M05M11084 最优化理论、算法与应用

6-2 不等式约束优化问题

不等式约束优化问题

参考：

1. 最优化导论（4版），第21章，Edwin K.P.Chong, Stanislaw H. Żak著，孙志强等译
2. Linear and Nonlinear Programming, Chapter 11, 3rd ed., David G. Luenberger, Yinyu Ye
3. Practical Optimization Algorithms and Engineering Applications, § 10.7, A. Antoniou, W. LU

注意：[1]中定义梯度算子为 $\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_1} \end{bmatrix}^T$ [2]中定义梯度算子为 $\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_1} \end{bmatrix}$

本课采用

1. Karush-Kuhn-Tucker (KKT) 条件
2. 二阶条件
3. KKT乘子的意义　敏感度分析

1. Karush-Kuhn-Tucker (KKT) 条件
    1.1 KKT条件
    1.2 Lagrangian 函数与KKT条件
2. 二阶条件
3. KKT乘子的意义　敏感度分析

积极约束与非积极约束

考虑 $\min f(\boldsymbol{x})$
s.t. $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$
$\boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}$

$\boldsymbol{x} \in \mathcal{R}^n, f: \mathcal{R}^n \to \mathcal{R}, \boldsymbol{h}: \mathcal{R}^n \to \mathcal{R}^m$，$\boldsymbol{g}: \mathcal{R}^n \to \mathcal{R}^p$，$m \le n$

$\mathcal{S} = \{\boldsymbol{x} \in \mathcal{R}^n | \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}\}$ 可行域

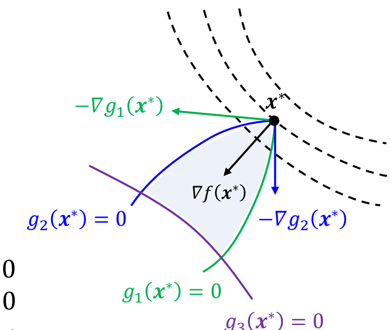
$\mathcal{E} = \{1, \dots, m\}$，$\mathcal{I} = \{1, \dots, p\}$

积极约束：$g_j(\boldsymbol{x}) \le 0$，对于$\boldsymbol{x}^* \in \mathcal{S}$，有$g_j(\boldsymbol{x}^*) = 0$

非积极约束：$g_j(\boldsymbol{x}) \le 0$，对于$\boldsymbol{x}^* \in \mathcal{S}$，有$g_j(\boldsymbol{x}^*) < 0$

积极约束的下标集：$\mathcal{J}(\boldsymbol{x}^*) \triangleq \{j \in \mathcal{I}: g_j(\boldsymbol{x}^*) = 0\}$

等式约束$h_i(\boldsymbol{x}) = 0$可看作为积极约束



active $g_1(\boldsymbol{x}^*) = 0$
$g_2(\boldsymbol{x}^*) = 0$

$\mathcal{J}(\boldsymbol{x}^*) \triangleq \{1,2\}$ *inactive* $g_3(\boldsymbol{x}^*) < 0$

正则点

设$\boldsymbol{x}^* \in \mathcal{S}$，$\mathcal{J}(\boldsymbol{x}^*)$是起作用的不等式约束的下标集，即，$\mathcal{J}(\boldsymbol{x}^*) \triangleq \{j: g_j(\boldsymbol{x}^*) = 0\}$
如果向量$\nabla h_i(\boldsymbol{x}^*), i \in \mathcal{E}; \nabla g_j(\boldsymbol{x}^*), j \in \mathcal{J}(\boldsymbol{x}^*)$线性无关，则称$\boldsymbol{x}^*$是正则点

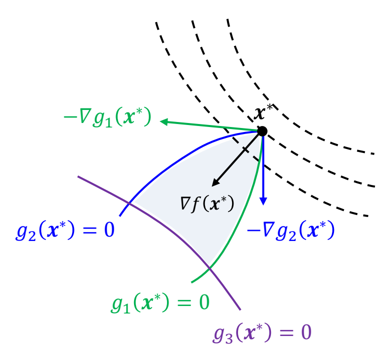
正则点$\Leftrightarrow \begin{bmatrix} D\boldsymbol{h}(\boldsymbol{x}^*) \\ Dg_j(\boldsymbol{x}^*), j \in \mathcal{J}(\boldsymbol{x}^*) \end{bmatrix}$ 满秩

active $g_1(\boldsymbol{x}^*) = 0$
$g_2(\boldsymbol{x}^*) = 0$

inactive $g_3(\boldsymbol{x}^*) < 0$

$\mathcal{J}(\boldsymbol{x}^*) \triangleq \{1,2\}$

$\nabla g_1(\boldsymbol{x}^*)$和$\nabla g_2(\boldsymbol{x}^*)$线性无关 $\boldsymbol{x}^*$是正则点



Karush-Kuhn-Tucker (KKT) 定理

设$f, \boldsymbol{h}, \boldsymbol{g} \in C^1$. 设$\boldsymbol{x}^*$是在约束条件$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$，$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$下极小化$f$问题的正则点和局部极值点，那么，必然存在 $\boldsymbol{\lambda}^* \in \mathcal{R}^m$和$\boldsymbol{\mu}^* \in \mathcal{R}^p$，使得，

1. $\boldsymbol{\mu}^* \geq \boldsymbol{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.

$D_x l(x^*, \lambda^*, \mu^*) = 0$

$\nabla f(\boldsymbol{x}^*) + \nabla \boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{\lambda}^* + \nabla \boldsymbol{g}(\boldsymbol{x}^*)\boldsymbol{\mu}^* = \boldsymbol{0}$.

称 $\boldsymbol{\lambda}^*$ 为Lagrange 乘子向量　　称 $\boldsymbol{\mu}^*$ 为KKT 乘子向量

λ_i^* 为 Lagrange 乘子　　μ_i^* 为 KKT 乘子

说明 1. $\boldsymbol{\mu}^* \geq \boldsymbol{0} \Leftrightarrow \mu_j^* \geq 0, j \in \mathcal{J}$

2. $Df(\boldsymbol{x}^*) + \sum_{i\in\mathcal{E}} \lambda_i^* Dh_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{J}} \mu_j^* Dg_j(\boldsymbol{x}^*) = \boldsymbol{0}^T$

2. $\nabla f(\boldsymbol{x}^*) + \sum_{i\in\mathcal{E}} \lambda_i^* \nabla h_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{J}} \mu_j^* \nabla g_j(\boldsymbol{x}^*) = \boldsymbol{0}$

3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0 \Leftrightarrow \sum_{j\in\mathcal{J}} \mu_j^* g_j(\boldsymbol{x}^*) = 0 \Leftrightarrow \mu_j^* g_j(\boldsymbol{x}^*) = 0, j \in \mathcal{J}$

因为 $g_j(\boldsymbol{x}^*) \leq 0$
$\mu_j^* \geq 0$

$\forall j \in \mathcal{J}(\boldsymbol{x}^*),\quad g_j(\boldsymbol{x}^*) = 0 \Rightarrow \mu_j^* \geq 0$

$\forall j \in \mathcal{J}\backslash\mathcal{J}(\boldsymbol{x}^*),\quad g_j(\boldsymbol{x}^*) < 0 \Rightarrow \mu_j^* = 0$

KKT定理的几何解释

例 21.1　点$\boldsymbol{x}^*$是问题的极小点. 不等式约束 $g_j(\boldsymbol{x}) \leq 0$，$j = 1,2,3$

其中，约束$g_3(\boldsymbol{x}) \leq 0$是不起作用的，$g_3(\boldsymbol{x}^*) < 0$；所以 $\mu_3^* = 0$

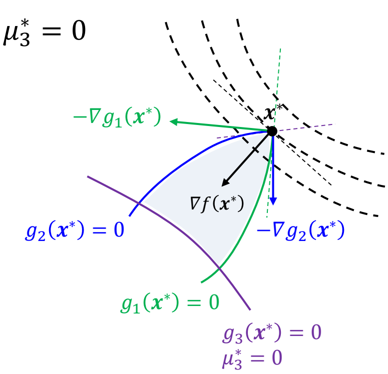
根据KKT定理，

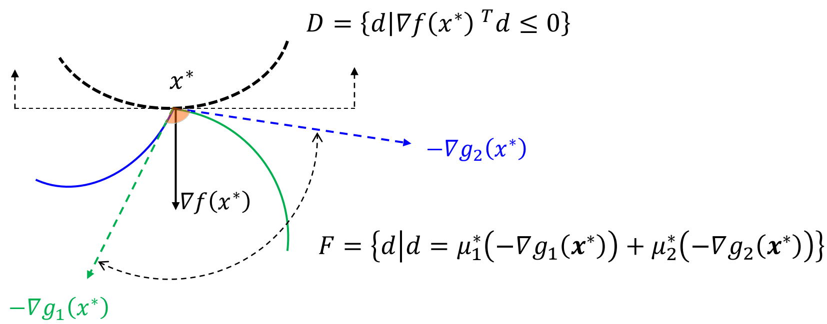
$$\nabla f(\boldsymbol{x}^*) + \mu_1^* \nabla g_1(\boldsymbol{x}^*) + \mu_2^* \nabla g_2(\boldsymbol{x}^*) + \mu_3^* \nabla g_3(\boldsymbol{x}^*) = \boldsymbol{0}$$

$$\mu_1^* > 0, \mu_2^* > 0, \mu_3^* = 0$$

$$\nabla f(\boldsymbol{x}^*) = -\mu_1^* \nabla g_1(\boldsymbol{x}^*) - \mu_2^* \nabla g_2(\boldsymbol{x}^*)$$

$\nabla f(\boldsymbol{x}^*)$ 是向量$-\nabla g_1(\boldsymbol{x}^*)$和$-\nabla g_2(\boldsymbol{x}^*)$的正线性组合（锥组合）





点x^*是极小点， x^*邻域内且属于可行域的点的函数值都较大，即，周边没有能使函数值下降的路径（方向）

也就是，没有可行下降方向 $\Leftrightarrow$ $F \cap D = \emptyset$

$\Updownarrow$ Farks引理

KKT条件，$\nabla f(x^*) \in F$

KKT定理的证明

设$f, \boldsymbol{h}, \boldsymbol{g} \in C^1$. 设$\boldsymbol{x}^*$是在约束条件$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$, $\boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}$下极小化$f$问题的正则点和局部极值点，那么，必然存在 $\boldsymbol{\lambda}^* \in \mathcal{R}^m$和$\boldsymbol{\mu}^* \in \mathcal{R}^p$，使得，

1. $\boldsymbol{\mu}^* \ge \boldsymbol{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.

分析： 不等式约束的优化问题 等式约束的优化问题

$$\begin{array}{ll} \min f(x) & \min f(x) \\ \text{s.t. } h(x)=0 & \text{s.t. } x\in S \\ \quad g(x)\le 0 & \end{array} \quad (P) \qquad\qquad \begin{array}{l} \min f(x) \\ \text{s.t. } x\in \tilde{S} \end{array} \quad (\tilde{P})$$

$x^* \in S=\{x|h(x)=0, g(x)\le 0\}$ $\qquad$ $x^*\in \tilde{S}=\{x\,|\,h(x)=0, g_j(x)=0, j\in \mathcal{J}(x^*)\}$

等式约束 $\longrightarrow$ 等式约束

不等式约束 $\begin{cases}\text{积极约束} \longrightarrow \text{积极约束} \\ \text{非积极约束}\end{cases}$

等式约束 + 积极约束 $\Big\}$ 等式约束

x^*是(P)的正则的极小解 $\qquad$ ① 如果x^*也是$(\tilde{P})$的正则的极小解 ？

$$Df(x^*)+\lambda^{*T}Dh(x^*)+\sum_{j\in \mathcal{J}(x^*)}\mu_j^* Dg_j(x^*)=0 \Longleftarrow \text{由Lagrange定理}$$

$$Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0$$

② $\mu^*\ge 0 \xLeftrightarrow{\mu^*=\{\mu_j^*, j\in \mathcal{J}\}} \begin{array}{l}\mu_j^*=0, j\in \mathcal{J}\backslash \mathcal{J}(x^*) \\ \mu_j^*\ge 0, \forall j\in \mathcal{J}(x^*)\end{array}$ ？

$$Df(x^*)+\lambda^{*T}Dh(x^*) \overbrace{\underline{\qquad\qquad} + \sum_{j\in \mathcal{J}(x^*)}\mu_j^* Dg_j(x^*)}^{\mu^{*T}Dg(x^*)} = 0$$

$$\sum_{j\in \mathcal{J}\backslash \mathcal{J}(x^*)}\mu_j^* Dg_j(x^*)=0$$

$\updownarrow$

$\mu_j^*=0, j\in \mathcal{J}\backslash \mathcal{J}(x^*)$ $\qquad$ $\begin{array}{l}\mu_j^* g_j(x^*)=0, j\in \mathcal{J}\backslash \mathcal{J}(x^*) \\ \mu_j^* g_j(x^*)=0, j\in \mathcal{J}(x^*)\end{array}$

$$Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0$$

$$\mu^{*T}g(x^*)=0$$

① 引理

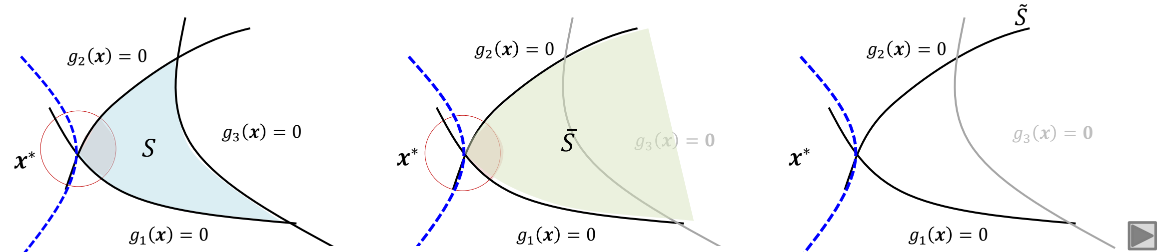
设$f,h,g \in C^1$. 如果x^*是(P)问题的正则的局部极小点，
那么， x^*也是$(\tilde{P})$问题的正则的局部极小点

$$\min f(x) \quad \text{s.t. } x \in S \qquad (P) \qquad\qquad \min f(x) \quad \text{s.t. } x \in \tilde{S} \qquad (\tilde{P})$$

$$S = \{x | h(x) = 0, g(x) \le 0\} \qquad \tilde{S} = \{x | h(x) = 0, g_j(x) = 0, j \in \mathcal{I}(x^*)\}$$

$$B = \{x | \|x - x^*\| < \varepsilon, \varepsilon > 0\} \qquad \bar{S} = \{x | h(x) = 0, g_j(x) \le 0, j \in \mathcal{I}(x^*)\}$$

