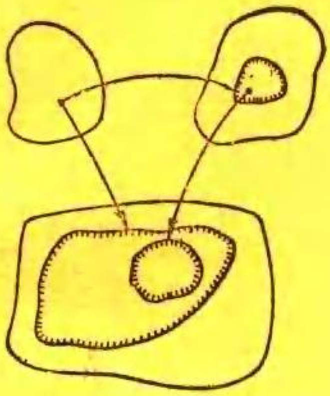
陈开明 编著



大学应用数学丛书

非线性规划

复旦大学出版社

大学应用数学丛书

非 线 性 规 划

陈开明　编著

复 旦 大 学 出 版 社

内 容 简 介

本书系统地阐述了非线性规划的基本理论和方法，主要内容有：最优性条件，算法理论，线性搜索，基本的无约束优化法，共轭梯度法与变尺度法，直接搜索法，乘子法，处理线性约束和非线性约束的优化方法等。本书既重视非线性规划的理论基础，同时也介绍了优化方法的实际应用和近年来在这一领域里的研究成果。本书可作为应用数学类或管理类专业的教材或参考书，也可作为工科类一些专业的高年级或研究生教材。对于从事优化方法研究和应用的工程技术人员，本书也有参考的价值。

非线性规划

陈开明　编著

复旦大学出版社出版

（上海国权路579号）

新华书店上海发行所发行　复旦大学印刷厂印刷

开本850×1168　1/32　印张10　字数297,000

1991年8月第1版　1991年8月第1次印刷

印数 1—3,000

ISBN7-309-00604-6/O·88

定价：3.30元

《大学应用数学丛书》

编审委员会

出版者的话

近年来，随着我国现代化建设事业的发展，许多高等学校创办了一大批重视数学理论和应用的专业和系科，如：应用数学、应用力学、计算数学、控制科学、信息科学、系统科学、运筹学、统计学、计算机科学、应用物理、管理科学等。为了满足这类专业数学教学的需要，我们组织编写和出版了一套“大学应用数学丛书”。本书即为这套丛书中的一本。

“大学应用数学丛书”重视现代数学的基本理论，强调数学的实际应用，反映现代科技的先进成果，并便于课堂教学和自学。我们希望，这套丛书的出版将有助于我国应用数学教学与研究的展开，促进数学更好地为国民经济和现代化建设服务。

在组织编写这套丛书的过程中，我们曾得到陈开明、陈有根、柳兆荣、徐公权等同志的热情帮助，在此特表谢忱。

复旦大学出版社

1987 年 1 月

序

非线性规划是运筹学的一个重要分支，研究求解极值与约束极值问题的有关理论和算法，在近三十多年迅速发展并日趋成熟。随着电子计算机的发展和普遍采用，作为一种有效的最优化方法，非线性规划在工程设计优化、管理优化、系统分析等多方面的应用日益开拓，愈来愈受到应用部门的重视，非线性规划自身，也在方法和有关理论的研究方面形成特色，取得很大的进展。

作为非线性规划的入门书，本书系统地介绍非线性规划算法的形成和发展，让读者了解研究非线性规划问题的方法并掌握必要的基础。本书既注重非线性规划的理论基础，在内容深度上又力求能为具有一般数学分析(或微积分)和线性代数基础的读者顺利地接受，使广大读者能较好地掌握非线性规划的方法以适应应用的需要。

考虑到在非线性规划领域不断出现新的研究成果，本书力求避免取材上的陈旧和繁琐，着重于使读者迅速具备从事非线性规划的方法与应用方面研究所必须的基础和能力。本书适宜于作为运筹学专业或一般应用数学类或管理类专业的非线性规划教材或参考书，全书的教学约需 120 学时。

