约束优化问题

翟卫欣 副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院

目录

目录

约束优化的最优性理论

- 一般约束最优化问题
- 约束规范条件
- 约束最优化问题的一阶最优性条件
- 约束最优化问题的二阶最优性条件

一般约束最优化问题

- 与无约束最优化问题相比，约束最优化问题的最优性理论要更为复杂。
- 因为在最优解处，除了要考虑目标函数之外，还需考虑约束函数，最优解是两者决定的。

一般约束最优化问题

原问题(P)

$$
\begin{aligned}
&\min f(\boldsymbol{x}) \\
s.t.\ &c_i(\boldsymbol{x}) = 0, i \in E = \{1, 2, \cdots, m_e\} \\
&c_i(\boldsymbol{x}) \geqslant 0, i \in I = \{m_e + 1, m_e + 2, \cdots, m\}
\end{aligned}
$$

- 满足约束条件的点成为可行点，所有可行点的集合为可行集
  $\boldsymbol{D} = \{c_i(\boldsymbol{x}) = 0, i \in E; c_i(\boldsymbol{x}) \geqslant 0, i \in I\}$
  $\boldsymbol{E}$是等式约束集合，$\boldsymbol{I}$是不等式约束集合。

一般约束最优化问题举例

$$min\, f(\boldsymbol{x}) = \frac{x_1^2}{2} + x_2^2$$
$$s.t.\; c_1(\boldsymbol{x}) = 1 - (x_1 - 1)^2 - x_2^2 \geqslant 0$$
$$c_2(\boldsymbol{x}) = -1 + x_1 - x_2^2 \geqslant 0$$

图中深灰色区域$\boldsymbol{D}$是可行域，实线代表$c_1(\boldsymbol{x})$、$c_2(\boldsymbol{x})$，$\boldsymbol{x}^*$是极小值点。

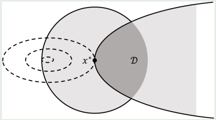


图 1: 可行域与最优解示例

局部最优解与全局最优解

- 对于$\boldsymbol{x}^* \in \boldsymbol{D}$，若存在$\varepsilon > 0$，当$\boldsymbol{x} \in \boldsymbol{D}$且$||\boldsymbol{x} - \boldsymbol{x}^*|| < \varepsilon$时，有$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*)$，则称$\boldsymbol{x}^*$为局部最优解。
- 对于$\boldsymbol{x}^* \in \boldsymbol{D}$，若存在$\varepsilon > 0$，当$\boldsymbol{x} \in \boldsymbol{D}$且$||\boldsymbol{x} - \boldsymbol{x}^*|| < \varepsilon$时，有$f(\boldsymbol{x}) > f(\boldsymbol{x}^*)$，则称$\boldsymbol{x}^*$为严格局部最优解。
- 对于$\boldsymbol{x}^* \in \boldsymbol{D}$，若$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*), \forall \boldsymbol{x} \in \boldsymbol{D}$，则称$\boldsymbol{x}^*$为全局最优解。
- 对于$\boldsymbol{x}^* \in \boldsymbol{D}$，若$f(\boldsymbol{x}) > f(\boldsymbol{x}^*), \forall \boldsymbol{x} \in \boldsymbol{D}$，则称$\boldsymbol{x}^*$为严格全局最优解。

局部最优解与全局最优解举例

$$min\ f(\boldsymbol{x}) = x_1^2 + x_2^2$$
$$s.t.\ c_1(\boldsymbol{x}) = 1 - (x_2 - 1)^2/4 + x_1^2 = 0$$

图中虚线为目标函数的等高线，实线为约束函数，空心点为局部最优解$(0,3)^T$，实心点为全局最优解$(0,-1)^T$。

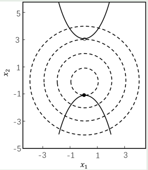


图 2: 局部最优解与全局最优解举例

起作用约束

- 在点$\boldsymbol{x} \in \boldsymbol{D}$处，若$c_i(\boldsymbol{x}) = 0$，则称该约束为在点$\boldsymbol{x}$处的起作用约束；若$c_i(\boldsymbol{x}) > 0$，则称该约束为在点$\boldsymbol{x}$处的不起作用约束。
- 等式约束都是起作用约束，不等式约束中是否是起作用约束要依点$\boldsymbol{x}$而定，我们用$\boldsymbol{I}(\boldsymbol{x})$表示$\boldsymbol{x}$处起作用的不等式约束集合，即$\boldsymbol{I}(\boldsymbol{x}) = \{i|c_i(\boldsymbol{x}) = 0, i \in \boldsymbol{I}\}$。
- 在点$\boldsymbol{x}$处的起作用约束下标集合（简称起作用约束集合）记为$\boldsymbol{A}(\boldsymbol{x}) = \{i|c_i(\boldsymbol{x}) = 0, i \in \boldsymbol{E} \cup \boldsymbol{I}\}$。或者写$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{E} \cup \boldsymbol{I}(\boldsymbol{x})$。特别地，记$\boldsymbol{A}^* = \boldsymbol{A}(\boldsymbol{x}^*), \boldsymbol{I}^* = \boldsymbol{I}(\boldsymbol{x}^*)$。
- 起作用约束是一个很重要的问题，因为在局部最优解$\boldsymbol{x}^*$处，若$\boldsymbol{A}^*$已知，则不起作用约束都可忽略，原问题可优化为：

