面向复杂优化问题求解的智能优化方法

Intelligent Optimization Methods for Solving Complex Optimization Problems

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

内 容 简 介

本书首先从一般优化问题求解的角度，论述了最优化问题的研究意义、优化研究中的基本概念与数学模型、优化问题的分类与求解方法运用原则，并介绍了智能优化的概念、方法分类以及发展历史。然后介绍了 8 类典型智能优化方法的基本思想、概念和原理、步骤流程、典型的算法变体以及算法的扩展与改进设计等内容。从共性理论与方法角度出发，介绍了智能优化方法的统一框架、探索–开发权衡理论，并介绍了典型的混合智能优化算法以及通用的混合策略分类法。在应用方面，针对多类典型的、具有不同难度特征的优化问题，分别介绍了智能优化方法在这些问题求解中的设计与运用方法。这些问题提炼于作者长期从事的与先进火力与指挥控制系统相关的科研实践。

本书可供自动化、计算机、系统工程、信息处理、运筹与管理、应用数学等专业的教师以及相关领域的技术开发人员参考，也可作为相关专业的高年级本科生和研究生的教材。

版权专有　侵权必究

图书在版编目（CIP）数据

面向复杂优化问题求解的智能优化方法/辛斌，陈杰著. —北京：北京理工大学出版社，2017.9

ISBN 978–7–5682–4891–4

Ⅰ. ①面…　Ⅱ. ①辛…　②陈…　Ⅲ. ①计算机算法–最优化算法–研究　Ⅳ. ①TP301.6

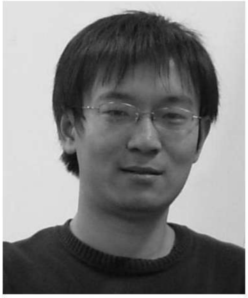
中国版本图书馆 CIP 数据核字（2017）第 241008 号

出版发行 / 北京理工大学出版社有限责任公司
社　　址 / 北京市海淀区中关村南大街 5 号
邮　　编 / 100081
电　　话 /（010）68914775（总编室）
　　　　　（010）82562903（教材售后服务热线）
　　　　　（010）68948351（其他图书服务热线）
网　　址 / http://www.bitpress.com.cn
经　　销 / 全国各地新华书店
印　　刷 /
开　　本 / 710 毫米×1000 毫米　1/16
彩　　插 / 2
印　　张 / 23
字　　数 / 447 千字
版　　次 / 2017 年 9 月第 1 版　2017 年 9 月第 1 次印刷
定　　价 / 76.00 元

责任编辑 / 梁铜华
文案编辑 / 郭贵娟
责任校对 / 周瑞红
责任印制 / 李志强

图书出现印装质量问题，请拨打售后服务热线，本社负责调换

作 者 简 介



辛斌，男，1982 年 1 月生，山东海阳人，博士。2004 年在北京理工大学电子工程系获信息工程学士学位，2012 年在北京理工大学自动化学院获模式识别与智能系统专业博士学位，毕业后留校任教，2015 年晋升副教授，2017 年晋升博士生导师。博士学位论文获 2013 年北京市优秀博士学位论文，并获 2014 年中国自动化学会首届优秀博士学位论文。现担任国际期刊 *JACIII* 的副主编，北京自动化学会理事，中国自动化学会会员、人工智能学会会员、指挥与控制学会会员、IEEE 计算智能学会会员、IEEE 控制系统学会会员、IEEE 系统–人–控制论学会会员。

主要研究方向为智能优化的理论和方法及其在先进火力与指挥控制系统中的应用，在 *IEEE-T-EC*、*IEEE-T-Cybern*、*IEEE-T-SMCA*、*IEEE-T-SMCC*、*IEEE-T-AES*、*ASOC*、*EO*、*JACIII*、*SCI CHINA INF SCI*、《自动化学报》、《控制理论与应用》等国内外著名学术刊物与会议上发表学术论文 60 余篇，其中 SCI 收录论文 17 篇，EI 收录论文 50 篇，出版学术专著 3 部。获教育部自然科学奖一等奖 1 项、国防科技进步二等奖 4 项，获第 18 届关肇直奖、国际会议 ISCIIA 2016 的青年学者奖。获评为国际期刊 *IEEE-T-Cybern* 2016 年度杰出评阅人。

