最优化理论与方法

第四章　约束规划的最优性条件

上海财经大学应用数学系

December 26, 2008

约束最优化问题的一般形式

约束最优化问题的一般形式如下:

$$
\begin{aligned}
min \quad & f(\mathbf{x}), \ \ \mathbf{x} \in R^n, \\
s.t. \quad & c_i(\mathbf{x}) = 0, i \in E = \{1, 2, \cdots, l\}, \\
& c_i(\mathbf{x}) \leq 0, i \in I = \{l+1, l+2, \cdots, l+m\}.
\end{aligned}
\tag{4.1}
$$

问题(4.1)的可行域

$$D = \{\mathbf{x} \mid c_i(\mathbf{x}) = 0, i \in E, c_i(\mathbf{x}) \leq 0, i \in I\}.$$

局部解、最优解

定义4.1.1

对于约束最优化问题(4.1)，若对$\mathbf{x}^* \in D$,存在$\varepsilon > 0$,使当$x \in D$且$\| \mathbf{x} - \mathbf{x}^* \| \leq \varepsilon$时，总有

$$f(\mathbf{x}) \geq f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题(4.1)的局部解，或简称$\mathbf{x}^*$为最优解。若当$\mathbf{x} \in D$且$0 < \| \mathbf{x} - \mathbf{x}^* \| \leq \varepsilon$，总有

$$f(\mathbf{x}) > f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题的严格局部解。

局部解、最优解

定义4.1.2

对于约束最优化问题(4.1)，若对$\mathbf{x}^* \in D$,总有

$$f(\mathbf{x}) \geq f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题(4.1)的全局最优解. 若当$\mathbf{x} \in D, x \neq x^*$，总有

$$f(\mathbf{x}) > f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题(4.1)的严格全局最优解。

显然，全局最优解必是局部最优解，但反之不然。特别地，若$D = R^n$,则以上有关最优解的定义便是无约束问题相应最优解的定义。

局部解、最优解

定义4.1.2

对于约束最优化问题(4.1)，若对$\mathbf{x}^* \in D$,总有

$$f(\mathbf{x}) \geq f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题(4.1)的全局最优解. 若当$\mathbf{x} \in D, x \neq x^*$，总有

$$f(\mathbf{x}) > f(\mathbf{x}^*),$$

则称$\mathbf{x}^*$为约束问题(4.1)的严格全局最优解。

显然，全局最优解必是局部最优解，但反之不然。特别地，若$D = R^n$,则以上有关最优解的定义便是无约束问题相应最优解的定义。

例题4.1.1

试确定约束问题

$$\begin{array}{ll} min & f(\mathbf{x}) = x_1^2 + x_2^2, \\ s.t. & c_1(\mathbf{x}) = (x_1 - 1)^2 - x_2^2 - 4 = 0, \\ & c_2(\mathbf{x}) = x_2 - 1 \leq 0 \end{array}$$

的局部解和全局解。

由图4.1.1可以看出问题的局部解$\mathbf{x}^* = (-1, 0)^T$和$\hat{\mathbf{x}}^* = (3, 0)^T$，其中$\mathbf{x}^*$是全局解.

例题4.1.1

试确定约束问题

$$\begin{array}{ll} min & f(\mathbf{x}) = x_1^2 + x_2^2, \\ s.t. & c_1(\mathbf{x}) = (x_1 - 1)^2 - x_2^2 - 4 = 0, \\ & c_2(\mathbf{x}) = x_2 - 1 \leq 0 \end{array}$$

的局部解和全局解。

由图4.1.1可以看出问题的局部解$\mathbf{x}^* = (-1, 0)^T$和$\hat{\mathbf{x}}^* = (3, 0)^T$，其中$\mathbf{x}^*$是全局解.

约束问题局部解的一阶必要条件

1、 设$\mathbf{x}^*$是约束问题(4.1)的可行点，$\{\mathbf{x}^{(k)}\}$是约束问题(4.1)的可行点列，并满足$\mathbf{x}^{(k)} \to \mathbf{x}^* (k \to \infty)$且$\mathbf{x}^{(k)} \neq \mathbf{x}^*$记

$$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}$$

其中$\mathbf{d}^{(k)}$有固定模长，而$\delta_k > 0$，由于$\mathbf{x}^{(k)} \to \mathbf{x}^*$，因此有$\delta_k \to 0$.(例如，$\mathbf{d}^{(k)} = \frac{\mathbf{x}^{(k)} - \mathbf{x}^*}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|}, \delta_k = \| \mathbf{x}^{(k)} - \mathbf{x}^* \|$)

2、 若$\{\mathbf{d}^{(k)}\}$有极限，即$\mathbf{d}^{(k)} \to \mathbf{d}$，则称$\{\mathbf{x}^{(k)}\}$为方向可行点列，$\mathbf{d}^{(k)}$为方向序列，而称$\mathbf{d}$为$\mathbf{x}^*$处的可行方向(Feasible Direction)。

3、 记

$$\mathbf{FD}^* = \mathbf{FD}(\mathbf{x}^*) = \{\mathbf{d} \mid \mathbf{d}\text{是}\mathbf{x}^*\text{处œ1}\}$$

为全体可行方向的集合，因此，由定义得知，若$\mathbf{d} \in \mathbf{FD}^*$的充分必要条件是：存在可行点列$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}, \| \mathbf{d}^{(k)} \| = \| \mathbf{d} \|, \mathbf{d}^{(k)} \to \mathbf{d}$和$\delta_k \to 0$。

