第七章　最优性条件

$$\text{非线性规划问题}\begin{cases}\text{无约束极值问题}\\ \text{有约束极值问题}\end{cases}$$

有约束极值问题：

$$\begin{aligned}\min\quad & f(x)\\ s.t.\quad & g_i(x)\ge 0,\quad i=1,2,\cdots,l\\ & h_j(x)=0,\quad j=1,2,\cdots,m.\end{aligned}$$

问题的提出

例:某公司经营两种设备。第一种设备每件售价为**30**元,第二种设备每件售价为**450**元。且知,售出第一、二种设备分别需时为每件约**0.5**小时和(**2+0.25x_2**)小时,其中x_2为第二种设备售出数量。公司的总营业时间为**800**小时。

求:公司为获取最大营业额(销售额)的最优营业计划。

[解]设公司应经营销售第一、二种设备数额分别为x_1件和x_2件,则有

$$
\begin{aligned}
\max:\ & f(X)=30x_1+450x_2 \\
s.t.\ & 0.5x_1+2x_2+0.25x_2^2\le 800 \\
& x_1\ge 0,\ x_2\ge 0
\end{aligned}
$$

- 定义１： 设$f(x)$为目标函数，S为可行域，$x^0 \in S$，若对$\forall x \in S$, **有**$f(x) \geqslant f(x^0)$, **则**x^0**称为极小化问题**$\min f(x)$， $x \in S$**的（全局）最优解**.

- **定义２：** 设$f(x)$为目标函数，S为可行域，若存在x^0的ε邻域

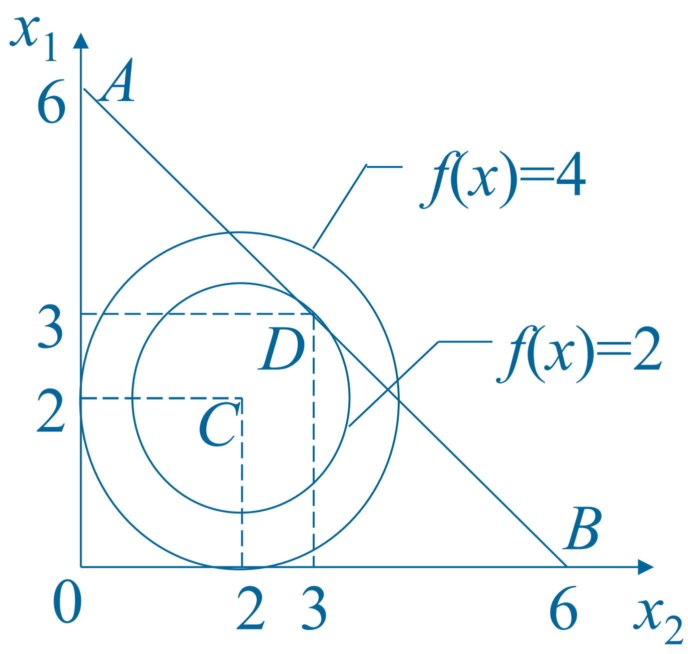
- $$N_\varepsilon(x^0) = \{x \mid \| x - x^0 \| < \varepsilon, \varepsilon > 0\}$$

- 使得对$\forall x \in S \cap N_\varepsilon(x^0)$，**有**$f(x) \geqslant f(x^0)$, **则**$x^0$**称为极小化问题**$\min f(x)$， $x \in S$**的局部最优解**.

[例]求解下述非线性规划

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

$$h(x)=x_1+x_2-6=0$$



显然，与直线AB相切的点必为最优解。

图中的D点即为最优点，此时目标函数值为：

$f(x^*)=2$，$x_1^*=x_2^*=3$

[例]非线性规划为

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

$$h(x)=x_1+x_2-6\le 0$$

最优解为$x_1^*=x_2^*=2$，$f(x^*)=0$，该点落在可行域内部，其边界约束失去作用。

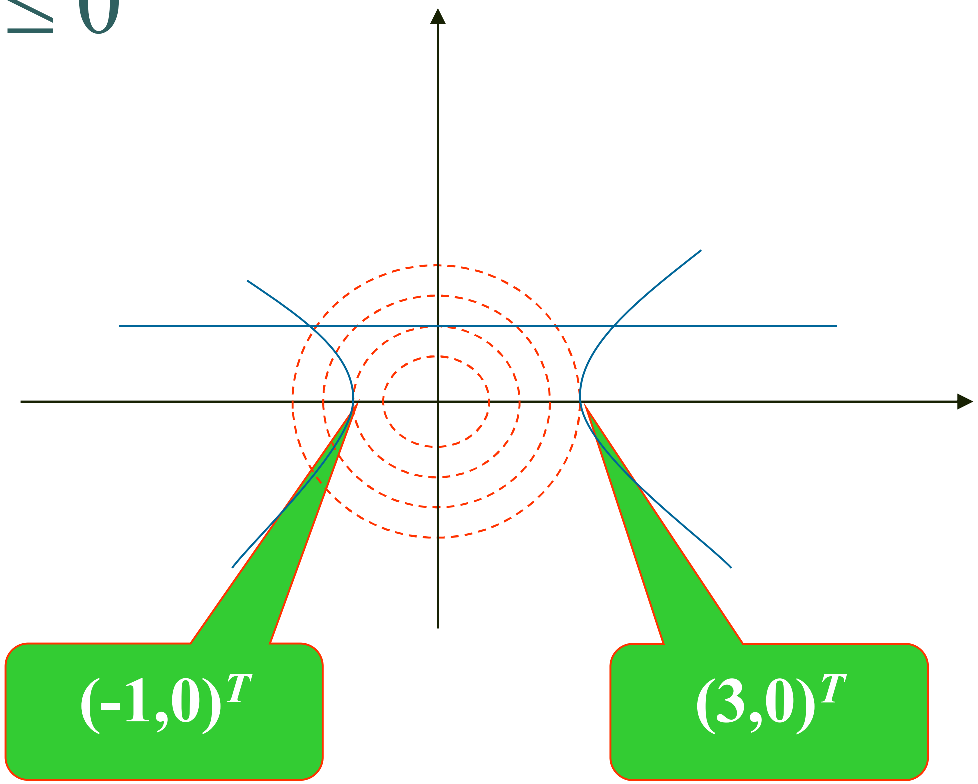
结论:非线性规划的最优解（如果存在）可在其可行域上任一点达到。

求下列约束问题的解：

$$\min\ x_1^2 + x_2^2$$

$$s.t.\quad (x_1 - 1)^2 - x_2^2 - 4 = 0$$

$$x_2 - 1 \le 0$$



无约束问题的极值条件

$$\begin{aligned} &\min \quad f(x) \\ &s.t. \quad x \in E^n \end{aligned}$$

定义： 对 $\min\limits_{x\in E^n} f(x)$，设 $\overline{x} \in E^n$ 是任给一点，$d \neq 0$，若存在 $\delta > 0$，使得对任意的 $\lambda \in (0,\delta)$，有 $f(\overline{x}+\lambda d) < f(\overline{x})$，则称 d 为 $f(x)$ 在点 $\overline{x}$ 处的下降方向。

引理：设函数$f(x)$在点$\bar{x}$可微，若存在$d \neq 0$使$\nabla f(\bar{x})^T d < 0$，则存在$\delta > 0$，使对$\forall \lambda \in (0, \delta)$，有$f(\bar{x} + \lambda d) < f(\bar{x})$。

证明：由泰勒展开公式

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^T d + \| \lambda d \| \alpha(\bar{x}, \lambda d)$$

$$= f(\bar{x}) + \lambda [\nabla f(\bar{x})^T d + \frac{|\lambda|}{\lambda} \| d \| \alpha(\bar{x}, \lambda d)]$$

其中当 $\lambda \to 0$时， $\alpha(\bar{x}, \lambda d) \to 0$

$\because \nabla f(\bar{x})^T d < 0$，当$|\lambda|$充分小时，有

$$\lambda [\nabla f(\bar{x})^T d + \frac{|\lambda|}{\lambda} \| d \| \alpha(\bar{x}, \lambda d)] < 0$$

$\therefore$ 存在 $\delta > 0$，使得当 $\lambda \in (0, \delta)$时，有

$$\lambda [\nabla f(\bar{x})^T d + \frac{|\lambda|}{\lambda} \| d \| \alpha(\bar{x}, \lambda d)] < 0$$

$\therefore f(\bar{x} + \lambda d) < f(\bar{x}).$

必要条件

定理1:(一阶必要条件)设函数$f(x)$在点$\bar{x}$处可微，若$\bar{x}$是局部极小点，则$\nabla f(\bar{x})=0$.

证明：设$\nabla f(\bar{x})\neq 0$，取$d=-\nabla f(\bar{x})\neq 0$,

则有$\nabla f(\bar{x})^T d=\nabla f(\bar{x})^T(-\nabla f(\bar{x}))$

$$=-\|\nabla f(\bar{x})\|^2<0$$

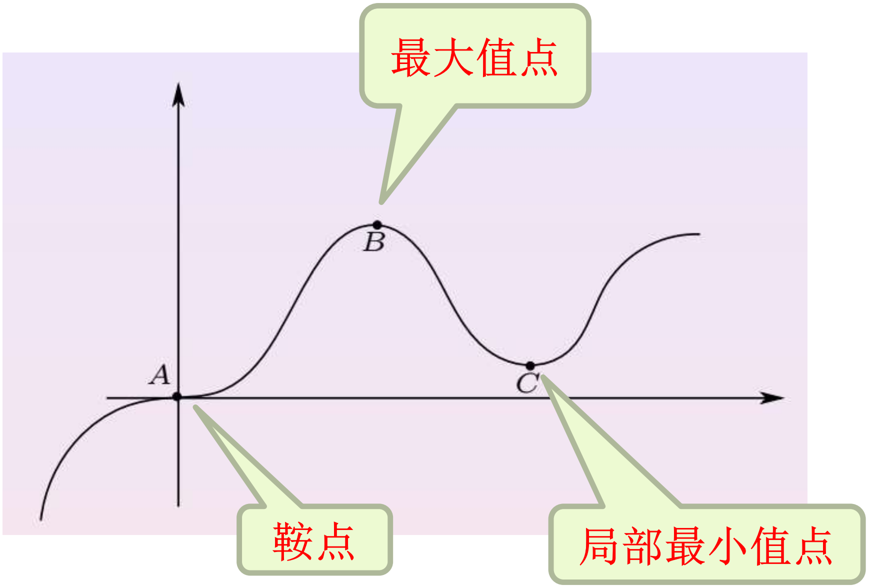
由引理，存在$\delta>0$,使当$\lambda\in(0,\delta)$时，有

$$f(\bar{x}+\lambda d)<f(\bar{x})$$

与$\bar{x}$是局部极小点矛盾。

定义: 若$f(x)$在点x^*可微，并且$\nabla f(x^*)=0$，则x^*称为$f(x)$的一个驻点(或者平稳点，stationary point)。既不是极小点，也不是极大点的驻点称为鞍点(saddle point)。

例: 对于二次函数$f(x)=x_1x_2,\ x^*=(0,0)^T$是它的驻点，但是该点既不是极小点，也不是极大点，所以x^*是$f(x)$的鞍点。



定理2:$\left(\text{二阶必要条件}\right)$设$f(x)$在$\bar{x}$处二阶可微，若$\bar{x}$是局部极小点，则$\nabla f(\bar{x})=0$，且$Hessian$矩阵$\nabla^2 f(\bar{x})$是半正定的。

证明：由定理1，$\nabla f(\bar{x})=0.$

设d是任意一个n维非零向量，

$\because f(x)$在$\bar{x}$处二阶可微,且$\nabla f(\bar{x})=0,$

$$\therefore f(\bar{x}+\lambda d)=f(\bar{x})+\lambda\nabla f(\bar{x})^T d+\frac{1}{2}\lambda^2 d^T\nabla^2 f(\bar{x})d+\lambda^2\|d\|^2\alpha(\bar{x},\lambda d)$$

$$\Rightarrow\quad \frac{f(\bar{x}+\lambda d)-f(\bar{x})}{\lambda^2}=\frac{1}{2}d^T\nabla^2 f(\bar{x})d+\|d\|^2\alpha(\bar{x},\lambda d)$$

其中当$\lambda\to 0$时,$\alpha(\bar{x},\lambda d)\to 0$

$\because \bar{x}$是局部极小点,当$|\lambda|$充分小时,必有$f(\bar{x}+\lambda d)\ge f(\bar{x})$

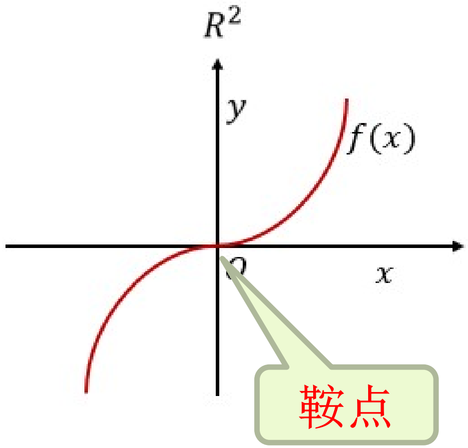
$\therefore$当$\lambda\to 0$时,有$d^T\nabla^2 f(\bar{x})d\ge 0$

即$\nabla^2 f(\bar{x})$为半正定的.

二阶必要条件的非充分性

函数 $f(x) = x^3$ 在零点满足 $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$, $\nabla^2 f(\boldsymbol{x}^*)$半正定

但零点不是其最优值点。



定理3：设函数$f(x)$在点$\bar{x}$处二次可微，若梯度$\nabla f(\bar{x})=0$,且$Hessian$矩阵$\nabla^2 f(\bar{x})$正定，则$\bar{x}$是严格局部极小点。

二阶充分条件

证明：对任意$x\neq\bar{x}$,

$\because f(x)$在$\bar{x}$处二阶可微且$\nabla f(\bar{x})=0$,

$$\therefore f(x)=f(\bar{x})+\frac{1}{2}(x-\bar{x})^T\nabla^2 f(\bar{x})(x-\bar{x})+o(\|x-\bar{x}\|^2)$$

$\because \nabla^2 f(\bar{x})$是正定的,$\therefore (x-\bar{x})^T\nabla^2 f(\bar{x})(x-\bar{x})>0$

因此，存在$\varepsilon>0$，当$x\in N_\varepsilon(\bar{x})$且$x\neq\bar{x}$时，有

$$f(x)>f(\bar{x})$$

$\therefore \bar{x}$是严格局部极小点.

二阶充分条件的非必要性

上述结论给出的二阶充分性条件不是必要的.

如零点是单元函数 $f(x) = x^4$

的严格最优解，但上述二阶充分性条件在该点并不成立.

无约束优化问题不存在充分必要的最优性条件！

推论：对于正定二次函数$f(x)=\frac{1}{2}x^T Ax+b^T x+c$

(A对称正定)，有唯一极小点$x^*=-A^{-1}b$.

证明：$\because f(x)=\frac{1}{2}x^T Ax+b^T x+c$

$\therefore \quad \nabla f(x)=Ax+b, \quad \nabla^2 f(x)=A$

令$\nabla f(x)=0, \quad \because A$正定，

$\therefore Ax+b=0$有唯一解$x^*=-A^{-1}b$.

显然$\nabla f(x^*)=0$且$\nabla^2 f(x^*)=\mathrm{A}$正定,

$\therefore x^*$为唯一极小点.

定理4：设$f(x)$是定义在E^n上的可微凸函数，$\bar{x} \in E^n$，则$\bar{x}$为整体极小点的充要条件是$\nabla f(\bar{x}) = 0.$

证明：只证充分性。

设$\nabla f(\bar{x}) = 0.$

$\because f(x)$是E^n上的可微凸函数，

$\therefore$ 对任意的$x \in E^n$，有

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) = f(\bar{x})$$

$\therefore \bar{x}$为整体极小点。

例：$\min\ f(x) = 5x_1^2 - 6x_1x_2 + 5x_2^2$

解：$\dfrac{\partial f}{\partial x_1} = 10x_1 - 6x_2, \dfrac{\partial f}{\partial x_2} = -6x_1 + 10x_2$

令$\nabla f(x) = 0$，得$x^* = \left(0,\ 0\right)^T$，

$\because \nabla^2 f(x) = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$正定，故$f(x)$为凸函数

$\therefore x^*$为整体极小点。

例：求$f(x)=\frac{1}{3}x_1^3+\frac{1}{3}x_2^3-x_2^2-x_1$的极小点。

解：$\frac{\partial f}{\partial x_1}=x_1^2-1,\quad \frac{\partial f}{\partial x_2}=x_2^2-2x_2$

令$\nabla f(x)=0$,得$\begin{cases} x_1^2-1=0 \\ x_2^2-2x_2=0 \end{cases}$，得驻点

$$x^{(1)}=\begin{pmatrix}1\\0\end{pmatrix}, x^{(2)}=\begin{pmatrix}1\\2\end{pmatrix}, x^{(3)}=\begin{pmatrix}-1\\0\end{pmatrix}, x^{(4)}=\begin{pmatrix}-1\\2\end{pmatrix}$$

$f(x)$的$Hessian$矩阵$\nabla^2 f(x)=\begin{pmatrix}2x_1 & 0\\ 0 & 2x_2-2\end{pmatrix}$

约束优化最优性条件

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}) \\
\text{s.t.} \quad & \boldsymbol{x} \in \Omega
\end{aligned}
$$

寻求和验证最优值点是一件非常棘手的问题。

对目标函数值两两比较？ 不可能！

如何验证最优值点？最优解的判定准则是什么？

约束优化问题的最优性条件

约束极值问题的最优性条件

$$\begin{cases} \min \quad f(x) \\ s.t. \quad g_i(x) \geq 0, \quad i = 1,2,\cdots,m \text{----不等式约束} \\ \qquad\quad h_j(x) = 0, \quad j = 1,2,\cdots,l. \text{----等式约束} \\ \qquad\quad x \in E^n \end{cases}$$

$$S = \left\{ x \left| \begin{matrix} g_i(x) \geq 0, & i = 1,2,\cdots,m \\ h_j(x) = 0, & j = 1,2,\cdots,l \end{matrix} \right. \right\}$$

------------可行集或可行域

定义： 对 $\min_{x\in E^n} f(x)$，设 $\bar{x}\in E^n$ 是任给一点，$d\neq 0$，若存在 $\delta>0$，使得对任意的 $\lambda\in(0,\delta)$，有 $f(\bar{x}+\lambda d)<f(\bar{x})$，则称 d 为 $f(x)$ 在点 $\bar{x}$ 处的下降方向(descent direction)。

$$F_0=\left\{d \,\middle|\, \nabla f(\bar{x})^T d<0\right\}$$

称为点 $\bar{x}$ 处的下降方向集。

定义： 设集合$S \subset E^n, \bar{x} \in clS, d$为非零向量，若存在数$\delta > 0$，使得对任意$\lambda \in (0, \delta)$，都有

$$\bar{x} + \lambda d \in S$$

则称d为集合S在$\bar{x}$的可行方向(feasible direction)。

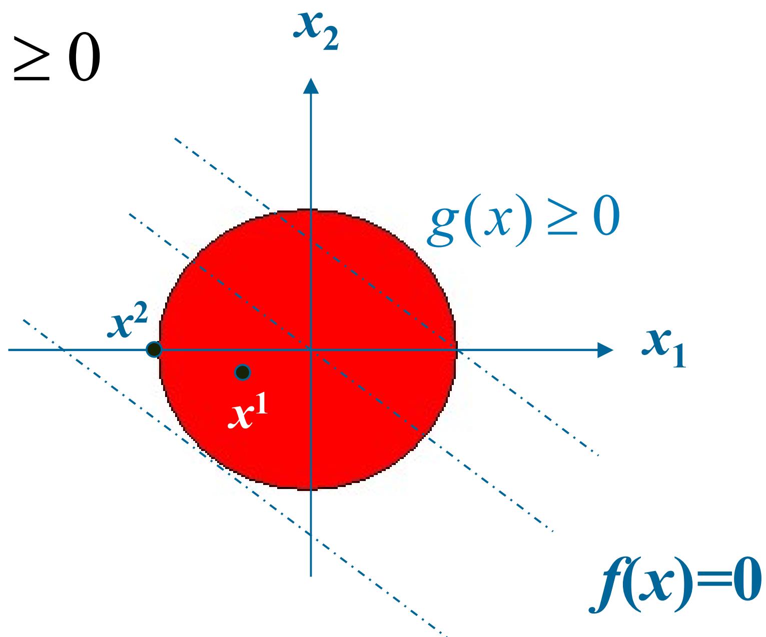
$$D = \left\{ d \middle| \, d \neq 0, \bar{x} \in clS, \exists \delta > 0, \forall \lambda \in (0, \delta), \text{有} \bar{x} + \lambda d \in S \right\}$$

$\bar{x}$处的可行方向锥。

例：考虑如下约束优化问题：

$$\min\ f(x)=x_1+x_2$$

$$s.t.\quad g(x)=1-x_1^2-x_2^2\geq 0$$



对于任意内点x^1，可行方向锥$D=R^2\setminus\{0\}$.

