

约束极值问题的最优性条件

一. 约束极值问题

二. 最优性条件

一. 约束极值问题

$$\begin{aligned} & \min f(x) \\ s.t. \quad & \begin{cases} h_i(x) = 0 & i = 1, 2, \dots, m \\ g_j(x) \geq 0 & j = 1, 2, \dots, l \end{cases} \end{aligned}$$

$$记 \mathbf{h}(x) = (h_1(x), h_2(x), \dots, h_m(x))^T,$$

$$\mathbf{g}(x) = (g_1(x), g_2(x), \dots, g_l(x))^T,$$

则约束极值问题可记为 $\min f(x)$

$$s.t. \quad \begin{cases} \mathbf{h}(x) = \mathbf{0} \\ \mathbf{g}(x) \geq \mathbf{0} \end{cases}$$

令 $Q = \{x | \mathbf{h}(x) = \mathbf{0}, \mathbf{g}(x) \geq \mathbf{0}\}$, 称 Q 为此约束极值问题的可行域。

$$\because h_i(x) = 0 \Leftrightarrow \begin{cases} h_i(x) \geq 0 \\ -h_i(x) \geq 0 \end{cases}$$

∴ 约束极值问题也可记为

$$\begin{aligned} & \min f(x) \\ s.t. \quad & g(x) \geq 0 \end{aligned}$$

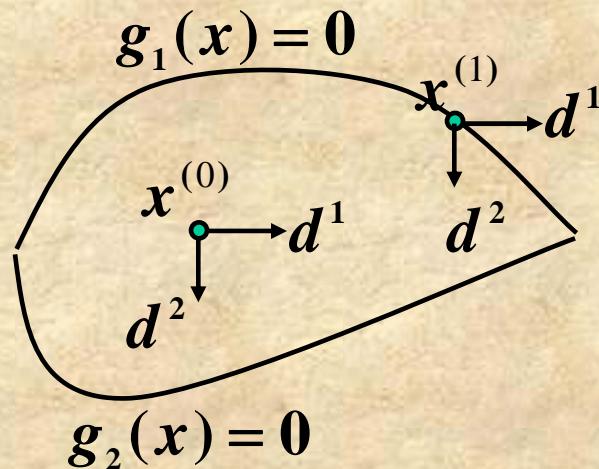
二. 最优性条件

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } g(x) \geq 0 \end{aligned} \tag{1}$$

可行域为 $Q = \{x | g(x) \geq 0\}$ 。

1. 可行方向和积极约束

可行方向: 设 $x^{(0)} \in Q, d$ 为一个向量。如果存在 实数 $\bar{\lambda} > 0$, 使得对任意的 $\lambda \in [0, \bar{\lambda}]$ 有 $x^{(0)} + \lambda d \in Q$, 则称 d 为 $x^{(0)}$ 处的一个可行方向。



积极约束: 设点 $\bar{x} \in Q$, 对于不等式约束 $g_i(x) \geq 0$, 如果 $g_i(\bar{x}) = 0$, 则称 $g_i(x) \geq 0$ 是点 \bar{x} 处的积极约束。

记 $I(\bar{x}) = \{i | g_i(\bar{x}) = 0, 1 \leq i \leq l\}$, 称 $I(\bar{x})$ 为点 \bar{x} 处的积极约束指标集。

例1 设 $g_1(x) = x_2 - \sqrt{2}x_1^2 \geq 0, g_2(x) = 1 - x_1^2 - x_2^2 \geq 0,$

$g_3(x) = x_1 \geq 0$ 。令 $\bar{x} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})^T$, 求点 \bar{x} 的积极约束

