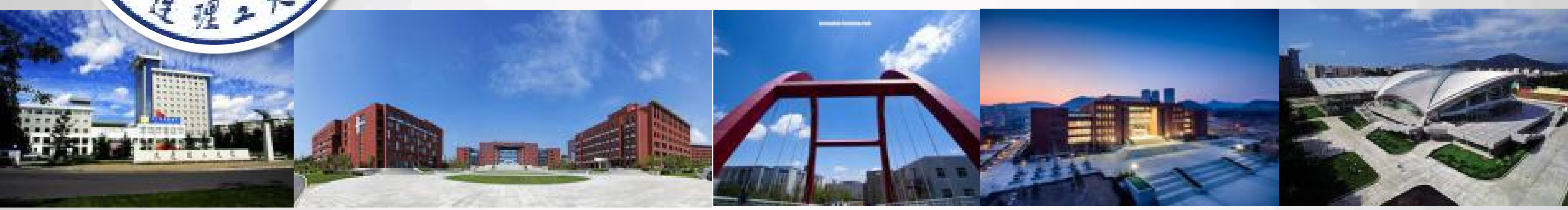
最优化理论
Optimality Theory



目 录 (CONTENT)

01 课程简介(Introduction)

02 线性规划(Linear Programming)

03 非线性规划(Non-Linear Programming)

04 整数规划(Integer Programming)

05 动态规划(Dynamic Programming)

07

PART SEVEN

最优性条件

Optimality Condition

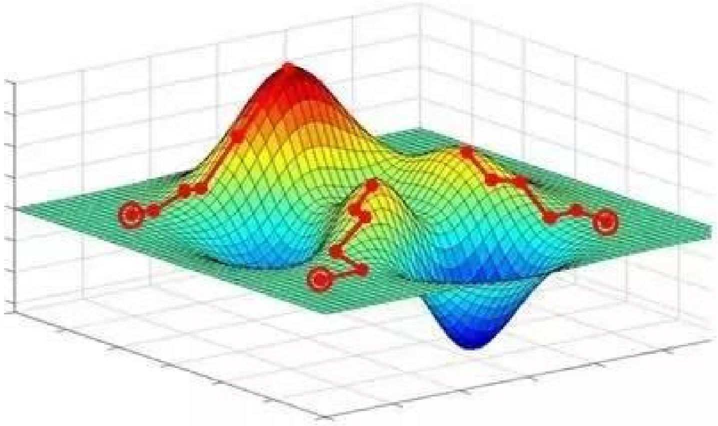
■ 基本概念

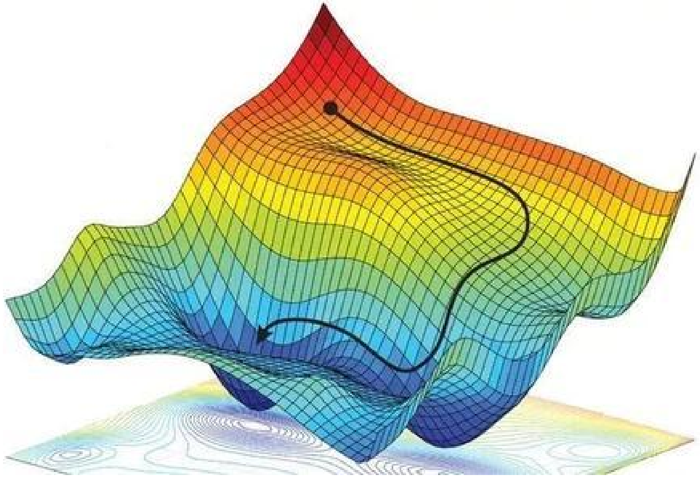
最优化问题可归结成如下数学形式：

$$\min_{x \in R^n} f(x) \qquad \text{---目标函数}$$

$$s.t. \quad g_i(x) \geq 0, i \in I$$

$$h_j(x) = 0, j \in E$$





大连理工大学
Dalian University Of Technology

■ 求函数的极小值

$$\phi(x)=\frac{x_1+x_2}{3+x_1^2+x_2^2+x_1x_2}$$

$$\begin{aligned}
&\min \quad (x_1-3)^2+(x_2-2)^2\\
&s.t. x_1^2+x_2^2\le 5,\\
&x_1+2x_2=4,\\
&x_1,x_2\ge 0.
\end{aligned}$$