$$min\, f(\boldsymbol{x})$$

$$s.t.\ c_i(\boldsymbol{x}) = 0, i \in \boldsymbol{A}^*$$

只有等式约束的优化问题举例

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

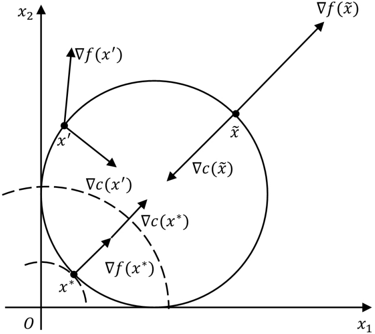


图 3: 在不同可行点处的$\nabla f(\boldsymbol{x}^*)$和$\nabla c(\boldsymbol{x}^*)$

只有等式约束的优化问题举例

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

- 在最优点$\boldsymbol{x}^* = (1 - \frac{\sqrt{2}}{2}, 1 - \frac{\sqrt{2}}{2})^T$处，有$\nabla f(\boldsymbol{x}) = (2 - \sqrt{2}, 2 - \sqrt{2})^T$，$\nabla c(\boldsymbol{x}^*) = (\sqrt{2}, \sqrt{2})^T$，其中$\nabla f(\boldsymbol{x}^*)$与$\nabla c(\boldsymbol{x}^*)$共线，有$\nabla f(\boldsymbol{x}^*) = \lambda \nabla c(\boldsymbol{x}^*), \lambda = \sqrt{2} - 1$。
- 在极大点$\tilde{\boldsymbol{x}} = (1 + \frac{\sqrt{2}}{2}, 1 + \frac{\sqrt{2}}{2})^T$处，$\nabla f(\tilde{\boldsymbol{x}})$与$\nabla c(\tilde{\boldsymbol{x}})$也共线。而在其他点处$\Delta f(\boldsymbol{x})$与$\nabla c(\boldsymbol{x})$不共线。
- 这说明$\nabla f(\boldsymbol{x})$与$\nabla c(\boldsymbol{x})$共线只是最优解的一个必要条件，是否为最优解还需进一步判断。

含不等式约束的优化问题举例

$$min\, f(\boldsymbol{x}) = x_1^2 + x_2^2,$$
$$s.t. 1 - (x_1 - 1)^2 - (x_2 - 1)^2 \geqslant 0$$

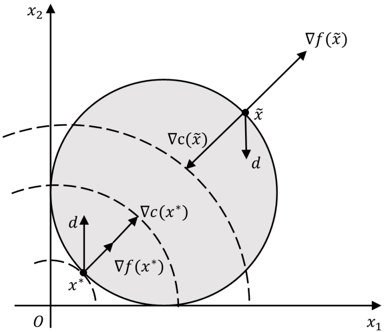


图 4: 在不同可行点处的$\nabla f(\boldsymbol{x}^*)$和$\nabla c(\boldsymbol{x}^*)$

含不等式约束的优化问题举例

$$
\begin{aligned}
&min\, f(\boldsymbol{x}) = x_1^2 + x_2^2,\\
&s.t. 1 - (x_1 - 1)^2 - (x_2 - 1)^2 \geqslant 0
\end{aligned}
$$

- 要求$\nabla f(\boldsymbol{x}) = \lambda \nabla c(\boldsymbol{x})$，其中$\lambda \geqslant 0$，也就是对$\lambda$的正负值有了要求，即约束条件的梯度方向需要指向可行域内部，该方向同样是目标函数的是梯度方向。
- 本例中只有一个约束，而对于一般的约束优化问题，一阶最优性条件为$\nabla f(\boldsymbol{x}^*) = \sum\limits_{i \in E \cup I} \lambda_i^* \nabla c(\boldsymbol{x}^*)$，其中当$i \in I^*$时，$\lambda_i^* \geqslant 0$，也就是说要求$\boldsymbol{x}^*$处的起作用的不等式约束对应的$\lambda_i$非负。
- 该条件称为一阶最优性条件。

一阶最优性条件不满足举例

- 在最优解x^*处，并非对任意约束函数，一阶最优性条件都会满足。

$$
\begin{aligned}
min\, f(x) &= x_2, \\
s.t. c_1 &= -x_1 - (x_2)^2 \geqslant 0, \\
c_2 &= x_1 = 0
\end{aligned}
$$

- 该问题的最优解为$x^* = (0,0)^T$，另外 $\nabla f(\boldsymbol{x}^*) = (0,1)^T, \nabla c_1(\boldsymbol{x}^*) = (-1,0)^T, \nabla c_2(\boldsymbol{x}^*) = (1,0)^T$ 显然在该处，无法找到一组$\{\lambda_i\}$使得$\nabla f(\boldsymbol{x}^*) = \sum_{i \in E \cup I} \lambda_i^* \nabla c(\boldsymbol{x}^*)$（其中当$i \in I^*$时，$\lambda_i^* \geqslant 0$）。

一阶最优性条件不满足举例

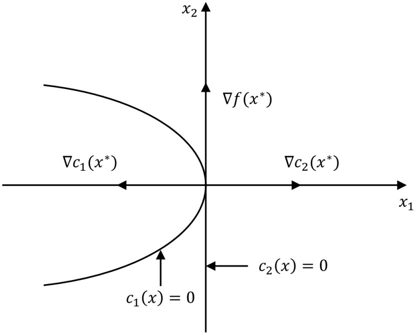


图 5: 在最优解处，一阶最优性条件不满足举例

本例说明，在最优解处，若要最优性条件满足，约束函数需满足某些条件。我们称这些条件为约束规范条件。

目录

中国农业大学 China Agricultural University

约束规范条件-可行方向

- 设x是约束最优化问题的可行点，存在可行点序列$\boldsymbol{x}_k$，有$\boldsymbol{x}_k \to \boldsymbol{x}, \boldsymbol{x}_k \neq \boldsymbol{x}$，记$\boldsymbol{x}_k = \boldsymbol{x} + \alpha_k \boldsymbol{d}_k$
- 其中$||\boldsymbol{d}_k|| = 1, \alpha_k > 0$。因为$\boldsymbol{x}_k \to \boldsymbol{x}$，所以$\alpha_k \to 0$。
- 如果$\boldsymbol{d} \to \boldsymbol{d}$，则称$\boldsymbol{d}_k$为可行方向点列，$\boldsymbol{d}$是$x$处的可行方向（或称序列可行方向）。记$\mathcal{F} = \mathcal{F}(\boldsymbol{x})$为$x$处全体可行方向组成的集合。

约束规范条件-可行方向举例

- 对于$\boldsymbol{x} = (1 - cos\theta, 1 - sin\theta)^T$，假设$\boldsymbol{x}_k = (1 - cos\theta_k, 1 - sin\theta_k)^T$是可行方向点列，其中$\theta_k = \frac{\pi}{4} + \frac{1}{k}$，$\theta = \frac{\pi}{4}$，求可行方向$\boldsymbol{d}$

