深度解析拉格朗日乘子法，让你成为高手

数学达人上官正申
信息技术行业 从业人员

又是一个”王“的男人——欧拉的徒弟拉格朗日，**他是欧洲数学界之光！法国最杰出的数学大师！数学分析的开拓者！**他说：”我把数学看作是一件有意思的工作，而不是想为自己建立什么纪念碑。“一腔热血、赤子之心！

什么是拉格朗日乘子法？它是用来求解条件极值的工具。今天从9个维度3000字教你真正读懂它！

本文对理论的研究与讨论**相当深邃**，大部分内容在任何教科书上都很难找到。文章会讨论**拉格朗日乘子法及其直观的数学含义**，以及**相应约束条件下取极值的必要条件要求二阶偏导数矩阵需要满足的条件，以及相应基矢及坐标系的重新构建。**

请记得收藏反复阅读，希望能帮助你在这方面的理解达到出神入化的境界。

---

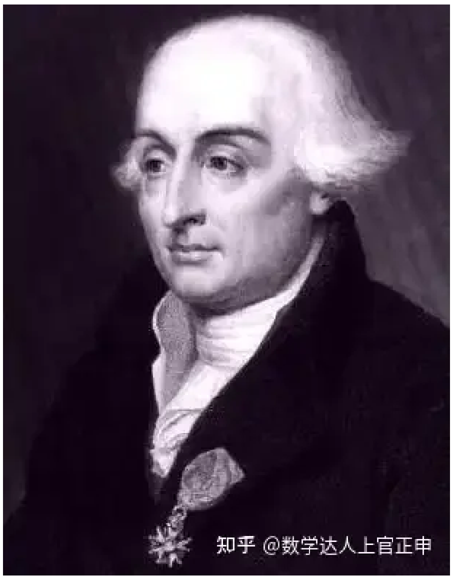
1. **数学家拉格朗日简介**

---

1. 数学家拉格朗日简介

约瑟夫·路易·拉格朗日（法语：Joseph-Louis Lagrange，1736年1月25日－1813年4月10日），出生时名为**朱塞佩·路易吉·拉格朗吉亚**或**朱塞佩·洛德维科·德·拉·格朗日·图尼尔**，是一位法国籍意大利裔数学家和天文学家。拉格朗日曾为普鲁士的腓特烈大帝在柏林工作了20年，被腓特烈大帝称做“欧洲最伟大的数学家”，后受法国国王路易十六的邀请定居巴黎直至去世。拉格朗日一生才华横溢，在数学、物理和天文等领域做出了很多重大的贡献。他的成就包括著名的拉格朗日中值定理，创立了拉格朗日力学等等。

拉格朗日是18世纪一位十分重要的科学家，在数学、力学和天文学三个学科中都有历史性的重大贡献，但他主要是数学家。他最突出的贡献是在把数学分析的基础脱离几何与力学方面起了决定性的作用，使数学的独立性更为清楚，而不仅是其他学科的工具。同时在使天文学力学化、力学分析化上也起了历史性作用，促使力学和天文学（天体力学）更深入发展。在他的时代，分析学等分支刚刚起步，欠缺严密性和标准形式，但这不足以妨碍他取得大量的成果。



拉格朗日肖像

2. 拉格朗日数乘法简介

拉格朗日乘数法（英語：Lagrange multiplier，以数学家约瑟夫·拉格朗日命名），在数学中的最优化问题中，是一种寻找多元函数在其变量受到一个或多个条件的约束时的局部极值的方法。这种方法可以将一个有n个变量与k个约束条件的最优化问题转换为一个解有$n+k$个变量的方程组的解的问题。这种方法中引入了一个或一组新的未知数，即**拉格朗日乘数**，又称**拉格朗日乘子**，或**拉氏乘子**，它们是在转换后的方程，即约束方程中作为梯度（gradient）的线性组合中各个向量的系数。