对于边界点$x^2=(-1,0)^T$,可行方向锥

$$D=\{d\in R^2\mid d_1>0\}.$$

定理1（几何最优性条件）：考虑问题

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & x \in S \end{aligned}$$

$$F_0 = \{d \mid \nabla f(\bar{x})^T d < 0\}$$

$$D = \{d \mid d \neq 0, \bar{x} \in clS, \exists \delta > 0, \forall \lambda \in (0,\delta), \text{有} \bar{x} + \lambda d \in S\}$$

设S是E^n的非空集合，$\bar{x} \in S$，$f(x)$在$\bar{x}$处可微，若$\bar{x}$是局部最优解，则$F_0 \bigcap D = \varnothing$。

证明：设存在的$d \in F_0 \bigcap D$，则$d \in F_0, d \in D$.

$\because d \in F_0, \therefore \exists \delta_1 > 0$，对$\forall \lambda \in (0,\delta_1)$，有$f(\bar{x}+\lambda d) < f(\bar{x})$;

$\because d \in D, \therefore \exists \delta_2 > 0$，对$\forall \lambda \in (0,\delta_2)$，有$\bar{x}+\lambda d \in S$.

取$\delta = \min(\delta_1, \delta_2)$，则当$\lambda \in (0,\delta)$，有

$\bar{x}+\lambda d \in S$且$f(\bar{x}+\lambda d) < f(\bar{x})$，

与$\bar{x}$为局部最优解矛盾。

不等式约束问题的一阶最优性条件

$$(1)\quad \begin{cases} \min \quad f(x) \\ s.t. \quad g_i(x) \geq 0, i = 1, \cdots, m. \end{cases}$$

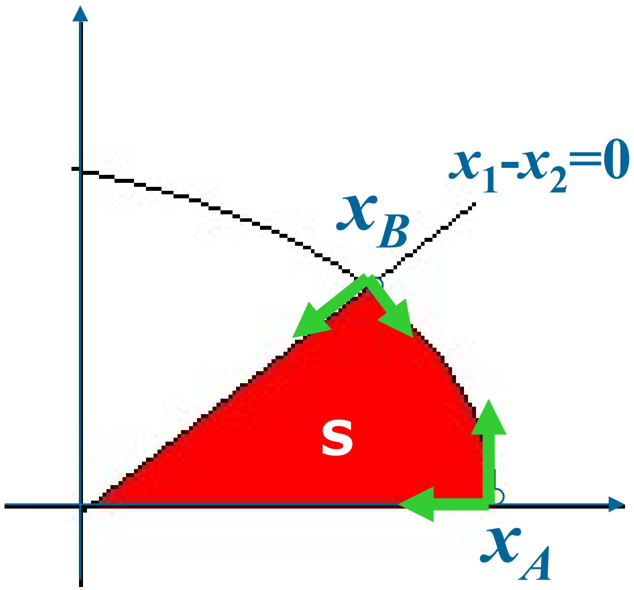
可行域 $S = \{x | g_i(x) \geq 0, i = 1, \cdots, m\}$

定义: 若问题(1)的一个可行点$\bar{x}$(即$\bar{x} \in S$)使某个不等式约束$g_i(x) \geq 0$变成等式，即$g_i(\bar{x}) = 0$，则该不等式约束称为关于可行点$\bar{x}$的起作用约束（或等式约束）；否则，若$\bar{x}$使得某个$g_i(\bar{x}) > 0$，则该不等式约束称为关于可行点$\bar{x}$的不起作用约束（或松约束）。

记 $\quad I = \{i | g_i(\bar{x}) = 0, \bar{x} \in S\}$

例：约束

$$
\begin{cases}
g_1(x) = -x_1^2 - x_2^2 + 1 \ge 0 \\
g_2(x) = x_1 - x_2 \ge 0 \\
g_3(x) = x_1 \ge 0 \\
g_4(x) = \quad x_2 \ge 0
\end{cases}
$$



$$x_A = (1,0)^T,\ x_B = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)^T \qquad I_A = \{1,4\}, \quad I_B = \{1,2\}$$

$$G_0 = \left\{ d \,\middle|\, \nabla g_i(\bar{x})^T d > 0, i \in I \right\}$$

称G_0为S在点$\bar{x}$处的局部约束方向锥（ 或内方向锥）。

$$G_0 = \left\{ d \middle| \nabla g_i(\overline{x})^T d > 0, i \in I \right\}$$

称G_0为S在点$\overline{x}$处的局部约束方向锥（ 或内方向锥）。

$$D = \left\{ d \middle| d \neq 0, \overline{x} \in clS, \exists \delta > 0, \forall \lambda \in (0, \delta), 有\overline{x} + \lambda d \in S \right\}$$

约束

$g(x) = x^2 \geq 0$

取 $\overline{x} = 0$

则 $G_0 = \varnothing$，但 $D = R \setminus \{0\}$

定理2：设$\bar{x} \in S$，$f(x)$和$g_i(x)(i \in I)$在$\bar{x}$处可微，$g_i(x)(i \notin I)$在$\bar{x}$处连续，如果$\bar{x}$是问题(1)的局部最优解，则$F_0 \cap G_0 = \varnothing$。

证明：由定理1，在$\bar{x}$处，有$F_0 \cap D = \varnothing$.

设$d \in G_0$，则$\nabla g_i(\bar{x})^T d > 0$，$i \in I$；令$\tilde{g}_i(x) = -g_i(x)$，$i \in I$.

则$\nabla \tilde{g}_i(\bar{x})^T d < 0$，$i \in I$;

由引理，存在$\delta_1 > 0$,当$\lambda \in (0, \delta_1)$时，有

$$\tilde{g}_i(\bar{x} + \lambda d) < \tilde{g}_i(\bar{x}),\ i \in I$$

即 $g_i(\bar{x} + \lambda d) > g_i(\bar{x}) = 0$，$i \in I$.

当$i \notin I$时，$g_i(\bar{x}) > 0$,$\because g_i(\bar{x})(i \notin I)$在$\bar{x}$连续,

$\therefore$存在$\delta_2 > 0$,当$\lambda \in (0, \delta_2)$时，有$g_i(\bar{x} + \lambda d) > 0, i \notin I$.

令$\delta = \min(\delta_1, \delta_2)$，则当$\lambda \in (0, \delta)$时，有$g_i(\bar{x} + \lambda d) > 0, \forall i$,

$\Rightarrow \bar{x} + \lambda d \in S \quad \Rightarrow d \in D$

$\Rightarrow \quad G_0 \subseteq D \quad \Rightarrow F_0 \cap G_0 = \varnothing$

性质 若d为约束问题在x点的可行方向，则 $\nabla g_i(x)^T d \geq 0, \forall i \in I(x)$

证明要点：用反证法。若存在 $\nabla g_{i_0}(x)^T d < 0 \quad i_0 \in I(x)$

则对充分小的 $\lambda > 0$

$$g_{i_0}(x+\lambda d) = g_{i_0}(x) + \lambda \nabla g_{i_0}(x)^T d + o(\lambda) < 0$$

特别提醒：上述条件不是充分的，除非约束函数是线性的。

定理3($Fritz\ John$条件) 设$\bar{x} \in S, I = \{i \mid g_i(\bar{x}) = 0\}$, $f(x), g_i(x)(i \in I)$在$\bar{x}$处可微， $g_i(x)(i \notin I)$在$\bar{x}$处连续，若$\bar{x}$是问题(1)的局部最优解，则存在 不全为零的数$w_0, w_i(i \in I)$,使得

$$\begin{cases} w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0 \\ w_0,\ w_i \geq 0, \quad i \in I. \end{cases}$$

$\bar{x}$ 称为**Fritz John** 点**(**即满足**Fritz John**条件的点**).**

证明：由定理2，在点$\bar{x}$, $F_0 \cap G_0 = \varnothing$,即不等式

$$\begin{cases} \nabla f(\bar{x})^T d < 0 \\ -\nabla g_i(\bar{x})^T d < 0, \quad i \in I \end{cases}$$ 无解。

$Ax < 0$有解的充要条件是不存在非零向量$y \geq 0$，使得$A^T y = 0$.

设$I = \{i_1, i_2, \cdots, i_s\}$, 令$A = \begin{pmatrix} \nabla f(\bar{x})^T \\ -\nabla g_{i_1}(\bar{x})^T \\ \vdots \\ -\nabla g_{i_s}(\bar{x})^T \end{pmatrix}$

有$Ad < 0$无解,由$Gordan$定理,存在$w = \left(w_0, w_{i_1}, \cdots, w_{i_s}\right)^T \geq 0, w \neq 0$,使得

$$A^T w = 0, 即\left(\nabla f(\bar{x}), -\nabla g_{i_1}(\bar{x}), \cdots, -\nabla g_{i_s}(\bar{x})\right)\begin{pmatrix} w_0 \\ w_{i_1} \\ \vdots \\ w_{i_s} \end{pmatrix} = 0$$

$$\Rightarrow w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

定理3'($Fritz\ John$条件) 设$\bar{x}\in S, f(x), g_i(x)$在$\bar{x}$处可微，若$\bar{x}$是问题(1)的局部最优解，则存在不全为零的数$w_0, w_1, \cdots, w_m$,使得

$$\begin{cases} w_0 \nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0 \\ w_i g_i(\bar{x}) = 0, \quad i = 1, 2, \cdots, m \\ w_0,\ w_i \geq 0, \quad i = 1, 2, \cdots, m. \end{cases}$$

互补松弛条件

问题

$$\begin{cases} \min \quad f(x) \\ s.t. \quad g_i(x) \le 0, i = 1, 2, \cdots, m \end{cases}$$

的*Fritz John*条件为

$$\begin{cases} w_0 \nabla f(\bar{x}) + \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0 \\ w_0, \ w_i \ge 0, \quad i \in I. \end{cases}$$

例：设非线性规划问题

$$\begin{cases} \min\ f(x)=(x_1-3)^2+(x_2-2)^2 \\ s.t.\quad g_1(x)=-x_1^2-x_2^2+5\ge 0 \\ \qquad\quad g_2(x)=-x_1-2x_2+4\ge 0 \\ \qquad\quad g_3(x)=x_1\ge 0 \\ \qquad\quad g_4(x)=x_2\ge 0 \end{cases}$$

判别点$x^{(1)}=(2,1)^T$和$x^{(2)}=(0,0)^T$是否是*Fritz John*点？

解：$\nabla f(x)=\big(2(x_1-3),2(x_2-2)\big)^T$

$$\nabla g_1(x)=(-2x_1,-2x_2)^T,\nabla g_2(x)=(-1,-2)^T$$

$$\nabla g_3(x)=(1,0)^T,\quad \nabla g_4(x)=(0,1)^T$$

例：设非线性规划问题

$$\begin{cases} \min\ f(x)=(x_1-3)^2+(x_2-2)^2 \\ s.t.\quad g_1(x)=-x_1^2-x_2^2+5\geq 0 \\ \qquad\ \ g_2(x)=-x_1-2x_2+4\geq 0 \\ \qquad\ \ g_3(x)=x_1\geq 0 \\ \qquad\ \ g_4(x)=x_2\geq 0 \end{cases}$$

判别点$x^{(1)}=(2,1)^T$和$x^{(2)}=(0,0)^T$是否是*Fritz John*点？

解： $\nabla f(x)=\left(2(x_1-3),2(x_2-2)\right)^T$

$$\nabla g_1(x)=(-2x_1,-2x_2)^T,\nabla g_2(x)=(-1,-2)^T$$

$$\nabla g_3(x)=(1,0)^T,\quad \nabla g_4(x)=(0,1)^T$$

$$\min\ f(x)=\left(x_1-1\right)^2+\left(x_2-1\right)^2$$

$$\begin{aligned} s.t.\quad & g_1(x)=\left(1-x_1-x_2\right)^3\geq 0 \\ & g_2(x)=x_1\geq 0 \\ & g_3(x)=x_2\geq 0 \end{aligned}$$

最优解为

$$x^*=\left(\frac{1}{2},\frac{1}{2}\right)^T.$$

$\because$ 直线$x_1+x_2=1$上所有可行点$\overline{x}$使$\nabla g_1(\overline{x})=0$,

$\therefore$ 取$w_0=0, w_1=a>0$，$w_2=w_3=0$，总有

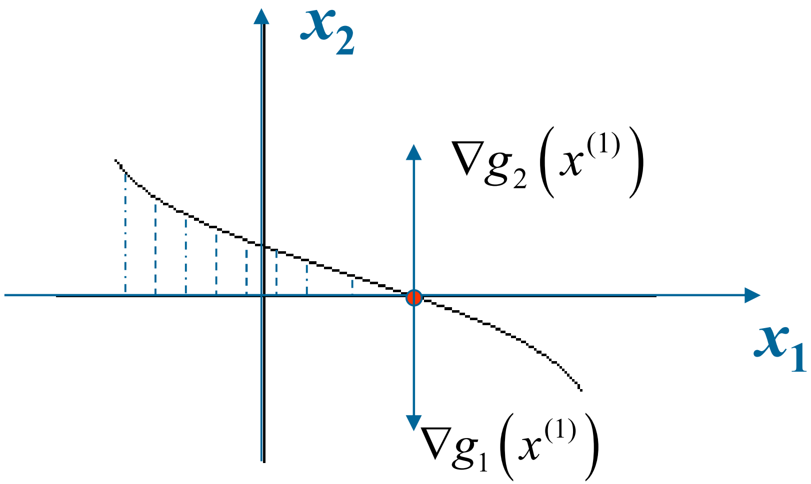
$$w_0\nabla f(\overline{x})-w_1\nabla g_1(\overline{x})-w_2\nabla g_2(\overline{x})-w_3\nabla g_3(\overline{x})=0$$

说明在直线$x_1+x_2=1$上每个可行点$\overline{x}$都是*Fritz John*点，但除x^*外，都不是最优解。

例2. 设非线性规划问题：

$$\begin{cases} \min -x_1 \\ s.t. \quad g_1(x) = -x_2 + (1-x_1)^3 \geq 0 \\ \qquad\quad g_2(x) = x_2 \geq 0 \end{cases}$$

判别点$x^{(1)} = (1, 0)^T$是否是$Fritz\ John$点？



解： $\because \nabla f(x)=(-1,0)^T$

$\nabla g_1(x)=(-3(1-x_1)^2,-1)^T, \nabla g_2(x)=(0,1)^T$

在点$x^{(1)}=(1,0)^T$处, $I=\{1,2\}$

$\nabla f(x^{(1)})=(-1,0)^T, \nabla g_1(x^{(1)})=(0,-1)^T$,

$\nabla g_2(x)=(0,1)^T$，设有

$$w_0\begin{pmatrix}-1\\0\end{pmatrix}-w_1\begin{pmatrix}0\\-1\end{pmatrix}-w_2\begin{pmatrix}0\\1\end{pmatrix}=0$$

$\Rightarrow w_0=0$, 取$w_1=w_2>0$

$\Rightarrow x^{(1)}=(1,0)^T$是$Fritz\ John$点。

Karush-Kuhn-Tucker(KKT)条件

定理2. 考虑问题$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \ge 0, i = 1, \cdots, m \end{cases}$

设$\bar{x} \in S, f, g_i(i \in I)$在$\bar{x}$处可微，$g_i(i \notin I)$在$\bar{x}$连续，$\{\nabla g_i(\bar{x}) \mid i \in I\}$线性无关，若$\bar{x}$是局部最优解，则存在非负数$w_i, i \in I$，使得

$$\nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

证明：由定理1，存在不全为零的非负数

$w_0, w_i', i \in I$，使得

$$w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i' \nabla g_i(\bar{x}) = 0.$$

显然$w_0 \neq 0$，否则$\nabla g_i(\bar{x})(i \in I)$线性相关，矛盾。

于是，令$w_i = \dfrac{w_i'}{w_0} \geq 0\ (i \in I)$，得

$$\nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) = 0.$$

Karush-Kuhn-Tucker(KKT)条件

定理2'. 考虑问题 $\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \ge 0, i = 1, \cdots, m \end{cases}$

