凸分析

史树中

上海科学技术出版社

凸 分 析

史 树 中

上海科学技术出版社

内 容 提 要

凸分析是六十年代末期形成和发展起来的一个新的数学分支。它在数学规划、最优控制、变分学、逼近论、数理经济学等许多方面都有广泛的应用。

本书以较少的篇幅、较高的观点系统地阐述凸分析的数学基础。对凸集和凸函数的性质、凸集分离定理、凸函数的次微分运算、凸规划的对偶性等，都有较详尽的论证。书中还介绍了凸分析的各种应用，尤其注意凸分析在经济学方面的应用背景。书中附有大量习题，既可启发读者进一步思考钻研，有些也是正文的补充。

本书取材翔实，推导严谨，行文流畅，深入浅出。既可供理工科大学高年级学生及研究生作教材用，也可供教学工作者、科研工作者参考。

责任编辑　赵序明

凸　分　析

史树中

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　浙江诸暨报印刷厂印刷
开本850×1156　1/32　印张6.75　字数163,000
1990年4月第1版　1990年4月第1次印刷
印数：1—2,50
ISBN 7-5323-1758-7/O·137　定价：4.00元

前 言

六十年代中期，由于数学规划论、对策论、数理经济学、逼近论、变分学、最优控制理论等多方面的需要，出现了一个新的数学分支——凸分析. 这一分支由于其基本内容相当初等，而应用又十分广泛，因此其许多结果很快就成为广大数学工作者手中的有力工具，甚至进入了大学应用数学课程. 本书以较少的篇幅但又较深入地介绍凸分析的基本理论及其应用.

凸分析的奠基著作是文献[1](Moreau J. J., Fonctionnelles Convexes, Seminaire sur les équations aux dérivées partielles, Collège de France, 1966～1967.) 和 [2] (Rockafellar R. T., Convex analysis, Princeton Univ. Press, Princeton, New York, 1969(俄译本 1973).)，前者是一份从未公开出版过的讲义，它对凸分析理论建立了一个非常一般的框架，并对凸分析的基本结果进行了根本的剖析. 因此，它具有较大的理论意义，但并不适合应用工作者. 后者则完全针对数学规划论的需要，对有限维空间上的凸分析作了详尽讨论. 由于该书所需的预备知识很少，因而在应用数学界影响较大. 不久，又出版了文献[3](Ekeland I. et Teman R., Analyse convexe et probèmes variationnels, Dunod-Gauthier-Villars, 1974 (英译本，1976; 俄译本，1979))，进一步讨论了无限维空间上的凸分析及其在变分问题上的应用，使凸分析的影响更为深远. 此后，有关凸分析的教程和专著不断出现(例如[4]～[7])，凸分析就变得更为普及.

编写本书所面临的第一个问题是在怎样的框架中讨论凸分析. 凸分析的基本研究对象是凸集和凸函数；基本工具是凸集分离定理. 这些概念和定理都可以纯代数地来研究，也就是说，可在一个不引入拓扑的线性空间中来研究. 因此，就凸分析的基本内

容而言，凸分析完全可表达为“凸代数”．有的数学家甚至走得更远，连线性空间的框架也被抛开，而在一个仅有抽象凸性的“凸空间”中讨论凸分析(例如[8])．但是过分抽象，使结果太一般，并无多大用处．较合适的框架还是拓扑线性空间，因为引进拓扑后，能使(无限维空间的)对偶空间缩小，有关对偶性的讨论就更为精确，这正是Moreau[1]所做的．然而，从应用的角度来看，不少人认为对于数学规划论等分支来说，有限维空间就足够了，而对于更多的分支来说，Hilbert空间或者至多是自反Banach空间也已足够了．因此，近年来出现的一系列有关凸分析的书往往都局限于这两种情况．也许人们认为进一步的理论基础已经由Moreau等人奠定，为了便于入门和应用，不必深究那些更一般的问题．对于本书来说，一种方便的做法，也是采用这样的框架．在这种框架中，作为凸分析的基本工具的凸集分离定理的证明，可以利用空间的距离和内积，并且避免公开出现选择公理(实质上还是用了)．但是考虑到国内的一般读者不但无处去觅Moreau的[1]这样未公开出版过的讲义，甚至连有关无限维线性空间和拓扑线性空间理论的合适的参考书也较难找到．为了使对凸分析有较深刻的了解，我们还是采用了象Moreau[1]采用的那样更一般的框架，即在对偶空间不起大作用时，我们在线性空间中讨论；而在对偶空间起本质作用时，我们在拓扑线性空间中讨论．为了使本书有较大的独立性，原则上只假定读者有初等分析和初等集合论的知识(但对抽象逻辑推理能力的要求则较高)．本书也概述了无限维线性空间和拓扑线性空间的基本内容，并给出了所需要的全部证明，使其与凸分析的内容融为一体．这样做也使本书有了自己的特色，而不至于成为已有的专著的翻版或编译，相当一部分结果的陈述形式是本书所特有的．