约束问题局部解的一阶必要条件

1、 设$\mathbf{x}^*$是约束问题(4.1)的可行点，$\{\mathbf{x}^{(k)}\}$是约束问题(4.1)的可行点列，并满足$\mathbf{x}^{(k)} \to \mathbf{x}^* (k \to \infty)$且$\mathbf{x}^{(k)} \neq \mathbf{x}^*$记

$$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}$$

其中$\mathbf{d}^{(k)}$有固定模长，而$\delta_k > 0$,由于$\mathbf{x}^{(k)} \to \mathbf{x}^*$，因此有$\delta_k \to 0$.(例如，$\mathbf{d}^{(k)} = \frac{\mathbf{x}^{(k)} - \mathbf{x}^*}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|}, \delta_k = \| \mathbf{x}^{(k)} - \mathbf{x}^* \|$)

2、 若$\{\mathbf{d}^{(k)}\}$有极限，即$\mathbf{d}^{(k)} \to \mathbf{d}$，则称$\{\mathbf{x}^{(k)}\}$为方向可行点列，$\mathbf{d}^{(k)}$为方向序列，而称$\mathbf{d}$为$\mathbf{x}^*$处的可行方向(Feasible Direction)。

3、 记

$$\mathrm{FD}^* = \mathrm{FD}(\mathbf{x}^*) = \{\mathbf{d} \mid \mathbf{d}\text{是}\mathbf{x}^*\text{处œ1}\}$$

为全体可行方向的集合，因此，由定义得知，若$\mathbf{d} \in \mathrm{FD}^*$的充分必要条件是：存在可行点列$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}, \| \mathbf{d}^{(k)} \| = \| \mathbf{d} \|, \mathbf{d}^{(k)} \to \mathbf{d}$和$\delta_k \to 0$。

约束问题局部解的一阶必要条件

1、设$\mathbf{x}^*$是约束问题(4.1)的可行点，$\{\mathbf{x}^{(k)}\}$是约束问题(4.1)的可行点列，并满足$\mathbf{x}^{(k)} \to \mathbf{x}^* (k \to \infty)$且$\mathbf{x}^{(k)} \neq \mathbf{x}^*$记

$$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}$$

其中$\mathbf{d}^{(k)}$有固定模长，而$\delta_k > 0,$由于$\mathbf{x}^{(k)} \to \mathbf{x}^*$，因此有$\delta_k \to 0.$(例如，$\mathbf{d}^{(k)} = \frac{\mathbf{x}^{(k)} - \mathbf{x}^*}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|}, \delta_k = \| \mathbf{x}^{(k)} - \mathbf{x}^* \|$)

2、 若$\{\mathbf{d}^{(k)}\}$有极限，即$\mathbf{d}^{(k)} \to \mathbf{d}$，则称$\{\mathbf{x}^{(k)}\}$为方向可行点列，$\mathbf{d}^{(k)}$为方向序列，而称$\mathbf{d}$为$\mathbf{x}^*$处的可行方向(Feasible Direction)。

3、记

$$\mathbf{FD}^* = \mathbf{FD}(\mathbf{x}^*) = \{\mathbf{d} \mid \mathbf{d}是\mathbf{x}^*处œ1\}$$

为全体可行方向的集合，因此，由定义得知，若$\mathbf{d} \in \mathbf{FD}^*$的充分必要条件是：存在可行点列$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}, \| \mathbf{d}^{(k)} \| = \| \mathbf{d} \|, \mathbf{d}^{(k)} \to \mathbf{d}$和$\delta_k \to 0$。

例4.2.1

考虑可行域

$$D=\{\mathbf{x}=(x_1,x_2)^T \mid x_2-x_1^3\le 0,-x_2\le 0\} \tag{4.2.1}$$

设$\mathbf{x}^*=(0,0)^T$, $\{x^{(k)}\}$为曲线$x_2=x_1^3$上满足$x_2^{(k)}>0$的一个趋于$\mathbf{x}^*$的点列，则

$$\mathbf{d}^{(k)}=\frac{\mathbf{x}^{(k)}-\mathbf{x}^*}{\|\mathbf{x}^{(k)}-\mathbf{x}^*\|}\to(1,0)^T,$$

故$\mathbf{d}=(1,0)^T\in\mathbf{FD}^*$.

约束问题局部解的一阶必要条件

定义4.2.1

设$\hat{\mathbf{x}}$是一般约束问题(4.1)的可行点，当$i \in I$时，对某个约束有$c_i(\hat{\mathbf{x}}) = 0$，则称$c_i(\mathbf{x}) \le 0$为$\hat{\mathbf{x}}$处的有效约束(Active Constraints)，若$c_i(\hat{\mathbf{x}}) < 0$,则称约束$c_i(\mathbf{x}) \le 0$为$\hat{\mathbf{x}}$处的非有效约束。

定义有效约束指标集

$$I(\hat{\mathbf{x}}) = \{i \mid c_i(\hat{\mathbf{x}}) = 0, i \in I\}$$

简称$I(\hat{\mathbf{x}})$为$\hat{\mathbf{x}}$处的有效集(Active Set)。 设$\mathbf{x}^*$是约束问题(4.1)的可行点，定义

$$LD^* = LD(\mathbf{x}^*) = \{\mathbf{d} \mid \mathbf{d} \neq 0, \mathbf{d}^T \nabla c_i(x^*) = 0, i \in E, \mathbf{d}^T \nabla c_i(\mathbf{x}^*) \le 0, i$$

显然，$LD^* \cup \{0\}$为一锥，称为$\mathbf{x}^*$处的线性化锥

约束问题局部解的一阶必要条件

定义4.2.1

设$\hat{\mathbf{x}}$是一般约束问题(4.1)的可行点，当$i \in I$时，对某个约束有$c_i(\hat{\mathbf{x}}) = 0$，则称$c_i(\mathbf{x}) \le 0$为$\hat{\mathbf{x}}$处的有效约束(Active Constraints)，若$c_i(\hat{\mathbf{x}}) < 0$,则称约束$c_i(\mathbf{x}) \le 0$为$\hat{\mathbf{x}}$处的非有效约束。

