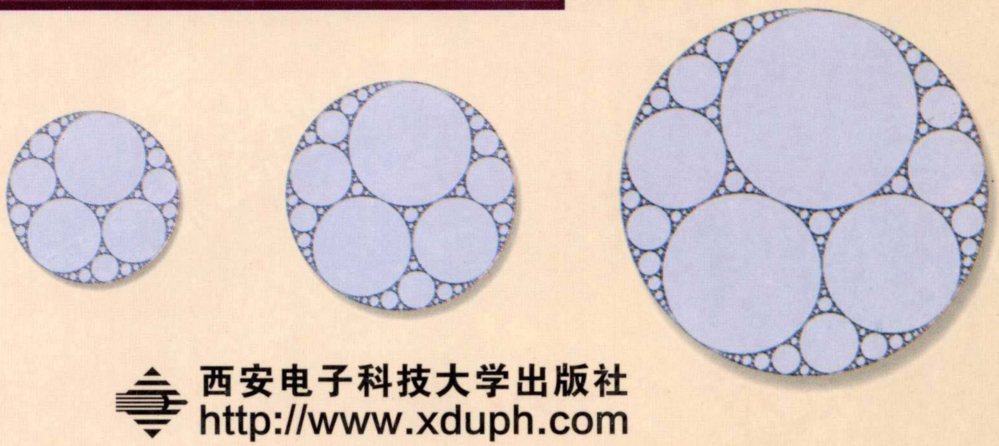
“十二五”重点图书

研究生系列教材

最优化方法

Optimization Methods

宋巨龙 王香柯 冯晓慧 编著



西安电子科技大学出版社
http://www.xduph.com

研究生系列教材

最优化方法

宋巨龙　王香柯　冯晓慧　编　著

西安电子科技大学出版社

内 容 简 介

本书是为工科研究生学习“最优化方法”课程而编写的。全书共七章，主要内容包括最优化方法的基础知识、一维搜索算法、无约束最优化方法、约束非线性最优化方法、线性规划、整数规划等。

本书起点低、跨度大，注重实用性，实例丰富，对算法的几何意义解释透彻，有利于读者掌握最优化方法的基本理论和基本算法。

本书可作为高等学校工科相关专业研究生或理科高年级本科生的教材或教学参考书，也可供工程技术领域的科研人员参考。

图书在版编目(CIP)数据

最优化方法/宋巨龙，王香柯，冯晓慧编著．—西安：西安电子科技大学出版社，2012.9
研究生系列教材
ISBN 978-7-5606-2886-8

Ⅰ．① 最… Ⅱ．① 宋… Ⅲ．① 最优化算法—研究生—教材 Ⅳ．① O242.23

中国版本图书馆 CIP 数据核字(2012)第 168125 号

策划编辑 李惠萍
责任编辑 王 瑛 李惠萍
出版发行 西安电子科技大学出版社(西安市太白南路 2 号)
电　　话 (029)88242885 88201467 邮　　编 710071
网　　址 www.xduph.com 电子邮箱 xdupfxb001@163.com
经　　销 新华书店
印刷单位 陕西天意印务有限责任公司
版　　次 2012 年 9 月第 1 版 2012 年 9 月第 1 次印刷
开　　本 787 毫米×960 毫米 1/16 印张 14
字　　数 278 千字
印　　数 1～2000 册
定　　价 23.00 元
ISBN 978-7-5606-2886-8 /O·0136
XDUP 3178001-1

“十二五”重点图书　　研究生系列教材

编审委员会名单

前　言

最优化方法是现代数学的一个重要组成部分，它在自然科学、社会科学中有着广泛的应用，特别是在工程设计和现代管理中有着很高的应用价值。由于许多工程问题最终都归结为优化问题，所以半个多世纪以来最优化方法得到了迅速的发展和广泛的应用。

