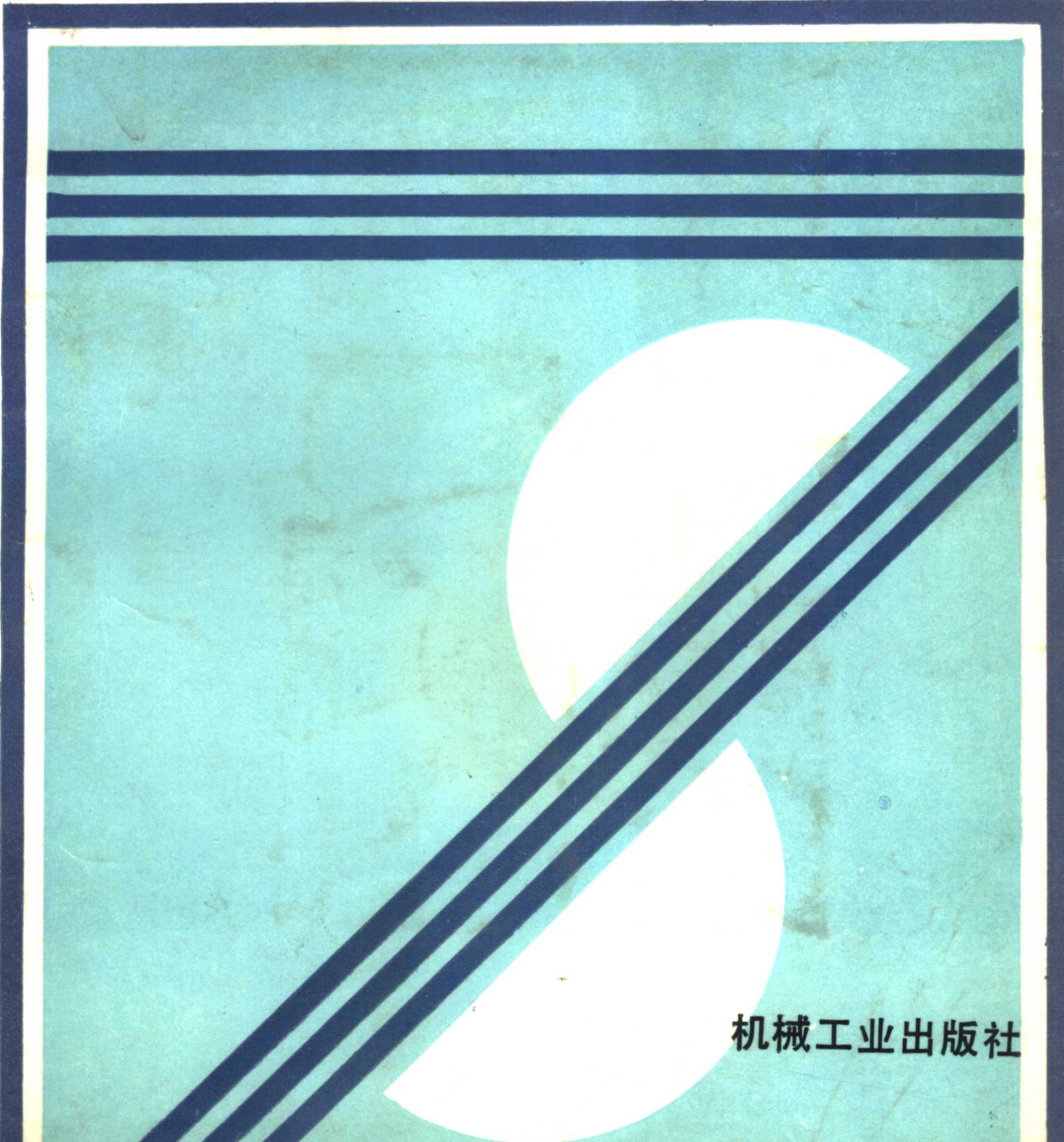


高等学校适用

机械优化设计

徐锦康 主编



机械工业出版社

高等学校适用

机械优化设计

徐锦康 主编

徐锦康 吕慧瑛 刘极峰 编



机械工业出版社

本书介绍机械优化设计的概念、理论、常用优化方法和机械优化设计实例。全书除绪论和附录外共八章,包括:机械优化设计概论、一维优化方法、无约束优化的直接搜索法、无约束优化的间接搜索法、约束优化的直接搜索法、约束优化的间接搜索法、机械优化设计实例和机械优化设计中需处理的几个问题。有关章节配有一定数量的习题,以便选用。附录给出常用优化方法源程序选例,供读者上机练习。

本书可作为高等工业学校机械类各专业大学生、研究生的教材或参考书,也可供工程技术人员参考。

机械优化设计

徐锦康 主编

*

责任编辑:钱飒飒 邵蔚林 版式设计:王颖

封面设计:姚毅 责任校对:陈立耘

责任印制:路琳

*

机械工业出版社出版(北京阜成门外百万庄南街一号)

邮政编码:100037

(北京市书刊出版业营业许可证出字第117号)

机械工业出版社印刷厂印刷

新华书店北京发行所发行·新华书店经售

*

开本 787×1092¹/₁₆·印张 11.25·字数 270千字

1996年10月第1版第1次印刷

印数 0 001-4 500 定价:12.00元

*

ISBN 7-111-05139-4/TH·687 (课)

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

前 言

机械优化设计是以电子计算机为工具、依据最优化的原理与方法、寻求机械最优设计参数的近代先进的设计方法之一，目前已发展为以机械、最优化理论与方法、算法语言三门学科为基础的一门综合性学科。在我国，随着科学技术的不断发展和社会主义市场竞争的日趋激烈，机械优化设计方法正日益为机械产品设计开发人员所重视与采用，并已经取得和正在取得可观的经济、技术效益和显著的社会效益。因此，作为一门课程，机械优化设计不仅是高等工业学校机械设计类各专业的一门必修课，而且是机械类其它各专业的一门重要的选修课。

