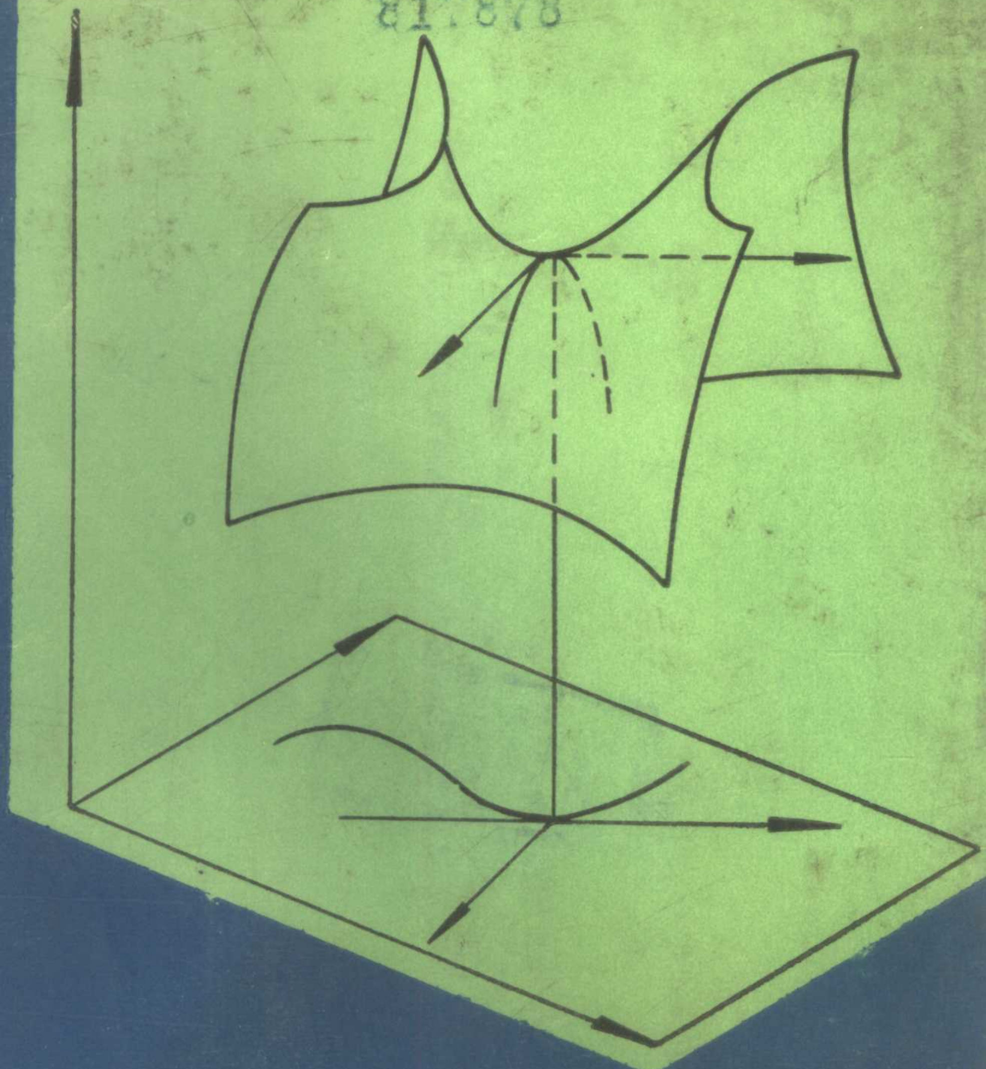

最优化方法

邓乃扬·诸梅芳 著

辽宁教育出版社

最优化方法

邓乃扬 诸梅芳 著

辽宁教育出版社
1987年·沈阳

最优化方法

邓乃扬　诸梅芳　著

辽宁教育出版社出版　辽宁省新华书店发行
（沈阳市南京街6段1里2号）沈阳新华印刷厂印刷

字数：250,000　开本：$850\times1168^{1}/_{32}$　印张：11
印数：1—5,000
1987年2月第1版　1987年2月第1次印刷

责任编辑：俞晓群　责任校对：杨力　李晓晶
封面设计：安今生　插　图：潘智倩

统一书号：7371·172　定价：2.60元

序　言

最优化方法是一个新兴的数学分支，它是选择“最优方案”的一种有力工具，能用于多种工程技术和经济管理部门，并产生直接的经济效益，因而已日益受到人们的重视.

