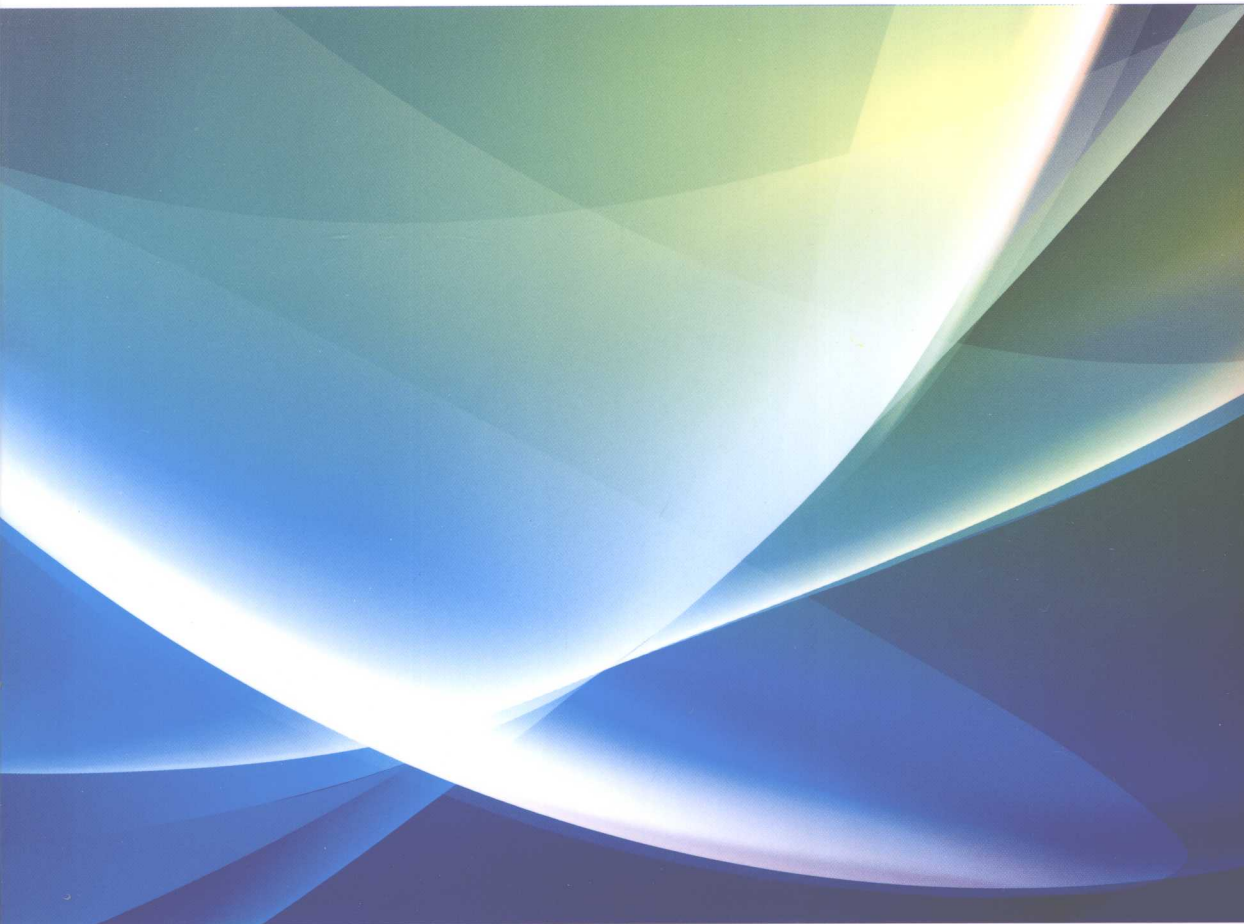


杜一平 潘铁英 张玉兰 编著

化学计量学应用



化学工业出版社

本书可作为高等院校化学、环境、生物、材料、医药、能源等专业的教材，也可供从事相关工作的工程技术人员参考。
责任编辑：张金全
封面设计：张金全
ISBN 978-7-122-03220-5
定价：38.00元

