# A Multi-Objective Optimization Method Combining Generalized Data Envelopment Analysis and Genetic Algorithms

**Y.B. Yun*, H. Nakayama**, T. Tanino* & M. Arakawa*****

* **Department of Electronics and Information Systems, Osaka University, Suita, Osaka 565-0871, Japan**
yeboon@eie.eng.osaka-u.ac.jp tanino@eie.eng.osaka-u.ac.jp

** **Department of Applied Mathematics, Konan University, Kobe 658-8501, Japan**
nakayama@koana-u.ac.jp

*** **Department of Reliability Information System Engineering, Kagawa University, Takamatsu, Kagawa 760-8526, Japan**
arakawa@eng.kagawa-u.ac.jp

## ABSTRACT

The present paper describes a method using generalized data envelopment analysis (GDEA) and genetic algorithms (GA) for generating efficient frontiers in multi-objective optimization problems. The purpose of GDEA is to measure the relative efficiency of decision making units and reflects the various preferences of decision makers. In addition, GA is used for directly finding Pareto optimal solutions of multi-objective optimization problems. In this paper, we suggest to combine GDEA and GA and search for Pareto optimal solutions. It will be shown that the proposed method overcomes shortcomings of existing methods and yields desirable efficient frontiers even in the problems with non-convex constraints as well as convex constraints, through several numerical examples.

## 1. INTRODUCTION

Many decision making problems can be formulated as multi-objective optimization problems (MOP):

$$\text{(MOP)}\quad \min_{\boldsymbol{x}} \; \boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))^T$$
$$\text{s.t.}\quad \boldsymbol{x} \in X = \{\boldsymbol{x} \mid g_j(\boldsymbol{x}) \leqq 0,\ j = 1, \cdots, l\},$$

where $\boldsymbol{x} = (x_1, \cdots, x_n)^T$. There does not necessarily exist the optimal solution that minimizes all objective functions $f_i(\boldsymbol{x})$ $(i = 1, \cdots, m)$ in (MOP), and then the concept of Pareto optimal solution (or efficient solution) is introduced [7]. Moreover, the value of objective function corresponding to Pareto optimal solution in (MOP) is called Pareto optimal value. Usually, there exist a number of Pareto optimal solutions, which are considered as candidates of solution to the decision making problem [6]. It is difficult for decision makers to choose one from those as a final solution. Up to now, several methods using interactive multi-objective optimization methods have been suggested so as to find out a final solution to the decision making problem. However, the function form of various criteria functions cannot be given explicitly in many practical problems of engineering design. Under this circumstance, the value of criteria functions for each value of design variables is obtained through some analyses such as structural analysis, thermodynamical analysis and fluid mechanical analysis. In general, it requires considerable time to carry out these analyses, and then it is hard to apply interactive multi-objective optimization methods to practical engineering design problems. Even in these cases, decision making can be easily performed by figuring out the set of Pareto optimal values on the objective space, i.e. the efficient frontier in decision making with two or three objective functions. To this end, several methods have been devised for figuring out the efficient frontier. So, we give brief explanations about methods using genetic algorithms (GA) under the circumstance that the value of objective function has been determined by analyses such as structural analysis, and in addition, it is hard to use differential information. In order to generate the efficient frontier by GA, there have been developed several methods, for example, ranking methods [4], [5] and Tamaki *et al.*'s method [9], and so on. However, it is difficult to generate smooth efficient frontier by ranking methods. Moreover, non-dominated individuals obtained at intermediate generation are not necessarily exact Pareto optimal solutions.

On the other hand, a method using data envelopment analysis (DEA) with GA was proposed by Arakawa *et al.* [1]. DEA, which was originally suggested by Charnes-Cooper-Rhodes, is a method to measure the relative efficiency of decision making units (DMUs)

0-7803-5731-0/99/$10.00 ©1999 IEEE

performing similar tasks in a production system that consumes multiple inputs to produce multiple outputs. So far, representative model examples are CCR model [3], BCC model [2] and FDH model [10]. These models are classified by how to determine the production possibility set; a convex cone, a convex hull and a free disposable hull (FDH) of observed data set is considered to be production possibility set, respectively. Almost of all non-dominated individuals obtained at intermediate generation by the method using DEA become Pareto optimal solutions. However, it produces only convex curves (or surfaces) of efficient frontier, because CCR model or BCC model is used for DEA there.

