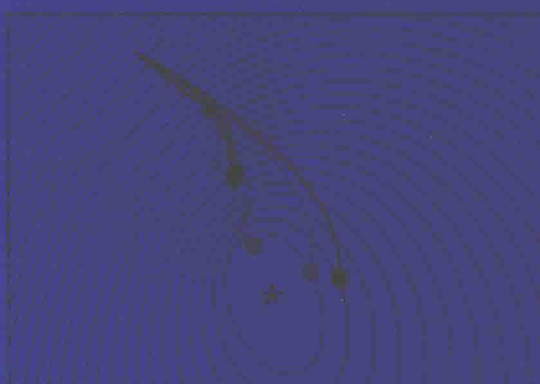
工科研究生数学类基础课程应用系列丛书

现代优化理论与方法

（上册）

黄庆道　吕显瑞　李晓峰　王彩玲　编



科学出版社

工科研究生数学类基础课程应用系列丛书

现代优化理论与方法

(上册)

黄庆道 吕显瑞 李晓峰 王彩玲 编

科学出版社

北京

内 容 简 介

本书分上、下两册, 共 11 章, 包括最优化问题、线性规划、无约束非线性规划问题、有约束非线性规划、多目标规划、全局最优化问题、二次规划、整数规划、动态规划及优化求解的软件实现等问题.

本书可作为最优化及相关专业的研究生教材和高年级本科生的选修教材, 也可供从事相关专业的科研人员和工程技术人员参考.

图书在版编目(CIP)数据

现代优化理论与方法. 上册/黄庆道等编. —北京：科学出版社, 2017.7
(工科研究生数学类基础课程应用系列丛书)

ISBN 978-7-03-053961-8

I. ①现… II. ①黄… III. ①最佳化-数学理论-研究生-教材 IV. ①O224

中国版本图书馆 CIP 数据核字(2017) 第 168338 号

责任编辑: 张中兴 / 责任校对: 邹慧卿
责任印制: 白 洋 / 封面设计: 迷底书装

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

保定市中画美凯印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*

2017 年 7 月第 一 版 开本: 720 × 1000 1/16
2017 年 7 月第一次印刷 印张: 12 1/2
字数: 252 000

定价: 36.00 元

(如有印装质量问题, 我社负责调换)

序　言

研究生公共数学系列教材是根据国家教育部关于研究生培养指导规划和目标、结合当前研究生教育改革的实际情况、借鉴国内外研究生教育的最新研究成果、旨在规范和加强研究生公共基础课教学的一套研究生公共数学系列教材. 本系列教材经过对研究生公共数学课程整合、优化, 共编写教材 13 册, 具体包括:《现代分析基础》(上、下)、《代数学基础》(上、下)、《现代统计学基础》(上、下)、《现代微分方程概论》(上、下)、《现代计算方法》(上、下)、《现代优化理论与方法》(上、下)、《应用泛函分析》. 其中上册为非数学类硕士研究生教材, 下册为非数学类博士研究生教材.

本系列教材的编写体现了时代的特征, 本着加强基础、淡化证明、强调应用的原则, 力争做到科学性、系统性和实用性的统一, 着眼于传授教学知识和培养学生数学素养的高度结合.

本系列教材吸取国内外同类教材的精华, 参考近年来出版的一些新教材, 结合当前研究生公共数学教学改革的实际, 特别是综合性大学非数学类研究生公共数学的实际需求.

本系列教材体例科学、结构合理、内容经典而追求创新, 既是作者多年教学经验的总结, 又是作者长期教学研究和科学研究成果的体现. 每章后面配备巩固基本概念、基本理论、基本运算的基本题目, 又有提高学生抽象思维、逻辑推理和综合运用基础知识解题的提高题目, 为学生掌握教材基本内容, 运用教材基本知识开发创新思维提供了可行条件.

本系列教材适用面广、涉及专业全、教学内容新, 可作为综合性大学非数学专业研究生公共数学教材和数学参考书, 在教材体系与内容的编排上认真考虑不同专业, 不同学时的授课对象的需求, 可选择不同的教学模块, 以满足广大读者的实际需要.

