



西安交通大学
XI'AN JIAOTONG UNIVERSITY

Systems Engineering Institute
Ministry of Education Key Lab for Intelligent Networks and Network Security

一维搜索

One-dimensional search

电信学部·自动化科学与工程学院
系统工程研究所
吴江

Outline

- ▶ 一维搜索简介
- ▶ 精确一维搜索（不利用导数）
- ▶ 精确一维搜索（利用导数）
- ▶ 非精确一维搜索

一维搜索

非线性规划的迭代下降算法框架：

$$x^{(k+1)} = x^{(k)} + t_k p^{(k)}, \quad f(x^{(k+1)}) < f(x^{(k)})$$

怎样确定步长因子 t_k ?

可以看作是寻找一元函数 $\varphi(t) = f(x^{(k)} + t p^{(k)})$ 的最小值点。

One-dimensional search

or

Line search

$$(LS) \begin{cases} \min_t \varphi(t) = f(x^{(k)} + t p^{(k)}) \\ s.t. \quad t \geq 0; \end{cases}$$

$$(LS) \begin{cases} \min_t \varphi(t) = f(x^{(k)} + t p^{(k)}) \\ s.t. \quad 0 \leq t \leq t_{\max} \end{cases}$$

一维搜索问题的一般形式

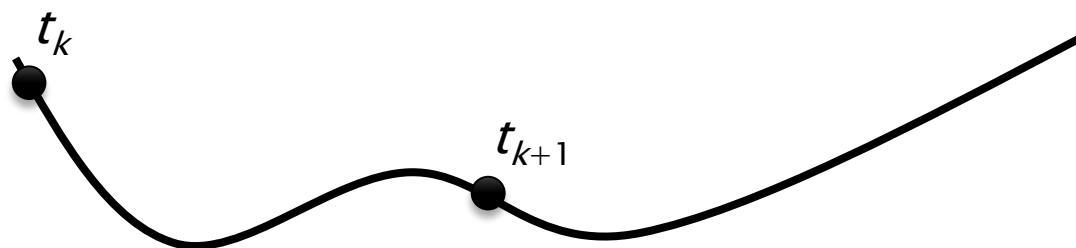
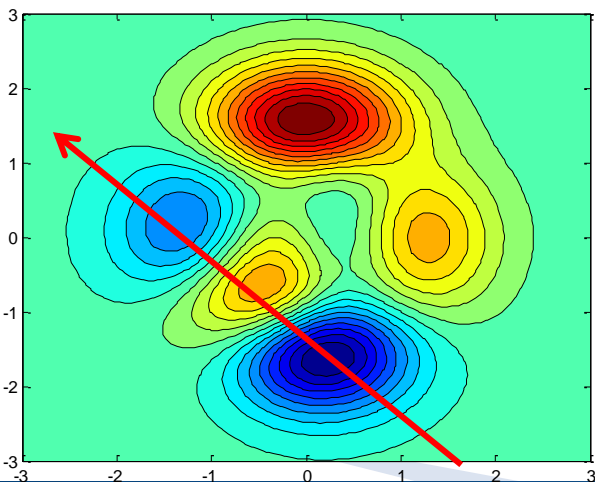
步长

搜索方向

$$\min \quad \varphi(t) = f(x^k + tp^k)$$

$$s.t \quad t \geq 0, \text{ or } 0 \leq t \leq t_{\max}$$

有效一维
搜索问题



几点说明

- ▶ 使用范围：一元函数极小点、NLP的搜索步长
- ▶ 一维搜索问题 = 一维NLP问题 \rightarrow 迭代: $t_1, t_2, \dots, t_l, \dots$
- ▶ 局部最优解, 全局最优解, 满意解
- ▶ 区间 $[a, b]$ 上的单谷函数 $\varphi(t)$:
 - $\exists t^* \in [a, b]$, 使得 $\varphi(t)$ 在 $[a, t^*]$ 上严格递减, 且在在 $[t^*, b]$ 上严格递增.
 - 单谷函数与凸函数