In this paper, we propose a method using GA with a generalized data envelopment analysis [11], by which CCR efficiency, BCC efficiency and FDH efficiency can be measured in a unified way. Additionally, we try to apply the proposed method to figure out the efficient frontier. Finally, through several numerical examples, it will be shown that shortcomings of existing methods are overcome by the proposed method.

## 2. GENETIC ALGORITHMS FOR MULTI-OBJECTIVE OPTIMIZATION

To begin with, we give a brief explanation on the ranking method [4]. Consider an individual $\boldsymbol{x}^o$ at a generation which is dominated by $n$ individuals in the current population, then his/her rank is given by $(1+n)$. All non-dominated individuals are assigned rank 1 and remain at the next generation. In Fig. 1, each numbers in parentheses represents the rank of each individual and the curve represents the exact efficient frontier. The ranking method based on the relation of domination among individuals has a merit to be computationally simple. However, the ranking method has a shortcoming to need to assess until a large number of generation, since non-dominated individuals in the current generation are often kept alive long, even though they are not Pareto optimal solutions in the final generation. Moreover, it is difficult to generate a smooth efficient frontier by the stated ranking method.

On the other hand, a method was suggested to divide the population into a number of small groups and try to maintain good features in each small group [8]. Subsequently, Tamaki *et al.* proposed the method to combine the above two approaches and to obtain good efficient frontiers for multi-objective optimization problems with non-convex objective function, too [9]. However, in many cases, non-dominated individuals generated at the current generation are far from the exact efficient frontier.

Arakawa *et al.* suggested a method using DEA [1] in order to overcome the shortcomings of the methods stated above. In DEA, the efficiency $\theta$ of an individual $\boldsymbol{x}^o$ $(o=1,\cdots,p)$ is given by solving the following linear programming problem:

$$\begin{array}{ll} \underset{\theta,\boldsymbol{\lambda}}{\text{Min}} & \theta \\ \text{s.t.} & \left[\boldsymbol{f}(\boldsymbol{x}^1),\cdots,\boldsymbol{f}(\boldsymbol{x}^p)\right]\boldsymbol{\lambda} - \theta \boldsymbol{f}(\boldsymbol{x}^o) \leqq 0, \\ & \boldsymbol{\lambda} \geqq 0,\ \ \boldsymbol{\lambda} \in \mathbb{R}^p. \end{array}$$

The degree of efficiency $\theta$ represents the degree how far DEA efficient frontier is, and when $\theta$ is equal to one, it means that an individual $\boldsymbol{x}^o$ is located on DEA efficient frontier. (A efficient frontier generated by CCR method, or BCC method is said *DEA efficient frontier.*) Selection in GA is performed by taking the degree of efficiency $\theta$ for fitness. In other words, this method is to investigate the relation of domination among individuals with respect to the shaded region (See Fig. 2). In Fig. 2, the solid curve represents the exact efficient frontier and the dotted line represents DEA efficient frontier at a generation. As the figure shows, individuals like $C$, $D$ and $G$ are removed fast, and then efficient frontier can be obtained efficiently.