许多读者是把非线性规划作为一种重要的最优化技术看待的，他们虽然希望了解非线性规划的内涵，以期在非线性规划的应用方面具有较高水平，但不希望或暂时不希望在目前的学习中过多地涉及非线性规划中较为理论的部分。为适应读者的不同要求，本书注意了编排合理、结构分块，书写中也注意内容简明。着重于非线性规划应用的读者，阅读中删去一些理论方面的深入探讨和若干定理的证明，照样可顺利地系统学习全书。那些整节可暂不阅读的内容，书中在节号上以记号“*”示意。在非线性规划的应用中会出现各式各样的问题，应用工作者只是掌握一些常用非线性规划算法往往还是不够的，编者希望本书有

助于他们提高选用算法和适当修改算法的能力。将内容适当作些删选，本书也可作为工科一些专业的研究生教材，教学一般控制在80学时左右。

本书是在对一些专业的本科生、研究生和进修班的教学实践基础上编写并反复修改而成的，编写过程中得到俞文鲰教授和李元熹副教授很大的帮助，编者在此表示衷心的感谢。限于作者水平，书中错误在所难免，欢迎有关专家和广大读者批评指正。

编著者

1991 年 1 月

目　录

第一章 引 言

非线性规划(Nonlinear Programming) 是运筹学的一个重要分支,主要研究极值问题和约束极值问题的理论和算法,也是近三十年来迅速发展起来的一门新学科分支. 在理论方面,非线性规划从数学的一些分支汲取营养,逐步形成自身的学科特色;在应用方面,非线性规划为系统的优化和管理提供了有用的工具,因而在自然科学、工程、经济、管理等诸多领域获得广泛的应用和重视.

§1.1 非线性规划概述

1. 什么是非线性规划

在应用领域内提出的许多**最优化**(Optimazition)问题中,有一些所对应的数学模型具有如下结构:

第一是最优化问题的**变量**,即所考察的问题可归结为优选若干个称为参数或变量的量 $x_1, x_2, \cdots, x_n$. 这些变量都取实数值,可用 R^n 中的 n 维向量 $x=(x_1, x_2, \cdots, x_n)^T$ 简记,这里 R^n 表示普通的 n 维 Euclid 空间,记号"T"表示矩阵运算转置.

第二是最优化问题中的**约束**,设问题中对变量 $x_1, x_2, \cdots, x_n$所加的限制或约束可以用这些变量满足的如下等式和不等式表示出来:

$$g_i(x_1, x_2, \cdots x_n) \geqslant 0, \qquad i=1, 2, \cdots, m,$$

$$h_j(x_1, x_2, \cdots, x_n)=0, \quad j=1,2, \cdots, p.$$

第三是最优化问题中的**目标**,设最优化问题的目的是在约束限制下确定变量 $x_1, x_2, \cdots, x_n$ 的取值,使得某个 n 个实函数 $f(x_1, x_2, \cdots, x_n)$取到极大值或极小值,或是最大值或最小值. $f(x_1, x_2, \cdots, x_n)$称为这一最优化问题中的目标函数,最优化就是优选变量 x 使$f(x)$极大化(Maximize)或极小化(Minimize).

注意到极大化 $f(x_1, x_2, \cdots, x_n)$ 相当于极小化 $-f(x_1, x_2, \cdots, x_n)$，我们只讨论如下形式的**数学规划**(Mathematical Programming)：

$$
\begin{aligned}
\text{(MP)} \qquad & \min f(x) \\
& \text{s.t. } g_i(x) \geqslant 0, \quad 1 \leqslant i \leqslant m \\
& \qquad h_j(x) = 0, \quad 1 \leqslant j \leqslant p,
\end{aligned}
$$

其中 f, $g_i(1\leqslant i\leqslant m)$, $h_j(1\leqslant j\leqslant p)$ 定义于某个 n 维区域 $D\subset R^n$，记号 min 读作 minimize，含意为极小化，s.t. 是subject to 的缩写，其意为“受限制于”或“受约束于”.

当 f, g_i, h_j 全为 x 的线性函数时，数学规划(MP)便是**线性规划**(Linear Programming)；否则，即目标函数、约束函数中至少有一个不是 x 的线性函数时，数学规划(MP)便称为**非线性规划**.

