

# 紫外光谱重叠的多组分混合物体系的同时定量测定

刘 慧, 胡仰栋, 卢彦越

中国海洋大学化学化工学院, 山东 青岛 266003

**摘 要** 提出一种可将紫外吸收光谱严重重叠的混合物体系准确定量分析的方法。该方法利用系统聚类结合加和性检验而得的优化波长集合有效地降低了混合物光谱的相关性。通过选取优化波长集合内的波长点的信息, 把吸光系数和待测浓度同时作为自变量, 建立非线性优化模型, 对于组分间浓度差较小的混合物体系经一次优化分析计算可得待测浓度; 混合物中组分浓度差较大的体系经四次左右优化迭代, 逐步降低误差的分析结果, 相对误差可控制在 3.63% 之内。分析结果表明该法稳定、准确、简便快速, 可对多组分混合物体系进行定量测定。

**关键词** 紫外光谱; 系统聚类; 多组分定量分析; 迭代; 非线性规划

**中图分类号:** O657.3 **文献标识码:** A **文章编号:** 1000-0593(2007)07-1416-04

## 引 言

各种光谱分析方法是多组分同时测定的重要方法。在分子光谱中, 不同有机化合物常因其化学结构相似使其紫外光谱严重重叠, 仪器和实验环境易使吸光度和浓度之间的关系偏离朗伯-比耳定律导致其呈现非线性特点。这些使得用常规分光光度方法同时测定这类化合物比较困难。采用化学计量学方法可以克服光谱定量分析中出现的多重共线性、光谱重叠和组分的相互干扰等问题, 并且不需分离可直接同时测定混合体系的各组分。因此发展了如主成分回归(PCR)<sup>[1]</sup>、小波包变换<sup>[2]</sup>、人工神经网络(ANN)<sup>[3]</sup>以及遗传算法<sup>[4]</sup>等多种线性或非线性的多元校准方法, 但是化学计量学方法可能会因为冗余数据的影响而得不到有效的校准模型<sup>[5]</sup>, 因此波长选择是光谱分析中各种多元校准方法的关键步骤。

本文在光谱数据预处理过程中用选取优化波长集合的方法, 即采用加和性检验和模式识别的系统聚类分析方法对波长进行优化选择, 突出各组分之间的差异, 降低波长点所含信息的相关性。然后把吸光系数和待测浓度同时作为自变量, 对分析问题建立有约束的非线性数学规划模型, 对于组分浓度差较小的混合物体系只需经一次优化计算可得待测浓度; 混合物中浓度差较大的体系经多次优化迭代, 逐步降低误差得到分析结果。利用本法对四组分光谱严重重叠的食用色素混合体系实验分析结果表明, 该法稳定、准确度高, 相对误差在 3.63% 内。

## 1 实验部分

### 1.1 实验仪器与试剂

仪器: Shimadzu2550 紫外-可见分光光度计(日本岛津公司); 试剂: 日落黄(Sunset Yellow)、苋菜红(Amaranth)、胭脂红(Ponceau 4R)和诱惑红(Allure red)。纯度都是 85%。

### 1.2 实验方法

用分析天平分别称取 0.250 0 g 以上试剂于 250 mL 容量瓶中, 用蒸馏水定容配制成 1.000 mg · mL<sup>-1</sup> 标准储备液。分别移取一定体积的储备液于 100 mL 的容量瓶中, 蒸馏水定容、摇匀、静置 15~30 min。用紫外-可见分光光度计, 参比溶液为蒸馏水, 波长 200~600 nm(间隔 1 nm)扫描其吸收光谱, 读取数据用于处理分析。

## 2 方法原理

### 2.1 优化波长集合的选择

选择加和性良好的连续波长段  $\lambda_p \sim \lambda_q$ , 运用系统聚类法分类, 分类的步骤是: 首先各样品自成一类; 以相似性度量计算各样品之间的距离, 用最近距离法将最近的两个样品并成一类; 计算新类与其余各类的距离, 再将距离最近的两类合并, 直到所有的样品归为一类, 则停止<sup>[6-11]</sup>。本文以相似系数距离作为相似性度量, 计算公式见下页。

