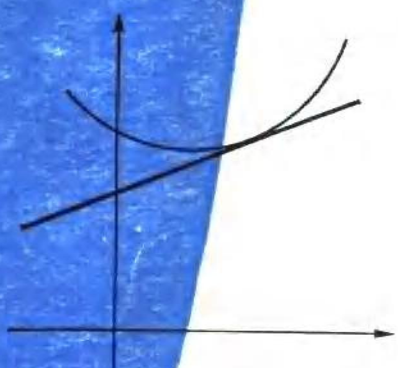
M·阿佛里耳著

非线性规划 上册

——分析与方法



上海科学技术出版社

非 线 性 规 划

—分析与方法—

上 册

M. 阿佛里耳 著

李元熹 陈开明 魏国华 等译

金福临 俞文魮 校

上海科学技术出版社

Nonlinear Programming
Analysis and Methods
M. Avriel
Prentice-Hall, Inc., 1976

非 线 性 规 划

——分析与方法——

上 册

〔美〕M. 阿佛里耳 著
李元熹 陈开明 魏国华 等译
金福临 俞文鮆 校

上海科学技术出版社出版
(上海瑞金二路 450 号)
新华书店上海发行所发行 上海市印刷六厂印刷
开本 850×1156 1/32 印张 6.75 字数 175,000
1979 年 10 月第 1 版 1979 年 10 月第 1 次印刷
印数 1—10,000

书号: 13119·775 定价: 0.78 元

译校者序

在20多年中发展起来的非线性规划，是运筹学的数学理论中特别重要而且比较活跃的一个分支，可认为它是有限维最优化经典理论的再创造，其新的特征主要在于：所研究的最优化问题允许复杂的约束，对最优性、对偶性诸方面进行深入的分析，并强调提出可行的算法和进行理论概括．以后者为重点的那部分内容亦常称为最优化方法．在应用上，在现今电子计算机日趋发展的情况下，非线性规划正越来越多地用于最优设计、管理科学、系统识别、数学物理变分问题等方面，并由此推动非线性规划的进一步发展；在理论上，它从数学的其他一些分支中吸取营养，特别联系着凸分析、线性代数、计算方法等方面，并逐步形成本身的特色．

本书相当全面地介绍了非线性规划，内容包括分析与算法，取材较新．如共轭函数与对偶性的理论、变尺度方法的统一处理等，在同类的著作中很少见到．同时，在一些章节中，还提到某些有待继续研究的问题．

本书对应用数学、运筹学、计算数学、工程设计、管理科学等方面的研究工作者、教师和学生，具有较好的参考价值．我们希望，这个中译本能对我国在非线性规划的理论方面与应用方面的发展有一定的推动作用．

参加本书译校工作的同志有：金福临、俞文魮、李元熹、陈开明、魏国华、吴廷芳、阎振中、王芬、李贤平、张国梁．我们在分头译出的基础上，经过几次统一的校阅．限于水平，一定还有不妥之处，欢迎读者提出宝贵意见．

1978年9月

前　言

最优化是在复杂的实际环境中遇到的许多可能决策中挑选最好决策的科学. 本书的主题是非线性最优化，或者就是更常说的非线性规划，在诸如运筹学、管理科学、工程、经济、系统分析和计算机科学等这样不同的领域中，它是越来越显得重要的一门学科.

为了作出决策，有条理的做法通常包括三个紧密相关的阶段. 第一阶段的目标是建立一个体现所考察决策问题的数学模型. 决策者建立模型时首先确定变量；然后收集有关数据；列出有待最优化(极大化或极小化)的目标函数；把变量和数据安排到一组称为约束的数学关系式(例如方程和不等式)中去. 第二阶段继续上一过程，对数学模型进行分析，并选择一个适当的求最优解的数值方法. 如果模型的一个或几个约束，或目标函数是非线性的，这个最优化问题是一个非线性规划. 非线性规划的分析，对于问题的结构提供了有价值的观察，并回答了关于能行决策和最优决策的存在性与特征性质问题. 非线性规划的数值解法依赖于目标函数和约束的性质以及最优化问题的结构. 在分析的基础上，决策者在现有的数值解法中确定哪一种(如果有的话)能够求得最优解. 第三阶段是求最优解，多数是用计算机. 在求解过程中，决策者或者采用一种可用的计算机程序或者建立一个新程序，实施在第二阶段选用的算法. 为完成整个过程还需要用决策来解释数值解，最后，要对所用算法的收敛性、效率和误差分析等实用方面作出评价.

