

SSS

中国石化集团公司职工培训系列教材
技-SBI/123-P-0004

石油化工工艺工程师必读

第一册

物性、数据处理和优化

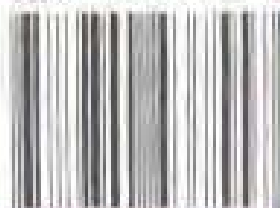
臧福录 应金良 主编



中国石化出版社

责任编辑 潘向阳
主 题 词
封面设计 阮 晓
责任校对 王 红

ISBN 7-80043-755-8



9 787800 437557 >

ISBN 7-80043-755-8/TQ-442

定价: 15.00元 (含邮费)

中国石化集团公司职工培训系列教材

石油化工工艺工程师必读

第一册

物性、数据处理和优化

臧福录 应金良 主编

中国石化出版社

内 容 提 要

《石油化工工艺工程师必读》共分三册,第一册物性、数据处理和优化(带磁盘一张),第二册多组分精馏的计算和分析,第三册化学反应器的计算和分析。

本书是第一册,流体物性估算部分介绍了气体和液体的密度、焓、热容、粘度、导热系数估算,蒸汽压的估算,石油馏分性质的估算及在自编程序中调用 BIPTCHEM 软件包物性数据库的方法。数据处理部分介绍了实验数据处理的常用手段、实验数据回归分析及化工过程数学模型的参数估计和检验等。化工过程优化介绍线性规划和非线性规划的求解方法、化工过程的优化方法和化工过程的优化实例。

本书所带软件是可用于工程实际计算的 BIPTCHEM 软件包,不仅可进行物性估算、数据处理、优化计算,还可配合第二、三分册进行精馏的计算、换热器的计算、板式塔的计算和化学反应器的计算。

本书适合于石油化工技术人员使用,也可供大专院校在校师生参考。

图书在版编目(CIP)数据

石油化工工艺工程师必读 第1册:物性、数据处理和优化/臧福录主编.
北京:中国石化出版社,1998
中国石油化工集团公司职工培训系列教材
ISBN 7-80043-755-8

I.石… II.臧… III.①流体-物性学-技术培训-教材
②流体-实验数据-数据处理-技术培训-教材③流体-化工
过程-最佳化-技术培训-教材 IV.TQ022.1

中国版本图书馆 CIP 数据核字(98)第 25577 号

中国石化出版社出版发行
地址:北京市东城区安定门外大街 58 号
邮编:100011 电话:(010) 64241850
海丰印刷厂排版印刷
新华书店北京发行所经销

*

787×1092 毫米 15 开本 10.25 印张 259 千字 印 1—2500

1998 年 12 月第 1 版 1998 年 12 月第 1 次印刷

定价:45.00 元(含磁盘)

序

由臧福录教授和应金良副教授主编的《石油化工工艺工程师必读》正式出版发行，为从事石油化工及相近专业工作的技术人员接受继续工程教育提供了一套难得的好教材，为他们分析和解决生产中的技术问题提供了一套极有参考价值的工具书，也为同类专业的高校师生提供了一套很好的教学参考书。

世界科学技术飞速进步，科技知识几乎以几何级数迅速增加，自然科学、社会科学各学科之间互相渗透交融，一系列边缘学科、综合学科相继出现，科技成果更新速度的不断缩短，必将把对在职工程技术人员进行继续工程教育的课题客观而严肃地摆到了我们面前。80年代中期，大家对继续工程教育的必要性、重要性的认识开始转化成具体行动。1984年8月茅以升主编的《现代工程师手册》一书是我国工程技术人员接受继续工程教育的第一部通用教科书。但工程技术专业门类众多，不仅需要通用教材解决技术人员普遍需要的知识更新问题，还必须配有各类专业教材，使继续工程教育能体现专业特点，适合专业需求，更具有针对性。臧福录教授敏锐地觉察到了这一问题。他搞过教学，在石油化工专业方面有很深的理论功底；做过科研，组织指导过许多研究课题；多年在燕山石化公司工作，熟悉石油化工生产，利用这些优势，通过认真调查，针对石油化工工艺工程师的知识水平及需求，组织编写了这套供石油化工工艺工程师继续教育使用的教材。近十年来，在燕山石化公司、大庆石化总厂及扬子石化公司等单位领导的支持下，开办了20余期石油化工工艺工程师继续教育进修班，进行教学实践。由于教材与教学有很强的针对性、实用性，知识有相当的新颖性，使学员受益匪浅、反映良好。编者虚心听取学员的意见，做到教学相长，不断对教材进行修改、补充，在大量的教学实践中提炼升华，现在交付正式出版。因此，本书是编者从事石油化工继续教育实践经验的结晶，汇集了学员们的建议，更凝结了编者的心血和智慧。

本套丛书内容包括了流体的物性估算、实验数据处理、石油化工最优化技术和应用、多组分精馏的计算及分析、典型化学反应器的设计计算和操作分析。实验数据处理部分包含逐步回归和非线性回归；流体物性估算部分包含有624种物质的基础数据库，纯组分和混合物气、液相物性的计算，石油馏分物性的计算；石油化工优化部分包括统计调优法的应用实例；多组分精馏的计算部分包括了单级平衡的计算，多组分精馏的简捷计算和严格计算，冷凝器和再沸器计算，板式塔的计算；典型化学反应器的计算部分包括间歇和连续操作釜式反应器，管式反应器，多段绝热固定床反应器，列管式固定床反应器等各种类型的反应器计算。本套丛书把建立数据模型、用计算机软件计算和解决工程实际问题作为重点，并且附有便于自学的源程序和可用于工程实际计算的应用软件。书中各篇既有利用软件计算，又有结合实际的例题和习题。这种应用软件与书本有机结合的读物，在国内外石油化工专业读物中尚属少见，十分适合石油化工类专业技术人员的自学和提高，这不仅包括没有接受过计算机应用技术教育的老工程技术人员，还包括接受过计算机应用技术教育、刚从大学毕业的技术人员，因为他们并不了解在实际的技术工作中如何应用计算机技术，本套丛书恰恰在这方面可以给予具体的指导和帮助。在燕山石化公司，目前能应用计算机及相应软件对生产过程进

行流程模拟、优化计算，提出改进操作和进行脱瓶颈改造建议并取得明显技术成果的科技人才几乎都曾接受过本教材的培训和学习。

我相信本套丛书一定会受到广大石油化工类专业技术人员的欢迎，成为大家经常会去翻阅和研读的最喜爱的读物。我认为，作为一名石油化工专业的技术人员，拥有这本书定会终生受益。

曹湘洪

1997 年 10 月于北京燕山石化公司

前 言

在市场经济深入发展和改革力度不断加大的新形势下，石化行业面临着严峻的挑战。为迎接这一挑战，面向市场经济，必须加强对企业人员的继续工程教育，向科技要生产力。《石油化工工艺工程师必读》正是为适应这一要求而编写的。

《石油化工工艺工程师必读》是编者在从事近十年石油化工继续工程教育实践经验的基础上编写的，重点介绍数学模型的建立及使用计算机软件计算和解决工程实际问题。为强化计算机的应用，强调定量地分析石油化工生产中的典型问题，本书还有与之配套使用的计算机软件。读者通过本书的学习及所附软件的使用，可分析和解决石油化工工程中常见的许多问题。

全书共分三册：第一册介绍流体物性的估算、实验数据处理及化工优化技术和应用；第二册为多组分精馏的计算和分析；第三册为典型化学反应器的设计计算和操作分析。本书由臧福录教授，应金良主编：第一册由王琴、曹长兴、程宏远、应金良和李红文编写，第二册由应金良和李翠清编写，第三册由臧福录和盖旭东编写；全书由臧福录和应金良统稿，由郭天民教授主审。与本书配套使用的 BIPTCHEM 软件包，由应金良主编。

《石油化工工艺工程师必读》第一册编写分工如下：第一篇程宏远和李红文，第二篇王琴，第三篇曹长兴和程宏远，第四篇应金良。

本册由郭天民教授和金海林副教授审稿。他们对本书提出了许多宝贵的意见和建议，我们对此表示衷心的感谢。

本书在编写和出版过程中，得到了中国石化集团公司人教部培训处、北京石油化工学院及中国石化出版社等许多单位领导和有关同志的关心、鼓励和支持。燕山石化集团公司的曹湘洪总经理还在百忙中抽出时间为本书作序。我们还要特别感谢李秋燕同志，她为我们做了许多协调工作。在此，我们向所有关心和支持本书工作的同志表示衷心的感谢。

由于作者水平有限，书中不妥或错误之处在所难免，恳请同行及广大读者批评指正。

编著者

目 录

第一篇 流体物性

第1章 基础物性数据	(1)
1.1 查询纯物质常数	(1)
1.2 BIPTCHEM 软件物性数据库结构及调用	(3)
1.2.1 BIPTCHEM 软件物性数据库的结构	(3)
1.2.2 BIPTCHEM 软件物性数据库的调用	(5)
第2章 流体的热力学性质	(7)
2.1 状态方程	(7)
2.1.1 RK 方程	(7)
2.1.2 SRK 方程	(8)
2.1.3 PR 方程	(8)
2.1.4 Lee-Kesler 方程	(9)
2.1.5 状态方程的求解	(12)
2.2 液体密度	(13)
2.3 蒸气压	(15)
2.3.1 Lee-Kesler 方程	(16)
2.3.2 Gomez-Thodos 蒸气压方程	(16)
2.4 汽化热、比热容和生成热	(17)
2.4.1 汽化热	(17)
2.4.2 理想气体比热容	(17)
2.4.3 生成热	(18)
2.4.4 液体比热容	(18)
2.5 焓的计算	(18)
2.5.1 焓差的计算	(18)
2.5.2 理想气体焓	(20)
第3章 流体的传递性质	(21)
3.1 粘度	(21)
3.1.1 低压下纯气体粘度	(21)
3.1.2 低压下气体混合物粘度	(22)
3.1.3 高压下纯气体粘度	(24)
3.1.4 高压下气体混合物粘度	(25)
3.1.5 液体粘度	(26)
3.2 导热系数	(27)
3.2.1 低压下气体导热系数	(27)
3.2.2 高压下气体导热系数	(28)

3.2.3 低压下气体混合物导热系数	(29)
3.2.4 高压下气体混合物导热系数	(30)
3.2.5 液体导热系数	(30)
3.2.6 液体混合物导热系数	(31)
第4章 石油和石油馏分物性计算	(32)
习 题	(36)
参考文献	(37)
附录 计算程序	(38)

第二篇 实验数据处理

第5章 随机误差及其统计分布	(43)
5.1 误差的概念	(43)
5.1.1 误差定义	(43)
5.1.2 误差分类	(44)
5.1.3 误差的表示方法	(44)
5.1.4 等精度测量与非等精度测量	(47)
5.1.5 加权平均	(47)
5.2 随机误差的统计分布	(48)
5.2.1 频数分布	(48)
5.2.2 正态分布	(50)
5.2.3 概率计算	(51)
习 题	(54)
第6章 离群值的检验与取舍	(55)
6.1 离群值的舍弃准则	(55)
6.2 离群值的检验方法	(55)
6.2.1 格鲁布斯 (Grubbs) 检验法	(55)
6.2.2 肖维奈特 (Chauvenet) 准则	(56)
习 题	(58)
第7章 回归分析与数据处理	(59)
7.1 回归分析和曲线拟合的概念	(59)
7.1.1 回归分析	(59)
7.1.2 曲线拟合	(59)
7.2 一元线性回归	(60)
7.2.1 经验公式	(60)
7.2.2 最小二乘法	(60)
7.2.3 相关性检验	(61)
7.2.4 回归系数的区间估计	(65)
7.2.5 预测与控制	(66)
7.3 多元线性回归	(69)
7.3.1 多元线性回归的数学模型	(69)

7.3.2 最小二乘法	(69)
7.3.3 相关性检验	(70)
7.3.4 逐步回归	(71)
7.4 非线性回归	(73)
7.4.1 可线性化的一元非线性回归	(74)
7.4.2 多项式回归	(76)
7.4.3 可线性化的多元非线性回归	(77)
7.4.4 不可线性化的非线性回归	(77)
习 题	(78)
参考文献	(80)

第三篇 化工过程优化技术及应用

第8章 最优化方法	(81)
8.1 概论	(81)
8.1.1 优化问题的性质和数学描述	(81)
8.1.2 最优化方法的分类	(85)
8.1.3 优化问题求解的一般步骤	(86)
8.2 线性规划	(87)
8.2.1 线性规划问题的形成	(87)
8.2.2 线性规划问题的数学模型	(88)
8.2.3 线性规划问题的解和图解法	(90)
8.2.4 线性规划问题的单纯形法	(93)
8.2.5 线性规划问题的应用	(98)
8.3 非线性规划	(101)
8.3.1 非线性规划基本概念	(101)
8.3.2 无约束问题最优化	(105)
8.3.3 有约束问题最优化	(111)
习 题	(116)
参考文献	(120)
第9章 调优操作法	(121)
9.1 概述	(121)
9.2 统计调优法	(122)
9.2.1 统计调优策略	(122)
9.2.2 统计调优实例	(126)
9.3 模拟调优法概要	(133)
参考文献	(134)

第四篇 配套软件介绍

第10章 化工过程计算软件包	(135)
10.1 软件包的安装、运行及主菜单的打开	(135)

10.1.1	安装.....	(135)
10.1.2	运行.....	(136)
10.1.3	主菜单的打开.....	(136)
10.2	物性计算.....	(136)
10.2.1	基础数据库.....	(137)
10.2.2	估算纯物质的性质.....	(138)
10.2.3	估算混合物的性质.....	(139)
10.2.4	估算石油馏分的性质.....	(139)
10.3	数据处理.....	(140)
10.3.1	非线性回归.....	(140)
10.3.2	逐步回归.....	(142)
10.4	优化计算.....	(142)
10.4.1	线性规划.....	(142)
10.4.2	非线性规划 (拉格朗日乘因子法和单纯形法)	(144)
10.4.3	非线性规划 (复合形法)	(145)
附录	化工过程计算机软件中的 624 个组分.....	(145)

第一篇 流体物性

化工设计和计算中需要可靠的流体物性。对常见的工业物质可查到丰富的实验数据^{[1][2]}，但是，在许多情况下还需要流体物性的关联式和估算式。

流体的物理性质取决于组成流体物质的分子特征，人们已了解了许多分子行为的谜团，可至今还无法完全掌握分子运动的所有特征^[3]。因此，目前计算流体性质的方法多数为半经验半理论，或纯经验的计算式。

本部分内容为常用流体热力学性质和传递性质的计算方法。它们包括流体的密度、蒸气压、热容、焓、粘度和导热系数等性质。这些性质的计算方法选自文献^[3]和^[4]，适用于常见流体物性的工程计算。多数性质的计算方法已编入“化工过程计算 (BIPTCHEM)”软件包，因此本篇举例说明某性质计算方法时，多借助该软件包来完成。一些物性计算方法、特别是一些经验方法在形成时沿用的是非 SI 单位，它们在广泛使用中，对不同的物质积累了大量来自实验数据的方程参数，故此，仅对这些方程计算的最终物理量给出 SI 单位换算系数。

有关其他的流体物性的计算方法，可参考相关文献^{[1]--[5]}。

第 1 章 基础物性数据

基础物性数据指纯物质常数和与物质有关的关联方程参数，如相对分子量、临界温度、临界压力、热容关联式参数等等。本章介绍利用 BIPTCHEM 软件包查询基础数据以及在自编程序中调用 BIPTCHEM 软件包数据库中数据的方法。具体计算式可参见文献^{[1][2]}和^[5]。

1.1 查询纯物质常数

启动 BIPTCHEM，在屏幕中选择“物性”一项，屏幕出现项目菜单：

基础数据库
估算纯物质性质
估算混合物的性质
石油馏分性质的估算
单位换算
退出

选择“基础数据库”一项，屏幕出现弹出菜单：

查询数据库中的基础数据
 改变数据库中的基础数据
 在数据库中增加一种物质基础数据
 查询或改变 Wilson 二元交互作用能量参数
 退出

选择“查询数据库中的基础数据”一项，屏幕出现：

数据库中有 624 种物质
 输入组分的标号或分子式—

在此可输入某物质的标号或分子式。各物质的标号可从第十章的附录中查到。

例 1.1 查询 1-戊烯的纯物质常数。

按上述方法运行软件，直到屏幕出现：

数据库中有 624 种物质
 输入组分的标号或分子式—

可键入“C5H10”，按回车，屏幕出现：

环戊烷
 1-戊烯
 2-戊烯（顺）
 2-戊烯（反）
 2-甲基-1-丁烯
 2-甲基-1 丁烯
 3-甲基-1-丁烯
 EXIT

选择“1-戊烯项”，按回车，屏幕显示出（排列方式与此处不同）：

标号：294 分子式：C5H10 名称：1-戊烯
 相对分子量：70.134
 凝固点：107.9 K
 沸点：303.1K
 临界温度：464.8 K
 临界压力：35.3 bar
 临界体积：300 cm³/mol
 临界压缩因子：0.31
 偏心因子：0.233
 偶极矩：0.4 debye
 293K 时的液相密度：640 kg/m³
 理想气体摩尔热容计算式：C_p = a + b * T + C * T² + d * T³ [J/mol·K]

各系数如下:

$$a = -0.134 \quad b = 0.4329 \quad c = -0.0002317 \quad d = 4.681E-08$$

生成焓: -20930 J/mol 生成 Gibbs 能: 79170 J/mol

汽化热: 25196.65 J/mol

理想气体焓计算式: $H^0 = a + bT + cT^2 + dT^3 + eT^4 + fT^5 \text{ [J/mol]}$

理想气体熵: $S^0 = g + b \ln T + 2cT + 1.5dT^2 + 4eT^3/3 + 1.25fT^4 \text{ [J/mol} \cdot \text{K]}$

各系数如下:

$$a = 4919.244 \quad b = -2.018426 \quad c = 0.2225423 \quad d = 8.641272E-05$$

$$e = 1.718969E-08 \quad f = -9.737707E-13 \quad g = 235.8965$$

粘度 μ 计算式: $\ln(\mu) = a + b/T$ (μ ——cP, T ——K), 其中:

$$a = -4.023 \quad b = 702.9, \text{ 温度从 } 183.2\text{K} \text{ 到 } 273.2\text{K}$$

温度 273.2K 时的粘度为: 0.24 cP

饱和蒸气压方程:

$$\lg(P_s/P_c) = a(1 - T_r) + b(1 - T_r)^{1.5} + c(1 - T_r)^3 + d(1 - T_r)^6/T_r$$

$$a = -7.04875 \quad b = 1.17813 \quad c = -2.45105 \quad d = -2.21727$$

温度从 190K 到 464.8K

液体导热系数计算式: $DI = a + b(T - c) \text{ W/m} \cdot \text{K}$

$$a = 0.120952 \quad b = -4.4194E-04 \quad c = 293.2$$

SRK 方程偏心因子: $0.2842 \quad V^* = 0.2951 \quad Z_{RA} = 0.2899$

25°C 时的溶解度参数: $14.5 \text{ MPa}^{0.5}$

25°C 时的摩尔体积: $V_L = 110.4 \text{ cm}^3/\text{mol}$

含有 1 个 CH_3 , 1 个 CH_2 , 1 个 $\text{CH}_2 = \text{CH}$ UNIFAC 基团

有关改变数据库中的基础数据、在数据库中增加一种物质基础数据和查询或改变 Wilson 二元交互作用能量参数等项目的详细使用方法请参见第十章软件使用说明。

1.2 BIPTCHEM 软件物性数据库结构及调用

利用简单的程序可调用 BIPTCHEM 软件数据库中的基础数据, 用于自编程序。通过了解该软件基础物性数据库的结构, 可达到上述目的。

1.2.1 BIPTCHEM 软件物性数据库的结构

BIPTCHEM 软件物性数据库均由 Quick BASIC 语言编制的随机数据文件组成, 它包括 9 个数据文件。

1. 纯物质临界参数文件 (Y1.DAT)

纯物质临界参数文件的文件名为 Y1.DAT。该文件存放纯物质的临界性质, 每一记录中含有 6 个域, 依次存放纯物质的相对分子量、临界温度 (K)、临界压力 (MPa)、临界体积 (cm^3/mol)、临界压缩因子和偏心因子。每个域均为单精度型数。

2. 熔点、沸点和汽化热数据文件 (Y2.DAT)

熔点、沸点和汽化热数据文件的文件名为 Y2.DAT。该文件存放纯物质的熔点 (K)、

沸点 (K) 和汽化热 (J/mol) 数据, 每一记录中含有 3 个域, 依次存放纯物质的熔点、沸点和汽化热。每个域均为单精度型数。

3. 理想气体热容参数文件 (Y3.DAT)

理想气体热容参数文件的文件名为 Y3.DAT。该文件存放理想气体热容关联式参数, 每一记录中含有 4 个域, 依次存放参数 a 、 b 、 c 、 d 。每个域均为单精度型数。该关联式为:

$$C_p = a + bT + cT^2 + dT^3 \quad (1.1)$$

式中 C_p ——理想气体热容, J/mol·K;

T ——温度, K;

a 、 b 、 c 、 d ——参数。

4. 理想气体的焓和熵数据文件 (Y4.DAT)

BIPTCHEM 采用的理想气体焓和熵关联式为:

$$H^0 = a + bT + cT^2 + dT^3 + eT^4 + fT^5 \quad (1.2)$$

$$S^0 = g + b \ln T + 2cT + 1.5dT^2 + \frac{4}{3}eT^3 + 1.25fT^4 \quad (1.3)$$

式中 H^0 ——理想气体的焓, J/mol;

S^0 ——理想气体的熵; J/mol·K;

T ——温度, K;

a 、 b 、 c 、 d 、 e 、 f 、 g ——常数。

Y4.DAT 的每一记录有 7 个域, 依次存放参数 a 、 b 、 c 、 d 、 e 、 f 、 g 。每个域均为单精度型数。

5. 标准生成焓和吉布斯 (Gibbs) 自由能数据文件 (Y5.DAT)

标准生成焓和吉布斯 (Gibbs) 自由能数据文件的文件名为 Y5.DAT。该文件存放标准生成焓 (J/mol) 和吉布斯 (Gibbs) 自由能 (J/mol) 数据, 每个域均为单精度型数。

6. 蒸气压关联式参数文件 (Y6.DAT)

BIPTCHEM 采用的蒸气压关联式为以下三个关联式之一:

$$(1) \lg \frac{P_s}{P_c} = [a(1 - T_r) + b(1 - T_r)^{1.5} + c(1 - T_r)^3 + d(1 - T_r)^6] / T_r \quad (1.4)$$

式中 P_s ——饱和蒸气压, MPa;

P_c ——临界压力, MPa。

$$(2) \lg P_s = a + \frac{b}{T} + c \ln T + d \frac{P_s}{T^2} \quad (1.5)$$

$$(3) \lg P_s = a - \frac{b}{T + c} \quad (1.6)$$

式中 P_s ——饱和蒸气压, mmHg;

$T_r = T / T_c$;

T ——温度, K;

T_c ——临界温度, K。

蒸气压关联式参数文件的文件名为 Y6.DAT。该文件存放蒸气压关联式参数数据, 依次存放关联式编号 (1、2 或 3), 关联式参数 a 、 b 、 c 、 d , 关联式适用温度范围 (最低温度和最高温度 T_{low} 、 T_{max}), 每个记录有 7 个域。关联式编号为整型数, 其他均为单精度型数。

7. 粘度关联式参数文件 (Y7.DAT)

BIPTCHEM 采用的粘度关联式为以下四个关联式之一：

$$(1) \mu = aT^b \quad (1.7)$$

$$(2) \lg \mu = a + \frac{b}{T} \quad (1.8)$$

$$(3) \lg \mu = a + \frac{b}{T} + cT + dT^2 \quad (1.9)$$

$$(4) \lg \mu = a \left(\frac{1}{T} - \frac{1}{b} \right) \quad (1.10)$$

式中 μ ——液体粘度，cP；

T ——温度，K。

粘度关联式参数文件的文件名为 Y7.DAT。该文件存放粘度关联式参数，依次存放关联式编号（1、或 2、或 3、或 4），关联式参数 a 、 b 、 c 、 d ，关联式适用温度范围（最低温度和最高温度 T_{low} 、 T_{max} ），粘度的实验值（cP），实验值的温度条件（K）。每个记录有 9 个域，关联式编号为整型数，其他均为单精度型数。

8. 液体导热系数关联式参数文件（Y8.DAT）

BIPTCHEM 采用的液体导热系数关联式为以下两个关联式之一：

$$(1) \lambda_L = a + \frac{b}{T - c} \quad (1.11)$$

$$(2) \lambda_L = a + bT + cT^2 \quad (1.12)$$

式中 λ_L ——液体导热系数，W/m·K；

T ——温度，K。

导热系数关联式参数文件的文件名为 Y8.DAT。该文件存放导热系数关联式参数数据，依次存放关联式编号（1 或 2），关联式参数 a 、 b 、 c 。每个记录有 4 个域，关联式编号为整型数，其他均为单精度型数。

9. 液相密度（Y9.DAT）

液相密度文件的文件名为 Y9.DAT。该文件依次存放液体密度的实验值（ ρ ），实验值的温度条件，SRK 方程的偏心因子 ω_{SRK} ，液体密度计算式的特征体积值 V^* ，液体密度计算式的 Z_{RA} 值。每个记录有 5 个域，均为单精度型数。

1.2.2 BIPTCHEM 软件物性数据库的调用

例 1.2 调用 BIPTCHEM 软件基础数据库中水的临界参数的 Quick BASIC 程序。

解：利用 Quick BASIC 语言编制的程序如下：

```
TYPE cP
  M AS SINGLE
  TC AS SINGLE
  PC AS SINGLE
  VC AS SINGLE
  ZC AS SINGLE
  W AS SINGLE
END TYPE
DIM JY1 AS cP
```



```

OPEN "\DAT\Y1.DA" FOR RANDOM AS #1 LEN=LEN (JY1)
INPUT "INPUT COMPONENT'S ID(RECORD ID)";ID
GET #1, ID, JY1
TC =JY1.TC
PC =JY1.PC
W =JY1.W
CLOSE #1
PRINT "TC="; TC; [K]
PRINT "PC="; PC; [MPa]
PRINT "W="; W
END

```

水的标号为 77，可由第十章附录查得。打开文件路径中的盘符由 BIPTCHEM 所在硬盘的盘符确定。

第2章 流体的热力学性质

2.1 状态方程

状态方程描述了流体的压力、体积和温度的行为。流体的热力学性质，如流体密度、焓、逸度和相平衡关系，可用状态方程来计算。状态方程可用于除临界区以外非常宽的温度和压力范围。对理想和轻度非理想系，状态方程仅需少量组分数据便可计算汽相和液相的热力学性质。

工程上使用的状态方程主要有两类：立方型状态方程和维里型状态方程。立方型状态方程主要有 Redlich-Kwong (RK)、Redlich-Kwong-Soave (SRK)、Peng-Robinson (PR) 和它们的一些修正式。维里型状态方程主要有 Hayden-O'Connell、BWR-Lee-Starling、Lee-Kesler 和它们的一些修正式。立方型状态方程可描述气（汽）体和液体的许多热力学性质，但不能准确表达液体的摩尔体积，一些状态方程（如 Schwartzentruber-Renon）经过体积校正可实现上述目的。维里型状态方程可以可靠地描述汽相和液相性质，它在许多时候优于立方型状态方程，但它不适用极性混合物。此处主要介绍 RK、SRK、PR 和 Lee-Kesler 状态方程

2.1.1 RK 方程

RK 方程可用于计算汽相的热力学性质，它不适用液体密度的计算。压缩因子形式的 RK 方程为：

$$Z^3 - Z^2 + (A - B - B^2)Z - AB = 0 \quad (2.1)$$

式中

$$Z = \frac{PV}{RT} \quad (2.2)$$

$$A = 0.42748 \frac{P_r}{T_r^{2.5}} \quad (2.3)$$

$$B = 0.08664 \frac{P_r}{T_r} \quad (2.4)$$

其中 Z ——压缩因子，对于理想气体， $Z=1$ ；

P ——压力，Pa；

T ——温度，K；

V ——摩尔体积， m^3/kmol ；

R ——气体常数；

P_r ——对比压力， $P_r = P/P_c$ ；

T_r ——对比温度， $T_r = T/T_c$ ；

P_c ——临界压力，Pa；

T_c ——临界温度, K。

RK 方程应用于混合物时, 相应的常数按以下混合规则计算:

$$A_m = \sum_{i=1}^n \sum_{j=1}^n y_i y_j A_{ij} \quad (2.5)$$

$$B_m = \sum_{i=1}^n y_i B_i \quad (2.6)$$

式中 A_m 、 B_m ——混合物的常数;

A_i 、 B_i ——组分 i 的常数, 分别由式 (2.3) 和式 (2.4) 计算;

y_i 、 y_j ——混合物中组分 i 、 j 的摩尔分数。

$$A_{ij} = (A_i A_j)^{0.5} (1 - k_{ij})$$

为保证计算结果的可靠, 应由相平衡实验数据确定二元参数 k_{ij} 。对烃类物质, k_{ij} 可近似取零。

2.1.2 SRK 方程

SRK 方程可用于烃加工过程, 如石油炼制和石油化工过程。压缩因子形式的 SRK 方程为:

$$Z^3 - Z^2 + (A - B - B^2)Z - AB = 0 \quad (2.7)$$

式中常数 B 和 RK 方程相同, 但在计算常数 A 时引进了第三参数 ω , 其关系式如下:

$$A = 0.42748 \alpha_i \frac{P_r}{T_r^2} \quad (2.8)$$

$$\alpha_i = [1 + m(1 - T_r^{0.5})]^2 \quad (2.9)$$

$$m = 0.48 + 1.574\omega - 0.176\omega^2 \quad (2.10)$$

$$B = 0.08664 \frac{P_r}{T_r} \quad (2.11)$$

式中 ω 为偏心因子。SRK 方程采用以下混合规则:

$$A_m = \sum_{i=1}^c \sum_{j=1}^c y_i y_j A_{i,j} \quad (2.12)$$

$$A_{i,j} = (A_i A_j)^{0.5} (1 - k_{ij}) \quad (2.13)$$

为保证计算结果的可靠, 应由相平衡实验数据确定二元参数 k_{ij} 。对烃类物质, k_{ij} 可近似取零。除临界点附近, 此方程可用于任一温度和压力。

2.1.3 PR 方程

PR 方程可用于烃加工过程, 如气体加工、石油炼制和石油化工过程。其形式为:

$$Z^3 - (1 - B)Z^2 + (A - 2B - 3B^2)Z - (AB - B - B^2) = 0 \quad (2.14)$$

式中常数 A 和 B 按下式计算:

$$A = 0.45724 \alpha_i \frac{P_r}{T_r^2} \quad (2.15)$$

$$\alpha_i = [1 + m(1 - T_r^{0.5})]^2 \quad (2.16)$$

$$m = 0.37464 + 1.54226\omega - 0.26992\omega^2 \quad (2.17)$$

$$B = 0.0778 \frac{P_c}{T_r} \quad (2.18)$$

式中各量含义同 RK 方程, PR 方程采用和 SRK 方程相同的混合规则。为保证计算结果的可靠, 应由相平衡实验数据确定二元参数 k_{ij} 。对烃类物质, k_{ij} 可近似取零。除临界点附近, 此方程可用于任一温度和压力。

2.1.4 Lee-Kesler 方程

Lee-Kesler 方程可以计算包括非烃类在内的许多物质的热力学性质, 对于烃类的热力学性质计算, 它获得了高度的评价, 可用于烃加工过程, 如气体加工、石油炼制和石油化工过程, 其基本形式如下:

$$\frac{P_r V_r}{T_r} = 1 + \frac{B}{V_r} + \frac{C}{V_r^2} + \frac{D}{V_r^3} + \frac{c_4}{T_r^3 V_r^2} \left[\beta + \frac{\gamma}{V_r^2} \right] \exp \left[-\frac{\gamma}{V_r^2} \right] \quad (2.19)$$

式中

$$B = b_1 - \frac{b_2}{T_r} - \frac{b_3}{T_r^2} - \frac{b_4}{T_r^3} \quad (2.20)$$

$$C = c_1 - \frac{c_2}{T_r} + \frac{c_3}{T_r^3} \quad (2.21)$$

$$D = d_1 + \frac{d_2}{T_r} \quad (2.22)$$

$$V_r = \frac{P_c V}{RT_c} \quad (2.23)$$

对于实际流体的压缩因子可表示为:

$$Z = Z^{(0)} + \left[\frac{\omega}{\omega^{(R)}} \right] (Z^{(R)} - Z^{(0)}) \quad (2.24)$$

式中 $\omega^{(R)}$ 为参照流体 (正辛烷) 的偏心因子, $\omega^{(R)} = 0.3978$; $Z^{(0)}$ 为简单流体的压缩因子, $Z^{(R)}$ 为参照流体的压缩因子, 其定义式分别为:

$$Z^{(0)} = \frac{P_r V_r^{(0)}}{T_r} \quad (2.25)$$

$$Z^{(R)} = \frac{P_r V_r^{(R)}}{T_r} \quad (2.26)$$

它们均可由式 (2.19) 计算。式 (2.19) 中, 简单流体和参照流体的各常数值见表 1.1

表 1.1 Lee-Kesler 常数

常 数	简单流体	参照流体	常 数	简单流体	参照流体
c_1	0.1181193	0.2026579	c_3	0.0	0.016901
b_1	0.265728	0.331511	c_4	0.042724	0.041577
c_2	0.154790	0.027665	$d_1 \times 10^4$	0.155488	0.48736
b_2	0.030323	0.203488	$d_2 \times 10^4$	0.623689	0.0740336
d_1	0.0236744	0.0313385	β	0.65392	1.226
c_4	0.0186984	0.0503618	γ	0.060167	0.03754

LK 方程采用以下混合规则：

$$T_{cm} = \frac{1}{8V_{cm}} \sum_{i=1}^c \sum_{j=1}^c y_i y_j (V_{ci}^{1/3} + V_{cj}^{1/3})^3 (T_{ci} T_{cj})^{0.5} \quad (2.27)$$

$$V_{cm} = \frac{1}{8} \sum_{i=1}^c \sum_{j=1}^c y_i y_j (V_{ci}^{1/3} + V_{cj}^{1/3})^3 \quad (2.28)$$

$$V_{ci} = (0.2905 - 0.085\omega_i) \frac{RT_{ci}}{P_{ci}} \quad (2.29)$$

$$P_{cm} = (0.2905 - 0.085\omega_m) \frac{RT_{cm}}{V_{cm}} \quad (2.30)$$

$$\omega_m = \sum_{i=1}^c y_i \omega_i \quad (2.31)$$

以上各式中， V_c 为临界体积，下标 i 表示组分 i ， m 表示混合物。Lee-Kesler 方程可很好地描述非极性和轻极性混合物的体积性质和热力学性质。除临界点附近，此方程可用于任一温度和压力。

例 2.1 求等摩尔 CO_2 和 C_3H_6 （丙烯）混合物在 30°C 和 2.58MPa 的压缩因子。实验值为 0.737 。

解：利用 BIPTCHEM 软件包求解。

启动软件后，在软件中选择“物性”菜单，按回车，屏幕弹出菜单：

```
基础数据库
估算纯物质性质
估算混合物的性质
石油馏分性质的估算
单位换算
退出
```

用箭头键选择“估算混合物的性质”一项，屏幕显示出：

```
请输入系统的组分数：
```

根据本题目，键入“2”，按回车，屏幕显示出：

```
数据库中有 624 种物质，请输入各组分标号或分子式：
序号 分子式    名称
1
```

在数字 1 后键入字母及数字“ CO_2 ”，屏幕弹出菜单：

```
二氧化碳
Exit
```

选择“二氧化碳”一项后，在“1. CO_2 ”下一行键入“ C_3H_6 ”，按回车，屏幕弹出菜单：

环丙烷
丙烯
Exit

选择“丙烯”一项，屏幕显示：

是否打印？(Y/N)

键入 Y 或 N，屏幕显示：

输入组分 1 的摩尔分数：

键入“0.5”后，屏幕显示：

输入温度：[K]

键入“303”，注意：此处温度单位是“K”。屏幕显示：

压力的单位：

1. kg/cm²

2. atm

3. kPa

4. mmHg

根据本题选择 3 后键入“2580”，按回车。屏幕显示出可以进行计算的性质包括：

1. 液体粘度
2. 液体导热系数
3. 液体比热容
4. 液体密度
5. 液相中各组分活度系数
6. 液相中各组分逸度系数
7. 气体的粘度
8. 气体的导热系数
9. 气体的比热容
- A. 气体的密度
- B. 气相中各组分的逸度系数
- C. 气相焓
- D. 液相焓
- 退出

根据本题选择“A”项。屏幕显示出计算结果为：

二氧化碳-丙烯气体密度			
方程	压缩因子	体积 m ³ /kmol	密度 kg/m ³
1) RK	0.7419801	(略)	(略)
2) PR	0.7159036	(略)	(略)
3) SRK	0.7332642	(略)	(略)
4) LK	0.7045089	(略)	(略)

计算结束。与实验值比较计算误差为：

方程	误差 %
RK	0.66
PR	2.86
SRK	0.51
LK	4.41

通过与实验值比较得知，此时 RK 和 SRK 方程计算结果比较准确。

除在临界区内，上述状态方程对烃类和含轻质气体 CO₂、N₂、H₂S 的烃类物系准确度在 5% 以内。在乙烯生产过程中，利用 SRK 和 PR 方程计算汽油汽提塔、凝液汽提塔、脱甲烷塔、脱乙烷塔、乙烯塔、脱丙烷塔和丙烯塔都获得了成功^[7]。

2.1.5 状态方程的求解

立方型 (RK、SRK 及 PR) 状态方程有三个根，最大根为气相根，最小根为液相根，中间的根没有物理意义。此类方程常用牛顿迭代法求解，也可用解析法求解。SRK 方程计算压缩因子的计算步骤如下：

- (1) 输入各组分的临界参数、摩尔分数及体系的温度和压力： T_{ci} 、 P_{ci} 、 ω_i 、 x_i 、 T 及 P ；
- (2) 计算各组分的对比温度和对比压力： $T_{ri} = T/T_{ci}$ ； $P_{ri} = P/P_{ci}$ ；
- (3) 按式 (2.8) 和式 (2.11) 计算各组分的常数： A_i 、 B_i ；
- (4) 按式 (2.12) 和式 (2.13) 计算混合物的常数： A_m 、 B_m ；
- (5) 解三次方程求得压缩因子。

附录中的程序 P1 为 Quick BASIC 语言编写的用 SRK 方程计算气相或液相压缩因子的源程序。程序中 C 为组分数， T_c 、 P_c 、 ω 、 x 、分别为组分的临界温度、临界压力、偏心因子及组分的摩尔分数， T 、 P 为温度和压力。

程序中 fz 为解析法解立方方程的子程序（在 Quick BASIC 中称函数），其参数 A 、 B 、 C 为三次方程 $Z^3 + AZ^2 + BZ + C = 0$ 的三个系数；参数 $y\$$ 表示是气相还是液相（当参数 $y\$$ 的传递值为“V”时，返回最大根（即气相压缩因子），传递值为“L”时返回最小根（即液相压缩因子））。

利用 P1 程序，输入组分数及各组分的临界温度、临界压力、偏心因子、摩尔分数及温度和压力，分别选择气相或液相计算，可求得气相或液相的压缩因子

例 2.2 试计算含丙烯的摩尔分数为 60%，丙烷的摩尔分数为 40% 的丙烯-丙烷混合气在温度为 313K、压力为 1500 kPa 时的压缩因子。

解：利用 BIPTCHEM 软件包求解。

得气相压缩因子为 0.7743。

例 2.3 试用 SRK 方程求例 2.2 中的丙烯-丙烷溶液在温度为 313K，压力 1700 kPa 时的

液相压缩因子。

解：启动 Quick BASIC，调入 P1 程序，运行该程序后根据屏幕提示输入组分数、各组分临界温度、临界压力、偏心因子及温度和压力后，选择液相计算，屏幕显示液相压缩因子为 $Z_L = 0.06538$ 。

前已述及，LK 方程可预示非烃类的气、液相性质，但它计算复杂。LK 方程计算压缩因子的步骤如下：

- (1) 按 LK 方程的混合规则计算混合物的 T_{cm} 、 V_{cm} 、 P_{cm} 及 ω_m ；
 - (2) 计算对比温度和对比压力；
 - (3) 由式 (2.19) 用简单流体的常数计算简单流体的 $V_r^{(0)}$ ；
 - (4) 由式 (2.25) 计算简单流体的压缩因子 $Z^{(0)}$ ；
 - (5) 由式 (2.19) 用参照流体的常数计算参照流体的 $V_r^{(R)}$ ；
 - (6) 由式 (2.25) 计算参照流体的压缩因子 $Z^{(R)}$ ；
 - (7) 由式 (2.24) 计算实际流体的压缩因子 Z 。
- 由式 (2.19) 计算 $V_r^{(0)}$ 和 $V_r^{(R)}$ 时，需用牛顿迭代法求解。

2.2 液体密度

对大多数纯组分，在知道至少一个实测饱和液体密度的情况下，可由 Spencer 等修正的 Rackett 方程计算某温度下的饱和液体密度。

$$\rho^{\text{sat}} = \frac{P_c}{RT_c Z_{RA} (1 + (1 - T_r)^{2/7})} \quad (2.32)$$

式中 ρ^{sat} ——饱和液体密度， kg/m^3 ；

P_c ——临界压力，Pa；

R ——气体常数；

T_c ——临界温度，K；

Z_{RA} ——由实验数据而定的常数；

T_r ——对比温度， $T_r = T/T_c$ 。

除了对乙醇和酸的误差小于 2% 外，此式推算误差小于 1%。Spencer 等列出了 165 种化合物的 Z_{RA} 值，其值参见文献[3]。若 Z_{RA} 不能由实验决定，则可用下式估算：

$$Z_{RA} = 0.29056 - 0.08775\omega \quad (2.33)$$

式中 ω 为物质的偏心因子，然而使用上式计算的 Z_{RA} 将导致误差增大。

例 2.4 试计算苯在 423.15K 时饱和液体密度 ρ^{sat} 。文献值为 0.7295 g/cm^3 。

解：利用 BIPTCHEM 软件包求解。按前例相同的方法输入已知化合物名称及体系的状态参数，计算结果为：

Spencer 和 Danner - Rackett 方法：	734 kg/m^3
Hankinson - Thomson 方程：	730 kg/m^3
Bhirud 方程：	722 kg/m^3
Tait 法（被压缩流体）：	732.73 kg/m^3

计算误差：

Spencer 和 Danner - Rackett 方法:	0.62 %
Hankinson - Thomson 方程:	0.07 %
Bhirud 方程:	1.0 %
Tait 法 (被压缩流体):	0.4 %

若没有实验数据, 可利用 Bhirud 关联式, 该式可用于非极性化合物。Bhirud 关联式:

$$\lg \frac{P_c V_s}{RT} = \lg V^{(0)} + \omega \lg V^{(1)} \quad (2.34)$$

$$\lg V^{(0)} = 1.39644 - 24.076 T_r + 102.615 T_r^2 - 255.719 T_r^3 \\ + 355.805 T_r^4 - 256.617 T_r^5 + 75.1088 T_r^6 \quad (2.35)$$

$$\lg V^{(1)} = 13.4412 - 135.7437 T_r + 533.380 T_r^2 - 1091.453 T_r^3 \\ + 1231.43 T_r^4 - 728.227 T_r^5 + 176.737 T_r^6 \quad (2.36)$$

式中 V_s ——饱和液体体积, m^3 ;

P_c ——临界压力, Pa;

R ——气体常数;

T ——温度, K;

ω ——偏心因子。

当 $0.3 < T_r < 1.0$ 时, 此式对碳氢化合物的 752 个数据点进行计算时, 平均误差为 0.76 %。

例 2.5 计算正己烷在 423.15 K 的饱和液体密度, 实验值为 521.7 kg/m^3 。

解: 利用 BIPTCHEM 软件包求解, 按前例相同的方法输入已知化合物名称及体系的状态参数, 进行计算得到液体密度为: 520 kg/m^3 。误差为: 0.33 %。

同样, 在没有实验数据可以利用时, 也可用 Hankinson - Thomson 方程, 即

$$\frac{V_s}{V^*} = V_R^{(0)} (1 - \omega_{\text{SRK}} V_R^{(1)}) \quad (2.37)$$

$$V_R^{(0)} = 1 + a(1 - T_r)^{1/3} + b(1 - T_r)^{2/3} + c(1 - T_r) + d(1 - T_r)^{4/3} \quad 0.25 < T_r < 0.95 \quad (2.38)$$

$$V_R^{(1)} = (e + fT_r + gT_r^2 + hT_r^3) / (T_r - 1.00001) \quad 0.25 < T_r < 1.0 \quad (2.39)$$

$$\begin{aligned} a &= -1.52816 & b &= 1.43907 \\ c &= -0.81446 & d &= 0.190454 \\ e &= -0.296123 & f &= 0.389614 \\ g &= -0.0427258 & h &= -0.0480645 \end{aligned}$$

T_r , V^* , ω_{SRK} 由参考文献 [3] 中表 3-10 查得。

高压下液体密度可用 Tait 方程估算。

Tait 方程:

$$V = V_s \left(1 - C \lg \frac{\beta + P}{\beta + P_s} \right) \quad (2.40)$$

式中 V ——液体体积, m^3 ;

V_s ——在蒸汽压 P_s 下饱和液体体积, m^3 ; 由 Hankinson-Thomson 方程计算;

P ——压力, Pa;

β , C ——常数。

$$\beta/P_c = -1 + a(1 - T_r)^{1/3} + b(1 - T_r)^{2/3} + d(1 - T_r) + e(1 - T_r)^{4/3} \quad (2.41)$$

$$e = \exp(f + g\omega_{\text{SRK}} + h\omega_{\text{SRK}}^2) \quad (2.42)$$

$$c = j + K\omega_{\text{SRK}} \quad (2.43)$$

ω_{SRK} 可由文献 [3] 中表 3-10 查得:

$$a = -9.070217 \quad b = 62.45326$$

$$d = -135.1102 \quad f = 4.79594$$

$$g = 0.250047 \quad h = 1.14188$$

$$j = 0.0861488 \quad k = 0.0344483$$

液体混合物密度可用修正的 Rackett 方程计算:

$$V_m = R \left\{ \sum_i \frac{x_i T_{ci}}{P_{ci}} \right\} Z_{\text{RAM}}^{[1+(1-T_r)^{2/3}]} \quad (2.44)$$

$$Z_{\text{RAM}} = \sum_i x_i Z_{\text{RA},i} \quad (2.45)$$

$$T_r = T/T_{\text{cm}} \quad (2.46)$$

$$T_{\text{cm}} = \sum_i \sum_j \Phi_i \Phi_j T_{cij} \quad (2.47)$$

$$\Phi_i = \frac{x_i V_{ci}}{\sum_i X_i V_{ci}} \quad (2.48)$$

$$T_{cij} = (1 - K_{ij})(T_{ci} T_{cj})^{1/2} \quad (2.49)$$

$$1 - K_{ij} = \frac{8(V_{ci} V_{cj})^{1/2}}{(V_{ci}^{1/8} + V_{cj}^{1/2})^3} \quad (2.50)$$

式中 V_m ——液体混合物体积, m^3 ;

x_i —— i 组分的摩尔分数;

T_{ci} —— i 组分的临界温度, K;

P_{ci} —— i 组分的临界压力, Pa;

V_{ci} —— i 组分的临界体积, m^3 。

2.3 蒸气压

广泛使用的纯物质的蒸气压关联式是 Antoine 方程, 即

$$\lg P^{\text{sat}} = A - B/(T + C) \quad (2.51)$$

式中, P^{sat} 是蒸气压 (mmHg, 1mmHg = 133.322Pa); T 为温度 (K); A , B , C 为方程常数, 由实验数据回归得到, 对许多物质已有公开发表的数据, 可由 BIPTCHEM 软件包查得

上式应用温度的范围不大, 对应压力应用范围为 $0.0133 \times 10^5 \text{Pa} \sim 2.0 \times 10^5 \text{Pa}$ 。方程不能在规定的温度范围外应用。此外, 还可由下述方程计算蒸气压。

2.3.1 Lee-Kesler 方程

$$\lg P_{br} = f^{(0)}(T_r) + \omega f^{(1)}(T_r) \quad (2.52)$$

$$f^{(0)} = 592714 - 609480/T_r - 128862 \lg T_r + 16934.7 T_r^6 \quad (2.53)$$

$$f^{(1)} = 1525180 - 1568750/T_r - 1347210 \lg T_r + 43577 T_r^6 \quad (2.54)$$

式中 $P_{br} = P_s/P_c$;

P_s ——蒸气压, Pa;

P ——临界压力, Pa;

T_r —— T/T_c ;

T ——温度, K;

T_c ——临界温度, K。

此式在正常沸点 T_b 和 T_c 之间, 误差通常为 1%~2%, 在 T_b 以下时, 误差低于 10%。

2.3.2 Gomez-Thodos 蒸气压方程:

$$\lg P_{sr} = \beta \left[\frac{1}{T_r^m} - 1 \right] + \gamma [T_r^* - 1] \quad (2.55)$$

式中 $P_{sr} = P_s/P_c$;

β , m , γ ——系数。

在临界点附近使用上式也可以获得满意的结果。在正常沸点时, β , m , γ 用下列关联式算出:

$$\gamma = ah + b\beta \quad (2.56)$$

$$a = \frac{1 - 1/T_{br}}{T_{br}^7 - 1} \quad (2.57)$$

$$b = \frac{1 - 1/T_{br}^m}{T_{br}^7 - 1} \quad (2.58)$$

$$h = T_{br} \frac{\lg(P_c/1.01325)}{1 - T_{br}} \quad (2.59)$$

式中 $T_{br} = T_b/T_c$

T_b ——正常沸点, K。

对非极性化合物 (有机物和无机物):

$$\beta = -4.26700 - \frac{221.79}{h^{2.5} \exp(0.0384h^{2.5})} + \frac{3.8126}{\exp(2272.44/h^3)} + \Delta^* \quad (2.60)$$

$$m = 0.78425 \exp(0.089315) - \frac{8.5217}{\exp(0.74826h)} \quad (2.61)$$

式中

对一般物质	$\Delta^* = 0$
对 He	$\Delta^* = 0.41815$
对 H ₂	$\Delta^* = 0.19904$
对 Ne	$\Delta^* = 0.02319$

对不含氢键的极性化合物 (包含氨和乙酸)

$$m = 0.466 T_c^{0.166} \quad (2.62)$$

$$\gamma = 0.08594 \exp(7.462 \times 10^{-4} T_c) \quad (2.63)$$

对含氢键化合物 (水和醇)

$$m = 0.0052 M^{0.29} T_c^{0.72} \quad (2.64)$$

$$\gamma = (2.464/M) \exp(9.8 \times 10^{-6} M T_c) \quad (2.65)$$

M ——相对分子量。

对水和醇

$$\beta = \gamma/b - ah/b \quad (2.66)$$

例 2.6 估算乙苯在 74.1℃ 时的蒸气压, 已知实验值为: 13.332 kPa。

解: 利用 BIPTCHEM 软件包求解, 计算结果为:

$$\text{Lee-Kesler 方程:} \quad 12.640 \text{ kPa}$$

$$\text{Gomez-Thodos 方法:} \quad 12.098 \text{ kPa}$$

误差为:

$$\text{Lee-Kesler 方程:} \quad 5.2\%$$

$$\text{Gomez-Thodos 方法:} \quad 9.2\%$$

2.4 汽化热、比热容和生成热

2.4.1 汽化热

Pitzer 推荐用对应状态方法计算汽化热。在 $0.6 < T_r \leq 1.0$ 时, 其式为:

$$\Delta H_V / RT_c = 7.08(1 - T_r)^{0.354} + 10.95(1 - T_r)^{0.456} \quad (2.67)$$

式中 T_c ——临界温度, K;

T_r ——对比温度, T/T_c ;

R ——气体常数。

例 2.7 对丙醛 ($\text{CH}_3\text{CH}_2\text{CHO}$), 已知 $T_c = 496.0\text{K}$, $\omega = 0.313$ 。计算在温度 $T = 302.69\text{K}$ 的汽化热 ΔH_V , 报道的实验值 29374 J/mol 。

解: 利用 BIPTCHEM 软件求解, 计算结果为:

Pitzer 对比态关联式: 30111.95 J/mol

两参考流体法: 30491.97 J/mol

与实验值比较计算误差为:

Pitzer 对比态关联式: 2.5%

两参考流体法: 3.8%

2.4.2 理想气体比热容

理想气体比热容是温度的函数, 它与压力无关 ($P^0 \rightarrow 0$)。用来关联 C_p^0 和 T 关系的解热式有许多种, 其中最常用如下的简单多项式:

$$C_p^0 = a + bT + cT^2 + dT^3 \quad (2.68)$$

式中 C_p^0 ——J/mol·K。

常见物质的参数可从 BIPTCHEM 查得。对混合物比热容 ($C_{p,\text{mix}}^0$) 按摩尔分数平均进行估算。

若遇软件中没有 C_p^0 参数值, 可借助化合物的分子结构进行推算。相应方法可见文献 [3]。

2.4.3 生成热

确定反应热需要理想气体的标准生成热 $\Delta H_{f,298}^0$, Demalski 对此进行了文献综述, 并推荐了 700 种有机物的 $\Delta H_{f,298}^0$ 值。从 BIPTCHEM 中查得 $\Delta H_{f,298}^0$, 其单位是 J/mol。