■ 二阶条件

例 考虑如下约束优化问题,

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

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

对于可行点$x^*=(0,1)^T$, $I(x^*)=\{1\}$, $\nabla f(x^*)=(0,2)^T$, $\nabla g(x^*)=(0,1)^T$. 易见x^*是满足KKT条件. 此时线性独立约束规格(LICQ)成立，即

$$\{\nabla g_i(x^*), \nabla h_j(x^*), i\in I(x^*), j\in E\}$$

线性无关.但我们无法利用一阶最优性条件判断x^*是否为问题的局部极小点。

■ 二阶条件

为此，我们考虑函数的二阶导数，首先给出如下定义

定义2.3 设S是R^n中的一个非空集合，点$\bar{x}\in clS$集合

$$T=\{d\,|\,\exists x^{(k)}\in S, x^{(k)}\to \bar{x}\text{及}\lambda_k>0,\text{使得}d=\lim_{k\to\infty}\lambda_k(x^{(k)}-\bar{x})\}$$

则称T为集合S在点$\bar{x}$的切锥。

根据上述定义，如果序列$\{x^{(k)}\}\subset S, x^{(k)}\to\bar{x}$

$$x^{(k)}\neq\bar{x},\text{使得}\lim_{k\to\infty}\frac{x^{(k)}-\bar{x}}{\left\|x^{(k)}-\bar{x}\right\|}=d$$

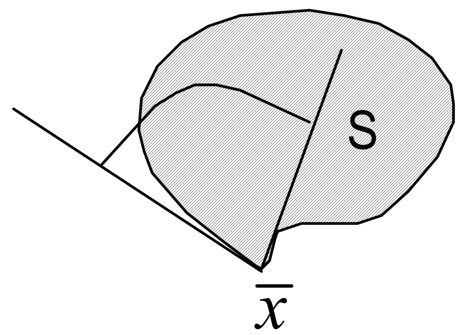
则$d\in T$。

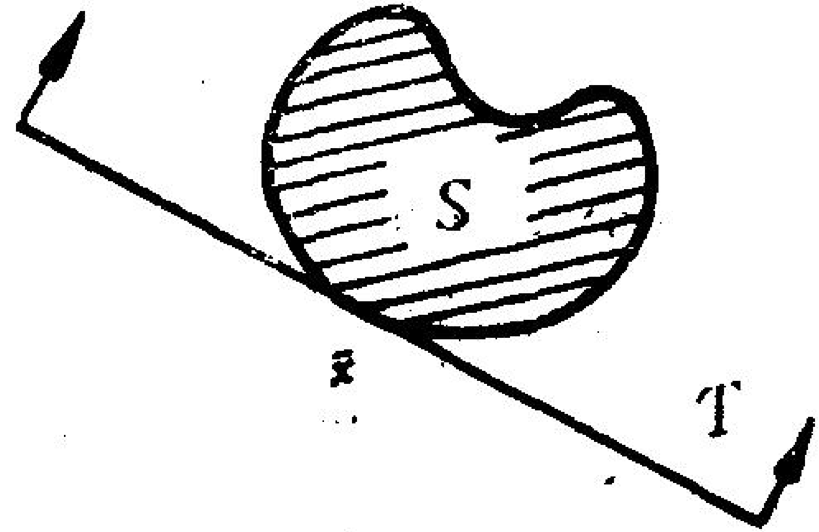
如果$\bar{x}\in clS$，则S在$\bar{x}$的切锥$T=R^n$.