设$\bar{x} \in S, f, g_i$在$\bar{x}$可微,$\{\nabla g_i(\bar{x}) \mid i \in I\}$线性无关,

若$\bar{x}$是局部最优解，则存在数$w_i, i = 1,2,\cdots,m,$使得

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) = 0$$

$$w_i g_i(\bar{x}) = 0 \quad i = 1, 2, \cdots, m$$

互补松弛条件

$$w_i \ge 0 \quad i = 1, 2, \cdots, m.$$

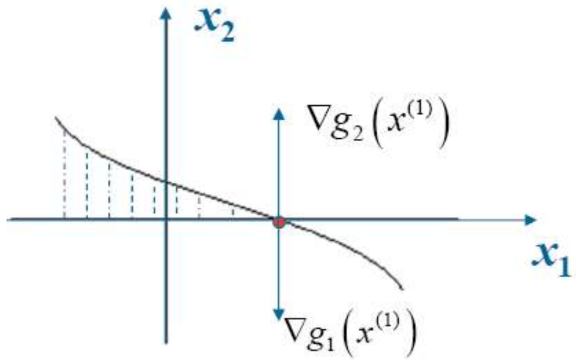
结论注释：在一定条件下，在最优值点，目标函数的梯度可以表示成起作用约束函数梯度的非负线性组合。

无约束优化问题 $\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$ 最优性条件 $\nabla f(\boldsymbol{x}^*) = \mathbf{0}.$ 的推广

约束规格 起作用约束函数的梯度在最优值点向量组线性无关

不唯一，却是必须的！

$$\begin{cases} \min -x_1 \\ s.t. \quad g_1(x) = -x_2 + (1-x_1)^3 \geq 0 \\ \qquad\quad g_2(x) = x_2 \geq 0 \end{cases}$$

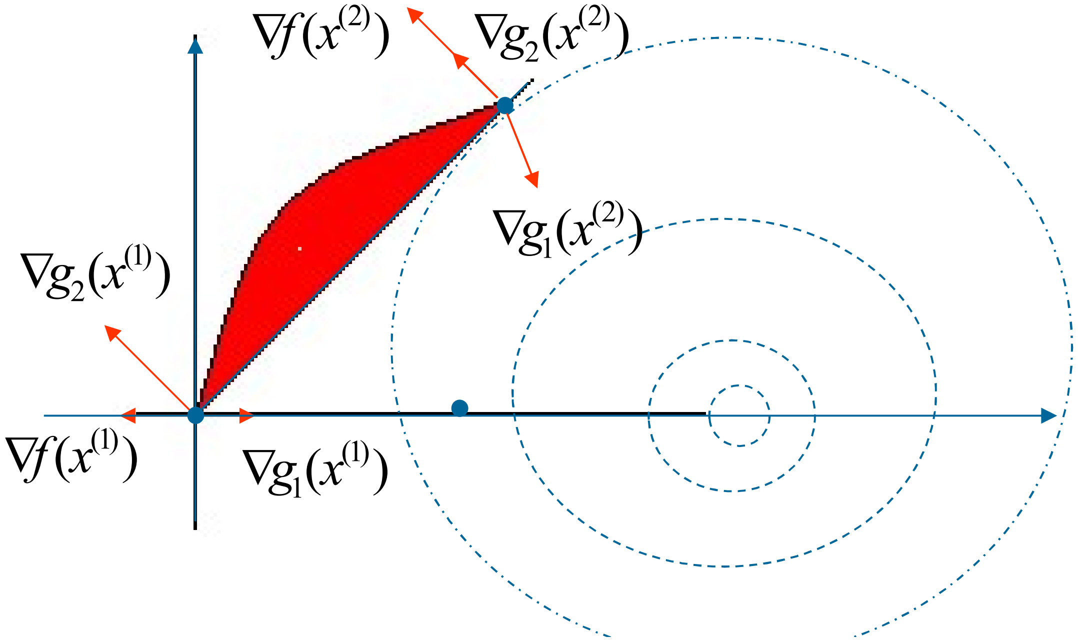


最优解 $(1, 0)^T$ 但定理结论不成立

例：给定非线性规划问题：

$$\begin{cases} \min\ (x_1-2)^2+x_2^2 \\ s.t.\quad x_1-x_2^2\geq 0 \\ \qquad\quad -x_1+x_2\geq 0 \end{cases}$$

验证$x^{(1)}=(0,0)^T, x^{(2)}=(1,1)^T$是否为$KKT$点。



例：给定非线性规划问题$\begin{cases} \min\ f(x)=(x_1-1)^2+x_2 \\ s.t.\quad g_1(x)=-x_1-x_2+2\geq 0 \\ \qquad\quad g_2(x)=x_2\geq 0 \end{cases}$

求满足KKT条件的点。

解：$\nabla f(x)=(2(x_1-1),1)^T,\nabla g_1(x)=(-1,-1)^T,\nabla g_2(x)=(0,1)^T$

设x为满足KTT条件的点，则有

$$\begin{cases} \begin{pmatrix} 2(x_1-1) \\ 1 \end{pmatrix}-w_1\begin{pmatrix} -1 \\ -1 \end{pmatrix}-w_2\begin{pmatrix} 0 \\ 1 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ w_1(-x_1-x_2+2)=0 \\ w_2x_2=0,\quad -x_1-x_2+2\geq 0 \\ w_1,w_2\geq 0,\quad x_2\geq 0 \end{cases}$$

$\Rightarrow x_1=1,x_2=0,w_1=0,w_2=1$

即KKT点为$(1,0)^T$.

例:求下列非线性规划问题的**KKT**点.

$$\min\ f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$$

$$s.t.\quad g_1(x) = -x_1^2 - x_2^2 + 5 \ge 0$$

$$g_2(x) = -3x_1 - x_2 + 6 \ge 0$$

$$\nabla f(x) = \begin{pmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{pmatrix}$$

$$\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} \quad \nabla g_2(x) = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$$

设x为满足KKT条件的点，则有

$$
\begin{cases}
4x_1+2x_2-10+2w_1x_1+3w_2=0 \\
2x_1+2x_2-10+2w_1x_2+w_2=0 \\
w_1(-x_1^2-x_2^2+5)=0 \\
w_2(-3x_1-x_2+6)=0 \\
-x_1^2-x_2^2+5\geq 0 \\
-3x_1-x_2+6\geq 0 \\
w_1,w_2\geq 0
\end{cases}
$$

$$x_1=1, x_2=2,$$

$$w_1=1, w_2=0$$

凸规划

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0, i = 1,2,\cdots,m \\ \qquad\quad h_j(x) = 0, \ j = 1,2,\cdots,l. \end{cases}$$

其中是$f(x)$凸函数，$g_i(x)$是凹函数，$h_j(x)$是线性函数。

定理3.(一阶充分条件)

设问题$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0 \quad i = 1,2,\cdots,m \end{cases}$

中，f是凸函数，$g_i(i = 1,2,\cdots,m)$是凹函数，S为可行域，$\bar{x} \in S$，$I = \{i \mid g_i(\bar{x}) = 0\}$。$f$和$g_i(i \in I)$在点$\bar{x}$可微，$g_i(i \notin I)$在点$\bar{x}$连续，且在$\bar{x}$处$KKT$条件成立，则$\bar{x}$为整体极小点。

证明：显然S为凸集，

$\because \overline{x} \in S, f$为凸函数且在$\overline{x}$可微，$\therefore$对$\forall x \in S$

$$f(x) \geq f(\overline{x}) + \nabla f(\overline{x})^T (x - \overline{x}) -----(1)$$

又点$\overline{x}$处KKT条件成立，所以存在$w_i (i \in I), w_i \geq 0$

使得 $\nabla f(\overline{x}) = \sum_{i \in I} w_i \nabla g_i(\overline{x}) -------(2)$

代入(1)得

$$f(x) \geq f(\overline{x}) + \sum_{i \in I} w_i \nabla g_i(\overline{x})^T (x - \overline{x}) ---(3)$$

$\because g_i$是凹函数，$\therefore$当$i \in I$有

$$g_i(x) \leq g_i(\overline{x}) + \nabla g_i(\overline{x})^T (x - \overline{x})$$

$$\Rightarrow \nabla g_i(\overline{x})^T (x - \overline{x}) \geq g_i(x) - g_i(\overline{x}) = g_i(x) \geq 0 --(4)$$

将(4)代入(3)，得

$f(x) \geq f(\overline{x}) \Rightarrow \overline{x}$是整体最优解.

一般约束问题的一阶最优性条件

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

$$\begin{cases}\min\ f(x) \\ s.t.\quad g(x)\geq 0 \\ \qquad\quad h(x)=0\end{cases} \qquad g(x)=\begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x)\end{bmatrix}, h(x)=\begin{bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_l(x)\end{bmatrix}$$

对任意的 $\boldsymbol{x} \in \Omega$，记

不等式积极约束指标集

$$\mathcal{I}(\boldsymbol{x}) \triangleq \{i \in \mathcal{I} \mid c_i(\boldsymbol{x}) = 0\}$$

积极约束

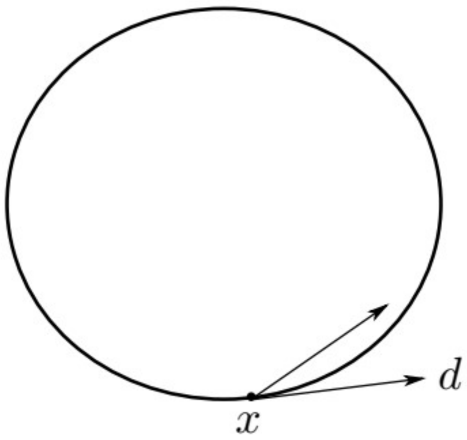
积极约束指标集

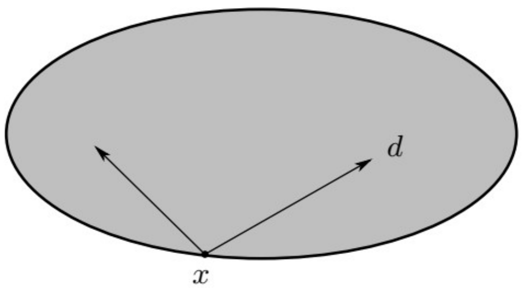
$$\mathcal{A}(\boldsymbol{x}) \triangleq \mathcal{E} \cup \mathcal{I}(\boldsymbol{x})$$

所有的等式约束都是积极约束。

可行方向 设 $\boldsymbol{x} \in \Omega$， $\boldsymbol{d} \in \mathbb{R}^n$. 若存在 $\delta > 0$, 使对任意的 $\alpha \in [0, \delta]$, 都有 $\boldsymbol{x} + \alpha \boldsymbol{d} \in \Omega$, 则称 $\boldsymbol{d}$ 为约束优化问题在 $\boldsymbol{x}$ 点的可行方向.

有的可行域没有可行方向，如 $\Omega = \{\boldsymbol{x} \in R^n \mid \boldsymbol{x}^\top \boldsymbol{x} = 1\}$





定义：设$\bar{x}$为可行点，不等式约束中在$\bar{x}$起作用约束下标集记为I，如果向量组

$$\left\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}) \mid i \in I, j = 1,2,\cdots,l\right\}$$

线性无关，则称$\bar{x}$为约束$g(x) \geq 0$和$h(x) = 0$的正则点。

定义：称点集$\left\{x = x(t) \mid t_0 \leq t \leq t_1\right\}$为曲面$S = \left\{x \mid h(x) = 0\right\}$上的一条曲线，如果对所有$t \in \left[t_0, t_1\right]$均有$h(x(t)) = 0$.

如果导数$x'(t) = \dfrac{\mathrm{d}x(t)}{\mathrm{d}t}$存在，则称曲线是可微的.

设$\bar{x}$是曲面$S=\{x|h(x)=0\}$上的一个点。

定义：称点集$\{x=x(t)\,|\,t_0\le t\le t_1\}$为曲面$S=\{x\,|\,h(x)=0\}$过点$\bar{x}$的一条曲线，如果对所有$t\in[t_0,t_1]$均有$h(x(t))=0$，而且存在$t^*\in[t_0,t_1]$，使得$x(t^*)=\bar{x}$.

曲面S上在点$\bar{x}$处所有可微曲线的切向量组成的集合，称为曲面S在点$\bar{x}$的切平面，记为$T(\bar{x})$.

定义子空间

$$H=\left\{d \mid \nabla h(x)^T d=0\right\}$$

其中$\nabla h(x)=\left(\nabla h_1(x),\nabla h_2(x),\cdots,\nabla h_l(x)\right)$

结论：若向量$d\in T(\bar{x})$，则有

$$d\in H=\left\{d \mid \nabla h(\bar{x})^T d=0\right\}$$

证明： 设 $d \in T(\bar{x})$，则有曲线 $x(t)$ 使得 $h(x(t)) = 0 \left(t_0 \le t \le t_1\right)$，

且存在 $t_0 \le t^* \le t_1$ 使得 $x(t^*) = \bar{x}, x'(t^*) = d$

由于

$$\frac{\mathrm{d}h(x(t))}{\mathrm{d}t} = \begin{pmatrix} \dfrac{\mathrm{d}h_1(x(t))}{\mathrm{d}t} \\ \dfrac{\mathrm{d}h_2(x(t))}{\mathrm{d}t} \\ \cdots \\ \dfrac{\mathrm{d}h_l(x(t))}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \nabla h_1(x(t))^T x'(t) \\ \nabla h_2(x(t))^T x'(t) \\ \vdots \\ \nabla h_l(x(t))^T x'(t) \end{pmatrix}$$

所以 $\nabla h(\bar{x})^T d = 0$

$\Rightarrow \quad d \in H$

补充知识：隐函数定理

将线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的系数矩阵和变量分块成

$$\boldsymbol{A}=(\boldsymbol{B}\quad \boldsymbol{N}),\quad \boldsymbol{x}=\begin{pmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{pmatrix}$$

则该线性方程组可以写成 $\boldsymbol{B}\boldsymbol{x}_B+\boldsymbol{N}\boldsymbol{x}_N=\boldsymbol{b}$

进一步，若分块矩阵 $\boldsymbol{B}$ 非奇异，则由该线性方程组可得

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}(\boldsymbol{b}-\boldsymbol{N}\boldsymbol{x}_N)$$

即在一定条件下，线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 确定隐函数

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}(\boldsymbol{b}-\boldsymbol{N}\boldsymbol{x}_N)$$

定理：设$\bar{x}$是曲面$S=\{x \mid h(x)=0\}$上的一个正则点，则

$$T(\bar{x})=H=\{d \mid \nabla h(\bar{x})^T d=0\}.$$

证明：设$d \in H$,考虑非线性方程组$h(\bar{x}+td+\nabla h(\bar{x})y)=0$

其中$t \in R^1, y \in R^l$.该方程组有解$(y,t)=(0,0)$.

在$t=0$处，h关于y的$Jacobi$矩阵为$\nabla h(\bar{x})^T \nabla h(\bar{x})$

由隐函数定理，在$t=0$的邻域，存在连续可微函数

$y=y(t)(y(0)=0)$,使$h(\bar{x}+td+\nabla h(\bar{x})y(t))=0$成立

令$x(t)=\bar{x}+td+\nabla h(\bar{x})y(t)$,则$x(t)$为曲面$S$上过$x(0)$的一条曲线。在点$x(0)=\bar{x}$，切向量为

$$x'(0)=d+\nabla h(\bar{x})y'(0)$$

又 $\nabla h(\bar{x})^T d+\nabla h(\bar{x})^T \nabla h(\bar{x})y'(0)=0$

而$d \in H \Rightarrow \nabla h(\bar{x})^T d=0, \because \bar{x}$是正则点$, \therefore \nabla h(\bar{x})^T \nabla h(\bar{x})$可逆，

$\Rightarrow y'(0)=0 \Rightarrow x'(0)=d \Rightarrow d \in T(\bar{x}).$

定理：设$\bar{x} \in S$，$f(x)$和$g_i(x)(i \in I)$在$\bar{x}$处可微，$g_i(x)(i \notin I)$在$\bar{x}$处连续，$h_j(j=1,2,\cdots,l)$在$\bar{x}$处连续可微，且$\bar{x}$是$S=\{x \mid h(x)=0\}$的正则点。如果$\bar{x}$是问题(NP)的局部最优解，则在$\bar{x}$处，有

$$F_0 \cap G_0 \cap H_0 = \varnothing,$$

其中

$$F_0 = \left\{d \mid \nabla f(\bar{x})^T d < 0\right\},$$

$$G_0 = \left\{d \mid \nabla g_i(\bar{x})^T d > 0, i \in I\right\},$$

$$H_0 = \left\{d \mid \nabla h_j(\bar{x})^T d = 0, j = 1,2,\cdots,l\right\}.$$

证明：设存在向量$y \in F_0 \cap G_0 \cap H_0$,则有$\nabla f(\overline{x})^T y < 0$,
$\nabla g_i(\overline{x})^T y > 0, i \in I$和$\nabla h_j(\overline{x})^T y = 0, j = 1,2,\cdots,l$.由定理，
有$y \in T(\overline{x})$,即在曲面$S = \{x \mid h(x) = 0\}$上存在经过点
$\overline{x} = x(0)$的可微曲线$x(t)$,使得$\dfrac{dx(0)}{dt} = y$.

(1)$i \in I$时，有$\dfrac{dg_i(x(t))}{dt}\Big|_{t=0} = \nabla g_i(\overline{x})^T \dfrac{dx(0)}{dt} = \nabla g_i(\overline{x})^T y > 0$,
因此存在$\delta_1 > 0$,当$t \in [0, \delta_1)$时，$g_i(x(t)) \geq 0$;
(2)$i \notin I$时，由于$g_i(\overline{x}) > 0$且g_i在$\overline{x}$连续，因此存在$\delta_2 > 0$,当
$t \in (0, \delta_2)$时，$g_i(x(t)) \geq 0$;
(3)因为$\dfrac{df(x(t))}{dt}\Big|_{t=0} = \nabla f(\overline{x})^T y < 0$,因此存在$\delta_3 > 0$,当
$t \in [0, \delta_3)$时，有$f(x(t)) < f(\overline{x})$;
取$\delta = \min\{\delta_1, \delta_2, \delta_3\}$,则当$t \in (0, \delta)$时$x(t)$是可行解，且
$f(x(t)) < f(\overline{x})$，矛盾。

引理：若系统$Ax<0, Bx=0$无解，则系统

$A^T y + B^T z = 0, y \ge 0$，且$y \ne 0$或$z \ne 0$有解.