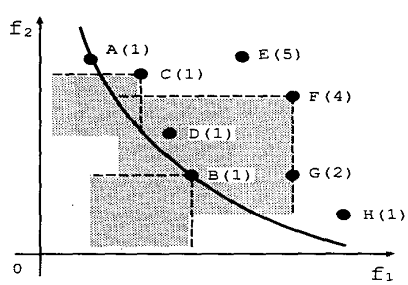



Fig. 1. Ranking method

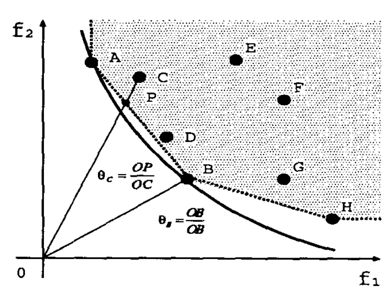



Fig. 2. GA with DEA method

That is to say, when the efficient frontier is convex[1], non-Pareto solutions can be removed at a young generation. However, when the efficient frontier is non-convex, the sunken part of it can not be generated by DEA method.

## 3. THE PROPOSED METHOD

Whereas DEA efficient can provides only convex, the efficient frontier in FDH model [10] is non-convex, since free disposable hull [2] of the given data set becomes non-convex. Subsequently, Yun-Nakayama-Tanino [11] suggested a generalized data envelopment (GDEA) which includes the existing DEA methods by varying the value of a parameter.

In GDEA, the efficiency of DMU$o$ ($o = 1, \cdots, p$) with inputs $x_{io}$ ($i = 1, \cdots m$) and outputs $y_{ko}$ ($k = 1, \cdots n$) is judged by solving the following problem:

$$
\begin{aligned}
\text{(GDEA)} \quad & \max_{\Delta,\, \mu,\, \nu} \quad \Delta \\
\text{s.t.} \quad & \Delta \leqq \tilde{d}_j + \alpha \left( \sum_{k=1}^{n} \mu_k (y_{ko} - y_{kj}) + \sum_{i=1}^{m} \nu_i (-x_{io} + x_{ij}) \right), \; j = 1, \cdots, p, \\
& \sum_{k=1}^{n} \mu_k + \sum_{i=1}^{m} \nu_i = 1, \\
& \mu_k \geqq \varepsilon, \;\; k = 1, \cdots, n, \\
& \nu_i \geqq \varepsilon, \;\; i = 1, \cdots, m,
\end{aligned}
$$

where $\tilde{d}_j := \max_{\substack{k=1,\cdots,n \\ i=1,\cdots,m}} \{\nu_k (y_{ko} - y_{kj}),\, \mu_i (-x_{io} + x_{ij})\}$, $\alpha$ is a constant and $\varepsilon$ is sufficiently small number.

For a given $\alpha$, DMU$o$ is defined to be *$\alpha$-efficient* if the optimal value of the problem (GDEA) is equal to zero. We have the following properties between DEA efficiencies and $\alpha$-efficiency [11].

**Theorem 1.** *DMUo is BCC efficient if and only if DMUo is $\alpha$-efficient for sufficiently large $\alpha > 0$.*

**Theorem 2.** *In the problem (GDEA), add the equation $\sum_{k=1}^{p} \mu_k y_{ko} = \sum_{i=1}^{m} \nu_i x_{io}$ to constraints. Then DMUo is CCR efficient if and only if DMUo is $\alpha$-efficient for sufficiently large $\alpha > 0$.*

**Theorem 3.** *DMUo is FDH efficient if and only if DMUo is $\alpha$-efficient for sufficiently small $\alpha > 0$.*

[1]Let $E$ be a efficient frontier set in $\mathbb{R}^n$ and $\mathbb{R}^n_+$ be a positive orthant in objective space. Then we call *the efficient frontier convex* if $(E + \mathbb{R}^n_+)$ is convex set. Otherwise, efficient frontier is non-convex.

[2]The free disposable hull is the set consisting of points that perform less output with the same amount of input as observed points, and/or those that perform more input with the same amount of output.