定义有效约束指标集

$$I(\hat{\mathbf{x}}) = \{i \mid c_i(\hat{\mathbf{x}}) = 0, i \in I\}$$

简称$I(\hat{\mathbf{x}})$为$\hat{\mathbf{x}}$处的有效集(Active Set)。 设$\mathbf{x}^*$是约束问题(4.1)的可行点，定义

$$LD^* = LD(\mathbf{x}^*) = \{\mathbf{d} \mid \mathbf{d} \neq 0, \mathbf{d}^T \nabla c_i(x^*) = 0, i \in E, \mathbf{d}^T \nabla c_i(\mathbf{x}^*) \le 0, i$$

显然，$LD^* \cup \{0\}$为一锥，称为$\mathbf{x}^*$处的线性化锥。

引理4.2.1

$$\mathbf{FD}^* \subset LD^*$$

证明：设$\mathbf{d} \in LD^*$,则存在$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}$是可行点，且$\delta_k \to 0, \mathbf{d}^{(k)} \to \mathbf{d}$. 由Taylor展式

$$c_i(\mathbf{x}^{(k)}) = c_i(\mathbf{x}^*) + \delta_k \mathbf{d}^{(k)T} \nabla c_i(\mathbf{x}^*) + o(\delta_k) \tag{4.2.2}$$

当$i \in E$时，$c_i(\mathbf{x}^{(k)}) = 0, c_i(\mathbf{x}^*) = 0$,所以(4.2.2)式化简为

$$\delta_k \mathbf{d}^{(k)T} \nabla c_i(\mathbf{x}^*) + o(\delta_k) = 0 \tag{4.2.3}$$

在(4.2.3)式两端同除δ_k,并令$k \to \infty$，得到

$$\mathbf{d}^T \nabla c_i(\mathbf{x}^*) = 0$$

约束问题局部解的一阶必要条件

当$i \in I(\mathbf{x}^*)$时，$c_i(\mathbf{x}^{(k)}) \le 0, c_i(\mathbf{x}^*) = 0$，所以(4.2.2)式化简为

$$\delta_k \mathbf{d}^{(k)T} \nabla c_i(\mathbf{x}^*) + o(\delta_k) \le 0 \tag{4.2.4}$$

在(4.2.4)式两端同除δ_k，并令$k \to \infty$，得到

$$\mathbf{d}^T \nabla c_i(\mathbf{x}^*) \le 0$$

因此，$\mathbf{d} \in LD^*$.

但命题反之并不成立。

约束问题局部解的一阶必要条件

当$i \in I(\mathbf{x}^*)$时，$c_i(\mathbf{x}^{(k)}) \leq 0, c_i(\mathbf{x}^*) = 0$，所以(4.2.2)式化简为

$$\delta_k \mathbf{d}^{(k)T} \nabla c_i(\mathbf{x}^*) + o(\delta_k) \leq 0 \tag{4.2.4}$$

在(4.2.4)式两端同除δ_k，并令$k \to \infty$，得到

$$\mathbf{d}^T \nabla c_i(\mathbf{x}^*) \leq 0$$

因此，$\mathbf{d} \in LD^*$.
但命题反之并不成立。

例4.2.2

继续考虑例4.2.1中的可行域$\mathbf{D}$，可行点$\mathbf{x}^* = (0,0)^T$，则$\nabla c_1(\mathbf{x}^*) = (0,1)^T, \nabla c_2(\mathbf{x}^*) = (0,-1)^T$，故

$$\begin{aligned} LD^* &= \{\mathbf{d} = (d_1,d_2)^T \mid \mathbf{d} \neq 0, \mathbf{d}^T\nabla c_1(\mathbf{x}^*) \leq 0, \mathbf{d}^T\nabla c_2(\mathbf{x}^*) \leq 0\} \\ &= \{\mathbf{d} = (d_1,d_2)^T \mid d_1 \neq 0, d_2 = 0\} \\ &= \{(d_1,0)^T \mid d_1 \neq 0\} \end{aligned}$$

约束问题局部解的一阶必要条件

显然，$\mathbf{d}=(-1,0)^T \in LD^*$，但由式(4.2.1)确定的可行域D对于所有的可行点$\mathbf{x}^{(k)}$有$x_1^{(k)} \geq 0$，故不存在这样的点列$\mathbf{x}^{(k)}=\mathbf{x}^*+\delta_k\mathbf{d}^{(k)}=\delta_k\mathbf{d}^{(k)}$，使得$\mathbf{d}^{(k)} \to \mathbf{d}=(-1,0)^T$，因此，$\mathbf{d} \notin \mathbf{FD}^*$.

因此，要使$LD^*=FD^*$，需要对约束附加条件，通常称任何一个保证$LD^*=FD^*$成立的条件为约束限制条件（Constraint Qualification）.

约束问题局部解的一阶必要条件

显然，$\mathbf{d} = (-1, 0)^T \in LD^*$，但由式(4.2.1)确定的可行域D对于所有的可行点$\mathbf{x}^{(k)}$有$x_1^{(k)} \geq 0$，故不存在这样的点列$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)} = \delta_k \mathbf{d}^{(k)}$，使得$\mathbf{d}^{(k)} \to \mathbf{d} = (-1, 0)^T$，因此，$\mathbf{d} \notin \mathbf{FD}^*$.

因此，要使$LD^* = FD^*$，需要对约束附加条件，通常称任何一个保证$LD^* = FD^*$成立的条件为约束限制条件（Constraint Qualification）.