本书着重介绍线性规划和非线性规划中适用面较广、使用方便、具有实效的方法，并力求反映先进成果. 无约束问题的解法包括使用目标函数一阶导数的变度量法和改进的共轭梯度法（第四章），以及只使用目标函数值的 Powell 方法（第五章）；约束问题的解法包括线性规划的有效集法和单纯形法（第七章），以及求解一般约束问题的两种方法——约束问题的变度量法和乘子法（第八章和第九章）.

我们认为在最优化这一领域中有许多概念、方法和理论都有着十分清晰的几何背景. 本书尝试着尽可能使用直观的、具有启发性的图象引进和说明它们，然后再给出必要的理论分析，以便使读者较快地和较深入地理解各种方法的实质.

本书既可用于自学，又可用做教材. 全部讲授本书约需70～80学时. 少学时的课程使用本书时，可以删去书中的部分内容和某些定理的证明. 50～60学时的课程可以讲授下列内容：第一章；第二章；第三章；第四章§1，§2，§3，§4：4.1和4.2；第六章；第七章§1，§2：2.1和2.2的1，§3；第八章（也可考虑删去）；第九章. 30～40学时的课程可讲授：第一章；第二章；第三章；第四章§1，§2，§3，§4：4.1

和4.2；第六章；第九章.

对于主要关心使用最优化方法解决实际问题的读者，可以略去书中大部分定理的证明细节.

在本书出版之际，谨向曾经给予我们帮助的桂湘云、席少霖、白雪林先生和高旅端、史明仁、陈志、薛毅、王本浩、刘军、李伟生同志表示衷心的谢意.

由于水平所限，书中错误或不妥之处在所难免，欢迎批评指正。

作　者

符　号

$f(\boldsymbol{x})$	目标函数
$\boldsymbol{g}(\boldsymbol{x}), \nabla f(\boldsymbol{x})$	目标函数 $f(\boldsymbol{x})$ 的梯度向量，$\boldsymbol{g}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = \left(\frac{\partial}{\partial x_1} f(\boldsymbol{x}), \cdots, \frac{\partial}{\partial x_n} f(\boldsymbol{x})\right)^T$
$G(\boldsymbol{x}), \nabla^2 f(\boldsymbol{x})$	目标函数 $f(\boldsymbol{x})$ 的 Hesse 矩阵，其第 i 行第 j 列的元素为 $\frac{\partial^2}{\partial x_i \partial x_j} f(\boldsymbol{x})$
$c_i(\boldsymbol{x})$	第 i 个约束函数
$\boldsymbol{c}(\boldsymbol{x})$	由约束函数构成的向量函数，例如 $\boldsymbol{c}(\boldsymbol{x}) = (c_1(\boldsymbol{x}), \cdots, c_l(\boldsymbol{x}))^T$
$\nabla c_i(\boldsymbol{x})$	第 i 个约束函数 $c_i(\boldsymbol{x})$ 的梯度向量
$\boldsymbol{x}^*$	无约束问题或约束问题的解
$\boldsymbol{x}^{(k)}$	解 $\boldsymbol{x}^*$ 的第 k 次近似
$f^{(k)}$	$f^{(k)} = f(\boldsymbol{x}^{(k)})$
$\boldsymbol{g}^{(k)}$	$\boldsymbol{g}^{(k)} = \boldsymbol{g}(\boldsymbol{x}^{(k)}) = \nabla f(\boldsymbol{x}^{(k)})$
$G^{(k)}$	$G^{(k)} = G(\boldsymbol{x}^{(k)}) = \nabla^2 f(\boldsymbol{x}^{(k)})$
$\boldsymbol{\lambda}^*$	乘子向量
R^n	n 维欧氏空间
$(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$	以向量 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 为列的矩阵
$\langle \cdot, \cdot \rangle$	向量的数量积
$\lVert \cdot \rVert$	向量的范数或矩阵的范数

