《数学中的小问题大定理》丛书（第二辑）

拉格朗日乘子定理

——从一道 2005 年全国高中联赛试题的高等数学解法谈起

刘培杰数学工作室 编

◎约束极值的最优性条件

◎数学规划的拉格朗日乘子

◎凸规划的拉格朗日乘子法则

◎线性规划和拉格朗日乘子的经济解释

◎科学中的数学化

◎第二次世界大战与美国数学的发展

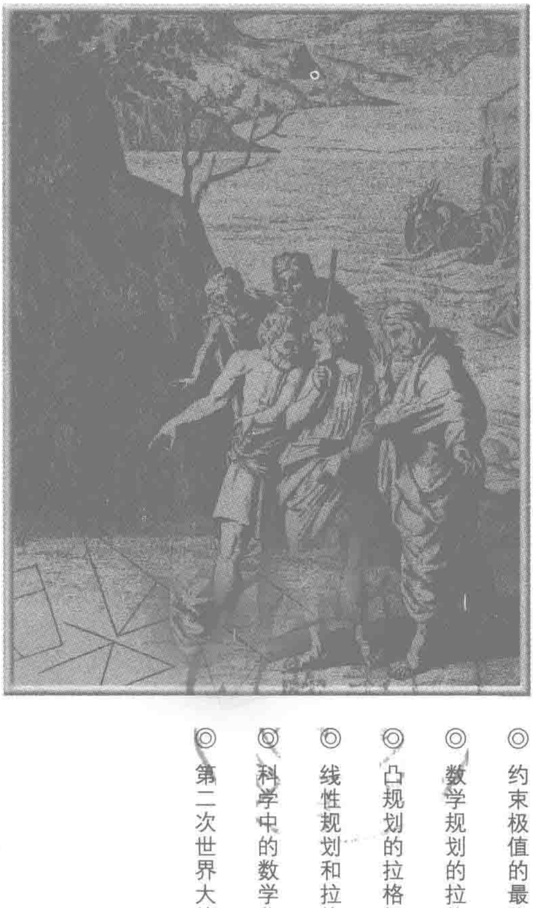
哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

《数学中的小问题大定理》丛书（第二辑）

拉格朗日乘子定理

——从一道2005年全国高中联赛试题的高等数学解法谈起

刘培杰数学工作室 编



◎约束极值的最优性条件

◎数学规划的拉格朗日乘子

◎凸规划的拉格朗日乘子法则

◎线性规划和拉格朗日乘子的经济解释

◎科学中的数学化

◎第二次世界大战与美国数学的发展

哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内 容 简 介

本书从一道2005年全国高中联赛试题的高等数学解法谈起，详细介绍了拉格朗日乘子定理的相关知识及应用. 全书共9章，读者可以较全面地了解这一类问题的实质，并且还可以认识到它在其他学科中的应用.

本书适合大学生及数学爱好者参考阅读.

图书在版编目(CIP)数据

拉格朗日乘子定理：从一道2005年全国高中联赛试题的高等数学解法谈起/刘培杰数学工作室编. ——哈尔滨：哈尔滨工业大学出版社，2015.8

ISBN 978-7-5603-5345-6

Ⅰ. ①拉… Ⅱ. ①刘… Ⅲ. ①中学数学课-高中-教学参考资料 Ⅳ. ①G634.603

中国版本图书馆CIP数据核字(2015)第087127号

策划编辑 刘培杰 张永芹
责任编辑 张永芹 聂兆慈
封面设计 孙茵艾
出版发行 哈尔滨工业大学出版社
社　　址 哈尔滨市南岗区复华四道街10号 邮编150006
传　　真 0451-86414749
网　　址 http://hitpress.hit.edu.cn
印　　刷 哈尔滨工业大学印刷厂
开　　本 787mm×960mm 1/16 印张11 字数118千字
版　　次 2015年8月第1版 2015年8月第1次印刷
书　　号 ISBN 978-7-5603-5345-6
定　　价 28.00元

目录

引　言

第1章

1.1　一道2005年全国高中联赛试题的高等数学解法

华南师范大学数学科学学院2005级研究生曲政在一本杂志中发表了一篇题为利用拉格朗日（Lagrange）乘数法求解一个条件极值问题的短文①.