精确与非精确一维搜索

- ▶ 定义: 若一种一维搜索算法产生的试探点到 $\{t_j\}_{j=1}^{\infty}$ 是收敛的($t_j \rightarrow t^*$), 而 t^* 是该一维搜索问题的某个局部最优解, 则称这样的一维搜索算法是精确一维搜索.
- ▶ 试探点列收敛于该一元函数的一个局部极小点
- ▶ 精确一维搜索, 最优一维搜索
- ▶ 非精确一维搜索, 可接受一维搜索
- ▶ 有限步终止
- ▶ 划界问题

常用一维搜索算法

	用导数($\varphi'(t)$)	不用导数
精确搜索	插值法	Fibonacci搜索
	二分法	黄金分割法
	Newton法	
非精确搜索	Goldstein法 _[65']	Goldstein法
	Armijo法 _[66']	Armijo法
	Wolfe法 _[69']	

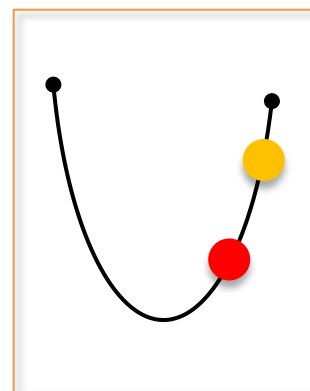
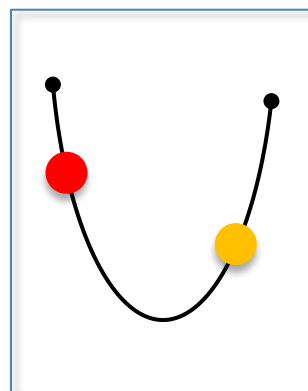
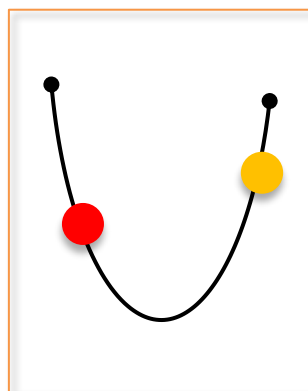
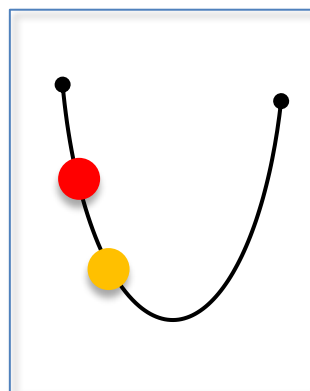
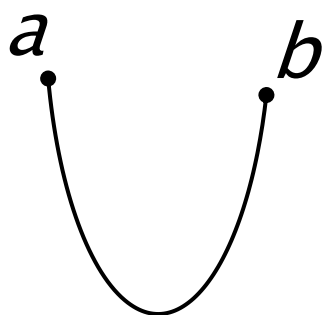
在绝大部分一维搜索问题中, 非精确搜索算法的效果明显好于精确搜索算法

精确一维搜索: 0.618与Fibonacci

- 问题: 已知 $\varphi(t)$ 在 $[a, b]$ 上为单谷函数, 如何通过不断增加试探点 $t_i (i=1, 2, \dots)$, 尽快得到其全局极小的一个足够好的近似?