非线性规划的一般形式也可记为

$$\text{(NP)} \quad \min\{f(x) \mid g_i(x)\geqslant 0,\ 1\leqslant i\leqslant m;\ h_j(x)=0,\ 1\leqslant j\leqslant p\},$$

其中 f, $g_i(1\leqslant i\leqslant m)$, $h_j(1\leqslant j\leqslant p)$ 不全为 x 的线性函数. 如果再引入向量值的约束函数

$$h(x)=\begin{pmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_p(x) \end{pmatrix}, \qquad g(x)=\begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix},$$

其中 h: $D\to R^p$, g: $D\to R^m$, $D\subset R^n$，那么(NP)可简记为

$$\text{(NP)} \quad \min\{f(x) \mid g(x)\geqslant 0,\ h(x)=0\}. \tag{1.1}$$

作为特例，不带有约束的极小化问题

$$\text{(UP)} \qquad \min f(x)$$

称为**无约束极小化问题**(Unconstrained Minimization Problem)，而

$$\text{(EP)} \qquad \min\{f(x) \mid h(x)=0\}$$

称为**等式约束下的极小化问题**(Minimization Problem with Equality Constraints).类似地，**不等式约束下的极小化问题**(Minimization Problem with Inequality Constraints)是指

$$\text{(IP)} \qquad \min\{f(x) \mid g(x)\geqslant 0\}.$$

定义1.1 对于非线性规划(1.1)，称点集

$$\Omega=\{x|g(x)\geqslant 0,\ h(x)=0\} \tag{1.2}$$

为它的**能行集**或**可行集**(Feasible Set),也称**能行域**或**可行域**(Feasible Field)，Ω 中的任一点都称为规划(NP)的能行点或可行点(Feasible Point)．

定义 1.2 对于非线性规划(NP)，设 $x^*\in\Omega$，存在正数 δ，使当 $x\in\Omega$ 且 $\|x-x^*\|<\delta$ 时,总成立

$$f(x)\geqslant f(x^*), \tag{1.3}$$

就称 x^* 是 f 在 Ω 上的一个**局部极小点**(Local Minimum),或称x^*是(NP) 的一个局部约束极小点或局部最优解．如果对一切 $x\in\Omega$，$\|x-x^*\|<\delta$，$x\neq x^*$ 成立

$$f(x)>f(x^*), \tag{1.4}$$

则称 x^*是 f 在Ω上的**严格的局部极小点** (Strong local minimum)，或称为(NP)的严格局部极小点．

定义 1.3 对于非线性规划(NP),设有点$x^*\in\Omega$使(1.3)式对一切 $x\in\Omega$ 成立，就称 x^* 是 f 在 Ω 上的**整体极小点**(Global Minimum)，或称 x^* 是(NP)的整体约束极小点．如果 $x^*\in\Omega$ 使 (1.4) 式对一切 $x\in\Omega$，$x\neq x^*$ 成立，则进一步称 x^* 是 f 在 Ω 上的**严格的**整体极小点,或称 x^*为 (NP)的严格的整体极小点．

求解一个非线性规划，通常是寻求它的局部和整体极小点．我们先用实例给出非线性规划最优解的几何直观．

例 1 设有定义在 R^2 上的函数 $f(x)=(x_1)^2+x_1$，$g(x)=-x_1-x_2+1$，$h(x)=(x_1)^2+(x_2)^2-9$,试考察非线性规划

$$\min\{f(x)|g(x)\geqslant 0,\ h(x)=0\}.$$

解 为了书写的简便,我们约定,这样的非线性规划可写成

$$\begin{aligned}&\min f(x)=(x_1)^2+x_2,\\ &\text{s.t. } g(x)=-x_1-x_2+1\geqslant 0,\\ &\quad\ \ h(x)=(x_1)^2+(x_2)^2-9=0.\end{aligned}$$

