普通高等学校信息与计算科学专业系列丛书

教育科学"十五"国家规划课题研究成果

最优化方法

孙文瑜 徐成贤 朱德通

高等教育出版社
HIGHER EDUCATION PRESS

普通高等学校信息与计算科学专业系列丛书

教育科学“十五”国家规划课题研究成果

最优化方法

孙文瑜　徐成贤　朱德通

高等教育出版社

内容提要

本书既系统地介绍最优化方法的基本理论和有效算法，又反映了目前该学科的发展动态。主要内容包括：基本概念，线性规划，线性搜索和信赖域方法，无约束最优化，线性与非线性最小二乘问题，二次规划，约束最优化等。全书深入浅出，理论、计算与实际应用相结合，尽可能避免较深的数学推导。每章后都有一个小结，并附有习题，以利于教学。

本书可作为信息与计算科学、数学与应用数学、统计、运筹学、管理科学与工程、计算机、经济与金融，以及有关理工科专业的本科生作为教材或教学参考书。具有微积分和高等代数基础的科技人员可自学本书。

图书在版编目(CIP)数据

最优化方法 / 孙文瑜，徐成贤，朱德通. —北京：高等教育出版社，2004. 8

ISBN 7-04-014375-5

Ⅰ. 最… Ⅱ. ①孙… ②徐… ③朱… Ⅲ. 最优化-高等学校-教材 Ⅳ. O224

中国版本图书馆 CIP 数据核字(2004)第 054929 号

出版发行	高等教育出版社	购书热线	010-64054588
社　　址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010-82028899		http://www.hep.com.cn
经　　销	新华书店北京发行所		
印　　刷	涿州市星河印刷有限公司		
开　　本	787×960　1/16	版　　次	2004 年 8 月第 1 版
印　　张	14	印　　次	2004 年 8 月第 1 次印刷
字　　数	260 000	定　　价	17.90 元

信息与计算科学专业系列教材编委会

总　序

根据教育部1998年颁布的普通高等院校专业目录，“信息与计算科学”专业被列为数学类下的一个新专业(它覆盖原有的计算数学及其应用软件、信息科学与运筹控制等专业)。这一新专业的设置很好地适应了新世纪以信息技术为核心的全球经济发展格局下的数学人才培养与专业发展的需要。然而，作为一个新专业，对其专业内涵、专业规范、教学内容与课程体系等有一个自然的认识与探索过程。教育部数学与统计学教学指导委员会数学类专业教学指导分委员会(下称教指委)经过过去两年艰苦细致的工作，对这些问题现在已有了比较明确的指导意见，发表了《关于信息与计算科学专业办学现状与专业建设相关问题的调查报告》及《信息与计算科学专业教学规范》(讨论稿)(见《大学数学》第19卷1期(2003))。为此，全国高等学校教学研究中心在承担全国教育科学“十五”国家级规划课题——“21世纪中国高等教育人才培养体系的创新与实践”研究工作的基础上，根据教指委所颁布的新的教学规范，组织国内各高校的专家教授，进行其子项目课题“21世纪中国高等学校信息与计算科学专业教学内容与课程体系的创新与实践”的研究与探索。为推动本专业的教材建设，该项目课题小组与高等教育出版社联合成立了“信息与计算科学专业系列教材编委会”，邀请有多年教学和科研经验的教师编写系列教材，由高等教育出版社独家出版，并冠以教育科学“十五”国家规划课题研究成果。

按照新的《信息与计算科学专业教学规范》(讨论稿)，信息与计算科学专业是以信息技术和计算技术的数学基础为研究对象的理科类专业。其目标是培养学生具有良好的数学基础和数学思维能力，掌握信息与计算科学基础理论、方法与技能，受到科学研究的训练，能解决信息技术和科学与工程计算中的实际问题的高级专门人才。毕业生能在科技、教育、信息产业、经济与金融等部门从事研究、教学、应用开发和管理工作，能继续攻读研究生学位。根据这一专业目标定位和落实“强基础、宽口径、重实际、有侧重、创特色”的办学指导思想，我们认为，本专业在数学基础、计算机基础、专业基础方面应该得到加强，而各学校在这三个基础方面可大体一致，但专业课(含选修课)允许各校自主选择、体现各自特点。考虑到已有大量比较成熟的数学基础与计算机基础课程教材，本次教材编写主要侧重于专业基础课与专业课(含选修课)方面。