$\boldsymbol{a}^T$	向量 $\boldsymbol{a}$ 的转置
A^T	矩阵A的转置
$\mathrm{sign}(x)$	符号函数；当 $x>0$ 时，$\mathrm{sign}(x)=1$；当 $x<0$时，$\mathrm{sign}(x)=-1$；$\mathrm{sign}(0)=0$
$\min\{f(\boldsymbol{x})\mid \boldsymbol{x}\in D\}$	函数$f(\boldsymbol{x})$在集合D上的最小值
$A\backslash B$	集合A和集合B的差集，即属于集合 A 但不属于集合 B 的所有元素组成的集合
$A\cup B$	集合A和集合B的并集，即集合 A 与集合B中所有元素组成的集合
$A\cap B$	集合A和集合B的交集，即集合 A 与集合B中所有公共元素组成的集合
$A-\{q\}$	从集合A中删除元素 q 后得到的集合
$A+\{p\}$	对集合 A 添加元素 p 后得到的集合
$\det Q$	矩阵Q的行列式

目　录

第一章 概　　论

本章从具体实例出发，介绍最优化问题的数学形式及其几何解释。

§1 最优化问题实例

什么是最优化？用数学语言来说，就是找出一个多变量函数的极小点，我们通过三个实例予以说明.

例1.1.1 曲线拟合问题

设有两个物理量ξ和η，据根某一物理定律知道它们满足下列函数关系：

$$\eta = a + b\xi^c \tag{1.1.1}$$

其中a, b, c是三个常数，在不同情况下它们取不同的数值. 假定现在由实验测得了(ξ, η)的m组数据

$$(\xi_1, \eta_1), (\xi_2, \eta_2), \cdots, (\xi_m, \eta_m) \tag{1.1.2}$$

试选择a, b, c的值，使得曲线$\eta = a + b\xi^c$尽可能靠近所有的实验点$(\xi_i, \eta_i), i = 1, 2, \cdots, m$（参看图１.１.１）.

这个问题可用最小二乘原理求解，即选择a, b, c的一组值，使偏差的平方和

$$\delta(a, b, c) = \sum_{i=1}^{m}(a + b\xi_i{}^c - \eta_i)^2 \tag{1.1.3}$$

取最小值．换句话说，欲求出三个变量的函数$\delta(a,b,c)$的极小点，以此作为问题的解．

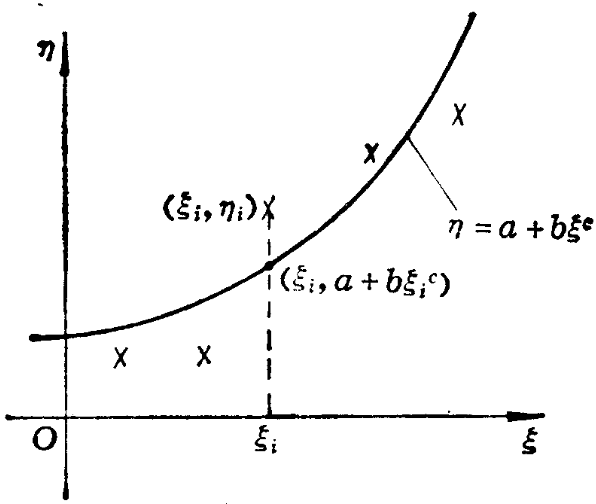

图1.1.1

例1.1.2 生产安排问题

某工厂生产D、E两种产品，每种产品均经三道工序加工而成．假定每生产1米³D种产品需用A种机器加工 7 小时，用B种机器加工 3 小时，用C种机器加工 4 小时．而每生产1米³E种产品需用A种机器加工2.8小时，用B种机器加工 9 小时，用C种机器加工 4 小时． 又已知每生产 1 米³D种产品可盈利500元，每生产 1 米³E种产品可盈利800 元．现设一个月中A种机器工作时间不能超过560小时，B种机器不能超过450小时，C种机器不能超过336小时．问每月D、E两种产品各生产多少可盈利最多？