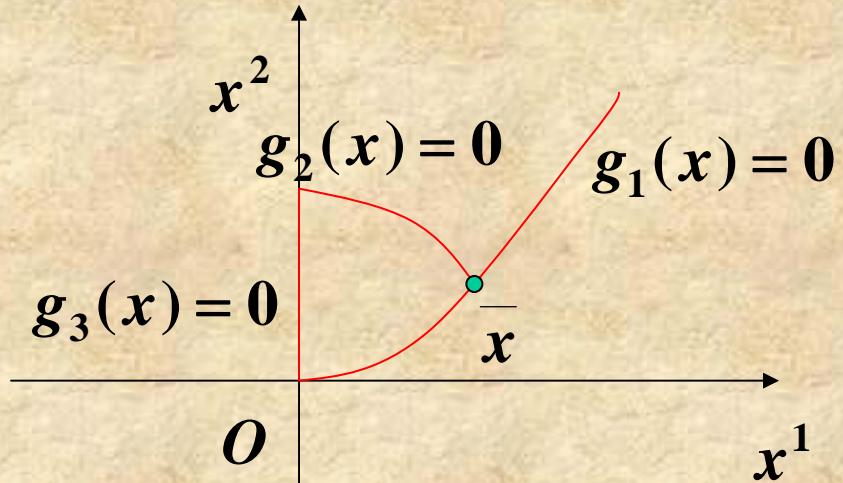
指标集。

$$\text{解: } \because g_1(\bar{x}) = \frac{\sqrt{2}}{2} - \sqrt{2} \times \left(\frac{\sqrt{2}}{2}\right)^2 = 0,$$

$$g_2(\bar{x}) = 1 - \left(\frac{\sqrt{2}}{2}\right)^2 - \left(\frac{\sqrt{2}}{2}\right)^2 = 0,$$

$$g_3(\bar{x}) = \frac{\sqrt{2}}{2} > 0.$$

$$\therefore I(\bar{x}) = \{1, 2\}.$$



如何判断一个向量是否 是可行方向?

定理1 给定点 $\bar{x} \in Q$, 记点 \bar{x} 的积极约束指标集为 $I(\bar{x})$ 。给定向量 d , 如果对任意的 $i \in I(\bar{x})$ 有 $\nabla g_i(\bar{x})^T d > 0$, 则 d 是点 \bar{x} 的可行方向。

证明：令 $x' = \bar{x} + t d$, $t > 0$ 。则对任意的 $i \in I(\bar{x})$, 有

$$\begin{aligned} g_i(x') &= g_i(\bar{x}) + t \nabla g_i(\bar{x})^T d + o(\|td\|^2) \\ &= t \nabla g_i(\bar{x})^T d + o(\|td\|^2) \\ &> 0 \end{aligned}$$

$\therefore x' \in Q$, 即 d 为可行方向。

可行下降方向：设点 $\bar{x} \in Q$, 给定向量 d , 如果 d 既是点 \bar{x} 处的可行方向, 又是该点 \bar{x} 的下降方向, 则称 d 为点 \bar{x} 处的可行下降方向。

定理 2 给定点 $\bar{x} \in Q$, 记点 \bar{x} 的积极约束指标集为 $I(\bar{x})$ 。给定向量 d , 如果 d 满足

$$\begin{cases} \nabla g_i(\bar{x})^T d > 0 & i \in I(\bar{x}) \\ \nabla f(\bar{x})^T d < 0 \end{cases}$$

则向量 d 是点 \bar{x} 处的可行下降方向。

极值点的必要条件:

定理 3 设 $x^* \in Q$, $I(x^*)$ 是其积极约束指标集。 $f(x)$ 和 $g_i(x)$ ($i \in I(x^*)$) 在点 x^* 处可微, $g_i(x)$ ($i \notin I(x^*)$) 在点 x^* 处连续。如果 x^* 是约束极值问题 (1) 的局部极小点, 则在点 x^* 处没有可行下降方向。

2. K-T 条件（库恩 - 塔克条件）

设点 x^* 是约束极值问题 (1) 的局部极小点，
 $I(x^*)$ 是其积极约束指标集。

则由定理 2 可知，不存在向量 d ，使下式成立

$$\begin{cases} \nabla g_i(x^*)^T d > 0 & i \in I(x^*) \\ \nabla f(x^*)^T d < 0 \end{cases}.$$