让我们来研究以下的条件极值问题.

例1　设 n 为正整数，$n \geqslant 2$，正数 $a_1, \cdots, a_n, x_1, \cdots, x_n$ 满足

$$\begin{cases} a_n x_2 + a_{n-1} x_3 + \cdots + a_2 x_n = a_1 \\ a_1 x_3 + a_n x_4 + \cdots + a_3 x_1 = a_2 \\ \qquad\qquad \vdots \\ a_{n-1} x_1 + a_{n-2} x_2 + \cdots + a_1 x_{n-1} = a_n \end{cases} \tag{1.1}$$

① 引自《中学数学研究》2006年第6期第46页.

求函数$f(x_1,\cdots,x_n)=\frac{x_1^2}{1+x_1}+\cdots+\frac{x_n^2}{1+x_n}$的最小值.

解 设函数$g(a_1,\cdots,a_n,x_1,\cdots,x_n)=f(x_1,\cdots,x_n)$,问题等价于求函数$g$满足条件(1.1)时的最值.将(1.1)的$n$个式子相加得

$$(a_1+\cdots+a_n)(x_1+\cdots+x_n)-(a_1x_1+\cdots+a_nx_n)-(a_1+a_2+\cdots+a_n)=0$$

应用拉格朗日乘数法,令

$$\begin{aligned}&L(a_1,\cdots,a_n,x_1,\cdots,x_n,k)\\=&\frac{x_1^2}{1+x_1}+\cdots+\frac{x_n^2}{1+x_n}+\\&k[(a_1+\cdots+a_n)(x_1+\cdots+x_n)-\\&(a_1x_1+\cdots+a_nx_n)-(a_1+\cdots+a_n)]\end{aligned}$$

对L求一阶偏导数,并令它们都等于0,有

$$\begin{aligned}L'_{a_i}&=k(x_1+\cdots+x_n-x_i-1)=0\\L'_{x_i}&=\frac{x_i^2+2x_i}{(1+x_i)^2}+k(a_1+\cdots+a_n-a_i)\\&=0,i=1,\cdots,n\\L'_k&=(a_1+\cdots+a_n)(x_1+\cdots+x_n)-\\&\quad(a_1x_1+\cdots+a_nx_n)-\\&\quad(a_1+\cdots+a_n)=0\end{aligned}$$

由对称性知函数L的稳定点为$(a',\cdots,a',x',\cdots,x',k)$,其中$x'=\frac{1}{n}-1$,则$f(x',\cdots,x')$是$f(x_1,\cdots,x_n)$的最小值,而

$$f(x',\cdots,x')=\frac{nx'^2}{1+x'}=\frac{1}{n-1}$$

此即所求.

上述问题的简单形式为：设正数 a,b,x,y 满足 $bx=a,ax=b$，求函数 $f(x,y)=\frac{x^2}{1+x}+\frac{y^2}{1+y}$ 的最小值. 易见 $f(x,y)$ 的最小值为 1.

例 2 是 2005 年全国高中数学联赛加试第二题.

例 2 设正数 a,b,c,x,y,z 满足 $cy+bz=a,az+cx=b,bx+ay=c$. 求函数 $f(x,y,z)=\frac{x^2}{1+x}+\frac{y^2}{1+y}+\frac{z^2}{1+z}$ 的最小值.

本题有多种初等解法，利用上述结论立即可得 $f(x,y,z)=\frac{x^2}{1+x}+\frac{y^2}{1+y}+\frac{z^2}{1+z}$ 的最小值为 $\frac{1}{2}$.

1.2 几个例子

问题 1 在 $D=\{(x,y,z)\in\mathbf{R}^3\mid x\geqslant 0,y\geqslant 0,z\geqslant 0\}$ 中确定

$$f:D\subset\mathbf{R}^3\to\mathbf{R},f(x,y,z)=x^m y^n z^p,m,n,p>0$$

在约束条件 $x+y+z-a=0(a>0)$ 下的极值.

