

第十讲：优化问题

凸优化问题及其性质

杨 林

大 纲

1. 优化问题及标准形式
2. 最优性条件
3. 优化问题的等价性
4. 拟凸优化

大 纲

1. 优化问题及标准形式

2. 最优性条件

3. 优化问题的等价性

4. 拟凸优化

1 优化问题及标准形式

■ 定义优化问题

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, \quad i = 1, 2, \dots, m$
	$h_i(x) = 0, \quad i = 1, 2, \dots, p$

- $x \in \mathbb{R}^n$ 是优化变量
- $f_0: \mathbb{R}^n \rightarrow \mathbb{R}$ 是目标函数或代价函数
- $f_i: \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, 2, \dots, m$ 是不等式约束函数
- $h_i: \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, 2, \dots, p$ 是等式约束函数

1 优化问题及标准形式

■ **定义1 (隐式约束、显式约束)**: 标准形式优化问题有一个隐式约束

$$x \in \mathcal{D} = \bigcap_{i=0}^m \text{dom } f_i \cap \bigcap_{i=0}^p \text{dom } h_i$$

我们将 \mathcal{D} 称为问题的**定义域**

约束条件 $f_i(x) \leq 0, h_i(x) = 0$ 是**显式约束**

如果问题没有显式约束 ($m = p = 0$) 则称其为**无约束问题**

□ **例 1**:

$$\text{最小化 } f_0(x) = - \sum_{i=1}^k \log(b_i - a_i^T x)$$

上述问题是一个具有隐式约束条件 $a_i^T x \leq b_i$ 的无约束问题

1 优化问题及标准形式

- **定义2(可行解、最优点和局部最优点)**: x 是可行的 (feasible), 如果 $x \in (\cap \text{dom } f_i) \cap (\cap \text{dom } h_i)$ 且它满足所有约束条件, 所有可行的 x 构成的集合称为**可行集**或**约束集**; 可行的 x 是**最优的** (**最优点、最优解**), 如果 $f_0(x) = p^*$; 最优点可以不唯一, 最优点集合定义为 X_{opt} ; x 是**局部最优的**, 如果存在 $R > 0$ 使得 x 对于如下问题是最优的

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$h_i(x) = 0, i = 1, 2, \dots, p$
	$\ z - x\ _2 \leq R$

1 优化问题及标准形式

■ 最优值和最优解

$$p^* = \inf\{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$$

$p^* = \infty$ 如果问题不可行 (没有 x 满足约束条件)

$p^* = -\infty$ 如果问题无下界

□ 例 2: (其中 $n = 1, m = p = 0$)

1. $f_0(x) = 1/x, \mathbf{dom} f_0 = \mathbb{R}_{++}$: $p^* = 0$, 无最优点
2. $f_0(x) = -\log x, \mathbf{dom} f_0 = \mathbb{R}_{++}$: $p^* = -\infty$, 无下界
3. $f_0(x) = x \log x, \mathbf{dom} f_0 = \mathbb{R}_{++}$: $x = 1/e$ 是最优点,
 $p^* = -1/e$
4. $f_0(x) = x^3 - 3x$: 在 $x = 1$ 处局部最优, $p^* = -\infty$, 无下界

1 优化问题及标准形式

■ 定义3 (凸优化问题的标准形式):

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$a_i^T x = b_i, i = 1, 2, \dots, p$

f_0, f_1, \dots, f_m 是凸函数; 等式约束是仿射的

如果 f_0 是拟凸函数 (并且 f_1, \dots, f_m 是凸函数), 则问题是**拟凸的**, 通常写作

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$Ax = b, i = 1, 2, \dots, p$

■ 凸优化问题的可行集是凸集

1 优化问题及标准形式

■ 定义4 (凹最大化问题):

最大化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$a_i^T x = b_i, i = 1, 2, \dots, p$

f_0 是凹函数, f_1, \dots, f_m 是凸函数; 等式约束是仿射的

■ 上述问题也是凸优化问题

1 优化问题及标准形式

■ 定义5(可行性问题):