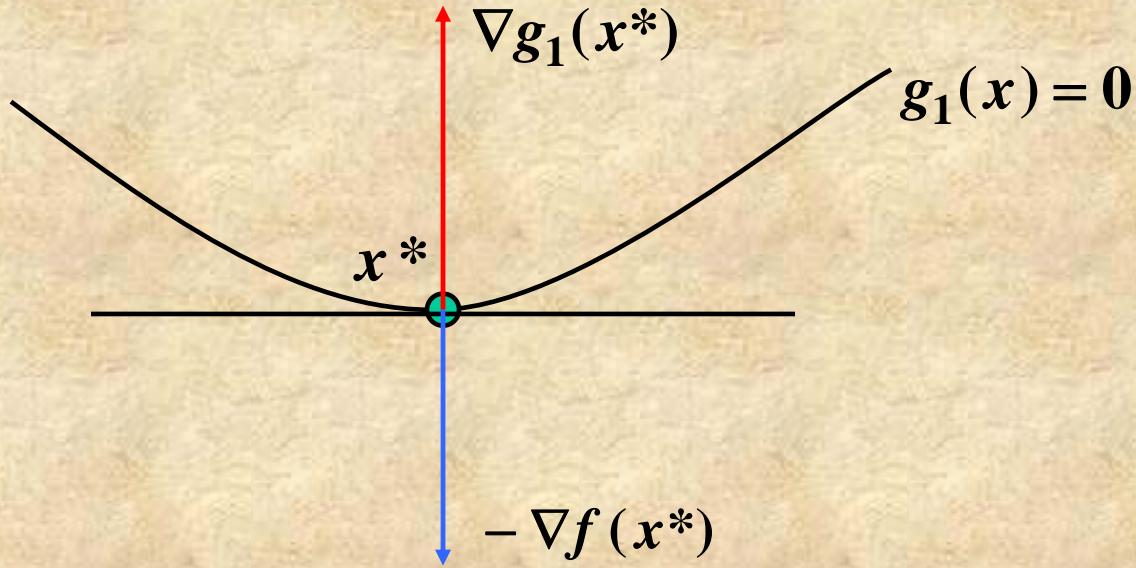
分析：

(1) 如果 $I(x^*)$ 中只有一个指标，不妨设 $g_1(x)$ 为积极约束。

则不存在向量 d 使得

$$\begin{cases} \nabla f(x^*)^T d < 0 \\ \nabla g_1(x^*)^T d > 0 \end{cases}$$

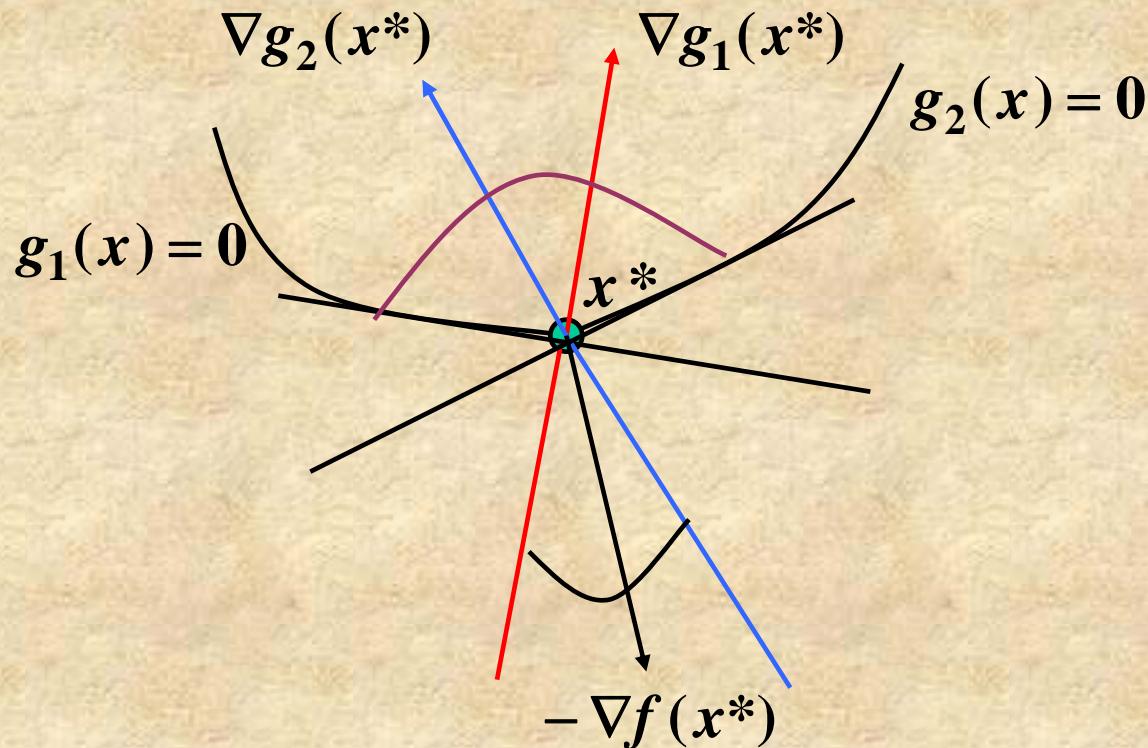
成立。



则有 $\nabla f(x^*) = \lambda \nabla g_1(x^*), \lambda \geq 0.$

即 $\nabla f(x^*) - \lambda \nabla g_1(x^*) = 0.$

(2) 如果 $I(x^*)$ 中有两个指标，不妨设 $g_1(x)$ 和 $g_2(x)$ 为积极约束。并设 $\nabla g_1(x^*)$ 和 $\nabla g_2(x^*)$ 线性无关。



\therefore 存在 $\lambda_1, \lambda_2 \geq 0$, 使得 $\nabla f(x^*) = \lambda_1 \nabla g_1(x^*) + \lambda_2 \nabla g_2(x^*)$ 。

$$\therefore \nabla f(x^*) - \lambda_1 \nabla g_1(x^*) - \lambda_2 \nabla g_2(x^*) = \mathbf{0}.$$

(3) 一般情况：设 $\{\nabla g_i(x^*) | i \in I(x^*)\}$ 线性无关。

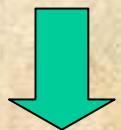
则存在非负实数 $\lambda_i (i \in I(x^*))$, 使得

$$\nabla f(x^*) - \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = \mathbf{0} \quad (2)$$

(2) 式可改写为

