第 6 章 约束非线性优化

第一节 最优性条件

一、起作用约束与可行下降方向

一般形式的约束非线性规划模型

$$\begin{aligned}&\min f(\mathbf{x})\\ &\text{s.t. } h_i(\mathbf{x})=0, \quad i=1,2,\dots,m\\ &\qquad g_j(\mathbf{x})\geq 0, \quad j=1,2,\dots,l\end{aligned} \tag{1}$$

约定：函数$f(\mathbf{x})$，$h_i(\mathbf{x})$，$g_j(\mathbf{x})$有一阶连续偏导数。

对约束极小化问题，下降迭代算法除了要使目标函数每次迭代有所下降而外，还要满足约束条件，保证可行性。

✧起作用约束与不起作用约束

考虑仅含不等式约束的问题

$$\begin{aligned} &\min f(\mathbf{x}) \\ &\text{s.t. } g_j(\mathbf{x}) \geq 0, \qquad j = 1,2,\dots,l \end{aligned} \tag{2}$$

记可行域：$\boldsymbol{R} = \{\mathbf{x} | g_j(\mathbf{x}) \geq 0, j = 1,2,\dots,l\}$。

设$\mathbf{x}^{(0)}$是一个可行解，对于某不等式约束条件，$\mathbf{x}^{(0)}$满足该条件有两种可能：

$$g_j(\mathbf{x}^{(0)}) > 0 \text{ 或 } g_j(\mathbf{x}^{(0)}) = 0$$

情况（1）：$g_j(\mathbf{x}^{(0)}) > 0$，即严格不等式成立，则$\mathbf{x}^{(0)}$不在该约束形成的可行域边界上，约束“$g_j(\mathbf{x}) \geq 0$”对$\mathbf{x}^{(0)}$的微小摄动不起限制作用，称该约束条件是$\mathbf{x}^{(0)}$点的不起作用约束（或称无效约束/非积极约束；**inactive**）。

情况（2）：$g_j(\mathbf{x}^{(0)}) = 0$，即等式成立，则$\mathbf{x}^{(0)}$处于该约束形成的可行域边界上，称该约束是$\mathbf{x}^{(0)}$点的起作用约束（有效约束/积极约束；**active**）。

注 1：等式约束对所有可行点来说都是起作用约束。

✧可行方向

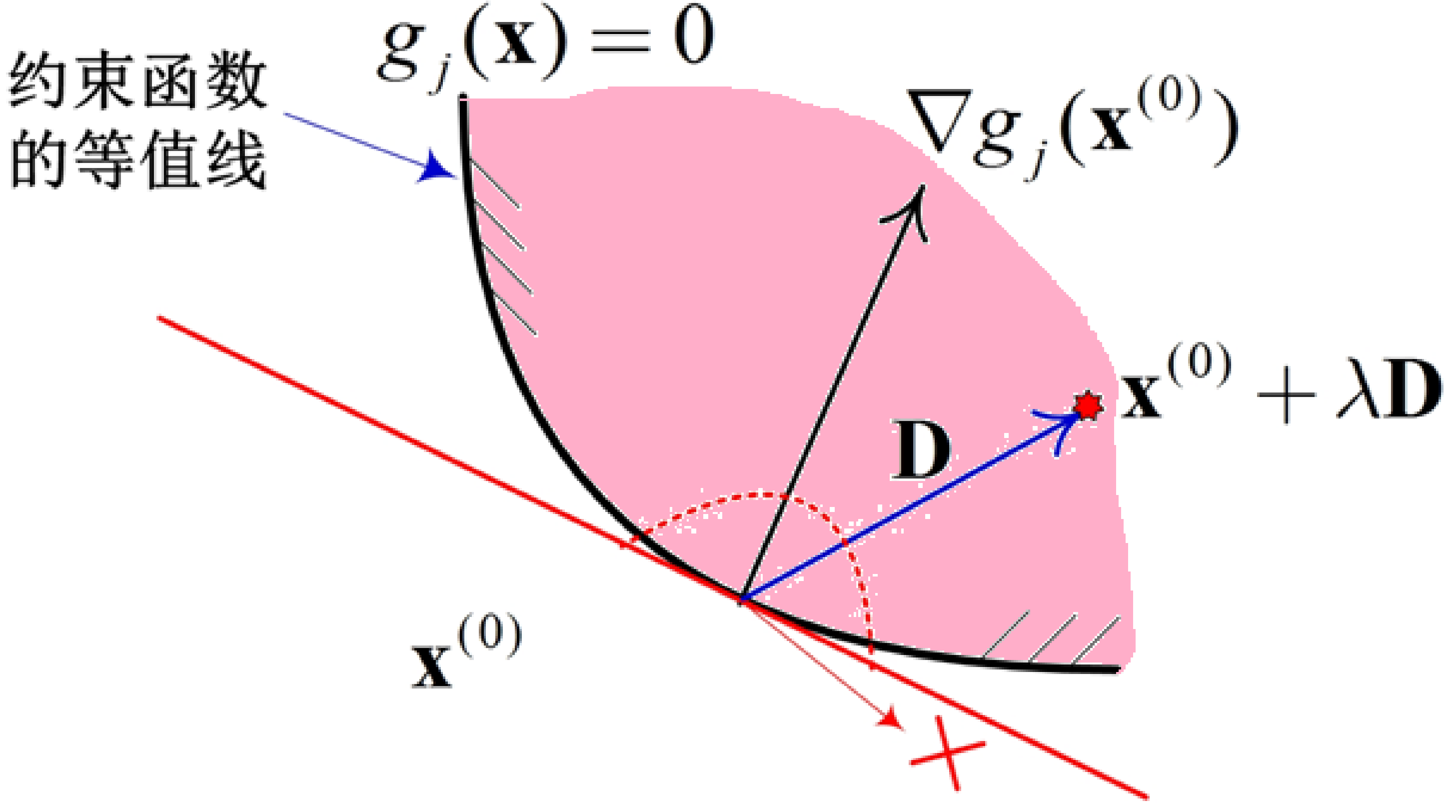
设$\mathbf{x}^{(0)}$是非线性规划问题（2）的一个可行点，$\boldsymbol{R}$为可行域，考虑从该点出发的某一方向向量 $\mathbf{D}$。