鉴于现在各高校日益扩大的研究生招生规模，不同领域、不同类型的研究生的需求有所不同，作者在多年“最优化方法”课程的教学基础上，根据高等学校工科研究生的教学需求编写了本书。本书不循大而全的编写模式，不追求遍历所有的优化方法和优化理论，而是紧紧围绕工科研究生的基本需求，本着起点低、跨度大，注重实用性，适当兼顾理论体系的原则来编写的。考虑到有不少研究生是工作多年后再次进入学校深造，许多基础知识已经淡忘，所以本书适当地加强了对线性代数及高等数学等先修课程内容的回顾性介绍。本书注重用实例来解析算法原理，书中的例题和习题都经过了严格的验算，力求做到准确无误。

本书以算法的实用性为主，详细地介绍了最优化方法的基本理论和基本算法。对于大多数算法，本书都给出了实例，以对算法进行说明；对于少数算法，则完全通过例题来阐述其原理和方法。书中特别对基本算法的原理都尽量给出几何解释，有利于读者对算法的理解。本书对算法的理论部分做了适当的介绍，对主要定理进行了证明，理论性过强的定理则略去，并且简单而不加证明地介绍了算法的收敛性。每章末均配有适当数量的习题，便于读者通过练习来更好地掌握所学内容，书末还附有部分习题参考答案。

全书由七章和两个附录构成。为方便读者对最优化问题做进一步深入的研究，附录一列举了一些常用的测试函数；为方便学生编写算法程序，附录二给出了几个基本算法的程序，供学生参考。

本书由宋巨龙统稿，并编写了第一章～第五章、第六章的大部分内容以及附录部分；王香柯编写了第六章的6.7节以及第七章并参与了全书的校对工作；冯晓慧参与了第三章和第五章的编写工作。

本书可作为高等学校工科相关专业研究生或理科高年级本科生的教材，也可供工程技术领域的科研人员参考。本书建议安排54学时，各学校也可根据学生的具体情况增减学时。

本书是在西安电子科技大学出版社李惠萍编辑和王瑛编辑的大力帮助下完成的，在此向她们表示衷心的感谢。

由于编者水平有限，加之编写时间仓促，书中难免有不妥之处，敬请广大读者批评指正。

编　者

2012年5月

本书主要符号

$\mathbf{R}^n$——n 维实向量空间。

$\boldsymbol{x}$，$\boldsymbol{x}^*$，$\boldsymbol{x}^k(k=1,2,\cdots,n)$——$n$ 维实列向量：

$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$；

$\boldsymbol{x}^*=(x_1^*,x_2^*,\cdots,x_n^*)^{\mathrm{T}}$；

$\boldsymbol{x}^k=(x_1^k,x_2^k,\cdots,x_n^k)^{\mathrm{T}}$。

$\boldsymbol{x}^{\mathrm{T}}$——向量的转置。

$\boldsymbol{x}>\mathbf{0}$——向量 $\boldsymbol{x}$ 的各分量均为正数。

$\boldsymbol{x}\geqslant\mathbf{0}$——向量 $\boldsymbol{x}$ 的各分量均为非负数。

$\boldsymbol{x}<\mathbf{0}$——向量 $\boldsymbol{x}$ 的各分量均为负数。

$\boldsymbol{x}\leqslant\mathbf{0}$——向量 $\boldsymbol{x}$ 的各分量均为非正数。

$\boldsymbol{x}\leqslant\boldsymbol{y}$——向量 $\boldsymbol{x}$ 的各分量均小于等于向量 $\boldsymbol{y}$ 的各分量。

$\boldsymbol{x}<\boldsymbol{y}$——向量 $\boldsymbol{x}$ 的各分量均小于向量 $\boldsymbol{y}$ 的各分量。