第二个问题是取材．除了文献[1]、[2]以外，几乎没有别的“纯凸分析”专著，甚至在[2]中也大量讨论了凸规划问题．在文献[3]～[7]中，真正的凸分析内容实际上都不到一半，甚至更少．本书当然也不是只写“纯凸分析”，而是兼顾介绍它的各方面的应用、

但是为了避免涉及面太广和篇幅太大，且限于作者本人的知识面，本书没有像[3]～[7]那样，大面积铺叙凸分析的各种应用，而只是以通常的凸规划作为主要应用课题贯穿全书．仅在最后一节列举了一些凸分析在各分支中的应用例子，但也只是为了说明应用的可能性，无意得出任何较有意义的结果．至于凸分析本身的内容，本书旨在阐明凸分析的理论基础、主要概念和主要结果．根据作者的看法，凸分析的整个理论几乎都是建立在凸集分离定理上．因此书中作了较细致的讨论．由此出发，我们认为凸分析的最主要的结果应该是以下几个：(1)（在分离局部凸空间中）每个闭凸集是闭半空间族的交；(2)每个下半连续凸函数是连续仿射函数族的上包络；(3)凸函数的次微分概念和 Moreau-Rockafellar 定理，即

$$\partial(f_1+f_2)(x)=\partial f_1(x)+\partial f_2(x)$$

在一定条件下成立；(4)凸规划的一般对偶理论．其中(1)、(2)应该是没有异议的，因为应用最多的凸集和凸函数的对偶性都起源于此．对于(3)、(4)，由于应用起来不很方便，故不少作者都更强调 Fenchel 对偶定理（包括[2]），但实质上还是一样的．本书的取材也就紧扣这条主线，而与此关系不大的概念和结果则少讨论，甚至不讨论．例如，书中没有讨论有关凸集端点的Крейн–Мильман 定理和 Choquet 积分表示定理，没有引入凸集的渐近锥、凸函数的 inf-卷积等重要概念，没有深入讨论凸锥理论，更没有涉及凸函数的极值存在定理、凸集上的不动点理论等等．有一点需要指出的是：书中把定理 2.5.1 放在相当中心的位置．它虽然也为人们所熟知，甚至还被 Berge[9] 称为“凸函数的基本定理”．但多数有关凸分析的著作对此并不很重视．而本书利用它证明许多重要结果．

本书共分四章．第一章讨论线性空间中的凸集．中心内容是代数形式的凸集分离定理．我们给出了它的最一般叙述．证明采用了文献[10]中利用超锥的证法．在这一章中，顺便也讨论了无限维线性空间的特征．第二章的前三节讨论了单变量凸函数和线性空间上的凸函数的性质，尤其是得出了凸函数所表示为仿射函

数族的上包络的代数陈述．后三节讨论了凸规划及其经济解释，它们可以说是凸分析发展的主要推动力之一．第三章首先是简要概述了有关拓扑空间和拓扑线性空间的必要知识．我们这里基本上是自成体系的，但只能满足于仅仅想粗知拓扑空间和拓扑线性空间知识的读者，而代替不了这方面的简要教程．然后是把凸集和凸函数放到拓扑线性空间中讨论．第四章的内容是凸分析构成为一门新数学分支的标志．对于基础较好的读者，可以从这里读起，而把前三章作为备查的内容．

每一章后面都配备一定数量的习题，供读者练习和进一步思考(大多数习题都不难)．

在正文前面，附了若干重要记号、定义和主要结果，以便于读者掌握本书中的基本内容．

本书在很大程度上是作为一本教材来编写的，它并不追求大而全．自从 Ekeland 变分原理 (1972)、Clarke 广义梯度 (1975) 等问世以来，比凸分析更为广阔的研究领域，即所谓“非凸分析”或“非光滑分析”正在展现．有兴趣的读者可参看刚出版的新书[15]和[16]，从中可看到凸分析以后的新发展．

本书是作者在 1982～1984 年间关于凸分析的多次教学和讲学的讲稿基础上的整理和总结．在最后的定稿中，吸收了许多同志提出的各种宝贵意见．姚景齐同志仔细审阅了本书的初稿，提出了宝贵的意见和评注，在此表示衷心感谢．尽管如此，限于作者的水平，本书中还可能存在不妥和错误之处，期待读者批评指正．至于本书中处理凸分析的方法，更需要同行同志们作出鉴定．

作者

于南开大学数学所

记号、定义和主要结果

实数域　$\mathbf{R}$. $\mathbf{R}_+$ 则表示非负实数全体.

n 维实线性空间　$\mathbf{R}^n$. $\mathbf{R}^n_+$ 则表示 $\mathbf{R}^n$ 中有非负分量的向量全体.

