Kuhn-Tucker定理

➢ 若$\boldsymbol{x}^*$是局部极小点，起作用约束的$\nabla g_j(\boldsymbol{x}^*)$线性无关（约束规格），则下列条件成立：

Kuhn-Tucker条件

$$\begin{cases} \nabla f(\boldsymbol{x}^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(\boldsymbol{x}^*) = \boldsymbol{0} \\ \mu_j^* g_j(\boldsymbol{x}^*) = 0 & j = 1, 2, \ldots, l \\ \mu_j^* \geq 0 & j = 1, 2, \ldots, l \end{cases}$$

Karush(1939) Kuhn-Tucker (1951)

约束规格(constraint qualification)

- Kuhn-Tucker点：可行解中满足KT条件的点。
- 约束规格：是保证极小值为KT点的前提条件。
- 约束规格有多种情形，例如Fritz John条件中的$\mu_0>0$。
- 正则条件：起作用约束$\nabla g_{j\in J}(\boldsymbol{x}^{\#})$线性无关。
- 正则KT点：以正则条件为规格约束的KT点。
- 由于正则条件是最常见的KT条件规格约束，所以无特殊说明时，KT点都是指正则KT点。

正则点的性质

- 正则条件：起作用约束$\nabla g_{j\in J}(\boldsymbol{x}^{\#})$线性无关。
- 正则点：满足正则条件的点。
- 正则点的性质

1、正则点一定存在可行方向$\boldsymbol{p}$，满足$\nabla g_{j\in J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}>0$。

2、正则点对应的$\nabla g_{j\in J}(\boldsymbol{x})$必严格分布在某一超平面的同一侧（不包括超平面）。

3、有可行方向的点不一定是正则点。

（所有$\nabla g_{j\in J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}\geq 0$时，$\boldsymbol{p}$仍可能是可行方向）。

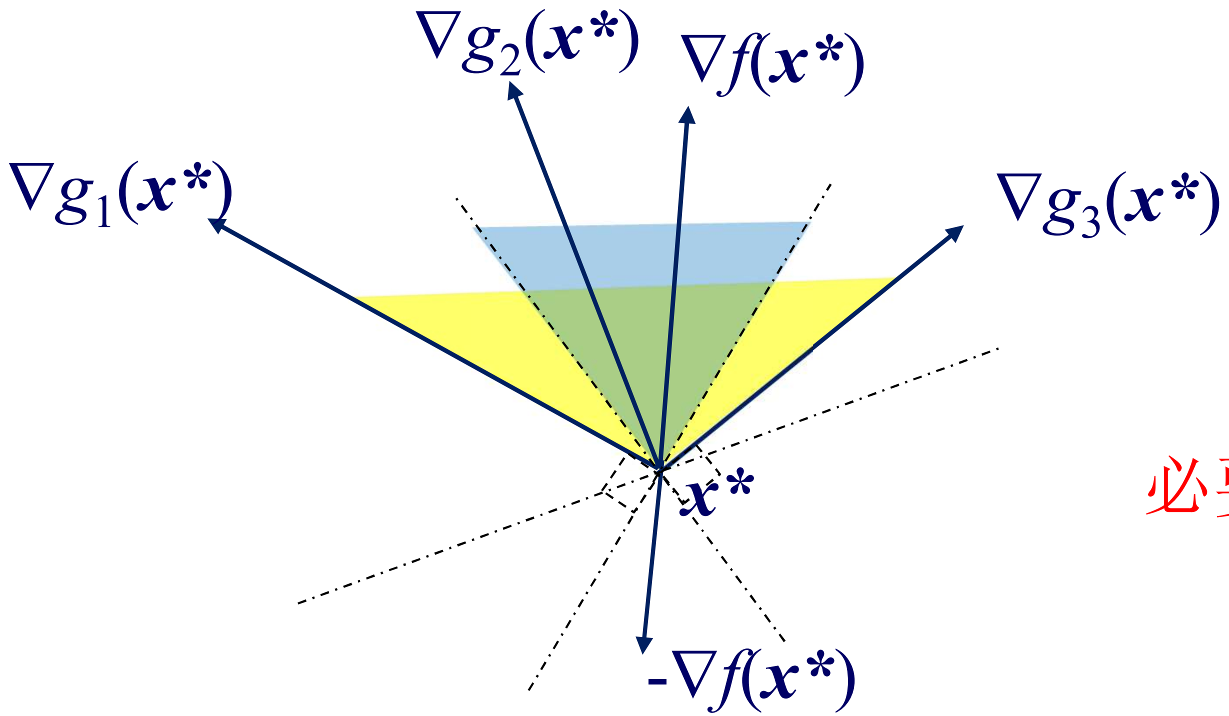
$$g(\boldsymbol{x}+\lambda\boldsymbol{p})=g(\boldsymbol{x})+\lambda\nabla g(\boldsymbol{x})^{T}\boldsymbol{p}+\frac{1}{2}\lambda^{2}\boldsymbol{p}^{T}\nabla^{2}g(\boldsymbol{x})\boldsymbol{p}+\mathrm{O}(\lambda^{2})$$