$\boldsymbol{x}\in D$——向量 $\boldsymbol{x}$ 属于向量集合 D。

$\boldsymbol{x}\notin D$——向量 $\boldsymbol{x}$ 不属于向量集合 D。

$\mathrm{int}D$——由向量集合 D 的内点所构成的集合。

$k\in K$——元素 k 属于集合 K。

$k\notin K$——元素 k 不属于集合 K。

$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$——在没有说明的情况下，大写字母表示矩阵。

$\boldsymbol{O}$——零矩阵。

$\boldsymbol{A}>0$——矩阵 $\boldsymbol{A}$ 正定。

$\boldsymbol{A}\geqslant0$——矩阵 $\boldsymbol{A}$ 半正定。

$\boldsymbol{A}<0$——矩阵 $\boldsymbol{A}$ 负定。

$\boldsymbol{A}\leqslant0$——矩阵 $\boldsymbol{A}$ 半负定。

$\forall\boldsymbol{x}\in\mathbf{R}^n$——属于 $\mathbf{R}^n$ 的任意 n 维实向量。

$\exists\boldsymbol{x}\in D$——存在 $\boldsymbol{x}$ 属于 D。

$\|\boldsymbol{x}\|$——向量 $\boldsymbol{x}$ 的欧氏范数。

s. t.——“满足约束”，“subject to”的缩写。

$N_\delta(\boldsymbol{x}^0)$或 $N(\boldsymbol{x}^0,\delta)$——以$\boldsymbol{x}^0$ 为中心，以 δ 为半径的开邻域。

$\nabla f(\boldsymbol{x})$——$f(\boldsymbol{x})$在 $\boldsymbol{x}$ 处的梯度。

$\nabla^2 f(\boldsymbol{x})$——$f(\boldsymbol{x})$在 $\boldsymbol{x}$ 处的 Hesse 矩阵。

$\boldsymbol{h}(\boldsymbol{x})$——向量函数$\boldsymbol{h}(\boldsymbol{x})=(h_1(\boldsymbol{x}),h_2(\boldsymbol{x}),\cdots,h_n(\boldsymbol{x}))^{\mathrm{T}}$。

目 录

第一章 绪 论

最优化的思想方法最早可以追溯到牛顿(Newton，1643—1727)、拉格朗日(Lagrange，1736—1813)和柯西(Cauchy，1789—1857)时代。牛顿和莱布尼茨(Leibniz，1646—1716)所创立的微积分中已经部分地包含有优化思想；而伯努利(Bernoulli，1654—1705)、欧拉(Euler，1707—1783)和拉格朗日等人则对变分法的建立作出了各自不同的贡献；柯西最早提出求多元函数最优值的最速下降法。

现代最优化方法起源于第一次世界大战时期。随着社会的进步，生产力的不断发展，优化问题在生产实践中的重要性日益突显。许多军事领域的优化问题，如搜索潜艇、护航、布雷、轰炸、运输管理等问题的提出，促使优化问题在第一次世界大战以后得到人们日益强烈的关注。到 20 世纪 60 年代，最优化方法最终发展成为一门新兴的基础学科。

1930 年，前苏联数学家康托罗维奇(Cantolovch)提出二维线性规划问题的图上作业法。1947 年，美国数学家丹茨(Dantzig)提出了轰动数学界的单纯形法，为解决高维线性规划问题提供了一种有效的工具。1950 年至 1965 年，匈牙利数学家库恩(Kuhn)和塔克(Tucker)建立了线性规划的对偶理论，为求解鞍点问题提供了数学工具，这两位年轻的数学家建立的约束极值问题的最优性条件被称为 K－T 条件，为求解非线性规划问题奠定了基础。

随着科学技术的日益发展，许多工程的核心问题最终都归结为优化问题。因此，最优化方法已经成为工程技术人员必不可少的计算工具。在计算机已经广为普及的今天，一些大规模的优化问题的求解可以在一台普通的计算机上实现，使得最优化方法得到了比以往任何时候都更加广泛的应用。如今，最优化方法已成为工程技术人员所必需具备的研究工具。

最优化方法主要是研究在一定条件限制下，选取某种方案以使某目标达到最优的一门学科。使目标达到最优的方案称为最优方案，而获取最优方案的方法称为最优化方法。这种方法的数学理论则称为最优化理论。最优化方法在当今的军事、工程、管理等领域有着极其广泛的应用。

1.1 最优化问题举例

【例 1.1】 (配料问题) 某工厂生产 A 和 B 两种产品，所用原料均为甲、乙、丙三种。每生产一个产品具体用料情况如表 1.1 所示。已知产品 A 每件可获利 9000 元，产品 B 每件可获利 15 000 元。如果该厂现有原料：甲 450，乙 150，丙 300。在现有条件下，A、B 各生产多

少才可使该厂获利最大?

