高等学校教材

优化方法及其在水工中的应用

浙江大学 汪树玉 主编

高 等 学 校 教 材

优化方法及其在水工中的应用

浙江大学 汪树玉 主编

水利电力出版社

（京）新登字115号

内 容 提 要

本书作为水利、土建、力学等理工科大学本科学生用的教材，首重介绍了数学规划和工程优化设计习用方法，以及如何塑造优化的数学模型等。全书共8章，系统论述了优化方法的基本概念和基础理论，无约束优化问题和约束优化问题的各种解法，结构优化的准则法和其他类型的优化问题，优化模型塑造及应用示例。本书概念叙述清楚，公式推导论证简明扼要，理论紧密结合实际应用。

本书除作为高校教材外，还可供从事水利、建筑、交通等的工程设计人员学习参考。

高等学校教材
优化方法及其在水工中的应用
浙江大学 汪树玉 主编
*
水利电力出版社出版
（北京三里河路6号）
新华书店北京发行所发行 · 各地新华书店经售
北京师范大学印刷厂印刷
*
787×1092毫米 16开本 12.75印张 285千字
1993年6月第一版 1993年6月北京第一次印刷
印数0001—1140册
ISBN 7-120-01757-8/TV·632
定价3.35元

前　言

优化问题是古典极值问题的延拓。人们在实践中一直向往着以最小的代价获得最大的收益，如材料消耗最少而产值最大，等等。就工程问题而言，早在17世纪中叶，伽利略就曾寻求变高度梁的最优形状。18世纪拉格朗日研究轴向荷载下柱的最轻设计。虽然在高等数学中，关于简单函数的极值问题，早已有所解决，但实际问题往往远为复杂，用简单的极值理论无法处理。40年代以来，由于电子计算机的出现和迅猛发展，以及数学规划论的长足进展，才使复杂的实际问题的优化成为可能。当前，全球资源日益短缺，为了适应生产的发展与生活水准的提高，更加要求用最少的资源和最优的方式去完成各种社会经济、生产、科技任务。因此，优化理论与方法这一有力工具，必然将引起人们普遍的关注。

本书作为水利、土建、力学等理工科大学本科学生用教材，着重介绍数学规划几个重要分支的内容、常用算法、工程优化设计的习用方法，以及如何塑造优化的数学模型等。全书有较宽的覆盖面，目的是使未来的科技工作者与工程师们对优化有一初步的但较全面的了解，熟悉优化的基本概念与重要原理，掌握主要的寻优迭代方法，理解如何将实际问题抽象成优化模型，以便能在今后工作中具体应用。

作为教材，书中比较注意概念与原理的阐述，但不作过多的数学推导，以讲清意义与相互关系为主；对于优化算法，除介绍迭代步骤外，大多列有详细算例；各章还尽可能附有适量的习题，便于读者自学与巩固。

全书共8章。第一、二章是基本概念、术语与原理，为了叙述紧凑，有关最优性条件分别在相应章内叙述；第三章是无约束优化问题的各种解法，包括一维搜索方法；第四章是线性规划问题，包括整数线性规划；第五章是一般的非线性约束问题的解法；第六章是工程结构优化设计中常用的准则法；第七章是其他几种类型的优化问题，如动态规划、几何规划与多目标优化等；最后一章讲述优化数学模型的塑造原则、算法的选择以及若干应用示例。

本书由汪树玉主编。第一、二、四、五章，第七章第一、二节以及第八章第二节由汪树玉编写；第三章和第七章第三节由杨德铨编写；第六章和第八章除第二节外由杜本钊编写。全书由杨仲侯主审。

限于作者水平，书中难免有错误与缺点，恳请读者批评指正。

编　者

1992.7.

符 号 说 明

min.　求极小

max.　求极大

s.t.　受约束于

$\boldsymbol{x}=[x_1, x_2, \cdots, x_n]^T$　表示由n个分量$x_1, \cdots, x_n$组成的n维向量，或是坐标值为$x_1, \cdots, x_n$的n维空间点

$\boldsymbol{x}^*$　最优点或最优解

f^*　目标函数$f(\boldsymbol{x})$的最优值

$g_j(\boldsymbol{x})\geqslant 0$　第j个不等式约束条件

$h_r(\boldsymbol{x})=0$　第r个等式约束条件

$A\subset B$　集合A被包含于集合B中

$A\supset B$　集合A包含集合B

$x\in A$　x属于集合A，或x是集合A的元素

$x\notin A$　x不属于集合A

$A\cup B$　集合A与集合B的和集（或并集）

$A\cap B$　集合A与集合B的交集

ϕ　空集

R　实数域

E^n　n维欧氏空间

$N_\varepsilon(\boldsymbol{x}^*)$　$\boldsymbol{x}^*$的ε邻域

$\triangleq$　定义为（或记为）

(a, b)　ab开区间

$[a, b]$　ab闭区间

$f\in C^p$　函数f具有p阶连续导数

$f: S\to R$　f是定义在集S上的实值函数

$\min\limits_{x\in S\subset E^n} f(\boldsymbol{x})$　在E^n空间的子集S上求函数$f(\boldsymbol{x})$的极小

