非线性规划及其理论

应玖茜 魏权龄 编著

中国人民大学出版社

非线性规划及其理论

应玫茜　魏权龄　编著

中国人民大学出版社

（京）新登字 156 号

非线性规划及其理论

应玫茜 魏权龄 编著

出版者：中国人民大学出版社
发行者：中国人民大学出版社
（北京海淀路 39 号 邮码 100872）
印刷者：北京市丰台区印刷厂
经销者：新华书店总店北京发行所
开 本：850×1168 毫米 32 开
字 数：356 000
印 张：14.375
版 次：1994 年 9 月第 1 版
印 次：1994 年 11 月第 2 次印刷
册 数：2 501—5 000
书 号：ISBN7-300-01643-X/F・442
定 价：11.70 元

前　言

非线性规划是运筹学的一个重要分支，它在军事、经济、工程、管理以及生产过程自动化等方面都有重要的应用。非线性规划作为一个独立的学科是在本世纪50年代才开始形成的。大型电子计算机的产生和使用大大地促进了它的发展。到70年代，这门学科开始处于兴旺发展时期。在国际上，这方面的专门性研究机构、刊物以及书籍犹如雨后春笋般地出现，国际学术会议召开的次数大大地增加。在我国，随着电子计算机日益广泛地使用，非线性规划的理论和方法也逐渐地引起很多部门的重视。关于非线性规划理论和应用方面的学术交流活动也日益频繁，我国的科学工作者和广大群众在这领域内已取得了一些可喜的成果。

本书侧重于介绍非线性规划的理论。作为非线性规划基础的凸分析部分在本书中占较大的比重；非线性规划本身的基本理论部分也是本书的重点之一；关于非线性规划方法，书中除了讲述计算方法之外，更着重于方法的收敛性讨论。同时，书中还收集了国内的一些成果。

全书共分四大部分。第一部分，凸分析。包括仿射集、凸集、凸锥、凸函数和广义凸函数等内容。鉴于凸分析与其他数学分支的密切关系，这一部分对于从事其他有关分支研究与应用的读者也可作为参考。第二部分，非线性规划基本理论。包括最优解的性质，最优解的充分必要条件，对偶理论以及稳定性理论等内容，它们对于深入研究非线性规划是不可缺少的。第三部分和第四部

分是计算方法。包括一维搜索法、直接法、无约束方法、带约束方法等。除一维搜索法外，几乎所有方法的收敛性都有严格的证明。对于所有方法，我们都提供了完整的计算步骤和大致思路。

鉴于篇幅以及使用的数学工具所限，还有一些国内的成果未能收集在本书中。例如：关于求总体极值的随机方法，可参见[22]；关于直接最优化方法可参见[20]、[21]；关于最小平方和问题，可参见[23]、[24]。此外，与非线性规划有密切关系的多目标最优化问题，可参见[16]、[18]、[19]、[34]、[35]、[36]、[37]。

在酝酿和编写本书的过程中，我们得到有关高等院校教师、从事实际工作的有关人员、中国科学院计算中心五室以及原数学研究所运筹室等同志的大力支持和帮助，在此表示衷心的感谢。

应玫茜　魏权龄

目　　录

Ⅰ. 凸　分　析

Ⅱ. 基 本 理 论

Ⅲ. 无条件极值问题

Ⅳ. 条件极值问题

I. 凸分析

第一章 仿射集

本章的目的是介绍一些仿射集的基本性质。我们假定读者对向量空间的基本性质已有所了解，因而对一些熟知的概念（例如子空间、维数、交集、并集、和集等）不作任何说明。

1.1 仿射集和子空间

我们记 E_n 为在实数域上的 n 维列向量空间（欧氏空间），即

$$E_n=\{(x_1,\cdots,x_n)^T \mid x_1,\cdots,x_n \text{ 为任意实数}\}$$

其中符号 T 为向量的转置号。

1.1.1 定义 集合 $M\subseteq E_n$ 称为**仿射集**，若对任何 $x,y\in M$，有 $\lambda x+(1-\lambda)y\in M$ 对一切 $\lambda\in E_1$ 成立。

空集 $\varnothing$，E_n 本身以及只包含一个点 a 的集合 $\{a\}$ 显然都是仿射集。由归纳法可以证明：M 是仿射集，当且仅当，对任意正整数 $k\geqslant 2$，当 $x^i\in M(i=1,\cdots,k)$ 时，有 $\sum\limits_{i=1}^{k}\lambda_i x^i\in M$ 对一切 $\lambda_i\in E_1(i=1,\cdots,k)$，$\sum\limits_{i=1}^{k}\lambda_i=1$ 成立。