- $||\boldsymbol{x}_k - \boldsymbol{x}||^2 = (cos\theta - cos\theta_k)^2 + (sin\theta - sin\theta_k)^2 = 4sin^2\frac{cos\theta - cos\theta_k}{2}$
  $\boldsymbol{d}_k = \frac{\boldsymbol{x}_k - \boldsymbol{x}}{||\boldsymbol{x}_k - \boldsymbol{x}||} = (sin\frac{\theta + \theta_k}{2}, -cos\frac{\theta + \theta_k}{2})^T$

- 当$k \to \infty$时，有$\boldsymbol{d}_k \to \boldsymbol{d} = (sin\theta, -cos\theta)^T$，知$d \in \mathcal{F}(\boldsymbol{x})$，此处不再一一计算$x$处其余可行方向

约束规范条件-线性化可行方向

- 设$\boldsymbol{x}$是约束最优化问题的可行点，定义$F=F(\boldsymbol{x})$ $=\{\boldsymbol{d} \mid \|\boldsymbol{d}\|=1, \nabla c_i(\boldsymbol{x})^T\boldsymbol{d}=0, i\in E; \nabla c_i(\boldsymbol{x})^T\boldsymbol{d}\geqslant 0, i\in I(\boldsymbol{x})\}$ 为$\boldsymbol{x}$处的线性化可行方向集合，$\boldsymbol{d}$是$\boldsymbol{x}$处的线性化可行方向。
  例如，考虑$\boldsymbol{x}\in \boldsymbol{D}=\{\boldsymbol{x} \mid c_1(\boldsymbol{x})\geqslant 0, c_2(\boldsymbol{x})=0\}$处的线性化可行方向集合：$F(\boldsymbol{x})=\{\boldsymbol{d} \mid \nabla c_1(\boldsymbol{x})^T\boldsymbol{d}\geqslant 0, \nabla c_2(\boldsymbol{x})^T\boldsymbol{d}=0\}$。灰色区域表示满足$\nabla c_1(\boldsymbol{x})^T\boldsymbol{d}\geqslant 0$的$\boldsymbol{d}$集合，粗黑线表示$F$。

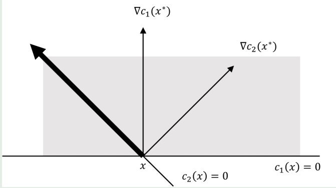


图 6: $\boldsymbol{x}$处的线性化可行方向

约束规范条件-线性化可行方向

线性化可行方向由泰勒公式得,对等式约束

$c_i(\boldsymbol{x}+\boldsymbol{d}) \approx c_i(\boldsymbol{x}) + \nabla c_i(\boldsymbol{x})^T\boldsymbol{d}, i \in \boldsymbol{E}$,

由于$c_i(\boldsymbol{x}+\boldsymbol{d}) = c_i(\boldsymbol{x}) = 0$，所以$\nabla c_i(\boldsymbol{x})^T\boldsymbol{d} = 0, i \in \boldsymbol{E}$。

对不等式约束

$c_i(\boldsymbol{x}+\boldsymbol{d}) \approx c_i(\boldsymbol{x}) + \nabla c_i(\boldsymbol{x})^T\boldsymbol{d}, i \in \boldsymbol{I}(\boldsymbol{x})$,

由于$c_i(\boldsymbol{x}+\boldsymbol{d}) \geqslant 0, c_i(\boldsymbol{x}) = 0$，所以$\nabla c_i(\boldsymbol{x})^T\boldsymbol{d} \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})$。

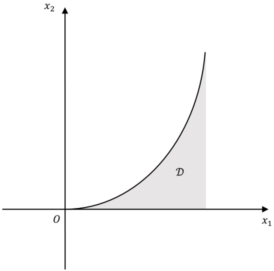
两类约束规范条件的关系

- $\boldsymbol{x} \in \boldsymbol{D}$处的可行方向集合$\mathcal{F}(\boldsymbol{x}) \subset F(\boldsymbol{x})$。
- 证明：
  设$\boldsymbol{d} \in \mathcal{F}(\boldsymbol{x})$，则由可行点列$\{\boldsymbol{x}_k\}$，满足$\boldsymbol{x}_k = \boldsymbol{x} + \alpha_k \boldsymbol{d}_k \to \boldsymbol{x}$，且有$\alpha_k \to 0, \boldsymbol{d}_k \to \boldsymbol{d}$。下面我们证明$\boldsymbol{d} \in F(\boldsymbol{x})$。
- 由泰勒公式，有：
  $c_i(\boldsymbol{x}_k) = c_i(\boldsymbol{x}) + \alpha_k \nabla c_i(\boldsymbol{x})^T \boldsymbol{d}_k + o(\alpha_k)$
  因为$c_i(\boldsymbol{x}_k) = c_i(\boldsymbol{x}) = 0, \boldsymbol{x} \in \boldsymbol{E}$，且有$c_i(\boldsymbol{x}_k) \geqslant c_i(\boldsymbol{x}) = 0, \boldsymbol{x} \in \boldsymbol{I}(\boldsymbol{x})$
- 于是可得$\nabla c_i(\boldsymbol{x})^T \boldsymbol{d}_k + o(1) = 0, i \in \boldsymbol{E}$
  且有$\nabla c_i(\boldsymbol{x})^T \boldsymbol{d}_k + o(1) \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})$
- 令$k \to \infty$，则$d_k \to d, o(1) \to 0$，得$\nabla c_i(\boldsymbol{x})^T \boldsymbol{d}_k = 0, i \in \boldsymbol{E}$
  且有 $\nabla c_i(\boldsymbol{x})^T \boldsymbol{d}_k \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})$
- 即$\boldsymbol{d} \in F(\boldsymbol{x})$

