

关于Lagrange乘数法的再探讨

孙慧静¹, 马启建², 刘晓燕¹

¹海军航空大学航空基础学院, 山东 烟台

²烟台文化旅游职业学院, 山东 烟台

收稿日期: 2024年11月24日; 录用日期: 2024年12月18日; 发布日期: 2024年12月26日

摘要

Lagrange乘数法是求条件极值的重要方法。教材中仅仅针对目标函数为二元函数, 约束条件为一个二元方程时, 给出了Lagrange乘数法的基本思想与详细的做法, 但对于自变量多余两个、约束条件多余一个的情形的Lagrange乘数法只是简单提及, 没有给出详尽的推导过程。本文分别从横向和纵向两个维度, 通过层层递进的方式, 按照五种情形, 给出了Lagrange乘数法的一般推广, 并进行了详细的理论推导以及给出了Lagrange乘数法的几何意义。研究结果不论对于一线的科研工作者还是初学者都有一定的启发与借鉴意义。

关键词

条件极值, Lagrange乘数法, 梯度, 法向量

Further Exploration of Lagrange Multiplier Method

Huijing Sun¹, Qijian Ma², Xiaoyan Liu¹

¹School of Basic Sciences for Aviation, Naval Aviation University, Yantai Shandong

²Yantai Vocational College of Culture and Tourism, Yantai Shandong

Received: Nov. 24th, 2024; accepted: Dec. 18th, 2024; published: Dec. 26th, 2024

Abstract

Lagrange multiplier method is an important method for finding conditional extremum. The textbook only provides the basic idea and detailed method of Lagrange multiplier method when the objective function is a binary function and the constraint condition is a binary equation. However, for cases where there are more than two independent variables and more than one constraint, the Lagrange multiplier method is only briefly mentioned without providing a detailed derivation

process. This article provides a general extension of the Lagrange multiplier method from both horizontal and vertical dimensions, using a progressive approach in five different scenarios. Moreover, detailed theoretical derivation and geometric significance of the Lagrange multiplier method are also presented. The research results of this article have certain inspirations and reference significance for both frontline researchers and beginners.

Keywords

Conditional Extremum, Lagrange Multiplier Method, Gradient, Normal Vector

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

Lagrange 乘数法是《高等数学》多元函数部分中的一个重要内容。在求解多元函数条件极值问题时，用到了该方法[1]-[4]。教材中仅仅针对目标函数为二元函数，约束条件为一个二元方程时，给出了 Lagrange 乘数法的基本思想与详细的做法，但对于自变量多余两个、约束条件多余一个的情形的 Lagrange 乘数法只是简单提及，没有给出详尽的推导过程[1]。至于 Lagrange 乘数的几何意义也没有给出解释。

许多文献从不同的角度给出了 Lagrange 乘数法的几何意义[3]-[7]。这些文献要么从低维的角度做了几何解释，要么仅仅通过例子加以说明，而没有给出推导过程。因此，本文主要对于更一般情形——自变量多余两个、约束条件多余一个条件极值的 Lagrange 乘数法进行推广，同时根据推导结果给出了 Lagrange 乘数法的几何意义。

2. Lagrange 乘数法的推广及几何意义

(一) 目标函数为二元函数 $z = f(x, y)$ ，约束条件为二元方程 $\varphi(x, y) = 0$ 。

1. Lagrange 乘数法

定理 1 [1]: 要找目标函数二元函数 $z = f(x, y)$ ，约束条件为二元方程 $\varphi(x, y) = 0$ 下的可能极值点，可以先作 Lagrange 函数：

$$L(x, y) = f(x, y) + \lambda\varphi(x, y),$$

其中 λ 为参数，求其对 x, y 与 λ 一阶偏导数，并使之为零，即

$$\begin{cases} f_x(x, y) + \lambda\varphi_x(x, y) = 0, \\ f_y(x, y) + \lambda\varphi_y(x, y) = 0, \\ \varphi(x, y) = 0, \end{cases} \quad (1)$$

由方程组解出 x, y 与 λ ，这样得到的 (x, y) 就是函数 $f(x, y)$ 在附加条件下 $\varphi(x, y) = 0$ 的可能极值点。

2. 几何意义

下面通过两种方式来解释 Lagrange 乘数法的几何意义。

方式一：假设函数 $z = f(x, y)$ 在 (x_0, y_0) 处取得极值，则 $\varphi(x_0, y_0) = 0$ 。假设 $\varphi(x, y) = 0$ 在 (x_0, y_0) 附近邻域内可以确定隐函数 $y = \psi(x)$ ，则函数目标函数变为 $z = f(x, \psi(x))$ 。又 $z = f(x, \psi(x))$ 在 (x_0, y_0) 处取