由于现实世界情况的精确表示常常使数学模型包含非线性函数，非线性规划的分析和有效的数值方法的发展，对于涉及到最优决策各个不同学科的学生和实际工作者进行学习来说，已经成为紧要的事项. 本书对非线性规划给以范围广泛的综览，向读者介

绍前面提到的决策过程第二阶段的技艺状况.

本书包含两部分. 第 I 部分, **分析**, 首先导出数学规划最优性的必要与充分条件, 其次介绍凸集与凸函数并导出凸规划理论. 再用近代的共轭函数方法阐述非线性凸对偶性. 然后, 以广义凸性和对几种非线性规划的分析来结束第 I 部分. 第 **II** 部分, **方法**, 讨论非线性最优化问题的数值解法. 首先介绍无约束最优化方法的研究, 包括不用导数的方法, 共轭方向法和变尺度方法等几章. 对于在这部分研究的有约束非线性规划来说, 惩罚函数法、无约束方法的扩充和近似型算法, 皆属于最为广泛使用的或者最近发展起来并被认为有发展前途的算法.

在一本书里要包括非线性规划的每个方面几乎是不可能的. 因此, 对某些重要的主题如互补性理论、离散最优化和随机规划或者未予列入或者仅仅提及一下. 同样, 也略去数值方法的某些计算方面的细节, 如截断误差、病态条件和计算机实现. 有兴趣的读者能够在更专门的著作中找到这些论题.

本书的材料曾作为一些大学高年级学生和研究生进行非线性最优化教学的基础. 需要的数学基础是微分学(偏导数、级数展开)、初等线性代数(矩阵和向量运算)和实分析(集、极限和数列、连续性). 本书不要求预先了解最优化, 特别是并不需要预先知道线性规划. (以下略)

MORDECAI AVRIEL

目　录

第 1 章

引　言

数学规划是最优化理论的一个分支. 在数学规划中, 对 n 个实变量 $x_1, \cdots, x_n$ 的一个单值目标函数 f 进行极小化(或极大化), 这些变量可能受限制于有限个写成不等式或方程的约束. 一般, 我们定义一个极小化的数学规划(Mathematical Program)为

$$\text{(MP)} \qquad \min f(x) \tag{1.1}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{1.2}$$

$$h_j(x) = 0, \quad j=1, \cdots, p, \tag{1.3}$$

其中 x 是分量为 $x_1, \cdots, x_n$ 的列向量. 换句话说, (MP)是求一个满足(1.2)和(1.3)的向量 x^*, 使 $f(x)$ 取极小值——也就是最优值. 如果出现在(MP)中的函数有一个或多个是 x 的非线性函数, 我们称之为**非线性规划**. 这与线性规划相反, 那里的一切函数都必须是线性的. 非线性规划某些基本方面的研究是本书的主题.

非线性规划问题, 出现在如工程、经济、商业管理、物理科学和数学这样一些不同的学科中, 或者出现在具有下述性质的其他领域中, 那里需要在某些复杂 (或相冲突) 的情况下作出(在广义下的)决策, 而这些情况又能用数学模型表示. 为了说明某些类型的非线性规划, 下面介绍几个例子.

非线性曲线拟合

在某些科学研究中, 例如在生物学或物理学中, 假设某种现象 f 作为时间的函数在实验室中被测量, 还假设这现象的数学模型已经给定, 并且根据模型知道 f 之值随时间 t 的变化为

$$f(t) = x_1 + x_2 \exp(-x_3 t). \tag{1.4}$$

实验室试验的目的是根据在时刻 $t^1, t^2, \cdots, t^M$ 测得的 f 之值来求未知参数 x_1, x_2 和 x_3. 这一决策过程包含要指定参数值, 于是自

然地要求 x_1, x_2 和 x_3 的数值在某种意义下为最优. 例如,我们可以在最小二乘方意义下求参数的最优值, 这就是说, 求那些数值, 使实验曲线与理论曲线的偏差的平方和为最小. 形式上, 我们有非线性规划

$$\min F(x_1,\ x_2,\ x_3)=\sum_{i=1}^{M}[f(t^i)-x_1-x_2\exp(-x_3t^i)]^2. \tag{1.5}$$

