“十三五”国家重点出版物出版规划项目

世界名校名家基础教育系列

Textbooks of Base Disciplines from World's Top Universities and Experts

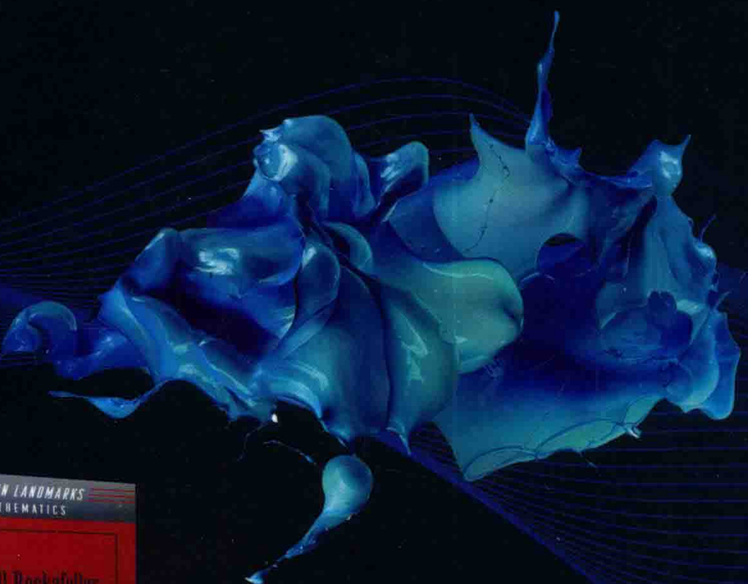
普林斯顿分析译丛

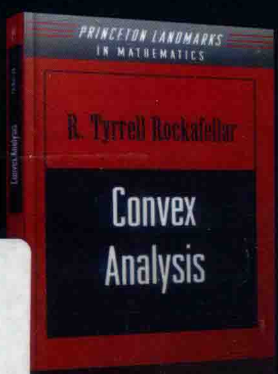
Convex Analysis

凸分析

[美] R.T. 洛克菲勒 (R. T. Rockafellar) 著

盛宝怀 译





机械工业出版社
CHINA MACHINE PRESS

Convex Analysis

凸分析

“十三五”国家重点出版物出版规划项目

世界名校名家基础教育系列
Textbooks of Base Disciplines from World's Top Universities and Experts

普林斯顿分析译丛

凸 分 析

［美］R. T. 洛克菲勒（R. T. Rockafellar） 著
盛宝怀 译

机 械 工 业 出 版 社

这是有关“凸分析”的较早的名著，是对凸分析理论进行系统总结和论述的经典之作，也是学习凸分析理论的必读之书。以“凸分析”为内容的教材、论文、论著，甚至在凸分析教学中的许多概念、内容，或来源于此，或以此为范本。

本书对与凸分析相关的许多概念均进行了严格定义，重点突出了“凸性”，如“凸集”“凸函数”“凸锥”，以及为刻画凸性所需用到的“超平面”“凸集分离”“方向导数”“次梯度”“相对内部”“共轭”“对偶”等。对与“凸性”有关的“Kuhn-Tucker最优性”条件、“鞍点最优性”条件均有详细的论述和证明。书中始终贯穿和应用了凸性是对线性推广的思想。本书是最早出现“多值映射”“凸过程”“双重函数”的著作之一。

本书是基础数学、应用数学、计算数学、计算机科学甚至物理学等学科研究生的理想的凸分析教材，也是从事数学理论和应用研究的科技工作者的经典参考书。

图书在版编目（CIP）数据

凸分析/(美) R. T. 洛克菲勒（R. T. Rockafellar）著；盛宝怀译. —北京：机械工业出版社，2017.11

（世界名校名家基础教育系列）

书名原文：Convex Analysis（Princeton Landmarks in Mathematics and Physics）

“十三五”国家重点出版物出版规划项目

ISBN 978-7-111-58182-6

Ⅰ.①凸…　Ⅱ.①R…②盛…　Ⅲ.①凸分析-高等学校-教材

Ⅳ.①O174.13

中国版本图书馆CIP数据核字（2017）第245099号

机械工业出版社（北京市百万庄大街22号　邮政编码100037）

策划编辑：汤　嘉　责任编辑：汤　嘉　李　乐　任正一

责任校对：刘秀芝　封面设计：张　静

责任印制：孙　炜

保定市中画美凯印刷有限公司印刷

2018年1月第1版第1次印刷

169mm×239mm · 20.5印张 · 2插页 · 420千字

标准书号：ISBN 978-7-111-58182-6

定价：78.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务	网络服务
服务咨询热线：010-88379833	机 工 官 网：www.cmpbook.com
读者购书热线：010-88379649	机 工 官 博：weibo.com/cmp1952
	教育服务网：www.cmpedu.com
封面无防伪标均为盗版	金 书 网：www.golden-book.com