$\min\limits_{x} f(x, y)$　函数$f(x, y)$对x求极小

$\max\limits_{y} f(x, y)$　函数$f(x, y)$对y求极大

0　零向量（各分量值均为零）

b, c　向量b与向量c

A　矩阵

$\{\boldsymbol{u}^j\}^T$　向量$\boldsymbol{u}^j$转置

$[\mathbf{G}^k]^{-1}$　矩阵$\mathbf{G}^k$的逆阵

$\mathbf{x}^k$　第k次迭代点（向量）

$\mathbf{d}^k$　第k次搜索方向

e　第i个分量为1，其余分量为零的单位向量

$\{\mathbf{X}^k\}_{1:m}$　m个迭代点列组成的集合

$\{\mathbf{d}^k\}_{1:m}$　m个搜索方向组合的向量集

$\{M_k\}$　数列$M_1, M_2, \cdots, M_k, \cdots$

$\|\boldsymbol{x}\|$　点$\boldsymbol{x}$的欧氏范数，

$$\|\boldsymbol{x}\|=\left(\sum_{i=1}^n x_i^2\right)^{1/2}$$

$\boldsymbol{x}^T\boldsymbol{y}$　向量（或点）$\boldsymbol{x}$与$\boldsymbol{y}$的内积

$$\boldsymbol{x}^T\boldsymbol{y}=\sum_{i=1}^n x_i y_i$$

$|\boldsymbol{A}|$或$\det(\boldsymbol{A})$　矩阵$\boldsymbol{A}$的行列式

$r(\boldsymbol{A})$　矩阵$\boldsymbol{A}$的秩

$\nabla f(\boldsymbol{x})$　$f(\boldsymbol{x})$的梯度向量

$H(\boldsymbol{x})$或$\nabla^2 f(\boldsymbol{x})$　$f(\boldsymbol{x})$的海森（Hesse）矩阵，即二阶导数矩阵

$\nabla_x L(x, \nu)$　函数$L(x, \nu)$关于x的梯度

$\nabla_x^2 L(x, \nu)$　函数$L(x, \nu)$关于x的二阶导数矩阵

I　单位矩阵

$I(\boldsymbol{x}^*)$　在点$\boldsymbol{x}^*$处的作用约束指标集

$\underline{x_i}$, $\overline{x_i}$　设计变量x_i取值的下界与上界

目　录

第一章 概 述

人们在社会实践与生产活动中，遇到的问题往往有多种可能的处理办法、方案或解答，这就需要寻求最合适的或最优的一种方案或解答。例如购置物品，满足所要求功能的品种在市场上可能有许多，采购者面临的任务就是如何来选择质量可靠、寿命较长而价格较低的一种。又如工厂生产中常需要解决在给定的资源(人力、物力、财力)条件下，如何安排生产以达到成本最低或利润最大。制作某种产品，如何合理选择参数使产品性能最优。设计一个工程，如何在满足各种功能要求和保证工程安全可靠的条件下，使投资或造价最低，等等。总之，如果一个问题有多种可行的处理方案或解答，应从所有可行的解决方案或解答中选择最合适的一个，使预先规定的目标达到最优，这称为最优化问题，简称优化。这里所谓可行是指满足问题所要求的各种规定和条件。所获得的具有最优目标的那个方案称为最优方案，数学术语是最优解。搜索或寻求最优解的方法称优化方法。对于简单的优化问题，可根据经验，加上必要的调查与计算，凭分析与判断即可解决。但在社会经济、生产与管理领域中，多数问题比较复杂且涉及因素多，如运输中的最优调度、大型工程的最优规划与设计，以及国土的最优开发方案等等，它们或者由于问题重要和影响深远，一旦处理方案选择不当，后果十分严重；或者由于问题复杂，光凭经验判断不能满意解决，甚至无法解决。对于这类问题必须采用基于坚实理论的较可靠的方法，和相应的计算手段与工具，才能完成。优化的理论与方法，正是为适应这种社会需要而产生，并随着电子计算机的迅速发展而发展，目前已广泛应用于各个领域。

第一节 优化问题的数学模型

例 1-1 设某砖厂生产两种铺地砖，每种使用的材料（砂、小石、水泥）量相等，但其中一种是彩色的，应添加颜料，且工艺稍复杂，当然利润高些。假设每吨有色砖需耗费 2 个机时、3 个人工工时，2 升颜料，创利润300元；无色砖相应的为 1 个机时、3 个工时，不用颜料，创利润200元。工厂能利用的资源每日为10机时、24个工时和 8 升颜料。现在的问题是如何安排两种砖的生产量，使工厂利润最高。