收稿日期: 2006-03-10, 修订日期: 2006-06-16

基金项目: 国家自然科学基金项目(20376078)资助

作者简介: 刘 慧, 女, 1979 年生, 中国海洋大学化学化工学院研究生 e-mail: huizi421@126.com

$$d_{rs} = 1 - \frac{(x_r - \bar{x}_r)(x_s - \bar{x}_s)'}{[(x_r - \bar{x}_r)(x_r - \bar{x}_r)']^{\frac{1}{2}} [(x_s - \bar{x}_s)(x_s - \bar{x}_s)']^{\frac{1}{2}}}$$

其中

$$\bar{x}_r = \frac{1}{n} \sum_j x_{rj}, \bar{x}_s = \frac{1}{n} \sum_j x_{sj}$$

选取加和性良好的波长段 270~460 nm, 将这 191 个波长点聚为 30 类, 每类中采用加和性检验选取加和性最好的一个点作为这一类的代表, 这样选出的 30 个波长点差异性较大, 最大限度地克服了光谱重叠引起的相关性。

### 2.2 基本数学原理

待测混合物体系中有 4 种纯组分, 第  $i$  种纯组分的浓度设为  $c_{0i}$ , 在波长  $\lambda_j$  处的吸光系数设为  $\epsilon_{ij}$ , 通过实验得到其吸光度为  $A_{0ij}$ 。根据 Beer 定律(石英比色皿 1 cm)

$$A_{0ij} = \epsilon_{ij} c_{0i} \quad (1)$$

$\Delta A_{0ij}$  是浓度为  $c_{0i}$  的纯组分在波长  $\lambda_j$  处的吸光度的实验值  $A_{0ije}$  和  $A_{0ij}$  计算值之差, 即

$$\Delta A_{0ij} = A_{0ije} - A_{0ij} = A_{0ije} - \epsilon_{ij} c_{0i} \quad (2)$$

根据吸收定律加和性, 多组分混合物在波长  $\lambda_j$  处的吸光度  $A_{mj}$  ( $c_i^*$  为混合物中纯组分的真实浓度)为

$$A_{mj} = \sum_{i=1}^4 \epsilon_{ij} c_i^* \quad (3)$$

类似地, 命  $\Delta A_{mj}$  为多组分混合物在波长  $\lambda_j$  处吸光度的实验值  $A_{mje}$  和  $A_{mj}$  计算值之差, 即

$$\Delta A_{mj} = A_{mje} - A_{mj} = A_{mje} - \sum_{i=1}^4 \epsilon_{ij} c_i^* \quad (4)$$

命

$$\omega_j = \sum_{i=1}^4 (\Delta A_{0ij})^2 + (\Delta A_{mj})^2 \quad (5)$$

$\lambda_1 \sim \lambda_{30}$  为经加和性检验的聚类后所选取的 30 个波长点, 命

$$f = \sum_{j=1}^{30} \omega_j \quad (6)$$

将  $f$  做为目标函数, 则数学优化模型为

$$\text{Min } f = \sum_{j=1}^{30} \left[ \sum_{i=1}^4 (A_{0ij} - \epsilon_{ij} c_{0i})^2 + (A_{mje} - \sum_{i=1}^4 \epsilon_{ij} c_i^*)^2 \right] \quad (7)$$

Subjected to :  $c_i \geq 0, \epsilon_{ij} \geq 0$

经实验证明对于各组分浓度相近的混合物体系由估计值  $c_{0i}$  分析计算可得一次优化结果  $c_{1i}$ , 其误差在 1.83% 之内, 结果令人满意。但对于组分浓度差较大的混合物体系经一次优化计算后误差较大, 而经下面的迭代优化后, 误差会大大降低。

### 2.3 迭代优化

因一次优化结果得到的体系  $\sum_{i=1}^4 c_{1i}$  与  $\sum_{i=1}^4 c_i^*$  体系近似, 所以用此体系去继续优化评估待测混合物可提高方法准确度。

