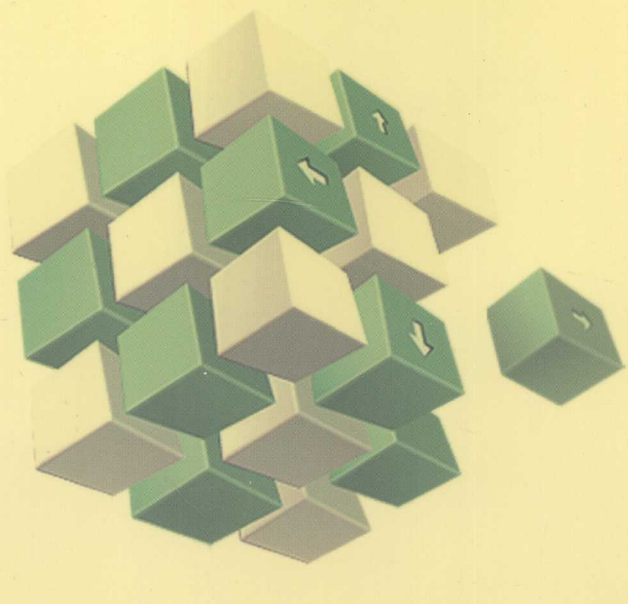
普通高等教育规划教材

运筹学与最优化MATLAB编程

吴祈宗 郑志勇 邓 伟 等编著



机械工业出版社
CHINA MACHINE PRESS

普通高等教育规划教材

运筹学与最优化 MATLAB 编程

吴祈宗　郑志勇　邓　伟　等编著

机械工业出版社

前　言

运筹学在自然科学、社会科学、工程技术、生产实践、经济建设及现代化管理中有着重要的意义。随着科学技术和社会经济建设的不断发展，运筹学得到了迅速的发展和广泛的应用。作为运筹学的重要组成部分——线性规划、非线性规划、目标规划、整数规划、层次分析法、遗传算法等内容成为管理、经济类以及大多数工科类学生所应具备的知识和学习其他相应课程的重要基础。本书根据管理、经济类以及大多数工科类学生知识结构的需要，利用 MATLAB 软件的特性，在理论知识与实际应用目标间建立桥梁。

本书是一本有关对运筹学与最优化理论、方法知识的理解、认识与提高的参考教材，可以独立使用。同时，它也可以作为《运筹学与最优化方法》（吴祈宗编著，机械工业出版社出版）的重要补充参考教材。由于运筹学与最优化方法涉及的数学基础较多，所以对于工科、管理、经济类的硕士研究生来说，完全从理论方面掌握这些必要的基础难度较大。考虑到运筹学与最优化方法理论性及应用性密切结合的特征，要学好这门课程必须注重对运筹学本质性知识的掌握，并需在实践中能够灵活运用这些运筹学与最优化方面的知识。

本书的编写原则是，注重教育思想和教育内容的改革，注意激发学生独立思考问题和创新的意识；把基础理论的研究、方法构造的思路、应用前景与利用 MATLAB 编程有机地结合起来；注重强调运筹学与实践的紧密联系，遵循“实践-理论-实践”的发展过程。

本书利用算法编程分析、算法语言实现、程序模块源码与教材文字配合，注意对学生知识结构的构建，把学科特征、较新发展成果、发展趋势与提高学生的研究、开创能力有机结合起来考虑，能使教学和自学收到较好的效果。

在教育、教学中，培养学生自学能力是十分重要的，本书在这方面作了有益的探索。本书在编写过程中，注意让学生感受、理解知识产生和发展的过程，培养学生的科学精神和创新思维习惯，重视培养学生收集处理信息的能力、获取新知识的能力、分析和解决问题的能力等。

本书的编著由吴祈宗、郑志勇、邓伟、侯福均、朱世慧、刘颖、陈明超、吴

明灯等多人协作完成。我们在编著过程中参考了大量的国内外有关文献，它们对本书的成文起了重要作用。在此对一切给予我们支持和帮助的朋友、同事、有关人员以及参考文献的作者一并表示衷心的感谢。