KT条件的几何意义

$$\nabla f(\boldsymbol{x}^*)=\sum_{j=1}^{l}\mu_j^*\nabla g_j(\boldsymbol{x}^*)\in\left\{\sum_{j=1}^{l}\mu_j\nabla g_j(\boldsymbol{x}^*)\,\middle|\,\mu_j\ge 0,\ j=1,2,\cdots,l\right\}$$

闭锥包

锥组合



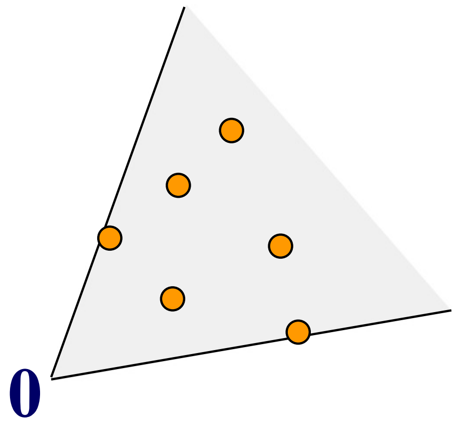
必要不充分条件

锥、凸锥与锥包

- 锥：$\forall \boldsymbol{x}\in C$和$\mu\geq 0$，均有$\mu\boldsymbol{x}\in C$，则集合C称为锥。
- 锥组合 $\sum_{j=1}^{l}\mu_j\boldsymbol{x}_j \quad \mu_j\geq 0$ $\boldsymbol{x}_j\in C$的非负线性组合
- 凸锥：$\forall \boldsymbol{x}_1,\boldsymbol{x}_2\in C$和$\mu_1,\mu_2\geq 0$，均有$\mu_1\boldsymbol{x}_1+\mu_2\boldsymbol{x}_2\in C$
  - 性质1：凸锥既是凸集又是锥。
  - 性质2：凸锥中任意有限个元素的锥组合仍在凸锥中（凸锥判定的充要条件）
- 集合C的锥包：集合C中所有锥组合的集合

$$\left\{\sum_{j=1}^{l}\mu_j\boldsymbol{x}_j \,\middle|\, \boldsymbol{x}_j\in C,\mu_j\geq 0,\ j=1,2,\cdots,l\right\}$$

包含了C的最小凸锥。



Kuhn-Tucker条件的使用

$$\nabla f(\boldsymbol{x}^{\#}) - \sum_{j=1}^{l} \mu_j^{\#} \nabla g_j(\boldsymbol{x}^{\#}) = \mathbf{0}$$

$$\mu_j^{\#} g_j(\boldsymbol{x}^{\#}) = 0$$

$$\mu_j^{\#} \geq 0 \qquad j = 1, 2, \ldots, l$$

1、实际应用KT条件时，无需事先检验$\nabla g_{j\in J}(\boldsymbol{x}^{\#})$是否线性无关，此时可找到所有对应$\mu_0>0$的Fritz John点，包括$\nabla g_{j\in J}(\boldsymbol{x}^{\#})$线性相关的情况。所以应用KT条件时，可以得到不满足正则条件的KT点（$\mu_0>0$）。

