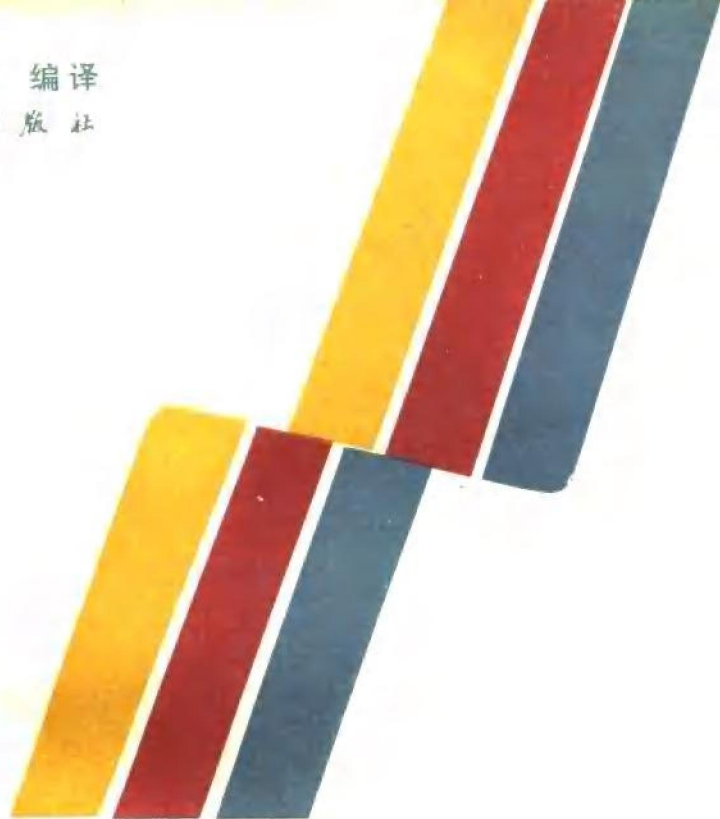


姚松年 编译
科学出版社



信息论

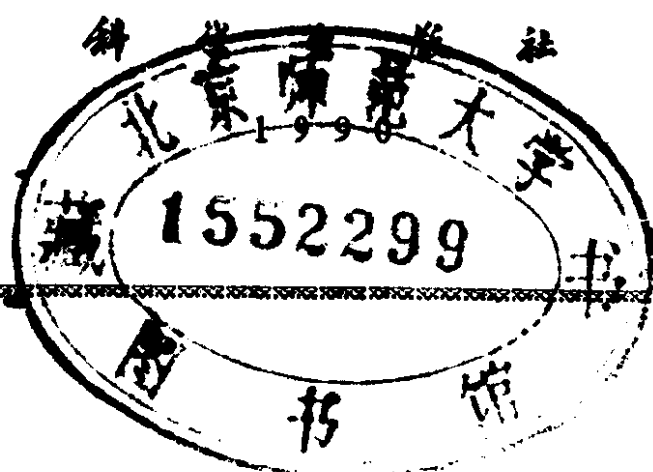
在分析化学中的应用

信息论

在分析化学中的应用

姚松年 编译

J211.72/09



内 容 简 介

本书着重介绍信息论的基本原理和概念在分析化学中的应用,把分析和测试作为获得信息的过程,用信息论、概率论和统计推断的方法讨论各种分析方法、分析过程中信息的产生与信息量的计算以及如何用信息量作为优化分析方法和过程的定量指标。

全书共八章:第一、二章论述分析是获取信息的过程;第三至五章介绍概率论、信息论、数理统计的基本概念、原理和方法;第六至八章阐述信息论在分析化学及分析测试中的应用。全书内容丰富,观点新颖,避免冗繁数学推导,论述简明扼要,颇具特色。