若存在实数$\lambda_0>0$，使对任意$\lambda\in[0,\lambda_0]$均有

$$\mathbf{x}^{(0)}+\lambda\mathbf{D}\in\boldsymbol{R} \qquad (3)$$

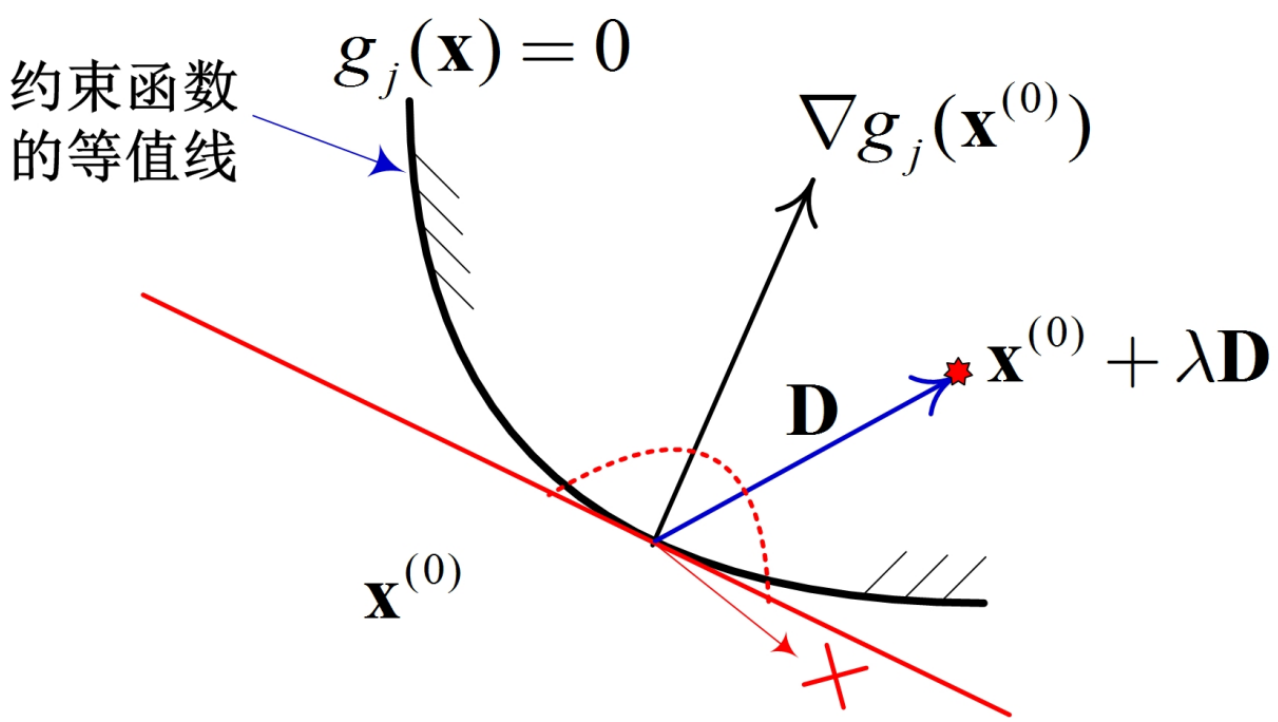
就称方向$\mathbf{D}$是$\mathbf{x}^{(0)}$点的一个可行方向。

若$\mathbf{x}^{(0)}$位于约束边界上，即$g_j(\mathbf{x}^{(0)})=0$为起作用约束，那么从$\mathbf{x}^{(0)}$沿方向$\mathbf{D}$移动时，就必须要求函数$g_j(\mathbf{x})$的值至少不减小，才能有$g_j(\mathbf{x}^{(0)}+\lambda\mathbf{D})\geq0$，从而继续保持可行。



在 $\mathbf{x}^{(0)}$ 点，可行移动方向 $\mathbf{D}$ 和约束函数 $g_j(\mathbf{x})$ 的梯度 $\nabla g_j(\mathbf{x}^{(0)})$的夹角就不应为钝角；否则$g_j(\mathbf{x})$就会从0减小从而小于0，结果新的点$\mathbf{x}^{(0)}+\lambda\mathbf{D}$不可行。

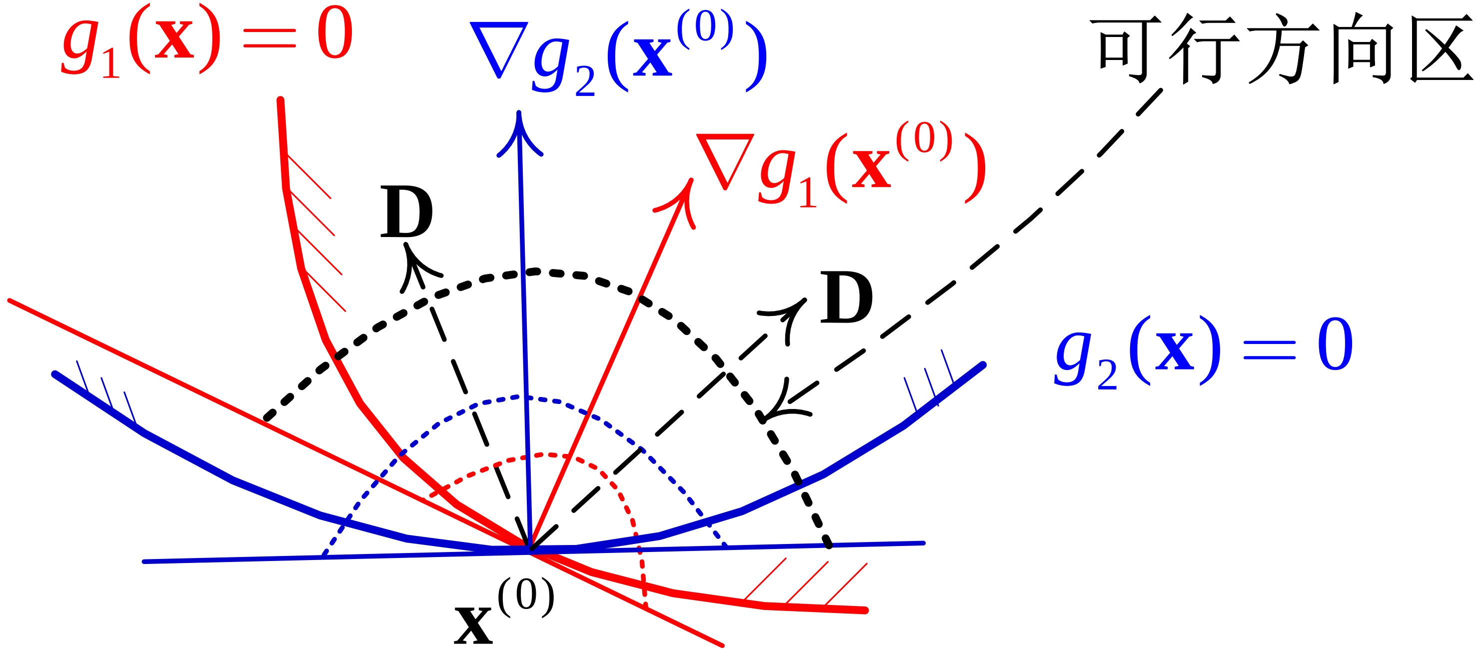
如果 $g_j(\mathbf{x})$ 是非线性函数，则夹角也不能是直角，于是下述严格不等式必须成立：



$$\nabla g_j(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D} > 0 \qquad (4)$$