陈杰，男，1965 年 7 月生，福建福清人，博士。北京理工大学教授、博士生导师。1986 年、1996 年、2001 年在北京理工大学自动化系分别获学士、硕士和博士学位。“复杂系统智能控制与决策”国家重点实验室主任、国家杰出青年科学基金获得者、教育部长江学者奖励计划特聘教授、国家自然科学基金创新研究群体学术带头人、教育部创新团队带头人，973 项目首席，享受国务院特殊津贴，新世纪百千万人才工程国家级人才，全国优秀科技工作者，北京市优秀博士学位论文指导教师。现担任国务院学位委员会学科评议组控制科学与工程组成员、教育部科学技术委员会委员、中国自动化学会副理事长、控制理论专业委员会副主任、中国人工智能学会常务理事、《系统与控制》丛书副主编、国际刊物 *JSSC* 执行编辑，*IEEE-T-Cybern*、*Int J Robust & Nonlinear Control*、*SCI*

China Inf Sci、*CTT* 等国内外著名刊物的副主编。主要研究方向为复杂系统的多指标优化与协调控制。近年来，以第一完成人获国家自然科学二等奖 1 项、国家科技进步二等奖 2 项、省部级一等奖 3 项，出版专著和教材 4 部，发表学术论文 200 余篇，获授权发明专利 100 多项、软件著作权 30 多项。

前　言

进化计算研究的先驱、进化策略的提出者 Schwefel 教授曾指出："无论是关于工程学、经济学、管理学、数学、物理学，还是关于社会科学的现代杂志，几乎每一本都在它的主题索引上有'优化'的概念。"优化问题普遍存在于作战运筹与规划、企业生产、社会管理等不同领域的科学研究中。对各种优化问题的求解需求促生了各种优化方法，优化问题的多样性和复杂性使得我们难以通过单一的方法来有效地解决所有问题，因此优化方法也呈现出明显的多样性。另一方面，人类对问题求解性能的追求和设计思想的多样化也促进了优化方法的多元化发展。

与数学规划领域建立的各种传统优化方法相比，智能优化方法是一类受不同自然规律启发设计而成的优化方法，主要用于解决各种复杂难解的优化问题。科学研究和工程实践中的大量优化问题都呈现出大规模、变量混杂、高复杂度、多目标、强约束、非线性、多极值、不确定性、动态时变，甚至建模困难等难点，很多问题还带有较强的计算实时性要求，这使得传统方法往往难以适用。智能优化方法对问题性质几乎无任何要求，并具有全局搜索优化能力，尤其适合求解具有上述各种难解特征的复杂优化问题，作为一类具有很强通用性的优化方法被广泛应用于不同领域的工程实际中。

由于设计思想和具体应用的多样性，智能优化方法门类众多，研究成果非常分散，不利于初学者在短时间内系统地掌握这方面的方法和技术。因此，本书尽量挑选具有代表性的多种经典和热点智能优化方法向读者进行介绍，以求兼顾内容的基础性和前沿性。同时，本书融入作者多年来从事智能优化方法与理论研究的经验和成果，对典型智能优化方法以及共性理论进行了系统的介绍，并从先进火力与指挥控制系统的研究中提炼出多类典型优化问题。这些优化问题各具特点，反映了不同的求解难点和挑战性。本书以这些典型优化问题为例，详细介绍了如何把智能优化方法以及先进的设计思想融入具体问题的求解中。为了便于读者深入学习，各章之后附有思考题和相关的参考文献，书后附有术语的中英文对照表。

本书包括 7 章内容。第 1 章为绪论，论述了最优化问题的研究意义、优化研究中的基本概念与数学模型、优化问题的分类与求解方法运用原则，并介绍了智能优化的概念、方法分类以及发展历史。第 2 章首先介绍了伪随机数的生成方法，然后介绍了 8 类典型的智能优化方法（混沌优化算法、模拟退火算法、禁忌搜索

试读结束：需要全本请在线购买： www.ertongbook.com

算法、遗传算法、差分进化算法、分布估计算法、蚁群优化算法、粒子群优化算法），包括各种方法的基本概念和原理、步骤流程、典型的算法变体以及算法的扩展与改进设计等。第 3 章从智能优化方法的共性角度出发，从统一框架和共性理论层面对智能优化方法进行分析，介绍了算法的收敛性分析方法和作者关于一类共性核心问题——“探索–开发权衡问题”的研究。第 4 章介绍了一类典型的混合智能优化算法——“文化基因算法”，以及作者建立的一种通用的混合策略分类法，这种分类法可以作为高级混合优化算法的设计指导。第 5～7 章每一章都以一类具体的优化问题为研究对象和主题线索，具体介绍了智能优化方法的设计和应用，并对相关背景领域的研究进行了系统的介绍和分析。