最优性条件

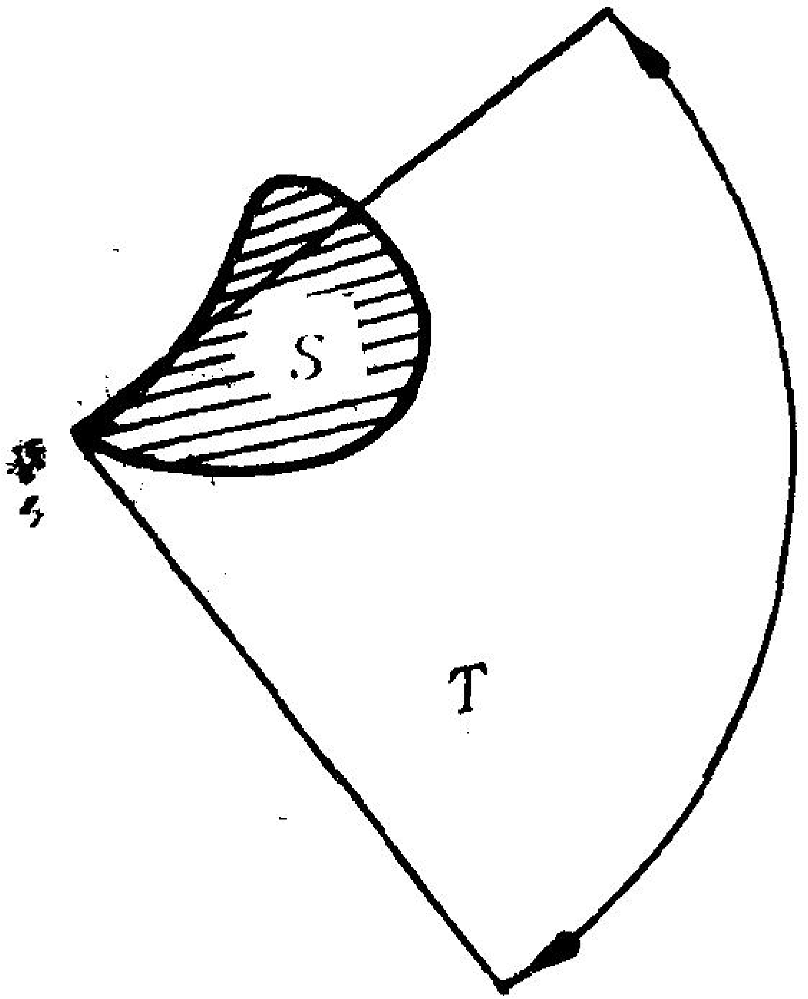
■ 二阶条件

例





■ 二阶条件



■ 二阶条件

现在我们考虑问题(7.2.1).

设在可行点$\bar{x}$，对应不等式约束中的起作用约束和等式约束的Lagrange乘子分别为：$\bar{w}_i \geq 0, i \in I$; $\bar{v}_j, j=1,2,..,l$.定义一个集合。

$$\bar{S}=\left\{x \left| \begin{array}{c} g_i(x)=0, i \in I \text{且} \bar{w}_i>0 \\ g_i(x) \geq 0, i \in I \text{且} \bar{w}_i=0 \\ h_i(x)=0, j=1,2,\ldots,l \end{array} \right.\right\}$$

■ 二阶条件

设集合$\bar{S}$在点$\bar{x}$的切锥为$\bar{T}$。再定义一个集合

$$\bar{G}=\left\{d\left|\begin{array}{l}\nabla g_i(\bar{x})'d=0, i\in I\text{且}\bar{w}_i>0\\ \nabla g_i(\bar{x})'d\geq 0, i\in I\text{且}\bar{w}_i=0\\ \nabla h_i(\bar{x})'d=0, j=1,2,\ldots,l\end{array}\right.\right\}$$

