薛毅　编著

# 最优化原理与方法

ZuiYouHua YuanLi Yu FangFa

北京工业大学出版社

# 最优化原理与方法

薛 毅 编著

北京工业大学出版社

**图书在版编目（CIP）数据**

最优化原理与方法/薛毅编著. —北京：北京工业大学出版社，2001. 2

ISBN 7-5639-0961-3

Ⅰ. 最… Ⅱ. 薛… Ⅲ. ①最佳化—数学理论②最佳化—数学方法 Ⅳ. 0224

中国版本图书馆 CIP 数据核字（2001）第 03796 号

**最优化原理与方法**

薛 毅 编著

*

北京工业大学出版社出版发行

邮编：100022 电话：(010) 67392308

各地新华书店经销

徐水宏远印刷厂印刷

*

2001 年 2 月第 1 版　2001 年 2 月第 1 次印刷

850mm×1168mm 32 开本 11.75 印张 280 千字

印数：1～2000 册

ISBN7-5639-0961-3/T·172

定价：20.00 元

# 序 言

最优化，是指从问题的许多可能的解答中，选择依某种指标最好的解答，它是数学的一个重要分支．在50年代后，随着计算机的发展和生产的需要，逐渐形成了最优化理论，以及相应的求解方法——最优化方法．目前这些方法已在生产管理、工程设计、系统控制、经济规划等领域内得到了广泛的应用．

本书是为应用数学系本科生、工科硕士研究生所写的有关最优化知识的一本教材．作为教材，本书的基本观点是：采用简单、基本直观的方法，向学生介绍最优化的有关理论、基本原理和相应的算法，并试图让学生了解算法的来龙去脉，以便使他们在解决实际问题的过程中，更好地运用这些方法．

本书的基础是“数学分析”和“线性代数”，即学生只需具备“数学分析”和“线性代数”知识就可读懂全部内容．对于工科学生，只需具备“高等数学”和“线性代数”知识就可读懂大部分内容．

本书共有十一章．第一章最优化方法，作为引言，向读者介绍最优化的基本概念和本书欲解决的问题．为了便于读者更好地学习本书的知识，特增加了两节数学预备知识．第二章线性规划与第三章线性规划的对偶问题主要涉及线性规划的基本内容．第四章无约束最优化问题的一般结构，第五章一维搜索，第六章使用导数的最优化方法和第七章直接方法，主要讨论无约束最优化问题的求解方法．第八章约束问题的最优性条件，第九章二次规划，第十章可行方向法和第十一章乘子法，主要讨论约束最优化问题的求解方法．作为教材，在每章的后面均列有习题，便于学

生复习和巩固该章所学的知识.

本书在编排上采用分块式，即本书分为三大块，线性规划、无约束最优化问题和约束最优化问题. 各块相对独立，教师可根据情况作一定的删减，可作为不同类型的教材. 如选用无约束最优化问题和约束最优化问题部分可作为“非线性规划”教材，选用线性规划、无约束最优化问题和约束最优化问题（或一部分）可作为“最优化方法”教材. 去掉较难的定理证明和较繁的算法可作为工科学生“最优化方法”方面的教材.

具体授课方法可作如下安排：

1. 作为应用数学系本科生“最优化计算方法”课程的教材，讲授全部内容，大约80学时.

2. 作为应用数学系本科生“非线性规划”课程的教材，可去掉本书的第二章、第三章，大约教授60学时.

3. 作为工科硕士生“最优化计算方法”课程的教材，可去掉较难的理论证明和第九章二次规划的内容，大约讲授60学时.

4. 作为工科本科生“最优化计算方法”课程的教材，可去掉较难的理论证明以及第八章约束问题的最优性条件和第九章二次规划及第四章无约束最优化问题的一般结构的部分内容，大约讲授40学时.

当然，任课教师可根据授课内容和对象，对教材进行适当的删减或增加.

本书除作为教材外，也可供从事优化方面的科技工作者和工程技术人员学习和参考.

在编写本书之前，编者曾为我校应用数学系本科生的“非线性规则”课程和工科研究生的“最优化方法”课程编写过讲义，本书是在此基础上，通过多次的教学实践修改而成的. 在编写过程中，得到了邓乃扬教授、诸梅芳教授的很大帮助，陈志教授、杨中华副教授在使用过程中曾提出过许多宝贵修改意见，在此向

他们表示衷心的感谢.

