A Global Optimization Approach to Non-Convex Problems

Z.F. Lu, Y Lan

School of Management, Nanjing University of Posts & Telecommunications
Nanjing, P. R. China

Abstract —**In this paper, a novel approach to find the global optimal solution of the special non-convex problems is proposed. The non-convex objective function is first decomposed into two convex sub-functions. Then a generalized gradient is introduced to determine a search direction and the evolution equation is built to obtain a global minimum point. By the approach, we can prevent the search process from some local minima and search a global minimum point. Two numerical examples are given to prove the approach to be effective.**

Keywords-global optimization; non-convex function; generalized gradient; evolution equation

I. INTRODUCTION

This paper is mainly concerned with the following optimization problem:

$$Min \quad J(X) \tag{1.1}$$

S. t.

$$X = (x_1, x_2, ..., x_n)' \in \Omega \subset \Re^n \tag{1.2}$$

Where X is a decision-making vector; the prime denotes transposition of the vector X; Ω represents the feasible domain and is a Banach space; $\Re^n$ is the n-dimensional Euclidean space; J(X) represents a Lipschitz differential non-convex function, which contains some disjunctive convex functions. We want to find the global minimum of J(X).

The formulated problem is often observed and studied in the real world. The methods or algorithms for solving the non-convex function are presented in the previous literatures, which may be classified into two categories: probabilistic and deterministic. The deterministic approach is effective and efficient on the condition that initial guess is near the global minimum, such as multidimensional geometric method [1], tight convex relaxations [2], parametric linearization relaxation algorithm[3], Global Barrier Exclusion algorithm[4]. The probabilistic methods are genetic algorithms[5], tunneling method[6] and auxiliary function approach [7], beta algorithm [8], heuristic Kalman algorithm[9], Which may find the optimal answer at the expense of requiring a high computation burden and offering a probabilistic convergence as well as be sensitive to the variation of the parameters. Because a non-convex function has many minima, some researchers present more global optimal algorithms and methods.

In this paper, we consider a special optimization problem. The problem has a special disjunctive structure in which the non-convex objective function is composed of two convex objective functions. So we can decompose the objective function and transform it a convex optimization problem.

II. DYNAMIC EVOLUTION EQUATION OF THE PROBLEM

A. Decomposition of the Non-Convex Problem

Definition 2.1 A vector $X^* \in \Omega$ is called a local minimum of the problem, if there exist a neighborhood $S(X^*)$ of X^*, such as

$$J(X^*) \le J(V) \quad \text{For all } X^*, V \in S(X^*) \subset \Omega \tag{2}$$

Definition 2.2 A vector $X^* \in \Omega$ is called a strong local minimum of the problem, if there exist a neighborhood $S(X^*)$ of X^*, such as

$$J(X^*) < J(V) \quad \text{For all } X^*, V \in S(X^*) \subset \Omega \tag{3}$$

Definition 2.3 A vector $X^0 \in \Omega$ is called a global minimum of the problem, it will be satisfied the following form:

$$J(X^0) = \underset{X \in \Omega}{Inf} (\quad J(X)) \tag{4}$$

Lemma 2.1 A vector $X^* \in \Omega$ is a strong local minimum of the non-convex problem, which must be satisfied the following necessary conditions:

$$\frac{\partial J}{\partial x_i^*} = 0 \qquad i = 1, 2, ..., n \tag{5}$$

According to Lemma 2.1, we can gain the analytic solution of the problem. Because of the complexity of the problem, it is very difficulty to gain the analytic solution. So many computation algorithms are proposed, such as developing a computation algorithm by the following equation:

$$\frac{dx_i}{dt} = -\lambda \frac{\partial J(X)}{\partial x_i} \qquad i = 1,2,...,n \tag{6}$$

Where λ represents a parameter, $\lambda > 0$; t denotes nominal variable. With the development of the time, the search process evolves in direction towards a local minimum or a global minimum that locates nearby starting point. If we want to find a global minimum or make certain that the minimum obtained is a global minimum, we must make many guess points. Obviously it leads to inefficiency of the approach. Considering our special problem, we can decompose the original problem into two convex sub-problems, that is

$$Min \quad J(X) = f_1(X) - f_2(X) \tag{7.1}$$

S.t.

