

■ 高等学校教材

# 最优化方法 及其应用

■ 郭科 陈聆 魏友华



高等教育出版社  
HIGHER EDUCATION PRESS

0224/55

2007

高等学校教材

# 最优化方法及其应用

郭 科 陈 聆 魏友华

高等教育出版社

## 内容简介

本书系统地介绍了最优化的理论和计算方法,在编写上遵循循序渐进、由浅入深、注重概念、突出方法的原则。本书将最优化技术与计算机技术融为一体,对最优化技术的理论作了适当深度的讨论,重点在于对概念和方法的论述;在应用方面,着重强调方法与应用的有机结合。

全书共十章,包括最优化问题总论、最优化问题数学基础、线性规划及其对偶问题、一维搜索法、常用无约束最优化方法、常用约束最优化方法、动态规划、多目标优化、现代优化算法简介、最优化问题程序设计方法等,其中前八章为传统优化算法,也是本书重点介绍的内容,最后一章还给出了部分优化问题的设计实例。

本书可作为数学类各专业本科高年级学生教材,也可供一般工科研究生以及数学建模竞赛参赛人员和工程技术人员参考。

### 图书在版编目(CIP)数据

最优化方法及其应用/郭科,陈聆,魏友华. —北京:高等教育出版社,2007.7

ISBN 978-7-04-021930-2

I. 最... II. ①郭...②陈...③魏... III. 最佳化-高等学校-教材 IV. O224

中国版本图书馆 CIP 数据核字(2007)第 071196 号

策划编辑 张长虹 责任编辑 崔梅萍 封面设计 王凌波  
责任绘图 尹莉 版式设计 张岚 责任校对 王雨  
责任印制 毛斯璐

---

出版发行	高等教育出版社	购书热线	010-58581118
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	<a href="http://www.hep.edu.cn">http://www.hep.edu.cn</a>
总 机	010-58581000		<a href="http://www.hep.com.cn">http://www.hep.com.cn</a>
经 销	蓝色畅想图书发行有限公司	网上订购	<a href="http://www.landaco.com">http://www.landaco.com</a>
印 刷	北京宏伟双华印刷有限公司		<a href="http://www.landaco.com.cn">http://www.landaco.com.cn</a>
		畅想教育	<a href="http://www.widedu.com">http://www.widedu.com</a>
开 本	787×960 1/16	版 次	2007 年 7 月第 1 版
印 张	14.5	印 次	2007 年 7 月第 1 次印刷
字 数	260 000	定 价	18.50 元

---

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 21930-00

# 前 言

今天,最优化问题几乎已经渗透到管理、经济和工程技术等领域的各个方面。现代科学技术特别是计算机技术的迅速发展,为求解最优化问题提供了雄厚的基础和有效的手段.因此,掌握好最优化技术具有十分重要的现实意义。

本书系统地介绍了最优化的理论和计算方法,在内容上遵循循序渐进、由浅入深、注重概念、突出方法的原则。本书将最优化技术与计算机技术融为一体,对最优化技术的理论作了适度的讨论,重点在于对概念和方法的论述。在应用方面,本书着重强调方法与应用的有机结合。

全书共十章。第一章介绍最优化问题一般理论;第二章介绍最优化问题的数学基础;第三章介绍线性规划及其对偶问题;第四章介绍求解一维最优化问题的一维搜索法;第五章介绍常用的无约束最优化方法,包括最速下降法、Newton法、修正 Newton 法、共扼方向法、共扼梯度法、变尺度法、坐标轮换法和单纯形法;第六章介绍常用的约束最优化方法,包括外点罚函数法、内点罚函数法、混合罚函数法、约束坐标轮换法和复合形法;第七章介绍动态规划;第八章介绍多目标优化问题;第九章介绍现代优化算法,包括模拟退火算法、遗传算法、禁忌搜索算法和人工神经网络算法;第十章介绍优化问题应用案例,包括最优化问题建模的一般步骤。其中前八章为传统优化算法,也是本书重点介绍的内容。

本书主要为数学类各专业本科高年级学生编写,也可供一般工科研究生、数学建模竞赛参赛人员以及工程技术人员参考。使用本书时,根据学时安排,可将第三、七、八、九章作为选学内容,第十章的应用实例可部分选学。

本书由成都理工大学郭科教授担任主编,郭科、陈聆、魏友华执笔编写。曹金文教授认真审阅了全书,提出了不少改进意见。在此,编者向他表示衷心的感谢,同时还要感谢魏贵民教授在本书的编写工作中给予的支持和帮助。本书编写过程中参阅了许多专家和学者的论著文献,并引用了部分文献中的实例,恕不一一指明出处,在此一并向有关作者致谢!