利用查到 $\Delta H_{f,298}^0$, 对其他温度下的取值可由下式计算:

$$\Delta H_{f,T}^0 = \Delta H_{f,298}^0 + \int_{298}^T C_p^0 dT \quad (2.69)$$

C_p^0 的确定方法前面已讨论过, 若遇软件中没有 C_p^0 参数值可借助基团贡献方法进行推算, 推算方法见文献 [3]。

2.4.4 液体比热容

对纯液体 Rowlinson 和 Bondi 推荐的关联式为:

$$\frac{C_{pL} - C_p^0}{R} = 1.45 + 0.45(1 - T_r)^{-1} + 0.25\omega \\ [17.11 + 25.2(1 - T_r)^{1/3} T_r^{-1} + 1.742(1 - T_r)^{-1}] \quad (2.70)$$

式中 C_{pL} ——温度 T 下液体比热容, J/mol·K;

C_p^0 ——温度 T 下理想气体比热容, J/mol·K;

ω ——偏心因子;

T_r ——对比温度。

对有机液体混合物, 通常用其纯组分比热容的摩尔分数平均值, 这种方法忽略混合热, 即:

$$C_{pm} = \sum_i C_{pi} x_i \quad (2.71)$$

式中 C_{pm} ——混合液体比热容, J/mol·K;

C_{pi} ——纯组分 i 的比热容, J/mol·K;

x_i —— i 组分的摩尔分数。

2.5 焓的计算

工程中计算焓的常用方法是采用温度的多项式计算出理想气体的焓, 然后选用适当的状态方程计算出实际流体和理想气体的焓差, 这样即可求得实际流体的焓。

2.5.1 焓差的计算

由热力学关系式可导出实际流体和理想气体的焓差:

$$H - H^0 = \int_{V_0}^V \left[T \left(\frac{\partial P}{\partial T} \right)_V - P \right] dV + PV - RT \quad (2.72)$$

式中 H ——实际流体的摩尔焓, J/mol;

H^0 ——理想气体的摩尔焓, J/mol;

V_0 —— $P \rightarrow 0$ 时 (理想气体) 的摩尔体积, m^3/mol ;

V ——实际流体的摩尔体积, m^3/mol 。

将流体的 PVT 性质代入式 (2.72), 即可导出焓差的计算式。下面给出几个重要方程的焓差计算式。

1. RK 方程

RK 方程导得出焓差计算式为:

$$\frac{H - H^0}{RT} = Z - 1 - \frac{3}{2} \frac{A}{B} \ln \left(1 + \frac{B}{Z} \right) \quad (2.73)$$

式中 Z 为压缩因子, 由 RK 方程求解; A 、 B 为 RK 方程的常数。式 (2.73) 只适用于气体焓差的计算, 用于液体焓差的计算时将有很大的误差。

2. SRK 方程

SRK 方程导出的焓差计算式为:

$$\frac{H - H^0}{RT} = Z - 1 - \frac{A(1+E)}{B} \ln \left(1 + \frac{B}{Z} \right) \quad (2.74)$$

$$E = -\frac{1}{A} \sum_{i=1}^c x_i m_i \left(0.42748 \frac{P_{ri}}{T_{ri}} \right)^{0.5} \quad (2.75)$$

式中 Z 为压缩因子, A 、 B 、 m_i 为 SRK 方程的常数, 按 SRK 方程的相应式计算。

由 SRK 方程导出的焓差计算式 (2.74) 既能用于气体, 也可用于液体。

3. PR 方程

PR 方程导出的焓差计算式为:

$$\frac{H - H^0}{RT} = Z - 1 - \frac{A(1+E)}{2\sqrt{2}B} \ln \frac{Z + 2.4142B}{Z - 0.4142B} \quad (2.76)$$

$$E = \frac{1}{A} \sum_{i=1}^c x_i m_i \left(0.45724 \frac{P_{ri}}{T_{ri}} \right)^{0.5} \quad (2.77)$$

式中压缩因子 Z 及常数 A 、 B 、 m_i 由 PR 方程的相应式计算。

PR 方程的焓差计算式 (2.76) 的计算精度和 SRK 方程的焓差计算式 (2.74) 相近。

4. LK 方程

LK 方程计算焓差的关系式为:

$$\frac{H - H^0}{RT_c} = \left(\frac{H - H^0}{RT_c} \right)^{(0)} + \frac{\omega}{\omega^{(R)}} \left[\left(\frac{H - H^0}{RT_c} \right)^{(R)} - \left(\frac{H - H^0}{RT_c} \right)^{(0)} \right] \quad (2.78)$$

对于简单流体, 其焓差按式 (2.79) 计算; 参照流体的焓差仍按式 (2.79) 计算, 但式中常数换成参照流体的常数。参照流体的偏心因子 $\omega^{(R)} = 0.3978$ 。

$$\left(\frac{H - H^0}{RT_c} \right)^{(0)} = T_r \left[Z^{(0)} - 1 - \frac{b_2 + \frac{2b_3}{T_r} + \frac{3b_3}{T_r^2}}{T_r V_r^{(0)}} - \frac{c_2 - \frac{3c_3}{T_r^2}}{2T_r (V_r^{(0)})^2} + \frac{d_2}{5T_r (V_r^{(0)})^5} + E \right] \quad (2.79)$$

$$E = \frac{c_4}{2T_r^3\gamma} \left\{ \beta + 1 - \left[\beta + 1 + \frac{\gamma}{(V_r^{(0)})^2} \right] \exp \left[-\frac{\gamma}{(V_r^{(0)})^2} \right] \right\} \quad (2.80)$$

Lee-Kesler 计算焓差的方法是目前比较准确的方法。

例 2.8 试用 SRK 方程分别计算含丙烯摩尔分数为 50% 的丙烯-丙烷混合物在 313K、1600 kPa 时的气相焓差和液相焓差。

解：运行 BIPTCHEM 中程序 P2（见附录），根据屏幕提示输入组分数、各组分序号及摩尔分数、温度和压力，选择气相计算，屏幕显示气相焓差为 -1954 kJ/kmol。

重新运行该程序，根据屏幕提示输入有关数据，选择液相计算，屏幕显示液相焓差为 -15122 kJ/kmol。

2.5.2 理想气体焓

理想气体混合物的摩尔焓可写成

$$H^0 = \sum_{i=1}^C y_i H_i^0 \quad (2.81)$$

式中 y_i 为组分 i 的摩尔分率， H_i^0 为组分 i 的理想气体摩尔焓。理想气体的摩尔焓 H_i^0 可用温度的多项式表示，即

$$H_i^0 = a_i + b_i T + c_i T^2 + d_i T^3 + e_i T^4 + f_i T^5 \quad (2.82)$$

式中 $a_i, b_i, c_i, d_i, e_i, f_i$ 为常数。

对于轻烃气体， H_i^0 的焓也可由下式计算

$$H_i^0 = 0.55566 (a_i + b_i T + c_i T^2 + d_i T^3) \quad (2.83)$$

式中

$$\left. \begin{aligned} a_i &= 100M_i + 800 \\ b_i &= 32.15M_i + 186.5 \\ c_i &= 3.708M_i - 26.42 \\ d_i &= -0.0859M_i + 40 \\ e &= 0.081T - 4.6 \end{aligned} \right\} \quad (2.84)$$

式中 M_i 为相对分子量， T 为温度。

第3章 流体的传递性质

在非常低的压力下,粘度和导热系数是温度的递增线性函数。温度恒定时,粘度和导热系数随密度的增加而增加。低压下粘度和导热系数的分子理论是成熟的,如 Chapman-Enskog-Brokaw 和 Chung-Lee-Starling 低压气体粘度模型,Stiel-Thodos 低压导热系数模型。压力(或密度)的影响常用剩余性质模型说明。另一类模型是利用分子参数和状态变量计算高压性质。如 Chung-Lee-Starling (Chung) 模型使用临界参数、偏心因子和偶极矩作分子参数,温度和密度作状态变量。

液体的粘度和导热系数常用经验式和关联式描述。在实验数据的温度和压力范围内,它们是准确的。对比态模型可同时描述气体和液体。虽然其精度比关联式低,但是它们的温度和压力外推性较好。

3.1 粘度

粘度单位换算:

$$1\text{P} = 1.000 \times 10^3 \text{cP} = 1.000 \times 10^6 \mu\text{P} = 0.1 \text{Pa}\cdot\text{s} \text{ (SI 制)}$$

粘度估算常用 Lucas 方法和 Chung 方法,此时需要知道纯物质的基本参数,这些参数可从 BIPTCHEM 软件包中查询。

3.1.1 低压下纯气体粘度

1. Lucas 方法

$$\eta\zeta = [0.807 T_r^{0.618} - 0.357 \exp(-0.449 T_r) + 0.340 \exp(-4.058 T_r) + 0.018] F_P^0 F_Q^0 \quad (3.1)$$

$$\zeta = 1.76 \times 10^{-8} (T_c / M^3 P_c^4)^{1/6} \quad (3.2)$$

式中 ζ ——对比粘度;

P_c ——临界压力, Pa;

T_c ——临界温度, K;

M ——相对分子量;

η ——粘度, $\text{Pa}\cdot\text{s}$;

$T_r = T / T_c$;

$F_P^0 = 1$

$$0 \leq \mu_r \leq 0.022 \quad (3.3)$$

$F_P^0 = 1 + 30.55 (0.292 - Z_C)^{1.72}$

$$0.022 \leq \mu_r \leq 0.075 \quad (3.4)$$

$$F_P^0 = 1 + 30.55 (0.292)^{1.72} [0.96 + 0.1 (T_r - 0.7)]; \quad 0.075 \leq \mu_r \quad (3.5)$$

$$\mu_r = 52.46 \mu^2 P_c / T_c^2 \quad (3.6)$$

μ ——偶极矩, 得拜 (debye)。

对量子气体 He, H₂, D₂ 按下式计算 F_Q^0 因子:

$$F_Q^0 = 1.22 Q^{0.15} \{1 + 0.00385[(T_r - 12)^2]^{1/M} \text{sgn}(T_r - 12)\} \quad (3.7)$$

其中, 对 He, $Q = 1.38$

H_2 , $Q = 0.76$

D_2 , $Q = 0.52$

此法对非极性物系误差为 0.5% ~ 1.5%, 对极性物系误差为 2% ~ 4%。

2. Chung 和 Reichenberg 法

$$\eta = 4.0785 \times 10^{-6} \frac{F_c (MT)^{1/2}}{V_c^{2/3} \Omega_V} \quad (3.8)$$

式中 η ——气体粘度, $\text{Pa}\cdot\text{s}$;

M ——相对分子量;

Ω_V ——碰撞积分常数;

$$\Omega_V = A(T^*)^{-B} + C[\exp(-DT^*)] + E[\exp(-FT^*)] \quad (3.9)$$

$$T^* = 1.2595 T_V \quad (3.10)$$

$$A = 1.16145 \quad B = 0.14874 \quad C = 0.52487$$

$$D = 0.77320 \quad E = 2.1678 \quad F = 2.43787$$

$$F_c = 1 - 0.2756\omega + 0.059035\mu_c^4 + \kappa$$

$$\mu_c = 131.3\mu / V_c T_c^{1/2}$$

μ ——偶极矩, 得拜 (debye);

κ ——缔合因子。

μ 、 κ 可由 BIPTCHEM 软件数据库查得。

例 3.1 计算正辛烷蒸气在 311K 时的粘度, 实验值为: $5.82 \times 10^{-6} \text{Pa}\cdot\text{s}$ 。

解: 利用 BIPTCHEM 软件包求解, 计算结果为:

$$\text{Chung 等法:} \quad 5.6624 \times 10^{-6} \text{Pa}\cdot\text{s}$$

$$\text{Lucas 方法:} \quad 5.96 \times 10^{-6} \text{Pa}\cdot\text{s}$$

与实验值比较计算误差为:

$$\text{Chung 等法:} \quad 2.7\%$$

$$\text{Lucas 方法:} \quad 2.4\%$$

3.1.2 低压下气体混合物粘度

1. Lucas 方法

在 Lucas 计算纯气体粘度计算式基础上, 借助下列混合规则, 便可计算混合气体性质

$$T_{cm} = \sum_{i=1}^n y_i T_{ci} \quad (3.11)$$

$$P_{cm} = RT_{cm} \frac{\sum_{i=1}^n y_i Z_{ci}}{\sum_{i=1}^n y_i V_{ci}} \quad (3.12)$$

$$M_m = \sum_{i=1}^n y_i M_i \quad (3.13)$$

$$F_{Q_m}^0 = \left(\sum_{i=1}^n y_i F_{Q_i}^0 \right) / A \quad (3.14)$$

$$F_{P_m}^0 = \sum_{i=1}^n y_i F_{P_i}^0 \quad (3.15)$$

$$A = 1 - 0.01 \left[\frac{M_H}{M_L} \right]^{0.87} \quad \frac{M_H}{M_L} > 9, 0.05 < y_H < 0.7 \quad (3.16)$$

$$A = 1$$

其它情况

式中 M_H ——混合物中最高相对分子量组分的相对分子量；

M_L ——混合物中最低相对分子量组分的相对分子量。

用 T_{cm} , P_{cm} , M_m , $F_{Q_m}^0$, $F_{P_m}^0$ 代入 Lucas 计算纯气体粘度计算式, 便可计算混合物的粘度。此种方法的计算误差通常在上 5% 以内。

2. Chung 法

$$\eta_m = \frac{2.669 \times 10^{-6} F_{cm} (M_m T)^{1/2}}{\sigma_m^2 \Omega_V} \quad (3.17)$$

式中 η_m ——混合气粘度, Pa·s;

T ——温度, K;

$$\Omega_V = [A(T_m^0)^{-B}] + C[\exp(-DT_m^0)] + E[\exp(-FT_m^0)] \quad (3.18)$$

$$T_m^0 = T/(\epsilon/\kappa)^m \quad (3.19)$$

$$(\epsilon/\kappa) = \frac{\sum_i \sum_j y_i y_j (\epsilon_{ij}/\kappa) \sigma_{ij}^3}{\sigma_m^3} \quad (3.20)$$

$$\sigma_m^3 = \sum_i \sum_j y_i y_j \sigma_{ij}^3 \quad (3.21)$$

$$M_m = \left[\frac{\sum_i \sum_j y_i y_j (\epsilon_{ij}/\kappa) \sigma_{ij}^2 M_{ij}^{1/2}}{(\epsilon/\kappa)_m \sigma_m^2} \right]^2 \quad (3.22)$$

$$\omega_m = \frac{\sum_i \sum_j y_i y_j \omega_{ij} \sigma_{ij}^3}{\sigma_m^3} \quad (3.23)$$

$$\mu_m^4 = \sigma_m^3 \sum_i \sum_j \frac{y_i y_j \mu_{ij}^2 \mu_{ij}^2}{\sigma_{ij}^3} \quad (3.24)$$

$$K_m = \sum_i \sum_j y_i y_j k_{ij} \quad (3.25)$$

$$F_{cm} = 1 - 0.275 \omega_m + 0.059035 \mu_m^4 + K_m \quad (3.26)$$

$$\mu_{cm} = 131.3 \mu_m / (V_{cm} T_{cm})^{1/2} \quad (3.27)$$

$$V_{cm} = (\sigma_m / 0.809)^3 \quad (3.28)$$

$$T_m = 1.2593 (\epsilon/R)^m \quad (3.29)$$

混合规则为:

$$\sigma_{ii} = \sigma_i = 0.809 V_{ci}^{1/3} \quad (3.30)$$

$$\sigma_{ij} = \zeta_{ij} (\sigma_i \sigma_j)^{1/2} \quad (3.31)$$

$$\frac{\epsilon_{ij}}{\kappa} = \frac{\epsilon_i}{\kappa} = \frac{T_{ci}}{1.2593} \quad (3.32)$$

$$\frac{\epsilon_{ij}}{\kappa} = \zeta_{ij} \left(\frac{\epsilon_i}{\kappa} \frac{\epsilon_j}{\kappa} \right)^{1/2} \quad (3.33)$$

$$\omega_{ij} = \omega_i \quad (3.34)$$

$$\omega_{ij} = \frac{1}{2}(\omega_i + \omega_j) \quad (3.35)$$

$$\kappa_{ii} = \kappa_i \quad (3.36)$$

$$K_{ij} = (K_i K_j)^{1/2} \quad (3.37)$$

$$M_{ij} = \frac{2M_i M_j}{M_i + M_j} \quad (3.38)$$

此方法应用时误差常常小于 5%。

例 3.2 计算甲烷和正丁烷在 293.2K 时的粘度，其中正丁烷的摩尔分数为 0.303。实验值为 $9.335 \times 10^{-6} \text{Pa}\cdot\text{s}$ 。

解：利用 BIPTCHEM 软件包求解，计算结果为：

甲烷-正丁烷气体粘度

Lucas 法： $9.826 \times 10^{-6} \text{Pa}\cdot\text{s}$

Chung 法： $9.386 \times 10^{-6} \text{Pa}\cdot\text{s}$

与实验值比较计算误差：

Lucas 法：5.3%

Chung 法：0.5%

3.1.3 高压下纯气体粘度

1. Lucas 方法

$$\eta = Z_2 F_P F_Q / \zeta \quad (3.39)$$

式中各项计算如下：

$$\zeta = 0.176 (T_c / M^3 P_c^4)^{1/6} \quad (3.40)$$

当 $T_r \leq 1.0$, $P_r < P_s / P_c$

$$Z_2 = 0.600 + 0.760 P_r^\alpha + (6.990 P_r^\beta - 0.6)(1 - T_r) \quad (3.41)$$

$$\alpha = 3.262 + 14.98 P_r^{5.508} \quad (3.42)$$

$$\beta = 1.390 + 5.746 P_r \quad (3.43)$$

当 $1 < T_r < 40$, $0 < P_r \leq 100$

$$Z_2 = Z_1 \left[1 + \frac{a P_r^\alpha}{b P_r^\beta + (1 + c P_r^d)^{-1}} \right] \quad (3.44)$$

$$Z_1 = [0.807 T_r^{0.618} - 0.357 \exp(-0.449 T_r) + 0.340 \exp(-4.058 T_r) + 0.018] F_P^0 F_Q^0 \quad (3.45)$$

$Z_1 = \eta^0 \zeta$, η^0 是低压气体粘度。

$$a = \frac{a_1}{T_r} \exp a_2 T_r^\gamma \quad (3.46)$$

$$b = a(b_1 T_r - b_2) \quad (3.47)$$

$$c = \frac{c_1}{T_r} \exp(c_2 T_r^3) \quad (3.48)$$

$$d = \frac{d_1}{T_r} \exp(d_2 T_r) \quad (3.49)$$

$$e = 1.3088$$

$$f = f_1 \exp f_2 T_r^3 \quad (3.50)$$

$$\begin{aligned} a_1 &= 1.245 \times 10^{-3} & a_2 &= 5.1726 & \gamma &= -0.3286 \\ b_1 &= 1.6553 & b_2 &= 1.2723 \\ c_1 &= 0.4489 & c_2 &= 3.0578 & \delta &= -37.7332 \\ d_1 &= 1.7368 & d_2 &= 2.2310 & \varepsilon &= -7.6351 \\ f_1 &= 0.9425 & f_2 &= -0.1853 & \zeta &= 0.4489 \end{aligned}$$

$$Y = \frac{Z_2}{Z_1} \quad (3.51)$$

$$F_P = \frac{1 + (F_P^0 - 1) Y^{-3}}{F_P^0} \quad (3.52)$$

$$F_Q = \frac{1 + (F_Q^0 - 1) [Y^{-1} - (0.007)(\ln Y)^4]}{F_Q^0} \quad (3.53)$$

式中 F_P^0 , F_Q^0 ——常数计算方法同纯气体低压下粘度中的方法。

此式的计算对极性和非极性化合物的误差小于 5%。

2. Chung 法

该方法是推算低压气体粘度 Chung 方法的扩展，其关联式为：

$$\mu = \mu^* \frac{36.344(MT_c)^{1/2}}{V_c^{2/3}} \quad (3.54)$$

$$\mu^* = \frac{(T^*)^{1/2}}{\Omega_V} \{ F_c [(G_2)^{-1} + E_6 y] \} + \mu^{**} \quad (3.55)$$

$$y = \frac{\rho V_c}{6} \quad (3.56)$$

$$G_1 = \frac{1 - 0.5y}{(1 - y)^3} \quad (3.57)$$

$$G_2 = \frac{E_1 \{ [1 - \exp(-E_4 y)] / y \} + E_2 G_1 \exp(E_5 y) + E_3 G_1}{E_1 E_4 + E_2 + E_3} \quad (3.58)$$

$$\mu^{**} = E_7 y^2 G_2 \exp[E_8 + E_9 (T^*)^{-1} + E_{10} (T^*)^{-2}] \quad (3.59)$$

Ω_V 和 T^* 计算方法和低压气体粘度 Chung 方法相同。

3.1.4 高压下气体混合物粘度

1. Lucas 方法

高压下气体混合物粘度的基本计算式与高压下纯气体的粘度计算式相同，但其 T_c , P_c , F_P^0 , F_Q^0 和 M 是用低压下气体混合物粘度的混合规则， T_m , P_m , M_m , F_{Qm}^0 , F_{Pm}^0 校正。该方法对极性和非极性化合物的计算误差为 8%~9%。

2. Chung 方法

高压下气体混合物粘度的基本计算式与高压下纯气体的粘度计算式相同, 但参数 T_c , V_c , ω , M , μ 和 κ 是用低压下气体混合流体粘度的混合规则 T_{cm} , V_{cm} , ω_m , M_m , μ_m 和 κ_m 校正。

例 3.3 计算 423K 和 120atm 下含乙烯摩尔分数为 18.65% 的乙烯-乙烷混合物的粘度, 实验值为: $1.882 \times 10^{-5} \text{Pa}\cdot\text{s}$ 。

解: 利用 BIPTCHEM 软件包求解, 计算结果为:

$$\text{Chung 法: } 1.9064 \times 10^{-5} \text{Pa}\cdot\text{s}$$

$$\text{Lucas 法: } 1.9297 \times 10^{-5} \text{Pa}\cdot\text{s}$$

与实验值比较计算误差为:

$$\text{Chung 法: } 1.3\%$$

$$\text{Lucas 法: } 2.5\%$$

3.1.5 液体粘度

1. 纯物质液体粘度

液体粘度随温度的升高而降低。在对比温度大约低于 0.75 时, 可给出一个相当准确的关联式:

$$\lg \eta_L = A + B/(T + C) \quad (3.60)$$

式中 T 为流体的温度 (K), 常数 A 、 B 、 C 可由液体粘度 η_L 实验数据回归得出。其中 A 、 B 对分子结构敏感。若将 C 置为 0, 此式简化为:

$$\lg \eta_L = A + B/T \quad (3.61)$$

式中粘度 η_L 的单位是 cP ($1 \text{ cP} = 0.001 \text{Pa}\cdot\text{s}$)。其他计算液体粘度的经验关联式有:

$$\eta_L = AT^B \quad (3.62)$$

$$\lg \eta_L = A + B/T + CT + DT^2 \quad (3.63)$$

$$\lg \eta_L = A \left\{ \frac{1}{T} - \frac{1}{b} \right\} \quad (3.64)$$

在高温下的液体, 可用下式作为近似估算液体粘度的方法:

$$\eta_L \zeta = (\eta_L \zeta)^0 + \omega (\eta_L \zeta)^1 \quad (3.65)$$

$$(\eta_L \zeta)^0 = (2.648 \times 10^{-6} - 3.725 \times 10^{-6} T_r + 1.309 \times 10^{-6} T_r^2) \quad (3.66)$$

$$(\eta_L \zeta)^1 = (7.425 \times 10^{-6} - 13.39 \times 10^{-6} T_r + 5.933 \times 10^{-6} T_r^2) \quad (3.67)$$

$$\zeta = 0.176 (T_c/M^3 P_c^4)^{1/6} \quad (3.68)$$

式中 η_L ——粘度, $\text{Pa}\cdot\text{s}$;

P_c ——临界压力, Pa;

T_c ——临界温度, K;

M ——相对分子量;

$T_r = T/T_{c0}$

此式仅在 $0.75 < T_r < 0.98$ 范围内适用。

在对比温度较低时, 液体结构更加特别, 没有简单、可靠的方法能估算液体的粘度。

例 3.4 计算苯在 383K 时的粘度, 实验值为: $2.49 \times 10^{-4} \text{Pa}\cdot\text{s}$ 。

解: 利用 BIPTCHEM 软件包求解, 计算结果为:

Przedziedzki - Sridhar 法: $2.27 \times 10^{-4} \text{Pa}\cdot\text{s}$

$\mu = \exp(a + b/T + c/T + dT^2)$ 方法: $2.404 \times 10^{-4} \text{Pa}\cdot\text{s}$

与实验值比较计算误差:

Przedziedzki - Sridhar 法: 8.8%

$\mu = \exp(a + b/T + c/T + dT^2)$ 方法: 3.45%

2. 液体混合物粘度

一般条件下, 可用立方根加和规律近似计算液体混合物粘度。

$$\mu_m^{1/3} = \sum x_i \mu_i^{1/3} \quad (3.69)$$

式中 μ_m ——液体混合物粘度;

μ_i ——纯组分 i 的粘度;

x_i —— i 组分的摩尔分数。

此式的平均误差为 7.6%, 各组分性质相近时精确性较好。

例 3.5 计算苯-环己烷混合物在 298K 时的粘度, 其中苯的摩尔分数为 0.6, 实验值为: $0.593 \times 10^{-3} \text{Pa}\cdot\text{s}$ 。

解: 利用 BIPTCHEM 软件求解, 计算结果: $0.7098 \times 10^{-3} \text{Pa}\cdot\text{s}$, 与实验值比较计算误差: 16.5%。

3.2 导热系数

3.2.1 低压下气体导热系数

对低压下烃类气体, Chung 等人给出了如下关联式。

1. Chung 方法

$$\frac{\lambda M}{\eta C_V} = \frac{3.75\varphi}{C_V/R} \quad (3.70)$$

式中 λ ——导热系数, $\text{W}/(\text{m}\cdot\text{K})$;

M ——分子量;

η ——低压气体粘度, $\text{Pa}\cdot\text{s}$;

C_V ——理想气体恒容比热容, $\text{J}/(\text{mol}\cdot\text{K})$;

R ——气体常数。

$$\varphi = 1 + \alpha \{ [0.215 + 0.28288\alpha - 1.061\beta + 0.26665Z] / [0.6366 + \beta Z + 1.061\alpha\beta] \} \quad (3.71)$$

$$\alpha = (C_V/R) - 1.5 \quad (3.72)$$

$$\beta = 0.7862 - 0.7109\omega + 1.3168\omega^2 \quad (3.73)$$

$$Z = 2.0 + 10.5T_c^{-1} \quad (3.74)$$

此关系式只适用于非极性物质, 对极性物质每种物质的 β 均有特定值。Chung 等仅给出了几种物质的 β 值, 对无 β 值的物质, 可用常数 0.758 代替。应用该方法的误差一般不超

过 5%~7%。

2. Eucken 方法

$$\frac{\lambda M'}{\eta C_V} = 1 + \frac{9/4}{(C_P/R) - 1} \quad (3.75)$$

式中 λ ——导热系数, W/(m·K);

M' ——相对分子量;

C_V ——摩尔定容热容, J/(mol·K);

η ——低压气体粘度, Pa·s;

C_P ——摩尔定压热容, J/(mol·K);

R ——气体常数。

3. 修正 Eucken 方法

$$\frac{\lambda M'}{\eta C_V} = 1.15 + \frac{2.03}{(C_P/R) - 1} \quad (3.76)$$

式中符号含义同上式。

4. Stiel-Thodos 方法

对非极性气体

$$\frac{\lambda M'}{\eta C_V} = 1.15 + \frac{2.03}{(C_V/R)^{-1}} \quad (3.77)$$

式中符号含义同前。

例 3.6 试估算 2-甲基丁烷 (异戊烷) 蒸气在 0.1 MPa 和 373 K 时的导热系数。文献报导值为 2.2×10^{-2} W/(m·K)。

解:

利用 BIPTCHEM 软件包求解, 计算结果为:

Chung 法: 2.30E-2 W/(m·K)

Eucken 法: 1.88E-2 W/(m·K)

修正 Eucken 法: 2.36E-2 W/(m·K)

Stiel-Thodos 法: 2.10E-2 W/(m·K)

与文献值比较误差:

Chung 法: 4.5%

Eucken 法: 14.5%

修正 Eucken 法: 7.3%

Stiel-Thodos 法: 4.5%

3.2.2 高压下气体导热系数

所有气体的导热系数是随着压力的增加而增加的, 在 0.1MPa 到 1MPa 的范围内, 这种压力影响可忽略。在低于 0.01MPa 后, 几乎正比于压力。在非常接近临界点处 λ_G 急剧增长到一最大值。Stiel 和 Thodos 推荐式如下。

对非极性物质:

$$\begin{aligned} (\lambda - \lambda^0) T Z_c^S &= (1.22 \times 10^{-2}) [\exp(0.535 \rho_r) - 1] \quad \rho_r < 0.5 \\ &= 1.14 \times 10^{-2} [\exp(0.67 \rho_r) - 1.069] \quad 0.5 \leq \rho_r \leq 2.0 \end{aligned} \quad (3.78)$$

$$= 2.60 \times 10^{-3} [\exp(1.155\rho_r) + 2.016] \quad 2.0 \leq \rho_r \leq 2.8$$

式中 λ ——高压下导热系数, W/(m·K);

λ^0 ——低压下(0.1MPa左右)导热系数, W/(m·K);

$$\Gamma = 210 \left(\frac{T_c M^3}{P_c^4} \right)^{1/6} \quad (3.79)$$

Z_c ——临界压缩因子;

ρ_r ——相对密度, ρ/ρ_c 。

此关联式的误差一般为 10% ~ 20%, 并且不能用于极性物质或氢气、氦气, 对这些物质没有通用的计算方法可以利用。

3.2.3 低压下气体混合物导热系数

低压下气体混合物的导热系数可由下面的关系式计算:

$$\lambda_m = \frac{\sum_{i=1}^n y_i \lambda_i}{\sum_{j=1}^n y_j A_{ij}} \quad (3.80)$$

式中 λ_m ——低压下气体混合物导热系数;

λ_i ——纯组分 i 低压下导热系数;

y_i ——组分 i 摩尔分数。

参数 A_{ij} 可由以下几种方法来估算。

对于非极性, 或非极性~极性二元气体, 可用 Mason-Saxena 计算式。

Mason-Saxena 计算式

$$A_{ij} = \frac{[1 + (g_i/g_j)^{-1/2} (M_i/M_j)^{1/4}]^2}{[8(1 + M_i/M_j)]^{1/2}} \quad (3.81)$$

其中

$$\frac{g_i}{g_j} = \left(\frac{\Gamma_i}{\Gamma_j} \right) = \left[\frac{\exp(0.046 T_{ri}) - \exp(-0.2412 T_{ri})}{\exp(0.0464 T_{rj}) - \exp(-0.2412 T_{rj})} \right] \quad (3.82)$$

式中 M ——相对分子量

$$\Gamma = 210 \left(\frac{T_c M^3}{P_c^4} \right)^{1/6} \quad (3.83)$$

对极性和非极性气体混合物, 用 Lindsay 和 Bromley 计算式。

Lindsay 和 Bromley 计算式

$$A_{ij} = \frac{1}{4} \left[1 + \left[\frac{\eta_i}{\eta_j} \left(\frac{M_i}{M_j} \right)^{3/4} \left(\frac{T + S_i}{T + S_j} \right) \right]^{1/2} \right]^2 \left(\frac{T + S_{ij}}{T + S_i} \right) \quad (3.84)$$

式中 η_i ——低压下纯组分 i 粘度;

S_i —— $1.5T_{bi}$;

$S_{ij} = S_{ji} = C(S_i S_j)^{1/2}$ 。

如果 i, j 中的一个或两个是强极性物, $C \approx 1$, 否则 C 取 0.73。 T, T_b (正常沸点) 取绝对温度 K。以上讨论的推算 λ_{mix} 方法的精确度一般在 2% ~ 5% 以内。此方法的适应性是它具有表达最大值和最小值的能力。

例 3.7 计算在 0.1MPa 和 368K 时甲烷的摩尔分数为 0.468 的甲烷-丙烷混合物的导热

系数, 实验值为: $3.20\text{E}-2 \text{ W}/(\text{m}\cdot\text{K})$

解: 利用 BIPTCHEM 软件包求解, 计算结果为:

Chung 法: $3.346\text{E}-2 \text{ W}/(\text{m}\cdot\text{K})$

与实验值比较计算误差: 4.56%。

3.2.4 高压下气体混合物导热系数

压力对混合气体导热系数的影响由前面 Stiel 和 Thodos 推荐式计算, 混合物的临界性质可由下面方程给出的混合规则来计算。

$$T_{\text{cm}} = \sum_{i=1}^n y_i T_{ci} \quad (3.85)$$

$$Z_{\text{cm}} = \sum_{i=1}^n y_i Z_{ci} \quad (3.86)$$

$$V_{\text{cm}} = \sum_{i=1}^n y_i V_{ci} \quad (3.87)$$

$$P_{\text{cm}} = \frac{Z_{\text{cm}} R T_{\text{cm}}}{V_{\text{cm}}} \quad (3.88)$$

式中 Z_c ——临界压缩因子。

y_i ——组分 i 的摩尔分数。

此种方法经实验验证, 其误差一般小于 10%。

例 3.8 计算在 17.48MPa 和 370.8K 时甲烷-二氧化碳混合物 (含甲烷摩尔分数为 0.755) 的导热系数, 实验值为: $5.08\text{E}-2 \text{ W}/(\text{m}\cdot\text{K})$

解: 利用 BIPTCHEM 软件包求解, 计算结果为: $5.59\text{E}-2 \text{ W}/(\text{m}\cdot\text{K})$

与实验值比较计算误差: 10%

3.2.5 液体导热系数

在正常沸点以下, 大多数常见有机液体的导热系数在 $0.1 \sim 0.17 \text{ W}/(\text{m}\cdot\text{K})$ 之间。但是像水和氨这类强极性分子物质的导热系数是上述值的 2~3 倍。温度对导热系数的影响不大, 一般情况下导热系数的值随温度增加而减小。常见液体的导热系数可在文献 [1] 和 [2] 中查得。液体导热系数的估算方法有关联式方法和 Latini 方法两种。

1. 关联式方法

对许多物质, BIPTCHEM 使用了有实验基础的经验关联式, 它们包括:

$$\lambda_L = a - \frac{b}{T} \quad (3.89)$$

$$\lambda_L = a + bT + cT^2 \quad (3.90)$$

式中 a 、 b 和 c 为关联常数, 可由 BIPTCHEM 软件包查询。

2. Latini 方程

$$\lambda_L = \frac{A(1 - T_r)^{0.38}}{T_r^{1/6}} \quad (3.91)$$

式中 λ_L ——液体导热系数, $\text{W}/(\text{m}\cdot\text{K})$;

$$A = \frac{A^* T_c^\gamma}{M^\beta T_c^\alpha} \quad (3.92)$$

$T_r = T/T_c$ ——对比温度；

T_b ——0.1MPa 下正常沸点温度，K；

M ——相对分子量（ $50 < M < 250$ 内应用）；

A^* 、 α 、 β 、 γ ——常数，由文献 [5] 中表 9-5 查得。

例 3.9 计算 CCl_4 在 293K 时的导热系数，实验值为：0.103 W/(m·K)。

解：利用 BIPTCHEM 软件包求解，计算结果为：

Sato-Riedel 法：0.1017 W/(m·K)

Tcon = a + b(T - c)法：0.1025 W/(m·K)

与实验值比较计算误差：

Sato-Riedel 法：1.26%

Tcon = a + b(T - c)法：0.45%

3.2.6 液体混合物导热系数

利用纯液体的导热系数值和下面的计算式，可计算液体混合物的导热系数：

$$\lambda_{Lm} = \sum_{i=1}^n \sum_{j=1}^n \Psi_i \Psi_j \lambda_{ij} \quad (3.93)$$

$$\lambda_{ij} = 2(\lambda_{ij}^{(1)} + \lambda_{ij}^{(2)})^{-1} \quad (3.94)$$

$$\Psi_i = \frac{x_i V_i}{\sum_{k=1}^n x_k V_k} \quad (3.95)$$

式中 x_i ——组分 i 的摩尔分数；

V_i ——组分 i 的摩尔体积；

$\lambda_{ij}^{(1)}$ ——纯组分 i 的导热系数。

此方法的误差对含水和不含水的混合物都很少超过 3%~4%。

例 3.10 计算甲醇和苯液体混合物在 273K 时的导热系数，甲醇的质量分数为 0.40。实验值为：0.170 W/(m·K)。

解：利用 BIPTCHEM 软件包求解，计算结果为：0.1569 W/(m·K)

与实验值比较计算误差：7.66%

第 4 章 石油和石油馏分物性计算

石油和石油馏分是化学组成非常复杂的流体，其中的组分数多到无从确知其详细化学成分的程 度。石油馏分的性质包括平均沸点、相对分子量、临界参数等。此处仅介绍一些基本性质^[8]。

1. 平均沸点

石油流体的体积平均沸点 t_V 通常根据恩氏蒸馏数据计算。

$$t_V = \frac{t_{10} + t_{30} + t_{50} + t_{70} + t_{90}}{5} \quad (4.1)$$

式中 t_V ——体积平均沸点, $^{\circ}\text{C}$;

t_{10} 、 t_{30} 、 t_{50} 、 t_{70} 、 t_{90} ——恩氏蒸馏 10%、30%、50%、70% 和 90% 点的温度, $^{\circ}\text{C}$ 。

其他平均沸点用下式计算。

重量平均沸点

$$t_W = t_V + \Delta_1 \quad (4.2)$$

$$\lg \Delta_1 = -3.64991 - 0.027060 t_V^{0.6667} + 5.16388 S^{0.25} \quad (4.3)$$

分子平均沸点

$$t_M = t_V - \Delta_2 \quad (4.4)$$

$$\lg \Delta_2 = -1.15158 - 0.011810 t_V^{0.6667} + 3.70684 S^{0.3333} \quad (4.5)$$

立方平均沸点

$$t_{CU} = t_V - \Delta_3 \quad (4.6)$$

中平均沸点

$$\lg \Delta_3 = -0.82368 - 0.089970 t_V^{0.45} + 2.45679 S^{0.45}$$

$$t_{ME} = \frac{t_M + t_{CU}}{2} = t_V - \Delta_4 \quad (4.7)$$

$$\lg \Delta_4 = -1.53181 - 0.012800 t_V^{0.6667} + 3.64678 S^{0.3333} \quad (4.8)$$

式中 t_W 、 Δ_1 ——重量平均沸点及其校正值, $^{\circ}\text{C}$;

t_M 、 Δ_2 ——分子平均沸点及其校正值, $^{\circ}\text{C}$;

t_{CU} 、 Δ_3 ——立方平均沸点及其校正值, $^{\circ}\text{C}$;

t_{ME} 、 Δ_4 ——中平均沸点及其校正值, $^{\circ}\text{C}$;

S ——恩氏蒸馏斜率, $^{\circ}\text{C}/\%$ 。

对沸程小于 30°C 的窄馏分, 可认为其各种平均沸点接近相等, 并以中沸点代替而不致引起大的误差。

2. 相对分子量

石油馏分的相对分子量可由下式计算, 该式为 Winn 列线图的拟合表达式。

$$\log M = \sum \sum A_{ij} (1.8 t_{ME} + 32)^i K_W^j \quad (4.9)$$

式中 K_W ——特性因数;

A_{ij} ——拟合参数, 见文献 [8]。

3. 临界温度

石油馏分的临界温度可按式 (Roess 式) 计算:

$$t_c = 85.66 + 0.9259D - 0.3959 \times 10^{-3}D^2 \quad (4.10)$$

$$D = d_{15.6}^{15.6} \times (1.8t_v + 132) \quad (4.11)$$

式中 t_c ——临界温度, $^{\circ}\text{C}$;

$d_{15.6}^{15.6}$ ——相对密度 ($15.6^{\circ}\text{C}:15.6^{\circ}\text{C}$)。

4. 偏心因子

石油馏分的偏心因子用 Edmister 式计算:

$$\omega = \frac{3}{7} \left(\frac{\log P_c}{T_c/T_B - 1} \right) - 1 \quad (4.12)$$

式中 ω ——偏心因子;

P_c ——临界压力;

T_B ——对宽馏分用立方平均沸点, 对窄馏分用中沸点, K。

5. 临界压力

石油馏分 ($S > 0.4$) 的临界压力可由下式估算:

$$\log P_c = -4.09457 - 2.04237 \lg M + 2.78765 \lg T_c + 0.65541 \lg \omega \quad (4.13)$$

式中 M ——相对分子量;

T_c ——临界温度, K。

6. 临界体积

Hall 和 Yarborough 建议用下式计算石油馏分的临界体积:

$$V_c = 0.0015606 \left(\frac{M}{(d_{15.6}^{15.6})^{0.69}} \right) \quad (4.14)$$

式中 V_c ——临界体积, m^3/mol 。

7. 汽化热

常压沸点时的汽化热 ΔH_{VB} 由下式计算:

$$\Delta H_{VB} = 4.18(8.75 + 4.571 \log T_B) T_B \quad (4.15)$$

25°C 时的汽化热 ΔH_{VN} 可按式计算:

$$\Delta H_{VN} = 4.18 \Delta H_{VB} \left(\frac{1 - 298.16/T_c}{1 - T_B/T_c} \right)^{0.38} \quad (4.16)$$

ΔH_{VN} , ΔH_{VB} ——常压沸点和 25°C 时的汽化热, J/mol。

其他石油流体物性的计算方法请见文献 [9、10]。

利用 BIPTCHEM 可计算石油流体的基本物性, 计算原油的物性适用于石蜡级原油 ($UOPK > 12$) 和常减压蒸馏塔的馏分油, 不适用发生化学变化后的产品油性质的计算, 计算准确程度与“炼油设计手册”相仿。

石油大学 (北京) 郭天民和其合作者开发了计算石油流体性质的计算软件 UP-PACK, 该软件具有全部石油和石油馏分常见物性的计算功能, 经实验数据检验, 计算结果比较可靠。

例 4.1 已知某原油的馏程为:

10%	173.1 K
30%	297.0 K
50%	345.0 K
70%	388.0 K
90%	446.0 K

20℃时的相对密度为 0.8350，原油的温度为 50℃、压力为 0.1 MPa，计算其物性。

解：

启动 BiPTCHEM 软件包后，选择“物性”菜单，按回车，屏幕弹出菜单，

```

基础数据库
估算纯物质性质
估算混合物的性质
石油馏分性质的估算
  
```

选择“石油馏分性质的估算”一项，按回车，屏幕显示：

```

请选择输入数据：
1. 恩氏蒸馏数据和 293K 时的密度（或折射率）
2. 中平均沸点 [K] 和 293K 时的密度（或折射率）
3. 323K 时的运动粘度 [m2/s] 和 293K 时的密度（或折射率）
4. 373K 时的运动粘度 [m2/s] 和 293K 时的密度（或折射率）
5. 立方平均沸点 [K] 和 293K 时的密度（或折射率）
6. 323K 和 373K 时的运动粘度 [m2/s]
Q. 退出
  
```

根据本题已知条件，选择 1，按回车，屏幕显示：

```

请输入温度 [K]:
  
```

根据本题输入 323，屏幕显示：

```

请输入压力 [kPa]:
  
```

根据本题输入 100，屏幕显示：

```

请输入 293K 时的密度（或折射率）:
  
```

根据本题输入 0.8350，屏幕显示：

```

请输入恩氏蒸馏 10%，30%，50%，70%和 90%点的温度 [K]:
  
```

根据本题输入 446.25，570.15，618.15，661.15，719.15。按回车，特别需注意，此处输入数据时需在一行内输入五个数据点，每个数据点之间需键入一个逗号。屏幕显示计算结果如下：

温度 323.15K	压力 100kPa	
293K 时的密度 835kg/m ³	密度指数 API 37.08	特性因数 12.04
临界温度 775.8 K	临界压力 1375.86kPa	偏心因子 0.6128
分子量 253.844	体积平均沸点 603 K	重量平均沸点 611 K
立方平均沸点 594 K	分子平均沸点 555.3 K	中平均沸点 574.7 K
蒸汽压 1.238E-3kPa	平衡常数 1.221E-5	表面张力 0.0295N/m
汽相焓 347.4 kJ/kmol	液相焓 94.1 kJ/kmol	汽化热 253.3kJ/kmol
导热系数 0.136 W/(m·K)	汽相粘度 0.006342Pa·s	
汽相比热容 1.665kJ/(kg·K)	液相比热容 1.991kJ/(kg·K)	
323K 时的运动粘度 4.077E-6m ² /s	373 时的运动粘度 1.756E-6m ² /s	

计算结束。

例 4.2 已知某原油 20℃ 时的相对密度为 0.8017, 中沸点温度为 270.0℃, 原油的温度为 100℃、压力为 0.1MPa, 计算其物性。

解:

启动 RIPTCHEM 软件后, 选择“物性”菜单, 按回车, 屏幕弹出菜单:

```

基础数据库
估算纯物质性质
估算混合物的性质
石油馏分性质的估算

```

选择“石油馏分性质的估算”一项, 按回车, 屏幕显示:

```

请选择输入数据:
1. 恩氏蒸馏数据和 293K 时的密度 (或折射率)
2. 中平均沸点 [K] 和 293K 时的密度 (或折射率)
3. 323K 时的运动粘度 [m2/s] 和 293K 时的密度 (或折射率)
4. 373K 时的运动粘度 [m2/s] 和 293K 时的密度 (或折射率)
5. 立方平均沸点 [K] 和 293K 时的密度 (或折射率)
6. 323K 和 373K 时的运动粘度 [m2/s]
Q. 退出

```

根据本题目已知条件, 选择 2, 按回车。屏幕显示:

```

请输入温度 [K]:

```

根据本题输入 373。屏幕显示:

```

请输入压力 [kPa]:

```

根据本题输入 1。屏幕显示:

```

请输入 293K 时的密度 (或折射率):

```

根据本题输入 0.8017。屏幕显示:

请输入中平均沸点温度 [K]:

根据本题输入 543.15。按回车。屏幕显示出计算结果:

温度 373.15K	压力 100kPa	
293K 时的密度 801.7kg/m ³	密度指数 API 44.01	特性因数 12.31
临界温度 764.9 K	临界压力 1198.95kPa	偏心因子 0.6288
相对分子量 229.5	体积平均沸点 603 K	重量平均沸点 611 K
立方平均沸点 594 K	分子平均沸点 555.3 K	中平均沸点 573.2 K
蒸汽压 2.3017E-1kPa	平衡常数 2.302E-3	表面张力 0.0247N/m
汽相焓 447.5kJ/kmol	液相焓 205.8kJ/kmol	汽化热 241.4kJ/kmol
导热系数 0.138W/(m·K)	汽相粘度 7E-5Pa·s	
汽相比热容 1.594kJ/(kg·K)	液相比热容 2.277kJ/(kg·K)	
323K 时的运动粘度 2.023E-6m ² /s	373 时的运动粘度 1.146E-6m ² /s	

计算结束。

习 题

1. 查询下列物质的纯物质常数。

二氧化硫气, 氩气, 甲烷, 丙酮, 顺丁烯, 丁烯, 1-戊烯和顺-2-戊烯。

2. 估算异丁烷饱和液体在 310.93K、 139.9×10^5 Pa 时的摩尔体积。实验值为 102.7cm³/mol。

3. 估算氮气和氨混合物 (混合摩尔比 1:3), 在 405.3×10^5 Pa, 300℃ 时的摩尔体积。文献值: 压缩因子 $Z_m = 1.1155$ 。

4. 估算饱和液体混合物的体积。

已知液体组成: CO₂ = 50.2% (摩尔分数)

正丁烷: 49.8% (摩尔分数)

计算:

(1) $T = 344.26$ K, $P = 64.3 \times 10^5$ Pa 时, 混合液体积, 实验值为 99.13cm³/mol。

(2) $T = 344.26$ K, $P = 345 \times 10^5$ Pa 时, 混合液体积, 实验值为 78.22cm³/mol。

5. 估算 $T = 349.8$ K, 顺-2-丁烯的液体热容, 实验值为 152.7J/mol·K。

6. 估算 SO₂ 气体在常压下 300℃ 时的粘度, 实验值为 246×10^{-7} Pa·s。

7. 估算甲醇蒸汽在 550K, 1×10^5 Pa 下的粘度, 实验值为 181×10^{-7} Pa·s。

8. 估算醋酸蒸汽在 125℃, 常压下粘度。实验值为 101×10^{-7} Pa·s。

9. 估算 N₂, CHClF₂ 混合气在 50℃ 下的粘度。

已知: 氮气的摩尔分数为 0.286, 50℃ 时, 氮气粘度为 188×10^{-7} Pa·s, CHClF₂ 粘度为 134×10^{-7} Pa·s。实验值为 145×10^{-7} Pa·s。

10. 估算甲烷, 正丁烷在 293K, 常压下的粘度。

已知: 正丁烷摩尔分数为: 0.303。293K 时, 甲烷粘度: 109.4×10^{-7} Pa·s, 正丁烷粘度: 72.7×10^{-7} Pa·s。实验值为: 93.35×10^{-7} Pa·s。

11. 估算正戊烷蒸汽在 500K, 101×10^5 Pa 下的粘度, 实验值为 546×10^{-7} Pa·s。

12. 估算氨气在 420K, 300×10^5 Pa 下的粘度, 实验值为 571×10^{-7} Pa·s。

13. 估算丙酮在常压, 30℃下的粘度, 实验值 $0.292 \times 10^{-3} \text{Pa} \cdot \text{s}$ 。
14. 估算丁醇在常压, 120℃粘度, 实验值为 $0.394 \times 10^{-3} \text{Pa} \cdot \text{s}$ 。
15. 估算液体混合物正己烷, 十六烷在 298K, 常压下粘度。已知正己烷摩尔分数为 0.40, 298K 时, 正己烷粘度: $0.298 \times 10^{-3} \text{Pa} \cdot \text{s}$, 十六烷粘度: $3.078 \times 10^{-3} \text{Pa} \cdot \text{s}$, 实验值为 $1.51 \times 10^{-3} \text{Pa} \cdot \text{s}$ 。
16. 估算乙酸和丙酮液体混合物在常压, 50℃粘度。已知乙酸摩尔分数为: 0.70, 50℃时, 乙酸粘度为: $0.798 \times 10^{-3} \text{Pa} \cdot \text{s}$, 丙酮粘度为: $0.241 \times 10^{-3} \text{Pa} \cdot \text{s}$, 实验值为 $0.587 \times 10^{-3} \text{Pa} \cdot \text{s}$ 。
17. 估算 2-甲基丁烷蒸汽在常压, 100℃时的导热系数。实验值为 $2.2 \text{W}/(\text{m} \cdot \text{K})$ 。
18. 估算醋酸在 184℃, 常压下的导热系数。实验值为 $2.38 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$ 。
19. 估算丙烯在 473K, $150 \times 10^5 \text{Pa}$ 下的导热系数。实验值为 $6.64 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$ 。
20. 估算苯和氩混合气在 100.6℃, 常压下的导热系数。已知苯的摩尔分数为 0.25, 苯的导热系数为: $1.66 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$, 氩的导热系数为: $2.14 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$ 实验值为 $1.92 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$ 。
21. 估算甲烷和 CO_2 混合气在 370.8K, $174.8 \times 10^5 \text{Pa}$ 时的导热系数, 已知甲烷的摩尔分数为: 0.755, 实验值为: $5.08 \times 10^{-2} \text{W}/(\text{m} \cdot \text{K})$ 。
22. 估算叔丁醇在 318K, 常压下的导热系数。实验值为 $0.128 \text{W}/(\text{m} \cdot \text{K})$ 。
23. 估算庚烷在 313K, $490 \times 10^5 \text{Pa}$ 下的导热系数。实验值为 $0.136 \text{W}/(\text{m} \cdot \text{K})$ 。
24. 估算甲醇和苯混合液在 273K 时的导热系数, 已知: 甲醇的摩尔分数为: 0.40, 甲醇的导热系数为: $0.210 \text{W}/(\text{m} \cdot \text{K})$, 苯的导热系数为: $0.152 \text{W}/(\text{m} \cdot \text{K})$ 。实验值为: $0.170 \text{W}/(\text{m} \cdot \text{K})$ 。

参 考 文 献

- 1 马沛生等. 石油化工基础数据手册续编. 化学工业出版社, 1993
- 2 卢焕章等. 石油化工基础数据手册. 化学工业出版社, 1982
- 3 R. C. Reid, J. M. Prausnitz and B. E. Poling. The Properties of Gases and Liquids. 4th Ed. McGraw - Hill, 1987
- 4 ASPEN PLUS®, Reference Manual. Physical Property and Methods and Models. Vol 3, 1993
- 5 Perry, R. H., D. W. Green and J. O. Maloney. Perry's Chemical Engineer's Handbook. 6th ed., McGraw - Hill, 1984
- 6 童景山. 流体的热物理性质. 中国石化出版社, 1996
- 7 王松汉等. 乙烯装置技术. 中国石化出版社, 1994
- 8 郭天民等. 多元汽-液平衡和精馏. 化学工业出版社, 1983
- 9 [美] K. S. 佩德森等著. 郭天民等译. 石油和天然气的性质. 中国石化出版社, 1992
- 10 林世雄等. 石油炼制工程. 石油大学出版社