找到 (Find) x

约束条件 $f_i(x) \leq 0, i = 1, 2, \dots, m$
 $h_i(x) = 0, i = 1, 2, \dots, p$

可以被视为 $f_i(x) = 0$ 的一般问题的一个特例:

最小化 0

约束条件 $f_i(x) \leq 0, i = 1, 2, \dots, m$
 $h_i(x) = 0, i = 1, 2, \dots, p$

$p^* = 0$ 如果可行解存在, 任何可行 x 都是最优的

$p^* = \infty$ 如果可行解不存在

大 纲

1.标准形式的优化问题

2. 最优性条件

3. 等价性

4.拟凸优化

2 最优性条件

■ 定理1 (局部与全局最优):

一个凸优化问题的局部最优点也是(全局)最优的

■ 证明:

假设 x 是局部最优的而 y 是最优的, 则 $f_0(y) < f_0(x)$.

x 局部最优意味着存在一个 $R > 0$, 使得

$$z \text{ 可行}, \|z - x\|_2 \leq R \implies f_0(z) \geq f_0(x)$$

考虑 $z = \theta y + (1 - \theta)x$, 其中 $\theta = R/(2\|y - x\|_2)$

因为 $\|y - x\|_2 > R$ (否则 $f_0(y) \geq f_0(x)$), 所以 $0 < \theta < 1/2$

z 是两个可行点的凸组合, 因此也是可行的. 另外, $\|z - x\|_2 = R/2$, 且有

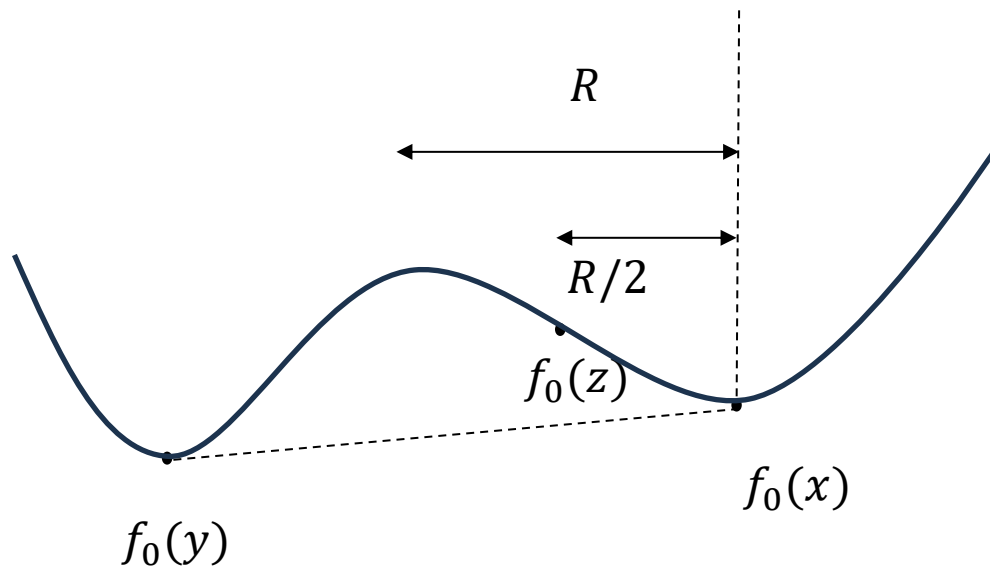
$$f_0(z) \leq \theta f_0(x) + (1 - \theta)f_0(y) < f_0(x)$$

这与我们假设 x 是局部最优的相矛盾

2 最优性条件

■ 定理1 (局部与全局最优):