分别以浓度  $c_{1i}$  配制纯组分溶液, 紫外光谱扫描得其吸光度值  $A_{0ije_1}$ , 方程(1)~(7)中的  $c_{0i}, A_{0ije}$  被  $A_{0ije_1}$  替代, 则数学优化模型为:

$$\text{Min } f = \sum_{j=1}^{30} \left[ \sum_{i=1}^4 (A_{0ije_1} - \epsilon_{ij} c_{1i})^2 + (A_{mje} - \sum_{i=1}^4 \epsilon_{ij} c_i^*)^2 \right]$$

$$\text{s. t. } \begin{cases} c_i \in (0, c_{\max}) \\ \epsilon_{ij} \geq 0 \end{cases}$$

经计算得一次迭代优化结果  $c_{2i}$ , 同理以  $c_{2i}$  代替  $c_{1i}$  配制混合溶液重复以上步骤, 依此类推, 可得  $c_i^*$  的逼近数列  $\{c_{ki}\}$ , 直到  $c_{ki}$  与  $c_{(k-1)i}$  的数值一致在读数允许误差范围之内, 迭代结束。

### 2.4 建模求解

调用 Matlab 多元统计工具箱进行聚类分析得计算所需优化波长集合。将上述数学模型采用 GAMS 语言编程, 根据所建模型非线性约束多、自由度小的特点, 选择广义简约梯度优化算法 (GRG), 它是求解非线性规划最有效的算法之一。GRG 算法的基本思想是: 利用简约梯度直接构造一个使目标函数改善的可行方向, 然后沿此方向进行搜索, 从而找到一个新点, 这样逐步逼近原问题的最优解。经计算可得优化结果  $\epsilon_{ij}$  和  $c_{ki}$ 。模型中良好初始值的设定可以使求解既快又准, 而随着迭代过程的进行, 变换的初始值在逐渐逼近真值, 所以优化计算的准确度也在逐步提高。

## 3 结果与讨论

### 3.1 紫外吸收光谱

分别配制 0.02 g · L<sup>-1</sup> 的日落黄、苋菜红、胭脂红和诱惑红的水溶液 1 份, 扫描其吸收光谱, 选择加和性良好的连续波长段 270~460 nm, 绘制吸收曲线, 如图 1。

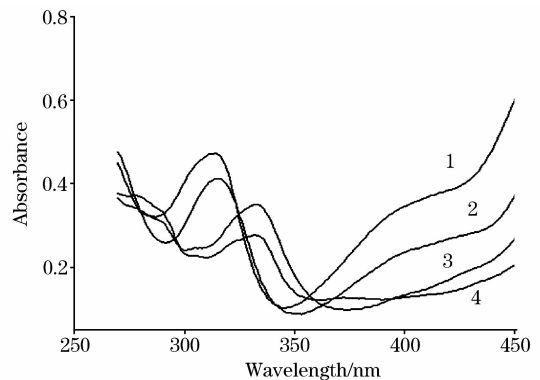


Fig. 1 Absorption spectra of the four edible pigments

1; Sunset yellow; 2; Allure red; 3; Ponceau 4R; 4; Amaranth

### 3.2 优化波长集合的选择

由图 1 可见, 四种食用人工合成色素的紫外吸收光谱重叠严重, 具有强的相关性, 直接用常规方法测定该混合体系中各组分的浓度难度较大。经聚类和加和性检验后选取的 30 个波长点为: 271, 275, 276, 277, 278, 279, 280, 281, 282, 283, 287, 291, 292, 293, 296, 323, 324, 325, 326, 327, 328, 349, 355, 356, 357, 358, 359, 360, 361, 401 nm。

### 3.3 样品分析结果

本文选择了  $0 \sim 0.050 \text{ g} \cdot \text{L}^{-1}$  浓度范围的人工合成食用色素日落黄(Sunset yellow)、苋菜红(Amaranth)、胭脂红

