最优化简介

一、最优化问题

最优化： 在一定条件下极大化或极小化某函数. 对应于实际问题，就是在所有可行方案中选取最佳方案，使效益最大或成本最低。

现实生活中的优化问题

问题：一座巨型建筑高宽成何比例时，最美观？

问题：一根绳子围成何种形状，包围的面积最大？

结构设计问题：机头坡度为多少时，才能使飞机飞行时、高铁运行时稳定性最强？

股市有风险，投资需谨慎！

投资组合：一种理财渠道。

利用现有资金在股票、债券等之间进行组合，以期获得丰厚回报。

投资组合可以分散风险，避免“把鸡蛋放在一个篮子里”

如何进行投资组合，才能使可能的收益最大，可能遭受的损失最小？——最优化问题

应 用

- **经济金融：**投资组合

诺贝尔经济学奖得主马尔柯维茨(1990)提出关于投资组合问题的均值—方差二次规划模型。

- **交通运输：**列车运行时刻、物流运输等

如何在确保列车运行安全的前提下，组织列车运行计划，使运输效益最大。

- **人工智能、航空航天：**无人驾驶中的智能控制，航天器的结构设计等。

飞机结构设计中的减颤振问题就是一非线性优化问题。

- **管理科学：**管理学最常用的建模方法和工具。

它与数理统计、数值仿真构成管理学的三大基本方法。

- **信息科学与生命科学：**数据挖掘、信号处理、机器学习、模式识别，医学图像等。

- **运筹学**

最优化是运筹学的基础学科和重要分支，也是应用数学的一个重要研究分支。

二、模型与分类

数学模型

费用函数

$$\begin{aligned} &\min \quad f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{x} \in \Omega \end{aligned}$$

决策集

subject to

决策变量

效益函数

$$\begin{aligned} &\boxed{\max} \; f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{x} \in \Omega \end{aligned} \quad \longrightarrow \quad \begin{aligned} &\boxed{\min \quad -f(\boldsymbol{x})} \\ &\text{s.t. } \boldsymbol{x} \in \Omega \end{aligned}$$

$$\min \quad f(\boldsymbol{x})$$

决策集

$$\text{s.t. } \boldsymbol{x} \in \Omega$$

决策变量

可行解

不等式约束

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid c_i(\boldsymbol{x}) = 0, \ i \in \mathcal{E}; \quad c_i(\boldsymbol{x}) \geqslant 0, \ i \in \mathcal{I}\}$$

等式约束

等式约束指标集

不等式约束指标集

最优化问题分类

1、根据约束划分

无约束优化 $\min\limits_{\boxed{\boldsymbol{x}\in\mathbb{R}^n}} \quad f(\boldsymbol{x})$

约束优化

$$\begin{aligned} &\min \quad f(\boldsymbol{x}) \\ &\boxed{\text{ s.t. } \boldsymbol{x} \in \Omega} \end{aligned}$$

最优化问题分类

1、根据约束划分

无约束优化 $$\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad f(\boldsymbol{x})$$

约束优化 $$\begin{aligned} &\min \quad f(\boldsymbol{x}) \\ &\text{s.t.} \ \boldsymbol{x} \in \Omega \end{aligned}$$

约束优化问题一般比无约束优化问题难解，但有时可以相互转化

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1\} \Longleftrightarrow \min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$$

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) \Longleftrightarrow \min_{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}}\{t \mid t-f(\boldsymbol{x}) \geqslant 0\}$$

2、根据函数的线性度划分

线性规划： 目标函数及约束函数均线性

线性规划

$$
\begin{aligned}
\min\ & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\
\text{s.t. } & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geqslant \boldsymbol{0}
\end{aligned}
$$

非线性规划： 目标函数或约束中含有非线性函数

$$
\begin{aligned}
\min\ & -3x_1 + \boxed{\ln(x_2^2 + x_3^2)} \\
\text{s.t. } & x_1 + x_2 + \boxed{e^{x_3}} \le 20 \\
& -x_1 + 2x_2 + \boxed{x_3^2} = 0
\end{aligned}
$$

非线性项

一类常用的非线性规划

二次规划

目标函数二次

$$\begin{aligned}\min\ & \boxed{Q(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{x} + \boldsymbol{g}^{\mathrm{T}}\boldsymbol{x}} \\ \text{s.t.}\ & \boxed{\begin{aligned}&\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} = b_i,\ \ i \in \mathcal{E} \\ &\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} \geqslant b_i,\ \ i \in \mathcal{I}\end{aligned}}\end{aligned}$$

