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

3.4单变量极值问题的解法

在处理一般的非线性规划问题

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

的算法中常常会遇到情形：求一元函数极值问题

$$\min_{\lambda \geq 0} f(x_k + \lambda z_k)$$

这里 $x_k, z_k \in \mathbb{R}^n$ 分别表示第k步得到的点和进一步探寻极值的方向.

有必要讨论一元函数无约束极值问题的算法.

3.4.1 “成功-失败”法

$$\min_{\lambda \in \mathbb{R}^1} f(\lambda)$$

选取初始点λ_0和步长h.

如果 $f(\lambda_0+h)<f(\lambda_0)$, 则搜索“成功”. 当前点变为 λ_0+h , 步长放大为2h. 再比较 $f(\lambda_0+h)$ 和 $f((\lambda_0+h)+2h)$.

如果 $f(\lambda_0+h)\geq f(\lambda_0)$, 则搜索“失败”. 当前点仍然为$\lambda_0$, 搜索方向改变， 步长变为先前步长的$\frac{1}{4}$.再比较 $f(\lambda_0)$ 和 $f(\lambda_0-\frac{h}{4})$.

重复上述过程，直到步长小于一开始设定的允许误差$\epsilon>0$.

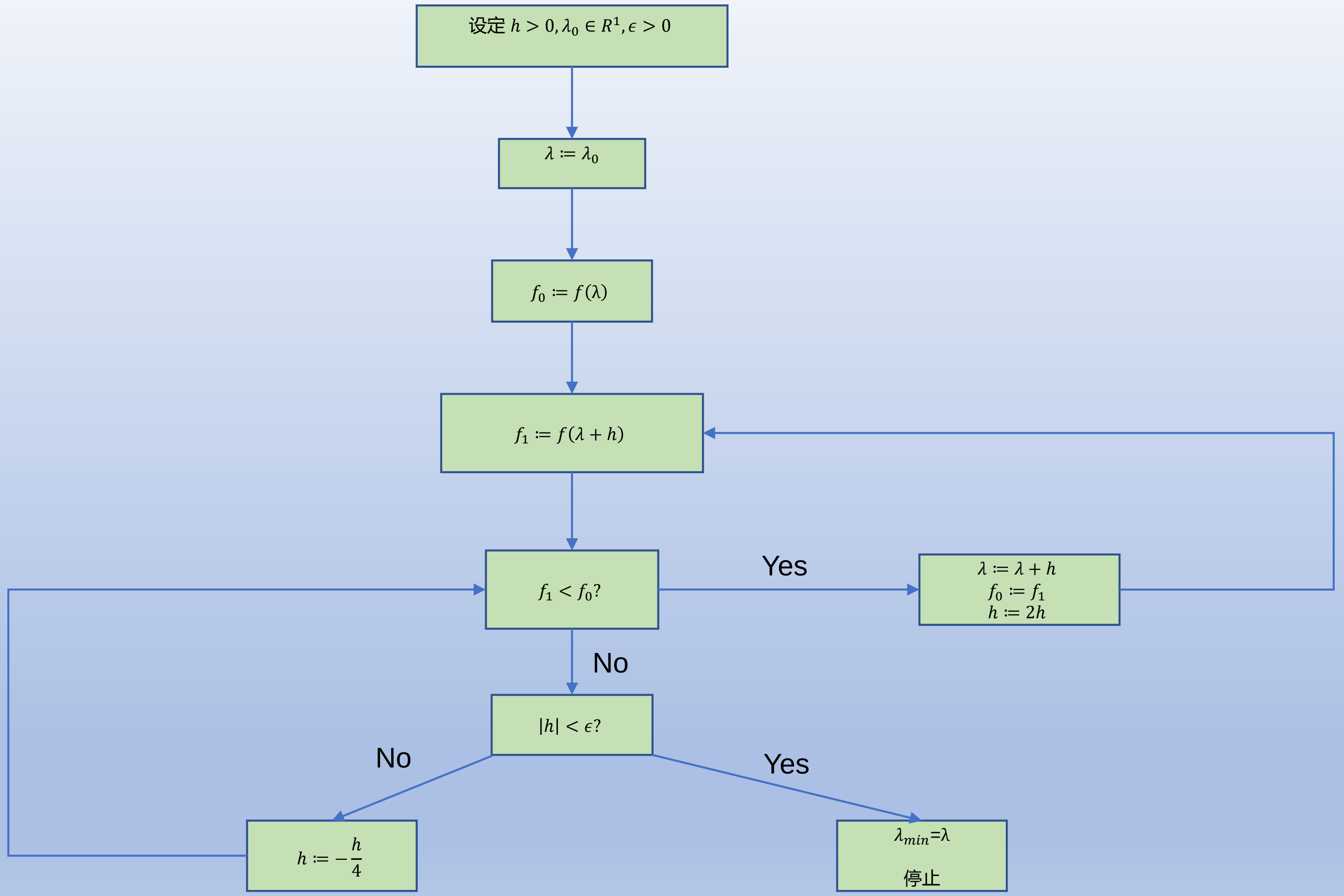

设定 $h > 0, \lambda_0 \in R^1, \epsilon > 0$
$\lambda := \lambda_0$
$f_0 := f(\lambda)$
$f_1 := f(\lambda + h)$
$f_1 < f_0?$
Yes
$\lambda := \lambda + h$
$f_0 := f_1$
$h := 2h$
No
$|h| < \epsilon?$
No
Yes
$h := -\frac{h}{4}$
λ_{min}=λ
停止