本系列教材的编写过程中, 得到了吉林大学研究生院、吉林大学数学学院和数学研究所的大力支持, 也得到了科学出版社的领导和编辑的鼎力帮助, 在此一并致谢.

由于编者水平有限, 书中的不妥之处在所难免, 恳请广大读者批评指正, 以期不断完善.

研究生公共数学课程系列教材编委会

2015 年 3 月于长春

前言

本课程属于非数学类研究生数学公共基础课程之一, 最优化方法是从所有可能方案中选择最合理的方案以达到最优目标的科学. 最优化方法随着电子计算机的普遍应用而迅猛发展, 已广泛应用于国民经济各部门和科学技术的各个领域中. 因此, 学习和掌握最优化的基本理论和方法, 对于将来从事工程技术工作的工科研究生来说是必不可少的. 本门课程旨在讲授最优化的基本理论和方法, 要求学生通过本课程的学习, 具有应用最优化方法解决一些实际问题的初步技能, 并为以后的学习和工作做必要的准备.

根据工科研究生课程指导委员会制定的"工科研究生最优化方法课程教学基本要求", 结合吉林大学为本校硕士研究生和本科生编写的最优化理论与方法教材及多年来教学实践的体会, 我们选取了线性规划、非线性规划、多目标规划、全局最优化和现代优化理论五部分, 即全书 1—7 章, 作为优化基本理论上册内容.

为了进一步优化理论的学习, 我们又选取了二次规划、整数规划、动态规划和优化求解的软件实现的内容, 即全书 8—11 章作为本书的下册内容. 每一部分内容着重阐明基本理论与基本方法, 以便给读者在该领域的深入学习和研究打下良好基础, 对于一些证明较冗长和复杂的定理, 我们只给出定理的内容, 证明从略.

本书力求深入浅出, 通俗易懂, 学过高等数学和线性代数的读者均能学习. 本书既可作为工科研究生和数学系高年级本科生学习最优化方法课程的教材, 也可以作为从事应用数学、管理学、系统工程及工程设计方面的广大科技工作者的参考书.

由于编者水平有限, 缺点和疏漏在所难免, 敬请读者予以批评指正.

黄庆道

2016 年 3 月

目　　录

第 1 章　最优化问题简介

最优化理论和方法是第二次世界大战后迅速发展起来的一门新学科, 随着现代科学技术的进步, 最优化理论和方法日益受到人们的重视, 它已经渗透到国防、工农业生产、交通运输、金融、贸易、管理、科学研究等各领域. 本章从几个实例出发, 说明最优化问题的数学模型及相关概念, 并介绍了凸集与凸函数的相关概念.

1.1　最优化问题的数学模型与基本概念

最优化问题的实例一般都比较复杂, 为了便于理解, 我们引入几个简单的例子.

例 1.1.1 (运输问题)　设有 m 个水泥厂 $A_1, A_2, \cdots, A_m$, 年产量分别为 $a_1, a_2, \cdots, a_m$ t. 有 k 个城市 $B_1, B_2, \cdots, B_k$ 用这些水泥厂生产的水泥, 年需求量各为 $b_1, b_2, \cdots, b_k$t. 再设由 A_i 到 B_j 每吨水泥的运价为 c_{ij} 元. 假设产销是平衡的, 即 $\sum\limits_{i=1}^{m} a_i = \sum\limits_{j=1}^{k} b_j$. 试设计一个调运方案, 在满足需要的同时使总运费最省.

