6-2 约束优化最优性条件

分离定理

引理6.2.1 超平面分离定理.

设C是m个n维向量$\boldsymbol{a}_i \in \mathbb{R}^n, i = 1, \ldots, m$生成的集合:

$$C = \left\{ \boldsymbol{v} \in \mathbb{R}^n \middle| \boldsymbol{v} = \sum_{i=1}^{m} \lambda_i \boldsymbol{a}_i, \lambda_i \geq 0, i = 1 \ldots, m \right\}$$

如果向量$\boldsymbol{g} \notin C$,则存在一个以$\boldsymbol{d}$为法向量的超平面将C与$\boldsymbol{g}$分离，使得:

$$\boldsymbol{g}^T \boldsymbol{d} < 0, \boldsymbol{a}_i^T \boldsymbol{d} \geq 0, i = 1, \ldots, m$$

成立。

引理6.2.2 Farkas引理.

给定任意n维向量$\boldsymbol{a}_i \in \mathbb{R}^n, i=1,\ldots,m$与$\boldsymbol{g}$，则集合:

$$\mathscr{D}_1 = \left\{ \boldsymbol{d} \in \mathbb{R}^n \middle| \ \boldsymbol{d}^T\boldsymbol{g} < 0; \boldsymbol{a}_i^T\boldsymbol{d} \geq 0, i=1,\ldots,m \ \right\}$$

为空集的充分必要条件为，存在$\lambda_i \geq 0 (i=1,\ldots,m)$,使得：

$$\boldsymbol{g} = \sum_{i=1}^{m} \lambda_i \boldsymbol{a}_i. \tag{6.2.1}$$

成立。

证明：**充分性** 对$\lambda_i \geq 0, i=1,\ldots,m$, 对满足$\boldsymbol{a}_i^T\boldsymbol{d} \geq 0, i=1,\ldots,m$的$\boldsymbol{d}$，有：

引理6.2.2证明

证明续：

$$\boldsymbol{g}^T\boldsymbol{d} = \sum_{i=1}^{m} \lambda_i \boldsymbol{a}_i^T \boldsymbol{d} \geq 0$$

故$\mathscr{D}_1$为空集。

必要性 假定$\boldsymbol{g} \notin C$，由C的定义以及分离定理知，存在以$\boldsymbol{d}$为法向量的超平面，使得：

$$\boldsymbol{g}^T\boldsymbol{d} < 0, \boldsymbol{a}_i^T\boldsymbol{d} \geq 0, i = 1, \ldots, m$$

这意味着$\mathscr{D}_1$非空，与条件矛盾，故$\boldsymbol{g} \in C$。 ■

下面的推论将Farkas引理推广至线性化可行下降方向集合。

推论6.2.3 Farkas引理推论.

设$\boldsymbol{d}, \boldsymbol{g}, \boldsymbol{a}_i \in \mathbb{R}^n, i \in \mathcal{E} \cup \mathcal{I}(\boldsymbol{x})$，集合：

$$\mathscr{D}_2 = \left\{ \boldsymbol{d} \middle| \boldsymbol{d}^T \boldsymbol{g} < 0; \begin{array}{l} \boldsymbol{a}_i^T \boldsymbol{d} = 0, i \in \mathcal{E} \\ \boldsymbol{a}_i^T \boldsymbol{d} \geq 0, i \in \mathcal{I}(\boldsymbol{x}) \end{array} \right\}$$

为空集的充分必要条件是，存在实数$\lambda_i\ (i \in \mathcal{E} \cup \mathcal{I}(\boldsymbol{x}))$使得：

$$\boldsymbol{g} = \sum_{i \in \mathcal{E} \cup \mathcal{I}(\boldsymbol{x})} \lambda_i \boldsymbol{a}_i, \lambda_i \geq 0, i \in \mathcal{I}(\boldsymbol{x}). \tag{6.2.2}$$

成立。

推论6.2.3证明

证明：因为$\boldsymbol{a}_i^T\boldsymbol{d}=0, i\in\mathcal{E}$可改写为：

$$\boldsymbol{a}_i^T\boldsymbol{d}\geq 0, -\boldsymbol{a}_i^T\boldsymbol{d}\geq 0, i\in\mathcal{E}.$$