约束问题局部解的一阶必要条件

定理4.2.2

若$\mathbf{x}^*$是约束问题(4.1)的局部解，则

$$\mathbf{FD}^* \cap \mathbf{DD}^* =$$

证明　任取$d \in FD^*$，则存在可行点列$\mathbf{x}^{(k)} = \mathbf{x}^* + \delta_k \mathbf{d}^{(k)}$，并且$\delta_k \to 0$和$\mathbf{d}^{(k)} \to \mathbf{d}$.有Taylor展式

$$f(\mathbf{x}^{(k)}) = f(\mathbf{x}^*) + \delta_k \mathbf{d}^{(k)T} \nabla f(\mathbf{x}^*) + o(\delta_k) \tag{4.2.5}$$

因为$\mathbf{x}^*$是局部解，存在着K，当$k \geq K$时，有$f(\mathbf{x}^{(k)}) \geq f(\mathbf{x}^*)$，由(4.2.5)式得到

$$\delta_k \mathbf{d}^{(k)T} \nabla f(\mathbf{x}^*) + o(\delta_k) \geq 0, \tag{4.2.6}$$

在(4.2.6)式两端同除δ_k,并令$k \to \infty$,得到

$$\mathbf{d}^T \nabla f(\mathbf{x}^*) \geq 0$$

故$\mathbf{d} \notin \mathbf{DD}^*$.

约束问题局部解的一阶必要条件

定理4.2.3 Fakars引理

设$A \in R^{m\times n}$和$w \in R^n$. 系统I:存在d满足

$$Ad \leq 0, w^T d > 0.$$

系统II: 存在非负向量$y = (y_1, y_2, \cdots, y_m)^T$，使得

$$w = A^T y$$

则两系统有且仅有一个有解。

约束问题局部解的一阶必要条件

证明：下面分两种情况来讨论。

(1)若系统II有解，则系统I无解。

设系统II有解，即存在$y=(y_1,y_2,\cdots,y_m)^T$ 且$y_i\geq 0, i=1,2,\cdots,m$，使得

$$w=A^Ty$$

若系统I有解，则有

$$0<w^Td=y^TAd\leq 0.$$

矛盾，因此系统I无解。

(2)若系统II无解，则系统I有解。设系统II无解，构造集合

$$C = \{v \mid v = A^T y,\ y_i \geq 0, i = 1, 2 \cdots, m\},$$

显然C是非空闭凸集。系统II无解表明，$w \notin C$, 由定理1.4.9，则存在d, β满足

$$d^T x \leq \beta < d^T w, \quad \forall x \in C$$

注意到$0 \in C$, 从上式知$\beta \geq 0$,因而$d^T w > 0$. 由此可得:

$$\beta \geq d^T x = d^T A^T y = (y^T A d)^T = y^T A d, \quad \forall y \geq 0$$

约束问题局部解的一阶必要条件

由于$y \geq 0$可以任意大，故$Ad \leq 0$. 又

$$0 \leq \beta < d^T w$$

知

$$d^T w > 0$$

所以d是系统I的解，产生矛盾。

约束问题局部解的一阶必要条件

定理4.2.4

设约束问题(4.1)中$f(\mathbf{x}), c_i(\mathbf{x})(i=1,2,\cdots,l+m)$ 具有连续的一阶偏导数，若$\mathbf{x}^*$是约束问题(4.1)的局部解，并且在$\mathbf{x}^*$处约束限制条件成立（即$LD^*=\mathbf{FD}^*$），则存在常数$\lambda^*=(\lambda_1^*,\lambda_2^*,\cdots,\lambda_{l+m}^*)^T$，使得

$$
\begin{aligned}
&\nabla_x L(\mathbf{x}^*,\lambda^*)=\nabla f(\mathbf{x}^*)+\sum_{i=1}^{l+m}\lambda_i^*\nabla c_i(\mathbf{x}^*)=0,\\
&c_i(\mathbf{x}^*)=0,\quad i\in E=\{1,2,\cdots,l\},\\
&c_i(\mathbf{x}^*)\le 0,\quad i\in I=\{l+1,l+2,\cdots,l+m\},\\
&\lambda_i^*\ge 0,\quad i\in I=\{l+1,l+2,\cdots,l+m\},\\
&\lambda_i^* c_i(\mathbf{x}^*)=0,\quad i\in I=\{l+1,l+2,\cdots,l+m\},
\end{aligned}
\tag{4.2.7}
$$

约束问题局部解的一阶必要条件

其中$L(\mathbf{x},\lambda)$为Lagrange函数

$$L(\mathbf{x},\lambda)=f(\mathbf{x})+\sum_{i=1}^{l+m}\lambda_i c_i(\mathbf{x})$$

证明:因为$\mathbf{x}^*$是约束问题(4.1)的局部解，由定理4.2.2得到$FD^*\cap DD^* =$ ，再由约束限制条件得到$LD^*\cap DD^* =$ ，因此系统

$$\begin{cases}\nabla c_i(x^*)^T d=0, & i\in E,\\ \nabla c_i^T d\le 0, & i\in I(\mathbf{x}^*),\\ \nabla f(x^*)^T d>0 & \end{cases}$$

无解

约束问题局部解的一阶必要条件

为了利用Farkas 引理，将第一个等式$\nabla c_i(x^*)^T d = 0$改写为：

$$\nabla c_i(x^*)^T d \le 0, \text{且} -\nabla c_i(x^*)^T d \le 0, i \in E.$$

则系统(I)可以写成

$$Ad \le 0, \quad -\nabla f(x^*)^T d > 0$$

其中A是以

$$\nabla c_i(x^*)^T (i \in E), -\nabla c_i(x^*)^T (i \in E), \nabla c_i(x^*)^T (i \in I(x^*))$$