证明：假设系统$Ax<0, Bx=0$无解，定义

$$S_1 = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \middle| d_1 = Ax, d_2 = Bx, x \in R^n \right\}, \quad S_2 = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \middle| d_1 < 0, d_2 = 0 \right\}$$

则S_1, S_2为非空凸集，且$S_1 \cap S_2 = \varnothing$,由凸集分离定理

存在非零向量$p = \begin{pmatrix} y \\ z \end{pmatrix}$,使得对$\forall x \in R^n, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \in clS_2$有

$$y^T Ax + z^T Bx \ge y^T d_1 + z^T d_2 = y^T d_1$$

令$x = 0, \because d_1 < 0, \therefore y \ge 0$.

令$d_1 \to 0$，得$y^T Ax + z^T Bx \ge 0$.

取$x = -\left(A^T y + B^T z\right)$,则有$-\left\| A^T y + B^T z \right\|^2 \ge 0$

$\Rightarrow A^T y + B^T z = 0$.

定理1($Fritz\ John$条件) 设$\bar{x} \in S, I = \{i \mid g_i(\bar{x}) = 0\}$，$f(x), g_i(x)(i \in I)$在$\bar{x}$处可微， $g_i(x)(i \notin I)$在$\bar{x}$处连续， $h_j(j = 1,2,\cdots,l)$在$\bar{x}$处连续可微，若 $\bar{x}$是问题(NP)的局部最优解，则存在 不全为零的数w_0，$w_i(i \in I)$和$v_j(j = 1,2,\cdots,l)$，使得

$$\begin{cases} w_0 \nabla f(\bar{x}) - \sum_{i \in I} w_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{x}) = 0 \\ w_0, \ w_i \geq 0, \quad i \in I. \end{cases}$$

证明：不妨假设$\nabla h_1(\overline{x}),\cdots,\nabla h_l(\overline{x})$线性无关，

则有　$F_0\cap G_0\cap H_0=\varnothing$

即不等式组
$$\begin{cases}\nabla f(\overline{x})^T d<0\\ \nabla g_i(\overline{x})^T d>0, \quad i\in I\\ \nabla h_j(\overline{x})^T d=0, \quad j=1,2,\cdots,l\end{cases}$$

无解. 令A是以$\nabla f(\overline{x})^T$，$-\nabla g_i(\overline{x})^T\left(i\in I\right)$为行组成的矩阵，

B是以$-\nabla h_j(\overline{x})^T\left(j=1,2,\cdots,l\right)$为行组成的矩阵,

则系统$Ad<0, Bd=0$无解，由引理，存在非零向量$P=\begin{bmatrix}p_1\\ p_2\end{bmatrix}$

且$p_1\geq 0$, 使得　$A^T p_1+B^T p_2=0$

记p_1的分量为$w_0, w_i\left(i\in I\right)$, p_2的分量为$v_j\left(j=1,\cdots,l\right)$,则有

$$w_0\nabla f(\overline{x})-\sum_{i\in I}w_i\nabla g_i(\overline{x})-\sum_{j=1}^{l}v_j\nabla h_j(\overline{x})=0$$

$w_0, w_i\left(i\in I\right)\geq 0$,且$w_0, w_i\left(i\in I\right), v_j\left(j=1,\cdots,l\right)$不全为零.

定理1'(*Fritz John*条件) 设$\bar{x}\in S, f(x), g_i(x)$在$\bar{x}$处可微，$h_j(j=1,\cdots,l)$在$\bar{x}$处连续可微，若$\bar{x}$是问题$(NP)$的局部最优解，则存在不全为零的数$w_0, w_1, \cdots, w_m$,和$v_j(j=1,\cdots,l)$使得

$$\begin{cases} w_0\nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) - \sum_{i=1}^{l} v_i \nabla h_i(\bar{x}) = 0 \\ w_i g_i(\bar{x}) = 0, \quad i = 1,2,\cdots,m \\ w_0,\ w_i \ge 0, \quad i=1,2,\cdots,m. \end{cases}$$

互补松弛条件

定理2(KKT必要条件）考虑问题$\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0, i=1,\cdots,m \\ \qquad\ h_j(x)=0,\ j=1,\cdots,l\end{cases}$

设$\overline{x}$为可行点,$I=\{i \mid g_i(\overline{x})=0\}$.$f,g_i(i\in I)$在$\overline{x}$处可微，

$g_i(i\notin I)$在$\overline{x}$连续，$h_j(j=1,\cdots,l)$在$\overline{x}$连续可微，向量集

$$\{\nabla g_i(\overline{x}),\ \nabla h_j(\overline{x}) \mid i\in I, j=1,\cdots,l\}$$

线性无关，若$\overline{x}$是局部最优解，则存在数$w_i, i\in I$

和$v_j(j=1,\cdots,l)$，使得

$$\nabla f(\overline{x})-\sum_{i\in I} w_i \nabla g_i(\overline{x})-\sum_{j=1}^{l} v_j \nabla h_j(\overline{x})=0.$$

$$w_i\geq 0 \quad (i\in I).$$

定理2'(*KKT*必要条件）考虑问题$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0, i=1,\cdots,m \\ \qquad\quad h_j(x)=0,\ j=1,\cdots,l \end{cases}$

设$\bar{x}$为可行点, f, g_i在$\bar{x}$处可微，$h_j(j=1,\cdots,l)$在$\bar{x}$连续可微，向量集

$$\{\nabla g_i(\bar{x}),\ \nabla h_j(\bar{x}) \mid i \in I, j=1,\cdots,l\}$$

线性无关，若$\bar{x}$是局部最优解，则存在数$w_i, i \in I$和$v_j (j=1,\cdots,l)$，使得

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{x}) = 0$$

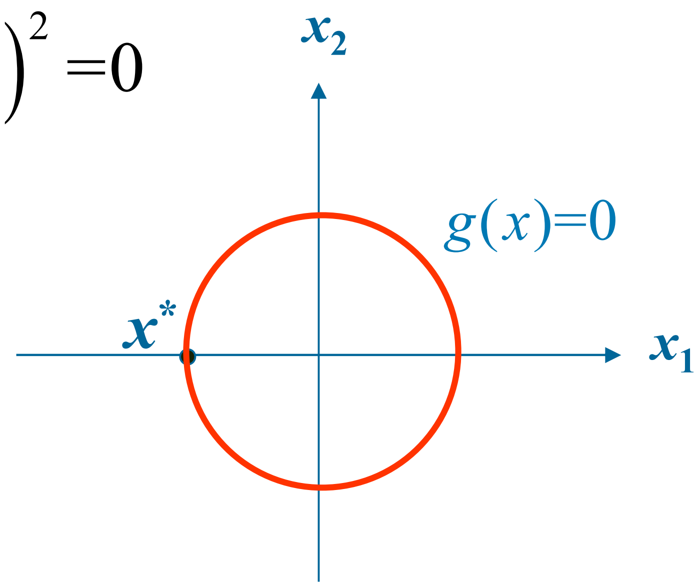
$$w_i g_i(\bar{x}) = 0, i=1,2,\cdots,m$$

$$w_i \geq 0 \quad i=1,2,\cdots,m.$$

例：考虑如下约束优化问题：

$$\min\ f(x) = x_1$$

$$s.t.\quad g(x) = \left(1 - x_1^2 - x_2^2\right)^2 = 0$$



约束规格 $\boxed{\text{约束函数的梯度在最优值点向量组线性无关}}$

不唯一，却是必须的！

定义广义的Lagrange函数:

$$L(x,w,v)=f(x)-\sum_{i=1}^{m}w_i g_i(x)-\sum_{j=1}^{l}v_j h_j(x)$$

$$=f(x)-w^T g(x)-v^T h(x)$$

其中 $w=(w_1,w_2,\cdots,w_m)^T$

$$v=(v_1,v_2,\cdots,v_l)^T$$

乘子向量

$$g(x)=(g_1(x),\cdots,g_m(x))^T$$

$$h(x)=(h_1(x),\cdots,h_l(x))^T.$$

定理2'(*KKT*必要条件）考虑问题$\begin{cases}\min\ f(x)\\ s.t.\quad g_i(x)\ge 0,\ i=1,\cdots,m\\ \qquad\ \ h_j(x)=0,\ j=1,\cdots,l\end{cases}$

设$\overline{x}$为可行点,$I=\{i\mid g_i(\overline{x})=0\}.f,g_i$在$\overline{x}$处可微，$h_j(j=1,\cdots,l)$在$\overline{x}$连续可微，向量集

$$\{\nabla g_i(\overline{x}),\ \nabla h_j(\overline{x})\mid i\in I, j=1,\cdots,l\}$$

线性无关，若$\overline{x}$是局部最优解，则存在乘子向量$\overline{w}\ge 0,\overline{v}$,使得

$$\nabla_x L(\overline{x},\overline{w},\overline{v})=0$$

$$\overline{w}^T g(\overline{x})=0$$

一般情形的一阶必要条件(**KKT**必要条件)可表示为**:**

$$\begin{cases} \nabla_x L(x,w,v)=0 \\ w_i g_i(x)=0,\ i=1,2,\cdots,m \\ g_i(x)\geq 0,\ i=1,2,\cdots,m \\ h_j(x)=0,\ j=1,2,\cdots,l \\ w_i\geq 0,\ i=1,2,\cdots,m \end{cases}$$

定理3.(一阶充分条件)

设问题$\begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \ge 0 \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,2,\cdots,l \end{cases}$

中，f是凸函数，$g_i(i=1,2,\cdots,m)$是凹函数，$h_j(j=1,2,\cdots,l)$是线性函数，S为可行域，$\overline{x}\in S$，$I=\{i \mid g_i(\overline{x})=0\}$。$f$和$g_i(i\in I)$在点$\overline{x}$可微，$h_j(j=1,2,\cdots,l)$在点$\overline{x}$连续可微，$g_i(i\notin I)$在点$\overline{x}$连续，且在$\overline{x}$处$KKT$条件成立，则$\overline{x}$为整体极小点。

证明：显然S为凸集,

$\because \bar{x} \in S, f$为凸函数且在$\bar{x}$可微,$\therefore$ 对$\forall x \in S$

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$$

而 $\nabla f(\bar{x}) = \sum_{i \in I} w_i \nabla g_i(\bar{x}) + \sum_{j=1}^{l} v_j \nabla h_j(\bar{x})$

$$\therefore f(x) \geq f(\bar{x}) + \sum_{i \in I} w_i \nabla g_i(\bar{x})^T (x - \bar{x}) + \sum_{j=1}^{l} v_j \nabla h_j(\bar{x})^T (x - \bar{x})$$

$\because g_i$是凹函数, $\therefore$ 当$i \in I$有

$$g_i(x) \leq g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x - \bar{x})$$

$$\Rightarrow \nabla g_i(\bar{x})^T (x - \bar{x}) \geq g_i(x) - g_i(\bar{x}) = g_i(x) \geq 0$$

$\because h_j$为线性函数 $\therefore h_j(x) = h_j(\bar{x}) + \nabla h_j(\bar{x})^T (x - \bar{x})$

$$\Rightarrow \quad \nabla h_j(\bar{x})^T (x - \bar{x}) = 0$$

$f(x) \geq f(\bar{x}) \Rightarrow \bar{x}$是整体最优解.

推论：问题$\begin{cases}\min cx \\ s.t. \quad Ax \geq b \\ \qquad\quad x \geq 0\end{cases}$,则$x^*$是最优解

$\Leftrightarrow$ 存在$w \in R^m, v \in R^n$,使得

$$\begin{cases} Ax^* \geq b \\ x^* \geq 0 \\ c - w^T A - v^T = 0 \\ w^T (Ax^* - b) = 0 \\ v^T x^* = 0 \\ w, v \geq 0 \end{cases}$$

（必要性）设x^*是(L)的最优解，则(1)成立。假设

$I_1 x^* > 0,\quad I_2 x^* = 0$，其中$I_1$，$I_2$分别为$p\times n,(n-p)\times n$矩阵

$$A_1 x^* = b_1,\quad A_2 x^* > b_2.$$

令$B=\begin{pmatrix} A_1 \\ I_2 \end{pmatrix}$，对任意满足$Bd\ge 0$的向量$d$以及充分小$\lambda>0$，有

$$A(x^*+\lambda d)\ge b,\quad x^*+\lambda d\ge 0,$$

$\Rightarrow x^*+\lambda d$为可行解，由于$x^*$为最优解，所以

$$cx^*\le c(x^*+\lambda d)\Rightarrow cd\ge 0$$

$\Rightarrow$ 系统$cd<0,\quad Bd\ge 0$无解。由$Farkas$引理，存在向量$u\ge 0$

使得$B^T u=(A_1^T, I_2^T)u=c^T$.记$u=(u_1,u_2)^T$，令

$$w=\begin{pmatrix} u_1 \\ 0 \end{pmatrix}\ge 0,\ v=\begin{pmatrix} 0 \\ u_2 \end{pmatrix}\ge 0$$

则 $A^T w+v=B^T u=c^T$，所以$c-w^T A-v^T=0$成立。

显然$v^T x^*=0$.注意到$Ax^*-b=\begin{pmatrix} 0 \\ A_2 x^*-b_2 \end{pmatrix}$，有$w^T(Ax^*-b)=0$，

即(3)成立。

证明：（充分性）设x^*是满足条件的点，则x^*是(L)的可行解；设x是任一可行解，则

$$0=(c^T-A^Tw-r)^T(x^*-x)$$

$$=c(x^*-x)-w^T(Ax^*-b)-r^Tx^*+w^T(Ax-b)+r^Tx$$

$$=c(x^*-x)+w^T(Ax-b)+r^Tx$$

$\because w\geq 0, r\geq 0, \therefore cx^*\leq cx\Rightarrow x^*$是最优解。

推论：对于标准形式的LP问题

$$\begin{cases} \min cx \\ s.t. \quad Ax = b \\ \qquad\quad x \geq 0 \end{cases}$$

则x是该问题的最优解，当且仅当存在向量w，r，使得

$$Ax = b, \quad x \geq 0;$$

$$A^T w + r = c^T, \quad r \geq 0;$$

$$r^T x = 0.$$

求解下列线性规划问题：

$$\begin{cases} \min -2x_1 + x_2 \\ s.t. \quad x_1 + x_2 + x_3 \geq 4 \\ \qquad\quad x_1 + 2x_2 + 2x_3 \leq 6 \\ \qquad\quad x_1, x_2, x_3 \geq 0 \end{cases}$$

$$\Rightarrow \begin{cases} \min -2x_1 + x_2 \\ s.t. \quad x_1 + x_2 + x_3 - 4 \geq 0 \\ \qquad\quad -x_1 - 2x_2 - 2x_3 + 6 \geq 0 \\ \qquad\quad x_1, x_2, x_3 \geq 0 \end{cases}$$

由**KKT**条件,得

$$
\begin{cases}
-2-w_1+w_2-w_3=0 & (1)\\
1-w_1+2w_2-w_4=0 & (2)\\
-w_1+2w_2-w_5=0 & (3)\\
w_1(x_1+x_2+x_3-4)=0 & (4)\\
w_2(-x_1-2x_2-2x_3+6)=0 & (5)\\
w_3x_1=0 \quad w_4x_2=0 \quad w_5x_3=0 & (6)\\
w_i\ge 0, x_i\ge 0 & (7)\\
x_1+x_2+x_3-4\ge 0 & (8)\\
-x_1-2x_2-2x_3+6\ge 0 & (9)
\end{cases}
$$

得到KKT点$(6, 0, 0)^T$.

$(6,0,0)^T$为整体最优解。

例:考虑问题

$$\begin{cases} \min \left(x_1 - \frac{9}{4}\right)^2 + (x_2 - 2)^2 \\ s.t. \quad x_2 - x_1^2 \geq 0 \\ \qquad\quad x_1 + x_2 \leq 6 \\ \qquad\quad x_1, x_2 \geq 0 \end{cases}$$

(1)写出KKT必要条件，并验证在点$x^* = \begin{pmatrix} \frac{3}{2} & \frac{9}{4} \end{pmatrix}^T$成立。

（2）证明x^*是全局最优解。

KKT必要条件为

$$\begin{cases} \begin{pmatrix} 2\left(x_1-\dfrac{9}{4}\right) \\ 2(x_2-2) \end{pmatrix} - w_1\begin{pmatrix} -2x_1 \\ 1 \end{pmatrix} - w_2\begin{pmatrix} -1 \\ -1 \end{pmatrix} - w_3\begin{pmatrix} 1 \\ 0 \end{pmatrix} - w_4\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \\ w_1(x_2-x_1^2)=0 \\ w_2(-x_1-x_2+6)=0 \\ w_3x_1=0 \\ w_4x_2=0 \\ x_2-x_1^2\geq 0 \\ -x_1-x_2+6\geq 0 \\ x_1,x_2,w_1,w_2,w_3,w_4\geq 0 \end{cases}$$

用**KKT**条件解下列问题:

$$\begin{cases} \min\ f = (x_1 - 1)^2 + (x_2 - 1)^2 \\ s.t. \quad -x_1 - x_2 + 2 \geq 0 \\ \qquad\quad x_2 - x_1 = 1 \\ \qquad\quad x_1 \geq 0 \\ \qquad\quad x_2 \geq 0 \end{cases}$$

$$\nabla f(x) = \left(2(x_1 - 1),\ 2(x_2 - 1)\right)^T \quad \nabla g_1(x) = (-1, -1)^T$$

$$\nabla g_2(x) = (1, 0)^T \quad \nabla g_3(x) = (0, 1)^T \quad \nabla h(x) = (-1, 1)^T$$

KKT条件为：

$$
\begin{cases}
2(x_1-1)-w_1(-1)-w_2-v(-1)=0 \\
2(x_2-1)-w_1(-1)-w_3-v=0 \\
-x_1-x_2+2\geq 0 \\
x_2-x_1=1 \\
x_1,x_2\geq 0 \\
w_1(-x_1-x_2+2)=0 \\
w_2x_1=0 \\
w_3x_2=0 \\
w_1,w_2,w_3\geq 0
\end{cases}
$$

$$x^* = \left(\frac{1}{2}, \frac{3}{2}\right)^T$$ 为KKT点.