本书大部分内容取自作者自身以及指导的研究生的科研成果，其中包括张兴博士以及博士生丁玉隆、丁舒忻、李娟、曾杰，硕士生王艺鹏、朱阳光等所做的研究工作。第 5 章由陈杰完成，其余章节由辛斌完成。第 2 章、第 5～7 章的内容由研究生参与撰写，包括丁玉隆、王艺鹏、朱阳光、李娟、陈璐、高冠强、高源、徐小桓、展娇杨、曾杰、漆鸣凤。最后由辛斌统一校阅、统调、定稿。在此感谢学生们的辛勤劳动和付出。由于作者水平有限，书中难免存在不足和错误，恳请广大读者批评指正，欢迎读者来信勘误和交流。

本书可供自动化、计算机、系统工程、信息处理、管理、应用数学等专业的教师以及相关领域的技术开发人员参考，也可作为相关专业的高年级本科生和研究生的教材。

最后，感谢北京理工大学研究生院为本书的出版创造了一个宝贵的机会，感谢北京理工大学出版社对本书的出版给予的支持和帮助，特别感谢王浩琪编辑的辛勤劳动。感谢作者的家庭成员在作者从事研究和撰写本书过程给予的关心和支持。感谢国家自然科学基金委员会信息科学部、教育部、北京市教育委员会的大力支持，本书的相关研究得到国家自然科学基金面上项目（61673058）、青年科学基金项目（61304215）、国家杰出青年科学基金（60925011）、国家自然科学基金重大国际合作研究项目（61120106010）、NSFC–浙江两化融合联合基金重点支持项目（U1209214）、教育部长江学者和创新团队支持计划（IRT1208）、高等学校博士点学科专项科研基金（20131101120033）、北京市优秀博士学位论文指导教师科技项目（20131000704）等项目的资助，在此表示衷心感谢。

辛斌　陈杰
E-mail: brucebin@bit.edu.cn
北京理工大学自动化学院
复杂系统智能控制与决策国家重点实验室
北京市智能机器人与系统高精尖创新中心
2017 年 5 月

目　录

第 1 章

绪　　论

1.1　最优化问题的研究意义

最优化问题（Optimization Problems）是从众多候选解中确定某种意义下的精确最优解的数学问题。在很多复杂实际问题的求解中，由于问题本身的复杂度和计算代价的限制，往往很难找到理论意义上的最优解，因此实际中通常会关注某种意义下的近似最优解或满意解。确定精确最优解、近似最优解或满意解的问题统称为优化问题。优化问题普遍存在于作战运筹与规划、企业生产、社会管理等不同领域的科学研究中。对于军队而言，从部队的日常训练到战时的作战筹划，从装备的采购、资源的部署与配置、物资的运输与保障，到作战过程中各种资源的分配与调度，所有环节都存在优化问题；对于企业而言，从产品的设计、研制、生产、存储，到产品的销售、运输、维护，整个供应链都涉及优化问题。进化计算领域的奠基人之一、进化策略的提出者 Schwefel 教授曾指出："无论是关于工程学、经济学、管理学、数学、物理学，还是关于社会科学的现代杂志，几乎每一本都在它的主题索引上有'优化'的概念。"

最优化问题与方法研究源于运筹学（Operations Research），通常又称为数学规划（Mathematical Programming），是运筹学的主要分支之一，因此最优化问题又称为数学规划问题。运筹学作为一门现代科学，是在第二次世界大战期间由英、美两国发展起来的。当时为了打败纳粹德国，迫切需要把各种紧缺的资源以有效的方式分配给各种不同的作战任务和军事活动，所以美国政府号召大批科学家运用科学手段来处理战略与战术问题，这些科学家成立的研究小组正是最早的运筹小组。第二次世界大战后，运筹学的研究方法广泛拓展到经济与社会管理的不同领域中，因此一般意义下的运筹学常被描述为对组织系统的各种运作（Operations）做出决策的科学手段。它使用许多数学工具和逻辑判断的方法来研究系统中人、财、物的组织管理，筹划调度等问题，以期发挥最大效益。而数学规划研究可以追溯至 1939 年，当时苏联的康托洛维奇（L. Kantorovich）等在生产组织管理和制定交通运输方案方面首先研究和应用了线性规划方法。1947 年，美国数学家丹齐格（G. B. Dantzig）提出的求解线性规划的单纯形法，为这门学科奠定了基础。