解 D 与 $M=\{(x,y,z)\mid x+y+z-a=0\}$ 的交是紧致的. f 在 $D'=D\cap M$ 上非负而且恰好在 D' 的边界上等于零，所以极小值就在边界上取得，而最大值在内部的某处取得.

因为在 $D^0=\{(x,y,z)\mid x>0,y>0,z>0\}$ 中 $f>0$，f 与 $\ln f$ 在同一点处 $(x_0,y_0,z_0)\in D^0$ 取得极值，所以只要研究 $\ln f$ 在 D^0 中在指定约束条件下的极值问题.

记 $g(x,y,z)=x+y+z-a$,按照拉格朗日乘因子法则,要讨论方程组(1.2),即

$$\begin{cases}\ln(f)_x-\lambda g_x\equiv\dfrac{m}{n}-\lambda=0\\ \ln(f)_y-\lambda g_y\equiv\dfrac{n}{y}-\lambda=0\\ \ln(f)_z-\lambda g_z\equiv\dfrac{p}{z}-\lambda=0\\ g\equiv x+y+z-a=0\end{cases}\tag{1.2}$$

所以有

$$\lambda=\frac{m}{x}$$

$$y=\frac{n}{m}x$$

$$z=\frac{p}{m}x$$

$$\left(\frac{n}{m}+\frac{p}{m}+1\right)x=a$$

即

$$x_0=\frac{m}{m+n+p}a$$

$$y_0=\frac{n}{m+n+p}a$$

$$z_0=\frac{p}{m+n+p}a$$

这个位置 (x_0,y_0,z_0) 是 D^0 中唯一要判别的位置.因为另一方面,f 与 $\ln f$ 在 D^0 中确实要取得极大值,而对于这个极大点,拉格朗日方程组确实是满足的(必要条件),所以 (x_0,y_0,z_0) 必定是极大点.

总之,有

$$
\begin{aligned}
0 &= \min\{f(D')\} \leqslant f(x,y,z) \leqslant \max\{f(D')\} \\
&= f(x_0,y_0,z_0) \\
&= \left(\frac{a}{m+n+p}\right)^{m+n+p} \cdot m^m \cdot n^n \cdot p^p
\end{aligned}
$$

其中,极小恰在 D' 的边界上取得,而极大正好在(x_0, y_0, z_0)处取得.

问题2　在平面上,给定3个不在一条直线上的点 P_1, P_2, P_3. 试确定棱锥 $Q - AP_2P_3$,使表面积最小. 其中 Q 对平面的高 h 一定,而以 $P_1P_2P_3$ 为底面的面积 $F_\triangle$ 一定.

提示　如果 P 是三角形上的一个点,x,y,z 是从 P 到边 a,b,c 上的垂线,而 $F_\triangle$ 是三角形的面积($x < 0$,如果以 a 边为准,点 P 不在三角形所在的一侧,等等),有

$$ax + by + cz = 2F_\triangle$$

解　设 $\bar{h}_i(i = 1,2,3)$ 是侧面三角形的高. 因为底面是事先给定的,所以侧面积 M 必须与表面积同时为极小. 我们应用拉格朗日乘因子法,来求

$$M = \frac{1}{2}(a\bar{h}_1 + b\bar{h}_2 + c\bar{h}_3)$$

的极小,其中

$$
\begin{aligned}
\bar{h}_1^2 &= h^2 + x^2 \\
\bar{h}_2^2 &= h^2 + y^2 \\
\bar{h}_3^2 &= h^2 + z^2
\end{aligned}
$$

并且有(约束条件),即

$$ax + by + cz = 2F_\triangle$$

或者

$$g(x,y,z) = ax + by + cz - 2F_\triangle = 0$$

从而,作为必要条件,得到

$$\frac{\partial M}{\partial x}+\lambda\frac{\partial g}{\partial x}=\frac{1}{2}a\cdot\frac{x}{\sqrt{h^2+x^2}}+\lambda a=0$$

$$\frac{\partial M}{\partial y}+\lambda\frac{\partial g}{\partial y}=\frac{1}{2}b\cdot\frac{y}{\sqrt{h^2+y^2}}+\lambda b=0$$

$$\frac{\partial M}{\partial z}+\lambda\frac{\partial g}{\partial z}=\frac{1}{2}c\cdot\frac{z}{\sqrt{h^2+z^2}}+\lambda c=0$$

