第三章 非线性规划

- 3.1 例子和预备知识
- 3.2 凸集、凸函数和凸规划
- 3.3 非线性规划的库恩-塔克定理
- 3.4 单变量极值问题的解法
- 3.5 无约束极值问题的解法
- 3.6 罚函数法
- 3.7 线性约束下线性逼近的方法

3.1 例子和预备知识

- 在科学管理和其他领域中，很多实际问题可归结为线性规划问题. 但也有很多问题，其目标函数或约束条件不是线性函数. 如果目标函数或约束条件中含有非线性函数，就称这种问题为非线性规划问题.
- 解这类问题需要用非线性规划方法. 非线性规划已成为运筹学一个重要分支，在最优设计、管理科学、系统控制等许多领域获得越来越广泛的应用.
- 一般说来，由于非线性函数的复杂性，解非线性规划问题远比解线性规划问题困难得多. 而且，也不像线性规划那样有单纯形法等通用方法. 非线性规划目前还没有一般性算法，各个方法都有自己特定的适用范围.

1. 非线性规划的标准形式

$$\text{(P)} \qquad \min f(x), x \in \mathbb{R}^n$$

$$\begin{cases} g_1(x) & \geq & 0 \\ g_2(x) & \geq & 0 \\ \cdots & \cdots & \\ g_m(x) & \geq & 0 \end{cases}$$

例1. 给定一根长400米的绳子，用来围成一块矩形菜地，问长和宽各为多少米时，可以使菜地的面积最大?

解：设长为x_1米，宽为x_2米. 规划问题为

$$\max x_1 x_2$$

$$\begin{cases} 2(x_1+x_2) & = & 400 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \end{cases}$$

$\Longleftrightarrow$

$$\min -x_1 x_2$$

$$\begin{cases} 2(x_1+x_2) & \geq & 400 \\ -[2(x_1+x_2)] & \geq & 400 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \end{cases}$$

2．极值点存在的必要条件

设 $f : R \subseteq \mathbb{R}^n \to \mathbb{R}$, $x_0 \in R$. 如果 f 在 x_0 取得局部极值且在 x_0 各偏导数存在，则必有

$$\frac{\partial f(x_0)}{\partial x_1} = \frac{\partial f(x_0)}{\partial x_2} = \cdots = \frac{\partial f(x_0)}{\partial x_n} = 0$$

或

$$\nabla f(x_0) = 0$$

上式中

$$\nabla f(x) = (\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n})^T$$

极值点存在的充分条件

设f在 $R \subseteq \mathbb{R}^n$ 上有连续的二阶偏导函数,$x_0 \in R$. 若 $\nabla f(x_0) = 0$ ，且对任何 $0 \neq y \in \mathbb{R}^n$ 有

$$y^T H(x_0) y > 0$$

则 x_0 为*f*的严格局部极小点. 类似的，若有 $y^T H(x_0) y < 0$, 则 x_0 为*f*的严格局部极大点.

这里 $H(x_0)$ 为*f*在点 x_0 处的海赛(Hesse)矩阵

$$H(x_0) = \left. \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial^2 x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial^2 x_n} \end{bmatrix} \right|_{x_0}$$

