WIKIPEDIA
The Free Encyclopedia

Karush–Kuhn–Tucker conditions
卡鲁什-库恩-塔克条件

In mathematical optimization, the **Karush–Kuhn–Tucker (KKT) conditions**, also known as the **Kuhn–Tucker conditions**, are first derivative tests (sometimes called first-order necessary conditions) for a solution in nonlinear programming to be optimal, provided that some regularity conditions are satisfied.

在数学优化中，Karush-Kuhn-Tucker (KKT) 条件，也称为 Kuhn-Tucker 条件，是非线性规划中最优解的一阶导数检验（有时称为一阶必要条件），前提是一些满足正则性条件。

Allowing inequality constraints, the KKT approach to nonlinear programming generalizes the method of Lagrange multipliers, which allows only equality constraints. Similar to the Lagrange approach, the constrained maximization (minimization) problem is rewritten as a Lagrange function whose optimal point is a global maximum or minimum over the domain of the choice variables and a global minimum (maximum) over the multipliers. The Karush–Kuhn–Tucker theorem is sometimes referred to as the saddle-point theorem.[1]

允许不等式约束，非线性规划的 KKT 方法推广了拉格朗日乘子方法，该方法仅允许等式约束。与拉格朗日方法类似，约束最大化（最小化）问题被重写为拉格朗日函数，其最佳点是选择变量域上的全局最大值或最小值以及乘数上的全局最小值（最大值）。卡鲁什-库恩-塔克定理有时称为鞍点定理。[1]

The KKT conditions were originally named after Harold W. Kuhn and Albert W. Tucker, who first published the conditions in 1951.[2] Later scholars discovered that the necessary conditions for this problem had been stated by William Karush in his master's thesis in 1939.[3][4]

KKT 条件最初以 Harold W. Kuhn 和 Albert W. Tucker 的名字命名，他们于 1951 年首次发表了该条件。[2] 后来学者发现，威廉·卡鲁什 (William Karush) 在他的著作中阐述了该问题的必要条件。1939年硕士论文。[3] [4]

Nonlinear optimization problem
非线性优化问题

Consider the following nonlinear optimization problem in standard form:

考虑以下标准形式的非线性优化问题：

minimize $f(\mathbf{x})$ 最小化 $f(\mathbf{x})$
subject to 受

$$g_i(\mathbf{x}) \leq 0,$$
$$h_j(\mathbf{x}) = 0.$$

where $\mathbf{x} \in \mathbf{X}$ is the optimization variable chosen from a convex subset of $\mathbb{R}^n$, f is the objective or utility function, $g_i\ (i = 1, \ldots, m)$ are the inequality constraint functions and $h_j\ (j = 1, \ldots, \ell)$ are the equality constraint functions. The numbers of inequalities and equalities are denoted by m and

ℓ respectively. Corresponding to the constrained optimization problem one can form the Lagrangian function

其中 $\mathbf{x} \in \mathbf{X}$ 是从 $\mathbb{R}^n$ 的凸子集中选择的优化变量， f 是目标函数或效用函数， $g_i\ (i = 1, \ldots, m)$ 是不等式约束函数和 $h_j\ (j = 1, \ldots, \ell)$ 是等式约束函数。不等式和等式的数量分别用 m 和 ℓ 表示。对应约束优化问题可以构成拉格朗日函数

$$\mathcal{L}(\mathbf{x}, \mu, \lambda) = f(\mathbf{x}) + \mu^\top \mathbf{g}(\mathbf{x}) + \lambda^\top \mathbf{h}(\mathbf{x}) = L(\mathbf{x}, \alpha) = f(\mathbf{x}) + \alpha^\top \begin{pmatrix} \mathbf{g}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}$$

where 在哪里

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_i(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix}, \quad \mathbf{h}(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ \vdots \\ h_j(\mathbf{x}) \\ \vdots \\ h_\ell(\mathbf{x}) \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_m \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_j \\ \vdots \\ \lambda_\ell \end{bmatrix} \quad \text{and} \quad \alpha = \begin{bmatrix} \mu \\ \lambda \end{bmatrix}.$$

The **Karush–Kuhn–Tucker theorem** then states the following.

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_i(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix}, \quad \mathbf{h}(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ \vdots \\ h_j(\mathbf{x}) \\ \vdots \\ h_\ell(\mathbf{x}) \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_m \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_j \\ \vdots \\ \lambda_\ell \end{bmatrix} \quad \text{and} \quad \alpha = \begin{bmatrix} \mu \\ \lambda \end{bmatrix}.$$

卡鲁什-库恩-塔克定理如下所述。

Theorem — (sufficiency) If $(\mathbf{x}^*, \alpha^*)$ is a saddle point of $L(\mathbf{x}, \alpha)$ in $\mathbf{x} \in \mathbf{X}$, $\mu \geq \mathbf{0}$, then $\mathbf{x}^*$ is an optimal vector for the above optimization problem.