信息与计算科学，就其范畴与研究内容而言，是数学、计算机科学和信息工

程等学科的交叉,已远远超出数学学科的范畴。但作为数学学科下的一个理科专业,信息与计算科学专业则主要研究信息技术的核心基础与运用现代计算工具高效求解科学与工程问题的数学理论与方法(或更简明地说,研究定向于信息技术与计算技术的数学基础),这一专业定位明显地与计算机科学与信息工程专业构成区别。基于这一定位,信息与计算科学专业可包括信息科学与科学计算(计算数学)两个大的方向。科学计算方向在我国已有长期的办学经验,通常被划分为偏微分方程数值解、最优化理论与方法、数值逼近与数值代数、计算基础等学科子方向。然而,对于信息科学,它到底应该怎样划分学科子方向?应该怎样设置专业与专业基础课?所有这些都仍是正在探索的问题。

任何技术都可以认为是延伸与扩展人的某种功能的方式与方法,所以信息技术可以认为是扩展人的信息器官功能的技术。人的信息器官主要包括感觉器官、传导器官(传导神经网络)、思维器官和效应器官四大类型,其功能则主要是信息获取、信息传输、信息处理和信息应用(控制),因而感测技术、通信技术、智能技术与控制技术通常被认为是最基本的信息技术(常称之为信息技术的四基元),其他信息技术可认为是这四种基本技术的高阶逻辑综合或分解衍生。所以可以把信息科学理解为是“有关信息获取、信息传输、信息处理与信息控制基础的科学”。从这个意义上,我们认为:信息处理(包括图像处理、信号分析等)、信息编码与信息安全、计算智能(人工智能、模式识别等)、自动控制等可构成信息科学的主要学科子方向。这一认识也是教指委设置信息与计算科学专业信息科学方向课程的基本依据。

本系列教材正是基于以上认识,为落实新的《信息与计算科学专业教学规范》(讨论稿)而组织编写的。我们相信,该系列教材的出版对缓解本专业教材的紧缺局面,对推动信息与计算科学专业的快速与健康发展会大有裨益。

从长远的角度看,为适应不同类型院校和不同层次要求的课程需求,本系列教材编委会还将不断组织教材的修订和编写新的教材,从而使本专业的教学用书做到逐步充实、完善和多样化。我们诚恳希望采用本系列教材的教师、同学们及广大读者对书中存在的问题及时指正并提出修改意见和建议。

信息与计算科学专业系列教材编委会

2003 年 8 月 31 日

前　言

最优化是一门应用十分广泛的学科，它研究在有限种或无限种可行方案中挑选最优方案，构造寻求最优解的计算方法。由于最优化在科学、工程、国防、交通、管理、经济、金融、计算机等领域的广泛应用，现在许多高校理科、工科、管理科学、经济与金融等学科都把最优化开设为一门必修或选修课程。

为了给我国高等院校理工科和管理类本科生提供一本现代、实用、合适的"最优化方法"教材，我们根据教育部信息与计算科学教材编写组的要求，结合我们自己在最优化领域的科研和教学体会，把这个领域最基本、最重要、最实用、最有效的现代最优化方法的内容介绍给学生，使本科生对最优化方法有一个基本的和较全面的了解，为进一步从事最优化方法的理论、算法、软件与应用打下一个较好的基础。

