

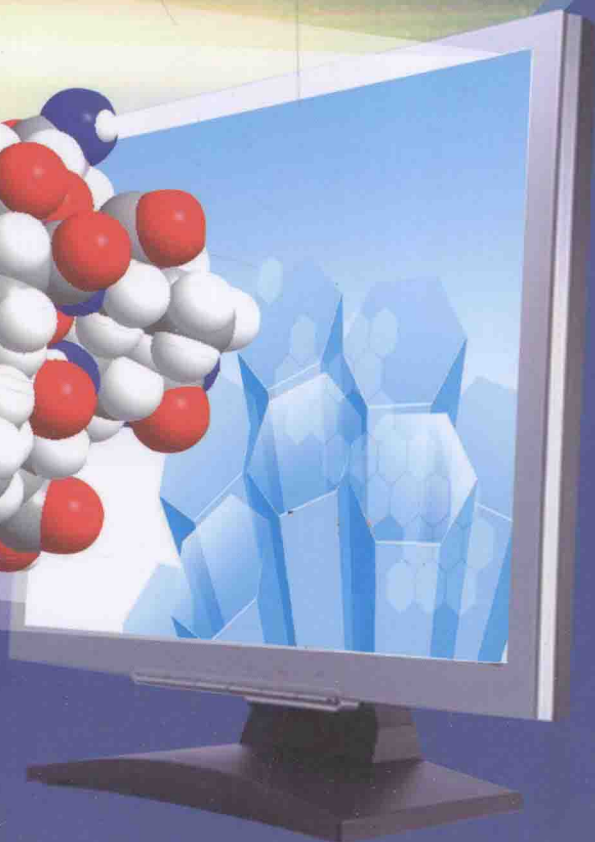
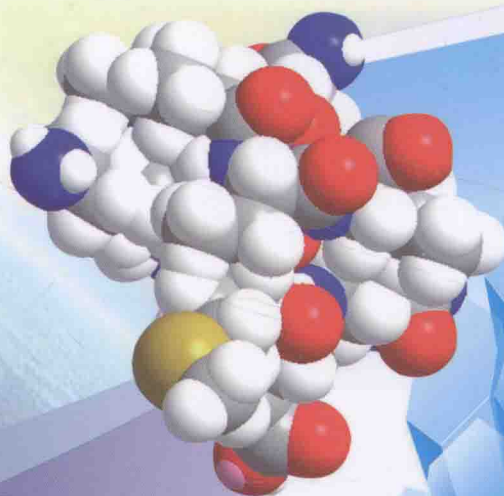
普通高等教育“十二五”规划教材

计算机化学

同济大学化学系

丛培盛 朱仲良 主编

许新华 曹同成 柴晓兰 编



化学工业出版社

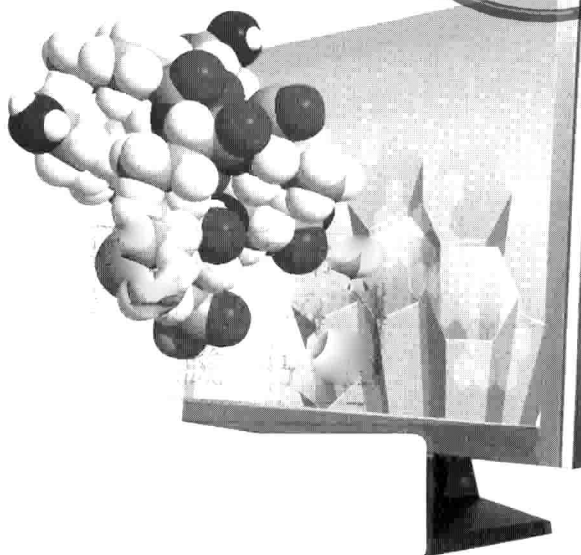
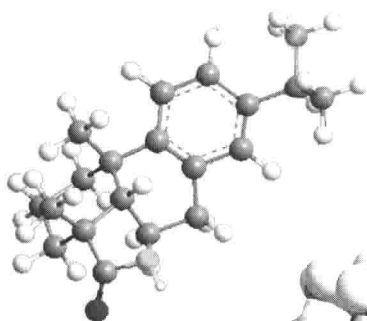
普通高等教育“十二五”规划教材

计算机化学

同济大学化学系

丛培盛 朱仲良 主编

许新华 曹同成 柴晓兰 编



化学工业出版社

· 北京 ·

本书参照高校计算机化学课程教材的内容设置,结合高校计算机基础课程设置及化学专业计算需求,适应社会发展,经过多年的教学实践,逐步调整、编排而成。全书采用当前流行的.net C#平台为编程环境,内容分为编程基础、多元校正、模式识别、信号处理、实验设计、优化、计算机模拟、常用软件介绍等。全书力求层次清晰、难易结合,既提供基础编程,也提供大量可直接调用的类库,辅助以大量的实际实验数据,从而可以从开发、应用两个方面锻炼、提高学习者利用计算机解决化学问题的能力。