由Farkas引理知，存在$\lambda_i^+\geq 0, \lambda_i^-\geq 0(i\in\mathcal{E})$, $\lambda_i\geq 0, (i\in\mathcal{I}(\boldsymbol{x}))$使得：

$$\begin{aligned}\boldsymbol{g}&=\sum_{i\in\mathcal{E}}\lambda_i^+\boldsymbol{a}_i-\sum_{i\in\mathcal{E}}\lambda_i^-\boldsymbol{a}_i+\sum_{i\in\mathcal{I}(\boldsymbol{x})}\lambda_i\boldsymbol{a}_i\\&=\sum_{i\in\mathcal{E}}(\lambda_i^+-\lambda_i^-)\boldsymbol{a}_i+\sum_{i\in\mathcal{I}(\boldsymbol{x})}\lambda_i\boldsymbol{a}_i\\&=\sum_{i\in\mathcal{A}(\boldsymbol{x})}\lambda_i\boldsymbol{a}_i\end{aligned}$$

这里: $\mathcal{A}(\boldsymbol{x})=\mathcal{E}\cup\mathcal{I}(\boldsymbol{x})$, $\lambda_i=\lambda_i^+-\lambda^-, i\in\mathcal{E}$. ■

KKT一阶最优性条件

定理6.2.4 KKT最优性.

设$\boldsymbol{x}^*$是问题(6.1.1)的局部极小点，设$f(\boldsymbol{x}), c_i(\boldsymbol{x})(i \in \mathcal{E} \cup \mathcal{I})$在$\boldsymbol{x}^*$的邻域内一阶连续可微。如果约束规范条件:

$$\mathscr{F}_S^* = \mathscr{F}_L^* \tag{6.2.3}$$

成立，则存在$\lambda_i^*(i \in \mathcal{E} \cup \mathcal{I})$使得:

$$\boldsymbol{g}^* = \nabla f(\boldsymbol{x}^*) = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*), \tag{6.2.4a}$$

$$c_i(\boldsymbol{x}^*) = 0, \quad i \in \mathcal{E}, \tag{6.2.4b}$$

$$c_i(\boldsymbol{x}^*) \geq 0, \quad i \in \mathcal{I}, \tag{6.2.4c}$$

$$\lambda_i^* \geq 0, \quad i \in \mathcal{I}, \tag{6.2.4d}$$

$$\lambda_i^* c_i(\boldsymbol{x}^*) = 0, \quad i \in \mathcal{I} \cup \mathcal{E}. \tag{6.2.4e}$$

定理6.2.4证明

证明： 由于$\boldsymbol{x}^*$是局部极小点，故$\boldsymbol{x}^*$可行，从而(6.2.4b)-(6.2.4c)成立。设$\boldsymbol{d} \in \mathscr{F}_S^*$,故从几何最优性条件引理6.1.2得:

$$\boldsymbol{d}^T \boldsymbol{g}^* \geq 0, \ \forall \boldsymbol{d} \in \mathscr{F}_S^*.$$

由约束规范条件(6.2.3)有$\boldsymbol{d} \in \mathscr{F}_L^*$. 故如下方程组无解：

$$\begin{aligned}
&\nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} = 0, i \in \mathcal{E}; \\
&\nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} \geq 0, i \in \mathcal{I}(\boldsymbol{x}^*) \\
&\boldsymbol{g}^{*T} \boldsymbol{d} < 0
\end{aligned}$$

利用Farkas引理的推论6.2.3可得，存在$\lambda_i^* \in \mathbb{R}(i \in \mathcal{E})$和$\lambda_i^* \geq 0(i \in \mathcal{I}(\boldsymbol{x}^*))$使得:

$$\boldsymbol{g}^* = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) + \sum_{i \in \mathcal{I}(\boldsymbol{x}^*)} \lambda_i^* \nabla c_i(\boldsymbol{x}^*).$$

定理6.2.4证明

证明续：再令$\lambda_i^* = 0$，$i \in \mathcal{I} \backslash \mathcal{I}(\boldsymbol{x}^*)$，便得

$$\begin{aligned} \boldsymbol{g}^* &= \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) + \sum_{i \in \mathcal{I}(\boldsymbol{x}^*)} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) + \sum_{i \in \mathcal{I} \backslash \mathcal{I}(\boldsymbol{x}^*)} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) \\ &= \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*). \end{aligned}$$