本书深入浅出，通俗易懂。我们尝试努力讲清每种方法的动机、算法、性质、收敛性及数值例子，尽可能避免较深较难的数学推导，力争把所需要的数学知识降到最低。对于个别难度较大的章节，为了全书的连贯性，我们仍然给出，但用 * 号表示，仅供教师和感兴趣的学生参考。例如§4.4.3 拟牛顿法的收敛性，§5.1.3 解线性不等式约束的线性最小二乘问题等。对于这些章节，教师可根据具体情况予以跳过。对于§1.2 凸集与凸函数中后面部分的理论内容教师可以选讲。

本书为本科生一学期用教材，材料覆盖最优化领域的核心内容。第一章介绍最优化问题的基本概念；第二章介绍在实际上有最广泛应用的线性规划；第三章讨论线性搜索与信赖域方法，它们提供了最优化方法的总体收敛策略；第四章研究无约束最优化方法，这一章是全书的一个中心，这不仅因为无约束最优化本身的重要性，而且众多约束最优化问题也都是转化为无约束问题来处理；第五章讨论一类特殊的最优化问题——线性与非线性最小二乘问题。这些问题在实际中广泛存在。我们向大家提供解这类问题的基本的现代方法，相信这会对大家今后的工作是相当有益的。第六和第七两章讨论约束最优化，其中我们将一类特殊的和重要的约束优化类型——二次规划单独构成一章，对约束最优化的理论与算法我们进行了简单的讨论。

由于这本书是作为本科生教材编写的，我们不希望本书内容太多、太全、太深，但也必须使学生了解这门学科的全貌，了解与掌握必要的方法、理论与算法

软件。应该说，本书包含的内容是最优化的核心部分，特别是实际中用得较多的内容，我们力求多讲一些。这本书可供一学期48课时（每周3课时×16周）～72课时（每周4课时×18周）的各类学校采用。教师可根据本校具体情况适当增删内容。本书适用面较广，既适用于重点院校，也适用于一般院校；既适用于信息与计算科学、数学与应用数学、统计与运筹等数学类专业，也适用于计算机、理学类、工程类、管理类、金融经济等学科。我们建议这门课在三年级讲授，使学生预先具有高等数学和线性代数的基本知识。同时，在讲授这门课时，最好要求学生上机编程序做练习，因为用最优化方法解题都需要在计算机上完成。

本书每章后面都有一个小结，概述了这一章的主要内容、背景，及一些进展。同时每章给出了一些习题供教师选择使用。全书后面附有一个附录和参考文献。附录中给出了求解无约束最优化和约束最优化的一些标准试验函数，这些试验函数供数值试验用。书末列出的参考文献供教师和感兴趣的学生作进一步深入研究探讨时参考。

在本书写作和录入的过程中，作者的研究生们提供了不少建议，并协助做了许多工作，谨致谢意。

南京航空航天大学理学院院长倪勤教授认真审阅了全书手稿，并提出了宝贵的修改意见。作者谨向他致以衷心的感谢。

教育部信息与计算科学教材编委会、高等教育出版社的领导及王瑜编辑对本书的写作给予了热情指导和关心，谨向他们致谢。

作者还要感谢国家自然科学基金委员会多年来的资助，感谢南京师范大学、西安交通大学和上海师范大学对作者工作的一贯支持。

尽管本书作者多年来一直从事最优化方法的研究和教学，但限于水平和时间，本书难免有不妥和错误之处，欢迎读者批评指正，以便再版时修改更正。

孙文瑜　徐成贤　朱德通

2004年2月15日

目 录

第1章 基本概念 …… 1
§1.1 最优化问题简介 …… 1
§1.2 凸集和凸函数 …… 5
§1.3 最优性条件 …… 16
§1.4 最优化方法概述 …… 27
小结 …… 31
习题 …… 32

第2章 线性规划 …… 34
§2.1 基本性质 …… 34
§2.2 单纯形方法 …… 45
§2.3 线性规划的对偶与对偶单纯形法 …… 59
§2.4 线性规划的内点算法 …… 71
小结 …… 79
习题 …… 81

