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

电信学院·自动化科学与技术系
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**

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

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

一维搜索问题的一般形式

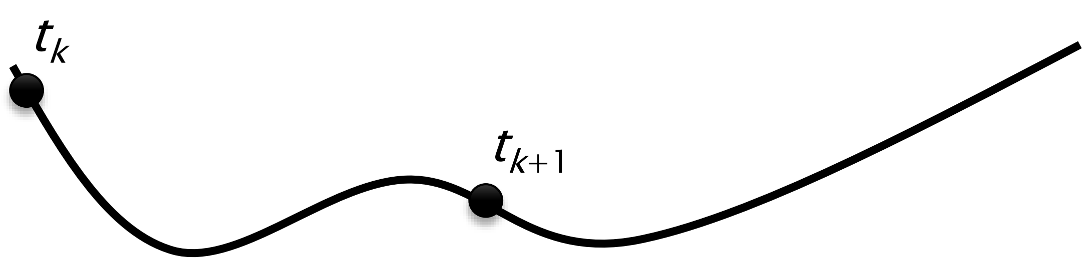
步长

搜索方向

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

$$s.t \quad t \ge 0, or \quad 0 \le t \le t_{\max}$$

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－＞迭代: t_1, $t_2, \ldots, t_l, \ldots$
- 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$\{t_l\}_{l=1}^{\infty}$是收敛的($t_l -> t^*$), 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

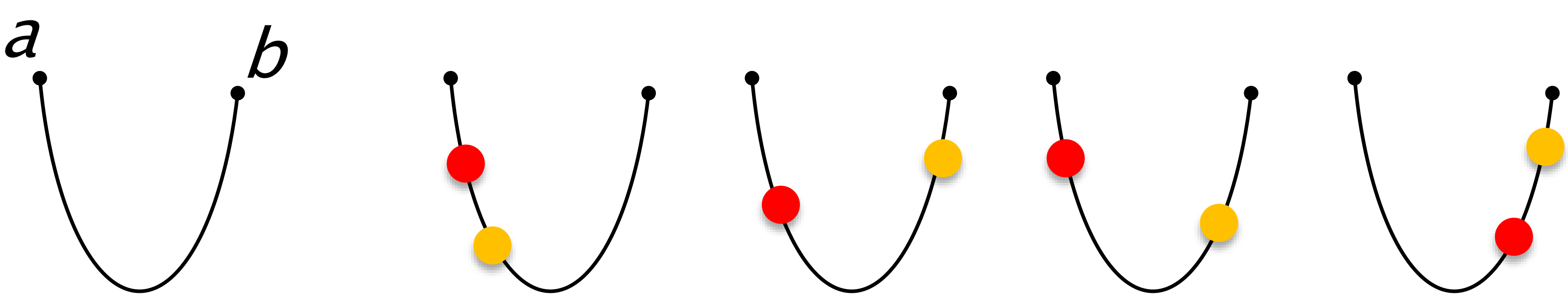
	用导数($\varphi'(t)$)	不用导数
精确搜索	插值法	**Fibonacci搜索**
	二分法	**黄金分割法**
	Newton法	
非精确搜索	**Goldstein法**	Goldstein法
	Armijo法	Armijo法
	Wolfe法	

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,…), 尽快得到其全局极小的一个足够好的近似?

$$a<t_1<t_2<b \longrightarrow \text{缩小区间}$$



算法基本思想

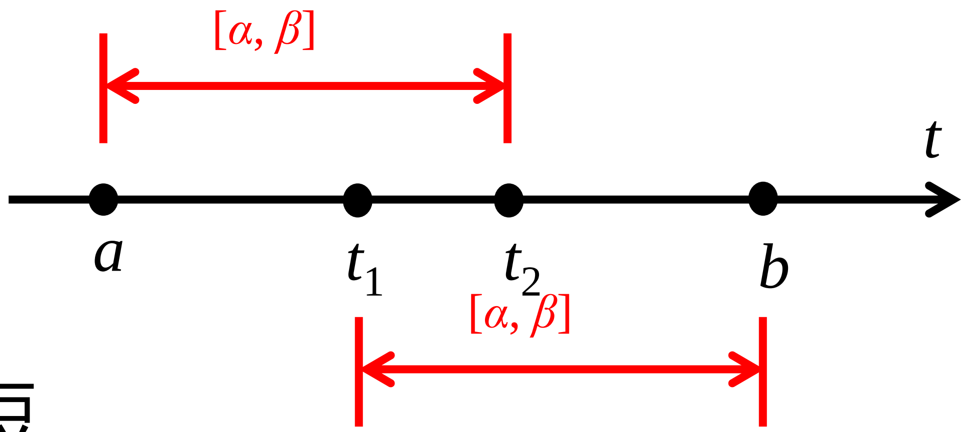
- **缩小区间**
  - Step1: 取t_1, $t_2 \in [a, b]$, 且$a < t_1 < t_2 < b$. 置$l=2$,
    - 当$\varphi(t_1) \le \varphi(t_2)$时, 令$\alpha=a$, $\beta=t_2$, $\gamma=t_1$
    - 当$\varphi(t_2) \le \varphi(t_1)$时, 令$\alpha=t_1$, $\beta=b$, $\gamma=t_2$
  - Step2: 若β-$\alpha \le \varepsilon$, 停, 输出γ. 否则更新t_{l+1}:

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

t_{l+1}的选取

- 方式一：保持区间缩小比

0.618

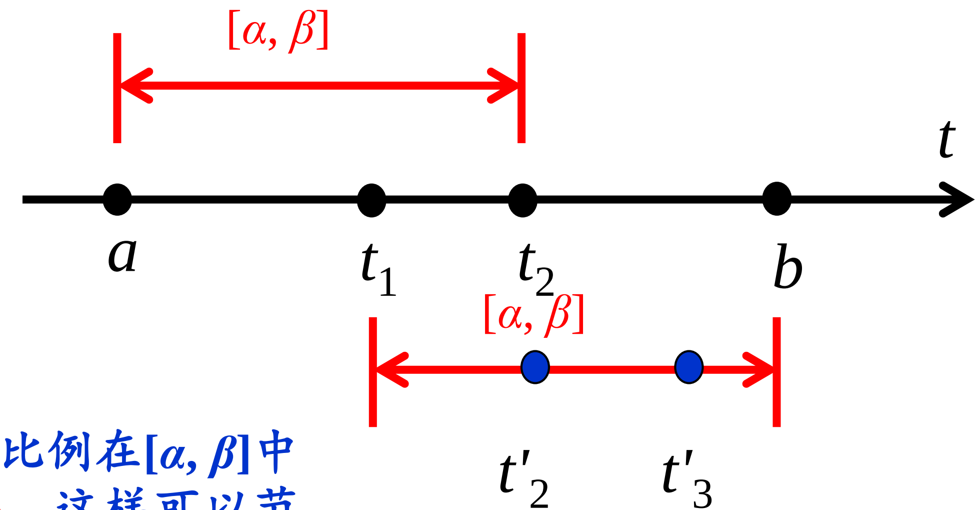


- 方式二：最终区间最短

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$\varphi_1 \le \varphi_2$，转第 4 步。否则转第 5 步;

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

第 5 步　若$b - t_1 \le \varepsilon$， 停止迭代，输出t_2 。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,…), 使最终包含全局极小 t^* 的待探索区间长度最小?**