(Ponceau 4R)和诱惑红(Allure red)。纯度都是 85%。作为实例进行了分析。组分浓度差较小的混合物体系实验分析结果如表 1 所示,浓度差较大的体系实验分析结果如表 2 所示。

**Table 1 Analysis results of solution in which the concentration of every component is close to each other**

序号	样品	加入量/ $(\mu\text{g} \cdot \text{mL}^{-1})$	检出量/ $(\mu\text{g} \cdot \text{mL}^{-1})$	相对误差/%
1	A	20.00	19.64	-1.83
	B	20.00	20.15	0.74
	C	20.00	20.02	0.10
	D	20.00	19.82	-0.91
2	A	50.00	49.61	-0.79
	B	35.00	35.11	0.31
	C	45.00	45.0	80.18
	D	25.00	24.67	-1.34
3	A	20.00	20.21	1.04
	B	30.00	30.08	0.26
	C	40.00	40.14	0.35
	D	50.00	49.76	-0.48
4	A	8.00	8.14	1.72
	B	35.00	35.18	0.51
	C	40.00	39.77	-0.58
	D	15.00	15.05	0.33

**Table 2 Analysis results of solution in which the concentration of every component is not close to each other**

序号	样品	加入量/ $(\mu\text{g} \cdot \text{mL}^{-1})$	$c_{1i}/(\mu\text{g} \cdot \text{mL}^{-1})$	$c_{2i}/(\mu\text{g} \cdot \text{mL}^{-1})$	$c_{3i}/(\mu\text{g} \cdot \text{mL}^{-1})$	$c_{4i}/(\mu\text{g} \cdot \text{mL}^{-1})$	$c_{5i}/(\mu\text{g} \cdot \text{mL}^{-1})$	相对误差/%
1	A	4.00	3.63	3.98	4.00	4.01		0.25
	B	40.00	40.21	39.65	39.91	40.00		0.00
	C	3.00	4.27	3.50	2.95	3.01		0.33
	D	45.00	44.50	44.81	44.92	44.99		-0.02
2	A	45.00	46.09	45.16	45.04	44.80	44.79	-0.47
	B	40.00	38.87	39.91	40.11	39.96	40.23	0.57
	C	3.50	4.57	3.61	3.41	3.64	3.61	3.05
	D	4.00	3.49	3.82	3.96	4.07	4.05	1.23
3	A	45.00	43.85	45.06	45.17	45.03	45.07	0.16
	B	3.00	2.64	3.09	3.08	3.02	3.06	1.96
	C	4.00	3.89	4.12	3.83	4.12	4.14	3.38
	D	3.00	3.99	2.75	2.88	2.93	2.91	-3.09
4	A	3.00	2.69	3.02	2.86	2.96	2.97	-1.01
	B	4.00	3.40	3.61	3.72	3.91	3.86	-3.63
	C	45.00	47.50	45.87	45.75	45.44	44.96	-0.09
	D	45.00	44.72	45.36	45.58	45.47	45.36	0.79

A: 日落黄; B: 苋菜红; C: 胭脂红; D: 诱惑红

由样品分析结果可得,经过系统聚类分析特征筛选和加和性检验选定的优化波长集合,使原始数据的多重共线性得到了较好的改善,组分谱图之间有足够的差异性,能对光谱严重重叠的食用色素混合体系进行准确定量分析,说明系统聚类法和加和性检验选择优化波长集合的技术是可行的。

由表 2 可见,组分浓度差在 10 倍以上的混合物体系经一次优化结果基本不能满足混合物的定量分析要求,而随着迭代过程的进行,即初始值逐步逼近真值的过程,误差在逐步下降,经过 1~3 次迭代之后,误差可降到 3.63%,结果令人满意。