得极值, 则 $\frac{dz}{dx}\bigg|_{x=x_0} = 0$ 。

即

$$f_x(x_0, y_0) + f_y(x_0, y_0) \frac{dy}{dx}\bigg|_{(x_0, y_0)} = 0,$$

或

$$(f_x(x_0, y_0), f_y(x_0, y_0)) \cdot \left(1, \frac{dy}{dx}\bigg|_{(x_0, y_0)} \right) = 0, \quad (2)$$

亦或

$$\text{grad}f\bigg|_{(x_0, y_0)} \cdot \mathbf{T} = 0, \quad (3)$$

其中 $\text{grad}f\bigg|_{(x_0, y_0)} = (f_x(x_0, y_0), f_y(x_0, y_0))$, $\mathbf{T} = \left(1, \frac{dy}{dx}\bigg|_{(x_0, y_0)} \right)$ 。

$\text{grad}f\bigg|_{(x_0, y_0)}$ 一方面可以理解为函数 $z = f(x, y)$ 在 (x_0, y_0) 处的梯度, 另一方面也可以理解为在 xOy 平面内 $z = f(x, y)$ 的某等值线 $f(x, y) = C_0$ 在 (x_0, y_0) 处的法向量, 其中 $C_0 = f(x_0, y_0)$, $\mathbf{T} = \left(1, \frac{dy}{dx}\bigg|_{(x_0, y_0)} \right)$ 为曲线 $\varphi(x, y) = 0$ 在 (x_0, y_0) 处的切向量。于是, 等值线 $f(x, y) = C_0$ 在 (x_0, y_0) 处的法线向量与曲线 $\varphi(x, y) = 0$ 在 (x_0, y_0) 处的切向量是互相垂直, 或者说成等值线 $f(x, y) = C_0$ 与曲线 $\varphi(x, y)$ 相切于点 (x_0, y_0) , 或者等值线 $f(x, y) = C_0$ 在 (x_0, y_0) 处的法向量与曲线 $\varphi(x, y) = 0$ 在 (x_0, y_0) 处的法向量是平行的, 即目标函数的梯度方向与约束条件的法向量平行。用线性代数的术语可表述为梯度向量能用约束条件的法向量线性表示。

由此可见, 若 (x_0, y_0) 为目标函数 $z = f(x, y)$ 在约束条件 $\varphi(x, y) = 0$ 的极值点, 其必要条件的几何意义为目标函数 $z = f(x, y)$ 的某等值线 $f(x, y) = C_0$ 与曲线 $\varphi(x, y) = 0$ 相切于 (x_0, y_0) 。

方式二:

对(2)式作进一步的整理, 因为 $\frac{dy}{dx}\bigg|_{(x_0, y_0)} = -\frac{\varphi_x(x_0, y_0)}{\varphi_y(x_0, y_0)}$, 所以(2)式可变为

$$f_x(x_0, y_0) + f_y(x_0, y_0) \left(-\frac{\varphi_x(x_0, y_0)}{\varphi_y(x_0, y_0)} \right) = 0, \quad (4)$$

令 $\frac{f_y(x_0, y_0)}{\varphi_y(x_0, y_0)} = \lambda$, 即

$$f_y(x_0, y_0) - \lambda \varphi_y(x_0, y_0) = 0, \quad (5)$$

则(4)式可变为

$$f_x(x_0, y_0) - \lambda \varphi_x(x_0, y_0) = 0 \quad (6)$$

即极值点一定满足(5)、(6)两式。(5)、(6)两式也可以用向量形式来表达, 即

$$(f_x(x_0, y_0), f_y(x_0, y_0)) - \lambda (\varphi_x(x_0, y_0), \varphi_y(x_0, y_0)) = 0,$$

亦即

$$\text{grad}f - \lambda \mathbf{n} = 0, \quad (7)$$

其中 $\mathbf{n} = (\varphi_x(x_0, y_0), \varphi_y(x_0, y_0))$, 表示曲线 $\varphi(x, y)$ 在 (x_0, y_0) 处的法向量。 (7) 式表示的意义是 $\text{grad}f$ 与 \mathbf{n} 共线(平行)或者用线性代数的术语表述成: $\text{grad}f$ 可用 \mathbf{n} 线性表示。若 (x_0, y_0) 为目标函数 $z = f(x, y)$ 在约束条件 $\varphi(x, y) = 0$ 的极值点, 其必要条件的几何意义为目标函数 $z = f(x, y)$ 的某等值线 $f(x, y) = C_0$ 与曲线 $\varphi(x, y) = 0$ 相切于 (x_0, y_0) 。

综上, 几何上, 约束条件 $\varphi(x, y) = 0$ 在 xOy 面内是一条平面曲线。目标函数 $z = f(x, y)$ 在 xOy 坐标面内有一系列的等值线, 若 $z = f(x, y)$ 在约束条件 $\varphi(x, y) = 0$ 下取得极值, 则必在某等值线与约束曲线相切时取得。若目标函数 $z = f(x, y)$ 的某等值线与曲线 $\varphi(x, y) = 0$ 相交于 (x_0, y_0) 点, 则该点 (x_0, y_0) 一定不可能为极值点。至于等值线与曲线 $\varphi(x, y) = 0$ 相离就更谈不上存在极值点的可能性。