$$Min \quad f_1(X) \tag{7.2}$$

$$Max \quad f_2(X) \tag{7.3}$$

Where $X = (x_1, x_2,...,x_n)' \in \Omega \subset \Re^n$; $f_j(X)$, $f_j : \Re^n \to \Re, j = 1,2$ are all Lipschitz differential convex function. In terms of (7), we get

$$\frac{dx_i}{dt} = -\lambda \{ \frac{\partial f_1(X)}{\partial x_i} - \frac{\partial f_2(X)}{\partial x_i} \} \qquad i = 1,2,...,n \tag{8}$$

Clearly, when $X^* \in \arg\{X \mid \frac{\partial f_1(X)}{\partial x_i} = \frac{\partial f_2(X)}{\partial x_i} \quad i = 1,2,...,n, X \in \Omega\}$, a local minimum is obtained. As above states, the computation would take much computer time. To get round of this difficulty, we can select the gradient direction of one convex objective function of two convex objective functions, which is the most influential objective at time t , as the search direction in search process. If so, the search direction varies discontinuously. Thus it is necessary to define a generalized gradient concept.

B. Generalized Gradient and Evolution Equation

Definition 2.4 For each point $X \in \Re^n$ and direction $V \in \Re^n$, the Clarke's generalized directional derivative $\partial^0 \Psi(X,V)$ and the generalized gradient of function $\partial f_M(X)$ are defined as follows.

$$\partial^0 \Psi(X,V) = \lim_{(Y,a)\to(X,0)} \sup\{ \frac{f_j(Y+aV) - f_j(Y)}{a}, j = 1,2\} \tag{9}$$

Where

$$\partial f_M(X) = co\{Z \in \Re^n \mid Z = \lim_{Y\to X} \nabla f_M(Y), \nabla f_M(Y) = \sup[\lim_{a\to 0} \frac{f_j(Y+aV) - f_j(Y)}{a}, j=1,2, Y \in \Omega]\} \tag{10}$$

. $\partial^0 \Psi(X,V)$ is a nonempty and compact set. The symbol co stands for a convex hull.

Proposition 2.1 If a g^* is selected as a effective generalized gradient, the g^* must satisfy the following condition

$$< X - X^t, \operatorname{sgn}(f_M(X))g^* > \le 0 \qquad g^*, X, X^t \in \Omega \tag{11}$$

Where X^t represents the position of vector X at time t .

Proof: we construct the following normal cone:

$$N(X^t) = \{u \mid < X - X^t, u \ge 0 \qquad X \in \Omega\}$$

$$M_g(X) = \{u \mid u = \operatorname{sgn}(f_M(X))g^*, g^* \in \partial f_M(X)\}$$

Where g^* is selected as a generalized gradient.

$$\operatorname{sgn}(f_M(X)) = \begin{cases} 1 & M = 1 \\ -1 & M = 2 \end{cases}.$$

Clearly, $N(X^t)$ is included by M_g

$$-\operatorname{sgn}(f_M(X)g^* \in -M_g(X)$$

Therefore

$$< X - X^t, -\operatorname{sgn}(f_M(X)g^* > \quad \ge 0 \qquad i = 1,2,...,n \tag{12}$$

That is $< X - X^t, \operatorname{sgn}(f_M(X)g^* > \le 0$. Qed.

Definition 2.5 If $X \in \Omega$ is called as monotone with respect to module α , the following inequality holds.

$$< X^t - X, \operatorname{sgn}(f_M(X))(g - g^*) > \ge \alpha(\|X - X^t\|) \tag{13}$$

Where
$X \in S(X^t); \mathrm{sgn}(f_M(X))g^* \in M_g(X); \mathrm{sgn}(f_M(X^t))g \in M_g(X^t)$;
$\|\bullet\|$ stands for norms notation.