试读结束：需要全本请在线购买： www.ertongbook.com

同时，电子计算机的出现和日益完善使数学规划研究得到迅速发展，电子计算机可用于处理包含成千上万个约束条件和变量的大规模线性规划问题，从单纯的工程技术开发到工业、农业、商业、交通运输业以及决策分析部门都可以发挥作用。非线性规划的基础性工作则是在 1951 年由库恩（H. W. Kuhn）和塔克（A. W. Tucker）等完成的。到了 20 世纪 70 年代，数学规划无论是在理论和方法上，还是在应用的深度和广度上都得到进一步的发展。

关于最优化问题的研究，在更早的时候，大数学家费马（Fermat）、拉格朗日（Lagrange）就提出基于微积分的公式来确定函数的最优解，牛顿（Newton）、高斯（Gauss）则建立了以渐近方式得到最优解的数值迭代方法。在中国古代，虽然没有形成系统的科学研究，但运筹规划的思想很早就有，例如田忌赛马、丁渭主持皇宫修复等故事都明确地体现了运筹规划对产生正确的决策或设计方案的重要作用。1955 年，我国学者从《史记 • 高祖本纪》中的名言“运筹帷幄之中，决胜千里之外”中摘取“运筹”二字，将“Operations Research”正式译作“运筹学”。

1.2 优化问题的基本概念与数学模型

最优化（Optimization）是指确定规划或设计方案使问题的目标最佳化的方法和过程。它需要根据选定的性能指标去评价和分析不同方案的有效性、成本、效益和效率等，同时还要考虑方案是否满足给定的要求和各种现实条件。因此，优化问题可以归结为 3 个要素：解、目标和约束。

（1）解（Solution）：即问题的解决方案，包括解决问题的步骤、操作和（或）设计实现的相关参数设置。在不同情形下，又称为决策方案、规划方案或设计方案等。为了便于分析和求解，在进行问题求解之前需要先根据问题的性质建立解的适当表达。例如，函数优化问题通常以连续变量的向量形式表达解，而很多组合优化问题则以符号的排列顺序（Permutation）来表达解。从几何角度看，可以把一个解抽象为问题解空间中的一个点。

（2）目标（Objective）：指用于定量评价解的质量（优劣）的指标，通常采用函数形式，故又称为目标函数（Objective Function）或指标函数（Index Function）。根据目标的性质和优化问题的特定背景，可以形成不同类型的目标，例如利润函数（Profit Function）、收益函数（Payoff Function）、代价函数（Cost Function）和能量函数（Energy Function）等。其中，利润函数是经济学中刻画企业利润的目标函数，收益函数是博弈论中用于刻画博弈方获得收益的目标函数，代价函数普遍用于度量某种解决方案带来的各种成本（例如时间、费用等），能量函数则常见于物理学、化学、分子生物学甚至图像处理等领域。

（3）约束（Constraint）：指限定解的可行性的条件，通常以关于解或解分量

的不等式或等式的形式表示。根据约束是否必须满足，可以分为硬约束和软约束；根据约束满足的难易程度，可以分为强约束和弱约束。

优化问题研究的目标是确定能够使目标函数达到最佳值的可行解。根据优化问题的性质，最佳目标函数值意味着目标函数的最小化或最大化。但从优化问题的表达形式角度看，最大化问题和最小化问题可以相互转换，因为 $\min f = \max\{-f\}$。因此，通常采用以下形式建立优化问题的数学表达：

$$\min \boldsymbol{F}(\boldsymbol{x}),\ \text{s.t. } \boldsymbol{G}(\boldsymbol{x}) \leqslant 0,\ \boldsymbol{H}(\boldsymbol{x}) = 0 \tag{1.1}$$