(二) 目标函数为三元函数 $u = f(x, y, z)$, 约束条件为三元方程 $\varphi(x, y, z) = 0$ 。

1. Lagrange 乘数法

定理 2: 要找目标函数二元函数 $u = f(x, y, z)$, 约束条件为二元方程 $\varphi(x, y, z) = 0$ 下的可能极值点, 可以先作 Lagrange 函数:

$$L(x, y, z) = f(x, y, z) + \lambda \varphi(x, y, z),$$

其中 λ 为参数, 求其对 x, y, z 与 λ 一阶偏导数, 并使之为零, 即

$$\begin{cases} f_x(x, y, z) - \lambda \varphi_x(x, y, z) = 0, \\ f_y(x, y, z) - \lambda \varphi_y(x, y, z) = 0, \\ f_z(x, y, z) - \lambda \varphi_z(x, y, z) = 0, \\ \varphi(x, y, z) = 0, \end{cases} \quad (8)$$

由方程组解出 x, y, z 与 λ , 这样得到的 (x, y, z) 就是函数 $u = f(x, y, z)$ 在附加条件下 $\varphi(x, y, z) = 0$ 的可能极值点。

证明: 假定 $\varphi(x, y, z) = 0$ 可确定隐函数 $z = z(x, y)$, 则目标函数变为 $u = f(x, y, z(x, y))$, 假定 $u = f(x, y, z)$ 在 (x_0, y_0, z_0) 处取得极值, 则

$$u_x(x_0, y_0, z_0) = 0, \quad u_y(x_0, y_0, z_0) = 0.$$

即

$$\begin{cases} f_x(x_0, y_0, z_0) + f_z(x_0, y_0, z_0) \frac{\partial z}{\partial x} \Big|_{(x_0, y_0, z_0)} = 0, \\ f_y(x_0, y_0, z_0) + f_z(x_0, y_0, z_0) \frac{\partial z}{\partial y} \Big|_{(x_0, y_0, z_0)} = 0, \end{cases}$$

$$\text{又 } \frac{\partial z}{\partial x} \Big|_{(x_0, y_0, z_0)} = -\frac{\varphi_x(x_0, y_0, z_0)}{\varphi_z(x_0, y_0, z_0)}, \quad \frac{\partial z}{\partial y} \Big|_{(x_0, y_0, z_0)} = -\frac{\varphi_y(x_0, y_0, z_0)}{\varphi_z(x_0, y_0, z_0)},$$

即

$$\begin{cases} f_x(x_0, y_0, z_0) + f_z(x_0, y_0, z_0) \left(-\frac{\varphi_x(x_0, y_0, z_0)}{\varphi_z(x_0, y_0, z_0)} \right) = 0, \\ f_y(x_0, y_0, z_0) + f_z(x_0, y_0, z_0) \left(-\frac{\varphi_y(x_0, y_0, z_0)}{\varphi_z(x_0, y_0, z_0)} \right) = 0, \end{cases}$$

令 $\frac{f_z(x_0, y_0, z_0)}{\varphi_z(x_0, y_0, z_0)} = \lambda$ ，即 $f_z(x_0, y_0, z_0) - \lambda \varphi_z(x_0, y_0, z_0) = 0$ ，故

$$\begin{cases} f_x(x_0, y_0, z_0) - \lambda \varphi_x(x_0, y_0, z_0) = 0, \\ f_y(x_0, y_0, z_0) - \lambda \varphi_y(x_0, y_0, z_0) = 0, \\ f_z(x_0, y_0, z_0) - \lambda \varphi_z(x_0, y_0, z_0) = 0, \\ \varphi(x_0, y_0, z_0) = 0. \end{cases}$$

写成向量形式

$$(f_x, f_y, f_z) \Big|_{(x_0, y_0, z_0)} - \lambda (\varphi_x, \varphi_y, \varphi_z) \Big|_{(x_0, y_0, z_0)} = 0,$$

即

$$\text{grad}f \Big|_{(x_0, y_0, z_0)} - \lambda \mathbf{n} = 0, \quad (9)$$

2. 几何意义

三元函数 $u = f(x, y, z)$ 在约束条件为二元方程 $\varphi(x, y, z) = 0$ 下在 (x_0, y_0, z_0) 取得极值的必要条件(9)式的几何意义为：目标函数 $u = f(x, y, z)$ 的梯度方向与约束条件 $\varphi(x, y, z) = 0$ 在 (x_0, y_0, z_0) 处法向量平行，或梯度向量能用约束条件的法向量线性表示，即 $\text{grad}f \Big|_{(x_0, y_0)} = \lambda \mathbf{n}$ ，其中 $\text{grad}f \Big|_{(x_0, y_0, z_0)} = (f_x, f_y, f_z) \Big|_{(x_0, y_0, z_0)}$ 表示 $u = f(x, y, z)$ 在 (x_0, y_0, z_0) 梯度方向， $\mathbf{n} = (\varphi_x, \varphi_y, \varphi_z) \Big|_{(x_0, y_0, z_0)}$ 表示曲面 $\varphi(x, y, z) = 0$ 在 (x_0, y_0, z_0) 处法向量。