如此，才可能在λ足够小时必然保证：$g_j(\mathbf{x}^{(0)}+\lambda)\geq 0$成立，因而方向**D**才是可行方向。

推广到多个积极约束的情况：



若**D**是可行点$\mathbf{x}^{(0)}$处的可行方向，则在该点，所有起作用约束$g_j(\mathbf{x}^{(0)})=0$，它们的梯度方向均应满足

$$\nabla g_j(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D}>0, j\in \boldsymbol{J} \qquad (5)$$

其中，$\boldsymbol{J}$为所有起作用约束下标的集合。

对于不起作用约束：$g_j(\mathbf{x}^{(0)}) > 0$，考虑泰勒展开

$$g_j(\mathbf{x}^{(0)} + \lambda\mathbf{D}) = g_j(\mathbf{x}^{(0)}) + \lambda\nabla g_j(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D} + o(\lambda)$$

当 $\lambda > 0$ 足够小时，因为 $g_j(\mathbf{x}^{(0)}) > 0$，故必能保证 $g_j(\mathbf{x}^{(0)} + \lambda\mathbf{D}) \geq 0$成立。

因此，只须方向**D**和积极约束的梯度满足（**5**）式：$\nabla g_j(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D} > 0, j \in \boldsymbol{J}$，就必然能保证沿方向**D**以某个步长前进之后仍然得到可行解。

满足（**5**）式的方向**D**是$\mathbf{x}^{(0)}$点的可行方向（非积极约束可暂不考虑）。

✧ 下降方向与可行下降方向

考虑某可行点$\mathbf{x}^{(0)}$和某方向$\mathbf{D}$，若存在实数$\lambda_0' > 0$，使对任意$\lambda \in (0, \lambda_0']$，均有$f(\mathbf{x}^{(0)} + \lambda \mathbf{D}) < f(\mathbf{x}^{(0)})$，则称方向$\mathbf{D}$为$\mathbf{x}^{(0)}$点的下降方向。

下降方向应和目标函数负梯度方向成锐角，即

$$-\nabla f(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D} > 0 \Leftrightarrow \nabla f(\mathbf{x}^{(0)})^{\mathrm{T}}\mathbf{D} < 0 \qquad (6)$$

满足（6）式的方向$\mathbf{D}$必为$\mathbf{x}^{(0)}$点的下降方向。

可行下降方向：如果方向$\mathbf{D}$在$\mathbf{x}^{(0)}$点既是可行方向，又是下降方向，就称它是该点的可行下降方向[1]。

---

[1] 显然，若某点存在可行下降方向，它就不会是极小点；若为极小点，则在该点就不应存在可行下降方向。

定理 1 如果$\mathbf{x}^*$是非线性规划（2）的局部极小点，目标函数$f(\mathbf{x})$在$\mathbf{x}^*$处可微。且对于$j \in \boldsymbol{J}$（积极约束集），$g_j(\mathbf{x})$在$\mathbf{x}^*$处可微；当$j \notin \boldsymbol{J}$（非积极约束集），$g_j(\mathbf{x})$在$\mathbf{x}^*$处连续。

那么，在局部极小点$\mathbf{x}^*$处必然不存在可行下降方向，即不存在向量D同时满足：

$$\begin{aligned} &\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{D} < 0 \\ &\nabla g_j(\mathbf{x}^*)^{\mathrm{T}}\mathbf{D} > 0, j \in \boldsymbol{J} \end{aligned} \quad (7)$$