3.2 凸函数和凸规划

- 定义

设$f(x)$ 为定义在n维欧式空间$\mathbb{R}^n$中的某个凸集S上的函数，若对任何实数0< α<1以及 $x,y\in S$ ，恒有

$$f(\alpha x+(1-\alpha)y)\leq \alpha f(x)+(1-\alpha)f(y)$$

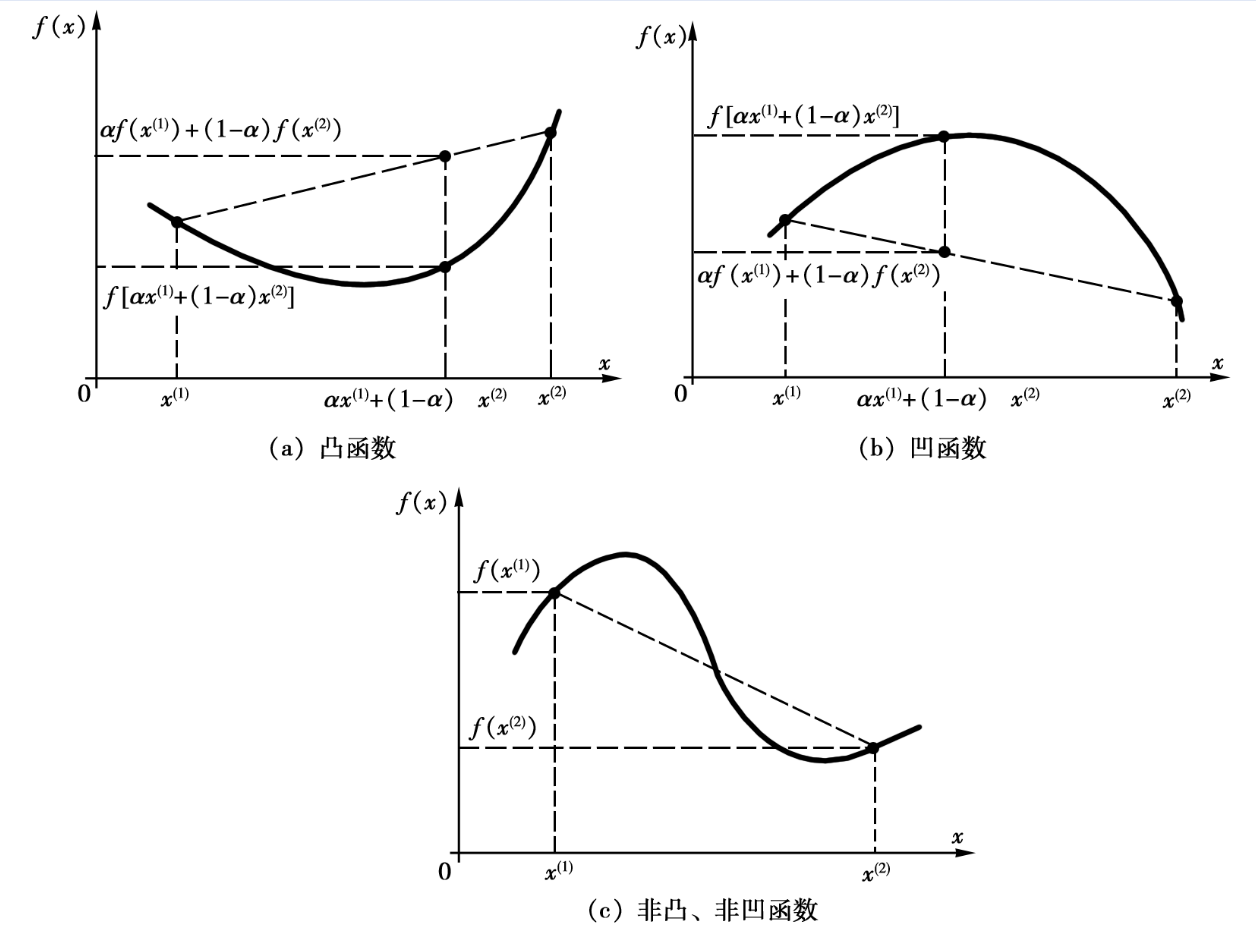
则称 $f(x)$为定义在S上的凸函数.

若对任何实数0< α<1以及 $x\neq y\in S$, 恒有

$$f(\alpha x+(1-\alpha)y)< \alpha f(x)+(1-\alpha)f(y)$$

则称 $f(x)$为定义在S上的严格凸函数.

类似可以定义凸集S上的凹函数.　显然, f是凸函数 $\Longleftrightarrow$ $-f$是凹函数.



凸函数的基本性质:

给定凸集 $S \subseteq E^n$

1. f(x)为S上的凸函数当且仅当对于$x, y \in S$， 一元函数

$$h(\alpha) := f(\alpha x + (1-\alpha)y) = f(y + \alpha(x-y))$$

是[0, 1]上的凸函数.