也可以表述为：若 $u = f(x, y, z)$ 在约束条件 $\varphi(x, y, z) = 0$ 下取得极值，则必在某等值面 $f(x, y, z) = C_0$ 与约束曲面 $\varphi(x, y, z) = 0$ 相切时取得。

(三) 目标函数为 n 元函数 $u = f(x_1, x_2, \dots, x_n)$ ，约束条件为 n 元方程 $\varphi(x_1, x_2, \dots, x_n) = 0$ 。

1. Lagrange 乘数法

定理 3：要找目标函数 n 元函数 $u = f(x_1, x_2, \dots, x_n)$ ，约束条件为 n 元方程 $\varphi(x_1, x_2, \dots, x_n) = 0$ 下的可能极值点，可以先作 Lagrange 函数：

$$L(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) + \lambda(x_1, x_2, \dots, x_n),$$

其中 λ 为参数，求其对 x_1, x_2, \dots, x_n 与 λ 一阶偏导数，并使之为零，即

$$\begin{cases} f_{x_1}(x_1, x_2, \dots, x_n) - \lambda \varphi_{x_1}(x_1, x_2, \dots, x_n) = 0, \\ f_{x_2}(x_1, x_2, \dots, x_n) - \lambda \varphi_{x_2}(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ f_{x_n}(x_1, x_2, \dots, x_n) - \lambda \varphi_{x_n}(x_1, x_2, \dots, x_n) = 0, \\ \varphi(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (10)$$

由方程组解出 x_1, x_2, \dots, x_n 与 λ ，这样得到的 (x_1, x_2, \dots, x_n) 就是函数 $u = f(x_1, x_2, \dots, x_n)$ 在附加条件下 $\varphi(x_1, x_2, \dots, x_n) = 0$ 的可能极值点。

证明：假定 $\varphi(x_1, x_2, \dots, x_n) = 0$ 可确定隐函数 $x_n = x_n(x_1, x_2, \dots, x_{n-1})$ ，则目标函数变为 $u = f(x_1, x_2, \dots, x_n(x_1, x_2, \dots, x_{n-1}))$ ，假定 $u = f(x_1, x_2, \dots, x_n)$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 处取得极值，则

$$\frac{\partial u}{\partial x_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0, \frac{\partial u}{\partial x_2} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0, \dots, \frac{\partial u}{\partial x_n} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0.$$

即

$$\begin{cases} f_{x_1}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \frac{\partial x_n}{\partial x_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0, \\ f_{x_2}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \frac{\partial x_n}{\partial x_2} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0, \\ \dots \\ f_{x_{n-1}}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \frac{\partial x_n}{\partial x_{n-1}} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0, \end{cases}$$

又

$$\begin{cases} \frac{\partial x_n}{\partial x_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = -\frac{\varphi_{x_1}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})}, \\ \frac{\partial x_n}{\partial x_2} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = -\frac{\varphi_{x_2}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})}, \\ \dots \\ \frac{\partial x_n}{\partial x_{n-1}} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = -\frac{\varphi_{x_{n-1}}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})}, \end{cases}$$

即

$$\begin{cases} f_{x_1}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \left(-\frac{\varphi_{x_1}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})} \right) = 0, \\ f_{x_2}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \left(-\frac{\varphi_{x_2}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})} \right) = 0, \\ \dots \\ f_{x_{n-1}}(x_{10}, x_{20}, \dots, x_{n0}) + f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) \left(-\frac{\varphi_{x_{n-1}}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})} \right) = 0, \end{cases}$$

$$\Leftrightarrow \frac{f_{x_n}(x_{10}, x_{20}, \dots, x_{n0})}{\varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0})} = \lambda, \text{ 即}$$

$$f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) - \lambda \varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) = 0,$$

故

$$\begin{cases} f_{x_1}(x_{10}, x_{20}, \dots, x_{n0}) - \lambda \varphi_{x_1}(x_{10}, x_{20}, \dots, x_{n0}) = 0, \\ f_{x_2}(x_{10}, x_{20}, \dots, x_{n0}) - \lambda \varphi_{x_2}(x_{10}, x_{20}, \dots, x_{n0}) = 0, \\ \dots \\ f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) - \lambda \varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0}) = 0, \end{cases}$$

写成向量形式

$$gradf \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} - \lambda \mathbf{n} = 0. \quad (11)$$

