'94 KACC (1994. 10. 17 ~ 20)

An analysis and modification of a unified phase I-phase II semi-infinite constrained optimization algorithm

Yang, Hyun Suk

Dept. of Electronic Engineering, Hongik University
72-1 Sangsoodong Mapoku Seoul, 121 Korea

<ABSTRACT>

In this paper, we analize the effect of a steering parameter used in a unified phase I-phase II semi-infinite constrained optimization algorithm and present a new algorithm based on the facts that when the point x is far away from the feasible region where all the constraints are satisfied, reaching to the feasible region is more important than minimizing the cost function and that when the point x is near the region, it is more efficient to try to reach the feasible region and to minimize the cost function concurrently. Also, the angle between the search direction vector and the gradient of the cost function is considered when the steering parameter value is computed. Even though changing the steering parameter does not change the rate of convergence of the algorithm, we show through some examples that the proposed algorithm performs better than the other algorithms.

1. Introduction

The difficulty of an optimization problem is very much a function of the constraints. Optimization problems with max-functional inequality constraints rank close to the top in terms of difficulty. Recent work on phase I-phase II methods of centers, methods of centers based on barrier functions, and barrier function methods for semi-infinite minimax problems constructs a reasonably promising optimization algorithm ([1-7,9-11]).

A phase I-phase II method of centers was proposed by Polak [3] and modified by Polak and He [5]. Since the modified version of the algorithm combines the original one which is consisted of two parts, it is called a unified phase I-phase II semi-infinite optimization algorithm with a steering parameter. It is important to choose a good steering parameter since it may result in saving as much as 90% of the computing time for some problems as shown in [5]. Unfortunately, no method is given to find a good steering parameter.

In this paper, we will analize the effect of a steering parameter and give a guideline on how to choose a steering parameter value. This paper is consisted that in section 2, basic optimization theory is mentioned and in section 3, the effect of the steering parameter is analized and a method of choosing a good steering parameter value is introduced. In section 4, the proposed algorithm is tested with some examples and compared with a unified phase I-phase II algorithm. This paper is concluded in section 5.

2. Preliminary

We consider a semi-infinite constrained optimization problem of the form:

$$\min_{x\in R^n} \{ f^0(x) \mid \psi^j(x) \le 0,\ j\in M \} \tag{1}$$

where $f^0(\cdot)$ is continuously differentiable and $M = \{1, \cdots, m\}$. For all $j\in M$ the constraints are defined by $\psi^j(x) = \max_{y\in Y_j} \phi^j(x, y)$ where $\phi^j(\cdot,\cdot)$ is continuously differentiable and $Y_j = [a_j, b_j] \subset R$ is a compact set. The reason why it is called semi-infinite constrained optimization problem is that even though it has m constraints, each constraint $\psi^j(x)$ is the maximum of infinite number of functions. The constraints can be written in a compact form, $\psi(x) \le 0$, where $\psi(x)$ is defined by $\psi(x) = \max_{j\in M} \psi^j(x)$. Polak [3] proposed a phase I-phase II semi- infinite optimization algorithm that solves the problem (1), which is consisted of two parts: one that when a point x_i is not in the feasible region where all the constraints are satisfied, it tries to find a point in the feasible reagion and one that when a point x_i is in the feasible region, it tries to find an optimal solution within the region. One drawback of the method is that when a point x_i is outside the feasible region, it is possible that a next point x_{i+1} is farther from an optimal solution since the search directon vector is computed only using the constraint functions. Polak and He [5] published an improved version of the method by combining two parts together. Before the modified algorithm is mentioned, let us define a parametrized function, $F_{x_i}(\cdot)$, at a point x_i by

$$F_{x_i}(x) = \max \{ f^0(x) - f^0(x_i) - \gamma\psi_+(x_i),\ \psi^i(x) - \psi_+(x_i),\ i\in M \} \tag{2}$$

where $\psi_+(x) = \max\{0, \psi(x)\}$ and a steering parameter

$\gamma > 0$. Since the parametrized function $F_x(\cdot)$ is not differentiable, computing an optimality function value and a search direction vector is very difficult. To simplify computations, let us define a linear first order convex approximation of $F_x(\cdot)$ by

$$\hat{F}_x(z) = \max \{ \hat{f}_x^0(z) - f^0(x) + \gamma\psi_+(x), \; \hat{\psi}_x^j(z) - \psi_+(x), \; j \in M \} \quad (3)$$

where $\hat{\psi}_x^j(z)$ is the linear first order convex approximation of $\psi^j(z)$ at a point x and is defined by

$$\hat{\psi}_x^j(z) = \max_{y_j \in Y_j} (\phi^j(x, y_j) + \langle \nabla_x \psi^j(x, y_j), (z - x) \rangle + (1/2) \|z - x\|^2). \quad (4)$$

