

化学因子分析

潘忠孝 司圣柱 聂圣哲 张懋森 编著

中国科学技术大学出版社

化 学 因 子 分 析

潘忠孝 司圣柱 聂圣哲 张懋森 编著

中国科学技术大学出版社

1992 年 · 合肥

(皖)新登字 08 号

化 学 因 子 分 析

潘忠孝 司圣柱 编著
聂圣哲 张懋森

★

中国科学技术大学出版社出版

(安徽省合肥市金寨路 96 号, 230026)

中国科学技术大学印刷厂印刷

安徽省新华书店发行

★

开本: 850×1168/32 印张: 11.25 字数: 300 千

1993 年 1 月第 1 版 1993 年 1 月第 1 次印刷

印数: 1—3000 册

ISBN7-312-00363-X/O·117 定价: 8.00 元

(凡购买中国科大出版社图书, 如有白页、缺页、倒页者, 由本社发行部负责调换)

内 容 简 介

因子分析技术已被实践证明是计量化学中最具潜力的技术之一。除了传统的抽象因子分析方法之外,本书还较全面地介绍了目标变换因子分析、渐进因子分析、秩消因子分析和对应因子分析等几种适合于处理化学问题的因子分析技术。这些技术对吸收光谱、发射光谱、色谱、质谱、核磁共振谱和电子能谱等谱学数据进行解析、研究的原理和收获以及这些技术应用于各种多元体系的定性定量解析和应用于化学的其它基础理论研究的内容,本书也做了较详尽的介绍。

本书可供从事计量化学、分析化学、仪器分析、物理化学、有机化学、物质结构、表面物理和表面化学、环境化学、生物化学、药物学、食品加工和地球化学的教学和科研人员参考。

PREFACE

Chemometrics is the art of extracting meaningful information from chemical data by the combination of mathematics, statistics and computers. Factor analysis, a strategy for deciphering matrices of multivariate data, is one of the most powerful techniques in the chemometric arsenal. Although the method has been used in the behavioral sciences since the early 1930's, it went unnoticed by the chemical profession until 1970, even though the methodology was ideally suited to solving chemical problems.

I became interested in factor analysis after reading a chapter in "Applied Group Theoretical and Matrix Methods", a book written by Higman, published by the Oxford University Press in 1955, and recommended to me by my thesis advisor, Prof. Luigi Pollara. This chapter concerned linear free energy, historically representing the first application of factor analysis to a chemical problem. In 1961 I defended my doctoral dissertation, part of which was devoted to the development and applications of target factor analysis to chemistry. The laborious calculations were carried out with a Marchant hand-operated electric calculator, a prototype of which is on display in the Science Museum in London, England.

My first publication on factor analysis, "Application of Factor Analysis to the Prediction of Activity Coefficients of Non-Electrolytes", appeared in *Separation Science*, **1**, 661(1966), representing a joint effort with Phil Funke, Dan Matire and Lui Pollara, my colleagues at Steven Institute. This work required the use of an IBM 1620 Computer, which often devoured the computer cards.

In 1975 I began to compile a list of publications from the chemical literature that could, in some sense, be classified as involving factor analytical principles. In collaboration with Prof. Howery of Brooklyn College and with helpful suggestions and contributions from Paul Weiner, this effort produced a research text, "Factor Analysis in Chemistry", which was published by Wiley Interscience in 1980.

Thus the seeds of factor analysis were planted in the field of chemistry. The story that followed is akin to "soft" drama, something profound taking place so quietly and in such small steps that it is hardly noticeable, as opposed to "hard" drama which is dramatic (like superconductivity or cold fusion) appearing overnight in public newspapers. The soft drama of factor analysis in chemistry has slowly but steadily unfolded. The seeds of the early investigations are just beginning to blossom and bear fruit. So many new, inventive and ingenious developments and applications of factor analysis have recently appeared in the chemical literature that it is impossible to keep abreast of the field.

