

知识点

前言

对于一元函数的极值问题相信大家都十分熟悉，但是对于多元函数的极值问题可能就会比较陌生。大家都学过淑芬怎么可能陌生呢

对于没有限制条件的多元函数来说，只需要对函数求导即可，但是若有了限制条件，即函数的值要在一定条件下才能取到，则需要用到拉格朗日乘子法。

引理

设函数 $f(\vec{x})$

$\varphi(\vec{x}) = (\varphi_1(\vec{x}), \varphi_2(\vec{x}), \dots, \varphi_m(\vec{x}))$

在区域 $D \subset \mathbb{R}^n$ ($m < n$) 内具有各个连续偏导数，再设

$\vec{x}_0 = (x_1^0, x_2^0, \dots, x_n^0) \in D$ 为 $f(\vec{x})$ 在约束条件

$$\begin{cases} \varphi_1(\vec{x}) = 0 \\ \varphi_2(\vec{x}) = 0 \\ \dots \\ \varphi_m(\vec{x}) = 0 \end{cases}$$
 下的极值点，并且 $\varphi'_i(\vec{x}_0)$ 的秩为 m 则存在

常数 $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$ 使得在 $\vec{x}_0$ 处成立下述

等式
$$\begin{cases} \frac{\partial f(\vec{x}_0)}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j(\vec{x}_0)}{\partial x_i} = 0 \quad (i=1, 2, \dots, n) \\ \varphi_j(\vec{x}_0) = 0 \quad (j=1, 2, \dots, m) \end{cases}$$

$$\frac{\partial f(\vec{x}_0)}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j(\vec{x}_0)}{\partial x_i} = 0 \quad (i=1, 2, \dots, n) \\ \varphi_j(\vec{x}_0) = 0 \quad (j=1, 2, \dots, m)$$

证明

由于 $\varphi'_i(\vec{x}_0)$ 的秩为 m 我们不妨设行列

$$\begin{vmatrix} \frac{\partial \varphi_1}{\partial x_{n-m+1}} & \frac{\partial \varphi_1}{\partial x_{n-m+2}} & \dots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_{n-m+1}} & \frac{\partial \varphi_2}{\partial x_{n-m+2}} & \dots & \frac{\partial \varphi_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \varphi_m}{\partial x_{n-m+1}} & \frac{\partial \varphi_m}{\partial x_{n-m+2}} & \dots & \frac{\partial \varphi_m}{\partial x_n} \end{vmatrix}$$
 在

$\vec{x}_0$ 处不为零。因此，在 $\vec{x}_0$ 的某个邻域内唯一确定一组具有各个连续偏导数的隐函

数
$$\begin{cases} x_{n-m+1} = g_1(x_1, x_2, \dots, x_{n-m}) \\ x_{n-m+2} = g_2(x_1, x_2, \dots, x_{n-m}) \\ \dots \\ x_n = g_m(x_1, x_2, \dots, x_{n-m}) \end{cases}$$
 满足

$x_j^0 = g_j(x_1^0, x_2^0, \dots, x_{n-m}^0)$ ($j = n-m+1, n-m+2, \dots, n$) 且有 $\varphi_k(x_1, \dots, x_{n-m}, g_1(x_1, x_2, \dots, x_{n-m}), \dots, g_m(x_1, x_2, \dots, x_{n-m})) = 0$ 将隐函数组代入 $f(\vec{x}_0)$

得 $f(x_1, \dots, x_{n-m}, g_1(x_1, x_2, \dots, x_{n-m}), \dots, g_m(x_1, x_2, \dots, x_{n-m}))$ 因此 $\vec{x}_0$ 是

条件极值点转化为 $(x_1^0, x_2^0, \dots, x_{n-m}^0)$ 为上述函数的通常极值点。

令 $\vec{x}_0$ 则对 $i = 1, 2, \dots, n-m$

有
$$\frac{\partial f(\vec{x}_0)}{\partial x_i} + \frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial g_1(\vec{x}_0)}{\partial x_i} + \dots + \frac{\partial f(\vec{x}_0)}{\partial x_n} \frac{\partial g_m(\vec{x}_0)}{\partial x_i} = 0$$
 令