约束线性

线性规划和二次规划问题是两类特殊的最优化问题。目前已有比较完善的理论和求解方法。

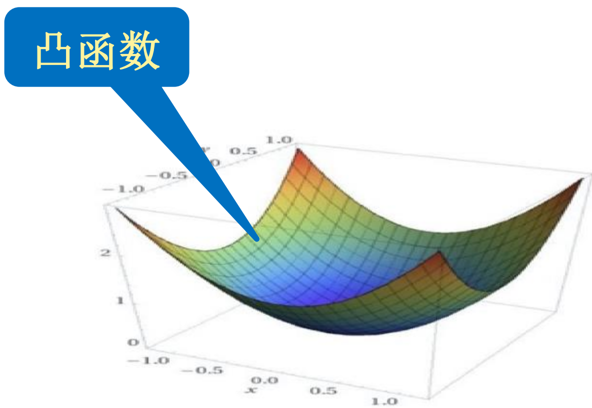
3、根据目标函数和可行域的凸性划分：

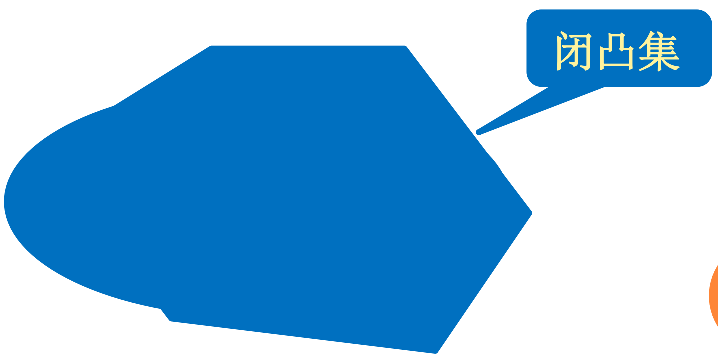
凸函数

凸规划

$$\min \quad f(\boldsymbol{x})$$
$$\text{s.t.} \ \boldsymbol{x} \in \Omega$$

闭凸集





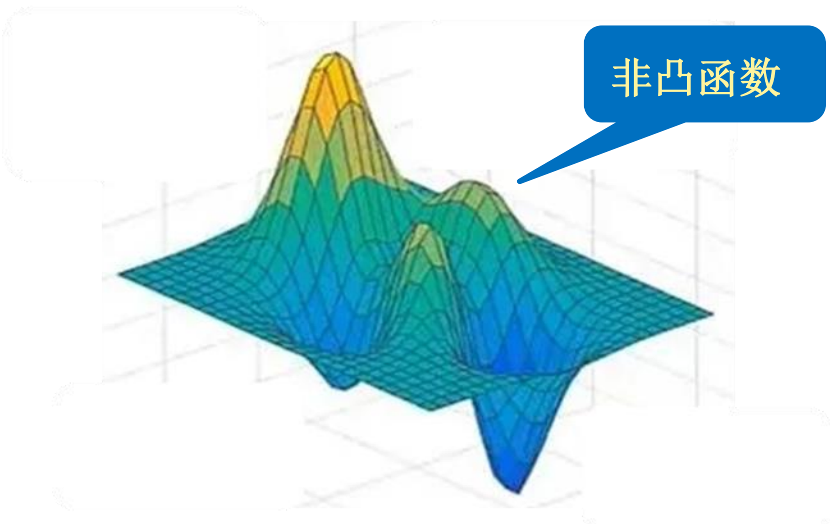
3、根据目标函数和可行域的凸性划分：

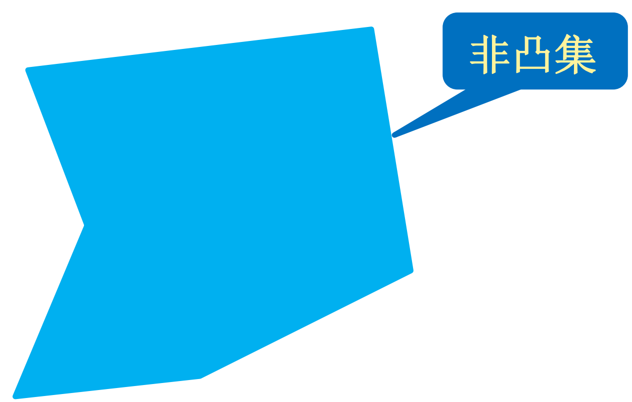
凸规划：目标函数为凸函数，可行域为闭凸集。

非凸优化：目标函数非凸或可行域非凸

凸规划问题：局部最优解即全局最优解，求解相对容易。

非凸优化问题：难于求解，特别是全局最优解。





4、根据函数的解析性质划分

光滑优化：所有函数都连续可微, 如多项式优化

$$\begin{aligned} \min\ & \left(x_1 - \frac{4}{9}\right)^2 + (x_2 - 2)^2 \\ \text{s.t.}\ & -x_1^2 + x_2 \geqslant 0 \\ & x_1 + x_2 \leqslant 6 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

非光滑优化：函数含不可微项

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - b\|^2 + \boxed{\lambda\|\boldsymbol{x}\|_1}$$

5、根据可行域中可行点的个数划分

连续优化：可行域含有无穷多个不可数的点且可行域中的点连续变化。

$$\min\ f(x)$$
$$\text{s.t.}\ \boxed{x \in [a,b]}$$

机头坡度为多少时，才能使高铁运行时稳定性最强？

离散优化：可行域含有有限(可数)个点,即在由有限个点或可数个点组成的可行域中寻求最优解。

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t.}\ \boxed{x \in \{0,1,2,3,\cdots\}} \end{aligned}$$

班里谁个子最高?