容易证明$\bar{G}\supset\bar{T}$

■ 二阶条件

设$d \in \bar{T}$,则存在可行序列$\{x^{(k)}\} \subset \bar{S}$和正数列$\{\lambda_k\}$,使得

$$\lim_{k\to\infty} \lambda_k (x^{(k)} - \bar{x}) = d$$

把$g_i(x)$和$h_i(x)$在$\bar{x}$展开,得到

$$g_i(x^{(k)}) = g_i(\bar{x}) + \nabla g_i(\bar{x})'(x^{(k)} - \bar{x}) + \left\| x^{(k)} - \bar{x} \right\| \alpha(\bar{x}, x^{(k)} - \bar{x})$$

$$h_j(x^{(k)}) = h_j(\bar{x}) + \nabla h_j(\bar{x})'(x^{(k)} - \bar{x}) + \left\| x^{(k)} - \bar{x} \right\| \alpha(\bar{x}, x^{(k)} - \bar{x})$$

■ 二阶条件

由于当$i \in I$时，$g_i(\bar{x})=0$，当$i \in I$且$\bar{w}_i > 0$时$g_i(x^{(k)})=0$，当$i \in I$且$\bar{w}_i=0$时$g_i(x^{(k)}) \geq 0$，以及$h_j(x^{(k)})=\mathrm{h}_j(\bar{x})=0$，故

当$i \in I$且$\bar{w}_i > 0$时

$$\nabla g_i(\bar{x})'(x^{(k)}-\bar{x})+\left\|x^{(k)}-\bar{x}\right\|\alpha(\bar{x},x^{(k)}-\bar{x})=0$$

当$i \in I$且$\bar{w}_i = 0$时

$$\nabla g_i(\bar{x})'(x^{(k)}-\bar{x})+\left\|x^{(k)}-\bar{x}\right\|\alpha(\bar{x},x^{(k)}-\bar{x})\geq 0$$

当$j=1,2,\ldots,l$时有

$$\nabla h_j(\bar{x})'(x^{(k)}-\bar{x})+\left\|x^{(k)}-\bar{x}\right\|\alpha(\bar{x},x^{(k)}-\bar{x})=0$$

二阶条件

把以上各式两端乘以λ_k，令$k \to \infty$，得

$$\nabla g_i(\bar{x})'d \geq 0, i \in I \text{ 且 } \bar{w}_i = 0$$

$$\nabla g_i(\bar{x})'d = 0, i \in I \text{ 且 } \bar{w}_i > 0$$

$$\nabla h_j(\bar{x})'\mathrm{d} = 0, j = 1, 2, \dots, l$$

即$d \in \bar{G}$.所以$\bar{G} \supset \bar{T}$.

由以上分析知道，切锥$\bar{T}$必包含于$\bar{G}$，但是反之不真。

下面，在$\bar{G} \subset \bar{T}$也成立的假设下，给出关于问题(7.2.1)的局部最优解的二阶必要条件。

■ 二阶条件

Theorem 2.10(二阶必要条件) 设$\bar{x}$是问题(7.2.1)的局部最优解, $f_i, g_i (i=1,\dots,m)$和$h_j (j=1,\dots,l)$二次连续可微，并存在满足(7. 2. 43)的乘子$\bar{w}=(\bar{w}_1,\dots,\bar{w}_m)$和$\bar{v}=(\bar{v}_1,\dots,\bar{v}_l)$, 再假设在点$\bar{x}$约束规格$\bar{G}=\bar{T}$成立，则对每一个向量$d\in\bar{G}$, 都有

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

其中$\nabla_x^2 L(\bar{x},\bar{w},\bar{v}) = \nabla^2 f(\bar{x}) - \sum_{i=1}^{m}\bar{w}_i\nabla^2 g_i(\bar{x}) - \sum_{j=1}^{l}\bar{v}_j\nabla^2 h_j(\bar{x})$

是Lagrange函数$L(x,w,v)$在点$\bar{x}$关于x的Hessian矩阵.