$B \cap \bar{S} = B \cap S$ ？ $\qquad S \subseteq \bar{S} \qquad \tilde{S} \subseteq \bar{S}$

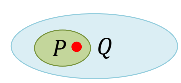


$x^* = \bar{x}^*$? $\qquad \tilde{x}^* = \bar{x}^*$?

① 引理

若$P \subseteq Q$，则 $\min_{x \in P} f(x) \ge \min_{y \in Q} f(y)$

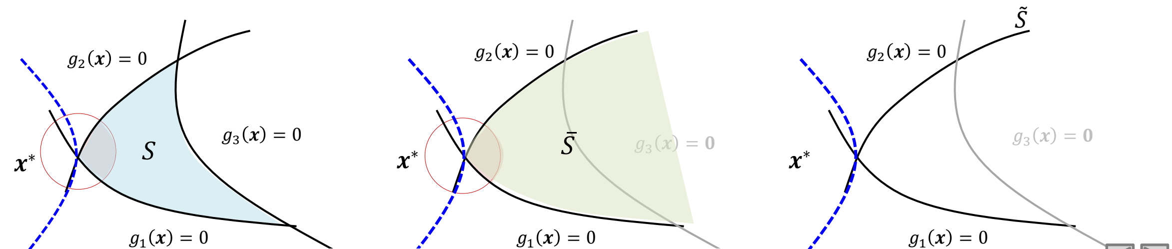
若$x^* = \operatorname{argmin}_{x \in Q} f(x) \in P$，则 $\quad \operatorname{argmin}_{x \in P} f(x) = x^*$



S上的局部极小点 $\Longrightarrow S \cap B \subseteq S$的局部极小点
$\Longrightarrow \bar{S} \cap B = S \cap B$的局部极小点 $\Longrightarrow \tilde{S} \subseteq \bar{S}$上正则局部极小点

$$B = \{x | \|x - x^*\| < \varepsilon, \varepsilon > 0\} \qquad \bar{S} = \{x | h(x) = 0, g_j(x) \le 0, j \in \mathcal{I}(x^*)\}$$

$B \cap \bar{S} = B \cap S \qquad S \subseteq \bar{S} \qquad \tilde{S} \subseteq \bar{S}$



$x^* = \bar{x}^* \qquad \tilde{x}^* = \bar{x}^*$

① 引理

(1) 先证明 $\boxed{S\cap B=\bar{S}\cap B}$，即，$S\cap B\subset \bar{S}\cap B$ 且 $S\cap B\supset \bar{S}\cap B$

构造 x^* 的小邻域 $B=\{x|\|x-x^*\|<\varepsilon\}$，　B存在

使得 x^* 的非积极约束仍然是"B中属于可行域的点"的非积极约束

即，$\forall x\in B$，$\mathcal{I}\backslash\mathcal{I}(x^*)\subseteq\mathcal{I}\backslash\mathcal{I}(x)$

由 $g_i(x^*)<0,\forall i\in\mathcal{I}\backslash\mathcal{I}(x^*)$，$g\in C^1,\forall i\in\mathcal{I}$，$g_i$ 都是连续的

$\Longrightarrow$ ∃足够小的 $\varepsilon>0$，使 $g_i(x)<0,\forall i\in\mathcal{I}\backslash\mathcal{I}(x^*),\forall x\in B$

$S\subset\bar{S}\quad\Longrightarrow\quad \underline{S\cap B\subset\bar{S}\cap B}$

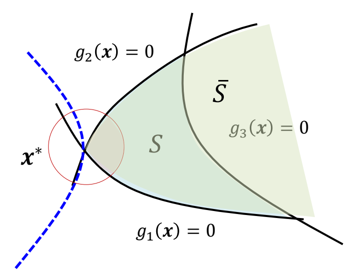
设 $x\in\bar{S}\cap B$，即 $x\in\bar{S}$，$x\in B$

由 $\bar{S}$ 和 B 的定义，有

$h(x)=0,g_j(x)\le 0,\ \forall j\in\mathcal{I}(x^*)$ 和 $g_i(x)<0,\forall i\in\mathcal{I}\backslash\mathcal{I}(x^*)$

所以，$x\in S\cap B$

$\Longrightarrow\quad \underline{\bar{S}\cap B\subset S\cap B}$



① 引理

设 $f,h,g\in C^1$. 如果 x^* 是 (P) 问题的正则的局部极小点，

那么，x^* 也是 $(\tilde{P})$ 问题的正则的局部极小点

$$\begin{aligned}&\min f(x)\\&\text{s.t. } x\in S\end{aligned}\qquad (P)\qquad\qquad \begin{aligned}&\min f(x)\\&\text{s.t. } x\in \tilde{S}\end{aligned}\qquad (\tilde{P})$$

$$S=\{x|h(x)=0,g(x)\le 0\}\qquad\qquad \tilde{S}=\{x|h(x)=0,g_j(x)=0,j\in\mathcal{I}(x^*)\}$$

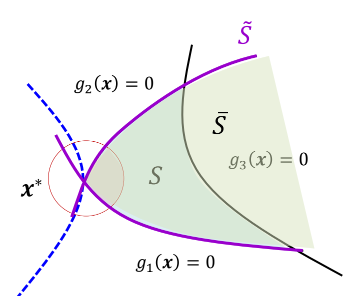
(2) ∵ $\boldsymbol{x}^*$ 是 f 集合 S 上的局部极小点

而 $S\cap B\subset S$，$\boldsymbol{x}^*$ 也是 f 在集合 $S\cap B=\bar{S}\cap B$ 上的局部极小点

∴ $\boldsymbol{x}^*$ 也是 f 集合 $\bar{S}$ 上的正则局部极小点

又，$\tilde{S}\subset\bar{S}$ 和 $\boldsymbol{x}^*\in\tilde{S}$

因此，$\boldsymbol{x}^*$ 是 f 集合 $\tilde{S}$ 上的正则局部极小点



KKT定理的证明

设$f,\boldsymbol{h},\boldsymbol{g}\in C^1$. 设$\boldsymbol{x}^*$是在约束条件$\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0}$, $\boldsymbol{g}(\boldsymbol{x})\le\boldsymbol{0}$下极小化$f$问题的正则点和局部极值点，那么，必然存在 $\boldsymbol{\lambda}^*\in\mathcal{R}^m$和$\boldsymbol{\mu}^*\in\mathcal{R}^p$，使得，

1. $\boldsymbol{\mu}^*\ge\boldsymbol{0}$. ?
2. $Df(\boldsymbol{x}^*)+\boldsymbol{\lambda}^{*T}D\boldsymbol{h}(\boldsymbol{x}^*)+\boldsymbol{\mu}^{*T}D\boldsymbol{g}(\boldsymbol{x}^*)=\boldsymbol{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*)=0$.

① 如果x^*是(P)问题的正则的局部极小点，那么， x^*也是$(\tilde{P})$问题的正则的局部极小点

$$\begin{aligned}&\min f(x)\\&\text{s.t. } x\in S\end{aligned}\qquad (P)\qquad\qquad \begin{aligned}&\min f(x)\\&\text{s.t. } x\in \tilde{S}\end{aligned}\qquad (\tilde{P})$$

$S=\{x|h(x)=0,g(x)\le 0\}$ $\qquad$ $\tilde{S}=\{x|h(x)=0,g_j(x)=0,j\in\mathcal{J}(x^*)\}$

$$Df(x^*)+\lambda^{*T}Dh(x^*)+\qquad\qquad+\sum_{j\in\mathcal{J}(x^*)}\mu_j^*Dg_j(x^*)=0 \;\Leftarrow\text{Lagrange定理}$$

$\mu_j^*=0,j\in\mathcal{I}\backslash\mathcal{J}(x^*)$

$\mu_j^*g_j(x^*)=0,j\in\mathcal{I}\backslash\mathcal{J}(x^*)$

$\mu_j^*g_j(x^*)=0,j\in\mathcal{J}(x^*)$

② $\mu^*\ge 0$

思路 $\begin{cases}\mu_j^*=0,j\in\mathcal{I}\backslash\mathcal{J}(x^*)\\ \mu_j^*\ge 0,\forall j\in\mathcal{J}(x^*)\end{cases}$

证明 $\forall j\in\mathcal{J}(x^*)$， $\mu_j^*\ge 0$ $\qquad$ 此时，只考虑某一个μ_j^*

对 $Df(x^*)+\lambda^{*T}Dh(x^*)+\mu_j^*Dg_j(x^*)+\sum_{i\in\mathcal{J}(\boldsymbol{x}^*)\backslash\{j\}}\mu_i^*Dg_i(x^*)=0$

找一个方向y，左乘

$$Df(x^*)y+\underbrace{\lambda^{*T}Dh(x^*)y}_{0}+\mu_j^*Dg_j(x^*)y+\sum_{i\in\mathcal{J}(\boldsymbol{x}^*)\backslash\{j\}}\underbrace{\mu_i^*Dg_i(x^*)\,y}_{0}=0$$

满足此条件的方向y的集合

$\hat{T}(x^*)=\{y|Dh(x^*)y=0,Dg_i(x^*)y=0,i\in\mathcal{J}(\boldsymbol{x}^*)\backslash\{j\}\}$

$\hat{S}=\{x|h(x)=0,g_i(x)=0,i\in\mathcal{J}(\boldsymbol{x}^*)\backslash\{j\}\}$ $\qquad$ 对应的除了j的积极约束集合

y不能是f的下降方向
因为x^*是局部极小点 $\quad 0<Df(x^*)y=-\mu_j^*Dg_j(x^*)y$ $\begin{cases}Dg_j(x^*)y<0 & (1)\ ?\\ \mu_j^*\ge 0 & (2)\ ?\end{cases}$

② $\mu^* \geq 0$ $\begin{cases} \mu_j^* = 0, j \in \mathcal{I} \backslash \mathcal{I}(x^*) \\ \mu_j^* \geq 0, \forall j \in \mathcal{I}(x^*) \end{cases}$

构造两个集合：$\hat{T}(x^*) = \{y | Dh(x^*)y = 0, Dg_i(x^*)y = 0, i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$ $\forall j \in \mathcal{I}(x^*)$

$\hat{S} = \{x | h(x) = 0, g_i(x) = 0, i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$

(1) 先证明：$Dg_j(x^*)\, y < 0$，$j \in \mathcal{I}(x^*)$

(a) $Dg_j(x^*)y \leq 0$

设$x = x^* + \delta y, \delta > 0$ $g_j(x) = g_j(x^*) + \delta Dg_j(x^*)y + o(\delta) \leq 0$

(b) $Dg_j(x^*)\, y \neq 0$

对正则点x^*，存在 $y \in \hat{T}(x^*)$

反证法. 假设$\forall\, y \in \hat{T}(x^*)$，$Dg_j(x^*)\, y = 0 \iff \nabla g_j(x^*)^T y = 0 \Leftrightarrow \nabla g_j(x^*) \in \hat{T}(x^*)^{\perp}$

由“切空间和法空间互为正交补的引理”可知，

$\nabla g_j(x^*) \in \mathrm{span}\{\nabla h_k(x^*), k \in \mathcal{E}; \nabla g_i(x^*), i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$

即，$\nabla g_j(x^*)$可由等式约束的$\nabla h_k\,(x^*)$和其它积极约束的$\nabla g_i(x^*)$线性组合表示

与x^*是正则点的矛盾

② $\mu^* \geq 0$ $\begin{cases} \mu_j^* = 0, j \in \mathcal{I} \backslash \mathcal{I}(x^*) \\ \mu_j^* \geq 0, \forall j \in \mathcal{I}(x^*) \end{cases}$