限于编著者水平，书中难免有不当或失误之处，敬请广大读者批评指正。

编　者

2009 年 7 月

目录

第1章 运筹学概述

运筹学（Operations Research，OR）作为科学名词出现在20世纪30年代末。第二次世界大战期间，运筹学的研究与应用范围主要是战略、战术方面。随着世界性战争的结束，各国开始快速发展经济，世界范围内的剧烈竞争也体现在经济、技术方面，运筹学的研究也向这些方面拓展。运筹学为了适应时代的要求，在近几十年中，无论从理论上还是应用上都得到了快速的发展。在应用方面，今天运筹学已经涉及了服务、管理、规划、决策、组织、生产、建设等诸多方面，甚至可以说，很难找出它涉不到的领域。在理论方面，由于运筹学的需要和刺激而发展起来的一些数学分支，如数学规划，应用概率与统计、应用组合数学、对策论、数理经济学、系统科学等，都得到了迅速发展。

运筹学是一门应用科学，很难给出一个确切的定义。根据运筹学工作者的一些论述，我们可以较深切地理解这门科学的内涵。运筹学工作的先驱、诺贝尔奖金获得者、英国著名物理学家 P. M. S. Blackett 在1940年就开始从事运筹学方面的研究与应用。他曾多次指出：运筹学的一个明显的特征，正如目前所实践的，是它有或应该有一个严格且实际的性质，其目标是帮助人们找出一些方法，来改进正在进行中的或计划在未来进行的作战效率。为了达到这一目的，要研究过去的作战来明确事实，要得出一些理论来解释事实，最后利用这些事实和理论对未来的作战作出预测。我们可以逻列出一些论述：运筹学是“为决策机构在对其控制下业务活动进行决策时，提供以数量化为基础的科学方法。”“运筹学是一门应用科学，它广泛运用现有的科学技术知识和数学方法，解决实际中提出的专门问题，为决策者选择最优决策提供定量依据。”“运筹学是一种给出问题的坏答案的艺术，否则，问题的结果会更坏。”实质上，运筹学的基本目的是找到“优”的方案、途径，而在实际中，最优只能是一种理想追求。由于问题的复杂性，各种确定与不确定因素的综合影响，运筹学目标准确（或有保障的）定位，应该是通过研究使我们避开更坏的结果。从哲学角度讲，我们可以说运筹学就是用科学方法去了解和解释运行系统的现象。它在自然界范围内所选择的研究对象就是这些系统。这些系统时常包含着人和在自然环境中运行的“机器”，这个广义的“机器”可以推广为按照公认的规则运行的复杂社会结构。

我们称运筹学是一门科学，是因为它用科学方法来创建其知识。它与其他科学的不同之处在于，它研究的是运行系统的现象，这在其他科学中往往被忽略。

根据上述讨论，我们可以体会到，运筹学的思想就是对我们所关心的对象（“机器”或运行系统）进行深入的了解、分析和研究，特别是进行量化分析，得到信息，再合理运用有关的数学工具、系统方法等进行研究，提出有效方案来解决问题。

1.1 运筹学的特点及其应用

1.1.1 朴素运筹学思想及其深刻内涵

自从 1956 年引入以来，运筹学在我国已有四十多年的历史。经过这四十多年，运筹学在我国有了很大的发展，确立了它在经济建设中的地位。但是，运筹学在我国的发展状况与世界其他国家相比，尚有不小的差距，其中最主要的是认识与基础的问题。人们公认，将“Operations Research”译为“运筹学”最恰当。我国历史上，在军事和科学技术方面对运筹思想的运用是世界著名的，从春秋战国和三国时期的战争中就可举出很多运用运筹思想取得战争胜利的例子。这反映出运筹学注意系统数据采集、分析并研究优化方案的思想是一种朴素、自然的思想。实际上，很多人都在自觉不自觉地运用这种思想。另一方面，我们常说“道高一尺，魔高一丈”，在竞争中，各方共同运用这些思想解决问题时，就表现为对运筹学内涵研究、运用能力的提升。随着科学技术的发展，特别是信息社会的到来，运筹学内涵不断扩大，涉及的数学及其他基础科学的知识越来越多，于是熟练掌握并运用这门学科有效解决实际问题的难度也逐渐加大。根据运筹学的发展，数学、计算机科学及其他新兴学科的最新知识、技术都能很快融入其中，特别是人的直接参与决策，使得运筹学发展更进入一个崭新的阶段。这一切与我国科技发展的现状决定了我们必须加强对运筹学的研究与应用，只有这样才能逐步缩短与世界发达国家的距离。

