

# 第六章 约束优化问题的最优性条件



## 本章内容:

### ✓ 5.1 约束最优化问题的最优性条件

# 6.1 约束最优化问题的最优性条件

信息与计算机科学系

邵建峰



*Optimization*

邵建峰



## 本节内容:

- ✓ 一. 等式约束最优化问题的最优性条件
- ✓ 二. 不等式约束最优化问题最优性条件
- ✓ 三. 一般约束最优化问题的最优性条件

## 约束最优化问题的一般形式

$$\min \quad f(x)$$

$$\text{s.t.} \quad g_i(x) \geq 0 \quad i = 1, 2, \dots, m$$

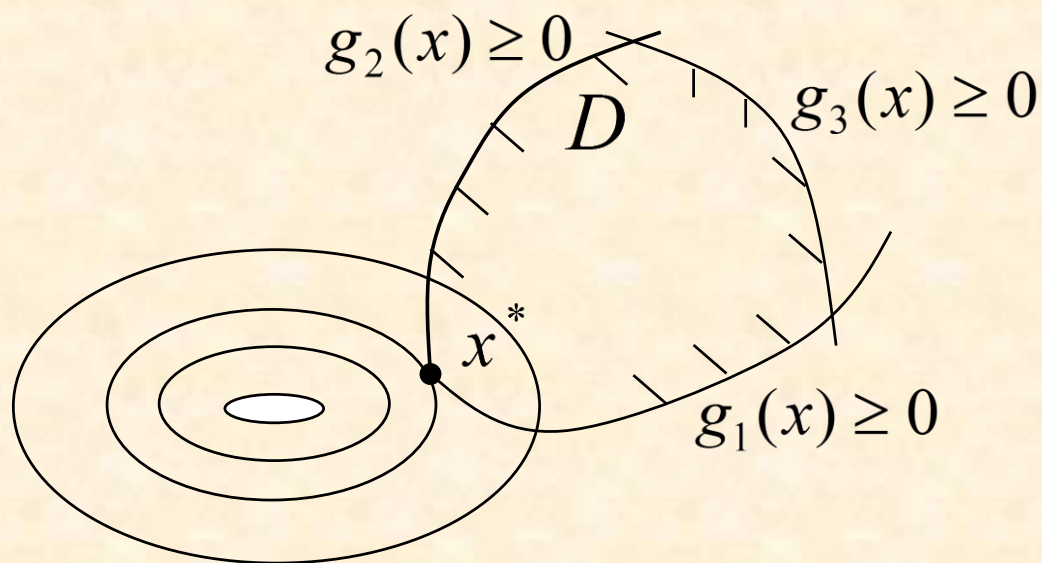
$$h_j(x) = 0 \quad j = 1, 2, \dots, l$$

$$x = (x_1, x_2, \dots, x_n)^T \in R^n$$

其可行（容许）解集

$$S = \left\{ x \mid g_i(x) \geq 0, i = 1, 2, \dots, m, h_j(x) = 0, j = 1, 2, \dots, l, \right. \\ \left. x = (x_1, x_2, \dots, x_n)^T \in R^n \right\}$$

下面研究约束优化问题的**最优解**应该**满足**什么条件：



## 一.等式约束最优化问题的最优性条件

### 等式约束优化问题 (I) :

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_j(x) = 0 \quad j = 1, 2, \dots, l \\ & x = (x_1, x_2, \dots, x_n)^T \in R^n \end{aligned}$$

这样的问题是《微积分》中条件极值问题的推广。

其拉格朗日 (Lagrange) 函数

$$L(x, \lambda) = f(x) - \lambda^T \cdot h(x) = f(x) - \sum_{j=1}^l \lambda_j \cdot h_j(x)$$

其中  $h(x) = (h_1(x), h_2(x), \dots, h_l(x))^T$ .