第3章 线性搜索与信赖域方法 …… 85
§3.1 线性搜索 …… 85
§3.2 0.618法和Fibonacci法 …… 86
3.2.1 0.618法 …… 87
3.2.2 Fibonacci法 …… 88
3.2.3 二分法 …… 90
§3.3 逐次插值逼近法 …… 90
§3.4 精确线性搜索方法的收敛性 …… 94
§3.5 不精确线性搜索方法 …… 96
3.5.1 Goldstein准则 …… 97
3.5.2 Wolfe准则 …… 98
3.5.3 Armijo准则 …… 99
3.5.4 不精确线性搜索的收敛性 …… 99

§3.6 信赖域方法的思想和算法框架 …… 102
§3.7 信赖域方法的收敛性 …… 103
§3.8 解信赖域子问题 …… 107
小结 …… 110
习题 …… 111

第4章 无约束最优化方法 …… 112
§4.1 最速下降法 …… 112
§4.2 牛顿法 …… 115
§4.3 共轭梯度法 …… 119
4.3.1 共轭方向法 …… 119
4.3.2 共轭梯度法 …… 121
4.3.3 对于非二次函数的共轭梯度法 …… 125
§4.4 拟牛顿法 …… 128
4.4.1 拟牛顿条件 …… 128
4.4.2 DFP 校正和 BFGS 校正 …… 130
4.4.3 拟牛顿法的收敛性* …… 135
小结 …… 139
习题 …… 139

第5章 线性与非线性最小二乘问题 …… 141
§5.1 线性最小二乘问题的解法 …… 142
5.1.1 解线性最小二乘问题 …… 142
5.1.2 解线性等式约束的线性最小二乘问题 …… 145
5.1.3 解线性不等式约束的线性最小二乘问题* …… 149
§5.2 非线性最小二乘的 Gauss-Newton 法 …… 151
§5.3 信赖域方法 …… 155
§5.4 对 Gauss-Newton 矩阵的拟牛顿修正 …… 162
小结 …… 168
习题 …… 169

第6章 二次规划 …… 172
§6.1 二次规划 …… 172
§6.2 等式约束二次规划问题 …… 174
§6.3 凸二次规划的有效集方法 …… 180

小结 ······ 184
习题 ······ 184

第7章　约束最优化的理论与方法 ······ 186
§7.1　约束最优化问题与最优性条件 ······ 186
§7.2　二次罚函数方法 ······ 191
§7.3　内点障碍罚函数法 ······ 193
§7.4　序列二次规划方法 ······ 196
小结 ······ 202
习题 ······ 202

附录　试验函数 ······ 203
§1　无约束最优化问题的试验函数 ······ 203
§2　约束最优化问题的试验函数 ······ 204

参考文献 ······ 207

第 I 章
基本概念

§1.1 最优化问题简介

最优化技术在国民经济的许多领域如国防、工农业生产、交通运输、金融、贸易、管理、科学研究中有广泛的应用. 所谓最优化就是在众多可行的方案或方法中找到最好的方案或方法. 例如在确定投资项目时希望选择期望收益最大或风险最小的项目；空间飞船的设计要求在有限的空间放置尽可能多的设备；一个建筑在满足设计要求的条件下建筑费用应尽可能少；两地之间的输送管道或运输路线在满足要求的条件下应尽可能短. 为应用最优化技术确定最优的方案，需要针对具体的实际问题建立相应的最优化模型，再根据模型的具体形式和特性选择适当的最优化方法求解. 本书旨在介绍最优化的基本理论和常用的最优化方法.

最优化问题的数学模型一般形式为

$$\begin{cases} \min & f(x), \\ \text{s.t.} & c_i(x)=0, i=1,2,\cdots,m, \\ & c_i(x)\geqslant 0, i=m+1,\cdots,p, \end{cases} \tag{1.1.1}$$

其中 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$，$f:\mathbf{R}^n\to\mathbf{R}^1$，$c_i:\mathbf{R}^n\to\mathbf{R}^1$ $(i=1,2,\cdots,p)$ 为连续函数，通常还要求连续可微. x 称为决策变量，$f(x)$ 为目标函数，$c_i(x)$，$i=1,2,\cdots,p$ 为约束函数，$c_i(x)=0$，$i=1,2,\cdots,m$ 为等式约束，$c_i(x)\geqslant 0$，$i=m+1,\cdots,p$ 为不等式约束. min 和 s. t. 分别是英语单词 minimize(极小化)和 subject to(受约束)的缩写.