译　者　序

R. T. Rockafellar 的《凸分析》(*Convex Analysis*)是名著，能够有机会将这本书翻译成为中文，我是十分荣幸的。我读书时曾经参阅过这本书，感到对书中有些章节的内容是熟悉的，认为完成这本书的翻译应该不会有问题。然而，当我接受这个任务之后很快就发现，自己其实犯下了一个大错。

作者将其对凸分析的深刻理解、深厚感情都渗透在了字里行间，要想将这种高深的感悟用中文很好地表达出来，我很难做到，因为离这样的境界还相差很远。

凸分析是非线性泛函分析中最艰深的核心内容，作者的学术水平、数学修养远远在我等之上，要想把书中内容用专业的、流畅的、通俗的中文完全表达出来，我自认为我没有这个水平，我的理解也是很难到位的。翻译这本书，我自己其实是很不够格的，因为读过的凸分析书很少，对博大精深的“凸分析”仅仅见过为数不多的几个概念，书中的多数内容、结果对我来说都是新的，需要从头学起。

书中有些概念在国内现有的著作、教材中叫法不一，有些概念尚未翻译过来，要想对这些概念给出中文定义，感到自己的经验不够，只好结合内容来理解或直译。

由于上述原因，我十分惭愧地告诉读者，我的翻译不一定是成功的。由于是边学、边翻译，所以译稿中肯定存在许多问题甚至错误，希望读者谅解并提出批评。

译　者

前　言

近年来，凸性在应用数学领域有关极值问题的研究中所发挥的作用越来越重要。本书是有关凸集和凸函数理论的系统阐述，这些理论在极值问题的研究中发挥着中心作用。不等式系统，定义在凸集上的凸函数的最大值或最小值、Lagrange乘子、极小极大定理以及有关凸集的结构、凸函数与鞍函数的连续性和可微性的基本结果，均是本书所要涉猎的内容。全书均涉及对偶性，特别地，要涉及关于凸函数Fenchel型共轭的相关理论。

书中的许多材料是以前没有出版过的。例如，给出了一种推广的线性代数，按照此理论，“凸双重函数”为线性变换的类似物，凸集的“内积”以及函数用Fenchel型对偶定理中的极值来定义。每个凸双重函数均与广义凸规划相联系，引入了一种有关双重函数的伴随运算，由其产生了一种对偶规划理论。线性变换和双线性泛函之间的所有经典结论均被推广到凸双重函数和鞍函数的情形，并且在鞍函数和极小极大问题的分析中作为主要工具。

不动点定理等一些可被看作正常凸分析的专题被去掉了，并非这些内容缺少吸引力和应用性，而是因为它们需要一些技术改进，这些技术从某种程度上超出了本书的内容。

考虑到不仅仅纯数学家，而且经济学家、工程师等其他领域的专家已经对凸分析有兴趣，我们尽最大努力，使书的内容保持在基础知识的范围，并且提供了细节，这些细节，如果仅限制在数学圈子的作品中是可接受的，仅仅被作为“练习”来处理。一些讨论，如实数 n 元组空间，甚至许多能够在更广泛的环境下成立的结果，都限制在欧氏空间 $\mathbf{R}^n$ 中。在注释与参考中收集了一些有关改进和推广的文献，这部分内容放在参考文献的前面，两者都安排在书的最后。

关于预备知识，我们要求读者应该能够至少具有良好的线性代数和基础实分析（收敛序列、连续函数、开集和闭集、紧性等）基础，$\mathbf{R}^n$ 空间的知识也不可缺。虽然与较深的抽象数学分支没有可比性，但是从读者的角度来讲，书的风格的确试图表达数学的某些缜密性。

书的开头安排了导读，对每部分的内容和取材进行了描述，可以看成是对每节主题的引言。

本书来源于1966年春季我在普林斯顿大学所授课程的讲稿。这份讲稿在很大程度上来源于哥本哈根大学的Werner Fenchel教授15年前在普林斯顿大学所授类似课程的手稿。Fenchel的手稿从未出版，但是，以油印本的方式传阅，作为主要且本质上为唯一的有关凸函数理论的文献，在这漫长的时间里惠及了许多研究者。