本书可供化学、化工、地质、冶金、环保及国防战线从事分析化学、分析测试、理化检测的科技工作者及大专院校化学专业师生参考。

信息论

在分析化学中的应用

姚松年 编译

责任编辑 操时杰

科学出版社出版

北京东黄城根北街16号

邮政编码:100707

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

1990年12月第1次印刷 开本:787×1092 1/32

1990年12月第一次印刷 印张:8

印数:001—1 830

字数:179 000

ISBN 7-03-001881-8/O·375

定价:6.80元

编 译 者 序

信息论在分析化学和分析测试中的应用是一个引人瞩目的新课题，也是化学计量学的一个分支。目前，信息论已用来研究、评价、优化分析方法、分析操作和分析结果，可以预料，随着科学技术的发展，它将具有越来越广泛的实际应用价值。捷克斯洛伐克科学院 K. Eckschlager 教授和布拉格经济研究中心 V. Štěpánek 教授在这个领域的研究工作在世界上处于领先地位。他们所著的“Information Theory as Applied to Chemical Analysis” (1979) (应用于分析化学中的信息论)和“Analytical Measurement and Information”(1984) (分析测试与信息)两书，不仅系统地阐述了信息论在分析化学和分析测试中应用的基本概念和一般原理，并且概括地介绍了作者和其他科学工作者在这个领域中最新的研究成果。第一本书偏重于基础，详细介绍信息论在分析化学中的应用，对概率统计和信息论的基本概念和原理也作了简要论述。第二本书是第一本的续集。它着重介绍作者及各国科学工作者在这一领域中的最新研究成果。

目前我国系统地介绍信息论在分析化学中应用的书籍还很少见。我们将这两本书编译成本书，旨在引起广大化学工作者对这个领域的注意和兴趣。但由于编译者的水平所限，失当之处在所难免，欢迎读者批评指正。

为使读者阅读方便，并避免重复和冗繁，编译者对这两本书的内容进行重新编写，使之更加系统、合理。

武汉大学化学系程介克教授对本书的编排方式和内容进

行过认真的审校，张华山副教授在本书定稿过程中提出了许多宝贵意见，在此谨向他们表示真挚的谢意。

编译者

目 录

第一章 绪论	1
参考文献.....	5
第二章 分析是获得信息的过程	8
2.1 分析过程	8
2.2 定性和定量的分析信息	10
2.3 样品成分信息产生过程的性质	12
2.3.1 灵敏度	12
2.3.2 专一性	13
2.3.3 选择性	13
2.3.4 线性性	14
2.3.5 时间独立性	15
2.4 分析信号的性质	15
2.4.1 一维信号	16
2.4.2 二维信号	16
2.5 样品成分信息的解析	19
2.6 分析结果的性质	23
2.7 从时间和经济观点评价获得的信息	24
参考文献.....	25
第三章 概率论的基本概念	26
3.1 概率论介绍	26
3.1.1 样本空间	26
3.1.2 概率的定义	27
3.1.3 概率的性质	31
3.1.4 条件概率	33

3.1.5 随机独立	33
3.2 数学模型	35
3.2.1 随机变量	35
3.2.2 描述函数	36
3.2.3 二维描述函数	43
3.2.4 期望和矩	46
3.3 某些特殊的概率分布	49
3.3.1 二项分布	50
3.3.2 泊松分布	52
3.3.3 正态分布	57
3.3.4 对数-正态分布	60
3.3.5 均匀分布	61
参考文献	62
第四章 信息论	63
4.1 Hartley 公式	63
4.2 Shannon 公式	64
4.3 信息的增益	67
4.4 信息定义的扩展	72
4.5 化学分析常用的几种信息量	75
4.5.1 Brillouin 量	76
4.5.2 Shannon 量	77
4.5.3 偏差量	81
参考文献	83
第五章 统计推断的若干方法	85
5.1 随机样本	85
5.2 统计样本的概率分布	86
5.2.1 均值的样本分布	86
5.2.2 样本方差的样本分布	87
5.2.3 Student 比	88

5.3	置信区间	89
5.4	试验的统计假设	91
5.4.1	正态分布均值的统计显著性试验	93
5.4.2	正态分布中两个均值的比较	97
5.4.3	正态分布中两个方差的比较	98
5.5	线性回归分析	99
5.6	单纯形优化	102
	参考文献	104