The seeds have now flown to China. Those capable of reading English are able to digest the wealth of information, the others must struggle for the knowledge. The efforts of researchers, such as Prof. Pan Zhongxiao and his colleagues, in compiling, translating and writing versions in the Chinese language is needed to nurture and feed Chinese scientists, including geologists, pharmacologists, biochemists, food specialists, chemical engineers, statisticians, *etc.*, as well as chemists. Such a fertilizing rain of information will undoubtedly produce new flowers in the field of factor analysis.

August 1991

Edmund R. Malinowski

前 言

经广大化学和统计学研究者近 30 年的不懈努力, 因子分析技术在化学理论研究实践中得到了很大的丰富和发展. 尽管因子分析模型和理论在数学上尚待完善, 但它的用途已为愈来愈多的化学研究实践所证实. 我国化学工作者中, 熟悉计算机的越来越多, 国内化学实验室中配备计算机的也越来越多, 同时, 想了解一下到底什么是因子分析以及如何采用因子分析技术来进行化学理论和应用研究的人员也越来越多. 为了在我国化学界大力传播因子分析这一强有力的计量化学手段, 为了在这一传播过程中尽我们的一分微薄的力量和热心, 在借鉴大量国内和国外前辈和同行卓有成效的研究工作的基础上, 结合自己的体会, 我们编写了这本书.

本书较全面和系统地介绍了化学界业经多年的努力已形成颇具化学特色的几种因子分析方法及其应用. 对于每一种方法, 我们只对与其有关的数学原理进行较简单扼要的介绍, 目的是想让读者在对方法的原理本身有一个基本了解的前提下不受过多的数学繁琐推导和证明所纠缠. 同时, 还尽量列举每种方法的化学研究应用实例或数据实例, 希望借此可加深读者对所学方法的了解. 对于因子分析的误差理论, 因子分析在组分分析中的应用和因子分析在化学基础研究中的应用等综合性内容则分别归纳成单独章节来介绍.

全书分 10 章. 第一章对计量化学和因子分析做一个简单的介绍; 第二章介绍抽象因子分析方法, 该方法是所有其它因子分析方法的基础; 第三章介绍由 E.R. Malinowski 在 60 年代发展起来的目标因子分析法, 该法在化学问题的研究中具有独特的优点, 它可在不管其它共存因子的情况下, 通过对一个目标因子进行检验而加以

鉴别；第四章介绍由 C.N. Ho 等提出的秩消因子分析法，该法特别适用于处理二维数据信息，能帮助研究人员在不考虑样品中自己不感兴趣的其它组分的情况下能获得所要分析的已知组分的定量信息；第五章介绍一类总称为渐进因子分析的方法，这类方法属于一种无模型技术，可以在不借助任何化学理论模型的情况下，对所含组分不明、组分含量及其组分数目未知的混合物进行分析；第六章介绍对应因子分析法，该法可有效地描述样品之间、变量之间以及样品和变量之间的双重关系，进而可对所研究的对象进行有效的判别和分类；第七章系统地介绍了因子分析的误差理论及其应用，这是因子分析的一个重要内容，如时间允许应争取对它们有一个全面的了解；第八章至第十章则综合介绍因子分析在化学中的各种应用情况，相信这些内容能给化学工作者许多有益的启发。

在编写本书的过程中，我们大量地借鉴了国内外老前辈和同行们的研究成果。著名的化学因子分析专家 E.R. Malinowski 教授 [美国] 高兴地给本书写了序并给我们寄来许多有价值的参考资料和程序，D.G. Howery 教授 [美国]，A.D. Zuberbuhler 教授 [瑞士]，P.J. Gemperline 教授 [美国] 和 P. Robert 教授 [法国] 都在本书编写期间给我们寄来了他们的最新研究成果或计算机程序。他们的热情支持和帮助大大地丰富了本书的内容。本书引用的作者研究室的某些研究成果系国家自然科学基金资助的工作。对于本书的编写，中国科学技术大学近代化学系的俞书勤教授、中国科学院化工冶金研究所的许志宏教授和长春应用化学研究所的许禄教授都曾给予热情的关怀和鼓励。中国科学院化工冶金研究所的石乐明博士、湖南大学的李志良博士、重庆大学的刘信安副教授和河南师范大学的夏四清老师曾提出过许多有益的建议。王月香同志在本书的整个编写过程中给予了热情的支持，席英、潘德川、邓浩明等同志对本书书稿的整理和打印付出了艰辛的劳动，中国科学技术大学近代化学系计算机化学研究室对于书稿的整理提供了良好的工作条件，江苏无锡堰桥防腐工程设备厂的吕正荣先生对本书的付梓予以慷慨的支持。在此，