二、最优性条件——KKT 点

✧ 最优点与可行域边界

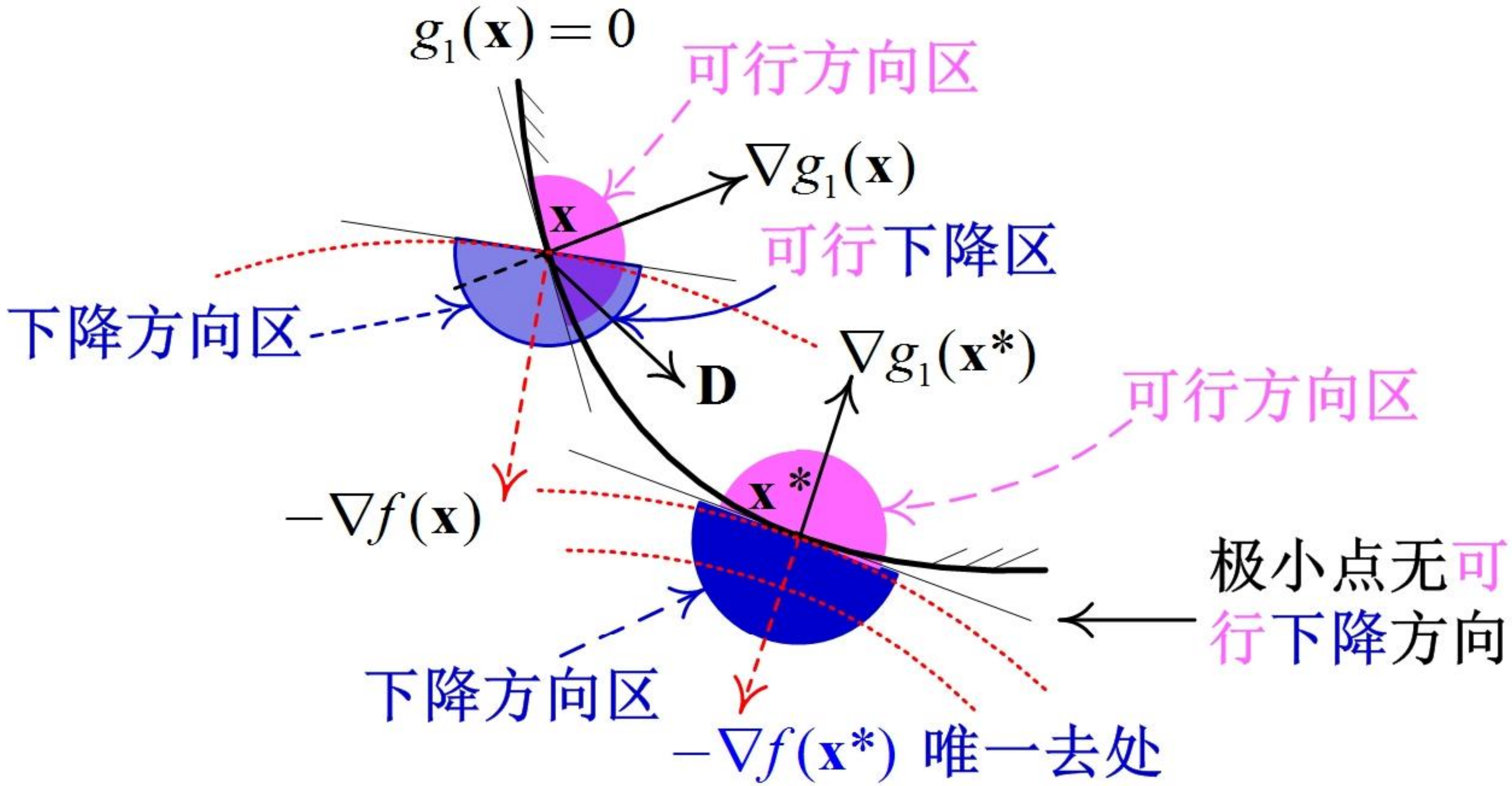
若$\mathbf{x}^*$为极小点，但起作用约束有 0 个，则必有

$$\nabla f(\mathbf{x}^*) = \mathbf{0}\text{——即无约束的情况}$$

若在极小点$\mathbf{x}^*$处，只有 1 个积极约束，不妨设积极约束为

$$g_1(\mathbf{x}^*) = 0$$

则该积极约束的梯度$\nabla g_1(\mathbf{x}^*)$必然与目标函数的负梯度$-\nabla f(\mathbf{x}^*)$在一条直线上且方向相反；否则，在该点就一定存在可行下降方向：

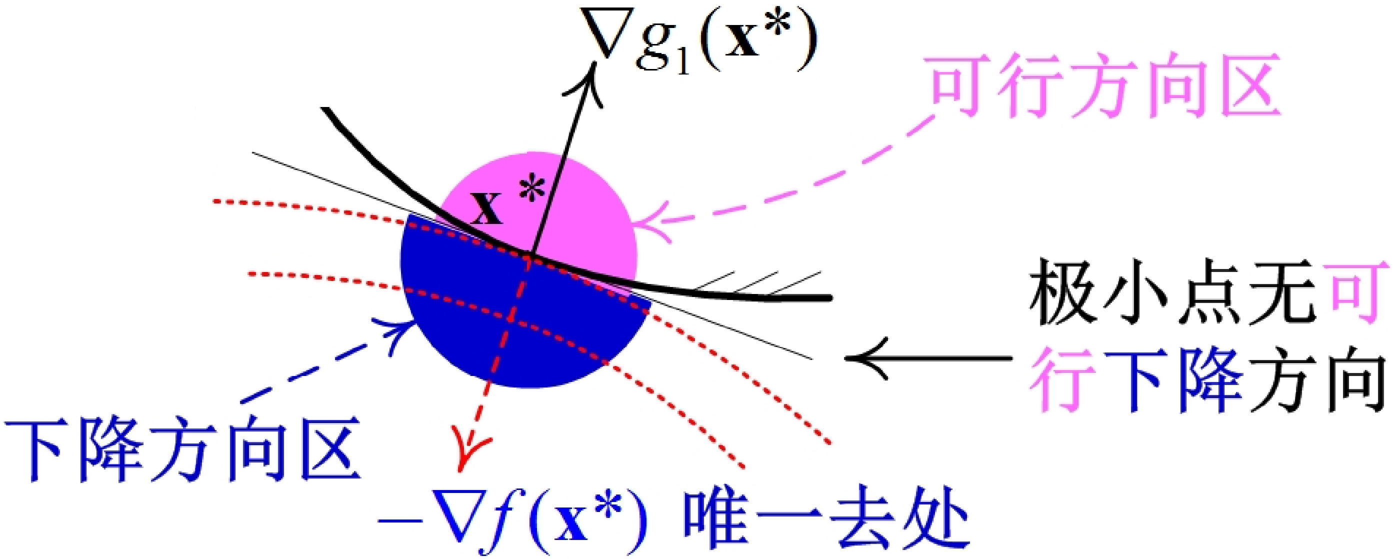


图中 $\mathbf{x}$ 点不满足上述要求，不是极小点。

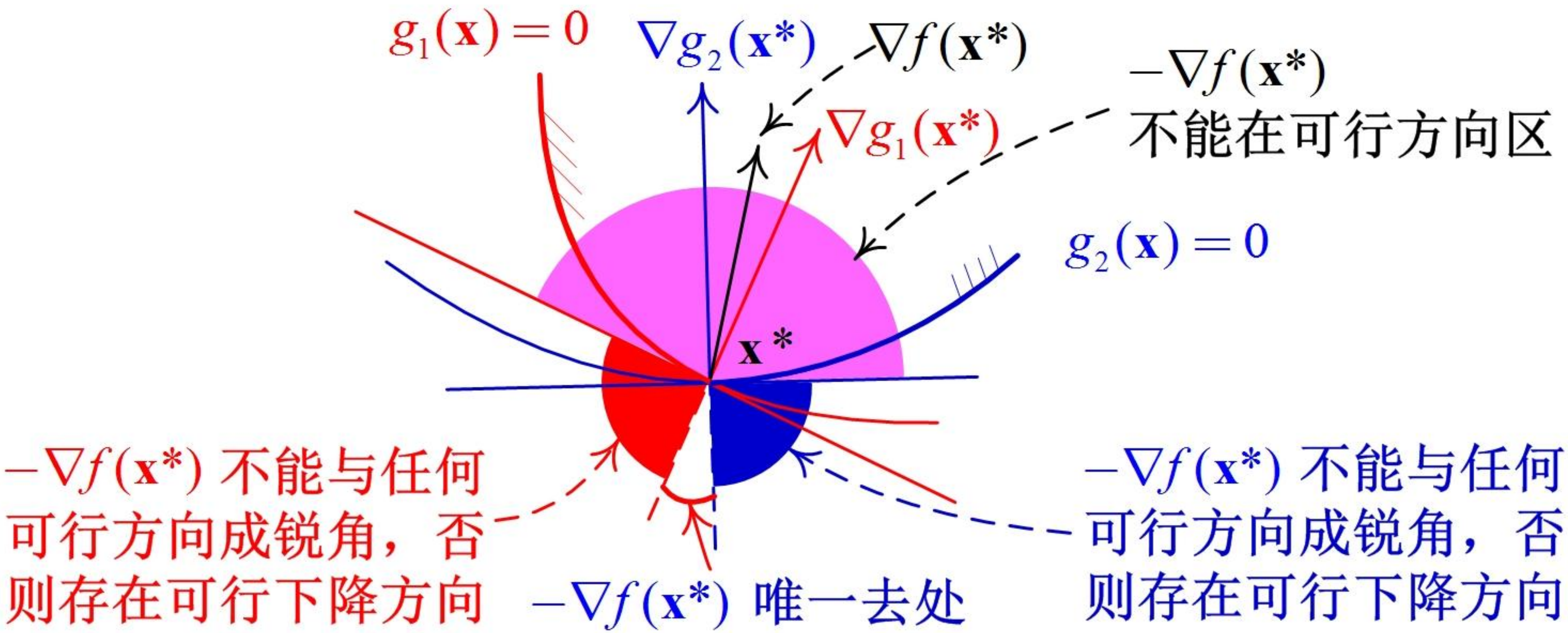
而在$\mathbf{x}^*$：$\nabla g_1(\mathbf{x}^*)$与$-\nabla f(\mathbf{x}^*)$共线且方向相反，因此可行方向区和下降方向区无交集，所以不存在可行下降方向，因此$\mathbf{x}^*$是极小点。