第六章 分析结果和分析方法的评价与优化..... 105

6.1	信息论在分析化学中的应用	105
6.2	定性检验的信息量	106
6.3	定量测定的信息量	108
6.3.1	定量测定中信息量的计算	108
6.3.2	分析条件对结果信息量的影响	112
6.4	测定次数,精密度和准确度对信息量的影响 ...	113
6.5	分析方法的校正对信息量的影响	117
6.5.1	单标准和多标准校正法	117
6.5.2	非零平均误差对信息量的影响	119
6.6	高精密度分析结果的信息量	121
6.7	痕量分析结果的信息量	124
6.8	二维仪器方法的信息量	127
6.8.1	二维仪器方法中信息量计算	127
6.8.2	二维仪器方法中信息的剩余度	129
6.9	联用分析方法的信息量	130
6.10	质量控制分析中结果的信息量	131
6.11	分析方法的评价	134
6.12	重复测定的有效次数	136
6.13	最优分析方法的选择	139

6.14	分析操作的优化	142
6.15	图形识别法	145
6.16	信息论在分析化学中的其它应用	149
	参考文献	153
第七章	信息论对各种分析操作的探讨	157
7.1	定性分析和鉴定	157
7.1.1	定性分析和鉴定的信息论模型	157
7.1.2	影响信息论评价指标的几个因素	161
7.2	单组分定量分析	164
7.2.1	信息论模型及影响评价的因素	165
7.2.2	单组分定量分析的其它评价指标	170
7.3	多组分定量分析	173
7.3.1	信息论模型及影响评价的因素	174
7.3.2	多组分定量分析中信息的相关性	181
7.4	痕量分析	187
7.4.1	评价痕量分析的信息论模型	187
7.4.2	截尾效应对痕量分析的影响	188
7.5	质量控制分析	193
	参考文献	195
第八章	分析体系的研究	198
8.1	不定度的来源	198
8.2	分析化学体系的描述	200
8.3	分析体系中的不定度规则	203
8.4	分析体系的定量特征	205
8.5	校正的影响	207
8.5.1	校正曲线法和标准加入法	207
8.5.2	两种校正方法的比较	211
8.6	化学分析取样	213
8.6.1	分层次取样	214

8.6.2 分层次取样与总体随机取样的比较	218
参考文献	221
附录 A 几种分布	223
A.1 χ^2 -分布	223
A.2 Student t -分布	224
A.3 F -分布	226
附录 B 几种信息量的推导	227
B.1 信息量的定量证明	227
B.2 定量测定的信息量	227
B.3 高精密度分析中的信息量	228
B.4 分析方法的准确度对信息量的影响	229
B.5 分析方法的校正对信息量的影响	230
B.6 痕量分析结果的信息量	232
附录 C 随机变量的熵	235
C.1 一个随机变量的熵	235
C.2 二个随机变量的熵	235
C.3 传输信息及其性质	237
C.4 f -偏差与 f -信息	238
符号表	244

第一章 绪 论

分析化学是一门用化学和物理学方法描述物质成分的基础学科。在本世纪初，它是建立在化学反应基础上的分析方法和分析操作，仅是实验化学的一部分。这门学科发展的初期，通常把“分析化学”作为分析方法的收集来理解。自从1977年 W. Fresenis^[1]提出“分析化学是应用自然科学的方法获得物质系统的信息，并解释和研究这些信息的一门科学”的新观点后，给分析化学赋予了新的定义。在这个定义中，把获得信息作为化学分析和仪器分析的目的，把分析操作作为获得物质化学成分信息的过程。当然，分析化学从“收集操作方法”到“获得信息的科学”，它经历了一个缓慢的发展过程。

分析化学概念的重要发展是在1936年。H. Kaiser^[2]根据一个未知组分的测定结果是具有某种概率分布的随机变量这一事实，用概率论的观点阐明了在当时用发射光谱测得的分析结果重现性差的实验现象。这就使得用统计方法评价分析结果和分析方法成为可能。自H. Kaiser^[2]第一篇论文发表以后10到15年间，分析工作中使用统计方法愈来愈普遍。分析结果和分析方法的统计评价随着仪器分析方法，特别是痕量分析的发展而发展。于是也出现了纯分析概念的统计定义，譬如H. Kaiser提出了检测限和测定下限的统计定义；尔后，J. D. Winefordner和L. A. Currie又发展了这个定义，例如信号、信噪比、信号的检测等通讯理论中的概念在痕量分析的文献中被普遍地采用。在最近十几年，分析化学已朝着借助于以物理方法为主来实现分析目的的方向发展。