两类约束规范条件的关系

- $F(\boldsymbol{x}) \subset \mathcal{F}(\boldsymbol{x})$不一定成立。
- 例如对如下约束：

$$
\begin{aligned}
x_1^3 &\geqslant x_2 \\
x_2 &\geqslant 0
\end{aligned}
$$



两类约束规范条件的关系

- 设$\boldsymbol{x}=(0,0)^T$，$\{\boldsymbol{x}^{(k)}\}$为$x_1^3=x_2$上满足$x_2>0$的点列，$\boldsymbol{x}^{(k)}\to\boldsymbol{x}$，则$\boldsymbol{d}_k=\frac{\boldsymbol{x}^{(k)}-\boldsymbol{x}}{||\boldsymbol{x}^{(k)}-\boldsymbol{x}||}\to(1,0)^T=\boldsymbol{d}$，即$\mathcal{F}=(1,0)^T$
- 在$\boldsymbol{x}=(0,0)^T$处，有$\nabla c_1(\boldsymbol{x})=(0,-1)^T,\nabla c_2(\boldsymbol{x})=(0,1)^T$，则$F=\{\boldsymbol{d}|\nabla c_1(\boldsymbol{x})^T\boldsymbol{d}\geqslant 0,\nabla c_2(\boldsymbol{x})^T\boldsymbol{d}\geqslant 0,||\boldsymbol{d}||=1\}=\{(1,0)^T,(-1,0)^T\}$。
- 因为$\nabla c_1(\boldsymbol{x})=(0,-1)^T,\nabla c_2(\boldsymbol{x})=(0,1)^T$，所以$\boldsymbol{d}=(-1,0)^T\in F$ 又由于任意可行点$\boldsymbol{x}^{(k)}\geqslant 0$，所以不存在可行点列$\{\boldsymbol{x}^{(k)}\},\boldsymbol{x}^{(k)}=\boldsymbol{x}+\alpha_k\boldsymbol{d}_k$使得$\boldsymbol{d}_k\to\boldsymbol{d}$，所以$\boldsymbol{d}=(-1,0)^T\notin\mathcal{F}$。

约束规范条件

- 引入可行方向和线性化可行方向的目的是建立约束规范条件，从而建立最优性条件。
- 约束规范条件有多种，常用的一种为KT(Kuhn-Tucker)约束规范条件：
  $F = \mathcal{F}$
- 下列两条件之一成立时，KT约束规范条件成立：
  $\boldsymbol{A}$中的所有约束为线性约束
  $\{\nabla c_i(\boldsymbol{x}), i \in \boldsymbol{A}\}$线性无关

约束规范条件

- 另一种约束规范条件是正则性假设，它是建立在可行方向和下降方向的基础之上的。
- 定义$\mathcal{D} = \mathcal{D}(\boldsymbol{x}) = \{\boldsymbol{d} | \boldsymbol{d}^T \nabla f(\boldsymbol{x}) < 0, \boldsymbol{d} \in \mathbb{R}^n\}$为$f(\boldsymbol{x})$在$\boldsymbol{x}$处的下降方向集合，其中元素$\boldsymbol{d}$称为$\boldsymbol{x}$处的下降方向。
- 正则性假设为：$F \cap \mathcal{D} = \mathcal{F} \cap \mathcal{D}$
- 正则性假设只考虑可行方向集合中下降方向的部分，显然这个条件比KT约束规范条件弱，即若KT约束规范条件成立，则正则性假设成立，反之不一定。

约束规范条件举例

$$min\ x_1^2 + x_2$$
$$s.t. x_1^3 \geqslant x_2,$$
$$x_2 \geqslant 0$$

- 最优解为$\boldsymbol{x}^* = (0,0)^T$。$F = \{(\boldsymbol{d}_1)^T \| \|\boldsymbol{d}_1\| = 1\}$，$\mathcal{F} = \{(\boldsymbol{d}_1)^T \| \|\boldsymbol{d}_1\| = 1\}$。

- 另外，由$\nabla f(\boldsymbol{x}^*)^T = (0,1)$，$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} < 0$可知 $\mathcal{D} = \{\boldsymbol{d} | \boldsymbol{d} = (d_1, d_2)^T, d_2 < 0\}$

- 此时$F \cap \mathcal{D} = \varnothing, F \neq \mathcal{F}$，正则假设成立而KT约束规范条件不成立。

局部最优解的一阶必要条件

- 若$\boldsymbol{x}^*$是问题（P）的局部最优解，则$\mathcal{F}^* \cap \mathcal{D}^* = \varnothing$
- 证明：只需证明任取$\boldsymbol{d} \in \mathcal{F}$，都有$\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geqslant 0$即可。
- 由泰勒公式，有：
  $f(\boldsymbol{x}^{(k)}) = f(\boldsymbol{x}^*) + \alpha_k \nabla f(\boldsymbol{x}^*)^T \boldsymbol{d}_k + o(\alpha_k)$
- 由于$\boldsymbol{x}^*$是局部最优解，所以当k充分大时，有$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d}_k + o(1) \geqslant 0$
- 令$k \to \infty$，得$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} \geqslant 0$，所以$\boldsymbol{d} \notin \mathcal{D}^*$
- 该定理说明在最优解$\boldsymbol{x}^*$处，不存在可行的下降方向。

目录

分离定理

- 设C是m个n维向量$\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$生成的集合：$C = \{\boldsymbol{v} | \boldsymbol{v} = \sum_{i=1}^{m} \lambda_i \boldsymbol{a}_i, \lambda_i \geqslant 0, i = 1, 2, \cdots, m\}$

- 如果n维向量$\boldsymbol{g} \notin C$，则存在一个法向量为$\boldsymbol{d}$的超平面Π分离$\boldsymbol{g}$与C，使得
  $\boldsymbol{g}^T\boldsymbol{d} < 0$且$\boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, i = 1, 2, \cdots, m$

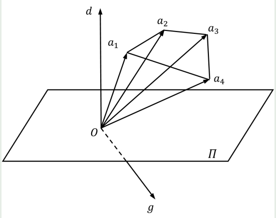


图 8: 分离定理的几何意义

Farkas定理

- 给定m个n维向量$\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$以及另一个$n$维向量$\boldsymbol{g}$，则集合$\mathcal{D}_1 = \{\boldsymbol{d} | \boldsymbol{g}^T\boldsymbol{d} < 0, \boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, i = 1, 2, \cdots, m\}$为空集的充分必要条件是:
  存在$\lambda_i \geqslant 0, (i = 1, 2, \cdots, m)$使得$\boldsymbol{g} = \sum_{i=1}^m \lambda_i \boldsymbol{a}_i$

