最优化概述

Liu Zheng

同济大学电信学院

April 22, 2014

概述

- 最优化理论和方法是近二十多年来发展十分迅速的一个数学分支
- 在数学上，最优化是一种求极值的方法
- 最优化已经广泛的渗透到工程、经济、电子技术等领域

概述

- 在实际生活中，人们做任何事情，不管是分析问题，还是进行决策，都要用一种标准衡量以下是否达到来最优。（比如基金人投资）
- 在各种科学问题、工程问题、生产管理、社会经济问题中，人们总是希望在有限的资源条件下，用尽可能小的代价，获得最大的收益。（比如保险）

几个概念

- 最优化是从所有可能方案中选择最合理的一种，以达到最优目标的学科
- 最优方案是达到最优目标的方案
- 最优化方法是搜寻最优方案的方法
- 最优化理论就是最优化方法的理论

经典极值问题

包括

- 无约束极值问题

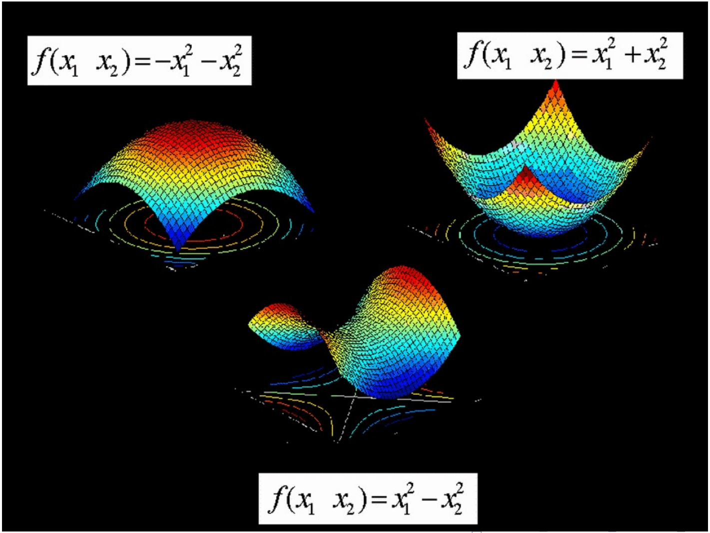
$$\min_{x} f(x)$$

- 约束条件下的极值问题

$$\begin{aligned} &\min_{x} f(x) \\ s.t.\quad & g_i(x) \leq 0, i = 1, 2, \cdots, m \\ & h_i(x) = 0, i = 1, 2, \cdots, n \end{aligned}$$

其中，极大值问题可以转换为极小值问题来进行求解。如求：$\max_{x} f(x)$ 可以转换为：$\min_{x} -f(x)$

无约束极值问题



最优化方法分类

(1) 线性最优化：目标函数和约束条件都是线性的则称为线性最优化。
非线性最优化：目标函数和约束条件如果含有非线性的，则称为非线性最优化。

(2) 静态最优化：如果可能的方案与时间无关，则是静态最优化问题。
动态最优化：如果可能的方案与时间有关，则是动态最优化问题

有约束最优化模型一般具有以下形式：

$$\min_{x} f(x) \quad \text{或} \quad \max_{x} f(x)$$
$$s.t. \quad \cdots\cdots \qquad\qquad s.t. \quad \cdots\cdots$$

其中 $f(x)$ 为目标函数，省略号表示约束式子，可以是等式约束，也可以是不等式约束。

最优化方法主要内容

根据目标函数，约束条件的特点将最优化方法包含的主要内容大致如下划分：

- 线性规划
- 整数规划
- 非线性规划
- 智能优化方法
- 变分法与动态规划

线性规划与整数规划

在许多线性规划问题中,要求最优解必须取整数.例如 所求的解是机器的台数、人数车辆船只数等.如果所得的解 中决策变量为分数或小数则不符合实际问题的要求.

对于一个规划问题,如果要求全部决策变量都取整数, 称为纯(或全)整数规划;如果仅要求部分决策变量取整数, 称为混合整数规划问题.有的问题要求决策变量仅取0或1两 个值,称为0-1规划问题.