线性空间及其代数对偶　一般表示为 X 和 X^*. 从第三章起也用它们来表示拓扑线性空间及其拓扑对偶.

线性空间的特殊子集类:

子空间 L:

$$\lambda_1 L+\cdots+\lambda_n L\subset L,\ \forall\lambda_1,\ \cdots,\ \lambda_n\in\mathbf{R}.$$

仿射集 A:

$$\lambda_1 A+\cdots+\lambda_n A\subset A,\ \forall\lambda_1,\ \cdots,\ \lambda_n\in\mathbf{R},\ \sum_{i=1}^n\lambda_i=1.$$

凸集 K:

$$\lambda_1 K+\cdots+\lambda_n K\subset K,\ \forall\lambda_1,\ \cdots,\ \lambda_n\in[0,\ 1],\ \sum_{i=1}^n\lambda_i=1.$$

凸锥 C:

$$\lambda_1 C+\cdots+\lambda_n C\subset C,\ \forall\lambda_1,\ \cdots,\ \lambda_n>0.$$

以上 n 为任何 $\geqslant 2$ 的整数.

集合 E 的各种包(或称 E 张成的子空间、仿射集、锥和凸集):

线性包: $\operatorname{lin} E=\{x\mid x=\sum_{y\in E}\lambda_y y,\ \lambda_y\in\mathbf{R}$, 有限个不为零$\}$.

仿射包: $\operatorname{aff} E=\{x\mid x=\sum_{y\in E}\lambda_y y,\ \lambda_y\in\mathbf{R}$, 有限个不为零, $\sum_{y\in E}\lambda_y=1\}$.

锥包: $\operatorname{cone} E=\{x\mid x=\lambda y,\ y\in E,\lambda>0\}=\bigcup_{\lambda>0}\lambda E$.

凸包: $\operatorname{co} E=\{x\mid x=\sum_{y\in E}\lambda_y y,\ \lambda_y\in[0,\ 1]$, 有限个不为零, $\sum_{y\in E}\lambda_y=1\}$.

集合 A 的代数内部:

$$A^{i}=\{x \mid \forall h \in X,\ \exists \varepsilon>0,\ \forall t \in[0,\ \varepsilon],\ x+th \in A\}.$$

集合 A 的相对代数内部:

$$A^{ri}=\{x \mid \forall h \in \operatorname{aff} A-x,\ \exists \varepsilon>0,\ \forall t \in[0,\ \varepsilon],\ x+th \in A\},$$

当 A 为有限维非空凸集时,

$$A^{ri} \neq \varnothing.$$

集合 A 的代数闭包:

$$A^{c}=\{x \mid \exists h \in X,\ \forall \varepsilon>0,\ \exists t \in[0,\ \varepsilon],\ x+th \in A\}.$$

当 A 为凸集, 且 $A^{ri} \neq \varnothing$ 时, A^{ri}、A^{c} 的表示: 对于任何 $x_0 \in A^{ri}$, 有

$$A^{ri}=\bigcup_{x_1 \in A \setminus\{x_0\}}[x_0,\ x_1)=\bigcup_{x_1 \in A^{ri} \setminus\{x_0\}}[x_0,\ x_1)=\bigcup_{x_1 \in A^{c} \setminus\{x_0\}}[x_0,\ x_1);$$

$$A^{c}=\{x \mid [x_0,\ x) \subset A\}=\{x \mid [x_0,\ x) \subset A^{ri}\}.$$

无限维线性空间 X 的特征(充要条件):

i) $X \neq X^{**}$;

ii) 存在凸集 $A \subset X$, $\operatorname{aff} A=X$, $A^{i}=\varnothing$;

iii) 存在凸集 $A \subset X$, $(A^{c})^{c} \neq A^{c}$;

iv) 存在凸集 $A \subset X$, $A \neq X$, $A^{c}=X$.

凸集分离定理: 设 A、B 为线性空间 X 中的凸集, $A^{ri} \neq \varnothing$, $B^{ri} \neq \varnothing$(或 $A^{i} \neq \varnothing$), 且 $A^{ri} \cap B^{ri}=\varnothing$ (或 $A^{i} \cap B=\varnothing$). 则 A、B 可用超平面分离, 即存在 $x^{*} \in X^{*}$, $x^{*} \neq 0$,

$$\langle x^{*},\ x\rangle \leqslant\langle x^{*},\ y\rangle, \quad \forall x \in A,\ \forall y \in B.$$

A^{ri}、B^{ri}(或 A^{i}, B)可用超平面严格分离, 即存在 $x^{*} \in X^{*}$, $x^{*} \neq 0$,

$$\langle x^{*},\ x\rangle<\langle x^{*},\ y\rangle, \quad \forall x \in A^{ri}(A^{i}),\ \forall y \in B^{ri}(B).$$