我们力图使本书成为一本具有改革新意和易教易学的教材,但由于编者水平所限,疏漏及不当之处,请读者批评指正。

编 者  
2006 年 2 月

# 目 录

<b>第一章</b>	<b>最优化问题总论</b> .....	1
§ 1.1	最优化问题数学模型 .....	1
§ 1.2	最优化问题的算法 .....	6
§ 1.3	最优化算法分类 .....	12
§ 1.4	组合优化问题简介 .....	13
	习题一 .....	16
<b>第二章</b>	<b>最优化问题的数学基础</b> .....	18
§ 2.1	二次型与正定矩阵 .....	18
§ 2.2	方向导数与梯度 .....	19
§ 2.3	Hesse 矩阵及泰勒展式 .....	22
§ 2.4	极小点的判定条件 .....	26
§ 2.5	锥、凸集、凸锥 .....	28
§ 2.6	凸函数 .....	32
§ 2.7	约束问题的最优性条件 .....	36
	习题二 .....	45
<b>第三章</b>	<b>线性规划及其对偶问题</b> .....	46
§ 3.1	线性规划数学模型基本原理 .....	46
§ 3.2	线性规划迭代算法 .....	50
§ 3.3	对偶问题的基本原理 .....	58
§ 3.4	线性规划问题的灵敏度 .....	65
	习题三 .....	68
<b>第四章</b>	<b>一维搜索法</b> .....	71
§ 4.1	搜索区间及其确定方法 .....	72
§ 4.2	对分法 .....	75
§ 4.3	Newton 切线法 .....	76
§ 4.4	黄金分割法 .....	78
§ 4.5	抛物线插值法 .....	81
	习题四 .....	84

<b>第五章</b>	<b>常用无约束最优化方法</b> .....	86
	§ 5.1 最速下降法 .....	86
	§ 5.2 Newton 法 .....	90
	§ 5.3 修正 Newton 法 .....	93
	§ 5.4 共轭方向法 .....	94
	§ 5.5 共轭梯度法 .....	99
	§ 5.6 变尺度法 .....	103
	§ 5.7 坐标轮换法 .....	111
	§ 5.8 单纯形法 .....	115
	习题五 .....	119
<b>第六章</b>	<b>常用约束最优化方法</b> .....	120
	§ 6.1 外点罚函数法 .....	120
	§ 6.2 内点罚函数法 .....	125
	§ 6.3 混合罚函数法 .....	128
	§ 6.4 约束坐标轮换法 .....	130
	§ 6.5 复合法 .....	132
	习题六 .....	136
<b>第七章</b>	<b>动态规划</b> .....	138
	§ 7.1 动态规划基本原理 .....	138
	§ 7.2 动态规划迭代算法 .....	142
	§ 7.3 动态规划有关说明 .....	145
	习题七 .....	145
<b>第八章</b>	<b>多目标优化</b> .....	146
	§ 8.1 多目标最优化问题的基本原理 .....	146
	§ 8.2 评价函数法 .....	150
	§ 8.3 分层求解法 .....	157
	§ 8.4 目标规划法 .....	159
	习题八 .....	161
<b>第九章</b>	<b>现代优化算法简介</b> .....	163
	§ 9.1 模拟退火算法 .....	163
	§ 9.2 遗传算法 .....	168
	§ 9.3 禁忌搜索算法 .....	177
	§ 9.4 人工神经网络 .....	182

---

第十章 最优化问题程序设计方法 .....	192
§ 10.1 最优化问题建模的一般步骤 .....	192
§ 10.2 常用最优化方法的特点及选用标准 .....	193
§ 10.3 最优化问题编程的一般过程 .....	196
§ 10.4 优化问题设计实例 .....	198
参考文献 .....	219

# 第一章 最优化问题总论

无论做什么事,人们总希望以最少的代价取得最大的效益,也就是力求最好,这就是优化问题.最优化就是在一切可能的方案中选择一个最好的方案以达到最优目标.例如,从甲地到乙地有公路、水路、铁路、航空四种走法,如果我们追求的目标是省钱,那么只要比较一下这四种走法的票价,从中选择最便宜的那一种走法就达到目标.这是最简单的最优化问题,实际优化问题一般都比较复杂.