构造两个集合：$\hat{T}(x^*) = \{y | Dh(x^*)y = 0, Dg_i(x^*)y = 0, i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$ $\forall j \in \mathcal{I}(x^*)$

$\hat{S} = \{x | h(x) = 0, g_i(x) = 0, i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$

(1) 先证明：$Dg_j(x^*)\, y < 0$，$j \in \mathcal{I}(x^*)$

(a) $Dg_j(x^*)y \leq 0$

设$x = x^* + \delta y, \delta > 0$ $g_j(x) = g_j(x^*) + \delta Dg_j(x^*)y + o(\delta) \leq 0$

(b) $Dg_j(x^*)\, y \neq 0$

对正则点x^*，存在 $y \in \hat{T}(x^*)$

反证法. 假设$\forall\, y \in \hat{T}(x^*)$，$Dg_j(x^*)\, y = 0 \iff \nabla g_j(x^*)^T y = 0 \Leftrightarrow \nabla g_j(x^*) \in \hat{T}(x^*)^{\perp}$

由“切空间和法空间互为正交补的引理”可知，

$\nabla g_j(x^*) \in \mathrm{span}\{\nabla h_k(x^*), k \in \mathcal{E}; \nabla g_i(x^*), i \in \mathcal{I}(\boldsymbol{x}^*) \backslash \{j\}\}$

即，$\nabla g_j(x^*)$可由等式约束的$\nabla h_k\,(x^*)$和其它积极约束的$\nabla g_i(x^*)$线性组合表示

与x^*是正则点的矛盾

② $\mu^* \geq 0$

$$\begin{cases} \mu_j^* = 0, j \in \mathcal{I}\backslash\mathcal{J}(x^*) \\ \mu_j^* \geq 0, \forall j \in \mathcal{J}(x^*) \end{cases}$$

(1) 已证明：$Dg_j(x^*)\, y < 0$，$j \in \mathcal{J}(x^*)$

(2) 现证明：$\mu_j^* \geq 0$，$j \in \mathcal{J}(x^*)$

由①得 $$Df(x^*) + \lambda^{*T} Dh(x^*) + \mu_j^* Dg_j(x^*) + \sum_{i \in \mathcal{J}(x^*)\backslash\{j\}} \mu_i^* Dg_i(x^*) = 0^T$$

右乘向量$y \in \hat{T}(x^*) = \{y | Dh(x^*)y = 0, Dg_i(x^*)y = 0, i \in \mathcal{J}(\boldsymbol{x}^*)\backslash\{j\}\}$

$$Df(x^*)y + \lambda^{*T} \underbrace{Dh(x^*)y}_{0} + \mu_j^* Dg_j(x^*)y + \sum_{i \in \mathcal{J}(x^*)\backslash\{j\}} \mu_i^* \underbrace{Dg_i(x^*)\, y}_{0} = 0^T$$

$$\Rightarrow \quad Df(x^*)y = -\mu_j^* Dg_j(x^*)y$$

$\forall x \in S \cap B$ $\quad f(x) = f(x^*) + Df(x^*)\dot{x}(t^*)\delta + o(\delta)$

$x = x(t^* + \delta)$

$\delta \geq 0$

$$\begin{aligned} 0 \leq f(x) - f(x^*) &\approx Df(x^*)\dot{x}(t^*)\delta \\ &= \delta Df(x^*)y \\ &= -\delta\mu_j^* Dg_j(x^*)y \end{aligned}$$

记号说明

$$D\boldsymbol{h}(\boldsymbol{x}) = \begin{bmatrix} Dh_1(\boldsymbol{x}) \\ \vdots \\ Dh_m(\boldsymbol{x}) \end{bmatrix}$$ Jacobian of $\boldsymbol{h}(\boldsymbol{x})$

$$\boldsymbol{h}(\boldsymbol{x}) = \begin{bmatrix} h_1(\boldsymbol{x}) \\ \vdots \\ h_m(\boldsymbol{x}) \end{bmatrix}$$

$$\begin{aligned} [D\boldsymbol{h}(\boldsymbol{x})]^T &= [[Dh_1(\boldsymbol{x})]^T \quad \cdots \quad [Dh_m(\boldsymbol{x})]^T] \\ &= [\nabla h_1(\boldsymbol{x}) \quad \cdots \quad \nabla h_m(\boldsymbol{x})] \\ &= \nabla \boldsymbol{h}(\boldsymbol{x}) \end{aligned}$$

$$\begin{aligned} [D\boldsymbol{h}(\boldsymbol{x})]^T \boldsymbol{\lambda} &= [\nabla h_1(\boldsymbol{x}) \quad \cdots \quad \nabla h_m(\boldsymbol{x})] \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} \\ &= \lambda_1 \nabla h_1(\boldsymbol{x}) + \cdots + \lambda_m \nabla h_m(\boldsymbol{x}) \\ &= \nabla \boldsymbol{h}(\boldsymbol{x}) \boldsymbol{\lambda} \\ &= \sum_{i=1}^{m} \lambda_i \nabla h_i(\boldsymbol{x}) \end{aligned}$$

$$\begin{aligned} \boldsymbol{\lambda}^T D\boldsymbol{h}(\boldsymbol{x}) &= [\lambda_1 \quad \cdots \quad \lambda_m] \begin{bmatrix} Dh_1(\boldsymbol{x}) \\ \vdots \\ Dh_m(\boldsymbol{x}) \end{bmatrix} \\ &= \lambda_1 Dh_1(\boldsymbol{x}) + \cdots + \lambda_m Dh_m(\boldsymbol{x}) \\ &= \sum_{i=1}^{m} \lambda_i Dh_i(\boldsymbol{x}) \end{aligned}$$

$$D\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} Dg_1(\boldsymbol{x}) \\ \vdots \\ Dg_p(\boldsymbol{x}) \end{bmatrix}$$ Jacobian of $\boldsymbol{g}(\boldsymbol{x})$

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_p(\boldsymbol{x}) \end{bmatrix}$$

$$\begin{aligned}[D\boldsymbol{g}(\boldsymbol{x})]^T &= \left[[Dg_1(\boldsymbol{x})]^T \quad \cdots \quad \left[Dg_p(\boldsymbol{x})\right]^T\right] \\ &= [\nabla g_1(\boldsymbol{x}) \quad \cdots \quad \nabla g_p(\boldsymbol{x})] \\ &= \nabla \boldsymbol{g}(\boldsymbol{x})\end{aligned}$$

$$\begin{aligned}\boldsymbol{\mu}^T D\boldsymbol{g}(\boldsymbol{x}) &= [\mu_1 \quad \cdots \quad \mu_p] \begin{bmatrix} Dg_1(\boldsymbol{x}) \\ \vdots \\ Dg_p(\boldsymbol{x}) \end{bmatrix} \\ &= \mu_1 Dg_1(\boldsymbol{x}) + \cdots + \mu_p Dg_p(\boldsymbol{x}) \\ &= \sum_{j=1}^{p} \mu_j Dg_j(\boldsymbol{x})\end{aligned}$$

$$\begin{aligned}[D\boldsymbol{g}(\boldsymbol{x})]^T \boldsymbol{\mu} &= [\nabla g_1(\boldsymbol{x}) \quad \cdots \quad \nabla g_p(\boldsymbol{x})] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} \\ &= \mu_1 \nabla g_1(\boldsymbol{x}) + \cdots + \mu_p \nabla g_p(\boldsymbol{x}) \\ &= \nabla \boldsymbol{g}(\boldsymbol{x}) \boldsymbol{\mu} \\ &= \sum_{j=1}^{p} \mu_j \nabla g_j(\boldsymbol{x})\end{aligned}$$

1. Karush-Kuhn-Tucker (KKT) 条件

   1.1 KKT条件

   1.2 Lagrangian 函数与KKT条件

2. 二阶条件

3. KKT乘子的意义　敏感度分析

Lagrangian 函数与KKT条件

$$l(x,\lambda,\mu) \triangleq f(x) + \lambda^T h(x) + \mu^T g(x)$$

$D_x l(x,\lambda,\mu) \triangleq D_x f(x) + \lambda^T D_x h(x) + \mu^T D_x g(x)$

$D_\lambda l(x,\lambda,\mu) \triangleq h(x)$

$D_\mu l(x,\lambda,\mu) \triangleq g(x)$

$Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$

$h(x^*) = 0$

$$D_\mu l(x^*,\lambda^*,\mu^*) = 0 \begin{cases} g(x^*) \le 0 \\ \mu^{*T} g(x^*) = 0, \mu^* \ge 0 \end{cases}$$

积极约束$g_j = 0$
非积极约束$\mu_i = 0$

Karush-Kuhn-Tucker (KKT) 定理.设$f,h,g \in C^1$.设x^*是在约束条件$h(x) = 0, g(x) \le 0$下极小化f问题的正则点和局部极小点，那么，必然存在 $\lambda^* \in \mathcal{R}^m$和$\mu^* \in \mathcal{R}^p$，使得，

1. $\mu^* \ge 0.$
2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T.$
3. $\mu^{*T} g(x^*) = 0.$

满足KKT 条件的点是候选极小点 FONC

KKT条件由5条组成 (3个方程和2个不等式):

1. $\mu^* \ge 0$
2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$ $\Leftrightarrow$ $\nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0$
3. $\mu^{*T} g(x^*) = 0$ 互补松弛Complementary Slack
4. $h(x^*) = 0$
5. $g(x^*) \le 0$

对偶可行
Dual
Feasible

原问题可行
Primal Feasible

$\Updownarrow$

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j=1}^{p} \mu_j \nabla g_j(x) = 0$$

$$\mu^{*T} g(x^*) = 0 \Leftrightarrow \mu_j g_j(x^*) = 0$$

$\because \mu_j \ge 0,\quad g_j(x^*) \le 0$

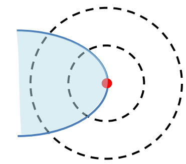
积极约束$g_j = 0$，$\mu_j \ge 0$
非积极约束$g_i < 0$，$\mu_i = 0$

例

$$\min f(x) = (x_1 - 1)^2 + (x_2 - 2)^2$$
$$\text{s.t. } g_1(x) = -(x_1 - 1) - (x_2 - 2)^2 \le 0$$
$$x^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$\mu^* = 0$$

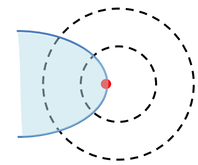
$$l(x, \mu) = f(x) + \mu g_1(x)$$

$$D_x l(x, \mu) = \begin{bmatrix} 2(x_1 - 1) - \mu \\ 2(x_2 - 2)\,(1 - \mu) \end{bmatrix} = 0$$



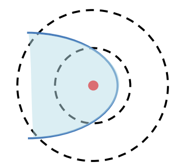
$$\min f(x) = (x_1 - 1)^2 + (x_2 - 2)^2$$
$$\text{s.t. } g_2(x) = -(x_1 - 1) - (x_2 - 2)^2 + 1 \le 0$$
$$x^* = \begin{bmatrix} 0.5 \\ 2 \end{bmatrix}$$
$$\mu^* = 1$$

$$x_1 = 1 + \frac{1}{2}\mu$$
$$x_2 = 2, \text{or}, \mu = 1$$
$$-(x_1 - 1) - (x_2 - 2)^2 \le 0$$
$$\mu(-(x_1 - 1) - (x_2 - 2)^2) = 0$$
$$\mu \ge 0$$



$$\min f(x) = (x_1 - 1)^2 + (x_2 - 2)^2$$
$$\text{s.t. } g_3(x) = -(x_1 - 1) - (x_2 - 2)^2 - 1 \le 0$$
$$x^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$\mu^* = 0$$

$$-(x_1 - 1) - (x_2 - 2)^2 + 1 \le 0$$
$$\mu(-(x_1 - 1) - (x_2 - 2)^2 + 1) = 0$$
$$\mu \ge 0$$
$$-(x_1 - 1) - (x_2 - 2)^2 - 1 \le 0$$
$$\mu(-(x_1 - 1) - (x_2 - 2)^2 - 1) = 0$$
$$\mu \ge 0$$



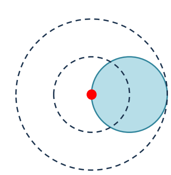
例

$$f(x) = \frac{1}{2}(x_1^2 + x_2^2)$$
$$g(x) = \frac{1}{2}\left((x_1 - 1)^2 + x_2^2\right) - \frac{1}{2} \le 0$$

$$l(x, \mu) = f(x) + \mu g(x)$$

$$\mu \ge 0$$
$$D_x l(x, \mu) = \begin{bmatrix} x_1 + \mu(x_1 - 1) \\ x_2 + \mu x_2 \end{bmatrix} = 0$$
$$\mu\left(\frac{1}{2}\left((x_1 - 1)^2 + x_2^2\right) - \frac{1}{2}\right) = 0$$
$$\frac{1}{2}\left((x_1 - 1)^2 + x_2^2\right) - \frac{1}{2} \le 0$$

$$x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\mu^* = 0$$



例 21.3

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

可行点 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$, 且有$g(\boldsymbol{x}_1) \geq 0$ 和 $g(\boldsymbol{x}_2) \geq 0$,
并满足KKT条件

$g(\boldsymbol{x}_1) \geq 0$

$\max f(\boldsymbol{x})$
$\text{s.t.}\, g(\boldsymbol{x}) \geq 0$ → $\min -f(\boldsymbol{x})$
$\text{s.t.} -g(\boldsymbol{x}) \leq 0$

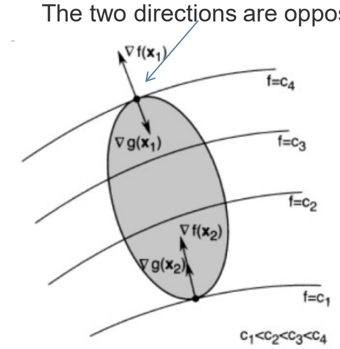
→
1. $\mu \geq 0$
2. $D(-f(\boldsymbol{x}_1)) + \mu D(-g(\boldsymbol{x}_1)) = 0$
3. $\mu(-g(\boldsymbol{x}_1)) = 0$
4. NIL
5. $-g(\boldsymbol{x}_1) \leq 0$

→
1. $\mu \geq 0$
2. $\nabla f(\boldsymbol{x}_1) + \mu\nabla g(\boldsymbol{x}_1) = 0$
3. $\mu g(\boldsymbol{x}_1) = 0$
4. NIL
5. $g(\boldsymbol{x}_1) \geq 0$

↓

$\nabla f(\boldsymbol{x}_1) = -\mu\nabla g(\boldsymbol{x}_1)$

The two directions are opposite.