凸集强分离定理: 设 A、B 为线性空间 X 中的两个集合, $A-B$ 凸, $(A-B)^{ri} \neq \varnothing$, $0 \notin(A-B)^{c}$. 则 A、B 可用超平面强分离, 即存在 $x^{*} \in X^{*}$, 使得

$$\sup_{x \in A}\langle x^{*},\ x\rangle<\inf_{y \in B}\langle x^{*},\ y\rangle.$$

拓扑线性空间中的凸集分离定理: 设 A、B 为拓扑线性空间 X 中的两个集合, $A-B$ 为凸集, $\operatorname{int}(A-B) \neq \varnothing$, 且 $0 \notin \operatorname{int}(A-B)$ (或 A、B 凸, $\operatorname{int} A \neq \varnothing$, $\operatorname{int} A \cap B=\varnothing$). 则 A、B 可用闭超平

面分离；0 与 $\operatorname{int}(A-B)$（或 $\operatorname{int}A$, B）可用闭超平面严格分离. 如果此外还有 $0\notin\overline{(A-B)}$，则 A、B 可用闭超平面强分离.

局部凸空间：是具有凸零邻域基的拓扑线性空间. 它的拓扑可用一族半范数表示.

局部凸空间中的凸集分离定理：设 A、B 为局部凸空间 X 中的凸集，$riA\neq\varnothing$，$riB\neq\varnothing$，且

$$riA\cap riB=\varnothing.$$

则 A、B 可用闭超平面分离，riA、riB 可用闭超平面严格分离.

分离局部凸空间中的凸集强分离定理：设 K 为分离局部凸空间 X 中的闭凸集，C 为 X 中的紧凸集，$K\cap C=\varnothing$. 则存在闭超平面，使 K、C 强分离.

分离局部凸空间中的集合的闭凸包：设 K 为分离局部凸空间 X 中的任意集合，

$$\sigma_K(x^*)=\sup_{x\in K}\langle x^*,\ x\rangle$$

为其承托函数. 则 K 的闭凸包为

$$\overline{co}\,K=\{x\in X\,|\,\langle x^*,\ x\rangle\leqslant\sigma_K(x^*),\ \forall x^*\in X^*\}.$$

对偶系 $(X,\ X^*)$：设 X、X^* 为两个线性空间，$\langle\cdot,\ \cdot\rangle$: $X^*\times X\to\mathbf{R}$ 为双线性形式. 如果

$$\forall x_1、x_2\in X,\ x_1\neq x_2,\ \exists x^*\in X^*,\ \langle x^*,\ x_1\rangle\neq\langle x^*,\ x_2\rangle;$$

$$\forall x_1^*、x_2^*\in X^*,\ x_1^*\neq x_2^*,\ \exists x\in X,\ \langle x_1^*,\ x\rangle\neq\langle x_2^*,\ x\rangle.$$

则 $(X,\ X^*)$ 称为对偶系. 当 X 为分离局部凸空间，X^* 为其拓扑对偶，$(X,\ X^*)$ 满足对偶系的条件.

极化拓扑：设 $(X,\ X^*)$ 为对偶系，$\{A_i^*\}_{i\in I}\subset X^*$ 为 X^* 中的子集族，且满足

$$\bigcup_{\lambda>0}\bigcup_{i\in I}\lambda A_i^*=X^*,$$

则由
$$p_i(x)=\max_{x^*\in A_i^*}\langle x^*,\ x\rangle,\ i\in I,$$

定义的半范数族 $\{p_i\}_{i\in I}$ 所决定的 X 的局部凸拓扑，称为 X 的极化拓扑.

弱拓扑或 $\sigma(X,\ X^*)$-拓扑：$\{A_i^*\}$ 为所有 X^* 中的对称有限

非零点集(可只取两个点)时的 X 的极化拓扑.

$\sigma(X^*, X)$-拓扑: 对换 X、X^* 的地位后在 X^* 上定义的弱拓扑.

强拓扑或 $\beta(X, X^*)$-拓扑: $\{A_i^*\}$为所有 X^* 中的对称 $\sigma(X^*, X)$-有界集时的 X 的极化拓扑.

Mackey 拓扑或 $\tau(X, X^*)$-拓扑: $\{A_i^*\}$ 为所有 X^* 中的对称 $\sigma(X^*, X)$-紧集时的 X 的极化拓扑.

Mackey-Arens 定理: 任何使 X 的拓扑对偶为 X^* 的局部凸拓扑一定在 $\sigma(X, X^*)$和 $\tau(X, X^*)$之间.

桶和桶形空间: 局部凸空间 X 的吸收对称闭凸集称为 X 的桶. 当 X 的所有桶都是零邻域时, X 称为桶形空间. 这时设 X 分离, 且 X 的拓扑对偶为 X^* 时, X 上的拓扑一定是强拓扑 $\beta(X, X^*)$, 从而必定有

$$\beta(X, X^*)=\tau(X, X^*).$$

Fréchet 空间、Banach 空间都是桶形空间.