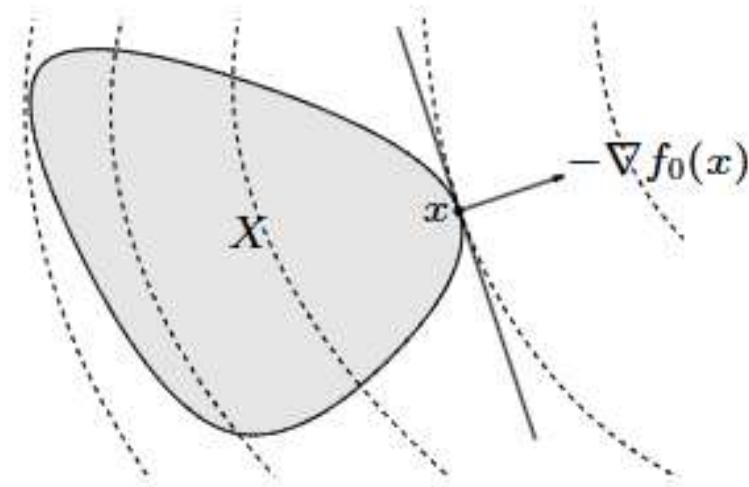
一个凸优化问题的局部最优点也是(全局)最优的



2 最优性条件

- **定理2(可微函数 f 的最优准则)**: x 是最优的当且仅当它是可行的并且对所有可行 $y \in X$ 有

$$\nabla f_0(x)^T (y - x) \geq 0$$



2 最优性条件

■ 定理2的证明:

如果非零, $\nabla f_0(x)^T$ 定义了 X 在 x 处的一个支撑超平面

$$\Rightarrow: f_0(y) \geq f_0(x) + (\nabla f_0(x))^T (y - x) \geq f_0(x)$$

\Leftarrow : 若存在 y , 使得 $(\nabla f_0(x))^T (y - x) < 0$, 则由 $(\nabla f_0(x))^T$ 和 x 确定的超平面与集合 X 的内部相交(过 x 的等高线也与其内部相交), 与 x 是最优点矛盾

2 最优性条件

□ 例 3: 证明 $x^* = (1, 1/2, -1)$ 是优化问题

$$\begin{array}{ll} \text{最小化} & (1/2)x^T Px + q^T x + r \\ \text{约束条件} & -1 \leq x_i \leq 1, i = 1, 2, 3 \end{array}$$

的最优解, 其中

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, r = 1$$

■ 证明: $\nabla f_0(x^*) = (-1, 0, 2)$. 由定理2知:

$$\nabla f_0(x^*)^T (y - x^*) = -(y_1 - 1) + 2(y_3 + 1) \geq 0$$

这对于任意满足 $-1 \leq y_i \leq 1, i = 1, 2, 3$ 的 y 都成立

2 最优性条件

□ 例 4 (无约束条件): x 是最优解当且仅当

$$x \in \mathbf{dom} f_0, \nabla f_0(x) = 0$$

□ 例 5 (等式约束条件): 对于 具有等式约束条件的优化问题:

$$\text{最小化 } f_0(x) = 0 \quad \text{约束条件 } Ax = b$$

x 是最优解当且仅当存在一个 v 使得

$$x \in \mathbf{dom} f_0, \quad Ax = b, \quad \nabla f_0(x) + A^T v = 0$$

其中, $A^T = (\alpha_1, \alpha_2, \dots)$

2 最优性条件

■ 例题5的证明:

x 是最优当且仅当对所有满足 $Ay = b$ 的 y , 有

$$\nabla f_0(x)^T (y - x) \geq 0$$

注意到 $y = x + v$, 其中 $v \in \mathcal{N}(A)$ (零空间)

因此, 上述条件等价于 $\nabla f_0(x)^T v \geq 0, \forall v \in \mathcal{N}(A)$

等价于 $\nabla f_0(x)^T v = 0, \forall v \in \mathcal{N}(A)$ (v 与所有的 A 的行向量垂直)

换句话说, $\nabla f_0(x)^T \in \mathcal{R}(A^T)$, 或 $\exists v, \nabla f_0(x) + A^T v = 0$

2 最优性条件

□ 例 6 (非负正交区域上的最小化):

最小化 $f_0(x)$ 约束条件 $x \geq 0$

x 是最优解当且仅当

$$x \in \mathbf{dom} f_0, x \geq 0, \begin{cases} \nabla f_0(x)_i \geq 0, x_i = 0 \\ \nabla f_0(x)_i = 0, x_i > 0 \end{cases}$$

■ 证明: (充分性证明略)

