Research on The Optimal Solution of Lagrangian Multiplier Function Method in Nonlinear Programming

Xue Jin*
Nanjing University of Finance & Economics Hongshan College, Jiangsu Zhenjiang, 212413
*Corresponding author: xuejin2020@nufe.edu.cn

Abstract. For the general nonlinear constrained optimization model, this article will propose a new nonlinear Lagrange function, discuss the properties of the function at the KKT point, and prove that under appropriate conditions, the iterative point generated by the dual algorithm based on the function is locally convergent, and then an error estimate of the solution related to the penalty parameter is given. This provides a new way to solve nonlinear constrained optimization problems.

Keywords: Nonlinear programming, Lagrangian algorithm, function, optimal solution problem.

1. Introduction
The nonlinear optimization problem has been one of the most important branches of optimization problems since the 1970s. It is a problem of finding the extreme value of a single-valued function in a finite-dimensional space. The independent variables of the function may be limited to finite equality or inequality constraint.

In recent years, theories and algorithms on nonlinear optimization problems have continuously appeared, and almost every research branch has corresponding monographs. For general nonlinear optimization problems, a direct idea is to use mathematical tools such as differential calculus, variational method and Lagrange multiplier method to obtain the solution or solution expression of the problem through logical derivation and analysis [1]. This is the so-called analysis method. This method is only effective for some nonlinear optimization problems with special structures. The second type of methods for solving nonlinear optimization problems are graphical methods and experimental methods. Although this type of method is simple to operate, it can only handle low-dimensional situations. In view of the limitations of the above two types of methods, for some practical problems, people try to use the function value information or derivative information of the objective function and the constraint function in a certain local area to construct an iterative numerical solution method, that is, from the current approximate solution point, step by step tune to produce new and better approximate solutions until they can't be improved again. This paper studies the nonlinear constrained optimization problem, discusses the smooth approximation of the low-order exact penalty function and the use of augmented Lagrangian exact penalty function to solve the cone optimization problem.
2. Penalty function method
The main idea of the penalty function method is to transform the nonlinear constrained optimization problem into a nonlinear unconstrained optimization problem [2]. According to how to combine the objective function and the constraint function, we derived many different forms of penalty function.

2.1. External penalty function method
Consider nonlinear programming problems with equality and inequality constraints

\[
\begin{align*}
\min f(x) \\
s.t. \quad g_i(x) &\leq 0, i = 1, \ldots, p \\
&\quad h_j(x) \leq 0, j = 1, \ldots, p
\end{align*}
\] (1)

Among them, \( f: \mathbb{R}^n \to \mathbb{R} \), \( g_i: \mathbb{R}^n \to \mathbb{R}^l (i = 1, L, p) \), \( h_j: \mathbb{R}^n \to \mathbb{R}^q (j = 1, L, q) \) and \( g(x) = (g_1(x), L, g_p(x))^T \), \( h(x) = (h_1(x), L, h_q(x))^T \) [3]. Specifically, a large number \( \delta \) is selected in advance, and the following penalty function is applied to the feasible region \( X \)

\[
P(x) = \begin{cases} 
0, & x \in X \\
\delta, & x \notin X 
\end{cases}
\] (2)

Then, use \( P(x) \) to construct an augmented objective function of (1)

\[
F(x) = f(x) + P(x).
\] (3)

Since the value of \( F \) at the feasible point is the same as the value of \( f \), and the value of \( F \) at the infeasible point is very large, the corresponding unconstrained nonlinear programming problem with the augmented objective function as the objective function

\[
\min F(x) = f(x) + P(x)
\] (4)

The optimal solution of must also be the optimal solution of constrained nonlinear programming problem (1). For (1) we consider introducing the following external penalty function with parameter \( c \):

\[
P_c(x) = c \sum_{i=1}^{p} \left[ \max(g_i(x),0) \right]^2 + \frac{c}{2} \sum_{j=1}^{q} \left[ h_j(x) \right]^2,
\]

where \( c \) is called the penalty parameter. The above external penalty function can also be simply written as

\[
P_c(x) = c \left\| g^+(x) \right\|^2 + \frac{c}{2} \left\| h(x) \right\|^2
\] (5)

