Lagrange Multipliers and their Applications

Huijuan Li

Department of Electrical Engineering and Computer Science
University of Tennessee, Knoxville, TN 37921 USA
(Dated: September 28, 2008)

This paper presents an introduction to the Lagrange multiplier method, which is a basic mathematical tool for constrained optimization of differentiable functions, especially for nonlinear constrained optimization. Its application in the field of power systems economic operation is given to illustrate how to use it.

INTRODUCTION

Optimization problems, which seek to minimize or maximize a real function, play an important role in the real world. It can be classified into unconstrained optimization problems and constrained optimization problems. Many practical uses in science, engineering, economics, or even in our everyday life can be formulated as constrained optimization problems, such as the minimization of the energy of a particle in physics;[1] how to maximize the profit of the investments in economics.[2]

In unconstrained problems, the stationary points theory gives the necessary condition to find the extreme points of the objective function $f(x_1, \cdots, x_n)$. The stationary points are the points where the gradient ∇f is zero, that is each of the partial derivatives is zero. All the variables in $f(x_1, \cdots, x_n)$ are independent, so they can be arbitrarily set to seek the extreme of f. However when it comes to the constrained optimization problems, the arbitration of the variables does not exist. The constrained optimization problems can be formulated into the standard form as:[3]

$$min\ f(x_1, \cdots, x_n) \qquad (1)$$
$$\text{Subject to: } \mathbf{G}(x_1, \cdots, x_n) = 0 \qquad (2)$$
$$\mathbf{H}(x_1, \cdots, x_n) \leq 0 \qquad (3)$$

Where, $\mathbf{G}$, $\mathbf{H}$ are function vectors. The variables are restricted to the feasible region, which refers to the points satisfying the constraints.

Substitution is an intuitive method to deal with optimization problems. But it can only apply to the equality-constrained optimization problems and often fails in most of the nonlinear constrained problems where it is difficult to get the explicit expressions for the variables needed to be eliminated in the objective function. The Lagrange multipliers method, named after Joseph Louis Lagrange, provide an alternative method for the constrained nonlinear optimization problems. It can help deal with both equality and inequality constraints.

In this paper, first the rule for the lagrange multipliers is presented, and its application to the field of power systems economic operation is introduced.

LAGRANGE MULTIPLIERS METHOD

In this section, first the Lagrange multipliers method for nonlinear optimization problems only with equality constraints is discussed. The mathematical proof and a geometry explanation are presented. Then the method is extended to cover the inequality constraints.

Without the inequality constraints, the standard form of the nonlinear optimization problems can be formulated as:

$$min\ f(x_1, \cdots, x_n) \qquad (4)$$
$$\text{Subject to: } \mathbf{G}(x_1, \cdots, x_n) = 0 \qquad (5)$$

Where, $\mathbf{G} = [G_1(x_1, \cdots, x_n) = 0, \cdots, G_k(x_1, \cdots, x_n) = 0]^T$, the constraints function vector.

The Lagrange function F is constructed as:[4]

$$F(\mathbf{X}, \lambda) = f(\mathbf{X}) - \lambda \mathbf{G}(\mathbf{X}) \qquad (6)$$

Where, $\mathbf{X} = [x_1, \ldots, x_n]$, the variable vector, $\lambda = [\lambda_1, \cdots, \lambda_k]$, $\lambda_1, \cdots \lambda_k$ are called Lagrange multipliers.

The extreme points of the f and the lagrange multipliers λ satisfy:

$$\nabla F = 0 \qquad (7)$$

that is:

$$\frac{\partial f}{\partial x_i} - \sum_{m=1}^{k} \lambda_m \frac{\partial G_m}{x_i} = 0,\ i = 1, \ldots n \qquad (8)$$

and

$$\mathbf{G}(x_1, \cdots, x_n) = 0 \qquad (9)$$

Lagrange multipliers method defines the necessary conditions for the constrained nonlinear optimization problems.

Mathematical Proof for Lagrange Multipliers Method

The proving is illustrated on the the nonlinear optimization problem (10)-(12), which has four variables and

two equality constraints.

$$u = f(x, y, z, t) \tag{10}$$
$$\text{Subject to: } \Phi(x, y, z, t) = 0 \tag{11}$$
$$\Psi(x, y, z, t) = 0 \tag{12}$$

Assuming that at the point $P(\xi, \eta, \zeta, \tau)$ the function takes an extreme value when compared with the values at all neighboring points satisfying the constraints. Also assuming that at the point P the Jacobian

$$\frac{\partial(\Phi, \Psi)}{\partial(z, t)} = \Phi_z \Psi_t - \Phi_t \Psi_z \neq 0 \tag{13}$$

Thus at the point of P, we can represent two of the variables, say z and t, as functions of the other two, x and y, by means (11) and (12). Substitute the functions $z = g(x, y)$ and $t = h(x, y)$ in (10), then we get an objective function of the two independent variables x and y, and this function must have a free extreme value at the point $x = \xi, y = \eta$; that is its two partial derivatives must vanish at that point. The two equations (14) and (15) must therefore hold.