- 证明：充分性
  设$\lambda_i \geqslant 0 (i = 1, 2, \cdots, m)$，对满足$\boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0 (i = 1, 2, \cdots, m)$的$\boldsymbol{d}$，有$\boldsymbol{g}^T\boldsymbol{d} = \sum_{i=1}^m \lambda_i \boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0$，所以$\mathcal{D}_1$为空集。

- 必要性
  假定$\boldsymbol{g} \notin \boldsymbol{C}$，$\boldsymbol{C} = \{\boldsymbol{v} | \boldsymbol{v} = \sum_{i=1}^m \lambda_i \boldsymbol{a}_i, \lambda_i \geqslant 0, i = 1, 2, \cdots, m\}$。由分离定理知存在以$\boldsymbol{d}$为法向量的超平面，使得$\boldsymbol{g}^T\boldsymbol{d} < 0, \boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, i = 1, 2, \cdots, m$
  所以$\boldsymbol{d} \in \mathcal{D}_1$，即$\mathcal{D}_1$非空，与已知条件矛盾，所以$\boldsymbol{g} \in \boldsymbol{C}$。

Farkas引理

- 设$A \in \mathbb{R}^{m\times n}$和$\omega \in \mathbb{R}^n$
  系统I：$\boldsymbol{Ad} \leqslant \boldsymbol{0}, \boldsymbol{g}^T\boldsymbol{d} > 0, \boldsymbol{d} \in \mathbb{R}^n$
  系统II：$\boldsymbol{g} = \boldsymbol{A}^T\boldsymbol{\lambda}, \boldsymbol{\lambda} \geqslant \boldsymbol{0}$
  则两系统有且仅有一个有解。
- 分两种情况讨论：
  假设系统II有解，则存在$\boldsymbol{\lambda} \geqslant 0$使得$\boldsymbol{g} = \boldsymbol{A}^T\boldsymbol{\lambda}$，若系统I有解，则$0 < \boldsymbol{g} = \boldsymbol{\lambda}^T\boldsymbol{Ad} \leqslant 0$，矛盾，因此系统I无解。
  假设系统II无解，构造集合$\boldsymbol{C} = \{\boldsymbol{v}|\boldsymbol{v} = \boldsymbol{A}^T\boldsymbol{\lambda}, \boldsymbol{\lambda} \geqslant \boldsymbol{0}\}$，$\boldsymbol{C}$是非空闭凸集。系统II无解表明$\boldsymbol{g} \notin \boldsymbol{C}$，由点与凸集的分离定理可知存在$\boldsymbol{d}, \beta$满足 $\boldsymbol{d}^T\boldsymbol{v} \leqslant \beta < \boldsymbol{d}^T\boldsymbol{g}, \forall \boldsymbol{v} \in \boldsymbol{C}$
- 注意到$\boldsymbol{0} \in \boldsymbol{C}$，则可知$0 \leqslant \beta < \boldsymbol{d}^T\boldsymbol{g}$
  同时$\beta \geqslant \boldsymbol{d}^T\boldsymbol{v} = \boldsymbol{d}^T\boldsymbol{A}^T\boldsymbol{\lambda} = (\boldsymbol{\lambda Ad})^T, \forall \boldsymbol{\lambda} \geqslant \boldsymbol{0}$
- 由于$\boldsymbol{\lambda}$可以无限大，所以必须有$\boldsymbol{Ad} \leqslant 0$，所以$\boldsymbol{d}$是系统I的解。

Farkas引理的推论

- $\mathcal{D}_2 = \{\boldsymbol{d} | \boldsymbol{g}^T\boldsymbol{d} < 0; \boldsymbol{a}_i^T\boldsymbol{d} = 0, i \in \boldsymbol{E}; \boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})\}$为空集的充分必要条件是存在$\lambda_i, i \in \boldsymbol{A} = \boldsymbol{E} \cup \boldsymbol{I}(\boldsymbol{x})$，使得$\boldsymbol{g} = \sum_{i \in \boldsymbol{A}} \lambda_i \boldsymbol{a}_i, \lambda_i \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})$,此处$\boldsymbol{g}$即为$\nabla f(\boldsymbol{x})$，$\boldsymbol{a}_i$即为$\nabla c_i(\boldsymbol{x})$。
- 证明：由于$\boldsymbol{a}_i^T\boldsymbol{d} = 0, i \in \boldsymbol{E}$可以写成$\boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, -\boldsymbol{a}_i^T\boldsymbol{d} \geqslant 0, i \in \boldsymbol{E}$，由Farkas引理知，存在$\lambda_i^+ \geqslant 0, i \in \boldsymbol{E}; \lambda_i^- \geqslant 0, i \in \boldsymbol{E}; \lambda_i^+ \geqslant 0, i \in \boldsymbol{I}(\boldsymbol{x})$使得

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

其中，当$i \in \boldsymbol{E}$时，$\lambda_i = \lambda_i^+ - \lambda_i^-$。

目录

约束最优化问题解的一阶必要条件

- 如果$\boldsymbol{x}^*$是问题(P)的局部最优解，且在$\boldsymbol{x}^*$处正则性假设成立，则存在Lagrange乘子$\lambda^* \in \mathbb{R}^m$，使得$\boldsymbol{x}^*$、$\lambda^*$满足KKT条件：

$$a)\nabla_x L(\boldsymbol{x}^*, \lambda^*) = 0 \rightarrow \nabla f(\boldsymbol{x}^*) = \sum_{i \in \boldsymbol{A}(\boldsymbol{x}^*)}^{m} \lambda_i^* \nabla c_i(\boldsymbol{x}^*)$$

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

$$c)c_i(\boldsymbol{x}^*) \geqslant 0, i \in \boldsymbol{I}$$

$$d)\lambda_i^* \geqslant 0, i \in \boldsymbol{I}$$

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

- 其中，$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i c_i(\boldsymbol{x})$为Lagrange函数。