由定义可以直接证明下述定理

1.1.2 定理 设仿射集 $M_j\subseteq E_n(j\in J)$，则

(i) $\sum\limits_{j\in J}M_j=\{\sum\limits_{j\in J}x^j \mid \text{一切 } x^j\in M_j, j\in J\}$ 为仿射集（J 为有限

指标集)；

(ii) $\bigcap_{j\in J} M_j$ 为仿射集(J 为任意指标集)。

1.1.3 定理 E_n 的任意子空间必为仿射集。任何包含零向量的仿射集必为子空间。

证 定理的第一个断言显然成立。设 M 为包含零向量的仿射集。任取 $x\in M,\lambda\in E_1$，于是 $\lambda x=(1-\lambda)0+\lambda x\in M$。任取 $x,y\in M$，于是 $\frac{1}{2}(x+y)=\frac{1}{2}x+(1-\frac{1}{2})y\in M$，$x+y=2[\frac{1}{2}(x+y)]\in M$，故 M 为一子空间。 证毕

1.1.4 定理 $M\subseteq E_n$ 为仿射集，当且仅当，存在 $x\in M$，以及子空间 L，使

$$M=x+L=\{x+y\mid y\in L\}$$

证 设 M 为仿射集，任取 $x\in M$，作集合

$$L=M-x=\{z-x\mid z\in M\}$$

欲证 L 为一子空间，任取 $y^1,y^2\in L$，于是存在 $x^1,x^2\in M$，使 $y^1=x^1-x,y^2=x^2-x$。因此，对任意 $\lambda\in E_1$，有

$$\lambda y^1+(1-\lambda)y^2=\lambda x^1+(1-\lambda)x^2-x\in M-x=L$$

即 L 为仿射集。又因 $0=x-x\in L$，由 1.1.3，L 为子空间，则 $M=x+L$。

设存在子空间 L 和 $x\in M$ 使 $M=x+L$。任取 $x^1,x^2\in M$，于是存在 $y^1,y^2\in L$，使 $x^1=x+y^1,x^2=x+y^2$。故对任何 $\lambda\in E_1$，有 $\lambda x^1+(1-\lambda)x^2=x+\lambda y^1+(1-\lambda)y^2\in x+L=M$，即 M 为一仿射集。

证毕

我们称仿射集 $M=x+L$ 是子空间 L 的一个**平移**，或称仿射集 M **平行于**子空间 L。

1.1.5 定理 设仿射集 $M\subseteq E_n$，则存在唯一的子空间 L，使之对任何 $x\in M$，有 $M=x+L$，且

$$L=M-M=\{x-y\mid x\in M,y\in M\}$$

证 设有 $x^1, x^2 \in M$ 以及子空间 L_1, L_2 使

$$M = x^1 + L_1 = x^2 + L_2$$

因 $0 \in L_1$，故 $x^1 - x^2 \in L_2$，$x^2 - x^1 \in L_2$，$L_1 = x^2 - x^1 + L_2 \subseteq L_2$。同理，$L_2 \subseteq L_1$。故 $L_1 = L_2$，即存在唯一子空间 L 使

$$M = x^1 + L = x^2 + L$$

任取 $x \in M$，于是存在 $y \in L$，使 $x = x^1 + y$。由于 $-y + L = L$，故 $M = x^1 + L = x - y + L = x + L$。

任取 $z \in L$，则必存在 $x, y \in M$，使 $z = y - x$，即 $L \subseteq M - M$。反之，对任何 $x, y \in M$，有 $z \in L$，使 $y = x + z$，故 $y - x = z \in L$，即 $M - M \subseteq L$。因而，$L = M - M$。 证毕

1.1.6 定义 设 $M = x + L$ 为仿射集，L 为子空间，我们称 L 的维数 $\dim L$ 为 M 的**维数**，记作 $\dim M$。特别，若 $\dim M$ 分别为 0，1 和 2 时，分别称 M 为**点**、**直线**和**平面**；若 $\dim M = n-1$，称 M 为 E_n 中的**超平面**。