离散优化：可行域含有有限(可数)个点，即在由有限个点或可数个点组成的可行域中寻求最优解。

多数情况下，该问题的可行域是通过某些元素的排列组合产生的．因此又称组合优化。

简单的组合优化问题：

烧水泡茶：洗水壶，洗茶壶，洗茶杯，烧水，放茶叶。

如何安排工序才能在最短的时间喝上茶？（华罗庚《统筹方法》）

经典的组合优化问题

装箱问题、 旅行商问题(销售商问题)、中国邮路问题。

有多个结构相同、大小相等的箱子及多个大小不同的物品。如何用最少的箱子将全部物品装入箱内，或就现有的箱子装入尽可能多的物品.

给定多个城市和每两城市间的距离。寻求一条最短线路，使得每座城市访问一次并最终回到起始城市

给定多条街道(村庄)及两两之间的距离。寻求一条最短路，从邮局出发，走遍所有街道（村庄），并最终回到邮局(数学家管梅谷1950s）

求解算法的差异：

- **连续优化：** 可借助函数的梯度信息建立梯度类算法。

- **离散优化：** 不能有效利用函数的梯度信息.

  若将整数变量松弛为实数，则求解后者得到的最优解无论如何取整都不能保证它是原问题的最优解.

- **问题联系：** 有些离散优化问题可通过松弛为连续优化问题求解；连续优化问题的优化技术，如对偶，可移植到离散优化问题的研究中.

根据变量的特殊要求，离散优化又分离出

整数规划：所有变量都取整数（离散优化）

0-1规划：所有变量取0或1（离散优化）

混合整数规划:部分变量为整数变量,其余变量为连续变量

混合0-1规划：部分变量取0或1,其余变量为连续变量

稀疏优化：要求最优解稀疏，即非零个数尽可能少,即

$$\|\boldsymbol{x}\|_0 \leqslant k$$

6、根据模型参数的确定性划分

确定规划：所有参数都是确定的

不确定规划：含有不确定取值的参数

随机规划：不确定参数服从某种概率分布

模糊规划：源自对事物的不确定性判断，如高矮、黑白、胖瘦、好坏等。（**研究工具：**隶属度）

三、最优解的概念

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \Omega \end{aligned}$$

可行解： $\boldsymbol{x} \in \Omega$

全局最优解： $\boldsymbol{x}^* \in \Omega$，对任意 $\boldsymbol{x} \in \Omega$ 都有 $f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x})$