化学计量学应用

杜一平 潘铁英 张玉兰 编著

图书在版编目(CIP)数据

化学计量学应用 / 杜一平, 潘铁英, 张玉兰编著. — 北京: 化学工业出版社, 2008.8

ISBN 978-7-122-03220-5

I. 化… II. ①杜…②潘…③张… III. 化学计量学

V. 06-04

中国版本图书馆(CIP)数据核字(2008)第032200号

责任编辑：张金全
封面设计：张金全

出版发行：化学工业出版社
社址：北京

地址：北京市东城区青年湖南街13号 邮政编码：100011
印 刷：天津滨海新区泰达印刷有限公司
发 行：北京市丰台区

电话：010-63922688（总机），010-63922689（发行），010-63922688（传真），010-63922688（客服）

网址：<http://www.cip.com.cn>
编辑部：北京青年湖南街13号 印刷部：天津滨海新区泰达印刷有限公司

定价：38.00元



化学工业出版社

定价：38.00元

· 北京 ·

本书首先简要介绍了化学计量学的发展历史、研究内容和应用领域,数学和统计学基本知识,计算机编程语言 MATLAB,分析化学中数理统计和数据处理方法。在此基础上,介绍了目前应用较多的一些化学计量学方法,包括实验设计、最优化方法、多元校正、多元分辨、化学模式识别、化学构效关系等。在阐述清楚各种方法基本理论的前提下,淡化数学理论,重点讲述化学意义,突出算法的使用,并对重要方法给出应用实例。

本书可以作为分析化学和其他相关专业本科生和研究生化学计量学课程的教学用书,也可以作为从事化学计量学研究和应用人员的参考书。

化学计量学应用

图书在版编目 (CIP) 数据

化学计量学应用/杜一平,潘铁英,张玉兰编著. —北京:化学工业出版社,2008.3
ISBN 978-7-122-02259-2

I. 化… II. ①杜…②潘…③张… III. 化学计量学
IV. 06-04

中国版本图书馆 CIP 数据核字 (2008) 第 025890 号

责任编辑:任惠敏
责任校对:吴静

文字编辑:林媛
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:大厂聚鑫印刷有限责任公司

装订:三河市宇新装订厂

720mm×1000mm 1/16 印张17½ 字数299千字 2008年5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: 38.00 元

版权所有 违者必究

前言

化学计量学是一门运用数学、统计学、计算机科学以及其他相关学科的理论与方法,优化化学量测过程,并从化学量测数据中最大限度地获取有用化学信息的科学。伴随着分析化学,尤其是仪器分析方法的迅速发展,化学计量学从诞生到逐步发展,引进、提出和发展了很多算法和方法,解决了分析化学中很多重要问题。其方法贯穿了分析量测中“采样—测量—数据处理”的各个部分,如采样理论、实验设计、最优化、信号处理、多变量校正、多变量分辨、化学模式识别、化学构效关系研究等,为分析化学的发展做出了重要贡献。近年来,随着各类分析仪器自身的发展,仪器所能提供数据量迅速增加,为以“从化学量测数据中最大限度地获取有用化学信息”为目标的化学计量学提供了广阔的发展空间,使分析化学家们可以更多、更好、更有效地利用量测数据来探索未知世界,获得更多重要信息,解决传统方法不易或不能解决的问题。另一方面,“获取有用化学信息”不只是分析化学的要求,也是其他化学以及化学相关学科的要求,因此近一二十年以来,化学计量学在有机化学、物理化学、高分子化学、农业科学、医药、环境、材料等领域的应用研究越来越广泛和深入,展现了化学计量学非常独特的魅力。

本书收集了比较重要的、目前应用较多的一些化学计量学方法,内容涉及实验设计、最优化方法、多元校正、多元分辨、化学模式识别、化学构效关系等。在各类方法选择上,重点介绍比较流行的现代方法,但为了保持内容的延续性,也介绍了基本概念、基础理论和一些传统方法。对每种方法的讲解重点强调方法的化学意义,用较多篇幅介绍了各种重要方法的应用,这对于有化学背景的读者来说是非常必要的。不过,化学计量学毕竟以数学和统计学方法和理论为基础,以计算机为研究工具,必要的数学和统计学知识对理解化学计量学方法大有益处,因此,本书专门开辟章节来介绍数学和统计学基本知识。同时也简要介绍了化学计量学常用的计算机编程语言 MATLAB,以期使读者能达到入门水平,为进一步学习 MATLAB,并利用其编