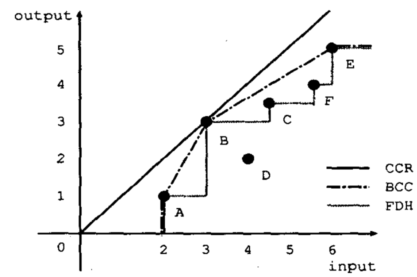



Fig. 3. Efficient frontiers with variation of $\alpha$

Consequently, several kinds of DEA efficient frontiers are conducted by changing the value of a parameter $\alpha$ in the problem (GDEA). For our discussion, we employ the example presented in Fig. 3 consisting of six DMUs consuming a single input to produce a single output. The figure indicates the efficient frontiers with varying the value of $\alpha$, and DMUs on the line are $\alpha$-efficient. It can be seen in Fig. 3 that as $\alpha$ becomes sufficiently large, the efficient frontier changes from the form of steps to the straight line.

In this paper, we propose a method combining GDEA and GA to overcome the shortcomings of the ranking methods and the DEA method. In applying GA to problems with constraints, we introduce an augmented objective function using penalty functions. Here, an augmented objective function to $f_i$ ($i = 1, \cdots, m$) in (MOP) is given by

$$
F_i(\boldsymbol{x}) = f_i(\boldsymbol{x}) + \sum_{j=1}^{l} p_j \times [P(g_j(\boldsymbol{x}))]^a ,
$$

where $p_j$ is a penalty coefficient, $a$ is a penalty exponent and $P(y) = \max\{y,\, 0\}$. To the end, the initial problem (MOP) can be converted into a problem to minimize the augmented objective function ($F_1(\boldsymbol{x})$, $\cdots, F_m(\boldsymbol{x})$). Here, we need to prepare the data set in order to evaluate the degree of GDEA efficiency[3] of a individual $\boldsymbol{x}^o$ in the current populations. Let inputs and outputs in GDEA be substituted by the value of $F_i(\boldsymbol{x}^o)$ and the unit, respectively. Then the problem (GDEA) reduces to the following problem (P).

$$
\begin{aligned}
\text{(P)} \quad & \max_{\Delta,\, \nu} \quad \Delta \\
\text{s.t.} \quad & \Delta \leqq \tilde{d}_j - \alpha \sum_{i=1}^{m} \nu_i (F_i(\boldsymbol{x}^o) - F_i(\boldsymbol{x}^j)), \\
& \qquad j = 1, \cdots, p, \\
& \sum_{i=1}^{m} \nu_i = 1, \\
& \nu_i \geqq \varepsilon, \;\; i = 1, \cdots, m,
\end{aligned}
$$

[3]We call the GDEA efficiency instead of the $\alpha$-efficiency

where $\tilde{d}_j = \max_{i=1,\cdots,m} \{\nu_i(-F_i(\boldsymbol{x}^o) + F_i(\boldsymbol{x}^j))\}$. $\varepsilon$ is a sufficiently small number. $\alpha$ is a monotonically decreasing function with respect to the number of generations.

The degree of the GDEA efficiency of an individual $\boldsymbol{x}^o$ in the current population is given by the optimal value $\Delta^*$ in the problem (P), and is considered as the fitness in GA. Therefore, the selection of an individual is determined by the degree of GDEA efficiency, i.e. if $\Delta^*$ equals to zero, the individual will remain at the next generation.

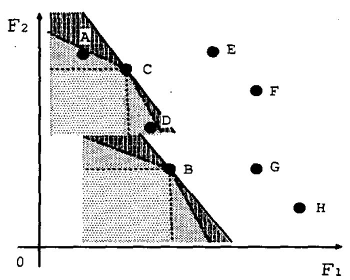



Fig. 4. Geometric interpretation of $\alpha$ in (P)

With making the best use of the stated properties of GDEA, it is possible to keep the merits, and at the same time, to overcome the shortcomings of existing methods. Namely, taking a large $\alpha$ can yield that the non-dominated individuals which are not Pareto optimal solutions are removed fast, and taking a small $\alpha$ can generate non-convex efficient frontiers.

## 4. EXAMPLES ; TWO-OBJECTIVE OPTIMIZATION PROBLEM

We consider the following examples with two objective functions.