这极大地影响到我的思想，这方面的一个例子就是共轭凸函数占据了书的大部分。因此，将本书以荣誉合著者的形式奉献给 Fenchel 是十分合适的。

我十分希望表达我对普林斯顿大学 A. W. Tucker 教授的深深谢意，自从学生时代起，他的鼓励和支持就已经成为我的精神支柱。事实上，就是按 Tucker 的建议给出了本书的书名。进一步还要感谢 Torrence D. Parsons 博士、Norman Z. Shapiro 博士和 Lynn McLinden 先生，他们仔细阅读了书稿并提出了十分有用的建议。我也要感谢我在普林斯顿大学和华盛顿大学的学生们，当书稿在教学中使用时，他们的建议使书稿在许多表达方面得到了改进。同时感谢 Janet Parker 夫人耐心称职的秘书工作。

本书的初稿为 1966 年在普林斯顿的演讲笔记，当时得到美国海军研究办公室基金 NONR1858（21）基金项目 NR-047-002 的资助。随后，空军科学研究局在华盛顿大学以基金 AF-AFOSR-1202-67 的形式给予了热情的资助。如果没有这些资助，本书的书写工作也许会延缓、间断。

R. T. 洛克菲勒

目　　录

写在前面：导读

本书并不要求读者从头到尾进行阅读，即使有人可能有这样足够的雄心。相反，素材是尽可能按照章节主题来组织的，例如，关于凸集相对内部的所有相关内容，不管其重要性如何，均收集到一块（第 6 节），而不是随着内容的进展而出现在这里、那里的。这种组织方式方便查找基本结论，至少当对主题有所熟悉之后是这样的，但是妨碍了初学者将本书作为入门读物。书中内容按照逻辑关系展开，但是，前几节的末尾多处缺少细节，可能使读者感到困难。

然而，如果能够对材料进行适当的取舍，本书能够被用作入门读物。导读的内容告知读者哪些内容是十分必要的，哪些内容可以略过不看，至少暂时不看并不会影响对内容和证明的理解。

第 1 部分：基 本 概 念

定义了凸集和凸函数的概念，并且讨论两种概念之间的关系。重点在于凸性的判别。给出了多个有用的例子，并且论述如何由这些例子以及求和与取凸包运算产生新的例子。基本思想是 $\mathbf{R}^n$ 上定义的凸函数能够被 $\mathbf{R}^{n+1}$ 中的某些凸子集［它们的上镜图（epigraph）］所确定，同时，$\mathbf{R}^n$ 中的凸集能够被某些定义在 $\mathbf{R}^n$ 上的凸函数［它们的指示函数（indicator）］所确定。这些方法使得几何方法和解析方法之间的转换变得容易。在处理函数时，通常会想到函数的图形，但是，在处理凸函数时，应该牢记凸函数的上镜图。

书中大部分素材虽然很初级，对于本书其余部分来说是基础性的，但是，部分内容对于首次接触此课题的读者来说应被忽略。虽然在第 1 节（仿射集）中用到了线性代数，但是，概念并不是完全类似的；第 1 节应该以重心坐标系的定义（在定理 1.6 之前的内容）作为引入凸性的背景来仔细阅读。第 1 节的其余部分涉及仿射变换（affine transformation），对于初学者来说，并不是理解的关键。整个第 2 节（凸集与锥）和第 3 节的前半部分是重要的，但是第 3 节的后半部分，从定理 3.5 开始，处理一些不太重要的运算。第 4 节（凸函数）除去一些例子以外均不应该跳过。但是，第 5 节（函数运算）中定理 5.7 之后的内容，在后续的各节中是不需要的。

第 2 部分：拓 扑 性 质

第 1 部分所考虑的凸性性质是基本代数：凸集和函数所形成的类在多次组合与

扩张运算下保持不变。第 2 部分反倒考虑了与内部、闭包及连续的拓扑概念相关联的凸性。

凸集和凸函数的拓扑性质虽然著名，但并不复杂，它能够追溯到一个直观事实：在凸集 C 中的线段，一个端点在 C 的内部，另外一个端点位于 C 的边界，则 C 的所有中间的点都将位于 C 的内部。能够引入“相对”内部的概念，以便这个事实能够作为基本工具用来处理内部为空集的状态的情形。这在第 6 节（凸集的相对内部）中得到讨论。每位学生都应该知道的凸性的主要结论被嵌入在第 6 节的前四个定理中。第 6 节的其余部分，从定理 6.5 开始，主要贡献在于采用不同方法由其他凸集所构造的凸集给出了相对内部的公式，得到了许多有用的结论（特别如推论 6.5.1 和推论 6.5.2，它们在本书里经常被引用，以及推论 6.6.2，在第 11 节被用来证明一个重要的分离定理），但是，这些都能够被暂时忽略，并在需要时再提起。