Among them, \( g^+(x) = (g_1^+(x), \ldots, g_p^+(x))^T \), \( g_i^+(x) = \max(g_i(x),0)(i = 1, \ldots, p) \). can see that function \( P_c(x) \) can maintain the continuous or differentiable function shape of \( g(x) \) and \( h(x) \) on the boundary of \( X \). Since when a point moves from the boundary of \( X \) to the outside of \( X \), the value of the penalty function (5) will gradually increase from 0. Therefore, only when the penalty parameter \( c \) increases infinitely, the corresponding unconstrained non-constraint of the augmented objective function The optimal solution of linear programming problem \( \min F_c(x) = f(x) + P_c(x) \) may be
close to the optimal solution of (1). Based on this, to use the external penalty function (5) to solve (1), it is generally necessary to first select a positive penalty parameter sequence \( \{C_k\} \) \( (k = 1, 2, \cdots) \) that increases and tends to infinity, and then make the external penalty function sequence

\[
P_{c_k}(x) = c_k \left\| g^+(x) \right\|^2 + \frac{c_k}{2} \left\| h(x) \right\|^2
\]

On the basis of, construct the augmented objective function sequence of (1)

\[
F_{c_k}(x) = f(x) + P_{c_k}(x),
\]

Reduce the problem as solving a series of unconstrained nonlinear programming

\[
\min_{x} F_{c_k}(x).
\]

Then, the point sequence \( \{x^k\} \) formed by the optimal solution \( x^k \) of (8) approaches the optimal solution of (MP). This is the process of solving the problem (MP) with the external penalty function method.

Step 1: Select initial data. Pick the initial point \( x^0 \) and the penalty parameter column \( C_k (k = 1, 2, \cdots) \). Given the error \( \varepsilon > 0 \) of the test termination condition, let \( k_0 = 1 \).

Step 2: Construct an augmented objective function. Follow (6) as the outer penalty function, and then follow (6) to construct the augmented objective function of (1), namely

\[
F_{c_k}(x) = f(x) + c_k \left\| g^+(x) \right\|^2 + \frac{c_k}{2} \left\| h(x) \right\|^2.
\]

Step 3: Solve unconstrained nonlinear programming. Choose some kind of unconstrained nonlinear programming method, Remember \( S(x) = \frac{1}{c_k} P_{c_k} (x) = \left\| g^+(x) \right\|^2 + \frac{1}{2} \left\| h(x) \right\|^2 \), \( S(x^k) \leq \varepsilon \) is often used as the termination condition. You can also make \( g_{max}^{k} = \max_{1 \leq j \leq p} \{ g_j(x^k) \} \), \( h_{max}^{k} = \max_{1 \leq j \leq q} \{ h_j(x^k) \} \) and use \( \max \{ g_{max}^{k}, h_{max}^{k} \} \leq \varepsilon \).

2.2. Internal penalty function method
For nonlinear programming problems with only inequalities

\[
\min_{x} f(x) \quad \text{s.t. } g_i(x) \leq 0
\]

Another penalty function method can also be considered, that is, the internal penalty function method is used to solve the problem. Denote \( g(x) = (g_1(x), \cdots, g_p(x))^T \), then the inside of the feasible region \( X \) of (10) can be denoted as \( X^0 = \{ x \in \mathbb{R}^n | g(x) < 0 \} \). We introduce another penalty function, its role is to give a “punishment” to the point that attempts to leave the feasible region [4]. Because this penalty function keeps the iteration point of the solution within the feasible region, it is called an internal penalty function.
Step 1: Select initial data. Pick the initial point \( x^0 \in X^0 \), and the penalty parameter column \( \{d_k\} (k=1,2,\cdots) \). Given the error \( \varepsilon > 0 \) of the test termination condition, set \( k_0 = 1 \).

Step 2: Construct an augmented objective function. According to (9), make the inner penalty function \( B_{d_k}(x) \), and then construct the augmented objective function of (1) according to (10), namely

\[
F_{d_k}(x) = f(x) + B_{d_k}(x).
\]

Step 3: Solve unconstrained nonlinear programming. Choose some kind of unconstrained nonlinear programming method, solve \( \min F_{d_k}(x) \) with \( x^{k-1} \) as the initial point, and suppose that the optimal solution \( x^k \) is obtained. If \( x^k \) has met a certain termination condition, stop iterative output \( x^k \), otherwise, let \( k_0 = k + 1 \) go to step 2.