Points satisfying the KKT condition
($\boldsymbol{x}_1$ is a maximizer and $\boldsymbol{x}_2$ is a minimizer)

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

可行点 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$, 且有$g(\boldsymbol{x}_1) \geq 0$ 和 $g(\boldsymbol{x}_2) \geq 0$,
并满足KKT条件

$g(\boldsymbol{x}_2) \geq 0$

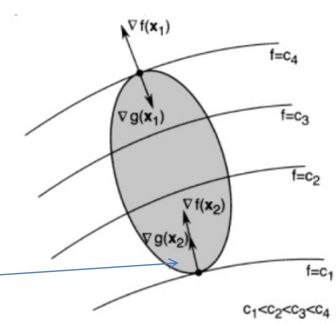
$\min f(\boldsymbol{x})$
$\text{s.t.}\, g(\boldsymbol{x}) \geq 0$ → $\min f(\boldsymbol{x})$
$\text{s.t.} -g(\boldsymbol{x}) \leq 0$

→
1. $\mu \geq 0$
2. $Df(\boldsymbol{x}_2) + \mu D(-g(\boldsymbol{x}_2)) = 0$
3. $\mu(-g(\boldsymbol{x}_2)) = 0$
4. NIL
5. $-g(\boldsymbol{x}_2) \leq 0$

→
1. $\mu \geq 0$
2. $\nabla f(\boldsymbol{x}_2) - \mu\nabla g(\boldsymbol{x}_2) = 0$
3. $\mu g(\boldsymbol{x}_2) = 0$
4. NIL
5. $g(\boldsymbol{x}_2) \geq 0$

↓

$\nabla f(\boldsymbol{x}_2) = \mu\nabla g(\boldsymbol{x}_2)$



Points satisfying the KKT condition
($\boldsymbol{x}_1$ is a maximizer and $\boldsymbol{x}_2$ is a minimizer)

例 21.4

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

考虑问题 $\min f(x) = x_1^2 + x_2^2 + x_1x_2 - 3x_1$

$\text{s.t.}\, x \geq 0$

$$g(x) = \begin{bmatrix} -x_1 \\ -x_2 \end{bmatrix}$$

$$Df(x) = [2x_1 + x_2 - 3 \quad x_1 + 2x_2]$$

$$Dg(x) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\mu^T Dg(x) = [\mu_1 \quad \mu_2] \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = [-\mu_1 \quad -\mu_2]$$

$$Df(x) + \mu^T Dg(x) = [2x_1 + x_2 - 3 - \mu_1 \quad x_1 + 2x_2 - \mu_2] = 0$$

1. $\mu \geq 0$
2. $Df(x) + \mu^T Dg(x) = 0^T$
3. $\mu^T g(x) = 0$
4. NIL
5. $g(x) \leq 0$

1. $\mu_1, \mu_2 \geq 0$
2. $[2x_1 + x_2 - 3 - \mu_1 \quad x_1 + 2x_2 - \mu_2] = 0^T$
3. $\mu_1(-x_1) = 0, \mu_2(-x_2) = 0$
4. NIL
5. $-x_1 \leq 0, -x_2 \leq 0$

$\longrightarrow$

$\mu_1^* = 0, \mu_2^* = 1.5$
$x_1^* = 1.5, x_2^* = 0$ only a candidate for a minimizer √

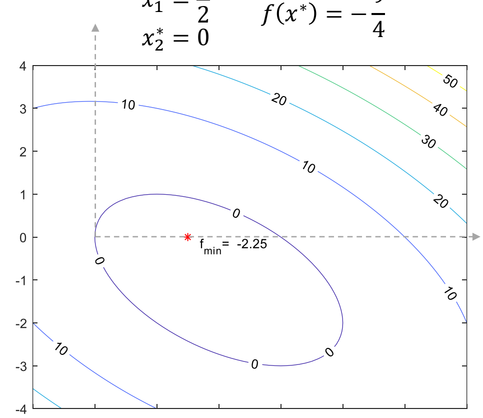
$\mu_1^* = -3, \mu_2^* = 0$
$x_1^* = 0, \quad x_2^* = 0$ violates the nonpositivity constraint X

```
close;
clear;
clc;

f=@(x1,x2) x1.^2+x2.^2+x1.*x2-3*x1;
t=-1:0.1:6;
s=-4:0.1:4;
[X1,X2]=meshgrid(t,s);
F=f(X1,X2);
contour(X1,X2,F,'ShowText','on');
hold on

xx1=1.5;xx2=0;
fx=f(xx1,xx2)
plot(xx1,xx2,'r*')
xxx1=xx1+0.2;xxx2=xx2-0.2;
fm=['f_m_i_n=',num2str(fx)]
text(xxx1,xxx2,fm)
```

$$x_1^* = \frac{3}{2} \qquad f(x^*) = -\frac{9}{4}$$
$$x_2^* = 0$$



例 21.5

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

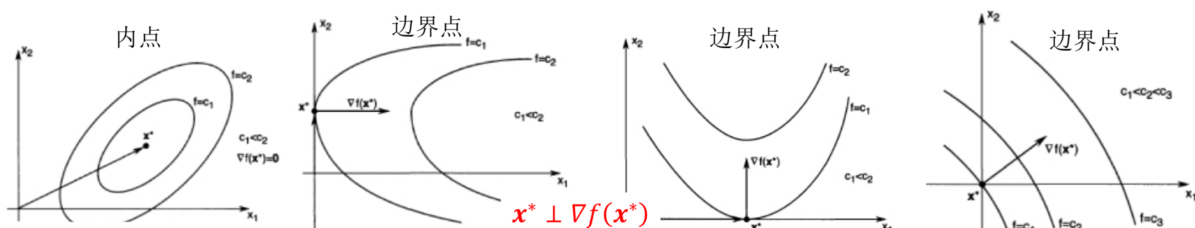
A special case of a more general problem of the form

$$\min f(\boldsymbol{x})$$
$$\text{s.t. } \boldsymbol{x} \geq 0$$

$$g(\boldsymbol{x}) = -\boldsymbol{x} \qquad Dg(\boldsymbol{x}) = -I_n$$

$$\nabla f(\boldsymbol{x}) = -\boldsymbol{\mu} \geq \mathbf{0}$$
$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, d_i \geq 0$$

$$\begin{cases} 1.\ \boldsymbol{v} \geq \boldsymbol{0} \\ 2.\ Df(\boldsymbol{x}) + \boldsymbol{v}^{\mathrm{T}} D\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}^{\mathrm{T}} \\ 3.\ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}) = 0 \\ 4.\ NIL \\ 5.\ \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{cases} \longrightarrow \begin{cases} 1.\ \boldsymbol{v} \geq \boldsymbol{0} \\ 2.\ Df(\boldsymbol{x}) - \boldsymbol{v}^{\mathrm{T}} = \boldsymbol{0}^{\mathrm{T}} \\ 3.\ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = 0 \\ 4.\ NIL \\ 5.\ \boldsymbol{x} \geq \boldsymbol{0} \end{cases} \xrightarrow{\boldsymbol{\mu} = -\boldsymbol{v}} \begin{cases} 1.\ \boldsymbol{\mu} \leq \boldsymbol{0} \\ 2.\ Df(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}} = \boldsymbol{0}^{\mathrm{T}} \\ 3.\ \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{x} = 0 \\ 4.\ NIL \\ 5.\ \boldsymbol{x} \geq \boldsymbol{0} \end{cases} \longrightarrow \begin{cases} \nabla f(\boldsymbol{x}) \geq 0 \\ \boldsymbol{x}^{\mathrm{T}}\nabla f(\boldsymbol{x}) = 0 \\ \boldsymbol{x} \geq \boldsymbol{0} \end{cases}$$



$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ Some possible points satisfying the KKT condition for problems with positive constraints, $\mathcal{R}^2$. $\boldsymbol{x}^* = \mathbf{0}$

M05M11084 最优化理论、算法与应用

6-2 不等式约束优化问题

1. Karush-Kuhn-Tucker (KKT) 条件

2. 二阶条件

   2.1 二阶必要条件

   2.2 二阶充分条件

3. KKT乘子的意义　敏感度分析

Lagrange函数的Hesse阵

$$\begin{aligned} &\min f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ &\qquad \boldsymbol{g}(\boldsymbol{x}) \le \mathbf{0} \end{aligned}$$

其中， $\boldsymbol{x} \in \mathcal{R}^n, f: \mathcal{R}^n \to \mathcal{R}, \boldsymbol{h}: \mathcal{R}^n \to \mathcal{R}^m$, $\boldsymbol{g}: \mathcal{R}^n \to \mathcal{R}^p$, $m \le n$

$$l(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x})$$

Lagrange函数的Hesse阵: $\nabla_{xx}^2 l(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = F(\boldsymbol{x}) + [\boldsymbol{\lambda} H(\boldsymbol{x})] + [\boldsymbol{\mu} G(\boldsymbol{x})]$$

$F(\boldsymbol{x})$函数f在点$\boldsymbol{x}$处的Hesse阵

$$[\boldsymbol{\lambda} H(\boldsymbol{x})] = \sum_{j=1}^{m} \lambda_j H_j(\boldsymbol{x}) = \lambda_1 H_1(\boldsymbol{x}) + \cdots + \lambda_m H_m(\boldsymbol{x})$$

$H_j(\boldsymbol{x})$ h_j 在点$\boldsymbol{x}$处的Hesse阵

$$H_j(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial^2 h_j(\boldsymbol{x})}{\partial x_1^2} & \cdots & \dfrac{\partial^2 h_j(\boldsymbol{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 h_j(\boldsymbol{x})}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 h_j(\boldsymbol{x})}{\partial x_n^2} \end{bmatrix}$$

$$[\boldsymbol{\mu} G(\boldsymbol{x})] = \sum_{k=1}^{p} \mu_k G_k(\boldsymbol{x}) = \mu_1 G_1(\boldsymbol{x}) + \cdots + \mu_p G_p(\boldsymbol{x})$$

$G_k(\boldsymbol{x})$ g_k在点$\boldsymbol{x}$处的Hesse阵

$$G_k(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial^2 g_k(\boldsymbol{x})}{\partial x_1^2} & \cdots & \dfrac{\partial^2 g_k(\boldsymbol{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 g_k(\boldsymbol{x})}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 g_k(\boldsymbol{x})}{\partial x_n^2} \end{bmatrix}$$

二阶必要条件 Second-Order Necessary Conditions

设x^*是约束条件$h(x)=0$，$g(x)\le 0$下目标函数$f:\mathcal{R}^n\to\mathcal{R}$的局部极小点，其中，$h:\mathcal{R}^n\to\mathcal{R}^m$，$m\le n$, $g:\mathcal{R}^n\to\mathcal{R}^p$, 和 $f,h,g\in C^2$. 若x^*是正则点，则，存在$\lambda^*\in\mathcal{R}^m$和$\mu^*\in\mathcal{R}^p$，使得:

1. $\mu^*\ge 0$

   $Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0^T$

   $\mu^{*T}g(x^*)=0$

   KKT 条件 $D_x l(x^*,\lambda^*,\mu^*)=0$

2. 对所有 $y\in T(x^*)$, 有$y^T L(x^*,\lambda^*,\mu^*)y\ge 0$

   Lagrange function is convex on the tangent space of active constraint functions at $\boldsymbol{x}^*$

证明思路:

x^*是S上局部极小点 $\Leftrightarrow$ x^*是$\tilde{S}$上的局部极小点 $\Leftrightarrow$ 等式约束SONC

积极（等式）约束极小点

$T(x^*)=\{y\in\mathcal{R}^n: Dh(x^*)y=0, Dg_j(x^*)y=0, j\in\mathcal{J}(x^*)\}$

$\mathcal{J}(x^*)\triangleq\{j: g_j(x^*)=0\}$ index set of active inequality constraints