全局最优

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \Omega} f(\boldsymbol{x})$$

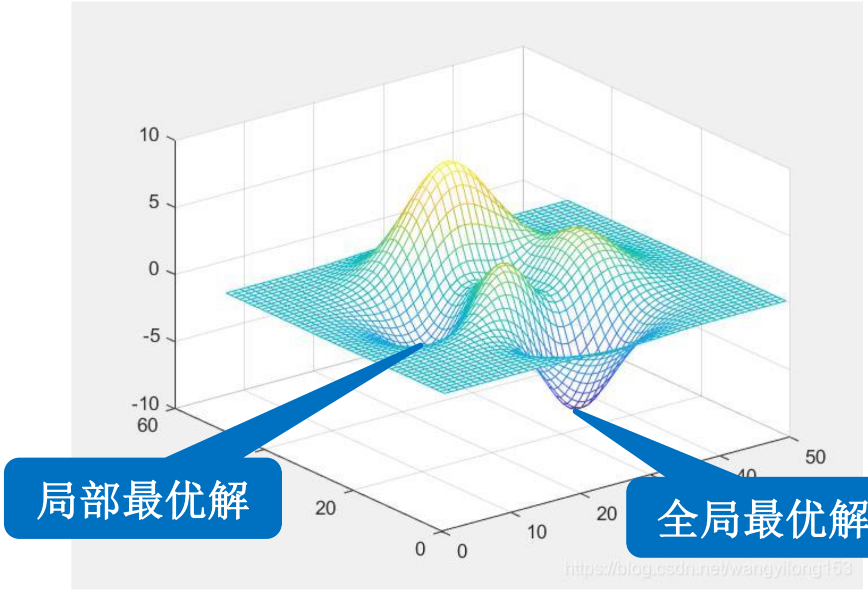
the argument of minimum

局部最优解： $\boldsymbol{x}^* \in \Omega$ ，存在邻域 $N(\boldsymbol{x}^*, \delta)$, 使对任意的

$\boldsymbol{x} \in N(\boldsymbol{x}^*, \delta) \cap \Omega$, 均有 $f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x})$

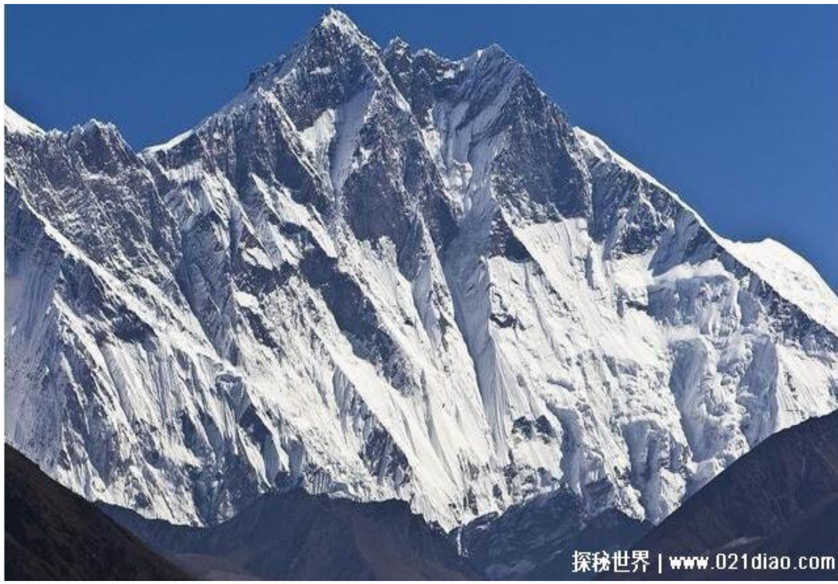
严格局部最优解： 对任意 $\boldsymbol{x} \in N(\boldsymbol{x}^*, \delta) \cap \Omega$, $\boldsymbol{x} \neq \boldsymbol{x}^*$