设每月生产D种产品x_1米³，生产E种产品x_2米³．这两种产品共盈利$500x_1+800x_2$元，我们要选择x_1和x_2，使盈利尽可能地大．然而由于受到机器加工时间的限制，x_1和x_2都不能太大．事实上，这两种产品共需用A种机器加工$7x_1+2.8x_2$小

时，它不能超过560小时；需用B种机器加工$3x_1+9x_2$小时，它不能超过450小时；需用C种机器加工$4x_1+4x_2$小时，它不能超过336小时。总之，我们应该在约束条件

$$7x_1+2.8x_2\leqslant 560$$

$$3x_1+9x_2\leqslant 450$$

$$4x_1+4x_2\leqslant 336$$

$$x_1\geqslant 0,\quad x_2\geqslant 0$$

下，寻找使函数$500x_1+800x_2$取最大值的x_1^*和x_2^*。

例1.1.3　人字架最优设计问题

考虑如图1.1.2所示的钢管构成的人字架，设钢管壁厚$t=\bar{t}$和半跨度$s=\bar{s}$已经给定，试求能承受负荷$2P$的最轻设计。

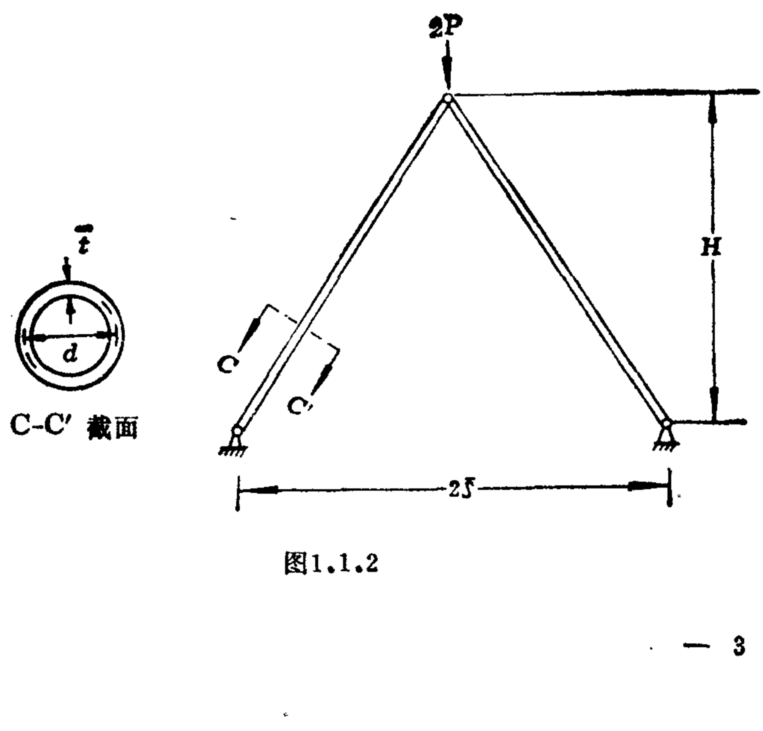

图1.1.2

在这个问题中，需要选择的参数有两个：钢管的平均直径d（d等于内径D_1与外径D_2的平均值）和人字架的高H。我们先看看用定性分析的方法能否解决它。既然希望人字架轻，自然会想到钢管要尽可能短些，但是由于跨度$2\overline{S}$已经取定，这样做将使人字架的张角很大（参看图1.1.3），负荷$2P$就会在钢管上引起很大的分力，进而要求选用较粗的钢管。钢管变粗会增加人字架的重量，所以这未必是最优的方案。由此可见，进行以下的定量分析是必要的。