KKT Condition

1. $\mu^*\ge 0$
2. $Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0^T$
3. $\mu^{*T}g(x^*)=0$
4. $h(x^*)=0$
5. $g(x^*)\le 0$

2.2 二阶充分条件 Second-Order Sufficient Conditions

设$f,h,g\in C^2$，且存在可行点 $x^*\in\mathcal{R}^n$ 和向量 $\lambda^*\in\mathcal{R}^m$ 和 $\mu^*\in\mathcal{R}^p$，使得:

1. $\mu^*\ge 0$

   $Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0^T$

   $\mu^{*T}g(x^*)=0$

   KKT Condition $D_x l(x^*,\lambda^*,\mu^*)=0$

2. 对于所有$y\in\tilde{T}(x^*)$, $y\ne 0$, 有$y^T L(x^*,\lambda^*,\mu^*)y>0$ L^*在非零KKT乘子的积极集的切空间上正定

则 x^* 是在约束条件$h(x)=0$, $g(x)\le 0$下目标函数f的严格局部极小点

注：若定理中$\tilde{T}(x^*)=\emptyset$，KKT点x^*是最优解

$\tilde{\mathcal{J}}(x^*)\triangleq\{j: g_j(x^*)=0, \mu_j^*>0\}\subset\mathcal{J}(x^*)$

$\tilde{T}(x^*)=\{y\in\mathcal{R}^n: Dh(x^*)y=0, Dg_j(x^*)y=0, j\in\tilde{\mathcal{J}}(x^*)\}\supset T(x^*)$

二阶充分条件的证明

思路：反证法. 假设x^*不是严格局部极小点
(1) 找到满足条件1的下降方向y，且$y \in \tilde{T}(x^*)$
(2) 使条件2不成立

证明 (1) 假定x^*不是严格局部最优解，则存在可行点列$\{y_k\}$，$y_k \to x^*$，使得

$$f(y_k) \le f(x^*)$$

取$y_k = x^* + \delta_k s_k$，$\|s_k\| = 1$，当$k \to \infty$时，$\delta_k \to 0$

由于$\|s_k\|$有界，故$\{s_k\}$有收敛子列. 不妨设该子列就是$\{s_k\}$，有$s_k \to s^*$

由Taylor 展式得

$k \to \infty$

$$f(y_k) = f(x^*) + \delta_k Df(x^*)s_k + o(\delta_k)$$

$$0 \ge \frac{f(y_k) - f(x^*)}{\delta_k} = Df(x^*)s_k + o(1) \qquad 0 \ge Df(x^*)s_k + o(1) \qquad Df(x^*)s^* \le 0$$

$$0 = h_i(y_k) = h_i(x^*) + \delta_k Dh_i(x^*)s_k + o(\delta_k),\ i \in \mathcal{E}$$

$$0 = \frac{h_i(y_k) - h_i(x^*)}{\delta_k} = Dh_i(x^*)s_k + o(1) \qquad 0 = Dh_i(x^*)s_k + o(1) \qquad Dh_i(x^*)s^* = 0$$

二阶充分条件的证明

证明 (1) 假定x^*不是严格局部最优解，则存在可行点列$\{y_k\}$，$y_k \to x^*$，使得

$$f(y_k) \le f(x^*)$$

取$y_k = x^* + \delta_k s_k$，$\|s_k\| = 1$，当$k \to \infty$时，$\delta_k \to 0$

由于$\|s_k\|$有界，故$\{s_k\}$有收敛子列. 不妨设该子列就是$\{s_k\}$，有$s_k \to s^*$

已证 $Df(x^*)s^* \le 0$ $\qquad$ s^*是下降方向

$$\left.\begin{aligned} &Dh_i(x^*)s^* = 0,\ i \in \mathcal{E} \\ &Dg_j(x^*)s^* = 0,\ j \in \tilde{\mathcal{J}}(x^*) \end{aligned}\right\} s^* \in \tilde{T}(x^*) = \left\{y \in \mathcal{R}^n : Dh(x^*)y = 0, Dg_j(x^*)y = 0, j \in \tilde{\mathcal{J}}(x^*)\right\}$$

由Taylor 展式得

$k \to \infty$

$$0 \ge g_i(y_k) = \underbrace{g_i(x^*)}_{0} + \delta_k Dg_i(x^*)s_k + o(\delta_k),\ i \in \tilde{\mathcal{J}}(x^*)$$

$$0 \ge \frac{g_i(y_k) - g_i(x^*)}{\delta_k} = Dg_i(x^*)s_k + o(1) \qquad 0 \ge Dg_i(x^*)s_k + o(1) \qquad Dg_i(x^*)s^* \le 0$$

若存在某$j \in \tilde{\mathcal{J}}(x^*)$，使 $Dg_j(x^*)s^* < 0$ 由条件1, $D_x l(x^*, \lambda^*, \mu^*) = 0$，得

$$Df(x^*)s^* = -\lambda^{*T} Dh(x^*)s^* - \mu^{*T} Dg(x^*)s^* = -\mu_j^* Dg_j(x^*)s^* > 0$$ 与已证$Df(x^*)s^* \le 0$矛盾

$$\Longrightarrow Dg_j(x^*)s^* = 0$$

证明 (1) 假定x^*不是严格局部最优解，则存在可行点列$\{y_k\}$，$y_k \to x^*$，使得

$$f(y_k) \le f(x^*)$$

取$y_k = x^* + \delta_k s_k$，$\|s_k\| = 1$，当$k \to \infty$时，$\delta_k \to 0$

由于$\|s_k\|$有界，故$\{s_k\}$有收敛子列. 不妨设该子列就是 $\{s_k\}$，有$s_k \to s^*$

已证 $Df(x^*)s^* \le 0$ $\qquad y = s^*$是下降方向

$s^* \in \tilde{T}(x^*) = \{y \in \mathcal{R}^n : Dh(x^*)y = 0, Dg_j(x^*)y = 0, j \in \tilde{\mathcal{J}}(x^*)\}$

---

(2) 拉格朗日函数 $l(y_k, \lambda^*, \mu^*) = f(y_k) + \lambda^{*T} h(y_k) + \mu^{*T} g(y_k) \le f(y_k) \qquad \mu^* \ge 0$

Taylor 展式

$$l(y_k, \lambda^*, \mu^*) = l(x^*, \lambda^*, \mu^*) + \delta_k \underbrace{D_x l(x^*, \lambda^*, \mu^*)}_{0} s_k + \frac{1}{2}\delta_k^2 s_k^T L(x^*, \lambda^*, \mu^*) s_k + o(\delta_k^2)$$

$$= f(x^*) + \frac{1}{2}\delta_k^2 s_k^T L(x^*, \lambda^*, \mu^*) s_k + o(\delta_k^2) \le f(y_k)$$

$$s_k^T L(x^*, \lambda^*, \mu^*) s_k \le 0 \quad \Longleftarrow \quad \frac{1}{2}\delta_k^2 s_k^T L(x^*, \lambda^*, \mu^*) s_k + o(\delta_k^2) \le f(y_k) - f(x^*) \le 0$$

$k \to \infty \qquad y^T L(x^*, \lambda^*, \mu^*) y \le 0 \qquad y = s^* \qquad$ 与条件2 矛盾

因此，x^*是严格局部最优解

$\min f(x), \text{s.t.} \begin{cases} c_i(x) = 0, i \in \mathcal{E} \\ c_i(x) \le 0, i \in \mathcal{J} \end{cases}$ Equality and inequality constrained problem

$L(x, \lambda) = f(x) + \lambda \cdot c(x)$

$\mathcal{E} = \{1,2,..,m\}, m \le n$
$\mathcal{J} = \{m+1, m+2, .., m+p\}$
$\mathcal{J}(x^*) = \{i \in \mathcal{J} | c_i(x^*) = 0\}$
$\tilde{\mathcal{J}}(x^*) = \{j \in \mathcal{J}(x^*) | c_j(x^*) = 0, \mu_j^* > 0\}$

KKT condition

primal feasibility	$c_i(x^*) = 0, i \in \mathcal{E}$
primal feasibility	$c_i(x^*) \le 0, i \in \mathcal{J}$
dual feasibility	$\nabla_x L(x^*, \lambda^*) = 0$
dual feasibility	$\lambda_i^* \ge 0, i \in \mathcal{J}$
complementary feasibility	$\lambda_i^* c_i(x^*) = 0, i \in \mathcal{J}$

$y^T \nabla_{xx}^2 L(x^*, \lambda^*) y \ge 0, \forall y \in T(x^*)$ is positive semi-definite (PSD) → necessary conditions

$y^T \nabla_{xx}^2 L(x^*, \lambda^*) y > 0, \forall y \in \tilde{T}(x^*), y \ne 0$ is positive definite (PD) → sufficient conditions

$$T(x^*) = \left\{ y \,\middle|\, \begin{matrix} \nabla c_i(x^*)^T y = 0, i \in \mathcal{E} \\ \nabla c_i(x^*)^T y = 0, i \in \mathcal{J}(x^*) \end{matrix} \right\}$$

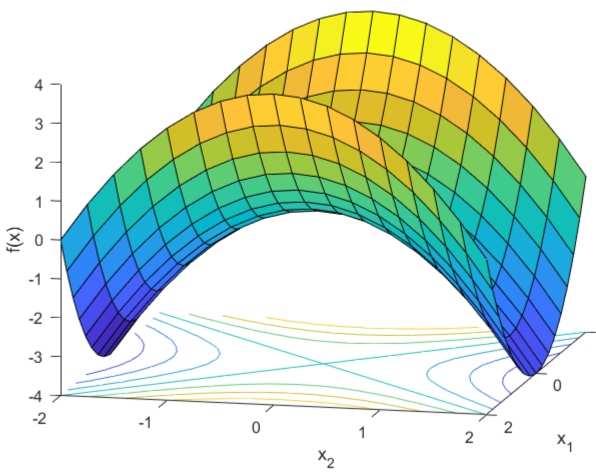
$$\tilde{T}(x^*) = \left\{ y \,\middle|\, \begin{matrix} \nabla c_i(x^*)^T y = 0, i \in \mathcal{E} \\ \nabla c_i(x^*)^T y = 0, i \in \tilde{\mathcal{J}}(x^*) \end{matrix} \right\}$$

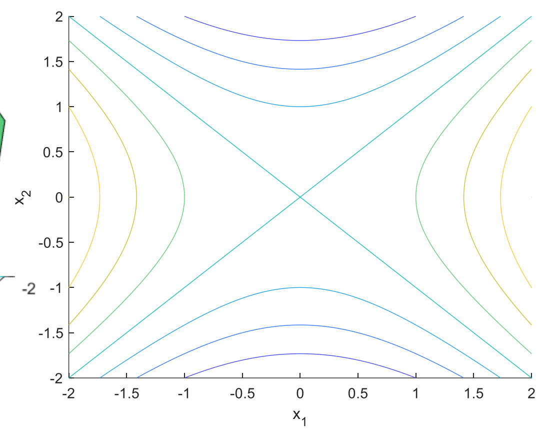
necessary conditions, sufficient conditions → x^* is a local minimizer

例

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

- 对于无约束优化，马鞍面无极值解





例

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

$\text{s.t. } g(x) = -x_2 + x_1^2 \le 0$

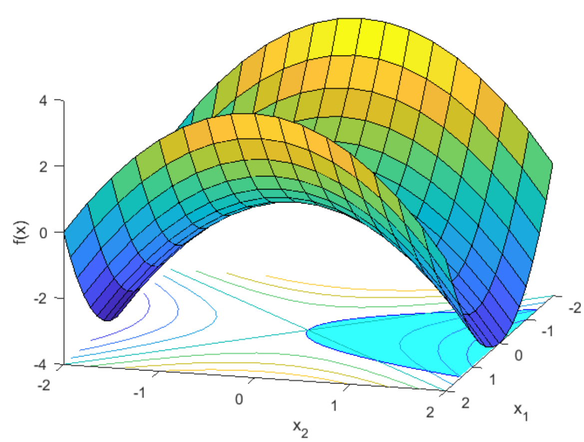
$$l(x,\mu) = x_1^2 - x_2^2 + \mu(-x_2 + x_1^2)$$

$$D_x l(x,\mu) = \begin{bmatrix} 2x_1 + 2\mu x_1 \\ -2x_2 - \mu \end{bmatrix} = 0$$

$$\nabla_{xx}^2 l(x,\mu) = \begin{bmatrix} 2+2\mu & 0 \\ 0 & -2 \end{bmatrix}$$

$$Dg(x) = [2x_1 \quad -1]$$

$$\mu(-x_2 + x_1^2) = 0$$
$$\mu \ge 0$$



例

$$D_x l(x,\mu) = \begin{bmatrix} 2x_1 + 2\mu x_1 \\ -2x_2 - \mu \end{bmatrix} = 0$$

$$\nabla_{xx}^2 l(x,\mu) = \begin{bmatrix} 2+2\mu & 0 \\ 0 & -2 \end{bmatrix}$$

$$Dg(x) = \begin{bmatrix} 2x_1 & -1 \end{bmatrix}$$

$$\mu(-x_2 + x_1^2) = 0$$
$$\mu \geq 0$$

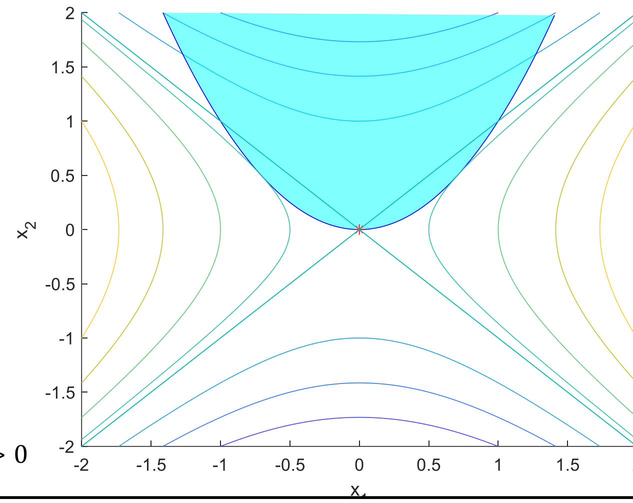
$Dg(x^*) = \begin{bmatrix} 0 & -1 \end{bmatrix}$ 正则点

$$T(x^*) = \left\{ y \middle| \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0 \right\}$$
$$= \left\{ y \middle| \begin{bmatrix} \alpha \\ 0 \end{bmatrix}, \alpha \in \mathcal{R} \right\}$$