有界性: 设 A 为拓扑线性空间 X 中的集合. 如果对于任何 X 的零邻域 U, 存在 $\lambda>0$, 使得 $\lambda A\subset U$, 则称 A 为 X 的有界集. 当(X, X^*)为对偶系时, $\sigma(X, X^*)$ 有界与 $\tau(X, X^*)$有界是一致的(Mackey 定理). 当 X 为分离桶形空间时, X^* 中的 $\sigma(X^*, X)$有界与$\beta(X^*, X)$有界也是一致的(Banach-Steinhaus 定理).

单变量凸函数: 函数 $f:(a, b)\to\mathbf{R}$ 称为凸函数, 是指

$$\forall x_1, x_2\in(a, b),\ \forall\lambda\in[0, 1],$$

$$f((1-\lambda)x_1+\lambda x_2)\leqslant(1-\lambda)f(x_1)+\lambda f(x_2).$$

其充要条件为:

i) $F(x, y)=\dfrac{f(x)-f(y)}{x-y}$, $x\neq y$ 作为 x 或 y 的函数在(a, b)上不减;

ii) 左、右导数 f'_-、f'_+ 处处存在, 且

$$\forall x_1、x_2\in(a, b),\ x_1<x_2,$$

$$f'_-(x_1)\leqslant f'_+(x_1)\leqslant\frac{f(x_2)-f(x_1)}{x_2-x_1}$$
$$\leqslant f'_-(x_2)\leqslant f'_+(x_2).$$

当 f 可导时，凸 $\Leftrightarrow f'$ 不减.

当 f 二次可导时，凸 $\Leftrightarrow \forall x\in(a, b), f''(x)\geqslant 0$.

线性空间上的凸函数：设 X 为线性空间，$f: X\to\mathbf{R}\cup\{\pm\infty\}$.

$$\operatorname{epi}f=\{(x, \alpha)\in X\times\mathbf{R}\,|\,f(x)\leqslant\alpha\},$$

则 f 称为 X 上的凸函数，是指 $\operatorname{epi}f$ 为 $X\times\mathbf{R}$ 中的凸集. 它也可定义为

$$\forall x_1, x_2\in X, \forall\lambda\in[0, 1],$$
$$f((1-\lambda)x_1+\lambda x_2)\leqslant(1-\lambda)f(x_1)+\lambda f(x_2),$$

其中当右端为 $\infty-\infty$ 时，认为 $\infty-\infty=\infty$.

$$\operatorname{dom}f=\{x\in X\,|\,f(x)<\infty\}$$

称为 f 的有效域. 不取 $-\infty$，且 $\operatorname{dom}f\neq\varnothing$ 的凸函数称为真凸函数.

Hahn-Banach 定理：设 L 为线性空间 X 的子空间，f 是 X 上的真凸函数，且

$$(\operatorname{dom}f)^{ri}\cap L\neq\varnothing.$$

如果 L 上的线性形式 $x_L^*\in L^*$，满足

$$\forall x\in L, \langle x_L^*, x\rangle\leqslant f(x),$$

那么存在 $x^*\in X^*$，满足

$$\forall x\in L, \langle x_L^*, x\rangle=\langle x^*, x\rangle;$$
$$\forall x\in X, \langle x^*, x\rangle\leqslant f(x).$$

次线性函数：$f: X\to\mathbf{R}\cup\{+\infty\}$ 满足：i) $\forall x\in X, \forall t\geqslant 0, f(tx)=tf(x)$；ii) $\forall x_1, x_2\in X, f(x_1+x_2)\leqslant f(x_1)+f(x_2)$. 就称为次线性函数. 当 $0\in(\operatorname{dom}f)^{ri}$ 时，存在 $A^*\subset X^*$，使得

$$\forall x\in X, f(x)=\sup_{x^*\in A^*}\langle x^*, x\rangle.$$

Minkowski 函数：非负的次线性函数. 设

$$p: X\to\mathbf{R}_+\cup\{+\infty\}$$

为 Minkowski 函数，则对于满足下列条件的凸集 A

$$\{x \mid p(x)<1\} \subset A \subset \{x \mid p(x) \leqslant 1\},$$

有

$$p(x)=p_A(x)=\inf\{\alpha>0 \mid [0,\ x/\alpha] \subset A\}.$$

反之，如果凸集 A 满足 $0 \in A^{ri}$，则上式定义的 p_A 是 Minkowski 函数，且

$$A^{ri}=\{x \mid p_A(x)<1\},$$
$$A^c=\{x \mid p_A(x) \leqslant 1\}.$$

不取 $+\infty$ 的 Minkowski 函数称为**规范函数**．对称的规范函数称为**半范数**．仅在原点为 0 的半范数称为**范数**．用范数和半范数族都可定义线性空间上的拓扑．它们分别形成赋范空间和局部凸空间．