3. 多元函数求导及梯度的几何意义

在讨论拉格朗日乘子法之前，我们有必要先讲解一下梯度相关的知识。我们知道，对于一个多元可微函数

$$f(x_1, x_2, \cdots, x_n)$$

有

$$\mathrm{d}f(x_1, x_2, \cdots, x_n) = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\mathrm{d}x_2 + \cdots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathrm{d}x_i\ ,$$

我们称向量

$$\nabla f = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n})^T$$

为**梯度向量。**利用**柯栖不等式**，有

$$\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathrm{d}x_i \leq \sqrt{\sum_{i=1}^{n}(\frac{\partial f}{\partial x_i})^2}\sqrt{\sum_{i=1}^{n}\mathrm{d}x_i^2}\ ,$$

当且仅当

$$\frac{\partial f}{\partial x_1}/\mathrm{d}x_1 = \frac{\partial f}{\partial x_2}/\mathrm{d}x_2 = \cdots = \frac{\partial f}{\partial x_n}/\mathrm{d}x_n$$

时等号成立，上式请允许零比零型的分式出现。

其实柯栖不等式的数学意义非常直观，其实就是向量

$$\nabla f = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n})^T$$

与向量

$$\mathrm{d}\boldsymbol{x} = (\mathrm{d}x_1, \mathrm{d}x_2, \cdots, \mathrm{d}x_n)$$

的内积（即一个向量模长乘以另一个向量在该向量上的投影）小于等于两个向量的模长相乘，当且仅当两个向量共线时等号成立。实际上我们可以定义两个向量夹角余弦值

$$\cos\theta = \frac{\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathrm{d}x_i}{\sqrt{\sum_{i=1}^{n}(\frac{\partial f}{\partial x_i})^2}\sqrt{\sum_{i=1}^{n}\mathrm{d}x_i^2}}$$

即

$$\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathrm{d}x_i = \sqrt{\sum_{i=1}^{n}(\frac{\partial f}{\partial x_i})^2}\sqrt{\sum_{i=1}^{n}\mathrm{d}x_i^2}\cos\theta$$

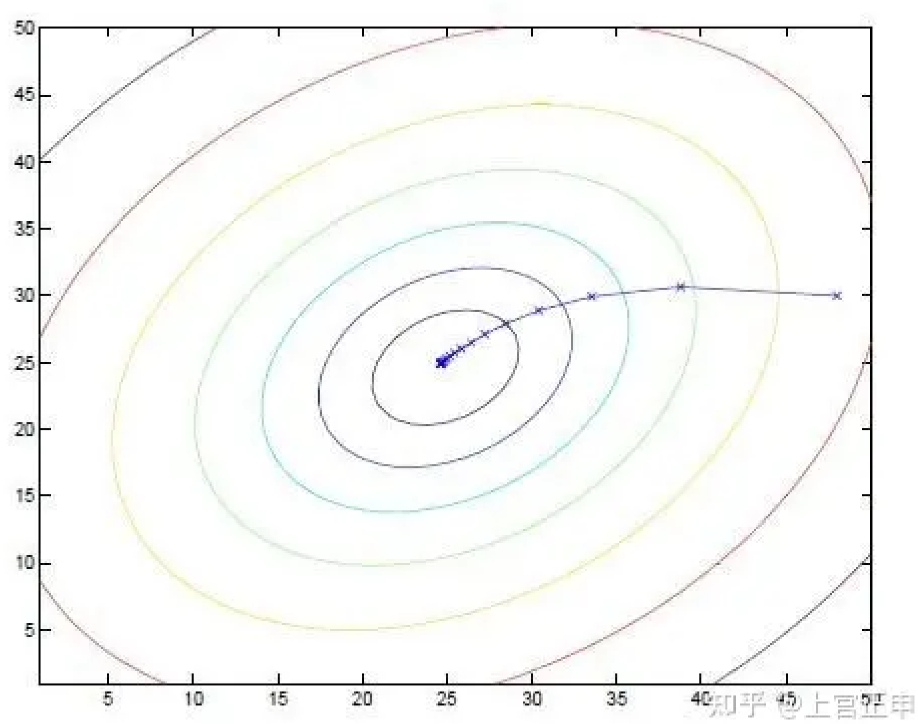
由柯栖不等式，我们知道，在自变量变化一定长度时，即向量

$$\mathrm{d}\boldsymbol{x} = (\mathrm{d}x_1, \mathrm{d}x_2, \cdots, \mathrm{d}x_n)^T$$

