

库恩-塔克条件教学中两个容易忽略的问题

董云达

陈铁生

摘要

库恩-塔克条件是同时含有等式与不等式约束函数的优化问题的最优性条件，已成为优化、控制、运筹与管理的一个核心内容。笔者通过多年的教学实践，发现了两个容易忽略的问题：第一，对比经典的拉格朗日乘子法（针对仅含等式约束函数的优化问题），它的推导方法并不是库恩-塔克条件推导方法的一个推广；第二，库恩-塔克条件中的拉格朗日乘子（存在的话）未必是唯一的。本文通过分析它们各自不同的证明手法，并结合实例加以说明。

关键词：库恩-塔克条件；教学；问题

在大学数学的教材体系中，库恩-塔克条件是理论与实践意义兼备的一个知识点。它实际上是约束优化的最优性条件，揭示了约束优化在一个局部极小点（如果存在的话）处，目标函数的梯度与约束函数的梯度之间应满足的关系。库恩-塔克条件现在已成为优化与控制、运筹与管理的一个核心内容。

库恩-塔克条件针对于同时包含等式约束函数和不等式约束函数的约束优化问题。在讲解这部分内容时，让人很容易联想到仅仅包含等式约束函数的情形以及相应的拉格朗日乘子法。并且常常想当然地断言库恩-塔克条件为拉格朗日乘子法的一个推广。本文的主要目的之一在于通过它们各自不同的证明手法的分析，以及举例说明，这样的观点应该仅仅停留在结论的形式之上。本文的主要目的之二在于结合一个实际例子说明，库恩-塔克条件中的拉格朗日乘子（存在的话）未必是唯一的。

库恩-塔克条件针对如下约束优化问题：

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, \dots, l, \\ & h_j(x) = 0, \quad j = 1, \dots, m, \end{aligned}$$

其中， f, g_i, h_j 均为欧式空间 \mathbb{R}^n 上的连续可微实值函数。 f 为目标函数，每个 g_i 和 h_j

为不等式约束函数和等式约束函数。由约束函数确定的点集

$$X := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, l, h_j(x) = 0, j = 1, \dots, m\}$$

被称为约束优化问题的可行点集。与之相关，有下面三个重要概念。

- (1) 可行方向：已知 X 为上述约束优化问题的可行点集。设非零向量 $d \in \mathbb{R}^n$, $x \in X$ 。若存在一个正数 $\delta > 0$, 使得对于任意给定的 $\alpha \in [0, \delta]$, 总有 $x + \alpha d \in X$, 则称 d 为集合 X 在 x 处的一个可行方向。
- (2) 下降方向：设 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在 x 处连续, $d \in \mathbb{R}^n$ 。若存在某个正数 $\bar{\alpha} > 0$, 使得对于所有 $\alpha \in (0, \bar{\alpha})$, 有 $f(x + \alpha d) < f(x)$, 则称 d 是 f 在 x 处的一个下降方向。
- (3) 积极约束：已知 $x \in X$ 为上述约束优化问题的一个可行点。若 $g_i(x) = 0$, 则称不等式约束函数 g_i 在该点处是一个积极约束函数；否则, 有 $g_i(x) > 0$, 这时称之为非积极约束函数。

为了给出库恩-塔克条件，先回顾以下梯度的定义。对于一个欧氏空间 \mathbb{R}^n 上的连续可微实值函数 f , 它的梯度为

$$\nabla f(x) = (\partial f / \partial x_1, \dots, \partial f / \partial x_n)^T$$

其中 $\partial f / \partial x_i$, $i = 1, \dots, n$, 表示了关于第 i 个分量的偏导数，而 T 表示转置。

在适当的假设条件下，库恩-塔克条件可以叙述为：

（库恩-塔克条件）设 z 为上述约束优化问题的一个局部极小点。则在适当的假设条件下，必然存在一组非负数 λ_i 和一组实数 μ_j , 使得

$$\begin{aligned} \nabla f(z) &= \sum_{i=1}^l \lambda_i \nabla g_i(z) + \sum_{j=1}^m \mu_j \nabla h_j(z), \\ \lambda_i &\geq 0, \quad g_i(z) \geq 0, \quad \lambda_i g_i(z) = 0, \quad i = 1, \dots, l, \\ h_j(z) &= 0, \quad j = 1, \dots, m. \end{aligned}$$