设有色砖日产为x_1吨，无色砖日产为x_2吨，此时每天的利润为（$300x_1+200x_2$），需占用的机时为（$2x_1+x_2$），其值不应超过给定的10个机时，即$2x_1+x_2\leqslant 10$；类似地，工时条件为 $3x_1+3x_2\leqslant 24$；颜料用量的限制条件为 $2x_1\leqslant 8$。当然，生产的每种砖数不可能为负，即$x_1\geqslant 0$，$x_2\geqslant 0$。故上述问题可归纳为：

寻求x_1、x_2，在满足条件

$$\begin{cases} 2x_1 + x_2 \leqslant 10 \\ 3x_1 + 3x_2 \leqslant 24 \\ 2x_1 \qquad \leqslant 8 \\ x_1 \geqslant 0 \quad x_2 \geqslant 0 \end{cases}$$

下，使（$300x_1 + 200x_2$）趋极大。

例 1-2 考虑图1-1中两杆对称的平面桁架，顶点处受有竖向荷载$2P$，两杆支座水平距离为$2B$。设已选定桁架为空心圆杆，壁厚为t，所用材料的弹性模量为E、比重ρ、容许应力$[\sigma]$，要求在满足强度与稳定条件下，使桁架重量最轻。

问题中待选的参数仅是桁架高度H与圆杆平均直径D。两桁杆总长度为$2\sqrt{H^2+B^2}$，故桁架重量为：

$$W = 2\rho\pi Dt\sqrt{(B^2+H^2)}$$

桁杆中承受的轴压力为$P\sqrt{(H^2+B^2)}/H$，相应的应力为$P\sqrt{H^2+B^2}/(\pi DtH)$，其值应小于容许应力。同时为了避免压屈失稳，还应不超过压杆稳定的临界应力：

$$[\sigma_c] = \frac{\pi^2 E(D^2+t^2)}{8(B^2+H^2)}$$

当然，H不可能为负，D不可能小于t。故问题可归纳成：寻求变量H与D，在满足条件下，使W趋极小。

$$\begin{cases} \dfrac{P\sqrt{H^2+B^2}}{\pi DtH} \leqslant [\sigma] & \text{（强度约束条件）} \\ \dfrac{P\sqrt{H^2+B^2}}{\pi DtH} \leqslant \dfrac{\pi^2 E(D^2+t^2)}{8(B^2+H^2)} & \text{（稳定约束条件）} \\ H \geqslant 0,\ D \geqslant t & \text{（界限约束条件）} \end{cases}$$

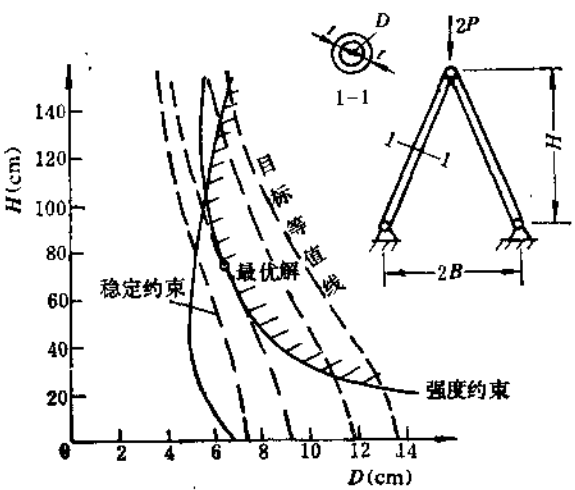

图 1-1　例1-2的计算分析示意图

上面的例子都是在探讨这样一类问题：怎样来选择一组可变参数，在满足规定的条件下，使问题的某种指标或目标最优（极大或极小）。上述待求的可变参数称为设计变量，衡量问题解决得好坏或设计的优劣指标称为目标函数，是设计变量的函数。规定要满足的条件称约束条件，也是设计变量的函数。设计变量、目标函数与约束条件是构造优化问题模型的三要素。下面再作进一步说明。