在极小点$\mathbf{x}^*$处，因为$\nabla g_1(\mathbf{x}^*)$与$-\nabla f(\mathbf{x}^*)$共线，所以存在正实数γ_1，有：

$$-\nabla f(\mathbf{x}^*) = -\gamma_1 \nabla g_1(\mathbf{x}^*) \Leftrightarrow \nabla f(\mathbf{x}^*) - \gamma_1 \nabla g_1(\mathbf{x}^*) = \mathbf{0}$$



对于多个积极约束，在极小点$\mathbf{x}^*$处：



极小点$\mathbf{x}^*$的负梯度方向$-\nabla f(\mathbf{x}^*)$就只能位于约束函数梯度的反向延长线夹角内。

在极小点$\mathbf{x}^*$处，若$\nabla g_1(\mathbf{x}^*)$和$\nabla g_2(\mathbf{x}^*)$线性无关（称为正则条件，约束函数梯度线性无关的点称为正则点），那么二者夹角内的$\nabla f(\mathbf{x}^*)$必能表示为

$$\nabla f(\mathbf{x}^*) = \gamma_1 \nabla g_1(\mathbf{x}^*) + \gamma_2 \nabla g_2(\mathbf{x}^*), \gamma_1 \geq 0, \gamma_2 \geq 0$$

即，

$$\nabla f(\mathbf{x}^*) - \gamma_1 \nabla g_1(\mathbf{x}^*) - \gamma_2 \nabla g_2(\mathbf{x}^*) = \mathbf{0}$$

一般，对问题（2）的所有起作用约束，在极小点$\mathbf{x}^*$处，有

$$\begin{aligned} \nabla f(\mathbf{x}^*) - \textstyle\sum_{j \in \boldsymbol{J}} \gamma_j \nabla g_j(\mathbf{x}^*) = \mathbf{0} \\ \gamma_j \geq 0, \text{for } j \in \boldsymbol{J} \end{aligned} \tag{8}$$

✧库恩-塔克条件（约束问题最优点一阶必要条件）

为统一起见，考虑将不起作用约束也包括进来，条件(8)可等价修改为

$$\begin{cases} \nabla f(\mathbf{x}^*) - \sum_{j=1}^{l} \gamma_j \nabla g_j(\mathbf{x}^*) = \mathbf{0} \\ \gamma_j g_j(\mathbf{x}^*) = 0 \\ \gamma_j \geq 0 \end{cases} \tag{9}$$

其中，若$g_j(\mathbf{x}^*) = 0$（积极约束），则γ_j可不为零；若$g_j(\mathbf{x}^*) > 0$（非积极约束），则必有$\gamma_j = 0$，一阶必要条件就约简为（8）式。

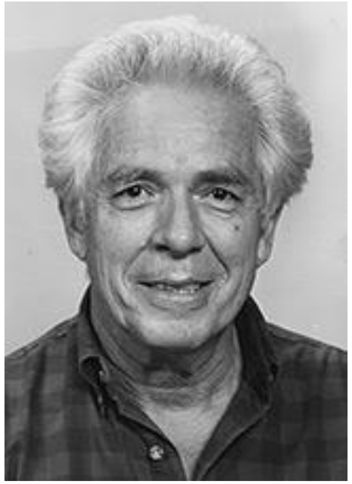
上述条件就是约束非线性优化问题最优点处成立的一阶必要条件，由 **Kuhn** 和 **Tucker** 提出（**1951**），称为：库恩-塔克（**Kuhn-Tucker**，**K-T**）条件。

只要极小点是正则点（即该处起作用约束的梯度线性无关），则在最优点就必成立 **K-T** 条件。

注 **2**：**K-T** 条件并不是约束极小化问题最优解的充分条件，即满足这个条件的点不一定是最优点；

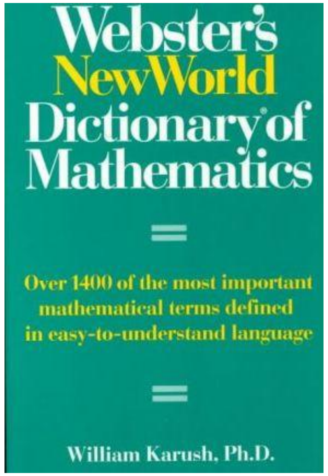
注 **3**：对于凸规划，**K-T** 条件既是最优点存在的必要条件，同时也是充分条件。

K-T 条件的历史



Harold W. Kuhn（1925-2014）

Albert W. Tucker（1905-1995）



William Karush（1917-1997）

1951 年,K-T 条件由 Kuhn 和 Tucker 二人发表在 Berkley 举办的一次学术会议上，之后就命名为库恩-塔克条件（Kuhn-Tucker conditions，K-T 条件）;而早在 1939 年，上述结果就由 Karush 在其硕士论文中提出，但当时并未引起重视；为纪念 Karush 的贡献，K-T 条件也合称为 Karush-Kuhn-Tucker conditions（简称 KKT 条件）。

K-T 条件（KKT 条件）的完整表述：

设$\mathbf{x}^*$是非线性规划问题（2）的极小点，且$\mathbf{x}^*$为正则点（各起作用约束的梯度线性无关），则下述条件成立：

$$\begin{cases} \nabla f(\mathbf{x}^*) - \sum_{j=1}^{l} \gamma_j^* \nabla g_j(\mathbf{x}^*) = \mathbf{0} \\ \gamma_j^* g_j(\mathbf{x}^*) = 0,\ j = 1,2,\dots,l \\ \gamma_j^* \geq 0,\ j = 1,2,\dots,l \end{cases} \qquad (10)$$

满足条件（10）的点称为 **K-T 点（KKT 点）**。

对于包含等式约束的非线性规划，用$h_i(\mathbf{x}) \geq 0$和$-h_i(\mathbf{x}) \geq 0$等价替换$h_i(\mathbf{x}) = 0$，套用（10）式，即可得到同时包含等式和不等式约束的 K-T 条件。

带等式约束的K-T条件（KKT条件）：

设$\mathbf{x}^*$是非线性规划（1）的极小点，且$\mathbf{x}^*$处所有起作用约束的梯度 $\nabla h_i(\mathbf{x}^*), i=1,2,\dots,m$ 和 $\nabla g_j(\mathbf{x}^*), j\in \boldsymbol{J}$ 线性无关，则存在向量

$$\boldsymbol{\Gamma}^* = \{\gamma_1^*, \gamma_2^*, \dots, \gamma_l^*\}\text{和}\boldsymbol{\Lambda}^* = \{\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*\}$$

