上海科技专著出版基金资助

多目标规划的理论方法及其应用研究

刘三明 著

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

上海科技专著出版基金资助

多目标规划的
理论方法及其应用研究

刘三明 著

上海交通大學出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书作者以多年来潜心研究的成果为主线，结合国内外相关研究的前沿思想和成果，介绍了多目标规划的理论、方法及其在电力系统多目标最优潮流问题中的应用．内容包括多目标规划的指数罚函数法、构造多目标规划罚函数的统一框架、各种约束的多目标分式规划问题以及广义凸多目标规划问题、多目标规划的正则化牛顿法、广义minmax问题的熵正则化方法和指数罚函数法的对偶性、多目标分式规划中非本质目标函数的特征、集合函数多目标规划的最优性条件及对偶性、电力系统多目标最优潮流问题．

本书可作为普通高等院校相关专业研究生的教学或辅导用书，亦可作为相关领域的科研及工程技术人员的参考用书．

图书在版编目(CIP)数据

多目标规划的理论方法及其应用研究 / 刘三明著．
—上海：上海交通大学出版社，2013
ISBN 978-7-313-10574-5

Ⅰ．①多…　Ⅱ．①刘…　Ⅲ．①多目标(数学)—数学规划—研究　Ⅳ．①O221

中国版本图书馆 CIP 数据核字(2013)第 266030 号

多目标规划的理论方法及其应用研究

著　　者：刘三明
出版发行：上海交通大学出版社
地　　址：上海市番禺路 951 号
邮政编码：200030
电　　话：021-64071208
出 版 人：韩建民
印　　制：上海交大印务有限公司
经　　销：全国新华书店
开　　本：787 mm×1092 mm　1/16
印　　张：11.5
字　　数：252 千字
版　　次：2014 年 7 月第 1 版
印　　次：2014 年 7 月第 1 次印刷
书　　号：ISBN 978-7-313-10574-5/O
定　　价：32.00 元

前　言

人们在解决实际问题的过程中常常需要作出决策.要作出好的决策,首先需要给出判别决策好坏的标准.当只用一个目标作为判断决策好坏的标准时,人们应设法选择使这一目标在某种意义下达到“最优”的决策.这类决策问题就是传统的单目标规划问题.但在现实世界中,衡量一个决策优劣的指标往往不止一个,一般需要多个指标来衡量,而这些指标之间又常常是相互矛盾的.如何协调这些矛盾,兼顾各个指标,选出最满意的方案,是人们在现实生活中经常遇到的难题,多目标规划因此应运产生.与单目标规划问题不同,多目标规划问题的本质在于,大多数情况下,某目标的改善可能引起其他目标性能的降低,同时使多个目标均达到最优是不可能的,只能在各目标之间进行权衡和折中处理,使所有目标函数尽可能达到最优.

本书主要介绍了多目标规划的理论、方法及其在工程中的应用.多目标规划属于数学与工程的交叉学科,在工程设计与优化控制、经济计划、交通运输及国防等领域有着相当广泛的应用.多目标规划的理论要涉及多门现代数学学科,如凸分析,非光滑分析和随机分析等.它的求解方法,则常借助于线性规划、非线性规划、随机规划以及数值计算的手段和技巧.目前,多目标规划无论在理论研究、求解方法,还是应用方面都取得了重要的进展,使它成为一门重要的学科.

多目标规划问题的最优性条件、对偶性、求解方法一直是研究的热点问题,本书旨在这一研究领域作一些有益的工作.

本书力图全面总结作者在多目标规划的理论、方法及其在工程中的应用等方面取得的一系列研究成果.全书共分为 12 章,各章具体内容如下:第 1 章是绪论,介绍本书所需的最优化、凸分析、多目标规划的基础知识;第 2 章研究了具有不等式约束多目标规划的指数罚函数法和多目标规划问题的拉格朗日正则化方法,并给出一种构造具有不等式约束的多目标规划问题的罚函数的统一框架;第 3 章对具有不等式约束的非光滑多目标分

式规划问题，在较弱的条件下建立了该非光滑多目标分式规划的Pareto有效解的最优性条件；第4章利用第3章对具有不等式约束的非光滑多目标分式规划问题的Pareto有效解所建立的最优性条件，研究了其对偶问题；第5章研究了一类带有抽象约束的非光滑多目标分式规划问题在Geoffrion意义下的真有效解的充分条件和必要条件；第6章研究了一类具有不可微凸不等式约束、线性等式约束和抽象约束的不可微多目标分式规划问题的最优性条件和对偶问题；第7章研究了具有F-凸目标函数和F-凸约束函数的多目标规划问题的最优性条件；第8章研究了多目标正则化牛顿法，该方法不需把原来的多目标优化问题标量化，对于任何一个具有紧的水平集的向量凸函数，对任意选择的初始点，正则牛顿法生成的序列均收敛到弱有效解：第9章研究了广义minmax问题的熵正则化方法和指数罚函数法的对偶性及多目标分式规划中非本质目标函数的特征，证明了目标函数是非本质目标函数的充分条件和必要条件，提供了判别多目标分式规划中目标函数是非本质目标函数的有效方法；第10章讨论了集合函数多目标规划问题有效解的一阶充分条件、弱有效解的一阶必要条件以及强有效解的一阶充分条件，以及集合函数多目标规划问题弱有效解的二阶必要条件、局部弱有效解的二阶充分条件和强有效解的二阶充分条件；第11章给出了向量值集合函数的上图象、向量值集合凸函数、向量值集合凹函数以及它们的共轭函数等基本概念，最后我们导出了集合函数多目标规划的Fenchel对偶定理；第12章将电力系统碳排放量加入到最优潮流的目标函数中，构建了一个把电力系统经济性和碳排放量同时作为优化目标的多目标最优潮流模型，用多目标优化问题的等距离算法，求解该多目标最优潮流模型，并给出了一个算例.

本书由上海市科学技术委员会、上海市新闻出版局、上海市科学技术专著出版基金项目资助出版.

本书在编写过程中曾得到上海市自然科学基金“大规模新能源电力多时空动力学特性及安全防御策略研究”(编号12ZRl411600)、国家自然科学基金(青年)“一类非线性切换系统的多目标优化理论与算法”(编号:11201267)、上海市教委科研创新重点资助项目(12ZZl97)、上海市教委创新基金(08DZ2210503)及国家电网公司山东电力公司、上海电气电站集团、上海电气风电设备有限公司、上海电气输配电股份有限公司等单位的支持，在此一并致以诚挚的感谢.

由于作者水平有限，书中不妥和错误之处在所难免，恳请读者批评指正。

作　者

2014年8月

目　录

第1章　绪　论

摘要： 本章主要阐述多目标规划的发展概况，并介绍研究本文所需的预备知识.

1.1　引言

人们在政治、经济和日常生活中常常需要作出决策. 要作出好的决策，首先需要给出判别决策好坏的标准. 当只用一个目标作为判断决策好坏的标准时，人们应设法选择使这一目标在某种意义下达到“最优”的决策. 这类决策问题就是传统的单目标规划问题. 但在现实世界中，衡量一个决策优劣的指标往往不止一个，一般需要多个指标来衡量，而这些指标之间又常常是相互矛盾的. 如何协调这些矛盾，兼顾各个指标，选出最满意的方案，是人们在现实生活中经常遇到的难题，多目标规划因此应运产生. 多目标规划问题出现在许多领域中，例如，在水资源分配这个范围内，多种用途的水库的经营，可以要求灌溉和为附近居民点提供电力，同时还要保持水库本身和下游的最低水位，以适应环境和娱乐的要求. 而这些目标实际上可能是互相冲突的，要同时满足它们往往是矛盾的. 所以，多目标规划问题，总是以牺牲一部分目标的利益来换取另一些目标利益的改善，正如 von neuman 和 Morgenstern[1, 2]指出的：“这种多目标的情形，肯定是无最优值的问题，而是几个相互冲突的最优问题特有的和扰人的混合……这一类问题不能用传统数学方法来处理……”这就是多目标规划问题的基本性质之一.

多目标规划的理论要涉及多门现代数学学科，如凸分析、非光滑分析和随机分析等. 它的求解方法，则常借助于线性规划、非线性规划、随机规划以及数值计算的手段和技巧. 目前，多目标规划无论在理论研究、求解方法，还是应用方面都取得了重要的进展，使它成为一门庞大的学科. 多目标规划包括广泛和丰富的内容. 多目标规划问题也称为：多目标优化问题或向量优化问题或向量极值问题，它与只含有一个数值目标函数的单目标规划问题不同，在一般情况下并不存在最优解. 因此，多目标规划理论的研究包括以下几个方面：① 解的定义. 由于多目标规划的目标不止一个，如何定义多目标最优化解的概念成为

首要问题;② 对于各类解的最优性条件的研究,是多目标规划中一个重要和基本的课题.所谓解的最优性条件,就是在某种意义下所考虑的解要满足的必要和充分条件.它不仅可以用来判断一个可行解是否为所考虑意义下的最优解(因此被广泛地用于设计求解的算法),还可以通过它来讨论稳定性理论等;③ 解的连通性.多目标规划的解通常不唯一,而是一个集合,所以讨论解集的连通性,成为一个重要的研究方向;④ 对偶性.对偶理论的内容非常丰富,是多目标规划这门学科的重要组成部分,它在多目标规划的理论研究和应用研究中都扮演着十分重要的角色.但到目前为止,对偶理论的整个内容还不成熟,正处在充实和发展阶段,有许多课题尚待研究;⑤ 稳定性.稳定性是多目标规划理论中一个重要的研究领域.由于多目标规划具有不同于单目标规划的某些特殊性,使得这方面的研究呈现复杂多样的特点.众所周知,非线性规划的扰动主要是指扰动解的连续性、可微性等性质.而在多目标规划中,解的概念要依赖于控制结构的选择,因而,其扰动除了涉及可行集和目标函数外,还要考虑控制结构的扰动.此外,又由于多目标规划解的多样性(如有效解,弱有效解,真有效解,较多有效解等)以及稳定性的多样性(如上、下半连续,次可微,Lipschitz 连续等),使这方面的研究成为一个多姿多彩的领域;⑥ 求解方法的研究.

本文只对多目标规划的某些方面进行研究,主要是多目标规划的最优性条件和对偶性,以及多目标规划的求解方法.

1.2 多目标规划的发展概况

多目标规划的思想萌芽于 1776 年经济学中关于效用理论的研究.1896 年,经济学家 Pareto[3] 首先在经济平衡的研究中提出了多目标规划问题,并给出了后来被称为 Pareto 最优解的朴素思想.1947 年,von neuman 和 Morgenstern[4] 在对策论的著作中提到了多目标规划问题,引起了人们对多目标规划的重视.1951 年,Koopmans[5] 在生产与分配的活动分析中提出了多目标最优化问题,并第一次提出了 Pareto 最优解的概念.同年,Kuhn 和 Tucker[6] 从数学规划的角度,给出了向量极值问题的 Pareto 最优解的概念,并研究了这种解存在的充分条件和必要条件.1954 年 Debreu[7] 有关评价均衡的讨论,以及 1958 年 Harwicz[8] 对拓扑向量空间中的多目标最优化问题的研究,都为这门学科的建立奠定了坚实的基础.1968 年,Johnsen[9] 出版了关于多目标决策模型的第一部专著.从此,不少学者先后转入这一研究领域,并取得了许多成果.直到 20 世纪 70,80 年代,经过众多学者的努力,终于建立起多目标规划的基本理论,使之成为应用数学的一个新学科分支.下面介绍本文涉及的多目标规划有关方面的发展概况.

1.2.1 解的最优性条件

在 Kuhn 和 Tucker[6] 的开拓性工作之后,国内外许多学者致力于多目标规划最优性条件的研究,取得了许多重要的成果.1976 年,对于有限维多目标规划问题的 Pareto 有效解,Lin[10] 首先给出了它的 Fritz John 必要条件和 Kuhn-Tucker 必要条件.此后,在 1977 年,Borwein[11] 利用切锥给出了一个最优性条件.在 Lin 工作的基础上,林锉云[12, 13, 15] 从

1982 年开始,利用特定的向量函数和次线性泛函等工具,给出了在广义拟凸和广义伪凸下的若干充要条件. 应玫茜[16]利用伪凸性,建立了解的若干充要条件和判别准则. 在 1989 年,Combini 和 Martein[17]给出了关于凸锥的有效解的 Fritz-John 型条件,考虑的空间是实的线性空间. 同年,刘三阳[66]利用 Dini 导数,建立了非光滑多目标规划的 Fritz-John 型条件. 在 1990 年,胡毓达和孙尔江[18]在对目标函数和约束函数上加上适当凸性的条件下,得到了几个 Fritz-Jhon 型充分条件. 在 F. J. 型条件中,为了保证关于目标函数的梯度的系数不为零以使之确实具备最优性条件的功能,应对目标函数附加一定的约束规格. Lin[10]给出了 K-T 约束规格下多目标规划的最优性 Kuhn-Tucker 条件. Sighn[20]利用收敛向量集的概念,给出了多目标可微规划一种非 K-T 约束规格的 Kuhn-Tucker 条件. Wang 和 Dong[21]用 Sighn 的方法,给出了非光滑多目标规划的一种约束规格. Maeda[22]利用线性锥和切锥给出了可微多目标规划 K-T 必要条件成立的约束规格. Li[23]通过引进有限个凸集并的核的概念,给出了非光滑多目标强 K-T 条件成立的约束规格. 胡毓达[24~26]在讨论较多有效解类时也得到了几个最优性条件. 关于无限维空间的情况,对最优性条件的研究成果则较少. 对 Banach 空间多目标规划问题锥有效解的最优性条件,陈广亚[27]曾得到若干结果. 对于拓扑向量空间的情况,李泽民[29]借助 Jeyacumar[28]的次拟凸概念,讨论了 Fritz-John 型必要条件,并给出了一个鞍点条件. 之后,胡毓达[30]引进广义约束规格,进一步得到了拓扑向量空间中锥弱有效解的 Kuhn-Tucker 最优性必要条件,同时还在锥凸的假设下证明了锥有效解和锥弱有效解的几个充分条件. Brando, Rojas-Mmedar 和 Silva[31]给出了 Banach 空间中非凸非光滑向量最优化问题的最优性条件.

1.2.2 对偶性

在 1951 年,Gale, Kuhn 和 Tucker[32]第一次给出了关于向量最优化的对偶性的有关结果. 他们建立了线性多目标规划的对偶定理. 后来,Kornbluth[33], Rodder[34]和 Isermann[35~36]进一步发展了线性多目标规划的对偶定理.

Isermann[37]首先给出了 Gale, Kuhn 和 Tucker, Isermann 和 Kornbluth 的对偶概念之间的关系. Ivanov 和 Nehse[38]研究了线性多目标最优化中一些对偶概念之间的关系,即 Gale, Kuhn 和 Tucker[32], Schonefeld[39], Isermann[35~36], Rodder[34]和 Fiala[40]所提出的对偶概念之间的关系.

最近 Galperin 和 Jimenez[41]用另一种方法研究了线性多目标最优化的对偶问题,该方法是以 Galperin[42]的平衡集方法为基础的.

关于非线性向量最优化问题,其对偶理论是很丰富的. 目前讨论较多的有 Lagrange 对偶模型、Mond-Weir 对偶模型、共轭对偶模型等. 在 Lagrange 对偶理论方面,1979 年,Tanino[43]等人根据单目标规划情形 Wolfe 对偶模型的框架,通过引入所谓的向量 Lagrange 函数提出了多目标规划的 Lagrange 对偶模型,并证明了当原问题是凸规划问题时,原问题和 Lagrange 对偶模型的有效解是对偶的. 此后,林锉云[44~46]进一步推广了上述结果,证明了条件减弱为广义凸性条件(如伪凸、拟凸、Invex 凸等)时,原问题和对偶

问题的有效解的对偶性仍然成立. 1993 年李仲飞和汪寿阳[47]则讨论了有限维空间中基于尖闭凸锥内部(即锥弱有效解)的 Lagrange 对偶问题. 1981 年, Mond 和 Weir[48]给出了多目标规划问题的另一种对偶模型,这种模型一出来,就引起了广大学者的极大兴趣,他们对 Lagrange 对偶模型进行了修改,简化了目标函数,增加了约束条件,并证明了有效解之间的对偶性. 我们把这种对偶模型称为 Mond-Weir 对偶模型. 在多目标规划的共轭对偶理论方面,在 1979 年, Tanino 和 Sawaragi[43]用共轭映射的概念研究了向量最优化问题的对偶. 实际上,他们通过对向量映射和集值映射引入新的概念——共轭映射和次梯度,把由 Rockafellar[75]所得到的单目标规划的对偶理论扩展到多目标规划. 此后,许多学者对集值最优化问题(集值最优化问题是向量最优化问题的推广)引入了不同的共轭对偶理论. 1980 年 Tanino 和 Sawaragi[50]给出了有限维空间中基于正锥(即 Pareto 有效解)的共轭对偶结果,其特点是: 利用共轭函数及性质来建立对偶问题,并证明了对偶定理. 1981 年和 1982 年 Kawasaki[51, 52]讨论了有限维空间中基于正锥内部(即 Pareto 弱有效解)的共轭对偶问题, 1988 年汪寿阳[53]建立了 Banach 空间中基于尖闭凸锥(即锥有效解)的共轭对偶理论. 最近, Huang 和 Yang[54]以及 Singh, Bhatia 和 Rueda[55]通过增广拉格朗日函数建立了其对偶模型,并证明了对偶定理; Osuna-Gomez, Rufian-Lizana 和 Ruiz-Canales[56]讨论了具有广义次似凸函数的不可微多目标规划问题的 Lagrange 对偶; Suneja, Lalitha, Seema Khurana[57]建立了多目标规划的二阶对称对偶模型,并在 η-二阶凸和 η-伪二阶凸的条件下,证明了弱对偶定理、强对偶定理和逆对偶定理; Gert 和 Radu[58]用 Fenchel-Rockerfellar 对偶理论导出了约束条件为线性不等式约束、目标函数为凸函数的平方与凹函数之商的多目标规划的对偶问题; Tarvainen[59]给出了凸多目标规划偏爱解的对偶理论.

1.2.3 多目标最优化方法

在求解多目标规划问题的方法方面,按 Huang 和 Masud[60]于 1979 年提出的由决策者的偏爱信息的表达方式来划分的原则,可分成以下三类: 事先评价法、事中评价法和事后评价法.

事先评价法要求决策者在求解多目标规划问题之前,一次性提供全部偏爱信息,它主要有评价函数法和目标规划法. 评价函数就是决策者事先给出关于多个目标的评价函数,从而将多目标最优化问题转化为单目标规划问题. 最典型的评价函数法有线性加权法、极大极小法和理想点法,参见文献[61]. 目标规划法[62]则是由决策者事先给出每个目标函数的目标函数值组成目标向量,然后将多目标规划问题,转化为对向量目标函数与目标向量之间在某种意义下的距离的极小化问题. 事先评价法是根据决策者事先给出的偏爱信息来构成多目标规划问题的求解,求解方法主要是普遍被人们采用的评价函数法,以及由 Charns 和 Cooper 提出的目标规划法.

事中评价法即交互规划算法,其特点是: 在求解过程中决策人通过与分析人(或计算机)之间的对话逐步给出其偏好信息,直至决策人在某步迭代时对分析人提供的某个解满意为止. 自 1971 年 Benayoun[63]等提出第一个交互规划算法以来,关于这类算法的研究已

得到了快速的发展.

事后评价法的特点是,寻求多目标问题的所有(或大多数)的有效解,然后把这些解送给决策人,由他从其中选择一个最合意的解. 事后评价法是其他两种方法的基础. 它的主要方法有:变权线性加权法[64],ε-约束法[65],多目标线性规划法[66],同伦法[67],广义同伦法[68]与边界垂直相交法[69]. 特别值得一提的是我国学者在多目标算法方面做了许多工作,取得了不少引人注目的成果. 主要有:施保昌、于寅[70]的"一类不可微多目标决策问题的可行方向法";张乐友、刘三阳[71]的"多目标规划的光滑同伦方法";孟红云、刘三阳[120]的"求解多目标优化问题的多智能体遗传算法",还有施保昌、陈铤[72]的"多目标的一类基于精确罚函数的交互式方法"以及刘庆怀、林正华[73]的"求解多目标最小弱有效解的同伦内点法"等.

1.3 预备知识

1.3.1 连续函数的基本概念及性质

本节介绍连续函数的一些性质及方向导数、次微分等概念,这是研究不可微优化问题的基本工具.

定义 1.3.1 设 ν 是实赋范空间

(1) $\{x_i\}_{i=0}^{\infty}$ 是 ν 中的点列,如果

$$\lim_{i\to\infty}\| x_i - x^* \| = 0$$

那么称 $\{x_i\}_{i=0}^{\infty}$ 收敛于 x^*,即当 $i\to\infty$ 时,$x_i\to x^*$, x^* 称为 $\{x_i\}_{i=0}^{\infty}$ 的极限点;

(2) $\{x_i\}_{i=0}^{\infty}$ 是 ν 中的点列,如果存在子集 $K\subset \mathbf{N}=\{1, 2, \cdots\}$,使得对每个子列 $\{x_i\}_{i\in K}$,

$$\lim_{i\to\infty,\ i\in K}\| x_i - x^* \| = 0$$

那么称 x^* 为 $\{x_i\}_{i=0}^{\infty}$ 的聚点,记为当 $i\to\infty$ 时,$x_i \xrightarrow{K} x^*$,或 $\lim_K x_i = x^*$.

定理 1.3.1 假设 $X\subset \mathbf{R}^n$ 是闭有界集,那么 X 是紧集,$\mathbf{R}^n$ 是 n 维实域欧氏空间.

推论 1.3.1 假设 $\{x_i\}_{i=0}^{\infty}$ 是 $\mathbf{R}^n$ 中的有界点列,如果 $\{x_i\}_{i=0}^{\infty}$ 仅有一个聚点 x^*,那么它收敛于 x^*.

定理 1.3.2 假设 $\{x_i\}_{i=0}^{\infty}$ 是 $\mathbf{R}$ 中的点列,且 $x_0 \geqslant x_1 \geqslant x_2 \geqslant \cdots$(即点列是单调递减的). 如果 $\{x_i\}_{i=0}^{\infty}$ 有聚点 x^*,那么当 $i\to\infty$ 时,$x_i\to x^*$,即 x^* 是极限点.

证明: 见文献[121]中的命题 5.1.16.

推论 1.3.2 假设 $\{x_i\}_{i=0}^{\infty}$ 是 $\mathbf{R}$ 中单调下降的点列,如果存在 $b\in \mathbf{R}$,使得对所有的 $i\in \mathbf{N}$, $\boldsymbol{x}_i\geqslant b$,那么 $\{x_i\}_{i=0}^{\infty}$ 收敛于 $\boldsymbol{x}^*\in \mathbf{R}$.

定义 1.3.2 设 ν 是实赋范空间

(1) 函数 f: $\nu\to \mathbf{R}$ 在点 $x\in \nu$ 处是上半连续的(upper semicontinuous),如果对 $\forall\, \delta >$

0,存在 $\rho>0$,使得

$$f(x')-f(x)<\delta,\ \forall x'\in \mathring{B}(x,\rho) \tag{1.3.1}$$

这里 $B(x,\rho)\xlongequal{\text{def}}\{x'\in\nu \mid \|x'-x\|<\rho\}$;

(2) 函数 $f:\nu\to\mathbf{R}$ 在点 $x\in\nu$ 处是下半连续的(lower semicontinuous),如果对 $\forall\delta>0$,存在 $\rho>0$,使得

$$f(x')-f(x)>-\delta,\ \forall x'\in \mathring{B}(x,\rho) \tag{1.3.2}$$

定理 1.3.3　设 ν 是实赋范空间

(1) 函数 $f:\nu\to\mathbf{R}$ 在 x^* 上半连续,当且仅当点列 $\{x_i\}_{i=0}^{\infty}\subset\nu$,当 $i\to\infty$ 时,$x_i\to x^*$,$\overline{\lim}f(x_i)\leqslant f(x^*)$;

(2) 函数 $f:\nu\to\mathbf{R}$ 在 x^* 下半连续,当且仅当点列 $\{x_i\}_{i=0}^{\infty}\subset\nu$,当 $i\to\infty$ 时,$x_i\to x^*$,$\underline{\lim}f(x_i)\geqslant f(x^*)$.

定义 1.3.3　设 $\mathscr{B}$ 是实 Banach 空间,$f:\mathscr{B}\to\mathbf{R}^m$,

(1) $x,h\in\mathscr{B}$,

$$f(x;h)\xlongequal{\text{def}}\lim_{t\downarrow 0}\frac{f(x+th)-f(x)}{t} \tag{1.3.3}$$

这里 $t>0$. 若极限存在,称 $f(x;h)$为 $f(\cdot)$在 x 处沿方向 h 的方向导数;

(2) 若对 $\forall h\in\mathscr{B}$,方向导数 $f(x;h)$存在,则称 $f(\cdot)$在 x 处方向可微;

(3) $x^*,h\in\mathscr{B}$,

$$f^{\circ}(x;h)\xlongequal{\text{def}}\overline{\lim}_{t\downarrow 0,\,x'\to x^*}\ \frac{f(x'+th)-f(x')}{t} \tag{1.3.4}$$

若上式极限存在,则称 $f(\cdot)$ 在 x^* 处沿方向 h 的 Clarke 方向导数或称为广义方向导数.

当 $f:\mathscr{B}\to\mathbf{R}$ 是局部 Lipschitz 连续的,它的 Clarke 方向导数存在,但是方向导数不一定存在. 若 $f:\mathscr{B}\to\mathbf{R}$ 的方向导数与 Clarke 方向导数都存在,则 $f^{\circ}(x;h)\geqslant f(x;h)$. 又若 $\forall x,h\in\mathscr{B}$, $f^{\circ}(x;h)=f(x;h)$,则称 $f(\cdot)$ 是 Clarke 正则的,简称为正则的.

定义 1.3.4　$\mathscr{B}$ 是实 Banach 空间,假设 $f:\mathscr{B}\to\mathbf{R}$, $\forall x,h\in\mathscr{B}$,Clarke 方向导数 $f^{\circ}(x;h)$ 存在,定义 $f(\cdot)$ 在 $x\in\mathscr{B}$ 的 Clarke 次微分 $\partial^{\circ}f(x)\subset\mathscr{B}$ 如下:

$$\partial^{\circ}f(x)\xlongequal{\text{def}}\{\xi\in\mathscr{B}\mid f^{\circ}(x;h)\geqslant\langle\xi,h\rangle,\ \forall h\in\mathscr{B}\} \tag{1.3.5}$$

下面我们给出极大值函数

$$\psi(x)\xlongequal{\text{def}}\max_{j\in \boldsymbol{q}}f^j(x) \tag{1.3.6}$$

的一些性质. 其中 $f^j:\mathbf{R}^n\to\mathbf{R}$, $j=1,2,\cdots q$, $Q\xlongequal{\text{def}}\{1,2,\cdots q\}$. 对 $\forall x\in\mathbf{R}^n$,定义

$$Q(x)\xlongequal{\text{def}}\{j\in Q\mid f^j(x)=\psi(x)\} \tag{1.3.7}$$

定理 1.3.4 函数 $\psi(\cdot)$如式(1.3.6)定义,这里 f^j: $\mathbf{R}^n\to\mathbf{R}$, $j\in Q$,是局部 Lipschitz 连续的,且 $\forall x, h\in\mathbf{R}^n$,方向导数 $f^j(x; h)$存在. 那么

(1) $\psi(\cdot)$是局部 Lipschitz 连续的;

(2) $\partial^\circ\psi(x)\subset co\{\partial^\circ f^j(x), j\in\hat{Q}(x)\}$.

这里$\hat{\boldsymbol{q}}(x)$如式(1.3.7)定义.

推论 1.3.3 函数 $\psi(\cdot)$如式(1.3.6)定义,f^j: $\mathbf{R}^n\to\mathbf{R}$, $j\in Q$ 是连续可微的,那么

(1) $\forall x, h\in\mathbf{R}^n$,广义方向导数 $\psi^\circ(x; h)$和方向导数 $\psi(x; h)$均存在且

$$\psi^\circ(x; h)=\psi(x; h)=\max_{j\in\hat{\boldsymbol{q}}(x)}<\nabla f^j(x), h> \tag{1.3.8}$$

(2) 对 $\forall x\in\mathbf{R}^n$,方向导数 $\psi(\cdot;\cdot)$是上半连续的,函数 $\psi(x;\cdot)$是正齐次、次可加的 Lipschitz 连续函数.

1.3.2 凸分析的基本知识

下面我们简单介绍与本文内容相关的凸分析中的一些定义和定理.

定义 1.3.5 设 f: $\mathbf{R}^n\to(-\infty, +\infty]$是一凸函数,如果它是下半连续的,则称为闭凸函数;若对每一个 $x\in\mathbf{R}^n$,都有 $f(x)>-\infty$,同时至少存在一点$\hat{x}$使得 $f(\hat{x})<+\infty$,则称为正常凸函数. 记 $\mathrm{dom}f=\{x\in\mathbf{R}^n\mid f(x)<+\infty\}$为 f 的有效域.

定义 1.3.6 设 f: $\mathbf{R}^n\to\{-\infty, +\infty]$是一凸函数,$f$ 的凸共轭 f^* 定义为

$$f^*(x^*)=\sup_{x\in\mathbf{R}^n}\{\langle x, x^*\rangle-f(x)\}=\sup_{x\in\mathrm{dom}f}\{\langle x, x^*\rangle-f(x)\}.$$

f 的单调凸共轭 f^+ 定义为

$$f^+(x^*)=\sup_{x\geqslant0}\{\langle x, x^*\rangle-f(x)\}.$$

定理 1.3.5 若 f 是 $\mathbf{R}^n$ 上的闭正常凸函数,则 f^* 也是 $\mathbf{R}^n$ 上的闭正常凸函数.

定义 1.3.7 设 $C\subset\mathbf{R}^n$ 是非空凸集,定义 C 的支撑函数σ_C: $\mathbf{R}^n\to\mathbf{R}$ 如下

$$\sigma_C(h)\xlongequal{\text{def}}\sup\{\langle h, x\rangle\mid x\in C\} \tag{1.3.9}$$

定义 1.3.8 设 $C\subset\mathbf{R}^n$ 是非空集合,定义 C 上的指示函数$\delta(x|C)$为

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

定理 1.3.6 $\delta(x|C)$为凸函数的充分必要条件为 C 是凸集.

定理 1.3.7 如果 C 是凸集,则指示函数的共轭函数为支撑函数,即

$$\delta^*(x^*\mid C)=\sigma_C(x^*).$$

定义 1.3.9 设 f: $\mathbf{R}^n\to\{-\infty, +\infty]$是一凸函数,则向量 x^* 被称为是 f 在点 x 处的次梯度,如果

$$f(z) \geqslant f(x) + \langle x^*, z - x \rangle, \qquad \forall z \in \mathbf{R}^n.$$

f 在点 x 处的次梯度构成的集合称为 f 在点 x 处的次微分，记为 $\partial f(x)$，即

$$\partial f(x) = \{x^* \mid f(z) \geqslant f(x) + \langle x^*, z - x \rangle, \forall z \in \mathbf{R}^n\}.$$

定理 1.3.8　如果 f 是正常凸函数，则 $\forall x \in \mathrm{ri}(\mathrm{dom} f)$，有 $\partial f(x) \neq \varnothing$，即正常凸函数在其有效域的相对内部总是次可微的.

定理 1.3.9　如果凸函数 f 在点 x 处是局部 Lipschitz 连续的，则

$$\partial^\circ f(x) = \partial f(x).$$

定义 1.3.10　如果对每一个满足 $\| x_k \| \to +\infty$ 的序列 $\{x_k\}$，都有

$$\lim_{k \to \infty} f(x_k) = \infty,$$

则称函数 $f: \mathbf{R}^n \to (-\infty, +\infty]$ 是 Coercive.

定义 1.3.11　设 $f: \mathbf{R}^n \to (-\infty, +\infty]$ 是闭的正常凸函数，如果 f 的上图象 $\mathrm{epi} f$ 不含有非垂直的半直线，即

$$f_\infty(y) = +\infty, \ \forall y \neq 0$$

或

$$\lim_{\varepsilon \to 0} \varepsilon f(\varepsilon^{-1} y) = +\infty, \ 0 \neq y \in \mathrm{dom} f,$$

则称 f 是 cofinite.

注　$f_\infty(y)$ 表示函数 f 的回收函数.

定义 1.3.12　设 $f: \mathbf{R}^n \to (-\infty, +\infty]$ 是正常的凸函数，如果它满足条件

(1) $C \overset{\mathrm{def}}{=\!=} \mathrm{int}(\mathrm{dom}\ f) \neq \varnothing$；

(2) f 在集合 $\mathrm{int}(\mathrm{dom} f)$ 上是可微的；

(3) 对每一个收敛于集合 $\mathrm{int}(\mathrm{dom} f)$ 边界点 x 的序列 $\{x_i\} \in C$，都有 $\lim\limits_{i \to \infty} \| \nabla f(x_i) \| = +\infty$，

则称函数 f 是基本光滑的.