注意到这是一个无约束规划,解出的参数值或许是不能接受的.为了避免这一情况,可以加上一些约束形式的限制. 例如,参数 x_3 限定为非负,即

$$x_3\geqslant 0. \tag{1.6}$$

对考察的特殊现象,还可假设所提出的数学模型,只在参数选为能使 $t=0$ 时 $f(0)=1$, 才是可以接受的. 因此必须加上约束

$$x_1+x_2=1, \tag{1.7}$$

受限制于(1.6)和(1.7)而求解(1.5),则是一个约束非线性规划问题,具有非线性的目标函数及线性的不等式和等式约束.

定位问题

现在假设要选定一个供应中心的位置, 这中心向城市中有固定空间位置的 m 个用户提供服务. 中心供应的商品可以是电、水、牛奶或其他货物. 供应中心的定位准则是使从中心到用户的某个距离函数最小. 例如, 很可能我们感兴趣的是使中心到任何一个用户的**最大距离**为最小. 由于在这个城市里货物必须沿垂直的路线(如街道)供应, 合适的距离函数是所谓**矩形**距离. 作为数学模型列出来, 设 $(x_1,\ x_2)$ 表示供应中心的待定位置(坐标), 而 $(a^i,\ b^i)$ 是第 i 个用户的给定位置. 我们的问题则是

$$\min_{x_1,x_2}\{\max_{1\leqslant i\leqslant m}[|a^i-x_1|+|b^i-x_2|]\}. \tag{1.8}$$

这个式子意味着, 首先对 $(x_1,\ x_2)$ 的每个可能值, 必须求指标 i 使方括弧中的矩形距离最大,其次在依赖于 $(x_1,\ x_2)$ 的所有最大距离中求出最小的. 另外, 如果每一位置 $(x_1,\ x_2)$ 都可以接受, 那末问题是无约束的. 但是,有时为了简化某些表示式,以添加额外变量

和约束为代价是有利的. 例如,定义一个新变量 x_0:

$$x_0=\max_{1\leqslant i\leqslant m}[|a^i-x_1|+|b^i-x_2|] \tag{1.9}$$

或

$$x_0\geqslant|a^i-x_1|+|b^i-x_2|,\quad i=1,\cdots,m, \tag{1.10}$$

我们就得到一个有三个变量 x_0, x_1, x_2 的非线性规划:

$$\min f(x)=x_0 \tag{1.11}$$

受限制于

$$g_i(x)=x_0-|a^i-x_1|-|b^i-x_2|\geqslant 0,\quad i=1,\cdots,m. \tag{1.12}$$

读者容易验证,问题(1.8)和问题(1.11)至(1.12)在以下意义下等价:(x_1^*, x_2^*)为(1.8)的最优解的充要条件是(x_0^*, x_1^*, x_2^*)为(1.11)和(1.12)的最优解,而且对某个 $k(1\leqslant k\leqslant m)$成立

$$x_0^*=|a^k-x_1^*|+|b^k-x_2^*|. \tag{1.13}$$

读者还能证明,通过引进更多的变量,(1.8)可以化为一个线性规划问题.

过程设计

在连续搅拌的槽式(回收混合物)反应器中,以反应物 A 的水溶液为原料,考察每小时生产一定量 F_B 克分子的化学产品 B 的问题. 化学反应为

$$2A\rightarrow B, \tag{1.14}$$

且符合由实验建立的,以反应液的单位体积为基准的经验速率方程:

$$-\frac{dC_A}{dt}=8.4(C_A)^2=8.4[C_A^0(1-x_A)]^2\left(\frac{\text{克分子}}{\text{升-小时}}\right), \tag{1.15}$$

其中 C_A——反应器中 A 的浓度(克分子/升),

C_A^0——原料中 A 的浓度(克分子/升),

t——时间(小时),

x_A——转换因子,反应物转化为产品的份数.