为行的矩阵，由Farkas引理知，存在向量

$$y = (u^{*+T}, u^{*-T}, \sigma^{*T})^T$$

满足

$$A^T y = -\nabla f(x^*)$$

约束问题局部解的一阶必要条件

即

$$-\nabla f(x^*) = \sum_{i\in E} {u_i}^{*+}\nabla c_i(x^*) + \sum_{i\in E} u_i^{*-}[-\nabla c_i(x^*)] + \sum_{i\in I(x^*)} \sigma_i^*\nabla c_i(x^*)$$

也即

$$\nabla f(x^*) + \sum_{i\in E}({u_i}^{*+} - u_i^{*-})\nabla c_i(x^*) + \sum_{i\in I(x^*)} \sigma_i^*\nabla c_i(x^*) = 0$$

令

$$\begin{array}{l} {\lambda_i}^* = {u_i}^{*+} - u_i^{*-}, \quad i \in E \\ {\lambda_i}^* = \sigma_i^*, \quad i \in I(x^*) \\ {\lambda_i}^* = 0, \quad i \in I \setminus I(x^*) \end{array}$$

显然满足式(4.2.7)，故定理得证.

约束问题局部解的一阶必要条件

K-T条件

由于这一定理是Kuhn和Tucker（1951）给出的，因此称上述一阶必要条件为Kuhn-Tucker条件（Kuhn-Tucker Conditions），或简称为K-T条件。称满足式（4.2.7）的点为Kuhn-Tucker点，或简称为K-T点，称λ^*为Lagrange函数，称λ^*为x^*处的Lagrange乘子（Lagrang Multiplies）。

请注意，在定理4.2.4中增加了约束限制条件，若无此条件，则局部解不一定是K-T点。

约束问题局部解的一阶必要条件

K-T条件

由于这一定理是Kuhn和Tucker（1951）给出的，因此称上述一阶必要条件为Kuhn-Tucker条件（Kuhn-Tucker Conditions），或简称为K-T条件。称满足式（4.2.7）的点为Kuhn-Tucker点，或简称为K-T点，称λ^*为Lagrange函数，称λ^*为x^*处的Lagrange乘子（Lagrang Multiplies）。

请注意，在定理4.2.4中增加了约束限制条件，若无此条件，则局部解不一定是K-T点。

约束问题局部解的一阶必要条件

反例1

考虑约束问题

$$\begin{aligned} min \quad & f(\mathbf{x}) = x_1, \\ s.t. \quad & c_1(\mathbf{x}) = x_2 - x_1^3 \leq 0, \\ & c_2(\mathbf{x}) = -x_2 \leq 0, \end{aligned}$$

因为$x_1^3 \geq x_2 \geq 0$，所以$\mathbf{x} = (0,0)^T$是约束问题的局部解。由例4.2.2知$LD^* \neq FD^*$，约束限制条件不成立，下面验证x^*不是K-T点. 考虑Lagrange函数

$$\begin{aligned} L(\mathbf{x}, \lambda) &= f(\mathbf{x}) + \lambda_1 c_1(x) + \lambda_2 c_2(x) \\ &= x_1 + \lambda_1(x_2 - x_1^3) - \lambda_2 x_2, \end{aligned}$$

约束问题局部解的一阶必要条件

因此，对一切$\lambda = (\lambda_1, \lambda_2)^T$，

$$\nabla_x L(\mathbf{x}^*, \lambda) = \nabla f(x^*) + \lambda_1 \nabla c_1(x^*) + \lambda_2 \nabla c_2(x^*) = \begin{pmatrix} 1 \\ \lambda_1 - \lambda_2 \end{pmatrix} \neq 0$$

所以x^*不是K-T点

反例2

考虑

$$\begin{array}{ll} min & f(\mathbf{x}) = x_2, \\ s.t. & c(\mathbf{x}) = -x_1^2 - x_2 \le 0, \end{array}$$

K-T点应满足方程组

$$\begin{array}{ll} -2\lambda x_1 = 0, & (a) \\ 1 - \lambda = 0, & (b) \\ -x_1^2 - x_2 \le 0, & (c) \\ \lambda \ge 0, & (d) \\ \lambda(x_1^2 + x_2) = 0, & (e) \end{array}$$

由(b)得到$\lambda = 1$,由(e)得到$x_1^2 + x_2 = 0$,再由(a)得到$x_1 = 0$,因此K-T点为$\mathbf{x}^* = (0,0)^T$,相应的乘子为$\lambda^* = 1$. 但$\mathbf{x}^* = (0,0)^T$不是约束问题的局部解。

定理4.2.5

若在约束问题(4.1)的局部解$\mathbf{x}^*$处下述两条件之一成立：(1)$c_i(\mathbf{x})(i \in E \cup I(\mathbf{x}^*))$是线性函数；
(2)$\nabla c_i(\mathbf{x}^*)(i \in E \cup I(\mathbf{x}^*))$线性无关；则在$\mathbf{x}^*$处有

$$FD^* = LD^*$$

证明:由引理(4.2.1)有

$$FD(x^*, D) \subset LD(x^*, D),$$

故只需在(1)(2)两种条件下证明

$$LD(x^*, D) \subset FD(x^*, D).$$

我们只证明情况(1).