$$a < t_1 < t_2 < b$$

→ 缩小区间



算法基本思想

► 缩小区间

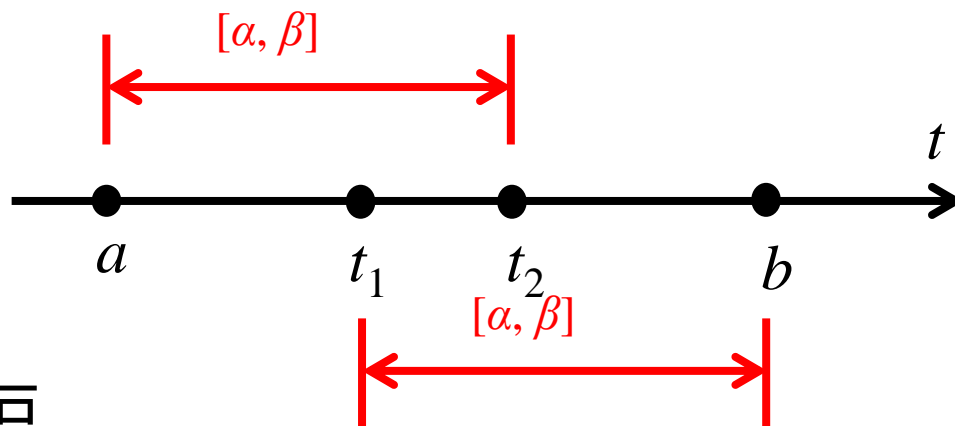
- Step1: 取 $t_1, t_2 \in [a, b]$, 且 $a < t_1 < t_2 < b$. 置 $l=2$,
 - 当 $\varphi(t_1) \leq \varphi(t_2)$ 时, 令 $\alpha=a, \beta=t_2, \gamma=t_1$
 - 当 $\varphi(t_2) \leq \varphi(t_1)$ 时, 令 $\alpha=t_1, \beta=b, \gamma=t_2$
- Step2: 若 $\beta-\alpha \leq \varepsilon$, 停, 输出 γ . 否则更新 t_{l+1} :

	$\varphi(t_{l+1}) \leq \varphi(\gamma)$	$\varphi(t_{l+1}) \geq \varphi(\gamma)$
$t_{l+1} < \gamma$	α 不变, $\beta=\gamma, \gamma=t_{l+1}$	$\alpha=t_{l+1}, \beta, \gamma$ 不变
$t_{l+1} > \gamma$	$\alpha=\gamma, \gamma=t_{l+1}, \beta$ 不变	α, γ 不变, $\beta=t_{l+1}$

t_{i+1} 的选取

- 方式一：保持区间缩小比

0.618



- 方式二：最终区间最短

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584.....

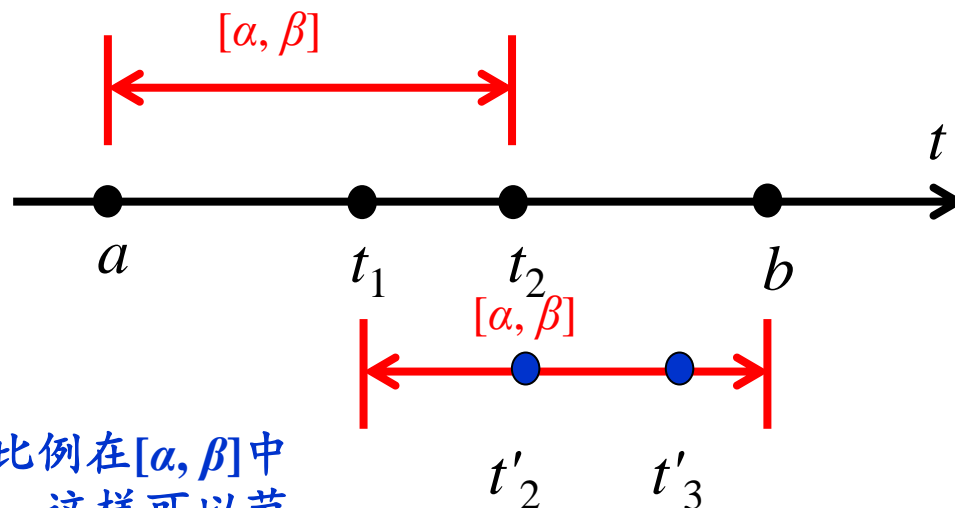
0.618法(黄金分割法)-保持区间缩小比

对称性：出于合理的考虑，
要求 $t_1 - a = b - t_2$

缩小比：第一步后区间缩小比

$$\omega = (t_2 - a) / (b - a) \\ = (b - t_1) / (b - a)$$

以 $[\alpha, \beta] = [t_1, b]$ 为例，若按同样比例在 $[\alpha, \beta]$ 中取两个点，则要求有一个点与 t_2 重合，这样可以节省一个试探点。



$$\omega \approx 0.618$$



$$t_2' = t_2 \Rightarrow \omega = \frac{b - t_2'}{b - t_1} = \frac{b - t_2}{b - t_1} = \frac{(1 - \omega)(b - a)}{\omega(b - a)} \Rightarrow \omega^2 = 1 - \omega \Rightarrow \omega = \frac{-1 \pm \sqrt{5}}{2}$$

$$t_3' = t_2 \Rightarrow 1 - \omega = \frac{b - t_3'}{b - t_1} = \frac{b - t_2}{b - t_1} = \frac{(1 - \omega)(b - a)}{\omega(b - a)} \Rightarrow \omega^2 = 1 \Rightarrow \omega = \pm 1$$