第 7 节（凸函数的闭包）的主题是下半连续。在凸函数的情况下，这个性质在许多方面比连续性重要，因为它与上镜图相关联：函数下半连续当且仅当其上镜图为闭的。不为下半连续的凸函数能够（以唯一确定的方式）简单地通过重新定义其在有效域的某些边界上的值而使其成为下半连续的。这便导致了凸函数的闭包运算的概念，当函数为正常函数时对应于上镜图（作为 $\mathbf{R}^{n+1}$ 的子集）的闭包运算。要想理解后续内容，除定理 7.6 之外，第 7 节的所有内容都是基本的。

所有第 8 节（回收锥及其无界性）的内容，虽然不如第 6 节和第 7 节的内容那样普遍需要，但是在整个书中都是需要的。第 8 节的前半部分阐明一种思想：除去某些“无穷远的点”之后，无界凸集与有界凸集是一样的。第 8 节的后半部分将这种思想应用于上镜图而获得关于凸函数增长性的结论。这些性质在形成若干散布在整本书中的基本存在性定理方面是重要的，第一个存在性定理出现在第 9 节（闭性准则）中。

第 9 节试图回答的问题是：什么时候闭集在线性变换下的像为闭的？这个问题在各种极值问题解的存在性方面是基本的。第 9 节的主要结论见定理 9.1 和定理 9.2（及其推论）。然而，读者在首次遇到第 9 节时跳过去，以后再回头阅读也是行得通的，如果需要可与第 16 节的内容联系起来。

就总体而言，第 10 节（凸函数的连续性）中仅仅第一个定理对于凸分析来说是基本的。更加引人注目的连续和收敛性定理是一个重点。它们仅在第 24 节和第 25 节中被用来获得关于凸函数梯度映射和次微分的连续性和收敛性定理，并且在鞍函数的情况下，在第 35 节获得类似的结论。

第 3 部分：对 偶 对 应

点与超平面之间的对偶在分析学中发挥着重要作用，但是，也许在凸分析中所发挥

的作用更为人所知。从几何的观点来说，凸性理论中的对偶基是指凸集是所有包含它的闭半空间的交。然而，从函数的观点来看，闭凸函数为所有略小于其的仿射函数的点态上确界。当用上镜图表达时这两个事实为等价的，从直觉的方面来考虑几何表达通常更可取，但是，在这种情况下两种表示都是重要的。对偶基的第二种表示的优点在于导致了 Fenchel 共轭对应这样的有关凸函数之间对称的一一对应的对偶对应。

从某种意义上说，共轭性是一种特殊情况，包含了闭凸锥（极）之间的对称的一一对应，但是，一般的闭凸集类中没有对称结果。在后文中，既有凸集，也有正齐次凸函数（它们的支撑函数）之间的类似对应。由于这个原因，在牵扯到对偶的应用中，用凸函数，而不是用凸集表达给定情形通常较好。当然，一旦这样做，几何推理仍能够应用于上镜图。

对偶理论的基础放在第 11 节（分离定理）。除定理 11.7 之外，本节的所有材料都是基本的。在第 12 节（凸函数的共轭）中，定义了共轭对应，并且给出了许多共轭对应函数的例子。定理 12.1 和定理 12.2 为基础性结论，应该知道；第 12 节的其余部分不是必要的。

共轭性在第 13 节（支撑函数）中得到了应用，得到了有关凸集和正齐次凸函数之间的对偶性。有效域的支撑函数以及凸函数的水平集用共轭函数 f^* 及其回收函数来计算。主要事实叙述在定理 13.2、定理 13.3 和定理 13.5 中，后面的两个定理需要预先熟悉第 8 节中的内容，其余定理及所有推论可以跳过，如果需要时再参考。

在第 14 节（凸集的极）中，凸函数的共轭对应特别针对凸锥的极对应，于是后者被推广到了关于含有原点的任意闭凸集的极对应的情况。在本书的其他地方有几次应用了凸锥的极，但是，除去第 15 节（凸函数的极）中讨论了它与范数理论的关系之外，并没有提及更一般的极。除得到了有关范数的 Minkowski 对偶性对应及其某些推广之外，第 15 节（凸函数的极）的目的是进一步给出（在定理 15.3 和推论 15.3.1 中）共轭凸函数的例子。然而，第 14 节和第 15 节的内容中，只要不是特意对逼近问题感兴趣，仅仅阅读定理 14.1 就够了。