Then, the optimality function value $\theta(x)$ and the search direction vector $h(x)$ can be defined by

$$\theta(x) = \min_{h \in R^n} \hat{F}_x(x+h) \quad (5)$$

$$h(x) = \arg \min_{h \in R^n} \hat{F}_x(x+h) \quad (6)$$

Polak and He [5] proposed a unified phase I-phase II semi-infinite constrained optimization algorithm with a steering parameter as follows:

Algorithm I [5]

Parameters: $\gamma > 0$, $\alpha, \beta \in (0, 1)$.

Data: $x_0 \in R^n$.

Step 0: Set $i = 0$.

Step 1: Compute the optimality function value $\theta_i = \theta(x_i)$ and the search direction vector $h_i = h(x_i)$ using (8) and (9), respectively.

Step 2: Compute the step size λ_i by

$$\lambda_i = \max \{ \beta^k \mid k \in N, \; F_{x_i}(x_i + \beta^k h_i) \le \beta^k \alpha \theta_i \} \quad (7)$$

Step 3: Let $x_{i+1} = x_i + \lambda_i h_i$, replace i with $i + 1$, and go to STEP 1. ■

Following results can be found in a reference [5].

Assumption 2.1: Suppose that $\phi^j(\cdot, \cdot)$, $j \in M_0 = \{0\} \cup M$, satisfy that

(a) there exist $0 < c < 1 < C < \infty$ such that for any $x, z \in R^n$, and $y \in Y_j$, $j \in M$, it is satisfied that

$$c\|z\|^2 \le \langle z, (\partial^2 \phi^j(x, y) / \partial x^2) \rangle \le C\|z\|^2. \quad (8)$$

(b) a set $\{ x \mid \psi(x) < 0 \}$ is not empty. ■

Theorem 2.2: Suppose that assumption 2.1 is satisfied, $\{x_i\}_{i=0}^{\infty}$ is a sequence of vectors generated by Algorithm I, and x^* is the only solution of the problem (1). Then,

(a) a sequence, $\{x_i\}_{i=0}^{\infty}$, converges to x^*.

(b) for any $\varepsilon \in (0, 1)$, there exists $\rho > 0$ such that

(i) if $x_i \in B^0(x^*, \rho)$ for all $i \ge 0$, then there exists $\delta_1(\varepsilon) \in (0, 1)$, such that $\psi(x_{i+1}) \le \delta_1(\varepsilon)\psi(x_i)$.

(ii) if there exists an $i_0 \ge 0$ such that $\psi(x_i) > 0$ and $x_i \in B^0(x^*, \rho)$ for all $i \ge i_0$, then it is satisfied that $\psi^0(x_{i+1}) - \psi^0(x_i) \le \delta_2(\varepsilon)[\psi^0(x_i) - \psi^0(x^*)]$ for some $\delta_2(\varepsilon) \in (0, 1)$.

Here, $B^0(x, \rho)$ is an open ball with a radius of ρ at a center x. ■

The first part of Theorem 2.2 implies that algorithm I finds an optimal solution and the second part deals with the rate of convergence of the algorithm.

3. The effect of a steering parameter and its modification

In this section, the effect of a steering parameter is analized and a rule on choosing a steering parameter is given. Based on this analysis, a modified unified phase I-phase II optimization algorithm is proposed.

Let us consider four cases based on a steering parameter value and the location of a vector x_i.

Case 1: Suppose that the value of a steering parameter is large and at some point x_i, all the constraints are satisfied. Then, since a point x_i is in a feasible region and as a result $\psi_+(x_i) = 0$, the steering parameter has no effect on the algorithm.