设 A_i 调往 B_j 的水泥为 x_{ij} t, 则问题化为求总运费

$$s = \sum_{i=1}^{k} \sum_{j=1}^{k} c_{ij} x_{ij}$$

的极小值, 且满足下面的条件:

$$\begin{aligned}
&\sum_{j=1}^{k} x_{ij} = a_i, \quad i = 1, 2, \cdots, m; \\
&\sum_{i=1}^{m} x_{ij} = b_j, \quad j = 1, 2, \cdots, k; \\
&x_{ij} \geqslant 0, \quad i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, k.
\end{aligned}$$

例 1.1.2 (生产计划问题)　设某工厂有 m 种资源 $B_1, B_2, \cdots, B_m$, 数量各为 $b_1, b_2, \cdots, b_m$. 用这些资源生产 n 中产品 $A_1, A_2, \cdots, A_n$. 每生产一个单位的 A_j 产品需要消耗资源 B_i 的量为 a_{ij}. 根据合同规定, 产品 A_j 的量不少于 d_j. 再设 A_j 的单价为 c_j. 问如何安排生产计划, 才能既完成合同, 又使该厂总收入最多?

设产品 A_j 的计划产量为 x_j, 总产值 $y = \sum_{j=1}^{n} c_j x_j$, 则问题化为求总产值

$$y = \sum_{j=1}^{n} c_j x_j$$

的极大值, 且满足条件:

$$\begin{aligned} &\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \quad i = 1, 2, \cdots, m; \\ &x_j \geqslant d_j, \quad j = 1, 2, \cdots, n. \end{aligned}$$

例 1.1.3 (指派问题) 设有四项任务 B_1, B_2, B_3, B_4, 派四个人 A_1, A_2, A_3, A_4 去完成. 每个人都可以承担四项任务中的任何一项, 但所耗费的资金不同. 设 A_i 完成 B_j 所需资金为 c_{ij}. 如何分配任务, 使总支出最少?

设变量

$$x_{ij} = \begin{cases} 1, & \text{指派 } A_i \text{ 去完成 } B_j, \\ 0, & \text{不派 } A_i \text{ 去完成 } B_j. \end{cases}$$

总支出为 $s = \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij} x_{ij}$, 则问题化为使总支出 s 最小且满足条件

$$\begin{aligned} &\sum_{j=1}^{4} x_{ij} = 1, \quad i = 1, 2, 3, 4; \\ &\sum_{i=1}^{4} x_{ij} = 1, \quad j = 1, 2, 3, 4. \end{aligned}$$

其中

$$x_{ij} = 0 \text{ 或 } 1,$$

这里的变量 x_{ij} 叫 0-1 变量.

例 1.1.4 (数据拟合问题) 在实验数据处理或统计资料分析中常遇到如下问题. 设两个变量 x 和 y, 已知存在函数关系, 但其解析表达式或者是未知的或者虽然为已知的但过于复杂. 设已取得一组数据

$$(x_i, y_i), \quad i = 1, 2, \cdots, m.$$

根据这一组数据导出函数 $y = f(x)$ 的一个简单而近似的解析表达式.

取一个简单的函数序列 $\varphi_0(x),\varphi_1(x),\cdots,\varphi_n(x)$, 比如取幂函数列 $1,x,x^2,\cdots,x^n$ 作为基本函数系. 求 $\varphi_0,\varphi_1,\cdots,\varphi_n$ 的一个线性组合 $\sum\limits_{j=0}^{n}\alpha_j\varphi_j(x)$ 作为函数 $f(x)$ 的近似表达式, 而系数 $\alpha_0,\alpha_1,\cdots,\alpha_n$ 的选取要使得平方和

$$Q=\sum_{i=1}^{m}\left[y_i-\sum_{j=0}^{n}\alpha_j\varphi_j(x_i)\right]^2$$

最小. 此问题的变量为 $\alpha_0,\alpha_1,\cdots,\alpha_n$, 对这些变量没有限制. 这种问题又叫最小二乘问题.

1.2 最优化问题的分类

最优化问题根据变量取值可分为: 连续最优化问题 (continuous optimization problem) 和离散最优化问题 (discrete optimization problem), 而离散最优化问题也称为组合优化问题 (combinatorial optimization problem).

连续最优化问题又分为:

线性规划 (linear program) (要求目标函数与约束条件均为线性的);

非线性规划 (nonlinear program) (目标函数与约束条件未必是线性的).

非线性规划又可分为:

二次规划 (quadratic program) (目标函数是二次, 约束条件是线性);

凸规划 (convex program) (目标函数是凸函数, 约束条件是凸集).