由于受编者水平限制，可能在内容的取材、结构的编排以及课程的讲法上存在着不妥之处，我们希望使用本书的老师、同学以及同行专家和其他读者提出宝贵的批评和建议.

在本书出版之际，谨向为本书出版予以大力支持和帮助的北工大教材建设委员会和北工大出版社表示衷心的感谢.

编　者

# 目　录

# 第一章　绪　　论

## §1.1　引　　言

**最优化**（Optimization），就是在复杂环境中遇到的许多可能的决策中，挑选“最好”的决策的科学．最优化方法——即求最优决策的方法．在本世纪30年代末，由于军事和工业生产发展的需要，提出了一些不能用古典微分法和变分法解决的问题．在许多学者和广大科技工作者的共同努力下，逐渐产生、发展和形成了一些新的方法．

前面提到，最优化就是选择“最好”决策．那么什么是最好呢？在不同的意义下，有不同的标准，因此它包括的范围很广，我们无法在短时间内全面而又系统地进行介绍．这里我们首先介绍一个名词——**数学规划**（Mathematical Programming）．数学规划是最优化理论的一个重要分支．数学规划，是指对 $n$ 个变量对单目标（或多目标）函数求极小（或极大）．而这些变量也可能受到某些条件（等式方程或不等式方程）的限制．其数学表达式为：

$$\begin{aligned} \min \quad & f(\boldsymbol{x}), \quad \boldsymbol{x}\in R^n, \\ \text{s.t.} \quad & c_i(\boldsymbol{x})=0, i\in E=\{1,2,\cdots,l\}, \\ & c_i(\boldsymbol{x})\leqslant 0, i\in I=\{l+1,l+2,\cdots,l+m\}, \end{aligned} \tag{1.1.1}$$

这里 min 表示求极小，s. t. ＝subject to 意思是受限制于…，$\boldsymbol{x}$ 是 $n$ 维向量，其分量为 $x_1$，$x_2$，…，$x_n$．在问题（1.1.1）中称 $f(\boldsymbol{x})$ 为**目标函数**（objective function），称 $c_i(\boldsymbol{x})=0,(i\in E)c_i(\boldsymbol{x})\leqslant 0(i\in I)$ 为**约束条件**（constraint condition）．若求极大，可

以将目标函数写成 min（$-f(\boldsymbol{x})$），若不等式约束 $c_i(\boldsymbol{x})\geqslant 0$ 可以写成 $-c_i(\boldsymbol{x})\leqslant 0$.

在问题（1.1.1）中，若 $f(\boldsymbol{x})$、$c_i(\boldsymbol{x})(i\in E\cup I)$ 均是线性函数，则得到的规划称为**线性规划**（Linear Programming），其一般表达式为

$$\begin{aligned}
\min\quad & f(\boldsymbol{x})=c^T\boldsymbol{x},\boldsymbol{x}\in R^n,\\
\text{s.t.}\quad & c_i(\boldsymbol{x})=a_i^T\boldsymbol{x}-b_i=0, i\in E=\{1,2,\cdots,l\}\\
& c_i(\boldsymbol{x})=a_i^T\boldsymbol{x}-b_i\leqslant 0, i\in I=\{l+1,l+2,\cdots,l+m\}.
\end{aligned}\tag{1.1.2}$$

在问题（1.1.1）中，$f(\boldsymbol{x})$、$c_i(\boldsymbol{x})(i\in E\cup I)$ 中之一是非线性函数，则称为**非线性规划**（Nonlinear Programming）. 若去掉问题（1.1.1）的约束条件，则得到

$$\min\quad f(\boldsymbol{x}),\ \boldsymbol{x}\in R^n,\tag{1.1.3}$$

称为**无约束最优化问题**（unconstraint optimmization problem）. 因此，也称问题（1.1.1）为**约束最优化问题**（constraint optimization problm）.

## §1.2 最优化问题

为进一步说明最优化问题，我们先举例说明.

### 1.2.1 无约束最优化问题

**【例 1.2.1】** 曲线拟合问题.

设有两个物理量 $\xi$ 和 $\eta$，根据某一物理定律得知它们满足如下关系：

$$\eta=a+b\xi^c,$$

其中 $a$、$b$、$c$ 三个常数在不同情况下取不同的值. 现由实验得到一给数据

$$(\xi_1,\ \eta_1),(\xi_2,\ \eta_2),\cdots,(\xi_m,\ \eta_m),$$

试选择 $a$、$b$、$c$ 的值，使曲线 $\eta=a+b\xi^c$ 尽可能靠近所有的实验点（$\xi_i$，$\eta_i$），$i=1$，2，…，$m$.

