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

非线性规划概述
Nonlinear Programming

电信学部·自动化科学与工程学院
系统工程研究所
吴江

Outline

- 非线性规划问题
- 非线性规划的基本概念
- 非线性规划问题的一般求解框架

非线性规划问题

$$\begin{aligned} \min \quad & z = f(x) \\ s.t \quad & x \in D, D \subset R^n \end{aligned}$$

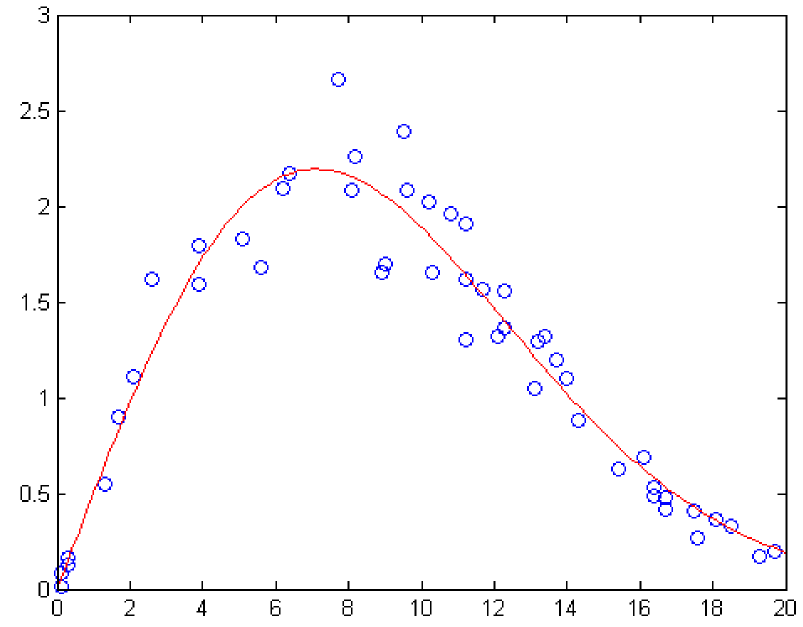
$$\begin{aligned} \min \quad & z = f(x) \\ s.t. \quad & A^{(1)}x \le b^{(1)} \\ & A^{(2)}x = b^{(2)} \end{aligned}$$

$$\begin{aligned} \min \quad & z = f(x) \\ s.t. \quad & A^{(1)}x = b^{(1)} \\ & x_j \in I \end{aligned}$$

$$\begin{aligned} \min \quad & z = f(x) \\ s.t. \quad & g_i(x) \le 0 \\ & h_j(x) = 0 \end{aligned}$$

$f(x)$, $g_i(x)$, $h_j(x)$**至少有一个**是非线性函数

例：曲线的最优拟合



- 理论分析表明，物理量y与另外n个物理量x_1, …, x_n之间具有如下函数关系：

$$y = g(x_1, x_2, \ldots, x_n; c_1, c_2, \ldots, c_k)$$

- 其中c_1, …, c_k为k个未知物理常数或参数。现在为了获得完整的上述函数关系，进行了m次实验，得到以下观测结果：

$$(x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}) \rightarrow y^{(i)}, i = 1, 2, \ldots, m$$

- 问题：如何由上述实验结果确定出c_1, …, c_k的**最优估计**?

非线性规划：例子

$$y=g(x_1,x_2,\dots,x_n;c_1,c_2,\dots,c_k)\,;(x_1^{(i)},x_2^{(i)},\dots,x_n^{(i)})\to y^{(i)};i=1,2,\dots,m$$

模型一

$$\begin{cases}\min\limits_{c_1,c_2,\cdots,c_k}\left\{\max\limits_{1\le i\le m}\left|y^{(i)}-g(x_1^{(i)},x_2^{(i)},\dots,x_n^{(i)};c_1,c_2,\dots,c_k)\right|\right\}\\ s.t.\quad (c_1,c_2,\dots,c_k)\in D\end{cases}$$

