实用非线性规划

〔美〕D. M. 希梅尔布劳 著

科学出版社

实用非线性规划

〔美〕D. M. 布梅尔布劳 著

张义燊 等 译

科 学 出 版 社

1 9 8 1

内 容 简 介

本书主要论述非线性规划的实用技术．全书共分三篇．第一篇包括前两章，叙述一般的非线性规划问题的基本知识；第二篇包括三章，论述了无约束最优化的各种算法，其中包含有利用导数的有关技术、搜索技术以及对各种较为流行的算法的评价；第三篇包括四章，论述了约束最优化的各种算法，其中包含有线性化法、罚函数法、可变容差法及对有关方法的评价．作者力图使本书方便于实际部门的读者，附了大量的图、详细的计算步骤以及计算框图和程序．

本书适于系统分析、运筹学、数值分析、管理科学、特别是工程科学的研究人员、研究生，高等学校的教师和高年级学生参考或自学．

D. M. Himmelblau

APPLIED NONLINEAR PROGRAMMING

McGraw-Hill Book Company, 1972

实用非线性规划

〔美〕D. M. 希梅尔布劳 著

张义燊 等 译

*

科学出版社 出版

北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1981 年 5 月第 一 版　开本：787×1092　1/32

1981 年 5 月第一次印刷　印张：17

印数：0001—7,660　字数：390,000

统一书号：13031・1557

本社书号：2137・13—1

定价： 2.60 元

前　　言

本书的目的是尽可能简单地叙述和评价非线性规划中的几种比较有效的方法. 虽然已经提出许多关于解决一般非线性规划的算法,但是将这些算法应用于大型问题时,只有少数被证明是有效的. 没有一种算法能够归结为是解决非线性规划的万能妙方.

本书旨在介绍一些流行的非线性规划算法，而对它们用于某类问题的收敛性并不给出数学上的证明. 虽然这些证明是重要的,但它们只能用于非常有限的几类问题,所以顶多不过是某一种算法的使用者的基础知识. 与线性规划相反,对于非线性规划,对某一个特殊策略来说,即使不容易得到收敛的证明,但却是有效的. 反之亦然,即已知一种算法对某些特殊情形是收敛的,但是对于更为复杂的问题,却很少能提供令人满意的策略上的见解.

只有那些在实践中证明是多少有点效果的非线性规划算法,才予以较为详细的叙述. 对那些比较起来好得多的算法,本书解释得更详细些，并对相应的计算机程序作了尽可能详细的评价. 比较程序所用的标准是:

(1) 可靠性(求解成功);

(2) 求解速度;

(3) 使用者准备时间;

(4) 解的准确度;

(5) 约束满足程度.

所考虑方法是用来优化一个非线性函数的，所受约束是

非线性等式和（或）不等式，这些约束全部是含有大量独立变量的函数．所有这些变量同随机变量相反，都是确定的．所处理的问题既不包含整数规划或分式规划，也不包含动态问题的最优化，即以时间为参量的问题．所有方法只有借助大型高速数字计算机或复合式计算机才能计算，不考虑用模拟计算机．

我们并不试图将各最优化方法研究得很深，但是我们还是足够详细地提供了使读者能够遵循的每一个方法的基本步骤．然而经常看到，编制算法程序的巧妙的细节，特别是常用的那些步骤，对算法的性能有着重要的影响．每一节所举的详细解出的例子说明了各种算法的式样，画出了大量的框图以说明其逻辑关系．附录 B 列出几个较好的算法的计算机程序，这些程序在公开的文献中或商业组织中不容易得到．因为对所有的算法都采用了统一的符号，因而从它们彼此之间的结构关系以及它们的共同特点中，比从相应的文献的评论中，能得到更多了解．

像 Gaul 的书一样，本书共分三篇．第一篇包括前两章，第一章是一个简单引言，第二章叙述了一般的非线性规划问题及其建立，以及它同其他类型的规划之间的关系，还有最优解存在的必要充分条件．

第二篇讲述无约束非线性规划的算法．第三章叙述了梯度、二阶导数和利用导数的有关策略．第四章包括搜索策略。第五章评价各种无约束算法．

第三篇讲述约束最优化算法．第六章讲述线性化方法．第七章讲述罚函数法．第八章讲了可变容差法．第九章对约束算法进行评价．附录 A 包括大量考题及其解．除第一章外，每一章都有留给读者解决的问题．

