教案

条件极值问题与 Lagrange 乘数法

1. 教学内容

讲解 Lagrange 乘数法的原理，并介绍如何应用 Lagrange 乘数法求解条件极值问题。

2. 指导思想

条件极值问题是实践中经常遇到的应用问题， Lagrange 乘数法是解决条件极值问题的一个有效的工具，也是数学分析课程教学上的一个难点，讲好这一节课程，对提高学生分析问题、并利用微积分这一工具解决问题的能力具有重要意义。

3. 教学安排

1. 在考虑函数的极值或最值问题时，经常需要对函数的自变量附加一定的条件。例如，求原点到直线

$$\begin{cases} x+y+z=1, \\ x+2y+3z=6 \end{cases}$$

的距离，就是在限制条件 $x+y+z=1$ 和 $x+2y+3z=6$ 的情况下，计算函数 $f(x,y,z)=\sqrt{x^2+y^2+z^2}$ 的最小值。这就是所谓的**条件极值**问题。

以三元函数为例，条件极值问题的提法是：求**目标函数**

$$f(x,y,z)$$

在**约束条件**

$$\begin{cases} G(x,y,z)=0, \\ H(x,y,z)=0 \end{cases}$$

下的极值。

假定 f, F, G 具有连续偏导数，且 Jacobi 矩阵

$$J=\begin{pmatrix} G_x & G_y & G_z \\ H_x & H_y & H_z \end{pmatrix}$$

在满足约束条件的点处是满秩的，即 $\operatorname{rank} J=2$。

先考虑取到条件极值的必要条件。上述约束条件实际上是空间曲线的方程。设曲线上一点 (x_0,y_0,z_0) 为条件极值点，由于在该点 $\operatorname{rank} J=2$，不妨假设在 (x_0,y_0,z_0) 点 $\dfrac{\partial(G,H)}{\partial(y,z)}\neq 0$，则由隐函数存在定理，在 (x_0,y_0,z_0) 附近由该方程可以唯一确定

$$y=y(x),\quad z=z(x),\quad x\in O(x_0,\rho)\quad（y_0=y(x_0), z_0=z(x_0)）。$$

它是这个曲线方程的参数形式。

将它们代入目标函数，原问题就转化为函数

$$\Phi(x)=f(x,y(x),z(x)),\quad x\in O(x_0,\rho)$$

的无条件极值问题，x_0 是函数 $\Phi(x)$ 的极值点，因此 $\Phi'(x_0)=0$，即

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

这说明向量

$$\operatorname{grad} f(x_0,y_0,z_0)=f_x(x_0,y_0,z_0)\boldsymbol{i}+f_y(x_0,y_0,z_0)\boldsymbol{j}+f_z(x_0,y_0,z_0)\boldsymbol{k}$$

与向量 $\boldsymbol{\tau}=\left(1,\dfrac{dy}{dx},\dfrac{dz}{dx}\right)$ 正交，即与曲线在 (x_0,y_0,z_0) 点的切向量正交，因此 $\operatorname{grad} f(x_0,y_0,z_0)$ 可看作是曲线在 (x_0,y_0,z_0) 点处的法平面上的向量。由定理 12.5.1，这个法平面是由 $\operatorname{grad} G(x_0,y_0,z_0)$ 与 $\operatorname{grad} H(x_0,y_0,z_0)$ 张成的，因此 $\operatorname{grad} f(x_0,y_0,z_0)$ 可以由 $\operatorname{grad} G(x_0,y_0,z_0)$ 和 $\operatorname{grad} H(x_0,y_0,z_0)$ 线性表出，或者说，存在常数 λ_0,μ_0，使得

$$\operatorname{grad} f(x_0,y_0,z_0)=\lambda_0 \operatorname{grad} G(x_0,y_0,z_0)+\mu_0 \operatorname{grad} H(x_0,y_0,z_0),$$