Example 1

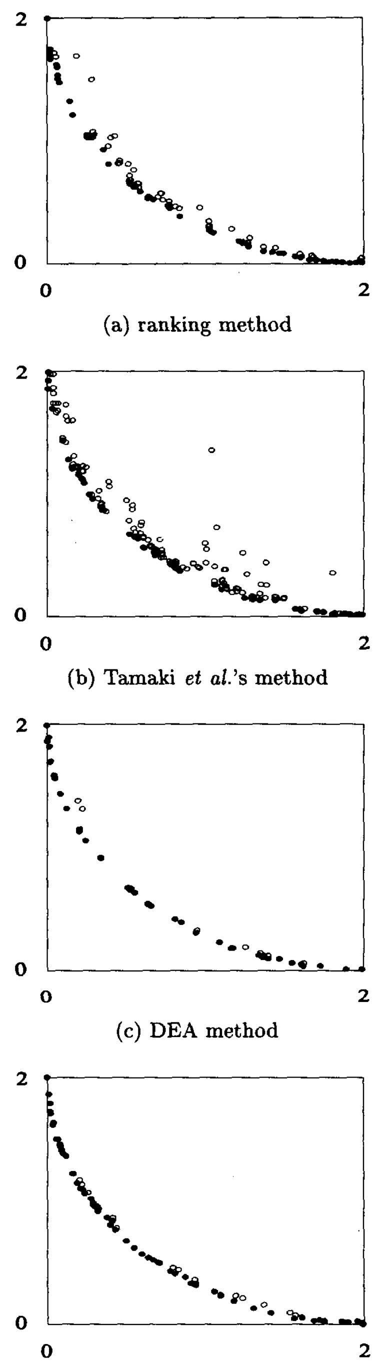
$$\begin{array}{ll} \min_{\boldsymbol{x}} & (f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})) = (x_1,\ x_2) \\ \text{s.t.} & (x_1-2)^2 + (x_2-2)^2 - 4 \leqq 0, \\ & x_1 \geqq 0,\ \ x_2 \geqq 0. \end{array}$$

Example 2 [9]

$$\begin{array}{ll} \min_{\boldsymbol{x}} & (f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})) = (-2x_1 + x_2,\ x_1) \\ \text{s.t.} & (x_1-1)^3 + x_2 \leqq 0, \\ & x_1 \geqq 0,\ \ x_2 \geqq 0. \end{array}$$

Example 3

$$\begin{array}{ll} \min_{\boldsymbol{x}} & (f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})) = (x_1,\ x_2) \\ \text{s.t.} & x_1^3 - 3x_1 - x_2 \leqq 0, \\ & x_1 \geqq -1,\ \ x_2 \leqq 2. \end{array}$$