最优性条件: $x \geq 0, \nabla f_0(x)^T (y - x) \geq 0$, 对 $\forall y \geq 0$ 成立

可以得出 $\nabla f_0(x)^T \geq 0$, 因为我们总能找到 $y - x$ 某一维趋向正无穷, 使得 $\nabla f_0(x)^T (y - x)$ 无下界

令 $y = 0$ (取下界), 最优性条件转化为 $-\nabla f_0(x)^T x \geq 0$

考虑 $x \geq 0$ 和 $\nabla f_0(x)^T \geq 0$, 我们得到 $\nabla f_0(x)^T x = 0$, 或 $(\nabla f_0(x))_i x_i = 0$, 对于 $\forall i$ 必须成立. 这完成了证明

大 纲

1.标准形式的优化问题

2. 最优性条件

3. 等价性

4.拟凸优化

3 等价性

□ 例 7:

最小化	$f_0(x) = x_1^2 + x_2^2$
约束条件	$f_1(x) = x_1/(1 + x_2^2) \leq 0$
	$h_1(x) = (x_1 + x_2)^2 = 0$

f_0 是凸函数; 可行集 $\{(x_1, x_2) | x_1 = -x_2, x_1 \leq 0\}$ 是凸集
不是凸问题 (根据我们的定义): f_1 不是凸函数, h_1 不是仿射函数
等价于 (但不相同于) 凸优化问题:

最小化	$x_1^2 + x_2^2$
约束条件	$x_1 \leq 0$
	$x_1 + x_2 = 0$

3 等价性

- 两个问题(非正式)等价, 如果其中一个问题的解可以轻易地从另一个问题的解中得出, 反之亦然
- 下面是一些常见的等价变换:
 1. 变量变换
 2. 目标函数和约束函数变换
 3. 松弛变量
 4. 消除等式约束
 5. 引入等式约束
 6. 极小化某些变量

3 等价性

1. 变量变换:

设 $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是一一映射, 其像包含了问题的定义域 \mathcal{D} , 即 $\phi(\text{dom } \phi) \supseteq \mathcal{D}$. 我们定义函数 \tilde{f}_i 和 \tilde{h}_i (令 $x = \phi(z)$) 为

$$\tilde{f}_i(z) = f_i(\phi(z)), i = 1, \dots, m \quad \tilde{h}_i(z) = h_i(\phi(z)), i = 1, \dots, p$$

那么标准形式问题等价于

最小化(对 z)	$\tilde{f}_0(z)$
约束条件	$\tilde{f}_i(z) \leq 0, i = 1, \dots, m$
	$\tilde{h}_i(z) = 0, i = 1, \dots, p$

- 如果 x 解决了标准形式的问题, 那么 $z = \phi^{-1}(x)$ 也解决了上述问题; 反之, z 是上述问题的最优解, $x = \phi(z)$ 也解决了标准形式的问题

3 等价性

2. 目标函数和约束函数变换:

设 $\psi_0: \mathbb{R} \rightarrow \mathbb{R}$ 单增; $\psi_1, \dots, \psi_m: \mathbb{R} \rightarrow \mathbb{R}$ 满足: 当且仅当 $u \leq 0$ 时 $\psi_i(u) \leq 0$; $\psi_{m+1}, \dots, \psi_{m+p}: \mathbb{R} \rightarrow \mathbb{R}$ 满足: 当且仅当 $u = 0$ 时 $\psi_i(u) = 0$. 我们定义函数 \tilde{f}_i 和 \tilde{h}_i 为复合函数 (ψ 为任意实现函数变换的算子)

$$\tilde{f}_i(z) = \psi_i(f_i(x)), i = 1, \dots, m$$

$$\tilde{h}_i(z) = \psi_{m+i}(h_i(x)), i = 1, \dots, p$$

那么标准形式问题等价于

最小化

$$\tilde{f}_0(x)$$

约束条件

$$\tilde{f}_i(x) \leq 0, i = 1, \dots, m$$

$$\tilde{h}_i(x) = 0, i = 1, \dots, p$$

3 等价性

□ 例 8 (最小函数和最小范数平方问题) :