2、如果$\nabla g_{j\in J}(\boldsymbol{x}^{\#})$正线性相关时，KT条件不能找到所有的极小值点，所以KT条件并不是所有极小值点的必要条件。

孤立点举例

孤点

***x**=0**

非驻点孤点1

$$f_1(\boldsymbol{x})=x_1$$
$$\text{s.t.}\quad x_2 \ge x_1^2$$
$$x_2 \le 0$$

非驻点孤点2

$$f_2(\boldsymbol{x})=x_2$$
$$\text{s.t.}\quad x_2 \ge x_1^2$$
$$x_2 \le 0$$

驻点孤点

$$f_3(\boldsymbol{x})=x_1^2+x_2^2$$
$$\text{s.t.}\quad x_2 \ge x_1^2$$
$$x_2 \le 0$$

$$g_1(\boldsymbol{X})=-x_1^2+x_2 \quad\Rightarrow\quad \nabla g_1(\boldsymbol{x})=\begin{bmatrix}-2x_1\\1\end{bmatrix}$$

$$g_2(\boldsymbol{X})=-x_2 \quad\Rightarrow\quad \nabla g_2(\boldsymbol{x})=\begin{bmatrix}0\\-1\end{bmatrix}$$

在***x**=0**正线性相关

$$f_1(\boldsymbol{X})=x_1 \quad\Rightarrow\quad \nabla f_1(\boldsymbol{x})=\begin{bmatrix}1\\0\end{bmatrix}$$

无解

$$f_2(\boldsymbol{X})=x_2 \quad\Rightarrow\quad \nabla f_2(\boldsymbol{x})=\begin{bmatrix}0\\1\end{bmatrix}$$

无穷多个非负解

满足$\mu_1-\mu_2=1$

$$f_3(\boldsymbol{X})=x_1^2+x_2^2 \quad\Rightarrow\quad \nabla f_2(\boldsymbol{x})=\begin{bmatrix}2x_1\\2x_2\end{bmatrix}$$