表 1.1 例 1.1 的原料构成

	甲	乙	丙
A	8	4	13
B	5	9	6

解 设需生产 A 产品 x_1 件,B 产品 x_2 件,则总获利为

$$s = 0.9x_1 + 1.5x_2 \text{ (万元)}$$

其中 x_1,x_2 应满足约束条件:

$$\begin{cases} 8x_1 + 5x_2 \leqslant 450 \\ 4x_1 + 9x_2 \leqslant 150 \\ 13x_1 + 6x_2 \leqslant 300 \\ x_1 \geqslant 0,\ x_2 \geqslant 0 \end{cases}$$

该问题可表示为

$$\begin{cases} \max & s = 9x_1 + 15x_2 \\ \text{s. t.} & 8x_1 + 5x_2 \leqslant 450 \\ & 4x_1 + 9x_2 \leqslant 150 \\ & 13x_1 + 6x_2 \leqslant 300 \\ & x_1 \geqslant 0,\ x_2 \geqslant 0 \end{cases}$$

【例 1.2】 (投资问题)某公司现有资金 a 万元,目前有 $n(n \geqslant 2)$ 个投资项目可供选择,设投资第 i 个项目要用资金 a_i 万元,可获利 b_i 万元,$i = 1,2,\cdots,n$。如何投资收益最大?

解 设第 i 个项目的选择变量为 x_i,即

$$x_i = \begin{cases} 1 & \text{(如果选择投资第 } i \text{ 个项目)} \\ 0 & \text{(如果不选择投资第 } i \text{ 个项目)} \end{cases}$$

这里 $i = 1, 2, \cdots, n$。则总的投资金额应满足 $\sum\limits_{i=1}^{n} a_i x_i \leqslant a$,总的收益为 $\sum\limits_{i=1}^{n} b_i x_i$。由于 x_i 只能取 0 和 1,所以应有 $x_i(x_i - 1) = 0$,$i = 1,2,\cdots,n$。投资的目标是:在投资额不超过现有资金的条件下使收益最大。即本投资问题表示为

$$\begin{cases} \max \sum\limits_{i=1}^{n} b_i x_i \\ \text{s. t.} \sum\limits_{i=1}^{n} a_i x_i \leqslant a \\ \quad\ x_i(x_i - 1) = 0 \qquad (i = 1,2,\cdots,n) \end{cases}$$

【例 1.3】 欲将一个宽为 a、长为 l 的矩形铁皮折成一个截面为等腰梯形的水槽,如何取

水槽的底长以及水槽侧边折起来离开底面的夹角可使水槽的截面积最大？

解　如图 1－1 所示，设水槽的底长为 x，水槽侧边与底面的夹角为 θ，则水槽的截面积为

$$S=\frac{2x+(a-x)\cos\theta}{4}(a-x)\sin\theta$$

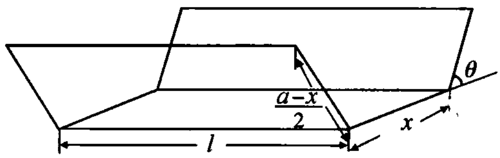


图 1－1　例 1.3 的水槽示意图

这里 $0\leqslant x\leqslant a$，$0\leqslant\theta\leqslant\frac{\pi}{2}$。合并起来就是

$$\begin{cases}\max\quad S=\dfrac{2x+(a-x)\cos\theta}{4}(a-x)\sin\theta\\ \text{s.t.}\quad 0\leqslant x\leqslant a\\ \qquad\ \ 0\leqslant\theta\leqslant\dfrac{\pi}{2}\end{cases}$$