考虑无约束的Euclid范数极小化问题

$$\text{最小化} \quad \|Ax - b\|_2$$

■ 证明：因为范数总是非负的,标准形式问题等价于

$$\text{最小化} \quad \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b)$$

3 等价性

3. 松弛变量:

通过观察可以得到一个简单的变换,即 $f_i(x) \leq 0$ 等价于存在一个 $s_i \geq 0$ 满足 $f_i(x) + s_i = 0$. s_i 称为对应元不等式约束 $f_i(x) \leq 0$ 的**松弛变量**

那么标准形式问题等价于

最小化
约束条件

$$\begin{aligned} & f_0(x) \\ & s_i \geq 0, i = 1, \dots, m \\ & f_i(x) + s_i = 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{aligned}$$

■ 什么时候是保凸变换? (LP)

3 等价性

4. 消除等式约束（保凸变换）：

如果我们可以用一些参数 $z \in \mathbb{R}^k$ 来显示的参数化等式约束

$$h_i(x) = 0, i = 1, \dots, p$$

的解,那么我们可以从原问题中消除等式约束.设函数 $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^n$ 是这样的函数,存在一些 $z \in \mathbb{R}^k$ 使得 $x = \phi(z)$, 那么标准形式问题等价于

最小化

$$\tilde{f}_0(z) = f_0(\phi(z))$$

约束条件

$$\tilde{f}_i(z) = f_i(\phi(z)) \leq 0, i = 1, \dots, m$$

如果 z 是新问题的最优解, 那么 $x = \phi(z)$ 是原问题的最优解

■ 与变量代换的区别：通过函数 ϕ 的像构建约束

3 等价性

■ 消除等式约束的保凸性：

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$Ax = b, i = 1, 2, \dots, p$

等价于

最小化(在 z 上)	$f_0(Fz + x_0)$
约束条件	$f_i(Fz + x_0) \leq 0, i = 1, 2, \dots, m$

其中 $F \in \mathbb{R}^{n \times k}$ 和 x_0 满足

$$Ax = b \Leftrightarrow x = Fz + x_0 \text{ 对于一些 } z$$

(我们可以选择 F 是满秩矩阵, 如此, 我们有 $k = n - \mathbf{rank} A$)

3 等价性

5. 引入等式约束（保凸变换）：

最小化

$$f_0(A_0x + b_0)$$

约束条件

$$f_i(A_ix + b_i) \leq 0, \quad i = 1, 2, \dots, m$$

$$h_i(x) = 0, \quad i = 1, \dots, p$$

等同于

最小化

$$f_0(y_0)$$

约束条件

$$f_i(y_i) \leq 0, \quad i = 1, 2, \dots, m$$

$$y_i = A_ix + b_i, \quad i = 1, 2, \dots, m$$

$$h_i(x) = 0, \quad i = 1, \dots, p$$

3 等价性

6. 极小化某些变量（保凸变换）：

最小化(在 x 上)	$f_0(x_1, x_2)$
约束条件	$f_i(x_1) \leq 0, \quad i = 1, 2, \dots, m$

等同于

最小化(在 x_1 上)	$\tilde{f}_0(x_1)$
约束条件	$f_i(x_1) \leq 0, \quad i = 1, 2, \dots, m$

其中 $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

3 等价性

□ 例 9 (在约束下优化二次函数的部分变量):

考虑具有严格凸的二次目标问题,其中某些变量不受约束

$$\text{最小化} \quad x_1^T P_{11} x_1 + 2x_1^T P_{12} x_2 + x_2^T P_{22} x_2$$

$$\text{约束条件} \quad f_i(x_1) \leq 0, i = 1, \dots, m$$

■ 分析: 这里我们可以解析地优化 x_2 :

$$\inf_{x_2} (x_1^T P_{11} x_1 + 2x_1^T P_{12} x_2 + x_2^T P_{22} x_2) = x_1^T (P_{11} - P_{12} P_{22}^{-1} P_{12}^T) x_1$$