In the above internal penalty function method, the initial point must be the internal point of the feasible region. The penalty parameter list \( \{d_k\} \) is usually generated by the following form: first take an initial penalty parameter \( d_0 > 0 \) and a reduction coefficient \( \beta \geq 2 \), let \( d_{k+1} = \frac{d_k}{\beta} \), usually take \( \beta \in [4,10] \). The termination condition can be \( B_{d_k}(x^k) \leq \varepsilon \) or \( \min_{1 \leq i \leq p} g_i(x^k) \leq \varepsilon \).

3. Lagrangian multiplier algorithm for nonlinear programming

Consider the nonlinear programming problem with equality constraints (9), namely

\[
\min f(x) \\
\text{s.t. } h_j(x) = 0, \ j = 1,\cdots, q.
\]

From the last analysis of the previous section, we know that \( \nabla f(x^\star) \neq 0 \) is a big reason why the penalty parameter \( c_k \) must be obtained in the external penalty function method [5]. This enlightens us to consider replacing the objective function \( f(x) \) in (11) with a function whose gradient is zero at the optimal solution \( x^\star \). This function does not seem to be difficult to find, because for the optimal solution \( x^\star \) of (11), there must be a multiplier \( u^\star \in R^q \), such that \( \nabla L(x^\star,u^\star) = 0 \), where \( L(x,u) \) is the Lagrange function \( L(x,u) = f(x) + \sum_{j=1}^q u_j h_j(x) \) of (11). Based on this, imagine replacing \( f(x) \) in (11) with \( L(x,u^\star) \). We can construct a problem equivalent to problem (11)

\[
\min L(x,u^\star) \\
\text{s.t. } h_j(x) = 0, \ j = 1,\cdots, q
\]

Let \( \{c_k\} \) be a sequence of positive penalty parameters that increases and tends to infinity [6]. For problem (12), take the augmented objective function column \( F_{c_k}(x) = L(x,u^\star) + \frac{c_k}{2} \|h(x)\|^2 \). Then the
external penalty function method can be used to solve the problem (11) into solving a series of unconstrained nonlinear programming problems \( \min F_{c_k}(x) = L(x, u^*) + \frac{c_k}{2} \| h(x) \|^2, \ k = 1, 2, \cdots. \)

However, the Lagrange multiplier \( u^* \) cannot be obtained in advance before solving (3.1), which requires us to study the method of establishing an iterative point sequence \( \{u^k\} \) gradually approaching \( u^* \). In the following, we use the dual method to derive the rules for finding the iteration point sequence \( \{u^k\} \), so that the problem of solving (3.1) through (3.2) can be converted to solving the following series of unconstrained nonlinear programming problems:

\[
\min F_{c_k}(x) = L(x, u^k) + \frac{c_k}{2} \| h(x) \|^2, \ k = 1, 2, \cdots.
\]

The basic dual method is that the strong convexity of positive definite is achieved under the requirements.

Let \( x(u) \) be

\[
\min_x L(x, u)
\]

Its local optimal solution, namely

\[
L(x(u), u) = \min_x L(x, u)
\]

Reset

\[
q(u) = \min_x L(x, u)
\]

Then \( q(u) \) is called the dual function of (11), (16) is called the dual program of (11), and \( u^* \) is called the dual optimal solution of (11). In terms of further constructing the augmented Lagrange function, Fletcher's approach gives a relatively new pattern [7]. We can clearly know that for the nonlinear programming problem with equality constraints (11), it is

\[
\min f(x) \quad \text{s.t.} \quad h_j(x) = 0, \ j = 1, \cdots, q.
\]

The form of the augmented Lagrange function used by Hestenes is (15), namely

\[
L_c(x, u) = f(x) + \frac{\xi}{2} \| h(x) \|^2 + u^T h(x).
\]

The augmented Lagrange function in the Fletcher multiplier method proposed by Fletcher (1975) has the form

\[
F_c(x) = L_c(x, u(x))
\]

Under the assumption that \( \nabla h(x) \) has rank \( q \), \( u(x) \) defined by the above formula is

\[
u(x) = -[\nabla h(x) \nabla h(x)^T]^{-1} \nabla h(x) \nabla f(x).
\]
This means that Fletcher uses (16) and (17) to give a new Lagrange function $F(x, c)$ without the Lagrange multiplier $u$. Polak et al. extended Fletcher's approach to solving general constrained nonlinear programming problems (1), namely

$$
\min f(x) \\
\text{s.t.} \quad \begin{align*}
    g_i(x) &\leq 0, \quad i = 1, \ldots, p \\
    h_j(x) &= 0, \quad j = 1, \ldots, q.
\end{align*}
$$

(19)

