Research of Multi-level Optimization Algorithm in Large-Scale system

Qian Kun Dong Xinmin Qu Zhihong

(The Air Force Engineering University, Dept of Automatic Control Engineering, Xi'an, China.710038 email: qiankun_0306@163.com)

Abstract--**A three-level hierarchical optimization algorithm in separable non-convex large-scale system is presented in this paper, which converts the original problem into a separable multi-objective optimization problem. And then the non-inferior frontier is convexified and the global optimal solution of the original problem is selected from the set of non-inferior solutions. A theory based on this algorithm is established and its convergence is proved. Digital simulation demonstrates that this algorithm brings about good control results.**

Key words**—multi-level optimization; large-scale system; global optimal solution; multi-objective optimization**

1. INTRODUCTION

MUlti-level optimization theory has been developing rapidly in the past three decades. Its important contribution lies in that it has solved the problems of incidence and coupling among subsystems as it is successfully applied in the field of large-scale system. Substantially, multi-level optimization algorithm is in mathematics defined as how to decompose and solve large-scale optimization control problem provided with a special configuration. The definition of decomposition is to decompose the whole system into several simple subsystems to be solved independently. As incidence exists among subsystems, popularly, the solution to each subsystem fails to be the solution to the whole system. Sometimes they even contradict each other. Consequently, in order to obtain the solution to the whole system, we have to harmonize these solutions to the subsystems according to the whole objective and incidence constraints in large-scale system[1].

Global optimization was a challenging question confronted by many scientists and technicians all along. As the multi-level optimization algorithm uses decisive search in both up-level and down-level, it is unavoidable to run into local optimization as for non-convex problems. Thus many convexification methods were put forward to solve the difficult problem. However, all these methods have affected the separability of the objective function and made the problem more complicated. In this paper, the author proposed a three-level hierarchical optimization algorithm, which converts the original problem into a separable multi-objective optimization problem. The non-inferior frontier is then convexified and the global optimal solution of the original problem is selected from the set of non-inferior solutions of multi-objective optimization problem.

2. PROBLEM DESPICTION

Consider separable optimization problem in large-scale system as follows:

$$\min \sum_{i=1}^{N} f_i(x_i), \quad \text{s.t.} \quad \sum_{i=1}^{N} g_i(x_i) = 0 \qquad (1)$$

where

x_i decision variable of the *i*th subsystem

f_i objective function of the *i*th subsystem

g_i convex vector function as constraint

Some problems, like complex industrial process, management system and resource distribution, can be expressed the same form as problem (1). So it is important to study the solution algorithm of problem (1).

Primal-dual decomposition can decompose problem (1) into N subsystems provided with minimum dimensions and carry out parallel solution. Assume that Lagrange function of problem (1) are given as follows:

$$L(x,\lambda) = \sum_{i=1}^{N} f_i(x_i) + \sum_{i=1}^{N} \lambda_i^T g_i(x_i)$$

$$= \sum_{i=1}^{N} [f_i(x_i) + \lambda_i^T g_i(x_i)] \qquad (2)$$

When objective function and its constraints are provided with convex configuration, optimal solution of problem (1) can be rewritten as

$$\max_{\lambda}\min_{x}\sum_{i=1}^{N}[f_i(x_i)+\lambda_i^T g_i(x_i)] \qquad (3)$$

By all appearances, two-level optimization algorithm can be used to obtain the optimal solution for problem (1~3). For a fixed Lagrange coefficient λ, let down-level algorithm solve N subsystems independently as follows:

$$\min[f_i(x_i)+\lambda_i^T g_i(x_i)]$$

Then the results are fed back up-level and up-level algorithm will update Lagrange coefficient λ. In this way, we can obtain the global optimal solution of the original problem through continual information exchange between up-level and down-level. It is the convergent condition of algorithm that Lagrange function is convex to x at the optimal point (x^*, λ^*). For non-convex problem, Bertsekas D P[2] thought of the problem in his paper as follows:

$$\min\sum_{i=1}^{N}f_i(x_i)+\frac{c}{2}\|\sum_{i=1}^{N}g_i(x_i)\|^2,$$

$$\text{s.t.} \quad \sum_{i=1}^{N}g_i(x_i)=0 \qquad (4)$$

where $\|g_i(x_i)\|$ is Archimedean normal number. Bertsekas D P has proved that Lagrange function of problem (4) is strictly convex to x in the arbitrarily small domain of x^* when c trends infinity. As a result, primal-dual decomposition can be used to approach the optimization. But crossing items caused by convexification method affected separability in large-scale system and separable action of primal-dual decomposition didn't bright into play. Tanikawa A and Mukai H[3] put forward a class of algorithm which approach crossing items with linear function in their paper. But it can't ensure algorithmic convergence. Tatjewski P and Engelmann B[4] made use of sequence convexification method in their paper. For Lagrange function is convexified locally in these methods, the solution is only the local optimal one. In this paper, Lagrange function isn't convexified. We adopt multi-objective optimization technique and take advantage of the pth power[5] for convexification of non-inferior frontier. As the worth of this method, it doesn't affect separability of large-scale system and ensures that global optimal solution can be obtained in the case of convex constraints domain.

3. MULTI-OBJECTIVE MODEL OF ORIGINAL PROBLEM

Formulate multi-objective optimization problem as follows:

$$\min[f_1(x_1), f_2(x_2),\dots, f_N(x_N)]^T,$$

$$\text{s.t.} \quad \sum_{i=1}^{N}g_i(x_i)=0 \qquad (5)$$

Theorem 1: The global optimal solution of problem (1) must be non-inferior solution of problem (5).

Proof: Assume that $x^*=(x_1^*,x_2^*,\dots,x_N^*)$ is the global optimal solution to problem (1). But it is not non-inferior solution to problem (5). So there is a feasible solution $\hat{x}=(\hat{x}_1,\hat{x}_2,\dots,\hat{x}_N)$ to make $f_i(\hat{x}_i)\le f_i(x_i^*)$ (i=1,2,…,N) and one inequality strictly comes into existence here at least. Due to $\sum_{i=1}^{N}f_i(\hat{x}_i)<\sum_{i=1}^{N}f_i(x_i^*)$, as the global optimal solution to problem (1), x^* contradicts this conclusion. As a result, theorem 1 is proved.

Theorem 1 indicates that the global optimal solution to problem (1) must exist in the set of non-inferior solutions to multi-objective optimization problem (5). Thereby we must obtain non-inferior solution to problem (5). The effective method of obtaining non-inferior solutions is to solve single objective optimization problem as follows:

$$\min\sum_{i=1}^{N}\omega_i f_i(x_i), \quad \text{s.t.} \quad \sum_{i=1}^{N}g_i(x_i)=0 \qquad (6)$$

The total non-inferior solutions can be obtained by searching optimal weight w_i in whole function domain. This conclusion requires that problem (5) is convex non-inferior frontier. When objective function $f_i(x_i)$ is non-convex, not all the points in non-inferior frontier have their supporting plane. Thus it is impossible for problem (6) to produce all the non-inferior solutions. Li Duan[6] put forward a method of using the *p*th power and proved that there is a positive number *q*. When p>q, multi-objective optimization problem is convex non-inferior frontier as follows:

$$\min[f_1^p(x_1), f_2^p(x_2),..., f_N^p(x_N)]$$

$$\text{s.t.} \quad \sum_{i=1}^{N} g_i(x_i) = 0 \qquad (7)$$

By all appearances, problem (7) is equivalent to problem (5). As its remarkable worth, problem (7) is convex non-inferior frontier. So weighting algorithm can be used to obtain all the non-inferior solutions as follows:

$$\min \sum_{i=1}^{N} \omega_i f_i^p(x_i), \quad \text{s.t.} \quad \sum_{i=1}^{N} g_i(x_i) = 0 \qquad (8)$$

Where $\omega_1=1, \omega_i \geqslant 0$, i=2,3,…,N.

4. SELECTING GLOBAL OPTIMAL SOLUTION

When $\omega>0$, problem (7) is called non-degeneration. Degeneration problem denotes some weight factors ω_i are 0, namely, corresponding objective functions in problem (5) don't conflict each other. Getting rid of these objective functions, we can convert degeneration problem into non-degeneration problem. Only non-degeneration problem is discussed in this paper. According to a given weight ω, the solution to problem (8) can be obtained as a function $\hat{x}(\omega)$. ω is replaced by x_i in function $f_i(x_i)$. We define as follows:

$$f_i(\omega) = f_i(\hat{x}_i(\omega)) \qquad \text{(i=1,2,…,N)}$$

For weight ω is a vector of *N-1* dimensions, non-inferior frontier belongs to N-1 dimensions objective space.

Theorem 2: Assume that the solution associated with weight ω^* in problem (8) is the global optimal solution, the optimization condition must come into existence as follows:

$$f_1^{p-1}|_{\omega^*} = \omega_2 f_2^{p-1}|_{\omega^*} = ... = \omega_2 f_N^{p-1}|_{\omega^*} \qquad (9)$$

proof Let $f(\omega) = \sum_{i=1}^{N} f_i(\omega)$. Due to the function $f(\omega)$ being minimized at ω^*, so $\frac{\partial f}{\partial \omega_j} = 0$ as follows:

$$\sum_{i=1}^{N} \frac{\partial f_i(\omega)}{\partial \omega_j}\bigg|_{\omega^*} = 0 \quad \text{(j=1,2,…N)} \qquad (10)$$

Apparently, given ω^* can reach the minimum of problem (8). Used analytic differential, problem (10) can be rewritten by:

$$\sum_{i=1}^{N} \omega_i f_i^{p-1} \frac{\partial f_i}{\partial \omega_j}\bigg|_{\omega^*} = 0 \qquad (11)$$

We define matrix *A* composed of row vectors $\left[\frac{\partial f_1}{\partial \omega_j}, \frac{\partial f_2}{\partial \omega_j},..., \frac{\partial f_N}{\partial \omega_j}\right]$. Its order A=N-1. According to problem (10) and (11), both $[1,1,…,1]^T$ and $[f_1^{p-1}, \omega_2 f_2^{p-1},..., \omega_N f_N^{p-1}]$ are included in the solutions space of *Ax=0*. And dimensions of the solutions space are N-order(A)=1. So two above vectors are in proportion and theorem 2 has been proved.

Let ∇f denote grads of the function f at non-inferior frontier $\{f_1^p, f_2^p,..., f_N^p\}$ as follows:

$$\nabla f = \left[\frac{\partial f}{\partial f_1^p}, \frac{\partial f}{\partial f_2^p},..., \frac{\partial f}{\partial f_N^p}\right]$$

$$= \left[\frac{1}{pf_1^{p-1}}, \frac{1}{pf_2^{p-1}},..., \frac{1}{pf_N^{p-1}}\right]^T \qquad (12)$$

And then, we formulate a direction vector as follows:

$$V(\omega) = [V_1(\omega), V_2(\omega),..., V_N(\omega)]^T$$

$$= -\nabla f(\omega) + (\omega^T \nabla f / \omega^T \omega)\omega \qquad (13)$$

According to Cauchy-Schwarz inequality, we have

$$\nabla f^T(\omega)V(\omega) =$$

$$-\|\nabla f(\omega)\|^2 + (\omega^T \nabla f)^2 / \omega^T \omega \le 0 \qquad (14)$$

It indicates V(ω) is descendent direction of the function *f*. It is obvious that the optimization condition of theorem 2 is satisfied when V(ω)=0.

Assume that weight ω used by the *s*th iteration denotes ω^s, the problem is considered as follows:

$$\min f_1^p(x_1), s.t. f_j^p(x_j) \le f_j^p(\omega^s) + \beta V_j(\omega^s)$$

$$\text{(j=2,3,…,N)}, \quad \sum_{i=1}^{N} g_i(x_i) = 0 \qquad (15)$$

So Lagrange function of equation (12) is given by

$$\min L(x,\lambda) = f_1^p(x_1) +$$

$$\sum_{j=2}^{N} \lambda_j [f_j^p(x_j) - f_j^p(\omega^s) - \beta V_j(\omega^s)] \qquad (16)$$

Where Lagrange multiplier $\lambda_j \geqslant 0$. The dual function of equation (16) is expressed as

$$H(\lambda) = \min_x L(x,\lambda) \qquad (17)$$

For a given λ, how to solve problem (17) is equivalent to how to solve problem (8). Therefore, λ and ω are same in essence and λ can be replaced by ω. According to adjustment formula associated with λ in equation (17), the *(s+1)*th iterative weight ω can be obtained as follows:

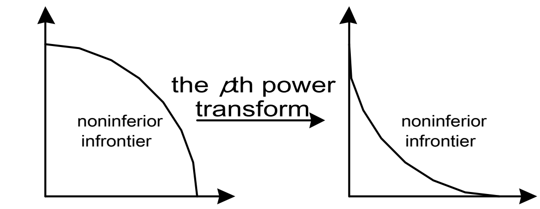
$$\omega_j^{s+1} = \omega_j^s + \beta_1 \frac{\partial H(\omega^s)}{\partial \omega_j} = \omega_j^s - \beta\beta_1 V_j(\omega^s)$$

$$= \omega_j^s - \alpha V_j(\omega^s) \qquad (18)$$

Where $\alpha = \beta\beta_1$. Finally $\partial H / \partial \omega = 0$, namely, V(ω)=0. The solution associated with ω is the global optimal solution. Broadly speaking, in order to obtaining algorithm of problem (1), firstly, we use primal-dual decomposition to solve problem (8). And then weight ω is adjusted by equation (18).