约束限制条件

任取

$$d \in LD(x^*, D) = \{d | d \neq 0, d^T \nabla c_i(x^*) = 0,\ \ i \in E;$$

$$d^T \nabla c_i(x^*) \geq 0, i \in I(x^*)\}.$$

令

$$x^k = x^* + \alpha_k d^k, \text{其中} d^k = d, \alpha_k > 0, \alpha_k \rightarrow 0 (k \longrightarrow \infty).$$

由于$c_i(x)(i \in E \cup I(x^*))$为线性函数，将$c_i(x^*)$在$x^*$处一阶泰勒展开，便得到$x^k$是可行点，故

$$d \in FD(x^*, D).$$

约束限制条件

定理4.2.6

设x^*是问题(4.1)的一个局部最优解，$c_i(\mathbf{x})(i \in E \cup I(\mathbf{x}^*))$是线性函数或$\nabla c_i(\mathbf{x}^*)(i \in E \cup I(\mathbf{x}^*))$线性无关，则必存在$\lambda^*$使得K-T条件(4.2.7) 成立.

约束限制条件

例4.2.3 考虑如下约束问题

$$\begin{aligned} min \quad & f(\mathbf{x}) = x_1^2 + x_2, \\ s.t. \quad & c_1(\mathbf{x}) = x_1^2 + x_2^2 - 9 \leq 0, \\ & c_2(\mathbf{x}) = x_1 + x_2 - 1 \leq 0 \end{aligned}$$

已知$x^* = (0, -3)^T$是上述问题的最优解，则$I(x^*) = \{1\}$, 向量$\nabla c_1(x^*)$显然是线性无关，故约束限制条件成立，下面可以验证当$\lambda_1 = \frac{1}{6}, \lambda_2 = 0$时，下面K-T条件成立

$$\begin{cases} \nabla f(x^*) + \lambda_1 \nabla c_1(x^*) + \lambda_2 \nabla c_2(x^*) = 0 \\ \lambda_1, \lambda_2 \geq 0, \\ \lambda_i c_i(x^*) = 0, \ i = 1, 2 \end{cases}$$

故x^*为K-T点.

约束限制条件

例4.2.4 考虑如下约束问题

$$
\begin{aligned}
min \quad & f(\mathbf{x}) = (x_1^2 - 3)^2 + (x_2 - 2)^2, \\
s.t. \quad & c_1(\mathbf{x}) = x_1^2 + x_2^2 \le 5, \\
& c_2(\mathbf{x}) = -x_1 \le 0 \\
& c_3(x) = -x_2 \le 0 \\
& c_4(x) = x_1 + 2x_2 = 4.
\end{aligned}
$$

已知$x^* = (2,1)^T$是上述问题的最优解，则$I(x^*) = \{1\}$，向量$\nabla c_1(x^*)$显然是线性无关，故约束限制条件成立，下面可以验证当$\lambda_1 = \frac{1}{3}, \lambda_2 = \lambda_3 = 0, \lambda_4 = \frac{2}{3}$时，下面K-T条件成立.

$$
\begin{cases}
\nabla f(x^*) + \sum_{i=1}^4 \lambda_i \nabla c_i(x^*) = 0 \\
\lambda_i \ge 0, i = 1,2,3,4 \\
\lambda_i c_i(x^*) = 0, \ \ i = 1,2,3
\end{cases}
$$

故x^*为K-T点.

二阶充分条件

定理4.3.1 约束问题局部解的二阶充分条件

考虑一般约束问题(4.1)，设$f(\mathbf{x})$、$c_i(\mathbf{x})(i \in E \cup I)$具有连续的二阶偏导数，若存在$x^*$满足下列条件：(1)K-T条件成立，即存在$\lambda^* = (\lambda_1^*, \lambda_2^*, \cdots, \lambda_{l+m}^*)^T$ 使得

$$
\begin{aligned}
&\nabla_x L(\mathbf{x}^*, \lambda^*) = \nabla f(\mathbf{x}^*) + \sum_{i=1}^{l+m} \lambda_i^* \nabla c_i(\mathbf{x}^*) = 0,\\
&c_i(\mathbf{x}^*) = 0, \quad i \in E = \{1, 2, \cdots, l\},\\
&c_i(\mathbf{x}^*) \le 0, \quad i \in I = \{l+1, l+2, \cdots, l+m\},\\
&\lambda_i^* \ge 0, \quad i \in I = \{l+1, l+2, \cdots, l+m\},\\
&\lambda_i^* c_i(\mathbf{x}^*) = 0, \quad i \in I = \{l+1, l+2, \cdots, l+m\},
\end{aligned}
$$

二阶充分条件

约束问题局部解的二阶充分条件

且λ_i^*和$c_i(x^*)(i \in I)$不同时为0（称为严格松弛互补条件）.
(2)对于任意的$d \in M$,有

$$d^T \nabla_x{}^2 L(x^*, \lambda^*) d > 0, \tag{4.3.1}$$

其中

$$M = \{d | d \neq 0, \nabla c_i(x^*)^T d = 0, i \in E \cup I(x^*)\},$$

$I(x^*)$是x^*处的有效约束指标集，则x^*是约束问题（4.1）的严格局部解

证明:反证法.若x^*不是约束问题的严格局部解，则存在可行点列$\{x^{(k)}\}, x^{(k)} \longrightarrow x^*$, 使得

$$f(x^{(k)}) \leq f(x^*). \tag{4.3.2}$$

二阶充分条件

令

$$x^{(k)} = x^* + \delta_k d^{(k)},$$

其中

$$\delta_k = \|x^{(k)} - x^*\|, \mathbf{d}^{(k)} = \frac{\mathbf{x}^{(k)} - \mathbf{x}^*}{\| \mathbf{x}^{(k)} - \mathbf{x}^* \|},$$

因此，$\|d^{(k)}\| = 1$和$\delta_k \longrightarrow 0$.
因为$d^{(k)}$有界，必有收敛子列，不妨仍记为$d^{(k)}$，即$d^{(k)} \longrightarrow d$.
由Taylor展开式和式(4.3.2),得到

$$0 \geq f(x^{(k)}) - f(x^*) = \delta_k \nabla f(x^*)^T d^{(k)} + o(\delta_k), \tag{4.3.3}$$