0.618法步骤

第1步 确定单谷区间 $[a,b]$ ，给定最后区间精度 $\varepsilon > 0$ ；

第2步 计算最初两个探索点

$$t_1 = a + 0.382(b - a) = b - 0.618(b - a)$$

$$t_2 = a + 0.618(b - a)$$

并计算 $\varphi_1 = \varphi(t_1)$ ， $\varphi_2 = \varphi(t_2)$ ；

第3步 若 $\varphi_1 \leq \varphi_2$ ，转第4步。否则转第5步；

第4步 若 $t_2 - a \leq \varepsilon$ ，停止迭代，输出 a 。否则令 $a := t_2$ ， $t_2 := t_1$ ， $t_1 := b - 0.618(b - a)$ ， $\varphi_2 := \varphi_1$ ，计算 $\varphi_1 = \varphi(t_1)$ ，转第3步；

第5步 若 $b - t_1 \leq \varepsilon$ ，停止迭代，输出 b 。否则令 $a := t_1$ ， $t_1 := t_2$ ， $t_2 := a + 0.618(b - a)$ ， $\varphi_1 := \varphi_2$ ，计算 $\varphi_2 = \varphi(t_2)$ ，转第3步。

Fibonacci搜索(Fibonacci法)-最终区间最短

- 问题进一步明确化: 已知 $\varphi(t)$ 在 $[a, b]$ 上为单谷函数, 按缩小区间法框架, 若只允许试探 N 个点, 怎样选取试探点 $t_l (l=1, 2, \dots)$, 使最终包含全局极小 t^* 的待探索区间长度最小?

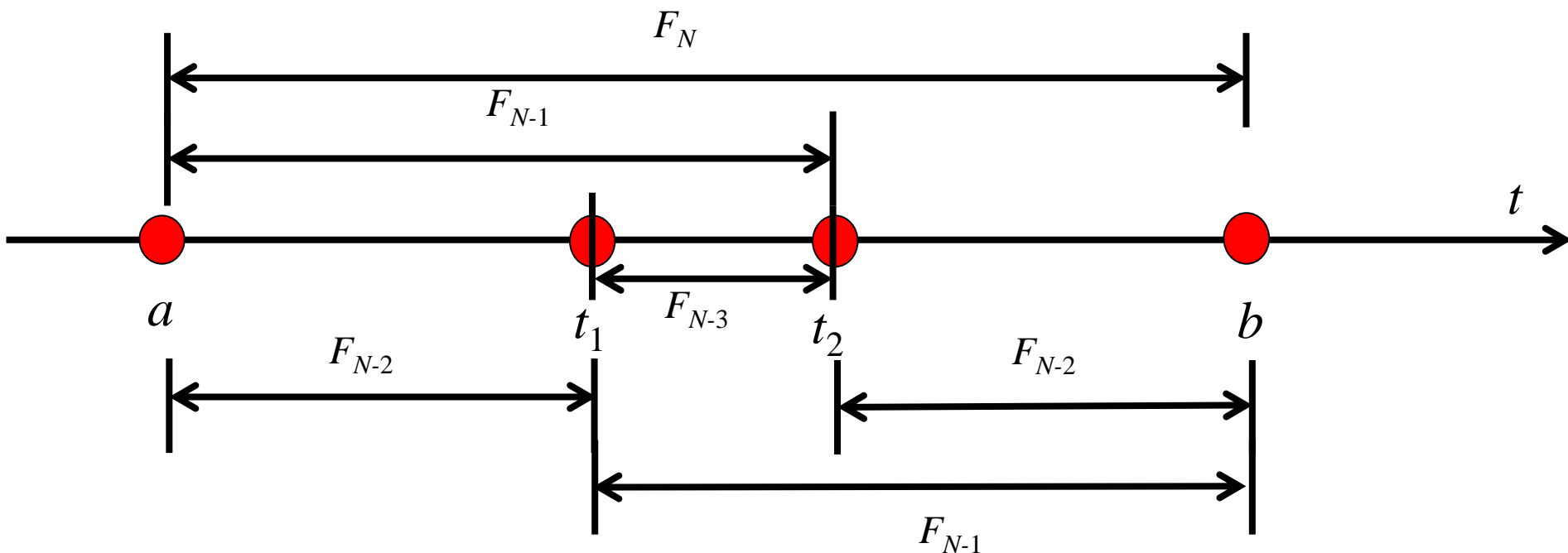
结论: 可描述为非线性规划问题, 可获得全局最优解!