Theorem 2.1 Given arbitrary initial guess $X_0 \in \Omega$, a global minimum is obtained by the following evolution equation

$$\frac{dX}{dt} = -\lambda g^* \mathrm{sgn}(f_M(X)) \tag{14}$$

Where $\lambda > 0, g^* \in \partial f_M(X)$, g_i is a projection of generalized gradient g on x_i, also is called generalized differentiation of $J(X)$ with respect to x_i, $g = (g_1, g_2, ..., g_n)'$

Proof: From the function $\Phi(X, X^*) = J(X) - J(X^*)$, we get

$$\Phi(X, X^*) > 0 \qquad \forall X, X \in \Omega \in \Re^n \tag{15}$$

If X^* is a global solution of the problem

$$\Phi(X^*, X^*) = J(X^*) - J(X^*) = 0 \tag{16}$$

So $\Phi(X, X^*)$ is a Lyapunov function.

Therefore

$$\Delta\Phi(X, X^*) = \Phi(X, X^*) - \Phi(X^t, X^*)$$
$$= \Phi(X^t + aV, X^*) - \Phi(X^t, X^*) \quad V \in S(X^t) \subset \Omega$$
$$\Delta\Phi(X, X^*) = \Phi(X^t + aV, X^*) - \Phi(X^t, X^*)$$
$$= [J(X^t + aV) - J(X^*)] - [J(X^t) - J(X^*)]$$
$$= \lim_{a \to 0} \sum_{j=1}^{2} \{ \mathrm{sgn}(f_j)(\frac{\partial J(X)}{\partial f_j})' \Delta f_j(X) \} \tag{17}$$

Where $\Delta f_j = <\frac{\partial f_j}{\partial X}, X - X^t>$, $< \bullet, \bullet >$ denotes the inner product of the vector. Substituting it into (17), we get

$$\Delta\Phi(X, X^*) = < \mathrm{sgn}(f_M)g, X - X^t > + 0_L$$

Where 0_L stands for the residual of $\Delta\Phi(X, X^*)$. g is selected as a gradient at time t.

$$g \in S(X^t) \tag{18}$$

Thus $\Delta\Phi(X, X^*) = < \mathrm{sgn}(f_M)g, X - X^t > + 0_L$

$= < \mathrm{sgn}(f_M)(g - g^*), X - X^t > + < \mathrm{sgn}(f_M)g^*, X - X^t > + 0_L$

$\leq -\alpha(\|X - X^t\|) \qquad \frac{d\Phi(X, X^*)}{dt} \leq -a\lambda\|\mathrm{sgn}(f_M)g^*\|$

So

$$\frac{d\Phi(X, X^*)}{dt} \leq 0 \tag{19}$$

Remarks: (19) implies that $\Phi(X, X^*)$ is a monotone descent function. With the development of the nominal time, the search process defined by dynamic equation (18) evolves towards the direction where a global minimum locates. When $\frac{d\Phi(X, X^*)}{dt} = 0$, we obtain a global minimum.

III. NUMERICAL EXAMPLES

Example 1 (one dimensional optimization problem)

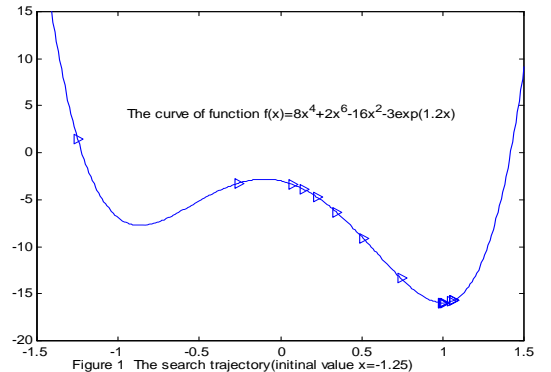*Min* $\quad J(x) = 8x^4 + 2x^6 - 16x^2 - 3e^{1.2x}$

S. t.

$x = [-1.5, 1.5]$

We can decompose the objective function into two sub-objective functions

$$f_1(x) = 8x^4 + 2x^6$$

$$f_2(x) = 16x^2 + 3e^{1.2x}$$

Selecting initial points at random and applying the method of this paper to this optimization problem, we obtain satisfactory results. The global minimum is $x = 1; J(x) = -15.999$. The search process is illustrated in figure 1.

