绪　　论

人们做任何事情都希望用最少的付出得到最佳的效果，工程设计人员总是力求取得工程问题的一组最合理的设计参数，使得由这组设计参数确定的设计方案既满足各种设计标准、设计规范和技术要求，又使其某一项或多项技术经济指标达到最佳，如结构最紧凑、用料最省、成本最低、工作性能最好等，这就是最优化设计。传统的工程设计，由于设计手段和设计方法的限制，设计者不可能在一次设计中得到某个项目的多个方案，不可能进行多方案的分析比较，更不可能寻求最优的设计方案。于是，人们只能在漫长的设计、实施和使用过程中，通过不断地认识、试验与改进，逐步使项目的设计方案趋于完善。现代电子计算机的发展和普及，以计算机为基础的最优化数值计算方法的成熟和应用，使工程问题的最优化设计成为可能，并在其应用和发展过程中形成了一套完整的工程最优化设计理论和方法。

工程最优化设计是把工程设计问题转化为与之对应的条件极值问题，然后利用最优化数值计算方法和计算机程序，借助计算机求得最优设计方案的过程和方法。进行工程最优化设计，首先必须将实际问题加以数学描述，形成一组代表该问题的数学表达式，称为设计问题的数学模型；然后选择一种最优化数值计算方法和计算机程序；最后在计算机上运算求解，得到一组代表最优设计方案的最佳的设计参数，称为设计问题的最优解。可见，最优化设计是一种先进的设计理念和方法，它同CAD设计、可靠性设计、动态设计、有限元分析等构成了一整套现代设计的理论和方法。在近四五十年内，这些现代设计理论和方法的推广和应用，使得各种工程设计的质量和速度得到了极大的改进和提高，从而在工程设计领域引起了一场巨大的变革。

20世纪50年代以前，用于解决最优化问题的数学方法仅限于古典的微分法和变分法。20世纪50年代末，由于计算机的出现，数学规划法即数值迭代法被用于求解工程最优化问题，并于其后的20～30年间，迅速地得到发展和应用，形成了一门新的应用数学的分支。

工程最优化问题可以分为函数最优化问题和组合最优化问题两大类。函数最优化问题就是通常所说的连续变量最优化问题，一般的工程设计问题都属于此类问题，用一般的最优化数值迭代方法即可求解。组合最优化是指一类离散变量和整数变量的最优化问题，这种问题的解是在一个有限或无限集合中，既满足各种设计要求，又使定义在该集合上的某个函数达到极值的子集。典型的组合最优化问题有旅行商（TSP）问题、加工调度问题、背包问题、装箱问题、着色问题和聚类问题等。

组合最优化问题有很强的工程代表性,从理论上讲可以用原始的穷举法求解,但该类问题的解集合的大小随着问题的复杂化会发生急剧地膨胀乃至“爆炸”。如已知 N 个城市中两两之间的距离,要求计算遍历每个城市一次的最短距离,这就是著名的 TSP 问题。这里存在$(N-1)!/2$条不同的路径,其计算量正比于$(N-1)!$,显然一般的穷举法或迭代搜索法都无法承受如此大的计算量。

随着人工智能学科的出现和发展,20 世纪 60 年代以后出现了一类模仿人类和生物繁衍、进化以及信息传播过程的最优化计算方法,如遗传算法、神经网络算法、蚁群算法等,简称智能最优化方法或进化方法。这类算法不仅能够求解一般的函数最优化问题,而且在解决全局最优解问题和组合最优化问题中具有独特的优势,因此在近 40 年来,得到了迅速的发展和应用。

本教材第 1 章介绍工程最优化设计的数学模型,第 2～6 章介绍各种最优化数值迭代方法,第 7 章介绍智能最优化方法,包括遗传算法和神经网络算法,第 8 章介绍 MATLAB 软件包中有关最优化计算的工具箱的使用,并列举了部分工程最优化设计的典型实例。