离散最优化问题可分为:

整数规划 (integer program) ;

0-1 规划 (0-1 program) ;

网络流问题 (network flow problem) ;

旅行商问题 (traveling salesman problem) .

上述线性规划、非线性规划、整数规划统称为数学规划 (mathematical programming problem) .

1.3 凸集与凸函数

凸性在最优化理论和方法的研究中起着重要作用. 本节扼要地介绍凸集和凸函数的基本概念和基本结果. 对凸分析感兴趣的读者可参阅这方面的专著, 如 Rockafellar (1970) 的 *Convex Analysis*, Eggleston (1958) 的 *Convexity* 等.

1.3.1 凸集

定义 1.3.1 设集合 $S \subset \mathbf{R}^n$, 如果对任意 $x_1, x_2 \in S$, 有

$$\alpha x_1 + (1-\alpha)x_2 \in S, \quad \forall \alpha \in [0,1], \tag{1.1}$$

则称 S 是凸集.

这个定义表明, 如果 $x_1, x_2 \in S$, 则连接 x_1 和 x_2 的线段属于 S.

归纳地可以证明, $\mathbf{R}^n$ 的子集 S 为凸集当且仅当对任意 $x_1, x_2, \cdots, x_m \in S$, 有

$$\sum_{i=1}^{m} \alpha_i x_i \in S, \tag{1.2}$$

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

(1.1) 中的 $x = \alpha x_1 + (1-\alpha)x_2$ 称为 x_1 和 x_2 的凸组合, (1.2) 中的 $x = \sum\limits_{i=1}^{m} \alpha_i x_i$ 称为 $x_1, x_2, \cdots, x_m$ 的凸组合.

例 1.3.1 超平面 $H = \{x \mid p^{\mathrm{T}}x = \alpha\}$ 是凸集, 其中 $p \in \mathbf{R}^n$ 是非零向量, 称为超平面的法向量, α 为实数.

事实上, 对任意 $x_1, x_2 \in H$ 和每一个 $\theta \in [0,1]$, 有

$$p^{\mathrm{T}}[\theta x_1 + (1-\theta)x_2] = \alpha,$$

故 $\theta x_1 + (1-\theta)x_2 \in H$.

例 1.3.2 闭半空间 $H^- = \{x|p^{\mathrm{T}}x \leqslant \beta\}$ 和 $H^+ = \{x|p^{\mathrm{T}}x \geqslant \beta\}$ 为凸集. 开半空间 $\dot{H}^- = \{x|p^{\mathrm{T}}x < \beta\}$ 和 $\dot{H}^+ = \{x|p^{\mathrm{T}}x > \beta\}$ 为凸集.

事实上, 注意到对任意 $x_1, x_2 \in H^-$ 和 $\theta \in [0,1]$ 有

$$p^{\mathrm{T}}[\theta x_1 + (1-\theta)x_2] = \theta p^{\mathrm{T}}x_1 + (1-\theta)p^{\mathrm{T}}x_2 \leqslant \beta.$$

故 $\theta x_1 + (1-\theta)x_2 \in H^-$. 其余类似.

例 1.3.3 射线 $S = \{x|x_0 + \lambda d, \lambda \geqslant 0\}$ 为凸集, 其中, d 是给定的非零向量, x_0 为定点.

对任意 $x_1, x_2 \in S$ 和每一个数 $\lambda \in [0,1]$, 有

$$x_1 = x_0 + \lambda_1 d, \quad x_2 = x_0 + \lambda_2 d,$$

其中 $\lambda_1, \lambda_2 \in [0,1]$. 因而

$$\begin{aligned}
\lambda x_1 + (1-\lambda)x_2 &= \lambda(x_0 + \lambda_1 d) + (1-\lambda)(x_0 + \lambda_2 d) \\
&= x_0 + [\lambda\lambda_1 + (1-\lambda)\lambda_2]d,
\end{aligned}$$

注意到 $\lambda\lambda_1+(1-\lambda)\lambda_2\geqslant 0$, 故 $\lambda x_1+(1-\lambda)x_2\in S$.