由这三个方程中的第一个,得到

$$\lambda=-\frac{1}{2}\frac{x}{\sqrt{h^2+x^2}}$$

由第二个,得到

$$\lambda=-\frac{1}{2}\frac{y}{\sqrt{h^2+y^2}}$$

即

$$\frac{x}{\sqrt{h^2+x^2}}=\frac{y}{\sqrt{h^2+y^2}}$$

或者

$$x^2h^2+x^2y^2=y^2h^2+y^2x^2$$

即 $x^2=y^2$ 或者 $x=y$,因为 $x=-y$ 不满足方程

$$\frac{x}{\sqrt{h^2+x^2}}=\frac{y}{\sqrt{h^2+y^2}}$$

同样得到$x=z$. 这就是说,只有当$x=y=z$时,才可能是一个极值点,从而由 $g(x,y,z)=0$,得到

$$x=\frac{2F_{\triangle}}{a+b+c}$$

我们立刻考虑到,没有一个极大点,因为当棱锥足够“斜”时,M 就会很大.

M 是点 Q 的,从而是 x,y,z 的连续函数.

因为一个在紧致的定义域上定义的连续函数(为什么在这里可以限制在紧致集合 D 上),在定义域上取得最大与最小,所以 M 必然当

$$x = y = z = \frac{2F_{\triangle}}{a + b + c}$$

时,成为最小的.

问题 3 在平面上,设有三个顶点 $P_1(a_1, b_1)$, $P_2(a_2, b_2)$ 与 $P_3(a_3, b_3)$,它们不在一条直线上. 试确定平面上一个点 P,使得它离三个点的距离之和 a 为极小.

解 对于平面的任意一点 P, 设 $r_i = \sqrt{(x - a_i)^2 + (y - b_i)^2}$ 是 P 离点 $p_i = (a_i, b_i)$ 的距离,于是问题就在于审查函数

$$a = \sum_{i=1}^{3} r_i = \sum_{i=1}^{3} \sqrt{(x - a_i)^2 + (y - b_i)^2}$$

的极值点. 在这种点处,偏导数必然为零,即

$$\frac{\partial a}{\partial x} = \sum_{i=1}^{3} \frac{x - a_i}{r_i} = \sum_i \cos \theta_i = 0 \qquad (1.3)$$

$$\frac{\partial a}{\partial y} = \sum_{i=1}^{3} \frac{y - b_i}{r_i} = \sum_i \sin \theta_i = 0 \qquad (1.4)$$

θ_i 是介于直线 P_iP 与 x 轴间的夹角($0 \leqslant \theta_i \leqslant \pi$). 作为极值点来审查的是导数不存在的点 P_e 以及由上述条件得到的点 P_e. 因为当 $X = (x, y) \to \infty$ 时,$a(x, y)$ 任意变大,所以,我们可以取适当的 $\mathbf{R} > 0$,而限于在 $\| X \| < \mathbf{R}$ 内的 X 处寻求极小值. 用 $\sin \theta_2$ 乘式(1.3),用 $\cos \theta_2$ 乘式(1.4),然后相减,得到 $\theta_1 - \theta_2 = \theta_2 - \theta_3$,类似地得到 $\theta_2 - \theta_3 = \theta_3 - \theta_1$. 按照这样,介于每两条直线 P_1P_e, P_2P_e, P_3P_e 的夹角都必须等于 $\frac{2\pi}{3}$. 如果在