2. 若 $f_1, f_2, \cdots, f_m$ 为S上的凸函数，那么它们的任意非负线性组合

$$k_1 f_1 + k_2 f_2 + \cdots + k_m f_m$$

$k_i \geq 0$, 仍然是S上的凸函数.

3. 设f(x)为定义在凸集S上的凸函数，则对任意实数β，集合

$$S_\beta = \{x \in S | f(x) \leq \beta\}$$

是凸集.

可微凸函数的判别准则

给定凸区域 $S \subseteq \mathbb{R}^n$，

定理1. 假设f是S上的有连续一阶偏导数的函数，那么f是凸函数的充分必要条件是：对于任意 $x, y \in S$，有

$$f(x) + \nabla f(x) \cdot (y - x) \le f(y).$$

- $z = f(x)$ 和 $z = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$ 分别给出函数图像和函数在 $(x_0, f(x_0))$处切面方程. 定理是说对于凸函数，切面总在函数图像的下方.

定理2．设f有连续的二阶偏导函数，那么f是S上的凸函数的充分必要条件是：对于任意$x \in S$，f的Hesse矩阵

$$H(x_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial^2 x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial^2 x_n} \end{bmatrix} \Bigg|_{x_0}$$

是半正定的.

- 一个n阶方阵A称为半正定的，如果对所有$x \in \mathbb{R}^n$, 有 $x^T A x \geq 0$ 成立.

一般而言，函数的局部极小值并不一定等于它的最小值，前者只不过反映了函数的局部性质。而最优化的目的，往往是要求函数在整个域中的最小值(或最大值). 为此, 必须将所得的全部极小值进行比较(如果有边界，还需考虑边界值)，以便从中选出最小者. 然而，对于定义在凸集上的凸函数来说，则用不着进行这种麻烦的工作，它的局部极小值就等于其最小值.

定理3. 若$f(X)$为定义在凸集R上的凸函数，则它的任一极小点也是它在R上的最小点（全局极小点），而且极小点集为凸集.

证明：设x_0是一个局部极小点，则对于充分小的邻域 $B_\delta(x_0)$ 中的x，均有

$$f(x_0) \leq f(x)$$

对于任意 $y \in R$，对于充分小的λ，有 $\lambda y + (1-\lambda)x_0 \in B_\delta(x_0)$. 结合凸函数性质，有

$$\lambda f(y) + (1-\lambda)f(x_0) \geq f(\lambda y + (1-\lambda)x_0) \geq f(x_0)$$

整理、移项后除以λ，得

$$f(y) \geq f(x_0).$$

这就是说, x_0 是全局极小点.

定义. 考虑标准形式的规划问题(P):

$$\min f(x), x \in E^n$$

$$\begin{cases} g_1(x) & \geq & 0 \\ g_2(x) & \geq & 0 \\ \cdots & \cdots & \\ g_m(x) & \geq & 0 \end{cases}$$

如果f为凸函数，$g_i(x), i = 1, 2, \cdots, m$ 为凹函数，那么称规划问题(P) 为凸规划.

称集合 $R = \{x \in \mathbb{R}^n | g_i(x) \geq 0, i = 1, 2, \cdots, m\}$
为约束集合或可行解集.

凸规划是一类简单而又具有重要意义的非线性规划. 线性规划也属于凸规划.

例2. 分析非线性规划问题

$$\min f(x) = x_1^2 + 2x_2^2 - 4x_1 + 6$$

$$\begin{cases} g_1(x) & = & 2x_1 - x_2 + 1 \geq 0 \\ g_2(x) & = & -x_1^2 + x_2 - 3 \geq 0 \\ x_1, x_2 & \geq & 0 \end{cases}$$

解：目标和约束函数都是光滑函数.

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial^2 x_2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

是正定的，故f是凸函数.

$$H(g_2) = \begin{bmatrix} \frac{\partial^2 g_2}{\partial^2 x_1} & \frac{\partial^2 g_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 g_2}{\partial x_2 \partial x_1} & \frac{\partial^2 g_2}{\partial^2 x_2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$$

是半负定，故 g_2是凹函数. 所以考虑问题是凸规划.