这就是点 (x_0,y_0,z_0) 为条件极值点的必要条件。

将这个方程按分量写开就是

$$\begin{cases} f_x(x_0,y_0,z_0)-\lambda_0 G_x(x_0,y_0,z_0)-\mu_0 H_x(x_0,y_0,z_0)=0,\\ f_y(x_0,y_0,z_0)-\lambda_0 G_y(x_0,y_0,z_0)-\mu_0 H_y(x_0,y_0,z_0)=0,\\ f_z(x_0,y_0,z_0)-\lambda_0 G_z(x_0,y_0,z_0)-\mu_0 H_z(x_0,y_0,z_0)=0. \end{cases}$$

于是，如果我们构造 **Lagrange 函数**

$$L(x,y,z)=f(x,y,z)-\lambda G(x,y,z)-\mu H(x,y,z)$$

（λ,μ 称为 **Lagrange 乘数**），**则条件极值点就在方程组**

$$\begin{cases} L_x=f_x-\lambda G_x-\mu H_x=0,\\ L_y=f_y-\lambda G_y-\mu H_y=0,\\ L_z=f_z-\lambda G_z-\mu H_z=0,\\ G=0,\\ H=0 \end{cases}$$

的所有解 $(x_0,y_0,z_0,\lambda_0,\mu_0)$ 所对应的点 (x_0,y_0,z_0) 中。用这种方法来求可能的条件极值点的方法，称为 **Lagrange 乘数法**。

2．作为一个例子，现在用 Lagrange 乘数法来解决本节开始提出的问题，即求函数

$$F(x,y,z)=x^2+y^2+z^2$$

在约束条件

$$\begin{cases} x+y+z=1,\\ x+2y+3z=6 \end{cases}$$

下的最小值（最小值的平方根就是距离）。为此，作 Lagrange 函数

$$L(x,y,z,\lambda,\mu)=x^2+y^2+z^2-\lambda(x+y+z-1)-\mu(x+2y+3z-6),$$

在方程组

$$\begin{cases} L_x=2x-\lambda-\mu=0,\\ L_y=2y-\lambda-2\mu=0,\\ L_z=2z-\lambda-3\mu=0,\\ x+y+z-1=0,\\ x+2y+3z-6=0. \end{cases}$$

把方程组中的第一、第二和第三式分别乘以 x、y、z 后相加，再利用第四、第五式得到

$$2(x^2+y^2+z^2)=\lambda+6\mu\text{。}$$

请同学思考，从上式我们能得出什么结论？

答案：从方程组解出λ和μ，如果只有一组解，则$\dfrac{\lambda+6\mu}{2}$就是原点到直线距离的平方！

为此我们只要从方程组解出λ和μ即可。

把方程组中的第一、第二和第三式相加，再利用第四式得

$$3\lambda+6\mu=2\text{；}$$

把第一式、第二式的两倍和第三式的三倍相加，再利用第五式得

$$6\lambda+14\mu=12\text{。}$$

从以上两个方程解得

$$\lambda=-\frac{22}{3},\ \mu=4\text{。}$$

于是原点到直线$\begin{cases}x+y+z=1,\\ x+2y+3z=6\end{cases}$的距离为$\sqrt{\dfrac{1}{2}(-\dfrac{22}{3}+24)}=\dfrac{5}{\sqrt{3}}$。

注 解出λ和μ后，容易得到本题的唯一可能条件极值点为$x=-\dfrac{5}{3},\ y=\dfrac{1}{3},\ z=\dfrac{7}{3}$，因此原点到直线$\begin{cases}x+y+z=1,\\ x+2y+3z=6\end{cases}$的距离为$\sqrt{F\left(-\dfrac{5}{3},\dfrac{1}{3},\dfrac{7}{3}\right)}=\sqrt{\dfrac{25}{3}}=\dfrac{5}{\sqrt{3}}$。