写化学计量学算法程序打下基础。

书中大多数应用实例取材于作者近年的科研工作，少部分来源于文献。本书由杜一平编写，潘铁英和张玉兰参加了部分章节的文字修改工作，并提供了部分应用实例的数据。本书编写过程中得到了恩师梁逸曾教授的鼓励和大力支持，也获得了华东理工大学“十一五”规划教材出版基金的资助，在此表示诚挚的感谢。

作者对化学计量学的理解有限，书中难免有错误和不当之处，敬请读者批评指正。

作者

2007年12月于上海

目 录

第 1 章 概论	1
1.1 化学计量学的诞生和发展历史	1
1.2 化学计量学的研究内容	3
1.3 化学计量学的应用领域	4
参考文献	9
第 2 章 数学和 MATLAB 编程基础	10
2.1 数学和统计学基础	10
2.1.1 矢量和矩阵运算	11
2.1.2 线性相关	19
2.1.3 范数	21
2.1.4 统计学基础	22
2.2 MATLAB 编程基础	30
2.2.1 MATLAB 基本命令和基本运算	31
2.2.2 MATLAB 的基本绘图功能	43
2.2.3 M 文件	47
2.2.4 MATLAB 编程举例	50
参考文献	54
第 3 章 分析化学中数理统计和数据处理方法	55
3.1 分析数据的数理统计方法	55
3.1.1 分析化学中数理统计的基本概念	55
3.1.2 分析数据的置信区间估计	57
3.1.3 分析数据的显著性检验	60

3.2 分析信号的平滑	63
3.2.1 平均法	64
3.2.2 多项式拟合法	65
3.3 分析信号的求导	71
3.3.1 差分法	71
3.3.2 多项式拟合法	72
参考文献	80

第4章 化学实验设计 81

4.1 基本概念	82
4.2 拉丁方和拉丁方设计	84
4.3 因子设计和部分因子设计	85
4.3.1 因子设计	85
4.3.2 部分因子设计	89
4.4 正交实验设计	91
4.4.1 正交表和正交实验设计	91
4.4.2 用正交表安排实验	94
4.4.3 实验结果的分析	95
4.5 均匀实验设计	97
4.5.1 均匀设计表	98
4.5.2 用均匀设计表安排实验	102
4.5.3 拟水平均匀设计	102
参考文献	104

第5章 最优化方法 105

5.1 最优化的基本概念	105
5.2 线性规划	108
5.2.1 线性规划的标准型	108
5.2.2 线性规划的图解法	109
5.2.3 单纯形优化法	110
5.3 非线性规划	111
5.3.1 消去法	111

5.3.2	变量轮换法	112
5.3.3	一阶梯度法	113
5.4	模拟退火算法	114
5.5	遗传算法	115
5.5.1	遗传算法的基本步骤	116
5.5.2	遗传算法的计算过程	117
5.5.3	遗传算法的实例	119
5.6	人工神经网络	121
5.6.1	生物神经网络	122
5.6.2	人工神经网络	123
5.6.3	反向误差传播人工神经网络	125
5.6.4	人工神经网络应用实例	130
	参考文献	132

第6章 多元校正 133

6.1	黑白灰分析体系	135
6.2	多元线性回归方法	138
6.2.1	相关与回归	138
6.2.2	一元线性回归方法	140
6.2.3	多元线性回归方法	143
6.2.4	多元线性回归解的统计性质	144
6.3	K -矩阵法和 P -矩阵法	145
6.3.1	K -矩阵法	146
6.3.2	P -矩阵法	147
6.4	主成分回归法	148
6.4.1	主成分分析	148
6.4.2	特征值和特征向量的计算方法	154
6.4.3	主成分数的确定	156
6.4.4	主成分回归	157
6.5	偏最小二乘法	159
6.5.1	偏最小二乘法基本原理	159
6.5.2	偏最小二乘法算法	161
6.5.3	隐变量数目确定方法	163

6.6	多元校正应用实例	165
6.6.1	用多元线性回归法进行多组分同时光谱定量分析	165
6.6.2	主成分回归法用于多组分近红外光谱定量分析	167
6.6.3	用偏最小二乘法进行近红外光谱定量分析	170
	参考文献	171