显然有:$\lambda_i^* \geq 0, \quad \forall i \in \mathcal{I}$; $\lambda_i^* c_i(\boldsymbol{x}^*) = 0, \quad \forall i \in \mathcal{I}$. ■

注意：结论(6.2.4a)也可以由Lagrange函数来表示：

$$\nabla L_{\boldsymbol{x}}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$$

其中: $L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(\boldsymbol{x})$, 称$\lambda_i$为Lagrange乘子。

KKT条件

定理6.2.4是由Kuhn和Tucker(1951)给出的，由于Karush (1939)也类似地考虑了约束优化的最优性条件，故它常称为Karush-Kuhn-Tucker定理，简称KKT定理。(6.2.4a)-(6.2.4e)称为KKT条件。

在KKT条件中：

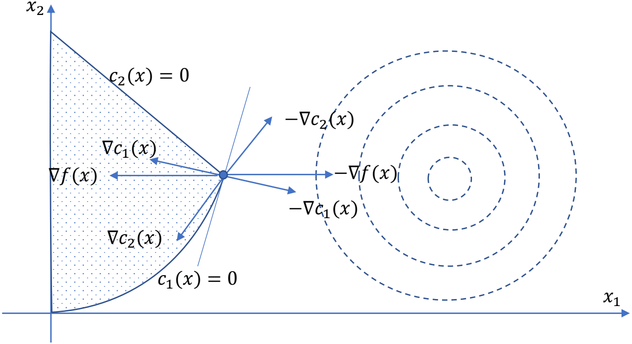
- (6.2.4a)称为驻点条件, 即$\nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$. $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$也称为KKT对;
- (6.2.4b)-(6.2.4c)称为可行性条件;
- (6.2.4d)称为乘子非负条件;
- (6.2.4e)称为互补松弛条件。

若正则性假设成立，则局部极小点必是KKT点；若正则性假设不成立，则局部极小点不一定是KKT点。

驻点条件示例

考虑以下约束优化问题：

$$\begin{cases} \min & f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{s.t.} & c_1(x) = x_2 - x_1^2 \geq 0 \\ & c_2(x) = 2 - x_1 - x_2 \geq 0 \end{cases}$$



判断KKT点

例题 6.2.1 考虑以下问题：

$$\begin{cases} \min & f(x) = -3x_1^2 - x_2^2 - 2x_3^2 \\ \text{s.t.} & -x_1 + x_2 \geq 0 \\ & x_1^2 + x_2^2 + x_3^2 = 3 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$$

试验证：$\boldsymbol{x}^* = (1, 1, 1)$是KKT点.

解： 令

$$\begin{aligned} c_1(x) &= -x_1 + x_2 \geq 0, \\ c_2(x) &= x_1 \geq 0, \\ c_3(x) &= x_2 \geq 0, \\ c_4(x) &= x_3 \geq 0, \\ c_5(x) &= x_1^2 + x_2^2 + x_3^2 - 3 = 0 \end{aligned}$$

KKT点验证示例

在点$\boldsymbol{x}^*$处, 不等式约束取等号的约束集:

$$\mathcal{I}(\boldsymbol{x}^*) = \{1\}.$$

计算:

$$\nabla f(\boldsymbol{x}^*) = (-6, -2, -4)^T$$
$$\nabla c_1(\boldsymbol{x}^*) = (-1, 1, 0)^T, \quad \nabla c_5(\boldsymbol{x}^*) = (2, 2, 2)^T$$

由KKT定理:

$$\begin{bmatrix} -6 \\ -2 \\ -4 \end{bmatrix} - \lambda \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} - \mu \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

显然有$\lambda = 2, \mu = -2$满足以上方程组。这里$\lambda > 0$, $\lambda c_1(\boldsymbol{x}^*) = 0, c_i(\boldsymbol{x}^*) \geq 0 (i = 1, 2, 3, 4), c_5(\boldsymbol{x}^*) = 0$, 故$\boldsymbol{x}^*$为KKT点。

定理6.2.5 一阶充分条件.

