# 最优化方法

张立卫　单　锋　编

科学出版社
www.sciencep.com

# 最优化方法

张立卫　单　锋　编

科学出版社
北　京

## 内 容 简 介

本书介绍最优化模型的理论与计算方法，其中理论包括对偶理论、非线性规划的最优性理论、非线性半定规划的最优性理论、非线性二阶锥优化的最优性理论；计算方法包括无约束优化的线搜索方法、线性规划的单纯形方法和内点方法、非线性规划的序列二次规划方法、非线性规划的增广 Lagrange 方法、非线性半定规划的增广 Lagrange 方法、非线性二阶锥优化的增广 Lagrange 方法以及整数规划的 Lagrange 松弛方法. 本书注重知识的准确性、系统性和算法论述的完整性，是学习最优化方法的一本入门书.

本书可用作高等院校数学系高年级本科生和管理专业研究生的教材，也可作为相关工程技术人员的参考用书.

**图书在版编目(CIP)数据**

最优化方法/张立卫，单锋编. —北京：科学出版社，2010

ISBN 978-7-03-027649-0

Ⅰ. 最… Ⅱ. 张… 单… Ⅲ. 最优化算法 Ⅳ. O242.23

中国版本图书馆 CIP 数据核字(2010) 第 091712 号

责任编辑：姚莉丽 房 阳 / 责任校对：鲁 素
责任印制：张克忠 / 封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京市安泰印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2010 年 6 月第 一 版 开本：B5(720 × 1000)
2010 年 6 月第一次印刷 印张：13 3/4
印数：1—3 500 字数：270 000

**定价：27.00 元**

(如有印装质量问题，我社负责调换)

# 前 言

在非线性优化计算方法方面, 已有许多好的专著和教材出版, 如袁亚湘的专著 [1] 深入系统地介绍了非线性优化的算法理论, 内容涵盖了最前沿的成果. 本书的侧重点在于基础理论和经典方法, 尽量从经典论文和书籍中直接取材, 做到基础扎实, 脉络清晰, 也期望能为读者在研究非线性优化问题时提供基础工具.

本书分为 7 章, 书中多处给出了素材的出处, 以便读者比照阅读.

第 1 章以较大篇幅给出了变分分析的相关素材, 包括集值映射的极限, 集合的切锥、法锥与二阶切集, 非线性系统的稳定性等, 主要的素材取自 Bonnans 和 Shapiro[2], Rockafellar 和 Wets[3] 及 Ruszczyński[4] 的专著.

第 2 章中无约束优化的素材参考了袁亚湘 [1] 和 Ruszczyński[4] 等的专著, 其中 DFP 方法与限制 Broyden 类 (DFP 除外) 的收敛性证明基本上从文献 [5] 与 [6] 中选取素材, BFGS 结合 Wolfe 条件的收敛性从文献 [7] 中选取素材. 信赖域方法的素材取自于 Conn 等的专著 [8].

由于线性规划的理论非常成熟, 中文书籍也很多, 本书在第 3 章中用很短的篇幅介绍这部分内容, 但选材又不失先进性. 从多面体几何出发描述单纯形方法, 而表格形式的单纯形方法则视为矩阵的行变换. 作者从叶荫宇教授的专著 [9] 选取了 Karmarkar 内点算法, 给出了多项式复杂性的详细分析.

对偶理论是以凸分析的共轭函数理论为基础建立起来的. 在第 4 章, 作者想引领读者作这样一些探索: 什么是对偶问题? 对偶间隙在什么条件下为 0? 怎样得到一个一般问题的对偶? 素材大部分从 Bonnans 和 Shapiro 的专著 [2] 中选取.

对于非线性规划的最优性条件, 本书利用切锥、二阶切集和对偶理论分别得到一阶必要性条件和二阶必要性条件, 用反证法证得二阶充分条件. 注意第 5 章中二阶条件的描述和大部分中文书籍中给出的形式有所不同.

第 6 章和第 7 章介绍约束非线性规划的求解方法, 选取了增广 Lagrange 方法和序列二次规划 (SQP) 方法, 其中增广 Lagrange 方法取材于 Ruszczyński[4] 和 Bertsekas[10] 的专著, 序列二次规划方法取材于 Bonnans 等的专著 [11]. 约束非线性规划的信赖域方法以及近年来兴起的滤子 (filter) 方法可以参见袁亚湘的专著 [1]

中的 9.5 节和 9.6 节.

本书是根据作者多年来为大连理工大学数学科学学院本科生、运筹学与控制论专业硕士生讲授相关课程的讲稿, 以及在沈阳航空航天大学的讨论班上的专题素材整理而成的. 作者在此特别感谢大连理工大学数学科学学院的夏尊铨教授、施光燕教授、冯恩民教授和已故的唐焕文教授, 是他们引领作者进入最优化的研究领域. 同时感谢沈阳航空航天大学的党委书记兼校长王维教授对沈阳航空航天大学运筹学研究所工作的大力支持.