$$\left\{ \begin{array}{l} \nabla f(x^*) - \sum_{i=1}^l \lambda_i \nabla g_i(x^*) = \mathbf{0} \\ \lambda_i g_i(x^*) = \mathbf{0}, i = 1, 2, \dots, l \\ \lambda_i \geq 0, i = 1, 2, \dots, l \end{array} \right. \quad (3)$$

$$\begin{cases} \lambda_i g_i(x^*) = 0, i = 1, 2, \dots, l \\ \lambda_i \geq 0, i = 1, 2, \dots, l \end{cases}$$



$$\begin{cases} \lambda_i = 0, & g_i(x^*) > 0; \\ \lambda_i \geq 0, & g_i(x^*) = 0; \end{cases}$$

定理 4(**K-T**条件) 设 $x^* \in Q$, $f(x)$ 和 $g_i(x)$ ($i \in I(x^*)$) 在点 x^* 处可微, $g_i(x)$ ($i \notin I(x^*)$) 在点 x^* 处连续, $\{\nabla g_i(x^*)|i \in I(x^*)\}$ 线性无关。若 x^* 是约束极值问题 (1) 的局部极小点, 则存在 一组实数 λ_i 使其满足

$$\left\{ \begin{array}{l} \nabla f(x^*) - \sum_{i=1}^l \lambda_i \nabla g_i(x^*) = \mathbf{0} \\ \lambda_i g_i(x^*) = \mathbf{0}, i = 1, 2, \dots, l \\ \lambda_i \geq \mathbf{0}, i = 1, 2, \dots, l \end{array} \right. (*)$$

(*)式称为**K-T**条件(库恩-塔克条件), 满足(*)式的点称为**K-T**点。

(4) 对于有等式约束的极值 问题

$$\begin{aligned} & \min f(x) \\ s.t. \quad & \begin{cases} h(x) = 0 \\ g(x) \geq 0 \end{cases} \end{aligned}$$