大连理工大学 Dalian University Of Technology

■ 二阶条件

证明：设向量$d \neq 0, d \in \bar{G}$.由于约束规格$\bar{G}=\bar{T}$成立，因此$d \in \bar{T}$，则存在可行序列$\{x^{(k)}\} \subset \bar{S}$和正数列$\{\lambda_k\}$，使得

$$\lim_{k\to\infty} \lambda_k (x^{(k)} - \bar{x}) = d$$

将$L(x, \bar{w}, \bar{v})$在$\bar{x}$展开，则

$$L(x^{(k)}, \bar{w}, \bar{v}) = L(\bar{x}, \bar{w}, \bar{v}) + \nabla_x L(\bar{x}, \bar{w}, \bar{v})'(x^{(k)} - \bar{x}) + \frac{1}{2}(x^{(k)} - \bar{x})' \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v})(x^{(k)} - \bar{x}) + \left\| x^{(k)} - \bar{x} \right\|^2 \alpha(\bar{x}, x^{(k)} - \bar{x}) \tag{7.2.50}$$

其中当$x^{(k)} \to \bar{x}$时，$\alpha(\bar{x}, x^{(k)} - \bar{x}) \to 0$。

■ 二阶条件

由于$x^{(k)} \in \bar{S}$，故有$h_j(x^{(k)}) = 0, \bar{w}_i g_i(x^{(k)}) = 0.$

根据Lagrange函数的定义，有

$$L(x^{(k)}, \bar{w}, \bar{v}) = f(x^{(k)}) \qquad (7.2.51)$$

$$L(\bar{x}, \bar{w}, \bar{v}) = f(\bar{x}) \qquad (7.2.52)$$

将上两式代入(7.2.50)，并注意到$\bar{x}$是局部最优解，

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

■ 二阶条件

$$f(x^{(k)})=f(\bar{x})+\frac{1}{2}(x^{(k)}-\bar{x})'\nabla_x{}^2L(\bar{x},\bar{w},\bar{v})(x^{(k)}-\bar{x})$$

$$+\left\|x^{(k)}-\bar{x}\right\|^2\alpha(\bar{x},x^{(k)}-\bar{x}) \qquad (7.2.53)$$

注意到$\bar{x}$是局部最优解，当k充分大时必有$f(x^{(k)})\geq f(\bar{x})$。

因此对充分大的k，由（7. 2. 53）得

$$\frac{1}{2}(x^{(k)}-\bar{x})'\nabla_x{}^2L(\bar{x},\bar{w},\bar{v})(x^{(k)}-\bar{x})+\left\|x^{(k)}-\bar{x}\right\|^2\alpha(\bar{x},x^{(k)}-\bar{x})\geq 0$$

上式两端乘以$\lambda_k{}^2$，并去取极限，则$d'\nabla_x{}^2L(\bar{x},\bar{w},\bar{v})d\geq 0$

证毕。

■ 二阶条件

为给出局部最优解的二阶充分条件，我们定义集合

$$G=\left\{d\left|\begin{array}{c} d\neq 0 \\ \nabla g_i(\bar{x})'d=0, i\in I \text{且} \bar{w}_i>0 \\ \nabla g_i(\bar{x})'d\geq 0, i\in I \text{且} \bar{w}_i=0 \\ \nabla h_i(\bar{x})'d=0, j=1,2,\ldots,l \end{array}\right.\right\}$$

Theorem 2.11(二阶充分条件)　设在问题7.2.1中，

$f_i, g_i(i=1,\ldots,m)$和$h_j(j=1,\ldots,l)$二次连续可微，并存在满足

(7. 2. 43) 的乘子$\bar{w}=(\bar{w}_1,\ldots,\bar{w}_m)$和$\bar{v}=(\bar{v}_1,\ldots,\bar{v}_l)$, $\bar{x}$为可行点，