在本问题中,变量 $x=(x_1,x_2)^{\mathrm T}$ 的允许取值范围按(1.2)式应为

$$\Omega=\{x|g(x)\geqslant 0,\quad h(x)=0\}.$$

在图 1.1 中，$g(x)\geqslant 0$ 是指直线 AB 的下侧，$h(x)=0$ 是以坐标原点

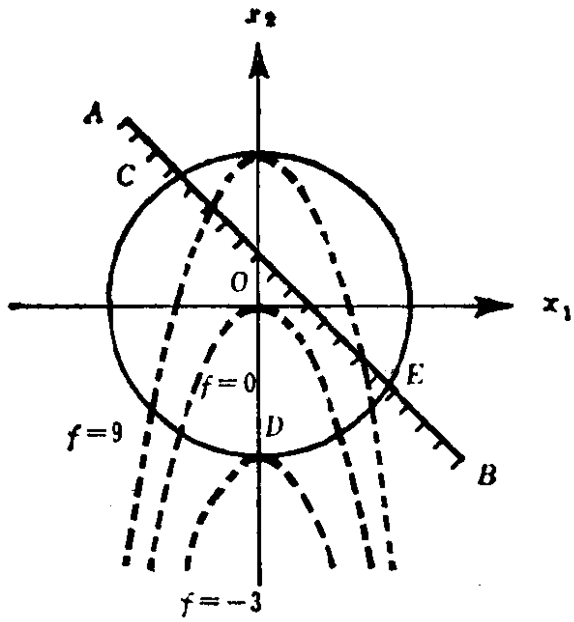


图 1.1

O 为中心的一个圆，于是变量 x 只能在圆弧 $\widehat{CDE}$ 上取值，其中 C 和 E 是直线 AB 与圆 $h(x)=0$ 的两个交点。随着点 x 在圆弧 $\widehat{CDE}$ 上的位置不同，对应的目标函数值 $f(x)$ 也不同。图 1.1 中画了几根目标函数 f 的等值线。注意到等值线 $f=c$ 全体结构成平行的抛物线族，容易明白，当 x 移动到圆弧 $\widehat{CDE}$ 的最低位置 D（坐标为 $x_1=0$，$x_2=-3$）时，$f(x)$ 取到约束范围内的最小值 $f=-3$. 这就利用图形直观找出了变量的最优取值为

$$x^*=(0,-3)^T,$$

它就是这一非线性规划的严格的整体极小点.

2. 建立非线性规划模型

在工程设计、管理决策等许多应用领域内，通过建立数学模型的技巧，有一大类最优化问题能归结为前面提到的非线性规划问题，下面仅举几个简单的例子.

例 2 二杆桁架的最优设计模型.

由两根等长钢管 AB 和 AC 构成的二杆桁架见示意图 1.2，其中 D-D' 是钢管的截面图. 设桁架的跨度 $2s$ 已经选定，钢管的厚度 t、抗压强度 σ_0、弹性模量 E 也是已知的. 所谓最优设计问题，就是要确定钢

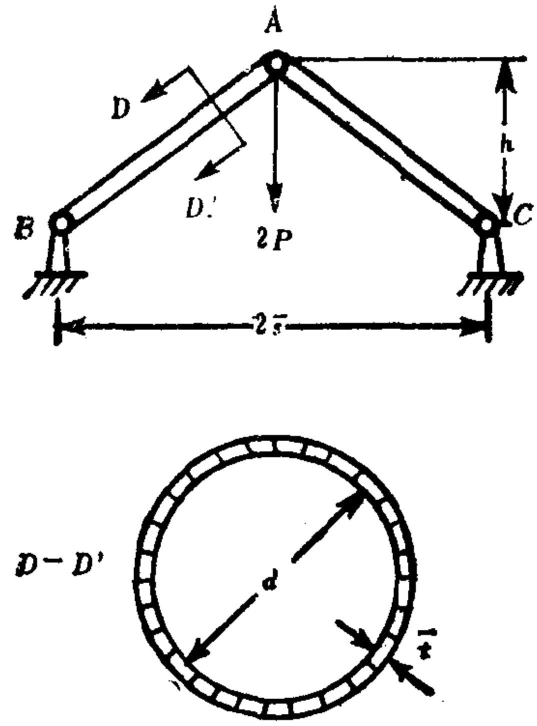


图 1.2

管的直径 d 和桁架的高度 h，使得桁架在点 A 处承受垂直载荷 $2P$ 时不出现屈曲或弹性变形，并使得桁架尽可能轻.