定理4. 设规划问题(P)为凸规划，那么

- 可行解集R为凸集.
- 任何局部最优解即为全局最优解.
- 其最优解的集合形成一个凸子集R*.
- 当目标函数f(X)为严格凸函数, 且R*非空，那么最优解必定唯一.

定理5. 设(P)为凸规划问题且目标函数f有连续一阶偏导数，则 $x_0 \in R$ 是最优解的充分必要条件是：对于任意 $x \in R$ ，有

$$\nabla f(x_0) \cdot (x - x_0) \geq 0$$

3.3 库恩-塔克定理

- 非线性规划中的Kuhn-Tucker定理是非线性规划中重要工具. 它是Lagrange乘数法的推广.
- Kuhn-Tucker条件是某点为非线性规划最优解的一个必要条件（在一定正则性假设下）.
- 对于凸规划，它既是最优点存在的必要条件，同时也是充分条件.

Lagrange乘数法回顾

求满足约束的极值问题

$$\begin{cases} \min(\max) f(x), & x \in \mathbb{R}^n \\ g_i(x) = 0, & i = 1, 2, \cdots, m \end{cases}$$

定义辅助函数 $\varphi(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$. 在一定正则性假设下，$\bar{x}$ 是该问题解的一个必要条件是：

存在 $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \cdots, \bar{\lambda}_m)^T$ 满足Lagrange条件

$$\begin{cases} \nabla_x \varphi(\bar{x}, \bar{\lambda}) = 0 \\ \nabla_\lambda \varphi(\bar{x}, \bar{\lambda}) = 0 \end{cases}$$

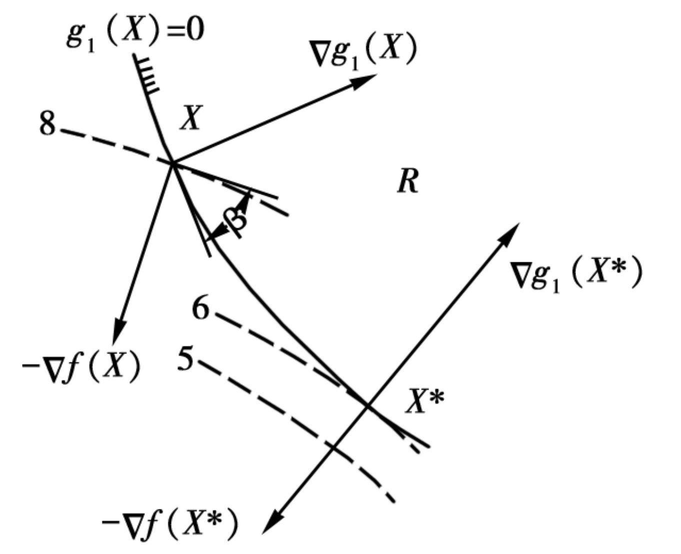
即

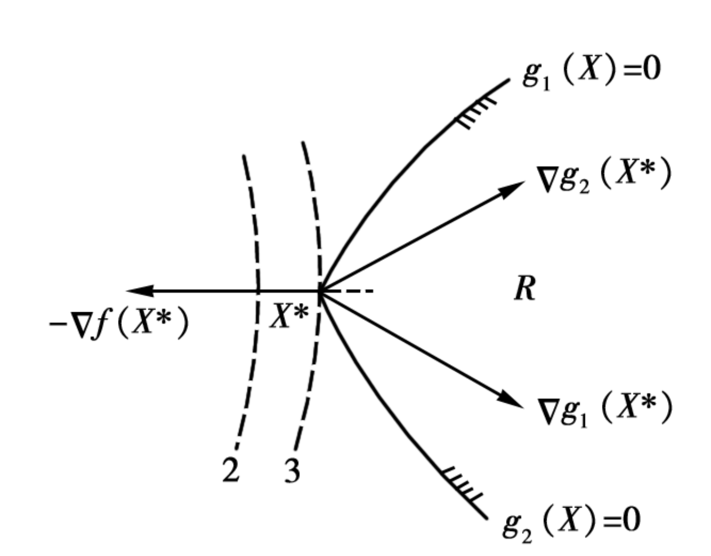
$$\begin{cases} \nabla f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}) = 0 \\ g_i(\bar{x}) = 0 \qquad i = 1, 2, \cdots, m \end{cases}$$