K - T 条件可写为

$$\left\{ \begin{array}{l} \nabla f(x^*) + \sum_{j=1}^m u_j \nabla h_j(x^*) - \sum_{i=1}^l \lambda_i \nabla g_i(x^*) = 0 \\ \lambda_i g_i(x^*) = 0, i = 1, 2, \dots, l \\ \lambda_i \geq 0, i = 1, 2, \dots, l \end{array} \right. \quad (**)$$

例2 试写出下述问题的 $K-T$ 条件。

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_1x_2 - 2x_2^2 - 2x_2 \\ s.t. \quad & \left\{ \begin{array}{l} 3x_1^2 - 2x_2^2 + 2x_1 \leq 8 + 3x_2 \\ x_1^3 + 2x_1x_2 = 4x_2^2 - 2x_1 + 1 \\ x_1 \geq 1 \end{array} \right. \end{aligned}$$

解: $\nabla f(x) = [2x_1 + x_2, x_1 - 4x_2 - 2]^T$

$$\because g_1(x) = 3x_2 + 8 - 3x_1^2 + 2x_2^2 - 2x_1$$

$$\therefore \nabla g_1(x) = [-6x_1 - 2, 3 + 4x_2]^T$$

$$\because g_2(x) = x_1^3 + 2x_1x_2 - 4x_2^2 + 2x_1 - 1$$

$$\therefore \nabla g_2(x) = [3x_1^2 + 2x_2 + 2, 2x_1 - 8x_2]^T$$

$$\because g_3(x) = x_1 - 1 \quad \therefore \nabla g_3(x) = [1, 0]^T$$

$\therefore K-T$ 条件为

$$\left\{ \begin{array}{l} \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 4x_2 - 2 \end{bmatrix} - \lambda_1 \begin{bmatrix} -6x_1 - 2 \\ 3 + 4x_2 \end{bmatrix} - \mu \begin{bmatrix} 3x_1^2 + 2x_2 + 2 \\ 2x_1 - 8x_2 \end{bmatrix} - \lambda_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1(3x_2 + 8 - 3x_1^2 + 2x_2^2 - 2x_1) = 0 \\ \lambda_2(x_1 - 1) = 0 \\ \lambda_1, \lambda_2 \geq 0 \end{array} \right.$$

例 给定非线性规划问题

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

检验下列两点 $x^{(1)} = (0,0)^T$ 和 $x^{(2)} = (1,1)^T$ 是否为K-T点。

解: $\nabla f = [2(x_1 - 2), 2x_2]^T$, $\nabla g_1 = [1, -2x_2]^T$, $\nabla g_2 = [-1, 1]^T$ 。

(1) 检验点 $x^{(1)}$:

$$\nabla f(x^{(1)}) = [-4, 0]^T, \nabla g_1(x^{(1)}) = [1, 0]^T, \nabla g_2(x^{(1)}) = [-1, 1]^T$$

所以K-T条件为:

$$\left\{ \begin{array}{l} \begin{bmatrix} -4 \\ 0 \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1(0 - 0^2) = 0 \\ \lambda_2(0 - 0) = 0 \\ \lambda_1, \lambda_2 \geq 0 \end{array} \right.$$

解方程可得: $\lambda_1 = -4, \lambda_2 = 0$.

因为 $\lambda_1 = -4 < 0$, 所以 $x^{(1)}$ 不是 K-T 点。

(2) 检验点 $x^{(2)}$:

$$\nabla f(x^{(2)}) = [-2, 2]^T, \nabla g_1(x^{(2)}) = [1, -2]^T, \nabla g_2(x^{(2)}) = [-1, 1]^T$$

所以K-T条件为:

$$\begin{cases} \begin{bmatrix} -2 \\ 2 \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} - \lambda_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1(1 - 1^2) = 0 \\ \lambda_2(1 - 1) = 0 \\ \lambda_1, \lambda_2 \geq 0 \end{cases}$$

解方程可得: $\lambda_1 = 0, \lambda_2 = 2$.