3．一般地，考虑目标函数$f(x_1,x_2,\cdots,x_n)$在m个约束条件

$$g_i(x_1,x_2,\cdots,x_n)=0\quad(i=1,2,\cdots,m;\ \ m<n)$$

下的极值，这里$f,\ g_i\ (i=1,2,\cdots,m)$具有连续偏导数，且 Jacobi 矩阵

$$J=\begin{pmatrix}\dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} & \cdots & \dfrac{\partial g_1}{\partial x_n}\\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} & \cdots & \dfrac{\partial g_2}{\partial x_n}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial g_m}{\partial x_1} & \dfrac{\partial g_m}{\partial x_2} & \cdots & \dfrac{\partial g_m}{\partial x_n}\end{pmatrix}$$

在满足约束条件的点处是满秩的，即$\operatorname{rank}J=m$。那么我们有下述类似的结论：

定理 1（条件极值的必要条件）若点$\boldsymbol{x}_0=(x_1^0,x_2^0,\cdots,x_n^0)$为函数$f(\boldsymbol{x})$满足约束条件的条件极值点，则必存在$m$个常数$\lambda_1,\lambda_2,\cdots,\lambda_m$，使得在$\boldsymbol{x}_0$点成立

$$\operatorname{grad}f=\lambda_1\operatorname{grad}g_1+\lambda_2\operatorname{grad}g_2+\cdots+\lambda_m\operatorname{grad}g_m\text{。}$$

于是可以将 Lagrange 乘数法推广到一般情形。同样地构造 Lagrange 函数

$$L(x_1,x_2,\cdots,x_n,\lambda_1,\lambda_2,\cdots,\lambda_m)=f(x_1,x_2,\cdots,x_n)-\sum_{i=1}^{m}\lambda_i g_i(x_1,x_2,\cdots,x_n)\text{，}$$

那么条件极值点就在方程组

$$(*)\qquad\begin{cases}\dfrac{\partial L}{\partial x_k}=\dfrac{\partial f}{\partial x_k}-\displaystyle\sum_{i=1}^{m}\lambda_i\dfrac{\partial g_i}{\partial x_k}=0,\\ g_l=0,\end{cases}\qquad(k=1,2,\cdots,n;\ l=1,2,\cdots,m)$$

的所有解$(x_1,x_2,\cdots,x_n,\lambda_1,\lambda_2,\cdots,\lambda_m)$所对应的点$(x_1,x_2,\cdots,x_n)$中。

判断如上所得的点是否为极值点有以下的一个充分条件，我们不加证明地给出，请有兴趣的读者将证明补上。

定理 2 设点 $\boldsymbol{x}_0=(x_1^0,x_2^0,\cdots,x_n^0)$ 及 m 个常数 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 满足方程组（*），则当方阵

$$\left(\frac{\partial^2 L}{\partial x_k \partial x_l}(\boldsymbol{x}_0,\lambda_1,\lambda_2\cdots,\lambda_m)\right)_{n\times n}$$

为正定（负定）矩阵时，$\boldsymbol{x}_0$ 为满足约束条件的条件极小（大）值点，因此 $f(\boldsymbol{x}_0)$ 为满足约束条件的条件极小（大）值。

注意，当这个定理中的方阵为不定时，并不能说明 $f(\boldsymbol{x}_0)$ 不是极值。例如，在求函数 $f(x,y,z)=x^2+y^2-z^2$ 在约束条件 $z=0$ 下的极值时，构造 Lagrange 函数 $L(x,y,z)=x^2+y^2-z^2-\lambda z$，并解方程组

$$\begin{cases} L_x = 2x=0, \\ L_y = 2y=0, \\ L_z = -2z-\lambda=0, \\ z=0 \end{cases}$$

得 $x=y=z=\lambda=0$。而在 $(0,0,0,0)$ 点，方阵

$$\begin{pmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{yx} & L_{yy} & L_{yz} \\ L_{zx} & L_{zy} & L_{zz} \end{pmatrix}=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

是不定的。但在约束条件 $z=0$ 下，$f(x,y,z)=x^2+y^2\geq f(0,0,0)=0$，即 $f(0,0,0)$ 是条件极小值。

4．例题

在实际问题中往往遇到的是求最值问题，这时可以根据问题本身的性质判定最值的存在性（如前面的例子）。这样的话，只要把用 Lagrange 乘数法所解得的点的函数值加以比较，最大的（最小的）就是所考虑问题的最大值（最小值）。