定理 — （充分性）如果 $(\mathbf{x}^*, \alpha^*)$ 是 $L(\mathbf{x}, \alpha)$ 在 $\mathbf{x} \in \mathbf{X}$ 、 $\mu \geq \mathbf{0}$ 中的鞍点，则 $\mathbf{x}^*$ 是上述优化问题的最优向量。

(necessity) Suppose that $f(\mathbf{x})$ and $g_i(\mathbf{x})$, $i = 1, \ldots, m$, are convex in $\mathbf{X}$ and that there exists $\mathbf{x}_0 \in \operatorname{relint}(\mathbf{X})$ such that $\mathbf{g}(\mathbf{x}_0) < \mathbf{0}$ (i.e., Slater's condition holds). Then with an optimal vector $\mathbf{x}^*$ for the above optimization problem there is associated a vector $\alpha^* = \begin{bmatrix} \mu^* \\ \lambda^* \end{bmatrix}$ satisfying $\mu^* \geq \mathbf{0}$ such that $(\mathbf{x}^*, \alpha^*)$ is a saddle point of $L(\mathbf{x}, \alpha)$.[5]

（必然性）假设 $f(\mathbf{x})$ 和 $g_i(\mathbf{x})$ 、 $i = 1, \ldots, m$ 在 $\mathbf{X}$ 中是凸的，并且存在 $\mathbf{x}_0 \in \operatorname{relint}(\mathbf{X})$ 使得 $\mathbf{g}(\mathbf{x}_0) < \mathbf{0}$ （即，斯莱特条件成立）。然后，对于上述优化问题，最优向量 $\mathbf{x}^*$ 会关联一个满足 $\mu^* \geq \mathbf{0}$ 的向量 $\alpha^* = \begin{bmatrix} \mu^* \\ \lambda^* \end{bmatrix}$ ，使得 $(\mathbf{x}^*, \alpha^*)$ 是鞍

点 $L(\mathbf{x}, \alpha)$ 。[5]

Since the idea of this approach is to find a supporting hyperplane on the feasible set $\mathbf{\Gamma} = \{\mathbf{x} \in \mathbf{X} : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\}$, the proof of the Karush–Kuhn–Tucker theorem makes use of the hyperplane separation theorem.[6]

由于该方法的思想是在可行集 $\mathbf{\Gamma} = \{\mathbf{x} \in \mathbf{X} : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\}$ 上找到一个支持超平面，因此 Karush-Kuhn-Tucker 定理的证明利用了超平面分离定理。[6]

The system of equations and inequalities corresponding to the KKT conditions is usually not solved directly, except in the few special cases where a closed-form solution can be derived analytically. In general, many optimization algorithms can be interpreted as methods for numerically solving the KKT system of equations and inequalities.[7]

与 KKT 条件相对应的方程组和不等式组通常不直接求解，除非在少数特殊情况下可以通过解析导出封闭式解。一般来说，许多优化算法可以解释为数值求解 KKT 方程组和不等式组的方法。[7]

Necessary conditions 必要条件

Suppose that the objective function $f\colon \mathbb{R}^n \rightarrow \mathbb{R}$ and the constraint functions $g_i\colon \mathbb{R}^n \rightarrow \mathbb{R}$ and $h_j\colon \mathbb{R}^n \rightarrow \mathbb{R}$ have subderivatives at a point $x^* \in \mathbb{R}^n$. If x^* is a local optimum and the optimization problem satisfies some regularity conditions (see below), then there exist constants $\mu_i\ (i = 1, \ldots, m)$ and $\lambda_j\ (j = 1, \ldots, \ell)$, called KKT multipliers, such that the following four groups of conditions hold:[8]

假设目标函数 $f\colon \mathbb{R}^n \rightarrow \mathbb{R}$ 以及约束函数 $g_i\colon \mathbb{R}^n \rightarrow \mathbb{R}$ 和 $h_j\colon \mathbb{R}^n \rightarrow \mathbb{R}$ 在点 $x^* \in \mathbb{R}^n$ 处有导数。如果 x^* 是局部最优且优化问题满足某些规律性条件（见下文），则存在常数 $\mu_i\ (i = 1, \ldots, m)$ 和 $\lambda_j\ (j = 1, \ldots, \ell)$ ，称为 KKT 乘子，例如满足以下四组条件：[8]

Stationarity 平稳性

For minimizing $f(x)$: $\partial f(x^*) + \sum_{j=1}^{\ell} \lambda_j \partial h_j(x^*) + \sum_{i=1}^{m} \mu_i \partial g_i(x^*) \ni \mathbf{0}$

为了最小化 $f(x)$ ： $\partial f(x^*) + \sum_{j=1}^{\ell} \lambda_j \partial h_j(x^*) + \sum_{i=1}^{m} \mu_i \partial g_i(x^*) \ni \mathbf{0}$

For maximizing $f(x)$: $-\partial f(x^*) + \sum_{j=1}^{\ell} \lambda_j \partial h_j(x^*) + \sum_{i=1}^{m} \mu_i \partial g_i(x^*) \ni \mathbf{0}$

为了最大化 $f(x)$ ： $-\partial f(x^*) + \sum_{j=1}^{\ell} \lambda_j \partial h_j(x^*) + \sum_{i=1}^{m} \mu_i \partial g_i(x^*) \ni \mathbf{0}$