在式(4.3.3)两端同除δ_k,并令$k \longrightarrow \infty$,得到

$$\nabla f(x^*)^T d \leq 0. \tag{4.3.4}$$

二阶充分条件

类似的方法可以得到

$$\nabla c_i(x^*)^T d = 0, i \in E, \tag{4.3.5}$$

$$\nabla c_i(x^*)^T d \leq 0, i \in I(x^*). \tag{4.3.6}$$

(4.3.6)表明，(1)对一切$i \in I(x^*)$,$\nabla c_i(x^*)^T d = 0$.或者，(2)存在$g \in I(x^*)$,使得$\nabla c_g(x^*)^T d < 0$.
下面证明(1)、(2)两种情况均不会出现。
若(1)成立，则$d \in M$.考虑Lagrange函数

$$L(x, \lambda) = f(x) + \sum_{i=1}^{l+m} \lambda_i c_i(x),$$

在x^*处的Taylor展开式

$$L(x^{(k)}, \lambda^*) = L(x^*, \lambda^*) + \delta_k \nabla_x L(x^*, \lambda^*)^T d^{(k)} + \frac{1}{2}\delta_k^2 d^{(k)T} {\nabla_x}^2 L(x^*, \lambda^*) d^{(k)} \tag{4.3.7}$$

二阶充分条件

注意到

$$
\begin{aligned}
L(x^{(k)},\lambda^*) &= f(x^{(k)}) + \sum_{i=1}^{l+m} \lambda_i^* c_i(x^{(k)}) \\
&= f(x^{(k)}) + \sum_{i\in I(x^*)} \lambda_i^* c_i(x^{(k)}) \\
&\le f(x^{(k)})
\end{aligned}
$$

和

$$
L(x^*,\lambda^*) = f(x^*) + \sum_{i=1}^{l+m} \lambda_i^* c_i(x^*) = f(x^*),
$$

因此，（4.3.7）式可写成

$$
\begin{aligned}
&0 \ge f(x^{(k)}) - f(x^*) \ge L(x^{(k)},\lambda^*) - L(x^*,\lambda^*) \\
&= \tfrac{1}{2}\delta_k^2 d^{(k)T} \nabla_x{}^2 L(x^*,\lambda^*) d^{(k)} + o(\delta_k^2)
\end{aligned}
\tag{4.3.8}
$$

二阶充分条件

在(4.3.8)式两端同除δ_k^2,并令$k \longrightarrow \infty$,得到

$$d^T {\nabla_x}^2 L(x^*, \lambda^*) d \le 0,$$

与(4.3.1)式矛盾。
再假设(2)式成立。由一阶必要条件和(4.3.5)式得到

$$\begin{aligned} \nabla f(x^*)^T d &= -\sum_{i=1}^{l+m} \lambda_i^* \nabla c_i(x^*)^T d \\ &= -\sum_{i \in I(x^*)} \lambda_i^* \nabla c_i(x^*)^T d \\ &\ge -\lambda_g^* \nabla c_i(x^*)^T d > 0, \end{aligned}$$

与(4.3.4)式矛盾。

二阶充分条件

例4.3.1

用约束问题局部解的一阶必要条件和二阶充分条件求解约束问题

$$\begin{array}{ll} min & f(\mathbf{x}) = x_1 x_2, \\ s.t. & c(\mathbf{x}) = x_1^2 + x_2^2 - 1 = 0, \end{array}$$

解： 由约束问题的Lagrange函数得

$$L(x,\lambda) = x_1 x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

根据约束问题的一阶必要条件得

$$\begin{cases} x_2 + 2\lambda x_1 = 0 & (1) \\ x_1 + 2\lambda x_2 = 0 & (2) \\ x_1^2 + x_2^2 - 1 = 0 & (3) \end{cases}$$

二阶充分条件

求解该方程，分情况讨论：1)当$\lambda = 0$时，有$x_1 = x_2 = 0$，与(3)矛盾
2)当$\lambda \neq 0$时，由(1)与(2)知

$$\begin{cases} (1-4\lambda^2)x_1 = 0 \\ (1-4\lambda^2)x_2 = 0 \end{cases}$$

解之，得到$\lambda = -\frac{1}{2}$,或$\lambda = \frac{1}{2}$
相应的K-T点为$(x_1, x_2) = (-\frac{\sqrt{\pi}}{2}, -\frac{\sqrt{\pi}}{2})$, $(x_1, x_2) = (\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2})$, $(x_1, x_2) = (-\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2})$, $(x_1, x_2) = (\frac{\sqrt{\pi}}{2}, -\frac{\sqrt{\pi}}{2})$

下面验证二阶充分条件:
(a) $\lambda = -\frac{1}{2}$时, Lagrange函数在(x^*, λ^*)处的Hesse矩阵为

$$\nabla_x^2 L(\mathbf{x}^*, \lambda^*) = \begin{pmatrix} -1, & 1 \\ 1, & -1 \end{pmatrix}$$

不是正定矩阵.考虑集合

$$\begin{aligned} M &= \{(\alpha, \beta)^T | (\alpha, \beta)^T \neq 0, \sqrt{2}\alpha + \sqrt{2}\beta = 0\} \\ &= \{(\alpha, -\alpha)^T | \alpha \neq 0\} \end{aligned}$$

对于$d \in M$,有

$$d^T \nabla_x{}^2 L(x^*, \lambda^*) d = (\alpha, -\alpha) \begin{pmatrix} -1, & 1 \\ 1, & -1 \end{pmatrix} (\alpha, -\alpha)^T = -4\alpha^2 < 0,$$

因此$(x_1, x_2) = (\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2})$, $(x_1, x_2) = (-\frac{\sqrt{\pi}}{2}, -\frac{\sqrt{\pi}}{2})$不是约束问题的局部解.