例 1 要制造一个容积为 a 立方米的无盖长方形水箱，问这个水箱的长、宽、高为多少米时，用料最省？

解 设水箱的长为 x、宽为 y、高为 z（单位：米），那么问题就变成在水箱容积

$$xyz=a$$

的约束条件下，求水箱的表面积

$$S(x,y,z)=xy+2xz+2yz$$

的最小值。

作 Lagrange 函数

$$L(x,y,z,\lambda)=xy+2xz+2yz-\lambda(xyz-a),$$

从方程组

$$\begin{cases} L_x = y+2z-\lambda yz=0, \\ L_y = x+2z-\lambda xz=0, \\ L_z = 2x+2y-\lambda xy=0, \\ \quad xyz-a=0 \end{cases}$$

得到唯一解

$$x=\sqrt[3]{2a},\quad y=\sqrt[3]{2a},\quad z=\frac{\sqrt[3]{2a}}{2}\text{。}$$

由于问题的最小值必定存在，因此它就是最小值点。也就是说，当水箱的底为边长是$\sqrt[3]{2a}$米的正方形，高为$\sqrt[3]{2a}/2$米时，用料最省。

例 2 求平面$x+y+z=0$与椭球面$x^2+y^2+4z^2=1$相交而成的椭圆的面积。

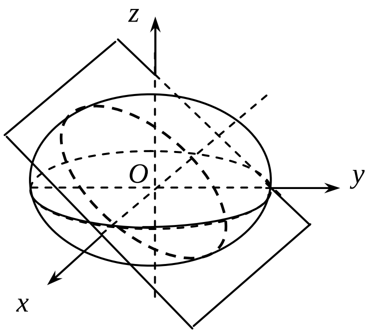

图 12.7.1

解 椭圆的面积为πab，其中a,b分别为椭圆的两个半轴，因为椭圆的中心在原点，所以a,b分别是椭圆上的点到原点的最大距离与最小距离。

于是，可以将问题表述为，求

$$f(x,y,z)=x^2+y^2+z^2$$

在约束条件

$$\begin{cases}x+y+z=0,\\ x^2+y^2+4z^2=1\end{cases}$$

下的最大值与最小值。

作 Lagrange 函数

$$L(x,y,z,\lambda,\mu)=x^2+y^2+z^2-\lambda(x+y+z)-\mu(x^2+y^2+4z^2-1)\text{，}$$

得到相应的方程组

$$\begin{cases}L_x=2(1-\mu)x-\lambda=0,\\ L_y=2(1-\mu)y-\lambda=0,\\ L_z=2(1-4\mu)z-\lambda=0,\\ x+y+z=0,\\ x^2+y^2+4z^2-1=0.\end{cases}$$

将方程组中的第一式乘以x，第二式乘以y，第三式乘以z后相加，再利用$x+y+z=0$和$x^2+y^2+4z^2=1$得到

$$f(x,y,z)=x^2+y^2+z^2=\mu\text{。}$$

请同学思考，从上式我们能得出什么结论？

答案：从方程组解出μ，如果只有二个解μ_1和μ_2，则它们就是该椭圆的半长轴与半短轴的平方！

所以问题转化为求μ的值。

将以上方程组中的第一式乘以$1-4\mu$，第二式乘以$1-4\mu$，第三式乘以$1-\mu$后相加，得到

$$3\lambda(1-3\mu)=0\text{。}$$

分两种情况：

（1） 当$1-3\mu=0$时，得 $\mu=\dfrac{1}{3}$。

（2） 当$\lambda=0$时，原方程组就是

$$\begin{cases}(1-\mu)x=0,\\(1-\mu)y=0,\\(1-4\mu)z=0,\\x+y+z=0,\\x^2+y^2+4z^2-1=0.\end{cases}$$