约束最优化问题解的一阶必要条件

- 证明：
  局部最优解的一阶必要条件要求$\mathcal{F}^* \cap \mathcal{D}^* = \varnothing$，正则性假设要求$\mathcal{F} \cap \mathcal{D} = F \cap \mathcal{D}$，即在$\boldsymbol{x}^*$处，无线性化的可行下降方向。由Farkas引理的推论可知，在$\boldsymbol{x}^*$处，有 $\nabla f(\boldsymbol{x}^*) = \sum\limits_{i \in \boldsymbol{A}^*} \lambda_i^* \nabla c_i(\boldsymbol{x}^*), \lambda_i^* \geqslant 0, i \in \boldsymbol{I}^*$
  令$\lambda_i^* = 0, i \in \boldsymbol{I} - \boldsymbol{I}^*$
  得到$\nabla f(\boldsymbol{x}^*) = \sum\limits_{i \in \boldsymbol{E} \cup \boldsymbol{I}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*)$，同时得到
  $\lambda_i^* c_i(\boldsymbol{x}^*) = 0, i \in \boldsymbol{E} \cup \boldsymbol{I}$
  所以(a)、(d)、(e)成立。又由于最优解的原因，所以(b)、(c)成立
- 其中，$\boldsymbol{A}$是起作用约束，$\boldsymbol{I}^*$是$\boldsymbol{x}^*$处起作用的不等式约束，$\boldsymbol{E}$是等式约束，$\boldsymbol{A}^*$是$\boldsymbol{x}^*$处起作用的所有约束，$\boldsymbol{A}^* = \boldsymbol{E} \cup \boldsymbol{I}^*$。

约束最优化问题解的一阶必要条件

- KKT条件全称为Karush-Kuhn-Tucker条件。满足KKT条件的$\boldsymbol{x}^*$点称为KKT点，相应的$\boldsymbol{\lambda}^*$称为Lagrange乘子，$\boldsymbol{x}^*$和$\boldsymbol{\lambda}^*$统称为KKT对。
- (e)成立的原因是
  当$i \in \boldsymbol{A}^* = \boldsymbol{E} \cup \boldsymbol{I}^*$时，$c_i(\boldsymbol{x}) = 0$
  当$i \in \boldsymbol{I} - \boldsymbol{I}^*$时，$\lambda_i^* = 0$
- （e）条件称为互补条件，当$\lambda_i^* > 0, i \in \boldsymbol{I}^*$时称为严格互补条件。
- 互补条件说明λ_i^*和c_i不可能同时非0。

互补条件的几何意义

图中实心点是最优解，虚线是目标函数$f(\boldsymbol{x})$的等高线，灰色区域是可行区域。

（a）是约束强有效的情形；（b）是约束弱有效的情形；（c）是约束无效的情形。

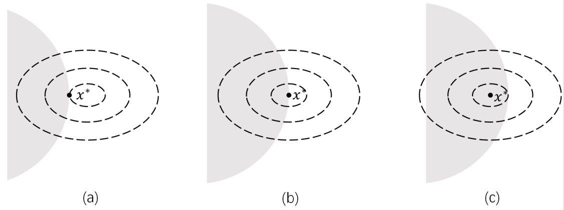


图 9: 互补条件的几何意义

互补条件的几何意义举例

$$
\begin{aligned}
min \quad & \left(x_1 - \frac{9}{4}\right)^2 + (x_2 - 2)^2 \\
s.t. \quad & -x_1^2 + x_2 \geqslant 0 \\
& x_1 + x_2 \leqslant 6
\end{aligned}
$$

请判断如下 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ 三个点是否满足 KKT 条件，写明必要计算过程，$\boldsymbol{x}_1 = \left[\frac{3}{2}, \frac{9}{4}\right]^T$,$\boldsymbol{x}_2 = \left[\frac{9}{4}, 2\right]^T$,$\boldsymbol{x}_3 = [0, 2]^T$ 。

互补条件的几何意义举例

$$L(\boldsymbol{x},\boldsymbol{\lambda}) = (x_1-9/4)^2-(x_2-2)^2-\lambda_1(-x_1^2+x_2)-\lambda_2(-x_1-x_2+6)$$

KKT条件为

$$
\begin{aligned}
&a)\ \frac{\partial L}{\partial x_1} = 2x_1-9/2+2\lambda_1 x_1+\lambda_2 = 0,\\
&b)\ \frac{\partial L}{\partial x_2} = 2x_2-4-\lambda_1+\lambda_2 = 0,\\
&c)\ -x_1^2+x_2 \geqslant 0,\\
&d)\ x_1+x_2-6 \leqslant 0,\\
&e)\ \lambda_1(-x_1^2+x_2) = \lambda_2(x_1+x_2-6) = 0\\
&f)\ \lambda_i \geqslant 0, i=1,2.
\end{aligned}
$$

互补条件的几何意义举例

分别验证三个点是否满足KKT条件：

1. 对于$\boldsymbol{x}_1=\left[\frac{3}{2},\frac{9}{4}\right]^T$，代入e)得$\lambda_2=0$，再代入b)得$\lambda_1=1/2$，再代入其余各式均成立，故$\boldsymbol{x}_1=\left[\frac{3}{2},\frac{9}{4}\right]^T$是KKT点；
2. 对于$\boldsymbol{x}_2=\left[\frac{9}{4},2\right]^T$，因其不满足c)，故$\boldsymbol{x}_2=\left[\frac{9}{4},2\right]^T$不是KKT点；
3. 对于$\boldsymbol{x}_3=[0,2]^T$，代入e)得$\lambda_1=\lambda_2=0$，但该结果不满足a)，故$\boldsymbol{x}_3=[0,2]^T$不是KKT点；

约束最优化问题的一阶充分条件

- 设$\boldsymbol{x}^* \in \boldsymbol{D}$，如果$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} > 0, \boldsymbol{d} \in \mathcal{F}$，则$\boldsymbol{x}^*$是问题(P)的严格局部最优解。
- 证明：如果$\boldsymbol{x}^*$不是严格局部最优解，则存在一个序列$\{\boldsymbol{x}_k\}$，其中$\boldsymbol{x}_k \in \boldsymbol{D}$，使得$f(\boldsymbol{x}_k) \leqslant f(\boldsymbol{x}^*)$，其中$\boldsymbol{x}_k \to \boldsymbol{x}^*, \boldsymbol{x}_k \neq \boldsymbol{x}^*$
- 定义$\boldsymbol{d}_k = \frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{||\boldsymbol{x}_k - \boldsymbol{x}^*||}$，$\boldsymbol{d}_k \to \boldsymbol{d}, \boldsymbol{d} \in \mathcal{F}^*$
  $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) = \nabla f(\boldsymbol{x}^*)(\boldsymbol{x}_k - \boldsymbol{x}^*) + o(||\boldsymbol{x}_k - \boldsymbol{x}^*||) \leqslant 0$
  从而$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d}_k + o(1) \leqslant 0$
- 令$k \to \infty$，得$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} \leqslant 0$，与原假设矛盾。
- 所以，设$\boldsymbol{x}^* \in \boldsymbol{D}$，如果$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} > 0, \boldsymbol{d} \in F$，则$\boldsymbol{x}^*$是问题(P)的严格局部最优解。