附录 计算程序

P1 用 SRK 方程计算压缩因子

```

DECLARE FUNCTION fz! (A!, B!, c!, y$)
CLS
INPUT "输入组分数: ", c
DIM Tc(c), Pc(c), W(c), x(c), Tr(c), Pr(c), A(c), B(c)
x(c) = 1
FOR i = 1 TO c
    PRINT "输入组分": ; IF c < > 1 THEN PRINT i;
    PRINT "的:"
    INPUT "临界温度 [K] = ", Tc(i)
    INPUT "临界压力 [kPa] ", Pc(i)
    INPUT "偏心因子 ", W(i)
    IF i < c THEN INPUT "摩尔分数 ", x(i): x(c) = x(c) - x(i)
NEXT
INPUT "输入温度 [K] = ", T: INPUT "输入压力 [kPa] = ", P
PRINT "计算气相或是液相? (V/L) ";
DO: y$ = UCASE$(INPUT$(1)): LOOP WHILE INSTR("VL", y$) = 0
A(0) = 0: B(0) = 0
FOR i = 1 TO c
    Tr(i) = T / Tc(i): Pr(i) = P / Pc(i)
    mi = .48505 + W(i) * (1.55171 - .15613 * W(i))
    IF Tr(i) < 1.8 THEN
        at = (1 + mi * (1 - Tr(i) ^ .5)) ^ 2
    ELSE
        at = EXP(2 * mi * (1 - Tr(i) ^ .5))
    END IF
    A(i) = .42748 * at * Pr(i) / Tr(i) ^ 2
    B(i) = .08664 * Pr(i) / Tr(i)
    A(0) = A(0) + SQR(A(i)) * x(i): B(0) = B(0) + B(i) * x(i)
NEXT i: A(0) = A(0) ^ 2
Z = fz(-1, A(0) - B(0) - B(0) ^ 2, -A(0) * B(0), y$)
PRINT
IF y$ = "V" THEN PRINT "气"; ELSE PRINT "液";
PRINT "相压缩因子为": Z

```

FUNCTION fz (A, B, c, y\$)

```

CONST pi2 = 2 * 3.14159
m = B - A^2 / 3
n = c + 2 * A^3 / 27 - A * B / 3
a0 = n^2 / 4: a1 = m^3 / 27: a2 = a0 + a1
IF a2 > 0 THEN
    a3 = SQR(a2): x1 = -n / 2 + a3: x2 = -n / 2 - a3
    x = SGN(x1) * ABS(x1)^(1 / 3) + SGN(x2) * ABS(x2)^(1 / 3)
ELSE
    a3 = SQR(-a0 / a1)
    IF n > 0 THEN a3 = -a3
    fi = ATN(SQR(1 - a3 * a3) / a3)
    IF y$ = "V" THEN x = -10 ELSE x = 100
    FOR k = 0 TO 2
        zz = 2 * SQR(-m / 3) * COS((fi + pi2 * k) / 3)
        IF y$ = "V" THEN
            IF zz > x THEN x = zz
        ELSE
            IF zz < x THEN x = zz
        END IF
    NEXT k
END IF
fz = x - A / 3
END FUNCTION

```

P2 SRK 方程计算焓差

```

DECLARE FUNCTION fz! (A!, b!, c!, y$)
DECLARE FUNCTION SRKdh! (T!, P!, y$)
TYPE Z
    n AS INTEGER
    f AS STRING* 10
    na AS STRING* 35
END TYPE
TYPE y1
    m AS SINGLE
    Tc AS SINGLE
    Pc AS SINGLE
    Vc AS SINGLE
    Zc AS SINGLE
    w AS SINGLE

```

```

END TYPE
CLS
INPUT "组分数: ", c
DIM id(c), jz AS Z, jy1 AS y1
DIM SHARED Tc(c), Vc(c), Pc(c), w(c), na$(c), x(c)
x(c) = 1
FOR i = 1 TO c:
    PRINT "输入组分"; i; : INPUT "的标号 ", id(i)
NEXT
FOR i = 1 TO c - 1:
    PRINT "输入组分"; i; : INPUT "的摩尔分数 ", x(i)
    x(c) = x(c) - x(i)
NEXT
    OPEN "\ dat \ Y1.dat" FOR RANDOM AS #1 LEN = LEN(jy1)
    FOR i = 1 TO c
        GET #1, id(i), jy1
        Tc(i) = jy1.Tc: Pc(i) = jy1.Pc
        w(i) = jy1.w
    NEXT i
    CLOSE
    OPEN "\ dat \ z.dat" FOR RANDOM AS #1 LEN = LEN(jz)
    FOR i = 1 TO c
        GET #1, id(i), jz
        na$(i) = RTRIM$(jz.na): IF i > 1 THEN na$ = na$ + "-"
        na$ = na$ + RTRIM$(jz.na)
    NEXT i
    CLOSE
    INPUT "输入温度[K] ", T
    INPUT "输入压力[kPa] ", P: P = P / 100
    PRINT "计算气相焓差还是计算液相焓差? (V/L)"
    DO: y$ = UCASE$(INPUT$(1)): LOOP UNTIL y$ = "V" OR y$ = "L"
    dH = SRKdh(T, P, y$)
    CLS
    LOCATE 10, 5: PRINT na$; "溶液在"; T; "K、"; P * 100; "KPa 时";
    IF y$ = "V" THEN PRINT "气"; ELSE PRINT "液";
    PRINT "相焓差 $\Delta H$ 为:"
    LOCATE 12, 20: PRINT dH; "kJ/kmol"

FUNCTION fz(A, b, c, y$)
    CONST pi2 = 2 * 3.14159

```

```

m = b * A^2 / 3
n = c + 2 * A^3 / 27 - A * b / 3
a0 = n^2 / 4: a1 = m^3 / 27: a2 = a0 + a1
IF a2 > 0 THEN
    a3 = SQR(a2): x1 = -n / 2 + a3: x2 = -n / 2 - a3
    x = SGN(x1) * ABS(x1)^(1 / 3) + SGN(x2) * ABS(x2)^(1 / 3)
ELSE
    a3 = SQR(-a0 / a1)
    IF n > 0 THEN a3 = -a3
    fi = ATN(SQR(1 - a3 * a3) / a3)
    IF y$ = "V" THEN x = -10 ELSE x = 100
    FOR k = 0 TO 2
        zz = 2 * SQR(-m / 3) * COS((fi + pi2 * k) / 3)
        IF y$ = "V" THEN
            IF zz > x THEN x = zz
        ELSE
            IF zz < x THEN x = zz
        END IF
    NEXT k
END IF
fz = x - A / 3
END FUNCTION

FUNCTION SRKdh (T, P, y$ )
    SHARED c
    DIM A(c), b(c), Tr(c), Pr(c)
    FOR i = 1 TO c
        Tr(i) = T / Tc(i): Pr(i) = P / Pc(i)
        mi = .48508 + w(i) * (1.55171 - .15613 * w(i))
        IF Tr(i) < 1.8 THEN
            at = (1 + mi * (1 - SQR(Tr(i))))^2
        ELSE
            at = EXP(2 * mi * (1 - SQR(Tr(i))))
        END IF
        A(i) = .42748 * at * Pr(i) / Tr(i)^2
        b(i) = .08664 * Pr(i) / Tr(i)
        A(0) = A(0) + SQR(A(i)) * x(i): b(0) = b(0) + b(i) * x(i)
        e = e + x(i) * mi * SQR(.42748 * Pr(i) / Tr(i))
    NEXT i
    A(0) = A(0)^2

```

```

Z = fz( -1, A(0) - b(0) - b(0)^2, -A(0)* b(0), y$ )
dH = Z - 1 - A(0) / b(0) * (1 + e / SQR(A(0))) * LOG(1 + b(0) / Z)
SRKdh = 8.314 * T * dH
END FUNCTION

```

第二篇 实验数据处理

随着科学技术的发展,实验数据处理的重要性日益突出,已成为测量技术学科的重要内容。近年来由于数理统计的广泛应用及电子计算机的大力普及,使实验数据的处理有了坚实的基础。本篇以数理统计为理论基础,力求简明扼要、通俗易懂地阐述误差理论及实验数据的处理方法,并介绍应用非常广泛的回归分析及软件的使用,以方便分析测试人员及科技工作者参考使用。

第5章 随机误差及其统计分布

对自然科学而言,特别是在科学技术和工程中,信息的收集其主要含义就是测量(取得观测和实验数据)。测量不可避免地会有误差,不管我们测得多么仔细,所使用的仪器多么精密,以及所采用的测量方法多么可靠,测量结果也都或多或少地存在误差。误差常常会掩盖以至歪曲客观事物本来的面目,如果我们对这个事实没有正确的认识,将会妨碍我们去正确地鉴别事物,甚至会使我们做出错误的结论,使产品报废,资源浪费,造成经济损失以及科学研究上的失败。因此,我们除了在测量技术上要做到精益求精,尽可能地使测量结果准确可靠外,还要对误差进行分析、判断。弄清误差产生的原因,误差的种类和性质,而后对误差进行科学地处理,使结果反映出事物的真实面目。

5.1 误差的概念

对测量结果进行分析、研究、判断、处理,必须采用经过长期实践积累形成的误差理论,它是我们认识客观世界的有力工具。为此,我们首先应该了解误差的一些基本概念,包括误差的定义、误差的分类、误差的表示方法以及误差的性质。

5.1.1 误差定义

测量对象的测量值 x_i 与真值 α 之间的差值 Δ_i 称为误差。即

$$\text{误差} = \text{测量值} - \text{真值}$$

或

$$\Delta_i = x_i - \alpha \quad (5.1)$$

很显然,当测量值大于真值时,误差为正;当测量值小于真值时,误差为负。真值在一般情况下是未知的,但在下面三种情况下可以认为真值是预先知道的。

(1) 理论真值 理论上肯定其存在的值。例如,三角形三个内角之和为 180° 。

(2) 约定真值 国际计量大会上通过的基准量值。例如,质量单位“千克”被国际计量大会规定为铂铱合金国际千克原器的质量。

(3) 相对真值 有时也将精度高一个数量级仪器的测量值作为低精度仪器的真值。由于这种“真值”是相对比较而言的, 故称为相对真值。

5.1.2 误差分类

根据误差的性质及产生的原因, 可将误差分为系统误差、随机误差和过失(粗大)误差。

1. 系统误差(可测误差)

由于某种固定原因所造成的误差。在重复测量中, 它会重复出现。对测量结果的影响比较恒定, 使测量结果不是偏高就是偏低, 而且大小有一定规律。因此, 系统误差的大小和正负是可以测定的, 至少从理论上来说是可以测定的, 所以又称为可测误差。主要由以下原因所造成:

(1) 仪器误差 由于测量仪器本身不够精密所引起的误差。如仪表刻度不准、砝码重量不够等;

(2) 试剂误差 由于试剂不纯所引起的误差。例如, 试剂中含有被测物质等;

(3) 方法误差 由于测量方法的错误或测量所依据的理论本身不完善等原因所导致的误差。例如, 用钢卷尺测出轴的圆周长 s , 再由公式 $d = \frac{s}{\pi}$ 来计算轴的直径 d 时所产生的误差;

(4) 操作误差 由于测量者个人的习惯与偏向所引起的误差。例如, 对准标志读数时总是偏上或偏下。

由于系统误差的产生一般具有规律性, 其产生的原因往往可知或能掌握。因此, 对于系统误差应根据其产生的原因设法消除, 对未被消除的系统误差也应设法确定或估计其值。

2. 随机误差(偶然误差)

由于某些难以控制的随机(偶然)因素所引起的误差。例如, 测量条件(环境、温度、气压等)微小变化。随机误差出现的大小、正负难以控制, 所产生的原因也难以确定, 处理方法是:

(1) 测量时仔细操作, 尽可能地减小测量误差;

(2) 进行统计处理: 人们发现随机误差虽在一次测量中无法预言它的大小和正负, 但在大量地重复测量中它却呈现出一定的规律性, 称为统计规律, 对于这种规律性可以用数理统计的方法进行分析、研究, 也是误差理论和数据处理中主要研究的内容。

3. 过失误差(粗大误差)

这是一种与事实不符的误差。主要是由于测量人员粗枝大叶、过度疲劳和操作不当所引起的。例如, 读错刻度尺、反读游标尺、记录错误、计算错误等等。此类误差只要测试人员增强责任心、细心操作是完全可以避免的。

5.1.3 误差的表示方法

反映测量值与真值接近程度的量, 称为精度。常用精度来描述误差, 它可分为准确度、精密度。

1. 准确度

指测量结果与被测量真值之间的偏离程度。显然, 它是由测量误差所反映的。误差愈

小、测量值愈接近真值，测量结果的准确度越高；反之，误差愈大，表示测量值越偏离真值，测量结果的准确度越低。因此，误差是衡量测量准确度高低的尺度。

误差又分为绝对误差和相对误差。

(1) 绝对误差：用被测量本身的误差所直接表示的误差（式 5.1）又称为绝对误差。

(2) 相对误差：绝对误差与真值的比值称为相对误差。即

$$\text{相对误差} = \frac{\text{绝对误差}}{\text{真值}} \quad (5.2)$$

通常以百分数表示。实践中，常用相对误差分析测量结果的准确度。

2. 精密度

实际测量中，真值常常是不知道的，因此也无法分析测量结果的准确度。此时可以用另外一种方法来判断测量结果的好坏，即在相同条件下重复测量多次，然后分析这多次测量结果的相互吻合程度，称为精密度。由于精密度表现了测量结果的再现性，常用偏差来表示。偏差愈小说明测量结果的精密度也愈高。

偏差也分为绝对偏差与相对偏差。

$$\text{绝对偏差} = \text{个别测得值} - \text{测得平均值} \quad (5.3)$$

$$\text{相对偏差} = \frac{\text{绝对偏差}}{\text{真值}} \quad (5.4)$$

相对偏差通常也以百分数表示。这里还有几点说明。

(1) 如果对同一试样做了多次测量，为了更好地分析测量结果的精密度，可使用平均偏差进行分析。即若对某一试样，在相同条件下重复测量 n 次，测得结果为 x_1, x_2, \dots, x_n ，则它们的平均值 \bar{x} 、平均偏差 \bar{d} 与相对平均偏差分别可由下列各式求得：

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (5.5)$$

$$\bar{d} = \frac{|d_1| + |d_2| + \dots + |d_n|}{n} = \frac{1}{n} \sum_{i=1}^n |d_i| \quad (5.6)$$

式中 $d_i = x_i - \bar{x}$ ($i = 1, 2, \dots, n$)

$$\text{相对平均偏差} = \frac{\bar{d}}{\bar{x}} \quad (5.7)$$

注意：平均偏差 \bar{d} 不记正负号，而个别测量值的偏差 d_i ($i = 1, 2, \dots, n$) 要记正负号

(2) 当一批测量数据分散程度较大时，往往仅从平均偏差还看不出或很难看出精密度的好坏。此时还可利用标准偏差（又称均方根偏差，简称均方差）来衡量精密度。其公式为：

$$s = \sqrt{\frac{d_1^2 + d_2^2 + \dots + d_n^2}{n-1}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n d_i^2} \quad (5.8)$$

s 值越大，精密度越差。

例 5.1 有甲、乙两组数据，对其各次测量的偏差分别为

甲组：+0.1, +0.4, 0.0, -0.3, +0.2, -0.3, +0.2, -0.2, -0.4, +0.3

乙组：-0.1, -0.2, +0.9, 0.0, +0.1, +0.1, 0.0, +0.1, -0.7, -0.2

试分析哪组数据的精密度好？

解：先计算平均偏差

$$\bar{d}_{\text{甲}} = \frac{1}{10} (0.1 + 0.4 + 0.0 + 0.3 + 0.2 + 0.3 + 0.2 + 0.2 + 0.4 + 0.3) = 0.24$$

$$\bar{d}_{\text{乙}} = \frac{1}{10} (0.1 + 0.2 + 0.9 + 0.0 + 0.1 + 0.1 + 0.0 + 0.1 + 0.7 + 0.2) = 0.24$$

两组数据的平均偏差相同，反映不出精密度的好坏，再分别计算它们的标准偏差得：

$$\bar{d}_{\text{甲}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n d_i^2} = \sqrt{\frac{(0.1)^2 + (0.4)^2 + \cdots + (0.3)^2}{10-1}} = 0.28$$

$$\bar{d}_{\text{乙}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n d_i^2} = \sqrt{\frac{(-0.1)^2 + (-0.2)^2 + \cdots + (-0.2)^2}{10-1}} = 0.40$$

可见甲组数据的精密度较好。

由于将单次测量的偏差平方之后，较大的偏差能更显著地反映出来。所以，一般情况下用标准偏差表示精密度往往比平均偏差更好。实际计算中，为方便有时也用 $\frac{1}{n}$ 代替式（式 5.8）中的 $\frac{1}{n-1}$ 。

3. 准确度与精密度的关系

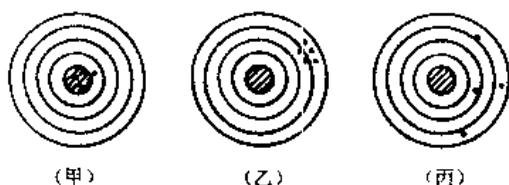


图 5.1 准确度与精密度的关系

由上述讨论可知，准确度是指测量值与真值之间的相互吻合程度，一般它是由系统误差所决定的；精密度则是指多次测量值之间相互吻合程度，它是由随机误差所决定的。两者含义不同，不可混淆。但准确度与精密度之间又存在着一定的关系，现以打靶为例予以说明。

我们视靶心为真值，弹孔相对靶心的距离为测量值，则由图 5.1 可以看出，甲的准确度高，精密度也好；乙的准确度较差，但精密度较好；丙的准确度差，精密度也不好，由此易得下列结论：

- (1) 一个好的测量结果，既要精密度好又要准确度高；
- (2) 准确度高一定需要精密度好，而精密度好不一定准确度高。

例 5.2 两位化验员对一样品各做四次测定，以确定其含量，结果如下：

甲 75.00%，74.41%，76.53%，77.04%

乙 75.31%，75.27%，75.40%，75.45%

样品标准含量为 75.00%，试比较甲、乙两人的测量结果，并指出他们各自在测量中存在的问题及改进工作的建议。

解：由计算他们各自的误差与偏差，寻找测量中存在的问题。

对甲所测结果的计算如下：

$$\text{平均值 } \bar{x}_{\text{甲}} = \frac{1}{4} (75.00 + 74.41 + 76.53 + 77.04) = 75.75$$

$$\text{绝对误差 } \Delta_{\text{甲}} = 75.75 - 75.00 = +0.75$$

$$\text{相对误差} = \frac{+0.75}{75} \times 100\% = +1.00\%$$

单次测量值的绝对偏差

$$d_1 = 75.00 - 75.75 = -0.75$$

$$d_2 = 74.41 - 75.75 = -1.34$$

$$d_3 = 76.53 - 75.75 = +0.78$$

$$d_4 = 77.04 - 75.75 = +1.29$$

$$\text{绝对平均偏差 } \overline{d}_{\text{甲}} = \frac{0.75 + 1.34 + 0.78 + 1.29}{4} = 1.00$$

$$\text{相对平均偏差} = \frac{\overline{d}_{\text{甲}}}{x_{\text{甲}}} = \frac{1.00}{75.75} = 1.32\%$$

对乙所测结果的计算如下:

$$\text{平均值 } \overline{x}_{\text{乙}} = \frac{1}{4} (75.31 + 75.27 + 75.40 + 75.45) = 75.36$$

$$\text{绝对误差 } \Delta_{\text{乙}} = 75.36 - 75.00 = +0.36$$

$$\text{相对误差} = \frac{0.36}{75.00} \times 100\% = 0.50\%$$

单次测量值的绝对偏差

$$d_1 = 75.31 - 75.36 = -0.05$$

$$d_2 = 75.27 - 75.36 = -0.09$$

$$d_3 = 75.40 - 75.36 = +0.04$$

$$d_4 = 75.45 - 75.36 = +0.09$$

由以上计算结果可以看出:甲测量结果的误差、偏差都比较大,表明准确度和精确度都不好,这是由于系统误差和随机误差都偏大的缘故,应该重做。重做时应注意校正系统误差,增强责任心。乙的测量结果较好,但误差还偏大,显然存在着系统误差,应对系统误差进行校正,以便进一步提高测量水平。

5.1.4 等精度测量与非等精度测量

按测量的精度情况,测量又可分为等精度测量和非等精度测量两种。

(1) 等精度测量 在整个测量过程中,影响和决定误差大小的全部因素(条件)始终不变,称为等精度测量。如在同样的环境条件下,由同一个观测者,使用同一台仪器、同样的方法,对同一个工件做相同次数的测量。很显然,在实际当中很难做到真正地等精度测量。一般情况下,只是某种程度上的近似地等精度测量。

(2) 非等精度测量 在整个测量过程中,影响和决定误差大小的因素不全相同的测量称为非等精度测量。如由不同的观测者,用不同的仪器、不同的方法,在不同的环境条件下对工件做不同次数的测量。

5.1.5 加权平均

在等精度测量中,各个测量值可以认为是同样可靠的,因此可取所有测量值的算术平均值作为最后的测量结果。而在不等精度测量中,由于测量条件不完全相同(如不同人员,不同仪器,不同地点,不同测量次数等),往往可靠性(精密度)并不相同。因此,对于不等精度测量值的评价就不能简单地取各次测量值的算术平均值衡量,而应考虑各测量值的大小,同时还应权衡它们的可信赖程度,使可靠程度大的测量值在最后的测量结果中所占的比重更大些。各测量值的可靠程度一般用“权”来表示,记作 p 。“权”的选择可根据测量条件

而定。例如，当其他的测量条件相同的情况下，其“权”也越大，此时可用测量次数 n 来确定“权”。带有“权”的平均称为加权平均。

例 5.3 设用同一测角仪，测量某一角度三回。第一回测 3 次，第二回测 6 次，第三回测 12 次（每取一读数为一读）结果如下：

$$\bar{x}_1 = 32^\circ 41' 20'', \quad \bar{x}_2 = 32^\circ 41' 30'', \quad \bar{x}_3 = 32^\circ 41' 25''$$

求加权平均值。

解：其“权”（测量次数比）为：

$$p_1 : p_2 : p_3 = 3 : 6 : 12 = 1 : 2 : 4$$

此角度的平均值（加权）为：

$$\begin{aligned}\bar{x}_p &= \frac{32^\circ 41' 20'' \times 3 + 32^\circ 41' 30'' \times 6 + 32^\circ 41' 25'' \times 12}{3 + 6 + 12} \\ &= 32^\circ 41' 20'' \times \frac{1}{7} + 32^\circ 41' 30'' \times \frac{2}{7} + 32^\circ 41' 25'' \times \frac{4}{7} \\ &= 32^\circ 41' 26''\end{aligned}$$

显然，加权平均不同于一般的算术平均（此例的算术平均值为 $32^\circ 41' 25''$ ），加权平均实际上是取值及取此值“权”的乘积之和。为此，算术平均值可以看作是加权平均值的一个特例，即等权平均值。

5.2 随机误差的统计分布

从测量实践得知，在排除了系统误差和过失误差的情况下，对某一物理量进行等精度的多次测量时，其测量值中还会含有随机误差。随机误差的出现具有偶然性，也就是说这种误差出现的大小、符号不能预先知道。但随测量次数增大，这类误差就显现出其统计规律性，且测量次数愈多，这种规律性表现也愈明显。掌握了这一规律就可以对测量结果进行恰当的分析、处理。

随机误差的统计规律性是用数理统计的方法来进行研究的。数理统计方法的特点是从大量或一批（足够数量）的测量值中寻求规律。被研究对象的全体称为总体，组成总体的每个基本单位称为个体（或样品）。从总体中随机地抽取的一部分个体称为一个样本（或子样），样本中所含测量值的个数称为样本容量。

下面我们就用数理统计的方法来研究测量值中随机误差的统计规律性。

5.2.1 频数分布

当我们对某一试样进行了多次重复测量后，就会获得一大批杂乱无章的数据。为了揭示这些数据的内在规律，需要绘制它们的频数直方图。

例 5.4 测量镍合金试样中镍的百分含量。在相同条件下，共测量 90 次，得到数据如下：

1.60	1.67	1.67	1.64	1.58	1.64	1.67	1.62	1.57	1.60
1.59	1.64	1.74	1.65	1.64	1.61	1.65	1.69	1.64	1.63
1.65	1.70	1.63	1.62	1.70	1.65	1.68	1.66	1.69	1.70
1.70	1.63	1.67	1.70	1.70	1.63	1.57	1.59	1.62	1.60
1.53	1.56	1.58	1.60	1.58	1.59	1.61	1.62	1.55	1.52
1.49	1.56	1.57	1.61	1.61	1.61	1.50	1.53	1.53	1.59
1.66	1.63	1.54	1.66	1.64	1.64	1.64	1.62	1.62	1.65
1.60	1.63	1.62	1.61	1.65	1.61	1.64	1.63	1.54	1.61
1.60	1.64	1.65	1.59	1.58	1.59	1.60	1.67	1.68	1.69

绘制直方图的步骤如下：

1. 算极差

从测量值中，找出最大值与最小值，最大值与最小值之差为极差。此例的极差 $R = 1.74 - 1.49 = 0.25$ 。

2. 确定组数和组距

由极差及测定值的样本容量可以确定出频数直方图中的组数与组距。

组数即分组的个数，依样本容量大小而定。一般容量较小的 ($n < 100$)，可分为 5~10 组，容量大时可适当增加组数，但无严格的划定。如本例，我们可将数据分为 9 组。

组距：即组与组之间的距离。为方便，常按等距分组。如本例组距为 $(1.74 - 1.49) / 9 \approx 0.03$ 。

这里应该指出的是为了使每一个数据只分在一组内（有时测量值正好与组界相等，出现了所谓的“骑墙”现象），分组时通常是将组界多取一位。为此，本例分成的 9 组应为：

1.485~1.515 1.515~1.545 1.545~1.575
 1.575~1.605 1.605~1.635 1.635~1.665
 1.665~1.695 1.695~1.725 1.725~1.755

3. 列频数分布表

分组后，数出应该落在每组中的数据个数（称为频数），计算频数与样本容量之比（称为相对频数。若以百分比表示，则称为频率）。根据以上统计、计算值列成下面的频数分布表见表 5.1。

表 5.1 频 数 分 布 表

分 组	频 数	相对频数
1.485~1.515	2	0.022
1.515~1.545	6	0.067
1.545~1.575	6	0.067
1.575~1.605	17	0.189
1.605~1.635	22	0.244
1.635~1.665	20	0.222
1.665~1.695	10	0.111
1.695~1.725	6	0.067
1.725~1.755	1	0.011
Σ	90	1.000

4. 绘制直方图

取横坐标为组界值、纵坐标为相对应的频数值，作出以组距为底和以频数为高的矩形，

就得到频数直方图（见图 5.2）。若取纵坐标为相对频数，则所得到的矩形图为相对频数直方图（见图 5.3）。

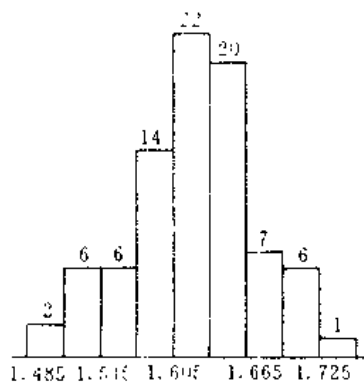


图 5.2 频数直方图

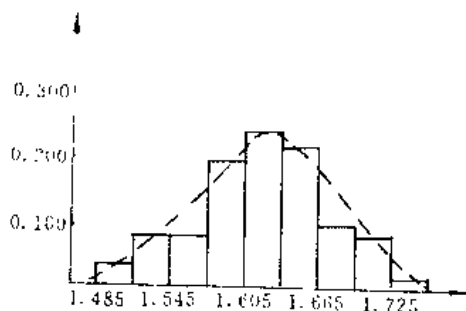


图 5.3 相对频数直方图

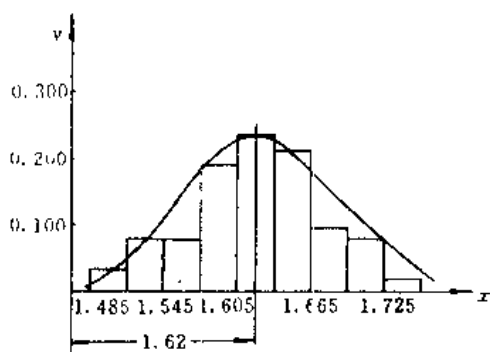


图 5.4 理论曲线

从频数分布表与频数直方图可以看出，全部测量值虽然是分散的，但又呈现出了明显地集中趋势。即大多数测量值都集中在平均值 1.62 附近。为进一步研究测量数据的这种既分散又集中的特性，我们再从相对频数直方图进行一些分析。由于相对频数的总和为 1，所以相对频数直方图上，小长方形面积的总和为 1。连接各矩形上边的中点，可以得到一条曲线（见图 5.3 中的虚折线），可以设想，如果测量的次数不断增加，分组愈来愈细、组距愈来愈小时，相对频数直方图就愈来愈接近于一

条连续曲线（称理论曲线，见图 5.4）。这条连续曲线的特点是具有单峰性，中间大、两头小，呈现钟形。在概率论中称其为正态分布曲线。

5.2.2 正态分布

正态分布最早是在十八世纪后半期由高斯（Gauss）在研究误差理论时发现的，所以正态分布又称为高斯分布。大量实践表明，在测量过程中所产生的随机误差，往往是观测者不能控制的大量偶然因素作用的结果，它们是服从正态分布的。概率论的中心极限定理从理论上说明了上述现象的必然性。该定理指出：“若随机变量是由为数众多的相互独立的随机因素的影响叠加而成，而每一个别因素在总的影响中所起的作用又都非常微弱，那么这类随机变量应该是服从或近似服从正态分布的。

正态分布的数学表达式为：

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < +\infty) \quad (5.9)$$

式中 $p(x)$ ——正态概率密度函数；

μ ——总体平均值（作为真值），相当于曲线最高点的横坐标；

σ ——正态分布的标准偏差（简称标准差）， $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$ ；

x ——个别测定值。

其图形关于 $x = \mu$ 对称 (见图 5.5), 在 $x = \mu$ 处有最大值 $\frac{1}{\sqrt{2\pi}\sigma}$ 。

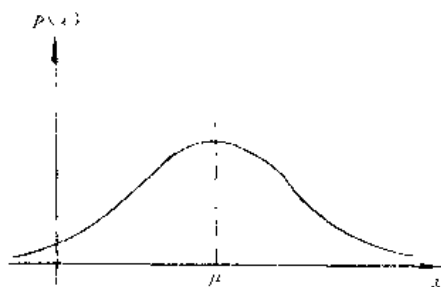


图 5.5 正态分布密度曲线

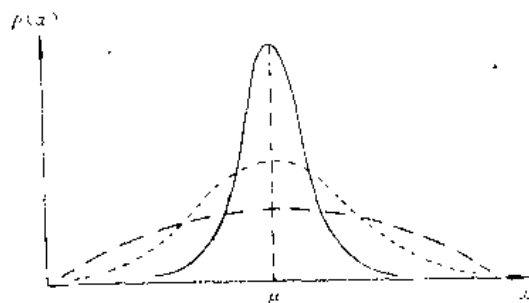


图 5.6 μ 固定, σ 变化

固定平均值 μ , 而使 σ 变化, 则曲线的对称轴不变, 而 σ 越大, 曲线最高点越低, 图形越平缓; σ 越小, 曲线最高点越高, 图形也越陡峭 (见图 5.6)。

从分析测试的实用观点来看, σ 越大, 意味着测量精密度越差, 测量值越分散; 反之, σ 越小, 表示测量精密度越好, 测量值越集中。当精密度不变, 平均值 μ 发生变化时, 则图形依 μ 的不同沿 x 轴左、右移动, 而形状不变 (见图 5.7)。

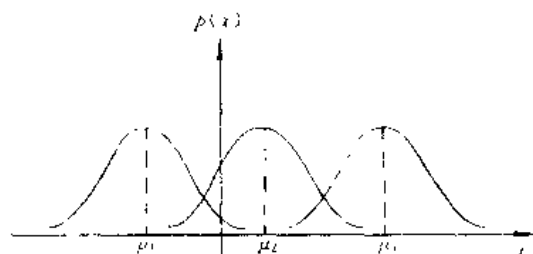


图 5.7 σ 固定, μ 变化

综上所述, 可以看出利用正态分布来描述测量中随机误差的分布规律是非常合适的。而平均值 μ 和标准差 σ 是正态分布的两个基本参数, 只要给定 μ 和 σ , 正态分布曲线就被完全确定。平均值为 μ , 标准差为 σ 的正态分布常记作 $N(\mu, \sigma^2)$ 。

5.2.3 概率计算

正态分布在误差理论和计算中都有十分重要的作用。下面我们将介绍怎样利用正态分布曲线, 计算测量值落在某一给定区间上的概率。

由前面分析可知: 不管 μ 、 σ 为何值, 正态分布曲线与横坐标轴之间所夹的面积, 即正态分布密度函数 $p(x)$ 在区间 $-\infty < x < +\infty$ 上的积分值。它也代表了具有各种大小偏差的样本值出现概率 (可能性大小的一种度量) 的总和, 其值为 1。也就是

$$P(-\infty < x < +\infty) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 \quad (5.10)$$

样本值 X 落在任一区间 $[a, b]$ 上的概率 ($a \leq x \leq b$) 就等于横坐标在 $x = a$, $x = b$ 区间上的曲线与横坐标轴之间所夹的面积, 即

$$P(a \leq x \leq b) = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (5.11)$$

这个积分的计算与 μ 和 σ 有关, 计算起来比较麻烦。为方便计算, 常经过一个变换,

$$\frac{x - \mu}{\sigma} = t \quad (5.12)$$

则计算式变为

$$\begin{aligned} P(a \leq x \leq b) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} \sigma dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt \end{aligned} \quad (5.13)$$

相当于均值 $\mu = 0$, 标准差 $\sigma = 1$ 的正态分布在区间 $\left[\frac{a-\mu}{\sigma}, \frac{b-\mu}{\sigma}\right]$ 上的概率 (积分)。称均值 $\mu = 0$, 标准差 $\sigma = 1$ 的正态分布为标准正态分布, 记作 $N(0, 1)$, 其密度函数为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (-\infty < x < +\infty) \quad (5.14)$$

其图形 (见图 5.8) 关于纵坐标对称, 最大值为 $\frac{1}{\sqrt{2\pi}}$ 。

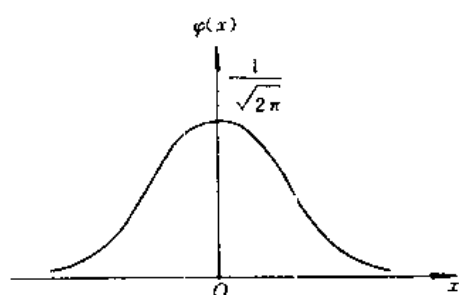


图 5.8 标准正态密度曲线

基于标准正态分布在实际应用上的重要性, 人们已将 (5.14) 式的积分值作成表 (见表 5.2), 使用时可查此表。表中给出的积分值是

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{x^2}{2}} dx \quad (5.15)$$

它表示测量值 X 落在区间 $(-\infty, u]$ 上的概率值。例如, 求 $\Phi(1.64)$ 的值。由表 5.2 可查得其值为 0.94950, 这意味着测量值落在 $(-\infty, 1.64]$ 内的概率为 0.94950。

这里还应该指出几点:

- (1) 当 $u > 0$ 时, $\Phi(u) = P(-\infty < x \leq u)$ 的值可以从表 5.2 直接查得;
- (2) 当 $u < 0$ 时, 由于标准正态密度函数曲线关于 $\mu = 0$ ($x = \mu$) 对称。因此

$$\Phi(u) = 1 - \Phi(-u)$$

查表 5.2 先找到 $\Phi(-u)$, 再计算 $\Phi(u)$ 即可;

- (3) 若样本值 X 遵循标准正态分布 $N(0, 1)$, 则 X 落在区间 $[a, b]$ 上的概率为

$$P(a \leq x \leq b) = P(x \leq b) - P(x \leq a) = \Phi(b) - \Phi(a)$$

可通过表 5.2 查得 $\Phi(b)$ 、 $\Phi(a)$ 后, 再计算其值;

- (4) 若样本值 X 遵循正态分布 $N(\mu, \sigma^2)$, 则 X 落在区间 $[a, b]$ 上的概率为

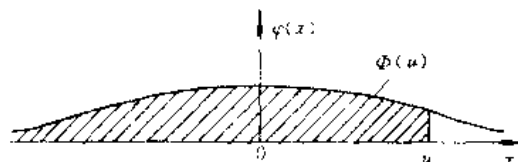
$$\begin{aligned} P(a \leq x \leq b) &= \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt \quad (t = \frac{x-\mu}{\sigma}) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

可查表 5.2 计算其值;

- (5) X 在一点处的概率值应该认为是零。

表 5.2 标准正态分布函数值

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{\tau^2}{2}} d\tau \quad (u \geq 0)$$



$\Phi(u)$	0.00	0.02	0.04	0.05	0.06	0.08	0.09
0.0	0.5000	0.5080	0.5160	0.5199	0.5239	0.5319	0.5359
0.4	0.6554	0.6628	0.6700	0.6736	0.6772	0.6844	0.6879
0.5	0.6915	0.6985	0.7054	0.7088	0.7123	0.7190	0.7224
0.9	0.8159	0.8212	0.8264	0.8289	0.8315	0.8365	0.8389
1.0	0.8413	0.8461	0.8508	0.8531	0.8554	0.8599	0.8621
1.2	0.8849	0.8888	0.8925	0.8944	0.8962	0.8997	0.9017
1.4	0.91924	0.92220	0.92507	0.92647	0.92785	0.93056	0.93189
1.5	0.93319	0.93574	0.93822	0.93943	0.94062	0.94295	0.94408
1.6	0.94520	0.94738	0.94950	0.95053	0.95154	0.95352	0.95449
1.9	0.97128	0.97257	0.97381	0.97441	0.97500	0.97615	0.97670
2.0	0.97725	0.97831	0.97932	0.97982	0.98030	0.98124	0.98169
2.5	0.993790	0.994132	0.994457	0.994614	0.994766	0.995060	0.995201
2.6	0.995339	0.995604	0.995855	0.995975	0.996093	0.996319	0.996427
3.0	0.998650	0.998736	0.998817	0.998856	0.998893	0.998965	0.998999
3.5	0.9997674	0.9997842	0.9997999	0.9998074	0.9998146	0.9998282	0.9998347
4.0	0.99996833	0.99997090	0.99997327	0.99997439	0.99997546	0.99997748	0.99997843
4.5	0.999996602	0.999996908	0.999997187	0.999997313	0.999997442	0.999997675	0.999997784
4.8	0.999999201	0.999999282	0.999999351	0.999999383	0.999999413	0.999999466	0.999999496
4.9	0.999999521	0.999999567	0.999999609	0.999999629	0.999999648	0.999999682	0.999999692

例 5.5 已知样本值 X 遵循正态分布 $N(\mu, \sigma^2)$ 。试求 X 落在区间 $(\mu - \sigma, \mu + \sigma)$ 、 $(\mu - 2\sigma, \mu + 2\sigma)$ 、 $(\mu - 3\sigma, \mu + 3\sigma)$ 内的概率。

解: $P(\mu - \sigma < x < \mu + \sigma)$

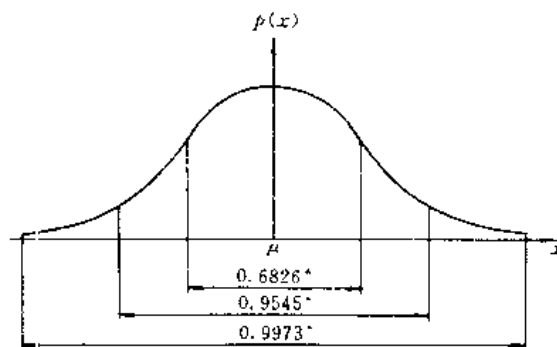
$$= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) = \Phi(1) - \Phi(-1) \\ = 2\Phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826$$

同理可得:

$$P(\mu - 2\sigma < x < \mu + 2\sigma) = 2\Phi(2) - 1 = 0.95450$$

$$P(\mu - 3\sigma < x < \mu + 3\sigma) = 2\Phi(3) - 1 = 0.99730$$

此例表明: 在全部测量值中, 偏差大于三倍标准差的测定值应该很少, 不超过全部测量值的 3%。即平均 1000 次测量中, 最多只有三次出现的机会, 称作三倍标准差规则 (或 3σ 规则) (见图 5.9)。在企业管理中经常应用 3σ 规则进行质量检查和工艺过程控制。例如, 当对某物理量进行测量时, 仅作了少数几次测量就发现了误差较大的测量值, 应该认为这是不正常的, 需考虑系统误差的存在。



注：图 5.9 中 0.6826 指区间 $(\mu - \sigma, \mu + \sigma)$ 内的概率值，其它意义类同。

图 5.9 例 5.5 图

习 题

1. 下列情况引起的误差是系统误差，还是随机误差？

- (1) 砝码被腐蚀；
- (2) 天平零点稍有变动；
- (3) 天平的两臂不等长；
- (4) 读取数据时，最后一位数字估计不准；
- (5) 试剂中含有微量待测组分。

2. 已知矿石标样中含有 Fe_2O_3 为 50.36%，现有甲、乙、丙三位化验员同时对矿样进行测量，各测四次，测量结果如下：

甲 50.20%，50.20%，50.18%，50.17%；

乙 50.40%，50.30%，50.20%，50.10%；

丙 50.36%，50.35%，50.34%，50.33%。

试比较三人的测量结果，并指出他们各自在测量中存在的问题。

3. 将一种样品送给两位化验员分析，两个的报告分别为：

甲 30.15%，30.15%，30.14%，30.16%；

乙 30.251%，30.007%，30.101%，30.241%。

试计算平均值及平均偏差。当其他条件相同时，分析哪一个更可靠。

4. 由某台机器生产的螺栓长度，服从参数 $\mu = 10.05$ ， $\sigma = 0.06$ 的正态分布。若规定其长度在范围 $10.05\text{cm} \pm 0.12\text{cm}$ 内的为合格品，求螺栓的合格品率。

5. 一工厂生产的电子管的寿命 x （以小时计）服从参数为 $\mu = 160$ 的正态分布，若要求

$$P(120 < x \leq 200) = 0.8$$

问允许 σ 最大为多少？

第六章 离群值的检验与取舍

当人们对同一个量进行多次重复测量后,有时会发现:在这一系列测量值中,有一、二个值明显地偏离其他值,我们称这样的测量值为离群值。有些人,主观地认为“离群值”的出现是不应该的,随意地将其舍弃。其实这是极不恰当的,因为有些离群值可能是测量值随机波动的极度表现,或说是在统计上处于所允许的误差范围内的极值,它的存在反而能表现出事物的某些特性。因此,对于离群值的取舍一定要慎重。一般应先从测量技术上设法弄清其出现的原因,如果离群值的出现确实是由于测量技术上的失误而造成的,则必须将其舍弃。但在很多情况下,未必能从测量技术上找出离群值出现的原因。此时应利用数理统计方法检验其是否为应该舍弃的异常值。

6.1 离群值的舍弃准则

在第五章中已经述及测量误差是服从正态分布的,根据正态分布的特性得知:在一组测量值中,出现大偏差测量值的概率是非常小的。例如,偏差大于两倍标准差的测量值,其出现的概率不超过5%;偏差大于三倍标准差的测量值,其出现的概率更小,不超过3‰。离群值的舍弃准则正是依赖于这个随机抽样理论的。

离群值的检验思想是:规定一个置信水平(或说可信水平,一般取90%、95%和99%),再根据这个置信水平及样本容量 n ,确定出一个合理的误差限。凡是超出这个限度的误差,就认为它已不属于随机误差的范围,而应予以剔除,否则,则无理由剔除,而应予以保留。例如,根据 3σ 准则知,当测量次数较大时,有

$$p(|X - \alpha| \leq 3\sigma) \approx 99.7\% \quad (6.1)$$

式中 x 为测量值;

α 为被测物理量真值。

它表明有99.7%的可信性,测量误差应不超过 3σ ,即在1000次测量值中,测量误差大于 3σ 的情况大约为3次。可见,在少数若干次测量中,若出现了 $|x - \alpha| > 3\sigma$ 的情况,应认为是不可能的,对出现了这种情况的测量值,应认为是异常值而予以舍弃。

6.2 离群值的检验方法

根据离群值的舍弃准则,人们总结出了一些检验离群值是否异常的方法。下面我们介绍两种常用的检验方法。

6.2.1 格鲁布斯(Grubbs)检验法

这种检验法是在1950年由F.E. Grubbs提出来的。使用的统计量为:

$$G_{it} = \frac{|X_{\alpha} - \bar{X}|}{S} \quad (6.2)$$

式中 X_a ——离群值；

\bar{X} 与 S ——分别为 n 次测量值的平均值和标准差。

检验步骤如下：

(1) 将全部测量值由小到大顺序排列为：

$$X_1 \leq X_2 \leq \cdots \leq X_n$$

(2) 计算统计量 $G_{\text{计}}$ ；

(3) 根据测量次数 n 及选定的检验水平 α 查表 6.1 (格鲁布斯检验临界值 $G_{\alpha}(n)$ 表)，得出临界值 $G_{\text{表}}$ ；

(4) 比较 $G_{\text{计}}$ 与 $G_{\text{表}}$ ，当 $G_{\text{计}} \leq G_{\text{表}}$ 为异常值， X_a 应剔除；否则应予以保留。

表 6.1 格鲁布斯检验临界值 $G_{\alpha}(n)$ 表

n	α		n	α		n	α	
	0.01	0.05		0.01	0.05		0.01	0.05
3	1.15	1.5	12	2.55	2.29	21	2.91	2.58
4	1.49	1.46	13	2.61	2.33	22	2.94	2.60
5	1.75	1.678	14	2.66	2.37	23	2.96	2.62
6	1.94	1.82	15	2.70	2.41	24	2.99	2.64
7	2.10	1.94	16	2.74	2.44	25	3.01	2.66
8	2.22	2.03	17	2.78	2.47	30	3.10	2.74
9	2.32	2.11	18	2.82	2.50	35	3.18	2.81
10	2.41	2.18	19	2.85	2.53	40	3.24	2.87
11	2.48	2.24	20	2.88	2.56	50	3.34	2.96

例 6.1 测量碱灰总灰量 w (Na_2O) 得到 6 个数据为：40.02, 40.12, 40.16, 40.18, 40.20, 40.18，问 40.02 这个数据应否剔除？

解：因为已指定被检验数据，可以不再排列测量数据顺序。

由样本值计算得： $\bar{X} = 40.14$ ， $S = 0.066$

$$G_{\text{计}} = \frac{|X_a - \bar{X}|}{S} = \frac{|40.02 - 40.14|}{0.066} = 1.862$$

又由 $n = 6$ ， $\alpha = 0.05$ 查表 6.1 得 $G_{\text{表}} = 1.82$ ， $G_{\text{计}} > G_{\text{表}}$ ，40.02 为异常值，应剔除。

格鲁布斯法检验离群值取舍的理论依据是：若实验数据 X 服从正态分布 $N(\mu, \sigma^2)$ ，则有

$$P\left\{\frac{|X_a - \bar{X}|}{S} \geq G_{\text{表}}\right\} = \alpha \quad (6.3)$$

成立。为此，在检验中引进了两个重要参数 \bar{X} 和 S (S 代替 σ)，所以准确度较高，是合理而普遍适用的一种方法。

6.2.2 肖维奈特 (Chauvenet) 准则

此准则指出：在 n 次等精度的测量中，对于测量列 X_1, X_2, \cdots, X_n ，若已不含有系统误差，且 X_i 服从正态分布，则测得值残差 δ_i 的绝对值超出某误差限 $\delta = k\sigma$ ，即

$$|\sigma_i| > k\sigma \quad (6.4)$$

式中

$$\sigma_i = X_i - \bar{X} \quad (6.5)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N \sigma_i^2}{n-1}} \quad (6.6)$$

时, 则可以认为 X_i 是离群值而予以剔除。式中 k 值可由表 6.2 查得。

表 6.2 肖维奈特准则的 n 与 k 值表

n	k	n	k	n	k	n	k
5	1.65	15	2.13	25	2.33	80	2.74
6	1.73	16	2.16	26	2.34	100	2.81
7	1.79	17	2.18	27	2.35	150	2.93
8	1.86	18	2.20	28	2.37	185	3.00
9	1.92	19	2.22	29	2.38	200	3.02
10	1.96	20	2.24	30	2.39	250	3.11
11	2.00	21	2.26	35	2.45	500	3.29
12	2.04	22	2.28	40	2.50	1000	3.48
13	2.07	23	2.30	50	2.58	2000	3.66
14	2.10	24	2.32	60	2.64	5000	3.89

利用肖维奈特准则检验离群值时, 应从最大或最小的数据开始。若发现不只一个值离群, 则当将最大(或最小)的离群值剔除后, 应在余下的测量值中重新计算 \bar{X}' 和 σ' , 根据这新的 \bar{X}' 和 σ' 值, 再次对另一离群值进行检验。如此类推, 每次剔除一个, 直至余下的测量值全部符合肖维奈特准则为止。

例 6.2 对某一温度进行 15 次等精度测量, 测量值为 ($^{\circ}\text{C}$): 20.42, 20.43, 20.40, 20.43, 20.42, 20.43, 20.39, 20.30, 20.40, 20.43, 20.42, 20.41, 20.39, 20.39, 20.40, 试用肖维奈特准则判断是否有应舍弃的离群值。

解: 由测量数据计算得:

$$\bar{X} = 20.404$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{0.01496}{14}} = 0.033$$

$$|\sigma_8| = 0.104$$

又由表 6.2 查得 $k = 2.13$, $k\sigma = 2.13 \times 0.033 = 0.07$, 而 $|\sigma_8| = 0.104 > 0.07 = k\sigma$, 故 $t_8 = 20.30$ 是离群值, 应予以剔除。

剔除 $t_8 = 20.30$ 后, 又计算:

$$\bar{t} = 20.411$$

$$\sigma' = 0.016$$

再次检验余下的 14 个测量值, 未再发现应舍弃的离群值, 故均应予以保留。

这里还应指出几点:

- (1) 当被剔除的离群值占全部数据的比例较大时, 应从测量方法和技术上探究原因;
- (2) 肖维奈特准则是在测量频率趋近于概率的前提下推导出来的。因此, 当测量次数 n 较小时, 不易使用;

(3) 两种检验方法各有优、劣，应根据实际问题恰当选用。但无论使用哪种方法进行检验，在实际测量中都应首先在技术上寻找原因，当找不出存在离群值的测量技术上的原因时，才能借助统计方法进行处理。

(4) 在回归分析（见第七章）中常用肖维奈特准则检验离群值，此时（6.4）式应为

$$\sigma_i = Y_i - \hat{Y}_i \quad (6.6)$$

式中 Y_i ——试验值；

\hat{Y}_i ——相应地回归值（详见第七章）。

使用这个准则时，可根据回归结果，对全部实验值进行逐组检查，把属于可疑数据的实验值选出。若发现不止一个可疑数据，则应把其中残差绝对值最大者剔除。然后重新计算标准差 σ 值。根据新的值，再次用肖维奈特准则进行检查。每次只剔除一个可疑数据，其余数据重新进行回归，直至回归所用的数据中不再含有可疑数据为止。

习 题

用格鲁布斯、肖维奈特准则，检验下列各题中是否存在异常值，并比较两种检验法的结果。

1. 用 1, 10-邻菲咯啉分光光度法测定硫酸锌中的 Fe，5 次测量值分别为：2.63, 2.50, 2.65, 2.59, 2.62 ($\times 10^{-6}$) ($\alpha = 0.01$)。

2. 用原子吸收法测定土壤样品中砷含量，9 个样品的测量结果为：7.76, 8.96, 8.82, 10.98, 8.85, 7.79, 8.20, 9.18, 9.52 ($\times 10^{-6}$)。

3. 用分子发射腔分析法，测定磷酸在氢——空气扩散火焰中的 HPO 分子发射带强度，得到如下一组测量值：73.5, 69.0, 69.5, 67.0, 63.5, 69.5, 70.0, 70.5, 67.0 ($\alpha = 0.05$)。

第7章 回归分析与数据处理

回归分析是应用数理统计方法,对实验数据进行分析、处理,从而得出反映变量间相互关系的近似表达式——经验公式的一种常用方法。回归分析的应用十分广泛,例如新产品的研制、零部件的质量分析、新标准的制定、工艺因素的分析与控制、气象预报、系统工程中的预测与规划等,都可以利用回归分析的方法分析解决。

7.1 回归分析和曲线拟合的概念

7.1.1 回归分析

当我们观察自然现象时,会发现许多相互制约着的量,它们之间存在着一定的联系,这些联系一般可以分为两大类:确定性关系与非确定性关系。

(1) 确定性关系 即一个量的数值可以由另外一个量或几个量的数值来确定。如长方形的面积 S 与它的长 a 与宽 b 之间的关系是 $S = ab$; 又如在固定体积 V 的容器中,气体的温度 T 与压力 P 之间的关系是 $P = RT/V$ (R 、 V 为常数) 等。

(2) 非确定性关系 即变量之间的关系不能用普通函数式来表达。例如身高与体重的关系,小麦单位面积产量与施肥量之间的关系,废品率与成本之间的关系等。

对于具有非确定性关系的变量,通过大量的观察和试验,人们发现它们之间又呈现出某种规律性,称统计规律性。这种变量之间具有统计相关关系,简称相关关系。

回归分析是处理变量之间相关关系的一个有力工具。由一个或一组非随机变量来估计或预测某一随机变量的观测值时,所建立的数学模型及所进行的统计分析称为回归分析。

回归分析一般是这样进行的:

- (1) 通过实验取得一组相互独立的近似值——实验数据;
- (2) 拟定出含有若干个待定参数的关系式(经验公式),再用回归分析法处理实验数据,以确定待定参数,得到回归方程;
- (3) 对所得到的回归方程进行相关性统计检验,以确定其可信程度;
- (4) 利用所得到的已达到所需要可信程度的回归方程,进行预测与控制。