设 L 为 E_n 中任一子空间，则其正交补 $L^{\perp} = \{y \mid x^T y = 0, \text{一切 } x \in L\}$ 也是子空间，且 $\dim L + \dim L^{\perp} = n$。设 $L^{\perp}$ 的正交补为 $L^{\perp\perp}$，则 $L^{\perp\perp} = L$。若 $\dim L^{\perp} = m \geqslant 1$，任取 $L^{\perp}$ 的一组基 $a^1, \cdots, a^m$，则

$$L = L^{\perp\perp} = \{y \mid a^{i^T} y = 0, i = 1, \cdots, m\}$$

由此，对一个 $n-m$ 维的仿射集 M 而言，存在 $c \in M$，以及子空间 L，使

$$\begin{aligned} M &= c + L = c + \{y \mid a^{i^T} y = 0, i = 1, \cdots, m\} \\ &= \{c + y \mid a^{i^T} y = 0, i = 1, \cdots, m\} \\ &= \{x \mid a^{i^T} x = b^i, i = 1, \cdots, m\} \\ &\quad (b^i = a^{i^T} c, i = 1, \cdots, m) \end{aligned}$$

又易见线性方程组 $a^{i^T} x = b^i (i = 1, \cdots, m)$ 的解集为仿射集，也就是 m 个超平面的交集。总结上述，可得下述二定理。

1.1.7 定理 设 $a \in E_n$，$a \neq 0$，$b \in E_1$，则

$$H = \{x \mid a^T x = b, x \in E_n\}$$

为 E_n 中的一超平面；反之，任何 E_n 中的超平面可以表示成这种形式。

1.1.8 定理 设 $b\in E_m$，A 为 $m\times n$ 矩阵，则

$$M=\{x\mid Ax=b,x\in E_n\}$$

为 E_n 中的仿射集；反之，任何 E_n 中 $n-m$ 维（$1\leqslant m\leqslant n$）的仿射集可以表示成这种形式。

由此可见，任何子空间 L 可表成 $L=\{x\mid Ax=0,x\in E_n\}$。

1.2 仿射包

1.2.1 定义 给定集合 $S\subseteq E_n$，我们称 $M(S)=\{\sum_{i=1}^{k}\lambda_i x^i\mid k$ 为任意正整数，$x^i\in S,\lambda_i\in E_1(i=1,\cdots,k),\sum_{i=1}^{k}\lambda_i=1\}$ 为由 S 生成的**仿射包**。（易见 $M(S)$ 为一仿射集）。若对仿射集 M 而言，存在集合 S，使 $M=M(S)$，则称 S 为 M 的**生成集**。称向量 $a^0,a^1,\cdots,a^m$ 为**仿射无关**，若 $\dim M(\{a^0,a^1,\cdots,a^m\})=m$。

由于

$$\begin{aligned}&M(\{a^0,a^1,\cdots,a^m\})\\&=\{\sum_{i=0}^{m}\lambda_i a^i\mid\lambda_i\in E_1(i=0,1,\cdots,m),\sum_{i=0}^{m}\lambda_i=1\}\\&=a^0+\{\sum_{i=1}^{m}\lambda_i(a^i-a^0)\mid\lambda_i\in E_1(i=1,\cdots,m)\}\end{aligned}$$

$L=\{\sum_{i=1}^{m}\lambda_i(a^i-a^0)\mid\lambda_i\in E_1(i=1,\cdots,m)\}$ 为包含 $a^i-a^0(i=1,\cdots,m)$ 的最小子空间，$\dim L=m$ 当且仅当 $a^1-a^0,\cdots,a^m-a^0$ 线性无关。因此，$a^0,a^1,\cdots,a^m$ 仿射无关，当且仅当 $a^1-a^0,\cdots,a^m-a^0$ 线性无关。

若仿射集 M 的生成集为 $S=\{a^0,a^1,\cdots,a^m\}$，则其平移子空间

L 可由 $a^1-a^0,\cdots,a^m-a^0$ 生成。于是 M 中的任意 x 可唯一地表示成

$$x=\sum_{i=0}^{m}\lambda_i a^i,\quad \sum_{i=0}^{m}\lambda_i=1$$

当且仅当 $a^0,a^1,\cdots,a^m$ 仿射无关，这时 $\lambda_1,\cdots,\lambda_m$ 为 x 在子空间 L 有坐标分量。称 $\lambda_0,\lambda_1,\cdots,\lambda_m$ 为 x 在 M 上的**重心坐标分量**。