因此原问题等价于

$$\text{最小化} \quad x_1^T (P_{11} - P_{12} P_{22}^{-1} P_{12}^T) x_1$$

$$\text{约束条件} \quad f_i(x_1) \leq 0, i = 1, \dots, m$$

3 等价性

7. 上境图问题形式（保凸变换）：

标准形式的凸问题等价于

最小化(在 x, t 上) t

约束条件

$$f_0(x) - t \leq 0$$

$$f_i(x) \leq 0, \quad i = 1, \dots, m$$

$$h_i(x) = 0, \quad i = 1, \dots, p$$

大 纲

1. 标准形式的优化问题

2. 最优性条件

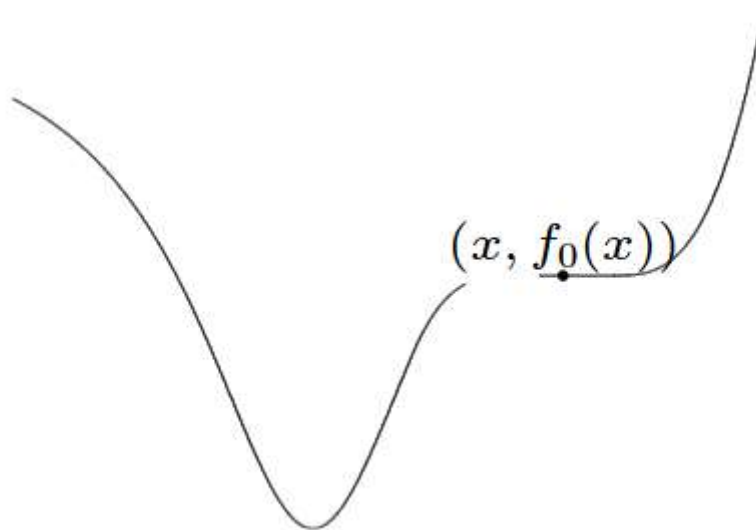
3. 等价性

4. 拟凸优化

4 拟凸优化

最小化	$f_0(x)$
约束条件	$f_i(x) \leq 0, i = 1, 2, \dots, m$
	$Ax = b$

其中 $f_0: \mathbb{R}^n \rightarrow \mathbb{R}$ 是拟凸函数, f_1, \dots, f_m 是凸函数
可能存在局部最优点, 这些点并非 (全局) 最优



4 拟凸优化

■ 函数 f_0 的下水平集的凸表示:

如果 f_0 是拟凸的,则存在一个函数族 ϕ_t ,使得:

对于固定的 t , $\phi_t(x)$ 在 x 上是凸的

f_0 的 t -下水平集是 ϕ_t 的 0-下水平集,即

$$f_0(x) \leq t \Leftrightarrow \phi_t(x) \leq 0$$

■ 例 10:

$$f_0(x) = \frac{p(x)}{q(x)}$$

其中 p 是凸函数, q 是凹函数, 且在 $\text{dom } f$ 上 $p(x) \geq 0$, $q(x) > 0$, f 可以取 $\phi(x) = p(x) - tq(x)$:

1. 当 $t \geq 0$ 时, ϕ_t 关于 x 是凸的
2. $p(x)/q(x) \leq t$ 当且仅当 $\phi_t \leq 0$

4 拟凸优化

■ 通过凸可行性问题进行拟凸优化:

$$\phi_t(x) \leq 0, \quad f_i(x) \leq 0, i = 1, \dots, m, \quad Ax = b$$

1. 对于固定的 t, x 中的凸可行性问题
2. 如果可行,我们可以得出 $t \geq p^*$;如果不可行, $t \leq p^*$

Bisection method for quasiconvex optimization

given $l \leq p^*, u \geq p^*$, tolerance $\epsilon > 0$.

repeat

1. $t := (l + u)/2$.

2. Solve the convex feasibility problem (1).

3. **if** (1) is feasible, $u := t$; **else** $l := t$.

until $u - l \leq \epsilon$.

需要恰好 $\left\lceil \log_2 \left(\frac{u-l}{\epsilon} \right) \right\rceil$ 次迭代(其中 u, l 是初始值)

谢谢！