7.1.2 曲线拟合

科学实验中,常常需要从一组测量的实验数据, n 对 (x_i, y_i) , 求得自变量 x 与因变量 y 之间的一个近似函数关系式 $y = f(x)$, 这就是所谓的曲线拟合问题。由图形上看,也就是这 n 个实验数据点可以近似落在怎样的一条曲线上。这里应当指出的是,拟合曲线不一定都通过所有实验点 (x_i, y_i) 。依据某种数学原理,或根据所研究问题的机理(或经验)来建立一条近似曲线(实验方程),使它能反映出实验数据的一般趋势,而尽量不让误差出现较大波动。因此,拟合曲线一般来自两个途径:

- (1) 根据实验数据直接作图,观察大致变化趋势,进行拟合曲线;

(2) 通过研究各变量的物理概念, 确定出函数的基本类型。

总之, 做曲线拟合时, 对未知函数的类型的选择和假设是个十分复杂而重要的问题。我们将在以下几节中分类进一步进行曲线拟合的分析研究。

7.2 一元线性回归

一元线性回归处理的是两个变量之间的线性相关关系问题。

7.2.1 经验公式

例 7.1 为了控制钢的强度, 需要了解它与钢水中碳的含量之间的关系。为此收集了 10 组不同的碳含量 x 及所对应的钢强度 y 数据 $(x_i, y_i) (i = 1, 2, \dots, 10)$ 如下:

$x/\%$	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.12	0.15	0.17	0.20
y/MPa	40.5	39.5	41.0	41.5	43.0	42.0	45.0	47.5	45.0	47.5	53.0	56.0

试选 x 与 y 之间的关系式。

解: 将上列数据点画在坐标纸上 (见图 7.1), 称为散点图。

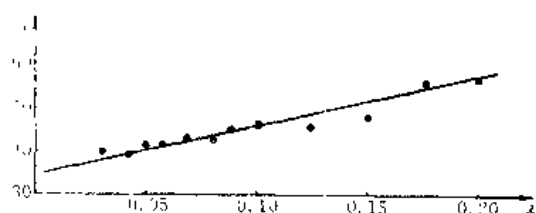


图 7.1 散点图

由上面的散点图可以看出, 这些点虽然是散乱着的, 但它们大体上是散布在一直线周围, 呈带状。我们用直线拟合, 认为 x 与 y 之间的关系近似是线性的, 设关系式为

$$\hat{y} = a + bx \quad (7.1)$$

这里, \hat{y} 是为了区别带有随机误差的观测值 y_i 。由于 y 与 x 之间实际上并不具有确定的函数关系, 式 (7.1) 只反映了 x 与 y 之间的一种

近似关系, 故称式 (7.1) 为经验公式, 其中 a 为截距, b 为一次项系数。

7.2.2 最小二乘法

显然, 当 a 、 b 取不同的数值时, 我们可以得到不同的方程。 a 、 b 怎样选取时, 所得到的方程才能最确切地反映出 x 与 y 之间的相关关系则是人们十分关心的问题。

给定平面上 n 个点 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 对于平面上任一条直线

$$\hat{y} = a + bx$$

我们可以用数量

$$[y_i - (a + bx_i)]^2 \quad i = 1, 2, \dots, n$$

来表示点 (x_i, y_i) 到直线 $\hat{y} = a + bx$ 的远近程度。那么, 欲使所找的直线最接近这 n 个点, 就要求出使得

$$Q(a, b) = \sum_{i=1}^n [y_i - (a + bx_i)]^2 \quad (7.2)$$

达到最小的 a 、 b 值, 分别记作 \hat{a} 、 \hat{b} , 此时方程 $\hat{y} = \hat{a} + \hat{b}x$ 称为回归方程。由于 $Q(a, b)$ 是 n 个平方数之和, 所以使 $Q(a, b)$ 达到最小的原则, 称为平方和最小原则或最小

乘法规则。

根据二元函数极值存在的必要条件, 有

$$\begin{cases} \frac{\partial Q}{\partial a} = \sum_{i=1}^n 2[y_i - (a + bx_i)](-1) = 0 \\ \frac{\partial Q}{\partial b} = \sum_{i=1}^n 2[y_i - (a + bx_i)](-x_i) = 0 \end{cases} \quad (7.3)$$

解这个方程组, 得

$$\begin{cases} \hat{b} = \frac{S_{XY}}{S_{XX}} \\ \hat{a} = \bar{Y} - \hat{b}\bar{X} \end{cases} \quad (7.4)$$

其中

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{i=1}^n x_i \\ \bar{Y} &= \frac{1}{n} \sum_{i=1}^n y_i \\ S_{XX} &= \sum_{i=1}^n (x_i - \bar{X})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \\ S_{XY} &= \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) = \sum_{i=1}^n x_i y_i - \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right) \end{aligned}$$

例7.2 用最小二乘法求例7.1的回归方程。

解: 由样本值计算得 $\bar{X} = 0.102$, $\bar{Y} = 44.9$

$$\begin{aligned} \hat{b} &= \frac{\sum_{i=1}^{10} x_i y_i - \frac{1}{10} \left(\sum_{i=1}^{10} x_i \right) \left(\sum_{i=1}^{10} y_i \right)}{\sum_{i=1}^{10} x_i^2 - \frac{1}{10} \left(\sum_{i=1}^{10} x_i \right)^2} = 92.64 \\ \hat{a} &= \bar{Y} - \hat{b}\bar{X} = 35.45 \end{aligned}$$

故得回归方程为

$$\hat{y} = 35.45 + 92.64x$$

7.2.3 相关性检验

由前面的分析可知, 应用最小二乘法找回归直线时, 并不依赖于 x 与 y 之间是否存在线性相关关系, 也就是说, 只要有 n 对数据, 通过式 (7.4), 就能找到一条回归直线。但是, 当 x 与 y 之间不存在线性相关关系时, 所求出的回归方程是毫无意义的。因此, 在求出回归方程时, 我们有必要判断 x 与 y 之间是否确实存在线性相关关系, 或说有必要检验 x 与 y 之间的线性相关关系是否显著。不难看出, 式 (7.1) 中, b 反映了 y 随 x 线性变化的变化率, 若 $b \neq 0$, 则意味着 y 随 x 作线性变化, x 与 y 之间线性关系显著; 反之, 若 $b = 0$, 则 y 不随 x 作线性变化, 所求出的回归方程没有实际意义。因此, 对回归方程进行线性相关显著性检验, 就是要检验假设

$$H_0: b = 0 \quad (7.5)$$

是否成真, 为此我们需要找一个用来作检验的统计量, 且当它在式 (7.5) 为真时服从一个

分布。根据所构造统计量的不同，回归方程的相关性检验一般有下列两种方法。

1. F 检验法

F 检验是从研究因变量 y 的变化规律开始的。通常，在 n 对测量数据中， y 总是变动的， y 值的这种变动称为变差。它的大小可以通过各次的测量与其均值的偏差 $(y_i - \bar{Y})$ 来表示，而全部 n 次测量值的总偏差，则可由这些偏差的平方和 $\sum_{i=1}^n (y_i - \bar{Y})^2$ 来表示。引起 y 产生变差的原因不外乎是：1) 式 (7.5) 不真；2) 其他一些因素所造成的随机误差影响。为弄清哪个因素对 y 产生变差影响更大，需对偏差平方和进行分解。

偏差平方和分解公式：对于任意 n 对实验数据 (x_i, y_i) ，恒有

$$\sum_{i=1}^n (y_i - \bar{Y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{Y})^2 \quad (7.6)$$

其中 $\hat{y}_i = \hat{a} + \hat{b}x_i$ 。令

$$L = \sum_{i=1}^n (y_i - \bar{Y})^2, Q = \sum_{i=1}^n (y_i - \hat{y}_i)^2, U = \sum_{i=1}^n (\hat{y}_i - \bar{Y})^2,$$

则有

$$L = Q + U$$

显然， L 是 y 的偏差平方和，它的大小描述了 y 的 n 个实验数据的分散程度。 Q 则是各次测量值与由回归直线计算出的各相应回归值之偏差平方和，它的大小不是由于 x 不同而产生的，而是由试验误差与无法控制的因素所引起的偏差，称为剩余平方和（或残差平方和）。 U 是回归值 \hat{y}_i 与它的平均值 \bar{Y} 之差的平方和，称为回归平方和，它的大小描述了 $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ 的分散程度。由于 $\hat{y}_i = \hat{a} + \hat{b}x_i$ ， \bar{Y} 是常数，所以 U 所描述的 $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ 的分散程度主要来自 x_1, x_2, \dots, x_n 的分散性，可见 U 可以用来反映 x 对 y 的线性影响的大小。

通过以上分析，我们知道，可以把总的偏差平方和分解为 U 和 Q 两部分， Q 反映了随机误差所引起的观测值的波动， U 反映了由于 x, y 之间的线性相关关系所引起的波动。不难想到：要由样本值 (x_i, y_i) ($i=1, 2, \dots, n$) 分析 x, y 之间是否存在线性相关关系，可以用 U 与 Q 相比较，如果 $\frac{U}{Q}$ 充分大，则可以认为 x, y 之间存在着线性相关关系，反之，则无理由认为 x, y 之间存在着线性相关关系。

F 检验方法：可以证明，在假设 $H_0: b=0$ （表示线性关系不显著）成立的条件下，统计量

$$F = \frac{U}{Q/(n-2)} \quad (7.7)$$

服从自由度为 1 和 $n-2$ 的 F 分布，记作

$$F = \frac{U}{Q/(n-2)} \sim F(1, n-2) \text{ 分布}$$

可利用 F 分布，根据 F 值的大小进行检验。为此，已制成 F 分布上侧临界值表，表 7.1 是该表的一部分。

F 检验法的具体步骤：

(1) 提出待定假设 $H_0: b=0$ ；

(2) 对于给定的检验水平 α （通常取 $\alpha=0.05, 0.01$ 或 0.10 ），由

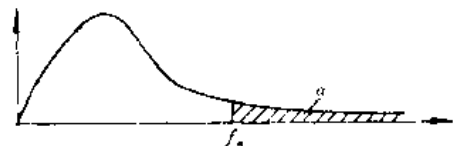
$$P\{F > f_{\alpha}(1, n-2)\} = \alpha \quad (7.8)$$

查第一自由度 $f_1 = 1$, 第二自由度 $f_2 = n - 2$ 时 F 分布表中的上侧临界值 $f_{\alpha}(1, n - 2)$;

(3) 由给定的样本值计算 $F = \frac{U}{Q/(n-2)}$ 之值。

表 7.1 F 分布上侧临界值表

$P\{F > f_{\alpha}(f_1, f_2)\} = \alpha$



$f_1 \backslash f_2$	1	2	3	4	5	6	8	10	12	20	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	241.9	243.9	248.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.40	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.85	8.79	8.74	8.66	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.96	5.91	5.80	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.74	4.68	4.56	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.06	4.00	3.87	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.64	3.57	3.44	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.35	3.28	3.15	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.14	3.07	2.94	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.98	2.91	2.77	2.54
11	4.84	3.98	3.59	3.36	3.20	3.07	2.95	2.85	2.79	2.65	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.85	2.75	2.69	2.54	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.77	2.67	2.60	2.46	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.60	2.53	2.39	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.67	2.54	2.48	2.33	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.49	2.42	2.28	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.45	2.38	2.23	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.41	2.34	2.19	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.38	2.31	2.16	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.35	2.28	2.12	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.32	2.25	2.10	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.30	2.23	2.07	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.37	2.27	2.20	2.05	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.25	2.18	2.03	1.73
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.16	2.09	1.92	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.08	2.00	1.84	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.10	1.99	1.92	1.75	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.91	1.82	1.66	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	1.94	1.83	1.75	1.57	1.00

若 $F > f_{\alpha}(1, n - 2)$, 则否定 $H_0: b = 0$, 表明 x 、 y 之间线性关系显著; 反之, 则应接受 $H_0: b = 0$, 表明 x 、 y 之间线性关系不显著。

通常, 直接按定义计算 U 和 Q 很麻烦, 实践中常采用变形后的式子简化计算, 其计算式为

$$Q = S_{YY} - U$$

$$U = bS_{XY} \quad (7.9)$$

其中 $S_{YY} = \sum_{i=1}^n (y_i - \bar{Y})^2 = \sum_{i=1}^n y_i^2 - \frac{1}{n} (\sum_{i=1}^n y_i)^2$

请读者完成取 $\alpha = 0.05$ 时, 对例 7.2 中回归方程的检验。

2. R 检验法 (或相关系数检验法)

为了检验相关性, 有时也选用样本相关系数

$$R = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad (7.10)$$

为统计量, R 的绝对值越接近 1, 表示两者线性关系越密切, 且当 $R > 0$ 时称为正相关, $R < 0$ 时称为负相关; R 的绝对值趋于零, 两者没有线性相关关系。

在实际应用中还应该注意, R 接近 1 的程度与样本容量 n 有很大关系, 特别当 $n = 2$ 时, 由 2 点决定一条直线, 所以必有 $|R| = 1$, 但这并不意味着相关关系十分密切, 因而应进一步作相关显著性检验。在许多统计书中都有相关系数表可查 (见表 7.2), 当 R 的绝对值大于表中相应临界值时, 才认为 x 、 y 之间相关关系显著。

表 7.2 检验相关系数的临界值表

$P (R > r_\alpha) = \alpha$					
$\alpha \backslash f$	0.10	0.05	0.02	0.01	0.001
1	0.98769	0.99692	0.999507	0.999877	0.9999988
2	0.90000	0.95000	0.98000	0.99000	0.99900
3	0.8054	0.8783	0.93433	0.95873	0.99116
4	0.7293	0.8114	0.8822	0.91720	0.97406
5	0.6694	0.7545	0.8329	0.8345	0.95074
6	0.6215	0.7067	0.7887	0.8743	0.92493
7	0.5822	0.6664	0.7498	0.7977	0.8982
8	0.5494	0.6319	0.7155	0.7646	0.8721
9	0.5214	0.6021	0.6851	0.7348	0.8471
10	0.4933	0.5760	0.6581	0.7079	0.8233
11	0.4762	0.5529	0.6339	0.6835	0.8010
12	0.4575	0.5324	0.6120	0.6674	0.7800
13	0.4409	0.5139	0.5923	0.6411	0.7603
14	0.4259	0.4973	0.5742	0.6226	0.7420
15	0.4124	0.4821	0.5577	0.6055	0.7246
16	0.4000	0.4683	0.5425	0.5897	0.7084
17	0.3887	0.4555	0.5285	0.5751	0.6932
18	0.3783	0.4438	0.5155	0.5614	0.6787
19	0.3687	0.4329	0.5034	0.5487	0.6652
20	0.3598	0.4227	0.4921	0.5368	0.6524
25	0.3233	0.3809	0.4451	0.4869	0.5974
30	0.2960	0.3494	0.4093	0.4487	0.5541
35	0.2746	0.3246	0.3810	0.4182	0.5189
40	0.2573	0.3044	0.3578	0.3932	0.4896
45	0.2428	0.2875	0.3384	0.3721	0.4648
50	0.2306	0.2732	0.3218	0.3541	0.4433
60	0.2108	0.2500	0.2948	0.3248	0.4078
70	0.1954	0.2319	0.2737	0.3017	0.3799
80	0.1829	0.2172	0.2565	0.2830	0.3568
90	0.1726	0.2050	0.2422	0.2673	0.3375
100	0.1638	0.1946	0.2301	0.2540	0.3211

例 7.3 对例 7.2 中的回归方程进行 R 检验 ($\alpha = 0.05$)。

解: 由样本值计算得 $R = 0.95$, 又由 $\alpha = 0.05$, 自由度 $f = n - 2 = 8$, 查表 7.2, 得临

界值 $\gamma_{0.05} = 0.632$

由于 $|R| > \gamma_{\alpha}$, 应否定 $H_0: b = 0$, 即这个温度与杂质含量间存在着较好的线性关系。事实上, R 与 F 之间存在着密切的关系, 这是由于

$$F = \frac{U}{Q/(n-2)} = \frac{(n-2)S_{XY}^2/S_{XX}}{S_{YY} - S_{XY}^2/S_{XX}} = \frac{(n-2)R^2}{1-R^2} \quad (7.11)$$

可见, F 的值较大, 等价于 $|R|$ 较大, 两种检验结果是一致的。

7.2.4 回归系数的区间估计

在一元线性回归函数中, \hat{b} 表示 \hat{y} 随 x 变化的变化率, 即回归系数, \hat{a} 表示在 $x = 0$ 时 \hat{y} 的值, 即回归常数。前面我们通过最小二乘法得到的 \hat{a} 、 \hat{b} 又称为经验公式 $\hat{y} = a + bx$ 中 a 、 b 的点估计量, 它随着样本值的不同会产生波动。也就是说, 单一参数的点估计量很可能与参数的真值不同, 因而有必要对点估计量的可靠性进行分析。其中一个方法就是对回归参数进行区间估计, 或说求回归参数的置信区间。

定义: 若用 $\hat{\theta}$ 作为参数 θ 的估计值 (或近似值), 给定一个非常小的正数 ϵ , 则当

$$P(|\hat{\theta} - \theta| \leq \epsilon) = 1 - \alpha \quad (7.12)$$

即

$$P(\hat{\theta} - \epsilon \leq \theta \leq \hat{\theta} + \epsilon) = 1 - \alpha$$

成立时, 称区间 $[\hat{\theta} - \epsilon, \hat{\theta} + \epsilon]$ 为参数 θ 置信区间, $1 - \alpha$ 为参数 θ 相对于置信区间 $[\hat{\theta} - \epsilon, \hat{\theta} + \epsilon]$ 的置信概率 (或置信度)。或简称区间 $[\hat{\theta} - \epsilon, \hat{\theta} + \epsilon]$ 是参数 θ 关于置信概率为 $1 - \alpha$ 的置信区间。

由数理统计得知

$$\hat{a} \sim N\left(a, \left(\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}}\sigma^2\right)\right) \quad (7.13)$$

$$\hat{b} \sim N\left(b, \frac{\sigma^2}{S_{XX}}\right) \quad (7.14)$$

为叙述方便, 记 \hat{a} 、 \hat{b} 的标准差分别为 σ_a 与 σ_b , 则

$$\sigma_a = \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}}}\sigma \quad \sigma_b = \frac{\sigma}{\sqrt{S_{XX}}}$$

其估计值为

$$\hat{\sigma}_a = \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}}}S \quad \hat{\sigma}_b = \frac{S}{\sqrt{S_{XX}}}$$

其中 $S = \sqrt{\frac{Q}{n-2}}$, 由数理统计得知,

$$\frac{\hat{a} - a}{\hat{\sigma}_a} \quad (7.15)$$

服从自由度为 $n-2$ 的 t 分布 (见表 7.3), 由

$$P\left(\left|\frac{\hat{a} - a}{\hat{\sigma}_a}\right| \leq t_{\alpha}\right) = 1 - \alpha \quad (7.16)$$

得到 a 的关于置信度为 $1 - \alpha$ 的置信区间为:

$$[\hat{a} - t_{\alpha}(n-2)\hat{\sigma}_a, \hat{a} + t_{\alpha}(n-2)\hat{\sigma}_a] \quad (7.17)$$

同理, 可得到 b 的关于置信度为 $1-\alpha$ 的置信区间为:

$$[\hat{b} - t_{\alpha}(n-2)\hat{\sigma}_b, \hat{b} + t_{\alpha}(n-2)\hat{\sigma}_b] \quad (7.18)$$

例 7.4 求例 7.2 中 a 、 b 的关于置信度为 $1-\alpha=0.95$ 的置信区间。

解: 由样本值计算得:

$$\hat{a} = 34.45$$

$$\hat{b} = 92.64$$

$$S = \sqrt{\frac{Q}{n-2}} = 1.8450$$

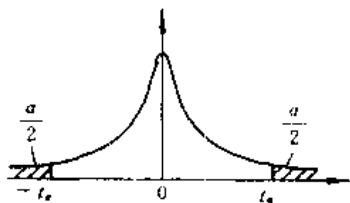
$$\hat{\sigma}_a = 1.24$$

$$\hat{\sigma}_b = 10.75$$

当 $1-\alpha=0.95$ 时, 由表 7.3 查得 $t_{0.05}(8) = 2.306$, 从而 a 与 b 关于置信度为 $1-\alpha=0.95$ 的置信区间分别为

$$[32.59, 38.31] \text{ 及 } [67.85, 117.43]$$

表 7.3 t 分布双侧临界值表

$P(|t| > t_{\alpha}) = \alpha$


t	α	0.5	0.1	0.05	0.02	0.01	0.001
1		1.000	6.314	12.706	31.821	63.657	636.619
2		0.816	2.920	4.303	6.965	9.925	31.598
3		0.765	2.353	3.182	4.541	5.841	12.924
4		0.741	2.132	2.776	3.747	4.604	8.610
5		0.727	2.015	2.571	3.365	4.032	6.859
6		0.718	1.943	2.447	3.143	3.707	5.959
7		0.711	1.895	2.365	2.998	3.499	5.405
8		0.706	1.806	2.306	2.896	3.355	5.041
9		0.703	1.833	2.262	2.821	3.250	4.781
10		0.700	1.812	2.228	2.764	3.169	4.587
11		0.697	1.796	2.201	2.718	3.106	4.437
12		0.695	1.782	2.179	2.681	3.055	4.318
13		0.694	1.771	2.160	2.650	3.012	4.221
14		0.692	1.761	2.145	2.624	2.977	4.140
15		0.691	1.753	2.131	2.602	2.947	4.073
16		0.690	1.746	2.120	2.583	2.921	4.015
17		0.689	1.740	2.110	2.567	2.898	3.965
18		0.688	1.734	2.101	2.552	2.878	3.922
19		0.688	1.729	2.093	2.539	2.861	3.883
20		0.687	1.725	2.086	2.528	2.845	3.850

7.2.5 预测与控制

建立了回归方程, 并确认其有意义后, 就可以利用这个方程进行预测与控制, 这是建立回归方程的主要目的。

1. 预测

对于回归方程 $\hat{y} = \hat{a} + \hat{b}x$, 如果再给定一个值 x_0 , 将其代入方程, 就可以得到相应的 \hat{y}_0 , 这就是对 y 的预测值。但是, 由于这个回归方程是根据样本数据确定的, 所得到的预测值会有抽样误差。因此, 预测时一般都是通过估计值 \hat{y} 找出 y 真值的置信区间。

可以证明:

$$\frac{y_0 - \hat{y}_0}{\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} S} \quad (7.19)$$

服从自由度为 $n-2$ 的 t 分布。对于给定的置信度 $1-\alpha$, 查自由度为 $n-2$ 的 t 分布表, 可以得到临界值 t_α , 使

$$P \left\{ \left| \frac{y_0 - \hat{y}_0}{\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} S} \right| \leq t_\alpha \right\} = 1 - \alpha \quad (7.20)$$

由此, 可以得到 y_0 的关于置信度为 $1-\alpha$ 的置信区间为:

$$\left[\hat{y}_0 - t_\alpha \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} S, \hat{y}_0 + t_\alpha \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} S \right] \quad (7.21)$$

由式 (7.21) 可以看到, 当 $x_0 = \bar{X}$ 时, 区间长度最短, 即 x_0 越靠近 \bar{X} , 则预测值 y_0 的精确度就越高。反之 x_0 离 \bar{X} 越远, 置信区间越长, 预测值 y_0 的精确度也越低。预测区间的上限与下限的曲线对称地落在回归直线的两侧, 而呈喇叭形 (见图 7.2)。

特别, 当 n 较大, 且 x_0 较接近 \bar{X} 时,

$$\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} \approx 1$$

预测区间式 (7.21) 近似为:

$$[\hat{y}_0 - t_\alpha S, \hat{y}_0 + t_\alpha S] \quad (7.22)$$

又当 n 较大时, 由数理统计知, 自由度为 $n-2$ 的 t 分布近似于 $N(0, 1)$, 故 y_0 的置信度为 0.95 的预测区间为:

$$[\hat{y}_0 - 1.96S, \hat{y}_0 + 1.96S] \quad (7.23)$$

同理可有 y_0 的置信度为 0.99、0.90 的置信区间分别为:

$$[\hat{y}_0 - 1.58S, \hat{y}_0 + 1.58S] \quad (7.24)$$

及

$$[\hat{y}_0 - 1.64S, \hat{y}_0 + 1.64S] \quad (7.25)$$

例7.5 在例7.2中, 当 $x_0 = 0.10$ 时, 试预测相应的 y 值。($1-\alpha=0.95$)

解: 将 $x_0 = 0.10$ 代入例7.2 所求出的回归方程中, 得

$$\hat{y}_0 = 35.45 + 92.64 \times 0.10 = 44.714$$

又 $S = 1.8540$

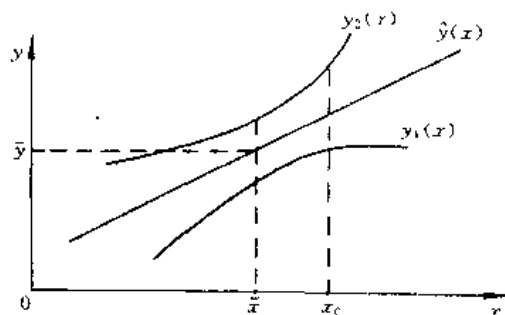


图 7.2 预测曲线

$$\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} = \sqrt{1 + \frac{1}{10} + \frac{(0.10 - 0.102)^2}{0.02976}} = 1.048$$

查 t 分布临界值得 $t_{0.05}(8) = 2.306$, 于是

$$t_{\alpha} \cdot S \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{S_{XX}}} = 2.306 \times 1.8540 \times 1 = 4.485$$

得产品杂质含量关于置信度为 0.95 的置信区间为:

$$[44.714 - 4.485, 44.714 + 4.485]$$

即

$$[40.229, 49.119]$$

2. 控制

控制是预测的反问题。即当 y 的值落在某个指定区间 $[y_1, y_2]$ 内时, 寻找应将 x 控制在什么范围内的问题。

由于式 (7.21) 区间两端点计算过于繁冗, 限制了它的推广和应用。为简化计算, 便于应用, 当 n 较大时, 可根据给定的置信度 (如 $1 - \alpha = 0.95$), 采取相应的计算式 (如 7.23 式), 且由

$$\begin{cases} y_1 = a + bx_1 - 1.96S \\ y_2 = a + bx_2 + 1.96S \end{cases}$$

解得

$$\begin{cases} x_1 = \frac{1}{b}(y_1 - a + 1.96S) \\ x_2 = \frac{1}{b}(y_2 - a + 1.96S) \end{cases} \quad (7.26)$$

当 $b > 0$ 时, x 的控制范围是区间 $[x_1, x_2]$; $b < 0$ 时, x 的控制范围是区间 $[x_2, x_1]$ 。在实际应用时, 常在平面上作两条平行于回归线的直线 (见图 7.3、图 7.4)。利用这两条直线就可以进行有关的预测与控制。显然, 要实现控制, 在此处必须使区间 $[y_1, y_2]$ 的长度不小于 $3.92S$ 。

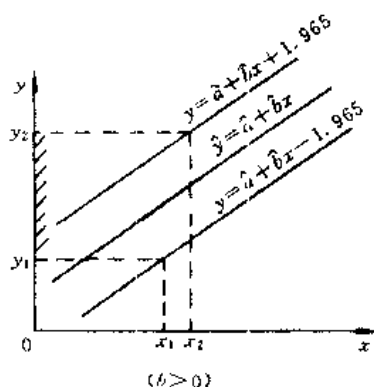


图 7.3

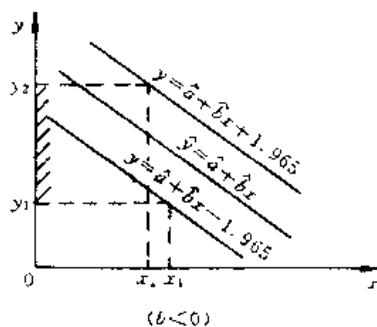


图 7.4

7.3 多元线性回归

前面介绍了一元线性回归问题，然而客观现象是比较复杂的，社会经济现象尤其如此，一种现象往往是多种因素综合作用的结果。例如，某种商品的销售量可能受人口、收入水平、消费习惯、产品质量、价格等因素的影响，某种化工厂产品的质量可能受原材料质量、配方比例、温度、湿度、压力等因素的影响，农产品的质量受栽培技术、品种、土质、肥料、气温等因素的影响。

一般来说,在进行回归预测时,如果能够考虑到多种有关因素的影响,预测效果往往会更好。研究两个以上变量的回归问题称为多元回归。和一元回归类似,多元回归也分为多元线性回归和多元非线性回归两大类。由于许多多元非线性回归问题都可以化为多元线性回归问题来解决。因此,本节首先讨论多元线性回归,具体分析方法与原理和一元线性回归相同,只是表达式更复杂,计算量更大。

7.3.1 多元线性回归的数学模型

如果因变量 y 与自变量 x_1, x_2, \dots, x_m 之间有线性相关关系存在, 即

$$y = b_0 + b_1x_1 + \cdots + b_mx_m + \epsilon \quad (7.27)$$

则由 n 组试验数据:

$$(x_{i1}, x_{i2}, \dots, x_{im}, y_i) \quad i = 1, 2, \dots, n$$

可以得到多元线性回归模型:

[illegible]

其中 b_0, b_1, \dots, b_m 是 $m + 1$ 个待定参数; $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是误差项, 它是 n 个独立且服从同一分布 $N(0, \sigma^2)$ 的随机变量。

我们的目的,是要通过这 n 组样本数据,将参数 b_0, b_1, \dots, b_m 估计出来。

7.3.2 最小二乘法

一般使用最小二乘法求 b_0, b_1, \dots, b_m 的估值 $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_m$ 。为此, 对于给定的样本数据 $(y_i; x_{i1}, x_{i2}, \dots, x_{im})$, $i = 1, 2, \dots, n$, 设其回归方程为

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_mx_m$$

利用最小二乘法求出回归系数 b_0, b_1, \dots, b_m , 并且应该使全部观测值 y_i 与回归值 \hat{y} 的偏差平方和 Q 达到最小, 即使

$$Q = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{i1} - b_m x_{im})^2 \quad (7.29)$$

达到最小。

由数学分析中的极值原理可知, 连续可导函数如果有极值存在, 则在极值点的导数等于零。参数 b_0, b_1, \dots, b_m 应是下列方程组的解。

作相关性检验。显然, R 越接近 1, 线性关系越显著; R 越接近于零, 线性关系越不显著。线性关系显著的程度, 可与表 7.2 中的临界值 γ_α 相比较。此时, 对于指定的检验水平 α , 取自由度 $f = n - m - 1$, 查得临界值 γ_α , 且当 $R > \gamma_\alpha$ 时, 则否定 H_0 , y 与 x_1, x_2, \dots, x_m 间线性关系显著; 当 $R \leq \gamma_\alpha$ 时, 则接受 H_0 , y 与 x_1, x_2, \dots, x_m 间线性关系不显著。

7.3.4 逐步回归

在许多实际问题中, 涉及的变量个数常常很多, 给计算及方程的稳定性都带来了不利因素。为此, 我们总希望在回归方程中包含的只是那些必不可少的变量 (即对因变量影响最显著的变量)。这样的回归方程用起来比较方便, 预报的精度也比较高, 称作“最优”回归方程。

对回归方程调优, 选择“最优”回归方程的方法很多, 而其中最好的方法就是逐步回归法。逐步回归法的基本思想是用偏回归平方和 (即去掉某变量 x_i 后 y 与 $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_m$ 间的回归平方和, 称为关于 x_i 的偏回归平方和) 选择最优回归方程。其具体方法是由大到小逐个将因子引入回归方程, 且随时用偏回归平方和作 F 检验, 显著时才引入, 而一旦新的变量引入后又要对老的变量重新检验, 若它变成不显著的就要将它从方程中剔除。直到没有变量可剔除, 也没有变量可引入时为止。最后, 对所选上的变量建立回归方程。

为方便读者使用, 已将逐步回归编入化工过程计算 (BIPTCHEM) 软件包, 使用方法见第四篇及例 7.6。

例 7.6 测的一组实验数据如下表:

r	-8	-7	-6	-5	-4	-3	-2	-1	
y	.000335	.00091	.00247	.00669	.01799	.047	.119	.269	
r	0	1	2	3	4	5	6	7	8
y	.5	.7311	.881	.953	.982	.993	.998	.999	.99966

试用 BIPTCHEM 软件包进行回归计算。

解: 先在 DOS 状态下启动计算机, 再运行 UC DOS, 然后进入 BIPTCHEM 软件包所在的子目录 (/BIPTCHEM 目录), 键入 RUN, 按 <Enter> 键。屏幕显示主页后按 “C” 键, 屏幕将出现以下菜单:

物性	数据处理	优化	精馏	换热器	板式塔	反应器

用方向键 (←键或→键) 将光标移至数据处理上, 使其变红, 按 <Enter> 键, 数据处

理菜单打开，屏幕显示：

非线性回归
逐步回归

用方向键（↑键或↓键）将光标移至逐步回归上，并按<Enter>键。屏幕询问是否打印（若要打印按Y键，否则按N键）。接着询问是否学习例题（若要学习按Y键，否则按N键），屏幕显示逐步回归主菜单：

1. 输入数据
2. 数据预处理
3. 回归计算
4. 输出计算结果
Q. 退出

用方向键（↑键或↓键）菜单，并按<Enter>键执行菜单。选择并执行输入数据后，屏幕显示：

1. 磁盘输入
2. 键盘输入
Q. 回到主菜单

选择键盘输入。根据屏幕提示，输入数据组数（17）、独立变量数（1）、因变量名（ y ），自变量名（ x ）。然后，依次输入各 x 、 y 的值。回至主菜单。

选择主菜单中的数据预处理，屏幕显示：

1. 显示数据
2. 检查数据
3. 打印数据
4. 数据存盘
5. 图示数据
Q. 回到主菜单

选择数据存盘，机器将询问要存的文件名，输入文件名（如 e1）。回到主菜单（其他各菜单请读者试用）。

选择主菜单中的回归计算，屏幕显示可选的三种方程。选择第三方程，询问 F 检验临界值：

1. 取 0.01
2. 0.05
3. 保持方程中每一自变量
Q. 退出

本例只有一个自变量，因此 F 检验临界值意义不大，可任选。计算机即行回归计算，计算完后回到主菜单。

选择主菜单中的输出计算结果，屏幕显示输出菜单。选择输出方程时，输出以下结果：

$$1/y = 1 + \exp[-(a+bx)]$$

$$a = 3.3365E-3 \quad -3.1506E-2$$

$$b = 1.009 \quad -6.4317E-3$$

复相关系数 $R = 0.9999318$ 均方差 $S = 0.06096$

其他各项输出由读者自己试做。

例 7.7 一组试验数据已存入磁盘中 YL1.SSS 的文件中（用键盘输入和存入磁盘的方法见例 7.6）。试用 BIPTCHEM 软件完成回归计算。

解：按例 7.6 方法运行 BIPTCHEM 软件包，选择磁盘输入，键入文件名（YL1）。选择主菜单的回归计算后，选择方程 1；屏幕上将询问“最高幂级数？”时，输入“1”；屏幕显示：

变换因变量

1. 退出(不变)
2. 取对数, 即 $y' = \lg y$
3. 取倒数, 即 $y' = 1/y$
- Q. 退出

选择“1. 退出(不变)”；屏幕又显示：

变换自变量：

1. 退出(不变)
2. 取 $x' = \lg x$
3. 取 $x' = 1/x$
4. 取 $x' = \exp(-x)$
5. 变换任一自变量
- Q. 退出

选择“1. 退出(不变),”在 F 检验临界值中，可选择不同的临界值进行计算比较。例如：选择第 1 项（即取 $F = 0.01$ ）后，可得到计算结果如下：

$$y = a + bx_1 + ex_4$$

$$a = 103.0974 \quad -4.214917$$

$$b = 1.439958 \quad -0.295989$$

$$e = -0.6139535 \quad 0.1050649$$

复相关系数: $R = 0.9861394$ 均方差: $S = 2.73427$

若 F 检验临界值选第 4 项时，则结果中将保留所有自变量。

7.4 非线性回归

实践中，变量之间的相关关系并不仅是线性的，更广泛存在的是非线性相关关系。在解决非线性相关关系的问题中，我们会遇到两类问题：一类是可线性化的非线性回归问题；另一类则是不可线性化的非线性回归问题。对于可线性化的非线性回归问题，一般是通过适当的变量代换即可将其非线性关系化为线性关系，而后按线性回归处理。

7.4.1 可线性化的一元非线性回归

对于一元函数可由散点图拟合的函数类型来选择可线性化的变量代换。下面我们介绍几种常见类型。

1. 对数函数型

若散点图如图 7.5 所示, 可用对数函数 $Y = a + b \log x$ ($x > 0$) 型曲线拟合。令 $X = \log x$, 原函数可线性化为

$$Y = a + bX$$

2. 幂函数型

若散点图如图 7.6 所示, 可用幂函数 $y = ax^b$, ($x > 0$, a, b 为常数, 且 $a > 0$) 型曲线拟合。

将原式两边取对数, 可得 $\ln y = \ln a + b \ln x$, 令 $Y = \ln y$, $A = \ln a$, $X = \ln x$, 原函数可线性化为

$$Y = A + bX$$

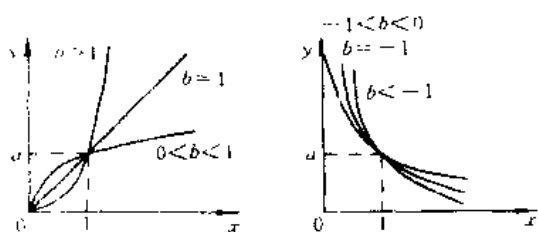


图 7.6 幂函数型曲线

4. 倒指数函数型

若散点图如图 7.8 所示, 可用倒指数函数 $y = ae^{b/x}$ 型曲线拟合。

将原式两边取对数, 可得 $\ln y = \ln a + b/x$ 。令 $Y = \ln y$, $A = \ln a$, $X = 1/x$, 则原函数可线性化为

$$Y = A + bX$$

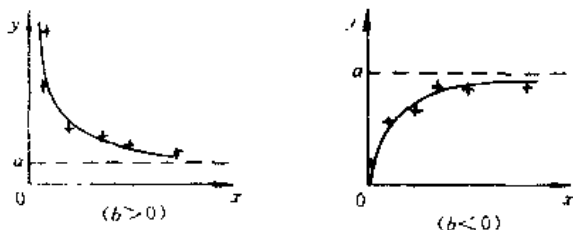


图 7.8 倒指数函数型曲线

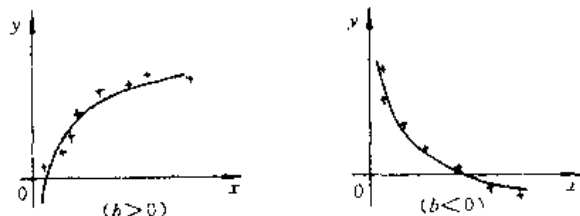


图 7.5 对数函数曲线

3. 指数函数型

若散点图如图 7.7 所示, 可用指数函数 $y = ae^{bx}$ ($a > 0$) 型曲线拟合。将原式两边取对数, 有 $\ln y = \ln a + bx$, 令 $Y = \ln y$, $A = \ln a$, 则原函数可线性化为

$$Y = A + bx$$

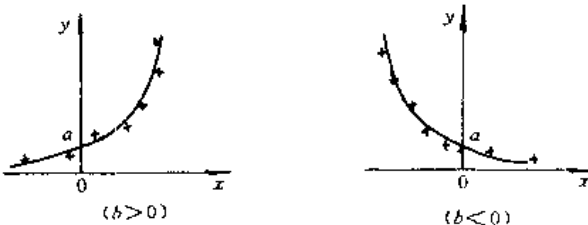


图 7.7 指数函数型曲线

5. 双曲线函数

若散点图如图 7.9 所示, 可用双曲线函数 $1/y = a + b/x$ 型曲线拟合。

若令 $Y = 1/y$, $X = 1/x$, 则原函数可线性化为

$$Y = a + bX$$

6. S 型曲线函数

若散点图如图 7.10 所示, 可用 S 型曲线函数 $y = 1 / (a + be^{-x})$ 曲线拟合。原方程可写为: $1/y = (a + be^{-x})$ 。令 $Y = 1/y$, X

$= e^{-t}$, 则原函数可线性化为

$$Y = a + bX$$

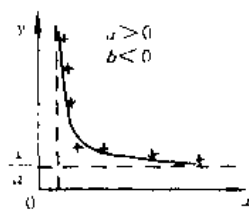


图 7.9 双曲线函数型曲线

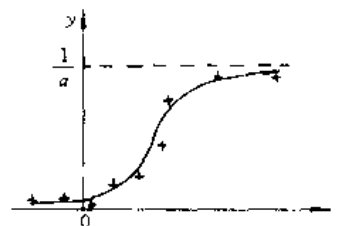
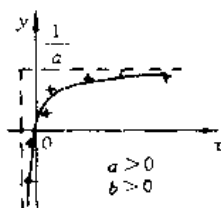


图 7.10 S型曲线

例 7.8 在绝对温度 T 时, 溶于 100g 水中的无水氯化铵为 S , 得到饱和溶液, 根据实验 T 与 S 的关系如下表:

T/K	273	283	288	293	313	333	353	373
S/g	29.4	33.3	35.2	37.2	45.8	55.2	65.6	77.3

求 T 与 S 的经验公式。

解: 运行 BIPTCHEM 软件包, 选择数据处理中的逐步回归计算 (所有线性回归及可线性化的非线性回归问题, 都可用逐步回归计算)。

用键盘输入数据 (方法同例 7.6) 后, 选择并执行数据预处理菜单中的图示数据, 分别用 $x-y$ 坐标系和 $\lg x-\lg y$ 坐标系进行图示, 它们的散点图见图 7.11, 可以看出经验曲线为幂函数型。

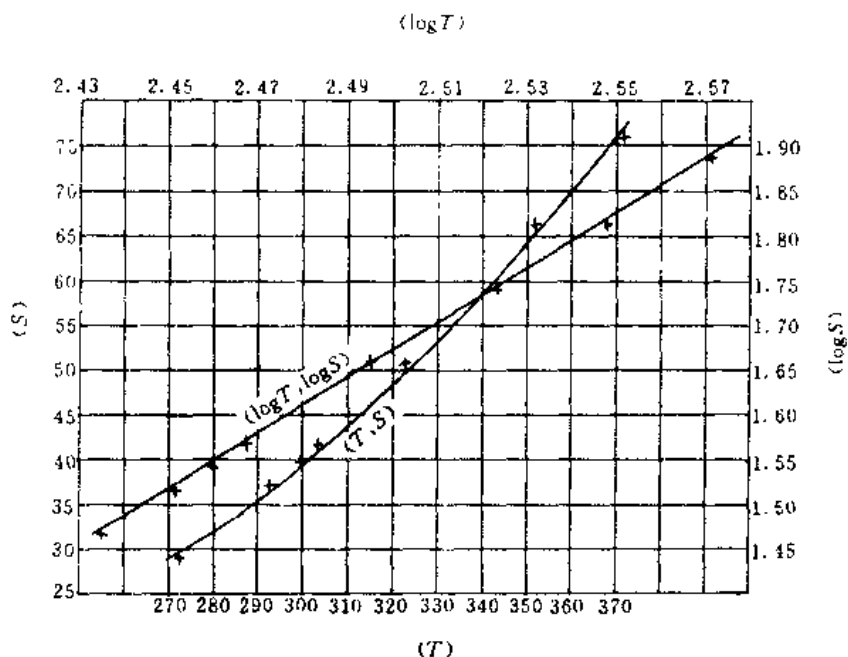


图 7.11 散点图

选择主菜单的回归计算, 输入最高幂级数 (1), 变换因变量时选择取对数 ($y' = \lg y$), 变换自变量时选择取 $x' = \lg x$ 。选择主菜单的输出计算中的方程后, 输出:

$$\begin{aligned}\ln S &= a + b \ln T \\ a &= -13.88015 + -0.3679 \\ b &= 3.0794 + -6.4066E-2 \\ \text{复相关系数 } R &= 0.99978 \quad \text{均方差 } S = 7.8145E-3\end{aligned}$$

7.4.2 多项式回归

在一元回归问题中, 如果回归方程是一个 K 次多项式

$$y = b_0 + b_1x + b_2x^2 + b_kx^k \quad (7.32)$$

则可令 $x_1 = x$, $x_2 = x^2$, $x_k = x^k$, 原 k 次多项式回归就转化为 k 元线性回归。即

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

这就是多项式回归。一般来说, 对于任意 n 对数据 (x_i, y_i) , 总可以用一个 k 次多项式来逼近。

例 7.9 已知某种成品在生产过程中的废品率 y 与它的化学成分 x 有关。现共有 16 组数据, 对此进行回归处理。

y	1.30	1.0	0.73	0.9	0.81	0.7	0.6	0.5	0.44	0.56	0.3	0.42	0.35	0.4	0.41	0.6
x	34	36	37	38	39	39	39	40	40	41	42	43	43	45	47	48

解: 运行 BIPTCHEM 软件包, 选择数据处理中的逐步回归计算。用键盘输入数据后, 选择并执行数据预处理菜单中的图示数据, 用 $x-y$ 坐标系进行图示, 散点图见图 7.12。从图中可以看出废品率最初随着成分的增加而降低; 当成分超过一定值后, 废品率又开始上升, 可用二次多项式(抛物线)拟合。

选择主菜单的回归计算, 选择方程 (1), 输入最高幂级数 (2), 变换因变量和自变量时均选择不变, F 检验临界值取 0.01。选择主菜单的输出计算中的方程后, 输出:

$$\begin{aligned}y &= a + bx + cx^2 \\ a &= 18.26349 + -0.6627 \\ b &= -0.3097 + -1.4541E-2 \\ c &= 9.1699E-3 + -1.7588E-4 \\ \text{复相关系数 } R &= 0.9387 \quad \text{均方差 } S = 0.10018\end{aligned}$$

多项式回归又称为多元线性的多项式回归, 除此而外还有多元非线性的多项式回归, 这种回归处理起来更为复杂, 实践中一般只用二次多项式来回归。例如, 有 m 个变量, 样本数据为 $y_i; x_{i1}, x_{i2}, \cdots, x_{im}$, 若利用二次多项式来回归时, 就要建立回归方程

$$\begin{aligned}y &= b_0 + b_1x_1 + \cdots + b_mx_m + b_{11}x_1^2 + \\ &\quad b_{12}x_1x_2 + \cdots + b_{1m}x_1x_m + \\ &\quad b_{22}x_2^2 + b_{23}x_2x_3 + b_{2m}x_2x_m + \\ &\quad \cdots b_{mm}x_m^2\end{aligned} \quad (7.33)$$

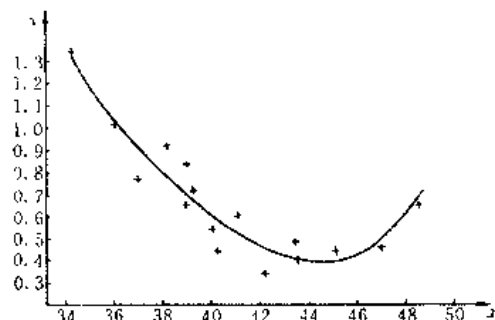


图 7.12 散点图

BIPTCHEM 中逐步回归中的方程 1 类似于式 (7.33)，只不过软件中方程的幂级数可以大于 2。随着变量个数 m 及幂级数的增加，回归方程中的项数增加非常快，计算量相当大。

7.4.3 可线性化的多元非线性回归

多元非线性回归的处理方法与一元函数类似，只是计算过程更复杂。

7.4.4 不可线性化的非线性回归

实践中，回归分析还常遇到一些不可线性化的非线性函数型。设非线性函数为 $y = f(x_i; a_1, a_2, \dots, a_k)$ ，进行回归分析时需要通过一组测量值 (x_i, y_i) ， $i = 1, 2, \dots, n$ ，来对该非线性函数中的参数 a_1, a_2, \dots, a_k 进行最小二乘估计。即应该寻找 a_1, a_2, \dots, a_k 使

$$Q = \sum_{i=1}^n [y_i - f(x_i; \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k)]^2$$

达到最小，需要解正规方程 (7.34)。

$$\sum_{i=1}^n [y_i - f(x_i; \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k)] \cdot \frac{\partial f(x_i; \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k)}{\partial a_j} = 0 \quad (j = 1, 2, \dots, k) \quad (7.34)$$

一般情况下，由于 a_j 在函数中以非线性形式出现，式 (7.34) 也是参数 a_j 的非线性方程组，不能直接求解。数学上可以将函数 $y = f(x_i; a_1, a_2, \dots, a_k)$ 对参数 a_j 作近似泰勒展开，使式 (7.34) 线性化，再用逐次迭代（牛顿-高斯迭代）法求解。由于计算过程十分繁杂，需要借助于电子计算机。

例 7.10 在化工生产中获得的氯气的组分 y 随生产时间 x 下降，现收集了 44 组数据，试求回归方程 $y = 1/(1 + e^{(a+bx)})$ 中的待定系数 a 和 b 。

i	x	y	i	x	y
1	8	0.49	23	22	0.41
2	8	0.49	24	22	0.40
3	10	0.18	25	24	0.42
4	10	0.47	26	24	0.40
5	10	0.48	27	24	0.40
6	10	0.47	28	26	0.41
7	12	0.46	29	26	0.40
8	12	0.46	30	26	0.41
9	12	0.45	31	28	0.41
10	12	0.43	32	28	0.40
11	14	0.45	33	30	0.40
12	14	0.43	34	30	0.40
13	14	0.43	35	30	0.38
14	16	0.44	36	32	0.41
15	16	0.43	37	32	0.40
16	16	0.43	38	34	0.40
17	18	0.46	39	36	0.41
18	18	0.45	40	36	0.38
19	20	0.42	41	38	0.40
20	20	0.42	42	38	0.40
21	20	0.43	43	40	0.39
22	20	0.41	44	42	0.39

解：运行 BIPTCHEM 软件包，由键盘输入各对 (x, y) 值后，根据屏幕提示输入待定

系数个数“2”，第1待定系数“a”，第2待定系数“b”。屏幕显示：

输入预选方程：

y=

在等号“=”后键入“1/(1+exp(a+b*x))”，按<Enter>键，按屏幕提示确认(预选方程)是否正确；屏幕再次显示：

输入预选方程：

y=1/(1+exp(a+b*x))

是否正确？(Y/N)

经确认后，输入第1待定系数初值(0.05)，输入第2待定系数初值(0.01)。回归计算结果如下：

方程 y=1/(1+exp(a+b*x))

a=5.44913E-2

b=1.119E-2

均方差 S=0.0152156

待定系数的初值可根据经验估计，若离真值太远，可能不收敛。

必须指出，以上回归问题都是在假设 x 没有误差，或相对 y 而言误差微小得可忽略不计。否则，其回归处理将更复杂，此处不过多研究，如若需要可参考有关书籍。

习 题

1. 在温度为 40℃ 时，活性炭吸收氯气量 A 与压强 P 的关系如下表，求经验公式。

P/mmHg	0.06	0.15	0.56	1.65	5.14	12.88	20.93	35.91	47.58
A/mg	33.91	50.87	70.31	99.87	122.13	140.87	152.42	166.44	173.66

2. 用比色法测定 SiO_2 含量所得的数据如下：

x/mg	0	0.02	0.04	0.06	0.08	0.10	0.12	未知液
吸光度 y	0.032	0.135	0.187	0.268	0.359	0.435	0.511	0.242

试求回归方程及未知液的含量。

3. 由实验得到下表中的数据，其中 C_A 是在时间 t 后，一个反应体系中未反应的反应物质的相对量（原料为 100），求 C_A 与 t 的关系。

t	2	5	8	11	14	17	27	31	35	44
C_A	94.8	87.9	81.3	74.9	68.7	64.0	49.3	44.0	39.1	31.6

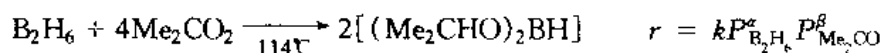
4. 间二甲苯与 Br_2 在 17℃ 进行反应，预想其化学反应速度方程式为：

$$r = kC_{\text{Br}}^m$$

试利用下列实验数据确定 k 和 m 。

$C_{Br}/(\text{mol/l})$	0.3150	0.2812	0.2555	0.2353	0.2153	0.198	0.1852	0.1713	0.1566
$r \times 10^3/(\text{mol/l} \cdot \text{min})$	16.44	13.56	11.48	11.68	9.11	8.00	7.73	7.71	5.87
$C_{Br}/(\text{mol/l})$	0.1465	0.1295	0.1107	0.0942	0.0799	0.0736	0.0692	0.0615	0.0518
$r \times 10^3/[\text{mol}/(\text{l} \cdot \text{min})]$	4.06	3.64	3.23	2.79	2.10	1.55	1.35	1.25	1.18

5. 下列反应的预选反应速度方程和实验数据如下, 试确定参数 k , α , β 。



(注: Me 代表 $-\text{CH}_3$)

$P_{\text{B}_2\text{H}_6}/\text{mmHg}$	6.0	8.0	10.0	16.0	10.0	10.0	10.0	10.0	10.0
$P_{\text{Me}_2\text{CO}}/\text{mmHg}$	20.0	20.0	20.0	20.0	10.0	20.0	40.0	60.0	100.0
$r \times 10^3/(\text{mol/l} \cdot \text{s})$	0.5	0.63	1.00	1.28	0.33	0.80	1.50	2.21	3.33

6. 由下列数据, 回归得出 Antoine 蒸汽压方程式中的常数 A 、 B 、 C

$$\lg P_{VP} = A - \frac{B}{T + C} \quad (T-\text{K})$$

(1) 苯的饱和蒸汽压 P_{VP} 与温度的关系

$t/^\circ\text{C}$	10	20	30	40	50	60	70	80	90	100
P_{VP}/mmHg	45.529	75.198	119.33	182.77	271.27	391.45	550.80	757.62	1020.9	1350.4

(初值: 设 $A = 15$, $B = 2800$, $C = -50$)