本书可以作为高等学校化学专业“计算机化学”课程教材,也可以作为初涉计算机化学的研究人员的参考资料,特别是本书提供的算法类库,可以为解决实际问题编程提供很多方便。

图书在版编目(CIP)数据

计算机化学 / 丛培盛, 朱仲良主编. —北京: 化学工业出版社, 2014.7

普通高等教育“十二五”规划教材

ISBN 978-7-122-20656-5

I. ①计… II. ①丛… ②朱 III. ①计算机应用-化学-高等学校-教材 IV. ①06-39

中国版本图书馆 CIP 数据核字(2014)第 097215 号

责任编辑: 刘俊之

文字编辑: 向 东

责任校对: 蒋 宇

装帧设计: 韩 飞

出版发行: 化学工业出版社(北京市东城区青年湖南街 13 号 邮政编码 100011)

印 刷: 北京永鑫印刷有限责任公司

装 订: 三河市宇新装订厂

787mm×1092mm 1/16 印张 19¼ 字数 510 千字 彩插 2 2014 年 9 月北京第 1 版第 1 次印刷

购书咨询: 010-64518888(传真: 010-64519686) 售后服务: 010-64518899

网 址: <http://www.cip.com.cn>

凡购买本书,如有缺损质量问题,本社销售中心负责调换。

定 价: 66.00 元

版权所有 违者必究



前言

计算机和化学的关系,是随着计算机程序设计的应用普及、计算机采样理论的发展、化学检测手段的现代化而逐步发展起来的。从 1928 年 James 和 Cooleedge 用手动计算机求解薛定谔方程完成氢分子电子结构算起,到现在,计算机已经渗透到了化学研究、教学的各个方面,包括量子化学计算、溶液平衡、定量构效关系、定性定量分析、计算机模拟、实验设计等诸多方面。而现代生物、医药、蛋白质结构与功能研究的兴起,更使得计算机计算体现出越来越重要的作用。

计算化学是大学计算机课程体系建设中的专业计算课程。近年来全国出现了不少优质的计算化学教材,其采用专业化学问题驱动的编排方式使得大学生更易于理解计算机的重要性,这其中的经典之作有张常群教授等编写的《计算化学》,这得益于化学科研工作者多年来在化学计算中的积累,大量的科研计算实例得以引入教材,极大地扩大了学生的视野。

不过,从课程体系建设而言,计算化学教材的建设仍急需加强,这主要表现在两个方面:

(1) 教材中的多数案例采用的程序设计语言不统一,与现代计算机基础教学脱节,例如案例大多用 FORTRAN、BASIC、Matlab 等实现,一本教材的不同案例用不同的语言实现,而目前全国绝大多数高校里的基础计算机语言都以 C/C++、C#、JAVA 等为主,导致教材与现代大学计算机教育脱节较严重,学生看不懂、难理解,教学效率较低。

(2) 多数案例采用基础编程的方式,需要学生花费大量的时间进行编程,使其注意力难以集中到化学问题的求解,但写程序不是化学专业学生的强项,导致课程难以施教。

鉴于以上问题,通过借鉴其他教材的优点,同济大学计算化学教学团队重新编排了教材,主要表现在以下五点:

(1) 理顺教材建设 参考同济大学计算机 2+X 系列课程建设的框架,即大学计算机基础、编程语言、软件开发或数据库应用框架,采用.net 编程环境统一解决教材中案例的编程。

(2) 加强化学计算思维教学 教材中提供大量的程序段及它们的算法构件库,从两个层次提高学生的学习兴趣。一方面通过提供基础编码,提高学生的编程能力,另一方面通过提供类库构件,使学生得以将精力集中于化学问题求解的步骤上,降低问题求解难度,每个步骤简单调用函数即可,例如整个小波库程序代码要 2000 多条,构件化后,学生只需要 3 句话就可以实现小波滤噪。

(3) 问题驱动 教材中从化学实际问题出发,提出需求而后编写算法求解。

(4) 网站计算 作为教材的辅助部分,课题组为本教材提供了配套的教学质

量监督管理网站(样例: <http://cal.tongji.edu.cn/chem>)。网站采取围绕知识点的教学法, 提供有课程章节知识点, 每个知识点采用 2~5min 的短视频深化, 以突出重点、深化重点的形式展开教学, 方便教师进行教学管理。需要教材、管理网站程序及课件的请联系: 丛培盛: pshcong@tongji.edu.cn, 朱仲良: zhuzl@tongji.edu.cn。