Primal feasibility 初步可行性

$h_j(x^*) = 0, \text{ for } j = 1, \ldots, \ell$

$g_i(x^*) \leq 0, \text{ for } i = 1, \ldots, m$

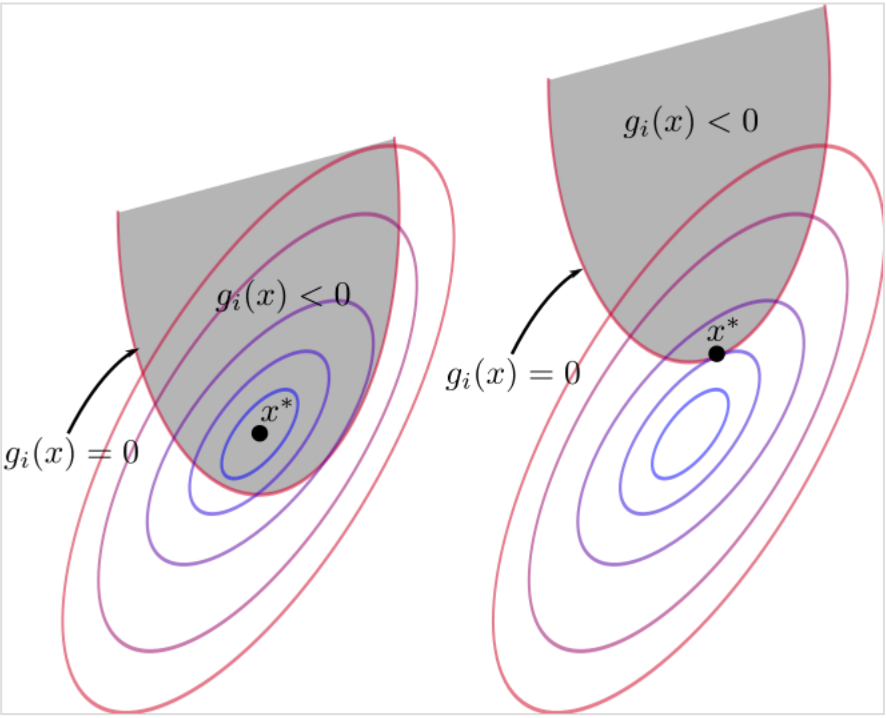
Dual feasibility 双重可行性

$\mu_i \geq 0, \text{ for } i = 1, \ldots, m$

Complementary slackness 互补松弛

$$\sum_{i=1}^{m} \mu_i g_i(x^*) = 0.$$

The last condition is sometimes written in the equivalent form:



Inequality constraint diagram for optimization problems

优化问题的不等式约束图

$\mu_i g_i(x^*) = 0, \text{ for } i = 1, \ldots, m.$

最后一个条件有时会写成等效的形式：$\mu_i g_i(x^*) = 0, \text{ for } i = 1, \ldots, m.$

In the particular case $m = 0$, i.e., when there are no inequality constraints, the KKT conditions turn into the Lagrange conditions, and the KKT multipliers are called Lagrange multipliers.

在特定情况 $m = 0$ 下，即不存在不等式约束时，KKT 条件变为拉格朗日条件，KKT 乘子称为拉格朗日乘子。

Proof 证明

Theorem — (sufficiency) If there exists a solution x^* to the primal problem, a solution (μ^*, λ^*) to the dual problem, such that together they satisfy the KKT conditions, then the problem pair has strong duality, and $x^*, (\mu^*, \lambda^*)$ is a solution pair to the primal and dual problems.

定理 —（充分性）如果存在原问题的解 x^* ，对偶问题的解 (μ^*, λ^*) ，使得它们一起满足 KKT 条件，则问题对有强对偶性， $x^*, (\mu^*, \lambda^*)$ 是原问题和对偶问题的解对。

(necessity) If the problem pair has strong duality, then for any solution x^* to the primal problem and any solution (μ^*, λ^*) to the dual problem, the pair $x^*, (\mu^*, \lambda^*)$ must satisfy the KKT conditions.[9]

（必然性）如果问题对具有强对偶性，那么对于原始问题的任何解 x^* 和对偶问题的任何解 (μ^*, λ^*) ，对 $x^*, (\mu^*, \lambda^*)$ 必须满足KKT条件。[9]

Proof 证明

First, for the $x^*, (\mu^*, \lambda^*)$ to satisfy the KKT conditions is equivalent to them being a Nash equilibrium.

首先， $x^*, (\mu^*, \lambda^*)$ 满足KKT条件就相当于纳什均衡。

Fix (μ^*, λ^*), and vary x: equilibrium is equivalent to primal stationarity.

固定 (μ^*, λ^*) ，并改变 x ：均衡相当于原始平稳性。

Fix x^*, and vary (μ, λ): equilibrium is equivalent to primal feasibility and complementary slackness.

固定 x^* ，并改变 (μ, λ) ：平衡相当于原始可行性和互补松弛。

Sufficiency: the solution pair $x^*, (\mu^*, \lambda^*)$ satisfies the KKT conditions, thus is a Nash equilibrium, and therefore closes the duality gap.

充分性：解对 $x^*, (\mu^*, \lambda^*)$ 满足KKT条件，因此是纳什均衡，因此闭合对偶间隙。

Necessity: any solution pair $x^*, (\mu^*, \lambda^*)$ must close the duality gap, thus they must constitute a Nash equilibrium (since neither side could do any better), thus they satisfy the KKT conditions.

必然性：任何解对 $x^*, (\mu^*, \lambda^*)$ 都必须闭合对偶间隙，因此它们必须构成纳什均衡（因为双方都无法做得更好），从而满足KKT条件。

Interpretation: KKT conditions as balancing constraint-forces in state space

解释：KKT 条件作为状态空间中的平衡约束力

The primal problem can be interpreted as moving a particle in the space of x, and subjecting it to three kinds of force fields:

主要问题可以解释为在 x 空间中移动粒子，并使其承受三种力场：

- f is a potential field that the particle is minimizing. The force generated by f is $-\partial f$.

  f 是粒子正在最小化的势场。 f 产生的力是 $-\partial f$ 。

- g_i are one-sided constraint surfaces. The particle is allowed to move inside $g_i \leq 0$, but whenever it touches $g_i = 0$, it is pushed inwards.

  g_i 是单侧约束曲面。粒子可以在 $g_i \leq 0$ 内部移动，但每当它接触 $g_i = 0$ 时，它就会被向内推。

- h_j are two-sided constraint surfaces. The particle is allowed to move only on the surface h_j.

  h_j 是两侧约束面。粒子只允许在表面 h_j 上移动。

Primal stationarity states that the "force" of $\partial f(x^*)$ is exactly balanced by a linear sum of forces $\partial h_j(x^*)$ and $\partial g_i(x^*)$.

原始平稳性表明 $\partial f(x^*)$ 的“力”完全由力 $\partial h_j(x^*)$ 和 $\partial g_i(x^*)$ 的线性和平衡。

Dual feasibility additionally states that all the $\partial g_i(x^*)$ forces must be one-sided, pointing inwards into the feasible set for x.