其中

$$gradf|_{(x_{10}, x_{20}, \dots, x_{n0})} = (f_{x_1}(x_{10}, x_{20}, \dots, x_{n0}), f_{x_2}(x_{10}, x_{20}, \dots, x_{n0}), \dots, f_{x_n}(x_{10}, x_{20}, \dots, x_{n0}))$$

$\mathbf{n} = (\varphi_{x_1}(x_{10}, x_{20}, \dots, x_{n0}), \varphi_{x_2}(x_{10}, x_{20}, \dots, x_{n0}), \dots, \varphi_{x_n}(x_{10}, x_{20}, \dots, x_{n0}))$, 表示超曲面 $\varphi(x_1, x_2, \dots, x_n) = 0$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 处的法向量。

2. 几何意义

n 元函数 $u = f(x_1, x_2, \dots, x_n)$ 在约束条件为 n 元方程 $\varphi(x_1, x_2, \dots, x_n) = 0$ 下在 $(x_{10}, x_{20}, \dots, x_{n0})$ 取得极值的必要条件(11)式的几何意义为: 目标函数 $u = f(x_1, x_2, \dots, x_n)$ 的梯度方向与约束条件 $\varphi(x_1, x_2, \dots, x_n) = 0$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 处法向量平行, 或者用线性代数的术语表述成: $gradf$ 可用 \mathbf{n} 线性表示。也可以表述为: 若 $u = f(x_1, x_2, \dots, x_n)$ 在约束条件 $\varphi(x_1, x_2, \dots, x_n) = 0$ 下取得极值, 则必在某等值超曲面 $f(x_1, x_2, \dots, x_n) = C_0$ 与约束超曲面 $\varphi(x_1, x_2, \dots, x_n) = 0$ 相切点 $(x_{10}, x_{20}, \dots, x_{n0})$ 取得。

(四) 目标函数为三元函数 $u = f(x, y, z)$, 约束条件为由 2 个 3 元方程组成的三元方程组 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 。

1. Lagrange 乘数法

定理 4: 要找目标函数二元函数 $u = f(x, y, z)$, 约束条件为一个由 2 个方程组成的三元方程组 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 下的可能极值点, 可以先作 Lagrange 函数:

$$L(x, y, z) = f(x, y, z) + \lambda\varphi(x, y, z) + \mu\psi(x, y, z),$$

其中 λ, μ 为参数, 求其对 x, y, z 与 λ, μ 一阶偏导数, 并使之为零, 即

$$\begin{cases} f_x(x, y, z) + \lambda\varphi_x(x, y, z) + \mu\psi_x(x, y, z) = 0, \\ f_y(x, y, z) + \lambda\varphi_y(x, y, z) + \mu\psi_y(x, y, z) = 0, \\ f_z(x, y, z) + \lambda\varphi_z(x, y, z) + \mu\psi_z(x, y, z) = 0, \\ \varphi(x, y, z) = 0, \\ \psi(x, y, z) = 0, \end{cases} \quad (12)$$

由方程组解出 x, y, z 与 λ, μ , 这样得到的 (x, y, z) 就是函数 $f(x, y, z)$ 在附加条件下 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 的可能极值点。

证明: 设 (x_0, y_0, z_0) 为目标函数 $u = f(x, y, z)$ 在约束条件 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 的极值点, 并设 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 能确定隐函数 $y = y(x)$, $z = z(x)$, 则目标函数 $u = f(x, y, z)$ 变为 $u = f(x, y(x), z(x))$ 。又 $u = f(x, y, z)$ 在 (x_0, y_0, z_0) 处取得极值, 则 $\frac{du}{dx}\bigg|_{x=x_0} = 0$ 。

即

$$f_x(x_0, y_0, z_0) + f_y(x_0, y_0, z_0) \frac{dy}{dx}\bigg|_{(x_0, y_0, z_0)} + f_z(x_0, y_0, z_0) \frac{dz}{dx}\bigg|_{(x_0, y_0, z_0)} = 0,$$

或

$$(f_x(x_0, y_0, z_0), f_y(x_0, y_0, z_0), f_z(x_0, y_0, z_0)) \cdot \left(1, \frac{dy}{dx}\bigg|_{(x_0, y_0, z_0)}, \frac{dz}{dx}\bigg|_{(x_0, y_0, z_0)}\right) = 0,$$

亦或

$$\text{grad}f|_{(x_0, y_0, z_0)} \cdot \mathbf{T} = 0, \quad (13)$$

其中 $\text{grad}f|_{(x_0, y_0, z_0)} = (f_x(x_0, y_0, z_0), f_y(x_0, y_0, z_0), f_z(x_0, y_0, z_0))$, $\mathbf{T} = \left(1, \frac{dy}{dx}|_{(x_0, y_0, z_0)}, \frac{dz}{dx}|_{(x_0, y_0, z_0)}\right)$ 表示空间曲线