【例 1.4】（运输问题）已知某省煤炭有 m 个产地 $A_1,A_2,\cdots,A_m$，产量分别记为 $a_1,a_2,\cdots,a_m$（吨）；有 n 个销售地 $B_1,B_2,\cdots,B_n$，每个销售地的需求量记为 $b_1,b_2,\cdots,b_n$（吨）。假定产销是平衡的，即 $\sum_{i=1}^{m}a_i=\sum_{j=1}^{n}b_j$，由 A_i 地到 B_j 地的运费为每吨 $c_{ij}\,(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 元。由每个产地向各销售地运多少吨煤时，既可保证满足需求量，又可使运费最省？

解　设应从 A_i 地到 B_j 地运送 x_{ij} 吨煤，则总的运费为

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

这里 x_{ij} 应满足

$$\begin{cases}\sum\limits_{j=1}^{n}x_{ij}=a_i & (i=1,2,\cdots,m)\\ \sum\limits_{i=1}^{m}x_{ij}=b_j & (j=1,2,\cdots,n)\\ x_{ij}\geqslant 0 & (i=1,2,\cdots,m;j=1,2,\cdots,n)\end{cases}$$

所以该问题可表示为

$$\begin{cases} \min \quad s = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \\ \text{s.t.} \quad \sum_{j=1}^{n} x_{ij} = a_i \qquad (i = 1, 2, \cdots, m) \\ \qquad \sum_{i=1}^{m} x_{ij} = b_j \qquad (j = 1, 2, \cdots, n) \\ \qquad x_{ij} \geqslant 0 \qquad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n) \end{cases}$$

【例 1.5】 设有非线性方程组

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \qquad \vdots \\ f_m(x_1, x_2, \cdots, x_n) = 0 \end{cases}$$

求该非线性方程组的解。

解 可将该问题等价地转化为优化问题：

$$\min f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{m} f_i^2(x_1, x_2, \cdots, x_n)$$

如果原方程组有解，则本优化问题应有最小值 0；反之，若本优化问题有最小值 0，则本优化问题的解就是原方程组的解。如果本优化问题的最小值不为 0，则说明原方程组没有解。

1.2 最优化问题的数学模型及其分类

以上列举了几个最优化问题的例子，它们的共同特点是：求满足一定条件的变量 x_1，x_2，…，x_n，使某函数 $f(x_1, x_2, \cdots, x_n)$ 取得最大值或者最小值。由于 $f(x_1, x_2, \cdots, x_n)$ 的最大问题可以转化为 $-f(x_1, x_2, \cdots, x_n)$ 的最小问题，所以较多时候只讨论最小问题。这里的函数 $f(x_1, x_2, \cdots, x_n)$ 称为**目标函数**或者**评价函数**；变量 $x_1, x_2, \cdots, x_n$ 称为**决策变量**；需要满足的条件称为**约束条件**；用以构成约束条件的函数称为**约束函数**。

对于优化问题通常采用如下方法进行分类。

1. 根据优化问题约束类型的不同进行分类

1）无约束问题

求 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 使函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$ 达到最小值，记为 $\min f(\boldsymbol{x})$。

2）约束问题

根据约束函数的类型又可进一步分为以下几类。

（1）等式约束问题：求 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 使其在满足 l 个等式约束条件 $h_j(\boldsymbol{x}) = 0$，$j = 1, 2, \cdots, l$ 的情况下，使函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$ 达到最小值，记为

$$\begin{cases}\min & f(\boldsymbol{x}) \\ \text{s.t.} & h_j(\boldsymbol{x}) = 0 \qquad (j = 1,2,\cdots,l)\end{cases}$$

(2) 不等式约束问题：求 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 使其在满足 m 个不等式约束条件 $g_i(\boldsymbol{x}) \geqslant 0$，$i=1,2,\cdots,m$ 的情况下，使函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$ 达到最小值，记为