概括地说,凡是追求最优目标的数学问题都属于最优化问题.作为最优化问题,一般要有三个要素:第一是目标;第二是方案;第三是限制条件.而且目标应是方案的“函数”.如果方案与时间无关,则该问题属于静态最优化问题;否则称为动态最优化问题.

## § 1.1 最优化问题数学模型

最简单的最优化问题实际上在高等数学中已遇到,这就是所谓函数极值,我们习惯上又称之为经典极值问题.

**例 1.1** 对边长为  $a$  的正方形铁板,在四个角处剪去相等的正方形以制成方形无盖水槽,问如何剪法使水槽的容积最大?

**解** 设剪去的正方形边长为  $x$ ,由题意易知,与此相应的水槽容积为

$$f(x) = (a - 2x)^2 x.$$

$$\text{令 } f'(x) = 2(a - 2x)(-2)x + (a - 2x)^2 = (a - 2x)(a - 6x) = 0,$$

得两个驻点:  $x = \frac{1}{2}a, x = \frac{1}{6}a$ .

第一个驻点不合实际,这是因为剪去 4 个边长为  $\frac{a}{2}$  的正方形相当于将铁板全部剪去.现在来判断第二个驻点是否为极大点.

$$\text{因为 } f''(x) = 24 - 8a, \quad f''\left(\frac{a}{6}\right) = -4a < 0,$$

所以  $x = \frac{a}{6}$  是极大点.

因此,每个角剪去边长为  $\frac{a}{6}$  的正方形可使所制成的水槽容积最大.

**例 1.2** 求侧面积为常数  $6a^2 (a > 0)$ , 体积最大的长方体体积.

**解** 设长方体的长、宽、高分别为  $x, y, z$ , 体积为  $V$ , 则依题意知体积为



$$V = f(x, y, z) = xyz,$$

条件为

$$\varphi(x, y, z) = 2(yz + zx + xy) - 6a^2 = 0.$$

由拉格朗日乘数法, 考虑函数

$$F(x, y, z) = xyz + \lambda(2yz + 2zx + 2xy - 6a^2),$$

$$F'_x = yz + 2\lambda(y + z) = 0,$$

$$F'_y = xz + 2\lambda(z + x) = 0,$$

$$F'_z = xy + 2\lambda(x + y) = 0,$$

由题意可知  $x, y, z$  应是正数, 由此  $\lambda \neq 0$ , 将上面三个等式分别乘以  $x, y, z$  并利用条件  $yz + zx + xy = 3a^2$ , 得到

$$\begin{cases} xyz + 2\lambda(3a^2 - yz) = 0, \\ xyz + 2\lambda(3a^2 - zx) = 0, \\ xyz + 2\lambda(3a^2 - xy) = 0. \end{cases}$$

比较以上三式可得

$$3a^2 - yz = 3a^2 - zx = 3a^2 - xy,$$

从而  $x = y = z = a$ . 又侧面积固定的长方体的最大体积客观存在, 因此侧面积固定的长方体中以正方体体积最大, 其最大值为  $a^3$ .

**例 1.3** 某单位拟建一排四间的停车房, 平面位置如图 1.1 所示. 由于资金及材料的限制, 围墙和隔墙的总长度不能超过 40 m, 为使车房面积最大, 应如何选择长、宽尺寸?

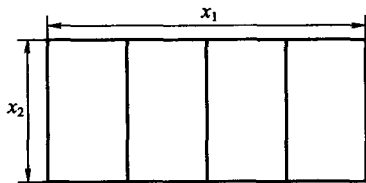


图 1.1

**解** 设四间停车房长为  $x_1$ , 宽为  $x_2$ . 由题意可知面积为

$$f(x_1, x_2) = x_1 x_2,$$

且变量  $x_1, x_2$  应满足

$$2x_1 + 5x_2 \leq 40,$$

$$x_1 \geq 0, x_2 \geq 0.$$

即求

$$\max f(x_1, x_2) = x_1 x_2,$$

$$\begin{cases} 2x_1 + 5x_2 \leq 40, \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

以上三个例子,虽然简单,但是它们代表了三种类型的最优化问题.