1.1.2 运筹学研究的工作步骤

为了有效地应用运筹学，应当遵循下列六条原则：合伙原则、催化原则、独立原则、互相渗透原则、宽容原则和平衡原则。这些原则反映了运筹学工作者与其他各种因素的横向和纵向的联系。运筹学研究的工作步骤可以归纳为以下九个内容：

（1）目标的确定。即确定决策者期望从方案中得到什么。这个目标不应限制在过分狭小的范围内，但也要避免不必要的扩大；目标可以是一个，也可以是多个，随着和谐社会的发展，实际决策中都要考虑各个群体的利益，以至于多目

标规划目前在现实生活中越来越多地被使用了。

（2）方案计划的研制。实施一项运筹学研究的过程常常是一个创造性的过程，计划的实质是规定出要完成某些子任务的时间，然后创造性地按时完成这一系列任务。这样做能够推动运筹学分析者得出结论，有助于方案的成功。对计划的任意延期和误时会导致分析者的消极工作和管理者的漠不关心。

（3）问题的表述。这项工作的开展需要与管理人员进行深入讨论，经常包括与其他职员和业务人员的接触，采集必要的数据，以便了解问题的本质、历史及未来，以及问题各个变量之间的关系。这项任务的目的是为研究中的问题提供一个模型框架，并为以后的工作确立方向。在这里，第一要考虑问题是否能够分解为若干串行或并行的子问题；第二要确定模型建立的细节，如问题尺度的确定、可控制决策变量的确定、不可控制状态变量的确定、有效性度量的确定和各类参数、常数的确定。

（4）模型的研制。模型是对各变量关系的描述，是成功解决问题的关键。构成模型的关系有几种类型，常用的有定义关系、经验关系和规范关系。

（5）计算手段的拟定。在模型研制的同时，需要研究如何用数值方法求解模型，其中包括对问题变量性质（确定性、随机性、模糊性）、关系特征（线性、非线性）、手段（模拟、优化）及使用方法（现有的、新构造的）等的确定。例如，对于一个模型可能有许多算法可以适用，有的算法计算时间长但计算结果质量高，有的算法计算时间快但计算结果的质量相对较差，如何根据实际需要选择有效的算法，也是运筹学问题求解中重要的一环。

（6）程序明细表的编制，程序设计和调试。对于计算过程需要编制程序来实现计算机运算的，运筹学研究应包含算法过程的描述、计算流程框图的绘制。程序的实现及调试可以交由程序员完成，或会同程序员完成。

（7）数据收集。把有效性试验和实行方案所需的数据收集起来加以分析，研究输入的灵敏性，从而可以更准确地估计得到的结果。

（8）方案验证。验证在运筹学的研究与应用中的重要性无论怎样强调都不会过分。验证包括两个方面：第一是确定验证模型，包括为验证一致性、灵敏性、似然性和工作能力而设计的分析与实验；第二是验证的进行，即把前一步收集到的数据用来对模型作完全试验。这样一种试验的结果，往往使模型必须重新设计，并要求与之相联系的程序重编。

（9）方案实施。运筹学分析者往往认为，在模型验证后，任务就完成了，这是不对的。事实上，一项研究的真正困难往往在方案的最后一步，即在实施和维护时才暴露出来。因此要使得整个研究有效，必须和那些与所研究的决策问题相关的各职能的各级管理人员进行合作，并让他们参与其中。