图3.4-1

3.4.2. 0.618法(优选法)

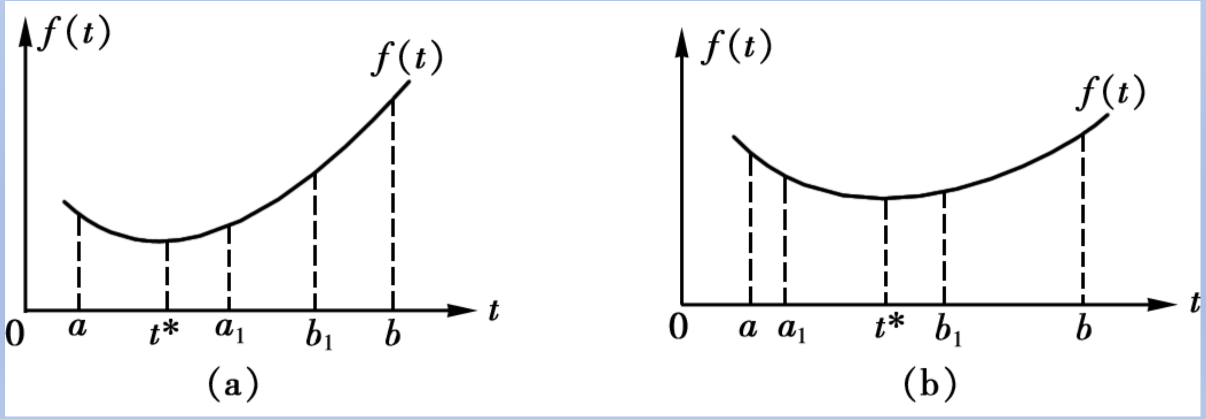
考虑问题 $$\min_{t\in[a,b]} f(t)$$

假设函数f是区间$[a,b]$上的下单峰函数, 它有唯一的极小点t^*. 函数在 $[a,t^*]$严格单调下降， 在$[t^*,b]$严格单调递增.

若在此区间内任取两点a_1和b_1 $a_1<b_1$ 有如下情形:

1. $f(a_1)<f(b_1)$ (图(a)), 这时极点t^*在区间$[a,b_1]$中.

2. $f(a_1)\geq f(b_1)$ (图(b)), 这时极点t^*在区间$[a_1,b]$中.