第一个例子是无约束极值问题,一般可表示为

$$\min f(x_1, x_2, \dots, x_n) \text{ 或 } \max f(x_1, x_2, \dots, x_n).$$

这里,  $f(x_1, x_2, \dots, x_n)$  是定义在  $\mathbf{R}^n$  上的可微函数.

求极值的方法是从如下含有  $n$  个未知数  $x_1, x_2, \dots, x_n$  的非线性方程组

$$\begin{cases} f'_{x_1}(x_1, x_2, \dots, x_n) = 0, \\ f'_{x_2}(x_1, x_2, \dots, x_n) = 0, \\ \dots\dots\dots \\ f'_{x_n}(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

中解出驻点,然后判定或验证这些驻点是不是极值点.

第二个例子是具有等式约束的极值问题,一般可表示为

$$\begin{cases} \min f(x_1, x_2, \dots, x_n) \text{ 或 } \max f(x_1, x_2, \dots, x_n), \\ h_j(x_1, x_2, \dots, x_n) = 0, j = 1, 2, 3, \dots, m (m < n). \end{cases}$$

该问题的求解通常采用拉格朗日乘数法,即把这个问题转化为求

$$L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) = f(x_1, x_2, \dots, x_n) - \sum_{j=1}^m \lambda_j h_j(x_1, x_2, \dots, x_n)$$

的无约束极值问题.

第三个例子是具有不等式约束的极值问题.

下面具体分析上述三种类型的最优化问题中按经典极值问题解法可能出现不能解决的情况:

(1) 当变量个数较多且方程组又是非线性的,此时方程组只有在相当特殊情况下才能人工解出.正因为如此,通常高等数学中的求极值问题的变量个数一般不超过三个.

(2) 当限制条件出现不等式,无论变量数多少,按经典极值方法求解几乎是不可能的.

直到 20 世纪 50 年代,最优化理论的建立以及电子计算机的迅速发展,才为求解各种最优化问题提供了雄厚的基础和有效手段.不过最优化方法作为一门崭新的应用学科,有关理论和方法还没有完善,许多问题还有待解决,目前正处于迅速发展之中.

最优化问题的数学模型包含三个要素:变量(又称设计变量)、目标函数、约束条件.

## 一、变量

一个优化设计方案是用一组设计参数的最优组合来表示的.这些设计参数

可概括地划分为两类：一类是可以根据客观规律、具体条件、已有数据等预先给定的参数，统称为**常量**；另一类是在优化过程中经过逐步调整，最后达到最优值的独立参数，称为**变量**。优化问题的目的就是使各变量达到最优组合。变量的个数称为优化问题的**维数**。例如有  $n$  个变量  $x_1, x_2, \dots, x_n$  的优化问题就是在  $n$  维空间  $\mathbf{R}^n$  中寻优。 $n$  维空间  $\mathbf{R}^n$  中的点  $X = [x_1, x_2, \dots, x_n]^T$  就代表优化问题中的一个方案。当变量为连续量时，称为**连续变量**；若变量只能离散取值，称为**离散变量**。

## 二、目标函数

反映变量间相互关系的数学表达式称为**目标函数**。其值的大小可以用来评价优化方案的好坏。按照规范化的形式，一般把优化问题归结为求目标函数的极小化问题，换句话说，目标函数值越小，优化方案越好。对于某些追求目标函数极大的问题，可以转化成求其负值最小的问题，即

$$\max f(X) = \max f(x_1, x_2, \dots, x_n) = -\min[-f(X)].$$

因此在本书的叙述中，一律把优化问题描述为目标函数的极小化问题，其一般形式为

$$\min f(X) = \min f(x_1, x_2, \dots, x_n).$$

如果优化问题只有一个目标函数，称为**单目标函数**，如果优化问题同时追求多个目标，则该问题的目标函数称为**多目标函数**。

多目标优化问题的目标函数通常表示为

$$V - \min F(X) = [f_1(X), f_2(X), \dots, f_m(X)]^T,$$

其中  $X = [x_1, x_2, \dots, x_n]^T$ 。这时的目标函数在目标空间中是一个  $m$  维矢量，所以又称之为**矢量优化问题**（一般用  $\min$  加一个前缀“ $V -$ ”来表示矢量极小化）。