一、设计变量

这是指在优化过程中所要选择的基本参数，它们对问题的解或设计方案的优劣有重要影响，优化的目的就是要寻找这些参数值的最优组合。例如，生产安排中各种产品的产量

可以选作设计变量，以保证工厂获利最大。结构设计中，各构件的截面尺寸可作设计变量，使结构用料最省，如果要寻求形状最佳，那么外形尺寸或结点坐标可取为设计变量等等。设计变量的数目，称为优化问题的维数，例如有n个设计变量，那么该优化问题称为n维的优化问题。设计变量数目越多，问题的规模就越大，求解就越困难，出现"维数灾难"。因此不宜将优化问题中所有的待定参数都当作要优化的设计变量，而可仅取能反映问题的主要矛盾、主要因素的参数作设计变量，而其它参数预先给定，作为常数不参与寻优，以减小问题的规模。另外选择哪些参数作为设计变量还要视问题性质、结合目前的计算分析能力与水平而定，有的参数虽然对设计方案的好坏影响很大，例如结构类型或所用材料种类，但它们涉及范围过于宽广，如作为变量寻优，有时优化难度与工作量过于庞大，以致无法进行，只能当作固定参数，改选低一层次的参数如构件截面、结点坐标、外形尺寸等作为设计变量。详细讨论见第八章。设计变量常记为$\boldsymbol{x}$，如果有多个变量，可用向量表示，即$\boldsymbol{x}=[x_1, x_2, \cdots, x_i\cdots, x_n]^T=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$，其中下标$i$表示分量序号，如$x_2$代表第2个设计变分量，$T$为转置符号。本书采用黑体字母来代表向量或矩阵，当有一组向量时，则用上标记录它在向量组中的序号，如$\boldsymbol{x}^k$表示第k次迭代的设计变量，而x_i^k为$\boldsymbol{x}^k$中的第i个分量。

二、目标函数

这用于衡量一个设计或一种解决问题方案优劣程度的标准或指标，亦称评价函数，是设计变量的函数，常记为$f(\boldsymbol{x})$。选用什么样的指标作目标函数取决于问题的性质与要求。如上例安排工厂生产计划时常以利润或利润率最大为目标，亦可以成本或投资额最低为目标。从数学角度看，求函数极小[记为$\min.f(\boldsymbol{x})$]与求函数极大[$\max.f(\boldsymbol{x})$]是类同的。函数的极大点就是相应负函数的极小点，同时负函数极小值再取负号就是原函数的极大值，即$\max.f(\boldsymbol{x})=-[\min.(-f(\boldsymbol{x}))]$。故不失一般性，下文如不特别指明均针对求函数的极小值问题。

三、约束条件

这指对所考虑问题施加的种种限制，每一个问题的解答或设计方案均应满足这些规定的要求，以保证所得的解或方案是合理的、实际可行的。从形式上看，约束条件可能是等式方程（设有l个），常记为$h_r(\boldsymbol{x})=0$，$r=1, 2, \cdots, l$；也可能为不等式方程（设有m个），记为$g_j(\boldsymbol{x})\geqslant 0$或$g_j(\boldsymbol{x})\leqslant 0$，$j=1, 2, \cdots, m$。其中$h_r(\boldsymbol{x})$和$g_j(\boldsymbol{x})$是设计变量的函数，称约束函数。由于可以用乘以（$-1$）来改变不等式约束中不等号的方向，如$g_j(\boldsymbol{x})\leqslant 0$等价于$-g_j(\boldsymbol{x})\geqslant 0$，而不改变约束函数$g_j(\boldsymbol{x})$的形式与属性，为便于统一处理，本书一般均采用$g_j(\boldsymbol{x})\geqslant 0$的形式。这是从形式上分，如按约束的性质看，一类是界限约束，它直接规定设计变量可能取值的上下限；另一类是性态约束，是对系统功能或外界因素作用后产生的响应所规定的限值，如例1-2中的圆杆强度与稳定约束。工程结构设计中常见的应

力约束、位移约束以及对频率或稳定的要求均属于这一类，它们往往比较复杂，甚至仅是设计变量的隐函数，无法写出显表达式。

归纳起来，一个优化问题可抽象为如下的数学形式，称为优化的数学模型：

$$\left.\begin{array}{ll}\text{寻求} & \boldsymbol{x}=[x_1,\ x_2,\ \cdots,\ x_n]^T \\ \text{min.（或max.）} & f(\boldsymbol{x}) \\ \text{s.t.} & g_j(\boldsymbol{x})\geqslant 0\quad j=1,\ 2,\ \cdots,\ m \\ & h_r(\boldsymbol{x})=0\quad r=1,\ 2,\ \cdots,\ l\end{array}\right\}\tag{1.1.1}$$

式中，min.（或max.）表示求极小（或极大），s.t.是subject to的缩写，表示“约束为”或“受约束于”。

一个优化问题可能只有等式约束条件，亦可能只有不等式约束条件或两者兼有。由于等式约束$h_r(\boldsymbol{x})=0$可用两个不等式约束条件

$$h_r(\boldsymbol{x})\geqslant 0\qquad\qquad h_r(\boldsymbol{x})\leqslant 0$$

等价表示，故有时为讨论方便亦可考虑仅具有不等式约束的情况，即寻求$\boldsymbol{x}=[x_1,\ \cdots,\ x_n]^T$

$$\left.\begin{array}{l}\text{min.}\quad f(\boldsymbol{x}) \\ \text{s.t.}\quad g_j(\boldsymbol{x})\geqslant 0\quad j=1,\ 2,\ \cdots,\ m\end{array}\right\}\tag{1.1.2}$$