因为$f(x)$为凸函数，$g_i(x)$，$h(x)$为线性函数，所以本问题为凸规划问题，

$\Rightarrow x^*$为全局最优解

$$f_{\min} = \frac{1}{2}$$

$$\min\ f(x) = x_2$$
$$s.t.\quad g(x) = x_1^2 + x_2 \geq 0$$

KKT点应满足方程组

$$\begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - w \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} = 0 \\ w(x_1^2 + x_2) = 0 \\ x_1^2 + x_2 \geq 0 \\ w \geq 0 \end{cases} \Rightarrow \begin{matrix} w = 1 \\ x_1 = x_2 = 0 \end{matrix}$$

$x^* = (0, 0)^T$ 不是极小点。

二阶条件

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

$$g(x)=(g_1(x),\cdots,g_m(x))^T$$

$$h(x)=(h_1(x),\cdots,h_l(x))^T$$

Lagrange 函数

$$L(x,w,v)=f(x)-w^T g(x)-v^T h(x)$$

其中$w\in E^m,v\in R^l$,称为$Lagrange$乘子。

$Lagrange$乘子问题：是否存在$x\in E^n,w\in R^m,v\in R^l$满足方程组：

$$(LM)\begin{cases}\nabla_x L(x,w,v)=\nabla f(x)-\nabla g(x)w-\nabla h(x)v=0\\ \nabla_w L(x,w,v)=-g(x)\le 0\\ \nabla_v L(x,w,v)=h(x)=0\\ w\ge 0\\ w_i g_i(x)=0,\ i=1,2,\cdots,m.\end{cases}$$

定理(KKT必要条件）考虑问题$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \ge 0,\ i = 1, \cdots, m \\ \qquad\quad h_j(x) = 0,\ j = 1, \cdots, l \end{cases}$

设$\bar{x}$为可行点,$I = \{i \mid g_i(\bar{x}) = 0\}$.$f, g_i$在$\bar{x}$处可微，$h_j(j = 1, \cdots, l)$在$\bar{x}$连续可微，向量集

$$\{\nabla g_i(\bar{x}),\ \nabla h_j(\bar{x}) \mid i \in I,\ j = 1, \cdots, l\}$$

线性无关，若$\bar{x}$是局部最优解，则存在$\bar{w}, \bar{v}$,使得$(\bar{x}, \bar{w}, \bar{v})$为$(LM)$的解。

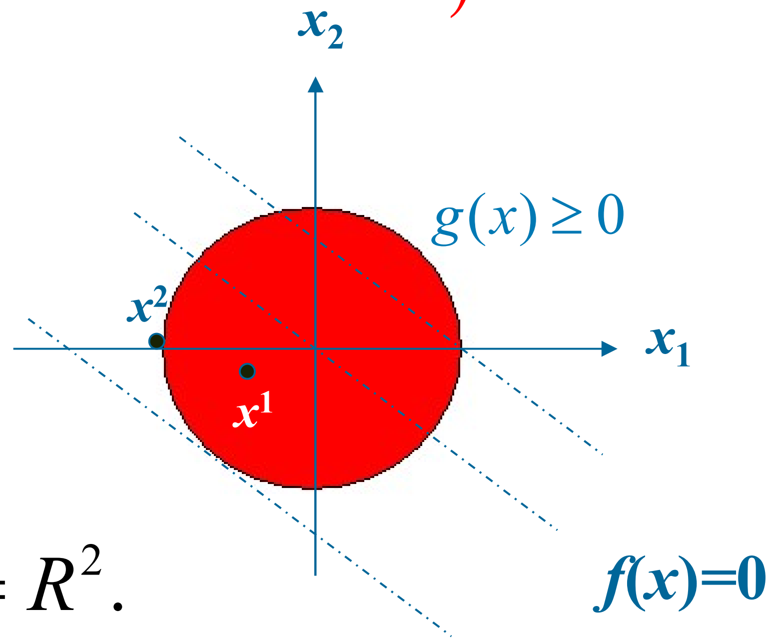
定义：设S是R^n中一个非空集合，点$\bar{x} \in clS$，集合

$$T=\left\{d \mid x^{(k)} \in S,\, x^{(k)} \to \bar{x},\, \lambda_k > 0,\, d=\lim_{k\to\infty} \lambda_k \left(x^{(k)}-\bar{x}\right)\right\},$$

则称T为集合S在点$\bar{x}$的切锥(tangent cone)或称为序列化可行方向锥(sequential feasible cone)。

$$\min\ f(x) = x_1 + x_2$$

$$s.t.\quad g(x) = 1 - x_1^2 - x_2^2 \geq 0$$



对于任意内点x^1，切锥$=R^2$.

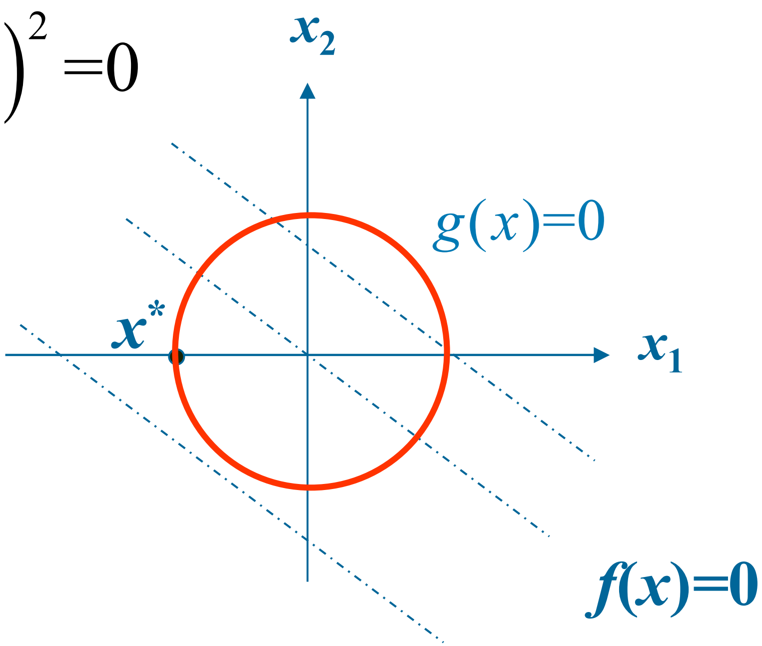
对于边界点$x^2=(-1,0)^T$,切锥

$$D=\{d \in R^2 \mid d_1 \geq 0\}.$$

例：考虑如下约束优化问题：

$$\min\ f(x) = x_1 + x_2$$

$$s.t.\quad g(x) = \left(1 - x_1^2 - x_2^2\right)^2 = 0$$



在$x^* = (-1,0)^T$，可行方向锥 $= \varnothing$.

在点$x^* = (-1, 0)^T$,切锥$D = \{d \in R^2 \mid d_1=0\}$.

命题：假设确定集合S的所有约束函数在$x \in S$处连续可微，则有

$$D(x,S) \subseteq SFD(x,S)$$

其中$D(x,S)$为x点的可行方向锥，$SFD(x,S)$为x点的切锥。

证明：$\forall d \in D(x,S)$,则存在$\delta > 0$,使得当$\lambda \in \left(0，\delta\right)$,有$x + \lambda d \in S$.

令$d^k = d, \lambda_k = \dfrac{2^k}{\delta} > 0, x^k = x + \dfrac{1}{\lambda_k} d^k \in S$,

则$x^k \to x$,且$d^k \to d$,

$\therefore d \in SFD(x,S)$.

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

设$\overline{x}$是(NP)的一个KKT点，$I=\{i\mid g_i(\overline{x})=0\}$.

即存在数$\overline{w}_i, i\in I$和$\overline{v}_j\,(j=1,\cdots,l)$，使得

$$\nabla f(\overline{x})-\sum_{i\in I}\overline{w}_i\nabla g_i(\overline{x})-\sum_{j=1}^{l}\overline{v}_j\nabla h_j(\overline{x})=0.$$

$$\overline{w}_i\geq 0, i\in I$$

定义 $$\bar{S}=\left\{x\left|\begin{array}{ll}x\in R^n & \\ g_i(x)=0, & i\in I\text{且}\bar{w}_i>0\\ g_i(x)\ge 0, & i\in I\text{且}\bar{w}_i=0\\ h_j(x)=0, & j=1,2,\cdots,l\end{array}\right.\right\}$$

设集合$\bar{S}$在点$\bar{x}$的切锥为$\bar{T}$.

定义 $$\bar{G}=\left\{d\left|\begin{array}{l}d\in R^n\\ \nabla g_i(\bar{x})^T d=0, i\in I\text{且}\bar{w}_i>0\\ \nabla g_i(\bar{x})^T d\ge 0, i\in I\text{且}\bar{w}_i=0\\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l\end{array}\right.\right\}.$$

结论：$\bar{G}\supseteq\bar{T}$.

例：给定非线性规划问题：

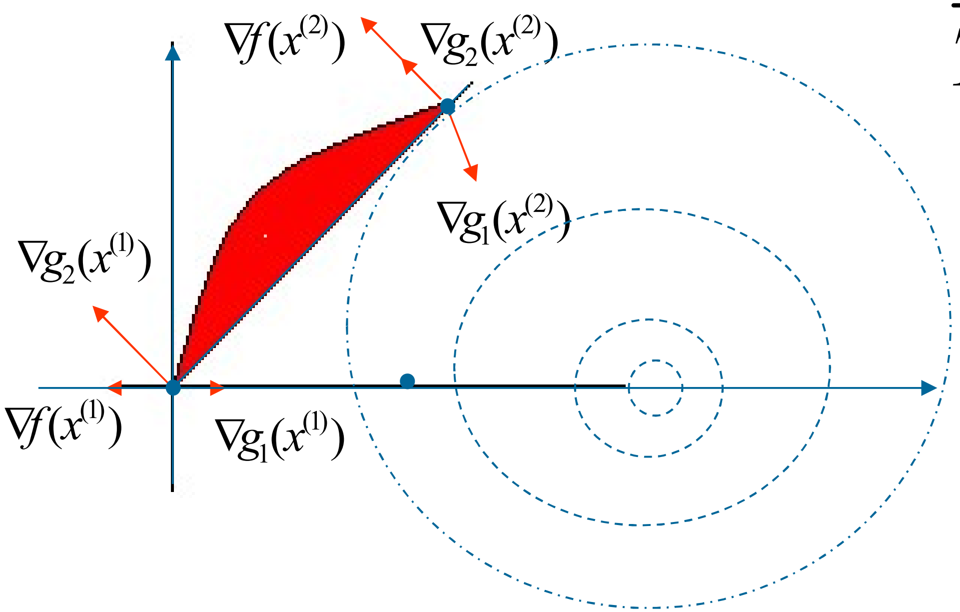
$$\begin{cases} \min\ f(x) = (x_1 - 2)^2 + x_2^2 \\ s.t.\quad g_1(x) = x_1 - x_2^2 \geq 0 \\ \qquad\quad g_2(x) = -x_1 + x_2 \geq 0 \end{cases}$$

$$\bar{S} = \{x \mid x_1 - x_2^2 \geq 0, -x_1 + x_2 = 0\}$$

$$= \left\{ \begin{pmatrix} t \\ t \end{pmatrix} \mid 0 \leq t \leq 1 \right\}$$

$x^{(2)} = (1,1)^T$ 是 KKT 点，且

$$\nabla f\left(x^{(2)}\right) - w_2 \nabla g_2\left(x^{(2)}\right) = 0$$



$$\bar{T} \subseteq \bar{G} = \left\{ \begin{pmatrix} t \\ t \end{pmatrix} \mid t \leq 0 \right\}$$

证明：设$d\in\overline{T}$,则存在$\left\{x^{(k)}\right\}\subseteq\overline{S}$和正数列$\{\lambda_k\}$,

使得$\lim\limits_{k\to\infty}\lambda_k\left(x^{(k)}-\overline{x}\right)=d$

$$g_i\left(x^{(k)}\right)=g_i\left(\overline{x}\right)+\nabla g_i\left(\overline{x}\right)^T\left(x^{(k)}-\overline{x}\right)+o\left(\left\|x^{(k)}-\overline{x}\right\|\right),$$

$$h_j\left(x^{(k)}\right)=h_j\left(\overline{x}\right)+\nabla h_j\left(\overline{x}\right)^T\left(x^{(k)}-\overline{x}\right)+o\left(\left\|x^{(k)}-\overline{x}\right\|\right).$$

当$i\in I$且$w_i>0$时，有　$\nabla g_i\left(\overline{x}\right)^T\left(x^{(k)}-\overline{x}\right)+o\left(\left\|x^{(k)}-\overline{x}\right\|\right)=0;$

当$i\in I$且$w_i=0$时，有　$\nabla g_i\left(\overline{x}\right)^T\left(x^{(k)}-\overline{x}\right)+o\left(\left\|x^{(k)}-\overline{x}\right\|\right)\geq 0;$

当$j=1,2,\cdots,l$时，有　$\nabla h_j\left(\overline{x}\right)^T\left(x^{(k)}-\overline{x}\right)+o\left(\left\|x^{(k)}-\overline{x}\right\|\right)=0.$

把以上各式两端乘以λ_k，令$k\to\infty$，得到

$$\begin{aligned}
&\nabla g_i\left(\overline{x}\right)^T d=0,\quad i\in I\text{且}w_i>0\\
&\nabla g_i\left(\overline{x}\right)^T d\geq 0,\quad i\in I\text{且}w_i=0\quad\Rightarrow d\in\overline{G}\Rightarrow\overline{G}\supseteq\overline{T}\\
&\nabla h_j\left(\overline{x}\right)^T d=0,\quad j=1,2,\cdots,l
\end{aligned}$$

定理（二阶必要条件）：设$\bar{x}$是(NP)的局部最优解，f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$二次连续可微，且存在$\bar{w},\bar{v}$，使$(\bar{x},\bar{w},\bar{v})$为$(LM)$的解, 又假设在$\bar{x}$处，约束规格$\bar{G}=\bar{T}$成立, 则对任意的$d\in\bar{G}$，都有

$$d^T\nabla_x^2 L(\bar{x},\bar{w},\bar{v})d\geq 0,$$

即$\nabla_x^2 L(\bar{x},\bar{w},\bar{v})$在子空间$\bar{G}$上是半正定的.

$$\bar{G}=\left\{d\left|\begin{array}{l}d\in R^n\\ \nabla g_i(\bar{x})^T d=0, i\in I\text{且}\bar{w}_i>0\\ \nabla g_i(\bar{x})^T d\geq 0, i\in I\text{且}\bar{w}_i=0\\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l\end{array}\right.\right\}.$$

证明：设向量$d \in \overline{G} \setminus \{0\}$，则存在$\{x^{(k)}\} \subseteq \overline{S}$和$\{\lambda_k\}$

使得 $\lim\limits_{k \to \infty} \lambda_k (x^{(k)} - \overline{x}) = d$。

$$\boxed{L(x,w,v) = f(x) - \sum_{i=1}^{m} w_i g_i(x) - \sum_{j=1}^{l} v_j h_j(x)}$$

$$L(x^{(k)}, \overline{w}, \overline{v}) = L(\overline{x}, \overline{w}, \overline{v}) + \nabla_x L(\overline{x}, \overline{w}, \overline{v})^T (x^{(k)} - \overline{x})$$

$$+ \frac{1}{2}(x^{(k)} - \overline{x})^T \nabla_x^2 L(\overline{x}, \overline{w}, \overline{v})^T (x^{(k)} - \overline{x}) + o(\| x^{(k)} - \overline{x} \|)$$

$\because x^{(k)} \in \overline{S}$, $\overline{x}$为局部最优解，

$\therefore f(x^{(k)}) = f(\overline{x})$

$$+ \frac{1}{2}(x^{(k)} - \overline{x})^T \nabla_x^2 L(\overline{x}, \overline{w}, \overline{v})^T (x^{(k)} - \overline{x}) + o(\| x^{(k)} - \overline{x} \|).$$

$\because \overline{x}$为局部最优解，$\therefore f(x^{(k)}) \ge f(\overline{x})$

$$\Rightarrow \quad d^T \nabla_x^2 L(\overline{x}, \overline{w}, \overline{v})^T d \ge 0.$$

定理（二阶必要条件）：设$\bar{x}$是(NP)的局部最优解，f，$g_i\ (i=1,\cdots,m)$和$h_j\ (j=1,\cdots,l)$二次连续可微，且在$\bar{x}$处，$\{\nabla g_i(\bar{x}), i\in I, \nabla h_j(\bar{x}), j=1,\cdots,l\}$为线性无关组，则存在$\bar{w},\bar{v}$，使$(\bar{x},\bar{w},\bar{v})$为$(LM)$的解且矩阵$\nabla_x^2 L(\bar{x},\bar{w},\bar{v})$在子空间$\bar{G}$上是半正定的，其中

$$\bar{G}=\left\{ d \,\middle|\, \begin{array}{l} \nabla g_i(\bar{x})^T d=0, i\in I \text{ 且 } \bar{w}_i>0 \\ \nabla g_i(\bar{x})^T d\ge 0, i\in I \text{ 且 } \bar{w}_i=0 \\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l \end{array} \right\}.$$

定理（二阶充分条件）：设f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$是二次连续可微函数，$\overline{x}$为可行解，若存在$\overline{w},\overline{v}$，使$(\overline{x},\overline{w},\overline{v})$为$(LM)$的解且矩阵$\nabla_x^2 L(\overline{x},\overline{w},\overline{v})$在子空间$\overline{G}$上是正定的，则$\overline{x}$是严格局部极小点。

其中
$$\overline{G}=\left\{d\left|\begin{array}{l}\nabla g_i(\overline{x})^T d=0, i\in I\text{且}\overline{w}_i>0\\ \nabla g_i(\overline{x})^T d\geq 0, i\in I\text{且}\overline{w}_i=0\\ \nabla h_j(\overline{x})^T d=0, j=1,2,\cdots,l\end{array}\right.\right\}.$$

证明： 假设 $\bar{x}$ 不是问题的严格局部极小点, 则存在序列 $\{x^{(k)}\}\subset S$, 使得 $x^{(k)}\neq\bar{x}, x^{(k)}\to\bar{x}$, 并且 $f(x^{(k)})\leq f(\bar{x})$, 即

$$f(x^{(k)})-f(\bar{x})=\nabla f(\bar{x})^T(x^{(k)}-\bar{x})+o(\|x^{(k)}-\bar{x}\|)\leq 0.$$