(2) 醋酸的饱和蒸汽压 P_{VP} 与温度的关系

$t/^\circ\text{C}$	50	60	70	80	90	100	110	120	130
P_{VP}/mmHg	57.053	90.476	138.94	207.28	301.29	427.71	594.28	809.77	1083.9

(初值: 设 $A = 16$, $B = 3400$, $C = -50$)

7. 利用下式

$$C_P = A + BT + CT^2 + DT^3$$

来表示 N_2 、 C_2H_4 的恒感热容 $[\text{cal}/\text{mol} \cdot ^\circ\text{C}]$ 与温度 T 的关系。用下列数据回归出 A 、 B 、 C 、 D 之值。

T/K	100	150	200	250	300	350	400	450	500	550	600	650	700
C_{P, N_2}	6.664	6.666	6.684	6.714	6.756	6.807	6.866	6.930	6.998	7.068	7.137	7.205	7.268

T/K	150	200	250	300	350	400	450	500	550	600	650	700
$C_{P, \text{C}_2\text{H}_4}$	6.08	7.62	9.08	10.45	11.74	12.95	14.08	15.15	16.14	17.08	17.95	18.76

8. 包面包的塑料薄膜强度试验中, 得知强度与两处操作温度及薄膜中聚乙烯的质量分数有关, 试验数据如下:

$y/(\text{g/m})$	$x_1/^{\circ}\text{F}$	$x_2/^{\circ}\text{F}$	$x_3/\%$
6.6	225	46	0.5
6.9	285	46	0.5
7.9	225	64	0.5
6.1	285	64	0.5
9.2	225	46	1.7
6.8	285	46	1.7
10.4	225	64	1.7
7.3	285	64	1.7
9.8	225	55	1.1
5.0	285	55	1.1
6.9	204.5	39.9	1.1
6.3	305.5	70.1	1.1
4.0	255	55	0.09
8.6	255	55	2.11
10.1	255	55	1.1
9.9	255	55	1.1
12.2	255	55	1.1
9.7	255	55	1.1
9.7	255	55	1.1
9.6	255	55	1.1

预选方程为

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

求其回归方程。

参 考 文 献

- 1 滕敏康主编. 实验误差与数据处理. 南京大学出版社, 1989
- 2 肖明耀著. 误差理论与应用. 计量出版社, 1985
- 3 梁晋文, 陈林才, 何贡编. 误差理论与数据处理. 中国计量出版社, 1989
- 4 王梓坤. 概率论基础及其应用. 科学出版社, 1989

第三篇 化工过程优化技术及应用

第8章 最优化方法

8.1 概述

在工程设计、生产技术、科学试验和计划管理等许多方面，人们总想采取种种措施以最小的代价获得最大的效果。这一类问题，长期以来经过数学归纳与提炼，就是范围很广的最优化问题。

简言之，最优化（Optimization）就是寻求最好的方式来解决实际问题。它不仅要找出问题的最优解，而且要提高求解的效率。这显然具有重要的现实意义，特别是对于高度复杂性和紧迫性的问题，效率的差别将对整个事件的成功与否发生影响。

对于化工过程中的最优化而言，最优化是以降低生产成本或增加利润为主要目标的，同时还要考虑保护自然环境、过程的安全性和可靠性等因素。例如，当我们在设计一个设备或一个工厂时，总是希望得到的产品成本最低或获利最大，这就是最优设计问题。对于现存设备或工厂，我们总是设法对其工况加以调节和控制，使产量最高或获利最大，这就是最优控制问题。还存在范围更加广泛的企业管理问题，如生产计划（包括产品的种类、产量的决定，原料供应，产品销售，设备维修，催化剂更新等）的制定；工厂各部门（如生产部门，运输部门，水、电、汽等公用工程部门，原料和产品仓库等）的协调配合。所有这些问题都必须用科学的方法加以解决。要使一个工厂或联合企业有效地、高水平地运转，上述三方面的优化是缺一不可的。

最优化方法，从学科上讲，它既是应用数学的一个分支，又是系统工程学的重要内容。最优化效果的大小和处理的对象系统有关，一般说来，对象系统范围愈大，优化后的效果也愈大。例如化工企业中，某个单元过程的优化效果比不上全过程的优化效果。然而，后者的优化比前者的优化更为困难和复杂。

8.1.1 优化问题的性质和数学描述

在工程问题中，常会遇到设备费和操作费之间的矛盾，如何在设备费和操作费之间进行平衡，使总费用最小，这就是优化要解决的问题。优化的目标是确定系统中各单元的结构参数和操作参数，使系统的经济指标达到最优。

例如，化工中确定设备和管道保温层的最经济厚度问题。

设备和管道的保温能减少热量损失，但购置保温材料需要消耗一定的费用。保温层的厚

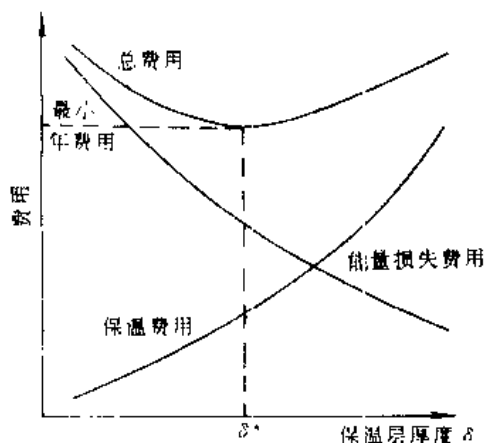


图 8.1 保温层厚度与总费用的关系

度应为多少才是最合适的，这就是一个优化问题。若增加保温层厚度，则保温费用增加，但能耗费用可进一步下降。保温费用、能耗费用与总费用和保温层厚度 δ 的关系曲线如图 8.1 所示。由图 8.1 可知，随着保温层厚度的增加，总费用先下降，然后从某一点开始上升。因为此时增加保温层厚度所节省的能耗费用已经很小。因此存在一个最优保温层厚度 δ^* ，此时总费用最小。该最优点可通过优化得到。

再例如，最优生产安排问题。

某大公司拥有 m 个生产厂（均生产同一产品），如何安排生产计划也是一个优化问题。若产品提供给 n 个用户，安排生产时就要确定每个厂的生产任务 (Y_1, Y_2, \dots, Y_m) 以及每个厂的产品如何分配给用户，即 Y_j 的分配 $(Y_{j1}, Y_{j2}, \dots, Y_{jn})$ 问题。最小费用不仅同产品由生产厂送货给用户的运费有关，也同每个厂的生产费用与产量的关系有关。各厂的设计能力各不相同，生产成本也因工厂水平不同而有差异。由于运费的不同，如果用户离老厂很近的话，可能会让生产成本较高的老厂（与现代化的新厂相比）生产更多的产品。另一方面，如果老厂的产量大大超过设计能力而使生产费用大幅度提高，这时安排离用户较远的新厂进行生产可能会更经济。这就是生产的最优安排问题。

化工过程的设计和操作问题一般具有多解，甚至无穷解。优化就是用定量的方法从众多的解中找出最优的一组解。为此，需要建立过程的数学模型，确定合适的优化目标。

各种优化问题在数学上具有相同的结构，抽象成数学问题后就有其共性。因此，优化问题的数学描述是解决优化问题的关键步骤。优化问题的数学描述包括：（1）目标函数；（2）约束条件。

1. 目标函数

凡是最优化问题，都有要达到的“最优”目标，把目标写成为数学形式的表达式称为目标函数。由于问题和场合的不同，目标函数有时称为性能指数或评价函数。

一般情况下，优化目标常选用经济指标，故目标函数也称为经济模型，它描述了经济指标与过程变量之间的函数关系。常用的经济指标有利润、费用、能耗、单耗、产率等。

目标函数的一般表达式为

$$J = f(\vec{x}), \text{求其最小(大)值}$$

其中，过程变量 $\vec{x} = (x_1, x_2, \dots, x_n)^T$ ，上标 T 表示向量的转置。

2. 约束条件

当 \vec{x} 的各分量 x_i ($i=1, 2, \dots, n$) 为一组特定的数值时，称为一个决策或一个解。因场合的不同有时也称为一个设计或一个控制。实际上有些决策在技术上、工程上或实践中是不现实的或明显地不合理，甚至是违反某些原则的。因此，过程变量 \vec{x} 的取值范围通常都有一个限制，这种对变量取值的限制称为约束条件。当这种约束条件以不等式表达时称为不等式约束。化工过程中通常由热力学、动力学所得到的系统模型方程（物料衡算式、热量衡算式、动量衡算式等）属于等式约束，而一般的安全条件，例如过程操作时间的压力上限、温度上限等则属于不等式约束。

等式约束把过程或系统状态刻画得十分清楚而准确，因此，等式约束式又被称为过程或

系统的状态方程。

3. 决策变量和状态变量、系统自由度

在最优化问题的 n 个过程变量中, 可根据在最优化决策中的特性和所引起的作用划分为决策变量和状态变量。状态变量是最能够描述过程或系统的特征或行为的 (即状态的) 最少一组变量, 其值是不能任意选取的。决策变量根据最优化任务的不同又称为设计变量, 控制变量、操作变量, 它是由决策者 (设计者、控制者、操作者等) 根据目标和约束条件的要求而确定的。所谓决策, 就是选择一组决策变量以满足任务的要求。决策变量一经确定, 状态变量也就随之确定, 从而过程或系统的状态也就被确定。在面临某个具体问题时, 究竟哪些变量选为决策变量应根据具体情况而定, 但一般应选择那些能观察, 能检验, 能控制的变量作为决策变量, 例如化工过程中压力、温度、流量等易于观察和控制, 宜于作决策变量。工程上决策变量和状态变量有较明显的特征上的差别, 但在数学处理上却并不严格区别, 它是以数学上如何处理更方便作为一般准则。

在确定状态变量和决策变量时必须遵循以下原则:

$$\text{状态变量数目} = \text{状态方程数目}$$

且有

$$\text{过程变量的总数} - \text{状态方程数} = \text{决策变量数}$$

过程或系统中最优化问题上的决策变量数又称为系统的自由度。系统自由度标志着系统在最优化设计中可供选择或设计的难易程度。对于系统的设计来说显然有:

- (1) 当自由度等于零时, 系统没有最优化问题, 即没有决策变量, 解是唯一的;
- (2) 当自由度较大时, 系统的最优设计较困难;
- (3) 当自由度较小时, 系统最优设计较容易, 显然, 系统的自由度必须大于零, 当自由度为 1 时, 实际为单变量的最优化问题。

例 8.1 生产计划最优安排数学模型的建立。

某化工厂生产甲、乙两种产品, 要用 A、B、C 三种不同原料, 从工艺资料知道该问题的有关条件如表 8.1 表示, 问如何安排生产, 使一天的总利润为最大?

表 8.1 原料-产品-利润表

原 料 \ 单位产品所需原料	产 品		原料供应量 (单位)
	甲	乙	
A	1	1	6
B	1	2	8
C	0	1	3
单位利润 (¥)	3	4	

解: 如何用数学语言描述上述问题, 即如何建立该问题的数学模型? 通常有以下三步:

(1) 确定决策变量。将生产决策中的关键的量设为决策变量, 即可设产品甲、乙的日产量为决策变量, 分别用 x_1 和 x_2 表示。

(2) 明确实际问题所追求的目标是什么, 并根据它同决策变量 \vec{x} 之间的关系, 把它写为决策变量的函数。本例所追求的目标是“工厂一天的总利润最大”, 设其为 J , 则目标函数

$$J = f(\vec{x}) = 3x_1 + 4x_2 \quad (1)$$

(3) 找出问题中所有对决策变量的限制条件, 并用方程或不等式表示。本例中, 目标

函数 J 虽然是随产品的产量增加而增加, 但两种产品都要受到三种原料供应的限制, 即有以下三个不等式约束

$$x_1 + x_2 \leq 6 \quad (2)$$

$$x_1 + 2x_2 \leq 8 \quad (3)$$

$$x_2 \leq 3 \quad (4)$$

是否还有约束? 显然, x 必须大于零, 因为 x 的负值没有意义的。因此有

$$x_1 \geq 0 \quad (5)$$

$$x_2 \geq 0 \quad (6)$$

这样, 该优化问题可描述为: 找出 x 的最优值, 其应满足 (2) ~ (6) 式表示的约束条件, 且使 $f(x)$ 的值最大。

例 8.2 冷凝器最优设计。

现需设计苯-甲苯精馏塔的塔冷凝器 (图 8.2), 要求冷凝器的总费用最小。根据该精馏塔和公用工程的工艺要求, 塔顶蒸汽流量 W_1 及温度 T_1 、冷却水进口温度 t_1 和传热系数 K 为已知, 试建立目标函数和约束方程。

解: 这个问题的经济指标为总费用。冷凝器的总费用 C_T 包括折旧费 C_e/N (C_e 为设备费, N 为折旧年限) 和操作费 C_{opr} 。 C_{opr} 包括冷却水费和维修费。目标函数可表示为

$$J = C_T = C_e/N + C_{opr} \quad (1)$$

该过程的变量共有 10 个。

状态变量: A 、 G_1 、 G_2 、 W_2 、 T_2 , 5 个;

决策变量 (设计变量): W_1 、 T_1 、 t_1 、 t_2 、 K , 5 个。而 W_1 、 T_1 、 t_1 、 K 为设计给定变量。

其中 A 为冷凝器的换热面积, G_1 和 G_2 为冷却水的进出口流量, T_2 为冷凝液温度, t_2 为冷却水出口温度。目标函数可表达为上述变量的函数, 因为

$$C_e = f(A, T_1, t_1, t_2, W_1, G_1) \quad (2)$$

$$C_{opr} = f(W_2, t_1, t_2, C_e) \quad (3)$$

该过程的数学模型为

$$\begin{cases} T_2 = T_b \\ W_2 = W_1 \\ G_2 = G_1 \\ W_1 \gamma = G_1 c_p (t_2 - t_1) \\ W_1 \gamma = KA \Delta t_m \end{cases} \quad (4)$$

式中 T_b 为泡点温度, γ 为汽化热, c_p 为冷却水比热容, Δt_m 为平均温差。模型的自由度 $f=5$, 即该问题中决策变量数为 5。对于设计型问题, 其中的 4 个变量可以预先指定, 因此参与决策的变量只有 1 个, 选冷却水出口温度 t_2 为决策变量。

该优化问题可描述为: 在满足约束方程, 即式 (4) 的条件下, 使总费用为最小时冷却水的最优出口温度 t_2^* 。该问题的图解见图 8.3。在解得 t_2^* 后, 可由模型方程确定冷凝器的最优传热面积 A^* 。

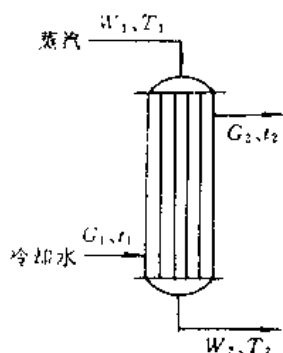


图 8.2 冷凝器

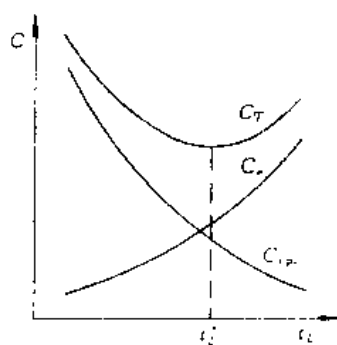


图 8.3 冷凝器优化图解

综上所述，优化问题数学表达的一般形式为

$$\text{目标函数: } J = f(\vec{x}) \quad (8.1a)$$

约束条件:

$$\text{等式约束: } H(\vec{x}) = \vec{0} \quad (8.1b)$$

$$\text{不等式约束: } \vec{g}(\vec{x}) \geq \vec{0} \quad (8.1c)$$

式中 $\vec{x} = (x_1, x_2, \dots, x_n)^T$

$$H(\vec{x}) = (h_1(\vec{x}), h_2(\vec{x}), \dots, h_l(\vec{x}))^T$$

$$\vec{g}(\vec{x}) = (g_1(\vec{x}), g_2(\vec{x}), \dots, g_m(\vec{x}))^T$$

在式 (8.1) 下求得的 $\vec{x}^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ ，称为该最优化问题的最优解，在 \vec{x}^* 对应下的目标函数 $J^* = f(\vec{x}^*)$ 称为目标函数的最优值。最优值可以是最小值或是最大值

优化问题可表达为：求出满足约束条件 (8.1b、8.1c 式)，使目标函数 J 达到最小（或最大）时变量的值 \vec{x}^* 。

在确定型系统的静态最优化中，一般不严格区分状态变量和决策变量，用 \vec{x} 表示系统最优化的所有变量。

8.1.2 最优化方法的分类

最优化问题的求解方法称为最优化方法。最优化问题的性质不同，求解的方法也将不同。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。近 20 几年来，由于实际需要和电子计算机的应用，促进了最优化理论和方法的发展。

根据优化问题有无约束条件，可分为无约束优化和有约束优化问题。根据优化问题的目标函数及约束条件的数学特点，即函数形式，可分为线性规划问题和非线性规划问题。当目标函数及约束条件均为线性时，称为线性规划问题；当目标函数或约束条件中至少有一个为非线性时，称为非线性规划问题。

求解线性规划问题的优化方法已相当成熟，通常采用单纯形法。

求解非线性规划问题的方法是很多的，近年来，这个领域的研究取得了许多进展。归纳起来，非线性规划问题的优化方法大致可分为两大类。

(1) 间接优化方法 间接优化方法就是解析法，即按照目标函数极值点的必要条件，用

数学分析的方法求解。再按照充分条件或者问题的实际物理意义，间接地确定最优解，因此，称为间接法，它适用于求解目标函数具有简单而明确的数学形式的非线性规划问题。而对于目标函数比较复杂，或无明确的数学表达式的情况，难以解析处理时，间接法失效，就得借助于下述的直接寻优法。

(2) 直接寻优方法 又称搜索法。它是一种数值方法，利用函数在某一局部区域的性质或一些已知点的数值，来确定下一步计算的点，这样一步步搜索、逼近，最后达到最优点。间接法是化工问题的主要求解方法。

8.1.3 优化问题求解的一般步骤

目前尚没有一种优化方法能有效地适应所有的优化问题。针对某一特定问题选择优化方法的主要根据为：

- (1) 目标函数的特性；
- (2) 约束条件的性质；
- (3) 决策变量和状态变量的数目。

优化问题求解的一般步骤为：

- (1) 对过程进行分析，列出全部变量；
- (2) 确定优化指标，建立指标同过程变量之间的关系，即目标函数关系（经济模型）；
- (3) 建立过程的数学模型和外部约束（包括等式约束及不等式约束），确定自由度和决策变量。一个过程的模型可以有多种，应根据需要，选择简繁程度合适的模型；
- (4) 如果优化问题过于复杂，则将系统分成若干子系统分别优化；或者对目标函数和模型进行简化；
- (5) 选用合适的优化方法进行求解；
- (6) 对得到的解检验，考察解对参数和简化假定的灵敏度。

前三部为优化问题进行数学描述。第四步建议对过程作尽可能的简化，而又不歪曲问题的本质。首先，可以忽略那些对目标函数影响不大的变量。这可以根据工程判断来确定，也可以通过灵敏度分析作出判断。

目前，尽管化工中的优化问题求解成功的实例屡见不鲜，但是仍有许多化工中的优化问题求解是十分困难的，有些还不能找到一种适宜的优化方法。

当目标函数和约束条件均为线性时，优化问题可用线性规划有效地求解。但对于化工中大多数优化问题来说，都是非线性的。因此，在求解优化问题时，常将过程模型简化成线性模型。化工中优化问题求解的主要困难在于：

- (1) 目标函数和（或）约束函数不连续。例如压缩机的价格不像压力、温度那样连续地变化。这会造成参数在某一个范围内发生变化时对费用没有影响，而在另一范围内发生微小变化就会使费用产生突变。
- (2) 目标函数和（或）约束条件为非线性。这时虽然可以用线性来近似，但对于线性近似引起的后果必须十分小心。
- (3) 目标函数和（或）约束条件中的变量具有很强的交联性。例如，压力容器设计中的压力和温度。若目标函数为 $f = 15.5x_1 \sqrt{x_2}$ ， x_1 同 x_2 的交联造成了多解。
- (4) 目标函数和（或）约束条件在某些范围内接近平坦，而在另一些范围内呈指数的剧变关系。这意味着在某些区间对变量不灵敏，而在另一些区间则极为灵敏。

(5) 目标函数存在许多局部最优点。这时用优化方法找到的往往不是最优点，而是次优点。要找出更优的解必须从另一起始点开始搜索。

8.2 线性规划

线性规划 (Linear Programming, 缩写 LP) 是应用十分广泛, 较为成熟的最优化方法。线性规划所研究的问题主要有两个方面: 一是如何统筹安排一项任务, 以尽量用最少的资源来完成它; 二是如何利用一定量的人力、物力和资金等来完成最多的任务。化工领域的有关规划、排产、资源最佳利用等也都涉及到线性规划问题。本节我们将简明介绍有关的基础知识, 重点放在利用计算机软件求解方面。

8.2.1 线性规划问题的形成

线性规划渊源于生产实践, 因此在生产实践中有很多问题是它的研究对象。本章例 8.3 就是一个典型的线性规划问题, 其目标函数和约束均为线性函数。下面再举一例以进一步说明什么是线性规划。

例 8.3 某制药厂生产 A、B、C 三种药物, 可供选择的原料有甲、乙、丙、丁, 四种原料的每千克成本分别为 5 元, 6 元, 7 元, 8 元。每千克不同原料所能提取各药物的数量 (g/kg) 如下所示, 药厂要求每天生产 A 种药物 100g, B 种药至少 350g, C 种药最多 160g, 按上述要求下使总成本为最小的最佳原料选配。

药 物 名		A	B	C
各原料所含 药物量 (g/kg)	甲	1	5	2
	乙	1	4	1
	丙	1	5	1
	丁	1	6	2

解: 设四种原料的每天选配数量 (kg) 分别为 x_1 、 x_2 、 x_3 、 x_4 , 则依题意可得下列数学模型, 即在

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 100 \\ 5x_1 + 4x_2 + 5x_3 + 6x_4 &\geq 350 \\ 2x_1 + x_2 + x_3 + 2x_4 &\leq 160 \\ x_j &\geq 0, \quad j = 1, 2, 3, 4 \end{aligned}$$

约束下求 $J = 5x_1 + 6x_2 + 7x_3 + 8x_4$ 的最小值。

可以看出, 目标函数和约束均为线性的, 属线性规划问题。

在讨论线性规划的优化方法之前, 先介绍一下化工中线性约束的一般构成, 这对我们建立实际问题的线性规划模型十分重要。

对于化工生产过程来说, 可能存在以下各种不等式约束。

- (1) 由于设备能力、仓库贮存容量或市场销路限制产生的对生产量的约束。
- (2) 由于供应来源限制对原料的约束。若原料由本公司其他厂家提供, 则原料要受到这些厂生产能力的限制。
- (3) 由于操作压力、温度和流量等限制对操作条件的约束。

线性等式约束则由化工过程模型方程的线性化得到。这一点很重要，拓宽了线性规划的应用范围。

8.2.2 线性规划问题的数学模型

由例 8.1 和 8.3 中的数学模型可以看出，线性规划数学模型有如下三个特征：

- (1) 含有不等式或等式在内的线性约束方程组。
- (2) 约束方程组中的所有变量均是非负的。
- (3) 求一线性目标函数为最小（或最大）。

根据以上三个特征，我们给出线性规划数学模型的一般形式。

1. 线性规划数学模型的一般形式

$$\min_{s.t.} J = \sum_{j=1}^n c_j x_j \quad (8.2a)$$

$$\sum_{j=1}^n a_{ij} x_j \leq (\text{或} = \text{或} \geq) b_i, \quad i = 1, 2, \dots, m \quad (8.2b)$$

$$x_j \geq 0, j = 1, 2, \dots, n \quad (8.2c)$$

式中，“s. t.”为英文 subject to（约束于）的缩写。我们称这样的线性规划模型具有 n 个变量和 m 个约束条件。

显然，式（8.2）可表达为向量矩阵形式，则可写成

$$\min_{s.t.} J = \vec{c}^T \vec{x} \quad (8.3a)$$

$$A\vec{x} = (\text{或} \leq \text{或} \geq) \vec{b} \quad (8.3b)$$

$$\vec{x} \geq 0 \quad (8.3c)$$

其中

A ——线性规划约束方程组的系数矩阵，即

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix};$$

\vec{x} ——线性规划的决策变量， $\vec{x} = (x_1, x_2, \dots, x_n)^T$ ；

\vec{b} ——线性方程组的右端向量， $\vec{b} = (b_1, b_2, \dots, b_m)^T$ ；

\vec{c} ——目标函数的系数向量， $\vec{c} = (c_1, c_2, \dots, c_n)^T$ 。

若实际问题中是求目标函数最大值，如例 8.1，则只要改变目标函数的符号，

$$\max f(\vec{x}) = -\min [-f(\vec{x})] \quad (8.4)$$

最小化方法可用于解决求最大点的问题。

如果约束方程（8.2b）全部为“=”，即：

$$\min_{s.t.} J = \sum_{j=1}^n c_j x_j \quad (8.5a)$$

(2) 线性方程组 (8.5b) 满足有解的充分必要条件, 即系数矩阵 A 的秩与其增广矩阵 $[A|b]$ 的秩相等, 即

$$K(A) = K(A|b)$$

(3) 式 (8.5b) 所描述的系统, 其系统自由度大于零, 即在满足秩条件下, 变量数 n 大于方程数 m 。

此外, 在求解过程中出现变量 $x_j < 0$ ($j=1, 2, \dots, n$) 的情况在数学上是允许的, 它不影响线性规划的最后结果。如不另外说明, 本章所讨论的线性规划问题, 均满足上述求解条件。

8.2.3 线性规划问题的解和图解法

本节以式 (8.5) 标准化模型为基础, 通过对例 8.1 的标准化模型 (8.8) 式运用线性代数的消去法求解, 阐明线性规划解的有关概念, 介绍线性规划问题的图解方法。

1. 线性规划的解和有关概念

由线性代数知, 使 8.1 的标准化模型 (8.8) 满足有解的条件, 且系统自由度 $f = n - m = 5 - 3 = 2 > 0$, 即存在两个自由变量, 现以 x_1, x_2 作为自由变量。将其移到约束方程式右端则可以得一组解, 即

$$\left. \begin{aligned} x_1 &= x_1 \\ x_2 &= x_2 \\ x_3 &= 6 - x_1 - x_2 \\ x_4 &= 8 - x_1 - 2x_2 \\ x_5 &= 3 - x_2 \end{aligned} \right\} \quad (8.9)$$

x_1, x_2 的取值不同, 则得线性规划的不同解。因而, 上式是以 x_1, x_2 为自由变量的一般解形式, 现取 $x_1 = x_2 = 0$, 则可得一个特解为:

$$\vec{w}_1 = (x_1, x_2, x_3, x_4, x_5)^T = (0, 0, 6, 8, 3)^T$$

若以 x_1, x_3 为自由变量, 采用同样的方法, 并取 $x_1 = x_3 = 0$, 则又可以得到一个特解为:

$$\vec{w}_2 = (0, 6, 0, -4, -3)^T$$

如此, 继续选择不同的两个变量为自由变量, 令其为零, 可得其他的不同的解, $\vec{w}_3, \vec{w}_4, \dots$ 。

由式 (8.9) 看出, 当两个自由变量一旦确定, 则其余三个变量均可由自由变量线性表出。由此, 可以引出线性规划的几个术语和概念。

(1) 基础变量、非基础变量及其数目

将标准化模型中所有变量划分为基础变量和非基础变量两类。非基础变量即是位于右端的变量, 简称非基变量, 其数目等于系统的自由度 ($n - m$)。基础变量则是由非基变量线性表达的那些变量, 或位于左端的变量, 简称基变量, 其数目等于方程数 m 。显然, 这两种变量只有数目上的限制, 对任意一个变量来说既可作基变量也可作非基变量, 但在一组解中是确定的。两种变量个数的和始终等于变量的总数。本例的变量总数为 5, 其中基变量数为 3, 非基变量数为 2。

(2) 基础解及个数的确定

线性规划中一组基变量在非基变量为零时所得到的一个解称为线性规划的一个基础解(或称基本解)。基变量不同所得的基础解也不同,例如 $\vec{w}_1, \vec{w}_2, \vec{w}_3$ 等均为基础解。每一个基础解中后 $(n-m)$ 个非基变量为零元素,前 m 个基变量一般应有非零元素,如果出现有零元素则称为退化的基础解。此外,基础解中基变量的值并不保证非负,它只保证基础解满足约束方程组 (8.5b)。

基础解的数目等于从 n 中抽出 m 的组合数,即 $C_n^m = \frac{n!}{m!(n-m)!}$, 所有的基础解构成线性规划的基础解系。

例如本例的基础解数目为 $C_5^3 = 10$ 。所有的基础解 $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_{10}$ 构成该线性规划问题的基础解系。

(3) 可行解

线性规划中既满足约束方程组 (8.5b) 又满足非负约束 (8.5c) 的解称为线性规划的可行解。注意,可行解并不保证非基变量都为零。

(4) 基础可行解

当线性规划的解即是基础解又是可行解时,称为线性规划的基础可行解。显然,它是基础解和可行解的交集所组成。基础可行解的非零元素的个数不大于 m , 且都是非负的。

(5) 最优解

线性规划基础可行解中满足目标函数为最优的解称为线性规划的最优解。

综上所述,线性规划各种解之间存在如下集合关系:

$$\text{最优解} \subset \text{基础可行解} \subset \text{基础解}$$

因而形成一种求解思路,即最优解的寻求可以通过先找出所有的基础解,再找出其中的基础可行解,再找目标函数值为最小(或最大)的解,则得最优解。按照这一思路,继续完成例 8.1 的求解。令非基变量为零,分别算出 $10-1=9$ (x_2, x_5 不能同时为零) 个基础解,再用 $\vec{x} \geq 0$ 选出可行解,即可得到其中目标函数值为 20 的最优解为

$$\vec{x}^* = (4, 2, 0, 0, 1)^T$$

其计算过程列入表 8.2 中。

表 8.2 例 8.1 消去法计算表

序号	基本量	非基变量	基 础 解					是否可行解	目标函数值	是否最优解	可行解号
			x_1	x_2	x_3	x_4	x_5				
1	x_3, x_4, x_5	$x_1=0, x_2=0$	0	0	6	8	3	是	0	/	A
2	x_2, x_4, x_5	$x_1=0, x_3=0$	0	6	0	-4	-3	/	24	/	G
3	x_2, x_3, x_5	$x_1=0, x_4=0$	0	4	2	0	-1	/	16	/	F
4	x_2, x_3, x_4	$x_1=0, x_5=0$	0	3	3	2	0	是	12	/	E
5	x_1, x_4, x_5	$x_2=0, x_3=0$	6	0	0	2	3	是	18	/	B
6	x_1, x_3, x_5	$x_2=0, x_4=0$	8	0	-2	0	3	/	24	/	I
7	x_1, x_2, x_5	$x_3=0, x_4=0$	4	2	0	0	1	是	20	是	C
8	x_1, x_2, x_4	$x_3=0, x_5=0$	3	3	0	-1	0	/	21	/	H
9	x_1, x_2, x_3	$x_4=0, x_5=0$	2	3	1	0	0	是	18	/	D

可见,按这种思路求解 LP 问题,当 n 和 m 较大时,计算工作量是很大的,而且非常繁琐,不可行。

2. 线性规划的图解法

当线性规划的初始模型是一个二维问题时,即决策变量为 2 个时,还可以采用图解法,

直接绘上图形进行求解，现仍以例 8.1 为例，在以 x_1, x_2 为横轴与纵轴的坐标系中绘出初始模型式 (8.7) 中三个约束方程所对应的①、②、③三条线，再由变量值应是非负的要求，可得本例的可行域 R 为阴影部分的凸五边形，其可行解和最优解必在此凸五边形（包括边界）中，如图 8.4 所示。

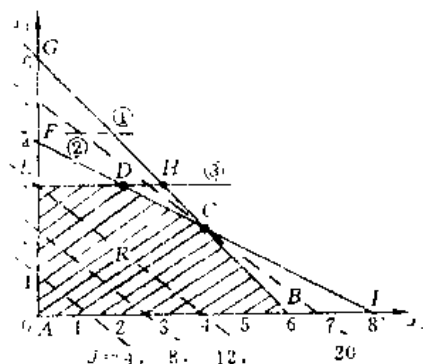


图 8.4 LP 的图解法

由图 8.4，可行域 R 为一五边形，其五个顶点（端点） A, B, C, D, E 与表 8.2 中所列的五个基础可行解相对应，它们对应的 x_1, x_2 坐标也正是这基础可行解中 x_1, x_2 的值。这五个顶点的（目标）函数值以 A 点 $(0, 0)$ 为最小， C 点 $(4, 2)$ 为最大。因此 C 点即是本例所要求的最大点，与消去法的结果相同。从 A 点到 C 点，各顶点的函数值由小到大依次排列，可行域 R 中的平行虚线表示了不同的由小到大目标函数等值（高）线。图中， R 以外的四个点 F, G, H, I 对

应的是表 8.2 中另外四个基础解，它们不满足变量非负的要求，位于可行域 R 之外。所以，此凸五边形正是以基础可行解为顶点所组成的一个凸可行集，或说，线性规划基础可行解就是其可行解集的顶点，最优解就是存在于这些顶点中。

3. 解的特征和几种特殊情况

由上述图解结果和出现的一些现象，并非本例特有，它反映的是线性规划问题解的一般特征，并由以下基本定理所证明。

定理 1 LP 可行解集 R 中，点 \bar{x} 为顶点的充分必要条件是为基础可行解。

定理 2 LP 的最优解，存在于可行解集 R 的顶点，或者说 LP 的最优值一定在可行解集 R 的顶点上找到。

在一般情况下，线性规划有单一的最优解，但有时会遇到一些特殊情况，简介如下。

(1) 有多重最优解情况

设线性规划的初始模型为：

$$\max J = 6x_1 + 10x_2$$

$$\begin{cases} 2x_1 + x_2 \leq 60 \\ 3x_1 + 5x_2 \leq 150 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

此题求得目标函数最大值为 300，最优解的顶点有两个，它们是：

$$\bar{x}^{*(1)} = (0, 30)^T, \bar{x}^{*(2)} = \left(21 \frac{3}{7}, 17 \frac{1}{7}\right)^T$$

而且在 $\bar{x}^{*(1)}$ 和 $\bar{x}^{*(2)}$ 的连续上的所有各点，均满足上述约束条件，目标函数值同为 300，因此，最优解有无穷多个，即有多重最优解。

(2) 可行域无界

设线性规划的初始模型为：

$$\begin{aligned} \max_{x_1, x_2} J &= 2x_1 - x_2 \\ \begin{cases} x_1 + x_2 \geq 1 \\ x_1 + 3x_2 \geq -3 \end{cases} \\ x_1 &\geq 0, x_2 \geq 0 \end{aligned}$$

由上述模型看出, 由于 x_1 可以无限增大, 目标函数的最大值成为正无穷大, 或说目标函数在可行域中无上界, 也不存在最优解。如果模型改为求目标函数最小值, 约束条件不变, 则可得目标函数的最小值为 -1 , 其最优解为 $\bar{x}^* = (x_1^*, x_2^*)^T = (0, 1)^T$ 。因此, 可行域无界并不一定意味着目标函数无极值, 主要看是求何种极值。

(3) 约束条件不构成可行域

设线性规划的初始模型为:

$$\begin{aligned} \max_{x_1, x_2} J &= x_1 + 3x_2 \\ \begin{cases} x_1 + x_2 \leq 10 \\ 2x_1 + x_2 \geq 30 \end{cases} \\ x_1 &\geq 0, x_2 \geq 0 \end{aligned}$$

由两个约束方程和变量的非负性条件看出, 没有任何一组变量可以同时满足这些约束条件, 即约束条件不构成可行域, 显然, 这时的线性规划是无解的。

上述三种情况由于都属于二维问题, 读者不妨用图解法予以验证。产生 (2), (3) 两种情况的主要原因, 一般是建立数学模型发生错误, 故此, 建模时应予注意。

8.2.4 线性规划问题求解的单纯形法

我们已经知道, 当变量 $n \geq 3$ 时, 图解法已不适用, 线性方程组的消去法也仅对低维数的 LP 问题可以求解, 因而必须寻求更为有效的计算方法。

二次大战期间, 美国乔治·丹茨格 (George Dantzig) 经长期的研究, 于 1947 年提出并创立了求解线性规划的一般方法——单纯形法 (Simplex Method), 它是普遍适用于求解一般线性规划问题的一种系统而有效的方法。

1. 单纯形法求解思路

由上述定理可知线性规划的基础可行解即是可行解集 R 的各顶点, 而最优解也必须在 R 的顶点上得到, 因而求解线性规划的思路是: 先找出线性规划的一个初始基础可行解, 即可行解集 R 的一个初始顶点, 再判别其是否为最优解, 如果是, 则不用继续求解; 如果不是, 则由初始顶点转移到另一个顶点, 再进行判别是否为最优解, 这样不断的进行, 最后必然得到所需的最优解和最优值。简言之, 单纯形法的基本思路是从约束可行域的一个顶点转移到另一个顶点, 使目标函数的值在转移的过程中得到改善。这种迭代方法的构思是使得每次转移时目标函数的值都能得到改善。它是对单纯线性图解 (可行域) 的顶点进行搜索寻优, 因而称为单纯形法。

如图 (8.5) 所示, 单纯形法的三个关键步骤是:

- (1) 初始顶点即初始基础可行解的寻找;
- (2) 最优性的判别;
- (3) 顶点的转移及其最佳途径。

2. 单纯形法计算要点

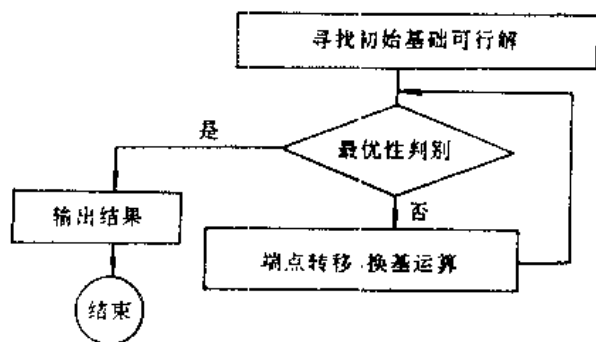


图 8.5 单纯形法示意图

(1) 将线性规划初始模型化为单纯形法要求的求解形式，即除了引入附加变量将不等式约束化为等式约束之外，还要求约束方程的常数项 $b_i \geq 0$ ($i = 1, 2, \dots, m$)。

(2) 取初始顶点（基础可行解）。

最明显的一个特解的求法是：把原约束方程组中系数为 1 的变量作为基变量，并令非基变量为零，便得到一个初始顶点。

系数为 1 的变量获得并不困难，对于非标准型的 LP 问题，当加入 m 个附加变量后，它们的系数均为 1，所以可直接选取附

加变量 x_{n+i} ($i = 1, 2, \dots, m$) 为基变量，决策变量为非基变量。

(3) 进行最优性判别。

经过严格的数学理论推导，单纯形法给出了顶点最优性判别准则^[1]。如果某一顶点满足判别准则，则计算终止，以此顶点为最优；如果某一顶点不满足，则进行下一步，顶点转移。

(4) 顶点转移及其最佳途径。

从线性代数消去法和图解法均可看出，各顶点（基础可行解）之间的差异是其基变量不同。因此，顶点转移是通过基变量的变换来实现的，亦称换基运算。即把原顶点的某一个基变量换出，并用另一个不同的非基变量代替。至于选取哪一个变量作为换出变量，哪一个变量作为调入变量，是顶点转移的途径问题。这也正是单纯形法的关键步骤。

按单纯形法的设计，顶点的转移是按照最佳的途径进行的。所谓转移的最佳途径是指：当上一个顶点不是最优解时，应进行换基，即转移到下一个顶点，下一个顶点的目标函数值必须进一步减小（增大），且减小（增大）的幅度应为最大。即以最少的转移次数到达目标函数为最优的顶点这样一种转移途径。

关于单纯形法的求解原理和计算步骤，文献[1] [2]中均有详述，本书不再重复。本节的重点是线性规划的基本概念和如何用单纯形法计算软件求解线性规划问题。

与本书配套的 BIPTCHEM 软件包可用于求解线性规划问题，软件使用方法见第十章。

例 8.4 解下列线性规划问题。

$$\begin{aligned} \max J &= x_1 - x_2 \\ \text{s.t.} \quad &\begin{cases} 2x_1 - x_2 \geq -2 \\ -x_1 + 3x_2 \geq -2 \\ -x_1 - x_2 \geq -4 \\ x_1 \geq 0, x_2 \geq 0 \end{cases} \end{aligned}$$

解：按使用说明运行 BIPTCHEM 软件包，打开优化菜单，选择线性规划，按〈Enter〉键。根据屏幕提示，输入基本变量数“2”及约束数“3”，屏幕显示：

目标函数形式为 $F = c_1x_1 + c_2x_2$
按以上格式输入目标函数
F =

在“F=”后键入“ $x_1 - x_2$ ”，按〈Enter〉键。屏幕显示：

是否正确？(Y/N)

键入“Y”，屏幕接着显示：

约束方程的形式为：
 $a(j, 1)x_1 + a(j, 2)x_2 \geq (=, \text{或} <=) a(j, 4)$
按以上格式输入约束方程
方程 1:

在“方程 1”后键入“ $2x_1 + x_2 \geq -1$ ”，按〈Enter〉键后，屏幕显示：

是否正确？(Y/N)

键入“Y”，屏幕接着显示：

方程 2:

以相同方式输入约束方程 2 “ $-x_1 + 3x_2 \geq -2$ ”和约束方程 3 “ $-x_1 - x_2 \geq -4$ ”，回答是否打印后，屏幕显示：

1. 目标函数求最大值
2. 目标函数求最小值
请选择!

键入“1”，屏幕显示：

目标函数: $F = x_1 - x_2$
约束条件: $2x_1 - x_2 \geq -2$
 $-x_1 + 3x_2 \geq -2$
 $-x_1 - x_2 \geq -4$
 $x_i \geq 0$

计算结果:

目标函数最大值为 3
 $x_1 = 3.5 \quad x_2 = 0.5$

例 8.5 用 BIPTCHEM 软件求解例 8.1。

解：例 8.1 的初始模型为式 (8.7)，求解步骤如下：

(1) 和例 8.4 相同方式运行 BIPTCHEM，打开优化菜单，选择线性规划，按〈Enter〉键

(2) 按屏幕提示输入基本变量数“2”，约束数“3”。在“F=”后边输入目标函数
“ $3x_1 + 4x_2$ ”

输入约束方程：

$$\begin{aligned}x_1 + x_2 &\leq -6 \\x_1 + 2x_2 &\leq 8 \\x_2 &\leq 3\end{aligned}$$

(3) 回答是否打印, 选择目标函数求最大值后, 屏幕显示:

$$\begin{aligned}\text{目标函数: } F &= 3x_1 + 4x_2 \\ \text{约束条件: } x_1 + x_2 &\leq -6 \\ x_1 + 2x_2 &\leq 8 \\ x_2 &\leq 3 \\ x_i &\geq 0\end{aligned}$$

计算结果:

$$\begin{aligned}\text{目标函数最大值为 } 20 \\ x_1 = 4 \quad x_2 = 2\end{aligned}$$

例 8.6 用 BIPTCHEM 软件包求解例 8.3。

解: 例 8.3 的线性规划初始模型为:

$$\begin{aligned}\max J &= 5x_1 + 6x_2 + 7x_3 + 8x_4 \\ \text{s.t.} \quad &\begin{cases} x_1 + x_2 + x_3 + x_4 = 100 \\ 5x_1 + 4x_2 + 5x_3 + 6x_4 \geq 350 \\ 2x_1 + x_2 + x_3 + 2x_4 \leq 160 \\ x_j \geq 0, j = 1, 2, 3, 4 \end{cases}\end{aligned}$$

求解步骤如下:

- (1) 运行 BIPTCHEM, 打开优化菜单, 选择线性规划, 按〈Enter〉键;
- (2) 按屏幕提示输入基本变量数“4”, 约束数“3”, 在“F=”后边输入目标函数
“5x1 + 6x2 + 7x3 + 8x4”

输入约束方程:

$$\begin{aligned}x_1 + x_2 + x_3 + x_4 &= 100 \\ 5x_1 + 4x_2 + 5x_3 + 6x_4 &\geq 350 \\ 2x_1 + x_2 + x_3 + 2x_4 &\leq 160\end{aligned}$$

(3) 回答是否打印, 选择目标函数求最小值后, 屏幕显示:

$$\begin{aligned}\text{目标函数: } F &= 5x_1 + 6x_2 + 7x_3 + 8x_4 \\ \text{约束条件: } x_1 + x_2 + x_3 + x_4 &= 100 \\ 5x_1 + 4x_2 + 5x_3 + 6x_4 &\geq 350 \\ 2x_1 + x_2 + x_3 + 2x_4 &\leq 160 \\ x_i &\geq 0\end{aligned}$$

计算结果:

$$\begin{aligned} &\text{目标函数最小值为 } 540 \\ &x_1 = 60 \quad x_2 = 40 \quad x_3 = 0 \quad x_4 = 0 \end{aligned}$$

例 8.7 解下列带无约束变量的线性规划问题。

$$\begin{aligned} \max J &= x_1 + 4x_2 \\ \begin{cases} x_1 + x_2 \leq 3 \\ -x_1 + x_2 \leq 1 \end{cases} \\ x_1 &\text{ 无约束, } x_2 \geq 0 \end{aligned} \quad (1)$$

解：在线性规划模型中，变量 x_j ($j = 1, 2, \dots, n$) 恒要求非负。但是实际问题中有些变量从物理意义上说并不要求非负，也就是说某些变量可以任意取值，我们称这些变量为无约束变量。本例模型中，变量 x_1 可取正，也可取负，即为无约束变量。对于此类 LP 问题，为了能施用单纯形法求解，可以采用以下方法进行无约束变量代换。

引入两个变量 $x_3 \geq 0, x_4 \geq 0$ ，令

$$x_1 = x_3 - x_4 \quad (2)$$

这样 x_1 可为正值，也可为负值。将式 (2) 代到目标函数和约束方程中，就将无约束变量 x_1 代换了，得

$$\begin{aligned} \max J &= 4x_2 + x_3 - x_4 \\ \begin{cases} x_2 + x_3 - x_4 \leq 3 \\ x_2 - x_3 + x_4 \leq 1 \end{cases} \\ x_2 &\geq 0, x_3 \geq 0, x_4 \geq 0 \end{aligned} \quad (3)$$

待求出该线性规划的最优点后，再代入式 (2) 即可求出 x_1 。

利用 BIPTCHEM 软件包求解步骤如下：

- (1) 运行 BIPTCHEM，打开优化菜单，选择线性规划，按〈Enter〉键；
- (2) 按屏幕提示输入基本变量数“3”，约束数“2”，在“F=”后边输入目标函数“ $4x_1 + x_2 - x_3$ ”

输入约束方程：

$$\begin{aligned} x_1 + x_2 - x_3 &\leq 3 \\ x_1 - x_2 + x_3 &\leq 1 \end{aligned}$$

- (3) 回答是否打印，选择目标函数求最小值后，屏幕显示：

$$\begin{aligned} &\text{目标函数: } F = 4x_1 + x_2 - x_3 \\ &\text{约束条件: } x_1 + x_2 - x_3 \leq 3 \\ &\quad \quad \quad x_1 - x_2 + x_3 \leq 1 \\ &\quad \quad \quad x_i \geq 0 \end{aligned}$$

计算结果：

$$\begin{aligned} &\text{目标函数最小值为 } -1 \\ &x_1 = 0 \quad x_2 = 0 \quad x_3 = 1 \quad (x_2 = 0 \quad x_3 = 0 \quad x_4 = 1) \end{aligned}$$

将 x_3, x_4 代入式 (1), 得: $x_1 = 0 - 1 = -1$

8.2.5 线性规划问题的应用

炼油厂和化工厂中有相当一部分优化问题是用线性规划求解的。文献中已发表的用线性规划求解的优化问题有:

(1) 多套装置生产计划的安排⁽⁴⁾⁽⁵⁾;

(2) 汽油配制;⁽⁴⁾⁽⁶⁾

(3) 炼油厂的操作⁽³⁾⁽⁷⁾。

下面再举 2 个化工生产中线性规划应用的简单实例。

例 8.8 炼油厂生产计划的安排问题。

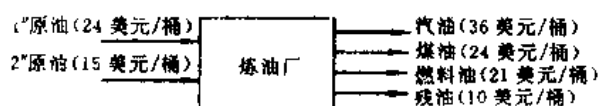


图 8.6 炼油厂输入和输出图

图 8.6 所示为一炼油厂原料输入和产品输出的情况, 括号内的数字为原料进厂和产品出厂的价格。表 8.3 列出了有关的数据。

表 8.3 炼油厂原料和产品数据

产 品 名 称	得 率 (%)		最高产量 (桶/天)
	1# 原油	2# 原油	
汽 油	80	44	24000
煤 油	5	10	2000
燃料油	10	36	6000
残 油	5	10	
加工费 (美元/桶)	0.50	1.00	

问如何安排生产才能使炼油厂利润最大。

解: 设定变量 (单位均为桶/天):

x_1 ——1#原油耗用量;

x_2 ——2#原油耗用量;

x_3 ——汽油产量;

x_4 ——煤油产量;

x_5 ——燃料油产量;

x_6 ——残油产量

若以利润为经济指标, 目标函数为:

$$\max J = f(\vec{x}) = \text{产值} - \text{原料费} - \text{加工费}$$

$$\text{产值} = 36x_3 + 24x_4 + 21x_5 + 10x_6$$

$$\text{原料费} = 24x_1 + 15x_2$$

$$\text{加工费} = 0.5x_1 + x_2$$

约束条件: 根据每个产品的得率 (物料衡算) 可列出 4 个等式约束

$$\text{汽 油} \quad 0.80x_1 + 0.44x_2 = x_3$$

$$\text{煤 油} \quad 0.05x_1 + 0.10x_2 = x_4$$

$$\text{燃料油} \quad 0.10x_1 + 0.36x_2 = x_5$$

$$\text{残 油} \quad 0.05x_1 + 0.05x_2 = x_6$$

为减少变量数, 可将上述约束方程代入目标函数, 消去 x_3, x_4, x_5, x_6 ,

$$\text{产值} = 36(0.80x_1 + 0.44x_2) + 24(0.05x_1 + 0.10x_2)$$

$$+ 21(0.10x_1 + 0.36x_2) + 10(0.05x_1 + 0.05x_2) \\ = 32.6x_1 + 26.8x_2$$

最后得到 $J = f(\vec{x}) = 8.1x_1 + 10.8x_2$

是否存在其他约束？表 8.3 列出了每种产品的最大允许产量，即

$$A: x_3 \leq 24000 \quad 0.80x_1 + 0.44x_2 \leq 24000$$

$$B: x_4 \leq 2000 \quad 0.05x_1 + 0.10x_2 \leq 2000$$

$$C: x_5 \leq 6000 \quad 0.10x_1 + 0.36x_2 \leq 6000$$

另外还有隐含的和非负的约束： $x_1 \geq 0, x_2 \geq 0$

于是，本规划问题的 LP 模型为：

$$\begin{aligned} \max J &= 8.1x_1 + 10.8x_2 \\ \text{s.t.} \quad &\begin{cases} 0.80x_1 + 0.44x_2 \leq 24000 \\ 0.05x_1 + 0.10x_2 \leq 2000 \\ 0.10x_1 + 0.36x_2 \leq 6000 \\ x_1 \geq 0, x_2 \geq 0 \end{cases} \end{aligned}$$

计算结果：

最优解：最优值 $\max J = 286758.6$ ，最优
点 $x_1 = 26206.9, x_2 = 6896.552$

当按照每天加工 1[#]原油 26207 桶，2[#]原油 6879 桶的指标组织生产时，炼油厂可获最大利润，即 286759 元/日。

例 8.9 化工管道设计的最优化问题。

如图 8.7 所示，某地水源取自某水库，水库涵洞底标高为 45m，水送到调节水池距离为 1470m，调节水池最高水位为 35m（高差为 10m），该段距离中要求输水 $Q = 0.174\text{m}^3/\text{s}$ 。

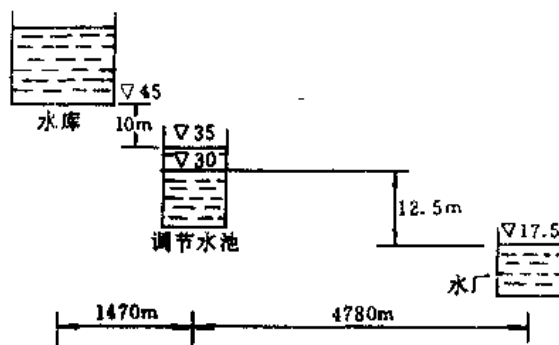


图 8.7 输水情况示意图

另一段，从调节水池输水到某水厂，调节水池低水位标高为 30m，水厂水池标高为 17.5m，高差为 12.5m，要求输水 $Q = 0.116\text{m}^3/\text{s}$ 。

可供铺设的输水管有四种不同口径，它们的单位长度造价及水头损失列于表 8.4 中。

表 8.4 输水管单位长度造价及水头损失数据

管径 D	单 价 (元/m)	单位长度水头损失 $h \times 10^3$ (m)	
		$Q = 0.174\text{m}^3/\text{s}$	$Q = 0.116\text{m}^3/\text{s}$
$\Phi 600$	110	0.873	0.419
$\Phi 500$	70	2.160	1.030
$\Phi 400$	54	6.760	3.120
$\Phi 300$	36	31.000	13.800