## 三、约束条件

变量间本身应该遵循的限制条件的数学表达式称为**约束条件**或**约束函数**。

约束条件按其表达式可分为不等式约束和等式约束两种，即

$$\text{s. t. } \begin{cases} g_i(X) \geq 0, & i = 1, 2, \dots, l, \\ h_j(X) = 0, & j = 1, 2, \dots, m. \end{cases}$$

上式中“s. t.”为“subject to”的缩写，意即“满足于”或“受限于”。按约束条件的作用还可将约束条件划分为起作用的约束（**紧约束**、**有效约束**）和不起作用的约束（**松约束**、**消极约束**）。等式约束相当于空间里一条曲线（曲面或超曲面），点  $X$  必须为该曲线（曲面或超曲面）上的一点，因而总是紧约束。有一个独立的等式约束，就可用代入法消去一个变量，使优化问题降低一维。因此，数学模型中独立的

等式约束个数应小于变量个数;如果相等,就不是一个待定优化系统,而是一个没有优化余地的既定系统.不等式约束通常是以其边界  $g(X)=0$ (或  $g(X)\approx 0$ )表现出约束作用的,它只限制点  $X$  必须落在允许的区域(包括边界上),因而不等式约束的个数与变量个数无关.不带约束条件的优化问题称为无约束最优化问题;带约束条件的优化问题称为约束最优化问题.

#### 四、带约束条件的优化问题数学模型的表示形式

综上所述,全书所要讨论的问题是如下的(静态)最优化问题,其表示形式有三种:

第一种最优化问题表示形式为

$$\begin{aligned} & \min_{[x_1, x_2, \dots, x_n]^T \in \Omega} f(x_1, x_2, \dots, x_n), \\ & \text{s. t.} \begin{cases} g_i(x_1, x_2, \dots, x_n) \geq 0, i = 1, 2, \dots, l, \\ h_j(x_1, x_2, \dots, x_n) = 0, j = 1, 2, \dots, m (m < n). \end{cases} \end{aligned}$$

第二种最优化问题表示形式为

$$\begin{aligned} & \min_{X \in \Omega} f(X), \\ & \text{s. t.} \begin{cases} g_i(X) \geq 0, i = 1, 2, \dots, l, \\ h_j(X) = 0, j = 1, 2, \dots, m (m < n). \end{cases} \end{aligned}$$

第三种最优化问题表示形式为

$$\begin{aligned} & \min_{X \in \Omega} f(X), \\ & \text{s. t.} \begin{cases} G(X) \geq 0, \\ H(X) = 0, \end{cases} \end{aligned} \quad (1.1)$$

其中  $G(X) = [g_1(X), \dots, g_l(X)]^T, H(X) = [h_1(X), \dots, h_m(X)]^T$ .

上述三种表示形式中,  $X \in \Omega$  称为集约束.在所讨论的最优化问题中,集约束是无关紧要的.这是因为一般有  $\Omega \equiv \mathbf{R}^n$ , 不然的话,  $\Omega$  通常也可用不等式约束表达出来.因此今后一般不再考虑集约束.

满足所有约束的点  $X$  称为容许点或可行点.容许点的集合称为容许集或可行域,可用

$$D = \{X \mid g_i(X) \geq 0, i = 1, 2, \dots, l; h_j(X) = 0, j = 1, 2, \dots, m (m < n)\}$$

表示.

一般地,对于最优化问题(1.1)的求解,是指在可行域内找一点  $X^*$ ,使得目标函数  $f(X)$  在该点取得极小值,即

$$\begin{aligned} & f(X^*) = \min f(X), \\ & \text{s. t.} \begin{cases} G(X^*) \geq 0, \\ H(X^*) = 0. \end{cases} \end{aligned}$$

这样的  $X^*$  称为问题(1.1)的最优点,也称极小点,而相应的目标函数值  $f(X^*)$  称为最优值;合起来  $(X^*, f(X^*))$  称为最优解,但习惯上常把  $X^*$  本身称为最优解.最优点的各分量和最优值必须是有限数.