1.2 运筹学建模

1.2.1 运筹学建模的一般思路

运筹学建模在理论上应属于数学建模的一个部分。因此，运筹学建模所采用的手段、途径就是数学建模中所采用的。本节所要介绍的，是根据运筹学本身的特征来处理建模问题的一般思路。

经过长期、深入的研究和发展，把运筹学处理的问题归纳成一系列具有较强背景和规范特征的典型问题。因此，运筹学建模就要把相当的精力放在将实际问题合理地描述为某种典型的运筹模型。在这个过程中，要求运筹学工作者具有以下几个方面的知识和能力：

（1）熟悉典型运筹模型的特征和它的应用背景，如线性规划、非线性规划、整数规划、层次分析法（AHP）等。

（2）有理解实际问题的能力，包括广博的知识、搜集信息、资料和数据的能力。

（3）有抽象分析问题的能力，包括抓主要矛盾、逻辑思维、推理、归纳、联想、类比等创造能力。

（4）有运用各类工具知识的能力，包括运用数学知识、计算机、自然科学和工程技术等的能力。

（5）有试验校正、维护修正模型的能力。

根据问题本身的情况，按照上节的讨论，我们在建模时一般有如下思路：

（1）直接方法。当我们熟悉问题的内在关系、特征以及运筹学的典型模型特点时，常常可以直接得到一些问题的模型或问题归类，即确定问题是属于线性规划、非线性规划、整数规划、排队模型等的哪一种。有时模型的参数也可直接从问题本身得到。

（2）类比方法。通过类比把新遇到的问题用已知类似问题的模型来建立该问题的模型。这种情况往往得到的是模型归类，而模型参数需用其他方法取得。

（3）模拟方法。利用计算机程序实现对问题的实际运行模拟，可得到有用的数据。这些数据常用来求得模型参数或对所建立模型的合理性、正确性的检验。

（4）数据分析法。利用数据处理的方法分析各数据变量之间的关系是确定关系还是相关关系，以及是何种相关等。这种方法还可以用回归分析找出变量的变化趋势，从而得到合理的数学模型。大量的模型参数求得也常常使用数据处理

的统计方法。另外，回归模型常常就是一个无约束最优化模型。

(5) 试验分析法。通过试验分析建模是工程管理中常用的方法。以局部的试验产生数据，经过统计处理得到总体的模型或模型归类。试验分析更多地用于产生模型参数。

1.2.2 运筹学模型的评价

一个好的运筹学模型，不仅能比较真实地反映实际问题，还要具备以下几个优点：

(1) 易于理解。模型应力求简明。这里要强调一点，模型越大越复杂，不一定意味着越好。应当把实际问题中那些不重要的因素删去。这样，一方面，形成模型以后，由于变量和约束个数较少，便于计算求解；另一方面，也更易于揭示主要因素对问题的影响以及它们之间的关系。

模型中的变元、函数符号要接近所代表的实际因素、资源和目标的原义，这样易于理解模型所表示的实际问题的结构，而且当变量和约束个数很多时，对于模型建立和解释也是便利的。

(2) 易于探查错误。如果上面一点做得比较好，那么模型也易于探查错误。模型的错误一般有两种：①书写错误；②模型与实际问题不符。后一种错误在建立模型时应尽量避免，在评价模型及其解时，也可以找出错误并改正。前一种错误的避免，一方面要求细心，另一方面要求模型的书写形式要规范，变量次序最好固定不变。如果在某个函数关系中不出现某个变量，那么该变量的位置最好以空白代替；在约束条件中，把等式约束与不等式约束分成两组，分开来写，而在不等式约束中又可以把“≥”不等式与“≤”不等式分开来写。为便于清楚地表明每个约束所代表的对实际资源的限制，不必把两种形式的不等式统一成一种。

(3) 易于计算。运筹学模型问题是否易于求解，取决于问题的规模、复杂程度、当前的计算技术水平和解该问题的算法。可以通过以下几个途径降低问题的复杂程度和规模：

1) 采用简单的函数。与其在建模时花费许多时间、精力，寻找很好地反映现实情况但非常复杂的函数，不如在误差允许的范围内，采用简单的函数。在这一点上，尤其应该注意的是，非线性函数在计算方面的困难远比线性函数大得多。