第1章

最优化问题的数学模型

数学模型是对实际问题的数学描述和概括，是进行最优化设计的基础。根据设计问题的具体要求和条件建立完备的数学模型是最优化设计成败的关键。这是因为最优化问题的计算求解完全是针对数学模型进行的。也就是说，最优化计算所得最优解实际上只是数学模型的解，至于是否是实际问题的解，则完全取决于数学模型与实际问题符合的程度。

工程设计问题通常是相当复杂的，欲建立便于求解的数学模型，必须对实际问题加以适当的抽象和简化。不同的简化方法得到不同的数学模型和计算结果，而且一个完善的数学模型，往往需要在计算求解过程中进行反复地修改和补充才能最后得到。由此可见，建立数学模型是一项重要而复杂的工作：一方面希望建立一个尽可能完善的数学模型，以求精确地表达实际问题，得到满意的设计结果；另一方面又要力求建立的数学模型尽可能简单，以方便计算求解。要想正确地协调这两方面的要求，就必须对实际问题及其相关设计理论和设计知识有深入的理解，并且善于将一个复杂的设计问题分解为多个子问题，抓住主要矛盾逐个加以解决。

本章通过几个简单的最优化设计简例，说明数学模型的一般形式、结构及其有关的基本概念。

1.1 设计简例

下面是3个最优化设计简例，可以看作几个复杂工程设计问题的子问题，虽然比较简单，但却具有一定的代表性。

例 1-1 用一块边长3m的正方形薄板，在四角各裁去一个大小相同的方块，做成一个无盖的箱子。试确定如何裁剪可以使做成的箱子具有最大的容积。

解：设裁去的4个小方块的边长为x，则做成的箱子的容积为

$$f(x)=x(3-2x)^2$$

于是，上述问题可描述为

求变量　x

使函数　$f(x)=x(3-2x)^2$　极大化

这样就把该设计问题转化成为一个求函数 $f(x)$ 最大值的数学问题。其中,x 是待定的求解参数,称为设计变量;函数 $f(x)$ 代表设计目标,称为目标函数。由于目标函数是设计变量的三次函数,并且不存在任何限制条件,故称此类问题为非线性无约束最优化问题。

根据一元函数的极值条件,令 $f'(x)=0$,解得 $x=0.5$,$f(x)=2.0$,记作 $x^*=0.5$,$f^*(x)=2.0$,称为原设计问题的最优解。

例 1-2 某工厂生产甲、乙两种产品,生产每种产品所需的材料、工时、用电量和可以获得的利润,以及每天能够提供的材料、工时、用电量见表 1-1,试确定该厂两种产品每天的生产计划,以使得每天获得的利润最大。

表 1-1 生产条件基本数据

产品	材料/kg	工时/h	用电能量/(kW·h)	利润/元
甲	9	3	4	60
乙	4	10	5	120
供应量	360	300	200	

解:这是一个简单的生产计划问题,可归结为在满足各项生产条件的基础上,合理安排两种产品每天的生产量,以使利润最大化的最优化设计问题。

设每天生产甲产品 x_1 件,乙产品 x_2 件,每天获得的利润用函数 $f(x_1,x_2)$ 表示,即

$$f(x_1,x_2)=60x_1+120x_2$$

每天实际消耗的材料、工时和电力分别用函数 $g_1(x_1,x_2)$、$g_2(x_1,x_2)$ 和 $g_3(x_1,x_2)$ 表示,即

$$g_1(x_1,x_2)=9x_1+4x_2$$
$$g_2(x_1,x_2)=3x_1+10x_2$$
$$g_3(x_1,x_2)=4x_1+5x_2$$

于是,该生产计划问题可归结为以下数学问题

求变量 x_1,x_2

使函数 $f(x_1,x_2)=60x_1+120x_2$ 极大化

并满足条件

$$g_1(x_1,x_2)=9x_1+4x_2\leqslant 360$$
$$g_2(x_1,x_2)=3x_1+10x_2\leqslant 300$$
$$g_3(x_1,x_2)=4x_1+5x_2\leqslant 200$$
$$g_4(x_1,x_2)=x_1\geqslant 0$$
$$g_5(x_1,x_2)=x_2\geqslant 0$$