的模长

$$\sqrt{\sum_{i=1}^{n}\mathrm{d}x_i^2}$$

一定时，两个向量的夹角越小，则 $\mathrm{d}f$ 的值越大。这说明函数 f 沿梯度的方向增加最快，沿梯度的反方向减小最快，沿着与梯度垂直的方向变化率为零，因此我们立即可以得到**梯度与等值“面”垂直的结论**。注意这个等值面是超曲面，维度为 $n-1$, 并不是指二维曲面。等值面越密集的地方梯度的模长越大，越稀疏的地方梯度的模长越小。



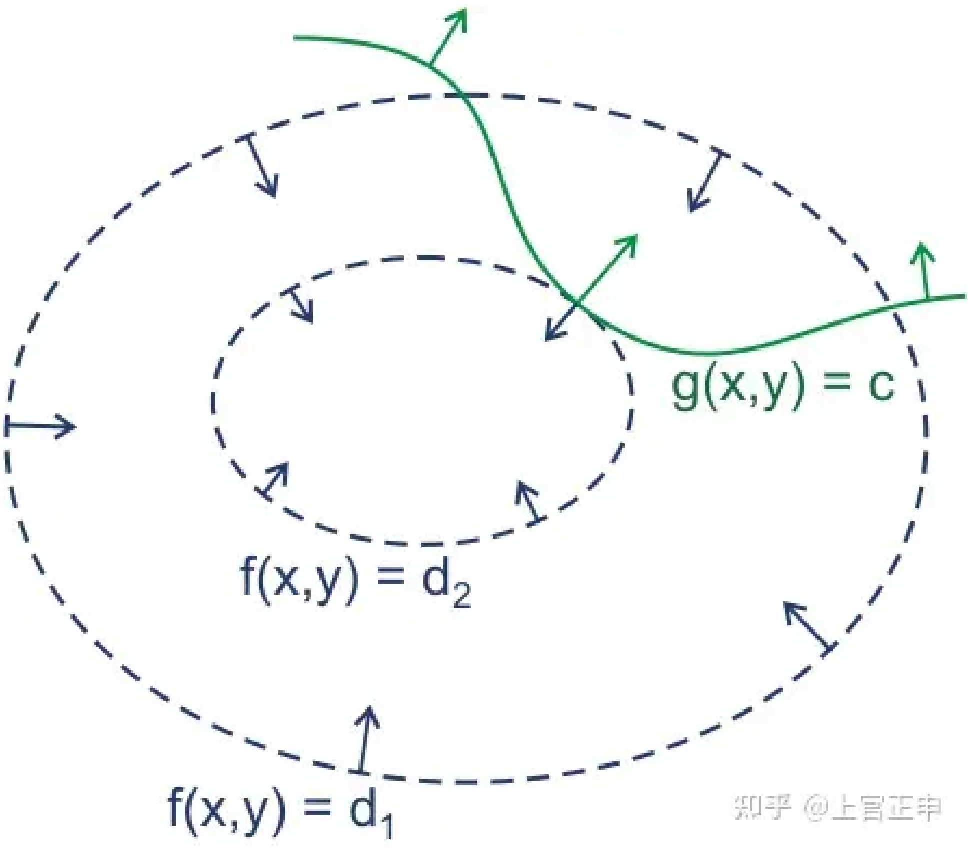
等值面图示

4. 无约束问题函数取极值一阶导数须满足的条件

没有任何约束条件的情况下，一个函数取极值的必要条件是梯度为零。因为梯度不为零的点，函数在各个方向上的变化率不可能全部为零，因此它不可能是极值。额外说一下，梯度为零并不是极值的充分条件，一个函数在某一点的梯度为零，也不一定是极值，例如 x^3 在 $x=0$ 处的一阶导数（梯度）为零，但是却不是极值。一阶导数为零我们不能直接确定是否是极值，我们还需要继续看二阶导数，如果二阶导数也为零，我们还要看三阶导数……