## § 1.2 最优化问题的算法

### 一、二维最优化问题的图解法

二维最优化问题具有鲜明的几何解释,这对于理解有关理论和掌握所述的方法是很有益处的.下面讨论的二维最优化问题为

$$\begin{aligned} & \min f(x_1, x_2), \\ & \text{s. t. } \begin{cases} g_i(x_1, x_2) \geq 0, & i = 1, 2, \dots, l, \\ h(x_1, x_2) = 0. \end{cases} \end{aligned}$$

#### (一) 约束集合

当约束函数为线性时,等式约束在坐标平面  $x_1Ox_2$  上为一条直线;不等式约束在坐标平面  $x_1Ox_2$  上为一半平面.当约束函数(例如  $h(x_1, x_2)$ )为非线性时,则等式约束条件  $h(x_1, x_2) = 0$  在坐标平面  $x_1Ox_2$  上为一条曲线(如图 1.2 所示).当约束函数(例如  $g_1(x_1, x_2)$ )为非线性时,则不等式约束  $g_1(x_1, x_2) \geq 0$  在坐标平面  $x_1Ox_2$  上为曲线  $g_1(x_1, x_2) = 0$  把坐标平面分成两部分当中的一部分(如图 1.3 所示).

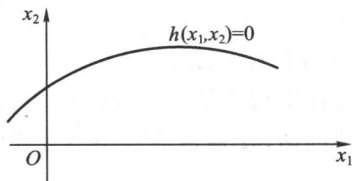


图 1.2

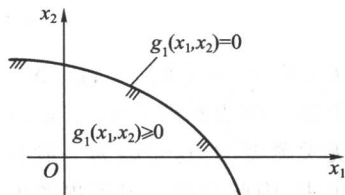


图 1.3

综上所述,当把约束条件中的每一个等式所确定的曲线,以及每一个不等式所确定的部分在坐标平面  $x_1Ox_2$  上画出之后,它们相交的公共部分即为约束集合  $D$ .

**例 1.4** 在  $x_1, x_2$  坐标平面上画出约束集合

$$D = \{[x_1, x_2]^T \mid x_1^2 + x_2^2 \leq 1, x_1 \geq 0, x_2 \geq 0\}.$$

**解** 满足  $x_1^2 + x_2^2 \leq 1$  的区域为以原点为圆心,半径为 1 的圆盘;满足  $x_1 \geq 0, x_2 \geq 0$  的区域为第一象限中的扇形(如图 1.4 所示).

## (二) 等高线

我们知道

$$t = f(x_1, x_2)$$

在三维空间中表示一张曲面.  $t=c$  (其中  $c$  为常数) 在三维空间中表示平行于  $x_1Ox_2$  平面的一个平面, 这个平面上任何一点的高度都等于常数  $c$  (如图 1.5 所示).

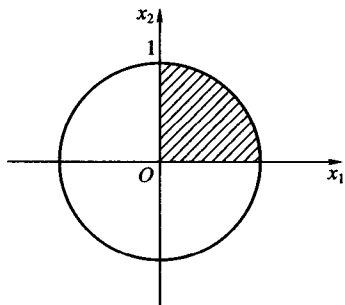


图 1.4

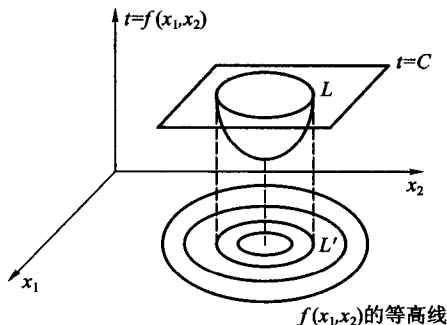


图 1.5

现在, 在三维空间中曲面  $t = f(x_1, x_2)$  与平面  $t = c$  有一条交线  $L$ . 交线  $L$  在  $x_1Ox_2$  平面上的投影曲线是  $L'$ , 可见曲线  $L'$  上的点  $[x_1, x_2]^T$  到平面  $t = c$  的高度都等于常数  $c$ , 也即曲线  $L'$  上的  $[x_1, x_2]^T$  的函数值  $f(x_1, x_2)$  都具有相同的值  $c$ . 当常数  $c$  取不同的值时, 重复上面的讨论, 在  $x_1Ox_2$  平面上得到一族曲线——等高线. 不难看出, 等高线的形状完全由曲面  $t = f(x_1, x_2)$  的形状所决定; 反之, 由等高线的形状也可以推测出曲面  $t = f(x_1, x_2)$  的形状. 在以后的讨论中, 不必具体画出曲面  $t = f(x_1, x_2)$  的图形, 只需在  $x_1Ox_2$  平面上变动常数  $c$  画出曲线簇  $f(x_1, x_2) = c$ .