如采用向量函数，记

$$\boldsymbol{g}(\boldsymbol{x})=[g_1(\boldsymbol{x}),\ g_2(\boldsymbol{x}),\ \cdots,\ g_m(\boldsymbol{x})]^T$$
$$\boldsymbol{h}(\boldsymbol{x})=[h_1(\boldsymbol{x}),\ h_2(\boldsymbol{x}),\ \cdots,\ h_l(\boldsymbol{x})]^T$$

来表示约束条件，则问题可写作如下形式：

$$\left.\begin{array}{l}\text{min.}\quad f(\boldsymbol{x}) \\ \text{s.t.}\quad \boldsymbol{g}(\boldsymbol{x})\geqslant \boldsymbol{0} \\ \qquad\ \ \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0}\end{array}\right\}\tag{1.1.3}$$

式中，黑体0代表由若干个数字零组成的向量。

第二节　设计空间、可行域与目标等值面

以各个设计变量为坐标轴张成的空间称设计空间。对于仅有2个设计变量的问题（二维），每一设计方案$\boldsymbol{x}=[x_1,\ x_2]^T$可看作是设计平面上一个点；三个变量问题，每一个设计方案代表三维空间中一个点；推而广之，n个变量$\boldsymbol{x}=[x_1,\ x_2,\ \cdots,\ x_n]^T$或称有$n$个分量的设计方案，代表$n$维设计空间中的一个点。在平面上每一个等式条件，$h_r(x_1,\ x_2)=0$，代表一条曲线，而三维空间内，$h_r(x_1,\ x_2,\ x_3)=0$，代表一个曲面。同样，推广到$n$维设计空间，等式约束$h_r(\boldsymbol{x})=0$代表一个超曲面，位于该超曲面上的设计点才算满足该约束条件。不等式约束$g_j(\boldsymbol{x})\geqslant 0$代表由$g_j(\boldsymbol{x})=0$的超曲面（又称约束界面）所分割出的一个半空间，只有位于该半空间内的设计点才满足$g_j(\boldsymbol{x})\geqslant 0$的要求。当问题有很多约束

时，这些等式约束构成的超曲面和不等式约束构成的半空间将在整个设计空间内分割出一个区域，落在该区域内的设计点满足全部约束条件，是一可行点，故该区域称可行域。以两个设计变量 x_1、x_2的优化问题为例，图 1-2(a)中的ABC区域（斜线部分）代表由 3 个不等式约束$g_1(\boldsymbol{x})\geqslant 0$、$g_2(\boldsymbol{x})\geqslant 0$和$g_3(\boldsymbol{x})\geqslant 0$所组成的可行域，图1-2($b$)中的$DF$曲线（粗黑线）则是由等式约束$h_1(\boldsymbol{x})=0$和两个不等式约束$g_2(\boldsymbol{x})\geqslant 0$ $g_3(\boldsymbol{x})\geqslant 0$组成的可行域。同样，目标函数等于某一定值（如 $f(\boldsymbol{x})=c_i$）的点的集合，将在n维设计空间中形成一条曲面，称为函数值为c_i的等值线。c_i取不同的数值可得到该函数的一族等值线，如图1-2 中的虚线所示。

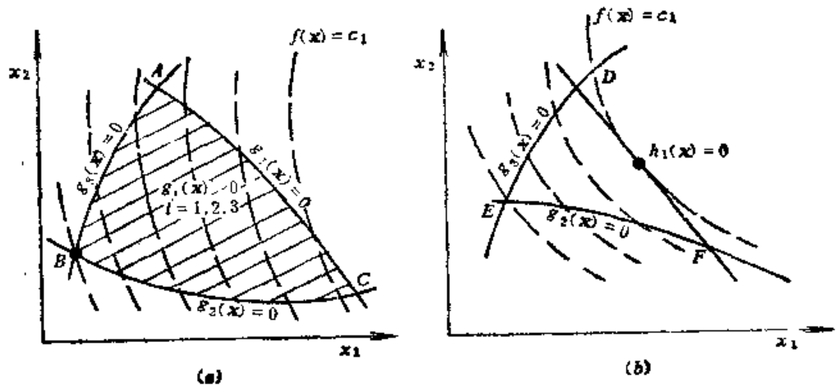

图 1-2 二维问题的可行域

利用上述概念，对于n 维优化问题（以求极小为例）的几何描述是："在n维设计空间中由等式约束超曲面以及不等式约束半空间所构成的可行域内，寻求位于最小目标等值面上的点"。这个点就是问题的最优解，常记为$\boldsymbol{x}^*$。

如用集合论符号来表示可行域S，则

$$S=\{\boldsymbol{x}\mid \boldsymbol{x}\in E^n,\ g_j(\boldsymbol{x})\geqslant 0\quad j=1,\ 2,\ \cdots,\ m$$
$$h_r(\boldsymbol{x})=0\quad r=1,\ 2,\ \cdots,\ l\} \tag{1.2.1}$$