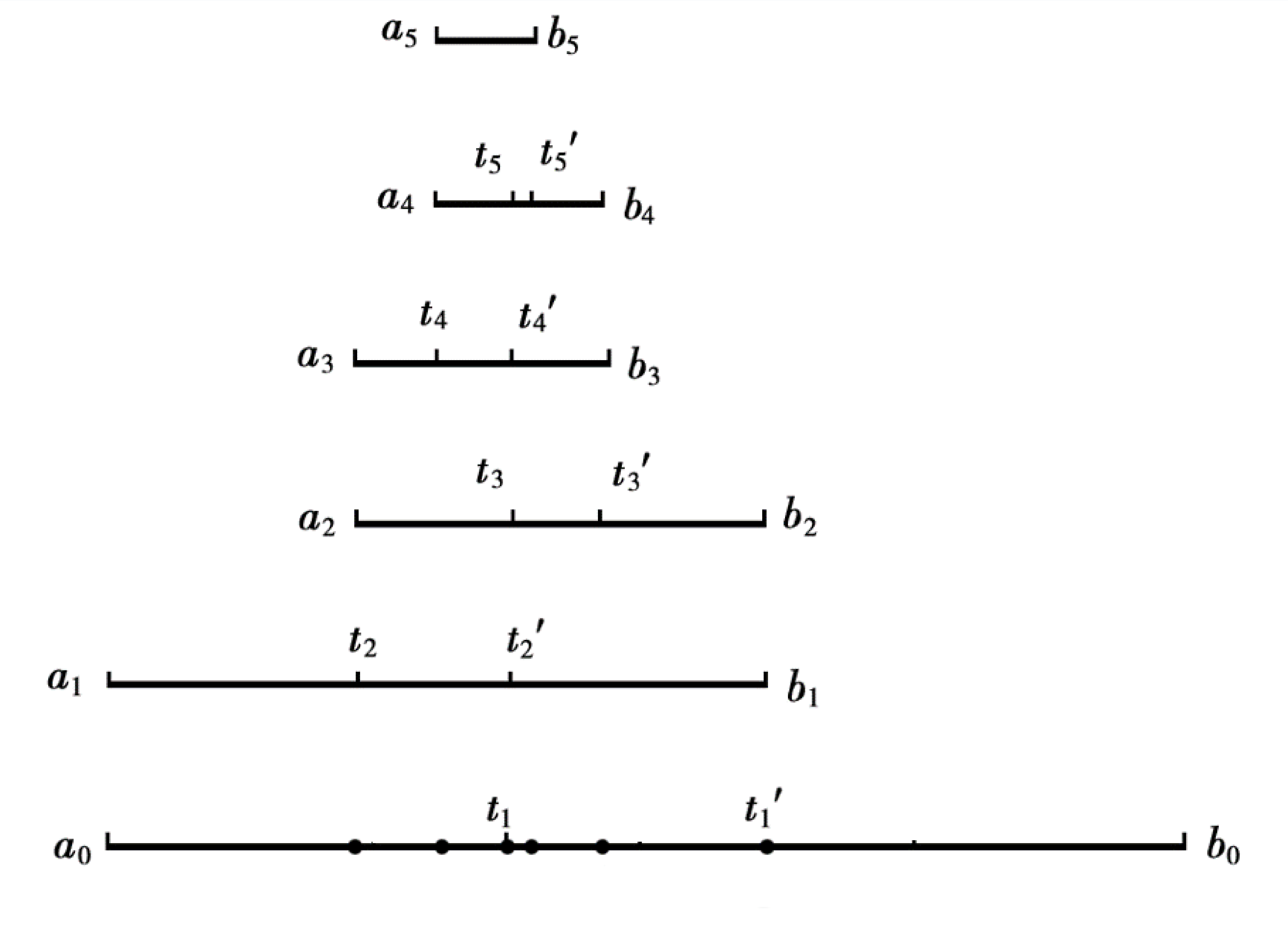


这说明，只要在区间 $[a,b]$ 内取两个不同点，并算出它们的函数值加以比较，就可以把搜索区间 $[a,b]$ 缩小成 $[a,b_1]$ 或 $[a_1,b]$ (缩小后的区间仍需包含极小点).

如果要继续缩小搜索区间 $[a,b_1]$ 或 $[a_1,b]$ ，就只需在上述区间内再取一点算出其函数值，并与 $f(a_1)$ 或 $f(b_1)$ 加以比较即可. 只要缩小后的区间包含极小点 t^*，则区间缩小得越小，就越接近于函数的极小点，但计算函数值的次数也就越多. 自然希望以恰当的方式选择点，使得以较少的计算次数，能够尽快地缩短区间.