$\vec{g}(\vec{x}) = (g_1(\vec{x}), g_2(\vec{x}), \dots, g_m(\vec{x}))^T$ 其中

$\vec{x} = (x_1, x_2, \dots, x_{n-m})$ 将上述 $n-m$ 个等式写成向量形式，

有
$$\left(\frac{\partial f(\vec{x}_0)}{\partial x_1}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_{n-m}} \right) + \left(\frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_n} \right) \vec{g}'(\vec{x}_0) = 0$$

由于 $\vec{g}'(\vec{x}_0) = - \begin{pmatrix} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} & \dots & \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \\ \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_1} & \dots & \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m+1}} \\ \dots & \dots & \dots \\ \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_1} & \dots & \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m+1}} \end{pmatrix}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$\frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} + \dots + \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n}$

$$\frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m+2}} \& \cdots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_n} \backslash \vdots \& \ddots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m+1}} \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m+2}} \& \cdots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_n} \end{array} \right)^{-1} \left(\begin{array}{c} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m}} \backslash \vdots \& \ddots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m}} \backslash \vdots \& \ddots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m}} \end{array} \right) \triangleq -A^{-1}B \quad \text{注意到} \left(\frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_n} \right) \cdot A^{-1} \text{ 是一个 } m \text{ 维行向量, 我们可以将其记为} \left(\frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_n} \right) \cdot A^{-1} = \left(\lambda_1, \lambda_2, \dots, \lambda_m \right) \quad \text{将} \left(\frac{\partial f(\vec{x}_0)}{\partial x_1}, \frac{\partial f(\vec{x}_0)}{\partial x_2}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_{n-m}} \right) \text{ 代入之前的式子} \left(\frac{\partial f(\vec{x}_0)}{\partial x_1}, \frac{\partial f(\vec{x}_0)}{\partial x_2}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_{n-m}} \right) + \left(\lambda_1, \lambda_2, \dots, \lambda_m \right) \left(\begin{array}{c} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m}} \backslash \vdots \& \ddots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m}} \backslash \vdots \& \ddots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_1} \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_2} \& \cdots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m}} \end{array} \right) = 0 \quad \text{另外我们可以将} \left(\frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_n} \right) + \left(\lambda_1, \lambda_2, \dots, \lambda_m \right) \left(\begin{array}{c} \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+1}} \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_{n-m+2}} \& \cdots \& \frac{\partial \varphi_1(\vec{x}_0)}{\partial x_n} \backslash \vdots \& \ddots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m+1}} \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_{n-m+2}} \& \cdots \& \frac{\partial \varphi_2(\vec{x}_0)}{\partial x_n} \backslash \vdots \& \ddots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m+1}} \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_{n-m+2}} \& \cdots \& \frac{\partial \varphi_m(\vec{x}_0)}{\partial x_n} \end{array} \right) = 0 \quad \text{将} \left(\frac{\partial f(\vec{x}_0)}{\partial x_{n-m+1}}, \dots, \frac{\partial f(\vec{x}_0)}{\partial x_n} \right) \text{ 写成分量形式再加上约束条件即可证明。}$$

拉格朗日乘子法

构造函数 $F(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(\vec{x}) + \sum_{j=1}^m \lambda_j \varphi_j(\vec{x})$ 则上述求条件极值点的必要条件形式转化为 F 的通常极值的必要条件
$$\begin{cases} \frac{\partial F(\vec{x}_0)}{\partial x_i} = 0 \quad (i=1, 2, \dots, n) \end{cases}$$

$$\frac{\partial F(\vec{x}_0)}{\partial \lambda_j} = 0 \quad (j=1,2,\cdots,m)$$
 此即拉格朗日乘子法

From:
<https://wiki.cvbbacm.com/> - CVBB ACM Team

Permanent link:
https://wiki.cvbbacm.com/doku.php?id=2020-2021:teams:farmer_john:2sozx:%E6%95%B0%E5%AD%A6:%E9%81%97%E7%95%99%E7%9A%84%E7%9F%A5%E8%AF%86%E7%82%B9

Last update: 2020/05/17 21:29