## 定理1 (Lagrange) 设

- (1)  $x^*$  是问题 ( I ) 的局部最优点;
- (2)  $f(x), h_1(x), h_2(x), \dots, h_l(x) : R^n \rightarrow R$  在  $x^*$  附近的某邻域内可微;
- (3)  $\nabla h_1(x), \nabla h_2(x), \dots, \nabla h_l(x)$  线性无关。

则必存在实数  $\lambda_1^*, \lambda_2^*, \dots, \lambda_l^*$  , 使得

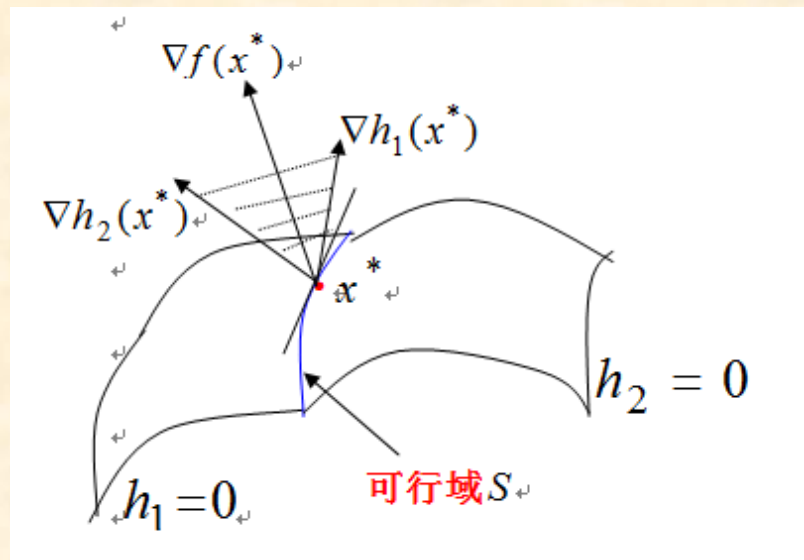
即 
$$\nabla f(x^*) - \sum_{j=1}^l \lambda_j^* \cdot \nabla h_j(x^*) = 0$$

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - (\lambda^*)^T \cdot \nabla h(x^*) = 0$$

其中  $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_l^*)^T$  。

**证明** 对  $n=3, l=2$  证明。

如果  $l=2$ ，约束曲线为三维空间中曲面  $h_1(x) = 0$  和曲面  $h_2(x) = 0$  的**交线**。



同样可以说明  $(-)\mathbf{g}^*$  与曲线的**切线**垂直。因此曲面在  $\mathbf{x}^*$  处的法向量  $\nabla h_1(\mathbf{x}^*), \nabla h_2(\mathbf{x}^*)$  与梯度向量  $\mathbf{g}^*$  共面。于是存在数  $\lambda_1, \lambda_2$ ，使得

$$\nabla f(\mathbf{x}^*) - \lambda_1 \nabla h_1(\mathbf{x}^*) - \lambda_2 \nabla h_2(\mathbf{x}^*) = 0$$

这个定理是等式约束最优化问题的局部最优解的  
一阶必要条件。

下面不予证明而给出相应的二阶充分条件。

**定理2 (二阶充分条件)** 设在问题 ( I ) 中

(1)  $f(x), h_1(x), h_2(x), \dots, h_l(x)$  **二阶连续**可微;

(2) 存在  $x^* \in R^n$ ,  $\lambda^* \in R^l$  使得 Lagrange 函数的梯度为零, 即

$$\nabla L(x^*, \lambda^*) = \nabla f(x^*) - (\lambda^*)^T \cdot \nabla h(x^*) = 0$$

(3) 对任意  $v \in R^n$ , 且使得  $v^T \cdot \nabla h(x^*) = 0$  的非零向量  $v$ , 总有

$$v^T \nabla L_x^2(x^*, \lambda^*) v > 0$$

那么,  $x^*$  是问题 ( I ) 的严格局部最优点。

### 例1 试用最优化条件求

$$\min \left\{ x_1^2 + x_2^2 \mid x_1 \cdot x_2 = 8 \right\}$$

的最优解。

**解** 因为拉格朗日（Lagrange）函数

$$\begin{aligned} L(x, \lambda) &= f(x) - \lambda^T \cdot h(x) = f(x) - \sum_{j=1}^l \lambda_j \cdot h_j(x) \\ &= (x_1^2 + x_2^2) + \lambda \cdot (x_1 \cdot x_2 - 8) \end{aligned}$$