第7章 多元分辨

7.1	多元分辨基本原理	173
7.2	渐进因子分析法	175
7.3	固定尺寸移动窗口渐进因子分析法	179
7.4	交替最小二乘法	181
7.5	直观推导式演进特征投影法	182
7.5.1	色谱二维数据的特点	183
7.5.2	演进特征投影图	183
7.5.3	秩图	186
7.5.4	背景校正	187
7.5.5	满秩分辨	190
7.6	直观推导式演进特征投影法应用实例	193
	参考文献	198

第8章 化学模式识别

8.1	模式识别的基本概念	199
8.2	模式识别中数据预处理方法	201
8.3	特征抽取	202
8.4	空间中距离和相似性的表达方法	205
8.4.1	空间中距离的表达方法	205
8.4.2	空间中相似性的表达方法	206
8.5	有监督的模式识别方法	207
8.5.1	距离判别法	207
8.5.2	Fisher 判别法	208
8.5.3	K 最近邻法	210
8.6	无监督的模式识别方法	212

8.6.1	系统聚类法及实例	212
8.6.2	最小生成树方法及实例	218
8.7	模式识别的显示方法	219
8.7.1	主成分分析投影显示法及实例	220
8.7.2	偏最小二乘特征投影法	221
8.7.3	非线性投影方法	222
	参考文献	223
第9章 定量构效关系		224
9.1	化学结构的表征	225
9.1.1	图论的基本概念	226
9.1.2	化学图	227
9.1.3	邻接矩阵和距离矩阵	228
9.2	拓扑指数	230
9.2.1	Wiener 指数	230
9.2.2	Hosoya 指数 Z	231
9.2.3	分子拓扑指数	232
9.2.4	分子连接性指数	232
9.2.5	含不饱和键和杂原子化合物的分子连接性指数	239
9.3	拓扑指数与化合物性质和生物活性之间的关系	241
9.3.1	分子拓扑指数与化合物沸点之间的关系	241
9.3.2	拓扑指数与色谱保留指数之间的关系	243
9.3.3	拓扑指数与生物活性之间的关系	251
9.4	定量构效关系的建模方法	252
9.4.1	变量选择	252
9.4.2	定量构效关系中多元线性回归建模方法及应用实例	255
9.4.3	定量构效关系中的人工神经网络方法及应用实例	265
9.4.4	定量构效关系中的模式识别方法	266
	参考文献	267

第1章 概论

1.1 化学计量学的诞生和发展历史

1971年,瑞典化学家 S. Wold 在申请一项基金项目时首次提出“化学计量学”(Chemometrics)这一名词。他参考生物计量学(Biometrics)和经济计量学(Econometrics)两个名词的词尾,与化学(Chemistry)的词头结合定义了 Chemometrics。Wold 提出的化学计量学得到美国科学家 Kowalski 的赞赏,他们于 1974 年共同发起成立了国际化学计量学学会,之后,化学计量学得到了快速发展。在化学计量学发展过程中,尤其是发展初期,分析化学工作者做出了重要贡献,同时这门分支学科也对分析化学的发展产生了深远的影响。所以,很多人认为化学计量学是隶属于分析化学学科的。但是,今天的化学计量学所研究的内容已经不仅仅局限于分析化学,它在整个化学学科及其相关学科,如生命科学、环境科学等领域,以及很多工业部门都发挥着重要作用。

按照国际化学计量学学会的定义,化学计量学是一门运用数学、统计学、计算机科学以及其他相关学科的理论与方法,优化化学量测过程,并从化学量测数据中最大限度地获取有用化学信息的学科。分析化学中的数理统计方法,如标准偏差、置信区间、有效数字、显著性检验等,就是早期化学计量学研究的内容。那时,Chemometrics 一词尚未出现,许多统计学与数学方法在化学与分析化学中逐步得到了应用。这一阶段的化学计量学只是简单地把统计学和数学方法应用到分析化学中,范围比较有限,也没有自己独特的方法。从 20 世纪 70 年代开始,科学家们除了使用统计学和数学方法外,还吸收了如行为科学(心理学等)、经济计量学、信息科学等领域已发展起来的数据与信号分析方法,将它们成功地应用于分析化学研究。