“0.618法”是一种高效地缩短包含极值点区间的算法. 每次把点 a_1, b_1 分别取在当前区间[a, b]的0.382和0.618处, 即

$$a_1 = a + 0.382(b-a)$$
$$b_1 = a + 0.618(b-a)$$



注：黄金分割数0.618满足方程

$$x^2 = 1 - x$$

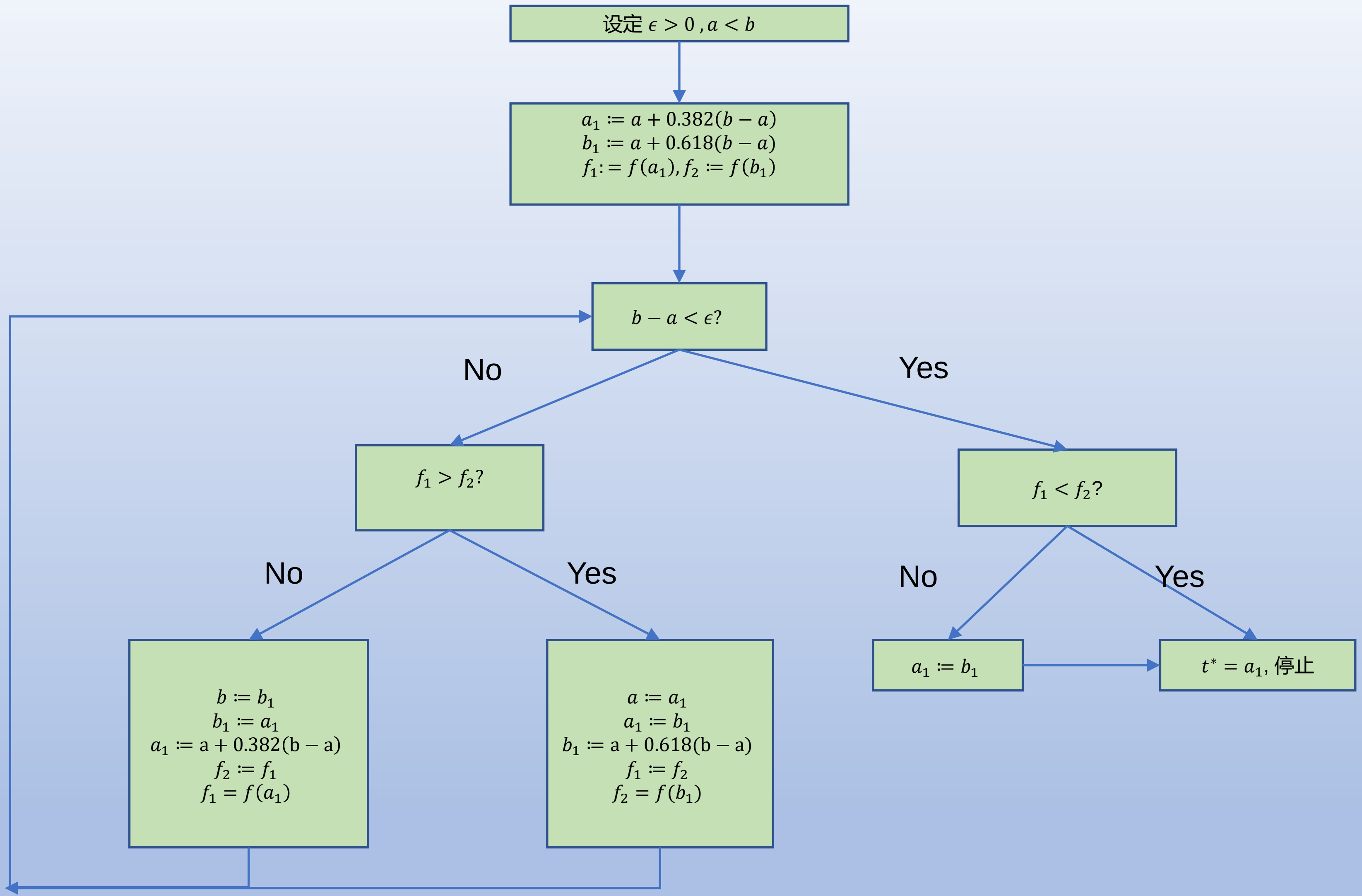

设定 $\epsilon > 0, a < b$
$a_1 := a + 0.382(b - a)$
$b_1 := a + 0.618(b - a)$
$f_1 := f(a_1), f_2 := f(b_1)$
$b - a < \epsilon?$
No
Yes
$f_1 > f_2?$
$f_1 < f_2?$
No
Yes
No
Yes
$b := b_1$
$b_1 := a_1$
$a_1 := a + 0.382(b - a)$
$f_2 := f_1$
$f_1 = f(a_1)$
$a := a_1$
$a_1 := b_1$
$b_1 := a + 0.618(b - a)$
$f_1 := f_2$
$f_2 = f(b_1)$
$a_1 := b_1$
$t^* = a_1$, 停止

图3.4-2

3.4.3 二次插值法

想法：用一个二次函数近似原函数，用二次函数的极点近似原函数的极点.

考虑 $$\min_{\lambda\geq 0} f(\lambda)$$

这里假设 $\frac{df}{d\lambda}(0)<0$. 寻找二次函数 $g(x)=a_0+a_1\lambda+a_2\lambda^2$ 近似f. 那么g的极小点为

$$\bar{\lambda}=-a_1/2a_2$$

为了找到合理的二次函数系数a_0,a_1,a_2,我们需要找到$\lambda_0>0$ 满足:

$$f(0)\geq f(\lambda_0), f(2\lambda_0)\geq f(\lambda_0)$$

那么对三个点0，$\lambda_0, 2\lambda_0$ 进行二次插值:

$$\begin{cases} a_0 & = & f(0) \\ a_0+a_1\lambda_0+a_2{\lambda_0}^2 & = & f(\lambda_0) \\ a_0+2a_1\lambda_0+4a_2{\lambda_0}^2 & = & f(2\lambda_0) \end{cases}$$

求得

$$\lambda_{\min} \sim \bar{\lambda} = \frac{4f(\lambda_0) - 3f(0) - f(2\lambda_0)}{2(2f(\lambda_0) - f(0) - f(2\lambda_0))}\lambda_0$$

