The Graphic Derivation Process of Lagrange Multiplier Method

Wenbing Wu, Xiaojian Yuan

School of Big Data, Fuzhou University of Foreign Studies and Trade, Fuzhou, China
Email: wwbysq@fjnu.edu.cn

Abstract

The Lagrange multiplier method is an effective method for finding the extreme values of functions, which can solve some complex mathematical and engineering problems. The introduction of this method in the textbook is only a simple expression, but it does not explain how this expression is derived. This article provides a detailed derivation of the Lagrangian multiplier method expression through graphical methods combined with the gradient principle. The Lagrange multiplier method has important applications in various engineering methods, and one of the closely related applications is the gradient descent method. Therefore, after understanding the working principle of the Lagrange multiplier method, one can deepen their understanding and mastery of related issues in the subsequent learning process.

Keywords

Lagrange Multiplier Method, Gradient, The Gradient Descent Method, Extreme Values

1. Introduction

In mathematical optimization problems, the Lagrange multiplier method (named after mathematician Joseph Louis Lagrange) is a method of finding the extreme values of multivariate functions with variables constrained by one or more conditions. This method transforms an optimization problem with n variables and k constraints into an extreme value problem with $n + k$ variables of an equation system, with no constraints on the variables. This method introduces a new scalar unknown, namely the Lagrange multiplier: the coefficient of each vector in the linear combination of the gradient of the constraint equation. The proof of this method involves partial differentiation, total differentiation, or

chain method, in order to find the value of an unknown number that can make the differential of the set implicit function zero.

Find the function $f(x, y, z)$ under the condition φ, the extreme value at $(x, y, z) = 0$.

The method (step) is:

1) Does the Lagrangian function $L = f(x, y, z) + λφ(x, y, z)$, λ Weigh the Lagrangian multiplier;

2) Find L for x, y, z, respectively, λ finds the partial derivative, obtains the system of equations, and finds the stationary point P (x, y, z);

If the maximum or minimum value of this practical problem exists, generally speaking, there is only one stationary point, so the maximum value can be obtained.

The conditional extremum problem can also be transformed into an unconditional extremum solution, but some conditional relationships are complex and substitution and operation are complicated. Relatively speaking, the "Lagrange multiplier method" does not require substitution and is simpler in operation, which is the advantage.

Condition $φ(x, y, z)$ must be an equation, let's set it as $φ(x, y, z) = m$.

Then create another function $g(x, y, z) = φ(x, y, z) – m$.

$G(x, y, z) = 0$ is replaced by $g(x, y, z)φ(x, y, z)$.

In many extreme value problems, the independent variable of the function is often limited by some conditions, such as designing a rectangular open water tank with a volume of V, determining the length, width, and height to minimize the surface area of the tank. If the length, width, and height of the water tank are x, y, and z, then the volume of the water tank $V = xyz$.

The steel plate area used for welding the water tank is $S = 2xz + 2yz + xy$.

This is actually a problem of finding the minimum value of function S under the V constraint.

This type of extreme value problem with conditional constraints is called a conditional extreme value problem, which generally takes the form of finding the extreme value of function F under conditional constraints.

2. The Difference between Conditional Extremum and Unconditional Extremum

Conditional extremum is an extreme value limited to a submanifold. When a conditional extremum exists, an unconditional extremum may not necessarily exist, and even if there are two, they may not necessarily be equal.

For example, find the lowest point on the curve where the saddle surface $z = x^2 – y^2 + 1$ is cut by the plane XOZ.

From its geometric shape, it can be seen that there are no extreme points on the entire saddle surface, but it is limited to the curve where the saddle surface is cut by the plane. There is a minimum value of 1, which is called the conditional extreme value.

3. Necessary Condition

Find the extreme value of a function under constraints. The point that satisfies the constraint condition is the conditional extreme point of the function, and when the point function satisfies the existence condition of the implicit function, the implicit function is determined by the equation, so the point is the limit point of the unary function.

4. Lagrange Multiplier Method

From the above discussion, it can be seen that the conditional extreme point of the function under the constraint conditions should be the solution of the equation system.

By introducing the so-called Lagrange function (where the real number is referred to as the Lagrange multiplier), the above system of equations is called the system of equations.

Therefore, there are usually three methods to solve conditional extremum (Reichel, Schrder, & Xu, 2023; Zhao & Wang, 2022; Ambartsumyan, 2018; Chen, Xu, Wang, Huang, & Chen, 2022; Salmalian, Alijani, & Azarboni, 2020):

1) The direct method is to solve the system of Equation (1) and represent it as a function that is substituted and eliminated as a variable, transforming the problem into an unconditional extremum problem of the function.