我们在多年从事机械优化设计课程教学的基础上编写了此书，编写中参考了全国高等工业学校机械设计及制造专业教学指导委员会 1987 年 11 月制定的教学计划与教学大纲，同时兼顾了不同层次的教学要求和不同读者的要求。全书除绪论和附录外共八章：第一章介绍机械优化设计的基本概念及数学基础；第二~六章分别介绍一维优化方法、无约束优化的直接法与间接法、约束优化的直接法与间接法；第七、八两章介绍机械优化设计实例和机械优化设计中应处理的几个问题。有关章节配有适量习题，以便选用。附录给出常用优化方法源程序选例，供上机练习。本书可作为高等工业学校机械类各专业大学生、研究生的教材或参考书，也可供有关专业教师及工程技术人员阅读与参考。

本书编写中坚持“理论联系实际、便于工程应用”的原则，并注意反映国内近几年在优化方法、优化理论研究与应用方面的新发展、新成果，努力精选内容。在章节体系及内容的编排上，本书特别注意由浅入深，突出重点和强调物理概念与几何解释，同时又不失逻辑性与系统性。本书叙述上力求深入浅出、主次分明、层次清晰。书中采用我国法定计量标准规定的单位、名称及符号，优化方法中采用国际惯用的术语、符号及表示方法。

参加本书编写工作的有徐锦康、吕慧瑛、刘极峰。编写分工如下：绪论、第一章、第五章、第七章第三~七节及附录由徐锦康编写；第二、四、八章由吕慧瑛编写；第三章、第七章第一、二节由刘极峰编写；第六章由徐锦康、刘极峰共同编写；全书由徐锦康任主编。

本书承南京理工大学沈守范教授精心审阅全稿并提出很多宝贵的意见与建议，谨此表示衷心的感谢！编写中查阅和参考了众多文献资料，在此向参考文献的作者们致以诚挚的谢意！限于编者水平，书中缺点、错误在所难免，敬希读者随时批评指正。

编者

1995 年 5 月

目 录

前言		
绪论	1	
第一章 机械优化设计概论	3	
第一节 空心等截面管支柱的优化设计	3	
第二节 优化设计的数学模型	4	
第三节 数学模型某些性态的分析与讨论	6	
第四节 优化问题的几何解释	15	
第五节 优化方法的简单分类	16	
第六节 函数的无约束极值和有约束极值	17	
第七节 迭代方法及其收敛准则	22	
第八节 机械优化设计的一般步骤	26	
习题	27	
第二章 一维优化方法	29	
第一节 进退法确定初始搜索区间	29	
第二节 黄金分割法 (0.618 法)	31	
第三节 二次插值法	33	
第四节 有理插值法	37	
习题	40	
第三章 无约束优化的直接搜索法	41	
第一节 坐标轮换法	41	
第二节 模式搜索法	44	
第三节 鲍威尔 (Powell) 法	45	
习题	51	
第四章 无约束优化的间接搜索法	52	
第一节 梯度法	52	
第二节 牛顿法	54	
第三节 共轭梯度法	56	
第四节 变尺度法 (DFP 法和 BFGS 法)	59	
第五节 常用无约束优化方法的特点及应用	63	
习题	65	
第五章 约束优化的直接搜索法	66	
第一节 网格法	67	
第二节 正交网格法	69	
第三节 约束随机方向搜索法	73	
第四节 复合形法	77	
习题	82	
第六章 约束优化的间接搜索法	84	
第一节 惩罚函数法 (SUMT 法)	84	
第二节 增广乘子法 (PHR 法)	92	
第三节 广义简约梯度法和序列二次规划算法简介	100	
第四节 常用约束优化方法的特点及应用	103	
习题	104	
第七章 机械优化设计实例	106	
第一节 平面连杆机构的优化设计	106	
第二节 凸轮机构的优化设计	109	
第三节 普通 V 带传动的优化设计	112	
第四节 变位圆柱齿轮机构的优化设计	115	
第五节 单级直齿圆柱齿轮减速器的优化设计	121	
第六节 圆柱形螺旋压缩弹簧的优化设计	125	
第七节 机器人操作机大臂弹簧式自重平衡装置的优化设计	127	
第八章 机械优化设计中需处理的几个问题	132	
第一节 多目标优化	132	
第二节 整型变量和离散型变量	135	
第三节 数学模型的尺度变换	138	
第四节 数表与线图的程序化	140	
第五节 优化结果分析和全局最优解问题	142	
附录 常用优化方法源程序选例	143	
附录 A 进退法源程序	143	
附录 B 黄金分割法 (0.618 法) 源程序	144	

附录 C DFP 变尺度法源程序	146
附录 D 复合形法源程序	153

附录 E 混合惩罚函数法源程序	160
参考文献	173

绪 论

本部分将简要介绍优化、优化设计和机械优化设计的含义及发展概况，简述本课程的主要内容和目的。

一、优化、优化设计和机械优化设计的含义

优化是万物演化的自然选择和必然趋势。优化作为一种观念与意向，人类从很早开始就一直在自觉与不自觉地追求与探索。而优化作为一门学科与技术，则是一切科学与技术所追求的永恒的主题，旨在从处理各种事物的一切可能的方案中，寻求最优的方案。优化的原理与方法，在科学的、工程的和社会的实际问题中的应用，这便是优化设计。

优化设计是60年代初发展起来的一门新学科，它为工程设计提供了一种重要的科学设计方法，因而是构成和推进现代设计方法的产生与不断发展的重要内容。优化设计是最优化原理和计算技术相结合的产物，并借助于计算机性能的发展与完善而得到充分的发展和广泛的应用。而机械优化设计则是优化设计在一切工程设计应用中的一个方面，其本质是依据最优化的原理与方法，且通常借助计算技术与计算机这一强有力的手段，对某项机械设计，在规定的各种设计限制条件下，优选设计参数，使某项或某几项设计指标获得最优值。它所追求的是最优结果、最佳设计等等，所遵循的是一套日益发展与完善的最优化理论与方法。所以，整个设计过程是一个科学寻优的过程。而传统设计追求的设计结果一般不是最优的，因为它并不以“最优”作为其终极目标。传统设计中也希望设计尽可能好，也讲求设计质量，但主要凭主观经验和直观判断，所以结果一般不可能是最优的。当然，实际工程问题的优化设计方案，一般不等于函数的自然极值，因为实际问题往往受到许多条件的限制，因而所谓最优还只是相对最优，绝对最优往往只具有数学的与物理的意义。