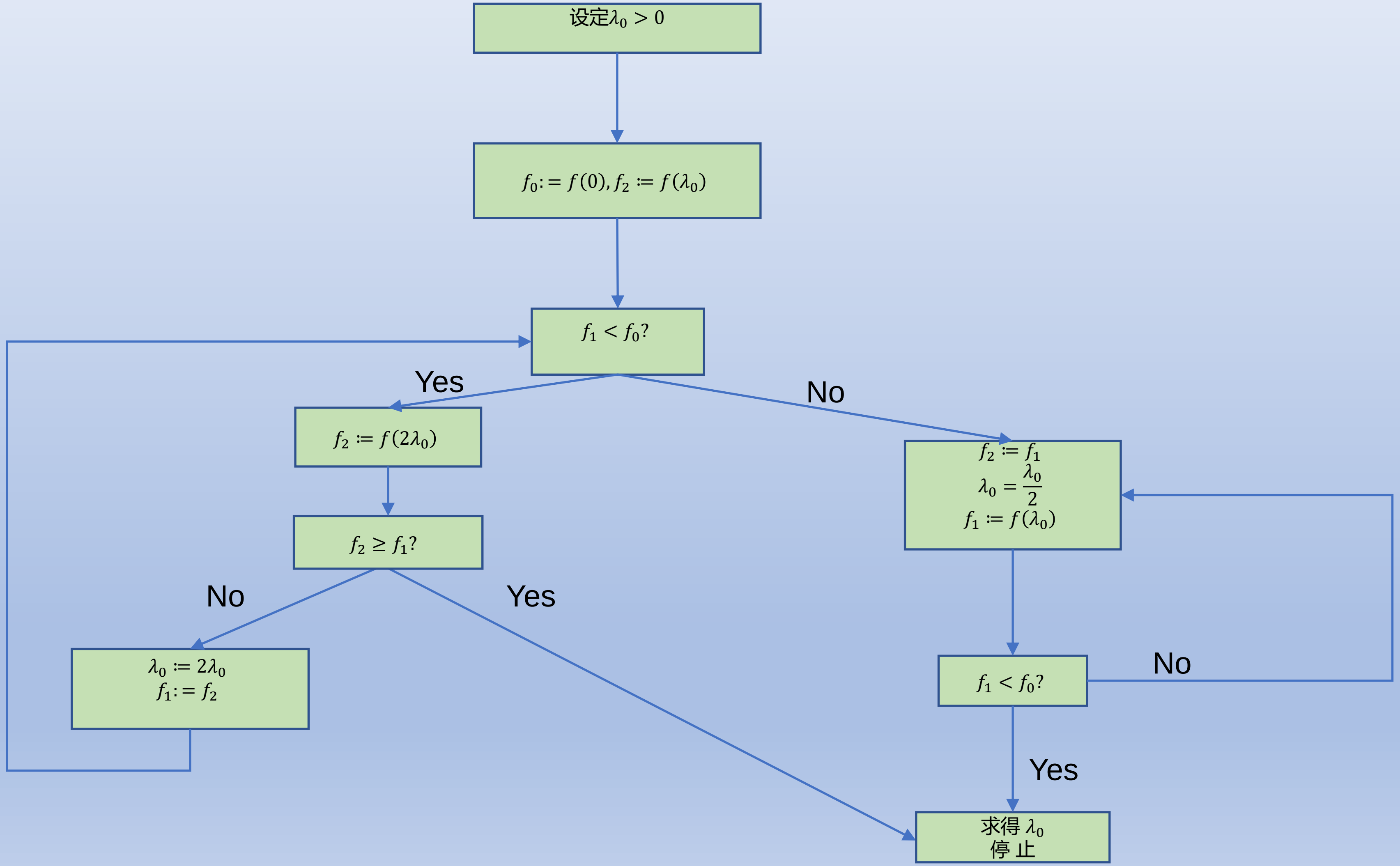


图3.4-3

3.5 无约束极值问题的解法

3.5.1 最速下降法

考虑问题无约束极值问题

$$\min_{x\in\mathbb{R}^n} f(x)$$

假设目标函数具有连续的一阶偏导函数。那么：

- 任何局部极点x_0 一定是临界点，即$\nabla f(x_0)$=0.
- 若$\nabla f(x_0) \neq$0, 那么向量$-\nabla f(x_0)$给出函数值下降最快的方向(局部上)

基于如上考虑，最速下降法就是在一点处求梯度向量，再延起负梯度方向，求对应一元函数的最小值。不断迭代，得到趋于极值点的序列.