对偶可行性还指出，所有 $\partial g_i(x^*)$ 力必须是单向的，向内指向 x 的可行集。

Dual slackness states that if $g_i(x^*) < 0$, then the $\partial g_i(x^*)$ force must be zero, since the particle is not on the boundary, the one-sided constraint force cannot activate.

对偶松弛表明，如果 $g_i(x^*) < 0$ ，则 $\partial g_i(x^*)$ 力必须为零，因为粒子不在边界上，因此单侧约束力无法激活。

Matrix representation 矩阵表示

The necessary conditions can be written with Jacobian matrices of the constraint functions. Let $\mathbf{g}(x) : \mathbb{R}^n \to \mathbb{R}^m$ be defined as $\mathbf{g}(x) = (g_1(x), \ldots, g_m(x))^\top$ and let $\mathbf{h}(x) : \mathbb{R}^n \to \mathbb{R}^\ell$ be defined as $\mathbf{h}(x) = (h_1(x), \ldots, h_\ell(x))^\top$. Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)^\top$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_\ell)^\top$. Then the necessary conditions can be written as:

必要条件可以用约束函数的雅可比矩阵来写。将 $\mathbf{g}(x) : \mathbb{R}^n \to \mathbb{R}^m$ 定义为 $\mathbf{g}(x) = (g_1(x), \ldots, g_m(x))^\top$ 并将 $\mathbf{h}(x) : \mathbb{R}^n \to \mathbb{R}^\ell$ 定义为 $\mathbf{h}(x) = (h_1(x), \ldots, h_\ell(x))^\top$ 。让 $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)^\top$ 和 $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_\ell)^\top$ 。那么必要条件可以写为：

Stationarity 平稳性

For maximizing $f(x)$: $\partial f(x^*) - D\mathbf{g}(x^*)^\top \boldsymbol{\mu} - D\mathbf{h}(x^*)^\top \boldsymbol{\lambda} = \mathbf{0}$

为了最大化 $f(x)$ ： $\partial f(x^*) - D\mathbf{g}(x^*)^\top \boldsymbol{\mu} - D\mathbf{h}(x^*)^\top \boldsymbol{\lambda} = \mathbf{0}$

For minimizing $f(x)$: $\partial f(x^*) + D\mathbf{g}(x^*)^\top \boldsymbol{\mu} + D\mathbf{h}(x^*)^\top \boldsymbol{\lambda} = \mathbf{0}$

为了最小化 $f(x)$ ： $\partial f(x^*) + D\mathbf{g}(x^*)^\top \boldsymbol{\mu} + D\mathbf{h}(x^*)^\top \boldsymbol{\lambda} = \mathbf{0}$

Primal feasibility 初步可行性

$\mathbf{g}(x^*) \le \mathbf{0}$
$\mathbf{h}(x^*) = \mathbf{0}$

Dual feasibility 双重可行性

$\boldsymbol{\mu} \ge \mathbf{0}$

Complementary slackness 互补松弛

$\boldsymbol{\mu}^\top \mathbf{g}(x^*) = 0.$

Regularity conditions (or constraint qualifications) 正则条件（或约束条件）

One can ask whether a minimizer point x^* of the original, constrained optimization problem (assuming one exists) has to satisfy the above KKT conditions. This is similar to asking under what conditions the minimizer x^* of a function $f(x)$ in an unconstrained problem has to satisfy the condition $\nabla f(x^*) = 0$. For the constrained case, the situation is more complicated, and one can state a variety of (increasingly complicated) "regularity" conditions under which a constrained minimizer also satisfies the KKT conditions. Some common examples for conditions that guarantee this are tabulated in the following, with the LICQ the most frequently used one:

人们可能会问，原始约束优化问题（假设存在）的最小化点 x^* 是否必须满足上述 KKT 条件。这类似于询问在无约束问题中函数 $f(x)$ 的最小化器 x^* 在什么条件下必须满足条件 $\nabla f(x^*) = 0$ 。对于受约束的情况，情况更加复杂，可以陈述多种（越来越复杂的）“正则性”条件，在这些条件下，受约束的最小化器也满足 KKT 条件。下表列出了保证这一点的一些常见条件示例，其中 LICQ 是最常用的一个：