2) In general, it is difficult or even impossible to solve from equation System (1), so the above solution method is often not feasible. The commonly used Lagrange multiplier method eliminates the difficulty of solving equation System (1) and transforms the problem of finding the conditional extremum into the problem of finding the stable points of the following Lagrange function. Then, based on the characteristics of the actual problem being discussed, it is determined which stable points are the extremum sought.

3) Under given conditions, if the unknown number can be substituted or solved, the conditional extremum can be transformed into an unconditional extremum, thereby avoiding the trouble of introducing Lagrangian multipliers.

Given the binary function $z = f(x, y)$ and additional conditions $\varphi(x, y) = 0$, in order to find the extreme point of $z = f(x, y)$ under additional conditions, first do the Lagrangian function

$$F(x, y, \lambda) = f(x, y) + \lambda \varphi'_X(x, y)$$

where λ is a parameter.

Let $F(x, y, \lambda)$, for x and y, the first partial derivative of is equal to zero, that is,

$$F'_x = f'_x(x, y) + \lambda \varphi'_X(x, y) = 0$$

$$F'_y = f'_y(x, y) + \lambda \varphi'_Y(x, y) = 0$$

$$F'_\lambda = \varphi(x, y) = 0$$

Solve x, y, and λ, the (x, y) obtained in this way is the function $z = f(x, y)$ under additional conditions φ possible extreme points at $(x, y) = 0$. If there is only

one such point, the actual problem can directly determine the desired point.

The biggest point of the Lagrange multiplier method is how to obtain the equation: $F(x, y, \lambda) = F(x, y) + \lambda \varphi(x, y)$.

To understand this issue, we start with gradients:

The mountain slope in **Figure 1** will project a contour map of a concentric circle on the ground.

Our current problem is to derive the equation for the Lagrange multiplier method from the contour line.

As shown in **Figure 2**. Assuming starting from any point O, the shortest distance from point O to the curve passing through point A in function $g(x)$ is required. We can do this:

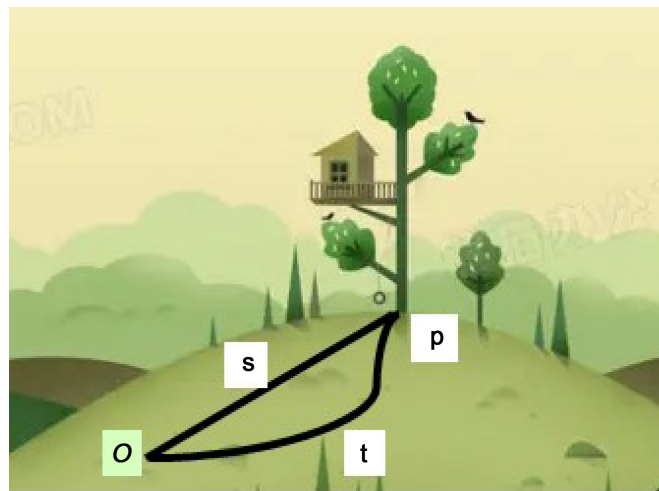Draw circles with a continuously expanding radius centered on point O.
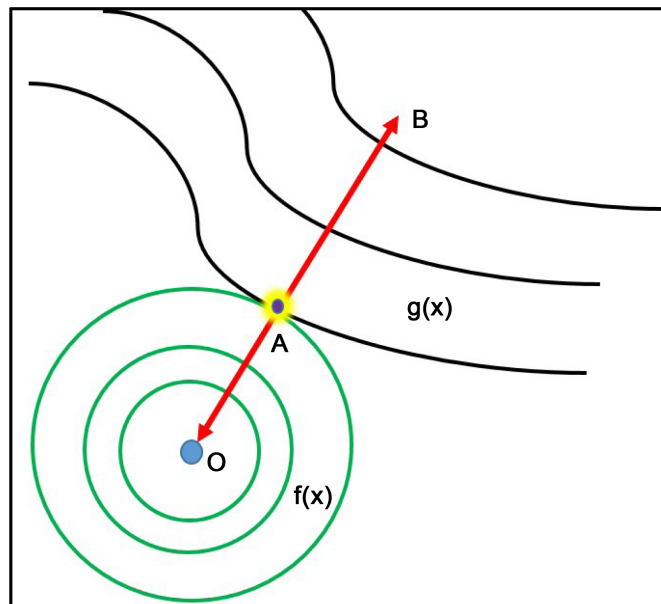

Figure 1. A mountain slope.

Figure 2. The distance from point O to a curve.