与此同时实验室的实验技术(仪器和计算技术)也迅猛发展。随着痕量分析的发展,分析过程所需试样的量也愈来愈少,过去认为无法分析和鉴别的物质现在已成为可能。借助于微区分析,我们已经能建立起在以往看来是完全均匀而实际上是多相组成的材料的分析方法。这些都与经典的分析化学有显著的差异。在今天,它已由原来个别的方法论发展成完全独立的体系,有自己的“理论基础”、工作方法和术语,或者说是实验室的“行话”。而且来自不同领域的“方法”论文都发表在相关专业的杂志上,用新的理论和数学表达式探讨有关分析问题的论文也不断在分析化学文献中出现。也正因为这种差异又大大加速了分析学科的发展,使它成为一个多学科的领域。分析工作者原来感兴趣的领域变得愈来愈狭小,有时竟到了这样的程度,以致于要解决的问题已失去了“分析处理”的意义。因此需要寻找一种普遍有效的分析概念来弥补这种差异,信息论在分析化学中的应用正好起到这样的作用。

自第二次世界大战末以来,信息论得到很大发展,开始它仅是概率论的一部分,而今已成为一个独立的分支。它起源于1920年由R. A. Fisher介绍的某些基本定义。1945年以后通过N. Wiener, C. E. Shannon和S. Kullback等人的努力在美国得到广泛的传播。信息论以概率论作为理论基础,用概率论的方法研究、处理测量的不定度。它并不起源于经验事实,而是从抽象的定义出发进行研究的一门纯数学学科。因此,无论哪里所应用的信息论,它的叙述方式都具有同样的数学表达式,并不依赖于任何经验关系。

信息论在分析化学中的应用是一个既发展迅速又引人瞩目的新课题。1970年初,“Lindau小组”的研究者从系统论和信息论的观点出发,根据分析化学的需要,导出了它的一般

概念和定义^[3]，他们也认识到分析化学已发展成为如何得到和如何解释有关材料的化学组成的信息，如何处理携带这些信息的信号这样一门独立的学科^[1]。尽管 H. Kaiser 不是把信息论应用于分析化学的第一个人，但他于 1969 年在乔治亚大学的演讲以及一年后发表的文章^[4]对于把信息论的观点引伸到分析化学中有着奠基性的意义。在对分析结果的信息量的理解和应用中引入测量偏差作为信息增益，这就比以前直接用由通讯理论变换来的量更为确切。今天信息量已被进一步用来评价、比较和优化分析方法、分析操作及分析仪器^[5-20]。另外根据所用的分析操作的化学和物理内容或仪器的性质，用信息系统的观点，对“黑箱”进行了探讨。信息论在分析信号的解析，例如在红外光谱 [IR] 或质谱 [MS] 中也很重要^[21,22]。1980 年 Kowalski 又将信息论在分析化学中的应用列入了化学计量学的领域^[23]。

信息论在分析化学中应用的文献愈来愈多。作者等在 1979 年发表的一篇论文^[5]中对此已作全面阐述。另外也发现信息论在寻求分析方法最优决策方面有诸多应用^[24]，还有许多评论性文章^[6-8,19]评述了信息论在分析化学某些领域应用的可能性。这些论文大致可分为三类：(1)从 70 年代开始，大部分的原始论文都指出了信息论在分析化学中应用的重要性。用语言文字或者信息论的关系式或者系统论的关系式对分析化学中某些概念和量作出了定义。这些论文为分析化学作为一门独立的现代交叉学科的概念作出了贡献。“Lindan 小组”的研究成员 Malissa 编写的教科书中的各类论文^[3]值得一读。(2)第二部分的论文是把信息测试和信息量引入分析化学中，并着重把它们应用于评价和优化分析操作的分析工作中。较早的有 Brillouin 量^[25,10-12,9,16]，较近期的有从 Shannon 通讯理论中转引而来的量^[7]，Cleij 和 Dijkstra

在 1979 年介绍了模糊度^[8], Frank 等^[19]又导出了评价分析结果和分析方法的传递信息. 偏差量是在测试中得到的一个综合信息量^[2], Eckschlager 和 Vajda 首先把它用于分析化学中^[26], 以后又由 Eckschlager 等把它的应用推广和总结^[17,5,27,6]. (3) 用信息论的方法和关系式评价与比较分析的结果、方法和操作, 处理分析数据, 优选分析方法和分析操作, 对分析信号进行编译、解析等方面的研究工作也发表了一百多篇论文.