设$\boldsymbol{x}^* \in \mathcal{D}$，如果$f(\boldsymbol{x})$和$c_i(\boldsymbol{x})$ $(i = 1, \cdots, m)$都在$\boldsymbol{x}^*$处可微，并且有

$$\boldsymbol{d}^T \boldsymbol{g}^* > 0, \ \forall 0 \neq \boldsymbol{d} \in \mathscr{F}_S^* \tag{6.2.5}$$

成立，则$\boldsymbol{x}^*$是问题(6.1.1)的一个严格局部极小点。

证明： **反证法** 假定$\boldsymbol{x}^*$不是严格局部极小点，则存在$\boldsymbol{x}_k \in \mathcal{D}$使得

$$f(\boldsymbol{x}_k) \leq f(\boldsymbol{x}^*), \tag{a}$$

且有$\boldsymbol{x}_k \to \boldsymbol{x}^*$, $\boldsymbol{x}_k \neq \boldsymbol{x}^*$ $(k = 1, 2, \cdots)$. 不失一般性，假定

$$\boldsymbol{d}_k = \frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2} \to \boldsymbol{d}. \tag{b}$$

定理6.2.5证明

令$\boldsymbol{d}_k = (\boldsymbol{x}_k - \boldsymbol{x}^*)/\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2$, $\delta_k = \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2$. 根据序列可行方向定义即知

$$\boldsymbol{d} \in \mathcal{F}_{\mathcal{S}}{}^*. \tag{c}$$

由(a)-(b)得：

$$f(\boldsymbol{x}_k) = f(\boldsymbol{x}^*) + \delta_k \boldsymbol{d}_k^T \nabla f(\boldsymbol{x}^*) + o(\delta_k^2).$$

两边同除以δ_k, 取极限，得:

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

从而(c)-(d)与(6.2.5)相矛盾。 ■

二阶条件讨论

在可行点$\boldsymbol{x}^* \in \mathcal{D}$处：

- **情形1**: 对于所有$\boldsymbol{d} \in \mathscr{F}_S^*$, 满足${\boldsymbol{g}^*}^T\boldsymbol{d} > 0$, 由一阶充分条件知，$\boldsymbol{x}^*$是一个严格局部极小点;
- **情形2**: 若存在$\boldsymbol{d} \in \mathscr{F}_S^*$, 满足${\boldsymbol{g}^*}^T\boldsymbol{d} < 0$, 由一阶必要条件，$\boldsymbol{x}^*$不是局部极小点。
- **情形3**: 对于${\boldsymbol{g}^*}^T\boldsymbol{d} = 0$无法判断$\boldsymbol{x}^*$是否为最优解。

在KKT点$\boldsymbol{x}^*$处, 由可行性条件有：

$$
{\boldsymbol{g}^*}^T\boldsymbol{d} = \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*)^T\boldsymbol{d} \geq 0,\ \ \boldsymbol{d} \in \mathscr{F}_L^*
$$

由于$\mathscr{F}_S^* \subseteq \mathscr{F}_L^*$, 只需考虑符合${\boldsymbol{g}^*}^T\boldsymbol{d} = 0, \boldsymbol{d} \in \mathscr{F}_L^*$的$\boldsymbol{d}$。

定义6.2.1 线性化零约束方向集合.

设$\boldsymbol{x}^*$为问题(6.1.1)的KKT点，$\boldsymbol{\lambda}^*$为对应的Lagrange乘子，在$\boldsymbol{x}^*$处定义$\mathscr{F}_L^*$的子集为:

$$Q_L^* = \left\{ \boldsymbol{d} \middle| \boldsymbol{d} \in \mathscr{F}_L^*; \text{且} \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} = 0, \lambda_i^* > 0, \forall i \in \mathcal{I}(\boldsymbol{x}^*) \right\}. \tag{6.2.6}$$

注意对不起作用约束，有：

$$\lambda_i^* = 0, i \in \mathcal{I} \backslash \mathcal{I}^*$$

由KKT条件得：

$$\boldsymbol{g}^{*T} \boldsymbol{d} = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} = 0,\ \boldsymbol{d} \in Q_L^*.$$

类似地，可以定义序列可行方向集的子集：

定义6.2.2 序列零约束方向.