众所周知,20 世纪 50 年代以后,仪器分析方法迅速崛起,各类分析仪器所能提供的数据容量迅速增加。一般的化学分析方法,一次实验往往只能得到有限的几个数据,如滴定分析的实验数据无非是消耗滴定剂体积,重量分析中沉淀物的重量等。而仪器分析方法的一个实验往往可以得到一组数据,

如一个色谱图、一个极谱图、一个光谱图等，它们都包含一系列数据。不过，分析化学中传统的数据处理方法一般只是利用了一部分数据，如紫外光谱定量分析中只是使用最大吸收波长处的吸光度数据，色谱分析也是利用某个或某几个保留时间，以及它们的峰高或峰面积数据。由于分析数据容量的迅速增加让分析化学家们去探索如何更多、更好、更有效地充分利用量测数据。化学计量学针对这样的数据，提出了一些方法能从中获取更多的有用信息，为分析化学展现了一个非常独特的美好前景。

例如利用多个光谱点可以进行多组分同时测定。设一个光谱包含 n 个波长点 $\lambda_i (i=1, 2, \dots, n)$ ，如果样品中含有 m 个组分，则可以从 n 个波长中选择 m 个（假设 $n > m$ ，一般情况下都能满足这个条件），按 Beer 定律列出方程组，

$$A_1 = k_{11}c_1 + k_{12}c_2 + \dots + k_{1m}c_m$$

$$A_2 = k_{21}c_1 + k_{22}c_2 + \dots + k_{2m}c_m$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$A_m = k_{m1}c_1 + k_{m2}c_2 + \dots + k_{mm}c_m$$

式中， A_i 为对应波长 λ_i 下样品的吸光度； c_1, \dots, c_m 为第 1 到第 m 个组分的浓度； k_{ij} 是第 j 个组分在第 i 个波长下的吸光系数，为常数。很显然，如果利用各个组分的标准溶液测定它们在各个波长下的吸光系数 k_{ij} ，并测定未知混合样品的吸光度，通过解上述方程组，很容易计算出 m 个组分的浓度。这种分析方法，在吸光系数已知的情况下，对未知样品进行一次测量就可以进行多组分含量的同时测定，显然测试效率得到了提高。不过，该方法不能保证每个组分都是在最佳条件下（如最大吸收波长下）进行测定，而且波长的选择也是要考虑的问题。上述方法只是最简单的多组分同时测定方法，实际上化学计量学方法利用全部波长的数据，用最小二乘方法来计算各组分含量，准确度更高。

类似上述多组分同时测定方法非常新颖，给分析化学界带来了新气象。这个时代的化学计量学已经具有了一定的规模，发展了一些独特的方法，能够解决一些用传统分析化学方法无法解决的问题。

到了 20 世纪 90 年代，随着分析化学的进一步发展，大量现代分析仪器，尤其是联用分析仪器（如色谱-质谱联用、色谱-红外联用、色谱-紫外联用等）的广泛使用，产生大量化学量测数据，为了更好地利用这些数据，需要合适的方法来处理数据，这就更加促进了化学计量学的发展和应用。化学分析方法的分析过程只有几个数据，一个仪器分析产生的谱是一组数据，而联用分析仪器分析产生的数据是很多组数据，数据容量大大增加了。比如，

一台气相色谱-质谱联用仪器,一次分析产生的数据可以达到几兆(M),甚至几十兆、上百兆。如此大量的数据中必然含有大量有用信息,如何从中提取有用信息,传统分析化学方法很难解决这类问题。化学计量学作为一门新兴学科,其目的就是解决从化学数据中提取有用信息的问题。化学计量学解决问题的基本策略是:使用数学和统计学方法和理论,利用计算机技术为手段,实现化学数据的解析。

经过几十年的发展,化学计量学已经成为一门比较成熟的学科,全球范围内有一支稳定的队伍从事该方向的研究和应用工作。值得一提的是,我国学者也为化学计量学的发展做出了重要贡献,目前在国际化学计量学界也活跃着一批来自中国的研究小组,比如湖南大学、中南大学、同济大学、中国科技大学、南开大学、西北大学、四川大学、首都师范大学、中国科学院长春应用化学研究所、石油化工科学研究院等单位都有很多研究人员从事化学计量学研究。