解 本问题中需要最优化的变量是设计变量 d 和 h，而 $\bar{s}$，$\bar{t}$，σ_0，E，P 都是已知量.

设计的目标是使桁架的重量最轻，这等价于使桁架的体积最小，所以目标函数可取为

$$f(d,h)=2\pi d\bar{t}(h^2+\bar{s}^2)^{\frac{1}{2}}.$$

余下只要考察问题对 d，h 的约束限制.

当桁架于 A 点承受垂直载荷 $2P$ 时，每根钢管所受的压力是 $P(h^2+\bar{s}^2)^{\frac{1}{2}}/h$，因而杆件的应力（即单位截面积所受的压力）为

$$\sigma=\frac{P(h^2+\bar{s}^2)^{\frac{1}{2}}}{\pi d\bar{t}h}.$$

根据结构力学原理，对于选定的钢管，不出现屈曲的条件是 $\sigma\leqslant\sigma_0$（σ_0 是钢管最大许可的抗压强度），而不出现弹性弯曲的条件可导出为

$$\sigma\leqslant\frac{\pi^2E(d^2+\bar{t}^2)}{8(h^2+\bar{s}^2)}.$$

这样，若再注意到 d,h 的选择还受到尺寸的限制，整个最优设计模型可用以下非线性规划表示：

$$\min\ f(d,h)=2\pi dE(h^2+\bar{s}^2)^{\frac{1}{2}},$$
$$\text{s.t.}\ g_1(d,h)=\sigma_0-P(h^2+\bar{s}^2)^{\frac{1}{2}}/(\pi d\bar{t}h)\geqslant 0,$$
$$g_2(d,h)=\frac{\pi^2E(d^2+\bar{t}^2)}{8(h^2+\bar{s}^2)}-\frac{P(h^2+\bar{s}^2)^{\frac{1}{2}}}{\pi d\bar{t}h}\geqslant 0,$$
$$g_3(d,h)=d-d_1\geqslant 0,$$
$$g_4(d,h)=d_2-d\geqslant 0,$$
$$g_5(d,h)=h-h_1\geqslant 0,$$
$$g_6(d,h)=h_2-h\geqslant 0.$$

例 3 曲线的最优拟合模型.

在实验数据处理或统计资料分析中，常常利用有关变量的数据资料去确定这些变量之间的关系.例如，设已取得有关变量 x_1，x_2，…，x_n 与 y 取值的以下资料：

数据资料序号	x_1	x_2	…	x_n	y
1	x_{11}	x_{12}	…	x_{1n}	y_1
2	x_{21}	x_{22}	…	x_{2n}	y_2
⋮	⋮				
i	x_{i1}	x_{i2}	…	x_{in}	y_i
⋮	⋮	⋮			⋮
p	x_{p1}	x_{p2}	…	x_{pn}	y_p

并据经验估计变量 x_1，x_2，…，x_n 与 y 之间具有以下形式的函数关系：

$$y=f(x_1,x_2,\cdots,\ x_n,\ c_1,\ c_2,\ \cdots,\ c_m), \tag{1.5}$$

其中 c_1，c_2，…，c_m 是待定参数，问题是如何确定这些参数.

解 本问题需要最优化的变量是

$$c=(c_1,\ c_2,\ \cdots,\ c_m)^{\mathrm{T}}.$$

把每一组数据资料看作关于变量 x_1，x_2，…，x_n 和 y 的一次试验结果.譬如说，对于第 i 次试验，变量 x_1，x_2，…，x_n 的取值为

$$x_1=x_{i1},\quad x_2=x_{i2},\ \cdots,\quad x_n=x_{in}, \tag{1.6}$$