5. 有约束问题函数取极值一阶导数须满足的条件

对于有约束条件的极值的必要条件，我们远不需要梯度为零这样强烈，因为存在约束面导致自变量的变化并非完全自由的，我们只需要在自变量可以自由运动的各个方向上，函数的变化率为零即可，也就是说我们保证约束面在该点是等值面，即梯度与约束面垂直即可。如下图所示，绿线为约束面，梯度与约束面垂直，由于变量只能在约束面内即与梯度垂直的方向上运动，因此 f 在该点有可能取到极值。



约束条件下的极值问题

例如当我们添加约束条件，

$$
\begin{aligned}
&g_1(x_1,x_2,\cdots,x_n)=0,\\
&g_2(x_1,x_2,\cdots,x_n)=0,\\
&\vdots,\\
&g_m(x_1,x_2,\cdots,x_n)=0
\end{aligned}
$$

我们已经知道与等值面垂直的向量是梯度向量，因此与约束面 $g_i(x_1,x_2,\cdots,x_n)=0$ 垂直的向量是

$$\nabla g_i=(\frac{\partial g_i}{\partial x_1},\frac{\partial g_i}{\partial x_2},\cdots,\frac{\partial g_i}{\partial x_n})^T$$

最终的约束面是这 m 个约束面的交集，各个约束面的梯度的线性叠加

$$\sum_{i=1}^{m}\lambda_i\nabla g_i(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{m}\lambda_i(\frac{\partial g_i}{\partial x_1},\frac{\partial g_i}{\partial x_2},\cdots,\frac{\partial g_i}{\partial x_n})^T$$

均与最终的约束面垂直。因此函数 $f(x_1,x_2,\cdots,x_n)$ 取极值的必要条件便弱化为

$$\nabla f=\sum_{i=1}^{m}\lambda_i\nabla g_i$$

将此式写成分量形式并配上约束条件便是

$$\begin{cases} \frac{\partial f}{\partial x_1} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_1} = 0, \\ \frac{\partial f}{\partial x_2} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_2} = 0, \\ \qquad\qquad \vdots, \\ \frac{\partial f}{\partial x_n} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_n} = 0, \\ g_1(x_1, x_2, \cdots, x_n) = 0, \\ g_2(x_1, x_2, \cdots, x_n) = 0, \\ \qquad\qquad \vdots, \\ g_n(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

如果我们引入拉格朗日函数

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

则上述方程组可以重新被表述为

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0, \\ \frac{\partial L}{\partial x_2} = 0, \\ \quad \vdots, \\ \frac{\partial L}{\partial x_n} = 0, \\ \frac{\partial L}{\partial \lambda_1} = 0, \\ \frac{\partial L}{\partial \lambda_2} = 0, \\ \quad \vdots, \\ \frac{\partial L}{\partial \lambda_m} = 0. \end{cases}$$

6. 二阶偏导矩阵的引入

下面我们再来探讨一下当一阶导数向量即梯度为零时，取极值的必要条件要求二阶偏导矩阵满足什么样的条件。我们将函数展开到二阶

$$\Delta f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i + \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k} \Delta x_j \Delta x_k$$

写成向量形式

$$\Delta f(x_1, x_2, \cdots, x_n) = \nabla f \cdot \Delta \boldsymbol{x} + \Delta \boldsymbol{x}^T H \Delta \boldsymbol{x}$$

其中

$$H=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}\\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}.$$

7. 无约束问题函数取极值一阶导数须满足的条件

显然当 $\nabla f=\mathbf{0}$ 时，

$$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^T$$

是极值的必要条件是 H **半正定（极小值）或半负定（极大值）**。当然这只是必要条件，并非充分条件，因为即便满足半正定或半负定，一旦有某个特征值为零，即行列式为零，我们还需要继续考查 f 的更高阶导数。