- 去除冗余信息, 问题的等价转化: 若只允许试探 N 个点, 且最终区间长度要求为 1, 最初区间的最大长度是多少?

Fibonacci数列

$$F_0 = 1, F_1 = 1, F_N = F_{N-1} + F_{N-2}$$

Fibonacci搜索(Fibonacci法)



计算过程：按比例缩放至 F_N ,

$$t_1 = a + \frac{F_{N-2}}{F_N} (b - a), \quad t_2 = b - \frac{F_{N-2}}{F_N} (b - a)$$

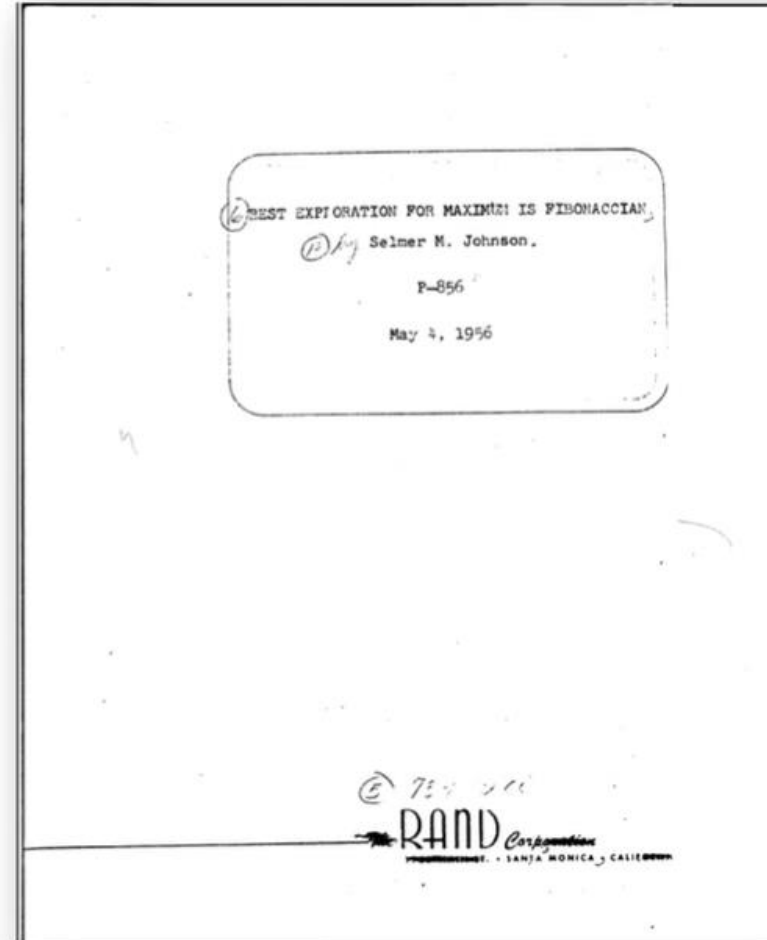
$$\lim_{N \rightarrow +\infty} \frac{F_{N-1}}{F_N} = \frac{\sqrt{5} - 1}{2}$$

0.618法是Fibonacci法的极限形式

Best Exploration for Maximum is Fibonacci by S. M. Johnson

“The F_n form the sequence 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ..., with $F_{20} > 10000$. Thus a maximum can always be located within 10^{-4} of the original interval length with 20 calculations. Using the more conventional technique of computing $f(x)$ and $f(x + \epsilon)$, where x is at the midpoint of the interval [aka Dichotomous Search], may take as many as 28 computations.”