趁本书出版之际，我们一并向他们表示衷心的感谢。

限于我们的水平和经验，加之时间仓促，书中难免会有错误、欠妥和不足之处，敬祈前辈和广大同行们批评指正。

编著者

1991 年于中国科学技术大学

目次

序言	E.R. Malinowski (i)
前言	(iii)
1 导论	(1)
1.1 计量化学的兴起	(1)
1.2 因子分析在计量化学中的地位	(3)
1.3 化学因子分析理论的发展	(5)
2 抽象因子分析	(9)
2.1 因子分析的主要步骤	(9)
2.1.1 主要步骤概述	(9)
2.1.2 主要步骤的数学提要	(11)
2.2 预备	(13)
2.2.1 问题的选择	(13)
2.2.2 数据的选择	(16)
2.2.3 数据的预处理	(17)
2.3 数据的分解与复原	(20)
2.3.1 协方差矩阵	(20)
2.3.2 分解的原理	(22)
2.3.3 主因子分析	(24)
2.3.4 JACOBI 法	(34)
2.3.5 计算列矩阵	(39)
2.3.6 计算行矩阵	(40)
2.3.7 因子的数目和复原	(41)
2.4 向量的诠释	(44)
2.5 主因子分析解的变换	(49)
2.6 数据例解	(53)

3	目标因子分析	(57)
3.1	概述	(57)
3.2	目标检验与目标变换	(57)
3.2.1	目标检验	(58)
3.2.2	总体考虑	(60)
3.2.3	构造检验向量	(61)
3.2.4	目标变换	(64)
3.3	自由浮动与迭代目标检验	(67)
3.3.1	自由浮动	(67)
3.3.2	迭代目标检验	(69)
3.4	目标检验结果的诠释	(72)
3.5	组合及有关的预测	(74)
3.5.1	组合概述	(75)
3.5.2	典型向量的组合及有关的预测	(76)
3.5.3	基础向量的组合及有关的预测	(83)
3.6	独特性检验和单位向量检验	(86)
3.7	数据例解	(88)
4	秩消因子分析	(93)
4.1	问题的提出与基本思想	(93)
4.2	秩消因子分析的早期模型	(94)
4.3	秩消因子分析的非迭代求解	(97)
4.3.1	奇异值分解	(97)
4.3.2	非迭代求解	(98)
4.3.3	秩消因子分析法与标准加入法的结合	(100)
4.4	广义秩消因子分析	(101)
4.4.1	单一组分的定量	(102)
4.4.2	几种组分的同时定量	(104)
4.4.3	以校准作为基础	(104)
4.4.4	通用模型	(105)
4.5	双线性目标因子分析	(106)
4.6	秩消因子分析法应用于荧光数据	(107)

4.7	秩消因子分析法在色谱中的应用	(109)
4.7.1	液相色谱	(109)
4.7.2	薄层色谱	(113)
4.8	秩消因子分析法在其它方面的应用	(114)
5	渐进因子分析	(115)
5.1	概述	(115)
5.1.1	Gemperline 法	(116)
5.1.2	VDK 法	(118)
5.1.3	GMMZ 法	(119)
5.1.4	几种方法的比较	(120)
5.2	渐进因子分析的原理	(121)
5.2.1	方法的数学描述	(121)
5.2.2	初级渐进因子分析	(123)
5.2.3	终级渐进因子分析	(127)
5.3	其它无模型技术及其与渐进因子分析的比较	(129)
5.4	渐进因子分析的应用	(131)
5.4.1	浓度、光谱及平衡常数的计算	(131)
5.4.2	色谱中的峰分辨	(133)
5.4.3	具有非理想行为的平衡混合物的分析	(135)
5.4.4	产品质量控制	(137)
5.4.5	光谱数据的半定量分析及模型选择	(138)
5.4.6	在圆二向谱中的应用	(140)
5.5	渐进秩消因子分析简介	(144)
6	对应因子分析	(149)
6.1	一些典型问题介绍	(149)
6.2	对应因子分析的数学描述	(151)
6.2.1	数据预处理	(151)
6.2.2	降维处理	(153)
6.3	对应因子分析的数据实例	(157)
6.4	对应因子分析的应用	(161)