非线性规划示意：设n=m=2

$$\begin{cases} \min f(x), & x \in \mathbb{R}^2 \\ g_1(x) \geq 0, & g_2(x) \geq 0 \end{cases}$$

假设x^*是最优解.





a. 若 $g_1(x^*) = 0, g_2(x^*) \neq 0$ 那么有满足 $\mu_1^* \geq 0$,

$$\nabla f(x^*) = \mu_1^* \nabla g_1(x^*)$$

另一种情形类似.

b. 若 $g_1(x^*) = g_2(x^*) = 0$，那么存在 $\mu_1^* \geq 0, \mu_2^* \geq 0$，满足

$$\nabla f(x^*) = \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*)$$

Kuhn-Tucker (K-T) 条件

考虑辅助函数 $\varphi(x,\mu)=f(x)-\sum\limits_{i=1}^{m}\mu_i g_i(x)$. 称 $(\bar{x},\bar{\mu})\in\mathbb{R}^n\times\mathbb{R}^m$ 满足 (K-T) 条件，如果

$$\begin{cases} \nabla_x\varphi(\bar{x},\bar{\mu}) & = & 0 \\ \nabla_\mu\varphi(\bar{x},\bar{\mu}) & \leq & 0 \\ \bar{u} & \geq & 0 \\ \bar{u}^T\nabla_\mu\varphi(\bar{x},\bar{\mu}) & = & 0 \end{cases}$$

成立.

等价的，

$$\begin{cases} \nabla f(\bar{x})-\sum_{i=1}^{m}\bar{\mu}_i\nabla g_i(\bar{x})=0 & & \\ g_i(\bar{x})\geq 0 & i= & 1,2,\cdots,m \\ \bar{\mu_i}\geq 0, & i= & 1,2,\cdots,m \\ \bar{\mu_i}g_i(\bar{x})=0 & i= & 1,2,\cdots,m \end{cases}$$

- 有的文献也称Karush-Kuhn-Tucker (KKT)条件.

考虑标准形式的凸规划

$$\begin{cases} \min f(x) & , \quad x \in \mathbb{R}^n \\ g_i(x) \geq 0 & , \quad i = 1, 2, \cdots, m \end{cases}$$

定理6. 假设 $f, g_i, i = 1, 2, \cdots, m$ 有连续的一阶偏导函数. 如果 $(\bar{x}, \bar{\mu})$ 满足(K-T) 条件，那么 $\bar{x}$ 是凸规划问题的最优解.

定理7. 假设如上. R为可行解集. 如果存在 $x_0 \in R$ 满足

$$g_i(x_0) > 0, i = 1, 2, \cdots, m$$

那么如果 $\bar{x}$ 是凸规划问题的最优解, 则一定存在 $\bar{u}$，使得 $(\bar{x}, \bar{\mu})$ 满足(K-T) 条件.

- 上述定理说明：在一定的正则性假设下，如函数的可导性、可行解集非”蜕化”性, 那么 $\bar{x}$ 是凸规划的最优解的充分必要条件是它满足(K-T)条件.

定理6证明：假设 $(\bar{x},\bar{\mu})$ 满足(K-T) 条件. 令

$$E(\bar{x})=\{i\mid g_i(\bar{x})=0,1\leq i\leq m\}$$

若$i\notin E(\bar{x})$, 有 $g_i(\bar{x})>0$. 由(K-T) 条件, 得 $\bar{\mu_i}=0$, 以及

$$\nabla f(\bar{x})=\sum_{i=1}^{m}\bar{\mu_i}\nabla g_i(\bar{x})=\sum_{i\in E(\bar{x})}^{m}\bar{\mu_i}\nabla g_i(\bar{x})$$

对$x\in R$, 有

$$\nabla f(\bar{x})\cdot(x-\bar{x})=\sum_{i\in E(\bar{x})}^{m}\bar{\mu_i}\nabla g_i(\bar{x})\cdot(x-\bar{x})$$

注意到g_i为凹函数，有

$$0\leq g_i(x)\leq g_i(\bar{x})+\nabla g_i(\bar{x})\cdot(x-\bar{x})=\nabla g_i(\bar{x})\cdot(x-\bar{x}).$$

从而

$$\nabla f(\bar{x})\cdot(x-\bar{x})\geq 0$$

由定理5，$\bar{x}$ 为凸规划的最优解.