8. 有约束问题函数取极值二阶导数须满足的条件

对于有约束的问题，当梯度为零时，函数取极值的必要条件对于二阶导数的要求没有这样强烈。因为自变量只能在最终的约束面内活动，因此我们只需要在最终约束面的线性子空间里满足正定或半正定即可。如果 m 个约束条件在该点线性无关，则最终在该点的约束面的维度是 $n-m$，如果线性相关，最终的维度会比这个数值大，我们不妨假定线性无关，线性相关时也是同理。我们在该约束面内选定一组基矢

$$\hat{\boldsymbol{e}}'_1,\hat{\boldsymbol{e}}'_2,\cdots,\hat{\boldsymbol{e}}'_{n-m}$$

将该向量组继续扩充，可以扩充为全空间上的一组基矢

$$\hat{\boldsymbol{e}}'_1,\hat{\boldsymbol{e}}'_2,\cdots,\hat{\boldsymbol{e}}'_n$$

假定原基矢为

$$\hat{\boldsymbol{e}}_1,\hat{\boldsymbol{e}}_2,\cdots,\hat{\boldsymbol{e}}_n$$

假定从原基矢到新基矢的过渡矩阵为 S，即

$$\begin{pmatrix}\hat{\boldsymbol{e}}'_1 & \hat{\boldsymbol{e}}'_2 & \cdots & \hat{\boldsymbol{e}}'_n\end{pmatrix}=\begin{pmatrix}\hat{\boldsymbol{e}}_1 & \hat{\boldsymbol{e}}_2 & \cdots & \hat{\boldsymbol{e}}_n\end{pmatrix}S,$$

写成矩阵形式为

$$\begin{pmatrix}\hat{\boldsymbol{e}}'_1 & \hat{\boldsymbol{e}}'_2 & \cdots & \hat{\boldsymbol{e}}'_n\end{pmatrix}=\begin{pmatrix}\hat{\boldsymbol{e}}_1 & \hat{\boldsymbol{e}}_2 & \cdots & \hat{\boldsymbol{e}}_n\end{pmatrix}\begin{pmatrix}s_{11} & s_{12} & \cdots & s_{1n}\\ s_{21} & s_{22} & \cdots & s_{2n}\\ \vdots & \vdots & & \vdots\\ s_{n1} & s_{n2} & \cdots & s_{nn}\end{pmatrix},$$

则

$$\Delta\boldsymbol{x}=S\Delta\boldsymbol{x}'$$

写成矩阵形式为

$$\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} \Delta x'_1 \\ \Delta x'_2 \\ \vdots \\ \Delta x'_n \end{pmatrix}$$

则

$$\Delta \boldsymbol{x}^T H \Delta \boldsymbol{x} = \Delta \boldsymbol{x}'^T S^T H S \Delta \boldsymbol{x}'$$

则在新的基矢下的二阶偏导矩阵为

$$H' = S^T H S = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

其中 A 为 $(n-m)\times(n-m)$ 阶对称矩阵，B 为 $(n-m)\times m$ 阶矩阵，C 为 $m\times m$ 阶矩阵。**则取极值的必要条件只要求 A 半正定或半负定，并不要求 H' 半正定或半负定。**

9. 约束子空间基矢的重新构建方法

下面我们再来讨论一下我们在特定的点上如何找到 $\Delta \boldsymbol{x}'$.对约束条件求偏导可得

$$\begin{cases} \sum_{i=1}^{n} \frac{\partial g_1}{\partial x_i} \Delta x_i = 0, \\ \sum_{i=1}^{n} \frac{\partial g_2}{\partial x_i} \Delta x_i = 0, \\ \quad \vdots, \\ \sum_{i=1}^{n} \frac{\partial g_m}{\partial x_i} \Delta x_i = 0. \end{cases}$$