Constraint	Acronym	Statement
Linearity constraint qualification 线性约束条件	LCQ	If g_i and h_j are affine functions, then no other condition is needed. 如果 g_i 和 h_j 是仿射函数，则不需要其他条件。
Linear independence constraint qualification 线性独立约束条件	LICQ	The gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at x^*. 主动不等式约束的梯度和等式约束的梯度在 x^* 处线性无关。
Mangasarian-Fromovitz constraint qualification Mangasarian-Fromovitz 约束条件	MFCQ	The gradients of the equality constraints are linearly independent at x^* and there exists a vector $d \in \mathbb{R}^n$ such that $\nabla g_i(x^*)^\top d < 0$ for all active inequality constraints and $\nabla h_j(x^*)^\top d = 0$ for all equality constraints.[10] 等式约束的梯度在 x^* 处是线性无关的，并且存在一个向量 $d \in \mathbb{R}^n$ ，使得所有有效不等式约束的 $\nabla g_i(x^*)^\top d < 0$ 和 $\nabla h_j(x^*)^\top d = 0$ 对于所有平等约束。 [10]
Constant rank constraint qualification 恒定等级约束条件	CRCQ	For each subset of the gradients of the active inequality constraints and the gradients of the equality constraints the rank at a vicinity of x^* is constant. 对于有效不等式约束的梯度和等式约束的梯度的每个子集， x^* 附近的等级是恒定的。
Constant positive linear dependence constraint qualification 恒定正线性相关约束条件	CPLD	For each subset of gradients of active inequality constraints and gradients of equality constraints, if the subset of vectors is linearly dependent at x^* with non-negative scalars associated with the inequality constraints, then it remains linearly dependent in a neighborhood of x^*. 对于主动不等式约束的梯度和等式约束的梯度的每个子集，如果向量子集在 x^* 处与与不等式约束关联的非负标量线性相关，则它在邻域中保持线性相关 x^* 。
Quasi-normality constraint qualification 拟正态性约束条件	QNCQ	If the gradients of the active inequality constraints and the gradients of the equality constraints are linearly dependent at x^* with associated multipliers λ_j for equalities and $\mu_i \geq 0$ for inequalities, then there is no sequence $x_k \to x^*$ such that $\lambda_j \neq 0 \Rightarrow \lambda_j h_j(x_k) > 0$ and $\mu_i \neq 0 \Rightarrow \mu_i g_i(x_k) > 0$. 如果主动不等式约束的梯度和等式约束的梯度在 x^* 处与关联的等式乘数 λ_j 和不等式的相关乘数 $\mu_i \geq 0$ 线性相关，则没有序列 $x_k \to x^*$ 使得 $\lambda_j \neq 0 \Rightarrow \lambda_j h_j(x_k) > 0$ 和 $\mu_i \neq 0 \Rightarrow \mu_i g_i(x_k) > 0$.
Slater's condition 斯莱特的病情	SC	For a convex problem (i.e., assuming minimization, f, g_i are convex and h_j is affine), there exists a point x such that $h_j(x) = 0$ and $g_i(x) < 0$. 对于凸问题（即，假设最小化， f, g_i 是凸的， h_j 是仿射的），存在一个点 x 使得 $h_j(x) = 0$ 和 $g_i(x) < 0$.

The strict implications can be shown

可以显示严格的含义

LICQ ⇒ MFCQ ⇒ CPLD ⇒ QNCQ

and

LICQ ⇒ CRCQ ⇒ CPLD ⇒ QNCQ

In practice weaker constraint qualifications are preferred since they apply to a broader selection of problems.

在实践中，较弱的约束条件是优选的，因为它们适用于更广泛的问题选择。

Sufficient conditions 充分条件

In some cases, the necessary conditions are also sufficient for optimality. In general, the necessary conditions are not sufficient for optimality and additional information is required, such as the Second Order Sufficient Conditions (SOSC). For smooth functions, SOSC involve the second derivatives, which explains its name.

在某些情况下，必要条件也足以实现最优性。一般来说，必要条件不足以实现最优性，需要额外的信息，例如二阶充分条件 (SOSC)。对于平滑函数，SOSC 涉及二阶导数，这也解释了它的名字。

The necessary conditions are sufficient for optimality if the objective function f of a maximization problem is a differentiable concave function, the inequality constraints g_j are differentiable convex functions, the equality constraints h_i are affine functions, and Slater's condition holds.[11] Similarly, if the objective function f of a minimization problem is a differentiable convex function, the necessary conditions are also sufficient for optimality.

如果最大化问题的目标函数 f 是可微凹函数，不等式约束 g_j 是可微凸函数，等式约束 h_i 类似地，如果最小化问题的目标函数 f 是可微凸函数，则必要条件也足以实现最优性。

It was shown by Martin in 1985 that the broader class of functions in which KKT conditions guarantees global optimality are the so-called Type 1 **invex functions**.[12][13]

Martin 在 1985 年指出，KKT 条件保证全局最优性的更广泛的一类函数是所谓的 1 型凹函数。[12][13]

Second-order sufficient conditions

二阶充分条件

For smooth, non-linear optimization problems, a second order sufficient condition is given as follows.

对于平滑的非线性优化问题，二阶充分条件如下。

The solution x^*, λ^*, μ^* found in the above section is a constrained local minimum if for the Lagrangian,

上一节中找到的解 x^*, λ^*, μ^* 是一个受约束的局部最小值，如果对于拉格朗日函数，

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x) + \sum_{j=1}^{\ell} \lambda_j h_j(x)$$

then, 然后，

$$s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) s \geq 0$$

where $s \neq 0$ is a vector satisfying the following,

其中 $s \neq 0$ 是满足以下条件的向量，

$$\left[\nabla_x g_i(x^*), \nabla_x h_j(x^*)\right]^T s = 0$$

where only those active inequality constraints $g_i(x)$ corresponding to strict complementarity (i.e. where $\mu_i > 0$) are applied. The solution is a strict constrained local minimum in the case the inequality is also strict.

其中仅应用那些与严格互补性相对应的主动不等式约束 $g_i(x)$ （即 $\mu_i > 0$ ）。在不等式也严格的情况下，解决方案是严格约束的局部最小值。

If $s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) s = 0$, the third order Taylor expansion of the Lagrangian should be used to verify if x^* is a local minimum. The minimization of $f(x_1, x_2) = (x_2 - x_1^2)(x_2 - 3x_1^2)$ is a good counter-example, see also Peano surface.