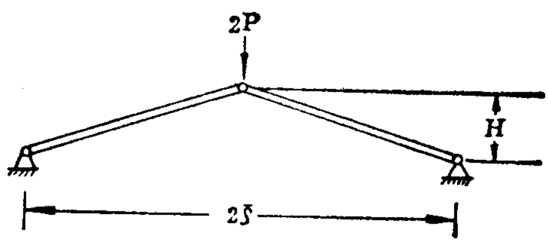

图1.1.3

给定一组d和H的值后，立刻可以算出钢管的横截面积A和单根钢管的长L

$$A=\frac{1}{4}\pi(D_2^2-D_1^2)=\frac{\pi}{4}(D_2+D_1)(D_2-D_1)=\pi d\overline{t}$$

$$L=(\overline{S}^2+H^2)^{1/2}$$

因而人字架的重量为

$$w=w(d,H)=\rho\cdot 2\pi d\overline{t}\,(\overline{S}^2+H^2)^{1/2} \qquad (1.1.4)$$

（其中ρ是钢管材料的比重）。现在要调整d和H使w取最小值。由于整个结构必须能够承受负荷$2P$，所以d和H不能随意取值，它们应该满足如下两个条件：

(i)压应力不能超过材料的许用应力σ_y。容易算出，在负荷$2P$作用下，杆件上的压应力为

$$\sigma(d,H)=\frac{P}{\pi\bar{t}}\cdot\frac{(\bar{S}^2+H^2)^{1/2}}{Hd}$$

故应有

$$\frac{P}{\pi\bar{t}}\cdot\frac{(\bar{S}^2+H^2)^{1/2}}{Hd}\leqslant\sigma_y \tag{1.1.5}$$

这个限制称为屈服约束.

(ii)压应力不能超过杆件的临界应力σ_e. 临界应力σ_e可由Euler公式算得

$$\sigma_e=\sigma_e(d,H)=\frac{\pi^2E}{8}\cdot\frac{d^2+\bar{t}^2}{\bar{S}^2+H^2}$$

（其中E是钢管材料的杨氏模量），故应有

$$\frac{P}{\pi\bar{t}}\cdot\frac{(\bar{S}^2+H^2)^{1/2}}{Hd}\leqslant\frac{\pi^2E}{8}\cdot\frac{d^2+\bar{t}^2}{\bar{S}^2+H^2} \tag{1.1.6}$$

这个限制称为屈曲约束.

总之，我们是要求出d和H的一组值d^*和H^*，使得在满足屈服约束（1.1.5）和屈曲约束（1.1.6）的条件下，重量（1.1.4）取最小值.

§2 最优化问题

2.1 最优化问题的数学形式

前节考虑的三个例题中都包含着若干个需要选择或调整的量，而调整的目的都是使这些量的某个函数达到其最小值或最大值. 这类问题称为最优化问题. 由于函数$f(x)$的最大值完全

与函数 $-f(\boldsymbol{x})$ 的最小值相对应，所以能把最优化问题统一归结为达到最小值的问题．然而上述三个最优化问题之间也有不同之处：第一个例题中要选择的量是不受限制的，而后两个例题中要选择的量只能在某个限定的范围内取值．它们各自代表着一类最优化问题，前者称为无约束问题，后者称为约束问题．下面分别给出它们的数学表达形式．