记 $$d^{(k)}=\frac{x^{(k)}-\bar{x}}{\|x^{(k)}-\bar{x}\|},$$

显然, $\forall k\geq 1,\quad \|d^{(k)}\|\equiv 1.$

根据有界序列的性质可知, 存在一个收敛子列 $\{d^{(k_i)}\}\to d\neq 0$

不妨假设 $d^{(k)}\to d$ 则， $d\in\bar{T}(\bar{x},S)$

因为 $\nabla f(\bar{x})^T d^k+o(1)\leq 0$

令 $k\to\infty$

于是有 $d^T\nabla f(\bar{x})\leq 0$

以下证明 $\nabla f(\bar{x})d^T=0$

将$g_i(x)$在点$\bar{x}$展开，再令$x=x^{(k)}$,得到

$$g_i(x^{(k)})=g_i(\bar{x})+\nabla g_i(\bar{x})^T(x^{(k)}-\bar{x})+o(\|x^{(k)}-\bar{x}\|)$$

$$\Rightarrow\quad \nabla g_i(\bar{x})^T(x^{(k)}-\bar{x})+o(\|x^{(k)}-\bar{x}\|)\geq 0\ (i\in I)$$

$$\Rightarrow\quad \nabla g_i(\bar{x})^T d\geq 0\quad (i\in I)$$

同理可证 $\nabla h_j(\bar{x})^T d=0,\quad j=1,2,\cdots,l$

利用**KKT**条件, 得到

$$\nabla f(\bar{x})^T d=\sum_{i\in I}\bar{w}_i\nabla g_i(\bar{x})^T d+\sum_{j=1}^{l}\bar{v}_j\nabla h(\bar{x})^T d\geq 0,$$

因此, $\nabla f(\bar{x})^T d=0.$

由于 $\nabla f(\bar{x})^T d=\sum_{i\in I}\bar{w}_i\nabla g_i(\bar{x})^T d=0$

$$\therefore \nabla g_i(\bar{x})^T d=0,\quad i\in I,\bar{w}_i>0\qquad\Rightarrow\quad d\in\overline{G}.$$

$\because\ x^{(k)} \in S$ 由**Taylor**展开公式, 有

$$L(\bar{x}, \bar{w}, \bar{v}) = f(\bar{x}) \geq f(x^{(k)})$$

$$\geq f(x^{(k)}) - \sum_{i \in I} \bar{w}_i g_i (x^{(k)})^T - \sum_{j=1}^{l} \bar{v}_j h_j (x^{(k)})^T = L(x^{(k)}, \bar{w}, \bar{v})$$

而$L(x^{(k)}, \bar{w}, \bar{v}) = L(\bar{x}, \bar{w}, \bar{v}) + \nabla_x L(\bar{x}, \bar{w}, \bar{v})^T (x^{(k)} - \bar{x})$

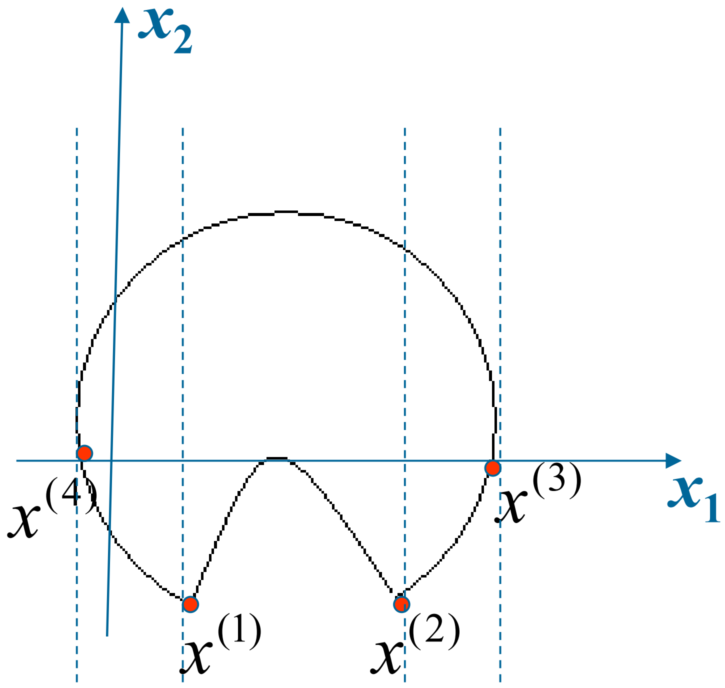
$$+\frac{1}{2}(x^{(k)} - \bar{x})^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v})(x^{(k)} - \bar{x}) + o(\| (x^{(k)} - \bar{x}) \|^2).$$

$$\therefore \frac{1}{2}(x^{(k)} - \bar{x})^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v})(x^{(k)} - \bar{x}) + o(\| (x^{(k)} - \bar{x}) \|^2) \leq 0.$$

$$\therefore \quad d^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v}) d \leq 0.$$

矛盾。

例:考虑下列非线性规划问题:



$$\begin{cases} \min x_1 \\ s.t. \quad 3(x_1-3)^2+x_2 \geq 0 \\ \qquad (x_1-3)^2+x_2^2-10=0 \end{cases}$$

检验$x^{(1)}=\begin{pmatrix} 2 \\ -3 \end{pmatrix}, x^{(2)}=\begin{pmatrix} 4 \\ -3 \end{pmatrix}, x^{(3)}=\begin{pmatrix} 3+\sqrt{10} \\ 0 \end{pmatrix},$

$x^{(4)}=\begin{pmatrix} 3-\sqrt{10} \\ 0 \end{pmatrix}$是否为局部最优解。

$$\begin{cases} \min x_1 \\ s.t. \quad 3(x_1-3)^2+x_2 \geq 0 \\ \qquad\quad (x_1-3)^2+x_2^2-10=0 \end{cases}$$

解：$\nabla f(x)=(1,0)^T$

$$\nabla g(x)=(6(x_1-3),1)^T, \nabla h(x)=(2(x_1-3),2x_2)^T$$

$$L(x,w,v)=f(x)-wg(x)-vh(x)$$

$$\nabla_x L=\begin{pmatrix}1\\0\end{pmatrix}-w\begin{pmatrix}6(x_1-3)\\1\end{pmatrix}-v\begin{pmatrix}2(x_1-3)\\2x_2\end{pmatrix}$$

$$\nabla_x^2 L=\begin{pmatrix}-6w-2v & 0\\0 & -2v\end{pmatrix}$$

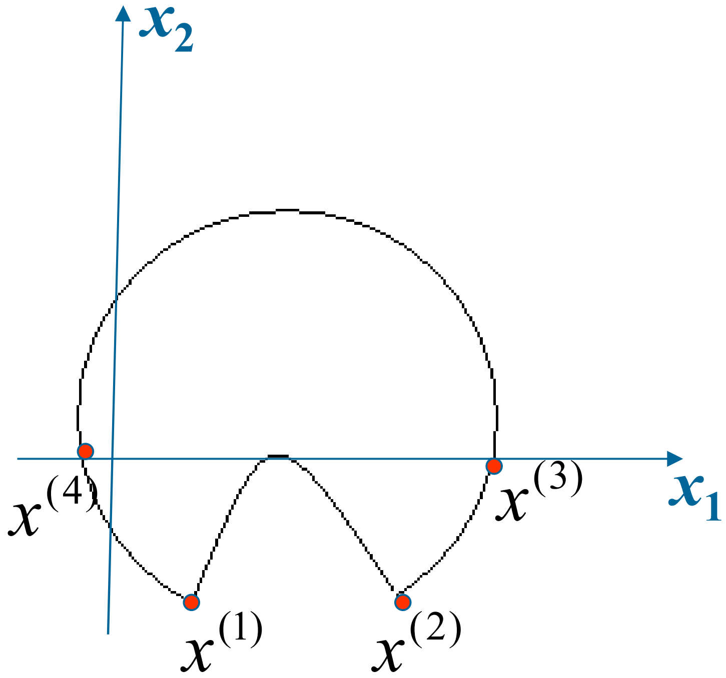
(1) $x^{(1)}=(2,-3)^T$是可行点，且$g(x)$是紧约束，设有

$$\begin{pmatrix}1\\0\end{pmatrix}-w\begin{pmatrix}6(2-3)\\1\end{pmatrix}-v\begin{pmatrix}2(2-3)\\2(-3)\end{pmatrix}=0$$

$$\Rightarrow v=-\frac{1}{38},\ w=-\frac{3}{19}<0$$

$\therefore x^{(1)}$不是KKT点，因此$x^{(1)}$一定不是局部最优解。

$$\begin{cases}\min x_1\\ s.t.\quad 3(x_1-3)^2+x_2\geq 0\\ \qquad\quad (x_1-3)^2+x_2^2-10=0\end{cases}$$



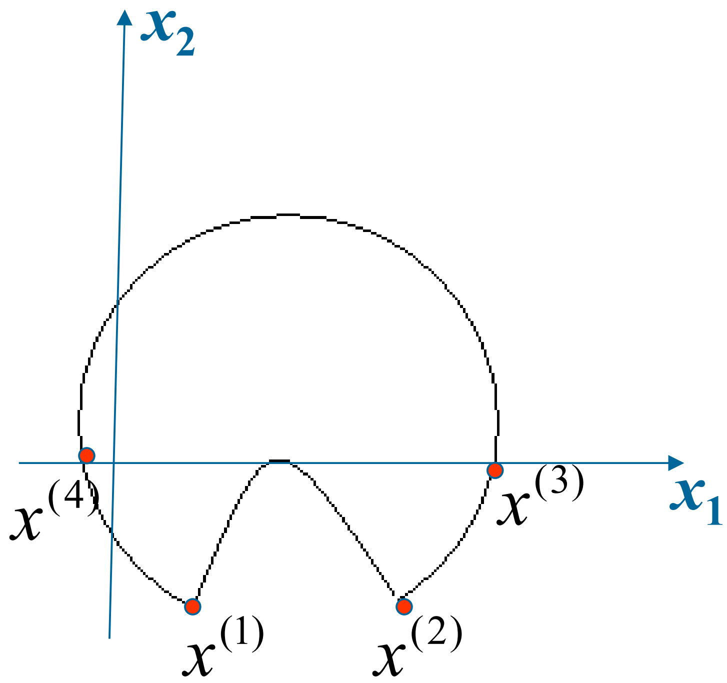
(2) $x^{(2)}=(4,-3)^T$是可行点，且$g(x)$是紧约束，设有

$$\begin{pmatrix}1\\0\end{pmatrix}-w\begin{pmatrix}6(4-3)\\1\end{pmatrix}-v\begin{pmatrix}2(4-3)\\2(-3)\end{pmatrix}=0$$

$$\Rightarrow v=\frac{1}{38},\ w=\frac{3}{19}\ \therefore x^{(2)}\text{是}KKT\text{点.}$$

$$\nabla_x^2 L\left(x,\frac{3}{19},\frac{1}{38}\right)=\begin{pmatrix}-1 & 0\\ 0 & -\frac{1}{19}\end{pmatrix}$$

但$\overline{G}=\{0\}$， $\therefore x^{(2)}$是局部极小点。



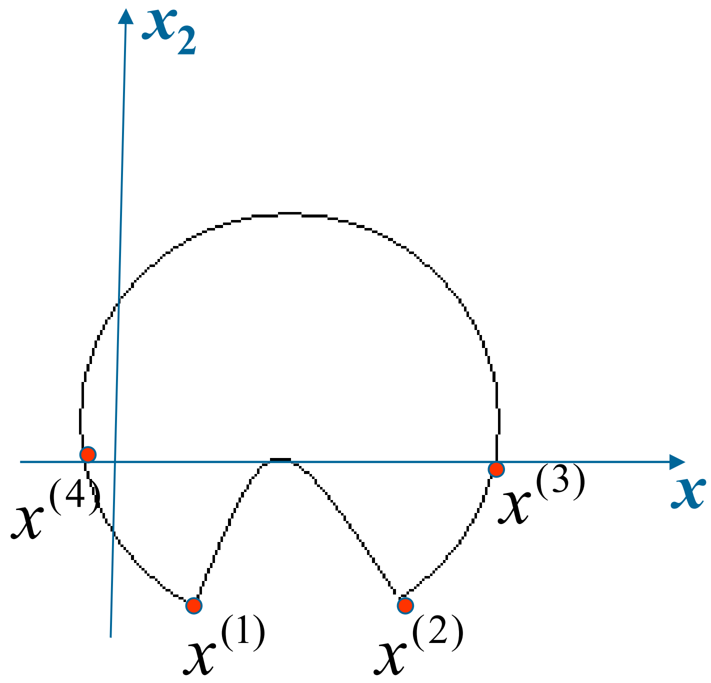
$$\begin{cases}\min x_1\\ s.t.\quad 3(x_1-3)^2+x_2\ge 0\\ \qquad\ (x_1-3)^2+x_2^2-10=0\end{cases}$$

(3) $x^{(3)}=(3+\sqrt{10},0)^T$是可行点，且$g(x)$是不起作用约束，

设有 $\begin{pmatrix}1\\0\end{pmatrix}-v\begin{pmatrix}2(3+\sqrt{10}-3)\\0\end{pmatrix}=0$

$\Rightarrow v=\dfrac{1}{2\sqrt{10}}\ \therefore x^{(3)}$是$KKT$点.

$$\nabla_x^2 L\left(x,0,\frac{1}{2\sqrt{10}}\right)=\begin{pmatrix}-\frac{1}{\sqrt{10}} & 0\\ 0 & -\frac{1}{\sqrt{10}}\end{pmatrix}$$



但$\overline{G}=\left\{d=(0,d_2)^T\mid d_2\in R\right\}$， 此时对$d\in\overline{G}$且$d\neq 0$,

$$d^T\nabla_x^2 L\left(x,0,\frac{1}{2\sqrt{10}}\right)d=-\frac{1}{\sqrt{10}}d_2^2<0$$

$\therefore x^{(3)}$不是局部极小点。

(4) $x^{(4)}=(3-\sqrt{10},0)^T$是可行点，且$g(x)$是不起作用约束，

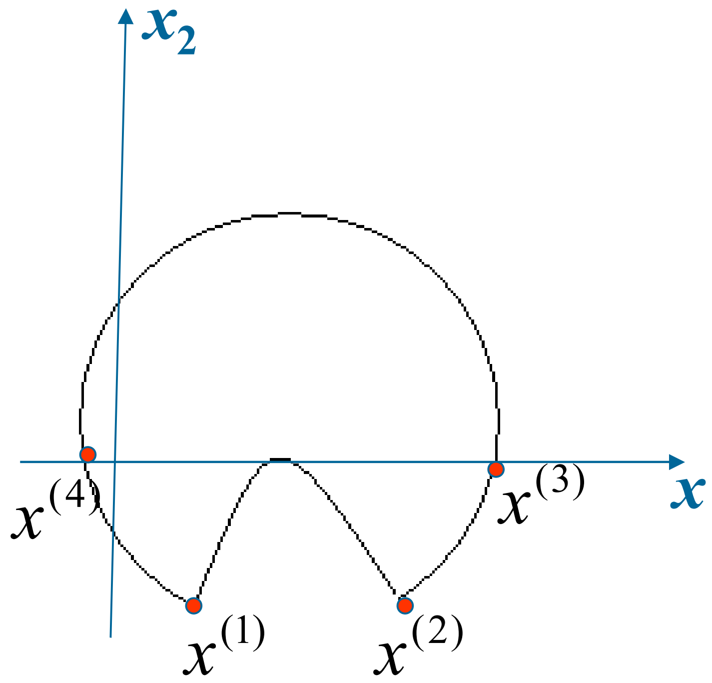
设有 $\begin{pmatrix}1\\0\end{pmatrix}-v\begin{pmatrix}2(3-\sqrt{10}-3)\\0\end{pmatrix}=0$

$\Rightarrow v=-\dfrac{1}{2\sqrt{10}}$ $\therefore x^{(4)}$是KKT点.

$$\nabla_x^2 L\left(x,0,-\frac{1}{2\sqrt{10}}\right)=\begin{pmatrix}\frac{1}{\sqrt{10}} & 0\\ 0 & \frac{1}{\sqrt{10}}\end{pmatrix}$$

是正定矩阵，

$\therefore x^{(4)}$是局部极小点。



例：考虑下列非线性规划问题

$$\min\ x_1^2+(x_2-2)^2$$
$$s.t.\quad \beta x_1^2-x_2=0$$

其中β为某个实数，讨论点$x^{(0)}=(0,0)^T$是否为局部最优解？

解： $L(x,v)=x_1^2+(x_2-2)^2-v(\beta x_1^2-x_2)$

$$\nabla_x L(x,v)=\begin{pmatrix}2x_1-v\cdot 2\beta x_1\\ 2(x_2-2)+v\end{pmatrix},\quad \nabla_x^2 L(x,v)=\begin{pmatrix}2-2v\beta & 0\\ 0 & 2\end{pmatrix}.$$

设 $\nabla_x L(x^{(0)},v)=0$，得$v=4$

$\therefore\quad x^{(0)}$是KKT点，且

$$\nabla_x^2 L(x,v)|_{v=4}=\begin{pmatrix}2-8\beta & 0\\ 0 & 2\end{pmatrix}.$$

求集合$\overline{G}$的元素$d=(d_1,d_2)^T$，令

$$\nabla h(x^{(0)})^T d=0 \quad \Rightarrow \quad d_2=0$$

$\therefore \quad \overline{G}=\left\{(d_1,0)^T \middle| d_1 \in R\right\}$.此时对$d\in\overline{G}\setminus\{0\}$,

$$d^T\nabla_x^2 L(x^{(0)},4)d=(2-8\beta)d_1^2$$

$\therefore$当$2-8\beta>0$，即$\beta<\dfrac{1}{4}$时，$x^{(0)}$是局部最优解。

当$2-8\beta<0$，即$\beta>\dfrac{1}{4}$时,对$\forall d\in\overline{G}\setminus\{0\}$,

$$d^T\nabla_x^2 L(x^{(0)},4)d<0,$$

所以，$x^{(0)}$不是局部最优解。

当$\beta=\dfrac{1}{4}$时，原问题为：

$$\min\ x_1^2+(x_2-2)^2$$
$$s.t.\quad \frac{1}{4}x_1^2-x_2=0$$

由$\dfrac{1}{4}x_1^2-x_2=0 \quad \Rightarrow x_1^2=4x_2$

$\therefore$ 原问题变为： $\min\ 4x_2+(x_2-2)^2$.