$\triangle P_1P_2P_3$ 中，只有小于$\frac{2\pi}{3}$的顶角，具有这个性质的一个点 P_e 确实存在的. 在这种情形，P_e 位于三角形的内部. 因为 a 是连续的，a 至少取得一次极小（甚至考虑满足 $\|X\| \leqslant \mathbf{R}$ 的 $X=(x,y)$ 就够了，而 $\mathbf{R}$ 取得足够大）. 因为 a 在$\frac{\mathbf{R}^2}{\{P_1,P_2,P_3\}}$是可微的，指明 $a(P_e) < a(P_1), a(P_2), a(P_3)$ 就够了；按照余弦定理，有 $\cos\theta = -\frac{1}{2}$ 即

$$\begin{aligned}\overline{P_1P_2}^2 &= \overline{P_eP_1}^2 + \overline{P_eP_2}^2 + \overline{P_eP_1}\cdot\overline{P_eP_2}\\ &> \left(\overline{P_eP_2} + \frac{1}{2}\overline{P_eP_1}\right)^2\end{aligned}$$

与

$$\overline{P_1P_3}^2 > \left(\overline{P_eP_3} + \frac{1}{2}\overline{P_eP_1}\right)^2$$

所以

$$\begin{aligned}a(P_1) &= \overline{P_1P_2} + \overline{P_1P_3} > \overline{P_eP_1} + \overline{P_eP_2} + \overline{P_eP_3}\\ &= a(P_e)\end{aligned}$$

类似地得到

$$a(P_2) > a(P_e), a(P_3) > a(P_e)$$

这就是说，P_e 是使 a 有绝对极小的点.

如果有一个角大于或等于$\frac{2\pi}{3}$，那么极小就在一个角度最大的顶点处取得；因为条件 $\theta_1-\theta_2=\theta_2-\theta_3=\theta_3-\theta_1$，在平面上没有一个点能满足（这从几何上各种的情形来看就立刻得到），作为极值点，剩下的只有角点了. 因为在三角形中，最大边对最大角，所以 P_e 必然是最大角的顶点.

极小点在各种情况下都是唯一确定的.

问题4　在约束条件

$$\sum_{v=1}^{n} x_v^2 = 1$$

与

$$\sum_{v=1}^{n} a_v x_v = 0$$

其中 $\sum_{v=1}^{n} a_v^2 > 0$ 下,试确定

$$\left(\sum_{v=1}^{n} b_v x_v\right)^2$$

的极值.

解　记 $X = (x_1, \cdots, x_n)$, $A = (a_1, \cdots, a_n) \neq 0$ 及 $B = (b_1, \cdots, b_n)$,于是,约束条件及函数就成为

$$X^2 = 1, (A \mid X) = 0, \text{其中 } A^2 > 0$$

与

$$f(X) = (B \mid X)^2$$

所以供讨论的 X 与 $A(\neq 0)$ 相垂直,且有长度为1. 这只有当 $\mathbf{R}^n$ 的维数 $n \geqslant 2$ 时才可能. 当 $n = 2$ 时,X 本身按约束条件除去符号是唯一确定的,所以寻找极值,当 $n > 2$ 时才有意义.

有 B 分解成一个与 A 成比例的分量与一个与 A 相垂直的分量

$$B = \lambda A + Y$$

由于$(A \mid Y) = 0$,所以

$$\lambda = \frac{(A \mid B)}{A^2}, A^2 > 0$$

与

$$Y = B - \frac{(A \mid B)}{A^2} A$$

$$\begin{aligned} f(X) &= (B \mid X)^2 = (\lambda A + Y \mid X)^2 \\ &= [\lambda (A \mid X) + (Y \mid X)]^2 = (Y \mid X)^2 \\ &= \left(B - \frac{(A \mid B)}{A^2} A \mid X\right)^2 \end{aligned}$$

这是由于$(A \mid X) = 0$.

因为$(X_1 \mid X_2) = |X_1| |X_2| \cos \varphi$,$f(X)$将是极大的,如果$|\cos \varphi| = 1$,即

$$X = \mu \cdot Y = \mu \left(B - \frac{(A \mid B)}{A^2} A\right)$$

将是极小的,如果$\cos \varphi = 0$,即

$$\left(X \mid B - \frac{(A \mid B)}{A^2} A\right) = (X \mid B) = 0$$

Y等于0的充要条件是B线性相关于A($B = cA$代入).当$B = cA$,有$f(X) = (B \mid X)^2 \equiv 0$.所以特殊情形$B = cA$是不值得注意的.