根据实际问题的不同要求，最优化模型有不同的形式，但经过适当的变换都可以转换成上述一般的形式. 例如，对于求目标函数 $f(x)$ 极大的问题

$$\max \quad f(x)$$

可转换成求 $-f(x)$ 极小的问题

$$\min \quad \phi(x)$$

其中 $\phi(x)=-f(x)$. 又如对于形如

$$c_i(x)\leqslant 0$$

的不等式约束，可同样转换成上述形式的不等式约束

$$h_i(x)\geqslant 0,$$

其中 $h_i(x)=-c_i(x)$. 还有像

$$a(x)\leqslant b(x)+c$$

的不等式约束，可通过令 $h(x)=b(x)-a(x)+c$ 转换成

$$h(x)\geqslant 0$$

的不等式约束形式. 通过后面进一步深入的学习，我们会看到转换在最优化理论和方法的研究和应用中有重要作用.

问题(1.1.1)是最优化问题的一般数学表现形式. 只要在问题中存在任何约束条件，就称为约束最优化问题. 只有等式约束时，

$$\begin{cases}\min & f(x),\\ \text{s. t.} & c_i(x)=0,\ i=1,2,\cdots,m\end{cases}$$

称为等式约束最优化问题. 只有不等式约束时，

$$\begin{cases}\min & f(x),\\ \text{s. t.} & c_i(x)\geqslant 0,\ i=1,2,\cdots,m\end{cases}$$

称为不等式约束最优化问题. 如果既有等式约束，又有不等式约束，则称为混合约束问题. 如果问题中无任何约束条件，则称为无约束最优化问题. 无约束最优化问题的数学模型为

$$\min \quad f(x),\quad x\in \mathbf{R}^n, \tag{1.1.2}$$

一般简记为

$$\min \quad f(x).$$

无约束最优化问题是最优化的基础，一则很多实际的最优化问题本身就是无约束最优化问题，二则许多约束最优化方法都是通过变换把约束最优化问题转换成无约束最优化问题后，用适当的无约束优化方法求解.

根据模型(1.1.1)中函数的具体性质和复杂程度，最优化问题又有许多不同的类型. 根据决策变量的取值是离散的还是连续的分为离散最优化和连续最优化，离散最优化通常又称组合最优化，如整数规划、资源配置、邮路问题、生产安排等问题都是离散最优化问题的典型例子. 离散最优化问题的求解较之连续最优化问题的求解难度更大，本书只介绍连续最优化的理论与方法. 根据连续最优化模型中函数的光滑与否又分为光滑最优化与非光滑最优化. 如果模型(1.1.1)中的所有函数都连续可微，则称为光滑最优化；只要有一个函数非光滑，则相应的优化问题就是非光滑优化问题. 本书只研究光滑最优化问题的求解方法，即在本书的大多数章节我们都假定模型中的函数连续可微，有时二阶或更高阶连续可微. 对于连续光滑的最优化问题，如果所有函数都是变量 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的线性函数，则称(1.1.1)为线性规划问题. 线性规划问题的一般形式为

$$
\begin{aligned}
\min\quad & c_1x_1+c_2x_2+\cdots+c_nx_n,\\
\text{s.t.}\quad & a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\
& \cdots\cdots\cdots\cdots\\
& a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m,\\
& a_{m+1,1}x_1+\cdots+a_{m+1,n}x_n\geqslant b_{m+1},\\
& \cdots\cdots\cdots\cdots\\
& a_{p1}x_1+a_{p2}x_2+\cdots+a_{pn}x_n\geqslant b_p.
\end{aligned}
$$