写成矩阵形式

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

由于这 m 个约束可能线性相关，所以矩阵的列秩可能 $l \le m$, 通过解这个线性方程组，我们能够找到 $n-l$ 个自由变量，不妨假定这 $n-l$ 个独立自由变量为

$$x_{i_{l+1}}, x_{i_{l+2}}, \cdots, x_{i_n}$$

这里 $i_1, i_2, \cdots, i_n$ 是 $1, 2, \cdots, n$ 的一个重排列，我们不妨选取

$$\begin{pmatrix} \hat{\boldsymbol{e}}_{i_1} & \hat{\boldsymbol{e}}_{i_2} & \cdots & \hat{\boldsymbol{e}}_{i_n} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{e}}_1 & \hat{\boldsymbol{e}}_2 & \cdots & \hat{\boldsymbol{e}}_n \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$$

为新基矢，则旧坐标与新坐标的变换关系为

$$\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix} \begin{pmatrix} \Delta x_{i_1} \\ \Delta x_{i_2} \\ \vdots \\ \Delta x_{i_n} \end{pmatrix}$$

其中

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$$

是置换矩阵（每一行只有一个元素是 1 ，其余元素均是 0；每一列也是如此）。解上述方程我们最终可以解出

$$\begin{cases} \Delta x_{i_1} = \sum\limits_{j=l+1}^{n} k_{1j} \Delta x_{i_j}, \\ \Delta x_{i_2} = \sum\limits_{j=l+1}^{n} k_{2j} \Delta x_{i_j}, \\ \qquad\qquad \vdots, \\ \Delta x_{i_l} = \sum\limits_{j=l+1}^{n} k_{lj} \Delta x_{i_j}. \end{cases}$$

可见 $\Delta x_{i_1}, \Delta x_{i_2}, \cdots, \Delta x_{i_l}$ 并非独立变量，而是被 $\Delta x_{i_{l+1}}, \Delta x_{i_{l+2}}, \cdots, \Delta x_{i_n}$ 唯一确定。则我们可以选定向量

$$\begin{cases} (k_{1,l+1}, k_{2,l+1}, \cdots, k_{l,l+1}, 1, 0, 0, \cdots, 0), \\ (k_{1,l+2}, k_{2,l+2}, \cdots, k_{l,l+2}, 0, 1, 0, \cdots, 0), \\ \qquad\qquad\qquad \vdots, \\ (k_{1,n}, \quad k_{2,n}, \quad \cdots, \quad k_{l,n}, \quad 0, 0, 0, \cdots, 1). \end{cases}$$

为新的基矢

$$\begin{cases} \hat{e}'_1 \\ \hat{e}'_2, \\ \vdots, \\ \hat{e}'_{n-l}. \end{cases}$$

的方向。如此我们便找到了前文中满足条件的

$$\hat{e}'_1, \hat{e}'_2, \cdots, \hat{e}'_{n-l}$$

我们将它继续扩充为一组基矢，即

$$\begin{cases} (k_{1,l+1}, k_{2,l+1}, \cdots, k_{l,l+1}, 1, 0, 0, \cdots, 0), \\ (k_{1,l+2}, k_{2,l+2}, \cdots, k_{l,l+2}, 0, 1, 0, \cdots, 0), \\ \qquad\qquad\qquad \vdots, \\ (k_{1,n}, \quad k_{2,n}, \quad \cdots, \quad k_{l,n}, \quad 0, 0, 0, \cdots, 1), \\ (k_{1,1}, \quad k_{2,1}, \quad \cdots, \quad k_{l,1}, \quad 0, 0, 0, \cdots, 0), \\ (k_{1,2}, \quad k_{2,2}, \quad \cdots, \quad k_{l,2}, \quad 0, 0, 0, \cdots, 0), \\ \qquad\qquad\qquad \vdots, \\ (k_{1,l}, \quad k_{2,l}, \quad \cdots, \quad k_{l,l}, \quad 0, 0, 0, \cdots, 0). \end{cases}$$