凸函数的连续性：设 f 是拓扑线性空间 X 上的真凸函数．如果 f 具有上有界的点，则 f 在 $\mathrm{int}(\mathrm{dom}f)$ 中连续．当 X 为赋范空间时，f 在 $\mathrm{int}(\mathrm{dom}f)$ 中还是局部 Lipschitz 函数．当 X 为桶形空间(包括 Fréchet 空间、Banach 空间等)时，f 在 X 上下半连续是 f 在 $\mathrm{int}(\mathrm{dom}f)$ 中连续的充分条件．此外，f 具有有限连续点等价于 $\mathrm{int}(\mathrm{epi}f) \neq \varnothing$．

凸函数的下半连续性：设 f 是拓扑线性空间 X 上的真凸函数．那么 f 下半连续 $\Leftrightarrow \mathrm{epi}f$ 是闭集．如果 f 是有有限连续点的下半连续真凸函数，那么 f 是连续仿射函数族的上包络．如果 X 是分离局部凸空间，或 $(X,\ X^*)$ 是对偶系，那么 f 下半连续凸 $\Leftrightarrow f$ 是连续仿射函数族的上包络 $\Leftrightarrow f=f^{**}$，这时，f 的下半连续性可对 $\sigma(X,\ X^*)$ 和 $\tau(X,\ X^*)$ 之间的任何拓扑而言．

共轭函数：设 $(X,\ X^*)$ 为对偶系，f: $X \to \mathbf{R} \cup \{\pm\infty\}$ 为任意函数．

$$f^*:\ X^* \to \mathbf{R} \cup \{\pm\infty\},$$
$$f^*(x^*)=\sup_{x \in X}\{\langle x^*,\ x\rangle - f(x)\}$$

称为 f 的**共轭函数**．

$$f^{**}:\ X \to \mathbf{R} \cup \{\pm\infty\},$$
$$f^{**}(x)=\sup_{x^* \in X^*}\{\langle x^*,\ x\rangle - f^*(x^*)\}$$

称为f的二次共轭函数. 一般地总有

$$f(x)\geqslant f^{**}(x).$$

$f=f^{**}\Leftrightarrow f$ 下半连续凸.

凸函数的次微分: 设 X 为拓扑线性空间. X^* 为它的拓扑对偶. $f: X\to\mathbf{R}\cup\{\pm\infty\}$ 为任意函数. f 在 $x\in X$ 上的次微分定义为

$$\partial f(x)=\{x^*\in X^*|\langle x^*, y-x\rangle\leqslant f(y)-f(x),\ \forall y\in X\}.$$

这个定义当 f 为真凸函数时较有意义. 如果这时 f 在 $x\in\operatorname{dom}f$ 上连续, 那么

$$\partial f(x)=\{x^*\in X^*|\langle x^*, h\rangle\leqslant\delta_+ f(x; h),\ \forall h\in X\},$$

其中
$$\delta_+ f(x; h)=\lim_{\substack{t\to0\\t>0}}\frac{f(x+th)-f(x)}{t}=\inf_{t>0}\frac{f(x+th)-f(x)}{t},$$

它称为 f 在 x 上的关于方向 h 的右方向导数, 且一定是 h 的连续次线性函数. 在这种情况下, $\partial f(x)$是 $\sigma(X^*, X)$-紧凸集.

次微分与共轭函数的关系:

$$x^*\in\partial f(x)\Leftrightarrow f(x)+f^*(x^*)=\langle x^*, x\rangle.$$

当 $f(x)=f^{**}(x)$, 还等价于 $x\in\partial f^*(x^*)$.

次微分的性质: i) $\forall\lambda\geqslant0$, $\partial(\lambda f)(x)=\lambda\partial f(x)$; ii) f 在 x_0 上达到最小值$\Leftrightarrow 0\in\partial f(x_0)$; iii) $\partial f_1(x)+\partial f_2(x)\subset\partial(f_1+f_2)(x)$. 当存在 $x_0\in(\operatorname{dom}f_1)\cap(\operatorname{dom}f_2)$, 使 f_1 在 x_0 上连续时,

$$\partial f_1(x)+\partial f_2(x)=\partial(f_1+f_2)(x)$$

(Moreau–Rockafellar 定理 4.2.3). iv) 当凸函数 $f_1, \cdots, f_n$ 都在 x_0 上有限连续时,

$$\partial(\max_{1\leqslant i\leqslant n}f_i)(x_0)=\operatorname{co}\bigcup_{i\in I(x_0)}\partial f_i(x_0),$$

$$I(x_0)=\{i|f_i(x_0)=\max_{1\leqslant i\leqslant n}f_i(x_0)\}.$$

v) 当(X, X^*)为对偶系. $f: X\to\mathbf{R}\cup\{+\infty\}$为真凸函数. 那么 $\partial f: X\to 2^{X^*}$ 关于 $\sigma(X, X^*)$和 $\sigma(X^*, X)$ 的拓扑是上半连续的单调映射.