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

(3) 混和约束问题或称一般约束问题：求 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 使其在满足 m 个不等式约束条件 $g_i(\boldsymbol{x}) \geqslant 0$，$i = 1,2,\cdots,m$ 以及 l 个等式约束条件 $h_j(\boldsymbol{x}) = 0$，$j = 1,2,\cdots,l$ 的情况下，使函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$ 达到最小值，记为

$$\begin{cases}\min & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) \geqslant 0 \qquad (i = 1,2,\cdots,m) \\ & h_j(\boldsymbol{x}) = 0 \qquad (j = 1,2,\cdots,l)\end{cases}$$

以上各问题中的函数 $f(x_1, x_2, \cdots, x_n)$ 称为**目标函数**，函数 $g_i(\boldsymbol{x})$、$h_j(\boldsymbol{x})$ 称为约束函数。满足约束条件的点 $\boldsymbol{x}$ 构成的集合，称为**可行解集合**，亦称**可行区**或**可行域**。例如：

$$D_1 = \{\boldsymbol{x} \mid g_i(\boldsymbol{x}) \geqslant 0, i = 1,2,\cdots,m\},\ D_2 = \{\boldsymbol{x} \mid h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l\}$$

分别是上述不等式约束问题和等式约束问题的可行域。

以上给出的是标准的优化问题。实际上某些不标准的问题可以化为标准的问题。例如：

(1) 约束条件 $g_i(\boldsymbol{x}) \leqslant 0$ 可以化为 $-g_i(\boldsymbol{x}) \geqslant 0$；

(2) 最大值问题可以化为最小值问题：$\max\limits_{x \in D} f(\boldsymbol{x}) = -\min\limits_{x \in D}[-f(\boldsymbol{x})]$；

(3) 等式约束可以化为不等式约束：$h_j(\boldsymbol{x}) = 0 \Leftrightarrow h_j(\boldsymbol{x}) \geqslant 0$ 且 $-h_j(\boldsymbol{x}) \geqslant 0$。

因此上述分类可以简化为下面两类：

(1) 无约束最优化问题：$\min f(\boldsymbol{x})$；

(2) 约束优化问题：$\begin{cases}\min & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) \geqslant 0 \qquad (i = 1,2,\cdots,m)\end{cases}$。

2. 根据目标函数及约束函数的类型进行分类

最优化问题也称为规划问题。

如果最优化问题的目标函数为 $f(\boldsymbol{x})$，约束条件为 $g_i(\boldsymbol{x}) \geqslant 0$，$i = 1,2,\cdots,m$，则：

当 $f(\boldsymbol{x})$ 和 $g_i(\boldsymbol{x})$ 均为线性函数时，称此最优化问题为**线性规划**；

当 $f(\boldsymbol{x})$ 和 $g_i(\boldsymbol{x})$ 不全为线性函数时，称此最优化问题为**非线性规划**；

当 $f(\boldsymbol{x})$ 为二次函数，而 $g_i(\boldsymbol{x})$ 全为线性函数时，称此最优化问题为**二次规划**。

3. 根据变量的类型进行分类

对于最优化问题，如果变量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 的各分量只能取整数，则相应的最优

化问题称为**整数规划**。如果变量$\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^{\mathrm{T}}$的部分分量只能取整数,则相应的最优化问题称为**混合整数规划**。如果变量$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$的各分量只能取0和1，则相应的最优化问题称为**0-1规划**。

4. 其他分类方法

如果最优化问题的目标函数不止一个，则称该优化问题为多目标规划。

如果最优化问题需要根据其特性将决策过程按时间或者空间分为若干个相互联系又相互区别的阶段，在它的每一个阶段都需要做出决策，从而使整个决策过程达到最好的效果，则这样的最优化问题称为动态规划。