设$\boldsymbol{x}^*$为问题(6.1.1)的KKT点，$\boldsymbol{\lambda}^*$为对应的Lagrange乘子，存在可行点序列$\{\boldsymbol{x}_k\}$, $\boldsymbol{x}_k \to \boldsymbol{x}^*$, $\boldsymbol{x}_k \neq \boldsymbol{x}^*$; $\boldsymbol{x}_k = \boldsymbol{x}^* + \alpha_k \boldsymbol{d}_k$满足:

$$\begin{aligned}
c_i(\boldsymbol{x}_k) &\geq 0, \quad \lambda_i^* = 0, \quad i \in \mathcal{I}^*, \\
c_i(\boldsymbol{x}_k) &= 0, \quad \lambda_i^* > 0, \quad i \in \mathcal{I}^*, \\
c_i(\boldsymbol{x}_k) &= 0, \quad \qquad\qquad i \in \mathcal{E}
\end{aligned} \tag{6.2.7}$$

且$\alpha_k \to 0, \boldsymbol{d}_k \to \boldsymbol{d}$, 称$\boldsymbol{d}$为$\boldsymbol{x}^*$处的序列零约束可行方向。所有这样的方向$\boldsymbol{d}$构成的集合记为：$Q_S^*$。

和$\mathscr{F}_S^* \subseteq \mathscr{F}_L^*$关系一样，显然有：$Q_S^* \subseteq Q_L^*$。

二阶必要条件

定义6.2.3.

二阶约束规范条件定义为：

$$Q_S^* = Q_L^*. \tag{6.2.8}$$

定理6.2.6 二阶必要性条件.

设$\boldsymbol{x}^*$是问题(6.1.1)的局部极小点，在$\boldsymbol{x}^*$处正则性假设成立，从而存在Lagrange乘子$\boldsymbol{\lambda}^*$，使得KKT条件满足。若对该乘子$\boldsymbol{\lambda}^*$有二阶约束规范条件(6.2.8)成立，则：

$$\boldsymbol{d}^T \nabla^2_{\boldsymbol{xx}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d} \geq 0, \quad \forall \boldsymbol{d} \in Q_L^*. \tag{6.2.9}$$

其中：拉格朗日函数关于$\boldsymbol{x}$的二阶导数矩阵：

$$\nabla^2_{\boldsymbol{xx}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \nabla^2 f(\boldsymbol{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla^2 c_i(\boldsymbol{x}^*).$$

定理6.2.6证明

证明：设$\boldsymbol{d} \in Q_L^*$, 则$\boldsymbol{d} \in Q_S^*$，于是存在点列$\{\boldsymbol{x}_k\}$, 满足：

$$\boldsymbol{x}_k = \boldsymbol{x}^* + \alpha_k \boldsymbol{d}_k, \quad \alpha_k \to 0, \quad \boldsymbol{d}_k \to \boldsymbol{d}.$$

由Q_S^*的定义知：

$$L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*) = f(\boldsymbol{x}_k) - {\boldsymbol{\lambda}^*}^T \boldsymbol{c}(\boldsymbol{x}_k) = f(\boldsymbol{x}_k)$$

另一方面，由KKT条件：

$$\begin{aligned} L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*) &= L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \alpha_k \nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)^T \boldsymbol{d}_k + \frac{1}{2}\alpha_k^2 \boldsymbol{d}_k^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d}_k + o(\alpha_k^2) \\ &= f(\boldsymbol{x}^*) + \frac{1}{2}\alpha_k^2 \boldsymbol{d}_k^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d}_k + o(\alpha_k^2) \end{aligned}$$

定理6.2.6证明

证明续：因为$\boldsymbol{x}^*$为局部最优解，当k充分大时，有：

$$f(\boldsymbol{x}_k) \geq f(\boldsymbol{x}^*).$$

由以上两个表达式可得：

$$\boldsymbol{d}_k^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d}_k + o(1) \geq 0$$

当$k \to \infty$时，则有:

$$\boldsymbol{d}^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d} \geq 0, \forall \boldsymbol{d} \in Q_L^*$$

■

定理6.2.7 二阶充分性条件.