Johnson, Selmer M. *Best exploration for maximum is Fibonacci*. No. P856. RAND CORP SANTA MONICA CA, 1956.



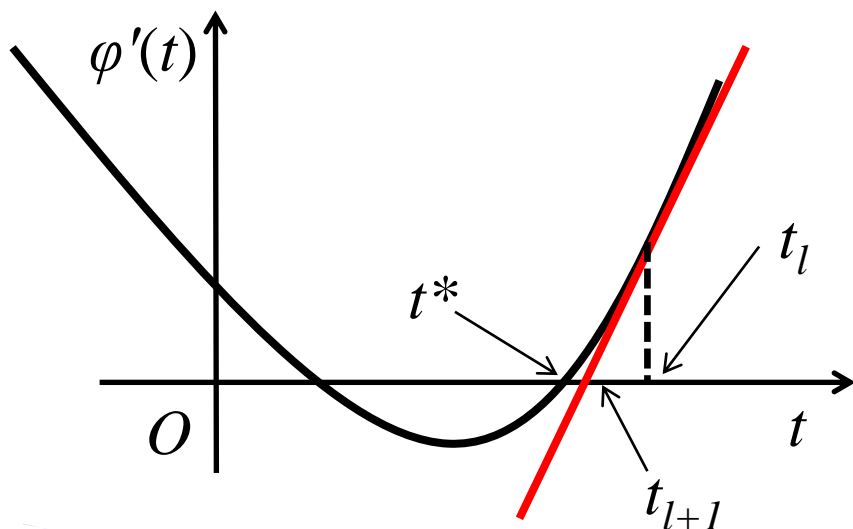
精确一维搜索: Newton法

$$\min \quad \varphi(t)$$

$$s.t \quad t \geq 0$$

寻找 $\varphi'(t^*)=0$ 的点 t^*

几何解释: 切线法



$$y - \varphi'(t_l) = \varphi''(t_l)(t - t_l)$$

$$y = 0 \Rightarrow t = t_l - \frac{\varphi'(t_l)}{\varphi''(t_l)}$$

$$t_{l+1} = t_l - \frac{\varphi'(t_l)}{\varphi''(t_l)}$$

优点: 若收敛, 则收敛很快

缺点: 要求 t_0 非常接近 t^*

要用到高阶导数

Newton法举例

例1. 给定 $t_0 = 2$, 求解:

$$\begin{cases} \min \varphi(t) = \frac{1}{3}t^3 - 2t + 1 \\ s.t. \quad t \geq 0 \end{cases}$$

$$t_{l+1} = t_l - \frac{\varphi'(t_l)}{\varphi''(t_l)}$$

解:
$$t_{l+1} = t_l - \frac{\varphi'(t_l)}{\varphi''(t_l)} = t_l - \frac{t_l^2 - 2}{2t_l} = \frac{t_l}{2} + \frac{1}{t_l}$$

$$t_1 = \frac{3}{2}$$

$$t_2 = \frac{17}{12}$$

$$t_3 = \frac{577}{408} \approx 1.4142157$$

实质: 解方程 $t^2 - 2 = 0$ 或求 $\sqrt{2}$

非精确一维搜索

$$(LS) \begin{cases} \min_t \varphi(t) = f(x^{(k)} + t p^{(k)}) \\ s.t. \quad t \geq 0; \end{cases}$$

基本假设:

$\varphi(t)$ 连续可导

$\varphi'(0) < 0$

非精确一维搜索方法并不试图寻找 (LS) 问题的全局或局部最优解，而是致力于获得一个满意解，因此，它的方法包含**两个结构要素**：

1. 满意步长的判别准则
2. 获得满意步长的方法

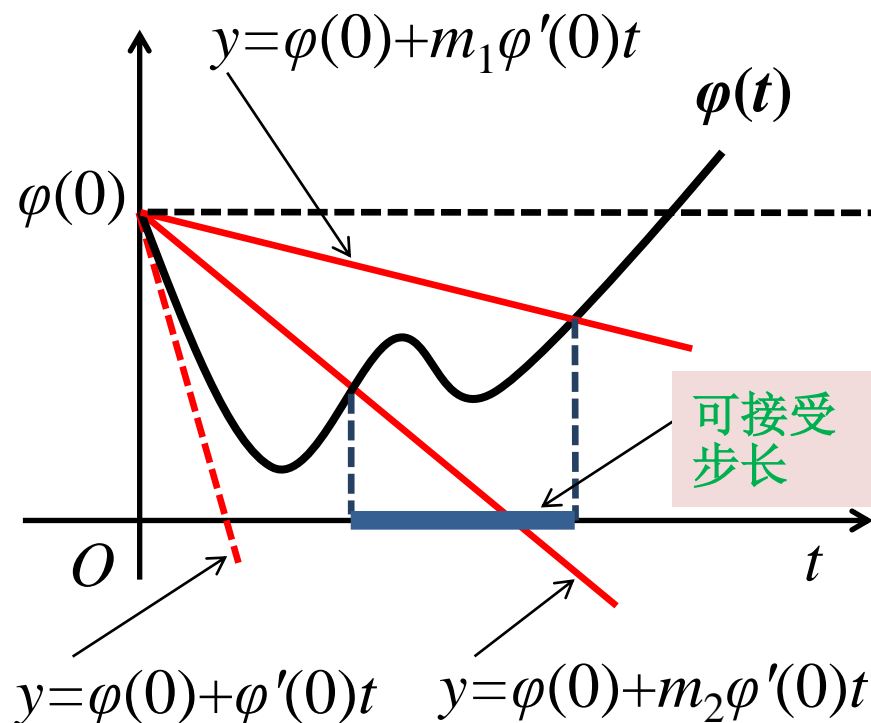
Goldstein准则

选 m_1, m_2 满足 $0 < m_1 < m_2 < 1$, 则步长因子 t_k 可接受当且仅当:

$$\begin{cases} \varphi(t_k) \leq \varphi(0) + m_1 t_k \varphi'(0) \\ \varphi(t_k) \geq \varphi(0) + m_2 t_k \varphi'(0) \end{cases}$$