这些表达式是对该生产计划问题的数学描述,称为数学模型。其中,函数 $f(x_1,x_2)$ 代表设计目标,称为目标函数。x_1,x_2 是待求解参数,称为设计变量。$g_u(x_1,x_2)(u=1,2,\cdots,5)$ 代表 5 个已知的生产指标,称为约束函数;对应的 5 个不等式代表 5 个生产条件,称为约束条件。由于目标函数和所有约束函数都是设计变量的线性函数,故称此类问题为线性约束最优化问题。此问题虽然比较简单,但却无法用高等数学中的极值条件直接求解。

例 1-3 一种承受纯扭矩的空心传动轴,已知需传递的转矩为 T,见图 1-1,试设计确定此传动轴的尺寸,以使其用料最省。

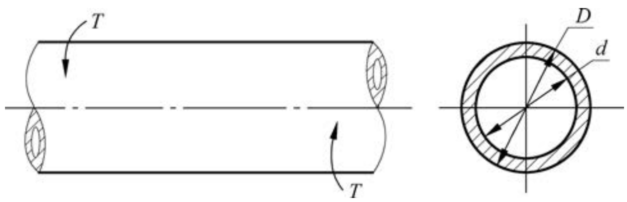

图 1-1 空心传动轴简图

解：由机械设计理论知，传动轴是只承受纯扭矩载荷的轴，一般采用空心圆截面轴。当传动轴的长度一定时，这种轴的体积和重量与轴的截面积成正比。为了承受一定的扭矩而又不发生失效，要求传动轴必须具备一定的强度和刚度。令轴的外径和内径分别用 D 和 d 表示，由设计资料知，空心传动轴的截面积、强度条件和刚度条件分别为

$$S=\frac{\pi}{4}(D^2-d^2)$$

$$\tau=\frac{16DT}{\pi(D^4-d^4)}\leqslant[\tau]$$

$$\theta=\frac{32T}{\pi G(D^4-d^4)}\leqslant[\theta]$$

式中，τ 为轴截面上的最大扭剪应力，$[\tau]$为轴用材料的许用扭剪应力；θ 和$[\theta]$为轴的扭转角和许用扭转角；G 为剪切弹性模量。

用 x_1 表示外径 D，用 x_2 表示内径 d，则上述传动轴设计问题可转化为如下数学模型所代表的最优化设计问题：

求设计变量 x_1,x_2

使目标函数 $f(x_1,x_2)=\frac{\pi}{4}(x_1^2-x_2^2)$ 极小化

满足约束条件

$$g_1(x_1,x_2)=\frac{16T}{\pi}\cdot\frac{x_1}{x_1^4-x_2^4}-[\tau]\leqslant 0$$

$$g_2(x_1,x_2)=\frac{32T}{\pi G}\cdot\frac{1}{x_1^4-x_2^4}-[\theta]\leqslant 0$$

$$g_3(x_1,x_2)=x_1\geqslant 0$$

$$g_4(x_1,x_2)=x_2\geqslant 0$$

显然，这是一个含有 4 个约束条件、两个设计变量的非线性约束最优化设计问题，同样无法直接用极值条件求解。

1.2 数学模型的一般形式

从以上 3 个例子可以看出，最优化设计的数学模型是对实际问题的数学描述，由设计变量、目标函数和约束条件 3 部分组成。可以概括为如下的一般形式：

求设计变量 $x_1,x_2,\cdots,x_n$

极小化目标函数 $f(x_1,x_2,\cdots,x_n)$

满足约束条件

$$g_u(x_1,x_2,\cdots,x_n)\leqslant 0 \quad (u=1,2,\cdots,p)$$
$$h_v(x_1,x_2,\cdots,x_n)=0 \quad (v=1,2,\cdots,m)$$

其中,$g_u(x_1,x_2,\cdots,x_n)\leqslant 0$ 称为不等式约束条件,简称不等式约束,$h_v(x_1,x_2,\cdots,x_n)=0$ 称为等式约束条件,简称等式约束。p 和 m 分别表示两种约束条件的个数。