在此我们还想综述属于第二部分论文的主要结果, 把某些不同的定义和术语进行分类, 对形成今天化学计量学领域的数学基础进行评述. Bochev^[28]在化学计量学丛书第五卷的绪言中对信息论在化学各个分支学科中的应用作了评述, 在本书中也讨论了信息论在分析化学中应用的各种可能性. 但是值得指出的是分析化学的化学计量学与化学的其它分支学科之间还存在一些区别. 数学和统计方法在物理化学、有机化学、无机化学或者在生物化学中的应用主要是对所获得的数据进行整理、分类或者寻求它们相互间的关系, 可是它们在分析化学中的用途除了获得数据外, 几乎涉及获得化学成分信息的全过程. 然而, 我们亦认为在分析化学中应用化学计量学的宗旨是同时使用几种数学方法, 这些方法是互相补充、互相结合的. 如果要同时应用几个数学方法解决某一问题, 对不同内容我们可以从不同的基本概念出发来研究它们的有利条件, 协同现象等, 即对给定的课题研究各种方法的最佳合作趋势. 尽管在分析工作和最佳合作方案的研究中, Derffel 和 Eckschlager^[24]在 1981 年的论文中已注意到统计方法与信息论的结合, 但是这种观点在化学计量学的文献中尚不多见.

在此尚需强调指出, 本书主要是介绍如何用信息论来描述、评价、优化在分析化学和分析测试中获得的信息过程. 不

着重于自动化和数据处理方面的应用，尽管信息论能有效地用于这些方面。我们感兴趣的课题已在 Venn 图(图 1.1)中用黑实线表示。

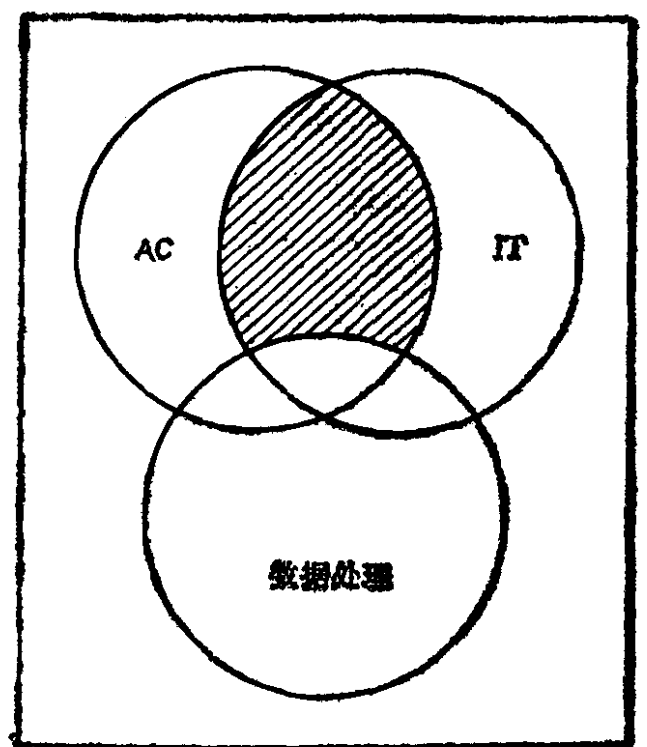


图 1.1 Venn 图
AC. 分析化学; IT. 信息论。

同时要指出，整个信息论是以信息为基本概念，而不定度是它的同义语。因为不定度是概率论的数学研究课题，它是信息论的基础。没有它，信息论的核心概念就不能被采纳。

参 考 文 献

- [1] Fresenius. W., "Reviews on Anal. Chem." p. 11, Akad, Kiado, budapest (1977).
- [2] Kaiser. H., *Z. Tec. phys.*, 17, 219, 277 (1936).
- [3] Malissa. H., "Automation in und mit der Fnalytischen Chemie. Verlag der Wiener Medizinischen Akademie", Wien (1972).
- [4] Kaiser, H., *Anal. Chem.*, 42, No. 2, 24A; No. 4, 26A (1970).