此外, 若 A 是 $m\times n$ 矩阵, $b\in\mathbf{R}^n$, 则集合

$$S=\{x\in\mathbf{R}^n|Ax=b\}$$

是凸集.

由有限个闭半空间的交组成的集合 S 叫多面集,

$$S=\{x|p_i^{\mathrm{T}}x\leqslant\beta_i, i=1,\cdots,m\},$$

其中 p_i 是非零向量, β_i 是数. 多面集是闭凸集. 由于等式可以用两个不等式表示, 所以下面的集合都是多面集的例子.

$$S=\{x\mid Ax=b, x\geqslant 0\},$$

$$S=\{x\mid Ax\geqslant 0, x\geqslant 0\}.$$

下面的引理叙述了凸集的性质, 即两个凸集的交集是凸集, 两个凸集的代数和是凸集. 这个引理的证明留给读者练习.

引理 1.3.1 设 S_1 和 S_2 是 $\mathbf{R}^n$ 中的凸集, 则

(1) $S_1\cap S_2$ 为凸集;

(2) $S_1\pm S_2=\{x_1\pm x_2\mid x_1\in S_1, x_2\in S_2\}$ 是凸集.

从这个引理可知, 线性规划和二次规划中的可行域是凸集, 因为它是超平面和半空间的交集.

设 $S\subset\mathbf{R}^n$, 包含子集 S 的所有凸集的交叫 S 的凸包, 记作 $\mathrm{conv}(S)$, 它是包含 S 的唯一的最小的凸集. 凸包 $\mathrm{conv}(S)$ 由 S 中元素的所有凸组合组成,

$$\mathrm{conv}(S)=\left\{x\ \middle|\ x=\sum_{i=1}^m\alpha_i x_i, x_i\in S, \sum_{i=1}^m\alpha_i=1, \alpha_i\geqslant 0, i=1,\ \cdots,m\right\}. \tag{1.3}$$

$\mathbf{R}^n$ 的子集叫锥, 如果它关于正的数乘运算是封闭的, 即当 $x\in K$ 和 $\lambda>0$ 时, $\lambda x\in K$. 如果锥 K 也是凸集, 则称之为凸锥. 例如,

$$\{x=(\xi_1,\cdots,\xi_n)\mid \xi_1\geqslant 0,\cdots,\xi_n\geqslant 0\},$$

$$\{x=(\xi_1,\cdots,\xi_n)\mid \xi_1>0,\cdots,\xi_n>0\}$$

和

$$\{x\in\mathbf{R}^n\mid x^{\mathrm{T}}b_i\leqslant 0,\ i\in I\}$$

均是凸锥, 在上式中, $b_i\in\mathbf{R}^n$, I 是一个任意指标集.

$\mathbf{R}^n$ 的一个子集是凸锥当且仅当它关于加法和正的数乘运算是封闭的. 包含凸集 C 的最小凸锥是

$$K = \{\lambda x \mid \lambda > 0, x \in C\}.$$

下面叙述一下开集、闭集、开凸集和闭凸集.

设 $x \in \mathbf{R}^n$, 开球 $B(x,r)$ 定义为

$$B(x,r) = \{y \in \mathbf{R}^n \mid \|y - x\| < r\},$$

这个开球以 x 为中心, 以 r 为半径.

设 $S \subset \mathbf{R}^n$, 如果存在 $r > 0$, 使得 $B(x,r) \subset S$, 则称 $x \in \mathbf{R}^n$ 是 S 的内点. S 的所有内点的集合叫 S 的内部, 用 $\text{int}(S)$ 表示. 显然, $\text{int}(S) \subset S$.

如果子集 S 的每一点都是 S 的内点, 即 $\text{int(S)}= S$, 则称 S 为开子集. 特别地, 空集 $\varnothing$ 和 n 维空间 $\mathbf{R}^n$ 是 $\mathbf{R}^n$ 的开子集.

设 $S \subset \mathbf{R}^n$, 如果

$$S \cap B(x,r) \neq \varnothing, \quad \forall r > 0,$$

则 x 称为属于 S 的闭包, 即 $x \in \overline{S}$. 显然, $S \subset \overline{S}$.