假设 A 的浓度在某一连续范围内时,原料溶液是可用的,并

且单位价格 p_A 由下列关系给出:

$$p_A=4(C_A^0)^{1.4} \text{ (美元/升)}, \tag{1.16}$$

连续搅拌槽式反应器(CSTR)的操作费用为

$$p_{\mathrm{CSTR}}=0.75(V)^{0.6} \text{ (美元/小时)}. \tag{1.17}$$

其中 V(升)是反应器的容积. 假定产品 B 的售价为 10 美元/克分子, 我们的问题是去决定: 原料溶液的速率 F_A^0(升/小时)、它的浓度 C_A^0、反应器的容积 V 和转换因子 x_A, 以达到最优操作. 也就是说, 使下列每小时的总收益为最大:

$$p_T=10F_B-p_AF_A^0-p_{\mathrm{CSTR}} \text{ (美元/小时)}. \tag{1.18}$$

反应器中物质的平衡给出

$$F_A^0C_A^0=F_A^0C_A^0(1-x_A)-\left(\frac{dC_A}{dt}\right)V. \tag{1.19}$$

由(1.14)得到

$$\frac{1}{2}F_A^0C_A^0x_A=F_B, \tag{1.20}$$

并由(1.15)和(1.19)得到

$$8.4C_A^0(1-x_A)^2V-F_A^0x_A=0. \tag{1.21}$$

于是, 设计问题成为

$$\max p_T=5F_A^0C_A^0x_A-4(C_A^0)^{1.4}F_A^0-0.75(V)^{0.6}, \tag{1.22}$$

受限制于约束(1.21). 这是变量 C_A^0, F_A^0, V, x_A 的非线性规划, 目标函数和约束函数都是非线性的. 注意, 这里是使目标函数极大化.

在整个这本书中, 我们将主要关心目标函数极小化的非线性规划. 这并不算任何限制, 因为每一个 $\max f(x)$型的问题, 通过考察 $\min \bar{f}(x)$, 其中 $\bar{f}(x)=-f(x)$, 总可以等价地进行分析和求解.

最后, 对以后各章所用的记号和术语说几句. 没有用特别的符号来表示向量. 向量的维数如不特别说明, 总可以从它出现的

式子中看出．所有的向量假定是列向量．一个向量的分量以下标表示，这样，x_1, x_2, $\cdots$, x_n 是 n 维向量 x 的分量．向量的上标用来区分不同的向量，这样，x^1, x^2, $\cdots$, x^m 表示 m 个不同的向量．为了避免混淆起见，出现实数的指数时将带有括弧，就是说 $(a)^2$ 是数 a 的平方．记号 x^T 用来表示一个行向量，即列向量的转置．实直线，即所有实数的集合用 R 来记，n 维实 Euclid 空间记为 R^n．向量 $x\in R^n$ 也常认为是 R^n 中的点 x．R^n 中特定的点常以它的坐标来表示，例如 $x^0=(-1, 2)$．记号 $x\geqslant 0$ 表示 x 的每个分量都是非负的．这样，如果 $0\in R^n$，那末 0 也是一个 n 维向量，它的每一个分量为零．记号 $x\neq 0$ 表示 x 至少有一个分量异于零．

仅仅用到具有实元素的矩阵．和向量类似，对矩阵也没有特别的符号(虽然一般用大写字母)．矩阵的维数从它出现的式子中唯一地确定．如果 A 是 $m\times n$ 阵(m 行 n 列)，那末 A 的转置 A^T 有 n 行 m 列．矩阵 A 的逆阵记为 A^{-1}，即 $AA^{-1}=A^{-1}A=I$，其中 I 是恒等(单位)阵．

向量 $x\in R^n$ 的模定义为

$$\|x\|=[(x_1)^2+(x_2)^2+\cdots+(x_n)^2]^{1/2}. \tag{1.23}$$

我们常用邻域的概念．集合

$$N_\delta(x^0)=\{x:x\in R^n,\ \|x-x^0\|<\delta\} \tag{1.24}$$

称为点 x^0 的一个(球形)邻域，这里 δ 是一正数．

矩阵的模在本书少数地方也用到，而它们是由向量诱导出的模．形式地说，如果 A 是一个 $m\times n$ 阵，x 是 n 维向量，那么

$$\|A\|=\sup\left\{\frac{\|Ax\|}{\|x\|}:\|x\|\neq 0\right\}=\sup\{\|Ax\|:\|x\|=1\}. \tag{1.25}$$

函数总是单值的；如以后将看到的．它们有时可取值 $+\infty$ 或 $-\infty$．在几处需要较深数学概念的地方，或是在书中定义它；或是在这个定义要求更多的背景材料时，我们就放弃完整性，要求读者去参看适当的文献．读者如对初等线性代数，实分析或拓扑不熟悉，希望去查阅这些方面的入门性教科书，诸如 Apostol[1]，Bartle[2]，Hall、Spencer[3] 与 Noble[4]．