The curve of function f(x)=8x⁴+2x⁶-16x²-3exp(1.2x)

Figure 1 The search trajectory(initial value x=-1.25)

Example 2 (two-dimensional optimization problem)

Considering the Hump Camel-Bak function:

451

$$Min \quad f(X) = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 + 2$$

S.t.

$$X = (x_1, x_2)' \in [-4, 4]$$

The function has 6 local minima including two global ones and four local minima. Here we decompose the function into two sub-functions, that is

$$f_1(X) = (4 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + 4x_2^4 + 2$$

$$f_2(X) = 2.1x_1^4 + 4x_2^2$$

By using the approach proposed in this paper, the global minima are $X^{(1)} = (-0.008984, 0.712656)^T$ and $X^{(2)} = (0.008984, -0.712656)^T$. The optimum solution is $f(X^*) = 0.9684$. The search process is illustrated in figure 2.

Figure 2 The search trajectory(initial value x=[-1.9,2.4])
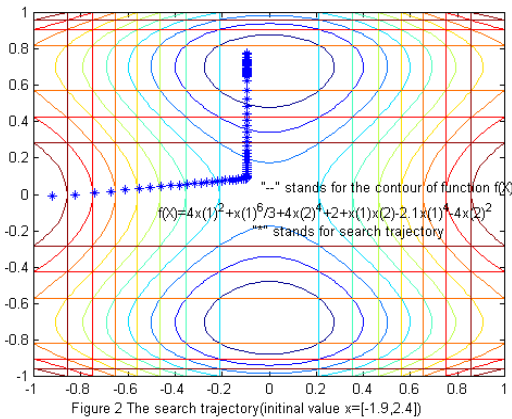
IV. CONCLUDING REMARKS

In this paper, a novel approach to find a global optimal solution of the special non-convex problems is proposed. Because the approach behaves escape mechanism, it can help us to prevent the process from a local minimum.

ACKNOWLEDGMENTS

REFERENCES

[1] Dmitri E. Kvasov, Yaroslav D. Sergeyev.Lipschitz gradients for global optimization in a one-point-based partitioning scheme[J]. Journal of Computational and Applied Mathematics, 236(16),pp4042-4054,2012

[2] Juan P. Ruiz, Ignacio E. Grossmann. Using convex nonlinear relaxations in the global optimization of nonconvex generalized disjunctive programs[J].Computers & Chemical Engineering, 4(11),pp70-84,2013

[3] Huizhuo Wu, KeCun Zhang. A new accelerating method for global non-convex quadratic optimization with non-convex quadratic constraints[J].Applied Mathematics and Computation, 197(2),pp810-818,2008

[4] Nikolaos Kazazakis, Claire S. Adjiman .Globie. An algorithm for the deterministic global optimization of box-constrained NLPs[J].Computer Aided Chemical Engineering, 34,pp669-674, 2014

[5] J. M.Renders and S P Flasse. Hybrid method using genetic algorithms for global optimization[J]. IEEE Trans. Syst. Man Cybernetics , part B, 26,pp243~258,1996

[6] R C Pinaki, Y P Sigh and R Chansarkar. Hybridization of gradient descent algorithms with tunneling methods for global optimization[J]. IEEE Trans. Syst. Man Cybernetics, part A, 30,pp 384-390,2000

[7] M Y Zon and K zon. Global optimization: An auxiliary cost function approach[J]. IEEE Trans. Syst. Man Cybernetics, part A, 30，pp347-354,2000

[8] M. Delgado Pineda. Non-convex global optimization by the beta algorithm: A MAPLE code[J].Nonlinear Analysis: Theory, Methods & Applications, 64(5-7),pp769-777,2005

[9] R. Toscano, P. Lyonnet. A new heuristic approach for non-convex optimization problems[J].Information Sciences, 180(10),pp1955-1966,2010