7	因子分析的误差理论	(171)
7.1	误差理论	(171)
7.1.1	实验误差的干扰	(172)
7.1.2	主要因子和次要因子	(176)
7.1.3	数字实例	(179)
7.2	误差与分析数据的改善	(181)
7.3	确定因子数目的方法	(185)
7.3.1	依赖实验误差的方法	(185)
7.3.2	经验方法与统计学方法	(190)
7.4	误差判据的其它应用	(206)
7.4.1	推断实验误差	(206)
7.4.2	检验可因子分析性	(207)
7.5	目标检验向量中的误差	(207)
7.5.1	理论	(208)
7.5.2	可靠性函数	(213)
7.5.3	损害函数	(214)
7.5.4	目标检验的统计学 F 检验	(219)
7.6	因子载荷中的误差	(223)
7.7	数据实例	(226)
8	因子分析在组分分析中的应用	(230)
8.1	吸收光谱	(230)
8.1.1	紫外 - 可见光谱	(230)
8.1.2	红外光谱	(232)
8.1.3	近红外光谱	(235)
8.2	发射光谱	(238)
8.2.1	喇曼光谱	(238)
8.2.2	荧光光谱	(240)
8.2.3	诱导耦合等离子发射光谱	(242)
8.3	色谱	(243)
8.4	质谱	(245)
8.4.1	抽象因子分析研究	(245)

8.4.2	目标因子分析研究	(246)
8.5	动力学	(250)
8.6	旋光色散	(252)
8.7	X 射线方法	(255)
8.8	表面光谱学	(256)
9	因子分析在化学基础研究中的应用	(258)
9.1	核磁共振理论研究	(258)
9.1.1	质子溶剂位移的研究	(258)
9.1.2	TMS 和环己烷的 ^1H , ^{13}C 和 ^{29}Si 溶剂位移	(271)
9.1.3	氟溶剂位移	(273)
9.1.4	取代基对 ^{13}C 位移的影响	(274)
9.1.5	其它机理研究	(275)
9.2	色谱性质的研究	(275)
9.2.1	色谱与因子分析	(275)
9.2.2	活度系数预测和溶剂的分类	(277)
9.2.3	因子数目的确定及其应用	(278)
9.2.4	独特性检验与单位值检验	(280)
9.2.5	典型向量关键集的应用	(282)
9.2.6	目标检验	(286)
9.2.7	基础向量关键集的应用	(288)
9.3	线性自由能关系的研究	(289)
9.4	质谱性质的研究	(293)
9.5	在分析化学中的比较	(297)
9.5.1	仪器比较	(298)
9.5.2	方法比较	(298)
9.5.3	介质比较	(300)
9.6	其它基础研究	(301)
9.6.1	溶解度与溶液特性	(301)
9.6.2	极谱	(304)
9.6.3	稳定常数	(305)
9.6.4	键能	(306)

9.6.5	物质状态	(307)
9.7	在化学其它学科中的应用	(308)
9.7.1	生物医学化学	(308)
9.7.2	环境化学	(313)
9.7.3	其它相关领域	(316)
10	多组分同时测定	(317)
10.1	一般原理	(318)
10.2	氨基酸混合体系的同时测定	(321)
10.2.1	实验手续	(322)
10.2.2	目标因子分析全过程	(322)
10.3	计算一个完整的模拟数值实例	(326)
10.3.1	抽象因子分析	(326)
10.3.2	目标检验与组合变换	(328)
10.4	目标因子分析法的一种改进	(329)
10.4.1	方法概述	(330)
10.4.2	实际应用	(331)
参考文献		(333)