1.3 最优化问题的最优解及最优值

设最优化问题为

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

即

$$(\mathrm{P})\ \min_{\boldsymbol{x}\in D} f(\boldsymbol{x}), D=\{\boldsymbol{x}\mid g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m\}$$

定义1.1　如果有$\boldsymbol{x}^*\in D$使得$f(\boldsymbol{x}^*)=\min\limits_{\boldsymbol{x}\in D}f(\boldsymbol{x})$，即$\exists \boldsymbol{x}^*\in D$，使得对$\forall \boldsymbol{x}\in D$有$f(\boldsymbol{x})\geqslant f(\boldsymbol{x}^*)$，则称$\boldsymbol{x}^*$为问题(P)的**全局最优解**，称$f(\boldsymbol{x}^*)$为**全局最优值**。

在上述定义中，如果当$\forall \boldsymbol{x}\in D$且$\boldsymbol{x}\neq\boldsymbol{x}^*$时恒有$f(\boldsymbol{x})>f(\boldsymbol{x}^*)$，则称$\boldsymbol{x}^*$为问题(P)的**严格全局最优解**，称$f(\boldsymbol{x}^*)$为**严格全局最优值**。

定义1.2　如果有$\boldsymbol{x}^*\in D$及$\delta>0$，使得当$\boldsymbol{x}\in D\cap N_\delta(\boldsymbol{x}^*)$时恒有$f(\boldsymbol{x})\geqslant f(\boldsymbol{x}^*)$，则称$\boldsymbol{x}^*$为问题(P)的**局部最优解**，称$f(\boldsymbol{x}^*)$为**局部最优值**。

这里的$N_\delta(\boldsymbol{x}^*)=\{\boldsymbol{x}\mid \|\boldsymbol{x}-\boldsymbol{x}^*\|<\delta\}$为$\boldsymbol{x}^*$的$\delta$邻域。范数$\|\cdot\|$指的是$\|\boldsymbol{x}\|=\sqrt{x_1^2+\cdots+x_n^2}$，本书后面所提到的范数均为此范数。

同样，上述定义中，如果当$\boldsymbol{x}\neq\boldsymbol{x}^*$时可将“$\geqslant$”改为“$>$”，则称$\boldsymbol{x}^*$为问题(P)的**严格局部最优解**，称$f(\boldsymbol{x}^*)$为**严格局部最优值**。

习　题　一

1.1　(指派问题)有n项工作$B_1,B_2,\cdots,B_n$，要分配给n个人$A_1,A_2,\cdots,A_n$，已知第i个人A_i做第j项工作B_j的费用为$c_{ij}(i,j=1,2,\cdots,n)$。如何分配工作可使费用最小？试写出此问题的数学模型。

1.2　设有一圆形材料，其半径为r，今欲将其截去一个圆心角为α的小扇形，而将剩

余的部分做成一个锥形容器，截去的小扇形的圆心角为多大时可使做成的容器容积最大？试写出此问题的数学模型。

1.3　三条边分别为 a、b、c 的三角形的面积为 $A=\sqrt{s(s-a)(s-b)(s-c)}$，今有长为 50 m 的栅栏，如何选取三条边可使三角形面积最大？

1.4　某公司计划花 1000 万元做广告，现有四种广告媒体：广播、电视、报纸、户外，若第 i 种媒体每花 x_i 万元可得收益由表 1.2 给出，如何分派广告费可使受益最大？

表 1.2　习题 1.4 的收益构成

广播	$8x_1^{0.5}$	报纸	$3x_3^{0.6}$
电视	$2x_2^{0.3}$	户外	$11x_4^{0.4}$

1.5　某省欲在甲、乙、丙三个城市中间建一个机场，如果三个城市的位置分别为 $\boldsymbol{x}^1=(60,30)^{\mathrm{T}}$，$\boldsymbol{x}^2=(0,0)^{\mathrm{T}}$，$\boldsymbol{x}^3=(0,50)^{\mathrm{T}}$，则机场建在什么位置可使机场到三个城市的距离之和最小？