Here $L_c(x, \lambda, u)$ is the augmented objective function of the form (19) given by Rockafellar, namely

$$
L_c(x, \lambda, u) = f(x) + \frac{1}{2c} \left( \|cg(x) + \lambda^+\| - \|\lambda\|^2 \right) + u^T h(x) + \frac{c}{2} \|h(x)\|^2.
$$

(20)

In order to make $\lambda(x)$ and $u(x)$ in (20) satisfy $\lambda^+ = \lambda(x^*)$ and $u^+ = u(x^*)$, Polak et al. let $\lambda(x)$ and $u(x)$ be the optimal solution of the following quadratic function minimization problem:

$$
\min_{\lambda, u} \|\nabla f(x) + \nabla g(x)^T \lambda + \nabla h(x)^T u\|_2^2 + \sum_{i=1}^p \left( \lambda_i^g_i(x) \right)^2.
$$

(21)

4. Case analysis

For constrained nonlinear programming problem $\min x^2$ s.t. $1 - x \leq 0$, write the expression $F_c(2)$ of the augmented Lagrange function when $x=2$ according to the Polak multiplier method. Solution: According to (17) and (18), the augmented Lagrange function of the problem is

$$
F_c(x) = x^2 + \frac{1}{2c} \left( c - cx + \lambda(x) \right)^2 - \frac{1}{2c} \|\lambda(x)\|^2.
$$

(22)

It can be known from (22) that $\lambda(x)$ is the optimal solution to $\lambda$ for the following unconstrained nonlinear programming problem: $\min \lambda (2x - \lambda)^2 + \lambda^2 (1 - x)^2$. When $x=2$, the above formula is $\min_\lambda (4 - \lambda)^2 + \lambda^2$. Solving the above minimization problem, you can get the optimal solution $\lambda(x) = 2$. Substituting it into (19), you can get

$$
F_c(2) = 4 + \frac{1}{2c} \left( c - 2c + 2 \right)^2 - \frac{1}{2c} \cdot 4 = 4 + \frac{1}{2c} \left( 2 - c \right)^2 - \frac{2}{c}
$$

(23)

Therefore, $F_c(2) = \begin{cases} 2 + \frac{c}{2}, & 0 \leq c \leq 2, \\
4 - \frac{2}{c}, & c > 2. \end{cases}$
5. Conclusion
This paper has drawn many conclusions through the theoretical explanation of the exact penalty function and the calculation and solution of nonlinear constrained optimization problems, such as the realization of some properties of approximating the exact penalty function. More importantly, the accuracy of the original penalty function in nonlinear constrained optimization problems solving, on the basis of keeping the value of the non-differentiable part of the original penalty function problem unchanged, it satisfies the differentiability and smoothness of the penalty function formula, thereby improving the efficiency and quality of the precise penalty function algorithm, and improving the algorithm The convergence rate of, proves the effectiveness and practicability of the exact penalty function method in solving nonlinear constrained optimization problems.

References
[1] Shariat, M., Shariati, M., Madadi, A., & Wakil, K. Computational Lagrangian Multiplier Method by using for optimization and sensitivity analysis of rectangular reinforced concrete beams. Steel Compos Struct, 29(2) (2018) 243-256.
[2] Xia, Y., Wang, J., & Guo, W. Two projection neural networks with reduced model complexity for nonlinear programming. IEEE transactions on neural networks and learning systems, 31(6) (2019) 2020-2029.
[3] Wu, J. Real-time magnetometer disturbance estimation via online nonlinear programming. IEEE Sensors Journal, 19(12) (2019) 4405-4411.
[4] Franke, M. B. Design of Dividing-Wall Columns by Mixed-Integer Nonlinear Programming Optimization. Chemie Ingenieur Technik, 89(5) (2017) 582-597.
[5] Zhang, Y., Li, X., Zhao, G., Lu, B., & Cavalcante, C. C. Signal reconstruction of compressed sensing based on alternating direction method of multipliers. Circuits, Systems, and Signal Processing, 39(1) (2020) 307-323.
[6] Kanno, Y., & Kitayama, S. Alternating direction method of multipliers as a simple effective heuristic for mixed-integer nonlinear optimization. Structural and Multidisciplinary Optimization, 58(3) (2018) 1291-1295.
[7] Fukuda, E. H., & Lourenço, B. F. Exact augmented Lagrangian functions for nonlinear semidefinite programming. Computational Optimization and Applications, 71(2) (2018) 457-482.