5. EXAMPLES

Example 1 Consider a quarter of circularity in the first quadrant $\{(x_1,x_2) \mid x_1^2 + x_2^2 = 1, x_1 \ge 0, x_2 \ge 0\}$ is a multi-objective optimization problem of non-inferior frontier shown in Fig.1. By all appearances, it is non-convex in objective space $\{x_1,x_2\}$. But non-inferior frontier $\{(x_1^2,x_2^2) \mid [x_1^4]^{1/2} + [x_2^4]^{1/2} = 1\}$ is convex in objective space $\{x_1^4, x_2^4\}$ through the *p*th power transform, where *p*=2.



(a) nonconvex noninferior infrontier (b) convex noninferior infrontier

Fig.1 nonconvex and convex noninferior frontier

Example 2 Assume the objective function is given by

$$f(x)=f_1(x_1)+f_2(x_2)$$

Where $f_1(x_1) = 32x_{11}^2 - 16x_{11} + (2x_{11} + x_{12} - 1)^2 + 6$,

$f_2(x_2) = 10x_{21}^2 + 4x_{21}x_{22} - 0.5(2x_{21} + 2x_{22}^2) + 20$,

$x_1 = (x_{11}, x_{12})^T, x_2 = (x_{21}, x_{22})^T$

The constraints are expressed as follows:

$$x_{22}-2x_{11}-x_{12}=0;\ x_{12}-0.5(x_{21}+x_{22})=0;\ 2x_{11}+x_{12}-2.25\leqslant 0$$

Through introduction of slack variable, we can convert inequality constraints into equation constraints. Used analytical method, two optimal solutions of above problem are obtained as: x_1=(0.28,0.3), x_2=(-0.26,0.86); x_1=(0.68,0.89), x_2=(-0.47,2.25). The first solution's performance index is 24.56 and the second is 35.77. Thus it is obvious that the first solution is the global optimal one. For a given *p*=1.7, firstly we convert problem (1) into problem (8) and adopt primal-dual decomposition obtaining the solutions. And then we adjust ω with equation (18), where iterative error $e = | f_1^p - \omega f_2^p |$, initial value ω_0=0.5, step length s=0.8 and other values are 0. The system error e<0.001 by five iterations and the global optimal solution is obtained eventually. Simulation indicates we can still obtain the global optimal solution when we change the initial values of ω and other variables at random.

6. CONCLUSION

In this paper, an algorithm of searching global optimal solution with multi-objective optimization technique has been presented for separable optimization problem whose objective is non-convex and constraints are convex in large-scale system. In fact, it is a three-level hierarchical optimization algorithm. Weight ω is adjusted at supreme level and Lagrange Multiplier λ is obtained at intermediate level. N subproblems with minimum dimensions are solved independently for given λ and ω at undermost level. And the analytical action of primal-dual decomposition is exerted adequately at this level. Broadly speaking, even if there are incidence constraints in large-scale system, they don't affect the convexification of constraints domain. Consequently, the above algorithm is still applicable to this class of large-scale system.

REFERENCES

[1] Lasdon L S. Optimization theory for large scale system[M]. London: Macmillan Company,1975.

[2] Bertsekas D P. Constrained optimization and Lagrange multiplier methods[M]. Academic Press, 1982

[3] Tanikawa A, Mukai H. A new technique for nonconvex primal-dual decomposition of a large scale separable optimization problem[J]. IEEE Transaction on automatic control, 1985, 30(2): 133~143.

[4] Tatjewski P, Engelmann B. Two level primal-dual decomposition technique for large scale nonconvex optimization problems with constraints[J]. Journalof Optimization Theory and Applications, 1995, 81(1): 306~326.

[5] Li Duan. Zero duality gap for a class of nonconvex optimization problems[J]. Journal of Optimization Theory and Applications, 1995, 85(2): 309~327

[6] Li Duan. Convexification of noninferior frontier[J]. Journal of Optimization Theory and Applications, 1996, 88(1): 177~196