现在设$Y \neq 0$,于是由$X^2 = 1$,得到

$$\mu = \pm \frac{1}{\left(B - \frac{(A \mid B)}{A^2} A\right)^2}$$

所以

$$X_{\max} = \pm \frac{A^2 B - (A \mid B) A}{\sqrt{(A^2 B - (A \mid B) A)^2}}$$

与

$$f(X_{\max}) = \left| \frac{A^2 B^2 - (A \mid B)^2}{\sqrt{(A^2 B - (A \mid B) A)^2}} \right|^2$$

对于极小点,得到一个线性方程组,即

$$(B \mid X_{\min}) = b_1 x_1 + \cdots + b_n x_n = 0$$

$$(A \mid X_{\min}) = a_1 x_1 + \cdots + a_n x_n = 0$$

由于 $A \neq \lambda B$,解空间是 $n-2$ 维的. 通过规范化条件 $X^2 = 1$,终于得到一个解,它与 $n-3$ 个参数相关,且有 $f(X_{\min}) = 0$.

(当 $n = 3$ 时,除了符号,解 $X_{\min}$ 是唯一的:$X_{\min} = \pm \dfrac{A \times B}{|A \times B|}$).

在上述讨论得到结果的过程中,我们没有用到微分法.

为了要得到

$$X = \pm \frac{A^2 B - (A \mid B)A}{\sqrt{(A^2 B - (A \mid B)A)^2}}$$

我们也可以应用拉格朗日乘因子的必要条件,即

$$\frac{\partial}{\partial x_v}(B \mid X)^2 - \lambda_1 \frac{\partial}{\partial x_v}(A \mid X) - \lambda_2 \frac{\partial}{\partial x_v}(X^2 - 1) = 0, v = 1, \cdots, n$$

用向量来写,得到

$$2(B \mid X)B - \lambda_1 A - \lambda_2(2X) = 0 \tag{1.5}$$

由这个方程,来确定 λ_1, λ_2 与 $(B \mid X)$ 是合适的.

用 X, A 以及 B 对方程(1.5)作数乘,就得到

$$(X \mid 2(B \mid X)B - \lambda_1 A - \lambda_2(2X)) = (X \mid O)$$

即

$$2(B \mid X)^2 - \lambda_1(A \mid X) - \lambda_2(2X^2) = 0$$

其中 $(A \mid X) = 0$,以及 $X^2 = 1$,所以有

$$\lambda_2 = (B \mid X)^2 \tag{1.6}$$

$$(A \mid 2(B \mid X)B - \lambda_1 A - \lambda_2(2X)) = (A \mid O)$$

$$2(B \mid X)(A \mid B) - \lambda_1 A^2 - 2\lambda_2(A \mid X) = 0$$

即有

$$\lambda_1 = \frac{2(B \mid X)(A \mid B)}{A^2} \tag{1.7}$$

$$(B \mid 2(B \mid X)B - \lambda_1 A - \lambda_2(2X)) = (B \mid O)$$

$$2(B \mid X)B^2 - \lambda_1(A \mid B) - 2\lambda_2(B \mid X) = 0$$

所以有

$$2(B \mid X)(B^2 - \lambda_2) = \lambda_1(A \mid B) \tag{1.8}$$

式(1.6),(1.7),(1.8) 是三个未知数 λ_1,λ_2 与 $(B \mid X)$ 的三个方程. 对于$(B \mid X) = 0$,方程(1.6),(1.7),(1.8) 都满足(这就得到前面讲的 $X_{\min}$). 现在设$(B \mid X) \neq 0$. 于是得到

$$\lambda_1 = \pm 2 \frac{(A \mid B)}{A^2} \cdot \sqrt{\frac{A^2B^2 - (A \mid B)^2}{A^2}}$$

$$\lambda_2 = \frac{A^2B^2 - (A \mid B)^2}{A^2}$$

$$(B \mid X) = \mp \sqrt{\frac{A^2B^2 - (A \mid B)^2}{A^2}} \neq 0$$

而且按方程(1.5) 就得到

$$X = \mp \frac{A^2B - (A \mid B)A}{\sqrt{A^2(A^2B^2 - (A \mid B)^2)}}$$

通过规范化,最后就得到 $X_{\max}$.