库恩-塔克条件表明，在适当的假设下，目标函数、不等式约束函数、等式约束函数的梯度。库恩-塔克条件成立所需要的假设主要有三类。第一，线性约束即所有的等式约束函数与不等式约束函数都是线性的。第二，线性无关约束：即在该局部极小点处所有的积极约束函数与等式约束函数的梯度都是线性无关的。第三，Slater 条件：目标函数 f , 不等式约束函数 $-g_i$ ($i = 1, \dots, l$) 都是连续可微凸函数（注意是 $-g$ 而不是 g 本身），

等式约束函数 h_j ($j = 1, \dots, m$) 都是线性的, 并且至少存在一个可行点 $\hat{x} \in X$ 使得

$$g_i(x) > 0, \quad i = 1, \dots, l; \quad h_j(x) = 0, \quad j = 1, \dots, m.$$

也就是说, 在三者之一成立的前提下, 结合线性等式与不等式系统中的 Motzkin 定理, 可以证明库恩–塔克条件。

注意, 在库恩–塔克条件中, 形如 $\lambda_i \geq 0$, $g_i(x) \geq 0$, $\lambda_i g_i(x) = 0$ 的关系被称为互补条件 (complementarity condition)。

这里, 我们顺便也提一下。库恩–塔克条件是在 1951 年由数学家库恩和塔克在一篇合作论文中提出的, 它的英文名字为 Kuhn–Tucker conditions, 在一些文献中也被称为 Karush–Kuhn–Tucker conditions。

显然, 对于下面仅仅包含等式约束函数的优化问题

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & h_j(x) = 0, \quad j = 1, \dots, m, \end{aligned}$$

相应的库恩–塔克条件可以简写为:

(库恩–塔克条件的一个特殊形式) 设 z 为上述约束优化问题的一个局部极小点。则在适当的假设条件下, 必然存在一组实数 μ_j , 使得

$$\begin{aligned} \nabla f(z) &= \sum_{j=1}^m \mu_j \nabla h_j(z), \\ h_j(z) &= 0, \quad j = 1, \dots, m. \end{aligned}$$

对比经典的拉格朗日乘子法, 所对应的最优性条件在形式上完全相同。然而, 值得强调的是它们的证明方法截然不同。库恩–塔克条件的推导源于这样的事实: 在约束优化问题的一个局部极小点处, 不存在一个非零向量 d , 它既是目标函数一个下降方向, 也是可行点集在该点处的一个可行方向。而经典的拉格朗日乘子法呢, 用的则是隐函数定理。

下面, 我们举个例子进一步说明本文的第一个目的。为此, 考虑

$$\begin{aligned} \min \quad & x_1 - x_2, \\ \text{s.t.} \quad & x_1^2 + x_2^2 = 1. \end{aligned}$$

可以验证，经典的拉格朗日乘子法适用。然而，对于上述的库恩–塔克条件，从它的推导方法来看，是不适用的。这是因为对于单位圆 $X = \{x = (x_1, x_2) : x_1^2 + x_2^2 = 1\}$ 来说，它在任何一个可行点处都不存在可行方向（注意，所讨论的是单位圆而不是单位圆面），根本不满足推导库恩–塔克条件时所要求的一些假设。

在本文的第二部分，我们结合一个实际例子说明，库恩–塔克条件中的拉格朗日乘子（存在的话）未必是唯一的。而人们往往想当然地认为它要么不存在，要么存在且唯一。实际上，这是一种错觉。

让我们考虑下面的约束优化问题

$$\begin{aligned} \min \quad & f(x) = x_1^2 - x_2, \\ \text{s.t.} \quad & g_1(x) = x_1 - 1 \geq 0, \quad g_2(x) = -x_1^2 - x_2^2 + 26 \geq 0, \\ & h(x) = x_1 + x_2 - 6 = 0. \end{aligned}$$