Case 2: Suppose that the value of a steering parameter is large and at a point x_i there exists a $j \in M$ such that $\psi^j(x_i) > 0$, i.e. some of the constraints are not satisfied. In this case we have that $\psi_+(x_i) = \psi(x_i) > 0$. Since γ is large, the value of $\gamma\psi_+(x_i)$ is large. Also, since $F_{x_i}(x_i) = 0$, there exists an $\varepsilon > 0$ such that for all $x \in B^0(x_i, \varepsilon)$, $F_{x_i}(x) = \max_{j \in M} \{ \psi^j(x) - \psi_+(x_i) \}$. It implies that the search direction vector computed by the equation (6) does not depend on the cost function.

Case 3: Suppose that the value of a steering parameter is close to zero and at a point x_i all the constraints are satisfied. Then, by the same reason as case 1, γ does not have any effect on generating a point x_{i+1}.

Case 4: Suppose that the value of a steering parameter is close to zero and at a point x_i, there exists a $j \in M$ such that $\psi^j(x_i) > 0$. If the value of $\psi_+(x_i)$ is very large, we have the same result as case 2 and if it is not large, since the value of $F_{x_i}(x)$ in the neighbor of a vector x_i depends on both the cost function and the constraint functions, the search direction vector h_i is affected by those functions. More specifically speaking, when a point x_i is far from the feasible region, it is likely that Algorithm I generates a search direction that leads the constraints towards the feasible region and when a point x_i is near the feasible region, it computes a search direction that not only leads the constraints to the feasible region but also leads, if possible, to an optimal solution.

It is noted that when the value of $\gamma\psi_+(x)$ is large, Algorithm I is not different from the original phase I-phase II optimization algorithm, which implies that it is possible to obtain a faster convergence if a proper value of a steering parameter is chosen. Unfortunately, there exist no rules on choosing a proper value of a steering parameter since it is torally problem-dependent. Here, let us propose a general rule that improves the performance of the algorithm likely, but not always. First, when a vector x_i is far from the feasible region, it is more efficient to pick a large value of γ since all the constraints have to be satisfied, firstly. One drawback of choosing a large value of γ is that a cost function value may increase a lot. More analysis on this case will be followed later. Second, when a vector x_i is near the feasible region, it is better to choose a small value of a steering parameter so that the cost function is considered when a search direction vector h_i is calculated. Third, when a vector x_i is in the feasible region, Algorithm I is independent of the steering parameter γ since we have that $\psi_+(x_i) = 0$. Lastly, when a value of $\psi(x_{i+1})/\psi(x_i)$ or $\psi(x_{i+1})/\psi(x_0)$ becomes small, it implies that a sequence x_{i+1} is close to the feasible region compared to the point x_i or the initial point, respectively. Then, it may be efficient to decrease the value of γ so that the cost function has more effect on determining the search direction vector h_i.

So far, we have analized effects of the steering parameter based on the location of a vector x_i. One aspect we have to consider is the angle between the search direction vector and the gradient of the cost function at a point x_i because if the angle is grater than 90^o, the search direction vector forces the cost function value to increase. In this case, a steering parameter value should be decreased so that the cost function is considered more in calculating a search direction vector. And if the angle is less than 90^o, then the search direction leads a next point x_{i+1} not only to the feasible region but also to an optimal solution. Based on these analysis, we propose that the steering paramater value is computed by $\gamma_i = \Gamma_i e^{c\cos\theta}$ where Γ_i is obtained by the analysis given in the previous paragraph, c is chosen by a user, and θ is the angle between the gradient of the cost function and the search direction vector. Let us define a parametrized function with a varing steering parameter as follow.

$$F_{x_i,\gamma_i}(x) = \max\{\psi^0(x) - \psi^0(x_i) - \gamma_i\psi_+(x_i),\ \psi^j(x) - \psi_+(x_i),\ j \in M\} \quad (9)$$

A modified unified phase I-phase II optimization algorithm is presented below.

Algorithm II:

Parameters: $0 < \Gamma_{min} \le \Gamma_0 \le \Gamma_{max} < \infty$, $\alpha, \beta \in (0, 1)$, $\rho, \delta \in (0, 0.5)$, and $c > 0$.

Data: $x_0 \in R^n$.

Step 0: Set $i = 0$.

Step 1: Compute the optimality function value $\theta_i = \theta(x_i)$ and the search direction vector $h_i = h(x_i)$.