第 16 节（对偶运算）中的定理说明第 15 节中的各种函数运算都变成有关共轭对应的对偶对。最重要的结果为定理 16.4，它刻画了凸函数的加法和卷积下确界。这个结果对于不等式系统（第 21 节）、次梯度法（第 23 节）有重要的影响，因此对于第 6 部分中极值问题的相关定理也有重要的作用。定理 16.3 的后半部分、定理 16.4 及定理 16.5（这些定理给出了使各自最小值取到的条件且在对偶公式中不需要闭包运算）依赖于第 9 节。这些材料，连同引理 16.2 及其推论都可以在首次阅读第 16 节时跳过。

第 4 部分：表述与不等式

本部分的目的是得到有关作为点集和方向凸包的凸集的表达式，并将这些结果

应用于线性和非线性不等式系统的研究。大部分材料牵扯到凸性理论的加细，利用了维数及某种程度的线性性。读者可以跳过第 4 部分，而不会对本书剩余部分的理解造成影响。或者，作为折中，按照下面所指出的，仅仅阅读第 4 部分中比较基础的内容。

第 17 节（Carathéodory 定理）探讨了维数在凸包产生中的作用，主要结论在定理 17.1 和 定理 17.2 中给出。第 18 节（极值点与凸集的面）考虑用极点、暴露点、端方向、暴露方向以及切超平面表示给定凸集的问题。整个第 18 节都在第 19 节（多面体凸集与函数）中得到了应用；在梯度和凸函数的最大值（第 32 节）的研究中也得到应用。第 19 节中最主要的结论为定理 19.1、定理 19.2 和定理 19.3 及其推论。

第 20 节（多面体凸性的应用）阐述凸分析中的一些通常定理在有些，但非全部，涉及凸集或函数为多面体的情况下是如何得到加强的。定理 20.1 和定理 20.2 在第 21 节中被用以建立 Helly 定理以及在第 27 节和第 28 节中应用于 Lagrange 乘子的存在性和凸规划的最优解的某些其他相对困难的结果的加细。虽然定理 20.2 不依赖于第 9 节，但定理 20.1 却要依赖第 9 节。然而，在不需要第 20 节，甚至第 18 节和第 19 节中知识的情况下理解第 21 节（Helly 定理与不等式系统）中的基本结论及证明是可能的。在这种情况下，可以简单地略去定理 21.2、定理 21.4 和定理 21.5。

方程和线性不等式的有限系统，不管是弱的或严格的，都是第 22 节（线性不等式）中的内容。结论是特殊的，在书的其他任何地方都没有再被提及。开头陈述了第 21 节中定理的许多推论，然后，借助于仅仅用到线性代数和非凸性理论的基本而独立的方法得到了同样的结果，并有所改进。

第 5 部分：微 分 理 论

当经典光滑表面意义下的切超平面不存在时，可以借助有关凸集的支撑超平面。同样，在通常梯度不存在的地方，关于凸集的次梯度常常有用，这种次梯度是基于上镜图，而不是基于图的切超平面。

第 23 节（方向导数与次梯度）中详细说明的有关凸函数的次微分理论在极值问题的分析方面是一种基本的工具，首先应该掌握。定理 23.6、定理 23.7、定理 23.9，以及定理 23.10 都可以省略，但是，在非多面体的情况下，读者应该肯定地知道定理 23.8，关于它有另外一个更基本的证明。第 23 节中的大部分内容与第 4 部分无关。

定理 25.1 的主要结论是有关凸函数的次梯度与通常梯度之间的关系，可以在第 23 节之后就紧跟着阅读。除在第 35 节中给出了鞍函数定理的证明之外，书中其他部分对第 24 节、第 25 节或第 26 节的其他结果均无特殊需要。第 5 部分剩余部

分的内容自给自足。

第 24 节（微分的连续性和单调性）中建立了有关闭正常单变量凸函数左右导数的基本理论。证明这样的函数的次微分的图可以被完整地刻画为“完备的非减曲线”。一维情形下的单调性因此而被推广到了 n 维情形。除去上面已经提到的定理之外，第 25 节（凸函数的可微性）的主要贡献是证明定义在开集上的有限凸函数是连续的，并且几乎处处存在通常的梯度映射。诸如在什么情况下梯度映射构成整个次微分映射，以及在什么时候是一一对应的问题在第 26 节（Legendre 变换）中得以考虑。第 26 节的主要目的是解释通过转化梯度使得共轭凸函数能够以经典方式进行计算，也讨论了光滑性和严格凸性之间的对偶性。从某种程度上讲，第 25 节和第 26 节中的讨论在某种程度上依赖于第 4 部分第 18 节，但不依赖于第 18 节以后各节。

