带有不等式约束的多元函数极值问题简介*

2023-2024 学年秋季学期 数学分析 (III)

作为对教材内容的拓展, 我们讨论带有不等式约束的多元函数极值问题.

设 $f:\mathbb{R}^n\to\mathbb{R}$, $\varphi:\mathbb{R}^n\to\mathbb{R}$, 且 f 和 φ 都是 $\mathbb{R}^n$ 上的 C^1 函数. 考虑以下具有不等式约束的极值/最值问题

$$\min_{\text{s.t. }\varphi(\mathbf{x})\le 0} f(\mathbf{x}).$$

这里的记号是要求可行域 $K:=\{\mathbf{x}\in\mathbb{R}^n:\ \varphi(\mathbf{x})\le 0\}$ 中 f 的最小值点, 但我们也可以考虑极小值点——它的定义是, 如果对于 $\mathbf{x}_*\in K$, 存在 $\varepsilon>0$, 使得

$$f(\mathbf{x}_*)\le f(\mathbf{x})\quad \forall\,\mathbf{x}\in U(\mathbf{x}_*,\varepsilon)\cap K,$$

那么就称 $\mathbf{x}_*$ 为条件极小值点. 同理可讨论条件最大值点或极大值点.

下面我们想要研究, 如果已知 $\mathbf{x}_*\in K$ 是一个条件极值点, 那它需要满足什么必要条件——我们期待能得到一些方程和不等式.

不妨假设 $\mathbf{x}_*\in K$ 是一个条件极小值点. 我们分两种情况考虑.

1. 如果 $\mathbf{x}_*$ 为 K 的内点, 由定义, 存在 $\varepsilon_*>0$ 使得 $U(\mathbf{x}_*,\varepsilon_*)\subset K$. 所以由上面极小值的定义, 对于充分小的 ε, $\varepsilon<\varepsilon_*$, 都有

   $$f(\mathbf{x}_*)\le f(\mathbf{x})\quad \forall\,\mathbf{x}\in U(\mathbf{x}_*,\varepsilon).$$

   因而 $\mathbf{x}_*$ 是 f 的一个一般极小值点, 故 $\nabla f(\mathbf{x}_*)=0$.

2. 如果 $\mathbf{x}_*$ 为 K 的边界点, 那么必然有 $\varphi(\mathbf{x}_*)=0$. 事实上, $\mathbf{x}_*\in K$ 保证了 $\varphi(\mathbf{x}_*)\le 0$, 而严格不等号不可能成立, 因为如果 $\varphi(\mathbf{x}_*)<0$, 由 φ 的连续性, 存在 $\mathbf{x}_*$ 的一个邻域, 使得 φ 在其中取值均为负, 因而这个邻域包含于 K, 这与 $\mathbf{x}_*$ 是 K 的边界点矛盾.

   于是 $\mathbf{x}_*$ 也是等式约束 $\varphi(\mathbf{x})=0$ 下 f 的条件极值点. 由 Lagrange 乘子法, 存在 $\lambda_*\in\mathbb{R}$ 使得

   $$\nabla f(\mathbf{x}_*)+\lambda_*\nabla\varphi(\mathbf{x}_*)=0,\quad \varphi(\mathbf{x}_*)=0.$$

   进一步地, 在适当的条件下 (此处略去讨论) 可以证明 $\lambda_*\ge 0$. 我们仅做如下直观解释: 由于 K 中 φ 取值非正, K^c 上 φ 的取值为正, 故在 $\mathbf{x}_*$ 处, $\nabla\varphi(\mathbf{x}_*)$ 指向 K 的外侧. 由上面的条件知 $\nabla f(\mathbf{x}_*)$ 与 $\nabla\varphi(\mathbf{x}_*)$ 平行. 如果 $\lambda_*<0$, 那么它们指向相同, $\nabla f(\mathbf{x}_*)$ 也指向 K 的外侧. 故如果沿着 $-\nabla f(\mathbf{x}_*)$ 方向从 $\mathbf{x}_*$ 移动到 $\mathbf{x}:=\mathbf{x}_*-\delta\nabla f(\mathbf{x}_*)$ (这里 $0<\delta\ll 1$ 是任意的), 那么就总会有 $\mathbf{x}\in K$ 且 $f(\mathbf{x})<f(\mathbf{x}_*)$, 这是因为负梯度方向是 f 下降的方向. 这与 $\mathbf{x}_*$ 是 K 中 f 的条件极小点矛盾. 综上, $\lambda_*\ge 0$.

下面我们尝试将这两种情形综合起来. 定义 Lagrange 增广函数

$$F(\mathbf{x},\lambda):=f(\mathbf{x})+\lambda\varphi(\mathbf{x}).$$

*最后更新日期：2023 年 10 月 23 日

考虑以下这一族条件

$$\begin{aligned}
&\nabla_{\mathbf{x}} F(\mathbf{x},\lambda) = \nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0,\\
&\varphi(\mathbf{x}) \le 0,\\
&\lambda \ge 0,\\
&\lambda \varphi(\mathbf{x}) = 0.
\end{aligned} \tag{1}$$

不难发现, 对于条件极小值点 $\mathbf{x}_*$, 总存在 $\lambda_* \ge 0$, 使得 $(\mathbf{x}_*, \lambda_*)$ 满足以上所有条件. 事实上, 如果出现上面的第一种情形, 只需补充定义 $\lambda_* = 0$ 即可.