Step 2: Compute the step size λ_i by

$$\lambda_i = \max\{\beta^k \mid F_{x_i,\gamma_i}(x_i + \beta^k h_i) \le \beta^k \alpha \theta_i\}. \quad (10)$$

Step 3: Let $x_{i+1} = x_i + \lambda_i h_i$.

Step 4: If $\psi_+(x_{i+1}) = 0$ or both $\psi_+(x_{i+1})/\psi_+(x_0) < \delta$ and $\psi_+(x_0) \ne 0$ are satisfied, set $\Gamma_{i+1} = \Gamma_i$,

if $\psi_+(x_{i+1})/\psi(x_i) < \rho$, set $\Gamma_{i+1} = \max\{\Gamma_{min}, \Gamma_i - 0.1 * \min(\Gamma_0, \Gamma_i)\}$, and

otherwise, set $\Gamma_{i+1} = \min\{\Gamma_{max}, \Gamma_i + \Gamma_0 * 0.1\}$.

Step 5: Replace i with $i+1$ and go to step 1. ■

The parameter δ in step 4 is the measure of how close a vector x_{i+1} is to the feasible region compared to the initial point. For example, when $\delta = 0.01$, it implies that if the value of $\psi_+(x_{i+1})$ is within 1% of the initial value $\psi_+(x_0)$, we can assume that the vector x_{i+1} is very close to the feasible region. Then, since the steering parameter has very little effect on the performance of the algorithm, the steering parameter value is not changed. In this algorithm, the amount of the change of a steering parameter value is arbitrarily picked as a 10% of the previous value.

We have following results on convergence of the algorithm.

Lemma 3.1: Suppose that assumption 2.1 is satisfied and a sequence of vectors $\{x_i\}_{i=0}^{\infty}$ is generated by Algorithm II. Then, for all $i \ge 0$, we have that

$$\psi^0(x_{i+1}) \le \psi^0(x_i) + \gamma_i\psi_+(x_i), \qquad \psi_+(x_{i+1}) \le \psi_+(x_i). \quad (11)$$

■

Theorem 3.2: Suppose that $\{x_i\}_{i=0}^{\infty}$ is a sequence of vectors generated by Algorithm II and x^* is an accumulation point of the sequence. Then, we have that $\theta(x^*) = 0$, i.e., when $x_i \to x^*$, $\theta(x_i) \to 0$. ■

4. Examples

In this section, Algorithm I and Algorithm II are compared. Both algorithms are written in C-language and were run in SUN SPAC II workstation.

Example 1: (Rozen-Suzuki problem [8]) Consider a following optimization problem.

$$\begin{aligned} \min\ & x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 - 7x_4, \\ \text{with}\ & 2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 \le 0, \\ & x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 \le 0, \\ & x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \le 0. \end{aligned} \quad (12)$$

In this problem, all functions are convex and its minimum occurs at $(0, 1, 2, -1)$ and the minimum value of the cost function is -44. A feasible initial point, $(0, 0, 0, 0)$, and an infeasible initial point, $(2, 4, 8, 1)$, are used. Variables used in Algorithm II are $\Gamma_0 = 2.0$, $\Gamma_{min} = 0.3$, $\Gamma_{max} = 4.0$, $c = 1.0$, $\alpha = 0.7$, $\beta = 0.6$, $\delta = 0.01$, and $\rho = 0.05$. For the feasible initial point, since $\psi_+(x_0) = 0$, Algorithm I and Algorithm II generate the same sequence of vectors. The cost function values versus iterations are shown in figure 1. Figures 2 and 3 are the value of the cost function and the constraints versus iterations, respectively, with the infeasible initial point. The computing time taken is 0.16 seconds for Algorithm I and 0.11 seconds for Algorithm II.

Example 2: (Wong problem [8]) Consider a following optimization problem.

$$\begin{aligned} \min\ & (x-10)^2 + 5(x_2-12)^2 + x_3^4 + 3(x_4-11)^2 \\ & + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7, \\ \text{with}\ & 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 - 127 \le 0, \\ & 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 - 282 \le 0, \\ & 23x_1 + x_2^2 + 6x_6^2 - 8x_7 - 196 \le 0, \\ & 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \le 0. \end{aligned} \quad (13)$$

We note that all the constraint functions are convex but the cost function is nonconvex. The optimal value is 680.63 at (2.33, 1.95, −0.48, 4.37, −0.62, 1.04, 1.59). An infeasible initial point of (3, 3, 0, 5, 1, 3, 0) is used. The variables used in Algorithm II are $\Gamma_0 = 2.0$, $\Gamma_{min} = 0.3$, $\Gamma_{max} = 4.0$, and $c = 2.0$. Figures 4 and 5 are the cost and the constraint versus iterations, respectively. The computing times are 0.47 seconds for Algorithm I and 0.39 seconds for Algorithm II.