二、优化设计和机械优化设计发展概况

正如前文所述，优化作为一种现象，是自然界万物演化的必然选择，而优化的简单意向，乃是人类在很早以前就有的。但是，科学地表达优化的概念，还是从17世纪时随着数学的发展而开始的。古典的优化方法，主要是应用微分法和变分法。直到本世纪40年代初，由于军事上的需要产生了运筹学，并使优化技术首先应用于解决战争中的实际问题，例如轰炸机最佳俯冲轨迹的设计等。50年代发展起来的数学规划理论，为优化设计奠定了理论基础。而60年代计算技术和计算机的发展，为优化设计的发展与应用提供了强有力的手段，使工程技术人员从大量繁琐的计算工作中解放出来，把主要精力转到优化模型的建立和优化方法的选择方面来，优化设计的应用从此得到飞速发展。近20多年来，优化设计方法已在工业、农业、交通、运输、能源、军事等各条战线上得到应用，并发挥着重要的作用。尤其在科技发达国家，优化的应用可以说是无处不在。

相对来说，优化方法在机械设计中的应用稍晚一些，直至60年代后期才开始有较成功的应用，但发展却十分迅速。20多年来，在机构综合、机械零部件设计、专用机械设计和工艺设计等方面，都获得应用并取得丰硕的成果。

机构运动参数的优化设计是机械优化设计中发展较早的领域，不仅研究了连杆机构、凸

轮机构等再现函数和轨迹的优化设计问题，而且还提出了一些标准程序。机构动力学优化设计方面也有很大进展，如惯性力的最优平衡，原动件力矩波动最小的优化设计等。机械零部件的优化设计最近十几年也有很大发展，主要是研究各种减速器的优化设计、液压轴承和滚动轴承的优化设计，以及轴、弹簧、制动器等结构参数优化。此外，在机床、锻压设备、压延设备、起重运输设备、汽车等的基本参数、基本工作机构及主体结构方面，也进行了优化设计工作。

优化设计的应用实例举不胜举，效果十分显著。例如美国 Bell 飞机公司对大型机翼用 138 个设计变量进行优化设计，使质量减少了三分之一；大型运输舰用 10 个变量进行优化设计，使造价降低约 10%。甚至运动员运动成绩的提高，也可借助优化技术。

我国对优化理论和方法的研究与应用虽然起步较晚，但经过高等学校、科研部门和工矿企业单位的多方努力，也取得了较快的进展和可观的经济效益。例如轨距 24m、起重量 3.5MN（质量约 550t）的桥式起重机的主梁，优化后的质量平均减少 14%；矿山用单级圆柱齿轮减速器，优化后质量减少 12%。由此可见，优化技术对我国四化建设有很大的作用。

值得一提的是，由华中理工大学、合肥工业大学等单位在“六五”计划期间研制成功的优化方法程序库 OPB 及其微机版本 PC-OPB，为工程技术人员提供了一个求解优化设计数学模型的强有力的工具。目前该程序库在国内已有成千上万的用户，正在广泛应用，并产生着巨大的经济、技术和社会效益。

三、本课程的主要内容和目的

本课程研究优化方法在机械设计中的应用，即机械优化设计的有关问题。

机械优化设计包括建立优化设计的数学模型和选择恰当的优化方法与程序两方面的内容，即建立数学模型和求解数学模型两方面的内容。建立数学模型除必须遵循一定的规范外，主要靠对机械学科的专业知识，所以不是本书所能解决的。求解数学模型一靠相应的优化方法，二靠计算技术、程序软件，并一般通过计算机实现。其中计算技术、优化程序编制及计算机的使用，也由专门学科解决。所以，本课程的主要内容概括地包括优化设计的基本概念、常用优化方法和典型机械优化设计实例三大部分。具体来讲，绪论及第一章介绍机械优化设计的基本概念；第二~六章分别介绍一维优化、无约束优化和约束优化的原理和算法；第七、八两章介绍典型机械优化设计实例和机械优化设计中应处理的几个问题；附录给出常用优化方法源程序选例（结构化 FORTRAN 77 语言），供读者上机练习。

希望通过对本课程的学习，了解优化设计的基本概念，掌握常用优化方法的原理、搜索步骤及应用特点，初步树立正确的优化设计观点，具备处理一般机械优化设计问题的能力。

第一章 机械优化设计概论

本章首先介绍空心等截面管支柱的优化设计,使读者对机械优化设计的基本概念和基本过程有一个初步的印象;接着介绍优化设计的数学模型、目标函数和约束函数的某些特性、优化设计的几何描述、优化方法分类、函数的无约束极值和有约束极值问题及优化搜索中的迭代方法和收敛准则;最后给出机械优化设计的一般步骤。

第一节 空心等截面管支柱的优化设计

如图 1-1 所示,一空心等截面管支柱两端承受轴向载荷 $F=22680\text{N}$,柱高 $h=2540\text{mm}$,材料为铝合金,弹性模量 $E=7.03\times 10^4\text{MPa}$,密度 $\rho=2768\text{kg/m}^3$,许用应力 $[\sigma]=140\text{MPa}$ 。要求管截面的平均直径 $D_m\leq 89\text{mm}$,壁厚 $\delta\geq 1\text{mm}$ 。试设计该管支柱,使其质量最小。

首先给出管支柱的质量表达式

$$m = \rho h \pi D_m \delta = 0.022 D_m \delta \quad (\text{a})$$

式中, m 的单位为 kg ; D_m 、 δ 的单位为 mm 。

根据题意,希望选择合适的管柱平均直径 D_m 及壁厚 δ ,使设计的管支柱既满足题目给出的有关限制条件,又具有最小的质量。显然, D_m 与 δ 是有待确定的变量,且 D_m 与 δ 均应为正值,可用图 1-2 表示 D_m 与 δ 间的设计关系。题目要求 $D_m\leq 89\text{mm}$ 、 $\delta\geq 1\text{mm}$,这可由图 1-2 中 $a-a$ 线以左、 $b-b$ 线以上区域来表示,它是参数 D_m 与 δ 选择时的边界限制区域。除此以外,管支柱还应满足强度与压杆稳定性要求。根据强度要求,应使管柱内的工作应力 $\sigma = F/(\pi D_m \delta)$ 小于许用应力 $[\sigma]$,即