目录

约束最优化问题的二阶条件

- 在$\boldsymbol{x}^* \in \boldsymbol{D}$处，由一阶充分条件知，对$\forall \boldsymbol{d} \in \mathcal{F}^*$，若$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d} > 0$成立，则$\boldsymbol{x}^*$是一个严格局部最优解；由一阶必要条件知，若$\exists \boldsymbol{d} \in \mathcal{F}^*$满足$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d} < 0$，即在$\boldsymbol{x}^*$处存在可行下降方向，则$\boldsymbol{x}^*$不是局部最优解。

- 这就是说，我们根据$\boldsymbol{x}^*$处可行方向均是上升方向还是有下降方向，我们可以判断$\boldsymbol{x}^*$是否是最优解。然而对于满足$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d} = 0$的可行方向，我们无法判断$\boldsymbol{x}^*$是否是最优解。

- 注意到在KKT点$\boldsymbol{x}^*$处有 $\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d} = \sum\limits_{i\in \boldsymbol{E}\cup \boldsymbol{I}} \lambda_i^* c_i(\boldsymbol{x}^*)^T\boldsymbol{d} \geqslant 0, \boldsymbol{d} \in F^*$，又因已知$\mathcal{F}^* \subset F^*$，我们知道只需在KKT处，对满足$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d} = 0$的$\boldsymbol{d}$，利用Lagrange函数的二阶信息进一步讨论$\boldsymbol{x}^*$的最优性问题。

约束最优化问题的二阶条件

- 我们要确定集合$\{\boldsymbol{d}|\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d}=0,\boldsymbol{d}\in F^*\}$，它是$F^*$的子集。
- 定义线性化可行方向子集F_1^*为

$$F_1^* = \{\boldsymbol{d}|\boldsymbol{d}\neq\boldsymbol{0},\nabla c_i(\boldsymbol{x}^*)^T\boldsymbol{d}=0,\lambda_i^*=0,i\in\boldsymbol{I}^*;$$
$$\nabla c_i(\boldsymbol{x}^*)^T\boldsymbol{d}=0,\lambda_i^*>0,i\in\boldsymbol{I}^*;\nabla c_i(\boldsymbol{x}^*)^T\boldsymbol{d}=0,i\in\boldsymbol{E}\}$$

- 注意到对于不起作用的约束，有$\lambda_i^*=0,i\in\boldsymbol{I}-\boldsymbol{I}^*$，由KKT条件得$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d}=\sum\limits_{i\in\boldsymbol{E}\cup\boldsymbol{I}}\lambda_i^*\nabla c_i(\boldsymbol{x}^*)^T\boldsymbol{d}=0,\boldsymbol{d}\in F_1^*$

约束最优化问题的二阶条件

- 类似的，定义可行方向子集$\mathcal{F}_1^*$。
- 设$\boldsymbol{x}^*$为最优化问题的可行点，存在可行点序列$\{\boldsymbol{x}_k\}$，$\boldsymbol{x}_k \to \boldsymbol{x}^*, \boldsymbol{x}_k \neq \boldsymbol{x}^*; \boldsymbol{x}_k = \boldsymbol{x}^* + \alpha_k \boldsymbol{d}_k$为满足

$$
\begin{aligned}
c_i(\boldsymbol{x}_k) \geqslant 0, \lambda_i^* = 0, i \in \boldsymbol{I}^*, \\
c_i(\boldsymbol{x}_k) = 0, \lambda_i^* > 0, i \in \boldsymbol{I}^*, \\
c_i(\boldsymbol{x}_k) = 0, i \in \boldsymbol{E}
\end{aligned}
$$

的可行点列，且$||\boldsymbol{d}_k|| = 1, \alpha_k \to 0, \boldsymbol{d}_k \to \boldsymbol{d}$，称$\boldsymbol{d}$为$\boldsymbol{x}^*$处的可行方向，由这样的$\boldsymbol{d}$构成的集合$\mathcal{F}_1^*$为$\boldsymbol{x}^*$处的可行方向集合。

- 同理可证明$\mathcal{F}_1^* \subset F_1^*$。进一步，我们可以定义下面的二阶约束规范条件。
- $\mathcal{F}_1^* = F_1^*$

约束最优化问题的二阶必要条件

- 设$\boldsymbol{x}^*$是P问题的局部最优解，在$\boldsymbol{x}^*$处正则性假设成立，从而存在$\boldsymbol{\lambda}^*$，使得KKT条件满足。若对该乘子$\boldsymbol{\lambda}^*$，满足二阶约束规范条件$\mathcal{F}_1^* = F_1^*$，则有：
  $\boldsymbol{d}^T \boldsymbol{W}^* \boldsymbol{d} \geqslant 0, \boldsymbol{d} \in F_1^*$
  其中$\boldsymbol{W}^* = \nabla_{\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \nabla^2 f(\boldsymbol{x}^*) - \sum_{i=1}^{m} \lambda_i^* \nabla^2 c_i(\boldsymbol{x}^*)$
- 证明：设$\boldsymbol{d} \in F_1^*$，则$\boldsymbol{d} \in \mathcal{F}_1^*$，存在可行点列$\{\boldsymbol{x}_k\}$，其中$\boldsymbol{x}_k = \boldsymbol{x}^* + \alpha_k \boldsymbol{d}_k, \alpha_k \to 0, \boldsymbol{d}_k \to \boldsymbol{d}$。
  一方面，由$\mathcal{F}_1^*$的定义知，$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i^* c_i(\boldsymbol{x}) = f(\boldsymbol{x})$
  另一方面，