使下述条件成立：

$$\begin{cases} \nabla f(\mathbf{x}^*) - \sum_{j=1}^{l}\gamma_j^*\nabla g_j(\mathbf{x}^*) - \sum_{i=1}^{m}\lambda_i^*\nabla h_i(\mathbf{x}^*) = \mathbf{0} & \text{(11a)} \\ \gamma_j^* g_j(\mathbf{x}^*) = 0,\ j=1,2,\dots,l & \text{(11b)} \\ \gamma_j^* \geq 0,\ j=1,2,\dots,l & \text{(11c)} \end{cases}$$

（10）式、（11）式中的$\gamma_1^*, \gamma_2^*, \dots, \gamma_l^*$和$\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*$，称为拉格朗日（Lagrange）乘子——KKT 条件就是拉格朗日乘子法在考虑不等式约束时的推广。

（11a）式称为梯度条件；（11b）称为互补松弛条件；（11c）为非负条件。

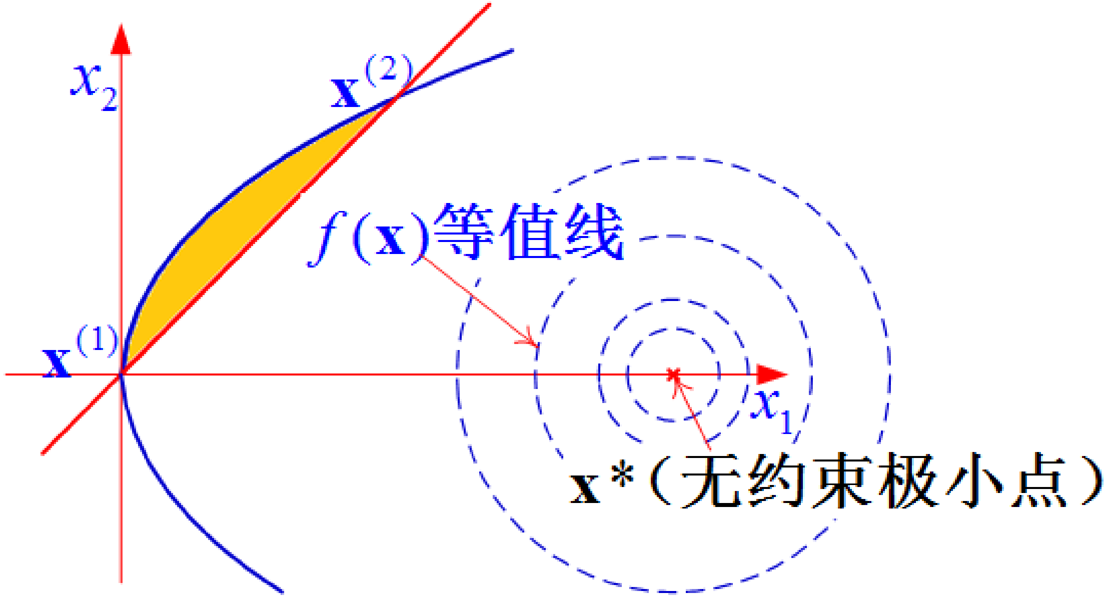
注意，$\geq$类不等式约束的拉氏乘子γ_j^*需要≥ 0，等式约束的拉氏乘子λ_i^*没有正负限制。

例 1，考虑问题

$$\min f(\mathbf{x}) = (x_1 - 2)^2 + x_2^2$$

$$\text{s.t.}\ g_1(\mathbf{x}) = x_1 - x_2^2 \ge 0$$

$$g_2(\mathbf{x}) = -x_1 + x_2 \ge 0$$



验证点$\mathbf{x}^{(1)} = (0,0)^{\mathrm{T}}$和点$\mathbf{x}^{(2)} = (1,1)^{\mathrm{T}}$是否满足 **KKT 条件**？

解，注意到上述模型已经是标准形式，直接求目标函数和约束函数的梯度：$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 - 4 \\ 2x_2 \end{bmatrix}$，$\nabla g_1(\mathbf{x}) = \begin{bmatrix} 1 \\ -2x_2 \end{bmatrix}$，$\nabla g_2(\mathbf{x}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$。

（1）验证$\mathbf{x}^{(1)}$。在$\mathbf{x}^{(1)}$点，$g_1(\mathbf{x}) \geq 0$和$g_2(\mathbf{x}) \geq 0$都是起作用约束，目标函数和约束函数的梯度为

$$\nabla f(\mathbf{x}^{(1)}) = \begin{bmatrix} -4 \\ 0 \end{bmatrix},\ \nabla g_1(\mathbf{x}^{(1)}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \nabla g_2(\mathbf{x}^{(1)}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}。$$

令，

$$\begin{bmatrix} -4 \\ 0 \end{bmatrix} - \gamma_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \gamma_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

得到

$\gamma_1 = -4$，$\gamma_2 = 0$。由于$\gamma_1 < 0$，故$\mathbf{x}^{(1)}$不是 **KKT 点**。

（2）验证$\mathbf{x}^{(2)}$。在$\mathbf{x}^{(2)}$，$g_1(\mathbf{x}) \geq 0$和$g_2(\mathbf{x}) \geq 0$都是起作用约束，目标函数和约束函数的梯度为

$$\nabla f(\mathbf{x}^{(1)}) = \begin{bmatrix} -2 \\ 2 \end{bmatrix},\ \nabla g_1(\mathbf{x}^{(1)}) = \begin{bmatrix} 1 \\ -2 \end{bmatrix},\ \nabla g_2(\mathbf{x}^{(1)}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}。$$

令，

$$\begin{bmatrix} -2 \\ 2 \end{bmatrix} - \gamma_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} - \gamma_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

得到$\gamma_1 = 0 \geq 0$，$\gamma_2 = 2 \geq 0$。故$\mathbf{x}^{(2)}$是 **KKT 点** [2]。

---

[2] KKT 点要看起作用约束，其他约束皆浮云。

例 2，求下述问题的 **K-T 点**，并判断是否为全局极小？

$$\begin{aligned} \min f(\mathbf{x}) &= x_1^2 + (x_2-4)^2 \\ \text{s.t.}\quad -x_2 + (x_1-2)^2 &\le 0 \\ 2x_1 - x_2 - 1 &= 0 \end{aligned}$$

解，注意第一个约束，须先化<u>标准形</u>

$$g_1(\mathbf{x}) = x_2 - (x_1-2)^2 \ge 0$$

对于标准形式，写出目标函数和约束函数的梯度：

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 - 8 \end{bmatrix}, \nabla g(\mathbf{x}) = \begin{bmatrix} -2x_1 + 4 \\ 1 \end{bmatrix}, \nabla h(\mathbf{x}) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