二阶充分条件

(b)$\lambda = \frac{1}{2}$时,Lagrange函数在(x^*, λ^*)处的Hesse矩阵为

$$\nabla_x^2 L(\mathbf{x}^*, \lambda^*) = \begin{pmatrix} 1, & 1 \\ 1, & 1 \end{pmatrix}$$

是半正定矩阵,考虑集合

$$M = \{(\alpha, \alpha)^T | \alpha \neq 0\}$$

对于$d \in M$,有

$$d^T \nabla_x{}^2 L(x^*, \lambda^*) d = (\alpha, \alpha) \begin{pmatrix} 1, & 1 \\ 1, & 1 \end{pmatrix} (\alpha, \alpha)^T = 4\alpha^2 > 0,$$

因此$(x_1, x_2) = (\frac{\sqrt{\pi}}{2}, -\frac{\sqrt{\pi}}{2})$, $(x_1, x_2) = (-\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2})$是约束问题的最优解.

凸规划的最优性条件

凸规划

对于一般的非线性规划(4.1)，若目标函数$f(x)$是凸函数，约束集合D是凸集，则称非线性规划(4.1)是凸规划。显然，在(4.1)中如果只含不等式约束，又$c_i(x)(i \in I)$是凸函数，则约束集D是凸集。对于混合约束问题，若$c_i(x)(i \in E)$是线性函数，$c_i(x)(i \in I)$是凸函数，则D是凸集。

凸规划的最优性条件

定理4.4.1

凸规划的局部最优解必是全局最优解。

证明：设x^*是凸规划的一个局部最优解，由局部最优解的定义知，存在x^*的一个$\delta > 0$邻域$N_\delta(x^*)$,使得

$$f(x^*) \le f(x), \quad \forall x \in N_\delta(x^*) \cap D$$

成立。假定x^*不是全局最优解，则存在$\bar{x} \in D$,使得

$$f(\bar{x}) < f(x^*).$$

令$\tilde{x} = \alpha\bar{x} + (1-\alpha)x^*(0 < \alpha < 1)$,由约束集合$D$是凸集知$\tilde{x} \in D$.令取$\alpha$满足$0 < \alpha < \frac{\delta}{\|\bar{x}-\mathbf{x}^*\|}$, 则有

$$\|\tilde{x} - x^*\| = \|\alpha\bar{x} + (1-\alpha)x^* - x^*\| = \alpha\|\bar{x} - x^*\| < \delta.$$

凸规划的最优性条件

从而$\tilde{x} \in N_\delta(x^*)$, 故有$\tilde{x} \in N_\delta(x^*) \cap D$. 再由$f(x)$是凸函数，有

$$\begin{aligned} f(\tilde{x}) &= f(\alpha\bar{x} + (1-\alpha)x^*) \\ &\le \alpha f(\bar{x}) + (1-\alpha)f(x^*) \\ &< \alpha f(x^*) + (1-\alpha)f(x^*) \\ &= f(x^*) \end{aligned}$$

这与x^*是局部最优解矛盾，定理得证。

凸规划的最优性条件

定理4.4.2

设目标函数$f(x)$和约束函数$c_i(x)$一阶连续可微，若凸规划的可行点x^*是K-T点，则x^*必是最优解。

证明：设x^*是K-T点，λ^*是相应的拉格朗日乘子.由$f(x)$是凸函数，根据定理(1.4.5) 有

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*),$$

由$c_i(x)(i \in E)$是线性函数，则有

$$c_i(x) = c_i(x^*) + \nabla c_i(x^*)^T (x - x^*), i \in E,$$

由$c_i(x)(i \in I)$是凹函数，即$-c_i(x)$是凸函数，则有

$$-c_i(x) \geq -c_i(x^*) - \nabla c_i(x^*)^T (x - x^*), i \in I,$$

凸规划的最优性条件

对于任意的$x \in D$,$\sum_{i=1}^{m} \lambda^* c_i(x) \geq 0$,所以

$$
\begin{aligned}
f(x) &\geq f(x) - \sum_{i=1}^{m} \lambda_i^* c_i(x) \\
&\geq f(x^*) + \nabla f(x^*)^T (x - x^*) - \sum_{i=1}^{m} \lambda_i^* [c_i(x^*) + \nabla c_i(x^*)^T (x - x^*) \\
&= f(x^*) - \sum_{i=1}^{m} \lambda_i^* c_i(x^*) + [\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla c_i(x^*)]^T (x - x^*) \\
&= f(x^*) - \sum_{i=1}^{m} \lambda_i^* c_i(x^*) \\
&= f(x^*),
\end{aligned}
$$

即有

$$f(x) \geq f(x^*), \forall x \in D,$$

所以x^*是全局最优解。

凸规划的最优性条件

例4.4.1

求如下约束问题的K-T点，并验证它是否为该约束问题的全局最优解。

$$
\begin{aligned}
min \quad & f(\mathbf{x}) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2, \\
s.t. \quad & c_1(\mathbf{x}) = x_1^2 + x_2^2 \le 5, \\
& c_2(\mathbf{x}) = 3x_1 + x_2 \le 6
\end{aligned}
$$

解：已知$x^* = (1,2)^T$是问题的可行解，同时满足K-T条件，故该点为约束问题的K-T点。又由于目标函数为凸函数，而约束函数$c_1(\mathbf{x})$和$c_2(\mathbf{x})$一阶连续可微，从而由定理(4.4.2)可知，点$x^* = (1,2)^T$也是约束问题的全局极小点。