无穷多个非负解

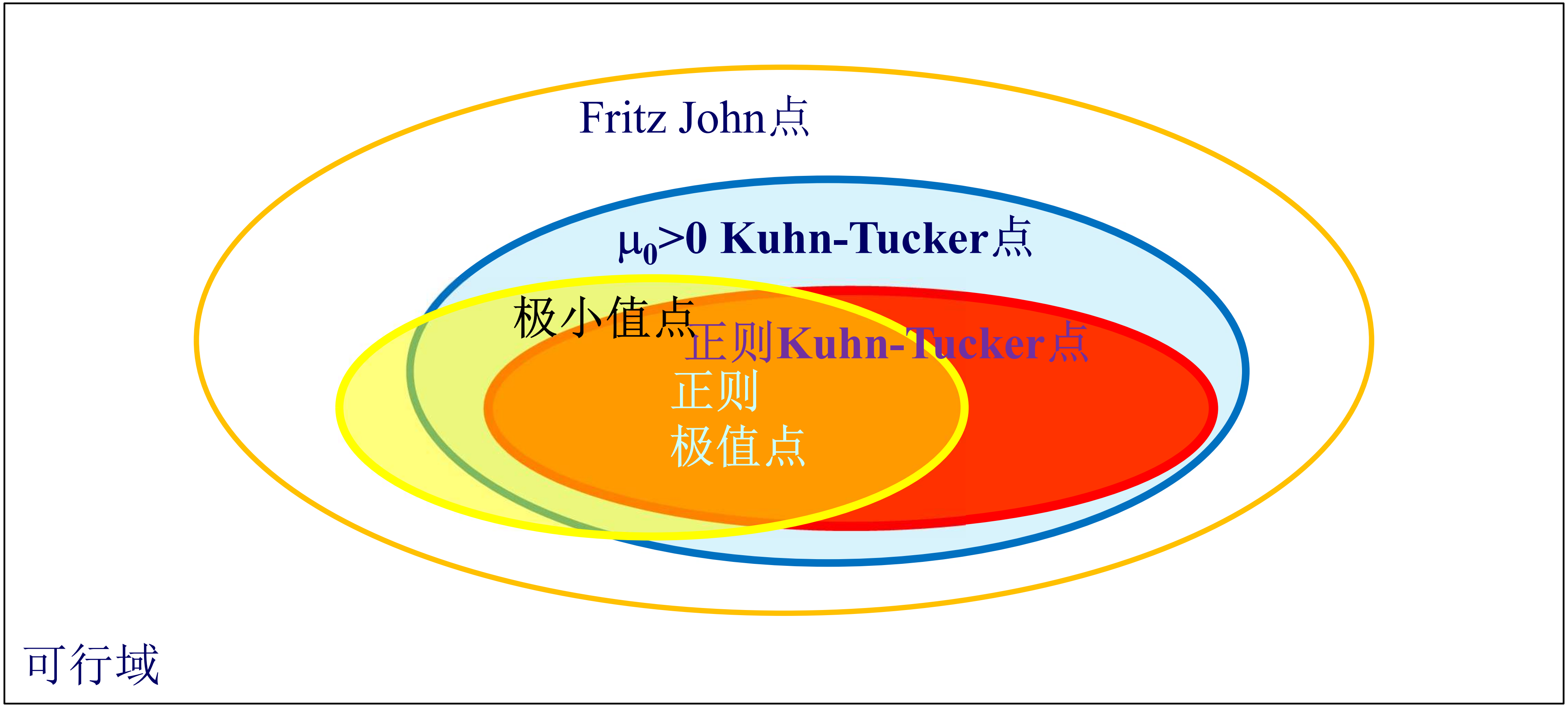
满足$\mu_1=\mu_2$

包括 $\mu_1=0$， $\mu_2=0$

Kuhn-Tucker条件：

$$\mu_1\nabla g_1(\boldsymbol{0})+\mu_2\nabla g_2(\boldsymbol{0})=\mu_1\begin{bmatrix}0\\1\end{bmatrix}+\mu_2\begin{bmatrix}0\\-1\end{bmatrix}=\nabla f(\boldsymbol{0})$$

不同点集的关系



问题：KT条件会丢失哪些极小值点？

FJ点与KT点关系

1、$\mu_0=0$（缺$f(\boldsymbol{x})$信息）

$\Rightarrow \nabla g_{j\in J}(\boldsymbol{x})$正线性相关

$\Rightarrow$不存在所有$\nabla g_{j\in J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}>0$的可行方向

1.1、无可行方向：孤立点（有部分也同时是**KT**点）

1.2、若存在可行方向$\boldsymbol{p}$，则所有$\nabla g_{j\in J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}\geq 0$，若$\boldsymbol{x}$是极值点，此时$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p}\geq 0$，可在2.1.1情形中找到等价KT点解。

$$\mu_0 \nabla f(\boldsymbol{x}^{\#})^T \boldsymbol{p} - \sum_{j\in J} \mu_j \nabla g_j(\boldsymbol{x}^{\#})^T \boldsymbol{p} = \boldsymbol{0}$$

2、$\mu_0>0$：**KT点**

2.1、$\nabla g_{j\in J}(\boldsymbol{x})$线性相关：非正则**KT**点

2.1.1 $\nabla g_{j\in J}(\boldsymbol{x})$正线性相关：同$\mu_0=0$分析。

2.1.2 $\nabla g_{j\in J}(\boldsymbol{x})$其他线性相关：存在约束冗余，可将冗余约束j按不起作用约束对待，即$\mu_j=0$，约简为$\nabla g_{j\in J}(\boldsymbol{x})$线性无关。

2.2、$\nabla g_{j\in J}(\boldsymbol{x})$线性无关：正则**KT**点

结论

1、KT点仅丢失了部分孤立极小值点。

2、正则KT点丢失了$\nabla g_{j\in J}(\boldsymbol{x})$线性相关部分的极小值点。

3、KT条件可找到少数孤点外的所有极小点外。

4、KT定理仅可找到正则KT点，是正则极小点的必要条件。