都有 $f(\boldsymbol{x}^*) < f(\boldsymbol{x})$



全球最高的峰：

全球最深的沟：



珠峰：8848m

马里亚纳海沟：11034m

最优值：目标函数在最优解处的值。

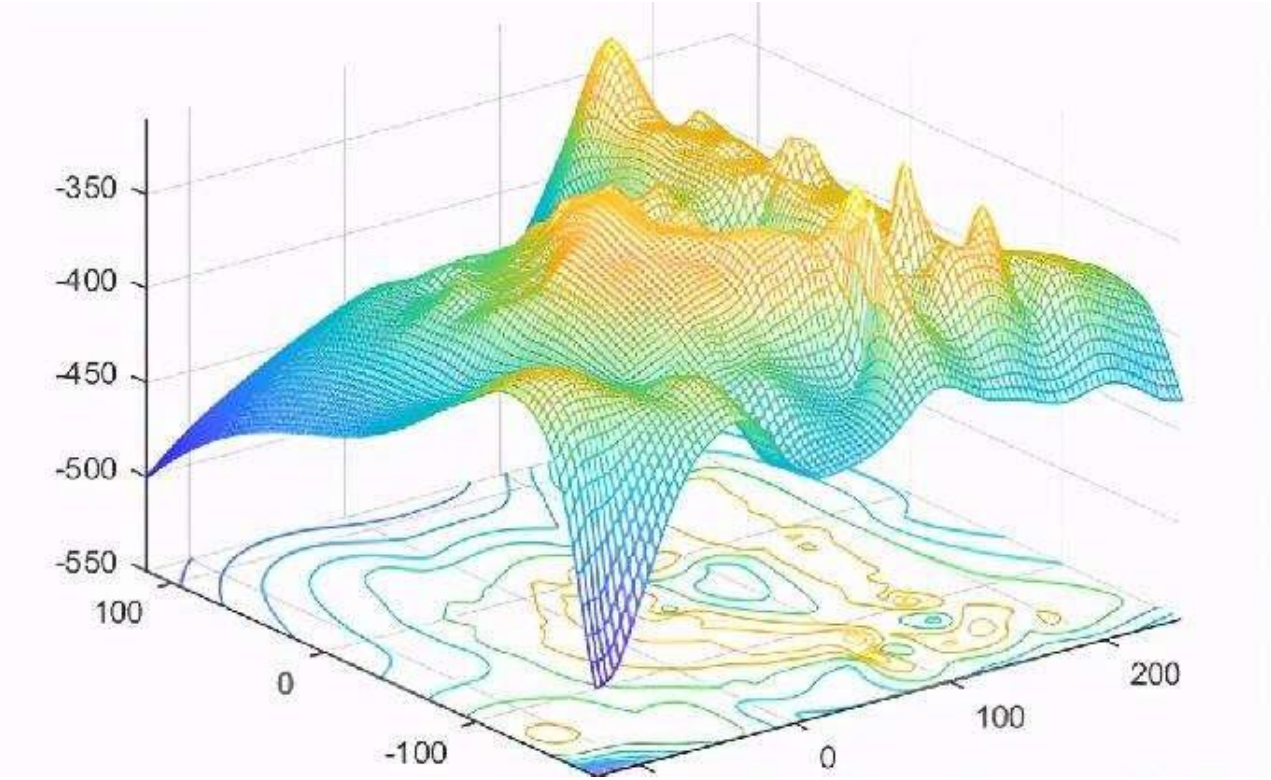
若目标函数在可行域上有下界，但无最优解，则目标函数在可行域上的下确界称为优化问题的最优值。

$$\inf\{f(\boldsymbol{x}) \mid \boldsymbol{x} \in \Omega\}$$

二元函数 $f(x) = x_1^2 + (1 - x_1 x_2)^2$ 的最优值为0，

但只能在 $x_1 = 1/x_2$ 且 $x_2 \to \infty$ 时达到。

最优值 $f^* = \inf\limits_{x \in \Omega} f(x)$



目标函数结构极其复杂。

一般情况下，得到局部最优解就OK了。

方法概述与算法评价

一、常见的优化方法

$$x^* \stackrel{?}{=} \arg\min_{x\in\Omega} f(x)$$

1. **解析法**　借助微分学、变分学等数学工具通过逻辑推理与分析运算给出问题最优解的**解析式**。

- **二次优化问题**　$\min\limits_{x\in\mathbb{R}^n} \frac{1}{2}x^{\mathrm{T}}Ax - b^{\mathrm{T}}x$

若Hesse阵正定，最优解：$x = A^{-1}b.$

可分离的优化问题

变量可分离

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{a}\|^2 + \lambda\|\boldsymbol{x}\|_1 \quad \text{其中 } \boldsymbol{a}\in\mathbb{R}^n, \lambda > 0.$$

$$\sum_{i=1}^{n}\left(\min_{\boldsymbol{x}_i\in\mathbb{R}} \frac{1}{2}(x_i-a_i)^2 + \lambda|x_i|\right)$$