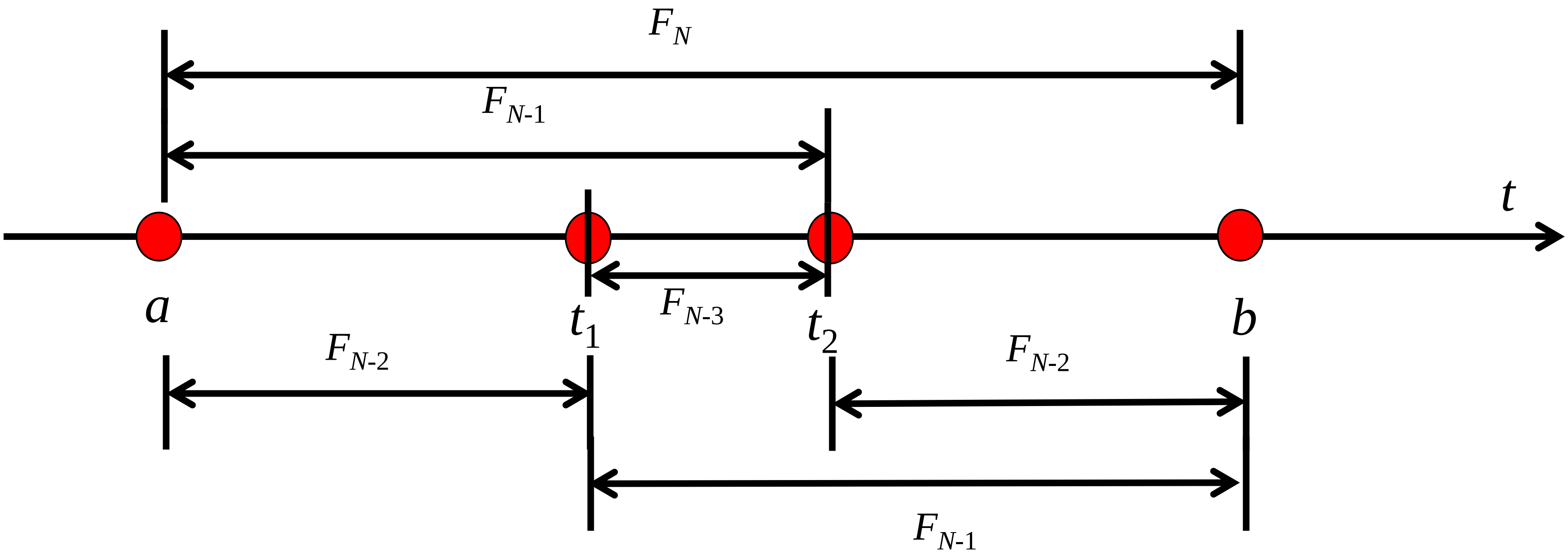
结论：可描述为非线性规划问题，可获得全局最优解！

- **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，……

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

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

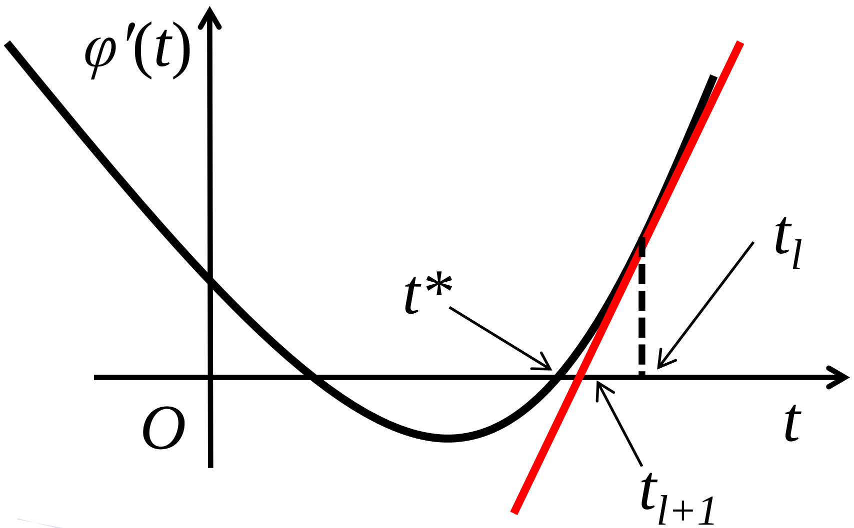
0.618法是Fibonacci法的极限形式

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



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

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

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

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：

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

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

解：

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

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

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

$$t_3=\frac{577}{408}\approx 1.41421\color{red}{57}$$

实质：解方程 $t^2-2=0$ **或求** $\sqrt{2}$

非精确一维搜索

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

基本假设：

$\varphi(t)$连续可导

$\varphi'(0) < 0$

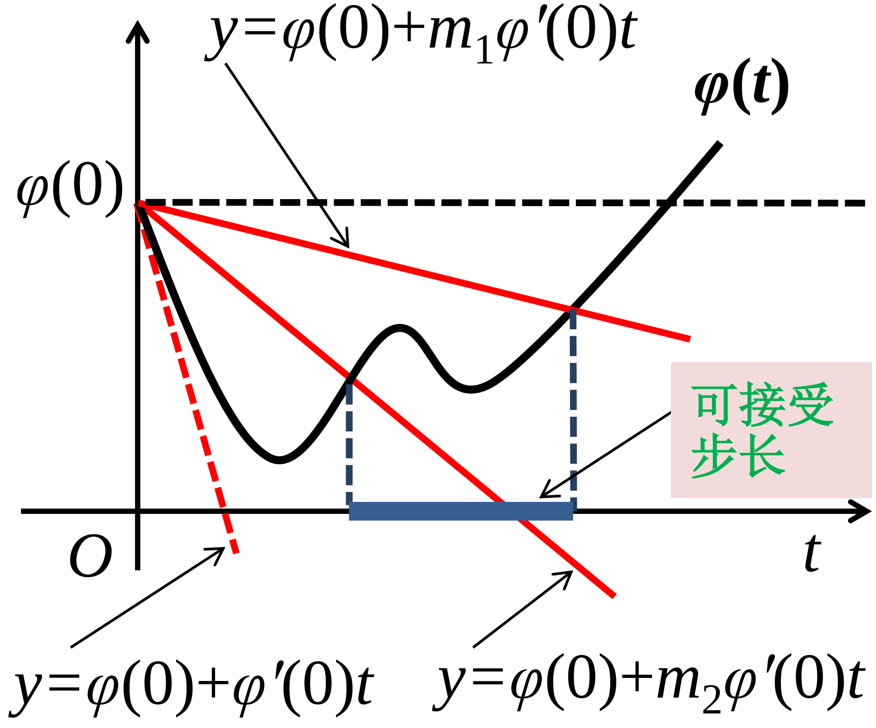
非精确一维搜索方法并不试图寻找(LS)问题的全局或局部最优解，而是致力于获得一个满意解，因此，它的方法包含两个结构要素：

