拉格朗日乘数法中一类问题的解决方法

张圣泽
山东大学（威海）数学与统计学院

摘　要：本文通过构建齐次线性方程组并利用克拉默法则求出二次齐次函数在特定约束条件下的极值、最值。此方法不仅避免了求解复杂方程组，并且避免了增值的讨论问题。

关键词：条件极值；二次齐次函数；拉格朗日乘数法

【DOI】10.12252/j.issn.2096-6288.2024.12.198

引言

由拉格朗日乘数法所得的拉格朗日方程组往往由于变量较多、方程复杂等原因难以求解；并且在求解方程组的过程中，往往会出现增值或是需要通过分类讨论等手段解决问题：这些原因都使得大部分条件极值问题的解答过程十分复杂。

在数学分析教材中，总能看到某些题目可以通过特定的变量代换将拉格朗日方程组变形为关于自变量的齐次线性方程组，再利用克拉默法则对该方程组的系数矩阵进行分析，从而快速求解。这种特殊的方法不仅能够做到不解出拉格朗日方程组而求得函数的条件极值，而且避免了对于方程组的增值的讨论；但是教材中往往不会明确提出这种方法的具体适用范围。本文的任务是明确此方法的适用范围并规范此方法的一般解题步骤。

一、主要结论及证明

设 $X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ 为 n 维列向量；

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\quad B=\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&&\vdots\\b_{n1}&b_{n2}&\cdots&b_{nn}\end{pmatrix},$$

为 $n\times n$ 对称矩阵，即 $a_{ij}=a_{ji},b_{ij}=b_{ji},i,j=1,2,\cdots,n$；$\boldsymbol{\xi}=\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}$ 为 n 维列向量．则有

命题　求二次齐次函数 $f(X)=X^TAX$ 在条件

$$\begin{cases}X^TBX=k\\\xi^TX=0\end{cases}$$

$k\neq 0$ 下的条件极值可以使用上述方法．

证明　(1) 当 $\boldsymbol{\xi}=0$ 时，实际上约束条件只有

$$X^TBX=k \tag{1.2}$$

作拉格朗日函数

$$L=X^TAX-\lambda\left(X^TBX-k\right),$$

得拉格朗日方程组

$$\begin{cases}2AX-2\lambda BX=0,\\X^TBX=k.\end{cases}$$

将第一式左乘 X^T 得 $2X^TAX-2\lambda X^TBX=0$,

即 $f=\lambda k$，即函数 f 的条件极值为 λ 的 k 倍。

考虑齐次线性方程组 $AX-\lambda BX=0$，它的系数矩阵为 $A-\lambda B$。

容易得到此方程组的系数行列式不恒等于 0，且仅含参数 λ。由于 $x=y=z=0$ 不满足 $X^TBX=k$，所以方程组存在非零解，系数行列式等于 0。由此求出 λ 可能取到的值，它们的 k 倍即为函数 f 在约束条件 (1.2) 下的极值。

(2) 当 $\boldsymbol{\xi}\neq 0$ 时，作拉格朗日函数

$$L=X^TAX-\lambda\left(X^TBX-k\right)-\mu\left(\boldsymbol{\xi}^TX\right),$$

得拉格朗日方程组

$$\begin{cases}2AX-2\lambda BX-\mu\boldsymbol{\xi}=0,\\X^TBX=k,\\\boldsymbol{\xi}^TX=0.\end{cases}$$

将第一式左乘 X^T 得

$$2X^TAX-2\lambda X^TBX-\mu\boldsymbol{\xi}^TX=0,$$

即 $f=\lambda k$，即函数 f 的条件极值为 λ 的 k 倍。

将第一式左乘 $\boldsymbol{\xi}^T$ 得

$$2\boldsymbol{\xi}^T(A-\lambda B)X-\mu\boldsymbol{\xi}^T\boldsymbol{\xi}=0,$$

即 $\mu=\dfrac{2\boldsymbol{\xi}^T(A-\lambda B)X}{\boldsymbol{\xi}^T\boldsymbol{\xi}}$.

作线性替换

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}\xi_1&\xi_2&\cdots&\xi_n\\0&1&&\\\vdots&&\ddots&\\0&&&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

又由第三式有

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}0\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}0&0&\cdots&0\\0&1&&\\\vdots&&\ddots&\\0&&&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

结合以上两式可得

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}0&-\frac{\xi_2}{\xi_1}&\cdots&-\frac{\xi_n}{\xi_1}\\0&1&&\\\vdots&&\ddots&\\0&&&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

令， $$C=\begin{pmatrix}0&-\frac{\xi_2}{\xi_1}&\cdots&-\frac{\xi_n}{\xi_1}\\0&1&&\\\vdots&&\ddots&\\0&&&1\end{pmatrix}$$