$\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 在 (x_0, y_0, z_0) 处的切向量。

空间曲线 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 在 (x_0, y_0, z_0) 处的切向量也可以表示为

$$\mathbf{n}_1 \times \mathbf{n}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \varphi_x & \varphi_y & \varphi_z \\ \psi_x & \psi_y & \psi_z \end{vmatrix}_{(x_0, y_0, z_0)},$$

空间曲线的切向量既垂直于 $\mathbf{n}_1 = (\varphi_x, \varphi_y, \varphi_z)_{(x_0, y_0, z_0)}$ 又垂直于 $\mathbf{n}_2 = (\psi_x, \psi_y, \psi_z)_{(x_0, y_0, z_0)}$, 显然梯度一定与 $\mathbf{n}_1 = (\varphi_x, \varphi_y, \varphi_z)_{(x_0, y_0, z_0)}$ 和 $\mathbf{n}_2 = (\psi_x, \psi_y, \psi_z)_{(x_0, y_0, z_0)}$ 共面, 即梯度可以用 $\mathbf{n}_1 = (\varphi_x, \varphi_y, \varphi_z)_{(x_0, y_0, z_0)}$ 和 $\mathbf{n}_2 = (\psi_x, \psi_y, \psi_z)_{(x_0, y_0, z_0)}$ 线性表示, 即必要条件如下

$$\begin{cases} f_x(x, y, z) + \lambda \varphi_x(x, y, z) + \mu \psi_x(x, y, z) = 0, \\ f_y(x, y, z) + \lambda \varphi_y(x, y, z) + \mu \psi_y(x, y, z) = 0, \\ f_z(x, y, z) + \lambda \varphi_z(x, y, z) + \mu \psi_z(x, y, z) = 0, \\ \varphi(x, y, z) = 0, \\ \psi(x, y, z) = 0. \end{cases} \quad (14)$$

2. 几何意义

三元函数 $u = f(x, y, z)$ 在约束条件 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 下在 (x_0, y_0, z_0) 取得极值的必要条件(14)式的几何意义为: 目标函数 $u = f(x, y, z)$ 的梯度向量在满足约束条件 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 的点 (x_0, y_0, z_0) 处与空间曲线的切向量垂直, 则与法向量平行, 或者说梯度向量能用约束条件的法向量线性表示, 即

$$\text{grad}f|_{(x_0, y_0, z_0)} = \lambda \mathbf{n}_1 + \mu \mathbf{n}_2,$$

其中 $\text{grad}f|_{(x_0, y_0, z_0)} = (f_x, f_y, f_z)|_{(x_0, y_0, z_0)}$ 表示 $u = f(x, y, z)$ 在 (x_0, y_0, z_0) 梯度方向, $\lambda \mathbf{n}_1 + \mu \mathbf{n}_2$ 表示空间曲线 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 的法向量。

几何上, 约束条件 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 表示空间直角坐标系 $Oxyz$ 中的一条空间曲线, 而目标函数在空间直角坐标系 $Oxyz$ 中表示有一系列的等值面, 若 $u = f(x, y, z)$ 在约束条件 $\begin{cases} \varphi(x, y, z) = 0 \\ \psi(x, y, z) = 0 \end{cases}$ 下取得极值, 则必在某等值面与约束曲线相切时取得。

(五) 若目标函数为 n 元函数 $u = f(x_1, x_2, \dots, x_n)$, 约束条件为 $n-1$ 个 n 元方程

$$\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

组成的方程组。

1. Lagrange 乘数法

定理 5: 要找目标函数 n 元函数 $u = f(x_1, x_2, \dots, x_n)$, 约束条件为一个由 $n-1$ 个方程组成的 n 元方程

组 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 下的可能极值点, 可以先作 Lagrange 函数:

$$L(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_1(x_1, x_2, \dots, x_n) + \dots + \lambda_{n-1} \varphi_{n-1}(x_1, x_2, \dots, x_n),$$

其中 λ_k , ($k = 1, \dots, n-1$) 为参数, 求其对 x_1, x_2, \dots, x_n 与 λ_k , ($k = 1, \dots, n-1$) 一阶偏导数, 并使之为零, 即

$$\begin{cases} f_{x_1}(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_{1,x_1}(x_1, x_2, \dots, x_n) + \dots + \lambda_{n-1} \varphi_{n-1,x_1}(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ f_{x_n}(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_{1,x_{n-1}}(x_1, x_2, \dots, x_n) + \dots + \lambda_{n-1} \varphi_{n-1,x_{n-1}}(x_1, x_2, \dots, x_n) = 0, \\ \varphi_1(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (15)$$

由方程组解出 x_1, x_2, \dots, x_n 与 λ_k , ($k = 1, \dots, n-1$), 这样得到的 (x_1, x_2, \dots, x_n) 就是函数 $u = f(x_1, x_2, \dots, x_n)$

在附加条件下 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 的可能极值点。

证明: 设 $(x_{10}, x_{20}, \dots, x_{n0})$ 为目标函数 $u = f(x_1, x_2, \dots, x_n)$ 在约束条件 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 下的极值点,