显然$(0,0)^T$是驻点，且$f''(x_2)=2>0$,

$\therefore x^{(0)}=(0,0)^T$是局部极小点。

即当$\beta\le\dfrac{1}{4}$时，$x^{(0)}=(0,0)^T$是局部极小点；

当$\beta>\dfrac{1}{4}$时，$x^{(0)}=(0,0)^T$不是局部极小点

对偶及鞍点问题

Lagrange对偶问题

$$\min\ f(x)$$
$$s.t.\quad g_i(x)\ge 0,\ i=1,\cdots,m$$
$$h_j(x)=0,\ j=1,\cdots,l \qquad \textbf{(1)}$$
$$x\in D$$

集约束

定义(1)的对偶问题:

$$\max\ \theta(w,v)$$
$$s.t.\quad w\ge 0 \qquad \textbf{(2)}$$

其中$\theta(w,v)=\inf\left\{f(x)-\sum_{i=1}^{m}w_i g_i(x)-\sum_{j=1}^{l}v_j h_j(x)\,\middle|\,x\in D\right\}$

若上式不存在有限下界时，令$\theta(w,v)=-\infty.$

$\theta(w,v)$称为$Lagrange$对偶函数。

$$\max \theta(w,v)$$
$$s.t. \quad w \geq 0$$

其中$\theta(w,v)=\inf\left\{ f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x) \middle| x\in D \right\}$

$$L(x,w,v)=f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x)$$

Lagrange函数

对于任意的$x\in D$，$Lagrangr$函数$L(x,w,v)$是w,v的线性函数，于是对偶函数$\theta(w,v)$作为线性函数的逐点下确界，必然是一个凹函数，所以，对偶问题是一个凸规划问题。

例：考虑线性规划问题

$$
\begin{aligned}
&\min\ cx \\
&s.t.\quad A_1 x \geq b_1 \\
&\qquad\ \ A_2 x = b_2 \\
&\qquad\ \ x \geq 0
\end{aligned}
$$

若取集合约束$\boldsymbol{D=\{x|x\geq 0\}}$，则该线性规划问题的**Lagrange**函数为

$$
\begin{aligned}
\theta(w,v) &= \inf\left\{cx - w^T(A_1 x - b_1) - v^T(A_2 x - b_2) \mid x \in D\right\} \\
&= \inf\{(c - w^T A_1 - v^T A_2)x + w^T b_1 + v^T b_2 \mid x \in D\} \\
&= \begin{cases} w^T b_1 + v^T b_2 & \text{若} c - w^T A_1 - v^T A_2 \geq 0 \\ -\infty & \text{若} c - w^T A_1 - v^T A_2 \not\geq 0. \end{cases}
\end{aligned}
$$

线性规划的对偶问题为：

$$
\begin{aligned}
&\max\ w^T b_1 + v^T b_2 \\
&s.t.\quad w^T A_1 + v^T A_2 \leq c \\
&\qquad\ \ w \geq 0
\end{aligned}
$$

求下列非线性规划问题的对偶问题：

$$\begin{cases}\min\ x_1^2+x_2^2 \\ s.t.\quad x_1+x_2-4\geq 0 \\ \qquad\quad x_1,x_2\geq 0\end{cases}$$

解1：把变量的非负限制作为集约束，即

$$x\in D=\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\middle| x_1\geq 0, x_2\geq 0\right\},$$

则 $\theta(w)=\inf\left\{x_1^2+x_2^2-w(x_1+x_2-4)\mid x\in D\right\}.$

$$\begin{aligned}\theta(w) &= \inf\left\{x_1^2 + x_2^2 - w(x_1+x_2-4) \mid x \in D\right\} \\ &= \inf\left\{x_1^2 - wx_1 + x_2^2 - wx_2 + 4w \mid x_1 \geq 0, x_2 \geq 0\right\} \\ &= \inf\left\{x_1^2 - wx_1 \mid x_1 \geq 0\right\} + \inf\left\{x_2^2 - wx_2 \mid x_2 \geq 0\right\} + 4w\end{aligned}$$

当$w \geq 0$时,

$$\inf\left\{x_1^2 - wx_1 \mid x_1 \geq 0\right\} = \left(\frac{w}{2}\right)^2 - w \times \frac{w}{2} = -\frac{w^2}{4}.$$

$$\inf\left\{x_2^2 - wx_2 \mid x_2 \geq 0\right\} = \left(\frac{w}{2}\right)^2 - w \times \frac{w}{2} = -\frac{w^2}{4}.$$

$$\therefore \theta(w) = -\frac{w^2}{4} - \frac{w^2}{4} + 4w = -\frac{w^2}{2} + 4w.$$

对偶问题为:

$$\begin{cases} \max -\dfrac{w^2}{2} + 4w \\ s.t. \quad w \geq 0 \end{cases}$$

求下列非线性规划问题的对偶问题:

$$\begin{cases} \min\ x_1^2 + x_2^2 \\ s.t. \quad x_1 + x_2 - 4 \geq 0 \\ \qquad\quad x_1, x_2 \geq 0 \end{cases}$$

解2：若取集约束$D = E^n$,则

$$\theta(w) = \inf\left\{x_1^2 + x_2^2 - w_1(x_1+x_2-4) - w_2x_1 - w_3x_2 \mid x \in E^n\right\}.$$

$$\theta(w)=\inf\left\{x_1^2+x_2^2-w_1(x_1+x_2-4)-w_2x_1-w_3x_2 \mid x\in E^n\right\}.$$

$$=\inf\left\{x_1^2-w_1x_1-w_2x_1\right\}+\inf\left\{x_2^2-w_1x_2-w_3x_2\right\}+4w_1$$

当$w_i\geq 0$时，有

$$\theta(w_1,w_2,w_3)=\left(\frac{w_1+w_2}{2}\right)^2-w_1\times\frac{w_1+w_2}{2}-w_2\times\frac{w_1+w_2}{2}$$

$$+\left(\frac{w_1+w_3}{2}\right)^2-w_1\times\frac{w_1+w_3}{2}-w_3\times\frac{w_1+w_3}{2}+4w_1$$

$$=-\frac{(w_1+w_2)^2}{4}-\frac{(w_1+w_3)^2}{4}+4w_1$$

对偶问题为: $\begin{cases}\max -\dfrac{(w_1+w_2)^2}{4}-\dfrac{(w_1+w_3)^2}{4}+4w_1\\ s.t.\quad w_i\geq 0,\ i=1,2,3.\end{cases}$

求下列非线性规划问题的对偶问题:

$$\begin{cases} \min x_1^2 + x_2^2 \\ s.t. \quad x_1 + x_2 - 4 \geq 0 \\ \qquad\quad x_1, x_2 \geq 0 \end{cases}$$

解3：若集约束为

$$x \in D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_1 + x_2 - 4 \geq 0 \right\},$$

则 $\theta(w_1, w_2) = \inf\left\{ x_1^2 + x_2^2 - w_1 x_1 - w_2 x_2 \mid x \in D \right\}.$

$$\theta(w_1, w_2) = \inf\left\{x_1^2 + x_2^2 - w_1 x_1 - w_2 x_2 \mid x \in D\right\}$$

令$f(x) = x_1^2 + x_2^2 - w_1 x_1 - w_2 x_2, g(x) = x_1 + x_2 - 4.$

当$w_1 + w_2 - 8 < 0$时，$f(x)$与$g(x)$相切时达最小

$$\Rightarrow \begin{pmatrix} 2x_1 - w_1 \\ 2x_2 - w_2 \end{pmatrix} = k\begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow x_1 = \frac{w_1 + k}{2}, x_2 = \frac{w_2 + k}{2}$$

$\because x_1, x_2$满足$x_1 + x_2 - 4 = 0$，

$$\therefore \text{有} \frac{w_1 + k}{2} + \frac{w_2 + k}{2} - 4 = 0 \quad \Rightarrow k = \frac{8 - (w_1 + w_2)}{2}$$

$$\therefore \quad x_1 = 2 + \frac{w_1 - w_2}{4}, x_2 = 2 + \frac{w_2 - w_1}{4}$$

$$\Rightarrow \theta(w_1, w_2) = 8 - 2w_1 - 2w_2 - \frac{(w_1 - w_2)^2}{8}$$

$$\theta(w_1,w_2)=\inf\left\{x_1^2+x_2^2-w_1x_1-w_2x_2 \mid x\in D\right\}$$

令$f(x)=x_1^2+x_2^2-w_1x_1-w_2x_2$

$$=\left(x_1-\frac{w_1}{2}\right)^2+\left(x_2-\frac{w_2}{2}\right)^2-\frac{w_1^2+w_2^2}{4},$$

$g(x)=x_1+x_2-4.$

当$w_1+w_2-8\geq 0$时，$f(x)$在$x_1=\dfrac{w_1}{2},x_2=\dfrac{w_2}{2}$时达最小

$$\Rightarrow \theta(w_1,w_2)=-\frac{w_1^2}{4}-\frac{w_2^2}{4}.$$

$$\therefore \theta(w_1,w_2)=\begin{cases} 8-2w_1-2w_2-\dfrac{(w_1-w_2)^2}{8} & w_1+w_2<8 \\ -\dfrac{w_1^2}{4}-\dfrac{w_2^2}{4} & w_1+w_2\geq 8 \end{cases}$$

$$\begin{cases} \min\ x_1^2 + x_2^2 \\ s.t. \quad x_1 + x_2 - 4 \geq 0 \\ \qquad\quad x_1, x_2 \geq 0 \end{cases}$$

解3：若集约束为

$$x \in D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_1 + x_2 - 4 \geq 0 \right\},$$

则对偶问题为：

$$\max\ 8 - 2w_1 - 2w_2 - \frac{(w_1 - w_2)^2}{8}$$

$$s.t. \quad w_1 + w_2 < 8$$

$$w_1, w_2 \geq 0$$

对偶定理

$$
\begin{aligned}
\min\ & f(x) \\
s.t.\ \ & g(x) \geq 0 \\
& h(x) = 0 \\
& x \in D
\end{aligned}
$$

$$
\begin{aligned}
\max\ & \theta(w, v) \\
s.t.\ \ & w \geq 0
\end{aligned}
$$

$$g(x) = \left(g_1(x), \cdots, g_m(x)\right)^T$$

$$h(x) = \left(h_1(x), \cdots, h_l(x)\right)^T$$

$$\theta(w, v) = \inf\left\{ f(x) - w^T g(x) - v^T h(x) \mid x \in D \right\}$$

定理1(弱对偶定理)

设$\bar{x}$和$(\bar{w},\bar{v})$分别是原问题和对偶问题的可行解，则

$$f(\bar{x}) \geq \theta(\bar{w},\bar{v})。$$

证明：$\because$ $\bar{x}$和$(\bar{w},\bar{v})$是可行解，

$$\therefore \quad g(\bar{x}) \geq 0, h(\bar{x}) = 0, \bar{w} \geq 0$$

$$\begin{aligned}\therefore \theta(\bar{w},\bar{v}) &= \inf\left\{ f(x) - \bar{w}^T g(x) - \bar{v}^T h(x) \mid x \in D \right\} \\ &\leq f(\bar{x}) - \bar{w}^T g(\bar{x}) - \bar{v}^T h(\bar{x}) \\ &\leq f(\bar{x}).\end{aligned}$$

推论1: 对于原问题和对偶问题，必有

$$\inf\{f(x) \mid g(x) \geq 0, h(x) = 0, x \in D\} \geq \sup\{\theta(w,v) \mid w \geq 0\}.$$

推论2: 若$f(\bar{x}) \leq \theta(\bar{w},\bar{v})$，其中$\bar{x}$为原问题的可行解，$\bar{w} \geq 0$，则$\bar{x}$和$(\bar{w},\bar{v})$分别是原问题和对偶问题的最优解。

推论3: 若$\inf\{f(x) \mid g(x) \geq 0, h(x) = 0, x \in D\} = -\infty$，则对$\forall w \geq 0$，有$\theta(w,v) = -\infty$。

推论4: 如果$\sup\{\theta(w,v) \mid w \geq 0\} = +\infty$，则原问题没有可行解。

$$\inf\{f(x) \mid g(x) \geq 0, h(x) = 0, x \in D\} \underset{\text{记}}{=} f_{\min}$$

$$\sup\{\theta(w,v) \mid w \geq 0\} \underset{\text{记}}{=} \theta_{\max}$$

对偶间隙： $\delta = f_{\min} - \theta_{\max} \geq 0$

问题： $\delta = 0$成立的条件.

引理 设D为非空的凸集，$f: D \rightarrow R$ 为凸函数，
$g_i: D \rightarrow R$ 为凹函数， $h(x)$ 为线性函数.

对于下面的两个不等式系统：

(1) 存在 $x \in D$，使得

$$f(x) < 0, \quad g(x) \geq 0, \quad h(x) = 0;$$

(2) 存在$(\lambda_0, \lambda^T, \mu^T) \neq 0$，使得$(\lambda_0, \lambda^T) \geq 0$，并且

$$\lambda_0 f(x) - \lambda^T g(x) - \mu^T h(x) \geq 0 \quad (\forall x \in D).$$

若系统(1)无解，则系统(2)有解；若系统(2)有满足 $\lambda_0 > 0$ 的解，则系统(1)无解.

证明 先证结论的第一部分: (1)无解 $\Rightarrow$ (2)有解.

令集合 $C=\{(p,q^T,r^T)\mid \exists x\in D,\ p>f(x),\ g(x)\ge q,\ h(x)=r\}$.

根据函数的凸(凹)性和凸集的定义, 易证 C 是非空的凸集.

若(1)无解, 则 $(0,0,0)\notin C$. 由凸集分离定理可知, 存在

$(\lambda_0,\lambda^T,\mu^T)\neq 0$, 使得 $\forall (p,q^T,r^T)\in cl(C)$,

$$\lambda_0 p-\lambda^T q-\mu^T r\ge 0.$$

分别令 $p\to+\infty,\ q\to-\infty$

于是有 $\lambda_0\ge 0,\ \lambda\ge 0.$

特别地, 上式对于

$$p=f(x),\ q=g(x),\ r=h(x)(\forall x\in D)$$

也成立, 因此(2)有解.

(2)有解 $\Rightarrow$ (1)无解.

假设(2)有满足 $\lambda_0 > 0$的解, 由于

$$\forall x \in D, \lambda_0 f(x) \geq \lambda^T g(x) + \mu^T h(x)$$

所以对于任何满足 $g(\bar{x}) \geq 0, h(\bar{x}) = 0$

的点 $\bar{x} \in D$

利用 $\lambda \geq 0$

(2) 存在$(\lambda_0, \lambda^T, \mu^T) \neq 0$,使得$(\lambda_0, \lambda^T) \geq 0$, 并且

$$\lambda_0 f(x) - \lambda^T g(x) - \mu^T h(x) \geq 0 \quad (\forall x \in D).$$

可知, $\lambda_0 f(\bar{x}) \geq 0$

也就是说, $f(\bar{x}) \geq 0$

存在 $x \in D$, 使得

$$f(x) < 0, \quad g(x) \geq 0, \quad h(x) = 0;$$

因此, 系统(1)无解.

强对偶定理：

设D为非空开凸集，f和$g_i(i=1,\cdots,m)$分别是E^n上的凸函数和凹函数，$h_j(j=1,\cdots,l)$是E^n上的线性函数，即$h(x)=Ax-b$。又设存在$\hat{x}\in D$，使得

$$g(\hat{x})>0,\ h(\hat{x})=0,\ 0\in \operatorname{int} H(D)$$

其中$H(D)=\{h(x)\,|\,x\in D\}$，则

$$f_{\min}\overset{def}{=}\inf\{f(x)\,|\,g(x)\geq 0,\ h(x)=0,\ x\in D\}=\sup\{\theta(w,v)\,|\,w\geq 0\}\overset{def}{=}\theta_{\max};$$

而且，若inf为有限值，则存在某个$(\bar{w},\bar{v})$,使得

$$\sup\{\theta(w,v)\,|\,w\geq 0\}$$

在$(\bar{w},\bar{v})$达到，$\bar{w}\geq 0$；如果inf在点$\bar{x}$达到，则$\bar{w}^T g(\bar{x})=0.$

证明：根据弱对偶定理的推论3，不妨假设$f_{\min} > -\infty$.
则，不等式组

$$f(x) - f_{\min} < 0, g(x) \geq 0, h(x) = 0$$

在非空凸集D内无解. 由引理，存在$(w_0, w, v) \neq 0$,
$(w_0, w) \geq 0$,使得对$\forall x \in D$,有

$$w_0\left(f(x) - f_{\min}\right) - w^T g(x) - v^T h(x) \geq 0.$$

若$w_0 = 0$，则$w^T g(x) + v^T h(x) \leq 0$

由假设，存在$\hat{x} \in D$使得$g(\hat{x}) > 0, h(\hat{x}) = 0 \quad \Rightarrow w = 0$

$\Rightarrow v^T h(x) \leq 0 \quad \forall x \in D$

$\because 0 \in \operatorname{int} H(D), \therefore$ 可取$x \in D$和$\rho > 0$,使得$h(x) = \rho v$

$\Rightarrow \quad 0 \geq v^T h(x) = v^T \rho v = \rho \| v \|^2 \quad \Rightarrow v = 0$

$\Rightarrow (w_0, w, v) = 0$，矛盾。所以$w_0 \neq 0$。

令 $\bar{w}=\dfrac{w}{w_0}$， $\bar{v}=\dfrac{v}{w_0}$，则对任意$x\in D$

$$f(x)-\bar{w}^T g(x)-\bar{v}^T h(x)\geq f_{\min}$$

$$\therefore \theta\left(\bar{w},\bar{v}\right)=\inf\{f(x)-\bar{w}^T g(x)-\bar{v}^T h(x)\mid x\in D\}\geq f_{\min},$$

由弱对偶定理的推论2，知$\left(\bar{w},\bar{v}\right)$为对偶问题的最优解，

并且$\theta_{\max}=\theta\left(\bar{w},\bar{v}\right)=f_{\min}$.

若$\exists\bar{x}\in D, f(\bar{x})=f_{\min}$，则利用$\bar{w}\geq 0, g(\bar{x})\geq 0, h(\bar{x})=0$

$\Rightarrow \quad \bar{w}^T g(\bar{x})=0$

考虑如下约束优化问题

$$\min\ f(x)=\frac{1}{x}$$
$$s.t.\quad g(x)=x-1\geq 0$$
$$x\in D=(0,+\infty)$$

该问题没有最优解，但目标函数的下确界 $f_{\min}=0.$

该问题的**Lagrange**函数为

$$\theta(w)=\inf\left\{\frac{1}{x}-w(x-1)\mid x\in D\right\}=\begin{cases}-\infty & w>0\\ 0 & w=0\end{cases}$$

对偶问题 $\max\{\theta(w)\mid w\geq 0\}$ 的最优解为 $\bar{w}=0$,

最优值 $\theta_{\max}=0=f_{\min}$。

对偶规划的意义

1. 利用对偶可将一个约束优化问题转化成另一个约束优化问题并建立两问题解之间的某种联系，进而揭示原规划问题最优解的存在性等理论性质，如对偶规划不但可以给出原规划问题最优值的一个下界，而且在特殊情况下可得到原规划问题的最优值.

2. 基于对偶规划还可建立原规划问题的对偶类算法.