此时$\mu=1$（否则从以上方程组的第一，第二和第四式得到$x=y=z=0$，这不是椭圆上的点）。

于是得到该椭圆的半长轴为 1，半短轴为$\dfrac{1}{\sqrt{3}}$，面积为$\dfrac{\pi}{\sqrt{3}}$。

许多实际问题并不需要完全解出方程组来求得最值，上述解法是一种常用的方法，可以使解决问题的方法与计算简化。

例 3 求函数 $f(x,y)=ax^2+2bxy+cy^2$（$b^2-ac<0;\ a,b,c>0$）在闭区域$\mathbf{D}=\{(x,y)\,|\,x^2+y^2\le 1\}$上的最大值和最小值。

解 首先考察函数 f 在 $\mathbf{D}$ 的内部$\{(x,y)\,|\,x^2+y^2<1\}$的极值，这是无条件极值问题。为此解线性方程组

$$\begin{cases}f_x=2ax+2by=0,\\f_y=2bx+2cy=0.\end{cases}$$

由假设$b^2-ac<0$知道方程组的系数行列式不等于零，因此只有零解$x=0,y=0$，即$(0,0)$点是驻点。易计算在$(0,0)$点

$$f_{xx}(0,0)=2a,\quad f_{xy}(0,0)=2b,\quad f_{yy}(0,0)=2c\text{，}$$

因此 $f_{xx}(0,0)f_{yy}(0,0)-{f_{xy}}^2(0,0)=4(ac-b^2)>0$。而 $f_{xx}>0$，所以$(0,0)$点是函数 f 的极小值点，极小值为 $f(0,0)=0$。

再考察函数 f 在 $\mathbf{D}$ 的边界$\{(x,y)\,|\,x^2+y^2=1\}$上的极值，这是条件极值问题。为此作 Lagrange 函数

$$L(x,y,\lambda)=ax^2+2bxy+cy^2-\lambda(x^2+y^2-1)\text{，}$$

并得方程组

$$\begin{cases}(a-\lambda)x+by=0,\\bx+(c-\lambda)y=0,\\x^2+y^2-1=0\,.\end{cases}$$

将方程组中的第一式乘以x，第二式乘以y后相加，再用第三式代入就得到

$$f(x,y)=ax^2+2bxy+cy^2=\lambda(x^2+y^2)=\lambda\text{，}$$

这说明 $f(x,y)$在$\{(x,y)\,|\,x^2+y^2=1\}$上的极大值与极小值包含在方程组关于λ的解中。下面来求λ的值。

由联立方程组中的$x^2+y^2-1=0$，可知二元一次方程组$\begin{cases}(a-\lambda)x+by=0\\bx+(c-\lambda)y=0\end{cases}$有

非零解，因此系数行列式等于零，即

$$\lambda^2-(a+c)\lambda+ac-b^2=0\text{。}$$

解这个关于λ的方程，得到

$$\lambda=\frac{1}{2}\left[(a+c)\pm\sqrt{(a+c)^2-4(ac-b^2)}\right]$$

（注意根号中$(a+c)^2-4(ac-b^2)=(a-c)^2+4b^2>0$）。

由于连续函数f在紧集$\{(x,y)\,|\,x^2+y^2=1\}$上必可取到最大值与最小值，因此f在$\mathbf{D}$的边界上的最大值为

$$\lambda_1=\frac{1}{2}\left[(a+c)+\sqrt{(a+c)^2-4(ac-b^2)}\right];$$

最小值为

$$\lambda_2=\frac{1}{2}\left[(a+c)-\sqrt{(a+c)^2-4(ac-b^2)}\right]\text{。}$$

再与f在$\mathbf{D}$内部的极值$f(0,0)=0$比较，就得到f在$\mathbf{D}$上的最大值为

$$\max\{\lambda_1,0\}=\frac{1}{2}\left[(a+c)+\sqrt{(a+c)^2-4(ac-b^2)}\right];$$

最小值为

$$\min\{\lambda_2,0\}=0\text{。}$$

例 4 设$a>0,\ a_i>0$（$i=1,2,\cdots,n$）。求n元函数

$$f(x_1,x_2,\cdots,x_n)=x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$$

在约束条件$x_1+x_2+\cdots+x_n=a$（$x_i>0,\ i=1,2,\cdots,n$）下的最大值。

解 作辅助函数

$$g(x_1,x_2,\cdots,x_n)=\ln f(x_1,x_2,\cdots,x_n)=a_1\ln x_1+a_2\ln x_2+\cdots+a_n\ln x_n,$$