问怎样适当地选择这四种不同口径的输水管进行铺设，既能保证供水而又使造价最低。

解：对于第一段（即水库—调节水池），在 1470m 距离中，我们设：

$\Phi 600$ 口径的输水管铺设长度为 x_1, m ；

$\Phi 500$ 口径的输水管铺设长度为 x_2, m ；

(1) 400 口径的输水管铺设长度为 x_3 , m;

(2) 300 口径的输水管铺设长度为 x_4 , m。

为了保证供水, 要求

$$x_1 + x_2 + x_3 + x_4 = 1470 \quad (1)$$

另外, 还要求输水 $Q = 0.174 \text{ m}^3/\text{s}$ 时, 通过该段时总的水头损失不超过 10m, 即

$$0.873x_1 + 2.16x_2 + 6.76x_3 + 31x_4 \leq 10 \times 1000 \quad (2)$$

式 (1) 及式 (2) 即为两个约束方程。

而要求输水管道铺设的总造价最低就是目标函数, 即

$$J = 110x_1 + 70x_2 + 54x_3 + 36x_4 \rightarrow \min \quad (3)$$

这是一个线性规划问题, LP 模型为:

$$\begin{aligned} \max_{s.t.} J &= 110x_1 + 70x_2 + 54x_3 + 36x_4 \\ \begin{cases} x_1 + x_2 + x_3 + x_4 &= 1470 \\ 0.873x_1 + 2.16x_2 + 6.76x_3 + 31x_4 &\leq 10000 \\ x_j &\geq 0, (j = 1, 2, 3, 4) \end{cases} \end{aligned}$$

计算结果:

最优解: 最优值 $\min J = 79333.4$

最优点 $x_1 = 0, x_2 = 0, x_3 = 1467.4, x_4 = 2.59$

对第一段 1470m 长的部分:

(1) 400 口径的输水管铺设 1467.4m;

(2) 300 口径的输水管铺设 2.59m;

而总造价为 79333.4 元。

同样, 对于第二段 (即调节水池—水厂), 在 4780m 距离中, 我们设:

(1) 600 口径的输水管铺设长度为 x_1 , m;

(2) 500 口径的输水管铺设长度为 x_2 , m;

(3) 400 口径的输水管铺设长度为 x_3 , m;

(4) 300 口径的输水管铺设长度为 x_4 , m;

与上述同样道理, 我们又可列出一个 LP 问题:

$$\begin{aligned} \max_{s.t.} J &= 110x_1 + 70x_2 + 54x_3 + 36x_4 \\ \begin{cases} x_1 + x_2 + x_3 + x_4 &= 4780 \\ 0.419x_1 + 1.03x_2 + 3.12x_3 + 13.84x_4 &\leq 12500 \\ x_j &\geq 0 \quad (j = 1, 2, 3, 4) \end{cases} \end{aligned} \quad (4)$$

计算结果:

最优解: 最优值 $\min J = 276597.3$

最优点 $x_1 = 0, x_2 = 1154.83, x_3 = 3625.17, x_4 = 0$

即对第二段 4780m 长的部分:

(2) 500 口径的输水管铺设 1154.8m;

(3) 400 口径的输水管铺设 3625.2m;

而总造价为 276597.3 元。

8.3 非线性规划

8.3.1 非线性规划基本概念

1. 非线性规划的一般形式

先看两个实例。

例 8.10 某石化公司有 n 个工厂排放污水，拟集中于污水处理厂进行污水处理。已知工厂 j 的污水排放量为 Q_j ，由工厂 j 到污水处理厂输送每

一立方米污水的每公里污水费用为 C_j 。工厂到污水处理厂的路程简化为如图 8.8 所示。要决定污水处理厂的位置，使总输水投资最小。

可见，从工厂 j 到位于 (x_1, x_2) 的污水处理厂的输水费用为

$$Q_j \cdot C_j \sqrt{(x_1 - a_j)^2 + (x_2 - b_j)^2}$$

因此， n 个工厂的总费用是

$$f(x_1, x_2) = \sum_{j=1}^n Q_j \cdot C_j \sqrt{(x_1 - a_j)^2 + (x_2 - b_j)^2}$$

该问题对变量没有什么约束条件， $f(x_1, x_2)$ 为非线性函数，归结为求 (x_1, x_2) ，使 $f(x_1, x_2)$ 达到最小值。

例 8.11 有一笔资金 P 向某化工厂内的三个车间投资，设每个车间获得资金的数额 x_j 与所取的收益 J_j 有下列函数关系

$$J_j = A_j \cdot x_j^2 \quad (j = 1, 2, 3)$$

此外，由于 1、2 两车间的产品要配套生产，所以要求投资额 x_1 与 x_2 之间有下列非线性的关系：

$$x_1 \geq g(x_2)$$

为使投资获得最大收益，这笔资金 P 应如何分配？

根据问题的要求，容易看出，上述资金分配问题的数学模型为

$$\begin{aligned} \max_{x_i} J &= \sum_{j=1}^3 J_j = \sum_{j=1}^3 A_j x_j^2 \\ x_1 + x_2 + x_3 &\leq P \\ x_1 &\geq g(x_2) \\ x_j &\geq 0 \quad (j = 1, 2, \dots, 3) \end{aligned}$$

从上述例子可以看出，当目标函数或约束条件中有一个或多个为非线性函数，就称这样的规划问题为非线性规划 (Nonlinear Programming, 缩写 NLP)。科学研究和工程技术中所遇到的问题大量是非线性的，其数学模型一般为

$$\max_{\vec{x}} J = f(\vec{x}) \quad (8.10a)$$

$$h_k(\vec{x}) = 0 \quad (k = 1, 2, \dots, l) \quad (8.10b)$$

$$g_i(\vec{x}) \geq 0 \quad (i = 1, 2, \dots, m) \quad (8.10c)$$

其中， $\vec{x} = (x_1, x_2, \dots, x_n)^T$ 。式 (8.10a)、(8.10b) 和 (8.10c) 至少有一函数为非线

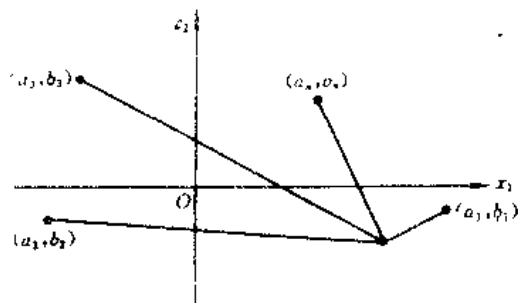


图 8.8 工厂到污水处理厂路程示意图

性函数。

因为 $\max f(\vec{x}) = -\min [-f(\vec{x})]$, 所以求目标函数的极小或极大, 其求解方法都是相同的。

若问题对 \vec{x} 没有约束条件, 即求 \vec{x}^* 使 $f(\vec{x})$ 达到最小值, 则称之为无约束非线性规划问题 (例 8.10), 否则就称之为有约束非线性规划问题。

2. 可行域和最优解

为了对非线性规划问题的可行域、目标函数等高线、最优解等有一个初步的认识, 考虑下面的例子。

例 8.12 讨论非线性规划问题。

$$\begin{aligned} \max_{s.t.} J &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ x_2 - x_1^2 &\geq -3 \\ x_2 &\leq 1 \\ x_1, x_2 &\geq 0 \end{aligned}$$

此问题的可行域是抛物线 $x_2 = x_1^2 - 3$ 、直线 $x_2 = 1$ 及 x_1 轴和 x_2 轴所围区域, 目标函数的等高线 $(x_1 - 3)^2 + (x_2 - 2)^2 = c$ 表示以点 $(3, 2)$ 为圆心的同心圆。因此, 上述问题即要在可行域内找一点, 使通过此点的同心圆半径最小。由图 8.9 可以看出, 所求的点 (即最优解) 为 $x_1 = 2, x_2 = 1$ 。

例 8.13 讨论非线性规划问题。

$$\begin{aligned} \max_{s.t.} J &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ x_2 - x_1^2 &\geq 0 \\ x_2 &\leq 4 \\ x_1, x_2 &\geq 0 \end{aligned}$$

此问题的可行域是抛物线 $x_2 = x_1^2$ 、直线 $x_2 = 4$ 及 x_1 轴和 x_2 轴所围区域, 目标函数的等高线与例 8.12 相同。由图 8.10 可以看出, 最优点在抛物线 $x_2 = x_1^2$ 与某条等高线的切点, 即 $x_1 = 1.57, x_2 = 2.46$ 。

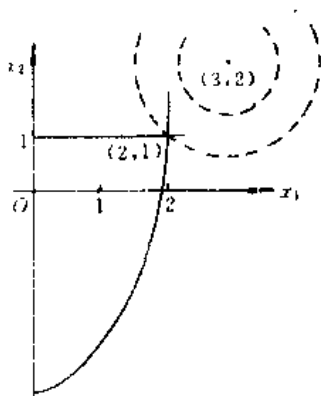


图 8.9 例 8.12 解的情况示意图

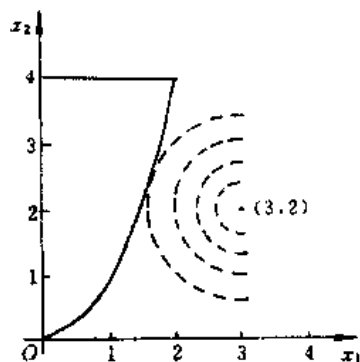


图 8.10 例 8.13 解的情况示意图

例 8.14 讨论非线性规划问题。

$$\max_{s.t.} J = (x_1 - 1)^2 + (x_2 - 3)^2$$

$$x_2 - x_1^2 \geq 0$$

$$x_2 \leq 4$$

$$x_1, x_2 \geq 0$$

此问题的可行域是抛物线 $x_2 - x_1^2 = 0$ 、直线 $x_2 = 4$ 及 x_1 轴和 x_2 轴所围区域。容易证实，目标函数的等高线的圆心 $(1, 3)$ 在可行域内。因此，最优解为等高线的圆心，即 $x_1 = 1, x_2 = 3$ ，如图 8.11 所示。

从上面的例子可以看出，对于只有两个变量的问题来说，非线性规划问题的可行域是由曲线围成的区域，其目标函数等高线是一族曲线，最优解可以取在可行域边界的一条曲线上，也可以取在可行域边界曲线的交点上，还可以取在可行域的内部。对于多变量的问题，也有类似的情形。可见，非线性规划问题远比线性规划问题复杂。这主要是在非线性情况下问题的几何性质起了根本的变化；而且非线性规划问题可能有局部最优解，它不同于全局最优解。

3. 局部最优解和凸规划

例 8.15 考虑 NLP 问题。

$$\begin{aligned} \max_{x \in I} J = f(x) &= \sin x + \frac{1}{2}x \\ 0 &\leq x \leq 2\pi \end{aligned}$$

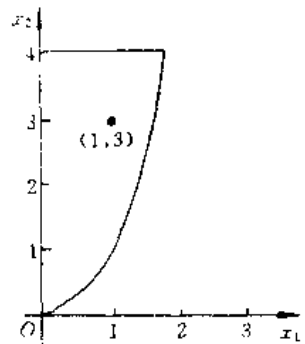


图 8.11 例 8.14 解的情况示意图

由图 8.12 可以看出，在整个可行域内，问题的最优解是 $x=0$ ，但是对 $x = \frac{4}{3}\pi$ 附近的任意 x ，有都大于 $x = \frac{4}{3}\pi$ 处的目标函数值，即 $f(x) \geq f(\frac{4}{3}\pi)$ ，因此 $x = \frac{4}{3}\pi$ 是一个局部极小点。这时称 $x=0$ 为整体最优解， $x = \frac{4}{3}\pi$ 为局部最优解。

当变量的个数不止一个时，也有局部最优解的问题。如图 8.13 所示，是一组目标函数的等高线，若点 P_1 和点 P_2 都在可行域内，则它们都是局部最优解。因此，只有求出所有的局部最优解，将它们加以比较，才能求出整体最优解。

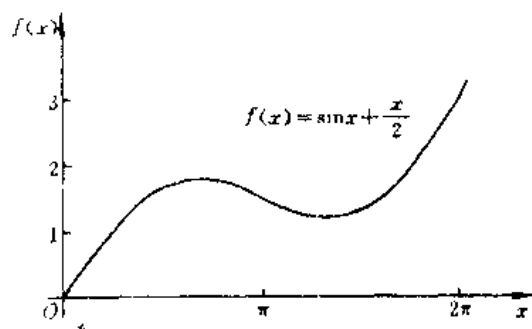


图 8.12 单变量函数局部最优解示意图

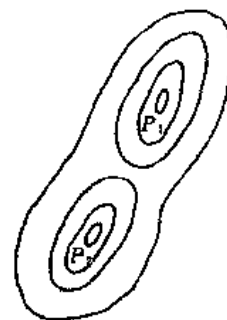


图 8.13 多变量函数局部最优解示意图

一般说来，若点 P^* 在可行域内，对于可行域内点 P^* 附近任意点 P ，相应目标函数值有下列关系

$$f(P) \geq f(P^*)$$

则称 P^* 为局部极小点，它是目标函数极小问题的一个局部最优解。若对可行域内任意点

P , 上述不等式成立, 则 P^* 是目标函数极小问题的整体最优解。显然, 对于目标函数的极大问题, 也可类似地定义它的局部最优解和整体最优解。

实际上, 局部最优和整体最优的概念是相对的, 由图 8.12 可以看出, 当可行域为 $0 \leq x \leq 2\pi$ 时, $x = \frac{4}{3}\pi$ 是局部最优解, 当可行域为 $\pi \leq x \leq 2\pi$ 时, $x = \frac{4}{3}\pi$ 不仅是局部最优解而且是整体最优解。因此, 人们关心的是, 目标函数和可行区域满足什么条件, 非线性规划问题的局部最优解就是整体最优解? 明确了这个条件, 不仅有助于求解局部最优就是整体最优的非线性规划问题。而且对于有若干个局部最优解的问题, 也可以看作在若干小区域内求整体最优解的问题。

对于一个变量的函数, 函数曲线朝下弯时, 称为凸函数, 如图 8.14 (a) 所示; 函数曲线朝上弯时, 称为凹函数, 如图 8.14 (b) 所示。显然, 目标函数是凸(凹)函数极小(大)问题, 局部最优解就是整体最优解。

对于多变量的函数 $f(\vec{x})$, 如果对任意 $\vec{x}^{(1)}, \vec{x}^{(2)}$ 及 $0 \leq \lambda \leq 1$, 有关系式

$$f(\lambda \vec{x}^{(1)} + (1 - \lambda) \vec{x}^{(2)}) \leq \lambda f(\vec{x}^{(1)}) + (1 - \lambda) f(\vec{x}^{(2)}) \quad (8.11)$$

则 $f(\vec{x})$ 称为凸函数。同理, 若 $-f(\vec{x})$ 是凸函数, 则 $f(\vec{x})$ 称为凹函数。

局部最优与整体最优的关系, 不仅依赖于目标函数, 而且也依赖于可行域的性质。

一个点的集合 C , 如果 $\vec{x}^{(1)}, \vec{x}^{(2)}$ 是 C 中任意两点, 过这两点所连的直线段均在 C 中, 即 $\lambda \vec{x}^{(1)} + (1 - \lambda) \vec{x}^{(2)} \in C$, 其中 $0 \leq \lambda \leq 1$, 则 C 称为凸集, 如图 8.15(a) 是凸集, b 不是凸集。

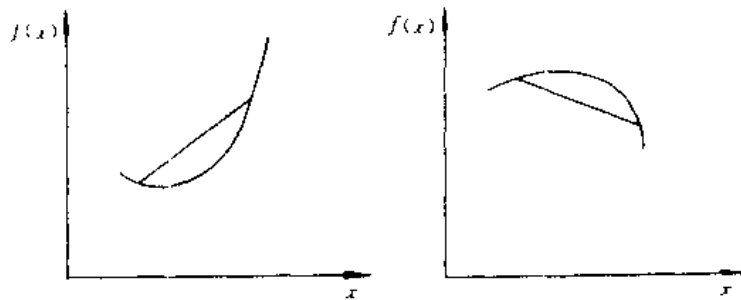


图 8.14

(a) 凸函数

(b) 凹函数

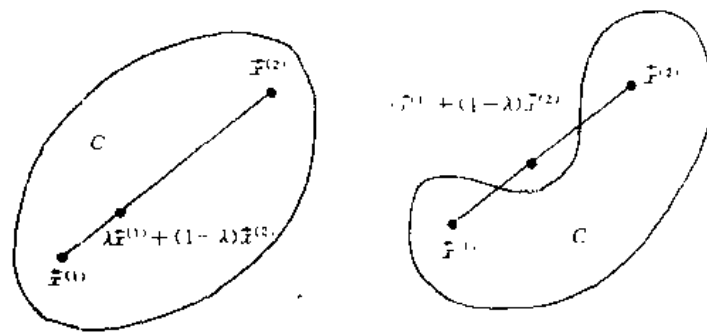


图 8.15

(a) 凸集

(b) 凹集

可以证明, 凸函数在凸可行域内的局部最小也是整体最小; 凹函数在凸可行域内的局部最大也是整体最大。

对于非线性规划问题(8.10),若目标函数 $f(\vec{x})$ 为凸函数, $h_k(\vec{x})$ 、 $g_i(\vec{x})$ 都是凸函数,这样的非线性规划问题称为凸规划。可以证明,凸规划的可行域为凸集,其局部最优解即为整体最优解。

由于线性函数既是凸函数,又是凹函数,所以可以说,线性规划问题是一个特殊的凸规划,它的局部最优解就是整体最优解。

为确定起见,下面我们总假设所研究的非线性规划问题是寻求目标函数的最小值,且所讨论的目标函数为严格的凸函数,即在所考虑的搜索区间内,函数有唯一的最小点。

8.3.2 无约束问题最优化

我们知道无约束最优化问题的一般形式是:求目标函数

$$\min f(\vec{x}), \vec{x} \in E^n \quad (8.12)$$

$\vec{x} \in E^n$ 表示 \vec{x} 为 n 维欧氏空间中的一个点,或称为 n 维向量,可表示为

$$\vec{x} = (x_1, x_2, \dots, x_n)^T$$

无约束最优化问题是非线性规划问题(8.10)的一种特殊情形。其求解方法的基本思想可以推广到有约束最优化问题上。

1. 单变量函数的黄金分割法

处理单变量函数的最优化方法除古典微分法(解析法)外,普遍采用的是直接法,或称一维搜索法。在处理多变量函数的非线性规划问题时,一旦搜索方向被确定之后,往往转化成一维搜索问题。因此,单变量函数的最优化方法是求解非线性规划问题最基本的方法。下面介绍单变量函数寻优的主要方法之一,黄金分割法。

黄金分割法是一维搜索法中消去法的一种。所谓消去法就是不断消去部分搜索区间,逐步缩小最优点的存在范围,以至最后求得最优点。

黄金分割法适用于区间 $[a, b]$ 上任何单值函数求极小点的问题。其基本思想是:在搜索区间 $[a, b]$ 内适当插入两点 x_1 和 x_2 ,它们将 $[a, b]$ 分为三段,通过比较这两点的函数值而确定删去最左端或最右端,这算迭代一次;然后再在保留下来的区间上作同样的处置。如此迭代下去,可将搜索区间无限缩小。现在提出一个问题,在每次迭代中如何确定两个插入点的位置?

设区间 $[a, b]$ 的长为 1。在与点 a 相距分别为 α 和 β 的点处插入 x_1 和 x_2 。为确定 α 和 β ,提出如下条件:

(1) x_1 和 x_2 两点在区间 $[a, b]$ 中的位置是对称的。这样,无论删去哪一段,总是保留长为 β 的区间,即有

$$\overline{ax_1} = \overline{x_2b}$$

即

$$\alpha + \beta = 1 \quad (8.13)$$

式中 $\overline{ax_1}$ 、 $\overline{x_2b}$ 分别表示区间 $[a, x_1]$ 、 $[x_2, b]$ 的长。

(2) 无论删掉哪一段,例如删去 $[x_2, b]$,在留下的长区间 $[a, x_2]$ 里,再插入一点 x_3 ,使得 x_3, x_1 在 $[a, x_2]$ 中的位置与 x_1, x_2 在 $[a, b]$ 中的位置具有相同的比例。这就保证每次迭代

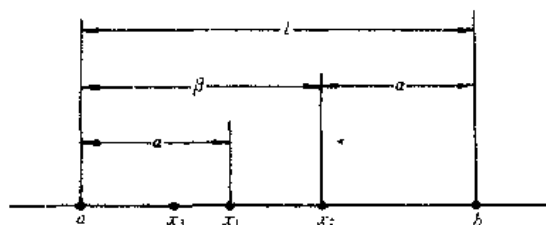


图 8.16 黄金分割法取点示意图

都以同一 β 的比率缩短区间, 根据这一条件有

$$\frac{\frac{ax_1}{ax_2}}{\frac{ax_2}{ab}} = \frac{\frac{ax_2}{ab}}{\frac{ax_2}{ab}}$$

即

$$\frac{\alpha}{\beta} = \frac{\beta}{1}$$

化简上式, 得

$$\alpha = \beta^2 \quad (8.14)$$

将式 (8.14) 代入式 (8.13), 得关于 β 的一元二次方程

$$\beta^2 + \beta - 1 = 0$$

解此方程共有两个根, 取正根得

$$\beta = \frac{-1 + \sqrt{5}}{2} = 0.6180339\cdots \approx 0.618$$

可见黄金分割法的分割点位于从搜索区间左端算起的 0.618 处, 因此黄金分割法又称为 0.618 法。

2. 多变量函数梯度法

解决多变量函数的最优化问题主要是用迭代法。迭代法的搜索过程实质上由两部分组成: 一是选定目标函数值下降的搜索方向; 二是在确定的方向上搜索。由于选取搜索方向与迭代的方式不同, 构成了不同的迭代法。梯度法是一种迭代法, 由于其有局部下降得最快的特性, 因此又称最速下降法。

梯度法主要包括以下三步:

(1) 选择初始点 $\vec{x}^{(0)}$ (越靠近最优解越好);

(2) 如已得出第 K 步迭代点 $\vec{x}^{(K)}$, $K \geq 0$, 且 $\vec{x}^{(K)}$ 为非最优解, 我们要建立一套规律以产生方向 \bar{P}_K , 使得目标函数 $f(\vec{x})$ 从 $\vec{x}^{(K)}$ 出发, 沿 \bar{P}_K 方向可以找到 $\vec{x}^{(K+1)}$, 有所下降。

我们知道函数 $f(\vec{x})$ 于一点 $\vec{x}^{(K)}$ 的梯度方向 $\nabla f(\vec{x}^{(K)})$ 是函数值上升最快的方向; 负梯度方向 $-\nabla f(\vec{x}^{(K)})$ 是 $f(\vec{x})$ 于点 $\vec{x}^{(K)}$ 附近下降得最快的方向。

这里梯度

$$\nabla f(\vec{x}^{(K)}) = \left[\frac{\partial f(\vec{x}^{(K)})}{\partial x_1}, \frac{\partial f(\vec{x}^{(K)})}{\partial x_2}, \dots, \frac{\partial f(\vec{x}^{(K)})}{\partial x_n} \right]^T$$

是一个 n 维向量。

从点 $\vec{x}^{(K)}$ 出发, 我们便可以构造出无约束极值问题的梯度法, 其迭代公式为

$$\vec{x}^{(K+1)} = \vec{x}^{(K)} - h_K \nabla f(\vec{x}^{(K)}) \quad (8.15)$$

式 (8.15) 中只有 h_K 为一待定的量, 我们称 h_K 为迭代法中第 K 步的步长因子。即在迭代过程中, 为了得到下一个近似极小点, 在选定了搜索方向之后, 还要确定步长 h_K , 关于步长 h_K 的选取, 可以有多种方法, 一种方法是取 h 为某一常数, 如, $h=1$ 进行试算 (称为定步长), 检验是否满足不等式

$$f(\vec{x}^{(K+1)}) = f[\vec{x}^{(K)} - h_K \nabla f(\vec{x}^{(K)})] < f(\vec{x}^{(K)}) \quad (8.16)$$

若上述不等式成立, 表示可以迭代; 否则, 缩小 h 使满足不等式 (8.16), 因为采用负梯度方向, 满足式 (8.16) 的 h 总是存在的。另一种方法是沿负梯度方向对函数进行一维搜索, 即

$$\min f[\vec{x}^{(K)} - h \nabla f(\vec{x}^{(K)})], h \geq 0 \quad (8.17)$$

来确定使 $f(\vec{x})$ 最小的 h_k (称为最优步长)。

(3) 检验新得出的迭代点 $\vec{x}^{(K+1)}$ 是否为最优解, 或是否为最优解的近似解, 检验的办法, 若梯度的长度小于某个预先指定的正数 ϵ , 即

$$\|\nabla f(\vec{x}^{(K+1)})\| \leq \epsilon \quad (8.18)$$

我们就认为 $\vec{x}^{(K+1)}$ 是近似最优解, 迭代过程也就终止了; 否则继续进行迭代。

梯度法收敛是比较慢的, 其原因在于该法的最速下降方向仅仅是指某点附近而言, 是一种局部性质。从局部看, 在一点附近函数值下降快了, 但从整体来讲, 则走了许多弯路, 反而收敛缓慢。尽管如此, 梯度法仍不失为一种基本的最优化方法。这不仅因为它具有算法简单, 使用方便等特点, 而且是其他一些重要方法的基础。

3. 多变量函数单纯形法

单纯形是指在一定的空间中最简单的图形。如一维空间中的单纯形是一条直线, 二维空间中是三角形, 三维空间中是四面体。由于维数太高时无法用几何图形直观地加以描述, 但可以说, 在 n 维空间中, 单纯形是指以 $n+1$ 个顶点所构成的最简单图形。

单纯形法是一种直接搜索法。它与解析法的区别在于, 它不是沿某个方向进行搜索, 而是对 n 维空间的 $n+1$ 个顶点 (构成一个初始单纯形的顶点) 上的函数值进行比较, 丢掉其中最坏的点, 代之以新的点, 从而构成一个新的单纯形, 逐步逼近极小点。因而单纯形法由两部分组成: 一是初始单纯形的形成, 即找出初始的 $n+1$ 个点; 二是该方法的迭代过程。

为了区别于线性规划的单纯形法, 有时我们称单纯形法为序贯单纯形法。

下面以求二元函数的极小点为例, 说明单纯形法的基本原理:

考虑无约束二元函数的最优化问题

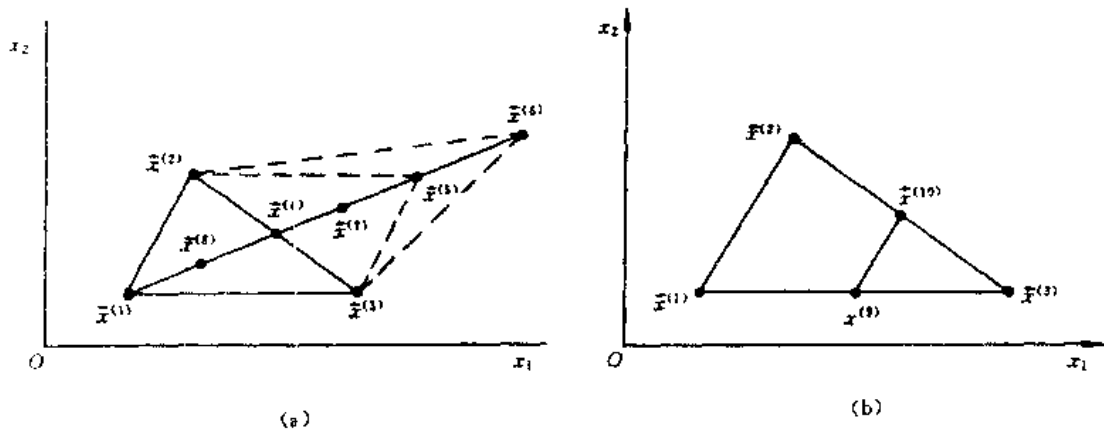
$$\min f(\vec{x}) = f(x_1, x_2)$$

在 $x_1 - x_2$ 平面上取不在同一条直线上的三个点 $\vec{x}^{(1)}$ 、 $\vec{x}^{(2)}$ 和 $\vec{x}^{(3)}$, 并以它们为顶点构成一单纯形——三角形。求出各顶点的函数值 $f(\vec{x}^{(1)})$ 、 $f(\vec{x}^{(2)})$ 和 $f(\vec{x}^{(3)})$, 并比较其大小, 如果有下面结果:

$$f(\vec{x}^{(1)}) > f(\vec{x}^{(2)}) > f(\vec{x}^{(3)})$$

这表明 $\vec{x}^{(1)}$ 点最差, $\vec{x}^{(3)}$ 点最好。为了寻找极小点, 一般应向最差点的反对称方向进行搜索。以 $\vec{x}^{(4)}$ 记 $\vec{x}^{(2)}\vec{x}^{(3)}$ 的中点 (图 8.17(a)), 在 $\vec{x}^{(1)}\vec{x}^{(4)}$ 的延长线上取点 $\vec{x}^{(5)}$, 使

$$\vec{x}^{(5)} = \vec{x}^{(4)} + (\vec{x}^{(4)} - \vec{x}^{(1)}) = 2\vec{x}^{(4)} - \vec{x}^{(1)} \quad (8.19)$$



• 图 8.17 单纯形法示意图

算出 $f(\vec{x}^{(5)})$ 的函数值 $f(\vec{x}^{(5)})$, 可能出现以下几种情形。

$$(1) f(\vec{x}^{(5)}) < f(\vec{x}^{(3)})$$

这说明搜索方向正确, 还可以进一步扩大效果, 沿 $\vec{x}^{(1)}\vec{x}^{(5)}$ 方向向前搜索, 取

$$\vec{x}^{(6)} = \vec{x}^{(4)} + \alpha(\vec{x}^{(5)} - \vec{x}^{(1)}) \quad (8.20)$$

其中 α 为延伸因子, 一般取 $\alpha = 1.2 \sim 2.0$ 。

如果 $f(\vec{x}^{(6)}) < f(\vec{x}^{(5)})$, 说明延伸有利, 就以 $\vec{x}^{(6)}$ 点代替 $\vec{x}^{(1)}$ 点, 构成新的单纯形 $\{\vec{x}^{(2)}, \vec{x}^{(3)}, \vec{x}^{(6)}\}$;

如果 $f(\vec{x}^{(6)}) > f(\vec{x}^{(5)})$, 表示不利, 舍弃 $\vec{x}^{(6)}$, 仍以 $\vec{x}^{(5)}$ 代替 $\vec{x}^{(1)}$ 构成新的单纯形 $\{\vec{x}^{(2)}, \vec{x}^{(3)}, \vec{x}^{(5)}\}$ 。

$$(2) f(\vec{x}^{(2)}) < f(\vec{x}^{(5)}) < f(\vec{x}^{(1)})$$

表示 $\vec{x}^{(5)}$ 走得太远, 应缩回一些, 若以 β 表示压缩因子, 则有

$$\vec{x}^{(7)} = \vec{x}^{(4)} + \beta(\vec{x}^{(5)} - \vec{x}^{(4)}) \quad (8.21)$$

$\beta < 1$, 通常取 $\beta = 0.5$ 。

$$(3) f(\vec{x}^{(5)}) > f(\vec{x}^{(1)})$$

表示 $\vec{x}^{(5)}$ 点比原来的 $\vec{x}^{(1)}$ 还差, 这时应压缩得更多一些, 即将新点压缩至 $\vec{x}^{(1)}$ 与 $\vec{x}^{(4)}$ 之间:

$$\vec{x}^{(8)} = \vec{x}^{(4)} - \beta(\vec{x}^{(4)} - \vec{x}^{(1)}) = \vec{x}^{(4)} + \beta(\vec{x}^{(1)} - \vec{x}^{(4)}) \quad (8.22)$$

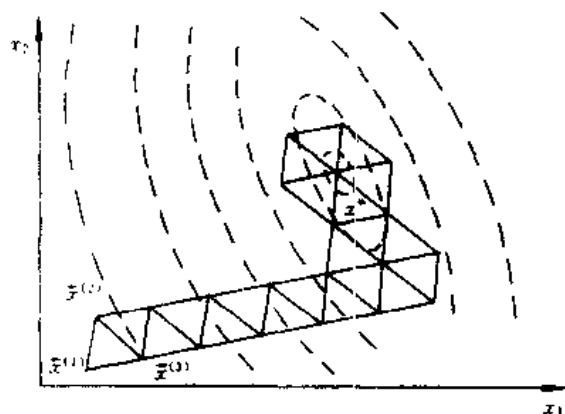


图 8.18 单纯形法搜索路径示意图

(4) 若 $\vec{x}^{(1)}\vec{x}^{(4)}$ 方向上所有点的函数值 $f(\vec{x})$ 都大于 $f(\vec{x}^{(1)})$, 则不能沿此方向搜索。这时, 可以以 $\vec{x}^{(3)}$ 为中心进行缩边, 即使顶点 $\vec{x}^{(1)}$ 和 $\vec{x}^{(2)}$ 向 $\vec{x}^{(3)}$ 移近一半距离 (图 8.17(b)) 得新的单纯形 $\{\vec{x}^{(3)}, \vec{x}^{(9)}, \vec{x}^{(10)}\}$ 。以此单纯形为基础进行寻优。

以上说明, 不管哪种情况, 我们都能得到一个新的单纯形, 其中至少有一个顶点的函数值比原单纯形为小, 并以此为基础寻优, 完成一步搜索过程。搜索路径示意图如图 8.18 所示。

对 n 维变量, 因为有 $n+1$ 个顶点, 计算工作量较大, 但原理和上述二维情况相同, 不再赘述。

BIPTCHEM 在求解无约束最优化问题时是直接调用单纯形法子程序进行计算。

例 8.16 用 BIPTCHEM 软件包求解。

$$\min_{x_1, x_2} f(x_1, x_2) = 4(x_1 - 5)^2 + (x_2 - 6)^2$$

解: 上机求解, 操作步骤如下。

(1) 按第十章软件使用说明运行 BIPTCHEM, 打开优化菜单, 选择非线性规划 (拉格朗日乘子法和单纯形法), 按 <Enter> 键。

(2) 屏幕显示:

1. 无约束
2. 有约束

时,用方向键(↓键或↑键)选择无约束,按〈Enter〉键。

(3)根据屏幕提示,输入基本变量数“2”。

(4)当屏幕显示:

```

请输入目标函数:F=f(x1,x2,...,x2)
如:F=x1^2+x2^2-x1*x2-10*x1-4*x2+60
目标函数:F=
  
```

时,在F=后面键入“ $4 * (x1 - 5)^2 + (x2 - 6)^2$ ”,按〈Enter〉键。屏幕询问“是否正确?”时进行确认。

(5)根据屏幕提示,选择目标函数求最小值。

(6)屏幕显示压缩因子(默认值为0.75)时,按〈Enter〉键。同样方法处理屏幕显示的其它几个量。

(7)回答是否打印。输入各基本变量的初值(如取默认值1,可直接按〈Enter〉键)。

计算完毕,屏幕显示:

```

目标函数:F=4*(x1-5)^2+(x2-6)^2
最优解为:x1(基本变量)=5
          x2(基本变量)=6
目标函数最小值=5.96e-8
  
```

例 8.17 换热器系列的最优设计问题

某化工厂欲利用本厂废热,采用三个换热器将某物料温度由 100°C 加热到 500°C ,其流程和有关数据如图 8.19 所示,问各温度如何选取才能使换热器系列的总传热面 F 为最小?已知情况为(除温度外,均为国际单位):

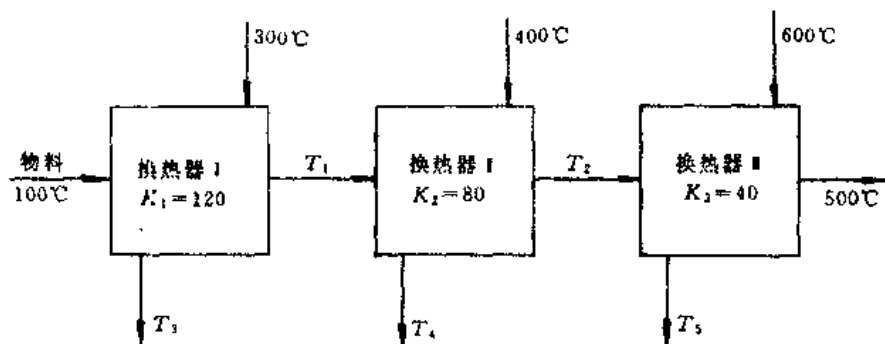


图 8.19 换热器系列示意图

- (1) 所有物流其 $w_i c_{pi} = 10^5$, 其中 w 是流体流量, c_p 为流体比热, i 为流体序号, $i = 1, 2, \dots$;
- (2) 三个换热器的总传热系数分别为 $K_1 = 120$, $K_2 = 80$, $K_3 = 40$;
- (3) 为简化起见,换热器采用逆流换热时,其温差 Δt_m 采用算术平均值。各数值略去量纲。

解：根据已知条件和要求，设计计算步骤如下。

1. 建立数学模型

由系统的热平衡方程可得

$$wc_p \Delta T = FK \Delta t_m$$

各物流的 wc_p 均相等，因此

对换热器 I 有： $wc_p(T_1 - 100) = wc_p(300 - T_3)$

得
$$T_1 - 100 = 300 - T_3, T_3 = 400 - T_1$$

对换热器 II 有： $wc_p(T_2 - T_1) = wc_p(400 - T_4)$

得
$$T_2 - T_1 = 400 - T_4, T_4 = 400 - T_2 + T_1$$

对换热器 III 有： $wc_p(500 - T_2) = wc_p(600 - T_5)$

得
$$500 - T_2 = 600 - T_5, T_5 = 100 + T_2$$

各换热器的温差均取算术平均值，则有

$$\Delta t_{m1} = \frac{(300 - T_1) + (T_3 - 100)}{2} = 300 - T_1$$

$$\Delta t_{m2} = \frac{(400 - T_2) + (T_4 - T_1)}{2} = 400 - T_2$$

$$\Delta t_{m3} = \frac{(600 - 500) + (T_5 - T_2)}{2} = 100$$

各换热器的面积由热平衡方程得

$$F_1 = \frac{10^5(T_1 - 100)}{120(300 - T_1)} \quad F_2 = \frac{10^5(T_2 - T_1)}{80(400 - T_2)} \quad F_3 = \frac{10^5(500 - T_2)}{40 \times 100}$$

则本例可化为如下无约束问题：

$$\min F = \sum_{i=1}^3 F_i = \frac{10^5(T_1 - 100)}{120(300 - T_1)} + \frac{10^5(T_2 - T_1)}{80(400 - T_2)} + \frac{10^5(500 - T_2)}{4000}$$

由目标函数式看出，此例属于二维问题，用单纯形法进行求解。

2. 上机求解

上机求解过程如下。

(1) 运行 BIPTCHEM，打开优化菜单，选择非线性规划（拉格朗日乘子法和单纯形法），按〈Enter〉键。

(2) 根据屏幕提示，选择无约束；输入基本变量数“2”。

(3) 在屏幕提示“F=”后面键入

“ $1e5 * (x1 - 100) / 120 (300 - x1) + 1e5 * (x2 - x1) / 80 (400 - x2) + 1e5 * (500 - x2) / 400$ ”

按〈Enter〉键，并进行确认。

(4) 选择目标函数求最小值，输入压缩因子等量，回答是否打印。

(5) 输入各基本变量的初值后，屏幕显示：

目标函数:

$$F = 1e5 * (x_1 - 100)/120/(300 - x_1) + 1e5 * (x_2 - x_1)/80/(400 - x_2) + 1e5 * (500 - x_2)/4000$$

最优解为: x_1 (基本变量) = 182.0

x_2 (基本变量) = 295.6

目标函数最小值 = 7049.25

因此, 本例计算结果为

$$F = 7049.25$$

8.3.3 有约束问题最优化

前面我们已讨论了无约束条件下的最优化方法, 但是大多数实际问题都是有约束的。一般说来, 求解有约束非线性规划问题是一件非常复杂的工作, 从目前来看求解方法大致可以分为两类: 把有约束问题转化为一系列无约束问题, 把有约束非线性规划问题转化为线性规划问题求解。

1. 最优性条件

首先考虑仅包含等式约束的 NLP 问题

$$\max_{\vec{x}} J = f(\vec{x}) \quad (8.23a)$$

$$h_k(\vec{x}) = 0 \quad (k = 1, 2, \dots, l) \quad (8.23b)$$

这个问题的最优性条件与求解方法在数学分析中已从理论上得到解决。在这里仅把拉格朗日 (Lagrange) 定理重述如下。

假设:

- (1) \vec{x}^* 是 NLP 问题 (8.23) 的局部最优解;
- (2) f, h_1, h_2, \dots, h_l 在 \vec{x}^* 的某一领域内连续可微;
- (3) $\nabla h_1(\vec{x}^*), \nabla h_2(\vec{x}^*), \dots, \nabla h_l(\vec{x}^*)$ 线性无关。那么, 存在一组实数 $\lambda_1^*, \lambda_2^*, \dots, \lambda_l^*$ 使得

$$\nabla f(\vec{x}^*) - \sum_{k=1}^l \lambda_k^* \nabla h_k(\vec{x}^*) = \vec{0} \quad (8.24)$$

式 (8.24) 是问题 (8.23) 的必要条件。

对于含有等式和不等式约束的 NLP 问题 (8.10), 其最优性条件是库恩-塔克 (Kuhn-Tucker) 条件, 它是非线性规划领域中最重要理论成果之一, 是确定某一点为极值点的必要条件, 但一般说来它并不是充分条件, 对于凸规划来说, 库恩-塔克条件不仅是必要条件, 而且也是充分条件。现将这一条件叙述如下。

假设在问题 (8.10) 中:

- (1) \vec{x}^* 是一局部最优解;
- (2) 记 I 为使得 $g_i(\vec{x}^*) = 0$ 的那些下标 i 的集合, 即 $I = \{i | g_i(\vec{x}^*) = 0, i = 1, 2, \dots, m\}$;
- (3) $f(\vec{x}), g_1(\vec{x}), \dots, g_m(\vec{x}), h_1(\vec{x}), \dots, h_l(\vec{x})$ 在点 \vec{x}^* 处可微;
- (4) 对于所有 $i \in I$ 的 $\nabla g_i(\vec{x}^*)$ 与 $\nabla h_1(\vec{x}^*), \dots, \nabla h_l(\vec{x}^*)$ 线性无关。

那么, 存在不全为零的一组数 $\mu_1, \mu_2, \dots, \mu_m, \lambda_1, \dots, \lambda_l$ 使得

$$\nabla f(\vec{x}^*) = \sum_{i=1}^m \mu_i \nabla g_i(\vec{x}^*) - \sum_{k=1}^l \lambda_k \nabla h_k(\vec{x}^*) = \vec{0} \quad (8.25a)$$

$$\mu_i g_i(\vec{x}^*) = 0 \quad i = 1, 2, \dots, m \quad (8.25b)$$

$$\mu_i \geq 0 \quad i = 1, 2, \dots, m \quad (8.25c)$$

式(8.25)即为问题(8.10)的库恩-塔克条件的结论式。式(8.25a)表明,若 \vec{x}^* 是极小点,则目标函数的梯度 $\nabla f(\vec{x}^*)$ 可以用起作用的约束函数的梯度 $\nabla g_i(\vec{x}^*)$ 和 $\nabla h_k(\vec{x}^*)$ ($i \in I$ 和 $k = 1, 2, \dots, l$)线性表出;式(8.25b)是一个互补条件,表明 $g_i(\vec{x}^*)$ 如果是一个起作用约束,则 $\mu_i \geq 0$,否则 $\mu_i = 0$,式(8.25c)表明 μ_i 不能取负值。而 λ_i 则可正可负。

以上定理和条件是有约束NLP问题最优化方法的理论基础。

2. 拉格朗日乘子法

拉格朗日定理的理论意义还在于,它将等式约束问题(8.23)的求解转化为无约束问题的求解。

假如我们构造一个 $n+l$ 的函数

$$L(\vec{x}, \lambda_1, \lambda_2, \dots, \lambda_l) = f(\vec{x}) - \sum_{k=1}^l \lambda_k h_k(\vec{x}) \quad (8.26)$$

这一函数称之为拉格朗日函数,其中 $\lambda_1, \lambda_2, \dots, \lambda_l$ 称为拉格朗日乘子,拉格朗日函数(8.26)的梯度为:

$$\nabla L = \begin{bmatrix} \nabla_{\vec{x}} L \\ \nabla_{\vec{\lambda}} L \end{bmatrix}$$

其中

$$\nabla_{\vec{x}} L = \nabla f(\vec{x}) - \sum_{k=1}^l \lambda_k \nabla h_k(\vec{x})$$

$$\nabla_{\vec{\lambda}} L = (-h_1(\vec{x}), -h_2(\vec{x}), \dots, -h_l(\vec{x}))^T$$

由此可以看到,无约束问题

$$\min L(\vec{x}, \lambda_1, \lambda_2, \dots, \lambda_l) \quad (8.27)$$

的最优性必要条件

$$\nabla L(\vec{x}^*, \lambda_1^*, \lambda_2^*, \dots, \lambda_l^*) = \vec{0} \quad (8.27)$$

恰好给出了约束问题(8.23)的最优性条件(8.24)以及

$$h_k(\vec{x}^*) = 0 \quad k = 1, 2, \dots, l \quad (8.28)$$

式(8.28)这组等式方程表示,最优解 \vec{x}^* 必须满足约束。若将式(8.24)和式(8.28)中的 \vec{x}^* 换成 \vec{x} 后所得到的 $n+l$ 个方程构成以 $x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_l$ 为未知数的非线性方程组,这个方程组的解 $(x_1^*, x_2^*, \dots, x_n^*, \lambda_1^*, \lambda_2^*, \dots, \lambda_l^*)^T$ 就是拉格朗日是函数的驻点,而 $\vec{x}^* = (x_1, x_2, \dots, x_n)^T$ 就有可能是问题(8.10)的最优点。

因此,拉格朗日乘子法根据最优性条件将等式约束问题最后转化成 $n+l$ 个未知数 $n+l$ 个非线性方程所构成的方程组求解问题。此非线性方程组的求解也是十分困难的,为此,我们可以引入一个新的函数,便于在计算机上用迭代法求解,新函数形式为:

$$\begin{aligned} Z &= (\nabla_{\vec{x}} L)^2 + (\nabla_{\vec{\lambda}} L)^2 \\ &= \sum_{j=1}^n \left(\frac{\partial L}{\partial x_j} \right)^2 + \sum_{k=1}^l (h_k(\vec{x}))^2 \end{aligned} \quad (8.29)$$

然后对 Z 函数求极小。 Z 函数求极小可以采用无约束问题的最优化方法。

此外，用拉格朗日乘子法也可以处理具有不等式约束条件的优化问题。对于不等式约束条件可以采用线性规划中处理不等式的方法，引入相应的剩余变量，将不等式化为等式，然后再用拉格朗日法求解。

例如对不等式约束

$$g_i(\vec{x}) \geq 0 \quad i = 1, 2, \dots, m$$

我们可以引入 m 个剩余变量 $(x_{n+1}, x_{n+2}, \dots, x_{n+i}, \dots, x_{n+m})$ ，由于此处的变量没有非负的要求，因此，为了保证不等式成立，对引入的剩余变量均用平方项，以保证该引入项非负。即

$$g_i(\vec{x}) - (x_{n+i})^2 \geq 0 \quad i = 1, 2, \dots, m \quad (8.30)$$

同样，根据库恩-塔克条件我们可以构造出不等式约束下的拉格朗日函数，即

$$L(\vec{x}, \vec{\mu}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^m \mu_i g_i(\vec{x}) - \sum_{k=1}^l \lambda_k h_k(\vec{x}) \quad (8.31)$$

我们称 $L(\vec{x}, \vec{\mu}, \vec{\lambda})$ 为问题(8.10)的广义拉格朗日函数， $\vec{\mu}, \vec{\lambda}$ 为拉格朗日乘子。

例 8.18 求解下列不等式约束 NLP 问题。

$$\begin{aligned} \min_{s.t.} J = f(\vec{x}) &= 2x_1^2 - 2x_1x_2 + 2x_2^2 - 6x_1 \\ &3x_1 + 4x_2 \leq 6 \\ &-x_1 + 4x_2 \leq 2 \end{aligned}$$

解：

(1) 运行 BIPTCHEM，打开优化菜单，选择非线性规划（拉格朗日乘子法和单纯形法），按〈Enter〉键。

(2) 根据屏幕提示，选择有约束；输入基本变量数“2”和约束数“2”。

(3) 在屏幕提示“F=”后面键入“ $2 * x1^2 - 2 * x1 * x2 + 2 * x2^2 - 6 * x1$ ”，按〈Enter〉键，并进行确认。

(4) 根据屏幕提示，分别输入约束方程 1 和约束方程 2，约束方程 1 为：

$$3 * x1 + 4 * x2 - 6 < = 0$$

约束方程 2 为：

$$-x1 + 4 * x2 - 2 < = 0$$

(5) 选择目标函数求最小值，输入压缩因子等量，回答是否打印。

(6) 输入各变量的初值，计算结束后屏幕显示：

目标函数: $F = 2 * x1^2 - 2 * x1 * x2 + 2 * x2^2 - 6 * x1$

约束条件:

$$3 * x1 + 4 * x2 - 6 < = 0$$

$$-x1 + 4 * x2 - 2 < = 0$$

最优解为: $x1 = (\text{基本变量}) = 1.459$

$$x2(\text{基本变量}) = 0.4054$$

目标函数最小值 = -5.351

3. 序贯二次规划法 (SQP 法)

求解有约束非线性规划问题 (8.10), Wilson-Han-Powell 的序贯二次规划法是当前公认的最有效的方法之一, 在实际工程中 (如电力、冶金、化工等) 的应用越来越广泛, 特别在石油化工领域中, 序贯二次规划法给复杂大系统的模拟与优化计算带来突破性的进展。

序贯二次规划法的发展可以追溯到 Wilson (1963) 的研究。以后, 经过 Biggs (1975)、Han (1975, 1977) 的发展和改进, 最后由 Powell (1977) 完成, 形成目前应用的基本形式。由于这种方法每一步迭代均需要解一个二次规划子问题, 因而称作序贯二次规划方法 (Successive Quadratic Programming), 简称 SQP 法。

SQP 法的基本思想

SQP 法的基本思想是: 在某个 $\vec{x}^{(K)}$ 近似解处, 将原非线性规划问题 (8.10) 化成如下的规划问题。

$$\min_{\vec{x}} J = \nabla f(\vec{x}^{(K)}) \nabla \vec{x} + \frac{1}{2} \nabla \vec{x}^T Q^{(K)} \nabla \vec{x} \quad (8.32a)$$

$$h_k(\vec{x}^{(K)}) + \nabla \vec{x}^T \nabla h_k(\vec{x}^{(K)}) = 0 \quad (8.32b)$$

$$g_i(\vec{x}^{(K)}) + \nabla \vec{x}^T \nabla g_i(\vec{x}^{(K)}) \geq 0 \quad (8.32c)$$

$$k = 1, 2, \dots, l \quad i = 1, 2, \dots, m$$

式中 $\nabla \vec{x} = \vec{x} - \vec{x}^{(K)}$

$Q^{(K)}$ 是拉格朗日函数 (8.31) 在点 $\vec{x}^{(K)}$ 处的赫森矩阵 $\nabla_{\vec{x}}^2 L$ 的近似。问题 (8.32) 为一个目标函数是二次的, 约束条件是线性的二次规划问题。这样, 原来的问题就用这个二次规划来近似地代替, 求解这一二次规划问题所得到的最优解就作为原问题的一个新的近似解。如果 $\vec{x}^{(K+1)}$ 与 $\vec{x}^{(K)}$ 充分接近, 比如, 它们的距离

$$\|\vec{x}^{(K+1)} - \vec{x}^{(K)}\| = \sqrt{\sum_{j=1}^n (\vec{x}_j^{(K+1)} - \vec{x}_j^{(K)})^2} \leq \varepsilon$$

(式中 ε 为某个给定的小数), 则以 $\vec{x}^{(K+1)}$ 为最优解, 否则以 $\vec{x}^{(K+1)}$ 代替 $\vec{x}^{(K)}$ 构造新的二次规划问题继续进行迭代。

二次规划子问题的求解可以运用前已述及的拉格朗日技术。其使用方法本书不作详细讨论。有关内容可参见文献 [1], [8]。读者需要此法的程序时可向本篇作者索取。

4. 复合形法 (直接法)

在 8.3.2 中, 曾经介绍了求解无约束问题的单纯形法。复合形法是单纯形法对有约束问题的推广, 是求解有约束 NLP 问题的直接搜索法, 它用以求解形式为

$$\min_{\vec{x}} J = f(\vec{x}) \quad (8.33a)$$

$$g_i(\vec{x}) \geq 0, \quad i = 1, 2, \dots, m \quad (8.33b)$$

$$a_j \leq x_j \leq b_j, \quad j = 1, 2, \dots, n \quad (8.33c)$$

一般来说, 满足变量的上下界并不一定与满足约束方程 $g_i(\vec{x}) \geq 0$ 相对应。另外, 复合形法不能处理非线性的等式约束问题。

在 n 维空间中, 由 $k = n + 1$ 个顶点可构成最简单的几何图形, 这是单纯形法最基本的概念。为了克服单纯形法容易产生退化的缺点, 复合形法采用了 $k \geq n + 2$ 个顶点。由于顶点的个数 $k > n + 1$, 所以叫做复合形法 (Complex 法)。复合形法要求有 k 个可行点, 这是很困难的, 一般是给出一个可行点, 再随机地产生其他 $k - 1$ 个顶点。