学生在了解算法后, 直接准备数据提交到网站, 在浏览器中即可方便完成计算。

(5) 建立教学管理网站 <http://cal.tongji.edu.cn/chem>, 采取围绕知识点的教学法, 教师可以采用突出重点、深化重点的形式展开教学, 在网站中提供课程章节知识点, 每个知识点采用 2~5min 的短视频深化, 在线批阅作业、提出不足, 每节课都有要点回顾, 网站实现作业的自动统计、每个学生参与度的自动统计, 及时了解学生的动态。同济大学计算化学教学团队向采用教材的院校免费提供网站程序并提供适当的技术支持。

通过这些手段, 教师可以采用多条路线教学, 可以采用基础算法编程, 也可以采用调用算法类库, 甚至更简单的只学习如何准备数据, 提交网站计算, 三种路线进行教学, 从而可以极大降低计算化学教学中的编程难度, 提高学生的兴趣和教学效率。

教材中的内容安排如下:

(1) 基础内容 包括第 1 章, 线性代数基础运算及矩阵、向量类库的调用。

(2) 数值计算 第 2、3 章, 内容包括代数方程求解、积分、微分方程求解等问题。

(3) 实验设计 第 4 章主要讲解实验设计及计算获取最优正交设计的方法, 正交实验设计的数据分析等。

(4) 多元校正、信号处理、化学模式识别问题 第 5~7 章结合实例讲解求解该类问题的部分算法及类库调用。

(5) 优化算法 第 8 章, 主要讲解单纯形、遗传算法两种优化算法在化学中的应用。

(6) 计算机模拟 第 9 章讲解蒙特卡洛方法模拟的应用。

(7) 应用软件 第 10 章讲解化学常用软件、计算网站(<http://cal.tongji.edu.cn>)的使用。

本教材共分十章, 第 1、5、6、8、9 章及附录由丛培盛编写, 第 7 章由朱仲良、丛培盛编写, 第 10 章由许新华编写, 第 2 章由曹同成编写, 第 4 章由柴晓兰、丛培盛编写, 第 3 章由许新华、丛培盛编写。全书由朱仲良统稿。

在教材的编写过程中, 得到了同济大学很多教师和同学的悉心指导与帮助, 特别是很多硕士、博士研究生提供了不少程序代码, 包括王志恒、丁长钦、章启元、包嵩同学, 在此表示感谢。

教材中提供的所有程序都在微软的 Visual Studio 2010 上调试运行通过, 有兴趣的读者可以到<http://cal.tongji.edu.cn/chem>中下载。

教材虽然在 3 年多的试用中不断完善修改, 但疏漏仍难避免, 希望广大读者给予宝贵的意见和建议。

编 者

2014 年 1 月

目 录

第 1 章 矩阵运算基础及算法类库	1
1.1 化学测量数据表达	1
1.2 向量、矩阵	3
1.2.1 向量	3
1.2.2 矩阵	4
1.3 向量、矩阵类设计	5
1.3.1 向量类	5
1.3.2 矩阵类	5
1.4 矩阵类库及调用	6
习题	14
第 2 章 化学问题的代数求解	15
2.1 二分法	16
2.1.1 二分法原理	16
2.1.2 二分法程序设计	16
2.1.3 二分法求解弱酸溶液浓度	17
2.2 牛顿迭代法	18
2.2.1 算法原理	18
2.2.2 牛顿迭代法程序设计	18
2.2.3 牛顿迭代法程序改进	21
2.2.4 牛顿迭代法求解二元弱碱溶液浓度	22
习题	22
第 3 章 实验数据的微积分	23
3.1 数据平滑与求导	23
3.1.1 典型化学实验数据	23
3.1.2 数据平滑方法	25
3.1.3 数据求导方法	28

3.2 插值	30
3.2.1 线性插值	30
3.2.2 一元三点拉格朗日插值	32
3.2.3 插值求解乙醇-水汽液平衡	34
3.3 数值积分	38
3.3.1 梯形法	39
3.3.2 辛普森积分	41
3.3.3 离散积分	43
3.3.4 离散积分求解氨合成反应焓变	44
3.4 常微分方程数值解	45
3.4.1 欧拉法	45
3.4.2 龙格-库塔法	47
3.4.3 龙格-库塔法求解复杂反应动力学	50
3.5 数值微分	53
3.5.1 差商法原理	53
3.5.2 程序实现	54
3.5.3 差商法求解反应动力学的级数和速率	54
习题	56