因为函数$\ln u$严格单调，所以只要考虑函数g的极值就可以得到f的极值。

作 Lagrange 函数

$$L=a_1\ln x_1+a_2\ln x_2+\cdots+a_n\ln x_n-\lambda(x_1+x_2+\cdots+x_n-a)\text{。}$$

由极值的必要条件得到

$$\begin{cases}\dfrac{\partial L}{\partial x_i}=\dfrac{a_i}{x_i}-\lambda=0, & i=1,2,\cdots,n,\\ x_1+x_2+\cdots+x_n=a.\end{cases}$$

由前n个方程得到$x_i=\dfrac{a_i}{\lambda}$，$i=1,2,\cdots,n$，再代入最后一个方程得到

$$\lambda=\frac{a_1+a_2+\cdots+a_n}{a},$$

所以

$$x_i=\frac{aa_i}{a_1+a_2+\cdots+a_n},\quad i=1,2,\cdots,n\text{。}$$

于是$(x_1,x_2,\cdots,x_n)$是函数g的唯一可能条件极值点。由于

$$\left(\frac{\partial^2 L}{\partial x_k \partial x_l}(x_1, x_2, \cdots, x_n, \lambda)\right)_{n\times n} = \begin{pmatrix} -\frac{a_1}{x_1^{\ 2}} & \cdots & 0 & 0 \\ \vdots & -\frac{a_2}{x_2^{\ 2}} & & \vdots \\ 0 & & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{a_n}{x_n^{\ 2}} \end{pmatrix}$$

为负定矩阵，由定理 2 可知$(x_1, x_2, \cdots, x_n)$为 g 的条件极大值点。它也是 f 的唯一条件极大值点，显然它就是 f 的条件最大值点。于是 f 在约束条件下的最大值为

$$\prod_{i=1}^{n}\left(\frac{aa_i}{a_1 + a_2 + \cdots + a_n}\right)^{a_i} = a_1^{\ a_1} a_2^{\ a_2} \cdots a_n^{\ a_n}\left(\frac{a}{a_1 + a_2 + \cdots + a_n}\right)^{a_1 + a_2 + \cdots + a_n}。$$

特别地，当 $a_1 = a_2 = \cdots = a_n = 1$ 及 $a = 1$ 时，f 的最大值为$\left(\frac{1}{n}\right)^n$，即当 $x_1 + x_2 + \cdots + x_n = 1$及 $x_i > 0\ \ (i = 1,2,\cdots,n)$时成立

$$x_1 x_2 \cdots x_n \le \left(\frac{1}{n}\right)^n。$$

对于任意正数$y_1, y_2, \cdots, y_n$，只要令

$$x_i = \frac{y_i}{y_1 + y_2 + \cdots + y_n} \quad （i = 1,2,\cdots,n），$$

就得到

$$\prod_{i=1}^{n}\frac{y_i}{y_1 + y_2 + \cdots + y_n} \le \left(\frac{1}{n}\right)^n，$$

即

$$\sqrt[n]{y_1 y_2 \cdots y_n} \le \frac{y_1 + y_2 + \cdots + y_n}{n}。$$

这就是熟知的平均值不等式。

4. 注意点

应用 Lagrange 乘数法求解条件极值问题，产生的方程组变量个数可能比较大，似乎解这个方程组往往是很困难的。但注意我们可以利用变量之间的关系（也就是问题给出的条件），找到解方程组的简便的方法，而不要用死板的方法去解方程组。