有时对变量 x 还有非负的约束

$$x_i\geqslant 0,\ i=1,2,\cdots,n,$$

或有界的约束

$$l_i\leqslant x_i\leqslant u_i,\ i=1,2,\cdots,n,$$

但这些约束都可以包括在上述模型的不等式约束中. 上述线性规划模型的矩阵向量表示为

$$
\begin{cases}
\min & c^{\mathrm{T}}x,\\
\text{s.t.} & A_1x=b^1,\\
 & A_2x\geqslant b^2,
\end{cases}
$$

其中 $c=(c_1,c_2,\cdots,c_n)^{\mathrm{T}},b^1=(b_1,b_2,\cdots,b_m)^{\mathrm{T}},b^2=(b_{m+1},\cdots,b_p)^{\mathrm{T}}$,

$$
A_1=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix},\quad
A_2=\begin{bmatrix} a_{m+1,1} & a_{m+1,2} & \cdots & a_{m+1,n}\\ \vdots & \vdots & & \vdots\\ a_{p1} & a_{p2} & \cdots & a_{pn}\end{bmatrix}.
$$

当模型(1.1.1)中的目标函数 $f(x)$是变量 x 的二次函数,而所有约束都是 x 的线性函数时称为二次规划问题. 二次规划问题的一般形式为

$$
\begin{cases}
\min & q(x)=\dfrac{1}{2}x^{\mathrm{T}}Gx+c^{\mathrm{T}}x+d,\\
\text{s.t.} & A_1x=b^1,\\
 & A_2x\geqslant b^2,
\end{cases}
$$

其中 A_1,A_2 的表示同线性规划模型类似,$c=(c_1,c_2,\cdots,c_n)^{\mathrm{T}}$,$d$ 为纯量,G 为 $n\times n$阶对称矩阵

$$
G=\begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1n}\\ G_{21} & G_{22} & \cdots & G_{2n}\\ \vdots & \vdots & & \vdots\\ G_{n1} & G_{n2} & \cdots & G_{nn}\end{bmatrix},
$$

满足 $G_{ij}=G_{ji},\forall\, i\neq j$. 只要模型(1.1.1)的函数中有一个关于 x 是非线性的,就称为非线性最优化问题. 非线性最优化问题是最一般的最优化问题,而线性规划和二次规划问题却是相当重要的特殊的最优化问题,这同样因为在实际中形成

的许多最优化问题是线性规划问题或二次规划问题，而且在用迭代法求非线性最优化问题的最优解时我们常常用线性规划或二次规划来局部近似原非线性最优化问题，并通过求所得近似问题的最优解来对已有最优解的估计进行改进.

如果点 $x\in\mathbf{R}^n$ 满足最优化模型(1.1.1)中的所有约束条件，就称其为可行点(feasible point)，所有可行点的全体称为可行域(feasible region)，在本书中我们用 $\mathcal{F}$ 表示可行域，即

$$\mathcal{F}=\{x\,|\,c_i(x)=0, i=1,2,\cdots,m,\quad c_i(x)\geqslant 0, i=m+1,\cdots,p\}.$$

对于一个可行点 $\bar{x}$，考虑不等式约束 $c_i(x)\geqslant 0$，如果有 $c_i(\bar{x})=0$，就称约束 $c_i(x)\geqslant 0$ 在点 $\bar{x}$ 是有效约束或起作用约束(active constraint)，并称可行点 $\bar{x}$ 位于约束 $c_i(x)\geqslant 0$ 的边界. 如果有 $c_i(\bar{x})>0$，就称不等式约束 $c_i(x)\geqslant 0$ 在点 $\bar{x}$ 是无效约束或不起作用约束(inactive constraint)，并称 $\bar{x}$ 是约束 $c_i(x)\geqslant 0$ 的内点. 在任意可行点，所有的等式约束都被认为是有效约束. 在一个可行点，所有有效约束的全体被称为该可行点的有效集，并记为

$$\mathcal{A}_{\bar{x}}=\{i\,|\,i=1,2,\cdots,m,m+1,\cdots,p,c_i(\bar{x})=0\}.$$

对于一可行点 $\bar{x}$，如果没有一个不等式约束是有效的，就称 $\bar{x}$ 是可行域的内点. 不是内点的可行点就是可行域的边界点. 显然，在边界点至少有一个不等式约束是有效约束. 当存在等式约束时，任何可行点都要满足等式约束，因此不可能是等式约束的内点.