式中，E^n 表示n维设计空间（每一分量为实数，且有距离与内积的定义，称欧氏空间）；$\in$表示属于。公式的意义是：可行域S是所有可行点$\boldsymbol{x}$的集合，其中每一点属于 n 维设计空间（即有n个分量），且该点满足所有的等式与不等式约束条件。S亦称为可行解集合。这样，优化问题可用下式更加简洁地表达为：

$$\mathbf{x}^*:\ f(\boldsymbol{x}^*)=\min_{x\in S}.\ f(\boldsymbol{x}) \tag{1.2.2}$$

含义是寻求点$\mathbf{x}^*$，它的目标值是所有可行点的目标值中的极小值。或简单地表达为

$$\min_{x\in S}.f(\boldsymbol{x}) \tag{1.2.3}$$

即在集合S上求$f(\mathbf{x})$的极小值问题。对于无约束优化问题，整个设计空间都是可行域，故可表达为

$$\min_{x \in E^n} f(x) \tag{1.2.4}$$

以上两种描述方法，式（1.1.1）和式（1.2.2）或（1.2.3）含义完全一样，都概括了优化问题的数学实质。

第三节　问题的分类与求解方式

为了便于分析与研究，常将优化问题分类，可以有多种分类方法。

（1）按目标函数的个数分：如果问题中寻优的目标函数有多个称多目标优化问题，否则称单目标优化问题。

（2）按约束条件的有无分：有约束优化问题与无约束优化问题。

（3）按式（1.1.1）中所含的函数类型分：当目标与约束均为设计变量的线性函数时称为线性规划。这是优化中发展最早最成熟的一支，有很完善的算法，且可获得全局最优解。如果约束仍为线性函数，而目标是设计变量的二次函数，则称二次规划。当目标与约束至少有一个为非线性函数时，统称非线性规划。显然二次规划是非线性规划的特例。如果目标与约束是设计变量幂的乘积函数组成的多项式，由于这种问题的非线性程度很高，需采用特殊的求解技术，形成所谓几何规划。

（4）按变量的性质分：一般优化问题所考虑的变量是连续型的实数，如果某些问题规定变量只能取离散值或整数时，则称离散规划或整数规划。如果限制变量只能取0或1时，又称0-1规划。近年来发展很快的"组合优化"亦是一种离散规划。如果变量本身不是单一的参数值，而是某些参数（如空间坐标）的连续函数，则称为分布参数优化问题，或轨迹优化问题，它在形状优化、最优控制问题中应用甚广。

线性与非线性规划所处理的问题，通常都是静态的，与时间因素无关。适于处理含时间序列的优化问题有动态规划，它将问题分为若干阶段，利用递推关系逐步优化，最后达到整个过程的优化。

此外，一般问题中的设计变量和有关的参数都假定为确定性的，如果它们中有非确定性的数（随机的或模糊的），则成为随机规划与模糊优化问题。

优化问题的求解方式，主要有如下几种。

（1）解析法：利用微分学与变分学等经典方法直接找出目标函数极值与极值点的一类方法。对于有约束问题可利用拉格朗日乘子法或约束变分法来求解。此法适用于函数有良好的性质（如存在二阶导数），且规模不大的问题。对于大型复杂问题，可以作为辅助手段，用于对问题的前期简化或后期分析方面。

（2）图解法：直接绘出目标函数与约束函数的图形曲线，通过图的分析来获得最优解点的方法。此法的优点是直观全面，但局限于仅有1～2个设计变量的问题。不过对于高维问题亦可固定其它变量，借助图解法研究一两个主要变量对问题与最优解的影响。

（3）数值法：利用电子计算机通过迭代算法生成一系列解点来逼近最优解（点）的一类方法。以下讨论的优化方法，大多属于这类。迭代算法除可根据数学规划理论来制定

外，还可按工程力学概念来建立，如结构优化中的最优性准则法。

（4）试验法：它是直接对实际变化过程进行试验，或通过变更设计变量数值，模拟过程变化规律，以寻求问题的最优解的方法。从后一方面看，它也属于数值法，此时一般只需求算函数值而不需用导数信息，故又称为直接法。这类方法简单易用，但收敛一般较慢。