$$f_x + f_z \frac{\partial z}{\partial x} + f_t \frac{\partial t}{\partial x} = 0 \tag{14}$$

$$f_y + f_z \frac{\partial z}{\partial y} + f_t \frac{\partial t}{\partial y} = 0 \tag{15}$$

We can determine two numbers λ and μ in such a way that the two equations

$$f_z - \lambda \Phi_z - \mu \Psi_z = 0 \tag{16}$$
$$f_t - \lambda \Phi_t - \mu \Psi_t = 0 \tag{17}$$

are satisfied at the point where the extreme value occurs. Take the partial derivative with respect to x on (11) and (12) and get

$$\Phi_x + \Phi_z \frac{\partial z}{\partial x} + \Phi_t \frac{\partial t}{\partial x} = 0 \tag{18}$$

$$\Psi_x + \Psi_z \frac{\partial z}{\partial x} + \Psi_t \frac{\partial t}{\partial x} = 0 \tag{19}$$

Multiply (18) and (19) by $-\lambda$ and $-\mu$ respectively and add them to (14), and we get

$$f_x - \lambda \Phi_x - \mu \Psi_x + (f_z - \lambda \Phi_z - \mu \Psi_z) \frac{\partial z}{\partial x}$$
$$+ (f_t - \lambda \Phi_t - \mu \Psi_t) \frac{\partial t}{\partial x} = 0 \tag{20}$$

Hence by the definition of λ and μ, we can get

$$f_x - \lambda \Phi_x - \mu \Psi_x = 0 \tag{21}$$

Similarly, we can get

$$f_y - \lambda \Phi_y - \mu \Psi_y = 0 \tag{22}$$

Thus, We arrive at the same conclusion as shown in (7) or (8) and (9). The proving method can be similarly extended to the objective functions with more variables and more constraints.
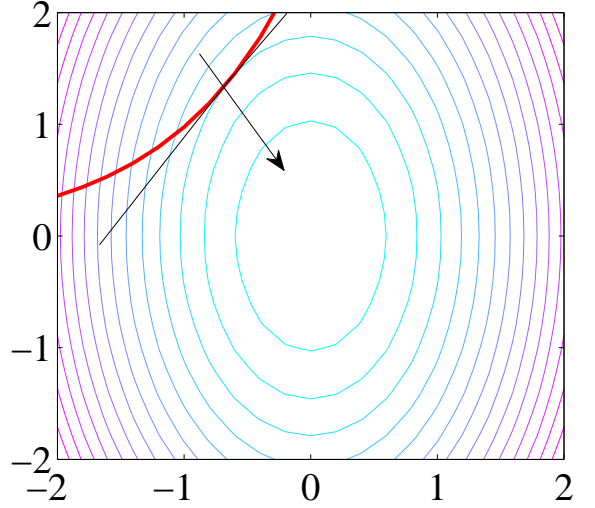

FIG. 1: Contour Curves of the Objective Function and the Feasible Region.

Geometric explanation of the method of Lagrange multipliers

The rule for the Lagrange method can also be explained from a geometric point of view. Fig.1 shows the contour curves of an objective function and the feasible area of the solutions. The circle curves are the contour of the objective function f; the thick line is the points which satisfy the constraint $g = 0$. The arrow indicates the descending direction of the objective function. Along the direction, f keeps dropping and intersects the feasible line until it reaches a level where they are still touching and the two curves ar tangent. But beyond this level, the contour curves do not intersect the feasible line any more. At the touching point, f get the extreme value and the normal vectors of f and g are parallel, and hence:

$$\nabla f(p) = \lambda \nabla g(p) \tag{23}$$

which is equivalent to (8).

Extension to Inequality Constraints

The Lagrange multipliers method also covers the case of inequality constraints, as (3). In the feasible region, $H(x_1, \cdots, x_n) = 0$ Or $H(x_1, \cdots, x_n) < 0$. When $H_i = 0, H$ is said to be active, otherwise H_i is inactive. The augmented Lagrange function is formulated as:[5]

$$F(\mathbf{X}, \lambda, \mu) = f(\mathbf{X}) - \lambda \mathbf{G}(\mathbf{X}) - \mu \mathbf{H}(\mathbf{X}) \tag{24}$$

Where: $\mu = [\mu_1, \cdots, \mu_m]$,
$\mathbf{H} = [H_1(x_1, \cdots, x_n), \cdots, H_m(x_1, \cdots, x_n)]^T$
When H_i is inactive, we can simply remove the constraint by setting $\mu_i = 0$. If $\nabla f < 0$, it points to the

descending direction of f and when H_i is active, this direction points out of the feasible region and towards the forbidden side, which means $\nabla H_i > 0$. This is not the solution direction. We can enforce $\mu_i \leq 0$ to keep the seeking direction still in the feasible region.