图 1.1 给出了由下述约束条件给出的可行域 $\mathcal{F}$：

$$c_1(x)=2x_1+3x_2+x_3-6=0,$$
$$x_1\geqslant 0,$$
$$x_2\geqslant 0,$$
$$x_3\geqslant 0.$$

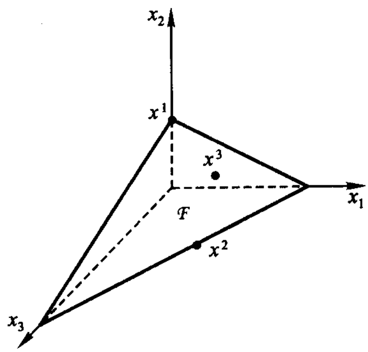

图 1.1　可行域与可行域的边界

对于可行点 x^1，约束 $x_1\geqslant 0$ 和 $x_3\geqslant 0$ 是有效约束，而 $x_2\geqslant 0$ 是无效约束. 对于可行点 x^2，则刚好相反，约束 $x_2\geqslant 0$ 是有效约束，而 $x_1\geqslant 0$ 和 $x_3\geqslant 0$ 是无效约束. 而对于可行点 x^3，三个不等式约束都是无效约束. 图中可行域的边界由粗线表示.

一个可行点 $x^*\in\mathcal{F}$ 称为问题(1.1.1)的(全局或总体)最优解，如果有

$$f(x^*)\leqslant f(x),\quad \forall x\in\mathcal{F};$$

如果上述不等式对所有不同于 x^* 的可行点 x 严格成立，即

$$f(x^*)<f(x),\quad \forall x\in\mathcal{F}, x\neq x^*,$$

则称 x^* 为严格(全局或总体)最优解. 对于可行点 x^*，如果存在一个邻域

$$\mathcal{N}(x^*)=\{x \mid \|x-x^*\| \leqslant \delta\}$$

使得成立

$$f(x^*)\leqslant f(x),\ \forall x\in \mathcal{N}(x^*)\cap \mathcal{F},$$

则称 x^* 为优化问题(1.1.1)的局部最优解，其中 $\delta>0$ 是一个小的正整，范数 $\|\cdot\|$ 可以是任意向量范数，但一般常用 l_2 范数

$$\|x\| = \left[\sum_{i=1}^{n} x_i^2\right]^{\frac{1}{2}}.$$

如果上述不等式对所有 $x\in \mathcal{N}(x^*)\cap \mathcal{F}$，$x\neq x^*$ 严格成立，则称 x^* 为严格局部极小点.

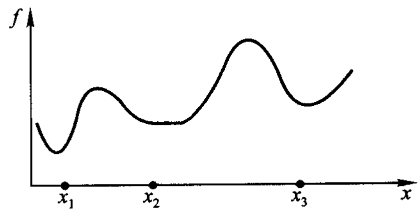

图 1.2　全局极小点与局部极小点

图 1.2 就单变量函数的情形给出了全局最优解和局部最优解的一个例子. 在这个图中，点 x_1 是所示函数的严格全局极小解，x_2 和 x_3 则是局部极小解，其中 x_3 是严格局部极小解，而 x_2 是非严格局部极小解. 需要注意的是并非所有的连续可微函数都有极小解，单变量函数 $f(x)=x^3$ 就是一个最简单而又容易理解的例子. 即使问题有最优解，最优解未必唯一，也未必是全局最优解. 本书介绍的方法一般只能用于确定最优化问题的局部最优解，有关确定全局最优解的最优化方法属于最优化研究的另一个领域——全局最优化. 然而，如果最优化问题的目标函数是凸的，而可行域是凸集，则问题的任何最优解(不一定唯一)必是全局最优解，这样的最优化问题又称凸规划.