应该指出，优化获得的解答，从数学上说，应是最优的解答。但由于在构造模型时往往略去一些次要因素和不可量化的条件，它只是实际问题的一种近似，因此所得解只是在给定条件下相对的"优"，而不能看作问题的绝对的"最优解答"，一般还应根据原实际问题作必要调整与修改，才能获得符合实际的满意解答。采用"优化"一词就是说明这是一个寻优的过程，它比"最优"要确切些。优化技术的应用包括三部分工作：一是将实际问题抽象为优化数学模型；二是求解所得的数学问题；三是对结果进行分析加工，如果需要的话，进一步完善模型重新求解，以使结果更切合实际，并付之实施与应用。目前在数学规划论和计算方法两方面都已取得了很大进展，已有一批较好的优化算法与软件可供使用。作为工程技术和管理人员虽然其主要工作在于塑造优化模型与分析优化成果，但选择合适的有效算法与软件也十分重要，因为每一个算法都有它的适应范围与限度，因而，需对优化理论与方法有较深的了解。此外应对所要处理的问题作透彻的剖析，塑造出合理的、性态较好的模型，进而挑选合适的算法与相应软件，再在运算中根据情况作出调整，只有如此才可望得到满意的优化效果。

习　　题

1-1　用图作出以下各组约束所表示的可行域：

（1）$0.5x_1+x_2\leqslant 4$
$3x_1+x_2\leqslant 15$
$x_1+x_2\geqslant 1$
$x_1\geqslant 0,\ x_2\geqslant 0$

（2）$-x_1^2+x_2-1\geqslant 0$
$x_1-x_2+2\geqslant 0$
$x_1\geqslant 0,\ x_2\geqslant 0$

（3）$x_1-x_2\geqslant 0$
$x_1+x_2\leqslant 4$
$x_1\leqslant 3,\ x_2\geqslant 0$

（4）$x_1^2+(x_2-1)^2-1\leqslant 0$
$(x_1-1)^2+x_2^2-1\leqslant 0$
$x_1\geqslant 0,\ x_2\geqslant 0$

（5）$x_1^2+x_2^2-1\leqslant 0$
$x_1+x_2-0.5\leqslant 0$
$x_1\geqslant 0,\ x_2\geqslant 0$

（6）$(x_1-1)^3-x_2^2=0$
$x_1\leqslant 2$

1-2　建造一容积为V（m^3）的长方形蓄水池（无盖），要求选择其长、宽、高，使表面积最小，从而建筑用料最省。试写出此问题的数学模型。

1-3　某公司有资金a万元，可供选择购置的设备有n个，已知相应于第i种设备所需资金为b_i万元，可得收益为c_i万元，要求收益最大的投资安排。试写出其数学模型。

1-4　某城市要建造一供应服务中心，向该市m个用户提供服务，设第i个用户的位

置为（a_i, b_i），需要货物量为w_i吨，寻求这个中心最经济的位置使运输量（吨公里数）最小。

1-5 有两动点A与B，A点坐标(a_1, b_1, c_1)应在曲面$f_1(x_1, x_2, x_3)=0$上移动，$B(a_2, b_2, c_2)$应在曲面$f_2(x_1, x_2, x_3)=0$上移动，问A、B两点分别在什么位置时，它们之间的距离为最小。试构造出它的数学模型。

第二章 基 本 知 识

第一节 函数及可微函数的性质

由于后面经常用到函数、可微函数和它们的若干性质，如梯度、海赛矩阵和泰勒展开式等，这些都是高等数学中有关概念的延拓，为便于应用与统一符号，此处作简要的归纳。

一、函数

如果存在一种规则f，使E^n空间的子集S上每一点$\boldsymbol{x}$，有相应的实数$f(\boldsymbol{x})\in R$与之对应，则称f是定义在S上的实值函数，记为：

$$f: S\rightarrow R$$

如果存在规则$\boldsymbol{F}$，使S上的每一点$\boldsymbol{x}$，有m个实数$f_1(x)$、$f_2(x)$、…、$f_m(x)$与之对应，则称$F=[f_1, f_2, \cdots, f_m]^T$是定义在$S$上的向量值函数，记为：

$$\boldsymbol{F}: S\rightarrow E^m$$

而f_j为第j个分量$f_j(\boldsymbol{x})$的简写。

函数在定义域S上可以是连续的、分段连续的（在S的某些点上不连续），或者是离散的。连续函数又可能是不可微的（不存在导数），也可能是一阶可微的（存在一阶导数）或二阶可微的（存在二阶导数）。可微性反映函数所描述的曲面的光滑程度。

如果函数f在$\boldsymbol{x}^0$处一阶可微，那么f在该点处关于$\boldsymbol{x}$各分量的偏导数存在，记为$\partial f/\partial x_i$，而

$$\frac{\partial f}{\partial x_i}=\lim_{t_i\to 0^+}\frac{f(\boldsymbol{x}^0+t_i\boldsymbol{e}^i)-f(\boldsymbol{x}^0)}{t_i} \tag{2.1.1}$$

式中，$\boldsymbol{e}^i=[0\cdots0\overset{i}{1}\ 0\cdots0]^T$，仅第$i$个分量为1，其余分量为0的向量，又称单位向量，$t_i\in R$。

二、梯度与方向导数

1.梯度

可微函数f在点$\boldsymbol{x}\in S$处的梯度定义为：

$$\text{grad}\ f(\boldsymbol{x})=\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}\right]^T \tag{2.1.2}$$