$\forall y \in T(x^*),\ \ y \neq 0$

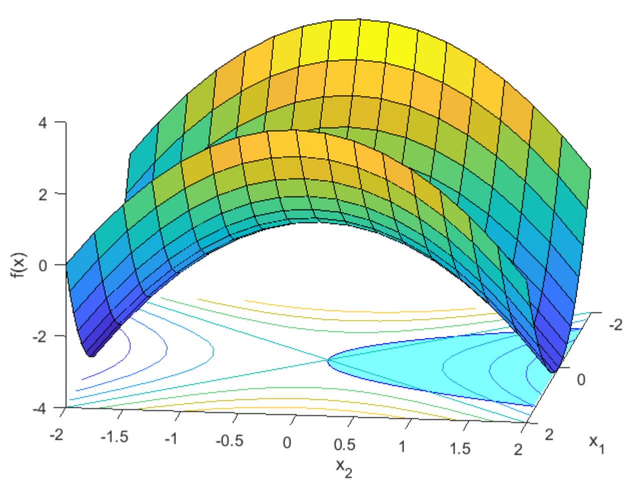
$$y^T L(x^*,\mu^*) y = \begin{bmatrix} \alpha & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = 2\alpha^2 > 0$$

序号	x^*		μ^*	
①	0	0	0	严小




```
figure(2)
hold on
[x,y]=meshgrid(-2:0.25:2);
f=x.^2-y.^2;
surfc(x,y,f)

x1=-sqrt(2):0.01:sqrt(2);
y1=x1.^2;
z1=-4*ones(size(x1));
x1=-sqrt(2):0.01:sqrt(2);
hold on
pic02 = fill3(x1,y1,z1,'c');
%改变边缘和阴影透明度
set(pic02,'edgealpha', 0, 'facealpha', 0.5);
plot3(x1,y1,z1,'-.b')
xlabel('x_1');ylabel('x_2');zlabel('f(x)');
```

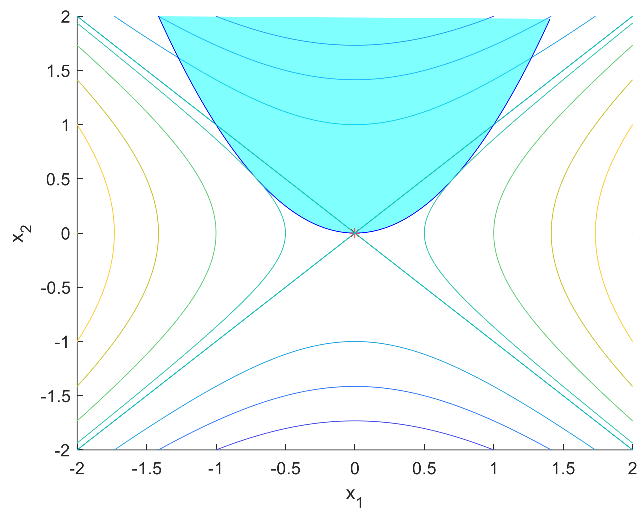


例

```
[x,y]=meshgrid(-2:0.02:2);
f=x.^2-y.^2;
figure(1)
hold on
contour(x,y,f)

x1=-sqrt(2):0.01:sqrt(2);
y1=x1.^2;
hold on
pic01 = fill(x1,y1,'c');
%改变边缘和阴影透明度
set(pic01,'edgealpha', 0, 'facealpha', 0.5);
plot(x1,y1,'b')

x2=0.5;x1=sqrt(x2);
plot(0,0,'*r')
plot(x1,x2,'*r')
plot(-x1,x2,'*r')
vf=x1^2-x2^2;
v=[vf,0];
contour(x,y,f,v)
xlabel('x_1');ylabel('x_2');
```



$$\min f(x) = x_1^2 - x_2^2$$
$$\text{s.t. } g(x) = -x_2 + x_1^2 \le 0$$

使用MATLAB中的约束优化求解函数fmincon

```
function test
fun=@(x)x(1).^2-x(2).^2;
x0=[1;1]; % feasible point
A=[];   % A*x <= b        linear inequality constraints
Aeq=[]; % Aeq*x = beq     linear equality constraints
b=[];
beq=[];
lb=[];  % lb <= x <= ub     bound
ub=[];
options = optimoptions('fmincon','Display','iter','Algorithm','interior-point');
[x,fval,exitflag,output,lambda]=fmincon(fun,x0,A,b,Aeq,beq,lb,ub,@confun,options);
lambda.eqlin
lambda.eqnonlin
lambda.ineqlin
lambda.lower
lambda.upper
lambda.ineqnonlin
end

function [c,ceq]=confun(x)
c = -x(2) + x(1).^2; % nonlinear inequality constraint g(x)<=0
ceq=[];              % nonlinear equality constraint
end
```

Lagrange multipliers at the solution *x*:

- lambda.lower for LB
- lambda.upper for UB
- lambda.ineqlin is for the linear inequalities
- lambda.eqlin is for the linear equalities
- lambda.ineqnonlin is for the nonlinear inequalities
- lambda.eqnonlin is for the nonlinear equalities

$$\min f(x) = x_1^2 - x_2^2$$
$$\text{s.t. } h(x) = -x_2 + x_1^2 = 0$$

使用MATLAB中的约束优化求解函数fmincon

```
function test
fun=@(x)x(1).^2-x(2).^2;
x0=[1;1]; % feasible point
A=[];   % A*x <= b         linear inequality constraints
Aeq=[]; % Aeq*x = beq      linear equality constraints
b=[];
beq=[];
lb=[];  % lb <= x <= ub     bound
ub=[];
options = optimoptions('fmincon','Display','iter','Algorithm','interior-point');
[x,fval,exitflag,output,lambda]=fmincon(fun,x0,A,b,Aeq,beq,lb,ub,@confun,options);
lambda.eqlin
lambda.eqnonlin
lambda.ineqlin
lambda.lower
lambda.upper
lambda.ineqnonlin
End

% nonlinear constraint
function [c,ceq]=confun(x)
c = []; % nonlinear inequality constraint: g(x)<=0
ceq= -x(2) + x(1).^2; % nonlinear equality constraint: h(x) = 0
end
```

Lagrange multipliers at the solution *x*:

- lambda.lower for LB,
- lambda.upper for UB,
- lambda.ineqlin is for the linear inequalities,
- lambda.eqlin is for the linear equalities,
- lambda.ineqnonlin is for the nonlinear inequalities
- lambda.eqnonlin is for the nonlinear equalities

例 切空间$T(x^*) = \emptyset$， KKT点x^*是最优解

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

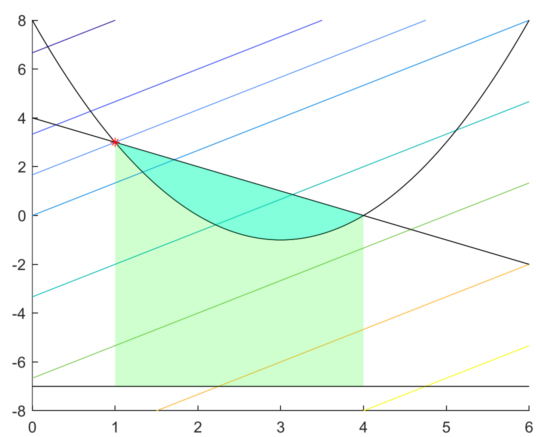
$$l(x,\mu) = 4x_1 - 3x_2 + \mu_1(x_1 + x_2 - 4) + \mu_2(-x_2 - 7) + \mu_3\big((x_1-3)^2 - x_2 - 1\big)$$

$$\nabla_x l(x,\mu) = \begin{bmatrix} 4 + \mu_1 + 2\mu_3(x_1 - 3) \\ -3 + \mu_1 - \mu_2 - \mu_3 \end{bmatrix} = 0$$

$$\begin{aligned} x_1 + x_2 - 4 &\le 0 \\ -x_2 - 7 &\le 0 \\ (x_1-3)^2 - x_2 - 1 &\le 0 \end{aligned}$$

$$\begin{aligned} \mu_1(x_1 + x_2 - 4) &= 0 \\ \mu_2(-x_2 - 7) &= 0 \\ \mu_3\big((x_1-3)^2 - x_2 - 1\big) &= 0 \end{aligned}$$

$$\mu_1, \mu_2, \mu_3 \ge 0$$



$$x^* = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \qquad \mu^* = \begin{bmatrix} 16/3 \\ 0 \\ 7/3 \end{bmatrix}$$

$$x^* = \begin{bmatrix}1\\3\end{bmatrix}, \qquad \mu^* = \begin{bmatrix}16/3\\0\\7/3\end{bmatrix}$$

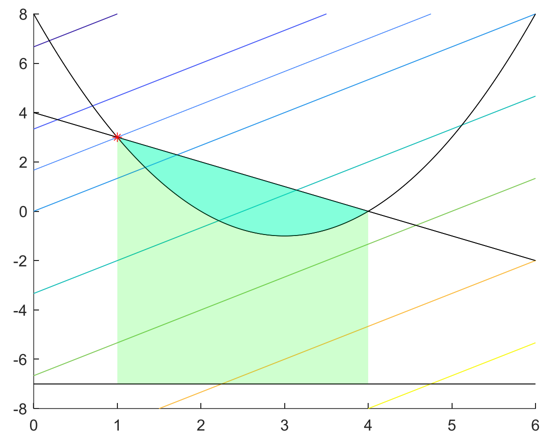
$$\tilde{J}(x^*) = \{1,3\}, \mu_1^*, \mu_3^* > 0$$

$$\begin{bmatrix}Dg_1(x^*)\\Dg_3(x^*)\end{bmatrix} = \begin{bmatrix}-1 & -1\\4 & 1\end{bmatrix}$$

$$\tilde{T}(x^*) = \left\{y \in \mathcal{R}^2 : \begin{bmatrix}-1 & -1\\4 & 1\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix} = 0\right\} = \{\mathbf{0}\} = \emptyset$$

所以， x^*为严格局部最优解

$$L(x^*,\mu^*) = 0 + 0 + 0 + \frac{7}{3}\begin{bmatrix}2 & 0\\0 & 0\end{bmatrix} \succcurlyeq 0$$

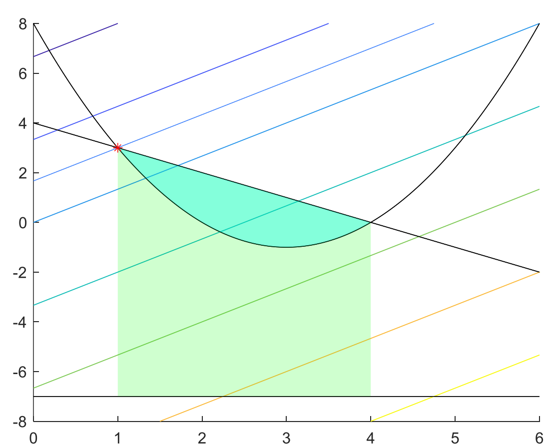



```
clear
clc
figure(1)
hold on
u=0:0.01:6;
[s,t]=meshgrid(u,-8:0.02:8);
f=4*s-3*t;
contour(s,t,f)
xx=1;yy=3;
ff=4*xx-3*yy;
vf=[ff,ff];
contour(s,t,f,vf)

y1=4-u;
y2=(u-3).^2-1;
n=length(y1);
y3=-7*ones(n,1)';
plot(u,y3,'k')
plot(u,y1,'k');
plot(u,y2,'k');

x=1:0.01:4;
y1=4-x; % top-curve
n=length(y1);
y3=-7*ones(n,1)'; % down-curve
pic02 = fill([x,fliplr(x)],[y3,fliplr(y1)],'g');
set(pic02,'edgealpha', 0, 'facealpha', 0.1);
%改变边缘和阴影透明度
y2=(x-3).^2-1; % down-curve
pic01 = fill([x,fliplr(x)],[y2,fliplr(y1)],'c');
set(pic01,'edgealpha', 0, 'facealpha', 0.2);

plot(xx,yy,'*r');
```



例 21.5

考虑 $\min x_1 x_2$

$\text{s.t. } x_1 + x_2 \geq 2$

$x_2 \geq x$□1

a. 写出问题的KKT条件.

b. 找出所有满足KKT条件的点 (及 KKT 乘子)

注意：确定这些点是正则的.

c. 从b 中找出满足 SONC的点.

d. 从c 中找出满足 SOSC的点.

e. 从c 中找出局部极小点.