了解各算法所必须的基础是：微积分知识、粗通矩阵和

矢量符号及它们的运算，以及简单熟悉线性规划的解法．对于那些不熟悉矩阵符号的人来说，在附录C中概述了本书所用到的一些基本性质．为了阅读计算机程序，熟悉一些Fortran 程序是必须的，但还是附有说明，以便那些只知道怎样给卡片穿孔的使用者可以按说明使用程序．这种运用已被证明是成功的，自然对那些粗心大意的人来讲，将会摔跤的．

最后，我要感谢A. R. Colville、J. Abadie 及 M. R. Anderberg 三位博士的帮助；Darcy A. Paviani 博士给予极好的援助，第八章就是根据他的学位论文写成的；David C. Stocker 和 Desmond Bond 的大量的计算分析对我是极好的帮助．他们的所有成就使我们有可能对许多约束和无约束的算法进行评价．

D. M. 希梅尔布劳

目　　录

第三篇 约束非线性规划方法

第一篇　基础知识

本篇叙述了非线性规划问题的概况，阐明了它和真实问题的关联，并定义与非线性规划有关的某些术语．另外，还叙述了能告诉我们一个假想的最优解是不是实际最优的方法．

第一章 引 言

从前，人们举行各种精心安排的仪式来帮助人们对某些事情做出决定．他们曾往地上倾注美酒或屠宰动物来祭神，有的则看星相，看鸟飞等等．他们把信念编成谚语以及靠已有的经验对生计做些推测．可是今天，支配人们对事物做出决定的是靠新的多半是更科学的活动，那就是利用计算机．如果没有任何工具的帮助，人们的智力无法处理各种各样的复杂问题，例如企业的管理、导弹的设计、交通路线的安排中的问题．由于利用了高效率的数字计算机以及复合式计算机，发展了许许多多的数学最优化方法．数学规划是属于这些最优化方法，而在数学规划中非线性规划是一种特殊情况．

数学规划这一术语是由 Robert Dorfman 在 1950 年前后提出来的，现在包含整数规划、线性规划、凸规划、非线性规划、网络流动理论、动态规划、不定过程中的规划等等的一般术语．非线性规划研究非线性或线性函数在线性与(或)非线性约束条件下的最优化问题．典型的应用领域是预报、生产流程的安排、库存控制、质量控制、保养和维修、过程设计、会计过程以及资金预算等．全由线性函数组成的最优化问题，总有单纯形算法来解它们，从这个意义上说，对非线性规划的最优化问题则没有普遍适用的解法．因此，在某种程度上非线性规划存在着研究的一个经验领域．迄今为止，通过特殊算法的提出和规划，检验把这些算法用于感兴趣的问题所得的结果，并且在这些经验的基础上来构造一些更好的算法，这样得到向前发展．

在过去二十年中，数学规划领域中大部分研究工作集中在线性规划范围内．这方面的贡献如此之多，以致使得当时的技术水平能在很精彩的形式下解决大多数的线性规划问题．

另一方面，已经提出了解决非线性规划问题的许多策略，但是提出的比应用成功的要多得多．现行的非线性规划算法的适用范围是有限的．随着计算机的进一步改进和更精确地描述现实世界中存在的问题的日益增长的需要，确实需要具有更广泛应用的解非线性规划的方法．

现实世界中存在的大多数问题都有好几个解，有些还可以有无限多个解．最优化的目的是，对于一个给定问题，在这些可能的解中，按某种有效性或可使用性的准则来找到最好的可能的解．只具备一个解的问题，不必化为最优化问题．最优化可通过许多策略来实现，包括从非常复杂的解析的和数值的数学方法到简单算术的巧妙应用．假定需要最优化的问题已经用某种方法确定下来（不一定以某种数学形式给出），则一般最优化的各种方法宜于分类如下：

（1）**解析法**．它利用微分学和变分法的经典方法，通过寻找使函数 $f[\mathbf{x}]$ 对 $\mathbf{x}$ 的导数等于零的 $\mathbf{x}$ 的值来求函数的极值．如果欲求函数在约束条件下的极值，那么就利用拉格朗日乘子法及约束变分法．为了应用解析方法，最优化问题必须用数学语言来描述，因而函数和变量均可按熟知的法则进行运算．对于大型高度非线性问题，解析法是不令人满意的，本书不予讨论．