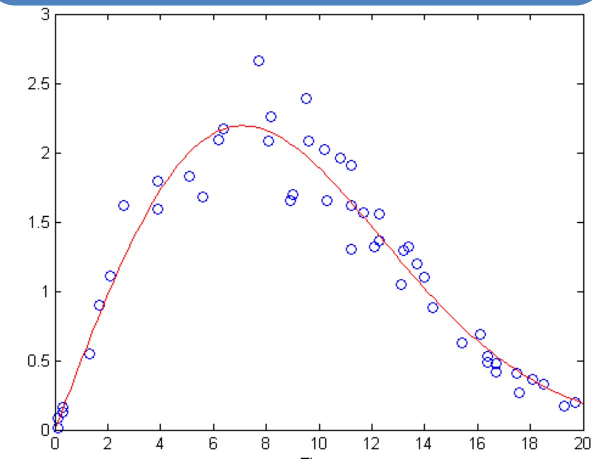
是否容易求解？

参数允许取值范围、参数的自然约束

$$\begin{cases}\min\limits_{c_1,c_2,\cdots,c_k,\varepsilon}\ \varepsilon\\ s.t.\quad \left|y^{(i)}-g(x_1^{(i)},x_2^{(i)},\dots,x_n^{(i)};c_1,c_2,\dots,c_k)\right|\le\varepsilon;\\ \qquad i=1,2,\cdots,m\\ (c_1,c_2,\dots,c_k)\in D\end{cases}$$

模型二(模型一的转化)

目标函数与约束的灵活建模



$$\begin{cases}\min\limits_{c_1,c_2,\cdots,c_k}\ \sum\limits_{i=1}^{m}\left(y^{(i)}-g(x_1^{(i)},x_2^{(i)},\dots,x_n^{(i)};c_1,c_2,\dots,c_k)\right)^2\\ s.t.\quad (c_1,c_2,\dots,c_k)\in D\end{cases}$$

模型三

应用最广泛的最小二乘模型

非线性规划的基本概念

$$\min \quad z = f(x)$$ 目标函数

$$s.t \quad x \in D, D \subset R^n$$ 可行域

无约束优化

$$\min \quad z = f(x)$$
$$s.t \quad x \in R^n$$

约束优化

$$\min \quad z = f(x)$$
$$s.t \quad x \in D, D \subset R^n$$

完整形式：(NLP)

$$\boxed{\begin{aligned} \min \quad & z = f(x) \\ s.t. \quad & g_i(x) \le 0, i = 1,2,\ldots p \\ & h_j(x) = 0, j = 1,2,\ldots q \\ & x \in R \end{aligned}}$$