**例 1.5** 在  $x_1Ox_2$  坐标平面上画出目标函数  $f(x_1, x_2) = x_1^2 + x_2^2$  的等高线.

**解** 因为当取  $c > 0$  时, 曲线  $x_1^2 + x_2^2 = c$  表示是以原点为圆心, 半径为  $\sqrt{c}$  的圆. 因此等高线是一簇以原点为圆心的同心圆 (如图 1.6 所示).

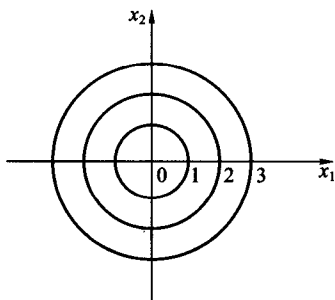


图 1.6

## (三) 几何意义及图解法

在  $x_1Ox_2$  坐标平面上分别画出约束集合  $D$  以及目标函数  $f(x_1, x_2)$  的等高线后, 不难求出二维最优化问题的最优解. 下面举例说明.

## 例 1.6 用图解法求解二维最优化问题

$$\min [(x_1+2)^2 + (x_2+2)^2],$$

$$\text{s.t.} \begin{cases} x_1^2 + x_2^2 \leq 1, \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

解 由例 1.4 得到约束集合  $D$  (如图 1.7 所示). 目标函数的等高线是以  $[-2, -2]^T$  为圆心的同心圆, 并且这簇同心圆的外圈比内圈的目标函数值大. 因此, 这一问题成为在约束集合中找一点  $[x_1, x_2]^T$ , 使其落在半径最小的那个同心圆上. 不难看出, 问题的最优解  $X^* = [x_1, x_2]^T = [0, 0]^T$ .

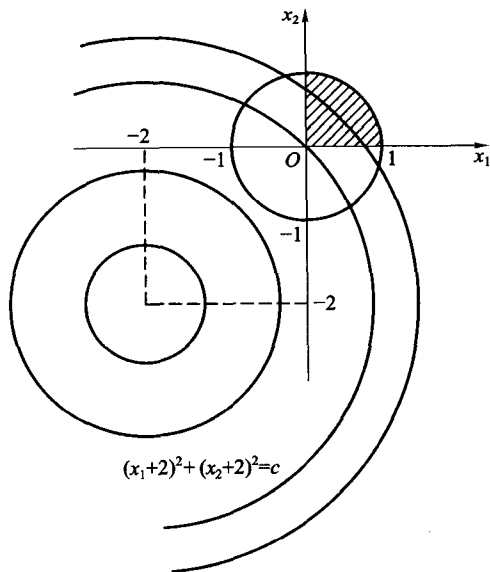


图 1.7

## 二、最优化问题的迭代解法

## (一) 迭代方法

在经典极值问题中, 解析法虽然具有概念简明、计算精确等优点, 但因只能适用于简单或特殊问题的寻优, 对于复杂的工程实际问题通常无能为力, 所以极少使用.

最优化问题的迭代算法是指: 从某一选定的初始点出发, 根据目标函数、约束函数在该点的某些信息, 确定本次迭代的一个搜索方向和适当的步长, 从而达到一个新点, 用式子表示即为

$$X_{k+1} = X_k + t_k P_k, \quad k = 0, 1, 2, \dots \quad (1.2)$$

式中  $X_k$  代表前一次已取得的迭代点, 在开始计算时为迭代初始点  $X_0$ ,  $X_{k+1}$  代表新的迭代点,  $P_k$  代表第  $k$  次迭代计算的搜索方向,  $t_k$  代表第  $k$  次迭代计算的步长因子.

按照式(1.2)进行一系列迭代计算所依据的思想是所谓的“爬山法”, 就是将寻求函数极小点(无约束或约束极小点)的过程比喻为向“山”的顶峰攀登的过程, 始终保持向“高”的方向前进, 直至达到“山顶”. 当然“山顶”可以理解为目标函数的极大值, 也可以理解为极小值, 前者称为上升算法, 后者称为下降算法. 这两种算法都有一个共同的特点, 就是每前进一步都应该使目标函数有所改善, 同时还要为下一步移动的搜索方向提供有用的信息. 如果是下降算法, 则序列迭代点的目标函数值必须满足下列关系

$$f(X_0) > f(X_1) > \dots > f(X_k) > f(X_{k+1}) > \dots$$

如果是求一个约束的极小点, 则每一次迭代的新点  $X_1, X_2, \dots$  都应该在约束可行域内, 即

$$X_k \in D, \quad k = 0, 1, 2, \dots$$

图 1.8 为迭代过程示意图.