用向量 $\boldsymbol{X}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$ 表示 n 个设计变量,用 $\boldsymbol{X}\in\mathbf{R}^n$ 表示向量 $\boldsymbol{X}$ 属于 n 维实欧氏空间,用 min 和 max 分别表示极小化和极大化,用 s. t. (subject to)表示“满足于”,则最优化设计的数学模型可简写为如下向量形式:

$$\begin{aligned} &\min f(\boldsymbol{X}) \quad \boldsymbol{X}\in\mathbf{R}^n \\ &\text{s. t. } g_u(\boldsymbol{X})\leqslant 0 \quad (u=1,2,\cdots,p) \\ &\qquad h_v(\boldsymbol{X})=0 \quad (v=1,2,\cdots,m) \end{aligned} \tag{1-1}$$

由于工程设计中所要求的解都是实数解,故式(1-1)中的 $\boldsymbol{X}\in\mathbf{R}^n$ 可以省略。

式(1-1)是数学模型的一般形式,本书后面所推导出的算法和相关公式,都是以此一般形式为基础给出的。当实际问题与此形式不一致时,应首先将其转化为一般形式。如设计问题要求目标函数 $f(\boldsymbol{X})$ 极大化时,只要将目标函数以 $-f(\boldsymbol{X})$ 代替即可,因为对同一个问题,$\min f(\boldsymbol{X})$ 和 $\max[-f(\boldsymbol{X})]$ 具有相同的解。同样,当约束条件中的不等号为不小于($\geqslant$)时,只需将不等式两端同乘以“-1”即可。

例 1-2 的数学模型经上述转化成为如下一般形式:

$$\begin{aligned} &\min f(\boldsymbol{X})=-60x_1-120x_2 \\ &\text{s. t. } 9x_1+4x_2-360\leqslant 0 \\ &\qquad 3x_1+10x_2-300\leqslant 0 \\ &\qquad 4x_1+5x_2-200\leqslant 0 \\ &\qquad -x_1\leqslant 0 \\ &\qquad -x_2\leqslant 0 \end{aligned}$$

最优化问题也称数学规划问题。根据数学模型中目标函数和约束函数的性质可将最优化问题分为线性最优化(规划)问题和非线性最优化(规划)问题。

当数学模型中的目标函数和约束函数全部是设计变量的线性函数时,称此问题为线性最优化问题或线性规划问题;当目标函数和约束函数中至少有一个是非线性函数时,称这样的问题为非线性最优化问题或非线性规划问题。

线性规划和非线性规划是数学规划的两个重要分支,生产计划和经济管理方面的问题一般可归结为线性规划问题,工程设计问题可归结为非线性规划问题。

1.3 数学模型的组成

1.3.1 设计变量与设计空间

在最优化问题的数学模型中,设计变量是一组待定的未知数,它对应于实际工程问题的

一组特征主参数，它的任意一组确定的数值代表该工程问题的一个特定的设计方案。因此，在建立工程问题的数学模型时，应该首先选取那些能够代表设计方案的主参数作为设计变量。

工程问题的设计参数一般是相当多的，包括常量和变量，变量又分独立变量和因变量。建立数学模型时，为了使数学模型尽量简单并且易于求解，通常只选取独立变量作为设计变量。如例1-3的空心传动轴设计中有3个设计参数：内径 d、外径 D 和壁厚 δ，其中只有两个参数是独立的。当选内径 d、外径 D 为设计变量时，壁厚可表示为 $\delta=0.5(D-d)$，当选外径 D、壁厚 δ 为设计变量时，内径可表示为 $d=D-2\delta$。

同一设计问题，当设计要求或设计条件发生变化时，设计变量的确定也应随之变化。如前述空心传动轴设计，若将传动轴改为转轴，则轴上不仅受扭矩作用，而且受弯矩作用，这时轴的长度也成为决定轴的强度和刚度的参数，故也应选作设计变量。

对于比较复杂的问题，可以先把那些较次要的参数或者变化范围较窄的参数暂时作为常量，建立简化的数学模型，以减少设计变量的数目，加快最优化求解的速度。当确定这种简化的模型计算无误时，再逐渐增加设计变量的个数，逐步提高求解的准确性与完整性。