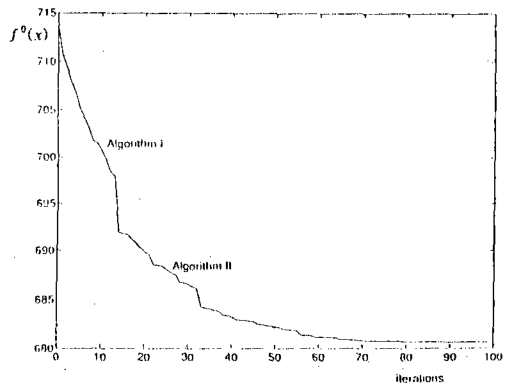

Fig. 1. Cost versus iteration plot with a feasible initial point.

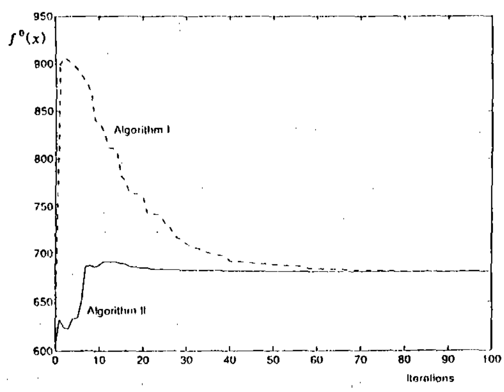

Fig. 2. Cost versus iteration plot with an infeasible initial point.

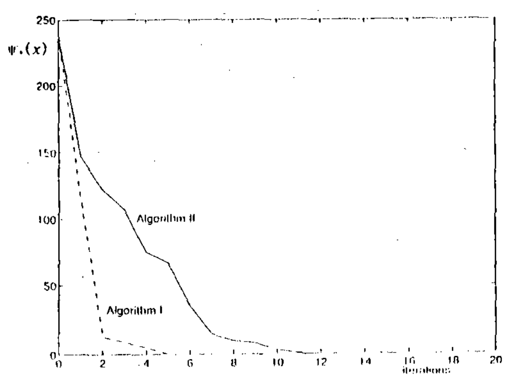

Fig. 3. Constraint versus iteration plot with an infeasible initial point.

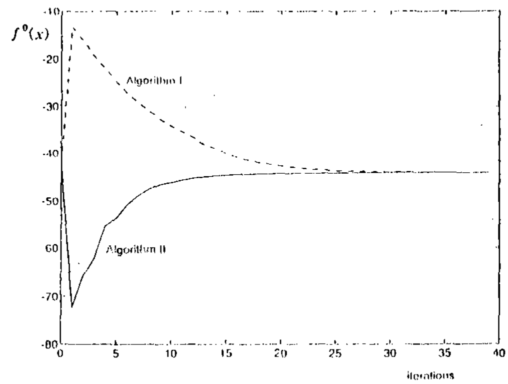

Fig. 4. Cost versus iteration plot with an infeasible initial point.

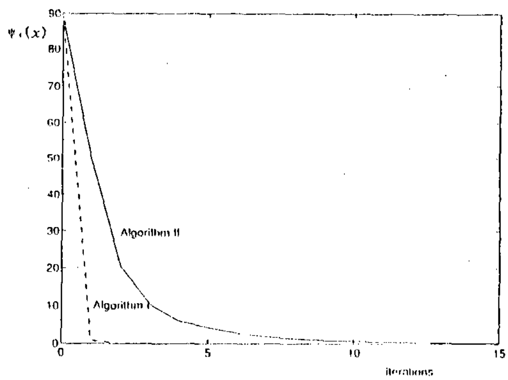

Fig. 5. Constraint versus iteration plot with an infeasible initial point.