1.6　欲设计一个圆柱形容器，其高和底半径都不超过 20 cm，底半径至少为 5 cm。高和底半径的比值应介于 2.4～3.4 之间，而侧面积不应超过 900 cm^2。如何取高和底半径可使容器容积最大？

第二章　最优化方法的基础知识

鉴于最优化问题的求解过程中要用到一些线性代数和高等数学的知识，而有些读者是工作多年后再接触最优化问题的，所以对最优化问题所涉及的一些基础知识作一个回顾是很有必要的。本章介绍在最优化方法中用到的线性代数和高等数学的一些基础知识，对大部分结论不给证明过程。以下讨论均限定在实数范围内，除非有特殊说明，数都是实数，向量都是实向量，函数也都是实函数，不再一一说明。

2.1　二次型和正定矩阵

向量和矩阵是我们讨论最优化问题的必要工具。向量在许多课程中都要用到，相信读者应该很熟悉了，这里就不作介绍了。下面我们对矩阵的正定性作一个回顾。

定义 2.1　含有 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次多项式：

$$\begin{aligned}f(x_1,x_2,\cdots,x_n) &= a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n\\&\quad+a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n\\&\quad\cdots\\&\quad+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\end{aligned}$$

其中 $a_{ij}=a_{ji}(i,j=1,2,\cdots,n)$ 为实数，称为一个 n 元实**二次型**。

如果记 $\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$，$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$，则 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，从而二次型可写为

$$f(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$$

称实对称矩阵 $\boldsymbol{A}$ 为实二次型的矩阵。实二次型和实对称矩阵之间是一一对应的。每给出一个实二次型，就能决定一个实对称矩阵；反之，每给出一个实对称矩阵，就能决定一

个实二次型。

定义 2.2　设有二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，如果对任意非零向量 $\boldsymbol{x}$ 都有：

(1) $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$，则称该二次型为**正定的**；

(2) $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geqslant 0$，则称该二次型为**半正定的**；

(3) $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} < 0$，则称该二次型为**负定的**；

(4) $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \leqslant 0$，则称该二次型为**半负定的**；

(5) 对于满足(1)～(4)情形的二次型，称其为**有定的**，如果不是上述四种情形，则称该二次型为**不定的**。

由于实二次型和实对称矩阵之间具有一一对应关系，所以：

(1) 当二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定时，称实对称矩阵 $\boldsymbol{A}$ 为**正定的**，记为 $\boldsymbol{A} > 0$；

(2) 当二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为负定时，称实对称矩阵 $\boldsymbol{A}$ 为**负定的**，记为 $\boldsymbol{A} < 0$；

(3) 当二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定时，称实对称矩阵 $\boldsymbol{A}$ 为**半正定的**，记为 $\boldsymbol{A} \geqslant 0$；

(4) 当二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半负定时，称实对称矩阵 $\boldsymbol{A}$ 为**半负定的**，记为 $\boldsymbol{A} \leqslant 0$；

(5) 当二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为不定时，称实对称矩阵 $\boldsymbol{A}$ 为**不定的**。

从定义不难看出：当 $\boldsymbol{A}$ 正定时，$-\boldsymbol{A}$ 必为负定的；当 $\boldsymbol{A}$ 半正定时，$-\boldsymbol{A}$ 必为半负定的。

对于实对称矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称行列式

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} \qquad (k = 1,2,\cdots,n)$$

为 $\boldsymbol{A}$ 的 k 阶顺序主子式，而称行列式

$$\begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{vmatrix} \qquad (1 \leqslant i_1, i_2, \cdots, i_k \leqslant n; k = 1,2,\cdots,n)$$

为 $\boldsymbol{A}$ 的 k 阶主子式，亦即 $\boldsymbol{A}$ 中行列位置相同的子式称为主子式。

定理 2.1　设 $\boldsymbol{A}$ 为 n 阶实对称矩阵，下列命题是等价的：

(1) 对任意 n 维非零向量 $\boldsymbol{x}$，恒有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$；