设计变量有连续变量和离散变量之分。工程问题中很多变量要求取整数或标准系列值，这就是离散变量问题。通常所说的最优化理论和算法都是对连续变量问题提出的，对于离散变量最优化问题，目前直接的求解算法还不够成熟，通常的处理方法是先将离散变量当作连续变量，用连续变量最优化算法求出连续最优解后，再作适当的离散化处理，如某种方式的圆整或取标准值等。

由线性代数知，分别以 n 个设计变量 $x_1, x_2, \cdots, x_n$ 为坐标轴，可以形成一个 n 维实欧氏空间，记作 $\mathbf{R}^n$。称这样的空间为设计空间，称 n 为空间的维数，称空间中的点为设计点。于是每一个设计点都对应设计变量的一组确定的值，都代表设计问题的一个确定的解。可见，设计空间就是最优化问题的解空间。最优化问题的目的就是要在设计空间内无穷多个设计点中，找到一个既满足所有约束条件，又使目标函数取得极小值的点，称为最优点，它所代表的解称为设计问题的最优解。

记设计空间中的一个设计点为 $\boldsymbol{X}=[x_1, x_2, \cdots, x_n]^{\mathrm{T}}$，其中 $x_1, x_2, \cdots, x_n$ 代表各个坐标方向上的坐标值。则 $\boldsymbol{X}$ 同时也代表一个以坐标原点为起点，以 $\boldsymbol{X}$ 为终点的向量(矢量)。这样的几个设计点间可以进行向量运算。如图1-2和图1-3所示，两个设计点 $\boldsymbol{X}^1$ 和 $\boldsymbol{X}^2$ 的连线构成的第三个向量，可以用 $\boldsymbol{X}^1-\boldsymbol{X}^2$ 表示。

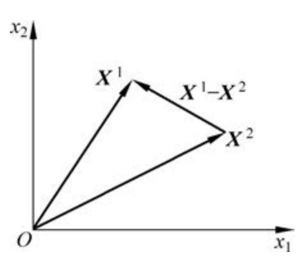

图1-2 二维设计空间

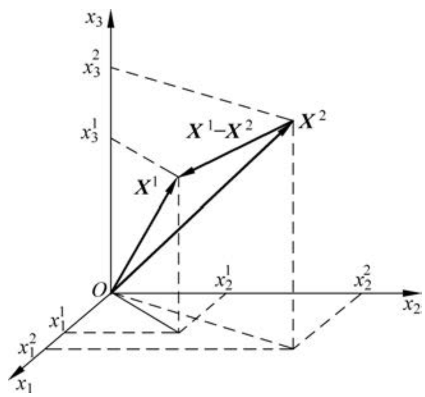

图1-3 三维设计空间

1.3.2 约束条件与可行域

任何设计问题都附带大量的设计要求和限制条件,将这样的要求和限制表示成设计变量 $\boldsymbol{X}$ 的函数 $h_v(\boldsymbol{X})$和 $g_u(\boldsymbol{X})$,进而构成如下的数学不等式或等式:

$$g_u(\boldsymbol{X}) \leqslant 0 \quad (u=1,2,\cdots,p)$$

$$h_v(\boldsymbol{X}) = 0 \quad (v=1,2,\cdots,m)$$

则这样的一组表达式就称为该设计问题的约束条件。

约束条件除有等式约束和不等式约束之分外,还可分为边界约束和性能约束,起作用约束和不起作用约束等。

边界约束是对设计变量本身所加的直接限制,如下面的约束

$$a_i - x_i \leqslant 0$$

$$x_i - b_i \leqslant 0$$

就限定了设计变量 x_i 的取值范围为闭区间$[a_i, b_i]$,因此属于边界约束。

性能约束从形式上看,是对设计问题的某一项技术性能指标或性能参数所加的限制,但实际上仍是对设计变量所加的间接限制。如例 1-3 中关于材料、用电量和工时的约束条件都属于性能约束条件。