2) 删去不必要的变量和约束条件。在建立模型前，分析问题时就应该注意把那些不重要的因素和资源限制简化掉。建立模型后，则应注意删去那些多余的约束条件。多余的约束条件是指该约束条件被删去后，可行解集没有改变。一个

简单的例子是：$t<1$ 和 $t<6$，这里 $6>1$，那么条件 $t<6$ 就是多余的。尤其要删去那些非线性的约束条件。但是没有实用的一般方法能识别模型中的多余约束条件。尽管如此，在求解运筹学模型前，对它进行一番数学上的分析、讨论它的性质仍是非常必要的。

3）函数变换。通过将复杂的函数进行代数处理，可以达到降低模型复杂程度的目的。

例如：

$$(1)\ \text{s.t.}\begin{cases}(x_1+x_2)^2-(x_1-x_2)^2\leqslant 10\\ x_1,\ x_2\leqslant 0\end{cases}\Rightarrow \text{s.t.}\begin{cases}4x_1x_2\leqslant 10\\ x_1,\ x_2\leqslant 0\end{cases}$$

$$(2)\ \prod_n x_n^{a_n}\leqslant c\Rightarrow \sum_n a_n\ln x_n\leqslant \ln c$$

上面两个例子分别根据因式分解方法和连乘函数转为对数的连加函数的方法，从计算的角度对约束函数进行了简化。

1.2.3 运筹学模型的求解

解决一类优化问题，一般会有多种算法可供选择，某些特别的问题也有专门的算法处理。某个算法对某类问题特别有效，但对于其他问题也许根本不起作用。建立模型时，要注意到哪些算法对求解该类问题是有效的，在计算误差与计算时间允许的范围内，选择那些相对比较有效的算法去求解该问题。因此，模型建立者应对优化算法进行系统的了解，熟悉每种算法的优势与缺点。例如以下两类问题对算法的速度要求较高：

例 1-1：投资组合优化配置就是优化算法在金融中的主要应用。由于市场的行情时刻在变化，原有投资组合在当前的市场情况中是否最优，直接决定着该组合的收益。对于投资组合包含投资品种较多的时候（例如 J. P. 摩根管理着四千多种资产），优化问题的计算量巨大，在有限的时间内得到全局最优似乎不可能，而且市场的信息多数也是不精确的，对于这些不精确的参数即使求出全局最优解，解的实际精度也会遭到质疑。如果将投资组合优化系统做成实时系统似乎也不太现实，一般只要求在原有投资组合的配置下根据市场变化有所改进就可接受。

例 1-2：炼油设备中的压力控制系统在相关参数基本确定的情况下，管道内压力的大小决定着炼油效率的高低，但压力的变化又同时影响着其他参数。这样的一个压力优化控制系统必须是实时的，但该优化问题的目标函数中涉及了复杂的流体力学的偏微分方程求解问题。这类问题不可能在短时间内求得最优解，只能在规定的反应时间内求得较有效解。

在实践中，最优化算法是求解运筹学问题时使用的主要算法。根据优化算法理论发展与算法原型将现有的最优化算法分为：经典优化算法和启发式优化算法两大类。经典优化算法和启发式优化算法都是迭代算法，但是，它们又有很大区别，如：

（1）经典算法以一个可行解为迭代的初始值，而启发式算法以一组可行解为初始值。

（2）经典算法的搜索策略为确定型的，而启发式算法的搜索策略是结构化和随机化的。

（3）经典算法大多都需要导数信息，而启发式算法仅用到目标函数值的信息。

（4）经典算法对函数性质有着严格要求，而启发式算法对函数性质没有太大要求。

（5）经典算法的计算量要比启发式算法小很多，比如，对于规模较大且函数性质比较差的优化问题，经典算法的效果不好，但一般的启发式算法的计算量太大。

在计算时，优化算法的迭代过程主要由找寻搜索方向和确定搜索步长组成。搜索方向和搜索步长的选取决定了优化算法的搜索广度和搜索深度。经典优化算法和启发式优化算法的区别主要是其搜索机制不同。经典算法的搜索方向和搜索步长是由局部信息（如导数）决定的，所以一般只能对局部进行有效的深度搜索，而不能进行有效的广度搜索，所以经典优化算法很难跳出局部最优；启发式优化算法为了避免像经典优化算法那样陷入局部最优，采用了相对有效的广度搜索，但是在问题规模较大的时候，这样做的结果往往使得计算量难以承受。