整数规划简称为IP问题(Integer Programming,IP)或ILP问题(Integer Linear Programming).这里主要讨论的是整数线性规划。

整数规划

- 最优化问题中的所有变量均为整数时，这类问题称为整数规划问题。
- 如果线性规划中的所有变量均为整数时，称这类问题为线性整数规划问题。
- 整数规划可分为线性整数规划和非线性整数规划 ，以及混合整数规划等。
- 如果决策变量的取值要么为0，要么为1，则这样的规划问题称为0 - 1规划。

列子

　　某钢厂两个炼钢炉同时各用一种方法炼钢。第一种炼法每炉用 a 小时，第二种用 b 小时（包括清炉时间）。假定这两种炼法，每炉出钢都是 k 公斤，而炼 1 公斤钢的平均燃料费第一法为 m 元，第二法为 n 元。若要求在 c 小时内炼钢公斤数不少于 d ，试列出燃料费最省的两种方法的分配方案的数学模型。

解法

设用第一种炼法炼钢 x_1 炉，第二种炼钢 x_2 炉

$$\max z = k(mx + ny)$$

$$s.t. \begin{cases} ax_1 & \leq & c \\ bx_2 & \leq & c \\ k(x_1 + x_2) & \geq & d \\ x_1, x_2 & \geq & 0\text{且为整数} \end{cases}$$

具体求解运算在此不作过多叙述，都是矩阵运算，相信大家对纯公式推导演算没多大兴趣。

单纯形法

单纯形法是优美国数学家G.B.Dantzig提出的，该方法的基本出发点就是在可行域（凸集）的定点中搜索最优点。搜索的过程是一个迭代的过程。首先找到一个基本可行解，判别它是否为最优解，如不是就找一个更好的基本可行解，再进行判别。如此迭代，直至找到最优解，或者判定该问题无界。

凸集

定义： $\mathbf{X}$ 是 $\mathbf{E}^n$ 空间的点集， x_1 和 x_2 是 $\mathbf{X}$ 上任意两点，若对任意 $0 \leq \lambda \leq 1$ 使得

$$\lambda x_1 + (1-\lambda)x_2 \in \mathbf{X}$$

则称 $\mathbf{X}$ 为凸集。

Figure : 凸集

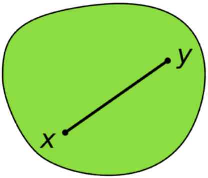
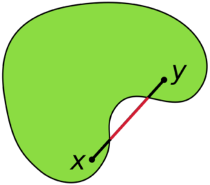


Figure : 非凸集



单纯形法基本形式

用单纯形法求解时，常将标准形式化为：

$$
\begin{aligned}
\min \quad & f = cx \\
s.t. \quad & Ax = b \\
& x \geq 0
\end{aligned}
$$

这里

$$
\begin{aligned}
A &= (a_{ij})_{m,n} \\
x &= (x_1, x_2, \cdots, x_n)^T \\
b &= (b_1, b_2, \cdots, b_n)^T \\
c &= (c_1, c_2, \cdots, c_n)
\end{aligned}
$$

解法

一般步骤如下：

（1）寻找一个初始的基本可行解。

（2）检查现行的基本可行解是否最优，如果为最优， 则停止迭代，已找到最优解，否则下一步。

（3）移至目标函数值有所改善的另一个基本可行解， 然后转回到步骤（2）。

算法思路如上，具体迭代、矩阵运算过于数学化，在此不作展开。

多目标规划

多目标规划法也是运筹学中的一个重要分支，它是在线性规划的基础上，为解决多目标决策问题而发展起来的一种科学管理的数学方法

多目标规划的概念是 1961年由美国数学家查尔斯和库柏首先提出的。

多目标规划标准型

求 $\mathbf{x} = x_1, x_2, \cdots, x_n \in D \subset E^n$ ，使

$$
\begin{aligned}
& \min \mathbf{f}(x) = \{f_1(x), f_2(x), \cdots, f_N(x)\} \\
s.t. \quad & h_l(\mathbf{x}) = 0, l = 1, 2, \cdots, L \\
& g_m(\mathbf{x}) \leq 0, m = 1, 2, \cdots, M
\end{aligned}
$$