（2）**数值法**．它利用已有的资料，通过迭代程序来产生最优化问题的更好的解．数值法能解决解析法所不能解决的问题，因为实际问题证明容易用数值法处理，因而非线性规划的数值法是本书所叙述的一种方法．

解最优化问题比较有效而本书又没介绍的其他常用的方法有:

(3) **图解法**. 它就是将求解极大值或极小值的函数作为单变量或多变量函数,画出它的图形. 函数的极值,在图形上通过直接观察来得到. 图解法的优点是,比较初等,而且立刻就可看出解是否存在. 不过,这个方法限于只有一个变量或最多有两个自变量的情况下适用.

(4) **实验法**. 一个函数的极值可通过实际变化过程的直接实验的办法来得到,而不做数学上的运算. 一次实验结果用来安排下一次实验,以便不断改进操作结果.

(5) **情况研究法**. 这个方法是把同一问题的许多典型解做估计,以确定可能的最好解. 于是,用情况研究方法得到的这个"最好"的解,很可能是次最优解.

在将真实过程通过数学方法实现最优化的过程中,会遇到许多困难,为讨论方便,我们将这些问题分成两类. 一类是将真实问题用数学模型表示出来的问题,而另一类就是数字解法问题. 我们在这里只讲清这些困难,并且联系到特定的算法指出怎样把它们解出来.

数学模型包含着最优化过程中所要处理的函数. 不言而喻,数学模型应当合理地表示出真实过程的重要特性,如果我们想求解有意义的极值的话. 即使这个条件被满足,数学模型所遇到的典型困难也还有:

(1) 最优化的判别式对独立(决定)变量的改变可能不敏感,因此无法十分明确地判断其极值.

(2) 在极值搜索范围内,要优化的判别式或者一个或多个约束可能成为无界,或者是在此模型中函数的偏导数变成无界的. 在分母中含有多项式的模型容易出现这种情况,如:

$$y = \frac{b_0 + b_1x_1}{b_2x_1 + b_3x_2}$$

在这里,函数 y 和 y 对 x_1 的一阶偏导数

$$\frac{\partial y}{\partial x_1} = \frac{-b_0b_2 + b_1b_2 - b_1b_2x_2 + b_1b_3x_3}{(b_2x_1 + b_3x_2)^2}$$

都在 $b_2x_1 = -b_3x_2$ 时变为无界的. 克服这个困难的方法是给问题加上限制,或者是重新提出数学模型,以适当限制独立变量变化范围.

(3) 这些变量中有可能出现病态标度. 例如在判别式中的某一项和其他项相比，从各项的有效数字来看其数量级相差太大时,这种情况有可能发生. 于是,判别式对较小的项的那些变量的变化很不敏感. 例如,目标函数

$$y = 100x_1^2 - 0.010x_2^2$$

的值,在 x_2 变化时无影响,除非 x_2(由于它的物理单位)比 x_1 大得很多.若 x_2 的数量级与 x_1 的相同,将一个或两个变量同时乘以标度因子,以改变方程式右端两项到数量级粗略地相等.令

$$\tilde{x}_1 = 10x_1 \qquad x_1^2 = 10^{-2}\tilde{x}_1^2$$

$$\tilde{x}_2 = 10^{-1}x_2 \qquad x_2^2 = 10^2\tilde{x}_2^2$$

于是目标函数中的各项变成相同数量级. 在

$$y = \tilde{x}_1^2 - \tilde{x}_2^2$$

的极值求得以后, x_1 和 x_2 的值可由 $\tilde{x}_1$ 和 $\tilde{x}_2$ 的值来确定. 当然,要重新标度一个数学模型的各函数,不可能象刚才所说的那么容易.