式中，$\boldsymbol{x}$ 为优化问题的解；$\boldsymbol{F}(\boldsymbol{x})$ 为目标函数；min $\boldsymbol{F}(\boldsymbol{x})$ 为最小化目标函数（min 是 minimize 的缩写，表示“使…最小化”）；s.t.为 subject to 的缩写，表示“受限于……”；$\boldsymbol{G}(\boldsymbol{x})$ 为不等式约束函数；$\boldsymbol{H}(\boldsymbol{x})$ 为等式约束函数。$\boldsymbol{F}(\boldsymbol{x})$，$\boldsymbol{G}(\boldsymbol{x})$ 和 $\boldsymbol{H}(\boldsymbol{x})$ 可以是标量函数，也可以是向量函数。

在连续变量优化研究中，下面的单目标无约束优化问题是研究各种复杂优化问题的基础：

$$\min f(\boldsymbol{x}),\ \text{s.t. } x \in X \tag{1.2}$$

式中，$\boldsymbol{x} = [x_1, x_2, \cdots, x_n] \in X \subset \mathbb{R}^n$；$X$ 为 $\mathbb{R}^n$ 上的开集，通常 $X = \mathbb{R}^n$。

最简单的优化问题就是评价指标唯一且计算方法正确可靠的情况下少数方案的选择（决策）问题，可以根据计算出的指标值直接评价出方案的优劣，进而做出最终选择，这类问题普遍存在于现实生活中。当候选方案（问题的解）数量巨大，难以在可接受时间内对所有解实现穷举，而同时又无法通过与问题相关的知识来确定最优解时，就不得不考虑对解空间的搜索策略，这是一般意义下的优化问题求解，即通过迭代搜索寻找最优解。但是，在数学意义上有可能存在没有最优解的情况，例如函数 $\min f(x) = -x^2$ 在开区间（−1, 1）上没有最优解。另外，也存在最优解对应于无穷值的情况，例如函数 $\min f(x) = -1/x^2$ 在闭区间 $[-1, 1]$ 上的最优解为 $x^* = 0$，对应的最优值为 $f(x^*) = -\infty$。根据 Weierstrass 函数定理，连续函数在有界闭区域内必有最大值和最小值。对于不连续函数，如果其在闭区域内有下确界（上确界），且下确界（上确界）等于某个点处的函数值，那么该函数就有最小值（最大值）。而在实际优化问题中，由于建模误差和各种不确定性等因素，某种数学意义下的最优解未必是实际问题的最优解，因此在工程优化中通常并不严格地追求数学意义上的最优解。关于最优解的存在性问题，通常假设实际优化问题都存在最优解。

从一般意义上讲，优化研究（Optimization Research）是一个由“问题”（Problem）、“方法”（Methodology）和“人”（Human）构成的三元体系。该领域内的研究存在“问题驱动”（Problem-driven）和“方法驱动”（Method-driven）两种典型范式，前者以具体问题求解为出发点，侧重具体问题的高效求解，而后者以方法设计为出发点，侧重方法的通用性和高效性。虽然问题和方法这两个层面

的研究更加客观，但“人”的主观因素对优化研究的影响也是不可忽略的。其重要性至少反映在以下 3 个方面：

1）优化问题的建模依赖于人的领域知识

建模是优化问题研究的前提，甚至在很多情况下需要把建模与优化过程作为一个整体来统一考虑。在优化问题的模型（包括目标、约束，甚至解的定义、表达和计算方法等）难以精确建立的情况下，经常利用人的知识和经验来构建近似的优化问题模型（例如基于规则的推理模型），甚至直接由人来对解做出评价和选择（例如在很多与人类的审美或感性相关的设计优化中）[1]。另外，在某些情况下，即使优化问题的模型已知，但为了便于问题的求解，往往需要通过近似或转换的方法对优化问题的模型进行处理，而这种方法也明显依赖于人对优化问题的分析和理解。因此，人的知识和经验将直接影响优化问题模型的合理性和准确性。由于优化研究的根本目标是为决策者制定决策提供定量的参考依据，因此人对优化问题建模过程的影响也必然会对最终优化结果的可信度产生影响。

2）优化方法的设计依赖于人的智慧和经验

在一般意义上很难对“什么样的优化问题的最优求解方法是什么”这一问题做出准确的回答。而且，对于复杂优化问题的求解，知识的运用是提高优化方法效率的关键，因此在算法设计时往往需要依赖人的精妙思维和灵活设计才能实现优化问题的高效求解。目前，在优化问题的分类谱系里（见 1.3 节），人类只对数量有限的几类特殊优化问题建立了理论相对完善的高效算法，例如求解线性规划问题的单纯形法、求解凸优化（Convex Optimization）问题的内点方法（Interior Point Method）等。而对于很多困难的优化问题，至今也没有找到完善的方法。很多优化算法的控制参数、停止准则都需人为设定。