1．无约束问题

考虑例 1．1．1 所代表的问题．为了导出它们的标准形式，我们把例 1．1．1 改写一下．分别记需要选择的量 a, b, c 为 x_1, x_2, x_3，并用一个3维欧氏空间 R^3 中的向量 $\boldsymbol{x}=(x_1, x_2, x\ )^T$ 表示它们．再把欲求其最小值的函数 $\delta(a,b,c)=\delta(x_1, x_2, x_3)$ 记作 $f(\boldsymbol{x})=f(x_1, x_2, x_3)$．最后把例 1．1．1 所述的问题表示为

$$\min f(\boldsymbol{x}), \qquad \boldsymbol{x}=(x_1, x_2, x_3)^T \in R^3$$

其中

$$f(x)=\sum_{i=1}^{m}(x_1+x_2\xi_i{}^{x^3}-\eta_i)^2$$

该问题的含义是求定义在 3 维欧氏空间 R^3 上的函数 $f(\boldsymbol{x})$ 的最小点．可以想象，如果有 n 个需要选择或调整的量时，我们可以用 n 维欧氏空间的向量 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T$ 表示它们，而把欲求其最小点的函数记为 $f(\boldsymbol{x})$，从而问题可表示为

$$\min f(\boldsymbol{x}), \qquad \boldsymbol{x}\in R^n \tag{1.2.1}$$

这就是对变量没有限制的最优化问题的一般形式，即无约束问题的一般形式．我们称其中的 $f(\boldsymbol{x})$ 为目标函数．因此无约束问题就是寻求一个定义在 n 维欧氏空间 R^n 上的函数 $f(\boldsymbol{x})$ 的最小点．

定义1.2.1 若存在着 $\boldsymbol{x}^*\in R^n$，使得对任意的 $\boldsymbol{x}\in R^n$，都有

$$f(\boldsymbol{x})\geqslant f(\boldsymbol{x}^*) \tag{1.2.2}$$

则称 $\boldsymbol{x}^*$是 $f(\boldsymbol{x})$的最小点，或无约束问题（1.2.1）的整体解.
若对任意的$\boldsymbol{x}\in R^n$，$\boldsymbol{x}\neq\boldsymbol{x}^*$，都有

$$f(\boldsymbol{x})>f(\boldsymbol{x}^*) \tag{1.2.3}$$

则称$\boldsymbol{x}^*$是$f(\boldsymbol{x})$的严格最小点，或无约束问题（1.2.1）的严格整体解.

2．约束问题

考虑例1.1.2．若记$\boldsymbol{x}=(x_1,x_2)^T$，并令

$$f(\boldsymbol{x})=-(500x_1+800x_2)$$

则使$500x_1+800x_2$取最大值等价于使$f(\boldsymbol{x})$取最小值．另外若再令

$$c_1(\boldsymbol{x})=7x_1+2.8x_2-560$$

$$c_2(\boldsymbol{x})=3x_1+9x_2-450$$

$$c_3(\boldsymbol{x})=4x_1+4x_2-336$$

$$c_4(\boldsymbol{x})=-x_1$$

$$c_5(\boldsymbol{x})=-x_2$$

我们可把问题简单地表为

$$\min f(\boldsymbol{x}),\quad \boldsymbol{x}=(x_1,x_2)^T\in R^2$$

$$s.t.\ c_i(\boldsymbol{x})\leqslant 0,\quad i=1,2,3,4,5$$

类似地，若在例1.1.3中分别用 x_1 和 x_2 代替 d 和 H，令$\boldsymbol{x}=(x_1,x_2)^T$，再把函数$w$记作$f$

$$f(\boldsymbol{x})=2\pi\rho\,\overline{t}\,x_1(\overline{S}^2+x_2^2)^{1/2}$$

并令

$$c_1(\boldsymbol{x})=\frac{P}{\pi\overline{t}}\cdot\frac{(\overline{S}^2+x_2^2)^{1/2}}{x_1x_2}-\sigma_y$$

$$c_2(\boldsymbol{x})=\frac{P}{\pi\overline{t}}\cdot\frac{(\overline{S}^2+x_2^2)^{1/2}}{x_1x_2}-\frac{\pi^2E}{8}\ \frac{\overline{t}^2+x_1^2}{\overline{S}^2+x_2^2}$$