且对每一个向量$d\in G$, 都有$d'\nabla_x{}^2L(\bar{x},\bar{w},\bar{v})d\geq 0$

则$\bar{x}$是严格的局部最优解。

■ 二阶条件

证明：反证法　　设$\bar{x}$不是严格的局部最优解，则存在收敛于$\bar{x}$的可行序列$\{x^{(k)}\}$，使得 $f(x^{(k)}) \le f(\bar{x})$　　(7.2.54)

令　　$d^{(k)} = \dfrac{x^{(k)} - \bar{x}}{\left\| x^{(k)} - \bar{x} \right\|}$　　(7.2.55)

因$\{d^{(k)}\}$为有界序列，故必有收敛子列$\{d^{(k_j)}\}$，设其极限为$d^{(0)}$.

把$g_i(x)$在$\bar{x}$展开，得到

$$g_i(x^{(k_j)}) = g_i(\bar{x}) + \nabla g_i(\bar{x})'(x^{(k_j)} - \bar{x}) + \left\| x^{(k_j)} - \bar{x} \right\| \alpha(\bar{x}, x^{(k_j)} - \bar{x})$$

其中当$k_j \to \infty$时$\alpha(\bar{x}, x^{(k_j)} - \bar{x}) \to 0$　　(7.2.56)

■ 二阶条件

当$i\in I$时$g_i(\bar{x})=0$，又$x^{(k_j)}$是可行点，$g_i(x^{(k_j)})\geq 0$，故由(7.2.56)

得到$\nabla g_i(\bar{x})'(x^{(k_j)}-\bar{x})+\left\|x^{(k_j)}-\bar{x}\right\|\alpha(\bar{x},x^{(k_j)}-\bar{x})\geq 0$

上式两端除以$\left\|x^{(k_j)}-\bar{x}\right\|$，令$k_j\to\infty$，则

$$\nabla g_i(\bar{x})'d^{(0)}\geq 0, i\in I. \qquad (7.2.57)$$

类似可得$\nabla h_j(\bar{x})'d^{(0)}\geq 0, j=1,\ldots,l.$ (7.2.58)

及 $$\nabla f(\bar{x})'d^{(0)}\leq 0, \qquad (7.2.59)$$

■ 二阶条件

下面分两种情况讨论

1) $d^{(0)} \notin G$

此时，由(7. 2. 57)和集合G的定义可知，必存在下标 $i \in I$ 使得 $\overline{w}_i > 0$ 和 $\nabla g_i(\overline{x})' d^{(0)} > 0$。于是利用 KKT 条件，

$$\text{必有} \nabla f(\overline{x})' d^{(0)} = (\sum_{i \in I} \overline{w}_i \nabla g_i(\overline{x}) + \sum_{j=1}^{l} \overline{v}_j \nabla h_j(\overline{x}))' d^{(0)}$$

$$= \sum_{i \in I} \overline{w}_i \nabla g_i(\overline{x})' d^{(0)} > 0$$

此与(7.2.59)矛盾。

大连理工大学 Dalian University Of Technology

■ 二阶条件

1) $d^{(0)} \in G$

此时，把Lagrange函数$L(x,\bar{w},\bar{v})$在$\bar{x}$展开，则

$$L(x^{(k_j)},\bar{w},\bar{v})=L(\bar{x},\bar{w},\bar{v})+\nabla_x L(\bar{x},\bar{w},\bar{v})'(x^{(k_j)}-\bar{x})$$

$$+\frac{1}{2}(x^{(k_j)}-\bar{x})'\nabla_x^2 L(\bar{x},\bar{w},\bar{v})(x^{(k_j)}-\bar{x})$$

$$+\left\|x^{(k_j)}-\bar{x}\right\|^2\alpha(\bar{x},x^{(k_j)}-\bar{x}) \qquad (7.2.60)$$