将不等式约束中的不等号改成等号后得到的方程

$$g_i(\boldsymbol{X}) = 0$$

称为约束方程,对应的图形称为约束边界。一个约束边界把设计空间一分为二,其中一部分区域内的所有点均满足原不等式约束,而另一部分区域内的点都不满足原不等式约束。

等式约束本身也是一种约束方程和约束边界,不过此时只有约束边界上的点满足该等式约束,边界之外的任何点都不满足该等式约束。如图 1-4 所示。

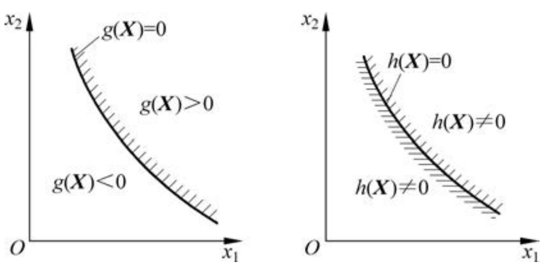

图 1-4 约束边界

可见,每一个不等式约束和等式约束都将设计空间分为满足约束和不满足约束的两个区域。对于一个最优化问题,满足所有约束条件的部分一般是由多个约束边界所围成的一个封闭区域。这个区域内的每一个点都同时满足所有的约束条件,称这种区域为最优化问题的约束可行域,记作 $\mathcal{D}$。可行域也可以看作满足所有约束条件的设计点的集合,于是也可用集合的方式表示如下:

$$\mathcal{D} = \{\boldsymbol{X} \mid g_u(\boldsymbol{X}) \leqslant 0, h_v(\boldsymbol{X}) = 0 \ (u=1,2,\cdots,p;\ v=1,2,\cdots,m)\} \tag{1-2}$$

由例 1-2 得到的 5 个约束方程分别是

$$g_1(\boldsymbol{X})=9x_1+4x_2-360=0$$
$$g_2(\boldsymbol{X})=3x_1+10x_2-300=0$$
$$g_3(\boldsymbol{X})=4x_1+5x_2-200=0$$
$$g_4(\boldsymbol{X})=x_1=0$$
$$g_5(\boldsymbol{X})=x_2=0$$

它们在二维设计平面中形成的约束边界和约束可行域如图 1-5 所示。可以看出这个问题的约束可行域是由 5 条约束边界(直线)围成的封闭五边形 $OABCD$。

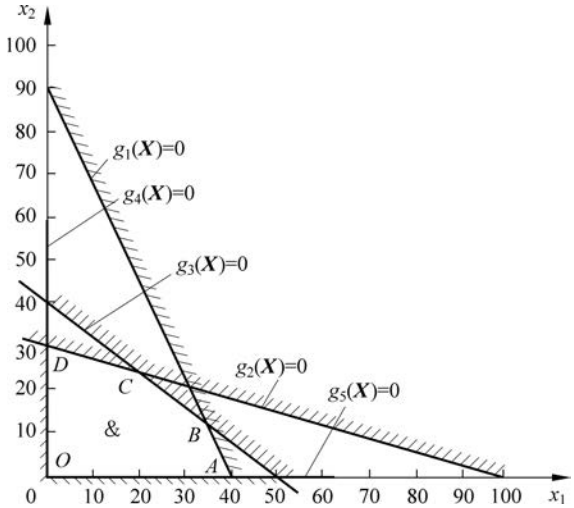

图 1-5　例 1-2 的可行域

再看下面的一组约束条件

$$g_1(\boldsymbol{X})=-x_1+x_2-2\leqslant 0$$
$$g_2(\boldsymbol{X})=x_1^2-x_2+1\leqslant 0$$
$$g_3(\boldsymbol{X})=-x_1\leqslant 0$$

它们的 3 条约束边界线所围成的可行域如图 1-6 所示。

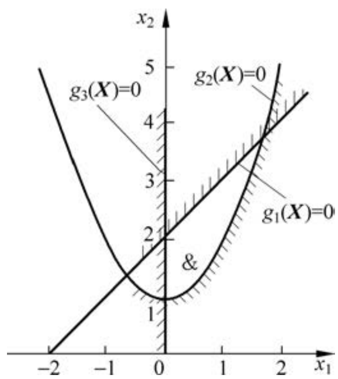

图 1-6　约束可行域

根据是否满足约束条件可以把设计点分成可行点(也称内点)和非可行点(也称外点),根据约束边界是否通过某个设计点,又可将约束条件分成该设计点的起作用约束和不起作用约束。