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：

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



步长规则的本质含义：

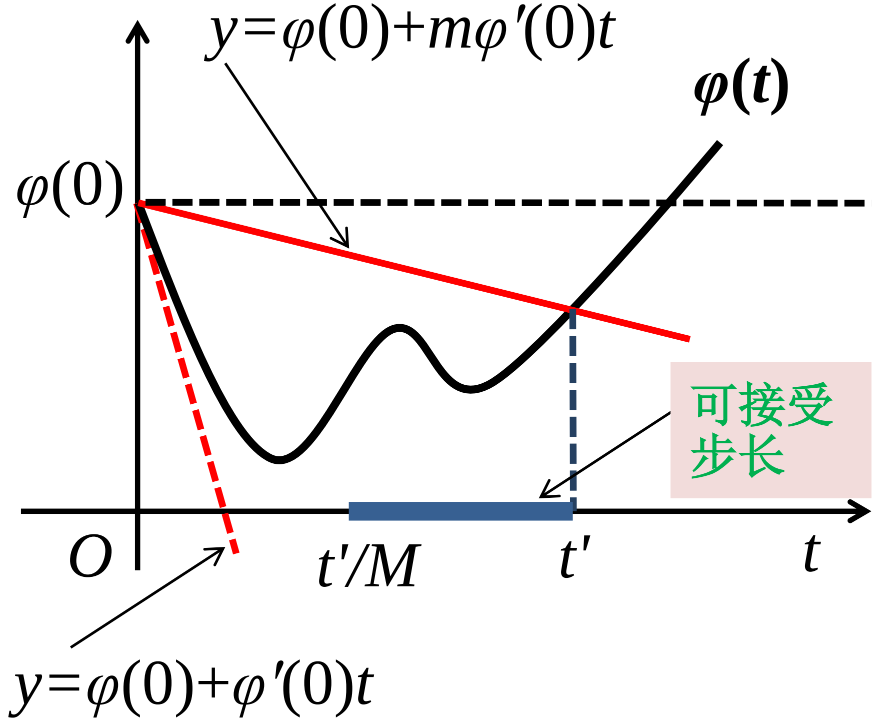
有足够下降，不过分靠近 0

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：

$$\begin{cases} \varphi(t_k) \le \varphi(0) + m t_k \varphi'(0) \\ \varphi(M t_k) \ge \varphi(0) + m M t_k \varphi'(0) \end{cases}$$



步长规则的本质含义：

有足够下降，不过分靠近 0

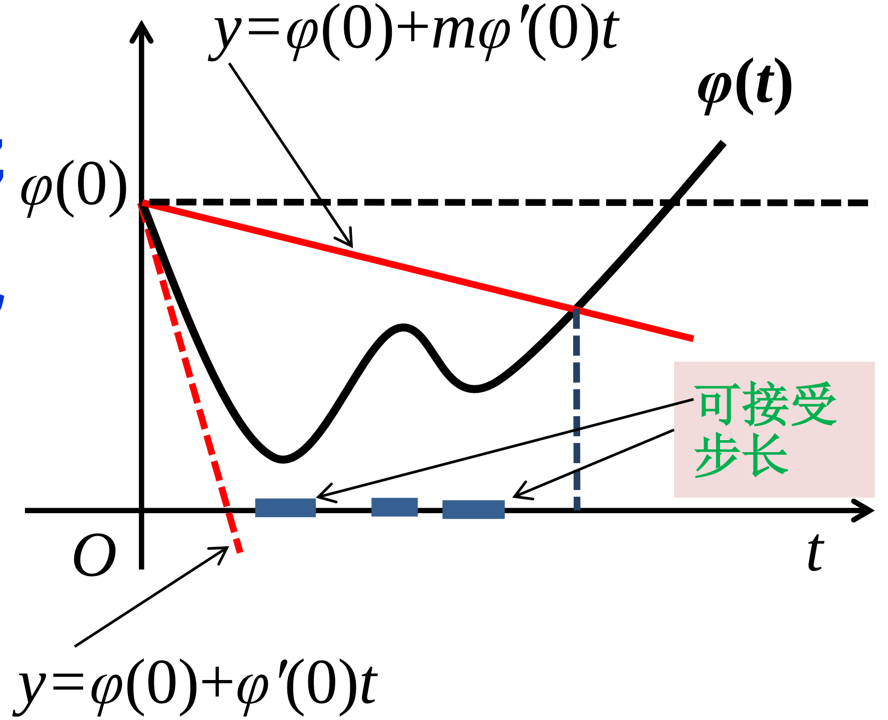
定理：若$\varphi(t)$在$[0, +\infty)$有下界，则满足Armijo准则的步长存在。

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：

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

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

- P152 10