$$\frac{F}{\pi D_m \delta} - [\sigma] \leq 0$$

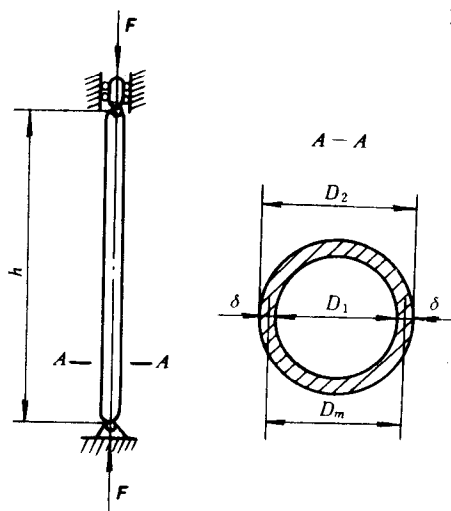


图 1-1 空心等截面管支柱

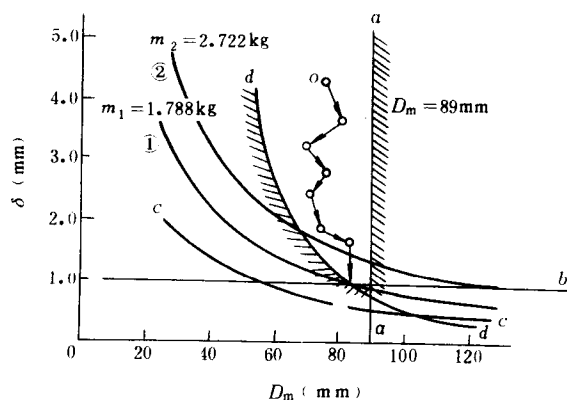


图 1-2 空心管支柱限制条件的几何关系

代入有关数值并整理为

$$51.6 - D_m \delta \leq 0$$

方程 $51.6 - D_m \delta = 0$ 的曲线为图 1-2 中 $c-c$ 曲线, 要使管柱满足强度要求, D_m 与 δ 应取 $c-c$ 曲线右上方的值。

根据压杆稳定性要求, 管柱内的工作应力 σ 应小于管柱的稳定临界应力 $\sigma_c = \frac{\pi^2 E}{8h^2} (D_m^2 + \delta^2) \approx \frac{\pi^2 E}{8h^2} D_m^2$ (假定壁厚 δ 远小于平均直径 D_m)。于是得

$$\frac{F}{\pi D_m \delta} - \frac{\pi^2 E D_m^2}{8h^2} \leq 0$$

代入有关数值并整理为

$$537000 - D_m^3 \delta \leq 0$$

根据这一不等式, D_m 与 δ 只可在图 1-2 中曲线 $d-d$ 的右上方取值。综合考虑以上尺寸限制条件、强度条件及压杆稳定性条件, D_m 与 δ 便只能在图 1-2 中阴影线所围区域内取值。为了在该区域内求得使管柱质量最小的设计点 $[D_m \ \delta]^T$, 可赋予管柱不同的质量, 例如 $m = 1.788\text{kg}, 2.722\text{kg}, \dots$, 则由质量表达式 (a), 可以在图 1-2 上画出相应的等质量曲线①、②等, 其中曲线①与压杆稳定的极限曲线 $d-d$ 和管柱壁厚的下限曲线 $b-b$ 交于 e 点。显然 e 点就是我们所要的设计点, 其坐标值可通过求 $b-b$ 曲线与 $d-d$ 曲线的交点得到: $D_m = 81.28\text{mm}$ 、 $\delta = 1\text{mm}$ 。所以最优的管柱设计方案为: $D_m = 81.28\text{mm}$ 、 $\delta = 1\text{mm}$ 、 $m = 1.788\text{kg}$ 。

就一般的工程实际问题而言, 最优设计方案的求解并不如此简单, 通常需要用优化设计方法并利用计算机作反复迭代搜索。即从一个不是很好的设计方案 (例图 1-2 中 o 点) 开始, 沿着使管柱质量下降的方向, 不断地进行迭代搜索, 直至找到管柱质量最小而又不违反所有限制条件的最优方案 (即 e 点或接近 e 点的点) 为止, 其搜索路线如图 1-2 所示。

通过上例, 我们对机械优化设计可以形成如下初步印象: ①优化设计是用于解决设计方案参数的最佳选择的方法; ②优化选择的最佳参数不仅使设计指标达到最优, 而且应满足各种限制条件; ③求解最优设计方案时一般要用优化方法并借助计算机。简单地说, 优化设计是在一些等式或不等式约束条件下, 求多变量函数的极大值或极小值的方法。

通过上例还可以看出, 处理优化设计问题时, 首先需要确定问题所包含的独立变量、尽可能用数学表达式表达优化意向即优化目标以及各种限制条件, 这就是我们在下一节中要讨论的优化设计的数学模型的概念, 这是优化设计中极其重要的基本概念。建立了优化问题的数学模型, 接着才是寻求适当的优化方法并借助计算机求解最优设计方案。此外我们还看到, 可以借助函数等值面 (线) (本例中又称等质量线) 的概念, 从几何上描述优化设计问题的最优设计方案及迭代求优过程。

第二节 优化设计的数学模型

用一组设计变量描述优化设计对象的设计内容、即描述优化意图和有关限制条件的数学表达式, 称为优化设计的数学模型。它是优化设计的基础及优化设计成败的关键。正确的数学模型加合适的优化方法, 才能获得满意的优化结果。

优化设计的数学模型包含三个要素, 即设计变量、目标函数和约束条件。

一、设计变量

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、具体条件或已有数据等预先给定的参数，称为设计常量，如材料的力学性能、机器的工作情况系数等；另一类是在优化过程中不断变化、最后使设计目标达到最优的独立参数，称为设计变量。优化设计的目的，就是寻找设计变量的一种组合，使某项或某几项设计指标最优。

设计变量的个数就是优化问题的维数。例如有 2 个设计变量的优化问题的维数为 2，有 3 个设计变量的优化问题的维数为 3。一般地说，有 n 个设计变量 x_1, x_2, \dots, x_n 的优化问题，其维数为 n 。由 n 个设计变量为坐标所组成的实空间称为设计空间。设计空间中，点 x 就代表一个设计方案（或称设计点），以向量表示时，可记为

$$\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T \text{ 或 } \mathbf{x} \in R^n$$