设$\boldsymbol{x}^*$是问题(6.1.1)一个KKT点，$\boldsymbol{\lambda}^*$是相应的Lagrange乘子。若

$$\boldsymbol{d}^T\nabla^2_{xx}L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)\boldsymbol{d}>0,\quad \forall\boldsymbol{d}\in Q_L^*,\boldsymbol{d}\neq 0, \tag{6.2.10}$$

则$\boldsymbol{x}^*$是问题(6.1.1)的严格局部极小点。

证明：假定$\boldsymbol{x}^*$不是严格局部最优解，则存在可行点列$\{\boldsymbol{x}_k\},\boldsymbol{x}_k\to\boldsymbol{x}^*$，使得：$f(\boldsymbol{x}_k)\leq f(\boldsymbol{x}^*)$，记：$\boldsymbol{x}_k=\boldsymbol{x}^*+\alpha_k\boldsymbol{d}_k$. 当$k\to\infty$时，$\alpha_k\to 0$。不失一般性，假定：

$$\frac{\boldsymbol{x}_k-\boldsymbol{x}^*}{\|\boldsymbol{x}_k-\boldsymbol{x}^*\|_2}\to\boldsymbol{d}$$

定理6.2.7证明

证明续：因为$\boldsymbol{d} \in \mathscr{F}_S^*$，有：

$$\boldsymbol{d}^T \boldsymbol{g}^* \leq 0 \tag{a}$$

又由于$\mathscr{F}_S^* \subseteq \mathscr{F}_L^*$，知：

$$\boldsymbol{d}^T \boldsymbol{g}^* = \sum_{i=1}^{m} \lambda_i^* \boldsymbol{d}^T \nabla c_i(\boldsymbol{x}*) \geq 0 \tag{b}$$

由(a)(b)得：

$$\boldsymbol{d}^T \boldsymbol{g}^* = 0; \lambda_i^* \boldsymbol{d}^T \nabla c_i(\boldsymbol{x}^*) = 0, \forall i \in \mathcal{I}(\boldsymbol{x}^*) \tag{c}$$

由$\boldsymbol{d} \in \mathscr{F}_S^*$与(c)知：$\boldsymbol{d} \in Q_L^*$。

证明续：由$f(\boldsymbol{x}^*) \geq f(\boldsymbol{x}_k)$知：

$$L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \geq L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*) = L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \frac{1}{2}\delta_k^2 \boldsymbol{d}_k^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d}_k + o(\delta_k^2)$$

其中：$\delta_k = \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2$.

当$k \to \infty$, 于是有：

$$\boldsymbol{d}^T \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \boldsymbol{d} \leq 0,$$

与定理条件矛盾，故$\boldsymbol{x}^*$为严格局部极小点。 ■

判断约束优化极小点

例6.2.2 求以下问题：

$$\begin{cases} \min & f(\boldsymbol{x}) = 4x_1 - 3x_2 \\ \text{s.t.} & 4 - x_1 - x_2 \geq 0 \\ & x_2 + 7 \geq 0 \\ & -(x_1-3)^2 + x_2 + 1 \geq 0 \end{cases}$$

的KKT点，并判断其是否为最优解。

解： 该问题的拉格朗日函数为：

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}) &= f(\boldsymbol{x}) - \boldsymbol{\lambda}^T c(\boldsymbol{x}) \\ &= 4x_1 - 3x_2 - \lambda_1(4 - x_1 - x_2) \\ &\quad - \lambda_2(x_2 + 7) - \lambda_3(-(x_1-3)^2 + x_2 + 1) \end{aligned}$$

判断约束优化极小点

解续： KKT条件为：

$$
\begin{cases}
\frac{\partial L}{\partial x_1} = 4 + \lambda_1 + 2\lambda_3(x_1 - 3) = 0 \\
\frac{\partial L}{\partial x_2} = -3 + \lambda_1 - \lambda_2 - \lambda_3 = 0 \\
4 - x_1 - x_2 \geq 0 \\
x_2 + 7 \geq 0 \\
-(x_1 - 3)^2 + x_2 + 1 \geq 0 \\
\lambda_1(4 - x_1 - x_2) = \lambda_2(x_2 + 7) = \lambda_3(-(x_1 - 3)^2 + x_2 + 1) = 0 \\
\lambda_i \geq 0, i = 1, 2, 3
\end{cases}
$$

解方程组得KKT点为：

$$\boldsymbol{x}^* = (1, 3)^T, \boldsymbol{\lambda}^* = (\frac{16}{3}, 0, \frac{7}{3})$$

注意： $Q_L^* = \emptyset$, 则由二阶判断条件知该点为严格局部极小点。