式中， n 为自变量 $\mathbf{x}$ 的维数； L 为等式约束的数目； M 为不等式约束的数目

非线性规划问题的一般数学模型

$$\min f(x)$$

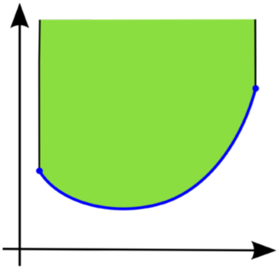
$$s.t. \begin{array}{rcl} g_i(x) & \leq & 0, i=1,2,\cdots,m \\ h_j(x) & = & 0, j=1,2,\cdots,l \end{array}$$

其中， $x \in E^n$ ， $f(x)$ 为目标函数， $g_i(x), h_j(x)$ 为约束函数，这些函数中至少有一个是非线性函数。

凸函数

定义：在某个开区间 C 内的凸函数 f 在 C 内连续，且在除可数个点之外的所有点可微。如果 C 是闭区间，那么 f 有可能在 C 的端点不连续。

Figure：凸函数



函数（蓝色）是凸的，当且仅当其上方的区域（绿色）是一个凸集。

下降迭代法的基本思想

任取一个初始迭代点 $\mathbf{x}^{(0)}$ ，在 $\mathbf{x}^{(0)}$ 处找一个下降方向 $\mathbf{p}^{(0)}$ ，移动 $\mathbf{x}^{(0)}$ 到 $\mathbf{x}^{(0)}+\lambda_0\mathbf{p}^{(0)}$ 处，令 $\mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\lambda_0\mathbf{p}^{(0)}$ ，显然 $f(\mathbf{x}^{(1)})<f(\mathbf{x}^{(0)})$ 。然后判断 $\mathbf{x}^{(0)}$ 是否为极小值点，若是，则停止迭代；否则，再从 $\mathbf{x}^{(1)}$ 出发，找比 $\mathbf{x}^{(1)}$ 更好的点 $\mathbf{x}^{(2)},\cdots$，如此继续，就产生了一个解点的序列 $\{\mathbf{x}^{(k)}\}$ ，满足

$$f(\mathbf{x}^{(0)})>f(\mathbf{x}^{(1)})>\cdots>f(\mathbf{x}^{(k)})>\cdots$$

变分法和动态规划

变分法涉及泛函，泛函又基于实变，以我的数学功底暂时无法完全理解，曾经在为专业课《工程光学》里提到过一次变分
原文为：

费马原理指出，光线从A点传到B点，经过任意多次的折射或反射，其光程为极值（极大值或极小值），可以用光程的一次变分为零来表示，即：

$$\delta\Delta = \delta\int_{A\to B} n(x, y, z)\,ds = 0$$

动态规划就不用多赘述了，

斐波那契数列(Fibonacci polynomial)

```
function fib(n)
   if n = 0 or n = 1
       return 1
   return fib(n − 1) + fib(n − 2)
```

当n=5时，fib(5)的计算过程如下:

1. fib(5)
2. fib(4) + fib(3)
3. (fib(3) + fib(2)) + (fib(2) + fib(1))
4. ((fib(2) + fib(1)) + (fib(1) + fib(0))) + ((fib(1) + fib(0)) + fib(1))
5. (((fib(1) + fib(0)) + fib(1)) + (fib(1) + fib(0))) + ((fib(1) + fib(0)) + fib(1))

$O(n)$


```
array map [0...n] = 0 => 0, 1 => 1
fib( n )
  if ( map m does not contain key n )
    m[n] := fib(n – 1) + fib(n – 2)
      return m[n]
```

将前n个已经算出的数保存在数组map中，这样在后面的计算中可以直接应用前面的结果，从而避免了重复计算。算法的运算时间变为$O(n)$

- 神经网络
- 模拟退火
- 遗传算法
- 蚁群算法

Overview