第 6 部分：约束极值问题

当然，极值问题理论是本书中许多结论的动力和源泉。关于此理论系统的应用从第 27 节（凸函数的最小值）开始。这个阶段由定理 27.1 所建立，定理概括了前面几节中所证明的一些相关事实。第 27 节中的所有定理都涉及使得凸函数相对一个给定的凸集而取到其最小值的方式，所有这些，也许除去利用多面体凸性加细的内容之外，都为首次阅读的内容。

凸函数在有限个凸不等式约束下的最小值问题将在第 28 节（常见凸规划及 Lagrange 乘子）中进行讨论。重点在存在性、证明、以及被称为 Kuhn-Tucker 向量的某些 Lagrange 乘子向量的特征。内容可以通过删去线性方程约束项而得到简化，定理 28.2 可以用属于其特殊情况的推论 28.2.1（有一种特别简单的证明）所代替，除此之外，没有其他更应该删除的例子了。

Lagrange 乘子理论得以推广，且某些方面在第 29 节（双重函数及广义凸规划）中得到了加强。凸双重函数的概念，能够被看成线性变换的推广，能够被用来构造最小化问题的扰动理论。广义 Kuhn-Tucker 向量刻画了扰动的影响。定理 29.1、定理 29.3 及其推论包含了后续所有的事实。

第 30 节（伴随双重函数及对偶规划）详细论述了极值对偶问题。实际上，定理 30.5 中的一切都是基本的，但是，第 30 节的剩余部分由例子组成，可以根据需要进行删减。第 31 节（Fenchel 对偶定理）中继续讨论对偶理论。第 31 节的主要目的是提供一些有益于应用的例子。后面的章节，除去第 38 节之外，均不依赖于第 31 节中的材料。

第 32 节（凸函数的最大值）描述了相当不同的特征结论。这些结论的证明不涉及第 6 部分以前的各节，但是，必须熟悉第 18 节和第 19 节中的内容。本书后面的内容也不参考第 32 节。

第 7 部分：鞍函数与极小极大理论

鞍函数是指某些变量为凸函数，而其余变量为凹函数的函数，与它们相关的极值问题自然牵扯到“极小极大”，而不是简单的极小或者极大。这种极小极大问题的理论能够与凸函数的最小化情况类似得到。关于鞍函数的一般极小极大问题（适当地正则化）就是与推广的（闭）凸规划相关的 Lagrange 乘子型鞍点问题。因此，可以理解，凸双重函数是鞍函数讨论的核心内容，读者如果对第 29 节和第 30 节中的基本思想不熟悉就不要将其跳跃过去。

定义在 $\mathbf{R}^m \times \mathbf{R}^n$ 上的鞍函数对应于映 $\mathbf{R}^m$ 到 $\mathbf{R}^n$ 的凸双重函数，就像定义在 $\mathbf{R}^m \times \mathbf{R}^n$ 上的双线性函数对应于映 $\mathbf{R}^m$ 到 $\mathbf{R}^n$ 的线性函数一样。这是第 33 节（鞍函数）的本质内容。第 34 节（闭包和等价类）研究了类似于凸函数的有关鞍函数的一些闭包运算。证明了定义在 $\mathbf{R}^m \times \mathbf{R}^n$ 中凸集的乘积上的每个有限鞍函数都唯一确定一个定义在整个 $\mathbf{R}^m \times \mathbf{R}^n$ 上的闭的鞍函数的等价类，但是，在没有热衷于极小极大理论本身以前，没有必要阅读比较靠后的事实（嵌入到定理 34.4 和定理 34.5 中的）。

第 35 节（鞍函数的连续性与可微性）中所证明的有关鞍函数的结论基本是类似的，或是第 10 节、第 24 节和第 25 节中关于凸函数结论的推广，对于后续内容而言，它们并不是先决条件。

第 36 节（极小极大问题）讨论鞍点和鞍值。那时将解释这些研究如何划归为相互对偶的凸凹规划的研究。因此，在第 37 节（共轭鞍函数与极小极大定理）中借助于鞍函数的共轭对应和双重函数的“逆”运算得到鞍点和鞍值存在性定理。