一元优化问题

求解子问题

$$\min_{\boldsymbol{x}_i\in\mathbb{R}} \frac{1}{2}(x_i-a_i)^2 + \lambda|x_i|$$

$$= \begin{cases} \min\limits_{\boldsymbol{x}_i\in\mathbb{R}} \frac{1}{2}(x_i-a_i)^2 + \lambda x_i, & a_i \geqslant 0 \\ \min\limits_{\boldsymbol{x}_i\in\mathbb{R}} \frac{1}{2}(x_i-a_i)^2 - \lambda x_i, & a_i < 0 \end{cases}$$

最优解

$$x_i^* = \begin{cases} 0, & \text{若 } |a_i| < \lambda \\ a_i - \operatorname{sgn}(a_i)\lambda, & \text{若 } |a_i| > \lambda \end{cases} = \operatorname{sgn}(a_i)(|a_i|-\lambda)_+.$$

某些一元优化问题，如四次以下的优化问题

$$\begin{aligned} \min \quad & ax^4 + bx^3 + cx^2 + dx + e \\ \text{s.t.} \ & 1 \leqslant x \leqslant 10 \end{aligned}$$

分析与求解： 计算目标函数的导数零点，即利用盛金公式求解一元三次方程

$$4ax^3 + 3bx^2 + 2cx + d = 0$$

得目标函数的稳定点 x_1，x_2，x_3

依次计算可行域内的稳定点的函数值，及端点处的函数值。最小者即为优化问题的最优值。

解析法：

优点：精确、简洁、直观，便于理论分析；

缺 陷：计算量大，只限于结构特殊的优化问题.

2. 图解法与实验法(手工作坊)

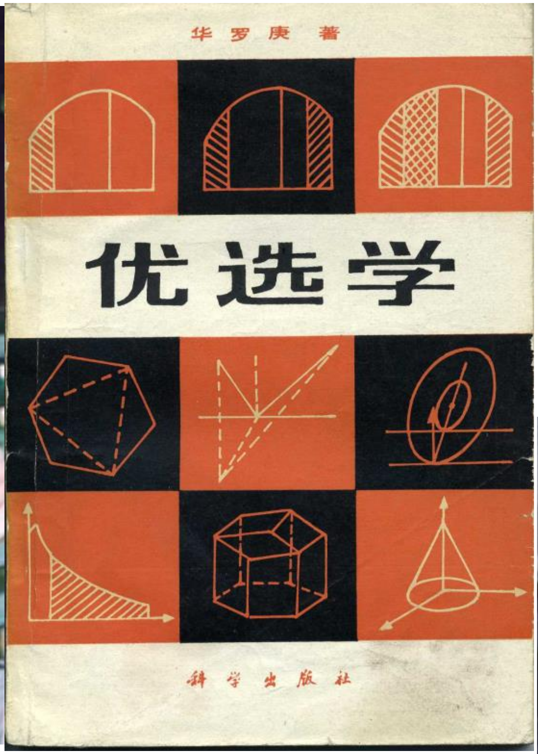
图解法：根据目标函数的图像求解

实验法：依据一定规则对变量取不同的值，通过实验观察目标函数值的变化规律，找出问题的最优解。

优点：操作简单、通俗易懂

缺点：效率低、要求变量维数小。

华罗庚的优选法



麦场设置

五粮液生产工艺

3. **形式转化法** 利用问题结构或最优性条件将其转化为有容易求解的另一类数学问题，然后对后者套用现有的方法求解。

最优性条件

最优化问题

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}) \\
\text{s.t.} \quad & c_i(\boldsymbol{x}) = 0, \quad i \in \mathcal{E}
\end{aligned}
$$

方程组问题

$$
\begin{cases}
\nabla_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{0}, \\
c_i(\boldsymbol{x}) = 0.
\end{cases}
$$

优点： 为问题的解决打开了一个思路

缺点： 未提供具体算法，转化需要条件。

4. **智能算法：**受自然界规律启迪并根据其原理来模拟某些自然现象而建立的一种算法。

蚂蚁寻找食物

常见算法：遗传算法、蚁群算法、模拟退火算法、神经网络算法等。属于启发式算法.