第4章 实验设计 58

4.1 实验设计方法	59
4.2 正交实验设计	59
4.2.1 正交实验设计方法及其特点	59
4.2.2 简单比较法	60
4.2.3 单一正交实验设计	60
4.2.4 基于 J_2 -Optimality 混合正交实验设计	61
4.2.5 正交实验设计分析方法	64
4.2.6 海水中丁基锡含量测定的正交分析	65
4.3 均匀实验设计	74
习题	76

第5章 多元校正 77

5.1 直接校正法	77
5.2 间接校正法	78
5.2.1 K 矩阵法	78
5.2.2 K 矩阵法求解色素混合物的组成	79
5.2.3 P 矩阵法	80
5.3 主成分分析 PCA	80
5.3.1 PCA 算法	81

5.3.2	PCA 算法用途	82
5.3.3	PCA 算法程序	83
5.3.4	PCA 评估光谱测定的可靠性	85
5.4	主成分回归 PCR	87
5.4.1	PCR 算法原理	87
5.4.2	PCR 求解色素混合物浓度	88
5.5	偏最小二乘 PLS	89
5.6	交叉验证	91
5.6.1	交叉验证的原理	91
5.6.2	交叉验证的应用	93
5.7	抽样验证	96
	习题	96

第 6 章 信号处理 97

6.1	多项式拟合去谱图基线	98
6.2	傅里叶变换去除白噪声	102
6.3	小波变换去噪	109
6.3.1	小波定义	110
6.3.2	小波变换	110
6.3.3	小波分析的应用	113
6.3.4	C#小波变换包及应用	116
	习题	118

第 7 章 化学模式识别 119

7.1	数据	119
7.1.1	数据表达	119
7.1.2	数据预处理	120
7.1.3	样本的均衡性	121
7.2	无监督模式识别	121
7.2.1	主成分分析 PCA 及应用	121
7.2.2	K 均值聚类法	128
7.3	有监督模式识别	133
7.3.1	K 最近邻	133
7.3.2	基于主成分分析的 SIMCA 方法	136
7.3.3	偏最小二乘——线性判别分析	139
7.4	选择显著变量	141
	习题	143

第 8 章 化学计算中的优化算法 ----- 144

8.1 优化问题的基本原理 -----	144
8.2 单纯形 -----	145
8.2.1 单纯形算法原理 -----	145
8.2.2 单纯形算法程序 -----	146
8.2.3 单纯形应用 -----	154
8.3 遗传算法 -----	158
8.3.1 遗传算法组成要素 -----	159
8.3.2 遗传算法程序 -----	161
8.3.3 遗传算法在化学中的应用 -----	166
习题 -----	168

第 9 章 蒙特卡洛模拟 ----- 169

9.1 基本原理 -----	170
9.2 应用案例 -----	170
9.2.1 蒙特卡洛法计算定积分 -----	170
9.2.2 蒙特卡洛模拟动力学 -----	173
9.2.3 蒙特卡洛法模拟线性缩聚反应 -----	186
习题 -----	193

第 10 章 化学常用软件 ----- 194

10.1 Origin 软件 -----	194
10.1.1 数据输入 -----	195
10.1.2 图形绘制 -----	196
10.1.3 线性拟合 -----	201
10.1.4 图形输出 -----	204
10.1.5 在一张图中绘制多条曲线 -----	208
10.1.6 多坐标图和图层叠合技术 -----	210
10.1.7 实例——同步热分析曲线的 Origin 输出 -----	213
10.1.8 非线性拟合及自设函数 -----	218
10.1.9 信号处理和峰形分析 -----	225
10.1.10 综合应用实例——求差热峰的起始温度点 -----	232
10.1.11 其他——三角形相图的绘制 -----	234
10.1.12 小结 -----	236
10.2 ChemBioOffice 软件 -----	236
10.2.1 ChemBioDraw 模块 -----	237

10.2.2 ChemBio3D 模块	254
10.2.3 小结	264
10.3 化学计算网站介绍	265
10.3.1 已有算法	265
10.3.2 系统建立	266
10.3.3 多元校正	267
10.3.4 模式识别	272
10.3.5 动力学模拟	277
10.3.6 通用酸碱滴定	279
10.3.7 正交实验设计	280
10.3.8 正交实验极差及方差分析	282
习题	285

附录 ----- 286

附录 1 算法类库	286
附录 2 实用类库	291
附录 3 ZedGraph 简介	292
附录 4 算法语言基础	295

参考文献 ----- 304

· 第1章 ·

➤ 矩阵运算基础及算法类库

化学实验中获得的数据，很多情况下可以表达为向量与矩阵，例如，一条色谱或光谱，可以表达为一个向量；多个样本组成的一组混合物溶液，对每个样本测定光谱后，可以组成一个矩阵。所以，很多问题求解的过程中不可避免地涉及矩阵运算。本章就矩阵与向量运算的基础概念作简单介绍，重点讨论这些基础运算的类库方法及如何调用，为后面的具体问题求解打下基础。