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

约束最优化问题的二阶必要条件

- 因为$\boldsymbol{x}^*$是P问题的局部最优解，当k充分大时，有$f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x}_k)$，由前述$L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*)$的两个表达式得：$\boldsymbol{d}_k^T \boldsymbol{W}^* \boldsymbol{d}_k + o(1) \geqslant 0$
- 令$k \to \infty$，得$\boldsymbol{d}_k^T \boldsymbol{W}^* \boldsymbol{d}_k \geqslant 0, \boldsymbol{d} \in F_1^*$。

约束最优化问题的二阶充分条件

- 设$\boldsymbol{x}^*$，$\boldsymbol{\lambda}^*$是问题(P)的KKT对，若有$\boldsymbol{d}^T\boldsymbol{W}^*\boldsymbol{d} > 0, \boldsymbol{d} \in F_1^*$，则$\boldsymbol{x}^*$是问题的严格局部最优解。

- 证明：假设$\boldsymbol{x}^*$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$f(\boldsymbol{x}_k) \leqslant f(\boldsymbol{x}^*)$。
  记$\boldsymbol{x}_k = \boldsymbol{x}^* + \alpha_k\boldsymbol{d}_k$，其中$||\boldsymbol{d}_k|| = 1$，当$k \to \infty$时，$\alpha_k \to 0, \boldsymbol{d}_k \to \boldsymbol{d}, \boldsymbol{d} \in \mathcal{F}^*$
  由泰勒展开得$f(\boldsymbol{x}_k) = f(\boldsymbol{x}^*) + \alpha_k\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d}_k + o(\alpha_k)$
  由于$f(\boldsymbol{x}_k) \leqslant f(\boldsymbol{x}^*)$，则$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d}_k + o(1) \leqslant 0$，

- 令$k \to \infty$，得$\nabla f(\boldsymbol{x}^*)^T\boldsymbol{d}_k \leqslant 0$

约束最优化问题的二阶充分条件

- 由于$\boldsymbol{d} \in \mathcal{F}^*$，所以
  $c_i(\boldsymbol{x}_k) = c_i(\boldsymbol{x}^*) + \alpha_k \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d}_k + o(\alpha_k) = 0, i \in \boldsymbol{E}$
  $c_i(\boldsymbol{x}_k) = c_i(\boldsymbol{x}^*) + \alpha_k \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d}_k + o(\alpha_k) \geqslant 0, i \in \boldsymbol{I}^*$
- 令$k \to \infty$，得条件
  $\nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} = 0, i \in \boldsymbol{E}$
  $\nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} \geqslant 0, i \in \boldsymbol{I}^*$

约束最优化问题的二阶充分条件

- 下面就$\boldsymbol{d}$的两种情形进行讨论：

- 1) 如果$\boldsymbol{d} \notin F_1^*$，则存在$i \in \boldsymbol{I}^*$，使得$\lambda_i^* > 0, \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} > 0$，从而由条件*得到$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} = \sum_{i=1}^{m} \lambda_i^* \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} > 0$，与$\nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} \leqslant 0$相矛盾。

- 2) 如果$\boldsymbol{d} \in F_1^*$，由$\boldsymbol{x}_k \in \boldsymbol{D}$可知，$c_i(\boldsymbol{x}_k) \geqslant 0, i = 1, 2, \cdots, m$，再考虑到$\lambda_i \geqslant 0, i = 1, 2, \cdots, m$
  所以$L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*) = f(\boldsymbol{x}_k) - \sum_{i=1}^{m} \lambda_i^* c_i(\boldsymbol{x}_k) \leqslant f(\boldsymbol{x}_k)$
  由KKT条件，有$L(\boldsymbol{x}_k, \boldsymbol{\lambda}^*) = f(\boldsymbol{x}^*) + \frac{1}{2}\alpha_k^2 \boldsymbol{d}_k^T \boldsymbol{W}^* \boldsymbol{d}_k + o(\alpha_k^2)$
  因为$f(\boldsymbol{x}_k) \leqslant f(\boldsymbol{x}^*)$，令$k \to \infty$，得$\boldsymbol{d}^T \boldsymbol{W}^* \boldsymbol{d} \leqslant 0$，与定理假设矛盾。
  另外，如果$F_1^* = \varnothing$，$\boldsymbol{x}^*$是最优解。

约束最优化问题举例

求解如下问题的最优解

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

思路：

1. 找到KKT点；
2. 判断KKT点是否是最优解。

约束最优化问题举例

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

KKT条件为

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

约束最优化问题举例

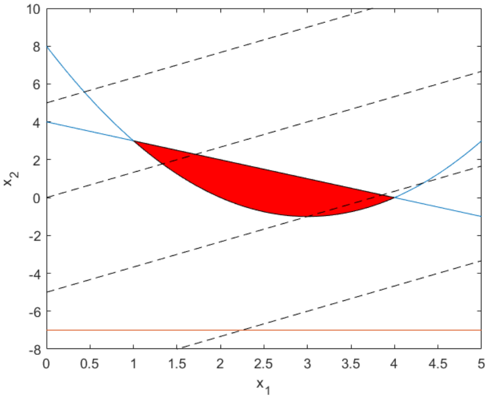


图 10: 目标函数的等高线和可行域

约束最优化问题举例

举例：

- 解方程组，得到KKT点$\boldsymbol{x}^* = (1,3)^T, \boldsymbol{\lambda}^* = (\frac{16}{3}, 0, \frac{7}{3})^T$
- 下面我们根据二阶条件，判断$\boldsymbol{x}^*$是否是最优解。在$\boldsymbol{x}^*$处，有$\boldsymbol{I}^* = \{1,3\}, \nabla c_1(\boldsymbol{x}^*) = (-1,-1)^T, \nabla c_3(\boldsymbol{x}^*) = (4,1)^T$
- 由$F_1^* = \{\boldsymbol{d} | \boldsymbol{d} \neq \boldsymbol{0}, \nabla c_i(\boldsymbol{x}^*)^T \boldsymbol{d} = 0, i = 1,3\}$
  得$F_1 = \varnothing$，所以$\boldsymbol{x}^*$是严格局部最优解。

Thank you for your attention!

翟卫欣 副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院