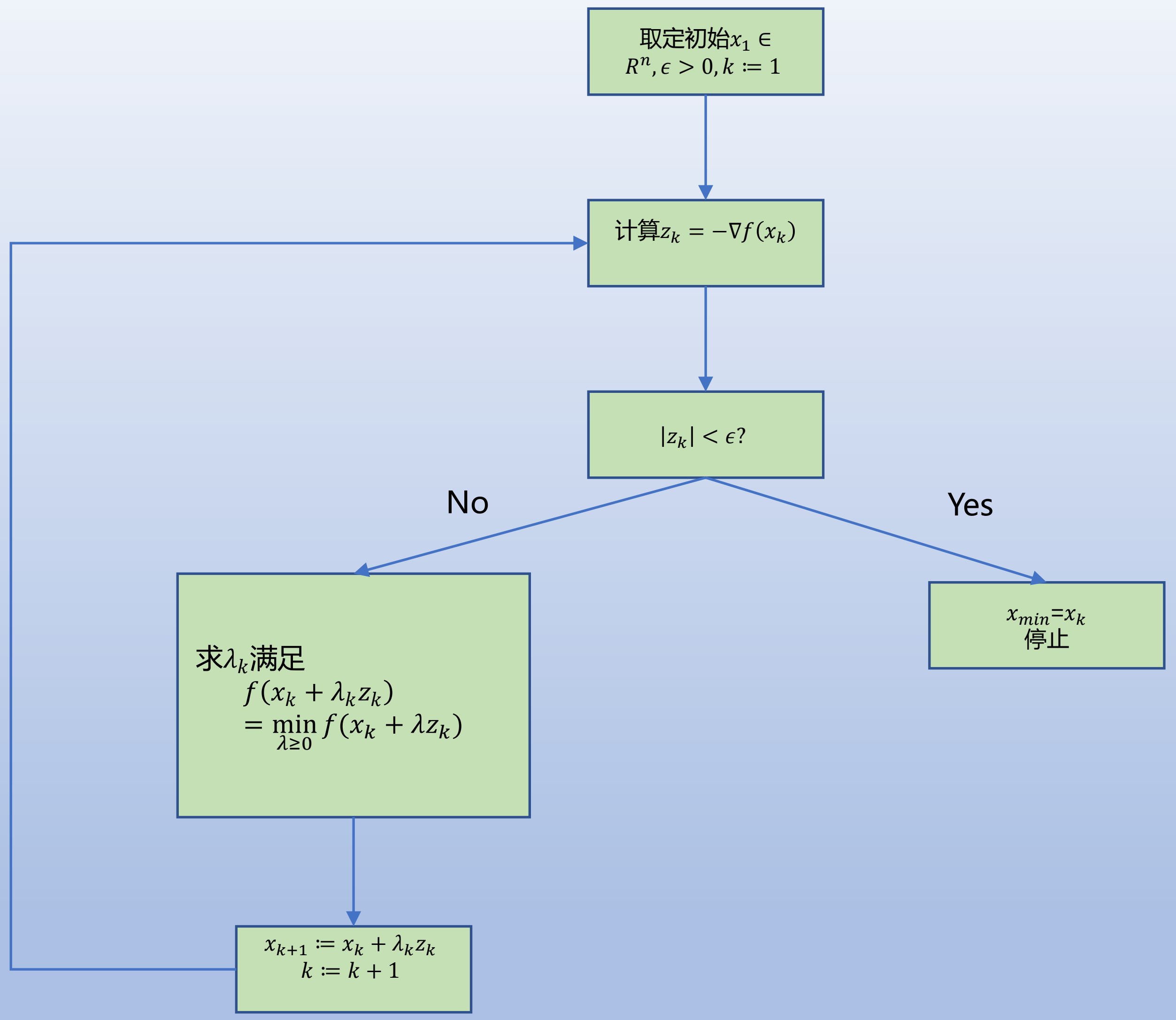

取定初始$x_1 \in R^n, \epsilon > 0, k := 1$
计算$z_k = -\nabla f(x_k)$
$|z_k| < \epsilon$?
No
Yes
求λ_k满足 $f(x_k + \lambda_k z_k) = \min_{\lambda \ge 0} f(x_k + \lambda z_k)$
$x_{min} = x_k$
停止
$x_{k+1} := x_k + \lambda_k z_k$
$k := k + 1$

图3.5-1

3.5.2 广义牛顿法

考虑优化问题

$$\min_{x\in\mathbb{R}^n} f(x)$$

假设f 有连续的二阶偏导函数, 且极小值点x_0处海赛矩阵$H(f)=[\frac{\partial^2 f}{\partial x_i \partial x_j}]$ 为对称正定矩阵.

首先考虑

$$\min_{x\in\mathbb{R}^n} f(x)=\frac{1}{2}x^TCx+p^Tx$$

这里C是对称正定n阶矩阵, $p\in R^n$.

f任意点处的海赛矩阵是C,所以是严格凸函数， 所以其局部最优点也是全局最优点. 令

$$\nabla f=Cx+p=0$$

得最优解

$$x_0=-C^{-1}p$$

对于任意$x \in R^n$,最优解到x的位移向量为

$$\begin{aligned} x_0 - x &= -C^{-1}p - x \\ &= -C^{-1}(p + Cx) \\ &= -C^{-1}\nabla f(x) \\ &= -(Hf(x))^{-1}\nabla f(x) \end{aligned}$$

注意：向量 $-(Hf(x))^{-1}\nabla f(x)$ 给出了从x出发，到最优解x_0的方向.