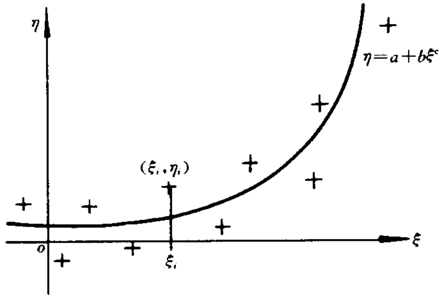



图 1.2.1　曲线拟合问题

这个问题可用最小二乘原理求解，即选择 $a$、$b$、$c$ 的一组值，使得偏差的平方和

$$\delta(a,b,c)=\sum_{i=1}^{m}(a+b\xi_i^c-\eta_i)^2 \tag{1.2.1}$$

达到最小．换句话说，就是求三个变量的函数 $\delta$（$a$，$b$，$c$）的极小点作为问题的解.

为了便于今后的讨论，我们把它写成统一的形式，把 $a$、$b$、$c$ 换成 $x_1$、$x_2$、$x_3$，记为 $\boldsymbol{x}=(x_1,\ x_2,\ x_3)^T$，把 $\delta$ 换成 $f$，这样（1.2.1）式改写为

$$f(\boldsymbol{x})=f(x_1,x_2,x_3)=\sum_{i=1}^{m}(x_1+x_2\xi_i^{x_3}-\eta_i)^2, \tag{1.2.2}$$

问题归纳为

$$\min\quad f(\boldsymbol{x}),\ \boldsymbol{x}\in R^3,$$

即无约束最优化问题.

无约束最优化问题的一般形式为

$$\min\quad f(\boldsymbol{x}),\quad \boldsymbol{x}\in R^n. \tag{1.2.3}$$

那么什么是它的最优解呢？我们给出严格的定义.

**定义 1.2.1** 若存在 $\boldsymbol{x}^* \in R^n$，使得对任意的 $\boldsymbol{x} \in R^n$，均有

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*), \tag{1.2.4}$$

则称 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的**全局极小点**（global minimum point）或称 $\boldsymbol{x}^*$ 是无约束问题**全局解**（global solution）. 若对任意的 $\boldsymbol{x} \in R^n$，$\boldsymbol{x} \neq \boldsymbol{x}^*$，均有

$$f(\boldsymbol{x}) > f(\boldsymbol{x}^*), \tag{1.2.5}$$

则称 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的严格全局极小点，或称 $\boldsymbol{x}^*$ 是无约束问题严格全局解.

由此可知，所谓无约束最优化问题的**最优解**（optimal solution），就是满足（1.2.4）式或（1.2.5）式的 $\boldsymbol{x}^*$. •

### 1.2.2 约束最优化问题

**【例 1.2.2】** 生产安排问题.

某工厂生产 $A$、$B$、$C$ 三种产品，每件产品所消耗的材料、工时以及盈利见表 1.2.1.

**表 1.2.1**

| 产 品 | $A$ | $B$ | $C$ |
|---|---|---|---|
| 材料（千克/件） | 4 | 4 | 5 |
| 工时（小时/件） | 4 | 2 | 3 |
| 盈利（元/件） | 7 | 3 | 6 |

已知该工厂每天的材料消耗不超过 600 千克，工时不得超过 1400 小时，问每天生产 $A$、$B$、$C$ 三种产品各多少可使得盈利最大？

**解** 设每天生产 $A$、$B$、$C$ 三种产品分别是 $x_1$、$x_2$、$x_3$ 件，因此共盈利

$$7x_1 + 3x_2 + 6x_3,$$

其相应的材料限制为

$$4x_1+4x_2+5x_3\leqslant 600,$$

工时限制为

$$4x_1+2x_2+3x_3\leqslant 1400,$$

再考虑自然限制

$$x_1\geqslant 0,\ x_2\geqslant 0,\ x_3\geqslant 0,$$

因此生产安排问题就是在上述限制条件下，使得其盈利达到最大，其数学表达式为

$$\begin{aligned}
\min\quad & f(\boldsymbol{x})=-7x_1-3x_2-6x_3,\\
\text{s.t.}\quad & c_1(\boldsymbol{x})=4x_1+4x_2+5x_3-600\leqslant 0,\\
& c_2(\boldsymbol{x})=4x_1+2x_2+3x_3-1400\leqslant 0,\\
& c_3(\boldsymbol{x})=-x_1\leqslant 0,\\
& c_4(\boldsymbol{x})=-x_2\leqslant 0,\\
& c_4(\boldsymbol{x})=-x_3\leqslant 0.
\end{aligned}$$