3）多目标优化涉及人的偏好

实际中的很多优化问题涉及多种评价指标，例如成本和收益，因此一般属于多目标优化问题。在这种情况下，多个目标之间往往存在矛盾，不能同时实现最优。因此，多目标优化问题的“最优解”最终需要通过决策者对多个指标做出权衡来确定，这种情况下的最优解被称为“最佳偏好解”（Most Preferred Solution）。由于最后的决策过程涉及决策者的主观偏好，因此属于决策科学和认知科学的范畴。在交互式多目标优化研究中，决策者的偏好表达与利用是算法设计的关键问题之一。

下面以目标函数最小化问题为例给出优化研究中经常用到的一些基本概念。

定义 1.1　可行解。即满足所有约束的解，对于式（1.1）表示的一般优化问题，可行解可以表示为 $\boldsymbol{x} \in S = \{\boldsymbol{x} \mid \boldsymbol{G}(\boldsymbol{x}) \leqslant 0, \boldsymbol{H}(\boldsymbol{x}) = 0\}$，其中 S 为可行域，即所有可行解构成的子空间。只满足部分约束的解称为部分可行解。

定义 1.2　最优可行解。即一般意义上的最优解，可表示为 $\boldsymbol{x}^* = \underset{x \in S}{\arg\min}\{f(\boldsymbol{x})\}$，

其中 S 为可行域，最优可行解即全局最优解。

定义 1.3　近似最优解。即使目标函数值接近最小值的可行解，可表示为 $\boldsymbol{x}' \in \{\boldsymbol{x} \mid f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \varepsilon, \boldsymbol{x} \in S\}$，其中 $\varepsilon > 0$ 是一个较小的正数。

定义 1.4　满意解。即使目标函数值达到满意水平的可行解，可表示为 $x^\# \in \{\boldsymbol{x} \mid f(\boldsymbol{x}) \leqslant f_\varepsilon, \boldsymbol{x} \in S\}$，其中 f_ε 是可接受的目标函数值水平。

定义 1.5　局部最优解。即在解所属的某个邻域内使目标函数值达到最小的解，可表示为 $\boldsymbol{x}_{\mathrm{L}}^* = \arg\min\limits_{x \in N}\{f(\boldsymbol{x})\}$，其中 $N \in S$ 是解空间的一个子集。

定义 1.6　局部极值解。即在任意小的可行邻域内都能取得该邻域内最小值的解。局部极值解是连续变量优化中的概念，是局部最优解的特殊情况。用以当前解 x 为中心、ε 为半径的球 $B_\varepsilon(\boldsymbol{x})$ 来表达邻域，即 $B_\varepsilon(\boldsymbol{x}) = \{\boldsymbol{y} \mid d(\boldsymbol{x}, \boldsymbol{y}) < \varepsilon\}$，其中 $d(\boldsymbol{x}, \boldsymbol{y})$ 表示两个解 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的欧氏距离。如果 $\exists \varepsilon > 0$ 使得 $f(\boldsymbol{x}_{\mathrm{E}}^*) = \min\limits_{x \in B_\varepsilon(\boldsymbol{x}_{\mathrm{E}}^*) \cap S}\{f(\boldsymbol{x})\}$，则称 $\boldsymbol{x}_{\mathrm{E}}^*$ 为局部极值解（又称局部极小解）。对于最大化问题而言，局部最优解为在任意小的可行邻域内都能取得该邻域内最大值的解。

定义 1.7　Pareto 最优解。即多目标优化意义下的一种最优解，又称为非可支配解（Non-dominated Solution）、有效解（Efficient Solution）、非劣解（Non-inferior Solution）等。对于多目标优化问题 $\min \boldsymbol{F}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x})]$，如果不存在解 $\boldsymbol{x}$ 能够使 $\forall i \in \{1,2,\cdots,m\}$，$f_i(\boldsymbol{x}) \leqslant f_i(\boldsymbol{x}_{\mathrm{P}}^*)$ 且 $\exists j \in \{1,2,\cdots,m\}$ 使得 $f_j(\boldsymbol{x}) < f_j(\boldsymbol{x}_{\mathrm{P}}^*)$，则称 $\boldsymbol{x}_{\mathrm{P}}^*$ 为 Pareto 最优解。Pareto 最优性（Pareto Optimality）意味着不存在能够支配（Dominate）当前解的解，即不存在所有目标都不劣于当前解且至少一个目标比当前解更优的解。所有的 Pareto 最优解构成 Pareto 解集。Pareto 解集在目标空间中的映像称为 Pareto 前沿（Pareto Front）。