复合形法的计算步骤与单纯形法类似, 其不同点是在求取反射点时, 不仅要求其目标函

数值下降。而且要满足显示约束 (8.33c)；由于约束的存在，不再作扩张运算，只作反射和收缩。计算步骤请参阅文献[1]。

例 8.19 解下列 NLP 问题：

$$\min J = f(X) = (x_1 - 5)^2 + (x_2 - 5)^2 + 3 \quad (1a)$$

$$16 - x_1^2 - x_2^2 \geq 0 \quad (1b)$$

$$0 \leq x_1 \leq 10 \quad (1c)$$

$$0 \leq x_2 \leq 10 \quad (1d)$$

解：

(1) 运行 BIPTCHEM，打开优化菜单，选择非线性规划（复合形法），按〈Enter〉键。

(2) 根据屏幕提示，选择有约束；输入基本变量数“2”和约束数“1”（除变量上、下限约束）。

(3) 输入第一变量的约束下限“0”，上限“10”；输入第二变量的约束下限“0”，上限“10”。

(4) 在屏幕提示“F=”后面键入“(x1-5)^2+(x2-5)^2+3”，按〈Enter〉键，并进行确认。

(5) 根据屏幕提示，分别输入约束方程“16-x1^2-x2^2>=0”，并进行确认。

(6) 选择目标函数求最小值，回答是否打印。

(7) 输入各变量的初值，计算结束后屏幕显示：

目标函数：F=(x1-5)^2+(x2-5)^2+3

约束条件：

$$16 - x_1^2 - x_2^2 \geq 0$$

$$0 \leq x_1 \leq 10$$

$$0 \leq x_2 \leq 10$$

收敛精度：7.965E-5

最优解：x1=2.8276 x2=2.8292

目标函数最小值：F=12.43148

和无约束多变量非线性规划的单纯形法一样，复合形法的收敛速度是很慢的，特别是约束条件多时，为了得到一个较好的新顶点，往往要向中心点多次收缩，工作量较大。但对维数不高，约束个数不多，且目标函数与约束条件函数很复杂时，使用复合形法自然有它的优越性。

习 题

1. 试写出下列问题的数学模型并求解。

(1) 某产品重量为 75kg，要用 A、B 两种原料制成，每单位 A 种原料成本为 2 元，每单位 B 种原料成本为 8 元，该产品至少需要含 14 单位 B 种原料，最多含 20 单位 A 种原料。每单位 A 种原料重 2.5kg，每单位 B 种原料重 5kg，为使成本最小，该产品中 A、B 两种原料应各占多少？

(2) 某化工厂生产 A 和 B 两种产品，已知制造产品 A，每公斤要用煤 9000kg、电力 16kW、劳力 3 个；制造产品 B，每公斤要用煤 4000kg、电力 20kW、劳力 10 个。又知制成

产品 A, 每公斤的产值是 7 万元; 制成产品 B, 每公斤的产值是 12 万元, 现该工厂只有煤 360,000kg、电力 800kW, 劳力 300 个, 问在这种条件下, 应该生产 A、B 产品各多少, 才能使产值为最高。

2. 分别用图解法和 BIPTCHEM 软件求解下列线性规划问题:

$$(1) \min J = -x_1 + x_2$$

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

$$(2) \max J = 30x_1 + 20x_2$$

$$\begin{cases} x_1 + x_2 \geq 1 \\ x_1 - x_2 \geq -1 \\ 3x_1 + 2x_2 \leq 6 \\ x_1 - 2x_2 \leq 1 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

3. 求解下列线性规划问题:

$$(1) \max J = 10x_1 + 25x_2 + 30x_3$$

$$\begin{cases} x_1 + x_2 + x_3 \leq 12 \\ 2x_1 + 3x_2 \leq 20 \\ x_3 \leq 5 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(2) \max J = 3x_1 + 2x_2 + x_3$$

$$\begin{cases} 6x_1 - x_2 + 4x_3 \geq 21 \\ 2x_1 + 3x_2 - 5x_3 \geq 35 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(3) \max J = 2x_1 + 3x_2 + 5x_3$$

$$\begin{cases} x_1 + x_2 \leq 10 \\ 3x_1 + 4x_2 \leq 30 \\ x_1 - x_3 \leq -1 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(4) \min J = -x_1 + 2x_2 + x_3$$

$$\begin{cases} 3x_1 - x_2 + 4x_3 \leq -8 \\ x_1 - x_3 = 2 \\ x_1 - x_2 \leq 4 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(5) \max J = -x_1 + 2x_2 + x_3$$

$$\begin{cases} 2x_1 - x_2 - x_3 \geq -4 \\ x_1 + 2x_2 \leq 6 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(6) \min J = -x_1 + 2x_2 + x_3$$

$$\begin{cases} 2x_1 - x_2 + x_3 \geq -4 \\ x_1 + 2x_2 \leq 6 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(7) \max J = 3x_1 + 2x_2$$

$$\begin{cases} 2x_1 + 4x_2 - x_3 \leq 5 \\ -x_1 + x_2 - x_3 \geq -1 \\ x_2 - x_3 \geq -1 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3)$$

$$(8) \min J = 3x_1 + x_2 + x_3 + x_4$$

$$\begin{cases} -2x_1 + 2x_2 + x_4 = 4 \\ 3x_1 + x_2 + x_4 = 6 \end{cases}$$

$$x_j \geq 0 \quad (j=1, 2, 3, 4)$$

4. 求解非线性规划问题

$$(1) \min J = f(\vec{x}) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2$$

$$\text{初始点: } \vec{x}^{(0)} = (1, 1)^T.$$

$$(2) \min J = f(\vec{x}) = 4(x_1 - 5)^2 + (x_2 - 6)^2$$

$$\text{初始点: } \vec{x}^{(0)} = (8, 9)^T.$$

$$(3) \min J = f(\vec{x}) = (x_1 + x_2 - x_3 - 1)^2 + (x_1 + x_2)^2 + 5x_1^2$$

$$2x_1 + x_3 = 0$$

5. 求解非线性规划问题

$$(1) \min_{s.t.} J = f(X) = 2x_1^2 - 4x_1x_2 + 4x_2^2 - 6x_1 - 3x_2$$

$$x_1 + x_2 \leq 3$$

$$4x_1 + x_2 \leq 9$$

$$x_1, x_2 \geq 0$$

$$(2) \min_{s.t.} J = f(X) = -x_1^2 + x_1x_2 - 2x_2^2 + x_1 + x_2$$

$$x_1 - x_2 \geq 3$$

$$x_1 + x_2 = 4$$

$$(3) \min_{s.t.} J = f(\vec{x}) = 3x_1e^{-0.1x_1x_6} + 4x_2 + x_3^2 + 7x_4 + \frac{10}{x_5} + x_6$$

$$-x_4 + x_5 - x_6 = 0.1$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 10$$

$$2x_1 + x_2 + x_3 + 3x_4 \geq 2$$

$$8x_1 + 3x_2 + 4x_3 + x_4 + x_5 \leq 10$$

$$2x_1 + 6x_2 + x_3 + 3x_4 + x_6 \leq 13$$

$$x_1 + 4x_2 + 5x_3 + 2x_4 \leq 18$$

$$-20 \leq x_j \leq +20 \quad j=1, 2, \dots, 6$$

$$\text{初始点 (不可行): } x_j^{(0)} = 1.0 \quad j=1, 2, \dots, 6$$

$$(4) \min_{s.t.} J = f(\vec{x}) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_4^2 + 5x_5^2$$

$$2x_1 + x_2 - 4x_3 + x_4 - x_5 = 0$$

$$5x_1 - 2x_3 + x_4 - x_5 = 0$$

$$x_1 + 2x_2 + x_3 \geq 6$$

$$4x_3 + x_4 - 2x_5 \leq 0$$

$$\text{初始点 (不可行): } x_j^{(0)} = 1 \quad j=1, 2, \dots, 5$$

$$(5) \min_{s.t.} J = f(\vec{x}) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2 + (x_4 - x_5)^2$$

$$x_1 + 2x_2 + 3x_3 - 6 = 0$$

$$x_2 + 2x_3 + 3x_4 - 6 = 0$$

$$x_3 + 2x_4 + 3x_5 - 6 = 0$$

$$\text{初始点 (可行): } \vec{x}^{(0)} = (35, -31, 11, 5, -5)^T$$

$$(6) \min_{s.t.} J = f(\vec{x}) = (x_1 - 1)^2 + (x_1 - x_2)^2 + 2(x_3 - 1)^2 \\ + (x_4 - 1)^4 + (x_5 - 1)^6$$

$$x_1 + x_2^2 + x_3^2 - 2 - 3\sqrt{2} = 0$$

$$x_2 - x_3^2 + x_4 + 2 - 2\sqrt{2} = 0$$

$$x_1x_5 - 2 = 0$$

$$\text{初始点: } \vec{x}^{(0)} = (2, 2, 2, 2, 2)^T$$

6. 求解非线性规划问题

$$\min_{s.t.} J = f(X) = 60 - 10x_1 - 4x_2 + x_1^2 + x_2^2 - x_1x_2$$

$$-x_1 - x_2 + 16 \geq 0$$

$$0 \leq x_1 \leq 9$$

$$0 \leq x_2 \leq 9$$

选取初始复合形的第一个顶点 $X^{(0)} = (1, 2)^T$ 。收敛精度为 10^{-4} 。

7. 精馏塔的设备费可用下式计算

$$C = C_p AN + C_f HAN + C_i + C_d + C_b + C_L + C_x \quad (a)$$

式中 C ——总费用, 元

C_p ——单位面积塔板费用, 元/ m^2

A ——塔的横截面积, m^2 ;

N ——塔板数;

N_{\min} ——最小塔板数;

C_s ——塔体费用, 元/ m^3 ;

H ——板间距, m;

C_f ——进料泵的费用, 元;

C_d ——馏出泵的费用, 元;

C_b ——釜液泵的费用, 元;

C_L ——回流泵的费用, 元;

C_x ——其他设备费用, 元。

假定已知塔的处理量和分离要求, 已知各泵的费用, 塔的结构材料已经确定, 即 C_f 、 C_d 、 C_L 、 C_b 、 C_p 、 C_s 、 C_x 均为已知, 问题是使塔的总费用最小。

过程变量间具有下述两个经验关系:

$$\frac{L}{D} = \left[\frac{1}{1 - (N_{\min}/N)} \right]^a \left(\frac{L}{D} \right)_{\min} \quad (b)$$

$$A = K(L + D)^{\beta} \quad (c)$$

式中 L ——回流量, kg/h;

D ——塔顶馏出量, kg/h。

为简化起见, 可令 $\alpha = \beta = 1$, 则

$$\frac{L}{D} = \left[\frac{1}{1 - (N_{\min}/N)} \right] \left(\frac{L}{D} \right)_{\min} \quad (b')$$

$$A = K(L + D) \quad (c)$$

对某一精馏塔, 各已知参数如下

$$C_p = 2,000 \text{ 元}/m^2 \quad C_x = 50,000 \text{ 元}$$

$$C_s = 2,200 \text{ 元}/m^2 \quad H = 0.6 \text{ m}$$

$$D = 500 \text{ kg/h} \quad C_f = 25,000 \text{ 元}$$

$$C_d = 20,000 \text{ 元} \quad C_b = 14,000 \text{ 元}$$

$$N_{\min} = 5 \quad K = 0.002 \frac{hm^2}{kg}$$

$$\left(\frac{L}{D} \right)_{\min} = 1$$

回流泵的费用为

$$C_L = 30000 + 8L \quad (d)$$

- (1) 试确定该过程的决策变量和状态变量。
- (2) 找出最小总费用和各变量相应的值。

本章符号说明

a	区间右端点、约束方程组系数、 约束上限	A	换热器面积 (m^2)
A	约束方程组系数矩阵	b	区间左端点、约束方程右端常数 项、约束下限
B	约束方程组右端系数向量	c	目标函数系数
\vec{c}	目标函数系数向量	C	点的集合
C_e	设备折旧费用 (元)	C_{opr}	操作费用 (元)
C_p	比热容 ($kJ/kg \cdot ^\circ C$) ⁻¹	C_T	总费用 (元)
E^n	n 维欧氏空间	f	目标函数、自由度
g	不等式约束函数	\vec{g}	不等式约束函数向量
G	物流量 (kg/h)	h	等式约束函数或搜索步长
\vec{h}	等式约束函数向量	I	单位矩阵
J	目标函数	k	矩阵的秩数
K	传热系数 ($W/m^2 \cdot ^\circ C$)、矩阵的 秩	L	拉格朗日函数
m	方程数	n	变量数
N	折旧年限	\vec{P}	搜索方向
Q	体积流量 (m^3/h)、赫森矩阵的 近似矩阵	R	回流比或可行域
t	时间 (s)、温度 ($^\circ C$)	T	温度 ($^\circ C$)
V	容积 (m^3)	\vec{w}	约束方程组的特解
w, W	物流量 (kg/h)	x	决策变量
\vec{x}	决策变量向量	y, Y	因变量
Z	函数		
上 标			
K	迭代序次	T	向量转秩
*	量优 (大或小) 值		
下 标			
i	方程序号	j	变量序号
k	方程序号	l	方程序号
希腊字符			
α	分割系数、延伸因子、搜索步长	β	分割系数或压缩因子
γ	汽化热 (kJ/kg)	δ	厚度 (m)
ϵ	收敛精度	λ	搜索步长、拉格朗日乘子

μ	拉格朗日乘子	$\vec{\lambda}$	拉格朗日乘子向量
$\vec{\mu}$	拉格朗日乘子向量	Δt_m	传热温差
∇	梯度		

参 考 文 献

1. 邓正龙. 化工中的优化方法. 化学工业出版社, 1992
2. 韦鹤平. 最优化技术应用. 同济大学出版社, 1987
3. Beightler, C. S., D. T. Phillips, and D. J. Wilde, Foundations of Optimization, Prentice - Hall, Englewood Cliffs, New Jersey, 1979
4. Gravin, W., W., Introduction to Linear Programming, McGraw - Hill, New York, 1960
5. Hadley, G., Linear Programming, Addison - Wesley, Reading, Massachusetts, 1962
6. Johnson, J., D., and C. Q. Williamson, "In - line Gasoline Blending at Suntide Refinery", IEEE Trans. Industry and General Applications, March/April, 159 ~ 167, 1967
7. Pike, R. W., Optimization for Engineering System, Van Nostrand Reinhold, New York, 1986
8. T. F. Edgar, D. M. Himmelblau, Optimization of Chemical Processes, McGraw - Hill, 1988

第9章 调优操作法

调优操作法是当今世界上广泛采用的以提高工业生产率的一种生产操作方法。它是将生产工艺规程、数学方法、最优化方法以及计算机技术合为一体的先进操作方法。这种方法适应性强,不仅适用于化学化工行业,而且还适用于其他的部门和行业。

9.1 概 述

调优操作的目的是解决设计环境与操作环境的差异引起的化工过程系统性能下降的问题。传统的生产操作方法可称为“固定操作方法”,通常是根据设计的结果和工厂的经验制定一套操作规程(工艺配方),作为企业的生产操作法规,生产操作人员必须按此法规来进行正常的生产操作,而生产管理人员也是按照这个法规来检查生产操作的优劣。

由于种种原因,一些化工企业的生产装置往往不是在最优状况下运行的,使其能源和原材料消耗偏高,综合经济效益低,总结大致有如下几种原因:①原设计裕度偏大,不优化;②生产量或生产比率偏离设计值;③原材料性能发生变化;④设备性能发生变化;⑤生产装置经过改造,改变了优化操作条件;⑥不同生产负荷的最优条件不同;⑦新的生产工艺有待于进一步优化。

因此,根据生产现场的实际条件,建立相应的数学模型,用计算机进行模拟或优化生产条件,找到节能降耗、增加产量的途径,于是便产生了“调优操作法”的概念。以期达到:①提高产量;②降低原材料单耗;③节约能源;④提高产品质量;⑤减少环境污染;⑥提出需要改进的关键部分;⑦提高经济效益。

调优操作法是六十年代美国威斯康星大学的 Box 教授提出的, Evolutionary Operation——逐次逼近操作方法,简称 EVOP 法^[8]。由于当时计算机应用技术尚不普及,Box 所提出的方法只能通过逐次试验的手段来完成。七十年代以后,特别是八十年代,计算机的飞速发展促使了 EVOP 法的实现方式朝着科学化和多样化方向迈进^{[5][1]}。目前常用调优操作方法有:统计调优法、模拟调优法和模式识别调优法三种。

1. 统计调优法

统计调优操作法,是指应用数理统计方法、最优化技术等数学工具寻求实际过程系统操作条件的改进方案,以获得显著经济效益的系统工程方法。该方法对经典 EVOP 法作了两点改进,一是操作条件的数据主要来自生产报表,仅在必要时做少量试验取得一些补充数据;二是引进了一些简单的设计模型,即用数理统计的方法由操作数据建立目标函数,然后做它的等值线图或采用数值优化方法来寻找最佳操作条件。统计调优法具有方法简单、易于掌握、运用范围广、见效较快等优点。其不足之处是:①对现场数据质量要求较高,要求生产数据稳定、可靠;②所得数学关联式不能外推;③通常只能用于单元操作过程优化,不能用于多参数、多变元条件下,过程复杂的系统优化。

2. 模拟调优法

模拟调优法是应用化工原理、数学模拟理论及计算机技术对化工单元或系统建立数学模

型，在模型上对生产操作条件进行分析和优化。模拟调优法与统计调优法比较，它有如下优点：①对数据要求的数量少；②通常不需要或极少需要安排现场试验，生产操作的优化条件可在计算机上模拟得到；③由于对生产过程的描述是机理或半机理性的，因此所得到的模型用来指导生产比较可靠，效益较大，且可适当外推。但是，这种方法比较复杂，需要对生产过程有比较深刻的了解，对工程技术人员的要求较高。

3. 模式识别调优法

所谓模式识别 (Pattern recognition) 又可以译为图形识别，是近十年来计算机技术发展起来的一类信息处理技术。它是根据性质相近者基本分布在一起的原理，将生产操作中好点与差点分别记入同一个超平面或超曲面将两种点分开，则可将 n 维空间分为好区或差区。若能在生产实际中将各操作控制在好区，则可使生产有所改善，达到调优的目的。

此种方法主要优点在于它不仅可以用于数值系数的优化，还可以用于非数值化的函数的优化 (如各种酒、香料评定等)。

以上三种调优技术各有优缺点，表 9.1 中列出了它们的对照表，在实际工作中，采用何种技术，应结合实际工作需要和实际条件灵活选用，或多者兼而有之，综合应用。

表 9.1 三种调优技术的比较^[1]

项 目 \ 调优方法	统计调优	模拟调优	模式识别调优
工作基础	数 据	过程物理、化学本质	数 据
数学模型	统 计	机理或半机理	
数据量	多	少	中
优化速度	较 快	快	中
适用范围	窄	宽	中
工作难度	小	大	中
改进幅度	小	大	中

9.2 统计调优法

9.2.1 统计调优策略^{[2][3]}

统计调优操作法的技术路线如图 9.1 所示。确定调优目标实际上就是确定优化问题性能指标的内容；因素分析是确定优化问题的决策变量；建立统计调优模型是用数理统计的方法确定系统性能指标与决策变量间的函数关系，建立目标函数和约束方程的统计模型。显而易见，这三部分是统计调优操作的关键技术。

1. 调优目标的确定

确定化工过程系统操作的调优目标，往往是在经营目标的指导下完成的。从经营角度看，经营总决策对系统的运行要求达到一些指标，例如产量、质量、能耗、环境要求等等，都有一些规定的目标。在调优模型中，可选择一个或几个主要指标作为调优目标 (多个指标可以通过加权合并成一个综合调优目标)，其余的指标处理为约束条件。

在确定调优目标之前，还需要划分调优对象系统的范围，通过过程系统分析，选择对经营目标有重大影响的子系统或单元操作过程作为调优对象，往往可以收到事半功倍的效果。

例如,选择过程系统的“瓶颈”(即卡脖子的单元)进行调优操作,常常可以提高系统的运行效益。

2. 调优因素分析

调优因素是指操作变量中选作调优变量的部分。

调优因素分析,就是根据对象系统的特性分析影响系统性能指标的各种可控因素,从中选择一些关键因素作为调优因素的过程。

调优因素分析包括两方面的工作。

(1) 调优变量的选择

调优变量的选取应遵循以下原则:

①按照过程系统的物理化学原理对系统进行详尽的分析,确定出对调优目标的诸影响因素。

②调优变量一定要可调,并对调优目标的影响足够大。系统中那些难以控制的变量,或影响不够灵敏的变量,都不宜选为调优变量。

③调优变量不宜过多,理论指出调优变量以不超过五、六个为宜。这是因为:一方面,调优变量将作为优化的决策变量,过多则会降低用统计方法得出的调优模型的精度,而且实施时优化操作中各个变量难以同时调整到最佳值;另一方面,过多的调优变量可能使结果数据淹没在误差之中,达不到精度要求,失去实际意义。

因素分析时要抓住主要矛盾,那些可忽略的因素和难以控制的因素应断然舍弃。

(2) 原始数据收集、整理和筛选(数据处理)

在确定调优因素以后,收集真实而准确的数据是建立统计调优模型的关键。收集生产数据有两种方法,第一种直接采集生产操作数据,即以操作记录 and 产品质量分析记录为依据,分组收集数据。用这种方法的先决条件是原始生产数据要准确,并且有一定的变动范围,否则得到的规律性不反映实际情况。为了保证搜集到的生产数据准确可信,在搜集数据时必须进行装置的物料平衡,以此来检验装置有没有泄漏,仪表和分析结果是否可靠等。另一种方法是在逐步回归生产数据的基础上,筛选变量,然后在生产装置上进行试验。为了减少试验次数,经常采用正交试验法安排试验。然后根据试验结果进行多元非线性回归。由于试验所得的数据比较准确并且有一定的变动范围。所以回归方程更接近实际情况。但是,变量的数目不能太多,否则无法组织试验。这种方法一般多用于某些中小型化工生产装置,特别是间歇生产过程。具体采用那种方法则要根据具体情况来决定。

生产数据的准确性还取决于仪表的精度、读取数据的准确性和数据的有效位数是否足够等等。但是,获取的数据总是带有误差的。应从误差理论分析生产数据的误差。

我们把生产数据中含有过失误差和测量误差的数据称为可疑数据。这种数据在建立模型

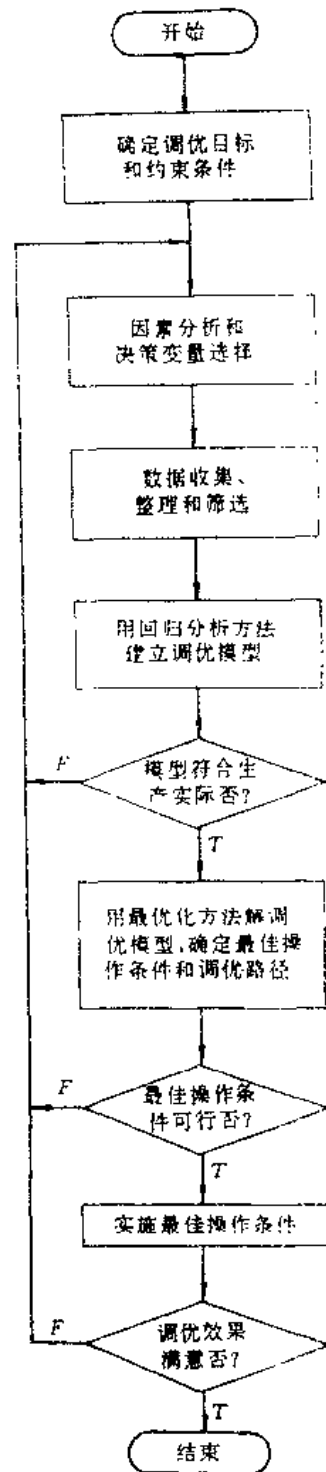


图 9.1 统计调优技术路线

时必须予以剔除,剔除方法有以下两种:

①直接观察法

筛选数据第一步总是要把那些明显不合理和误差大的数据通过观察直接去掉,这种剔除可疑数据的方法称为直接观察法。例如,记录有误差的数据;生产工艺条件不稳定、开停车、有故障时的数据;产品不合格、组成分析不统一的数据;物料不平衡的数据等。这些不合理的数据是明显的、易判断的,仅仅是可疑数据的一小部分。大部分可疑数据的剔除主要还得利用统计判别法在回归计算时完成。

②统计判别法^[6]

线性回归计算可靠的前提是每个原始数据的观测值 y_i 与回归计算值 \hat{y}_i 之间的残差 $\delta_i = y_i - \hat{y}_i$ 均在随机误差允许的范围之内。亦即,残差中不应含有测量误差和过失误差的成分。根据这一原则,如果某一数据的残差超过随机误差的允许范围,那么就认定该数据为可疑数据,予以剔除。本书第二篇介绍的 Chauvenet 准则是认定可疑数据有效的判据。使用这个准则时,可根据回归结果,对全部实验值进行逐组检查,把属于可疑数据的实验值选出。若发现不止一个可疑数据,则应把其中残差绝对值最大者剔除。然后重新计算标准差值。根据新的值,再次用 Chauvenet 准则进行检查。每次只剔除一个可疑数据,其余数据重新进行回归,直至回归所用的数据中不再含有可疑数据为止。

3. 统计调优模型

对于化工生产装置调优操作的数学模型,目标函数通常选用经营指标,其一般形式表示如下:

$$\text{经济效益 } J = \text{产值} - \text{成本}$$

$$\text{产值} = \sum_i p_i P_i$$

$$\text{成本} = \sum_j r_j R_j + \sum_m \sum_n u_m U_{mn} + \alpha I + \sum_k m_k M_k$$

式中

J ——生产装置的经济效益(利润和税收),元/年。

p_i ——第 i 种化工产品的单价,元/t。

P_i ——第 i 种化工产品的产量,t/年。

r_j ——第 j 种原料的单价,元/t。

R_j ——第 j 种原料的消耗量,t/年。

u_m ——第 m 种公用工程的单价,元/t(公用工程包括水、电、气等)。

U_{mn} ——第 n 种设备对第 m 种公用工程的消耗量,t/年。

α ——装置的折旧率。

I ——装置的固定投资,元。

m_k ——第 k 类生产人员的工资,元。

M_k ——生产雇用第 k 类人员数量。

对一个正在生产的装置,装置折旧费用和各类生产人员的工资在一定时期内一般不会变动,那么生产过程优化的目标是提高目的产物的产量和质量,降低原料和能源消耗。当二者其中之一不为主要矛盾时,调优目标可简化为:

①产值最大, $\max(\sum_i p_i P_i)$

或

②原料和能源消耗最少, $\min(\sum_j r_j R_j + \sum_m \sum_n u_m U_{mn})$

以上目标函数中的各生产量(如 P_i , R_j , U_{mn} 等)可通过回归分析方法与调优因素进行关联。

建立统计调优模型的理论和方法是成熟的,但实施起来并非驾轻就熟之事。在建模过程中比较困难的是确定模型的数学形式。到目前为止,还没有简单、可靠的办法能快速地确定化工过程统计优化模型的数学形式,即使在计算机上用一些现成的软件也是要花费不少的时间,有时还得凭经验试估模型的形式,采用边算边修改的方法,可用逐步回归建立统计调优的目标函数。

统计调优模型建立以后,必须进行检验。检验工作可按以下三步进行。

(1) 检验模型是否符合生产实际的变化趋势。可以改变模型中的决策变量,考察其对状态变量和目标函数的影响是否符合生产原理。

(2) 检验模型的相关程度、回归系数的置信度和残差,并对残差进行分析。

常见几种情况的残差($\delta = y - \hat{y}$)图如下:

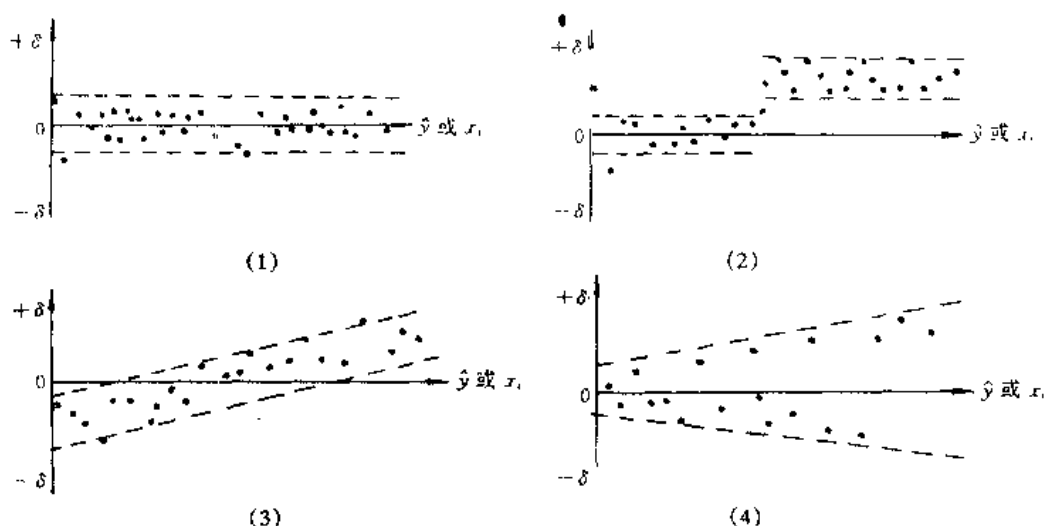


图 9.2 残差图

图 9.2 (1) 为正常残差图,图 9.2 (2) ~ (4) 为异常残差图。

(3) 用模型预测实际生产装置生产情况。

若条件允许,应用模型预测正在运行的生产装置的生产结果,以此最后判断模型的准确性。在进行预测过程中允许有 10% 左右的误差。

4. 调优搜索

应用最优化方法求解调优模型的最优解时,需要考虑调优操作过程的一些特点。首先,调优操作是一个逐渐改进的过程,其中每一步追求的是有所改进且不扰乱生产。因此,对于操作改进数据的优惠解,感兴趣的不仅是最优点,更重要的是通往最优点的路径,路径上的点可为调优过程提供更多的信息,便于以调节策略实施优惠解。这种路径称之为调优路径。其次,调优对操作条件的改变应是微小的,因此求解总是在一些现行操作点附近进行的,即在给定初始点的领域内寻求优惠解。调优搜索的上述两点特性决定了调优路径上的点列应具有以下两点性质:

(1) 它是通向最优点的,一般后续点总比先行点好;

(2) 相近两点相距不远。

以上我们概括地介绍了统计调优的基本策略和实施步骤。实际生产过程中各种因素的关系十分复杂，外界的干扰十分庞杂。因此，要想借助统计调优法寻求优惠操作条件，必须具有丰富的实际经验，这一点是不容忽视的。

9.2.2 统计调优实例

以国内某厂甲苯歧化装置的操作参数调优^[3]为例，进一步阐明统计调优操作法的应用技术。

【装置概况和调优背景】

国内某厂甲苯歧化装置是七十年代末期从国外引进的，原料处理能力为 21 万吨/年，生产方法采用的是二甲苯增产法 (Xylene-Plus)，原料为甲苯和 C_9 芳烃，产品为苯和二甲苯。装置于 1980 年投产。开工试车时催化剂活性，甲苯转化率和选择性均偏低，达不到设计要求。几年来工厂做了许多改进工作，到 1984 年上半年，二甲苯产量仍满足不了聚酯生产的需求，不得不外购混合二甲苯 2000 吨，至使装置经济效益下降。为此，有必要对装置进行调优分析，力图用统计调优法确定装置的最佳操作条件，以提高装置经济效益。

1. 确定调优对象

从了解生产工艺入手，分析装置生产状况，找出问题所在，确定调优对象。

甲苯歧化生产工艺如图 9.3 所示。甲苯和 C_9 芳烃混合后（混合比 11:1）泵入换热器与反应后物料换热到 315°C ，经加热炉汽化加热到 490°C 进入反应器上部。

反应器为移动床式，内装 33t 稀土 Y 型 Si—Al 分子筛小球催化剂，催化剂经烧焦再生后返回反应器，循环量 40t/h，进入反应器时的温度为 515°C ，反应器内压力为常压。汽化的物料 (490°C) 进入反应器与热催化剂 (515°C) 相遇发生歧化和烷基转移反应，主产物为苯、二甲苯，副产物为 C_1 — C_2 烃化物、稠环芳烃等。

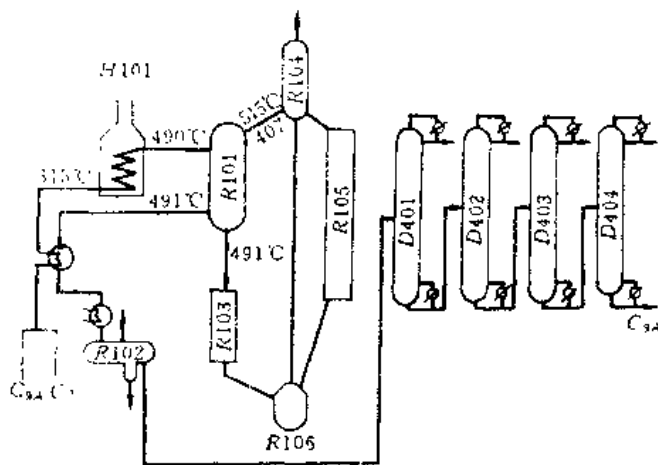


图 9.3 甲苯歧化装置工艺流程

反应产物从反应器底部出来 (491°C) 与原料换热后进入气液分离罐，从此罐出来后进入稳定塔，接着进入三个精馏塔分出苯、甲苯、二甲苯和 C_9 芳烃。

结焦催化剂在反应器底部经汽提后进入再生器烧焦，烧焦后催化剂进入提升罐，在提升罐内经空气提升到分离罐后进入反应器。

从工艺设计资料和生产情况对比来看，整个装置的关键设备是甲苯歧化反应器，而问题又恰出于此，即歧化反应催化剂活性、选择性和转化率偏低，工

艺参数变化幅度较大，至使二甲苯产量下降，不能满足聚酯生产要求，每年需外购数百吨二甲苯。可以看出，影响二甲苯产量下降的主要原因是歧化反应器未按最佳操作条件操作所致。因此，抓住主要矛盾，选定甲苯歧化反应器为调优关键设备，作为调优对象。

2. 确定调优目标和调优变量

根据装置生产状况分析和经济分析, 采取以利税最高为调优目标, 于是目标函数为:

$$J = \sum p_i P_i - \sum r_j R_j - \sum \sum u_m U_{mn} \quad (9.1)$$

对甲苯歧化装置, 上式中各项分别为:

产品有:

P_1 ——二甲苯产量	p_1 ——700 元/t
P_2 ——苯产量	p_2 ——650 ~ 670 元/t

原料消耗主要有:

R_1 ——甲苯用量	r_1 ——800 元/t
R_2 ——芳烃用量	r_2 ——700 元/t
R_3 ——燃油用量	r_3 ——137 元/t

(原料和产品的单价为 1984 年的价格。)

能耗主要有:

$U_{m,1}$ ——反应器
$U_{m,2}$ ——加热炉
$U_{m,3}$ ——精馏塔

m ——指水、电、汽、风

此式包含的因素很多, 我们应从诸多因素中抓住主要矛盾, 略去一些次要因素。根据这个原则, 结合生产实际, 我们可进行如下简化。

第一步简化:

假设 $\sum r_j R_j$, $\sum \sum u_m U_{mn}$, 两项变化很小, 对经济效益的影响不大, 可以略去。

这样我们只需求取 $\sum p_i P_i$ 项最大就能达到 J 最大的目的。

第二步简化:

尽管二甲苯产量的增加会导致苯产量下降, 但考虑到企业对苯、二甲苯的需求对比和装置的具体情况, 只求二甲苯产量最大即可提高经济效益。这样我们确定的调优目标最终简化为求二甲苯产量最大, 即

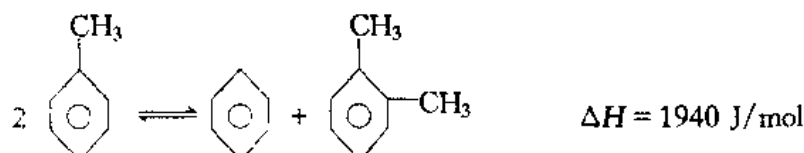
$$\max J = P_1 \quad (9.2)$$

这个调优目标是在特定时期和特定环境下产生的, 至于目标是否合理, 要经调优实施前的经济评价加以验证。

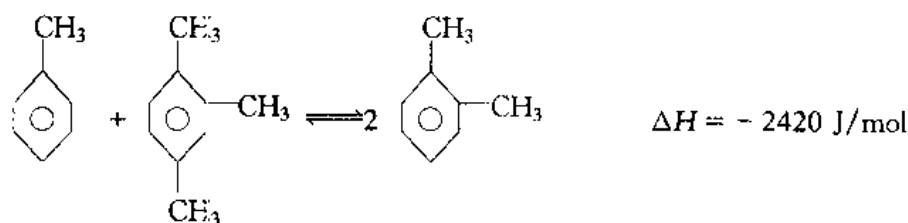
以下围绕以二甲苯产量最大进行因素分析, 分析影响调优目标的主要因素。

歧化反应器内主要进行以下化学反应:

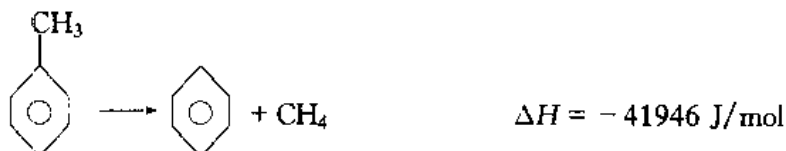
(1) 甲苯歧化



(2) 甲苯和三甲苯烷基转移



(3) 甲苯脱烷基



根据装置的原始实验资料分析影响反应过程的因素。

从装置的设计资料查到甲苯歧化反应条件影响产品收率的关系曲线^[4]，见图 9.4 至 9.6。分析这几条曲线后，我们得出如下四点结论。

(1) 反应温度

众所周知，反应温度是化学反应的最积极因素，改变反应温度将影响歧化反应，烷基转移反应，裂解生成气体的反应以及缩合结焦反应等等。甲苯歧化反应是一个转化率低，需要大量循环的工艺。因此，提高生成目的产品的选择性是降低损失提高产品收率的关键。

在较低的温度范围内，提高温度可加速甲苯歧化和烷基转移反应过程，可提高甲苯和 C₉ 芳烃的转化率，以及苯、二甲苯的单程收率。

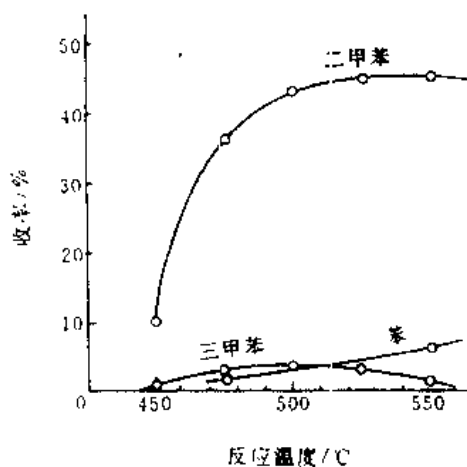


图 9.4 反应温度的影响

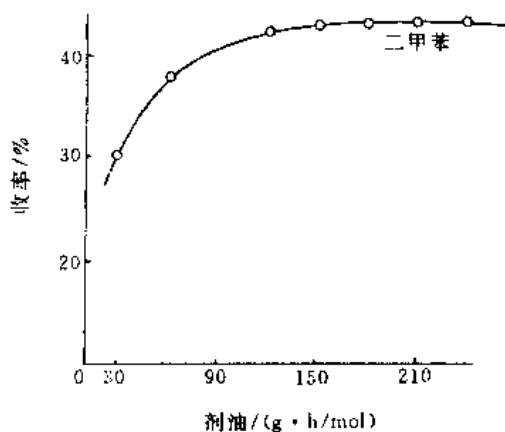


图 9.5 空速的影响

但是，在较高的温度范围内，提高温度将有利于歧化反应而不利于烷基转移反应的进行，二甲苯的单程收率将下降，苯的单程收率将继续增加。

如果温度过高，副反应加剧，催化剂表面结炭量也进一步增加。不但选择性降低，而且催化剂在参与反应过程中的平均活性也会下降，因而反应温度不宜过高。

(2) 压力

由于反应是在常压下进行的，压力变化很小，因此压力因素对反应产物分布影响很小。

(3) 空速

空速是指单位床层的催化剂（重量）在单位时间内所通过的物料量。

本反应器设计床层催化剂装填量为 33t, 进料量为 21.585t/h, 则重量空速为 0.654 h^{-1} 。

在实际操作中用改变进料量来调节空速, 在其他条件不变的情况下, 提高空速, 接触时间减少, 反应深度降低, 转化率和收率都将降低, 选择性将有所提高。

降低空速, 物料与催化剂的接触时间增加, 反应深度增大, 转化率提高, 副反应增加, 选择性降低。

一般地要尽量避免在低空速下进行。如果一旦处于低空速下运转, 就要适当降低催化剂循环来降低反应速度, 以提高选择性, 减少物料损失。

(4) 原料组成

进料中若没有 C_9 芳烃存在, 只有甲苯歧化反应进行, 而烷基转移反应不能进行。

本装置所用 C_9 芳烃来自重整装置。

C_9 芳烃存在, 特别是三甲苯存在, 甲苯和三甲苯之间的烷基转移反应便可发生。提高进料中 C_9 芳烃含量, 可使烷基转移反应的速度加快, 二甲苯生成量增加。

乙基甲苯, 丙基苯的存在使甲苯歧化反应受到抑制, 影响苯、二甲苯收率, 所以要控制重整油的馏分。

茚满存在会使催化剂表面积灰量增加。使催化剂活性下降。

从以上分析可以看出, 反应温度、空速和进料 C_9 芳烃含量对产品苯、二甲苯含量有显著影响, 而且在装置上有调节手段。因此, 选择这三个影响因素作为调优变量, 即①反应温度; ②空速; ③进料 C_9 芳烃含量。

3. 数据的收集与筛选

从甲苯歧化装置上取到 1984 年上半年的生产数据, 并输入计算机。首先对这些数据作装置的物料平衡计算。结果表明, 1984 年上半年装置的总物料平衡率高达 99.1%, 损失和误差为 0.9%。这样高的物料平衡表明装置没有泄漏, 流量测量和分析结果是可信的, 收集的数据是可信的。

在收集到的工艺数据中, 若有异常数据, 则它们必然会歪曲装置的运行情况, 我们应仔细检查它们出现的原因并判断是否应当剔除, 剔除方法采用前以述及的两步法进行。

经剔除可疑数据, 筛选出 166 天的数据用于建立统计调优模型。现将收集到的原始数据和计算数据用符号说明, 见表 9.2、9.3。

4. 模型建立

建立模型就是利用收集到的工艺数据, 依据数理统计原理, 用逐步回归分析, 确定调优目标和调优变量之间的函数关系。

下面推导甲苯歧化一次通过时苯和二甲苯产量与操作条件的关联式。

二甲苯和苯的产量等于产品油量和其中二甲苯和苯含量的乘积。产品油量则等于进油量乘以液收率。现把液收率以及二甲苯、苯含量与操作条件的关系分别回归, 然后合成整个数学模型。

在回归之前, 先对变量进行同数量级化处理, 以免造成计算误差。即

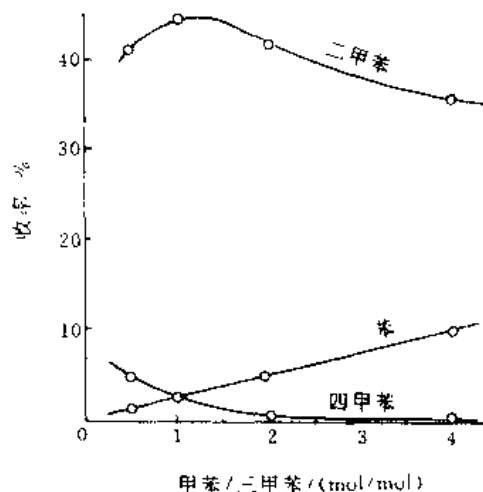


图 9.6 原料组成对烷基转移反应的影响

表 9.2 原始数据符号对照表

符 号	符 号 含 义	符 号	符 号 含 义
DA	日 期	X_{C_9A}	进料 C_9 芳烃含量, %
F_1	C_9 芳烃进料量仪表指示值	F_A	进料量, kg/h
F_2	甲苯进料量仪表指示值	T_1	进料温度, $^{\circ}C$
F_3	产油量仪表指示值	K	催化剂循环量, t/h
T_R	反应温度, $^{\circ}C$	F_B	进料二甲苯量, t/h
Y_{NA}	产品非芳烃组成, %	Y_B	产品苯组成, %
Y_T	产品甲苯组成, %	Y_v	产品二甲苯组成, %
Y_{C_9A}	产品 C_9 芳烃组成, %		

表 9.3 计算数据符号对照表

符 号	符 号 含 义	符 号	符 号 含 义
DA	日 期	T_R	反应温度, $^{\circ}C$
X_{C_9A}	进料 C_9 芳烃含量, %	S_v	空速, 1/h
R	剂油比	F_A	甲苯进料量, kg/h
F_{C_9A}	C_9 芳烃进料量, kg/h	F	总进料量, kg/h
W	产品量, kg/h	Y_W	重量液收率, %
Y_M	分子液收率, %	C_T	甲苯转化率, %
C_{C_9A}	C_9 芳烃转化率, %	S	选择性, %
P_B	苯产量, kg/h	P_T	甲苯产量, kg/h
P_X	二甲苯产量, kg/h	P_{C_9A}	C_9 芳烃产量, kg/h
DI	歧化率, %		

$$x_1 = \frac{T_R - 491}{20}$$

$$x_2 = S_v - 0.66$$

$$x_3 = X_{C_9A}$$

先回归产品油中二甲苯、苯含量与操作条件的关联式。

通过对 1984 年上半年甲苯歧化生产装置上 166 天的数据在计算机上进行回归, 得出如下回归方程:

$$y_X = 0.13024 + 0.01389x_1^2 - 1.13685x_2^2 + 0.0051776x_3 \quad (9.3)$$

相关系数 $r = 0.7134$

$$y_B = 0.09395 + 0.04066x_1^2 + 0.14998x_1x_2 - 0.0030846x_1x_3 - 0.01418x_2x_3 \quad (9.4)$$

相关系数 $r = 0.7676$

从式 (9.3) 看出, 产品油二甲苯含量随着反应温度的升高而增加, 随着空速的提高而降低, 随着进料 C_9 芳烃含量的提高而增加; 从式 (9.4) 看出, 产品油苯含量随着反应温度升高而增加, 随着空速和进料 C_9 芳烃含量的增加而降低。这些规律与理论推断是完全一致的。

从式 (9.3) 及式 (9.4) 可算出歧化率, 即二甲苯产量/苯产量。歧化率随着进料 C_9 芳烃含量的提高而增加, 反应温度上升, 歧化率略有降低, 但与进料空速无关。

从操作条件变化可以看出, 随着反应温度的上升, 液收率略有降低。现将 1984 年上半年 180 天的数据用计算机回归得到

$$Y_W = 0.98279 - 0.035503x_1 \quad (9.5)$$

把产品油的苯、二甲苯含量和液收率的关联式合在一起,便得到一次通过时苯和二甲苯产量的数学模型如下:

$$\text{二甲苯产量:} \quad P_X = F \cdot Y_W \cdot y_X \quad (9.6)$$

$$\text{苯产量:} \quad P_B = F \cdot Y_W \cdot Y_B \quad (9.7)$$

式中

$$F = 33000(x_2 + 0.66) \quad (9.8)$$

Y_W 、 y_X 和 Y_B 分别决定于 (9.5)、(9.3) 和 (9.4) 式。

影响一次通过二甲苯和苯产量的变量有反应温度、空速和进料 C_9 芳烃含量,这些变量不是可以任意改变的,要受到一定的约束。这些变量最大可变动的范围要受到化学反应,化工操作以及设备等的制约,目前还不太清楚,需要进一步由理论和生产实践来决定。但目前计算最优化暂时把约束条件取到 1984 年上半年运转数据边缘上。这样约束条件为:

$$T_R \leq 511^\circ\text{C}, S_V \leq 0.76\text{h}^{-1}, X_{C_9A} \leq 16\%$$

即

$$20x_1 + 491 \leq 511 \quad (9.9)$$

$$x_2 + 0.66 \leq 0.76 \quad (9.10)$$

$$x_3 \leq 16 \quad (9.11)$$

经过上述推导,整理式 (9.3) ~ (9.11), 得甲苯歧化反应装置统计调优模型:

$$\begin{aligned} \max_{s.t.} P_X &= 33000(x_2 + 0.66)(0.98279 - 0.035503x_1) \\ &\quad (0.13024 + 0.01389x_1^2 - 1.13685x_2^2 + 0.0051776x_3) \end{aligned} \quad (9.12)$$

$$g_1(\vec{x}) = 1.0 - x_1 \geq 0 \quad (9.13)$$

$$g_2(\vec{x}) = 0.1 - x_2 \geq 0 \quad (9.14)$$

$$g_3(\vec{x}) = 16 - x_3 \geq 0 \quad (9.15)$$

5. 模型求解与讨论

(1) 优化计算

对式 (9.12) ~ (9.15) 所示的优化模型,采用最优化方法求解。

决策变量初值: $x_1^0 = 0.5$ $x_2^* = 0.04$ $x_3^* = 15$

计算精度: $\varepsilon = 0.001$

计算结果:

$$x_1^* = 0.134 \quad x_2^* = 0.110 \quad x_3^* = 14.95$$

转换成温度、空速和 C_9 芳烃含量后为:

$$T_R = 493^\circ\text{C}$$

$$S_V = 0.77\text{hr}^{-1}$$

$$X_{C_9A} = 15\%$$

P_X 的最大值为 4828kg/h

(2) 结果讨论

①考察优化结果与工艺情况是否相符。

表 9.4 优化计算结果与设计值对比

操 作 变 量	设 计 值	优 化 值	变 化 幅 度
T_R	491℃	493℃	增加 0.41%
S_V	$0.65h^{-1}$	$0.77h^{-1}$	增加 17.7%
X_{C_9A}	9.1%	15%	增加 64.8%

比较表 9.4 中数据可以看到, 优化结果是反应温度上升, 空速提高, 进料中 C_9 芳烃含量增加, 二甲苯产量增大, 这与生产上的变化情况相符。

观察优化计算结果中各变量增大的幅度后可看到: 在三个操作变量中温度增加幅度不大, 它只比设计值高 2℃, 仅增加 0.41%。

在甲苯歧化工艺过程的讨论中, 我们分析影响反应因素时指出: 在较低温度范围内, 增加温度可提高二甲苯产量。优化结果与上述分析结论相符。

通过上述讨论我们可看到优化计算结果是可信的。

②利用优化计算结果检验建模时的进行的简化是否合理。

在优化计算结果的基础上, 分别计算出 $\sum r_j R_j$, $\sum \sum u_m U_m$ 二项的数值, 检验它们是否变化小, 是否可以忽略。

6. 经济评价和调优实施

(1) 对优化计算结果进行经济评价

以优化计算结果为依据作出全车间的经济核算表, 预测实施调优后可获得的经济效益。根据经济效益的好坏、大小决定是否实施调优。

(2) 实施调优操作

当认定优化计算结果可行时, 我们可到工业装置上进行调优实施。

实施过程中注意, 在调整工艺参数时应以小幅变化向预定的优化条件调整。每调整一小步, 观察调优目标的变化情况以确定下一步的调整方向和幅度。

甲苯歧化装置调优实施情况:

为了验证数学模型和最优化计算结果, 并取得实际的经济效益。先后在该装置上进行两次最优化的工业试验。

第一次调整工艺参数, 于 1984 年 10 月 11 至 13 日三天。

控制工艺参数为: $T_R = 505^\circ\text{C}$, $S_V = 0.7h^{-1}$, $X_{C_9A} = 16\%$

装置运行结果见表 9.5。

甲苯歧化装置所在企业每年仅缺数百吨二甲苯, 若维持第一次调整的工艺参数生产, 二甲苯每小时增产 1039kg, 苯每小时减产 572kg, 二甲苯增产过多, 苯产量损失较大。应对工艺参数重新进行调整, 降低进料中 C_9 芳烃含量, 减少烷基转移反应的速度能使二甲苯产量下降。为此进行第二次工艺参数调整。

第二次调整工艺参数, 1984 年 10 月 14 至 16 日。

控制工艺参数为: $T_R = 505^\circ\text{C}$, $S_V = 0.7h^{-1}$, $X_{C_9A} = 14.6\%$

装置运行结果见表 9.5。

每月增产: $(4570 - 4080) \times 24 \times 30 = 352.8t$ 二甲苯。

表 9.5 调优实施结果与设计值、计算值对比

	第一次试验		第二次试验	
	二甲苯产量	苯产量	二甲苯产量	苯产量
试验值, kg/h	5119	1624	4570	1902
设计值, kg/h	4080	2196	4080	2196
计算值, kg/h	4945	1698	4679	1820
调优增长率, %	-25.5	-26	+12.0	-13.4
计算误差, %	+3.52	-4.36	-2.33	+4.51

维持此工艺条件生产一个月可基本满足该化工厂对二甲苯的需求量。

此外, 把试验得到的二甲苯和苯产量与数学模型 (9.12) 计算的结果进行比较, 试验值与计算值的误差全在 5% 以内 (表 9.5), 这表明数学模型能表征实际的运行结果, 这也进一步验证了数学模型。

调优工作结束。

9.3 模拟调优法概要

模拟调优法是应用化学工程基本理论、数学模型理论及计算机技术对化工单元或系统建立数学模型, 在模型上对生产操作条件进行分析和优化。模拟调优法与统计调优法的主要区别在于优化模型的建立方法不同, 所建立的模型是纯机理模型, 它的约束方程一般是通过分析过程的物理化学本质和机理, 利用化学工程学的基本理论 (如质量守恒, 化工热力学, 化学反应动力学等基本规律) 建立一套描述过程特性的数学模型及边界条件。