显然，目标函数、不等式约束函数、等式约束函数的梯度为

$$\begin{aligned} \nabla f(x) &= (2x_1, -1)^T, \quad \nabla g_1(x) = (1, 0)^T, \\ \nabla g_2(x) &= (-2x_1, -2x_2)^T, \quad \nabla h(x) = (1, 1)^T. \end{aligned}$$

并且，可行点 $\hat{x} = (3, 3)^T$ 满足 Slater 条件，从而，对应的库恩–塔克条件为

$$\begin{aligned} 2x_1 - \lambda_1 + \lambda_2(-2x_1 + 2) - \mu &= 0, \\ -1 + \lambda_2 x_2 - \mu &= 0, \\ \lambda_1 \geq 0, \quad x_1 - 1 \geq 0, \quad \lambda_1(x_1 - 1) &= 0, \\ \lambda_2 \geq 0, \quad -x_1^2 - x_2^2 + 26 \geq 0, \quad \lambda_2(-x_1^2 - x_2^2 + 26) &= 0, \\ x_1 + x_2 - 6 &= 0. \end{aligned}$$

现在面临的问题是：如何来得到它的解 $x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \mu^*$ 呢？我们分以下四种情形。

情形 1 $\lambda_1 = 0, \lambda_2 = 0$ 。在此情况下，第一个和第二个关系式分别变为

$$2x_1^* - \mu^* = 0, \quad -1 - \mu^* = 0.$$

于是，我们得到

$$x_1^* = -0.5.$$

但是，它不满足第一个不等式约束条件。所以，这时无解。

情形 2 $\lambda_1^* > 0, \lambda_2^* = 0$ 。由于已知 $\lambda_1 > 0$ ，所以由上面的第一个互补条件，必然有 $x_1^* = 1$ 。这样，我们容易得到它的解

$$x_1^* = 1, \quad x_2^* = 5, \quad \lambda_1^* = 3, \quad \lambda_2^* = 0, \quad \mu^* = -1.$$

情形 3 $\lambda_1^* = 0, \lambda_2^* > 0$ 。由于已知 $\lambda_2 > 0$ ，所以由上面的第二个互补条件，必然有 $-(x_1^*)^2 - (x_2^*)^2 + 26 = 0$ 。结合最后一个关系式，可以得到

$$x_1^* = 1, \quad x_2^* = 5, \quad \lambda_1^* = 0, \quad \lambda_2^* = 3/8, \quad \mu^* = 11/4.$$

另外容易验证这也是唯一的解。

情形 4 $\lambda_1^* > 0, \lambda_2^* > 0$ 。在此情况下，利用互补条件，相应的库恩-塔克条件变为

$$2x_1 - \lambda_1 + 2\lambda_2 x_1 - \mu = 0,$$

$$-1 + \lambda_2 x_2 - \mu = 0,$$

$$x_1 - 1 = 0,$$

$$-x_1^2 - x_2^2 + 26 = 0,$$

$$x_1 + x_2 - 6 = 0.$$

这时，对应的解为 $x_1^* = 1, x_2^* = 5$ 。相应的 Lagrange 乘子为

$$\lambda_1^* = (11 - 4t)/5, \quad \lambda_2^* = (t + 1)/10, \quad \mu^* = t,$$

其中参数 t 满足 $-1 < t < 11/4$ 。

通过这个例子，我们清楚地看到：第四种情形说明了 Lagrange 乘子（存在的话）不一定是唯一的。

本文的讨论，也给我们高校的数学教师带来启示。在讲解一些较为重要的定理时，一定要真正弄清楚它的假设条件、推导过程以及应用范围。有时结论本身较为抽象，往往还需要用一些简明的例子加以进一步解释。这样的话，就会收到事半功倍的效果。

参考文献

- [1] 席少霖, 赵凤治. 最优化计算方法 [M]. 上海: 上海科学技术出版社, 1983.
- [2] 马建国. 数学分析 [M]. 北京: 科学出版社, 2011.
- [3] J. Nocedal, S. Wright. Numerical Optimization[M]. Springer New York, 2006.
- [4] G. Cornuéjols, R. Tütüncü. Optimization Methods in Finance[M]. Cambridge University Press, 2007.