K-T 条件（注意**将约束条件也考虑进来**）：

$$
\begin{cases}
2x_1-\gamma(-2x_1+4)-2\lambda=0\\
2x_2-8-\gamma+\lambda=0\\
\gamma[x_2-(x_1-2)^2]=0\\
x_2-(x_1-2)^2\geq 0\\
2x_1-x_2-1=0\\
\gamma\geq 0
\end{cases}
$$

分两种情况讨论：

情况 1：若$\gamma > 0$，根据互补松弛条件和约束条件，有

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

解得$\begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$和$\begin{cases} x_1 = 5 \\ x_2 = 9 \end{cases}$。

再利用梯度条件，解得$\begin{cases} \gamma = -5/2 \\ \lambda = \quad 7/2 \end{cases}$和$\begin{cases} \gamma = -15/2 \\ \lambda = -35/2 \end{cases}$，$\gamma$不满足非负要求，故这两个点不是 **K-T 点**。

情况 2：若$\gamma=0$，解下列联立方程组

$$\begin{cases}2x_1-2\lambda=0\\2x_2-8+\lambda=0\\2x_1-x_2-1=0\end{cases}$$

得到$x_1=2,x_2=3,\lambda=2$，且满足不等式约束

$$x_2-(x_1-2)^2=3>0$$

故$\begin{cases}x_1=2\\x_2=3\end{cases}$为 **K-T 点**，相应的 **Lagrange 乘子**为$\begin{cases}\gamma=0\\\lambda=2\end{cases}$。

上述问题的目标函数为凸函数，“≥”形式的约束函数为凹函数，等式约束为线性函数，因此原问题为凸规划，其 **K-T** 点就是全局最小点。

例 3，求下述问题的 KKT 点。

$$\begin{aligned} \max\ & -(x_1-2)^2-(x_2-1)^2 \\ \text{s.t.}\ & -{x_1}^2+x_2 \geq 0 \\ & x_1+x_2-2 \leq 0 \end{aligned}$$

解，注意目标函数和第二个约束，须先化<u>标准形</u>：

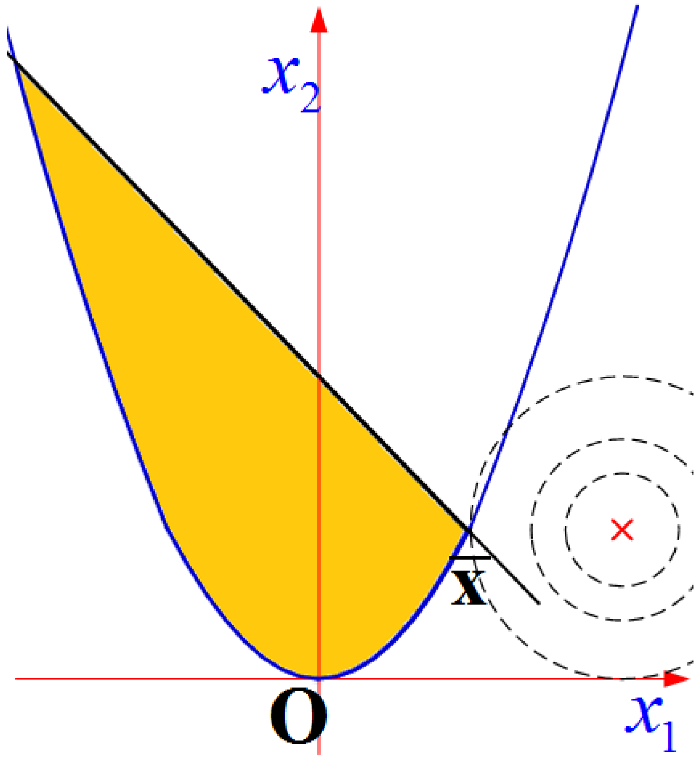
$$\rightarrow \quad \begin{aligned} \min\ & (x_1-2)^2+(x_2-1)^2 \\ \text{s.t.}\ & -{x_1}^2+x_2 \geq 0 \\ & -x_1-x_2+2 \geq 0 \end{aligned}$$

标准型中，目标函数和约束函数的**梯度**：

$$\nabla f(\mathbf{x})=\begin{bmatrix} 2(x_1-2) \\ 2(x_2-1) \end{bmatrix},\ \nabla g_1(\mathbf{x})=\begin{bmatrix} -2x_1 \\ 1 \end{bmatrix},\ \nabla g_2(\mathbf{x})=\begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

KKT 条件（注意将约束条件也考虑进来）：

$$\begin{cases} 2(x_1-2)+2\gamma_1 x_1+\gamma_2=0 \\ 2(x_2-1)-\gamma_1+\gamma_2=0 \\ \gamma_1(-{x_1}^2+x_2)=0 \\ \gamma_2(-x_1-x_2+2)=0 \\ -{x_1}^2+x_2\geq 0 \\ -x_1-x_2+2\geq 0 \\ \gamma_1,\gamma_2\geq 0 \end{cases}$$



分情况讨论，可解得：$x_1=1, x_2=1, \gamma_1=2/3, \gamma_2=2/3$。

故$\bar{\mathbf{x}}=(1,1)^{\mathrm{T}}$为 **KKT 点**。

✧最优点 vs. KKT 点

[1] 上述问题是凸规划，因此 **KKT** 点同时也是全局最优点。

[2] 对于非凸规划，**KKT** 点不一定是全局最优，甚至不一定是局部最优；局部最优点在满足正则条件时，才是 **KKT** 点，否则也不一定是 **KKT** 点。

[3] **KKT** 点应该是可行点，因此在验证和求解 **KKT** 点时，必须考虑满足约束条件。

三、最优性条件的应用

（一）投资组合问题

三种股票，各有两种风险：独立风险（标准差）和交叉风险（协方差）。要求以总方差表示的风险最小，且期望回报不低于 **18%**。求最佳投资组合比例。

股票	期望回报率	独立风险（标准差）	交叉风险（协方差）
1	21%	25%	1 and 2： 0.040
2	30%	45%	1 and 3：–0.005
3	8%	5%	2 and 3：–0.010

解，设投资比例为x_i，回报率为r_i，则投资组合的方差为

$$
\begin{aligned}
&\operatorname{var}(x_1 r_1 + x_2 r_2 + x_3 r_3) \\
&\quad = (0.25x_1)^2 + (0.45x_2)^2 + (0.05x_3)^2 + 2(0.04)x_1x_2 \\
&\quad +2(-0.005)x_1x_3 + 2(-0.01)x_2x_3
\end{aligned}
$$

整理上式并考虑最低期望回报率约束，决策模型为

$$
\begin{aligned}
\min \text{Risk} = {} & 0.0625x_1^2 + 0.2025x_2^2 + 0.0025x_3^2 \\
& +0.08x_1x_2 - 0.01x_1x_3 - 0.02x_2x_3
\end{aligned}
$$

$$
\text{s.t.}\begin{cases}
21\%x_1 + 30\%x_2 + 8\%x_3 \geq 18\% \\
x_1 + x_2 + x_3 = 100\% \\
x_1, x_2, x_3 \geq 0
\end{cases}
$$