式中 R^n 代表 n 维实空间，又称 n 维欧氏空间。这里， \mathbf{x} 既表示从坐标系原点到以 n 个变量为坐标的点 x 的一个 n 维向量，又表示该向量的端点 x 。当 $n > 3$ 时， \mathbf{x} 即为一个超空间向量。

当变量为连续量时，称为连续变量，如杆件的长度、计算载荷、计算功率、转速等。当变量只能在离散量中取值时，称为离散变量，如齿轮的模数、齿数，滚动轴承的内径及 V 带的根数等。其中只能取整数的量，又称为整数变量，如 V 带根数，齿轮的齿数等。

二、目标函数

目标函数 $f(\mathbf{x})$ 是反映优化意向的关于设计变量的数学表达式，可用来直接评价优化方案的好坏，所以又称为评价函数。

为了规范化，优化设计数学模型中通常规定求目标函数的极小值，即

$$\min f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) \quad (1-1)$$

但实际机械优化设计中，还有追求目标值 $F(\mathbf{x})$ 的最大值的情况，如弹簧的减振能力最大，减速器的输出功率最大等。此时可处理为如下的极小化模式

$$\min f(\mathbf{x}) = -F(\mathbf{x}) \quad (1-2)$$

如果优化问题只有一个目标函数，称为单目标优化问题。例如减速器优化设计中，只要求体积最小，这就是单目标优化问题。若把目标函数值看作变量，单目标优化问题在目标空间中就只有一个变量，所以又称单目标优化问题为标量优化问题。如果优化问题有几个目标函数，则称为多目标优化问题，其综合优化目标函数通常表示为

$$f(\mathbf{x}) = [f_1(\mathbf{x}) \ f_2(\mathbf{x}) \ \dots \ f_m(\mathbf{x})]^T \quad \mathbf{x} \in R^n$$

这时的目标函数 $f(\mathbf{x})$ 在目标空间中已经是一个 m 维向量，所以又称多目标优化问题为向量优化问题。例如减速器优化设计中，如再要求输出功率最大，便是一个多目标优化问题。

本书主要讨论单目标优化问题。

三、约束条件

约束条件也称设计约束，它是设计变量间或设计变量本身应该遵循的限制条件的数学表达。约束条件按其表达方式可分为不等式约束和等式约束两种，即

$$g_j(\mathbf{x}) \leq 0 \quad (j=1, 2, \dots, m)$$

$$h_v(\mathbf{x}) = 0 \quad (v=1, 2, \dots, p)$$

按约束条件的性质，又可将约束分为性能约束和边界约束。性能约束是针对设计对象的某种性能或设计要求给出的限制，如零件的工作应力应小于或等于许用应力，轴的扭转变形

应小于许用扭转角等。边界约束（又称区域约束）是对某些变量的取值范围的限制，可能只给出上限或下限，也可能同时给出上、下限。例如 $a_i \leq x_i \leq b_i$ ($i=1, 2, \dots, n$)，写成小于等于即 $g_j(\mathbf{x}) \leq 0$ 的统一形式则为： $a_i - x_i \leq 0$ 、 $x_i - b_i \leq 0$ 。其中 a_i 、 b_i 为常量，所以也称边界约束为常量约束。

按约束条件的作用，还可将约束划分为起作用约束（又称紧约束、有效约束）和不起作用约束（又称松约束、无效约束）。若优化问题的极值点 \mathbf{x}^* 落在某个约束的边界上 ($g_k(\mathbf{x}^*) = 0$)，这个约束 ($g_k(\mathbf{x}^*) \leq 0$) 便称为起作用的约束。所以，等式约束一定是紧约束。有一个等式约束，就有可能用代入法消去一个设计变量，使优化问题降低一维。因此，数学模型中独立等式约束的数目应小于设计变量的个数，如果相等，就成了没有优化余地的既定系统。不等式约束通常是以其边界 $g_j(\mathbf{x}) = 0$ （或 $g_j(\mathbf{x}) \approx 0$ ）表现出约束作用的，所以 $g_j(\mathbf{x}) \leq 0$ 只限制设计变量必须落在允许的区域（包括边界上），因而不等式约束的数目与设计变量的个数无关。设计空间中，诸设计约束允许的设计点取值区域（包括边界）称为可行域，用字母 D 表示。可行域以外的区域称为非可行域，用 \bar{D} 表示。

带有约束条件的优化问题称为约束优化问题，又称为数学规划；不带约束条件时称为无约束优化问题。工程实际中的优化问题均属约束优化问题，因而设计点 \mathbf{x} 需落在诸约束围成的可行域 D 上，而 D 为包含于 n 维欧氏实空间的一个子集，即 $\mathbf{x} \in D \subset R^n$ 。

四、数学模型的基本表达方式

对无约束优化问题，数学模型的基本表达方式为

$$\min f(\mathbf{x}) \quad \mathbf{x} \in R^n \quad (1-3)$$

约束优化问题分等式约束优化、不等式约束优化和既有等式约束又有不等式约束的一般约束优化问题。一般约束优化问题数学模型的基本表达方式为

$$\left. \begin{array}{l} \min f(\mathbf{x}) \quad \mathbf{x} \in R^n \\ \text{S. t. } \quad g_j(\mathbf{x}) \leq 0 \quad (j = 1, 2, \dots, m) \\ \quad \quad h_v(\mathbf{x}) = 0 \quad (v = 1, 2, \dots, p) \end{array} \right\} \quad (1-4)$$

式中“S. t.”为“Subject to”的缩写，意即“满足于”或“受约束于”。当上式中 $j=0$ 时即为等式约束优化问题；当 $v=0$ 时即为不等式约束优化问题；当 $j=v=0$ 时，便退化为无约束优化问题（见式 (1-3)）。

利用某种优化方法求解式 (1-3) 或式 (1-4) 后，即可得到一组优化的设计变量，即一个优化设计方案 $\mathbf{x}^* = [x_1^* \quad x_2^* \quad \dots \quad x_n^*]^T$ 。称 \mathbf{x}^* 为优化问题的最优点；对应的目标函数值 $f^* = f(\mathbf{x}^*)$ 称为最优值； \mathbf{x}^* 与 f^* 合称为最优解。