Example 3: Consider a following semi-infinite optimization problem.

$$\min \; x_1^2 + x_2^2 + x_3^2$$
$$\text{with} \; x_1 + x_2 e^{x_3 t} + e^{2t} - 2\sin(4t) \le 0, \; t \in [0, 1] \quad (14)$$

The minimum value for the problem is 5.3928 at (−0.1848, −1.3404, 1.8873). The infeasible initial point (1.5, 1.5, 1.5) were used. Twenty-one equispaced points were used to discretize the interval [0, 1]. Cost and Constraints versus iterations are shown in

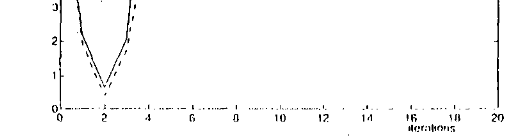

Fig. 6. Cost versus iteration plot with an infeasible initial point.

figure 6 and 7, respectively. The computing time taken is 0.11 seconds for Algorithm I and 0.09 seconds for Algorithm II.

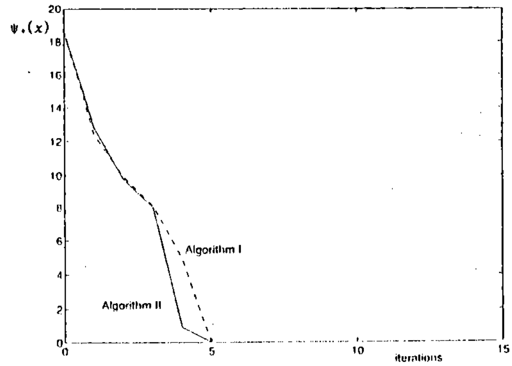

Fig. 7. Constraint versus iteration plot with an infeasible initial point.

5. Conclusion

In this paper, the effects of a steering parameter of a unified phase I-phase II semi-infinite optimization algorithm are analized and a new algorithm is proposed. It is more efficient that when a point x_i is far from the feasible region, a large value of a steering parameter is chosen and when it is near the feasible region, a small value is chosen so that the cost function is considered in calculating the search direction vector. Also, it is known that the angle between the gradient of the cost function and the search direction vector has a great role in determining the steering parameter value. Based on these analysis, a modified unified phase I-phase II optimization algorithm is proposed. Even though we can not claim that the new algorithm is better than the previous algorithms for all problems, it is shown through some examples that the proposed one is better than others. When the user has more information on the problem, it is possible to obtain better performance by choosing proper parameter values.

<참고문헌>

[1] Conn, A.R., "Constrained Optimization Using a Nondifferentiable Penalty Function", SIAM Journal on Numerical Analysis, Vol.10, No.4, pp.760-784, 1973.

[2] Klessig, R., and Polak, E., "An Adaptive Precision Gradient Methods for Optimal Control", SIAM Joutnal on Control, Vol.10, pp.80-93, 1973.

[3] Polak, E., "On the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design", SIAM Review, Vol.29, No.1, pp.21-91, 1987.

[4] Polak, E., and He, L., "Rate-Preserving Discretization Strategies for Semi-Infinite Programming and Optimal Control", UCB/ERL Memo M89/112, Electric Research Laboratory, University of California, Berkeley, 1989.

[5] Polak, E., and He, L., "Unified Steerable Phase I-Phase II Method of Feasible Directions for Semi-Infinite Optimization", JOTA Vol.69, No.1, pp.83-107, April 1991.

[6] Polak, E., Trahan, R., and Mayne, D.Q., "Combined Phase I-Phase II Methods of Feasible Directions", Mathematical Programming, Vol.17, pp.61-73, 1979.

[7] Pironneau, O., and Polak, E., "On the Rate of Convergenc of Certain Methods of Centers", Mathematical Programming, Vol.2, pp.230-258, 1972.

[8] Tanaka, Y., Fukushima, M., and Ibaraki, T.A., "A Comparative Study of Several Semi- Infinite Nonlinear Programming Algorithms", European Journal of Operational Research, Vol.36, pp.92-100, 1988.

[9] Topkis, D.M. and Veinott, A F.Jr., "On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming", SIAM Journal on Control, Vol.5, pp.268-279, 1967.

[10] Zoutendijk, G., "Methods of Feasible Directions", Elsevier, Amsterdam, Holland, 1960.

[11] Zukhovitskii, S.I., Polyak, R.A., and Primak, M.E., "An Algorithm for the Solution of Convex Problems", Doklady Akademii Nauk USSR, Vol.153, pp.991-1000, 1963.