定义 1.8　凸集（Convex Set）。一个集合 C 是凸集，当且仅当 $\forall \boldsymbol{x}, \boldsymbol{y} \in C$，$\forall \lambda \in [0,1]$：$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C$。

定义 1.9　凸函数（Convex Function）。X 是实向量空间上的凸集，f：$X \to \mathbb{R}$ 是一个函数。f 为凸函数，当且仅当 $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in X$，$\forall \lambda \in [0,1]$：$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leqslant \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$。

f 为严格凸函数，当且仅当 $\forall \boldsymbol{x}_1 \neq \boldsymbol{x}_2 \in X$，$\forall \lambda \in (0,1)$：$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) < \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$。

定义 1.10　凸优化问题（Convex Optimization Problem）。凸集上的凸函数的最小化问题称为凸优化问题。

对于优化问题 $\min f(\boldsymbol{x})$, s.t. $g_i(\boldsymbol{x}) \leqslant 0, i = 1,2,\cdots,l; h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,k$，如果 $f, g_1, g_2, \cdots, g_l$：$\mathbb{R}^n \to \mathbb{R}$ 都是凸函数，$h_1, h_2, \cdots, h_k$：$\mathbb{R}^n \to \mathbb{R}$ 都是仿射函数，则该问题为凸优化问题。凸优化问题的任何局部极值解都是全局最优解。对于严格凸函数，如果优化问题有局部极值解，那么这个局部极值解是唯一的。

定义 1.11 下降方向（Descent Direction）。向量 $\boldsymbol{d}$ 是 $\boldsymbol{x}=\boldsymbol{x}_0$ 处的下降方向，当且仅当对于任意小的正数 ε，存在 $f(\boldsymbol{x}+\varepsilon\boldsymbol{d})<f(\boldsymbol{x})$。

定义 1.12 梯度（Gradient）。在连续变量优化中，梯度用于刻画函数的最速变化方向。正梯度表示最速上升方向，负梯度表示最速下降方向。对于优化问题（1.2），假设目标函数 $f(\boldsymbol{x})$ 在解空间 X 上处处可微，则函数 $f(\boldsymbol{x})$ 在解 $\boldsymbol{x}_0$ 处的梯度可以定义为：

$$\nabla f(\boldsymbol{x})\Big|_{\boldsymbol{x}=\boldsymbol{x}_0}=\left[\frac{\partial f(\boldsymbol{x})}{\partial x_1},\frac{\partial f(\boldsymbol{x})}{\partial x_2},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right]^{\mathrm{T}}\Bigg|_{\boldsymbol{x}=\boldsymbol{x}_0}$$

相对应地，负梯度方向为 $-\nabla f(x)$。

定义 1.13 次梯度（Subgradient）。假设 f：$X\to\mathbb{R}$ 是一个实变量凸函数，且 $X\in\mathbb{R}^n$ 是一个开凸集，如果对于 X 内的所有 $\boldsymbol{x}$，都有 $f(\boldsymbol{x})-f(\boldsymbol{x}_0)\geqslant\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_0)$，则向量 $\boldsymbol{v}$ 称为函数在点 $\boldsymbol{x}_0$ 处的次梯度。所有次梯度的集合称为次微分，记为 $\partial f(\boldsymbol{x}_0)$。次微分总是非空的凸紧集。与梯度相比，次梯度适用于不可微的函数。

定义 1.14 Jacobian 矩阵。假设 $\boldsymbol{f}$：$\mathbb{R}^n\to\mathbb{R}^m$ 是一个向量函数，则其 Jacobian 矩阵定义为：

$$\boldsymbol{J}=\left[\frac{\partial \boldsymbol{f}}{\partial x_1},\frac{\partial \boldsymbol{f}}{\partial x_2},\cdots,\frac{\partial \boldsymbol{f}}{\partial x_n}\right]=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}\end{bmatrix}$$