建立实际工程优化问题的数学模型的工作，简称建模工作。不同的工程优化问题，其优化设计的数学模型的具体内容显然是各不相同的，但它们的数学模型均应写成式 (1-3) 或式 (1-4) 的形式。

第三节 数学模型某些性态的分析与讨论

一、二次函数、二次型及正定矩阵

二次函数、二次型及正定矩阵的概念，是最优化方法中十分有用的数学概念。

1. 二次函数、二次型及正定矩阵的概念

形式为

$$f(x) = \frac{1}{2}a_{11}x_1^2 + \frac{1}{2}a_{22}x_2^2 + \cdots + \frac{1}{2}a_{nn}x_n^2 + a_{12}x_1x_2 + \cdots + a_{n-1,n}x_{n-1}x_n + b_1x_1 + \cdots + b_nx_n + C \quad (1-5)$$

的函数称为二次函数，用矩阵表示为

$$f(x) = \frac{1}{2}x^T Ax + B^T x + C \quad (1-6)$$

式中

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

且 $a_{ij} = a_{ji}$ ($i, j = 1, 2, \dots, n$)，故 A 为对称矩阵。

若二次函数中只含有变量的二次项，则称 $f(x)$ 为二次齐次函数或二次型。此时

$$f(x) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{n-1,n}x_{n-1}x_n = \sum_{i,j=1}^n a_{ij}x_i x_j \quad (1-7)$$

用矩阵表示可简记为

$$f(x) = x^T Ax = [x_1 \ x_2 \ \cdots \ x_n] A [x_1 \ x_2 \ \cdots \ x_n]^T \quad (1-8)$$

如果对于任何 $x = [x_1 \ x_2 \ \cdots \ x_n]^T \neq 0$ ，恒有 $f(x) = x^T Ax > 0$ ，则称对称矩阵 A 是正定矩阵，并称此二次型为正定二次型；若 $f(x) = x^T Ax \geq 0$ ，则称 A 为半正定矩阵。

类似地，对于任意不为零的 $x = [x_1 \ x_2 \ \cdots \ x_n]^T$ ，若恒有 $f(x) = x^T Ax < 0$ ，则称 A 是负定矩阵；若对于有些 x 有 $f(x) = x^T Ax > 0$ ，而对于另一些 x 又有 $f(x) = x^T Ax < 0$ ，则称 A 为不定矩阵。

一般二次函数 $f(x) = \frac{1}{2}x^T Ax + B^T x + C$ ，总可以通过变量的线性变换转化成为二次型 $f(y) = y^T A' y$ ，若其中 A' 矩阵是正定的，则相应的二次函数称为正定二次函数。这一概念在优化方法的讨论中也常常用到。

2. 矩阵正定性的判别

若矩阵 A 正定，其必要与充分条件是矩阵行列式 $|A|$ 各阶主子式的值均大于零。若矩阵 A 负定，其充分必要条件是矩阵行列式 $|A|$ 各阶主子式的值负、正相间，也即 $-A$ 为正定矩阵。若矩阵行列式 $|A|$ 各阶主子式值的正、负不符合前面两种规律，称矩阵 A 为不定矩阵。

例 1-1 判别矩阵

$$A = \begin{bmatrix} 10 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 4 \end{bmatrix} \quad \text{的正定性。}$$

解

$$\because a_{11} = 10 > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 10 & 2 \\ 2 & 5 \end{vmatrix} = 46 > 0,$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 10 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 4 \end{vmatrix} = 101 > 0$$

各阶主子式的值均大于零，故矩阵 A 正定。

例 1-2 判别矩阵

$$A = \begin{bmatrix} -2 & 1 & 2 \\ -10 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix} \text{ 的正定性。}$$

解 $\because a_{11} = -2 < 0$, $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} -2 & 1 \\ -10 & 1 \end{vmatrix} = 8 > 0$,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} -2 & 1 & 2 \\ -10 & 1 & 3 \\ 1 & 3 & 2 \end{vmatrix} = -25 < 0$$

各阶主子式的值负、正相间，故矩阵 A 负定。

例 1-3 判定矩阵

$$A = \begin{bmatrix} 9 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 4 & -2 \end{bmatrix} \text{ 的正定性。}$$

解 $a_{11} = 9 > 0$, $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 9 & 2 \\ 2 & -1 \end{vmatrix} = -13 < 0$

至此已可断定 A 矩阵为不定矩阵。实际上，给定 $\mathbf{x}_1 = [1 \ 1 \ 1]^T$ 和 $\mathbf{x}_2 = [1 \ -1 \ 5]^T$ ，则 $\mathbf{x}_1^T A \mathbf{x}_1 = 24$ ， $\mathbf{x}_2^T A \mathbf{x}_2 = -56$ ，由定义亦知 A 矩阵确为不定矩阵。

二、目标函数的近似表达

在机械优化设计中，许多实际优化设计课题的目标函数往往是极为复杂的多元非线性函数。为简化对问题的描述，通常将目标函数在所讨论的点附近展开成泰勒级数，用一阶或二阶泰勒级数近似表达原函数。下面先介绍目标函数泰勒展开中要用到的函数的梯度及海赛矩阵的概念，再讨论函数的泰勒展开。

1. 梯度

设目标函数 $f(\mathbf{x})$ 是一阶连续可微的，则函数 $f(\mathbf{x})$ 在某点 $\mathbf{x}^{(k)}$ 处对 $x_i (i=1, 2, \dots, n)$ 的一阶偏导数的列矩阵（也称列向量），称为函数 $f(\mathbf{x})$ 在 $\mathbf{x}^{(k)}$ 点的梯度，记为 $\nabla f(\mathbf{x}^{(k)})$ ，则

$$\nabla f(\mathbf{x}^{(k)}) = \begin{bmatrix} \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_1} \\ \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_1} \quad \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_2} \quad \dots \quad \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_n} \right]^T \quad (1-9)$$

梯度 $\nabla f(\mathbf{x}^{(k)})$ 的模为：

$$\|\nabla f(\mathbf{x}^{(k)})\| = \sqrt{\left(\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_1}\right)^2 + \left(\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_2}\right)^2 + \cdots + \left(\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_n}\right)^2}$$

下面介绍几种常用的特殊类型函数的梯度表达式。

(1) 函数 $f(\mathbf{x}) = \mathbf{B}^T \mathbf{x}$ 的梯度

$$\because f(\mathbf{x}) = \mathbf{B}^T \mathbf{x} = b_1 x_1 + b_2 x_2 + \cdots + b_n x_n$$