因为 $\lambda_1, \lambda_2 \geq 0$, 所以 $x^{(2)}$ 是K—T点。

3. $K-T$ 点的计算

例 求约束极值问题

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2^2 - 6x_1 - 6x_2 + 8 \\ s.t. \quad & \left\{ \begin{array}{l} x_1 + x_2 \leq 4 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{array} \right. \end{aligned}$$

的 $K-T$ 点。

解: $\nabla f(x) = 2[x_1 - 3, x_2 - 3]^T$ 。

$$\because g_1(x) = 4 - x_1 - x_2 \quad \therefore \nabla g_1(x) = [-1, -1]^T$$

$$g_2(x) = x_1, \quad \nabla g_2(x) = [1, 0]^T.$$

$$g_3(x) = x_2, \quad \nabla g_3(x) = [0, 1]^T.$$

由 $K - T$ 条件得

$$\begin{bmatrix} x_1 - 3 \\ x_2 - 3 \end{bmatrix} - \lambda_1 \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

由 $K - T$ 条件及约束条件得

$$\left\{ \begin{array}{l} x_1 + \lambda_1 - \lambda_2 = 3 \\ x_2 + \lambda_1 - \lambda_3 = 3 \\ \lambda_1(4 - x_1 - x_2) = 0 \\ \lambda_2 x_1 = 0 \\ \lambda_3 x_2 = 0 \\ x_1 + x_2 \leq 4 \\ \lambda_1, \lambda_2, \lambda_3, x_1, x_2 \geq 0 \end{array} \right.$$

以下分情况讨论：

(1) 若 $x_1 = x_2 = 0$:

由 $\lambda_1(4 - x_1 - x_2) = 0$ 可得 $\lambda_1 = 0$ 。

$$\therefore \lambda_1 - \lambda_2 = 3 \quad \Rightarrow \lambda_2 = -3$$

这与 $\lambda_2 \geq 0$ 矛盾。

(2) 若 $x_1 = 0, x_2 \neq 0$:

$$\therefore \lambda_3 = 0$$

$$\therefore \begin{cases} x_2 + \lambda_1 = 3 \\ \lambda_1 - \lambda_2 = 3 \end{cases} \Rightarrow x_2 + \lambda_2 = 0 \quad \Rightarrow \lambda_2 = -x_2 < 0$$

这与 $\lambda_2 \geq 0$ 矛盾。

(3) 若 $x_1 \neq 0, x_2 = 0$:

$$\therefore \lambda_2 = 0$$

$$\therefore \begin{cases} x_1 + \lambda_1 = 3 \\ \lambda_1 - \lambda_3 = 3 \end{cases} \Rightarrow x_1 + \lambda_3 = 0 \Rightarrow \lambda_3 = -x_1 < 0$$

这与 $\lambda_3 \geq 0$ 矛盾。

(4) 若 $x_1 \neq 0, x_2 \neq 0$:

$$\therefore \lambda_2 = \lambda_3 = 0$$

$$\therefore \begin{cases} x_1 + \lambda_1 = 3 \\ x_2 + \lambda_1 = 3 \end{cases} \Rightarrow x_1 = x_2$$

$$\text{若 } x_1 + x_2 < 4 \Rightarrow \lambda_1 = 0 \Rightarrow x_1 = x_2 = 3$$

$$\Rightarrow x_1 + x_2 = 6 > 4 \Rightarrow \text{矛盾。}$$

$$\therefore x_1 + x_2 = 4 \Rightarrow x_1 = x_2 = 2 \Rightarrow \lambda_1 = 1$$

$\therefore [2, 2]^T$ 为 $K - T$ 点。

4. 其它最优化条件

定理 5(*Fritz John* 条件) 设 $x^* \in Q$, $f(x)$ 和 $g_i(x)(i \in I(x^*))$ 在点 x^* 处可微, $g_i(x)(i \notin I(x^*))$ 在点 x^* 处连续。若 x^* 是约束极值问题 (1) 的局部极小点, 则存在一组不全为零的非负实数 λ_i , 使得

$$\lambda_0 \nabla f(x^*) - \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = \mathbf{0}$$