When extend to cover the inequality constraints, the rule for the Lagrange multipliers method can be generalized as:

$$\nabla f(\mathbf{X}) - \sum_{i=1}^{k} \lambda_i \nabla G_i(\mathbf{X}) - \sum_{j=1}^{m} \mu_j \nabla H_j(\mathbf{X}) = 0 \quad (25)$$

$$\mu_i \leq 0 \qquad i = 1 \cdots m \quad (26)$$

$$\mu_i H_i = 0 \qquad i = 1 \cdots m \quad (27)$$

$$G(\mathbf{X}) = 0 \quad (28)$$

In summary, for inequality constraints, we add them to the Lagrange function just as if they were equality constraints, except that we require that $\mu_i \leq 0$ and when $H_i \neq 0$, $\mu_i = 0$. This situation can be compactly expressed as (27).

APPLICATION TO THE POWER SYSTEMS ECONOMIC OPERATION

The Lagrange multipliers method is widely used to solve extreme value problems in science, economics, and engineering. In the cases where the objective function f and the constraints G have specific meanings, the Lagrange multipliers often has an identifiable significance. In economics, if you're maximizing profit subject to a limited resource, λ is the resource's marginal value, sometimes called shadow price. Specifically, the value of the Lagrange multiplier is the rate at which the optimal value of the objective function f changes if you change the constraints. An important application of Lagrange multipliers method in power systems is the economic dispatch, or λ-dispatch problem, which is the cross fields of engineering and economics. In this problem, the objective function to minimize is the generating costs, and the variables are subjected to the power balance constraint. This economic dispatch method is illustrated in the following example.[6]

Three generators with the following cost functions serve a load of 952Mw, Assuming a lossless system, calculate the optimal generation scheduling.

$$f_1: \ x_1 + 0.0625x_1^2 \ \$/\text{hr} \quad (29)$$

$$f_2: \ x_2 + 0.0125x_2^2 \ \$/\text{hr} \quad (30)$$

$$f_3: \ x_3 + 0.0250x_3^2 \ \$/\text{hr} \quad (31)$$

Where, x_i is the output power of the ith generator; f_i is the cost per hour of the ith generator. The cost function f_i with respect to x_i is generated by polynomial curve fitting based on the generator operation data. x_i has the unit Mw. Since $w = \frac{J}{S}$ and the costs to produce $1J$ has the unit \$, we have $[w] \equiv [\$/s] \equiv [\$/hr]$. Hence, the cost f_i has the unit \$/hr.

The first step in determining the optimal scheduling of the generators is to express the problem in mathematical form. Thus the optimization statement is:

$$min: \ f = f_1 + f_2 + f_3 =$$
$$x_1 + 0.0625x_1^2 + x_2 + 0.0125x_2^2 + x_3 + 0.0250x_3^2 \quad (32)$$
$$\text{subject to:} \qquad G = x_1 + x_2 + x_3 - 952 = 0 \quad (33)$$

The corresponding Lagrange function is:

$$F = x_1 + 0.0625x_1^2 + x_2 + 0.0125x_2^2$$
$$+x_3 + 0.0250x_3^2 - \lambda(x_1 + x_2 + x_3 - 952) \quad (34)$$

Setting $\nabla F = 0$ and yields the following set of linear equations:

$$\begin{pmatrix} 0.125 & 0 & 0 & -1 \\ 0 & 0.025 & 0 & -1 \\ 0 & 0 & 0.05 & -1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 952 \end{pmatrix} \quad (35)$$

Solving (33) yields:

$$x_1 = 112Mw$$
$$x_2 = 560Mw$$
$$x_3 = 280Mw$$
$$\lambda = \$15/Mwhr$$

and the constrained cost f is \$7616/hr.

This is the generation scheduling that minimizes the hourly cost of production. The value of λ is the incremental of break-even cost of production. This gives a company a price cut-off for buying or selling generation: if they can purchase generation for less than λ, then their overall costs will decrease. Likewise, if generation can be sold for greater than λ, their overall costs will decrease. Also note that at the optimal scheduling, the value of λ and x_i satisfy:

$$\lambda(\$/Mwhr) = 1 + 0.125x_1 = 1 + 0.025x_2$$
$$= 1 + 0.05x_3 \quad (36)$$

Since λ is the incremental cost for the system, this point is also called the point of "equal incremental cost criterion." Any deviation in generation from the equal incremental cost scheduling will result in an increase in the production cost f.

SUMMARY

The Lagrange multipliers method is a very efficient tool for the nonlinear optimization problems, which is

capable of dealing with both equality constrained and inequality constrained nonlinear optimization problems. Many computational programming methods, such as the barrier and interior point method, penalizing and augmented Lagrange method,[5] have been developed based on the basic rules of Lagrange multipliers method. The Lagrange multipliers method and its extended methods are widely applied in science, engineering, economics and our everyday life.

[1] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Elsevier academic, 2005), 5th ed.

[2] N. Schofield, *Mathematical Methods in Economics and Social Choice* (Springer, 2003), 1st ed.

[3] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, 1982), 1st ed.

[4] R. Courant, *Differential and Integral Calculus* (Interscience Publishers, 1937), 1st ed.

[5] D. P. Bertsekas, *Nonlinear Programming* (Athena Scientfic, 1999), 2nd ed.

[6] M. Crow, *Computational Methods for Electric Power Systems* (CRC, 2003), 1st ed.