1.2 化学计量学的研究内容

化学计量学研究着眼于发展化学数据解析的新理论和新方法,同时也注重在相关学科中的应用研究。很多化学计量学方法来源于其他学科,吸取其他学科的精华,并结合化学数据的特点,来发展适合于自己特点的方法。如统计学、数学和信息科学的很多方法,被引入化学计量学中。数理统计方法、概率方法、最优化方法、图论方法、稳健算法、回归分析、试验设计方法等都来源于数学和统计学;而人工神经网络方法、模拟退火算法、遗传算法、小波分析、图像分析方法等,都是从信息科学等领域引进的。这些方法经过改造和发展,解决了很多化学实际问题,引起了化学家的极大重视。还有很多化学计量学方法,就是化学计量学家提出的。这些方法往往起源于化学问题(很多是分析化学问题),具有强烈的化学特征。在多元校正、多元分辨、化学模式识别、定量构效关系研究等方面,就有很多这样的方法,如偏最小二乘法、SIMICA方法、秩消失因子分析法、渐近因子分析方法、固定尺寸移动窗口渐近因子分析方法、直观推导式演进特征投影法等。

化学计量学的研究内容已经贯穿了分析化学量测的全过程。一个分析任务一般从采样开始,然后是样品前处理、制样,随后进行测试,测试手段往往是分析仪器,最后的步骤就是对测试产生的数据进行数据处理,获得分析量测所要获得的信息。针对上述“采样—测量—数据处理”这一分析化学量测过程,化学计量学发展了相应的各类方法,如采样理论、实验设计、最优

化、信号处理、多变量校正、多变量分辨、化学模式识别、化学构效关系研究等。其中,采样理论是根据统计学理论,经过发展而建立的理论和方法,用来指导试样的科学采集;实验设计方法已经在很多领域得到应用,在分析测试过程中,可以用来进行实验前处理中某些实验的科学设计、仪器分析操作条件选择等;最优化方法提供一些算法,可以从复杂的实验数据中选择最佳的条件等参数;信号处理方法可以在测量过程中,对数据进行处理,以消除噪声、干扰等;多变量校正和多变量分辨是非常重要的化学计量学方法,前者是用多通道数据(传统方法往往只用一个通道的数据,如紫外分光光度法只使用最大吸收波长数据进行定量分析等)进行定量分析,后者是针对分析中经常遇到的组分之间的干扰而提出的,它可以利用多通道的数据分别计算出复杂混合体系中各种组分的浓度等有用信息;化学模式识别是一种有用化学信息的提取方法,它能从复杂的数据中发现数据之间的内部联系、隐含的规律等,为研究者做出决定提供依据;化学构效关系研究通过建立化合物结构数据与化合物的药效或性质等之间的关系,来研究和预测新化合物的药效或性质,该方法在药物设计、分子设计等领域应用非常广泛。

如图 1-1 所示,化学计量学的基本内容包括采样理论和方法、化学中实验设计方法、化学最优化方法、分析检测的信号处理方法、多变量校正方法、多变量分辨方法、化学模式识别、化学构效关系研究、化学专家系统和计算机模拟方法等。本书将重点介绍实验设计、最优化方法、信号处理、多变量校正、多变量分辨、化学模式识别和定量构效关系研究等内容。

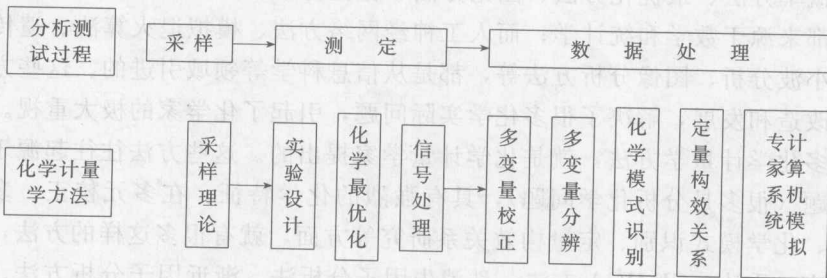


图 1-1 化学计量学研究内容

1.3 化学计量学的应用领域

化学计量学的内容已经包含了分析测试的各个部分,在分析化学中已经广泛应用了化学计量学各种方法。在化学各个分支学科,如有机化学、物理

化学、高分子化学等,以及化学的相关学科,如生命科学、环境科学、食品科学、农业科学、医药科学等的科学研究中,化学计量学的应用也取得很多重要成果,应用面越来越广泛。实际上,在各个学科的科学实验中必定要获得各种数据,一般情况下,数据量越大,所含的信息应该越多,从大量实验数据中获取更多有用信息,是研究者共同的心愿。化学计量学的任务就是从各类实验产生的大量数据中最大限度地提取有用信息,所以,以解析数据以获得最有价值信息为己任的化学计量学,在很多学科中都具有美好的应用前景。