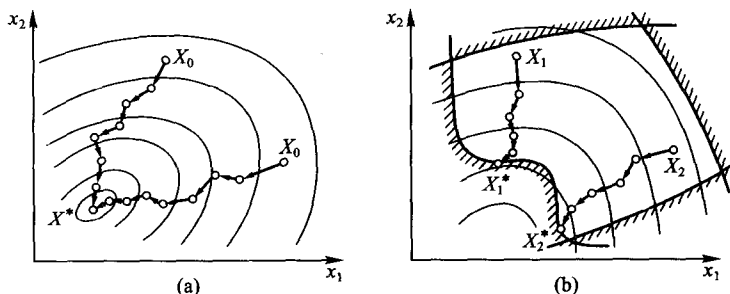


图 1.8

由上面的迭代过程可知, 在迭代过程中有两个规则需要确定: 一个是搜索方向  $P_k$  的选取; 一个是步长因子  $t_k$  的选取. 一旦  $P_k$  和  $t_k$  的选取方法确定, 则一种迭代算法就确定, 即不同的规则就对应不同的最优化方法.

## (二) 收敛速度与计算终止准则

### 1. 收敛速度

作为一个算法, 能够收敛于问题的最优解当然是必要的, 但仅收敛还不够, 还必须能以较快的速度收敛, 这才是好的算法.

**定义 1.1** 设由算法  $A$  产生的迭代点列  $\{X_k\}$  在某种“ $\|\cdot\|$ ”的意义下收敛



于点  $X^*$ , 即  $\lim_{k \rightarrow \infty} \|X_k - X^*\| = 0$ , 若存在实数  $\alpha > 0$  及一个与迭代次数  $k$  无关的常数  $q > 0$ , 使得

$$\lim_{k \rightarrow \infty} \frac{\|X_{k+1} - X^*\|}{\|X_k - X^*\|^\alpha} = q,$$

则称算法  $A$  产生的迭代点列  $\{X_k\}$  具有  $\alpha$  阶收敛速度, 或称算法  $A$  为  $\alpha$  阶收敛的, 特别地:

(1) 当  $\alpha = 1, q > 0$  时, 称迭代点列  $\{X_k\}$  具有线性收敛速度或称算法  $A$  为线性收敛的.

(2) 当  $1 < \alpha < 2, q > 0$ , 或  $\alpha = 1, q = 0$  时, 称迭代点列  $\{X_k\}$  具有超线性收敛速度或称算法  $A$  是超线性收敛.

(3) 当  $\alpha = 2$  时, 迭代点列  $\{X_k\}$  叫做具有二阶收敛速度或算法  $A$  是二阶收敛的.

一般认为, 具有超线性收敛或二阶收敛的算法是较快速的算法.

**例 1.7** 设一算法  $A$  产生迭代点列  $\left\{X_k = \frac{1}{k}\right\}$ , 它收敛于点  $X^* = 0$ , 试判定算法  $A$  的收敛速度.

解 因为

$$\lim_{k \rightarrow \infty} \frac{\left|\frac{1}{k+1} - 0\right|}{\left|\frac{1}{k} - 0\right|} = 1,$$

即取

$$\alpha = 1, q = 1 > 0.$$

所以算法  $A$  具有线性收敛速度.

**例 1.8** 设一算法  $A$  产生迭代点列  $\left\{X_k = \frac{1}{k^k}\right\}$ , 它收敛于  $X^* = 0$ , 试确定  $A$  的收敛速度.

解 因为

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\left|\frac{1}{(k+1)^{k+1}} - 0\right|}{\left|\frac{1}{k^k} - 0\right|} &= \lim_{k \rightarrow \infty} \frac{k^k}{(k+1)^{k+1}} \\ &= \lim_{k \rightarrow \infty} \left(\frac{k}{k+1}\right)^{k+1} \cdot \frac{1}{k} = 0, \end{aligned}$$

即取

$$\alpha = 1, q = 0.$$

所以  $A$  是超线性收敛的.

## 2. 计算终止准则

用迭代方法寻优时, 其迭代过程不能无限制地进行下去, 那么什么时候截断这种迭代呢? 这就是迭代什么时候终止的问题.