法向锥和切向锥：设 K 为拓扑线性空间 X 中的凸集，$x\in K$．

$$N_K(x)=\partial\delta_K(x)=\{x^*\in X^*\mid\langle x^*,\ y-x\rangle\leqslant 0,\ \forall y\in K\}$$

称为 K 在 x 的法向锥．

$$T_K(x)=N_K^-(x)=\{x\in X\mid\langle x^*,\ x\rangle\leqslant 0,\ \forall x^*\in N_K(x)\}$$

称为K在 x 的切向锥．当 X 为分离局部凸空间时，

$$T_K(x)=\overline{\bigcup_{\lambda>0}\lambda(K-x)}.$$

这时，如果 $K=\{x\in X\mid g(x)\leqslant 0\}$，$g$ 为 X 上的真凸函数，存在 $x_1\in X$，$g(x_1)<0$，且在 $x\in X$ 上有限连续，那么

$$N_K(x)=\begin{cases}\bigcup\limits_{\lambda\geqslant 0}\lambda\partial g(x), & \text{当 } g(x)=0;\\ \{0\}, & \text{当 } g(x)<0.\end{cases}$$

数学规划的 Lagrange 函数和 Lagrange 乘子：设 X 为线性空间．$f,\ g_i,\ h_j(i=1,\ \cdots,\ p;\ j=1,\ \cdots,\ q)$：$X\to\mathbf{R}\cup\{\pm\infty\}$为任意函数，考虑下列数学规划：

$$(\mathscr{P})\qquad\begin{cases}f(x)\to\min,\\ g_i(x)\leqslant 0, & i=1,\ \cdots,\ p;\\ h_j(x)=0, & j=1,\ \cdots,\ q.\end{cases}$$

则

$$L:\ X\times\mathbf{R}_+^{p*}\times\mathbf{R}^{q*},$$

$$\begin{aligned}L(x,\ \lambda,\ \mu)&=f(x)+\langle\lambda,\ g(x)\rangle+\langle\mu,\ h(x)\rangle\\&=f(x)+\sum_{i=1}^{p}\lambda_i g_i(x)+\sum_{j=1}^{q}\mu_j h_j(x)\end{aligned}$$

称为问题($\mathscr{P}$)的 Lagrange 函数．问题($\mathscr{P}$)这时等价于：

$$(\mathscr{P})\quad\sup_{\lambda,\mu}L(x,\ \lambda,\ \mu)\to\min,\ x\in X.$$

如果存在

$$\hat\lambda=(\hat\lambda_1,\ \cdots,\ \hat\lambda_p)\in\mathbf{R}_+^{p*},$$
$$\hat\mu=(\hat\mu_1,\ \cdots,\ \hat\mu_q)\in\mathbf{R}^{q*},$$

使得问题($\mathscr{P}$)可转化为

$$L(x,\ \hat\lambda,\ \hat\mu)=f(x)+\sum_{i=1}^{p}\hat\lambda_i g_i(x)+\sum_{j=1}^{q}\hat\mu_j h_j(x)\to\min,\ \ x\in X.$$

那么($\hat\lambda,\ \hat\mu$)称为($\mathscr{P}$)的 Lagrange 乘子．$\hat x$ 是($\mathscr{P}$)的解，($\hat\lambda,\ \hat\mu$)

是($\mathscr{P}$)的 Lagrange 乘子 $\Leftrightarrow(\hat{x}, \hat{\lambda}, \hat{\mu})$是 $L(x, \lambda, \mu)$的鞍点，即

$$L(\hat{x}, \lambda, \mu)\leqslant L(\hat{x}, \hat{\lambda}, \hat{\mu})\leqslant L(x, \hat{\lambda}, \hat{\mu}),$$
$$\forall x\in X,\ \forall\lambda\in\mathbf{R}_+^{p*},\ \forall\mu\in\mathbf{R}^{q*}.$$

通常凸规划的 Lagrange 乘子法则：当 $f, g_i, \cdots, g_p$ 都是 X 上的真凸函数，$h_1, \cdots, h_q$ 都是 X 上的仿射函数时，上述($\mathscr{P}$)称为凸规划．这时有下列结果：

i) 代数结果：如果

$$V=\inf_{\substack{g(x)\leqslant 0\\ h(x)=0}} f(x)\neq\pm\infty,$$

那么存在不全为零的 $\hat{\lambda}_0, \hat{\lambda}_1, \cdots, \hat{\lambda}_p\geqslant 0$, $\hat{\mu}_1, \cdots, \hat{\mu}_q\in\mathbf{R}$，使得

$$\hat{\lambda}_0 V\leqslant\hat{\lambda}_0 f(x)+\sum_{i=1}^{p}\hat{\lambda}_i g_i(x)+\sum_{j=1}^{q}\hat{\mu}_j h_j(x),$$
$$\forall x\in\operatorname{dom} f\cap\Big(\bigcap_{i=1}^{p}\operatorname{dom} g_i\Big).$$