并设 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 能确定隐函数 $\begin{cases} x_2 = x_2(x_1) \\ x_3 = x_3(x_1) \\ \dots \\ x_n = x_n(x_1) \end{cases}$ 则目标函数 $u = f(x_1, x_2, \dots, x_n)$ 变为

$u = f(x_1, x_2(x_1), \dots, x_n(x_1))$ 。又 $u = f(x_1, x_2, \dots, x_n)$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 处取得极值, 则 $\frac{du}{dx_1} \bigg|_{x_1=x_{10}} = 0$ 。

即

$$f_{x_1}(x_1, x_2, \dots, x_n) + f_{x_2}(x_1, x_2, \dots, x_n) \frac{dx_2}{dx_1} \bigg|_{(x_{10}, x_{20}, \dots, x_{n0})} + \dots + f_{x_n}(x_1, x_2, \dots, x_n) \frac{dx_n}{dx_1} \bigg|_{(x_{10}, x_{20}, \dots, x_{n0})} = 0,$$

或

$$\left(f_{x_1}, f_{x_2}, \dots, f_{x_n} \right)_{(x_{10}, x_{20}, \dots, x_{n0})} \cdot \left(1, \frac{dx_2}{dx_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})}, \dots, \frac{dx_n}{dx_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} \right) = 0,$$

亦或

$$gradf \Big|_{(x_0, y_0, z_0)} \cdot \mathbf{T} = 0,$$

其中

$$gradf \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} = \left(f_{x_1}, f_{x_2}, \dots, f_{x_n} \right)_{(x_{10}, x_{20}, \dots, x_{n0})},$$

$$\mathbf{T} = \left(1, \frac{dx_2}{dx_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})}, \dots, \frac{dx_n}{dx_1} \Big|_{(x_{10}, x_{20}, \dots, x_{n0})} \right)$$

表示空间曲线 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 处的切向量。

空间曲线的切向量 \mathbf{T} 垂直于 $\mathbf{n}_1 = (\varphi_{1x_1}, \varphi_{1x_2}, \dots, \varphi_{1x_n})$, $\mathbf{n}_2 = (\varphi_{2x_1}, \varphi_{2x_2}, \dots, \varphi_{2x_n})$, \dots , $\mathbf{n}_{n-1} = (\varphi_{n-1x_1}, \varphi_{n-1x_2}, \dots, \varphi_{n-1x_n})$, 显然, $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{n-1}$ 是 n 维空间中 $n-1$ 个线性无关的向量, 梯度可以用 $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{n-1}$ 线性表示。

2. 几何意义

几何上, 约束条件 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 表示 n 维空间中的一条空间曲线, 而目标函数在 n 维空间中

表示有一系列的等值超曲面, 若 $u = f(x_1, x_2, \dots, x_n)$ 在约束条件 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 下取得极值, 则必在

某等值超曲面 $u = f(x_1, x_2, \dots, x_n) = C_0$ (某常数) 与约束曲线 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_{n-1}(x_1, x_2, \dots, x_n) = 0 \end{cases}$ 相切时取得。

(六) 若目标函数为 n 元函数 $u = f(x_1, x_2, \dots, x_n)$, 约束条件为 m 个 n 元方程组成的一个方程组

$$\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0, \\ \varphi_2(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (m < n).$$

1. Lagrange 乘数法

定理 6: 要找目标函数二元函数 $u = f(x_1, x_2, \dots, x_n)$, 约束条件为 m 个 n 元方程组成的一个方程组

$$\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0, \\ \varphi_2(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (m < n)$$

下的可能极值点, 可以先作 Lagrange 函数:

$$L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) = f(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_1(x_1, x_2, \dots, x_n) + \dots + \lambda_m \varphi_m(x_1, x_2, \dots, x_n),$$

其中 $\lambda_k, k = 1, 2, \dots, m$ 为参数, 求其对 $x_i, i = 1, 2, \dots, n$ 与 $\lambda_k, k = 1, \dots, m$ 一阶偏导数, 并使之为零, 即

$$\begin{cases} f_{x_1}(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_{1x_1}(x, y, z) + \dots + \lambda_m \varphi_{mx_1}(x, y, z) = 0, \\ f_{x_2}(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_{1x_2}(x, y, z) + \dots + \lambda_m \varphi_{mx_2}(x, y, z) = 0, \\ \dots \\ f_{x_n}(x_1, x_2, \dots, x_n) + \lambda_1 \varphi_{1x_n}(x, y, z) + \dots + \lambda_m \varphi_{mx_n}(x, y, z) = 0, \\ \varphi_1(x_1, x_2, \dots, x_n) = 0, \\ \varphi_2(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (16)$$