如果 $S = \overline{S}$, 则 S 称为闭子集. 特别地, 空集 $\varnothing$ 和 n 维空间 $\mathbf{R}^n$ 是 $\mathbf{R}^n$ 的闭子集. 直观地说, 如果一个子集包含它所有的边界点, 则它是闭的. 例如, 闭球 $\overline{B}(x,r) = \{y \in \mathbf{R}^n | \ \|y - x\| \leqslant r\}$ 是闭集.

显然, 一个子集是闭的, 当且仅当它的补是开的.

根据上面的定义, 闭包 $\overline{S}$ 可以写为

$$\overline{S} = \left\{ x \in \mathbf{R}^n \;\middle|\; \lim_k \| x_k - x \| = 0, x_k \in S \right\}.$$

设 $S \subset \mathbf{R}^n$ 是凸集, 若它是开的, 则称为开凸集; 若它是闭的, 则称为闭凸集.

下面的定理表明凸集的闭包是凸集.

定理 1.3.1　如果 $C \subset \mathbf{R}^n$ 是凸集, 那么 C 的闭包 $\overline{C}$ 也是凸集.

证明　我们必须证明对任何 $x, y \in \overline{C}$, $0 \leqslant \lambda \leqslant 1$, 有

$$\lambda x + (1-\lambda) y \in \overline{C}.$$

选择 $\{x_k\}, \{y_k\} \subset C$, 使得

$$\lim_k \| x_k - x \| = 0, \quad \lim_k \| y_k - y \| = 0,$$

于是,

$$\begin{aligned}
&\|[\lambda x_k + (1-\lambda) y_k] - [\lambda x + (1-\lambda) y]\| \\
=&\|\lambda(x_k - x) + (1-\lambda)(y_k - y)\| \\
\leqslant&\lambda\|x_k - x\| + (1-\lambda)\|y_k - y\|.
\end{aligned}$$

因此, 取极限便有

$$\lim_{k} \|[\lambda x_k + (1-\lambda)y_k] - [\lambda x + (1-\lambda)y]\| = 0,$$

从而 $\lambda x + (1-\lambda)y \in \overline{C}$. □

在凸集的研究中另一个有用的概念为凸集的极值点和极值方向.

定义 1.3.2 设 $S \subset \mathbf{R}^n$ 是非空凸集, $x \in S$, 若 x 不在 S 中任何线段的内部, 即, 若假设 $x = \theta x_1 + (1-\theta)x_2$, $x_1, x_2 \in S, \theta \in (0,1)$ 必推出 $x = x_1 = x_2$, 则称 x 是凸集 S 的极值点.

显然, 多边形的顶点和圆周上的任意点都是极限点.

定义 1.3.3 设 $S \subset \mathbf{R}^n$ 为闭凸集, d 为非零向量, 如果对每一个 $x \in S$, $x + \lambda d \in S$, $\forall \lambda \geqslant 0$, 则称向量 d 为 S 的方向. 又设 d_1 和 d_2 为 S 的两个方向, 如果 $d_1 \neq \alpha d_2$, $\forall \alpha > 0$, 则称 d_1 和 d_2 是 S 的两个不同方向. 如果 S 的方向 d 不能表示成该集合的两个不同方向的正的线性组合, 即如果 $d = \lambda_1 d_1 + \lambda_2 d_2$, $\lambda_1, \lambda_2 > 0$, 必推得 $d_1 = \alpha d_2$, 其中 $\alpha > 0$, 则称 d 为 S 的极值方向.

今考虑多面集

$$S = \{x \mid Ax = b, x \geqslant 0\},$$

其中 A 是 $m \times n$ 矩阵, $\operatorname{rank}(A) = m$, $b \in \mathbf{R}^m$. 不失一般性, 设

$$A = [B, N],$$

其中 B 是 $m \times m$ 非奇异矩阵, N 是 $m \times (n-m)$ 矩阵. 设 x_B 和 x_N 分别是对应于 B 和 N 的向量,

$$Bx_B + Nx_N = b, \quad x_B \geqslant 0, \ x_N \geqslant 0.$$

于是, x 是多面集 S 的极值点的充分必要条件是

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix},$$

其中, $B^{-1}b \geqslant 0$.