设$\boldsymbol{\xi}^T\boldsymbol{\xi}=\xi$，整理得系数矩阵为

$$\begin{pmatrix}
a_{11}-\lambda b_{11} & a_{12}-\lambda b_{12}-\frac{\xi_1}{\xi}\left(-\frac{\xi_2}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{k2}-\lambda b_{k2}\right)\right) & \cdots & a_{1n}-\lambda b_{1n}-\frac{\xi_1}{\xi}\left(-\frac{\xi_n}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{kn}-\lambda b_{kn}\right)\right)\\
a_{21}-\lambda b_{21} & a_{22}-\lambda b_{22}-\frac{\xi_2}{\xi}\left(-\frac{\xi_2}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{k2}-\lambda b_{k2}\right)\right) & \cdots & a_{2n}-\lambda b_{2n}-\frac{\xi_2}{\xi}\left(-\frac{\xi_n}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{kn}-\lambda b_{kn}\right)\right)\\
\vdots & \vdots & & \vdots\\
a_{n1}-\lambda b_{n1} & a_{n2}-\lambda b_{n2}-\frac{\xi_n}{\xi}\left(-\frac{\xi_2}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{k2}-\lambda b_{k2}\right)\right) & \cdots & a_{nn}-\lambda b_{nn}-\frac{\xi_n}{\xi}\left(-\frac{\xi_n}{\xi_1}\sum_{k=1}^{n}\xi_k\left(a_{k1}-\lambda b_{k1}\right)+\sum_{k=1}^{n}\xi_k\left(a_{kn}-\lambda b_{kn}\right)\right)
\end{pmatrix}$$

(1.4)

容易得到此方程组的系数行列式不恒等于0，且仅含参数λ。由于$x=y=z=0$不满足条件$X^TBX=k$，所以方程组存在非零解，系数行列式等于0。由此求出λ可能取到的值，它们的k倍即为函数f在约束条件（1.1）下的极值。

二、例题

首先，我们通过对一个题的两种解法进行对比来感受命题中所给出的方法的优越性。

例2 1[1] 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。

解 因为椭圆的中心在原点，所以仅需求椭圆上的点到原点的距离的最值。问题即求

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

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

下的最大值和最小值。

作拉格朗日函数

$L(x,y,z,\lambda,\mu)=x^2+y^2+z^2-\lambda\left(x^2+y^2+4z^2-1\right)-\mu(x+y+z)$，得到拉格朗日方程组

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

解法一 由上述方程组的前四式得到$3\mu(1-3\lambda)=0$，因此$\mu=0$或$1-3\lambda=0$。

（1）当$\mu=0$时，方程组前三式相加得$6\lambda z=0$。又$\lambda\neq0$，故$z=0$。代入$x+y+z=0$，$x^2+y^2+4z^2=1$得到(x,y,z)的两组解

$$\left(\frac{1}{\sqrt2},-\frac{1}{\sqrt2},0\right),\left(-\frac{1}{\sqrt2},\frac{1}{\sqrt2},0\right),$$

f在这两点的值都是1。

（2）当$1-3\lambda=0$时，由方程组前三式及第四式得到(x,y,z)的两组解

$$\left(\frac{\sqrt2}{6},\frac{\sqrt2}{6},-\frac{\sqrt2}{3}\right),\left(-\frac{\sqrt2}{6},-\frac{\sqrt2}{6},\frac{\sqrt2}{3}\right),$$

f在这两点的值都是$\frac{1}{3}$。

由于椭圆的长轴与短轴必存在，所以f在条件$\begin{cases}x+y+z=0,\\x^2+y^2+4z^2=1\end{cases}$下的最大值和最小值分别为1和$\frac{1}{3}$。

所以椭圆面积为$\frac{\pi}{\sqrt3}$。

解法二 显然，此题满足命题中的条件，属于其中（2）的情形，可以使用命题中的方法。

将方程组第一式乘x，第二式乘y，第三式乘z，得$2\left(x^2+y^2+z^2\right)-2\lambda\left(x^2+y^2+4z^2\right)-\mu(x+y+z)=0$，再由$x+y+z=0,x^2+y^2+4z^2=1$得$f(x,y,z)=\lambda$.

这说明函数$f(x,y,z)$的条件极值包含在方程组关于的解中。将方程组前三式相加，代入第四式得$\mu=-2\lambda z$，

代入前三式得 $\begin{cases}2(1-\lambda)x+2\lambda z=0,\\2(1-\lambda)y+2\lambda z=0,\\2(1-3\lambda)z=0.\end{cases}$

由$x^2+y^2+4z^2=1$可知上述方程组有非零解，因此系数行列式等于零，即 $(1-\lambda)^2(1-3\lambda)=0$,

解得$\lambda=1$或$\lambda=\frac{1}{3}$．由于连续函数f在紧集$\left\{(x,y)\,\middle|\,x+y+z=0,x^2+y^2+4z^2=1\right\}$上必可以取

则有

$$\mu=\frac{2\boldsymbol{\xi}^T(A-\lambda B)CX}{\boldsymbol{\xi}^T\boldsymbol{\xi}}.$$

将（1.3）代入第一式得关于$x_1,x_2,\cdots,x_n$的齐次线性方程组

$$(A-\lambda B)X-\frac{\boldsymbol{\xi}^T(A-\lambda B)CX}{\boldsymbol{\xi}^T\boldsymbol{\xi}}\boldsymbol{\xi}=0.$$

