(\boldsymbol{x}^{o1}), f_2(\boldsymbol{x}^{o1}))$ are obtained as (45.0, 21.3).

Next, based on the above values, the derivatives $\partial \boldsymbol{f}(\boldsymbol{x}^{o1})/\partial \boldsymbol{x}$ and so on, we construct the direction-finding problem (18) for the central, and find a direction $(y_1^1, y_2^1, s_1^1)=(1.00, 1.00, -1.00)$ improving the current central objective value $\tilde{\Phi}(\alpha_1^1, \alpha_2^1, \varepsilon_1^1)=7050$. Then, the objective value of the direction-finding problem is -130.0. A change of f_2, $s_2^1=-2.5$, caused by a displacement (y_1^1, y_2^1, s_1^1), is calculated by the eqn. (19). By the linear search (20) in the direction $(y_1^1, y_2^1, s_1^1, s_2^1)$, a new point $(\alpha_1^2, \alpha_2^2, \varepsilon_1^2)=(7.50, 7.50, 7.50)$ is generated.

After the parameters $(\alpha_1, \alpha_2, \varepsilon_1)$ were updated five times in the similar manner, the following values were obtained.

$$\varepsilon_1^5=-0.05 \quad \boldsymbol{\alpha}^5=(\alpha_1^5, \alpha_2^5)=(10.02, 4.99)$$

$$\boldsymbol{x}^{o5}=(x_1^{o5}, x_2^{o5})=(10.02, 4.99)$$

$$\boldsymbol{f}(\boldsymbol{x}^{o5})=(f_1(\boldsymbol{x}^{o5}), f_2(\boldsymbol{x}^{o5}))=(-0.05, 15.05)$$

$$\Phi(\boldsymbol{f}(\boldsymbol{x}^{o5}), \boldsymbol{\alpha}^5)=400.0$$

They nearly coincide with the true values. And each of the solutions $\boldsymbol{x}^{ok}$ for the lower level problem in iteration process could be confirmed to be one of the noninferior solutions $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ for the multi-objective problem.

5. Conclusion

We have studied the hierarchical planning problem for multi-objective systems in which there exist plural semi-autonomous local systems subordinated to a central system that determines optimal resource allocations. In this problem, the coordinating center allocates scarce resources so as to optimize the central objective function and the local ones optimize their own local objective with use of the given resources.

At first, we formulated two-level planning problems for multi-objective systems in which the lower level composes a vector minimization problem and proposed a computational method such as a feasible direction method by use of directional derivatives. Next, we specialized this to the decentralized systems where the local systems were separated with respect to local decision variables.

The proposed methods can be extended to more general cases such as an objective function $\boldsymbol{f}$ and constraint $\boldsymbol{g}$ are arbitrary functions of $(\boldsymbol{x}, \boldsymbol{\alpha})$ and/or the resource constraint $\boldsymbol{G}(\boldsymbol{\alpha})$ is nonlinear.

The systems as newly formulated here exist in many real and important organizations.

REFERENCES

DA CUNHA, N.C. and POLAK, E. (1967): Constrained minimization under vector-valued criteria in finite-dimensional spaces, J. Math. Anal. & Appl., **19**, 103–124.

GEOFFRION, A.M. (1970): Primal resource-directive approaches for optimizing nonlinear decomposable systems, Opns. Res., **18**, 375–403.

GEOFFRION, A.M. and HOGAN, W.W. (1972): Coordination of two-level organizations with multiple objectives, in Balakrishnan A.V. (ed.), Techniques of optimization, Academic Press, 455–468.

HAIMES, Y.Y., HALL, W.A. and FREEDMAN, H.T. (1975): Multi-objective optimization in water resources systems, Elsevier Scientific Publishing Company, 141–169.

KORNAI, J. (1965): Two-level planning, Econometrica, **33**, 141–169.

LASDON, L.S. (1970): Optimization theory for large systems, Macmillan, Ch. 9 and Appendix 2.

LIN, J.G. (1977): Proper inequality constraints and maximization of index vectors, J. Opt. Th. & Appl., **21**, 505–521.

MANGASARIAN, O.L. (1969): Nonlinear programming, McGraw-Hill.

PAYNE, H.J., POLAK, E., COLLINS, D.C. and MEISEL, W.S. (1975): An algorithm for bicriteria optimization based on the sensitivity function, IEEE Trans. Autom. Contr., **AC–20**, 546–548.

ROCKAFELLAR, R.T. (1970): Convex analysis, Princeton University Press, Sect. 23.

SHIMIZU, K. (1975): Optimization theory for multiple-objective systems (vector objective function), J. of the Society of Instrument and Control Engineering, **14**, 94–103 (in Japanese).

SHIMIZU, K. (1976): Optimization algorithms for multiple objective programs and decentralized hierarchical systems, Proc. of IFAC–LSSTA Italy (Udine), 489–497.

SHIMIZU, K. (1977): Hierarchical decentralized systems, Systems and Control, **20**, 58–60 (in Japanese).

SHIMIZU, K., AIYOSHI, E. and UENO, T. (1977): Decentralized optimization systems and their application to a class of transportation problem, Trans. of the Society of Instrument and Control Engineers, **13**, 29–36 (in Japanese).

SILVERMANN, G.J. (1972): Primal decomposition of mathematical programs by resource allocation, Opens. Res., **20**, 58–74.

TOPKIS, D.M. and VEINOTT, Jr A.F. (1969): On the convergence of some feasible direction algorithms for nonlinear programming, J. SIAM Control, **5**, 268–279.

YU, P.L. (1974): Cone convexity, cone extreme points and nondominated solutions in decision problems with multi-objective, J. Opt. Th. & Appl., **14**, 319–377.