5. **数值迭代法** 利用函数值或梯度信息按一定规则从一个点产生一个新的点，直到不能改进为止。得到的是数值解，是近似解。

常见方法：模式搜索法、梯度算法

（1）模式搜索法：根据函数值变化规律探测函数的下降方向并沿着该方向寻求更优的点。

优点：简单、直观、无需计算函数梯度

缺点：适用于变量较少、约束简单的情形，效率较低

常见算法：坐标轮转法、Hooke-Jeeves法、Powell共轭方向法、单纯形法

(2)梯度方法：最优解、最优值未知。

从一个点出发，利用目标函数梯度信息，按照某种策略逐步调优，得到一个更好的点。然后依次重复，直至不能改进。

优点：收敛速度快，容易建立其理论性质；

缺点：解析性质要求较高，如函数光滑、梯度容易计算等；

种类：线搜索方法、信赖域方法。

A：线搜索方法：在每一步迭代，首先基于目标函数梯度信息产生搜索方向，然后沿该方向寻求目标函数的(近似)最小值点。重复上述过程，直到不能改进。

算法框架

初始步

计数器

步1、取初始点 $\boldsymbol{x}_0 \in \mathbb{R}^n$，及有关参数． 令 $k=0$

步2、验证停机准则. 到家了吗？

行进方向

步3、计算当前迭代点 $\boldsymbol{x}_k$ 的搜索方向 $\boldsymbol{d}_k \in \mathbb{R}^n$

行进路程

步4、计算步长 $\alpha_k > 0$ 使满足

行进要求

$$f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) < f(\boldsymbol{x}_k)$$

步5、令 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$， $k = k+1$，转步2.

以旧换新

重复上述过程

步1、取初始点 $\boldsymbol{x}_0 \in \mathbb{R}^n$ 及有关参数．令 $k=0$

步2、验证停机准则.

步3、计算当前迭代点 $\boldsymbol{x}_k$ 的搜索方向 $\boldsymbol{d}_k \in \mathbb{R}^n$

步4、计算步长 $\alpha_k > 0$ 使满足 $f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) < f(\boldsymbol{x}_k)$

步5、令 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$，$k=k+1$，转步2.

有关说明

初始点： 对算法效率和数值结果产生影响，原点或随机选取；

算法参数： 有理论取值范围，但会严重影响算法效率；经验值；

停机准则： 最优性条件准则、点距准则、函数下降量准则。

攀登珠峰从南路，还是北路？

数值计算常用的停机准则

迭代停滞不前

- **点距准则** $\|x_{k+1}-x_k\|\leqslant\varepsilon$ 或 $\dfrac{\|x_{k+1}-x_k\|}{\|x_k\|}\leqslant\varepsilon$

迭代没有成效

- **函数下降量准则** $f(x_k)-f(x_{k+1})\leqslant\varepsilon$ 或 $\dfrac{f(x_k)-f(x_{k+1})}{|f(x_k)|}\leqslant\varepsilon$

满足基本要求

- **最优性条件准则** $\|\nabla f(x_k)\|\leqslant\varepsilon$
- **最大步数准则** 迭代次数限制在一定次数之内

迭代次数无法忍受

理论分析时，只能用最优性条件准则！

搜索方向：从当前迭代点到下一迭代点的行进方向。

一般为目标函数的下降方向。

迭代步长：从当前迭代点到下一步迭代点的行进距离。

通过线搜索产生，使目标函数有所下降。

核　心：搜索方向、迭代步长.

影响力：搜索方向大于迭代步长。选择比努力重要！

B. 信赖域方法：基于当前迭代点信息建立目标函数在该点邻域内的一个近似二次模型，然后将该模型小邻域内的最小值点作为新的迭代点。

$$\begin{aligned} f(\boldsymbol{x}_k+\boldsymbol{d}) &= f(\boldsymbol{x}_k)+\boldsymbol{d}^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{d}+o(\|\boldsymbol{d}\|^2) \\ &\approx f(\boldsymbol{x}_k)+\boldsymbol{d}^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{d} \end{aligned}$$

$$m_k(\boldsymbol{d}) = f(\boldsymbol{x}_k)+\boldsymbol{d}^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}$$

$\boldsymbol{B}_k$为$\nabla^2 f(\boldsymbol{x}_k)$或近似

$$\min_{\boldsymbol{d}\in\mathbb{R}^n} f(\boldsymbol{x}_k+\boldsymbol{d}) \Longrightarrow \min\ \{m_k(\boldsymbol{d}) \mid \boldsymbol{d}\in\mathbb{R}^n,\ \|\boldsymbol{d}\|\leqslant\Delta_k\}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k+\boldsymbol{d}_k$$