定义 1.3.13　设 $f: \mathbf{R}^n \to (-\infty, +\infty]$ 是严格凸函数，如果它满足条件

(1) $C \overset{\mathrm{def}}{=\!=} \mathrm{int}(\mathrm{dom} f) \neq \varnothing$ 且 $\mathrm{int}(\mathrm{dom} f)$ 为开凸集；

(2) f 在集合 $\mathrm{int}(\mathrm{dom} f)$ 上是可微的；

(3) 对每一个收敛于集合 $\mathrm{int}(\mathrm{dom} f)$ 边界点 x 的序列 $\{x_i\} \in C$，都有 $\lim\limits_{i \to \infty} \| \nabla f(x_i) \| = +\infty$，

则称函数 f 是 Legendre 型函数. Legendre 型函数的全体记为 $L(C)$.

定理 1.3.10[75]　若 $f \in L(C)$，则 f 的共轭函数 f^* 也是 Legendre 型凸函数，且

$$(f^*)' = (f')^{-1}.$$

特别地，定义在 $\mathbf{R}^n$ 上的光滑凸函数也是基本光滑的，并将相应的 Legendre 型凸函数记为 $L(\mathbf{R}^n)$.

定理 1.3.11[76] 设 $f: \mathbf{R} \to (-\infty, +\infty]$是一闭函数，如果下列条件之一成立：

(1) $\mathrm{dom}f$ 是一紧集；

(2) $\mathrm{dom}f$ 是闭的且 f 是 coercive；

(3) 存在实数 $\gamma \in \mathbf{R}$ 使得集合 $\{x \mid f(x) \leqslant \gamma\}$ 是一非空紧集，

则 f 在 $\mathbf{R}$ 上达到最小值.

定理 1.3.12[75] 设 f 是一正常凸函数，则 f 的回收函数 f_∞是一正齐次正常凸函数，并且

$$f_\infty(y) = \sup\{f(x+y) - f(x) \mid x \in \mathrm{dom}f\}, \ \forall\, y.$$

如果 f 是闭的，则 f_∞ 也是闭的，并且对任意 $x \in \mathrm{dom}f$，有

$$f_\infty(y) = \sup_{\lambda>0} \frac{f(x+\lambda y) - f(x)}{\lambda} = \lim_{\lambda\to\infty} \frac{f(x+\lambda y) - f(x)}{\lambda}.$$

推论 1.3.4[75] 如果 f 是任一闭正常凸函数，则对每一个 $y \in \mathrm{dom}f$，都有

$$f_\infty(y) = \lim_{\lambda\downarrow 0}(f\lambda)(y) = \lim_{\lambda\downarrow 0} \lambda f(\lambda^{-1} y), \ \forall\, y \in \mathrm{dom}f.$$

如果 $0 \in \mathrm{dom}f$，则此公式对每一 $y \in \mathbf{R}^n$ 都成立.

定理 1.3.13[75] 设 f 是一正常凸函数，则 $\mathrm{dom}f$ 上的支撑函数是 f^* 的回收函数 $(f^*)_\infty$；如果 f 是闭的，则 $\mathrm{dom}f^*$ 上的支撑函数是 f 的回收函数 f_∞.

推论 1.3.5[75] 设 f 是 $\mathbf{R}^n$ 上的闭凸函数，则当且仅当 f 是 co-finite 函数时 f^* 处有限，即 $\mathrm{dom}f^* = \mathbf{R}^n$.

定理 1.3.14[75] 假设 $f_1, \cdots f_m$ 是 $\mathbf{R}^n$ 上的正常凸函数，则有

$$(f_1 \Diamond \cdots \Diamond f_m)^* = f_1^* + \cdots + f_m^*,$$

$$(\mathrm{cl}f_1 + \cdots \mathrm{cl}f_m)^* = \mathrm{cl}(f_1^* \Diamond \cdots \Diamond f_m^*).$$

若集合 $\mathrm{ri}(\mathrm{dom}f_i)(i=1, 2, \cdots m)$有公共点，则第二式子中的闭包运算可去掉且

$$(f_1 + \cdots + f_m)^*(x^*) = \inf\{f_1^*(x_1^*) + \cdots + f_m^*(x_m^*) \mid x_1^* + \cdots + x_m^* = x^*\}$$

其中对每一个 x^*，下确界是可以达到的.

注 符号$\Diamond$表示最小卷积.

定理 1.3.15[75] 设 $f_1, \cdots f_m$ 是 $\mathbf{R}^n$ 上的正常凸函数，令 $f = f_1 + \cdots + f_m$，则

$$\partial f(x) \supset \partial f_1(x) + \cdots + \partial f_m(x), \ \forall\, x.$$

如果凸集 $\mathrm{ri}(\mathrm{dom}f_i)$，$i=1, 2, \cdots m$ 有公共点，则

$$\partial f(x) = \partial f_1(x) + \cdots + \partial f_m(x), \ \forall\, x.$$