(4) 在粗糙的数学模型中,各变量之间有可能相互影响. 参量之间的相互影响，可通过检验一个具有两个变量相乘的极其简单的判别式

$$y = 2x_1x_2 + 10$$

来说明. 对乘积 x_1x_2 的一给定值, x_1 和 x_2 的个别数值可以

取任意一系列数值．在有相互影响的情况下，作标度是较为困难的．二次函数可以变换为标准形式，使其不含变量的乘积项．可以定义新的坐标轴，叫做**主轴**，二次曲面对于该轴是对称的．例如，曲面

$$y = 7x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 4x_2x_3 - 6x_1 - 24x_2 + 18x_3 + 18$$

可以通过坐标轴的平移和旋转变换为

$$y - 18 = 3\tilde{x}_1^2 + 6\tilde{x}_2^2 + 9\tilde{x}_3^2$$

在新的坐标系里，每一项的标度比原来坐标系无疑是更清楚的．只是在把数学模型用某一适当的变换作些变动的时候，或用截尾泰勒级数逼近的时候，非线性函数才变成二次函数．一个巧妙的但是易受变量间相互影响的一个例子，包含着如

$$y = x_1 e^{b_1 x_2}$$

的一个模型，其中 x_1 和 $e^{b_1 x_2}$ 相乘．

（5）数学模型中可能有**零效应**．零效应可用下列判别式加以说明：

$$y = x_1^2 + 2x_1x_2 + x_2^2 + 2$$
$$= (x_1 + x_2)^2 + 2$$

在作变换 $x_1 + x_2 = \tilde{x}_1$ 之后，我们得到

$$y = \tilde{x}_1^2 + 2$$

我们看到，只剩下唯一的为求 y 的极值需要变化的变量 $\tilde{x}_1$．

第二类困难与最优化问题的数值解法相关联．

（1）如何能获得对独立变量的适当的初始推测呢？因为问题包含着非线性函数，所以可能有一个以上的极值，而这一个特点是线性分析所不能具备的．因此，如果对变量的初始推测离极值太远，那么最优化过程可能达不到全局（最好的）极值，而是某些其他的极值．独立变量的近似最优值常常从前面的研究或从物理理由中得到．最后办法是在可行域中试

验几个起始矢量，判别在极值点处它们是否都使判别式产生相同的数值，但是这种解决办法也有危险，如同在关于构成合适的模型所遇到的困难时所作注释中见到过的那样.

（2）怎样处理实际变量的随机特性呢？我们在本书中忽略了数学模型中系数和变量可能是随机变量这个非常真实的可能性.

（3）怎样减小数字计算误差？截断函数的误差减低了许多算法的效力. 稳定性与近似非线性规划问题的解是否收敛到原来问题的解作为极限有关. 在最优化过程中舍入误差也十分令人讨厌，特别当导数用差分格式取近似时尤其如此.

象所有其他数学工具一样，非线性规划方法不能事先不加任何考虑而盲目地应用于给定的问题. 象所有工匠的工具一样，非线性规划的应用要求使用者有某些熟练的技巧. 不可避免地要求仔细设计数学模型，而且要求采用适当的数字方法.

本书其余各章基本上分成三篇. 第一篇叙述非线性规划问题及其演变. 第二篇叙述并比较无约束非线性规划的方法. 第三篇提出并评价约束非线性规划的实用算法. 附录中包括许多试验问题及其解，以及市面上得不到的非线性规划的某些计算机算法程序.

第二章　非线性规划问题及其最优解

在工程学、经济学、自然科学以及其他领域中，分析工作者经常面临着的是使设备、操作、电路或生产程序的复杂安排最优化的问题. 他所希望的是要使得表示价值、重量、产量或其他同类东西的某种称为目标函数的函数，在某些约束下取得最小值或最大值. 用数学语言来描述时，这些最优化问题的很大一类可归入称之为非线性规划问题的范畴；解决这样一些问题的方法称为非线性规划. 在本章中，我们首先形式化地描述一般的非线性规划问题以及某些特殊的子问题. 然后用一个例子来描述实际过程的最优化目的与最优化的数学表示之间的关系. 其次，扼要叙述一些定义和术语，最后，说明最优性的条件.

2.1　线性规划问题

线性规划问题是一种规划问题，其中要极小化或极大化的判别函数是线性函数，判别函数所受的约束也是些线性函数. 一般用 X_i 表示的一些标量或矢量的组合称为**线性的**，只要这些标量或矢量可以组合成形式

$$c_1X_1 + c_2X_2 + \cdots + c_nX_n \tag{2.1-1}$$

其中 c_i 全为常数. 例如，函数 $4x_1 + 3x_2 + 5x_3 + 2$对变量 x_1, x_2, x_3 是线性的，而函数 $2x_1^2 + x_1x_2 + 3e^{x_3}$ 对于同是这些变量来说却是非线性的. 一组 X_i 称为**线性相关**，只要有一组常数 c_i（假定 c_i 不全为零）使下式成立: