

硝基苯类化合物的岭回归

——可见分光光度法同时测定

林 君 邱雪梅 邵明武 李惕川

(北京工业大学环境与能源工程学院, 北京, 100022)

摘 要 根据萘乙二胺与硝基苯类发生偶合显色反应而生成有不同颜色的稳定配合物的特点, 测定了水中硝基苯类混合体系的吸收光谱, 结合聚类分析、因子分析和岭回归技术对合成水样中硝基苯、对硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯等多组分硝基苯的光谱数据进行解析, 成功地确定了合成水样中硝基苯类的物种数及各物种的含量。

关键词 硝基苯类化合物, 化学计量学, 因子分析, 岭回归

分类号 X832, O635

工业废水中存在着多种硝基苯类化合物, 它们会引起水质感官性状的长期严重恶化, 并对人体产生毒性作用。目前工业废水中的5种硝基苯类化合物: 硝基苯、对硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯已被列入优先污染物^[1]。然而由于这些物质的物理和化学性质相近, 互相干扰严重, 如果对它们进行分别测定, 为了保证测定的准确性和精密性, 就必须对样品进行前处理预分离, 过程繁琐冗长。因此, 不需要经过分离直接同时测定多组分硝基苯类化合物已逐渐引起环境分析工作者的关注^[2~4]。本文根据萘乙二胺与硝基苯、对硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯发生还原-偶氮反应, 用可见分光光度法进行了实验测定, 实验数据运用化学计量学方法进行实验条件优化、定性、定量分析, 并取得了较好的效果。

1 基本原理

1.1 优化波长集合的选择

对于光谱复杂的体系, 尤其对于病态体系, 直观的选择分辨较好的波长点是不容易的。为此, 我们可采用一种优化波长集合的方法——聚类优化波长点方法^[5]。该方法包括以下两个判据。

1.1.1 加和性检验

分别对 A 、 B 、 C 、 D 、 E 5种化合物和一个它们的混合样品 H 进行处理, 并在一定范

围内扫描。设吸收曲线分别为 A 、 B 、 C 、 D 、 E 、 H ，通过运算可绘制出可见光区的加和性曲线 $(A + B + C + D + E) - H$ ，从曲线中选取加和性好的波长范围，进行下面的检验。

1.1.2 组分谱差异性检验

以各种组分在不同波长点处的摩尔吸光系数值为指标，在给定的波长范围内进行合并，直到合并为一大类为止。然后根据要选定的波长点数，确定最优波长集合。

1.2 岭回归^[6]

当一组解释变量中存在多重共线性时，个别回归系数的普通最小二乘估计趋向于不稳定，从而可能导致错误的判断。但若采用岭回归的估计方法，可以比普通最小二乘法提供更多的数据分析信息。

回归系数的岭回归估计可以通过解一个略微变更的正规方程而得到。假定回归模型的标准形式为

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + u \quad (1)$$

关于估计岭回归系数的方程组是

$$\begin{cases} (1+k)\hat{\beta}_1 + r_{12}\hat{\beta}_2 + \cdots + r_{1p}\hat{\beta}_p = r_{1y} \\ r_{21}\hat{\beta}_1 + (1+k)\hat{\beta}_2 + \cdots + r_{2p}\hat{\beta}_p = r_{2y} \\ \cdots \\ r_{p1}\hat{\beta}_1 + r_{p2}\hat{\beta}_2 + \cdots + (1+k)\hat{\beta}_p = r_{py} \end{cases} \quad (2)$$

这里， r_{ij} 是第 i 个与第 j 个解释变量之间的单相关系数， r_{iy} 是第 i 个解释变量与相应变量 y 之间的单相关系数。(2)式的解是一组所估计的岭回归系数。

设 $0 \leq k < \infty$ ，称

$$\hat{\beta}(k) = (\hat{\beta}_1(k), \hat{\beta}_2(k), \cdots, \hat{\beta}_p(k)) = (X'X + kI)^{-1}X'Y \quad (3)$$

为 β 的岭回归。由 $\hat{\beta}$ 的岭估计求得的回归方程便为岭回归。

2 实验

2.1 试剂与仪器

所用硝基苯、对硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯、氨基磺酸铵、N-(1-萘基)乙二胺、亚硝酸钠、硫酸铜、盐酸等均为分析纯试剂。样品在 25 mL 比色管中处理；测定用 TU-1221 紫外-可见分光光度计。数据处理使用 Pentium 166 计算机 (IBM)。

2.2 测定步骤

在 25 mL 具塞比色管中，加入硝基苯、对硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯 (均为 $\rho = 10 \mu\text{g} / \text{L}$) 标准溶液，再加入 $6.0 \text{ mol} / \text{L}$ 的盐酸 2.0 mL 和 $\rho = 50 \text{ g} / \text{L}$ 的硫酸铜溶液 0.2 mL，用 0.2 g 锌粉还原。还原 20 min 后，用滤纸过滤于另外的比色管中，然后加入 $\rho = 10 \text{ g} / \text{L}$ 的亚硝酸钠溶液 0.2 mL，摇匀，放置 5 min，加入 $\rho = 25 \text{ g} / \text{L}$ 氨基磺酸铵溶液 0.5 mL，摇匀 3 次，赶走气泡，待气泡消失后，加 2.0 mL $\rho = 5 \text{ g} / \text{L}$ 的 N-(1-萘基)乙二胺盐酸溶液，显色 40 min 后，在 500~650 nm 范围测定吸光度。

3 结果与讨论

3.1 可见光吸收光谱

5种硝基苯与二盐酸N-(1-萘基)乙二胺发生偶合反应生成5种紫红色染料的可见光吸收光谱(如图1)。从图中可以看到:5种硝基苯类化合物在可见光区的吸收曲线重叠严重。要想不作任何分离,直接用常规方法测定混合体系中的各吸收组分的浓度是不可能的。

3.2 优化波长集合的选择

3.2.1 加合性检验

由图2所示,在540~600nm的波长范围内加合性曲线趋于零,加合性良好。

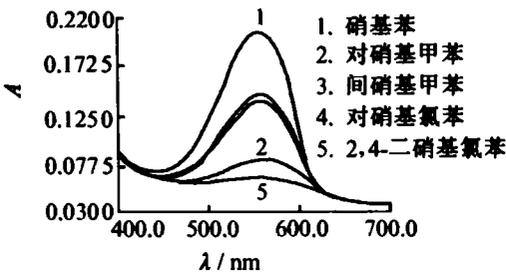


图1 5种硝基苯的可见光吸收曲线

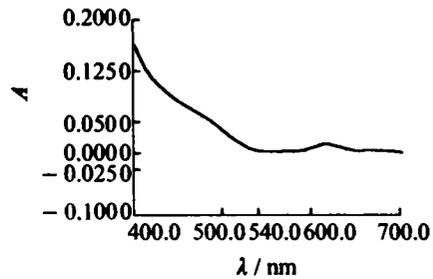


图2 5种硝基苯的可见光区加和性曲线

3.2.2 组分谱差异性检验

对于可见光区5种硝基苯类物质的数据矩阵由5个样品和61个指标(即540~600nm)组成,数据就是它们在61个波长点处的摩尔吸光系数值。根据模糊聚类的方法,对61个指标进行R型聚类。最终得到的优化波长点为:598,591,589,587,586,584,583,582,581,580,579,578,577,576,575,574,573,572,571,570,557,548,546,543nm,共24个。

3.3 合成水样的测定结果

选取这5种物质不同配比的混合水样10个,分别测定500~650nm范围内151个波长点吸光度值。进行定性、定量分析,结果见表1~表3。

表1 合成水样的抽象因子分析结果

NF	特征值 $\lambda/\times 10^3$	RE/ $\times 10^3$	IE/ $\times 10^3$	IND/ $\times 10^3$	ER	REV/ $\times 10^3$
1	359 483.00	39.943	12.631	0.493	760.665	998.564
2	472.59	12.406	10.548	0.394	34.721	1.500
3	13.61	8.040	9.047	0.325	1.364	0.500
4	9.98	3.797	8.196	0.272	1.670	0.432
5	5.98	1.754	6.402	0.262	1.206	0.031
6	4.96	1.246	6.388	0.515	1.332	0.032
7	3.75	0.778	6.256	0.831	1.358	0.031
8	2.76	0.744	6.032	1.686	1.3 7	0.032
9	2.05	0.627	5.528	5.827	1.679	0.037
10	1.22					

表2 合成水样的目标因子分析结果

物类	AET 误差的均方根	RET 真实误差	REP 目标向量均方根	SPOIL 损害函数
硝基苯	29 141.92	21 665.65	19 489.77	0.90
对硝基甲苯	15 890.87	6 277.59	14 598.34	2.32
间硝基甲苯	25 484.78	16 555.33	19 375.12	1.17
对硝基氯苯	24 559.56	15 341.36	19 178.50	1.25
2,4-二硝基氯苯	18 207.69	8 207.65	16 252.82	1.98

表3 合成水样实验计算结果

样品 编号	硝基苯			对硝基甲苯		
	实验值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	计算值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	相对 误差/%	实验值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	计算值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	相对 误差/%
1	0.080	0.083	3.50	0.160	0.164	2.75
2	0.160	0.165	3.00	0.240	0.245	2.17
3	0.240	0.245	2.17	0.320	0.327	2.12
4	0.320	0.328	2.62	0.400	0.416	3.90
5	0.400	0.398	0.40	0.080	0.083	3.50
6	0.160	0.164	2.25	0.160	0.166	3.75
7	0.160	0.58	1.50	0.240	0.248	3.17
8	0.240	0.242	0.83	0.240	0.244	1.83
9	0.240	0.234	1.33	0.320	0.327	2.25
10	0.320	0.308	3.88	0.160	0.165	3.25
平均相对误差/%	2.15			2.87		
总平均相对误差/%	3.11					