如果 $s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) s = 0$ ，则应使用拉格朗日的三阶泰勒展开来验证 x^* 是否为局部最小值。 $f(x_1, x_2) = (x_2 - x_1^2)(x_2 - 3x_1^2)$ 的最小化是一个很好的反例，另请参见皮亚诺曲面。

Economics 经济学

Often in mathematical economics the KKT approach is used in theoretical models in order to obtain qualitative results. For example,[14] consider a firm that maximizes its sales revenue subject to a minimum profit constraint. Letting Q be the quantity of output produced (to be chosen), $R(Q)$ be sales revenue with a positive first derivative and with a zero value at zero output, $C(Q)$ be production costs with a positive first derivative and with a non-negative value at zero output, and $G_{\min}$ be the positive minimal acceptable level of profit, then the problem is a meaningful one if the revenue function levels off so it eventually is less steep than the cost function. The problem expressed in the previously given minimization form is

在数理经济学中，KKT 方法经常用于理论模型中以获得定性结果。例如， [14] 考虑一家在最小利润约束下最大化其销售收入的公司。令 Q 为产出量（待选择）， $R(Q)$ 为一阶导数为正且零产出时的销售收入， $C(Q)$ 是具有正一阶导数且在零产出时具有非负值的生产成本，并且 $G_{\min}$ 是正的最小可接受利润水平，那么如果收入函数趋于平稳，则该问题是有意义的所以它最终比成本函数陡峭。以前面给出的最小化形式表达的问题是

Minimize $-R(Q)$ 最小化 $-R(Q)$
subject to 受
$G_{\min} \leq R(Q) - C(Q)$
$Q \geq 0,$

and the KKT conditions are

KKT 条件为

$$\left(\frac{\mathrm{d}R}{\mathrm{d}Q}\right)(1+\mu)-\mu\left(\frac{\mathrm{d}C}{\mathrm{d}Q}\right)\leq 0,$$

$$Q\geq 0,$$

$$Q\left[\left(\frac{\mathrm{d}R}{\mathrm{d}Q}\right)(1+\mu)-\mu\left(\frac{\mathrm{d}C}{\mathrm{d}Q}\right)\right]=0,$$

$$R(Q)-C(Q)-G_{\min}\geq 0,$$

$$\mu\geq 0,$$

$$\mu[R(Q)-C(Q)-G_{\min}]=0.$$

Since $Q=0$ would violate the minimum profit constraint, we have $Q>0$ and hence the third condition implies that the first condition holds with equality. Solving that equality gives

由于 $Q=0$ 会违反最小利润约束，因此我们有 $Q>0$，因此第三个条件意味着第一个条件相等。解决这个平等给出

$$\frac{\mathrm{d}R}{\mathrm{d}Q}=\frac{\mu}{1+\mu}\left(\frac{\mathrm{d}C}{\mathrm{d}Q}\right).$$

Because it was given that $\mathrm{d}R/\mathrm{d}Q$ and $\mathrm{d}C/\mathrm{d}Q$ are strictly positive, this inequality along with the non-negativity condition on μ guarantees that μ is positive and so the revenue-maximizing firm operates at a level of output at which marginal revenue $\mathrm{d}R/\mathrm{d}Q$ is less than marginal cost $\mathrm{d}C/\mathrm{d}Q$ — a result that is of interest because it contrasts with the behavior of a profit maximizing firm, which operates at a level at which they are equal.

因为已知 $\mathrm{d}R/\mathrm{d}Q$ 和 $\mathrm{d}C/\mathrm{d}Q$ 严格为正，所以这种不等式以及 μ 上的非负条件保证了 μ 是正数，因此收入最大化的公司在边际收入 $\mathrm{d}R/\mathrm{d}Q$ 小于边际成本 $\mathrm{d}C/\mathrm{d}Q$ 的产出水平上运营——这个结果很有趣，因为它与利润最大化企业的行为，该企业在平等的水平上运作。

Value function 价值函数

If we reconsider the optimization problem as a maximization problem with constant inequality constraints:

如果我们将优化问题重新考虑为具有恒定不等式约束的最大化问题：

$$\text{Maximize }\ f(x)$$

$$\text{subject to}$$

$$g_i(x)\leq a_i, h_j(x)=0.$$

The value function is defined as

价值函数定义为

$$V(a_1,\ldots,a_n)=\sup_x f(x)$$

$$\text{subject to}$$

$$g_i(x) \leq a_i, h_j(x) = 0$$

$$j \in \{1, \ldots, \ell\}, i \in \{1, \ldots, m\},$$

so the domain of V is $\{a \in \mathbb{R}^m \mid \text{for some } x \in X, g_i(x) \leq a_i, i \in \{1, \ldots, m\}\}$.

所以 V 的域是 $\{a \in \mathbb{R}^m \mid \text{for some } x \in X, g_i(x) \leq a_i, i \in \{1, \ldots, m\}\}$.

Given this definition, each coefficient μ_i is the rate at which the value function increases as a_i increases. Thus if each a_i is interpreted as a resource constraint, the coefficients tell you how much increasing a resource will increase the optimum value of our function f. This interpretation is especially important in economics and is used, for instance, in utility maximization problems.

根据此定义，每个系数 μ_i 是价值函数随着 a_i 增加而增加的速率。因此，如果每个 a_i 被解释为资源约束，则系数会告诉您增加资源将增加函数 f 的最佳值多少。这种解释在经济学中尤其重要，例如用于效用最大化问题。

Generalizations 概括

With an extra multiplier $\mu_0 \geq 0$, which may be zero (as long as $(\mu_0, \mu, \lambda) \neq 0$), in front of $\nabla f(x^*)$ the KKT stationarity conditions turn into

使用额外的乘数 $\mu_0 \geq 0$，该乘数可能为零（只要 $(\mu_0, \mu, \lambda) \neq 0$ ），在 $\nabla f(x^*)$ 之前，KKT 平稳性条件变为

$$\mu_0 \nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{\ell} \lambda_j \nabla h_j(x^*) = 0,$$

$$\mu_j g_i(x^*) = 0, \quad i = 1, \ldots, m,$$

which are called the Fritz John conditions. This optimality conditions holds without constraint qualifications and it is equivalent to the optimality condition *KKT or (not-MFCQ)*.

这被称为弗里茨约翰条件。该最优性条件在没有约束条件的情况下成立，并且等同于最优性条件 KKT或（非MFCQ）。

The KKT conditions belong to a wider class of the first-order necessary conditions (FONC), which allow for non-smooth functions using subderivatives.