1.1 化学测量数据表达

化学分析的很多测量数据，都可以形成由行和列组成的表，可以看作是由向量和矩阵组成的。这点可以从如下的实例中得到理解。

首先来观察一个简单的日落黄、柠檬黄、胭脂红三种色素混合物溶液的紫外光谱测量实验，设实验设计时准备配制6个混合物溶液，如果将这6个混合物的浓度组成书写成一个表，则可以表达如表1.1所示。

表 1.1 色素混合物紫外吸收光谱实验的浓度

色素	样 1	样 2	样 3	样 4	样 5	样 6
胭脂红	0.00	2.00	4.00	6.00	7.00	10.00
柠檬黄	4.00	5.00	6.00	3.00	8.00	2.00
日落黄	10.00	7.00	5.00	6.00	0.00	3.00

用分光光度计测量表1.1中每个混合物在不同波长下的吸光度，假设从300nm开始测量，每间隔2nm测量一次，到400nm结束，这样对每个样本都可以得到50个测量数据，把它们放在一行，就得到一个向量。如果以这50个数据为纵坐标(y)，以波长为横坐标(x)，就可以得到一条曲线，就是混合物的吸收光谱曲线。如果将这6个混合物的光谱数据逐行排列在一起，也会得到一个有6行、50列的数据表。其特点是：每行是一个混合物在一系列波长下的



吸光度，每列是不同混合物在一个特定的波长下的吸光度。

再来看一个小麦的品质判断的例子，从纯物理测量的角度看，可以取如下的参数来表征：小麦颗粒的长径、短径、颗粒饱满度、色泽度。因此对每一粒小麦，经过以上 4 个参数的测量后，可以得到由 4 个数字组成的一个向量。对多粒小麦进行同样的测量，然后按行排列，就可以得到一个矩阵，从物理意义上来说，每行就是一个样本，每列就是一个测量特征。如果测量了 50 粒小麦，则可以得到一个 50×4 的矩阵。

再如，区分不同类别的翡翠，可以将这些翡翠样品在给定的波数范围内测定其拉曼光谱，设每种类别测定 5 个样本，共 3 个类别，从 4001.05cm^{-1} 到 100.396cm^{-1} ，一个样本被测量了 2815 次，则可以得到一个 15×2815 的数据矩阵。

以上种类的数据，过去一直被称为 OV 型的数据，其中 O 代表 Object，V 代表 Variable，即一行一个样本，一列一个属性。

随着化学分析测量仪器的发展，现在通过化学分析仪器，往往可以得到更多的数据信息，仪器的联用就是很常见的获得海量数据的方式。例如色谱-DAD(二极管阵列检测器)就是色谱光谱联用的仪器，可以同时测得样本的色谱和光谱信息。其原理是，在色谱上流出物质的同时，用二极管阵列(每个二极管发出不同波长的光)在瞬间测得物质在指定波长下的吸光度。这样的数据特征是：一行数据就是在一个色谱保留时间下，在全波长扫描范围内的不同波长下的光吸收结果。把不同的保留时间下的吸光值按保留时间从小到大排列，就形成了一个矩阵。

这样的数据矩阵，它的行是光谱信息，列是色谱信息，都是量测特征，所以被称为高维数据。或者更典型地，被称为 VV 型的数据，即其每一维都是特征。

对于上述的 VV 型的数据，很明显，任意一行，都是一个光谱向量，即在某个保留时间下留出的物质的光谱。而每一列，就是一条色谱。典型的 VV 型测量数据的三维表达参见图 1.1，在 Excel 中，很容易实现这种显示效果。