- [5] Vajda. I. and Eckschlager. K., *Analysis of a Measurement Information Kybernetika*, **16**, 120—144 (1980).
- [6] Eckschlager. K. and Štěpánek. V., "Information theory in Analytical Chemistry", *Anal. Chem.*, **54**, 1115A—1127A. (1982).
- [7] Liteanu, C. and Rica. I., "Utilization of the Amount of Information in Evaluation of Analytical Methods", *Anal. Chem.*, **51**, 1986—1995 (1979).
- [8] Cleij. P. and Dijkstra. A., "Information Theory Applied to Qualitative Analysis", *Fresenius Z. Anal. Chem.*, **298**, 97—109 (1979).
- [9] Danzer. K. and Eckschlager. K., "Information Efficiency of analytical of Analytical Methods", *Talanta*, **25**, 725—726 (1978).
- [10] Danzer. K., "Zu Einigen Informationstheoretischen Aspekt der Analytik," *Z. Chem.*, **13**, 20—21 (1973a).
- [11] Danzer. K., "Die Information Smenge Als Kenngrösse Zweri Dimensionaler Analytischen Aussagen", *Z. Chem.*, **13**, 69—70 (1973b).
- [12] Danzer. K., "Ermittlung der Informationsmenge Qualitativer Analysengänge", *Z. Chem.*, **13**, 229—231 (1973c).
- [13] Danzer. K., "Information Stheoretische Charakterisierung der Verteiligngsanalysen", *Z. Chem.*, **14**, 73—75 (1974).
- [14] Danzer. K., "Zur Ermittlung der Informationsmenge bei spektrochemischen Analysenverfahren", *Z. Chem.*, **15**, 158—159 (1975).
- [15] Danzer. K., "Kennzeichnung der Leichstungs Vermogen des Rentabalen Einsatzes Von Analysenverfahren mit Hilfe der Informationstheorie", *Z. Chem.*, **15**, 326—327 (1975b).
- [16] Danzer. K. and Marx. G., "Zn Einigen Informationstheoretischen Grundlagen der Strukturanalytik", *Chem. Analis*, **24**, 33—42; 43—50 (1979).
- [17] Eckschlager. K., "Informationsgehalt analytischer Ergebnisse Fresenius", *Z. Anal. Chem.*, **277**, 1—8 (1975).
- [18] Eckschlager. K., "Zur Ermittlung der Informationsmenge bei Zweidimensionalen Analysenverfahren", *Z. Chem.*, **15**, 326—327 (1975b).
- [19] Frank, J., Veress. G. and Pugor. E., Some problems of the Application of Information Theory in Analytical Chemistry", *Hungar. Sci Instr.* **54**, 1—9 (1982).
- [20] Danzer. K., Hopfe. V. and Marx. G., "Möglichkeiten der Erhöhung der Informationsmenge Spektroskopischer Analysenmethode mit Hilfe der Rechentechnik", *Z. Chem.*, **22**, 332—328 (1982).
- [21] Van. G. Marlen and Dijkstra. A., "Information Theory Applied to Selection of Peaks for Retrieral of Mass Spectra", *Anal. Chem.*, **48**, 595—598 (1976).
- [22] Dupuis. P. F. and Dijkstra. A., "An Information Theoretical Evaluation of the Applicability of the ASTM Infrared Data Base for Retrieval

- Purposes", *Fresenius Z. Anal. Chem.*, **290**, 357—368. (1978).
- [23] Kowalski. B. R., Chemometrics, *Anal. Chem.*, **52**, 112R—122R. (1980).
- [24] Doerffel. K. and Eckschlager. K., "Optimale Strategien in der Analytik", Verlag Harri Deutsch, Frankfurt/m (1982).
- [25] Eckschlager. K., "The Amount of Information Obtained by Analysis", *Coll. Czechoslov. Chem. Commun.*, **36**, 3016—3022 (1971).
- [26] Eckschlager. K. and Vajda. I., "Amount of Information of Repeated Higher Precision Analyses", *Coll. Czechoslov. Chem. Commun.*, **39**, 3076—3081 (1974).
- [27] Eckschlager. K. and Štěpánek. V., "Accuracy of Analytical Results", *Coll. Czechoslov. Chem. Commun.*, **45**, 2516—2523 (1980).
- [28] Bonchev. D., "Information Theoretic Indices for Characterization of Chemical Structures" (1983).