(a) ranking method

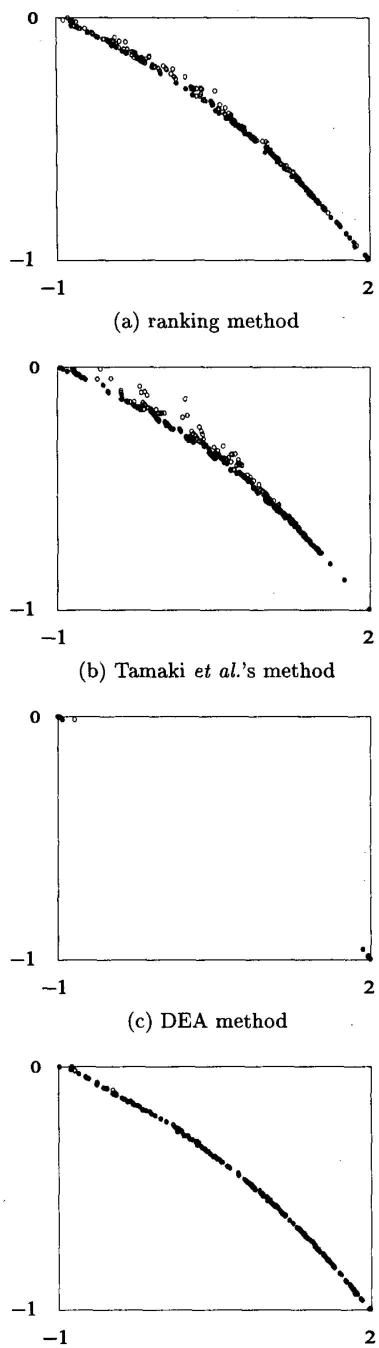
(b) Tamaki *et al.*'s method

(c) DEA method

(d) GDEA method

Fig. 5. Efficient frontiers for example-1

The efficient frontier in Example 1 is convex, and non-convex in both Example 2 and Example 3. In order to show the effectiveness of GDEA method, we compare the results by (a) ranking method, (b) Tamaki *et al.*'s



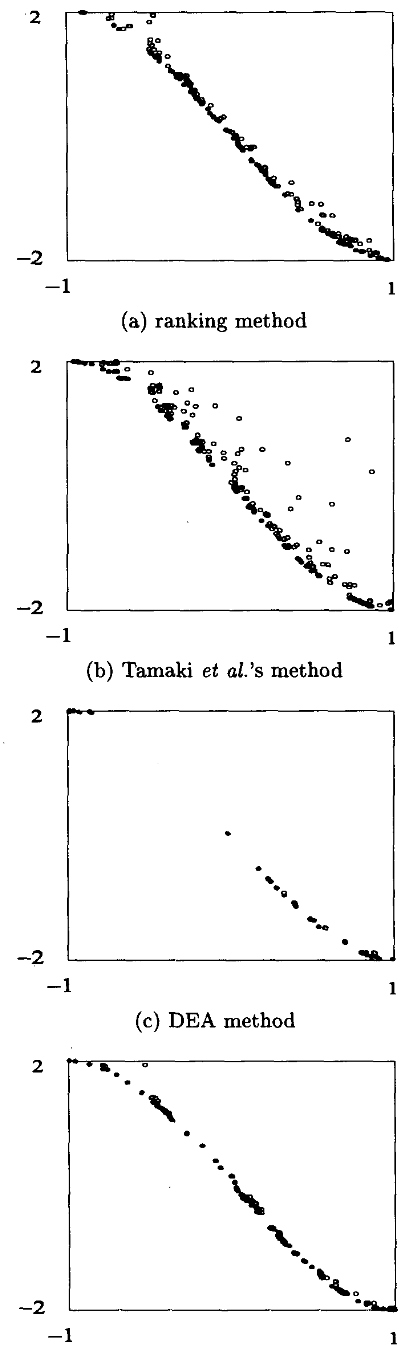

(a) ranking method

(b) Tamaki *et al.*'s method

(c) DEA method

(d) GDEA method

Fig. 6. Efficient frontiers for example-2




(a) ranking method

(b) Tamaki *et al.*'s method

(c) DEA method

(d) GDEA method

Fig. 7. Efficient frontiers for example-3

method, (c) DEA method, (d) GDEA method. The numbers of generation are 15 (Example 1), 20 (Example 2), 30 (Example 3), and $\alpha$ in the problem (P) is given by $10 \times \exp(-0.2 \times$ numbers of generation).

We accumulate non-dominated individuals obtained at each generation, and finally examine the Pareto optimality among those individuals in objective space. The results are shown in Fig. 5-7. The horizontal axis and the vertical axis indicate the values of objective function $f_1$ and $f_2$, respectively. The symbols $\bullet$ represents a Pareto optimal solution, and $\circ$ does a non-Pareto optimal solution.

(a) Ranking method

We obtained relatively many Pareto optimal solutions. However, there are also many non-Pareto optimal solutions among non-dominated individuals at each generation. Moreover, it is usually difficult to generate smooth efficient frontiers as shown in (a) of Fig. 5-7.

(b) Tamaki *et al.*'s method

A large number of Pareto optimal solutions are obtained by this method. Efficient frontiers generated by this method are smoother than the ones by ranking method. However, it is seen in (b) of Fig. 5-7 that many of non-dominated individuals at each generation are not finally Pareto optimal solutions.