模拟调优法是过程系统稳态模拟和最优化方法的有机结合。在优化模型中把过程系统模型作为等式约束条件。因此, 它们一般都具有过程系统模型的大型稀疏性特点, 在求解算法上也要采用一些特殊技术。本节仅概括性地阐明模拟调优的基本原理, 对于涉及的稳态模拟优化方法, 由于本书篇幅有限, 不做详述。此部分内容如读者感兴趣, 请参考有关化工过程模拟或化工系统工程等方面的书籍。

9.3.1 稳态模拟优化模型

稳态又称做定常态, 稳态优化模型中的过程变量、目标函数和约束条件均不随时间的变化而改变。过程示意图如图 9.7 所示, 稳态集中参数优化模型由代数方程组成; 是目前使用最广泛的模型, 可以表述为下列普遍化的数学形式:

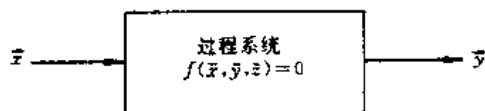


图 9.7 过程示意图

$$\begin{aligned}
 & \min_{x, y} F(\vec{x}, \vec{y}) = 0 & (9.11) \\
 & \text{s.t.} \\
 & f(\vec{x}, \vec{y}, \vec{z}) = 0 & (\text{流程描述方程}) \\
 & c(\vec{x}, \vec{y}, \vec{z}) = 0 & (\text{尺寸, 成本方程}) \\
 & h(\vec{x}, \vec{y}) = 0 & (\text{等式约束}) \\
 & g(\vec{x}, \vec{y}) \geq 0 & (\text{不等式约束})
 \end{aligned}$$

式中 \vec{x} ——调优变量 (决策变量);
 \vec{y} ——响应变量 (状态变量);
 \vec{z} ——单元内部变量。

9.3.2 模型求解方法

稳态模拟优化问题的求解方法, 可以看作是最优化方法与稳态模拟方法的结合。通常, 可以利用稳态模拟方法求解优化模型的等式约束方程, 利用最优化方法寻求满足约束的目标函数最优解。

化工过程的稳态模拟方法有以下三种: (1) 序贯模块法; (2) 面向方程法; (3) 联立模块法。

以上三种模拟方法与最优化方法的结合以及计算策略的不同就产生了不同类型的模拟优化策略。目前对于化工过程的模拟优化法, 特别是大系统的模拟优化, 国内外有许多专著^{[2][7]}发表, 一些问题还正处于研究探索之中。

本章符号说明

c	尺寸、成本函数	R	原料消耗量 (T/y)
f	流程描述函数	u	公用工程单价 (元/ T)
F	目标函数	U	公用工程消耗量 (T/y)
g^*	不等式约束函数	\vec{x}	调优变量或决策变量向量
h	等式约束函数	y	实验观测值
I	装置的固定投资 (元)	\hat{y}	回归计算值
J	目标函数	\vec{y}	响应变量或状态变量向量
m	生产人员的工资 (元)	\vec{z}	单元内部变量向量
M	生产人员的数量	希腊字符	
p	产品单价 (元/ T)	α	装置折旧率
P	产品产量 (T/y)	δ	残差
r	原料单价 (元/ T)、相关系数	ΔH	反应热 (J/mol)

参 考 文 献

- 1 刘鲜京, 姜家轩, 张克诚, 计算机应用研究, 1994, 2 (51)
- 2 杨冀宏, 麻德贤, 过程系统工程导论, 烃加工出版社, 1989
- 3 韦士平, 现有化工装置生产调优操作, 内部出版资料, 1987
- 4 金海林, 程宏远, 最优化方法及其在化工中应用, 内部出版资料, 1990
- 5 目崎令司, 现有工厂的最佳化方法, 化学技术志(日文), 1980:18(3)
- 6 中国科学院数学研究所数理统计组编, 回归分析方法, 科学出版社, 1974
- 7 T. F. Edgar, D. M. Himmelblau, Optimization of Chemical Processes, McGraw - Hill, 1988
- 8 Box, G. E. P, et al. Evolutionary Operation, John Wiley & Sons Inc., N. Y. 1966

第四篇 配套软件介绍

第 10 章 化工过程计算软件包

化工过程计算软件包是与本书配套的自编软件包（与第一册书一起发行），读者通过该软件包的使用及对计算结果的分析，既可解决许多工程实际问题，掌握分析和解决实际问题的基本方法，又有助于掌握大型软件的使用，提高利用计算机处理实际问题的能力。本篇简要介绍化工过程计算软件包（以 BIPTCHEM 表示，BIPT 指北京石油化工学院，CHEM 指化工过程模拟计算）的功能及使用。

化工过程计算软件包（BIPTCHEM）能进行物性计算，数据处理，优化计算，精馏的计算，换热器计算，板式塔的计算和化学反应器的计算。物性计算中包括一个 624 种物质的基础数据库（可以查询、修改和调用），纯组分性质的计算，混合物性质的计算，石油馏分性质的计算；数据处理中包含逐步回归和非线性回归；优化计算包括线性规划，单纯形法和拉格朗日乘子法的非线性规划，复合形法非线性规划；精馏计算包括泡、露点计算和闪蒸计算，多组分精馏的简捷计算，多组分精馏的严格计算；换热器计算包括卧式冷凝器计算，卧式再沸器计算，立式再沸器计算，无相变管壳式换热器的计算；板式塔计算有浮阀塔和筛板塔，反应器的计算包括常见反应器的计算。

BIPTCHEM 软件包由应金良主编，臧福录为顾问。其中数据处理和优化部分由工程师班 BASIC 源程序改编而来，反应器部分由应金良和盖旭东共同开发。

10.1 软件包的安装、运行及主菜单的打开

10.1.1 安装

本软件在 DOS 下安装和运行，大约需占 3M 的硬盘空间。安装过程如下：

- (1) 在 DOS 下启动机器；
- (2) 将软盘插入驱动器 A；
- (3) 键入 "A:"，并按〈Enter〉键；
- (4) 键入 "install"，并按〈Enter〉键；

根据屏幕提示，键入要安装的硬盘盘符（如 C，或 D）后，即自动进行安装。安装完后退出安装程序。

10.1.2 运行

化工过程计算软件包必须有中文平台的支持（如 UC DOS），使用时应先启动 UC DOS 系

统，然后进入该软件所在的子目录（如 d: \ BIPTCHEM 目录），在该子目录下键入“BIPTCHEM”，并按〈Enter〉键，BIPTCHEM 开始运行，屏幕显示如下所示：

化工过程计算

应金良

北京石油化工学院化工教研室

10.1.3 主菜单的打开

过 20 秒（或按任意键）后，屏幕显示如下主菜单：

物性	数据处理	优化	蒸馏	换热器	板式塔	反应器

用方向键（→键或←键）选择主菜单（若某菜单的颜色为红色，则该菜单被选），选好菜单后，按〈Enter〉键，该菜单即被打开。根据用户的需要，在主菜单打开之后，即可根据菜单选择和运行与菜单相关的软件。以下介绍软件包中各软件的使用。

10.2 物性计算

选择“物性”菜单，并按〈Enter〉键之后，屏幕显示的菜单为：

基础数据库
估算纯物质的性质
估算混合物的性质
石油馏分性质的估算
退出

用户用方向键 (↑或↓) 进行选择, 按 〈Enter〉 键执行。若选择“退出”, 按 〈Enter〉 键后退出 BIPTCHEM, 回到 DOS 状态。

10.2.1 基础数据库

选择基础数据库, 并按 〈Enter〉 键之后, 屏幕给出菜单:

查询基础数据
改变数据库中的基础数据
在数据库中增加一种物质的基础数据
查询或改变 Wilson 二元交互作用能量参数
退出

1. 查询基础数据

若选择“查询基础数据”, 根据屏幕提示, 回答是否打印后, 屏幕显示:

数据库中有 624 种物质
输入组分的标号或分子式:

组分的标号可从表 10.1 中查得, 为避免查书的麻烦, 也可输入分子式。若输入了组分的分子式, 且数据库中存在该组分, 则屏幕将显示出它的化学名称 (若有同分异构体存在, 也将同时给出), 并请用户确认。用户一经确认, 屏幕立即显示出该物质的相对分子量、凝固点、沸点、临界温度、临界压力、临界体积、临界压缩因子、偏心因子、偶极矩、液相密度、理想气体热容计算式、生成焓、生成 Gibbs 能、汽化热、理想气体焓计算式、理想气体熵、粘度计算式、饱和蒸汽压计算式、液体导热系数计算式、SRK 方程偏心因子、298K 时的溶解度参数及液相摩尔体积、UNIFAC 官能团等。

2. 改变数据库中的数据

根据屏幕提示, 用上述同样方式输入要改变的组分的标号或分子式, 屏幕显示以下内容供用户修改:

- (1) 分子式, 名称, 分类号
- (2) 临界参数
- (3) 熔点、沸点、汽化热
- (4) 理想气体热容
- (5) 理想气体的焓和熵
- (6) 理想气体的标准生成热、标准 Gibbs 生成自由能和偶极矩
- (7) 蒸汽压
- (8) 液体粘度和溶解度参数
- (9) 液体的导热系数
- (10) 液体的 PVT 性质
- (11) UNIFAC 基团的划分
- Q 退出

用户可根据需要, 选择相应菜单, 并输入新的数据, 以替代数据库中的原有数据 (若非必要, 用户不要轻易修改数据库中的各项数据, 以免将原本正确的数据改成错误的数, 而影响各项计算)。

3. 增加一种物质的基础数据

数据库中有 624 种物质的基础数据, 如用户需有增加新物质的数据, 则所增物质的序号从 625 开始, 依次往后增加。根据屏幕提示, 从键盘依次输入所增物质的各项基础数据。

4. 查询或改变 Wilson 二元交互作用能量参数

根据屏幕提示输入一对组分的标号或分子式, 屏幕将显示该二元组分间的 Wilson 交互作用能量参数, 并询问是否要改变 (Y/N)? 若要改变, 键入 “Y”, 并输入新值, 以替换原有的值。

10.2.2 估算纯物质的性质

选择估算纯物质的性质后, 根据屏幕提示, 回答是否打印, 然后输入组分的标号或分子式, 并确认后, 根据屏幕提示输入温度 [K] 和压力 [bar] ($1\text{bar}=100\text{kPa}$), 屏幕显示如下菜单:

蒸汽压
汽化热
液体密度
液体粘度
液体导热系数
液体比热容
液体逸度系数

气体密度
气体逸度系数
气体比热
气体焓
气体熵
气体粘度
气体导热系数
水和水蒸气表
液体焓
退出

用方向键（↑或↓）选择所需计算的物性，按〈Enter〉键后即给出计算结果。

在物性的计算中，本软件采用了几种常见的方法或模型，如在蒸汽压的估算中，采用了 Lee-Kesler 方程、Gomez-Thodos 方程、用两参考流体推算、Wagner 方程；在汽化热的估算中，采用了 Pitzer 对比态关联式、两参考流体方法。

10.2.3 估算混合物的性质

选择估算混合物的性质后，根据屏幕提示，回答是否打印后输入系统的组分数、各组分的分子式或标号，并确认。输入各组分的摩尔分率及混合物的温度和压力，屏幕给出以下菜单：

液体粘度
液体导热系数
液体比热容
液体密度
液相中组分的活度系数
液相中组分的逸度系数
气体的粘度
气体的导热系数
气体的比热容
气体的密度
气相中各组分的逸度系数
气相焓
液相焓
退出

用户根据需要，用方向键（↑或↓）选择计算项目。

10.2.4 估算石油馏分的性质

在估算石油馏分的性质时，根据屏幕提示输入温度（℃）和压力后，屏幕给出以下菜单：

给定：

恩氏蒸馏数据和 20 度时的相对密度（或折射率）

中平均沸点和 20 度时的相对密度（或折射率）

50 度时的运动粘度和 20 度时的相对密度（或折射率）

100 度时的运动粘度 [厘斯] 和 20 度时的相对密度（或折射率）

立方平均沸点和 20 度时的相对密度（或折射率）

退出

用户根据事先已经给定（已知）的数据用方向键（↑或↓）进行选择菜单，在输入相应的数据后，屏幕显示计算结果。

10.3 数据处理

打开的数据处理菜单如下：

非线性回归
逐步回归

根据需要，用方向键（↑或↓）进行选择，使所选菜单命令变红。选好菜单后，按〈Enter〉键。用户还应根据屏幕提示，输入有关数据。下面分别对非线性回归和逐步回归进行说明

10.3.1 非线性回归

非线性回归程序运行后，屏幕显示：

是否打印？(Y/N)

若要将计算结果在打印机输出，键入“Y”，否则，键入“N”。接着屏幕显示：

是否要学习例题？Y/N

如要学习例题（磁盘中有简单例题，可供初学者参考），键入“Y”，否则键入“N”。接着屏幕显示：

键盘输入还是磁盘输入？(K/D)

若准备从键盘输入数据，则键入“K”；若选择磁盘输入，则键入“D”（选择磁盘输入时，要输入文件名，同时计算前应在该文件中存有有关数据）。如选择键盘输入，根据屏幕提示，从键盘输入：数据组数、自变量数（本程序最多允许 4 个自变量）、因变量

(函数)名、各自变量名,依次输入各组数据每组数据中先输第1自变量的值、第2自变量的值、...,最后是因变量的值。输入完毕后屏幕显示:

- 1. 显示数据
- 2. 检查数据
- 3. 打印数据
- 4. 数据存盘
- Q. 退出

用户可键入数字键1~4,或字母键Q。根据用户的选择,可将所输入的数据(数组序号、因变量数、自变量数)进行显示,检查修改、打印、或存盘(根据屏幕提示,很容易操作,为节省篇幅,此处从略)。选择Q,则退出。若自变量只有一个,屏幕将询问:

是否图示数据?(Y/N)

如用户要观察输入数据的图形,键入“Y”,否则键入“N”。本软件可在9种不同的坐标体系中图示数据(必须在定义域范围之内),用户可根据需要进行选择。选好坐标系后,屏幕显示自变量与因变量之间的关系图。按C键继续,可根据需要重新选择坐标体系。如要退出图示、按字符键“Q”。

退出图示之后,屏幕给出了预选方程,并询问是否正确(Y/N)?如预选方程不正确,请从键盘输入预选方程。输入预选方程后,再次询问预选方程是否正确。如此反复,直至确认输入的预选方程正确。然后根据屏幕提示,输入方程中的待定系数的初值。如初值合理,可很快计算出结果,然后屏幕显示如下菜单,供用户选择:

- 1. 方程
- 2. 结果比较
- 3. 误差分析图
- 4. 函数图形
- Q. 退出

选择1,屏幕将显示回归方程及方程中的各常数。选择2,显示结果比较,其中有因变量的输入值、计算值,绝对偏差和相对偏差。选择3,屏幕将显示误差分析图。选择4(若自变量只有一个时,该项可选),屏幕输出回归方程的图形,并标出输入的数据点。选择Q,回到上一级菜单。

当选择3(显示误差分析图)时,用户可进行如下选择:

- 1. 绝对偏差
- 2. 相对偏差
- Q. 退出

选择1或2,输出误差分析图(横坐标为因变量,纵坐标为绝对偏差或相对偏差),并给出最大偏差和平均偏差。选择Q,回到上一级菜单。

10.3.2 逐步回归

运行逐步回归程序之后, 根据屏幕提示, 回答是否打印、是否学习例题后, 屏幕显示一级菜单如下:

- 1. 输入数据
- 2. 数据预处理
- 3. 回归计算
- 4. 输出计算结果
- Q. 退出

选择数据输入及数据预处理时, 其过程和方法大致和非线性回归相同 (参见非线性回归)。

选择回归计算, 屏幕显示三种预选方程的形式, 并请用户选择其一。选择了预选方程之后, 根据屏幕提示, 输入最高幂级数 (如果输入不正确, 可以进行修改)。用户可根据需要, 变换因变量 (取对数、倒数) 和自变量。选定了 F 检验临界值后, 可很快算出结果, 回到一级菜单。

选择计算结果 (必须在完成回归计算之后), 屏幕显示如下菜单:

- 1. 方程
- 2. 结果比较
- 3. 误差分析图
- 4. 函数图形
- Q. 退出

输出计算结果的操作与非线性回归相同, 此处不再重复。

10.4 优化计算

打开的优化计算菜单如下:

- 线性规划
- 非线性规划 (拉格朗日乘因子法和单纯形法)
- 非线性规划 (复合形法)

用方向键 (\uparrow 或 \downarrow) 进行选择, 使所选菜单命令变红。选好菜单后, 按 $\langle \text{Enter} \rangle$ 键。

10.4.1 线性规划

选择线性规划, 并按 $\langle \text{Enter} \rangle$ 键后, 屏幕提示:

输入基本变量数:

时，请输入基本变量的数目。屏幕接着显示：

输入约束数：

在输入约束条件数之后，屏幕给出：

目标函数形式为 $F = c_1x_1 + c_2x_2$

按以上格式输入目标函数：

F =

时，在“F=”后键目标函数，并按〈Enter〉键；屏幕接着询问：

是否正确 (Y/N)

如输入正确，键入“Y”；否则键入“N”，并重新输入目标函数。输入目标函数并经确认正确后，屏幕接着显示：

约束方程的形式为：

$a(j,1)x_1 + a(j,2)x_2 \geq (=, \text{或} <) a(j,4)$

按以上格式输入约束方程

方程 1：

时，键入第一约束方程，按〈Enter〉键后，屏幕显示：

是否正确？(Y/N)

和目标函数同样方式进行确认；以同样方式输入并确认各约束方程。当屏幕出现：

是否打印后？(Y/N)

时，键入 Y 或 N 键进行回答。屏幕接着显示：

1. 目标函数求最大值

2. 目标函数求最小值

请选择！

若目标函数求最大值，则键入 1；否则，键入 2。输入完毕，进行计算，计算结束后在屏幕输出如下内容：

目标函数: $F = \dots$

约束条件: \dots

\dots

$$x_i \geq 0$$

计算结果:

目标函数最大值为 \dots

$$x_1 = \dots \quad x_2 = \dots \quad \dots$$

若在询问是否答应时, 键入了 Y, 则屏幕输出的内容同样在打印机上输出。

10.4.2 非线性规划 (拉格朗日乘因子法和单纯形法)

选择非线性规划 (拉格朗日乘因子法和单纯形法), 并按 $\langle \text{Enter} \rangle$ 键后, 屏幕显示:

1. 无约束

2. 有约束

时, 若求解无约束的非线性规划问题, 则键入 1; 若求解有约束的非线性规划问题, 键入 2。然后根据屏幕提示, 输入基本变量数。若为有约束问题, 还要输入约束条件数。接着屏幕显示:

请输入目标函数: $F = f(x_1, x_2, \dots, x_n)$

如: $F = x_1^2 + x_2^2 - x_1 * x_2 - 10 * x_1 - 4 * x_2 + 60$

目标函数: $F =$

时, 在“ $F =$ ”后面键入目标函数, 并按 $\langle \text{Enter} \rangle$ 键。屏幕询问“是否正确?”时进行确认。然后, 根据屏幕提示, 选择目标函数求最小值或最大值。当屏幕显示:

压缩因子 (默认值为 0.75)?

时, 默认压缩因子为 0.75, 则按 $\langle \text{Enter} \rangle$ 键; 否则输入压缩因子值。以同样方法处理屏幕接着询问的其他几个量。回答是否打印后, 屏幕显示:

输入 x_1 (基本变量) 的初值 (默认值为 1)

时, 若用默认值, 则按 $\langle \text{Enter} \rangle$ 键; 否则输入初值。以同样方式输入各变量的初值。计算机即行计算, 计算完毕屏幕输出:

目标函数: $F =$

最优解为: x_1 (基本变量) $= \dots$

\dots

目标函数最小(大)值 $= \dots$

10.4.3 非线性规划(复合形法)

选择非线性规划(复合形法),按〈Enter〉键。根据屏幕提示,选择有、无约束,输入基本变量数,(若为有约束问题,输入约束数,这里输入的约束数不包括变量的上、下限约束;依次输入各变量的约束下限和上限)。然后和10.4.2相同的方式根据屏幕提示输入目标函数和约束方程,并进行确认。选择目标函数求最小值或输入压缩因子等,回答是否打印。

输入各变量的初值,计算结束后屏幕显示:

```

目标函数: ...
约束条件: ...
...
最优解为: x1 = ...
          x2 = ...
          ...
目标函数最小值: F = ...
    
```

附录 化工过程计算软件中的624个组分

标号	分子式	名称	标号	分子式	名称
1	AlBr ₃	三溴化铝	21	ClNO	亚硝酸氯
2	AlCl ₃	三氯化铝	22	Cl ₂	氯
3	AlI ₃	三碘化铝	23	Cl ₂ FP	一氟二氯化磷
4	Ar	氩	24	Cl ₃ P	三氯化磷
5	As	砷	25	Cl ₄ Si	四氯化硅
6	AsCl ₃	三氯化砷	26	Cl ₄ Ti	四氯化钛
7	BBr ₃	三溴化硼	27	Cl ₅ P	五氯化磷
8	BCl ₃	三氯化硼	28	D ₂	氘(顺)
9	BF ₃	三氟化硼	29	D ₂	氘(正常)
10	BI ₃	三碘化硼	30	D ₂ O	重水
11	Br ₂	溴	31	FNO ₂	硝酸氟
12	BrI	溴化碘	32	F ₂	氟
13	Br ₃ P	三溴化磷	33	F ₂ N ₂	顺-二氟化二氮
14	Br ₄ Si	四溴化硅	34	F ₂ N ₂	反-二氟化二氮
15	Br ₄ Ti	四溴化钛	35	F ₂ O	氧化氟
16	ClFO ₃	高氯氧化氟	36	F ₂ Xe	二氟化氙
17	ClF ₂ N	二氟一氯化氮	37	F ₃ N	三氟化氮
18	ClF ₂ P	二氟一氯化磷	38	F ₃ NO	氧化三氟化氮
19	ClF ₂ PS	二氟一氯硫化磷	39	F ₃ P	三氟化磷
20	ClF ₅	五氟化氯	40	F ₃ PS	三氟硫化磷

标号	分子式	名 称	标号	分子式	名 称
41	F_4N_2	四氟化肼	81	H_3P	磷化氢
42	F_4S	四氟化硫	82	H_4ClN	氯化氮
43	F_4Si	四氟化硅	83	H_4ClP	氯化磷
44	F_4Xe	四氟化氙	84	H_4N_2	肼
45	F_6S	六氟化硫	85	H_4Si	硅烷
46	F_6U	六氟化铀	86	H_6B_2	二甲硼烷
47	He	氦 ³	87	$CBrClF_2$	溴氯二氟甲烷
48	He	氦 ⁴	88	CBr_2F_2	二溴二氟甲烷
49	Hg	汞	89	$CBrF_3$	三氟溴甲烷
50	I_2	碘	90	$CClF_3$	三氟氯甲烷
51	I_4Si	四碘化硅	91	CCl_2F_2	二氟二氯甲烷
52	I_4Ti	四碘化钛	92	CCl_2O	光气
53	Kr	氪	93	CCl_3F	三氯氟甲烷
54	NO	氧化氮	94	CCl_4	四氯化碳
55	NO_2	二氧化氮	95	CD_4	甲氘烷
56	N_2	氮	96	CF_4	四氟化碳
57	N_2O	氧化亚氮	97	CO	一氧化碳
58	Ne	氖	98	COS	羰基硫
59	O_2	氧	99	CO_2	二氧化碳
60	SO_2	二氧化硫	100	CS_2	二硫化碳
61	O_3	臭氧	101	$CHClF_2$	二氟氯甲烷
62	O_3S	三氧化硫	102	$CHCl_2F$	二氯氟甲烷
63	P	磷	103	$CHCl_3$	三氯甲烷
64	Rn	氡	104	CHF_3	三氟甲烷(氟仿)
65	S	硫	105	CHN	氰化氢
66	Se	硒	106	CH_2Br_2	二溴甲烷
67	T_2	氡	107	CH_2Cl_2	二氯甲烷
68	X_e	氙	108	CH_2F_2	二氟代甲烷
69	HBr	溴化氢	109	CH_2O	甲醛
70	HCl	氯化氢	110	CH_2O_2	甲酸
71	HD	氘化氢	111	CH_3Br	溴甲烷
72	HF	氟化氢	112	CH_3Cl	氯甲烷
73	HI	碘化氢	113	CH_3F	氟甲烷
74	HF_2N	二氟化胺	114	CH_3I	碘甲烷
75	H_2	氢(正常)	115	CH_3NO_2	硝基甲烷
76	H_2	氢	116	CH_4	甲烷
77	H_2O	水	117	CH_4O	甲醇
78	H_2S	硫化氢	118	CH_4S	甲硫醇
79	H_3As	三氢化砷	119	CH_3N	甲胺
80	H_3N	氨	120	CH_6N_2	甲基肼

标号	分子式	名 称	标号	分子式	名 称
121	CH_3Si	甲基甲硅烷	161	$\text{C}_2\text{H}_4\text{O}$	环氧乙烷
122	$\text{C}_2\text{Br}_2\text{ClF}_3$	1,2-二溴-1-氯-三氟乙烷	162	$\text{C}_2\text{H}_4\text{O}_2$	乙酸
123	$\text{C}_2\text{Br}_2\text{F}_4$	1,2-二溴-四氟乙烷	163	$\text{C}_2\text{H}_4\text{O}_2$	甲酸甲酯
124	C_2ClF_3	三氟氯乙烷	164	$\text{C}_2\text{H}_5\text{Br}$	溴乙烷
125	C_2ClF_5	五氟氯乙烷	165	$\text{C}_2\text{H}_5\text{Cl}$	氯乙烷
126	$\text{C}_2\text{Cl}_2\text{F}_4$	1,2,2,2-四氟二氯乙烷	166	$\text{C}_2\text{H}_5\text{F}$	氟乙烷
127	$\text{C}_2\text{Cl}_2\text{F}_4$	1,1,2,2-四氟二氯乙烷	167	$\text{C}_2\text{H}_5\text{I}$	碘乙烷
128	$\text{C}_2\text{Cl}_3\text{F}_3$	1,2,2-三氯三氟乙烷	168	$\text{C}_2\text{H}_5\text{N}$	亚乙基亚胺
129	C_2Cl_4	四氯乙烷	169	C_2H_6	乙烷
130	$\text{C}_2\text{Cl}_4\text{F}_2$	1,1,2,2-四氯二氟乙烷	170	$\text{C}_2\text{H}_6\text{O}$	甲醚
131	$\text{C}_2\text{F}_3\text{N}$	三氟代乙腈	171	$\text{C}_2\text{H}_6\text{O}$	乙醇
132	C_2F_4	四氟乙烷	172	$\text{C}_2\text{H}_6\text{O}_2$	乙二醇
133	C_2F_6	全氟乙烷	173	$\text{C}_2\text{H}_6\text{S}$	乙硫醇
134	C_2N_2	氰	174	$\text{C}_2\text{H}_6\text{S}$	甲硫醚
135	C_2HClF_2	1-氟-2,2-二氟代乙烷	175	$\text{C}_2\text{H}_7\text{N}$	乙胺
136	C_2HClF_2	1,1,2,2-四氟氯乙烷	176	$\text{C}_2\text{H}_7\text{N}$	二甲胺
137	C_2HCl_3	三氯乙烷	177	$\text{C}_2\text{H}_7\text{NO}$	2-胺基乙醇
138	C_2HCl_5	五氯乙烷	178	$\text{C}_2\text{H}_8\text{N}_2$	1,2-二胺基乙烷
139	$\text{C}_2\text{HF}_3\text{O}_2$	三氟乙酸	179	$\text{C}_3\text{ClF}_5\text{O}$	氯五氟丙酮
140	C_2H_2	乙炔	180	$\text{C}_2\text{F}_6\text{O}$	全氟丙酮
141	$\text{C}_2\text{H}_2\text{Cl}_2$	顺-1,2-二氯乙烷	181	C_2F_8	全氟丙烷
142	$\text{C}_2\text{H}_2\text{Cl}_2$	反-1,2-二氯乙烷	182	$\text{C}_3\text{H}_5\text{F}_3$	三氟丙烷
143	$\text{C}_2\text{H}_2\text{Cl}_4$	1,1,2,2-四氯乙烷	183	$\text{C}_2\text{H}_6\text{F}_5$	1,1,1,2,2-五氟丙烷
144	$\text{C}_2\text{H}_2\text{F}_2$	1,1-二氟乙烷	184	$\text{C}_2\text{H}_3\text{N}$	丙烯腈
145	$\text{C}_2\text{H}_2\text{O}$	乙炔酮	185	$\text{C}_3\text{H}_5\text{NO}$	异恶唑
146	$\text{C}_2\text{H}_3\text{Cl}$	氯乙烯	186	C_3H_4	丙二烯
147	$\text{C}_2\text{H}_3\text{ClF}_2$	1,1-二氟氯乙烷	187	C_3H_6	丙炔
148	$\text{C}_2\text{H}_3\text{ClO}$	乙炔氯	188	$\text{C}_2\text{H}_6\text{O}$	丙烯醛
149	$\text{C}_2\text{H}_3\text{Cl}_3$	1,1,1-三氯乙烷	189	$\text{C}_2\text{H}_6\text{O}_2$	丙烯酸
150	$\text{C}_2\text{H}_3\text{Cl}_3$	1,1,2-三氯乙烷	190	$\text{C}_2\text{H}_4\text{O}_2$	甲酸乙烯酯
151	$\text{C}_2\text{H}_3\text{F}$	氟乙烷	191	$\text{C}_3\text{H}_2\text{Cl}$	3-氯丙烯
152	$\text{C}_2\text{H}_3\text{F}_3$	1,1,1-三氟乙烷	192	$\text{C}_3\text{H}_6\text{Cl}_3$	1,2,3-三氯丙烷
153	$\text{C}_2\text{H}_3\text{N}$	乙腈	193	$\text{C}_2\text{H}_5\text{N}$	丙腈
154	$\text{C}_2\text{H}_3\text{NO}$	异氰酸甲酸	194	C_3H_6	环丙烷
155	C_2H_6	乙烷	195	C_2H_6	丙烯
156	$\text{C}_2\text{H}_4\text{Br}_2$	1,2-二溴乙烷	196	$\text{C}_2\text{H}_6\text{Cl}_2$	1,2-二氯丙烷
157	$\text{C}_3\text{H}_4\text{Cl}_2$	1,1-二氯乙烷	197	$\text{C}_3\text{H}_6\text{O}$	丙酮
158	$\text{C}_2\text{H}_4\text{Cl}_2$	1,2-二氯乙烷	198	$\text{C}_2\text{H}_6\text{O}$	丙烯醇
159	$\text{C}_2\text{H}_6\text{F}_2$	1,1-二氟乙烷	199	$\text{C}_2\text{H}_6\text{O}$	1-丙醇
160	$\text{C}_2\text{H}_4\text{O}$	乙醚	200	$\text{C}_2\text{H}_4\text{O}$	环氧丙烷

标号	分子式	名 称	标号	分子式	名 称
201	C_3H_6O	乙烯甲醚	241	C_4H_8	环丁烷
202	$C_3H_6O_2$	丙酸	242	C_4H_8	2-甲基丙烯
203	$C_3H_6O_2$	甲酸乙酯	243	C_4H_8O	正丁醛
204	$C_3H_6O_2$	乙酸甲酯	244	C_4H_8O	异丁醛
205	C_3H_7Cl	氯丙烷	245	C_4H_8O	甲乙酮
206	C_3H_7Cl	2-氯丙烷	246	C_4H_8O	四氢呋喃
207	C_3H_8	丙烷	247	C_4H_8O	乙烯基乙基醚
208	C_3H_8O	正丙醇	248	$C_4H_8O_2$	正丁酸
209	C_3H_8O	异丙醇	249	$C_4H_8O_2$	异丁酸
210	C_3H_6O	甲乙醚	250	$C_4H_8O_2$	1,4-二噁烷
211	$C_3H_8O_2$	二甲氧基甲烷	251	$C_4H_8O_2$	乙酸乙酯
212	$C_3H_8O_2$	1,2-丙二醇	252	$C_4H_8O_2$	丙酸甲酯
213	$C_3H_8O_2$	1,3-丙二醇	253	$C_4H_8O_2$	甲酸正丙酯
214	$C_3H_8O_3$	1,2,3-丙三醇	254	C_4H_8S	四氢噻吩
215	C_3H_8S	甲乙硫醚	255	C_4H_9Cl	1-氯丁烷
216	$C_3H_8BO_3$	硼酸三甲酯	256	C_4H_9Cl	2-氯丁烷
217	C_3H_9N	正丙胺	257	C_4H_9Cl	2-氯-2-甲基丙烷
218	C_3H_9N	异丙胺	258	C_4H_9N	四氢吡咯
219	C_3H_9N	三甲胺	259	C_4H_9NO	四氢噁唑
220	C_4F_8	全氟环丁烷	260	C_4H_{10}	正丁烷
221	C_4F_{10}	全氟丁烷	261	C_4H_{10}	异丁烷
222	$C_4H_2O_3$	顺丁烯二酸酐	262	$C_4H_{10}O$	正丁醇
223	C_4H_4	乙烯基乙炔	263	$C_4H_{10}O$	仲丁醇
224	C_4H_8O	呋喃	264	$C_4H_{10}O$	异丁醇
225	C_4H_4S	噻吩	265	$C_4H_{10}O$	叔丁醇
226	C_4H_5N	α -丁烯腈	266	$C_4H_{10}O$	乙醚
227	C_4H_5N	吡咯	267	$C_4H_{10}O$	甲丙醚
228	C_4H_6	1-丁炔	268	$C_4H_{10}O$	甲基异丙基醚
229	C_4H_6	2-丁炔	269	$C_4H_{10}O_2$	1,2-二甲氧基乙烷
230	C_4H_6	1,2-丁二烯	270	$C_4H_{10}O_2$	二甘醇
231	C_4H_6	1,3-丁二烯	271	$C_4H_{10}S$	乙硫醚
232	$C_4H_6O_3$	乙酸乙烯酯	272	$C_4H_{10}S_2$	乙二硫醚
233	$C_4H_6O_2$	丙烯酸甲酯	273	$C_4H_{11}N$	正丁胺
234	$C_4H_6O_2$	乙酸酐	274	$C_4H_{11}N$	异丁胺
235	$C_4H_6O_4$	草酸二甲酯	275	$C_4H_{11}N$	二乙胺
236	$C_4H_6O_4$	丁二酸	276	C_5F_{12}	全氟戊烷
237	C_4H_7N	丁腈	277	$C_5H_2F_6O_2$	六氟代乙硫丙酮
238	C_4H_8	1-丁烯	278	$C_5H_4O_2$	糠醛
239	C_4H_8	顺-2-丁烯	279	C_5H_5N	吡啶
240	C_4H_8	反-2-丁烯	280	$C_5H_6N_2$	2-甲基吡嗪

标号	分子式	名 称	标号	分子式	名 称
281	C_5H_6O	2-甲基呋喃	321	$C_5H_{12}O$	2,2-二甲基-1-丙醇
282	C_5H_8	环戊烯	322	$C_5H_{12}O$	乙丙醚
283	C_5H_8	1,2-戊二烯	323	$C_5H_{12}O$	丁甲醚
284	C_5H_8	反-1,3-戊二烯	324	$C_5H_{12}O$	叔丁基甲基醚
285	C_5H_8	1,4-戊二烯	325	C_6BrF_5	一溴五氟代苯
286	C_5H_6	1-戊炔	326	C_6ClF_5	一氯五氟代苯
287	C_5H_8	异戊二烯	327	$C_6Cl_2F_4$	二氯四氟代苯
288	C_5H_8	2-甲基-1,2-丁二烯	328	$C_6Cl_3F_3$	1,3,5-三氯三氟代苯
289	C_5H_8	3-甲基-1,2-丁二烯	329	C_6F_6	六氟苯
290	C_5H_8O	环戊酮	330	C_6F_{12}	全氟环己烷
291	$C_5H_8O_2$	二氢吡嗪	331	C_6F_{14}	全氟正己烷
292	$C_5H_8O_2$	丙烯酸乙酯	332	C_6F_{14}	全氟代-2-甲基戊烷
293	C_5H_{10}	环戊烷	333	C_6F_{14}	全氟代-3-甲基戊烷
294	C_5H_{10}	1-戊烯	334	C_6F_{14}	全氟代-2,3-二甲基丁烷
295	C_5H_{10}	顺-2-戊烯	335	C_6HF_5	五氟代苯
296	C_5H_{10}	反-2-戊烯	336	C_6HF_5O	五氟代酚
297	C_5H_{10}	2-甲基-1-丁烯	337	$C_6H_2F_4$	1,2,3,4-四氟代苯
298	C_5H_{10}	2-甲基-2-丁烯	338	$C_6H_2F_4$	1,2,3,5-四氟代苯
299	C_5H_{10}	3-甲基-1-丁烯	339	$C_6H_2F_4$	1,2,4,5-四氟代苯
300	$C_5H_{10}O$	戊醛	340	$C_6H_4Cl_2$	邻二氯苯
301	$C_5H_{10}O$	甲基正丙基酮	341	$C_6H_4Cl_2$	间二氯苯
302	$C_5H_{10}O$	甲基异丙基酮	342	$C_6H_4Cl_2$	对二氯苯
303	$C_5H_{10}O$	二乙基酮	343	$C_6H_6F_2$	二氟代苯
304	$C_5H_{10}O$	2-甲基-四氢呋喃	344	C_6H_5Br	溴苯
305	$C_5H_{10}O$	四氢吡嗪	345	C_6H_5Cl	氯苯
306	$C_5H_{10}O_2$	正戊酸	346	C_6H_5F	氟苯
307	$C_5H_{10}O_2$	异戊酸	347	C_6H_5I	碘苯
308	$C_5H_{10}O_2$	甲酸异丁酯	348	C_6H_6	苯
309	$C_5H_{10}O_2$	乙酸正丙酯	349	C_6H_6O	苯酚
310	$C_5H_{10}O_2$	丙酸乙酯	350	C_6H_7N	苯胺
311	$C_5H_{10}O_2$	丁酸甲酯	351	C_6H_7N	2-甲基吡啶
312	$C_5H_{10}O_2$	异丁酸甲酯	352	C_6H_7N	3-甲基吡啶
313	$C_5H_{11}N$	六氢吡啶	353	C_6H_7N	4-甲基吡啶
314	C_5H_{12}	正戊烷	354	C_6H_{10}	1,5-己二烯
315	C_5H_{12}	2-甲基丁烷	355	C_6H_{10}	环己烯
316	C_5H_{12}	2,2-二甲基丙烷	356	$C_6H_{10}O$	环己酮
317	$C_5H_{12}O$	正戊醇	357	$C_6H_{11}N$	己腈
318	$C_5H_{12}O$	2-甲基-1-丁醇	358	C_6H_{12}	环己烷
319	$C_5H_{12}O$	3-甲基-1-丁醇	359	C_6H_{12}	甲基环戊烷
320	$C_5H_{12}O$	2-甲基-2-丁醇	360	C_6H_{12}	1-己烯

标号	分子式	名 称	标号	分子式	名 称
361	C_6H_{12}	顺-2-己烯	401	$C_7H_3F_5$	2,3,4,5,6-五氟代甲苯
362	C_6H_{12}	反-2-己烯	402	C_7H_5N	苯甲腈
363	C_6H_{12}	顺-3-己烯(顺)	403	C_7H_6O	苯甲醛
364	C_6H_{12}	反-3-己烯	404	$C_7H_6O_2$	苯甲酸
365	C_6H_{12}	2-甲基-2-戊烯	405	C_7H_8	甲苯
366	C_6H_{12}	3-甲基-2-戊烯	406	C_7H_8O	苯甲醚
367	C_6H_{12}	3-甲基-2-戊烯	407	C_7H_8O	苯醇
368	C_6H_{12}	4-甲基-2-戊烯	408	C_7H_6O	邻甲酚
369	C_6H_{12}	4-甲基-2-戊烯	409	C_7H_8O	间甲酚
370	C_6H_{12}	2,3-二甲基-1-丁烯	410	C_7H_8O	对甲酚
371	C_6H_{12}	2,3-二甲基-2-丁烯	411	C_7H_8N	2,3-二甲基吡啶
372	C_6H_{12}	3,3-二甲基-1-丁烯	412	C_7H_9N	2,4-二甲基吡啶
373	$C_6H_{12}O$	环己醇	413	C_7H_8N	2,5-二甲基吡啶
374	$C_6H_{12}O$	乙基丙基甲酮	414	C_7H_9N	2,6-二甲基吡啶
375	$C_6H_{12}O$	甲基丁基甲酮	415	C_7H_9N	3,4-二甲基吡啶
376	$C_6H_{12}O$	甲基异丁基酮	416	C_7H_9N	3,5-二甲基吡啶
377	$C_6H_{12}O_2$	乙酸正丁酯	417	C_7H_9N	甲基苯胺
378	$C_6H_{12}O_2$	乙酸异丁酯	418	C_7H_9N	邻甲苯胺
379	$C_6H_{12}O_2$	丁酸乙酯	419	C_7H_9N	间甲苯胺
380	$C_6H_{12}O_2$	异丁酸乙酯	420	C_7H_9N	对甲苯胺
381	$C_6H_{12}O_2$	丙酸正丙酯	421	C_7H_{14}	环庚烯
382	$C_6H_{12}O_2$	甲酸正戊酯	422	C_7H_{14}	1,1-二甲基环戊烷
383	$C_6H_{12}O_2$	甲酸异戊酯	423	C_7H_{14}	顺-1,2-二甲基环戊烷
384	C_6H_{14}	正己烷	424	C_7H_{14}	反-1,2-二甲基环戊烷
385	C_6H_{14}	2-甲基戊烷	425	C_7H_{14}	乙基环戊烷
386	C_6H_{14}	3-甲基戊烷	426	C_7H_{14}	甲基环己烷
387	C_6H_{14}	2,2-二甲基丁烷	427	C_7H_{14}	1-庚烷
388	C_6H_{14}	2,3-二甲基丁烷	428	C_7H_{14}	2,3,3-三甲基-1-丁烯
389	$C_6H_{14}O$	1-己醇	429	C_7H_{14}	甲基戊基甲酮
390	$C_6H_{14}O$	2-己醇	430	$C_7H_{14}O_2$	n-丁酸丙酯
391	$C_6H_{14}O$	乙丁醚	431	$C_7H_{14}O_2$	异丁酸丙酯
392	$C_6H_{14}O$	甲基戊基醚	432	$C_7H_{14}O_2$	乙酸异戊酯
393	$C_6H_{14}O$	二丙醚	433	$C_7H_{14}O_2$	丙酸异丁酯
394	$C_6H_{14}O$	二异丙醚	434	C_7H_{16}	正庚烷
395	$C_6H_{15}N$	二丙胺	435	C_7H_{16}	2-甲基己烷
396	$C_6H_{15}N$	二异丙胺	436	C_7H_{16}	3-甲基己烷
397	$C_6H_{15}N$	三乙胺	437	C_7H_{16}	2,2-二甲基戊烷
398	C_7F_8	全氟代甲苯	438	C_7H_{16}	2,3-二甲基戊烷
399	C_7H_{14}	全氟甲基环己烷	439	C_7H_{16}	2,4-二甲基戊烷
400	C_7H_{16}	全氟正庚烷	440	C_7H_{16}	3,3-二甲基戊烷

标号	分子式	名 称	标号	分子式	名 称
441	C_7H_{16}	3-乙基戊烷	480	C_8H_{16}	异丙基环戊烷
442	C_7H_{16}	2,2,3-三甲基丁烷	481	C_8H_{16}	环辛烷
443	$C_7H_{16}O$	1-庚醇	482	C_8H_{16}	1-辛烯
444	$C_8H_4O_3$	苯二甲苯酐	483	C_8H_{16}	2-辛烯(反)
445	C_8H_8	苯乙烯	484	$C_8H_{16}O_2$	丙酸异戊酯
446	C_8H_8O	苯乙酮	485	$C_8H_{16}O_3$	丁酸异戊酯
447	$C_8H_8O_2$	苯甲酸甲酯	486	$C_8H_{16}O_2$	异丁酸异丁酯
448	$C_8H_8O_2$	水杨酸甲酯	487	$C_8H_{16}O_2$	异戊酸丙酯
449	C_8H_{10}	邻二甲苯	488	C_8H_{18}	正辛烷
450	C_8H_{10}	间二甲苯	489	C_8H_{18}	2-甲基庚烷
451	C_8H_{10}	对二甲苯	490	C_8H_{18}	3-甲基庚烷
452	C_8H_{10}	乙苯	491	C_8H_{18}	4-甲基庚烷
453	$C_8H_{10}O$	邻乙基酚	492	C_8H_{18}	2,2-二甲基己烷
454	$C_8H_{10}O$	间乙基酚	493	C_8H_{18}	2,3-二甲基己烷
455	$C_8H_{10}O$	对乙基酚	494	C_8H_{18}	2,4-二甲基己烷
456	$C_8H_{10}O$	苯乙醚	495	C_8H_{18}	2,5-二甲基己烷
457	$C_8H_{10}O$	2,3-二甲苯酚	496	C_8H_{18}	3,3-二甲基己烷
458	$C_8H_{10}O$	2,4-二甲苯酚	497	C_8H_{18}	3,4-二甲基己烷
459	$C_8H_{10}O$	2,5-二甲苯酚	498	C_8H_{18}	3-乙基己烷
460	$C_8H_{10}O$	2,6-二甲苯酚	499	C_8H_{18}	2,2,3-三甲基戊烷
461	$C_8H_{10}O$	3,4-二甲苯酚	500	C_8H_{18}	2,2,4-三甲基戊烷
462	$C_8H_{10}O$	3,5-二甲苯酚	501	C_8H_{18}	2,3,3-三甲基戊烷
463	$C_8H_{11}N$	<i>n,n</i> -二甲苯胺	502	C_8H_{18}	2,3,4-三甲基戊烷
464	$C_8H_{11}N$	<i>N</i> -乙基苯胺	503	C_8H_{18}	2-甲基-3-乙基戊烷
465	$C_8H_{14}O_4$	丁二基酸二乙酯	504	C_8H_{18}	3-甲基-3-乙基戊烷
466	C_8H_{16}	1,1-二甲基环己烷	505	C_8H_{18}	2,2,3,3-四甲基丁烷
467	C_8H_{16}	顺-1,2-二甲基环己烷	506	$C_8H_{18}O$	1-辛醇
468	C_8H_{16}	反-1,2-二甲基环己烷	507	$C_8H_{18}O$	2-辛醇
469	C_8H_{16}	顺-1,3-二甲基环己烷	508	$C_8H_{18}O$	4-甲基-3-庚醇
470	C_8H_{16}	反-1,3-二甲基环己烷	509	$C_8H_{18}O$	5-甲基-3-庚醇
471	C_8H_{16}	顺-1,4-二甲基环己烷	510	$C_8H_{18}O$	2-乙基己醇
472	C_8H_{16}	反-1,4-二甲基环己烷	511	$C_8H_{18}O$	丁醚
473	C_8H_{16}	乙基环己烷	512	$C_8H_{18}O$	二叔丁基醚
474	C_8H_{16}	1,1,2-三甲基环戊烷	513	$C_8H_{19}N$	二丁胺
475	C_8H_{16}	1,1,3-三甲基环戊烷	514	$C_8H_{19}N$	二异丁胺
476	C_8H_{16}	顺顺反-1,2,4-三甲基环戊烷	515	C_9H_7N	噻啉
477	C_8H_{16}	顺反顺-1,2,4-三甲基环戊烷	516	C_9H_7N	异噻啉
478	C_8H_{16}	1-甲基-1-乙基-环戊烷	517	C_9H_{10}	茚
479	C_8H_{16}	正丙基环戊烷	518	C_9H_{10}	α -甲基苯乙烯
			519	$C_9H_{10}O_2$	苯甲酸乙酯
			520	C_9H_{12}	正丙基苯

标号	分子式	名 称	标号	分子式	名 称
521	C_9H_{12}	异丙苯	561	$C_{10}H_{15}N$	正丁苯胺
522	C_9H_{12}	1-甲基-2-乙基苯	562	$C_{10}H_{18}$	顺萘烷
523	C_9H_{12}	1-甲基-3-乙基苯	563	$C_{10}H_{18}$	反萘烷
524	C_9H_{12}	1-甲基-4-乙基苯	564	$C_{10}H_{18}$	1,3-癸二烯
525	C_9H_{12}	1,2,3-三甲基苯	565	$C_{10}H_{19}N$	癸腈
526	C_9H_{12}	1,2,4-三甲基苯	566	$C_{10}H_{20}$	正丁基环己烷
527	C_9H_{12}	1,3,5-三甲基苯	567	$C_{10}H_{20}$	异丁基环己烷
528	$C_9H_{13}N$	<i>n</i> , <i>n</i> -二甲基-邻-甲苯胺	568	$C_{10}H_{20}$	仲丁基环己烷
529	C_9H_{18}	正丙基环己烷	569	$C_{10}H_{20}$	叔丁基环己烷
530	C_9H_{18}	异丙基环己烷	570	$C_{10}H_{20}$	1-癸烯
531	C_9H_{18}	1,反-3,5-三甲基环己烷	571	$C_{10}H_{20}O$	薄荷醇
532	C_9H_{18}	1-壬烯	572	$C_{10}H_{22}$	正癸烷
533	$C_9H_{18}O$	二丁基甲酮	573	$C_{10}H_{22}$	3,3,5-三甲基庚烷
534	C_9H_{20}	正壬烷	574	$C_{10}H_{22}$	2,2,3,3-四甲基己烷
535	C_9H_{20}	2-甲基辛烷	575	$C_{10}H_{22}$	2,2,5,5-四甲基己烷
536	C_9H_{20}	2,2-二甲基庚烷	576	$C_{10}H_{22}O$	1-癸醇
537	C_9H_{20}	2,2,3-三甲基己烷	577	$C_{11}H_{10}$	1-甲基萘
538	C_9H_{20}	2,2,4-三甲基己烷	578	$C_{11}H_{10}$	2-甲基萘
539	C_9H_{20}	2,2,5-三甲基己烷	579	$C_{11}H_{14}O_2$	苯甲酸丁酯
540	C_9H_{20}	3,3-二乙基戊烷	580	$C_{11}H_{16}$	五甲基苯
541	C_9H_{20}	2,2,3,3-四甲基戊烷	581	$C_{11}H_{22}$	正己基环戊烷
542	C_9H_{20}	2,2,3,4-四甲基戊烷	582	$C_{11}H_{22}$	1-十一烯
543	C_9H_{20}	2,2,4,4-四甲基戊烷	583	$C_{11}H_{24}$	正十一烷
544	C_9H_{20}	2,3,3,4-四甲基戊烷	584	$C_{12}H_{10}$	联苯
545	$C_9H_{20}O$	1-壬醇	585	$C_{12}H_{10}O$	二苯醚
546	$C_{10}F_8$	全氟萘	586	$C_{12}H_{18}$	六甲基苯
547	$C_{10}F_{18}$	全氟萘烷	587	$C_{12}H_{24}$	正庚基环戊烷
548	$C_{10}H_8$	萘	588	$C_{12}H_{24}$	1-十二烯
549	$C_{10}H_{12}$	1,2,3,4-四氢化萘	589	$C_{12}H_{26}$	正十二烷
550	$C_{10}H_{14}$	正丁苯	590	$C_{12}H_{26}O$	十二醚
551	$C_{10}H_{14}$	异丁苯	591	$C_{12}H_{26}O$	十二醇
552	$C_{10}H_{14}$	仲丁苯	592	$C_{12}H_{27}N$	三丁胺
553	$C_{10}H_{14}$	叔丁苯	593	$C_{13}H_{12}$	二苯基甲烷
554	$C_{10}H_{14}$	1-甲基-2-异丙苯	594	$C_{13}H_{26}$	正辛基环戊烷
555	$C_{10}H_{14}$	1-甲基-3-异丙苯	595	$C_{13}H_{26}$	1-十三烯
556	$C_{10}H_{14}$	1-甲基-4-异丙苯	596	$C_{13}H_{28}$	正十三烷
557	$C_{10}H_{14}$	1,4-二乙基苯	597	$C_{14}H_{10}$	蒽
558	$C_{10}H_{14}$	1,2,3,5-四甲基苯	598	$C_{14}H_{10}$	菲
559	$C_{10}H_{14}$	1,2,4,5-四甲基苯	599	$C_{14}H_{28}$	正壬基环戊烷
560	$C_{10}H_{14}O$	5-甲基-2-异丙基苯酚	600	$C_{14}H_{28}$	1-十四烯

标号	分子式	名 称	标号	分子式	名 称
601	$C_{14}H_{30}$	正十四烷	613	$C_{18}H_{14}$	间三苯
602	$C_{15}H_{30}$	正癸基环戊烷	614	$C_{18}H_{14}$	对三苯
603	$C_{15}H_{30}$	1-十五烯	615	$C_{18}H_{36}$	1-十八烯
604	$C_{15}H_{32}$	正十五烷	616	$C_{18}H_{36}$	正十三基环戊烷
605	$C_{16}H_{22}O_4$	邻苯二甲酸二丁酯	617	$C_{18}H_{38}$	正十八烷
606	$C_{16}H_{32}$	正癸基环己烷	618	$C_{18}H_{38}O$	1-十八醇
607	$C_{16}H_{32}$	1-十六烯	619	$C_{18}H_{38}$	正十四基环戊烷
608	$C_{16}H_{34}$	正十六烷	620	$C_{19}H_{40}$	正十九烷
609	$C_{17}H_{34}$	正十二基环戊烷	621	$C_{20}H_{40}$	正十五基环戊烷
610	$C_{17}H_{36}O$	十七醇	622	$C_{20}H_{42}$	正二十烷
611	$C_{17}H_{36}$	正十七烷	623	$C_{20}H_{42}O$	1-二十醇
612	$C_{18}H_{14}$	邻三苯	624	$C_{21}H_{42}$	正十六基环戊烷