KKT 条件属于更广泛的一阶必要条件 (FONC)，它允许使用导数的非光滑函数。

See also 也可以看看

- Farkas' lemma 法卡斯引理
- Lagrange multiplier 拉格朗日乘子
- The Big M method, for linear problems, which extends the simplex algorithm to problems that contain "greater-than" constraints.

  Big M 方法，用于解决线性问题，它将单纯形算法扩展到包含“大于”约束的问题。
- Interior-point method a method to solve the KKT conditions.

  内点法是一种求解KKT条件的方法。

- Slack variable 松弛变量
- Slater's condition 斯莱特的病情

References 参考

1. Tabak, Daniel; Kuo, Benjamin C. (1971). *Optimal Control by Mathematical Programming*. Englewood Cliffs, NJ: Prentice-Hall. pp. 19–20. ISBN 0-13-638106-5.

   塔巴克，丹尼尔；郭本杰明 (1971)。通过数学规划进行最优控制。新泽西州恩格尔伍德悬崖：Prentice-Hall。第 19-20 页。 ISBN 0-13-638106-5。

2. Kuhn, H. W.; Tucker, A. W. (1951). "Nonlinear programming" (http://projecteuclid.org/euclid.bsmsp/1200500249). *Proceedings of 2nd Berkeley Symposium*. Berkeley: University of California Press. pp. 481–492. MR 0047303 (https://mathscinet.ams.org/mathscinet-getitem?mr=0047303).

   库恩，H.W.；塔克，A.W. (1951)。 “非线性规划”。第二届伯克利研讨会论文集。伯克利：加州大学出版社。第 481–492 页。先生 0047303。

3. W. Karush (1939). *Minima of Functions of Several Variables with Inequalities as Side Constraints* (http://pi.lib.uchicago.edu/1001/cat/bib/4111654) (M.Sc. thesis). Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois.

   W·卡鲁什 (1939)。以不等式作为侧面约束的多个变量函数的最小值（硕士论文）。大学数学系芝加哥，芝加哥，伊利诺伊州。

4. Kjeldsen, Tinne Hoff (2000). "A contextualized historical analysis of the Kuhn-Tucker theorem in nonlinear programming: the impact of World War II" (https://doi.org/10.1006%2Fhmat.2000.2289). *Historia Math*. **27** (4): 331–361. doi:10.1006/hmat.2000.2289 (https://doi.org/10.1006%2Fhmat.2000.2289). MR 1800317 (https://mathscinet.ams.org/mathscinet-getitem?mr=1800317).

   蒂尼·霍夫·凯尔德森 (2000)。 “非线性规划中库恩-塔克定理的情境历史分析：第二次世界大战的影响”。历史数学。 27（4）：331-361。 doi：10.1006/hmat.2000.2289。先生 1800317。

5. Walsh, G. R. (1975). "Saddle-point Property of Lagrangian Function" (https://books.google.com/books?id=K0EZAQAAIAAJ&pg=PA39). *Methods of Optimization*. New York: John Wiley & Sons. pp. 39–44. ISBN 0-471-91922-5.

   沃尔什，G.R.（1975）。 “拉格朗日函数的鞍点性质”。优化方法。纽约：约翰·威利父子公司。第 39-44 页。 ISBN 0-471-91922-5。

6. Kemp, Murray C.; Kimura, Yoshio (1978). *Introduction to Mathematical Economics* (https://archive.org/details/introductiontoma0000kemp/page/38). New York: Springer. pp. 38–44 (https://archive.org/details/introductiontoma0000kemp/page/38). ISBN 0-387-90304-6.

   肯普，默里·C.；木村义夫 (1978)。数理经济学导论。纽约：施普林格。第 38-44 页。 ISBN 0-387-90304-6。

7. Boyd, Stephen; Vandenberghe, Lieven (2004). *Convex Optimization*. Cambridge: Cambridge University Press. p. 244. ISBN 0-521-83378-7. MR 2061575 (https://mathscinet.ams.org/mathscinet-getitem?mr=2061575).

   博伊德，斯蒂芬；列文·范登堡 (2004)。凸优化。剑桥：剑桥大学出版社。 p。 244.ISBN 0-521-83378-7。先生2061575。

8. Ruszczyński, Andrzej (2006). *Nonlinear Optimization*. Princeton, NJ: Princeton University Press. ISBN 978-0691119151. MR 2199043 (https://mathscinet.ams.org/mathscinet-getitem?mr=2199043).

   安杰伊·鲁什琴斯基 (2006)。非线性优化。新泽西州普林斯顿：普林斯顿大学出版社。 ISBN 978-0691119151。先生 2199043。

9. Geoff Gordon & Ryan Tibshirani. "Karush-Kuhn-Tucker conditions, Optimization 10-725 / 36-725" (https://web.archive.org/web/20220617163414/http://www.cs.cmu.edu/~ggordon/10725-F12/slides/16-kkt.pdf) (PDF). Archived from the original (http://www.cs.cmu.edu/~ggordon/10725-F12/slides/16-kkt.pdf) (PDF) on 2022-06-17.

   杰夫·戈登和瑞安·提布希拉尼。 “Karush-Kuhn-Tucker 条件，优化 10-725 / 36-725” (PDF)。原始版本 (PDF) 存档于2022-06-17。

10. Dimitri Bertsekas (1999). *Nonlinear Programming* (2 ed.). Athena Scientific. pp. 329–330. ISBN 9781886529007.

    迪米特里·贝尔塞卡斯 (1999)。非线性规划（第二版）。雅典娜科学。第 329–330 页。国际标准书号 9781886529007。

11. Boyd, Stephen; Vandenberghe, Lieven (2004). *Convex Optimization*. Cambridge: Cambridge University Press. p. 244. ISBN 0-521-83378-7. MR 2061575 (https://mathscinet.ams.org/mathscinet-getitem?mr=2061575).