所以

$$\nabla_x L(x, \lambda) = \nabla f(x) - \lambda^T \cdot \nabla h(x) = \begin{pmatrix} 2x_1 - \lambda \cdot x_2 \\ 2x_2 - \lambda \cdot x_1 \end{pmatrix}$$

且

$$\nabla_x^2 L(x, \lambda) = \begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix}$$

令  $\nabla_x L(x, \lambda) = 0$ , 得

$$2x_1 - \lambda \cdot x_2 = 0$$

$$2x_2 - \lambda \cdot x_1 = 0$$

又

$$x_1 \cdot x_2 = 8$$

解得  $x^* = (\sqrt{8}, \sqrt{8})^T$ ,  $\lambda^* = 2$ . 而

$$\nabla h(x^*) = (x_2^*, x_1^*)^T = (\sqrt{8}, \sqrt{8})^T$$

而且对使  $v^T \cdot \nabla h(x^*) = 0$  的任意非零向量

$v = (k, -k)^T$ , ( $k > 0$ ), 都有

$$v^T \nabla_x^2 L(x^*, \lambda^*) v = v^T \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} v = 8k > 0$$

从而  $x^* = (\sqrt{8}, \sqrt{8})^T$  是原问题的严格局部最优点。

## 二. 不等式约束最优化问题的最优性条件

### 不等式约束优化问题（II）

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \geq 0 \quad i = 1, 2, \dots, m \\ & x = (x_1, x_2, \dots, x_n)^T \in R^n \end{aligned}$$

其可行（容许）解集

$$D = \left\{ x \mid g_i(x) \geq 0, i = 1, 2, \dots, m, x \in R^n \right\}$$

**定义1** 设  $x \in D$ 。  $\forall i$ ，若  $g_i(x)=0$ ，则称第  $i$  个约束是起作用约束或积极约束。并记

$$I(x) = \{ i \mid g_i(x)=0 \quad i=1,2,\dots,m \}$$

叫起作用约束集或积极约束集。

**例2** 设约束  $s_1(x) = 1 - x_1^2 - x_2^2 \geq 0$

$$s_2(x) = x_1 - x_2 \geq 0$$

$$s_3(x) = x_1 \geq 0, \quad s_4(x) = x_2 \geq 0$$

而  $x_A = (1, 0)^T$ ，  $x_B = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})^T$ ， 则

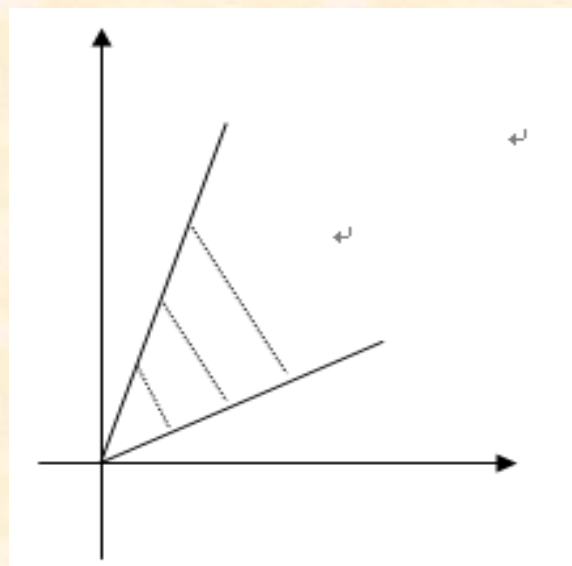
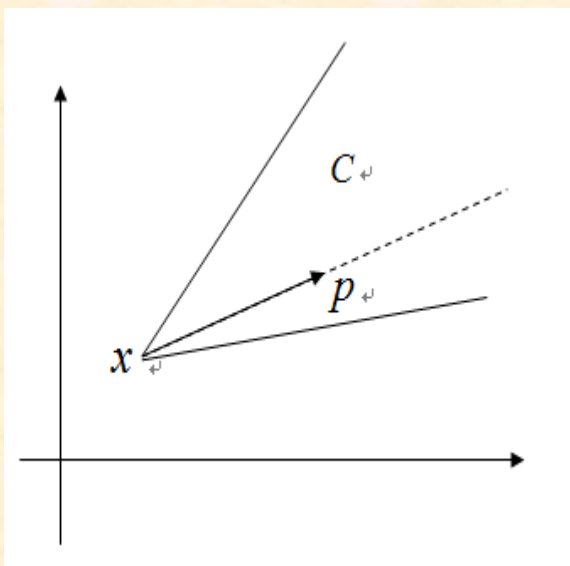
$$I(x_A) = \{1, 4\}, \quad I(x_B) = \{1, 2\}$$