限于作者的学识和水平, 书中不足之处在所难免, 敬请读者批评指正.

张立卫　单　锋

2009 年 7 月于沈阳

试读结束：需要全本请在线购买： www.ertongbook.com

# 目　录

# 第1章

# 变分分析的相关素材

## 1.1 凸分析素材

本节仅列出后继讨论中需要的凸分析的相关内容, 为方便起见, 主要引述文献 [12] 的定义和结论, 并指出它们在专著 [12] 中的位置.

### 1.1.1 凸集合

对于 $x, y \in \mathbf{R}^n$, 经过 $x$ 与 $y$ 的直线可以表示为

$$(1-t)x + ty, \quad t \in \mathbf{R}.$$

连接 $x$ 与 $y$ 的线段, 记为 $[x, y]$, 可表示为

$$[x, y] = \{(1-t)x + ty \,|\, t \in [0, 1]\}.$$

设 $M \subset \mathbf{R}^n$ 是一子集合, 若对任意 $x \in M$, $y \in M$, $t \in \mathbf{R}$, 均有 $(1-t)x + ty \in M$, 则称 $M$ 是 $\mathbf{R}^n$ 的一仿射集合.

**命题 1.1.1** ([12, Theorem 1.2]) *每一非空的仿射集合 $M$ 均平行于唯一的线性子空间 $L$. 这一子空间 $L$ 可表示为*

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

**命题 1.1.2** ([12, Theorem 1.4]) *设 $B \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$, 集合*

$$M = \{x \in \mathbf{R}^n \,|\, Bx = b\}$$

*是 $\mathbf{R}^n$ 的一仿射子集, 而且 $\mathbf{R}^n$ 的每一仿射集合均可写成这样的形式.*

设 $C \subset \mathbf{R}^n$ 是一集合, 如果对任意的 $x \in C, y \in C$, 均有 $[x, y] \subset C$, 则称 $C$ 是凸集合.

设 $K \subset \mathbf{R}^n$ 是一集合, 如果对任意的 $x \in K, t > 0$, 均有 $tx \in K$, 则称 $K$ 是一锥. 如果 $K$ 还是凸集合, 则称 $K$ 是一凸锥.

设 $C \subset \mathbf{R}^n$ 是一集合, 包含 $C$ 的最小的仿射集合称为 $C$ 的仿射包, 记为 $\mathrm{aff}(C)$; 若 $C$ 包含原点, 被 $C$ 包含的最大的线性子空间称为 $C$ 的线空间, 记为 $\mathrm{lin}(C)$.

设 $C \subset \mathbf{R}^n$ 是一凸集合, 则它的闭包、内部、相对内部、相对边界分别表示为

$$\begin{aligned}\mathrm{cl}C &= \bigcap_{\varepsilon>0}(C+\varepsilon\mathbb{B}),\\ \mathrm{int}C &= \{x \mid \exists\varepsilon>0, x+\varepsilon\mathbb{B}\subset C\},\\ \mathrm{ri}C &= \{x\in \mathrm{aff}(C) \mid \exists\varepsilon>0, (x+\varepsilon)\mathbb{B}\cap \mathrm{aff}(C)\subset C\},\\ \mathrm{rbd}C &= (\mathrm{cl}C)\setminus(\mathrm{ri}C),\end{aligned}$$

其中 $\mathbb{B}$ 表示单位球.

**命题 1.1.3** ([12, Theorem 6.1]) 设 $C \subset \mathbf{R}^n$ 是一凸集合. 若 $x \in \mathrm{ri}\,C, y \in \mathrm{cl}\,C$, 则对任意 $t \in [0,1)$ 有 $(1-t)x+ty \in \mathrm{ri}\,C$.

**命题 1.1.4** ([12, Theorem 6.2, Theorem 6.3]) 设 $C \subset \mathbf{R}^n$ 是一凸集合, 则

$$\mathrm{aff}\,(\mathrm{cl}\,C) = \mathrm{aff}\,(\mathrm{ri}\,C).$$

特别地, 若 $C \neq \varnothing$, 则 $\mathrm{ri}\,C \neq \varnothing$,

$$\mathrm{cl}\,(\mathrm{ri}\,C) = \mathrm{cl}\,C, \quad \mathrm{ri}\,(\mathrm{cl}\,C) = \mathrm{ri}\,C.$$

**推论 1.1.1** ([12, Corollary 6.3.1]) 设 $C_1$ 与 $C_2$ 是 $\mathbf{R}^n$ 的凸集合, 则 $\mathrm{cl}\,C_1 = \mathrm{cl}\,C_2$ 当且仅当 $\mathrm{ri}\,C_1 = \mathrm{ri}\,C_2$. 这些条件等价于 $\mathrm{ri}\,C_1 \subset C_2 \subset \mathrm{cl}\,C_1$.

**注 1.1.1** 对有穷维 Hilbert 空间, 只要 $C$ 非空, 就有 $\mathrm{ri}\,C \neq \varnothing$, 这一结论在无穷维空间不成立.

**命题 1.1.5** ([12, Theorem 6.4]) 设 $C \subset \mathbf{R}^n$ 是一凸集合, 则 $z \in \mathrm{ri}\,C$ 的充分必要条件是对每一 $x \in C$, 存在 $\mu > 1$ 满足 $(1-\mu)x+\mu z \in C$.

**证明** 若 $z \in \mathrm{ri}\,C$, 则 $C$ 中每一以 $z$ 为一端点的线段均可以延长至 $z$ 点以外而不离开集合 $C$, 因此, 对每一 $x \in C$, 存在 $\mu > 1$ 满足 $(1-\mu)x+\mu z \in C$. 相反地, 设 $z$ 满足所给的条件. 由于 $C \neq \varnothing$, $\mathrm{ri}\,C \neq \varnothing$ (由命题 1.1.4), 存在 $x \in \mathrm{ri}\,C$. 设 $y$ 为满足条件的 $(1-\mu)x+\mu z$ $(\mu > 1)$, 则 $y \in C$, 于是可得 $z = (1-\lambda)x+\lambda y$, 其中 $0 < \lambda = \mu^{-1} < 1$. 由命题 1.1.3 得 $z \in \mathrm{ri}\,C$. ■

**推论 1.1.2** ([12, Corollary 6.4.1]) 设 $C \subset \mathbf{R}^n$ 是一凸集合, 则 $z \in \mathrm{int}\,C$ 当且仅当对每一 $y \in \mathbf{R}^n$, 存在 $\varepsilon > 0$ 满足 $z+\varepsilon y \in C$.

**命题 1.1.6** ([12, Theorem 6.6]) 设 $C \subset \mathbf{R}^n$ 是一凸集合, $A$ 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的线性映射, 则

$$\mathrm{ri}\,(AC) = A(\mathrm{ri}C), \quad \mathrm{cl}\,(AC) \supset A(\mathrm{cl}C).$$

**推论 1.1.3** ([12, Corollary 6.6.2]) 设 $C_1$ 与 $C_2$ 是 $\mathbf{R}^n$ 的凸集合, 则

$$\mathrm{ri}\,(C_1+C_2) = \mathrm{ri}\,C_1+\mathrm{ri}\,C_2, \quad \mathrm{cl}\,(C_1+C_2) \supset \mathrm{cl}\,C_1+\mathrm{cl}\,C_2.$$

### 1.1.2　凸函数的闭包

设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是增广实值函数, $f$ 的上图定义为

$$\operatorname{epi} f=\{(x,\alpha)\in\mathbf{R}^{n+1}\,|\,f(x)\leqslant\alpha\};$$

$f$ 的有效域定义为

$$\operatorname{dom} f=\{x\,|\,f(x)<+\infty\}.$$

称函数 $f$ 是正常的, 如果存在 $x\in\operatorname{dom} f$ 且对任意 $x$ 均有 $f(x)>-\infty$. 如果 $f$ 不是正常的, 则称它是非正常的.

增广实值函数 $f$ 被称为凸函数, 如果上图 $\operatorname{epi} f$ 是 $\mathbf{R}^{n+1}$ 中的凸子集.

**命题 1.1.7** ([12, Theorem 4.1])　*设 $C\subset\mathbf{R}^n$ 是一凸集合, $f:C\to(-\infty,+\infty]$ 是一增广实值函数, 则 $f$ 是 $C$ 上的凸函数的充分必要条件是对任意 $x\in C, y\in C$,*

$$f((1-t)x+ty)\leqslant(1-t)f(x)+tf(y),\quad 0<t<1.$$

**命题 1.1.8** ([12, Theorem 4.2])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是一增广实值函数, 则 $f$ 是凸函数的充分必要条件是只要 $f(x)<\alpha, f(y)<\beta$, 必有*

$$f((1-t)x+ty)\leqslant(1-t)\alpha+t\beta,\quad 0<t<1.$$

设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是增广实值函数, 称 $f$ 在 $x$ 处是下半连续的, 如果

$$f(x)\leqslant\liminf_{y\to x} f(y).$$

**命题 1.1.9** ([12, Theorem 7.1])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是一增广实值函数, 则下述条件等价:*

(1) *$f$ 在 $\mathbf{R}^n$ 上是下半连续的;*

(2) *对每一 $\alpha\in\mathbf{R}$, $\{x\,|\,f(x)\leqslant\alpha\}$ 是闭集合;*

(3) *$\operatorname{epi} f$ 是 $\mathbf{R}^{n+1}$ 的闭子集.*

上图是 $\operatorname{cl}\operatorname{epi} f$ 的函数, 记为 $\operatorname{lsc} f$, 称为 $f$ 的下半连续包, 即

$$\operatorname{epi}(\operatorname{lsc} f)=\operatorname{cl}(\operatorname{epi} f).$$

$f$ 的闭包, 记为 $\operatorname{cl} f$, 定义为

$$(\operatorname{cl} f)(x)=\begin{cases}(\operatorname{lsc} f)(x), & \operatorname{lsc} f>-\infty,\\ -\infty, & \text{其他}.\end{cases}$$

由此可见, 如果 $\operatorname{lsc} f$ 是正常函数, 则 $\operatorname{cl} f=\operatorname{lsc} f$.

**命题 1.1.10** ([12, Theorem 7.2])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是非正常凸函数, 则*

$$f(x)=-\infty,\quad x\in\operatorname{ri}(\operatorname{dom} f).$$

试读结束：需要全本请在线购买： www.ertongbook.com

**证明** 若 $\operatorname{dom} f = \mathbf{R}^n$, 则必有 $f \equiv -\infty$. 若 $\varnothing \neq \operatorname{dom} f \neq \mathbf{R}^n$, 则存在 $u \in \operatorname{dom} f$, 使得 $f(u) = -\infty$. 令 $x \in \operatorname{ri}(\operatorname{dom} f)$. 由命题 1.1.5, 存在 $\mu > 1$ 满足 $y = (1-\mu)u + \mu x \in \operatorname{dom} f$, 于是可得 $x = (1-t)u + ty, 0 < t = \mu^{-1} < 1$. 根据命题 1.1.8 知对任何 $\alpha > f(u), \beta > f(y)$,

$$f(x) = f((1-t)u + ty) < (1-t)\alpha + t\beta.$$

由于 $f(u) = -\infty$, $f(y) < +\infty$, 所以 $f(x) = -\infty$. ■

**推论 1.1.4** ([12, Corollary 7.2.1]) *若凸函数 $f$ 是下半连续的非正常凸函数, 则 $f$ 不取有限值.*

**证明** $f$ 的下半连续性使得 $f(x) = -\infty$ 的点 $x$ 的集合必包含 $\operatorname{cl}\operatorname{ri}(\operatorname{dom} f) = \operatorname{cl}(\operatorname{dom} f) \supset \operatorname{dom} f$, 可见必有

$$f(x) = \begin{cases} +\infty, & x \notin \operatorname{dom} f, \\ -\infty, & x \in \operatorname{dom} f. \end{cases}$$ ■

**命题 1.1.11** ([12, Lemma 7.3]) *设 $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ 是凸函数, 则*

$$\operatorname{ri}(\operatorname{epi} f) = \{(x,\mu) \mid x \in \operatorname{ri}(\operatorname{dom} f), f(x) < \mu < +\infty\}.$$

**证明** 不妨设 $(\operatorname{dom} f)$ 的维数是 $n$. 只需证明

$$\operatorname{int}(\operatorname{epi} f) = \{(x,\mu) \mid x \in \operatorname{ri}(\operatorname{dom} f), f(x) < \mu < +\infty\}.$$

包含关系 “ $\subset$ ” 显然, 只需证明包含关系 “ $\supset$ ”. 令 $\bar{x} \in \operatorname{int}(\operatorname{dom} f)$, $\bar{\mu} \in \mathbf{R}$ 满足 $\bar{\mu} > f(\bar{x})$. 设 $a_1, \cdots, a_r$ 是 $\operatorname{dom} f$ 中的点且满足 $\bar{x} \in \operatorname{int} P$, 其中

$$P = \operatorname{conv}\{a_1, \cdots, a_r\}.$$

令

$$\alpha = \max\{f(a_i) \mid i = 1, \cdots, r\}.$$

对于 $x \in P$, 存在 $\lambda_1, \cdots, \lambda_r$ 满足

$$x = \sum_{i=1}^{r} \lambda_i a_i, \quad \sum_{i=1}^{r} \lambda_i = 1, \quad \lambda_i \geqslant 0, i = 1, \cdots, r,$$

则

$$f(x) \leqslant \sum_{i=1}^{r} \lambda_i f(a_i) \leqslant \alpha.$$

可见开集

$$\{(x,\mu) \mid x \in \operatorname{int} P, \alpha < \mu < +\infty\}$$

包含在 $\operatorname{epi} f$ 中. 特别地, 对任意 $\mu > \alpha$,

$$(\bar{x}, \mu) \in \operatorname{int}(\operatorname{epi} f),$$

可见 $(\bar{x}, \bar{\mu})$ 可视为 $\operatorname{epi} f$ 中与 $\operatorname{int}(\operatorname{epi} f)$ 相交的“竖直”线段的一内部点. 这意味着 $(\bar{x}, \bar{\mu}) \in \operatorname{int}(\operatorname{epi} f)$. ■

**命题 1.1.12** ([12, Theorem 7.4])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是正常凸函数, 则 $\mathrm{cl}\,f$ 是一闭的正常凸函数. 进一步, 除了 $\mathrm{dom}\,f$ 的相对边界点外, $\mathrm{cl}\,f$ 与 $f$ 的取值是相同的.*

**证明**　因为 $f$ 是正常凸函数, $\mathrm{cl}\,f=\mathrm{lsc}\,f$, 从而 $\mathrm{epi}\,(\mathrm{cl}\,f)=\mathrm{cl}\,(\mathrm{epi}\,f)$, $\mathrm{epi}\,f$ 是凸集, $\mathrm{cl}\,(\mathrm{epi}\,f)$ 是闭凸集, 从而 $\mathrm{cl}\,f$ 是下半连续的. 根据推论 1.1.4 得 $\mathrm{cl}\,f$ 是正常函数, 从而是闭函数, 因为 $f$ 在 $\mathrm{dom}\,f$ 上取有限值. 对任意 $x\in\mathrm{ri}\,(\mathrm{dom}\,f)$, 考虑竖直直线 $M=\{(x,\mu)\,|\,\mu\in\mathbf{R}\}$. 由命题 1.1.12, $M$ 与 $\mathrm{epi}\,f$ 是相交的. 因此, 直接由命题 1.1.3 得到

$$M\cap\mathrm{cl}\,(\mathrm{epi}\,f)=\mathrm{cl}\,(M\cap\mathrm{epi}\,f)=M\cap\mathrm{epi}\,f.$$

这意味着 $(\mathrm{cl}\,f)(x)=f(x)$. 另一方面, 如果设 $x\notin\mathrm{cl}\,(\mathrm{dom}\,f)$, 注意到

$$(\mathrm{cl}\,f)(y)=(\mathrm{lsc}\,f)(y)=\liminf_{z\to y}f(z),$$

可以得到

$$\mathrm{cl}\,(\mathrm{dom}\,f)\supset\mathrm{dom}\,(\mathrm{cl}\,f)\supset\mathrm{dom}\,f,$$

从而 $(\mathrm{cl}\,f)(x)=+\infty=f(x)$. ■

**推论 1.1.5** ([12, Corollary 7.4.1])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是正常凸函数, 则 $\mathrm{dom}(\mathrm{cl}\,f)$ 与 $\mathrm{dom}\,f$ 相差的点至多为 $\mathrm{dom}\,f$ 的相对边界点. 特别地,*

$$\mathrm{cl}\,\mathrm{dom}(\mathrm{cl}\,f)=\mathrm{cl}(\mathrm{dom}\,f),\quad \mathrm{ri}\,\mathrm{dom}(\mathrm{cl}\,f)=\mathrm{ri}(\mathrm{dom}\,f).$$

**推论 1.1.6** ([12, Corollary 7.4.2])　*设 $f:\mathbf{R}^n\to\overline{\mathbf{R}}$ 是正常凸函数, $\mathrm{dom}\,f$ 是一仿射集合, 则 $f$ 是闭的.*

**证明**　由于 $\mathrm{dom}\,f$ 没有相对边界点, $f=\mathrm{cl}\,f$. ■

### 1.1.3　共轭函数

设 $X$ 是有限维的 Hilbert 空间, 则 $X$ 与 $X^*$ 是成对空间, 令 $f:X\to\overline{\mathbf{R}}$ 是增广实值 (可能非凸的) 函数, 共轭函数 $f^*:X^*\to\overline{\mathbf{R}}$ 定义为

$$f^*(x^*):=\sup_{x\in X}\{\langle x^*,x\rangle-f(x)\},$$

$f$ 的双共轭函数记为 $f^{**}=(f^*)^*$, 定义为

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

由 $f^*(x^*)$ 的定义可以得到

$$f(x) \geqslant \langle x^*, x\rangle - f^*(x^*),$$

这就是著名的 Young-Fenchel 不等式.

进一步, 对所有的 $x \in X$, 由 Young-Fenchel 不等式可推出 $f(x) \geqslant f^{**}(x)$. 因为 $f^*$ 的上图是半空间

$$\{(x^*, c) \in X^* \times \mathbf{R} | c \geqslant \langle x^*, x\rangle - f(x)\}$$

的交, 所以 $f^*$ 是凸的且下半连续的. 类似地, $f^{**}$ 是凸的且下半连续的.

**命题 1.1.13** ([2, Proposition 2.112]) *若 $f : X \to \overline{\mathbf{R}}$ 是正常的下半连续的凸函数, 则 $f^*$ 也是正常的凸函数.*

**证明** 因为 $f^*$ 是一共轭函数, 所以它是凸的. 因为 $f$ 是正常的, 则对所有的 $x \in X$, $f(x) > -\infty$ 且存在点 $x_0$, $f(x_0)$ 是有限的, 则对任意 $x^* \in X^*$,

$$f^*(x^*) \geqslant \langle x^*, x_0\rangle - f(x_0) > -\infty.$$

因为 $f$ 是下半连续的、凸的, 它的上图是 $X \times \mathbf{R}$ 的闭凸子集, 所以有 epi $f$ 可表示为 $X \times \mathbf{R}$ 的一族闭半空间的交. 由于 $f$ 是正常的, 这些闭半空间中至少有一者必是非竖直的, 即存在 $\bar{x}^* \in X^*$, $c \in \mathbf{R}$ 满足: 对所有的 $x \in X$, $f(x) \geqslant \langle \bar{x}^*, x\rangle - c$, 则

$$f^*(\bar{x}^*) \leqslant \sup_{x\in X}\{\langle \bar{x}^*, x\rangle - (\langle \bar{x}^*, x\rangle - c)\} = c,$$

因此, $\bar{x}^* \in \operatorname{dom} f^*$. 这证得 $f^*$ 是正常的. ■

下面是关于共轭函数的基本对偶性定理.

**定理 1.1.1** (Fenchel-Moreau-Rockafellar)([2, Theorem 2.113]) *设 $f : X \to \overline{\mathbf{R}}$ 是一增广实值函数, 则*

$$f^{**} = \text{cl}\ (\text{conv}\ f). \tag{1.1.1}$$

**证明** 由 $f^*$ 的定义, $(x^*, \beta) \in \text{epi}\ f^*$ 当且仅当

$$f(x) \geqslant \langle x^*, x\rangle - \beta, \quad \forall x \in X.$$

换言之, $(x^*, \beta) \in \text{epi}\ f^*$ 当且仅当 $f$ 的上图被包含在闭的非竖直的半空间中,

$$\{(x, \alpha) \in X \times \mathbf{R} | \alpha \geqslant \langle x^*, x\rangle - \beta\}.$$

于是可得 $f^{**}$ 的上图由闭的非竖直半空间族

$$\{(x, \alpha) \in X \times \mathbf{R} | \alpha \geqslant \langle x^*, x\rangle - \beta\}, \quad (x^*, \beta) \in \text{epi}\ (f^*)$$

的交给出, 或换言之, $f^{**}$ 是不超过 $f$ 的连续的仿射函数的上确界. 因此, epi $(f^{**})$ 是包含 epi $(f)$ 的非竖直半空间的交. 另一方面, lsc(conv$f$) 的上图是包含 epi $(f)$ 的 (可能是竖直的) 半空间的交, 因为它的上图是包含 epi $(f)$ 的最小闭凸子集.

若对至少一点 $x \in X$ 有 $\mathrm{lsc}(\mathrm{conv}f)(x) = -\infty$, 则由定义有 $\mathrm{cl}(\mathrm{conv}f)(\cdot) = -\infty$, 没有不超过 $f$ 的连续的仿射函数存在, 因而 (1.1.1) 成立. 若 $f(\cdot) = +\infty$, 则 $f^{**}$ 与 $\mathrm{cl}(\mathrm{conv}f)$ 等于 $f$. 只剩下考虑 $\mathrm{lsc}(\mathrm{conv}f)$ 是正常的, 因而等于 $\mathrm{cl}(\mathrm{conv}f)$ 的情形. 由上面的讨论, 只需证明若 $\mathrm{lsc}(\mathrm{conv}f)$ 是正常的, 点 $(x_0, \alpha_0)$ 不属于 $\mathrm{lsc}(\mathrm{conv}f)$ 的上图, 则该点可以被一非竖直的半空间与上图分离开来.

设 $\mathrm{lsc}(\mathrm{conv}f)$ 是正常的, 则 $f$ 至少有一不超过它的仿射函数. 事实上, 闭凸集 $\mathrm{epi}\ (\mathrm{lsc}(\mathrm{conv}f))$ 由包含它的半空间的交给出. 若这些半空间均是竖直的, 则 $\mathrm{lsc}(\mathrm{conv}f)$ 在其定义域上取值 $-\infty$. 因为 $\mathrm{lsc}(\mathrm{conv}f)$ 是正常的, 它的定义域非空, 因此, 这与假设 $\mathrm{lsc}(\mathrm{conv}f)(\cdot) > -\infty$ 矛盾, 从而存在 $(x^*, \beta) \in X^* \times \mathbf{R}$ 满足

$$\alpha \geqslant \langle x^*, x\rangle - \beta, \quad \forall (x, \alpha) \in \mathrm{epi}\ (\mathrm{lsc}(\mathrm{conv}f)). \tag{1.1.2}$$