定义一

- 定义1: 对于(NLP), 若 $x^* \in D$, 满足

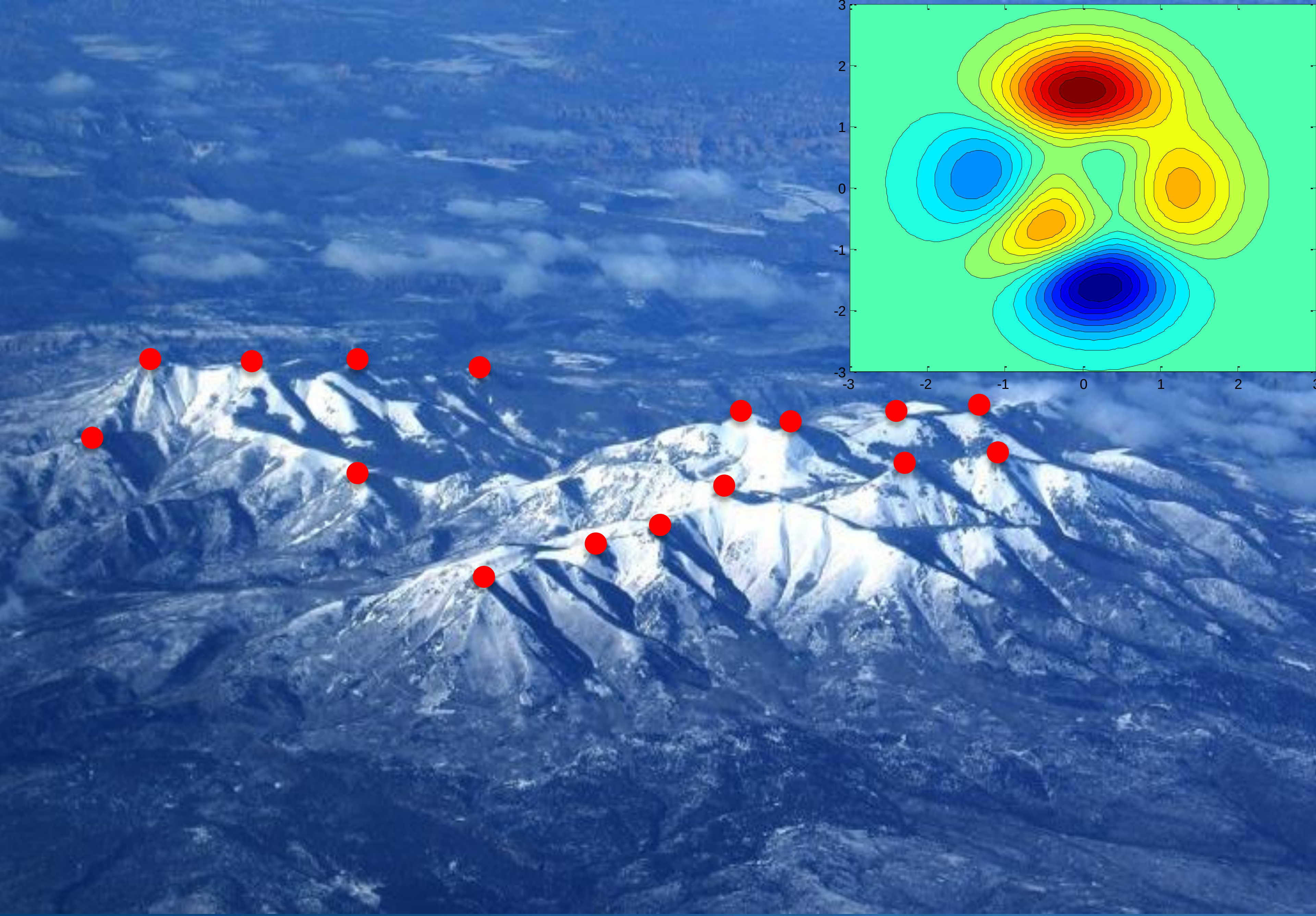
$$\forall x \in D \Rightarrow f(x^*) \le f(x)$$

则称x^*是该(NLP)的一个全局最优解. 并称$f(x^*)$为该(NLP)问题的全局最优值.

- 定义2: 对于(NLP), 若 $x^* \in D$, 且存在$\delta>0$使得:

$$\forall x \in D, \| x - x^* \| < \delta \Rightarrow f(x^*) \le f(x)$$

则称x^*是该(NLP)的一个局部极优解. 并称$f(x^*)$为该(NLP)问题的局部极优值.