如果这时还存在 $x_0\in\operatorname{dom} f$，使得

$$g_i(x_0)<0\ \ (i=1, \cdots, p),\ h_j(x_0)=0(j=1, \cdots, q)$$

且
$$0\in\Big(h\Big(\operatorname{dom} f\cap\Big(\bigcap_{i=1}^{p}\operatorname{dom} g_i\Big)\Big)\Big)^{ri},$$

那么上式中 $\hat{\lambda}_0$ 可取为 1；特别是，当

$$\bigcap_{i=1}^{p}\operatorname{dom} g_i=X$$

时，($\mathscr{P}$) 的 Lagrange 乘子存在．

ii) 拓扑结果：当 X 为拓扑线性空间时，如果 $f, g_1, \cdots, g_p$, $h_1, \cdots, h_q$ 都在 $\hat{x}\in X$ 上连续，且

$$h(\hat{x})=(h_1(\hat{x}), \cdots, h_q(\hat{x}))=0,$$

那么 $\hat{x}$ 是($\mathscr{P}$)的解等价于存在 $\hat{\lambda}_0\geqslant 0$, $\hat{\lambda}_1, \cdots, \hat{\lambda}_p\geqslant 0$，且不全为零，以及 $\hat{\mu}_1, \cdots, \hat{\mu}_q\in\mathbf{R}$，使得

$$0\in\hat{\lambda}_0\partial f(\hat{x})+\sum_{i=1}^{p}\hat{\lambda}_i\partial g_i(\hat{x})+\sum_{j=1}^{q}\hat{\mu}_j\partial h_j(\hat{x}),$$
$$\hat{\lambda}_i g_i(\hat{x})=0,\ (i=1, \cdots, p).$$

如果还存在 $x_0\in X$，使得

$$g_i(x_0)<0,\ (i=1, \cdots, p);$$

$$h_j(x_0)=0 \quad (j=1, \cdots, q),$$

那么上面的 $\hat{\lambda}_0$ 可取为 1. 这时

$$(\hat{\lambda}, \hat{\mu})=(\hat{\lambda}_1, \cdots, \hat{\lambda}_p; \hat{\mu}_1, \cdots, \hat{\mu}_q)$$

即 $(\mathscr{P})$ 的 Lagrange 乘子.

凸规划的一般对偶理论: 设 X、Y 为两个分离局部凸空间, X^*、Y^* 为它们的拓扑对偶.

$$F: X\times Y \to \mathbf{R}\cup\{\pm\infty\},$$

$$F^*: X^*\times Y^* \to \mathbf{R}\cup\{\pm\infty\}$$

为其共轭函数. $A: X\to Y$ 为线性映射. 极值问题

$$(\mathscr{F})\begin{cases} F(x, Ax)\to \min \\ x\in X \end{cases}$$

的对偶问题是指

$$(\mathscr{F}^*)\qquad\begin{cases} -F^*(-A^*y^*, y^*)\to \max \\ y^*\in Y^* \end{cases}$$

这里 A^* 为 A 的共轭映射;

二次对偶问题是指

$$(\mathscr{F}^{**})\begin{cases} F^{**}(x, Ax)\to \min \\ x\in X \end{cases}$$

当 F 在 $X\times Y$ 上下半连续真凸时, $(\mathscr{F})=(\mathscr{F}^{**})$.

如果存在 $\hat{y}^*$, 满足

$$-F^*(-A^*\hat{y}^*, \hat{y}^*)=\inf_{x\in X} F(x, Ax),$$

则 $\hat{y}^*$ 称为 $(\mathscr{F})$ 的 Lagrange 乘子. $\hat{x}$ 是 $(\mathscr{F})$ 的解, $\hat{y}^*$ 是 $(\mathscr{F})$ 的 Lagrange 乘子 $\Leftrightarrow(-A^*\hat{y}^*, \hat{y}^*)\in\partial F(\hat{x}, A\hat{x})$ (Euler-Lagrange 方程). 当 X, Y 有限维(一般可为自反 Banach 空间)时, 下列条件是 $(\hat{x}, \hat{y}^*)$ 存在的充分条件:

$$0\in\operatorname{int}(A\oplus(-1))\operatorname{dom}F, \quad 0\in\operatorname{int}(1\oplus A^*)\operatorname{dom}F^*.$$

这里

$$(A\oplus(-1))(x, y)=Ax-y,$$

$$(1\oplus A^*)(x^*, y^*)=x^*+A^*y^*.$$

$$H(x, y^*)=\sup_{y\in Y}\{\langle y^*, y\rangle-F(x, y)\}$$