参 考 文 献

1. Apostol, T. M., *Mathematical Analysis*, 2nd ed., Addison-Wesley Publishing Co., Reading, Mass., 1974.

2. Bartle, R. G., *The Elements of Real Analysis*, 2nd ed., John Wiley & Sons, New York, 1976.

3. Hall, D. W., and G. L. Spencer, *Elementary Topology*, John Wiley & Sons, New York, 1955.

4. Noble, B., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1969.

第Ⅰ部分　分　　析

如不作特别说明，求解一个非线性规划问题的涵义是求一个最优解向量 x^*，而不是求可能存在的所有最优解．判定一个最优解 x^*，并研究它的性质，构成本书第Ⅰ部分的中心论题，它处理非线性规划问题的某些分析方面的内容．我们将会看到，如果向量 x 是最优解的候选者，它必需满足某些最优性必要条件．但是，不幸的是可能有其他异于最优解的向量也满足这些条件．这样，必要条件主要是在否定意义下使用的：如果向量 x 不满足条件，它就不是最优解．因此，为了验证最优性，我们希望找出最优性的充分条件，当它与必要条件一起被满足时，就清楚地指明了所考虑的具体解向量的性质．这两种类型的最优性条件是要稍详细讨论的第一个主题．特别地，把经典的 Lagrange 乘子法扩充到具有不等式约束的最优化问题．在一般的非线性规划中，向量可能满足一类最优性条件而不满足另一类．最优性必要条件也是充分条件的规划是很重要的，因为这时在所考察的规划中，表示最优性必要条件的一组方程或不等式的解，必然是这规划所求的最优解．

称为凸规划的一类非线性规划就具有上述这样好的性质，它们以某种结构方式含有凸函数和凹函数．这种凸规划特别便于分析．与每一个这样的规划相联系，存在所谓对偶规划，与线性规划类似，它具有某些有趣的理论性质．对偶性关系将借助于基于共轭函数的近代方式进行推导．

凸规划属于每个局部极小都是整体极小的那一类非线性规划．由于有许多非凸的实际问题，我们还得考虑一个重要问题，就

是找出具有这种局部-整体最优性质的、较一般的非凸函数和非凸规划. 作为本书分析部分的结尾，选择了一些非线性规划，对它们阐明某些理论性的结果.

第 2 章

经典最优化——无约束和等式约束问题

求实函数的极值，即极小值或极大值的问题，在数学最优化中处于中心地位．我们从最简单的无约束问题开始这个极值课题，然后进入带有等式约束的极小和极大的论题．这里讨论经典的 Lagrange 乘子理论以及可微函数极值的某些必要和充分条件．这些课题的处理可以回溯到几个世纪以前，所以有"经典的"称呼．后面几章将讨论有不等式约束的最优化问题．这个问题得到的所有值得注意的结果可以算是"近代的"，因为它们是近二三十年来对不等式约束问题的强烈兴趣所引出的结果．所有"经典的"结果可以看成更一般的"近代"理论的特殊情况．我们先介绍经典的结果，是因为它们可以作为桥梁，把多数在第一和第二学年开设的大学微积分或实分析课程的内容，与数学规划更深一些的主题连接起来．此外，经典理论比近代理论在这样的意义下更简单，即诸如关于极值的充要条件等结果，不会象不等式约束情形那样被更复杂的要求弄得模糊不清．

2.1 无约束极值

考虑 R^n 中区域 D 上的一个实值函数 f，称 f 在一点 $x^*\in D$ 有**局部极小值**，如果存在一个实数 $\delta>0$，使得

$$f(x)\geqslant f(x^*) \tag{2.1}$$

对适合 $\|x-x^*\|<\delta$ 的一切 $x\in D$ 成立，同样可定义**局部极大值**，只要倒转(2.1)中的不等号．如果不等式(2.1)换为严格不等式

$$f(x)>f(x^*),\quad x\in D,\quad x\neq x^*, \tag{2.2}$$

就称 f 在 x^* 有**严格**局部极小值；而如果(2.2)的不等号反向，则得到严格局部极大值．如果(2.1)(或(2.2))对一切 $x\in D$ 成立，称