解：

$$f(x) = x_1 x_2$$
$$g_1(x) = 2 - x_1 - x_2$$
$$g_2(x) = x_1 - x_2$$

$$Df(x) = \begin{bmatrix} x_2 & x_1 \end{bmatrix}$$
$$Dg(x) = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{aligned} &Df(x) + \mu^T Dg(x) \\ &= \begin{bmatrix} x_2 & x_1 \end{bmatrix} + \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} x_2 - \mu_1 + \mu_2 & x_1 - \mu_1 - \mu_2 \end{bmatrix} \end{aligned}$$

KKT 条件

1. $\mu \geq 0$
2. $Df(x) + \mu^T Dg(x) = 0^T$
3. $\mu^T g(x) = 0$
4. NULL
5. $g(x) \leq 0$

$$\begin{aligned} \mu_1, \mu_2 &\geq 0 \\ x_2 - \mu_1 + \mu_2 &= 0 \\ x_1 - \mu_1 - \mu_2 &= 0 \\ \mu_1(2 - x_1 - x_2) &= 0 \\ \mu_2(x_1 - x_2) &= 0 \\ 2 - x_1 - x_2 &\leq 0 \\ x_1 - x_2 &\leq 0 \end{aligned}$$

KKT 条件

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

---

b.找出所有满足KKT条件的点 (及 KKT 乘子) .

注意：确定这些点是正则的.

$$\begin{aligned} \mu_1, \mu_2 &\geq 0 \\ x_2 - \mu_1 + \mu_2 &= 0 \\ x_1 - \mu_1 - \mu_2 &= 0 \\ \mu_1(2 - x_1 - x_2) &= 0 \\ \mu_2(x_1 - x_2) &= 0 \\ 2 - x_1 - x_2 &\leq 0 \\ x_1 - x_2 &\leq 0 \end{aligned}$$

It is easy to check that $\mu_1 \neq 0$ and $\mu_2 \ngtr 0$.

$x_1^* = x_2^* = 1, \mu_1^* = 1, \mu_2^* = 0$. (only one solution)

$Dg_1(\boldsymbol{x}^*) = \begin{bmatrix} -1 & -1 \end{bmatrix}$ and $Dg_2(\boldsymbol{x}^*) = \begin{bmatrix} 1 & -1 \end{bmatrix}$.

Hence, $\boldsymbol{x}^*$ is regular.

c. 从b 中找出满足 SONC的点.

$x_1^* = x_2^* = 1, \mu_1^* = 1, \mu_2^* = 0$

$\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}$，both constraints are active.

Hence, because $\boldsymbol{x}^*$ is regular, $T(\boldsymbol{x}^*) = \{\boldsymbol{0}\}$.

$$\boldsymbol{y}^T L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} = 0$$

This implies that the SONC is satisfied.

$J(\boldsymbol{x}^*) = \{1,2\}$

$J(\boldsymbol{x}^*) \triangleq \{j: g_j(\boldsymbol{x}^*) = 0\}$

$\nabla g_1(\boldsymbol{x}^*) = \begin{bmatrix}-1\\-1\end{bmatrix}$ and $\nabla g_2(\boldsymbol{x}^*) = \begin{bmatrix}1\\-1\end{bmatrix}$ are linearly independent.

$$T(\boldsymbol{x}^*) = \left\{\boldsymbol{y} \in \mathcal{R}^2: \begin{bmatrix}-1 & -1\\ 1 & -1\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix} = 0\right\}$$
$$= \{\boldsymbol{y} \in \mathcal{R}^2: \boldsymbol{y} = \boldsymbol{0}\}$$

$\boldsymbol{y}^T L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} = 0$

SONC

1. $\boldsymbol{\mu}^* \geq \boldsymbol{0}$.
   $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$.
   $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}$.
2. for all $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, we have $\boldsymbol{y}^T L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} \geq 0$.

$T(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathcal{R}^n: D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = 0, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in J(\boldsymbol{x}^*)\}$

d. 从c 中找出满足 SOSC的点.

$$\boldsymbol{F}(x) = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \quad \boldsymbol{G}_1(x) = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \quad \boldsymbol{G}_2(x) = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

$\boldsymbol{h}$ Null $H = 0$

$x_1^* = x_2^* = 1, \mu_1^* = 1, \mu_2^* = 0$

$$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \boldsymbol{F}(\boldsymbol{x}) + [\boldsymbol{\lambda H}(\boldsymbol{x})] + [\boldsymbol{\mu G}(\boldsymbol{x})] = \boldsymbol{F}(\boldsymbol{x}) + 0 + 0$$

$$\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

$$\tilde{T}(\boldsymbol{x}^*) = \left\{\boldsymbol{y} \in \mathcal{R}^2: Dg_1(\boldsymbol{x}^*)\boldsymbol{y} = [-1 \quad -1]\begin{bmatrix}y_1\\y_2\end{bmatrix} = 0\right\}$$
$$= \begin{bmatrix}\alpha\\-\alpha\end{bmatrix}, \alpha \in \mathcal{R}$$

$J(\boldsymbol{x}^*) = \{1,2\} \longrightarrow \mu_1^* = 1 > 0$

$\tilde{J}(\boldsymbol{x}^*) = \{1\}$

Let $\quad \boldsymbol{y} = \begin{bmatrix}\alpha\\-\alpha\end{bmatrix}, \alpha \neq 0$

$D\boldsymbol{g}_1 = [-1 \quad -1]$

but
$$\boldsymbol{y}^T L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y}$$
$$= [\alpha \quad -\alpha]\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\alpha\\-\alpha\end{bmatrix}$$
$$= -2\alpha^2 < 0$$

$\tilde{J}(\boldsymbol{x}^*) \triangleq \{j: g_j(\boldsymbol{x}^*) = 0, \mu_j^* > \boldsymbol{0}\} \subset J(\boldsymbol{x}^*) \triangleq \{j: g_j(\boldsymbol{x}^*) = 0\}$

$\tilde{T}(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathcal{R}^n: D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = 0, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in \tilde{J}(\boldsymbol{x}^*)\}$

不满足SOSC

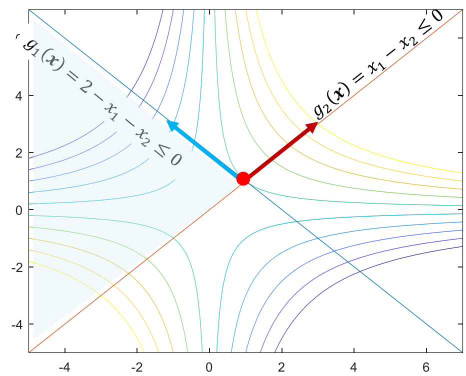
e.从c 中找出局部极小点.

In fact, the point $\boldsymbol{x}^*$ is not a local minimizer.

To see this, draw a picture of the constraint set and level sets of the objective function.

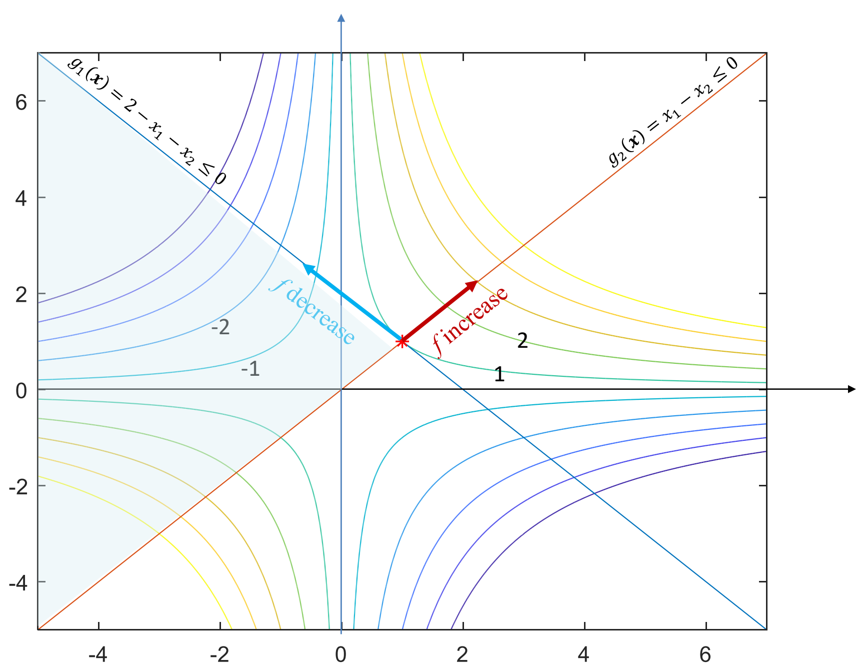
Moving in the feasible direction $[1, 1]^T$, the objective function increases;

but moving in the feasible direction $[-1, 1]^T$, the objective function decreases.



$x_1^* = x_2^* = 1, \mu_1^* = 1, \mu_2^* = 0$

$$\min x_1 x_2$$
$$\text{s.t. } x_1 + x_2 \ge 2$$
$$x_2 \ge x_1$$



1. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$.
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$.
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$.
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$.

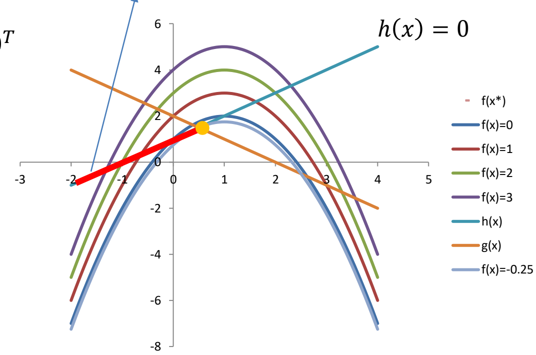
例 21.6 极小化 $f(\boldsymbol{x}) = (x_1 - 1)^2 + x_2 - 2$ 受约束于 $h(\boldsymbol{x}) = x_2 - x_1 - 1 = 0$, $g(\boldsymbol{x}) = x_1 + x_2 - 2 \leq 0$

$$Df(\boldsymbol{x}) = [2x_1 - 2 \quad 1]$$
$$Dh(\boldsymbol{x}) = [-1 \quad 1]$$
$$Dg(\boldsymbol{x}) = [1 \quad 1]$$

则, $\nabla h(\boldsymbol{x})$ 和 $\nabla g(\boldsymbol{x})$ 线性无关, 因而, 所有的可行点都是正则点.

$$Df(\boldsymbol{x}) + \boldsymbol{\lambda}^T D\boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\mu}^T D\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}^T$$

$$\begin{cases} \mu \geq 0 \\ 2x_1 - 2 - \lambda + \mu = 0 \\ 1 + \lambda + \mu = 0 \\ \mu(x_1 + x_2 - 2) = 0 \\ x_2 - x_1 - 1 = 0 \\ x_1 + x_2 - 2 \leq 0 \end{cases} \qquad \begin{cases} \mu^* = 0 \\ \lambda^* = -1 \\ x_1^* = \dfrac{1}{2} \\ x_2^* = \dfrac{3}{2} \end{cases}$$



满足SOSC条件

$$\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*) = \boldsymbol{F}(\boldsymbol{x}^*) + [\lambda^* \boldsymbol{H}(\boldsymbol{x}^*)] + [\mu^* \boldsymbol{G}(\boldsymbol{x}^*)]$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + (-1)\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + (0)\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{cases} \mu^* = 0 \\ \lambda^* = -1 \\ x_1^* = \dfrac{1}{2} \\ x_2^* = \dfrac{3}{2} \end{cases}$$

$$Dh(\boldsymbol{x}^*) = [-1 \quad 1]$$

$$\tilde{\mathcal{J}}(\boldsymbol{x}^*, \mu^*) = \emptyset$$

$\tilde{\mathcal{J}}(\boldsymbol{x}^*) \triangleq \{j : g_j(\boldsymbol{x}^*) = 0, \mu_j^* > \mathbf{0}\} \subset \mathcal{J}(\boldsymbol{x}^*) \triangleq \{j : g_j(\boldsymbol{x}^*) = 0\}$

$\tilde{T}(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathcal{R}^n : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = 0, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in \tilde{\mathcal{J}}(\boldsymbol{x}^*)\}$

$$\tilde{\boldsymbol{T}}(\boldsymbol{x}^*) = \{\boldsymbol{y} \in R^2 : Dh(\boldsymbol{x}^*)\boldsymbol{y} = 0\}$$
$$= \{\boldsymbol{y} \in R^2 : -y_1 + y_2 = 0\}$$
$$= \{[a \quad a]^{\mathrm{T}}, a \in R\}$$

$$\boldsymbol{y} \in \tilde{\boldsymbol{T}}(\boldsymbol{x}^*) = \{[a \quad a]^{\mathrm{T}}, a \in R\}$$

$$\boldsymbol{y}^{\mathrm{T}} \boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)\boldsymbol{y} = [a \quad a]\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ a \end{bmatrix} = 2a^2 > 0$$

根据SOSC, $\boldsymbol{x}^* = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$ 是严格局部极小点

1. Karush-Kuhn-Tucker (KKT) 条件
2. 二阶条件
3. KKT乘子的意义　敏感度分析

Sensitivity Theorem

$f, h, g \in C^2$

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } h(x) = u \quad (P) \\ &\quad\ g(x) \le v \end{aligned} \qquad \begin{aligned} &\min f(x) \\ &\text{s.t. } h(x) = u \quad (P_{uv}) \\ &\quad\ g(x) \le v \end{aligned}$$

设x^*是问题(P)的正则极小点、λ^*, μ^*是对应的KKT乘子，满足二阶充分条件.