(c) DEA method

When the efficient frontier is convex in (a) of Fig. 5, it is smooth, nevertheless the obtained Pareto optimal solutions are fewer than by the above methods. On the other hand, for non-convex efficient frontiers in (b) of Fig. 6 and Fig. 7, the sunken part of it can not be generated by this method. Therefore, DEA method cannot be applied to multi-objective optimization problems with non-convex functions.

(d) GDEA method

In (d) of Fig. 5-7, the largest number of Pareto optimal solutions are obtained among the stated methods. Moreover, efficient frontiers generated by the proposed method are smooth, even they are non-convex. In addition, it is seen that almost of all non-dominated individuals at each generation become the final Pareto optimal solutions.

In particular, it should be noted in the ranking methods and Tamaki *et al.*'s method that non-dominated individuals obtained at intermediate generation are often not Pareto optimal solutions. In practical problems, we don't know when to stop the computation in advance. Usually, the computation is terminated at a relatively early generation due to the time limitation. It is important, therefore, that non-dominated individuals at intermediate generations are finally Pareto optimal solutions. GDEA method has a desirable performance from this point of view.

## 5. CONCLUSION

In this paper, we have proposed a multi-objective optimization method combining GDEA and GA for generating efficient frontiers of multi-objective optimization problems. The proposed method overcomes the shortcomings of existing methods: it provides a lot of Pareto optimal solutions in a small number of generation, and can be applied to cases of multi-objective optimization problems with non-convex functions as well as convex. However it requires a certain amount of time to solve the problem (P) in order to evaluate the GDEA efficiency. Since the time required for analyses such as structural analysis, thermodynamical analysis and fluid mechanical analysis in engineering design problems is extremely long, the computational time for solving the problem (P) is not so serious.

## 6. REFERENCE


[1] M. Arakawa, I. Hagiwara, H. Nakayama and H. Yamakawa, "Multiobjective Optimization Using Adaptive Range Genetic Algorithms with Data Envelopment Analysis", *American Institute of Aeronautics and Astronautics*, 1998, pp. 1–9.

[2] M. Banker, A. Charnes, W.W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, Vol. 30, 1984, pp. 1078–1092.

[3] A. Charnes, W.W. Cooper, E. Rhodes, "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, Vol. 2, 1978, pp. 429–444.

[4] C.M. Fonseca and P.J. Fleming, "Genetic Algorithms for Multi-objective Optimization:Formulation, Discussion and Generalization", *Proceeding of the Fifth International Conference on Genetic Algorithms*, 1993, pp. 416–426.

[5] D.E. Goldberg, it Genetic Algorithms in Search, Optimization and Machine Learning, Massachusetts : Addison-Wesley, Inc., 1989.

[6] T.C. Koopmans, "Analysis of Production as an Efficient Combination of Activities", *Analysis of Production and Allocation* : John Wiley, Inc., 1951, pp. 33–97.

[7] Y. Sawaragi, H. Nakayama and T. Tanino, *Theory of Multiobjective Optimization* : Academic Press, Inc., 1985.

[8] J.D. Schaffer, "Multiple Objective Optimization with Vector Evaluated Genetic Algorithms", *Proceeding of the First International Conference on Genetic Algorithms*, 1985, pp. 93–100.

[9] H. Tamaki, "Multi-Criteria Optimization by Genetic Algorithms", *Genetic Algorithms 2* : Sangyotosho, 1995, pp. 71–81.

[10] H. Tulkens, "On FDH efficiency: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit", *Journal of Productivity Analysis*, Vol. 4, 1993, pp. 183–210.

[11] Y.B. Yun, H. Nakayama, T. Tanino, "On Efficiency of Data Envelopment Analysis", *to appear in Proceedings of the 14th International Conference on Multiple Criteria Decision Making*, Charlottesville, Virginia : Springer-Verlag.