鞍点最优性条件

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ \qquad\ \ h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

$$g(x)=(g_1(x),\cdots,g_m(x))^T$$

$$h(x)=(h_1(x),\cdots,h_l(x))^T$$

Lagrange
函数

$$L(x,w,v)=f(x)-w^Tg(x)-v^Th(x)$$

$$\max\ \theta(w,v)$$

$$s.t.\quad w\geq 0$$

$$\theta(w,v)=\inf\left\{f(x)-w^Tg(x)-v^Th(x)\mid x\in E^n\right\}$$

约束优化问题鞍点的引出

$$\min\ f(\boldsymbol{x})$$
$$\text{s.t.}\quad c_i(\boldsymbol{x})=0,\quad i\in\mathcal{E}$$

约束函数梯度线性无关

$$\begin{cases}\nabla_{\boldsymbol{x}}L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)=\boldsymbol{0},\\ \nabla_{\boldsymbol{\lambda}}L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)=\boldsymbol{c}(\boldsymbol{x}^*)=\boldsymbol{0}.\end{cases}$$

(x^*,λ^*)构成Lagrange函数的驻点

其中，$\boldsymbol{x}^*$最优解，$\boldsymbol{\lambda}^*$为最优Lagrange乘子。

$(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$是哪种类型驻点？

局部最优点？鞍点？

可以证明的结论

- 线性等式约束的凸规划问题

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}) \\
\text{s.t.} \quad & \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} + b_i = 0, \quad i \in \mathcal{E}
\end{aligned}
$$

KKT对 $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ 中的 $\boldsymbol{x}^*$ 为Lagrange函数在 $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ 时关于 $\boldsymbol{x}$ 的最小值点

$\boldsymbol{\lambda}^*$ 为Lagrange函数在 $\boldsymbol{x} = \boldsymbol{x}^*$ 时关于 $\boldsymbol{\lambda}$ 的最大值点。

定义： 设$L(x,w,v)$为$Lagrange$函数，$\bar{x}\in E^n$，$\bar{w}\in E^m$，$\bar{w}\geq 0$，$\bar{v}\in E^l$，如果对任意$x\in E^n$，$w\in E^m$，$w\geq 0$及$v\in E^l$，都有

$$L(\bar{x},w,v)\leq L(\bar{x},\bar{w},\bar{v})\leq L(x,\bar{w},\bar{v})$$

则称$(\bar{x},\bar{w},\bar{v})$为$L(x,w,v)$的鞍点。

结论：

Lagrange函数的鞍点必是Lagrange函数关于x的极小点及关于$(w, v)(w\geq 0)$的极大点.

$$\begin{cases} \min\ f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \\ s.t. \quad g_1(x) = -x_1 - x_2 + 2 \geq 0 \\ \qquad\ \ g_2(x) = x_1 \geq 0 \qquad g_3(x) = x_2 \geq 0 \\ \qquad\ \ h(x) = x_2 - x_1 - 1 = 0 \end{cases}$$

$$L(x, w, v) = f(x) - w_1 g_1(x) - w_2 g_2(x) - w_3 g_3(x) - vh(x)$$

鞍点为： $\bar{x} = \left(\frac{1}{2}, \frac{3}{2}\right)^T, \bar{w} = (0, 0, 0)^T, \bar{v} = 1$

$$L(\bar{x}, w, v) = \frac{1}{2} - \frac{1}{2} w_2 - \frac{3}{2} w_3 \quad L(\bar{x}, \bar{w}, \bar{v}) = \frac{1}{2}$$

$$L(x, \bar{w}, \bar{v}) = (x_1 - 1)^2 + (x_2 - 1)^2 - (x_2 - x_1 - 1)$$

鞍点定理:

(1)设$(\overline{x},\overline{w},\overline{v})$是原问题的$Lagrange$函数$L(x,w,v)$的鞍点，则$\overline{x}$和$(\overline{w},\overline{v})$分别是原问题和对偶问题的最优解。(2)假设$f$是凸函数，$g_i(x)(i=1,\cdots,m)$是凹函数，$h_j(x)(j=1,\cdots,l)$是线性函数，即$h(x)=Ax-b$，且$A$是行满秩矩阵，又设存在$\hat{x}$，使$g(\hat{x})>0$，$h(\hat{x})=0$，如果$\overline{x}$是原问题的最优解，则存在$(\overline{w},\overline{v})(\overline{w}\geq 0)$，使$(\overline{x},\overline{w},\overline{v})$是$Lagrange$函数的鞍点。

证明：（1）设$\left(\overline{x}, \overline{w}, \overline{v}\right)$为$Lagrange$函数的鞍点，

则对所有$w \in R^m, w \geq 0, v \in R^l$,有

$$L\left(\overline{x}, w, v\right) \leq L\left(\overline{x}, \overline{w}, \overline{v}\right)$$

$$\Rightarrow \quad \left(w - \overline{w}\right)^T g(\overline{x}) + \left(v - \overline{v}\right)^T h(\overline{x}) \geq 0.$$

若取$w = \overline{w}, v = \overline{v} + e^j$,有$h_j(\overline{x}) \geq 0,$

若取$w = \overline{w}, v = \overline{v} - e^j$,有$h_j(\overline{x}) \leq 0$

$$\Rightarrow \quad h(\overline{x}) = 0.$$

若取$w = \overline{w} + e^i$,有$g_i(\overline{x}) \geq 0$

$$\Rightarrow \quad g(\overline{x}) \geq 0$$

$\therefore \quad \overline{x}$是可行点。

以下证明$\overline{x}$和$\left(\overline{w},\overline{v}\right)$分别是原问题和对偶问题的最优解

$\because h(\overline{x})=0, \quad \therefore \left(w-\overline{w}\right)^T g(\overline{x})\geq 0$

取$w=0$,有$-\overline{w}^T g(\overline{x})\geq 0.$

$\because \overline{w}\geq 0, g(\overline{x})\geq 0, \quad \therefore \quad \overline{w}^T g(\overline{x})=0$

由鞍点的定义可知，对任意的$x\in D$，有

$$
\begin{aligned}
f(\overline{x}) &= L\left(\overline{x},\overline{w},\overline{v}\right)\leq L\left(x,\overline{w},\overline{v}\right) \\
&= f(x)-\overline{w}^T g(x)-\overline{v}^T h(x)
\end{aligned}
$$

$\Rightarrow \quad f(\overline{x})\leq \theta\left(\overline{w},\overline{v}\right)$

根据弱对偶定理的推论2，$\overline{x}$和$\left(\overline{w},\overline{v}\right)$分别是原问题和对偶问题的最优解。

(2)在假设条件成立时，根据强对偶定理，对原问题的最优解$\bar{x}$，存在$(\bar{w},\bar{v}),\bar{w}\geq 0$，使得

$$f(\bar{x})=\theta(\bar{w},\bar{v}),\quad \bar{w}^T g(\bar{x})=0$$

$\because \theta(w,v)=\inf\{f(x)-w^T g(x)-v^T h(x)\mid x\in E^n\}$

$\therefore$ 对任意的$x\in E^n$，有

$\theta(\bar{w},\bar{v})\leq f(x)-\bar{w}^T g(x)-\bar{v}^T h(x)=L(x,\bar{w},\bar{v})$

$\because \bar{x}$为可行解，所以

$L(\bar{x},\bar{w},\bar{v})=f(\bar{x})=\theta(\bar{w},\bar{v})\leq L(x,\bar{w},\bar{v})$

$\because L(\bar{x},w,v)=f(\bar{x})-w^T g(\bar{x})-v^T h(\bar{x})$

且$g(\bar{x})\geq 0,w\geq 0,h(\bar{x})=0$

$\therefore L(\bar{x},w,v)\leq f(\bar{x})=L(\bar{x},\bar{w},\bar{v})$

$\Rightarrow(\bar{x},\bar{w},\bar{v})$是$Lagrange$函数的鞍点。

例 $\min f(x)=x^3$

$s.t.\quad -x^2\geq 0, x\in R$

最优解$\bar{x}=0, f(\bar{x})=0$,相应的$Lagrange$函数

$$L(x,w)=x^3+wx^2.$$

求$\bar{w}\geq 0$,使得对任意$x\in R$，有

$$L(\bar{x},w)\leq L(\bar{x},\bar{w})\leq L(x,\bar{w})$$

$$\Rightarrow \bar{x}^3+w\bar{x}^2\leq \bar{x}^3+\overline{w}\overline{x}^2\leq x^3+\bar{w}x^2$$

$$\Rightarrow \quad x^3+\bar{w}x^2\geq 0 \quad \forall x\in R$$

若$\bar{w}=0$，则取$x=-1$，上式不成立；

若$\bar{w}>0$，则取$x=-2\bar{w}$，上式不成立；

$\therefore$不存在$\bar{w}\geq 0$，使$(\bar{x},\bar{w})$为鞍点。

鞍点可能不存在，即便存在，也很难求

鞍点与KKT条件之间的关系

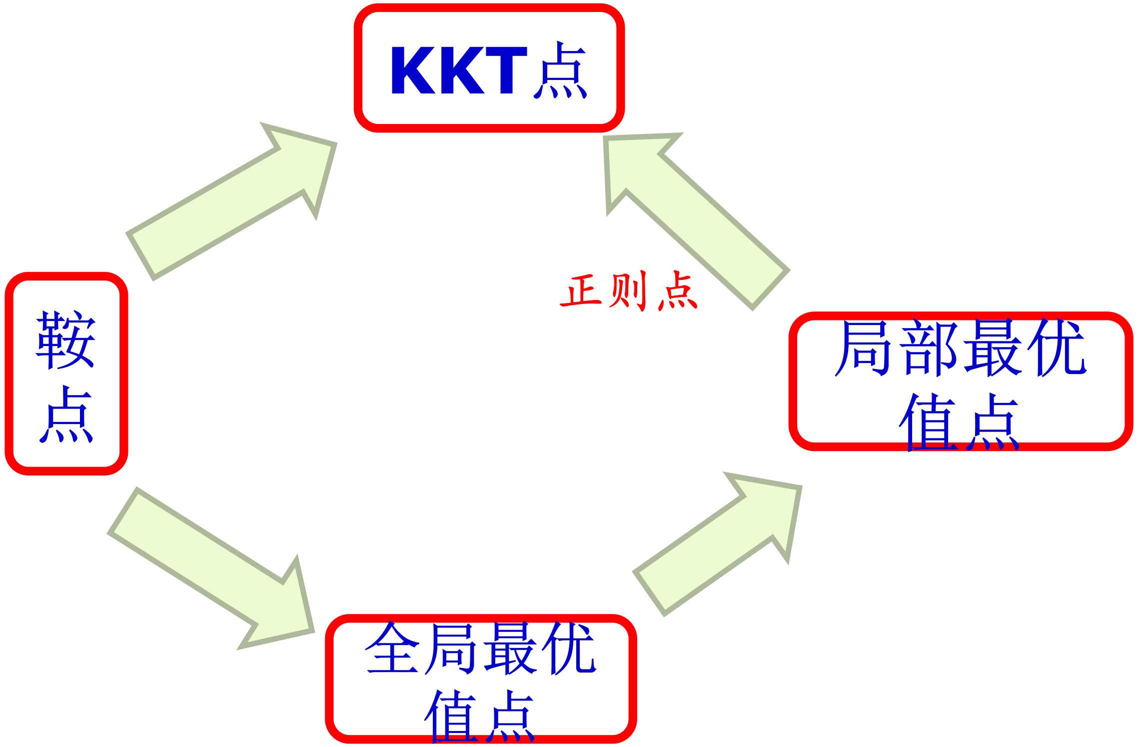
定理：

$\min\{f(x)\mid g(x)\geq 0, h(x)=0\}$中，可行域为$S$，$\bar{x}\in S$满足$KKT$条件，即存在$\bar{w}\geq 0$，$\bar{v}$使

$$\nabla f(\bar{x})-\nabla g(\bar{x})\bar{w}-\nabla h(\bar{x})\bar{v}=0,$$

且f为凸函数，$g_i(i\in I)$为凹函数，h_j为线性函数，则$(\bar{x},\bar{w},\bar{v})$为$Lagrange$函数$L(x,w,v)$的鞍点；反之，设$f,g_i,h_j$可微，若$(\bar{x},\bar{w},\bar{v})(\bar{w}\geq 0)$是$Lagrange$函数的鞍点，则$(\bar{x},\bar{w},\bar{v})$满足$KKT$条件。

最优性条件之间的关系



Lagrange乘子的意义

$$
(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\geq 0 \quad i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}
$$

设(NP)的局部最优解为x^*,相应的$Lagrange$乘子为(w^*,v^*),$w^*\geq 0$.

对约束的右端项进行扰动

$$\begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \ge \varepsilon_i \quad i = 1, 2, \cdots, m \\ \qquad\quad h_j(x) = \lambda_j \quad j = 1, 2, \cdots, l \end{cases}$$

扰动问题

令 $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_m), \lambda = (\lambda_1, \cdots, \lambda_l)$

设扰动问题的局部最优解为 $x^*(\varepsilon, \lambda)$，相应的 $Lagrange$ 乘子为 $(w^*(\varepsilon), v^*(\lambda))$，则当 $(\varepsilon, \lambda) = (0,0)$ 时，有 $x^*(0,0) = x^*$，$(w^*(0), v^*(0)) = (w^*, v^*)$.

只有一个等式约束

$$(NP)\begin{cases}\min\ f(x)\\ s.t.\quad h(x)=0\end{cases}$$

设局部最优解为x^*,相应的乘子为v^*.

扰动问题 $\begin{cases}\min\ f(x)\\ s.t.\quad h(x)=\lambda\end{cases}$

设局部最优解为$x^*(\lambda)$,相应的乘子为$v^*(\lambda)$.

$$\frac{d}{d\lambda}f\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x f\left(x^*(\lambda)\right)^T\frac{d}{d\lambda}x^*(\lambda)\Big|_{\lambda=0}$$

$$=\nabla_x f\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}.$$

由扰动问题的约束条件，得到

$$h\left(x^*(\lambda)\right)=\lambda$$

$$\therefore 1=\frac{d}{d\lambda}h\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x h\left(x^*(\lambda)\right)^T\frac{d}{d\lambda}x^*(\lambda)\Big|_{\lambda=0}$$

$$=\nabla_x h\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}.$$

由**KKT**条件 $\nabla f\left(x^*\right)-v^*\nabla h\left(x^*\right)=0$

得 $$\frac{d}{d\lambda}f\left(x^*(\lambda)\right)\Big|_{\lambda=0}=\nabla_x f\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}$$

$$=v^*\nabla_x h\left(x^*\right)^T\left[\frac{d}{d\lambda}x^*(\lambda)\right]_{\lambda=0}=v^*.$$

只有一个不等式约束

$$(NP)\begin{cases}\min\ f(x)\\ s.t.\quad g(x)\ge 0\end{cases}$$

设局部最优解为x^*,相应的乘子为w^*.

扰动问题 $\begin{cases}\min\ f(x)\\ s.t.\quad g(x)\ge \varepsilon\end{cases}$

设局部最优解为$x^*(\varepsilon)$,相应的乘子为$w^*(\varepsilon)$.

并假设$x^*(0)=x^*, w^*(0)=w^*$.

分两种情况讨论

(1) $g\left(x^*\right)=0$,即$g(x)\geq 0$在x^*处是起作用约束.

当$|\varepsilon|$很小时，可以假设有$g\left(x^*(\varepsilon)\right)=\varepsilon$,

即$g(x)$在$x^*(\varepsilon)$处为起作用约束，所以有

$$\frac{d}{d\varepsilon}f\left(x^*(\varepsilon)\right)\Big|_{\varepsilon=0}=w^*.$$

(2) $g\left(x^*\right)>0$,即$g(x)\geq 0$在x^*处是不起作用约束.

此时，x^*是无约束问题$\min f(x)$的局部最优解，因此当$|\varepsilon|$很小时，x^*也是扰动问题的局部最优解，所以有

$$\frac{d}{d\varepsilon}f\left(x^*(\varepsilon)\right)\Big|_{\varepsilon=0}=0=w^*.$$

$$(NP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\ge 0 \quad i=1,2,\cdots,m \\ \qquad\ \ h_j(x)=0 \quad j=1,2,\cdots,l\end{cases}$$

定理：设$f(x), g_i(x), h_j(x)$具有连续的二阶偏导数，x^*是(NP)的局部最优解，(w^*, v^*)是相应的$Lagrange$乘子向量.假设$x^*(\lambda,\varepsilon)$是扰动问题的局部最优解，$\left(w^*(\lambda), v^*(\varepsilon)\right)$是相应的乘子向量，则有

$$\nabla_\lambda f\left(x^*(\lambda)\right)\big|_{\lambda=0}=w^*$$

$$\nabla_\varepsilon f\left(x^*(\varepsilon)\right)\big|_{\varepsilon=0}=v^*.$$

例：某企业预算以2千元作为广告费，根据以往的经验，若以x_1千元作广播广告，x_2千元作报纸广告，销售金额为

$$-2x_1^2-10x_2^2-8x_1x_2+18x_1+34x_2\left(\text{千元}\right)$$

试问：

(1) 如何分配2千元广告费？

(2) 广告费预算作微小改变的影响如何？

解：最优化问题为

$$\min\ 2x_1^2+10x_2^2+8x_1x_2-18x_1-34x_2$$

$$s.t.\quad x_1+x_2-2=0$$

$$x_1\geq 0, x_2\geq 0$$

相应的KKT条件为

$$\begin{cases}4x_1+8x_2-18-w_1-v=0\\ 20x_2+8x_1-34-w_2-v=0\\ w_1x_1=0,\quad w_2x_2=0\\ x_1+x_2-2=0\\ x_1,x_2,w_1,w_2\geq 0\end{cases}$$

KKT点为$x^*=(1,1)^T$

$w_1{}^*=w_2{}^*=0$

$v=-6$

广告费作微小改动，考虑扰动问题

$$\min\ f(x)=2x_1^2+10x_2^2+8x_1x_2-18x_1-34x_2$$

$$s.t.\quad x_1+x_2-2=\varepsilon$$

$$x_1\geq 0, x_2\geq 0$$

有 $$\frac{df\left(x^*(\varepsilon)\right)}{d\varepsilon}\bigg|_{\varepsilon=0}=v^*=-6$$

当ε增加时，$f\left(x^*(\varepsilon)\right)$下降，即$-f\left(x^*(\varepsilon)\right)$上升，即当广告费增加后，销售金额也随着增加，而且销售金额的增加大约是广告费的6倍，可见适当增加广告费的预算是有利的。