则对于问题(P_{uv})存在以$(u, v) = (0,0)$为球心的开球S，使对于任意$(u, v) \in S$，存在问题(P_{uv})的极小值点$x(u, v)$和对应的KKT乘子$\lambda(u, v), \mu(u, v)$. 它们在S中是连续可微的，且 $x(0,0) = x^*, \lambda(0,0) = \lambda^*, \mu(0,0) = \mu^*$，

进一步地，对于任意$(u, v) \in S$，有

$$\begin{aligned} \nabla_u p(u, v) &= -\lambda(u, v) \\ \nabla_v p(u, v) &= -\mu(u, v) \end{aligned} \qquad \text{其中 } p(u, v) = f\big(x^*(u, v)\big)$$

分析

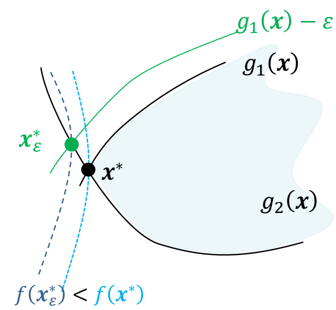
乘子$\boldsymbol{\lambda},\boldsymbol{\mu}$可反映约束条件右端项发生扰动时，
最优目标函数值的变化情况

不失一般性，
只改变第k个不等式约束

$$\begin{aligned} &\min f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ &\qquad \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0} \end{aligned} \qquad P\text{问题}$$

$$\begin{aligned} &\min f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ &\qquad \boldsymbol{g}_\varepsilon(\boldsymbol{x}) \le \boldsymbol{0} \end{aligned} \qquad P_\varepsilon\text{问题}$$

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) \\ \vdots \\ g_p(\boldsymbol{x}) \end{bmatrix} \qquad \boldsymbol{g}_\varepsilon(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) - \varepsilon \\ \vdots \\ g_p(\boldsymbol{x}) \end{bmatrix}$$



设$\boldsymbol{x}^*$是P问题的最优解，$\boldsymbol{x}_\varepsilon^*$是$P_\varepsilon$问题的最优解

$\varepsilon > 0$足够小，使得$\boldsymbol{x}^*$发生足够小的变动到$\boldsymbol{x}_\varepsilon^*$

设$\boldsymbol{x}^*$是P问题的最优解，$\boldsymbol{x}_\varepsilon^*$是$P_\varepsilon$问题的最优解

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j^* \nabla g_j(\boldsymbol{x}^*) = \boldsymbol{0} \qquad \text{KKT条件}$$

不等式积极约束的指标集

$$\mathcal{J} = \mathcal{J}(\boldsymbol{x}^*) \qquad \mathcal{J}_\varepsilon = \mathcal{J}_\varepsilon(\boldsymbol{x}_\varepsilon^*) \longrightarrow \mathcal{J} = \mathcal{J}_\varepsilon \quad \varepsilon > 0\text{足够小}$$

$$g_{j\in\mathcal{J}}(\boldsymbol{x}^*) = 0 \qquad g_{j\in\mathcal{J}_\varepsilon}(\boldsymbol{x}_\varepsilon^*) = 0 \longrightarrow \begin{cases} g_j(\boldsymbol{x}_\varepsilon^*) = 0 & j \in \mathcal{J}, j \ne k \\ g_k(\boldsymbol{x}_\varepsilon^*) - \varepsilon = 0 & j = k \end{cases} \qquad g_k(\boldsymbol{x}_\varepsilon^*) = \varepsilon$$

Taylor展开

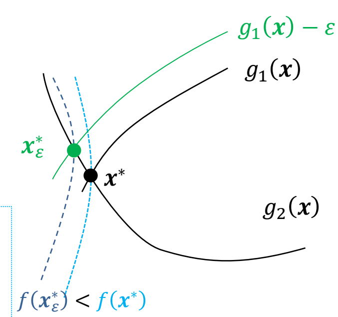
$$f(\boldsymbol{x}_\varepsilon^*) - f(\boldsymbol{x}^*) \approx [\nabla f(\boldsymbol{x}^*)]^T(\boldsymbol{x}_\varepsilon^* - \boldsymbol{x}^*)$$

$$= -\sum_{i=1}^{m} \lambda_i^* [\nabla h_i(\boldsymbol{x}^*)]^T(\boldsymbol{x}_\varepsilon^* - \boldsymbol{x}^*) - \sum_{j=1}^{p} \mu_j^* [\nabla g_j(\boldsymbol{x}^*)]^T(\boldsymbol{x}_\varepsilon^* - \boldsymbol{x}^*)$$

$$\approx -\mu_k^* \varepsilon$$

$$[\nabla h_i(\boldsymbol{x}^*)]^T(\boldsymbol{x}_\varepsilon^* - \boldsymbol{x}^*) \approx h_i(\boldsymbol{x}_\varepsilon^*) - h_i(\boldsymbol{x}^*) = 0 - 0 = 0$$

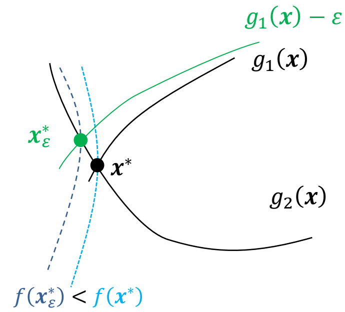
$$[\nabla g_j(\boldsymbol{x}^*)]^T(\boldsymbol{x}_\varepsilon^* - \boldsymbol{x}^*) \approx g_j(\boldsymbol{x}_\varepsilon^*) - g_j(\boldsymbol{x}^*) = \begin{cases} 0 & j \ne k \\ \varepsilon & j = k \end{cases}$$



$$\min f(\boldsymbol{x})$$
$$\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \qquad P\text{问题}$$
$$\boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}$$

$$\min f(\boldsymbol{x})$$
$$\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \qquad P_\varepsilon\text{问题}$$
$$\boldsymbol{g}_\varepsilon(\boldsymbol{x}) \le \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) \\ \vdots \\ g_p(\boldsymbol{x}) \end{bmatrix} \qquad \boldsymbol{g}_\varepsilon(\boldsymbol{x}) = \begin{bmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) - \varepsilon \\ \vdots \\ g_p(\boldsymbol{x}) \end{bmatrix}$$



设$\boldsymbol{x}^*$是P问题的最优解，$\boldsymbol{x}_\varepsilon^*$是$P_\varepsilon$问题的最优解

$\varepsilon > 0$足够小，使得$\boldsymbol{x}^*$发生小变动到$\boldsymbol{x}_\varepsilon^*$

$$f(\boldsymbol{x}_\varepsilon^*) - f(\boldsymbol{x}^*) \approx -\mu_k^* \varepsilon$$

权因子的意义：不等式约束右端项增加ε，目标函数最优值降低ε的μ倍

较大的拉格朗日乘子对应的约束条件发生扰动，
对最优目标函数值变化的影响也较大

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j^* \nabla g_j(\boldsymbol{x}^*) = \boldsymbol{0} \qquad \text{KKT条件}$$

$$\frac{\partial f}{\partial h_i} = -\lambda_i, \qquad \frac{\partial f}{\partial g_j} = -\mu_j$$

权因子的意义：不等式约束右端项增加ε，目标函数最优值降低ε的μ倍
等式约束右端项增加ε，目标函数最优值降低ε的λ倍

注：实际中，由于$\lambda(u,v)$，$\mu(u,v)$是(u,v)的函数
最优值的变化量与λ^*, μ^*的关系不能简单理解为线性关系

进一步说

$$f,h,g\in C^2\qquad \begin{aligned}&\min f(x)\\ &\text{s.t. } h(x)=u\\ &\quad g(x)\le v\end{aligned}\quad (P)\qquad\qquad \begin{aligned}&\min f(x)\\ &\text{s.t. } h(x)=u\\ &\quad g(x)\le v\end{aligned}\quad (P_{uv})$$

$p(u,v)$为问题(P_{uv})的最优值函数，其中$p(0,0)$为问题(P)最优值

1. 若问题(P)为凸，则$p(u,v)$为(u,v)的凸函数
2. 若问题(P)为凸，对偶间隙为零，λ^*,μ^*为问题(P)对偶最优解，则

$$p(u,v)\ge p(0,0)-\lambda^{*T}u-\mu^{*T}v$$

① 若μ_i^*很大，且加紧第i次不等式约束，即$v_i<0$则$p(u,v)$可能会急剧上升

② 若λ_i^*很大正值，且$u_i<0$；或λ_i^*为负值，即绝对值很大，且$u_i>0$，则$p(u,v)$可能会急剧上升

③ 若μ_i^*很小，且$v_i>0$，则最优值变化不大

④ 若λ_i^*为很小正值，且$u_i>0$或λ_i^*为绝对值很小的负值，且$u_i<0$，则最优值变化不大

进一步说

$$f,h,g\in C^2\qquad \begin{aligned}&\min f(x)\\ &\text{s.t. } h(x)=u\\ &\quad g(x)\le v\end{aligned}\quad (P)\qquad\qquad \begin{aligned}&\min f(x)\\ &\text{s.t. } h(x)=u\\ &\quad g(x)\le v\end{aligned}\quad (P_{uv})$$

$p(u,v)$为问题(P_{uv})的最优值函数，其中$p(0,0)$为问题(P)最优值

3. 若问题(P)为凸，对偶间隙为零，且$p(u,v)$在$(u,v)=(0,0)$处可微，则

$$\mu_i^*=-\frac{\partial p(0,0)}{\partial v_i},\qquad \lambda_i^*=-\frac{\partial p(0,0)}{\partial u_i}$$

即根据泰勒展开，得到

$$p(u,v)\approx p(0,0)-\lambda^{*T}u-\mu^{*T}v$$

EXERCISES

21.1 Consider the optimization problem

$$\begin{aligned} &\min x_1^2 + 4x_2^2 \\ &\text{s.t. } x_1^2 + 2x_2^2 \geq 4 \end{aligned}$$

a. Find all the points that satisfy the KKT conditions.

b. Apply the SOSC to determine the nature of the critical points from the previous part.

21.12 Consider the *quadratic programming problem*

$$\begin{aligned} &\min \frac{1}{2} \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \end{aligned}$$

where $\boldsymbol{Q} = \boldsymbol{Q}^T > 0$, $A \in \mathcal{R}^{m \times n}$, and $\boldsymbol{b} \geq \boldsymbol{0}$. Find all points satisfying the KKT condition.

21.5 Consider the problem

$$\begin{aligned} &\min\ x_2 - (x_1 - 2)^3 + 3 \\ &\text{s.t. } \ x_2 \geq 1 \end{aligned}$$

where x_1 and x_2 are real variables. Answer each of the following questions, making sure that you give complete reasoning for your answers.

a. Write down the KKT condition for the problem, and find all points that satisfy the condition. Check whether or not each point is regular.

b. Determine whether or not the point(s) in part a satisfy the second-order necessary condition.

c. Determine whether or not the point(s) in part b satisfy the second-order sufficient condition.

21.7 Consider the problem of optimizing (either minimizing or maximizing) $(x_1 - 1)^2 + (x_2 - 1)^2$ subject to

$$x_2 - x_1^2 \geq 0$$
$$2 - x_1 - x_2 \geq 0$$
$$x_1 \geq 0$$

The point $\boldsymbol{x}^* = \mathbf{0}$ satisfies the KKT conditions.

a. Does $\boldsymbol{x}^*$ satisfy the FONC for minimization or maximization? What are the KKT multipliers?

b. Does $\boldsymbol{x}^*$ satisfy the SOSC? Carefully justify your answer.

QUICK QUIZ

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

a. 写出问题的KKT条件.
b. 找出所有满足KKT条件的点 (及 KKT 乘子)
c. 从b 中找出满足 SONC的点.
d. 从c 中找出满足 SOSC的点.
e. 从c 中找出局部极小点.