对于一般的满足假设的函数f，在最优解x_0附近，利用多元函数的Taylor展开，有

$$\begin{aligned} & f(x) \\ = \quad & f(x_0) + (\nabla f(x_0))^T(x - x_0) + \frac{1}{2}(x - x_0)^T Hf(x_0)(x - x_0) + o(|x - x_0|^2) \\ \sim \quad & f(x_0) + (\nabla f(x_0))^T(x - x_0) + \frac{1}{2}(x - x_0)^T Hf(x_0)(x - x_0) \end{aligned}$$

利用情形 $\min_{x\in\mathbb{R}^n} f(x)=\frac{1}{2}x^TCx+p^Tx$ 得到的“指引”向量，结合一维搜索有广义牛顿算法:

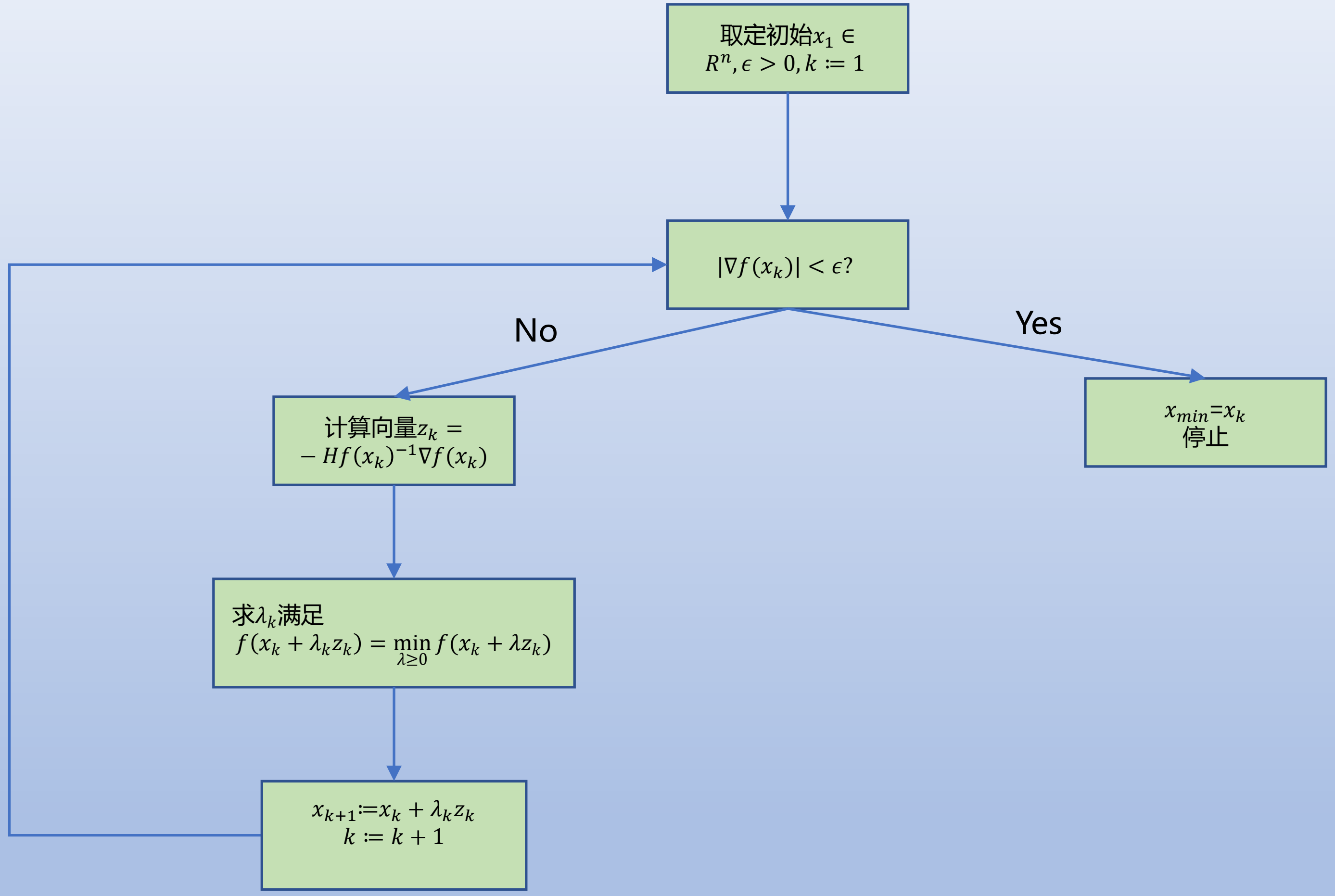


图3.5-2

3.6 罚函数方法

介绍求解非线性规划问题的罚函数法. 使用这种方法，可将非线性规划问题的求解，转化为求解一系列无约束极值问题.