即

$$\begin{pmatrix} \hat{\boldsymbol{e}}_1' & \hat{\boldsymbol{e}}_2' & \cdots & \hat{\boldsymbol{e}}_n' \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{e}}_{i_1} & \hat{\boldsymbol{e}}_{i_2} & \cdots & \hat{\boldsymbol{e}}_{i_n} \end{pmatrix} \begin{pmatrix} k_{1,l+1} & k_{1,l+2} & \cdots & k_{1,n} & k_{1,1} \\ k_{2,l+1} & k_{2,l+2} & \cdots & k_{2,n} & k_{2,1} \\ \vdots & \vdots & & \vdots & \vdots \\ k_{l,l+1} & k_{l,l+2} & \cdots & k_{l,n} & k_{l,1} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

则坐标的变换关系为

$$\begin{pmatrix} \Delta x_{i_1} \\ \Delta x_{i_2} \\ \vdots \\ \Delta x_{i_n} \end{pmatrix} = \begin{pmatrix} k_{1,l+1} & k_{1,l+2} & \cdots & k_{1,n} & k_{1,1} & k_{1,2} & \cdots & k_{1,l} \\ k_{2,l+1} & k_{2,l+2} & \cdots & k_{2,n} & k_{2,1} & k_{2,2} & \cdots & k_{2,l} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ k_{l,l+1} & k_{l,l+2} & \cdots & k_{l,n} & k_{l,1} & k_{l,2} & \cdots & k_{l,l} \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \Delta x_1' \\ \Delta x_2' \\ \vdots \\ \Delta x_n' \end{pmatrix}$$

令

$$K = \begin{pmatrix} k_{1,l+1} & k_{1,l+2} & \cdots & k_{1,n} & k_{1,1} & k_{1,2} & \cdots & k_{1,l} \\ k_{2,l+1} & k_{2,l+2} & \cdots & k_{2,n} & k_{2,1} & k_{2,2} & \cdots & k_{2,l} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ k_{l,l+1} & k_{l,l+2} & \cdots & k_{l,n} & k_{l,1} & k_{l,2} & \cdots & k_{l,l} \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

则有

$$\Delta \boldsymbol{x} = T \begin{pmatrix} \Delta x_{i_1} \\ \Delta x_{i_2} \\ \vdots \\ \Delta x_{i_n} \end{pmatrix} = TK \begin{pmatrix} \Delta x_1' \\ \Delta x_2' \\ \vdots \\ \Delta x_n' \end{pmatrix} = TK\Delta \boldsymbol{x}'$$

对比前文显然可见，矩阵 TK 就是我们要找的矩阵 S , 也就是说，我们作坐标变换

$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

在新的坐标系下的函数 $f(x_1', x_2', \cdots, x_n')$ 的海森矩阵就是我们在前文中需要的 H' .

注意，满足条件的 $\Delta \boldsymbol{x}'$ 并不是唯一的，因为 K 矩阵的前 $n-l$ 行作线性无关的线性组合组合出一个等价的线性无关组替换前 $n-l$ 行也可以作为一个新的 $\Delta \boldsymbol{x}'$, 另一方面，扩充向量组时， K

的后 l 行自由度是非常高的，甚至不要求后边的 $n-l$ 列的元素全为 0，只要保证最终的矩阵线性无关即可。

10. 拉格朗日乘子法举例

例 假定

$$x+y=1$$

求 x^2y 的极大值。

解 应用不等式知识，我们立即能够得到结论

$$x^2y=xxy=4\times\frac{1}{2}x\cdot\frac{1}{2}x\cdot y\leq 4(\frac{\frac{1}{2}x+\frac{1}{2}x+y}{3})^3=\frac{4}{27}$$

当且仅当

$$\frac{1}{2}x=y=\frac{1}{3}$$

时成立。

用拉格朗日乘子法 构造拉格朗日函数

$$L(x,y,\lambda)=x^2y-\lambda(x+y-1)$$

$$\begin{cases}\frac{\partial L}{\partial x}=2xy-\lambda=0\\ \frac{\partial L}{\partial y}=x^2-\lambda=0\\ \frac{\partial L}{\partial \lambda}=x+y-1=0\end{cases}$$

解得

$$\begin{cases}x=\frac{2}{3}\\ y=\frac{1}{3}\\ \lambda=\frac{4}{9}\end{cases}$$

编辑于 2023-05-01 13:36 • IP 属地北京