所谓起作用约束就是对某个设计点特别敏感的约束,或者说约束边界正好通过该设计点的约束。

具体地说,如果在点 $\boldsymbol{X}^k$ 上,某个不等式约束条件 $g_i(\boldsymbol{X})\leqslant 0$ 变成等式,即有

$$g_i(\boldsymbol{X}^k)=0$$

也就是说,该点位于这个约束边界上时,则称 $g_i(\boldsymbol{X})\leqslant 0$ 是点 $\boldsymbol{X}^k$ 的起作用约束。在图 1-7 中,点 $\boldsymbol{X}^1$ 位于约束边界 $g_1(\boldsymbol{X}^1)=0$ 上,故 $g_1(\boldsymbol{X})\leqslant 0$ 是点

$\boldsymbol{X}^1$ 的起作用约束,其余 3 个约束条件则是点 $\boldsymbol{X}^1$ 的不起作用约束。又如点 $\boldsymbol{X}^2$,它位于两个约束边界的交点上,故这两个约束条件 $g_1(\boldsymbol{X})\leqslant 0$ 和 $g_2(\boldsymbol{X})\leqslant 0$ 都是点 $\boldsymbol{X}^2$ 的起作用约束。

一个点 $\boldsymbol{X}^k$ 的起作用约束的个数和对应的约束条件序号可以用集合形式表示如下

$$I_k=\{u \mid g_u(\boldsymbol{X}^k)=0\ (u=1,2,\cdots,p)\} \tag{1-3}$$

式中,I_k 称为点 $\boldsymbol{X}^k$ 的起作用约束的下标集合。如图 1-7 中,$I_1=\{1\}$,$I_2=\{1,2\}$。

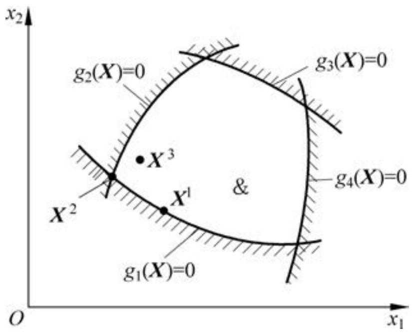

图 1-7 起作用约束

1.3.3 目标函数与等值线

要寻求某一问题的最优解,首先要有评判问题好坏的标准。在最优化设计的数学模型中,目标函数就是衡量设计方案优劣的定量标准。对于极小化问题,目标函数的值越小对应的设计方案越好。

不同的设计问题有不同的方案评价标准,甚至一个问题有几个不同的评价标准。一般情况下,应针对具体问题,选择设计问题的某个重要的技术经济指标作为最优化设计的目标函数,如利润、成本、功率、重量等。

通常一个设计问题只有一个目标函数,这就是单目标最优化问题,它是本书讨论的重点。

求解最优化问题的目的就是要找出最优解所代表的最优设计方案。对于一个具体的设计问题,是否有最优解?有几个最优解?最优解在什么位置?这些问题都决定于目标函数和约束函数的性态及其变化规律。对于简单的问题,函数的等值线(面)可以直观地描绘函数的变化趋势,成为判断和确定最优解的重要依据。

令函数 $f(\boldsymbol{X})$ 等于常数 c,即

$$f(\boldsymbol{X})=c \tag{1-4}$$

则满足式(1-4)的点 $\boldsymbol{X}$ 在设计空间定义了一个点集。当 $c=2$ 时,该点集是设计平面内的一条直线或曲线;当 $c\geqslant 3$ 时,该点集是设计空间内的一个平面、曲面或超曲面。在这样的一条线或一个面上,所有点的函数值均相等,因此,称这种线或面为函数的等值线或等值面。

当 c 取一系列不同的常数值时,式(1-4)定义了一组形态相似的等值线(面),称为函数 $f(\boldsymbol{X})$ 的等值线(面)族。如图 1-8 和图 1-9 所示。