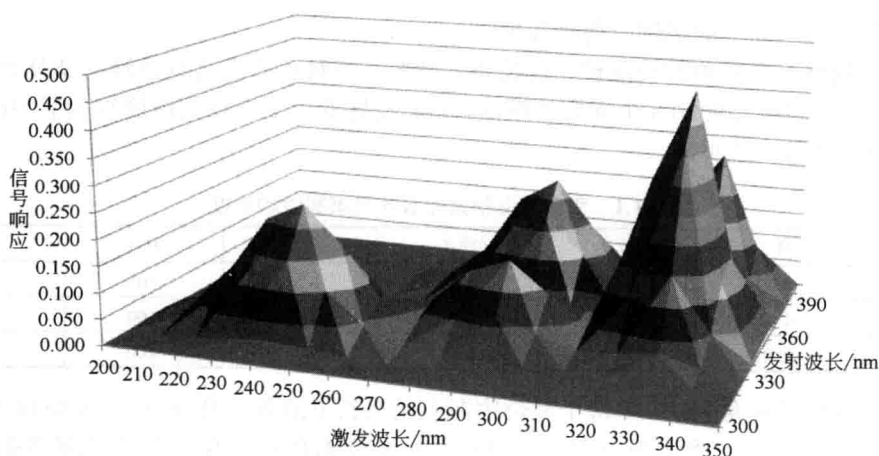


图 1.1 VV 型测量数据的三维展示样例

如果将多个样本都进行 VV 型的测量，并叠放在一起，就构成了三维的数据。

典型的 GC-MS/MS (气相色谱双质谱联用技术)，一个样本就得到一个三维矩阵。



1.2 向量、矩阵

为方便读者阅读起见,对本书矩阵和向量符号特别约定如下:

矩阵用大写黑斜体字母表示,例如 $\mathbf{X}$ 代表一个二维矩阵。

向量用小写黑斜体字母表示,例如 $\mathbf{h}$ 表示一个向量。且为公式表达方便,所有向量皆默认为列向量。

上标 T 或',表示矩阵或向量的转置。

1.2.1 向量

(1) 基本概念

① 定义 一组由 n 个数据构成的一行或一列数据。

一个由 n 个数据组成的向量称为 n 维向量,可以看作是 n 维空间中的一个点。

② 向量的模 向量所有元素平方和的开平方根,它是向量的长度。

③ 向量的标准化 将向量的每个元素除以它的模。很明显,向量标准化后,其模等于 1,称为单位向量。

(2) 向量运算

① 向量的内积 向量的内积又被称为点乘,它要求两个向量的长度一致,可以看作是一个行向量与一个列向量的相乘,得到一个数值。见图 1.2 (a)。

向量的内积常用来判别两条向量的相似性,例如两条光谱向量 $\mathbf{s}_1$ 和 $\mathbf{s}_2$,在两者都经过标准化后(模为 1),如果两条光谱是完全一样的,则内积为 1,如果完全不相似(即完全不交盖),则内积为零。

② 向量的外积 向量的外积并不要求两个向量的长度一样,它可以看作是一个列向量乘以一个行向量。

一个长度为 n 的向量和一个长度为 m 的向量的外积,得到一个 $n \times m$ 的矩阵。见图 1.2 (b)。

向量的外积在化学量测中也很常见,例如纯组分的色谱-DAD 测量,得到的二维矩阵,实际上是该物质的色谱向量和光谱向量的外积,该矩阵的秩为 1。从这点出发,通过判断秩,可以轻易判断 VV 型测量体系中的化学组分数。

③ 向量相加 两个等长向量对应元素相加,得到一个新向量。例如一个两组分混合物的光谱,本质上是每种组分的光谱向量的加和。

④ 向量相减 两个等长向量对应元素相减,得到一个新向量。

⑤ 向量乘以常数 向量的每个元素都乘该常数,得到一个新向量。例如一种纯组分的光谱,可由其单位浓度的光谱乘以浓度得到。

⑥ 向量加常数 向量的每个元素加上该常数,得到一个新向量。

⑦ 向量前乘矩阵 向量与矩阵相乘,向量在前,矩阵在后,得到的结果是一个向量。很明显,向量的长度与矩阵的行数相等。例如对于一个三组分混合物,将每种组分的纯光谱组成一个矩阵,将三种组分的浓度组成一个向量,则混合物光谱就是浓度向量与纯光谱矩阵的乘积。

⑧ 向量后乘矩阵 矩阵与向量相乘,矩阵在前,向量在后,得到的结果是一个向量,矩阵的列数要等于向量的长度。

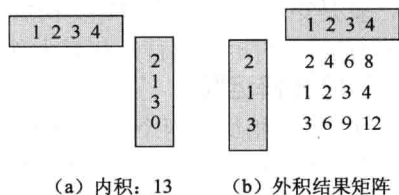


图 1.2 向量的内积和外积



1.2.2 矩阵

(1) 基本概念

① 定义 由 n 行 m 列的数据组成矩形数据表, 称为矩阵。元素为实数的矩阵称为实矩阵, 本课程中只讨论实矩阵。

② 行列 矩阵的行和列是向量, 对于 n 行 m 列的矩阵, 行向量的长度为 m , 列向量的长度为 n 。

③ 转置矩阵 矩阵的行变列, 列变行后, 元素 a_{ij} 转置后位于 a_{ji} 。矩阵 X 的转置用 X^T 表达。

④ 协方差矩阵 矩阵 X 的协方差矩阵, 表达为 $X^T X$, 它是一个对称方阵。

⑤ 特征值和特征向量 设矩阵 A 是 n 阶方阵, 如果存在常数 λ 和非 0 向量 x , 使得

$$Ax = \lambda x \quad (1.1)$$

则称 λ 是矩阵 A 的特征值(eigenvalue), 而 x 称为从属于 λ 的特征向量(eigenvector)。

式 (1.1) 可以改写为:

$$(A - \lambda I)x = 0 \quad (1.2)$$

由于 x 为非 0 向量, 所以行列式 $|A - \lambda I| = 0$, 它是 λ 的 n 次多项式方程, 所以方阵 A 会有多个特征值。

⑥ 矩阵的秩 矩阵的秩数, 就是矩阵的协方差矩阵的非零特征值的数目。从矩阵论上讲, 就是矩阵经不断的初等变换后, 矩阵中最后保留的有效行(行中有非 0 元素)的个数。对于 n 行 m 列的矩阵而言, 其最大秩为 n 和 m 中的小者。

(2) 矩阵运算

① 矩阵相加 两个尺寸相同的矩阵对应的元素相加, 得到一个新矩阵。

② 矩阵相减 两个尺寸相同的矩阵对应的元素相减, 得到一个新矩阵。

③ 矩阵转置 矩阵的第 i 行第 j 列的元素, 变成第 j 行第 i 列的元素, 一个 $n \times m$ 的矩阵转置后, 得到一个 $m \times n$ 的矩阵。

④ 矩阵相乘 矩阵相乘, 类似向量的内积, 可以看作是左边矩阵的第 i 行和右边矩阵的第 j 列的内积, 作为结果矩阵的第 i 行第 j 列的元素。从这点上看, 左边矩阵的列数必定等于右边的行数。这一点也是两个矩阵是否可以相乘的依据。

⑤ 矩阵的逆 很多情况下需要求解矩阵的逆, 只有方阵且满秩时矩阵才有逆。矩阵 X 的逆用 X^{-1} 表示, 矩阵和其逆的左乘和右乘都是单位矩阵 I 。即对角元素为 1, 其他的全部为 0。

⑥ 矩阵的伪逆 伪逆又称为非方阵的逆, 矩阵 X 的伪逆表示为 X^+ , 一个矩阵和其伪逆的乘积, 是其对角元上有数个 1 的元素, 其他元素则全部为 0。至于有几个 1, 则取决于 X 的秩的个数。

⑦ 矩阵的迹 方阵的迹定义为其对角元素的和。

⑧ 矩阵所有元素的平方和均值开根 将矩阵的每个元素平方, 然后加和, 求平均后, 开平方根, 该方法用于计算噪声的水平, 见后续的 PCA 算法解释。

⑨ 取矩阵的 1 行或 1 列, 得到一个向量。

⑩ 设置矩阵的 1 行或 1 列 用一个向量设置矩阵的某行或某列的值。

上述的矩阵、向量运算是后续很多化学计算方法的基础, 在进行程序设计时, 最好是将上述基础运算用库实现, Matlab 能给计算工作者带来巨大便利的原因也是其强大的矩阵计算功能。为此, 本教材中也将其类库化, 使后续算法的程序实现简洁化。



1.3 向量、矩阵类设计

1.3.1 向量类

向量的数据用一维数组就可以记录，因此可以围绕一维数组来实现，内容围绕前面定义的向量需要实现的计算来进行。根据对向量的定义和描述，考虑其属性和方法，可以定义如下的向量类：

```
public class Vector
{
    private int length;
    private double[] vec; //用来存储数据的一维数组
    public Vec(int n) //定义长度 n，构造一个数据元素都为 0 的向量
    {
        length = n;
        vec = new double[n];
    }
    public Vec(double[] data) //用给定的数组，构造一个向量
    {
        Length = data.GetLength();
        vec = data;
    }
    public int Length() //返回向量的长度
    {
        return length;
    }
    public double Model() // 计算向量的模
    {
        double sum=0;
        for (int i=0; i<length; i++)
        {
            sum += vec[i] * vec[i];
        }
        return Math.Sqrt(sum)
    }
}
```

以上给出了向量类的简单定义并实现了简单的求向量模的方法。向量类中还有很多方法需要编写程序实现，包括向量的其他运算，如内积、外积、向量与矩阵的运算等，有兴趣的读者可以逐一将其实现，提高自己基础编程能力。