例3. 用库恩－塔克条件解非线性规划

$$\begin{cases} \min\ f(x) = (x-3)^2 \\ \qquad 0 \le x \le 5 \end{cases}$$

解： 先将该非线性规划问题写成以下形式

$$\begin{cases} \min\ f(x) = (x-3)^2 \\ \qquad g_1(x) = x \ge 0 \\ \qquad g_2(x) = 5 - x \ge 0 \end{cases}$$

写出其目标函数和约束函数的梯度

$$\nabla f(x) = 2(x-3),$$
$$\nabla g_1(x) = 1,\ \nabla g_2(x) = -1$$

对第一个和第二个约束条件分别引入K-T乘子，设K-T点为X*，则可以得到该问题的K-T条件.

该问题的K-T条件为：

$$\begin{cases} 2(\bar{x}-3)-\bar{\mu}_1+\bar{\mu}_2 & = & 0 \\ \bar{x} & \geq & 0 \\ 5-\bar{x} & \geq & 0 \\ \bar{\mu}_1,\bar{\mu}_2 & \geq & 0 \\ \bar{\mu}_1\bar{x} & = & 0 \\ \bar{\mu}_2(5-\bar{x}) & = & 0 \end{cases}$$

为解上述方程组，考虑以下几种情形

(1) 若 $\bar{\mu}_1 \neq 0, \bar{\mu}_2 \neq 0,$ 无解.
(2) 若 $\bar{\mu}_1 \neq 0, \bar{\mu}_2 = 0,$ 得 $\bar{x}=0, \bar{\mu}_1=-6,$ 不是K-T点.
(3) 若 $\bar{\mu}_1 = 0, \bar{\mu}_2 \neq 0,$ 得 $\bar{x}=5, \bar{\mu}_2=-4,$ 不是K-T点.
(4) 若 $\bar{\mu}_1 = \bar{\mu}_2 = 0,$ 得 $\bar{x}=3$，此为K-T点，其目标函数值 $f(\bar{x})=0.$

由于该非线性规划问题为凸规划，故 $\bar{x}=3$ 就是其全局极小点. 该点是可行域的内点，它也可直接由梯度等于零的条件求出.

例4. 考虑凸规划问题

$$\begin{cases} \min\,(-x_1 - x_2) & \\ \qquad 3x_1^2 + x_2^2 & \leq \quad 1 \end{cases}$$

解:

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

K-T条件为

$$\begin{cases} -1 + 6\bar{\mu}\bar{x}_1 & = & 0 \\ -1 + 2\bar{\mu}\bar{x}_2 & = & 0 \\ 1 - 3\bar{x}_1^2 - \bar{x}_2^2 & \geq & 0 \\ \bar{\mu} & \geq & 0 \\ \bar{\mu}(1 - 3\bar{x}_1^2 - \bar{x}_2^2) & = & 0 \end{cases}$$

若 $\bar{\mu} = 0,$ 则 $-1 = 0$ ，矛盾.
若 $\bar{\mu} > 0,$ 则 $1 - 3\bar{x}_1^2 - \bar{x}_2^2 = 0$. 由前两式得 $\bar{x}_2 = 3\bar{x}_1 > 0$. 故可解得

$$\bar{x}_1 = \sqrt{\frac{1}{12}}, \bar{x}_2 = 3\sqrt{\frac{1}{12}}, \bar{\mu} = \frac{\sqrt{3}}{3}.$$

凸规划最优解为 $\bar{x}_1 = \sqrt{\frac{1}{12}}, \bar{x}_2 = 3\sqrt{\frac{1}{12}}.$