纵观优化算法的发展，完美的算法是不存在的。常用来评价算法好坏的标准有：

（1）算法收敛速度。

（2）算法使用范围（普适性）。

（3）算法的时间复杂度。

（4）算法得到解的质量（局部性或全局性，对绝对最优解的近似程度）。

（5）算法的可实现性。

可以说这些标准是不可公度的（不可能同时都好）。以全局最优问题为例，如果要求计算时间少，那么搜索广度就无法保证，解的质量就差；如果要求收敛速度快，就需要有效的搜索方向，有了搜索方向就降低了搜索广度，这样解的全局最优性无法保证。

计算复杂性理论，全局优化问题是 NP-complete 问题，一般根据实际问题采

用不同的启发式算法。算法的评价标准不可公度，而且在具体问题中，这些标准也不是等重要的。比如某些问题对解的要求降低 10%，它的计算时间就可以减少 50%，这样做是否值得，要根据实际情况而定。

1.3 基本概念和符号

1.3.1 空间与向量

本书的算法理论是建立在 n 维欧氏空间基础上的，书中使用下列符号：

$\mathbf{R}^n$ 表示 n 维欧氏空间：$\boldsymbol{x}\in\mathbf{R}^n$，表示 $\boldsymbol{x}$ 为 $\mathbf{R}^n$ 中的向量，若不进行特殊说明，本书中 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，其中 x_i 为 $\boldsymbol{x}$ 的第 i 维分量，符号“T”表示转置。

设 $\boldsymbol{x}$，$\boldsymbol{y}\in\mathbf{R}^n$，$\boldsymbol{x}$，$\boldsymbol{y}$ 的内积记为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=\sum\limits_{i=1}^{n}x_iy_i$，在不加说明时，我们用 $\|\cdot\|$ 表示 2 范数，$\|\boldsymbol{x}\|=\sqrt{\sum\limits_{i=1}^{n}x_i^2}$。

1.3.2 梯度向量与 Hesse 矩阵

设 $f(\boldsymbol{x}):\mathbf{R}^n\to\mathbf{R}$，$f$ 二次可微。

定义：设 $f(\boldsymbol{x})$ 在 $S\subset\mathbf{R}^n$ 上有定义，$\boldsymbol{x}^{(0)}\in\mathrm{int}S$，若 $\exists p\in\mathbf{R}^n$，使得 $\forall\boldsymbol{x}\in S$，有：

$$f(\boldsymbol{x})=f(\boldsymbol{x}^{(0)})+\boldsymbol{p}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^{(0)})+\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|\alpha(\boldsymbol{x}-\boldsymbol{x}^{(0)}) \tag{1-1}$$

其中，$\lim\limits_{\boldsymbol{x}\to\boldsymbol{x}^{(0)}}\alpha(\boldsymbol{x}-\boldsymbol{x}^{(0)})=0$，则称 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^{(0)}$ 处可微，向量 $\boldsymbol{p}$ 为 f 在 $\boldsymbol{x}^{(0)}$ 处的梯度，记为：$\nabla f(\boldsymbol{x}^{(0)})=\boldsymbol{p}$。

如果进一步存在对称的 n 阶方阵 $\boldsymbol{H}$，使得 $\forall\boldsymbol{x}\in S$，有：

$$f(\boldsymbol{x})=f(\boldsymbol{x}^{(0)})+p^{\mathrm{T}}(\dot{\boldsymbol{x}}-\boldsymbol{x}^{(0)})+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^{(0)})^{\mathrm{T}}H(\boldsymbol{x}-\boldsymbol{x}^{(0)})+\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|^2\beta(\boldsymbol{x},\boldsymbol{x}^{(0)}) \tag{1-2}$$