§ 1.2　凸集和凸函数

正如我们在上一节所指出的目标函数为凸函数，可行集为凸集的凸规划问题可确保任意最优解就是问题的全局最优解. 因此，凸集与凸函数在最优化理论与算法的研究与学习中起着重要作用. 本节我们介绍凸集与凸函数的某些基本概念和基本性质.

定义 1.2.1　集合 $D\subset \mathbf{R}^n$ 称为凸的，如果对于任意 $x,y\in D$ 有

$$\lambda x+(1-\lambda)y\in D,\quad \forall\, 0\leqslant\lambda\leqslant 1.$$

换句话说，如果 $x,y\in D$，则连接 x 与 y 的直线段上的所有点都在 D 内.

图 1.3 给出了平面上凸集与非凸集的例子.

有关凸集有下面这样一些有用的性质：两个凸集的交，和以及差仍然是凸集. 即如果 D_1，$D_2 \subset \mathbf{R}^n$ 是凸集，则 $D_1 \cap D_2 = \{x \mid x \in D_1, x \in D_2\}$，$D_1 + D_2 = \{x + y \mid x \in D_1, y \in D_2\}$ 和 $D_1 - D_2 = \{x - y \mid x \in D_1, y \in D_2\}$ 都是凸集；对于任意非零实数 α，集合 $\alpha D_1 = \{\alpha x \mid x \in D_1\}$ 也是凸集. 根据凸集的定义，我们还可以得到凸集的任意有限个点的凸组合仍属于凸集的结论，即我们有下述定理.

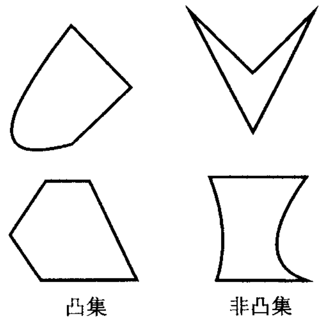

图 1.3 凸集与非凸集

定理 1.2.2 设 $D \subset \mathbf{R}^n$ 是凸集，则任意 m 个点 $x^{(i)} \in D (i=1,2,\cdots,m)$ 的凸组合仍属于 D，即有

$$\sum_{i=1}^{m} \alpha_i x^{(i)} \in D, \tag{1.2.1}$$

其中 $\alpha_i \geqslant 0, i=1,2,\cdots,m,\ \sum_{i=1}^{m} \alpha_i = 1$.

证明 定理的结论可以用归纳法证明. 根据凸集的定义，定理的结论在 $m=2$ 时显然成立. 设结论在 $m=k$ 时成立，我们证明结论在 $m=k+1$ 时也成立. 设 $x^{(i)} \in D, i=1,2,\cdots,k+1$，考虑任意一组实数 $\alpha_i \geqslant 0, i=1,2,\cdots,k+1,\ \sum_{i=1}^{k+1} \alpha_i = 1$，则

$$\begin{aligned}\sum_{i=1}^{k+1} \alpha_i x^{(i)} &= \sum_{i=1}^{k} \alpha_i x^{(i)} + \alpha_{k+1} x^{(k+1)} \\ &= (1-\alpha_{k+1}) \sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} x^{(i)} + \alpha_{k+1} x^{(k+1)}.\end{aligned}$$

由于 $1-\alpha_{k+1} = \sum_{i=1}^{k} \alpha_i$，我们有

$$\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} = 1.$$

由归纳假设有

$$\bar{x} = \sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} x^{(i)} \in D.$$

再由凸集的定义得

$$\sum_{i=1}^{k+1} \alpha_i x^{(i)} = (1-\alpha_{k+1})\bar{x} + \alpha_{k+1} x^{(k+1)} \in D.$$

这就完成了定理的证明. □

凸集除了在刻画可行域的特性方面起重要作用外，下面的凸集分离定理在分析最优解的最优性条件时起重要作用.