定义 1.15 Hessian 矩阵。在连续变量优化中，Hessian 矩阵用于描述多变量函数的局部曲率。假设目标函数 $f(\boldsymbol{x})$ 的所有二阶偏导数都存在且在解空间 X 上连续，则 Hessian 矩阵表示为：

$$\boldsymbol{H}_{\mathrm{e}}=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{bmatrix}$$

Hessian 矩阵与 Jacobian 矩阵的关系为 $\boldsymbol{H}_{\mathrm{e}}=\boldsymbol{J}(\nabla f(\boldsymbol{x}))^{\mathrm{T}}$。

定义 1.16 正定矩阵。矩阵 $\boldsymbol{M}$ 为正定矩阵，当且仅当 $\forall\boldsymbol{x}\in\mathbb{R}^n$ 且 $\boldsymbol{x}\neq 0$，

$x^{\mathrm{T}} M x > 0$。

定义 1.17 Lipschitz 常数。对于函数 f：$X \to \mathbb{R}, X \subset \mathbb{R}^n$，如果对于任意 x_1，$x_2 \in X$，存在常数 L 使得 $|f(x_1) - f(x_2)| < L|x_1 - x_2|$，则称 L 为 $f(x)$ 在 X 上的 Lipschitz 常数。对于连续变量优化问题，Lipschitz 常数是一种重要的全局信息。

下面是特定的连续变量优化情况下最优解应满足的必要性条件，这些条件是建立解析方法，实现这些问题求解的重要依据，也是优化研究中的经典理论。

（1）**一阶最优性条件**。假设 $f(x)$ 在 $x = x_0$ 处可微，如果 x_0 是一个局部最小解，则 $\nabla f(x_0) = 0$。

（2）**二阶最优性条件**。假设 $f(x)$ 在 $x = x_0$ 处二次连续可微，如果 $\nabla f(x_0) = 0$ 且 $H(x_0)$ 正定，则 x 是局部极小解。

（3）**凸函数的最优性条件**。对于凸函数 f：$X \to \mathbb{R}$，$x = x_0$ 为全局最小解，当且仅当 $\nabla f(x_0) = 0$。

（4）**Karush–Kuhn–Tucker 条件**。简称 KKT 条件，又称 Kuhn–Tucker 条件，是非线性规划中某种正则性条件下最优解应满足的一阶必要条件。假设目标函数 f：$\mathbb{R}^n \to \mathbb{R}$，不等式约束函数 g_i：$\mathbb{R}^n \to \mathbb{R}$（$i = 1,2,\cdots,l$）和等式约束函数 h_j：$\mathbb{R}^n \to \mathbb{R}$（$j = 1,2,\cdots,k$）在某个点 x^* 处连续可微，如果 x^* 是一个局部最优解，且最优化问题满足某些正则性条件（如下所述），则存在常数向量 $\boldsymbol{\mu}=[\mu_1,\mu_2,\cdots,\mu_l]$ 和 $\boldsymbol{\lambda} = [\lambda_1,\lambda_2,\cdots,\lambda_k]$（称为 KKT 乘子），使得：

① 平稳性（Stationarity）：

对于最小化问题：$\nabla f(x^*) = \sum_{i=1}^{l} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{k} \lambda_j \nabla h_j(x^*)$。

对于最大化问题：$-\nabla f(x^*) = \sum_{i=1}^{l} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{k} \lambda_j \nabla h_j(x^*)$。

② 原始可行性（Primal Feasibility）：

$$g_i(x^*) \leqslant 0, i = 1,2,\cdots,l \text{；} \quad h_j(x^*) = 0, j = 1,2,\cdots,k$$

③ 对偶可行性（Dual Feasibility）：

$$\mu_i \geqslant 0, i = 1,2,\cdots,l$$

④ 互补松弛性（Complementary Slackness）：

$$\mu_i g_i(x^*) = 0, i = 1,2,\cdots,l$$

正则性条件实际上是对约束的限制，下面是两种常见的正则性条件：

① 线性条件（Linearity Condition）：

g_i：$\mathbb{R}^n \to \mathbb{R}$（$i = 1,2,\cdots,l$）和 h_j：$\mathbb{R}^n \to \mathbb{R}$（$j = 1,2,\cdots,k$）都是仿射函数。

② Slater 条件：

对于一个凸优化问题，存在一个解 x 使得：

$$g_i(x) \leqslant 0, i = 1,2,\cdots,l \text{；} \quad h_j(x) = 0, j = 1,2,\cdots,k$$

试读结束：需要全本请在线购买： www.ertongbook.com