罚函数法的经济解释：

将目标函数$f(x)$看成“价格”，约束条件看成某种“规定”，采购人需要在规定范围内购置最便宜的东西. 管理者对违反规定制定了一种“惩罚”机制. 若符合规定，罚款为零. 否则，要收取罚款. 此时，采购人付出的总代价应是价格和罚款的总和. 采购者的目标是使总代价最小，这是一个无约束问题. 当罚款规定得很苛刻时，违反规定支付的罚款很高，这就迫使采购人符合规定. 数学上的表现为当惩罚因子M足够大时，无约束问题的最优解应满足约束条件，而成为约束问题的最优解.

简单情形

$$\begin{cases} \min f(x) & \\ g(x) & = 0 \end{cases}$$

可行解集 $R = \{x \mid g(x) = 0\}$. 取定非常大的整数$M > 0$. 直观上，如果$x \notin R$,那么x不太可能是无约束极值问题

$$\min_{x\in\mathbb{R}^n} T(x, M) = f(x) + Mg(x)^2$$

的最优解. 如果是最优解，也一定离可行集R很近.

注意到

$$T(x, M)\begin{cases} = f(x) & x \in R \\ > f(x) & x \notin R \end{cases}$$

如果无约束极值问题的最优解$x_0 \in R$,那么x_0也是原问题的最优解.

处理一般不等式约束，只需注意到

$$g(x) \geq 0 \Leftrightarrow \min(0, g(x)) = 0$$

罚函数法迭代步骤:

(1) 取 $M_1 > 0$ 以及允许误差 $\epsilon > 0$，并令 $k := 1$

(2) 求无约束极值问题的最优解:

$$\min_{x \in \mathbb{R}^n} T(x, M_k) = T(x^k, M_k)$$

其中

$$T(x, M_k) = f(x) + M_k \sum_{j=1}^{m} [\min(0, g_j(x))^2]$$

(3) 若对某一个 $j (1 \leq j \leq m)$ 有 $g_j(x^k) \leq -\epsilon$，则取 $M_{k+1} > M_k$

(例如 $M_{k+1} = cM_k, c = 10$)，令 $k := k+1$

并转向第2步. 否则，停止迭代，得近似解

$$x_{\min} \sim x^k$$

例1　求解非线性规划

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

解：构造罚函数

$$\begin{aligned} T(x, M) &= f(x) + M\{[\min(0, g_1(x))]^2 + [\min(0, g_2(x))]^2\} \\ &= x_1 + x_2 + M\{[\min(0, -x_1^2 + x_2)]^2 + [\min(0, x_1)]^2\} \end{aligned}$$

求 $\min\limits_{x \in \mathbb{R}^2} T(x, M)$

$$\begin{aligned} \frac{\partial T}{\partial x_1} &= 1 + 2M\{[\min(0, g_1(x))(-2x_1)] + \min(0, g_2(x))\} \\ \frac{\partial T}{\partial x_2} &= 1 + 2M[\min(0, -x_1^2 + x_2)] \end{aligned}$$

令 $\dfrac{\partial T}{\partial x_1} = \dfrac{\partial T}{\partial x_2} = 0$，得

$$X(M) = -\left(\frac{1}{2(1+M)}, \left(\frac{1}{4(1+M)^2} - \frac{1}{2M}\right)\right)^{\mathrm{T}}$$

取 $M=1,2,3,4$ ，可得出以下结果:

$M=1,\quad X=(-1/4,-7/16)^{\mathrm{T}}$

$M=2,\quad X=(-1/6,-2/9)^{\mathrm{T}}$

$M=3,\quad X=(-1/8,-29/192)^{\mathrm{T}}$

$M=4,\quad X=(-1/10,-23/200)^{\mathrm{T}}$

可知 $X(M)$ 从R的外面逐步逼近R的边界，当 $M\to\infty$ 时，$X(M)$ 趋于原问题的极小解 $X_{\min}=(0,0)^{\mathrm{T}}$

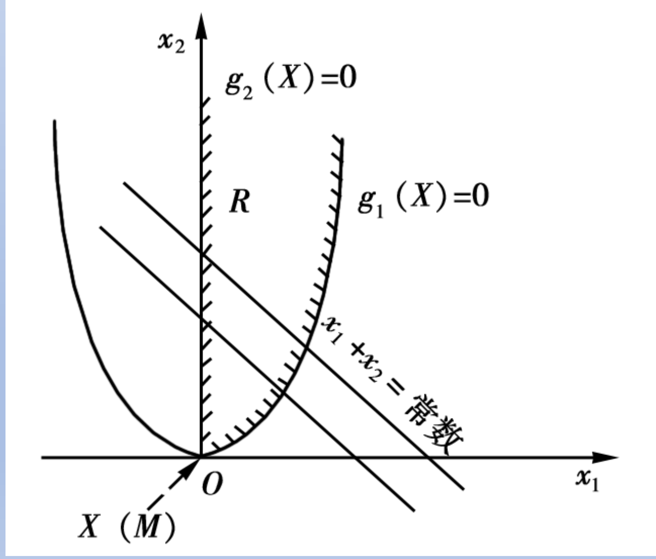


图3.6-1