其中当$x^{(k_j)}\to\bar{x}$时$\alpha(\bar{x},x^{(k_j)}-\bar{x})\to 0$。

由于$x^{(k_j)}$是可行点，$\bar{w}=(w_1,\ldots,w_m)\geq 0$, 根据Lagrange函数的定义，有

最优性条件

■ 二阶条件

$$L(x^{(k_j)},\bar{w},\bar{v})=f(x^{(k_j)})-\sum_{i=1}^{m}\bar{w}_i g_i(x^{(k_j)})-\sum_{k=1}^{l}\bar{v}_k h_k(x^{(k_j)})$$

故 $L(x^{(k_j)},\bar{w},\bar{v})\le f(x^{(k_j)})$ (7.2.61)

又知 $L(\bar{x},\bar{w},\bar{v})=f(\bar{x})$ (7.2.62)

由假设还有$\nabla_x L(\bar{x},\bar{w},\bar{v})=0$ (7.2.63)

$$f(x^{(k_j)})\le f(\bar{x}) \qquad (7.2.64)$$

将(7.2.61)—(7.2.64)代入(7.2.60)，则

$$\frac{1}{2}(x^{(k_j)}-\bar{x})'\nabla_x^2 L(\bar{x},\bar{w},\bar{v})(x^{(k_j)}-\bar{x})+\left\| x^{(k_j)}-\bar{x}\right\|^2\alpha(\bar{x},x^{(k_j)}-\bar{x})\le$$

■ 二阶条件

上式两端除以$\left\| x^{(k_j)} - \bar{x} \right\|^2$，令$k_j \to \infty$，则

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

此与$d' \nabla_x^{\,2} L(\bar{x}, \bar{w}, \bar{v}) d > 0 (d \in G)$的假设相矛盾。

■ 二阶条件

例7.2.7 考虑下列非线性规划问题

$$\min \quad x_1$$
$$s.t. 3(x_1-3)^2+x_2 \geq 0$$
$$(x_1-3)^2+x_2^2-10=0$$

检验以下各点是否为局部最优解

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

■ 二阶条件

记目标函数和约束函数分别为f(x),g(x),h(x),它们在点x处的梯度分别是

$$\nabla f(x)=\begin{Bmatrix}1\\0\end{Bmatrix},\nabla g(x)=\begin{Bmatrix}6(x_1-3)\\1\end{Bmatrix},\nabla h(x)=\begin{Bmatrix}2(x_1-3)\\2x_2\end{Bmatrix}$$

Lagrange函数是

$$L(x,w,v)=x_1-w[3(x_1-3)^2+x_2]-v[(x_1-3)^2+x_2^2-10]$$

Lagrange函数关于x的Hessian矩阵是

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

大连理工大学 Dalian University Of Technology

■ 二阶条件

检查$x^{(1)}$:是可行点，且两约束都是起作用约束。

$$\nabla f(x^{(1)})=\begin{Bmatrix}1\\0\end{Bmatrix},\nabla g(x^{(1)})=\begin{Bmatrix}-6\\1\end{Bmatrix},\nabla h(x^{(1)})=\begin{Bmatrix}-2\\-6\end{Bmatrix}$$

按照KKT条件，设

$$\begin{Bmatrix}1\\0\end{Bmatrix}-w\begin{Bmatrix}-6\\1\end{Bmatrix}-v\begin{Bmatrix}-2\\-6\end{Bmatrix}=0\Rightarrow w=-\frac{3}{19}.v=-\frac{1}{38}$$

不存在使$w\geq 0$的解，故它不是KKT点.