式中 b_1, b_2, \dots, b_n 为列矩阵 \mathbf{B} 的分量, 均为常量。则

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i} (\mathbf{B}^T \mathbf{x}) = b_i \quad (i=1, 2, \dots, n)$$

$$\therefore \nabla f(\mathbf{x}^{(k)}) = \nabla (\mathbf{B}^T \mathbf{x})|_{\mathbf{x}^{(k)}} = \mathbf{B} = [b_1 \ b_2 \ \cdots \ b_n]^T \quad (1-10)$$

(2) 函数 $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ 的梯度

$$\because f(\mathbf{x}) = \mathbf{x}^T \mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$$

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i} (\mathbf{x}^T \mathbf{x}) = 2x_i \quad (i=1, 2, \dots, n)$$

$$\therefore \nabla f(\mathbf{x}^{(k)}) = \nabla (\mathbf{x}^T \mathbf{x})|_{\mathbf{x}^{(k)}} = 2\mathbf{x}^{(k)} = 2[x_1^{(k)} \ x_2^{(k)} \ \cdots \ x_n^{(k)}]^T \quad (1-11)$$

(3) 函数 $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ 的梯度

$$\because f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \sum_{i,j=1}^n a_{ij} x_i x_j$$

式中 \mathbf{A} 为实对称矩阵。

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = \frac{\partial}{\partial x_i} \left[\sum_{j=1}^n a_{ij} x_i x_j \right] = 2 \sum_{j=1}^n a_{ij} x_j \quad (i=1, 2, \dots, n)$$

$$\therefore \nabla f(\mathbf{x}^{(k)}) = \nabla (\mathbf{x}^T \mathbf{A} \mathbf{x})|_{\mathbf{x}^{(k)}} = 2 \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j^{(k)} \\ \sum_{j=1}^n a_{2j} x_j^{(k)} \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j^{(k)} \end{bmatrix} = 2 \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix}$$

$$= 2\mathbf{A}\mathbf{x}^{(k)}$$

(1-12)

(4) 一般二次函数 $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{B}^T \mathbf{x} + C$ 的梯度根据式 (1-10)、(1-12) 可得

$$\nabla f(\mathbf{x}^{(k)}) = \mathbf{A}\mathbf{x}^{(k)} + \mathbf{B} \quad (1-13)$$

2. 海赛 (Hessian) 矩阵

设函数 $f(\mathbf{x})$ 在点 $\mathbf{x}^{(k)}$ 处存在连续的一阶、二阶偏导数 $\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_i} (i=1, 2, \dots, n)$ 和 $\frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_i \partial x_j}$

($i, j=1, 2, \dots, n$), 则函数 $f(x)$ 在 $x^{(k)}$ 点的 n^2 个二阶偏导数 $\frac{\partial^2 f(x^{(k)})}{\partial x_i \partial x_j}$ ($i, j=1, 2, \dots, n$) 所构成的 $n \times n$ 阶方阵

$$H_n(x^{(k)}) = \begin{bmatrix} \frac{\partial^2 f(x^{(k)})}{\partial x_1^2} & \frac{\partial^2 f(x^{(k)})}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x^{(k)})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x^{(k)})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x^{(k)})}{\partial x_2^2} & \dots & \frac{\partial^2 f(x^{(k)})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x^{(k)})}{\partial x_n \partial x_1} & \frac{\partial^2 f(x^{(k)})}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x^{(k)})}{\partial x_n^2} \end{bmatrix} \quad (1-14)$$

称为海赛矩阵, 用符号 $H_n(x^{(k)})$ 或 $H(x^{(k)})$ 表示, 也可用 $\nabla^2 f(x^{(k)})$ 表示, 并可简记为

$$\nabla^2 f(x^{(k)}) = H(x^{(k)}) = \left[\frac{\partial^2 f(x^{(k)})}{\partial x_i \partial x_j} \right] \quad (i, j=1, 2, \dots, n) \quad (1-15)$$

由数学分析知, 如果 $f(x)$ 的一阶偏导数在定义域内处处可微, 则二阶偏导数与求导次序无关, 即

$$\frac{\partial^2 f(x^{(k)})}{\partial x_i \partial x_j} = \frac{\partial^2 f(x^{(k)})}{\partial x_j \partial x_i} \quad (i, j=1, 2, \dots, n, i \neq j) \quad (1-16)$$

所以, 海赛矩阵 $H(x^{(k)})$ 通常是一个 $n \times n$ 阶的对称方阵。

显然, 前面二次函数中的系数矩阵 A , 就是二次函数的海赛矩阵 (常量矩阵)。

函数梯度和海赛矩阵的概念及其有关表达式, 不仅在近似表达目标函数时有直接应用, 而且在整个优化理论中占有很重要的位置。例如在讨论极值条件、收敛准则及构造搜索方向等方面, 均用到梯度和海赛矩阵的概念。

3. 函数的泰勒 (Taylor) 展开

(1) 一元函数的泰勒展开式 若一元函数 $f(x)$ 在点 $x^{(k)}$ 的某个开区间 (a, b) 内具有直到 $n+1$ 阶导数, 则当 x 在 (a, b) 内时, $f(x)$ 的泰勒展开式为

$$f(x) = f(x^{(k)}) + \frac{f'(x^{(k)})}{1!} (x - x^{(k)}) + \frac{f''(x^{(k)})}{2!} (x - x^{(k)})^2 + \dots \\ + \frac{f^{(n)}(x^{(k)})}{n!} (x - x^{(k)})^n + R_n \quad (1-17)$$

实际应用中常忽略二阶以上的高阶微量, 只取前三项, 且令 $\Delta x = x - x^{(k)}$, 则

$$f(x) \approx f(x^{(k)}) + f'(x^{(k)}) \Delta x + \frac{1}{2} f''(x^{(k)}) \Delta x^2 \quad (1-18)$$

或
$$\Delta f(x) = f(x) - f(x^{(k)}) \approx f'(x^{(k)}) \Delta x + \frac{1}{2} f''(x^{(k)}) \Delta x^2 \quad (1-19)$$