$d \neq 0$ 是 S 的一个方向, 当且仅当 $Ad = 0$, $d \geqslant 0$. $\overline{d}$ 是 S 的一个极值方向, 当且仅当

$$B^{-1}a_j \leqslant 0, \quad \text{对某个 } a_j \text{ 是 } N \text{ 的列},$$

$$\overline{d} = \alpha d = \alpha \begin{pmatrix} B^{-1}a_j \\ e_j \end{pmatrix},$$

其中 $\alpha > 0$, $e_j \in \mathbf{R}^{n-m}$ 是单位向量.

1.3.2 凸函数

定义 1.3.4 设 $S \subset \mathbf{R}^n$ 是非空凸集, $\alpha \in (0,1)$, f 是定义在 S 上的函数. 如果对任意 $x_1, x_2 \in S$, 有

$$f(\alpha x_1 + (1-\alpha)x_2) \leqslant \alpha f(x_1) + (1-\alpha)f(x_2), \tag{1.4}$$

则称函数 f 是 S 上的凸函数. 如果当 $x_1 \neq x_2$ 时, (1.4) 中严格不等式成立,

$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha)f(x_2), \tag{1.5}$$

则称 f 为 S 上的严格凸函数. 如果存在一个常数 $c > 0$, 使得对任意 $x_1, x_2 \in S$, 有

$$\begin{aligned} &\alpha f(x_1) + (1-\alpha)f(x_2) \\ \geqslant &f(\alpha x_1 + (1-\alpha)x_2) + c\alpha(1-\alpha)\|x_1 - x_2\|^2, \end{aligned} \tag{1.6}$$

则称 f 在 S 上是一致凸的.

如果 $-f$ 是 S 上的凸（严格凸）函数, 则称 f 为 S 上的凹（严格凹）函数.

凸函数的几何解释告诉我们, 一个凸函数的图形总是位于相应弦的下方. 由凸函数的定义易知, 线性函数 $f(x) = a^{\mathrm{T}}x + \beta$ $(a, x \in \mathbf{R}^n, \beta \in \mathbf{R}^n)$ 在 $\mathbf{R}^n$ 上既是凸函数也是凹函数.

凸函数有如下性质.

定理 1.3.2 (1) 设 f 是定义在凸集 S 上的凸函数, 实数 $\alpha \geqslant 0$, 则 αf 也是定义在 S 上的凸函数.

(2) 设 f_1, f_2 是定义在凸集 S 上的凸函数, 则 $f_1 + f_2$ 也是定义在 S 上的凸函数.

(3) 设 $f_1, f_2, \cdots, f_m$ 是定义在凸集 S 上的凸函数, 实数 $\alpha_1, \alpha_2, \cdots, \alpha_m \geqslant 0$, 则 $\sum_{i=1}^{m} \alpha_i f_i$ 也是定义在 S 上的凸函数.

证明 (2) 设 $x_1, x_2 \in S$ $0 < \alpha < 1$, 则

$$\begin{aligned} &f_1(\alpha x_1 + (1-\alpha)x_2) + f_2(\alpha x_1 + (1-\alpha)x_2) \\ \leqslant &\alpha[f_1(x_1) + f_2(x_1)] + (1-\alpha)[f_1(x_2) + f_2(x_2)]. \end{aligned}$$

性质 (1) 和性质 (3) 的证明留给读者. □

如果凸函数是可微的, 我们可以用下面的特征描述凸函数. 下面的定理刻画了凸函数的一阶特征.

定理 1.3.3 设 $S \subset \mathbf{R}^n$ 为非空开凸集, f 是定义在 S 上的可微函数, 则 f 为凸函数的充分必要条件是

$$f(y) \geqslant f(x) + \nabla f(x)^{\mathrm{T}}(y - x), \quad \forall x, y \in S. \tag{1.7}$$