下面将主要介绍化学计量学与分析化学之间的特殊关系,以及目前化学计量学在分析化学中比较活跃的研究和应用领域,以展现其独特魅力。

(1) 化学计量学是化学量测的基础理论与方法学

分析化学是一门历史悠久的学科,经历了三次巨大的变革,已经形成了自己特定的研究领域和研究方法。但是,分析化学学科还没有自己完善的科学理论,现在在分析化学学科各种书籍中所讲述的理论,实际上都是来源于其他学科的理论。比如,组成化学分析方法的四大滴定方法(酸碱滴定法、络合滴定法、氧化还原滴定法和沉淀滴定法)的理论是取自溶液化学的四大平衡理论,即酸碱电离平衡、配位-离解平衡、氧化-还原平衡和沉淀溶解平衡;光谱分析方法的理论都是从物理学的光学和光谱学理论引进的;电化学分析理论也是根据溶液电化学理论发展而来的;色谱分析的理论则是物质相互作用理论演变而来。这些理论实际上是各类分析方法的理论,并非是分析化学本身的理论。

什么才应该是分析化学的理论呢?按照 Valcarcel 对分析化学的定义,分析化学的任务是发展、优化、应用量测过程,以获取全局或局部性的化学品质信息,解决所提出的量测课题。Valcarcel 的定义实际就是把分析化学定义为量测的科学,其研究的内容是发展、优化、应用量测过程,目的就是获取被测物的有用信息。而化学计量学的定义是运用数学、统计学、计算机科学以及其他相关学科的理论与方法,优化化学量测过程,并从化学量测数据中最大限度地获取有用的化学信息的学科。从上述的描述中可以清楚地看出化学计量学与作为化学量测科学的分析化学之间的特殊关系:化学计量学就是化学量测的基础理论与方法学。化学计量学很多研究内容,都涉及分析化学的基础性问题,如样品的采样理论、分析方法的灵敏度、检出限等。因此,化学计量学的一个基本任务就是发展分析化学量测的基本理论和方法学体系。

(2) 化学计量学在近红外光谱分析中的应用

近红外光谱分析是近红外光谱技术与计算机和化学计量学有机结合而发

展起来的,是近几十年来发展最快、最引人注目的光谱分析技术。与其他光谱分析方法相比,近红外光谱分析有其独特之处,其主要特点有:①分析速度快;②多组分同时测定;③样品不需预处理;④非破坏性分析;⑤远距离测定和实时分析;⑥测定重现性好;⑦低分析成本和操作简单等。它的应用涉及石油化工、基本有机化工、精细化工、制药、农业、食品、质量监督、烟草、环境检测等很多领域。近一二十年来近红外光谱分析技术的优势越来越明显,应用越来越广泛,尤其是无损快速分析检测中优势更加突出。近红外光谱技术有别于其他常用分析技术的最大优点是,其操作非常简便,无需制样,不消耗试剂、不会造成环境污染,分析费用低、响应速度快,可以真正做到不损害样品。比如,用近红外光谱对农产品进行品质鉴别,可以用光纤探头直接插入粮食等产品中,测定其水分、淀粉、蛋白等成分的含量;也可以用光纤探头直接接触到水果表面来实现对其成熟度、酸度、糖度等的现场测定,以决定是否采摘;在工业生产中,可以用近红外光谱分析技术进行原材料的快速检测,对生产过程进行实时检测等。