(2) 二元函数的泰勒展开近似式 设二元函数 $f(x)$ 在点 $x^{(k)}$ 的某一邻域内有直到 $n+1$ 阶连续偏导数, 则 $f(x)$ 在点 $x^{(k)}$ 邻域内的泰勒展开近似式 (略去二阶以上的高阶微量) 为

$$f(x) \approx f(x^{(k)}) + f'_{x_1}(x^{(k)}) (x_1 - x_1^{(k)}) + f'_{x_2}(x^{(k)}) (x_2 - x_2^{(k)}) \\ + \frac{1}{2} [f''_{x_1^2}(x^{(k)}) (x_1 - x_1^{(k)})^2 + 2f''_{x_1 x_2}(x^{(k)}) (x_1 - x_1^{(k)}) (x_2 - x_2^{(k)}) \\ + f''_{x_2^2}(x^{(k)}) (x_2 - x_2^{(k)})^2] \quad (1-20)$$

用向量矩阵形式表示, 则上式为

$$f(x) \approx f(x^{(k)}) + \begin{bmatrix} \frac{\partial f(x^{(k)})}{\partial x_1} & \frac{\partial f(x^{(k)})}{\partial x_2} \end{bmatrix} \begin{bmatrix} x_1 - x_1^{(k)} \\ x_2 - x_2^{(k)} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x_1 - x_1^{(k)} & x_2 - x_2^{(k)} \end{bmatrix}$$

$$\begin{aligned} & \begin{bmatrix} \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_2^2} \end{bmatrix} \begin{bmatrix} x_1 - x_1^{(k)} \\ x_2 - x_2^{(k)} \end{bmatrix} \\ & = f(\mathbf{x}^{(k)}) + [\nabla f(\mathbf{x}^{(k)})]^T \Delta \mathbf{x} + \frac{1}{2} [\Delta \mathbf{x}]^T \nabla^2 f(\mathbf{x}^{(k)}) \Delta \mathbf{x} \end{aligned} \quad (1-21)$$

式中, $\mathbf{x} = [x_1 \ x_2]^T$; $\Delta \mathbf{x} = [x_1 - x_1^{(k)} \ x_2 - x_2^{(k)}]^T$ 。

(3) n 元函数 $f(\mathbf{x})$ 的泰勒展开近似式 仿照式(1-21), 可写出 n 元函数 $f(\mathbf{x})$ 在点 $\mathbf{x}^{(k)}$ 处的泰勒展开近似式

$$\begin{aligned} f(\mathbf{x}) & \approx f(\mathbf{x}^{(k)}) + [\nabla f(\mathbf{x}^{(k)})]^T \Delta \mathbf{x} + \frac{1}{2} [\Delta \mathbf{x}]^T \nabla^2 f(\mathbf{x}^{(k)}) \Delta \mathbf{x} \\ & = f(\mathbf{x}^{(k)}) + [\nabla f(\mathbf{x}^{(k)})]^T \Delta \mathbf{x} + \frac{1}{2} [\Delta \mathbf{x}]^T \mathbf{H}(\mathbf{x}^{(k)}) \Delta \mathbf{x} \end{aligned} \quad (1-22)$$

式中, $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$; $\Delta \mathbf{x} = [x_1 - x_1^{(k)} \ x_2 - x_2^{(k)} \ \cdots \ x_n - x_n^{(k)}]^T$ 。

三、目标函数的等值面(线)

设目标函数 $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$ 为一 n 元函数, 当给定一组设计变量 $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$, 就对应着一个确定的函数值 c , 即 $f(\mathbf{x}) = c$, 称具有这种性质的函数为可计算函数。反之, 给定 $f(\mathbf{x}) = c$, 将有无限多个设计点 \mathbf{x} 使上式成立, 这些设计点在 n 维设计空间中组成一个点集, 称此点集为等值曲面或等值超曲面, 统称等值面, 而在二维空间中此点集为等值线。当给定一系列函数值 c 时, 将在设计空间内得到一组等值面(线)簇。

图 1-3a 所示为一椭圆抛物面, 它是目标函数 $f(\mathbf{x}) = ax_1^2 + 2bx_1x_2 + cx_2^2$ ($a > 0, c > 0, ac - b^2 > 0$) 的图形(在 $n+1=2+1=3$ 维空间中), $f(\mathbf{x})$ 实际是一正定二次函数(见本节“一”)。将该椭圆抛物面上离 x_1, x_2 平面具有相同高度 c 的点投影到该平面上, 则得 $f(\mathbf{x}) = c$ 的等值线——平面椭圆曲线。显然, 给定一系列不同的 c 值, 可得到一组等值线簇——椭圆簇(图 1-3b), 该椭圆簇的中心即为坐标原点, 故又称这类曲线簇为有心曲线簇。分析图 1-3 可得到两点结论: 1) 等值线的分布规律反映了目标函数值的变化规律, 对有心曲线簇, 其中心就是目标函数的一个相对极小点; 2) 等值线的疏密反映了目标函数值变化率的大小, 等值线较密的部位, 目标函数值变化率较大。

二元二次函数的等值线簇还有双曲线簇和抛物线簇, 而二元线性函数的等值线簇为平行直线簇。这些曲线簇可称为无心曲线簇, 它们的相对极小点在无穷远处, 并且可能存在多个相对极小点。二元高次函数的等值线簇也分有心曲线簇和无心曲线簇, 但图形更复杂。图 1-4 为函数 $f(\mathbf{x}) = x_1^4 + x_1^2 - 2x_1^2x_2 - 2x_1x_2 + 4.5x_1 - 4x_2 + 2x_2^2 + 4$ 的等值线簇, 有两个中心, 即两个相对极小点。

当函数具有两个或两个以上的相对极小点时, 所有极小点中函数值最小的极小点称为全局极小点或全局最优点, 其他极小点则称为局部极小点或局部最优点。全局最优点的坐标及函数值, 称为全局最优解; 局部极小点的坐标及函数值, 则称为局部最优解。

如果一个严重非线性二次函数的等值线簇发生严重偏心和扭曲, 而且其分布也是疏密不一的, 情况严重时就成为所谓的“病态”函数, 这种情况将给优化设计带来极大的困难。

以上关于二维设计空间的等值线及其性质、特点的讨论, 完全可以推广到多维问题的分析中去。不过对于三维问题, 是设计空间中的等值面问题, 而高于三维的问题, 在设计空间