信赖域方法与线搜索方法：

线搜索方法：借助搜索方向将一个多元函数的极值问题化为一单元函数的极值问题；

$$\alpha_k = \arg\min_{\alpha \geqslant 0} f(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k).$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$$

信赖域方法： 在一个值得信赖的区域内用二次函数近似

二、算法评价体系

- 收敛性
- 收敛速度
- 稳定性
- 计算复杂性
- 数值效果
- 算法的研究焦点

1、收敛性

全局收敛：从任意的初始点出发，算法产生的迭代点列都收敛到问题的最优值点.

局部收敛：只有在初始点和最优值点具有某种程度的靠近时才能保证迭代点列收敛到最优值点.

弱收敛：迭代点列的某一聚点为优化问题的最优值点.

2、收敛速度

$$\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}^* \Longleftrightarrow \lim_{k\to\infty} \|\boldsymbol{x}_k - \boldsymbol{x}^*\| = 0$$

纵向比较 → 商收敛(Q-收敛)

横向比较 → 根收敛（R-收敛）

$$\frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|}$$

$$\limsup_{k\to\infty} \frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|}$$

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leqslant \kappa q^k,$$

$$\kappa > 0,\ 0 < q < 1$$

Q-线性收敛 前后两迭代点靠近最优值点的距离之比

$$\limsup_{k\to\infty}\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|}{\|\boldsymbol{x}_k-\boldsymbol{x}^*\|}\leqslant q.$$

Q-线性收敛： $0 < q < 1$

Q-超线性收敛： $q = 0$

前提：迭代点列 $\{\boldsymbol{x}_k\}$ 收敛到 $\boldsymbol{x}^*$.

Q-超线性收敛 $\limsup_{k\to\infty}\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|}{\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^r}\leqslant p,\quad 0\leqslant p<\infty,\ r>1$

r 越大，收敛速度越快。

R-线性收敛

借助一趋于零的等比数列度量迭代点收敛的快慢程度

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leqslant \kappa q^k, \quad \kappa > 0,\ 0 < q < 1$$

R-超线性收敛

存在$\kappa > 0$和收敛于零的正数列$\{q_k\}$使

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leqslant \kappa \prod_{i=0}^{k} q_i$$

Q-(超)线性收敛 $\Longrightarrow$ **R-(超)线性收敛**

二次终止性：对任意凸二次函数，从任意初始点出发，经有限步到达极小值点。

理论依据

（1）严格凸二次函数是非线性函数中形式最简单、条件最强的函数，一个好的算法理应在有限步内得到最优解；

（2）对一般的目标函数，它在最优值点附近可以用严格凸二次函数近似. 为此,对严格凸二次函数数值效果好的算法，对一般函数应有好的数值效果.

3、稳定性

初始数据的舍入误差会通过系列运算进行遗传和传播.

如果初始数据的误差对最终结果的影响较小, 即在计算过程中舍入误差增长缓慢，则称该算法是稳定的.

若输出结果的误差随初始数据的舍入误差呈恶性增长，则称该算法是不稳定的.

算法的稳定性还包括算法参数的取值对数值结果的影响。

原因分析：问题病态，算法缺陷；

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \boxed{\begin{matrix} b_1 \\ b_2 \\ b_3 \end{matrix}} \end{pmatrix} \Rightarrow \begin{cases} x_1 = b_1 \\ x_2 = \dfrac{b_2}{\epsilon} \\ x_3 = \dfrac{b_3}{\epsilon^2} \end{cases}$$

系数矩阵病态

解决办法：调比技术、正则化技术。

4、计算复杂性

算法中每次迭代需要的计算量和存储量.

收敛速度并不是衡量算法好坏的唯一标准！

每一步的计算量（计算复杂性）也很重要！

5、数值效果

数值效果是算法被接受的重要依据。有时会揭露某些可能的理论结果.

影响要素：参数和初始点的选取、程序的编制

评价标准：好的理论性质、诱人的数值效果.

6、算法的研究焦点

以问题为导向，针对一个特定的实际问题，设计专门的算法；

对一个既定算法，通过理论分析给出其适用的问题类，使其成为一个“指示”性算法.