3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3

非线性规划问题的求解

无约束优化 Euler, 1755

$$\min \ f(x_1,\ldots,x_n)$$

$$\nabla f(\boldsymbol{x}) = 0$$

约束优化 Lagrange, 1797

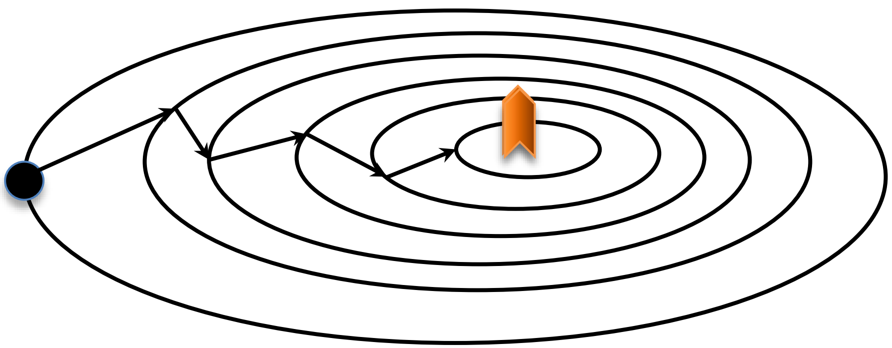
$$\min \quad f(x_1,\ldots,x_n)$$

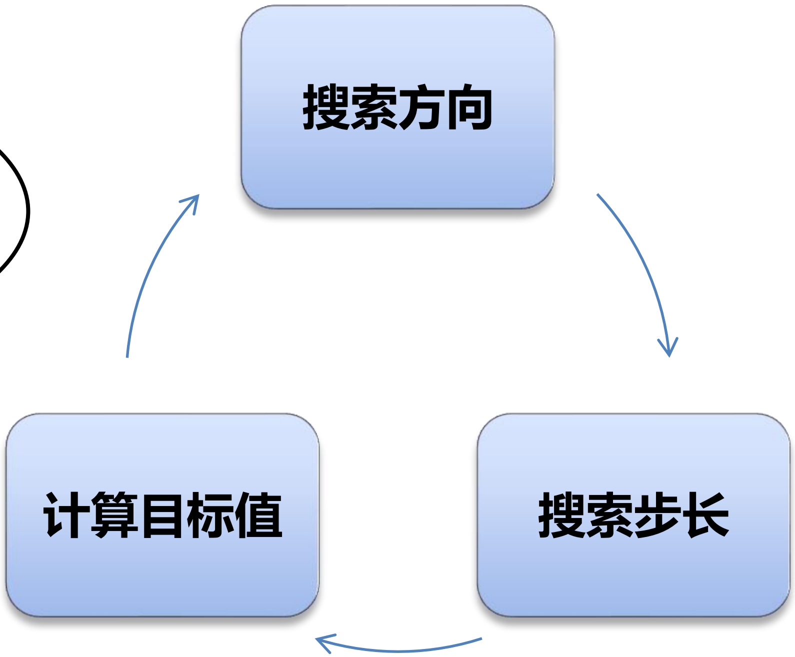
$$\text{s.t.} \quad g_k(x_1,\ldots,x_n) = 0 \qquad k = 1,\ldots,m$$

求解非线性方程组 ≈ 求解非线性规划

求解非线性规划问题的一般框架
——瞎子爬山法

华罗庚





求解非线性规划问题的一般框架

单纯形法的对比

步骤	内容
1	产生初始点: $x^0 \in D$或$x^0 \in R^n, k=0$
2	由x^k产生x^{k+1}, $k->k+1$ 一般$x^{k+1}=x^k+t_k p^k$ 构造 p^k (**搜索方向**) 确定恰当的t_k (**一维搜索**)
3	判断x^k是否可接受, 是->停 否->2

大部分算法讨论的核心问题

定义二

- 定义3: $f: R^n \to R, \bar{x} \in R^n, p \in R^n, p \neq 0$. 若存在 $\delta > 0$ 使 $\forall t \in (0, \delta)$ 有 $f(\bar{x} + tp) < f(\bar{x})$, 则向量$p$是$f(x)$在 $\bar{x}$ 的下降方向.
- 定义4: 对于一个NLP问题, 设D为其可行域. $\bar{x} \in D$, 且$p \in R^n, p \neq 0$. 若存在$t>0$使得 $\bar{x} + tp \in D$.则p是点 $\bar{x}$ 的一个可行方向.
- 定义5:对于一个NLP问题, 设D为其可行域. $\bar{x} \in D$, 且$p \in R^n, p \neq 0$. 若存在$\delta>0$使得 $\forall t \in [0, \delta], \bar{x} + tp \in D$ 则p是点 $\bar{x}$ 的一个可行下降方向.

一般算法构造的搜索方向是一个***可行下降***方向

问题

- (局部)最优解具有那些性质?怎样判定?
- 什么样的全局性质可以使得寻找全局最优解比较容易?
- 怎样构造(可行的)下降方向?
- 怎样确定步长?
- 怎样确定初始点?
- 算法的终止准则?