### 3.4 精密度的测定

取表 2 中的 4 号样品进行五次测定,并计算各次测定结果之间的相对标准偏差(RSD)结果如表 3 所示,本法具有较好的重现性。

**Table 3 Relative standard deviation (RSD)**

样品	日落黄	苋菜红	胭脂红	诱惑红
RSD%	3.02	2.23	1.18	2.71

## 4 结 论

采用聚类分析法对光谱数据进行优化选择,可有效减少数据矩阵的多重共线性,增加组分间的光谱差异性,提高了分析结果的准确度。将待测浓度和吸光系数同时作为决策变量,将定量分析问题表达为一个带约束的非线性数学规划问

题,优化分析出待测浓度。对于组分浓度差较大的体系,用与待测混合物体系相近的物系去定量分析评估计算,结果就愈接近真实浓度值。所以,在迭代过程中相对误差的降低是必然的。一般迭代4次左右数值趋于稳定,实例验证分析结果令人满意。证明本文方法是一种准确、易行的多组分定量分析方法。

## 参 考 文 献

- [1] ZHANG Li-qing, WU Xiao-hua, TANG Xi, et al(张立庆, 吴晓华, 唐 曦, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2002, 22(3): 427.
- [2] GAO Ling, REN Shou-xin(高 玲, 任守信). Journal of Analytical Science(分析科学学报), 2004, 20(6): 607.
- [3] SUN Xiu-yun, LI Yan, WANG Jun-de(孙秀云, 李 燕, 王俊德). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2003, 23(4): 739.
- [4] LI Zhi-liang, ZENG Ge-ming, XIA Zhi-ning, et al(李志良, 曾鹤鸣, 夏之宁, 等). Analytical Instrumentation(分析仪器), 1998, 2: 46.
- [5] Liu Shiqing, Wang Weiwen. Chemometrics and Intelligent Lab Systems, 1999, 45, 131.
- [6] FANG Kai-tai, PAN En-pei(方开泰, 潘恩沛). Cluster Analysis(聚类分析). Beijing: Geologic Press(北京: 地质出版社), 1982. 37.
- [7] SUN Xiao-yun, HU Yang-dong, LU Yan-yue(孙小云, 胡仰栋, 卢彦越). Computers and Applied Chemistry(计算机与应用化学), 2005, 22(9): 720.
- [8] DUAN Hui-qi, LU Xiao-hua(段慧琦, 陆晓华). Journal of Analytical Science(分析科学学报), 2003, 19(1): 86.
- [9] ZHANG Li-qing, LI Ping, MU Jie(张立庆, 李 萍, 穆 洁). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(10): 1238.
- [10] ZHENG Jing, LIN Kai-li, ZHOU Wei-liang(郑 静, 林开利, 周伟良). Analytical Laboratory(分析实验室), 2004, 23(11): 14.
- [11] YIN Chun-sheng, SHEN Yang, et al(印春生, 沈 阳, 等). Chemical Research in Chinese Universities(高等学校化学学报), 2000, 21(1): 49.

## Study of Simultaneous Quantitative Analysis of Overlapping UV Spectra

LIU Hui, HU Yang-dong, LU Yan-yue

College of Chemistry and Chemical Engineering, Ocean University of China, Qingdao 266003, China

**Abstract** The simultaneous multi-component quantitative determination was proposed without any chemical separation procedure. The method uses optimum wavelength gather obtained after hierarchical clustering and additive validation to reduce correlation of the component spectra. Information of 30 wavelength points after hierarchical clustering analysis was taken and, with the information of optimum wavelength gather, the absorptivity coefficient and component concentration of solutions were treated as independent variables, and a constrained optimization model was set up. The concentration awaiting test can be obtained after one optimization calculation for the mixed solution in which the concentration of every component is close to each other. But for some solutions in which component concentration is not close to each other, iterative optimization is necessary. The final results obtained are satisfying with a good accuracy and precision only after about 4 times of iterations, and the relative error can be controlled in 3.63%. So this method can be used for quantitative analysis of overlapping UV spectra.

**Keywords** Ultraviolet spectra; Hierarchical clustering; Multi-component quantitative analysis; Iteration; Nonlinear programming

(Received Mar. 10, 2006; accepted Jun. 16, 2006)