显然，目标函数为二次凸函数；而约束都是线性函数，因此是一个凸规划，其 **KKT** 点就是最优点。

目标函数梯度：

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 0.125x_1 + 0.08x_2 - 0.01x_3 \\ 0.08x_1 + 0.405x_2 - 0.02x_3 \\ -0.01x_1 - 0.02x_2 + 0.005x_3 \end{bmatrix}$$

约束函数梯度 [3]：

$$\nabla g(\mathbf{x}) = \begin{bmatrix} 0.21 \\ 0.30 \\ 0.08 \end{bmatrix}，\nabla h(\mathbf{x}) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

于是，问题的 KKT 条件方程为

---

[3] 为简化分析，先不考虑非负约束，待求解出来后再验证之。

$$
\begin{cases}
0.125x_1 + 0.08x_2 - 0.01x_3 - 0.21\gamma - \lambda = 0 \\
0.08x_1 + 0.405x_2 - 0.02x_3 - 0.30\gamma - \lambda = 0 \\
-0.01x_1 - 0.02x_2 + 0.005x_3 - 0.08\gamma - \lambda = 0 \\
\gamma(0.21x_1 + 0.30x_2 + 0.08x_3 - 0.18) = 0 \\
\gamma \geq 0 \\
0.21x_1 + 0.30x_2 + 0.08x_3 - 0.18 \geq 0 \\
x_1 + x_2 + x_3 - 100\% = 0
\end{cases}
$$

情况 1： if $\gamma = 0$，**KKT** 条件方程化为

$$
\begin{cases}
0.125x_1 + 0.08x_2 - 0.01x_3 - \lambda = 0 \\
0.08x_1 + 0.405x_2 - 0.02x_3 - \lambda = 0 \\
-0.01x_1 - 0.02x_2 + 0.005x_3 - \lambda = 0 \rightarrow \text{无可行解} \\
0.21x_1 + 0.30x_2 + 0.08x_3 - 0.18 \geq 0 \\
x_1 + x_2 + x_3 - 100\% = 0
\end{cases}
$$

情况 2： if $\gamma > 0$，**KKT** 条件方程化为

$$\begin{cases} 0.125x_1 + 0.08x_2 - 0.01x_3 - 0.21\gamma - \lambda = 0 \\ 0.08x_1 + 0.405x_2 - 0.02x_3 - 0.30\gamma - \lambda = 0 \\ -0.01x_1 - 0.02x_2 + 0.005x_3 - 0.08\gamma - \lambda = 0 \\ 0.21x_1 + 0.30x_2 + 0.08x_3 - 0.18 = 0 \\ x_1 + x_2 + x_3 - 100\% = 0 \end{cases}$$

得到 KKT 点即最优投资组合：

$$x_1 = 40.2\%, x_2 = 21.7\%, x_3 = 38.1\%, \gamma = 0.54 > 0$$

最小风险（总方差）：

$$\mathrm{Risk} = (x_1 r_1 + x_2 r_2 + x_3 r_3) = 15.4\%$$

（二）推导需求定律

设消费者消费两种产品，需求量分别为x, y，价格分别为p_x, p_y。已知效用函数为$Q(x,y) = Ax^{\alpha}y^{1-\alpha}$，$\alpha \in (0,1)$，消费者总支出预算为$C$，求价格与需求量的关系。

解，消费者的决策应是以给定预算实现效用最大，即

$$\begin{aligned} \max \quad & Q(x,y) = Ax^{\alpha}y^{1-\alpha}, \\ \text{s.t.} \quad & p_x x + p_y y = C \ \text{ and } \ x, y \geq 0 \end{aligned}$$

为分析方便，等价变换为最大化目标函数的自然对数：

$$\max \ln Q(x,y) = \ln A + \alpha \ln x + (1-\alpha) \ln y$$

化为最小化问题：

$$\min -\ln Q(x,y) = -\ln A - \alpha \ln x - (1-\alpha)\ln y$$

最小化问题的目标函数显然为凸函数，则配合线性等式约束就形成凸规划。运用 **KKT** 条件，设拉氏乘子为v，有

$$\begin{cases} \begin{bmatrix} -\alpha/x \\ -(1-\alpha)/y \end{bmatrix} - v\begin{bmatrix} p_x \\ p_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ p_x x + p_y y = C \\ x, y \geq 0 \end{cases}$$

求解之，得到价格与需求量的关系为

$$p_y = \frac{(1-\alpha)C}{y}, p_x = \frac{\alpha C}{x}, v = -\frac{Ax^{\alpha}y^{1-\alpha}}{C}$$

可见价格与需求量成反比，符合需求定律。

（三）生产要素最佳组合

设生产商用n种生产要素生产产品，各生产要素单价为p_i，总产出函数为$Q(x_1,\dots,x_n,)$；生产要素的边际规模递减，可假设总产出函数为凹函数。企业的资金约束为B，求总产量最大时生产要素的最佳组合条件。

解，产量最大的决策模型为 [4]

$$\begin{aligned}&\max Q(x_1,\dots,x_n)\\&\text{s.t. } p_1x_1+\cdots+p_nx_n\le B\end{aligned}$$

由于总产出$Q(\cdot)$为凹函数，因此原问题等价于下述凸规划：

[4] 为简化分析，这里先不考虑非负约束。

$$\min -Q(x_1, \dots, x_n)$$
$$\text{s.t.} -p_1 x_1 - \cdots - p_n x_n + B \geq 0$$

设要素x_i的边际产出为：

$$MP_i = \partial Q(x_1, x_2, \dots, x_n,)/\partial x_i$$

则目标函数的梯度为：

$$\nabla[-Q(x_1, \dots, x_n)] = \begin{bmatrix} -MP_1 \\ \vdots \\ -MP_n \end{bmatrix}$$

而约束函数的梯度为：

$$\nabla[-\sum_{i=1}^{n} p_i x_i + B] = \begin{bmatrix} -p_1 \\ \vdots \\ -p_n \end{bmatrix}$$

构造 **KKT** 条件，有

$$\begin{bmatrix} -MP_1 \\ \vdots \\ -MP_n \end{bmatrix} - \gamma \begin{bmatrix} -p_1 \\ \vdots \\ -p_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

即

$$\begin{gathered} -MP_1 + \gamma p_1 = 0 \rightarrow MP_1/p_1 = \gamma \\ \vdots \\ -MP_n + \gamma p_n = 0 \rightarrow MP_n/p_n = \gamma \end{gathered}$$

整理后，得到生产要素最佳组合（必要）条件为

$$\frac{MP_1}{p_1} = \cdots = \frac{MP_i}{p_i} = \cdots = \frac{MP_n}{p_n} = \gamma$$

更多关于数学规划方法在经济学中的应用，参见：

高山晟.《经济学中的分析方法》