1.3.2 矩阵类

矩阵的数据分行和列，因此可以用二维数组实现存储，并以其为核心实现矩阵的各种运算。类似向量类的设计，本节只对矩阵类的设计开个头，如构造函数、矩阵的乘法运算，代码如下：

```
public class Matrix
{
    private int row,col;
```



```
private double[,] matrix;
public Matrix(int row, int col)//构造元素都是 0 的矩阵
{
    this.row = row;
    this.col = col;
    matrix = new double[row,col];
}
public Matrix(double[,] data)//用二维数组构造矩阵
{
    this.row = data.GetLength(0);
    this.col = data.GetLength(1);
    matrix = data;
}
public static Matrix operator *(Matrix lhs, Matrix rhs) //定义矩阵的乘法运算符
{
    if (lhs.ColCount != rhs.RowCount)
    {
        throw new MatrixException(“相乘的两个矩阵维数不匹配。”);
    }
    Matrix rst = new Matrix(lhs.RowCount, rhs.ColCount); //乘积所得结果矩阵
    for (int i = 0; i < lhs.RowCount; i++)
    {
        for (int j = 0; j < rhs.ColCount; j++)
        {
            double sum = 0.0;
            for (int k = 0; k < lhs.ColCount; k++)
            {
                sum += lhs[i, k] * rhs[k, j];
            }
            rst[i, j] = sum;
        }
    }
    return rst;
}
```

上面定义的矩阵类中，只以 2 个矩阵相乘为例进行了编程。矩阵运算需要的各种方法较多，部分算法也较复杂，例如矩阵的逆和伪逆的求解，有兴趣的读者可以参照前面对矩阵的描述，查找相应的文献，用相应的方法实现。

1.4 矩阵类库及调用

同济大学化学系计算机与化学实验室用 C# 开发的矩阵运算库 **LinearAlgebra**，位于名字空间 **TJU.AC.LinearAlgebra** 中，在应用中通过不断改进优化而成，读者可以从 <http://cal.tongji.edu.cn/chem> 中下载。该类库已经实现了化学计量学诸多算法所需求的基本矩阵、向量运算，下面主要描述 **Matrix** 类和 **Vector** 类的主要方法。

(1) 矩阵类 **Matrix**

① 构造函数，**public Matrix(int row, int col)**，构造一个 row 行 col 列的矩阵，每个元素值



为 0。

例子:

`Matrix A = new Matrix(10,20);` 定义一个 10 行 20 列的矩阵, 每个元素的值为 0。

② 构造函数, `public Matrix(int row, int col, double value)`, 构造一个 row 行 col 列的矩阵且每个元素值都为 value。

例子:

`Matrix A = new Matrix(10,5,0.2);` 定义一个 10 行 5 列, 每个元素值都为 0.2 的矩阵。

③ 构造函数, `public Matrix(double[,] array2D)`, 以一个二维数组构造一个矩阵。

例子:

`double[,] arr2D = {{1,2,3},{4,5,6}};`

`Matrix A = new Matrix(arr2D);`

④ 构造函数, `public Matrix(Matrix matrix)`, 构造一个与 Matrix 一样的矩阵。但新矩阵和原矩阵引用的不是同一个数据块。

例子:

设已定义矩阵 A , 则可以定义 `Matrix B = new Matrix(A)`。

此时, B 矩阵数据的改变不影响 A 矩阵。

⑤ 返回矩阵的行数, `public int RowCount`。

例子:

`int row = A.RowCount;` 得到矩阵的行数。

⑥ 返回矩阵的列数, `public int ColCount`。

⑦ 返回矩阵中元素的个数, `public int Size`, 即矩阵的行数乘列数。

例子:

`int count = A.Size;`

⑧ 返回或设置矩阵第 row 行, 第 col 元素的值, `public double this[int row, int col]`。

例子

`A[0,2] = 3.2;` 将矩阵 A 的第 0 行第 2 列的元素值设置为 3.2。

⑨ 矩阵转置, `public Matrix Transpose()`。

例子:

`Matrix B = A.Transpose();` B 为 A 矩阵的转置。

⑩ 返回矩阵第 i 行, 结果是一个向量, `public Vector GetRow(int i)`。

例子:

`Vector v = A.GetRow(0);` 返回矩阵 A 的第 0 行。

⑪ 返回矩阵第 j 列, 结果是一个向量, `public Vector GetCol(int j)`。

⑫ 用向量 rowVector 设置矩阵第 i 行, `public void SetRow(int i, Vector rowVector)`。

例子:

设已定义了一个向量 v , 其元素的个数为矩阵 A 的列数, 则

`A.SetRow(0,v);`

将向量 v 放置在矩阵的第 0 行上。

⑬ 用向量 colVector 设置矩阵第 j 列, `public void SetCol(int j, Vector colVector)`。

⑭ 拷贝矩阵 Matrix 到当前矩阵, `public void Copy(Matrix matrix)`

例子:

`Matrix A;`

.....

`Matrix B = new Matrix();`