其中，$\lim\limits_{\boldsymbol{x}\to\boldsymbol{x}^{(0)}}\beta(\boldsymbol{x}-\boldsymbol{x}^{(0)})=0$，则称 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^{(0)}$ 处二阶可微，方阵 $\boldsymbol{H}$ 成为 f 在 $\boldsymbol{x}^{(0)}$ 处的 Hesse 矩阵，记为：$\nabla^2 f(\boldsymbol{x}^{(0)})=\boldsymbol{H}$。

不难得到，当 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^{(0)}$ 处二阶可微时：

$$\nabla f(\boldsymbol{x}^{(0)})=\left(\frac{\partial f(\boldsymbol{x}^{(0)})}{\partial x_1},\frac{\partial f(\boldsymbol{x}^{(0)})}{\partial x_2},\cdots,\frac{\partial f(\boldsymbol{x}^{(0)})}{\partial x_n}\right)^{\mathrm{T}}$$

$$\nabla^2 f(\boldsymbol{x}^{(0)}) = \begin{pmatrix} \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_1^2} & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_n \partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_1 \partial x_n} & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(\boldsymbol{x}^{(0)})}{\partial x_n^2} \end{pmatrix} \tag{1-3}$$

$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$处的一阶泰勒展开式为：

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla f(\boldsymbol{x}^{(0)})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^{(0)}) + o(\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|) \tag{1-4}$$

$f(\boldsymbol{x})$在$\boldsymbol{x}^{(0)}$处的二阶泰勒展开式为：

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla f(\boldsymbol{x}^{(0)})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^{(0)}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^{(0)})^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x}-\boldsymbol{x}^{(0)}) + o(\|\boldsymbol{x}-\boldsymbol{x}^{(0)}\|^2) \tag{1-5}$$

1.3.3　点和方向

在解数学规划问题时，常常要涉及迭代点和搜索方向，点和方向在R^n中都是向量，我们涉及的点是$\mathbf{R}^n$中的一个固定的向量，一般用$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\cdots$来标记；本书中谈到的方向，是指非零自由向量，它有固定的长度和方向，但起点可以移动，一般用$\boldsymbol{d},\boldsymbol{s},\cdots$来标记。

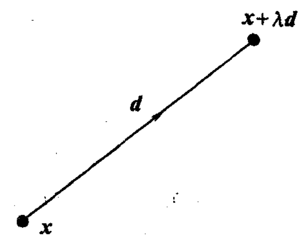


图　1-1

如$\boldsymbol{x}+\lambda\boldsymbol{d}$，表示从$\boldsymbol{x}$点出发沿方向$\boldsymbol{d}$，移动$\boldsymbol{d}$长度的$\lambda$倍所得到的向量，如图 1-1 所示。函数$f(\boldsymbol{x})$，表示沿这个方向的斜率为$f_\lambda'(\boldsymbol{x}+\lambda\boldsymbol{d})|_{\lambda=0} = \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}$，称关于$\lambda$的二阶导数为函数$f(\boldsymbol{x})$沿此方向的曲率。

$$f_\lambda''(\boldsymbol{x}+\lambda\boldsymbol{d})|_{\lambda=0} = \boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}$$

定义：$S\subset\mathbf{R}^n$，非空，$f:S\to\mathbf{R}$，$\boldsymbol{x}\in S$，$\boldsymbol{d}\in\mathbf{R}^n$，$\boldsymbol{d}\neq 0$，使当$\lambda>0$充分小时有$\boldsymbol{x}+\lambda\boldsymbol{d}\in S$，如果极限$\lim\limits_{\lambda\to 0^+}\frac{f(\boldsymbol{x}+\lambda\boldsymbol{d})-f(\boldsymbol{x})}{\lambda}$存在（包括取$\pm\infty$），则称$f$在$\boldsymbol{x}$点沿$\boldsymbol{d}$方向有导数，记为：

$$f'(\boldsymbol{x};\boldsymbol{d}) = \lim_{\lambda\to 0^+}\frac{f(\boldsymbol{x}+\lambda\boldsymbol{d})-f(\boldsymbol{x})}{\lambda}$$