    博伊德，斯蒂芬；列文·范登堡 (2004)。凸优化。剑桥：剑桥大学出版社。 p。 244.ISBN 0-521-83378-7。先生2061575。

12. Martin, D. H. (1985). "The Essence of Invexity". *J. Optim. Theory Appl.* **47** (1): 65–76. doi:10.1007/BF00941316 (https://doi.org/10.1007%2FBF00941316). S2CID 122906371 (https://api.semanticscholar.org/CorpusID:122906371).

    马丁，D.H.（1985）。 《烦恼的本质》。 J. Optim。理论应用47（1）：65-76。 doi：10.1007/BF00941316。 S2CID 122906371。

13. Hanson, M. A. (1999). "Invexity and the Kuhn-Tucker Theorem" (https://doi.org/10.1006%2Fjmaa.1999.6484). *J. Math. Anal. Appl.* **236** (2): 594–604. doi:10.1006/jmaa.1999.6484 (https://doi.org/10.1006%2Fjmaa.1999.6484).

    汉森，文学硕士（1999）。 “Invexity 和库恩-塔克定理”。 J.马斯。肛门。应用。 236（2）：594-604。 doi：10.1006/jmaa.1999.6484。

14. Chiang, Alpha C. *Fundamental Methods of Mathematical Economics*, 3rd edition, 1984, pp. 750–752.

    Jiang, Alpha C. 《数理经济学基本方法》，第 3 版，1984 年，第 750–752 页。

Further reading 进一步阅读

- Andreani, R.; Martínez, J. M.; Schuverdt, M. L. (2005). "On the relation between constant positive linear dependence condition and quasinormality constraint qualification". *Journal of Optimization Theory and Applications*. **125** (2): 473–485. doi:10.1007/s10957-004-1861-9 (https://doi.org/10.1007%2Fs10957-004-1861-9). S2CID 122212394 (https://api.semanticscholar.org/CorpusID:122212394).

  安德烈亚尼，R.；马丁内斯，J.M.；舒维尔特，M.L.（2005）。 “论常数正线性相关条件与拟正态约束条件之间的关系”。优化理论与应用杂志。 125（2）：473-485。 doi：10.1007/s10957-004-1861-9。 S2CID 122212394。

- Avriel, Mordecai (2003). *Nonlinear Programming: Analysis and Methods*. Dover. ISBN 0-486-43227-0.

  艾薇儿，莫迪凯 (2003)。非线性规划：分析和方法。多佛。 ISBN 0-486-43227-0。

- Boltyanski, V.; Martini, H.; Soltan, V. (1998). "The Kuhn–Tucker Theorem" (https://books.google.com/books?id=YD7UBwAAQBAJ&pg=PA78). *Geometric Methods and Optimization Problems*. New York: Springer. pp. 78–92. ISBN 0-7923-5454-0.

  博尔琴斯基，V.；马蒂尼，H.；索尔坦，V. (1998)。 “库恩-塔克定理”。几何方法和优化问题。纽约：施普林格。第 78–92 页。 ISBN 0-7923-5454-0。

- Boyd, S.; Vandenberghe, L. (2004). "Optimality Conditions" (https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf#page=255) (PDF). *Convex Optimization*. Cambridge University Press. pp. 241–249. ISBN 0-521-83378-7.

  博伊德，S.；范登堡，L.（2004）。 “最优性条件”(PDF)。凸优化。剑桥大学出版社。第 241–249 页。 ISBN 0-521-83378-7。

- Kemp, Murray C.; Kimura, Yoshio (1978). *Introduction to Mathematical Economics* (https://archive.org/details/introductiontoma0000kemp/page/38). New York: Springer. pp. 38–73 (https://archive.org/details/introductiontoma0000kemp/page/38). ISBN 0-387-90304-6.

  肯普，默里·C.；木村义夫 (1978)。数理经济学导论。纽约：施普林格。第 38-73 页。 ISBN 0-387-90304-6。

- Rau, Nicholas (1981). "Lagrange Multipliers". *Matrices and Mathematical Programming*. London: Macmillan. pp. 156–174. ISBN 0-333-27768-6.

  尼古拉斯·劳 (1981)。 “拉格朗日乘子”。矩阵和数学规划。伦敦：麦克米伦。第 156–174 页。 ISBN 0-333-27768-6。

- Nocedal, J.; Wright, S. J. (2006). *Numerical Optimization*. New York: Springer. ISBN 978-0-387-30303-1.

  诺塞达尔，J.；赖特，S.J.（2006）。数值优化。纽约：施普林格。 ISBN 978-0-387-30303-1。

- Sundaram, Rangarajan K. (1996). "Inequality Constraints and the Theorem of Kuhn and Tucker" (https://books.google.com/books?id=yAfug81P-8YC&pg=PA145). *A First Course in Optimization Theory*. New York: Cambridge University Press. pp. 145–171. ISBN 0-521-49770-1.

  Sundaram，Rangarajan K. (1996)。 “不等式约束以及库恩和塔克定理”。优化理论第一课。纽约：剑桥大学出版社。第 145–171 页。 ISBN 0-521-49770-1。

External links 外部链接

- Karush–Kuhn–Tucker conditions with derivation and examples

  卡鲁什-库恩-塔克条件及其推导和示例

  (http://www.onmyphd.com/?p=kkt.karush.kuhn.tucker)
- Examples and Tutorials on the KKT Conditions

  KKT 条件的示例和教程

(http://apmonitor.com/me575/index.php/Main/KuhnTucker)