到最大值和最小值，因此f在约束条件下的最大值为1，最小值为$\frac{1}{3}$。所以椭圆面积为$\frac{\pi}{\sqrt{3}}$。

正如我们在引言中所说的：在解法一中，使用一般方法求解拉格朗日方程组时，总会不可避免地出现需要分类讨论的情况，这无疑使题目的求解变得复杂。而解法二不仅避免了分类讨论，其解答过程也更程序化，使得问题变得易于解决。

我们当然也可以直接将参数值代入（1。4）式进行计算，得出 λ 的值；但是由于（1。3）式难以记忆，且这种方法的解题过程并不复杂，因而按照证明过程解题更加容易。

例 2.2[2] 试证明：二次型

$$f(x,y,z)=Ax^2+By^2+Cz^2+2Dyz+2Ezx+2Fxy$$

在单位球面$x^2+y^2+z^2=1$

上的最大值和最小值恰好是矩阵

$$\Phi=\begin{pmatrix}A & F & E\\ F & B & D\\ E & D & C\end{pmatrix}$$

的最大特征值和最小特征值 .

证明　显然，此题满足命题中的条件，属于其中（1）的情形，可以使用命题中的方法。

令

$$L(x,y,z,\lambda)=f(x,y,z)-\lambda\left(x^2+y^2+z^2-1\right),$$

得拉格朗日方程组

$$\begin{cases}\dfrac{\partial L}{\partial x}=2(A-\lambda)x+2Fy+2Ez=0,\\ \dfrac{\partial L}{\partial y}=2Fx+2(b-\lambda)y+2Dz=0,\\ \dfrac{\partial L}{\partial z}=2Ex+2Dy+2(C-\lambda)z=0,\\ x^2+y^2+z^2=1.\end{cases}$$

由$x\frac{\partial L}{\partial x}+y\frac{\partial L}{\partial y}+z\frac{\partial L}{\partial z}$得$f(x,y,z)=\lambda,$

即目标函数的条件极值即为拉格朗日方程组关于λ的解。

方程组前三式所组成的齐次线性方程组有非零解，因而系数行列式等于 0，即

$$2\left|\begin{matrix}A-\lambda & F & E\\ F & B-\lambda & D\\ E & D & C-\lambda\end{matrix}\right|=0,$$

即λ是矩阵Φ特征值。又因为满足约束条件的点集是连通紧集，且目标函数连续，所以必有最大值和最小值。所以最大值和最小值恰好是矩阵Φ的最大特征值和最小特征值。

事实上，此题相当于命题中A=Φ，B=E，$\boldsymbol{\xi}=0$的情况 . 此时A−λB=Φ−λE，自然λ是矩阵Φ的特征值。

例 2.3[1] 求二次型$\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}x_ix_j\,(a_{ij}=a_{ji})$在$n$维单位球面

$$\left\{(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n\,\middle|\,\sum_{k=1}^{n}x_k^2=1\right\}$$

上的最大值和最小值。

解　显然，此题满足命题中的条件，属于其中（1）的情形，可以使用命题中的方法。

令

$$L(x_1,x_2,\cdots,x_n,\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-\lambda\left(\sum_{k=1}^{n}x_k^2-1\right),$$

得拉格朗日方程组

$$\begin{cases}\dfrac{\partial L}{\partial x_k}=2\sum\limits_{i=1}^{n}a_{ik}x_i-2\lambda x_k=0,k=1,2,\cdots,n,\\ x_1^2+x_2^2+\cdots+x_n^2=1.\end{cases}$$

由

$$\sum_{k=1}^{n}x_k\frac{\partial L}{\partial x_k}=2\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-\lambda\sum_{k=1}^{n}x_k^2\right)=0$$

可知

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\lambda,$$

即目标函数的条件极值即为拉格朗日方程组关于λ的解。

记A $=(a_{ij})$，因为方程组前三式所组成的齐次线性方程组有非零解，因而系数行列式$|$A $-\lambda$E$|=0$，即λ是矩阵A特征值。由于A为实对称矩阵，所以特征值都是实数，将它们按照大小排序为$\lambda_1\le\lambda_2\le\cdots\le\lambda_n$，可得$f_{max}=\lambda_n, f_{min}=\lambda_1.$

此题和例 2.2 相同，只不过拓展到了n维。这也说明这种方法适用于n维的情况。

结语

本文就二次型函数在一类约束条件下的条件极值问题的简便解法的适用性和解题步骤做了严格的证明和规范。虽然证明过程中公式较为复杂，但通过实例我们发现，在题目中并不会出现过于复杂的情况，并且这种简便解法的解题过程比较简单，因此在实际解题中，按照证明过程进行求解是更好的选择。

参考文献

[1] 陈纪修，於崇华，金路．数学分析（第 3 版）[M]．北京：高等教育出版社，2019.

[2] 华东师范大学数学科学学院．数学分析（第 5 版）[M]．北京：高等教育出版社，2019.

作者简介：张圣泽，2004.08，男，汉，山东省德州市，本科在读，研究方向：数学与应用数学。