常记为$\nabla f(\boldsymbol{x})$。以下为几个常见函数的梯度。

（1）
$$f(\boldsymbol{x})=\boldsymbol{b}^T\boldsymbol{x}=\sum_{i=1}^{n}b_ix_i$$

由于
$$\frac{\partial f}{\partial x_i}=b_i$$

故 $$\nabla f(\boldsymbol{x})=\left[\frac{\partial f}{\partial x_1}\cdots\frac{\partial f}{\partial x_i}\cdots\frac{\partial f}{\partial x_n}\right]^T=[b_1\cdots b_i\cdots b_n]^T=\mathbf{b}$$

（2）
$$f(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{x}=\sum_{j=1}^{n}x_i^2$$

类似（1）有 $$\nabla f(\boldsymbol{x})=[2x_1,\ 2x_2,\ \cdots,\ 2x_n]^T=2\boldsymbol{x}$$

（3）
$$f(\boldsymbol{x})=a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n+a_{21}x_1x_2+a_{22}x_2^2+\cdots+a_{2n}x_nx_2$$
$$+\cdots a_{nn}x_n^2=\sum_{i,\ j=1}^{n}a_{ij}x_ix_j=\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$$

式中，$\boldsymbol{Q}$为$n\times n$矩阵：

$$\boldsymbol{Q}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ & \cdots\cdots\cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}$$

可证，$\nabla f(\boldsymbol{x})=2\boldsymbol{Q}\boldsymbol{x}$。

2.方向导数

设f是定义在E^n空间中子集S上的连续函数，从点$\boldsymbol{x}^0\in S$出发，对某一向量$\boldsymbol{d}\in E^n$及足够小的实数$t>0$，有$\boldsymbol{x}^0+t\boldsymbol{d}\in S$，如果极限

$$\lim_{t\to 0}\frac{f(\boldsymbol{x}^0+t\boldsymbol{d})-f(\boldsymbol{x}^0)}{t\|\boldsymbol{d}\|}$$

存在，则称该极限为f在$\boldsymbol{x}^0$处沿$\boldsymbol{d}$的方向导数，记为$f'(\boldsymbol{x}^0,\ \boldsymbol{d})$。方向导数代表函数沿该方向上的变化率。

可微函数的方向导数$f'(\boldsymbol{x}^0,\ \boldsymbol{d})$等于函数在$\boldsymbol{x}^0$处的梯度在$\boldsymbol{d}$上的投影

$$f'(\boldsymbol{x}^0,\ \boldsymbol{d})=\nabla f(\boldsymbol{x}^0)^T\boldsymbol{d}/\|\boldsymbol{d}\|=\sum_{i=1}^{n}\frac{\partial f(\boldsymbol{x}^0)}{\partial x_i}\frac{d_i}{\|\boldsymbol{d}\|}\qquad(2.1.3a)$$

式中，$\|\cdot\|$表向量的模（或范数）。当$\|\boldsymbol{d}\|=1$时上式可改写为

$$f'(\boldsymbol{x}^0,\ \boldsymbol{d})=\nabla f(\boldsymbol{x}^0)^T\boldsymbol{d}=\|\nabla f(\boldsymbol{x}^0)\|\cdot\cos\theta\qquad(2.1.3b)$$

式中，θ是向量$\boldsymbol{d}$与$\nabla f(\boldsymbol{x}^0)$间的夹角。显然：

当$0\leqslant\theta<90°$时，$f'(\boldsymbol{x}^0,\ \boldsymbol{d})>0$，表明当$\boldsymbol{d}$与函数梯度的交角是锐角时，$f(\boldsymbol{x})$在$\boldsymbol{x}^0$处沿方向$\boldsymbol{d}$是上升的。特别在$\boldsymbol{d}$与$\nabla f(\boldsymbol{x}^0)$重合、即$\theta=0$时，$f'(\boldsymbol{x}^0,\ \boldsymbol{d})$值达最大。

当$90°<\theta\leqslant 180°$时，$f'(\boldsymbol{x}^0,\ \boldsymbol{d})<0$，表明当$\boldsymbol{d}$与函数梯度的交角为钝角时，$f(\boldsymbol{x})$在$\boldsymbol{x}^0$处沿方向$\boldsymbol{d}$是下降的，此时称$\boldsymbol{d}$是$f(\boldsymbol{x})$在$\boldsymbol{x}^0$处的下降方向。在$\boldsymbol{d}$与$-\nabla f(\boldsymbol{x}^0)$重合，即$\theta=180°$时，$f'(\boldsymbol{x}^0,\ \boldsymbol{d})$值达最小。

因此函数在一点的梯度方向是该点上函数值增加最快的方向；而负梯度方向是该点处函数值下降最快的方向，称最速下降方向。