When a circle and our target curve happen to have a tangent point, then the line connecting point O and this tangent point is the shortest distance from point O to this curve.

In the above figure, when the circle and curve have a tangent point A, AO is the shortest distance from point O to this curve.

As mentioned earlier, a concentric circle $f(x)$ can be seen as a contour line on a hemispherical surface, so AO is the gradient direction of the concentric circle $f(x)$. Similarly, the curve cluster $g(x)$ can also be seen as the contour lines of a three-dimensional shape. It can be seen that at point A, the gradients of $f(x)$ and $g(x)$ at the same point A are equal in magnitude but opposite in direction. Meanwhile, because a gradient is a vector, its modulus varies in size:

$$grad f(x, y) = \frac{\partial f}{\partial x} \boldsymbol{i} + \frac{\partial f}{\partial y} \boldsymbol{j}$$

Therefore, the gradients of the two functions at point A should satisfy the following relationship (Deparis, Iubatti, & Pegolotti, 2019; Hu et al., 2021; Chaudhary, Raja, Khan et al., 2021; Hijazi, Kandil, Zaarour et al., 2019).

$\Delta f(x) = -\alpha \nabla G(x)$ where α is a constant, i.e.

$$\Delta f(x) + \alpha \nabla G(x) = 0$$

Consider $f(x)$ as $f(x, y)$, and $g(x)$ as $\varphi(x, y)$, we obtain the Lagrange multiplier method:

$$F(x, y, \lambda) = F(x, y) + \lambda \varphi(x, y)$$

The Lagrange multiplier method is an effective method for finding the extreme values of functions, which can solve some complex mathematical and engineering problems. In practical applications, the gradient descent method is closely related to the Lagrange multiplier method. Gradient descent method is a commonly used optimization algorithm for solving unconstrained optimization problems. The basic idea is to continuously update parameters along the gradient direction of the objective function through iteration until reaching the optimal solution or convergence. The core of the gradient descent method is to calculate the gradient of the objective function for each parameter, and then update the parameters in the direction of the negative gradient.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Ambartsumyan, I., Khattatov, E., Yotov, I. et al. (2018). A Lagrange Multiplier Method for a Stokes-Biot Fluid-Poroelastic Structure Interaction Model Mathematics Subject Classification. *Numerische Mathematik, 140,* 513-553. https://doi.org/10.1007/s00211-018-0967-1

Chaudhary, N. I., Raja, M. A. Z., Khan, Z. A. et al. (2021). Hierarchical Quasi-Fractional

Gradient Descent Method for Parameter Estimation of Nonlinear ARX Systems Using Key Term Separation Principle. *Mathematics, 9,* 3302. https://doi.org/10.3390/math9243302

Chen, L. X., Xu, Z. X., Wang, C. R., Huang, D. S., & Chen, Z. N. (2022). Intercomparison-Oriented Evaluation of Equipment's Contribution Rate to Armament System-of-Systems Construction. *Acta Armamentarii, 43,* 1208-1214.

Deparis, S., Iubatti, A., & Pegolotti, L. (2019). Coupling Non-Conforming Discretizations of PDEs by Spectral Approximation of the Lagrange Multiplier Space. *ESAIM: Mathematical Modelling and Numerical Analysis, 53,* 1667-1694.

Hijazi, H., Kandil, N., Zaarour, N. et al. (2019). Impact of Initialization on Gradient Descent Method in Localization Using Received Signal Strength. In *ITM Web of Conferences.*

Hu, J. Y. et al. (2021). Exploring the Measurement Accuracy of Flush Air Data Sensing Based on Normal Cloud Model and Multi-Objective Programming. *Journal of Northwestern Polytechnical University 39,* 987-994. https://doi.org/10.1051/jnwpu/20213950987

Reichel, M., Schrder, J., & Xu, B. X. (2023). Efficient Micromagnetic Finite Element Simulations Using a Perturbed Lagrange Multiplier Method. *PAMM, 22,* e202200016. https://doi.org/10.1002/pamm.202200016

Salmalian, K., Alijani, A., & Azarboni, H. R. (2020). A Lagrange Multiplier-Based Technique within the Nonlinear Finite Element Method in Cracked Columns. *Periodica Polytechnica Civil Engineering, 65,* 84-98 https://doi.org/10.3311/PPci.16395

Zhao, Y. X., & Wang, X. L. (2022). Multiple Robust Estimation of Parameters in Varying-Coefficient Partially Linear Model with Response Missing at Random. *Mathematical Modelling and Control, 2,* 24-33. https://doi.org/10.3934/mmc.2022004