若 $(x_0, \alpha_0)$ 不属于 $\mathrm{lsc}(\mathrm{conv}f)$ 的上图, 则存在一线性泛函 $(\bar{x}^*, a)$, 它强分离 $(x_0, \alpha_0)$ 与 $\mathrm{lsc}(\mathrm{conv}f)$. 若 $a \neq 0$, 则对应的半空间不是竖直的. 因此, 设 $a = 0$, 则存在 $\varepsilon > 0$,

$$\langle \bar{x}^*, x_0\rangle \geqslant \langle \bar{x}^*, x\rangle + \varepsilon, \quad \forall (x, \alpha) \in \mathrm{epi}\ (\mathrm{lsc}(\mathrm{conv}f)).$$

上述不等式乘以 $\gamma > 0$, 加到不等式 (1.1.2) 上得到

$$\alpha \geqslant \langle \hat{x}^*, x\rangle - \hat{\beta}, \quad \forall (x, \alpha) \in \mathrm{epi}\ (\mathrm{lsc}(\mathrm{conv}f)),$$

其中 $\hat{x}^* := x^* + \gamma\bar{x}^*$, $\hat{\beta} := \beta + \gamma\langle \bar{x}^*, x\rangle - \gamma\varepsilon$. 回顾 $\varepsilon > 0$, 对充分大的 $\gamma > 0$, 有不等式

$$\alpha_0 < \langle \hat{x}^*, x_0\rangle - \hat{\beta}$$

成立, 因为右端恰好是 $\langle x^*, x_0\rangle - \beta - \gamma\varepsilon$. 结果得到一非竖直的半空间, 它分离了 $(x_0, \alpha_0)$ 与 $\mathrm{lsc}(\mathrm{conv}f)$ 的上图, 这就完成了证明. ■

定理 1.1.1 表明, $f = f^{**}$ 当且仅当 $f$ 是凸的且是闭的. 特别地, 若对所有的 $x \in X$ 有 $f(x) > -\infty$, 则 $f = f^{**}$ 当且仅当 $f$ 是凸的且是下半连续的.

设 $C \subset X$ 是一集合, $C$ 的指示函数与支撑函数分别定义为

$$I_C(x) = \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C \end{cases}$$

与

$$\sigma(x^* \mid C) = \sup_{x \in C} \langle x^*, x\rangle.$$

**命题 1.1.14** ([12, Theorem 13.2])　*闭凸集的指示函数与它的支撑函数是互为共轭的. 函数是非空集合的支撑函数的充分必要条件是它是正齐次的闭的正常凸函数.*

**证明** 由定义得闭凸集合的指示函数与支撑函数互为共轭, 剩下的只需证明闭的正常凸函数的值只取 0 和 $+\infty$ 的充分必要条件是它的共轭函数是正齐次的. 前一性质等价于 $f(x)=\lambda f(x)(\forall x, \forall\lambda>0)$. 第二性质等价于

$$f^*(x^*)=\lambda f^*(\lambda^{-1}x^*)=(f^*\lambda)(x^*),\quad \forall x^*, \forall\lambda>0.$$

由于

$$\begin{aligned}(\lambda f)^*(x^*)&=\sup_x\{\langle x,x^*\rangle-\lambda f(x)\}\\&=\sup_x\{\lambda\langle x,\lambda^{-1}x^*\rangle-\lambda f(x)\},\end{aligned}$$

所以当 $f$ 是闭凸函数时, $f=\lambda f(\lambda>0)$ 的充分必要条件是 $f^*=f^*\lambda(\lambda>0)$. ■

**推论 1.1.7** ([12, Corollary 13.2.1]) *设 $f$ 是不恒等于 $+\infty$ 的正齐次凸函数, 则 $\mathrm{cl}\, f$ 是下述集合 $C$ 的支撑函数:*

$$C=\{x^*\,|\,\forall x, \langle x,x^*\rangle\leqslant f(x)\}.$$

**证明** $\mathrm{cl}\, f$ 或者是闭的正常的正齐次凸函数, 或者是常值函数 $-\infty$($\varnothing$ 的支撑函数). 因此, 存在一闭凸集合 $C$ 满足 $\mathrm{cl}\, f=\sigma(\cdot\,|\,C)$. 由定义, $f^*=(\mathrm{cl}\, f)^*=I_C(\cdot)$, $C=\{x^*\,|\,f^*(x^*)\leqslant 0\}$, 而 $f^*(x^*)\leqslant 0$ 当且仅当 $\langle x,x^*\rangle-f(x)\leqslant 0(\forall x)$. ■

**推论 1.1.8** ([12, Corollary 13.2.2]) *非空有界凸集合的支撑函数是有限的正齐次凸函数.*

**证明** 由推论 1.1.6 得有限正齐次凸函数必是闭的. 注意到命题 1.1.14 中支撑函数的刻画, 只需要观察到一凸集合 $C$ 是有界的当且仅当对每一 $x^*$ 均有 $\sigma(x^*\,|\,C)<+\infty$ 即可. 事实上, $\mathbf{R}^n$ 的子集合 $C$ 有界当且仅当它被包含在一立方体中, 从而 $C$ 有界当且仅当每一线性函数在 $C$ 上均是有界的. ■