近红外光谱属于振动光谱,分析化学中广泛用于结构测定的红外光谱也是振动光谱,属于中红外光谱。中红外光谱来源于多种化学基团的基频振动,而近红外光谱主要是由含 $X-H$ ($X=C, N, O, S$)基团的分子在近红外光谱区域的倍频和合频吸收,波长范围为 $800\sim 2500\text{nm}$ 。有机化合物中 $X-H$ 基团非常普遍,一个 $X-H$ 基团往往不止出现在一个分子上,一个有机物分子一般情况下也不止一个 $X-H$ 基团。因此不容易用某一特征峰进行近红外光谱分析,也就是说,近红外光谱的基团含量信息与其待测组分含量和性质指标之间难以用一简单数学关系式来定量描述。另外,近红外光谱为某些基团的倍频和合频吸收,吸收信号的信噪比往往比较低,从而导致分析灵敏度较低。这一点正是近红外光谱分析不需要样品前处理,可直接进行无损分析的原因。但另一方面,因为信噪比比较低,也导致分析灵敏度较低,某些低含量组分分析的准确度不高。为了尽量多地利用近红外光谱中 useful 信息对样品进行分析,化学计量学被应用到近红外光谱分析中解决近红外光谱分析中的关键问题。

化学计量学在近红外光谱分析中的应用主要包括两个方面内容。第一个方面是光谱预处理。光谱预处理包括选择合适的分析波长区间,对光谱进行校正,去除背景和共存组分的干扰等,从而提高信噪比。目前经常使用的方法有光谱平滑、求导、多元信号校正(Multiple Signal Correction, MSC)、正交信号校正(Orthogonal Signal Correction, OSC)等。第二个方面的应用就是建模。由于近红外光谱信号较弱,光谱干扰严重,很难像其他光谱定

量分析方法那样进行分析测定,必须充分利用各个波长范围的有用信号进行分析。为了达到这个目的,近红外光谱分析发展了利用数学模型来进行定性和定量分析的方法。通过大量样品的被测组分和近红外光谱的测定,用化学计量学方法建立数学模型(相当于传统定量分析的标准曲线),再利用模型对未知样品进行定性和定量分析。应用最广泛的化学计量学建模方法是偏最小二乘方法(Partial Least Squares, PLS),此外还有主成分回归(Principal Component Regression, PCR)、多元线性回归(Multiple Linear Regression, MLR)、人工神经网络(Artificial Neural Network, ANN)等。由于化学计量学的应用,使得近红外光谱分析方法焕然一新,在实际应用中发挥了很重要的作用,可以说,没有化学计量学的应用就没有现代近红外光谱分析的今天。

(3) 多元分辨方法解决复杂组分色谱重叠峰的分峰问题

多元分辨是化学计量学中的一种重要方法,它在色谱重叠峰的分峰上得到了很好的应用,这是化学计量学方法解决分析难题的一个成功例子。

色谱分析是分析化学中一类非常重要的仪器分析方法,它利用不同组分在固定相和流动相之间具有不同的保留性质这一特点,对多组分混合物样品进行分离分析。将样品中各种组分清晰地分离开来是色谱分析追求的目标,但对于复杂的样品,如天然产物的提取物、原油样品、复杂的工业产品等,色谱分离往往不能很好地分离。所得到的色谱图中,经常看到未分离完全的很多组分聚集在一起形成的“宽峰”,而且看起来一个“正常”的色谱峰也有可能是两个或更多组分混杂在一起的非纯峰,这些峰常称为重叠峰。如何鉴别一个色谱峰是否纯峰,对非纯峰如何进行适当处理以得到各个组分的纯色谱峰(即分峰),都是色谱研究中的重要课题。分峰是色谱分析的一个难题,一般做法是选择更理想的色谱柱、优化色谱分离条件等,从色谱分析的“硬件”方面寻求解决方法。“软件”方面的一个解决方法是对色谱峰进行拟合分峰,即利用高斯函数或洛仑兹函数,或它们的组合函数对多组分色谱进行拟合,使拟合后所得各个组分的色谱图的加合与混合物的非纯峰最大程度地“吻合”。这种方法虽然简单,但是由于色谱图并不完全符合高斯函数或洛仑兹函数,拟合时组分数也没有好方法来确定,所以它的准确度往往不高。化学计量学中的多元分辨也是一种“软件”方法,它是利用了色谱联用仪器产生的二维数据,并结合色谱分离的特点建立的方法,是一种建立在严格数学理论基础上的科学方法。

色谱联用仪器,如液相色谱与二极管阵列联用仪(HPLC-DAD)、气相色谱与质谱联用仪(GC-MS)、气相色谱与红外光谱联用仪(GC-IR)等产