**定义2** 设集合  $C \subset R^n$ ,  $x \in C$ 。若  $\forall p \in R^n$ , 有

$$x + p \in C \Rightarrow x + t \cdot p \in C \quad \forall t \geq 0$$

则称集合  $C$  是一个以  $x$  为顶点的锥。

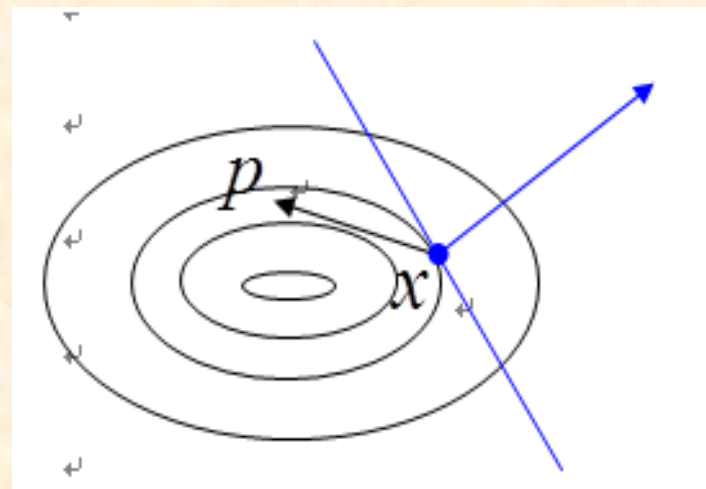
$C$  是凸集时, 又叫凸锥。  $x=0$  时, 就简称为锥。



## 1、下降方向 与 可行方向

又设  $x \in D$ 。则向量集：

$$S = \{ p \mid \nabla f(x)^T p < 0 \}$$



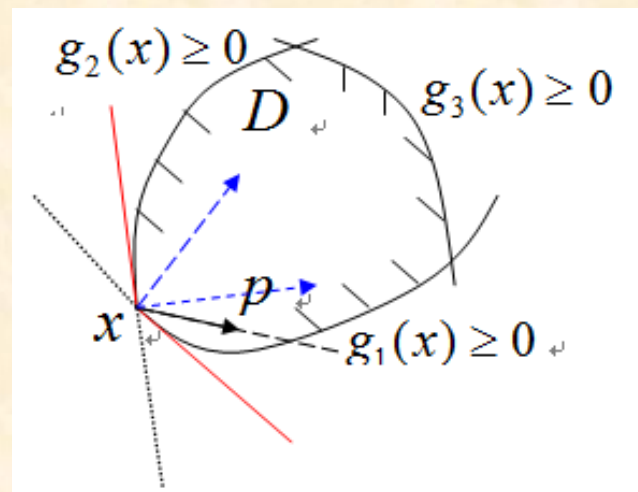
就是目标函数在  $x \in D$  处的下降方向集。

下降方向集是半空间，从而也是锥。

**定义3** 设  $D \subset R^n$  非空,  $x \in D$ 。对于非零向量  $p$ , 若  $\exists \delta > 0$ , 使  $\forall t \in (0, \delta)$ , 必有

$$x + t \cdot p \in D$$

则称  $p$  叫集合  $D$  在  $x$  处的一个  
**可行（容许）方向**。而集合

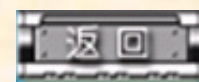
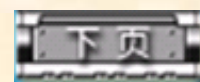


$$G(x) = \{ p \neq 0 \mid \exists \delta > 0, \text{ s.t. } \forall t \in (0, \delta), x + t \cdot p \in D \}$$