**【例 1.2.3】** 人字架最优设计问题.

考虑如图 1.2.2 所示的钢管构造的人字架，设钢管壁厚 $t=\bar{t}$ 和半跨度 $s=\bar{s}$ 已给定，试求能承受负荷 $2P$ 的最轻设计.

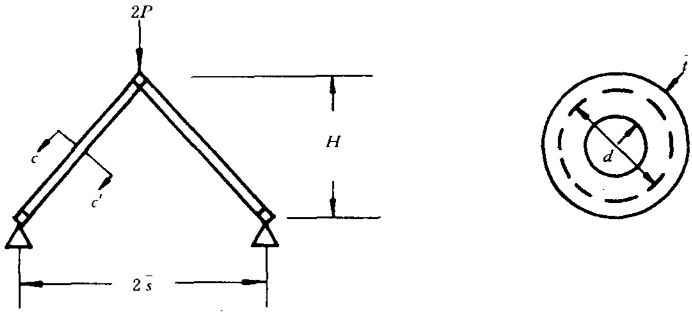
首先我们定性地分析此问题. 壁厚和跨度一定，欲求最轻设计，需要杆短. 这样做必使张角增大，则负荷 $2P$ 就会在钢管上有很大的张力. 为了能承受这样的应力，钢管需变粗，其结果是杆变重.

下面进行定量的分析. 给定一组 $d$ 和 $H$ 值后，可以计算出钢管的截面积 $A$ 和钢管的长度 $L$，即

$$A=\frac{1}{4}\pi(D_2^2-D_1^2)=\frac{\pi}{4}(D_2+D_1)(D_2-D_1)=\pi d\bar{t},$$

$$L=(\bar{s}^2+H^2)^{\frac{1}{2}},$$

因此钢管的重量为




人字架结构　　　　c—c′截面图

图 1.2.2　人字架最优设计问题

$$W(d,H)=2\rho\pi d\,\bar{t}(\bar{s}^2+H^2)^{\frac{1}{2}},$$

其中 $\rho$ 是比重.

下面考虑 $d$、$H$ 受到的限制，当负荷为 $2P$ 时，杆件受到的压力为

$$\sigma(d,H)=\frac{P}{\pi\bar{t}}\cdot\frac{(\bar{s}^2+H^2)^{\frac{1}{2}}}{Hd}.$$

根据结构力学原理，对于选定的钢管，不出现屈服的条件是

$$\sigma(d,H)\leqslant\sigma_y,$$

其中 $\sigma_y$ 是钢管最大许可的抗压强度. 而不出现弹性变曲（屈曲条件）为

$$\sigma(d,H)\leqslant\frac{\pi^2E(d^2+\bar{t}^2)}{8(\bar{s}^2+H^2)},$$

其中 $E$ 为钢管材料的扬氏模量. 因此人字架最优设计问题是在上述两个条件下使 $W$（$d$，$H$）达到最小.

为了方便起见，写成统一的数学表达式，分别用 $x_1$、$x_2$ 代替 $d$、$H$，记 $\boldsymbol{x}=(x_1,\ x_2)^T$，用 $f$ 代替 $W$，因此得到

$$f(\boldsymbol{x}) = 2\pi\rho\,\bar{t}x_1(\bar{s}^2 + x_2^2)^{\frac{1}{2}},$$

$$c_1(\boldsymbol{x}) = \frac{P}{\pi\,\bar{t}} \cdot \frac{(\bar{s}^2 + x_2^2)^{\frac{1}{2}}}{x_1x_2} - \sigma_y,$$

$$c_2(\boldsymbol{x}) = \frac{P}{\pi\,\bar{t}} \cdot \frac{(\bar{s}^2 + x_2^2)^{\frac{1}{2}}}{x_1x_2} - \frac{\pi^2E}{8} \cdot \frac{\bar{t}^2 + x_1^2}{\bar{s}^2 + x_2^2}.$$