■ 二阶条件

检查$x^{(2)}$：是可行点，且两约束都是起作用约束。

$$\nabla f(x^{(2)}) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}, \nabla g(x^{(2)}) = \begin{Bmatrix} 6 \\ 1 \end{Bmatrix}, \nabla h(x^{(1)}) = \begin{Bmatrix} 2 \\ -6 \end{Bmatrix}$$

按照KKT条件，设

$$\begin{Bmatrix} 1 \\ 0 \end{Bmatrix} - w\begin{Bmatrix} 6 \\ 1 \end{Bmatrix} - v\begin{Bmatrix} 2 \\ -6 \end{Bmatrix} = 0 \Rightarrow w = \frac{3}{19}, v = \frac{1}{38}$$

故它是KKT点, 此点Lagrange函数的Hessian矩阵为：

$$\nabla_x^2 L(x^{(2)}, w, v) = \begin{bmatrix} -1 & 0 \\ 0 & -\frac{1}{19} \end{bmatrix}$$

■ 二阶条件

求集合$\bar{G}$中的元素。由于$w>0$,根据$\bar{G}$的定义，令

$\nabla g(x^{(2)})'d=0,\ \nabla h(x^{(2)})'d=0$

上述方程组即$\begin{cases} 6d_1+d_2=0 \\ 2d_1-6d_2=0 \end{cases} \Rightarrow d=(0,0)'$

故G＝Φ。此情况表明在充分条件中对曲率的要求自然满足，因此该点是局部最优解。

后两点请自行验证之

■ 二阶条件

例7.2.8 考虑下列非线性规划问题

$$\min \quad x_1^2+(x_2-2)^2$$

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

其中β为某个实数。讨论点$x=(0,0)$是否为局部最优解。

记目标函数和约束函数分别为f(x), h(x),他们在点x
处的梯度分别是

$$\nabla f(x)=\begin{bmatrix}0\\-4\end{bmatrix},\nabla f(x)=\begin{bmatrix}0\\-1\end{bmatrix}$$

■ 二阶条件

设 $\begin{bmatrix} 0 \\ -4 \end{bmatrix} - v \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0 \Rightarrow$ v=4

*Lagrange*函数为

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

它关于x的Hessian矩阵是

$$\nabla_x^2 L = \begin{bmatrix} 2-2\beta v & 0 \\ 0 & 2 \end{bmatrix}$$

在点$x^{(0)}$处，有$\nabla_x^2 L(x^{(0)}, v) = \begin{bmatrix} 2-2\beta v & 0 \\ 0 & 2 \end{bmatrix}$

二阶条件

求集合G的元素d，令$(0,-1)\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}=0 \Rightarrow d_2=0$

解得$d=(d_1,0)', d_1$可取任意实数。此时有

$$d'\nabla_x^2 L(x^{(0)},v)d = 2(1-4\beta)d_1^2$$

当$\beta<1/4$时。对每一个向量$d\in G$，有

$$d'\nabla_x^2 L(x^{(0)},v)d > 0$$

故$x^{(0)}=(0,0)$是局部最优解。

大连理工大学 Dalian University Of Technology

■ 二阶条件

当$\beta > 1/4$时。对每一个向量$d \in G$,有

$d'\nabla_x^2 L(x^{(0)}, v)d < 0$

此时不满足局部最优解的二阶必要条件，故

$x^{(0)} = (0,0)$不是局部最优解。

当$\beta = 1/4$时利用二阶条件给不出结论，可用其他方法判断。

此时原问题即

$$\min \quad x_1^2 + (x_2 - 2)^2$$

$$s.t. \quad x_1^2 / 4 - x_2 = 0$$

利用约束条件，从目标函数中消去一个变量，把问题化为

无约束问题$\min 4x_2 + (x_2 - 2)^2$，显见$x^{(0)} = (0,0)$为局部最优解

■ 二阶条件

例7.2.8 考虑下列非线性规划问题

$$\min \quad x_1^2 + (x_2 - 2)^2$$

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

其中β为某个实数。讨论点$x=(0,0)$是否为局部最优解。

本例表明，研究约束问题的二阶条件时，只考虑目标函数的Hessian矩阵是不行的。

■ 第五次作业

第243页第七章2.3.4.5.6.8.9

感兴趣的同学学习7.3节--对偶及鞍点问题
(和线性规划的对偶问题类似)

11月23日--------习题课！
(相同时间答疑)

小结

- 无约束问题的极值条件
- 约束极值问题的最优性条件
- 对偶及鞍点