叫集合  $D$  在  $x$  处的**可行（容许）方向集**。

可行（容许）方向集是锥。

*Optimization*

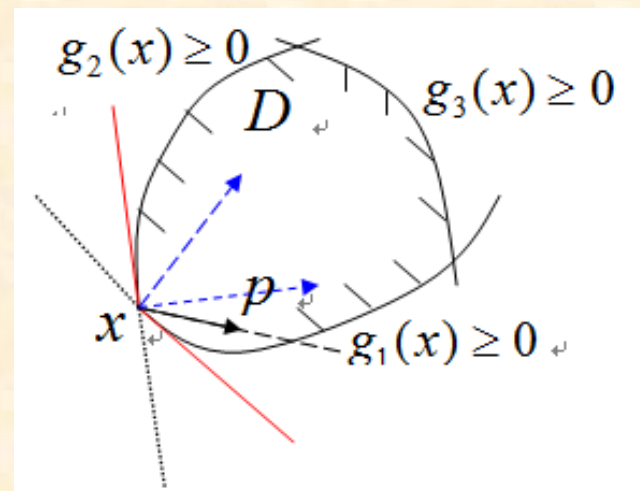


$$I(x) = \{ i \mid g_i(x) = 0 \quad i = 1, 2, \dots, m \}$$

**引理1** 设问题( II )的(不等式)约束函数在点  $x$  处:

①  $\forall i = 1, 2, \dots, m$ , 当  $i \notin I(x)$  时,  
 $g_i(x)$  在点  $x$  处**连续**;

②  $\forall i = 1, 2, \dots, m$ , 当  $i \in I(x)$  时,  
 $g_i(x)$  在点  $x$  处**可微**。



则对任意  $i \in I(x)$  满足:  $\nabla g_i(x)^T \cdot p > 0$

的方向  $p$  必定是集合  $D$  在  $x$  处的一个**可行 (容许) 方向**。  
 即

$$\{ p \mid \nabla g_i(x)^T \cdot p > 0, \forall i \in I(x) \} \subseteq G(x)$$

$$I(x) = \{ i \mid g_i(x) = 0 \quad i = 1, 2, \dots, m \}$$

**证明** 设方向  $p$  满足:

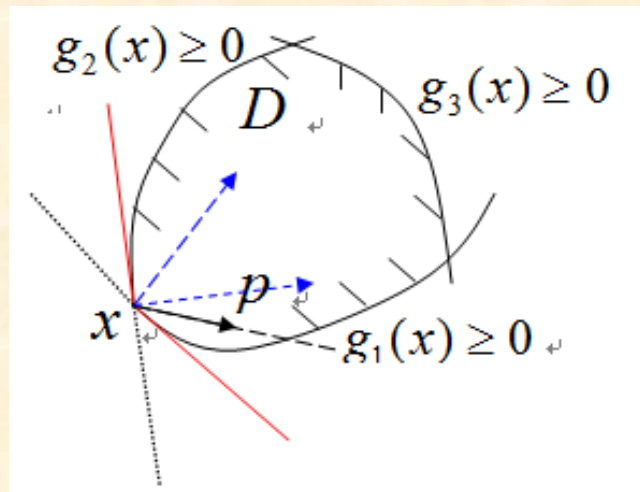
$$\nabla g_i(x)^T \cdot p > 0 \quad i \in I(x)$$

①  $\forall i = 1, 2, \dots, m$ , 当  $i \notin I(x)$  时,  $g_i(x) > 0$ ,  
由连续函数的性质知, 存在  $\delta_1 > 0$ ,  
使得当  $t \in (0, \delta_1)$ , 有

$$g_i(x + t \cdot p) > 0 \quad i \notin I(x)$$

②  $\forall i = 1, 2, \dots, m$ , 当  $i \in I(x)$  时, 由 Taylor 展式

$$\begin{aligned} g_i(x + t \cdot p) &= g_i(x) + \nabla g_i(x)^T \cdot t \cdot p + o(\|t \cdot p\|) \\ &= \nabla g_i(x)^T \cdot t \cdot p + o(\|t \cdot p\|) \end{aligned}$$



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/467164110102010032>