人字架最优设计问题的数学表达式为

$$\begin{aligned} &\min \quad f(\boldsymbol{x}), \quad \boldsymbol{x} \in R^n, \\ &\text{s.t.} \quad c_i(\boldsymbol{x}) \leqslant 0, \ i = 1,2. \end{aligned}$$

综上所述，约束最优化问题的一般形式为

$$\begin{aligned} &\min \quad f(\boldsymbol{x}), \qquad \boldsymbol{x} \in R^n, \\ &\text{s.t.} \quad c_i(\boldsymbol{x}) = 0, i \in E = \{1,2,\cdots,l\}, \\ &\qquad\ \ c_i(\boldsymbol{x}) \leqslant 0, i \in I = \{l+1, l+2, \cdots, l+m\}, \end{aligned} \tag{1.2.6}$$

称 $f(\boldsymbol{x})$ 为目标函数，称 $c_i(\boldsymbol{x}) = 0$ $(i \in E)$ 或 $c_i(\boldsymbol{x}) \leqslant 0$ $(i \in I)$ 为约束条件.

在例 1.2.2 中，所有函数均是线性的，因此是线性规划问题，而例 1.2.3 则属于非线性规划问题.

**定义 1.2.2** 称满足约束条件的点为**可行点**（feasible point），称可行点全体组成的集合为**可行域**（feasible region），记作 $D$，即

$$D = \{\boldsymbol{x} \mid c_i(\boldsymbol{x}) \leqslant 0, i \in E, c_i(\boldsymbol{x}) \leqslant 0, i \in I, \boldsymbol{x} \in R^n\}. \tag{1.2.7}$$

这样我们可以定义约束问题的最优解.

**定义 1.2.3** 若存在 $\boldsymbol{x}^* \in D$，使得对任意的 $\boldsymbol{x} \in D$，均有

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*), \tag{1.2.8}$$

则称 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 在可行域上的全局极小点，或称 $\boldsymbol{x}^*$ 是约束问题全局解. 若对任意的 $\boldsymbol{x} \in D$，$\boldsymbol{x} \neq \boldsymbol{x}^*$，均有

$$f(\boldsymbol{x}) > f(\boldsymbol{x}^*), \qquad (1.2.9)$$

则称 $\boldsymbol{x}^*$ 是 $f$（$\boldsymbol{x}$）在可行域上的严格全局极小点，或称 $\boldsymbol{x}^*$ 是约束问题严格全局解.

与无约束问题一样，满足（1.2.8）式或（1.2.9）式的 $\boldsymbol{x}^*$，称为约束问题的最优解.

### 1.2.3 最优化问题的图解法

**【例 1.2.4】** 根据图像求解无约束问题

$$\min f(\boldsymbol{x}) = (x_1-2)^2+(x_2-2)^2+4.$$

**解** 先画出曲面 $y=f(x)$ 的图像，该曲面的最低点在 $x_1 0 x_2$ 上的投影为问题的解. 为寻找该投影点，考虑 $f$ 的等高线，$f(\boldsymbol{x})=k$，即

$$(x_1-2)^2+(x_2-2)^2=k-4,$$

（见图 1.2.3）当 $k>4$ 时，等高线为同心圆，其同心圆的中心是 $(2,\ 2)^T$，即是问题的最优解 $\boldsymbol{x}^*$，其目标函数值为 $f$（$\boldsymbol{x}^*$）$=4$.

**【例 1.2.5】** 用图解法求解约束问题

$$\begin{aligned}
\min \quad & f(\boldsymbol{x}) = -3x_1-4x_2, \\
\text{s.t} \quad & c_1(\boldsymbol{x}) = x_1+2x_2-6 \leqslant 0, \\
& c_2(\boldsymbol{x}) = 3x_1+2x_2-12 \leqslant 0, \\
& c_3(\boldsymbol{x}) = x_2-2 \leqslant 0, \\
& c_4(\boldsymbol{x}) = -x_1 \leqslant 0, \\
& c_5(\boldsymbol{x}) = -x_2 \leqslant 0.
\end{aligned}$$

**解** 首先画出问题的可行域 $D$，它是由五条直线组成的区域（见图 1.2.4），然后再画出等高线

$$-3x_1-4x_2=k,$$

这是若干条直线. 由图形可知，在两条直线 $x_1+2x_2=6$、$3x_1+$