由方程组解出 x_1, x_2, \dots, x_n 与 $\lambda_1, \lambda_2, \dots, \lambda_m$, 这样得到的 (x_1, x_2, \dots, x_n) 就是函数 $f(x_1, x_2, \dots, x_n)$ 在附加条

$$\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad (m < n)$$

件下 (x_1, x_2, \dots, x_n) 的可能极值点。

证明略。

2. 几何意义

$$n \text{ 元函数 } u = f(x_1, x_2, \dots, x_n) \text{ 在约束条件 } \begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad (m < n) \text{ 下在 } (x_{10}, x_{20}, \dots, x_{n0}) \text{ 取得极值的}$$

必要条件 (16) 式的几何意义为: 目标函数 $u = f(x_1, x_2, \dots, x_n)$ 的梯度向量在满足约束条件

$$\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad (m < n) \text{ 的点 } (x_{10}, x_{20}, \dots, x_{n0}) \text{ 处与空间超曲面的切平面垂直, 即与空间超曲面的法向}$$

量平行, 或者说梯度向量能用约束条件的法向量线性表示, 即

$$\text{grad}f|_{(x_{10}, x_{20}, \dots, x_{n0})} = \lambda_1 \mathbf{n}_1 + \lambda_2 \mathbf{n}_2 + \dots + \lambda_m \mathbf{n}_m$$

其中 $\text{grad}f|_{(x_{10}, x_{20}, \dots, x_{n0})}$ 表示 $u = f(x_1, x_2, \dots, x_n)$ 在 $(x_{10}, x_{20}, \dots, x_{n0})$ 梯度方向, $\lambda_1 \mathbf{n}_1 + \lambda_2 \mathbf{n}_2 + \dots + \lambda_m \mathbf{n}_m$ 表示空间

超曲面 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} (m < n)$ 的法向量。

几何上, 约束条件 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} (m < n)$ 表示 n 维空间的 m 维的超曲面, 而目标函数 n 维空间表示有一系列的等值超曲面, 若 $u = f(x_1, x_2, \dots, x_n)$ 在约束条件 $\begin{cases} \varphi_1(x_1, x_2, \dots, x_n) = 0 \\ \varphi_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ \varphi_m(x_1, x_2, \dots, x_n) = 0 \end{cases} (m < n)$ 下取得极值,

则必在某等值超曲面与约束超曲面相切时取得。

3. 小结

针对教材中仅仅对目标函数为二元函数, 约束条件为一个二元方程时, 给出了 Lagrange 乘数法的基本思想与详细的做法, 但对于自变量多余两个、约束条件多余一个的情形 Lagrange 乘数法只是简单提及, 没有给出详尽的推导过程的现象, 本文主要探讨了 Lagrange 乘数法的一般意义的推广以及几何意义。

按照“横向”和“纵向”两个维度对 Lagrange 乘数法进行推广。Lagrange 乘数法的“横向”推广包括如下情形: 1. 目标函数为三元函数, 约束条件为一个三元方程; 2. 目标函数为 n 元函数, 约束条件为一个 n 元方程。Lagrange 乘数法的“纵向”推广包括如下情形: 1. 目标函数为三元函数, 约束条件为由 2 个 3 元方程组成的三元方程组; 2. 目标函数为 n 元函数, 约束条件为 $n-1$ 个 n 元方程组成的方程组; 最后得出了 Lagrange 乘数法一般推广: 目标函数为 n 元函数, 约束条件为 m 个 $n (m < n)$ 元方程组成的方程组。

至于 Lagrange 乘数法的几何意义可以统一理解为: 若目标函数在约束条件下取得极值, 则必在目标函数等于某常数时表示的等值超曲面与约束条件表示的超曲面相切时取得。或者用线性代数的术语叙述: 若目标函数在约束条件下在某点处取得极值, 则目标函数在该点处的梯度向量能用约束条件方程的法向量线性表示。

基金项目

海军航空大学科研自主立项项目(H2202301004)。

参考文献

- [1] 同济大学数学科学学院. 高等数学(下) [M]. 第八版. 北京: 高等教育出版社, 2023: 112-117.
- [2] 杨丽娜. 拉格朗日乘子法几何意义的教学设计[J]. 高等数学研究, 2023, 26(2): 49-51+60.
- [3] 陈建发. 关于拉格朗日乘数法的几何意义[J]. 高等数学研究, 2016, 19(2): 35-36.
- [4] 马明华, 李明, 孙宁. 关于微积分中多元函数极值教学的几点思考[J]. 科技风, 2023(23): 63-65.
- [5] 汪良辉. Lagrange 乘数法的几何意义[J]. 杭州师院学报(自然科学版), 1985(S1): 78-79.
- [6] 杨俊兴. 从几何角度给予拉格朗日乘数法新的推导思路[J]. 高等数学研究, 2016, 19(2): 54-55.
- [7] 李长青. 高等数学教学中应重视几何直观的作用[J]. 高等数学研究, 2007, 10(2): 25-27.