显然，当f在$\boldsymbol{x}$点可微时，$f'(\boldsymbol{x};\boldsymbol{d}) = \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}$。

第 2 章 基本概念和基本理论

凸分析是运筹学的支柱性理论，优化理论大多建立在凸集合与凸规划的基础之上。进而，根据优化问题可行解集合、目标与约束函数的凸性，将优化问题的解分为局部最优解与全局最优解。目前，优化算法种类繁多，根据其优化理论，将现有优化算法分为经典优化算法与启发式优化算法。运筹学与最优化的发展与其他科学理论的发展密切相关，如数值计算、计算复杂性等。

本章将主要对凸集合与凸规划、局部最优解与全局最优解、经典优化算法与启发式优化算法、数值计算与计算复杂性等作综述性的简单介绍。

2.1 基本概念

在现实生活中，许多重要的问题都涉及选取一个最好的目标，或者为达到这个目标而选择某些参数、确定某些值，这些问题都可以归结为最优化问题。对于一个最小值问题，其数学规划模型的一般形式为：

$$(fS)\ \begin{cases}\min f(\boldsymbol{x})\\ \text{s.t.}\ \ \boldsymbol{x}\in S\end{cases}\tag{2-1}$$

其中，$S\subset\mathbf{R}^n$ 为约束集合或可行集；f: $S\to\mathbf{R}$ 为目标函数；若 $\boldsymbol{x}\in S$，则称 $\boldsymbol{x}$ 为问题（fS）的可行解。显然，只要改变目标函数的符号，最大值问题就可以转变成最小值问题。因此，本书一般以最小值问题为标准。解决最优化问题的算法称为最优化算法，可以分为经典优化算法和启发式优化算法。

若 $\boldsymbol{x}^*\in S$，且满足 $\forall\boldsymbol{x}\in S$，$f(\boldsymbol{x}^*)\leqslant f(\boldsymbol{x})$，则称 $\boldsymbol{x}^*$ 是问题（fS）的最优解，记为 opt.。对于 $f^*=f(\boldsymbol{x}^*)$，称之为问题（fs）的最优值。

定义 2-1 考虑问题（fs），设 $\boldsymbol{x}^*\in S$，有：

(1) 若 $\forall\boldsymbol{x}\in S$，恒有 $f(\boldsymbol{x}^*)\leqslant f(\boldsymbol{x})$，则称 $\boldsymbol{x}^*$ 是问题(fs)的全局最优解，记为 g. opt. (global optimum)或 opt.。若当 $\boldsymbol{x}\neq\boldsymbol{x}^*$ 时，严格不等式 $f(\boldsymbol{x}^*)<f(\boldsymbol{x})$ 成立，则称 $\boldsymbol{x}^*$ 是问题(fS)的严格全局最优解。

(2) 若存在 $\boldsymbol{x}^*$ 的邻域 $N(\boldsymbol{x}^*)$，使得 $\forall\boldsymbol{x}\in S\cap N(\boldsymbol{x}^*)$ 恒有 $f(\boldsymbol{x}^*)\leqslant f(\boldsymbol{x})$，则称 $\boldsymbol{x}^*$ 是问题(fS)的局部最优解，记为 l. opt. (local optimum)；若当 $\boldsymbol{x}\neq\boldsymbol{x}^*$ 时，严格不等式 $f(\boldsymbol{x}^*)<f(\boldsymbol{x})$ 成立，则称 $\boldsymbol{x}^*$ 是问题(fs)的严格局部最优解。

定义 2-2 设 $S\subset\mathbf{R}^n$，若对于任意 $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)}\in S,\lambda\in[0,1]$，均有 $\lambda\boldsymbol{x}^{(1)}+(1-$