样品 编号	间硝基甲苯			对硝基氯苯			2,4-二硝基氯苯		
	实验值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	计算值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	相对 误差/%	实验值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	计算值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	相对 误差/%	实验值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	计算值/ ($\mu\text{g}\cdot\text{mL}^{-1}$)	相对 误差/%
1	0.240	0.240	3.00	0.320	0.328	2.62	0.400	0.443	10.70
2	0.320	0.329	2.75	0.400	0.412	3.10	0.080	0.082	3.00
3	0.400	0.409	2.30	0.080	0.079	1.50	0.160	0.167	4.25
4	0.080	0.084	4.50	0.160	0.165	3.75	0.240	0.249	3.83
5	0.160	0.165	3.25	0.240	0.249	3.83	0.320	0.328	2.50
6	0.240	0.246	2.50	0.240	0.263	9.67	0.320	0.326	2.00
7	0.240	0.247	2.83	0.320	0.331	3.38	0.160	0.165	3.00
8	0.320	0.333	4.00	0.160	0.162	1.25	0.160	0.154	4.00
9	0.160	0.164	3.25	0.160	0.156	2.75	0.240	0.248	3.17
10	0.160	0.165	3.25	0.240	0.234	2.75	0.240	0.246	2.50
平均相对 误差/%	3.06			3.56			3.90		
总平均相 对误差/%	3.11								

5种硝基苯类化合物的不同配比的10个合成水样的抽象因子分析结果见表1。从表1可看出:当 $n=5$ 时,RE(真实误差)的值恰好小于实验误差;IE(嵌入误差)基本趋于稳定;IND(指示函数)为最小;ER(特征值比)出现一最小值,即出现一折点;REV(约化特征值)显著变小且大体趋于稳定。综合考虑这些判据,可判断样品中有5种组分,与实验完全相符。

在确定了混合物的组分数之后,运用目标因子分析法对可能存在于体系中的各组分进行

目标检验,以确定所存在的物种种类.用在相同条件下测定的各组分在各波长点处的摩尔吸光系数作为目标检测向量,对这些目标进行检验,所得结果见表2.

由表2可见:SPOIL值都在0.0~3.0之间,根据Malinowski提出的SPOIL(损害函数判据)的经验规则,可以认为这5种物质确实存在,都是真实因子,这与实验完全相符.

从表3可以看到:用岭回归对混合样品进行定量分析,除有个别奇异值外,在可见光区的总平均相对误差为2.96%,具有较好的可靠性.

4 结论

采用聚类分析法对光谱数据进行特征筛选,减少了数据矩阵的多重共线性关系,增加了组分光谱间的差异,提高了定性定量分析的准确度.这种方法对于解决光谱重叠的病态体系具有较大的实用价值.用抽象因子和目标因子分析法成功的对硝基苯类五元混合体系进行了定性分析,正确确定了体系中的物种种数和物种种类.岭回归对于硝基苯、硝基甲苯、间硝基甲苯、对硝基氯苯、2,4-二硝基氯苯的混合体系的定量测定效果较好.

参 考 文 献

- 1 国家环保局著.水和废水监测分析方法.北京:中国环境科学出版社,1989.424
- 2 李新云,李惕川.模糊聚类——因子分析光度法同时测定工业废水中多组分酚.北京工业大学学报,1997,23(1):7~17
- 3 李惕川,黄敏.化学模式识别—岭回归分光光度测定混合酚.北京工业大学学报,1997,23(2):30~36
- 4 李惕川.水中几种硝基苯类化合物的同时测定.数理统计与管理,1997,(增刊):261~263
- 5 方开泰.实用多元统计分析.上海:华东师范大学出版社,1986.215~246
- 6 王学仁.应用回归分析.重庆:重庆大学出版社,1991.98~110
- 7 Hoerl A E, Kennard R W. Ridge regression. Biased estimation for nonorthogonal problems. Technometrics, 1970, 12: 69~82

Simultaneous Determination of Nitrobenzenes with Ridge Regression Spectrophotometry

Lin Jun Qiu Xuemei Shao Mingwu Li Tichuan

(College of Environment and Energy Engineering,

Beijing Polytechnic university, Beijing, 100022)

Abstract Based on that N-(1-Naphthyl)-aethylendiamin-dihydrochlorid can reaction with nitrobenes to form a magenta, the spectra data are analyzed by mathematical methods to optimize the wavelengths, and to determine the number of multicomponents, the quantity of each species as well as the concentrations of every compound involved.

Keywords nitrobenzene, chemometrics, factor analysis, ridge regression