步长规则的本质含义:
有足够下降, 不过分靠近 0

定理: 若 $\varphi(t)$ 在 $[0, +\infty)$ 有下界, 则满足Goldstein准则的步长存在。



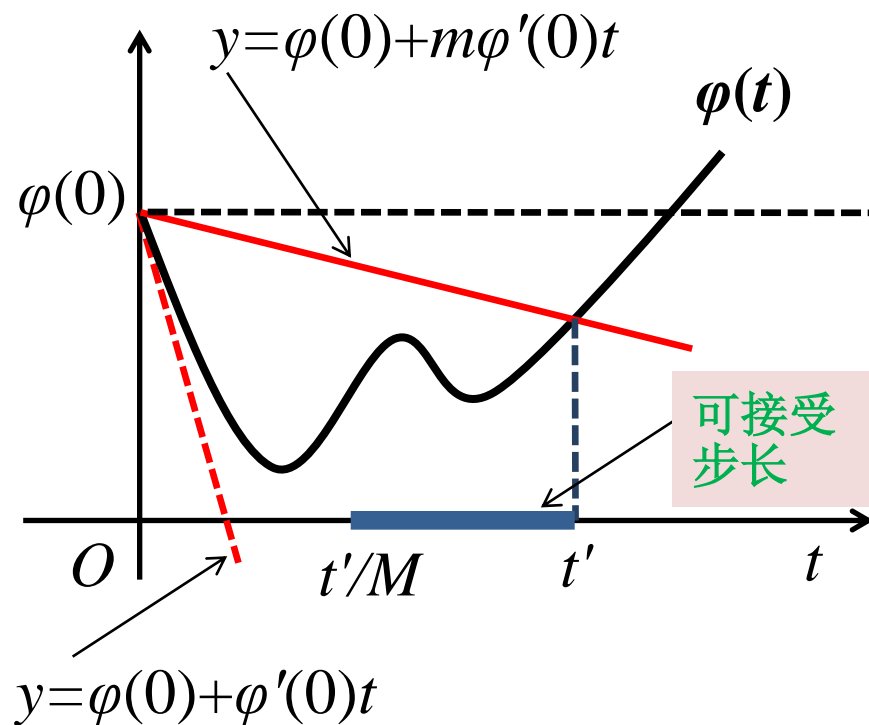
Armijo准则

选 m , M 满足 $0 < m < 1 < M$, 则步长因子 t_k 可接受当且仅当:

$$\begin{cases} \varphi(t_k) \leq \varphi(0) + mt_k \varphi'(0) \\ \varphi(Mt_k) \geq \varphi(0) + mMt_k \varphi'(0) \end{cases}$$

步长规则的本质含义:
有足够下降, 不过分靠近 0

定理: 若 $\varphi(t)$ 在 $[0, +\infty)$ 有下界, 则满足Armijo准则的步长存在。



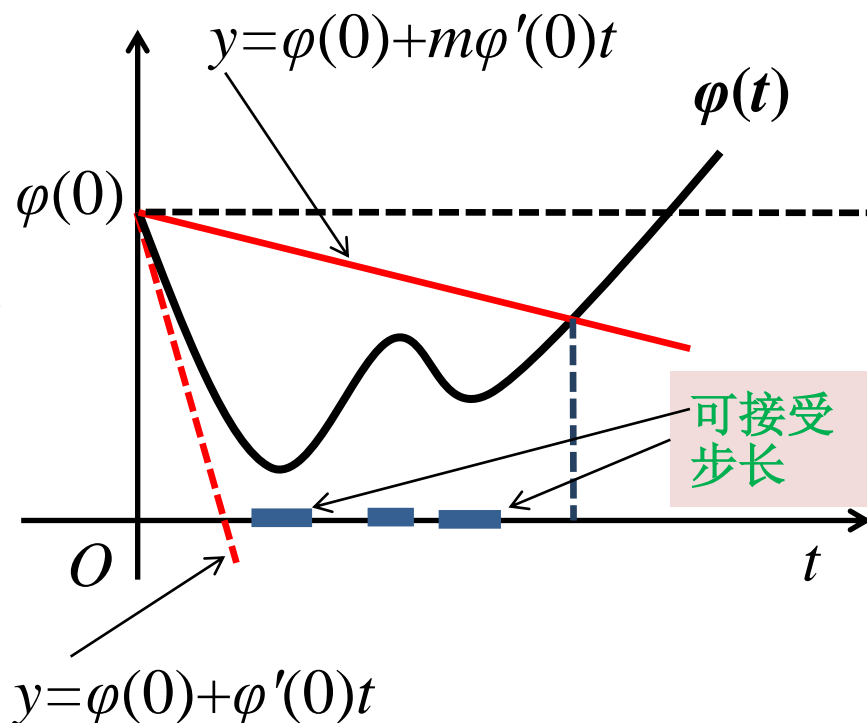
Wolfe准则

选 $m \in (0, 1/2)$, $\sigma \in (m, 1)$, 则步长因子 t_k 可接受当且仅当:

$$\begin{cases} \varphi(t_k) \leq \varphi(0) + mt_k \varphi'(0) \\ |\varphi'(t_k)| \leq -\sigma \varphi'(0) \end{cases}$$

步长规则的本质含义: 有足够下降, 不过分靠近 0, 尽量包含最优解

定理: 若 $\varphi(t)$ 在 $[0, +\infty)$ 有下界, 则满足 Wolfe 准则的步长存在。



非精确一维搜索的一般原则

- ▶ $\varphi(t_k)$ 和 $\varphi(t_0)$ 相比要有足够的下降
- ▶ 距离不要 t_0 太近
- ▶ 尽可能找到局部极小点

作业

▶ P154 10