### 1.1.4 次可微性

设 $X$ 是有限维的 Hilbert 空间, 泛函 $x^*\in X^*$ 称为 (可能非凸的) 函数 $f:X\to\overline{\mathbf{R}}$ 在点 $x$ 处的次梯度, 若 $f(x)$ 是有限的且

$$f(y)-f(x)\geqslant\langle x^*,y-x\rangle,\quad \forall y\in X.$$

这表明线性函数 $\ell(y):=\langle x^*,y-x\rangle+f(x)$ 是 $f$ 的上图在 $(x,f(x))$ 的支撑超平面上, 即 $\ell(x)=f(x)$ 且对所有 $y\in X$ 有 $f(y)\geqslant\ell(y)$.

函数 $f$ 在 $x$ 处的所有次梯度的集合称为 $f$ 在 $x$ 处的次微分, 即

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

作为闭半空间的交, $\partial f(x)$ 是一凸集, $\partial f(x)$ 在 $X^*$ 中是闭的. 称函数 $f$ 在 $x$ 处为次可微的, 若 $f(x)$ 是有限的且 $\partial f(x)\neq\varnothing$. 次可微函数显然是正常的. 然而, 并非每一正常凸函数在定义域的每一点处均是次可微的.

由共轭函数的定义得

$$x^* \in \partial f(x)\text{当且仅当} f(x) + f^*(x^*) = \langle x^*, x\rangle. \tag{1.1.3}$$

**命题 1.1.15** ([2, Proposition 2.118]) 设 $f : X \to \overline{\mathbf{R}}$ 是一 (可能非凸的) 函数, 则下述各结论是成立的:

(1) 若对某一 $x \in X$, $f^{**}(x)$ 是有限的, 则

$$\partial f^{**}(x) = \underset{x^* \in X^*}{\operatorname{argmax}}\{\langle x^*, x\rangle - f^*(x^*)\}; \tag{1.1.4}$$

(2) 若 $f$ 在 $x$ 处是次可微的, 则 $f^{**}(x) = f(x)$;

(3) 若 $f^{**}(x) = f(x)$ 且是有限的, 则 $\partial f(x) = \partial f^{**}(x)$.

**证明** 将 (1.1.3) 应用到 $f^{**}$, 则有 $x^* \in \partial f^{**}(x)$ 当且仅当

$$f^{**}(x) = \langle x^*, x\rangle - f^{***}(x^*).$$

由定理 1.1.1 得 $f^{***} = f^*$, 因而上述等式等价于

$$f^{**}(x) = \langle x^*, x\rangle - f^*(x^*). \tag{1.1.5}$$

由双重共轭的定义, $f^{**}(x)$ 等于 (1.1.5) 的右端在 $x^* \in X^*$ 取最大值的最大值点集合, 得到 (1.1.4).

若存在 $x^* \in \partial f(x)$, 则由 (1.1.3) 得 $f(x) \leqslant f^{**}(x)$. 因为总有 $f(x) \geqslant f^{**}(x)$, 性质 (2) 得证.

为证明 (3), 由 (1.1.3) 得 $x^* \in \partial f(x)$ 当且仅当 $f(x) = \langle x^*, x\rangle - f^*(x^*)$, $x^* \in \partial f^{**}(x)$ 当且仅当 (1.1.5) 成立, 这证得结论. ■

由 (1.1.3) 及命题 1.1.15, 若 $x$ 使得 $f(x)$ 是有限的, 则下述结论等价:

(1) $x^* \in \partial f(x)$;

(2) $x \in \partial f^*(x^*)$;

(3) $f(x) + f^*(x^*) = \langle x^*, x\rangle$.

**命题 1.1.16** ([2, Proposition 2.121]) 设 $K$ 是 $X^*$ 的一非空闭凸子集, 令 $f(x) := \sigma(x, K)$ 是相应的支撑函数. 若 $x \in \operatorname{dom} f$, 则

$$\partial f(x) = \operatorname{argmax}\{\langle x^*, x\rangle | x^* \in K\}. \tag{1.1.6}$$

**证明** 由支撑函数 $f(\cdot)$ 是正常的下半连续的凸函数得 $f^*(\cdot) = I_K(\cdot)$, 从而得到 $f^{**}(x) = f(x)$, 因而有 $\partial f^{**}(x) = \partial f(x)$. 由式 (1.1.4) 即得式 (1.1.6). ■

现在考虑定义在 $\mathbf{R}^n$ 上的凸函数的次微分.

**命题 1.1.17** ([12, Theorem 23.1]) 设 $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ 是凸函数, $x \in \mathbf{R}^n$ 满足 $f(x)$ 有限, 则对任何 $y \in \mathbf{R}^n$,

试读结束：需要全本请在线购买： www.ertongbook.com