1 导 论

1.1 计量化学的兴起

在当代科学技术革命浪潮的推动与冲击下,分析工作的仪器化、自动化和计算机化等得到迅猛的发展.现代分析仪器能迅速、准确地为人们提供大量的可靠的量测数据,因此,对于分析化学工作者来说,如何选择最合适的分析方法和最优的测量过程,以及如何对由实验得到的原始数据进行再加工以便从中提炼出新的、更多的有价值的化学信息就显得越来越重要.随着电子计算机科学、应用数学以及统计学方法在化学、尤其是在分析化学中应用的不断广泛和深入,一门崭新的化学分支学科——计量化学 (Chemometrics) 诞生了.计量化学的兴起,正在逐步改变以前被认为理论工作落后于迅猛发展的学科实践的分析化学的学科面貌.

计量化学是建立在多学科基础上的一门新兴学科,是化学的一个重要分支.计量化学是在 1970 年由瑞典的 S. Wold 教授首先提出的,并与美国的 B.R. Kowalski 教授一起于 1974 年在美国 Seattle 成立了计量化学学会,他们以改进化学和应用数学及统计学衔接为目标,并把计量化学的任务规定为:应用和发展统计学方法和其它数学方法,从实验量测中取得有用的化学信息.鉴于现代分析仪器的展,以及计量化学在计算机与分析仪器相结合时所显示的优越性, Kowalski 进一步认为,计量化学作为一门新的化学学科,应强调应用现代数学和统计学去改进分析仪器,并从化学量测中取得更有用的化学信息.在华盛顿成立的计量化学学会上,计量化学的初次定义为:“计量化学是一门化学分支学科,它应用数学和统计学方法(借助计算机技术),设计和选择最优的测量程序和实验方法,并且

通过解释化学数据而获得最大限度的信息。在分析化学领域中, 计量化学是应用数学和统计学方法, 用最佳方式获取关于物质系统的有关信息。”如果把现代分析仪器看作是体现现代分析化学功能强弱的硬件的话, 那么, 计量化学就可比作是反映现代分析化学水平高低的软件。强功能的硬件和高水平的软件的结合及开发和应用, 将使现代分析化学的面目焕然一新。

计量化学的发展, 不但开拓了对统计学的运用, 而且已涉及到光谱和波形分析、运筹学及控制论等计算机应用技术。计量化学所包括的内容相当广泛, 据目前较流行的资料概括, 主要有: ①统计学方法、②最优化方法、③信号处理、④因子分析、⑤曲线分辨、⑥校正、⑦模型化与参数估计、⑧结构与性能相关、数据库及其检索、⑨模式识别(包括图象分析)和⑩人工智能等。

计量化学在解决复杂的化学问题中显示出强大的生命力, 并导致了分析化学的又一次革命, 引起了化学工作者的极大关注和浓厚的兴趣。开展计量化学的研究, 可从实验量测数据中尽可能多地提取有用信息, 实现分析工作者由过去的单纯的“数据提供者”到“问题的解决者”的飞跃。计量化学在分析化学中已有多方面的应用, 它能提高分析测试的精密度和准确度。在分析化学中, 无论是选用已建立的方法, 还是开发新的分析方法, 都要进行一系列条件实验, 研究各种因素的影响, 以确定最佳的测试条件, 获得最佳的测试结果。在各因素间存在协同影响时, 单因素实验将不能获得最佳测试条件, 而这时计量化学方法却能提供有用的帮助, 能借助信号处理技术提高分析测试的灵敏度与选择性, 提高信噪比, 消除干扰, 使重叠信号得以分辨。如傅里叶变换及各种滤波技术的应用, 使分析测试发生了深刻的变革。计量化学能促进仪器联机与自动化及智能化, 由于电子计算机技术的发展, 特别是微型计算机的普及, 分析方法与分析实验室的自动化程度越来越高, 采用有效的计量化学手段可使数据的获取、处理及由分析数据加工成有用的分析信息的过程日趋自动化与智能化。计量化学还可帮助化学家发展许多新的测