以上这族条件被称为 Karush-Kuhn-Tucker 条件 (简称 KKT 条件). 其中的 (1) 被称为互补条件, 它表明 "约束不起作用, $\lambda = 0$", 和 "约束起作用, $\varphi(\mathbf{x}) = 0$", 这两者至少居其一. 所以上面的讨论说明, (在适当的条件下) KKT 条件是条件极值点需满足的一个必要条件. 所以如果要求原问题的条件极小值点, 可以先求满足 KKT 条件的点, 然后再从这些点中筛选.

注记. 如果要求的是条件极大值, KKT 条件中的 $\lambda \ge 0$ 需换为 $\lambda \le 0$.

更进一步地, 我们可以将上述方法推广到同时含有等式和不等式约束的最值/极值问题中. 考虑

$$\begin{aligned}
\min_{\mathbf{x}} f(\mathbf{x}), \quad \text{s.t. } &\varphi_i(\mathbf{x}) = 0, \quad i = 1, \cdots m,\\
&\psi_j(\mathbf{x}) \le 0, \quad j = 1, \cdots l.
\end{aligned}$$

定义 Lagrange 增广函数

$$F(\mathbf{x},\lambda,\mu) := f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j \varphi_j(\mathbf{x}) + \sum_{k=1}^{l} \mu_k \psi_k(\mathbf{x}),$$

这里 $\lambda := (\lambda_1, \cdots, \lambda_m)$, $\mu := (\mu_1, \cdots, \mu_l)$. 此时的 KKT 条件为

$$\begin{aligned}
&\nabla_{\mathbf{x}} F(\mathbf{x},\lambda,\mu) = 0,\\
&\varphi_j(\mathbf{x}) = 0, \quad j = 1, \cdots, m, \quad \text{也即 } \nabla_{\lambda} F(\mathbf{x},\lambda,\mu) = 0,\\
&\psi_k(\mathbf{x}) \le 0,\\
&\mu_k \ge 0,\\
&\mu_k \psi_k(\mathbf{x}) = 0, \quad \text{以上三式对 } k = 1, \cdots, l \text{ 成立}.
\end{aligned}$$

同样地, 我们可以先求满足以上 KKT 条件的点, 然后从中筛选出原问题的条件极小值点. 如果求的是极大值, 那么同样需要将 $\mu_k \ge 0$ 换为 $\mu_k \le 0$.

例 1. 在条件 $x^2 + y^2 = 1$, $x \ge \frac{1}{2}$ 下求函数 $f(x,y) = xy$ 的极大值点.

解. 定义

$$F(x,y,\lambda,\mu) := xy + \lambda(x^2 + y^2 - 1) + \mu\left(\frac{1}{2} - x\right).$$

注意为了应用上述方法, 需将不等式约束写成 $\frac{1}{2}-x\le 0$ 的形式. 由 KKT 条件,

$$\begin{aligned}
&\nabla_x F(x,y,\lambda,\mu)=0, \quad \Leftrightarrow \quad y+2\lambda x-\mu=0,\\
&\nabla_y F(x,y,\lambda,\mu)=0, \quad \Leftrightarrow \quad x+2\lambda y=0,\\
&\nabla_\lambda F(x,y,\lambda,\mu)=0, \quad \Leftrightarrow \quad x^2+y^2=1,\\
&\frac{1}{2}-x\le 0,\\
&\mu\le 0,\\
&\mu\left(\frac{1}{2}-x\right)=0.
\end{aligned}$$

注意因为求的是极大值, 所以 μ 的符号条件为 $\mu\le 0$.

可以从最后一个式子入手求解以上方程.

1. 如果 $\mu=0$, 那么最后两式自然成立, 而前面四式变成

$$y=-2\lambda x, \quad x=-2\lambda y, \quad x^2+y^2=1, \quad x\ge\frac{1}{2}.$$

将这里的第二式代入第一以及第三式得

$$(1-4\lambda^2)y=0, \quad (1+4\lambda^2)y^2=1.$$

由于后一式子表明 $y\ne 0$, 故 $\lambda^2=\frac{1}{4}$, 进而 $y^2=\frac{1}{2}$. 所以 $x^2=1-y^2=\frac{1}{2}$. 最终可以解得两组解

$$(x,y,\lambda,\mu)=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},-\frac{1}{2},0\right),\ \left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},\frac{1}{2},0\right).$$

2. 如果 $\mu\ne 0$, 那么 $x=\frac{1}{2}$. 代入 KKT 条件得

$$\mu=y+\lambda, \quad 1+4\lambda y=0, \quad y^2=\frac{3}{4}, \quad \mu\le 0.$$

最终可以得到一组解

$$(x,y,\lambda,\mu)=\left(\frac{1}{2},-\frac{\sqrt{3}}{2},\frac{1}{2\sqrt{3}},-\frac{1}{\sqrt{3}}\right).$$

从以上三组解中, 我们可以证明

$$(x,y)=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right),\left(\frac{1}{2},-\frac{\sqrt{3}}{2}\right)$$

为条件极大值点, 而另一组解给出的是条件极小值点. 这部分的论证可以通过直接验证极值点定义来完成, 我们留给读者补全细节.