第 8 部分：凸　代　数

凸双重函数和线性变换之间的类似性为第 6 部分和第 7 部分的特色，在第 38 节（双重函数代数）中得到进一步研究。双重函数的“加法”和“乘法”用借助于基于 Fenchel 对偶性定理的广义内积概念来研究。就像在线性代数中引入伴随之后一样，这样的有关双重凸函数的自然运算是值得保留的，并非没有价值。

第 38 节有关双重函数的结论在第 39 节（凸过程）中被特殊化到一类更加类似于线性变换的凸集值映射。

第 1 部分：基本概念

第 1 节 仿 射 集

书中用 $\mathbf{R}$ 表示实数系，$\mathbf{R}^n$ 为由实 n 元组 $\boldsymbol{x}=(\xi_1, \xi_2, \cdots, \xi_n)$ 所组成的向量空间。除非特别说明，认为 $\mathbf{R}^n$ 空间中一切性质均要求成立。两向量 $\boldsymbol{x}$ 与 $\boldsymbol{x}^*$ 在 $\mathbf{R}^n$ 中的内积表示为

$$\langle \boldsymbol{x}, \boldsymbol{x}^* \rangle = \xi_1\xi_1^* + \xi_2\xi_2^* + \cdots + \xi_n\xi_n^*$$

同一符号 $\boldsymbol{A}$ 既用来表示 $m\times n$ 阶实矩阵 $\boldsymbol{A}$，也表示映 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的线性变换 $\boldsymbol{x}\to\boldsymbol{A}\boldsymbol{x}$。转置矩阵和映 $\mathbf{R}^m$ 到 $\mathbf{R}^n$ 的相应伴随线性变换记为 $\boldsymbol{A}^*$。因此，有

$$\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y}^* \rangle = \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y}^* \rangle$$

（在表示向量的符号中，$*$ 没有运算方面的意义；所有向量都认为是列向量，以便进行矩阵乘法。涉及 $*$ 的向量符号常常既用来表示点向量的对偶，也用来表示 n 元线性函数系数向量的对偶）每个证明均以符号 ‖ 表示结束。

如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为 $\mathbf{R}^n$ 中不同的点，形如

$$(1-\lambda)\boldsymbol{x}+\lambda\boldsymbol{y}=\boldsymbol{x}+\lambda(\boldsymbol{y}-\boldsymbol{x}), \lambda\in\mathbf{R}$$

的点集被称为连接 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的直线。对于 $\mathbf{R}^n$ 中的子集 M，如果对于任意 $\boldsymbol{x}$，$\boldsymbol{y}\in M$ 和 $\lambda\in\mathbf{R}$ 都有 $(1-\lambda)\boldsymbol{x}+\lambda\boldsymbol{y}\in M$，则称 M 为 $\mathbf{R}^n$ 中的仿射集（affine set）。[别的书中所用到的与“仿射集”同义的名词有“仿射流形”（affine manifold）“仿射变量”（affine variety）“线性变量”（linear variety）或“平坦的”（flat）]。

空集$\varnothing$和空间本身为仿射集的极端的例子，而且，按照定义，当 M 由孤立点组成时也为仿射集。一般来说，含有两个点的仿射集必须包含通过这两个点的整个直线。直观图片是没有端点的不弯曲的结构，如直线或空间中的平面等。

仿射集的形式几何可以由线性代数中关于 $\mathbf{R}^n$ 中子空间的定理而得到。仿射集和子空间之间的精确对应在随后的定理中有描述。

定理 1.1 $\mathbf{R}^n$ 中含有原点的子空间都为仿射集。

证明 任何含有 $\mathbf{0}$ 且对于加法和数乘运算封闭的子空间都是特殊仿射集。

相反，如果 M 为含有 $\mathbf{0}$ 的仿射集。对于任何 $\boldsymbol{x}\in M$ 和 $\lambda\in\mathbf{R}$ 都有

$$\lambda\boldsymbol{x}=(1-\lambda)\mathbf{0}+\lambda\boldsymbol{x}\in M.$$

因此，在加法运算下 M 为闭的子空间。

如果 $x\in M$ 且 $y\in M$，我们有$\frac{1}{2}(x+y)=\frac{1}{2}x+\left(1-\frac{1}{2}\right)y\in M$，