1.2.2 定理 设集合 $S\subseteq E_n$，则 $M(S)$ 是包含 S 的最小仿射集，也是包含 S 的一切仿射集的交集。S 是仿射集，当且仅当 $M(S)=S$。

证 因任何包含 S 的仿射集 M 必有 $M\supseteq M(S)$，而 $M(S)$ 又为仿射集，故 $M(S)$ 为包含 S 的最小仿射集。由1.1.2，一切包含 S 的仿射集的交集为仿射集，且为包含 S 的最小仿射集，故此交集即为 $M(S)$。第二个结论显然成立。 证毕

第二章 凸 集

本章的目的是介绍一些凸集以及与之有关的基本性质。2.1节，引入凸锥、凸集以及凸包的定义，并阐明其间的关系；2.2节，介绍单纯形、集合的维数以及由集合 S 生成的凸包、凸锥和仿射集的一些性质；2.3节，介绍集合的相对内点、相对边界的性质；2.4节，介绍各种支撑和分离定理；2.5节，介绍凸集相交定理。

2.1 凸锥、凸集和凸包

2.1.1 定义 集合 $C\subseteq E_n$ 称为以 o **为顶点的凸锥**，若对一切 $x,y\in C$，有

$$\lambda_1 x+\lambda_2 y\in C \qquad \text{一切 } \lambda_1\geqslant 0,\lambda_2\geqslant 0,\lambda_1+\lambda_2>0$$

若 C 是以 o 为顶点的凸锥，则称 $K=x^0+C$ 为**以 x^0 为顶点的凸锥**。K 可看作 C 的平移。

2.1.2 定义 设集合 $S\subseteq E_n$，称集合

$$C(S)=\{\sum_{i=1}^{k}\lambda_i x^i \mid \text{任何正整数 } k\geqslant 1;x^1,\cdots,x^k\in S,$$

$$\lambda_i\geqslant 0(i=1,\cdots,k),\sum_{i=1}^{k}\lambda_i>0\}$$

为**由 S 生成的凸锥**。显见 $C(S)$ 不一定包含原点 o，$C(S)$ 是以 o 为顶点的凸锥。

2.1.3 定义 集合 $S\subseteq E_n$ 称为**凸集**，若对任何 $x,y\in S$，有

$$\lambda x+(1-\lambda)y\in S \qquad \text{一切 } \lambda\in[0,1]$$

2.1.4 定理 集合 $S\subseteq E_n$ 为凸集，当且仅当，对任何正整数 $k\geqslant 1$，当 x^1，…，$x^k\in S$ 时，有 $\sum_{i=1}^{k}\lambda_i x^i\in S$ 对一切 $\lambda_i\geqslant 0$ $(i=1,\cdots,k)$，$\sum_{i=1}^{k}\lambda_i=1$ 成立(试用归纳法证明)。

2.1.5 定义 设集合 $S\subseteq E_n$，集合

$$H(S)=\{\sum_{i=1}^{k}\lambda_i x^i \mid \text{任意正整数 } k\geqslant 1;\ x^1,\cdots,x^k\in S,\lambda_i\geqslant 0(i=1,\cdots,k),\ \sum_{j=1}^{k}\lambda_j=1\}$$

称为**由 S 生成的凸包**。易见，$H(S)$ 为凸集。

2.1.6 定理 设集合 $S\subseteq E_n$，则 $H(S)$ 为包含 S 的最小凸包，也是包含 S 的一切凸集的交集。

证 可与1.2.2类似地证明之。

2.1.7 定理 设集合 $S_1,S_2\subseteq E_n$，则

$$H(S_1\cap S_2)\subseteq H(S_1)\cap H(S_2)$$

证 因 $H(S_i)(i=1,2)$ 为凸集，而

$$S_1\cap S_2\subseteq S_i\subseteq H(S_i)\quad i=1,2$$

由2.1.6，即得欲证。

注 2.1.7中等号不一定成立，例如

$$S_1=\{(x_1,x_2)^T \mid 0\leqslant x_1<1,x_2=1\}\cup\{(x_1,x_2)^T \mid x_1=1,0\leqslant x_2<1\}$$

$$S_2=\{(x_1,x_2)^T \mid x_1>0,x_2>0,x_1+x_2=1\}$$

易见 $S_1\cap S_2=\varnothing$(空集)，$H(S_1\cap S_2)=\varnothing$，$H(S_1)\cap H(S_2)=S_2$。