由此得到

$$x+y=2\left[\frac{1}{2}(x+y)\right]\in M.$$

因此，在加法的运算下 M 也是闭的且为子空间。

对于 $M\subset\mathbf{R}^n$ 及 $\boldsymbol{a}\in\mathbf{R}^n$，$M$ 关于 $\boldsymbol{a}$ 的平移定义为集合

$$M+\boldsymbol{a}=\{\boldsymbol{x}+\boldsymbol{a}\mid\boldsymbol{x}\in M\}.$$

容易证明，仿射集的平移仍为仿射集。

仿射集 M 与仿射集 L 平行是指存在 $\boldsymbol{a}$ 使得 $M=L+\boldsymbol{a}$。显然，M 平行于 L 是定义在 $\mathbf{R}^n$ 中仿射子集的集合上的等价关系。注意到这种平行的定义比通常的定义要严格，因为并不包含线平行于平面的思想。我们必须在一个平面内说一条线平行于另外一条线等。

定理 1.2 每个非空仿射集 M 一定平行于唯一的子空间 L。这个 L 由

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

所给出。

证明 我们先证明 M 不能平行于两个不同的子空间。子空间 L_1 和 L_2 平行于 M，因而相互平行，所以，存在 $\boldsymbol{a}$ 使得 $L_2=L_1+\boldsymbol{a}$。因为 $\boldsymbol{0}\in L_2$，$-\boldsymbol{a}\in L_1$，因此，$\boldsymbol{a}\in L_1$，但此时 $L_1\supset L_1+\boldsymbol{a}=L_2$。类似可证 $L_2\supset L_1$，所以，$L_1=L_2$。这便证明了唯一性。注意到对于任意 $\boldsymbol{y}\in M$，$M-\boldsymbol{y}=M+(-\boldsymbol{y})$ 是 M 的平移，且含有 $\boldsymbol{0}$。由定理 1.1 和刚才所证知道这个仿射集一定为平行于 M 的唯一子空间 L。因为，无论选择哪个 $\boldsymbol{y}$，总有 $L=M-\boldsymbol{y}$，因此，确有 $L=M-M$。

非空仿射集的维数定义为与其平行的子空间的维数。（习惯上，空集$\varnothing$的维数为-1。）自然，维数为 0、1 和 2 的仿射集分别被称为点、线和平面。$\mathbf{R}^n$ 中的 $(n-1)$ 维仿射集被称为超平面。超平面十分重要，因为它们所扮演的角色与点在 n 维几何中的角色相对应。

超平面和其他的仿射集可以用线性函数和线性方程来表示。这可以由 $\mathbf{R}^n$ 中的正交性理论所推导。注意到，由定义知道，$\boldsymbol{x}\perp\boldsymbol{y}$ 指 $\langle\boldsymbol{x},\boldsymbol{y}\rangle=0$。给定 $\mathbf{R}^n$ 中的子空间 L，满足 $\boldsymbol{x}\perp L$，即对于一切 $\boldsymbol{y}\in L$ 均有 $\boldsymbol{x}\perp\boldsymbol{y}$ 的向量 $\boldsymbol{x}$ 的集合称为 L 的正交补 (orthogonal complement)，记为 $L^{\perp}$。当然，其为另外的子空间，并且

$$\dim L+\dim L^{\perp}=n.$$

$L^{\perp}$ 的正交补$(L^{\perp})^{\perp}$确实为 L。如果 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$，…，$\boldsymbol{b}_m$ 为 L 的基，则 $\boldsymbol{x}\perp L$ 等价于条件 $\boldsymbol{x}\perp\boldsymbol{b}_1$，$\boldsymbol{x}\perp\boldsymbol{b}_2$，…，$\boldsymbol{x}\perp\boldsymbol{b}_m$。特别地，$\mathbf{R}^n$ 中的 $n-1$ 维子空间为一维子空间的正交补，一维子空间是由单个非零向量（唯一的非零倍数）所构成基的子空间。因此，$n-1$ 维子空间是形如 $\{\boldsymbol{x}\mid\boldsymbol{x}\perp\boldsymbol{b}\}$ 的集合，其中，$\boldsymbol{b}\neq\boldsymbol{0}$。超平面为这些集合的平移。但是，

$$\begin{aligned}\{\boldsymbol{x}\mid\boldsymbol{x}\perp\boldsymbol{b}\}+\boldsymbol{a}&=\{\boldsymbol{x}+\boldsymbol{a}\mid\langle\boldsymbol{x},\boldsymbol{b}\rangle=0\}\\&=\{\boldsymbol{y}\mid\langle\boldsymbol{y}-\boldsymbol{a},\boldsymbol{b}\rangle=0\}=\{\boldsymbol{y}\mid\langle\boldsymbol{y},\boldsymbol{b}\rangle=\beta\}\end{aligned}$$